

Grundlagen der Mathematik

Thomas Markwig
Fachbereich Mathematik
Technische Universität Kaiserslautern

Vorlesungsskript

Sommer- und Wintersemester 2010/11

Inhaltsverzeichnis

Einleitung	1
Kapitel I Grundlegende Begriffsbildungen	3
§ 1 Etwas Logik	3
§ 2 Mengen	14
§ 3 Abbildungen	18
§ 4 Vollständige Induktion	25
§ 5 Mächtigkeit von Mengen	27
§ 6 Äquivalenzrelationen	34
§ 7 Gruppen und Körper	41
§ 8 Ordnungsrelationen	52
§ 9 Eigenschaften der reellen Zahlen \mathbb{R}	59
§ 10 Der Körper der komplexen Zahlen	67
Kapitel II Eindimensionale Analysis	73
§ 11 Folgen und ihre Grenzwerte	73
§ 12 Unendliche Reihen	89
§ 13 Grenzwerte von Funktionen	111
§ 14 Stetigkeit	121
§ 15 Konvergenz von Funktionenfolgen	134
§ 16 Exponentialfunktion, Logarithmus, trigonometrische Funktionen	138
§ 17 Differenzierbarkeit	153
§ 18 Der Mittelwertsatz und seine Anwendungen	163
§ 19 Das Riemann-Integral	189
§ 20 Hauptsatz der Differential- und Integralrechnung mit Anwendungen	206
§ 21 Uneigentliche Integrale	222
Kapitel III Lineare Algebra	227
§ 22 Rechnen mit Matrizen	227
§ 23 Vektorräume und lineare Abbildungen	233
§ 24 Basen von Vektorräumen	254

§ 25	Endlich-dimensionale Vektorräume	268
§ 26	Lineare Abbildungen und Matrizen	278
§ 27	Der Gauß-Algorithmus	292
§ 28	Lineare Gleichungssysteme	306
§ 29	Die symmetrische Gruppe	327
§ 30	Die Determinante	332
§ 31	Der Polynomring $K[t]$	352
§ 32	Endomorphismen und ihre Eigenwerte	360
§ 33	Diagonalisierbarkeit und Trigonalisierbarkeit	374
§ 34	Die Jordansche Normalform	392
§ 35	Lineare Algebra mit SINGULAR	413
§ 36	Bilinearformen und Sesquilinearformen	425
§ 37	Euklidische und unitäre Räume	440
§ 38	Spektralsatz und Hauptachsentransformation	456
Kapitel IV Mehrdimensionale Analysis		483
§ 39	Topologische Grundbegriffe in metrischen und in normierten Räumen	483
§ 40	Stetigkeit in metrischen und in normierten Räumen	507
§ 41	Differenzierbarkeit im \mathbb{R}^n	527
§ 42	Der Satz von Taylor und seine Anwendungen	544
§ 43	Der Satz über implizite Funktionen	562
§ 44	Das Riemann-Integral über n -dimensionalen Quadern	590
§ 45	Das Lebesguesche Integrabilitätskriterium	613
§ 46	Das Riemann-Integral über Jordan-meßbaren Mengen	625
§ 47	Der Transformationssatz für Integrale	643
Literaturverzeichnis		665

Einleitung

Die vorliegende Ausarbeitung zu den Vorlesungen Grundlagen der Mathematik 1 und 2 im Sommersemester 2010 und Wintersemester 2010/11 wird im wesentlichen wiedergeben, was während der Vorlesung an die Tafel geschrieben wird. Einige wenige Abschnitte werden ausführlicher sein. Die Ausarbeitung ersetzt somit in keiner Weise ein Lehrbuch oder ein Vorlesungsskript, wie es etwa von Andreas Gathmann oder Wolfram Decker in den vergangenen Semestern erstellt wurde.

KAPITEL I

Grundlegende Begriffsbildungen

Wir beginnen damit, grundlegende Begriffe einzuführen und zu besprechen, die für alle mathematischen Disziplinen gleich wichtig sind.

§ 1 Etwas Logik

Wie alle Wissenschaftler versuchen auch die Mathematiker *Aussagen* über die Objekte ihrer Forschungsarbeit aufzustellen und *als wahr nachzuweisen*. Anders aber als etwa in den Naturwissenschaften werden die zu untersuchenden Objekte nicht von außen an die Mathematiker herangetragen, vielmehr schaffen sie sie sich selbst durch die Vorgabe sogenannter *Axiome*. Wie hat man dies zu verstehen? Was ist ein Axiom? Was heißt es, eine Aussage als wahr nachzuweisen? Und was eigentlich ist eine Aussage?

Nun, sobald wir uns auf eine Sprache geeinigt haben, in der wir uns verständigen wollen, sind wir in der Lage, Sätze zu bilden, Sätze, wie etwa (in unserer Alltagssprache)

“Dieser Satz enthält fünf Worte.”

oder

“Löse die folgende Aufgabe.”

Ein solcher Satz stellt eine *Aussage* in unserem Sinne dar, wenn wir entscheiden können, ob er wahr oder falsch ist. Gemäß dieser Konvention ist der erste der obigen Sätze eine – wahre – Aussage, während beim zweiten Satz, einer Aufforderung, die Frage nach wahr oder falsch wenig Sinn ergibt. Er ist mithin keine Aussage. Wir halten fest:

Aussagen erkennen wir daran, daß ihnen ein Wahrheitswert zugeordnet ist, **w** für *wahr* oder **f** für *falsch*.

Im folgenden werden wir als Platzhalter für Aussagen meist Großbuchstaben verwenden: A, B, C, \dots

Eine Aussage als *wahr nachzuweisen*, soll bedeuten, daß wir sie durch logische Schlüsse auf andere, uns als wahr bekannte Aussagen zurückführen. Nehmen wir etwa den folgenden Satz:

A : Der Bundespräsident ist stets mindestens vierzig Jahre alt.

Wir stellen zunächst einmal fest, daß es sich um eine Aussage handelt – und zwar um eine *wahre* Aussage, wie wir aus Artikel 54 des Grundgesetzes ableiten. Dort nämlich finden wir zur Wahl des Bundespräsidenten folgende Aussage:

B : Wählbar ist jeder Deutsche, der das Wahlrecht zum Bundestage besitzt und das vierzigste Lebensjahr vollendet hat.

Weil nun das Grundgesetz gültig ist, ist Aussage A wahr. Wir haben Aussage A also auf eine uns bekannte wahre Aussage zurückgeführt.

Daß die von uns aus dem Grundgesetz zitierte Aussage B ihrerseits wahr ist, läßt sich nicht weiter auf andere Aussagen zurückführen. Vielmehr handelt es sich hierbei um eine Festlegung des Gesetzgebers, der das Gesetz erlassen und damit diese Aussage für wahr erklärt hat.

Eine Aussage, der der Wahrheitswert \mathbf{w} schlicht durch Festlegung zugewiesen wurde, nennen wir ein *Axiom*.

Man kann in diesem Sinne das Grundgesetz als eine Sammlung von Axiomen, oder ein Axiomensystem, auffassen – auch wenn der Vergleich in mancher Hinsicht hinken mag.

Eingangs haben wir erklärt, daß die Mathematiker sich die Welt, die sie untersuchen, und ihre Objekte selbst erschaffen. Sie tun dies, indem sie sich einige wenige Aussagen als Axiome vorgeben und sodann studieren, was sich aus diesen durch logisch korrekte Schlüsse ableiten läßt. Freilich, so wie der Gesetzgeber seine Gesetze nicht willkürlich erläßt, so wählen auch die Mathematiker die Axiome, die sie sich vorgeben, mit Bedacht, das heißt, mit dem Ziel, interessante Strukturen zu gewinnen – und die vielfältigen Anwendungen zeigen, daß die Mathematiker bei diesem Vorhaben nicht nur sehr kreativ, sondern auch sehr erfolgreich gewesen sind. Immer wieder haben sie sich von Fragestellungen der Alltagswelt inspirieren lassen, haben die Probleme auf wenige Kernpunkte reduziert und in ein (mathematisches) *Modell* übersetzt. Dabei bedeutet letzteres nichts anderes, als daß man die zu benutzende Sprache und die geltenden Axiome festlegt und daß man die Fragen in dieser neuen Sprache formuliert. Die Stärke dieser *Modellbildung* besteht nun darin, daß man innerhalb des Modells exakt und ohne Wenn und Aber feststellen kann, ob eine Aussage wahr ist oder nicht. Wahr ist sie stets dann, wenn sie durch eine ganze Reihe logisch korrekter Schlüsse aus den vorgegebenen Axiomen hervorgeht. Wann aber ist denn eine Aussage aus einer anderen durch einen *logisch korrekten Schluß* hervorgegangen?

Bevor wir uns dieser Frage erneut zuwenden, wollen wir klären, wie man aus gegebenen Aussagen überhaupt neue Aussagen gewinnen und so das Arsenal an Aussagen erweitern kann.

Eine ganz natürliche Möglichkeit ist die Verneinung oder *Negation* einer Aussage, etwa

$\neg A$: Der Bundespräsident ist *nicht* stets vierzig Jahre alt.

Wir wollen generell die Negation einer Aussage X mit dem Symbol $\neg X$ bezeichnen, und es sollte gelten, wenn X wahr ist, so ist $\neg X$ falsch, und umgekehrt. Das heißt insbesondere, der Wahrheitswert von $\neg X$ hängt nur vom Wahrheitswert von X ab. Dies erlaubt es uns, den Wahrheitswert von $\neg X$ in Abhängigkeit des Wahrheitswertes von X in einer Tabelle festzuhalten:

X	$\neg X$
w	f
f	w

Aus unserer Alltagssprache sind wir es gewohnt, mehrere Aussagen in auflistender Weise durch das Wort “und” miteinander zu verbinden. Betrachten wir etwa die folgenden Aussagen

C : Wählbar sind nur Deutsche, die das Wahlrecht zum Bundestag besitzen.

sowie

D : Wählbar sind nur Deutsche, die das vierzigste Lebensjahr vollendet haben.

Man erkennt unschwer, daß die Verknüpfung der Aussagen C und D durch “und” inhaltlich mit unserer obigen Aussage B übereinstimmt, und man spricht von der *Konjunktion* von C und D . Auch hier wollen wir wieder eine symbolische Schreibweise einführen. Sind X und Y zwei Aussagen, so schreiben wir für “ X und Y ” auch $X \wedge Y$. Wenn nun $X \wedge Y$ wieder eine Aussage ist, so muß ihr auch ein Wahrheitswert zugeordnet sein. Dabei sollte wohl $X \wedge Y$ nur dann wahr sein, wenn sowohl X als auch Y wahr sind. Wir können den Wahrheitswert von $X \wedge Y$ also wieder in Abhängigkeit von den Wahrheitswerten von X und Y in einer Tabelle, auch *Wahrheitstafel* genannt, festhalten.

X	Y	$X \wedge Y$
w	w	w
w	f	f
f	w	f
f	f	f

Ebenso ist uns aus unserem alltäglichen Gebrauch ein weiteres Bindewort bekannt, “oder”, welches wir hier instrumentalisieren wollen. Sind X und Y wieder Aussagen, so werden wir gewöhnlich $X \vee Y$ statt “ X oder Y ” schreiben. Die so entstandene neue Aussage nennt man die *Disjunktion* von X und Y , und damit sie wahr ist, soll es uns reichen, daß eine der Aussagen X und Y wahr ist. Dies führt zur folgenden

Wahrheitstafel:

X	Y	$X \vee Y$
w	w	w
w	f	w
f	w	w
f	f	f

Man beachte, daß *oder* hier nicht das ausschließende *entweder oder* ist!

Die Aussage etwa, daß die Kinder unserer Bundestagsabgeordneten stets die deutsche *oder* eine andere Staatsangehörigkeit haben, ist wahr, weil sie nicht ausschließt, daß sie die deutsche und eine andere Staatsangehörigkeit haben.

Im Absatz zur Konjunktion heißt es, daß die Aussage **B** mit der Konjunktion der Aussagen **C** und **D** inhaltlich übereinstimme. Sprachlich sind beide Aussagen aber deutlich verschieden. Anstatt sie *gleich* zu nennen, wollen wir deshalb nur davon sprechen, daß **B** und $C \wedge D$ *gleichwertig* oder *äquivalent* sind. Dies soll zum Ausdruck bringen, daß sie den gleichen Wahrheitswert besitzen. Gehen wir einen Schritt weiter, so können wir eine neue Verknüpfung zweier Aussagen **X** und **Y** einführen, die *Äquivalenz* von **X** und **Y**, in Symbolen $X \Leftrightarrow Y$. Sie soll genau dann wahr sein, wenn **X** und **Y** den gleichen Wahrheitswert besitzen. Dies führt zu folgender Wahrheitstafel:

X	Y	$X \Leftrightarrow Y$
w	w	w
w	f	f
f	w	f
f	f	w

Ein kurzer Blick auf die bislang eingeführten Operationen zur Gewinnung neuer Aussagen aus gegebenen zeigt, daß die Wahrheitswerte der neuen Aussagen stets allein von den Wahrheitswerten der gegebenen Aussagen abhängen, und nicht von deren konkretem Inhalt.

Wir erlauben uns deshalb, eine letzte Verknüpfung von Aussagen, die *Implikation*, dadurch einzuführen, daß wir bei gegebenen Aussagen **X** und **Y** den Wahrheitswert der Aussage “**X** impliziert **Y**” oder “wenn **X**, dann **Y**”, in Zeichen $X \Rightarrow Y$, festlegen:

X	Y	$X \Rightarrow Y$
w	w	w
w	f	f
f	w	w
f	f	w

(1)

Die Wortwahl legt nahe, daß die Aussage $X \Rightarrow Y$ es erlaubt, aus der Wahrheit von **X** Rückschlüsse auf die Wahrheit von **Y** zu ziehen. Dies kommt auch in den ersten beiden Zeilen der Wahrheitstafel zum Ausdruck, wird aber noch deutlicher, wenn wir zeigen, daß die Aussagen $X \Rightarrow Y$ und $\neg X \vee Y$ zueinander äquivalent sind. Ist

dann nämlich X wahr, so ist $\neg X$ falsch. Damit $\neg X \vee Y$ wahr sein kann, muß mithin Y wahr sein. Dies läßt sich so interpretieren, daß sich bei wahrer Aussage X und korrekter Implikation $X \Rightarrow Y$ für Y nur die Möglichkeit ergibt, ebenfalls wahr zu sein.

In dieser Weise werden wir die Implikation immer wieder anwenden. Wir werden mit einer wahren Aussage starten und mittels einer logisch korrekten Argumentationskette Y aus X ableiten – sprich wir werden $X \Rightarrow Y$ als wahr erweisen. Damit haben wir dann zugleich die Wahrheit von Y bewiesen.

Die Gültigkeit der behaupteten Äquivalenz leiten wir durch eine Betrachtung der Wahrheitstabellen her. Es reicht, festzustellen, daß die Werte in den Spalten von $X \Rightarrow Y$ und von $\neg X \vee Y$ übereinstimmen:

X	Y	$\neg X$	$\neg X \vee Y$	$X \Rightarrow Y$
w	w	f	w	w
w	f	f	f	f
f	w	w	w	w
f	f	w	w	w

Die bisherigen Betrachtungen erläutern die ersten beiden Zeilen der Wahrheitstafel der Implikation. Mysteriöser sind auf den ersten Blick zweifellos die beiden letzten, erlauben sie es doch, aus einer falschen Aussage eine beliebige andere Aussage herzuleiten und den vorgenommenen Schluß als korrekt anzusehen. Widerstrebt uns das nicht zutiefst? Wir möchten an einem Beispiel, das auf ein wenig Schulwissen in Mathematik zurückgreift, verdeutlichen, daß die obige Festlegung sehr wohl Sinn macht. Will man etwa die Lösungen der Gleichung

$$x^2 - 2x = -1$$

finden, so wird man auf beiden Seiten der Gleichung zunächst die Zahl 1 addieren, um so auf der linken Seite den Ausdruck $(x-1)^2$ zu erhalten, ein Verfahren, welches als *quadratische Ergänzung* bekannt ist. Man leitet aus der Aussage $x^2 - 2x = -1$ die Aussage $x^2 - 2x + 1 = 0$ her. Dieser Schluß läßt sich formulieren als die Implikation

$$(x^2 - 2x = -1) \Rightarrow (x^2 - 2x + 1 = 0).$$

Der Schluß, daß die Addition einer Zahl auf beiden Seiten einer Gleichung, die Gleichheit nicht zerstört, ist uns wohl vertraut und wir sehen ihn als korrekt an, unabhängig davon, was auf beiden Seiten der Gleichung steht. Wenden wir diesen Schluß nun auf eine andere Gleichung an, etwa auf die Gleichung $0 = 1$, so erhalten wir die Implikation

$$(0 = 1) \Rightarrow (0 + 1 = 1 + 1).$$

Die beiden Aussagen links und rechts des Implikationspfeiles sind offenbar falsch, der Schluß an sich ist jedoch nach dem eben Gesagten zulässig. Mithin sollte die Implikation den Wahrheitswert **w** tragen.

Ein Beispiel dafür, daß sich aus einer falschen Aussage durch einen korrekten Schluß auch eine wahre Aussage herleiten läßt, erhalten wir in analoger Weise, wenn wir uns vergegenwärtigen, daß die Gleichheit auch durch Multiplikation mit einer Zahl nicht zerstört wird. Dies führt dann zu der wahren Implikation

$$(0 = 1) \Rightarrow (0 \cdot 0 = 1 \cdot 0),$$

bei der die Aussage auf der linken Seite des Implikationspfeiles falsch ist, während die auf der rechten Seite wahr ist.

Wir halten fest:

Der Wahrheitswert der Implikation $X \Rightarrow Y$ bewertet nur die Korrektheit des Schließens, nicht jedoch die Wahrheit der Aussagen X und Y .

Es sei deshalb jedem ans Herz gelegt, die Voraussetzungen, auf die er seine Aussagen gründet, genauestens auf ihren Wahrheitsgehalt zu prüfen! Sonst nützt auch noch so sauberes Schließen gar nichts.

Wir wollen den eingeführten Begriffsapparat nun an zwei Beispielen testen, die uns einige wichtige Erkenntnisse liefern werden.

Beispiel 1.1

Es seien X und Y zwei Aussagen.

- a. Wir haben bereits bei der Definition der Äquivalenz davon gesprochen, daß $X \Leftrightarrow Y$ bedeuten solle, daß “ X genau dann wahr ist, wenn Y wahr ist”. Dies wollte verkürzt ausdrücken, “wenn X , dann Y ” und “wenn Y , dann X ”. Wir behaupten deshalb, daß die Aussagen “ $X \Leftrightarrow Y$ ” und “ $(X \Rightarrow Y) \wedge (Y \Rightarrow X)$ ” äquivalent sind, mit anderen Worten, die Aussagen X und Y sind genau dann äquivalent, wenn Y aus X folgt und umgekehrt.

Diese Tatsache werden wir immer wieder verwenden, wenn wir die Äquivalenz zweier Aussagen beweisen wollen. Ihre Gültigkeit leiten wir wieder durch eine Betrachtung der Wahrheitstabellen her.

X	Y	$X \Rightarrow Y$	$Y \Rightarrow X$	$(X \Rightarrow Y) \wedge (Y \Rightarrow X)$	$X \Leftrightarrow Y$
w	w	w	w	w	w
w	f	f	w	f	f
f	w	w	f	f	f
f	f	w	w	w	w

- b. Die Aussagen " $X \Rightarrow Y$ " und " $\neg Y \Rightarrow \neg X$ " sind ebenfalls äquivalent, wie die folgende Tabelle zeigt:

X	Y	$\neg X$	$\neg Y$	$X \Rightarrow Y$	$\neg Y \Rightarrow \neg X$
w	w	f	f	w	w
w	f	f	w	f	f
f	w	w	f	w	w
f	f	w	w	w	w

Man nennt diese Äquivalenz auch *Kontraposition*. Will man also zeigen, daß eine Aussage X eine Aussage Y impliziert, so kann man statt dessen beide Aussagen verneinen und zeigen, daß aus $\neg Y$ die Aussage $\neg X$ folgt.

□

Kehren wir nun zu der Frage zurück, wann eine Aussage Y aus einer Aussage X durch einen *logisch korrekten Schluß* hervorgegangen ist. Bedeutet dies nur, daß $X \Rightarrow Y$ den Wahrheitswert **w** besitzt? Ja ... und nein! Ist X wahr und hat die Implikation $X \Rightarrow Y$ den Wahrheitswert **w**, so folgt unmittelbar, daß Y wahr ist. In diesem Sinne gilt die Antwort *ja*. Aber damit haben wir das Problem nur verlagert, da die Frage bleibt, wie wir prüfen, ob $X \Rightarrow Y$ denn wahr ist, ohne den Wahrheitswert von Y zu kennen. Wir haben bereits weiter oben – sehr vage – angedeutet, daß wir hierzu meist eine Kette von logisch korrekten und in sich schlüssigen Argumenten verwenden, und viel deutlicher wollen wir hier auch nicht werden. Im Verlauf der folgenden Kapitel werden wir viele Beispiele dafür sehen, wie eine Implikation durch eine Reihe von Argumenten bewiesen – oder besser untermauert – wird; und es wird sicher immer wieder vorkommen, daß Euch diese auf den ersten Blick *nicht* wirklich schlüssig vorkommen, daß es eines genaueren Hinsehens und vielleicht auch der Ergänzung einiger Argumente bedarf, bis Ihr der Kette das Prädikat *logisch korrekt und in sich schlüssig* verleihen wollt. Und das ist eine wichtige Erkenntnis, ob ein Schluß als logisch korrekt erkannt wird, hängt vom Betrachter ab. Und deshalb ist die Frage, ob ein Schluß logisch korrekt ist, weit mehr als nur die Frage, ob $X \Rightarrow Y$ wahr ist.

Beispiel 1.2

Hier nun einige mathematische Aussagen.

- Jede gerade Zahl ist Summe zweier ungerader Zahlen.
- Es gibt unendlich viele Primzahlen.
- Jede gerade Zahl größer zwei ist Summe zweier Primzahlen.
- Zu jedem Kreis läßt sich, nur mit Zirkel und Lineal, ein Quadrat konstruieren, das den gleichen Flächeninhalt hat.
- Die Gleichung $x^n + y^n = z^n$ besitzt für $n > 2$ keine Lösung mit positiven ganzen Zahlen x, y, z .

- F. Gegeben sei eine Ansammlung nicht-leerer Mengen. Dann läßt sich aus jeder der Mengen ein Element auswählen.

Die Aussage A ist offensichtlich wahr, und auch die Aussage B ist richtig, allerdings ist dies keine triviale Aussage. Sie muß bewiesen werden. Die Aussage C ist die bekannte *Goldbachsche Vermutung* aus dem Jahre 1742. Sie ist bis heute weder bewiesen noch widerlegt.

Die Aussage D ist unter dem Begriff *Quadratur des Kreises* bekannt. Sie ist falsch, was sich daraus ableiten läßt, daß die Kreiszahl π transzendent ist (Lindemann 1882). Umgangssprachlich sollte man also die Quadratur des Kreises nicht als Synonym für etwas extrem Schwieriges verwenden, sondern für etwas Unmögliches.

Die Aussage E hat jahrhundertlang als *Fermatsche Vermutung* die Mathematiker beschäftigt. Sie wurde erst 1995 von dem englischen Mathematiker Andrew Wiles als wahr nachgewiesen. Für den Beweis wurden modernste und tiefste mathematische Methoden verwendet.

Die Aussage F, möchte man meinen, ist offensichtlich wahr, eher noch als Aussage A. In gewissem Sinne ist diese Aussage jedoch weder beweisbar noch widerlegbar. Sie ist im Axiomensystem der Mengenlehre von Zermelo und Fraenkel unabhängig von den anderen Axiomen. In der Tat kann man die Aussage F, die als *Auswahlaxiom* bezeichnet wird, als Axiom der Mengenlehre zulassen (was wir, wie die überwiegende Zahl der Mathematiker, tun wollen) oder auch nicht. Da das Auswahlaxiom, wenn überhaupt, so nur für sogenannte überabzählbare Ansammlungen strittig ist, sind Zustimmung oder Ablehnung in dieser Vorlesung kaum von praktischer Relevanz. \square

Wir wollen nun der besseren Übersichtlichkeit halber in einer Bemerkung zusammenfassen, was wir bisher gelernt haben.

Bemerkung 1.3

- Eine *Aussage* ist eine Äußerung, der eindeutig ein Wahrheitswert wahr (**w**) oder falsch (**f**) zugeordnet ist.
- Aus Aussagen X und Y können wir durch Anwenden *logischer Operatoren* neue Aussagen bilden:

Symbol	Bedeutung	Bezeichnung	Alternative Beschreibung
$\neg X$	nicht X	<i>Negation</i>	
$X \vee Y$	X oder Y	<i>Disjunktion</i>	
$X \wedge Y$	X und Y	<i>Konjunktion</i>	
$X \Rightarrow Y$	aus X folgt Y	<i>Implikation</i>	$(\neg X) \vee Y$
$X \Leftrightarrow Y$	genau dann X, wenn Y	<i>Äquivalenz</i>	$(X \Rightarrow Y) \wedge (Y \Rightarrow X)$

Neben Aussagen, die wahr oder falsch sein können, sind *Aussageformen* oder *Prädikate* wichtig.

Eine *Aussageform* ist eine Äußerung, die eine oder mehrere Variablen enthält und zu einer Aussage (d.h. wahr oder falsch) wird, wenn man zulässige Werte für diese Variablen einsetzt.

So ist etwa

$$a > b$$

eine Aussageform, die von den Variablen a und b abhängt, für die wir die ganzen Zahlen als zulässige Werte ansehen wollen. Setzen wir konkrete Werte ein, so entsteht eine Aussage, die wahr sein kann (z.B. für $a = 42$ und $b = 37$) oder falsch (z.B. für $a = 2$ und $b = 4$).

Aussageformen werden in der Praxis häufig mit *Quantoren* gebraucht:

- \forall : “für alle”.
- \exists : “es existiert ein”.
- \exists_1 : “es existiert genau ein”.
- \nexists : “es existiert kein”.

Ist P eine Aussageform, die von einer Variablen x abhängt, so bedeutet:

- $\forall x : P(x)$: “für alle x gilt $P(x)$ ”,
- $\exists x : P(x)$: “es gibt ein x , so daß $P(x)$ gilt”.

Mit Hilfe der Quantoren haben wir aus den Aussageformen neue Aussagen gebildet.

Beispiel 1.4

$$\forall x, \forall y, \forall z, \forall n : n > 2 \Rightarrow x^n + y^n \neq z^n.$$

Dies ist für positive natürliche Zahlen x , y , z und n die in Beispiel 1.2 formulierte Fermatsche Vermutung. □

Wichtig ist das richtige Verneinen einer Aussage.

$$\neg(\forall x : P(x)) \Leftrightarrow \exists x : (\neg P(x)).$$

Die Verneinung der Aussage “für alle x gilt die Aussage $P(x)$ ” ist gleichbedeutend mit “es gibt ein x , für das die Aussage $P(x)$ nicht gilt”.

$$\neg(\exists x : P(x)) \Leftrightarrow \forall x : (\neg P(x)).$$

Die Verneinung der Aussage “es gibt ein x , für das die Aussage $P(x)$ gilt” ist gleichbedeutend mit “für alle x gilt die Aussage $P(x)$ nicht” bzw. mit “für kein x gilt die Aussage $P(x)$ ”.

$$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A).$$

Die Aussage “aus A folgt B ” ist gleichbedeutend mit “aus nicht B folgt nicht A ”. Letzteres bezeichnet man auch als *Kontraposition* von ersterem.

Proposition 1.5

Es seien X, Y und Z Aussagen.

- a. Assoziativgesetze
- $(X \vee Y) \vee Z \iff X \vee (Y \vee Z).$
 - $(X \wedge Y) \wedge Z \iff X \wedge (Y \wedge Z).$
- b. Kommutativgesetze
- $X \vee Y \iff Y \vee X.$
 - $X \wedge Y \iff Y \wedge X.$
- c. Distributivgesetze
- $X \wedge (Y \vee Z) \iff (X \wedge Y) \vee (X \wedge Z).$
 - $X \vee (Y \wedge Z) \iff (X \vee Y) \wedge (X \vee Z).$

Beweis: Den Nachweis der Äquivalenzen überlassen wir dem Leser als Übungsaufgabe. □

Bemerkung 1.6 (Griechisches Alphabet)

Es hat sich in der Mathematik eingebürgert, neben den lateinischen auch griechische Buchstaben zu verwenden, um Objekte und Variablen zu bezeichnen, und das werden wir immer wieder mal tun. Deshalb füge ich hier das griechische Alphabet an:

A α Alpha	B β Beta	Γ γ Gamma	Δ δ Delta	E ϵ ε Epsilon	Z ζ Zeta	H η Eta	Θ θ ϑ Theta
I ι Iota	K κ Kappa	Λ λ Lambda	M μ My	N ν Ny	Ξ ξ Xi	O \omicron Omikron	Π π Pi
P ρ Rho	Σ σ Sigma	T τ Tau	Y υ Ypsilon	Φ ϕ φ Phi	X χ Chi	Ψ ψ Psi	Ω ω Omega

Aufgaben**Aufgabe 1.7**

- a. Negiere die folgenden Aussagen:
- (i) Jedes Auto, das am Samstag um 9:00 auf dem Parkplatz parkte, war rot.
 - (ii) Mindestens ein Auto, das am Samstag um 9:00 auf dem Parkplatz parkte, war rot.
 - (iii) Am Samstag um 9:00 parkten rote Autos auf dem Parkplatz.
 - (iv) Es gibt keine größte ganze Zahl.
 - (v) Keine Regel ohne Ausnahme.

Warum ist das Sprichwort „Keine Regel ohne Ausnahme“ in sich widersprüchlich?

- b. Beweise oder widerlege Aussage (iv).

Aufgabe 1.8

Es seien X und Y Aussagen. Zeige die folgenden Äquivalenzen:

- a. *De Morgansche Regeln*
- $\neg(X \vee Y) \iff \neg X \wedge \neg Y.$
 - $\neg(X \wedge Y) \iff \neg X \vee \neg Y.$
- b. $(\neg X \implies f) \iff X.$

Aufgabe 1.9

- a. Drücke die folgenden Aussagen in Worten aus und, falls eine Aussage falsch sein sollte, ersetze sie dabei durch ihre Negation.
- (i) $\forall m \in \mathbb{N}, \exists n \in \mathbb{N} : m = n + n,$
 - (ii) $\exists m \in \mathbb{N}, \exists n \in \mathbb{N} : (m \neq n) \wedge (m^n = n^m).$
- b. Drücke die folgenden Aussagen in Symbolen aus:
- (i) Zwischen je zwei verschiedenen reellen Zahlen gibt es eine weitere reelle Zahl.
 - (ii) Es gibt keine größte Primzahl in den natürlichen Zahlen.

Aufgabe 1.10

Welche der folgenden Schlußfolgerungen ist korrekt?

- a. Falls es anfängt zu regnen, wird die Straße naß. Aber, da die Straße nicht naß werden wird, wird es auch nicht regnen.
- b. Einige Politiker sind ehrlich. Einige Frauen sind Politiker. Also sind einige weibliche Politiker ehrlich.

Aufgabe 1.11

Drücke die folgende Aussage in Worten aus:

$$\forall m \in \mathbb{N}, \forall n \in \mathbb{N} : m \geq n \implies \exists l \in \mathbb{N} : m = n + l.$$

Aufgabe 1.12 a. Negiere die folgenden Aussagen:

- (i) Zu jedem Vorschlag gibt es jemanden, der den Vorschlag kritisiert.
 - (ii) In manchen Häusern haben nicht alle Wohnungen fließendes Wasser.
- b. Beweise oder widerlege die folgenden Aussagen:
- (i) Jede ganze Zahl ist ein Vielfaches von drei.
 - (ii) Die Summe von je zwei ungeraden Zahlen ist gerade.

§ 2 Mengen

Definitionsversuch 2.1 (Georg Cantor)

Eine *Menge* ist eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen. Die in einer Menge zusammengefaßten Objekte nennen wir die *Elemente* der Menge.

Notation 2.2

a. Mengen angeben durch Auflisten der Elemente:

$$\text{z.B. } \{1, 2, 5, 3, 4, 0\}$$

b. Mengen angeben durch Vorschreiben einer Eigenschaft:

$$\text{z.B. } \{x \mid x \text{ ist eine natürliche Zahl kleiner als } 6\}$$

c. Sei M eine Menge.

- $x \in M$ heißt “ x ist Element von M ”
- $x \notin M$ heißt “ x ist nicht Element von M ”

d. $\{\}$ und \emptyset bezeichnen die *leere Menge*, d.h. die Menge, die kein Element enthält.

Definition 2.3 (Inklusionsrelationen)

Für zwei Mengen M und N definieren wir:

- 1) $M \subseteq N \iff (x \in M \Rightarrow x \in N)$ “ M ist *Teilmenge* von N ”
- 2) $M = N \iff (M \subseteq N \wedge N \subseteq M)$
 $\iff (x \in M \Leftrightarrow x \in N)$
- 3) $M \neq N \iff \neg(M = N)$
 $\iff ((\exists x \in M : x \notin N) \vee (\exists x \in N : x \notin M))$
- 4) $M \subsetneq N \iff (M \subseteq N \wedge M \neq N)$ “ M ist *echte Teilmenge* von N ”

Beispiel 2.4

- a. $\{1, 2, 5, 3, 4, 0\} = \{x \mid x \text{ ist eine natürliche Zahl kleiner als } 6\}$.
- b. $\{1, 3\} \subsetneq \{1, 2, 3\}$.
- c. $\{1, 2, 1\} = \{1, 2\} = \{2, 1\}$.
- d. $1 \notin \{2, 3\}, 2 \in \{2, 3\}$.

Bemerkung 2.5 (Die Zahlbereiche)

Wir setzen die folgenden Mengen in unserer Vorlesung als bekannt voraus:

- $\mathbb{N} = \{0, 1, 2, 3, 4, \dots\}$ die Menge der *natürlichen Zahlen*,
- $\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$ die Menge der *ganzen Zahlen*,
- $\mathbb{Q} = \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z}, q \neq 0 \right\}$ die Menge der *rationalen Zahlen*,
- \mathbb{R} , die Menge der *reellen Zahlen*, d.h. der Dezimalbrüche.

Beachte:

$$\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}.$$

Im Verlauf der Vorlesung werden wir viele bekannte Eigenschaften dieser Mengen nochmals ausführlich thematisieren.

Definition 2.6 (Operationen von Mengen)

Es seien M, N, P sowie M_i für $i \in I$ Mengen.

- a. $M \cap N := \{x \mid x \in M \wedge x \in N\}$ heißt der *Durchschnitt* von M und N .
- b. $M \cup N := \{x \mid x \in M \vee x \in N\}$ heißt die *Vereinigung* von M und N .
- c. $M \setminus N := \{x \mid x \in M \wedge x \notin N\}$ heißt die *Differenzmenge* von M und N . Wir sagen auch *M ohne N*.
- d. $M \times N := \{(x, y) \mid x \in M \wedge y \in N\}$ heißt das *kartesische Produkt* von M und N . Dabei ist (x, y) ein *geordnetes Paar*, und für zwei geordnete Paare $(x, y), (u, v) \in M \times N$ gilt

$$(x, y) = (u, v) \iff (x = u \wedge y = v).$$

- e. M und N heißen genau dann *disjunkt*, wenn $M \cap N = \emptyset$, d.h. wenn sie kein Element gemeinsam besitzen.
- f. $P = M \cup N \iff (P = M \cup N \wedge M \cap N = \emptyset)$.
Wir sagen dann, P ist die *disjunkte Vereinigung* von M und N .

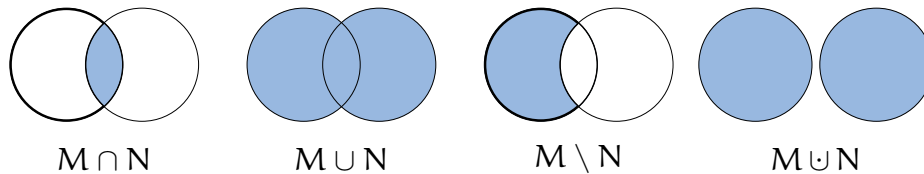


ABBILDUNG 1. Vereinigung, Durchschnitt, Differenzmenge, disjunkte Vereinigung

- g. $\bigcap_{i \in I} M_i := \{x \mid x \in M_i \forall i \in I\}$ heißt der *Durchschnitt* der M_i .
- h. $\bigcup_{i \in I} M_i := \{x \mid \exists i \in I : x \in M_i\}$ heißt die *Vereinigung* der M_i .
- i. $P = \bigcup_{i \in I} M_i \iff (P = \bigcup_{i \in I} M_i \wedge M_i \cap M_j = \emptyset \forall i, j \in I \text{ mit } i \neq j)$.
Wir nennen die $(M_i)_{i \in I}$ dann auch eine *disjunkte Zerlegung* von P , und wir sagen, die M_i sind *paarweise disjunkt*.

Beispiel 2.7

Ist $M = \{1, 2\}$ und $N = \{e, \pi, i\}$, so ist

$$M \times N = \{(1, e), (1, \pi), (1, i), (2, e), (2, \pi), (2, i)\}.$$

Proposition 2.8 (Einfache Rechengesetze für Mengenoperationen)

Es seien M, N, P Mengen.

- a. Assoziativgesetze
 - $(M \cup N) \cup P = M \cup (N \cup P)$.
 - $(M \cap N) \cap P = M \cap (N \cap P)$.

- b. Kommutativgesetze
- $M \cup N = N \cup M.$
 - $M \cap N = N \cap M.$
- c. Distributivgesetze
- $M \cap (N \cup P) = (M \cap N) \cup (M \cap P).$
 - $M \cup (N \cap P) = (M \cup N) \cap (M \cup P).$
- d. Identitätsgesetze
- $M \cup \emptyset = M.$
 - $M \subseteq N \implies M \cap N = M.$
- e. Komplementgesetze
- $M \subseteq N \implies M \cup (N \setminus M) = N.$
 - $M \subseteq N \implies M \cap (N \setminus M) = \emptyset.$

Beweis: a., d. und e. überlassen wir dem Leser als Übungsaufgabe.

b. Es gilt:

$$M \cup N = \{x \mid x \in M \vee x \in N\} \stackrel{1.5}{=} \{x \mid x \in N \vee x \in M\} = N \cup M$$

und

$$M \cap N = \{x \mid x \in M \wedge x \in N\} \stackrel{1.5}{=} \{x \mid x \in N \wedge x \in M\} = N \cap M.$$

c. Es gilt:

$$\begin{aligned} x \in M \cap (N \cup P) &\iff x \in M \wedge x \in N \cup P \\ &\iff x \in M \wedge (x \in N \vee x \in P) \\ &\stackrel{1.5}{\iff} (x \in M \wedge x \in N) \vee (x \in M \wedge x \in P) \\ &\iff x \in M \cap N \vee x \in M \cap P \\ &\iff x \in (M \cap N) \cup (M \cap P) \end{aligned}$$

und

$$\begin{aligned} x \in M \cup (N \cap P) &\iff x \in M \vee x \in N \cap P \\ &\iff x \in M \vee (x \in N \wedge x \in P) \\ &\stackrel{1.5}{\iff} (x \in M \vee x \in N) \wedge (x \in M \vee x \in P) \\ &\iff x \in M \cup N \wedge x \in M \cup P \\ &\iff x \in (M \cup N) \cap (M \cup P). \end{aligned}$$

□

Bemerkung 2.9 (Paradoxon von Russel)

Man muß bei der Definition von Mengen mittels Eigenschaften vorsichtig sein!

Betrachte die “Menge”

$$M = \{X \mid X \text{ ist Menge} \wedge X \notin X\}$$

aller Mengen, die sich nicht selbst als Element enthalten!

Angenommen, M wäre eine Menge. Dann sind zwei Fälle zu unterscheiden.

1. **Fall:** $M \notin M$: Dann ist M eine Menge, die sich nicht selbst als Element enthält. Mithin gilt $M \in M$ aufgrund der Definition von M . Dies ist ein Widerspruch.
2. **Fall:** $M \in M$: Dann ist M eine Menge, die sich selbst als Element enthält. Mithin gilt $M \notin M$ aufgrund der Definition von M . Dies ist ebenfalls ein Widerspruch.

Also kann keiner der beiden Fälle auftreten, und wir haben insgesamt einen Widerspruch hergeleitet.

Fazit: M ist *keine* Menge! Auch die *Menge aller Mengen* gibt es nicht!

Aufgaben

Aufgabe 2.10 (De Morgansche Regeln)

Es seien M und M_i , $i \in I$, Mengen. Zeige, die de Morganschen Regeln

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

und

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

§ 3 Abbildungen

In diesem Abschnitt wollen wir den für die Mathematik zentralen Begriff der Abbildung einführen.

Definition 3.1 (Abbildungen)

Es seien M und N zwei Mengen. Eine *Abbildung* oder *Funktion* f von M nach N ist eine *eindeutige Zuordnung*, die *jedem* Element $x \in M$ *genau ein* Element $f(x) \in N$ zuweist. Wir werden den Begriff *Funktion* nur dann verwenden, wenn $N = \mathbb{R}$ ist.

Wir nennen M den *Definitionsbereich* von f und N den *Ziel-* oder *Wertebereich*.

Notation:

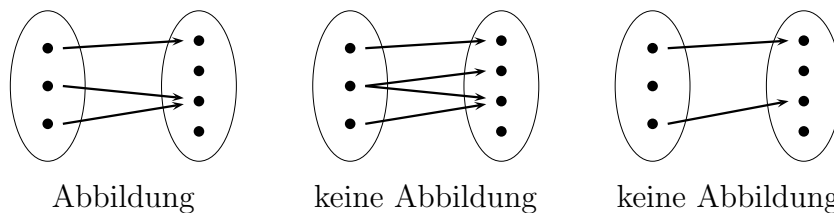
$$f : M \longrightarrow N : x \mapsto f(x).$$

Beachte, aufgrund der Definition einer Abbildung, gilt für zwei Abbildungen $f : M \longrightarrow N$ und $g : X \longrightarrow Y$:

$$f = g \iff (M = X \wedge N = Y \wedge \forall x \in M : f(x) = g(x)).$$

Beispiel 3.2

- a. Die folgenden Bilder sollen den Begriff der Abbildung graphisch veranschaulichen:



- b. $f : \mathbb{N} \longrightarrow \mathbb{N} : x \mapsto x^2$ und $g : \mathbb{Z} \longrightarrow \mathbb{N} : x \mapsto x^2$. Beachte: $f \neq g$, da ihre Definitionsbereiche nicht übereinstimmen.
- c. Sei $f : M \longrightarrow N$ eine Abbildung und $A \subseteq M$. Dann heißt die Abbildung

$$f|_A : A \longrightarrow N : x \mapsto f(x)$$

die *Einschränkung* von f auf A .

- d. Sei M eine Menge. Dann heißt die Abbildung

$$\text{id}_M : M \longrightarrow M : x \mapsto x$$

die *Identität* auf M .

Definition 3.3 (Bilder und Urbilder)

Es sei $f : M \longrightarrow N$ eine Abbildung, $A \subseteq M$ und $B \subseteq N$.

- a. $\text{Graph}(f) := \{(x, f(x)) \mid x \in M\} \subseteq M \times N$ heißt der *Graph* von f .
- b. $f(A) := \{f(x) \mid x \in A\} \subseteq N$ heißt das *Bild* von A unter f .
- c. $\text{Im}(f) := f(M) \subseteq N$ heißt das *Bild* von f .
- d. $f^{-1}(B) := \{x \in M \mid f(x) \in B\} \subseteq M$ heißt das *Urbild* von B unter f .

Beispiel 3.4

- a. Wir betrachten die Abbildung $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^2$.

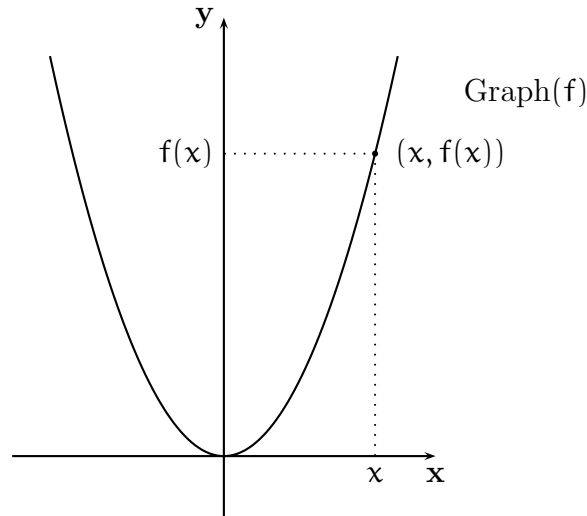


ABBILDUNG 2. Graph(f) für $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^2$

- Der Graph von f ist in Abbildung 2 zu sehen.
 - Für $A = \{-1, 0, 1, 2\}$ ist $f(A) = \{0, 1, 4\}$.
 - Für $B = \{0, 1\}$ ist $f^{-1}(B) = \{0, 1, -1\}$.
 - Für $B' = \{-1\}$ ist $f^{-1}(B') = \emptyset$.
 - $\text{Im}(f) = \{x \in \mathbb{R} \mid x \geq 0\}$.
- b. Die Abbildung $nf : \mathbb{N} \rightarrow \mathbb{N} : x \mapsto x + 1$ nennen wir die *Nachfolgerfunktion*. Es gelten

$$\text{Im}(nf) = \mathbb{N} \setminus \{0\}$$

und

$$\forall y \in \text{Im}(f) : nf^{-1}(\{y\}) = \{y - 1\}.$$

Bemerkung 3.5 (Abbildungen und ihre Graphen)

- a. Für zwei Abbildungen $f : M \rightarrow N$ und $g : P \rightarrow N$ gilt:

$$f = g \iff \text{Graph}(f) = \text{Graph}(g).$$

- b. Ist $\Gamma \subseteq M \times N$ so, daß

$$\forall x \in M \exists_1 y \in N : (x, y) \in \Gamma,$$

dann gibt es eine Abbildung $f : M \rightarrow N$ mit $\Gamma = \text{Graph}(f)$.

Fazit: Man hätte Abbildungen von M nach N auch als Teilmengen von $M \times N$ definieren können, die die Bedingung in b. erfüllen. So würde man vorgehen, wenn man die Mathematik ausgehend vom Begriff der Menge sukzessive aufbauen möchte.

Mit dieser Beschreibung sieht man übrigens sehr schön, daß es für jede Menge M genau eine Abbildung $f : \emptyset \rightarrow M$ gibt, und daß es für eine nicht-leere Menge M keine Abbildung $f : M \rightarrow \emptyset$ geben kann.

Definition 3.6 (Injektiv, surjektiv, bijektiv)

Es sei $f : M \rightarrow N$ eine Abbildung.

