

Florian Modler Martin Kreh

Tutorium Analysis 2 und Lineare Algebra 2

Mathematik von Studenten für
Studenten erklärt und kommentiert



Tutorium Analysis 2 und Lineare Algebra 2

Florian Modler • Martin Kreh

Tutorium Analysis 2 und Lineare Algebra 2

Mathematik von Studenten für Studenten
erklärt und kommentiert

Autoren

Florian Modler
Drosselweg 10
31157 Sarstedt
modler@mathestudium-tutor.de

Martin Kreh
Lortzingstr. 23
31228 Peine
kreh@mathestudium-tutor.de

Homepage: www.mathestudium-tutor.de

Wichtiger Hinweis für den Benutzer

Der Verlag und die Autoren haben alle Sorgfalt walten lassen, um vollständige und akkurate Informationen in diesem Buch zu publizieren. Der Verlag übernimmt weder Garantie noch die juristische Verantwortung oder irgendeine Haftung für die Nutzung dieser Informationen, für deren Wirtschaftlichkeit oder fehlerfreie Funktion für einen bestimmten Zweck. Der Verlag übernimmt keine Gewähr dafür, dass die beschriebenen Verfahren, Programme usw. frei von Schutzrechten Dritter sind. Die Wiedergabe von Gebrauchsnamen, Handelsnamen, Warenbezeichnungen usw. in diesem Buch berechtigt auch ohne besondere Kennzeichnung nicht zu der Annahme, dass solche Namen im Sinne der Warenzeichen- und Markenschutz-Gesetzgebung als frei zu betrachten wären und daher von jedermann benutzt werden dürften. Der Verlag hat sich bemüht, sämtliche Rechteinhaber von Abbildungen zu ermitteln. Sollte dem Verlag gegenüber dennoch der Nachweis der Rechtsinhaberschaft geführt werden, wird das branchenübliche Honorar gezahlt.

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Springer ist ein Unternehmen von Springer Science+Business Media
springer.de

© Spektrum Akademischer Verlag Heidelberg 2011
Spektrum Akademischer Verlag ist ein Imprint von Springer

11 12 13 14 15 5 4 3 2 1

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung außerhalb der engen Grenzen des Urheberrechtsgesetzes ist ohne Zustimmung des Verlages unzulässig und strafbar. Das gilt insbesondere für Vervielfältigungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Planung und Lektorat: Dr. Andreas Rüdinger, Anja Groth
Redaktion: Bernhard Gerl
Satz: Florian Modler, Martin Kreh, Marco Daniel
Herstellung: Crest Premedia Solutions (P) Ltd, Pune, Maharashtra, India
Umschlaggestaltung: SpieszDesign, Neu-Ulm
Titelbildmotiv: © Carolyn Hall
Fotos/Zeichnungen: Marco Daniel

ISBN 978-3-8274-2630-7

Vorwort

Endlich ist es soweit, unser zweites Buch ist erschienen, wenn ihr diese Zeilen lest. Nach dem großen Erfolg mit dem ersten Teil (siehe [MK09]) war es uns ein Vergnügen, eine Fortsetzung schreiben zu dürfen. Leider hat dies ein wenig gedauert, aber das Buch sollte ja auch gründlich verfasst und kontrolliert werden.

Das Konzept bleibt das bewährte

Das Konzept ist genau wie im ersten Band. Zunächst werden wir in jedem Kapitel die wichtigsten Definitionen und Sätze mit Beweisen geben. Im zweiten Teil findet ihr die Erklärungen zu den Definitionen, Sätzen und Beweisen mit vielen Abbildungen und Beispielen. Voraussetzung für das Verständnis sind Analysis 1 und Lineare Algebra 1. Solltet ihr da ein wenig Nachholbedarf haben, so können wir ein gutes Buch empfehlen :-).

Inhalt

Der Inhalt ist klassisch aufgebaut. So hoffen wir jedenfalls, denn bei den Vorlesungen Analysis 2 und der Linearen Algebra 2 gibt es größere Unterschiede von Universität zu Universität. Der wichtigste und größte Themenkomplex in der Analysis 2 ist jedoch die Analysis mehrerer Veränderlicher, die in den Kapiteln über stetige Abbildungen, differenzierbare Abbildungen, Extremwertberechnungen und implizite Funktionen behandelt wird. Wir starten dabei mit einem Kapitel über metrische und topologische Räume, um ein wenig die Grundlagen zu legen. Abgerundet wird der Analysis-Teil durch jeweils ein Kapitel über gewöhnliche Differentialgleichungen, Kurven und Untermannigfaltigkeiten. Wir geben daher also auch einen Ausblick in die (elementare) Differentialgeometrie und in die höhere Analysis. Solche Ausblicke sind im zweiten Semester wichtig, denn bald wird es ernst und ihr müsst euch entscheiden, in welche Richtung ihr euer Studium vertiefen wollt und wo eure Interessen liegen.

Der Lineare-Algebra-Teil sollte euch von den Themen her auch nicht überraschen: Wir starten mit euklidischen und unitären Vektorräumen, gehen weiter zu Bilinearformen und hermiteschen Formen, bis wir bei den Normalformen und der Jordan-Theorie angelangt sind, was man wohl als Höhepunkt jeder Linearen-Algebra-2-Vorlesung bezeichnen kann. Um den Weg bis dort hin ein wenig schmackhafter zu gestalten, geben wir noch Kapitel über zyklische Gruppen, Ringe und Quadriken und eins über die schönen Symmetriegruppen. Abge-

rundet wird der Lineare-Algebra-2-Teil durch das Tensorprodukt, das wir kurz anreißen werden.

Insgesamt umfasst dieses Buch 15 Kapitel, und wir hoffen, dass wir nichts vergessen haben und ihr damit neben der Vorlesung, dem Vorlesungsskript und anderen Büchern gut für diese beiden Vorlesungen gewappnet seid. Wir wünschen euch jedenfalls ganz viel Erfolg im zweiten (oder höheren) Semester. Bei Fragen könnt ihr gerne wieder unsere altbekannte Homepage

<http://www.mathestudium-tutor.de>

besuchen, um eure Fragen im Forum loszuwerden, um Zusatzmaterial zu erhalten oder um uns einfach nur eure Meinung zu schreiben.

Danksagungen

Zu guter Letzt sei den Menschen gedankt, ohne die es dieses Buch gar nicht geben würde: den Korrekturlesern, den seelischen Unterstützern, den Grafikern und allen, die zum Entstehen und Erscheinen dieses Buches beigetragen haben. Da hätten wir zum einen die großartigen Korrekturleser, die fast alles neben Studium und Beruf stehen und liegen gelassen haben, um Fehler auszubessern und uns auf Verbesserungswürdiges hinzuweisen. Na ja, ganz so extrem war es nicht, dennoch waren alle sehr bemüht und fix im Lesen, und dies waren: Dr. Florian Leydecker, Dominik Bilitewsk, Stefan Hasselmann, Christoph Fuest, Fabian Grünig und Susanne Hensel. Außerdem geht ein Dank an Carolin Liebisch und Bernhard Gerl für die Ausbesserung des einen oder anderen Rechtschreibfehlers! Carolyn Hall hat wieder ein wundervolles Cover von Florti in der Badewanne erstellt. Wir danken ihr sehr dafür, denn uns gefällt es ausgesprochen gut!

Und wer hat diese genialen und tollen ca. 70 Abbildungen in dem Buch erstellt? Natürlich, es war wieder Marco Daniel, dem wir zutiefst danken wollen, denn so reibungslos und problemlos klappt es mit keinem, außer mit ihm!

Ein weiteres sehr großes Dankeschön geht an unsere Lektoren Anja Groth und Dr. Andreas Rüdinger, ohne die ihr dieses Buch jetzt nicht in den Händen halten könntet. Ja, wir wissen, es ist deren Job, aber zum Job gehört es nicht unbedingt, dass die Atmosphäre und die Zusammenarbeit so gestaltet wird, dass es einfach nur Spaß macht, mit dem Verlag zusammenzuarbeiten.

Ein letzter Dank gilt unseren Familien, Freunden und Freundinnen, die uns immer unterstützt haben und ein paar Mal auf uns verzichten mussten, weil der

Termin des Erscheinens und Druckens des Buches dann doch plötzlich immer näher rückte. Danke, ihr seid die Besten und Wertvollsten!

Nun aber genug der Reden, genießt das Buch, und für Fehlerhinweise sind wir, wie immer, sehr dankbar!

Hannover und Kopenhagen, November 2010
Florian Modler und Martin Kreh

Inhaltsverzeichnis

1	Metrische und topologische Räume	1
1.1	Definitionen	1
1.2	Sätze und Beweise	6
1.3	Erklärungen zu den Definitionen	14
1.4	Erklärungen zu den Sätzen und Beweisen	38
2	Stetige Abbildungen	43
2.1	Definitionen	43
2.2	Sätze und Beweise	45
2.3	Erklärungen zu den Definitionen	50
2.4	Erklärungen zu den Sätzen und Beweisen	61
3	Differenzierbare Abbildungen	65
3.1	Definitionen	65
3.2	Sätze und Beweise	67
3.3	Erklärungen zu den Definitionen	70
3.4	Erklärungen zu den Sätzen und Beweisen	82
4	Extremwertberechnungen	93
4.1	Definitionen	93
4.2	Sätze und Beweise	94
4.3	Erklärungen zu den Definitionen	95
4.4	Erklärungen zu den Sätzen und Beweisen	96
5	Implizite Funktionen	107
5.1	Sätze und Beweise	107
5.2	Erklärungen zu den Sätzen und Beweisen	111
6	Gewöhnliche Differentialgleichungen	117
6.1	Definitionen	117
6.2	Sätze und Beweise	119
6.3	Erklärungen zu den Definitionen	124
6.4	Erklärungen zu den Sätzen und Beweisen	127
7	Kurven	133
7.1	Definitionen	133
7.2	Sätze und Beweise	136
7.3	Erklärungen zu den Definitionen	141
7.4	Erklärungen zu den Sätzen und Beweisen	152
8	Untermannigfaltigkeiten	153
8.1	Definitionen	154
8.2	Sätze und Beweise	155
8.3	Erklärungen zu den Definitionen	156
8.4	Erklärungen zu den Sätzen und Beweisen	161

9 Euklidische und unitäre Vektorräume	165
9.1 Definitionen	165
9.2 Sätze und Beweise	166
9.3 Erklärungen zu den Definitionen	167
9.4 Erklärungen zu den Sätzen und Beweisen	170
10 Bilinearformen und hermitesche Formen	171
10.1 Definitionen	171
10.2 Sätze und Beweise	176
10.3 Erklärungen zu den Definitionen	183
10.4 Erklärungen zu den Sätzen und Beweisen	198
11 Gruppen und Ringe II	219
11.1 Definitionen	219
11.2 Sätze und Beweise	221
11.3 Erklärungen zu den Definitionen	230
11.4 Erklärungen zu den Sätzen und Beweisen	234
12 Symmetriegruppen	243
12.1 Definitionen	243
12.2 Sätze und Beweise	244
12.3 Erklärungen zu den Definitionen	249
12.4 Erklärungen zu den Sätzen und Beweisen	252
13 Symmetrische Bilinearformen und Quadriken	257
13.1 Definitionen	257
13.2 Sätze und Beweise	259
13.3 Erklärungen zu den Definitionen	266
13.4 Erklärungen zu den Sätzen und Beweisen	267
14 Normalformen	275
14.1 Definitionen	275
14.2 Sätze und Beweise	276
14.3 Erklärungen zu den Definitionen	280
14.4 Erklärungen zu den Sätzen und Beweisen	283
15 Tensoren und das Tensorprodukt	297
15.1 Definitionen	297
15.2 Sätze und Beweise	299
15.3 Erklärungen zu den Definitionen	301
15.4 Erklärungen zu den Sätzen und Beweisen	304
Symbolverzeichnis	307
Literaturverzeichnis	309
Index	311

1 Metrische und topologische Räume

Übersicht

1.1	Definitionen	1
1.2	Sätze und Beweise	6
1.3	Erklärungen zu den Definitionen	14
1.4	Erklärungen zu den Sätzen und Beweisen	38

Meistens macht man in der Analysis 2 die ersten Bekanntschaften mit topologischen Grundbegriffen und lernt wichtige topologische und metrische Räume, den Begriff der Kompaktheit, die offenen und abgeschlossenen Kugeln und den Satz von Heine-Borel kennen. Diese Begriffe überfordern am Anfang einige Studenten, da sie sehr abstrakt und teilweise wenig anschaulich sind. Aber keine Panik: Wir werden diese wichtigen Grundbegriffe der Topologie einführen, erklären und anhand von einigen Abbildungen und Beispielen mit Leben füllen.

1.1 Definitionen

Definition 1.1 (Metrik)

Sei M eine Menge. Eine **Metrik** ist eine Abbildung

$$d : M \times M \rightarrow \mathbb{R}$$

auf $M \times M$, für die folgende drei Axiome erfüllt sind:

- i) Positive Definitheit: Für alle $x, y \in M$ gilt $d(x, y) \geq 0$. Gleichheit gilt genau dann, wenn $x = y$ ist.
- ii) Symmetrie: Es gilt $d(x, y) = d(y, x)$ für alle $x, y \in M$.
- iii) Dreiecksungleichung: Es gilt:

$$d(x, y) \leq d(x, z) + d(z, y) \quad \forall x, y, z \in M.$$

Das Paar (M, d) nennen wir einen **metrischen Raum**.

Definition 1.2 (offene, abgeschlossene Kugel)

Seien (M, d) ein metrischer Raum, $x_0 \in M$ und $r > 0$. Die Menge

$$U(x_0, r) := \{x \in M : d(x, x_0) < r\}$$

bezeichnen wir als **offene Kugel**. Die Menge

$$B(x_0, r) := \{x \in M : d(x, x_0) \leq r\}$$

wird als **abgeschlossene Kugel** bezeichnet. Ab und an sagen wir statt „Kugel“ auch „Ball“.

Definition 1.3 (Umgebung)

Sei (M, d) ein metrischer Raum. Eine Teilmenge $U \subset M$ heißt **Umgebung** eines Punktes $x \in M$, falls ein $\varepsilon > 0$ existiert, sodass $U(x, \varepsilon) \subset U$. Insbesondere ist $U(x, \varepsilon)$ selbst eine Umgebung von x . Man nennt $U(x, \varepsilon)$ die **ε -Umgebung** von x .

Definition 1.4 (offene Menge)

Eine Menge $\Omega \subset M$ eines metrischen Raums (M, d) heißt **offen**, genauer **d -offen**, wenn zu jedem $x \in \Omega$ ein $\varepsilon > 0$ existiert, sodass $U(x, \varepsilon) \subset \Omega$.

Definition 1.5 (abgeschlossene Menge)

Eine Menge $A \subset M$ eines metrischen Raums (M, d) heißt **abgeschlossen**, wenn das Komplement $M \setminus A$ offen ist. Für das Komplement schreibt man auch oft A^c .

Definition 1.6 (Konvergenz)

Sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in M . Dann nennt man die Folge **konvergent** gegen den Punkt $x \in M$ genau dann, wenn Folgendes gilt:

$$\forall \varepsilon > 0 \exists N \in \mathbb{N} : d(x_n, x) < \varepsilon \forall n \geq N.$$

x heißt in diesem Fall **Grenzwert** der Folge.

In Worten: Für alle $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$ mit der Eigenschaft, dass $d(x_n, x) < \varepsilon$ für alle $n \geq N$.

Definition 1.7 (Häufungspunkt)

Sei $(x_n)_{n \in \mathbb{N}} \subset M$ eine Folge. Ein Punkt $x \in M$ heißt **Häufungspunkt** der Folge $(x_n)_{n \in \mathbb{N}}$, wenn es eine konvergente Teilfolge von $(x_n)_{n \in \mathbb{N}}$ gibt, die gegen x konvergiert.

Definition 1.8 (Cauchy-Folge)

Sei $(x_n)_{n \in \mathbb{N}} \subset M$ eine Folge. Wir sagen $(x_n)_{n \in \mathbb{N}}$ ist eine **Cauchy-Folge**, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, sodass $d(x_m, x_n) < \varepsilon$ für alle $m, n \geq N$.

Definition 1.9 (Vollständigkeit)

Ist K eine beliebige Teilmenge eines metrischen Raums (M, d) . Dann heißt K **vollständig**, wenn jede Cauchy-Folge $(x_n)_{n \in \mathbb{N}} \subset K$ auch einen Grenzwert in K besitzt. Ist M selbst eine vollständige Menge, so heißt der metrische Raum vollständig.

Definition 1.10 (Rand)

Seien (M, d) ein metrischer Raum und $A \subset M$ eine Teilmenge. Ein Punkt $x \in M$ heißt **Randpunkt** von A , wenn für jedes $\varepsilon > 0$ sowohl $U(x, \varepsilon) \cap A \neq \emptyset$ als auch $U(x, \varepsilon) \cap (M \setminus A) \neq \emptyset$ gilt. Wir definieren den **Rand** ∂A durch

$$\partial A := \{x \in M : x \text{ ist Randpunkt von } A\}.$$

Anmerkung: Der Rand ist also die Menge aller Randpunkte.

Definition 1.11 (Inneres)

Seien (M, d) ein metrischer Raum und $A \subset M$ eine Teilmenge. Das **Innere** von A ist definiert als $\overset{\circ}{A} := A \setminus \partial A$.

Definition 1.12 (Abschluss)

Seien (M, d) ein metrischer Raum und $A \subset M$ eine Teilmenge. Der **Abschluss** von A ist definiert als $\overline{A} := A \cup \partial A$.

Definition 1.13 (offene Überdeckung)

Seien (M, d) ein metrischer Raum und $K \subset M$ eine beliebige Teilmenge. Sei weiterhin I eine beliebige Indexmenge, und mit \mathcal{O} bezeichnen wir die Menge aller offenen Teilmengen von M . Eine **offene Überdeckung** $(\Omega_i)_{i \in I}$ von K ist eine Familie von offenen Teilmengen Ω_i , deren Vereinigung die Menge K umfasst, das heißt, $K \subset \bigcup_{i \in I} \Omega_i$, wobei $\Omega_i \in \mathcal{O}$ für alle $i \in I$.

Definition 1.14 (überdeckungskompakt)

Seien (M, d) ein metrischer Raum und $K \subset M$ eine Teilmenge.

K heißt **überdeckungskompakt**, wenn es zu jeder beliebig vorgegebenen offenen Überdeckung $(\Omega_i)_{i \in I}$ eine endliche Teilüberdeckung gibt, das heißt eine endliche Teilmenge $E \subset I$, sodass $K \subset \bigcup_{i \in E} \Omega_i$.

Definition 1.15 (folgenkompakt)

Seien (M, d) ein metrischer Raum und $K \subset M$ eine beliebige Teilmenge. K nennt man **folgenkompakt**, wenn jede Folge $(x_n)_{n \in \mathbb{N}}$ in K eine konvergente Teilfolge in K besitzt, das heißt, wenn der Grenzwert wieder in K liegt.

Definition 1.16 (total beschränkt, präkompakt)

Seien (M, d) ein metrischer Raum und $K \subset M$ eine Teilmenge. K heißt **total beschränkt** oder **präkompakt**, wenn es zu jedem $\varepsilon > 0$ eine endliche Anzahl von Punkten $x_1, \dots, x_n \in K$ und $n = n(\varepsilon) \in \mathbb{N}$ gibt mit $K \subset \bigcup_{i=1}^n U(x_i, \varepsilon)$.

Definition 1.17 (beschränkt)

Eine Teilmenge $K \subset M$ eines metrischen Raums (M, d) heißt **beschränkt**, wenn ein $x \in M$ und ein $r \in \mathbb{R}$ existieren mit der Eigenschaft, dass $K \subset U(x, r)$.

Definition 1.18 (Kompaktheit)

Eine Teilmenge $K \subset M$ eines metrischen Raums (M, d) heißt **kompakt**, wenn K überdeckungskompakt ist.

Anmerkung: Im Satz von Heine-Borel (siehe Satz 1.10) werden wir sehen, dass in metrischen Räumen die Überdeckungskompaktheit zu anderen Kompaktheitsbegriffen, wie zum Beispiel der Folgenkompaktheit, äquivalent ist. In den Erklärungen zu Definition 1.18 werden wir zeigen, dass es aber durchaus sinnvoll ist, mehrere Kompaktheitsbegriffe zu haben, weil es je nach Aufgabe einfacher ist, die Folgenkompaktheit oder die Überdeckungskompaktheit oder Ähnliches nachzuweisen.

Es ist möglich, einen abstrakteren Begriff eines Raums zu definieren. Wir hatten ja durchaus gesehen, dass sich gewisse Eigenschaften von metrischen Räumen nur durch offene Mengen beschreiben lassen. Dies wollen wir nun ausführen und die oben angeführten Definitionen allgemeiner fassen.

Definition 1.19 (topologischer Raum)

Seien M eine Menge und $\mathcal{O} \subset \mathcal{P}(M)$ ein System von Teilmengen von M . \mathcal{O} heißt eine **Topologie** auf M und das Paar (M, \mathcal{O}) ein **topologischer Raum**, wenn folgende Axiome erfüllt sind:

- i) $\emptyset, M \in \mathcal{O}$.
- ii) $\Omega_1, \Omega_2 \in \mathcal{O} \Rightarrow \Omega_1 \cap \Omega_2 \in \mathcal{O}$.
- iii) Ist I eine beliebige Indexmenge, und sind $(\Omega_i)_{i \in I}$ Elemente von \mathcal{O} , dann ist auch $\bigcup_{i \in I} \Omega_i \in \mathcal{O}$.

Die Elemente der Topologie nennen wir **offen**.

Definition 1.20 (offene Menge)

Sei (M, \mathcal{O}) ein topologischer Raum. Dann nennt man eine Teilmenge $\Omega \subset M$ **offen**, wenn $\Omega \in \mathcal{O}$.

Definition 1.21 (abgeschlossene Menge)

Sei (M, \mathcal{O}) ein topologischer Raum. Eine Menge $A \subset M$ heißt **abgeschlossen**, wenn das Komplement $M \setminus A$ offen ist.

Definition 1.22 (Umgebung)

Sei (M, \mathcal{O}) ein topologischer Raum. Ist $p \in M$, so heißt $\Omega \in \mathcal{O}$ eine **offene Umgebung** von p , wenn $p \in \Omega$.

Definition 1.23 (zusammenhängend)

Sei (M, \mathcal{O}) ein topologischer Raum. Dann heißt dieser **zusammenhängend** genau dann, wenn es außer der leeren Menge \emptyset und M selbst keine zugleich offenen und abgeschlossenen Teilmengen von M gibt.

Definition 1.24 (Basis)

Ist $B \subset \mathcal{O}$ ein System offener Teilmengen, so nennen wir B eine **Basis** von (M, \mathcal{O}) , wenn sich jede offene Menge $\Omega \in \mathcal{O}$ als Vereinigung von Mengen aus B darstellen lässt.

Definition 1.25 (Subbasis)

Seien (M, \mathcal{O}) ein topologischer Raum und $B \subset \mathcal{P}(M)$ eine Familie von Teilmengen von M . Dann heißt B eine **Subbasis** für \mathcal{O} genau dann, wenn die Familie

$$D := \left\{ \bigcap_{i=1}^n S_i : n \in \mathbb{N}, S_i \in B \right\}$$

aller endlichen Durchschnitte von Elementen aus B eine Basis für \mathcal{O} ist.

Wir führen jetzt den Begriff der Norm ein und werden den Zusammenhang zur Topologie und Metrik herausarbeiten.

Definition 1.26 (Norm)

Sei V ein K -Vektorraum, wobei K in der Regel entweder \mathbb{R} oder \mathbb{C} sein soll. Eine Abbildung

$$\|\cdot\| : V \rightarrow K$$

heißt **Norm**, wenn folgende drei Axiome erfüllt sind.

- i) Positive Definitheit: $\|v\| \geq 0$ für alle $v \in V$ und Gleichheit gilt genau dann, wenn $v = 0$ gilt.
- ii) Homogenität: $\|\lambda \cdot v\| = |\lambda| \cdot \|v\|$ für alle $\lambda \in K$ und für alle $v \in V$.

iii) Dreiecksungleichung: $\|v + w\| \leq \|v\| + \|w\|$ für alle $v, w \in V$.

Ist $\|\cdot\|$ eine Norm auf V , so heißt $(V, \|\cdot\|)$ ein **normierter Vektorraum**.

Definition 1.27 (Normenäquivalenz)

Es seien V ein Vektorraum und $|\cdot|$ und $\|\cdot\|$ seien zwei Normen auf V . Dann heißen die Normen **äquivalent**, wenn zwei positive Konstanten μ und λ existieren, sodass

$$\lambda|v| \leq \|v\| \leq \mu|v| \quad \forall v \in V.$$

Definition 1.28 (Banach-Raum)

Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Ist V dann mit der durch $d(x, y) = \|x - y\|$ definierten Metrik (siehe Satz 1.14) ein vollständiger metrischer Raum, so nennt man $(V, \|\cdot\|)$ einen **Banach-Raum**.

Definition 1.29 (Exponentialfunktionen von linearen Abbildungen)

Seien V ein Banach-Raum und A ein beschränkter Homomorphismus. Dann setzen wir

$$e^{tA} := \sum_{k=0}^{\infty} \frac{t^k}{k!} A^k, \quad t \in \mathbb{R}.$$

1.2 Sätze und Beweise

Die Sätze 1.1–1.4 gelten für topologische und für metrische Räume. Wir führen den Beweis jeweils für metrische Räume. Für topologische Räume gelten die Sätze aufgrund der Definition.

Satz 1.1

Sei (M, d) ein metrischer Raum. Der Durchschnitt von endlich vielen offenen Mengen ist dann wieder offen.

Vorsicht: Der Durchschnitt von unendlich vielen offenen Mengen muss nicht wieder offen sein, siehe dazu die Erklärungen zu diesem Satz.

Beweis: Gegeben seien ein metrischer Raum (M, d) und zwei offene Mengen Ω_1 und Ω_2 . Wir wollen nun zeigen, dass $\Omega_1 \cap \Omega_2$ offen ist. Sei $x \in \Omega_1 \cap \Omega_2$. Da nach Voraussetzung Ω_1 und Ω_2 offen sind, existieren ε_1 und $\varepsilon_2 > 0$ mit $U(x, \varepsilon_1) \subset \Omega_1$ und $U(x, \varepsilon_2) \subset \Omega_2$. Wir wählen nun $\varepsilon := \min\{\varepsilon_1, \varepsilon_2\}$. Es gilt $U(x, \varepsilon) \subset \Omega_1 \cap \Omega_2$. Induktiv folgt nun die Behauptung. q.e.d.

Satz 1.2

Sei (M, d) ein metrischer Raum. Die Vereinigung von beliebig vielen offenen Mengen ist wieder offen.

Beweis: Sei I eine beliebige Indexmenge, und für jedes $\alpha \in I$ sei Ω_α eine offene Teilmenge. Ist $x \in \bigcup_{\alpha \in I} \Omega_\alpha$, so existiert ein $\alpha \in I$ mit $x \in \Omega_\alpha$ und dann auch ein $\varepsilon > 0$ mit $U(x, \varepsilon) \subset \Omega_\alpha$. Daraus folgt, dass

$$U(x, \varepsilon) \subset \bigcup_{\alpha \in I} \Omega_\alpha$$

und folglich die Behauptung.

q.e.d.

Satz 1.3

Sei (M, d) ein metrischer Raum. Dann ist die Vereinigung endlich vieler abgeschlossener Mengen wieder abgeschlossen.

Vorsicht: Die Vereinigung von unendlich vielen abgeschlossenen Mengen ist im Allgemeinen nicht wieder abgeschlossen, siehe dazu wieder die Erklärungen zu diesem Satz.

Beweis: Sei (M, d) ein metrischer Raum und seien U_0, U_1, \dots, U_n , $n \in \mathbb{N}$, abgeschlossene Mengen mit $U_i \subset M$ für alle $i = 1, \dots, n$. Dann sind die Komplemente $M \setminus U_0, M \setminus U_1, \dots, M \setminus U_n$ nach Definition offene Mengen. Nach Satz 1.1 wissen wir, dass der Durchschnitt von endlich vielen offenen Mengen wieder offen ist, das heißt $\bigcap_{i=1}^n M \setminus U_i$ ist eine offene Menge, und aus den Regeln von De Morgan folgt nun, dass

$$\bigcap_{i=1}^n M \setminus U_i = M \setminus \bigcup_{i=1}^n U_i$$

offen ist. Also ist die Menge $\bigcup_{i=1}^n U_i$ abgeschlossen. Alles gezeigt.

q.e.d.

Satz 1.4

Sei (M, d) ein metrischer Raum. Der Durchschnitt beliebig vieler abgeschlossener Mengen ist wieder abgeschlossen.

Beweis: Sei (M, d) ein metrischer Raum und seien U_0, U_1, \dots beliebig viele abgeschlossene Mengen mit $U_i \subset M$ für alle $i \in I$, wobei I eine Indexmenge ist. In unserem Fall ist $I = \mathbb{N}$. Dann sind die Komplemente $M \setminus U_0, M \setminus U_1, \dots$ offene Mengen. Wir wissen bereits, dass die Vereinigung von beliebig vielen offenen Mengen wieder offen ist (Satz 1.2), das heißt, $\bigcup_{i \in I} U_i$ ist eine offene Menge, und aus den Regeln von De Morgan folgt nun, dass

$$\bigcup_{i \in I} U_i = M \setminus \bigcap_{i \in I} U_i$$

offen ist. Also ist die Menge $\bigcap_{i \in I} U_i$ abgeschlossen; fertig. q.e.d.

Satz 1.5 (hausdorffsche Trennungseigenschaft)

Sei (M, d) ein metrischer Raum. Dann gibt es zu je zwei beliebigen Punkten $x, y \in M$, mit $x \neq y$ Umgebungen U von x und V von y , die punktfremd sind, das heißt, $U \cap V = \emptyset$.

Anmerkung: Wir nennen den Raum, der diese Eigenschaft besitzt, dann auch hausdorffsch. Viele Räume besitzen diese Eigenschaft. Beispielsweise die metrischen Räume, wie wir sehen werden. Für topologische Räume ist dies aber im Allgemeinen falsch.

Beweis: Sei $\varepsilon := \frac{1}{3} \cdot d(x, y)$. Dann ist $\varepsilon > 0$, und $U := U(x, \varepsilon)$, $V := U(y, \varepsilon)$ sind punktfremde Umgebungen von x bzw. y . Denn gäbe es einen Punkt $z \in U \cap V$, so würde aus der Dreiecksungleichung folgen, dass

$$3\varepsilon = d(x, y) \leq d(x, z) + d(z, y) < \varepsilon + \varepsilon = 2\varepsilon,$$

also $3\varepsilon < 2\varepsilon$. Dieser Widerspruch beweist die Behauptung. q.e.d.

Satz 1.6

Offene Bälle $U(x_0, r)$ in einem metrischen Raum (M, d) sind offen.

Beweis: Sei $x \in U(x_0, r)$. Wir definieren $\varepsilon := r - d(x, x_0)$. Dann ist $\varepsilon > 0$ und $U(x, \varepsilon) \subset U(x_0, r)$, denn für $y \in U(x, \varepsilon)$ gilt wegen der Dreiecksungleichung

$$d(y, x_0) \leq d(y, x) + d(x, x_0) < \varepsilon + d(x, x_0) = \varepsilon - \varepsilon + r = r.$$

Es gilt also $U(x, \varepsilon) \subset U(x_0, r)$ und damit ist gezeigt, dass $U(x_0, r)$ offen ist. q.e.d.

Satz 1.7

Sei (M, d) ein metrischer Raum. Eine Teilmenge $A \subset M$ ist genau dann abgeschlossen, wenn jedes $x \in M$, welches Grenzwert einer Folge $(x_n)_{n \in \mathbb{N}} \subset A$ ist, sogar schon in A liegt.

Beweis: Zwei Richtungen sind zu beweisen:

„ \Rightarrow “: Sei A abgeschlossen, das heißt, $M \setminus A$ ist offen. Nach Definition der Offenheit (Definition 1.4) gibt es zu jedem $x \in M \setminus A$ ein $\varepsilon > 0$ mit $U(x, \varepsilon) \subset M \setminus A$. Sei weiter $(x_n)_{n \in \mathbb{N}} \subset A$ eine Folge. Dann gilt $x_n \notin U(x, \varepsilon)$ für alle $n \in \mathbb{N}$, und die Folge kann daher nicht gegen x konvergieren. Also folgt die Hin-Richtung.

„ \Leftarrow “: Wir führen einen Widerspruchsbeweis: Wäre A nicht abgeschlossen, das heißt $M \setminus A$ nicht offen, so gäbe es ein $x \in M \setminus A$ mit der Eigenschaft, dass für kein $n \in \mathbb{N}$ die Inklusion $U(x, 1/n) \subset M \setminus A$ richtig wäre. Zu $n \in \mathbb{N}$ existiert also ein $x_n \in A \cap U(x, 1/n)$. Dann konvergiert die Folge $(x_n)_{n \in \mathbb{N}}$ gegen x , denn es gilt $d(x_n, x) < \frac{1}{n}$. Nach Voraussetzung liegt der Grenzwert aber wieder in A . Dieser Widerspruch beweist die Rückrichtung. q.e.d.

Satz 1.8

Jede endliche Teilmenge eines metrischen Raums (M, d) ist abgeschlossen.

Beweis: Folgt sofort Satz 1.7. q.e.d.

Satz 1.9

Seien (M, d) ein metrischer Raum und $A \subset M$ eine Teilmenge. Dann sind der Rand ∂A und der Abschluss \overline{A} abgeschlossen und das Innere $\overset{\circ}{A}$ offen.

Beweis:

- Um zu zeigen, dass ∂A abgeschlossen ist, bleibt zu zeigen, dass $M \setminus \partial A$ offen ist. Sei $x \in M \setminus \partial A$ beliebig, das heißt, x ist kein Randpunkt. Daher existiert ein offener Ball $U(x, r)$, sodass $U(x, r) \cap A = \emptyset$ oder $U(x, r) \cap (M \setminus A) = \emptyset$ gilt. Nun zeigen wir, dass für dieses r dann auch $U(x, r) \cap \partial A = \emptyset$ gilt. Dies sieht man so ein: Wäre $y \in U(x, r) \cap \partial A$, so würde wegen der Offenheit von $U(x, r)$ zunächst ein $\varepsilon > 0$ mit $U(y, \varepsilon) \subset U(x, r)$ existieren. Da dann aber $U(y, \varepsilon)$ wegen $y \in \partial A$ und nach Definition des Randes sowohl einen Punkt aus A als auch einen Punkt aus $M \setminus A$ enthielte, träfe dies auch auf $U(x, r)$ zu. Dies ist ein Widerspruch, daher gilt $U(x, r) \cap \partial A = \emptyset$. Es folgt, dass $U(x, r) \subset M \setminus \partial A$, also ist $M \setminus \partial A$ offen und folglich ∂A abgeschlossen.
- Um einzusehen, dass der Abschluss $\overline{A} = A \cup \partial A$ abgeschlossen ist, zeigen wir, dass $M \setminus (A \cup \partial A)$ offen ist. Dazu sei $x \in M \setminus (A \cup \partial A)$ beliebig. Es gilt nach einfachen mengentheoretischen Aussagen (man male sich ein Venn-Diagramm)

$$M \setminus (A \cup \partial A) \subset M \setminus A$$

und damit auch $x \in M \setminus A$. x ist aber kein Randpunkt, also muss eine offene Kugel $U(x, r)$ mit $U(x, r) \cap A = \emptyset$ existieren, also $U(x, r) \subset M \setminus A$. Wie im ersten Teil folgt $U(x, r) \cap \partial A = \emptyset$ und damit ist

$$U(x, r) \subset (M \setminus A) \setminus \partial A = M \setminus (A \cup \partial A).$$

Also ist $M \setminus (A \cup \partial A)$ offen und $A \cup \partial A = \overline{A}$ abgeschlossen.

- Zum Schluss zeigen wir noch, dass $\overset{\circ}{A}$ offen ist. Wir definieren $B := M \setminus A$. Auf diese Mengen wenden wir nun das eben Bewiesene an. Es ergibt sich, dass \overline{B} abgeschlossen ist. Da nun aber $\partial(M \setminus A) = \partial A$, ist

$$\overline{B} = B \cup \partial B = (M \setminus A) \cup \partial A = M \setminus \overset{\circ}{A}.$$

Also ist $\overset{\circ}{A} = M \setminus \overline{B}$ offen.

Wir sind fertig.

q.e.d.

Satz 1.10 (Heine-Borel (allgemeine Version))

Sei (M, d) ein metrischer Raum. Dann sind für eine Teilmenge $K \subset M$ die folgenden Aussagen äquivalent:

- i) K ist überdeckungskompakt.
- ii) K ist folgenkompakt.
- iii) K ist total beschränkt und vollständig.

Anmerkung: Dieser Satz ist in topologischen Räumen nicht mehr korrekt.

Beweis: Wir wollen einen Ringschluss durchführen, daher sind drei Implikationen zu zeigen.

„i) \Rightarrow ii)“: Sei K überdeckungskompakt, das heißt, zu jeder beliebig vorgegebenen offenen Überdeckung existiert eine endliche Teilüberdeckung. Wir haben zu zeigen, dass K folgenkompakt ist. Dazu sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in K . Angenommen $(x_n)_{n \in \mathbb{N}}$ besitzt keine in K konvergente Teilfolge, dann existiert zu jedem $x \in K$ ein $r = r(x) > 0$, sodass $U(x, r(x))$ nur noch endlich viele Folgenglieder von $(x_n)_{n \in \mathbb{N}}$ enthält. Da dann aber

$$K \subset \bigcup_{x \in K} U(x, r(x))$$

und K überdeckungskompakt ist, existiert eine endliche Teilmenge $E \subset K$ mit $K \subset \bigcup_{x \in E} U(x, r(x))$. Das kann aber nicht sein, da in wenigstens einer dieser endlich vielen Mengen unendlich viele Folgenglieder von $(x_n)_{n \in \mathbb{N}}$ liegen müssten. Oder anders: $\bigcup_{x \in E} U(x, r(x))$ enthält nur endlich viele Folgenglieder, aber K unendlich viele und $K \subset \bigcup_{x \in E} U(x, r(x))$. Damit ist die erste Implikation gezeigt.

„ii) \Rightarrow iii)“: Sei K folgenkompakt. Wir haben zu zeigen, dass K total beschränkt und vollständig ist. Und los: Erst einmal bemerken wir, dass aus der Dreiecksungleichung ganz allgemein folgt, dass Cauchy-Folgen schon dann konvergent sind, wenn sie eine konvergente Teilfolge besitzen, und dieser Grenzwert stimmt dann auch mit dem Grenzwert der konvergenten

Teilfolge überein. Sei $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge. Wegen der Folgenkompaktheit existiert eine in K konvergente Teilfolge. Mit obiger Bemerkung folgt, dass K vollständig ist. Wir zeigen nun noch, dass K total beschränkt ist und zwar durch einen Widerspruchsbeweis. Angenommen K wäre nicht total beschränkt, dann existiert ein $\varepsilon > 0$, sodass K nicht von endlich vielen $U(x_i, \varepsilon)$ mit $x_i \in K$, $i \in E$, wobei E eine endliche Menge ist, überdeckt werden kann. Sei $x_1 \in K$ beliebig. Induktiv definieren wir eine Folge $(x_n)_{n \in \mathbb{N}} \subset K$ mit $d(x_i, x_j) \geq \varepsilon$ für alle $i \neq j \in \mathbb{N}$. Dies können wir machen, denn nach Annahme existieren für jede Auswahl von n Punkten $x_1, \dots, x_n \in K$ noch Punkte $x \in K$ mit $x \notin \bigcup_{i=1}^n U(x_i, \varepsilon)$. Wir wählen jetzt für x_{n+1} irgendeinen Punkt aus der Menge

$$K \setminus \left(\bigcup_{i=1}^n U(x_i, \varepsilon) \right)$$

aus. Die auf diese Weise erhaltene Folge muss wegen der Folgenkompaktheit eine konvergente Teilfolge besitzen. Dies ist aber ein Widerspruch zu $d(x_i, x_j) \geq \varepsilon$. Dieser Widerspruch beweist, dass K total beschränkt ist, und damit ist $ii) \Rightarrow iii)$ gezeigt.

„ $iii) \Rightarrow i)$ “: Wollen wir einen Ringschluss durchführen, so bleibt jetzt noch eine letzte Implikation zu zeigen. Sei K total beschränkt und vollständig. Wir zeigen, ebenfalls wieder durch einen Widerspruchsbeweis, dass K dann auch überdeckungskompakt ist. Angenommen K wäre nicht überdeckungskompakt, dann existiert eine offene Überdeckung $(U_i)_{i \in I}$ von K , welche keine endliche Teilüberdeckung besitzt. Sei $\eta > 0$. Weil K total beschränkt ist, existieren eine endliche Menge E und Punkte $\psi_i \in K$ mit $i \in E$ und

$$K \subset \bigcup_{i \in E} B(\psi_i, \eta).$$

So existiert aber wenigstens ein $i \in E$, sodass auch $K \cap B(\psi_i, \eta)$ nicht durch endlich viele der U_i überdeckt wird, das heißt auch nicht überdeckungskompakt ist. Denn $(U_i)_{i \in I}$ ist natürlich eine offene Überdeckung von $K \cap B(\psi_i, \eta)$. Oder anders formuliert: Gebe es kein solches $i \in E$, dann wäre die Vereinigung der endlichen Teilüberdeckungen der $K \cap B(\psi_i, \eta)$ eine endliche Teilüberdeckung von K . Auf diese Weise erhalten wir nun iterativ eine Folge $(x_n)_{n \in \mathbb{N}} \subset K$ mit der Eigenschaft, dass $K \cap B(x_n, 1/2^n)$ nicht durch endlich viele der U_i überdeckt wird und

$$K \cap B(x_{n+1}, 1/2^{n+1}) \subset K \cap B(x_n, 1/2^n).$$

Wegen der Inklusion gilt dann insbesondere

$$d(x_n, x_{n+1}) \leq \frac{1}{2^n} \quad \forall n \in \mathbb{N}.$$

Hieraus folgt, dass $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge ist. Da K nach Voraussetzung vollständig ist, konvergiert sie gegen ein $x \in K$. Hierzu gibt es ein $\alpha_x \in I$ mit $x \in U_{\alpha_x}$. Da U_{α_x} offen ist, existiert ein $\varepsilon > 0$, sodass $B(x, \varepsilon) \subset U_{\alpha_x}$. Da $(x_n)_{n \in \mathbb{N}}$ gegen x konvergiert, existiert ein $n_0 \in \mathbb{N}$, so dass $d(x_n, x) < \frac{\varepsilon}{2}$ für alle $n \geq n_0$. Für $y \in B(x_n, 1/2^n)$ mit $n \geq n_0$ folgt aus der Dreiecksungleichung

$$d(y, x) \leq d(y, x_n) + d(x_n, x) \leq \frac{1}{2^n} + \frac{\varepsilon}{2}.$$

Wählen wir daher $n_1 \in \mathbb{N}$ so groß, dass $\frac{1}{2^{n_1}} < \frac{\varepsilon}{2}$, so gilt für alle $n \geq \max\{n_0, n_1\}$ die Inklusion

$$B(x_n, 1/2^n) \cap K \subset B(x, \varepsilon) \subset U_{\alpha_x}.$$

Dies ist nun aber der gesuchte Widerspruch zur Konstruktion der $B(x_n, 1/2^n)$. Damit folgt, dass K überdeckungskompakt ist.

Damit ist alles gezeigt.

q.e.d.

Satz 1.11

Seien (M, d) ein metrischer Raum und $K \subset M$ kompakt. Dann ist jede abgeschlossene Menge $A \subset K$ auch kompakt.

Beweis: Wir zeigen, dass A folgenkompakt ist. Sei $(x_n)_{n \in \mathbb{N}} \subset A$ eine Folge. Da K kompakt ist, also insbesondere folgenkompakt, und $A \subset K$ eine Teilmenge von K ist, existiert eine Teilfolge von $(x_n)_{n \in \mathbb{N}}$, die gegen ein $x \in K$ konvergiert. Da nun A abgeschlossen ist, folgt aus Satz 1.7, dass $x \in A$ ist. Damit ist A folgenkompakt und nach dem Satz von Heine-Borel (Satz 1.10) kompakt.

q.e.d.

Satz 1.12

Sei \mathbb{R}^n mit der Standardmetrik (euklidische Metrik) versehen. Dann ist eine Teilmenge $K \subset \mathbb{R}^n$ genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beweis:

„ \Rightarrow “: Sei K kompakt, insbesondere nach Satz 1.10 und Definition 1.9 vollständig und total beschränkt. Da vollständige Mengen abgeschlossen und total beschränkte Mengen beschränkt sind, folgt die Hin-Richtung.

„ \Leftarrow “: Sei $K \subset \mathbb{R}^n$ beschränkt und abgeschlossen. Dann existiert ein $R > 0$, sodass $K \subset B(0, R)$. Da K abgeschlossen ist, genügt es nach Satz 1.11 zu zeigen, dass $B(0, R)$ kompakt ist. Sei hierzu $(x_n)_{n \in \mathbb{N}} \subset B(0, R)$ eine Folge. Schreiben wir $x_k = (x_k^1, \dots, x_k^n)$, $k \in \mathbb{N}$, so erhalten wir wegen

$$\|x_k\|^2 = \sum_{i=1}^n (x_k^i)^2 \leq R^2$$

insgesamt n beschränkte Folgen in \mathbb{R} . Nach dem Satz von Bolzano-Weierstraß (siehe [MK09], Satz 8.5) existiert daher eine konvergente Teilfolge $(y_n)_{n \in \mathbb{N}}$ von $(x_n)_{n \in \mathbb{N}}$, sodass $(y_n^1)_{n \in \mathbb{N}}$, das heißt, die erste Komponenten der Folge $(y_n)_{n \in \mathbb{N}}$ konvergiert. Nach Auswahl einer weiteren Teilfolge $(z_n)_{n \in \mathbb{N}}$ von $(y_n)_{n \in \mathbb{N}}$ erhält man dann eine Teilfolge, bei der die ersten beiden Koordinatenfolgen $(z_n^1)_{n \in \mathbb{N}}, \dots$ konvergieren. Wir meinen damit, dass wir eine weitere Teilfolge finden können, für die beide Komponentenfolgen konvergieren. Iterativ erhalten wir eine Teilfolge $(\tilde{x}_n)_{n \in \mathbb{N}}$ von $(x_n)_{n \in \mathbb{N}}$, bei der sämtliche Koordinatenfolgen in \mathbb{R} konvergieren. Da eine Folge genau dann bezüglich $\|\cdot\|$ konvergiert, wenn die Koordinatenfolgen in \mathbb{R} konvergieren, ist $(\tilde{x}_n)_{n \in \mathbb{N}}$ somit eine konvergente Teilfolge von $(x_n)_{n \in \mathbb{N}}$. Ferner ist $\|\tilde{x}_k\|^2 \leq R^2$ für alle $k \in \mathbb{N}$ und daher auch mit den Grenzwertsätzen

$$\|\lim_{k \rightarrow \infty} \tilde{x}_k\|^2 = \lim_{k \rightarrow \infty} \|\tilde{x}_k\|^2 \leq R^2.$$

Dies impliziert $\lim_{k \rightarrow \infty} \tilde{x}_k \in B(0, R)$ und daher ist $B(0, R)$ kompakt.

q.e.d.

Satz 1.13

Seien (M, d) ein metrischer Raum und $\mathcal{O} := \{\Omega \subset M : \Omega \text{ ist } d\text{-offen}\}$. Dann ist M mit diesem System \mathcal{O} ein topologischer Raum, der hausdorffsch ist.

Beweis: Den Beweis haben wir eigentlich schon geführt, als wir gezeigt haben, dass der Durchschnitt von endlich vielen und die Vereinigung von beliebig vielen offenen Mengen wieder offen ist (siehe die Sätze 1.1 und 1.2 und deren Beweise). Dennoch führen wir den Beweis der Vollständigkeit halber nochmal, indem wir die Axiome einer Topologie aus Definition 1.19 nachprüfen:

- i) Die leere Menge \emptyset ist in \mathcal{O} enthalten, da es gar kein $x \in \emptyset$ gibt, für das ein $U(x, r) \subset \emptyset$ gefunden werden müsste. Auch die Menge M selbst liegt in \mathcal{O} , da sie Umgebung jedes ihrer Punkte und damit sicherlich offen ist. Dies zeigt, dass Offenheit eben keine intrinsische Eigenschaft ist.
- ii) Seien Ω_1 und Ω_2 offen. Dann existieren ε_1 und ε_2 und $x \in \Omega_1 \cap \Omega_2$ mit $U(x, \varepsilon_1) \subset \Omega_1$ und $U(x, \varepsilon_2) \subset \Omega_2$. Wähle $\varepsilon := \min\{\varepsilon_1, \varepsilon_2\}$. Dann ist auch

$$U(x, \varepsilon) \subset \Omega_1 \cap \Omega_2,$$

also ist $\Omega_1 \cap \Omega_2$ offen und insbesondere ist $\Omega_1 \cap \Omega_2 \in \mathcal{O}$.

- iii) Seien I eine beliebige Indexmenge und für $i \in I$ eine offene Teilmenge Ω_i gegeben. Ist $x \in \bigcup_{i \in I} \Omega_i$, so existiert ein $i \in I$ mit $x \in \Omega_i$ und dann auch ein $\varepsilon > 0$ mit $U(x, \varepsilon) \subset \Omega_i$. Daraus folgt

$$U(x, \varepsilon) \subset \bigcup_{i \in I} \Omega_i.$$

Nehmt euch für den Nachweis der Hausdorff-Eigenschaft zwei verschiedene Punkte $p, q \in M$ und setzt $r := \frac{1}{2}d(p, q) > 0$. q.e.d.

Satz 1.14

Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Dann wird durch

$$d : V \times V \rightarrow \mathbb{K} \text{ mit } d(x, y) := \|x - y\|$$

eine Metrik auf V definiert.

Beweis: Vergleiche Definition 1.1 und Definition 1.26.

q.e.d.

1.3 Erklärungen zu den Definitionen

Zur Definition 1.1 einer Metrik: Wie euch sicherlich in der Analysis 1 schon aufgefallen ist, untersucht man in der Analysis häufig Grenzwertfragen. Hierfür ist ein adäquater Konvergenzbegriff nötig, für den man aber Abstände bzw. Abweichungen messen muss, zum Beispiel wie weit bei einer Funktion der Form $f : M \rightarrow N$ zwischen zwei Mengen M und N die Funktionswerte $f(x)$ und $f(y)$ voneinander entfernt sind bzw. voneinander abweichen, wenn die Abweichungen von x und y in M bekannt sind. Genau dies leistet der Begriff der Metrik. Ihr alle kennt schon eine Metrik, nämlich die sogenannte euklidische Metrik (also den elementargeometrischen Abstandsbegriff), die wir gleich noch einmal in den Beispielen genauer betrachten werden. Nun gibt es aber, wie wir auch noch sehen werden, viele weitere interessante und nützliche Metriken.

Für den einen oder anderen mag die Definition 1.1 einer Metrik vielleicht etwas kompliziert klingen. Die Idee dahinter ist aber mehr als simpel. Füllen wir die Definition doch einmal mit Leben und übersetzen die Sprache der Mathematik in die deutsche Sprache, die ihr alle versteht (hoffen wir jedenfalls :-)). Stellt euch Folgendes vor: Ihr steht gerade an eurer Universität am Haupteingang, zum Beispiel am wunderschönen Welfenschloss der Leibniz Universität Hannover (nein, wir wollen keine Werbung machen :-P). Wir schauen uns nun an, was die einzelnen Axiome i)–iii) in der Definition bedeuten.

- i) Das erste Axiom der positiven Definitheit fordert, dass der Abstand von euch zu irgendeinem Ort immer eine positive reelle Zahl ist und dass der Abstand von euch zu euch selbst null ist.
- ii) Das Axiom der Symmetrie besagt nichts anderes, als dass der Abstand von eurer Universität, zum Beispiel Hannover, zu einer anderen Universität, zum Beispiel München, genauso groß ist, wie der Abstand von der Universität in München nach Hannover.
- iii) Die Dreiecksungleichung kann man sich so verdeutlichen: Wenn ihr auf dem Weg von Hannover nach München einen Umweg über Berlin macht, dann müsst ihr euch nicht wundern, wenn ihr länger unterwegs seid.

Wir bemerken noch: Eigentlich benötigen wir aus dem ersten Axiom nur, dass $d(x, y) = 0 \Leftrightarrow x = y$. Denn daraus folgt dann mit der Dreiecksungleichung und der Symmetrie

$$0 = d(x, x) \leq d(x, y) + d(y, x) = 2 \cdot d(x, y),$$

also $d(x, y) \geq 0$.

Beispiel 1

Wir wetten, dass jeder von euch einen metrischen Raum kennt (sollten wir diese Wette gerade verlieren, so schreibt uns eine Mail mit einem möglichen Wetteinsatz). Dass euch der Begriff der Metrik bzw. des metrischen Raums schon vertraut sein sollte, zumindest wenn ihr Analysis 1 gehört habt, zeigt das folgende erste einfache Beispiel:

- Die Menge der reellen Zahlen ist mit der Abstandsmetrik $d(x, y) := |x - y|$ ein metrischer Raum. Den Beweis habt ihr indirekt schon in [MK09] gelesen (solltet ihr dies getan haben), als wir die Dreiecksungleichung bewiesen und uns einige Eigenschaften des Betrags angeschaut haben.
- Die Metrik von oben kann man sehr leicht auf den $\mathbb{R}^n = \underbrace{\mathbb{R} \times \dots \times \mathbb{R}}_{n\text{-mal}}$ übertragen. Für $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ setzen wir

$$\|x\| := \sqrt{x_1^2 + \dots + x_n^2}.$$

Dann ist $d(x, y) := \|x - y\|$ eine Metrik. Als Übungsaufgabe solltet ihr die Axiome aus Definition 1.1 überprüfen. Die positive Definitheit und die Symmetrie sind fast geschenkt, nur die Dreiecksungleichung erfordert etwas mehr Arbeit. In der Erklärung zur Definition 1.26 der Norm präsentieren wir euch schon eine fast komplette Lösung. Anders geschrieben sieht die Metrik so aus:

$$d(x, y) := \sqrt{\sum_{k=1}^n (x_k - y_k)^2}.$$

Sie definiert auf der Menge $M := \mathbb{R}^n$ mit der Abbildung $d : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$ eine Metrik. Auch sie hat einen besonderen Namen. Wir nennen sie die **euklidische Metrik** und wir schreiben statt d auch oft $|\cdot|$. Der Punkt \cdot soll andeuten, dass wir in $|\cdot|$ etwas einsetzen.

- Sei M eine beliebige Menge. Die Abbildung

$$d : M \times M \rightarrow \mathbb{R},$$

$$d(x, y) := \begin{cases} 0, & \text{für } x = y \\ 1, & \text{für } x \neq y \end{cases}$$

definiert eine Metrik auf M . Man bezeichnet sie als **diskrete Metrik**.

- Sei $M := \mathbb{R}^n$. Die Metrik

$$d(x, y) := \max_{k=1, \dots, n} |x_k - y_k|$$

heißt **Maximumsmetrik**. Wenn wir den Begriff der Norm kennen (siehe Definition 1.26 und die entsprechende Erklärung), so werden wir eine „Verallgemeinerung“ dieser Metrik auf dem Vektorraum der stetigen reellen Funktionen kennenlernen, nämlich die Maximumsnorm.

- Mit dem folgenden Beispiel wollen wir darlegen, wie ihr bei Übungsaufgaben vorgehen müsst, um zu zeigen, dass eine gewisse Abbildung eine Metrik definiert. Es ist klar, was ihr dann zu tun habt. Ihr müsst die positive Definitheit, die Symmetrie und die Dreiecksungleichung nachweisen. Wir zeigen jetzt, dass auf $\mathbb{R}_{>0} := (0, \infty)$ durch

$$d : \mathbb{R}_{>0} \times \mathbb{R}_{>0} \rightarrow \mathbb{R}, \quad d(x, y) := \frac{|x - y|}{xy}$$

eine Metrik definiert wird. Die Positive Definitheit folgt, da $|x - y| > 0$ für alle $x, y \in \mathbb{R}_{>0}$, sowie $xy > 0$, da $x, y \in \mathbb{R}_{>0}$. Außerdem ist ja $|x - y| = 0$ genau dann gleich Null, wenn $x = y$. Die Symmetrie folgt aus der Symmetrie der euklidischen Metrik $|\cdot|$. Die Dreiecksungleichung erfordert etwas Arbeit. Diese sieht man so ein: Zu zeigen ist, dass $d(x, y) \leq d(x, z) + d(z, y)$ für $x, y, z \in \mathbb{R}_{>0}$. Dafür beginnen wir mit der rechten Seite und erhalten

$$\begin{aligned} d(x, z) + d(z, y) &= \frac{|x - z|}{xz} + \frac{|z - y|}{zy} \geq \left| \frac{x - z}{xz} + \frac{z - y}{zy} \right| \\ &= \left| \frac{xy - yz + xz - xy}{xzy} \right| = \frac{|x - y|}{xy} = d(x, y). \end{aligned}$$

Hierbei ging einfach nur die („normale“) Dreiecksungleichung ein.

- Noch ein etwas anderes Beispiel. Wir wollen zeigen, dass

$$d(m, n) := \left| \frac{1}{m} - \frac{1}{n} \right|$$

eine Metrik auf \mathbb{N} definiert. Die Positive Definitheit und die Symmetrie folgen wieder sofort (macht euch das klar!). Die Dreiecksungleichung sieht man so ein (Wir addieren geschickt Null): Seien $m, n, k \in \mathbb{N}$, so ergibt sich

$$\begin{aligned} d(m, n) &= \left| \frac{1}{m} - \frac{1}{n} \right| = \left| \frac{1}{m} - \frac{1}{k} + \frac{1}{k} - \frac{1}{n} \right| = \left| \frac{1}{m} - \frac{1}{k} + \left(\frac{1}{k} - \frac{1}{n} \right) \right| \\ &\leq \left| \frac{1}{m} - \frac{1}{k} \right| + \left| \frac{1}{k} - \frac{1}{n} \right| = d(m, k) + d(k, n). \end{aligned}$$

- Seien (M, d) und (M', d') zwei metrische Räume. Auf dem kartesischen Produkt $M \times M'$ wird eine Metrik, die sogenannte **Produktmetrik** durch

$$\tilde{d}((x, x'), (y, y')) = \max(d(x, y), d'(x', y'))$$

definiert.

- Auf \mathbb{R} werde ein δ definiert durch

$$\delta(x, y) := \arctan |x - y|.$$

Wir wollen beweisen, dass dies wirklich eine Metrik ist. Dazu weisen wir die entsprechenden Axiome nach.

- i) Die Arkus-Tangens-Funktion hat genau eine Nullstelle bei $x = 0$, denn sie ist streng monoton wachsend. Daher gilt für alle $(x, y) \in \mathbb{R}^2$:

$$\delta(x, y) = 0 \Leftrightarrow \arctan |x - y| = 0 \Leftrightarrow |x - y| = 0 \Leftrightarrow x = y.$$

- ii) Für alle $(x, y) \in \mathbb{R}^2$ gilt

$$\delta(x, y) = \arctan |x - y| = \arctan |y - x| = \delta(y, x).$$

- iii) Wir zeigen zunächst, dass $\arctan(x + y) \leq \arctan(x) + \arctan(y)$ für alle $x, y \geq 0$ gilt. Mit der Substitution $z := t - x$ erhalten wir schließlich:

$$\begin{aligned} \arctan(x + y) &= \int_0^{x+y} \frac{1}{1+t^2} dt \\ &= \int_x^{x+y} \frac{1}{1+t^2} dt + \int_x^x \frac{1}{1+t^2} dt \\ &= \arctan(x) + \int_0^y \frac{1}{1+(z+x)^2} dz \\ &\leq \arctan(x) + \int_0^y \frac{1}{1+z^2} dz \\ &= \arctan(x) + \arctan(y). \end{aligned}$$

Damit ist

$$\begin{aligned}
 \delta(x, z) &= \arctan |(x - y) + (y - z)| \\
 &\leq \arctan (|x - y| + |y - z|) \\
 &\leq \arctan |x - y| + \arctan |y - z| \\
 &= \delta(x, y) + \delta(y, z) \quad \forall x, y, z \in \mathbb{R}.
 \end{aligned}$$

■

Zur Definition 1.2 der offenen und abgeschlossenen Kugel: Die Begriffe der offenen und abgeschlossenen Kugel kann man sich relativ leicht im euklidischen Raum bzw. in der Ebene verdeutlichen. Stellt euch in der Ebene einfach einen Kreis mit Mittelpunkt x_0 und Radius $r > 0$ vor. Die offene Kugel beinhaltet alle Punkte, die von dem Mittelpunkt x_0 einen Abstand kleiner als r haben. Sie liegen also alle innerhalb des Kreises und nicht auf dem Rand. Geht man nun zum dreidimensionalen Raum über, so rechtfertigt dies auch den Begriff der Kugel. Analog macht man sich den Begriff der abgeschlossenen Kugel klar und zwar als Kreis bzw. Kugel, bei der der Rand „dazugehört“. Zur abgeschlossenen Kugel sagt man auch oft abgeschlossener Ball. Das kommt daher, dass im Englischen zum Beispiel im Dreidimensionalen begrifflich zwischen „sphere“ (Kugeloberfläche) und „ball“ (Vollkugel) unterschieden werden kann, während im Deutschen bei „Kugel“ es nicht so klar ist. „Ball“ ist also wohl eine Anleihe aus dem Englischen. Graphisch sieht das in etwa so aus wie in der Abbildung 1.1 dargestellt. Das Gestrichelte in Abbildung 1.1 soll andeuten, dass die Randpunkte nicht dazugehören. Wir wollen hier eine kleine Warnung aussprechen:

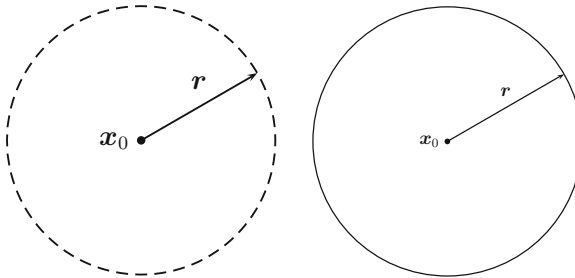


Abb. 1.1: Veranschaulichung einer offenen und abgeschlossenen Kugel.

Man darf sich die offenen bzw. abgeschlossenen Kugeln nicht als Kugeln in dem Sinne vorstellen, wie der Begriff in unserem Sprachgebrauch benutzt wird. Um das zu verdeutlichen, zeichnet doch einmal die Kugeln in der diskreten Metrik auf dem \mathbb{R}^2 . Wir wollen für die Normen (siehe dazu auch Definition 1.26)

$$|\cdot|_1 := \sum_{k=1}^n |x_k|, \quad |\cdot|_2 := \left(\sum_{k=1}^n x_k^2 \right)^{\frac{1}{2}} \quad \text{und} \quad |\cdot|_\infty := \max_{1 \leq k \leq n} |x_k|$$

im \mathbb{R}^2 die Einheitskugeln $B(0, 1) = \{x : |\cdot|_p \leq 1\}$ mit $p = 1, 2, \infty$ skizzieren. Siehe dazu die Abbildungen 1.2, 1.3 und 1.4. Wir arbeiten hier jetzt mit Normen, weil jede Norm auf einen Vektorraum einen metrischen Raum durch $d(x, y) := \|x - y\|$ bildet. Dies wird in der Erklärung zu Definition 1.26 auch deutlich werden.

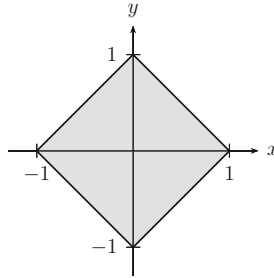


Abb. 1.2: Man sieht hier die „Einheitskugel“ der Metrik, die aus der Norm $\|\cdot\|_1 = \sum_{k=1}^n |x_k|$ hervorgeht. Das ist alles andere als eine Kugel, oder?

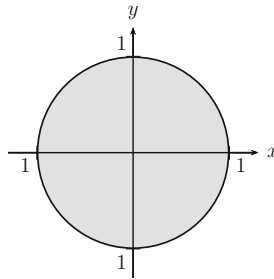


Abb. 1.3: Hier sieht man die „Einheitskugel“ zur Norm $\|\cdot\|_2 = \left(\sum_{k=1}^n x_k^2\right)^{\frac{1}{2}}$.

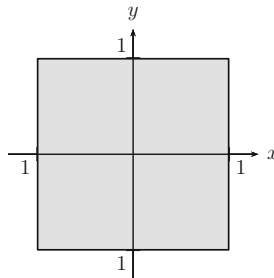


Abb. 1.4: Und schließlich die „Einheitskugel“ zur Norm $\|\cdot\|_\infty = \max_{1 \leq k \leq n} |x_k|$.

Zur Definition 1.3 der Umgebung: Zur Erklärung soll die Abbildung 1.5 dienen. Dies bedeutet also nur, dass man eine Teilmenge $U \subset M$ Umgebung von $x \in M$

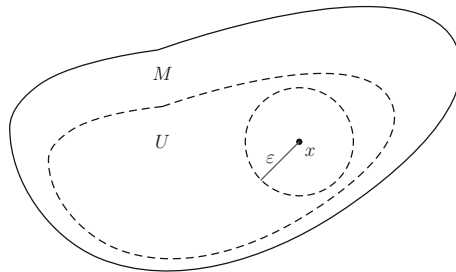


Abb. 1.5: Umgebung U eines Punktes $x \in M$.

nennt, wenn die entsprechende Kugel noch komplett in U enthalten ist.

Zur Definition 1.4 einer offenen Menge: Wie kann man sich die Definition 1.4 vorstellen? Dies ist eigentlich recht einfach, wenn man die Definition 1.3 verstanden hat. Malen wir uns ein Bildchen (siehe Abbildung 1.6) und schauen, was diese Definition aussagt: Sie sagt einfach nur, dass eine Menge M offen heißt, wenn zu jedem Punkt der Menge eine Umgebung in M existiert, das heißt, zu jedem Punkt kann man eine offene Kugel finden, sodass diese noch in der Menge M liegt, egal wie „nah“ der Punkt $x \in M$ am Rand von M liegt. Die leere Menge und die ganze Menge selbst sind beispielsweise offen: Der ganze Raum M ist offen, da M Umgebung jedes Punktes $x \in M$ ist. Die leere Menge ist offen, da es keinen Punkt $x \in \emptyset$ gibt, zu dem es eine ε -Umgebung $U(x, \varepsilon) \subset \emptyset$ geben müsste. Anschaulich ist eine Menge offen, wenn ihre Elemente nur von

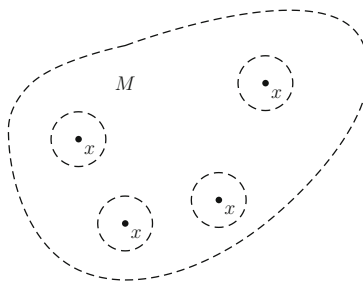


Abb. 1.6: Die Menge M ist offen.

Elementen dieser Menge umgeben sind, das heißt also, kein Element der Menge liegt auf dem Rand der Menge!

Beispiel 2

Aus der Analysis 1 wisst ihr, dass man ein Intervall $(0, 1)$ in den reellen Zahlen offen nennt. Dann sollte dieses Intervall aber natürlich auch wirklich offen sein. Dies gilt, denn jede reelle Zahl mit $0 < x < 1$ ist nur von den Zahlen mit

derselben Eigenschaft umgeben. Um dies einzusehen, wählt man beispielsweise als Umgebung die Menge

$$\left(\frac{x}{2}, \frac{1}{2} + \frac{x}{2}\right).$$

■

Man muss ein wenig mit dem Begriff der Offenheit aufpassen, denn ob eine Menge offen ist oder nicht, hängt ganz stark von dem umgebenden Raum ab, in dem sie liegt. Die rationalen Zahlen zwischen 0 und 1 bilden eine offene Menge in den rationalen Zahlen, aber nicht in den reellen Zahlen!

Beispiel 3

Wir wollen jetzt noch zeigen, dass die offenen Mengen bzgl. der Metrik δ aus dem letzten Punkt in Beispiel 1 dieselben sind wie bzgl. der Metrik $d(x, y) = |x - y|$. Dies geht so: Mit $\tilde{B}(\cdot, r)$ bezeichnen wir eine offene Kugel bzgl. der Metrik δ und mit $B(\cdot, r)$ eine offene Kugel bzgl. der Metrik d . Sei U eine offene Menge bzgl. der Metrik δ und $x \in U$. Dann gibt es nach Definition ein $\varepsilon_1 \in (0, \frac{\pi}{2})$ mit $\tilde{B}(\varepsilon_1, x) \subset U$. Mit $\varepsilon_2 := \tan(\varepsilon_1) > 0$ erhalten wir dann

$$\begin{aligned} B(\varepsilon_2, x) &= \{\xi \in \mathbb{R} : |x - \xi| < \tan(\varepsilon_1)\} \\ &= \{\xi \in \mathbb{R} : \arctan|x - \xi| < \varepsilon_1\} \\ &= \tilde{B}(\varepsilon_1, x) \subset U, \end{aligned}$$

also ist U offen bzgl. der Metrik d . Wieso man dieses ε_2 gerade als $\tan(\varepsilon_1)$ wählt, sieht man also eigentlich erst an der Rechnung, die wir gerade nachvollzogen haben. Man muss also ein wenig mit der Definition der Metrik spielen. Die andere Richtung geht fast genauso. Versucht euch erst dran und schaut erst danach in die folgenden Zeilen :-).

Sei nun umgekehrt U eine offene Menge bzgl. der Metrik d und $x \in U$. Dann existiert $\varepsilon_1 > 0$ mit $B(\varepsilon_1, x) \subset U$. Wir setzen $\varepsilon_2 := \arctan(\varepsilon_1) > 0$ und erhalten sofort

$$\begin{aligned} \tilde{B}(\varepsilon_2, x) &= \{\xi \in \mathbb{R} : \arctan|x - \xi| < \arctan(\varepsilon_1)\} \\ &= \{\xi \in \mathbb{R} : |x - \xi| < \varepsilon_1\} \\ &= B(\varepsilon_1, x) \subset U, \end{aligned}$$

also ist U offen bzgl. der Metrik δ .

■

Zur Definition 1.5 einer abgeschlossenen Menge: Diese Definition bedarf keiner großen Erklärung. Wenn man zeigen möchte, dass eine Menge abgeschlossen ist, zeigt man einfach, dass das Komplement offen ist. Und wie man das macht, werden wir uns noch etwas später bei den Erklärungen zu den Sätzen und Beweisen im Abschnitt 1.4 anschauen. Abgeschlossen ist eine Menge, wenn sie all ihre

Häufungspunkte enthält. Eine wichtige Anmerkung müssen wir noch machen: Wählen wir als Raum X zum Beispiel die Vereinigung der Intervalle $[0,1]$ und $[2,3]$ mit euklidischer Metrik, so sind in diesem Raum die Teilmengen $[0,1]$ und $[2,3]$ beide sowohl offen als auch abgeschlossen. Dabei ist wichtig, zu bemerken, dass Offenheit vom Raum abhängig ist! Außerdem ist in einem metrischen Raum die zugrundeliegende Gesamtmenge ebenso wie die leere Menge stets sowohl offen als auch abgeschlossen. Mengen können also durchaus beide Eigenschaften besitzen! Das heißt: Wenn eine Menge beispielsweise abgeschlossen ist, könnte sie auch noch offen sein! Schließt niemals, dass eine nicht offene Menge abgeschlossen sein muss oder Ähnliches. So ist zum Beispiel das Intervall $(1, 2]$ weder offen noch abgeschlossen. Ein weiteres Beispiel in diesem Zusammenhang.

Beispiel 4

Sei X eine beliebige Menge. Dann wird ja durch $d(x, y) := 0$, falls $x = y$ und $d(x, y) := 1$, falls $x \neq y$ auf X eine Metrik definiert, die sogenannte diskrete Metrik auf X . Jede Teilmenge von X ist bzgl. dieser Metrik zugleich offen und auch abgeschlossen. Es reicht zu zeigen, dass jede Teilmenge von X offen ist, da Abgeschlossenheit über die Komplementbildung definiert ist. Dies folgt aber sofort daraus, dass für alle $x \in X$ gilt:

$$B_{1/2}(x) = \{y \in X : d(x, y) < 1/2\} = \{y \in X : x = y\} = \{x\} \subset X.$$

■

Zur Definition 1.6 der Konvergenz einer Folge: Als Erläuterung stelle man sich Folgendes vor: Sei (M, d) ein metrischer Raum und sei $(x_n)_{n \in \mathbb{N}}$ eine Folge in M . Die Folge $(x_n)_{n \in \mathbb{N}}$ konvergiert gegen $x \in M$, wenn $\lim_{n \rightarrow \infty} d(x_n, x) = 0$. Der Abstand der Folgenglieder zum Grenzwert x selbst wird also immer kleiner und ist im Grenzfall null. Wir schreiben $\lim_{n \rightarrow \infty} x_n = x$. Hierzu sollten wir noch erwähnen, dass sich der Limes auf die Konvergenz bezüglich der euklidischen Metrik in \mathbb{R} bezieht. Die Definition sollte euch natürlich schon aus der Analysis 1 bekannt sein. Nur dass wir jetzt eine „allgemeinere“ Metrik, also einen allgemeineren Abstandsbegriff, verwenden, die aber durchaus die aus Analysis 1 bekannte euklidische Metrik sein kann. So haben wir es jetzt ja bei vielen Definitionen gemacht: Das, was wir schon kannten, verallgemeinert. So funktioniert Mathematik, man möchte immer allgemeinere und bessere Resultate erzielen.

Zu den Definition 1.7–1.9: Wir wollen hierzu nicht viel sagen, sondern nur, dass wir die Definition schon aus der Analysis 1 kennen und dass wir hier wieder eine allgemeine Metrik verwenden. Wir verweisen auf unseren ersten Band [MK09].

Zur Definition 1.10 des Randes: Diese Definition verdeutlicht man sich am besten mithilfe eines Bildes, wie in Abbildung 1.7 versucht. Ein Randpunkt ist also

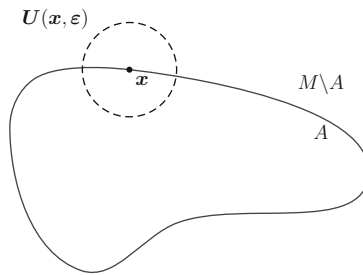


Abb. 1.7: Der Rand einer Menge $A \subset M$.

ein Punkt $x \in M$ von A , wenn der offene Ball um x sowohl mit A als auch mit $M \setminus A$ einen nichtleeren Schnitt besitzt. Vereinigen wir alle Randpunkte, so erhalten wir den Rand der Menge. Man sollte bei dem Begriff des Randes natürlich nicht immer nur Bilder von Gebieten des \mathbb{R}^2 im Kopf haben, denn es kann alles vorkommen: Der Rand ist gleich der Menge, der Rand ist eine echte Teilmenge der Menge, die Menge ist eine echte Teilmenge ihres Randes.

Beispiel 5

- Beispielsweise ist $\partial\mathbb{Q} = \mathbb{R}$, denn in der Umgebung eines jeden Punktes $x \in \mathbb{R}$ liegen sowohl rationale Zahlen als auch irrationale Zahlen. Dies kennen wir aus der Analysis 1; man sagt \mathbb{Q} liegt dicht in \mathbb{R} .
- Der Rand der Einheitskugel im \mathbb{R}^n

$$B(0, 1) = \{x \in \mathbb{R}^n : \|x\| \leq 1\}$$

ist gerade die Einheitssphäre

$$\partial B(0, 1) =: \mathbb{S}^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}.$$

Wir hoffen, dass der Begriff des Randes nun klarer geworden ist

■

Zur Definition 1.11 des Inneren: Das Innere einer Menge ist einfach nur die Menge ohne den Rand. Abbildung 1.8 soll dies noch einmal darstellen.

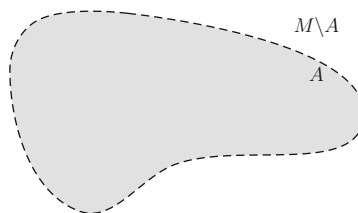


Abb. 1.8: Das Innere einer Menge A .

Zur Definition 1.12 des Abschlusses: Der Abschluss einer Menge A ist einfach nur das Innere und der Rand zusammen. Abbildung 1.9 zeigt dies noch einmal in visueller Form.

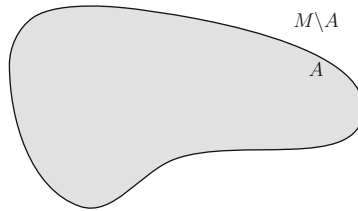


Abb. 1.9: Der Abschluss einer Menge A .

Zur Definition 1.13 der offenen Überdeckung: Die Definition 1.13 der offenen Überdeckung macht man sich am besten mit einem Bild klar, wie in Abbildung 1.10. Wenn es möglich ist, K durch offene Mengen zu überdecken, dann besitzt K eine offene Überdeckung.

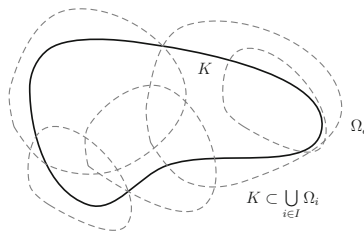


Abb. 1.10: Eine offene Überdeckung einer Menge.

Beispiel 6

Das folgende Beispiel soll eher eine versteckte Übungsaufgabe darstellen: Versucht aus der Überdeckung des offenen Intervalls $(0,1)$ durch Mengen der Form $(\frac{1}{n}, 1 - \frac{1}{n})$ mit $n \in \mathbb{N}$ eine endliche Teilüberdeckung auszuwählen.

Um euch die Übungsaufgabe nicht wegzunehmen, geben wir nun eine andere offene Überdeckung an und zeigen zunächst, dass diese eine offene Überdeckung des Intervalls $(0,1)$ ist und danach, dass diese Überdeckung keine endliche Teilüberdeckung besitzt, die das Intervall $(0,1)$ überdeckt. Hieraus ergibt sich, dass $(0,1)$ nicht überdeckungskompakt und damit nicht kompakt ist. Wir behaupten als erstes, dass

$$\bigcup_{n=1}^{\infty} \underbrace{\left(\frac{1}{n+2}, \frac{1}{n} \right)}_{=: U_n}$$

eine offene Überdeckung vom Intervall $(0, 1)$ ist. Dazu ist zu zeigen, dass für alle $x \in (0, 1)$ ein $n \in \mathbb{N}$ mit $x \in U_n$ existiert. Solch ein n existiert aber mit $\frac{1}{n+2} < x$. Wir verwenden nun das kleinste solche n , also $x \leq \frac{1}{n+1}$. Dann gilt:

$$\frac{1}{n+2} < x \leq \frac{1}{n+1} < \frac{1}{n} \Rightarrow x \in U_n.$$

Hieraus folgt nun, dass $\bigcup U_n$ eine Überdeckung von $(0, 1)$ ist. Dass diese auch noch offen ist, ist klar, da die U_n offene Intervalle sind.

Nun beweisen wir als zweites, dass von dieser offenen Überdeckung keine endliche Teilüberdeckung existiert. Dies ergibt unser Ziel, dass $(0, 1)$ nicht kompakt ist. Dazu wählen wir eine endliche Indexmenge $I \subset \mathbb{N}$ aus. So existiert ein $n \in \mathbb{N}$ mit $n-1, n \notin I$. Wegen

$$\left(\frac{1}{i+1}, \frac{1}{i}\right) \subset (0, 1) \setminus (U_{i-2} \cup U_{i+1}) \quad \text{und} \quad U_j \cap U_{j+i} = \emptyset$$

für jedes $j \in \mathbb{N}$ und $i \geq 2$ ist dann

$$\left(\frac{1}{n+1}, \frac{1}{n}\right) \subset (0, 1) \setminus \bigcup_{i \in I} U_i.$$

Das heißt, es fehlt mindestens das Intervall $\left(\frac{1}{n+1}, \frac{1}{n}\right)$, um $(0, 1)$ zu überdecken. Wir haben also eine offene Überdeckung angegeben, die keine endliche Teilüberdeckung besitzt. Daher ist $(0, 1)$ nicht überdeckungskompakt.

Dass $[0, 1]$ kompakt ist, zeigen wir in Beispiel 7. ■

Zur Definition 1.14 der Überdeckungskompaktheit: In den folgenden Definitionen gibt es einige Kompaktheitsdefinitionen. Der Begriff der Kompaktheit ist schon aus der Analysis 1 bekannt. Ein abgeschlossenes und beschränktes Intervall nannten wir dort (siehe [MK09]) kompakt. Hier sollte man aber etwas vorsichtig sein. Was wir unter Kompaktheit genau verstehen, sehen wir jetzt. Die Definition 1.14 der Überdeckungskompaktheit wird oft falsch verstanden. Überdeckungskompaktheit bedeutet nicht, dass man K überdeckungskompakt nennt, wenn man K durch eine endliche Anzahl von offenen Mengen überdecken kann. Denn das ist sowieso immer möglich, nämlich durch die Menge selbst. Überdeckungskompaktheit bedeutet, dass zu *jeder beliebig vorgegebenen* offenen Überdeckung eine endliche Teilüberdeckung existieren muss. Dies haben wir gerade in Beispiel 6 gesehen: Es ist bei dem Intervall $(0, 1)$ zu jeder offenen Überdeckung eine endliche Teilüberdeckung vorhanden, bei dem Intervall $(0, 1]$ aber beispielsweise nicht, wie wir gleich sehen werden! Dieses Intervall ist auf \mathbb{R} nicht überdeckungskompakt, da die Überdeckung

$$\mathcal{U} := \left\{ \left(\frac{1}{n}, 1 + \frac{1}{n} \right) : n \in \mathbb{N} \right\}$$

zwar offen ist, \mathcal{U} sich jedoch nicht auf eine endliche Teilüberdeckung reduzieren lässt. Beispielsweise kann auch die Überdeckung $\mathcal{U}' := \{[1/n, 1] : n \in \mathbb{N}\}$ nicht auf eine endliche Teilüberdeckung reduziert werden, da, nach der Wegnahme von beliebig vielen Elementen von \mathcal{U}' , noch unendlich viele Elemente benötigt werden, um das Intervall $(0, 1]$ zu überdecken.

Beispiel 7

Wenn man zeigen will, dass das Intervall $[0, 1]$ überdeckungskompakt ist, so muss man zeigen, dass jede offene Überdeckung eine endliche Teilüberdeckung besitzt. Da dies relativ schwierig ist, führt man einen Widerspruchsbeweis und verwendet Intervallschachtelung. Etwas genauer: Wir gehen davon aus, dass eine offene Überdeckung existiere, aus der keine endliche Teilüberdeckung ausgewählt werden kann. Teilen wir nun das Intervall in $[0, 1/2]$ und $[1/2, 1]$, dann kann man wenigstens eines dieser Intervalle nicht durch endlich viele Elemente aus der offenen Überdeckung überdecken. Sei zum Beispiel $[0, 1/2]$ dieses Intervall, dann kann man auch dieses wieder teilen, und eine Hälfte können wir wieder nicht voll überdecken. Dieses setzen wir fort und wir erhalten eine Intervallschachtelung. Nun nutzen wir diese, um einen Widerspruch zu konstruieren.

Genauer geht dies so (wir haben diesen Beweis in [KÖ9], Seite 315, gefunden): Angenommen, $\{I_i\}$ sei eine offene Überdeckung des Intervalls $[0, 1]$ derart, dass je endlich viele der I_i das Intervall nicht überdecken. Ausgehend von irgendeinem Intervall $[a_1, b_1]$ mit $[0, 1] \subset [a_1, b_1]$ kann durch sukzessives Halbieren eine Intervallschachtelung konstruiert werden, deren sämtliche Intervalle $[a_n, b_n]$ die Eigenschaft besitzen, dass $[0, 1] \cap [a_n, b_n]$ nicht durch endlich viele der I_i überdeckt wird.

Sei a der durch diese Intervallschachtelung definierte Punkt und α_n irgendein Punkt in $[0, 1] \cap [a_n, b_n]$. Dann ist a der Grenzwert der Folge (α_n) . Nun liegt auch a in $[0, 1]$. Folglich gibt es ein offenes Intervall I der Überdeckung mit $a \in I$. Für hinreichend großes N gilt damit $[a_n, b_n] \subset I$. Dies widerspricht aber der Eigenschaft, dass $[0, 1] \cap [a_n, b_n]$ nicht durch endlich viele der I_i überdeckt wird.

Ein anderer sehr kurzer Beweis, der uns auf dem Matheplaneten über den Weg gelaufen ist (<http://www.matheplanet.com>), ist der folgende: Es sei eine offene Überdeckung von $[0, 1]$ gegeben. Sei $a \in [0, 1]$ das Supremum aller Zahlen, für die das Intervall $[0, a]$ von endlich vielen der gegebenen offenen Mengen überdeckt werden kann. Die Zahl a selbst muss auch von irgendeiner der gegebenen offenen Mengen überdeckt werden. Der Fall $a < 1$ führt zu einem Widerspruch: Es gibt dann ein Intervall (b, c) , das a enthält und Teilmenge von einer Überdeckungsmenge ist (das folgt aus der Überdeckungseigenschaft und aus der Offenheit der überdeckenden Mengen). Somit ist $b < a < c \leq 1$, und es gibt endlich viele Mengen, die das Intervall $[0, b]$ überdecken, und nun haben wir eine weitere Menge, die auch das Intervall (b, c) noch überdeckt, und das bedeutet: Für jedes d mit

$a < d < c$ kann das Intervall $[0, d]$ von endlich vielen Mengen überdeckt werden. Dies ergibt wegen $d > a$ einen Widerspruch zur Definition von a . Also muss $a = 1$ sein. Dies hatten wir zu zeigen. ■

Zur Definition 1.16 der totalen Beschränktheit: Die Definition der totalen Beschränktheit klingt am Anfang vielleicht erst einmal recht kompliziert. Aber schauen wir uns genau an, was dort eigentlich steht und gesagt wird. Dort steht eigentlich nichts anderes, als dass man eine Menge K total beschränkt nennt, wenn es möglich ist, K durch eine endliche Anzahl an offenen Bällen mit vorgegebenem beliebigem Radius zu überdecken. Die Abbildung 1.11 soll zum Verständniss beitragen. Während bei der offenen Überdeckung (siehe Definition 1.13) beliebig viele offene Bälle die Menge überdecken können, wird hier gefordert, dass endlich viele ausreichen müssen. Totalbeschränktheit stellt eine

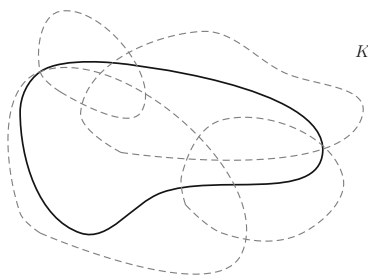


Abb. 1.11: Zur Erklärung des Begriffes der totalen Beschränktheit.

bestimmte Endlichkeitsbedingung an einen metrischen Raum dar. Da es sehr schwierig ist globale Eigenschaften von metrischen Räumen zu erhalten, gibt es den Begriff der Totalbeschränktheit. Dieser umgeht nämlich das Problem, dass eine Metrik durch eine äquivalente Metrik ersetzt werden kann, mit der der Raum dann endlichen Durchmesser besitzt. Die Totalbeschränktheit fordert nun, dass man den Raum in endlich viele Stücke unterteilen kann, von denen jedes eine vorgegebene Größe nicht überschreitet.

Zur Definition 1.17 der Beschränktheit: Bei der Definition der Beschränktheit gibt es einen Unterschied zur totalen Beschränktheit (siehe Definition 1.16). Zum Beispiel ist ein unendlicher Raum mit diskreter Metrik zwar beschränkt, aber niemals total beschränkt.

Zur Definition 1.18 der Kompaktheit: Diese Definition sagt, dass es verschiedene Kompaktheitsbegriffe gibt. Wir müssen aber keine Angst haben. Zumindest im \mathbb{R}^n sind nach dem Satz von Heine-Borel (siehe Satz 1.10) alle Kompaktheitsbegriffe, die wir so kennen, äquivalent. Im \mathbb{R}^n ist eine Menge genau dann kompakt, wenn sie abgeschlossen und beschränkt ist.

Beispiel 8

Wir wollen zeigen: Die Vereinigung von endlich vielen kompakten Teilmengen eines metrischen Raums ist wieder kompakt. Dies bewerkstelligen wir, indem wir die Überdeckungskompaktheit verwenden: Seien M ein metrischer Raum und A_1, \dots, A_n kompakte Teilmengen von M . Weiter seien $A := \bigcup_{i=1}^n A_i$ und $(\Omega_i)_{i \in \mathbb{N}}$ eine offene Überdeckung von A . Da

$$A \subset \bigcup_{i \in \mathbb{N}} \Omega_i \Rightarrow A_i \subset \bigcup_{i \in \mathbb{N}} \Omega_i \quad \forall i = 1, \dots, n,$$

ist $(\Omega_i)_{i \in \mathbb{N}}$ eine offene Überdeckung für jede Menge A_i mit $i = 1, \dots, n$. Da nach Voraussetzung aber alle A_i kompakt sind, gilt für alle $i \in \{1, \dots, n\}$, dass ein $M_i \subset \mathbb{N}$ mit $|M_i| < \infty$ und $A_i \subset \bigcup_{j \in M_i} \Omega_j$ existiert. Da nun

$$A = \bigcup_{i=1}^n A_i \subset \bigcup_{i=1}^n \bigcup_{j \in M_i} \Omega_j = \bigcup_{j \in \bigcup_{i=1}^n M_i} \Omega_j$$

und $|M_i| < \infty$ für alle $i \in \{1, \dots, n\}$, ist auch $|\bigcup_{i=1}^n M_i| < \infty$, also ist A kompakt. ■

Nun zu einem anderen Beispiel, bei dem wir mit der Folgenkompaktheit arbeiten werden.

Beispiel 9

Seien K und L kompakte Teilmengen von \mathbb{R}^n . Wir wollen nun beweisen, dass die Menge

$$K + L := \{x + y : x \in K, y \in L\}$$

dann ebenfalls kompakt ist. Dies geht so: Nach der Definition 1.15 der Folgenkompaktheit ist nur zu zeigen, dass jede Folge $(x_i)_{i \in \mathbb{N}}$ aus $K + L$ eine Teilfolge $(x_{i_k})_{k \in \mathbb{N}}$ besitzt, die gegen ein $c \in K + L$ konvergiert. Sei also $(x_i)_{i \in \mathbb{N}}$ eine beliebige Folge aus $K + L$. Wir können jedes Folgenglied x_i darstellen als

$$x_i = a_i + b_i \quad \text{mit } a_i \in K, b_i \in L.$$

$(a_i)_{i \in \mathbb{N}}$ ist eine Folge aus K und $(b_i)_{i \in \mathbb{N}}$ eine Folge aus L . Da K kompakt ist, existiert nach dem Satz von Bolzano-Weierstraß eine Teilfolge $(a_{i_k})_{k \in \mathbb{N}}$ von $(a_i)_{i \in \mathbb{N}}$, die gegen ein $a \in K$ konvergiert. Nun ist $(b_{i_k})_{k \in \mathbb{N}}$ eine Folge aus L und da L kompakt ist, gibt es eine Teilfolge $(b_{i_{k_l}})_{l \in \mathbb{N}}$ von $(b_{i_k})_{k \in \mathbb{N}}$, die gegen ein $b \in L$ konvergiert. Also haben wir:

- (1) $(b_{i_{k_l}})_{l \in \mathbb{N}}$ ist eine Teilfolge von $(b_i)_{i \in \mathbb{N}}$, die gegen ein $b \in L$ konvergiert.
- (2) $(a_{i_k})_{k \in \mathbb{N}}$ ist eine Teilfolge von $(a_i)_{i \in \mathbb{N}}$, die gegen ein $a \in K$ konvergiert, und damit ist auch $(a_{i_{k_l}})_{l \in \mathbb{N}}$ eine Teilfolge von $(a_i)_{i \in \mathbb{N}}$, die gegen ein a konvergiert.

Mit den Grenzwertsätzen folgt für die Teilfolge $(x_{i_{k_l}})_{l \in \mathbb{N}} = (a_{i_{k_l}} + b_{i_{k_l}})_{l \in \mathbb{N}}$ von (x_i)

$$\lim_{l \rightarrow \infty} x_{i_{k_l}} = \lim_{l \rightarrow \infty} (a_{i_{k_l}} + b_{i_{k_l}}) = \lim_{l \rightarrow \infty} a_{i_{k_l}} + \lim_{l \rightarrow \infty} b_{i_{k_l}} = a + b \in K + L.$$

Daraus folgt, dass $K + L$ kompakt ist. ■

Zur Definition 1.19 eines topologischen Raums: Schauen wir uns an, was die einzelnen Axiome *i)* – *iii)* aus der Definition eines topologischen Raums bedeuten.

Axiom i) sagt aus, dass die leere Menge und die Menge selbst zum topologischen Raum gehören.

Axiom ii) besagt: Sind zwei Mengen Elemente der Topologie, so gehört auch deren Durchschnitt dazu. Der Durchschnitt zweier offener Menge ist also wieder offen. Dies werden wir in Satz 1.1 für metrische Räume sehen.

Axiom iii) bedeutet nichts anderes, als dass die Vereinigung von beliebig vielen offenen Mengen wieder offen ist. Dies haben wir in Satz 1.2 für metrische Räume gezeigt.

Ein topologischer Raum oder eine Topologie ist also ein System von Teilmengen bzw. von offenen Mengen. Wir wollen uns noch ein paar Beispiele für Topologien bzw. topologischen Räumen ansehen.

Beispiel 10

- Das einfachste Beispiel eines topologischen Raums ist die Menge der reellen Zahlen mit folgender Topologie: Das System der offenen Teilmengen ist so erklärt, dass wir eine Menge $\Omega \subset \mathbb{R}$ offen nennen, wenn sie sich als Vereinigung von offenen Intervallen darstellen lässt.
- **Trivialtopologie:** Jede Menge M kann auf wenigstens zwei Arten zu einem topologischen Raum gemacht werden. Dazu definieren wir die **Klumpentopologie** $\mathcal{O}_1 := \{\emptyset, M\}$ und die **diskrete Topologie** $\mathcal{O}_2 := \mathcal{P}(M)$. In anderer Literatur wird \mathcal{O}_1 auch häufig die triviale oder indiskrete Topologie genannt. Der aufmerksame Leser möge überprüfen, ob die Axiome einer Topologie wirklich erfüllt sind.
- **Relativtopologie:** Ist (M, \mathcal{O}) ein topologischer Raum und $N \subset M$ eine Teilmenge, so induziert \mathcal{O} auch eine Topologie auf der Teilmenge N und zwar durch

$$\mathcal{O}|_N := \{\Omega \subset M : \Omega = N \cap \tilde{\Omega} \text{ mit } \tilde{\Omega} \in \mathcal{O}\}.$$

- **Produkttopologie:** Seien (M_1, \mathcal{O}_1) und (M_2, \mathcal{O}_2) zwei topologische Räume. Dann existiert auf

$$M := M_1 \times M_2 := \{(x_1, x_2) : x_1 \in M_1 \text{ und } x_2 \in M_2\}$$

eine induzierte Topologie $\mathcal{O}_1 \times \mathcal{O}_2$, die von der Basis

$$\{\Omega_1 \times \Omega_2 : \Omega_1 \in \mathcal{O}_1 \text{ und } \Omega_2 \in \mathcal{O}_2\}$$

erzeugt wird, die man Produkttopologie nennt. Das heißt $\mathcal{O}_1 \times \mathcal{O}_2$ besteht aus allen denjenigen Teilmengen von $M_1 \times M_2$, die sich als Vereinigung von kartesischen Produkten offener Mengen aus M_1 mit offenen Mengen aus M_2 darstellen lassen.

- **Allgemeine Produkttopologie:** Mit dem Begriff der Subbasis (siehe Definition 1.25) kann man die Produkttopologie verallgemeinern: Sei I eine Menge und für jedes $i \in I$ sei (M_i, \mathcal{O}_i) ein topologischer Raum. Dann nennt man die von der Subbasis

$$B := \left\{ \prod_{i \in I} L : (\forall i \in I : L_i \in \mathcal{O}_i) \wedge (\exists i_0 \in I : \forall i \neq i_0 : L_i = M_i) \right\}$$

definierte Topologie $\prod_{i \in I} \mathcal{O}_i$ die Produkttopologie auf $\prod_{i \in I} M_i$ bezüglich der gegebenen Räume (M_i, \mathcal{O}_i) . Sie ist also die Familie aller derjenigen Produkte offener Mengen aus den jeweiligen Räumen, bei denen höchstens ein Faktor nicht gleich dem jeweiligen Gesamttraum ist.

- **Sierpinski-Raum:** Sei $M := \{0, 1\}$. Dann ist

$$\mathcal{O} := \{\emptyset, M, \{0\}\}$$

eine Topologie auf M . Der topologische Raum (M, \mathcal{O}) heißt der Sierpinski-Raum.

- **Euklidische Topologie:** Sei $M := \mathbb{R}^n$ der euklidische Standardraum und d die übliche euklidische Metrik. Die dadurch erzeugte Topologie nennt man die euklidische Topologie. Diese euklidische Topologie auf \mathbb{R} ist natürlich dieselbe wie die in Beispiel 1.

■

Vielleicht zum Abschluss der Erklärung noch ein paar Worte: Die Topologie stellt im Grunde einen noch allgemeineren, noch abstrakteren Kontext für die für Metriken definierten Begriffe dar, wie beispielsweise offene, abgeschlossene Mengen und so weiter. Wenn man sich die Definition der Topologie anschaut und mit den Sätzen 1.1 und 1.2 über offene Mengen vergleicht, werden einen die Parallelen geradezu anspringen, und der clevere Leser bekommt vielleicht schon eine Idee, was genau eine Topologie macht. Eine Topologie verzichtet einfach

gänzlich auf eine Metrik und arbeitet nur noch auf den offenen Teilmengen, die ohne den Begriff der Metrik neu definiert werden müssen. Dies tun wir in Definition 1.20. Eine Teilmenge von M ist bezüglich der Topologie offen genau dann, wenn sie Element der Topologie ist.

Wir betonen hier noch einmal, dass es sich hierbei um eine Definition und keinen Satz handelt. Das heißt, wenn ihr eine Topologie gegeben habt, müsst ihr euch keine Gedanken darüber machen, ob die Elemente der Topologie offen sind. Das sind sie per Definition! Und so lässt sich auch verstehen, warum eine Topologie eine Metrik induziert und umgekehrt. Habt ihr eine Topologie definiert, legt ihr automatisch fest, welche Mengen offen und damit auch welche Mengen abgeschlossen zu sein haben. Findet ihr nun eine Metrik, für die dieselben Mengen nach Definition 1.4 offen sind, habt ihr die von der Topologie induzierte Metrik gefunden. Der umgekehrte Weg ist meist einfacher. Habt ihr eine Metrik gegeben, dann müsst ihr einfach nur alle Mengen finden, die nach Definition 1.4 offen sind, und in einer Obermenge sammeln. Diese Obermenge stellt dann die durch die Metrik induzierte Topologie dar. Die drei Axiome aus Definition 1.19 bekommt ihr hierbei geschenkt.

Zu den Definitionen 1.20, 1.21 und 1.22: Diese Definitionen wollen wir nicht weiter erklären. Die Vorstellung ist in topologischen Räumen sowieso schwer. Daher haben wir die Definitionen der offenen und abgeschlossenen Menge bzw. der Umgebung zuerst für metrische Räume eingeführt, da dort eine Metrik (Abstandsfunktion) existiert, die uns eine gute Vorstellung gibt. Vergleiche also zur Erklärung auch die Definitionen 1.4, 1.5 und 1.3. Wir bemerken: In anderen Büchern werden meistens die Begriffe zuerst für topologische Räume definiert und dann auf metrische Räume übertragen.

Zur Definition 1.23 des Zusammenhangs: Hier bemerken wir nur: Es gibt noch eine dazu äquivalente Definition des Zusammenhangs: Sind $\Omega_1, \Omega_2 \in \mathcal{O}$ zwei Mengen mit $\Omega_1 \cap \Omega_2 = \emptyset$ und $\Omega_1 \cup \Omega_2 = M$, so nennt man M zusammenhängend, wenn entweder $\Omega_1 = \emptyset$, $\Omega_2 = M$ oder $\Omega_1 = M$, $\Omega_2 = \emptyset$ gilt.

Zur Definition 1.25 der Subbasis: Wir bemerken hier die Beziehung zwischen Basis (Definition 1.24) und Subbasis: Jede Basis ist eine Subbasis, aber nicht umgekehrt. Mehr wollen wir dazu nicht sagen, denn so wichtig ist es nun auch nicht für die Analysis, sondern nur, wenn ihr Topologie vertieft.

Zur Definition 1.26 einer Norm: Für eine Vielzahl metrischer Räume gilt, dass sie zugleich Vektorräume sind und dass die Metrik durch eine Norm definiert ist. Zur Verdeutlichung der Norm-Definition ein paar Beispiele:

Beispiel 11

- Der Standardvektorraum \mathbb{R}^n mit der Norm

$$||x|| := \sqrt{\sum_{k=1}^n x_k^2}$$

und $x := (x_1, \dots, x_n)$ ist ein normierter Vektorraum. Diese Norm $|| \cdot ||$ nennt man die **euklidische Norm**.

- Für endlich-dimensionale Räume K^n sind die **p -Normen** definiert als

$$||x||_p := \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}.$$

Dabei ist $p \geq 1$ eine reelle Zahl. Wieso können wir hier nicht fordern, dass $p < 1$ ist? Na ja, dafür lassen sich keine Normen definieren, wie ihr euch einmal überlegen solltet. Die p -Norm ist eine Verallgemeinerung von dem, was wir schon kennen, denn offensichtlich gilt:

$$||x||_1 = \sum_{i=1}^n |x_i| \quad \text{und} \quad ||x||_2 = \sqrt{\sum_{i=1}^n |x_i|^2}.$$

- Wir betrachten für ein reelles Intervall $[a, b] \subset \mathbb{R}$ den Vektorraum der stetigen Funktionen $f : [a, b] \rightarrow \mathbb{R}$. Dann definiert man die **p -Norm** durch

$$||f||_p := \left(\int_a^b |f(x)|^p \right)^{\frac{1}{p}}.$$

Die Dreiecksungleichung beispielsweise zeigt man mithilfe der Minkowski-Ungleichung. Diese Begrifflichkeiten führen später in der Funktionalanalysis zum Begriff des Sobolev-Raums. Also noch genug Stoff für weitere Bücher :-). Bis dahin ist es aber noch ein Stückchen...

- Auf dem \mathbb{R}^n können noch weitere Normen definiert werden, zum Beispiel die **Maximumsnorm** durch

$$||x||_\infty := \max\{|x_1|, \dots, |x_n|\}$$

für $(x_1, \dots, x_n) \in \mathbb{R}^n$. Die Maximumsnorm ergibt sich aus der p -Norm durch $p \rightarrow \infty$.

- Seien M eine beliebige Menge und $B(M)$ der Vektorraum aller beschränkten reellwertigen Funktionen auf der Menge M . Dann ist durch

$$||f||_\infty := \sup\{|f(x)| : x \in M\} < \infty$$

eine Norm definiert, die man **Supremumsnorm** nennt. Weisen wir dies einmal nach. Die Eigenschaften *i*) (positive Definitheit) und *ii*) (Homogenität) aus Definition 1.26 sind trivial und geschenkt. Nur die Dreiecksungleichung macht etwas Arbeit:

$$\begin{aligned} \|f + g\|_\infty &= \sup\{|f(x) + g(x)| : x \in M\} \leq \sup\{|f(x)| + |g(x)| : x \in M\} \\ &\leq \sup\{|f(x)| : x \in M\} + \sup\{|g(x)| : x \in M\} \\ &= \|f(x)\|_\infty + \|g(x)\|_\infty. \end{aligned}$$

Man sollte sich jeden Schritt noch einmal klar machen, wieso er tatsächlich gilt und nicht einfach nur runterschreiben, ohne zu reflektieren. Bei der ersten Ungleichheit beispielsweise haben wir die Dreiecksungleichung des Betrages ausgenutzt.

Dies soll an Beispielen genügen. ■

Wir wollen jetzt noch die Begriffe der Norm und der Metrik in Beziehung setzen. Bei der Normdefinition hat man das Homogenitätsaxiom. Als Menge benötigen wir hier also einen linearen Raum, während man die Metrik auch für allgemeinere (metrische) Räume definieren kann. Aus jeder Norm lässt sich also eine Metrik erstellen, umgekehrt jedoch nicht (nicht einmal in linearen Räumen).

Zur Definition 1.27 der Normenäquivalenz: In einem etwas aufwendigen Beweis kann man zeigen, dass alle Normen im \mathbb{R}^n äquivalent sind (siehe für einen Beweis zum Beispiel [Forster]). Dies gilt für andere Vektorräume aber im Allgemeinen natürlich nicht! Die Normäquivalenz definiert man gerade so wie man sie definiert, weil daraus ein paar nette Dinge folgen, beispielsweise ein gleicher Konvergenzbegriff, gleiche Topologien etc. Wieso? Ganz einfach: Sind zwei Normen äquivalent, so bilden die ε -Kugeln bzgl. der einen Norm eine Umgebungsbasis (Ein System von Umgebungen eines Punktes x heißt *Umgebungsbasis* von x , wenn jede Umgebung U von x eine Umgebung aus diesem System der Umgebungen enthält). Demnach sind offene Mengen bzgl. der einen Norm auch offen bzgl. der anderen. Da solche Dinge wie Kompaktheit, Konvergenz, Stetigkeit, und so weiter nur von der Topologie abhängen, gelten sie dann auch bzgl. der äquivalenten Norm, wenn sie bzgl. einer Norm gelten.

Beispiel 12

Als Übung solltet ihr einmal die folgende Ungleichungskette für die Normen $\|x\|_p$ im \mathbb{R}^n beweisen:

$$\|x\|_\infty \leq \|x\|_p \leq \|x\|_1 \leq \sqrt{n}\|x\|_2 \leq n\|x\|_\infty \quad (x \in \mathbb{R}^n) \quad (1.1)$$

■

Beispiel 13

Wenn Normen äquivalent sind (und dies ist auf dem \mathbb{R}^n ja nach unserer obigen Bemerkung für alle Normen erfüllt), so gibt es ein paar nette Eigenschaften. Zum Beispiel, dass offene Mengen bzgl. der einen Norm auch offen bzgl. der anderen Norm sind. Dazu wollen wir ein Beispiel angeben! Wir behaupten, dass im \mathbb{R}^n die offenen Mengen bzgl. der euklidischen Norm $\|\cdot\|_2$ dieselben offenen Mengen liefert wie bzgl. der Maximumsnorm $\|\cdot\|_\infty$. Um dies zu zeigen, bezeichnen wir mit $B(x_0, r) := \{x \in \mathbb{R}^n : \|x - x_0\|_1 < r\}$ die offenen Kugeln bzgl. der euklidischen Norm und mit $\tilde{B}(x_0, r) := \{x \in \mathbb{R}^n : \|x - x_0\|_\infty < r\}$ diese bzgl. der Maximumsnorm. Es gilt dann:

$$\tilde{B}(x_0, r/\sqrt{n}) \subset B(x_0, r) \subset \tilde{B}(x_0, r).$$

$B(x_0, r) \subset \tilde{B}(x_0, r)$ haben wir uns mehr oder weniger für einen Spezialfall im \mathbb{R}^2 schon in Abbildung 1.3 und 1.4 überlegt. $\tilde{B}(x_0, r/\sqrt{n}) \subset B(x_0, r)$ folgt aus der Ungleichungskette (1.1). ■

Zur Definition 1.28 eines Banach-Raums: Die Definition mag am Anfang etwas kompliziert klingen. Schauen wir uns ein paar Beispiele an.

Beispiel 14

Wir behaupten, dass $(\mathbb{R}^n, \|\cdot\|_\infty)$ ein Banach-Raum ist.

Beweis: Es ist klar, dass $(\mathbb{R}^n, \|\cdot\|_\infty)$ einen normierten Vektorraum bildet. Bleibt nur noch die Vollständigkeit zu zeigen. Hierbei benutzen wir, dass $(\mathbb{R}, |\cdot|)$ ($|\cdot|$ bezeichnet euklidische Metrik bzw. die induzierte Norm) vollständig und damit ein Banach-Raum ist. Hieraus schließen wir jetzt, dass $(\mathbb{R}^n, \|\cdot\|_\infty)$ ebenfalls einen Banach-Raum bildet, denn eine Folge im \mathbb{R}^n ist genau dann bezüglich $\|\cdot\|$ eine Cauchy-Folge, wenn die Koordinatenfolgen Cauchy-Folgen in $(\mathbb{R}, |\cdot|)$ sind. q.e.d. ■

Beispiel 15

Seien $D \subset \mathbb{R}$ eine beliebige Menge und

$$B(D) := \{f : D \rightarrow \mathbb{R} : \|f\| = \sup |f(x)| < \infty\}$$

der Raum der beschränkten Funktionen mit der Supremumsnorm ausgestattet. Dann bildet $(B(D), \|\cdot\|)$ einen Banach-Raum.

Beweis: Wir zeigen nur die Vollständigkeit. Der Rest sollte klar sein. Sonst zum Beispiel uns fragen :-)! Sei $(f_n)_{n \in \mathbb{N}} \subset B(D)$ eine Cauchy-Folge bezüglich der Norm $\|\cdot\|$. Zu $\varepsilon > 0$ existiert ein $N \in \mathbb{N}$, sodass für alle $x \in D$ gilt:

$$|f_n(x) - f_m(x)| < \varepsilon \quad \forall n, m \geq N. \quad (1.2)$$

Für jedes feste $x \in D$ bildet dann die Folge $(f_n(x))_{n \in \mathbb{N}}$ eine Cauchy-Folge in \mathbb{R} , die wegen der Vollständigkeit der reellen Zahlen gegen eine reelle Zahl konvergiert. Damit existiert eine Funktion $f : D \rightarrow \mathbb{R}$, welche punktweiser Limes von $(f_n)_{n \in \mathbb{N}}$ ist. Bildet man in der obigen Ungleichung (1.2) den Grenzübergang $m \rightarrow \infty$, so ergibt sich

$$\|f_n - f\| < \varepsilon \quad \forall n \geq N.$$

Dies bedeutet gerade, dass f_n gleichmäßig gegen f konvergiert. Aus der Dreiecksungleichung folgt auch die Beschränktheit von f , denn

$$\|f\| = \|f - f_N + f_N\| \leq \|f - f_N\| + \|f_N\| < \varepsilon + \|f_N\| < \infty.$$

Damit ist alles gezeigt. Hierbei haben wir benutzt, dass der Limes gleichmäßig konvergenter stetiger Funktionen wieder stetig ist, siehe [MK09], Kapitel 13. q.e.d. ■

Beispiel 16

Seien $M \subset \mathbb{R}$ eine kompakte Menge und

$$\mathcal{C}^k(M) := \{f : M \rightarrow \mathbb{R} : f \text{ ist } k\text{-mal stetig differenzierbar}\}.$$

Auf $\mathcal{C}^k(M)$ definieren wir die Norm

$$\|f\| := \sum_{i=0}^k \|f^{(i)}\|_{\infty}.$$

Da stetige Funktionen auf kompakten Mengen ihr Supremum bzw. Infimum annehmen (siehe Kapitel 2, Satz 2.4), ist die Norm endlich. Man kann zeigen, dass $(\mathcal{C}^k(M), \|\cdot\|)$ einen Banach-Raum bildet.

Beweis: Wir wollen euch dies als Übungsaufgabe überlassen. Dem interessierten Leser verraten wir den Beweis auf Mail-Anfrage. Tipp: Induktion kann helfen :-)! q.e.d. ■

Beispiel 17

Die Riemann-integrierbaren Funktionen auf einem kompakten Intervall bilden mit der Supremumsnorm einen Banach-Raum.

Beweis: Riemann-integrierbare Funktionen sind beschränkt. Wir wissen das aus der Analysis 1. Damit folgt aus Beispiel 15, dass eine Cauchy-Folge gleichmäßig gegen eine beschränkte Funktion f konvergiert. Da der gleichmäßige Limes Riemann-integrierbarer Funktionen wieder Riemann-integrierbar ist, ist die Vollständigkeit gezeigt und damit die Behauptung bewiesen. q.e.d. ■

An den Beispielen haben wir gesehen, wieso Banach-Räume so interessant und wichtig sind. Diese sind halt vollständig und dies hat immer wieder nette Vorteile, die man ausnutzen kann, wie auch in einigen Beweisen bestimmt auffallen wird.

Zur Definition 1.29 der Exponentialfunktion von linearen Abbildungen: Es scheint auf den ersten Blick etwas ungewöhnlich, dass man in die Exponentialfunktion auch irgendwie Matrizen einsetzen kann (Den Brückenschlag von linearen Abbildungen zu Matrizen kennt ihr aus der Linearen Algebra 1, denn dort hatten wir gesehen, dass man jeder linearen Abbildung zwischen endlich-dimensionalen Vektorräumen eine Matrix, die sogenannte Darstellungsmatrix, zuordnen kann). In unserer Definition haben wir sogar unendlichdimensionale Vektorräume. Hier klappt dies aber trotzdem mit der Basiszuordnung. Aber ihr habt ja gesehen, dass man es durchaus sinnvoll definieren kann. Wir geben ein paar Eigenschaften ohne Beweis an, wobei der Leser natürlich, wie bei allen unbewiesenen Aussagen hier, aufgefordert ist, sich einen Beweis zu überlegen.

- Die Reihe für e^{tA} konvergiert für jedes $t \in \mathbb{R}$ absolut.
- Die Funktion $t \mapsto e^{tA}$ ist differenzierbar, damit auch stetig in \mathbb{R} , und es gilt für die Ableitung

$$(e^{tA})' = Ae^{tA}.$$

- Man kann zeigen, dass

$$e^{(s+t)A} = e^{sA} \cdot e^{tA}, \quad s, t \in \mathbb{R} \quad \text{und} \quad e^0 = \text{Id}_V.$$

- Ist B auch ein beschränkter Homomorphismus, so gilt $AB = BA$ genau dann, wenn

$$e^{tA} e^{tB} = e^{t(A+B)} \quad \text{mit } t \in \mathbb{R}.$$

Zusammenhang Topologie, Metrik, Norm und Skalarprodukt: Wir wollen in diesem Abschnitt noch einmal den Zusammenhang zwischen der Topologie, der Metrik und der Norm (und auch der Vollständigkeit halber des Skalarprodukts, das erst in Kapitel 9 bzw. 10 eingeführt wird) erklären und erläutern, denn diese Begriffe werden immer wieder verwechselt, bzw. man liest Sätze wie „die durch die Metrik induzierten Topologie“ oder ähnliches, weiß aber gar nicht genau, was dahinter steckt. Ein wenig hatten wir zwar immer schon bei den jeweiligen Definition 1.19, 1.1 und 1.26 gesagt, aber wir wollen das Geheimnis jetzt für ein- und alle mal aufdecken. Jeder metrische Raum ist auch ein topologischer Raum mit der durch die Metrik induzierte Topologie. Was genau bedeutet dies aber nun? Eine Topologie ist ja einfach nur ein System von offenen Mengen. Mithilfe der Metrik kann man sehr leicht offene Kugeln, sogenannte Umgebungen, definieren,

und diese offenen Mengen packt man dann quasi zu einem System zusammen und hat sofort eine Topologie. Dies nennt man dann die durch die Metrik induzierte Topologie. Es muss aber nicht notwendigerweise zu jeder Topologie eine Metrik geben, die diese induziert. Ein normierter Raum unterscheidet sich jetzt von einem metrischen Raum dadurch, dass es sich um einen Vektorraum handeln muss. Wir müssen also insbesondere ein Nullelement haben und außerdem die Addition und auch die skalare Multiplikation. Dieses braucht man bei einem metrischen Raum eben nicht. Normen bilden einen Punkt aus einem Vektorraum in einen Körper ab. Dagegen bildet eine Metrik aus einer nichtleeren, aber sonst beliebigen Menge X das Punktepaaar (x, y) in den Körper (bei uns \mathbb{R}) ab. Dennoch haben sie bis auf die Homogenität und die Symmetrie ähnliche Eigenschaften. Dass dies kein Zufall ist, zeigt dies, dass jede Norm eine Metrik durch $d(x, y) := \|x - y\|$ induziert. Jeder normierte Raum ist also auch ein metrischer Raum. Die Umkehrung ist falsch. Dies sieht man schon an den Voraussetzungen in der entsprechenden Definition, denn X muss nicht einmal ein Vektorraum sein. Es gibt also Metriken, die nicht aus einer Norm induziert werden können. Beispiel 18 liefert ein Beispiel hierfür.

Beispiel 18

Sei d eine Metrix. Dann ist

$$\tilde{d}(x, y) := \frac{d(x, y)}{1 + d(x, y)}$$

nicht von einer Norm induzierbar, aber eine Metrik. ■

Ähnliches kann man nun noch zwischen dem Skalarprodukt und der Norm herstellen. Durch $\|x\| := \sqrt{\langle x, x \rangle}$ lässt sich auf jedem Vektorraum eine Norm einführen. Man sagt, dass dies die durch das Skalarprodukt induzierte Norm ist. Jeder Vektorraum mit Skalarprodukt ist mit der durch das Skalarprodukt induzierten Norm ein normierter Raum und somit auch ein metrischer Raum. Die Umkehrung ist aber auch hier im Allgemeinen falsch. Eine Norm kann, muss aber nicht durch ein Skalarprodukt definiert sein.

Mithilfe der Parallelogrammgleichung kann man sich schnell überlegen, ob eine Norm durch ein Skalarprodukt induziert werden kann. In Abbildung 1.12 haben wir versucht, diese Beziehungen noch einmal zusammenzufassen. Zum Abschluss noch zu einem Gegenbeispiel, das zeigt, wieso einige Umkehrungen in Abbildung 1.12 nicht richtig sind.

Beispiel 19

Die diskrete Metrik aus Beispiel 1 wird nicht durch eine Norm induziert. Das Paradebeispiel für die Notwendigkeit einer Topologie ist die Konvergenz in $\mathbb{R}^{\mathbb{R}}$, also die punktweise Konvergenz von Funktionenfolgen. Man kann zeigen, dass diese Konvergenz nicht durch eine Metrik beschrieben werden kann. ■

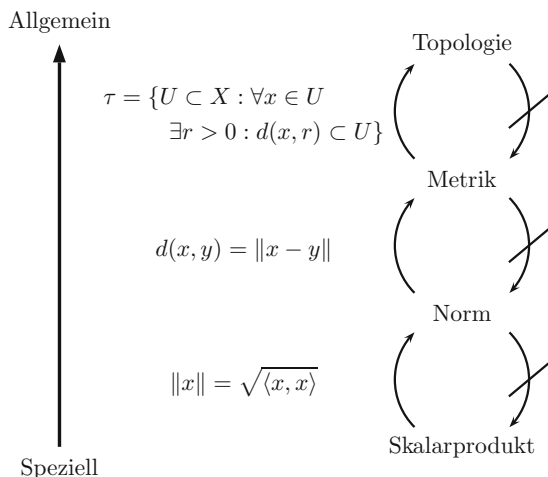


Abb. 1.12: Zusammenhang Topologie-Metrik-Norm-Skalarprodukt.

1.4 Erklärungen zu den Sätzen und Beweisen

Zum Satz 1.1: Um den Beweis zu verstehen, muss man sich eigentlich nur klar machen, wieso $\varepsilon := \min\{\varepsilon_1, \varepsilon_2\}$ das Gewünschte leistet. Schneiden sich die beiden Mengen Ω_1 und Ω_2 nicht, erhalten wir die leere Menge, die sowieso offen und abgeschlossen ist. Sonst gibt es eine Schnittmenge $\Omega_1 \cap \Omega_2 \neq \emptyset$, und wenn man für diese Schnittmenge das ε als das Minimum der beiden ε_1 und ε_2 wählt, so haben wir alles bewiesen und jeden Fall berücksichtigt. Im Beweis steht weiter, dass induktiv die Behauptung folgt. Dieses wollen wir noch einmal kurz erklären. Für zwei Mengen Ω_1, Ω_2 ist die Behauptung klar. Kommt nun noch eine dritte Menge Ω_3 hinzu, so wählen wir die Schnittmenge $\Omega_1 \cap \Omega_2$ einfach als eine Menge, und wir sind wieder im Fall zweier Mengen.

Die Frage, die man sich jetzt noch stellen könnte, ist, ob der Durchschnitt von beliebig vielen offenen Mengen, also zum Beispiel der Durchschnitt von unendlich vielen offenen Mengen wieder offen ist. Dies ist im Allgemeinen falsch, wie das folgende Beispiel zeigt.

Beispiel 20

Es gilt

$$\bigcap_{n \in \mathbb{N}} \left(-\frac{1}{n}, \frac{1}{n} \right) = \{0\},$$

aber $\{0\}$ ist nicht offen, denn anschaulich bedeutet Offenheit einer Menge ja gerade, dass ihre Elemente nur von Elementen dieser Menge umgeben sind, das heißt also kein Element der Menge liegt auf dem Rand der Menge, aber dies ist offensichtlich bei $\{0\}$ erfüllt, denn die Menge besteht ja nur aus dem Punkt 0, und der Rand von $\{0\}$ ist 0 selbst. ■

Daher ist der Durchschnitt von beliebig vielen offenen Mengen im Allgemeinen nicht wieder offen! Zu guter Letzt noch eine Anmerkung: Durch die vollständige Induktion erhält man also die Aussage für eine beliebig große endliche Anzahl von Mengen, aber nicht für eine beliebig große Anzahl von Mengen. Man darf nicht denken, dass man mit der Induktion bis ins „Unendliche“ kommt.

Zum Satz 1.2: Um den Beweis zu verstehen, muss man sich nur Folgendes klar machen: Wenn $U(x, \varepsilon) \subset \Omega_\alpha$, dann ist natürlich $U(x, \varepsilon)$ auch eine Teilmenge der großen Vereinigung $\bigcup_{\alpha \in I} \Omega_\alpha$.

Zum Satz 1.3: Wir hoffen, dass der Beweis ausführlich genug war. Man sollte sich nur noch einmal die Regeln von De Morgan (die Querstriche haben nichts mit Abschluss zu tun, sondern meinen hier natürlich das Komplement) verdeutlichen, die Folgendes besagen: Sei I eine beliebige Indexmenge. Dann gilt

$$\left(\bigcap_{i \in I} A_i \right)^c = \bigcup_{i \in I} A_i^c \quad \text{bzw.} \quad \left(\bigcup_{i \in I} A_i \right)^c = \bigcap_{i \in I} A_i^c.$$

Hierbei bezeichnet das oben stehende „ c “ die Komplementbildung.

Wir bemerken wieder, dass die Vereinigung von beliebig vielen abgeschlossenen Mengen nicht unbedingt wieder abgeschlossen sein muss. Dazu betrachten wir das folgende Beispiel.

Beispiel 21

Es gilt:

$$\bigcup_{n \in \mathbb{N} \setminus \{1\}} \left[\frac{1}{n}, 1 - \frac{1}{n} \right] = (0, 1),$$

aber $(0, 1)$ ist ein offenes Intervall, das bzgl. der Standardmetrik nicht abgeschlossen ist. ■

Zum Satz 1.4: Wir bilden uns ein, dass dieser Satz bzw. sein Beweis keine Erklärung benötigt :-). Sollte dies dennoch der Fall sein, dann schreibt uns eine E-Mail oder postet in unser Online-Angebot.

Zur Hausdorff'schen Trennungseigenschaft (Satz 1.5): Die Trennungseigenschaft verdeutlicht Abbildung 1.13. Der Satz sagt aus, dass jeder metrische Raum hausdorffsch ist. Für topologische Räume (siehe Definition 1.19) ist dies im Allgemeinen falsch.

Zum Satz 1.6, dass offene Bälle wieder offen sind: Es wäre durchaus blöd, wenn der Satz nicht gelten würde, denn dann würde eine offene Kugel den Namen „offen“ gar nicht verdienen. Der Beweis, den wir uns jetzt noch einmal anschauen wollen, rettet aber die Würde der offenen Bälle :-). Um den Beweis zu

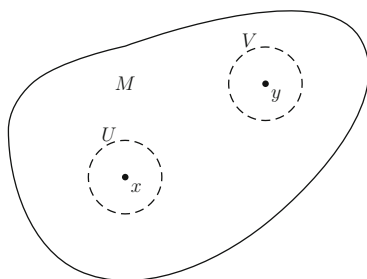


Abb. 1.13: Metrische Räume sind hausdorffsch.

verstehen, muss man sich nur klar machen, was wir eigentlich tun müssen. Wir müssen uns einen beliebigen Punkt x aus der offenen Kugel $U(x_0, r)$ vorgeben und eine weitere offene Kugel $U(x, \varepsilon)$ finden, sodass $U(x, \varepsilon) \subset U(x_0, r)$ ist. Das Einzige, was wir zu tun haben, ist also die Angabe von ε . Dies wurde im obigen Beweis mit $\varepsilon := r - d(x, x_0)$ getan. Wie sind wir aber darauf gekommen? Na, schaut euch mal folgende Abbildung 1.14 an, dann sollte es klar werden. Als Übungsaufgabe solltet ihr einmal zeigen, dass abgeschlossene Bälle wirklich abgeschlossen sind.

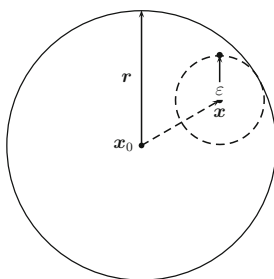


Abb. 1.14: Wählen wir ε als $r - d(x, x_0)$, so gilt auf jeden Fall $U(x, \varepsilon) \subset U(x_0, r)$.

Zum Satz 1.9: Dieser Satz wiederum ist auch irgendwie intuitiv klar, denn das Innere sollte offen und der Abschluss abgeschlossen sein.

Zum Satz von Heine-Borel (Satz 1.10): Dieser Satz ist eigentlich eine Verallgemeinerung des Satzes 1.12, der nur besagt, dass eine Teilmenge des \mathbb{R}^n genau dann kompakt ist, wenn sie abgeschlossen und beschränkt ist. Wir haben aber gleich den allgemeineren Fall angegeben, denn wir haben unter den Definitionen viele verschiedene Kompaktheitsbegriffe (siehe die Definitionen 1.14, 1.15 oder 1.17) kennengelernt. Der Satz von Heine-Borel (Satz 1.10) zeigt nun, dass diese ganzen Begriffe äquivalent sind. Wenn ihr also zeigen wollt, dass eine Menge kompakt ist, so könnt ihr euch aussuchen, ob ihr zeigt, dass die Menge folgenkompakt oder überdeckungskompakt ist. Dennoch sollte man von Fall zu Fall

unterscheiden, was einfacher ist. Siehe dazu beispielsweise den Beweis zu Satz 1.11.

Zum Satz 1.13: Jeder metrische Raum ist hausdorffsch (siehe Satz 1.5). Aber nicht jeder topologische Raum ist hausdorffsch. Dazu betrachtet $M := \{0, 1\}$ und die indiskrete Topologie oder auch die Sierpinski-Topologie.

Zum Satz 1.14: Jeder metrische Raum ist in kanonischer Weise ein topologischer Raum durch die von der Metrik induzierte Topologie.

2 Stetige Abbildungen

Übersicht

2.1	Definitionen	43
2.2	Sätze und Beweise	45
2.3	Erklärungen zu den Definitionen	50
2.4	Erklärungen zu den Sätzen und Beweisen	61

In diesem Kapitel wollen wir stetige Abbildungen in metrischen und topologischen Räumen einführen und untersuchen. Dazu müssen wir natürlich zuerst klären, wie die Stetigkeit einer Abbildung zwischen topologischen Räumen überhaupt definiert ist. Die Stetigkeit zwischen metrischen Räumen kennen wir aber eigentlich schon aus der Analysis 1.

Im Folgenden seien (M, d_M) und (N, d_N) zwei metrische Räume und $f : M \rightarrow N$ eine Abbildung. Dabei werden vor allem Abbildungen der Form $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ untersucht, wobei n und m natürliche Zahlen sind. Wir sagen dazu Funktionen mehrerer Veränderlicher. Die meisten Definitionen und Sätze sind für allgemeine metrische Räume definiert. In den Erklärungen und Beispielen verwenden wir diese Sätze dann aber eher für den \mathbb{R}^n , weil dies in der Praxis und in euren Übungsaufgaben am häufigsten getan wird. Um die folgenden Definitionen und Sätze zu verstehen, wollen wir uns zunächst anschauen, was man unter solch einer Funktion versteht und wie man sie sich graphisch verdeutlichen kann, um eine gute Vorstellung davon zu erhalten.

2.1 Definitionen

Definition 2.1 (Funktion mehrerer Veränderlicher)

Eine reellwertige Funktion in mehreren Veränderlichen ist eine Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Definition 2.2 (Grenzwert einer Funktion)

Seien M und N metrische Räume und d_M und d_N Metriken auf M bzw. N . Weiterhin sei $f : M \rightarrow N$ eine Abbildung. Der Limes (Grenzwert)

$$\lim_{x \rightarrow x_0} f(x) = a$$

existiert, wenn für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, sodass

$$d_N(f(x) - a) < \varepsilon \quad \text{für alle } x \text{ mit } d_M(x, x_0) < \delta.$$

Anmerkung: f besitzt keinen Grenzwert, wenn sich bei Annäherung an x_0 auf verschiedenen Kurven (zum Beispiel Geraden) verschiedene oder keine Grenzwerte ergeben.

Definition 2.3 (Folgenstetigkeit)

Eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ heißt im Punkt x_0 **(folgen)stetig**, wenn

$$\lim_{x \rightarrow x_0} f(x) = f(x_0)$$

ist. f heißt (folgen)stetig, wenn f in jedem Punkt aus dem Definitionsbereich (folgen)stetig ist.

Definition 2.4 (Stetigkeit in metrischen Räumen)

f heißt **(punktweise) stetig** im Punkt $x_0 \in M$, wenn gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 : \forall x \in M \text{ mit } d_M(x, x_0) < \delta \text{ gilt } d_N(f(x), f(x_0)) < \varepsilon.$$

f heißt **stetig**, wenn f in jedem Punkt $x_0 \in M$ stetig ist.

Definition 2.5 (gleichmäßige Stetigkeit in metrischen Räumen)

f heißt **gleichmäßig stetig**, wenn gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 : \forall x, x' \in M \text{ mit } d_M(x, x') < \delta \text{ gilt } d_N(f(x), f(x')) < \varepsilon.$$

Definition 2.6 (α -Hölder-stetig)

Für $0 < \alpha \leq 1$ heißt f auf M **α -Hölder-stetig**, wenn es eine Konstante $C \geq 0$ gibt, sodass

$$d_N(f(x), f(x')) \leq C \cdot d_M(x, x')^\alpha \quad \forall x, x' \in M.$$

Definition 2.7 (Lipschitz-Stetigkeit in metrischen Räumen)

f heißt **lipschitz-stetig**, wenn es eine Konstante $L \geq 0$ gibt, sodass

$$d_N(f(x), f(x')) \leq L \cdot d_M(x, x') \quad \forall x, x' \in M.$$

Definition 2.8 (Stetigkeit zwischen topologischen Räumen)

Seien (M, \mathcal{O}_M) und (N, \mathcal{O}_N) zwei topologische Räume. Dann nennt man eine Abbildung $f : M \rightarrow N$ zwischen diesen topologischen Räumen **stetig**, wenn die Urbilder von in N offenen Mengen offen in M sind, das heißt in Formelschreibweise: Wenn $f^{-1}(\Omega) \in \mathcal{O}_M$ für alle $\Omega \in \mathcal{O}_N$ gilt.

Definition 2.9 (Homöomorphismus)

Seien (M, \mathcal{O}_M) und (N, \mathcal{O}_N) zwei topologische Räume. Ist die Abbildung $f : M \rightarrow N$ bijektiv, und sind sowohl f als auch die Umkehrabbildung f^{-1} stetig, so nennt man f einen **Homöomorphismus**. Man sagt: Die topologischen Räume sind **homöomorph**, wenn ein Homöomorphismus zwischen ihnen existiert.

Definition 2.10 (Fixpunkt)

Seien (M, d) ein metrischer Raum und $f : M \rightarrow M$ eine Abbildung. Ein Punkt $m \in M$ heißt **Fixpunkt** von f , wenn $f(m) = m$ gilt.

Definition 2.11 (Kontraktion)

Sei (M, d) ein metrischer Raum. Eine Abbildung $f : M \rightarrow M$ heißt **Kontraktion**, wenn eine Konstante $C \in [0, 1)$ für alle x mit der Eigenschaft

$$d(f(x), f(y)) \leq C \cdot d(x, y)$$

existiert.

Definition 2.12 (Operatornorm)

Seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ zwei normierte Vektorräume und $L : V \rightarrow W$ eine lineare Abbildung. Die **Operatornorm** von L ist definiert als

$$\|L\| := \sup_{\|v\|_V=1} \|L(v)\|_W = \sup_{v \in V \setminus \{0\}} \frac{\|L(v)\|_W}{\|v\|_V}.$$

2.2 Sätze und Beweise

Satz 2.1 (Zusammensetzung stetiger Funktionen)

Seien (M, d_M) , (N, d_N) und (L, d_L) drei metrische Räume und $f : L \rightarrow M$ sei stetig in $x_0 \in L$ und $g : M \rightarrow N$ sei stetig in $y_0 := f(x_0)$. Dann ist auch die Funktion $g \circ f : L \rightarrow N$ stetig in x_0 . Weiterhin gelten für zwei metrische Räume (M, d_M) und (N, d_N) folgende Aussagen: Seien $f : M \rightarrow \mathbb{R}$ stetig in $x_0 \in M$ und $g : N \rightarrow \mathbb{R}$ stetig in $y_0 \in N$. Dann sind auch die Funktionen

$f(x) + g(y), f(x) - g(y), f(x) \cdot g(y)$ und, sofern $g(y) \neq 0$, $\frac{f(x)}{g(y)}$ bezüglich der Produktmetrik (siehe Beispiel 1, Seite 17) auf $M \times N$ stetig im Punkt (x_0, y_0) .

Beweis: Analog wie in Analysis 1, siehe [MK09]. q.e.d.

Satz 2.2 (Stetigkeitskriterium)

Seien (M, d_M) und (N, d_N) zwei metrische Räume und $f : M \rightarrow N$ eine Abbildung. Dann sind die folgenden Aussagen äquivalent:

- i) f ist stetig.
- ii) Die Urbilder d_N -offener Mengen sind d_M -offen.
- iii) Die Urbilder d_N -abgeschlossener Mengen sind d_M -abgeschlossen.

Beweis:

„i) \Rightarrow ii)“: Seien f stetig und $\Omega \subset N$ eine d_N -offene Menge. Ohne Einschränkung nehmen wir an, dass $\Omega \neq \emptyset$ und $f^{-1}(\Omega) \neq \emptyset$. Nun wählen wir ein $x_0 \in M$ mit $f(x_0) \in \Omega$, das heißt $x_0 \in f^{-1}(\Omega)$. Da Ω nach Voraussetzung d_N -offen ist, existiert ein $\varepsilon > 0$ mit $U(f(x_0), \varepsilon) \subset \Omega$. Da f aber stetig nach i) ist, existiert auch ein $\delta > 0$, sodass

$$d_N(f(x), f(x_0)) < \varepsilon$$

für alle $x \in M$ mit $d_M(x, x_0) < \delta$. Dies bedeutet aber nichts anderes als, dass für alle $x \in U(x_0, \delta)$ gilt, dass $f(x) \in U(f(x_0), \varepsilon)$. Da $U(f(x_0), \varepsilon) \subset \Omega$ und

$$f(U(x_0, \delta)) \subset U(f(x_0), \varepsilon) \subset \Omega,$$

folgt $U(x_0, \delta) \subset f^{-1}(\Omega)$, das heißt, $f^{-1}(\Omega)$ ist d_M -offen. Das war zu zeigen.

„ii) \Rightarrow i)“: Es seien $x_0 \in M$ und $\varepsilon > 0$. Da $U(f(x_0), \varepsilon)$ d_N -offen ist, gilt nach Voraussetzung, dass $f^{-1}(U(f(x_0), \varepsilon))$ d_M -offen ist. Dies bedeutet aber gerade, dass ein $\delta > 0$ existiert mit

$$U(x_0, \delta) \subset f^{-1}(U(f(x_0), \varepsilon)).$$

Dann gilt aber für alle $x \in M$ mit $d_M(x, x_0) < \delta$ auch $d_N(f(x), f(x_0)) < \varepsilon$. Daher ist das ε - δ -Kriterium erfüllt und folglich f stetig.

„iii) \Rightarrow ii)“: Dies folgt sofort aus Komplementbildung

$$f^{-1}(N \setminus \Omega) = M \setminus f^{-1}(\Omega).$$

Dies soll uns genügen.

q.e.d.

Satz 2.3

Seien (M, d_M) und (N, d_N) zwei metrische Räume und $f : M \rightarrow N$ eine stetige Abbildung. Dann sind die Bilder d_M -kompakter Menge wieder d_N -kompakt.

Beweis: Sei $K \subset M$ kompakt. Wir zeigen, dass $f(K)$ folgenkompakt ist. Nach dem Satz von Heine-Borel (Satz 1.10) ist dies äquivalent zu den anderen Kompaktheitsbegriffen, die wir in Kapitel 1 eingeführt hatten. Sei hierzu $(y_n)_{n \in \mathbb{N}} \subset f(K)$ eine Folge und $(x_n)_{n \in \mathbb{N}} \subset K$ so gewählt, dass $f(x_n) = y_n$. Da K kompakt ist, existiert eine Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$ von $(x_n)_{n \in \mathbb{N}}$ und ein $x \in K$ mit $\lim_{k \rightarrow \infty} x_{n_k} = x$. Da f stetig ist, folgt hieraus

$$y := f(x) = \lim_{k \rightarrow \infty} f(x_{n_k}) = \lim_{k \rightarrow \infty} y_{n_k}$$

und $y \in f(K)$. Daher existiert eine in $f(K)$ konvergente Teilfolge von $(y_n)_{n \in \mathbb{N}}$.
q.e.d.

Satz 2.4

Seien (M, d) ein metrischer Raum und $f : M \rightarrow \mathbb{R}$ eine stetige Abbildung. Dann ist f auf jeder kompakten Teilmenge $K \subset M$ beschränkt und nimmt ihr Supremum und Infimum an.

Beweis: Da das Bild kompakter Mengen nach Satz 2.3 unter stetigen Abbildung wieder kompakt ist, und kompakte Mengen insbesondere beschränkt sind, ist f auf K beschränkt. Wir setzen

$$\lambda := \inf_{x \in K} f(x) \quad \text{und} \quad \mu := \sup_{x \in K} f(x).$$

Ist nun $(x_n)_{n \in \mathbb{N}} \subset K$ eine Folge mit $\lim_{n \rightarrow \infty} f(x_n) = \lambda$, so existiert wegen der Kompaktheit von K eine Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$ von $(x_n)_{n \in \mathbb{N}}$, die gegen ein $x \in K$ konvergiert. Für dieses x gilt wegen der Stetigkeit von f nun insgesamt

$$f(x) = \lim_{k \rightarrow \infty} f(x_{n_k}) = \lim_{n \rightarrow \infty} f(x_n) = \lambda,$$

also nimmt f in x sein Infimum an. Der Beweis für das Supremum geht genauso. Übungsaufgabe :-).
q.e.d.

Satz 2.5

Seien (M, d_M) und (N, d_N) zwei metrische Räume und $f : M \rightarrow N$ stetig. Dann ist f auf jeder kompakten Teilmenge von M sogar gleichmäßig stetig.

Beweis: Wir führen den Beweis durch Widerspruch. Angenommen f wäre nicht gleichmäßig stetig. Sei weiterhin $k \subset M$. Dann gäbe es ein $\varepsilon > 0$, sodass für alle $n \in \mathbb{N}$ Punkte x_n und x'_n existieren mit

$$d_M(x_n, x'_n) < \frac{1}{n} \quad \text{und} \quad d_N(f(x_n), f(x'_n)) \geq \varepsilon.$$

Da K kompakt ist, existiert eine Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$, die gegen ein $x \in K$ konvergiert. Wegen

$$d_M(x_{n_k}, x'_{n_k}) < \frac{1}{n_k}$$

gilt dann aber auch $\lim_{k \rightarrow \infty} x'_{n_k} = x$. Da f andererseits stetig ist, folgt aus der Dreiecksungleichung

$$\lim_{k \rightarrow \infty} d_N(f(x_{n_k}), f(x'_{n_k})) \leq \lim_{k \rightarrow \infty} d_N(f(x_{n_k}), f(x)) + \lim_{k \rightarrow \infty} d_N(f(x), f(x'_{n_k})) = 0.$$

Dies ist aber ein Widerspruch zu $d_N(f(x_{n_k}), f(x'_{n_k})) \geq \varepsilon$ und beweist damit die Behauptung. q.e.d.

Satz 2.6 (Banach'scher Fixpunktsatz)

(M, d) sein ein vollständiger metrischer Raum und $f : M \rightarrow M$ eine Kontraktion. Dann besitzt f genau einen Fixpunkt in M .

Beweis:

Eindeutigkeit: Seien m_1 und m_2 zwei Fixpunkte von f . Wir zeigen, dass dann $m_1 = m_2$ gelten muss. Dies sieht man so: Es ist

$$d(m_1, m_2) = d(f(m_1), f(m_2)) \leq C \cdot d(m_1, m_2).$$

Hieraus folgt

$$\underbrace{(1 - C)}_{>0} \cdot d(m_1, m_2) \leq 0,$$

also

$$d(m_1, m_2) = 0 \Rightarrow m_1 = m_2.$$

Existenz: Wir wählen einen Punkt $m_0 \in M$ beliebig und eine Folge rekursiv definiert durch $m_{n+1} := f(m_n)$ mit $n \in \mathbb{N}$. Wir zeigen nun, dass $(m_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge ist. Wählen wir $n, k \in \mathbb{N}$, so müssen wir abschätzen

$$d(m_{n+k}, m_n) \leq \sum_{i=n+1}^{n+k} d(m_i, m_{i-1}).$$

Es gilt:

$$\begin{aligned} d(m_i, m_{i-1}) &= d(f(m_{i-1}), f(m_{i-2})) \leq C \cdot d(m_{i-1}, m_{i-2}) \\ &= C \cdot d(f(m_{i-2}), f(m_{i-3})) \leq C^2 \cdot d(m_{i-2}, m_{i-3}). \end{aligned}$$

Rekursiv und induktiv ergibt sich

$$\dots \leq C^{i-1} d(m_1, m_0) = C^{i-1} \cdot d(f(m_0), m_0).$$

Damit ist

$$\begin{aligned} d(m_{n+k}, m_n) &\leq \sum_{i=n+1}^{n+k} d(m_i, m_{i-1}) \leq \sum_{i=n+1}^{n+k} C^{i-1} \cdot d(f(m_0), m_0) \\ &= d(f(m_0), m_0) \cdot \sum_{i=n}^{n+k-1} C^i \\ &\leq d(f(m_0), m_0) \sum_{i=n}^{\infty} C^i = C^n d(f(m_0), m_0) \frac{1}{1-C}. \end{aligned}$$

Also ist $(m_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge. Da M nach Voraussetzung vollständig ist, gilt $m := \lim_{n \rightarrow \infty} m_n \in M$. Der Grenzwert existiert also in M . Da f eine Kontraktion ist, ergibt sich nun die Behauptung aus der Rekursionsvorschrift

$$\lim_{n \rightarrow \infty} f(m_n) = f(m) \Rightarrow f(m) = m.$$

q.e.d.

Satz 2.7

Eine lineare Abbildung $L : V \rightarrow W$ zwischen normierten Vektorräumen ist genau dann beschränkt, wenn sie stetig ist. Insbesondere sind stetige lineare Abbildungen zwischen normierten Vektorräumen auch lipschitz-stetig.

Beweis:

„ \Rightarrow “: Sei L ein beschränkter linearer Operator. Dann gilt für alle $v_1, v_2 \in V$ und $v_1 \neq v_2$

$$\frac{\|L(v_1) - L(v_2)\|_W}{\|v_1 - v_2\|_V} = \frac{\|L(v_1 - v_2)\|_W}{\|v_1 - v_2\|_V} < \|L\| < \infty.$$

Also ist L lipschitz-stetig, und insbesondere stetig.

„ \Leftarrow “: Sei L stetig. Dann ist L auch im Punkt $0 \in V$ stetig und daher existiert ein $\delta > 0$ mit $\|L(v') - L(0)\|_W = \|L(v')\|_W \leq 1$ für alle $v' \in V$ mit $\|v'\|_V < \delta$. Wir definieren $c := \frac{2}{\delta}$. Dann gilt für alle $v \in V$ mit $\|v\|_V = 1$ die Abschätzung

$$\|L(v)\|_W = \|c \cdot L(v/c)\|_W = c \cdot \|L(v/c)\|_W \leq c,$$

denn $\|v/c\|_V = \delta/2$. Demnach ist L beschränkt.

Wir haben nun alles gezeigt.

q.e.d.

2.3 Erklärungen zu den Definitionen

Zur Definition 2.1 einer Funktion mehrerer Veränderlicher: Eine reellwertige Funktion in mehreren Veränderlichen ist also einfach eine Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Als Funktionsgleichung schreiben wir $f(x_1, \dots, x_n)$. Was kann man sich darunter vorstellen? Das ist ganz einfach. Die Funktion hängt einfach von mehreren Variablen, wir sagen mehreren Veränderlichen x_1, \dots, x_n , ab. Als Funktionswert erhalten wir einen Vektor aus dem \mathbb{R}^m , also einen Vektor mit m Einträgen. In der Analysis 1 haben wir Funktionen einer Veränderlicher, also Funktionen der Form $f : \mathbb{R} \rightarrow \mathbb{R}$ (oder auch $f : \mathbb{R} \rightarrow \mathbb{R}^m$) untersucht, beispielsweise $f(x) = x^2$. Nun hindert uns doch aber nichts daran, Funktionen mit der Form $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ zu untersuchen, und dies ist Gegenstand der Analysis 2, also beispielsweise $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch $f(x, y) = \frac{xy}{x^2 + y^2}$. Diese Funktion besitzt zwei Veränderliche, nämlich x und y . Für x und y können wir entsprechend reelle Werte einsetzen. Solche Funktionen können stetig oder differenzierbar (siehe Kapitel 3) sein.

Wir werden nun zuerst ein paar Möglichkeiten geben, wie wir Abbildungen der Form $f : \mathbb{R}^n \rightarrow \mathbb{R}$ visualisieren können. Im Allgemeinen wird durch $f(x, y)$ eine Fläche im x, y, z -Raum beschrieben. Wir werden jetzt sogenannte Niveaulinien, Flächen im Raum und Blockbilder studieren. Um noch eine Anschauung zu haben, können wir natürlich nur Funktionen der Form $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ bzw. $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ verdeutlichen, aber alle Konzepte lassen sich leicht verallgemeinern, nur ist die Vorstellung dann wieder etwas schwieriger :-).

Höhenlinien (oder auch Niveaulinien)

Beispiel 22

Wir wollen die durch die Funktion

$$z := f(x, y) = \frac{y}{1 + x^2}$$

beschriebene Fläche im Bereich

$$B := \{(x, y) : -2 \leq x \leq 2, -2 \leq y \leq 2\}$$

durch Höhenlinien verdeutlichen. Um Höhenlinien bzw. Niveaulinien zu berechnen, setzen wir bestimmte Werte für $z = c$ ein und erhalten so eine Funktion in Abhängigkeit von x , die wir darstellen können. Vorstellen kann man sich die Höhenlinien als eine Landkarte in einem Atlas.

Sei zum Beispiel $z = c = 0$. Dann folgt für die Funktion, dass $y = 0$. Also erhalten wir für $z = c = 0$ die x -Achse. Sei jetzt $z = c = \frac{1}{2}$. Es ergibt sich sofort

$$\frac{1}{2} = \frac{y}{1 + x^2} \Rightarrow y = \frac{1}{2}(1 + x^2),$$

Tab. 2.1: Weitere Höhenlinien mit bestimmten Werten.

Höhe	definierte Gleichung	Beschreibung Normalform
$z = c = 0$	$y = 0$	x -Achse
$z = c = \frac{1}{2}$	$2 \cdot y = 1 + x^2$	Parabel $x^2 = 2 \cdot \left(y - \frac{1}{2}\right)$: Scheitel in $\left(0, \frac{1}{2}\right)$, Öffnung $p = 2$
$z = c = 1$	$y = 1 + x^2$	Parabel $x^2 = 1 \cdot (y - 1)$: Scheitel in $\left(0, \frac{1}{2}\right)$, Öffnung $p = -2$
$z = c = \frac{3}{2}$	$\frac{2}{3} \cdot y = 1 + x^2$	Parabel $x^2 = \frac{2}{3} \cdot \left(y - \frac{3}{2}\right)$: Scheitel in $\left(0, \frac{3}{2}\right)$, Öffnung $p = \frac{2}{3}$
$z = c = -\frac{1}{2}$	$2 \cdot y = -(1 + x^2)$	Parabel $x^2 = -2 \cdot \left(y + \frac{1}{2}\right)$: Scheitel in $\left(0, -\frac{1}{2}\right)$, Öffnung $p = -2$
$z = c = -1$	$y = -(1 + x^2)$	Parabel $x^2 = -1 \cdot (y + 1)$: Scheitel in $\left(0, -\frac{1}{2}\right)$, Öffnung $p = 2$
$z = c = -\frac{3}{2}$	$\frac{2}{3} \cdot y = -(1 + x^2)$	Parabel $x^2 = -\frac{2}{3} \cdot \left(y + \frac{3}{2}\right)$: Scheitel in $\left(0, -\frac{3}{2}\right)$, Öffnung $p = -\frac{2}{3}$

also eine Parabel. Dieses Spielchen können wir weiter treiben und erhalten so die Tabelle 2.1.

Die Abbildung 2.1 verdeutlicht die Funktion. Wir schauen sozusagen von oben auf die Funktion drauf. Erkennbar sind dann die Höhenlinien. ■

Blockbild

Das Blockbild (die folgende Darstellung wird jetzt von uns immer so bezeichnet) verdeutlichen wir uns ebenfalls an einem Beispiel.

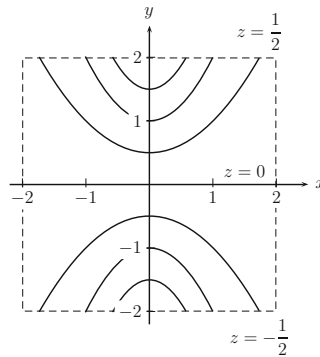


Abb. 2.1: Höhenlinien am Beispiel.

Beispiel 23

Wir betrachten die Funktion $z := f(x, y) = \frac{y}{1+x^2}$ und wollen nun ein sogenanntes Blockbild der Funktion zeichnen. Um dies anzudeuten, wird für jede Spante (Der Begriff des Spants stammt aus dem Schiffsbau. Wer dies nicht sofort versteht, der frage einfach einmal wikipedia.de. Dort gibt es auch ein nettes Bildchen, das den Begriff gut erklärt.) und für den Rand jeweils die Höhe $z = f(x_i, y)$ mit $i \in \{0, 1, 2, 3, 4\}$ berechnet. Wir erhalten

$$z(x_0, y) = f(-2, y) = \frac{y}{5}, \quad z(x_1, y) = f(-1, y) = \frac{y}{2}, \quad z(x_2, y) = f(0, y) = y, \\ z(x_3, y) = f(1, y) = \frac{y}{2}, \quad z(x_4, y) = f(2, y) = \frac{y}{5}.$$

Nun müssen wir noch für jede Spante und den Rand jeweils die Höhe $z = f(x, y_i)$ mit $i \in \{0, 1, 2, 3, 4\}$ berechnen. Es ergibt sich

$$z(x, y_0) = f(x, -2) = \frac{-2}{1+x^2}, \quad z(x, y_1) = f(x, -1) = \frac{-1}{1+x^2}, \\ z(x, y_2) = f(x, 0) = 0, \quad z(x, y_3) = f(x, 1) = \frac{1}{1+x^2}, \quad z(x, y_4) = f(x, 2) = \frac{2}{1+x^2}.$$

Diese Funktionen zeichnen wir nun nacheinander in ein sogenanntes Blockbild ein. Am besten man lässt sich dieses plotten. Die Abbildung 2.2 verdeutlicht das oben Berechnete. ■

Schnitt mit Ebenen

Als Letztes verdeutlichen wir uns Funktionen mehrerer Veränderlicher durch den Schnitt mit Ebenen.

Beispiel 24

Sei eine Funktion gegeben durch

$$z = f(x, y) = \frac{1}{x^2 + y^2}.$$

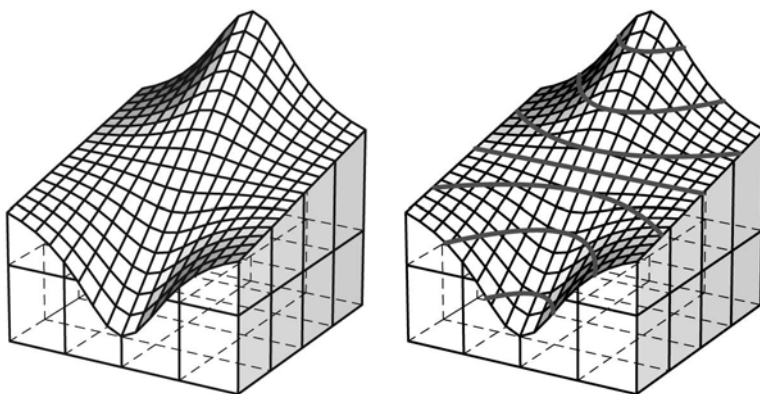


Abb. 2.2: Das Blockbild von Beispiel 23 (links ohne und rechts mit Höhenlinien).

Wir zeichnen den Schnitt mit einer Ebene, also setzen beispielsweise $x = \text{const.}$ bzw. $y = \text{const.}$ Genauer sogar $x = 0$ bzw. $y = 0$. In beiden Fällen erhalten wir

$$z = \frac{1}{y^2} \quad \text{bzw.} \quad z = \frac{1}{x^2}.$$

Graphisch ist dies völlig klar, wie Abbildung 2.3 zeigt.

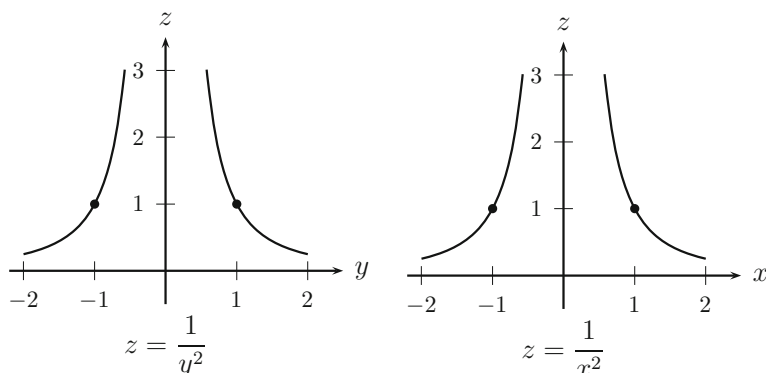


Abb. 2.3: Schnitt mit einer Ebene am Beispiel. Links mit $x = 0$ und rechts mit $y = 0$.

In Abbildung 2.4 geben wir noch einmal die Höhenlinien und das Blockbild an. Beide Abbildungen verdeutlichen, dass wir bei den Höhenlinien quasi auf die Ebene drauf schauen. ■

Zur Definition 2.2 des Grenzwertes einer Funktion Diese Definition entspricht ebenfalls dem, was wir aus der Analysis 1 für Funktionen einer Veränderlichen kennen. Dieses Konzept kann also ebenfalls leicht verallgemeinert werden. Aber wie weisen wir dies konkret nach? Betrachten wir ein Beispiel.

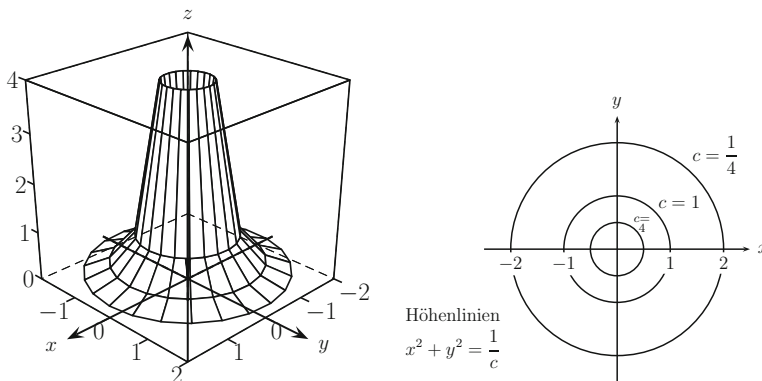


Abb. 2.4: Links die Rotationsfläche als Blockbild und rechts die Funktion f nochmals in Höhenlinien.

Beispiel 25

Existiert der Grenzwert

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{e^{x^2} - 1}?$$

Um zu zeigen, dass der Grenzwert nicht existiert, müssen wir uns dem Punkt $(0,0)$, also dem Nullpunkt im \mathbb{R}^2 , nur auf zwei verschiedenen Arten annähern und zeigen, dass die entsprechenden Grenzwerte nicht übereinstimmen und damit der Limes gar nicht existieren kann.

Sei zunächst $y = 0$. Dann erhalten wir

$$\lim_{(x,0) \rightarrow (0,0)} f(x,0) = \lim_{(x,0) \rightarrow (0,0)} \frac{x \cdot 0}{e^{x^2} - 1} = \lim_{(x,0) \rightarrow (0,0)} 0 = 0.$$

Nun sei $y = x$. So ergibt sich mit zweimaligem Anwenden der Regeln von Hospital (siehe [MK09], Seite 180)

$$\begin{aligned} \lim_{(x,x) \rightarrow (0,0)} f(x,x) &= \lim_{(x,x) \rightarrow (0,0)} \frac{x^2}{e^{x^2} - 1} = \lim_{(x,x) \rightarrow (0,0)} \frac{2x}{2x \cdot e^{x^2}} \\ &= \lim_{(x,x) \rightarrow (0,0)} \frac{2}{2e^{x^2} + 4x^2 \cdot e^{x^2}} = 1. \end{aligned}$$

Es kommen also verschiedene Grenzwerte heraus, damit existiert der Grenzwert $\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{e^{x^2} - 1}$ nicht. ■

Zu den Definitionen 2.4–2.7 der Stetigkeit in metrischen Räumen: Die Stetigkeit in metrischen Räumen überträgt sich aus der Definition, die ihr schon in der Analysis 1 kennengelernt habt (siehe wieder einmal [MK09]). Der Unterschied ist nur, dass wir für den Abstandsbegriff eine allgemeine Metrik verwenden. Auch hier gelten die bekannten Sätze, dass gleichmäßig stetige Funktionen auch stetig sind. Ebenso sind lipschitz-stetige oder α -Hölder-stetige Funktionen stetig. Weiterhin gilt auch hier, dass die Zusammensetzung von stetigen Funktionen wieder

stetig ist (siehe auch Satz 2.1). Alles nichts Neues. Wir wollen dennoch ein paar Beispiele geben und die Stetigkeit von Abbildungen der Form $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ nachweisen.

Ach, zuvor noch eine Anmerkung, wenn diese den meisten nicht sowieso schon klar war: Das d_M deutet an, dass damit die Metrik auf M bezeichnet werden soll. Analog für d_N .

Wie bei Funktionen einer Veränderlichen, also der Form $f : \mathbb{R} \rightarrow \mathbb{R}^n$, können wir auch bei Abbildungen mit mehreren Veränderlichen, das heißt $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ die Stetigkeit nachweisen. Die Definition 2.4 ist dafür leider ziemlich nutzlos bzw. unhandlich, da wir jetzt ja gerade Funktionen betrachten, in die wir einen Punkt aus dem \mathbb{R}^n einsetzen und aus denen einen Punkt aus dem \mathbb{R}^m erhalten. Vielmehr ist die äquivalente (die Äquivalenz haben wir nicht gezeigt, aber man kann sich dies einmal überlegen) Definition 2.3 wesentlich nützlicher, wie die folgenden Beispiele zeigen werden. Sinnvoll wird die Definition erst durch die Definition 2.2 des Grenzwertes.

Beispiel 26

Wir wollen nachweisen, dass die Funktion

$$f(x, y) := \begin{cases} \frac{1 - \cos(xy)}{y}, & \text{für } y \neq 0, \\ 0, & \text{für } y = 0 \end{cases}$$

überall stetig ist. Für $y \neq 0$ ist f nach Satz 2.1 stetig als Zusammensetzung stetiger Funktionen. Hierfür müssen wir nichts weiter zeigen. Nur die Stetigkeitsuntersuchung für $y = 0$ ist interessant. Wir zeigen also die Stetigkeit auf der x -Achse, also für $y = 0$. Wir zeigen nun für beliebiges $x_0 \in \mathbb{R}$

$$\lim_{(x,y) \rightarrow (x_0,0)} f(x, y) = 0.$$

Hierzu verwenden wir Kenntnisse aus der Analysis 1, genauer die Taylorreihe der Kosinusfunktion, also

$$\cos(y) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} y^{2k}.$$

Es ergibt sich zunächst

$$\begin{aligned} f(x, y) &= \frac{1 - \left(1 - \frac{1}{2!}x^2y^2 + \frac{1}{4!}x^4y^4 \mp \dots\right)}{y} = \frac{1 - 1 + \frac{1}{2!}x^2y^2 - \frac{1}{4!}x^4y^4 \pm \dots}{y} \\ &= \frac{\frac{1}{2!}x^2y^2 - \frac{1}{4!}x^4y^4 \pm \dots}{y} = \frac{1}{2!}x^2y - \frac{1}{4!}x^4y^3 \pm \dots \end{aligned}$$

Wir haben die Funktion nun also umgeschrieben und können so den Grenzwert leicht berechnen. Es gilt:

$$\lim_{(x,y) \rightarrow (x_0,0)} f(x, y) = \lim_{(x,y) \rightarrow (x_0,0)} \frac{1}{2!}x^2y - \frac{1}{4!}x^4y^3 \pm \dots = 0 = f(x_0, 0).$$

Damit ist f überall stetig.

Merke also: Um die Stetigkeit einer Funktion zu zeigen, ist es eventuell nötig, die Funktion zuerst, beispielsweise mittels bekannter Taylorreihen oder ähnliches, so umzuschreiben, dass man den Grenzwert und damit die Stetigkeit leicht ermitteln kann. ■

Für Grenzwertbestimmungen, also auch für Stetigkeitsuntersuchungen, ist es oft nützlich, die Funktion mittels Polarkoordinaten umzuschreiben. Vor allem bei rationalen Funktionen kann dies hilfreich sein. Schauen wir uns also noch einmal die Polarkoordinaten im \mathbb{R}^2 an (man kann sie also nur anwenden, wenn man im \mathbb{R}^2 arbeitet, sonst gibt es andere Koordinaten, wie Kugelkoordinaten), siehe dazu die Abbildung 2.5.

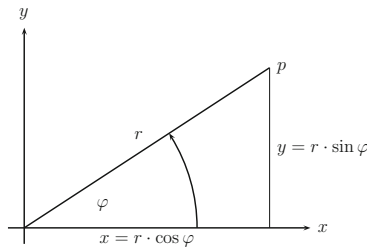


Abb. 2.5: Die Polarkoordinaten.

Man kann sich diese Koordinaten leicht herleiten. Wieso man jetzt aber zu anderen Polarkoordinaten übergeht, werden wir gleich sehen. Es kann nämlich sein, dass man so die Stetigkeit leichter nachweisen kann. Aus der Schulgeometrie wissen wir, dass

$$\begin{aligned}\cos \varphi &= \frac{x}{r} \Leftrightarrow x = r \cdot \cos \varphi, \\ \sin \varphi &= \frac{y}{r} \Leftrightarrow y = r \cdot \sin \varphi.\end{aligned}$$

Hierbei ist r die Länge des Vektors (x, y) und φ der Winkel, den (x, y) mit der x -Achse einschließt. Nun lassen wir die Länge r gegen Null konvergieren. Erhalten wir einen Grenzwert, der unabhängig vom Winkel ist, dann haben wir gezeigt, dass f im Nullpunkt $(0, 0)$ stetig ist. Zeigen wir dies am nächsten Beispiel!

Beispiel 27

Wir betrachten die Funktion

$$f(x, y) := \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2}, & \text{für } (x, y) \neq (0, 0), \\ 0, & \text{für } (x, y) = (0, 0). \end{cases}$$

Zunächst ist klar, dass die Funktion für $(x, y) \neq (0, 0)$ stetig ist, da sie sich nur aus stetigen Funktionen zusammensetzt, siehe Satz 2.1. Die Frage ist nun, was

passiert im Nullpunkt $(x, y) = (0, 0)$? Dazu schreiben wir die Funktion mit den Polarkoordinaten von oben um und erhalten

$$\begin{aligned} f(x, y) &= (r \cos \varphi)(r \sin \varphi) \frac{r^2 \cos^2 \varphi - r^2 \sin^2 \varphi}{r^2 \cos^2 \varphi + r^2 \sin^2 \varphi} \\ &= r^4 \cos \varphi \sin \varphi \frac{\cos^2 \varphi - \sin^2 \varphi}{r^2}. \end{aligned}$$

Nun gilt

$$|f(x, y) - f(0, 0)| = \left| \frac{1}{2} r^2 \sin(2\varphi) \cos(2\varphi) \right| \leq \frac{1}{2} r^2,$$

denn $\sin(2\varphi) = 2 \sin \varphi \cos \varphi$ bzw. $\cos(2\varphi) = \cos^2 \varphi - \sin^2 \varphi$, wenn man etwas nachdenkt oder in einer gut sortierten Formelsammlung nachschlägt. Viel einfacher sieht man dies natürlich daran, dass $\sin(2\varphi)$ und $\cos(2\varphi)$ jeweils durch 1 beschränkt sind. Dies haben wir gezeigt, um zu folgern, dass $\sin(2\varphi) \cos(2\varphi)$ beschränkt ist. Wir erhalten demnach

$$\lim_{(x, y) \rightarrow (0, 0)} f(x, y) = \lim_{r \rightarrow 0} \frac{1}{2} r^2 \sin(2\varphi) \cos(2\varphi) = 0.$$

Daraus ergibt sich die Stetigkeit von f im Nullpunkt und damit in jedem Punkt. ■

Merke hier: Um die Stetigkeit (oder auch Unstetigkeit) nachzuweisen, bzw. um generell Grenzwerte zu berechnen, ist es manchmal zweckmäßig, die Funktion mittels Polarkoordinaten umzuschreiben. Man lässt hier die Länge r gegen Null gehen. Erhalten wir einen Grenzwert unabhängig vom Winkel, dann ist gezeigt, dass f im Nullpunkt stetig ist. Anschaulich ist das auch klar, denn wenn die Länge des Vektor gegen Null geht, also $r \rightarrow 0$, dann nähert man sich ja dem Nullpunkt an. In den nächsten Beispielen wollen wir die Unstetigkeit von Funktionen der Form $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ zeigen. Hierbei geht man genauso vor, wie wir dies im Beispiel zum Grenzwert in den Erklärungen zu Definition 2.2 gemacht haben, denn es gilt: Die Funktion $f(x, y)$ ist bei (x_0, y_0) unstetig, falls es zu zwei verschiedenen Kurven, zum Beispiel Geraden, durch (x_0, y_0) bei Annäherung an (x_0, y_0) verschiedene (oder keine) Grenzwerte gibt.

Beispiel 28

- Wir behaupten, dass die Funktion $f(x, y) = \frac{xy}{e^{x^2} - 1}$ im Nullpunkt unstetig ist. Wir müssen uns nur mit zwei verschiedenen Geraden annähern. Die Rechnung haben wir schon in Beispiel 25 durchgeführt. Für $y = 0$ ergab sich

$$\lim_{(x, 0) \rightarrow (0, 0)} f(x, 0) = 0$$

und für $y = x$ entsprechend

$$\lim_{(x, x) \rightarrow (0, 0)} f(x, x) = 1.$$

Daraus folgt die Unstetigkeit im Nullpunkt.

■ Wir zeigen: Die Funktion

$$f(x, y) := \begin{cases} \frac{xy}{x^2 + y^2}, & \text{für } (x, y) \neq (0, 0) \\ 0, & \text{für } (x, y) = (0, 0) \end{cases}$$

ist im Nullpunkt unstetig. Die Stetigkeit für $(x, y) \neq (0, 0)$ ist wegen der Zusammensetzung stetiger Funktionen wieder klar. Um die Unstetigkeit im Nullpunkt nachzuweisen, nähern wir uns mithilfe verschiedener Geraden an und zeigen, dass die Grenzwerte nicht übereinstimmen.

- Annäherung auf der Geraden $y = 0$ (x -Achse):

$$\lim_{(x,0) \rightarrow (0,0)} f(x, 0) = \lim_{(x,0) \rightarrow (0,0)} \frac{x \cdot 0}{x^2 + 0^2} = 0.$$

- Annäherung auf der Geraden $x = 0$ (y -Achse):

$$\lim_{(0,y) \rightarrow (0,0)} f(0, y) = \lim_{(0,y) \rightarrow (0,0)} \frac{0 \cdot y}{0^2 + y^2} = 0.$$

- Annäherung auf der Geraden $y = x$:

$$\lim_{(x,x) \rightarrow (0,0)} f(x, x) = \lim_{(x,x) \rightarrow (0,0)} \frac{x^2}{x^2 + x^2} = \frac{1}{2}.$$

Also existiert der Grenzwert

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{x^2 + y^2}$$

nicht und damit folgt die Unstetigkeit von f im Nullpunkt.

Merke: Um Unstetigkeit zu zeigen, nähere dich der Funktion mittels verschiedener Geraden an und zeige, dass die Grenzwerte nicht übereinstimmen, da dies die Unstetigkeit impliziert.

Auch die Unstetigkeit kann mittels Polarkoordinaten gezeigt werden. Ab und zu ist dies einfacher. Wir betrachten noch einmal dieselbe Funktion. Es gilt:

$$\begin{aligned} f(x, y) &= f(r \cos \varphi, r \sin \varphi) = \frac{r^2 \sin \varphi \cos \varphi}{r^2 \sin^2 \varphi + r^2 \cos^2 \varphi} \\ &= \sin \varphi \cos \varphi = \frac{1}{2} \sin(2\varphi). \end{aligned}$$

Der Grenzwert

$$\lim_{r \rightarrow 0} f(r \cos \varphi, r \sin \varphi) = \lim_{r \rightarrow 0} \frac{1}{2} \sin(2\varphi) = \frac{1}{2} \sin(2\varphi)$$

ist abhängig von φ . Daher existiert der Grenzwert $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$ nicht. Daraus folgt die Unstetigkeit.

Merke: Auch die Unstetigkeit lässt sich mittels Polarkoordinaten sehr gut demonstrieren. Dazu muss gezeigt werden, dass der entsprechende Grenzwert nicht existiert. Mit Polarkoordinaten zu arbeiten, bietet sich immer dann an, wenn man so besser oder einfacher zeigen kann, dass ein gewisser Grenzwert nicht existiert. ■

Zur Definition 2.8 der Stetigkeit zwischen topologischen Räumen: Diese Definition wird durch Satz 2.2 motiviert und liefert eine Möglichkeit, wie man geschickt Stetigkeit zwischen topologischen Räumen definieren kann.

Dies gilt übrigens auch für metrische Räume M und N . Eine Abbildung $f : M \rightarrow N$ ist genau dann im Punkt $x \in M$ stetig, wenn zu jeder Umgebung V von $f(x)$ eine Umgebung U von x existiert mit $f(U) \subset V$. Und da dieses Stetigkeitskriterium nicht auf die Metriken von M und N Bezug nimmt, sondern nur mit Umgebungen und offenen Mengen arbeitet, kann man dies natürlich auch für topologische Räume definieren, und dies haben wir in Definition 2.8 getan.

Zur Definition 2.9 des Homöomorphismus: Um uns diese Definition zu erklären, geben wir ein paar Standardbeispiele an.

Beispiel 29

- Wir betrachten als erstes den \mathbb{R}^n und die offene Einheitskugel. Wir zeigen, dass der \mathbb{R}^n und die offene Einheitskugel

$U(0, 1) := \{x \in \mathbb{R}^n : \|x\| < 1\}$ homöomorph sind. Das Schwierige daran ist, den Homöomorphismus anzugeben. Beispielsweise ist ein Homöomorphismus $f : \mathbb{R}^n \rightarrow U(0, 1)$ gegeben durch

$$x \mapsto f(x) := \frac{x}{1 + \|x\|}.$$

Diese Abbildung ist sicherlich stetig, da sie eine Zusammensetzung aus stetigen Funktionen ist. Außerdem ist sie bijektiv. Die Umkehrabbildung ist gegeben durch

$$f^{-1} : U(0, 1) \rightarrow \mathbb{R}^n, \quad x \mapsto \frac{x}{1 - \|x\|}.$$

Diese ist als Zusammensetzung von stetigen Funktionen wieder stetig.

- Wir möchten einen Homöomorphismus von dem offenen Würfel

$$V := \{(x_1, \dots, x_n) \in \mathbb{R}^n : |x_i| < 1 \text{ für } i = 1, \dots, n\}$$

auf die Einheitskugel $U(0, 1) := \{x \in \mathbb{R}^n : \|x\| < 1\}$ konstruieren und lassen uns von dem ersten Punkt in diesem Beispiel leiten. Wir wissen ja bereits, dass $\varphi : (-1, 1) \rightarrow \mathbb{R}$ mit

$$\psi \mapsto \varphi(\psi) := \frac{\psi}{1 - |\psi|}$$

ein Homöomorphismus ist. Es ergibt sich ebenfalls, dass $\phi : V = (-1, 1)^n \rightarrow \mathbb{R}^n$ mit

$$(x_1, \dots, x_n) \mapsto (\varphi(x_1), \dots, \varphi(x_n))$$

ein Homöomorphismus ist. Nach dem ersten Punkt in diesem Beispiel ist $f : \mathbb{R}^n \rightarrow U(0, 1)$ mit $x \mapsto \frac{x}{1 + \|x\|}$ ein Homöomorphismus. Die Komposition $f \circ \phi$ bildet nun den gesuchten Homöomorphismus des Würfels V auf die Einheitskugel $U(0, 1)$. ■

Beispiel 30

Betrachten wir nun noch einmal ein Beispiel einer bijektiven stetigen Abbildung, deren Umkehrung nicht stetig ist, sodass die Abbildung damit kein Homöomorphismus ist. Seien

$$M := [0, 2\pi) \subset \mathbb{R} \quad \text{und} \quad N := \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \subset \mathbb{R}^2.$$

Dabei verstehen wir diese Mengen jeweils mit der von \mathbb{R} bzw. \mathbb{R}^2 induzierten Metrik. Die Abbildung

$$f : M \rightarrow N, \quad t \mapsto (\cos(t), \sin(t))$$

ist stetig und bijektiv. Die Umkehrabbildung $f^{-1} : N \rightarrow M$ ist aber unstetig in dem Punkt $(1, 0)$. Denn die Punktfolge

$$p_k := \left(\cos \left(2\pi - \frac{1}{k} \right), \sin \left(2\pi - \frac{1}{k} \right) \right)$$

mit $k \geq 1$ konvergiert für $k \rightarrow \infty$ gegen den Punkt $(1, 0) = f(0)$. Die Folge $f^{-1}(p_k) = 2\pi - \frac{1}{k}$, $k \geq 1$ konvergiert aber nicht gegen 0, sondern gegen 2π . Die Abbildung f rollt das halboffene Intervall $[0, 2\pi)$ zu einer Kreislinie S^1 zusammen, während die Umkehrabbildung den Kreis an dem Punkt $(1, 0)$ aufschneidet und entrollt. Das halboffene Intervall $[0, 2\pi)$ ist also nicht homöomorph zur Kreislinie S^1 . ■

Zur Definition 2.10 des Fixpunktes: Ein triviales und einfaches Beispiel ist das folgende.

Beispiel 31

Ist f die Identität auf M , so sind alle Punkte $m \in M$ Fixpunkte von f . Das ist klar. ■

Zur Definition 2.11 der Kontraktion: Offensichtlich sind Kontraktionen stetig. Dieses können wir sofort festhalten.

Beispiel 32

- Ist f die Identität, so ist dies keine Kontraktion. Man könnte jetzt sagen, dass es die Konstante $C = 1$ doch tut, denn es gilt

$$d(f(x), f(y)) = 1 \cdot d(x, y) = d(x, y),$$

aber laut Definition 2.11 muss C im halboffenen Intervall $[0, 1)$ liegen.

- Seien $M = \mathbb{R}$, $X = [1, 3]$ und $f(x) = \frac{1}{2} \left(x + \frac{2}{x}\right)$. Aus dem Mittelwertsatz, den ihr alle noch aus der Analysis 1 kennen solltet, ergibt sich

$$|f(x) - f(y)| = |f'(c)| \cdot |x - y|$$

mit $x \leq c \leq y$. Aus $f'(x) = \frac{1}{2} - \frac{1}{x^2}$ ergibt sich $|f'(c)| \leq \frac{1}{2}$ für $1 \leq c \leq 3$. Also ist f nach Definition 2.11 eine Kontraktion. ■

2.4 Erklärungen zu den Sätzen und Beweisen

Zum Satz 2.1, dass die Zusammensetzung stetiger Funktionen wieder stetig ist: Um uns diese Definition zu verdeutlichen, betrachten wir ein paar Beispiele.

Beispiel 33

Die Funktionen

$$\begin{aligned} f(x, y) &= x^2 + y^2, \quad g(x, y, z) = x^4 + \frac{2x^2}{y^2 + 1} + z, \\ h(x, y) &= 4x^3y - 3xy^2 + 2y + 11, \quad l(x, y) = \cos(xy) \end{aligned}$$

sind stetig, da sie aus stetigen Funktionen zusammengesetzt sind. ■

Merke also: Um eine Funktion mehrerer Veränderlicher auf Stetigkeit zu untersuchen, sollte man zunächst einmal überprüfen, ob sie aus stetigen Funktionen zusammengesetzt sind. Daraus folgt dann die Stetigkeit der Funktion selbst.

Zum Stetigkeitskriterium (Satz 2.2): Dieses Stetigkeitskriterium liefert ein nettes Kriterium, mit dem man ebenfalls Abbildungen auf Stetigkeit oder Nicht-Stetigkeit untersuchen kann und zeigt, dass die Ausführungen in Kapitel 1 nicht umsonst waren.

Beispiel 34

Die Stetigkeit einer Abbildung sagt nichts über die Bilder offener Mengen aus. Das heißt, die Bilder offener Mengen können durchaus abgeschlossen sein, so wie die Bilder abgeschlossener Mengen auch offen sein können. Dazu betrachte man beispielsweise die Funktionen $f(x) = 1$ und $g(x) = \arctan(x)$. Es ist falsch zu sagen, dass abgeschlossen das Gegenteil von offen ist. ■

Stetige Bilder offener Mengen brauchen keineswegs offen sein, sondern können alles Mögliche sein: offen, abgeschlossen oder auch keins von beiden. Bedenkt das bitte immer!

Zum Satz 2.3: Dieser Satz sagt einfach nur aus, dass die Bilder kompakter Mengen unter stetigen Abbildungen wieder kompakt sind.

Zum Satz 2.4: Hier nur soviel: Sowohl Satz 2.3 als auch Satz 2.4 gelten für topologische Räume.

Als kleine Übungsaufgabe zu diesem Satz beweist einmal Folgendes: Seien (M, d) ein metrischer Raum, $X \subset M$ eine Teilmenge und $x \in M$. Wir definieren den Abstand des Punktes x von der Menge X als die „Distanz“

$$\text{dist}(x, X) := \inf\{d(x, y) : y \in X\}.$$

Sei K eine weitere Teilmenge von M . Hierfür definieren wir

$$\text{dist}(K, X) := \inf\{\text{dist}(x, X) : x \in K\}.$$

Nun zeigt: Sind X abgeschlossen, K kompakt und $X \cap K = \emptyset$, so ist $\text{dist}(K, X) > 0$. Kleiner Tipp: Um Satz 2.4 anwenden zu können, so müsst ihr erst einmal die Stetigkeit der Abbildung $x \mapsto \text{dist}(x, X)$ nachweisen.

Zum Satz 2.5: Dieser Satz sagt aus, dass stetige Abbildungen auf kompakten Mengen sogar gleichmäßig stetig sind.

Zum Banach'schen Fixpunktsatz (Satz 2.6): Der Fixpunktsatz von Banach ist ein sehr mächtiges Werkzeug, das wir jetzt zur Verfügung haben. Die Aussage an sich ist klar. Der Beweis sollte, hoffentlich, auch verständlich gewesen sein. Merkt euch unter anderem, dass man bei Eindeutigkeitsbeweisen dies immer so macht: Man nimmt an, es gibt beispielsweise wie hier zwei Fixpunkte und zeigt, dass diese gleich sind. Mehr sei an dieser Stelle nicht gesagt. Wenn ihr das Buch aufmerksam weiter lest, so werdet ihr die Stellen finden, an denen wir den Fixpunktsatz verwenden.

Beispiel 35

■ Seien $M := (0, \infty) \subset \mathbb{R}$ und $f : M \rightarrow M$ definiert durch

$$f(x) = \frac{1}{2}x.$$

Die Abbildung besitzt offenbar keinen Fixpunkt auf $(0, \infty)$ (dies ist schon rein anschaulich klar, denn f ist eine Ursprungsgerade mit Steigung $1/2$), aber es gilt:

$$|f(x) - f(y)| = \left| \frac{1}{2}x - \frac{1}{2}y \right| = \frac{1}{2}|x - y| \quad \forall x, y \in M.$$

Woran liegt es jetzt aber, dass f keinen Fixpunkt besitzt? Na ja, ganz einfach. Es sind nicht alle Voraussetzung des Banach'schen Fixpunktsatzes erfüllt, denn M ist nicht abgeschlossen und damit nicht vollständig. Auf die Vollständigkeit des metrischen Raumes kann also nicht verzichtet werden.

- Seien nun $M = [0, \infty)$ und $f : M \rightarrow M$ definiert durch

$$f(x) = \frac{\pi}{2} + x + \arctan(x).$$

Der Mittelwertsatz aus der Analysis 1 sagt uns, dass f keine Kontraktion ist und daher Satz 2.6 nicht anwendbar ist, denn es gilt:

$$|f(x) - f(y)| = \left| 1 - \frac{1}{(1 + \xi)^2} \right| \cdot |x - y| < |x - y|.$$

Weitere Beispiele folgen im Buch. Geht auf Entdeckungstour. ■

Eine Anwendung in einigen Beweisen werden wir in diesem Buch ebenfalls noch entdecken.

Zum Satz 2.7: Sind V und W endlichdimensionale Vektorräume, so lässt sich jede lineare Abbildung als Matrix darstellen. Da die Matrix nur aus endlich vielen Konstanten besteht, folgt aus den Grenzwertsätzen sofort die Stetigkeit einer Abbildung $L : V \rightarrow W$, das heißt zusammengefasst, dass lineare Abbildungen zwischen endlichdimensionalen Vektorräumen stetig sind. Der Satz 2.7 gibt nun eine Verallgemeinerung auch auf unendlichdimensionale Vektorräume.

Beispiel 36

Um die Stetigkeit eines linearen Operators zu zeigen, reicht es also nachzuweisen, dass er beschränkt ist. Wir zeigen: Die Abbildung

$$D : \mathcal{C}^1[a, b] \rightarrow \mathcal{C}[a, b], \quad f \mapsto f',$$

wird stetig, wenn man $\mathcal{C}^1[a, b]$ mit der $\|\cdot\|_{\mathcal{C}^1}$, das heißt $\|f\|_{\mathcal{C}^1} := \sup\{|f(x)| + |f'(x)| : x \in [a, b]\}$ und $\mathcal{C}[a, b]$ mit der Supremumsnorm $\|f\|_{\infty} := \sup\{|f(x)| : x \in [a, b]\}$ versieht. Aus den Ableitungsregeln folgt sofort, dass D linear ist. Nach dem Satz 2.7 reicht es also zu zeigen, dass ein $C \geq 0$ existiert, sodass

$$\|D(f)\|_{\infty} \leq C\|f\|_{\mathcal{C}^1} \quad f \in \mathcal{C}^1[a, b].$$

Dies ist aber klar, denn für alle $f \in \mathcal{C}^1[a, b]$ gilt:

$$\begin{aligned} \|D(f)\|_\infty &= \sup\{|f'(x)| : x \in [a, b]\} \\ &\leq \sup\{|f(x)| + |f'(x)| : x \in [a, b]\} \\ &= \underbrace{1}_{=:C} \cdot \|f\|_{\mathcal{C}^1}. \end{aligned}$$

Also ist D stetig. ■

3 Differenzierbare Abbildungen

Übersicht

3.1	Definitionen	65
3.2	Sätze und Beweise	67
3.3	Erklärungen zu den Definitionen	70
3.4	Erklärungen zu den Sätzen und Beweisen	82

In diesem Kapitel wollen wir uns schrittweise anschauen, wie wir den Begriff der Differenzierbarkeit für reelle Funktionen der Form $f : \mathbb{R} \rightarrow \mathbb{R}$ aus der Analysis 1 auf die weitaus größere Klasse von Abbildungen zwischen Banach-Räumen, insbesondere Abbildungen der Form $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, verallgemeinern können. Wie im letzten Kapitel halten wir es hier auch wieder: Die Definitionen sind für allgemeine Banach-Räume definiert und bei den Beispielen in den Erklärungen verwenden wir nur den \mathbb{R}^n .

3.1 Definitionen

Einige der folgenden Definitionen führen wir für allgemeine Banach-Räume ein. Bis jetzt haben wir zwar meistens nur Abbildungen der Form $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ betrachtet, aber dies ist dann einfach nur ein Spezialfall unserer Definitionen. Man sollte sich schnell an die Allgemeinheit gewöhnen :-).

Definition 3.1 (differenzierbare Abbildung, Differential)

Seien $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ zwei Banach-Räume, $\Omega \subset V$ offen und $f : \Omega \rightarrow W$ eine Abbildung. Dann heißt f im Punkt $x_0 \in \Omega$ **(total) differenzierbar**, wenn es eine beschränkte (also stetige) lineare Abbildung $L : V \rightarrow W$ gibt, sodass

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0) - L(x - x_0)\|_W}{\|x - x_0\|_V} = 0.$$

In diesem Fall nennen wir die Abbildung L das **Differential** oder auch **Ableitung** an der Stelle x_0 und schreiben hierfür $Df(x_0)$. f heißt (total) differenzierbar in Ω , wenn sie in jedem $x_0 \in \Omega$ (total) differenzierbar ist. Ist f in jedem

Punkt $x_0 \in \Omega$ differenzierbar und hängt $Df(x_0)$ stetig von x_0 ab, so heißt f in Ω **stetig differenzierbar**. Wir setzen

$$\mathcal{C}^1(\Omega, W) := \{f : \Omega \rightarrow W : f \text{ ist in } \Omega \text{ stetig differenzierbar}\}.$$

Anmerkung: Für Abbildungen $f : V \rightarrow W$ zwischen endlichdimensionalen Banach-Räumen ist die Forderung der Beschränktheit redundant, da man zeigen kann, dass lineare Abbildungen in diesem Fall automatisch schon beschränkt sind. Genauer gilt dies nur für stetige Abbildungen, aber lineare Abbildungen zwischen endlichdimensionalen Banach-Räumen sind immer stetig.

Satz 3.2 sagt uns, dass die Ableitung, wenn sie existiert, eindeutig ist.

Definition 3.2 (Richtungsableitung)

Seien V, W Banach-Räume und $\Omega \subset V$ offen. Wir sagen eine Abbildung $f : \Omega \rightarrow W$ besitzt an der Stelle $x_0 \in \Omega$ eine **Richtungsableitung** in Richtung $v \in V$, wenn der Grenzwert

$$\lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t}$$

existiert. In diesem Fall bezeichnen wir den Grenzwert mit $D_v f(x_0)$, und dieser heißt die Richtungsableitung von f im Punkt x_0 in Richtung v .

Definition 3.3 (partielle Ableitung)

Ω sei eine offene Teilmenge des \mathbb{R}^n , W ein Banach-Raum und $f : \Omega \rightarrow W$ eine Abbildung. f heißt in $x_0 \in \Omega$ nach der j -ten Variablen **partiell differenzierbar**, wenn f an der Stelle x_0 eine Richtungsableitung in Richtung e_j besitzt. Hierbei ist (e_1, \dots, e_n) die Standardbasis des \mathbb{R}^n . Die entsprechende Ableitung wird dann mit $D_j f(x_0)$ oder auch mit $\frac{\partial f}{\partial x_j}(x_0)$ bezeichnet.

Anmerkung: Ab und an schreiben wir auch $f_{x_j}(x_0)$ für die partielle Ableitung nach x_j oder $\partial_{x_j} f(x_0)$.

Definition 3.4 (Gradient)

Seien $\Omega \subset \mathbb{R}^n$ offen und $f : \Omega \rightarrow \mathbb{R}$ in $x_0 \in \Omega$ partiell differenzierbar. Der **Gradient** von f an der Stelle x_0 ist der Vektor

$$\text{grad} f(x_0) := \nabla f(x_0) := \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0) e_i.$$

Definition 3.5 (Jakobi-Matrix)

Ist $f : \Omega \rightarrow \mathbb{R}^m$ mit $\Omega \subset \mathbb{R}^n$ in $x_0 \in \Omega$ differenzierbar, so nennt man die Matrix

$$J_f(x) := \left(\frac{\partial f_i}{\partial x_j}(x_0) \right)_{i=1, \dots, m, j=1, \dots, n}$$

die **Jakobi-Matrix** von f an der Stelle x_0 .

Definition 3.6 (zweimal differenzierbar)

V und W seien Banach-Räume, $\Omega \subset V$ offen und $f : \Omega \rightarrow W$ differenzierbar in Ω . Ist dann die Abbildung $Df : \Omega \rightarrow W$ in $x_0 \in \Omega$ differenzierbar, so heißt f an der Stelle x_0 **zweimal differenzierbar** und wir schreiben $D^2f(x_0)$ für $(D(Df))(x_0)$.

Definition 3.7 (höhere Ableitung)

Für zwei Banach-Räume V und W , eine offene Teilmenge $\Omega \subset V$ und $k \in \mathbb{N}_0$ setzen wir

$$C^k(\Omega, W) := \{f : \Omega \rightarrow W : f \text{ ist in } \Omega \text{ } k\text{-mal stetig differenzierbar}\}$$

sowie

$$C^\infty(\Omega, W) := \bigcap_{k \in \mathbb{N}_0} C^k(\Omega, W).$$

3.2 Sätze und Beweise

Satz 3.1 (Linearität der Ableitung)

Gegeben seien zwei Abbildungen $f, g : \Omega \rightarrow W$, die in $x_0 \in \Omega$ differenzierbar sind und eine Konstante $\lambda \in \mathbb{R}$. Dann sind auch die Abbildungen $f + g$, λf in x_0 differenzierbar mit

$$D(f + g)(x_0) = Df(x_0) + Dg(x_0), \quad D(\lambda f)(x_0) = \lambda Df(x_0).$$

Satz 3.2 (Eindeutigkeit der Ableitung)

$f : \Omega \rightarrow W$ sei wie in Definition 3.1. Dann ist L eindeutig bestimmt.

Beweis: Angenommen, es gibt zwei beschränkte lineare Operatoren L und L' mit dieser Eigenschaft. Wir zeigen nun, dass dann schon $L = L'$ gelten muss. Wir erhalten sofort

$$\begin{aligned} & \frac{\|(L - L')(x - x_0)\|_W}{\|x - x_0\|_V} \\ &= \frac{\|f(x) - f(x_0) - L'(x - x_0) - (f(x) - f(x_0) - L(x - x_0))\|_W}{\|x - x_0\|_V} \\ &\leq \frac{\|f(x) - f(x_0) - L'(x - x_0)\|_W}{\|x - x_0\|_V} + \frac{\|f(x) - f(x_0) - L(x - x_0)\|_W}{\|x - x_0\|_V}. \end{aligned}$$

Die rechte Seite strebt für $x \rightarrow x_0$ gegen Null. Daher existiert zu $\varepsilon > 0$ ein $\delta > 0$, sodass

$$\left\| (L - L') \left(\frac{x - x_0}{\|x - x_0\|_V} \right) \right\|_W \leq \varepsilon \quad \forall x \in \Omega \quad \text{mit } 0 < \|x - x_0\|_V < \delta.$$

Da Ω nach Voraussetzung offen ist, existiert ein $r > 0$ mit $U(x_0, r) \subset \Omega$. Sei $y \in V$ mit $\|y\|_V = 1$ beliebig. Dann finden wir hierzu ein x , zum Beispiel mit $x := x_0 + \min\{\delta/2, r/2\}y$, mit

$$x \in \Omega, \quad 0 < \|x - x_0\|_V < \delta, \quad \frac{x - x_0}{\|x - x_0\|_V} = y.$$

Somit ist

$$\|(L - L')(y)\|_W \leq \varepsilon \quad \forall y \in V, \quad \|y\|_V = 1.$$

Da $\varepsilon > 0$ beliebig war, und $L - L'$ auch linear ist, folgt hieraus die Behauptung, denn

$$L - L' = 0 \Leftrightarrow L = L'.$$

q.e.d.

Satz 3.3 (differenzierbare Abbildungen sind stetig)

Die Abbildung $f : \Omega \rightarrow W$ sei differenzierbar in $x_0 \in \Omega$. Dann ist f in x_0 auch stetig.

Beweis: Zum Beweis verwenden wir die Definition 3.1 und die der Stetigkeit (Definition 2.4). Sei f differenzierbar, das heißt, es gibt ein L mit

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0) - L(x - x_0)\|_W}{\|x - x_0\|_V} = 0.$$

Dann existiert ein $\delta > 0$ mit

$$\|f(x) - f(x_0) - L(x - x_0)\|_W \leq \|x - x_0\|_V \quad \forall x \in \Omega, \quad \|x - x_0\|_V < \delta.$$

Aus der Dreiecksungleichung (hier haben wir genauer die inverse Dreiecksungleichung verwendet, also $|a - b| \geq ||a| - |b||$) ergibt sich nun

$$\|f(x) - f(x_0)\|_W \leq \|x - x_0\|_V + \|L(x - x_0)\|_W, \quad \forall x \in \Omega, \quad \|x - x_0\|_V < \delta.$$

Die rechte Seite lässt sich mit der Operatornorm (Supremumsnorm, siehe auch das Beispiel 1 aus Kapitel 1) $\|L\|$ weiter abschätzen zu

$$\|f(x) - f(x_0)\|_W \leq (1 + \|L\|)\|x - x_0\|_V, \quad \forall x \in \Omega, \quad \|x - x_0\|_V < \delta.$$

Da nach Voraussetzung $\|L\| < \infty$ gilt, existiert eine Konstante $C > 0$ mit

$$\|f(x) - f(x_0)\|_W \leq C\|x - x_0\|_V, \quad \forall x \in \Omega, \quad \|x - x_0\|_V < \delta.$$

Dies ist gerade die Definition der Lipschitz-Stetigkeit und damit folgt die Behauptung. q.e.d.

Satz 3.4 (Kettenregel)

$(U, \|\cdot\|_U)$, $(V, \|\cdot\|_V)$ und $(W, \|\cdot\|_W)$ seien Banach-Räume, $\Omega \subset U$ und $\tilde{\Omega} \subset V$ jeweils offen. Sind $f : \Omega \rightarrow V$, $g : \tilde{\Omega} \rightarrow W$ Abbildungen mit $f(\Omega) \subset \tilde{\Omega}$, und sind f in $x_0 \in \Omega$ und g in $f(x_0) \in \tilde{\Omega}$ differenzierbar, so ist auch die Verkettung $g \circ f : \Omega \rightarrow W$ in x_0 differenzierbar, und es gilt die Gleichung

$$D(g \circ f)(x_0) = Dg(f(x_0)) \circ Df(x_0). \quad (3.1)$$

Satz 3.5 (Mittelwertungleichung)

W sei ein Banach-Raum und $f : [a, b] \rightarrow W$ eine stetige Abbildung, die auf (a, b) differenzierbar sei mit $\|Df(x)\| \leq M$ für alle $x \in (a, b)$. Dann gilt:

$$\|f(x) - f(y)\| \leq M|x - y| \quad \forall x, y \in (a, b).$$

Satz 3.6 (Zusammenhang zwischen Differential und Richtungsableitung)

V, W seien zwei Banach-Räume, $\Omega \subset V$ sei offen und $f : \Omega \rightarrow W$ in $x_0 \in \Omega$ differenzierbar. Dann existieren in x_0 sämtliche Richtungsableitungen, und es ist

$$Df(x_0)(v) = D_v f(x_0).$$

Satz 3.7

Es sei $\Omega \subset \mathbb{R}^n$ offen. Dann ist eine Abbildung $f : \Omega \rightarrow \mathbb{R}^m$ genau dann in Ω stetig differenzierbar, wenn in jedem Punkt $x_0 \in \Omega$ sämtliche partiellen Ableitungen $\frac{\partial f^i}{\partial x^j}(x)$, $i = 1, \dots, m$, $j = 1, \dots, n$ existieren und jeweils stetig von x abhängen. Hierbei ist $f^j : \Omega \rightarrow \mathbb{R}$ die j -te Koordinatenfunktion von f . Insbesondere gilt für das Differential $Df(x)$ die Gleichung

$$Df(x)(v) = J_f(x) \cdot v^T,$$

wobei wir mit $J_f(x)$ die Jakobi-Matrix von f bezeichnen (siehe Definition 3.5).