- a. f heißt genau dann *injektiv*, wenn

$$\forall x, x' \in M : f(x) = f(x') \implies x = x'.$$

- b. f heißt genau dann *surjektiv*, wenn

$$\forall y \in N \exists x \in M : f(x) = y,$$

d.h. wenn $\text{Im}(f) = N$.

- c. f heißt genau dann *bijektiv*, wenn f injektiv und surjektiv ist.

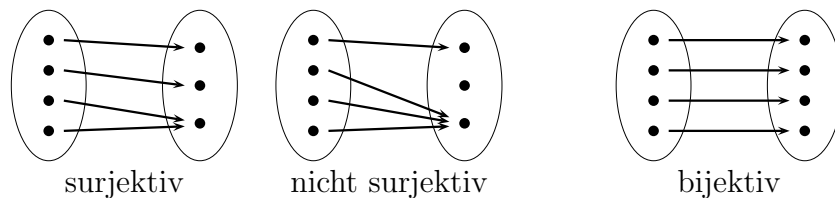
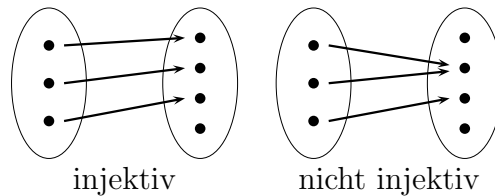
Bemerkung 3.7 (Injektiv, surjektiv, bijektiv)

Es sei $f : M \rightarrow N$ eine Abbildung.

- a. Ist $y \in N$ und $x \in M$ mit $f(x) = y$, so nennen wir x *ein Urbild* von y unter f .

- b. Es gelten:

- f ist injektiv \iff jedes $y \in N$ hat *höchstens* ein Urbild.
- f ist surjektiv \iff jedes $y \in N$ hat *mindestens* ein Urbild.
- f ist bijektiv \iff jedes $y \in N$ hat *genau* ein Urbild.

**Beispiel 3.8**

- a. Die Nachfolgerfunktion $nf : \mathbb{N} \rightarrow \mathbb{N} : x \mapsto x + 1$ ist injektiv, aber nicht surjektiv. Denn, $x + 1 = nf(x) = nf(y) = y + 1$ für $x, y \in \mathbb{N}$ impliziert $x = y$, und $0 \notin \text{Im}(f)$.
- b. $g : \mathbb{Z} \rightarrow \mathbb{N} : x \mapsto x^2$ ist nicht injektiv. Denn, für $x = 1 \neq -1 = y$ gilt $g(x) = g(1) = 1 = g(-1) = g(y)$.
- c. Die Abbildung id_M ist bijektiv für jede Menge M . Denn, für $y \in M$ gilt $\text{id}_M(y) = y$, so daß id_M surjektiv ist, und für $x, x' \in M$ mit $\text{id}_M(x) = \text{id}_M(x')$ gilt $x = x'$, so daß id_M injektiv ist.
- d. Ist $f : M \rightarrow N$ injektiv, so ist die Abbildung $M \rightarrow \text{Im}(f) : x \mapsto f(x)$ offenbar bijektiv.

Definition 3.9 (Komposition von Abbildungen)

Seien $f : M \rightarrow N$ und $g : N \rightarrow P$ zwei Abbildungen. Die Abbildung

$$g \circ f : M \rightarrow P : x \mapsto g(f(x))$$

heißt die *Komposition* oder *Verkettung* von f und g .

Beispiel 3.10

Seien $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^2$ und $g : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x + 1$. Dann gilt

$$(g \circ f)(x) = g(f(x)) = g(x^2) = x^2 + 1$$

und

$$(f \circ g)(x) = f(g(x)) = f(x + 1) = (x + 1)^2 = x^2 + 2x + 1.$$

Man beachte, daß die Abbildungen $g \circ f$ und $f \circ g$ nicht gleich sind, da $(g \circ f)(1) = 2 \neq 4 = (f \circ g)(1)$.

Proposition 3.11 (Assoziativität der Komposition)

Seien $f : M \rightarrow N$, $g : N \rightarrow P$ und $h : P \rightarrow Q$ Abbildungen. Dann gilt

$$(h \circ g) \circ f = h \circ (g \circ f).$$

Wir schreiben deshalb auch kurz $h \circ g \circ f$.

Beweis: Da die Definitions- und Zielbereiche der beiden Funktionen übereinstimmen, reicht es, die Abbildungsvorschrift zu überprüfen. Sei dazu $x \in M$. Dann gilt

$$((h \circ g) \circ f)(x) = (h \circ g)(f(x)) = h(g(f(x))) = h((g \circ f)(x)) = (h \circ (g \circ f))(x).$$

Dies zeigt die Behauptung. □

Satz 3.12 (Bijektivität = Existenz einer Umkehrabbildung)

Es sei $f : M \rightarrow N$ eine Abbildung.

- a. f ist genau dann bijektiv, wenn eine Abbildung $g : N \rightarrow M$ existiert, so daß $g \circ f = \text{id}_M$ und $f \circ g = \text{id}_N$.
- b. Die Abbildung g in Teil a. ist dann eindeutig bestimmt und bijektiv. Wir nennen sie die Inverse oder Umkehrabbildung von f und bezeichnen sie mit f^{-1} .

Beweis:

- a. **” \Leftarrow ”:** Wir wollen zunächst zeigen, daß f surjektiv ist. Sei dazu $y \in N$ gegeben. Setze $x := g(y) \in M$. Dann gilt

$$f(x) = f(g(y)) = (f \circ g)(y) = \text{id}_N(y) = y.$$

Also ist f surjektiv.

Dann wollen wir zeigen, daß f injektiv ist. Seien dazu $x, x' \in M$ mit $f(x) = f(x')$ gegeben. Dann gilt

$$x = \text{id}_M(x) = (g \circ f)(x) = g(f(x)) = g(f(x')) = (g \circ f)(x') = \text{id}_M(x') = x'.$$

Also ist f injektiv.

” \implies “: Da f bijektiv ist, gibt es für jedes $\mathbf{y} \in \mathbf{N}$ genau ein Urbild $\mathbf{x}_{\mathbf{y}} \in \mathbf{M}$ von \mathbf{y} unter f , d.h. $f(\mathbf{x}_{\mathbf{y}}) = \mathbf{y}$. Wir definieren nun eine Abbildung

$$\mathbf{g} : \mathbf{N} \longrightarrow \mathbf{M} : \mathbf{y} \mapsto \mathbf{x}_{\mathbf{y}}.$$

Dann gilt zunächst für $\mathbf{y} \in \mathbf{N}$

$$(f \circ \mathbf{g})(\mathbf{y}) = f(\mathbf{g}(\mathbf{y})) = f(\mathbf{x}_{\mathbf{y}}) = \mathbf{y} = \text{id}_{\mathbf{N}}(\mathbf{y}).$$

Also ist $f \circ \mathbf{g} = \text{id}_{\mathbf{N}}$.

Zudem gilt für $\mathbf{x} \in \mathbf{M}$ und $\mathbf{y} := f(\mathbf{x}) \in \mathbf{N}$

$$f(\mathbf{x}_{\mathbf{y}}) = \mathbf{y} = f(\mathbf{x}).$$

Da f injektiv ist, folgt daraus $\mathbf{x} = \mathbf{x}_{\mathbf{y}}$, und wir erhalten

$$(\mathbf{g} \circ f)(\mathbf{x}) = \mathbf{g}(f(\mathbf{x})) = \mathbf{g}(\mathbf{y}) = \mathbf{x}_{\mathbf{y}} = \mathbf{x} = \text{id}_{\mathbf{M}}(\mathbf{x}).$$

Damit ist auch $\mathbf{g} \circ f = \text{id}_{\mathbf{M}}$ gezeigt.

- b. Sei $\mathbf{h} : \mathbf{N} \longrightarrow \mathbf{M}$ eine zweite Abbildung mit $\mathbf{h} \circ f = \text{id}_{\mathbf{M}}$ und $f \circ \mathbf{h} = \text{id}_{\mathbf{N}}$. Dann gilt für $\mathbf{y} \in \mathbf{N}$

$$f(\mathbf{g}(\mathbf{y})) = (f \circ \mathbf{g})(\mathbf{y}) = \text{id}_{\mathbf{N}}(\mathbf{y}) = (f \circ \mathbf{h})(\mathbf{y}) = f(\mathbf{h}(\mathbf{y})).$$

Da f injektiv ist, folgt mithin $\mathbf{g}(\mathbf{y}) = \mathbf{h}(\mathbf{y})$, und somit $\mathbf{g} = \mathbf{h}$. Die Eindeutigkeit von \mathbf{g} ist also gezeigt. Außerdem ist \mathbf{g} nach Teil a. auch bijektiv.

□

Beispiel 3.13

Die Abbildung $f : \mathbb{R} \longrightarrow \mathbb{R} : \mathbf{x} \mapsto 2\mathbf{x} + 1$ ist bijektiv mit $f^{-1} : \mathbb{R} \longrightarrow \mathbb{R} : \mathbf{y} \mapsto \frac{1}{2} \cdot \mathbf{y} - \frac{1}{2}$.

Denn für $\mathbf{y} \in \mathbb{R}$ gilt

$$(f \circ f^{-1})(\mathbf{y}) = 2 \cdot \left(\frac{1}{2} \cdot \mathbf{y} - \frac{1}{2} \right) + 1 = \mathbf{y} = \text{id}_{\mathbb{R}}(\mathbf{y})$$

und für $\mathbf{x} \in \mathbb{R}$ gilt

$$(f^{-1} \circ f)(\mathbf{x}) = \frac{1}{2} \cdot (2\mathbf{x} + 1) - \frac{1}{2} = \mathbf{x} = \text{id}_{\mathbb{R}}(\mathbf{x}).$$

Die Behauptung folgt also aus Satz 3.12.

Proposition 3.14 (Injektivität, Surjektivität, Bijektivität unter Komposition)

Seien $f : \mathbf{M} \longrightarrow \mathbf{N}$ und $\mathbf{g} : \mathbf{N} \longrightarrow \mathbf{P}$ zwei Abbildungen.

- Sind f und \mathbf{g} injektiv, so ist $\mathbf{g} \circ f$ injektiv.
- Sind f und \mathbf{g} surjektiv, so ist $\mathbf{g} \circ f$ surjektiv.
- Sind f und \mathbf{g} bijektiv, so ist $\mathbf{g} \circ f$ bijektiv.

Beweis: a. Seien $x, x' \in M$ mit $(g \circ f)(x) = (g \circ f)(x')$. Dann gilt

$$g(f(x)) = (g \circ f)(x) = (g \circ f)(x') = g(f(x')).$$

Da g injektiv ist, ist $f(x) = f(x')$, und da f injektiv ist, ist auch $x = x'$. Also ist $g \circ f$ injektiv.

b. Sei $z \in P$. Da g surjektiv ist, gibt es ein $y \in N$ mit $g(y) = z$, und da f surjektiv ist, gibt es ein $x \in M$ mit $f(x) = y$. Die Surjektivität von $g \circ f$ folgt dann aus

$$(g \circ f)(x) = g(f(x)) = g(y) = z.$$

c. Wegen a. ist $g \circ f$ injektiv und wegen b. ist $g \circ f$ auch surjektiv, also bijektiv. □

Lemma 3.15

Es sei $f : M \rightarrow N$ eine injektive Abbildung, $x' \in M$ und $y' = f(x') \in N$.

- a. Dann ist $g : M \setminus \{x'\} \rightarrow N \setminus \{y'\} : x \mapsto f(x)$ eine injektive Abbildung.
- b. Ist f bijektiv, so ist g auch bijektiv.

Beweis: Da f injektiv ist, gilt $f(x) \in N \setminus \{y'\}$ für alle $x \in M \setminus \{x'\}$, da y' unter f nur das Urbild x' hat. Durch die Abbildungsvorschrift in a. und b. wird also eine Abbildung g von $M \setminus \{x'\}$ nach $N \setminus \{y'\}$ definiert.

Zudem gilt offenbar, daß jedes Element in $N \setminus \{y'\}$ höchstens ein Urbild unter g haben kann, wenn f injektiv ist. Somit ist g dann also auch injektiv.

Setzen wir nun zudem voraus, daß f bijektiv ist, und ist $y \in N \setminus \{y'\}$ gegeben, so gibt es ein $x \in M$ mit $f(x) = y$. Wegen $f(x) = y \neq y' = f(x')$ und f injektiv muß zudem $x \neq x'$ gelten, d.h. $x \in M \setminus \{x'\}$. Also ist $g(x) = f(x) = y$, und wir haben gezeigt, daß dann auch g bijektiv ist. □

Lemma 3.16

Es sei M eine Menge und $a, b \in M$. Die Abbildung $h : M \rightarrow M$ mit

$$h(x) = \begin{cases} a, & \text{wenn } x = b, \\ b, & \text{wenn } x = a, \\ x, & \text{sonst,} \end{cases}$$

ist bijektiv mit $h^{-1} = h$. Beachte, h vertauscht genau die beiden Elemente a und b (sofern sie verschieden sind).

Beweis: Wenn $a = b$, so ist h die Identität und somit bijektiv. Wenn $a \neq b$, dann gilt für $x \in M$

$$(h \circ h)(x) = h(h(x)) = \begin{cases} h(b), & \text{wenn } x = a, \\ h(a), & \text{wenn } x = b, \\ h(x), & \text{sonst.} \end{cases} = x = \text{id}_M(x).$$

Mithin ist $h \circ h = \text{id}_M$ und somit ist h nach Satz 3.12 bijektiv mit $h^{-1} = h$. □

Aufgaben

Aufgabe 3.17

Ist $f : M \rightarrow N$ eine surjektive Abbildung und $y \in N$, so ist

$$g : M \setminus f^{-1}(\{y\}) \rightarrow N \setminus \{y\} : x \mapsto f(x)$$

eine surjektive Abbildung.

Aufgabe 3.18

Für eine Abbildung $f : M \rightarrow N$, $M \neq \emptyset$, beweise man die folgenden Aussagen:

- a. f ist injektiv $\iff \exists g : N \rightarrow M$, so dass $g \circ f = \text{id}_M$.
- b. f ist surjektiv $\iff \exists g : N \rightarrow M$, so dass $f \circ g = \text{id}_N$.

Aufgabe 3.19

Untersuche ob die folgenden Abbildungen injektiv, surjektiv oder bijektiv sind:

- a. $f_1 : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto 3x + 2$
- b. $f_2 : \mathbb{Z} \rightarrow \mathbb{Z} : x \mapsto 3x + 2$
- c. $f_3 : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \times \mathbb{R} : (x, y) \mapsto (xy, x + 1)$
- d. $f_4 : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \times \mathbb{R} : (x, y) \mapsto (x - 2y, 2x + y)$

Aufgabe 3.20

Seien M, N zwei nicht-leere Mengen und $f : M \rightarrow N$ eine Abbildung. Formuliere die folgende Aussage in Quantorenschreibweise und beweise sie:

f ist genau dann surjektiv, wenn für alle nicht-leeren Mengen X und für alle Abbildungen $g : N \rightarrow X$ und $h : N \rightarrow X$ aus $g \circ f = h \circ f$ schon $g = h$ folgt.

Aufgabe 3.21

Seien L, M, N Mengen und $f : L \rightarrow M$, $g : M \rightarrow N$ Abbildungen. Beweise oder widerlege - durch Gegenbeispiel - die folgenden Aussagen:

- a. Ist $g \circ f$ injektiv, so ist g injektiv.
- b. Ist $g \circ f$ injektiv, so ist f injektiv.

Aufgabe 3.22

Seien M, N Mengen, $A_1, A_2 \subseteq M$ und $B, B_1, B_2 \subseteq N$ Teilmengen und $f : M \rightarrow N$ eine Abbildung. Beweise die folgenden Aussagen:

- a. $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$.
- b. $f(f^{-1}(B)) \subseteq B$.
- c. $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$.
- d. $f(A_1 \cap A_2) \subseteq f(A_1) \cap f(A_2)$.

Gib außerdem konkrete Beispiele dafür an, dass in b. und d. keine Gleichheit gilt.

§ 4 Vollständige Induktion

Bemerkung 4.1 (Prinzip der vollständigen Induktion)

Die folgende Eigenschaft der natürlichen Zahlen ist uns wohl vertraut:

Addiert man zur Zahl 0 sukzessive die Zahl 1, so erhält man nach und nach alle natürlichen Zahlen.

Man nennt sie das *Prinzip der vollständigen Induktion*.

Mit Hilfe der Nachfolgerfunktion $\text{nf} : \mathbb{N} \rightarrow \mathbb{N} : n \mapsto n + 1$ können wir die Eigenschaft auch wie folgt formulieren:

Ist $M \subseteq \mathbb{N}$ mit $0 \in M$ und $\forall n \in M : n + 1 = \text{nf}(n) \in M$, so ist $M = \mathbb{N}$.

Daraus leitet sich das im folgenden Satz formulierte Beweisprinzip für Aussagen ab, die von einer natürlichen Zahl abhängen.

Satz 4.2 (Prinzip der vollständigen Induktion)

Sei $\mathcal{A}(n)$ eine Aussageform mit zulässigen Werten $n \in \mathbb{N}$.

Falls $\mathcal{A}(0)$ wahr ist und $\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$ wahr ist, so ist $\mathcal{A}(n)$ wahr für alle $n \in \mathbb{N}$.

Beweis: Setze $M := \{n \in \mathbb{N} \mid \mathcal{A}(n) \text{ wahr}\}$. Nach Voraussetzung gilt dann $0 \in M$ und für $n \in M$ folgt $n + 1 \in M$. Aus dem Prinzip der Vollständigen Induktion in Bemerkung 4.1 folgt dann $M = \mathbb{N}$. Also ist $\mathcal{A}(n)$ wahr für alle $n \in \mathbb{N}$. \square

Bemerkung 4.3

Man beachte, um den Schluß $\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$ als wahr zu erweisen, reicht es, den Fall zu betrachten, daß $\mathcal{A}(n)$ wahr ist, da andernfalls der Schluß ohnehin den Wahrheitswert wahr trägt.

Wir nennen:

- “ $\mathcal{A}(0)$ wahr” den *Induktionsanfang*,
- “ $\mathcal{A}(n)$ wird als wahr vorausgesetzt” die *Induktionsvoraussetzung* und
- “ $\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$ ” den *Induktionsschluß*.

Beispiel 4.4

Die Zahl $n^3 - n$ ist für jedes $n \in \mathbb{N}$ durch 6 teilbar.

Beweis: Wir führen den Beweis durch vollständige Induktion und formulieren dazu zunächst unsere Aussageform $\mathcal{A}(n)$:

$$\mathcal{A}(n) : \text{Es gibt ein } k \in \mathbb{N} \text{ mit } n^3 - n = 6 \cdot k.$$

Induktionsanfang: $n = 0$: $0^3 - 0 = 0 = 6 \cdot 0$. Also ist $\mathcal{A}(0)$ wahr.

Induktionsvoraussetzung: Wir setzen voraus, daß $\mathcal{A}(n)$ wahr ist, d.h. es gibt ein $k \in \mathbb{N}$ mit $n^3 - n = 6 \cdot k$.

Induktionsschritt: $n \mapsto n + 1$: Man beachte, daß eine der beiden Zahlen n oder $n + 1$ gerade sein muß, und daß deshalb die Zahl $n \cdot (n + 1)$ gerade ist. Es gibt also eine natürliche Zahl $l \in \mathbb{N}$ mit $n \cdot (n + 1) = 2 \cdot l$. Damit erhalten wir dann

$$(n + 1)^3 - (n + 1) = (n^3 - n) + 3 \cdot n \cdot (n + 1) = 6k + 6l = 6 \cdot (k + l).$$

Wir haben also gezeigt, daß $\mathcal{A}(n + 1)$ wahr ist.

Also ist $\mathcal{A}(n)$ wahr für alle $n \in \mathbb{N}$, und das heißt, daß $n^3 - n$ stets durch 6 teilbar ist. \square

Bemerkung 4.5 (Varianten der vollständigen Induktion)

a. *Alternativer Induktionsanfang:*

Statt $n = 0$ als Induktionsanfang zu wählen, kann eine beliebige ganze Zahl $n_0 \in \mathbb{Z}$ als Induktionsanfang dienen. Man erhält dann, daß $\mathcal{A}(n)$ wahr ist für alle $n \geq n_0$. Denn, man erhält alle ganzen Zahlen $n \geq n_0$, indem man zu n_0 sukzessive 1 addiert.

b. *Alternative Induktionsvoraussetzung:*

Im Induktionsschritt schließen wir von $\mathcal{A}(n)$ auf $\mathcal{A}(n + 1)$, d.h. wir setzen nur $\mathcal{A}(n)$ als richtig voraus und schließen daraus die Korrektheit von $\mathcal{A}(n + 1)$. Stattdessen können wir auch $\mathcal{A}(k)$ für $k = n_0, \dots, n$ als richtig voraussetzen und auf $\mathcal{A}(n + 1)$ schließen (wobei $\mathcal{A}(n_0)$ der Induktionsanfang sein soll). Das ist manchmal hilfreich.

Aufgaben

Aufgabe 4.6

Zeige, daß $3^{n+1} - 3$ für jede natürliche Zahl $n \in \mathbb{N}$ durch 6 teilbar ist.

Aufgabe 4.7

Es sei $a \in \mathbb{N}$ eine natürliche Zahl. Zeige, daß dann $a^{2n+1} - a$ für jede natürliche Zahl $n \in \mathbb{N}$ durch 6 teilbar ist.

§ 5 Mächtigkeit von Mengen

Definition 5.1

Für zwei ganze Zahlen $m, n \in \mathbb{Z}$ bezeichnen wir mit

$$\{m, \dots, n\} := \{k \in \mathbb{Z} \mid m \leq k \leq n\}$$

die Menge der ganzen Zahlen zwischen m und n . Man beachte, daß $\{m, \dots, n\} = \emptyset$, wenn $m > n$.

Lemma 5.2 (Der Prozeß des Zählens)

Es seien $m, n \in \mathbb{N}$ zwei natürliche Zahlen.

- Genau dann gibt es ein injektives $f: \{1, \dots, n\} \rightarrow \{1, \dots, m\}$, wenn $n \leq m$.*
- Genau dann gibt es ein surjektives $f: \{1, \dots, n\} \rightarrow \{1, \dots, m\}$, wenn $n \geq m \geq 1$.*
- Genau dann gibt es eine bijektives $f: \{1, \dots, n\} \rightarrow \{1, \dots, m\}$, wenn $n = m$.*

Beweis:

- a. " \Leftarrow ": Ist $n \leq m$, so ist die Abbildung

$$f: \{1, \dots, n\} \rightarrow \{1, \dots, m\}: x \mapsto x$$

offensichtlich injektiv.

" \Rightarrow ": Wir führen den Beweis durch Induktion nach n und formulieren dazu zunächst eine Aussageform $\mathcal{A}(n)$:

$\mathcal{A}(n)$: Wenn für ein $m \in \mathbb{N}$ eine injektive Abbildung $g: \{1, \dots, n\} \rightarrow \{1, \dots, m\}$ existiert, dann ist $n \leq m$.

Induktionsanfang: $n = 0$: Hier ist nichts zu zeigen, da jede natürliche Zahl größer oder gleich 0 ist.

Induktionsvoraussetzung: Wir setzen voraus, daß $\mathcal{A}(n)$ wahr ist.

Induktionsschritt: Wir wollen zeigen, daß $\mathcal{A}(n+1)$ wahr ist. Dazu sei $m \in \mathbb{N}$ so, daß es eine injektive Abbildung

$$g: \{1, \dots, n+1\} \rightarrow \{1, \dots, m\}$$

gibt. Nach Lemma 3.16 gibt es eine bijektive Abbildung

$$h: \{1, \dots, m\} \rightarrow \{1, \dots, m\},$$

die die Zahlen m und $g(n+1)$ vertauscht. Nach Proposition 3.14 ist

$$h \circ g: \{1, \dots, n+1\} \rightarrow \{1, \dots, m\}$$

als Komposition zweier injektiver Abbildungen wieder injektiv. Außerdem gilt

$$(h \circ g)(n+1) = h(g(n+1)) = m,$$

und mit Lemma 3.15 ist dann auch

$$f: \{1, \dots, n\} \rightarrow \{1, \dots, m-1\}: x \mapsto (h \circ g)(x)$$

injektiv. Da nach Induktionsvoraussetzung $\mathcal{A}(n)$ wahr ist, impliziert dies, daß $n \leq m - 1$ gilt und somit auch $n + 1 \leq m$. Also ist auch $\mathcal{A}(n + 1)$ korrekt.

Aus dem Prinzip der vollständigen Induktion 4.2 folgt damit, daß $\mathcal{A}(n)$ korrekt ist für alle $n \in \mathbb{N}$.

- b. Man beachte zunächst, daß die Existenz von g bedingt, daß $\{1, \dots, m\}$ nicht-leer und mithin $m \geq 1$ ist, so daß dies auf alle Fälle vorausgesetzt werden kann.

Wegen Aufgabe 3.18 gibt es eine surjektive Abbildung

$$f : \{1, \dots, n\} \longrightarrow \{1, \dots, m\},$$

genau dann, wenn es dazu passend eine Abbildung

$$g : \{1, \dots, m\} \longrightarrow \{1, \dots, n\}.$$

mit

$$\text{id}_{\{1, \dots, m\}} = f \circ g$$

gibt, wobei diese Bedingung wegen der gleichen Aufgabe dazu gleichwertig ist, daß g injektiv ist. Wegen Teil a. ist letzteres aber wiederum gleichwertig dazu, daß $m \leq n$ gilt.

- c. " \Leftarrow ": Ist $m = n$, so ist $f = \text{id}_{\{1, \dots, n\}}$ eine bijektive Abbildung.
 " \Rightarrow ": Die Abbildung f ist injektiv und surjektiv, so daß mit Teil a. und b. auch $m = n$ folgt.

□

Definition 5.3 (Die Mächtigkeit von Mengen)

- a. Wir nennen eine Menge M *endlich*, wenn es eine natürliche Zahl $n \in \mathbb{N}$ und eine bijektive Abbildung

$$f : \{1, \dots, n\} \longrightarrow M$$

gibt, und in diesem Fall heißt die wegen Lemma 5.2 eindeutig bestimmte Zahl n die *Mächtigkeit* von M . Wir schreiben

$$\#M := |M| := n$$

Andernfalls nennen wir M *unendlich* und setzen $\#M := |M| := \infty$.

- b. Wir nennen zwei Mengen M und N *gleichmächtig*, wenn es eine bijektive Abbildung $f : M \longrightarrow N$ gibt.
 c. Eine Menge heißt *abzählbar unendlich*, wenn sie gleichmächtig zu \mathbb{N} ist.
 d. Eine Menge heißt *überabzählbar*, wenn sie weder endlich noch abzählbar unendlich ist.

Bemerkung 5.4 (Einfache Eigenschaften der Mächtigkeit endlicher Mengen)

- a. Die Mächtigkeit einer Menge ist also die *Anzahl der Elemente* der Menge. Die Idee dabei ist, daß die bijektive Abbildung $f : \{1, \dots, n\} \longrightarrow M$ dazu dient, die Elemente von M aufzuzählen:

$$M = \{f(1), f(2), f(3), \dots, f(n-1), f(n)\}.$$

- b. Ist M endlich und $A \subseteq M$, so ist auch A endlich und $|A| \leq |M|$.
- c. Ist $M = A \cup B$ eine endliche Menge, so gilt $|M| = |A| + |B|$.
- d. So, wie wir den Begriff der Endlichkeit und der Mächtigkeit definiert haben, müßte man die Aussagen b. und c. eigentlich beweisen. Wir verzichten an dieser Stelle aber auf den Beweis und überlassen ihn dem Leser als Aufgabe 5.16. Mit der Interpretation der Mächtigkeit in Teil a. sind wir ohnehin sofort bereit, die Aussagen zu glauben.

Satz 5.5 (Gleichmächtigkeit)

Seien M und N zwei endliche Mengen.

Genau dann sind M und N gleichmächtig, wenn $|M| = |N|$.

Beweis: Sei $m := |M|$ und $n := |N|$, dann gibt es nach Definition bijektive Abbildungen

$$g : \{1, \dots, m\} \longrightarrow M$$

und

$$h : \{1, \dots, n\} \longrightarrow N.$$

Setzen wir nun zunächst voraus, daß M und N gleichmächtig sind, dann gibt es eine bijektive Abbildung

$$f : M \longrightarrow N.$$

Nach Proposition 3.14 ist dann aber auch die Abbildung

$$h^{-1} \circ f \circ g : \{1, \dots, m\} \longrightarrow \{1, \dots, n\}$$

bijektiv, und nach Lemma 5.2 gilt dann $m = n$.

Setzen wir nun umgekehrt $m = n$ voraus, so ist

$$h \circ g^{-1} : M \longrightarrow N$$

nach Proposition 3.14 bijektiv, und somit sind M und N gleichmächtig. \square

Lemma 5.6

Seien M und N zwei endliche Mengen.

Gibt es eine surjektive Abbildung $f : M \rightarrow N$, so gilt $|M| \geq |N|$.

Beweis: Wir setzen $m := |M|$ und $n := |N|$ und finden mithin bijektive Abbildungen

$$g : \{1, \dots, m\} \longrightarrow M$$

und

$$h : \{1, \dots, n\} \longrightarrow N.$$

Nach Proposition 3.14 ist die Abbildung

$$h^{-1} \circ f \circ g : \{1, \dots, m\} \longrightarrow \{1, \dots, n\}$$

als Komposition surjektiver Abbildungen dann surjektiv. Die Behauptung folgt dann aus Lemma 5.2. \square

Satz 5.7 (Injektiv = surjektiv = bijektiv bei gleichmächtigen endlichen Mengen)
Es seien M und N endliche Mengen mit $|M| = |N|$. Dann sind die folgenden Aussagen für eine Abbildung $f : M \longrightarrow N$ äquivalent:

- a. f ist injektiv.
- b. f ist surjektiv.
- c. f ist bijektiv.

Beweis:

a. \implies b.: Angenommen, f wäre nicht surjektiv, dann gibt es ein

$$y \in N \setminus \text{Im}(f)$$

und mithin ist

$$\text{Im}(f) \subseteq N \setminus \{y\}.$$

Da f injektiv ist, ist $g : M \longrightarrow \text{Im}(f) : x \mapsto f(x)$ nach Beispiel 3.8 bijektiv, so daß mit Satz 5.5 und Bemerkung 5.4

$$|M| \stackrel{5.5}{=} |\text{Im}(f)| \stackrel{5.4}{\leq} |N| - 1 < |N| = |M|$$

folgt, was ein offensichtlicher Widerspruch ist. Mithin muß f surjektiv sein.

b. \implies c.: Wir müssen zeigen, daß f injektiv ist. Dazu nehmen wir an, f sei nicht injektiv. Dann gibt es $x, x' \in M$ mit $x \neq x'$ und $y := f(x) = f(x')$. Die Abbildung

$$h : M \setminus f^{-1}(\{y\}) \longrightarrow N \setminus \{y\} : z \mapsto f(z)$$

ist nach Aufgabe 3.17 surjektiv. Mithin gilt nach Lemma 5.6 und Bemerkung 5.4

$$|M| - 1 \stackrel{\text{Vor.}}{=} |N| - 1 \stackrel{5.4}{=} |N \setminus \{y\}| \stackrel{5.6}{\leq} |M \setminus f^{-1}(\{y\})| \stackrel{5.4}{\leq} |M \setminus \{x, x'\}| \stackrel{5.4}{=} |M| - 2,$$

was offenbar ein Widerspruch ist. Mithin muß f injektiv sein.

c. \implies a.: Jede bijektive Abbildung ist auch injektiv, also ist f injektiv.

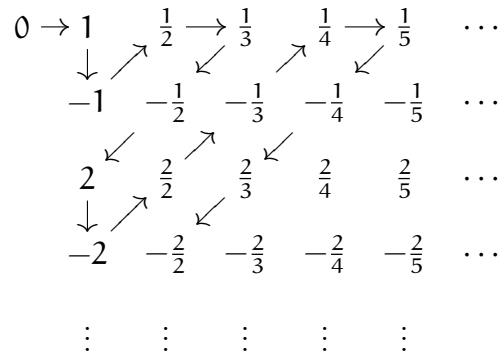
Damit haben wir die Aussage durch einen Ringschluß gezeigt. \square

Proposition 5.8 (Cantorsches Diagonalverfahren)

Die Menge \mathbb{Q} der rationalen Zahlen ist abzählbar unendlich.

Beweis: Wir zeigen, wie man mit Hilfe des Cantorschen Diagonalverfahrens, wie man eine bijektive Abbildung von \mathbb{N} nach \mathbb{Q} konstruiert.

Dazu listen wir die rationalen Zahlen zunächst wie folgt auf



und laufen sie dann wie angedeutet entlang der Pfeile ab. Dabei sammeln wir jede rationale Zahl, die mehrfach vorkommt, nur bei ihrem ersten Auftreten auf. Auf dem Weg erhalten wir eine bijektive Abbildung von \mathbb{N} nach \mathbb{Q} . □

Proposition 5.9 (\mathbb{R} ist überabzählbar.)

Die Menge \mathbb{R} der reellen Zahlen ist überabzählbar.

Beweis: Auch dies zeigen wir mit Hilfe einer Variante des Cantorschen Diagonalverfahrens.

\mathbb{R} ist sicherlich nicht endlich. Wäre \mathbb{R} abzählbar unendlich, so gäbe es eine bijektive Abbildung von $\varphi : \mathbb{N} \rightarrow \mathbb{R}$, und wir schreiben dann $\varphi(i)$, $i \in \mathbb{N}$, in Dezimaldarstellung (siehe auch Aufgabe 12.45):

$$\begin{array}{rcl}
 \varphi(0) & = & \mathbf{a}_{0,-p_0} \quad \mathbf{a}_{0,-p_0+1} \quad \dots \quad \underline{\mathbf{a}_{0,0}}, \quad \mathbf{a}_{01} \quad \mathbf{a}_{02} \quad \mathbf{a}_{03} \quad \dots \\
 \varphi(1) & = & \mathbf{a}_{1,-p_1} \quad \mathbf{a}_{1,-p_1+1} \quad \dots \quad \mathbf{a}_{1,0}, \quad \underline{\mathbf{a}_{11}} \quad \mathbf{a}_{12} \quad \mathbf{a}_{13} \quad \dots \\
 \varphi(2) & = & \mathbf{a}_{2,-p_2} \quad \mathbf{a}_{2,-p_2+1} \quad \dots \quad \mathbf{a}_{2,0}, \quad \mathbf{a}_{21} \quad \underline{\mathbf{a}_{22}} \quad \mathbf{a}_{23} \quad \dots \\
 & \vdots & & & & & & \ddots
 \end{array}$$

Dann setzen wir $\mathbf{a} := \mathbf{a}_{00}, \mathbf{a}_{11}\mathbf{a}_{22}\mathbf{a}_{33}\dots \in \mathbb{R}$, d. h. \mathbf{a} ist diejenige Zahl, die in obiger Aufzählung durch die unterstrichenen Diagonalelemente gegeben ist. Nun ändern wir jede der Ziffern von \mathbf{a} ab (etwa $\mathbf{b}_{ii} = 1$, falls $\mathbf{a}_{ii} = 0$ und $\mathbf{b}_{ii} = 0$ sonst) und erhalten eine Zahl

$$\mathbf{b} = \mathbf{b}_{00}, \mathbf{b}_{11}\mathbf{b}_{22}\mathbf{b}_{33}\dots \in \mathbb{R},$$

mit $\mathbf{a}_{ii} \neq \mathbf{b}_{ii}$ für alle $i \in \mathbb{N}$. Da φ bijektiv ist, gibt es ein $i \in \mathbb{N}$ mit $\varphi(i) = \mathbf{b}$, also $\mathbf{a}_{ii} = \mathbf{b}_{ii}$, im Widerspruch zur Konstruktion von \mathbf{b} . (Wir müssen noch berücksichtigen, daß $0,9999\dots = 1$, was aber die einzige Zweideutigkeit der Dezimaldarstellung ist, und dieser weichen wir durch unsere Wahl der \mathbf{b}_{ii} aus.) Also ist \mathbb{R} überabzählbar. □

Bemerkung 5.10 (Kontinuumshypothese)

Da \mathbb{Q} und \mathbb{R} nicht gleichmächtig sind und \mathbb{Q} eine Teilmenge von \mathbb{R} ist, stellt sich ganz natürlich die Frage, ob es eine Menge M mit $\mathbb{Q} \subsetneq M \subsetneq \mathbb{R}$ gibt, die weder

zu \mathbb{Q} noch zu \mathbb{R} gleichmächtig ist. Es hat lange gedauert, bis man feststellen mußte, daß die Frage auf der Grundlage des allgemein anerkannten Axiomensystems der Mengenlehre von Zermelo-Fränkel nicht entscheidbar ist. Man hat nun also die Wahl, als neues Axiom hinzuzufügen, daß es eine solche Menge gibt, oder auch, daß es keine solche Menge gibt. Die lange bestehende Vermutung, daß man schon mit den übrigen Axiomen beweisen könnte, daß es keine solche Menge gibt, ist als *Kontinuumshypothese* bekannt.

Definition 5.11 (Potenzmenge)

Es sei M eine Menge. Wir nennen die Menge

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

aller Teilmengen von M die *Potenzmenge* von M .

Beispiel 5.12

$$\mathcal{P}(\emptyset) = \{\emptyset\}, \quad \mathcal{P}(\{1\}) = \{\emptyset, \{1\}\}, \quad \mathcal{P}(\{1, 2\}) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}.$$

Proposition 5.13 (Potenzmengen endlicher Mengen)

Sei M eine endliche Menge mit $n = |M|$, so ist $|\mathcal{P}(M)| = 2^n$.

Beweis: Wir führen den Beweis durch Induktion nach n .

Induktionsanfang: $n = 0$: Dann ist $M = \emptyset$ und $\mathcal{P}(M) = \{\emptyset\}$ hat genau $1 = 2^0$ Elemente.

Induktionsschritt: $n \mapsto n + 1$: Sei also $|M| = n + 1$. Wir wählen ein $y \in M$ und setzen $N = M \setminus \{y\}$, so daß $|N| = |M| - 1 = n$. Die Potenzmenge $\mathcal{P}(M)$ läßt sich nun wie folgt disjunkt aufspalten:

$$\mathcal{P}(M) = \{A \subseteq M \mid y \notin A\} \cup \{A \subseteq M \mid y \in A\}.$$

Dabei ist

$$\{A \subseteq M \mid y \notin A\} = \{A \subseteq M \mid A \subseteq N\} = \mathcal{P}(N)$$

und

$$\{A \subseteq M \mid y \in A\} = \{B \cup \{y\} \subseteq M \mid B \subseteq N\} = \{B \cup \{y\} \subseteq M \mid B \in \mathcal{P}(N)\}.$$

Beide Mengen sind offenbar gleichmächtig zu $\mathcal{P}(N)$, und nach Induktionsvoraussetzung gilt $|\mathcal{P}(N)| = 2^n$. Insgesamt erhalten wir also

$$|\mathcal{P}(M)| = 2^n + 2^n = 2^{n+1}.$$

Damit folgt die Aussage mittels Induktion. □

Aufgaben

Aufgabe 5.14

Die Menge \mathbb{Z} der ganzen Zahlen ist abzählbar unendlich.

Aufgabe 5.15

Seien M, N Mengen. Beweise die folgenden Aussagen:

- a. Ist $f : M \rightarrow N$ eine surjektive Abbildung, dann gilt $|N| \leq |M|$.
- b. Falls $0 < |N| \leq |M| < \infty$, dann gibt es eine surjektive Abbildung $f : M \rightarrow N$.

Aufgabe 5.16

Es sei M eine endliche Menge. Zeige die folgenden Aussagen:

- a. Ist $A \subseteq M$, so ist auch A endlich und es gilt $|A| \leq |M|$.
- b. Ist $M = A \cup B$, so ist $|M| = |A| + |B|$.

§ 6 Äquivalenzrelationen

Äquivalenzrelationen stellen ein sehr wichtiges *Ordnungs-* und *Konstruktionsprinzip* innerhalb der Mathematik dar, auf das wir im Verlauf der Vorlesung an einigen zentralen Stellen zurückkommen werden, etwa im Zusammenhang mit Faktorräumen (siehe Bemerkung 23.29), der Äquivalenz von Matrizen (siehe Bemerkung 26.19) oder der Konjugation von Matrizen und der Jordanschen Normalform (siehe Bemerkung 32.3).

Definition 6.1 (Relation)

Seien M und N zwei Mengen, so nennen wir jede Teilmenge $R \subseteq M \times N$ eine *Relation* zwischen M und N .

Bemerkung 6.2

Ist R eine Relation zwischen M und N , $x \in M$ und $y \in N$, so wollen wir sagen x *steht in Relation zu y bezüglich R* , wenn $(x, y) \in R$. Die Menge R legt also fest, wann zwei Elemente in Relation zueinander stehen. Wir schreiben auch xRy statt $(x, y) \in R$.

Beispiel 6.3 (Abbildungen als Relationen)

- Der Graph einer Abbildung $f : M \rightarrow N$ ist ein Beispiel einer Relation, bei der jedes $x \in M$ zu genau einem $y \in N$ in Relation steht.
- Ist M die Menge der Hörer der Vorlesung und N die Menge der in Kaiserslautern studierbaren Fächer, so ist

$$R = \{(x, y) \in M \times N \mid x \text{ studiert } y\}$$

eine Relation zwischen M und N , die ganz sicher nicht Graph einer Funktion ist.

Bemerkung 6.4 (Motivation des Begriffs Äquivalenzrelation)

Der folgende Begriff der *Äquivalenzrelation* bereitet den Studenten oft extreme Schwierigkeiten. Dabei liegt auch ihm ein ganz einfaches Prinzip zugrunde, das wir zunächst an einem Beispiel erläutern wollen.

Die Gesamtheit aller Schüler einer Schule werden von der Schulleitung zwecks sinnvoller Organisation des Unterrichts in Schulklassen eingeteilt. Dabei achtet die Schulleitung darauf, daß jeder Schüler zu einer Schulklasse gehört und auch nur zu dieser einen. Etwas mathematischer ausgedrückt, die Schulleitung teilt die *Menge* S der Schüler in *paarweise disjunkte Teilmengen* K_i , $i = 1, \dots, k$, ein, so daß wir anschließend eine *disjunkte Zerlegung*

$$S = \bigcup_{i=1}^k K_i$$

der Menge S in die Schulklassen K_1, \dots, K_k haben. Dabei kann man für die Zugehörigkeit der Schüler Alfred, Ben und Christoph zu einer Schulklasse folgendes feststellen:

- 1) Alfred gehört zu einer Schulklasse.
- 2) Wenn Alfred in der gleichen Schulklasse ist wie Ben, dann ist Ben auch in der gleichen Schulklasse wie Alfred.
- 3) Wenn Alfred in der gleichen Schulklasse ist wie Ben und wenn zugleich Ben in der gleichen Schulklasse ist wie Christoph, dann ist auch Alfred in der gleichen Schulklasse wie Christoph.

Diese Aussagen sind so offensichtlich, daß man kaum glauben mag, daß es einen tieferen Sinn hat, sie zu erwähnen. Aber nehmen wir für einen Augenblick an, die Schulleitung hat ihre Einteilung der Schüler vorgenommen und für jede Schulklasse eine Liste mit den Namen der Schüler erstellt, die zu dieser Schulklasse gehören sollen. Nehmen wir ferner an, die Schulleitung hat noch nicht überprüft, ob jeder Schüler in genau einer Schulklasse eingeteilt ist. Dann behaupte ich, wenn man in den drei Aussagen 1)-3) die Schüler Alfred, Ben und Christoph durch beliebige Schüler ersetzt und die Aussagen richtig sind für jede Kombination der Schülernamen, dann ist sichergestellt, daß auch jeder Schüler in genau einer Schulklasse eingeteilt ist.

Als Mathematiker suchen wir nach möglichst einfachen Regeln, denen die Einteilung der Schulklassen genügen muß, um sicherzustellen, daß sie wirklich eine disjunkte Zerlegung von S ist, d.h. daß wirklich jeder Schüler in genau einer Schulklasse ist, und die Regeln 1)-3) sind genau die Regeln, die wir dazu brauchen. Wenn wir nun die Zugehörigkeit zweier Schüler x und y zur gleichen Klasse verstehen als “ x steht in Relation zu y ”, dann definieren uns die drei Regeln 1)-3) zudem eine Teilmenge von $S \times S$, nämlich die Relation

$$R = \{(x, y) \in S \times S \mid x \text{ ist in der gleichen Schulklasse wie } y\}.$$

Die Regeln 1)-3) lassen sich für Schüler $x, y, z \in S$ dann wie folgt formulieren:

- $(x, x) \in R$.
- Wenn $(x, y) \in R$, dann ist auch $(y, x) \in R$.
- Wenn $(x, y) \in R$ und $(y, z) \in R$, dann ist auch $(x, z) \in R$.

Eine solche Relation nennt man eine *Äquivalenzrelation*, man nennt Schüler der gleichen Schulklasse *äquivalent* und die Schulklassen nennt man dann auch *Äquivalenzklassen*.

Wir führen den Begriff der *Äquivalenzrelation* nun für beliebige Mengen ein.

Definition 6.5 (Äquivalenzrelation)

Es sei M eine Menge. Eine *Äquivalenzrelation* auf M ist eine Teilmenge $R \subseteq M \times M$, so daß für alle $x, y, z \in M$ gilt:

- | | |
|--|-------------------|
| R1: $(x, x) \in R$, | (“Reflexivität”) |
| R2: $(x, y) \in R \Rightarrow (y, x) \in R$, | (“Symmetrie”) |
| R3: $(x, y), (y, z) \in R \Rightarrow (x, z) \in R$. | (“Transitivität”) |

Bei Äquivalenzrelationen hat sich eine alternative Schreibweise zu $(x, y) \in R$ durchgesetzt, die auch wir im folgenden verwenden wollen.

Notation 6.6 (Schreibweise \sim für Äquivalenzrelationen)

Sei M eine Menge und R eine Äquivalenzrelation auf M . Wir definieren für $x, y \in M$

$$x \sim y \quad :\iff \quad (x, y) \in R,$$

und wir sprechen dann meist von der Äquivalenzrelation “ \sim ” statt R , sofern keine Mißverständnisse zu befürchten sind.

Mit dieser Schreibweise lassen sich die drei Axiome in Definition 6.5 wie folgt formulieren. Für $x, y, z \in M$ soll gelten:

- R1:** $x \sim x$, (“Reflexivität”)
R2: $x \sim y \Rightarrow y \sim x$, (“Symmetrie”)
R3: $x \sim y, y \sim z \Rightarrow x \sim z$. (“Transitivität”)

Definition 6.7 (Äquivalenzklassen)

Es sei M eine Menge und \sim eine Äquivalenzrelation auf M . Für $x \in M$ heißt die Menge

$$\bar{x} := \{y \in M \mid y \sim x\}$$

die *Äquivalenzklasse* von x . Jedes $y \in \bar{x}$ heißt ein *Repräsentant* der Klasse \bar{x} . Mit

$$M/\sim := \{\bar{x} \mid x \in M\}$$

bezeichnen wir die Menge der *Äquivalenzklassen modulo der Äquivalenzrelation \sim* .

Beispiel 6.8 (Der Abstand vom Ursprung als Äquivalenzrelation)

Wir betrachten die Menge $M = \mathbb{R}^2$ der Punkte in der reellen Zahlenebene und wir bezeichnen mit $|P|$ den Abstand von P zum Ursprung $(0, 0)$. Für zwei Punkte $P, Q \in M$ definieren wir

$$P \sim Q \iff |P| = |Q|,$$

d.h. wir nennen die Punkte *äquivalent*, falls ihr Abstand zum Ursprung gleich ist. Dann ist \sim eine Äquivalenzrelation.

- R1:** Sei $P \in M$, dann ist $|P| = |P|$, also $P \sim P$.
R2: Falls $P, Q \in M$ mit $P \sim Q$, dann ist $|P| = |Q|$ und somit auch $|Q| = |P|$.
 Damit gilt aber $Q \sim P$.
R3: Falls $P, Q, R \in M$ mit $P \sim Q$ und $Q \sim R$, dann gilt $|P| = |Q|$ und $|Q| = |R|$.
 Aber damit gilt auch $|P| = |R|$ und somit $P \sim R$.

Die Äquivalenzklasse

$$\bar{P} = \{Q \in M \mid |Q| = |P|\}$$

von $P \in M$ ist der Kreis um den Ursprung vom Radius $|P|$.

Wir haben anfangs behauptet, daß die drei Axiome einer Äquivalenzrelation sicherstellen, daß die zugehörigen Äquivalenzklassen eine disjunkte Zerlegung von M induzieren, und umgekehrt, daß jede disjunkte Zerlegung eine Äquivalenzrelation mit sich bringt. Dies wollen wir im Folgenden beweisen. Dazu sollten wir zunächst den Begriff disjunkt klären.

Proposition 6.9 (Die Äquivalenzrelation zu einer disjunkten Zerlegung)

Ist $(M_i)_{i \in I}$ eine disjunkte Zerlegung von M und definieren wir eine Relation auf M durch

$$x \sim y \iff \exists i \in I : x, y \in M_i,$$

dann ist \sim eine Äquivalenzrelation auf M .

Beweis: Ist $x \in M = \bigcup_{i \in I} M_i$, so gibt es ein $i \in I$ mit $x \in M_i$ und somit gilt $x \sim x$. \sim ist also reflexiv.

Sind $x, y \in M$ mit $x \sim y$, so gibt es ein $i \in I$ mit $x, y \in M_i$. Dann gilt aber auch $y \sim x$. Die Relation ist also symmetrisch.

Sind $x, y, z \in M$ mit $x \sim y$ und $y \sim z$, so gibt es $i, j \in I$ mit $x, y \in M_i$ und $y, z \in M_j$. Da die Zerlegung disjunkt ist und $y \in M_i \cap M_j$, folgt $M_i = M_j$. Also gilt $x, z \in M_i$ und somit $x \sim z$. \sim ist also auch transitiv. \square

Proposition 6.10 (Die disjunkte Zerlegung zu einer Äquivalenzrelation)

Es sei M eine Menge. Ist \sim eine Äquivalenzrelation auf M , dann bilden die Äquivalenzklassen eine disjunkte Zerlegung von M , d. h. jedes $x \in M$ liegt in genau einer Äquivalenzklasse.

Insbesondere gilt für Äquivalenzklassen \bar{x} und \bar{y} entweder $\bar{x} = \bar{y}$ oder $\bar{x} \cap \bar{y} = \emptyset$.

Beweis: Sei $x \in M$ beliebig. Aus $x \sim x$ folgt $x \in \bar{x} \subseteq \bigcup_{\bar{y} \in M/\sim} \bar{y}$. Mithin gilt

$$M = \bigcup_{\bar{y} \in M/\sim} \bar{y}.$$

Es bleibt also zu zeigen, daß die Äquivalenzklassen paarweise disjunkt sind.

Seien $\bar{x}, \bar{y} \in M/\sim$ mit $\bar{x} \cap \bar{y} \neq \emptyset$. Dann gibt es ein $z \in \bar{x} \cap \bar{y}$, und es gilt $z \sim x$ und $z \sim y$. Wegen der Symmetrie gilt aber auch $x \sim z$ und mittels der Transitivität dann $x \sim y$. Sei nun $u \in \bar{x}$ beliebig, dann gilt $u \sim x$ und wieder wegen der Transitivität $u \sim y$. Also $u \in \bar{y}$ und damit $\bar{x} \subseteq \bar{y}$. Vertauschung der Rollen von x und y in der Argumentation liefert schließlich $\bar{x} = \bar{y}$. \square

Korollar 6.11 (Äquivalenzrelationen auf endlichen Mengen)

Sei M eine endliche Menge, \sim eine Äquivalenzrelation auf M und M_1, \dots, M_s seien die paarweise verschiedenen Äquivalenzklassen von \sim . Dann gilt:

$$|M| = \sum_{i=1}^s |M_i|.$$

Beweis: Mit M sind auch alle M_i endlich und die Behauptung folgt aus Proposition 6.10 und Bemerkung 5.4. \square

Ein Beispiel aus dem Alltag für eine Äquivalenzrelation haben wir oben bereits gesehen. Ein weiteres wichtiges und wohlbekanntes Beispiel sind die rationalen Zahlen! Ein Bruch ist nichts weiter als die Äquivalenzklasse eines Tupels von ganzen Zahlen, und das Kürzen des Bruches, z.B. $\frac{1}{2} = \frac{2}{4}$, ist nur die Wahl eines möglichst einfachen Repräsentanten.

Beispiel 6.12 (Die rationalen Zahlen)

Man kann die rationalen Zahlen wie folgt als Äquivalenzklassen von Paaren ganzer Zahlen definieren. Für $(p, q), (p', q') \in M := \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ definiere

$$(p, q) \sim (p', q') \iff pq' = p'q.$$

Wir wollen nun zeigen, daß hierdurch wirklich eine Äquivalenzrelation auf M definiert wird. Seien dazu $x = (p, q), x' = (p', q'), x'' = (p'', q'') \in M$ gegeben:¹

R1: Für die Reflexivität müssen wir $x \sim x$ zeigen. Nun gilt aber $pq = pq$, woraus $x = (p, q) \sim (p, q) = x$ folgt.

R2: Für die Symmetrie nehmen wir an, daß $x \sim x'$ gilt und müssen $x' \sim x$ folgern. Wegen $x \sim x'$ gilt aber nach Definition $pq' = p'q$, und folglich auch $p'q = pq'$. Letzteres bedeutet aber, daß $x' = (p', q') \sim (p, q) = x$.

R3: Für die Transitivität nehmen wir schließlich an, daß $x \sim x'$ und $x' \sim x''$ gilt, und müssen daraus schließen, daß $x \sim x''$. Wegen $x \sim x'$ gilt nun aber $pq' = p'q$, und wegen $x' \sim x''$ gilt $p'q'' = p''q'$. Multiplizieren wir die erste der Gleichungen mit q'' und die zweite mit q , so erhalten wir

$$pq'q'' = p'q''q = p'q''q = p''q'q.$$

Da nach Voraussetzung $q' \neq 0$, können wir beide Seiten der Gleichung durch q' teilen und erhalten:

$$pq'' = p''q.$$

Das wiederum bedeutet, daß $x = (p, q) \sim (p'', q'') = x''$ gilt.

Die drei Axiome einer Äquivalenzrelation sind also erfüllt.

Wir setzen nun $\mathbb{Q} := M / \sim$ und für $(p, q) \in M$ setzen wir $\frac{p}{q} := \overline{(p, q)}$, d. h. die rationale Zahl $\frac{p}{q}$ ist die Äquivalenzklasse des Paares (p, q) unter der obigen Äquivalenzrelation. Dann bedeutet die Definition von \sim soviel wie, daß $\frac{p}{q}$ und $\frac{p'}{q'}$ gleich

¹Man sollte sich nicht dadurch verwirren lassen, daß die Elemente von M nun selbst schon Zahlenpaare sind! Wollte man die Relation als Teilmenge von $M \times M$ schreiben, so müßte man

$$R = \{((p, q), (p', q')) \in M \times M \mid pq' = p'q\}$$

betrachten. Das erläutert vielleicht auch, weshalb wir die *alternative* Schreibweise bevorzugen – solche Paare von Paaren werden doch leicht unübersichtlich.

sind, wenn die kreuzweisen Produkte von Zähler und Nenner, pq' und $p'q$, übereinstimmen, oder in der vielleicht etwas bekannteren Formulierung, wenn die Brüche nach *Erweitern* mit q' bzw. mit q übereinstimmen: $\frac{p}{q} = \frac{pq'}{qq'} \stackrel{!}{=} \frac{p'q}{q'q} = \frac{p'}{q'}$.

Auch die Rechenregeln für rationale Zahlen lassen sich mit Hilfe der Äquivalenzklassen definieren. Für $(p, q), (r, s) \in M$ definiere:

$$\begin{aligned}\overline{(p, q)} + \overline{(r, s)} &:= \overline{(ps + qr, qs)}, \\ \overline{(p, q)} \cdot \overline{(r, s)} &:= \overline{(pr, qs)}.\end{aligned}$$

In Anlehnung an unser erstes Beispiel, der Einteilung der Schüler in Schulklassen, kann man das obige Rechenprinzip als “Rechnen mit Klassen” bezeichnen. Will man zwei Klassen addieren (bzw. multiplizieren), so nimmt man aus jeder der Klasse ein Element, addiert (bzw. multipliziert) diese Elemente und schaut, in welche Klasse das Resultat gehört. Diese Klasse ist dann die Summe (bzw. das Produkt) der beiden Klassen.

Was man sich bei diesem Vorgehen allerdings klar machen muß, ist, daß das Ergebnis nicht von der Wahl der Repräsentanten (d.h. der Elemente aus den Klassen) abhängt. Man spricht davon, daß die Operation *wohldefiniert* ist. Wir führen das für die Addition der rationalen Zahlen vor.

Sind $(p', q') \in \overline{(p, q)}$ und $(r', s') \in \overline{(r, s)}$ andere Repräsentanten, dann gilt $p'q = q'p$ und $r's = s'r$. Es ist zu zeigen, daß $(p's' + q'r', q's') \in \overline{(ps + qr, qs)}$ gilt. Ausmultiplizieren liefert

$$(p's' + q'r')(qs) = p'qs's + q'qr's = q'ps's + q'qs'r = (ps + qr)(q's'),$$

was zu zeigen war. □

Aufgaben

Aufgabe 6.13

Wir definieren für zwei Punkte $(x, y), (x', y') \in \mathbb{R}^2$

$$(x, y) \sim (x', y') \quad :\iff \quad |x| + |y| = |x'| + |y'|.$$

Zeige, \sim ist eine Äquivalenzrelation auf \mathbb{R}^2 . Zeichne die Äquivalenzklassen zu $(1, 1)$ und zu $(-2, 3)$ in die Zahlenebene \mathbb{R}^2 ein.

Aufgabe 6.14 (Die ganzen Zahlen)

Es sei $M = \mathbb{N} \times \mathbb{N}$ und $m = (a, b) \in M$ und $m' = (a', b') \in M$ seien zwei Elemente in M . Wir definieren

$$m \sim m' \quad \iff \quad a + b' = a' + b.$$

Zeige, daß \sim eine Äquivalenzrelation ist und daß die folgende Abbildung bijektiv ist:

$$\Phi : \mathbb{Z} \longrightarrow M / \sim : z \mapsto \begin{cases} \overline{(z, 0)}, & \text{falls } z \geq 0, \\ \overline{(0, -z)}, & \text{falls } z < 0. \end{cases}$$

Aufgabe 6.15 (Die projektive Gerade)

Wir definieren für $v = (v_1, v_2), w = (w_1, w_2) \in \mathbb{R}^2 \setminus \{(0, 0)\}$

$$v \sim w \iff \exists \lambda \in \mathbb{R} \setminus \{0\} : v = \lambda \cdot w$$

wobei $\lambda \cdot w := (\lambda \cdot w_1, \lambda \cdot w_2)$.

- Zeige, daß \sim eine Äquivalenzrelation auf $M = \mathbb{R}^2 \setminus \{(0, 0)\}$ ist. Es ist üblich die Äquivalenzklasse $\overline{(v_1, v_2)}$ von (v_1, v_2) mit $(v_1 : v_2)$ zu bezeichnen, und man nennt die Menge M / \sim der Äquivalenzklassen die *projektive Gerade* über \mathbb{R} und bezeichnet sie mit $\mathbb{P}_{\mathbb{R}}^1$.
- Die Menge $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ ist Kreis vom Radius Eins um den Mittelpunkt $(0, 0)$. Zeige, daß die Abbildung

$$\Phi : S^1 \longrightarrow \mathbb{P}_{\mathbb{R}}^1 : (x, y) \mapsto \overline{(x, y)}$$

surjektiv ist.

- Wenn wir in der Definition von \sim alle Elemente $v, w \in \mathbb{R}^2$ zulassen, definiert \sim dann eine Äquivalenzrelation auf \mathbb{R}^2 ? Falls ja, was ist die Äquivalenzklasse von $(0, 0)$?

Aufgabe 6.16 (Kongruenz modulo n)

Ist $n \in \mathbb{Z}_{>0}$ eine positive ganze Zahl, so definieren wir für $x, y \in \mathbb{Z}$

$$x \equiv y \iff x - y \text{ ist ein Vielfaches von } n.$$

Zeige, daß \equiv eine Äquivalenzrelation ist mit genau den n paarweise verschiedenen Äquivalenzklassen $\overline{0}, \overline{1}, \dots, \overline{n-1}$.

Man nennt zwei äquivalente Zahlen x und y dann auch *kongruent modulo n* . Diese Äquivalenzrelation wird in der Vorlesung algebraische Strukturen genauer untersucht.

§ 7 Gruppen und Körper

A) Gruppen

Definition 7.1 (Gruppen)

- a. Eine *Gruppe* ist ein Paar $(G, *)$ bestehend aus einer *nicht-leeren* Menge G und einer zweistelligen Operation “*”, d. h. einer Abbildung

$$*: G \times G \rightarrow G : (g, h) \mapsto g * h,$$

so daß die folgenden *Gruppenaxiome* gelten:

$$\mathbf{G1:} \quad (g * h) * k = g * (h * k) \quad \forall g, h, k \in G, \quad (\text{“Assoziativgesetz”})$$

$$\mathbf{G2:} \quad \exists e \in G : \forall g \in G : e * g = g, \quad (\text{“Existenz eines Neutralen”})$$

$$\mathbf{G3:} \quad \forall g \in G \exists g^{-1} \in G : g^{-1} * g = e. \quad (\text{“Existenz von Inversen”})$$

Ein Element mit der Eigenschaft von e nennt man *neutrales Element* der Gruppe G . Ein Element mit der Eigenschaft von g^{-1} nennt man ein *Inverses zu g* .

- b. Eine Gruppe $(G, *)$ heißt *abelsch* oder *kommutativ*, wenn $(G, *)$ zudem noch dem folgenden Axiom genügt:

$$\mathbf{G4:} \quad g * h = h * g \quad \forall g, h \in G \quad (\text{“Kommutativgesetz”})$$

Beispiel 7.2

- a. $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$ und $(\mathbb{R}, +)$ mit der üblichen Addition als Gruppenoperation sind abelsche Gruppen. Die Zahl Null erfüllt jeweils die Rolle eines neutralen Elements, und zu einer Zahl g existiert mit $-g$ ein inverses Element.
- b. $(\mathbb{Q} \setminus \{0\}, \cdot)$ und $(\mathbb{R} \setminus \{0\}, \cdot)$ mit der üblichen Multiplikation als Gruppenoperation sind ebenfalls abelsche Gruppen. Die Zahl 1 ist jeweils ein neutrales Element, und zu einer Zahl g existiert als inverses Element die Zahl $\frac{1}{g}$.
- c. Ist M eine Menge, so ist die Menge

$$\text{Sym}(M) := \{f : M \longrightarrow M \mid f \text{ ist bijektiv}\}$$

mit der Komposition von Abbildungen als Gruppenoperation eine Gruppe. Die Assoziativität von “ \circ ” haben wir in Proposition 3.11 gezeigt, die Identität ist das neutrale Element und in Satz 3.12 haben wir gezeigt, daß jede bijektive Abbildung ein Inverses besitzt. Wir nennen $(\text{Sym}(M), \circ)$ die *symmetrische Gruppe* auf M . Enthält M mehr als zwei Elemente, so ist $\text{Sym}(M)$ nicht abelsch.

Bemerkung 7.3

Es sei $(G, *)$ eine Gruppe.

- a. Das neutrale Element $e \in G$ ist eindeutig bestimmt und hat die Eigenschaft:

$$e * g = g * e = g \quad \forall g \in G.$$

- b. Sei $g \in G$. Das inverse Element g^{-1} zu g ist eindeutig bestimmt und hat die Eigenschaft:

$$g^{-1} * g = g * g^{-1} = e.$$

- c. Für $g, h \in G$ gelten $(g^{-1})^{-1} = g$ und $(g * h)^{-1} = h^{-1} * g^{-1}$.
- d. Wird die Gruppenoperation als Multiplikation und mit “ \cdot ” bezeichnet, so schreiben wir für das Neutrale Element meist 1 und für das Inverse zu g weiterhin g^{-1} oder $\frac{1}{g}$.
Wird die Gruppenoperation als Addition und mit “ $+$ ” bezeichnet, so schreiben wir für das Neutrale Element meist 0 und für das Inverse zu g meist $-g$. Zudem schreiben wir statt $g + (-h)$ in aller Regel $g - h$.
- e. In Ermangelung eines besseren Namens nennen wir auch “ $*$ ” oft einfach die *Gruppenmultiplikation*.

Die Aussagen in der Bemerkung werden in der Vorlesung Algebraische Strukturen bewiesen. Für den interessierten Leser fügen wir hier einen Beweis ein.

Beweis von Bemerkung 7.3: Da wir für das Paar $(G, *)$ die Axiome G1-G3 aus Definition 7.1 voraussetzen, gibt es ein neutrales Element $e \in G$, und zu beliebigem, aber fest gegebenem $g \in G$ gibt es ein Inverses $g^{-1} \in G$.

Wir wollen zunächst zeigen, daß für dieses e und dieses g^{-1} die in a. und b. geforderten zusätzlichen Eigenschaften gelten.

Da $(G, *)$ eine Gruppe ist, gibt es ein $(g^{-1})^{-1} \in G$ mit

$$(g^{-1})^{-1} * g^{-1} = e. \quad (2)$$

Also folgt:

$$\begin{aligned} g * g^{-1} &\stackrel{G2}{=} e * (g * g^{-1}) \stackrel{(2)}{=} ((g^{-1})^{-1} * g^{-1}) * (g * g^{-1}) \stackrel{G1}{=} (g^{-1})^{-1} * (g^{-1} * (g * g^{-1})) \\ &\stackrel{G1}{=} (g^{-1})^{-1} * ((g^{-1} * g) * g^{-1}) \stackrel{G3}{=} (g^{-1})^{-1} * (e * g^{-1}) \stackrel{G2}{=} (g^{-1})^{-1} * g^{-1} \stackrel{(2)}{=} e. \end{aligned} \quad (3)$$

Damit ist gezeigt, daß g^{-1} die zusätzliche Eigenschaft in b. erfüllt, und wir erhalten:

$$g * e \stackrel{G3}{=} g * (g^{-1} * g) \stackrel{G1}{=} (g * g^{-1}) * g \stackrel{(3)}{=} e * g \stackrel{G2}{=} g. \quad (4)$$

Nun war aber g ein beliebiges Element in G , so daß damit die zusätzliche Eigenschaft von e in a. gezeigt ist.

Sei nun $\tilde{e} \in G$ irgendein Element mit der Eigenschaft des Neutralen, d.h.

$$\tilde{e} * h = h \quad (5)$$

für alle $h \in G$. Wir müssen zeigen, daß $e = \tilde{e}$ gilt. Da wir bereits wissen, daß e die zusätzliche Eigenschaft in a. erfüllt, können wir diese, d.h. (4), mit \tilde{e} in der Rolle von g anwenden, und anschließend (5) mit e in der Rolle von h :

$$\tilde{e} \stackrel{(4)}{=} \tilde{e} * e \stackrel{(5)}{=} e.$$

Schließlich müssen wir noch zeigen, wenn $\tilde{g}^{-1} \in G$ ein weiteres inverses Element zu g ist, d.h. wenn

$$\tilde{g}^{-1} * g = e \quad (6)$$

gilt, dann ist schon $g^{-1} = \tilde{g}^{-1}$. Wenden wir das bislang Gezeigte an, so gilt:

$$\tilde{g}^{-1} \stackrel{(4)}{=} \tilde{g}^{-1} * e \stackrel{(3)}{=} \tilde{g}^{-1} * (g * g^{-1}) \stackrel{G1}{=} (\tilde{g}^{-1} * g) * g^{-1} \stackrel{(6)}{=} e * g^{-1} \stackrel{G2}{=} g^{-1}.$$

Damit sind die Aussagen in Teil a. und b. gezeigt und es bleibt noch, die Aussagen in Teil c. zu zeigen.

Um die erste Gleichheit zu zeigen, reicht es wegen der Eindeutigkeit des Inversen zu g^{-1} zu zeigen, daß g die Eigenschaft *des* Inversen zu g^{-1} besitzt. Beim Beweis können wir die Gruppenaxiome sowie die in a. und b. bewiesenen zusätzlichen Eigenschaften des Inversen anwenden:

$$g * g^{-1} \stackrel{b.}{=} e.$$

Also ist g ein Inverses zu g^{-1} , und damit gilt wie angedeutet wegen der Eindeutigkeit des Inversen zu g^{-1} :

$$(g^{-1})^{-1} = g.$$

Analog ist nach Voraussetzung $(gh)^{-1}$ ein Inverses zu gh , und es reicht wegen der Eindeutigkeit des Inversen zu gh zu zeigen, daß $h^{-1}g^{-1}$ ebenfalls die Eigenschaft eines Inversen zu gh hat:

$$\begin{aligned} (h^{-1} * g^{-1}) * (g * h) &\stackrel{G1}{=} h^{-1} * (g^{-1} * (g * h)) \stackrel{G1}{=} h^{-1} * ((g^{-1} * g) * h) \\ &\stackrel{G3}{=} h^{-1} * (e * h) \stackrel{G2}{=} h^{-1} * h \stackrel{G3}{=} e. \end{aligned}$$

Mithin ist $h^{-1} * g^{-1}$ ein Inverses zu gh , und somit

$$(g * h)^{-1} = h^{-1} * g^{-1}.$$

Damit sind nun alle Aussagen der Bemerkung bewiesen. □

Lemma 7.4 (Kürzungsregeln)

Sei $(G, *)$ eine Gruppe, $g, a, b \in G$. Dann gelten die Kürzungsregeln:

- a. $g * a = g * b \Rightarrow a = b$, und
- b. $a * g = b * g \Rightarrow a = b$.

Beweis: Die erste Kürzungsregel folgt durch Multiplikation mit dem Inversen zu g von links:

$$\begin{aligned} a \stackrel{G2}{=} e * a \stackrel{G3}{=} (g^{-1} * g) * a \stackrel{G1}{=} g^{-1} * (g * a) \\ \stackrel{\text{Vor.}}{=} g^{-1} * (g * b) \stackrel{G1}{=} (g^{-1} * g) * b \stackrel{G3}{=} e * b \stackrel{G2}{=} b. \end{aligned}$$

Entsprechend folgt die zweite Kürzungsregel durch Multiplikation mit g^{-1} von rechts und unter Berücksichtigung der zusätzlichen Eigenschaft des Inversen in Bemerkung 7.3. Die Details überlassen wir dem Leser. □

B) Körper**Definition 7.5** (Körper)

Ein *Körper* ist ein Tripel $(K, +, \cdot)$ bestehend aus einer Menge K zusammen mit zwei zweistelligen Operationen

$$+ : K \times K \rightarrow K : (x, y) \mapsto x + y, \quad (\text{“Addition”})$$

und

$$\cdot : K \times K \rightarrow K : (x, y) \mapsto x \cdot y, \quad (\text{“Multiplikation”})$$

so daß folgende Axiome erfüllt sind:

- $(K, +)$ ist eine abelsche Gruppe mit neutralem Element 0 .
- $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe mit neutralem Element 1 .
- Es gilt das *Distributivgesetz* $x \cdot (y + z) = x \cdot y + x \cdot z$ für $x, y, z \in K$.

Ist eine Teilmenge $L \subseteq K$ eines Körpers mit den *gleichen* Operationen wieder selbst ein Körper, so nennen wir L einen *Teilkörper* von K .

Beispiel 7.6 (Die endlichen Körper \mathbb{F}_p)

- Die rationalen Zahlen $(\mathbb{Q}, +, \cdot)$ und die reellen Zahlen $(\mathbb{R}, +, \cdot)$ mit der üblichen Addition und Multiplikation sind Körper. \mathbb{Q} ist ein Teilkörper von \mathbb{R} .
- Die ganzen Zahlen $(\mathbb{Z}, +, \cdot)$ sind kein Körper, da z.B. der Zahl 2 ein multiplikatives Inverses fehlt.
- Auf der Menge $\mathbb{F}_2 := \{0, 1\}$ definieren wir zwei Operationen durch folgende Additions- und Multiplikationstabellen:

$$\begin{array}{c|c|c} + & 0 & 1 \\ \hline 0 & 0 & 1 \\ \hline 1 & 1 & 0 \end{array} \qquad \begin{array}{c|c|c} \cdot & 0 & 1 \\ \hline 0 & 0 & 0 \\ \hline 1 & 0 & 1 \end{array}$$

Mit ein wenig Aufwand kann man nachrechnen, daß alle Körperaxiome erfüllt sind und daß mithin \mathbb{F}_2 ein Körper ist. \mathbb{F}_2 ist der kleinstmögliche Körper, da nach Definition ein Körper stets mindestens zwei Elemente, nämlich ein Neutrales bezüglich der Addition und ein davon verschiedenes Neutrales bezüglich der Multiplikation enthalten muß. Man beachte auch, daß aufgrund von Lemma 7.8 keine andere Möglichkeit für die obigen Verknüpfungstabellen besteht, wenn man einen Körper mit genau zwei Elementen haben möchte.

- Allgemeiner zeigt man in der Vorlesung Algebraische Strukturen, daß man für eine Primzahl p die Menge

$$\mathbb{F}_p := \{0, 1, \dots, p-1\}$$

auf folgende Weise zu einem Körper machen kann. Für eine natürliche Zahl $a \in \mathbb{N}$ können wir Division mit Rest durch die Zahl p durchführen. Wir erhalten

dann eindeutig bestimmte Zahlen $q \in \mathbb{N}$ und $0 \leq r < p$ mit

$$a = q \cdot p + r.$$

Die Zahl r heißt der Rest von a bei Division mit Rest durch p , und wir bezeichnen sie $r(a : p)$.

Mit dieser Notation definieren wir für zwei Zahlen $a, b \in \mathbb{F}_p$

$$a + b := r(a + b : p)$$

und

$$a \cdot b := r(a \cdot b : p),$$

wobei das “+” bzw. das “.” auf der rechten Seite jeweils die Operation in den ganzen Zahlen bezeichnet, während das “+” und das “.” auf der linken Seite neu definierte Operationen sind. Formal wäre es besser, für diese neuen Operationen neue Symbole zu verwenden, etwa “ \oplus ” und “ \otimes ”, aber Mathematiker sind bequeme Menschen und schreiben nur ungerne mehr als nötig. Deshalb bleiben wir bei den bewährten Symbolen und müssen nur drauf achten, wo wir gerade rechnen. Jedenfalls gilt, daß \mathbb{F}_p mit diesen beiden Operationen ein Körper ist.

Man beachte auch, daß in \mathbb{F}_p für jede Primzahl p stets

$$\underbrace{1 + 1 + \dots + 1}_{p\text{-mal}} = r(p : p) = 0$$

gilt! Damit ist auch das Negative einer Zahl $a \in \mathbb{F}_p$ leicht zu berechnen als $p - a$, hingegen ist das multiplikative Inverse $\frac{1}{a}$ einer Zahl $0 \neq a \in \mathbb{F}_p$ nicht so ohne weiteres anzugeben. Man lernt in den Algebraischen Strukturen, wie man dieses mit Hilfe des Euklidischen Algorithmus’ berechnen kann.

Z.B., gilt in \mathbb{F}_5

$$3 + 4 = r(3 + 4 : 5) = r(7 : 5) = 2$$

und

$$3 \cdot 4 = r(3 \cdot 4 : 5) = r(12 : 5) = 2.$$

In den Algebraischen Strukturen schreibt man übrigens meist $\mathbb{Z}/p\mathbb{Z}$ oder \mathbb{Z}_p anstatt \mathbb{F}_p , und die Zahl a wird dort meist mit \bar{a} oder $[a]$ bezeichnet. Das liegt daran, daß man den Körper mit der Menge der Äquivalenzklassen der Kongruenz modulo p identifizieren kann (siehe Aufgabe 6.16).

Notation 7.7

Ist K ein Körper und sind $x, y, z \in K$ mit $z \neq 0$, so schreiben wir statt $x + (-y)$ in aller Regel $x - y$, und statt $x \cdot z^{-1}$ schreiben wir oft $\frac{x}{z}$. Außerdem schreiben wir statt $x \cdot y$ meist nur xy .

Lemma 7.8 (Rechenregeln)

Es sei K ein Körper, $x, y, z \in K$ und $u, v \in K \setminus \{0\}$.

a. $-(-x) = x,$

- b. $x + y = z \Leftrightarrow x = z - y$,
- c. $-(x + y) = -x - y$,
- d. $0 \cdot x = x \cdot 0 = 0$,
- e. $(-x) \cdot y = x \cdot (-y) = -(x \cdot y)$,
- f. $(-x) \cdot (-y) = x \cdot y$,
- g. $x \cdot (y - z) = x \cdot y - x \cdot z$.
- h. $(x^{-1})^{-1} = x$, für $x \neq 0$,
- i. $x \cdot y = 0 \Leftrightarrow x = 0$ oder $y = 0$,
- j. $z \cdot x = z \cdot y$, $z \neq 0 \Rightarrow x = y$,
- k. $\frac{x}{u} \cdot \frac{y}{v} = \frac{x \cdot y}{u \cdot v}$,
- l. $\frac{x}{u} + \frac{y}{v} = \frac{x \cdot v + y \cdot u}{u \cdot v}$.

Beweis: Die Aussagen a., b., c. und h. folgen unmittelbar aus Bemerkung 7.3 und Lemma 7.4.

d. Für $x \in K$ gilt $0 \cdot x = (0 + 0) \cdot x = 0 \cdot x + 0 \cdot x$, also folgt $0 \cdot x = 0$ mittels der Kürzungsregeln in $(K, +)$. Analog sieht man $x \cdot 0 = 0$.

e. Für $x, y \in K$ gilt wegen d.:

$$x \cdot y + (-x) \cdot y = (x - x) \cdot y = 0 \cdot y = 0,$$

also $-(x \cdot y) = (-x) \cdot y$. Die Gleichheit des Ausdrucks zu $x \cdot (-y)$ folgt analog.

f. Für $x, y \in K$ folgt unter Zuhilfenahme von a. und e.:

$$(-x) \cdot (-y) = -(x \cdot (-y)) = -(- (x \cdot y)) = x \cdot y.$$

g. Für $x, y, z \in K$ impliziert e.:

$$x \cdot (y - z) = x \cdot y + x \cdot (-z) = x \cdot y + (- (x \cdot z)) = x \cdot y - x \cdot z.$$

i. Ist $x = 0$ oder $y = 0$, so ist nach d. auch $x \cdot y = 0$. Ist $x \neq 0$ und $y \neq 0$, so ist $x \cdot y \in K \setminus \{0\}$, da $K \setminus \{0\}$ bezüglich der Multiplikation abgeschlossen ist.

j. Die Aussage zeigt man genau wie die Kürzungsregeln für Gruppen (siehe Lemma 7.4).

k. Unter Beachtung der Assoziativität und Kommutativität der Multiplikation sowie der Notation 7.7 gilt

$$\frac{x}{u} \cdot \frac{y}{v} = (x \cdot u^{-1}) \cdot (y \cdot v^{-1}) = (x \cdot y) \cdot (u \cdot v)^{-1} = \frac{x \cdot y}{u \cdot v}.$$

l. Dies geht analog zu k. mit etwas mehr Schreiarbeit.

□

Notation 7.9 (Produkte und Summen)

Es sei K ein Körper und $x_0, \dots, x_n \in K$ seien $n + 1$ Elemente in K , $n \in \mathbb{N}$. Wir schreiben

$$\prod_{i=0}^n x_i := x_0 \cdot \dots \cdot x_n$$

für das *Produkt* der Zahlen x_0, \dots, x_n und

$$\sum_{i=0}^n x_i := x_0 + \dots + x_n$$

für die *Summe* der Zahlen x_0, \dots, x_n .

Außerdem definieren wir für $x \in K$ und $n \in \mathbb{N}$ die *Potenzen* von x durch

$$x^n := \underbrace{x \cdot \dots \cdot x}_{n\text{-mal}} = \prod_{i=1}^n x$$

falls $n \geq 1$ sowie $x^0 := 1$. Ist zudem $x \neq 0$, so definieren wir

$$x^{-n} := (x^{-1})^n = \underbrace{x^{-1} \cdot \dots \cdot x^{-1}}_{n\text{-mal}} = \frac{1}{x^n}.$$

Analog dazu setzen wir

$$n \cdot x := \underbrace{x + \dots + x}_{n\text{-mal}} = \sum_{i=1}^n x$$

und

$$(-n) \cdot x := n \cdot (-x) = \underbrace{(-x) + \dots + (-x)}_{n\text{-mal}}$$

für $n \geq 1$, sowie $0 \cdot x = 0$.

Bemerkung 7.10 (Rekursionsprinzip)

Dem Prinzip der vollständigen Induktion ist das *Rekursionsprinzip* eng verwandt. Wollen wir einen Ausdruck für alle natürlichen Zahlen definieren, so definieren wir ihn für die Zahl 0 und führen die Definition für die Zahl n auf die Definition für die Zahl $n - 1$ zurück.

Die Notation mit Punkten “...” in Notation 7.9 ist stets eine versteckte Induktion oder Rekursion. Formal korrekt wäre es das Produkt rekursiv zu definieren durch $\prod_{i=0}^0 x_i := x_0$ und $\prod_{i=0}^n x_i := \left(\prod_{i=0}^{n-1} x_i\right) \cdot x_n$. Analog sollte man die Summe rekursiv definieren durch $\sum_{i=0}^0 x_i := x_0$ und $\sum_{i=0}^n x_i := \left(\sum_{i=0}^{n-1} x_i\right) + x_n$. Und für die Definition von x^n und $n \cdot x$ gilt Entsprechendes.

Beispiel 7.11 (Gauß)

Die Summe der natürlichen Zahlen bis zu einer gegebenen Zahl n ist

$$\sum_{k=0}^n k = \frac{n \cdot (n + 1)}{2}.$$

Man beweist die Aussage durch Induktion nach n , wobei sie für $n = 0$ offenbar richtig ist. Nehmen wir nun an, daß sie für n gilt, so folgt

$$\sum_{k=0}^{n+1} k = \sum_{k=0}^n k + (n+1) \stackrel{\text{Ind}}{=} \frac{n \cdot (n+1)}{2} + (n+1) = \frac{(n+1) \cdot (n+2)}{2}.$$

Also gilt die Aussage für alle $n \in \mathbb{N}$ nach dem Prinzip der vollständigen Induktion.

Satz 7.12 (Endliche geometrische Reihe)

Ist K ein Körper, $1 \neq q \in K$ und $n \in \mathbb{N}$, so gilt

$$\sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q}.$$

Beweis: Der Beweis ist eine einfache Anwendung des Prinzips der vollständigen Induktion. \square

Definition 7.13 (Fakultät)

Für eine natürliche Zahl $n \in \mathbb{N}$ definieren wir die *Fakultät* durch

$$n! := \prod_{i=1}^n i = 1 \cdot \dots \cdot n,$$

falls $n \geq 1$, und durch $0! := 1$.

Für zwei natürliche Zahlen $k, n \in \mathbb{N}$ erklären wir den *Binomialkoeffizienten* von n über k durch

$$\binom{n}{k} := \frac{n!}{(n-k)! \cdot k!} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-k+1)}{k \cdot (k-1) \cdot \dots \cdot 1},$$

falls $0 \leq k \leq n$, und durch $\binom{n}{k} := 0$ sonst.

Proposition 7.14 (Binomialkoeffizienten)

Es seien $n, k \in \mathbb{N}$ natürliche Zahlen. Dann gilt

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}.$$

Beweis: Wir unterscheiden mehrere Fälle.

1. Fall: $k = 0$:

$$\binom{n+1}{k} = \binom{n+1}{0} = 1 = 0 + 1 = \binom{n}{-1} + \binom{n}{0} = \binom{n}{k-1} + \binom{n}{k}.$$

2. Fall: $k = n + 1$:

$$\binom{n+1}{k} = \binom{n+1}{n+1} = 1 = 1 + 0 = \binom{n}{n} + \binom{n}{n+1} = \binom{n}{k-1} + \binom{n}{k}.$$

3. Fall: $k < 0$ oder $k > n + 1$:

$$\binom{n+1}{k} = 0 = 0 + 0 = \binom{n}{k-1} + \binom{n}{k}.$$

4. Fall: $1 \leq k \leq n$:

$$\begin{aligned} \binom{n}{k-1} + \binom{n}{k} &= \frac{n!}{(n+1-k)! \cdot (k-1)!} + \frac{n!}{(n-k)! \cdot k!} \\ &= \frac{n! \cdot k}{(n+1-k)! \cdot k!} + \frac{n! \cdot (n+1-k)}{(n+1-k)! \cdot k!} \\ &= \frac{n! \cdot (k+n+1-k)}{(n+1-k)! \cdot k!} = \frac{(n+1)!}{(n+1-k)! \cdot k!} = \binom{n+1}{k}. \end{aligned}$$

□

Satz 7.15 (Binomischer Lehrsatz)

Es sei K ein Körper, $x, y \in K$ und $n \in \mathbb{N}$, so gilt

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} \cdot x^k \cdot y^{n-k}.$$

Beweis: Wir führen den Beweis durch Induktion nach n .

Induktionsanfang: $n = 0$: Nach Definition gilt

$$(x + y)^0 = 1 = 1 \cdot 1 \cdot 1 = \sum_{k=0}^0 \binom{0}{k} \cdot x^k \cdot y^{0-k}.$$

Induktionsschluß: $n \mapsto n + 1$: Es gilt

$$\begin{aligned} (x + y)^{n+1} &= (x + y)^n \cdot (x + y) = (x + y)^n \cdot x + (x + y)^n \cdot y \\ &\stackrel{\text{Ind.}}{=} \sum_{k=0}^n \binom{n}{k} \cdot x^{k+1} \cdot y^{n-k} + \sum_{k=0}^n \binom{n}{k} \cdot x^k \cdot y^{n+1-k} \\ &= x^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k} \cdot x^{k+1} \cdot y^{n-k} + \sum_{k=1}^n \binom{n}{k} \cdot x^k \cdot y^{n+1-k} + y^{n+1} \\ &= x^{n+1} + \sum_{k=1}^n \binom{n}{k-1} \cdot x^k \cdot y^{n+1-k} + \sum_{k=1}^n \binom{n}{k} \cdot x^k \cdot y^{n+1-k} + y^{n+1} \\ &\stackrel{7.14}{=} x^{n+1} + \sum_{k=1}^n \binom{n+1}{k} \cdot x^k \cdot y^{n+1-k} + y^{n+1} \\ &= \sum_{k=0}^{n+1} \binom{n+1}{k} \cdot x^k \cdot y^{n+1-k} \end{aligned}$$

Die Aussage folgt damit aus dem Prinzip der vollständigen Induktion. □

Bemerkung 7.16 (Pascalsches Dreieck)

Man ordnet die Binomialkoeffizienten gerne in der folgenden Form an, die als Pascalsches Dreieck bekannt ist:

$$\begin{array}{cccccc}
 & & & & & \binom{0}{0} \\
 & & & & & \binom{1}{0} & \binom{1}{1} \\
 & & & & \binom{2}{0} & \binom{2}{1} & \binom{2}{2} \\
 & & \binom{3}{0} & \binom{3}{1} & \binom{3}{2} & \binom{3}{3} \\
 \binom{4}{0} & \binom{4}{1} & \binom{4}{2} & \binom{4}{3} & \binom{4}{4}
 \end{array}$$

Berechnet man die Werte der Binomialkoeffizienten, erhält man die folgende Gestalt:

$$\begin{array}{r}
 0. \text{ Zeile:} \qquad \qquad \qquad 1 \\
 1. \text{ Zeile:} \qquad \qquad \qquad 1 \quad 1 \\
 2. \text{ Zeile:} \qquad \qquad \qquad 1 \quad 2 \quad 1 \\
 3. \text{ Zeile:} \qquad \qquad \qquad 1 \quad 3 \quad 3 \quad 1 \\
 4. \text{ Zeile:} \quad 1 \quad 4 \quad 6 \quad 4 \quad 1
 \end{array}$$

Aufgrund von Proposition 7.14 kann man die Einträge der $n + 1$ -ten Zeile aus den Einträgen der n -ten Zeile berechnen. Graphisch im Pascalschen Dreieck nimmt die Proposition folgende Gestalt an:

$$\begin{array}{ccc}
 \binom{n}{k-1} & + & \binom{n}{k} \\
 \swarrow & & \searrow \\
 & \binom{n+1}{k} &
 \end{array}$$

D.h. die Summe zweier benachbarter Einträge der n -ten Zeile liefert den mittig unter ihnen stehenden Eintrag der $n + 1$ -ten Zeile.

Aufgrund des binomischen Lehrsatzes sind die Einträge der n -ten Zeile des Pascalschen Dreiecks genau die Koeffizienten, die wir erhalten, wenn wir $(x + y)^n$ aus schreiben. Z.B.

$$(x + y)^3 = 1 \cdot x^3 + 3 \cdot x^2y + 3 \cdot xy^2 + 1 \cdot y^3.$$

Aufgaben**Aufgabe 7.17**

Es sei K ein Körper und $x \in K$. Zeige, $x^2 = 1$ genau dann, wenn $x \in \{1, -1\}$.

Aufgabe 7.18

a. Auf der Menge $G := \mathbb{R} \times \mathbb{R}$ definieren wir eine zweistellige Operation

$$+ : G \times G \longrightarrow G : ((x, y), (u, v)) \mapsto (x + u, y + v).$$

Zeige, $(G, +)$ ist eine abelsche Gruppe mit neutralem Element $(0, 0)$.

b. Auf der Menge $H := (\mathbb{R} \times \mathbb{R}) \setminus \{(0, 0)\}$ definieren wir eine zweistellige Operation

$$\cdot : H \times H \longrightarrow H : ((x, y), (u, v)) \mapsto (xu - yv, xv + yu).$$

Zeige, (H, \cdot) ist eine abelsche Gruppe mit neutralem Element $(1, 0)$.

c. Zeige, daß $(\mathbb{R} \times \mathbb{R}, +, \cdot)$ ein Körper ist, wenn die Operationen “+” und “ \cdot ” wie in a. und b. definiert sind.

Aufgabe 7.19

Zeige durch vollständige Induktion, daß

$$\sum_{k=0}^n (k+1) \cdot \binom{n}{k} = 2^{n-1} \cdot (n+2)$$

für $n \in \mathbb{N}$ gilt.

Aufgabe 7.20 (Die projektive Gerade als Gruppe)

Wir haben in Aufgabe 6.15 die Projektive Gerade $\mathbb{P}_{\mathbb{R}}^1$ als Menge von Äquivalenzklassen auf $\mathbb{R}^2 \setminus \{(0, 0)\}$ eingeführt.

Zeige, daß die zweistellige Operation

$$(v_1 : v_2) \cdot (w_1 : w_2) := (v_1 \cdot w_1 - v_2 \cdot w_2 : v_1 \cdot w_2 + v_2 \cdot w_1).$$

wohldefiniert ist, d.h. nicht von der Wahl der Repräsentanten für die Äquivalenzklasse abhängt, und daß $\mathbb{P}_{\mathbb{R}}^1$ mit dieser Operation eine Gruppe ist.

§ 8 Ordnungsrelationen

A) Ordnungsrelationen

Definition 8.1 (Ordnungsrelation)

Es sei M eine Menge. Eine *Ordnungsrelation* auf M , auch *Halbordnung* oder *partielle Ordnung* genannt, ist eine Relation $R \subseteq M \times M$, so daß für alle $x, y, z \in M$ gilt:

- O1:** $(x, x) \in R$, (“Reflexivität”)
O2: $(x, y), (y, x) \in R \implies x = y$, (“Antisymmetrie”)
O3: $(x, y), (y, z) \in R \implies (x, z) \in R$. (“Transitivität”)

Beispiel 8.2

Es sei $M = \mathbb{N}$.

- a. Die übliche Größerrelation

$$R = \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid x \leq y\}$$

ist eine Ordnungsrelation auf \mathbb{N} .

- b. Die Relation

$$R = \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid x \text{ teilt } y\}$$

ist eine weitere Ordnungsrelation auf \mathbb{N} (siehe Aufgabe 8.21).

Notation 8.3 (Schreibweise \leq für Ordnungsrelationen)

Es sei M eine Menge und R eine Ordnungsrelation auf M . Wir definieren für $x, y \in M$

$$x \leq y :\Leftrightarrow (x, y) \in R,$$

und sprechen in aller Regel von der Ordnungsrelation “ \leq ” statt R , sofern keine Mißverständnisse zu befürchten sind. Ferner sprechen wir von der *partiell* oder *teilgeordneten Menge* (M, \leq) .

Mit dieser Schreibweise lassen sich die drei Axiome in Definition 8.1 wie folgt formulieren. Für $x, y, z \in M$ soll gelten:

- O1:** $x \leq x$, (“Reflexivität”)
O2: $x \leq y \wedge y \leq x \implies x = y$, (“Antisymmetrie”)
O3: $x \leq y \wedge y \leq z \implies x \leq z$. (“Transitivität”)

Gilt für $x, y \in M$, daß $x \leq y$ und $x \neq y$, so schreiben wir auch $x < y$.

Beispiel 8.4

Ist M eine Menge, so ist die Potenzmenge $\mathcal{P}(M)$ von M durch

$$A \leq B :\Leftrightarrow A \subseteq B, \text{ für } A, B \in \mathcal{P}(M),$$

partiell geordnet, aber im allgemeinen sind zwei Elemente von $\mathcal{P}(M)$ nicht unbedingt vergleichbar bezüglich dieser Ordnungsrelation. Z. B. sind im Fall $M = \mathbb{N}$ die Elemente $\{2\}$ und $\{3\}$ in $\mathcal{P}(\mathbb{N})$ nicht vergleichbar.