Satz 3.8 (Lemma von Schwarz)

V und W seien zwei Banach-Räume, $\Omega \subset V$ offen und $f : \Omega \rightarrow W$ sei in $x_0 \in \Omega$ zweimal stetig differenzierbar. Dann ist die stetige bilineare Abbildung $D^2 f(x_0) := V \times V \rightarrow W$ symmetrisch, das heißt, es gilt:

$$D^2 f(x_0)(v_1, v_2) = D^2 f(x_0)(v_2, v_1) \quad \forall v_1, v_2 \in V.$$

Satz 3.9

Eine Abbildung $f : \Omega \rightarrow W$ ist genau dann in Ω zweimal stetig differenzierbar, wenn sie dort zweimal partiell differenzierbar ist und sämtliche partiellen Ableitungen dort stetig sind.

Satz 3.10 (Taylor'sche Formel)

V sei ein Banach-Raum, $\Omega \subset V$ offen und $f \in \mathcal{C}^k(\Omega, \mathbb{R})$. Für ein $x_0 \in \Omega$ und $t \in V$ gelte

$$\{x_0 + rt \text{ für } r \in [0, 1]\} \subset \Omega.$$

Dann existiert ein $\theta \in [0, 1]$, sodass

$$\begin{aligned} f(x_0 + t) &= f(x_0) + Df(x_0)(t) + \frac{1}{2!} D^2 f(x_0)(t, t) + \dots \\ &\quad \dots + \frac{1}{(k-1)!} D^{k-1} f(x_0) \underbrace{(t, \dots, t)}_{(k-1)\text{-mal}} + \frac{1}{k!} D^k f(x_0 + \theta t) \underbrace{(t, \dots, t)}_{k\text{-mal}}. \end{aligned}$$

Beweis: Wir betrachten die Funktion

$$g : [0, 1] \rightarrow \mathbb{R}, \quad g(r) := f(x_0 + rt).$$

Die Voraussetzung $f \in \mathcal{C}^k(\Omega, \mathbb{R})$ und die Kettenregel implizieren per Induktion, dass g k -mal stetig differenzierbar ist. Induktiv erhalten wir für die Ableitungen

$$g^{(j)}(r) = D^j f(x_0 + rt)(t, \dots, t). \quad (3.2)$$

Da g nur von einer Variablen abhängt, können wir die aus der Analysis 1 bereits bekannte Taylor-Formel benutzen und erhalten so

$$g(1) = \sum_{j=0}^{k-1} \frac{1}{j!} g^{(j)}(0) + \frac{1}{k!} g^{(k)}(\theta) \quad (3.3)$$

für ein $\theta \in [0, 1]$. Kombinieren wir die beiden Gleichungen (3.2) und (3.3), so liefert dies die Behauptung. q.e.d.

3.3 Erklärungen zu den Definitionen

Zur Definition 3.1 einer differenzierbaren Abbildung: Wir erinnern uns zunächst einmal an die Definition einer differenzierbaren reellen Funktion. Eine Funktion $f : D \rightarrow \mathbb{R}$ mit $D \subset \mathbb{R}$ ist genau dann differenzierbar im Punkt $x_0 \in D$, wenn der Grenzwert

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. Diesen Grenzwert bezeichnen wir mit $f'(x_0)$ und nennen ihn die Ableitung von f an der Stelle x_0 . f ist also genau dann an der Stelle x_0 differenzierbar, wenn es eine Konstante $a \in \mathbb{R}$ gibt mit

$$\lim_{x \rightarrow x_0} \frac{|f(x) - f(x_0) - a(x - x_0)|}{|x - x_0|} = 0.$$

In diesem Fall ist a die Ableitung, das heißt, $a = f'(x_0)$. Die letzte Gleichung haben wir in [MK09] auch noch so interpretiert, dass sich die Funktion f an der Stelle x_0 durch die lineare Funktion

$$x \mapsto a(x - x_0)$$

approximieren lässt. Es gilt dann nämlich

$$f(x) = f(x_0) + a(x - x_0) + \rho(x)$$

mit einer Restfunktion $\rho : D \rightarrow \mathbb{R}$, welche

$$\lim_{x \rightarrow x_0} \frac{\rho(x)}{x - x_0} = 0$$

erfüllt. Dieses Konzept motiviert unsere Definition 3.1. Sei jetzt nämlich eine Abbildung

$$f : \mathbb{R}^n \rightarrow \mathbb{R}^m$$

gegeben. Ist (e_1, \dots, e_n) eine Basis des \mathbb{R}^n , so kann man mit Wissen aus der Linearen Algebra 1 jeden Vektor $x \in \mathbb{R}^n$ als Linearkombination der Basisvektoren darstellen, das heißt,

$$x = \sum_{i=1}^n x_i e_i,$$

wobei die Koeffizienten x_1, \dots, x_n reelle Zahlen sind. Wir nennen diese Koeffizienten die Koordinaten von x bzgl. der Basis (e_1, \dots, e_n) . Wählt man die Standardbasis, so nennen wir die Koordinaten die kartesischen Koordinaten von x . Der Koordinatenvektor von x schreibt sich demnach in Zeilen- oder Spaltenform, das heißt,

$$x = (x_1, \dots, x_n) \quad \text{oder} \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

je nachdem, ob wir Vektoren in Zeilen- oder Spaltenform schreiben wollen. Sind x_1, \dots, x_n Koordinaten eines Vektors $x \in \mathbb{R}^n$ bezüglich einer Basis (e_1, \dots, e_n) , so schreiben wir

$$f(x_1, \dots, x_n) := f\left(\sum_{i=1}^n x_i e_i\right) = f(x).$$

Ist $(\tilde{e}_1, \dots, \tilde{e}_m)$ eine Basis der Zielmenge \mathbb{R}^m , so lässt sich $f(x)$ auch in der Form

$$f(x) = \sum_{j=1}^m f_j(x) \tilde{e}_j$$

schreiben. Die Funktion

$$x \mapsto f_j(x)$$

nennen wir die j -te Koordinatenfunktion von f .

In Analogie zu der differenzierbaren Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ würde es jetzt also durchaus Sinn machen, wenn wir $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ genau dann im Punkt $x_0 \in \mathbb{R}^n$ differenzierbar nennen, wenn es eine reelle $(m \times n)$ -Matrix gäbe, für die

a)

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0) - A \cdot (x - x_0)\|}{\|x - x_0\|} = 0,$$

falls wir x und $f(x)$ jeweils als Spaltenvektoren auffassen, oder

b)

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0) - (x - x_0) \cdot A^T\|}{\|x - x_0\|} = 0,$$

falls wir x und $f(x)$ jeweils als Zeilenvektoren betrachten, oder

c)

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0) - A \cdot (x - x_0)^T\|}{\|x - x_0\|} = 0,$$

falls wir x als einen Zeilenvektor und $f(x)$ als einen Spaltenvektor auffassen. Das hat etwas mit der Definition der Matrizenmultiplikation zu tun, wie ihr euch sicherlich erinnert :-).

Mit dem Punkt „ \cdot “ meinen wir natürlich das übliche Matrizenprodukt aus der Linearen Algebra 1. Leider gibt es für beide Schreibweisen gute Gründe. Bei der Zeilenform beispielsweise ist die Schreibweise übersichtlicher und spart Platz. Also gut für uns, denn so können wir hier in diesem Buch mehr schreiben. Da wir das Differential einer differenzierbaren Abbildung aber eher durch eine Matrix darstellen werden, die von links mit dem Vektor in üblicher und bekannter Weise multipliziert wird, bietet sich andererseits die Schreibweise in Spaltenform an. Kurz gesagt: Wir werden zwischen allen drei Schreibweisen in diesem Buch hin und her springen und alle einmal oder mehrmals benutzen. Dieser Ansatz motiviert unsere Definition 3.1.

Wir wollen diese Definition noch an ein paar Beispielen einüben.

Beispiel 37

- Die Abbildung $f : V \rightarrow W$, wobei V und W die Banach-Räume aus Definition 3.1 sind, sei konstant, das heißt, $f(x) = c$ für ein festes $c \in W$. Dann behaupten wir, dass f überall differenzierbar ist und das Differential $Df(x) = 0$ lautet. Dies ist einfach und folgt sofort aus der obigen Definition, denn es gilt:

$$f(x) - f(x_0) - 0(x - x_0) = 0,$$

sodass dann

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0) - L(x - x_0)\|_W}{\|x - x_0\|_V} = 0$$

gilt. Bei uns ist also $L = Df(x) = 0$.

- Wir beweisen: Eine lineare Abbildung $L : V \rightarrow W$ ist genau dann differenzierbar, wenn sie beschränkt ist, und in diesem Fall gilt $DL(x) = L$ für alle $x \in V$. Dies sieht man so: Da L linear ist, erhalten wir

$$L(x) - L(x_0) - L(x - x_0) = L(x) - L(x_0) - L(x) + L(x_0) = 0,$$

und hieraus ergibt sich sofort die Behauptung, wenn man sich an die Definition 3.1 der Differenzierbarkeit erinnert.

- Wir betrachten nun das bekannte Beispiel

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = ax.$$

Bekanntlich gilt $f'(x) = a$ für alle $x \in \mathbb{R}$. Wie verträgt sich dies aber mit unserer verallgemeinerten Definition 3.1? Eigentlich ganz einfach, denn wir können die Konstante $a \in \mathbb{R}$ ja als eine (1×1) -Matrix auffassen und so ist wieder alles im Lot.

Dazu sei gesagt, dass $Df(x) = f$ für $f(x) = a \cdot x$. Denkt nicht fälschlicherweise, dass $Df(x) = f'(x)$. Das ist so nicht korrekt. Es gilt $f'(x) = J_f$, aber $Df(x) = J_f \cdot x$.

- Das nächste Beispiel greift ein wenig vor. Wir werden in Definition 3.5 und den entsprechenden Erklärungen dazu noch einmal erklären, wie wir auf die Matrix $J_f(x) = (x_2, x_1)$ kommen. Also noch etwas Geduld. Aber nun zum Beispiel:

Wir behaupten, dass die Abbildung

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \right) = x_1 x_2$$

differenzierbar ist und $Df(x)$ durch die Multiplikation von links mit der Matrix

$$J_f = (x_2, x_1)$$

dargestellt wird. Hierbei seien $x_1, x_2 \in \mathbb{R}$ die kartesischen Koordinaten von $x \in \mathbb{R}^2$. Schreiben wir

$$f(x_1, x_2) = x_1 x_2$$

dagegen in Zeilenform, so wird $Df(x)$ durch die Multiplikation von rechts mit der Matrix $J_f(x)^T$ dargestellt, denn es gilt:

$$\left(J_f(x) \cdot v^T \right)^T = v \cdot J_f(x)^T.$$

Dies wollen wir schnell nachweisen. Schreiben wir die Vektoren in Spaltenform, so ergibt sich

$$\begin{aligned}
 & f(x) - f(y) - J_f(y) \cdot (x - y) \\
 &= x_1x_2 - y_1y_2 - (y_2, y_1) \cdot \begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \end{pmatrix} \\
 &= x_1x_2 - y_1y_2 - y_2x_1 + y_2y_1 - y_1x_2 + y_1y_2 \\
 &= x_1x_2 + y_1y_2 - y_2x_1 - y_1x_2 \\
 &= (x_1 - y_1)(x_2 - y_2).
 \end{aligned}$$

Demnach gilt:

$$\begin{aligned}
 \frac{\|f(x) - f(y) - Df(y)(x - y)\|}{\|x - y\|} &\leq \frac{|x_1 - y_1| \cdot |x_2 - y_2|}{\|x - y\|} \\
 &\leq \frac{1}{2} \cdot \frac{(x_1 - y_1)^2 + (x_2 - y_2)^2}{\|x - y\|} \\
 &\leq \|x - y\| \rightarrow 0, \quad \text{für } x \rightarrow y.
 \end{aligned}$$

Damit ist alles gezeigt.

Dies soll an Beispielen genügen. ■

In den folgenden Erklärungen zu den Definitionen und Sätzen werden wir natürlich weitaus praktischere Werkzeuge kennenlernen, um eine Funktion auf Differenzierbarkeit zu untersuchen und dessen Differential zu bestimmen. Das ist so wie in Analysis 1. Wenn man nur die Definition der Differenzierbarkeit hat, muss man erst einmal damit arbeiten und sich zunächst einfachere Dinge überlegen.

Zur Definition 3.2 der Richtungsableitung: Die Definition ist wie folgt motiviert: Wir betrachten die Situation, dass $f : V \rightarrow W$ eine Abbildung zwischen Banach-Räumen und $x_0 \in V$ fest gewählt ist. Dann erhalten wir für jedes $v \in V$ eine Abbildung

$$u_v : \mathbb{R} \rightarrow W, \quad u_v(t) := f(x_0 + tv).$$

Ist u_v an der Stelle $t = 0$ differenzierbar, so beschreibt $\dot{u}_v(0)$ die Ableitung von f in Richtung v . Dies ist auch klar. Denn wir laufen quasi zeitabhängig auf der Abbildung in eine bestimmte Richtung v . Satz 3.6 zeigt, dass dies mit der oben definierten Ableitung (siehe Definition 3.1) übereinstimmt. Betrachten wir ein Beispiel.

Beispiel 38

Seien $a \in \mathbb{R}^3$ ein Vektor im \mathbb{R}^3 und $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ die Abbildung $f(v) := a \times v$. Hierbei ist für zwei Vektoren

$$a = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}, \quad v = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

das Kreuzprodukt \times definiert als

$$a \times v := \begin{pmatrix} a_2 v_3 - a_3 v_2 \\ a_3 v_1 - a_1 v_3 \\ a_1 v_2 - a_2 v_1 \end{pmatrix}.$$

Es gilt nun:

$$\begin{aligned} D_v f(x_0) &= \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} \\ &= \lim_{t \rightarrow 0} \frac{a \times (x_0 + tv) - a \times x_0}{t} \\ &= \lim_{t \rightarrow 0} \frac{ta \times v}{t} \\ &= a \times v. \end{aligned}$$

Wir haben also gesehen, dass die Richtungsableitung von f in Richtung eines Vektors v mit der eigentlichen Abbildung übereinstimmt. ■

Zur Definition 3.3 der partiellen Ableitung: Wir hatten ja gesehen, dass wir nun Funktionen mehrerer Veränderlicher betrachten. Das heißt, unsere Funktionen besitzen mehrere Variablen. Also können wir ja auch einfach nur nach einer Variablen ableiten. Dies bedeutet partiell abzuleiten. Ist eine Funktion beispielsweise von x und y abhängig und wir wollen partiell nach x ableiten, so betrachten wir y einfach als eine Konstante. Wenn ihr euch immer schon gefragt habt, wo der Unterschied zwischen $\frac{d}{dt}$ und $\frac{\partial}{\partial x_j}$ liegt, so klären wir euch nun auf :-): $\frac{d}{dt}$ schreibt man, wenn die Funktion nur von einer Variablen t abhängt. Hängt die Funktion von mehreren Veränderlichen ab, so schreiben wir für die partielle Ableitung $\frac{\partial}{\partial x_j}$. Dies ist wirklich wichtig. Für Physiker, die ja auch Analysis 2 hören müssen, gibt es noch einen weiteren kleinen „Unterschied“. Dieser ist physikalisch motiviert. Alle, die damit nichts zu tun haben wollen, springen gleich zu den Beispielen.

Wir betrachten die Bahn eines Punktes in der Ebene in Abhängigkeit von der Zeit t und eine Funktion f , die nicht nur von den Koordinaten x und y , sondern auch von der Zeit abhängt. Die Ortskoordinaten des sich bewegenden Punktes

beschreiben wir durch $x = g(t)$ und $y = h(t)$. Die Funktion, die sich ergibt, lautet $t \mapsto f(t, x, y) = f(t, g(t), h(t))$ und hängt in zweierlei Hinsicht von der Zeit t ab: Einerseits hängt f selbst von der Zeit ab, weil in der ersten Variable t auftaucht. Andererseits hängen die Ortskoordinaten von der Zeit t ab. Physiker sprechen dann von der partiellen Ableitung von f nach der Zeit t , wenn man die partielle Ableitung der ersten Funktion meint, genauer

$$\frac{\partial f}{\partial t}(t, x, y),$$

und von der totalen Ableitung von f nach der Zeit t , wenn man die Ableitung der zusammengesetzten Funktion meint, genauer

$$\frac{d}{dt}f(t, g(t), h(t)).$$

Diese beiden hängen natürlich auch miteinander zusammen. Es gilt:

$$\frac{d}{dt}f(t, g(t), h(t)) = \frac{\partial f}{\partial t} + \frac{\partial f}{\partial x} \frac{dx}{dt} + \frac{\partial f}{\partial y} \frac{dy}{dt}.$$

Schauen wir uns in diesem Zusammenhang noch ein paar Beispiele an.

Beispiel 39

Seien $c \in \mathbb{R}$ beliebig und

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad f(x, y) = \begin{pmatrix} \cos(x)(c + \cos(y)) \\ \sin(x)(c + \cos(y)) \\ \sin(y) \end{pmatrix}.$$

Wir können jetzt also nach x oder nach y partiell ableiten. Machen wir dies: Für die partiellen Ableitungen erhalten wir

$$D_1 f(x, y) = \frac{\partial f}{\partial x}(x, y) = \begin{pmatrix} -\sin(x)(c + \cos(y)) \\ \cos(x)(c + \cos(y)) \\ 0 \end{pmatrix}$$

und

$$D_2 f(x, y) = \frac{\partial f}{\partial y}(x, y) = \begin{pmatrix} -\cos(x)\sin(y) \\ -\sin(x)\sin(y) \\ \cos(y) \end{pmatrix}.$$

■

Zur Definition 3.4 des Gradienten: Der Gradient ist einfach nur ein Name, für das, was wir schon kennen. Denn partielle Ableitungen können wir schon berechnen, und wenn wir diese in einen Vektor eintragen, so erhalten wir den Gradienten.

Beispiel 40

Wir betrachten die Funktion $f(x, y) = x^2y^3 + xy^2 + 2y$. Wir wollen den Gradienten an der Stelle (x, y) und an der Stelle $(0, 1)$ berechnen. Zunächst benötigen wir die partiellen Ableitungen nach x und y . Wir schreiben jetzt auch f_x für $\frac{\partial f}{\partial x}$:

$$f_x(x, y) = 2y^3x + y^2, \quad f_y(x, y) = 3x^2y^2 + 2xy + 2.$$

Der Gradient ergibt sich damit als

$$\text{grad}f(x, y) = (2y^3x + y^2, 3x^2y^2 + 2xy + 2).$$

Wir können den Gradienten nun an der bestimmten Stelle $(0, 1)$ ausrechnen zu

$$\text{grad}f(0, 1) = (1, 2).$$

■

Anschaulich kann man sich den Gradienten auch ganz einfach vorstellen: Für eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ steht der Gradient senkrecht auf der Niveaulinie $f(x, y) = f(x_0, y_0)$. Der Gradient zeigt in Richtung maximaler Steigung. Das Negative des Gradienten zeigt damit in Richtung minimaler Steigung. Die Abbildung 3.1 verdeutlicht dieses. Wie kann man sich aber herleiten, dass der

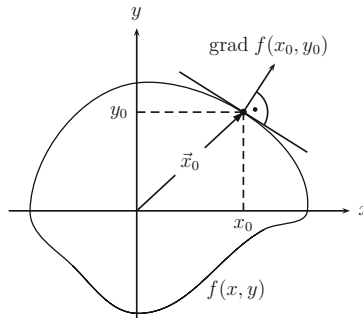


Abb. 3.1: Der Gradient steht senkrecht auf der Niveaulinie.

Gradient wirklich in Richtung des steilsten Anstiegs zeigt? Dies geht so:
Da für ein differenzierbares f

$$D_v f(x_0) = Df(x_0)(v) = \sum_{i=1}^n D_i f(x_0) v_i = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0) v_i$$

für $v = \sum_{i=1}^n v_i e_i$ gilt, gilt auch:

$$D_v f(x_0) = Df(x_0)(v) = \langle \nabla f(x_0), v \rangle. \quad (3.4)$$

Aus der Cauchy-Schwarz'schen Ungleichung

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|$$

folgt die Abschätzung

$$\|Df(x_0)\| = \sup_{\|v\|=1} \|Df(x_0)(v)\| = \sup_{\|v\|=1} \langle \nabla f(x_0), v \rangle \leq \|\nabla f(x_0)\|.$$

Ist $\nabla f(x_0) = 0$, so ist auch $\|Df(x_0)\| = 0$. Für den Fall, dass $\nabla f(x_0) \neq 0$, setzen wir den Einheitsvektor $v := \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|}$ in Gleichung (3.4) ein und erhalten für dieses v

$$Df(x_0) = \left\langle \nabla f(x_0), \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|} \right\rangle = \|\nabla f(x_0)\|,$$

sodass wegen $\|Df(x_0)\| \leq \|\nabla f(x_0)\|$ sogar immer

$$\|Df(x_0)\| = \|\nabla f(x_0)\| \quad (3.5)$$

gilt. Es ist also wirklich, dass der Gradient immer die Richtung des steilsten Anstiegs der Funktionswerte von f an der Stelle x_0 angibt.

Wem dieser Beweis nicht so zusagt, dem geben wir hier noch einen etwas anderen: Es gilt (wie oben) $\|Df(x_0)\| \leq \|\nabla f(x_0)\|$. Das bedeutet aber nur: Die Größe der „Steigung“ ist beschränkt durch die Länge des Gradienten. Es bleibt zu zeigen, dass der größte Wert auch in Richtung des Gradienten angenommen wird, aber dies zeigt gerade

$$D_v f(x_0) = \left\langle \nabla f(x_0), \frac{\nabla f(x_0)}{\|\nabla f(x_0)\|} \right\rangle = \|\nabla f(x_0)\|.$$

Oder anders formuliert: Das Skalarprodukt $\langle \nabla f(x_0), v \rangle$ wird maximal, wenn die beiden Vektoren $\nabla f(x_0)$ und v parallel sind. Damit zeigt der Gradient in Richtung des steilsten Anstiegs.

Zum Abschluss noch ein Beispiel zum Rechnen.

Beispiel 41

Seien a, b, c Konstanten. Wir betrachten die Funktion

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = ax^2 + bxy + cy^2.$$

Die partiellen Ableitungen sind

$$\frac{\partial f}{\partial x}(x, y) = 2ax + by \quad \text{und} \quad \frac{\partial f}{\partial y}(x, y) = 2cy + bx.$$

Der Gradient von f ist also gegeben durch

$$\nabla f(x, y) = \begin{pmatrix} 2ax + by \\ 2cy + bx \end{pmatrix}.$$

■

Zur Definition 3.5 der Jakobi-Matrix: Wie berechnet man die Jakobi-Matrix? Das ist ganz einfach: Man schreibt die partielle Ableitung (siehe Definition 3.3) einfach entsprechend in eine Matrix. Beispielsweise erhalten wir für eine Abbildung der Form $f : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ die Jakobi-Matrix als eine (3×2) -Matrix. Ist

$$f(x, y) := \begin{pmatrix} f_1(x, y) \\ f_2(x, y) \\ f_3(x, y) \end{pmatrix},$$

so gilt

$$J_f(x, y) = \begin{pmatrix} \partial_x f_1 & \partial_y f_1 \\ \partial_x f_2 & \partial_y f_2 \\ \partial_x f_3 & \partial_y f_3 \end{pmatrix}.$$

Für eine Abbildung der Form $f : \mathbb{R}^n \rightarrow \mathbb{R}$ stimmt der Gradient mit der Transponierten der Jakobi-Matrix überein. Schauen wir uns dies an Beispielen an. Zuvor noch eine kleine Anmerkung: Wir werden statt $J_f(x_1, \dots, x_n)$ auch häufig einfach nur $Df(x_1, \dots, x_n)$ schreiben, weil die Jakobi-Matrix dem Differential von f entspricht.

Beispiel 42

Wir betrachten die Abbildung

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^3, \quad f(x, y) = \begin{pmatrix} \cos(x) \cos(y) \\ \sin(x) \cos(y) \\ \sin(y) \end{pmatrix}.$$

Hier ist $f_1(x, y) = \cos(x) \sin(y)$, $f_2(x, y) = \sin(x) \cos(y)$ und $f_3(x, y) = \sin(y)$. Also erhalten wir beispielsweise

$$\frac{\partial f_1}{\partial x} = \frac{\partial}{\partial x}(\cos(x) \cos(y)) = -\sin(x) \cos(y).$$

Machen wir dies mit allen f_i , so ergibt sich für die Jakobi-Matrix eine (3×2) -Matrix. Die Größe der Jakobi-Matrix kann man also an der Dimension des Definitions- und Wertebereichs der Funktion sofort ablesen. Wir erhalten so

$$J_f(x, y) = \begin{pmatrix} -\sin(x) \cos(y) & -\cos(x) \sin(y) \\ \cos(x) \cos(y) & -\sin(x) \sin(y) \\ 0 & \cos(y) \end{pmatrix}.$$

In den Spalten der Jakobi-Matrix stehen die Richtungsableitungen von f in Richtung e_1 bzw. e_2 , das heißt,

$$\begin{aligned}\frac{\partial f}{\partial x}(x, y) &= D_1 f(x, y) = Df(x, y)(e_1) = \begin{pmatrix} -\sin(x) \cos(y) \\ \cos(x) \cos(y) \\ 0 \end{pmatrix}, \\ \frac{\partial f}{\partial y}(x, y) &= D_2 f(x, y) = Df(x, y)(e_2) = \begin{pmatrix} -\cos(x) \sin(y) \\ -\sin(x) \sin(y) \\ \cos(y) \end{pmatrix}.\end{aligned}$$

■

Zur Definition 3.6 und 3.7 der höheren Ableitungen: Diese Definitionen sind irgendwie klar. Dies sollte ja auch so sein, dass man eine Abbildung zwischen Banach-Räumen mehrmals ableiten kann. Betrachten wir ein paar Beispiele.

Beispiel 43

Die Abbildung

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = x^2 y e^y$$

ist überall differenzierbar, und $Df(x, y)$ ist durch die Jakobi-Matrix

$$\begin{aligned}J_f(x, y) &= (D_1 f(x, y), D_2 f(x, y)) \\ &= \left(\frac{\partial f}{\partial x}(x, y), \frac{\partial f}{\partial y}(x, y) \right) \\ &= (2xye^y, x^2(y+1)e^y)\end{aligned}$$

gegeben. Auch die Abbildung $J_f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ist überall differenzierbar und es gilt:

$$J_{J_f}(x, y) = \begin{pmatrix} \frac{\partial D_1 f}{\partial x}(x, y) & \frac{\partial D_1 f}{\partial y}(x, y) \\ \frac{\partial D_2 f}{\partial x}(x, y) & \frac{\partial D_2 f}{\partial y}(x, y) \end{pmatrix} = \begin{pmatrix} 2ye^y & 2x(y+1)e^y \\ 2x(y+1)e^y & x^2(y+2)e^y \end{pmatrix}.$$

Also ist die zweite Ableitung gegeben durch

$$D^2 f(x, y) = \begin{pmatrix} 2ye^y & 2x(y+1)e^y \\ 2x(y+1)e^y & x^2(y+2)e^y \end{pmatrix}.$$

Hier fällt auf, dass die Matrix symmetrisch ist. Dies ist kein Zufall. Wir werden in Satz 3.8 sehen, dass dies für zweimal (total) differenzierbare Funktionen immer gilt. ■

Im Folgenden werden wir auch häufig die zweite partielle Ableitung schreiben als

$$\frac{\partial D_2 f}{\partial x} = \frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) =: \frac{\partial^2 f}{\partial x \partial y}.$$

Ab und zu schreiben wir auch f_{xy} für $\frac{\partial^2 f}{\partial x \partial y}$. Analog f_{xx} , f_{yx} und f_{yy} . Geben wir noch ein paar weitere Beispiele.

Beispiel 44

- Sei zunächst einmal $z = f(x, y) = x^3 y + e^{xy^2}$. Wir berechnen die zweiten Ableitungen und erhalten entsprechend

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x} \left(\frac{\partial}{\partial x} (x^3 y + e^{xy^2}) \right) = \frac{\partial}{\partial x} (3yx^2 + y^2 e^{xy^2}) = 6yx + y^4 e^{xy^2},$$

$$\frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y} \left(\frac{\partial}{\partial y} (x^3 y + e^{xy^2}) \right) = \frac{\partial}{\partial y} (x^3 + 2xy e^{xy^2}) = 2x e^{xy^2} (1 + 2xy^2),$$

$$\begin{aligned} \frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial x} \left(\frac{\partial}{\partial y} (x^3 y + e^{xy^2}) \right) = \frac{\partial}{\partial x} (x^3 + 2xy e^{xy^2}) \\ &= 3x^2 + 2y e^{xy^2} (1 + y^2 x), \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 f}{\partial y \partial x} &= \frac{\partial}{\partial y} \left(\frac{\partial}{\partial x} (x^3 y + e^{xy^2}) \right) = \frac{\partial}{\partial y} (3yx^2 + y^2 e^{xy^2}) \\ &= 3x^2 + 2y e^{xy^2} (1 + y^2 x). \end{aligned}$$

Hier folgt die Symmetrie aus dem Lemma von Schwarz.

- Wir betrachten

$$f(x, y) := \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2}, & \text{für } (x, y) \neq (0, 0), \\ 0, & \text{für } (x, y) = (0, 0) \end{cases}$$

und berechnen $f_{xy}(0, 0)$ und $f_{yx}(0, 0)$. Da f auf beiden Achsen konstant ist, gilt $f_x(0, 0) = f_y(0, 0) = 0$. Partielles Differenzieren ergibt nun:

$$f_x(x, y) = \begin{cases} \frac{yx^4 - y^5 + 4x^2 y^3}{(x^2 + y^2)^2}, & \text{für } (x, y) \neq (0, 0), \\ 0, & \text{für } (x, y) = (0, 0) \end{cases}$$

und

$$f_y(x, y) = \begin{cases} \frac{x^5 - xy^4 - 4x^3 y^2}{(x^2 + y^2)^2}, & \text{für } (x, y) \neq (0, 0), \\ 0, & \text{für } (x, y) = (0, 0). \end{cases}$$

Wir erhalten damit

$$f_{xy}(0, 0) = \lim_{y \rightarrow 0} \frac{f_x(0, y) - f_x(0, 0)}{y} = \lim_{y \rightarrow 0} \frac{-y^5}{y^5} = -1$$

und

$$f_{yx}(0,0) = \lim_{x \rightarrow 0} \frac{f_y(x,0) - f_y(0,0)}{x} = \lim_{y \rightarrow 0} \frac{x^5}{x^5} = 1.$$

Also ist $f_{xy}(0,0) \neq f_{yx}(0,0)$. Die Funktion kann nach dem Lemma von Schwarz 3.8 nicht (total) differenzierbar sein.

- Die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = x^2 e^y$ ist überall differenzierbar, denn die partiellen Ableitungen existieren und sind sogar stetig. $Df(x, y)$ ist durch die (1×2) -Jakobi-Matrix

$$J_f(x, y) = (2xe^y, x^2 e^y)$$

gegeben. Da $Df : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ wieder überall differenzierbar ist, erhalten wir für die Jakobi-Matrix von Df die Matrix

$$J_{Df}(x, y) = \begin{pmatrix} \frac{\partial D_1 f}{\partial x}(x, y) & \frac{\partial D_1 f}{\partial y}(x, y) \\ \frac{\partial D_2 f}{\partial x}(x, y) & \frac{\partial D_2 f}{\partial y}(x, y) \end{pmatrix} = \begin{pmatrix} 2e^y & 2xe^y \\ 2xe^y & x^2 e^y \end{pmatrix}.$$

Wir hoffen, dass das schöne Lemma von Schwarz nun klar geworden ist. Weitere Beispiele in der entsprechenden Erklärung. ■

3.4 Erklärungen zu den Sätzen und Beweisen

Zum Satz 3.1: Dieser Satz sagt, dass die schon aus Analysis 1 bekannte Summenregel für zwei differenzierbare Funktionen gilt. Unsere definierte Ableitung (Definition 3.1) besitzt also vernünftige Eigenschaften. Der Beweis ist klar (hoffen wir), aber versucht die Linearität einmal aus der Definition 3.1 direkt herzuleiten.

Zum Satz 3.3: Die Aussage dieses Satzes kennen wir auch schon aus der Analysis 1. Auch dies zeigt wieder, dass unsere Definition 3.1 der Differenzierbarkeit durchaus sinnvoll war, da diese wirklich das Konzept der Ableitung aus der Analysis 1 verallgemeinert, denn eine differenzierbare Abbildung sollte auch stetig sein.

Zum Satz 3.4 der Kettenregel: Für unsere so definierte Ableitung gilt auch eine Kettenregel, so wie wir sie aus Analysis 1 kennen. Gleichung (3.1) wirkt am Anfang erst einmal kompliziert. Daher schauen wir uns ein paar Beispiele an, wie wir die Ableitung der Verkettung von Funktionen berechnen.

Beispiel 45

- Wir betrachten die Abbildungen

$$f : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2, f(x) = \frac{x}{||x||^2}$$

und

$$g : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad g(y) := \|y\|.$$

f ist auf $\mathbb{R}^2 \setminus \{0\}$ differenzierbar, und die Jakobi-Matrix in kartesischen Koordinaten (siehe Definition 3.5) ist die Matrix

$$J_f(x) = \frac{1}{\|x\|^4} \begin{pmatrix} (x_2)^2 - (x_1)^2 & -2x_1x_2 \\ -2x_1x_2 & (x_1)^2 - (x_2)^2 \end{pmatrix}.$$

g ist auf $\mathbb{R}^2 \setminus \{0\}$ differenzierbar und $Dg(y)$ wird in kartesischen Koordinaten durch die Multiplikation von links mit der Matrix

$$J_g(y) = \frac{1}{\|y\|} (y_1, y_2)$$

dargestellt. Folglich ist

$$g \circ f : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}, \quad (g \circ f)(x) = \frac{1}{\|x\|}$$

differenzierbar und die Matrixdarstellung von $D(g \circ f)(x)$ lautet

$$\begin{aligned} J_{(g \circ f)}(x) &= J_g(f(x)) \cdot J_f(x) \\ &= \frac{1}{\left\| \frac{x}{\|x\|^2} \right\|} \begin{pmatrix} x_1 & x_2 \\ \|x\|^2 & \|x\|^2 \end{pmatrix} \cdot \frac{1}{\|x\|^4} \begin{pmatrix} (x_2)^2 - (x_1)^2 & -2x_1x_2 \\ -2x_1x_2 & (x_1)^2 - (x_2)^2 \end{pmatrix} \\ &= \frac{1}{\|x\|^5} (x_1, x_2) \cdot \begin{pmatrix} (x_2)^2 - (x_1)^2 & -2x_1x_2 \\ -2x_1x_2 & (x_1)^2 - (x_2)^2 \end{pmatrix} \\ &= \frac{1}{\|x\|^5} \left(x_1(x_2)^2 - (x_1)^3 - 2x_1(x_2)^2, -2(x_1)^2x_2 + x_2(x_1)^2 - (x_2)^3 \right) \\ &= -\frac{1}{\|x\|^3} (x_1, x_2). \end{aligned}$$

- Seien $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(x, y) := e^{xy^2}$ und $g : \mathbb{R} \rightarrow \mathbb{R}^2$, $g(t) = (\cos t, \sin t)$. Zunächst berechnen wir die Ableitung von $f \circ g$ durch Differentiation. Es gilt $h(t) := (f \circ g)(t) = f(g(t)) = e^{\cos t \cdot \sin^2 t}$ und damit, dies können wir schon, das ist Analysis 1:

$$\begin{aligned} Dh(t) &= e^{\cos t \cdot \sin^2 t} \cdot \left(-\sin t \cdot \sin^2 t + \cos t \cdot 2 \sin t \cdot \cos t \right) \\ &= e^{\cos t \cdot \sin^2 t} \cdot \left(-\sin^3 t + 2 \sin t \cdot \cos^2 t \right). \end{aligned}$$

Ob unsere Kettenregel wirklich Sinn macht, überprüfen wir jetzt, indem wir auf die Verkettung noch einmal Gleichung (3.1) anwenden. Wir müssen

$$J_{(f \circ g)}(x) = J_f(g(x)) \cdot J_g(x)$$

berechnen. Dazu sind ein paar Jakobi-Matrizen von Nöten, die wir jetzt angeben. Die Jakobi-Matrix enthält als Einträge jeweils die partiellen Ableitungen. Demnach gilt:

$$J_f(x) = \left(y^2 e^{xy^2}, 2xy e^{xy^2} \right)$$

und

$$J_g(x) = (-\sin x, \cos x)^T.$$

Zu guter Letzt gilt:

$$J_f(g(x)) = \left(\sin^2(x) e^{\cos(x) \cdot \sin^2(x)}, 2 \sin(x) \cos(x) \cdot e^{\cos(x) \cdot \sin^2(x)} \right).$$

Nun können wir $J_{(f \circ g)}(x) = J_f(g(x)) \cdot J_g(x)$ berechnen:

$$\begin{aligned} J_{(f \circ g)}(x) &= J_f(g(x)) \cdot J_g(x) \\ &= \left(\sin^2(x) e^{\cos(x) \cdot \sin^2(x)}, 2 \sin(x) \cos(x) \cdot e^{\cos(x) \cdot \sin^2(x)} \right) \cdot \begin{pmatrix} -\sin(x) \\ \cos(x) \end{pmatrix} \\ &= -\sin(x) \cdot \sin^2(x) \cdot e^{\cos(x) \cdot \sin^2(x)} + 2 \sin(x) \cos(x) e^{\cos(x) \cdot \sin^2(x)} \cos(x) \\ &= e^{\cos(x) \cdot \sin^2(x)} \cdot \left(-\sin^3(x) + 2 \sin(x) \cdot \cos^2(x) \right). \end{aligned}$$

Auf beiden Wegen erhalten wir also dasselbe. So sollte es auch sein :-)

■ Seien $f, g : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}^2$ gegeben durch

$$f(x, y) := (x^2 - y^2, 2xy) \quad \text{und} \quad g(x, y) := \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right).$$

Wir wollen $J_{f \circ g}(1, 0)$ berechnen. Dazu berechnen wir zunächst $J_{(f \circ g)}(x) = J_f(g(x)) \cdot J_g(x)$. Wir benötigen also verschiedene Jakobi-Matrizen, die wir zunächst berechnen wollen. Es gilt:

$$J_f(x) = \begin{pmatrix} 2x & -2y \\ 2y & 2x \end{pmatrix}$$

und

$$\begin{aligned} J_g(x) &= \begin{pmatrix} \frac{1 \cdot (x^2 + y^2) - 2x^2}{(x^2 + y^2)^2} & -\frac{2xy}{(x^2 + y^2)^2} \\ \frac{2xy}{(x^2 + y^2)^2} & \frac{-1 \cdot (x^2 + y^2) + 2y^2}{(x^2 + y^2)^2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{x^2 + y^2 - 2x^2}{(x^2 + y^2)^2} & -\frac{2xy}{(x^2 + y^2)^2} \\ \frac{2xy}{(x^2 + y^2)^2} & \frac{-x^2 - y^2 + 2y^2}{(x^2 + y^2)^2} \end{pmatrix} = \begin{pmatrix} \frac{-x^2 + y^2}{(x^2 + y^2)^2} & -\frac{2xy}{(x^2 + y^2)^2} \\ \frac{2xy}{(x^2 + y^2)^2} & \frac{-x^2 + y^2}{(x^2 + y^2)^2} \end{pmatrix}. \end{aligned}$$

Für $J_f(g(x))$ erhalten wir nun durch Einsetzen

$$J_f(g(x)) = \begin{pmatrix} 2 \frac{x}{x^2 + y^2} & 2 \frac{y}{x^2 + y^2} \\ -2 \frac{y}{x^2 + y^2} & 2 \frac{x}{x^2 + y^2} \end{pmatrix} = 2 \begin{pmatrix} \frac{x}{x^2 + y^2} & \frac{y}{x^2 + y^2} \\ -\frac{y}{x^2 + y^2} & \frac{x}{x^2 + y^2} \end{pmatrix}.$$

Aus der Kettenregel folgt nun

$$\begin{aligned} J_{(f \circ g)}(x) &= J_f(g(x)) \cdot J_g(x) \\ &= \begin{pmatrix} 2 \frac{x}{x^2+y^2} & 2 \frac{y}{x^2+y^2} \\ -2 \frac{y}{x^2+y^2} & 2 \frac{x}{x^2+y^2} \end{pmatrix} \cdot \begin{pmatrix} \frac{-x^2+y^2}{(x^2+y^2)^2} & -\frac{2xy}{(x^2+y^2)^2} \\ \frac{2xy}{(x^2+y^2)^2} & \frac{-x^2+y^2}{(x^2+y^2)^2} \end{pmatrix} \\ &= \begin{pmatrix} \frac{6xy^2-2x^3}{(x^2+y^2)^3} & \frac{6x^2y+2y^3}{(x^2+y^2)^3} \\ \frac{6x^2y-2y^3}{(x^2+y^2)^3} & \frac{6xy^2-2x^3}{(x^2+y^2)^3} \end{pmatrix}. \end{aligned}$$

Nun gilt also:

$$J_{(f \circ g)}(1, 0) = \begin{pmatrix} \frac{6 \cdot 1 \cdot 0^2 - 2 \cdot 1^3}{(1^2 + 0^2)^3} & \frac{6 \cdot 1^2 \cdot 0 + 2 \cdot 0^3}{(1^2 + 0^2)^3} \\ \frac{6 \cdot 1^2 \cdot 0 - 2 \cdot 0^3}{(1^2 + 0^2)^3} & \frac{6 \cdot 1 \cdot 0^2 - 2 \cdot 1^3}{(1^2 + 0^2)^3} \end{pmatrix} = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.$$

■ Seien $f : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^2$ gegeben durch

$$(x, y) \mapsto \left(\frac{x}{x^2 + y^2}, \frac{y}{x^2 + y^2} \right)$$

und $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch

$$(u, v) \mapsto \sqrt{u^2 + v^2}.$$

Wir berechnen zunächst $D(g \circ f)(x_0)$ auf „herkömmlichem“ Weg. Es gilt:

$$(g \circ f)(x_0) = g(f(x_0)) = \sqrt{\frac{x^2}{(x^2 + y^2)^2} + \frac{y^2}{(x^2 + y^2)^2}} = \frac{1}{\sqrt{x^2 + y^2}}.$$

Damit ist also

$$D(g \circ f)(x_0) = \left(\frac{-x}{\sqrt{(x^2 + y^2)^3}}, \frac{-y}{\sqrt{(x^2 + y^2)^3}} \right).$$

Nun wollen wir $Dg(f(x_0)) \circ Df(x_0)$ berechnen und schauen, ob dasselbe Ergebnis herauskommt. Zunächst gilt:

$$\begin{aligned} Df(x) &= Df(x, y) = \begin{pmatrix} \frac{-x^2+y^2}{(x^2+y^2)^2} & \frac{-2xy}{(x^2+y^2)^2} \\ \frac{-2xy}{(x^2+y^2)^2} & \frac{x^2-y^2}{(x^2+y^2)^2} \end{pmatrix} \\ &= \frac{1}{(x^2 + y^2)^2} \begin{pmatrix} -x^2 + y^2 & -2xy \\ -2xy & x^2 - y^2 \end{pmatrix}. \end{aligned}$$

$Dg(f(x_0))$ berechnen wir zu

$$\begin{aligned} Dg(f(x)) &= \left(\frac{\frac{x}{x^2+y^2}}{\sqrt{\frac{x^2}{(x^2+y^2)^2} + \frac{y^2}{(x^2+y^2)^2}}}, \frac{\frac{y}{x^2+y^2}}{\sqrt{\frac{x^2}{(x^2+y^2)^2} + \frac{y^2}{(x^2+y^2)^2}}} \right) \\ &= \left(\frac{x\sqrt{x^2+y^2}}{x^2+y^2}, \frac{y\sqrt{x^2+y^2}}{x^2+y^2} \right) = \left(\frac{x}{\sqrt{(x^2+y^2)^3}}, \frac{y}{\sqrt{(x^2+y^2)^3}} \right). \end{aligned}$$

Nun wollen wir mittels Matrizenmultiplikation endlich $Dg(f(x_0)) \cdot Df(x_0)$ berechnen.

$$\begin{aligned}
 & Dg(f(x_0)) \cdot Df(x_0) \\
 &= \frac{1}{(x^2 + y^2)^2} \left(\frac{x}{\sqrt{x^2 + y^2}}, \frac{y}{\sqrt{x^2 + y^2}} \right) \cdot \begin{pmatrix} -x^2 + y^2 & -2xy \\ -2xy & x^2 - y^2 \end{pmatrix} \\
 &= \frac{1}{(x^2 + y^2)^2} \frac{1}{\sqrt{x^2 + y^2}} (x, y) \cdot \begin{pmatrix} -x^2 + y^2 & -2xy \\ -2xy & x^2 - y^2 \end{pmatrix} \\
 &= \frac{1}{\sqrt{x^2 + y^2}^5} (x(y^2 - x^2) - 2xy^2, -2x^2y + y(x^2 - y^2)) \\
 &= \frac{1}{\sqrt{x^2 + y^2}^5} (-x(y^2 + x^2), -y(y^2 + x^2)) \\
 &= \left(\frac{-x}{\sqrt{x^2 + y^2}^3}, \frac{-y}{\sqrt{x^2 + y^2}^3} \right).
 \end{aligned}$$

Das passt also :-).

Wir hoffen, dass die Beispiele ausführlich genug waren und ihr jetzt die Kettenregel anwenden könnt. ■

Zum Satz 3.5 von der Mittelwertungleichung: Die Mittelwertungleichung ist das Analogon zum Mittelwertsatz aus der Analysis 1.

Zum Satz 3.6: Der Satz sagt gerade, dass die Richtungsableitung mit der normalen Ableitung angewendet auf den Richtungsvektor, die wir in Definition 3.1 haben, übereinstimmt.

Beispiel 46

Für die Abbildung $f(v) = a \times v$ aus Beispiel 38 hatten wir

$$D_v(f(x_0)) = a \times v = f(v)$$

berechnet. Andererseits ist $f : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ linear und dann gilt, dass $Df(x_0) = f$. Wir sehen also an dem konkreten Beispiel, dass der Satz wirklich gilt. ■

Zum Satz 3.7: Es reicht nicht aus zu zeigen, dass eine Funktion partiell differenzierbar ist, um auf die (totale) Differenzierbarkeit zu schließen. Auch wenn alle partiellen Ableitungen existieren, heißt es nicht, dass die Funktion in dem von uns definiertem Sinne differenzierbar ist! Wir haben bei Funktionen mehrerer Veränderlicher also zwei Ableitungsbegriffe. Die (totale) Differenzierbarkeit ist wesentlich stärker als die der partiellen Differenzierbarkeit. Totale Differenzierbarkeit impliziert partielle Differenzierbarkeit, aber nicht umgekehrt, wie folgendes Beispiel zeigt.

Beispiel 47

Wir betrachten die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := \sqrt{|xy|}.$$

Die Funktion ist an der Stelle $(0, 0)$ partiell differenzierbar, denn die partiellen Ableitungen existieren, wie wir leicht nachrechnen können:

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{t \rightarrow 0} \frac{f(t, 0) - f(0, 0)}{t} = 0 \quad \text{und} \quad \frac{\partial f}{\partial y}(0, 0) = 0.$$

f ist aber an der Stelle $(0, 0)$ nicht differenzierbar, denn sonst müsste nach Satz 3.6 $Df(0, 0) = 0$ und $D_v f(0, 0) = 0$ für alle $v \in \mathbb{R}^2$ gelten. Allerdings ist für $t \neq 0$

$$\frac{f(t, t) - f(0, 0)}{t} = \frac{|t|}{t} = \operatorname{sgn}(t),$$

sodass $D_v f(0, 0)$ für $v = e_1 + e_2$ gar nicht eindeutig ist, denn es gilt:

$$D_{v=e_1+e_2} f(0, 0) = \lim_{t \rightarrow 0} \frac{f(t, t) - f(0, 0)}{t} = \operatorname{sign}(t).$$

Dies ist nun wirklich nicht eindeutig, weil es darauf ankommt, ob man für $t < 0$ oder $t > 0$ gegen Null läuft. Hierbei ist $\operatorname{sgn}(t)$ übrigens 1, wenn $t > 0$ und -1 , wenn $t < 0$. ■

Zum Satz 3.8 (Lemma von Schwarz): In Beispiel 43 haben wir gesehen, dass

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}.$$

Dies gilt im Allgemeinen nicht, sondern nur, wenn die Funktion zweimal (total) differenzierbar ist, aber beispielsweise nicht, wenn die Funktion nur zweimal partiell differenzierbar ist, wie das folgende Beispiel zeigt.

Beispiel 48

Wir betrachten die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) := \begin{cases} xy \frac{x^2 - y^2}{x^2 + y^2}, & \text{für } (x, y) \neq (0, 0), \\ 0, & \text{für } (x, y) = (0, 0). \end{cases}$$

In Beispiel 42 hatten wir gezeigt, dass die Funktion überall stetig ist. Zudem gilt für die partiellen Ableitungen für $(x, y) \neq (0, 0)$, dass

$$\frac{\partial f}{\partial x}(x, y) = y \frac{x^2 - y^2}{x^2 + y^2} + 4 \frac{x^2 y^3}{(x^2 + y^2)^2}, \quad \frac{\partial f}{\partial y}(x, y) = x \frac{x^2 - y^2}{x^2 + y^2} - 4 \frac{x^3 y^2}{(x^2 + y^2)^2}.$$

Demnach ist f auf $\mathbb{R}^2 \setminus \{0\}$ stetig differenzierbar. Man überzeugt sich, dass $Df(x, y)$ sogar im Ursprung stetig fortsetzbar ist. Es ergibt sich, dass Df überall partiell differenzierbar ist. Wir erhalten

$$\frac{\partial Df}{\partial x}(0, 0) = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \frac{\partial Df}{\partial y}(0, 0) = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

Somit ist

$$\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1 \neq -1 = \frac{\partial^2 f}{\partial y \partial x}(0, 0).$$

Die fehlende Symmetrie liegt daran, dass Df im Ursprung zwar partiell differenzierbar und somit f dort zweimal partiell differenzierbar, aber nicht (total) differenzierbar ist. ■

Das Lemma 3.8 von Schwarz liefert also auch ein Kriterium, um zu entscheiden, ob eine Funktion nicht (total) differenzierbar ist. Das ist nämlich genau dann der Fall, wenn die zweiten Ableitungen nicht symmetrisch sind, also wenn

$$\frac{\partial^2 f}{\partial x \partial y} \neq \frac{\partial^2 f}{\partial y \partial x}.$$

Zum Satz 3.9: Dieser Satz ist sehr nützlich, denn der Nachweis, dass eine Funktion total differenzierbar ist, kann durchaus schwer sein. Aber zu zeigen, dass sie partiell differenzierbar ist, dass also die partiellen Ableitungen existieren und dass diese stetig sind, ist meistens leichter. Vergleiche dazu das Beispiel 44.

Zum Satz 3.10 (Taylor-Formel): Dieser Satz ist wieder das Analogon zur Taylor-Formel für eine Funktion einer Veränderlichen. Am besten versteht man diese anhand von ein paar Beispielen, die wir jetzt geben wollen.

Beispiel 49

- Wir ermitteln das Taylor-Polynom erster Ordnung der Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := \cos(x) \cos(y)$$

im Punkt $(x_0, y_0) = (\pi/4, \pi/4)$. Die Taylor-Formel lautet in diesem Fall

$$T(x, y) = f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0).$$

Für die partiellen Ableitungen erhalten wir

$$f_x(x, y) = -\cos(y) \sin(x) \quad \text{und} \quad f_y(x, y) = -\cos(x) \sin(y).$$

Ein Einsetzen in die Taylor-Formel und Ausrechnen liefert nun das gesuchte Taylor-Polynom erster Ordnung:

$$\begin{aligned} T(x, y) &= f\left(\frac{\pi}{4}, \frac{\pi}{4}\right) + f_x\left(\frac{\pi}{4}, \frac{\pi}{4}\right)\left(x - \frac{\pi}{4}\right) + f_y\left(\frac{\pi}{4}, \frac{\pi}{4}\right)\left(y - \frac{\pi}{4}\right) \\ &= \frac{1}{2} + \left(-\sin \frac{\pi}{4} \cdot \cos \frac{\pi}{4}\right)\left(x - \frac{\pi}{4}\right) + \left(-\sin \frac{\pi}{4} \cdot \cos \frac{\pi}{4}\right)\left(y - \frac{\pi}{4}\right) \\ &= \frac{1}{2} - \frac{1}{2}\left(x - \frac{\pi}{4}\right) - \frac{1}{2}\left(y - \frac{\pi}{4}\right) \\ &= -\frac{1}{2}x - \frac{1}{2}y + \frac{1}{2} + \frac{\pi}{4}. \end{aligned}$$

- Zu bestimmen ist das Taylor-Polynom $T_2(x, y)$ im Punkt $(x_0, y_0) = (1, 1)$ zweiter Ordnung der Funktion

$$g: \mathbb{R}_{>0}^2 \rightarrow \mathbb{R}, \quad g(x, y) = y^x.$$

Jetzt ist die Formel für das Taylor-Polynom etwas komplizierter und lautet

$$\begin{aligned} T_2(x, y) &= f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0) \quad (3.6) \\ &+ \frac{1}{2!}(f_{xx}(x_0, y_0)(x - x_0)^2 + 2f_{xy}(x_0, y_0)(y - y_0)(x - x_0) \\ &+ f_{yy}(x_0, y_0)(y - y_0)^2). \end{aligned}$$

Wir benötigen also die partiellen Ableitungen. Um diese zu berechnen, schreibt man am besten erst einmal $y^x = e^{\ln y^x} = e^{x \cdot \ln y}$. So errechnet man

$$\begin{aligned} f(x, y) &= y^x \Rightarrow f(1, 1) = 1 \\ f_x(x, y) &= \ln y \cdot y^x \Rightarrow f_x(1, 1) = 0 \\ f_y(x, y) &= xy^{x-1} \Rightarrow f_y(1, 1) = 1 \\ f_{xx}(x, y) &= \ln^2 y \cdot y^x \Rightarrow f_{xx}(1, 1) = 0 \\ f_{yy}(x, y) &= x(x-1)y^{x-2} \Rightarrow f_{yy}(1, 1) = 0 \\ f_{xy}(x, y) &= f_{yx}(x, y) = y^{x-1}(1 + x \cdot \ln y) \Rightarrow f_{xy}(1, 1) = f_{yx}(1, 1) = 1. \end{aligned}$$

Einsetzen in (3.6) liefert nun

$$T_2(x, y) = xy - x + 1.$$

- Als Übungsaufgabe solltet ihr noch einmal das Taylor-Polynom $T_2(x, y)$ von $f(x, y) = \cos(xy) + xe^{y-1}$ im Punkt $(x_0, y_0) = (\pi, 1)$ berechnen. Wenn ihr das richtig macht, müsste Folgendes rauskommen:

$$T_2(x, y) = -1 + \frac{\pi}{2} + 2\pi^2 - 2\pi x - (\pi + 2\pi^2)y + \frac{1}{2}x^2 + (1 + \pi)xy + \frac{1}{2}(\pi + \pi^2)y^2.$$

Es ist etwas Rechnerei, aber wie sagt man so schön: Nur die Übung macht den Meister (Ja, wir wissen, „doofer Spruch“, aber es ist was dran!) :-).

■

Zusammenfassung: Fassen wir noch einmal zusammen, wie man eine Funktion auf Differenzierbarkeit im Punkt x_0 untersucht. Zunächst einmal können wir festhalten (nach Satz 3.9), dass eine Funktion (total) differenzierbar ist, wenn die partiellen Ableitungen überall stetig sind.

Gegeben sei eine Abbildung $f : \Omega \rightarrow W$ mit V und W als Banach-Räume und $\Omega \subset V$ offen. Zunächst sollte man überprüfen, ob f stetig in x_0 ist oder eben zeigen, dass

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0) - A(x - x_0)\|}{\|x - x_0\|} = 0.$$

wobei A die Jakobi-Matrix bezeichnet. Ist f noch nicht einmal stetig, dann kann f auch nicht differenzierbar sein. Wenn f stetig ist, muss man weiter untersuchen. Und zwar sollte man dann erst einmal schauen, ob f in x_0 partiell differenzierbar ist. Wenn dies nicht der Fall ist, so ist f nicht total differenzierbar. Sollte f partiell differenzierbar sein, sollte man überprüfen, ob die partiellen Ableitungen stetig sind. Ist dies der Fall, so folgt ebenfalls die Differenzierbarkeit im Punkt x_0 . Ist dies nicht der Fall, heißt das nicht, dass f nicht total differenzierbar sein kann. Jetzt müssen wir auf die Definition 3.1 zurückgreifen, um die Differenzierbarkeit zu zeigen. Gelingt dies auch nicht, so können wir sagen, dass f nicht in x_0 differenzierbar ist.

Beispiel 50

Zwei abschließende Beispiele wollen wir uns nun anschauen.

■ Sei

$$f(x, y) := \begin{cases} \frac{x^2 y^2}{x^2 + y^2}, & \text{für } (x, y) \neq (0, 0), \\ 0, & \text{für } (x, y) = (0, 0). \end{cases}$$

Es ist klar, dass f für alle $(x, y) \neq (0, 0)$ differenzierbar ist, da f aus differenzierbaren Funktionen zusammengesetzt ist. Um die Differenzierbarkeit von f in $(0, 0)$ zu untersuchen, verwenden wir Polarkoordinaten. Mit $x = r \cos \varphi$ und $y = r \sin \varphi$ ergibt sich

$$f(r \cos \varphi, r \sin \varphi) = \frac{r^4 \cos^2 \varphi \sin^2 \varphi}{r^2} = r^2 \cos^2 \varphi \sin^2 \varphi.$$

- f ist in $(0, 0)$ stetig, denn es gilt:

$$\lim_{(x, y) \rightarrow (0, 0)} f(x, y) = \lim_{r \rightarrow 0} r^2 \cos^2 \varphi \sin^2 \varphi = 0$$

und

$$f(0, 0) = 0.$$

- f ist in $(0, 0)$ differenzierbar, denn es gilt:

$$\text{grad} f(0, 0) = (0, 0),$$

da

$$\begin{aligned} f_x(0, 0) &= \lim_{(x, 0) \rightarrow (0, 0)} \frac{f(x, 0) - f(0, 0)}{x} = \lim_{x \rightarrow 0} \frac{0 - 0}{x} = 0, \\ f_y(0, 0) &= \lim_{(0, y) \rightarrow (0, 0)} \frac{f(0, y) - f(0, 0)}{y} = \lim_{y \rightarrow 0} \frac{0 - 0}{y} = 0. \end{aligned}$$

- f ist in $(0, 0)$ differenzierbar und es gilt:

$$f'(0, 0) = \nabla f(0, 0) = (0, 0).$$

Hierzu zeigen wir entweder die Stetigkeit der partiellen Ableitungen oder gehen auf die Definition der Differenzierbarkeit zurück.

$$\begin{aligned} \lim_{(x, y) \rightarrow (0, 0)} \frac{\frac{x^2 y^2}{x^2 + y^2} - 0 - 0x - 0y}{|(x, y)|} &= \lim_{r \rightarrow 0} \frac{r^4 \cos^2 \varphi \sin^2 \varphi}{r^3} \\ &= \lim_{r \rightarrow 0} r \cos^2 \varphi \sin^2 \varphi = 0. \end{aligned}$$

■ Sei

$$g(x, y) := \begin{cases} \frac{x^2 y}{x^2 + y^2}, & \text{für } (x, y) \neq (0, 0), \\ 0, & \text{für } (x, y) = (0, 0). \end{cases}$$

Auch g ist in den Punkten $(x, y) \neq (0, 0)$ differenzierbar, da g aus differenzierbaren Funktionen zusammengesetzt ist. Um die Differenzierbarkeit von g in $(0, 0)$ nachzuweisen, verwenden wir Polarkoordinaten. Wir erhalten so zunächst

$$g(r \cos \varphi, r \sin \varphi) = \frac{r^3 \cos^2 \varphi \sin \varphi}{r^2} = r \cos^2 \varphi \sin \varphi.$$

- g ist in $(0, 0)$ stetig, denn es gilt:

$$\lim_{(x, y) \rightarrow (0, 0)} g(x, y) = \lim_{r \rightarrow 0} r \cos^2 \varphi \sin \varphi = 0$$

und

$$g(0, 0) = 0.$$

- g ist in $(0, 0)$ partiell differenzierbar, denn es ist

$$\text{grad} g(0, 0) = (0, 0),$$

da

$$g_x(0, 0) = \lim_{(x, 0) \rightarrow (0, 0)} \frac{f(x, 0) - f(0, 0)}{x} = \lim_{x \rightarrow 0} \frac{0 - 0}{x} = 0.$$

Analog erhalten wir $f_y(0, 0) = 0$.

- Sind die partiellen Ableitungen von g in $(0,0)$ stetig? Nein, denn es ist

$$g_x(x, y) := \begin{cases} \frac{2xy^3}{(x^2 + y^2)^2}, & \text{für } (x, y) \neq (0, 0), \\ 0, & \text{für } (x, y) = (0, 0), \end{cases}$$

und der Grenzwert

$$\lim_{(x,y) \rightarrow (0,0)} g_x(x, y) = \lim_{r \rightarrow 0} \frac{2r^4 \cos \varphi \sin^3 \varphi}{r^4} = \lim_{r \rightarrow 0} 2 \cos \varphi \sin^3 \varphi$$

existiert, ist aber von φ abhängig und daher ist g_x in $(0,0)$ nicht stetig. Das heißt, wir müssen auf die Definition der Differenzierbarkeit zurückgehen, um zu zeigen, dass g in $(0,0)$ differenzierbar ist. Wir müssen also die Frage beantworten, ob

$$\lim_{(x,y) \rightarrow (0,0)} \frac{f(x, y) - f(0, 0) - f_x(0, 0)x - f_y(0, 0)y}{|(x, y)|} = 0$$

ist. Dies ist nicht der Fall, denn der Grenzwert

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\frac{x^2 y}{x^2 + y^2} - f(0, 0) - 0x - 0y}{|(x, y)|}$$

existiert nicht, da sich in Polarkoordinaten

$$\begin{aligned} \lim_{(x,y) \rightarrow (0,0)} \frac{\frac{x^2 y}{x^2 + y^2} - f(0, 0) - 0x - 0y}{|(x, y)|} &= \lim_{r \rightarrow 0} \frac{\frac{r^3 \cos \varphi \sin \varphi}{r^2}}{r} \\ &= \lim_{r \rightarrow 0} \cos^2 \varphi \sin \varphi \end{aligned}$$

ergibt. Der letzte Grenzwert existiert zwar, ist aber von φ abhängig. Daher ist g in $(0,0)$ nicht differenzierbar. ■

4 Extremwertberechnungen

Übersicht

4.1	Definitionen	93
4.2	Sätze und Beweise	94
4.3	Erklärungen zu den Definitionen	95
4.4	Erklärungen zu den Sätzen und Beweisen	96

Wen hätte es gewundert? So wie uns in Analysis 1 oder vielleicht viel mehr in der Schule Extremwerte von Abbildungen der Form $f : \mathbb{R} \rightarrow \mathbb{R}$ interessierten, wollen wir uns in diesem Kapitel anschauen, wie man Extremwerte von Abbildungen der Form $f : \mathbb{R}^n \rightarrow \mathbb{R}$ berechnen kann.

Im Folgenden seien $\Omega \subset \mathbb{R}^n$ stets offen und $f : \Omega \rightarrow \mathbb{R}$ differenzierbar. Weiterhin sei (e_1, \dots, e_n) die Standardbasis des \mathbb{R}^n .

4.1 Definitionen

Definition 4.1 (Hesse-Matrix)

Sei $f : \Omega \rightarrow \mathbb{R}^n$ zweimal partiell differenzierbar in x_0 , so heißt

$$D^2 f(x_0) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j=1,\dots,n}$$

die **Hesse-Matrix** von f an der Stelle x_0 .

Definition 4.2 (lokales Extremum)

Seien $\Omega \subset \mathbb{R}^n$ offen und $f : \overline{\Omega} \rightarrow \mathbb{R}$ eine Funktion. Wir sagen f besitzt an der Stelle $x_0 \in \overline{\Omega}$ ein **lokales Maximum**, wenn es ein $\varepsilon > 0$ gibt, sodass

$$f(x) \leq f(x_0) \quad \forall x \in \overline{\Omega} \cap U(x_0, \varepsilon).$$

Gilt sogar

$$f(x) < f(x_0) \quad \forall x \in \overline{\Omega} \cap U(x_0, \varepsilon), x \neq x_0$$

für genügend kleines $\varepsilon > 0$, so nennen wir das lokale Maximum **isoliert**. Ist $x_0 \in \Omega$, so spricht man von einem inneren (isolierten) lokalen Maximum. Analog sind lokale Minima definiert. Die Menge aller lokalen Minima und Maxima bildet die Menge aller Extremstellen von f . Die zugehörigen Funktionswerte nennen wir **Extremwerte**.

Anmerkung: In dieser Definition benötigen wir den Abschluss von Ω , also $\overline{\Omega}$, weil wir auch an Randpunkten interessiert sind.

Definition 4.3 (Definitheit)

Eine symmetrische $(n \times n)$ -Matrix A heißt **positiv definit** (bzw. **positiv semidefinit**), wenn

$$\langle Av, v \rangle > 0 \text{ (bzw. } \geq 0) \quad \forall v \neq 0.$$

Entsprechend heißt A **negativ (semi-)definit**, wenn

$$\langle Av, v \rangle < 0 \text{ (bzw. } \leq 0) \quad \forall v \neq 0.$$

A heißt **indefinit**, wenn A weder positiv noch negativ (semi-)definit ist.

4.2 Sätze und Beweise

Satz 4.1 (notwendiges Kriterium für das Vorhandensein von Extremstellen)

$f : \Omega \rightarrow \mathbb{R}$ sei in x_0 differenzierbar und besitze dort ein inneres Extremum. Dann gilt

$$\nabla f(x_0) = 0.$$

Satz 4.2 (hinreichendes Kriterium)

Seien $\Omega \subset \mathbb{R}^n$ offen, $f : \Omega \rightarrow \mathbb{R}$ zweimal stetig differenzierbar und $x_0 \in \Omega$ ein beliebiger Punkt in Ω . Dann gilt:

- i) Ist $\nabla f(x_0) = 0$ und die Hesse-Matrix $D^2 f(x_0)$ positiv definit, so besitzt f an der Stelle x_0 ein isoliertes Minimum.
- ii) Besitzt umgekehrt f an der Stelle x_0 ein lokales Minimum, so gilt $\nabla f(x_0) = 0$ und die Hesse-Matrix $D^2 f(x_0)$ ist positiv semidefinit.
- iii) Ist $\nabla f(x_0) = 0$ und die Hesse-Matrix $D^2 f(x_0)$ negativ definit, so besitzt f an der Stelle x_0 ein isoliertes Maximum.
- iv) Besitzt umgekehrt f an der Stelle x_0 ein lokales Maximum, so gilt $\nabla f(x_0) = 0$ und die Hesse-Matrix $D^2 f(x_0)$ ist negativ semidefinit.

v) Ist die Matrix indefinit, so liegt ein Sattelpunkt vor.

Satz 4.3 (Extrema unter Nebenbedingungen)

Seien $U \subset \mathbb{R}^n$ offen und sowohl $f : U \rightarrow \mathbb{R}$, als auch $F : U \rightarrow \mathbb{R}$ stetig differenzierbar. Weiterhin sei $\text{grad}F(x) \neq 0$ für alle $x \in U$. Ist x_0 ein lokaler Extremwert von f aus

$$B := \{x : F(x) = 0\},$$

so existiert ein $\lambda \in \mathbb{R}$ mit

$$\text{grad}f(x_0) = \lambda \cdot \text{grad}F(x_0)$$

Anmerkung: Das λ wird der zugehörige **Lagrange-Multiplikator** genannt. Die Funktion $L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$ (hier nur im 2-Dimensionalen dargestellt) heißt **Lagrange-Funktion**.

4.3 Erklärungen zu den Definitionen

Zur Definition 4.1 der Hesse-Matrix: In der Hesse-Matrix werden also einfach nur die zweiten partiellen Ableitungen eingetragen. Schauen wir uns dies konkret an einem Beispiel an.

Beispiel 51

Wir knüpfen an Beispiel 41 an und berechnen die Hesse-Matrix von $f(x, y) = ax^2 + bxy + cy^2$. Für die zweiten partiellen Ableitungen erhalten wir

$$\frac{\partial^2 f}{\partial x^2}(x, y) = 2a, \quad \frac{\partial f}{\partial x \partial y}(x, y) = \frac{\partial f}{\partial y \partial x}(x, y) = b, \quad \frac{\partial^2 f}{\partial y^2}(x, y) = 2c.$$

Damit erhalten wir die Hesse-Matrix

$$D^2 f(x, y) = \begin{pmatrix} 2a & b \\ b & 2c \end{pmatrix}.$$

■

Die Frage, die man sich noch stellen könnte, ist, wieso die Hesse-Matrix überhaupt immer symmetrisch ist? Denn nur dafür hatten wir ja die Definitheit (siehe Definition 4.3) erklärt. Naja, dies ist ganz einfach. Wenn wir Extrempunkte bestimmen, und die Funktion zweimal (total) differenzierbar ist, kann man das Lemma von Schwarz, siehe Satz 3.8, anwenden, das gerade besagt, dass die zweiten partiellen Ableitungen symmetrisch sind, wenn die Funktion zweimal differenzierbar ist.

Zur Definition 4.2 des Extrempunktes: Diese Erklärung entspricht dem eindimensionalen Fall aus der Analysis 1, siehe [MK09], Definition 11.4. Denn wenn in einer Umgebung der Funktionswert an einer bestimmten Stelle x_0 der größte ist, so spricht man hier von einem lokalen Maximum. Dies sollte klar sein :-).

Zur Definition 4.3 der Definitheit von Matrizen: Die Definitheit von Matrizen kann man noch wesentlich einfacher überprüfen als in Definition 4.3 dargestellt, und zwar können wir die Eigenwerte berechnen. Dann gelten die folgenden Aussagen: Eine quadratische symmetrische (bzw. hermitesche) Matrix ist genau dann

- positiv definit, falls alle Eigenwerte größer als Null sind,
- positiv semidefinit, falls alle Eigenwerte größer oder gleich Null sind,
- negativ definit, falls alle Eigenwerte kleiner als Null sind,
- negativ semidefinit, falls alle Eigenwerte kleiner oder gleich Null sind und
- indefinit, falls positive und negative Eigenwerte existieren.

Ein Kriterium über die sogenannten Hauptminoren werden wir noch in den Kapiteln 9 bzw. 10 kennenlernen, siehe auch Satz 10.7. Wir geben hierzu schon einmal ein Beispiel an.

Beispiel 52

Die symmetrische Matrix

$$A := \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

besitzt $\det A = ac - b^2$. Das heißt, A ist positiv definit, wenn $a > 0$ und $\det A > 0$. ■

4.4 Erklärungen zu den Sätzen und Beweisen

Zum Satz 4.1 (notwendiges Kriterium) und 4.2 (hinreichendes Kriterium): Auch dies sind Verallgemeinerungen von dem, was wir aus der Schule für Funktionen der Form $f : \mathbb{R} \rightarrow \mathbb{R}$, also einer Veränderlichen kennen. Denn der Gradient ist ja quasi die erste Ableitung und die Hesse-Matrix die zweite Ableitung, nur dass hier viele partielle Ableitungen auftauchen.

Beispiel 53

- Wir betrachten die Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^2$. Dann ist

$$\nabla f(x) = 2x, \quad D^2 f(x) = 2\text{Id}.$$

Offensichtlich ist D^2f positiv definit. Also besitzt f an der Stelle $x_0 = 0$ ein isoliertes Minimum.

- Für die Funktion $f(x, y) = x^2 - y^2$ errechnen wir

$$\nabla f(x, y) = 0 \Leftrightarrow x = y = 0$$

und

$$D^2f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Da $D^2f(0, 0)$ indefinit ist (denn die Eigenwerte dieser Matrix sind 2 und -2 , also einmal positiv und einmal negativ), besitzt f keine Extremstellen.

- Sei

$$f(x, y) = \sin(x) + \sin(y) + \sin(x + y)$$

mit $0 < x, y < \frac{\pi}{2}$ gegeben. Um die möglichen Extremstellen zu berechnen, benötigen wir die partiellen Ableitung oder anders formuliert, den Gradienten. Es gilt:

$$f_x(x, y) = \cos(x) + \cos(x + y) \quad \text{und} \quad f_y(x, y) = \cos(y) + \cos(x + y).$$

Der Gradient ergibt sich damit als

$$\nabla f(x, y) = \begin{pmatrix} \cos(x) + \cos(x + y) \\ \cos(y) + \cos(x + y) \end{pmatrix}.$$

Nun muss $\nabla f(x, y) = 0$ als notwendige Bedingung für das Vorhandensein einer Extremstelle stehen.

$$0 = \cos(x) + \cos(x + y)$$

gilt genau dann, wenn $x = y = \frac{\pi}{3}$ in unserem offenen Intervall $(0, \frac{\pi}{2})$. An der Stelle $(\frac{\pi}{3}, \frac{\pi}{3})$ liegt also ein mögliches Extremum vor.

Für die hinreichende Bedingung benötigen wir die Hesse-Matrix. Dazu sind erst einmal die zweiten partiellen Ableitungen nötig, die wir jetzt berechnen zu

$$\begin{aligned} f_{xx}(x, y) &= -\sin(x) - \sin(x + y), \\ f_{yy}(x, y) &= -\sin(y) - \sin(x + y), \\ f_{xy}(x, y) &= f_{yx}(x, y) = -\sin(x + y). \end{aligned}$$

Die Hesse-Matrix lautet folglich

$$D^2f(x, y) = \begin{pmatrix} -\sin(x) - \sin(x + y) & -\sin(x + y) \\ -\sin(x + y) & -\sin(y) - \sin(x + y) \end{pmatrix}.$$

Nun müssen wir den möglichen Extrempunkt $(\frac{\pi}{3}, \frac{\pi}{3})$ einsetzen, um zu entscheiden, ob wirklich einer vorliegt und wenn ja, ob dies ein Minimum oder Maximum ist. Es ergibt sich

$$D^2 f\left(\frac{\pi}{3}, \frac{\pi}{3}\right) = \begin{pmatrix} -\sqrt{3} & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\sqrt{3} \end{pmatrix}.$$

Wir berechnen die Eigenwerte von $D^2 f\left(\frac{\pi}{3}, \frac{\pi}{3}\right)$:

$$\det \begin{pmatrix} -\sqrt{3} - \lambda & -\frac{1}{2}\sqrt{3} \\ -\frac{1}{2}\sqrt{3} & -\sqrt{3} - \lambda \end{pmatrix} = \lambda^2 + 2\sqrt{3}\lambda + \frac{9}{4}.$$

Die Nullstellen dieses charakteristischen Polynoms sind die Eigenwerte, die sich ergeben zu

$$\lambda_1 = -\sqrt{3} + \frac{\sqrt{3}}{2} < 0 \quad \text{und} \quad \lambda_2 = -\sqrt{3} - \frac{\sqrt{3}}{2} < 0.$$

Die Matrix ist also negativ definit und damit liegt bei $(\frac{\pi}{3}, \frac{\pi}{3})$ ein Maximum vor. Der Funktionswert berechnet sich zu

$$f\left(\frac{\pi}{3}, \frac{\pi}{3}\right) = \frac{3}{2}\sqrt{3}.$$

■ Wir untersuchen nun die Funktion

$$f(x, y) := (x^2 + 2y^2)e^{-x^2 - y^2}$$

mit $x, y \in \mathbb{R}$ auf Extrema. Die Vorgehensweise ist analog wie oben.

Die notwendige Bedingung für das Vorhandensein eines Extrempunktes ist nach Satz 4.1, dass der Gradient verschwindet. Wir haben also zunächst einmal die partiellen Ableitungen zu bestimmen. Diese sind gegeben durch

$$\begin{aligned} f_x(x, y) &= 2x \cdot e^{-x^2 - y^2} + (x^2 + 2y^2) \cdot (-2x)e^{-x^2 - y^2} \\ &= 2xe^{-x^2 - y^2}(1 - x^2 - 2y^2) \\ f_y(x, y) &= 4ye^{-x^2 - y^2} + (x^2 + 2y^2)(-2y)e^{-x^2 - y^2} \\ &= 2ye^{-x^2 - y^2}(2 - x^2 - 2y^2). \end{aligned}$$

Der Gradient ist demnach gegeben durch

$$\nabla f(x, y) = \begin{pmatrix} 2xe^{-x^2 - y^2}(1 - x^2 - 2y^2) \\ 2ye^{-x^2 - y^2}(2 - x^2 - 2y^2) \end{pmatrix}. \quad (4.1)$$

Dieser wird Null, wenn die einzelnen Einträge Null sind. Dabei berücksichtigt man, dass ein Produkt genau dann Null wird, wenn einer der Faktoren Null wird. $e^{-x^2 - y^2}$ wird für kein x oder y Null. Dies lernt man schon in der Schule. $2x$ bzw. $2y$ wird Null, wenn $x = y = 0$. Dies ist also schon einmal ein möglicher Extrempunkt.

Das hinreichende Kriterium, siehe Satz 4.2, ist die Überprüfung der Hesse-Matrix auf Definitheit. Dazu muss diese bestimmt, also zuerst die zweiten partiellen Ableitungen berechnet werden. Diese sind gegeben durch

$$\begin{aligned}f_{xx}(x, y) &= e^{-x^2-y^2}(4x^4 - 10x^2 + 8y^2x^2 - 4y + 2) \\f_{yy}(x, y) &= e^{-x^2-y^2}(8y^4 - 20y^2 + 4x^2y^2 + 4 - 2x^2) \\f_{xy}(x, y) &= f_{yx}(x, y) = e^{-x^2-y^2}(-12xy + 4x^3y + 8xy^3).\end{aligned}$$

Die partiellen Ableitungen f_{xy} und f_{yx} sind nach dem Lemma von Schwarz symmetrisch, denn die Funktion ist eine Zusammensetzung differenzierbarer Funktionen und damit insbesondere total differenzierbar. Die Hesse-Matrix lautet

$$D^2f(x, y) = e^{-x^2-y^2} \begin{pmatrix} 4x^4 - 10x^2 + 8y^2x^2 - 4y + 2 & -12xy + 4x^3y + 8xy^3 \\ -12xy + 4x^3y + 8xy^3 & 8y^4 - 20y^2 + 4x^2y^2 + 4 - 2x^2 \end{pmatrix}$$

Es ist nun

$$D^2f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

Diese Matrix ist symmetrisch. Damit reell diagonalisierbar. Die Eigenwerte kann man sofort zu $\lambda_1 = 2$ und $\lambda_2 = 4$ ablesen. Die Matrix besitzt also nur positive Eigenwerte und ist damit positiv definit. Folglich liegt an der Stelle $(0, 0)$ ein Minimum vor.

Wer sagt uns aber, dass dies der einzige Extrempunkt ist? In unserem Gradienten (4.1) haben wir ja in jeder Komponente noch einen Faktor, der eventuell Null werden kann. Wir müssen uns also fragen, ob das Gleichungssystem, gegeben durch die beiden Gleichungen

$$x(1 - x^2 - 2y^2) = 0 \quad \text{und} \quad y(2 - x^2 - 2y^2) = 0$$

weitere Lösungen besitzt. Dies soll nun eine Übungsaufgabe an euch sein. Soviel sei verraten: Eine weitere Lösung ist $x = 0$ und $y = 1$.

- Wir wollen die absoluten und lokalen Extrema von

$$f(x, y) = x^4 + \frac{1}{2}y^2 + \cos(x^2 + y^2)$$

in der Kreisscheibe $x^2 + y^2 \leq \frac{\pi}{2}$ berechnen. Wegen der Symmetrie reicht es aus, f nur im Viertelkreis des 1. Quadranten zu untersuchen, denn es gilt:

$$f(x, y) = f(-x, y) = f(-x, -y) = f(x, -y).$$

Die partiellen Ableitungen sind gegeben durch

$$\begin{aligned}f_x(x, y) &= 2x(2x^2 - \sin(x^2 + y^2)), \\f_y(x, y) &= y(1 - \sin(x^2 + y^2)), \\f_{xx}(x, y) &= 12x^2 - 2\sin(x^2 + y^2) - 4x^2 \cos(x^2 + y^2), \\f_{yy}(x, y) &= 1 - 2\sin(x^2 + y^2) - 4y^2 \cos(x^2 + y^2), \\f_{xy}(x, y) &= f_{yx}(x, y) = -4xy \cos(x^2 + y^2).\end{aligned}$$

Es muss nun

$$2x(2x^2 - \sin(x^2 + y^2)) = 0 \quad (4.2)$$

und

$$y(1 - \sin(x^2 + y^2)) = 0 \quad (4.3)$$

gelten. Gleichung (4.2) wird Null, wenn $x = 0$ oder wenn $\sin(x^2 + y^2) = 2x^2$. Die Gleichung (4.3) wird Null, wenn $y = 0$ oder wenn $\sin(x^2 + y^2) = \frac{1}{2}$. Dies ist äquivalent zu

$$\begin{aligned}(x, y) &= 0, \\(x, y) &= \left(0, \sin y^2 = \frac{1}{2}\right) = \left(0, \sqrt{\frac{\pi}{6}}\right), \\(x, y) &= (\sin x^2 = 2x^2, 0) = (0, 0), \\(x, y) &= \left(\frac{1}{2}, \sqrt{\frac{\pi}{6} - \frac{1}{4}}\right).\end{aligned}$$

Unsere Untersuchungen wurden dadurch erleichtert, dass wir wegen der Symmetrie (siehe Bemerkung oben) das Ganze nur im ersten Quadranten untersuchen mussten. Schauen wir uns an, was im Nullpunkt passiert.

Für die Hesse-Matrix gilt:

$$D^2 f(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}.$$

Die Eigenwerte sind 0 und 1, das heißt, die Matrix ist positiv definit. Daher ist Satz 4.2 nicht anwendbar. Wenn dies der Fall ist, heißt dies nicht, das kein Extremum vorliegen kann! Wir müssen deshalb auf die Definition 4.2 eines lokalen Extremums zurückgreifen.

Für $0 < |x| < \frac{1}{4}$ und $0 < |y| < \frac{1}{4}$ ist $x^4 + \frac{1}{2}y^2 < \frac{1}{2}$. Hieraus folgt

$$\begin{aligned}f(x, y) - f(0, 0) &= x^4 + \frac{1}{2}y^2 + \cos(x^2 + y^2) - 1 \\&= x^4 + \frac{1}{2}y^2 - 2\sin^2\left(\frac{x^2 + y^2}{2}\right) \\&\geq x^4 + \frac{1}{2}y^2 - 2\left(\frac{x^2 + y^2}{2}\right)^2 > 0.\end{aligned}$$

Daher ist $f(x, y) - f(0, 0) > 0$, also $f(x, y) > f(0, 0)$ und damit liegt bei $(0, 0)$ ein relatives Minimum vor. Als kleine Ergänzung: In den letzten beiden Schritten haben wir ausgenutzt, dass $\sin^2(x) = \frac{1}{2}(1 - \cos(2x))$.

Zu den anderen beiden möglichen Extrempunkten: Dies macht man wieder wie in den obigen Beispielen. Einsetzen in die Hesse-Matrix, Überprüfen auf Definitheit und Folgerung. Wenn man dies macht, stellt man fest, dass bei $(0, \sqrt{\frac{\pi}{6}})$ ein relatives Maximum und bei $(\frac{1}{2}, \sqrt{\frac{\pi}{6} - \frac{1}{4}})$ ein Sattelpunkt vorliegt.

Merke also: Wenn das hinreichende Kriterium 4.2 mit der Hesse-Matrix nicht anwendbar ist, so gehe auf die Definition 4.2 des lokalen Extrempunktes zurück und untersuche separat :-).

- So, jetzt noch ein letztes Beispiel, das wir mit einer Zeichnung untermalen wollen. Gegeben sei die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = e^{-x^2 - y^2}.$$

Schauen wir uns die Funktion einfach einmal in Abbildung 4.1 an. Wir

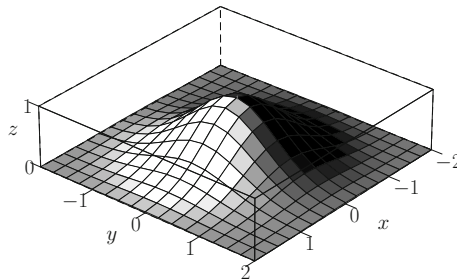


Abb. 4.1: Die Funktion $f(x, y) = e^{-x^2 - y^2}$.

sehen also schon, dass vermutlich ein Hochpunkt vorliegt. Bestimmen wir ihn also auch einmal! Wir berechnen zunächst den Gradienten. Da wir hier aber eine Abbildung der Form $f : \mathbb{R}^n \rightarrow \mathbb{R}$ betrachten, ist der Gradient mit der Jakobi-Matrix identisch. Diese ist gegeben durch

$$J_f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right) = -e^{-x^2 - y^2} (2x, 2y).$$

Die Hesse-Matrix lautet

$$D^2 f(x, y) = 2e^{-x^2 - y^2} \begin{pmatrix} 2x^2 - 1 & 2xy \\ 2xy & 2y^2 - 1 \end{pmatrix}.$$

Der Gradient bzw. die Jakobi-Matrix wird nur dann Null, wenn $x = y = 0$. Einsetzen in die Hesse-Matrix liefert

$$D^2 f(0, 0) = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Die Eigenwerte liest man sofort zu $\lambda_{1,2} = -1$ ab. Da alle Eigenwerte negativ sind, ist die Hesse-Matrix im Nullpunkt negativ definit und damit liegt bei $(0, 0)$ ein Hochpunkt vor.

Dies soll an Beispielen genügen. ■

Zum Satz 4.3 (Extrema unter Nebenbedingungen): Was genau bedeutet Extrema unter Nebenbedingungen? Dies wollen wir uns jetzt anschauen. Betrachten wir beispielsweise eine Polynomfunktion $f(x, y)$ vierten Grades. Uns interessiert nun das Verhalten der Funktion auf dem Einheitskreis. Also genau: Wo nimmt f auf dem Einheitskreis ihre Extremwerte an? Viele denken sich vielleicht nun: Gut, dann nehme ich mir doch einfach die Kreisgleichung und betrachte diese auf f . Ja, dies ist eine Möglichkeit, aber verkompliziert vielmehr die Rechnungen als dass sie einfacher werden. Und genau um einfachere Rechnungen durchführen zu können, verwendet man die sogenannte Lagrange-Methode, die wir uns jetzt an ein paar Beispielen anschauen wollen. Eine andere Motivation ist die Folgende: Bis jetzt haben wir immer Extremwerte von Funktionen gesucht bzw. berechnet, die auf offenen Mengen definiert sind. Wir haben den Gradienten berechnet, diesen Null gesetzt und das entstehende Gleichungssystem gelöst. Dies geht aber nicht mehr, wenn Ω kein Inneres besitzt. Beispielsweise, wenn wir den Extrempunkt einer Funktion auf der Kugeloberfläche (Sphäre S^1) berechnen wollen, die ja gegeben ist durch

$$K := \{(x, y, z) : x^2 + y^2 + z^2 = 1\}.$$

Bevor wir zu Beispielen kommen, wollen wir die Motivation von Satz 4.3 noch einmal ein wenig näherbringen. Wir müssen es ja irgendwie schaffen, dass wir das Problem so zurückführen, sodass wir wieder einen offenen Definitionsbereich erhalten. Betrachten wir ein Standardbeispiel:

Beispiel 54

Sei $U \in \mathbb{R}$ beliebig. Wir wollen ein Rechteck mit Umfang U so finden, dass der Flächeninhalt A maximal wird. Das Rechteck besitze die Seiten x und y . Dann müssen wir also die Funktion $f(x, y) = x \cdot y$ unter der Bedingung $2x + 2y = U$ maximieren. Oder anders formuliert: f muss auf der Menge

$$B := \{(x, y) \in \mathbb{R}^2 : 2x + 2y = U, x, y \geq 0\}$$

maximal werden. Übrigens beschreibt B eine Gerade, und B hat in \mathbb{R}^2 kein Inneres, daher können wir die bekannten Methoden mit dem Aufstellen des Gradienten und Nullsetzen nicht anwenden. Wenn man dies machen würde, so würden wir $(x, y) = (0, 0)$ erhalten, was uns aber auch nicht wirklich weiterbringen würde. Aber schon mit Schulkenntnissen könnt ihr diese Aufgabe lösen. Wir formen $2x + 2y = U$ einfach nach y um zu $y = \frac{U-2x}{2}$ und setzen dies in f ein. So erhalten wir eine Funktion, die nur noch von y abhängig ist, die also nur

eine Variable enthält, und von dieser können wir dann locker den Extrempunkt berechnen. Übrigens: Es kommt $x = y = \frac{U}{4}$ raus, das heißt wir erhalten ein Quadrat. Rechnet dies ruhig einmal nach! ■

Dieses Beispiel ist sehr lehrreich, wenn man sich die Idee anschaut. Die Idee war ja, die Gleichung zum Eliminieren einer Variable zu verwenden, die den Definitionsbereich beschrieben hat. Dies läuft einfach darauf hinaus, eine Variable als durch diese Gleichung implizit definiert aufzufassen. Aus dem Satz über die impliziten Funktionen aus dem noch folgenden Kapitel 5 wissen wir sogar, dass solch eine Auflösung in diesem Fall existiert und möglich ist. Leider müssen wir euch aber sagen, dass dies nicht immer zum Ziel führen wird, wenn die auftretenden Funktionen nicht explizit vorgegeben sind. Daher brauchen wir eine andere Idee: Gegeben seien eine offene Teilmenge $U \subset \mathbb{R}^n$ und eine stetig differenzierbare Funktion f , sowie $F : U \rightarrow \mathbb{R}$. Es soll

$$B := \{x : x \in U, F(x) = 0\}$$

gelten. Zu berechnen sind nun Extremwerte von f auf B . Die Bedingung $F(x) = 0$ nennt man auch die Nebenbedingung. Im Beispiel 54 war dies gerade $F(x, y) = 2x + 2y - U$. Man kann nun die folgende Beobachtung machen, dass lokale Extremwerte $x_0 \in B$ von f auf B zu erwarten sind, wenn $\text{grad}F(x_0)$ und $\text{grad}f(x_0)$ parallel sind, für die es also ein λ mit der Eigenschaft

$$\text{grad}F(x_0) = \lambda \text{grad}f(x_0)$$

gibt. Das kann man sich geometrisch überlegen: Wir betrachten den zweidimensionalen Fall mit einer Nebenbedingung. Wir wollen eine Funktion $f(x, y)$ unter der Nebenbedingung $F(x, y) = c$ für eine Konstante c maximieren. Verfolgen wir die Höhenlinien von $F(x, y) = c$, so berühren oder kreuzen wir sogar die Höhenlinien von $f(x, y)$. Wir können demnach nur einen gemeinsamen Punkt (x, y) der Nebenbedingung $F(x, y) = c$ und einer Höhenlinie $f(x, y) = d$ finden, der ein Extremwert von f ist, wenn die Bewegung auf der Höhenlinie $F(x, y) = c$ tangential zu $f(x, y) = d$ ist. Geometrisch übersetzen wir die Tangentenbedingung, indem wir fordern, dass der Gradient von f zum Gradienten von F parallel sein soll.

Dies führt schließlich auf Satz 4.3 und liefert die sogenannte Lagrange-Methode, um Extremwerte unter Nebenbedingungen zu berechnen. Schauen wir uns Beispiele an und üben dies ein!

Beispiel 55

- Der Abstand d zwischen dem Punkt $(0, 4)$ und einem Punkt (x, y) auf der Hyperbel $x^2 - y^2 = 12$ ist definiert durch

$$d(x, y) = \sqrt{x^2 + (y - 4)^2}.$$

Diese Funktion wollen wir unter der Nebenbedingung $x^2 - y^2 = 12$ minimieren. Da es lästig ist, $\sqrt{x^2 + (y - 4)^2}$ zu differenzieren (wegen der Wurzelausdrücke), ist es gleichbedeutend mit der Funktion

$$d^2(x, y) = x^2 + (y - 4)^2 \quad (4.4)$$

zu arbeiten. Wir geben im Folgenden zwei Lösungswege an, um zu zeigen, dass die Lagrange-Methode bei „kleinen“ Problemen nicht immer der schnellste Weg ist.

1. Lösungsweg: Wir formen $x^2 - y^2 = 12$ nach x^2 um. Dies ergibt gerade $x^2 = 12 + y^2$. Dies setzen wir in (4.4) ein und erhalten so

$$d(y) := (12 + y^2) + (y - 4)^2.$$

Diese Funktion hängt nur von einer Veränderlichen ab. Wir verfahren also wie in Analysis 1. Zunächst ist

$$d'(y) = 2y + 2(y - 4) = 4(y - 2) \quad \text{und} \quad d''(y) = 4.$$

Es gilt $d'(y) = 0$ genau dann, wenn $y = 2$ ist. Da $d''(2) = 4 > 0$, liegt bei $y = 2$ ein Minimum vor.

2. Lösungsweg: Nach der Lagrange-Methode existiert eine Funktion

$$L(x, y, \lambda) = f(x, y) + \lambda g(x, y)$$

mit der Nebenbedingung

$$g(x, y) = x^2 - y^2 - 12 = 0.$$

Wir erhalten demnach

$$L(x, y, \lambda) = x^2 + (y - 4)^2 + \lambda(x^2 - y^2 - 12).$$

Nun sind die Bedingungen

$$\frac{\partial}{\partial x} L = \frac{\partial}{\partial y} L = \frac{\partial}{\partial \lambda} L = 0$$

zu lösen. So ergeben sich die drei Gleichungen

$$\frac{\partial}{\partial x} L = 0 \Rightarrow 2x + 2\lambda x = 2x(1 + \lambda) = 0, \quad (4.5)$$

$$\frac{\partial}{\partial y} L = 0 \Rightarrow 2(y - 4) - 2y\lambda = 2y - 2\lambda y - 8 = 0, \quad (4.6)$$

$$\frac{\partial}{\partial \lambda} L = 0 \Rightarrow x^2 - y^2 - 12 = 0. \quad (4.7)$$

Aus (4.5) folgt sofort $\lambda = -1$. Dies setzen wir in (4.6) ein und erhalten $y = 2$. Setzen wir nun $y = 2$ in (4.7) ein, so ergibt sich $x = \pm 4$.

■ Gegeben sei ein Ellipsoid

$$E : \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1$$

mit $x, y, z \in \mathbb{R}$. Gesucht ist der achsenparallele Quader mit größtem Volumen, welcher zentriert in einem kartesischen Koordinatensystem liegt, der E eingeschrieben werden kann, siehe Abbildung 4.2.

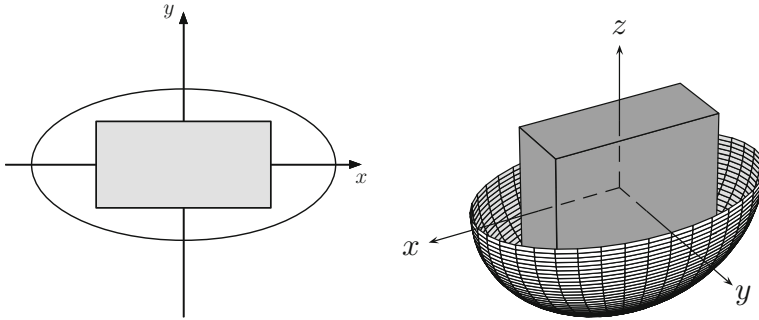


Abb. 4.2: Ein Ellipsoid mit einem achsenparallelen Quader mit größtem Volumen mit $a = 2, b = 4, c = 3$. Einmal im Querschnitt mit $z = 0$, einmal ein offenes 3D-Modell.

Das Volumen des beschriebenen Quaders lautet

$$V = 8 \cdot x \cdot y \cdot z.$$

Wir betrachten nun die Funktion V unter der Nebenbedingung von f , wobei

$$f(x, y, z) = \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} - 1.$$

Es gilt nun:

$$J_V(x, y, z) = (8yz, 8xz, 8xy)$$

und

$$J_f(x, y, z) = \left(2\frac{x}{a^2}, 2\frac{y}{b^2}, 2\frac{z}{c^2} \right).$$

Es gilt nun $\text{Rang}(J_f) = 1$ für alle $(x, y, z) \in E$. Daher gibt es ein eindeutig bestimmtes λ mit

$$J_V(x, y, z) + \lambda \cdot J_f(x, y, z) = 0 \Leftrightarrow (8yz, 8xz, 8xy) + \lambda \cdot \left(2\frac{x}{a^2}, 2\frac{y}{b^2}, 2\frac{z}{c^2} \right) = 0.$$

Dies entspricht dem Gleichungssystem

$$\begin{aligned}8yz + \lambda \cdot 2 \cdot \frac{x}{a^2} &= 0, \\8xz + \lambda \cdot 2 \cdot \frac{y}{b^2} &= 0, \\8xy + \lambda \cdot 2 \cdot \frac{z}{c^2} &= 0, \\8xyz + \lambda \cdot 2 \cdot \frac{x^2}{a^2} &= 0, \\8xyz + \lambda \cdot 2 \cdot \frac{y^2}{b^2} &= 0, \\8xyz + \lambda \cdot 2 \cdot \frac{z^2}{c^2} &= 0.\end{aligned}$$

Addieren wir alle Gleichungen, so erhalten wir

$$24xyz + 2 \cdot \lambda \cdot \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \right) = 0.$$

Die Nebenbedingung ist aber

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

also folgt

$$24xyz + 2 \cdot \lambda = 0 \Leftrightarrow \lambda = -12 \cdot xyz.$$

Einsetzen in die ersten beiden Gleichungen von oben liefert

$$\begin{aligned}0 &= 8yz - 24x^2y \cdot \frac{z}{a^2} \Rightarrow x = \sqrt{\frac{1}{3}} \cdot a, \\0 &= 8xz - 24y^2x \cdot \frac{z}{b^2} \Rightarrow y = \sqrt{\frac{1}{3}} \cdot b, \\0 &= 8xy - 24z^2x \cdot \frac{y}{c^2} \Rightarrow z = \sqrt{\frac{1}{3}} \cdot c.\end{aligned}$$

Damit ergibt sich das maximale Volumen als

$$V_{\max} = 8xyz = \frac{8}{3} \cdot \frac{1}{\sqrt{3}} abc,$$

denn die stetige Funktion V muss auf der kompakten Menge E nach Satz 2.4 aus Kapitel 2 ein Maximum annehmen. ■

Bei dem Lagrange-Verfahren finden wir immer nur mögliche Kandidaten. Woher wissen wir aber, dass dies tatsächlich Maxima bzw. Minima sind? Dies folgt aus dem Satz 2.4, dass stetige Funktionen auf kompakten Mengen Maxima und Minima annehmen. Unterscheiden, ob ein Maximum oder Minimum vorliegt, geht ganz einfach: Man setzt aus einer entsprechenden Umgebung einfach ein paar Werte in seine zu untersuchende Funktion ein und überprüft, ob die Funktionswerte größer oder kleiner werden.