Allgemeiner gilt, ist N eine Menge, deren Elemente wieder Mengen sind, so wird N mit der analogen Definition von “ \leq ” eine partiell geordnete Menge.

Definition 8.5 (Total- und Wohlordnungen)

Es sei M ein Menge.

- Eine Ordnungsrelation “ \leq ” auf M heißt *Totalordnung* oder *lineare Ordnung*, falls je zwei Elemente aus M vergleichbar sind, d. h. für je zwei Elemente $x, y \in M$ gilt $x \leq y$ oder $y \leq x$.
- Ist “ \leq ” eine Ordnungsrelation auf M , $A \subseteq M$ und $x \in A$, so heißt x *minimal* (bzw. *maximal*) in A , falls für alle $y \in A$ mit $y \leq x$ (bzw. $x \leq y$) gilt $x = y$.
- Eine Totalordnung heißt *Wohlordnung*, falls jede nicht-leere Teilmenge von M ein minimales Element besitzt.

Bemerkung 8.6 (Minimum und Maximum)

Das Minimum bzw. Maximum einer Menge M bezüglich einer Totalordnung ist offenbar eindeutig bestimmt, sofern es existiert. Wir bezeichnen es mit $\min(M)$ bzw. mit $\max(M)$.

Beispiel 8.7

- Die reellen Zahlen (\mathbb{R}, \leq) mit der üblichen Kleiner-Gleich-Relation \leq sind total geordnet, aber nicht wohlgeordnet.
- Gleiches trifft auf (\mathbb{Z}, \leq) mit der üblichen Kleiner-Gleich-Relation

$$\dots - 2 < -1 < 0 < 1 < 2 < \dots$$

zu. Allerdings definiert die “unübliche” Anordnung

$$0 < -1 < 1 < -2 < 2 < -3 < 3 < \dots$$

in der Tat eine Wohlordnung auf \mathbb{Z} .

Bemerkung 8.8 (Archimedisches Prinzip)

Die natürlichen Zahlen sind bezüglich der üblichen Ordnungsrelation “ \leq ” wohlgeordnet, d.h.:

Jede nicht-leere Menge natürlicher Zahlen enthält eine kleinste Zahl.

Diese wohlbekannte Eigenschaft der natürlichen Zahlen nennen wir auch das *archimedische Prinzip*.

Definition 8.9 (Charakteristik eines Körpers)

Es sei K ein Körper. Gibt es eine positive ganze Zahl $n \in \mathbb{Z}_{>0}$ mit $n \cdot 1_K = 0_K$, so definieren wir

$$\text{char}(K) := \min\{m > 0 \mid m \cdot 1_K = 0_K\} \in \mathbb{N},$$

sonst setzen wir $\text{char}(K) := 0$. Die Zahl $\text{char}(K)$ heißt die *Charakteristik* von K .

Proposition 8.10 (Die Charakteristik eines Körpers ist eine Primzahl oder Null.)
Ist K ein Körper mit $\text{char}(K) \neq 0$, so ist $\text{char}(K)$ eine Primzahl.

Beweis: Angenommen, $n := \text{char}(K)$ sei keine Primzahl. Dann gibt es zwei Zahlen $1 < a, b < n$ mit $n = a \cdot b$. Setzen wir $x = a \cdot 1_K$ und $y = b \cdot 1_K$, so gilt

$$x \cdot y = (a \cdot 1_K) \cdot (b \cdot 1_K) = (a \cdot b) \cdot 1_K = n \cdot 1_K = 0.$$

Aus Lemma 7.8 folgt dann aber $a \cdot 1_K = x = 0$ oder $b \cdot 1_K = y = 0$, im Widerspruch zur Minimalität von $n = \text{char}(K)$. Also muß n eine Primzahl sein. \square

Beispiel 8.11

Ist p eine Primzahl, so hat der Körper F_p aus Beispiel 7.6 die Charakteristik $\text{char}(F_p) = p$. Die Körper \mathbb{Q} , \mathbb{R} und \mathbb{C} haben Charakteristik null.

Definition 8.12

 (Supremum und Infimum)

Es sei " \leq " eine Totalordnung auf einer Menge M und $\emptyset \neq A \subseteq M$ eine nicht-leere Teilmenge von M .

- Wir nennen $s \in M$ eine *obere Schranke* von A , falls $s \geq x$ für alle $x \in A$.
- Wir nennen A *nach oben beschränkt*, falls A eine obere Schranke besitzt.
- Wir nennen $s \in M$ das *Supremum* von A , falls s das Minimum der Menge der oberen Schranken von A ist. Dieses Minimum ist eindeutig bestimmt, wenn es existiert, und wir bezeichnen es dann mit $\sup(A)$.
- Wir nennen $s \in M$ eine *untere Schranke* von A , falls $s \leq x$ für alle $x \in A$.
- Wir nennen A *nach unten beschränkt*, falls A eine untere Schranke besitzt.
- Wir nennen $s \in M$ das *Infimum* von A , falls s das Maximum der Menge aller unteren Schranken von A ist. Dieses Maximum ist eindeutig bestimmt, wenn es existiert, und wir bezeichnen es dann mit $\inf(A)$.
- Wir nennen A *beschränkt*, wenn A nach oben und nach unten beschränkt ist.

Beispiel 8.13

- Besitzt eine Teilmenge A einer totalgeordneten Menge M ein Maximum, so ist dieses offenbar auch das Supremum von A . Analog ist das Minimum einer Menge A auch ihr Infimum.
- Betrachten wir die reellen Zahlen mit ihrer üblichen Ordnung und die Menge $A = \{x \in \mathbb{R} \mid 0 < x \leq 1\}$, so ist $1 = \sup(A) = \max(A)$ das Supremum von A , das zugleich ein Maximum ist, und $0 = \inf(A)$ ist ein Infimum von A , das kein Minimum ist.
- Betrachten wir die rationalen Zahlen mit ihrer üblichen Ordnungsrelation, so ist

$$\{x \in \mathbb{Q} \mid x > 0 \text{ und } x^2 \leq 2\}$$

nach oben beschränkt, besitzt aber kein Supremum in \mathbb{Q} (siehe Satz 9.12).

Bemerkung 8.14 (Supremumsaxiom)

Die reellen Zahlen sind bezüglich ihrer üblichen Ordnungsrelation nicht wohlgeordnet, d.h. nicht jede nicht-leere Teilmenge besitzt ein kleinstes Element. Selbst, wenn wir voraussetzen, daß die Teilmenge nach unten beschränkt ist, muß sie kein kleinstes Element besitzen, d.h. kein Minimum enthalten, wie wir in Beispiel 8.13 gesehen haben. Es gilt aber, daß zu jeder nicht-leeren, nach unten beschränkten Teilmenge von \mathbb{R} ein Infimum in \mathbb{R} existiert. Äquivalent dazu ist die duale Aussage für das Supremum:

Jede nicht-leere, nach oben beschränkte Teilmenge von \mathbb{R} besitzt ein Supremum in \mathbb{R} .

Diese Eigenschaft ist als *Supremumsaxiom* der reellen Zahlen bekannt. Auch wenn sich die Korrektheit der Aussage nicht unmittelbar aus unserer Alltagserfahrung mit den reellen Zahlen als Dezimalzahlen erschließt, wollen wir sie ohne weiteren Beweis als gegeben voraussetzen.

B) Angeordnete Körper**Definition 8.15** (Angeordnete Körper)

Es sei K ein Körper und " \leq " eine Totalordnung auf K . Wir nennen das Quadrupel $(K, +, \cdot, \leq)$ einen *angeordneten Körper*, wenn die Totalordnung mit der Addition und der Multiplikation verträglich ist, d.h. wenn für alle $x, y, z \in K$

$$x < y \implies x + z < y + z$$

und

$$x < y, 0 < z \implies x \cdot z < y \cdot z$$

gilt. Ist $x \in K$ und $x > 0$, so nennen wir x *positiv*, ist $x < 0$, so nennen wir x *negativ*.

Beispiel 8.16

- a. Die rationalen Zahlen \mathbb{Q} und die reellen Zahlen \mathbb{R} mit der üblichen Ordnungsrelation sind Beispiele für angeordnete Körper. \mathbb{Q} erfüllt das Supremumsaxiom nicht (siehe Beispiel 8.13), \mathbb{R} erfüllt es.
- b. Es gibt keine Totalordnung auf \mathbb{F}_2 , durch die \mathbb{F}_2 ein angeordneter Körper würde. Denn würde es eine solche Totalordnung " \leq " geben, so wäre entweder $0 < 1$, was zum Widerspruch $1 = 0 + 1 < 1 + 1 = 0$ führt, oder es wäre $1 < 0$, was zum Widerspruch $0 = 1 + 1 < 0 + 1 = 1$ führt.

Lemma 8.17 (Rechenregeln in angeordneten Körpern)

Es sei $(K, +, \cdot, \leq)$ ein angeordneter Körper und $x, y, u, v \in K$.

- a. $x > 0 \iff -x < 0$.
- b. Ist $x \neq 0$, so ist $x^2 > 0$.
- c. $1 > 0$.

- d. Ist $0 < x < y$, so ist $0 < \frac{1}{y} < \frac{1}{x}$.
- e. Ist $x < y$ und $u < v$, so ist $x + u < y + v$.
- f. Ist $0 < x$ und $n \in \mathbb{N}$, so ist $0 < x^n$.
- g. Ist $0 \leq x, y$ und $n \in \mathbb{N}$ mit $n \geq 1$, so gilt

$$x < y \iff x^n < y^n.$$

Beweis:

- a. Aus $0 < x$ folgt durch Addition von $-x$

$$-x = 0 + (-x) < x + (-x) = 0.$$

Umgekehrt folgt aus $-x < 0$ durch Addition von x

$$0 = -x + x < 0 + x = x.$$

- b. Ist $x > 0$, so folgt unmittelbar

$$0 = 0 \cdot x < x \cdot x = x^2.$$

Ist $x < 0$, so ist $0 < -x$ und es gilt

$$0 = 0 \cdot (-x) < (-x) \cdot (-x) = x \cdot x = x^2.$$

- c. $1 = 1^2 > 0$.

- d. Nach Voraussetzung ist $y > 0$. Nehmen wir an, $\frac{1}{y} < 0$, so folgt

$$1 = \frac{1}{y} \cdot y < 0 \cdot y = 0$$

im Widerspruch zu Teil c., also ist $0 < \frac{1}{y}$. Entsprechend gilt $0 < \frac{1}{x}$, so daß auch

$$0 = 0 \cdot \frac{1}{y} < \frac{1}{x} \cdot \frac{1}{y} = \frac{1}{xy}$$

und somit wegen $x < y$ auch

$$\frac{1}{y} = x \cdot \frac{1}{xy} < y \cdot \frac{1}{xy} = \frac{1}{x}.$$

- e. Wir wenden die Verträglichkeit der Totalordnung mit der Addition mehrfach an:

$$x + u < y + u < y + v.$$

- f./g. Den Beweis überlassen wir dem Leser als Übungsaufgabe.

□

Proposition 8.18 (Charakterisierung des Supremums und Infimums)

Ist $(K, +, \cdot, \leq)$ ein angeordneter Körper, $A \subseteq K$ und $s \in K$, dann gelten

$$s = \sup(A) \iff \begin{array}{l} 1) \quad \forall x \in A : x \leq s \text{ und} \\ 2) \quad \forall 0 < \varepsilon \in K : \exists x \in A : s - \varepsilon < x \end{array}$$

sowie

$$s = \inf(A) \iff \begin{array}{l} 1) \quad \forall x \in A : x \geq s \text{ und} \\ 2) \quad \forall 0 < \varepsilon \in K : \exists x \in A : s + \varepsilon > x. \end{array}$$

Beweis: Ist $s = \sup(A)$, so ist s eine obere Schranke von A und somit gilt Bedingung 1). Sei also $0 < \varepsilon \in K$, so ist $s - \varepsilon < s$ und mithin ist $s - \varepsilon$ keine obere Schranke von A . Also gibt es ein $x \in A$ mit $x > s - \varepsilon$ und Bedingung 2) ist erfüllt.

Nehmen wir nun umgekehrt an, daß die Bedingungen 1) und 2) gelten. Wegen 1) ist s dann eine obere Schranke von A , und wir müssen nur noch zeigen, daß es keine kleinere obere Schranke geben kann. Dazu betrachten wir eine beliebige kleinere Zahl $t \in K$ mit $t < s$. Für $\varepsilon := s - t \in K$ gilt $\varepsilon > 0$ und wegen 2) gibt es dann ein $x \in A$ mit $x > s - \varepsilon = t$. Also ist t keine obere Schranke von A .

Die Aussage für das Infimum zeigt man analog. □

Das folgende Lemma ist interessant bei der Definition des Riemann-Integrals einer Funktion (siehe Definition 19.7).

Lemma 8.19

Seien $A, B \subseteq \mathbb{R}$ zwei nicht-leere Teilmengen von \mathbb{R} mit $a \leq b$ für alle $a \in A, b \in B$. Dann gilt

$$\sup(A) \leq \inf(B).$$

Beweis: Aus der Voraussetzung folgt unmittelbar, daß A nach oben und B nach unten beschränkt ist, so daß $\sup(A) \in \mathbb{R}$ und $\inf(B) \in \mathbb{R}$ existieren.

Angenommen, $\sup(A) > \inf(B)$, so ist $\varepsilon := \frac{\sup(A) - \inf(B)}{2} > 0$. Somit ist $\sup(A) - \varepsilon$ keine obere Schranke von A und $\inf(B) + \varepsilon$ keine untere Schranke von B . Es gibt also ein $a \in A$ und ein $b \in B$ mit

$$a > \sup(A) - \varepsilon = \frac{\sup(A) + \inf(B)}{2} = \inf(B) + \varepsilon > b,$$

was im Widerspruch zur Voraussetzung steht. □

Aufgaben

Aufgabe 8.20

Ist M eine endliche Menge, so gilt

$$|M| = \min\{n \in \mathbb{N} \mid \exists f : M \longrightarrow \{1, \dots, n\} \text{ injektiv}\},$$

und jede injektive Abbildung $f : M \longrightarrow \{1, \dots, |M|\}$ ist bijektiv.

Aufgabe 8.21

Zeige, daß durch

$$R := \{(x, y) \in \mathbb{N} \times \mathbb{N} \mid x \text{ teilt } y\}$$

eine Ordnungsrelation auf \mathbb{N} definiert wird. Ist R eine Totalordnung?

Aufgabe 8.22

Definiere auf $M = \mathbb{N} \times \mathbb{N}$ eine Relation durch

$$\begin{aligned}
 (m, n) \leq (k, l) \iff & \quad 1. \max\{m, n\} < \max\{k, l\} \text{ oder} \\
 & \quad 2. (\max\{m, n\} = \max\{k, l\} \text{ und } m < k) \text{ oder} \\
 & \quad 3. (\max\{m, n\} = \max\{k, l\} \text{ und } m = k \text{ und } n > l) \text{ oder} \\
 & \quad 4. (m, n) = (k, l).
 \end{aligned}$$

Zeige, daß “ \leq ” eine Totalordnung auf M definiert. Stelle graphisch in der Zahlenebene \mathbb{R}^2 dar, wie die Elemente (m, n) in M mit $\max\{m, n\} \leq 4$ angeordnet sind.

Aufgabe 8.23

Sei K ein angeordneter Körper und $A, B \subseteq K$ Teilmengen, so daß $\sup(A)$ und $\sup(B)$ existieren. Wir setzen $A + B := \{a + b \mid a \in A, b \in B\}$. Beweise, daß auch $\sup(A + B)$ existiert und $\sup(A + B) = \sup(A) + \sup(B)$ gilt.

Aufgabe 8.24

Bestimme Supremum, Infimum, Maximum und Minimum (sofern sie existieren) der Mengen:

$$A = \left\{ \frac{m+n}{m \cdot n} \mid m, n \in \mathbb{N}_{>0} \right\} \subseteq \mathbb{R}$$

und

$$B = \left\{ n + \frac{(-1)^n}{n} \mid n \in \mathbb{N}_{>0} \right\} \subseteq \mathbb{R}.$$

§ 9 Eigenschaften der reellen Zahlen \mathbb{R}

Theorem 9.1 (Charakterisierung der reellen Zahlen)

Der Körper \mathbb{R} der reellen Zahlen mit der üblichen Ordnungsrelation ist der einzige angeordnete Körper, in dem jede nicht-leere, nach oben beschränkte Menge ein Supremum besitzt.

Bemerkung 9.2

Die Aussage in Theorem 9.1 besagt zweierlei. Zum einen wird festgestellt, daß \mathbb{R} ein angeordneter Körper ist und dem Supremumsaxiom genügt. Zum anderen wird festgestellt, daß dies für keinen *anderen* angeordneten Körper gilt. Das soll heißen, wenn es einen anderen angeordneten Körper $(K, +, \cdot, \leq)$ mit diesen Eigenschaften gibt, dann gibt es eine *bijektive* Abbildung

$$f : \mathbb{R} \longrightarrow K,$$

so daß $f(x + y) = f(x) + f(y)$, $f(x \cdot y) = f(x) \cdot f(y)$ und

$$x \leq y \iff f(x) \leq f(y)$$

für alle $x, y \in \mathbb{R}$ gilt. In dem Fall kann man die beiden Körper nicht mehr unterscheiden. Man sagt deshalb auch, daß die reellen Zahlen durch die Eigenschaften in Theorem 9.1 charakterisiert sind, und man könnte die reellen Zahlen axiomatisch durch Angabe der Eigenschaften einführen.

Wir wollen Theorem 9.1 in dieser Vorlesung *nicht* beweisen. Stattdessen werden wir von den reellen Zahlen von nun an nur noch die im Satz angegebenen Eigenschaften wirklich verwenden. Wenn wir uns also \mathbb{R} als einen beliebigen angeordneten Körper mit Supremumsaxiom denken, dann wird alles, was wir von nun an beweisen, dort genauso gelten. Wir müßten die reellen Zahlen also noch gar nicht kennen, um die weitere Theorie betreiben zu können. Die wenigen oben gegebenen Axiome reichen uns aus. Insofern befinden wir uns von jetzt an auf wesentlich sichererem Grund und müssen nicht mehr immer wieder Bezug auf unser Vorwissen zu den Zahlssystemen nehmen.

Satz 9.3 (\mathbb{R} ist archimedisch angeordnet.)

Für $x, y \in \mathbb{R}$ mit $0 < x < y$ gibt es eine natürliche Zahl $n \in \mathbb{N}$, so daß $y < n \cdot x$.

Beweis: Wir betrachten die nicht-leere Teilmenge

$$A := \{n \cdot x \mid n \in \mathbb{N}\} \subsetneq \mathbb{R}$$

der reellen Zahlen und müssen zeigen, daß y keine obere Schranke dieser Menge ist.

Nehmen wir an, dies wäre doch der Fall, dann ist A nach oben beschränkt und somit existiert das Supremum

$$s := \sup(A).$$

Da $x > 0$ ist, ist $s - x < s$ und somit ist $s - x$ keine obere Schranke von A , d.h. es gibt eine natürliche Zahl $n \in \mathbb{N}$ mit

$$s - x < n \cdot x.$$

Dann ist aber auch

$$s = (s - x) + x < n \cdot x + x = (n + 1) \cdot x,$$

im Widerspruch dazu, daß s eine obere Schranke von A ist.

Damit haben wir gezeigt, daß A keine obere Schranke besitzt und insbesondere, daß y keine solche ist, d.h. es gibt eine natürliche Zahl $n \in \mathbb{N}$ mit $y < n \cdot x$. \square

Korollar 9.4 (Konsequenzen der archimedischen Anordnung)

- Für alle $x \in \mathbb{R}$ gibt es eine ganze Zahl n , so daß $n \leq x < n + 1$.
- Für alle $\varepsilon \in \mathbb{R}$ mit $\varepsilon > 0$ gibt es eine natürliche Zahl n , so daß $0 < \frac{1}{n} < \varepsilon$.

Beweis:

- Ist $0 \leq x < 1$, so ist $n = 0$. Ist $1 \leq x$, so gibt es nach Satz 9.3 eine Zahl $m \in \mathbb{N}$ mit $x < m \cdot 1 = m$. Nach dem Archimedischen Prinzip 8.8 besitzt dann die nicht-leere Menge

$$M := \{k \in \mathbb{N} \mid x < k\}$$

ein Minimum $m_0 = \min(M)$, und für $n := m_0 - 1 < m_0$ gilt mithin

$$n \leq x < m_0 = n + 1.$$

Ist $x < 0$, so ist $-x > 0$ und wir haben schon gezeigt, daß es eine natürliche Zahl $m \in \mathbb{N}$ mit $m \leq -x < m + 1$ gibt. Dann ist aber

$$-m - 1 < x \leq -m.$$

Falls $x = -m$, so setzen wir $n := -m$, und sonst setzen wir $n := -m - 1$.

- Wegen $\varepsilon > 0$ ist nach Lemma 8.17 auch $\frac{1}{\varepsilon} > 0$, und nach a. gibt es dann eine natürliche Zahl $n \in \mathbb{N}$ so, daß

$$0 < \frac{1}{\varepsilon} < n.$$

Mit Lemma 8.17 folgt dann

$$0 < \frac{1}{n} < \frac{1}{\frac{1}{\varepsilon}} = \varepsilon.$$

\square

Definition 9.5 (Intervalle)

Es seien $a, b \in \mathbb{R}$. Wir nennen eine Menge der Form

$$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$$

ein *abgeschlossenes Intervall*, eine Menge der Form

$$(a, b) := \{x \in \mathbb{R} \mid a < x < b\}$$

ein *offenes Intervall* und Mengen der Form

$$[a, b) := \{x \in \mathbb{R} \mid a \leq x < b\}$$

bzw.

$$(a, b] := \{x \in \mathbb{R} \mid a < x \leq b\}$$

halboffene Intervalle. Mengen der Form

$$[a, \infty) := \{x \in \mathbb{R} \mid a \leq x\},$$

$$(a, \infty) := \{x \in \mathbb{R} \mid a < x\},$$

$$(-\infty, a] := \{x \in \mathbb{R} \mid x \leq a\},$$

$$(-\infty, a) := \{x \in \mathbb{R} \mid x < a\}$$

$$(-\infty, \infty) := \mathbb{R}$$

heißen *uneigentliche Intervalle*.

Satz 9.6 (\mathbb{Q} liegt dicht in \mathbb{R} .)

Sind $a, b \in \mathbb{R}$ mit $a < b$, so gibt es eine rationale Zahl im Intervall (a, b) .

Beweis: Wegen $b - a > 0$ gibt es nach Korollar 9.4 eine natürliche Zahl $n \in \mathbb{N}$ mit

$$0 < \frac{1}{n} < b - a. \quad (7)$$

Zudem gibt es nach Korollar 9.4 eine ganze Zahl $m \in \mathbb{Z}$ mit

$$m \leq n \cdot a < m + 1. \quad (8)$$

Damit gilt dann

$$a < \stackrel{(8)}{\frac{m+1}{n}} = \frac{m}{n} + \frac{1}{n} \stackrel{(8)}{\leq} a + \frac{1}{n} \stackrel{(7)}{<} b$$

und $\frac{m+1}{n} \in \mathbb{Q}$ ist eine rationale Zahl. \square

Definition 9.7 (Absolutbetrag)

Für eine reelle Zahl $x \in \mathbb{R}$ definieren wir den *Betrag* oder *Absolutbetrag* von x durch

$$|x| := \begin{cases} x, & \text{falls } x \geq 0 \\ -x, & \text{falls } x < 0. \end{cases}$$

Damit erhalten wir dann die Betragsfunktion $|\cdot| : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto |x|$.

Lemma 9.8 (Eigenschaften des Betrages)

Es seien $x, y, z \in \mathbb{R}$ und $z \neq 0$.

a. Der Betrag ist multiplikativ, d.h.

$$|x \cdot y| = |x| \cdot |y|.$$

b. *Der Betrag genügt der Dreiecksungleichung, d.h.*

$$|\mathbf{x} + \mathbf{y}| \leq |\mathbf{x}| + |\mathbf{y}|.$$

c. *Außerdem gelten offenbar:*

- $|\mathbf{x}| = |-\mathbf{x}|.$
- $|\mathbf{x}| \geq 0.$
- $|\mathbf{x}| = 0 \iff \mathbf{x} = 0.$

d. $||\mathbf{x}| - |\mathbf{y}|| \leq |\mathbf{x} - \mathbf{y}|.$

e. $\left|\frac{\mathbf{x}}{\mathbf{z}}\right| = \frac{|\mathbf{x}|}{|\mathbf{z}|}.$

Beweis: Wir stellen zunächst fest, daß die Aussagen in c. unmittelbar aus der Definition des Betrages folgen.

a. Wir unterscheiden die Fälle, ob \mathbf{x} bzw. \mathbf{y} negativ oder nicht negativ ist und fassen unsere Überlegungen in einer Tabelle zusammen:

\mathbf{x}	\mathbf{y}	\mathbf{xy}	$ \mathbf{x} $	$ \mathbf{y} $	$ \mathbf{x} \cdot \mathbf{y} $	$ \mathbf{xy} $
≥ 0	≥ 0	≥ 0	\mathbf{x}	\mathbf{y}	\mathbf{xy}	\mathbf{xy}
≥ 0	< 0	≤ 0	\mathbf{x}	$-\mathbf{y}$	$\mathbf{x} \cdot (-\mathbf{y}) = -\mathbf{xy}$	$-\mathbf{xy}$
< 0	≥ 0	≤ 0	$-\mathbf{x}$	\mathbf{y}	$(-\mathbf{x}) \cdot \mathbf{y} = -\mathbf{xy}$	$-\mathbf{xy}$
< 0	< 0	> 0	$-\mathbf{x}$	$-\mathbf{y}$	$(-\mathbf{x}) \cdot (-\mathbf{y}) = \mathbf{xy}$	\mathbf{xy}

b. Aus der Definition des Betrages folgt unmittelbar, daß $\mathbf{x} \leq |\mathbf{x}|$ und $\mathbf{y} \leq |\mathbf{y}|$, und damit gilt dann auch

$$\mathbf{x} + \mathbf{y} \stackrel{8.17}{\leq} |\mathbf{x}| + |\mathbf{y}|.$$

Analog gilt $-\mathbf{x} \leq |-\mathbf{x}| = |\mathbf{x}|$ und $-\mathbf{y} \leq |-\mathbf{y}| = |\mathbf{y}|$ und damit auch

$$-(\mathbf{x} + \mathbf{y}) = (-\mathbf{x}) + (-\mathbf{y}) \stackrel{8.17}{\leq} |\mathbf{x}| + |\mathbf{y}|.$$

Wegen

$$|\mathbf{x} + \mathbf{y}| = \begin{cases} \mathbf{x} + \mathbf{y}, & \text{falls } \mathbf{x} + \mathbf{y} \geq 0, \\ -(\mathbf{x} + \mathbf{y}) & \text{falls } \mathbf{x} + \mathbf{y} < 0, \end{cases}$$

haben wir die Dreiecksungleichung gezeigt.

d. Aus der Dreiecksungleichung erhalten wir

$$|\mathbf{x}| = |(\mathbf{x} - \mathbf{y}) + \mathbf{y}| \leq |\mathbf{x} - \mathbf{y}| + |\mathbf{y}|,$$

und somit

$$|\mathbf{x}| - |\mathbf{y}| \leq |\mathbf{x} - \mathbf{y}|.$$

Analog folgt

$$-(|\mathbf{x}| - |\mathbf{y}|) = |\mathbf{y}| - |\mathbf{x}| \leq |\mathbf{y} - \mathbf{x}| = |-(\mathbf{y} - \mathbf{x})| = |\mathbf{x} - \mathbf{y}|.$$

Wegen

$$||x| - |y|| = \begin{cases} |x| - |y|, & \text{falls } |x| - |y| \geq 0, \\ -(|x| - |y|) & \text{falls } |x| - |y| < 0, \end{cases}$$

folgt dann $||x| - |y|| \leq |x - y|$.

e. Aus der Multiplikativität des Betrages folgt

$$|x| = \left| \frac{x}{z} \cdot z \right| = \left| \frac{x}{z} \right| \cdot |z|,$$

und da $|z| \neq 0$ wegen $z \neq 0$ folgt mithin auch

$$\frac{|x|}{|z|} = \left| \frac{x}{z} \right|.$$

□

Satz 9.9 (Bernoullische Ungleichung)

Es sei $x \in \mathbb{R}$ mit $x \geq -1$ und $n \in \mathbb{N}$, dann gilt

$$(1 + x)^n \geq 1 + n \cdot x.$$

Beweis: Wir führen den Beweis durch Induktion nach n .

Induktionsanfang: $n = 0$: $(1 + x)^0 = 1 = 1 + 0 \cdot x$.

Induktionsschluß: $n \mapsto n + 1$: Nach Lemma 8.17 b. ist $x^2 \geq 0$ und nach Voraussetzung gilt zudem $1 + x \geq 0$. Damit erhalten wir dann:

$$\begin{aligned} (1 + x)^{n+1} &= (1 + x)^n \cdot (1 + x) \stackrel{\text{Ind}}{\geq} \\ &(1 + n \cdot x) \cdot (1 + x) = 1 + (n + 1) \cdot x + n \cdot x^2 \stackrel{8.17b.}{\geq} 1 + (n + 1) \cdot x. \end{aligned}$$

Die Aussage ist damit also mittels Induktion gezeigt. □

Satz 9.10 (Existenz von n -ten Wurzeln in \mathbb{R})

Zu jeder reellen Zahl $x \in \mathbb{R}$ mit $x \geq 0$ und jeder natürlichen Zahl $n \in \mathbb{N}$ mit $n \geq 2$ gibt es genau eine reelle Zahl $a \in \mathbb{R}$ mit $a \geq 0$ und $a^n = x$.

Wir nennen diese Zahl die n -te Wurzel aus x und bezeichnen sie mit $\sqrt[n]{x}$ oder $x^{\frac{1}{n}}$.

Beweis: Wir wollen uns zunächst der Eindeutigkeit der Lösung zuwenden, sofern sie existiert. Nehmen wir also an, es würde zwei verschiedene nicht-negative reelle Zahlen $a, b \in \mathbb{R}$ mit $a^n = b^n = x$ geben. Dann ist eine der beiden echt kleiner als die andere und wir können ohne Einschränkung annehmen, daß dies a ist, d.h. $0 \leq a < b$. Aus Lemma 8.17 g. folgt dann $x = a^n < b^n = x$, was ein offensichtlicher Widerspruch ist. Mithin haben wir gezeigt, daß es höchstens eine nicht-negative Zahl $a \in \mathbb{R}$ mit $a^n = x$ geben kann.

Es bleibt noch zu zeigen, daß es auch wirklich eine solche nicht-negative Zahl a gibt. Ist $x = 0$, so ist $a = 0$ eine Lösung für $a^n = 0$. Wir können im weiteren Verlauf des Beweises also voraussetzen, daß $x > 0$.

Wir betrachten dann die Teilmenge

$$A := \{y \in \mathbb{R} \mid y \geq 0, y^n \leq x\}$$

der reellen Zahlen, und wir behaupten, daß $1+x$ eine obere Schranke für A ist. Dazu betrachten wir eine reelle Zahl $y \in \mathbb{R}$ mit $y \geq 1+x > 0$. Aus der Bernoullischen Ungleichung folgt dann

$$y^n \stackrel{8.17g.}{\geq} (1+x)^n \stackrel{9.9}{\geq} 1+n \cdot x > x,$$

und somit ist $y \notin A$. Also ist A nach oben beschränkt durch $x+1$. Wegen $0 \in A$ ist A zudem nicht-leer und deshalb existiert das Supremum

$$a := \sup(A) \geq 0.$$

Wir wollen nun zeigen, daß $a^n = x$ gilt.

Zeige: $a^n \geq x$: Nehmen wir an, es gelte $a^n < x$.

Idee: Finde eine reelle Zahl $\varepsilon > 0$, so daß $a + \varepsilon \in A$. – ζ

Wegen $a \geq 0$ ist

$$c := \sum_{k=1}^n \binom{n}{k} \cdot a^{n-k} \geq \binom{n}{n} = 1 > 0$$

und somit auch $\frac{1}{c} > 0$ nach Lemma 8.17. Aus unserer Annahme folgt dann

$$\frac{x - a^n}{c} > 0.$$

Somit ist auch

$$\varepsilon := \min \left\{ \frac{x - a^n}{c}, 1 \right\} > 0$$

und es folgt

$$a^n + c \cdot \varepsilon \leq x. \tag{9}$$

Wegen $0 < \varepsilon \leq 1$ ist $\varepsilon^k \leq \varepsilon$ für alle $k \geq 1$, und aus dem Binomischen Lehrsatz 7.15 folgt dann

$$\begin{aligned} (a + \varepsilon)^n &= a^n + \sum_{k=1}^n \binom{n}{k} \cdot a^{n-k} \cdot \varepsilon^k \\ &\leq a^n + \sum_{k=1}^n \binom{n}{k} \cdot a^{n-k} \cdot \varepsilon = a^n + c \cdot \varepsilon \stackrel{(9)}{\leq} x. \end{aligned}$$

Somit ist $a + \varepsilon \in A$ und $a + \varepsilon > a$ im Widerspruch dazu, daß a das Supremum von A ist. Mithin muß $a^n \geq x$ sein.

Zeige: $a^n \leq x$: Nehmen wir an, es gelte $a^n > x$.

Idee: Finde ein $\varepsilon > 0$ und ein $y \in A$, so daß $y^n > (a - \varepsilon)^n \geq x$. – ζ

Wegen $a^n > x$ ist $a > 0$ und dann ist auch die Zahl

$$\frac{a \cdot (a^n - x)}{n \cdot a^n} > 0$$

positiv. Wir setzen nun

$$\varepsilon := \min \left\{ \frac{\mathbf{a} \cdot (\mathbf{a}^n - \mathbf{x})}{\mathbf{n} \cdot \mathbf{a}^n}, \mathbf{a} \right\} > 0.$$

Aus der Definition von ε folgt zum einen

$$-\frac{\varepsilon}{\mathbf{a}} \geq -1 \tag{10}$$

und zum anderen unter Anwendung der Bernoullischen Ungleichung

$$\mathbf{x} \leq \mathbf{a}^n \cdot \left(1 + \mathbf{n} \cdot \frac{-\varepsilon}{\mathbf{a}} \right) \stackrel{9.9}{\leq} \mathbf{a}^n \cdot \left(1 - \frac{\varepsilon}{\mathbf{a}} \right)^n = (\mathbf{a} - \varepsilon)^n; \tag{11}$$

dabei beachten wir die Bernoullische Ungleichung wegen (10) anwenden können.

Da \mathbf{a} das Supremum von A ist und $\mathbf{a} - \varepsilon < \mathbf{a}$ ist, muß es eine Zahl $\mathbf{y} \in A$ geben mit

$$\mathbf{y} > \mathbf{a} - \varepsilon > 0.$$

Dann gilt nach Lemma 8.17 auch

$$\mathbf{y}^n > (\mathbf{a} - \varepsilon)^n \stackrel{(11)}{\geq} \mathbf{x},$$

im Widerspruch dazu, daß $\mathbf{y} \in A$. Also muß auch $\mathbf{a}^n \leq \mathbf{x}$ gelten.

Da sowohl $\mathbf{a}^n \geq \mathbf{x}$, als auch $\mathbf{a}^n \leq \mathbf{x}$ gilt, folgt aus der Antisymmetrie der Ordnungsrelation, daß $\mathbf{a}^n = \mathbf{x}$, und wir haben die \mathbf{n} -te Wurzel von \mathbf{x} gefunden. \square

Bemerkung 9.11

In \mathbb{R} besitzt also insbesondere jede nicht-negative Zahl eine Quadratwurzel. Dies gilt in den rationalen Zahlen nicht (siehe Satz 9.12), und man kann die reellen Zahlen als eine Erweiterung des Zahlbereichs der rationalen Zahlen ansehen, die unter anderem deshalb notwendig war. Negative Zahlen besitzen aber auch in \mathbb{R} noch keine Quadratwurzeln, und wir werden im folgenden Kapitel deshalb unseren Zahlbereich noch einmal erweitern zu den sogenannten komplexen Zahlen, die dieses Manko dann beheben.

Satz 9.12 ($\sqrt{2}$ ist irrational.)

Es gibt keine rationale Zahl $\mathbf{a} \in \mathbb{Q}$ mit $\mathbf{a}^2 = 2$.

Beweis: Nehmen wir an, es wäre $\mathbf{a} = \frac{p}{q} \in \mathbb{Q}$ eine solche Zahl. Wir können ohne weiteres annehmen, daß der Bruch in gekürzter Form vorliegt. Aus

$$\frac{p^2}{q^2} = \mathbf{a}^2 = 2$$

folgt dann

$$p^2 = q^2 \cdot 2.$$

Also ist p^2 eine gerade Zahl, und dann muß notwendigerweise auch p eine gerade Zahl sein. D.h. es gibt ein $\mathbf{b} \in \mathbb{Z}$ mit $p = 2 \cdot \mathbf{b}$. Also ist

$$4 \cdot \mathbf{b}^2 = p^2 = 2 \cdot q^2,$$

und somit

$$2 \cdot b^2 = q^2.$$

Mit dem gleichen Argument sind dann auch q^2 und q gerade Zahlen, und somit ist q von der Form $q = 2 \cdot c$. Aber das widerspricht der Voraussetzung, daß der Bruch $\frac{p}{q}$ in gekürzter Form vorgelegen hat. \square

Aufgaben

Aufgabe 9.13

Zeige durch vollständige Induktion, daß

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \dots + \frac{1}{\sqrt{n-1}} + \frac{1}{\sqrt{n}} > \sqrt{n}$$

für alle $n \geq 2$ gilt.

§ 10 Der Körper der komplexen Zahlen

Wir kommen jetzt zum Körper \mathbb{C} der komplexen Zahlen, dem neben \mathbb{R} wichtigsten Körper. Warum reichen eigentlich die reellen Zahlen nicht aus, wozu braucht man die komplexen Zahlen? Ja, man kann sogar fragen, warum wir überhaupt die reellen Zahlen benötigen, wenn wir doch ohnehin nur mit endlichen Dezimalbrüchen, also rationalen Zahlen, rechnen können? Die Antwort auf die zweite Frage ist schnell gegeben. Wir wissen alle, daß etwa ganz natürlich auftretende Größen wie die Länge der Diagonalen eines Quadrates mit Seitenlänge eins, sprich die Zahl $\sqrt{2}$, oder das Verhältnis von Umfang zum Durchmesser eines Kreises, sprich die Kreiszahl π , keine rationalen Zahlen sind. Sie sind aber reelle Zahlen und die reellen Zahlen sind in gewissem Sinne, eine Vervollständigung der rationalen Zahlen. Wir brauchen also die reellen Zahlen, da die rationalen Zahlen Lücken aufweisen. Die komplexen Zahlen werden nun deshalb eingeführt, um einen Mangel, den die reellen Zahlen immer noch haben, zu beheben. Hierbei geht es um das Lösen von Gleichungen, aber nicht mehr linearen, sondern quadratischen. Es ist bekannt, daß das Quadrat einer reellen Zahl stets nicht-negativ ist. Also kann es keine reelle Zahl x geben, die die Gleichung $x^2 = -1$ löst.

Als Lösung genau dieser Gleichung wird nun eine neue Größe eingeführt, die *imaginäre Einheit* i . Definitionsgemäß ist sie diejenige Zahl, für die $i^2 = -1$ gilt. Wenn man nun eine solche Größe i einführt, dann ist damit alleine gar nichts gewonnen. Man will ja mit i auch rechnen können, und zwar will man möglichst alle Rechenregeln von \mathbb{R} übertragen. Man will nicht nur $i^2 = i \cdot i$, sondern auch $i+i$ oder Ausdrücke wie $37+42i$ bilden können. Dabei sollen die so zu konstruierenden *komplexen Zahlen* die reellen Zahlen als Teilmenge enthalten.

Daß es wirklich ein solches Zahlssystem komplexer Zahlen, in unserer Sprache den Körper der komplexen Zahlen, gibt, ist überhaupt nicht klar und wurde historisch erst spät realisiert und auch akzeptiert.² Gauß hat die Zahlen geometrisch, als Punkte in der Ebene, eingeführt, weshalb die komplexen Zahlen heute noch *Gaußsche Zahlenebene* heißen. Wir führen die komplexen Zahlen ebenfalls als reelle Zahlenpaare ein, definieren die Addition und die Multiplikation aber algebraisch und werden die Definitionen erst im Anschluß daran geometrisch interpretieren.

Satz 10.1 (Der Körper der komplexen Zahlen)

Die Menge $\mathbb{C} := \{(x, y) \mid x, y \in \mathbb{R}\}$ zusammen mit der durch

$$(x, y) + (u, v) := (x + u, y + v), \quad \text{für } (x, y), (u, v) \in \mathbb{C},$$

und

$$(x, y) \cdot (u, v) := (xu - yv, xv + yu), \quad \text{für } (x, y), (u, v) \in \mathbb{C},$$

²Erstmals tauchte $\sqrt{-1}$ wohl um 1540 bei Cardano auf. Wirklich als Zahlssystem wurden die komplexen Zahlen aber erst durch Gauß, 1777-1855, etabliert. Hierzu und zu vielen weiteren interessanten Tatsachen um die komplexen Zahlen vgl. [Ebb92] § 3.

definierten Addition und Multiplikation ist ein Körper, der Körper der komplexen Zahlen.

Beweis: Dies folgt aus Aufgabe 7.18. \square

Bemerkung 10.2 (\mathbb{R} als Teilkörper von \mathbb{C})

Die Teilmenge $\mathbb{R} \times \{0\}$ von \mathbb{C} ist ein Teilkörper von \mathbb{C} , wie man leicht sieht, und wir können den Körper \mathbb{R} auf offensichtliche Weise mit diesem Teilkörper identifizieren, indem wir $x \in \mathbb{R}$ mit $(x, 0) \in \mathbb{C}$ gleichsetzen.

Notation 10.3 (Die Schreibweise $x + iy$)

Praktischer als das Rechnen mit Paaren von Zahlen ist die folgende Notation für komplexe Zahlen. Wir setzen $x := (x, 0)$ für $x \in \mathbb{R}$ und $i := (0, 1)$. Dann gilt für $z = (x, y) \in \mathbb{C}$

$$z = (x, y) = (x, 0) + (0, y) = (x, 0) + (0, 1) \cdot (y, 0) = x + iy.$$

Bemerkung 10.4 (Die imaginäre Einheit i)

Mit dieser Schreibweise gilt zunächst:

$$i^2 = (0, 1) \cdot (0, 1) = -1.$$

Ferner ergibt sich die etwas willkürlich anmutende Definition der Multiplikation ganz "natürlich" aus

$$(x + iy)(u + iv) = (xu + i^2yv) + i(xv + yu) = (xu - yv) + i(xv + yu).$$

Lemma 10.5 (\mathbb{C} ist nicht angeordnet.)

Es gibt keine Totalordnung " \leq " auf \mathbb{C} , die \mathbb{C} zu einem angeordneten Körper macht.

Beweis: Angenommen, es gäbe eine Totalordnung " \leq ", die \mathbb{C} zu einem angeordneten Körper macht. Dann muß entweder $0 < i$ oder $0 < -i$ gelten, und somit $0 < i^2 = -1$ oder $0 < (-i)^2 = -1$, was im Widerspruch zu $0 < 1$ steht. \square

Definition 10.6 (Die komplexe Konjugation)

a. Wir definieren die *Betragsfunktion* auf \mathbb{C} durch

$$|\cdot| : \mathbb{C} \longrightarrow \mathbb{R}_{\geq 0} : x + iy \mapsto \sqrt{x^2 + y^2}.$$

b. Wir definieren die *komplexe Konjugation* als

$$\bar{\cdot} : \mathbb{C} \longrightarrow \mathbb{C} : z = x + iy \mapsto \bar{z} := x - iy.$$

Für $z \in \mathbb{C}$ heißt \bar{z} die zu z *konjugiert komplexe Zahl*.

c. Für $z = x + iy \in \mathbb{C}$ heißt $\operatorname{Re}(z) := x$ der Realteil von z und $\operatorname{Im}(z) := y$ der Imaginärteil von z .

Bemerkung 10.7

Man beachte, daß die Betragsfunktion auf \mathbb{C} die Betragsfunktion auf \mathbb{R} fortsetzt, da für eine reelle Zahl x stets $\sqrt{x^2} = |x|$ gilt.

Lemma 10.8 (Einfache Rechenregeln in \mathbb{C})

Es seien $z, w \in \mathbb{C}$.

a. Der Betrag ist multiplikativ, d.h.

$$|z| \cdot |w| = |zw|.$$

b. Der Betrag erfüllt die Dreiecksungleichung, d.h.

$$|z + w| \leq |z| + |w|.$$

c. $z = 0 \iff |z| = 0$.

d. $z \cdot \bar{z} = |z|^2$.

e. Wenn $z \neq 0$, dann ist $z^{-1} = \frac{1}{z} = \frac{\bar{z}}{|z|^2}$.

f. $\bar{z} + \bar{w} = \overline{z + w}$.

g. $\bar{z} \cdot \bar{w} = \overline{z \cdot w}$.

h. $\overline{\bar{z}} = z$.

i. $\bar{z} = z \iff z \in \mathbb{R}$.

j. $\operatorname{Re}(z) = \frac{z + \bar{z}}{2} \leq |z|$.

k. $\operatorname{Im}(z) = \frac{z - \bar{z}}{2i} \leq |z|$.

l. $|z| = |\bar{z}|$.

Beweis: Wir wollen die Dreiecksungleichung unter Verwendung der übrigen Aussagen zeigen und überlassen den Beweis dieser dem Leser als eine einfache Übungsaufgabe. Seien also $z, w \in \mathbb{C}$ gegeben. Dann gilt:

$$\begin{aligned} |z + w|^2 &\stackrel{\text{d.}}{=} (z + w) \cdot \overline{(z + w)} \\ &\stackrel{\text{f.}}{=} z \cdot \bar{z} + (z \cdot \bar{w} + \bar{z} \cdot w) + w \cdot \bar{w} \\ &\stackrel{\text{d.,h.}}{=} |z|^2 + (z \cdot \bar{w} + \overline{z \cdot \bar{w}}) + |w|^2 \\ &\stackrel{\text{j.}}{=} |z|^2 + 2 \cdot \operatorname{Re}(z \cdot \bar{w}) + |w|^2 \\ &\stackrel{\text{j.}}{\leq} |z|^2 + 2 \cdot |z \cdot \bar{w}| + |w|^2 \\ &\stackrel{\text{a.}}{=} |z|^2 + 2 \cdot |z| \cdot |\bar{w}| + |w|^2 \\ &\stackrel{\text{l.}}{=} |z|^2 + 2 \cdot |z| \cdot |w| + |w|^2 \\ &= (|z| + |w|)^2. \end{aligned}$$

Da dies eine Ungleichung von nicht-negativen Zahlen in dem angeordneten Körper \mathbb{R} ist, folgt aus Lemma 8.17, daß

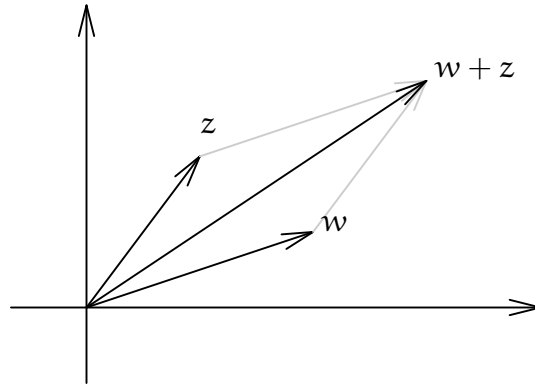
$$|z + w| \leq |z| + |w|.$$

□

Bemerkung 10.9 (Geometrische Deutung der komplexen Zahlen)

Wir betrachten $z = (x, y)$ als Richtungsvektor in der Zahlenebene \mathbb{R}^2 .