5 Implizite Funktionen

Übersicht

5.1	Sätze und Beweise	107
5.2	Erklärungen zu den Sätzen und Beweisen	111

Wir wollen uns in diesem Kapitel damit beschäftigen, unter welchen Umständen man eine implizite Funktion in zwei oder mehr Variablen (das heißt eine Funktion, in der nicht eine Variable in Abhängigkeit von der anderen explizit gegeben ist) nach einer Variablen auflösen kann. Dies ist wichtig, da nicht jede Gleichung explizit auflösbar ist, zum Beispiel können wir die Gleichung

$$\sin x \sin y + x \cos y + y \cos x = 0$$

nicht explizit nach y auflösen, wollen aber trotzdem Aussagen über die Abhängigkeit von y bezüglich x treffen, zum Beispiel ob y in einer Umgebung von x differenzierbar von x abhängt.

Wir formulieren die folgenden Sätze für Banach-Räume, weil die Sätze und Beweise so allgemein sind und im Spezialfall des \mathbb{R}^n auch nicht einfacher wären. Man kann sich für V und W aber jederzeit \mathbb{R}^n und \mathbb{R}^m denken.

5.1 Sätze und Beweise

Satz 5.1 (Banach'scher Fixpunktsatz für eine stetige Familie)

Seien V und W Banach-Räume, $\Omega \subset W$ offen und $A \subset V$ abgeschlossen. Für jedes $x \in \Omega$ sei $T(x) : A \rightarrow A$ eine stetige Abbildung mit

$$\|T(x)(y_1) - T(x)(y_2)\| \leq \eta |y_1 - y_2| \quad \forall y_1, y_2 \in A.$$

mit einem $0 < \eta < 1$, das nicht von x abhängt. Dann existiert für jedes $x \in \Omega$ genau ein $y(x) \in A$ mit

$$T(x)(y(x)) = y(x).$$

und y ist stetig in x .

Anmerkung: Was eine stetige Familie ist, könnt ihr in den Erklärungen nachlesen.

Beweis: Für eine beliebige stetige Funktion $y_0 : \Omega \rightarrow A$ definieren wir rekursiv die stetigen Funktionen $y_{n+1}(x) := T(x)(y_n(x))$. Dann ist

$$\begin{aligned} y_n(x) &= \sum_{k=1}^n (y_k(x) - y_{k-1}(x)) + y_0(x) \\ &= \sum_{k=1}^n (T^{k-1}(x)(y_1(x)) - T^{k-1}(x)(y_0(x))) + y_0(x), \end{aligned}$$

wobei T^{k-1} einfach die $(k-1)$ -malige Ausführung von T bedeutet. Wegen

$$\begin{aligned} \left\| \sum_{k=1}^n (T^{k-1}(x)y_1(x) - T^{k-1}(x)y_0(x)) \right\| &\leq \sum_{k=1}^n \eta^{k-1} \|y_1(x) - y_0(x)\| \\ &\leq \frac{\|y_1(x) - y_0(x)\|}{1 - \eta} \end{aligned}$$

(im letzten Schritt schätzen wir durch eine geometrische Reihe ab) konvergiert die Folge (y_n) absolut und gleichmäßig, und die Grenzfunktion

$$y(x) := \lim_{n \rightarrow \infty} y_n(x)$$

ist daher auch stetig. Da A abgeschlossen ist, ist auch $y(x) \in A$, das heißt $y : \Omega \rightarrow A$. Aus $y_{n+1}(x) = T(x)(y_n(x))$ und der Stetigkeit von $T(x)$ folgt nun wie im Banach'schen Fixpunktsatz 2.6

$$y(x) = T(x)(y(x)) \quad \forall x \in \Omega$$

und die Eindeutigkeit.

q.e.d.

Satz 5.2 (Satz über implizite Funktionen)

Seien V_1, V_2 und W Banach-Räume, $\Omega \subset V_1 \times V_2$ offen, $(x_0, y_0) \in \Omega$ und $f : \Omega \rightarrow W$ in Ω stetig differenzierbar. Es gelte $f(x_0, y_0) = 0$. Angenommen, die stetige lineare Abbildung

$$D_2 f(x_0, y_0) : V_2 \rightarrow W$$

ist invertierbar und die Inverse ebenfalls stetig.

Dann existieren offene Umgebungen Ω_1 von x_0 und Ω_2 von y_0 mit $\Omega_1 \times \Omega_2 \subset \Omega$ und eine differenzierbare Abbildung $g : \Omega_1 \rightarrow \Omega_2$ mit

$$f(x, g(x)) = 0 \quad \forall x \in \Omega_1.$$

Ferner ist $g(x)$ für alle $x \in \Omega_1$ die einzige in Ω_2 enthaltene Lösung dieser Gleichung. Für die Ableitung von g gilt die Gleichung

$$Dg(x) = -(D_2 f(x, g(x)))^{-1} \circ D_1 f(x, g(x)) \quad \forall x \in \Omega_1.$$

Beweis: Wir setzen der Einfachheit halber $L := D_2 f(x_0, y_0)$. Da L linear und nach Voraussetzung invertierbar ist, folgt

$$f(x, y) = 0 \Leftrightarrow L^{-1}f(x, y) = 0 \Leftrightarrow y = y - L^{-1}f(x, y).$$

Sei nun $G(x, y) := y - L^{-1}f(x, y)$. Wir wollen nun den Banach'schen Fixpunktsatz 5.1 auf G anwenden.

- Wegen $L^{-1} \circ L = \text{Id}_{V_2}$ gilt:

$$\begin{aligned} G(x, y_1) - G(x, y_2) &= y_1 - y_2 - L^{-1}(f(x, y_1) - f(x, y_2)) \\ &= L^{-1}(D_2 f(x_0, y_0)(y_1 - y_2) - (f(x, y_1) - f(x, y_2))). \end{aligned}$$

Da f differenzierbar und L^{-1} stetig ist, folgt die Existenz von $\delta_1 > 0, \eta > 0$, sodass für alle x, y_1, y_2 mit $\|x - x_0\| < \delta_1, \|y_0 - y_1\| < \eta, \|y_0 - y_2\| < \eta$

$$\|G(x, y_1) - G(x, y_2)\| \leq \frac{1}{2}\|y_1 - y_2\|$$

gilt. Wegen der Stetigkeit von G existiert dazu ein $\delta_2 > 0$ mit

$$\|G(x, y_0) - G(x_0, y_0)\| \leq \frac{\eta}{2} \quad \forall \|x - x_0\| < \delta_2.$$

Ist nun $\|y - y_0\| \leq \eta$, so ist wegen $G(x_0, y_0) = y_0$ für alle $\|x - x_0\| \leq \delta := \min\{\delta_1, \delta_2\}$

$$\begin{aligned} \|G(x, y) - y_0\| &= \|G(x, y) - G(x_0, y_0)\| \\ &\leq \|G(x, y) - G(x, y_0)\| + \|G(x, y_0) - G(x_0, y_0)\| \\ &\leq \frac{1}{2}\|y - y_0\| + \frac{\eta}{2} \leq \eta. \end{aligned}$$

Für jedes feste x mit $\|x - x_0\| \leq \delta$ bildet $G(x, \cdot)$ also die abgeschlossene Kugel $B(y_0, \eta)$ auf sich selbst ab, ist dort also wegen

$$\|G(x, y_1) - G(x, y_2)\| \leq \frac{1}{2}\|y_1 - y_2\|$$

eine Kontraktion.

Wir können also den Banach'schen Fixpunktsatz für eine stetige Familie anwenden und erhalten damit eine stetige Funktion $y : U(x_0, \delta) \rightarrow B(y_0, \eta)$ mit

$$G(x, y(x)) = y(x).$$

- Wir nennen die oben gefundene Funktion nun g und wollen noch die Differenzierbarkeit sowie die behauptete Formel beweisen. Sei dazu $(x_1, y_1) \in U(x_0, \delta) \times U(y_0, \eta)$ mit $y_1 = g(x_1)$. Es gilt wegen unserer ersten Überlegung im Beweis $f(x_1, y_1) = f(x_1, g(x_1)) = 0$, und da f in (x_1, y_1) differenzierbar ist, können wir die Taylor-Entwicklung um (x_1, y_1) betrachten. Diese lautet

$$f(x, y) = D_1 f(x_1, y_1)(x - x_1) + D_2 f(x_1, y_1)(y - y_1) + R(x, y)$$

mit

$$\lim_{(x,y) \rightarrow (x_1,y_1)} \frac{R(x,y)}{\|(x-x_1, y-y_1)\|} = 0.$$

Da L invertierbar und L^{-1} stetig ist, können wir wegen der Stetigkeit von f δ und η so klein wählen, dass auch $D_2f(x_1, y_1)$ invertierbar mit stetiger Inverser ist für alle $(x_1, y_1) \in U(x_0, \delta) \times U(y_0, \eta)$. Da aber für alle $x \in U(x_0, \delta)$ ja $f(x, g(x)) = 0$ gilt, folgt aus der Taylor-Entwicklung

$$\begin{aligned} g(x) &= - (D_2f(x_1, y_1))^{-1} \circ D_1f(x_1, y_1)(x - x_1) \\ &\quad + y_1 - (D_2f(x_1, y_1))^{-1}(R(x, g(x))). \end{aligned}$$

Aus den Eigenschaften von R folgt nun die Existenz von $\rho_1, \rho_2 > 0$ mit

$$\|R(x, y)\| \leq \frac{\|x - x_1\| + \|y - y_1\|}{2\|(D_2f(x_1, y_1))^{-1}\|} \quad \forall \|x - x_1\| \leq \rho_1, \|y - y_1\| \leq \rho_2,$$

das heißt insbesondere

$$\|R(x, y)\| \leq \frac{\|x - x_1\| + \|g(x) - g(x_1)\|}{2\|(D_2f(x_1, y_1))^{-1}\|} \quad \forall \|x - x_1\| \leq \rho_1, \|y - y_1\| \leq \rho_2,$$

und dies ergibt zusammen mit der vorherigen Gleichung

$$\begin{aligned} \|g(x) - g(x_1)\| &\leq \|(D_2f(x_1, y_1))^{-1} \circ D_1f(x_1, y_1)\| \|x - x_1\| \\ &\quad + \frac{1}{2} \|x - x_1\| + \frac{1}{2} \|g(x) - g(x_1)\|, \end{aligned}$$

also

$$\|g(x) - g(x_1)\| \leq (2\|(D_2f(x_1, y_1))^{-1} \circ D_1f(x_1, y_1)\| + 1) \|x - x_1\|.$$

Sei nun

$$r(x) := -(D_2f(x_1, y_1))^{-1}(R(x, g(x))).$$

Dann folgt

$$g(x) - g(x_1) = -(D_2f(x_1, y_1))^{-1} \circ D_1f(x_1, y_1)(x - x_1) + r(x).$$

Aus $\lim_{x \rightarrow x_1} \frac{R(x, g(x))}{\|x - x_1\|} = 0$ folgt nun

$$\lim_{x \rightarrow x_1} \frac{r(x)}{x - x_1} = 0,$$

und daraus folgt die Differenzierbarkeit von g in x_1 und die behauptete Formel.

q.e.d.

5.2 Erklärungen zu den Sätzen und Beweisen

Der Einfachheit halber sollte man sich in diesem Kapitel immer vorstellen, dass man als Banach-Räume den \mathbb{R}^n beziehungsweise \mathbb{R}^m vorliegen hat. Alle unsere Beispiele werden nur diesen Fall behandeln.

Zum Banach'schen Fixpunktsatz für eine stetige Familie (Satz 5.1): Dieser Satz ist ganz analog zum „normalen“ Banach'schen Fixpunktsatz 2.6, nur dass wir hier für jedes x eine andere Funktion $T(x)$ erhalten, die man dann auf y anwendet. Auch hier zeigt man wieder zunächst eine Konvergenzaussage, und der Rest folgt dann analog wie im Banach'schen Fixpunktsatz.

Was aber haben wir unter einer Familie stetiger Funktionen zu verstehen? Betrachten wir als Beispiel den Ausdruck $x + y$ für $x, y \in \mathbb{R}^n$. Dann könnten wir einerseits die Funktion $f(x, y) = x + y$ betrachten. Diese ist natürlich stetig. Hierbei sind x und y variabel. Andererseits könnten wir zunächst x frei wählen, zum Beispiel $x = \pi, \sqrt{2}, 1337$. Dann erhalten wir für jedes feste x eine Funktion $f_x(y) = x + y$, die dann nur noch von y abhängt, da x ja fest ist. Diese Funktion ist dann stetig in y . Mehr noch, diese Familie von Funktionen hängt auch stetig von x ab. Also ist eine stetige Familie stetiger Funktionen eine Menge von Funktionen, die alle stetig sind und von einem Parameter stetig abhängen.

Zum Satz über implizite Funktionen (Satz 5.2): Dieser Satz stellt nun das Resultat dar, das wir in diesem Kapitel beweisen wollten. Er gibt uns ein Kriterium, wann eine implizit gegebene Funktion (zumindest in einer Umgebung) nach einer der Variablen auflösbar ist. Dabei ist dieses Resultat so genial, wie auch einleuchtend, was wir kurz an folgendem Bild (siehe Abbildung 5.1) erläutern wollen. Betrachten wir den Punkt p_1 und einen kleinen Teil der Kurve, die die

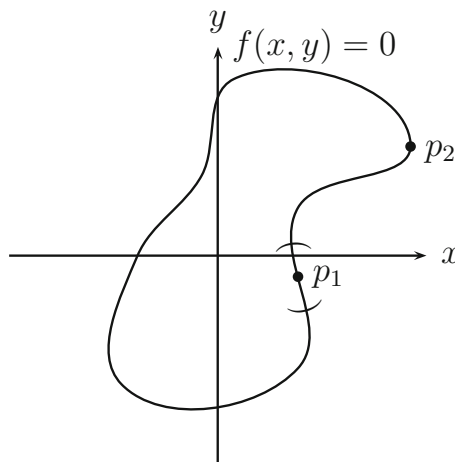


Abb. 5.1: Der Satz über implizite Funktionen graphisch veranschaulicht.

Gleichung $f(x, y) = 0$ beschreibt. Denken wir uns nun den Rest weg, so erhalten wir etwas, was aussieht wie eine Funktion, das heißt, jedem x -Wert wird genau ein y -Wert zugeordnet. Die implizite Funktion f ist an diesem Punkt also nach y auflösbar. Betrachten wir dagegen p_2 , so können wir keine Umgebung finden, sodass die Kurve dort wie eine Funktion aussieht. Damit können wir nichts über die Auflösbarkeit sagen. Es können Fälle auftreten, wie zum Beispiel $y = x^2$ im Punkt 0, wo die Funktion nicht auflösbar ist. Es kann aber auch passieren, wie bei der Funktion $y = x^3$ im Punkt 0, dass die Funktion dann doch auflösbar ist. Wenn D_2 invertierbar ist, kann man also nach y auflösen. Wenn D_2 nicht invertierbar ist, kann man für die Existenz der Umkehrfunktion zunächst nichts sagen.

Der Beweis gibt uns nur eine Existenzaussage, das Tolle ist aber, dass wir sogar Aussagen über die Ableitung treffen können. Leider ist dies oftmals so kompliziert, dass wir daraus nicht die Funktion selbst berechnen können.

Bevor wir nun näher auf den Beweis und Beispiele eingehen, noch eine kurze Anmerkung zu der Annahme $f(x_0, y_0) = 0$. Diese ist nicht notwendig für die Auflösbarkeit, das Kriterium behält seine Gültigkeit für $f(x_0, y_0) = c$, wobei c beliebig ist. Wieso? Nun, anschaulich folgt dies genau aus unseren obigen Bemerkungen. Mathematisch solltet ihr euch das einmal selber überlegen, bei Problemen schaut ins Forum ;).

Einige Anmerkungen zu dem etwas technischen und langwierigen Beweis. Wir fangen damit an, uns eine Abbildung zu definieren, auf die wir den Banach'schen Fixpunktsatz in der Version von Satz 5.1 anwenden wollen. Im ersten Punkt wird dann mithilfe von Stetigkeit und Differenzierbarkeit der betrachteten Funktionen gezeigt, dass wir den Satz hier tatsächlich anwenden können, das heißt dass eine Kontraktion vorliegt. Im zweiten Teil weisen wir dann mithilfe der Taylor-Entwicklung die Differenzierbarkeit und die Formel für die Ableitung nach.

Aber nun einmal zu einigen Beispielen.

Beispiel 56

- Wir beginnen mit der Funktion (siehe Abbildung 5.2)

$$f(x, y) = x^2 + y^2 - 1.$$

Wir untersuchen also die Gleichung $f(x, y) = 0$ (siehe Abbildung 5.3) und wollen zunächst nur den Satz über implizite Funktionen anwenden, ohne auf die Voraussetzungen zu achten, dies werden wir in den weiteren Beispielen tun. Dies lässt sich explizit auflösen und wir erhalten

$$y(x) = \pm \sqrt{1 - x^2}$$

und damit

$$y'(x) = \frac{-x}{\sqrt{1 - x^2}} = \frac{-x}{y(x)} \text{ für } y(x) = \sqrt{1 - x^2}$$

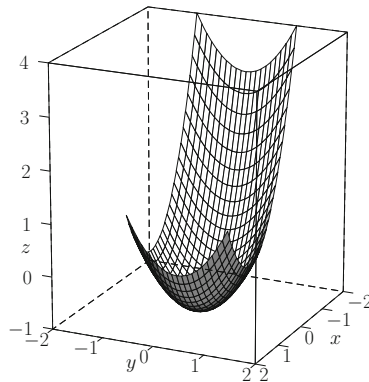


Abb. 5.2: Die Funktion $f(x, y) = x^2 + y^2 - 1$.

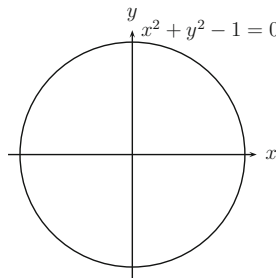


Abb. 5.3: Die Niveaulinie $f(x, y) = 0$.

Wenden wir stattdessen den Satz über implizite Funktionen an, so erhalten wir

$$y'(x) = -(D_2 f(x, y))^{-1} \circ D_1 f(x, y) = -(2y)^{-1} 2x = \frac{2x}{-2y} = \frac{-x}{y(x)},$$

also dasselbe Ergebnis, nur dass es mit der impliziten Methode schneller ging, wobei der Nachteil hier ist, dass wir nicht y direkt erhalten.

- Wir betrachten die Funktion (siehe Abbildung 5.4)

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = y^2 - x^2(1 - x^2).$$

Es gilt:

$$D_2 f(x, y) = \frac{\partial f}{\partial y}(x, y) = 2y.$$

Folglich ist $D_2 f(x, y)$ für alle $(x, y) \in \mathbb{R}^2$ mit $y \neq 0$ invertierbar, und für $(x_0, y_0) \in \mathbb{R}^2$ mit $y_0 \neq 0$ und $f(x_0, y_0) = 0$ können wir lokal nach y auflösen mit einer differenzierbaren Funktion $y(x)$, für die

$$\begin{aligned} Dy(x) &= y'(x) = -(D_2 f(x, y(x)))^{-1} \circ D_1 f(x, y(x)) \\ &= -(2y(x))^{-1} \cdot (-2x + 4x^3) = \frac{x - 2x^3}{y(x)} \end{aligned}$$

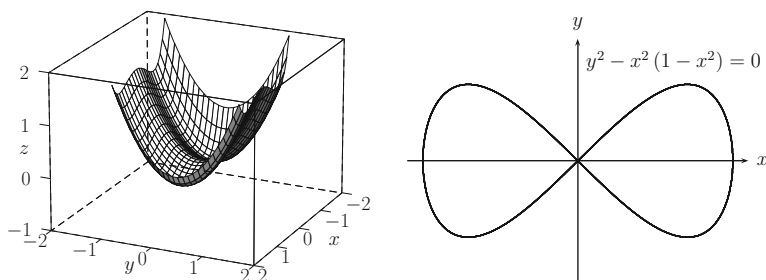


Abb. 5.4: Die Funktion und die Nullniveaulinie zu $f(x, y) = y^2 - x^2(1 - x^2)$.

gilt. In der Tat folgt aus $f(x, y) = 0$ die Gleichung

$$y^2(x) = x^2(1 - x^2)$$

(man kann also sogar explizit lösen) und hieraus nach der Kettenregel $2y(x)y'(x) = 2x - 4x^3$ und für $y(x) \neq 0$ dann auch

$$y'(x) = \frac{x - 2x^3}{y(x)}$$

als Bestätigung der obigen Formel.

■ Sei (siehe Abbildung 5.5)

$$f(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = (x^2 + y^2 - 1)^3 + 27x^2y^2.$$

Hier ist

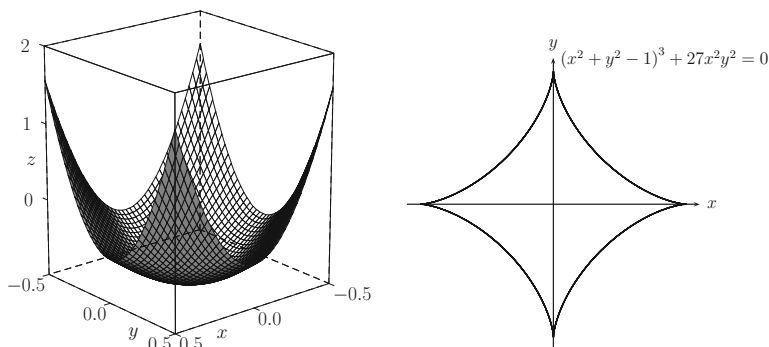


Abb. 5.5: Die Funktion und die Nullniveaulinie zu $f(x, y) = (x^2 + y^2 - 1)^3 + 27x^2y^2$.

$$D_2f(x, y) = \frac{\partial f}{\partial y}(x, y) = 6y((x^2 + y^2 - 1)^2 + 9x^2).$$

Dies verschwindet genau dann, wenn entweder $y = 0$ oder wenn $x = 0$ und $|y| = 1$. Die einzigen Punkte $(x_0, y_0) \in \mathbb{R}^2$, für die $f(x_0, y_0) = 0$ und $D_2f(x_0, y_0) = 0$ gilt, sind demnach die Punkte der Menge

$$\{(-1, 0), (1, 0), (0, -1), (0, 1)\}.$$

In allen anderen Punkten $(x_0, y_0) \in f^{-1}\{0\}$ existiert eine eindeutige Lösung von $f(x, y(x)) = 0$. Wegen

$$D_1 f(x, y) = 6x((x^2 + y^2 - 1)^2 + 9y^2)$$

ist dann

$$y'(x) = -\frac{x((x^2 + y(x)^2 - 1)^2 + 9y(x)^2)}{y(x)((x^2 + y(x)^2 - 1)^2 + 9x^2)}.$$

Wie schon in der Anmerkung zum Satz können wir allein anhand des Satzes nicht über die Auflösbarkeit in den Punkten $(0, 1)$ und $(0, -1)$ entscheiden. Hier ist es aber so, dass in einer Umgebung von $(0, 1)$ und $(0, -1)$ auch eine eindeutige Lösung $y(x)$ von $f(x, y(x)) = 0$ existiert, diese ist nur nicht differenzierbar.

- Wir haben bisher nur Gleichungen in zwei Variablen untersucht. Was ist, wenn es mehrere sind? Wir betrachten die Funktion

$$f : \mathbb{R}^4 \rightarrow \mathbb{R}^2, \quad f(w, x, y, z) := (w^2 + xy + e^z, 2w + y^2 - yz - 5).$$

Hier müssen wir leider auf ein Bild zur Veranschaulichung verzichten. Wir untersuchen den Punkt $(2, 5, -1, 0)$ und wollen nach (y, z) auflösen, das heißt, wir suchen eine Funktion $g(w, x) = (y(w, x), z(w, x))$. Es gilt:

$$f(2, 5, -1, 0) = (4 - 5 + 1, 2 + 1 - 0 - 5) = (0, 0).$$

und

$$J_f(w, x, y, z) = \begin{pmatrix} 2w & y & x & e^z \\ 2 & 0 & 2y - z & -y \end{pmatrix},$$

also

$$J_f(2, 5, -1, 0) = \begin{pmatrix} 4 & -1 & 5 & 1 \\ 2 & 0 & -2 & 1 \end{pmatrix}.$$

D_2 bedeutet hier einfach das Differential nach den Variablen, nach denen wir auflösen. Demnach gilt:

$$D_2(2, 5, -1, 0) = \begin{pmatrix} 5 & 1 \\ -2 & 1 \end{pmatrix}.$$

Wegen $\det(D_2(2, 5, -1, 0)) = 7 \neq 0$ ist dies invertierbar, und damit existiert die gewünschte lokale Auflösung. Wir können außerdem nach der bewiesenen Formel die Jakobi-Matrix von g an der Stelle $(2, 5)$ bestimmen, es gilt:

$$J_g(2, 5) = -\begin{pmatrix} 5 & 1 \\ -2 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 4 & -1 \\ 2 & 0 \end{pmatrix} = \frac{1}{7} \begin{pmatrix} 2 & -1 \\ -18 & -2 \end{pmatrix}.$$

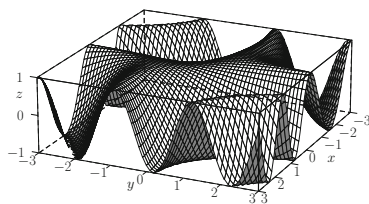


Abb. 5.6: Die Funktion $f(x, y) = \sin(\cos(x) - xy)$.

- Auch das schöne Bild auf unserem Cover (siehe auch Abbildung 5.6), dargestellt durch die Funktion

$$f(x, y) = \sin(\cos(x) - xy),$$

ist in jedem Punkt implizit auflösbar.

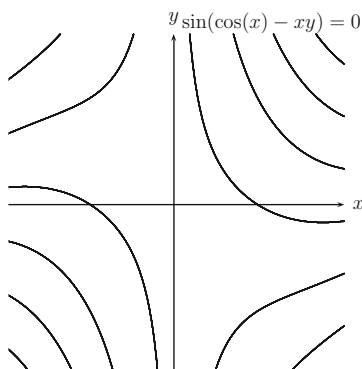


Abb. 5.7: Die Niveaulinien zur Funktion $f(x, y) = \sin(\cos(x) - xy) = 0$.

Natürlich erhalten wir dann aber für die Ableitung einen sehr komplizierten Term.

■

Wir haben also gesehen, dass wir Funktionen unter bestimmten Umständen lokal auflösen und eine Formel für die Ableitung erhalten können. Wie wir auch in zwei Beispielen gesehen haben, ist diese Formel eine Gleichung, die sowohl eine Funktion als auch deren Ableitung enthält. Es stellt sich nun die Frage, wie man eine solche Gleichung lösen kann. Dies wollen wir im nächsten Kapitel behandeln.

6 Gewöhnliche Differentialgleichungen

Übersicht

6.1	Definitionen	117
6.2	Sätze und Beweise	119
6.3	Erklärungen zu den Definitionen	124
6.4	Erklärungen zu den Sätzen und Beweisen	127

Wir wollen uns nun mit Gleichungen beschäftigen, in der sowohl eine Funktion als auch ihre Ableitungen vorkommen. Dies wird in vielen mathematischen Bereichen und in den Anwendungen der Physik, Technik, Biologie, Chemie und Wirtschaftswissenschaft gebraucht. Mit einfachen Formen solcher Differentialgleichungen wollen wir uns nun beschäftigen.

6.1 Definitionen

Definition 6.1 (Differentialgleichung)

Seien $I \subset \mathbb{R}$ ein Intervall und $k \in \mathbb{N}$. Eine **gewöhnliche Differentialgleichung** der Ordnung k ist eine Gleichung von der Form

$$F(t, x(t), \dots, x^{(k)}(t)) = 0.$$

Dabei ist F eine gegebene Funktion auf einer Teilmenge U von $I \times \mathbb{R}^n \times \dots \times \mathbb{R}^n$ mit Werten in \mathbb{R}^m .

Die Funktion $x : I \rightarrow \mathbb{R}^n$ heißt Lösung der Differentialgleichung, falls gilt:

- x ist k -mal differenzierbar auf I .
- $(t, x(t), \dots, x^{(k)}(t)) \in U$.
- $F(t, x(t), \dots, x^{(k)}(t)) = 0, t \in I$.

Definition 6.2 (explizit, homogen, autonom)

Eine Differentialgleichung heißt **explizit**, falls sie in der Form

$$x^{(k)} = f(t, x, \dots, x^{(k-1)})$$

gegeben ist, andernfalls **implizit**. Eine Differentialgleichung der Form

$$\sum_{j=0}^k A_j(t)x^{(j)} + f(t) = 0$$

mit linearen A_j heißt linear von der Ordnung k . Sie heißt **homogen**, falls $f \equiv 0$, sonst **inhomogen**. Ist $F(t, x, \dots, x^{(k)}) = 0$ nicht explizit von t abhängig, so heißt die Differentialgleichung **autonom**.

Definition 6.3 (Anfangswertaufgabe)

Ist $x^{(k)} = f(t, x, \dots, x^{(k-1)})$ eine Differentialgleichung, $t_0 \in I$, so besteht eine **Anfangswertaufgabe** im Finden einer Lösung x der Differentialgleichung, sodass außerdem die ersten $k - 1$ Ableitungen von x in t_0 die vorgegebenen Anfangswerte $x(t_0) = c_0, \dots, x^{(k-1)}(t_0) = c_{k-1}$ annehmen.

Definition 6.4 (Nullraum)

Seien J ein offenes Intervall in \mathbb{R} , $x : J \rightarrow \mathbb{R}^n$ eine differenzierbare Funktion und A eine auf J stetige matrixwertige Funktion, also eine Funktion, die jedem $t \in J$ eine Matrix aus $\mathbb{R}^{n \times n}$ zuordnet. Dann bezeichnen wir die Menge aller Lösungen von

$$x' - A(t)x = 0$$

mit \mathcal{N}_A und nennen dies den **Nullraum** der Differentialgleichung.

Anmerkung: Satz 6.3 zeigt, dass dies ein Vektorraum ist.

Definition 6.5 (Fundamentalsystem)

Seien $J \subset \mathbb{R}$ ein Intervall und A eine stetige, matrixwertige Funktion auf J . Eine Basis $\{x^1, \dots, x^n\}$ von \mathcal{N}_A nennt man ein **Fundamentalsystem** für die Differentialgleichung $x' = A(t)x$. Meistens nennt man auch die Matrix

$$\Phi := \begin{pmatrix} x_1^1 & \cdots & x_1^n \\ \vdots & & \vdots \\ x_n^1 & \cdots & x_n^n \end{pmatrix}$$

ein **Fundamentalsystem**.

Anmerkung: Mit x^k meinen wir in diesem Fall keine Potenzen. Jedes x^k ist eine Funktion von J nach \mathbb{R}^n , die wiederum aus n Komponenten x_1^k, \dots, x_n^k besteht.

6.2 Sätze und Beweise

Satz 6.1 (Satz von Peano)

Es seien $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^n$, $a > 0$, $b > 0$. Setze

$$D = \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : |t - t_0| \leq a, \|x - x_0\| \leq b\}.$$

Ist $f = f(t, x) : D \rightarrow \mathbb{R}^n$ stetig, so hat die Anfangswertaufgabe

$$x' = f(t, x), x(t_0) = x_0$$

eine Lösung auf dem Intervall $(t_0 - c, t_0 + c)$, wobei $c = \min\{a, \frac{b}{A}\}$ und $A = \max\{\|f(t, x)\| : (t, x) \in D\}$.

Satz 6.2 (Satz von Picard-Lindelöf)

Seien wieder $(t_0, x_0) \in \mathbb{R} \times \mathbb{R}^n$, $a > 0$, $b > 0$ und

$$D = \{(t, x) \in \mathbb{R} \times \mathbb{R}^n : |t - t_0| \leq a, \|x - x_0\| \leq b\}.$$

Ist nun $f = f(t, x) : D \rightarrow \mathbb{R}^n$ sogar Lipschitz-stetig in x mit Lipschitz-Konstante $L \geq 0$, dann hat die Anfangswertaufgabe $x' = f(t, x)$, $x(t_0) = x_0$ eine eindeutige Lösung auf dem Intervall $(t_0 - d, t_0 + d)$, wobei $d = \min\{a, \frac{b}{A}, \frac{1}{L}\}$.

Beweis: Es sei $I_0 = [t_0 - r, t_0 + r]$ für ein $0 < r < d$. Wir integrieren zunächst und sehen, dass x genau dann die Anfangswertaufgabe $x' = f(t, x)$, $x(t_0) = x_0$ auf I_0 löst, wenn x auf I_0 stetig ist und

$$x(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds, \quad t \in I_0$$

gilt. Wir wollen den Banach'schen Fixpunktsatz anwenden. Wir wählen als Banach-Raum $X = \mathcal{C}(I_0, \mathbb{R}^n)$ mit der Supremumsnorm. Darauf definieren wir die Abbildung T durch:

$$Tx(t) = x_0 + \int_{t_0}^t f(s, x(s)) ds, \quad x \in X, t \in I_0.$$

Nun ist eine Funktion x genau dann Lösung der Anfangswertaufgabe, wenn sie ein Fixpunkt der Abbildung T ist. Wir wählen als abgeschlossene Teilmenge von X die Menge

$$M = \{x \in X : \|x - x_0\| \leq b\}$$

und prüfen die Voraussetzungen des Banach'schen Fixpunktsatzes nach.

Für $x \in X$ ist die Funktion Tx stetig auf I_0 , da $\|Tx(t_1) - Tx(t_2)\| \leq A|t_1 - t_2|$. Also ist $Tx \in X$. Ferner ist für $t \in I_0$

$$\|Tx(t) - x_0\| = \left\| \int_{t_0}^t f(s, x(s)) ds \right\| \leq A|t - t_0| \leq Ar \leq b,$$

somit ist $\|Tx - x_0\| \leq b$ und T eine Abbildung von M nach M .

Für $x, y \in X$ und $t \geq t_0$ gilt nun:

$$\begin{aligned} \|Tx(t) - Ty(t)\| &= \left\| \int_{t_0}^t f(s, x(s)) - f(s, y(s)) ds \right\| \\ &\leq \int_{t_0}^t \|f(s, x(s)) - f(s, y(s))\| ds \\ &\leq \int_{t_0}^t L \|x - y\| ds \leq rL \|x - y\|. \end{aligned}$$

Dieselbe Abschätzung erhält man auch für $t \leq t_0$. Damit ist

$$\|Tx - Ty\| \leq rL \|x - y\|.$$

Weil $rL < dL \leq 1$ ist, ist T kontrahierend und besitzt nach dem Banach'schen Fixpunktsatz genau einen Fixpunkt. Also hat das Anfangswertproblem genau eine Lösung. q.e.d.

Satz 6.3

Seien $J \subset \mathbb{R}$ ein offenes Intervall und $A : J \rightarrow \mathbb{R}^{n \times n}$ stetig. Dann bildet die Menge aller Lösungen von

$$x' - A(t)x = 0$$

einen Vektorraum.

Beweis: Übungsaufgabe für euch.

q.e.d.

Satz 6.4

Sei $J \subset \mathbb{R}$ ein offenes Intervall. Sind $f \in \mathcal{N}_A$ und $f(t_0) = 0$ für ein $t_0 \in J$, so ist $f(t) = 0$ für alle t .

Beweis: Folgt direkt aus dem Satz von Picard-Lindelöf, denn gilt $f(t_0) = 0$ für ein t_0 , so ist die Funktion $f \equiv 0$ eine Lösung und wegen der Eindeutigkeit dann auch die einzige. q.e.d.

Satz 6.5

Es gilt $\dim \mathcal{N}_A = n$, das heißt, die Lösungen von $x' - A(t)x = 0$ bilden einen n -dimensionalen Vektorraum.

Beweis: Wir betrachten die Abbildung, die einer Funktion aus \mathcal{N}_A den Wert $x(t_0) \in \mathbb{R}^n$ zuordnet. Diese ist linear. Weil nach dem Satz von Picard-Lindelöf für jeden Wert $x(t_0)$ eine zugehörige Lösung x existiert, ist sie surjektiv. Weil diese Lösung sogar eindeutig ist, ist sie auch injektiv, und damit ist diese Abbildung ein Isomorphismus, also gilt $\mathcal{N}_A \cong \mathbb{R}^n$. q.e.d.

Satz 6.6

Seien $J \subset \mathbb{R}$ ein offenes Intervall und $x^1, \dots, x^n \in \mathcal{N}_A$. Sei

$$D(t) := \det \begin{pmatrix} x_1^1(t) & \cdots & x_1^n(t) \\ \vdots & & \vdots \\ x_n^1(t) & \cdots & x_n^n(t) \end{pmatrix}.$$

Dann sind äquivalent:

- i) $D(t_0) = 0$ für ein $t_0 \in J$.
- ii) $D(t) = 0$ für alle $t \in J$.
- iii) $x^1(t_0), \dots, x^n(t_0)$ sind als Vektoren im \mathbb{R}^n linear abhängig für ein $t_0 \in J$.
- iv) $x^1(t), \dots, x^n(t)$ sind als Vektoren im \mathbb{R}^n linear abhängig für alle $t \in J$.
- v) x^1, \dots, x^n sind linear abhängig als Funktionen.

Anmerkung: $D(t)$ nennt man auch Wronski-Determinante.

Beweis: Die Implikationen $v) \Rightarrow iv) \Rightarrow iii)$ und die Äquivalenzen $i) \Leftrightarrow iii)$ und $ii) \Leftrightarrow iv)$ sollten klar sein. Es reicht deshalb zu zeigen, dass die dritte Bedingung die fünfte impliziert.

Es existieren also nach Annahme $c_1, \dots, c_n \in \mathbb{R}$, die nicht alle 0 sind, mit

$$\sum_{j=1}^n c_j x^j(t_0) = 0.$$

Wir definieren dann eine neue Funktion x durch

$$x := \sum_{j=1}^n c_j x^j.$$

Dann ist $x \in \mathcal{N}_A$ und $x(t_0) = 0$, also gilt $x(t) = 0$ für alle t , und das bedeutet, dass die Funktionen x^1, \dots, x^n linear abhängig sind. q.e.d.

Satz 6.7

Φ ist genau dann ein Fundamentalsystem, wenn $\Phi' = A\Phi$ und $\det \Phi(t) \neq 0$ für ein $t \in J$ gilt.

Beweis: Folgt direkt aus Satz 6.6.

q.e.d.

Satz 6.8 (Lösung der homogenen Gleichung)

Sei Φ ein Fundamentalsystem für $x' = A(t)x$. Dann gilt:

- i) $x(t) = \Phi(t)\Phi(t_0)^{-1}x_0$ ist die Lösung des Anfangswertproblems $x' = A(t)x, x(t_0) = x_0$.
- ii) Eine matrixwertige Funktion Ψ ist genau dann ein weiteres Fundamentalsystem, wenn es eine invertierbare Matrix C gibt mit

$$\Psi(t) = \Phi(t)C.$$

Beweis:

- i) Wir setzen $c = \Phi(t_0)^{-1}x_0$. Dann ist

$$x'(t) = (\Phi(t)c)' = \Phi'(t)c = A(t)\Phi(t)c = A(t)x(t) \quad (6.1)$$

und $x(t_0) = x_0$.

- ii) Angenommen, es gibt eine solche Matrix C mit $\Phi(t)C = \Psi(t)$. Dann ist

$$(\Phi(t)C)' = \Phi'(t)C = A(t)\Phi(t)C,$$

und wegen $\det(\Phi C) = \det \phi \det C$ folgt die Aussage dann aus dem letzten Satz.

Ist andererseits Ψ ein weiteres Fundamentalsystem, so setzen wir $C(t) := \Phi(t)^{-1}\Psi(t)$. Dann folgt aus der Produktregel und weil Φ und Ψ Fundamentalsysteme sind

$$0 = \Psi' - A\Psi = \Phi'C + \Phi C' - A\Phi C = \Phi C'.$$

Aus der Invertierbarkeit von Φ folgt $C'(t) = 0$ für alle t , somit ist C also konstant.

q.e.d.

Satz 6.9 (Lösung der inhomogenen Gleichung)

Es sei Φ ein Fundamentalsystem für $x' = A(t)x$. Dann hat das Anfangswertproblem $x' = A(t)x + f, x(t_0) = x_0$ die Lösung

$$x(t) = \Phi(t)(\Phi(t_0)^{-1}x_0 + \int_{t_0}^t \Phi(s)^{-1}f(s)ds).$$

Beweis: Es gilt:

$$\begin{aligned} x'(t) &= A(t)\Phi(t)(\Phi(t_0)^{-1}x_0 + \int_{t_0}^t \Phi(s)^{-1}f(s)ds) + \Phi(t)(\Phi(t)^{-1}f(t)) \\ &= A(t)x(t) + f(t) \end{aligned}$$

und $x(t_0) = x_0$.

q.e.d.

Satz 6.10

Sei $A \in \mathbb{R}^{n \times n}$, also eine konstante Matrix. Sei $J = T^{-1}AT$ die Jordan-Normalform. Dann sind e^{tA} und Te^{tJ} Fundamentalsysteme für $x' = Ax$.

Beweis: Da e^{tA} für jedes $t \in \mathbb{R}$ invertierbar ist, verschwindet die Determinante nicht. Außerdem gilt $(e^{tA})' = Ae^{tA}$, also ist e^{tA} ein Fundamentalsystem. Weil T invertierbar ist, ist wegen $Te^{tJ} = e^{tA}T$ dann auch Te^{tJ} ein Fundamentalsystem. q.e.d.

Satz 6.11 (einige Lösungsverfahren)

i) *Separation der Variablen:* Eine Lösung der Differentialgleichung

$$x' = f(t)g(x), x(t_0) = x_0, g(x_0) \neq 0$$

ist gegeben durch

$$\int_{t_0}^t f(s)ds = \int_{t_0}^t \frac{x'(s)}{g(x(s))}ds = \int_{x(t_0)}^{x(t)} \frac{dy}{g(y)}.$$

ii) *Variation der Konstanten:* Ist $x_0 \neq 0$, so ist eine Lösung von

$$x' + a(t)x = 0, x(t_0) = x_0$$

gegeben durch

$$x(t) = x_0 e^{-\int_{t_0}^t a(s)ds}$$

und die von

$$x' + a(t)x = f(t), x(t_0) = x_0$$

durch

$$x(t) = \left(\int_{t_0}^t f(r) e^{\int_{t_0}^r a(s)ds} dr + x_0 \right) e^{-\int_{t_0}^t a(s)ds}.$$

Beweis:

- i) Falls eine Lösung existiert, so ist wegen der Stetigkeit von g auch $g(x(t)) \neq 0$ für t nahe t_0 . Dann gilt dort:

$$\int_{t_0}^t f(s)ds = \int_{t_0}^t \frac{x'(s)}{g(x(s))}ds = \int_{x(t_0)}^{x(t)} \frac{dy}{g(y)}.$$

Dies liefert eine implizite Gleichung der Form $G(t, x) = 0$. Wenn diese auflösbar ist, ist die Gleichung lösbar. Nun gilt:

$$\frac{\partial G}{\partial x}(t, x) = \frac{\partial}{\partial x} \left(\int_{t_0}^t f(s)ds - \int_{x_0}^x \frac{dy}{g(y)} \right) = -\frac{1}{g(x)} \neq 0.$$

Nach dem Satz über implizite Funktionen ist die Gleichung also auflösbar, also ist durch obige Formel eine Lösung gegeben.

ii) Wir betrachten zunächst die homogene Differentialgleichung.

Ist $x_0 = 0$, so ist $x(t) \equiv 0$ eine Lösung. Ist $x_0 \neq 0$, so gilt wegen der Stetigkeit zumindest für t nahe t_0

$$\frac{x(t)}{x_0} > 0$$

und damit folgt

$$\ln \frac{x}{x_0} = - \int_{t_0}^t a(s) ds \Rightarrow x(t) = x_0 e^{-\int_{t_0}^t a(s) ds}.$$

Somit existiert die Lösung für alle $t \in \mathbb{R}$, und $x(t) \neq 0$ für alle $t \in \mathbb{R}$.

Davon ausgehend wollen wir nun die inhomogene Differentialgleichung betrachten. Der Ansatz hierfür ist die sogenannte **Variation der Konstanten**. Wir setzen

$$F(t) = e^{-\int_{t_0}^t a(s) ds}.$$

Dann ist $F(t) \neq 0$ für alle t und $F(t_0) = 1$. Wir machen den Ansatz $x(t) = C(t)F(t)$. Es folgt $C(t_0) = x(t_0) = x_0$. Ferner ergibt sich

$$f(t) = x' + a(t)x = C'(t)F(t) + C(t)(F'(t) + a(t)F(t)) = C'(t)F(t),$$

denn F löst die homogene Gleichung. Also gilt:

$$\begin{aligned} C'(t) &= \frac{f(t)}{F(t)} = f(t)e^{\int_{t_0}^t a(s) ds}, \\ C(t) &= \int_{t_0}^t f(r)e^{\int_{t_0}^r a(s) ds} dr + C(t_0), \\ x(t) &= \left(\int_{t_0}^t f(r)e^{\int_{t_0}^r a(s) ds} dr + x_0 \right) e^{-\int_{t_0}^t a(s) ds}. \end{aligned}$$

q.e.d.

6.3 Erklärungen zu den Definitionen

Zur Definition 6.1 einer Differentialgleichung: Diese Definition hört sich zunächst einmal kompliziert an, ist sie aber eigentlich nicht. Es ist zum Beispiel $5t^3x + \frac{4xt}{x'}x^{(3)} - \cos(tx') = 7, t \in \mathbb{R}$ eine Differentialgleichung der Ordnung 3 auf \mathbb{R} . Eine Differentialgleichung ist also einfach eine Gleichung, in der eine Funktion, deren Ableitungen, die Variable, von der die Funktion abhängt, und Konstanten vorkommen. Die Ordnung bezeichnet dabei die höchste Ableitung, die vorkommt. Man schreibt dabei, wie zum Beispiel oben, meist x für die

Funktion, meint aber $x(t)$, sie hängt also von t ab. Man spricht auch von einem System von m Differentialgleichungen für die n Komponenten x_1, \dots, x_n von x . Ein Beispiel hierfür wäre die Differentialgleichung

$$\begin{aligned}x_1'' &= x_2, \\x_2'' &= x_1.\end{aligned}$$

Dies schreibt man auch oft als

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

Ist $n = 1$, so spricht man von einer skalaren Differentialgleichung (wie im Beispiel weiter oben).

Solche Gleichungen zu lösen, wird also unser Ziel in diesem Kapitel sein.

Dabei kann das Lösen solcher Gleichungen beliebig kompliziert werden. Könnt ihr zum Beispiel auf Anhieb eine Lösung für die obige Differentialgleichung erkennen? Es gibt aber auch sehr einfache Differentialgleichungen, zum Beispiel $x' = x$. Hier ist eine Lösung gegeben durch $x(t) = e^t$. Nun könnte man natürlich fragen, ob es noch andere Lösungen gibt. Auch damit wollen wir uns hier beschäftigen.

Zur Definition 6.2 von explizit, homogen und autonom: Diese Definitionen sind fast selbsterklärend. Eine explizite Differentialgleichung ist also einfach eine Gleichung, die nach der höchsten Ableitung aufgelöst ist. Außerdem ist zum Beispiel $x''' + x' + e^t = 0$ eine inhomogene lineare Differentialgleichung und $x' - ax = 0$ eine homogene autonome lineare Differentialgleichung.

Beispiel 57

Wir wollen nun einmal, ohne uns zunächst näher mit der Theorie von Differentialgleichungen zu beschäftigen, versuchen, eine zu lösen. Wir betrachten dafür ein physikalisch motiviertes Beispiel, das viele von euch aus der Schule noch kennen sollten, die Bewegung eines Massepunktes unter dem Einfluss der Schwerkraft. Wie noch bekannt sein sollte, beträgt der zurückgelegte Weg bei Anfangsgeschwindigkeit 0 und wenn zum Startzeitpunkt noch kein Weg zurückgelegt ist, gerade $\frac{1}{2}gt^2$, wobei g die Erdbeschleunigung $g \approx 9.8 \text{ m/s}^2$ bezeichnet. Dies wollen wir nun versuchen zu beweisen. Wir wissen etwas über die Beschleunigung, das heißt über die zweite Ableitung, nämlich

$$x''(t) = -g, t \in \mathbb{R},$$

wobei wir hier das Vorzeichen „-“ wählen, da es sich um eine Bewegung nach unten handelt. Für festes $t_0 \in \mathbb{R}$ liefert der Hauptsatz der Differential- und Integralrechnung:

$$x'(t) - x'(t_0) = \int_{t_0}^t x''(s) ds = - \int_{t_0}^t g ds = -g(t - t_0),$$

also $x'(t) = -gt + c$, $c = gt_0 + x'(t_0)$ und

$$x(t) - x(t_0) = \int_{t_0}^t x'(s) ds = \int_{t_0}^t (-gs + c) ds = -\frac{1}{2}g(t^2 - t_0^2) + c(t - t_0),$$

das heißt,

$$x(t) = -\frac{1}{2}gt^2 + ct + d \text{ mit } d = x(t_0) + \frac{1}{2}gt_0^2 - x'(t_0)t_0 - gt_0^2.$$

Also hat die Lösung der Differentialgleichung die Form $x(t) = -\frac{1}{2}gt^2 + ct + d$, wobei c und d sich aus den Startwerten $x(t_0)$ und $x'(t_0)$ berechnen lassen. Dabei erhalten wir hier das Minuszeichen vor $\frac{1}{2}gt^2$, weil wir in der Anfangsgleichung das Minuszeichen gewählt haben. Was ist aber nun mit den anderen beiden Termen? Das klären wir in der nächsten Erklärung. ■

Zur Definition 6.3 eines Anfangswertproblems: Wir haben im Beispiel eben gerade gesehen, dass die Differentialgleichung keine eindeutige Lösung hatte. Dies lag daran, dass wir keine Anfangswerte vorgegeben hatten. Daher kamen auch die Terme ct und d in der Lösung. Betrachten wir nun im Beispiel oben den Anfangszeitpunkt $t_0 = 0$, nehmen an, dass zu diesem Zeitpunkt kein Weg zurückgelegt ist, das heißt $x(0) = 0$, und dass wir mit Geschwindigkeit 0 starten, also $x'(0) = 0$, so erhalten wir für die Konstanten $c = 0, d = 0$, also genau die erwartete Lösung.

Wir sehen also schon einmal, dass eine Differentialgleichung im Allgemeinen unendlich viele Lösungen hat, es sei denn man gibt Anfangswerte vor. Dazu noch ein Beispiel.

Beispiel 58

Die Anfangswertaufgabe $x' = ax, x(t_0) = x_0$ hat für jede Vorgabe von t_0, x_0 eine eindeutige Lösung, nämlich

$$x(t) = x_0 e^{a(t-t_0)}.$$

Dass dies tatsächlich eine Lösung ist, rechnet man leicht nach, dass dies die einzige Lösung ist, folgt sofort aus dem Satz von Picard-Lindelöf 6.2. ■

Zur Definition 6.4 des Nullraums: Dieser Raum wird uns bei der Lösung von homogenen linearen Differentialgleichungen in höheren Dimensionen begegnen. Wieso? Nun, ein x ist ja genau dann in \mathcal{N}_A , wenn $x' = A(t)x$ gilt, also ist x Lösung einer homogenen linearen Differentialgleichung.

6.4 Erklärungen zu den Sätzen und Beweisen

Allgemeine Anmerkung: Die ersten beiden Sätze behandeln jeweils nur Differentialgleichungen erster Ordnung. Wieso? Ist

$$x^{(k)} = g(t, x, \dots, x^{(k-1)})$$

eine explizite Differentialgleichung, so können wir

$$x_1 = x, x_2 = x', \dots, x_k = x^{(k-1)}$$

setzen. Dann ist das Lösen von $x^{(k)} = f(t, x, \dots, x^{(k-1)})$ mit den Anfangswerten $x(t_0) = c_0, \dots, x^{(k-1)}(t_0) = c_{k-1}$ äquivalent zum Lösen von

$$x'_1 = x_2, \dots, x'_{k-1} = x_k, x'_k = g(t, x_1, \dots, x_k)$$

mit den Anfangswerten $x_1(t_0) = c_0, \dots, x_k(t_0) = c_{k-1}$. Fasst man x_1, \dots, x_k als Vektor x auf, so lautet das letzte System

$$x' = f(t, x), x(t_0) = c,$$

mit

$$f(t, x) = (x_2, \dots, x_k, g(t, x_1, \dots, x_k)) \text{ und } c = (c_0, \dots, c_{k-1}).$$

Das heißt, jede explizite Differentialgleichung ist äquivalent zu einem System 1. Ordnung. Es genügt also, Anfangswertaufgaben der Form $x' = f(t, x), x(t_0) = x_0$ zu studieren.

Zum Satz von Peano (Satz 6.1): Dieser Satz sichert uns nun schon einmal unter gewissen Umständen die Existenz einer Lösung. Dummerweise kann es sein, dass es mehrere Lösungen gibt, und dies ist für die Anwendung unpraktisch. Deshalb wollen wir uns mit dem Satz und dem Beweis nicht weiter beschäftigen und behandeln lieber den Satz von Picard-Lindelöf.

Zum Satz von Picard-Lindelöf (Satz 6.2): Dieser Satz enthält nun eine wesentlich schönere Aussage, nämlich nicht nur die Existenz, sondern auch die Eindeutigkeit. Damit macht dieser Satz auch numerische Berechnungen möglich, was wir hier allerdings nicht vertiefen wollen. Wie schon öfters in diesem Buch wird dabei der Banach'sche Fixpunktsatz als Beweismethode benutzt.

Wir merken noch an, dass die Lipschitz-Bedingung stets erfüllt ist, wenn die Funktion f auf einer Umgebung von D nach x stetig differenzierbar ist. Wegen der Kompaktheit von D ist dann nämlich $\|\partial_x f(t, x)\|$ beschränkt auf D und somit

$$\|f(t, x_1) - f(t, x_2)\| \leq \sup\{\|\partial_x f(t, x)\|\} \|x_1 - x_2\|.$$

Deshalb wird man diese Bedingung häufig durch die Differenzierbarkeit begründen können. Die Einschränkung $d \leq \frac{1}{L}$ ist außerdem nicht nötig, sie vereinfacht aber den Beweis.

Beispiel 59

In einigen Fällen ist es möglich, die Lösung von $x' = f(t, x)$ durch das im Beweis des Satzes von Picard-Lindelöf implizit benutzte Iterationsverfahren explizit zu bestimmen. Denn schließlich verwendet der Banach'sche Fixpunktsatz ja Funktionenfolgen.

Wir betrachten die Gleichung

$$x' = kt^{k-1}x$$

mit einem $k \in \mathbb{N}$. Sei $x_0 \in \mathbb{R}$ beliebig und $t_0 := 0$. Wir suchen also nach der eindeutigen Lösung zur Anfangsbedingung $x(0) = x_0$. Es sei $x_{(0)} := x_0$ und $x_{(n+1)}$ iterativ durch

$$x_{(n+1)}(t) := (Tx_{(n)})(t) = x_0 + \int_{t_0}^t f(s, x_{(n)}(s))ds = x_0 + \int_0^t ks^{k-1}x_{(n)}(s)ds$$

definiert. Nach dem Beweis des Satzes von Picard-Lindelöf ist nun die eindeutige Lösung der Differentialgleichung gerade der Grenzwert von $x_{(n)}$. Diesen wollen wir also bestimmen.

Wir erhalten durch direktes Integrieren

$$x_{(1)}(t) = x_0 + \int_0^t ks^{k-1}x_0ds = x_0(1 + t^k)$$

und

$$\begin{aligned} x_{(2)}(t) &= x_0 + \int_0^t ks^{k-1}x_0(1 + s^k)ds \\ &= x_0 + \int_0^t ks^{k-1}x_0ds + \int_0^t ks^{2k-1}x_0ds \\ &= x_0 + t^k x_0 + \frac{1}{2}t^{2k}x_0 = x_0\left(1 + t^k + \frac{(t^k)^2}{2}\right). \end{aligned}$$

Per Induktion beweist man

$$x_{(n)}(t) = x_0 \sum_{l=0}^n \frac{(t^k)^l}{l!},$$

also

$$x(t) = \lim_{n \rightarrow \infty} x_{(n)}(t) = x_0 \sum_{l=0}^{\infty} \frac{(t^k)^l}{l!} = x_0 e^{t^k}.$$

In der Tat ist

$$x'(t) = x_0 kt^{k-1} e^{t^k} = kt^{k-1} x(t), x(0) = x_0.$$

■

Wie das folgende Beispiel zeigt, kann man nicht auf die Lipschitz-Bedingung im zweiten Argument von f verzichten.

Beispiel 60

Wir betrachten die Funktion

$$f(t, x) = \sqrt{2|x|}.$$

Die gewöhnliche Differentialgleichung $x'(t) = \sqrt{2|x(t)|}$ besitzt zur Anfangsbedingung $x(0) = 0$ die beiden Lösungen $x_1(t) = 0$ und $x_2(t) = \frac{\operatorname{sgn}(t)t^2}{2}$, denn es ist $x_1(0) = x_2(0) = 0$ und

$$x_1'(t) = 0 = \sqrt{0} = \sqrt{2|x|}, \quad x_2'(t) = \operatorname{sgn}(t)t = |t| = \sqrt{2|x|}.$$

■

Zu Lösungsverfahren linearer Differentialgleichungen in höheren Dimensionen

(Satz 6.3–6.9): In diesen Sätzen beschäftigen wir uns nun mit dem Lösen einer linearen Differentialgleichung in höheren Dimensionen. Da diese Sätze zusammen die Richtigkeit der Methode beschreiben, wollen wir sie hier zusammen kommentieren.

Zunächst bildet die Menge der Lösungen der homogenen Gleichung einen Vektorraum, und wir treffen eine Aussage darüber, wann eine Funktion f dort die Nullfunktion ist, nämlich schon dann, wenn sie an einer Stelle den Wert 0 annimmt.

Und wenn wir einen Vektorraum gegeben haben, dann interessiert uns dessen Dimension; diese bestimmen wir dann in Satz 6.5.

Im nächsten Satz treffen wir nun einige Aussagen darüber, wann n Funktionen in diesem Vektorraum linear unabhängig sind. Dies ist wichtig, weil wir später beim Lösen eine Basis, das heißt n linear unabhängige Funktionen dieses Vektorraums benötigen.

Haben wir dies getan, so können wir nun in Definition 6.5 das sogenannte Fundamentalsystem definieren, das wir aus einer Basis von \mathcal{N}_A gewinnen.

In Satz 6.7 treffen wir nun noch eine Aussage, wann die Spalten einer Matrix ein Fundamentalsystem darstellen. Diese Charakterisierung benötigen wir dann im nächsten Satz, in dem wir nun endlich die homogene Differentialgleichung lösen können.

Wie schon im eindimensionalen Fall sieht die Lösung der inhomogenen Gleichung ein wenig komplizierter aus. Allerdings wird sie ganz analog zum eindimensionalen Fall aus homogener Lösung und spezieller inhomogener Lösung zusammengesetzt.

Zu Lösungsverfahren linearer Differentialgleichung in höheren Dimensionen im Falle einer konstanten Matrix (Satz 6.10):

Um nun die für uns wichtige Differentialgleichung mit einer konstanten Matrix lösen zu können, brauchen wir nur ein Fundamentalsystem. Dieses bekommen wir gerade durch Satz 6.10, wir müssen nur die Jordan-Form kennen.

Beispiel 61

Wir suchen nach Lösungen der Gleichung

$$x' = Ax, x(t_0) = x_0$$

mit

$$A = \begin{pmatrix} -8 & 47 & -8 \\ -4 & 18 & -2 \\ -8 & 39 & -5 \end{pmatrix}, \quad x_0 = \begin{pmatrix} 9 \\ 3 \\ 7 \end{pmatrix}.$$

Hier haben wir im Kapitel über die Jordan-Form in Beispiel 132 schon

$$Te^{Jt} = \begin{pmatrix} 6e^t & 17e^{2t} & -4e^{2t} + 17te^{2t} \\ 2e^t & 6e^{2t} & -e^{2t} + 6te^{2t} \\ 5e^t & 14e^{2t} & -3e^{2t} + 14te^{2t} \end{pmatrix}$$

berechnet. Als nächstes müssen wir nun $Tc = x_0$ lösen. Mithilfe des Gauß'schen Eliminationsverfahrens erhalten wir

$$c = \begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix},$$

und damit ist die Lösung der Differentialgleichung gegeben durch

$$x(t) = Te^{J(t-t_0)}c$$

oder ausgeschrieben

$$x(t) = -2 \begin{pmatrix} 6 \\ 2 \\ 5 \end{pmatrix} e^{t-t_0} + \begin{pmatrix} 17 \\ 6 \\ 14 \end{pmatrix} e^{2(t-t_0)} + \begin{pmatrix} 4 \\ 1 \\ 3 \end{pmatrix} e^{2(t-t_0)} - \begin{pmatrix} 17 \\ 6 \\ 14 \end{pmatrix} (t-t_0) e^{2(t-t_0)}.$$

■

Separation der Variablen: Diese Methode ist möglich für skalare Differentialgleichungen der Form

$$x' = f(t)g(x), x(t_0) = x_0,$$

mit stetigen Funktionen f, g . Ist nun $g(x_0) = 0$, so ist $x'(t_0) = 0$. Dann ist die konstante Funktion $x \equiv x_0$ anscheinend eine Lösung der Gleichung.

Dieses Verfahren mag so zunächst einmal kompliziert erscheinen, deshalb ein Beispiel.

Beispiel 62

Wir betrachten die Differentialgleichung

$$x' = \frac{t}{x}, x(t_0) = x_0 \neq 0,$$

also $f(t) = t, g(x) = \frac{1}{x}$. Es folgt Schritt für Schritt

$$\begin{aligned} \int_{t_0}^t s ds &= \int_{x_0}^x y dy, \\ \frac{1}{2}(t^2 - t_0^2) &= \frac{1}{2}(x^2 - x_0^2), \\ x^2 &= x_0^2 - t_0^2 + t^2, \\ x &= \pm \sqrt{x_0^2 - t_0^2 + t^2}, \end{aligned}$$

für $t^2 \geq t_0^2 - x_0^2$. ■

Hier erkennt man auch, woher der Name des Verfahrens stammt, wir stellen unsere Ausgangsfunktion dar als das Produkt zweier unabhängiger Funktionen, die jeweils nur von einer der Variablen abhängen, haben also die Variablen separiert.

Was ist nun, wenn dieses Verfahren nicht sofort anwendbar ist? Unter Umständen kann man die Lösung einer Differentialgleichung auf Separation der Variablen zurückführen.

Ist zum Beispiel

$$x' = f\left(\frac{x}{t}\right),$$

so setzen wir $u := \frac{x}{t}$. Es folgt

$$u' = \frac{x't - x}{t^2} = \frac{1}{t}\left(x' - \frac{x}{t}\right) = \frac{1}{t}(f(u) - u).$$

Bestimmen wir nun u mit Separation der Variablen, so erhalten wir auch x .

Beispiel 63

Sei

$$x' = \frac{2x - t}{t} = 2\frac{x}{t} - 1, x(t_0) = x_0, t_0 \neq 0.$$

Mit $u = \frac{x}{t}$ ist

$$u' = \frac{1}{t}(2u - 1 - u) = \frac{1}{t}(u - 1), \quad u(t_0) = \frac{x(t_0)}{t_0} = \frac{x_0}{t_0} =: u_0.$$

Ist nun $u(t_0) = 1$, so ist $u \equiv 1$ und damit $x = t$.

Ist $u(t_0) \neq 1$, so liefert Separation der Variablen für $\frac{t}{t_0} > 0, \frac{u-1}{u_0-1} > 0$ (diese Einschränkung ist für t nahe t_0 und u nahe u_0 immer erfüllt, da dies ja einfach nur bedeutet, dass t nahe t_0 dasselbe Vorzeichen hat, und dies gilt wegen $t_0 \neq 0$. Dasselbe gilt auch für $u - 1$ und $u_0 - 1$ wegen $u_0 \neq 1$ und der Stetigkeit von u . Diese Einschränkung benötigen wir gleich, um den Logarithmus anwenden zu können.)

$$\int_{t_0}^t \frac{1}{s} ds = \int_{u_0}^u \frac{1}{y-1} dy \Rightarrow [\ln |s|]_{t_0}^t = [\ln |y-1|]_{u_0}^u.$$

Es folgt $\ln \frac{u-1}{u_0-1} = \ln \frac{t}{t_0}$, also $u = \frac{t}{t_0}(u_0 - 1) + 1$ und $x = tu = t^2 \frac{u_0-1}{t_0} + t$.
 Hierzu noch eine kurze Anmerkung. Wäre hier nun $u_0 = 1$, so erhalten wir wie im zuerst behandelten Fall auch $x = t$. Warum dann den Fall einzeln betrachten? Wieder mal eine Aufgabe für euch ;-). Ein Tipp: Betrachtet die Integrale. ■

Die lineare skalare Differentialgleichung erster Ordnung, Variation der Konstanten:

Wir wollen hier die lineare skalare Differentialgleichung erster Ordnung, das heißt

$$x' + a(t)x = f(t), x(t_0) = x_0,$$

allgemein lösen, und zwar sowohl die homogene als auch die inhomogene.

Verzichtet man hier auf das Stellen einer Anfangsbedingung, so erhält man, dass die allgemeine Lösung der inhomogenen Gleichung gegeben ist als Summe der allgemeinen Lösung der homogenen Gleichung und der speziellen Lösung der inhomogenen Gleichung.

Auch zu dieser Methode ein Beispiel.

Beispiel 64

Wir betrachten

$$x' - \frac{x}{t} = t, \quad x(t_0) = x_0.$$

Es ist $a(t) = -t^{-1}$, $f(t) = t$. Für $\frac{t}{t_0} > 0$ gilt:

$$\int_{t_0}^t a(s)ds = -\ln \frac{t}{t_0}.$$

Also folgt

$$x(t) = \left(\int_{t_0}^t r \frac{t_0}{t} dr + x_0 \right) \frac{t}{t_0} = (t_0(t - t_0) + x_0) \frac{t}{t_0} = t^2 + t \left(\frac{x_0}{t_0} - t_0 \right).$$

■

Die Bernoulli'sche Differentialgleichung: Abgeleitet von der linearen Differentialgleichung können wir die sogenannte Bernoulli'sche Differentialgleichung lösen. Für $r \in \mathbb{R}$ betrachten wir

$$x' + a(t)x + b(t)x^r = 0, x(t_0) = x_0 > 0.$$

Wir wollen zunächst einige Spezialfälle ausschließen. Für $r = 0$ erhalten wir $x' + a(t)x + b(t) = 0$, das heißt die schon behandelte inhomogene lineare Gleichung. Ist $r = 1$, so lautet die Gleichung $x' + a(t)x + b(t)x = 0$. Auch diese homogene lineare Gleichung können wir schon lösen. Sei also $r \neq 0, 1$. Dann setzen wir $z := x^{1-r}$. Multiplikation der Differentialgleichung mit $(1-r)x^{-r}$ liefert

$$(1-r)x'x^{-r} + (1-r)a(t)x^{1-r} + (1-r)b(t) = 0, \quad z' + (1-r)a(t)z + (1-r)b(t) = 0.$$

Dies ist eine inhomogene lineare Differentialgleichung, die uns keine Schwierigkeiten bereitet.

7 Kurven

Übersicht

7.1	Definitionen	133
7.2	Sätze und Beweise	136
7.3	Erklärungen zu den Definitionen	141
7.4	Erklärungen zu den Sätzen und Beweisen	152

Was versteht man überhaupt unter einer Kurve im \mathbb{R}^n ? Wie ist sie mathematisch definiert? Und kann man sich Beispiele für Kurven überlegen und diese dann etwa von Maple[®] zeichnen lassen? Ja, das kann man natürlich. In diesem Kapitel wollen wir genau dies tun und uns vor allem einige Beispiele für ebene Kurven und Raumkurven anschauen.

Weiterhin werden wir verschiedene „Darstellungsarten“ von Kurven angeben. Danach werden wir über den Begriff der regulären Kurve und der Parametrisierung nach Bogenlänge sprechen. Wir werden zeigen, dass sich jede regulär parametrisierte Kurve nach Bogenlänge umparametrisieren lässt. Außerdem wollen wir Längen von Kurven berechnen und uns anschauen, wie die Krümmung von Kurven definiert ist, und was wir darunter verstehen wollen.

7.1 Definitionen

Definition 7.1 (Kurve)

Sei $I \subset \mathbb{R}$ ein Intervall. Eine **parametrisierte Kurve** ist eine unendlich oft differenzierbare Abbildung der Form

$$c : I \subset \mathbb{R} \rightarrow \mathbb{R}^n.$$

Anmerkung: Das Intervall I kann offen, abgeschlossen oder halbabgeschlossen sein. Ebenfalls kann das Intervall beschränkt oder unbeschränkt sein. Dies spielt keine Rolle.

Definition 7.2 (regulär parametrisiert)

Eine parametrisierte Kurve heißt **regulär parametrisiert**, wenn der Geschwindigkeitsvektor (erste Ableitung, wir schreiben hierfür $c'(t)$), nirgends verschwindet, das heißt, wenn

$$c'(t) \neq 0 \quad \forall t \in I.$$

Definition 7.3 (Umparametrisierung, Parametertransformation)

Sei $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ eine parametrisierte Kurve. Eine **Parametertransformation** von c ist eine bijektive Abbildung $\phi : J \rightarrow I$, wobei $J \subset \mathbb{R}$ ein weiteres Intervall ist, sodass sowohl $\phi : J \rightarrow I$ als auch $\phi^{-1} : I \rightarrow J$ unendlich oft differenzierbar sind. Die parametrisierte Kurve $\tilde{c} := c \circ \phi : J \rightarrow \mathbb{R}^n$ heißt **Umparametrisierung** von c .

Definition 7.4 (orientierungserhaltend, orientierungsumkehrend)

Eine Parametertransformation ϕ heißt **orientierungserhaltend**, falls $\phi'(t) > 0 \quad \forall t$ und **orientierungsumkehrend**, falls $\phi'(t) < 0 \quad \forall t$.

Definition 7.5 (Parametrisierung nach Bogenlänge)

Eine **nach Bogenlänge parametrisierte Kurve** ist eine reguläre Kurve $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ mit $|c'(t)| = 1 \quad \forall t \in I$. Dabei gilt $|c'(t)| := \sqrt{(c'_1(t))^2 + \dots + (c'_n(t))^2}$.

Anmerkung: Eine **proportional zur Bogenlänge parametrisierte Kurve** ist eine regulär parametrisierte Kurve $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$, für die $|c'(t)| = \text{const.}$ für alle t gilt.

Außerdem bemerken wir: Die Definition von $|c'(t)|$ ist eine Norm im Sinne von 1.26, aber aus Einfachheitsgründen schreiben wir für die Norm im Folgenden $|\cdot|$ statt $\|\cdot\|$.

Definition 7.6 (Spur)

Wird eine Kurve durch eine regulär parametrisierte Kurve $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ repräsentiert, dann nennt man das Bild $c(I)$ auch die **Spur** der Kurve.

Definition 7.7 (Orientierung)

Eine **orientierte Kurve** ist eine Äquivalenzklasse von parametrisierten Kurven, wobei diese als äquivalent angesehen werden, wenn sie durch orientierungserhaltende Parametertransformationen auseinander hervorgehen.

Definition 7.8 (Rektifizierbarkeit einer Kurve, Länge einer Kurve)

Eine parametrisierte Kurve $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ heißt auf dem Intervall $[a, b] \subset I$ **rektifizierbar**, falls

$$L[c] := L(c|_{[a,b]}) := \sup \left\{ \sum_{i=1}^k |c(t_i) - c(t_{i-1})| : k \in \mathbb{N}, a = t_0 < t_1 < \dots < t_k = b \right\}$$

endlich ist, also wenn $L(c|_{[a,b]}) < \infty$. In diesem Fall nennen wir $L(c|_{[a,b]})$ die **Länge der Kurve** c auf dem Intervall $[a, b]$.

Definition 7.9 (Länge einer Kurve)

Sei $c : [a, b] \subset I \rightarrow \mathbb{R}^n$ eine parametrisierte Kurve. Dann heißt

$$L[c] = \int_a^b |c'(t)| dt$$

die **Länge der Kurve** c im Intervall $[a, b]$.

Anmerkung: Diese weitere Definition der Länge einer Kurve wird durch Satz 7.4 plausibel.

Definition 7.10 (Periode, geschlossen)

Sei $c : \mathbb{R} \rightarrow \mathbb{R}^n$ eine parametrisierte Kurve. Diese Kurve heißt **periodisch mit Periode** L , falls für alle $t \in \mathbb{R}$ gilt, dass

$$c(t + L) = c(t) \quad \text{mit } L > 0.$$

Außerdem fordert man, dass es kein $0 < L' < L$ gibt mit $c(t + L') = c(t) \forall t \in \mathbb{R}$. Eine Kurve heißt **geschlossen**, wenn sie eine periodische reguläre Parametrisierung besitzt.

Definition 7.11 (einfach geschlossen)

Eine geschlossene Kurve c heißt **einfach geschlossen**, wenn sie eine periodische reguläre Parametrisierung c mit Periode L besitzt, sodass die Einschränkung auf das Intervall $[0, L]$, also $c|_{[0,L]}$ injektiv ist.

Wir betrachten nun noch die Krümmung einer Kurve im \mathbb{R}^2 .

Definition 7.12 (Normalenfeld)

Sei $c : I \rightarrow \mathbb{R}^2$ eine nach Bogenlänge parametrisierte Kurve. Das **Normalenfeld** ist definiert als

$$n(t) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \cdot c'(t). \quad (7.1)$$

Definition 7.13 (ebene Krümmung)

Sei $c : I \rightarrow \mathbb{R}^2$ eine ebene, nach Bogenlänge parametrisierte Kurve. Die Funktion $\kappa : I \rightarrow \mathbb{R}$, die der Gleichung (7.2) genügt, heißt **Krümmung** von c .

$$c''(t) = \kappa(t) \cdot n(t). \quad (7.2)$$

Anmerkung: Dass die Krümmung ein Skalar ist, ist nicht selbstverständlich und wird in den Erklärungen deutlich.

7.2 Sätze und Beweise

Satz 7.1

Die Notation sei wie in Definition 7.3. Gegeben sei eine reguläre Parametrisierung der Kurve c . Dann ist jede Umparametrisierung wieder regulär.

Beweis: Sei $\tilde{c} := c \circ \phi$ die Umparametrisierung einer regulären Kurve c . Es gilt $\phi(t) \neq 0$ für alle $t \in J$. Daraus ergibt sich nun mittels Differentiation auf beiden Seiten

$$1 = \frac{d}{dt} (\phi^{-1} \circ \phi)(t)$$

und nach der Kettenregel

$$\frac{d}{dt} \phi^{-1}(\phi(t)) \cdot \frac{d}{dt} \phi(t) \Rightarrow \frac{d}{dt} \phi(t) = \phi'(t) \neq 0 \quad \forall t \in J.$$

Dies war nur Vorbereitung. Jetzt ergibt sich also

$$\tilde{c}'(t) = (c \circ \phi)'(t) = c'(\phi(t)) \cdot \phi'(t) \neq 0$$

und damit die Regularität der umparametrisierten Kurve und folglich die Behauptung. q.e.d.

Satz 7.2

Jede rektifizierbare, regulär parametrisierte Kurve lässt sich so umparametrisieren, dass die Umparametrisierung nach Bogenlänge parametrisiert ist.

Beweis: Seien $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ regulär parametrisiert und $t_0 \in I$. Für den Beweis definieren wir $\psi(s) := \int_{t_0}^s |c'(t)| dt$. Nach dem Hauptsatz der Differential- und Integralrechnung ergibt sich nun $\psi'(s) = |c'(s)| > 0$. Das bedeutet wiederum, dass ψ streng monoton wachsend ist. Dies liefert, dass $\psi : I \rightarrow J := \psi(I)$ eine orientierungserhaltende Parametertransformation ist. Wir setzen $\phi := \psi^{-1} : J \rightarrow I$. Unter Anwendung der Kettenregel folgt

$$\phi'(t) = \frac{1}{\psi'(\phi(t))} = \frac{1}{|c'(\phi(t))|}.$$

Dies ergibt

$$|\tilde{c}'(t)| = |(c \circ \phi)'(t)| = |c'(\phi(t)) \cdot \phi'(t)| = \left| c'(\phi(t)) \cdot \frac{1}{|c'(\phi(t))|} \right| = 1.$$

Also ist \tilde{c} nach Bogenlänge parametrisiert.

q.e.d.

Satz 7.3

Sind $c_1 : I_1 \subset \mathbb{R} \rightarrow \mathbb{R}^n$ und $c_2 : I_2 \subset \mathbb{R} \rightarrow \mathbb{R}^n$ Parametrisierungen nach der Bogenlänge derselben Kurve c , so ist die zugehörige Parametertransformation $\phi : I_1 \rightarrow I_2$ mit $c_1 = c_2 \circ \phi$ von der Form $\phi(t) = t + t_0$ für ein $t_0 \in \mathbb{R}$, falls c_1 und c_2 gleich orientiert sind. Falls c_1 und c_2 entgegengesetzt orientiert sind, ist sie von der Form $\phi(t) = -t + t_0$.

Beweis: Seien die Bezeichnungen wie in dem Satz gewählt. Es gilt

$$|c'_1(s)| = 1 = |c'_2(t)| \quad \forall s \in I_1, t \in I_2,$$

da c_1, c_2 beide Parametrisierungen nach der Bogenlänge sind. Außerdem ist $c_1 = c_2 \circ \phi$, also gilt nach Kettenregel

$$c'_1(s) = (c_2 \circ \phi)'(s) = c'_2(\phi(s))\phi'(s),$$

in der Norm also

$$1 = |c'_1(s)| = |c'_2(\phi(s))\phi'(s)| = |\phi'(s)| \cdot |c'_2(\phi(s))| = |\phi'(s)|.$$

Ist ϕ orientierungserhaltend, so gilt $\phi'(s) > 0$ für alle $s \in I_1$, also gilt $1 = |\phi'(s)| = \phi'(s)$ für alle $s \in I_1$. Integriert ergibt das $\phi(t) = t + t_0 \quad \forall t \in I_1$ für ein $t_0 \in \mathbb{R}$. Ist ϕ orientierungsumkehrend, so gilt entsprechend $\phi'(s) < 0$, also $1 = |\phi'(s)| = -\phi'(s)$, oder anders $\phi'(s) = -1$ für alle $s \in I_1$. Integriert hat man wieder $\phi(t) = -t + t_0 \quad \forall t \in I_1$ für ein $t_0 \in \mathbb{R}$. Damit ist alles gezeigt. q.e.d.

Satz 7.4

Sei $c : [a, b] \rightarrow \mathbb{R}^n$ eine parametrisierte Kurve. Dann gibt es für jedes $\varepsilon > 0$ ein $\delta > 0$, sodass für jede Unterteilung $a = t_0 < t_1 < \dots < t_k = b$ mit $t_{i+1} - t_i < \delta$, wobei $i = 0, \dots, k$ gilt:

$$|L[c] - L[P]| < \varepsilon, \quad \text{wobei } P = (c(t_0), \dots, c(t_k)).$$

Anmerkung: $P = (c(t_0), \dots, c(t_k))$ bezeichnet einen sogenannten Polygon- oder Streckenzug. Man verbindet einfach die Punkte $c(t_0)$ bis $c(t_n)$. Ein Bildchen hierzu findet ihr in den Erklärungen.

Beweis: Der Beweis ist etwas umfangreich und verläuft in insgesamt fünf Schritten. Also tief durchatmen, Luft holen und los geht es. Sei $\varepsilon > 0$ vorgegeben. Wir wählen $\varepsilon' \in \left(0, \frac{\varepsilon}{1 + \sqrt{n(b-a)}}\right)$. Wieso das Sinn macht, werden wir gleich noch sehen.

1. Schritt: Wir behaupten zunächst einmal: Zu ε' existiert ein $\delta_0 > 0$, sodass für jede Unterteilung $a = t_0 < t_1 < \dots < t_k = b$ mit $t_{i+1} - t_i < \delta_0, i = 0, \dots, k$ gilt:

$$\left| L[c] - \sum_{i=0}^{k-1} |c'(t_{i+1}) \cdot (t_{i+1} - t_i)| \right| < \varepsilon'.$$

Dies zeigen wir so: Das Integral von riemann-integrierbaren Funktionen kann durch Riemann'sche Summen approximiert werden. Es gilt:

$$\left| \int_a^b f(t) dt - \sum_{i=0}^{k-1} f(t_{i+1}) \cdot (t_{i+1} - t_i) \right| < \varepsilon.$$

Hier ist $L[c] = \int_a^b |c'(t)| dt$, das heißt, $f(t) = |c'(t)|$.

2. Schritt: Wir zeigen: Zu ε' existiert ein $\delta_j > 0$, sodass $|c'_j(t) - c'_j(s)| < \varepsilon'$, falls $|t - s| < \delta_j$ mit $t, s \in [a, b]$. Der Beweis geht so: $c'_j : [a, b] \rightarrow \mathbb{R}$ sind stetig und $[a, b]$ kompakt (nach Heine-Borel (Satz 1.12), da abgeschlossen und beschränkt). Daraus folgt, dass $c'_j, j = 1, \dots, n$ gleichmäßig stetig sind.
3. Schritt: Wir definieren $\delta := \min\{\delta_0, \dots, \delta_n\}$. Sei nun eine Unterteilung $a = t_0 < t_1 < \dots < t_k = b$ der Feinheit kleiner als δ vorgegeben. Nach dem Mittelwertsatz der Analysis existiert ein $\tau_{i,j} \in (t_i, t_{i+1})$, sodass

$$c_j(t_{i+1}) - c_j(t_i) = c'_j(\tau_{i,j})(t_{i+1} - t_i).$$

4. Schritt: Wir behaupten: Es gilt:

$$|c(t_{i+1}) - c(t_i)| - |c'(t_{i+1}) \cdot (t_{i+1} - t_i)| < \sqrt{n} \cdot \varepsilon'(t_{i+1} - t_i).$$

Dies rechnen wir einfach nach:

$$\begin{aligned} & \left| |c(t_{i+1}) - c(t_i)| - |c'(t_{i+1}) \cdot (t_{i+1} - t_i)| \right| = \\ & \left| (c'_1(\tau_{i,1}), \dots, c'_n(\tau_{i,1})) \right| - \left| (c'_1(t_{i+1}), \dots, c'_n(t_{i+1})) \right| \cdot (t_{i+1} - t_i) \\ & \leq \left| c'_1(\tau_{i,1}) - c'_1(t_{i+1}), \dots, c'_n(\tau_{i,1}) - c'_n(t_{i+1}) \right| \cdot (t_{i+1} - t_i) \\ & = \sqrt{\sum_{j=1}^n (c'_j(\tau_{i,1}) - c'_j(t_{i+1}))^2} \cdot (t_{i+1} - t_i) \\ & \leq \sqrt{n} \cdot \varepsilon'(t_{i+1} - t_i). \end{aligned}$$

5. Schritt: Summation über i liefert nun:

$$\begin{aligned}
 & \left| L[P] - \sum_{i=0}^{k-1} |c'(t_{i+1})| \cdot (t_{i+1} - t_i) \right| \\
 &= \left| \sum_{i=0}^{k-1} |c(t_{i+1}) - c(t_i)| - \sum_{i=0}^{k-1} |c'(t_{i+1})| \cdot (t_{i+1} - t_i) \right| \\
 &\leq \sum_{i=0}^{k-1} ||c(t_{i+1}) - c(t_i)| - |c'(t_{i+1}) \cdot (t_{i+1} - t_i)|| \\
 &\leq \sum_{i=0}^{k-1} \sqrt{n} \cdot \varepsilon' \cdot (t_{i+1} - t_i) = \sqrt{n} \cdot \varepsilon' \sum_{i=0}^{k-1} (t_{i+1} - t_i) = \sqrt{n} \varepsilon' (b - a).
 \end{aligned}$$

Nun ergibt sich insgesamt die Behauptung und zwar folgt

$$\begin{aligned}
 & |L[P] - L[c]| \\
 &= \left| L[P] - \sum_{i=0}^{k-1} |c'(t_{i+1})| \cdot (t_{i+1} - t_i) + \sum_{i=0}^{k-1} |c'(t_{i+1})| \cdot (t_{i+1} - t_i) - L[c] \right| \\
 &\leq \left| L[P] - \sum_{i=0}^{k-1} |c'(t_{i+1})| \cdot (t_{i+1} - t_i) \right| + \left| \sum_{i=0}^{k-1} |c'(t_{i+1})| \cdot (t_{i+1} - t_i) - L[c] \right| \\
 &\leq \sqrt{n} \cdot \varepsilon' (b - a) + \varepsilon' = \varepsilon (1 + \sqrt{n}(b - a)) < \varepsilon,
 \end{aligned}$$

$$\text{denn } \varepsilon' < \frac{\varepsilon}{1 + \sqrt{n}(b - a)}.$$

Damit ist alles gezeigt, und jetzt versteht ihr auch, wieso wir am Anfang ε' so gewählt haben. q.e.d.

Satz 7.5 (Länge unabhängig von Wahl der Parametrisierung)

Sei $c : [a, b] \rightarrow \mathbb{R}^n$ regulär parametrisiert und $\tilde{c} = c \circ \phi : [a', b'] \rightarrow \mathbb{R}^n$ eine Umparametrisierung mit $\phi : [a', b'] \rightarrow [a, b]$. Dann gilt $L[c] = L[\tilde{c}]$.

Beweis: Dies folgt sofort aus der Kettenregel und dem Transformationssatz für Integrale, genauer: Seien $\tilde{c} = c \circ \phi$ die Umparametrisierung von c und $\phi : [a', b'] \rightarrow [a, b]$. O.B.d.A. nehmen wir an, dass $\phi'(t) > 0$, das heißt, ϕ ist orientierungserhaltend. Es folgt nun

$$\begin{aligned}
 L[\tilde{c}] &= \int_{a'}^{b'} |\tilde{c}'(t)| dt = \int_{a'}^{b'} |(c \circ \phi)'(t)| dt = \int_{a'}^{b'} |c'(\phi(t)) \cdot \phi'(t)| dt \\
 &= \int_{a'}^{b'} |c'(\phi(t))| \cdot |\phi'(t)| dt.
 \end{aligned}$$

Substitution $s := \phi(t)$ liefert $\frac{ds}{dt} = \phi'(t)$ und damit

$$L[\tilde{c}] = \dots = \int_a^b |c'(s)| ds = L[c].$$

Und da steht das Gewünschte $L[\tilde{c}] = L[c]$.

q.e.d.

Anmerkung: Man kann also von der Länge einer Kurve sprechen, da die Länge parametrisierter Kurven nicht von der speziellen Parametrisierung abhängt.

Satz 7.6

Sei $c : I \rightarrow \mathbb{R}^2$ eine ebene Kurve. Für die Krümmung $\kappa(t)$ gilt dann

$$\kappa(t) = \frac{\det(c'(t), c''(t))}{|c'(t)|^3}.$$

Anmerkung: $\det(c'(t), c''(t))$ ist so zu verstehen, dass wir in eine Matrix spaltenweise die Vektoren $c'(t)$ und $c''(t)$ schreiben und von dieser Matrix die Determinante berechnen.

Beweis: Seien

$$c(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}, \quad c'(t) = \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix}, \quad c''(t) = \begin{pmatrix} x''(t) \\ y''(t) \end{pmatrix}, \quad n(t) = \begin{pmatrix} n_1(t) \\ n_2(t) \end{pmatrix}.$$

Es gilt nun:

$$|c'(t)|' = \sqrt{\langle c'(t), c'(t) \rangle}' = \frac{\langle c''(t), c'(t) \rangle}{\sqrt{\langle c'(t), c'(t) \rangle}}.$$

Jetzt rechnen wir die Formel unter Anwendung der Definition 7.13 einfach nach:

$$\begin{aligned} \kappa(t) &= \frac{\langle n_1'(t), n_2(t) \rangle}{|c'(t)|} = \frac{1}{|c'(t)|} \left\langle \left(\frac{c'(t)}{|c'(t)|} \right)', n_2(t) \right\rangle \\ &= \frac{1}{|c'(t)|} \left\langle \frac{c''(t) \cdot |c'(t)| - c'(t) \cdot |c'(t)|'}{|c'(t)|^2}, n_2(t) \right\rangle \\ &= \frac{1}{|c'(t)|^3} \left\langle c''(t) \cdot |c'(t)| - \frac{c'(t) \cdot \langle c''(t), c'(t) \rangle}{|c'(t)|}, n_2(t) \right\rangle \\ &= \frac{1}{|c'(t)|^3} \left\langle \begin{pmatrix} x''(t) \\ y''(t) \end{pmatrix}, \begin{pmatrix} -y'(t) \\ x'(t) \end{pmatrix} \right\rangle - \frac{1}{|c'(t)|^5} \left\langle \begin{pmatrix} x''(t) \\ y''(t) \end{pmatrix}, \begin{pmatrix} x'(t) \\ y'(t) \end{pmatrix} \right\rangle \\ &= \frac{1}{|c'(t)|^3} \left\langle \begin{pmatrix} x''(t) \\ y''(t) \end{pmatrix}, \begin{pmatrix} -y'(t) \\ x'(t) \end{pmatrix} \right\rangle \\ &= \frac{1}{|c'(t)|^3} \left\langle \begin{pmatrix} -y'(t) \\ x'(t) \end{pmatrix}, \begin{pmatrix} x''(t) \\ y''(t) \end{pmatrix} \right\rangle \\ &= \frac{1}{|c'(t)|^3} \det(c'(t), c''(t)). \end{aligned}$$

q.e.d.

7.3 Erklärungen zu den Definitionen

Zur Definition 7.1 einer Kurve: Anschaulich ist eine Kurve nichts anderes als ein, in der Regel verbogenes, in den Raum gelegtes Geradenstück. Rein mathematisch ist eine Kurve nur ein wichtiger Spezialfall einer differenzierbaren Abbildung. Schauen wir uns ein paar Beispiele für Kurven in der Ebene (\mathbb{R}^2) und im Raum (\mathbb{R}^3) an.

Beispiel 65

- Einen *Kreis* kann man beispielsweise durch

$$c : \mathbb{R} \rightarrow \mathbb{R}^2, c(t) := (\cos(t), \sin(t))$$

parametrisieren. Er sieht dann so aus wie in Abbildung 7.1. Wie jede

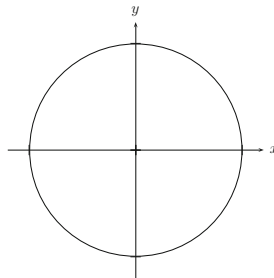


Abb. 7.1: Ein Kreis.

Kurve besitzt auch der Kreis unendlich viele Parametrisierungen. Wir könnten beispielsweise den Kreis auch schneller durchlaufen, also durch $c(t) = (\cos(2t), \sin(2t))$ parametrisieren. Aber natürlich interessiert man sich nur für solche Parametrisierungen, die „schöne“ Eigenschaften, wie beispielsweise die Parametrisierung nach Bogenlänge etc. besitzen. Wir gehen hierauf noch einmal in den Erklärungen zur Definition 7.3 ein.

- Die Parametrisierung der *Schleife* in der Abbildung 7.2 erfolgt durch

$$c : \mathbb{R} \rightarrow \mathbb{R}^2, c(t) := (t^2 - 1, t(t^2 - 1)).$$

- Eine *Gerade* könnten wir auch wie folgt parametrisieren:

$$c : \mathbb{R} \rightarrow \mathbb{R}^2, c(t) := (t^3, t^3).$$

- Sie sieht aus wie eine Schnecke. Mathematiker sagen zu ihr *logarithmische Spirale*. Eine Parametrisierung ergibt sich aus

$$c : \mathbb{R} \rightarrow \mathbb{R}^2, c(t) := (e^{-t} \cos(2\pi t), e^{-t} \sin(2\pi t)).$$

Dies sieht dann so aus, wie in Abbildung 7.3.

Kommen wir nun zu Raumkurven.

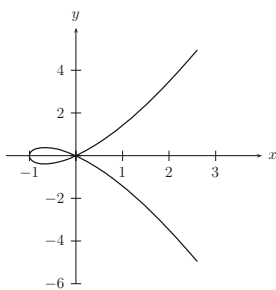


Abb. 7.2: Eine Schleife.

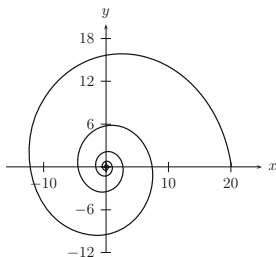


Abb. 7.3: Die logarithmische Spirale; eine Schnecke.

- Die *Helix* parametrisiert man wie folgt durch

$$c : \mathbb{R} \rightarrow \mathbb{R}^3, c(t) := (\cos(t), \sin(t), t).$$

Dies sieht dann so aus wie in Abbildung 7.4.

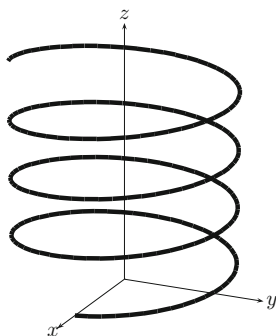


Abb. 7.4: Eine Helix. Man kennt sie vielleicht von der berühmten DNA-Struktur.

- Eine weitere schöne Raumkurve ist die *konische Spirale*. Die Parametrisierung ergibt sich aus

$$c : \mathbb{R} \rightarrow \mathbb{R}^3, c(t) := (t \cos(t), t \sin(t), t).$$

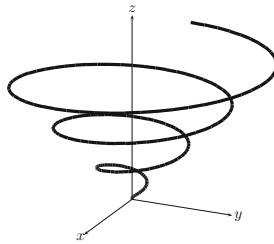


Abb. 7.5: Die konische Spirale.

Wir werden vorwiegend ebene Kurven oder Raumkurven betrachten, das heißt Abbildungen der Form $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^2$ oder $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^3$. Noch eine kurze Anmerkung zur Schreibweise: Für eine Abbildung der Form $c : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ gilt in Koordinatenschreibweise

$$c(t) = (c_1(t), c_2(t), \dots, c_n(t))$$

und für die erste Ableitung

$$c'(t) = (c'_1(t), c'_2(t), \dots, c'_n(t)).$$

Jede Komponente besteht also aus einer Funktion, die dann differenziert wird. Es gilt genauer

$$c'_j = \frac{d}{dt} c_j \quad \forall j = 1, \dots, n.$$

■

Zur Definition 7.2 einer regulär parametrisierten Kurve: Wir stellen mit dieser Bedingung sicher, dass sich beim Durchlauf von $t \in I$ der Kurvenpunkt $c(t)$ tatsächlich bewegt. Wir schließen damit konstante Abbildungen der Form $c(t) = d, d \in \mathbb{R}$ aus. Das ist sicherlich sinnvoll, denn das Bild dieser Abbildung besteht nur aus dem Punkt d , und das ist nicht gerade das, was man sich unter einer Kurve vorstellt. Den Ableitungsvektor $c'(t)$ bezeichnen wir auch als *Tangentenvektor*. Physikalisch wird er als *Geschwindigkeitsvektor* zum Zeitpunkt t interpretiert.

Beispiel 66 (Die Traktrix)

Die sogenannte *Schleppkurve* (*Traktrix*) kann parametrisiert werden durch

$$c : \left(0, \frac{\pi}{2}\right) \rightarrow \mathbb{R}^2, c(t) = 4 \left(\sin(t), \cos(t) + \log(\tan(t/2)) \right).$$

Graphisch veranschaulichen wir dies in Abbildung 7.6. Es gilt:

$$c'(t) = 4 \left(\cos(t), -\sin(t) - \frac{1}{2} \frac{1}{\tan(t/2)} \cdot \frac{1}{\cos^2(t/2)} \right).$$

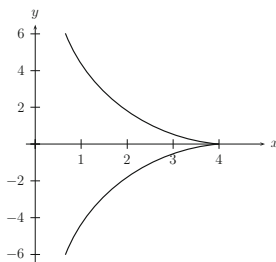


Abb. 7.6: Die Traktrix für $t \in (0, \pi)$. Der Wert $t = \pi/2$ entspricht der Spitze bei $(4, 0)$. Der untere Teil alleine, das heißt $t \in (0, \pi/2)$, ist regulär.

Wir könnten nun die Kurve c auf ganz $(0, \pi)$ definieren. Das Problem ist aber, dass für $t = \frac{\pi}{2}$ sich $c'(\pi/2) = (0, 0)$ und somit $|c'(\pi/2)| = 0$. Daher hätte die Kurve in diesem Punkt, also im Punkt $c(\pi/2) = (4, 0)$, eine Spitze. Und genau das wollen wir durch unsere Definition 7.2 der Regularität ausschließen. ■

Zur Übung wollen wir einige Geschwindigkeitsvektoren berechnen.