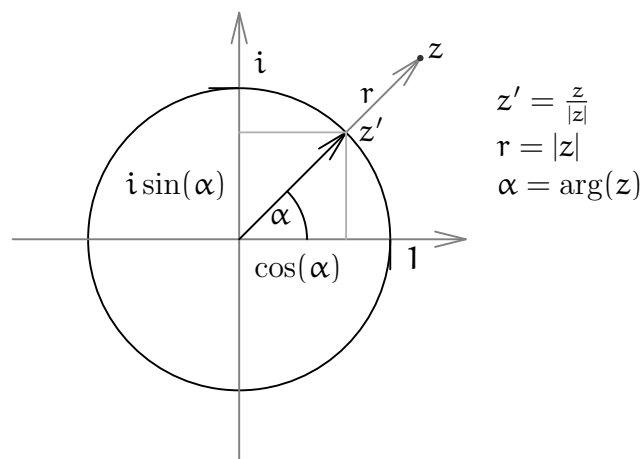
Die Addition ist einfach die komponentenweise Addition, also die Addition der Vektoren, und die Dreiecksungleichung besagt deshalb im wesentlichen, daß in einem Dreieck die Summe der Seitenlängen von zwei Seiten stets eine obere Schranke für die Seitenlänge der dritten Seite ist.



Zur geometrischen Interpretation der Multiplikation brauchen wir die Betragsfunktion. Der Betrag $r := |z|$ einer komplexen Zahl z ist die Länge des Vektors z (Pythagoras). Für $z \neq 0$ hat $z' := \frac{z}{|z|}$ die Länge eins, und es gilt

$$z = |z| \cdot z' = r \cdot z'.$$

D.h. z ist das Produkt eines Vektors von Länge eins mit einer nicht-negativen reellen Zahl. Dabei ist z' vollständig durch den Winkel α bestimmt, den z' mit der x -Achse einschließt, nämlich $z' = (\cos(\alpha), \sin(\alpha))$. Also ist jede komplexe Zahl $z \neq 0$ eindeutig durch ihren Betrag und den Winkel $\alpha =: \arg(z)$, das *Argument* von z , bestimmt. Das Paar $(r, \alpha) = (|z|, \arg(z))$ nennt man die *Polarkoordinaten* von z .



Wir wollen nun den folgenden Kapiteln etwas vorgreifen und die trigonometrischen Funktionen Sinus und Cosinus sowie die Exponentialfunktion verwenden, um die Multiplikation zweier komplexer Zahlen zu interpretieren.

Die komplexen Zahlen vom Betrag eins sind genau die Punkte auf dem Einheitskreis. Für $z' \in \mathbb{C}$ mit $|z'| = 1$ gibt es also genau ein $0 \leq \alpha < 2\pi$ mit

$$z' = \cos(\alpha) + i \sin(\alpha).$$

Wir werden später sehen, daß zwischen Sinus, Cosinus und der Exponentialfunktion der folgende Zusammenhang besteht,

$$\cos(\alpha) + i \sin(\alpha) = e^{i\alpha},$$

und daß die Exponentialfunktion den Potenzgesetzen genügt.

Damit gilt für ein beliebiges $c \in \mathbb{C}$

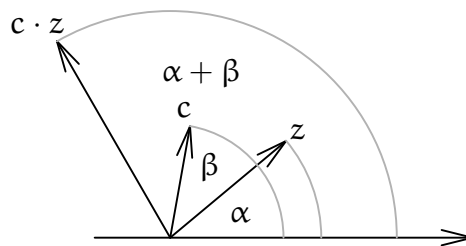
$$c = |c| \cdot (\cos(\beta) + i \sin(\beta)) = |c| \cdot e^{i\beta},$$

für $\beta = \arg(c)$.

Daraus ergibt sich für die Multiplikationsabbildung mit einer festen komplexen Zahl $z = |z| \cdot e^{i\alpha}$,

$$m_z : \mathbb{C} \longrightarrow \mathbb{C} : c \mapsto |z| \cdot |c| \cdot e^{i(\alpha+\beta)}.$$

Die Multiplikation mit z ist also eine *Drehstreckung*, daß heißt der Vektor c wird um den Winkel $\alpha = \arg(z)$ gedreht und um den Faktor $|z|$ gestreckt.



Bemerkung 10.10 (n-te Wurzeln in \mathbb{C})

Ist $z = r \cdot e^{i\alpha} \in \mathbb{C}$ und $n \geq 2$, so genügt $a = \sqrt[n]{r} \cdot e^{i\frac{\alpha}{n}} \in \mathbb{C}$ der Gleichung $a^n = z$. Insbesondere besitzt jede komplexe Zahl eine n-te Wurzel.

Aufgaben

Aufgabe 10.11

Bestimme für die folgenden komplexen Zahlen $\operatorname{Re} z$, $\operatorname{Im} z$, $\arg z$, $|z|$, \bar{z} und z^{-1} :

- (a) $z = i - 1$.
- (b) $z = \frac{4i}{1+i}$.
- (c) $z = \frac{(2+2i)^7}{(1-i)^3}$.

KAPITEL II

Eindimensionale Analysis

Im folgenden wollen wir die eindimensionale Analysis entwickeln, teilweise nur über den reellen Zahlen, teilweise parallel über den reellen und den komplexen Zahlen. Deshalb führen wir folgende Notation ein.

Im folgenden sei stets $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ einer der beiden Körper \mathbb{R} oder \mathbb{C} .

§ 11 Folgen und ihre Grenzwerte

A) Konvergente Folgen

Definition 11.1 (Folgen)

Eine *Folge* in \mathbb{K} ist eine Abbildung

$$\alpha : \mathbb{N} \longrightarrow \mathbb{K}$$

von den natürlichen Zahlen \mathbb{N} nach \mathbb{K} .

Notation 11.2 (Familienschreibweise für Folgen)

Eine Folge $\alpha : \mathbb{N} \longrightarrow \mathbb{K}$ ist eindeutig festgelegt durch ihre Funktionswerte $\mathbf{a}_n := \alpha(n)$ mit $n \in \mathbb{N}$. Wir schreiben deshalb statt $\alpha : \mathbb{N} \longrightarrow \mathbb{K}$ gemeinhin nur $(\mathbf{a}_n)_{n \in \mathbb{N}}$ oder $(\mathbf{a}_0, \mathbf{a}_1, \mathbf{a}_2, \dots)$.

Manchmal ist es angenehmer, eine Folge nicht bei 0 starten zu lassen, sondern bei einer anderen natürlichen Zahl k . Dann schreiben wir für die Folge schlicht $(\mathbf{a}_n)_{n \geq k}$. Formal würde dem dann die Abbildung

$$\mathbb{N} \longrightarrow \mathbb{K} : n \mapsto \mathbf{a}_{n+k}$$

entsprechen.

Beispiel 11.3

- a. Ist $c \in \mathbb{K}$, so heißt $\alpha : \mathbb{N} \longrightarrow \mathbb{K} : n \mapsto c$ eine *konstante Folge*. Es gilt $\mathbf{a}_n = c$ für $n \in \mathbb{N}$, und mithin $(\mathbf{a}_n)_{n \in \mathbb{N}} = (c)_{n \in \mathbb{N}}$.
- b. Für $q \in \mathbb{K}$ ist auch $\alpha : \mathbb{N} \longrightarrow \mathbb{K} : n \mapsto q^n$ eine Folge mit $\mathbf{a}_n = q^n$, also $(\mathbf{a}_n)_{n \in \mathbb{N}} = (q^n)_{n \in \mathbb{N}}$.
- c. $(\frac{1}{n})_{n \geq 1}$ ist ein Beispiel für eine Folge in \mathbb{K} , bei der der Folgenindex nicht bei 0 startet.

Definition 11.4 (Konvergenz und Grenzwert)

Es sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{K} und $\mathbf{a} \in \mathbb{K}$.

- a. Wir nennen \mathbf{a} genau dann einen *Grenzwert* von $(\mathbf{a}_n)_{n \in \mathbb{N}}$, wenn

$$\forall 0 < \varepsilon \in \mathbb{R} \quad \exists n_\varepsilon \in \mathbb{N} : \forall n \geq n_\varepsilon : |\mathbf{a}_n - \mathbf{a}| < \varepsilon.$$

In diesem Fall sagen wir auch, daß $(\mathbf{a}_n)_{n \in \mathbb{N}}$ *gegen \mathbf{a} konvergiert* und schreiben

$$\lim_{n \rightarrow \infty} \mathbf{a}_n = \mathbf{a}$$

oder

$$\mathbf{a}_n \longrightarrow \mathbf{a}.$$

- b. Wir nennen $(\mathbf{a}_n)_{n \in \mathbb{N}}$ genau dann *konvergent*, wenn es ein $\mathbf{a} \in \mathbb{K}$ gibt, so daß $(\mathbf{a}_n)_{n \in \mathbb{N}}$ gegen \mathbf{a} konvergiert. Andernfalls nennen wir $(\mathbf{a}_n)_{n \in \mathbb{N}}$ *divergent*.
- c. Wir nennen eine Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in \mathbb{K} eine *Nullfolge*, wenn $(\mathbf{a}_n)_{n \in \mathbb{N}}$ gegen 0 konvergiert, d.h. $\mathbf{a}_n \longrightarrow 0$.

Beispiel 11.5

- a. Die konstante Folge $(\mathbf{a}_n)_{n \in \mathbb{N}} = (\mathbf{c})_{n \in \mathbb{N}}$ konvergiert gegen \mathbf{c} , d.h. $\lim_{n \rightarrow \infty} \mathbf{c} = \mathbf{c}$.
Um das zu sehen, wählen wir für eine reelle Zahl $\varepsilon > 0$ die natürliche Zahl $n_\varepsilon = 0$, so daß für jedes $n \geq n_\varepsilon = 0$ gilt

$$|\mathbf{a}_n - \mathbf{c}| = |\mathbf{c} - \mathbf{c}| = 0 < \varepsilon.$$

- b. Die Folge $(\frac{1}{n})_{n \geq 1}$ konvergiert gegen 0 , d.h. $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$.
Denn: sei $0 < \varepsilon \in \mathbb{R}$ gegeben, so gibt es nach Korollar 9.4 eine natürliche Zahl n_ε , so daß $0 < \frac{1}{n_\varepsilon} < \varepsilon$. Ist nun $n \geq n_\varepsilon$, so folgt

$$\left| \frac{1}{n} - 0 \right| = \frac{1}{n} \leq \frac{1}{n_\varepsilon} < \varepsilon.$$

- c. Die Folge $(\mathbf{a}_n)_{n \in \mathbb{N}} = ((-1)^n)_{n \in \mathbb{N}}$ ist divergent.
Denn: nehmen wir an, $(\mathbf{a}_n)_{n \in \mathbb{N}}$ konvergiert gegen \mathbf{a} . Dann gibt es zu $\varepsilon = \frac{1}{2}$ ein n_ε , so daß $|\mathbf{a}_n - \mathbf{a}| < \varepsilon$ für $n \geq n_\varepsilon$. Insbesondere gilt dann wegen der Dreiecksungleichung

$$2 = |(-1)^{n_\varepsilon} - (-1)^{n_\varepsilon+1}| = |\mathbf{a}_{n_\varepsilon} - \mathbf{a}_{n_\varepsilon+1}| \leq |\mathbf{a}_{n_\varepsilon} - \mathbf{a}| + |\mathbf{a} - \mathbf{a}_{n_\varepsilon+1}| < \varepsilon + \varepsilon = 1,$$

was ein offensichtlicher Widerspruch ist.

Lemma 11.6 (Geometrische Folge)

Es sei $q \in \mathbb{K}$ mit $|q| < 1$, so ist $(q^n)_{n \in \mathbb{N}}$ eine Nullfolge.

Beweis: Wir können ohne Einschränkung annehmen, daß $q \neq 0$, da die Folge sonst sicher eine Nullfolge ist.

Sei $\varepsilon > 0$ gegeben. Wir betrachten die reelle Zahl

$$x := \frac{1}{|q|} - 1 > 0,$$

die positiv ist, da nach Voraussetzung $0 < |q| < 1$. Nach Korollar 9.4 gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$\frac{1}{n_\varepsilon} < x \cdot \varepsilon \quad (12)$$

Ist nun $n \geq n_\varepsilon$, so gilt wegen $|q| = \frac{1}{1+x}$ und der Bernoullischen Ungleichung auch

$$|q^n - 0| = |q|^n = \frac{1}{(1+x)^n} \stackrel{9.9}{\leq} \frac{1}{1+n \cdot x} < \frac{1}{n \cdot x} \leq \frac{1}{n_\varepsilon \cdot x} \stackrel{(12)}{<} \frac{x \cdot \varepsilon}{x} = \varepsilon.$$

□

Bemerkung 11.7

Für eine Folge $(a_n)_{n \in \mathbb{N}}$ in \mathbb{K} gilt offenbar:

$$a_n \longrightarrow a \iff a_n - a \longrightarrow 0 \iff |a_n - a| \longrightarrow 0.$$

Diese Feststellung ist in mancher Anwendung von Nutzen, um Argumente abzukürzen.

Proposition 11.8 (Eindeutigkeit des Grenzwertes von Folgen)

Der Grenzwert einer konvergenten Folge in \mathbb{K} ist eindeutig bestimmt.

Beweis: Nehmen wir an, eine Folge $(a_n)_{n \in \mathbb{N}}$ in \mathbb{K} besitze zwei verschiedene Grenzwerte $a, b \in \mathbb{K}$. Dann ist die reelle Zahl

$$\varepsilon := \frac{|a - b|}{2} > 0$$

positiv. Mithin gibt es zwei natürliche Zahlen $n_\varepsilon, n'_\varepsilon \in \mathbb{N}$, so daß

$$|a_n - a| < \varepsilon$$

für $n \geq n_\varepsilon$ und

$$|a_n - b| < \varepsilon$$

für $n \geq n'_\varepsilon$. Setzen wir nun $N := \max\{n_\varepsilon, n'_\varepsilon\}$, so gilt

$$|a - b| \leq |a - a_n| + |a_n - b| < \varepsilon + \varepsilon = |a - b|,$$

was ein offensichtlicher Widerspruch ist. □

B) Beschränkte Folgen

Definition 11.9 (Beschränkte Folgen)

Eine Folge $(a_n)_{n \in \mathbb{N}}$ in \mathbb{K} heißt *beschränkt*, wenn die Menge

$$\{|a_n| \in \mathbb{R} \mid n \in \mathbb{N}\}$$

beschränkt ist, d.h. wenn es eine Zahl $s \in \mathbb{R}$ gibt, so daß $|a_n| \leq s$ für alle $n \in \mathbb{N}$.

Man beachte dabei, daß die Menge stets durch 0 nach unten beschränkt ist, und wir nennen eine Zahl s wie oben eine Schranke für $(a_n)_{n \in \mathbb{N}}$.

Beispiel 11.10

- a. Die konvergente Folge $(\frac{1}{n})_{n \geq 1}$ ist beschränkt, da $|\frac{1}{n}| \leq 1$ für alle $n \geq 1$.

- b. Die divergente Folge $(\mathbf{a}_n)_{n \in \mathbb{N}} = ((-1)^n)_{n \in \mathbb{N}}$ ist ebenfalls beschränkt, da $\{|\mathbf{a}_n| \in \mathbb{R} \mid n \in \mathbb{N}\} = \{1\}$.

Satz 11.11 (Konvergente Folgen sind beschränkt.)

Jede konvergente Folge in \mathbb{K} ist beschränkt.

Beweis: Sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{K} und $\lim_{n \rightarrow \infty} \mathbf{a}_n = \mathbf{a}$. Dann gibt es zu $\varepsilon = 1$ eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$|\mathbf{a}_n - \mathbf{a}| < \varepsilon = 1$$

für $n \geq n_\varepsilon$. Setze

$$s := \max\{1 + |\mathbf{a}|, |\mathbf{a}_0|, |\mathbf{a}_1|, \dots, |\mathbf{a}_{n_\varepsilon - 1}|\},$$

wobei man beachte, daß das Maximum existiert, weil die Menge endlich ist.

Damit erhalten wir dann

$$|\mathbf{a}_n| \leq \begin{cases} s, & \text{falls } n < n_\varepsilon, \\ |\mathbf{a}_n - \mathbf{a}| + |\mathbf{a}| < 1 + |\mathbf{a}| \leq s, & \text{falls } n \geq n_\varepsilon. \end{cases}$$

Mithin ist s eine Schranke für $(\mathbf{a}_n)_{n \in \mathbb{N}}$. □

Beispiel 11.12

- Beispiel 11.10 zeigt, daß die Umkehrung von Satz 11.11 nicht gilt.
- Für $k \in \mathbb{N}$ mit $k \geq 1$ ist die Folge $(\mathbf{a}_n)_{n \in \mathbb{N}} = (n^k)_{n \in \mathbb{N}}$ nicht beschränkt, also auch nicht konvergent.
- Für $\mathbf{q} \in \mathbb{K}$ mit $|\mathbf{q}| > 1$ ist die Folge $(\mathbf{a}_n)_{n \in \mathbb{N}} = (\mathbf{q}^n)_{n \in \mathbb{N}}$ nicht beschränkt und somit divergent.

Um dies zu sehen, nehmen wir an, $s > 0$ sei eine Schranke für die Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ und setzen $\mathbf{x} := |\mathbf{q}| - 1 > 0$. Da \mathbb{R} archimedisch angeordnet ist (siehe Satz 9.3), gibt es eine natürliche Zahl $n \in \mathbb{N}$, so daß

$$s < n \cdot \mathbf{x}.$$

Aus der Bernoullischen Ungleichung erhalten wir damit

$$|\mathbf{q}|^n = (1 + \mathbf{x})^n \geq 1 + n \cdot \mathbf{x} > s,$$

was ein Widerspruch zur Wahl von s als Schranke von $(\mathbf{a}_n)_{n \in \mathbb{N}}$ ist. Dies zeigt, daß $(\mathbf{a}_n)_{n \in \mathbb{N}}$ nicht beschränkt ist.

Lemma 11.13

Ist $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Nullfolge in \mathbb{K} und $(\mathbf{b}_n)_{n \in \mathbb{N}}$ eine beschränkte Folge in \mathbb{K} , so ist $(\mathbf{a}_n \cdot \mathbf{b}_n)_{n \in \mathbb{N}}$ eine Nullfolge.

Beweis: Da $(\mathbf{b}_n)_{n \in \mathbb{N}}$ beschränkt ist, gibt es eine positive reelle Zahl $s \in \mathbb{R}_{>0}$ mit

$$|\mathbf{b}_n| \leq s$$

für alle $n \in \mathbb{N}$.

Sei nun $\varepsilon > 0$ gegeben. Wegen $\mathbf{a}_n \rightarrow 0$ gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$|\mathbf{a}_n - 0| < \frac{\varepsilon}{s}$$

für $n \geq n_\varepsilon$. Für $n \geq n_\varepsilon$ erhalten wir damit

$$|\mathbf{a}_n \cdot \mathbf{b}_n - 0| = |\mathbf{a}_n \cdot \mathbf{b}_n - 0 \cdot \mathbf{b}_n| = |\mathbf{a}_n - 0| \cdot |\mathbf{b}_n| \leq |\mathbf{a}_n - 0| \cdot s < \frac{\varepsilon}{s} \cdot s = \varepsilon.$$

Mithin konvergiert $(\mathbf{a}_n \cdot \mathbf{b}_n)_{n \in \mathbb{N}}$ gegen 0. \square

Beispiel 11.14

Da die Folge $(\frac{1}{n})_{n \geq 1}$ eine Nullfolge ist und da zudem die Folge $((-1)^n)_{n \geq 1}$ beschränkt ist, gilt

$$\frac{(-1)^n}{n} \rightarrow 0.$$

C) Grenzwertsätze und Konvergenzkriterien

Proposition 11.15 (Grenzwertsätze)

Seien $(\mathbf{a}_n)_{n \in \mathbb{N}}$ und $(\mathbf{b}_n)_{n \in \mathbb{N}}$ konvergente Folgen in \mathbb{K} mit $\mathbf{a}_n \rightarrow \mathbf{a}$ und $\mathbf{b}_n \rightarrow \mathbf{b}$.

- $\mathbf{a}_n + \mathbf{b}_n \rightarrow \mathbf{a} + \mathbf{b}$ und $\mathbf{a}_n - \mathbf{b}_n \rightarrow \mathbf{a} - \mathbf{b}$.
- $\mathbf{a}_n \cdot \mathbf{b}_n \rightarrow \mathbf{a} \cdot \mathbf{b}$.
- $|\mathbf{a}_n| \rightarrow |\mathbf{a}|$.
- Ist zudem $\mathbf{b} \neq 0$, so gibt es ein $n_0 \in \mathbb{N}$ mit $\mathbf{b}_n \neq 0$ für alle $n \geq n_0$ und die Folge $(\frac{\mathbf{a}_n}{\mathbf{b}_n})_{n \geq n_0}$ ist konvergent mit

$$\frac{\mathbf{a}_n}{\mathbf{b}_n} \rightarrow \frac{\mathbf{a}}{\mathbf{b}}.$$

Beweis:

- Sei $\varepsilon > 0$ gegeben. Dann gibt es natürliche Zahlen $n'_\varepsilon, n''_\varepsilon \in \mathbb{N}$, so daß

$$|\mathbf{a}_n - \mathbf{a}| < \frac{\varepsilon}{2}$$

für $n \geq n'_\varepsilon$ und

$$|\mathbf{b}_n - \mathbf{b}| < \frac{\varepsilon}{2}$$

für $n \geq n''_\varepsilon$. Mit $n_\varepsilon := \max\{n'_\varepsilon, n''_\varepsilon\}$ gilt dann

$$|(\mathbf{a}_n + \mathbf{b}_n) - (\mathbf{a} + \mathbf{b})| \leq |\mathbf{a}_n - \mathbf{a}| + |\mathbf{b}_n - \mathbf{b}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

für $n \geq n_\varepsilon$. Mithin konvergiert $(\mathbf{a}_n + \mathbf{b}_n)_{n \in \mathbb{N}}$ gegen $\mathbf{a} + \mathbf{b}$. Analog sieht man $\mathbf{a}_n - \mathbf{b}_n \rightarrow \mathbf{a} - \mathbf{b}$.

- Nach Satz 11.11 ist $(\mathbf{a}_n)_{n \in \mathbb{N}}$ als konvergente Folge beschränkt und nach Voraussetzung ist $(\mathbf{b}_n - \mathbf{b})_{n \in \mathbb{N}}$ eine Nullfolge, so daß

$$\mathbf{a}_n \cdot (\mathbf{b}_n - \mathbf{b}) \rightarrow 0$$

nach Lemma 11.13. Analog ist nach Voraussetzung $(\mathbf{a}_n - \mathbf{a})_{n \in \mathbb{N}}$ eine Nullfolge und die konstante Folge $(\mathbf{b})_{n \in \mathbb{N}}$ ist als konvergente Folge beschränkt, so daß

$$(\mathbf{a}_n - \mathbf{a}) \cdot \mathbf{b} \longrightarrow 0.$$

Aus a. folgt dann, daß die Summe der beiden Nullfolgen eine Nullfolge ist, d.h.

$$\mathbf{a}_n \cdot \mathbf{b}_n - \mathbf{a} \cdot \mathbf{b} = \mathbf{a}_n \cdot (\mathbf{b}_n - \mathbf{b}) + (\mathbf{a}_n - \mathbf{a}) \cdot \mathbf{b} \longrightarrow 0 + 0 = 0.$$

Also gilt auch $\mathbf{a}_n \cdot \mathbf{b}_n \longrightarrow \mathbf{a} \cdot \mathbf{b}$.

c. Ist $\varepsilon > 0$ gegeben, so gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$ mit

$$|\mathbf{a}_n - \mathbf{a}| < \varepsilon$$

für alle $n \geq n_\varepsilon$. Aber dann gilt nach Lemma 9.8 auch

$$||\mathbf{a}_n| - |\mathbf{a}|| \leq |\mathbf{a}_n - \mathbf{a}| < \varepsilon$$

für alle $n \geq n_\varepsilon$. Es folgt die Behauptung.

d. Wegen $\mathbf{b} \neq 0$ ist

$$\varepsilon := \frac{|\mathbf{b}|}{2} > 0$$

und es gibt ein $n_0 \in \mathbb{N}$ mit

$$|\mathbf{b}_n - \mathbf{b}| < \varepsilon = \frac{|\mathbf{b}|}{2}$$

für alle $n \geq n_0$. Mithin ist

$$|\mathbf{b}| = |\mathbf{b}_n + (\mathbf{b} - \mathbf{b}_n)| \leq |\mathbf{b}_n| + |\mathbf{b} - \mathbf{b}_n| = |\mathbf{b}_n| + |\mathbf{b}_n - \mathbf{b}| < |\mathbf{b}_n| + \frac{|\mathbf{b}|}{2},$$

so daß $0 < \frac{|\mathbf{b}|}{2} \leq |\mathbf{b}_n|$ für $n \geq n_0$. Insbesondere ist $\mathbf{b}_n \neq 0$ in diesen Fällen.

Aus Lemma 8.17 folgt zudem

$$0 < \frac{1}{|\mathbf{b}_n|} \leq \frac{2}{|\mathbf{b}|} \tag{13}$$

für $n \geq n_0$.

Ist nun $\varepsilon > 0$ beliebig gegeben, so gibt es eine natürliche Zahl $n_\varepsilon \geq n_0$ mit

$$|\mathbf{b}_n - \mathbf{b}| < \frac{\varepsilon \cdot |\mathbf{b}|^2}{2} \tag{14}$$

für alle $n \geq n_\varepsilon$. Für diese n erhalten wir damit

$$\left| \frac{1}{\mathbf{b}_n} - \frac{1}{\mathbf{b}} \right| = \frac{|\mathbf{b} - \mathbf{b}_n|}{|\mathbf{b}_n \cdot \mathbf{b}|} = \frac{1}{|\mathbf{b}_n|} \cdot \frac{1}{|\mathbf{b}|} \cdot |\mathbf{b}_n - \mathbf{b}| \stackrel{(13),(14)}{<} \frac{2}{|\mathbf{b}|^2} \cdot \frac{\varepsilon \cdot |\mathbf{b}|^2}{2} = \varepsilon.$$

Also gilt

$$\frac{1}{\mathbf{b}_n} \longrightarrow \frac{1}{\mathbf{b}},$$

und mit Teil b. folgt dann die Behauptung $\frac{\mathbf{a}_n}{\mathbf{b}_n} \longrightarrow \frac{\mathbf{a}}{\mathbf{b}}$.

□

Beispiel 11.16

a. Die Folge $(\frac{1}{n^2})_{n \geq 1}$ ist eine Nullfolge, da $\lim_{n \rightarrow \infty} \frac{1}{n^2} = \lim_{n \rightarrow \infty} \frac{1}{n} \cdot \lim_{n \rightarrow \infty} \frac{1}{n} = 0 \cdot 0 = 0$.

b. Die Folge $(a_n)_{n \in \mathbb{N}}$ mit

$$a_n = \frac{7n^2 + 3}{4n^2 + n + 1}$$

ist wegen der Grenzwertsätze konvergent, denn es gilt

$$a_n = \frac{7 + \frac{3}{n^2}}{4 + \frac{1}{n} + \frac{1}{n^2}} \longrightarrow \frac{7 + 0}{4 + 0 + 0} = \frac{7}{4}.$$

Proposition 11.17 (Einschachtelungssatz)

Seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ konvergente Folgen in \mathbb{R} mit $a_n \longrightarrow a$ und $b_n \longrightarrow b$, und sei $(c_n)_{n \in \mathbb{N}}$ eine weitere Folge reeller Zahlen.

- Ist $a_n \leq b_n$ für alle $n \geq n_0$, so ist $a \leq b$.
- Ist $a_n \leq c_n \leq b_n$ für alle $n \geq n_0$ und ist $a = b$, so gilt $c_n \longrightarrow a$.

Beweis:

a. Nehmen wir $b < a$ an, so gibt es für

$$\varepsilon := \frac{a - b}{2} > 0$$

natürliche Zahlen $n'_\varepsilon, n''_\varepsilon \in \mathbb{N}$ mit

$$a_n \in (a - \varepsilon, a + \varepsilon)$$

für alle $n \geq n'_\varepsilon$ und

$$b_n \in (b - \varepsilon, b + \varepsilon)$$

für alle $n \geq n''_\varepsilon$. Mithin gilt für $n = \max\{n_0, n'_\varepsilon, n''_\varepsilon\}$

$$a - \varepsilon < a_n \leq b_n < b + \varepsilon,$$

so daß

$$a - b < 2 \cdot \varepsilon = a - b. \quad \text{!}$$

b. Sei $\varepsilon > 0$ gegeben, dann gibt es natürliche Zahlen $n'_\varepsilon, n''_\varepsilon \in \mathbb{N}$ mit

$$|a_n - a| < \varepsilon$$

für alle $n \geq n'_\varepsilon$ und

$$|b_n - a| < \varepsilon$$

für alle $n \geq n''_\varepsilon$. Mithin gilt für $n \geq n_\varepsilon := \max\{n_0, n'_\varepsilon, n''_\varepsilon\}$ sicher

$$a - \varepsilon < a_n \leq c_n \leq b_n < a + \varepsilon,$$

d.h.

$$|c_n - a| < \varepsilon.$$

Also konvergiert $(c_n)_{n \in \mathbb{N}}$ gegen a .

□

Beispiel 11.18

Wegen $0 \leq \frac{1}{n^k} \leq \frac{1}{n}$ für alle $n \geq 1$ und $k \geq 1$ folgt aus $0 \rightarrow 0$ und $\frac{1}{n} \rightarrow 0$ auch

$$\frac{1}{n^k} \rightarrow 0.$$

Definition 11.19 (Monotone Folgen)

Es sei $(a_n)_{n \in \mathbb{N}}$ eine Folge reeller Zahlen. Wir nennen $(a_n)_{n \in \mathbb{N}}$ *monoton wachsend*, falls

$$a_n \leq a_{n+1}$$

für alle $n \in \mathbb{N}$.

Analog nennen wir $(a_n)_{n \in \mathbb{N}}$ *monoton fallend*, falls

$$a_n \geq a_{n+1}$$

für alle $n \in \mathbb{N}$.

Beispiel 11.20

Die Folge $(n)_{n \in \mathbb{N}}$ ist monoton wachsend und divergent, die Folge $(\frac{1}{n})_{n \geq 1}$ ist monoton fallend und konvergent.

Satz 11.21 (Monotoniekriterium)

Jede monoton wachsende oder fallende, beschränkte Folge in \mathbb{R} ist konvergent.

Beweis: Sei $(a_n)_{n \in \mathbb{N}}$ eine monoton wachsende, beschränkte Folge reeller Zahlen und sei $s > 0$ eine Schranke. Dann ist die Menge

$$A := \{a_n \mid n \in \mathbb{N}\}$$

nach oben beschränkt durch s , und somit existiert das Supremum

$$a := \sup(A).$$

Wir wollen zeigen, daß $(a_n)_{n \in \mathbb{N}}$ gegen a konvergiert.

Dazu sei $\varepsilon > 0$ gegeben. Dann ist $a - \varepsilon$ keine obere Schranke von A , so daß ein $n_\varepsilon \in \mathbb{N}$ existiert mit

$$a - \varepsilon < a_{n_\varepsilon}.$$

Da die Folge monoton wachsend ist, gilt dann aber für alle $n \geq n_\varepsilon$ auch

$$a - \varepsilon < a_{n_\varepsilon} \leq a_n \leq a < a + \varepsilon,$$

oder anders formuliert

$$|a - a_n| = |a_n - a| < \varepsilon.$$

Mithin haben wir $a_n \rightarrow a$ gezeigt. Der Fall einer monoton fallenden Folge wird analog mit Hilfe des Infimums bewiesen. \square

Bemerkung 11.22 (Supremum und Infimum sind Grenzwerte von Folgen.)

Es sei $\emptyset \neq A \subseteq \mathbb{R}$ eine nicht-leere Menge reeller Zahlen.

- a. Ist A nach oben beschränkt, so gibt es eine monoton wachsende Folge $(a_n)_{n \in \mathbb{N}}$ in A , die gegen $\sup(A)$ konvergiert.

- b. Ist A nach unten beschränkt, so gibt es eine monoton fallende Folge $(a_n)_{n \in \mathbb{N}}$ in A , die gegen $\inf(A)$ konvergiert.

Beweis: Sei zunächst A nach oben beschränkt. Wir wählen $a_0 \in A$ beliebig und setzen $a := \sup(A)$. Für $n \geq 1$ und $\varepsilon = \frac{1}{n}$ gibt es ein $b_n \in A$ mit $a - \varepsilon < b_n \leq a$. Setzen wir nun $a_n := \max\{b_n, a_{n-1}\} \in A$, so definieren wir auf diese Weise rekursiv eine offenbar monoton steigende Folge in A . Für diese gilt zudem

$$0 \leq |a_n - a| = a - a_n \leq a - b_n \leq \frac{1}{n} \longrightarrow 0,$$

woraus mit dem Einschachtelungssatz folgt, daß $(a_n)_{n \in \mathbb{N}}$ gegen a konvergiert.

Ist A nach unten beschränkt, so zeigt man die Aussage analog. \square

Beispiel 11.23 (Rekursive Folgen)

Es sei $c \in \mathbb{R}_{>0}$ eine positive reelle Zahl. Wir setzen $a_0 := 1$ und für $n \in \mathbb{N}$ definieren wir a_{n+1} durch die Rekursionsvorschrift

$$a_{n+1} := \frac{1}{2} \cdot \left(a_n + \frac{c}{a_n} \right) > 0.$$

Wir wollen zeigen, daß die Folge $(a_n)_{n \in \mathbb{N}}$ gegen \sqrt{c} konvergiert.

1. Schritt: $a_n^2 \geq c$ für alle $n \in \mathbb{N}$: Für $n \in \mathbb{N}$ gilt

$$0 \leq \left(a_n - \frac{c}{a_n} \right)^2 = a_n^2 - 2c + \frac{c^2}{a_n^2}.$$

Addieren wir auf beiden Seiten $4c$, so erhalten wir

$$0 \leq 4c \leq a_n^2 + 2c + \frac{c^2}{a_n^2} = \left(a_n + \frac{c}{a_n} \right)^2 = 4 \cdot a_{n+1}^2.$$

2. Schritt: $(a_n)_{n \geq 1}$ ist monoton fallend: Aus dem 1. Schritt wissen wir, daß $a_n^2 \geq c$ für $n \geq 1$ und mithin auch

$$a_n \geq \frac{c}{a_n}$$

für $n \geq 1$ gilt. Wir erhalten damit

$$a_{n+1} = \frac{1}{2} \cdot \left(a_n + \frac{c}{a_n} \right) \leq \frac{1}{2} \cdot (a_n + a_n) = a_n$$

für alle $n \geq 1$, so daß die Folge monoton fallend ist.

3. Schritt: $(a_n)_{n \geq 1}$ ist beschränkt: Denn $0 < a_n \leq a_1$ für alle $n \geq 1$.

4. Schritt: $a_n \longrightarrow \sqrt{c}$: Da die Folge $(a_n)_{n \geq 1}$ monoton fallend und beschränkt ist, folgt aus Satz 11.21 dann, daß sie konvergent ist, d.h. es gibt ein $a \in \mathbb{R}$ mit $a_n \longrightarrow a$. Den Grenzwert können wir nun mit Hilfe der Grenzwertsätze und der Eindeutigkeit des Grenzwertes bestimmen; es gilt nämlich

$$a \longleftarrow a_{n+1} = \frac{1}{2} \cdot \left(a_n + \frac{c}{a_n} \right) \longrightarrow \frac{1}{2} \cdot \left(a + \frac{c}{a} \right),$$

d.h.

$$a = \frac{1}{2} \cdot \left(a + \frac{c}{a} \right).$$

Lösen wir die Gleichung nach a auf, so erhalten wir

$$a^2 = c,$$

und da die Folgenglieder nie negativ sind, kann auch der Grenzwert nicht negativ sein (siehe Proposition 11.17). Damit ist also $a = \sqrt{c}$ nach Satz 9.10.

Beachte, daß man die Folge $(a_n)_{n \in \mathbb{N}}$ nutzen kann, um die Wurzel \sqrt{c} näherungsweise zu berechnen. Versucht dies einmal für $c = 2$ oder $c = 4$.

D) Der Satz von Bolzano-Weierstraß

Definition 11.24 (Teilfolge)

Es sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{K} und ist zudem

$$n_0 < n_1 < n_2 < n_3 < \dots$$

eine aufsteigende Folge natürlicher Zahlen, so nennen wir die Folge

$$(a_{n_k})_{k \in \mathbb{N}} = (a_{n_0}, a_{n_1}, a_{n_2}, a_{n_3}, \dots)$$

eine *Teilfolge* von $(a_n)_{n \in \mathbb{N}}$.

Beispiel 11.25

Die Folge $(\frac{1}{n^2})_{n \geq 1}$ ist eine Teilfolge von $(\frac{1}{n})_{n \geq 1}$.

Satz 11.26 (Bolzano-Weierstraß)

Jede beschränkte Folge in \mathbb{K} besitzt eine konvergente Teilfolge.

Beweis: Wir unterscheiden im Beweis die Fälle $\mathbb{K} = \mathbb{R}$ und $\mathbb{K} = \mathbb{C}$.

1. Fall: $\mathbb{K} = \mathbb{R}$: Da $(a_n)_{n \in \mathbb{N}}$ beschränkt ist, gibt es eine Zahl $s > 0$, so daß

$$-s \leq a_n \leq s$$

für alle $n \in \mathbb{N}$, d.h. das Intervall

$$[b_0, c_0] := [-s, s]$$

enthält unendlich viele Folgenglieder der Folge $(a_n)_{n \in \mathbb{N}}$ und wir wählen eines davon, a_{n_0} . Teilen wir das Intervall in zwei gleichgroße Hälften $[-s, 0]$ und $[0, s]$, so enthält mindestens eines der beiden neuen Intervalle wieder unendlich viele Folgenglieder. Wir wählen ein solches und nennen es $[b_1, c_1]$. Da es unendlich viele Folgenglieder enthält, enthält es auch ein a_{n_1} mit $n_1 > n_0$. Mit dem Intervall $[b_1, c_1]$ verfahren wir in der gleichen Weise und konstruieren so rekursiv eine Folge von Intervallen

$$[b_0, c_0] \supsetneq [b_1, c_1] \supsetneq [b_2, c_2] \supsetneq [b_3, c_3] \supsetneq \dots,$$

so daß jedes $[b_j, c_j]$ unendlich viele Folgenglieder von $(a_n)_{n \in \mathbb{N}}$ enthält. Zugleich konstruieren wir dabei eine Teilfolge

$$a_{n_0}, a_{n_1}, a_{n_2}, a_{n_3}, \dots$$

mit

$$b_j \leq a_{n_j} \leq c_j$$

und $n_0 < n_1 < n_2 < \dots$

Aufgrund der Konstruktion ist die Folge $(b_j)_{j \in \mathbb{N}}$ eine monoton wachsende beschränkte Folge und besitzt deshalb einen Grenzwert nach dem Monotoniekriterium 11.21, d.h.

$$b_j \longrightarrow b.$$

Analog besitzt $(c_j)_{j \in \mathbb{N}}$ als monoton fallende beschränkte Folge einen Grenzwert c . Da das Intervall $[b_n, c_n]$ aufgrund seiner Definition die Länge $\frac{2s}{2^n}$ hat, folgt dann

$$c - b \leftarrow c_n - b_n = \frac{2s}{2^n} = 2s \cdot \left(\frac{1}{2}\right)^n \longrightarrow 0,$$

wobei wir für die Konvergenz der rechten Seite die Eigenschaften der geometrischen Folge berücksichtigen (siehe Lemma 11.6). Wegen der Eindeutigkeit des Grenzwertes einer Folge gilt dann $b = c$, und aus dem Einschachtelungssatz folgt dann auch

$$\lim_{j \rightarrow \infty} a_{n_j} = b.$$

2. Fall: $\mathbb{K} = \mathbb{C}$: Aus Lemma 10.8 wissen wir, daß

$$|\operatorname{Re}(a_n)| \leq |a_n|,$$

so daß die Folge $(\operatorname{Re}(a_n))_{n \in \mathbb{N}}$ ebenfalls beschränkt ist. Da wir den Satz von Bolzano-Weierstraß für $\mathbb{K} = \mathbb{R}$ bereits bewiesen haben, gibt es also eine Teilfolge $(a_{n_k})_{k \in \mathbb{N}}$ und eine reelle Zahl b , so daß

$$\operatorname{Re}(a_{n_k}) \longrightarrow b.$$

Ebenfalls aus Lemma 10.8 folgt

$$|\operatorname{Im}(a_{n_k})| \leq |a_{n_k}|,$$

so daß auch die Folge $(\operatorname{Im}(a_{n_k}))_{k \in \mathbb{N}}$ beschränkt ist, und wieder folgt mittels des Satzes von Bolzano-Weierstraß für $\mathbb{K} = \mathbb{R}$, daß $(a_{n_k})_{k \in \mathbb{N}}$ eine Teilfolge $(a_{n_{k_j}})_{j \in \mathbb{N}}$ besitzt und daß es eine reelle Zahl c gibt, so daß

$$\operatorname{Im}(a_{n_{k_j}}) \longrightarrow c.$$

Aus Aufgabe 11.37 wissen wir, daß die Teilfolge $(\operatorname{Re}(a_{n_{k_j}}))_{j \in \mathbb{N}}$ von $(\operatorname{Re}(a_{n_k}))_{k \in \mathbb{N}}$ ebenfalls gegen b konvergiert, und aus Aufgabe 11.36 ergibt sich dann, daß auch die Folge $(a_{n_{k_j}})_{j \in \mathbb{N}}$ konvergent ist mit

$$a_{n_{k_j}} = \operatorname{Re}(a_{n_{k_j}}) + i \cdot \operatorname{Im}(a_{n_{k_j}}) \longrightarrow b + i \cdot c.$$

□

Beispiel 11.27

Die divergente beschränkte Folge $((-1)^n)_{n \in \mathbb{N}}$ besitzt als konvergente Teilfolge die konstante Folge $((-1)^{2k})_{k \in \mathbb{N}} = (1)_{k \in \mathbb{N}}$.

Satz 11.28 (Abgeschlossene Intervalle sind abgeschlossen.)

Ist $(a_n)_{n \in \mathbb{N}}$ eine konvergente Folge im abgeschlossenen Intervall $[a, b]$, so gilt

$$\lim_{n \rightarrow \infty} a_n \in [a, b].$$

Beweis: Nehmen wir an, $c := \lim_{n \rightarrow \infty} a_n \in \mathbb{R} \setminus [a, b]$. Ist $c > b$, so gibt es zu $\varepsilon := \frac{c-b}{2} > 0$ ein n_ε für das unter anderem $0 < c - b \leq c - a_{n_\varepsilon} = |a_{n_\varepsilon} - c| < \varepsilon = \frac{c-b}{2}$ gilt, was ein offensichtlicher Widerspruch ist. Analog sieht man auch, daß $c < a$ nicht möglich ist. \square

E) Das Cauchy-Kriterium

Definition 11.29 (Cauchy-Folge)

Eine Folge $(a_n)_{n \in \mathbb{N}}$ in \mathbb{K} heißt *Cauchy-Folge*, falls

$$\forall 0 < \varepsilon \in \mathbb{R} \exists n_\varepsilon \in \mathbb{N} : \forall m > n \geq n_\varepsilon : |a_m - a_n| < \varepsilon.$$

Satz 11.30 (Cauchy-Kriterium: \mathbb{K} ist vollständig.)

Eine Folge in \mathbb{K} ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.

Beweis:

\implies : Wir setzen voraus, daß $(a_n)_{n \in \mathbb{N}}$ eine konvergente Folge ist mit Grenzwert a . Sei nun $\varepsilon > 0$ gegeben, dann gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$|a_n - a| < \frac{\varepsilon}{2}$$

für alle $n \geq n_\varepsilon$. Für zwei natürliche Zahlen $m > n \geq n_\varepsilon$ folgt dann mit der Dreiecksungleichung

$$|a_m - a_n| \leq |a_m - a| + |a - a_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Also ist $(a_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge.

\impliedby : Sei nun umgekehrt $(a_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge. Wir wollen zeigen, daß $(a_n)_{n \in \mathbb{N}}$ konvergent ist und müssen dazu einen Grenzwert von $(a_n)_{n \in \mathbb{N}}$ finden, was nicht ganz leicht ist. Unsere Idee hierzu ist, daß wir eine konvergente Teilfolge von $(a_n)_{n \in \mathbb{N}}$ mit Hilfe des Satzes von Bolzano-Weierstraß finden und dann zeigen, daß deren Grenzwert auch ein Grenzwert von $(a_n)_{n \in \mathbb{N}}$ ist.

1. Schritt: Zeige, daß $(a_n)_{n \in \mathbb{N}}$ beschränkt ist.¹

Zu $\varepsilon = 1$ gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$|a_m - a_n| < \varepsilon = 1$$

¹Der Beweis geht wie der Beweis von Satz 11.11, wenn man dort den Grenzwert a durch a_{n_ε} ersetzt.

für alle $m > n \geq n_\varepsilon$. Setze

$$s := \max\{1 + |a_{n_\varepsilon}|, |a_0|, |a_1|, \dots, |a_{n_\varepsilon-1}|\}.$$

Damit erhalten wir dann

$$|a_n| \leq \begin{cases} s, & \text{falls } n < n_\varepsilon, \\ |a_n - a_{n_\varepsilon}| + |a_{n_\varepsilon}| < 1 + |a_{n_\varepsilon}| \leq s, & \text{falls } n \geq n_\varepsilon. \end{cases}$$

Mithin ist s eine Schranke für $(a_n)_{n \in \mathbb{N}}$.

2. Schritt: Aufgrund des Satzes von Bolzano-Weierstraß 11.26 besitzt $(a_n)_{n \in \mathbb{N}}$ also eine konvergente Teilfolge $(a_{n_k})_{k \in \mathbb{N}}$, und wir setzen

$$a := \lim_{k \rightarrow \infty} a_{n_k}.$$

3. Schritt: Zeige, $a_n \rightarrow a$.

Sei dazu $\varepsilon > 0$ gegeben. Da $(a_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge ist, gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$ mit

$$|a_m - a_n| < \frac{\varepsilon}{2}$$

für alle $m > n \geq n_\varepsilon$. Da zudem $a_{n_k} \rightarrow a$ existiert auch ein $k_\varepsilon \in \mathbb{N}$ mit

$$|a_{n_k} - a| < \frac{\varepsilon}{2}$$

für alle $k \geq k_\varepsilon$. Wir wählen nun eine Zahl $k \geq k_\varepsilon$ so, daß $n_k \geq n_\varepsilon$. Dann gilt für jedes $n \geq n_\varepsilon$ auch

$$|a_n - a| \leq |a_n - a_{n_k}| + |a_{n_k} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Also konvergiert $(a_n)_{n \in \mathbb{N}}$ gegen a .

□

Beispiel 11.31

Ist $1 \neq q \in \mathbb{K}$ mit $|q| = 1$, so ist die Folge $(q^n)_{n \in \mathbb{N}}$ keine Cauchy-Folge und mithin auch nicht konvergent.

Um dies zu sehen, betrachten wir $\varepsilon = |q-1| > 0$ und $n_\varepsilon \in \mathbb{N}$ beliebig. Für $m = n_\varepsilon + 1$ und $n = n_\varepsilon$ gilt dann

$$|q^m - q^n| = |q|^{n_\varepsilon} \cdot |q - 1| = 1^{n_\varepsilon} \cdot \varepsilon = \varepsilon.$$

Wäre die Folge eine Cauchy-Folge, so müßte der Ausdruck für ein geeignetes n_ε echt kleiner als ε werden.

Bemerkung 11.32 (\mathbb{Q} ist nicht vollständig.)

Eine Cauchy-Folge rationaler Zahlen muß in \mathbb{Q} nicht konvergent sein, d.h. ihr Grenzwert in \mathbb{R} muß keine rationale Zahl sein. Zum Beispiel ist $\sqrt{2}$ keine rationale Zahl (siehe Satz 9.12) und ist $\sqrt{2} = \sum_{i=1}^{\infty} c_i \cdot 10^{-i}$ ihre Dezimalzahldarstellung, so wird durch

$$a_n = \sum_{i=1}^n c_i \cdot 10^{-i}$$

eine Folge $(a_n)_{n \in \mathbb{N}}$ rationaler Zahlen definiert, die in \mathbb{R} gegen $\sqrt{2}$ konvergiert und mithin eine Cauchy-Folge ist, deren Grenzwert $\sqrt{2}$ aber nicht in \mathbb{Q} liegt.

Man sagt auch, die rationalen Zahlen sind nicht vollständig (siehe auch Abschnitt 39.G)). Dieses Manko der rationalen Zahlen erfordert den Übergang zu den reellen Zahlen. Mit dem gleichen Argument wie für $\sqrt{2}$ sieht man übrigens, daß jede reelle Zahl Grenzwert einer Folge rationaler Zahlen ist. Dies liegt daran, daß \mathbb{Q} dicht in \mathbb{R} liegt (siehe Satz 9.6).

F) Bestimmt divergente Folgen

Definition 11.33 (Bestimmte Divergenz)

Es sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{R} .

- a. Wir sagen, daß $(a_n)_{n \in \mathbb{N}}$ *bestimmt divergiert gegen* ∞ , falls

$$\forall s > 0 \exists n_s \in \mathbb{N} : \forall n \geq n_s : a_n > s.$$

In diesem Fall schreiben wir $a_n \rightarrow \infty$ oder $\lim_{n \rightarrow \infty} a_n = \infty$, und nennen ∞ auch den uneigentlichen *Grenzwert* von $(a_n)_{n \in \mathbb{N}}$.

- b. Analog sagen wir, daß $(a_n)_{n \in \mathbb{N}}$ *bestimmt divergiert gegen* $-\infty$, falls

$$\forall s < 0 \exists n_s \in \mathbb{N} : \forall n \geq n_s : a_n < s.$$

In diesem Fall schreiben wir $a_n \rightarrow -\infty$ oder $\lim_{n \rightarrow \infty} a_n = -\infty$, und nennen $-\infty$ auch den uneigentlichen *Grenzwert* von $(a_n)_{n \in \mathbb{N}}$.

Eine Folge, die bestimmt divergiert nennen wir *bestimmt divergent*.

Beispiel 11.34

Die Folge $(n)_{n \in \mathbb{N}}$ ist bestimmt divergent mit Grenzwert ∞ , die Folge $((-1)^n \cdot n)_{n \in \mathbb{N}}$ ist divergent, aber nicht bestimmt divergent.

Bemerkung 11.35 (Grenzwertsätze für uneigentliche Grenzwerte)

Wir einigen uns für $a \in \mathbb{R}$ auf die folgenden Rechenregeln:

- $a + \infty := \infty$ und $a - \infty := -\infty$.
- $a \cdot \infty := \infty$, $a \cdot -\infty := -\infty$ und $\frac{a}{0} := \infty$, falls $a > 0$.
- $a \cdot \infty := -\infty$, $a \cdot -\infty := \infty$ und $\frac{a}{0} := -\infty$, falls $a < 0$.
- $\frac{a}{\infty} := 0$ und $\frac{a}{-\infty} := 0$.
- $\frac{\infty}{0} := \infty$ und $\frac{-\infty}{0} := -\infty$.

Damit lassen sich die Grenzwertsätze für Folgen 11.15 verallgemeinern auf Fälle unter Einbeziehung von bestimmt divergenten Folgen. Wann immer man bei der Anwendung der Grenzwertsätze als Grenzwert einen der obigen Ausdrücke erhält, kann man den Grenzwert auf dem Weg berechnen. Die Beweise sind einfach, aber es gilt viele Fälle zu unterscheiden. Z.B. gelten:

- a. Wenn $a_n \rightarrow a > 0$ und $b_n \rightarrow 0$ mit $b_n > 0$, so gilt $\frac{a_n}{b_n} \rightarrow \frac{a}{0} = \infty$.

- b. Wenn $\mathbf{a}_n \longrightarrow \mathbf{a}$ und $\mathbf{b}_n \longrightarrow \infty$, so gilt $\frac{\mathbf{a}_n}{\mathbf{b}_n} \longrightarrow \frac{\mathbf{a}}{\infty} = 0$.

Aufgaben

Aufgabe 11.36

Es sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge komplexer Zahlen. Zeige, daß die folgenden Aussagen äquivalent sind:

- $\mathbf{a}_n \longrightarrow \mathbf{a}$.
- $\operatorname{Re}(\mathbf{a}_n) \longrightarrow \operatorname{Re}(\mathbf{a})$ und $\operatorname{Im}(\mathbf{a}_n) \longrightarrow \operatorname{Im}(\mathbf{a})$.

Aufgabe 11.37

Ist $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{K} und $\sigma: \mathbb{N} \longrightarrow \mathbb{N}$ bijektiv, so nennen wir die Folge

$$(\mathbf{a}_{\sigma(n)})_{n \in \mathbb{N}} = (\mathbf{a}_{\sigma(0)}, \mathbf{a}_{\sigma(1)}, \mathbf{a}_{\sigma(2)}, \mathbf{a}_{\sigma(3)}, \dots)$$

eine *Umordnung* von $(\mathbf{a}_n)_{n \in \mathbb{N}}$. Beweise die folgenden beiden Aussagen.

- Wenn $(\mathbf{a}_n)_{n \in \mathbb{N}}$ gegen \mathbf{a} konvergiert, so konvergiert jede Teilfolge von $(\mathbf{a}_n)_{n \in \mathbb{N}}$ gegen \mathbf{a} .
- Wenn $(\mathbf{a}_n)_{n \in \mathbb{N}}$ gegen \mathbf{a} konvergiert, so konvergiert jede Umordnung von $(\mathbf{a}_n)_{n \in \mathbb{N}}$ gegen \mathbf{a} .

Aufgabe 11.38

- Zeige, daß die Folge $(s_n)_{n \in \mathbb{N}}$ mit

$$s_n := \sum_{k=0}^n \frac{1}{k!}$$

konvergent ist.

- Zeige, daß die Folge $(t_n)_{n \in \mathbb{N}}$ mit

$$t_n := \left(1 + \frac{1}{n}\right)^n$$

konvergent ist.

- Zeige, daß die Grenzwerte von $(s_n)_{n \in \mathbb{N}}$ und $(t_n)_{n \in \mathbb{N}}$ übereinstimmen. Wir nennen den Grenzwert die *Eulersche Zahl* e .

Hinweis zu Teil c., zeige hierfür, daß der Grenzwert von $(t_n)_{n \in \mathbb{N}}$ nach unten durch s_m beschränkt ist.

Aufgabe 11.39

Untersuche die folgenden Folgen $(\mathbf{a}_n)_{n \geq 1}$ auf Konvergenz und berechne gegebenenfalls den Grenzwert:

- $\mathbf{a}_n = \frac{n^4 - 3n + 5}{3n^5 + 6n^3 + 11}$.
- $\mathbf{a}_n = \frac{(3n+1) \cdot (n+1)^2 - 5(n-1)}{1+n+n^2}$.
- $\mathbf{a}_n = \sum_{i=1}^n \frac{i}{n^2}$.

d. $\mathbf{a}_n = \frac{n}{n^2+1} + \dots + \frac{n}{n^2+n}$.

e. $\mathbf{a}_n = \frac{n}{2^n}$.

Aufgabe 11.40

Zeige, dass die rekursiv definierte Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ mit $\mathbf{a}_0 = 1$ und

$$\mathbf{a}_{n+1} := \sqrt{1 + \mathbf{a}_n}$$

konvergiert und bestimme ihren Grenzwert.

Hinweis: Prüfe die Folge (bzw. eine geeignete Teilfolge) auf Monotonie und Beschränktheit. Für die Berechnung des Grenzwertes können dann die Grenzwertsätze geeignet angewandt werden.

Aufgabe 11.41

Zeige, dass die rekursiv definierte Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ mit $\mathbf{a}_0 = 1$ und

$$\mathbf{a}_{n+1} := 1 + \frac{1}{\mathbf{a}_n}$$

konvergiert und bestimme ihren Grenzwert.

Hinweis: Prüfe die Folge (bzw. eine geeignete Teilfolge) auf Monotonie und Beschränktheit. Für die Berechnung des Grenzwertes können dann die Grenzwertsätze geeignet angewandt werden.

Aufgabe 11.42

- a. Sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge reeller Zahlen, so daß $|\mathbf{a}_{n+1} - \mathbf{a}_n| < \frac{1}{2^n}$ für alle $n \in \mathbb{N}$.
Zeige, daß es sich um eine konvergente Folge handelt.
- b. Bleibt die Behauptung aus Aufgabenteil a. korrekt, wenn wir die Bedingung $|\mathbf{a}_{n+1} - \mathbf{a}_n| < \frac{1}{n}$ voraussetzen? Gib einen Beweis oder ein Gegenbeispiel an.

§ 12 Unendliche Reihen

A) Konvergenz unendlicher Reihen

Definition 12.1 (Unendliche Reihen)

Ist $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{K} , so nennen wir die Folge $(s_n)_{n \in \mathbb{N}}$ mit

$$s_n := \sum_{k=0}^n \mathbf{a}_k$$

der *Partialsummen* von $(\mathbf{a}_n)_{n \in \mathbb{N}}$ auch die durch $(\mathbf{a}_n)_{n \in \mathbb{N}}$ definierte *Reihe*.

Die Reihe heißt *konvergent*, wenn $(s_n)_{n \in \mathbb{N}}$ eine konvergente Folge ist, und andernfalls heißt sie *divergent*.

Wir bezeichnen sowohl die Reihe $(s_n)_{n \in \mathbb{N}}$ selbst, als auch ihren Grenzwert, sofern er existiert, mit

$$\sum_{n=0}^{\infty} \mathbf{a}_n.$$

Beachte, wie stets bei Folgen müssen weder $(\mathbf{a}_n)_{n \geq n_0}$ noch $(s_m)_{m \geq n_0} = \sum_{n=n_0}^{\infty} \mathbf{a}_n$ (wobei $s_m = \sum_{k=n_0}^m \mathbf{a}_k$) mit dem Index 0 starten!

Beispiel 12.2 (Teleskopsumme)

Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n \cdot (n+1)}$ ist konvergent mit Grenzwert $\sum_{n=1}^{\infty} \frac{1}{n \cdot (n+1)} = 1$.

Dazu beachten wir, daß $\frac{1}{k \cdot (k+1)} = \frac{1}{k} - \frac{1}{k+1}$ gilt, so daß

$$\begin{aligned} s_n &= \sum_{k=1}^n \frac{1}{k \cdot (k+1)} = \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k+1} \right) \\ &= \left(\frac{1}{1} - \frac{1}{2} \right) + \left(\frac{1}{2} - \frac{1}{3} \right) + \left(\frac{1}{3} - \frac{1}{4} \right) + \dots + \left(\frac{1}{n-1} - \frac{1}{n} \right) + \left(\frac{1}{n} - \frac{1}{n+1} \right) \\ &= 1 - \frac{1}{n+1} \longrightarrow 1. \end{aligned}$$

Summen, die sich wie s_n auf zwei Summanden reduzieren, weil sich die übrigen Teile der Summe sukzessive auslöschen, nennt man *Teleskopsummen*.

Beispiel 12.3 (Harmonische Reihe)

Die *harmonische Reihe* $\sum_{n=1}^{\infty} \frac{1}{n}$ ist divergent.

Denn für $n_k = 2^k$ mit $k \in \mathbb{N}$ gilt

$$\begin{aligned} s_{n_k} &= \sum_{i=1}^{2^k} \frac{1}{i} \\ &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots + \left(\frac{1}{2^{k-1}+1} + \dots + \frac{1}{2^k}\right) \\ &\geq 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \dots + \left(\frac{1}{2^k} + \dots + \frac{1}{2^k}\right) \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \dots + \frac{1}{2} \\ &= 1 + \frac{k}{2} \longrightarrow \infty, \end{aligned}$$

so daß $(s_{n_k})_k \in \mathbb{N}$ eine divergente Teilfolge der Folge der Partialsummen ist, weshalb letztere nicht konvergent sein kann.

Lemma 12.4 (Grenzwertsätze für konvergente Reihen)

Seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ zwei konvergente Reihen in \mathbb{K} und $a \in \mathbb{K}$.

- $\sum_{n=0}^{\infty} (a_n + b_n) = \sum_{n=0}^{\infty} a_n + \sum_{n=0}^{\infty} b_n.$
- $\sum_{n=0}^{\infty} (a_n - b_n) = \sum_{n=0}^{\infty} a_n - \sum_{n=0}^{\infty} b_n.$
- $\sum_{n=0}^{\infty} a \cdot a_n = a \cdot \sum_{n=0}^{\infty} a_n.$
- $\mathbb{K} = \mathbb{R}$ und $a_n \leq b_n$ für alle $n \in \mathbb{N}$, so ist $\sum_{n=0}^{\infty} a_n \leq \sum_{n=0}^{\infty} b_n.$

Insbesondere, sind die Reihen in a.-c. konvergent.

Beweis: Die Aussagen folgen unmittelbar aus den Grenzwertsätzen für Folgen 11.15 sowie aus Proposition 11.17 angewendet auf die Folgen der Partialsummen. \square

B) Konvergenzkriterien für unendliche Reihen

Proposition 12.5 (Cauchy-Kriterium für Reihen)

Sei $\sum_{n=0}^{\infty} a_n$ eine Reihe in \mathbb{K} . Genau dann ist $\sum_{n=0}^{\infty} a_n$ konvergent, wenn

$$\forall \varepsilon > 0 \exists n_\varepsilon \in \mathbb{N} : \forall m > n \geq n_\varepsilon : \left| \sum_{k=n+1}^m a_k \right| < \varepsilon.$$

Beweis: Die Aussage folgt unmittelbar aus dem Cauchy-Kriterium für Folgen 11.30, da

$$s_m - s_n = \sum_{k=n+1}^m a_k.$$

\square

Lemma 12.6 (Restglieder einer konvergenten Reihe)

Ist die Reihe $\sum_{k=0}^{\infty} a_k$ konvergent, so ist die Folge der Restglieder eine Nullfolge, d.h.

$$\lim_{n \rightarrow \infty} \sum_{k=n}^{\infty} a_k = 0.$$

Beweis: Zu $\varepsilon > 0$ gibt es wegen des Cauchy-Kriteriums für Reihen ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $m > n \geq n_\varepsilon$

$$\left| \sum_{k=n+1}^m a_k \right| < \varepsilon.$$

Halten wir n fest und betrachten die linke Seite als eine Folge mit Index m , so erhalten wir aus dem Einschachtelungssatz 11.17

$$\left| \sum_{k=n+1}^{\infty} a_k \right| = \lim_{m \rightarrow \infty} \left| \sum_{k=n+1}^m a_k \right| \leq \varepsilon$$

für alle $n \geq n_\varepsilon$. Also ist die Folge der Restglieder eine Nullfolge. \square

Lemma 12.7 (Nullfolgekriterium)

Ist $\sum_{n=0}^{\infty} a_n$ eine konvergente Reihe in \mathbb{K} , so ist $(a_n)_{n \in \mathbb{N}}$ eine Nullfolge.

Beweis: Man beachte, daß die Partialsummen s_n der Reihe folgende Eigenschaft erfüllen:

$$a_n = s_n - s_{n-1}.$$

Aus den Grenzwertsätzen für Folgen 11.15 folgt deshalb, daß $(a_n)_{n \in \mathbb{N}}$ als Differenz zweier konvergenter Folgen konvergent ist und daß

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} s_n - \lim_{n \rightarrow \infty} s_{n-1} = \sum_{n=0}^{\infty} a_n - \sum_{n=0}^{\infty} a_n = 0.$$

\square

Wer statt der Grenzwertsätze lieber Epsilons verwenden möchte, der bevorzugt vielleicht den folgenden alternativen Beweis des Nullfolgekriteriums mit Hilfe des Cauchy-Kriteriums.

Alternativer Beweis des Nullfolgekriteriums: Sei $\varepsilon > 0$ gegeben. Dann gibt es wegen des Cauchy-Kriteriums 12.5 ein $n'_\varepsilon \in \mathbb{N}$, so daß für $m > n \geq n'_\varepsilon$ auch

$$\left| \sum_{k=n+1}^m a_k \right| < \varepsilon. \quad (15)$$

Für $N \geq n_\varepsilon := n'_\varepsilon + 1$ gilt mit $m := N$ und $n := N - 1$ mithin

$$|a_N - 0| = |a_N| = \left| \sum_{k=n+1}^m a_k \right| \stackrel{(15)}{<} \varepsilon.$$

Mithin konvergiert $(a_n)_{n \in \mathbb{N}}$ gegen Null. \square

Beispiel 12.8

- a. Die Reihe $\sum_{n=0}^{\infty} \left(1 + \frac{1}{n}\right)^n$ ist divergent, da $\left(\left(1 + \frac{1}{n}\right)^n\right)_{n \in \mathbb{N}}$ keine Nullfolge ist.
- b. Die Umkehrung von Lemma 12.7 gilt nicht, wie das Beispiel der harmonischen Reihe zeigt.

Satz 12.9 (Geometrische Reihe)

Es sei $q \in \mathbb{K}$.

- a. Ist $|q| < 1$, so ist die geometrische Reihe $\sum_{n=0}^{\infty} q^n$ konvergent mit Grenzwert

$$\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}.$$

- b. Ist $|q| \geq 1$, so ist die geometrische Reihe $\sum_{n=0}^{\infty} q^n$ divergent.

Beweis:

- a. Aus Satz 7.12 wissen wir

$$s_n = \sum_{k=0}^n q^k = \frac{1 - q^{n+1}}{1 - q},$$

und da $|q| < 1$ gilt dann wegen Lemma 11.6

$$s_n = \frac{1 - q^{n+1}}{1 - q} \longrightarrow \frac{1 - 0}{1 - q} = \frac{1}{1 - q}.$$

- b. Für $|q| \geq 1$ ist die Folge $(q^n)_{n \in \mathbb{N}}$ keine Nullfolge (siehe Beispiele 11.12 und 11.31), und somit ist die Reihe $\sum_{n=0}^{\infty} q^n$ aufgrund des Nullfolgenkriteriums 12.7 divergent.

□

Satz 12.10 (Leibniz-Kriterium für alternierende Reihen in \mathbb{R})

Ist $(a_n)_{n \in \mathbb{N}}$ eine monoton fallende Nullfolge in \mathbb{R} , so konvergiert die Reihe

$$\sum_{n=0}^{\infty} (-1)^n \cdot a_n.$$

Beweis: Es sei wieder $s_n = \sum_{k=0}^n (-1)^k \cdot a_k$ die n -te Partialsumme der Reihe. Wir betrachten nun zunächst die Teilfolge $(s_{2n})_{n \in \mathbb{N}}$ der *geraden* Partialsummen. Für $n \in \mathbb{N}$ gilt dann

$$s_{2 \cdot (n+1)} = s_{2n} - (a_{2n+1} - a_{2n+2}) \leq s_{2n},$$

da nach Voraussetzung $a_{2n+1} \geq a_{2n+2}$. Die Folge $(s_{2n})_{n \in \mathbb{N}}$ ist also monoton fallend.

Analog sieht man, daß die Folge $(s_{2n+1})_{n \in \mathbb{N}}$ der *ungeraden* Partialsummen monoton steigend ist, denn

$$s_{2 \cdot (n+1) + 1} = s_{2n+1} + (a_{2n+2} - a_{2n+3}) \geq s_{2n+1}.$$

Damit sind beide Folgen dann aber auch beschränkt, denn

$$s_1 \leq s_{2n+1} = s_{2n} - a_{2n+1} \leq s_{2n} \leq s_0$$

für $n \in \mathbb{N}$. Aufgrund des Monotoniekriteriums 11.21 sind also beide Folgen konvergent, d.h. es gibt reelle Zahlen $s, t \in \mathbb{R}$ mit

$$s_{2n} \longrightarrow s \quad \text{und} \quad s_{2n+1} \longrightarrow t.$$

Aus den Grenzwertsätzen für Folgen erhalten wir dann

$$s - t \longleftarrow s_{2n} - s_{2n+1} = a_{2n+1} \longrightarrow 0,$$

so daß $s = t$ gilt.

Sei nun $\varepsilon > 0$ gegeben. Dann gibt es natürliche Zahlen $n'_\varepsilon, n''_\varepsilon \in \mathbb{N}$, so daß

$$|s_{2n} - s| < \varepsilon$$

für alle $n \geq n'_\varepsilon$ und

$$|s_{2n+1} - s| < \varepsilon$$

für alle $n \geq n''_\varepsilon$. Setzen wir nun $n_\varepsilon = \max\{2 \cdot n'_\varepsilon, 2 \cdot n''_\varepsilon + 1\}$, so gilt für $n \geq n_\varepsilon$ offenbar

$$|s_n - s| < \varepsilon.$$

Also ist die Folge $(s_n)_{n \in \mathbb{N}}$ konvergent. □

Beispiel 12.11 (Alternierende harmonische Reihe)

Die alternierende harmonische Reihe $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ ist konvergent. Aus dem Beweis des Leibnizkriteriums wissen wir zudem, daß

$$-1 = s_1 \leq \sum_{n=1}^{\infty} \frac{(-1)^n}{n} \leq s_2 = -\frac{1}{2}.$$

Mehr können wir im Augenblick über den Grenzwert der Reihe nicht sagen (siehe dazu Beispiel 18.33).

Lemma 12.12 (Umklammern in Reihen)

Es sei $\sum_{n=0}^{\infty} a_n$ eine konvergente Reihe in \mathbb{K} und $0 = k_0 < k_1 < k_2 < \dots$ eine aufsteigende Folge natürlicher Zahlen. Setzen wir

$$b_n := \sum_{k=k_n}^{k_{n+1}-1} a_k = a_{k_n} + a_{k_n+1} + \dots + a_{k_{n+1}-1}$$

so ist die Reihe $\sum_{n=0}^{\infty} b_n$ konvergent und es gilt

$$\sum_{n=0}^{\infty} a_n = \sum_{n=0}^{\infty} b_n.$$

Beweis: Ist $(s_n)_{n \in \mathbb{N}}$ die Folge der Partialsummen zu $\sum_{n=0}^{\infty} a_n$ und $(t_n)_{n \in \mathbb{N}}$ die Folge der Partialsummen zu $\sum_{n=0}^{\infty} b_n$, so gilt

$$t_n = s_{k_{n+1}-1}$$

für $n \in \mathbb{N}$. Also ist $(t_n)_{n \in \mathbb{N}}$ eine Teilfolge von $(s_n)_{n \in \mathbb{N}}$ und konvergiert wegen Aufgabe 11.37 gegen den gleichen Grenzwert. □

C) Absolut konvergente Reihen

Definition 12.13 (Umordnung)

Es sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{K} und $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ bijektiv. Wir nennen die Folge

$$(\mathbf{a}_{\sigma(n)})_{n \in \mathbb{N}} = (\mathbf{a}_{\sigma(0)}, \mathbf{a}_{\sigma(1)}, \mathbf{a}_{\sigma(2)}, \mathbf{a}_{\sigma(3)}, \dots)$$

eine *Umordnung* von $(\mathbf{a}_n)_{n \in \mathbb{N}}$ und die Reihe

$$\sum_{n=0}^{\infty} \mathbf{a}_{\sigma(n)} = \mathbf{a}_{\sigma(0)} + \mathbf{a}_{\sigma(1)} + \mathbf{a}_{\sigma(2)} + \mathbf{a}_{\sigma(3)} + \dots$$

eine *Umordnung* der Reihe $\sum_{n=0}^{\infty} \mathbf{a}_n$.

Beispiel 12.14

Betrachten wir folgende Umordnung der alternierenden harmonischen Reihe

$$\left(-\frac{1}{1} + \frac{1}{2}\right) + \frac{1}{4} + \left(-\frac{1}{3} + \frac{1}{6}\right) + \frac{1}{8} + \left(-\frac{1}{5} + \frac{1}{10}\right) + \frac{1}{12} + \left(-\frac{1}{7} + \frac{1}{14}\right) + \frac{1}{16} + \dots,$$

d.h. in den Klammern sind sukzessive die ungeraden Folgenglieder \mathbf{a}_n zusammen jeweils mit dem zugehörigen Folgenglied \mathbf{a}_{2n} aufgeführt, und zwischen den Klammernausdrücken stehen der Reihe nach die Folgenglieder, deren Index durch 4 teilbar ist. Es ist klar, daß man auf dem Weg alle Glieder der harmonischen Reihe auflistet. Wenn diese Umordnung der harmonischen Reihe wieder konvergent ist, so können wir wegen Lemma 12.12 zur Berechnung des Grenzwertes auch die Klammern wie angegeben setzen. Der Grenzwert der Reihe ist dann

$$-\frac{1}{2} + \frac{1}{4} - \frac{1}{6} + \frac{1}{8} - \frac{1}{10} + \frac{1}{12} - \frac{1}{14} + \frac{1}{16} - \dots = \frac{1}{2} \cdot \sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

genau die Hälfte des Grenzwertes der harmonischen Reihe. Daraus ergibt sich folgende Erkenntnis:

Durch Umordnung einer konvergenten Reihe kann sich der Grenzwert ändern.

Definition 12.15

Eine Reihe $\sum_{n=0}^{\infty} \mathbf{a}_n$ in \mathbb{K} heißt *absolut konvergent*, wenn die Reihe ihrer *Absolutbeträge* $\sum_{n=0}^{\infty} |\mathbf{a}_n|$ konvergiert. Da die Folge der Partialsummen $\mathbf{t}_n := \sum_{k=0}^n |\mathbf{a}_k|$ monoton wächst, ist dies gleichwertig dazu, daß die Folge $(\mathbf{t}_n)_{n \in \mathbb{N}}$ beschränkt ist (siehe Monotoniekriterium 11.21 und Satz 11.11).

Beispiel 12.16

Die alternierende harmonische Reihe ist konvergent, aber nicht absolut konvergent.

Lemma 12.17

Ist $\sum_{n=0}^{\infty} \mathbf{a}_n$ in \mathbb{K} *absolut konvergent*, so ist sie auch konvergent.

Beweis: Sei $\varepsilon > 0$ gegeben. Dann gibt es wegen des Cauchy-Kriteriums für Reihen eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$\left| \sum_{k=n+1}^m |a_k| \right| < \varepsilon$$

für alle $m > n \geq n_\varepsilon$, da die Reihe $\sum_{n=0}^{\infty} |a_n|$ konvergiert. Aus der Dreiecksungleichung wissen wir nun aber, daß dann auch

$$\left| \sum_{k=n+1}^m a_k \right| \leq \sum_{k=n+1}^m |a_k| = \left| \sum_{k=n+1}^m |a_k| \right| < \varepsilon$$

für alle $m > n \geq n_\varepsilon$ gilt. Mithin ist die Reihe $\sum_{n=0}^{\infty} a_n$ nach dem Cauchy-Kriterium für Reihen konvergent. \square

Man beachte, daß Beispiel 12.14 zeigt, daß im folgenden Satz die Voraussetzung *absolut konvergent* nicht durch die Bedingung *konvergent* ersetzt werden kann.

Satz 12.18 (Umordnungssatz)

Jede Umordnung einer absolut konvergenten Reihe ist absolut konvergent und konvergiert gegen den gleichen Grenzwert.

Beweis: Sei $\sum_{n=0}^{\infty} a_n$ eine absolut konvergente Reihe und $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ sei bijektiv.

Wir wollen zunächst zeigen, daß die Reihe $\sum_{n=0}^{\infty} (a_n - a_{\sigma(n)})$ gegen Null konvergiert. Sei dazu $\varepsilon > 0$ gegeben. Da die Reihe $\sum_{n=0}^{\infty} a_n$ absolut konvergent ist, gibt es ein $n'_\varepsilon \in \mathbb{N}$, so daß

$$\sum_{k=n+1}^m |a_k| = \left| \sum_{k=n+1}^m |a_k| \right| < \varepsilon$$

für alle $m > n \geq n'_\varepsilon$ gilt. Da die Abbildung σ surjektiv und die Menge $\{0, 1, \dots, n'_\varepsilon\}$ endlich ist, gibt es eine Zahl $n_\varepsilon \geq n'_\varepsilon$ mit

$$\{0, 1, \dots, n'_\varepsilon\} \subseteq \{\sigma(0), \sigma(1), \dots, \sigma(n_\varepsilon)\}.$$

Für $n \geq n_\varepsilon$ heben sich deshalb in der Partialsumme

$$\sum_{k=0}^n (a_k - a_{\sigma(k)})$$

die a_i mit $i \leq n'_\varepsilon$ heraus, da sie einmal mit positivem und einmal mit negativem Vorzeichen auftreten. Die übrigen a_i können sich herausheben oder auch nicht; in letzterem Fall kommen sie in der Summe genau einmal (entweder mit positivem oder mit negativem Vorzeichen) vor. Setzen wir nun

$$m := \max\{\sigma(0), \sigma(1), \dots, \sigma(n), 0, 1, \dots, n\} + 1 > n'_\varepsilon,$$

so erhalten wir insgesamt

$$\left| \sum_{k=0}^n (a_k - a_{\sigma(k)}) - 0 \right| \leq \sum_{k=n'_\varepsilon+1}^m |a_k| < \varepsilon.$$

Also konvergiert die Reihe $\sum_{n=0}^{\infty} (\mathbf{a}_n - \mathbf{a}_{\sigma(n)})$ gegen Null.

Aus den Grenzwertsätzen für Reihen 12.4 erhalten wir deshalb, daß die Reihe

$$\sum_{n=0}^{\infty} \mathbf{a}_{\sigma(n)} = \sum_{n=0}^{\infty} \mathbf{a}_n - \sum_{n=0}^{\infty} (\mathbf{a}_n - \mathbf{a}_{\sigma(n)}) = \sum_{n=0}^{\infty} \mathbf{a}_n$$

konvergent ist mit dem Grenzwert $\sum_{n=0}^{\infty} \mathbf{a}_n$.

Wenden wir das Ergebnis auf die konvergente Reihe $\sum_{n=0}^{\infty} |\mathbf{a}_n|$ und ihre Umordnung $\sum_{n=0}^{\infty} |\mathbf{a}_{\sigma(n)}|$ an, so folgt auch, daß die Umordnung absolut konvergent ist. \square

D) Konvergenzkriterien für absolute Konvergenz

Satz 12.19 (Majorantenkriterium)

Es seien $\sum_{n=0}^{\infty} \mathbf{a}_n$ und $\sum_{n=0}^{\infty} \mathbf{b}_n$ zwei Reihen in \mathbb{K} . Ist $\sum_{n=0}^{\infty} \mathbf{b}_n$ absolut konvergent und

$|\mathbf{a}_n| \leq |\mathbf{b}_n|$ für alle $n \geq n_0$, so ist auch $\sum_{n=0}^{\infty} \mathbf{a}_n$ absolut konvergent.

Wir nennen $\sum_{n=0}^{\infty} \mathbf{b}_n$ dann eine konvergente Majorante von $\sum_{n=0}^{\infty} \mathbf{a}_n$.

Beweis: Die Folge von Partialsummen $(s_n)_{n \in \mathbb{N}}$ mit

$$s_n := \sum_{k=0}^n |\mathbf{a}_k|$$

ist beschränkt durch $s_{n_0} + \sum_{n=0}^{\infty} |\mathbf{b}_n|$. Also ist die Reihe absolut konvergent. \square

Proposition 12.20 (Minorantenkriterium)

Es seien $\sum_{n=0}^{\infty} \mathbf{a}_n$ und $\sum_{n=0}^{\infty} \mathbf{b}_n$ zwei Reihen in \mathbb{R} .

Ist $\sum_{n=0}^{\infty} \mathbf{b}_n$ divergent und $\mathbf{a}_n \geq \mathbf{b}_n \geq 0$ für alle $n \in \mathbb{N}$, so ist $\sum_{n=0}^{\infty} \mathbf{a}_n$ divergent.

Wir nennen $\sum_{n=0}^{\infty} \mathbf{b}_n$ dann eine divergente Minorante von $\sum_{n=0}^{\infty} \mathbf{a}_n$.

Beweis: Wegen $\mathbf{b}_n \geq 0$ ist die Folge der Partialsummen $(t_n)_{n \in \mathbb{N}}$ mit

$$t_n := \sum_{k=0}^n \mathbf{b}_k$$

monoton wachsend. Da die Folge $(t_n)_{n \in \mathbb{N}}$ nach Voraussetzung divergent ist, ist sie wegen des Monotoniekriteriums für Folgen 11.21 nicht beschränkt. Aber dann ist auch die Folge der Partialsummen $(s_n)_{n \in \mathbb{N}}$ mit

$$s_n := \sum_{k=0}^n \mathbf{a}_k$$

unbeschränkt, wegen $s_n \geq t_n$, und mithin ist sie divergent nach Satz 11.11. \square

Beispiel 12.21

Für $k \geq 2$ ist die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^k}$ konvergent.

Dazu betrachten wir zunächst den Fall $k = 2$. Wegen

$$a_n := \frac{1}{(n+1)^2} \leq \frac{1}{n \cdot (n+1)} =: b_n$$

ist wegen Beispiel 12.2 die Reihe $\sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} \frac{1}{n \cdot (n+1)}$ eine konvergente Majorante von $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} \frac{1}{(n+1)^2}$. Nehmen wir nun noch eine Indexverschiebung vor, so sehen wir, daß die Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \sum_{n=1}^{\infty} \frac{1}{(n+1)^2}$$

ebenfalls konvergent ist. Für den Fall $k > 2$ gilt nun wegen $0 \leq \frac{1}{n^k} \leq \frac{1}{n^2}$, daß die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$ ihrerseits eine konvergente Majorante von $\sum_{n=1}^{\infty} \frac{1}{n^k}$ ist.

Satz 12.22 (Wurzelkriterium)

Es sei $\sum_{n=0}^{\infty} a_n$ eine Reihe in \mathbb{K} .

- Existiert ein $q < 1$ mit $\sqrt[n]{|a_n|} \leq q$ für $n \geq n_0$, so ist $\sum_{n=0}^{\infty} a_n$ absolut konvergent.
- Ist $\sqrt[n]{|a_n|} \geq 1$ für alle $n \geq n_0$, so ist $\sum_{n=0}^{\infty} a_n$ divergent.

Beweis: Ist $q < 1$ mit $\sqrt[n]{|a_n|} \leq q$ für $n \geq n_0$, d.h. $|a_n| \leq q^n$, so ist die geometrische Reihe $\sum_{n=0}^{\infty} q^n$ nach Satz 12.9 eine konvergente Majorante von $\sum_{n=0}^{\infty} a_n$.

Falls $\sqrt[n]{|a_n|} \geq 1$ für alle $n \geq n_0$, d.h. $|a_n| \geq 1^n = 1$, so ist $(a_n)_{n \in \mathbb{N}}$ keine Nullfolge und mithin ist $\sum_{n=0}^{\infty} a_n$ wegen des Nullfolgekriteriums divergent. \square

Satz 12.23 (Quotientenkriterium)

Es sei $\sum_{n=0}^{\infty} a_n$ eine Reihe in \mathbb{K} mit $a_n \neq 0$ für alle $n \geq n_0$.

- Existiert ein $q < 1$ mit $\left| \frac{a_{n+1}}{a_n} \right| \leq q$ für $n \geq n_0$, so ist $\sum_{n=0}^{\infty} a_n$ absolut konvergent.
- Ist $\left| \frac{a_{n+1}}{a_n} \right| \geq 1$ für alle $n \geq n_0$, so ist $\sum_{n=0}^{\infty} a_n$ divergent.

Beweis: Wenn eine reelle Zahl $0 < q < 1$ existiert mit $\left| \frac{a_{n+1}}{a_n} \right| \leq q$ für $n \geq n_0$, so gilt

$$|a_{n+1}| \leq q \cdot |a_n|$$

für alle $n \geq n_0$, und mit Induktion sieht man dann, daß

$$|a_n| \leq q \cdot |a_{n-1}| \leq q^2 \cdot |a_{n-2}| \leq \dots \leq q^{n-n_0} \cdot |a_{n_0}|.$$

Also ist die geometrische Reihe

$$\frac{|a_{n_0}|}{q^{n_0}} \cdot \sum_{n=0}^{\infty} q^n = \sum_{n=0}^{\infty} q^{n-n_0} \cdot |a_{n_0}|$$

nach Satz 12.9 eine konvergente Majorante von $\sum_{n=0}^{\infty} \mathbf{a}_n$.

Ist $\left| \frac{\mathbf{a}_{n+1}}{\mathbf{a}_n} \right| \geq 1$ für alle $n \geq n_0$, so ist $|\mathbf{a}_{n+1}| \geq |\mathbf{a}_n| \neq 0$ für alle $n \geq n_0$. Mithin ist $(\mathbf{a}_n)_{n \in \mathbb{N}}$ keine Nullfolge und die Reihe $\sum_{n=0}^{\infty} \mathbf{a}_n$ ist wegen des Nullfolgekriteriums dann divergent. \square

Korollar 12.24 (Praktikables Quotienten-/Wurzelkriterium)

Sei $\sum_{n=0}^{\infty} \mathbf{a}_n$ eine Reihe in \mathbb{K} mit $\mathbf{a}_n \neq 0$ für alle $n \geq n_0$.

- Falls $\lim_{n \rightarrow \infty} \left| \frac{\mathbf{a}_{n+1}}{\mathbf{a}_n} \right| < 1$ oder $\lim_{n \rightarrow \infty} \sqrt[n]{|\mathbf{a}_n|} < 1$, so ist $\sum_{n=0}^{\infty} \mathbf{a}_n$ absolut konvergent.
- Falls $\lim_{n \rightarrow \infty} \left| \frac{\mathbf{a}_{n+1}}{\mathbf{a}_n} \right| > 1$ oder $\lim_{n \rightarrow \infty} \sqrt[n]{|\mathbf{a}_n|} > 1$, so ist $\sum_{n=0}^{\infty} \mathbf{a}_n$ divergent.
- Im Fall $\lim_{n \rightarrow \infty} \left| \frac{\mathbf{a}_{n+1}}{\mathbf{a}_n} \right| = 1$ oder $\lim_{n \rightarrow \infty} \sqrt[n]{|\mathbf{a}_n|} = 1$ wird keine Aussage getroffen!

Beweis: Falls $\lim_{n \rightarrow \infty} \left| \frac{\mathbf{a}_{n+1}}{\mathbf{a}_n} \right| < 1$ bzw. $\lim_{n \rightarrow \infty} \sqrt[n]{|\mathbf{a}_n|} < 1$, so kann man Satz 12.22 bzw. Satz 12.23 mit

$$q := \frac{1 + \lim_{n \rightarrow \infty} \sqrt[n]{|\mathbf{a}_n|}}{2} \quad \text{bzw.} \quad q := \frac{1 + \lim_{n \rightarrow \infty} \left| \frac{\mathbf{a}_{n+1}}{\mathbf{a}_n} \right|}{2}$$

anwenden.

Falls $\lim_{n \rightarrow \infty} \sqrt[n]{|\mathbf{a}_n|} > 1$ bzw. $\lim_{n \rightarrow \infty} \left| \frac{\mathbf{a}_{n+1}}{\mathbf{a}_n} \right| > 1$, so ist sicher $\sqrt[n]{|\mathbf{a}_n|} > 1$ bzw. $\left| \frac{\mathbf{a}_{n+1}}{\mathbf{a}_n} \right| > 1$ für n hinreichend groß, so daß die Aussage ebenfalls aus Satz 12.22 bzw. Satz 12.23 folgt. \square

Bemerkung 12.25

Man beachte, daß die harmonische Reihe $\sum_{n \geq 1} \mathbf{a}_n = \sum_{n \geq 1} \frac{1}{n}$ divergent ist, obwohl stets

$$\left| \frac{\mathbf{a}_{n+1}}{\mathbf{a}_n} \right| = \frac{1}{\frac{n+1}{n}} = 1 - \frac{1}{n} < 1$$

gilt. Aber, es gibt kein $q < 1$ mit

$$\left| \frac{\mathbf{a}_{n+1}}{\mathbf{a}_n} \right| = 1 - \frac{1}{n} < q$$

für alle hinreichend großen n . Das Quotientenkriterium ist deshalb nicht anwendbar. Beachte auch, daß in diesem Fall

$$\lim_{n \rightarrow \infty} \left| \frac{\mathbf{a}_{n+1}}{\mathbf{a}_n} \right| = \lim_{n \rightarrow \infty} 1 - \frac{1}{n} = 1$$

gilt.

Beispiel 12.26

Die Reihe $\sum_{n=1}^{\infty} \frac{n^2}{n!}$ ist absolut konvergent, da

$$\left| \frac{\frac{(n+1)^2}{(n+1)!}}{\frac{n^2}{n!}} \right| = \frac{(n+1)^2 \cdot n!}{n^2 \cdot (n+1)!} = \frac{n+1}{n^2} = \frac{1}{n} + \frac{1}{n^2} \rightarrow 0 + 0 = 0.$$

Satz 12.27 (Cauchy-Produkt)

Es seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ zwei absolut konvergente Reihen in \mathbb{K} . Für $n \in \mathbb{N}$ setzen wir

$$c_n := \sum_{k=0}^n a_k \cdot b_{n-k} = \sum_{i+j=n} a_i \cdot b_j.$$

Dann ist die Reihe $\sum_{n=0}^{\infty} c_n$ absolut konvergent und es gilt

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} a_n \cdot \sum_{n=0}^{\infty} b_n.$$

Beweis: Wir konstruieren zunächst eine bijektive Abbildung $\sigma : \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$, indem wir die Elemente in $\mathbb{N} \times \mathbb{N}$ in der in Abbildung 1 angegebenen Weise durchlaufen.

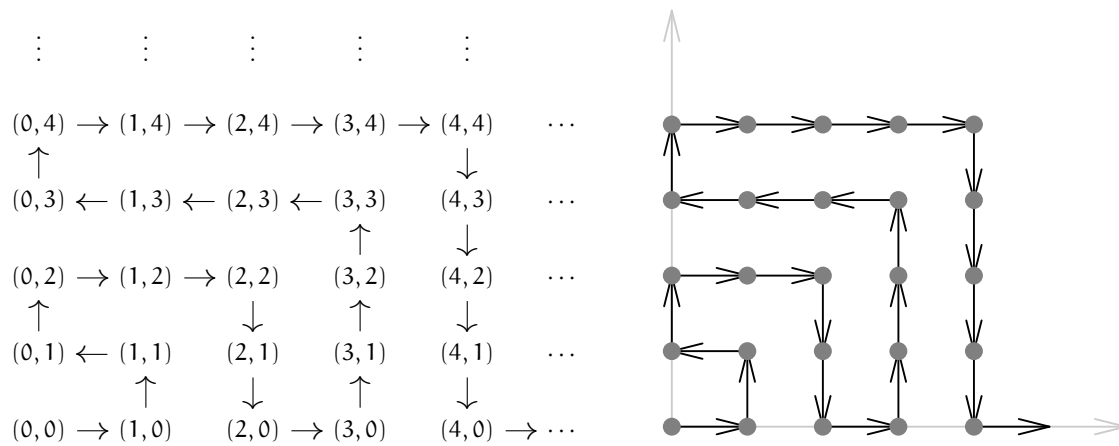


ABBILDUNG 1. Die Bijektion $\sigma : \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$

Aufgrund der Definition von σ gilt für $m \in \mathbb{N}$ offenbar

$$\{\sigma(0), \sigma(1), \dots, \sigma((m+1)^2 - 1)\} = \{(k, l) \mid 0 \leq k, l \leq m\} \quad (16)$$

Dann definieren wir uns eine Folge $(d_n)_{n \in \mathbb{N}}$ durch

$$d_n := a_k \cdot b_l, \quad \text{wenn } (k, l) = \sigma(n).$$

Wir wollen nun zunächst zeigen, daß die Reihe $\sum_{n=0}^{\infty} d_n$ absolut konvergent ist. Dazu beachten wir, daß für $m \in \mathbb{N}$ die Ungleichung

$$\sum_{n=0}^m |d_n| \leq \sum_{n=0}^{(m+1)^2-1} |d_n| \stackrel{(16)}{=} \sum_{k=0}^m \sum_{l=0}^m |a_k \cdot b_l| = \sum_{k=0}^m |a_k| \cdot \sum_{l=0}^m |b_l| \leq \sum_{k=0}^{\infty} |a_k| \cdot \sum_{l=0}^{\infty} |b_l|$$

erfüllt ist. Mithin ist die Folge der Partialsummen von $\sum_{n=0}^{\infty} |d_n|$ nach oben beschränkt. Damit ist $\sum_{n=0}^{\infty} d_n$ absolut konvergent und deshalb auch konvergent.