Beispiel 67

- Den Kreis hatten wir weiter oben parametrisiert durch $c(t) = (\cos(t), \sin(t))$. Für den Geschwindigkeitsvektor ergibt sich nun $c'(t) = (-\sin(t), \cos(t))$. Es gilt natürlich $c'(t) \neq 0$ für alle t , da Sinus und Kosinus niemals an derselben Stelle t Null werden. Dies macht man sich zum Beispiel klar, wenn man sich das Bild vom Sinus und Kosinus vor Augen führt. Das Beispiel des Kreises zeigt weiter, dass eine regulär parametrisierte Kurve nicht unbedingt injektiv sein muss. Man kann die Injektivität aber nach Einschränkung auf ein kleines Intervall fordern. Darauf gehen wir in Definition 7.11 bzw. den Erklärungen dazu ein.
- Die Gerade hatten wir parametrisiert durch $c(t) = (t^3, t^3)$. Es ergibt sich $c'(t) = (3t^2, 3t^2)$.
- Die *Neil'sche Parabel*, die durch

$$c : \mathbb{R} \rightarrow \mathbb{R}^2, c(t) = (t^2, t^3)$$

parametrisiert wird, ist für $t = 0$ nicht regulär (man sagt sie besitzt dort einen *singulären Punkt*), da dann $c'(0) = (0, 0)$, denn es gilt $c'(t) = (2t, 3t^2)$.

Kommen wir nun zu einem Beispiel für Raumkurven.

- Der Geschwindigkeitsvektor der Helix lautet $c'(t) = (-\sin(t), \cos(t), 1)$. Auch hier sieht man, dass die Helix regulär parametrisiert ist, da der letzte Eintrag des Geschwindigkeitsvektors immer konstant 1 ist. ■

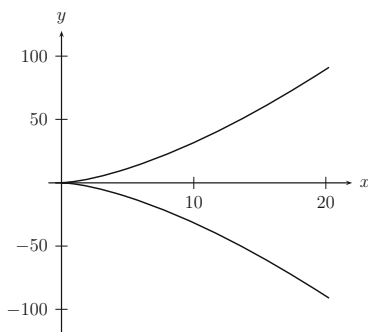


Abb. 7.7: Die Neil'sche Parabel.

Zur Definition 7.3 der Umparametrisierung: Folgende Abbildung 7.8 verdeutlicht Definition 7.3 ganz gut. Umparametrisierungen kann man sich so klar

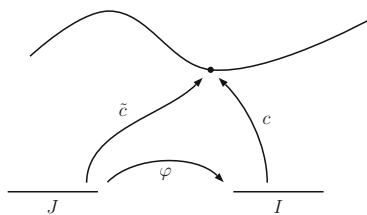


Abb. 7.8: Eine Umparametrisierung „anschaulich“. $\tilde{c}(t)$ entspricht also gerade $\varphi \circ c(t)$.

machen: Schauen wir uns das Konzept der Umparametrisierung am Beispiel eines Kreises an. Diesen hatten wir im Beispiel 65 parametrisiert als $c(t) = (\cos(t), \sin(t))$. Aber wir können ihn auch als $c(t) = (\cos(2t), \sin(2t))$ parametrisieren. Das Bild ist dasselbe, aber wir durchlaufen den Kreis schneller.

Beispiel 68

Betrachten wir nun noch ein Beispiel von regulär parametrisierten Kurven, die man so umparametrisieren kann und somit zeigen kann, dass sie dieselbe Kurve darstellen und beschreiben, also äquivalent sind. (Wir haben beispielsweise in Beispiel 65 angemerkt, dass eine Gerade mehrere Parametrisierungen besitzt, aber alle dasselbe Bild liefern.) Dazu betrachten wir die beiden regulär parametrisierten Kurven

$$c_1 : \mathbb{R} \rightarrow \mathbb{R}^2, c_1(t) = (t, t)$$

und

$$c_2 : \mathbb{R}_{>0} \rightarrow \mathbb{R}^2, c_2(t) = (\ln(t), \ln(t)).$$

Diese beiden Kurven sind äquivalent, denn es gilt mit der Parametertransformation $\phi(t) = e^t$ gerade

$$c_1(t) = (c_2 \circ \phi)(t) = c_2(\phi(t)) = \ln(e^t) = t.$$

Sie repräsentieren daher dieselbe Kurve. ■

Zur Definition 7.4 einer Orientierungserhaltenden Parametertransformation: Was eine Parametertransformation ändern kann, ist die Richtung, in der die Bildkurve durchlaufen wird. Sie kann sie entweder erhalten oder umkehren. Die triviale Parametertransformation $\phi(t) = t$ beispielsweise ändert nichts an der parametrisierten Kurve, die Parametertransformation $\psi(t) = -t$ dagegen ändert den Durchlaufsin. Es ist klar, dass eine Parametertransformation entweder Orientierungserhaltend oder Orientierungsumkehrend ist. Dies begründet man mit dem Zwischenwertsatz aus der Analysis. Denn angenommen, es gibt ein $t_1 \in I$ mit $\phi'(t_1) < 0$ und ein $t_2 \in I$ mit $\phi'(t_2) > 0$, so gäbe es nach dem Zwischenwertsatz ein $t_3 \in I$ mit $\phi'(t_3) = 0$. Dies ist aber nicht möglich.

Zur Definition 7.5 einer nach Bogenlänge parametrisierten Kurve: Nach Bogenlänge parametrisierte Kurven werden mit konstanter Geschwindigkeit 1 durchlaufen. Welche Vorteile dies zum Beispiel bei der Berechnung der Länge von Kurven hat, werden wir in Definition 7.9 und den entsprechenden Erklärungen dazu sehen.

Zur Definition 7.7 der Orientierung: Die Orientierung einer Kurve gibt einfach nur die Richtung an, in der die Kurve durchlaufen wird. Von vornherein können wir diese aber nicht für jede Parametrisierung definieren, sondern dies geht erst dann, wenn die Kurve nicht „umdreht“, also wenn der Geschwindigkeitsvektor an keiner Stelle t verschwindet.

Zur Definition 7.8 der Rektifizierbarkeit: Stellt man sich eine beliebige Kurve vor, so könnte man die Kurve doch durch Polygonzüge (Streckenzüge) approximieren und durch Grenzwertübergang würden wir dann die Länge der Kurve erhalten. Betrachten wir die folgende Abbildung 7.9.

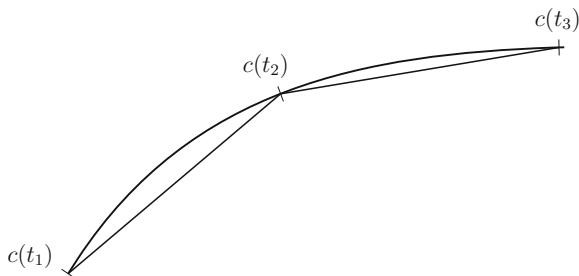


Abb. 7.9: Approximation der Länge einer Kurve durch Polygonzüge.

Unterteilen wir das entsprechende Intervall immer weiter, so bekommen wir eine noch genauere Näherung der eigentlichen Länge der Kurve (siehe Abbildung 7.10): Nun wollen wir dies etwas mathematisieren: Erst einmal schauen wir uns die Abbildungen 7.9 und 7.10 noch einmal genauer an. Um nun die Länge der Kurve zu bestimmen, approximieren wir also die Kurve durch Streckenzüge, das

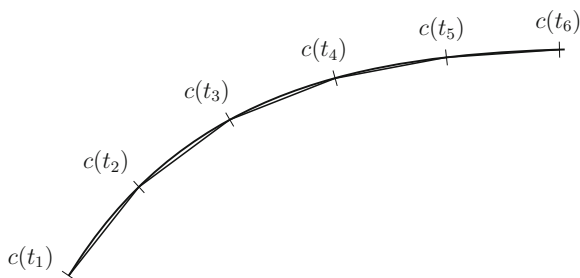


Abb. 7.10: Verkleinern der Streckenzüge liefert eine genauere Approximation der Kurvenlänge.

heißt, wir wählen eine Zerlegung des Intervalls $[a, b]$ durch $a = t_0 < t_1 < \dots < t_k = b$ mit $k \in \mathbb{N}$ und bilden damit den Streckenzug, indem wir bei $c(t_1)$ starten, geradlinig zu $c(t_2)$ laufen, von hier weiter geradlinig zu $c(t_3)$ und so weiter, bis wir schließlich bei $c(t_k)$ angekommen sind. Die Länge dieser Strecken ist also durch $|c(t_i) - c(t_{i-1})|$ der Punkte $c(t_{i-1})$ und $c(t_i)$ gegeben (man sollte sich klar machen, dass dies eigentlich Ortsvektoren sind). So erhalten wir die Länge der Verbindungsstrecken. Die Länge des gesamten Streckenzugs ist damit die Summe

$$\sum_{i=1}^k |c(t_i) - c(t_{i-1})|.$$

Wählt man also eine Verfeinerung der Zerlegung, so wird der Weg im Allgemeinen durch den feineren Streckenzug besser approximiert, und die Länge des feineren Streckenzugs wird aufgrund der Dreiecksungleichung höchstens länger. Es ist nun aber sehr umständlich, die Länge einer Kurve mittels Streckenzügen zu berechnen. Daher gibt es Definition 7.9, die durch Satz 7.4 plausibel gemacht wird, denn dort haben wir gezeigt, dass Definition 7.8 und 7.9 der Länge einer Kurve äquivalent sind.

Zur Definition 7.9 der Länge einer Kurve: Wir verwenden natürlich diese Definition der Länge einer Kurve, um sie konkret auszurechnen. Es ist jetzt klar, wieso man sich über den Satz 7.2 so freuen kann. Dort hatten wir gezeigt, dass jede regulär parametrisierte Kurve nach Bogenlänge umparametrisiert werden kann. *Eine nach Bogenlänge parametrisierte Kurve ist also gerade so lang wie das Parameterintervall.* Wir müssen jetzt nur noch einmal zeigen, dass sich die Länge einer Kurve aber unter Umparametrisieren nicht ändert. Dies ist aber die Aussage des Satzes 7.5. Um die Länge einer Kurve $c : [a, b] \rightarrow \mathbb{R}^n$ im Intervall $[a, b]$ zu berechnen, müssen wir also nur das Integral $L[c] = \int_a^b |c'(t)| dt$ berechnen. Schauen wir uns ein paar Beispiele an.

Beispiel 69

- Wir betrachten den Kreis $c : [0, 2\pi] \rightarrow \mathbb{R}^2, c(t) = (r \cos(t), r \sin(t))$, wobei r der Radius ist. Wir wollen die Länge der Kurve im Intervall $[0, 2\pi]$

berechnen. Dazu benötigen wir zunächst den Geschwindigkeitsvektor $c'(t)$. Dieser ist gerade gegeben durch

$$c'(t) = (-r \sin(t), r \cos(t)) = r(-\sin(t), \cos(t)).$$

Dadurch ergibt sich die folgende Länge der Kurve:

$$\begin{aligned} \int_0^{2\pi} |c'(t)| dt &= \int_0^{2\pi} |(-r \sin(t), r \cos(t))| dt \\ &= \int_0^{2\pi} \sqrt{r^2(\sin^2(t) + \cos^2(t))} dt \\ &= \int_0^{2\pi} r dt = 2\pi r. \end{aligned}$$

So sollte es auch sein. Wir kennen dies als den Umfang eines Kreises mit Radius r .

- Wir betrachten die Zykloide $c: [0, 2\pi] \rightarrow \mathbb{R}^2, c(t) := (t - \sin(t), 1 - \cos(t))$. Wir berechnen die Länge der Kurve im Intervall $[0, 2\pi]$. Zunächst ist $c'(t) = (1 - \cos(t), \sin(t))$, also

$$\begin{aligned} |c'(t)| &= |(1 - \cos(t), \sin(t))| = \sqrt{(1 - \cos(t))^2 + \sin^2(t)} \\ &= \sqrt{1 - 2\cos(t) + \cos^2(t) + \sin^2(t)} = \sqrt{2(1 - \cos(t))}. \end{aligned}$$

Die Länge der Kurve ist demnach gegeben durch

$$\int_0^{2\pi} |c'(t)| dt = \int_0^{2\pi} \sqrt{2(1 - \cos(t))} dt = \dots = 8.$$

- Nun aber endlich zu einem Beispiel einer Kurve, die nach Bogenlänge parametrisiert ist, und deren Länge man auf einem bestimmten Intervall direkt ablesen kann.

Wir verwenden als Beispiel einfach den Einheitskreis $c: \mathbb{R} \rightarrow \mathbb{R}^2, c(t) = (\cos(t), \sin(t))$ und wollen die Länge im Intervall $[0, \pi]$ berechnen. Wir wissen, dass $c(t)$ nach Bogenlänge parametrisiert ist, denn

$$|c'(t)| = |(-\sin(t), \cos(t))| = \sqrt{\sin^2(t) + \cos^2(t)} = 1.$$

Also folgt sofort

$$L[c] = \int_0^\pi |c'(t)| dt = \int_0^\pi 1 dt = \pi.$$

Der ein oder andere wird sich vielleicht fragen, wieso wir die Kurven aus den ersten Beispielen nicht einfach nach Bogenlänge umparametrisieren, damit wir das Integral leichter ermitteln können. Der Grund liegt einfach darin, dass solche Umparametrisierungen in der Realität nicht so leicht zu finden sind.

- Wir berechnen die Länge der logarithmischen Spirale $c : \mathbb{R} \rightarrow \mathbb{R}^2$, $c(t) := \mu e^{\lambda t} \cdot (\cos(t), \sin(t))$ mit $\mu < 0 < \lambda$ auf einem Intervall $[a, b]$. Da $c'(t) = \mu e^{\lambda t} (\lambda \cos(t) - \sin(t), \lambda \sin(t) + \cos(t))$, folgt

$$\begin{aligned} |c'(t)| &= \sqrt{(\mu e^{\lambda t})^2 \cdot |(\lambda \cos(t) - \sin(t), \lambda \sin(t) + \cos(t))|^2} \\ &= \mu e^{\lambda t} \sqrt{\lambda^2 \cos^2(t) + \lambda^2 \sin^2(t) + \sin^2(t) + \cos^2(t)} \\ &= \mu e^{\lambda t} \sqrt{1 + \lambda^2}. \end{aligned}$$

Wir erhalten also

$$L(c|_{[a,b]}) = \int_a^b \mu e^{\lambda t} \sqrt{1 + \lambda^2} = \frac{\mu \sqrt{1 + \lambda^2}}{\lambda} (e^{\lambda b} - e^{\lambda a}).$$

- Die Helix mit der Ganghöhe h und Radius $r > 0$ ist die Kurve

$$c : \mathbb{R} \rightarrow \mathbb{R}^3, c(t) := (r \cos(t), r \sin(t), \frac{ht}{2\pi}).$$

Der Geschwindigkeitsvektor von c ist $c'(t) = (-r \sin(t), r \cos(t), h/2\pi)$.
Damit besitzt $c'(t)$ die Länge

$$\begin{aligned} |c'(t)| &= \left| \left(-r \sin(t), r \cos(t), \frac{h}{2\pi} \right) \right| = \sqrt{r^2 \sin^2(t) + r^2 \cos^2(t) + \frac{h^2}{4\pi^2}} \\ &= \sqrt{r^2 (\sin^2(t) + \cos^2(t)) + \frac{h^2}{4\pi^2}} = \sqrt{r^2 + \frac{h^2}{4\pi^2}}. \end{aligned}$$

Daraus ergibt sich nun

$$L(c|_{[a,b]}) = (b - a) \cdot \sqrt{r^2 + \frac{h^2}{4\pi^2}}.$$

Für proportional nach Bogenlänge parametrisierte Kurven gilt demnach

$$L(c|_{[a,b]}) = (b - a) \cdot |c'(t)|.$$

■

Zur Definition 7.10 der Periode einer Kurve: Ein triviales Beispiel.

Beispiel 70

Der Kreis ist periodisch mit Periode $L = 2\pi$.

■

Zur Definition 7.11 der einfachen Geschlossenheit: Verdeutlichen wir uns die Definition anhand zweier Bilder. Siehe dazu die Abbildung 7.11.

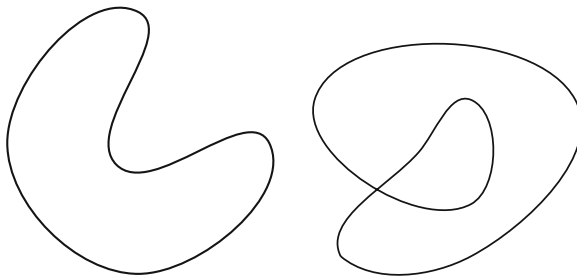


Abb. 7.11: Beispiel einer einfach geschlossenen Kurve im \mathbb{R}^2 (links) und einer nicht einfach geschlossenen Kurve im \mathbb{R}^2 (rechts).

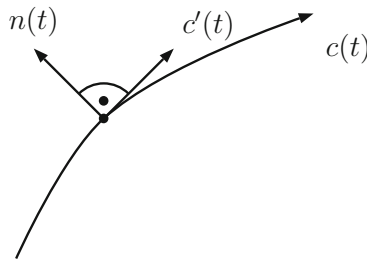


Abb. 7.12: Das Normalenfeld (der Normalenvektor) einer ebenen Kurve.

Zur Definition 7.12 des Normalenfelds: Anschaulich macht man sich diese Definition anhand der Abbildung 7.12 deutlich. Der Geschwindigkeitsvektor $c'(t)$ wird um 90 Grad im mathematisch positiven Sinne gedreht. Dies spiegelt auch die Definition 7.12 wider, denn die dort angegebene Matrix ist gerade eine Drehung um 90 Grad gegen den Uhrzeigersinn. Damit steht $n(t)$ senkrecht auf $c'(t)$ und $(c'(t), n(t))$ bildet eine Orthonormalbasis des \mathbb{R}^2 .

Zur Definition 7.13 der Krümmung einer ebenen Kurve: Was versteht man anschaulich unter der Krümmung einer ebenen Kurve? Dies ist eigentlich recht einfach: *Die Krümmung ist ein Maß dafür, wie stark die Kurve von einer Geraden abweicht.* Wir wollen die Formel (7.2) für die Krümmung einer Kurve in der Ebene herleiten. Sei dazu $c : I \rightarrow \mathbb{R}^2$ eine ebene, nach Bogenlänge parametrisierte Kurve. Da c eben nach Bogenlänge parametrisiert ist, gilt $\langle c'(t), c'(t) \rangle = 1$. Wenn wir diese Gleichung auf beiden Seiten differenzieren, so ergibt dies

$$\langle c''(t), c'(t) \rangle + \langle c'(t), c''(t) \rangle = 2 \langle c'(t), c''(t) \rangle = 0.$$

Demnach steht $c'(t)$ senkrecht auf $c''(t)$. Also ist $c''(t)$ ein Vielfaches des Normalenvektors $n(t)$. Es gilt:

$$c''(t) = \kappa(t) \cdot n(t).$$

Wie wir oben schon geschrieben haben, ist die Krümmung ein Maß dafür, wie stark die Kurve von einer Geraden abweicht. Wenn c eine nach Bogenlänge

parametrisierte Kurve ist, so ist c genau dann eine Gerade, wenn $c''(t) = 0$, das heißt wenn $\kappa(t) \equiv 0$. Im Zweidimensionalen ist es möglich, eine Krümmung positiv oder negativ zu nennen. Wir fragen uns jetzt, wann eine Krümmung positiv und wann negativ genannt wird. Die Krümmung heißt *positiv*, wenn sich die Kurve in Richtung des Normalenvektors krümmt, also in Durchlaufrichtung nach links. Sie ist *negativ*, wenn sie nach rechts gekrümmt ist.

1. Fall: Die Krümmung ist positiv. Es gilt $\kappa(t) > 0$.

2. Fall: Die Krümmung ist Null. Es gilt $\kappa(t) = 0$.

3. Fall: Die Krümmung ist negativ. Es gilt $\kappa(t) < 0$.

Betrachten wir ein Beispiel, das wir ausführlich durchrechnen wollen.

Beispiel 71

Dazu sei die ebene Kurve $c(t) = (r \cos(t/r), r \sin(t/r))$ gegeben. Es gilt nun für den Geschwindigkeitsvektor und die zweite Ableitung:

$$\begin{aligned} c'(t) &= \left(-r \cdot \frac{1}{r} \sin(t/r), r \cdot \frac{1}{r} \cos(t/r) \right) = (-\sin(t/r), \cos(t/r)), \\ c''(t) &= \frac{1}{r} (-\cos(t/r), -\sin(t/r)). \end{aligned}$$

Nun gilt doch offensichtlich

$$c'(t) \cdot c''(t) = 0.$$

Daher ist $c''(t) = \frac{1}{r} n(t)$. Also ist $\kappa(t) = \frac{1}{r}$ die Krümmung von c . Diese ist unabhängig vom Parameter t , also liegt eine konstante Krümmung vor. ■

Jetzt fragt sich der ein oder andere bestimmt, wie man die Krümmung für eine Kurve im \mathbb{R}^3 , also eine sogenannte Raumkurve, berechnen kann. Im Dreidimensionalen stehen wir jetzt aber vor einem Problem mit unserer Definition der Krümmung. Für eine räumliche Kurve ist der Normalenvektor nämlich nicht eindeutig. Wie wollen wir das Normalenfeld definieren? Es lässt sich zwar kein einzelner Normalenvektor bestimmen, dafür aber eine *Normalenebene*. Bei ebenen Kurven hatten wir zwei senkrecht stehende Normalenvektoren. Durch die Orientierung haben wir diesen dann eindeutig festgelegt. Aber wie machen wir das nun bei Raumkurven? Dort wird es schwer sein, über die Orientierung zu argumentieren. So viel sei verraten: Es ist auch hier möglich, vernünftig die Krümmung zu definieren, aber dies würde zu weit führen. Wie man dies macht und wie sich dies dann für Kurven im \mathbb{R}^n verallgemeinern lässt, entnehmt ihr bitte einer Vorlesung zur elementaren Differentialgeometrie oder lest in dem tollen Buch [B00] nach.

7.4 Erklärungen zu den Sätzen und Beweisen

Zum Satz 7.1: Der Satz 7.1 sagt einfach nur aus, dass die Umparametrisierung einer Kurve c die Eigenschaft der Regularität erhält.

Zum Satz 7.2: Diesen Satz lassen wir uns nochmal auf der Zunge zergehen: Jede regulär parametrisierte Kurve kann nach Bogenlänge umparametrisiert werden. Das ist ein sehr nützlicher Satz, wie wir später auch noch sehen werden. Man muss aber dazu sagen, dass es in der Realität nicht so leicht ist, die Umparametrisierungen zu finden.

Zum Satz 7.4: Dieser Satz besagt, dass die Definitionen 7.8 und 7.9 äquivalent sind und wir daher ganz einfach die Länge einer Kurve mittels des Integrals in Definition 7.9 berechnen können. Nur ist diese halt nicht so anschaulich. Man versteht nicht sofort, wieso dies gerade die Länge liefert. Daher haben wir zunächst mit Definition 7.8 angefangen. Aber woran liegt das? Die Idee ist, einfach erst einmal die Kurve durch einen Polygonzug zu approximieren und dessen Länge bei immer feinerer Approximation zu untersuchen. Verwenden wir die Notation aus Abbildung 7.9, so ist die Länge des Polygonzugs, also

$$L[P] = \sum_{i=1}^k |c(t_i) - c(t_{i-1})| \quad (7.3)$$

eine untere Schranke für die eigentliche Länge der Kurve. Fügt man weitere Punkte ein, so wird die neue Näherung wohl etwas größer werden, aber mit Sicherheit nicht kleiner. Dies liegt einfach an der Dreiecksungleichung. Konvergiert $L[P]$ nun, wenn $\sup |t_i - t_{i-1}|$ gegen Null strebt, so haben wir diesen Grenzwert als Länge der Kurve c definiert. Mit dem Mittelwertsatz können wir in der Summe aus Gleichung (7.3) sofort

$$|c(t_i) - c(t_{i-1})| = |\dot{c}(\xi_i)|(t_i - t_{i-1})$$

mit $\xi_i \in (t_{i-1} - t_i)$ setzen und erhalten die Länge im Grenzübergang einer unendlich feinen Unterteilung. Dies rechtfertigt Definition 7.9.

Zum Satz 7.5: Dieser Satz besagt, dass die Länge einer Kurve bei Umparametrisierungen erhalten bleibt. Wir können (Satz 7.2) also jede Kurve nach Bogenlänge parametrisieren, und damit vereinfacht sich das Integral bei der Längenberechnung erheblich, siehe auch die Erklärungen zu Definition 7.9.

Zum Satz 7.6: Dieser Satz gibt einfach nur eine weitere Formel, um die Krümmung einer Kurve im \mathbb{R}^2 zu berechnen, weil es mit unserer Definition 7.13 der Krümmung einer ebenen Kurve recht schwer sein kann, wie Beispiel 71 gezeigt hat.

8 Untermannigfaltigkeiten

Übersicht

8.1	Definitionen	154
8.2	Sätze und Beweise	155
8.3	Erklärungen zu den Definitionen	156
8.4	Erklärungen zu den Sätzen und Beweisen	161

In diesem Kapitel werden wir Untermannigfaltigkeiten, genauer Untermannigfaltigkeiten des \mathbb{R}^n , betrachten. Leider wird man damit den Mannigfaltigkeiten nicht ganz gerecht, denn diese sind wesentlich mehr als nur Untermannigfaltigkeiten vom \mathbb{R}^n . Daher sollte man auch immer die allgemeinere Version einer Mannigfaltigkeit, die komplett ohne den \mathbb{R}^n auskommt, im Kopf haben. Wir werden diese ebenfalls angeben, uns aber in dieser Ausführung, die ja nur einen Einblick in die Untermannigfaltigkeiten geben soll, auf den \mathbb{R}^n beschränken. Allen, die tiefer einsteigen wollen, sei eine Vorlesung zur Differentialgeometrie ans Herz gelegt.

Schon einmal so viel: Lokal können wir eine k -dimensionale Untermannigfaltigkeit des \mathbb{R}^n unter gewissen Voraussetzungen durch eine Parameterdarstellung mit k reellen Parametern, als Nullstellengebilde von $n - k$ unabhängigen differenzierbaren Funktion oder als Urbilder regulärer Werte darstellen. Wir werden uns jetzt alles in Ruhe anschauen.

Man fragt sich, wieso man Mannigfaltigkeiten überhaupt benötigt. Dies wird in diesem Übersichtsartikel leider nicht ganz so deutlich werden, weil wir kaum Anwendungen von Mannigfaltigkeiten geben, da dies den Rahmen des zweiten Semesters sprengen würde. Mannigfaltigkeiten braucht man aber nicht nur in der Differentialgeometrie (die Heimat von differenzierbaren Mannigfaltigkeiten), sondern auch in der theoretischen Physik und in vielen anderen Bereichen. In einem dritten Band unseres Tutoriums („Tutorium höhere Analysis“, Erscheinungstermin: Mai 2012) werden wir einige Anwendungen von Mannigfaltigkeiten in Bezug auf wichtige Integralsätze, wie den Integralsätzen von Gauß oder Stokes, kennenlernen. Wir werden uns dort dann anschauen, wie man über Mannigfaltigkeiten integriert. Zunächst einmal reicht es für euch aber, wenn ihr euch Mannigfaltigkeiten als Teilmengen des \mathbb{R}^n vorstellt.

8.1 Definitionen

Definition 8.1 (Immersion)

Sei $U \subset \mathbb{R}^k$ eine offene Menge. Wir nennen eine stetig differenzierbare Abbildung

$$\varphi : U \rightarrow \mathbb{R}^n, (u_1, \dots, u_k) \mapsto \varphi(u_1, \dots, u_k)$$

eine **Immersion**, wenn der Rang der Funktionalmatrix

$$D\varphi = \left(\frac{\partial \varphi_i}{\partial u_j} \right)_{1 \leq i \leq n, 1 \leq j \leq k} =: \frac{\partial(\varphi_1, \dots, \varphi_n)}{\partial(u_1, \dots, u_k)}$$

in jedem Punkt $u \in U$ gleich k ist.

Definition 8.2 (Untermannigfaltigkeit, Kodimension)

Eine Teilmenge $M \subset \mathbb{R}^n$ heißt k -dimensionale **Untermannigfaltigkeit** von \mathbb{R}^n , wenn zu jedem Punkt $p \in M$ eine offene Umgebung $V \subset \mathbb{R}^n$, sowie eine offene Teilmenge $U \subset \mathbb{R}^k$ und eine Immersion $\varphi : U \rightarrow \mathbb{R}^n$ existieren, sodass φ die Menge U homöomorph auf $\varphi(U) = M \cap V$ abbildet. $\varphi : U \rightarrow M \cap V$ heißt dann **Parameterdarstellung** oder **lokales Koordinatensystem** der Untermannigfaltigkeit in einer Umgebung von p . Ist $(u_1, \dots, u_k) \in U$, so heißen u_1, \dots, u_k die lokalen Koordinaten. Die Zahl $n - k$ heißt **Kodimension** der Untermannigfaltigkeit.

Anmerkung: Untermannigfaltigkeiten mit Kodimension 1 heißen **Hyperflächen**.

Definition 8.3 (Tangentialvektor, Tangentialraum)

Seien $M \subset \mathbb{R}^n$ eine Untermannigfaltigkeit und $p \in M$. Ein Vektor $v \in \mathbb{R}^n$ heißt **Tangentialvektor** an M im Punkt p , wenn eine stetig differenzierbare Kurve

$$\gamma : (-\varepsilon, \varepsilon) \rightarrow M \subset \mathbb{R}^n, \varepsilon > 0,$$

existiert mit

$$\gamma(0) = p \quad \text{und} \quad \gamma'(0) = v.$$

Die Menge aller Tangentialvektoren an M in p heißt **Tangentialraum** und wird mit $T_p M$ bezeichnet.

Definition 8.4 (Normalenvektor, Normalenraum)

Die Voraussetzungen seien wie in Definition 8.3. Ein **Normalenvektor** von M im Punkt $p \in M$ ist ein Vektor $w \in \mathbb{R}^n$, der auf allen Tangentialvektoren $v \in T_p M$ senkrecht steht. Die Menge aller Normalenvektoren heißt **Normalenraum**, und wir bezeichnen ihn mit $N_p M$.

Definition 8.5 (regulärer Punkt, regulärer Wert)

Ein Punkt $x \in U \subset \mathbb{R}^n$ heißt ein **regulärer Punkt** der differenzierbaren Abbildung $f : U \rightarrow Y$, wenn das Differential $df(x) : X \rightarrow Y$ surjektiv abbildet. Ferner heißt ein Punkt $y \in Y$ ein **regulärer Wert** von f , wenn alle $x \in f^{-1}(y)$ reguläre Punkte sind. Bildet $df(x)$ nicht surjektiv ab, so heißt x ein **singulärer Punkt** und $f(x)$ ein **singulärer Wert** von f .

8.2 Sätze und Beweise

Satz 8.1

Eine Teilmenge $M \subset \mathbb{R}^n$ ist genau dann eine k -dimensionale Untermannigfaltigkeit, wenn eine der folgenden äquivalenten Bedingungen erfüllt ist.

- i) Zu jedem Punkt $p \in M$ gibt es eine offene Umgebung $V \subset \mathbb{R}^n$ und $n - k$ stetig differenzierbare Funktionen $f_i : V \rightarrow \mathbb{R}$, $i = 1, \dots, n - k$, sodass

$$M \cap V = \{x \in V : f_1(x) = \dots = f_{n-k}(x) = 0\}$$

und der Rang von $\frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_1, \dots, x_n)}$ gerade $n - k$ für alle $x \in M \cap V$ ist.

- ii) Zu jedem Punkt $p \in M$ gibt es (nach eventueller Umnummerierung der Koordinaten) offene Umgebungen $U' \subset \mathbb{R}^k$ von $p' := (p_1, \dots, p_k)$ und $U'' \subset \mathbb{R}^{n-k}$ von $p'' := (p_{k+1}, \dots, p_n)$ und eine stetig differenzierbare Abbildung $g : U' \rightarrow U'' \subset \mathbb{R}^{n-k}$, sodass

$$M \cap (U' \times U'') = \{(x', x'') \in U' \times U'' : x'' = g(x')\}.$$

- iii) Zu jedem Punkt $p \in M$ existiert eine offene Umgebung $V \subset \mathbb{R}^n$, eine offene Menge $U \subset \mathbb{R}^n$ und eine C^1 -invertierbare Abbildung $\Phi : V \rightarrow U$, sodass

$$\Phi(M \cap V) = E_k \cap U.$$

Dabei ist $E_k \subset \mathbb{R}^n$ die k -dimensionale Ebene

$$E_k = \{x = (x_1, \dots, x_n) \in \mathbb{R}^n : x_{k+1} = \dots = x_n = 0\}.$$

Satz 8.2 (Satz vom regulären Wert)

Seien $U \subset \mathbb{R}^n$ und $f : U \rightarrow \mathbb{R}^k$ eine Abbildung. Ist a ein regulärer Wert von f , so ist $M := f^{-1}(a)$ eine $(n - k)$ -dimensionale Untermannigfaltigkeit des \mathbb{R}^n .

Satz 8.3 (Eigenschaften des Tangential- und Normalenraums)

Seien $M \subset \mathbb{R}^n$ eine k -dimensionale Untermannigfaltigkeit und $p \in M$ ein Punkt. Dann gelten die folgenden beiden Aussagen für Tangential- bzw. Normalenraum.

- i) Der Tangentialraum $T_p M$ ist ein k -dimensionaler Vektorraum. Eine Basis ergibt sich wie folgt: Sei $\varphi : U \rightarrow M \subset \mathbb{R}^n$ ein lokales Koordinatensystem von M in einer Umgebung von p . Sei weiter $u^* \in U$ ein Punkt mit $\varphi(u^*) = p$. Dann bilden die Vektoren

$$\frac{\partial \varphi}{\partial u_1}(u^*), \dots, \frac{\partial \varphi}{\partial u_k}(u^*)$$

eine Basis des Tangentialraums $T_p M$.

- ii) Der Normalenraum $N_p M$ ist ein $(n-k)$ -dimensionaler Vektorraum. Eine Basis des Normalenraums erhalten wir nun so: Seien $f_1, \dots, f_{n-k} : U \rightarrow \mathbb{R}$ stetig differenzierbare Funktionen in einer offenen Umgebung $V \subset \mathbb{R}^n$ von p mit

$$M \cap V = \{x \in V : f_1(x) = \dots = f_{n-k}(x) = 0\}.$$

Weiterhin sei der Rang von $\frac{\partial(f_1, \dots, f_{n-k})}{\partial(x_1, \dots, x_n)}(p)$ gerade $n-k$. Dann bilden die Vektoren

$$\text{grad} f_j(p), \quad j = 1, \dots, n-k$$

eine Basis des Normalenraums.

8.3 Erklärungen zu den Definitionen

Zur Definition 8.1 einer Immersion: Bevor wir uns ein paar Beispiele für Immersionen anschauen, noch einige Anmerkungen:

- Die Matrix $\left(\frac{\partial \varphi_i}{\partial t_j}\right)_{1 \leq i \leq n, 1 \leq j \leq k}$ in der Definition 8.1 ist eine $(n \times k)$ -Matrix. Es gilt natürlich notwendigerweise $n \geq k$, wenn der Rang gleich k sein soll. Man sagt auch der Rang der Matrix ist maximal.
- Die Bedingung, dass $\left(\frac{\partial \varphi_i}{\partial t_j}\right)$ maximalen Rang (also Rang k) besitzt, ist äquivalent dazu, dass die Vektoren

$$\frac{\partial \varphi}{\partial u_1}(u), \dots, \frac{\partial \varphi}{\partial u_k}(u) \in \mathbb{R}^n$$

linear unabhängig sind.

Beispiel 72

- Ist $k = 1$ zum Beispiel $U \subset \mathbb{R}$ ein offenes Intervall, dann ist eine Immersion $\varphi : U \rightarrow \mathbb{R}^n$ nichts anderes als eine reguläre Kurve.

- Sei $\varphi : \mathbb{R} \rightarrow \mathbb{R}^2$ gegeben durch

$$\varphi(u) = (u^2, u^3).$$

Diese Abbildung stellt keine Immersion dar, denn die Funktionalmatrix lautet

$$D\varphi(u) = \begin{pmatrix} 2u \\ 3u^2 \end{pmatrix}.$$

Sie besitzt im Nullpunkt $u = (0, 0)$ aber keinen maximalen Rang, sondern Rang Null.

- Sei nun $\varphi : \mathbb{R} \rightarrow \mathbb{R}^2$ gegeben durch

$$\varphi(u) = (u(u^2 - 1), u^2 - 1).$$

Dies ist eine Immersion, denn man kann sich leicht überlegen, dass die Funktionalmatrix

$$D\varphi(u) = \begin{pmatrix} 3u^2 - 1 \\ 2u \end{pmatrix}.$$

für alle u vollen Rang besitzt.

■

Zur Definition 8.2 einer Untermannigfaltigkeit: Erst einmal bemerken wir, dass man zeigen kann, dass so ein Homöomorphismus wirklich, wie in der Definition gefordert, existiert. Wir wollen an dieser Stelle aber auf diesen Nachweis verzichten und uns lieber ein paar Beispiele für Untermannigfaltigkeiten des \mathbb{R}^n anschauen. Zuvor soll die Abbildung 8.1 eine Untermannigfaltigkeit im \mathbb{R}^n verdeutlichen. Diese Abbildung 8.1 ist so zu verstehen: Wir nennen eine Teilmenge

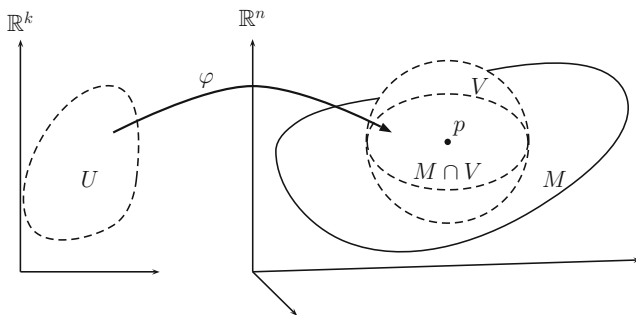


Abb. 8.1: So stellt man sich am besten eine Untermannigfaltigkeit im \mathbb{R}^n vor.

M des \mathbb{R}^n eine Untermannigfaltigkeit, wenn wir für jeden Punkt $p \in M$ eine offene Umgebung $V \subset \mathbb{R}^n$, eine offene Teilmenge $U \subset \mathbb{R}^k$ und ein $\varphi : U \rightarrow \mathbb{R}^n$ finden, welches eine Immersion sein muss, sodass φ die Menge U homöomorph

auf $\varphi(U) = M \cap V$ abbildet. Das φ ist dann eine Karte. Meistens reicht natürlich nicht nur eine Karte aus, um die gesamte Mannigfaltigkeit in den \mathbb{R}^n „abzubilden“. Man braucht mehrere Kartenabbildungen. Bei der Sphäre beispielsweise benötigt man zwei Karten, die sogenannte *stereographische Projektion*, siehe das nächste Beispiel.

Beispiel 73

- Wir betrachten die Sphäre mit Radius r im \mathbb{R}^3 . Wir erhalten die Sphäre anschaulich und rechnerisch durch die Rotation der Kreislinie

$$\alpha \mapsto \begin{pmatrix} x(\alpha) \\ z(\alpha) \end{pmatrix} = \begin{pmatrix} r \sin \alpha \\ r \cos \alpha \end{pmatrix}$$

um die z -Achse. Die zugehörige Parameterdarstellung lautet

$$(\alpha, \beta) \mapsto F(\alpha, \beta) = \begin{pmatrix} r \sin \alpha \cos \beta \\ r \sin \alpha \sin \beta \\ r \cos \alpha \end{pmatrix}.$$

Wir beschränken nun α auf den Bereich $0 < \alpha < \pi$ ein, damit DF den Rang 2 besitzt. Hierdurch werden der Nordpol $(0, 0, r)$ und der Südpol $(0, 0, -r)$ ausgeschlossen. (α, β) sind die Polarkoordinaten auf der Sphäre. Eine Ergänzung zum Rang von DF . Es gilt:

$$DF(\alpha, \beta) = \begin{pmatrix} r \cos \alpha \cos \beta & -r \sin \alpha \sin \beta \\ r \cos \alpha \sin \beta & r \sin \alpha \cos \beta \\ -r \sin \alpha & 0 \end{pmatrix}.$$

In dem oben genannten Bereich besitzt DF den Rang 2.

- Als neues Beispiel nehmen wir den Torus. Sei $R > r > 0$. Der Torus mit den Radien r und R entsteht durch Rotation der Kreislinie

$$t \mapsto \begin{pmatrix} x(t) \\ z(t) \end{pmatrix} = \begin{pmatrix} R + r \cos(t) \\ r \sin(t) \end{pmatrix}$$

um die z -Achse. Nun ist

$$(t, s) \mapsto \begin{pmatrix} (R + r \cos(t)) \cos(s) \\ (R + r \cos(t)) \sin(s) \\ r \sin(t) \end{pmatrix} =: F(t, s).$$

Das Bild $F(\mathbb{R} \times \mathbb{R})$ liefert nun den Torus und damit handelt es sich um eine Immersion, da der Kreis die z -Achse nicht schneidet.

■

Anmerkung: Wie im Einführungstext zu diesem Kapitel angedeutet, kann man eine Mannigfaltigkeit (diese nennt man dann eine topologische Mannigfaltigkeit) ohne den umgebenden Raum \mathbb{R}^n definieren. Dies geht so:

Eine *topologische Mannigfaltigkeit* M der Dimension m ist ein zusammenhängender, parakompakter Hausdorff-Raum, welcher lokal euklidisch ist. Lokal euklidisch heißt, dass zu jedem Punkt $p \in M$ eine Umgebung U existiert mit $p \in U$, eine offene Umgebung $\Omega \subset \mathbb{R}^m$ und ein Homöomorphismus $x : U \rightarrow \Omega$. Jedes $x : U \rightarrow \Omega$ heißt eine **lokale Karte (lokales Koordinatensystem)** von M um p . Eine topologische Mannigfaltigkeit kann man sich so vorstellen wie in Abbildung 8.2.

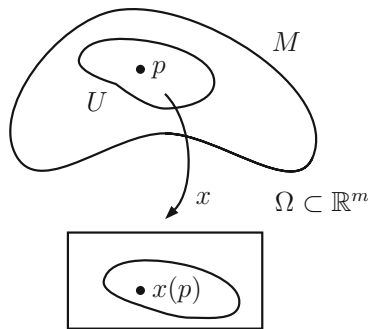


Abb. 8.2: Eine topologische Mannigfaltigkeit der Dimension m entspricht lokal dem \mathbb{R}^m . Wir sagen, M ist lokal euklidisch, sieht lokal also wie der \mathbb{R}^m aus.

Wir geben einige Anmerkungen, bevor wir zu konkreten Beispielen kommen.

- Existiert zu jeder offenen Überdeckung $(\Omega_\alpha)_{\alpha \in I}$ von M eine lokal endliche Verfeinerung, das heißt eine lokal endliche Überdeckung $(\Omega'_\beta)_{\beta \in I'}$ mit

$$\forall \beta \in I' : \exists \alpha \in I : \Omega'_\beta \subset \Omega_\alpha,$$

so heißt M *parakompakt*.

- In vielen Büchern werden topologische Mannigfaltigkeiten etwas anders eingeführt. Dort wird nicht die Parakompaktheit von M gefordert, sondern, dass M als topologischer Raum eine abzählbare Basis besitzt. Da aber topologische Mannigfaltigkeiten mit abzählbarer Basis stets parakompakt sind, ist unser Begriff allgemeiner.
- Es gibt in der Differentialgeometrie auch noch den Begriff der *differenzierbaren Mannigfaltigkeit*. Um einen Punkt $p \in M$ einer Mannigfaltigkeit M gibt es ja durchaus mehrere offene Umgebungen und damit auch verschiedene Karten. Man kann aber mit einem sogenannten Kartenwechsel

zwischen den Karten hin und her springen. Grob gesprochen: Eine topologische Mannigfaltigkeit ist eine differenzierbare Mannigfaltigkeit, wenn die Kartenwechsel alle differenzierbar sind.

Beispiel 74

Wir geben ein paar elementare Beispiele für topologische Mannigfaltigkeiten an.

- Die elementarste topologische Mannigfaltigkeit ist der \mathbb{R}^m mit $\dim \mathbb{R}^m = m$ und \mathbb{R}^m mit der Standardtopologie versehen, die durch die Standardmetrik (euklidische Metrik) induziert wird.
- Sind M eine topologische Mannigfaltigkeit und $U \subset M$ offen und zusammenhängend, so ist auch U mit der Relativtopologie eine topologische Mannigfaltigkeit derselben Dimension.
- Die Sphäre $S^m \subset \mathbb{R}^{m+1}$ ist mit der Relativtopologie auch eine topologische Mannigfaltigkeit. Wir benötigen genau zwei Karten wie bei der S^2 , die zusammen die Sphäre S^m komplett überdecken. Diese bezeichnet man als *stereographische Projektion*.
- Sind M_1 und M_2 zwei topologische Mannigfaltigkeiten, so auch $M := M_1 \times M_2$ zusammen mit der Produkttopologie der Dimension $\dim M = \dim M_1 + \dim M_2$. Insbesondere ist wegen des obigen Beispiels der Sphäre auch der Torus

$$T^m := \underbrace{S^1 \times \dots \times S^1}_{m\text{-mal}}$$

eine topologische Mannigfaltigkeit der Dimension m .

■

Zur Definition 8.3 des Tangentialvektors: Ein Tangentialvektor an eine Untermannigfaltigkeit ist einfach nur ein Tangentialvektor einer in der Untermannigfaltigkeit verlaufenden Kurve. Die Abbildung 8.3 soll dies verdeutlichen.

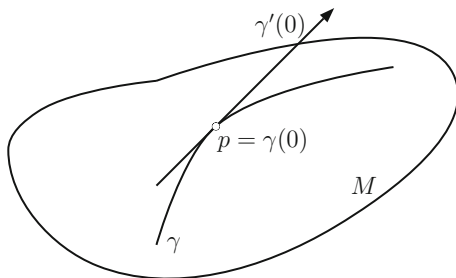


Abb. 8.3: Ein Tangentialvektor anschaulich.

Zur Definition 8.4 des Normalenvektors: Ein Normalenvektor ist einfach nur ein Vektor, der auf der Untermannigfaltigkeit senkrecht steht und das bzgl. der kanonisch gegebenen euklidischen Metrik im \mathbb{R}^n .

Beispiel 75

In Beispiel 79 (ganz am Ende des Kapitels) werden wir mithilfe des Satzes 8.2 vom regulären Wert zeigen, dass für

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = x^4 + y^4 - 4xy$$

$f^{-1}(1)$ eine eindimensionale Untermannigfaltigkeit des \mathbb{R}^2 ist. Wir wollen jetzt den Tangentialraum und den Normalenraum im Punkt $(1, 0)$ berechnen. Es gilt:

$$Df(x, y) = \nabla 4(x^3 - y, y^3 - x).$$

Daher ist

$$T_{(1,0)}M = \ker Df(1, 0) = \ker(4, -4) = \langle (1, 1) \rangle$$

und natürlich

$$N_{(1,0)}M = \langle (1, -1) \rangle,$$

denn der Vektor $(1, -1)^T$ ist offenbar (da das Skalarprodukt Null ist) senkrecht zum Vektor $(1, 1)^T$. In der Praxis bestimmt man genau so Tangential- bzw. Normalenraum. ■

Zur Definition 8.5 des regulären Wertes: Um die Definition zu verstehen, schauen wir uns ein Beispiel an. Weitere Beispiele folgen dann in den Erklärungen zum Satz 8.2 des regulären Wertes.

Beispiel 76

Übung: Betrachtet die Abbildung

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}^2, \quad f(x, y, z) = (x, y).$$

■

8.4 Erklärungen zu den Sätzen und Beweisen

Zum Satz 8.1: Dieser Satz zeigt, dass man Untermannigfaltigkeiten des \mathbb{R}^n nicht immer durch Parameterdarstellungen wie in Beispiel 73 darstellen muss, sondern es noch eine andere Möglichkeit gibt. Schauen wir uns an, wie wir uns das Kriterium zu Nutze machen können, um Untermannigfaltigkeiten zu entlarven.

Beispiel 77

- Die Funktionen $f_i : \mathbb{R}^4 \rightarrow \mathbb{R}$ mit $i = 1, 2, 3$ seien definiert durch

$$f_1(x_1, \dots, x_4) = x_1x_3 - x_2^2,$$

$$f_2(x_1, \dots, x_4) = x_2x_4 - x_3^2,$$

$$f_3(x_1, \dots, x_4) = x_1x_4 - x_2x_3.$$

Wir behaupten und beweisen, dass

$$M := \{x \in \mathbb{R}^4 \setminus \{0\} : f_1(x) = f_2(x) = f_3(x) = 0\}$$

eine 2-dimensionale Untermannigfaltigkeit des \mathbb{R}^4 ist.

Sei dazu $x = (x_1, \dots, x_4) \in M$. Mindestens eine der Konstanten x_i ist nach Definition ungleich Null. Wir wollen den Satz 8.1 anwenden und machen eine Fallunterscheidung.

1. Fall: Sei $x_1 \neq 0$. Dann gilt:

$$\begin{aligned} f_1(x) = 0 &\Rightarrow x_3 = \frac{x_2^2}{x_1}, \\ f_3(x) = 0 &\Rightarrow x_4 = \frac{x_2x_3}{x_1}, \\ &\Rightarrow x_2x_4 = \frac{x_2^2x_3}{x_1} = x_3^2 \Rightarrow f_2(x) = 0. \end{aligned}$$

Definieren wir also $U_1 := \{x \in \mathbb{R}^4 : x_1 \neq 0\}$, so gilt:

$$M \cap U_1 = \{x \in U_1 : f_1(x) = f_3(x) = 0\}.$$

Nun ist

$$\nabla f_1(x) = (x_3, -2x_2, x_1, 0), \quad \nabla f_3(x) = (x_4, -x_3, -x_2, x_1).$$

Da

$$\det \begin{pmatrix} x_1 & 0 \\ -x_2 & x_1 \end{pmatrix} = x_1^2 \neq 0,$$

sind ∇f_1 und ∇f_2 linear unabhängig. Daher ist M in einer Umgebung von x eine zweidimensionale Untermannigfaltigkeit.

2. Fall: Sei $x_2 \neq 0$. Wegen $f_1(x) = x_1x_3 - x_2^2 = 0$, folgt dann $x_1 \neq 0$, und wir sind im ersten Fall.
3. Fall: Sei $x_3 \neq 0$. Wegen $f_2(x) = x_2x_4 - x_3^2 = 0$, folgt dann $x_2 \neq 0$, also (2. Fall) auch $x_1 \neq 0$, und wir sind im ersten Fall.
4. Fall: Sei $x_4 \neq 0$. Dies geht analog wie im ersten Fall. Probiert euch daran!

■ Die Funktionen $f, g : \mathbb{R}^3 \rightarrow \mathbb{R}$ seien definiert durch

$$f(x, y, z) := x^2 + xy - y - z, \quad g(x, y, z) := 2x^2 + 3xy - 2y - 3z.$$

Wir behaupten und zeigen:

$$C := \{(x, y, z) \in \mathbb{R}^3 : f(x, y, z) = g(x, y, z) = 0\}$$

ist eine eindimensionale Untermannigfaltigkeit des \mathbb{R}^3 und

$$\varphi : \mathbb{R} \rightarrow \mathbb{R}^3, \varphi(t) = (t, t^2, t^3)$$

ist eine globale Parameterdarstellung von C . Wir definieren

$$C_1 := \{(t, t^2, t^3) : t \in \mathbb{R}\}.$$

„ $C_1 \subset C$ “: Da $f(t, t^2, t^3) = 0$ und $g(t, t^2, t^3) = 0$ für alle $t \in \mathbb{R}$ gilt, folgt sofort $C_1 \subset C$.

„ $C \subset C_1$ “: Es ist $3f(x, y, z) - g(x, y, z) = x^2 - y$ und $2f(x, y, z) - g(x, y, z) = -xy + z$. Für jeden Punkt $(x, y, z) \in C$ gilt also $y = x^2$ und $z = xy = x^3$. Der Punkt hat also die Gestalt (x, x^2, x^3) , das heißt, er liegt in C_1 , also gilt auch $C \subset C_1$.

Insgesamt ist demnach $C_1 = C$. Die Gradienten

$$\nabla f(x, y, z) = (2x + y, x - 1, -1), \quad \nabla g(x, y, z) = (4x + 3y, 3x - 2, -3)$$

sind in jedem Punkt $(x, y, z) \in C$ linear unabhängig, also ist C eine eindimensionale Untermannigfaltigkeit des \mathbb{R}^3 . ■

Zum Satz vom regulären Wert (Satz 8.2): Wir wollen den wichtigen Satz vom regulären Wert an einem Beispiel verdeutlichen.

Beispiel 78

Sei

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := (x^2 + y^2 + 2x)^2 - 4(x^2 + y^2).$$

Für welche $a \in \mathbb{R}$ ist $f^{-1}(a)$ eine Untermannigfaltigkeit?

Wir berechnen zunächst den Gradienten von f , weil wir nach kritischen bzw. regulären Punkten von f suchen. Es gilt:

$$\nabla f(x, y) = \left(2(x^2 + y^2 + 2x)(2x + 2) - 8x, 4(x^2 + y^2 + 2x)y - 8y \right).$$

Ist (x, y) ein kritischer Punkt von f , das heißt $\nabla f(x, y) = 0$, so ist $4(x^2 + y^2 + 2x)y - 8y = 0$, also $y = 0$ oder $x^2 + y^2 + 2x = 2 \Leftrightarrow (x^2 + y^2 + 2x) - 2 = 0$. Der zweite Fall $((x^2 + y^2 + 2x) - 2 = 0)$ führt wegen $2(x^2 + y^2 + 2x)(2x + 2) - 8x = 8 \neq 0$ zu einem Widerspruch. Der erste Fall ($y = 0$) ergibt $x = 0$ oder $x = -3$. Wir erhalten also die beiden kritischen Punkte $(0, 0)$ und $(-3, 0)$. Wegen $f(0, 0) = 0$ und $f(-3, 0) = -27$ ist somit jedes $a \neq 0, -27$ ein regulärer Wert (kein kritischer Wert) und somit ist für jedes solche a die Niveaumenge $f^{-1}(a)$ nach dem Satz 8.2 eine Untermannigfaltigkeit. ■

Beispiel 79

Sei

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = x^4 + y^4 - 4xy$$

Wir behaupten, dass $f^{-1}(1)$ eine eindimensionale Untermannigfaltigkeit des \mathbb{R}^2 ist. Dies geht natürlich mit dem Satz vom regulären Wert. Dazu berechnen wir zunächst die kritischen Punkte von f . Wegen

$$Df(x, y) = \nabla 4(x^3 - y, y^3 - x)$$

ist $Df(x, y) = (0, 0)$ genau dann, wenn

$$(x, y) \in \{(0, 0), (1, 1), (-1, -1)\}.$$

Da $f(0, 0) = 0$, $f(1, 1) = f(-1, -1) = -2$ ist, ist $a = 1$ ein regulärer Wert und folglich ist $f^{-1}(1)$ eine eindimensionale Untermannigfaltigkeit des \mathbb{R}^2 . ■

9 Euklidische und unitäre Vektorräume

Übersicht

9.1	Definitionen	165
9.2	Sätze und Beweise	166
9.3	Erklärungen zu den Definitionen	167
9.4	Erklärungen zu den Sätzen und Beweisen	170

In diesem Kapitel wollen wir unser Wissen über euklidische Vektorräume aus der Linearen Algebra 1 auffrischen und auf sogenannte unitäre Vektorräume verallgemeinern. Dies sind, grob gesprochen, Vektorräume, die mit einem komplexen Skalarprodukt ausgestattet sind. Was ein komplexes Skalarprodukt ist, werden wir uns jetzt anschauen.

9.1 Definitionen

Definition 9.1 (Standardskalarprodukt, euklidischer Vektorraum)

Sei $V = \mathbb{R}^n$ der Standardvektorraum der Dimension $n \in \mathbb{N}$ über den reellen Zahlen \mathbb{R} . Das **Standardskalarprodukt** auf V ist die Abbildung

$$V \times V \rightarrow \mathbb{R}, (x, y) \mapsto \langle x, y \rangle := x^T \cdot y = \sum_{i=1}^n x_i \cdot y_i.$$

Das Paar $(V, \langle \cdot, \cdot \rangle)$ heißt **euklidischer Vektorraum**.

Anmerkung: Diese Definition werden wir für die Definition eines allgemeinen Skalarprodukts heranziehen.

Definition 9.2 (Länge eines Vektors)

Sei $x \in V$ ein Vektor aus unserem Standardvektorraum. Die **Länge des Vektors** definieren wir als

$$||x|| := \sqrt{\langle x, x \rangle}.$$

Definition 9.3 (Winkel zwischen Vektoren)

Seien $x, y \in V$, $x, y \neq 0$ zwei Vektoren aus V . Dann ist der **Winkel** α , den die beiden Vektoren einschließen, gegeben durch

$$\cos \alpha = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}.$$

Definition 9.4 (Standardskalarprodukt im Komplexen)

Sei $V = \mathbb{C}^n$ der Standardvektorraum über \mathbb{C} der Dimension $n \in \mathbb{N}$. Das (**hermitesche**) **Standardskalarprodukt** zweier Vektoren $x, y \in V$ ist definiert als die komplexe Zahl

$$\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}, \quad (x, y) \mapsto \bar{x}^T \cdot y = \sum_{i=1}^n \bar{x}_i \cdot y_i.$$

9.2 Sätze und Beweise

Satz 9.1

Für $x, y, z \in V = \mathbb{C}^n$, $\lambda \in \mathbb{C}$ gelten die folgenden Aussagen:

Linearität im 2. Argument:

$$\langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle, \quad \langle x, \lambda \cdot y \rangle = \lambda \cdot \langle x, y \rangle.$$

Semilinearität im 1. Argument:

$$\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle, \quad \langle \lambda \cdot x, y \rangle = \bar{\lambda} \langle x, y \rangle.$$

Hermiteische Symmetrie:

$$\langle y, x \rangle = \overline{\langle x, y \rangle}.$$

Positive Definitheit:

$$\langle x, x \rangle > 0 \quad \forall x \neq 0 \quad \text{und} \quad \langle x, x \rangle = 0 \Leftrightarrow x = 0.$$

9.3 Erklärungen zu den Definitionen

Zur Definition 9.1 des Standardskalarprodukts: Das Standardskalarprodukt kennt ihr schon aus der Schule. Die Eigenschaften aus Satz Vektoren $x, y \in V = \mathbb{R}^n$ seien gegeben durch

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{bzw.} \quad y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}.$$

So erhalten wir sofort die gewöhnliche Definition

$$\langle x, y \rangle = x^T \cdot y.$$

Dass wir den Vektor x transponieren müssen, angedeutet durch das T , liegt einfach daran, dass wir eine Matrixmultiplikation durchführen. Wenn wir zwei Vektoren aus dem \mathbb{R}^n gegeben haben, würden wir eine $(n \times 1)$ -Matrix mit einer $(n \times 1)$ -Matrix multiplizieren. Dies geht nach Definition der Matrixmultiplikation aber nicht, denn die Spaltenanzahl der ersten Matrix muss mit der Zeilenanzahl der zweiten Matrix übereinstimmen. Transponieren wir aber die erste Matrix (in diesem Fall ist dies ein Vektor), so multiplizieren wir eine $(1 \times n)$ -Matrix mit einer $(n \times 1)$ -Matrix, was wiederum ohne Probleme funktioniert. Daher definieren wir das Standardskalarprodukt gerade so, wie wir dies getan haben.

Beispiel 80

Wir wollen die beiden Vektoren

$$x = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

miteinander multiplizieren. Wir wissen nun, wie das geht, und zwar so:

$$\langle x, y \rangle = x^T \cdot y = 1 \cdot 1 + (-1) \cdot 1 + 0 \cdot 0 = 1 - 1 = 0.$$

Dies bedeutet übrigens, wie ihr euch bestimmt erinnert, dass die beiden Vektoren senkrecht aufeinander stehen, wie wir auch in den Erklärungen zu Definition 9.3 sehen bzw. nachrechnen werden.

Einige von euch denken jetzt vielleicht, dass dies doch alles nichts Neues ist. Ja, das ist richtig. Rufen wir uns noch einmal die Eigenschaften des Standardskalarprodukts in Erinnerung. So werden wir in Kapitel 10 sehen, dass man durchaus viele weitere Skalarprodukte definieren kann. Dies ist so ähnlich wie bei den Metriken oder Topologien aus Kapitel 1. Es gibt halt viele, auch recht komplizierte Skalarprodukte, von denen wir uns später einige anschauen werden. Unser Standardskalarprodukt hat die folgenden Eigenschaften.

Bilinearität: Für alle $u, v, w \in V$ und $\lambda \in \mathbb{C}$ gilt:

$$\begin{aligned}\langle u + v, w \rangle &= \langle u, w \rangle + \langle v, w \rangle, \\ \langle \lambda \cdot v, w \rangle &= \lambda \cdot \langle v, w \rangle, \\ \langle u, v + w \rangle &= \langle u, v \rangle + \langle u, w \rangle, \\ \langle u, \lambda \cdot w \rangle &= \lambda \cdot \langle v, w \rangle.\end{aligned}$$

Symmetrie: Für alle $v, w \in V$ gilt:

$$\langle v, w \rangle = \langle w, v \rangle.$$

Definitheit: Für alle $v \in V$, $v \neq 0$ gilt:

$$\langle v, v \rangle > 0.$$

■

Zur Definition 9.2 der Länge eines Vektors: Vektoren haben eine gewisse Länge, und auch diese können wir mittels des Standardskalarprodukts ausrechnen. Die Länge von Vektoren bildet im Sinne der Definition 1.26 eine Norm.

Beispiel 81

Wir berechnen die Länge der Vektoren aus Beispiel 80. Wir erhalten

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{1^2 + (-1)^2 + 0^2} = \sqrt{2}$$

und

$$\|y\| = \sqrt{\langle y, y \rangle} = \sqrt{1^2 + 1^2 + 0^2} = \sqrt{2}$$

Noch eine Anmerkung: Die Norm, wie sie hier steht, wird natürlich durch das Skalarprodukt induziert. Dies sollte man immer im Hinterkopf behalten. ■

Zur Definition 9.3 des Winkels zwischen zwei Vektoren: Zwei Vektoren schließen einen Winkel ein, wie man sich anschaulich sehr leicht überlegt. Naja, wenn wir im \mathbb{R}^2 oder \mathbb{R}^3 sind. Im \mathbb{R}^{31} versagt die Vorstellung aber. Wir haben gesehen, dass das Standardskalarprodukt der beiden Vektoren aus Beispiel 80 Null ist und gesagt, dass dies gerade bedeuten soll, dass die Vektoren senkrecht aufeinander stehen. Dies errechnen wir so:

$$\cos \alpha = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} = \frac{0}{\sqrt{2} \cdot \sqrt{2}} = 0 \Rightarrow \alpha = (2k + 1) \cdot \frac{\pi}{2}, \quad k \in \mathbb{Z}.$$

Zur Definition 9.4 des Standardskalarprodukts im \mathbb{C}^n : Zunächst erinnern wir uns an die komplexen Zahlen. So wie wir in [MK09] definiert haben, ist

$$\mathbb{C} = \{z = x + i \cdot y : x, y \in \mathbb{R}\}$$

der Körper der komplexen Zahlen mit $i^2 = -1$. \mathbb{C} können wir mit dem \mathbb{R}^2 identifizieren. Damit kann jede komplexe Zahl als ein Vektor im $\mathbb{C} \cong \mathbb{R}^2$ aufgefasst werden und folglich besitzt jede komplexe Zahl eine Länge. Wir nennen dies den *Absolutbetrag* und definieren mit $z = x + i \cdot y$

$$|z| := \sqrt{x^2 + y^2} \in \mathbb{R}_{\geq 0}.$$

Das *komplex Konjugierte* einer komplexen Zahl $z = x + i \cdot y$ ist

$$\bar{z} := x - i \cdot y \in \mathbb{C}.$$

Es gelten die folgenden Rechenregeln für zwei komplexe Zahlen $z, w \in \mathbb{C}$

$$|z \cdot w| = |z| \cdot |w|, \quad \overline{z + w} = \bar{z} + \bar{w}, \quad \overline{z \cdot w} = \bar{z} \cdot \bar{w}, \quad |z|^2 = \bar{z} \cdot z. \quad (9.1)$$

Wir wollen nun ein Beispiel betrachten, um das Standardskalarprodukt einzüben.

Beispiel 82

Gegeben seien die beiden Vektoren

$$x = \begin{pmatrix} 1 \\ i \\ 2 \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} -i \\ 1 \\ -1 \end{pmatrix}.$$

Es gilt nun

$$\begin{aligned} \langle x, y \rangle &= \overline{x}^T \cdot y = \overline{\begin{pmatrix} 1 \\ i \\ 2 \end{pmatrix}}^T \cdot \begin{pmatrix} -i \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} \bar{1} \\ \bar{i} \\ \bar{2} \end{pmatrix}^T \cdot \begin{pmatrix} -i \\ 1 \\ -1 \end{pmatrix} \\ &= \begin{pmatrix} 1 \\ -i \\ 2 \end{pmatrix}^T \cdot \begin{pmatrix} -i \\ 1 \\ -1 \end{pmatrix} = 1 \cdot (-i) + (-i) \cdot 1 + 2 \cdot (-1) \\ &= -2i - 2 = -2(1 + i). \end{aligned}$$

Hierbei bedeutet der Strich über dem Vektor, dass wir jeden Eintrag komplex konjugieren. ■

9.4 Erklärungen zu den Sätzen und Beweisen

Zum Satz 9.1 über die Eigenschaften des komplexen Standardskalarprodukts:

Diese Eigenschaften folgen sofort aus den bekannten Gleichungen (9.1). Die Eigenschaften aus Satz 9.1 sind so ähnlich wie die, die auch für das Standardskalarprodukt aus Definition 9.1 gelten, nur dass wir hier wegen der Definition 9.4 des Standardskalarprodukts im Komplexen ab und an mit dem komplex Konjugierten arbeiten müssen. Auch dieses Skalarprodukt kann man verallgemeinern. Wir werden in Kapitel 10 sehen, dass ein Skalarprodukt, das dieser Eigenschaften genügt, eine hermitesche Form genannt wird, siehe dazu Definition 10.11. Also auch das Standardskalarprodukt im Komplexen kann auf andere schöne Skalarprodukte verallgemeinert werden.

Noch eine Anmerkung: In vielen Büchern, beispielsweise in [Fis08], wird das Standardskalarprodukt durch die Formel

$$\langle x, y \rangle := \sum_{i=1}^n x_i \cdot \bar{y}_i$$

definiert. Mit dieser Definition erhält man in Satz 9.1 die Linearität im 1. Argument und die Semilinearität im 2. Argument.

Die positive Definitheit erlaubt es uns, auf $V = \mathbb{C}^n$ eine Norm durch

$$||x|| := \sqrt{\langle x, x \rangle} = \sqrt{|x_1|^2 + \dots + |x_n|^2} \in \mathbb{R}_{\geq 0}$$

zu definieren. Ebenso macht es durchaus Sinn, zwei Vektoren $x, y \in V$ orthogonal zu nennen, wenn $\langle x, y \rangle = 0$. Wegen der hermiteschen Symmetrie ist dies eine symmetrische Relation. Dies geht natürlich auch für das reelle Skalarprodukt.

10 Bilinearformen und hermitesche Formen

Übersicht

10.1 Definitionen	171
10.2 Sätze und Beweise	176
10.3 Erklärungen zu den Definitionen	183
10.4 Erklärungen zu den Sätzen und Beweisen	198

Das Standardskalarprodukt, siehe Definition 9.1 für euklidische Vektorräume bzw. Definition 9.4 für unitäre Vektorräume, kann man auf die sogenannten Bilinearformen im euklidischen und auf die hermiteschen Formen im komplexen Fall verallgemeinern. Unser erstes wichtiges Ergebnis wird sein, dass jeder euklidische Vektorraum V eine Orthonormalbasis besitzt. Identifiziert man V mit \mathbb{R}^n vermöge so einer Orthogonal- oder Orthonormalbasis, so identifiziert sich das Skalarprodukt auf V mit dem Standardskalarprodukt auf dem \mathbb{R}^n . Analog werden wir sehen, dass jeder unitäre Vektorraum eine Orthonormalbasis besitzt.

10.1 Definitionen

Definition 10.1 (Bilinearform)

Sei V ein Vektorraum über einem beliebigen Körper K . Eine **Bilinearform** auf V ist eine Abbildung

$$f : V \times V \rightarrow K,$$

die die folgenden Axiome erfüllt:

$$f(u + v, w) = f(u, w) + f(v, w),$$

$$f(\lambda \cdot v, w) = \lambda \cdot f(v, w),$$

$$f(u, v + w) = f(u, v) + f(u, w),$$

$$f(v, \lambda \cdot w) = \lambda \cdot f(v, w)$$

für alle $u, v, w \in V$ und $\lambda \in K$.

Anmerkung: Ab und an schreiben wir für $f(v, w)$ auch $\langle v, w \rangle$.

Definition 10.2 (Symmetrische Bilinearform)

Eine Bilinearform heißt **symmetrisch**, wenn zusätzlich zu den Eigenschaften aus Definition 10.1 noch

$$f(v, w) = f(w, v)$$

für alle $v, w \in V$ gilt.

Anmerkung: Näheres zu den symmetrischen Bilinearformen gibt es in Kapitel 13 über die Quadriken.

Definition 10.3 (darstellende Matrix)

Seien V ein endlichdimensionaler K -Vektorraum, $f : V \times V \rightarrow K$ eine Bilinearform und $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V . Die **Darstellungsmatrix** oder auch **darstellende Matrix** ist die Matrix

$$A = \mathcal{M}_{\mathcal{B}}(f) = (a_{ij}) \in \mathcal{M}_{n,n}(K)$$

mit Einträgen

$$a_{ij} := f(v_i, v_j).$$

Anmerkung: Wir werden sehen, dass der Zusammenhang $f(x, y) = x^T \cdot A \cdot y$ bzgl. der Darstellungsmatrix existiert.

Definition 10.4 (Orthonormalbasis)

Seien V ein endlichdimensionaler K -Vektorraum der Dimension n und f eine Bilinearform auf V . Eine Basis $\mathcal{B} = (v_1, \dots, v_n)$ heißt **orthogonal** bezüglich f , wenn

$$\langle v_i, v_j \rangle = 0 \quad \forall i \neq j.$$

Die Basis \mathcal{B} ist eine **Orthonormalbasis**, wenn

$$\langle v_i, v_j \rangle = \delta_{ij},$$

wobei δ_{ij} das bekannte Kronecker-Delta bezeichnet, welches 1 ist, wenn $i = j$ und 0 ist, wenn $i \neq j$.

Definition 10.5 (positiv definit)

Ein symmetrische Bilinearform $f : V \times V \rightarrow \mathbb{R}$ auf einem reellen Vektorraum V heißt **positiv definit**, wenn für jeden nicht verschwindenden Vektor $v \in V$ gilt:

$$f(v, v) > 0.$$

Eine symmetrische Matrix $A \in \mathcal{M}_{n,n}(\mathbb{R})$ heißt **positiv definit**, wenn die zugehörige Bilinearform auf \mathbb{R}^n positiv definit ist, das heißt, wenn für alle $x \in \mathbb{R}^n$, $x \neq 0$ gilt:

$$x^T \cdot A \cdot x > 0.$$

Definition 10.6 (euklidischer Vektorraum, Norm)

Ein **euklidischer Vektorraum** ist ein endlichdimensionaler reeller Vektorraum V zusammen mit einer symmetrischen, positiv definiten Bilinearform $f : V \times V \rightarrow \mathbb{R}$.

Die **Norm** eines Vektors $v \in V$ in einem euklidischen Vektorraum ist definiert durch

$$||v|| := \sqrt{\langle v, v \rangle}.$$

Definition 10.7 (Summe, direkte Summe)

Sei V ein K -Vektorraum und $U_1, U_2 \subset V$ seien zwei Untervektorräume. Die **Summe** von U_1 und U_2 ist der von $U_1 \cup U_2$ aufgespannte Untervektorraum von V

$$U = U_1 + U_2 =: \langle U_1 \cup U_2 \rangle.$$

Oder anders formuliert:

$$U_1 + U_2 := \{u_1 + u_2 : u_1 \in U_1, u_2 \in U_2\}.$$

Die Summe $U = U_1 + U_2$ ist eine **direkte Summe**, in Zeichen

$$U = U_1 \oplus U_2,$$

wenn $U_1 \cap U_2 = \{0\}$.

Definition 10.8 (Komplement)

Seien V ein K -Vektorraum und $U \subset V$ ein Untervektorraum. Ein **Komplement** zu U in V ist ein Untervektorraum $W \subset V$ mit

$$V = U \oplus W.$$

Dabei ist also $U \cap W = \{0\}$, und $U + W$ besteht aus den Elementen

$$\{u + w : u \in U, w \in W\}.$$

Definition 10.9 (orthogonales Komplement)

Sei V ein K -Vektorraum, auf dem eine symmetrische (oder alternierende) Bilinearform (oder hermitesche Sesquilinearform) $\langle \cdot, \cdot \rangle$ gegeben ist. Für einen Unterraum $U \subset V$ schreiben wir

$$U^\perp := \{v \in V : \forall u \in U : \langle u, v \rangle = 0\}$$

und nennen dies das **orthogonale Komplement**.

Anmerkung: Wichtig ist zu bemerken, dass ein Komplement nicht notwendigerweise ein orthogonales Komplement sein muss, bzw. man kann das Komplement

für jeden Vektorraum definieren (es muss aber nicht immer existieren), während man das orthogonale Komplement nur dann definieren kann, wenn man einen Vektorraum mit einem Skalarprodukt hat.

Definition 10.10 (Projektion)

Seien V ein K -Vektorraum und

$$V = U \oplus W$$

eine Zerlegung von V in die direkte Summe von zwei Untervektorräumen U, W . Dann heißt die Abbildung

$$p : V \rightarrow U, \quad v = u + w \mapsto u,$$

die einem Vektor v die erste Komponente der (eindeutigen, da direkte Summe) Zerlegung $v = u + w$ mit $u \in U, w \in W$ zuordnet, die **Projektion** auf U bezüglich W .

Nun definieren wir alles noch einmal für den komplexen Fall in \mathbb{C} und werden einige Gemeinsamkeiten und Ähnlichkeiten entdecken.

Definition 10.11 (hermitesche Form)

Eine **hermitesche Form** auf einem \mathbb{C} -Vektorraum V ist eine Abbildung

$$f : V \times V \rightarrow \mathbb{C}$$

mit den folgenden Eigenschaften für alle $u, v, w \in V$ und $\lambda \in \mathbb{C}$:

Linearität im 2. Argument:

$$f(u, v + w) = f(u, v) + f(u, w), \quad f(u, \lambda \cdot v) = \lambda \cdot f(u, v).$$

Semilinearität im 1. Argument:

$$f(u + v, w) = f(u, w) + f(v, w), \quad f(\lambda \cdot u, v) = \bar{\lambda} f(u, v).$$

Hermitesche Symmetrie:

$$f(v, u) = \overline{f(u, v)}.$$

Positive Definitheit:

$$f(u, u) > 0 \text{ für } u \neq 0.$$

Definition 10.12 (selbstadjungiert)

Sei $A = (a_{ij}) \in \mathcal{M}_{m,n}(\mathbb{C})$ eine $(m \times n)$ -Matrix mit komplexen Einträgen. Die $(n \times m)$ -Matrix

$$A^* := \bar{A}^T = ((\bar{a})_{ji})_{ij}$$

heißt die **Adjungierte** von A . Im Fall $m = n$ heißt A **selbstadjungiert**, wenn $A^* = A$ gilt, also wenn

$$\bar{a}_{ij} = a_{ji} \quad \forall i, j.$$

Definition 10.13 (darstellende Matrix)

Sei V ein endlichdimensionaler Vektorraum der Dimension n . Seien weiter $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V und f eine hermitesche Form auf V . Die **darstellende Matrix** oder auch **Darstellungsmatrix** von f bzgl. \mathcal{B} ist die $(n \times n)$ -Matrix

$$\mathcal{M}_{\mathcal{B}}(f) := (f(v_i, v_j))_{ij}.$$

Definition 10.14 (unitärer Raum)

Ein **unitärer Raum** ist ein \mathbb{C} -Vektorraum V , zusammen mit einer positiv definiten hermiteschen Form $\langle \cdot, \cdot \rangle$.

Definition 10.15 (unitärer)

Eine quadratische Matrix $Q \in \mathcal{M}_{n,n}(\mathbb{C})$ heißt **unitär**, wenn

$$Q \cdot Q^* = Q^* \cdot Q = E_n.$$

Wir bezeichnen die Teilmenge von $\mathcal{M}_{n,n}(\mathbb{C})$ aller unitären Matrizen mit U_n .

Definition 10.16 (Adjungierte)

Sei $\phi : V \rightarrow V$ ein Endomorphismus von V . Eine **Adjungierte** von ϕ ist ein Endomorphismus $\phi^* : V \rightarrow V$ mit der Eigenschaft, dass

$$\langle \phi^*(v), w \rangle = \langle v, \phi(w) \rangle \quad \forall v, w \in V.$$

Anmerkung: Diese Definition gilt natürlich auch für Matrizen. Denn aus der Linearen Algebra 1 wissen wir ja, dass wir eine lineare Abbildung zwischen endlichdimensionalen Vektorräumen mit einer Matrix identifizieren können.

Definition 10.17 (unitär, selbstadjungiert, normal)

Ein Endomorphismus $\phi : V \rightarrow V$ mit einer Adjungierten ϕ^* heißt

- i) **unitär**, wenn $\phi^* \circ \phi = \phi \circ \phi^* = \text{Id}_V$,
- ii) **selbstadjungiert**, wenn $\phi^* = \phi$,
- iii) **normal**, wenn $\phi^* \circ \phi = \phi \circ \phi^*$.

10.2 Sätze und Beweise

Satz 10.1

Seien V ein endlichdimensionaler K -Vektorraum der Dimension n , $f : V \times V \rightarrow K$ eine Bilinearform und $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V . Seien $v, w \in V$ beliebige Vektoren und $x, y \in K^n$ die zugehörigen Koordinatenvektoren bezüglich der Basis \mathcal{B} , das heißt, es gilt $v = \sum_{i=1}^n x_i v_i$ und $w = \sum_{i=1}^n y_i v_i$. So gilt:

$$f(v, w) = x^T A y.$$

Hierbei bezeichnet $A = \mathcal{M}_{\mathcal{B}}(f)$ die darstellende Matrix von f bezüglich der Basis \mathcal{B} .

Anmerkung: Wir lassen im Folgenden den Matrizenmultiplikationspunkt weg. Schreiben also nicht $x^T \cdot A \cdot y$, sondern einfach nur $x^T A y$. Es ist nicht ausgeschlossen, dass wir dies an einigen Stellen aber schlichtweg vergessen haben, oder, weil es deutlicher ist, trotzdem hingeschrieben haben. Dies sei uns verziehen, und wir geben dann zu bedenken, dass der Malpunkt die Matrixmultiplikation, die Skalarproduktmultiplikation oder die herkömmliche Multiplikation bezeichnen kann.

Beweis: Die Behauptung folgt sofort, wenn man die Bilinearität von f benutzt, denn es gilt so

$$\begin{aligned} f(v, w) &= f\left(\sum_i x_i v_i, \sum_j y_j v_j\right) \\ &= \sum_{i,j} x_i y_j f(v_i, v_j) \\ &= x^T A y. \end{aligned}$$

q.e.d.

Satz 10.2

Ist $f : V \times V \rightarrow K$ eine Bilinearform auf dem Standardvektorraum $V = K^n$ und ist $A = (a_{ij}) \in \mathcal{M}_{n,n}(K)$ die darstellende Matrix von f bezüglich der Standardbasis, also

$$a_{ij} := f(e_i, e_j), \quad i, j = 1, \dots, n,$$

so gilt:

$$f = \langle \cdot, \cdot \rangle_A.$$

Anmerkung: Zur Definition von $\langle \cdot, \cdot \rangle_A$ siehe Beispiel 83. Dort definieren wir das wichtige Skalarprodukt $\langle x, y \rangle_A := x^T A y$.

Satz 10.3 (Basiswechselsatz für Bilinearformen)

Seien V ein endlichdimensionaler K -Vektorraum der Dimension n und $f : V \times V \rightarrow K$ eine Bilinearform. Seien \mathcal{A}, \mathcal{B} Basen von V . Die Beziehung zwischen den darstellenden Matrizen von f bzgl. \mathcal{A} und \mathcal{B} wird durch die Formel

$$\mathcal{M}_{\mathcal{B}}(f) = (T_{\mathcal{A}}^{\mathcal{B}})^T \cdot \mathcal{M}_{\mathcal{A}}(f) \cdot T_{\mathcal{A}}^{\mathcal{B}}$$

beschrieben. Hierbei ist $T_{\mathcal{A}}^{\mathcal{B}} \in GL_n(K)$ die Transformationsmatrix des Basiswechsels von \mathcal{B} nach \mathcal{A} .

Beweis: Wir nehmen zwei Vektoren $v, w \in V$ und bilden zuerst die zugehörigen Koordinatenvektoren $x, y \in K^n$ bezüglich der Basis $\mathcal{A} = (v_1, \dots, v_n)$ ab und erhalten

$$v = \sum_i x_i v_i, \quad w = \sum_i y_i v_i.$$

Dann sind die Koordinatenvektoren x', y' bezüglich der Basis $\mathcal{B} = (w_1, \dots, w_n)$ gegeben durch

$$v = \sum_i x'_i w_i, \quad w = \sum_i y'_i w_i.$$

Ist $Q = T_{\mathcal{A}}^{\mathcal{B}}$ die Transformationsmatrix des Basiswechsels von \mathcal{B} nach \mathcal{A} , so gilt:

$$x = Qx', \quad y = Qy'. \quad (10.1)$$

Ist nun $A = \mathcal{M}_{\mathcal{A}}(f)$ die Darstellungsmatrix von f bezüglich \mathcal{A} , so gilt:

$$f(v, w) = x^T A y. \quad (10.2)$$

Durch Einsetzen von (10.1) in (10.2), erhalten wir

$$f(v, w) = (Qx')^T A (Qy') = (x')^T (Q^T A Q) y'.$$

Es folgt nun mit Satz 10.2, dass

$$B := Q^T A Q = \mathcal{M}_{\mathcal{B}}(f)$$

die darstellende Matrix von f bezüglich der Basis \mathcal{B} ist.

q.e.d.

Satz 10.4

Seien V ein endlichdimensionaler Vektorraum der Dimension n , f eine Bilinearform auf V und $\mathcal{A} = (v_1, \dots, v_n)$ eine Basis von V . Sei $A := \mathcal{M}_{\mathcal{A}}(f)$ die darstellende Matrix. Dann gelten die folgenden Aussagen:

- i) Es gibt eine orthogonale Basis von V bzgl. f genau dann, wenn eine invertierbare Matrix $Q \in GL_n(K)$ existiert, sodass

$$Q^T A Q$$

eine Diagonalmatrix ist.

ii) Es gibt eine Orthonormalbasis von V bzgl. f genau dann, wenn es eine invertierbare Matrix $Q \in GL_n(K)$ gibt, sodass

$$Q^T A Q = E_n$$

die Einheitsmatrix ist. Äquivalent formuliert: Es gibt ein invertierbares $P \in GL_n(K)$ mit

$$A = P^T P.$$

Beweis:

i) Sei $\mathcal{B} = (w_1, \dots, w_n)$ eine Basis von V . Nach Definition 10.4 ist diese genau dann orthogonal bezüglich f , wenn für $i \neq j$ gilt:

$$b_{ij} := f(w_i, w_j) = 0.$$

Dies bedeutet aber, dass die Darstellungsmatrix

$$B = \mathcal{M}_{\mathcal{B}}(f) = (b_{ij}), \quad b_{ij} = f(w_i, w_j)$$

eine Diagonalmatrix ist. Es gilt nun

$$B = Q^T A Q \quad \text{mit } Q := T_{\mathcal{A}}^{\mathcal{B}} \in GL_n(K).$$

Damit ist die erste Aussage bewiesen.

ii) \mathcal{B} ist genau dann eine Orthonormalbasis, wenn

$$B = \mathcal{M}_{\mathcal{B}}(f) = E_n.$$

Dies ist aber genau dann der Fall, wenn

$$A = P^T \mathcal{M}_{\mathcal{B}} P = P^T E_n P = P^T P \quad \text{mit } P = T_{\mathcal{A}}^{\mathcal{B}} = Q^{-1}.$$

q.e.d.

Satz 10.5 (Orthogonalisierungsverfahren von Gram-Schmidt)

Sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum. Dann gibt es eine Orthonormalbasis, das heißt eine Basis $\mathcal{B} = (w_1, \dots, w_n)$ von V mit

$$\langle w_i, w_j \rangle = \delta_{ij}.$$

Anmerkung: Den Beweis wollen wir hier auf jeden Fall geben. Dieser wird konstruktiv sein, wie nicht anders zu erwarten war. Aus dem Beweis folgt das sogenannte *Orthogonalisierungsverfahren von Gram-Schmidt* (in unserer Form ist es sogar ein Orthonormalisierungsverfahren). Wer das in dieser abstrakten Form nicht sofort versteht, den verweisen wir auf die Erklärungen und Beispiele zu diesem Satz :-). Los geht es:

Beweis: Sei $\mathcal{A} = (v_1, \dots, v_n)$ eine beliebige Basis von V . Zum Beweis werden wir induktiv Vektoren w_1, \dots, w_n definieren, sodass für $k = 1, \dots, n$ die Vektoren w_1, \dots, w_k eine Orthonormalbasis des Untervektorraums

$$V_k := \langle v_1, \dots, v_k \rangle \subset V$$

bilden. Für $k = n$ erhalten wir dann die Behauptung. Für den Induktionsanfang setzen wir

$$w_1 := \frac{1}{\|v_1\|} v_1.$$

Offenbar ist (w_1) eine Orthonormalbasis von V_1 , denn so wurde dies ja gerade konstruiert. Wir nehmen nun an, dass $k > 1$ und dass wir bereits eine Orthonormalbasis (w_1, \dots, w_{k-1}) von V_{k-1} gefunden haben. Wir müssen also nur noch einen passenden Vektor w finden, der (w_1, \dots, w_{k-1}) zu einer Orthonormalbasis von V_k ergänzt. Unser erster Ansatz soll

$$w = v_k - a_1 \cdot w_1 - \dots - a_{k-1} \cdot w_{k-1} \quad (10.3)$$

lauten. Hierbei sind die a_i noch zu bestimmende Skalare. Durch diesen Ansatz wird auf jeden Fall gewährleistet, dass (w_1, \dots, w_k) eine Basis von V_k ist. Durch passende Wahl der a_i möchten wir erreichen, dass zusätzlich w zu w_i orthogonal ist und zwar für $i = 1, 2, \dots, k-1$. Nach der Induktionsvoraussetzung gilt $\langle w_i, w_j \rangle = 0$ für $i \neq j$ und $\langle w_i, w_i \rangle = 1$. Nach Einsetzen von (10.3), erhalten wir

$$\langle w, w_i \rangle = 0 \quad \text{für } i = 1, \dots, k-1.$$

Der Vektor w ist allerdings noch nicht normiert, hat also noch nicht die Länge 1. Deshalb setzen wir

$$w_k := \frac{1}{\|w\|} w.$$

Nun ist (w_1, \dots, w_k) eine Orthonormalbasis von V_k . Dies war zu zeigen. q.e.d.

Satz 10.6 (Korollar zu Satz 10.4)

Für eine Matrix $A \in \mathcal{M}_{n,n}(\mathbb{R})$ sind die folgenden Aussagen äquivalent.

- i) Der Standardvektorraum $V = \mathbb{R}^n$ besitzt bezüglich des durch A definierten Skalarprodukts $\langle \cdot, \cdot \rangle_A$ eine Orthonormalbasis.
- ii) Es gibt eine invertierbare Matrix $P \in GL_n(\mathbb{R})$ mit

$$A = P^T P.$$

- iii) Die Matrix A ist symmetrisch und positiv definit.

Satz 10.7 (Hauptminorenkriterium)

Eine reelle symmetrische Matrix $A = (a_{ij}) \in \mathcal{M}_{n,n}(\mathbb{R})$ ist genau dann positiv definit, wenn alle Hauptminoren von A positiv sind, das heißt,

$$\det A_i > 0$$

für $i = 1, \dots, n$, wobei die Hauptminoren gegeben sind durch

$$A_1 = (a_{11}), \quad A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad A_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \dots, \quad A_n = A.$$

Satz 10.8

Seien V ein K -Vektorraum und $U, W \subset V$ Untervektorräume. Dann sind die folgenden Bedingungen äquivalent, wobei wir für iii) annehmen, dass V endlichdimensional ist.

i) $V = U \oplus W$.

ii) Jeder Vektor $v \in V$ lässt sich auf eindeutige Weise als eine Summe

$$v = u + w$$

schreiben, wobei $u \in U$ und $w \in W$.

iii) Sei $\mathcal{A} = (u_1, \dots, u_r)$ eine Basis von U und $\mathcal{B} = (w_1, \dots, w_s)$ eine Basis von W . Dann ist

$$\mathcal{A} \cup \mathcal{B} := (u_1, \dots, u_r, w_1, \dots, w_s)$$

eine Basis von V .

Satz 10.9 (Dimensionsformeln)

Seien V und W zwei endlichdimensionale Vektorräume, so gelten die beiden Dimensionsformeln

$$\dim(V + W) + \dim(V \cap W) = \dim(V) + \dim(W),$$

$$\dim(V \oplus W) = \dim(V) + \dim(W).$$

Satz 10.10

Seien V ein endlichdimensionaler K -Vektorraum und $U \subset V$ ein Untervektorraum. Dann gelten die folgenden Eigenschaften

i) Es gibt ein Komplement $W \subset V$ zu U .

ii) Für jedes Komplement W zu U gilt:

$$\dim_K V = \dim_K U + \dim_K W.$$

iii) Es gilt:

$$(U^\perp)^\perp = U.$$

Satz 10.11

Seien $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum und $U \subset V$ ein Untervektorraum. Wir bezeichnen mit $p : V \rightarrow U$ die orthogonale Projektion auf U . Für alle Vektoren $v \in V$ gilt dann:

$$\|v - p(v)\| = \min_{u \in U} \|v - u\|.$$

Wir formulieren analoge Aussagen noch einmal in \mathbb{C} .

Satz 10.12

Sei $\mathcal{B} = (v_1, \dots, v_n)$ eine Basis von V , f sei eine hermitesche Form und $A := \mathcal{M}_{\mathcal{B}}(f)$ sei die darstellende Matrix. Dann gilt:

i) A ist selbstadjungiert.

ii) Sind $v, w \in V$ zwei Vektoren mit zugehörigen Koordinatenvektoren $x, y \in \mathbb{C}^n$, das heißt

$$v = \sum_i x_i v_i, \quad w = \sum_i y_i v_i,$$

so gilt

$$f(v, w) = x^* A y = \langle x, y \rangle_A$$

für alle $v, w \in V$.

Satz 10.13 (Basiswechsel)

Seien V ein endlichdimensionaler \mathbb{C} -Vektorraum, \mathcal{A} und \mathcal{B} zwei Basen von V und f eine hermitesche Form. Dann gilt:

$$\mathcal{M}_{\mathcal{B}}(f) = (T_{\mathcal{A}}^{\mathcal{B}})^* \mathcal{M}_{\mathcal{A}}(f) T_{\mathcal{A}}^{\mathcal{B}}.$$

Satz 10.14 (Gram-Schmidt im Komplexen)

Seien V ein endlichdimensionaler \mathbb{C} -Vektorraum und $f = \langle \cdot, \cdot \rangle$ eine hermitesche Form auf V . Dann besitzt V eine Orthonormalbasis bezüglich f genau dann, wenn f positiv definit ist.

Satz 10.15 (Korollar zum Gram-Schmidt-Verfahren)

Für eine Matrix $A \in \mathcal{M}_{n,n}(\mathbb{C})$ sind die folgenden Aussagen äquivalent:

- i) A ist selbstadjungiert und positiv definit, das heißt (also die positive Definitheit)

$$x^* Ax > 0$$

für alle $x \in \mathbb{C}^n$, $x \neq 0$.

- ii) Es gibt eine invertierbare Matrix $P \in GL_n(\mathbb{C})$, sodass

$$A = P^* P.$$

Satz 10.16

Sei $\phi : V \rightarrow V$ ein Endomorphismus. Dann gelten die folgenden Eigenschaften.

- i) Die Adjungierte ϕ^* von ϕ ist eindeutig, wenn sie existiert.
- ii) Ist V endlichdimensional, so gibt es eine Adjungierte ϕ^* . Ist $A := \mathcal{M}_{\mathcal{B}}(\phi)$ die darstellende Matrix von ϕ bzgl. einer Orthonormalbasis \mathcal{B} , so gilt:

$$\mathcal{M}_{\mathcal{B}}(\phi^*) = A^*.$$

Anmerkung: Die Frage ist, wie die Adjungierte eines Endomorphismus definiert ist. Dies ist aber recht leicht zu beantworten, denn in der linearen Algebra 1 haben wir gesehen, dass man lineare Abbildungen zwischen endlichdimensionalen Vektorräumen als Matrizen realisieren kann.

Satz 10.17 (Spektralsatz für normale Endomorphismen)

Seien $(V, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler unitärer Raum und $\phi : V \rightarrow V$ ein Endomorphismus. Dann sind die folgenden Aussagen äquivalent:

- i) ϕ ist normal.
- ii) V besitzt eine Orthonormalbasis, die aus Eigenwerten von ϕ besteht.

Satz 10.18

Für eine Matrix $A \in \mathcal{M}_{n,n}(\mathbb{C})$ sind die folgenden Bedingungen äquivalent:

- i) A ist normal, das heißt $A^* A = A A^*$.
- ii) Es gibt eine unitäre Matrix $S \in U_n$, sodass die Matrix

$$S^{-1} A S$$

Diagonalgestalt hat.

Satz 10.19

Eine selbstadjungierte (im reellen Fall symmetrische) Matrix $A \in \mathcal{M}_{n,n}(\mathbb{R})$ besitzt nur reelle Eigenwerte.

Beweis: Aus der Selbstadjungiertheit folgt für alle $x \in \mathbb{C}^n$

$$\overline{x^*Ax} = (x^*Ax)^* = x^*A^*x = x^*Ax.$$

Dies bedeutet aber, dass $\langle x, x \rangle_A = x^*Ax$ eine reelle Zahl ist. Nun sei $x \in \mathbb{C}^n$ ein Eigenvektor von A zum Eigenwert $\lambda \in \mathbb{C}$. Dann gilt:

$$x^*Ax = x^*(\lambda x) = \lambda(x^*x) \in \mathbb{R}.$$

Da x^*Ax und x^*x reelle Zahlen sind, ist λ ebenfalls reell. Dies war zu zeigen.
q.e.d.

Satz 10.20

Ist $A := \mathcal{M}_{\mathcal{B}}(\phi)$ die darstellende Matrix eines Endomorphismus ϕ bezüglich einer Orthonormalbasis \mathcal{B} , so gilt:

- i) ϕ ist unitär genau dann, wenn A unitär ist.
- ii) ϕ ist selbstadjungiert genau dann, wenn A selbstadjungiert ist.
- iii) ϕ ist normal genau dann, wenn A normal ist.

10.3 Erklärungen zu den Definitionen

Zur Definition 10.1 der Bilinearform: Diese Definition erinnert stark an das Standardskalarprodukt (siehe Definition 9.1 und seine Erklärungen). Dies ist kein Zufall. Denn ein Skalarprodukt ist nichts anderes als eine symmetrische, positiv definite Bilinearform, also nur ein Spezialfall von Definition 10.1.

Beispiel 83

Sei $V = K^n$ der Standardvektorraum der Dimension n über dem Körper K und $A \in \mathcal{M}_{n,n}(K)$ eine $(n \times n)$ -Matrix. Dann behaupten wir, dass durch

$$\langle x, y \rangle_A := x^T Ay \tag{10.4}$$

eine Bilinearform auf V erklärt und definiert wird. Dies müssen wir nun zeigen, indem wir die Eigenschaften aus obiger Definition getrennt nachweisen. Hierbei können wir aber Wissen aus der Linearen Algebra 1 benutzen. Die Bilinearität, also die obigen Eigenschaften aus der Definition, folgt sofort aus dem Distribu-

tivgesetz für die Matrixaddition und Matrixmultiplikation, denn etwas anderes machen wir da ja eigentlich gar nicht. Weisen wir dies einmal nach:

$$\begin{aligned}
 \langle x + y, z \rangle_A &= (x + y)^T A z = (x^T + y^T) A z = (x^T \cdot A + y^T \cdot A) \cdot z \\
 &= (x^T A z) + (y^T A z) = \langle x, z \rangle_A + \langle y, z \rangle_A \\
 \langle x, y + z \rangle_A &= x^T A (y + z) = x^T (A y + A z) = x^T A y + x^T A z \\
 &= (x^T A y) + (x^T A z) = \langle x, y \rangle_A + \langle x, z \rangle_A \\
 \langle \lambda \cdot x, y \rangle_A &= (\lambda x)^T A y = \lambda^T x^T A y \\
 &= \lambda x^T A y = \lambda \langle x, y \rangle_A.
 \end{aligned}$$

Der Rest zeigt sich analog und genauso einfach :-). Um weitere nette Bekanntschaften mit dieser Bilinearform zu machen, schreiben wir uns dies einmal in Koordinaten hin. Die Matrix A habe die Einträge $A = (a_{ij})$. Demnach gilt:

$$\begin{aligned}
 \langle x, y \rangle_A &= \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \\
 &= \sum_{i,j=1}^n a_{ij} x_i y_j. \quad (\text{Doppelsumme})
 \end{aligned}$$

Ist $A = E_n$ die Einheitsmatrix, so erhält man natürlich das Standardskalarprodukt

$$\langle x, y \rangle = x^T y = x_1 y_1 + \dots + x_n y_n.$$

Noch ein paar Bemerkungen: Ist (e_1, \dots, e_n) die Standardbasis des \mathbb{R}^n , so gilt offenbar

$$\langle e_i, e_j \rangle_A = e_i^T \cdot A \cdot e_j = a_{ij}.$$

Außerdem folgt hieraus sofort, dass die Matrix A durch die Bilinearform $\langle \cdot, \cdot \rangle_A$ eindeutig bestimmt ist. ■

Beispiel 84

Wir geben noch ein Beispiel zur obigen Bilinearform aus Beispiel 83 an. Seien

$$V = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 0\} \quad \text{und} \quad v_1 := \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad v_2 := \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

zwei Basisvektoren von V . Nach Definition der Basis ist es möglich, einen beliebigen Vektor als Linearkombination der Basisvektoren darzustellen. Es gilt demnach

$$V = \{v = x'_1 v_1 + x'_2 v_2 : x'_1, x'_2 \in \mathbb{R}\}.$$

Es ergibt sich nun

$$\langle v_1, v_1 \rangle = 2, \quad \langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle = 1, \quad \langle v_2, v_2 \rangle = 2.$$

Damit also

$$A = \begin{pmatrix} \langle v_1, v_1 \rangle & \langle v_1, v_2 \rangle \\ \langle v_2, v_1 \rangle & \langle v_2, v_2 \rangle \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.$$

Es gilt jetzt:

$$v = x'_1 v_1 + x'_2 v_2, \quad x' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} \in \mathbb{R}^2,$$

$$w = y'_1 v_1 + y'_2 v_2, \quad y' = \begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix} \in \mathbb{R}^2.$$

Mit dieser Definition der Vektoren erhalten wir nach Definition 9.1 des Standardskalarprodukts und den Eigenschaften der Bilinearform 10.1

$$\begin{aligned} \langle v, w \rangle &= \langle x'_1 v_1 + x'_2 v_2, y'_1 v_1 + y'_2 v_2 \rangle \\ &= x'_1 y'_1 \cdot \underbrace{\langle v_1, v_1 \rangle}_{=2} + x'_1 y'_2 \cdot \underbrace{\langle v_1, v_2 \rangle}_{=1} + x'_2 y'_1 \cdot \underbrace{\langle v_2, v_1 \rangle}_{=1} + x'_2 y'_2 \cdot \underbrace{\langle v_2, v_2 \rangle}_{=2} \\ &= \begin{pmatrix} x'_1 & x'_2 \end{pmatrix} \cdot A \cdot \begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix} \\ &= (x')^T A y'. \end{aligned}$$

Das Beispiel wird durch Definition 10.3 motiviert, wie die Erklärungen zu dieser Definition zeigen werden. ■

Zur Definition 10.2 der symmetrischen Bilinearform: Die Bilinearform $\langle \cdot, \cdot \rangle_A$ aus Beispiel 83 ist genau dann symmetrisch, wenn die Matrix A symmetrisch ist, das heißt wenn $A^T = A$. Dies folgt sofort aus

$$\langle x, y \rangle_A = (x^T A y)^T = y^T A x = \langle y, x \rangle_A.$$

Näheres in Kapitel 13.