Zudem folgt aus den Grenzwertsätzen

$$\sum_{i=0}^{\infty} d_i \longleftarrow \sum_{i=0}^{(n^2+1)-1} d_i \stackrel{(16)}{=} \sum_{k=0}^n \sum_{l=0}^n a_k \cdot b_l = \sum_{k=0}^n a_k \cdot \sum_{l=0}^n b_l \longrightarrow \sum_{k=0}^{\infty} a_k \cdot \sum_{l=0}^{\infty} b_l,$$

und wegen der Eindeutigkeit des Grenzwertes gilt dann zudem

$$\sum_{n=0}^{\infty} d_n = \sum_{n=0}^{\infty} a_n \cdot \sum_{n=0}^{\infty} b_n.$$

Diese absolut konvergente Reihe $\sum_{n=0}^{\infty} d_n$ werden wir nun umordnen. Dazu konstruieren wir uns nach dem Cantorschen Diagonalverfahren eine weitere bijektive Abbildung $\pi: \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$ wie in Abbildung 2 angedeutet.

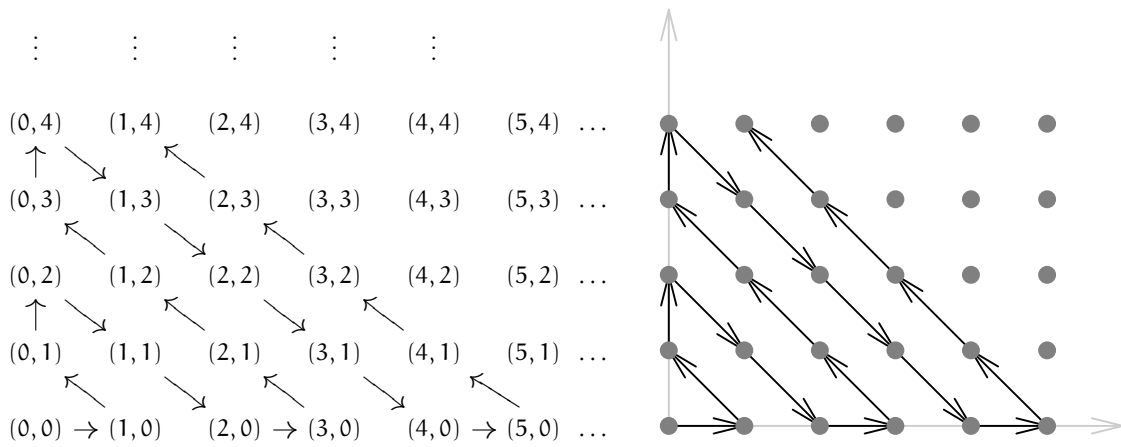


ABBILDUNG 2. Die Bijektion $\pi: \mathbb{N} \rightarrow \mathbb{N} \times \mathbb{N}$

Wir setzen nun

$$e_n := d_{\sigma^{-1}(\pi(n))} = a_k \cdot b_l, \quad \text{wenn } \pi(n) = (k, l)$$

für $n \in \mathbb{N}$ und erhalten so eine Umordnung $(e_n)_{n \in \mathbb{N}} = (d_{\sigma^{-1}(\pi(n))})_{n \in \mathbb{N}}$ der Folge $(d_n)_{n \in \mathbb{N}}$. Wegen des Umordnungssatzes 12.18 ist dann auch die Reihe $\sum_{n=0}^{\infty} e_n$ absolut konvergent mit dem gleichen Grenzwert

$$\sum_{n=0}^{\infty} e_n = \sum_{n=0}^{\infty} d_n.$$

Nun entsteht die Reihe $\sum_{n=0}^{\infty} c_n$ offenbar aus der Reihe $\sum_{n=0}^{\infty} e_n$ durch Einfügen von Klammern² im Sinne von Lemma 12.12. Mithin ist die Reihe nach eben diesem

²Wir wollen dies in der Fußnote etwas ausführen. Aufgrund der Definition von π und unter Verwendung der Formel in Beispiel 7.11 zur Berechnung der Summe der ersten n Zahlen sieht man, daß für $n \in \mathbb{N}$ folgende Gleichheit gilt

$$\{(k, l) \mid k + l = n\} = \left\{ \pi(i) \mid \frac{n \cdot (n+1)}{2} \leq i \leq \frac{(n+2) \cdot (n+1)}{2} - 1 \right\}. \quad (17)$$

Lemma ebenfalls konvergent mit dem gleichen Grenzwert, d.h.

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} e_n = \sum_{n=0}^{\infty} d_n = \sum_{n=0}^{\infty} a_n \cdot \sum_{n=0}^{\infty} b_n.$$

Genauso können wir Lemma 12.12 auch auf die Reihen $\sum_{n=0}^{\infty} |e_n|$ und $\sum_{n=0}^{\infty} |c_n|$ anwenden, um zu sehen, daß letztere konvergiert und mithin $\sum_{n=0}^{\infty} c_n$ absolut konvergent ist. \square

E) Potenzreihen

Definition 12.28

Es sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{K} , $a \in \mathbb{K}$ und t eine Veränderliche.

Wir nennen einen Ausdruck der Form $\sum_{n=0}^{\infty} a_n \cdot (t - a)^n$ eine *Potenzreihe* über \mathbb{K} in der Veränderlichen t mit *Entwicklungspunkt* a . Im folgenden beschränken wir uns im wesentlichen auf den Fall $a = 0$ und schreiben dann einfach $\sum_{n=0}^{\infty} a_n \cdot t^n$.

Unser Ziel ist es, für die Veränderliche t Werte $x \in \mathbb{K}$ einzusetzen und so eine Reihe zu erhalten, die konvergiert oder auch nicht.

Lemma 12.29

Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in \mathbb{K} und $y \in \mathbb{K}$, so daß die Reihe $\sum_{n=0}^{\infty} a_n \cdot y^n$ konvergiert. Dann ist die Reihe $\sum_{n=0}^{\infty} a_n \cdot x^n$ absolut konvergent für alle $x \in \mathbb{K}$ mit $|x| < |y|$.

Beweis: Da die Reihe $\sum_{n=0}^{\infty} a_n \cdot y^n$ konvergent ist, ist die Folge $(a_n \cdot y^n)_{n \in \mathbb{N}}$ nach dem Nullfolgekriterium eine Nullfolge, und mithin ist sie auch beschränkt. D.h. es gibt ein $s > 0$ mit

$$|a_n \cdot y^n| \leq s$$

für alle $n \in \mathbb{N}$. Für $x \in \mathbb{K}$ mit $|x| < |y|$ setzen wir $q := \frac{|x|}{|y|} < 1$ und erhalten dann

$$|a_n \cdot x^n| = |a_n \cdot y^n| \cdot \frac{|x|^n}{|y|^n} \leq s \cdot q^n.$$

Also ist die geometrische Reihe $\sum_{n=0}^{\infty} s \cdot q^n = s \cdot \sum_{n=0}^{\infty} q^n$ eine konvergente Majorante von $\sum_{n=0}^{\infty} a_n \cdot x^n$, so daß diese nach dem Majorantenkriterium absolut konvergiert. \square

Notation 12.30

Wir wollen den Begriff des Supremums etwas erweitern, indem wir $\sup(\emptyset) := -\infty$

Für c_n ergibt sich daraus

$$c_n = \sum_{k+l=n} a_k \cdot b_l = \sum_{i=\frac{n \cdot (n+1)}{2}}^{\frac{(n+2) \cdot (n+1)}{2} - 1} e_i = e_{\frac{n \cdot (n+1)}{2}} + e_{\frac{n \cdot (n+1)}{2} + 1} + \dots + e_{\frac{(n+2) \cdot (n+1)}{2} - 1},$$

d.h. $\sum_{n=0}^{\infty} c_n$ entsteht aus $\sum_{n=0}^{\infty} e_n$ durch Zusammenfassung von Summanden mittels Einfügen von Klammern.

setzen und $\sup(A) := \infty$, falls $A \subseteq \mathbb{R}$ nicht nach oben beschränkt ist. Damit gilt für jede Teilmenge $A \subseteq \mathbb{R}$

$$\sup(A) \in \mathbb{R} \cup \{\infty, -\infty\}.$$

Wir erinnern uns, daß wir in Bemerkung 11.35 bereits folgende Konventionen für $x \in \mathbb{R}$ eingeführt haben: $\frac{x}{\infty} := \frac{x}{-\infty} := 0$, und $\frac{x}{0} := \infty$ für $x > 0$ sowie $\frac{x}{0} := -\infty$ für $x < 0$.

Definition 12.31

Für eine Potenzreihe $\sum_{n=0}^{\infty} a_n \cdot t^n$ über \mathbb{K} nennen wir

$$r := \sup \left\{ |y| \mid y \in \mathbb{K}, \sum_{n=0}^{\infty} a_n \cdot y^n \text{ ist konvergent} \right\} \in \mathbb{R}_{\geq 0} \cup \{\infty\}$$

den *Konvergenzradius* der Potenzreihe.

Man beachte, daß die Potenzreihe zumindest für $y = 0$ konvergiert, so daß die angegebene Menge nicht-leer ist!

Satz 12.32 (Konvergenzradius)

Es sei $\sum_{n=0}^{\infty} a_n \cdot t^n$ eine Potenzreihe über \mathbb{K} mit Konvergenzradius r .

- Ist $x \in \mathbb{K}$ mit $|x| < r$, so ist $\sum_{n=0}^{\infty} a_n \cdot x^n$ absolut konvergent.
- Ist $x \in \mathbb{K}$ mit $|x| > r$, so ist $\sum_{n=0}^{\infty} a_n \cdot x^n$ divergent.

Setzen wir $U_r(0) := \{x \in \mathbb{K} \mid |x| < r\}$, so definiert die Potenzreihe also eine Abbildung

$$U_r(0) \longrightarrow \mathbb{K} : x \mapsto \sum_{n=0}^{\infty} a_n \cdot x^n.$$

Wir nennen $U_r(0)$ den Konvergenzbereich der Potenzreihe.

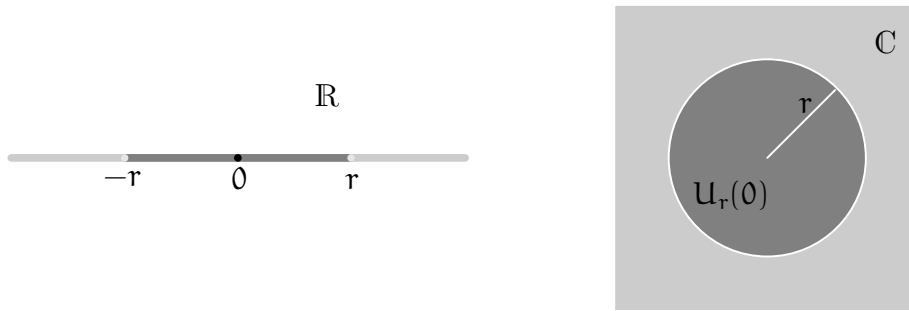
Bemerkung 12.33

- Über den Fall $|x| = r$ wird in Satz 12.32 keine Aussage getroffen! Wir nennen die Menge $\{x \mid |x| = r\}$ den *Rand* des Konvergenzbereiches.
- Konvergenzradius $r = \infty$ heißt, daß $\sum_{n=0}^{\infty} a_n \cdot x^n$ für alle $x \in \mathbb{K}$ absolut konvergent ist.
- Ist $\mathbb{K} = \mathbb{R}$, so ist die Menge $U_r(0) = (-r, r)$ ein offenes Intervall; ist $\mathbb{K} = \mathbb{C}$, so ist die Menge $U_r(0)$ ein Kreis mit Radius r um den Ursprung. In Abbildung 3 stellen wir den Konvergenzbereich der Reihe graphisch dar.

Beweis von Satz 12.32: a. Wir betrachten die Menge

$$A := \left\{ |y| \mid y \in \mathbb{K}, \sum_{n=0}^{\infty} a_n \cdot y^n \text{ ist konvergent} \right\},$$

so daß $r = \sup(A)$. Ist $r = \infty$, so ist A unbeschränkt und zu jedem $x \in \mathbb{K}$ gibt es ein $y \in A$ mit $|y| > |x|$. Ist $|x| < r < \infty$, so ist $\varepsilon := \frac{r-|x|}{2} > 0$ und

ABBILDUNG 3. Konvergenzbereich $U_r(0)$

$r - \varepsilon = \sup(A) - \varepsilon$ ist keine obere Schranke von A . Es gibt also ein $y \in A$ mit $|y| > r - \varepsilon = \frac{r+|x|}{2} > |x|$. In beiden Fällen $\sum_{n=0}^{\infty} a_n \cdot y^n$ ist konvergent und $|x| < |y|$, und nach Lemma 12.29 ist $\sum_{n=0}^{\infty} a_n \cdot x^n$ mithin absolut konvergent.

- b. Ist $|x| > r$, so ist $|x| \notin A$ und mithin muß $\sum_{n=0}^{\infty} a_n \cdot x^n$ divergent sein.

□

Satz 12.34 (Cauchy-Hadamard)

Sei $\sum_{n=0}^{\infty} a_n \cdot t^n$ eine Potenzreihe über \mathbb{K} .

- a. Falls der *eigentliche* oder *uneigentliche* Grenzwert $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ existiert, so ist der Konvergenzradius von $\sum_{n=0}^{\infty} a_n \cdot t^n$ gegeben durch

$$r = \frac{1}{\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|}.$$

- b. Falls der *eigentliche* oder *uneigentliche* Grenzwert $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ existiert, so ist der Konvergenzradius von $\sum_{n=0}^{\infty} a_n \cdot t^n$ gegeben durch

$$r = \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}}.$$

Beweis:

- a. Es sei $r = \frac{1}{\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|}$ und $x \in \mathbb{K}$. Ist $|x| < r$, so ist die Reihe $\sum_{n=0}^{\infty} a_n \cdot x^n$ nach dem Quotientenkriterium in Korollar 12.24 absolut konvergent, da

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1} \cdot x^{n+1}}{a_n \cdot x^n} \right| = \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| \cdot |x| < 1.$$

Analog ist die Reihe divergent, wenn $|x| > r$.

- b. Der Beweis geht analog zu a., wobei wir das Quotientenkriterium in Korollar 12.24 durch das dortige Wurzelkriterium ersetzen.

□

Beispiel 12.35

a. Die *geometrische Reihe* $\sum_{n=0}^{\infty} t^n$ hat den Konvergenzradius $r = \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{1}} = 1$.

Damit wissen wir, daß die Reihe absolut konvergiert für $|x| < 1$ und daß sie divergiert für $|x| > 1$. Wir haben aber in Beispiel 11.31 schon gesehen, daß sie zudem auch für alle $|x| = 1$ divergiert, d.h. sie divergiert für alle Punkte im Rand des Konvergenzbereiches.

b. Die Potenzreihe $\sum_{n=1}^{\infty} \frac{t^n}{n}$ hat ebenfalls den Konvergenzradius

$$r = \frac{1}{\lim_{n \rightarrow \infty} \frac{n}{n+1}} = \frac{1}{\lim_{n \rightarrow \infty} \left(1 - \frac{1}{n+1}\right)} = 1.$$

Aber für $x = -1$ erhalten wir die alternierende harmonische Reihe $\sum_{n=1}^{\infty} \frac{x^n}{n} = \sum_{n=1}^{\infty} \frac{(-1)^n}{n}$, die konvergiert, so daß die Potenzreihe nicht für alle x im Rand des Konvergenzbereiches divergiert.

Satz 12.36 (Exponentialfunktion)

Die Potenzreihe $\sum_{n=0}^{\infty} \frac{t^n}{n!}$ über \mathbb{K} hat Konvergenzradius $r = \infty$.

Die dadurch definierte Abbildung

$$\exp : \mathbb{K} \longrightarrow \mathbb{K} : x \mapsto \exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

nennen wir die Exponentialfunktion. Sie genügt der Funktionalgleichung

$$\exp(x + y) = \exp(x) \cdot \exp(y)$$

für $x, y \in \mathbb{K}$.

Beweis: Der Konvergenzradius ergibt sich als

$$r = \frac{1}{\lim_{n \rightarrow \infty} \frac{n!}{(n+1)!}} = \frac{1}{\lim_{n \rightarrow \infty} \frac{1}{n+1}} = \frac{1}{0} = \infty.$$

Zudem folgt aus dem Cauchy-Produkt für Reihen 12.27 und dem Binomischen Lehrsatz 7.15

$$\begin{aligned} \exp(x) \cdot \exp(y) &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \cdot \sum_{n=0}^{\infty} \frac{y^n}{n!} \\ &\stackrel{12.27}{=} \sum_{n=0}^{\infty} \sum_{k=0}^n \frac{x^k}{k!} \cdot \frac{y^{n-k}}{(n-k)!} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^n \binom{n}{k} \cdot x^k \cdot y^{n-k} \\ &\stackrel{7.15}{=} \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} \\ &= \exp(x+y). \end{aligned}$$

□

Bemerkung 12.37

Nach Aufgabe 11.38 gilt $\exp(1) = e$ und mit Induktion folgt aus der Funktionalgleichung leicht, daß $\exp(n) = e^n$ für $n \in \mathbb{Z}$ und $\exp(\frac{1}{n}) = e^{\frac{1}{n}} = \sqrt[n]{e}$ für $n \geq 2$. Wir setzen für $x \in \mathbb{K}$ deshalb allgemein

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

so daß die neue Notation mit der üblichen Potenzschreibweise und mit der Notation in Satz 9.10 übereinstimmt, und das Potenzgesetz $e^{x+y} = e^x \cdot e^y$ gilt.

Satz 12.38 (Sinus und Cosinus)

- a. Die Potenzreihe $\sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n}}{(2n)!}$ über \mathbb{K} hat Konvergenzradius $r = \infty$.
Die dadurch definierte Abbildung

$$\cos : \mathbb{K} \longrightarrow \mathbb{K} : x \mapsto \cos(x) := \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n}}{(2n)!}$$

nennen wir den Cosinus.

- b. Die Potenzreihe $\sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!}$ über \mathbb{K} hat Konvergenzradius $r = \infty$.
Die dadurch definierte Abbildung

$$\sin : \mathbb{K} \longrightarrow \mathbb{K} : x \mapsto \sin(x) := \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!}$$

nennen wir den Sinus.

- c. Für $x \in \mathbb{K}$ gelten

$$\sin(-x) = -\sin(x)$$

und

$$\cos(-x) = \cos(x).$$

Wir nennen den Sinus eine ungerade Funktion und den Cosinus eine gerade.

- d. Für $x \in \mathbb{K}$ gilt

$$e^{ix} = \exp(i \cdot x) = \cos(x) + i \cdot \sin(x).$$

- e. Für $x \in \mathbb{K}$ gilt

$$\cos(x)^2 + \sin(x)^2 = 1.$$

- f. Für $x \in \mathbb{K}$ gilt

$$\cos(x) = \frac{1}{2} \cdot (e^{ix} + e^{-ix})$$

und

$$\sin(x) = \frac{1}{2i} \cdot (e^{ix} - e^{-ix}).$$

- g. Für zwei reelle Zahlen $x, y \in \mathbb{R}$ gelten die Additionstheoreme

$$\cos(x + y) = \cos(x) \cdot \cos(y) - \sin(x) \cdot \sin(y)$$

und

$$\sin(x + y) = \cos(x) \cdot \sin(y) + \sin(x) \cdot \cos(y).$$

h. Für eine reelle Zahl $x \in \mathbb{R}$ gilt $|e^{ix}| = 1$.

Beweis:

a. Ist $x \in \mathbb{K}$, so setzen wir

$$a_n := \begin{cases} (-1)^m \cdot \frac{x^{2m}}{(2m)!}, & \text{falls } n = 2m \text{ gerade,} \\ 0, & \text{falls } n \text{ ungerade.} \end{cases}$$

Dann ist $\cos(x) = \sum_{n=0}^{\infty} a_n$ und $|a_n| \leq \left|\frac{x^n}{n!}\right|$. Mithin ist $\exp(x)$ eine konvergente Majorante von $\cos(x)$, und $\cos(x)$ ist absolut konvergent für alle $x \in \mathbb{K}$. Insbesondere ist der Konvergenzradius also

$$r = \sup \{|x| \mid x \in \mathbb{K}\} = \sup(\mathbb{R}_{\geq 0}) = \infty.$$

b. Ist $x \in \mathbb{K}$, so setzen wir

$$a_n := \begin{cases} (-1)^m \cdot \frac{x^{2m+1}}{(2m+1)!}, & \text{falls } n = 2m+1 \text{ ungerade,} \\ 0, & \text{falls } n \text{ gerade.} \end{cases}$$

Dann ist $\sin(x) = \sum_{n=0}^{\infty} a_n$ und $|a_n| \leq \left|\frac{x^n}{n!}\right|$. Mithin ist $\exp(x)$ eine konvergente Majorante von $\sin(x)$, und $\sin(x)$ ist absolut konvergent für alle $x \in \mathbb{K}$. Insbesondere ist der Konvergenzradius also

$$r = \sup \{|x| \mid x \in \mathbb{K}\} = \sup(\mathbb{R}_{\geq 0}) = \infty.$$

c. Für $x \in \mathbb{K}$ gilt

$$\sin(-x) = \sum_{n=0}^{\infty} (-1)^n \frac{(-x)^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} (-1)^n \frac{(-1) \cdot (x^{2n+1})}{(2n+1)!} = -\sin(x)$$

und

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

d. Wir beachten, daß für die imaginäre Einheit i stets $i^{2n} = (-1)^n$ und $i^{2n+1} = (-1)^n \cdot i$ gilt. Dadurch erhalten wir

$$\begin{aligned} \cos(x) + i \cdot \sin(x) &= \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n}}{(2n)!} + i \cdot \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} i^{2n} \cdot \frac{x^{2n}}{(2n)!} + \sum_{n=0}^{\infty} i^{2n+1} \cdot \frac{x^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} \frac{(i \cdot x)^n}{n!} = \exp(i \cdot x) \end{aligned}$$

unter Berücksichtigung der Grenzwertsätze für konvergente Reihen 12.4 und des Umordnungssatzes für absolut konvergente Reihen 12.18.

e. Für $x \in \mathbb{K}$ gilt

$$\begin{aligned} \cos(x)^2 + \sin(x)^2 &= (\cos(x) + i \cdot \sin(x)) \cdot (\cos(x) - i \cdot \sin(x)) \\ &\stackrel{c.}{=} (\cos(x) + i \cdot \sin(x)) \cdot (\cos(-x) + i \cdot \sin(-x)) \\ &\stackrel{d.}{=} \exp(ix) \cdot \exp(-ix) \\ &\stackrel{12.36}{=} \exp(ix - ix) = \exp(0) = 1. \end{aligned}$$

f. Für $x \in \mathbb{K}$ gilt

$$e^{ix} + e^{-ix} \stackrel{d.}{=} \cos(x) + i \cdot \sin(x) + \cos(-x) + i \cdot \sin(-x) \stackrel{c.}{=} 2 \cdot \cos(x)$$

und

$$e^{ix} - e^{-ix} \stackrel{d.}{=} \cos(x) + i \cdot \sin(x) - \cos(-x) - i \cdot \sin(-x) \stackrel{c.}{=} 2 \cdot i \cdot \sin(x).$$

g. Werten wir den Sinus oder den Cosinus an einer reellen Zahl aus, so erhalten wir eine reelle Zahl. Für $x, y \in \mathbb{R}$ betrachten wir nun die komplexe Zahl

$$\begin{aligned} \cos(x + y) + i \cdot \sin(x + y) &\stackrel{d.}{=} \exp(i \cdot (x + y)) \stackrel{12.36}{=} \exp(ix) \cdot \exp(iy) \\ &\stackrel{d.}{=} (\cos(x) + i \cdot \sin(x)) \cdot (\cos(y) + i \cdot \sin(y)) \\ &= (\cos(x) \cdot \cos(y) - \sin(x) \cdot \sin(y)) \\ &\quad + i \cdot (\cos(x) \cdot \sin(y) + \sin(x) \cdot \cos(y)). \end{aligned}$$

Durch einen Vergleich des Realteils bzw. des Imaginärteils der beiden Seiten der Gleichung, erhalten wir die gewünschten Formeln.

h. Für $x \in \mathbb{R}$ gilt $|\exp(ix)| = |\cos(x) + i \sin(x)| = \sqrt{\cos(x)^2 + \sin(x)^2} = 1$.

□

Bemerkung 12.39 (Alternativer Beweis für $\cos(x)^2 + \sin(x)^2 = 1$)

Wir wollen zunächst $\cos(x)^2$ mit Hilfe des Cauchy-Produktes 12.27 ausrechnen.

$$\begin{aligned} \cos(x)^2 &= \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n}}{(2n)!} \cdot \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n}}{(2n)!} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^n (-1)^k \cdot \frac{x^{2k}}{(2k)!} \cdot (-1)^{n-k} \cdot \frac{x^{2n-2k}}{(2n-2k)!} \\ &= \sum_{n=0}^{\infty} \sum_{k=0}^n (-1)^n \cdot \frac{x^{2n}}{(2k)! \cdot (2n-2k)!} \\ &= \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n}}{(2n)!} \cdot \sum_{k=0}^n \binom{2n}{2k} \\ &= 1 + \sum_{n=0}^{\infty} (-1)^{n+1} \cdot \frac{x^{2n+2}}{(2n+2)!} \cdot \sum_{k=0}^{n+1} \binom{2n+2}{2k}, \end{aligned}$$

wobei die letzte Gleichheit durch eine Indexverschiebung zustande kommt.

Dann wenden wir uns $\sin(x)^2$ zu und erhalten analog

$$\begin{aligned}
 \sin(x)^2 &= \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!} \cdot \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!} \\
 &= \sum_{n=0}^{\infty} \sum_{k=0}^n (-1)^k \cdot \frac{x^{2k+1}}{(2k+1)!} \cdot (-1)^{n-k} \cdot \frac{x^{2n-2k+1}}{(2n-2k+1)!} \\
 &= \sum_{n=0}^{\infty} \sum_{k=0}^n (-1)^n \cdot \frac{x^{2n+2}}{(2k+1)! \cdot (2n-2k+1)!} \\
 &= \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n+2}}{(2n+2)!} \cdot \sum_{k=0}^n \binom{2n+2}{2k+1}.
 \end{aligned}$$

Addieren wir die Gleichungen, so erhalten wir mit dem Binomischen Lehrsatz 7.15

$$\begin{aligned}
 \cos(x)^2 + \sin(x)^2 &= 1 + \sum_{n=0}^{\infty} (-1)^{n+1} \cdot \frac{x^{2n+2}}{(2n+2)!} \cdot \left(\sum_{k=0}^{n+1} \binom{2n+2}{2k} - \sum_{k=0}^n \binom{2n+2}{2k+1} \right) \\
 &= 1 + \sum_{n=0}^{\infty} (-1)^{n+1} \cdot \frac{x^{2n+2}}{(2n+2)!} \cdot \left(\sum_{l=0}^{2n+2} \binom{2n+2}{l} \cdot (-1)^l \cdot 1^{2n-l} \right) \\
 &\stackrel{7.15}{=} 1 + \sum_{n=0}^{\infty} (-1)^{n+1} \cdot \frac{x^{2n+2}}{(2n+2)!} \cdot (1-1)^{2n} \\
 &= 1
 \end{aligned}$$

unter Berücksichtigung der Grenzwertsätze für Reihen 12.4 sowie des Umordnungsatzes für absolut konvergente Reihen 12.18.

Bemerkung 12.40

Die Additionstheoreme

$$\cos(x+y) = \cos(x) \cdot \cos(y) - \sin(x) \cdot \sin(y)$$

und

$$\sin(x+y) = \cos(x) \cdot \sin(y) + \sin(x) \cdot \cos(y).$$

gelten in der Tat nicht nur für reelle Zahlen x und y , sondern auch für beliebige komplexe Zahlen $x, y \in \mathbb{C}$. Allerdings funktioniert dann der oben geführte Beweis nicht, da dann die Aufteilung in den Real- und Imaginärteil in der angegebenen Form nicht möglich ist. Stattdessen kann man die Formeln direkt aus der Definition von Sinus und Cosinus mittels Potenzreihen herleiten, wie wir das in Bemerkung 12.39 für die Gleichung $\cos(x)^2 + \sin(x)^2 = 1$ getan haben. Das zu tun, überlassen wir dem Leser als Übungsaufgabe.

Bemerkung 12.41 (Potenzreihen mit beliebigem Entwicklungspunkt a)

Für eine Potenzreihe $\sum_{n=0}^{\infty} a_n \cdot (t-a)^n$ über \mathbb{K} mit Entwicklungspunkt $a \in \mathbb{K}$ ist

der Konvergenzradius immer noch definiert als

$$r := \sup \left\{ |y| \mid y \in \mathbb{K}, \sum_{n=0}^{\infty} a_n \cdot y^n \text{ ist konvergent} \right\} \in \mathbb{R}_{\geq 0} \cup \{\infty\},$$

und wir erhalten dann aus Satz 12.32

- $\forall x \in \mathbb{K}$ mit $|x - a| < r$ ist $\sum_{n=0}^{\infty} a_n \cdot (x - a)^n$ absolut konvergent,
- $\forall x \in \mathbb{K}$ mit $|x - a| > r$ ist $\sum_{n=0}^{\infty} a_n \cdot (x - a)^n$ divergent,

und Satz 12.34 impliziert, daß sich der Konvergenzradius ggf. berechnen läßt als

$$r = \frac{1}{\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|}$$

bzw. als

$$r = \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}}.$$

Aufgaben

Aufgabe 12.42

Untersuche die folgenden Reihen auf Konvergenz. Die Berechnung der Grenzwerte im Falle der Konvergenz ist nicht erforderlich.

- a. $\sum_{n=1}^{\infty} \frac{1-n^4}{100n^4}$.
- b. $\sum_{n=1}^{\infty} \frac{n!}{n^{n+1}}$.
- c. $\sum_{n=1}^{\infty} \frac{(2n)^n}{(-3)^{n+1}}$.

Aufgabe 12.43

Es sei $q \in \mathbb{K}$ mit $|q| < 1$.

- a. Berechne das Cauchy-Produkt $\left(\sum_{n=0}^{\infty} q^n \right)^2$.
- b. Berechne den Wert der Reihe $\sum_{n=0}^{\infty} nq^n$.

Aufgabe 12.44

Seien $\sum_{n=0}^{\infty} a_n t^n$ und $\sum_{n=1}^{\infty} n a_n t^{n-1}$ Potenzreihen in \mathbb{K} . Zeige die folgenden Aussagen:

- a. Konvergiert $\sum_{n=0}^{\infty} a_n y^n$ für ein $y \in \mathbb{K}$, so konvergiert $\sum_{n=1}^{\infty} n a_n x^{n-1}$ absolut für alle $x \in \mathbb{K}$ mit $|x| < |y|$.
- b. Die gegebenen Potenzreihen haben denselben Konvergenzradius.
- c. Konvergieren die Potenzreihen vielleicht sogar stets für dieselben $x \in \mathbb{K}$?

HINWEIS: Schaut euch hilfestellend den Beweis von Lemma 12.29 an und verwendet Aufgabe 12.43.

Aufgabe 12.45 (Dezimalzahldarstellung)

Wir sind es gewohnt, reelle Zahlen als „Dezimalzahlen mit eventuell unendlich vielen Nachkommastellen“, wie z.B. $3,1415926\dots$, zu schreiben. Diese Aufgabe soll zeigen, warum das eigentlich möglich ist und welche Eigenschaften diese Dezimaldarstellung hat. Der Einfachheit halber beschränken wir uns dabei auf positive Zahlen ohne Stellen vor dem Komma.

Seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ Folgen mit $a_n, b_n \in \{0, 1, 2, \dots, 9\}$. Beweise die folgenden Aussagen.

- a. Die Reihe

$$\sum_{n=1}^{\infty} a_n \cdot 10^{-n}$$

ist konvergent und stellt somit eine reelle Zahl dar. Man schreibt diese Zahl dann in Dezimaldarstellung durch Hintereinanderschreiben der Ziffern als $0, a_1 a_2 a_3 \dots$

- b. Jede reelle Zahl im Intervall $[0, 1)$ besitzt eine Dezimaldarstellung wie in Aufgabenteil a..
- c. Ist $N \geq 2$ eine feste natürliche Zahl mit $a_j = b_j$ für $1 \leq j \leq N - 1$, $a_N < b_N$ und $0, a_1 a_2 a_3 \dots = 0, b_1 b_2 b_3 \dots$, so gilt $a_N + 1 = b_N$ und $a_j = 9$, $b_j = 0$ für alle $j > N$.

Aufgabe 12.46

Bestimme die Konvergenzradien folgender Potenzreihen.

a. $\sum_{n=0}^{\infty} n^k \cdot t^n$ für $k \in \mathbb{N}$.

b. $\sum_{n=1}^{\infty} \frac{n!}{(2n)^n} \cdot t^n$.

Aufgabe 12.47

Beweise für $x, y \in \mathbb{K}$ die Gleichung $\sin(x + y) = \sin(x) \cdot \cos(y) + \cos(x) \cdot \sin(y)$.

§ 13 Grenzwerte von Funktionen

Wir werden uns in den folgenden Paragraphen im wesentlichen dem Studium von Abbildungen

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

widmen und nennen diese eher *Funktionen* als Abbildungen. Wir könnten dabei viele Begriffe auch gleich wieder für den Körper der komplexen Zahlen \mathbb{C} statt \mathbb{R} einführen und untersuchen, wollen dies aber zurück stellen, um näher an dem Vorwissen aus der Schulzeit zu bleiben.

A) Häufungspunkte von Teilmengen von \mathbb{R}

Definition 13.1 (Häufungspunkte)

Es sei $U \subseteq \mathbb{R}$ eine Teilmenge von \mathbb{R} und $a \in \mathbb{R}$. Wir nennen a einen *Häufungspunkt* von U , wenn

$$\forall \varepsilon > 0 \quad \exists x \in U \setminus \{a\} : 0 < |x - a| < \varepsilon.$$

Man beachte, daß a kein Element von U sein muß.

Bemerkung 13.2 (ε -Umgebung)

Für $\varepsilon > 0$ und $a \in \mathbb{R}$ nennen wir das Intervall

$$U_\varepsilon(a) := (a - \varepsilon, a + \varepsilon) = \{x \in \mathbb{R} \mid |x - a| < \varepsilon\}$$

die ε -Umgebung von a .

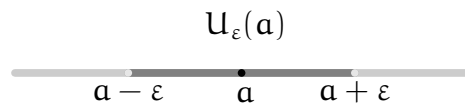


ABBILDUNG 4. Die ε -Umgebung $U_\varepsilon(a)$ von a .

Mit dieser Sprechweise gilt also:

Genau dann ist a ein Häufungspunkt von U , wenn jede ε -Umgebung von a einen von a verschiedenen Punkt aus U enthält.

Beispiel 13.3

Jede reelle Zahl ist Häufungspunkt von \mathbb{Q} .

Dazu seien $a \in \mathbb{R}$ und $\varepsilon > 0$ gegeben. Wir wenden Satz 9.6 an und finden eine rationale Zahl x im Intervall $(a, a + \varepsilon)$, d.h. $|x - a| < \varepsilon$ und $x \neq a$.

Proposition 13.4 (Folgenkriterium für Häufungspunkte)

Ein $a \in \mathbb{R}$ ist genau dann Häufungspunkt von $U \subseteq \mathbb{R}$, wenn es eine Folge $(a_n)_{n \in \mathbb{N}}$ mit $a_n \in U \setminus \{a\}$ und $\lim_{n \rightarrow \infty} a_n = a$ gibt, d.h. a ist Grenzwert einer Folge in $U \setminus \{a\}$.

Beweis: Ist \mathbf{a} ein Häufungspunkt von \mathbf{U} und $n \in \mathbb{N}$, so gibt es zu $\varepsilon = \frac{1}{n+1} > 0$ ein $x \in \mathbf{U}$ mit $0 < |x - \mathbf{a}| < \varepsilon = \frac{1}{n+1}$. Wir wählen \mathbf{a}_n als dieses x . Dann konvergiert die Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ offenbar gegen \mathbf{a} nach dem Einschließungssatz, da

$$0 < |\mathbf{a}_n - \mathbf{a}| < \frac{1}{n+1} \longrightarrow 0.$$

Gibt es umgekehrt eine Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in $\mathbf{U} \setminus \{\mathbf{a}\}$, die gegen \mathbf{a} konvergiert, so gibt es für jedes $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$ mit $0 < |\mathbf{a}_n - \mathbf{a}| < \varepsilon$ für alle $n \geq n_\varepsilon$. Setzen wir nun $x = \mathbf{a}_{n_\varepsilon} \in \mathbf{U}$, so folgt $0 < |x - \mathbf{a}| < \varepsilon$ und somit ist \mathbf{a} ein Häufungspunkt von \mathbf{U} . \square

Das folgende Beispiel zeigt, welche Art von Häufungspunkten einer Menge \mathbf{U} , die nicht bereits in \mathbf{U} liegen, wir typischerweise erwarten.

Beispiel 13.5

Sind $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ mit $\mathbf{a} < \mathbf{b}$, so enthält $[\mathbf{a}, \mathbf{b}]$ genau die Häufungspunkte von (\mathbf{a}, \mathbf{b}) , d.h. zu den Punkten im Intervall kommen noch die Randpunkte hinzu.

Die analogen Aussagen für halboffene, abgeschlossene und uneigentliche Intervalle gelten ebenfalls und mit analogem Beweis.

Beweis: Ist $\mathbf{c} \in \mathbb{R}$ ein Häufungspunkt von (\mathbf{a}, \mathbf{b}) , so gibt es nach dem Folgenkriterium 13.4 eine Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in $(\mathbf{a}, \mathbf{b}) \subset [\mathbf{a}, \mathbf{b}]$, die gegen \mathbf{c} konvergiert, und nach Satz 11.28 ist dann $\mathbf{c} \in [\mathbf{a}, \mathbf{b}]$.

Ist $\mathbf{a} \leq \mathbf{c} < \mathbf{b}$, so gilt für $n \geq 1$

$$(\mathbf{a}, \mathbf{b}) \ni \mathbf{c} + \frac{\mathbf{b} - \mathbf{c}}{2 \cdot n} \longrightarrow \mathbf{c},$$

also ist \mathbf{c} ein Häufungspunkt von (\mathbf{a}, \mathbf{b}) . Analog gilt $(\mathbf{a}, \mathbf{b}) \ni \mathbf{b} - \frac{\mathbf{b} - \mathbf{a}}{2 \cdot n} \longrightarrow \mathbf{b}$, so daß auch \mathbf{b} ein Häufungspunkt von (\mathbf{a}, \mathbf{b}) ist. \square

B) Grenzwerte von Funktionen

Definition 13.6 (ε - δ -Kriterium für Grenzwerte von Funktionen)

Sei $\mathbf{U} \subseteq \mathbb{R}$, $f : \mathbf{U} \longrightarrow \mathbb{R}$ eine Funktion und \mathbf{a} ein Häufungspunkt von \mathbf{U} .

Wir nennen $\mathbf{y} \in \mathbb{R}$ den *Grenzwert* von f in \mathbf{a} , falls

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : \forall x \in \mathbf{U} \text{ mit } 0 < |x - \mathbf{a}| < \delta_\varepsilon \text{ gilt } |f(x) - \mathbf{y}| < \varepsilon.$$

Wir schreiben dann

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

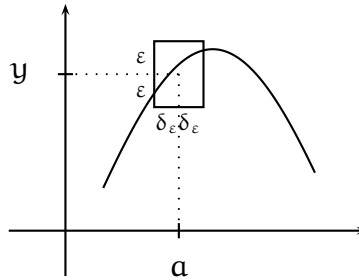
oder “ $f(x) \longrightarrow \mathbf{y}$ für $x \rightarrow \mathbf{a}$ ” und sagen, $f(x)$ *konvergiert* gegen \mathbf{y} für x gegen \mathbf{a} .

Proposition 13.7 (Folgenkriterium für Grenzwerte von Funktionen)

Es sei $\mathbf{U} \subseteq \mathbb{R}$, $f : \mathbf{U} \longrightarrow \mathbb{R}$ eine Funktion und \mathbf{a} ein Häufungspunkt von \mathbf{U} .

Dann sind die beiden folgenden Aussagen gleichwertig:

a. $\lim_{x \rightarrow \mathbf{a}} f(x) = \mathbf{y}.$

ABBILDUNG 5. ε - δ -Kriterium für Grenzwerte

b. $\forall (a_n)_{n \in \mathbb{N}}$ mit $a_n \in U \setminus \{a\}$ und $\lim_{n \rightarrow \infty} a_n = a$ gilt $\lim_{n \rightarrow \infty} f(a_n) = y$.

Beweis: a. \implies b.: Es sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in $U \setminus \{a\}$ mit $\lim_{n \rightarrow \infty} a_n = a$. Wir müssen $\lim_{n \rightarrow \infty} f(a_n) = y$ zeigen. Dazu sei $\varepsilon > 0$ gegeben.

Wegen $\lim_{x \rightarrow a} f(x) = y$ gibt es ein $\delta_\varepsilon > 0$, so daß aus $x \in U$ mit $0 < |x - a| < \delta_\varepsilon$ auch $|f(x) - y| < \varepsilon$ folgt.

Wegen $\lim_{n \rightarrow \infty} a_n = a$ gibt es zu δ_ε nun ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $n \geq n_\varepsilon$ auch $|a_n - a| < \delta_\varepsilon$ gilt.

Sei nun $n \geq n_\varepsilon$ dann erfüllt $a_n \in U$ die Bedingung $0 < |a_n - a| < \delta_\varepsilon$ und somit ist auch $|f(a_n) - y| < \varepsilon$. Damit ist $f(a_n) \rightarrow y$ gezeigt.

b. \implies a.: Wir nehmen an, y wäre nicht der Grenzwert von f in a . Dann gilt:

$\exists \varepsilon > 0 : \forall \delta_\varepsilon > 0 \exists x_{\delta_\varepsilon} \in U$ mit $0 < |x_{\delta_\varepsilon} - a| < \delta_\varepsilon$, aber $|f(x_{\delta_\varepsilon}) - y| \geq \varepsilon$.

Für $n \geq 1$ und $\delta_\varepsilon = \frac{1}{n}$ setzen wir $a_n := x_{\delta_\varepsilon} = x_{\frac{1}{n}} \in U \setminus \{a\}$. Dann gilt

$$0 < |a_n - a| < \frac{1}{n} \rightarrow 0,$$

so daß $a_n \rightarrow a$, und zugleich gilt

$$|f(a_n) - y| \geq \varepsilon$$

für alle $n \in \mathbb{N}$. Dies ist ein Widerspruch dazu, daß $f(a_n)$ gegen y konvergieren muß.

□

Beispiel 13.8

a. Betrachte $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^2$ und $a = 3$. Für eine Folge $(a_n)_{n \in \mathbb{N}}$ mit $a_n \rightarrow 3$ gilt dann wegen der Grenzwertsätze für Folgen 11.15

$$f(a_n) = a_n^2 = a_n \cdot a_n \rightarrow 3 \cdot 3 = 9.$$

Mithin ist 9 der Grenzwert von f in 3, d.h.

$$\lim_{x \rightarrow 3} x^2 = 9 = f(3).$$

b. Betrachte die Funktion

$$f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \begin{cases} 1, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0 \end{cases}$$

und $\mathbf{a} = 0$. Ist nun $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge mit $\mathbf{a}_n \rightarrow 0$ und $\mathbf{a}_n \neq 0$, dann gilt

$$f(\mathbf{a}_n) = 1 \rightarrow 1.$$

Mithin ist 1 der Grenzwert von f in 0, d.h.

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

c. Betrachte die Funktion

$$f : \mathbb{R} \setminus \{1\} \rightarrow \mathbb{R} : x \mapsto \frac{x^2 - 1}{x - 1}$$

und $\mathbf{a} = 1$. Da es in $\mathbb{R} \setminus \{1\}$ offenbar eine Folge gibt, die gegen 1 konvergiert, ist \mathbf{a} ein Häufungspunkt von $\mathbb{R} \setminus \{1\}$. Sei nun $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in $\mathbb{R} \setminus \{1\}$ mit $\mathbf{a}_n \rightarrow 1$, so gilt

$$f(\mathbf{a}_n) = \frac{\mathbf{a}_n^2 - 1}{\mathbf{a}_n - 1} = \mathbf{a}_n + 1 \rightarrow 2.$$

Mithin ist 2 der Grenzwert von f in 1, d.h.

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

Man beachte, daß in diesem Fall $\mathbf{a} = 1$ gar nicht im Definitionsbereich von f liegt.

d. Betrachte die Funktion

$$f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x < 0, \\ 1, & \text{falls } x > 0 \end{cases}$$

und $\mathbf{a} = 0$. Dann gilt für $\mathbf{a}_n := -\frac{1}{n} \rightarrow 0$ und $f(\mathbf{a}_n) = 0 \rightarrow 0$ sowie $\mathbf{b}_n := \frac{1}{n} \rightarrow 0$ und $f(\mathbf{b}_n) = 1 \rightarrow 1$. Mithin existiert der Grenzwert von f in $\mathbf{a} = 0$ nicht.

!!! Warnung !!!

Unsere Definition des Begriffes Grenzwert stimmt **nicht** mit der Definition in den Vorlesungsskripten von Andreas Gathmann oder Wolfram Decker überein! Wenn f im Punkt \mathbf{a} definiert ist, muß bei uns **nicht** notwendig $\lim_{x \rightarrow \mathbf{a}} f(x) = f(\mathbf{a})$ gelten (siehe Beispiel 13.8 b.), was nach deren Definition gelten muß!

Definition 13.9

Für zwei Funktionen $f : \mathbf{U} \rightarrow \mathbb{R}$ und $g : \mathbf{V} \rightarrow \mathbb{R}$ sowie $c \in \mathbb{R}$ definieren wir

$$c \cdot f : \mathbf{U} \rightarrow \mathbb{R} : x \mapsto c \cdot f(x),$$

$$f + g : \mathbf{U} \cap \mathbf{V} \rightarrow \mathbb{R} : x \mapsto f(x) + g(x),$$

$$f - g : \mathbf{U} \cap \mathbf{V} \rightarrow \mathbb{R} : x \mapsto f(x) - g(x)$$

und

$$f \cdot g : \mathbf{U} \cap \mathbf{V} \rightarrow \mathbb{R} : x \mapsto f(x) \cdot g(x).$$

Falls zudem $g(x) \neq 0$ für $x \in U \cap V$, so definieren wir

$$\frac{f}{g} : U \cap V \longrightarrow \mathbb{R} : x \mapsto \frac{f(x)}{g(x)}.$$

Proposition 13.10 (Grenzwertsätze für Funktionen)

Es seien $f : U \longrightarrow \mathbb{R}$ und $g : U \longrightarrow \mathbb{R}$ zwei Funktionen, a ein Häufungspunkt von U und $c \in \mathbb{R}$.