Zur Definition 10.3 der darstellenden Matrix: Das Konzept ist schon aus der Linearen Algebra 1 bekannt. Auch für Bilinearformen gibt es eine Darstellungsmatrix und auch einen Basiswechselsatz (siehe Satz 10.3). Um das Prinzip dahinter zu verstehen, fassen wir nochmals zusammen, was wir über Darstellungsmatrizen bei linearen Abbildungen zwischen Vektorräumen wissen und was wir dort gelernt haben. Dort konnte man zwischen zwei Vektorräumen, die durch eine

lineare Abbildung miteinander „verbunden“ waren (was wir natürlich meinen: die lineare Abbildung bildete von einem Vektorraum in den anderen ab), eine darstellende Matrix ermitteln, indem wir uns entsprechend eine Basis des einen und eine des anderen Vektorraums gewählt und danach die Basisvektoren der einen Basis unter der linearen Abbildung abgebildet haben, und diese Vektoren haben wir dann als Linearkombinationen der anderen Basisvektoren geschrieben. So erhielt man die darstellende Matrix. Auch eine Transformationsmatrix ist uns nicht unbekannt. Damit konnten wir quasi zwischen zwei Basen „hin und her springen“. Ein Beispiel für Bilinearformen schauen wir uns in den Erklärungen zu Satz 10.1 an.

Zur Definition 10.4 der Orthonormalbasis: Wir wissen aus der Linearen Algebra 1, dass man zu jedem Vektorraum eine Basis finden kann. Eine Orthonormalbasis ist nun eine „besondere“ Basis, bei der alle Basisvektoren paarweise senkrecht aufeinander stehen. Dies sagt gerade die Eigenschaft $\langle v_i, v_j \rangle = 0$. Wenn nur das gilt, so heißt die Basis orthogonal. Wenn zusätzlich noch gilt, dass $\langle v_i, v_i \rangle = 1$ für alle Vektoren v_i , so heißt die Basis eine Orthonormalbasis. Diese Eigenschaft bedeutet einfach nur, dass alle Vektoren normiert sind, also die Länge 1 besitzen, mehr nicht. Wir werden sehen, dass unter gewissen Voraussetzungen eine Orthonormalbasis (ab und zu kürzen wir dieses aus Faulheit mit ONB ab) in einem Vektorraum existiert. Die interessante Frage ist aber, wie man aus einer „normalen“ Basis eine ONB konstruiert. Dies geschieht mit dem sogenannten Orthogonalisierungsverfahren von Gram-Schmidt, siehe dazu Satz 10.5 und seine Erklärungen.

Zur Definition 10.5 der positiven Definitheit: Dies kennen wir schon von Matrizen. Diese Definition ist einfach nur die Verallgemeinerung auf Bilinearformen. Aber da jede Bilinearform eine Darstellungsmatrix besitzt, können wir das Konzept wieder auf die Definitheit von Matrizen zurückführen. Und dies haben wir in Kapitel 4 bei den Extremwertaufgaben geübt, denn dort mussten wir ja die sogenannte Hesse-Matrix auf Definitheit überprüfen, um entscheiden zu können, ob ein Maximum oder Minimum vorlag. Um das Problem einzuüben, eine kleine Aufgabe zum Lösen, die etwa Klausurstil hat :-).

Beispiel 85

Sei $A = (a_{ij}) \in \mathcal{M}_{n,n}(\mathbb{R})$ eine reelle symmetrische Matrix.

- i) Zeige: Ist A positiv definit, so gilt $a_{ii} > 0$ für alle $i = 1, \dots, n$.
- ii) Zeige, dass die Umkehrung in a) nicht gilt!
- iii) Sei

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix} \in \mathcal{M}_{2,2}(\mathbb{R}).$$

Beweise das Hauptminorenkriterium (vgl. auch den Satz 10.7):

$$A \text{ positiv definit} \Leftrightarrow a > 0 \text{ und } \det(A) = ac - b^2 > 0.$$

Die Lösung geht so:

- i) Erst einmal stellen wir fest, dass die Behauptung für beliebige quadratische Matrizen falsch ist. Denn dazu betrachte die Matrix

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

Diese Matrix ist nach dem Hauptminorenkriterium (Teil iii) dieser schönen Aufgabe positiv definit, denn $\det 1 = 1 > 0$ und

$$\det A = \det \begin{pmatrix} 1 & -2 \\ 2 & -3 \end{pmatrix} = -3 + 4 = 1 > 0.$$

(Wir benutzen diesen Teil schon einmal, aber für den Beweis von c) benötigen wir Teil i) und ii) nicht, daher ergibt sich kein Ringschluss.) Aber es gilt nicht $a_{ii} > 0$ für alle $i = 1, \dots, n$. Die Symmetrie, die an die Matrix vorausgesetzt wird, ist also irgendwie wichtig und muss im Beweis mit eingehen. Beginnen wir mit dem Beweis: Wenn A positiv definit ist, dann heißt das nach Definition $x^T A x > 0$ für alle $x \in \mathbb{R}^n, x \neq 0$. Es gilt demnach:

$$x^T A x = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix}^T \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i,j=1}^n a_{ij} x_i x_j > 0.$$

Da die Matrix symmetrisch ist, gilt $A^T = A$ und damit also

$$\sum_{i,j=1}^n a_{ij} x_i x_j = \sum_{i,j=1}^n a_{ji} x_j x_i > 0.$$

Folglich muss $a_{ij} > 0$ für alle $i = 1, \dots, n$ gelten. Zum Beispiel erfüllt der i -te Einheitsvektor diese Eigenschaft. Es gilt nämlich:

$$\begin{aligned}
 0 < e_i^T A e_i &= \begin{pmatrix} 0 & \cdots & 1 & \cdots & 0 \end{pmatrix}^T \cdot \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \\
 &= \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = a_{ii} \Rightarrow a_{ii} > 0.
 \end{aligned}$$

Die 1 steht dabei jeweils an der i -ten Stelle.

- ii) Nun zeigen wir durch ein konkretes Beispiel, dass die Umkehrung in i) falsch ist, das heißt, wenn alle Diagonaleinträge größer als Null sind, folgt eben nicht, dass die Matrix positiv definit ist. Dazu betrachten wir die Matrix

$$B := \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}.$$

Diese Matrix besitzt nur positive Diagonaleinträge. Nach dem Hauptminorenkriterium (benutzen wir hier, ohne es bewiesen zu haben, weil beim Beweis von Teil iii) die Teile i) und ii) nicht eingehen!) ist diese aber nicht positiv definit, denn es gilt:

$$\det B = \det \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} = 1 - 4 = -3 < 0.$$

- iii) Der Beweis der Behauptung besteht aus zwei Richtungen. Fangen wir mit der Hin-Richtung (Wortspiel :-P) an: Sei A positiv definit. Wir müssen zeigen, dass dann $a > 0$ und $\det A = ac - b^2 > 0$. Der erste Teil, also dass $a > 0$ folgt sofort aus Aufgabenteil i). Bleibt noch zu zeigen, dass $\det A = ac - b^2 > 0$. Dies wird klar, wenn wir uns die Definition der positiven Definitheit einer Matrix A vor Augen führen. Eine Matrix A heißt nach unserer Definition 10.5 positiv definit, wenn für alle $x \in \mathbb{R}^n$, $x \neq 0$

$$x^T A x > 0$$

gilt. Dies können wir aufgrund der einfachen Struktur der Matrix A leicht ausrechnen. Dazu sei $x := \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. Es gilt nun:

$$\begin{aligned} \begin{pmatrix} x_1 & x_2 \end{pmatrix} \cdot \begin{pmatrix} a & b \\ b & c \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \begin{pmatrix} x_1 & x_2 \end{pmatrix} \cdot \begin{pmatrix} ax_1 + bx_2 \\ bx_1 + cx_2 \end{pmatrix} \\ &= (ax_1 + bx_2)x_1 + (bx_1 + cx_2)x_2 = ax_1^2 + bx_2x_1 + bx_1x_2 + cx_2^2 \\ &= ax_1^2 + 2bx_2x_1 + cx_2^2 > 0. \end{aligned}$$

Angenommen, es würde $\det(A) = ac - b^2 < 0$ gelten, dann wäre $ac < b^2$. Dann wäre aber

$$\underbrace{ax_1^2}_{>0} + 2bx_2x_1 + \underbrace{cx_2^2}_{>0}$$

nicht unbedingt größer als Null, was ein Widerspruch zu unserer Voraussetzung ist. Folglich muss $\det(A) = ac - b^2 > 0$ gelten. Dies zeigt den ersten Teil.

Kommen wir zur Rückrichtung: Seien $a > 0$ und $\det(A) = ac - b^2 > 0$. Wir müssen nun zeigen, dass A positiv definit ist. Hierfür führen wir eine einfache quadratische Ergänzung durch. Es gilt ja, wie oben berechnet, $x^T Ax = ax_1^2 + 2bx_2x_1 + cx_2^2$. Nun schließen wir so:

$$\begin{aligned} x^T Ax &= ax_1^2 + 2bx_2x_1 + cx_2^2 \\ &= a \left(x_1^2 + 2\frac{b}{a}x_1x_2 + \frac{b^2}{a^2}x_2^2 \right) + cx_2^2 - \frac{b^2}{a}x_2^2 \\ &= a \left(x_1 + \frac{b}{a}x_2 \right)^2 + \left(c - \frac{b^2}{a} \right) x_2^2. \end{aligned}$$

Da nun $a > 0$ und $\det(A) = ac - b^2 > 0$, ergibt sich sofort, dass

$$a \left(x_1 + \frac{b}{a}x_2 \right)^2 + \left(c - \frac{b^2}{a} \right) x_2^2 > 0$$

und folglich $x^T Ax > 0$.

Zum Schluss bemerken wir noch: Die Rechnung der Rück-Richtung bei iii) impliziert eigentlich einen Beweis für die Hin-Richtung. Wir hätten also eigentlich nur diese Rechnung angeben müssen. Seht ihr das? Aufgabe: Verkürzt den Beweis unter iii). ■

Zur Definition 10.6 zum euklidischen Vektorraum und der Norm: Ein Vektorraum (bekannt aus der Linearen Algebra 1) ist also ein euklidischer Vektorraum, wenn wir hierauf eine Bilinearform definieren können. Die Norm eines Vektors gibt einfach seine Länge an. Auch nichts Neues, oder?

Zur Definition 10.7 der direkten Summe: Hier wollen wir bemerken, dass man die Definition genau lesen muss. Die Summe $U_1 + U_2$ wird für jedes Paar U_1, U_2 von Untervektorräumen als Untervektorraum von V definiert. Das heißt, für beliebig vorgegebene U_1, U_2 macht die Zuordnung $U := U_1 + U_2$ Sinn. Bei der direkten Summe dagegen wird von der Summe $U_1 + U_2$ eine zusätzliche Eigenschaft gefordert, die nicht notwendigerweise erfüllt sein muss, nämlich $U_1 \cap U_2 = \{0\}$. Die Zuordnung $U := U_1 \oplus U_2$ macht daher nicht im Allgemeinen Sinn.

Beispiel 86

Bei diesem Beispiel verwenden wir ab und an die wichtige Dimensionsformel. Wem dies nichts mehr sagt, der lese bitte in Satz 10.9 nach.

- Es sei V der Vektorraum aller $(n \times n)$ -Matrizen über den reellen Zahlen \mathbb{R} . Weiterhin definieren wir zwei Unterräume

$$U := \{A \in V : A^T = A\} \quad \text{und} \quad W := \{A \in V : A^T = -A\}.$$

Wir wollen jetzt untersuchen, ob $V = U \oplus W$ gilt. Zunächst müssten wir natürlich erst einmal zeigen, dass U und W wirklich Unterräume von V sind. Dazu sei nicht viel gesagt, denn das solltet ihr aus der Linearen Algebra 1 kennen. Nur noch einmal so viel: Seien $A, B \in U$, also $A^T = A$ und $B^T = B$. Daraus ergibt sich

$$(A + B)^T = A^T + B^T = A + B.$$

Den Rest prüfe man bitte selbst nach :-). Wir wollen jetzt überprüfen, ob V die direkte Summe von U und W ist. Dazu ist es immer nützlich, wenn wir die Dimensionen der einzelnen Unterräume kennen. Man kann sich überlegen (und auch dies soll eine Übungsaufgabe sein), dass

$$\dim(U) = \frac{1}{2}(n^2 + n) \quad \text{und} \quad \dim(W) = \frac{1}{2}(n^2 - n).$$

Nun an die eigentliche Arbeit: Der Vektorraum der $(n \times n)$ -Matrizen über \mathbb{R} hat die Dimension n^2 , klar, oder? Die Nullmatrix liegt sowohl in U als auch in W . Wegen (siehe Satz 10.9)

$$\dim(U) + \dim(W) = \frac{1}{2}(n^2 + n) + \frac{1}{2}(n^2 - n) = n^2,$$

und

$$U \cap W = \{0\}$$

(denn die Nullmatrix ist die einzige Matrix, die sowohl in U als auch in W enthalten ist) gilt also tatsächlich $V = U \oplus W$.

- Wir betrachten jetzt den Vektorraum aller Polynome vom Grad kleiner oder gleich 3 über \mathbb{R} und die Unterräume

$$U_1 = \langle x^3 + 2x^2, 2x^3 + 3x^2 + 1 \rangle \text{ und} \\ U_2 = \langle x^3 + 2x^2 + 1, x^3 - 1, x^2 + 1 \rangle.$$

Wir wollen jetzt jeweils die Dimensionen von $U_1, U_2, U_1 + U_2$ und $U_1 \cap U_2$ angeben und einen Unterraum W von V so konstruieren, dass $V = U_1 \oplus W$.

Da $x^3 + 2x^2$ und $2x^3 + 3x^2 + 1$ linear unabhängig sind, folgt $\dim(U_1) = 2$, und $x^3 + 2x^2, 2x^3 + 3x^2 + 1$ ist eine Basis von U_1 . Außerdem sind $x^3 - 1$ und $x^2 + 1$ linear unabhängig, und es gilt:

$$(x^3 - 1) + 2(x^2 + 1) = x^3 + 2x^2 + 1.$$

Demnach ist also $\dim(U_2) = 2$ und $x^3 - 1, x^2 + 1$ bildet eine Basis von U_2 . Für $u \in U_1 \cap U_2$ erhalten wir

$$u = \alpha(x^3 + 2x^2) + \beta(2x^3 + 3x^2 + 1) = \gamma(x^3 - 1) + \delta(x^2 + 1).$$

Hieraus ergibt sich ein lineares Gleichungssystem, das auf einen eindimensionalen Lösungsraum

$$U = \langle (0, 1, 2, 3) \rangle$$

führt. Daraus folgt nun

$$U_1 \cap U_2 = \langle 2x^3 + 3x^2 + 1 \rangle.$$

Die Dimensionsformel (Satz 10.9) liefert nun

$$\dim(U_1 + U_2) = 2 + 2 - 1 = 3.$$

Da $x^3 + 2x^2, 2x^3 + 3x^2 + 1, x^2 + 1 \in U_1 + U_2$ linear unabhängig sind, bilden sie eine Basis der Summe $U_1 + U_2$. Ergänzt man die obige Basis von U_1 durch 1 und x zu einer Basis von V , so gilt für $W := \langle 1, x \rangle$, wie gewünscht, $V = U_1 \oplus W$. ■

Zur Definition 10.8 des Komplements: Das Komplement ist keinesfalls eindeutig. Dazu betrachte man einfach eine Gerade im \mathbb{R}^2 durch den Ursprung. Dazu kann ein Komplement gebildet werden. Nämlich wieder eine Gerade durch den Ursprung. Davon gibt es aber unendlich viele.

Zur Definition 10.10 zur Projektion: Ist $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Vektorraum. In diesem Fall gibt es für jeden Untervektorraum $U \subset V$ eine ausgezeichnete Projektion $p_U : V \rightarrow V$, die *orthogonale Projektion*. Ist $U \subset V$ ein Untervektorraum, dann ist die Teilmenge

$$U^\perp := \{v \in V : v \perp u \ \forall u \in U\}$$

ein Komplement zu U . Der Unterraum U^\perp heißt das *orthogonale Komplement* von U ; wir haben dies in Definition 10.9 bereits angeführt. Die Projektion $p = p_U : V \rightarrow U$ bezüglich U^\perp heißt die *orthogonale Projektion*.

Ist (u_1, \dots, u_r) eine Orthonormalbasis eines Untervektorraums $U \subset V$, so wird die orthogonale Projektion $p : V \rightarrow U$ durch die Formel

$$p(v) := \sum_{i=1}^r \langle u_i, v \rangle \cdot u_i$$

gegeben. Diese Aussage macht auch für den Fall $U = V$ Sinn. Man erhält so die folgende Aussage: Ist $\mathcal{B} = (u_1, \dots, u_n)$ eine Orthonormalbasis von V , so gilt für jeden Vektor $v \in V$ die Formel

$$v = \sum_{i=1}^n \langle u_i, v \rangle \cdot u_i.$$

Der Kooordinatenvektor von v bezüglich der Orthonormalbasis \mathcal{B} ist also

$$x = \begin{pmatrix} \langle u_1, v \rangle \\ \vdots \\ \langle u_n, v \rangle \end{pmatrix}.$$

Zum Abschluss dieser Erklärungen wollen wir am folgenden Beispiel noch den Zusammenhang zwischen Projektion und Skalarprodukt erklären.

Beispiel 87

Das Skalarprodukt kann man auch als Länge eines Vektors multipliziert mit der projizierten Länge eines anderen Vektors sehen. Dies geht so: Wir betrachten zwei Vektoren aus dem \mathbb{R}^2 mit den Koordinaten $x = (x_1, x_2)^T$ und $y = (y_1, y_2)^T$. Das Standardskalarprodukt dieser beiden Vektoren ist gegeben durch

$$\langle x, y \rangle = x_1 y_1 + x_2 y_2.$$

Die Länge von x ist gerade $|x| = \sqrt{x_1^2 + x_2^2}$. Ohne Beschränkung der Allgemeinheit sei $x_2 = 0$. Dann entspricht die Projektion von y auf x dem Vektor $y_p = (y_1, 0)$. Demnach ist

$$\langle x, y \rangle = x_1 y_1 = |x| |y_p|.$$

Dies zeigt das Gewünschte. ■

Zur Definition 10.11 der hermiteschen Form: Dies ist quasi die Bilinearform (siehe Definition 10.1) im komplexen Fall. Diese bezeichnet man in der Literatur auch häufig als *Sesquilinearform* und meint damit eine Funktion, die zwei Vektoren einen Skalarwert zuordnet und die linear in einem und semilinear im anderen ihrer beiden Argumente ist. Mehr wollen wir dazu nicht sagen, nur eine Anmerkung: Ist f eine hermitesche Form auf V , so folgt aus der hermiteschen Symmetrie, dass

$$q(v) := f(v, v) \in \mathbb{R}$$

eine reelle Zahl ist und zwar für alle $v \in V$. Die Funktion $q : V \rightarrow \mathbb{R}_{\geq 0}$ hat die beiden Eigenschaften

$$q(\lambda v) = |\lambda|^2 q(v) \quad \text{und} \quad q(v + w) = q(v) + q(w) + 2\operatorname{Re} f(v, w).$$

Hierbei meinen wir mit $\operatorname{Re} f(v, w)$ den Realteil von $f(v, w)$. Betrachtet man V als \mathbb{R} -Vektorraum, so ist q eine quadratische Form. Sie ist genau dann positiv definit, wenn f positiv definit ist.

Beispiel 88

Seien $I = [a, b] \subset \mathbb{R}$ ein abgeschlossenes Intervall und $\mathcal{C}(I, \mathbb{C})$ der \mathbb{C} -Vektorraum der stetigen, komplexwertigen Funktionen $f : I \rightarrow \mathbb{C}$. Durch

$$\langle f, g \rangle_I := \int_a^b \overline{f(t)} g(t) dt$$

wird eine positiv definite hermitesche Form auf V definiert. ■

Zur Definition 10.12 der Selbstadjungiertheit: Für die Adjungierte kann man relativ leicht die folgenden Eigenschaften zeigen:

$$(A + B)^* = A^* + B^*, \quad (AB)^* = B^* A^*, \quad (\lambda A)^* = \bar{\lambda} A^*, \quad (A^*)^* = A.$$

Beispiel 89

Seien $V = \mathbb{C}^n$ und $A = (a_{ij}) \in \mathcal{M}_{n,n}(\mathbb{C})$ eine selbstadjungierte Matrix. Dann ist die durch

$$\langle x, y \rangle_A := x^* A y = \sum_{i,j} a_{ij} \bar{x}_i y_j \tag{10.5}$$

definierte Abbildung $\langle \cdot, \cdot \rangle_A : V \times V \rightarrow \mathbb{C}$ eine hermitesche Form. Ist $A = E_n$ die Einheitsmatrix, so ergibt sich natürlich das Standardskalarprodukt. Wir zeigen zur Übung noch einmal, dass (10.5) wirklich eine hermitesche Form definiert. Es gilt:

$$\begin{aligned} \langle x + y, z \rangle_A &= (x + y)^* A z = (x^* + y^*) A z \\ &= x^* A z + y^* A z \\ &= \langle x, z \rangle_A + \langle y, z \rangle_A \end{aligned}$$

und

$$\begin{aligned}\langle \lambda x, y \rangle_A &= (\lambda x)^* A y \\ &= \bar{\lambda} x^* A y = \bar{\lambda} \cdot \langle x, y \rangle_A.\end{aligned}$$

Die hermitesche Symmetrie folgt aus der Selbstadjungiertheit von A :

$$\begin{aligned}\overline{\langle x, y \rangle_A} &= (x^* A y)^* = y^* A^* x \\ &= y^* A x = \langle y, x \rangle_A.\end{aligned}$$

■

Letzte Anmerkung zu dieser Definition: Es gibt recht viele unterschiedliche Notationen für die Adjungierte von A . Beispielsweise A^+ oder A^\dagger etc. Da müsst ihr also immer schauen, mit welchem Buch ihr es zu tun habt, und euch dem Autor anpassen :-).

Zur Definition 10.13 der darstellenden Matrix im Komplexen: Dies ist nichts Neues, sondern funktioniert genauso wie im Reellen. Dazu sei nicht mehr so viel gesagt. Siehe die Definition 10.3 und die Erklärungen.

Zur Definition 10.14 und 10.15 des unitären Raums und der unitären Matrix: Ein unitärer Raum im Komplexen ist das Analogon zum euklidischen Raum im Reellen. Der unitäre Standardvektorraum der Dimension n ist $V = \mathbb{C}^n$ zusammen mit dem Standardskalarprodukt

$$\langle x, y \rangle = x^* y.$$

Ein paar weitere Anmerkungen:

- i) Unitäre Matrizen sind nach Definition invertierbar, und mit $Q, P \in U_n$ sind auch die Matrizen $Q^{-1} = Q^*$ und PQ unitär. Deshalb ist U_n eine Untergruppe von $\mathrm{GL}_n(\mathbb{C})$. Für eine reelle Matrix $Q \in \mathcal{M}_{n,n}(\mathbb{R})$ gilt $Q^* = Q^T$. Eine reelle Matrix ist daher genau dann unitär, wenn sie orthogonal ist. Mit anderen Worten: Es gilt:

$$\mathrm{O}_n(\mathbb{R}) = U_n \cap \mathrm{GL}_n(\mathbb{R}).$$

- ii) Seien $(V, \langle \cdot, \cdot \rangle)$ ein endlichdimensionaler unitärer Raum und \mathcal{A} eine Orthonormalbasis von V . Für jede Basis \mathcal{B} von V gilt dann: \mathcal{B} ist genau dann eine Orthonormalbasis, wenn die Basiswechselmatrix $Q = T_{\mathcal{A}}^{\mathcal{B}}$ unitär ist.
- iii) Wendet man ii) auf den unitären Standardraum \mathbb{C}^n an, so erhält man die folgende Aussage: Eine Matrix $Q \in \mathcal{M}_{n,n}(\mathbb{C})$ ist genau dann unitär, wenn die Spalten von Q eine Orthonormalbasis des \mathbb{C}^n bilden.

Beispiel 90

Beispielsweise sind die Matrizen für $t \in \mathbb{R}$

$$A := \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}, \quad B = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}, \quad C = \begin{pmatrix} e^{it} & 0 \\ 0 & e^{-it} \end{pmatrix}$$

unitär, wie ihr euch überlegen solltet! Tipp: Nicht invertieren! ■

Zur Definition 10.16 der Adjungierten eines Endomorphismus: Sind $\phi, \psi : V \rightarrow V$ Endomorphismen mit Adjungierten ϕ^* und ψ^* , so gilt:

$$(\phi + \psi)^* = \phi^* + \psi^*, \quad (\lambda\phi)^* = \bar{\lambda}\phi^* \quad \forall \lambda \in \mathbb{C}, \quad (\phi^*)^* = \phi.$$

Zur Definition 10.17 von unitär, selbstadjungiert und normal: Bevor wir ein paar Beispiele geben, drei Anmerkungen:

- i) Ein Endomorphismus $\phi : V \rightarrow V$ ist genau dann unitär, wenn für alle $v, w \in V$

$$\langle \phi(v), \phi(w) \rangle = \langle v, w \rangle$$

gilt.

- ii) Unitäre Endomorphismen sind normal, selbstadjungierte ebenfalls.

- iii) Geometrisch bedeutet unitär, dass das Skalarprodukt Längen und Winkel erhält.

Beliebte Übungsaufgaben auf Übungszetteln oder in Klausuren sind, dass man gewisse Behauptungen gibt und die Studenten entscheiden lässt, ob diese wahr (dann muss man es beweisen) oder falsch (dann muss man ein Gegenbeispiel geben) sind. Als Mathematiker solltet ihr nämlich im Laufe eures Studiums ein „Gefühl“ entwickeln, ob gewisse Aussagen richtig oder falsch sind. Also versuchen wir uns dran. Seien $A, B \in \mathcal{M}_{n,n}(\mathbb{C})$ zwei Matrizen.

- Wir behaupten: Mit A und B ist auch $A + B$ unitär. Diese Behauptung ist falsch. Beispielsweise sind E und $-E$ unitäre Matrizen, aber $-E + E = 0 \notin U_n$, denn die Nullmatrix ist nicht unitär. Ein anderes Gegenbeispiel liefert die unitäre Matrix

$$A = B := \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}.$$

Es gilt:

$$A + B = 2 \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix}.$$

Demnach gilt:

$$(A + B)^* = 2 \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}.$$

Also

$$\begin{aligned} (A + B)^* \cdot (A + B) &= 2 \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} 2 \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} \\ &= 4 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \notin E_2. \end{aligned}$$

- Wir behaupten: Mit A und B ist auch $A + B$ selbstadjungiert. Diese Behauptung ist richtig, denn da A und B selbstadjungiert sind, erhalten wir

$$(A + B)^* = A^* + B^* = A + B.$$

- Wir behaupten: Ist A selbstadjungiert, so ist $\det(A) \in \mathbb{R}$. Sei A selbstadjungiert, das heißt $A^* = A$. Wir wissen außerdem, dass $\det(A) = \det(A)^T$. Weiterhin gilt $\det(A^*) = \det(\overline{A})^T = \det(\overline{A})$. Insgesamt also

$$\det A = \det A^* = \det \overline{A} = \det A^T.$$

Aus der Leibniz-Formel der Determinante folgt $\det \overline{A} = \overline{\det A}$. Dies impliziert nun

$$\det A = \det \overline{A} = \overline{\det A}.$$

Insbesondere also $\det A = \overline{\det A}$. Eine Zahl ist aber nur dann gleich ihrem komplex Konjugierten, wenn sie reell ist. Daher ist die Behauptung richtig.

- Wir behaupten: Ist A unitär, so gilt $|\det(A)| = 1$. Auch diese Behauptung ist richtig. Dies müssen wir nur noch beweisen.

$$\begin{aligned} A \text{ unitär} &\Leftrightarrow AA^* = A\overline{A}^T = E_n \\ &\Leftrightarrow \det A \cdot \det \overline{A}^T = \det E_n = 1 \\ &\Leftrightarrow \det A \cdot \det \overline{A} = 1. \end{aligned}$$

Unter Benutzung der Leibniz-Formel der Determinante folgt, dass $\det \overline{A} = \overline{\det A}$. Demnach gilt $\det A \cdot \overline{\det A} = 1$, also $|\det A|^2 = 1$ und somit $|\det A| = 1$. Damit sind wir fertig.

Eine andere Möglichkeit des Beweises kann über den Spektralsatz (siehe Satz 10.17) geführt werden: Wir wissen, dass unitäre Matrizen insbesondere normal und damit unitär diagonalisierbar sind. Die Eigenwerte haben dann alle den Betrag 1. Damit folgt ebenfalls die Behauptung.

- Wir behaupten: Ist A selbstadjungiert, so folgt $\ker(A) = (\operatorname{im}(A))^\perp$. Auch diese letzte Aussage ist richtig. Man beweist sie am besten so: Seien $x \in \ker(A)$ und $y \in \mathbb{C}^n$ beliebig. Wir müssen zeigen, dass dann auch $x \in (\operatorname{im}(A))^\perp$. Da x im Kern von A liegt, gilt $\langle Ax, y \rangle = \langle 0, y \rangle = 0$. Da A selbstadjungiert ist, ergibt sich aber auch

$$\langle Ax, y \rangle = \langle x, A^* y \rangle = \langle x, Ay \rangle.$$

Nun ist $Ay \in (\operatorname{im}(A))$. Also ist $x \in (\operatorname{im}(A))^\perp$. Folglich gilt $\ker(A) \subset (\operatorname{im}(A))^\perp$.

Umgekehrt: Sei $z \in \operatorname{im}(A)$, das heißt $z = Ay$ für $y \in \mathbb{C}^n$, somit $x \in (\operatorname{im}(A))^\perp$. Zu zeigen ist, dass auch $x \in \ker(A)$ gilt. Man schließt nun

$$0 = \langle x, z \rangle = \langle x, Ay \rangle = \langle A^* x, y \rangle = \langle Ax, y \rangle.$$

Hieraus ergibt sich $Ax = 0$ und damit $x \in \ker(A)$, folglich $(\operatorname{im}(A))^\perp \subset \ker(A)$. Insgesamt folgt also die Behauptung $\ker(A) = (\operatorname{im}(A))^\perp$.

Beispiel 91

Und zum Schluss zur Erklärung dieser Definition noch einmal drei Behauptungen, die wir beweisen wollen, denn Beweisen macht Spaß oder etwa nicht? Sei $A \in \mathcal{M}_{n,n}(\mathbb{R})$ schiefssymmetrisch, das heißt $A^T = -A$.

- Wir behaupten: A ist normal. Wir müssen zeigen, dass $A^* A = A A^*$. Dabei ist $A^* = \overline{A}^T$. Da A reell ist, gilt $A^* = A^T$. Es gilt:

$$\begin{aligned} A^* A &= A^T A = -A A = -A^2, \\ A A^* &= A A^T = A(-A) = -A^2. \end{aligned}$$

Also ist jede schiefssymmetrische Matrix normal!

- Wir beweisen: Jeder Eigenwert von A ist rein imaginär, das heißt von der Form $t \cdot i$ mit $t \in \mathbb{R}$. Dies sieht man so: Sei λ ein Eigenwert der Matrix A zum Eigenvektor x . Dann gilt die folgende Gleichungskette:

$$\begin{aligned} \lambda \langle x, x \rangle &= \langle \lambda x, x \rangle = \langle Ax, x \rangle = \langle x, A^* x \rangle \\ &= \langle x, A^T x \rangle = \langle x, -Ax \rangle = \langle x, -\lambda x \rangle = -\bar{\lambda} \langle x, x \rangle. \end{aligned}$$

Nun kürzen wir auf beiden Seite $\langle x, x \rangle$ und erhalten $\lambda = -\bar{\lambda}$.

- Wir zeigen: Ist n ungerade, so hat A den Eigenwert 0. Es ist klar, dass eine quadratische Matrix mit ungerader Zeilen- und Spaltenanzahl immer mindestens einen reellen Eigenwert besitzt. Der Grund liegt im charakteristischen Polynom: Dessen Grad ist ungerade. Und ein Polynom mit maximalem ungeraden Grad hat immer genau eine Nullstelle im Reellen.

Dies folgt aus dem Zwischenwertsatz der Analysis. Aus dem letzten Aufgabenteil wissen wir aber, dass eine schiefsymmetrische Matrix rein imaginäre Eigenwerte besitzt. 0 ist aber die einzige Zahl, die gleichzeitig reell als auch rein imaginär ist. Dies folgt aus der Tatsache, dass sich (geometrisch gesehen) die reelle und die imaginäre Achse im Ursprung schneiden. ■

10.4 Erklärungen zu den Sätzen und Beweisen

Zum Satz 10.1 und 10.2 über die Darstellungsmatrix: Dieser Satz besagt einfach nur, dass f in dem durch \mathcal{B} definierten Koordinatensystem durch das in Beispiel 83 definierte Skalarprodukt $\langle \cdot, \cdot \rangle_A$ gegeben ist.

Beispiel 92

- Seien $V = \mathbb{Q}^2$ und $f := \langle \cdot, \cdot \rangle$ das Standardskalarprodukt. Dann ist die Darstellungsmatrix die Einheitsmatrix. Es gilt also:

$$\mathcal{M}_{\mathcal{E}}(f) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

wobei \mathcal{E} die Standardbasis bezeichnet.

- Sei $\mathcal{B} = (v_1, v_2)$ eine Basis mit

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

Dann gilt:

$$\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle = 1, \quad \langle v_1, v_1 \rangle = 2, \quad \langle v_2, v_2 \rangle = 5.$$

Demnach ist die Darstellungsmatrix gegeben durch

$$\mathcal{M}_{\mathcal{B}}(f) = \begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}.$$

Dies soll an Beispielen genügen. ■

Zum Satz 10.2 und 10.3 (Basiswechselsatz für Bilinearformen): Wir erklären den Satz an einem Beispiel.

Beispiel 93

Seien $V = \mathbb{Q}^2$ der Standardvektorraum der Dimension 2 über \mathbb{Q} und $f := \langle \cdot, \cdot \rangle$ das Standardskalarprodukt auf V . Ist $\mathcal{E} = (e_1, e_2)$ die Standardbasis von V , so gilt:

$$\mathcal{M}_{\mathcal{E}}(f) = E_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Nun sei $\mathcal{B} := (v_1, v_2)$ die Basis mit den Vektoren

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

Die Transformationsmatrix des Basiswechsels von \mathcal{B} nach \mathcal{E} ist also

$$Q := T_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}.$$

Aus Satz 10.3 folgt nun

$$B := \mathcal{M}_{\mathcal{B}}(f) = Q^T Q = \begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}.$$

Wie erwartet ist B eine symmetrische Matrix mit den Einträgen

$$\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle = 1, \quad \langle v_1, v_1 \rangle = 2, \quad \langle v_2, v_2 \rangle = 5.$$

Vergleiche auch das Beispiel 92. ■

Beispiel 94

Sei $\mathcal{B} = (v_1, v_2)$ eine Basis des Vektorraums \mathbb{Q}^2 mit

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad v_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

Dann haben wir schon errechnet, dass

$$\langle v_1, v_1 \rangle = 2, \quad \langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle = 1, \quad \langle v_2, v_2 \rangle = 5.$$

Folglich kann man die darstellende Matrix ablesen zu

$$\mathcal{M}_{\mathcal{B}}(f) = \begin{pmatrix} 2 & 1 \\ 1 & 5 \end{pmatrix}.$$

Wir wollen jetzt den Basiswechselssatz anwenden und stellen nun die Vektoren v_1 und v_2 als Linearkombinationen der Einheitsvektoren dar, also bezüglich der Standardbasis \mathcal{E} . Es gilt demnach:

$$v_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = e_1 + e_2 \quad \text{und} \quad v_2 = \begin{pmatrix} -1 \\ 2 \end{pmatrix} = -e_1 + 2e_2.$$

Die Basiswechselmatrix lautet folglich

$$Q := T_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 1 & -1 \\ 1 & 2 \end{pmatrix}.$$

Wir erhalten demnach

$$\begin{aligned} v &= x'_1 v_1 + x'_2 v_2 = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \Rightarrow x = Qx', \\ w &= y'_1 v_1 + y'_2 v_2 = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \Rightarrow y = Qy'. \end{aligned}$$

Hierbei sind $x' = \begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix}$ und $y' = \begin{pmatrix} y'_1 \\ y'_2 \end{pmatrix}$. Bei dieser Rechnung haben wir einfach ausgenutzt, dass man jeden Vektor des Vektorraums auf eindeutige Weise als Linearkombination der Basisvektoren darstellen kann. Weiterhin gilt:

$$\begin{aligned} f(v, w) &= x^T E y = x^T y = (Qx')^T Qy' \\ &= (x')^T Q^T Qy' = (x')^T (Q^T Q)y' \\ &\Rightarrow M_B(f) = Q^T Q. \end{aligned}$$

■

Zum Satz 10.4 und dem Gram-Schmidt-Verfahren (Satz 10.5): Mit Gram-Schmidt werdet ihr in jeder Vorlesung zur Linearen Algebra 2 und in den Übungen „gequält“ werden. Seht dies aber bitte nicht als Quälen, sondern genießt es (auch wenn es schwer fällt :-D), denn dieses Orthogonalisierungsverfahren ist sehr mächtig und muss eingeübt werden. Wir betrachten daher eine Menge an Beispielen.

Beispiel 95

Sei $I = [a, b] \subset \mathbb{R}$ ein abgeschlossenes und endliches Intervall. Mit $\mathcal{C}(I)$ bezeichnen wir den Vektorraum der stetigen Funktionen $f : I \rightarrow \mathbb{R}$. Dann wird durch

$$\langle f, g \rangle_I := \int_a^b f(x)g(x)dx$$

eine symmetrische, positiv definite Bilinearform auf $\mathcal{C}(I)$ erklärt. Wir nehmen an, dass $I = [0, 1]$, also $a = 0$ und $b = 1$. Sei

$$V = \langle 1, x, x^2 \rangle$$

der Untervektorraum von $\mathcal{C}(I)$ aller Polynomfunktionen vom Grad kleiner gleich 2. Sei $\mathcal{A} = (1, x, x^2)$ die Standardbasis von V . Wegen

$$\langle x^i, x^j \rangle_I = \int_0^1 x^{i+j} dx = \left[\frac{1}{i+j+1} \cdot x^{i+j+1} \right]_0^1 = \frac{1}{i+j+1}$$

ist die Darstellungsmatrix von $\langle \cdot, \cdot \rangle_I$ gerade gegeben durch

$$A = \mathcal{M}_{\mathcal{A}}(\langle \cdot, \cdot \rangle_I) = \begin{pmatrix} \langle f_1, f_1 \rangle & \langle f_1, f_2 \rangle & \langle f_1, f_3 \rangle \\ \langle f_2, f_1 \rangle & \langle f_2, f_2 \rangle & \langle f_2, f_3 \rangle \\ \langle f_3, f_1 \rangle & \langle f_3, f_2 \rangle & \langle f_3, f_3 \rangle \end{pmatrix} = \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix}.$$

Etwas abweichend zum Gram-Schmidt-Verfahren werden wir zunächst nur eine orthogonale Basis $\mathcal{B} := (f_1, f_2, f_3)$ von V bestimmen. Unser Ansatz lautet

$$f_1 = 1, \quad f_2 = x + a, \quad f_3 = x^2 + bx + c. \quad (10.6)$$

Die Zahlen $a, b, c \in \mathbb{R}$ sind noch zu bestimmen. Die erste Orthogonalitätsrelation lautet:

$$\langle f_2, f_1 \rangle_I = \int_0^1 (x + a) dx = \frac{1}{2} + a = 0.$$

Dies ist offenbar genau dann erfüllt, wenn $a = -1/2$. Damit ist also $f_2 = x - 1/2$. Die zweite Orthogonalitätsrelation lautet:

$$\langle f_3, f_1 \rangle_I = \int_0^1 (x^2 + bx + c) dx = \frac{1}{3} + \frac{1}{2}b + c = 0. \quad (10.7)$$

Die dritte lautet entsprechend:

$$\begin{aligned} \langle f_3, f_2 \rangle_I &= \int_0^1 (x^2 + bx + c)(x - 1/2) dx \\ &= \int_0^1 \left(x^3 + \left(b - \frac{1}{2}\right)x^2 + \left(c - \frac{1}{2}b\right)x - \frac{1}{2}c \right) dx \\ &= \frac{1}{4} + \left(b - \frac{1}{2}\right)\frac{1}{3} + \left(c - \frac{1}{2}b\right)\frac{1}{2} - \frac{1}{2}c \\ &= \frac{1}{12} + \frac{1}{12}b = 0. \end{aligned} \quad (10.8)$$

Die beiden Gleichungen (10.7) und (10.8) bilden ein lineares Gleichungssystem, das man lösen kann. Die Lösung ist $b = -1$ und $c = 1/6$. Die gesuchte orthogonale Basis besteht also aus den Funktionen

$$f_1(x) = 1, \quad f_2(x) = x - 1/2, \quad f_3(x) = x^2 - x + 1/6.$$

Die Basiswechselmatrix ist

$$Q := T_{\mathcal{A}}^{\mathcal{B}} = \begin{pmatrix} 1 & -1/2 & 1/6 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Es gilt demnach also für die Darstellungsmatrix

$$\mathcal{M}_{\mathcal{B}}(\langle \cdot, \cdot \rangle_I) = Q^T \cdot A \cdot Q = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/12 & 0 \\ 0 & 0 & 1/180 \end{pmatrix}.$$

Die Diagonaleinträge sind gerade die Normen der Funktionen f_i . Eine Orthonormalbasis $\tilde{\mathcal{B}} = (\tilde{f}_1, \tilde{f}_2, \tilde{f}_3)$ ist also gegeben durch die Funktionen

$$\tilde{f}_1(x) = 1, \quad \tilde{f}_2(x) = 2\sqrt{3} \left(x - \frac{1}{2} \right), \quad \tilde{f}_3(x) = 6\sqrt{5} \left(x^2 - x + \frac{1}{6} \right).$$

■

Beispiel 96

Sei $f = \langle \cdot, \cdot \rangle_A$ das symmetrische Skalarprodukt auf $V = \mathbb{R}^3$ zur Matrix

$$A := \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \in \mathcal{M}_{3,3}(\mathbb{R}).$$

Wir wollen nun eine Orthonormalbasis von V bzgl. f angeben und eine Matrix $P \in \text{GL}_3(\mathbb{R})$ finden mit $A = P^T P$.

Wir gehen einfach von der kanonischen Standardbasis des \mathbb{R}^3 aus, also von

$$\mathcal{E} = (e_1, e_2, e_3) = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right).$$

Also starten wir. Wir wenden das Gram-Schmidt-Verfahren an, um eine ONB (v_1, v_2, v_3) zu finden. Hierzu setzen wir als ersten Vektor einfach

$$v_1 = e_1.$$

Die Länge des Vektors ergibt sich durch die Darstellungsmatrix. Es gilt demnach

$$\langle v_1, v_1 \rangle_A = e_1^T A e_1 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 2.$$

Dieser Schritt ist wichtig! Viele denken vielleicht jetzt: „Hä? Der Einheitsvektor e_1 hat doch Länge 1?“ Das ist auch richtig, aber nur bzgl. des Standardskalarprodukts! Wir verwenden jetzt ein anderes Skalarprodukt, das durch die Matrix

A repräsentiert wird und damit ergibt sich auch eine andere Länge. Den zweiten Vektor der ONB bestimmen wir nach dem Orthogonalisierungsverfahren von Gram-Schmidt. Wir erhalten somit

$$\begin{aligned}
 \tilde{v}_2 &= e_2 - \frac{\langle v_1, e_2 \rangle_A}{\langle v_1, v_1 \rangle_A} v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\rangle_A}{\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle_A} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\
 &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{\begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}}{\begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\
 &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{\begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}}{\begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 1 \\ 0 \end{pmatrix}.
 \end{aligned}$$

Somit haben wir also den zweiten Vektor der Orthogonalbasis

$$\tilde{v}_2 = \begin{pmatrix} -1/2 \\ 1 \\ 0 \end{pmatrix}$$

berechnet. Um nicht mit Brüchen rechnen zu müssen, können wir diesen Vektor vervielfachen, ohne die Orthogonalitätseigenschaft zu verlieren. Wir wählen also

$$v_2 = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}.$$

Die Länge des Vektors ist nun

$$\langle v_2, v_2 \rangle_A = \left\langle \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \right\rangle_A = \begin{pmatrix} -1 & 2 & 0 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} = 6.$$

Den dritten, noch fehlenden Vektor bestimmen wir so:

$$\begin{aligned} \tilde{v}_3 &= e_3 - \frac{\langle v_2, e_3 \rangle_A}{\langle v_2, v_2 \rangle_A} v_2 - \frac{\langle v_1, e_3 \rangle_A}{\langle v_1, v_1 \rangle_A} v_1 \\ &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle_A}{\left\langle \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \right\rangle_A} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} - \frac{\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle_A}{\left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle_A} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} - \frac{2}{6} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} - \frac{0}{2} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/3 \\ -2/3 \\ 1 \end{pmatrix}. \end{aligned}$$

Da wir auch hier wieder nicht mit Brüchen arbeiten wollen, verwenden wir

$$v_3 = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}.$$

Die Länge dieses Vektors ist gegeben durch

$$\langle v_3, v_3 \rangle_A = \begin{pmatrix} 1 & -2 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} = 12.$$

Die gesuchte Orthonormalbasis lautet also

$$\mathcal{B} := \left(\frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{12}} \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} \right).$$

Wir wollen jetzt noch die invertierbare Matrix $P \in \text{GL}_n(\mathbb{R})$ mit $A = P^T P$ angeben. Hierzu gibt es zwei Möglichkeiten:

1. Möglichkeit: Wir wissen, dass

$$S^T AS = (\langle v_i, v_j \rangle_A)$$

die darstellende Matrix bzgl. der Basisvektoren aus der Orthonormalbasis ist, die wir gerade bestimmt haben, wobei in der Matrix S in den Spalten die orthogonalen Basisvektoren (noch nicht normiert) stehen. Weiterhin können wir diese darstellende Matrix natürlich auch so direkt angeben, denn es muss auf jeden Fall eine Diagonalmatrix sein, weil die Basisvektoren gerade so konstruiert sind, dass $\langle v_i, v_j \rangle_A = 0$ für alle $i \neq j$ ist. Und in den Diagonaleinträgen stehen die entsprechenden Längen der Vektoren. Es gilt also zusammengefasst:

$$\begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & 0 \\ 1 & -2 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 0 & 2 & -2 \\ 0 & 0 & 3 \end{pmatrix} = (\langle v_i, v_j \rangle_A) = \underbrace{\begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{pmatrix}}_{=:D}.$$

Es ist auch

$$DD = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{pmatrix} \cdot \begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{pmatrix}.$$

Was haben wir dadurch gewonnen? Überlegen wir allgemein:

$$\begin{aligned} S^T AS &= DD = D^T D \quad (\text{Jede Diagonalmatrix ist symmetrisch}) \\ \Rightarrow A &= (S^T)^{-1} D^T D S^{-1} = (S^{-1})^T D^T D S^{-1} = \underbrace{(DS^{-1})^T (DS^{-1})}_{=: P^T \cdot P}. \end{aligned}$$

Um P nun anzugeben, müssen wir S^{-1} errechnen. Mit dem Gauß-Verfahren, das ihr noch aus der Linearen Algebra 1 kennen solltet, erhält man

$$S^{-1} = \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1/3 \\ 0 & 0 & 1/3 \end{pmatrix}$$

und damit

$$P = DS^{-1} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 12 \end{pmatrix} \begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 1/2 & 1/3 \\ 0 & 0 & 1/3 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & \sqrt{2}/2 & 0 \\ 0 & \sqrt{6}/2 & \sqrt{6}/3 \\ 0 & 0 & 2\sqrt{3}/3 \end{pmatrix}.$$

Fassen wir die erste Möglichkeit noch einmal zusammen:

- i) Wähle die kanonische Standardbasis als „Ausgangsbasis“.

- ii) Beginne mit dem ersten kanonischen Einheitsbasisvektor und wende Gram-Schmidt an.
- iii) Wir verzichten auf die Normierung der Vektoren, da dies später zum Rechnen einfacher ist und wir nicht mit Wurzeln rechnen müssen.
- iv) Bestimme die „Länge“ der Vektoren mithilfe der darstellenden Matrix, bedenke, es gilt $\langle x, x \rangle_A = x^T A x$.
- v) Es gilt $S^T A S = D$, wobei A die darstellende Matrix ist, und in den Spalten von S stehen die gefundenen orthogonalen Basisvektoren. D ist eine Diagonalmatrix, die als Einträge die „Längen“ der Vektoren enthält.
- vi) Eine invertierbare Matrix $P \in \text{GL}_n(\mathbb{R})$ mit $A = P^T P$ finden wir nun, wenn wir $P := D S^{-1}$ berechnen.

Ohne Beweis merken wir an, dass man durch dieses Verfahren bei geeigneter Normierung die Legendre-Polynome erhält, die euch bestimmt einmal begegnen werden, beispielsweise in einer Numerik-Vorlesung.

2. Möglichkeit: Wir haben versprochen, dass wir noch eine zweite Möglichkeit geben, um eine Matrix $P \in \text{GL}_n(\mathbb{R})$ mit $A = P^T P$ anzugeben. Und da man Versprechen halten soll, machen wir dies nun auch. Zunächst bestimmen wir die Orthonormalbasis ganz analog wie in der ersten Möglichkeit geschehen ist, mithilfe des Gram-Schmidt-Verfahrens. Weiterhin wissen wir, dass

$$Q^T A Q = E \quad \text{für } Q := T_{\mathcal{E}}^{\mathcal{B}}$$

gilt, wobei \mathcal{E} die Standardbasis bezeichnet. Die Transformationsmatrix (Basiswechselmatrix) Q kann man sofort angeben. Es gilt:

$$Q = T_{\mathcal{E}}^{\mathcal{B}} = \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{12} \\ 0 & 2/\sqrt{6} & -2/\sqrt{12} \\ 0 & 0 & 3/\sqrt{12} \end{pmatrix}.$$

Nun gilt:

$$Q^T A Q = E \Rightarrow A = (Q^T)^{-1} E Q^{-1} = (Q^T)^{-1} Q^{-1} =: P^T P.$$

Um P zu berechnen, bestimmt man also $P = (T_{\mathcal{E}}^{\mathcal{B}})^{-1} = T_{\mathcal{B}}^{\mathcal{E}}$. Auch hier wollen wir die Vorgehensweise zusammenfassen:

- i) Bestimme die Orthonormalbasis wie bei der ersten Möglichkeit mithilfe von Gram-Schmidt.
- ii) Bestimme $Q := T_{\mathcal{E}}^{\mathcal{B}}$.
- iii) Es gilt $P = (T_{\mathcal{E}}^{\mathcal{B}})^{-1} = T_{\mathcal{B}}^{\mathcal{E}}$.

Wieso gerade zwei Möglichkeiten, um P zu berechnen? Naja, man muss schauen, was man lieber mag. Bei der zweiten Möglichkeit müssen wir oft mit Wurzeln und Brüchen rechnen, was viel fehleranfälliger ist, jedenfalls für uns Autoren :-). Entscheidet also selbst.

■

Beispiel 97

Zum Schluss noch ein letztes Beispiel mit einem etwas exotischeren Skalarprodukt (naja, so exotisch sind sie eigentlich gar nicht. Es ist das einfachste Skalarprodukt auf einem Funktionenraum, aber da ihr noch keine Funktionalanalysis gehört habt, ist es für euch schon noch exotisch). Seien $\mathcal{C}([-1, 1], \mathbb{R})$ der Vektorraum der stetigen Funktionen auf dem Intervall $I = [-1, 1]$ und $\langle \cdot, \cdot \rangle_I$ das Skalarprodukt

$$\langle f, g \rangle_I := \int_{-1}^1 f(x)g(x)dx.$$

Wir wollen kurz zeigen, dass dies wirklich ein Skalarprodukt definiert. Die Symmetrie ist klar. Nur über die positive Definitheit könnte man etwas stolpern. Aber mit ein wenig Analysis-Kenntnissen sieht man dies auch und zwar so: Zu zeigen ist, dass

$$\langle f, f \rangle_I = \int_{-1}^1 f(x)f(x)dx = \int_{-1}^1 f^2(x)dx > 0.$$

Aus der Analysis weiß man aber: Ist f stetig auf einem Intervall $[a, b]$ und $f(t) \neq 0$ für ein $t \in [a, b]$, so gilt:

$$\int_a^b f^2(t)dt > 0.$$

Damit ist alles gezeigt.

Sei nun V der Untervektorraum der Polynome vom Grad kleiner gleich 2. Zu bestimmen ist eine Orthonormalbasis von V und die Darstellungsmatrix des Skalarprodukts $\langle \cdot, \cdot \rangle_I$ bzgl. der Basis $\mathcal{B} = (1, x, x^2)$.

1. Möglichkeit: Wir sehen sofort, dass 1 und x schon orthogonal sind, denn es gilt:

$$\langle 1, x \rangle_I = \int_{-1}^1 1 \cdot x dx = \int_{-1}^1 x dx = 0.$$

Wir brauchen mit dem Gram-Schmidt-Verfahren also nur noch den dritten Vektor angeben, der jeweils paarweise orthogonal auf 1 bzw. x steht. Es gilt:

$$v_3 = x^2 - \frac{\langle x, x^2 \rangle_I \cdot x}{\langle x, x \rangle_I} - \frac{\langle 1, x^2 \rangle_I}{\langle 1, 1 \rangle_I} \cdot 1 = x^2 - \frac{1}{3}$$

Nun müssen wir diesen Vektor noch normieren. Also müssen wir seine „Länge“ berechnen. Nun könnten wir natürlich straight forward vorgehen und

$$\left\langle x^2 - \frac{1}{3}, x^2 - \frac{1}{3} \right\rangle_I = \int_{-1}^1 \left(x^2 - \frac{1}{3} \right)^2 dx$$

berechnen. Angenommen, wir hätten die Darstellungsmatrix des Skalarprodukts $\langle \cdot, \cdot \rangle_I$ gegeben, so können wir die Länge sofort berechnen. Die Darstellungsmatrix lautet

$$\begin{pmatrix} 2 & 0 & 2/3 \\ 0 & 2/3 & 0 \\ 2/3 & 0 & 2/5 \end{pmatrix} = \begin{pmatrix} \langle 1, 1 \rangle & \langle 1, x \rangle & \langle 1, x^2 \rangle \\ \langle x, 1 \rangle & \langle x, x \rangle & \langle x, x^2 \rangle \\ \langle x^2, 1 \rangle & \langle x^2, x \rangle & \langle x^2, x^2 \rangle \end{pmatrix}.$$

Die Länge berechnet sich zu

$$\left\langle x^2 - \frac{1}{3}, x^2 - \frac{1}{3} \right\rangle_I = \begin{pmatrix} -1/3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 2/3 \\ 0 & 2/3 & 0 \\ 2/3 & 0 & 2/5 \end{pmatrix} \begin{pmatrix} -1/3 \\ 0 \\ 1 \end{pmatrix} = \frac{8}{45}.$$

Die gesuchte Orthonormalbasis ergibt sich damit zu

$$\mathcal{B} = \left(\frac{1}{\sqrt{2}} \cdot 1, \frac{\sqrt{3}}{\sqrt{2}} \cdot x, \frac{\sqrt{45}}{\sqrt{8}} \cdot \left(x^2 - \frac{1}{3} \right) \right).$$

2. Möglichkeit: Wir verfolgen nun noch einmal den Ansatz in Beispiel 95. Wir bestimmen eine orthogonale Basis $\mathcal{B}_{\text{orth}} = (f_1, f_2, f_3)$ von V . Unser Ansatz lautet wieder

$$f_1 = 1, \quad f_2 = x + a, \quad f_3 = x^2 + bx + c$$

mit noch zu bestimmenden reellen Zahlen $a, b, c \in \mathbb{R}$. Die erste Orthogonalitätsrelation lautet

$$\langle f_2, f_1 \rangle_I = \int_{-1}^1 f_2(x) f_1(x) dx = \int_{-1}^1 (x + a) dx = 2a = 0.$$

Hieraus ergibt sich sofort $a = 0$, und damit ist $f_2(x) = x + 0 = x$ festgelegt. Die zweite Orthogonalitätsrelation lautet

$$\langle f_3, f_1 \rangle_I = \int_{-1}^1 f_3(x) f_1(x) dx = \int_{-1}^1 (x^2 + bx + c) dx = \frac{2}{3} + 2c = 0.$$

Hieraus ermittelt man $c = -1/3$. Mit der dritten Orthogonalitätsrelation erhalten wir nun auch b durch

$$\langle f_3, f_2 \rangle_I = \int_{-1}^1 f_3(x) f_2(x) dx = \int_{-1}^1 (x^2 + bx + c) x dx = \frac{2}{3} \cdot b = 0.$$

Demnach ist $b = 0$ und folglich $f_3(x) = x^2 - \frac{1}{3}$. Unsere orthogonale Basis lautet also

$$\mathcal{B}_{\text{orth}} = \left(1, x, x^2 - \frac{1}{3} \right).$$

Die Basiswechsellmatrix (Transformationsmatrix) ist

$$Q := T_{\mathcal{B}}^{\mathcal{B}_{\text{orth}}} = \begin{pmatrix} 1 & 0 & -1/3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Nach dem Basiswechselsatz 10.3 ergibt sich die Darstellungsmatrix bzgl. der orthogonalen Basis $\mathcal{B}_{\text{orth}}$ sofort durch

$$\mathcal{M}_{\mathcal{B}_{\text{orth}}}(\langle \cdot, \cdot \rangle_I) = Q^T A Q = (T_{\mathcal{B}}^{\mathcal{B}_{\text{orth}}})^T A T_{\mathcal{B}}^{\mathcal{B}_{\text{orth}}}.$$

Dies rechnen wir einfach nur aus. Der Vorteil davon ist, dass die Diagonaleinträge gerade die Normen der Funktionen f_i mit $i = 1, 2, 3$ sind. Wir erhalten

$$\begin{aligned} \mathcal{M}_{\mathcal{B}_{\text{orth}}} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1/3 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 2/3 \\ 0 & 2/3 & 0 \\ 2/3 & 0 & 2/5 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1/3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2/3 & 0 \\ 0 & 0 & 8/45 \end{pmatrix}. \end{aligned}$$

In den Diagonaleinträgen stehen die entsprechenden Normen. Die gesuchte Orthonormalbasis ergibt sich damit auch hier zu

$$\mathcal{B} = \left(\frac{1}{\sqrt{2}} \cdot 1, \frac{\sqrt{3}}{\sqrt{2}} x, \frac{\sqrt{45}}{\sqrt{8}} \left(x^2 - \frac{1}{3} \right) \right).$$

■

Nach diesen Beispielen hat man vielleicht schon eine kleine Vorstellung, was das Gram-Schmidt-Verfahren anschaulich eigentlich macht. Wir wollen dies noch einmal geometrisch deuten: Gegeben seien zwei Basisvektoren der Ebene, sprich des \mathbb{R}^2 , die schon orthogonal und normiert sind.

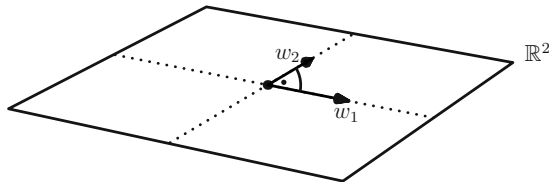


Abb. 10.1: Zwei schon normierte und orthogonale Vektoren im \mathbb{R}^2 .

Wir wollen nun einen dritten Basisvektor konstruieren, sodass die drei Vektoren eine Orthonormalbasis des \mathbb{R}^3 bilden. Also sei ein zusätzlicher Vektor v_k zum Beispiel wie in 10.2 gegeben. Wir wenden das Gram-Schmidt-Verfahren nun so

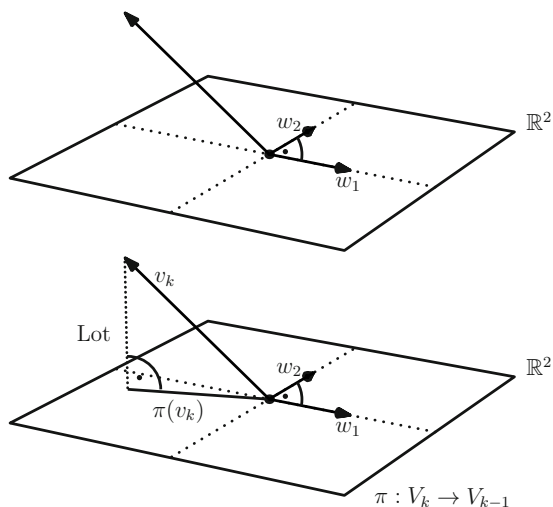


Abb. 10.2: Wir wenden Gram-Schmidt an.

an, dass wir zunächst das Lot vom Vektorenendpunkt fällen. Dabei bezeichnet $\pi : V_k \rightarrow V_{k-1}$ die senkrechte Projektion, siehe Abbildung 10.2. Nun verschieben wir das Lot auf den Anfangspunkt der beiden anderen Vektoren, siehe Abbildung 10.3. Zum Schluss wird dieser Vektor v_k noch normiert, also auf die Länge 1

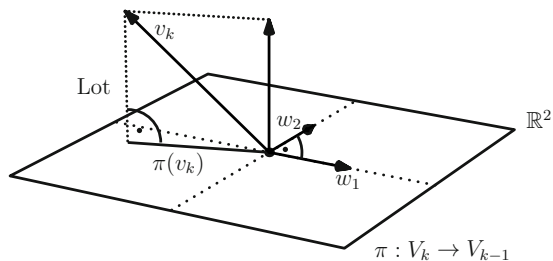


Abb. 10.3: Verschieben des Lotes.

gebracht, siehe Abbildung 10.4.

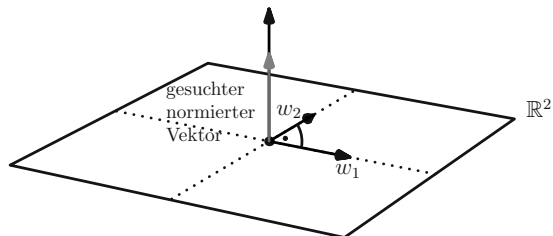


Abb. 10.4: Normieren des Vektors.

Wir fassen zusammen: Konkret sieht der Algorithmus nämlich so aus: Sei $\mathcal{A} = (v_1, \dots, v_n)$ die Ausgangsbasis. Dann ist $\mathcal{B} = (w_1, \dots, w_n)$ eine orthogonale

Basis, wenn man induktiv die Vektoren w_1, w_2, \dots folgendermaßen bestimmt. Hat man die Vektoren w_1, \dots, w_{k-1} bereits bestimmt, so setzt man zunächst

$$w'_k := v_k - \sum_{i=1}^{k-1} \langle v_k, w_i \rangle \cdot w_i$$

und anschließend

$$w_k := \frac{w'_k}{\sqrt{\|w'_k\|}}.$$

Zum Satz 10.6: Dieser Satz gibt uns keinen konkreten Algorithmus, um zu entscheiden, wann eine reelle symmetrische Matrix A positiv definit ist. Aber dafür haben wir das Hauptminorenkriterium, siehe Satz 10.7.

Zum Hauptminorenkriterium (Satz 10.7): Wie wir in Beispiel 85 gesehen haben, ist die Symmetrie wesentlich für die Aussage des Satzes.

Beispiel 98

- Wir betrachten die Matrix

$$A := \begin{pmatrix} 1 & 4 & 1 \\ 4 & 17 & 2 \\ 1 & 2 & 6 \end{pmatrix}.$$

Diese Matrix ist positiv definit wie aus dem Hauptminorenkriterium folgt, denn die Determinanten aller Hauptminoren sind positiv.

$$\det 1 = 1 > 0, \quad \det \begin{pmatrix} 1 & 4 \\ 4 & 17 \end{pmatrix} = 1 > 0, \quad \det \begin{pmatrix} 1 & 4 & 1 \\ 4 & 17 & 2 \\ 1 & 2 & 6 \end{pmatrix} = 1 > 0.$$

- Sei nun die Matrix

$$B := \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$

gegeben. Wir wollen zeigen, dass diese reelle symmetrische Matrix, positiv definit ist. Dies ist erfüllt, da alle Hauptminoren positive Determinante besitzen, denn

$$\det 2 = 2 > 0, \quad \det \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} = 3 > 0, \quad \det \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix} = 4 > 0.$$

Wir denken, dass dies an Beispielen genügen sollte. ■

Zu den Dimensionsformeln (Satz 10.9): Um die Formeln einzuüben, betrachten wir ein einfaches Beispiel.

Beispiel 99

- Wir wollen untersuchen, ob V die direkte Summe von U und W ist. Dazu sei zunächst einmal $V = \mathbb{R}^3$ und

$$U = \langle (1, -3, 1), (0, 1, 0) \rangle \quad \text{und} \quad W = \langle (2, 1, 1), (3, 2, 1) \rangle.$$

Wir können uns überlegen, dass die Vektoren in U als auch in W jeweils linear unabhängig sind (nachvollziehen!). Demnach ist also $\dim(U) = \dim(W) = 2$. Mit der Dimensionsformel (Satz 10.9) ist also V nicht die direkte Summe von U und W , denn $U \oplus W$ muss ein Untervektorraum vom \mathbb{R}^3 sein.

Betrachten wir nun noch einmal $V = \mathbb{C}^4$ und

$$U = \langle (1, i, i+1, -i), (i, -1, i-1, 1) \rangle \quad \text{und} \\ W = \langle (2+i, i, -i, 1), (-i, 3i, 1, i) \rangle.$$

Die Vektoren aus U sind nicht linear unabhängig, denn es gilt:

$$i \cdot \begin{pmatrix} 1 \\ i \\ i+1 \\ -i \end{pmatrix} = \begin{pmatrix} i \\ -1 \\ i-1 \\ 1 \end{pmatrix}.$$

Also ist $\dim(U) = 1$. Da aber $\dim(W) \leq 2$ und $\dim(\mathbb{C}^4) = 4$, ergibt sich wieder aus der Dimensionsformel, dass V nicht die direkte Summe von U und W sein kann. ■

Zum Satz 10.11: Dieser Satz besagt, dass der Vektor $p(v)$ die beste Approximation von v durch einen Vektor in U ist. Betrachten wir ein Beispiel.

Beispiel 100

Im \mathbb{R}^3 betrachten wir die Ebene E mit der Gleichung

$$E : 2x_1 + x_2 - x_3 = 2$$

und den Punkt $P := (2, 1, 1)$. Wir möchten jetzt den Abstand von P zur Ebene berechnen. Ihr kennt dies vielleicht schon aus der Schule. Wir wollen es aber ein wenig strukturierter und mathematischer machen, indem wir gleichzeitig den Fußpunkt des Lotes von P auf die Ebene E bestimmen. Wir bezeichnen mit $U \subset \mathbb{R}^3$ den Raum der Richtungsvektoren von E , das heißt den Lösungsraum der homogenen Gleichung

$$2x_1 + x_2 - x_3 = 0.$$

Durch eine kurze Rechnung findet man eine Orthonormalbasis (u_1, u_2) von U , etwa

$$u_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad u_2 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Wir wählen einen Aufpunkt von E , etwa $O = (1, 0, 0)$, und nun können wir einen allgemeinen Punkt von E in der Form

$$Q = O + u, \quad u = \lambda_1 \cdot u_1 + \lambda_2 \cdot u_2$$

schreiben. Entsprechend schreiben wir

$$P = O + v, \quad v = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Die Distanz von P zu Q ist damit

$$d(P, Q) = \|v - u\|.$$

Nach Satz 10.11 wird dieser Abstand minimiert durch die Wahl

$$Q := P^* = O + p_U(u),$$

wobei P^* der Fußpunkt des Lotes von P auf E ist. Nun verwenden wir die Formel für die orthogonale Projektion $p_U(v)$, die wir in den Erklärungen zur Definition 10.10 erläutert und hergeleitet haben. Es ergibt sich

$$\begin{aligned} p(v) &= \langle v, u_1 \rangle \cdot u_1 + \langle v, u_2 \rangle \cdot u_2 \\ &= \frac{2}{2} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + \frac{1}{3} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}. \end{aligned}$$

Wir erhalten

$$P^* = O + p_U(v) = \frac{2}{3} \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}$$

und

$$d(P, E) = \|v - p_U(v)\| = \left\| \begin{pmatrix} 2/3 \\ 1/3 \\ -1/3 \end{pmatrix} \right\| = \frac{\sqrt{6}}{3}.$$

■

Zum Basiswechselsatz im Komplexen (Satz 10.12 und 10.13): Diese Sätze kennen wir schon aus den Sätzen 10.1 und 10.2, nur dass wir diese nun für den komplexen Fall formuliert haben.

Zum Gram-Schmidt-Verfahren im Komplexen (Satz 10.14 und 10.15): Das Gram-Schmidt-Verfahren funktioniert für eine hermitesche Form ganz genauso wie im reellen Fall. Nur dass man beim Skalarprodukt ein wenig aufpassen muss, denn irgendwann muss einmal das komplex Konjugierte eines Vektors bestimmt werden und dies geht, indem man jeden Vektoreintrag komplex konjugiert.

Beispiel 101

Der \mathbb{C}^4 sei mit der hermiteschen Form $\langle x, y \rangle = \sum x_k \bar{y}_k$ versehen. Zu bestimmen ist eine Orthonormalbasis des Unterraums U , der von den Vektoren

$$\begin{pmatrix} 1 \\ 0 \\ i \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \\ -i \end{pmatrix}, \begin{pmatrix} i \\ -2 \\ -2 \\ 0 \end{pmatrix}$$

aufgespannt wird. Wir berechnen mit dem Gram-Schmidt-Verfahren:

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ i \\ -1 \end{pmatrix},$$

$$\langle v_1, v_1 \rangle = \left\langle \begin{pmatrix} 1 \\ 0 \\ i \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ i \\ -1 \end{pmatrix} \right\rangle = 1 + i \cdot (-i) + 1 = 3,$$

$$v_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \\ -i \end{pmatrix} - \frac{1}{3} \left\langle \begin{pmatrix} 0 \\ 2 \\ 1 \\ -i \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ i \\ -1 \end{pmatrix} \right\rangle \cdot \begin{pmatrix} 1 \\ 0 \\ i \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 1 \\ -i \end{pmatrix},$$

$$\langle v_2, v_2 \rangle = 4 + 1 + (-i) \cdot i = 6,$$

$$v_3 = \begin{pmatrix} i \\ -2 \\ -2 \\ 0 \end{pmatrix} - \frac{1}{3} \left\langle \begin{pmatrix} i \\ -2 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ i \\ -1 \end{pmatrix} \right\rangle \cdot \begin{pmatrix} 1 \\ 0 \\ i \\ -1 \end{pmatrix}$$

$$\begin{aligned}
&= -\frac{1}{6} \left\langle \begin{pmatrix} i \\ -2 \\ -2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \\ -i \end{pmatrix} \right\rangle \cdot \begin{pmatrix} 0 \\ 2 \\ 1 \\ -i \end{pmatrix} \\
&= \begin{pmatrix} -i \\ -2 \\ -2 \\ 0 \end{pmatrix} - \begin{pmatrix} i \\ 0 \\ -1 \\ -i \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ 1 \\ -i \end{pmatrix} \\
&= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.
\end{aligned}$$

Beispielsweise haben wir beim Skalarprodukt $\langle v_1, v_1 \rangle$ den einen Vektor komplex konjugiert, denn so ist das komplexe Skalarprodukt ja gerade definiert. Da sich

hier $v_3 = 0$ ergibt, folgt $v_3 \in \left\langle \begin{pmatrix} 1 \\ 0 \\ i \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \\ -i \end{pmatrix} \right\rangle$. Also bilden v_1 und v_2 eine

Orthogonalbasis. Normieren liefert die Orthonormalbasis

$$\mathcal{B} = \left(\frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 0 \\ i \\ -1 \end{pmatrix}, \frac{1}{\sqrt{6}} \begin{pmatrix} 0 \\ 2 \\ 1 \\ -i \end{pmatrix} \right).$$

■

Zum Spektralsatz für normale Endomorphismen (Satz 10.17 und 10.18): Der Beweis des Satzes ist recht aufwendig, aber auch lehrreich. Er kann in der Literatur aus dem Literaturverzeichnis nachgelesen werden, beispielsweise in [Bos08]. Wir betrachten ein Beispiel, um den Spektralsatz 10.17 und seine äquivalente Formulierung im Satz 10.18 für Matrizen einzuüben.

Beispiel 102

Gegeben sei die Matrix

$$A := \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Vorweg ein wichtiger Satz: Eine Matrix $A \in \mathcal{M}_{n,n}(\mathbb{C})$ ist genau dann unitär diagonalisierbar, wenn sie normal ist. Wir müssen jetzt eine unitäre Matrix $S \in U_n$ finden, sodass $S^{-1}AS$ Diagonalgestalt hat. Der Spektralsatz sagt uns, dass dies funktioniert. Man prüft nach, dass A wirklich normal ist, denn es gilt:

$$AA^* = A^*A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

Daher ist A unitär diagonalisierbar. Wir diagonalisieren einfach A (Wir haben ja gesagt, dass das Wissen aus der Linearen Algebra 1 vorausgesetzt wird :-P). Das charakteristische Polynom ergibt sich zu

$$P_A = \det(A - \lambda E_2) = \det \begin{pmatrix} 1 - \lambda & -1 \\ 1 & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2 + 1.$$

Die Nullstellen des charakteristischen Polynoms sind $\lambda_1 = 1 + i$ und $\lambda_2 = 1 - i$ und damit die Eigenwerte von A . Die Eigenvektoren berechnen sich über die Eigenraumberechnung zu

$$\begin{aligned} \text{Eig}(A, 1 + i) &= \ker \begin{pmatrix} -i & 1 \\ 1 & -i \end{pmatrix} = \ker \begin{pmatrix} -i & -1 \\ 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} -1 \\ i \end{pmatrix} \right\rangle \\ \text{Eig}(A, 1 - i) &= \ker \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} = \ker \begin{pmatrix} i & -1 \\ 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ i \end{pmatrix} \right\rangle. \end{aligned}$$

Die Matrix S ist daher gegeben durch

$$S = \frac{1}{\sqrt{1 + |i|^2}} \begin{pmatrix} -1 & 1 \\ i & i \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 1 \\ i & i \end{pmatrix}.$$

Sie besteht einfach aus den normierten Eigenvektoren. Außerdem ist sie unitär, wie man leicht nachprüft. Nun muss also $S^{-1}AS$ Diagonalgestalt haben. Dies ist aber ganz einfach, denn dies ist gerade die Diagonalmatrix mit den Diagonaleinträgen der Eigenwerte, die wir eben berechnet haben. Und hierbei sieht man, dass wir, was wir eben gemacht haben, irgendwie schon aus der Linearen Algebra 1 kennen, vergleiche dazu [MK09]. Insgesamt erhalten wir also

$$S^{-1}AS = \begin{pmatrix} 1 + i & 0 \\ 0 & 1 - i \end{pmatrix}.$$

■

Übung: Macht dasselbe wie in Beispiel 102 noch einmal mit der Matrix

$$A := \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

Zum Satz 10.19: Aus Satz 10.17 folgt, dass A diagonalisierbar über \mathbb{C} ist. Genauer: Es gibt eine Orthonormalbasis des unitären Standardraums \mathbb{C}^n , die aus Eigenwerten von A besteht. Der Satz 10.19 impliziert, dass alle Eigenwerte von A reell sind. Es ist nun nicht schwer zu zeigen, dass A sogar über \mathbb{R} diagonalisierbar ist und es eine Orthonormalbasis des euklidischen Standardvektorraums \mathbb{R}^n gibt, die aus Eigenvektoren von A besteht.

11 Gruppen und Ringe II

Übersicht

11.1 Definitionen	219
11.2 Sätze und Beweise	221
11.3 Erklärungen zu den Definitionen	230
11.4 Erklärungen zu den Sätzen und Beweisen	234

In diesem Kapitel (es heißt „Gruppen und Ringe II“ in Anlehnung an unseren ersten Band [MK09], dort gibt es schon ein Kapitel „Gruppen, Ringe, Körper“), wollen wir uns noch einmal genauer mit Gruppen und Ringen, insbesondere mit wichtigen Teilmengen von ihnen beschäftigen. Daraus lassen sich dann weitere wichtige Strukturen gewinnen.

11.1 Definitionen

Definition 11.1 (Konjugation)

Sei G eine Gruppe. Zwei Elemente $g, g_0 \in G$ heißen **konjugiert** in G , wenn es ein Element $h \in G$ gibt, sodass

$$g_0 = h^{-1} \circ g \circ h.$$

Wir schreiben dann $g_0 \sim_G g$.

Anmerkung: Häufig schreibt man dafür auch einfach $g_0 \sim g$ und nennt g und g_0 konjugiert, ohne zu erwähnen in welcher Gruppe G . Dies macht man aber nur, wenn klar ist, welche Gruppe gemeint ist.

Definition 11.2 (Gruppenerzeugendensystem)

Sei G eine Gruppe. Für jede Teilmenge A einer Gruppe G ist der Durchschnitt aller Untergruppen von G , die A enthalten,

$$\langle A \rangle := \bigcap_{A \subset U} U,$$

eine Untergruppe von G . Es ist $A \subset \langle A \rangle$ und $\langle A \rangle \subset U$ für jede Untergruppe U von G , die A enthält. Damit ist $\langle A \rangle$ die kleinste Untergruppe von G , die A enthält. Man nennt $\langle A \rangle$ die von A **erzeugte Untergruppe** und A ein **Erzeugendensystem** von U .

Definition 11.3 (Ordnung)

Es seien G eine Gruppe und $a \in G$. Gibt es ein $n \in \mathbb{N}$ mit $a^n = e$, so heißt das kleinste solche n die **Ordnung** von a , geschrieben $\text{ord } a$. Existiert kein solches n , so schreibt man oft formal $\text{ord } a = \infty$.

Definition 11.4 (zyklische Gruppe)

Eine Gruppe G heißt **zyklisch**, wenn sie von einem Element erzeugt werden kann, also wenn es ein $a \in G$ gibt mit

$$G = \langle a \rangle.$$

Definition 11.5 (Ideal)

Eine Teilmenge I eines Ringes R heißt ein **Ideal**, wenn gilt:

- i) $0 \in I$,
- ii) für alle $a, b \in I$ ist $a + b \in I$,
- iii) für alle $a \in I$ und $x \in R$ ist $ax \in I$.

Ist I ein Ideal von R , so schreiben wir $I \triangleleft R$.

Definition 11.6 (Ringerzeugendensystem)

Es sei M eine beliebige Teilmenge eines Ringes R . Wir setzen

$$\langle M \rangle_R := \bigcap_{\substack{I \triangleleft R \\ M \subset I}} I,$$

das heißt, $\langle M \rangle_R$ ist der Durchschnitt aller Ideale von R , die M enthalten. Ist $M = \{a_1, \dots, a_n\}$ eine endliche Menge, so schreibt man statt $\langle M \rangle_R$ auch $\langle a_1, \dots, a_n \rangle_R$. Man nennt $\langle M \rangle_R$ dann **Ringerzeugendensystem**.

11.2 Sätze und Beweise

Satz 11.1 (zyklische Gruppen sind abelsch)

Sei (G, \cdot) eine Gruppe. Sind die Elemente eines Erzeugendensystems von G paarweise vertauschbar, so ist G abelsch. Insbesondere ist jede zyklische Gruppe abelsch.

Beweis: Sei a ein zyklischer Erzeuger der Gruppe. Dann existiert für jedes $b \in G$ ein $k \in \mathbb{Z}$ mit $b = a^k$. Sei weiter $c \in G$ mit $c = a^l$ und $l \in \mathbb{Z}$. Dann gilt:

$$b \cdot c = a^k \cdot a^l = a^{k+l} = a^{l+k} = c \cdot b.$$

q.e.d.

Satz 11.2

Eine Gruppe ist genau dann zyklisch, wenn es einen surjektiven Gruppenhomomorphismus $\phi : \mathbb{Z} \rightarrow G$ gibt.

Satz 11.3

Jede Untergruppe einer zyklischen Gruppe ist selbst wieder zyklisch.

Beweis: Seien (G, \cdot) zyklisch mit Erzeuger a , das heißt $G = \langle a \rangle$, und U eine Untergruppe von G . Sei weiter $b = a^k$ dasjenige Element aus U mit dem kleinsten positiven Exponenten k . U enthält wegen der Abgeschlossenheit alle Potenzen von b , aber keine anderen Potenzen von a . Dies sieht man mittels der Division mit Rest.

q.e.d.

Satz 11.4

Eine endliche zyklische Gruppe (G, \cdot) der Ordnung m besitzt zu jedem Teiler d von m genau eine Untergruppe der Ordnung d .

Beweis: Sei $d \cdot k = m$, also d ein Teiler von m . Dann erzeugt das Element a^k nach Satz 11.3 eine zyklische Untergruppe der Ordnung d (da $(a^k)^d = a^m = e$) und es gilt:

$$d \cdot k = m = \text{kgV}(k, m).$$

Angenommen, es existiert eine weitere Untergruppe U' der Ordnung d . Diese enthalte a^j als Potenz mit kleinstem positiven Exponenten j . Dann besteht U' nach Satz 11.3 genau aus den d verschiedenen Potenzen $a^{j'}$ mit Exponent $j, 2j, 3j, \dots, dj$, und es ist

$$dj = m = \text{kgV}(j, m).$$

Daraus folgt nun $k = j$ und damit die Eindeutigkeit der Untergruppe von Ordnung d .

q.e.d.

Satz 11.5 (Untergruppen von \mathbb{Z})

Die Untergruppen von $(\mathbb{Z}, +)$ sind genau die Mengen

$$\langle n \rangle = n\mathbb{Z}$$

mit $n \in \mathbb{N}$.

Beweis: Wir wissen bereits, dass $n\mathbb{Z} \subset \mathbb{Z}$ eine Untergruppe ist. Wir müssen also nur noch zeigen, dass es zu einer beliebigen Untergruppe $U \subset \mathbb{Z}$ ein $n \in \mathbb{N}$ gibt mit $U = n\mathbb{Z}$.

U muss die Zahl 0 enthalten. Ist $U = \{0\}$, so ist natürlich $U = 0\mathbb{Z}$, und wir sind fertig. Andernfalls gibt es ein Element $a \in U$ mit $a \neq 0$. Da mit a auch $-a$ in U liegen muss, gibt es dann also sogar eine positive Zahl in U . Es sei n die kleinste positive Zahl in U . Wir behaupten, dass dann $U = n\mathbb{Z}$ gilt und zeigen diese Gleichheit, indem wir die beiden Inklusionen separat beweisen.

„ \supset “: Nach Wahl von n ist U eine Untergruppe von \mathbb{Z} , die das Element n enthält. Also muss U auch die von n erzeugte Untergruppe $n\mathbb{Z}$ enthalten.

„ \subset “: Es sei $a \in U$ beliebig. Indem wir die ganze Zahl a mit Rest durch n dividieren, können wir a schreiben als

$$a = kn + r,$$

wobei $k \in \mathbb{Z}$ und $r \in \{0, \dots, n-1\}$ gilt. Wir schreiben dies um als $r = a - kn$. Nun ist $a \in U$ nach Wahl von a und außerdem auch $-kn \in n\mathbb{Z} \subset U$. Wegen der Abgeschlossenheit von U liegt damit auch die Summe $r = a - kn$ dieser beiden Zahlen in U . Aber r war als Rest der obigen Division kleiner als n , und n war schon als die kleinste positive Zahl in U gewählt. Dies ist nur dann möglich, wenn $r = 0$ gilt. Setzen wir dies nun aber oben ein, so sehen wir, dass dann $a = kn + 0 \in n\mathbb{Z}$ folgt. Dies zeigt auch diese Inklusion.

q.e.d.

Satz 11.6 (Nebenklassen)

Es seien G eine Gruppe und U eine Untergruppe.

i) Die Relation

$$a \sim b :\Leftrightarrow a^{-1}b \in U$$

für $a, b \in G$ ist eine Äquivalenzrelation auf G .

ii) Für die Äquivalenzklasse a eines Elements $a \in G$ bezüglich dieser Relation gilt:

$$\bar{a} = aU := \{au : u \in U\} \quad (11.1)$$

Man nennt diese Klassen die **Linksnebenklassen** von U , weil man das Element $a \in G$ links neben alle Elemente von U schreibt. Die Menge aller Äquivalenzklassen dieser Relation, also die Menge aller Linksnebenklassen, wird mit

$$G/U := G/\sim = \{aU : a \in G\}$$

bezeichnet.

Beweis: Wir müssen die drei Eigenschaften von Äquivalenzrelationen zeigen. In der Tat entsprechen diese Eigenschaften in gewissem Sinne genau den drei Eigenschaften des Untergruppenkriteriums.

- Für alle $a \in G$ gilt $a^{-1}a = e \in U$ und damit $a \sim a$.
- Sind $a, b \in G$ mit $a \sim b$, also $a^{-1}b \in U$, so ist auch $b^{-1}a = (a^{-1}b)^{-1} \in U$ und damit $b \sim a$.
- Sind $a, b, c \in G$ mit $a \sim b$ und $b \sim c$, das heißt $a^{-1}b, b^{-1}c \in U$, so ist auch $a^{-1}c = (a^{-1}b)(b^{-1}c) \in U$ und damit $a \sim c$.
- Für $a \in G$ gilt:

$$\begin{aligned}\bar{a} &= \{b \in G : b \sim a\} = \{b \in G : a^{-1}b = u, u \in U\} \\ &= \{b \in G : b = au, u \in U\} = aU.\end{aligned}$$

q.e.d.

Satz 11.7

Es seien G eine endliche Gruppe und $U \subset G$ eine Untergruppe. Dann hat jede Links- und jede Rechtsnebenklasse von U genauso viele Elemente wie U .

Beweis: Für $a \in G$ betrachten wir die Abbildung

$$f : U \rightarrow aU, f(x) = ax.$$

Nach Definition von aU ist f surjektiv. Die Abbildung f ist aber auch injektiv, denn aus $f(x) = f(y)$, also $ax = ay$, folgt natürlich sofort $x = y$. Also ist f bijektiv, und damit müssen die Startmenge U und die Zielmenge aU gleich viele Elemente haben. Die Aussage für Ua ergibt sich analog. q.e.d.

Satz 11.8 (Satz von Lagrange)

Es seien G eine endliche Gruppe und $U \subset G$ eine Untergruppe. Dann gilt:

$$|G| = |U||G/U|. \quad (11.2)$$

Insbesondere ist die Ordnung jeder Untergruppe von G also ein Teiler der Ordnung von G .

Beweis: G ist die disjunkte Vereinigung aller Linksnebenklassen. Die Behauptung des Satzes folgt nun sofort daraus, dass nach Satz 11.7 jede Linksnebenklasse $|U|$ Elemente hat und es insgesamt G/U solcher Linksnebenklassen gibt. q.e.d.

Satz 11.9

Es seien G eine Gruppe und $a \in G$ mit $\text{ord } a =: n < \infty$. Dann ist $\langle a \rangle = \{a^0, \dots, a^{n-1}\}$ und $|\langle a \rangle| = n = \text{ord } a$.

Beweis: Zunächst ist $\langle a \rangle = \{a^k : k \in \mathbb{Z}\}$. Mit Division durch n mit Rest lässt sich aber jedes solche $k \in \mathbb{Z}$ schreiben als $k = pn + r$ mit $p \in \mathbb{Z}$ und $r \in \{0, \dots, n-1\}$. Wegen $a^n = e$ folgt daraus

$$a^k = (a^n)^p a^r = e^p a^r = a^r.$$

Alle a^k mit $k \in \mathbb{Z}$ lassen sich also bereits als ein a^r mit $r \in \{0, \dots, n-1\}$ schreiben. Also ist

$$\langle a \rangle = \{a^0, \dots, a^{n-1}\}.$$

Weiterhin sind diese n Elemente alle verschieden, denn wäre $a^i = a^j$ für gewisse $0 \leq i < j \leq n-1$, so hätten wir $a^{j-i} = e$, was wegen $0 < j-i < n$ ein Widerspruch dazu ist, dass n die kleinste positive Zahl ist mit $a^n = e$. Also sind die Elemente a^0, \dots, a^{n-1} alle verschieden, und es folgt $|\langle a \rangle| = n$. q.e.d.

Satz 11.10 (kleiner Satz von Fermat)

Es seien G eine endliche Gruppe und $a \in G$. Dann gilt:

i) $\text{ord } a$ ist ein Teiler von $|G|$.

ii) $a^{|G|} = e$.

Beweis: Nach Satz 11.9 hat die von a erzeugte Untergruppe $\langle a \rangle$ die Ordnung $\text{ord } a$. Da diese Ordnung nach dem Satz von Lagrange ein Teiler von $|G|$ sein muss, ergibt sich sofort der erste Teil. Weiterhin ist, wiederum nach dem Satz von Lagrange

$$a^{|G|} = a^{|\langle a \rangle| \cdot |G/\langle a \rangle|} = (a^{\text{ord } a})^{|G/\langle a \rangle|} = e,$$

und damit folgt auch der zweite Teil.

q.e.d.

Satz 11.11 (Faktorstrukturen und Normalteiler)

Es seien G eine Gruppe und $U \subset G$ eine Untergruppe. Dann sind die folgenden Aussagen äquivalent:

i) Für alle $a, a', b, b' \in G$ mit $\bar{a} = \overline{a'}$ und $\bar{b} = \overline{b'}$ gilt auch $\overline{a \circ b} = \overline{a' \circ b'}$, das heißt, die Vorschrift $\bar{a} \circ \bar{b} := \overline{a \circ b}$ bestimmt eine wohldefinierte Verknüpfung auf G/U .

ii) Für alle $a \in G$ und $u \in U$ ist $a \circ u \circ a^{-1} \in U$.

iii) Für alle $a \in G$ ist $aU = Ua$, das heißt, die Links- und Rechtsnebenklassen von U stimmen überein.

Eine Untergruppe, die eine dieser Eigenschaften erfüllt, wird **Normalteiler** genannt, die übliche Schreibweise hierfür ist $U \triangleleft G$.

Beweis:

„i) \Rightarrow ii)“ Es seien $a \in G$ und $u \in U$. Dann gilt $\bar{a} = \overline{a \circ u}$, denn $a^{-1} \circ a \circ u = u \in U$. Mit $a' = a \circ u$ und $b = b' = a^{-1}$ folgt dann

$$\bar{e} = \overline{a \circ a^{-1}} = \overline{a \circ u \bar{a}^{-1}}$$

und damit $a \circ u \circ a^{-1} \in U$.

„ii) \Rightarrow iii)“ „ \subset “ Ist $b \in aU$, also $b = a \circ u$ für ein $u \in U$, so können wir dies auch als $b = a \circ u \circ a^{-1} \circ a$ schreiben. Dann ist aber $a \circ u \circ a^{-1} \in U$ und damit $b \in Ua$.

„ \supset “ Ist $b = u \circ a$ für ein $u \in U$, so schreiben wir $b = a \circ a^{-1} \circ u \circ (a^{-1})^{-1}$. Dann ist $a^{-1} \circ u \circ (a^{-1})^{-1} \in U$ also $b \in aU$.

„iii) \Rightarrow i)“ Seien $a, a', b, b' \in G$ mit $\bar{a} = \overline{a'}$, $\bar{b} = \overline{b'}$. Also ist $a^{-1} \circ a', b^{-1} \circ b' \in U$. Dann folgt zunächst

$$(a \circ b)^{-1} \circ (a' \circ b') = b^{-1} \circ a^{-1} \circ a' \circ b'.$$

Wegen $a^{-1} \circ a' \in U$ liegt $a^{-1} \circ a' \circ b'$ in Ub' und damit auch in $b'U$. Also können wir $a^{-1} \circ a' \circ b' = b' \circ u$ für ein $u \in U$ schreiben. Dies liefert nun

$$(a \circ b)^{-1} \circ (a' \circ b') = b^{-1} \circ b' \circ u.$$

Wegen $b^{-1} \circ b' \in U$ und $u \in U$ ist also $(a \circ b)^{-1} \circ (a' \circ b') \in U$. Damit ist also $\bar{a} \circ \bar{b} = \overline{a' \circ b'}$.

q.e.d.

Satz 11.12 (Faktorgruppe)

Es sei G eine Gruppe und $U \triangleleft G$. Dann gilt:

i) G/U ist mit der Verknüpfung $\bar{a} \circ \bar{b} = \overline{a \circ b}$ eine Gruppe. Das neutrale Element ist \bar{e} , das zu \bar{a} inverse Element ist \bar{a}^{-1} .

ii) Ist G abelsch, so ist auch G/U abelsch.

iii) Die Abbildung $\pi : G \rightarrow G/U, \pi(a) = \bar{a}$ ist ein Epimorphismus mit Kern U .

Die Gruppe G/U wird als eine Faktorgruppe von G bezeichnet. Man liest G/U oft als G modulo U und sagt, dass man G/U aus G erhält, indem man U herausschneidet. Dementsprechend schreibt man für $a, b \in G$ statt $\bar{a} = \bar{b} \in G/U$ auch $a \equiv b \pmod{U}$, gesprochen a kongruent b modulo U . Die Abbildung π heißt Restklassenabbildung.

Beweis:

i) Nachdem wir die Wohldefiniertheit der Verknüpfung auf G/U schon nachgeprüft haben, rechnet man die Gruppenaxiome nun ganz einfach nach:

- Für alle $a, b, c \in G$ gilt:

$$(\bar{a} \circ \bar{b}) \circ \bar{c} = \overline{a \circ b \circ c} = \overline{(a \circ b) \circ c} = \overline{a \circ (b \circ c)} = \bar{a} \circ \overline{b \circ c} = \bar{a} \circ (\bar{b} \circ \bar{c}).$$

- Für alle $a \in G$ gilt $\bar{e} \circ \bar{a} = \overline{e \circ a} = \bar{a}$.
- Für alle $a \in G$ ist $\overline{a^{-1}} \circ \bar{a} = \overline{a^{-1} \circ a} = \bar{e}$.

ii) Genau wie die Assoziativität oben überträgt sich die Kommutativität sofort von G auf G/U , denn

$$\bar{a} \circ \bar{b} = \overline{a \circ b} = \overline{b \circ a} = \bar{b} \circ \bar{a}.$$

iii) Die Abbildung π ist nach Definition ein Morphismus, denn für alle $a, b \in G$ gilt:

$$\pi(a \circ b) = \overline{a \circ b} = \bar{a} \circ \bar{b} = \pi(a) \circ \pi(b).$$

Nach Definition von G/U ist π surjektiv und der Kern ist

$$\ker \pi = \{a \in G : \bar{a} = \bar{e}\} = \bar{e} = U.$$

q.e.d.

Satz 11.13 (Homomorphiesatz für Gruppen)

Es sei $f : G \rightarrow H$ ein Morphismus von Gruppen. Dann ist die Abbildung

$$g : G/\ker f \rightarrow \operatorname{im} f, \bar{a} \mapsto f(a) \quad (11.3)$$

zwischen der Faktorgruppe $G/\ker f$ von G und der Untergruppe $\operatorname{im} f$ von H ein Isomorphismus.

Beweis: Zunächst einmal ist $\ker f$ ein Normalteiler von G , sodass $G/\ker f$ also wirklich eine Gruppe ist. Die im Satz angegebene Abbildung g ist außerdem wohldefiniert, denn für $a, b \in G$ mit $\bar{a} = \bar{b}$, also $a^{-1}b \in \ker f$, gilt:

$$e = f(a^{-1} \circ b) = f(a)^{-1} \circ f(b) \quad (11.4)$$

und damit $f(a) = f(b)$. Weiterhin ist g ein Morphismus, denn für $a, b \in G$ gilt:

$$g(\bar{a} \circ \bar{b}) = g(\overline{a \circ b}) = f(a \circ b) = f(a) \circ f(b) = g(\bar{a}) \circ g(\bar{b}). \quad (11.5)$$

Wir müssen also nur noch zeigen, dass g surjektiv und injektiv ist. Beides folgt im Prinzip unmittelbar aus der Konstruktion von g .

- g ist surjektiv, denn ist b ein Element in $\operatorname{im} f$, so gibt es also ein $a \in G$ mit $f(a) = b$, das heißt mit $g(\bar{a}) = b$.
- g ist injektiv, denn ist $a \in G$ mit $g(\bar{a}) = f(a) = e$, so ist also $a \in \ker f$ und damit $\bar{a} = \bar{e}$. Also ist $\ker g = \{\bar{e}\}$.

q.e.d.

Satz 11.14 (Klassifikation zyklischer Gruppen)

Es sei G eine Gruppe.

- i) *Ist G zyklisch, so ist G isomorph zu \mathbb{Z} oder zu $\mathbb{Z}/n\mathbb{Z}$ für ein $n \in \mathbb{N}$.*
- ii) *Ist G endlich und $p := |G|$ eine Primzahl, dann ist G isomorph zu $\mathbb{Z}/p\mathbb{Z}$.*

Beweis:

- i) Es sei $a \in G$ mit $G = \langle a \rangle$. Wir betrachten die Abbildung

$$f: \mathbb{Z} \rightarrow G, \quad k \mapsto a^k, \quad (11.6)$$

die ein Morphismus ist. Dann ist

$$\operatorname{im} f = \{a^k : k \in \mathbb{Z}\} = \langle a \rangle = G, \quad (11.7)$$

das heißt, f ist surjektiv. Der Kern von f muss als Untergruppe von \mathbb{Z} die Form $n\mathbb{Z}$ für ein $n \in \mathbb{N}_0$ haben. Es ergeben sich also zwei Fälle

- Ist $n = 0$, also $\ker f = \{0\}$, so folgt aus dem Homomorphiesatz $\mathbb{Z}/\{0\} \cong G$, also $\mathbb{Z} \cong G$.
- Ist $n > 0$, so folgt aus dem Homomorphiesatz 11.13 $\mathbb{Z}/n\mathbb{Z} \cong G$.

- ii) Es sei $a \in G$ ein beliebiges Element mit $a \neq e$. Nach dem Satz von Lagrange (Satz 11.8) muss die Ordnung der Untergruppe $\langle a \rangle$ von G ein Teiler der Gruppenordnung p sein. Da p eine Primzahl ist, kommt hier also nur $|\langle a \rangle| = 1$ oder $|\langle a \rangle| = p$ in Frage. Weil aber bereits die beiden Elemente e und a in $\langle a \rangle$ liegen, ist $|\langle a \rangle| = p$, das heißt, es ist bereits $\langle a \rangle = G$. Also ist G zyklisch. Die Behauptung folgt damit aus dem ersten Teil.

q.e.d.

Satz 11.15

Für jede endliche Teilmenge M eines Ringes R ist $\langle M \rangle_R$ ein Ideal, das M enthält, und es gilt:

$$\langle M \rangle_R = \{a_1x_1 + \cdots + a_nx_n : n \in \mathbb{N}, a_1, \dots, a_n \in M, x_1, \dots, x_n \in R\}. \quad (11.8)$$

Man sagt, dass $\langle M \rangle_R$ aus den endlichen Linearkombinationen von Elementen aus M mit Koeffizienten in R besteht und nennt $\langle M \rangle_R$ das von M erzeugte Ideal.

Beweis: Die rechte Seite

$$J = \{a_1x_1 + \cdots + a_nx_n : n \in \mathbb{N}, a_1, \dots, a_n \in M, x_1, \dots, x_n \in R\}$$

der behaupteten Gleichung ist offensichtlich ein Ideal von R , die ersten beiden Eigenschaften sind klar nach Konstruktion, die dritte ergibt sich daraus, dass wir für alle $a_1x_1 + \cdots + a_nx_n \in J$ und $x \in R$ das Produkt dieser beiden Elemente als $a_1(xx_1) + \cdots + a_n(xx_n)$ schreiben können. Natürlich enthält J auch die Menge M , denn wir können ja jedes $a \in M$ als $a \cdot 1 \in J$ schreiben. Wir behaupten nun, dass $\langle M \rangle_R = J$ ist.

„ \subset “ Wir haben gerade gesehen, dass J ein Ideal von R ist, das M enthält. Also ist J eines der Ideale I , über das der Durchschnitt gebildet wird. Damit folgt sofort $\langle M \rangle_R \subset J$.

„ \supset “ Ist I ein beliebiges Ideal von R , das M enthält, so muss I natürlich auch alle in J enthaltenen Linearkombinationen von Elementen aus M enthalten. Also ist in diesem Fall $I \supset J$. Bilden wir nun den Durchschnitt $\langle M \rangle_R$ aller dieser Ideale, muss dieser natürlich immer noch J enthalten.

q.e.d.

Satz 11.16 (Ideale und Faktorringe)

Es sei I ein Ideal in einem Ring R . Dann gilt

i) Auf der Menge R/I sind die beiden Verknüpfungen

$$\bar{x} + \bar{y} := \overline{x + y} \quad \bar{xy} := \overline{xy} \quad (11.9)$$

wohldefiniert.

ii) Mit diesen Verknüpfungen ist R/I ein Ring.

iii) Die Restklassenabbildung $\pi : R \rightarrow R/I, \pi(x) = \bar{x}$ ist ein surjektiver Ringhomomorphismus mit Kern I .

Analog zum Fall von Gruppen wird R/I als ein Faktoring von R bezeichnet. Man liest R/I meistens als R modulo I und sagt, dass man R/I aus R erhält, indem man I herusteilt. Dementsprechend schreibt man für $x, y \in R$ statt $\bar{x} = \bar{y} \in R/I$ oft auch $x \equiv y \pmod{I}$.

Beweis:

- i) Die Wohldefiniertheit der Addition folgt direkt, da I ein Normalteiler von R ist. Für die Wohldefiniertheit der Multiplikation seien $x, x', y, y' \in R$ mit $\bar{x} = \bar{x'}$ und $\bar{x} = \bar{y'}$, also $x' - x =: a \in I$ und $y' - y =: b \in I$. Dann gilt

$$x'y' - xy = (x + a)(y + b) - xy = ay + bx + ab.$$

Da jeder der drei Summanden dieses Ausdrucks mindestens einen Faktor aus I enthält, liegen alle diese Summanden in I . Also liegt auch deren Summe in I , das heißt, es ist $x'y' - xy \in I$ und damit wie behauptet $\overline{xy} = \overline{x'y'}$.

- ii) Die Ringeigenschaften übertragen sich sofort von R auf R/I . Wir zeigen exemplarisch die Distributivität. Für alle $x, y, z \in R$ gilt:

$$(\bar{x} + \bar{y})\bar{z} = \overline{x + yz} = \overline{(x + y)z} = \overline{xz + yz} = \overline{xz} + \overline{yz} = \bar{xz} + \bar{yz}.$$

- iii) Wir wissen bereits, dass π ein surjektiver additiver Gruppenhomomorphismus mit Kern I ist. Wir müssen also nur noch die Verträglichkeit von π mit der Multiplikation nachprüfen. Für alle $x, y \in R$ ist

$$\pi(xy) = \overline{xy} = \overline{x} \overline{y} = \pi(x)\pi(y).$$

q.e.d.

Satz 11.17

Es sei $n \in \mathbb{N}_{\geq 2}$. Der Ring $\mathbb{Z}/n\mathbb{Z}$ ist genau dann ein Körper, wenn n eine Primzahl ist.

Beweis:

„ \Rightarrow “ Es sei $\mathbb{Z}/n\mathbb{Z}$ ein Körper. Angenommen, n wäre keine Primzahl. Dann gäbe es eine Faktorisierung $n = pq$ für gewisse $1 \leq p, q < n$, und es wäre in $\mathbb{Z}/n\mathbb{Z}$

$$\bar{0} = \bar{n} = \overline{pq} = \overline{p} \overline{q},$$

aber dies kann nicht sein, denn p und q müssen Einheiten sein und Nullteiler sind keine Einheiten.

„ \Leftarrow “ Es seien n eine Primzahl und $a \in \{1, \dots, n-1\}$. Die Ordnung der vom Element $\bar{a} \in \mathbb{Z}/n\mathbb{Z} \setminus \{\bar{0}\}$ erzeugten additiven Untergruppe

$$\langle \bar{a} \rangle = \{k\bar{a} : k \in \mathbb{Z}\} = \{\overline{ka} : k \in \mathbb{Z}\}$$

muss dann nach dem Satz von Lagrange (Satz 11.8) als Teiler von n gleich 1 oder n sein. Da aber bereits die beiden Elemente 0 und a in dieser Untergruppe liegen, ist $|\langle a \rangle| = 1$ ausgeschlossen, das heißt, es ist $|\langle a \rangle| = n$ und damit $\langle a \rangle = \mathbb{Z}/n\mathbb{Z}$ schon der gesamte Ring. Insbesondere ist also $1 \in \langle a \rangle$, das heißt, es gibt ein $k \in \mathbb{Z}$ mit $\overline{ka} = 1$. Also ist a eine Einheit in $\mathbb{Z}/n\mathbb{Z}$. Da $a \neq 0$ beliebig war, ist $\mathbb{Z}/n\mathbb{Z}$ damit ein Körper.

q.e.d.

Satz 11.18 (Homomorphiesatz für Ringe)

Es sei $f : R \rightarrow S$ ein Ringhomomorphismus. Dann ist die Abbildung

$$g : R/\ker f \rightarrow \text{im } f, \quad \bar{x} \mapsto f(x)$$

ein Ringisomorphismus.

Beweis: Da $\ker f$ ein Ideal von R ist, können wir den Faktorring $R/\ker f$ bilden. Wenden wir weiterhin den Homomorphiesatz für Gruppen auf den zugehörigen Gruppenhomomorphismus $f : (R, +) \rightarrow (S, +)$ an, so sehen wir, dass g wohldefiniert, mit der Addition verträglich und bijektiv ist. Außerdem ist g auch mit der Multiplikation verträglich, denn für alle $x, y \in R/\ker f$ ist

$$g(\overline{xy}) = g(\overline{x}\overline{y}) = f(xy) = f(x)f(y) = g(\overline{x})g(\overline{y})$$

Wegen $f(1) = 1$ gilt schließlich auch $g(\bar{1}) = f(1) = 1$, und damit ist g ein Ringisomorphismus.

q.e.d.

11.3 Erklärungen zu den Definitionen

Zur Definition 11.1 der Konjugation: Ist G eine Gruppe, so interessiert uns oftmals nicht ein einzelnes Element $g \in G$, sondern die Konjugationsklassen, das heißt alle Elemente, die zu g konjugiert sind. Dies werden wir vor allem bei der Klassifikation der Symmetrien sehen.

Beispiel 103

- Wir betrachten die symmetrische Gruppe S_3 . Das neutrale Element ist natürlich nur zu sich selbst konjugiert, denn ist $g = e$ das neutrale Element, so gilt:

$$g_0 = h^{-1} \circ e \circ h = h^{-1} \circ h = e.$$

Tab. 11.1: Die Verknüpfungstabelle der Klein'schen Vierergruppe.

\cdot	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

Ist $g \in S_3$ eine Transposition, das heißt vertauscht die Elemente i und j , für $i, j \in \{1, 2, 3\}$, so hat diese genau einen Fixpunkt l , nämlich das Element, das nicht permutiert wird. Außerdem gibt es eine Permutation h mit $h(k) = l$ für $k \in \{1, 2, 3\}$. Dann gilt aber:

$$g_0(k) = h^{-1}(g(h(k))) = h^{-1}(g(l)) = h^{-1}(l) = k,$$

also hat g_0 einen Fixpunkt, kann also nur eine Transposition oder die Identität sein. Da die Identität aber nur zu sich selbst konjugiert ist, ist g_0 eine Transposition. Wählen wir nun h_1, h_2, h_3 so, dass $h_m(m) = l$ für $m \in \{1, 2, 3\}$, so erhalten wir als g_0 die Transposition, die m festhält. Also sind alle Transpositionen konjugiert. Als Übung solltet ihr euch nun klarmachen, dass die verbleibenden zwei Permutationen konjugiert sind, wir also drei Konjugationsklassen haben.

- Ist G eine abelsche Gruppe, so ist

$$h^{-1} \circ g \circ h = g \circ h^{-1} \circ h = g,$$

das heißt, jedes Element ist nur zu sich selbst konjugiert. ■

Ihr solltet euch als Übung klarmachen, dass Konjugation eine Äquivalenzrelation darstellt.

Zur Definition 11.2 des Erzeugendensystems: Um den Begriff einzuüben, betrachten wir ein einfaches Beispiel und zwar die sogenannte Klein'sche Vierergruppe.

Beispiel 104

Durch die folgende Verknüpfungstabelle 11.1 wird eine kommutative Gruppe $V = \{e, a, b, c\}$ gegeben, die sogenannte *Klein'sche Vierergruppe*, wobei e das neutrale Element bezeichnet.

Für jedes $x \in V$ gilt offensichtlich $x \cdot x = e$, also ist jedes Element sein eigenes Inverses, das heißt, $x = x^{-1}$. Prüft einmal nach, dass dies wirklich eine Gruppe bildet. Okay, wir geben zu: Der Nachweis der Assoziativität ist etwas lästig, aber

sollte einmal im Leben gemacht werden :-). Man prüft nun leicht nach, dass die Klein'sche Vierergruppe wie folgt erzeugt wird und zwar durch

$$V = \langle a, b \rangle = \langle b, c \rangle = \langle a, c \rangle.$$

■

Zur Definition 11.3 der Ordnung einer Gruppe: Ist a ein Element der Gruppe G , so sind in $\langle a \rangle$ entweder alle Potenzen a^n verschieden und damit die Ordnung der Gruppe unendlich, oder es existieren $i, j \in \mathbb{Z}$ mit $i < j$ und $a^i = a^j$, das heißt, $a^{j-i} = e$.

Beispiel 105

- Für $G = S_3$ und $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}$ ist $\text{ord } \sigma = 2$, denn $\sigma \neq \text{Id}$ aber $\sigma^2 = \text{Id}$.
- In $G = \mathbb{Z}$ ist $\text{ord } 1 = \infty$, denn $n \cdot 1 \neq 0$ für alle $n \in \mathbb{N}$.

■

Zur Definition 11.4 einer zyklischen Gruppe: Zyklische Gruppen sind nach Definition endlich oder abzählbar unendlich. Betrachten wir gleich ein paar einfache Beispiele:

Beispiel 106

- Die Klein'sche Vierergruppe aus Beispiel 104 ist endlich erzeugt (zwei Elemente reichen aus, auf die aber nicht verzichtet werden kann), aber nicht zyklisch.
- Die Gruppe $(\mathbb{Z}, +)$ ist zyklisch, denn das Element 1 erzeugt die gesamte Gruppe. Durch mehrfache endliche Verknüpfung (hier Addition) der 1 kann man jede ganze Zahl erzeugen. Natürlich ist aber auch -1 ein Erzeuger, es gilt also:

$$(\mathbb{Z}, +) = \langle 1 \rangle = \langle -1 \rangle.$$

- Trivialerweise ist

$$\langle \emptyset \rangle = \{e\} = \langle e \rangle.$$

Und es ist auch $\langle U \rangle = U$ für jede Untergruppe U einer Gruppe mit neutralem Element $e \in G$.

- Die n -ten Einheitswurzeln sind für jedes $n \in \mathbb{N}$ eine zyklische Gruppe der Ordnung n . Sie wird durch die Einheitswurzel $e^{\frac{2\pi i}{n}}$ erzeugt. Beachte, dass es auch hier mehrere Erzeuger gibt. Für $n = 3$ ist sowohl $e^{\frac{2\pi i}{3}}$ als auch $e^{\frac{4\pi i}{3}}$ ein zyklischer Erzeuger.

- Eine Darstellung einer zyklischen Gruppe liefert die Addition modulo einer Zahl, die sogenannte Restklassenarithmetik. In der additiven Gruppe $(\mathbb{Z}/n\mathbb{Z}, +)$ ist die Restklasse der 1 ein Erzeuger. Zum Beispiel ist

$$\mathbb{Z}/4\mathbb{Z} = \{0, 1, 2, 3\}.$$

Dies soll an Beispielen genügen. ■

Zur Definition 11.5 von Idealen:

- Jedes Ideal I eines Ringes R ist eine additive Untergruppe von R , die ersten beiden Eigenschaften des Untergruppenkriteriums sind genau die ersten beiden Bedingungen für Ideale und die dritte Eigenschaft folgt mit $x = -1$. Da $(R, +)$ außerdem nach Definition eines Ringes abelsch ist, ist jedes Ideal von R sogar ein Normalteiler bezüglich der Addition in R .
- Ist I ein Ideal in einem Ring R mit $1 \in I$, so folgt aus der dritten Eigenschaft mit $a = 1$ sofort $x \in I$ für alle $x \in R$, das heißt, es ist dann bereits $I = R$.

Beispiel 107

- Im Ring $R = \mathbb{Z}$ ist $I = n\mathbb{Z}$ für $n \in \mathbb{N}$ ein Ideal. $0 \in I$ ist offensichtlich, für zwei Zahlen $kn, ln \in n\mathbb{Z}$ mit $k, l \in \mathbb{Z}$ ist auch $kn + ln = (k + l)n \in n\mathbb{Z}$ und für $kn \in n\mathbb{Z}$ und $x \in \mathbb{Z}$ mit $k \in \mathbb{Z}$ ist auch $kx = (kx)n \in n\mathbb{Z}$.

Da jedes Ideal eines Ringes auch eine additive Untergruppe sein muss und diese im Ring \mathbb{Z} alle von der Form $n\mathbb{Z}$ für ein $n \in \mathbb{N}$ sind, sind dies auch bereits alle Ideale von \mathbb{Z} . Insbesondere stimmen Untergruppen und Ideale im Ring \mathbb{Z} also überein. Dies ist aber nicht in jedem Ring so, so ist zum Beispiel \mathbb{Z} eine additive Untergruppe von \mathbb{Q} , aber kein Ideal, denn es ist ja $1 \in \mathbb{Z}$, aber $\mathbb{Z} \neq \mathbb{Q}$.

- In einem Ring R sind $\{0\}$ und R offensichtlich stets Ideale von R . Sie werden die trivialen Ideale von R genannt.
- Ist K ein Körper, so sind die trivialen Ideale $\{0\}$ und K bereits die einzigen Ideale von K . Enthält ein Ideal $I \triangleleft K$ nämlich ein beliebiges Element $a \neq 0$, so enthält es dann auch $1 = aa^{-1}$ und ist damit gleich K .
- Ist $f : R \rightarrow S$ ein Ringhomomorphismus, so ist $\ker f$ stets ein Ideal von f . Es ist $f(0) = 0$ und damit $0 \in \ker f$, für $a, b \in \ker f$ ist $f(a + b) = f(a) + f(b) = 0 + 0 = 0$ und damit $a + b \in \ker f$ und für $a \in \ker f$ und $x \in R$ ist außerdem $f(ax) = f(a)f(x) = 0 \cdot f(x) = 0$ und damit $ax \in \ker f$. ■

Zur Definition 11.6 eines Erzeugendensystems von Idealen:**Beispiel 108**

- Besteht die Menge M nur aus einem Element a , so ist offensichtlich $\langle a \rangle_R = \{ax : x \in R\}$ die Menge aller Vielfachen von a . Insbesondere gilt in $R = \mathbb{Z}$ also für $n \in \mathbb{N}$

$$\langle n \rangle_{\mathbb{Z}} = \{nx : x \in \mathbb{Z}\} = n\mathbb{Z} = \langle n \rangle \quad (11.10)$$

- Im Ring $R = \mathbb{Z} \times \mathbb{Z}$ ist das vom Element $(2,2)$ erzeugte Ideal

$$\langle (2,2) \rangle_{\mathbb{Z} \times \mathbb{Z}} = \{(2,2)(m,n) : m,n \in \mathbb{Z}\} = \{(2m,2n) : m,n \in \mathbb{Z}\}, \quad (11.11)$$

während die von diesem Element erzeugte additive Untergruppe gleich

$$\langle (2,2) \rangle = \{n(2,2) : n \in \mathbb{Z}\} = \{(2n,2n) : n \in \mathbb{Z}\} \quad (11.12)$$

ist. ■

Sind M eine Teilmenge eines Ringes R und I ein Ideal mit $M \subset I$, so gilt bereits $\langle M \rangle_R \subset I$, denn I ist ja dann eines der Ideale, über die der Durchschnitt gebildet wird. Diese triviale Bemerkung verwendet man oft, um Teilmengenbeziehungen für Ideale nachzuweisen. Wenn man zeigen möchte, dass das von einer Menge M erzeugte Ideal $\langle M \rangle_R$ in einem anderen Ideal I enthalten ist, so genügt es dafür zu zeigen, dass die Erzeuger M in I liegen. Diese Eigenschaft ist völlig analog zu der für Untergruppen.

11.4 Erklärungen zu den Sätzen und Beweisen

Zum Satz 11.6 über Nebenklassen: Es seien G eine Gruppe und $U \subset G$ eine Untergruppe.

- Für die oben betrachtete Äquivalenzrelation gilt also

$$\bar{a} = \bar{b} \Leftrightarrow a^{-1}b \in U.$$

Wenn wir im Folgenden mit dieser Äquivalenzrelation arbeiten, ist dies das Einzige, was wir dafür benötigen werden. Insbesondere ist also $\bar{b} = \bar{e}$ genau dann, wenn $b \in U$.

- Es war in der Definition etwas willkürlich, dass wir $a \sim b$ durch $a^{-1}b \in U$ und nicht umgekehrt durch $ba^{-1} \in U$ definiert haben. In der Tat könnten wir auch für diese umgekehrte Relation eine zu dem Satz analoge Aussage beweisen, indem wir dort die Reihenfolge aller Verknüpfungen umdrehen. Wir würden dann demzufolge als Äquivalenzklassen also auch nicht die Linksnebenklassen, sondern die sogenannten Rechtsnebenklassen

$$Ua = \{ua : u \in U\}$$

erhalten. Ist G abelsch, so sind Links- und Rechtsnebenklassen natürlich dasselbe. Im nicht abelschen Fall werden sie im Allgemeinen verschieden sein, wie wir im folgenden Beispiel sehen werden, allerdings wird auch hier später der Fall, in dem Links- und Rechtsnebenklassen übereinstimmen, eine besonders große Rolle spielen. Wir vereinbaren im Folgenden, dass wie in der Definition die Notationen \bar{a} beziehungsweise G/U stets für die Linksnebenklasse aU beziehungsweise die Menge dieser Linksnebenklassen stehen. Wollen wir zwischen Links- und Rechtsnebenklassen unterscheiden, müssen wir sie explizit als aU beziehungsweise Ua schreiben.

- Für jede Untergruppe U einer Gruppe G ist natürlich $U = eU = Ue$ stets sowohl eine Links- als auch eine Rechtsnebenklasse. In der Tat ist dies die einzige Nebenklasse, die eine Untergruppe von G ist, denn die anderen Nebenklassen enthalten ja nicht einmal das neutrale Element e .

Beispiel 109

- Sind $G = \mathbb{Z}$, $n \in \mathbb{N}$ und $U = n\mathbb{Z}$, so erhalten wir für $k, l \in \mathbb{Z}$, dass genau dann $k \sim l$ gilt, wenn $l - k \in n\mathbb{Z}$ ist und die Äquivalenzklassen, also die Linksnebenklassen, sind

$$\bar{k} = k + n\mathbb{Z} = \{\dots, k - 2n, k - n, k, k + n, k + 2n, \dots\},$$

also für $k \in \{0, \dots, n-1\}$ alle ganzen Zahlen, die bei Division durch n den Rest k lassen. Demzufolge ist die Menge aller Linksnebenklassen gleich

$$\mathbb{Z}/n\mathbb{Z} = \{\bar{0}, \bar{1}, \dots, \overline{n-1}\},$$

also die Menge aller möglichen Reste bei Division durch n . Beachte, dass wir hier mit \mathbb{Z} statt mit \mathbb{N} begonnen haben, was nötig war, da \mathbb{N} im Gegensatz zu \mathbb{Z} keine Gruppe ist, dass dies aber an den erhaltenen Äquivalenzklassen nichts ändert. Da dieses Beispiel besonders wichtig ist, hat die Menge $\mathbb{Z}/n\mathbb{Z}$ öfter eine besondere Bezeichnung. Man schreibt sie oft als \mathbb{Z}_n . Dies kann jedoch später Verwechslung mit den p -adischen Zahlen, die ihr vielleicht irgendwann kennenlernt, hervorrufen. Man schreibt auch manchmal \mathbb{F}_n , sie ist allerdings meist nur dann üblich, wenn n eine Primzahl p ist, dann schreibt man \mathbb{F}_p , da dies dann ein Körper ist. Wir werden hier allerdings keine dieser Schreibweisen nutzen.

- Wir betrachten die Gruppe $G = A_3$ und darin die Teilmenge

$$U = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right\},$$

die eine Untergruppe ist. Ist nun $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$, so gilt:

$$\begin{aligned} \sigma \circ U &= \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \right\}, \end{aligned}$$

$$\begin{aligned} U \circ \sigma &= \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \right\}. \end{aligned}$$

Es gilt also $\sigma \circ U \neq U \circ \sigma$, das heißt, Links- und Rechtsnebenklassen sind verschieden. Neben $\sigma \circ U$ ist natürlich auch U selbst eine Linksnebenklasse. Die noch fehlende Linksnebenklasse ist

$$\begin{aligned} &\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \circ U \\ &= \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \right\}. \end{aligned}$$

Wir hoffen, ihr habt alles verstanden. ■

Zum Satz 11.11 über Faktorstrukturen: Wir haben zu einer Untergruppe U einer gegebenen Gruppe G die Menge der Linksnebenklassen G/U untersucht und damit bereits einige interessante Resultate erhalten. Eine Menge ist für sich genommen allerdings noch keine besonders interessante Struktur. Wünschenswert wäre es natürlich, wenn wir G/U nicht nur als Menge, sondern ebenfalls wieder als Gruppe auffassen könnten, also wenn wir aus der gegebenen Verknüpfung in G auch eine Verknüpfung in G/U konstruieren könnten. Dafür legt dieser Satz den Grundstein. Bei der Definition von Faktorstrukturen müssen wir uns jedoch Gedanken um die Wohldefiniertheit von Abbildungen machen.

Beispiel 110 (Wohldefiniertheit)

Es sei \sim eine Äquivalenzrelation auf einer Menge M . Will man eine Abbildung $f : M/\sim \rightarrow N$ von der Menge der zugehörigen Äquivalenzklassen in eine weitere Menge N definieren, so ist die Idee hierfür in der Regel, dass man eine Abbildung $g : M \rightarrow N$ wählt und dann

$$f : M/\sim \rightarrow N, f(\bar{a}) := g(a)$$

setzt. Man möchte das Bild einer Äquivalenzklasse unter f also dadurch definieren, dass man einen Repräsentanten dieser Klasse wählt und diesen Repräsentanten dann mit g abbildet. Als einfaches konkretes Beispiel können wir einmal die Abbildung

$$f : \mathbb{F}_{10} \rightarrow \{0,1\}, f(\bar{n}) := \begin{cases} 1, & n \text{ gerade} \\ 0, & n \text{ ungerade} \end{cases}$$

betrachten, das heißt, wir wollen die Elemente $\bar{0}, \bar{2}, \bar{4}, \bar{6}$ und $\bar{8}$ auf 1 und die anderen auf 0 abbilden. Beachte, dass wir in dieser Funktionsvorschrift genau die oben beschriebene Situation haben. Um eine Äquivalenzklasse in \mathbb{F}_{10} abzubilden, wählen wir einen Repräsentanten n dieser Klasse und bilden diesen mit der Funktion

$$g : \mathbb{Z} \rightarrow \{0,1\}, g(n) := \begin{cases} 1, & n \text{ gerade} \\ 0, & n \text{ ungerade} \end{cases}$$

ab. Offensichtlich ist diese Festlegung so nur dann widerspruchsfrei möglich, wenn der Wert dieser Funktion g nicht von der Wahl des Repräsentanten abhängt. Mit anderen Worten muss

$$g(a) = g(b) \text{ für alle } a, b \in M \text{ mit } \bar{a} = \bar{b}$$

gelten, damit die Definition widerspruchsfrei ist. Statt widerspruchsfrei sagen Mathematiker in diesem Fall in der Regel, dass f durch die Vorschrift dann wohldefiniert ist. Die Wohldefiniertheit einer Funktion muss man also immer dann nachprüfen, wenn der Startbereich der Funktion eine Menge von Äquivalenzklassen ist und die Funktionsvorschrift Repräsentanten dieser Klassen benutzt. In unserem konkreten Beispiel sieht das so aus. Sind $m, n \in \mathbb{Z}$ mit $\bar{n} = \bar{m} \in \mathbb{F}_{10}$, so ist ja $n - m = 10k$ für ein $k \in \mathbb{Z}$. Damit sind n und m also entweder beide gerade oder beide ungerade, und es gilt in jedem Fall $g(n) = g(m)$. Die Funktion f ist also wohldefiniert. Im Gegensatz dazu ist die Vorschrift

$$h : \mathbb{F}_{10} \rightarrow \{0,1\}, h(\bar{n}) := \begin{cases} 1, & \text{falls } n \text{ durch 3 teilbar ist} \\ 0, & \text{falls } n \text{ nicht durch 3 teilbar ist} \end{cases} \quad (11.13)$$

nicht wohldefiniert, sie definiert also keine Funktion auf \mathbb{F}_{10} , denn es ist zum Beispiel $\bar{6} = \bar{16}$, aber 6 ist durch 3 teilbar und 16 nicht. Der Funktionswert von h auf dieser Äquivalenzklasse ist durch die obige Vorschrift also nicht widerspruchsfrei festgelegt. ■

Zum Satz 11.12 über Faktorgruppen: Ist U eine Untergruppe von G , so ist es natürlich sehr naheliegend, auf G/U eine Verknüpfung durch

$$\overline{ab} := \overline{ab}$$

definieren zu wollen. Um zwei Äquivalenzklassen in G/U miteinander zu verknüpfen, verknüpfen wir einfach zwei zugehörige Repräsentanten in G und nehmen dann vom Ergebnis wieder die Äquivalenzklasse.

Beispiel 111

Betrachten wir als konkretes Beispiel hierfür wieder die Menge \mathbb{F}_{10} , so würden wir also die Addition gerne von \mathbb{Z} auf \mathbb{F}_{10} übertragen wollen, indem wir zum Beispiel

$$\overline{6} + \overline{8} + \overline{6} + \overline{8} = \overline{14} = \overline{4}$$

rechnen, also genau wie bei einer Addition ohne Übertrag. Nach der Bemerkung müssen wir allerdings noch überprüfen, ob diese neue Verknüpfung auf G/U wirklich wohldefiniert ist. Im Beispiel hätten wir statt der Repräsentanten 6 und 8 ja zum Beispiel auch 26 beziehungsweise 48 wählen können. In der Tat hätten wir dann allerdings ebenfalls wieder dasselbe Endergebnis

$$\overline{6} + \overline{8} = \overline{26 + 48} = \overline{74} = \overline{4} \quad (11.14)$$

erhalten, in diesem Beispiel scheint die Situation also erst einmal in Ordnung zu sein. ■

In der Tat ist die Verknüpfung in diesem Fall wohldefiniert, wie wir noch sehen werden. Leider ist dies jedoch nicht immer der Fall, wie das folgende Beispiel zeigt.

Beispiel 112

Wir betrachten noch einmal die Untergruppe

$$U = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right\}$$

von S_3 mit der Menge der Linksnebenklassen

$$S_3/U = \left\{ \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right\}, \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \right\}, \right. \\ \left. \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} \right\} \right\}.$$

Angenommen, wir könnten auch hier die Nebenklassen dadurch miteinander verknüpfen, dass wir einfach Repräsentanten der beiden Klassen miteinander verknüpfen und vom Ergebnis wieder die Nebenklasse nehmen. Um die ersten beiden Klassen miteinander zu verknüpfen, könnten wir also zum Beispiel jeweils den ersten Repräsentanten wählen und

$$\overline{\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}} \circ \overline{\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}} = \overline{\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}} \circ \overline{\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}} = \overline{\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}}$$

rechnen, das heißt, das Ergebnis wäre wieder die zweite Nebenklasse. Hätten wir für die erste Nebenklasse statt $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ jedoch den anderen Repräsentanten

$\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}$ gewählt, so hätten wir als Ergebnis

$$\overline{\begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}} \circ \overline{\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}} = \overline{\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}} \circ \overline{\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}} = \overline{\begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}}$$

erhalten, also die dritte Nebenklasse. Die Verknüpfung auf der Menge der Nebenklassen ist hier also nicht wohldefiniert. ■

Im Satz ist die erste Eigenschaft eines Normalteilers in der Regel diejenige, die man benötigt; unser Ziel war es ja gerade, die Menge G/U zu einer Gruppe zu machen und somit dort insbesondere erst einmal eine Verknüpfung zu definieren. Um nachzuprüfen, ob eine gegebene Untergruppe ein Normalteiler ist, sind die anderen Eigenschaften in der Regel jedoch besser geeignet. Hier sind ein paar einfache Beispiele.

Beispiel 113

- Ist G abelsch, so ist jede Untergruppe von G ein Normalteiler, denn die dritte Eigenschaft aus dem Satz ist hier natürlich stets erfüllt.
- Die trivialen Untergruppen $\{e\}$ und G sind immer Normalteiler von G , in beiden Fällen ist die zweite Eigenschaft aus dem Satz offensichtlich.
- Die Untergruppe $U = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \right\}$ von S_3 ist kein Normalteiler, in der Tat haben wir bereits nachgeprüft, dass die erste und dritte Eigenschaft verletzt ist.
- Ist $f : G \rightarrow H$ eine Morphismus, so gilt stets $\ker f \triangleleft G$. Sind nämlich $a \in G$ und $u \in \ker f$, so gilt:

$$f(a \circ u \circ a^{-1}) = f(a) \circ f(u) \circ f(a^{-1}) = f(a) \circ f(a^{-1}) = f(e) = e$$

also $a \circ u \circ a^{-1} \in \ker f$, also ist die zweite Bedingung erfüllt.

- Als spezielles Beispiel hiervon ist für $n \in \mathbb{N}$ die alternierende Gruppe $A_n = \ker \text{sign}$ ein Normalteiler von S_n . ■

Beispiel 114

Es sei $n \in \mathbb{N}$. Die Untergruppe $n\mathbb{Z}$ von \mathbb{Z} ist natürlich ein Normalteiler, da \mathbb{Z} abelsch ist. Also ist $(\mathbb{Z}/n\mathbb{Z}, +)$ mit der bekannten Verknüpfung eine abelsche Gruppe. Wir können uns die Verknüpfung dort vorstellen als die gewöhnliche Addition in \mathbb{Z} , wobei wir uns bei der Summe aber immer nur den Rest bei Division durch n merken. Die Gruppen $(\mathbb{Z}/n\mathbb{Z}, +)$ sind sicher die wichtigsten Beispiele von Faktorgruppen. Für $k, l \in \mathbb{Z}$ schreibt man statt $k = l \pmod{n}$, also $k \equiv l \in \mathbb{Z}/n\mathbb{Z}$, oft auch $k = l \pmod{n}$. ■

Wir hatten gesehen, dass jeder Kern eines Morphismus ein Normalteiler ist. Nach dem dritten Teil des Satzes gilt hier auch die Umkehrung, jeder Normalteiler kann als Kern eines Morphismus geschrieben werden, nämlich als Kern der Restklassenabbildung.

Zum Homomorphiesatz für Gruppen (Satz 11.13):

Beispiel 115

- Wir betrachten für $n \in \mathbb{N}_{\geq 2}$ den Morphismus $\text{sign} : S_n \rightarrow \{1, -1\}$. Der Kern dieser Abbildung ist die alternierende Gruppe A_n . Andererseits ist sign natürlich surjektiv, da die Identität das Vorzeichen 1 und jede Transposition das Vorzeichen -1 hat. Also folgt aus dem Homomorphiesatz, dass die Gruppen S_n/A_n und $\{1, -1\}$ isomorph sind. Insbesondere haben diese beiden Gruppen also gleichviele Elemente, und wir erhalten mit dem Satz von Lagrange

$$\frac{|S_n|}{|A_n|} = |S_n/A_n| = 2.$$

Da S_n genau $n!$ Elemente besitzt, gilt also $|A_n| = \frac{n!}{2}$.

- Sind G eine beliebige Gruppe und $f = \text{Id} : G \rightarrow G$ die Identität, so ist natürlich $\ker f = \{e\}$ und $\text{im } f = G$. Nach dem Homomorphiesatz ist also $G/\{e\} \cong G$ mit der Abbildung $\bar{a} \mapsto a$. Dies ist auch anschaulich klar, wenn man aus G nichts herausschneidet, also keine nicht trivialen Identifizierungen von Elementen aus G vornimmt, so ist die resultierende Gruppe immer noch G .
 - Im anderen Extremfall, dem konstanten Morphismus $f : G \rightarrow G, a \mapsto e$, ist umgekehrt $\ker f = G$ und $\text{im } f = \{e\}$. Hier besagt der Homomorphiesatz also $G/G \cong \{e\}$ mit Isomorphismus $a \mapsto e$. Wenn man aus G alles herausschneidet, so bleibt nur noch die triviale Gruppe $\{e\}$ übrig. ■
-

Zum Satz 11.16 über Ideale und Faktorringer: Die naheliegendste Idee zur Konstruktion von Faktorstrukturen für Ringe ist sicher, einen Ring R und darin einen Unterring $S \subset R$ zu betrachten. Beachte, dass $(S, +)$ dann eine Untergruppe von $(R, +)$ ist. Da $(R, +)$ außerdem eine abelsche Gruppe ist, ist $(S, +)$ sogar ein Normalteiler von $(R, +)$. Wir können also in jedem Fall schon einmal die Faktorgruppe $(R/S, +)$ bilden, das heißt, wir haben auf R/S bereits eine wohldefinierte und kommutative Addition. Wir müssen nun also untersuchen, ob sich auch die Multiplikation auf diesen Raum übertragen lässt. Wir müssen also zunächst einmal überprüfen, ob die Vorschrift

$$\overline{ab} := \overline{a}\overline{b}$$

eine wohldefinierte Verknüpfung auf R/S definiert, das heißt ob für alle $a, a', b, b' \in R$ mit $\overline{a} = \overline{a'}$ und $\overline{b} = \overline{b'}$ auch $\overline{ab} = \overline{a'b'}$ gilt. Leider ist dies nicht der Fall, wie das folgende einfache Beispiel zeigt. Es seien $a = a' \in R$ beliebig, $b \in S$ und $b' = 0$. Wegen $b - b' = b \in S$ ist dann also $\overline{b} = \overline{b'}$. Damit müsste auch gelten, dass $\overline{ab} = \overline{a} \cdot \overline{0} = \overline{0}$ ist, also $ab \in S$. Wir brauchen für die Wohldefiniertheit der Multiplikation auf R/S also sicher die Eigenschaft, dass für alle $a \in R$ und $b \in S$ auch $ab \in S$ gilt. Dies ist eine gegenüber der Abgeschlossenheit der Multiplikation eines Unterrings stark verschärfte Bedingung, die Multiplikation eines Elements von S mit einem beliebigen Element von R und nicht nur einem von S muss wieder in S liegen. Für einen Unterring ist das aber praktisch nicht erfüllbar, es muss ja auch $1 \in S$ sein, also können wir $b = 1$ einsetzen und erhalten, dass jedes Element $a \in R$ bereits in S liegen muss, das heißt S müsste der ganze Ring R sein. Dieser Fall ist aber natürlich ziemlich langweilig.

Um nicht triviale Faktorstrukturen für Ringe konstruieren zu können, sehen wir also

- wir müssen für die herauszuteilende Teilmenge S statt der normalen multiplikativen Abgeschlossenheit die obige stärkere Version fordern, damit sich die Multiplikation wohldefiniert auf R/S überträgt und
- die 1 sollte nicht notwendigerweise in S liegen müssen, da wir sonst nur den trivialen Fall $S = R$ erhalten.

In der Tat werden wir sehen, dass dies die einzig notwendigen Abänderungen in der Definition eines Unterrings sind, um sicherzustellen, dass der daraus gebildete Faktorraum wieder zu einem Ring wird, und dies ist gerade, wie wir Ideale eingeführt haben.

Zum Satz 11.16 über Ideale und Faktorringer:

Beispiel 116

Da $n\mathbb{Z}$ ein Ideal in \mathbb{Z} ist, ist $\mathbb{Z}/n\mathbb{Z}$ ein Ring mit der Multiplikation $\overline{kl} = \overline{k}\overline{l}$. In \mathbb{F}_{10} ist also zum Beispiel $\overline{46} = \overline{24} = \overline{4}$, das heißt wir haben in $\mathbb{Z}/n\mathbb{Z}$ auch eine

Multiplikation, die wir uns als die gewöhnliche Multiplikation in \mathbb{Z} vorstellen können, bei der wir schließlich aber nur den Rest bei Division durch n behalten. Diese Ringe $\mathbb{Z}/n\mathbb{Z}$ sind sicher die mit Abstand wichtigsten Beispiele von Faktorringen. Wir wollen sie daher noch etwas genauer studieren und herausfinden, in welchen Fällen diese Ringe sogar Körper sind. Dies wurde in Satz 11.17 getan. ■

12 Symmetriegruppen

Übersicht

12.1 Definitionen	243
12.2 Sätze und Beweise	244
12.3 Erklärungen zu den Definitionen	249
12.4 Erklärungen zu den Sätzen und Beweisen	252

Wir wollen uns hier mit abstrakten Symmetriegruppen beschäftigen, wobei Symmetrien das sein werden, was man sich auch anschaulich darunter vorstellt. Dabei werden wir uns bei den wichtigen Sätzen der Einfachheit halber, und auch, weil wir nur diese Aussagen im nächsten Kapitel brauchen, auf den zweidimensionalen Fall beschränken. Wir wollen zunächst einmal einführen, was eine Symmetrie überhaupt ist und dann sehen, dass es im Wesentlichen gar nicht so viele verschiedene Symmetrien gibt. Wir werden sehen, dass die Symmetrien eine Gruppe bilden und am Ende noch kurz die endlichen Untergruppen dieser Gruppe bestimmen.

12.1 Definitionen

Definition 12.1 (Symmetrieoperation)

Sei E ein euklidischer Vektorraum. Wir identifizieren E mittels einer Basis mit dem \mathbb{R}^n . Eine **Symmetrieoperation** auf E ist dann eine Abbildung

$$f : E \rightarrow E, f(x) = Ax + b$$

mit $A \in O_n(\mathbb{R}), b \in \mathbb{R}^n$ bezüglich der gewählten Basis. Wir bezeichnen mit $\text{Sym}(E)$ die Menge aller Symmetrieoperationen auf E .

Anmerkung: Man kann leicht zeigen, dass die definierte Abbildung die Längen und Winkel invariant lässt. Man kann außerdem (mit mehr Aufwand) auch zeigen, dass jede stetige Abbildung, die die Längen und Winkel invariant lässt, sich

in der angegebenen Form schreiben lässt. Diese Variante ist auch eine mögliche Definition von Symmetrioperationen

Definition 12.2 (kongruent)

Zwei Teilmengen A, B des \mathbb{R}^2 heißen **kongruent** oder deckungsgleich, wenn sie durch eine Symmetrie aufeinander abgebildet werden.

Definition 12.3 (Dieder-Gruppe)

Seien $n \in \mathbb{N}, d = d_{2\pi/n} \in O_2(\mathbb{R})$ die Drehung um den Ursprung mit Winkel $2\pi/n$ und $s \in O_2(\mathbb{R})$ die Spiegelung an der x_1 -Achse. Wir definieren nun zwei endliche Untergruppen von $O_2(\mathbb{R})$.

- i) $C_n := \langle d \rangle = \{1, d, d^2, \dots, d^{n-1}\}$ und
- ii) $D_n := \langle s, d \rangle = \{1, d, \dots, d^{n-1}, s, sd, \dots, sd^{n-1}\}$.

Wir nennen D_n die **Dieder-Gruppe** der Ordnung $2n$.

12.2 Sätze und Beweise

Satz 12.1 (Klassifikation orthogonaler Matrizen)

Sei $A \in O_2(\mathbb{R})$.

- i) Falls $\det(A) = 1$, so gibt es eine eindeutige reelle Zahl $\alpha \in [0, 2\pi)$, sodass

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

- ii) Falls $\det(A) = -1$, so gibt es eine orthogonale Matrix $S \in O_2(\mathbb{R})$ mit $\det(S) = 1$ und

$$S^{-1}AS = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Beweis: Wir schreiben

$$A = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} =: (x|y)$$

mit $x := \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, y := \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$. Dass A orthogonal ist, bedeutet gerade $x \perp y$ und $\|x\| = \|y\| = 1$. Die Orthogonalitätsbedingung lautet ausgeschrieben

$$x_1y_1 + x_2y_2 = 0.$$

Fasst man diese Gleichung als lineares Gleichungssystem in den Unbestimmten y_1 und y_2 auf, so hat der Lösungsraum wegen $x \neq 0$ die Dimension 1. Eine Basis des Lösungsraums ist die Lösung $y_1 := -x_2, y_2 := x_1$. Es gibt also eine eindeutig bestimmte reelle Zahl $\lambda \neq 0$ mit $y_1 = -\lambda x_2, y_2 = \lambda x_1$. Wegen $\|x\| = \|y\| = 1$ folgt

$$1 = y_1^2 + y_2^2 = \lambda^2(x_1^2 + x_2^2) = \lambda^2,$$

also $\lambda = \pm 1$.

i) Zu $\lambda = 1$.

Hier gilt:

$$A = \begin{pmatrix} x_1 & -x_2 \\ x_2 & x_1 \end{pmatrix}$$

und $\det(A) = x_1^2 + x_2^2 = 1$. Dies bedeutet aber, dass (x_1, x_2) auf dem Einheitskreis liegt. Betrachten wir dies als komplexe Zahl, so sehen wir, dass es genau ein $\alpha \in [0, 2\pi)$ gibt mit $x_1 = \cos \alpha, x_2 = \sin \alpha$.

ii) Zu $\lambda = -1$.

Hier ist also

$$A = \begin{pmatrix} x_1 & x_2 \\ x_2 & -x_1 \end{pmatrix}$$

mit $\det A = -1$ und es gilt:

$$P_A = t^2 - (x_1^2 + x_2^2) = t^2 - 1 = (t - 1)(t + 1).$$

Also gibt es eine Basis $\mathcal{B} = (v, w)$ des \mathbb{R}^2 mit $Av = v, Aw = -w$. Wir können hier o.B.d.A. $\|v\| = \|w\| = 1$ annehmen. Aus der Orthogonalität von A folgt

$$\langle v, w \rangle = \langle Av, Aw \rangle = \langle v, -w \rangle = -\langle v, w \rangle,$$

also ist \mathcal{B} eine Orthonormalbasis aus Eigenvektoren. Wir setzen nun $S := (v|w)$. Dann ist S eine orthogonale Matrix und $S^{-1}AS$ ist eine Diagonalmatrix mit Diagonaleinträgen $1, -1$, und es gilt $\det(S) = \pm 1$. Ist $\det(S) = -1$, so folgt die Aussage durch Ersetzen von v durch $-v$.

q.e.d.

Satz 12.2 (Inverse und Komposition von Symmetrien)

Das Inverse einer Symmetrie $f(x) = Ax + b$ ist die Symmetrie $f^{-1}(x) = A^{-1}x - A^{-1}b$. Das Produkt von zwei Symmetrien $f_i(x) = A_i x + b_i, i = 1, 2$, ist die Symmetrie $(f_1 \circ f_2)(x) = (A_1 A_2)x + (A_1 b_2 + b_1)$.

Beweis: Diese Aufgabe überlassen wir euch als Übung.

q.e.d.

Satz 12.3 (Klassifikation der Symmetrien)

Jede Symmetrie von \mathbb{R}^2 ist konjugiert zu genau einer der folgenden Symmetrien.

i) Die Identität $\text{Id}_{\mathbb{R}^2}$ (das neutrale Element von $\text{Sym}(\mathbb{R}^2)$).

ii) Eine Translation um einen Vektor der Länge $a > 0$ in Richtung der x_1 -Achse, das heißt

$$t_a(x) = \begin{pmatrix} x_1 + a \\ x_2 \end{pmatrix}.$$

iii) Eine Drehung um den Ursprung mit einem Winkel $\alpha \in (0, \pi]$, das heißt

$$d_\alpha(x) = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

iv) Die Spiegelung an der x_1 -Achse, also

$$s(x) = \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}.$$

v) Eine Gleitspiegelung an der x_1 -Achse, also die Komposition

$$s_a(x) = t_a \circ s(x) = \begin{pmatrix} x_1 + a \\ -x_2 \end{pmatrix}$$

mit $a > 0$.

Beweis: Ist $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ eine Symmetrie, so ist also zu zeigen, dass es ein $\phi \in \text{Sym}(\mathbb{R}^2)$ gibt, sodass $f' = \phi^{-1} \circ f \circ \phi$ von einer der obigen Gestalten ist und dass dieses f' dann eindeutig ist. Ist $f = \text{Id}_{\mathbb{R}^2}$, so gilt für jedes ϕ automatisch $f' = \text{Id}_{\mathbb{R}^2}$, dieser Fall ist also fertig. Sei nun $f \neq \text{Id}_{\mathbb{R}^2}$ und f gegeben durch $f = Ax + b$. Wir betrachten nun drei Fälle.

1. $A = E_2$.

Es gelte also $f(x) = x + b$ mit $b = (b_1, b_2) \in \mathbb{R}^2$, $b \neq 0$. Wir setzen $a := \|b\|$.

Dann gibt es eine eindeutige Drehung $d_\alpha \in \text{Sym}(\mathbb{R}^2)$ mit $d_\alpha(b) = \begin{pmatrix} a \\ 0 \end{pmatrix}$.

Dann gilt für alle $x \in \mathbb{R}^2$:

$$(d_\alpha \circ f \circ d_\alpha^{-1})(x) = d_\alpha(d_\alpha^{-1}(x) + b) = x + d_\alpha(b) = x + \begin{pmatrix} a \\ 0 \end{pmatrix}.$$

Wir setzen also $\phi := d_\alpha^{-1}$.

2. $A \neq E_2, \det(A) = 1$.

Das bedeutet $A \in SO_2(\mathbb{R})$. Das heißt, es gibt ein eindeutiges $\alpha \in (0, 2\pi)$, sodass

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

gilt. Wegen

$$(\cos \alpha - 1)^2 + \sin^2 \alpha = 2 - 2 \cos \alpha \neq 0$$

ist die Matrix $(A - E_2)$ invertierbar, das heißt, das Gleichungssystem

$$(A - E_2)v = -b$$

hat eine eindeutige Lösung $v \in \mathbb{R}^2$. Mit $\phi_1(x) := x + v$ gilt dann:

$$(\phi_1^{-1} \circ f \circ \phi_1)(x) = \phi_1^{-1}(Ax + Av + b) = Ax + \underbrace{Av - v}_{=-b} + b = Ax.$$

Ist nun $\alpha \leq \pi$, so sind wir mit $\phi := \phi_1$ fertig. Andernfalls sei $\phi := \phi_1 \circ s$. Dann erhalten wir wegen

$$s^{-1} \circ d_\alpha \circ s = d_{-\alpha} = d_{2\pi-\alpha},$$

dass f zu $d_{2\pi-\alpha}$ konjugiert ist.

3. $A \neq E_2, \det A = -1$.

Wegen 12.1 wissen wir, dass ein $S \in O_2(\mathbb{R})$ existiert, sodass

$$A_1 := S^{-1}AS = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

gilt. Setzen wir $\phi_1(x) := Sx$, so folgt

$$\begin{aligned} f_1(x) &:= \phi_1^{-1} \circ f \circ \phi_1(x) = \phi_1^{-1}(f(Sx)) \\ &= \phi_1^{-1}(ASx + b) = S^{-1}ASx + S^{-1}b \\ &= A_1x + S^{-1}b. \end{aligned}$$

Mit $S^{-1}b = \begin{pmatrix} a \\ c \end{pmatrix}$ und $v := (0, \frac{c}{2})$ folgt

$$(A_1 - E_2)v + b_1 = \begin{pmatrix} 0 \\ -\frac{c}{2} \end{pmatrix} - \begin{pmatrix} 0 \\ \frac{c}{2} \end{pmatrix} + \begin{pmatrix} a \\ c \end{pmatrix} = \begin{pmatrix} a \\ 0 \end{pmatrix}.$$

Setzen wir $\phi_2(x) := x + v$, so folgt

$$\begin{aligned} f_2(x) &:= (\phi_2^{-1} \circ f_1 \circ \phi_2)(x) = \phi_2^{-1}(f_1(x + v)) = \phi_2^{-1}(A_1x + A_1v + S^{-1}b) \\ &= A_1x + A_1v - v + S^{-1}b = A_1x + \begin{pmatrix} a \\ 0 \end{pmatrix}. \end{aligned}$$

Ist $a = 0$, so ist also f konjugiert zu einer Spiegelung, im Fall $a > 0$ zu der gewünschten Gleitspiegelung. Ist $a < 0$, erhalten wir das gewünschte Ergebnis mit $\phi_3(x) = -x$ und $f_3 = \phi_3^{-1} \circ f_2 \circ \phi_3$.

q.e.d.

Satz 12.4 (Klassifikation der endlichen Symmetriegruppen der Ebene)

Sei $G \subset O_2(\mathbb{R})$ eine endliche Untergruppe. Dann gelten die folgenden beiden Aussagen.

- i) Ist G in der Untergruppe $SO_2(\mathbb{R})$ aller Drehungen enthalten, so gilt $G = C_n$, wobei $n = |G|$ die Ordnung von G ist. Insbesondere ist G zyklisch.
- ii) Enthält G mindestens eine Spiegelung $s' \in O_2(\mathbb{R}) \setminus SO_2(\mathbb{R})$, so gilt $G = \langle s', C_n \rangle$, wobei $n = |G|/2$. Nach Wahl eines geeigneten Koordinatensystems gilt $G = D_n$, wobei $n = |G|/2$.

Anmerkung: Da wir hier nur die endlichen Symmetriegruppen betrachten, kann man den Translationsanteil zu Null wählen und sich somit auf $O_2(\mathbb{R})$ beschränken.

Beweis:

- i) Wir betrachten den surjektiven Gruppenhomomorphismus

$$\psi : \mathbb{R} \rightarrow SO_2(\mathbb{R}), \quad \alpha \mapsto d_\alpha,$$

der einer reellen Zahl die Drehung um den Ursprung mit entsprechendem Winkel zuordnet. Es ist

$$H := \psi^{-1}(G) = \{\alpha \in \mathbb{R} : d_\alpha \in G\}$$

eine Untergruppe von \mathbb{R} , und weil G eine endliche Untergruppe ist, gilt:

- $\ker(\psi) = 2\pi\mathbb{Z} \subset H$,
- $H \cap (0, 2\pi)$ ist eine endliche Menge.

Sei o.B.d.A. $G \neq \{1\}$, das heißt, $H \cap (0, 2\pi)$ ist nichtleer. Wir setzen

$$\alpha_0 := \min\{\alpha \in (0, 2\pi) : d_\alpha \in G\}.$$

Wir behaupten, dass $H = \alpha_0 \mathbb{Z}$.

Sei also $\alpha \in H$ beliebig. Dann führen wir Division mit Rest durch und erhalten ein $k \in \mathbb{Z}$ und ein $\beta \in [0, \alpha_0)$ mit

$$\alpha = k\alpha_0 + \beta.$$

Weil α_0 und α in H liegen und H eine Untergruppe ist, liegt auch β in H . Dann muss aber, weil α_0 das kleinste positive Element war, $\beta = 0$ gelten, das heißt, $\alpha = k\alpha_0 \in \alpha_0 \mathbb{Z}$, und damit ist α_0 zyklischer Erzeuger von H . Wegen $2\pi \in H$ gibt es also eine natürliche Zahl n mit $\alpha_0 n = 2\pi$, das heißt, $\alpha_0 = \frac{2\pi}{n}$. Die Gruppe G wird also von der Drehung $d_{\frac{2\pi}{n}}$ erzeugt, das heißt, $G = C_n$.

- ii) Wir nehmen nach Wahl eines geeigneten Koordinatensystems an, dass die in G enthaltene Spiegelung s' die Spiegelung an der x_1 -Achse ist, das heißt, $s' = s$. Wir setzen

$$G_0 := G \cap SO_2(\mathbb{R}) \subset G.$$

Dann ist G_0 die Untergruppe der in G enthaltenen Drehungen. Nach dem ersten Teil gilt also $G_0 = C_n$ mit $n := |G_0|$. Sei nun $g \in G \setminus G_0$. Dann ist $sg \in G$ das Produkt von zwei Spiegelungen, also von zwei Matrizen mit Determinante -1 , das heißt eine Drehung, also gilt $sg = h = d^i$ für ein $i \in \{0, \dots, n-1\}$, und es gilt $g = sh = sd^i$. Damit folgt die Aussage.

q.e.d.

12.3 Erklärungen zu den Definitionen

Zur Definition 12.1 der Symmetrieoperation: Eine Symmetrieoperation ist nun einfach das, was man sich auch anschaulich unter einer Symmetrie vorstellt, nämlich eine Abbildung, die Längen und Winkel invariant lässt. Um dies zu veranschaulichen, zeichnen wir also mal Längen und Winkel, am besten anhand eines Musters. Hat man ein bestimmtes Muster, wie zum Beispiel in Abbildung 12.1, so ist eine Symmetrie also eine Abbildung, die dieses Muster von der Form her nicht verändert, sondern nur die Lage ändert, wie etwa in Abbildung 12.2.

Oder auch in Abbildung 12.3 Es ist also genau so, dass Längen und Winkel gleich bleiben. Eine solche Symmetrie ist dann eine Symmetrie des umgebenden Vektorraums, nicht jedoch eine Symmetrie der betrachteten Figur! Von einer solche Symmetrie würden wir auch noch fordern, dass sie die Figur als Teilmenge des Vektorraums invariant lässt, das heißt dass sich die Lage des Objekts bezüglich des Koordinatensystems nicht ändert.

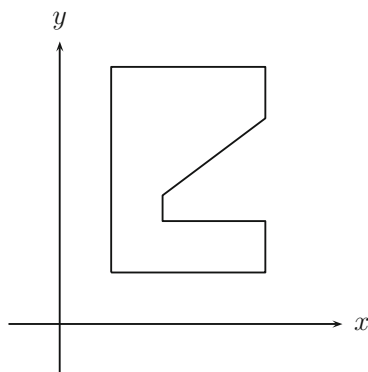


Abb. 12.1: Ein „Muster“.

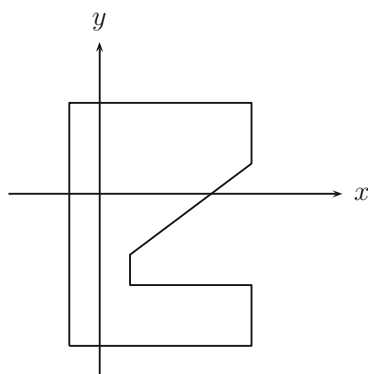


Abb. 12.2: Das „Muster“ verschoben.

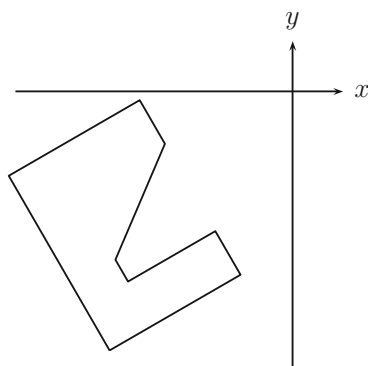


Abb. 12.3: Das „Muster“ verschoben und gedreht.

Zur Definition 12.3 der Dieder-Gruppe: Wie bereits angedeutet, kann man auch die Symmetriegruppe von Objekten betrachten, also die Symmetrien die neben Form auch die Lage nicht verändern, das heißt das ganze Objekt unverändert lassen. Damit kommen wir zu der Hauptinterpretationsweise der Dieder-Gruppen D_n und der Gruppen C_n . Und zwar ist die Gruppe D_n genau die Menge der

Symmetrien, die sowohl Form als auch Lage eines regelmäßigen n -Eckes nicht verändert. Die Gruppen C_n ist die Untergruppe hiervon, die auch noch die „Richtung“ der Eckpunkte nicht verändert. Dazu kurz eine Veranschaulichung für das 5-Eck (siehe Abbildungen 12.4-12.6).

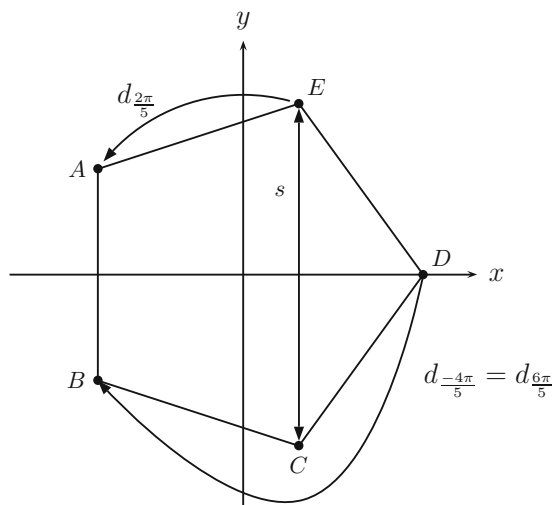


Abb. 12.4: Das regelmäßige 5-Eck und drei seiner Symmetrien.

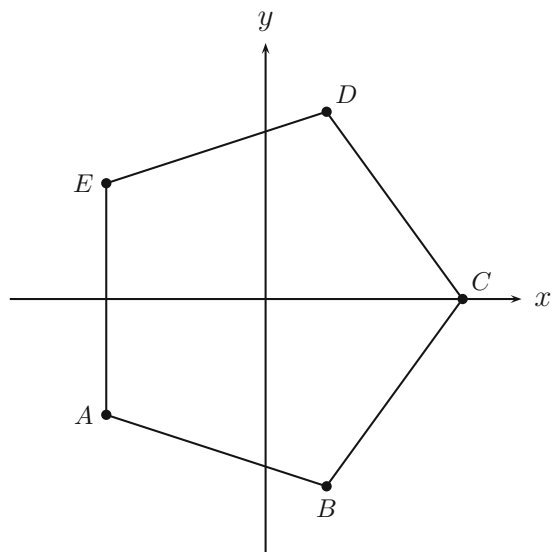


Abb. 12.5: Die Symmetrie $d_{\frac{2\pi}{5}}$, angewendet auf das regelmäßige 5-Eck.

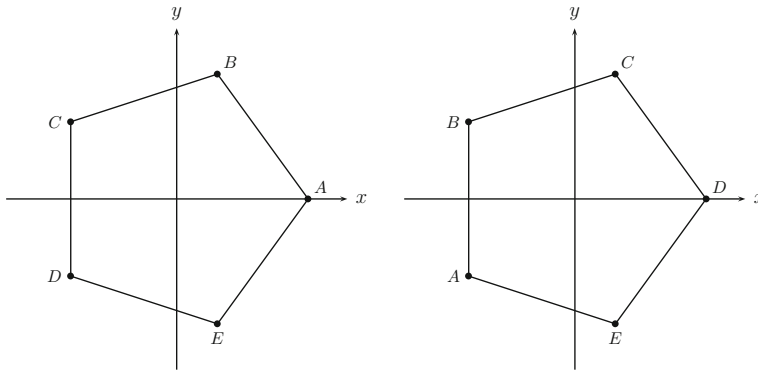


Abb. 12.6: Die Symmetrien $d_{\frac{6\pi}{5}}$ und S , angewendet auf das regelmäßige 5-Eck.

12.4 Erklärungen zu den Sätzen und Beweisen

Zum Satz 12.1 über die Klassifikation orthogonaler Matrizen: Dieser Satz klassifiziert orthogonale Matrizen, abhängig von ihrer Determinante. Sowohl Aussage und Beweis sollten selbsterklärend sein. Wir merken noch kurz an, dass man, indem man den Beweis auf eine gegebene Matrix durchführt, für jede Matrix diese Form auch leicht erhalten kann. Dies wollen wir an einem Beispiel illustrieren.

Beispiel 117

Sei

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Es ist $\det(A) = -1$, wir sind also im zweiten Fall, suchen also zunächst die normierten Eigenvektoren von A . Diese sind gerade

$$v = \begin{pmatrix} \frac{1}{2} \frac{\sqrt{2}}{\sqrt{2-\sqrt{2}}} \\ \frac{1}{2} \sqrt{2-\sqrt{2}} \end{pmatrix}, \quad w = \begin{pmatrix} -\frac{1}{2} \frac{\sqrt{2}}{\sqrt{2+\sqrt{2}}} \\ \frac{1}{2} \sqrt{2+\sqrt{2}} \end{pmatrix}$$

und damit ist

$$S = \begin{pmatrix} \frac{1}{2} \frac{\sqrt{2}}{\sqrt{2-\sqrt{2}}} & -\frac{1}{2} \frac{\sqrt{2}}{\sqrt{2+\sqrt{2}}} \\ \frac{1}{2} \sqrt{2-\sqrt{2}} & \frac{1}{2} \sqrt{2+\sqrt{2}} \end{pmatrix},$$

und es gilt $\det(S) = 1$. Mit

$$S^{-1} = \begin{pmatrix} \frac{1}{2} \sqrt{2+\sqrt{2}} & \frac{1}{2} \frac{\sqrt{2}}{\sqrt{2+\sqrt{2}}} \\ -\frac{1}{2} \sqrt{2-\sqrt{2}} & \frac{1}{2} \frac{\sqrt{2}}{\sqrt{2-\sqrt{2}}} \end{pmatrix}$$

gilt dann $S^{-1}AS = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$

■

Zum Satz 12.2 über Inverse und Komposition von Symmetrien: Hier ist nun der Satz von oben neu formuliert für den Fall, dass die Symmetrie in Matrixform gegeben ist. Der Beweis folgt sofort durch Nachrechnen, was wir euch überlassen ;-)

Beispiel 118

Seien

$$S_1(x) := \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} x + \begin{pmatrix} \frac{1}{2} \\ \frac{3}{2} \end{pmatrix}, \quad S_2(x) := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} x.$$

Dann sind

$$S_2^{-1}(x) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} x, \quad S_1^{-1}(x) = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} x - \begin{pmatrix} \sqrt{2} \\ \frac{1}{\sqrt{2}} \end{pmatrix}$$

und

$$S_2 \circ S_1(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ -1 & -1 \end{pmatrix} x + \begin{pmatrix} \frac{1}{2} \\ -\frac{3}{2} \end{pmatrix}.$$

Dies kann man aber auch leicht graphisch einsehen, siehe dazu Abbildung 12.7: ■

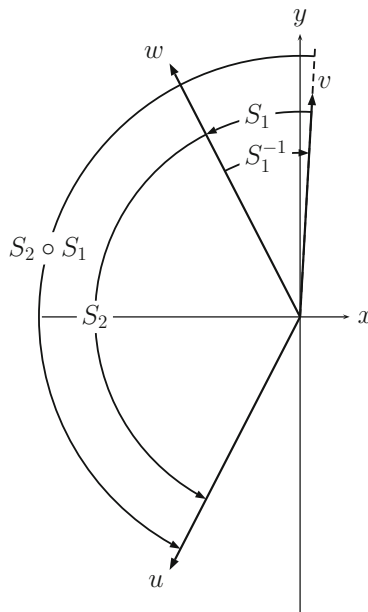


Abb. 12.7: Darstellung zu $S_2 \circ S_1$.

Zum Satz 12.3 über die Klassifizierung von Symmetrien: Dieser Satz klassifiziert nun alle möglichen Symmetrien der Ebene. Wir nennen eine Symmetrie nun eine Translation, Drehung, Spiegelung oder Gleitspiegelung, wenn sie zu der entsprechenden Symmetrie konjugiert ist.

Der Beweis ist konstruktiv, wir gehen ihn an einigen Beispielen durch.

Beispiel 119

Wir betrachten die 8 Symmetrien $S_i = A_i x + b_i : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ mit

$$S_1(x) = x + \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad S_2(x) = x + \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \quad S_3(x) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} x + \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$S_4(x) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad S_5(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} x + \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

$$S_6(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} x, \quad S_7(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} x + \begin{pmatrix} -2 \\ 2 \end{pmatrix},$$

$$S_8(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} x + \begin{pmatrix} 0 \\ -2\sqrt{2} \end{pmatrix}.$$

Wir wollen untersuchen, welche Symmetrie jeweils vorliegt, das heißt die Symmetrie aus der Liste finden, die zu der jeweiligen Symmetrie konjugiert ist. Weiterhin wollen wir wissen, ob einige dieser Symmetrien vielleicht sogar konjugiert zueinander sind.

Die ersten beiden Symmetrien sind natürlich beides Translationen. Wegen

$$\left\| \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\| = \left\| \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\| = 5$$

sind diese beiden sogar konjugiert, und die entsprechende Symmetrie aus der Liste ist

$$x + \begin{pmatrix} 5 \\ 0 \end{pmatrix}.$$

Als Nächstes sehen wir, dass $\det(A_3) = \det(A_4) = \det(A_5) = 1$ und $\det(A_6) = \det(A_7) = \det(A_8) = -1$. Es können also höchstens S_3, S_4 und S_5 beziehungsweise S_6, S_7 und S_8 zueinander konjugiert sein. Wir betrachten zunächst die ersten drei.

Nach dem Beweis müssen wir als Erstes ein α_i finden, sodass $A_i = \begin{pmatrix} \cos \alpha_i & -\sin \alpha_i \\ \sin \alpha_i & \cos \alpha_i \end{pmatrix}$. Es muss also gelten

$$\cos \alpha_3 = 0, \sin \alpha_3 = 1, \quad \cos \alpha_4 = 0, \sin \alpha_4 = -1, \quad \cos \alpha_5 = \frac{1}{2}\sqrt{2} = \sin \alpha_5.$$

Dies ergibt

$$\alpha_3 = \frac{\pi}{2}, \alpha_4 = \frac{3\pi}{2}, \alpha_5 = \frac{\pi}{4}.$$

Nun ist aber $\alpha_4 > \pi$, deswegen müssen wir stattdessen $\alpha_4^* = 2\pi - \alpha_4 = \frac{\pi}{2}$ nehmen. Also sind S_3 und S_4 konjugiert zu

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

also auch zueinander, und S_5 ist konjugiert zu

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Diese fünf Symmetrien waren noch relativ leicht zu handhaben, aber jetzt wird es ein wenig schwieriger, denn wir müssen als erstes die Matrix S finden, für die

$$S^{-1}AS = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

gilt. Genauer brauchen wir die Matrix S^{-1} . Diese hatten wir aber bereits in Beispiel 117 bestimmt, es gilt:

$$S^{-1} = \begin{pmatrix} \frac{1}{2}\sqrt{2+\sqrt{2}} & \frac{1}{2}\frac{\sqrt{2}}{\sqrt{2+\sqrt{2}}} \\ -\frac{1}{2}\sqrt{2-\sqrt{2}} & \frac{1}{2}\frac{\sqrt{2}}{\sqrt{2-\sqrt{2}}} \end{pmatrix}.$$

Wir berechnen nun also $S^{-1}b_i = \begin{pmatrix} a_i \\ c_i \end{pmatrix}$ und erhalten

$$\begin{pmatrix} a_6 \\ c_6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} a_7 \\ c_7 \end{pmatrix} = \begin{pmatrix} \frac{-2}{\sqrt{2+\sqrt{2}}} \\ \frac{2}{\sqrt{2-\sqrt{2}}} \end{pmatrix}, \begin{pmatrix} a_8 \\ c_8 \end{pmatrix} = \begin{pmatrix} \frac{-2}{\sqrt{2+\sqrt{2}}} \\ \frac{-2}{\sqrt{2-\sqrt{2}}} \end{pmatrix}.$$

Damit ist S_6 eine Spiegelung und S_7, S_8 sind Gleitspiegelungen, die zueinander konjugiert sind. ■

Zum Satz 12.4 über die Klassifizierung von endlichen Symmetriegruppen der Ebene:

Dieser Satz sagt nun einfach, dass es bis auf Isomorphie außer den bekannten endlichen Symmetriegruppen C_n und D_n keine weiteren gibt. Wir wollen euch nur auf drei Sachen im Beweis aufmerksam machen, die ihr euch selbst überlegen solltet (bei Fragen schaut ins Forum). Und zwar sollt ihr die beiden Punkte im ersten Beweisteil, das heißt $\ker(\varphi) = 2\pi\mathbb{Z} \subset H$ und die Tatsache, dass $H \cap (0, 2\pi)$ endlich ist, verstehen und euch überlegen, wieso das Minimum in $\alpha_0 = \min\{\alpha \in (0, 2\pi) : d_\alpha \in G\}$ existiert.

13 Symmetrische Bilinearformen und Quadriken

Übersicht

13.1 Definitionen	257
13.2 Sätze und Beweise	259
13.3 Erklärungen zu den Definitionen	266
13.4 Erklärungen zu den Sätzen und Beweisen	267

Quadriken in der Ebene $E = \mathbb{R}^2$ nennt man auch Kegelschnitte. Unser Ziel ist es nun, die Kongruenzklassen solcher Quadriken zu studieren. Quadriken lassen sich auch im \mathbb{R}^n definieren; wir beschränken uns aber auf den Fall $n = 2$, also auf ebene Quadriken.

Dafür studieren wir zuerst symmetrische Bilinearformen, sei also hier immer $\langle \cdot, \cdot \rangle$ eine symmetrische Bilinearform. Wir haben gesehen, dass durch eine positiv definite Bilinearform ein Skalarprodukt definiert werden kann. Hier geht es nun aber um nicht notwendigerweise positiv definite Bilinearformen.

13.1 Definitionen

Definition 13.1 (isotrop, anisotrop)

Sei V ein reeller Vektorraum, so nennen wir einen Vektor $v \in V$ **isotrop**, falls $\langle v, v \rangle = 0$ und andernfalls **anisotrop**.

Definition 13.2 (Signatur)

Sei V ein endlichdimensionaler reeller Vektorraum mit Orthogonalbasis $\mathcal{B} = (v_1, \dots, v_n)$. Wir definieren die **Signatur** der Bilinearform $\langle \cdot, \cdot \rangle$ als das Paar ganzer Zahlen (p, q) , wobei p die Anzahl der Basisvektoren v_i bezeichnet mit $\langle v_i, v_i \rangle > 0$ und q die Anzahl der v_i mit $\langle v_i, v_i \rangle < 0$.

Definition 13.3

Eine symmetrische Bilinearform auf einem reellen Vektorraum V heißt **nichtausgeartet**, wenn

$$V^\perp = \{v \in V : \langle v, w \rangle = 0 \ \forall w \in V\} = \{0\}.$$

Eine symmetrische Matrix $A \in \mathcal{M}_{n,n}(\mathbb{R})$ heißt nichtausgeartet, wenn die zugehörige Bilinearform auf \mathbb{R}^n nichtausgeartet ist.

Definition 13.4 (quadratische Form)

Sei V ein Vektorraum über einem Körper K , in dem $2 \neq 0$ gilt, und sei $s : V \times V \rightarrow K$ eine Bilinearform. Dann ordnen wir dieser die **quadratische Form**

$$q : V \rightarrow K, \quad q(v) := s(v, v)$$

zu.

Anmerkung: die Forderung $2 \neq 0$ brauchen wir, um durch 2 teilen zu können.

Definition 13.5 (selbstadjungiert)

Ein Endomorphismus $\phi : V \rightarrow V$ heißt **selbstadjungiert** bezüglich $\langle \cdot, \cdot \rangle$, wenn für alle $v, w \in V$

$$\langle \phi(v), w \rangle = \langle v, \phi(w) \rangle$$

gilt.

Definition 13.6 (Quadrik)

Eine **Quadrik** ist die (nichtleere) Lösungsmenge $Q \subset \mathbb{R}^n$ einer quadratischen Gleichung, das heißt einer Gleichung der Form

$$q(x) + l(x) + c = 0,$$

wobei $q : \mathbb{R}^n \rightarrow \mathbb{R}$ eine quadratische Form und $l : \mathbb{R}^n \rightarrow \mathbb{R}$ eine lineare Funktion ist.

Definition 13.7 (entartete Quadrik)

Eine ebene Quadrik $Q \subset E = \mathbb{R}^2$ heißt **entartet**, wenn sie die leere Menge ist oder nur aus einem Punkt, oder aus einer Geraden oder aus zwei Geraden besteht oder die gesamte Ebene ist. Andernfalls heißt Q nichtentartet.

13.2 Sätze und Beweise

Satz 13.1

Sei V ein reeller Vektorraum.

- i) Ist $\langle \cdot, \cdot \rangle$ nicht identisch 0, so gibt es einen anisotropen Vektor.
- ii) Ist v anisotrop und $U := \langle v \rangle$, so gilt:

$$V = U \oplus U^\perp.$$

Beweis:

- i) Nach Voraussetzung gibt es $v, w \in V$ mit $\langle v, w \rangle \neq 0$. Ist einer dieser beiden Vektoren anisotrop, so sind wir fertig. Andernfalls setzen wir $u := v + w$. Dann gilt:

$$\langle u, u \rangle = \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle = 2\langle v, w \rangle \neq 0.$$

Also ist u anisotrop.

- ii) Wir müssen $U \cap U^\perp = \{0\}$ und $V = U + U^\perp$ zeigen. Für jedes $u \in U$ gilt $U = \lambda v, \lambda \in \mathbb{R}$. Ist u nun auch in U^\perp , so muss $\langle u, v \rangle = 0$ gelten, das heißt,

$$0 = \langle u, v \rangle = \lambda \langle v, v \rangle$$

und da v anisotrop ist, muss $\lambda = 0$ und damit $u = 0$ gelten. Also folgt $U \cap U^\perp = \{0\}$. Sei nun $w \in V$ beliebig. Es ist zu zeigen, dass wir w schreiben können als $w = \lambda v + w'$ mit $\langle w', v \rangle = 0$. Dann gilt aber:

$$\langle w', v \rangle = \langle w - \lambda v, v \rangle = \langle w, v \rangle - \lambda \langle v, v \rangle,$$

und da v anisotrop ist, ist diese Gleichung nach λ auflösbar und damit ist w' bestimmt.

q.e.d.

Satz 13.2

Sei V ein endlichdimensionaler reeller Vektorraum. Dann gibt es eine orthogonale Basis $\mathcal{B} = (v_1, \dots, v_n)$ mit

$$\langle v_i, v_i \rangle \in \{0, 1, -1\}$$

Wählt man also die Reihenfolge geeignet, so gilt:

$$\langle v_i, v_i \rangle = \begin{cases} 1 & , i = 1, \dots, p, \\ -1 & , i = p+1, \dots, p+q, \\ 0 & , i = p+q+1, \dots, n \end{cases}$$

für geeignete $p, q \geq 0$.

Beweis: Ist $\langle \cdot, \cdot \rangle = 0$, so ist nichts zu zeigen. Sei also $\langle \cdot, \cdot \rangle \neq 0$. Wir wollen die Aussage durch Induktion über die Dimension von V beweisen. Für $n = 0$ ist ebenfalls nichts zu zeigen. Gelte also die Aussage für $n - 1$ und $\dim V = n$. Dann gibt es einen anisotropen Vektor $v \in V$, damit setzen wir

$$v_1 := \frac{1}{\sqrt{|\langle v, v \rangle|}} v,$$

und es gilt $\langle v_1, v_1 \rangle = \pm 1$. Mit $U := \langle v_1 \rangle$ und $W := U^\perp$ erhalten wir eine Zerlegung $V = U \oplus W$, und es gilt $\dim W = n - 1$. Nach Induktionsannahme existiert für W und die Bilinearform $\langle \cdot, \cdot \rangle|_W$ eine orthogonale Basis $\mathcal{B}' = (v_2, \dots, v_n)$ mit $\langle v_i, v_i \rangle \in \{0, 1, -1\}$ für $i = 2, \dots, n$. Wegen $W = U^\perp$ gilt $\langle v_1, v_i \rangle = 0$ für $i = 2, \dots, n$ und damit erfüllt die Basis $\mathcal{B} = (v_1, \dots, v_n)$ die gewünschten Eigenschaften. q.e.d.

Satz 13.3

Seien $\mathcal{B} = (v_1, \dots, v_n)$ eine orthogonale Basis von V und (p, q) die jeweilige Signatur. Wir nehmen o.B.d.A.

$$\langle v_i, v_i \rangle = \begin{cases} 1 & , i = 1, \dots, p, \\ -1 & , i = p + 1, \dots, p + q, \\ 0 & , i = p + q + 1, \dots, n \end{cases}$$

an. Dann bilden die Vektoren v_{p+q+1}, \dots, v_n eine Basis des Untervektorraums

$$V^\perp = \{v \in V : \langle v, w \rangle = 0 \quad \forall w \in V\}.$$

Es gilt also insbesondere $\dim V^\perp = n - p - q$.

Beweis: Zunächst sind diese Vektoren orthogonal zueinander, also auch insbesondere linear unabhängig. Dass sie in V^\perp liegen, folgt ebenfalls sofort. Wir wollen noch zeigen, dass sie ein Erzeugendensystem bilden. Sei dafür $v \in V^\perp$ beliebig:

$$v = \sum_{i=1}^n \lambda_i v_i.$$

Weil \mathcal{B} Orthogonalbasis ist, folgt wegen $v \in V^\perp$

$$0 = \langle v, v_j \rangle = \sum_{i=1}^n \lambda_i \langle v_i, v_j \rangle = \lambda_j \langle v_j, v_j \rangle \quad \forall j.$$

Für $j \leq p + q$ folgt nun wegen $\langle v_j, v_j \rangle \neq 0$, dass $\lambda_j = 0$, und damit wird v bereits von den Vektoren v_{p+q+1}, \dots, v_n erzeugt. q.e.d.

Satz 13.4 (Trägheitssatz von Sylvester)

Die Signatur (p, q) einer symmetrischen Bilinearform ist unabhängig von der Wahl der orthogonalen Basis \mathcal{B} .

Beweis: Seien $\mathcal{B} = (v_1, \dots, v_n)$ und $\mathcal{B}' = (v'_1, \dots, v'_n)$ zwei orthogonale Basen von V und $(p, q,)$ beziehungsweise (p', q') die jeweilige Signatur. Wir nehmen o.B.d.A.

$$\langle v_i, v_i \rangle = \begin{cases} 1 & , i = 1, \dots, p, \\ -1 & , i = p + 1, \dots, p + q, \\ 0 & , i = p + q + 1, \dots, n \end{cases}$$

und

$$\langle v'_i, v'_i \rangle = \begin{cases} 1 & , i = 1, \dots, p', \\ -1 & , i = p' + 1, \dots, p' + q', \\ 0 & , i = p' + q' + 1, \dots, n \end{cases}$$

an. Es ist $p = p'$ und $q = q'$ zu zeigen. Zunächst folgt aus dem vorherigen Lemma sofort $p + q = p' + q'$, da diese Aussage des Lemmas unabhängig von der Basiswahl ist. Nun betrachten wir die $n + p - p'$ Vektoren

$$\underbrace{(v_1, \dots, v_p)}_{p \text{ Vektoren}}, \underbrace{(v'_{p'+1}, \dots, v'_n)}_{n-p' \text{ Vektoren}}.$$

Wir wollen zeigen, dass diese linear unabhängig sind. Sind diese Vektoren linear abhängig, so erhalten wir durch Umstellen der Gleichung der linearen Abhängigkeit

$$\lambda_1 v_1 + \dots + \lambda_p v_p = \mu_1 v'_{p'+1} + \dots + \mu_{n-p'} v'_n$$

mit mindestens einem Skalar ungleich 0. Sei nun

$$v = \lambda_1 v_1 + \dots + \lambda_p v_p = \mu_1 v'_{p'+1} + \dots + \mu_{n-p'} v'_n.$$

Dann erhalten wir aufgrund der Voraussetzungen an \mathcal{B} und wegen

$$v = \lambda_1 v_1 + \dots + \lambda_p v_p$$

die Gleichung

$$\langle v, v \rangle = \lambda_1^2 \langle v_1, v_1 \rangle + \dots + \lambda_p^2 \langle v_p, v_p \rangle = \lambda_1^2 + \dots + \lambda_p^2 \geq 0.$$

Benutzen wir die andere definierende Gleichung und die Basis \mathcal{B}' , so erhalten wir

$$\langle v, v \rangle = \mu_1^2 \langle v'_{p'+1}, v'_{p'+1} \rangle + \dots + \mu_{n-p'}^2 \langle v'_n, v'_n \rangle = -(\mu_1^2 + \dots + \mu_{q'}^2) \leq 0,$$

und daraus folgt $\langle v, v \rangle = 0$, damit $\lambda_i = 0$ für $i = 1, \dots, p$ und $\mu_i = 0$ für $i = 1, \dots, q'$ und daraus wiederum

$$0 = v = \mu_{q'+1} v'_{p'+q'+1} + \dots + \mu_{n-p'} v'_n.$$

Aber $(v'_{p'+q'+1}, \dots, v'_n)$ ist Teil der Basis \mathcal{B}' und damit linear unabhängig, es folgt damit $\mu_i = 0$ für alle i . Damit sind obige $n + p - p'$ Vektoren linear unabhängig, also gilt $p \leq p'$. Da wir in dem Argument aber auch die Rollen vertauschen können, folgt hier sogar die Gleichheit, das heißt, $p = p'$ und wegen $p + q = p' + q'$ dann auch $q = q'$. q.e.d.

Satz 13.5 (selbstadjungiert = symmetrisch)

Seien $\phi : V \rightarrow V$ ein Endomorphismus, \mathcal{B} eine Orthonormalbasis von V und $A \in \mathcal{M}_{n,n}(\mathbb{R})$ die darstellende Matrix von ϕ bezüglich \mathcal{B} . Dann ist ϕ genau dann selbstadjungiert, wenn A symmetrisch ist.

Beweis: Wir identifizieren V durch \mathcal{B} mit \mathbb{R}^n . Dann ist $\langle x, y \rangle = x^T y$ und $\phi(x) = Ax$, das heißt, es gilt:

$$\langle x, Ay \rangle = x^T Ay \quad \text{und} \quad \langle Ax, y \rangle = x^T A^T y = x^T Ay$$

für alle $x, y \in \mathbb{R}^n$. Daraus folgt die Behauptung. q.e.d.

Satz 13.6 (selbstadjungierter Endomorphismus hat Eigenvektor)

Sei $\phi : V \rightarrow V$ ein selbstadjungierter Endomorphismus eines euklidischen Vektorraums der Dimension $n \in \mathbb{N}$. Dann besitzt ϕ mindestens einen Eigenvektor.

Beweis: Wir fassen V als einen normierten Vektorraum auf. Wir betrachten die Menge

$$S := \{v \in V : \|v\| = 1\}.$$

Diese Teilmenge ist beschränkt und abgeschlossen, also kompakt. Wir betrachten nun die quadratische Form

$$q : V \rightarrow \mathbb{R}, \quad v \mapsto q(v) := \langle \phi(v), v \rangle = \langle v, \phi(v) \rangle.$$

Diese Funktion ist stetig, nimmt also auf S ein Maximum an, das heißt, es existiert ein $u \in V$ mit $\|u\| = 1$ und $q(u) \geq q(v)$ für alle $v \in S$. Wir behaupten, dass dieses u ein Eigenvektor ist. Dafür betrachten wir $U := \langle u \rangle$. u ist genau dann Eigenvektor von ϕ , wenn $\phi(u) = \lambda u \in U$. Wir haben eine Zerlegung $V = U \oplus W$ mit $W := U^\perp$, das heißt, $U = W^\perp$. Damit reicht es

$$\langle \phi(u), w \rangle = 0 \forall w \in W$$

zu zeigen, denn dann ist $\phi(u) \in W^\perp = U$. Wir nehmen hier o.B.d.A $\|w\| = 1$ für alle zu untersuchenden w an. Für $t \in \mathbb{R}$ sei

$$v := v(t) := \cos(t)u + \sin(t)w.$$

Nun gilt wegen $\|u\| = \|w\| = 1$ und $\langle u, w \rangle = 0$

$$\|v\|^2 = \cos^2(t)\|u\|^2 + \sin^2(t)\|w\|^2 = 1,$$

das heißt, $v \in S$. Mit

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(t) := q(v(t))$$

folgt nun

$$f(t) = q(\cos(t)u + \sin(t)w) = \cos^2(t)q(u) + \sin^2(t)q(w) + 2\cos(t)\sin(t)\langle \phi(u), w \rangle,$$

wobei wir hier

$$q(\lambda v) = \lambda^2 q(v), \quad q(v + w) = q(v) + q(w) + 2s(v, w)$$

für die durch s definierte quadratische Form q (in unserem Fall $s(u, w) = \langle \phi(u), w \rangle$) ausgenutzt haben.

Offensichtlich ist f also differenzierbar, und es gilt:

$$\begin{aligned} f'(0) &= -2\cos(0)\sin(0)q(u) + 2\sin(0)\cos(0)q(w) \\ &\quad + 2(-\sin^2(0) + \cos^2(0))\langle \phi(u), w \rangle \\ &= 2\langle \phi(u), w \rangle \end{aligned}$$

Nach Wahl von u gilt aber

$$f(t) = q(v) \leq q(u) = f(0),$$

also nimmt f bei 0 ein Maximum an, und es folgt

$$\langle \phi(u), w \rangle = \frac{1}{2}f'(0) = 0$$

und damit ist der Satz bewiesen.

q.e.d.

Satz 13.7

Es seien $\phi : V \rightarrow V$ ein selbstadjungierter Endomorphismus und $V = U \oplus W$ eine Zerlegung von V in eine direkte Summe von zwei Untervektorräumen mit $U = W^\perp$. Dann gilt:

$$\phi(U) \subset U \Leftrightarrow \phi(W) \subset W.$$

Beweis: Es reicht, eine der zwei Implikationen zu beweisen. Angenommen es gilt $\phi(U) \subset U$. Dann gilt:

$$\langle u, \phi(w) \rangle = \underbrace{\langle \phi(u), w \rangle}_{\substack{\in U \\ \in U^\perp}} = 0,$$

also wie gewünscht $\phi(w) \in U^\perp = W$.

q.e.d.

Satz 13.8 (Spektralsatz)

Sei $\phi : V \rightarrow V$ ein selbstadjungierter Endomorphismus auf einem euklidischen Vektorraum V . Dann besitzt V eine Orthonormalbasis $\mathcal{B} = (v_1, \dots, v_n)$, die aus Eigenvektoren von ϕ besteht. Insbesondere ist jeder selbstadjungierte Endomorphismus von V diagonalisierbar.

Beweis: Wir beweisen den Satz durch Induktion über $n = \dim V$. Ist $\dim V = 1$, so ist die Aussage natürlich richtig. Gelte die Aussage also für $n - 1$ und habe V Dimension n . Dann wissen wir, dass ein Eigenvektor $v_1 \in V$ existiert, das heißt, $\phi(v_1) = \lambda_1 v_1$. Durch Normieren von v_1 nehmen wir $\|v_1\| = 1$ an. Seien nun $U := \langle v_1 \rangle$ und $W := U^\perp$. Dann haben wir eine Zerlegung $V = U \oplus W$, und es gilt offensichtlich $\phi(U) \subset U$. Also folgt auch $\phi(W) \subset W$. Wegen $\dim W = n - 1$ wenden wir die Induktionsannahme auf W und $\phi|_W$ an, es existiert also eine Orthonormalbasis $\mathcal{B}' = (v_2, \dots, v_n)$ von W aus Eigenvektoren von $\phi|_W$. Dann ist nach Konstruktion $\mathcal{B} = (v_1, \dots, v_n)$ eine Orthonormalbasis von V , die aus Eigenvektoren von ϕ besteht.

q.e.d.

Satz 13.9 (Klassifikation ebener Quadriken)

Eine nichtentartete ebene Quadrik $Q \subset E = \mathbb{R}^2$ ist kongruent zu genau einer der folgenden Typen:

i) Ellipse

$$a_{1,1}x_1^2 + a_{2,2}x_2^2 - 1 = 0,$$

ii) Hyperbel

$$a_{1,1}x_1^2 - a_{2,2}x_2^2 - 1 = 0,$$

iii) Parabel

$$a_{1,1}x_1^2 - x_2 = 0,$$

wobei jeweils $a_{1,1}, a_{2,2} > 0$ sind.

Beweis: Sei also Q eine nichtentartete Quadrik. Wir betrachten den quadratischen Anteil

$$q(x) = x^T A x.$$

Dann gibt es eine orthogonale Matrix $S \in O_2(\mathbb{R})$ (das heißt eine Symmetrie), sodass

$$A' := S^T A S = \begin{pmatrix} a_{1,1} & 0 \\ 0 & a_{2,2} \end{pmatrix}$$

eine Diagonalmatrix ist. Setzen wir also $x = Sy$, so gilt $q(x) = q(Sy) = a_{1,1}y_1^2 + a_{2,2}y_2^2$, und die gesamte Quadrik ergibt sich zu

$$q(y) = a_{1,1}y_1^2 + a_{2,2}y_2^2 + b'_1y_1 + b'_2y_2 + c = 0$$

mit Konstanten b'_1, b'_2 , die sich aus der Symmetrieoperation berechnen lassen. Wir unterscheiden nun zwei Fälle. Dabei werden wir im Folgenden immer wieder Transformationen durchführen, die die Koeffizienten $a_{i,i}$ verändern. Da aber zum Beispiel $a''_{i,i}$ in einer Gleichung unübersichtlich ist, schreiben wir weiterhin für diese Koeffizienten immer $a_{i,i}$, auch wenn sie sich ändern. Die Koeffizienten die man am Ende erhält, sind also *nicht* dieselben wie am Anfang.

- $a_{1,1}, a_{2,2} \neq 0$: In dem Fall können wir durch die weiteren Koordinatenwechsel

$$y_i = z_i - \frac{b'_i}{2a_{i,i}}, \quad i = 1, 2$$

den linearen Term beseitigen und erhalten dadurch die Form

$$a_{1,1}z_1^2 + a_{2,2}z_2^2 + c' = 0$$

für ein $c' \in \mathbb{R}^2$. Diese Substitution entspricht natürlich gerade der Translation um den Vektor

$$\begin{pmatrix} \frac{b'_1}{2a_{1,1}} \\ \frac{b'_2}{2a_{2,2}} \end{pmatrix}$$

Da wir angenommen hatten, dass die Quadrik nicht entartet ist, muss $c' \neq 0$ gelten, wir können also die Gleichung mit einer Konstanten multiplizieren, um c' auf -1 zu normieren und erhalten damit

$$a_{1,1}z_1^2 + a_{2,2}z_2^2 = 1.$$

Wir machen nun wieder eine Fallunterscheidung nach den Vorzeichen der $a_{i,i}$, wobei wir den Fall, dass beide negativ sind, nicht betrachten müssen, da in diesem Falle die Quadrik entartet wäre.

- $a_{1,1}, a_{2,2} > 0$ In dem Fall erhalten wir die Ellipse, also den ersten Fall in unserem Satz. Gilt sogar $a_{1,1} = a_{2,2}$, so erhalten wir als Spezialfall einen Kreis.
- $a_{1,1} \cdot a_{2,2} < 0$ Hier erhalten wir den zweiten Fall unseres Satzes, eine Hyperbel.

- $a_{1,1} = 0$ oder $a_{2,2} = 0$: Der Fall, dass beide Koeffizienten gleich 0 sind kann nicht auftreten, da dann die Quadrik entartet wäre. Es reicht nun einen der beiden möglichen Fälle zu betrachten. Sei also o.B.d.A. $a_{1,1} \neq 0$ und $a_{2,2} = 0$, das heißt, wir sind bei der Gleichung

$$a_{1,1}y_1^2 + b'_1y_1 + b'_2y_2 + c' = 0.$$

Wäre nun $b'_2 = 0$, so wäre die Quadrik entartet, da sie nicht mehr von y_2 abhängt. Also muss $b'_2 \neq 0$ gelten und wir beseitigen durch die Substitution

$$y_1 = z_1 - \frac{b'_1}{2a_{1,1}}$$

den Koeffizienten b'_1 und durch

$$y_2 = z_2 - \frac{c' + \frac{(b'_1)^2}{4a_{1,1}}}{b'_2}$$

den Term c' . Auch hier können wir anschließend die Gleichung so normieren, dass $b'_2 = -1$ gilt und erhalten insgesamt

$$a_{1,1}z_1^2 - z_2 = 0,$$

also für $a_{1,1} > 0$ den dritten Fall, die Parabel. Ist $a_{1,1} < 0$, so drehen wir durch die Substitution

$$w_1 = z_1, w_2 = -z_2$$

das Vorzeichen um und erhalten auch hier die Parabel.

q.e.d.

13.3 Erklärungen zu den Definitionen

Zur Definition 13.2 der Signatur: Dass diese Definition wohldefiniert ist, sagt uns der Trägheitssatz von Sylvester. Wie bestimmt man aber die Signatur einer Bilinearform? Ist $V = \mathbb{R}^n$, so lässt sich jede Bilinearform durch eine symmetrische Matrix M darstellen. Diese ist diagonalisierbar, es gibt also eine Basis \mathcal{B}' (die wir auch bestimmen können, denn wir können ja noch alle diagonalisieren oder?), für die $M_{\mathcal{B}'}$ eine Diagonalmatrix ist, die aus p positiven Eigenwerten, q negativen Eigenwerten besteht und der Rest ist Eigenwert 0. Teilen wir nun jeden Basisvektor v'_i aus \mathcal{B}' durch die Wurzel des Betrages des zugehörigen Eigenwertes (so wie es auch im Beweis des Satzes getan wird), so bilden die Vektoren v_i , die man erhält, genau die gesuchte Basis \mathcal{B} .

Ist allgemeiner $A \in \mathcal{M}_{n,n}(\mathbb{R})$ eine symmetrische Matrix mit Signatur (p, q) , so ist $p + q = \text{Rang}(A)$.

Zur Definition 13.3 einer nichtausgearteten Bilinearform: Was ist nun eine nichtausgeartete Matrix? Eine Matrix ist nichtausgeartet, wenn für alle $w \in V$ gilt:

$$v^T A w = 0$$

mit einem $v \neq 0$. $v^T A$ ist aber wieder ein Vektor v' . Wir wählen nun für w die n Standardbasisvektoren e_i . Dann muss also $v' e_i = 0$ gelten für alle i , und damit folgt $v' = 0$. Dies bedeutet $v \in \ker A$. Also bedeutet in diesem Fall nichtausgeartet dasselbe wie nicht singulär.

Zur Definition 13.5 der Selbstadjungiertheit: Diese Definition liefert uns nun eine weitere Eigenschaft, die Endomorphismen haben können. Dies erscheint zunächst einmal neu, Satz 13.5 sagt uns aber, dass dies im Falle reeller Vektorräume nichts anderes als Symmetrie bedeutet.

13.4 Erklärungen zu den Sätzen und Beweisen

Zum Satz 13.1: Die erste Aussage des Satzes erscheint zunächst sofort richtig. Wenn es Vektoren v, w gibt, deren Skalarprodukt nicht 0 ist, so muss es einen anisotropen Vektor geben. Diesen konstruiert man nun leicht aus v und w . Den zweiten Teil beweisen wir wieder, indem wir einzeln die beiden Eigenschaften des direkten Produkts zeigen.

Zum Satz 13.2: Dieser Satz legt die Grundlage für die Definition der Signatur. Wir beweisen ihn ähnlich wie das Gram-Schmidtsche Orthogonalisierungsverfahren durch Induktion und normieren der Vektoren. Wir nehmen allerdings immer nur Vektoren dazu, deren Skalarprodukt ± 1 ist. Wieso brauchen wir dann im Satz auch die 0? Schränken wir das Skalarprodukt immer weiter ein, so kann es passieren, dass diese Einschränkung irgendwann 0 wird. Und für $\langle \cdot, \cdot \rangle = 0$ gilt natürlich für den Vektor v , der dazugenommen wird, $\langle v, v \rangle = 0$. Für reelle Vektorräume bedeutet das gerade, dass für eine symmetrische Matrix $A \in \mathcal{M}_{n,n}(\mathbb{R})$ eine invertierbare Matrix $Q \in \text{GL}_n(\mathbb{R})$ existiert, sodass

$$Q^{-1} A Q = \begin{pmatrix} E_p & & \\ & -E_q & \\ & & 0 \end{pmatrix}.$$

Zum Satz 13.3: Diesen Satz benötigen wir zum Beweis des wichtigen Trägheitssatzes von Sylvester. Die wichtige Aussage hier ist, dass die Zahl $n - p - q$ eine Invariante der Bilinearform, das heißt nicht abhängig von der Wahl der Basis ist.

Zum Trägheitssatz von Sylvester (Satz 13.4): Dieser Satz zeigt uns, dass die Signatur wohldefiniert ist. Dabei zeigen wir einzeln zunächst, dass die Anzahl der Nullen gleich ist und dann durch ein Dimensionsargument, dass auch die anderen Anzahlen übereinstimmen.

Zum Satz 13.5 (selbstadjungiert = symmetrisch): Hier wird nun gezeigt, dass im reellen Fall symmetrisch dasselbe wie selbstadjungiert ist. Der Beweis folgt leicht durch Anwenden der Definitionen.

Zum Satz 13.6, dass jeder selbstadjungierte Endomorphismus einen Eigenvektor hat: Dieser Satz stellt den ersten Schritt für den Beweis des wichtigen Spektralsatzes dar. Dafür verwenden wir Methoden der Analysis, wir betrachten eine Funktion auf einer kompakten Menge, bestimmen dort das Maximum und es stellt sich heraus, dass wir hieraus leicht einen Eigenvektor erhalten.

Zum Spektralsatz (Satz 13.8): Dieser Satz sagt uns, auf den Fall $V = \mathbb{R}^n$ angewandt, dass es zu jeder reellen symmetrischen Matrix A eine orthogonale Matrix S gibt, sodass

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

eine Diagonalmatrix ist. Insbesondere ist also jede reelle symmetrische Matrix über \mathbb{R} diagonalisierbar.

Für den Beweis verwenden wir einfach die letzten beiden bewiesenen Sätze und machen dann eine vollständige Induktion.

Zur Klassifikation der Quadriken (Satz 13.9): Nun können wir uns endlich dem widmen, wozu das alles gut ist. Dieser Satz macht zunächst klar, warum Quadriken auch Kegelschnitte heißen (siehe Abbildungen 13.1 und 13.2). Dabei entsteht eine Parabel genau dann, wenn die schneidende Ebene parallel zu einer Mantelfläche ist. Der Beweis beruht darauf, dass jede symmetrische Matrix diagonalisierbar ist. Die weiteren Beweisschritte wollen wir anhand von drei Beispielen nachvollziehen.

Beispiel 120

■ Wir betrachten

$$Q := \left\{ (x_1, x_2) \in \mathbb{R}^2 : -2x_1^2 + 3x_1x_2 + 2x_2^2 - 3x_1 + x_2 - 4 = 0 \right\}.$$

Wir müssen zuerst die Matrix A und den linearen Teil l bestimmen.

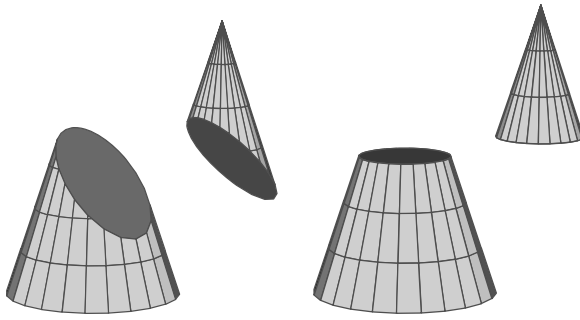


Abb. 13.1: Der Kegelschnitt, der eine Ellipse ergibt, und daneben der Spezialfall des Kreises.

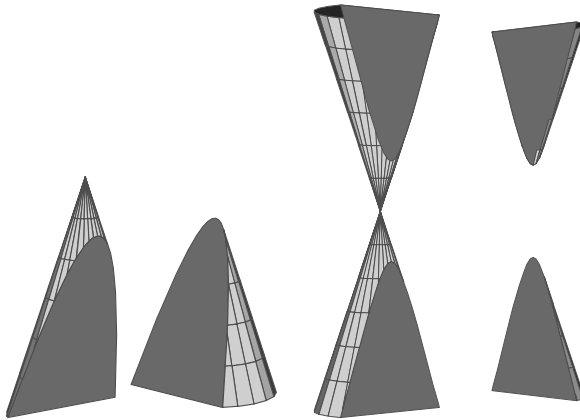


Abb. 13.2: Die beiden Kegelschnitte, die eine Parabel bzw. eine Hyperbel ergeben. Bei der Parabel (linke Hälfte) ist die schneidende Ebene parallel zu einer Mantellinie.

$-2x_1^2 + 3x_1x_2 + 2x_2^2 - 3x_1 + x_2 - 4 = 0$ lässt sich in der Form

$$(x_1, x_2)^T \begin{pmatrix} -2 & 3/2 \\ 3/2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (-3, 1)^T \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} - 4 = 0$$

schreiben. Mit

$$A := \begin{pmatrix} -2 & 3/2 \\ 3/2 & 2 \end{pmatrix}, b = \begin{pmatrix} -3 \\ 1 \end{pmatrix}, c = -4$$

erhalten wir

$$Q := \left\{ x = (x_1, x_2)^T \in \mathbb{R}^2 : x^T A x + b x + c = 0 \right\}.$$

Als Nächstes müssen wir A diagonalisieren.

Das charakteristische Polynom ist gegeben durch

$$P_A = \det \begin{pmatrix} -2 - \lambda & 3/2 \\ 3/2 & 2 - \lambda \end{pmatrix} = (-2 - \lambda)(2 - \lambda) - \frac{9}{4}.$$

Die Nullstellen des charakteristischen Polynoms sind die Eigenwerte von A :

$$(-2 - \lambda)(2 - \lambda) - \frac{9}{4} = 0 \Rightarrow -4 \underbrace{+ 2\lambda - 2\lambda}_{=0} + \lambda^2 - \frac{9}{4} = 0$$

$$\Leftrightarrow \lambda^2 = \frac{25}{4} \Rightarrow \lambda_{1,2} = \pm \frac{5}{2}.$$

Berechnung der Eigenräume bzw. Eigenvektoren ergibt:

$$\begin{aligned} \text{Eig}(A, \frac{5}{2}) &= \ker \begin{pmatrix} -2 - 5/2 & 3/2 \\ 3/2 & 2 - 5/2 \end{pmatrix} = \ker \begin{pmatrix} -9/2 & 3/2 \\ 3/2 & -1/2 \end{pmatrix} \\ &= \ker \begin{pmatrix} -9/2 & 3/2 \\ 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\rangle, \end{aligned}$$

$$\begin{aligned} \text{Eig}(A, -\frac{5}{2}) &= \ker \begin{pmatrix} -2 + 5/2 & 3/2 \\ 3/2 & 2 + 5/2 \end{pmatrix} = \ker \begin{pmatrix} 1/2 & 3/2 \\ 3/2 & 9/2 \end{pmatrix} \\ &= \ker \begin{pmatrix} 1/2 & 3/2 \\ 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right\rangle. \end{aligned}$$

Eine Orthonormalbasis des \mathbb{R}^2 ist also gegeben durch

$$\left(\frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \frac{1}{\sqrt{10}} \begin{pmatrix} 3 \\ -1 \end{pmatrix} \right).$$

Demnach ist

$$S := \frac{1}{\sqrt{10}} \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix}.$$

Nun müssen wir die Substitution $x = Sy$ durchführen. Es ist

$$(Sy)^T A(Sy) + b(Sy) - 4 = y^T \underbrace{S^T A S}_{=D} y + bSy - 4 = y^T D y + bSy - 4 = 0.$$

Hierbei ist

$$S^{-1}AS = S^TAS = \begin{pmatrix} -5/2 & 0 \\ 0 & 5/2 \end{pmatrix} =: D.$$

Wir erhalten

$$\begin{aligned}
 & \begin{pmatrix} y_1 & y_2 \end{pmatrix}^T \begin{pmatrix} -5/2 & 0 \\ 0 & 5/2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}^T \\
 & + \begin{pmatrix} -3 & 1 \end{pmatrix}^T \frac{1}{\sqrt{10}} \begin{pmatrix} 3 & 1 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} - 4 = 0, \\
 & \begin{pmatrix} y_1 & y_2 \end{pmatrix} \begin{pmatrix} (-5/2)y_1 \\ (5/2)y_2 \end{pmatrix} + \begin{pmatrix} -3 & 1 \end{pmatrix}^T \frac{1}{\sqrt{10}} \begin{pmatrix} 3y_1 + y_2 \\ -y_1 + 3y_2 \end{pmatrix} - 4 = 0, \\
 & -\frac{5}{2}y_1^2 + \frac{5}{2}y_2^2 - \sqrt{10}y_1 - 4 = 0.
 \end{aligned}$$

Es sind nun beide Koeffizienten vor den y_i^2 nicht Null, wir befinden uns also im ersten Fall des Beweises. Die Translation ist also

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} y_1 + \frac{\sqrt{10}}{5} \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} \frac{\sqrt{10}}{5} \\ 0 \end{pmatrix}.$$

Dies liefert nun

$$-\frac{5}{2}z_1^2 + \frac{5}{2}z_2^2 - 3 = 0 \Leftrightarrow -\frac{5}{6}z_1^2 + \frac{5}{6}z_2^2 - 1 = 0.$$

Es ist also eine Hyperbel (siehe Abbildung 13.3) gegeben durch

$$Q' := \left\{ z = (z_1, z_2)^T : -\frac{5}{6}z_1^2 + \frac{5}{6}z_2^2 - 1 = 0 \right\}.$$

Wir können nun noch die Symmetrie f angeben, die Q in die Normalform

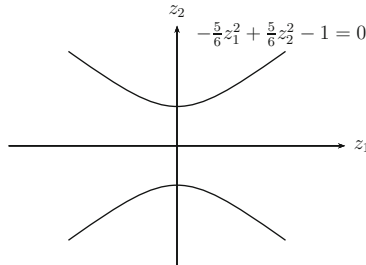


Abb. 13.3: Die Hyperbel $-\frac{5}{6}z_1^2 + \frac{5}{6}z_2^2 - 1 = 0$.

überführt. Wir erhalten

$$x = Sy = S \left(z - \begin{pmatrix} \frac{\sqrt{10}}{5} \\ 0 \end{pmatrix} \right) = Sz - S \begin{pmatrix} \frac{\sqrt{10}}{5} \\ 0 \end{pmatrix} = Sz - \begin{pmatrix} \frac{1}{5} \\ \frac{3}{5} \end{pmatrix}.$$

Für

$$\bar{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2, z \mapsto Sz - \begin{pmatrix} \frac{1}{5} \\ \frac{3}{5} \end{pmatrix}$$

gilt $\bar{f}(Q') = Q$. Daraus folgt: Für

$$\begin{aligned} f = \bar{f}^{-1} : \mathbb{R}^2 &\rightarrow \mathbb{R}^2, x \mapsto S^{-1} \left(x + \begin{pmatrix} \frac{1}{5} \\ \frac{3}{5} \end{pmatrix} \right) = S^T \left(x + \begin{pmatrix} \frac{1}{5} \\ \frac{3}{5} \end{pmatrix} \right) \\ &= S^T x + S^T \begin{pmatrix} \frac{1}{5} \\ \frac{3}{5} \end{pmatrix} = S^T x + \begin{pmatrix} \frac{\sqrt{10}}{5} \\ 0 \end{pmatrix} \end{aligned}$$

gilt $f(Q) = Q'$.

■ Wir betrachten die Quadrik

$$Q(x) = x_1^2 + 2x_1x_2 + x_2^2 + 2x_1 + 4x_2 - 1 = 0.$$

Hier ist mit

$$\begin{aligned} A &= \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, b = \begin{pmatrix} 2 \\ 4 \end{pmatrix}, c = -1, \\ Q(x) &= x^T A x + b^T x + c = 0. \end{aligned}$$

Die Eigenwerte von A sind 0 und 2 und mit

$$S := \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

gilt:

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

Die Substitution $x = Sy$ führt auf

$$2y_1^2 + 3\sqrt{2}y_1 - \sqrt{2}y_2 - 1 = 0$$

wir sind also im zweiten Fall des Beweises. Mit

$$y_1 = z_1 - \frac{3}{4}\sqrt{2}, y_2 = z_2 - \frac{13}{8}\sqrt{2}$$

ist dann

$$2z_1^2 - \sqrt{2}z_2 = 0$$

und Normieren führt auf

$$\sqrt{2}z_1^2 - z_2 = 0$$

also eine Parabel (siehe Abbildung 13.4). Wir wollen nun wieder die Symmetrie von f mit $f(Q) = Q'$ bestimmen. Es ist

$$x = Sy = S \left(z - \begin{pmatrix} \frac{3}{4}\sqrt{2} \\ \frac{13}{8}\sqrt{2} \end{pmatrix} \right) = Sz + \begin{pmatrix} -\frac{19}{8} \\ \frac{7}{8} \end{pmatrix}.$$

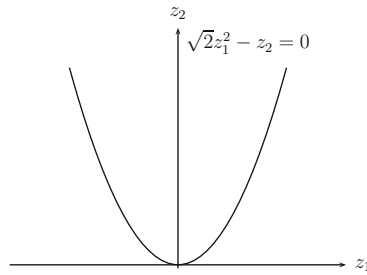


Abb. 13.4: Die Parabel $\sqrt{2}z_1^2 - z_2 = 0$.

Also ist die gesuchte Symmetrie

$$f : x \mapsto S^{-1}\left(x - \begin{pmatrix} -\frac{19}{8} \\ \frac{7}{8} \end{pmatrix}\right) = S^T x + \begin{pmatrix} \frac{3}{4}\sqrt{2} \\ \frac{13}{8}\sqrt{2} \end{pmatrix}$$

■ Sei nun Q die durch die Gleichung

$$3x_1^2 - 2x_1x_2 + 3x_2^2 + 8x_2 - 8x_1 + 6 = 0$$

definierte Quadrik. Die Matrix A ist also

$$A = \begin{pmatrix} 3 & -1 \\ -1 & 3 \end{pmatrix}.$$

Auch hier müssen wir wieder A diagonalisieren, die Durchführung überlassen wir euch. Man erhält dann

$$S^{-1}AS = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} \text{ mit } S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}.$$

Die Substitution $x = Sy$ in die Gleichung für Q und anschließende Division durch 2 führt uns auf die Gleichung

$$y_1^2 + 2y_2^2 - 4\sqrt{2}y_2 + 3 = 0.$$

Durch die Substitution $y_1 = z_1, y_2 = z_2 + \sqrt{2}$ erhalten wir schließlich eine Gleichung in Normalform

$$z_1^2 + 2z_2^2 - 1 = 0.$$

Die Quadrik Q ist also eine Ellipse (siehe Abbildung 13.5). Die Symmetrie, die Q in Normalform überführt, bestimmt man wie in den beiden obigen Fällen, dies solltet ihr zur Übung tun.

■

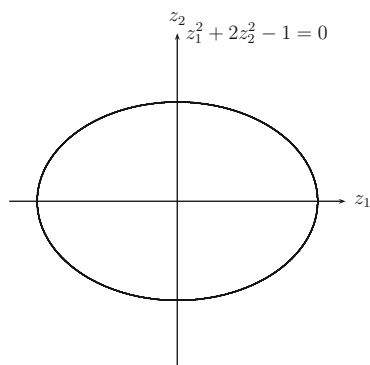


Abb. 13.5: Die Ellipse $z_1^2 + 2z_2^2 - 1 = 0$.

14 Normalformen

Übersicht

14.1 Definitionen	275
14.2 Sätze und Beweise	276
14.3 Erklärungen zu den Definitionen	280
14.4 Erklärungen zu den Sätzen und Beweisen	283

Wir wollen uns nun wieder einmal (wie in unserem ersten Buch) damit befassen, Matrizen in „einfache“ Form zu bringen. Wir beschränken uns dabei nun aber nicht mehr auf diagonalisierbare Endomorphismen. Es seien hier immer V ein endlichdimensionaler K -Vektorraum, n seine Dimension und ϕ ein Endomorphismus von V mit charakteristischem Polynom P_ϕ .

14.1 Definitionen

Definition 14.1 (Minimalpolynom)

Das Minimalpolynom ist das eindeutige normierte Polynom m_ϕ , das unter allen Polynomen f mit $f(\phi) = 0$ den kleinsten Grad hat.

Anmerkung: Dass diese Definition Sinn macht, sagt uns Satz 14.1.

Definition 14.2 (Hauptraum, algebraische Vielfachheit)

Sei $\lambda \in K$ ein Eigenwert von ϕ . Dann heißt

$$\mu(\phi, \lambda) := \max\{r : (x - \lambda)^r \mid P_\phi\}$$

die **algebraische** Vielfachheit von λ . Der Untervektorraum

$$\text{Hau}(\phi, \lambda) := \ker(\phi - \lambda \text{Id}_V)^{\mu(\phi, \lambda)} \subset V$$

heißt der **Hauptraum** zum Eigenwert λ .

Definition 14.3 (nilpotent)

Ein Endomorphismus $\phi : V \rightarrow V$ heißt **nilpotent**, wenn es ein $d \in \mathbb{N}$ gibt mit

$$\phi^d = 0.$$

Definition 14.4 (Jordan-Block)

Für $d, s \in \mathbb{N}, \lambda \in K$ definieren wir

$$J_d(\lambda) := \begin{pmatrix} \lambda & 1 & & \\ & \lambda & 1 & \\ & & \ddots & \ddots \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix} \in \mathcal{M}_{d,d}(K)$$

und

$$J_d^{(s)}(\lambda) := \begin{pmatrix} J_d(\lambda) & & \\ & \ddots & \\ & & J_d(\lambda) \end{pmatrix} \in \mathcal{M}_{ds,ds}(K)$$

Die Matrix $J_d(\lambda)$ heißt ein **Jordan-Block** der Länge d zum Eigenwert λ .

14.2 Sätze und Beweise

Satz 14.1

Das Minimalpolynom ist wohldefiniert und für jedes Polynom p mit $p(\phi) = 0$ gilt $m_\phi | p$. Außerdem kommt jeder Linearfaktor von P_ϕ auch in m_ϕ vor, und m_ϕ besitzt keine weiteren Linearfaktoren.

Beweis: Wir beweisen nur den ersten Teil, müssen also zeigen, dass es für jeden Endomorphismus genau ein Minimalpolynom gibt. Aus dem Satz von Caley-Hamilton folgt aber $P_\phi(\phi) = 0$, also gibt es ein Polynom, das die Bedingungen erfüllt. Wir müssen noch zeigen, dass dieses eindeutig ist. Angenommen, es gibt zwei solche Polynome. Dann müssen beide den gleichen Grad haben, sagen wir m und es seien

$$f(x) = x^m + \dots + a_0, \quad g(x) = x^m + \dots + b_0$$

diese Polynome. Dann ist

$$h(x) := f(x) - g(x) = c_{m-1}x^{m-1} + \dots + c_0$$

ein Polynom vom Grad höchstens $m - 1$. Hieraus folgt schon die Behauptung, wie ihr euch einmal selbst überlegen solltet. q.e.d.

Satz 14.2 (Trigonalisierung)

Wir nehmen an, dass das charakteristische Polynom in Linearfaktoren zerfällt, das heißt

$$P_\phi = (x - \lambda_1) \cdots (x - \lambda_n)$$

mit $\lambda_1, \dots, \lambda_n \in K$. Dann gibt es eine Basis \mathcal{B} von V , sodass die Darstellungsmatrix von ϕ bezüglich \mathcal{B} obere Dreiecksform hat.

Satz 14.3 (Hauptraumzerlegung)

Wir nehmen an, dass das charakteristische Polynom von ϕ in Linearfaktoren zerfällt, etwa

$$P_\phi = (x - \lambda_1)^{n_1} \cdots (x - \lambda_r)^{n_r}$$

mit $\lambda_i \neq \lambda_j$ für $i \neq j$, das heißt $n_i = \mu(\phi, \lambda_i)$. Dann haben wir eine Zerlegung

$$V = \text{Hau}(\phi, \lambda_1) \oplus \cdots \oplus \text{Hau}(\phi, \lambda_r)$$

und die folgenden Aussagen sind äquivalent:

i) ϕ ist diagonalisierbar.

ii) Für $i = 1, \dots, r$ gilt:

$$\text{Eig}(\phi, \lambda_i) = \text{Hau}(\phi, \lambda_i).$$

iii) Für $i = 1, \dots, r$ gilt:

$$\dim_K \text{Eig}(\phi, \lambda_i) = \mu(\phi, \lambda_i).$$

iv) $m_\phi = (x - \lambda_1) \cdots (x - \lambda_r)$.

Satz 14.4 (nilpotente Endomorphismen)

Für einen Endomorphismus ϕ sind die folgenden Bedingungen äquivalent:

i) Es gibt eine Basis \mathcal{B} von V , sodass

$$M_{\mathcal{B}}(\phi) = \begin{pmatrix} 0 & * & \cdots & * \\ & 0 & & \vdots \\ & & \ddots & * \\ & & & 0 \end{pmatrix}.$$

ii) $P_\phi = x^n$.

iii) Es gilt $\phi^d = 0$ für ein d mit $1 \leq d \leq n$.

iv) ϕ ist nilpotent.

Beweis: Die Richtungen $i) \Rightarrow ii) \Rightarrow iii) \Rightarrow iv)$ sind trivial, es bleibt zu zeigen, dass aus $iv)$ auch $i)$ folgt.

Ist ϕ nilpotent, so wählen wir ein e so, dass $\phi^e = 0$. Dann folgt aus der Definition des Minimalpolynoms $m_\phi | x^e$. Der einzige Linearfaktor von m_ϕ ist also das Monom x , also kann auch in P_ϕ kein anderer Linearfaktor auftreten, also folgt $P_\phi = x^n, m_\phi = x^d$ mit $1 \leq d \leq n$. Damit folgt die Behauptung dann aus der Trigonalisierung. q.e.d.

Satz 14.5

Sei $\phi : V \rightarrow V$ ein nilpotenter Endomorphismus mit Minimalpolynom $m_\phi = x^d, 0 < d \leq n$. Dann existieren eindeutig bestimmte nicht negative ganze Zahlen $s_1, \dots, s_d \geq 0$ mit $s_d \geq 1$ und

$$n = s_1 + 2s_2 + \dots + ds_d$$

und eine Basis \mathcal{B} von V , sodass

$$M_{\mathcal{B}}(\phi) = \begin{pmatrix} J_1^{(s_1)} & & \\ & \ddots & \\ & & J_d^{(s_d)} \end{pmatrix}.$$

Satz 14.6 (Jordan-Normalform)

Seien V ein K -Vektorraum der Dimension n und $\phi : V \rightarrow V$ ein Endomorphismus, dessen charakteristisches Polynom in Linearfaktoren zerfällt,

$$P_\phi = (x - \lambda_1)^{n_1} \dots (x - \lambda_r)^{n_r}$$

mit $\lambda_i \neq \lambda_j$ für $i \neq j$ und $n_i \geq 1$. Sei

$$m_\phi = (x - \lambda_1)^{d_1} \dots (x - \lambda_r)^{d_r}$$

das Minimalpolynom. Es gilt $1 \leq d_i \leq n_i$. Dann gibt es eindeutig bestimmte nicht negative ganze Zahlen

$$s_1^{(i)}, \dots, s_{d_i}^{(i)}, \quad i = 1, \dots, r$$

mit

$$s_1^{(i)} + 2s_2^{(i)} + \dots + d_i s_{d_i}^{(i)} = n_i, \quad s_{d_i}^{(i)} > 0$$

und eine Basis \mathcal{B} von V , sodass

$$M_{\mathcal{B}}(\phi) = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}$$

mit

$$A_i = \begin{pmatrix} J_1^{(s_1^{(i)})}(\lambda_i) & & \\ & \ddots & \\ & & J_{d_i}^{(s_{d_i}^{(i)})}(\lambda_i) \end{pmatrix}.$$

Beweis: Wir setzen $U_i := \text{Hau}(\phi, \lambda_i)$. Dann gilt $\dim_K U_i = n_i$ und nach dem Satz 14.3 über die Hauptraumzerlegung gilt:

$$V = U_1 \oplus \cdots \oplus U_n.$$

Aus dem (hier nicht gemachten) Beweis des Satzes über die Hauptraumzerlegung würde außerdem folgen, dass $\phi_i := \phi|_{U_i}$ wieder ein Endomorphismus von U_i ist. Dieser hat natürlich das charakteristische Polynom $P_{\phi_i} = (x - \lambda_i)^{n_i}$. Nach dem Satz von Caley-Hamilton gilt:

$$P_{\phi_i}(\phi_i) = (\phi_i - \lambda_i)^{n_i} = 0.$$

Also ist der Endomorphismus $\phi_i - \lambda_i$ nilpotent. Nach Satz 14.5 gibt es also Zahlen $s_j^{(i)}$ mit

$$n_i = s_1^{(i)} + \cdots + d_i s_{d_i}^{(i)}$$

und eine Basis \mathcal{B}_i von U_i , sodass die Matrix $B_i := M_{\mathcal{B}_i}(\phi_i - \lambda_i)$ die Form

$$M_{\mathcal{B}}(\phi) = \begin{pmatrix} J_1^{(s_1^{(i)})} & & \\ & \ddots & \\ & & J_{d_i}^{(s_{d_i}^{(i)})} \end{pmatrix}$$

hat. Das bedeutet aber gerade, dass die Matrizen $A_i := M_{\mathcal{B}_i}(\phi_i)$ die gewünschte Form haben. Da die Abbildungen ϕ_i Endomorphismen von U_i sind, also mit U_j für $i \neq j$ nichts zu tun haben, folgt die Aussage einfach, indem wir als Basis die Vereinigung der einzelnen Basen wählen, das heißt

$$\mathcal{B} := \bigcup_i \mathcal{B}_i.$$

q.e.d.

Satz 14.7 (Exponentialabbildung von Matrizen)

Seien $A \in \mathcal{M}_n(\mathbb{R})$ und J die Jordan-Normalform mit $J = T^{-1}AT$. Dann gilt

i) Sind J_i quadratische Matrizen mit

$$J = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_r \end{pmatrix},$$

dann ist

$$e^{tJ} = \begin{pmatrix} e^{tJ_1} & & \\ & \ddots & \\ & & e^{tJ_r} \end{pmatrix}$$

ii) Für $\lambda \in \mathbb{C}$ gilt:

$$e^{(\lambda E_n + B)t} = e^{\lambda t} e^{Bt}.$$

iii) Es gilt:

$$\exp\left(t \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ \vdots & & & \ddots & 1 \\ 0 & \cdots & \cdots & \cdots & 0 \end{pmatrix}\right) = \begin{pmatrix} 1 & t & \frac{t^2}{2} & \cdots & \frac{t^k}{k!} \\ 0 & 1 & t & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \frac{t^2}{2} \\ \vdots & & & \ddots & t \\ 0 & \cdots & \cdots & \cdots & 1 \end{pmatrix}.$$

iv) Es ist

$$e^{tA} = e^{(TJT^{-1})t} = Te^{tJ}T^{-1}.$$

Anmerkung: Für die Definition von e^A für eine Matrix A siehe Definition 1.29.

14.3 Erklärungen zu den Definitionen

Zur Definition 14.1 des Minimalpolynoms:

Beispiel 121

Wir wollen einmal das Minimalpolynom konkret berechnen. Sei ϕ gegeben durch die Matrix

$$A := \begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & -2 \\ -2 & 2 & 3 \end{pmatrix}.$$

Wie können wir nun das Minimalpolynom berechnen? Satz 14.1 gibt uns einen kleinen Hinweis: Wir berechnen zunächst das charakteristische Polynom der Matrix. Mit der Regel von Sarrus erhalten wir

$$P_A = (x+1)(x-1)^2.$$

(Das solltet ihr natürlich selbst nachrechnen). Nun gibt es also genau zwei Möglichkeiten für das Minimalpolynom, nämlich

$$M_1 = (x+1)(x-1)^2 \quad \text{oder} \quad M_2 = (x+1)(x-1) = x^2 - 1.$$

Hier müssen wir jetzt ausprobieren, wir überprüfen immer zuerst das Polynom, das kleinsten Grad hat. Wir berechnen also

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

also ist $A^2 - E_2 = 0$ und damit ist das Minimalpolynom $m_\phi = x^2 - 1$. ■

Besonders einfach ist es damit natürlich, wenn das charakteristische Polynom nur einfache Nullstellen hat. Ist zum Beispiel

$$A := \begin{pmatrix} 1 & 7 & 9 & 5 \\ 0 & 2 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 4 \end{pmatrix},$$

so ist das charakteristische Polynom

$$P_A = (x-1)(x-2)(x-3)(x-4),$$

und da das Minimalpolynom ein Teiler von P_A sein muss, der jeden Linearfaktor enthält, folgt hier sofort $m_A = P_A$.

Zur Definition 14.2 des Hauptraums: Den Begriff des Hauptraums wollen wir hier nicht weiter vertiefen, konkrete Berechnungen werden noch in Beispielen zur Jordan-Form folgen. Wir erwähnen nur, dass offenbar

$$\text{Eig}(\phi, \lambda) = \ker(\phi - \lambda \text{Id}_V) \subset \text{Hau}(\phi, \lambda)$$

gilt, und daher

$$\dim_K \text{Eig}(\phi, \lambda) \leq \mu(\phi, \lambda).$$

$\dim_K \text{Eig}(\phi, \lambda)$ nennen wir die **geometrische Vielfachheit** des Eigenwertes λ . Dazu ein kurzes Beispiel.

Beispiel 122

Wir betrachten die Matrix

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

Das charakteristische Polynom von A ist $P_A(x) = (x - 1)^2$. Also ist die algebraische Vielfachheit von 1 gleich 2. Suchen wir andererseits nach Eigenvektoren zum Eigenwert 1, so sind dies nur alle Vielfachen von $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Der Eigenraum ist also eindimensional und damit ist die geometrische Vielfachheit 1. Dieser Fall, nämlich dass algebraische und geometrische Vielfachheit nicht übereinstimmen, kann nur dann eintreten, wenn A nicht diagonalisierbar ist. ■

Zur Definition 14.3 von nilpotent: Ein Endomorphismus heißt also nilpotent, wenn er jedes Element, so lange man ihn nur oft genug anwendet, auf die 0 schickt. Hierzu ein Beispiel:

Beispiel 123

Seien $K = \mathbb{R}$ und V der Vektorraum der Polynome von Grad $< n$, also $V = \mathbb{R}[x]_{<n}$. Als Endomorphismus betrachten wir die Ableitung nach x , das heißt

$$\phi : V \rightarrow V, V \ni f \mapsto f'.$$

Da die Ableitung linear ist, das heißt $(\lambda f + g)' = \lambda f' + g'$, ist dies tatsächlich ein Endomorphismus. Wir behaupten, dass er nilpotent ist. Dafür betrachten wir, was mit dem Grad eines Polynoms passiert, wenn wir es ableiten. Dieser verringert sich natürlich um 1. Starten wir also mit einem Polynom vom Grad $k < n$ und leiten es k mal ab, so erhalten wir eine Konstante, $f^{(k)} = c$ für ein $c \in \mathbb{R}$. Leiten wir dies noch einmal ab, so ergibt sich 0. Es gilt also auf jeden Fall $\phi^n = 0$. ■

Das d in der Definition kann also in diesem Fall als die Dimension von V gewählt werden. Allgemeiner lässt sich bei nilpotenten Endomorphismen immer ein d finden, das höchstens $\dim V$ ist, siehe Satz 14.4.

Zu Definition 14.4 des Jordan-Blocks: Ein Spezialfall dieser Matrizen ist

$$J_d := \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix} \in \mathcal{M}_{d,d}(K)$$

und

$$J_d^{(s)} := \begin{pmatrix} J_d & & \\ & \ddots & \\ & & J_d \end{pmatrix} \in \mathcal{M}_{ds, ds}(K).$$

Die Form dieser Matrizen scheint zunächst einmal vom Himmel zu fallen. Betrachten wir allerdings die Matrix aus Satz 14.4, so fallen gewisse Ähnlichkeiten zu den Matrizen J_d auf, nämlich dass auf und unter der Diagonalen nur Nullen stehen. Die Matrizen $J_d^{(s)}$ werden wir nutzen, um Darstellungsmatrizen nilpotenter Endomorphismen in eine schöne Gestalt zu bringen, siehe Satz 14.5.

Da wir jedoch nicht nur an nilpotenten Endomorphismen interessiert sind, benötigen wir noch die Matrizen $J_d^{(s)}(\lambda)$. Genauere Erklärungen findet ihr bei den Erklärungen zum Satz 14.6.

14.4 Erklärungen zu den Sätzen und Beweisen

Zur Eindeutigkeit des Minimalpolynoms (Satz 14.1): Für diejenigen, die den Beweis nicht fertig gemacht haben: Es ist natürlich $h(\phi) = f(\phi) - g(\phi)$, und dies ist nach Voraussetzung 0. Also ist h ein Polynom kleineren Grades als f und g mit $h(\phi) = 0$. Falls h nicht das Nullpolynom ist, bedeutet das aber, dass f und g keine Minimalpolynome waren, im Widerspruch zur Annahme. Ist h dagegen das Nullpolynom, so waren f und g identisch, auch das war ausgeschlossen.

Diese Methode ist üblich, wenn man zeigen will, dass es nur ein Element gibt, das eine gewisse Minimalität erfüllt. Hat man nämlich zwei, so kann man daraus meistens ein noch minimaleres Element erzeugen, was ein Widerspruch ist.

Für die zweite Behauptung benötigt man etwas mehr Theorie, genauer Ideale. Dies könnt ihr in den Lineare-Algebra-Büchern unserer Literaturliste nachlesen. Die letzte Eigenschaft bedeutet gerade, dass wenn

$$P_\phi = (x - \lambda_1)^{r_1} \cdots (x - \lambda_n)^{r_n}$$

mit $r_i \geq 1$ gilt, dann

$$m_\phi = (x - \lambda_1)^{t_1} \cdots (x - \lambda_n)^{t_n}$$

mit $1 \leq t_i \leq r_i$ ist.

Zur Trigonalisierung (Satz 14.2): Schon wieder ein Satz, den wir nicht beweisen! Üblicherweise wird dieser Satz zusammen mit einem zusätzlichen Resultat gebracht, welches äquivalent zu der Existenz einer solchen Basis wie im Satz ist. über diese Aussage wird der Satz dann auch bewiesen. Da uns aber vor allem das Resultat des Satzes für die weitere Verwendung interessiert, überlassen wir es euch, den Beweis in einem Buch eurer Wahl nachzulesen.

Das Stichwort hierzu ist ϕ -invariante Unterräume.

Die Aussage hier ist etwas schwächer als bei der Diagonalisierung, dafür sind die Voraussetzungen aber auch wesentlich schwächer. Dort konnte man eine Basis finden, sodass die Matrix Diagonalgestalt hat, das heißt, obere und untere Dreiecksmatrix. Dies geht nun im Allgemeinen nicht mehr. Die Diagonalisierung ist also ein Spezialfall der Trigonalisierung. Natürlich stehen auch hier auf der Diagonale die Eigenwerte.

Noch eine kurze Anmerkung: Die Voraussetzung, dass das charakteristische Polynom in Linearfaktoren zerfällt, ist über den komplexen Zahlen natürlich immer erfüllt.

Zur Hauptraumzerlegung (Satz 14.3): Auch dieser Beweis funktioniert wieder über ϕ -invariante Unterräume. Hier ist wichtig, sich zu merken, dass ein Endomorphismus genau dann diagonalisierbar ist, wenn die Dimension des Hauptraumes eines Eigenwertes mit der Dimension des Eigenraums des gleichen Eigenwertes übereinstimmt. Kennen wir also bereits eine dieser äquivalenten Eigenschaften, so können wir das andere daraus ableiten und auch einfach das Minimalpolynom bestimmen. Ist andersrum das Minimalpolynom gegeben, so können wir daraus auf die Diagonalisierbarkeit schließen.

Beispiel 124

Wir hatten bereits als Minimalpolynom von

$$A := \begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & -2 \\ -2 & 2 & 3 \end{pmatrix}$$

das Polynom $m_A = (x-1)(x+1)$ erhalten. Wir bekommen nun also aus dem Satz sofort, dass A diagonalisierbar ist. ■

Zum Satz 14.4 über nilpotente Endomorphismen: Dieser Satz gibt uns nun einige Eigenschaften, mit denen wir feststellen können, ob ein Endomorphismus nilpotent ist. Dazu wollen wir wieder ein Beispiel durchrechnen.

Beispiel 125

Wir schließen an Beispiel 123 an. Wir haben bereits gesehen, dass die Ableitung nilpotent ist, und dass tatsächlich das d kleiner gleich (in diesem Fall sogar gleich) der Dimension von V ist. Wir wollen zeigen, dass die beiden anderen Bedingungen auch gelten. Zuerst wollen wir bezüglich einer geeigneten Basis die Darstellungsmatrix bestimmen. Wir nehmen die übliche Basis $\mathcal{B} = \{1, x, x^2, \dots, x^{n-1}\}$. Dann gilt:

$$\phi(1) = 0, \quad \phi(x) = 1, \quad \phi(x^2) = 2x, \quad \phi(x^3) = 3x^2, \dots, \quad \phi(x^{n-1}) = (n-1)x^{n-2}.$$

Wir erhalten also die Darstellungsmatrix

$$M_{\mathcal{B}}(\phi) = \begin{pmatrix} 0 & 1 & & & \\ 0 & 0 & 2 & & \\ 0 & 0 & 0 & 3 & \\ \vdots & \vdots & \vdots & \ddots & \ddots \\ 0 & 0 & 0 & \cdots & 0 & n-1 \end{pmatrix}.$$

Also hat die Matrix die gewünschte Form.

Daraus würde auch sofort folgen, dass $P_{\phi} = x^n$ gilt. Dies wollen wir aber unabhängig von der Wahl einer Basis nochmal zeigen. Sei λ ein Eigenwert von ϕ , es gelte also $f' = \lambda f$. Dies ist eine Differentialgleichung, vergleiche Kapitel 6, die Lösung ist gegeben durch

$$f = e^{\lambda x}.$$

Nun muss aber f ein Polynom sein. Dies ist natürlich nur für $\lambda = 0$ erfüllt, der einzige Eigenwert ist also 0. Und das einzige charakteristische Polynom, das zu einem solchen Endomorphismus gehört, ist $P_{\phi} = x^n$. ■

Zum Satz 14.5: Der Beweis dieses Satzes ist sehr technisch und erfordert einige Hilfssätze, deshalb wollen wir ihn hier nicht bringen. Stattdessen wollen wir genauer verstehen, was der Satz aussagt.

Haben wir einen nilpotenten Endomorphismus mit gegebenem Minimalpolynom, so können wir eine Basis finden, sodass die Darstellungsmatrix eine sehr einfache Form hat. Ist zum Beispiel $n = 10$ und $d = 4$, so könnte zum Beispiel

$$s_4 = s_3 = s_2 = s_1 = 1 \text{ oder auch } s_4 = 1, s_2 = 3$$

gelten. Im ersten Fall enthielte die Matrix $M_{\mathcal{B}}(\phi)$ die 4 Matrizen J_1, J_2, J_3, J_4 es wäre also

$$M_{\mathcal{B}}(\phi) = \begin{pmatrix} 0 & 0 & & & & & & & & \\ & 0 & 1 & & & & & & & \\ & & 0 & 0 & & & & & & \\ & & & 0 & 1 & & & & & \\ & & & & 0 & 1 & & & & \\ & & & & & 0 & 0 & & & \\ & & & & & & 0 & 1 & & \\ & & & & & & & 0 & 1 & \\ & & & & & & & & 0 & 1 \\ & & & & & & & & & 0 \end{pmatrix}.$$

Im zweiten Fall wäre

$$M_{\mathcal{B}}(\phi) = \begin{pmatrix} 0 & 1 & & & & \\ & 0 & 0 & & & \\ & & 0 & 1 & & \\ & & & 0 & 0 & \\ & & & & 0 & 1 \\ & & & & & 0 & 0 \\ & & & & & & 0 & 1 \\ & & & & & & & 0 & 1 \\ & & & & & & & & 0 & 1 \\ & & & & & & & & & 0 \end{pmatrix}.$$

Beispiel 126

Wir wollen nun einmal die Basis und Matrix für den Endomorphismus aus Beispiel 123 beziehungsweise Beispiel 125 bestimmen. Zunächst müssen wir das Minimalpolynom berechnen. Wir wissen bereits $P_{\phi} = x^n$, also kommt für das Minimalpolynom nur noch $x^d, 1 \leq d \leq n$ infrage, das bedeutet $\phi^d = 0$. Dies wiederum heißt nichts anderes, als dass ϕ^d angewendet auf jedes Element des Vektorraums 0 ergeben muss. Wir betrachten nun das Element $x^{n-1} \in V$ und schauen, was passiert, wenn wir dieses d -mal mit $d < n$ ableiten:

$$\begin{aligned} x^{n-1} &\mapsto (n-1)x^{n-2} \mapsto (x-1)(x-2)x^{n-3} \mapsto \dots \mapsto (n-1) \dots (n-d)x^{n-1-d} \\ &= \frac{(n-1)!}{(n-1-d)!} x^{n-1-d}. \end{aligned}$$

Es gilt also für $d < n$:

$$\phi^d(x^{n-1}) = \frac{(n-1)!}{(n-1-d)!} x^{n-1-d}$$

Nun ist aber x^{n-1-d} für $x \neq 0$ auch $\neq 0$, deshalb kann für $d < n$ nicht $\phi^d = 0$ gelten. Also muss x^n das Minimalpolynom sein. Als einzige Zerlegung

$$n = s_1 + \dots + ds_d$$

bleibt wegen $d = n$ und $s_n \geq 1$ nur $n = 1 \cdot n$ übrig, also finden wir eine Basis \mathcal{B} , sodass

$$M_{\mathcal{B}}(\phi) = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}.$$

Dies hatten wir in Beispiel 123 schon fast geschafft, es stimmen nur die Werte der Elemente über den Diagonalen nicht. Wir wollen also unsere dortige Basis nur leicht modifizieren und machen den Ansatz

$$\mathcal{B} = \{c_0, c_1x, c_2x^2, \dots, c_{n-1}x^{n-1}\}$$

mit $c_j \in \mathbb{R}$. Dann stimmt auf jeden Fall die Form überein, wir müssen nur dafür sorgen, die c_j so zu wählen, dass die Werte über der Diagonalen stimmen. Damit dort überall 1 steht, muss aber einfach nach Definition der Darstellungsmatrix

$$(c_jx^j)' = c_{j-1}x^{j-1}$$

gelten. Wegen $(c_jx^j)' = c_jjx^{j-1}$ führt dies auf die Rekursion

$$c_j = \frac{c_{j-1}}{j},$$

und da in Beispiel 125 der erste Eintrag schon gestimmt hatte, können wir $c_0 = 1$ wählen. Wir erhalten dann

$$c_0 = 1, c_1 = 1, c_2 = \frac{1}{2}, c_3 = \frac{1}{6}, \dots, c_j = \frac{1}{j!}.$$

Also ist die Basis \mathcal{B} , bezüglich der der Endomorphismus ϕ die obige Matrixgestalt hat

$$\mathcal{B} = \{1, x, \frac{1}{2}x^2, \frac{1}{6}x^3, \dots, \frac{1}{(n-1)!}x^{n-1}\}.$$

■

Wie man eine solche Basis oder die Zerlegung von n im Allgemeinen findet, wollen wir hier nicht besprechen, dies werden wir in der nächsten Erklärung in einem allgemeineren Fall tun.

Zum Satz 14.6 über die Jordan-Normalform: Dieser Satz stellt nun das Hauptergebnis dieses Abschnitts dar. Im für uns wichtigen Fall, das heißt $K = \mathbb{C}$, zerfällt natürlich jedes Polynom in Linearfaktoren, das heißt, jede Matrix besitzt eine Jordan-Normalform. Wir wollen uns nun damit beschäftigen, sie zu berechnen.

Beispiel 127

In einem Beispiel des ersten Bandes haben wir die Matrix

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

untersucht. Wir haben gesehen, dass das charakteristische Polynom durch $(x-1)^2(x+3)$ gegeben war, und die Matrix war nicht diagonalisierbar. Diese Informationen genügen uns, um die Jordan-Normalform anzugeben, denn für das charakteristische Polynom $(x-1)^2(x+3)$ gibt es als Jordan-Normalform nur die beiden Möglichkeiten

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix}, \quad \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -3 \end{pmatrix}$$

und da A nicht diagonalisierbar ist, ist die Jordan-Normalform

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

■

Nun wollen wir aber auch noch eine Basis mit den gewünschten Eigenschaften, beziehungsweise eine Basiswechselmatrix T mit $J = T^{-1}AT$ finden. Hierbei bezeichne J die Jordan-Normalform von A . Der Beweis gibt uns schon eine Hilfe, wir müssen Haupträume bestimmen. Dies wollen wir zunächst an einem Beispiel tun.

Beispiel 128

Wir betrachten die Matrix

$$A = \begin{pmatrix} -1 & -6 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & -3 & 2 \\ 1 & 2 & -2 & 1 \end{pmatrix} \in M_{4,4}(\mathbb{R})$$

und wollen eine Jordan-Form J von A und eine Matrix $T \in GL_{4,4}(\mathbb{R})$ mit $T^{-1}AT = J$ bestimmen.

Wir berechnen zunächst die Eigenwerte von A . Das charakteristische Polynom ist gegeben durch

$$P_A = \det \begin{pmatrix} x+1 & 6 & 0 & 0 \\ 0 & x-2 & 0 & 0 \\ 0 & 0 & x+3 & -2 \\ -1 & -2 & 2 & x-1 \end{pmatrix} = (x+1)^3(x-2).$$

Wir betrachten jetzt den Eigenwert $x = -1$. Es gilt:

$$\text{Eig}(A, -1) = \ker(A + E) = \ker \begin{pmatrix} 0 & -6 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & -2 & 2 \\ 1 & 2 & -2 & 2 \end{pmatrix} = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\rangle.$$

Somit gibt es $\dim(\ker(A + E)) = 1$ Jordan-Kästchen zum Eigenwert -1 (notwendigerweise der Ordnung 3). Wir berechnen die Kerne von $(A + E)^2$ und $(A + E)^3$.

$$\begin{aligned} \ker(A + E)^2 &= \ker \begin{pmatrix} 0 & -18 & 0 & 0 \\ 0 & 9 & 0 & 0 \\ 2 & 4 & 0 & 0 \\ 2 & 4 & 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\rangle. \\ \ker(A + E)^3 &= \ker \begin{pmatrix} 0 & -54 & 0 & 0 \\ 0 & 27 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle. \end{aligned}$$

Setze

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \in \ker(A + E)^3 \setminus \ker(A + E)^2,$$

$$v_2 = (A + E) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

$$v_3 = (A + E)^2 \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 2 \end{pmatrix}.$$

Der Eigenwert $x = 2$ ist einfach. Das zugehörige Jordan-Kästchen ist also der Ordnung 1. Es gilt:

$$\text{Eig}(A, 2) = \ker(A - 2E) = \ker \begin{pmatrix} -3 & -6 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -5 & -2 \\ 1 & 2 & -2 & -1 \end{pmatrix} = \left\langle \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\rangle.$$

Wir setzen $v_4 = \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$. Die Matrix $T := (v_3, v_2, v_1, v_4)$ erfüllt dann

$$J := \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} = T^{-1}AT.$$

■

Was haben wir also getan? Zuerst müssen wir natürlich das charakteristische Polynom berechnen und daraus die Eigenwerte ablesen. Dann bestimmen wir, wie schon beim Diagonalisieren, die Kerne der zugehörigen Matrizen. Kommt ein Eigenwert nur einfach vor, so sind wir damit fertig. Erhalten wir von einem mehrfach vorkommenden Eigenwert weniger Eigenvektoren, so müssen wir auch noch die Kerne höherer Potenzen berechnen, und zwar so lange, bis wir so viele Vektoren haben, wie der Eigenwert vorkommt. Für jeden Vektor v , der dann im zuletzt berechneten Kern, aber nicht in dem davor liegt, berechnen wir sukzessive die Vektoren $(A - \lambda E)v$, das ganze so lange, bis ein Vektor rauskommt, der im ersten Kern liegt. Erhalten wir so nicht alle Vektoren, die wir berechnet haben, müssen wir mit Elementen voriger Kerne weiter machen. Genauer: Angenommen wir haben eine (5×5) -Matrix mit charakteristischem Polynom $(x - 1)^5$. Weiter angenommen, wir erhalten

$$\begin{aligned} \ker(A + E) &= \langle v_1, v_2 \rangle, \ker(A + E)^2 = \langle v_1, v_2, v_3, v_4 \rangle, \ker(A + E)^3 \\ &= \langle v_1, v_2, v_3, v_4, v_5 \rangle. \end{aligned}$$

Wir wählen den Vektor v_5 und berechnen $(A + E)v_5$. Dies ergebe nun v_4 . Dann berechnen wir $(A + E)^2 v_5 = (A + E)v_4$. Sei dies v_3 . Dann ist $(A + E)^3 v_5 = (A + E)v_3 = 0$. Wir haben also die drei Vektoren v_1, v_4, v_5 erhalten. Diese korrespondieren zu einem Jordan-Block der Länge 3. Wir haben aber noch nicht genug Vektoren erhalten. Wir wählen also als nächstes v_3 und berechnen $(A + E)v_3$. Dies ergibt nun v_2 und liefert noch einen Jordan-Block der Länge 2.

Nun erhalten wir die Matrix T durch spaltenweises Schreiben der gefundenen Vektoren.

Die ganze Prozedur nochmal in Kurzform:

1. Charakteristisches Polynom bestimmen und Eigenwerte ablesen.
2. Kerne bestimmen.
3. Vektoren berechnen und dadurch Jordan-Blöcke bestimmen.
4. Jordan-Blöcke zur Jordan-Normalform und Vektoren zur Matrix T zusammenfügen.

Hierzu ein weiteres, leicht modifiziertes Beispiel.

Beispiel 129

Der Endomorphismus f des \mathbb{R}^4 sei gegeben durch

$$M_E^E(f) = \begin{pmatrix} -2 & 0 & 0 & 1 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

mit Standardbasis E des \mathbb{R}^4 . Wir wollen eine Jordan-Basis \mathcal{B} des \mathbb{R}^4 zu f , sowie die zugehörige Matrix $M_{\mathcal{B}}^{\mathcal{B}}(f)$ bestimmen. Hier haben wir analog wie oben zu verfahren, denn die Matrix T ist ja im Grunde ein Basiswechsel.

Man sieht sofort, dass $P_A = x^3(x+2)$.

$(A + 2E)x = 0$ liefert den Eigenraum

$$\text{Eig}(A, -2) = \text{span}\{(2, 1, 0, 0)^T\}.$$

Ferner sieht man

$$\text{Eig}(A, 0) = \text{span}\{e_2, e_3\}.$$

Damit ist der Hauptraum von A zum Eigenwert 0 der Lösungsraum $A^3x = 0$. Von A^3 benötigen wir nur die zweite Zeile, die sich zu $\begin{pmatrix} -4 & 0 & 0 & 2 \end{pmatrix}$ berechnet. Dies liefert sofort

$$\text{Hau}(A, 0) = \text{span}\{e_2, e_3, (1, 0, 0, 2)^T\}.$$

Mit

$$A \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

folgt die Jordan-Basis

$$\mathcal{B} = \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right\}$$

und

$$M_{\mathcal{B}}^{\mathcal{B}}(f) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}.$$

■

Das Prinzip ist also jedes Mal dasselbe. Zum Abschluss noch zwei Beispiele.

Beispiel 130

Es seien V ein vierdimensionaler Vektorraum und $P_{\phi} = (x-1)(x+1)^3$ das charakteristische Polynom des Endomorphismus ϕ . Wir wollen untersuchen, welche Fälle auftreten können.

Da der Eigenwert $\lambda = 1$ mit Multiplizität 1 auftritt, gilt $\text{Eig}(\phi, 1) = \text{Hau}(\phi, 1)$. Der zweite Hauptraum $\text{Hau}(\phi, -1)$ ist dreidimensional. Für die Dimension des zugehörigen Eigenraums gibt es also genau drei Möglichkeiten, $\dim_K \text{Eig}(\phi, -1) \in \{1, 2, 3\}$.

Angenommen, $\dim_K \text{Eig}(\phi, -1) = 3$. Dann gilt:

$$\text{Eig}(\phi, -1) = \text{Hau}(\phi, -1),$$

und ϕ ist diagonalisierbar, das heißt, es gibt eine Basis \mathcal{B} mit

$$M_{\mathcal{B}}(\phi) = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}.$$

Im Fall $\dim_K \text{Eig}(\phi, -1) < 3$ ist ϕ nicht diagonalisierbar. Die Jordan-Normalform von ϕ ist durch Zahlen d und s_1, \dots, s_d mit

$$1 \leq d \leq 3, s_1, \dots, s_{d-1} \geq 0, s_d > 0, s_1 + \dots + ds_d = 3$$

bestimmt.

Der Fall $d = 1, s_1 = 3$ entspricht dem schon behandelten Fall, dass ϕ diagonalisierbar ist. Im Fall $d = 2$ gibt es nur eine Möglichkeit, nämlich $s_1 = s_2 = 1$. Die Jordan-Normalform von ϕ ist dann

$$M_{\mathcal{B}}(\phi) = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & 1 \\ & & & -1 \end{pmatrix}.$$

In diesem Fall gilt $\dim_K \text{Eig}(\phi, -1) = 2$.

Im letzten Fall gilt $d = 3$ und dann notwendigerweise $s_1 = s_2 = 0, s_3 = 1$. Die Jordan-Normalform von ϕ ist

$$M_{\mathcal{B}}(\phi) = \begin{pmatrix} 1 & & & \\ & -1 & 1 & \\ & & -1 & 1 \\ & & & -1 \end{pmatrix},$$

und es gilt $\dim_K \text{Eig}(\phi, -1) = 1$. ■

An diesem Beispiel sehen wir, dass wir, um die Jordan-Normalform zu berechnen, die Eigenvektoren nicht brauchen. Wir wollen unsere Überlegungen auf eine gegebene Matrix anwenden.

Beispiel 131

Wir betrachten die Matrix

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

Das charakteristische Polynom von A ist

$$P_A = (x - 1)(x + 1)^3.$$

Wir sind also in der Situation von oben. Wir bestimmen

$$\dim_K \text{Eig}(A, -1) = 4 - \text{Rang}(A + 1) = 4 - 3 = 1.$$

Also können wir jetzt schon die Jordan-Normalform angeben, sie lautet

$$M_{\mathcal{B}}(\phi) = \begin{pmatrix} 1 & & & \\ & -1 & 1 & \\ & & -1 & 1 \\ & & & -1 \end{pmatrix}.$$

Wir sehen also, dass wir, wie schon beim Diagonalisieren, keine Eigenvektoren brauchen, um die Jordan-Form zu berechnen. Diese brauchen wir, um die Basiswechselmatrix zu finden. Das wollen wir noch ein letztes Mal kurz tun. Zunächst bestimmen wir einen Eigenvektor zum Eigenwert $\lambda = 1$. Es ist

$$v_1 := \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \text{Eig}(A, 1) = \langle v_1 \rangle.$$

Nun wählen wir uns einen Vektor $v_4 \in \ker(A + 1)^3 \setminus \ker(A + 1)^2$. Diese Kerne muss man dafür natürlich zuerst berechnen ;-). Wir geben gleich einen passenden Vektor an,

$$v_4 := \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Nun erhalten wir

$$v_3 := (A + 1)v_4 = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, v_2 := (A + 1)v_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix},$$

und die Basis, bezüglich der A die Jordan-Form hat, ist

$$\mathcal{B} = \{v_1, v_2, v_3, v_4\}.$$

■

Aber wir könnten hier so viele Beispiele geben wie wir wollen, das Wichtigste ist, dass ihr es einmal selbst probiert, und zwar an so vielen Beispielen, bis ihr es im Schlaf könnt ;-). Eine Anleitung habt ihr ja jetzt.

Zum Satz 14.7 über die Exponentialabbildung bei Matrizen: In Definition 1.29 hatten wir die Exponentialabbildung für Matrizen definiert. Mithilfe der Jordan-Normalform kann man diese nun leicht berechnen. Die Anleitung hierfür gibt der Satz. Wir betrachten dafür ein Beispiel.

Beispiel 132

Sei

$$A = \begin{pmatrix} -8 & 47 & -8 \\ -4 & 18 & -2 \\ -8 & 39 & -5 \end{pmatrix}.$$

Hiervon müssen wir zunächst die Jordan-Normalform und die Matrix T mit $J = T^{-1}AT$ berechnen. Dies wollen wir als Übung für euch lassen, nutzt dies am Besten auch gleich ;-). Als Ergebnis solltet ihr dann

$$J = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix}$$

erhalten. Als Nächstes unterteilen wir die Matrix J in quadratische Matrizen, das heißt in die einzelnen Jordan-Blöcke. Hier ist $J = J_1 + J_2$ mit

$$J_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad J_2 = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

Nun zerteilen wir jede dieser Matrizen zwei Anteile, der eine besteht aus der Einheitsmatrix, der zweite ist eine nilpotente Matrix. Es ist

$$J_1 = E_1, \quad J_2 = 2E_2 + N, \quad N := \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Für jede dieser Matrizen berechnen wir nun einzeln (für die nilpotente Matrix mithilfe des Satzes) die Exponentialabbildung. Man erhält

$$e^{E_1 t} = e^t, \quad e^{2E_2 t} = e^{2t}, \quad e^N = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}.$$

Nun fügen wir dies alles zusammen. Es ist

$$e^{J_1 t} = e^{E_1 t} = e^t, \quad e^{J_2 t} = e^{2E_2 t} e^N = e^{2t} \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix},$$

und damit ist

$$e^{Jt} = \begin{pmatrix} e^t & & \\ e^{2t} & te^{2t} & \\ & & e^{2t} \end{pmatrix}.$$

Nun berechnen wir noch

$$Te^{Jt} = \begin{pmatrix} 6e^t & 17e^{2t} & -4e^{2t} + 17te^{2t} \\ 2e^t & 6e^{2t} & -e^{2t} + 6te^{2t} \\ 5e^t & 14e^{2t} & -3e^{2t} + 14te^{2t} \end{pmatrix}$$

und schließlich ist

$$\begin{aligned}
 e^A &= T e^{Jt} T^{-1} = \begin{pmatrix} 6e^t & 17e^{2t} & -4e^{2t} + 17te^{2t} \\ 2e^t & 6e^{2t} & -e^{2t} + 6te^{2t} \\ 5e^t & 14e^{2t} & -3e^{2t} + 14te^{2t} \end{pmatrix} \begin{pmatrix} -4 & -5 & 7 \\ 1 & 2 & -2 \\ -2 & 1 & 2 \end{pmatrix} \\
 &= \begin{pmatrix} -24e^t + 25e^{2t} - 34te^{2t} & -30e^t + 30e^{2t} + 17te^{2t} & 42e^t - 42e^{2t} + 34te^{2t} \\ -8e^t + 8e^{2t} - 12te^{2t} & -10e^t + 11e^{2t} + 6te^{2t} & 14e^t - 14e^{2t} + 12te^{2t} \\ -20e^t + 20e^{2t} - 28te^{2t} & -25e^t + 25e^{2t} + 14te^{2t} & 35e^t - 34e^{2t} + 28te^{2t} \end{pmatrix} -
 \end{aligned}$$

Unschön, aber so ist das nunmal ;-)

■

Wir wollen nun nochmal die Schritte zusammenfassen, die nötig sind, um aus einer Matrix A die Matrix e^A zu bestimmen.

1. Bestimme die Jordan-Normalform J und die Matrix T mit $J = T^{-1}AT$. Berechne T^{-1} .
2. Unterteile J in die quadratischen Anteile (das heißt, in die Jordan-Blöcke) J_1, \dots, J_r .
3. Zerteile jeden Jordan-Block J_i in $J_i = \lambda_i E_{j_i} + N_i$, das heißt in eine Summe aus dem Vielfachen einer Einheitsmatrix und einer nilpotenten Matrix.
4. Berechne für jedes N_i mit dem Satz die Matrix e^{tN_i} . Bestimme außerdem $e^{t\lambda_i E_{j_i}}$.
5. Berechne e^{tJ_i} durch $e^{tJ_i} = e^{t\lambda_i E_{j_i}} e^{tN_i}$.
6. Füge die Matrizen e^{tJ_i} zusammen, um die Matrix e^{Jt} zu erhalten.
7. Berechne $e^A = T e^{Jt} T^{-1}$.

Abgesehen davon, dass es natürlich an sich schon nützlich ist, e^A bestimmen zu können, werden wir dieses beim Lösen von linearen Differentialgleichungen in mehreren Dimensionen gut gebrauchen können.

15 Tensoren und das Tensorprodukt

Übersicht

15.1 Definitionen	297
15.2 Sätze und Beweise	299
15.3 Erklärungen zu den Definitionen	301
15.4 Erklärungen zu den Sätzen und Beweisen	304

Wir wollen uns nun kurz mit dem Begriff des Tensors und des Tensorprodukts beschäftigen. Dabei wollen wir hier nicht tief in die Theorie eindringen.

In diesem Kapitel seien $V, W, V_1, \dots, V_p, W_1, \dots, W_q$ stets reelle endlichdimensionale Vektorräume. Dies kann man auch auf beliebige Körper verallgemeinern, werden wir hier aber nicht tun.

15.1 Definitionen

Definition 15.1 (p -Linearform, Tensoren)

Eine p -**Linearform** oder auch **Multilinearform** der Ordnung p auf V_1, \dots, V_p ist eine Abbildung $\phi : V_1 \times \dots \times V_p \rightarrow \mathbb{R}$, welche in jedem Argument linear ist. Oft nennt man Tensoren auch p -**Tensor** oder **Tensor p -ter Stufe**. Der Raum der p -Linearformen auf $V_1 \times \dots \times V_p$ wird durch die Operation

$$(\lambda\phi_1 + \mu\phi_2)(v_1, \dots, v_p) := \lambda\phi_1(v_1, \dots, v_p) + \mu\phi_2(v_1, \dots, v_p), \quad \lambda, \mu \in \mathbb{R}$$

selbst zu einem reellen Vektorraum, den wir mit $\text{Mult}(V_1, \dots, V_p)$ bezeichnen. Elemente $\phi \in \text{Mult}(V_1, \dots, V_p)$ nennt man auch **Tensoren** der Ordnung p .

Definition 15.2 (Tensorprodukt von Abbildungen)

Es seien $\phi \in \text{Mult}(V_1, \dots, V_p)$, $\psi \in \text{Mult}(W_1, \dots, W_q)$. Dann erhalten wir eine Abbildung $\phi \otimes \psi \in \text{Mult}(V_1, \dots, V_p, W_1, \dots, W_q)$ durch die Vorschrift

$$(\phi \otimes \psi)(v_1, \dots, v_p, w_1, \dots, w_q) = \phi(v_1, \dots, v_p)\psi(w_1, \dots, w_q).$$

Wir nennen $\phi \otimes \psi$ das **Tensorprodukt** von ϕ und ψ .

Definition 15.3 (Tensorprodukt von Vektorräumen)

Die Abbildung

$$\otimes : \text{Mult}(V_1, \dots, V_p) \times \text{Mult}(W_1, \dots, W_q) \rightarrow \text{Mult}(V_1, \dots, V_p, W_1, \dots, W_q)$$

ist bilinear. Es ist daher üblich, für $\text{Mult}(V, W)$ einfach $V^* \otimes W^*$ zu schreiben (dabei bezeichnet V^* wieder den Dualraum) und dies dann das Tensorprodukt von V^* und W^* zu nennen. Iterativ erhält man das Tensorprodukt V_1^*, \dots, V_p^* als

$$V_1^* \otimes \dots \otimes V_p^* := \text{Mult}(V_1, \dots, V_p).$$

Wegen $V_i \cong (V_i^*)^*$ im Endlichdimensionalen setzt man dann

$$V_1 \otimes \dots \otimes V_p := \text{Mult}(V_1^*, \dots, V_p^*).$$

Definition 15.4

Sind sämtliche V_i identisch, so setzen wir

$$\bigotimes^p V^* := \underbrace{V^* \otimes \dots \otimes V^*}_{p\text{-mal}}.$$

Definition 15.5 (symmetrisch, alternierend)

- i) Ein Tensor $\phi \in \bigotimes^p V^*$ heißt **symmetrisch**, wenn

$$\phi(v_{\sigma(1)}, \dots, v_{\sigma(p)}) = \phi(v_1, \dots, v_p) \forall \sigma \in S_p \forall v_1, \dots, v_p \in V$$

- ii) Ein Tensor $\phi \in \bigotimes^p V^*$ heißt **schief-symmetrisch** oder **alternierend**, wenn

$$\phi(v_{\sigma(1)}, \dots, v_{\sigma(p)}) = \text{sign}(\sigma) \phi(v_1, \dots, v_p) \forall \sigma \in S_p \forall v_1, \dots, v_p \in V.$$

Den Raum der symmetrischen p -Linearformen auf V bezeichnen wir mit $\bigvee^p V^*$ und den Raum der schief-symmetrischen mit $\bigwedge^p V^*$. Für $\sigma \in S_p$, $\phi \in \bigotimes^p V^*$ definieren wir $\phi_\sigma \in \bigotimes^p V^*$ durch

$$\phi_\sigma(v_1, \dots, v_p) := \phi(v_{\sigma(1)}, \dots, v_{\sigma(p)}).$$

Wir erhalten nun zwei Operatoren durch die Vorschriften

$$\text{Sym}_p : \bigotimes^p V^* \rightarrow \bigvee^p V^* \subset \bigotimes^p V^*, \text{Sym}_p(\phi) := \frac{1}{p!} \sum_{\sigma \in S_p} \phi_\sigma,$$

$$\text{Alt}_p : \bigotimes^p V^* \rightarrow \bigwedge^p V^* \subset \bigotimes^p V^*, \text{Alt}_p(\phi) := \frac{1}{p!} \sum_{\sigma \in S_p} \text{sign}(\sigma) \phi_\sigma.$$

Definition 15.6 (äußeres Produkt, Dachprodukt)

Die Abbildung

$$\wedge : \bigwedge^p V^* \times \bigwedge^q V^* \rightarrow \bigwedge^{p+q} V^*, (\phi, \psi) \mapsto \phi \wedge \psi := \frac{(p+q)!}{p!q!} \text{Alt}_{p+q}(\phi \otimes \psi)$$

heißt das **äußere Produkt** oder **Dachprodukt** von ϕ und ψ . Analog heißt

$$\vee : \bigvee^p V^* \times \bigvee^q V^* \rightarrow \bigvee^{p+q} V^*, (\phi, \psi) \mapsto \phi \vee \psi := \frac{(p+q)!}{p!q!} \text{Sym}_{p+q}(\phi \otimes \psi)$$

das **symmetrische Produkt** von ϕ und ψ .

15.2 Sätze und Beweise

Satz 15.1 (Dimension des Vektorraums der Multilinearformen)

Es sei m_i die Dimension V_i , $i = 1, \dots, p$. Dann besitzt $\text{Mult}(V_1, \dots, V_p)$ die Dimension $\prod_{i=1}^p m_i$.

Beweis: Wir wählen jeweils Basen $\{e_{i_1}, \dots, e_{i_{m_i}}\}, i = 1, \dots, p$ für V_i . Sind

$$v_i = \sum_{j=1}^{m_i} e_{i_j} \in V_i, i = 1, \dots, p$$

beliebig, so ist

$$\phi(v_1, \dots, v_p) = \sum_{j_1, \dots, j_p} x_1^{j_1} \cdots x_p^{j_p} \phi(e_{1_{j_1}}, \dots, e_{p_{j_p}}).$$

Daher ist ϕ vollständig durch die Koeffizienten $\phi(e_{1_{j_1}}, \dots, e_{p_{j_p}})$ bestimmt. Sind umgekehrt $a_{j_1}, \dots, a_{j_p} \in \mathbb{R}$ gegeben, so ist durch

$$\phi(v_1, \dots, v_p) := \sum_{j_1, \dots, j_p} x_1^{j_1} \cdots x_p^{j_p} a_{j_1} \cdots a_{j_p}$$

eine p -Linearform definiert.

q.e.d.

Satz 15.2 (Eigenschaften des Tensorproduktes)

Für das Tensorprodukt gilt:

$$\text{Distributivität} \quad (\lambda_1 \phi_1 + \lambda_2 \phi_2) \otimes \psi = \lambda_1 \phi_1 \otimes \psi + \lambda_2 \phi_2 \otimes \psi,$$

$$\phi \otimes (\lambda_1 \psi_1 + \lambda_2 \psi_2) = \lambda_1 \phi \otimes \psi_1 + \lambda_2 \phi \otimes \psi_2.$$

$$\text{Assoziativität} \quad (\phi \otimes \psi) \otimes \sigma = \phi \otimes (\psi \otimes \sigma),$$

$$(U \otimes V) \otimes W = U \otimes (V \otimes W).$$

Sind außerdem $(\phi_i)_{i=1,\dots,m}$ und $(\psi_j)_{j=1,\dots,n}$ Basen von $V_1^* \otimes \dots \otimes V_p^*$ beziehungsweise von $W_1^* \otimes \dots \otimes W_q^*$, so ist $(\phi_i \otimes \psi_j)_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$ eine Basis von $V_1^* \otimes \dots \otimes V_p^* \otimes W_1^* \otimes \dots \otimes W_q^*$.

Beweis: Dies folgt direkt aus der Konstruktion des Tensorprodukts.
q.e.d.

Satz 15.3 (Eigenschaften symmetrischer und alternierender Tensoren)

- i) ϕ ist genau dann symmetrisch, wenn $\text{Sym}_p \phi = \phi$ und genau dann alternierend, wenn $\text{Alt}_p \phi = \phi$.
- ii) Sym_p und Alt_p sind Projektionen, das heißt, $\text{Sym}_p^2 = \text{Sym}_p$, $\text{Alt}_p^2 = \text{Alt}_p$.
- iii)

$$\text{Sym}_p(\bigotimes^p V^*) = \bigvee^p V^* \quad , \quad \text{Alt}_p(\bigotimes^p V^*) = \bigwedge^p V^*. \quad (15.1)$$

Beweis: Übung.

q.e.d.

Satz 15.4 (Eigenschaften des Dachproduktes und des symmetrischen Produktes)

- i) Das Dachprodukt ist assoziativ und erfüllt

$$\alpha \wedge \beta = (-1)^{pq} \beta \wedge \alpha \quad \forall \alpha \in \bigwedge^p V^*, \beta \in \bigwedge^q V^*$$

- ii) Das symmetrische Produkt ist assoziativ und kommutativ.

Beweis: Übung.

q.e.d.

Satz 15.5 (Dimension des Raums der alternierenden Tensoren)

Sei V ein Vektorraum der Dimension m . Dann ist $\bigwedge^p V^* = \{0\} \forall p > m$. Ist η_1, \dots, η_m eine Basis von V^* , so ist

$$\{\eta_{i_1} \wedge \dots \wedge \eta_{i_p} : 1 \leq i_1 < \dots < i_p \leq m\}$$

eine Basis von $\bigwedge^p V^*$. Insbesondere ist $\dim(\bigwedge^p V^*) = \binom{m}{p}$. Hierbei sei $\bigwedge^0 V^* := \mathbb{R}$.

Beweis: Es sei e_1, \dots, e_m eine Basis von V . Ist $\phi \in \bigwedge^p V^*$ mit $p > m$, so ist $\phi(e_{i_1}, \dots, e_{i_p}) = 0$, da mindestens zwei der Vektoren $(e_{i_k})_{k=1,\dots,p}$ gleich sind und ϕ schiefsymmetrisch ist. Seien nun $0 \leq p \leq m$ und (η_1, \dots, η_m) die

Dualbasis von (e_1, \dots, e_n) . Da $\text{Alt}_p(\bigotimes^p V^*) = \bigwedge^p V^*$, spannt das Bild der Basis $(\eta_{i_1} \otimes \dots \otimes \eta_{i_p})_{\substack{i_k \in \{1, \dots, m\} \\ \forall k=1, \dots, p}}$ unter Alt_p den Raum $\bigwedge^p V^*$ auf und bildet eine Basis von $\bigwedge^p V^*$. Es ist

$$\text{Alt}_p(\eta_{i_1} \otimes \dots \otimes \eta_{i_p}) = \frac{1}{p!} \eta_{i_1} \wedge \dots \wedge \eta_{i_p}. \quad (15.2)$$

Die rechte Seite hängt bis auf Vorzeichen nicht von der Reihenfolge der Vektoren ab. Daher spannen die $\binom{m}{p}$ Elemente $\eta_{i_1} \wedge \dots \wedge \eta_{i_p}$, $1 \leq i_1 < \dots < i_p \leq m$ den Raum $\bigwedge^p V^*$ auf, und da sie linear unabhängig sind bilden sie auch eine Basis. q.e.d.

15.3 Erklärungen zu den Definitionen

Zur Definition 15.1 der p -Linearformen: Ein Tensor ist also einfach eine Abbildung, die in allen ihren Argumenten linear ist. Dies versteht man am besten anhand einiger schon bekannter Beispiele.

Beispiel 133

- Für $p = 1$ ist $\text{Mult}(V) = V^*$ natürlich einfach der Dualraum von V , denn $\text{Mult}(V)$ besteht ja, genauso wie V^* , gerade aus allen linearen Abbildungen $V \rightarrow \mathbb{R}$.
- Wir betrachten die Determinante von $(n \times n)$ -Matrizen. Diese ist linear in jeder Zeile, aber an sich erstmal keine Multilinearform. Fassen wir jedoch eine Matrix als einen Vektor von Vektoren auf, das heißt

$$A = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

mit $v_i \in \mathbb{R}^n = V$, so ist die Determinante eine Abbildung

$$\det : \mathbb{R}^n \times \dots \times \mathbb{R}^n \rightarrow \mathbb{R},$$

die in jedem Argument linear ist. Also ist die Determinante einer $(n \times n)$ -Matrix ein n -Tensor.

- Wir betrachten nun das anschaulich wichtigste Beispiel.

Ist $\phi : V \times V \rightarrow W$ eine Bilinearform, so ist, wie der Name schon sagt, ϕ ja bilinear. Also ist ϕ ein 2-Tensor. Nun haben wir gesehen, dass sich bezüglich einer gewählten Basis jede Bilinearform als Matrix schreiben lässt. Die Matrizen sind also gerade die 2-Tensoren. In gewissem Sinne kann man also Tensoren als Verallgemeinerung von Matrizen (oder genauer: der von Matrizen induzierten Abbildungen) ansehen.

So wäre dies in Abbildung 15.1 zum Beispiel die Darstellung eines 3-Tensors.

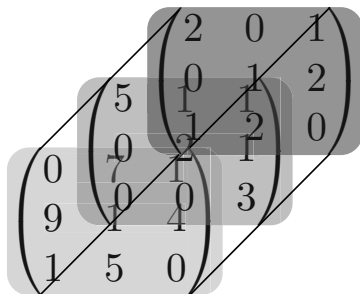


Abb. 15.1: Ein 3-Tensor.

■

Auch wenn sich Tensor nun erstmal kompliziert anhört, haben wir also gesehen, dass sich dahinter eigentlich nur eine Verallgemeinerung von schon Bekanntem versteckt.

Zur Definition 15.2 des Tensorprodukts von Abbildungen: Das Tensorprodukt von Abbildungen dient im Grunde genommen dazu, aus zwei Abbildungen, die linear in jedem Argument sind, eine einzige zu erhalten, die dann in jedem Argument, das heißt in denen der ersten und zweiten Abbildung, wieder linear ist. Hierzu ein Beispiel.

Beispiel 134

Wir wollen einen einfachen Fall betrachten und berechnen das Tensorprodukt für $V_1 = V_2 = W_1 = \mathbb{R}^3$. Es seien also ϕ eine (3×3) -Matrix und ψ ein Element des Raums $(\mathbb{R}^3)^*$. Seien zum Beispiel ϕ gegeben durch die Matrix

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

und ψ die Abbildung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto x_1 + x_2 + x_3.$$

Dann ist

$$\begin{aligned}
 \phi \otimes \psi(v, w, x) &= (v^T A w)(\psi(x)) \\
 &= \begin{pmatrix} 2v_1 + 2v_2 + 3v_3 \\ v_1 + 2v_2 \\ 3v_1 + 2v_2 + v_3 \end{pmatrix}^T \begin{pmatrix} w_1 \\ w_2 \\ w_3 \end{pmatrix} \cdot (x_1 + x_2 + x_3) \\
 &= ((2v_1 + 2v_2 + 3v_3)w_1 + (v_1 + 2v_2)w_2 \\
 &\quad + (3v_1 + 2v_2 + v_3)w_3)(x_1 + x_2 + x_3).
 \end{aligned}$$

Es ist also zum Beispiel

$$\phi \otimes \psi\left(\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right) = (30 + 0 + 20)3 = 150,$$

aber es ist

$$\psi \otimes \phi\left(\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right) = \phi \otimes \psi\left(\begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}\right) = 120.$$

Also ist das Tensorprodukt nicht kommutativ. ■

Das Tensorprodukt ist natürlich per Definition selbst eine bilineare Abbildung.

Zur Definition 15.3 des Tensorprodukts von Vektorräumen: Nun da wir mithilfe des Tensorprodukts von Abbildungen aus zwei Multilinearformen eine neue erhalten haben, wollen wir etwas über den Vektorraum erfahren, in dem die neue Linearform lebt. Dies ist gerade das Tensorprodukt $V \otimes W$ der beiden ursprünglichen Vektorräume. Das Tensorprodukt zweier Vektorräume hat nun eine sehr schöne Eigenschaft.

Haben wir eine bilineare Abbildung $\phi : V \times W \rightarrow \mathbb{R}$, so gibt es genau eine lineare Abbildung $\tilde{\phi} \in V \otimes W$ mit

$$\phi = \tilde{\phi} \circ \otimes,$$

das heißt, das folgende Diagramm kommutiert:

$$\begin{array}{ccc}
 V \times W & \xrightarrow{\phi} & \mathbb{R} \\
 \downarrow \otimes & \nearrow \tilde{\phi} & \\
 V \otimes W & &
 \end{array}$$

Man kann also sagen, dass das Tensorprodukt aus Multilinearformen einfach nur Linearformen macht (über einem anderen Vektorraum).

Zur Definition 15.5 von symmetrischen und alternierenden Tensoren: Die Definition an sich sollte klar sein. Fragen könnte man sich, wozu das dienen soll. Dies kann hier noch nicht klar werden, da man dafür tiefer in die Theorie einsteigen muss. Zwei kleine Gründe können wir aber nennen. Zunächst mal haben wir gesehen, dass Tensoren als Verallgemeinerung von Matrizen angesehen werden können, und auch dort spielen symmetrische und schiefsymmetrische eine besondere Rolle. Der zweite Grund wird sich später, nach Definition des symmetrischen Produkts und des Dachprodukts, zeigen.

Zur Definition 15.6 des symmetrischen und des Dachprodukts: Auch hier sollte die Definition klar sein. Wir betrachten ein Beispiel.

Beispiel 135

Es seien $\eta, \sigma, \tau \in V^* = \bigwedge^1 V^* = \bigotimes^1 V^*$. Dann ist

$$\begin{aligned}\eta \wedge \sigma &= \frac{2!}{1!1!} \text{Alt}_2(\eta \otimes \sigma) = 2\left(\frac{1}{2!}(\eta \otimes \sigma - \sigma \otimes \eta)\right) = \eta \otimes \sigma - \sigma \otimes \eta \\ \eta \wedge \sigma \wedge \tau &= c(\eta \otimes \sigma \otimes \tau - \eta \otimes \tau \otimes \sigma - \sigma \otimes \eta \otimes \tau - \tau \otimes \sigma \otimes \eta + \sigma \otimes \tau \otimes \eta + \tau \otimes \eta \otimes \sigma)\end{aligned}$$

■

Wir erkennen also, dass das Dachprodukt „fast“ kommutativ ist. Und dies ist der zweite kleine Grund für die Einführung dieser neuen Objekte, denn das Tensorprodukt war gar nicht kommutativ.

Eine Anmerkung noch zum Faktor $\frac{(p+q)!}{p!q!}$. Dieser Faktor ist nützlich, um eine schöne Formel für das Dachprodukt von mehr als nur 2 Tensoren zu erhalten.

15.4 Erklärungen zu den Sätzen und Beweisen

Zum Satz 15.1 über die Dimension des Vektorraums der Multilinearformen: Nach diesem Satz können schon aus Dimensionsgründen $V \times W$ und $V \otimes W$ in der Regel nicht isomorph sein. Der Beweis verläuft einfach, indem man aus den Basen der Vektorräume eine Basis des Raums der Multilinearformen beziehungsweise des Tensorprodukts konstruiert.

Als Spezialfall können wir schon bekannte Dimensionen wiederentdecken. Sei dafür $V_i = \mathbb{R}^n$. Dann hat der Dualraum, also der Raum der 1-Tensor Dimension n und der Raum der Matrizen, das heißt der 2-Tensor Dimension n^2 .

Zum Satz 15.2 über die Eigenschaften des Tensorprodukts: Die Eigenschaften hier folgen alle sofort, da wir es nur mit Linearformen zu tun haben. Hier sei nochmal gesagt, dass das Tensorprodukt im Allgemeinen nicht kommutativ ist.

Zur Übung könnt ihr die Eigenschaften einmal direkt für drei Tensoren nachweisen. Nehmt zum Beispiel die Tensoren ϕ und ψ aus unserem obigen Beispiel und dazu noch eine weitere (3×3) -Matrix.

Zum Satz 15.3 über die Eigenschaften symmetrischer und alternierender Tensoren:

Dieser Satz gibt uns eine schöne Charakterisierung von symmetrischen und alternierenden Tensoren. Er bietet außerdem eine sehr gute Übung für euch, die Aussagen zu beweisen. Als Tipp: Den ersten Teil kann man leicht über die Definition beweisen, und für die anderen beiden Teile kann man das schon Bewiesene nutzen.

Zum Satz 15.4 über die Eigenschaften des Dachprodukts und des symmetrischen Produkts:

Hier haben wir nun noch einmal die Eigenschaften des Dachprodukts und die des symmetrischen Produkts zusammengefasst, die man aus dem Beispiel von oben schon erahnen konnte. Auch hier solltet ihr euch wieder an einem Beweis versuchen.

Zum Satz 15.5 über die Dimension des Raums der alternierenden Tensoren:

Wie weiter oben schon für den Raum der Multilinearformen, sagen wir nun noch etwas über die Dimension des Raums der alternierenden Tensoren aus. Dabei benutzen wir im Beweis als Trick die Charakterisierung der alternierenden Tensoren aus Satz 15.3.

Bemerkungen: Wie schon am Anfang angekündigt, haben wir hier Tensoren nur recht oberflächlich behandelt. Je nach Interessen kann man dies noch sehr stark ausweiten, vor allem in der Algebra, der Differentialgeometrie oder auch in der Physik. Dann kann man zum Beispiel Tensorprodukte von Körpern und Ringen betrachten oder Tensorfelder, das heißt Abbildungen, die jedem Punkt in einem Raum einen anderen Tensor zuordnen. Dies findet man dann in der entsprechenden weiterführenden Literatur.

Symbolverzeichnis

$:=$	ist definiert als
$\binom{n}{k}$	Binomialkoeffizient
\cap	Durchschnitt
\cup	Vereinigung
δ_{ij}	Kronecker-Delta
$\dot{\cup}$	disjunkte Vereinigung
\exists	es existiert
$\exists!$	es existiert genau ein
\forall	für alle
$\frac{\partial f}{\partial x_i}$	partielle Ableitung von f nach der Variablen x_i
$\inf(A)$	Infimum von A
$\int_a^b f(x) \, dx$	Integral der Funktion f über das Intervall $[a, b]$
$\langle \cdot, \cdot \rangle$	Skalarprodukt
\liminf	Limes inferior
\limsup	Limes superior
\mathbb{C}	Menge der komplexen Zahlen
\mathbb{N}	Menge der natürlichen Zahlen ohne die Null
\mathbb{N}_0	Menge der natürlichen Zahlen mit der Null
\mathbb{Q}	Menge der rationalen Zahlen
\mathbb{R}	Menge der reellen Zahlen
\mathbb{Z}	Menge der ganzen Zahlen
\mathcal{O}	bezeichnet eine Topologie
\mathring{A}	Innere einer Menge A
\otimes	Tensorprodukt
\overline{A}	Abschluss einer Menge A
\bar{a}	Restklasse von a
\bar{z}	konjugiert komplexe Zahl zu z
∂A	Rand der Menge A
\prod	Produktzeichen
sign	Vorzeichenfunktion
\subset	Teilmenge von
\sum	Summenzeichen
$\sup(A)$	Supremum von A
$\text{grad } f(x_0) = \nabla f(x_0)$	Gradient von f im Punkt x_0
$\text{im}(f)$	Bild von f
$\text{Im}(z)$	Imaginärteil von z
$\ker(f)$	Kern von f

$\text{ord } a$	Ordnung von a
$\text{Re}(z)$	Realteil von z
\vee	Das logische „Oder“
\wedge	Dachprodukt
\wedge	Das logische „Und“
A^*	adjungierte Matrix
$B(x_0, r)$...	abgeschlossene Kugel mit Zentrum x_0
D_n	Dieder-Gruppe
f'	Ableitung von f
$J_d(\lambda)$	Jordan-Block
$J_f(x)$	Jakobi-Matrix von f im Punkt x
$n!$	Fakultät
$U(x_0, r)$...	offene Kugel mit Zentrum x_0
Korollar	..	ist eine Folgerung aus einem Satz.
Lemma	...	ist ein Hilfssatz, den man zum Beweis eines anderen Satzes benötigt.
O.B.d.A.	..	Ohne Beschränkung der Allgemeinheit

Literaturverzeichnis

- [AE06] H. Amann und J. Escher. *Analysis II (Grundstudium Mathematik)*. 2. Aufl. Birkhäuser, März 2006.
- [Ama95] H. Amann. *Gewöhnliche Differentialgleichungen*. 2. Aufl. de Gruyter Lehrbuch, März 1995.
- [Art98] M. Artin. *Algebra*. 1. Aufl. Birkhäuser, Mai 1998.
- [AZ03] M. Aigner und G. M. Ziegler. *Das Buch der Beweise: Buch über die Beweise für mathematische Sätze, z.B. Bertrandsches Postulat, Zwei-Quadrate-Satz von Fermat, Starrheitssatz von Cauchy, Borsuk-Vermutung, Satz von Turan*. 2. Aufl. Springer, Sep. 2003.
- [Bö0] C. Bär. *Elementare Differentialgeometrie*. 1. Aufl. de Gruyter, März 2000.
- [Bar07] Rene Bartsch. *Allgemeine Topologie I*. 1. Aufl. Oldenbourg Verlag, Sep. 2007.
- [Beh07] E. Behrends. *Analysis Band 2: Ein Lernbuch für den sanften Wechsel von der Schule zur Uni*. 2. Aufl. Vieweg und Teubner, Apr. 2007.
- [Beu03] A. Beutelspacher. *Lineare Algebra: Eine Einführung in die Wissenschaft der Vektoren, Abbildungen und Matrizen. Mit liebevollen Erklärungen, einleuchtenden Beispielen ... Nutzen der Studierenden der ersten Semester*. 6. Aufl. Vieweg und Teubner, Sep. 2003.
- [Beu06] A. Beutelspacher. *"Das ist o.B.d.A. trivial!": Eine Gebrauchsanleitung zur Formulierung mathematischer Gedanken mit vielen praktischen Tipps für Studierende der Mathematik und Informatik*. 8. überarb. Auflage. Vieweg und Teubner, Okt. 2006.
- [Bos08] A. Bosch. *Lineare Algebra*. 4. überarbeitete Auflage. Springer, März 2008.
- [Fis08] G. Fischer. *Lineare Algebra: Eine Einführung für Studienanfänger*. 16. überarb. u. erw. Auflage. Vieweg und Teubner, Aug. 2008.
- [For08] O. Forster. *Analysis 2: Differential- und Integralrechnung einer Veränderlichen*. 8. überarbeitete Auflage. Vieweg und Teubner, Aug. 2008.
- [Fri06] K. Fritzsche. *Grundkurs Analysis 2: Differentiation und Integration in einer Veränderlichen*. 1. Aufl. Spektrum Akademischer Verlag, Apr. 2006.

- [Fur95a] P. Furlan. *Das Gelbe Rechenbuch 1: für Ingenieure, Naturwissenschaftler und Mathematiker: Bd 1*. Verlag Martina Furlan, Sep. 1995.
- [Fur95b] P. Furlan. *Das Gelbe Rechenbuch 2: für Ingenieure, Naturwissenschaftler und Mathematiker: Bd 2*. Verlag Martina Furlan, Sep. 1995.
- [Fur95c] P. Furlan. *Das Gelbe Rechenbuch 3: für Ingenieure, Naturwissenschaftler und Mathematiker: Bd 3*. Verlag Martina Furlan, Sep. 1995.
- [Her07] N. Herrmann. *Mathematik ist überall*. 3. korr. Auflage. Oldenbourg, Jan. 2007.
- [Heu08] H. Heuser. *Lehrbuch der Analysis. Teil 2*. 14. durchges. Auflage. Vieweg und Teubner, Mai 2008.
- [Hö] A. Hölzle. *Metrische, normierte und topologische Räume*. URL: <http://www.mathering.de/pdf/SkalNormMetrTopo.pdf> (besucht am 01.11.2010).
- [Kö9] K. Königsberger. *Analysis 2*. 5. Aufl. Springer, Juni 2009.
- [KM08] C. Karpfinger und K. Meyberg. *Algebra*. 1. Aufl. Spektrum Akademischer Verlag, Okt. 2008.
- [MK09] F. Modler und M. Kreh. *Tutorium Analysis 1 und Lineare Algebra 1*. 1. Aufl. Spektrum Akademischer Verlag, Sep. 2009.
- [Sch] S. Schlesi. *Extrema reellwertiger Funktionen mehrerer Veränderlicher*. URL: <http://www.matheplanet.com/matheplanet/nuke/html/article.php?sid=1104> (besucht am 01.11.2010).
- [Sin00] S. Singh. *Fermats letzter Satz: Die abenteuerliche Geschichte eines mathematischen Rätsels*. 2. Aufl. Deutscher Taschenbuch Verlag, März 2000.
- [Tim97] S. Timmann. *Repetitorium der Analysis Teil 2*. 2. Aufl. Binomi Verlag, Juni 1997.
- [Wil98] D. Wille. *Repetitorium der Linearen Algebra Teil 2*. 2. Aufl. Binomi Verlag, Jan. 1998.
- [Woh] M. (alias matroid) Wohlgemuth. *Lineare Algebra für Dummies*. URL: <http://www.mathe-online.at/materialien/matroid/files/lafd1.pdf> (besucht am 01.11.2010).
- [Woh09] M. Wohlgemuth, Hrsg. *Mathematisch für Anfänger*, Spektrum Akademischer Verlag, Okt. 2009.
- [Woh10] M. Wohlgemuth, Hrsg. *Mathematisch für fortgeschrittene Anfänger*, Spektrum Akademischer Verlag, Okt. 2010.

Index

A

Abbildung	
differenzierbar	65
abgeschlossene Kugel	2
abgeschlossene Menge	2, 5, 7
Ableitung	65
partiell	66
Abschluss	3
Adjungierte	174, 175
Alpha-Hölder-stetig	44
Anfangswertaufgabe	118
anisotrop	257

B

Banach'scher Fixpunktsatz	48
für eine stetige Familie	107
Banach-Raum	6
Basis	5
Basiswechsel	181
Basiswechselsatz	177
Bernoulli'sche Differentialgleichung	132
beschränkt	4
Bilinearform	171, 172
symmetrisch	172
Bilinearformen	177
Bogenlänge	134

C

Cauchy-Folge	2
--------------	---

D

Dachprodukt	299
darstellende Matrix	172, 175
Darstellungsmatrix	175
Definitheit	94
Dieder-Gruppe	244
Differential	65
Differentialgleichung	117
Lösungsverfahren	123
differenzierbar	65
differenzierbare Mannigfaltigkeit	159
direkte Summe	173

E

Einheitssphäre	23
endliche Untergruppen	248
Erzeugendensystem	219, 220
euklidische Metrik	16
euklidischer Vektorraum	165, 173
Exponentialabbildung	
Matrizen	280
Exponentialfunktion	6
Extrema unter Nebenbedingungen	95

F

Faktorgruppe	225
Faktoring	228
Faktorstruktur	224, 228
Fixpunkt	45
Fixpunktsatz von Banach	48
Folge	
Cauchy-Folge	2
Grenzwert	2
Häufungspunkt	2
Konvergenz	2
Vollständigkeit	3
folgenkompakt	3
Folgenstetigkeit	44
Fundamentalsystem	118
Funktion mehrerer Veränderlicher	43

G

Geschwindigkeitsvektor	143
gleichmäßig stetig	44
Gradient	66
Gram-Schmidt	178, 181
Grenzwert	2, 43
Gruppenerzeugendensystem	219

H

Häufungspunkt	2
Hauptminorenkriterium	180, 187
Hauptraum	275
Hauptraumzerlegung	277
hausdorff'sche Trennungseigenschaft	8
Heine-Borel	10, 40
Helix	142
hermitesche Form	174
Hesse-Matrix	93
hinreichendes Kriterium	94
höhere Ableitung	67
Homöomorphismus	45
Homomorphiesatz für Gruppen	226
Homomorphiesatz für Ringe	230
Hyperflächen	154

I

Ideal	220
Immersion	154
implizite Funktionen	108
Inneres	3
isotrop	257

J

Jakobi-Matrix	66
Jordan-Block	276
Jordan-Normalform	278

K

Kettenregel	69
Klein'sche Vierergruppe	231
kleiner Satz von Fermat	224
Kodimension	154
kompakt	4
Komplement	173
kongruent	244
konische Spirale	142
Konjugation	219
Kontraktion	45
Konvergenz	2
Kugel	
abgeschlossen	2
offen	2
Kurve	133
Bogenlänge	134
Länge	134, 135, 137, 139
Orientierung	134
proportional parametrisiert	134
regulär parametrisiert	133
Rektifizierbarkeit	134
Spur	134
Umparametrisierung	134

L

Länge einer Kurve	134, 135, 137, 139
Limes	43
Lipschitz-stetig	44
logarithmische Spirale	141
lokales Extremum	93
lokales Koordinatensystem	154

M

Mannigfaltigkeit	159
Maximum	93
Menge	
abgeschlossen	2, 7
offen	2, 6
Metrik	1, 36
metrischer Raum	1
Minimalpolynom	275
Minimalpolynom	
Wohldefiniertheit	276
Minimum	93
Mittelwertungleichung	69
Multilinearformen	
Dimension	299

N

Nebenklasse	222
negativ definit	94
nilpotent	276
Äquivalenzen	277
Norm	5, 36, 173
normal	175
Normalenraum	154
Normalenvektor	154

Normalteiler	224
Normenäquivalenz	6
notwendiges Kriterium	94
Nullraum	118

O

offene Überdeckung	3
offene Kugel	2
offene Menge	2, 5, 6
Operatornorm	45
Ordnung	220
orientierungserhaltend	134
orientierungsumkehrend	134
orthogonale Matrizen	
Klassifikation	244
orthogonale Projektion	191
orthogonales Komplement	173, 191
Orthogonalisierungsverfahren	178
Orthonormalbasis	172

P

p -Form	32
p -Linearform	297
Parameterdarstellung	154
Parametertransformation	134
partielle Ableitung	66
Peano	119
Picard-Lindelöf	119
Polarkoordinaten	56
positiv definit	94, 172, 173
präkompakt	4
Produktmetrik	17
Projektion	174
proportional parametrisiert	134
punktweise stetig	44

Q

Quadrik	258
entartet	258
Klassifikation	264

R

Rand	3
regulär parametrisiert	133
regulärer Punkt	155
regulärer Wert	155
Rektifizierbarkeit	134
Restklassenabbildung	225
Richtungsableitung	66
Ringerzeugendensystem	220

S

Satz	
von Peano	119
von Picard-Lindelöf	119
Satz von Lagrange	223
Schleppkurve	143

selbstadjungiert 174, 175, 258, 262
 Eigenvektor 262
 Sesquilinearform 193
 Signatur 257
 singulärer Punkt 144, 155
 singulärer Wert 155
 Skalarprodukt 36
 Spektralsatz 182, 264
 Sphäre 23, 158
 Spur 134
 Standardskalarprodukt 165, 166
 stereographische Projektion 158
 stetig differenzierbar 65
 Stetigkeit 44, 45
 Stetigkeitskriterium 46
 Subbasis 5
 Summe 173
 Supremumsnorm 32
 Symmetrien
 endliche Untergruppen 248
 Klassifikation 246
 Symmetrieoperation 243
 symmetrische Bilinearform 172

T

Tangentialraum 154
 Tangentialvektor 143, 154
 Taylor'sche Formel 70
 Tensor 297
 alternierend 300
 alternierend 298
 symmetrisch 298
 Tensorprodukt
 Eigenschaften 299
 von Abbildungen 297
 von Vektorräumen 298
 Topologie 4, 36
 topologische Mannigfaltigkeit 159
 topologischer Raum 4
 Torus 158
 total beschränkt 4
 total differenzierbar 65
 Trägheitssatz von Sylvester 261
 Traktrix 143
 Trennungseigenschaft 8
 Trigonalisierung 277

U

überdeckungskompakt 3
 Umgebung 2, 5
 Umparametrisierung 134
 unitär 175
 unitärer Raum 175
 Untergruppen von \mathbb{Z} 222
 Untermannigfaltigkeit 154

V

Vektor

Länge 165
 Winkel 166
 Vektorraum 165
 Vollständigkeit 3

W

Winkel 166

Z

zusammenhängend 5
 zyklische Gruppe 220
 Klassifikation 227