- a. Der Grenzwert von f in a ist eindeutig bestimmt, d.h. falls $\lim_{x \rightarrow a} f(x) = y$ und $\lim_{x \rightarrow a} f(x) = z$, so ist $y = z$.
- b. Wenn $\lim_{x \rightarrow a} f(x)$ und $\lim_{x \rightarrow a} g(x)$ existieren, so gelten:
 - (i) $\lim_{x \rightarrow a} (c \cdot f)(x) = c \cdot \lim_{x \rightarrow a} f(x)$.
 - (ii) $\lim_{x \rightarrow a} (f + g)(x) = \lim_{x \rightarrow a} f(x) + \lim_{x \rightarrow a} g(x)$.
 - (iii) $\lim_{x \rightarrow a} (f - g)(x) = \lim_{x \rightarrow a} f(x) - \lim_{x \rightarrow a} g(x)$.
 - (iv) $\lim_{x \rightarrow a} (f \cdot g)(x) = \lim_{x \rightarrow a} f(x) \cdot \lim_{x \rightarrow a} g(x)$.
- c. Falls zudem $\lim_{x \rightarrow a} f(x) \neq 0$, so ist a ein Häufungspunkt der Menge $V = \{x \in U \mid f(x) \neq 0\}$ und es gilt

$$\lim_{x \rightarrow a} \frac{1}{f}(x) = \frac{1}{\lim_{x \rightarrow a} f(x)}.$$

Beweis: a. Dies folgt aus dem Folgenkriterium für Grenzwerte von Funktionen 13.7 und der Eindeutigkeit des Grenzwertes bei Folgen 11.8. Genauer, da a ein Häufungspunkt von U ist, gibt es nach Proposition 13.4 eine Folge $(a_n)_{n \in \mathbb{N}}$ in $U \setminus \{a\}$ mit $a_n \longrightarrow a$, und mit den eben erwähnten Sätzen folgt dann

$$y = \lim_{n \rightarrow \infty} f(a_n) = z.$$

- b. Analog folgen die Aussagen aus dem Folgenkriterium für Grenzwerte von Funktionen 13.7 und den Grenzwertsätzen für Folgen 11.15 unter Berücksichtigung von Proposition 13.4.
- c. Nach Proposition 13.4 gibt es eine Folge $(a_n)_{n \in \mathbb{N}}$ in $U \setminus \{a\}$, die gegen a konvergiert, und nach dem Folgenkriterium 13.7 gilt dann

$$f(a_n) \longrightarrow \lim_{x \rightarrow a} f(x) =: y.$$

Wegen $y \neq 0$ gibt es wegen der Grenzwertsätze für Folgen 11.15 ein n_0 , so daß $f(a_n) \neq 0$ für alle $n \geq n_0$, so daß $(a_n)_{n \geq n_0}$ eine Folge in V ist mit $a_n \longrightarrow a$. Nach Proposition 13.4 ist dann a ein Häufungspunkt von V . Die Aussage zum Grenzwert folgt dann wieder aus den Grenzwertsätzen für Folgen 11.15 und dem Folgenkriterium 13.7.

□

Definition 13.11

Ist t eine Veränderliche und sind $a_0, \dots, a_n \in \mathbb{R}$, so nennen wir einen Ausdruck der Form

$$\sum_{k=0}^n a_k \cdot t^k = a_n \cdot t^n + a_{n-1} \cdot t^{n-1} + \dots + a_1 \cdot t + a_0$$

ein *Polynom* in der Veränderlichen t mit Koeffizienten in \mathbb{R} . Ist $a_n \neq 0$, so heißt

$$\deg \left(\sum_{k=0}^n a_k \cdot t^k \right) := n$$

der *Grad* des Polynoms, und wir setzen zudem $\deg(0) := -\infty$. Mit

$$\mathbb{R}[t] := \left\{ \sum_{k=0}^n a_k \cdot t^k \mid n \in \mathbb{N}, a_0, \dots, a_n \in \mathbb{R} \right\}$$

bezeichnen wir die Menge aller Polynome in der Veränderlichen t mit Koeffizienten in \mathbb{R} , so daß der Grad eine Abbildung $\deg: \mathbb{R}[t] \rightarrow \mathbb{N} \cup \{-\infty\}$ ist.

Für ein Polynom $f = \sum_{k=0}^n a_k \cdot t^k \in \mathbb{R}[t]$ und ein $x \in \mathbb{R}$ setzen wir

$$f(x) := \sum_{k=0}^n a_k \cdot x^k.$$

Sind $f, g \in \mathbb{R}[t]$ zwei Polynome, $g \neq 0$ nicht das Nullpolynom, so nennen wir die Funktion

$$f: \mathbb{R} \rightarrow \mathbb{R}: x \mapsto f(x)$$

eine *Polynomfunktion* und die Funktion

$$\frac{f}{g}: \mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) = 0\} \rightarrow \mathbb{R}: x \mapsto \frac{f(x)}{g(x)}$$

nennen wir eine *rationale Funktion*.

Ist $h: \mathbb{R} \rightarrow \mathbb{R}$ irgendeine Funktion, so nennen wir eine reelle Zahl $x \in \mathbb{R}$ mit $h(x) = 0$ eine *Nullstelle* von h .

Bemerkung 13.12

Man zeigt in der Vorlesung Algebraische Strukturen, daß die Menge der *Nullstellen* von $0 \neq g \in \mathbb{R}[t]$ eine endliche Menge ist. Genauer zeigt man:

$$|\{x \in \mathbb{R} \mid g(x) = 0\}| \leq \deg(g) < \infty.$$

Beispiel 13.13

a. Ist $f = \sum_{k=0}^n a_k \cdot t^k$ ein Polynom und $a \in \mathbb{R}$, so gilt

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

Dies folgt aus den Grenzwertsätzen für Funktionen 13.10, da offenbar $\lim_{x \rightarrow a} \text{id}(x) = \lim_{x \rightarrow a} x = a$ und f sich als endliche Summe von Produkten dieser Funktion mit sich selbst und mit Konstanten schreiben läßt.

b. Für jede rationale Funktion

$$\frac{f}{g} : \mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) = 0\} \longrightarrow \mathbb{R} : x \mapsto \frac{f(x)}{g(x)}$$

und jedes $a \in \mathbb{R}$ mit $g(a) \neq 0$ folgt dann aus Teil a. und Satz 13.10 c., daß a ein Häufungspunkt von $\mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) \neq 0\}$ ist und daß

$$\lim_{x \rightarrow a} \frac{f}{g}(x) = \frac{f(a)}{g(a)} = \frac{f}{g}(a).$$

C) Uneigentliche Grenzwerte

Definition 13.14 (Grenzwerte für $x \rightarrow \pm\infty$)

Es sei $U \subseteq \mathbb{R}$, $f : U \rightarrow \mathbb{R}$ und $y \in \mathbb{R}$.

a. Wir nennen U *nach oben unbeschränkt* bzw. *nach unten unbeschränkt*, wenn die Menge $U \cap [0, \infty)$ bzw. $U \cap (-\infty, 0]$ nicht beschränkt ist.

b. Ist U nach oben unbeschränkt, so nennen wir y den *Grenzwert* von f in ∞ , wenn

$$\forall \varepsilon > 0 \exists s_\varepsilon > 0 : \forall x \in U \text{ mit } x > s_\varepsilon \text{ gilt } |f(x) - y| < \varepsilon.$$

Wir schreiben dann $\lim_{x \rightarrow \infty} f(x) = y$.

c. Ist U nach unten unbeschränkt, so nennen wir y den *Grenzwert* von f in $-\infty$, wenn

$$\forall \varepsilon > 0 \exists s_\varepsilon < 0 : \forall x \in U \text{ mit } x < s_\varepsilon \text{ gilt } |f(x) - y| < \varepsilon.$$

Wir schreiben dann $\lim_{x \rightarrow -\infty} f(x) = y$.

Bemerkung 13.15 (Folgenkriterium und Grenzwertsätze für Grenzwerte in $\pm\infty$)

Das Folgenkriterium für Grenzwerte von Funktionen gilt analog auch für die Grenzwerte in $\pm\infty$. D.h.

$$\lim_{x \rightarrow \infty} f(x) = y \iff \forall (a_n)_{n \in \mathbb{N}} \text{ mit } a_n \in U \text{ und } a_n \rightarrow \infty \text{ gilt } f(a_n) \rightarrow y$$

und

$$\lim_{x \rightarrow -\infty} f(x) = y \iff \forall (a_n)_{n \in \mathbb{N}} \text{ mit } a_n \in U \text{ und } a_n \rightarrow -\infty \text{ gilt } f(a_n) \rightarrow y.$$

Zudem gelten auch die Grenzwertsätze für Funktionen 13.10 für Grenzwerte in $\pm\infty$.

Definition 13.16 (Uneigentliche Grenzwerte)

Sei $U \subseteq \mathbb{R}$, $f : U \rightarrow \mathbb{R}$ und a ein Häufungspunkt von U .

a. Wir nennen ∞ den *uneigentlichen Grenzwert* von f in a , wenn

$$\forall s > 0 \exists \delta_s > 0 : \forall x \in U \text{ mit } 0 < |x - a| < \delta_s \text{ gilt } f(x) > s.$$

Wir schreiben dann $\lim_{x \rightarrow a} f(x) = \infty$.

- b. Wir nennen $-\infty$ den *uneigentlichen Grenzwert* von f in \mathbf{a} , wenn

$$\forall s < 0 \exists \delta_s > 0 : \forall x \in \mathbf{U} \text{ mit } 0 < |x - \mathbf{a}| < \delta_s \text{ gilt } f(x) < s.$$

Wir schreiben dann $\lim_{x \rightarrow \mathbf{a}} f(x) = -\infty$.

- c. Ist \mathbf{U} nach oben unbeschränkt, so nennen wir ∞ den *uneigentlichen Grenzwert* von f in ∞ , wenn

$$\forall s > 0 \exists t > 0 : \forall x \in \mathbf{U} \text{ mit } x > t \text{ gilt } f(x) > s.$$

Wir schreiben dann $\lim_{x \rightarrow \infty} f(x) = \infty$.

- d. Ist \mathbf{U} nach unten unbeschränkt, so nennen wir $-\infty$ den *uneigentlichen Grenzwert* von f in ∞ , wenn

$$\forall s < 0 \exists t > 0 : \forall x \in \mathbf{U} \text{ mit } x > t \text{ gilt } f(x) < s.$$

Wir schreiben dann $\lim_{x \rightarrow \infty} f(x) = -\infty$.

- e. Ist \mathbf{U} nach unten unbeschränkt, so nennen wir ∞ den *uneigentlichen Grenzwert* von f in $-\infty$, wenn

$$\forall s > 0 \exists t < 0 : \forall x \in \mathbf{U} \text{ mit } x < t \text{ gilt } f(x) > s.$$

Wir schreiben dann $\lim_{x \rightarrow -\infty} f(x) = \infty$.

- f. Ist \mathbf{U} nach unten unbeschränkt, so nennen wir $-\infty$ den *uneigentlichen Grenzwert* von f in $-\infty$, wenn

$$\forall s < 0 \exists t < 0 : \forall x \in \mathbf{U} \text{ mit } x < t \text{ gilt } f(x) < s.$$

Wir schreiben dann $\lim_{x \rightarrow -\infty} f(x) = -\infty$.

Bemerkung 13.17 (Folgenkriterium und Grenzwertsätze für uneigentliche GWe)
Auch für uneigentliche Grenzwerte gelten naheliegende Folgenkriterien:

- a. $\lim_{x \rightarrow \mathbf{a}} f(x) = \infty \iff \forall (\mathbf{a}_n)_{n \in \mathbb{N}}$ mit $\mathbf{a}_n \in \mathbf{U} \setminus \{\mathbf{a}\}$ und $\mathbf{a}_n \rightarrow \mathbf{a}$ gilt $f(\mathbf{a}_n) \rightarrow \infty$.
b. $\lim_{x \rightarrow \infty} f(x) = \infty \iff \forall (\mathbf{a}_n)_{n \in \mathbb{N}}$ mit $\mathbf{a}_n \in \mathbf{U}$ und $\mathbf{a}_n \rightarrow \infty$ gilt $f(\mathbf{a}_n) \rightarrow \infty$.

Die übrigen Fälle ergeben sich analog. Außerdem verallgemeinern sich auch die Grenzwertsätze für Funktionen 13.10 auf uneigentliche Grenzwerte in der naheliegenden Weise, wenn wir die Konventionen aus Bemerkung 11.35 berücksichtigen.

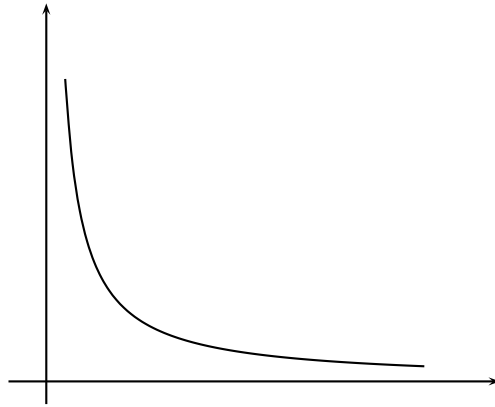
Beispiel 13.18

Für die Funktion $f : (0, \infty) \rightarrow \mathbb{R} : x \mapsto \frac{1}{x}$ ist 0 ein Häufungspunkt des Definitionsbereiches, und es gilt

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

Zudem ist $(0, \infty)$ nach oben unbeschränkt und

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



Beweis: Wenn $(a_n)_{n \in \mathbb{N}}$ eine Folge in $(0, \infty)$ mit $a_n \rightarrow 0$ ist und $s > 0$, so gibt es ein $n_s \in \mathbb{N}$ mit $a_n < \frac{1}{s}$ für $n \geq n_s$. Damit gilt dann für $n \geq n_s$ aber auch

$$f(a_n) = \frac{1}{a_n} > s,$$

d.h. $f(a_n) \rightarrow \infty$.

Wenn $(a_n)_{n \in \mathbb{N}}$ eine Folge in $(0, \infty)$ mit $a_n \rightarrow \infty$ ist und $\varepsilon > 0$, so gibt es ein $n_\varepsilon \in \mathbb{N}$ mit $a_n > \frac{1}{\varepsilon}$ für alle $n \geq n_\varepsilon$. Damit gilt dann für $n \geq n_\varepsilon$ aber auch

$$|f(a_n) - 0| = \frac{1}{a_n} < \varepsilon,$$

d.h. $f(a_n) \rightarrow 0$. □

Beispiel 13.19

Es sei $f = \sum_{k=0}^n a_k \cdot t^k \in \mathbb{R}[t]$ ein Polynom vom Grad $n \geq 1$. Dann gilt

$$\lim_{x \rightarrow \infty} f(x) = \begin{cases} \infty, & \text{falls } a_n > 0, \\ -\infty, & \text{falls } a_n < 0, \end{cases}$$

und

$$\lim_{x \rightarrow -\infty} f(x) = \begin{cases} \infty, & \text{falls } (a_n > 0 \text{ und } n \text{ gerade}) \text{ oder } (a_n < 0 \text{ und } n \text{ ungerade}), \\ -\infty, & \text{falls } (a_n < 0 \text{ und } n \text{ gerade}) \text{ oder } (a_n > 0 \text{ und } n \text{ ungerade}). \end{cases}$$

Wir beweisen die Aussage nur für $\lim_{x \rightarrow \infty} f(x)$ und $a_n > 0$, da der Rest sich analog zeigen läßt. Hierzu betrachten wir ein beliebiges $x \in \mathbb{R}$ mit

$$x \geq \max \left\{ \frac{-2 \cdot n \cdot a_0}{a_n}, \frac{-2 \cdot n \cdot a_1}{a_n}, \dots, \frac{-2 \cdot n \cdot a_{n-1}}{a_n}, 1 \right\}.$$

Dann gilt

$$\frac{a_n \cdot x^n}{2 \cdot n} \geq -a_k \cdot x^k$$

für alle $0 \leq k \leq n-1$, und mithin

$$\frac{a_n \cdot x^n}{2} = n \cdot \frac{a_n \cdot x^n}{2 \cdot n} \geq - \sum_{k=0}^{n-1} a_k \cdot x^k$$

oder alternativ

$$f(x) = \frac{a_n \cdot x^n}{2} + \left(\frac{a_n \cdot x^n}{2} + \sum_{k=0}^{n-1} a_k \cdot x^k \right) \geq \frac{a_n \cdot x^n}{2}.$$

Da zudem offenbar $\lim_{x \rightarrow \infty} \frac{a_n \cdot x^n}{2} = \infty$, muß auch $\lim_{x \rightarrow \infty} f(x) = \infty$ gelten. \square

Aufgaben

Aufgabe 13.20

Bestimme für die nachfolgenden Mengen jeweils die Menge aller ihrer Häufungspunkte:

- $M_1 = \left\{ (-1)^n + \left(\frac{-1}{n}\right)^{n+1} \mid n \in \mathbb{Z}_{\geq 1} \right\}$.
- $M_2 = \mathbb{N}$.

Aufgabe 13.21

Bestimme die folgenden Grenzwerte:

- $\lim_{x \rightarrow 2} \frac{x^3 - 2x^2 - x + 2}{x^2 - x - 2}$.
- $\lim_{x \rightarrow x_0} \frac{x^n - x_0^n}{x - x_0}$, wobei $n \in \mathbb{N}$ und $x_0 \in \mathbb{R}$ beliebig, aber fest vorgegeben sind.
- $\lim_{x \rightarrow \infty} (\sqrt{x+5} - \sqrt{x})$.

§ 14 Stetigkeit

A) Stetige Funktionen

Definition 14.1 (ε - δ -Kriterium für Stetigkeit)

Es sei $U \subseteq \mathbb{R}$, $f : U \rightarrow \mathbb{R}$ und $a \in U$.

Wir nennen f *stetig in a* , wenn

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : \forall x \in U \text{ mit } |x - a| < \delta_\varepsilon \text{ gilt } |f(x) - f(a)| < \varepsilon.$$

Die Funktion f heißt *stetig* (auf U), wenn sie stetig in jedem Punkt in U ist.

$\mathcal{C}(U, \mathbb{R}) := \{f : U \rightarrow \mathbb{R} \mid f \text{ stetig}\}$ ist die Menge der auf U stetigen Funktionen.

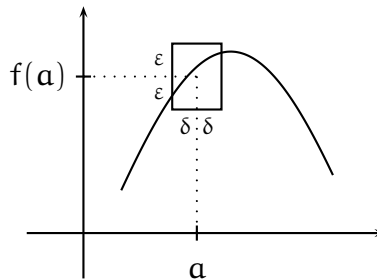


ABBILDUNG 6. ε - δ -Kriterium für Stetigkeit

Bemerkung 14.2

Für die Stetigkeit einer Funktion in einem Punkt a ist nur das Verhalten von f in einer kleinen ε -Umgebung $U_\varepsilon(a) = (a - \varepsilon, a + \varepsilon)$ von a maßgeblich. Wir sagen deshalb auch, daß die Stetigkeit eine *lokale* Eigenschaft ist!

Lemma 14.3 (Stetigkeit in Häufungspunkten)

Es sei $U \subseteq \mathbb{R}$, $f : U \rightarrow \mathbb{R}$ und $a \in U$ ein Häufungspunkt.

Genau dann ist f stetig in a , wenn $\lim_{x \rightarrow a} f(x) = f(a)$.

Beweis: Dies folgt unmittelbar aus den Definitionen 13.6 und 14.1. □

Beispiel 14.4 (Polynomfunktionen sind stetig.)

a. Jede Polynomfunktion $f : \mathbb{R} \rightarrow \mathbb{R}$ ist stetig.

Denn nach Beispiel 13.13 gilt für $a \in \mathbb{R}$ auch $\lim_{x \rightarrow a} f(x) = f(a)$.

b. Jede rationale Funktion $\frac{f}{g} : \mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) = 0\} \rightarrow \mathbb{R}$ ist stetig.

Denn nach Beispiel 13.13 ist $a \in \mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) \neq 0\}$ ein Häufungspunkt des Definitionsbereiches und $\lim_{x \rightarrow a} \frac{f}{g}(x) = \frac{f}{g}(a)$.

c. Die Funktion

$$f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \begin{cases} 1, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0 \end{cases}$$

aus Beispiel 13.8 b. ist nicht stetig in 0 , da $\lim_{x \rightarrow 0} f(x) = 1 \neq 0 = f(0)$. Aber, f ist stetig in jedem $a \neq 0$, wie man leicht sieht.

- d. Ist $f : \mathcal{U} \rightarrow \mathbb{R}$ stetig und $V \subseteq \mathcal{U}$, so ist die Einschränkung $f|_V : V \rightarrow \mathbb{R}$ von f auf V offenbar ebenfalls stetig.
- e. Ist $f : \mathbb{Z} \rightarrow \mathbb{R}$ irgendeine Funktion, so ist f stetig! (Kein nützliches Konzept!)
Denn, ist $\mathbf{a} \in \mathbb{Z}$ und $\varepsilon > 0$ wählen wir $\delta_\varepsilon := \frac{1}{2}$. Für $x \in \mathbb{Z}$ mit $|x - \mathbf{a}| < \delta_\varepsilon = \frac{1}{2}$ muß dann $x = \mathbf{a}$ gelten und somit auch $|f(x) - f(\mathbf{a})| = 0 < \varepsilon$.

Satz 14.5 (Folgenkriterium für Stetigkeit)

Es sei $\mathcal{U} \subseteq \mathbb{R}$, $f : \mathcal{U} \rightarrow \mathbb{R}$ und $\mathbf{a} \in \mathcal{U}$.

Genau dann ist f stetig in \mathbf{a} , wenn

$$\forall (\mathbf{a}_n)_{n \in \mathbb{N}} \text{ mit } \mathbf{a}_n \in \mathcal{U} \text{ und } \lim_{n \rightarrow \infty} \mathbf{a}_n = \mathbf{a} \text{ gilt } \lim_{n \rightarrow \infty} f(\mathbf{a}_n) = f(\mathbf{a}). \quad (18)$$

Beweis: Der Beweis geht genau wie der Beweis des Folgenkriteriums für Grenzwerte von Funktionen 13.7.

“ \implies ”:
Es sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in \mathcal{U} mit $\lim_{n \rightarrow \infty} \mathbf{a}_n = \mathbf{a}$. Wir müssen $\lim_{n \rightarrow \infty} f(\mathbf{a}_n) = f(\mathbf{a})$ zeigen. Dazu sei $\varepsilon > 0$ gegeben.

Da f stetig in \mathbf{a} ist, gibt es ein $\delta_\varepsilon > 0$, so daß aus $x \in \mathcal{U}$ mit $|x - \mathbf{a}| < \delta_\varepsilon$ auch $|f(x) - f(\mathbf{a})| < \varepsilon$ folgt.

Wegen $\lim_{n \rightarrow \infty} \mathbf{a}_n = \mathbf{a}$ gibt es zu δ_ε nun ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $n \geq n_\varepsilon$ auch $|\mathbf{a}_n - \mathbf{a}| < \delta_\varepsilon$ gilt.

Sei nun $n \geq n_\varepsilon$ dann erfüllt $\mathbf{a}_n \in \mathcal{U}$ die Bedingung $|\mathbf{a}_n - \mathbf{a}| < \delta_\varepsilon$ und somit ist auch $|f(\mathbf{a}_n) - f(\mathbf{a})| < \varepsilon$. Damit ist $f(\mathbf{a}_n) \rightarrow f(\mathbf{a})$ gezeigt.

“ \impliedby ”:
Wir nehmen an, f wäre nicht stetig in \mathbf{a} . Dann gilt:

$$\exists \varepsilon > 0 : \forall \delta_\varepsilon > 0 \exists x_{\delta_\varepsilon} \in \mathcal{U} \text{ mit } |x_{\delta_\varepsilon} - \mathbf{a}| < \delta_\varepsilon, \text{ aber } |f(x_{\delta_\varepsilon}) - f(\mathbf{a})| \geq \varepsilon.$$

Für $n \geq 1$ und $\delta_\varepsilon = \frac{1}{n}$ setzen wir $\mathbf{a}_n := x_{\delta_\varepsilon} = x_{\frac{1}{n}} \in \mathcal{U} \setminus \{\mathbf{a}\}$. Dann gilt

$$0 \leq |\mathbf{a}_n - \mathbf{a}| < \frac{1}{n} \rightarrow 0,$$

so daß $\mathbf{a}_n \rightarrow \mathbf{a}$, und zugleich gilt

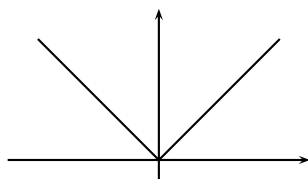
$$|f(\mathbf{a}_n) - f(\mathbf{a})| \geq \varepsilon$$

für alle $n \in \mathbb{N}$. Dies ist ein Widerspruch dazu, daß $f(\mathbf{a}_n)$ gegen $f(\mathbf{a})$ konvergieren muß.

□

Beispiel 14.6 (Die Betragsfunktion ist stetig)

Die Betragsfunktion $|\cdot| : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto |x|$ ist stetig.



Denn für $\mathbf{a} \in \mathbb{R}$ und $(\mathbf{a}_n)_{n \in \mathbb{N}}$ mit $\mathbf{a}_n \rightarrow \mathbf{a}$ gilt aufgrund der Grenzwertsätze für Folgen 11.15 auch $|\mathbf{a}_n| \rightarrow |\mathbf{a}|$.

Proposition 14.7 (Rechenregeln für stetige Funktionen)

Seien $f : \mathbf{U} \rightarrow \mathbb{R}$ und $g : \mathbf{U} \rightarrow \mathbb{R}$ Funktionen, die in $\mathbf{a} \in \mathbf{U}$ stetig sind, und $c \in \mathbb{R}$.

- $c \cdot f, f + g, f - g$ und $f \cdot g$ sind stetig in \mathbf{a} .
- Ist $g(\mathbf{a}) \neq 0$, so ist auch $\frac{f}{g} : \mathbf{U} \setminus \{x \in \mathbf{U} \mid g(x) = 0\} \rightarrow \mathbb{R}$ stetig in \mathbf{a} .

Beweis: Der Beweis folgt aus dem Folgenkriterium für Stetigkeit 14.5 und den Grenzwertsätzen für Folgen 11.15.

Z.B. sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in \mathbf{U} mit $\mathbf{a}_n \rightarrow \mathbf{a}$, dann gilt

$$(f + g)(\mathbf{a}_n) = f(\mathbf{a}_n) + g(\mathbf{a}_n) \rightarrow f(\mathbf{a}) + g(\mathbf{a}) = (f + g)(\mathbf{a}),$$

da f und g in \mathbf{a} stetig sind. Also ist auch $f + g$ stetig in \mathbf{a} . \square

Proposition 14.8 (Komposition stetiger Funktionen)

Es seien $f : \mathbf{U} \rightarrow \mathbb{R}$ und $g : \mathbf{V} \rightarrow \mathbb{R}$ Funktionen mit $\text{Im}(f) \subseteq \mathbf{V}$ und es sei $\mathbf{a} \in \mathbf{U}$. Ist f stetig in \mathbf{a} und g stetig in $f(\mathbf{a})$, so ist $g \circ f$ stetig in \mathbf{a} .

Beweis: Sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in \mathbf{U} mit $\mathbf{a}_n \rightarrow \mathbf{a}$, dann ist $(f(\mathbf{a}_n))_{n \in \mathbb{N}}$ eine Folge in \mathbf{V} und, da f stetig in \mathbf{a} ist, gilt zudem $f(\mathbf{a}_n) \rightarrow f(\mathbf{a})$. Nun ist auch g stetig in $f(\mathbf{a})$, so daß daraus

$$(g \circ f)(\mathbf{a}_n) = g(f(\mathbf{a}_n)) \rightarrow g(f(\mathbf{a})) = (g \circ f)(\mathbf{a})$$

folgt. Aufgrund des Folgenkriteriums für Stetigkeit 14.5 ist dann $g \circ f$ stetig in \mathbf{a} . \square

Beispiel 14.9

Ist $f : \mathbf{U} \rightarrow \mathbb{R}$ stetig in $\mathbf{a} \in \mathbf{U}$, so ist auch $|f| : \mathbf{U} \rightarrow \mathbb{R} : x \mapsto |f(x)|$ als Komposition stetiger Funktionen stetig in \mathbf{a} .

Definition 14.10 (Stetig fortsetzbar)

Es sei $f : \mathbf{U} \rightarrow \mathbb{R}$ eine stetige Funktion und $\mathbf{a} \in \mathbb{R} \setminus \mathbf{U}$ ein Häufungspunkt von \mathbf{U} . Wir nennen f in \mathbf{a} *stetig fortsetzbar*, wenn $\lim_{x \rightarrow \mathbf{a}} f(x)$ existiert.

In dieser Situation nennen wir

$$g : \mathbf{U} \cup \{\mathbf{a}\} \rightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } x \neq \mathbf{a}, \\ \lim_{z \rightarrow \mathbf{a}} f(z), & \text{falls } x = \mathbf{a}, \end{cases}$$

die *stetige Fortsetzung* von f , und g ist nach Lemma 14.3 stetig in \mathbf{a} und damit stetig auf $\mathbf{U} \cup \{\mathbf{a}\}$.

Beispiel 14.11

- Die Funktion

$$f : \mathbb{R} \setminus \{1\} \rightarrow \mathbb{R} : x \mapsto \frac{x^2 - 1}{x - 1}$$

aus Beispiel 13.8 c. ist in $\mathbf{a} = 1$ stetig fortsetzbar, und die stetige Fortsetzung ist

$$g : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x + 1.$$

b. Die Funktion

$$f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x < 0, \\ 1, & \text{falls } x > 0 \end{cases}$$

aus Beispiel 13.8 d. ist in $\mathbf{a} = 0$ nicht stetig fortsetzbar, da der Grenzwert von f in 0 nicht existiert.

c. Die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 1, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0 \end{cases}$$

aus Beispiel 13.8 b. ist nach unserer Definition in $\mathbf{a} = 0$ nicht stetig fortsetzbar, obwohl der Grenzwert von f in 0 existiert, da 0 bereits zum Definitionsbereich der Funktion gehört!

B) Wichtige Eigenschaften stetiger Funktionen

Satz 14.12 (Zwischenwertsatz)

Eine stetige Funktion $f : [\mathbf{a}, \mathbf{b}] \longrightarrow \mathbb{R}$ nimmt jeden Wert zwischen $f(\mathbf{a})$ und $f(\mathbf{b})$ an.

Beweis: Für den Beweis können wir $f(\mathbf{a}) \leq f(\mathbf{b})$ annehmen. Für $\mathbf{c} \in [f(\mathbf{a}), f(\mathbf{b})]$ definieren wir eine Funktion

$$g : [\mathbf{a}, \mathbf{b}] \longrightarrow \mathbb{R} : x \mapsto f(x) - \mathbf{c},$$

und diese ist aufgrund der Proposition 14.7 stetig auf $[\mathbf{a}, \mathbf{b}]$.

Wir müssen zeigen, daß g eine Nullstelle in $[\mathbf{a}, \mathbf{b}]$ besitzt.

Dazu wenden wir wie im Beweis des Satzes von Bolzano-Weierstraß 11.26 ein Intervallschachtelungsverfahren an. Wir setzen

$$[\mathbf{a}_0, \mathbf{b}_0] := [\mathbf{a}, \mathbf{b}]$$

und betrachten den Punkt

$$x_0 = \frac{\mathbf{a}_0 + \mathbf{b}_0}{2} \in [\mathbf{a}, \mathbf{b}].$$

Ist $g(x_0) = 0$, so sind wir fertig. Andernfalls gilt entweder $g(x_0) > 0$ und wir setzen $[\mathbf{a}_1, \mathbf{b}_1] := [\mathbf{a}_0, x_0]$, oder es gilt $g(x_0) < 0$ und wir setzen $[\mathbf{a}_1, \mathbf{b}_1] := [x_0, \mathbf{b}_0]$.

Mit dem neuen Intervall verfahren wir wie mit dem vorherigen. Auf dem Weg finden wir entweder nach endlich vielen Schritten einen Punkt $x_n \in [\mathbf{a}, \mathbf{b}]$ mit $g(x_n) = 0$, oder wir konstruieren rekursiv eine monoton steigende, beschränkte Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in $[\mathbf{a}, \mathbf{b}]$ und eine monoton fallende, beschränkte $(\mathbf{b}_n)_{n \in \mathbb{N}}$ in $[\mathbf{a}, \mathbf{b}]$ mit

$$\mathbf{b}_n - \mathbf{a}_n = \frac{\mathbf{b} - \mathbf{a}}{2^n} \longrightarrow 0. \quad (19)$$

Aufgrund des Monotoniekriteriums für Folgen konvergiert $(\mathbf{a}_n)_{n \in \mathbb{N}}$ gegen einen Wert x und $(\mathbf{b}_n)_{n \in \mathbb{N}}$ gegen einen Wert y , und wegen (19) gilt dann

$$x = y \in [\mathbf{a}, \mathbf{b}].$$

Da das Intervall $[\mathbf{a}, \mathbf{b}]$ abgeschlossen ist, gilt zudem nach Satz 11.28

$$x \in [\mathbf{a}, \mathbf{b}].$$

Man beachte auch, daß aufgrund der Konstruktion von $(\mathbf{a}_n)_{n \in \mathbb{N}}$ und $(\mathbf{b}_n)_{n \in \mathbb{N}}$ stets

$$g(\mathbf{a}_n) < 0 \quad \text{und} \quad g(\mathbf{b}_n) > 0.$$

Für die stetige Funktion g folgt dann aus dem Folgenkriterium 14.5 und Satz 11.17

$$g(x) = \lim_{n \rightarrow \infty} g(\mathbf{a}_n) \leq 0 \leq \lim_{n \rightarrow \infty} g(\mathbf{b}_n) = g(x),$$

also $g(x) = 0$. □

Beispiel 14.13 (Nullstellen von Polynomfunktionen)

Ist $f \in \mathbb{R}[t]$ ein Polynom von ungeradem Grad, so besitzt f eine Nullstelle.

Denn nach Beispiel 13.19 gilt, daß $\lim_{x \rightarrow \infty} f(x)$ und $\lim_{x \rightarrow -\infty} f(x)$ verschiedene Vorzeichen haben, so daß es $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ mit $f(\mathbf{a}) > 0$ und $f(\mathbf{b}) < 0$ geben muß. Wenden wir dann den Zwischenwertsatz auf $f|_{[\mathbf{a}, \mathbf{b}]}$ bzw. $f|_{[\mathbf{b}, \mathbf{a}]}$ an, so folgt die Behauptung.

Definition 14.14 (Beschränkte Funktionen)

Eine Funktion $f: \mathcal{U} \rightarrow \mathbb{R}$ heißt *beschränkt*, wenn $\text{Im}(f)$ beschränkt ist.

Proposition 14.15 (Beschränktheit stetiger Funktionen)

Eine stetige Funktion $f: [\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ ist beschränkt.

Beweis: Nehmen wir an, f wäre nicht beschränkt. Dann gibt es für jedes $n \in \mathbb{N}$ ein $\mathbf{a}_n \in [\mathbf{a}, \mathbf{b}]$ mit

$$|f(\mathbf{a}_n)| > n.$$

Die Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ ist beschränkt, da sie im abgeschlossenen Intervall $[\mathbf{a}, \mathbf{b}]$ liegt, und nach dem Satz von Bolzano-Weierstraß 11.26 gibt es also eine konvergente Teilfolge $(\mathbf{a}_{n_k})_{k \in \mathbb{N}}$ mit Grenzwert c , d.h.

$$\mathbf{a}_{n_k} \rightarrow c.$$

Da das Intervall $[\mathbf{a}, \mathbf{b}]$ abgeschlossen ist, gilt nach Satz 11.28

$$c \in [\mathbf{a}, \mathbf{b}].$$

Da f und somit nach Beispiel 14.9 auch $|f|$ stetig auf $[\mathbf{a}, \mathbf{b}]$ ist, folgt

$$|f(c)| \leftarrow |f(\mathbf{a}_{n_k})| \geq n_k \rightarrow \infty,$$

was ein offensichtlicher Widerspruch ist. □

Satz 14.16 (Maximum / Minimum stetiger Funktionen)

Eine stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ nimmt ihr Maximum und ihr Minimum an, d.h. es gibt $c, d \in [a, b]$, so daß für alle $x \in [a, b]$ gilt

$$f(c) \leq f(x) \leq f(d).$$

Beweis: Nach Proposition 14.15 ist die Menge

$$A := \text{Im}(f) = \{f(x) \mid x \in [a, b]\}$$

beschränkt und somit existiert

$$y := \sup(A) \in \mathbb{R}.$$

Da y die kleinste obere Schranke von A ist, gibt es für jedes $n \geq 1$ ein $a_n \in [a, b]$ mit

$$y - \frac{1}{n} < f(a_n) \leq y.$$

Die Folge $(a_n)_{n \geq 1}$ ist beschränkt, da sie im abgeschlossenen Intervall $[a, b]$ liegt, also besitzt sie nach dem Satz von Bolzano-Weierstraß 11.26 eine konvergente Teilfolge $(a_{n_k})_{k \in \mathbb{N}}$ mit Grenzwert d . Dann gilt aber

$$y \leftarrow y - \frac{1}{n_k} < f(a_{n_k}) \leq y \rightarrow y,$$

so daß aufgrund des Einschachtelungssatzes 11.17 auch

$$f(a_{n_k}) \rightarrow y$$

gilt. Da f aber stetig ist, folgt dann

$$f(d) = \lim_{k \rightarrow \infty} f(a_{n_k}) = y.$$

Die Existenz von c zeigt man analog mit Hilfe von $\inf(A)$. □

Beispiel 14.17

- a. Die Funktion $f : [-1, 1] \rightarrow \mathbb{R} : x \mapsto x^2$ ist beschränkt, und es gilt $f(0) = 0$ ist das Minimum und $f(1) = f(-1) = 1$ ist das Maximum von $\text{Im}(f)$.
- b. Die Funktion $f : (0, \infty) \rightarrow \mathbb{R} : x \mapsto \frac{1}{x}$ ist nicht beschränkt und nimmt weder ihr Minimum noch ihr Maximum an.

C) Umkehrsatz für streng monotone stetige Funktionen**Definition 14.18** (Monotone Funktionen)

Es sei $f : U \rightarrow \mathbb{R}$ eine Funktion.

- a. f heißt *monoton wachsend*, wenn für $x, y \in U$ aus $x \leq y$ stets $f(x) \leq f(y)$ folgt.
- b. f heißt *streng monoton wachsend*, wenn für $x, y \in U$ aus $x < y$ stets $f(x) < f(y)$ folgt.
- c. f heißt *monoton fallend*, wenn für $x, y \in U$ aus $x \leq y$ stets $f(x) \geq f(y)$ folgt.

- d. f heißt *streng monoton fallend*, wenn für $x, y \in U$ aus $x < y$ stets $f(x) > f(y)$ folgt.

Beispiel 14.19

- a. Die Funktion $f : [0, \infty) \rightarrow \mathbb{R} : x \mapsto x^n$ ist für jedes $n \geq 1$ streng monoton wachsend, da nach Lemma 8.17 aus $0 \leq x < y$ stets $x^n < y^n$ folgt.
- b. Die Funktion

$$f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x < 0, \\ 1, & \text{falls } x > 0 \end{cases}$$

aus Beispiel 13.8 d. ist monoton wachsend, aber nicht streng monoton wachsend.

Bemerkung 14.20

Ist $f : U \rightarrow \mathbb{R}$ streng monoton wachsend oder fallend, so ist f injektiv.

Denn, für $x, y \in U$ mit $x \neq y$ gilt $x < y$ oder $x > y$ und somit $f(x) < f(y)$ oder $f(x) > f(y)$, aber in jedem Fall $f(x) \neq f(y)$.

Satz 14.21 (Umkehrsatz für streng monotone stetige Funktionen)

Es seien $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$ mit $a < b$, $f : (a, b) \rightarrow \mathbb{R}$ sei eine Funktion und es seien $c := \inf(\text{Im}(f)) \in \mathbb{R} \cup \{-\infty\}$ und $d := \sup(\text{Im}(f)) \in \mathbb{R} \cup \{\infty\}$.

- a. Ist f streng monoton wachsend und stetig, so gelten:
- (i) $f : (a, b) \rightarrow (c, d)$ ist bijektiv.
 - (ii) $f^{-1} : (c, d) \rightarrow (a, b)$ ist streng monoton wachsend und stetig.
- b. Ist f streng monoton fallend und stetig, so gelten:
- (i) $f : (a, b) \rightarrow (c, d)$ ist bijektiv.
 - (ii) $f^{-1} : (c, d) \rightarrow (a, b)$ ist streng monoton fallend und stetig.

Beweis: Wir beweisen nur den Fall, daß f streng monoton wachsend ist, da der Beweis für streng monoton fallende Funktionen analog geht.

Zeige: $c, d \notin \text{Im}(f)$: Wäre $d \in \text{Im}(f)$, so würde es ein $x \in (a, b)$ geben mit $f(x) = d$. Wegen $x < b$ gilt dann auch $x < \frac{x+b}{2} \in (a, b)$ und somit

$$d = f(x) < f\left(\frac{x+b}{2}\right) \in \text{Im}(f),$$

im Widerspruch dazu, daß d das Supremum von $\text{Im}(f)$ ist. Analog sieht man, daß $c \notin \text{Im}(f)$.

Zeige: $\text{Im}(f) = (c, d)$: Nach Definition von $c = \inf(\text{Im}(f))$ und $d = \sup(\text{Im}(f))$ sowie nach der obigen Vorüberlegung folgt für $y \in \text{Im}(f)$ sofort $c < y < d$, d.h.

$$\text{Im}(f) \subseteq (c, d).$$

Sei nun $y \in (c, d)$. Wegen $y < d = \sup(\text{Im}(f))$ gibt es ein $x_1 \in (a, b)$ mit $y < f(x_1)$, und wegen $y > c = \inf(\text{Im}(f))$ gibt es ein $x_2 \in (a, b)$ mit $y > f(x_2)$.

Nach Voraussetzung ist die Einschränkung von f

$$f|_{[x_1, x_2]} \longrightarrow \mathbb{R}$$

auf das Intervall $[x_1, x_2]$ stetig als Einschränkung einer stetigen Funktion, und nach dem Zwischenwertsatz 14.12 gibt es wegen $f(x_1) < y < f(x_2)$ dann ein $x \in [x_1, x_2] \subset (c, d)$ mit $y = f(x)$, d.h.

$$(c, d) \subseteq \text{Im}(f).$$

Zeige: $f : (a, b) \longrightarrow (c, d)$ **ist bijektiv:** Nach Bemerkung 14.20 ist die streng monotone Funktion f injektiv, und wie eben gezeigt, ist f surjektiv auf (c, d) .

Zeige: $f^{-1} : (c, d) \longrightarrow (a, b)$ **ist streng monoton wachsend:** Seien $y_1, y_2 \in (c, d)$ mit $y_1 < y_2$. Dann gibt es $x_1, x_2 \in (a, b)$ mit $f(x_1) = y_1 < y_2 = f(x_2)$, und da f streng monoton wachsend ist, muß notwendigerweise auch $x_1 < x_2$ gelten. Dann ist aber

$$f^{-1}(y_1) = x_1 < x_2 = f^{-1}(y_2),$$

und f^{-1} ist streng monoton wachsend.

Zeige: $f^{-1} : (c, d) \longrightarrow (a, b)$ **ist stetig:** Seien $y_0 \in (c, d)$ und $\varepsilon > 0$ gegeben. Wir setzen $x_0 := f^{-1}(y_0) \in (a, b)$ und

$$r_\varepsilon := \min \left\{ \frac{\varepsilon}{2}, \frac{b - x_0}{2}, \frac{x_0 - a}{2} \right\} > 0.$$

Damit gilt

$$a < x_0 - r_\varepsilon < x_0 < x_0 + r_\varepsilon < b$$

und somit

$$f(x_0 - r_\varepsilon) < y_0 < f(x_0 + r_\varepsilon),$$

da f streng monoton wachsend ist. Für

$$\delta_\varepsilon := \min\{y_0 - f(x_0 - r_\varepsilon), f(x_0 + r_\varepsilon) - y_0\} > 0$$

gilt dann offenbar

$$f(x_0 - r_\varepsilon) \leq y_0 - \delta_\varepsilon < y_0 < y_0 + \delta_\varepsilon \leq f(x_0 + r_\varepsilon),$$

und da f^{-1} streng monoton wachsend ist, folgt für $y \in (y_0 - \delta_\varepsilon, y_0 + \delta_\varepsilon) \subset (c, d)$ deshalb

$$x_0 - r_\varepsilon = f^{-1}(f(x_0 - r_\varepsilon)) < f^{-1}(y) < f^{-1}(f(x_0 + r_\varepsilon)) = x_0 + r_\varepsilon$$

d.h.

$$|f^{-1}(y) - f^{-1}(y_0)| = |x_0 - f^{-1}(y)| < 2 \cdot r_\varepsilon \leq \varepsilon.$$

Also ist f^{-1} stetig in y_0 , und damit stetig auf (c, d) .

□

Bemerkung 14.22 (Umkehrsatz für streng monotone stetige Funktionen)

Ist die Abbildung $f : (a, b) \rightarrow \mathbb{R}$ im Umkehrsatz 14.21 streng monoton wachsend, so ist

$$c = \inf(\text{Im}(f)) = \lim_{x \rightarrow a} f(x) \quad \text{und} \quad d = \sup(\text{Im}(f)) = \lim_{x \rightarrow b} f(x),$$

und ist f streng monoton fallend, so ist

$$c = \inf(\text{Im}(f)) = \lim_{x \rightarrow b} f(x) \quad \text{und} \quad d = \sup(\text{Im}(f)) = \lim_{x \rightarrow a} f(x).$$

Außerdem, falls f stetig in $a \in \mathbb{R}$ bzw. in $b \in \mathbb{R}$ fortgesetzt werden kann, so ist $\lim_{x \rightarrow a} f(x) \in \mathbb{R}$ bzw. $\lim_{x \rightarrow b} f(x) \in \mathbb{R}$ und f^{-1} wird durch

$$f^{-1}\left(\lim_{x \rightarrow a} f(x)\right) = a \quad \text{bzw.} \quad f^{-1}\left(\lim_{x \rightarrow b} f(x)\right) = b$$

stetig fortgesetzt. D.h. die Aussagen im Umkehrsatz 14.21 gelten für *halboffene* und *abgeschlossene* Intervalle entsprechend.

Beweis: Wir betrachten nur den Fall f streng monoton wachsend und

$$d := \sup(\text{Im}(f)) \in \mathbb{R} \cup \{\infty\}.$$

1. Fall: $d \in \mathbb{R}$: Zu $\varepsilon > 0$ gibt es ein $y \in \text{Im}(f)$ mit $y > d - \varepsilon$ und es gibt ein $x_0 \in (a, b)$ mit $f(x_0) = y$.

Fall 1.1: $b \in \mathbb{R}$: Wir setzen nun $\delta_\varepsilon := b - x_0$ und erhalten für $x \in (a, b)$ mit $b - x = |x - b| < \delta_\varepsilon = b - x_0$ notwendigerweise $x_0 < x$ und somit auch $y = f(x_0) < f(x)$, d.h.

$$|f(x) - d| = d - f(x) < d - y < \varepsilon.$$

Fall 1.2: $b = \infty$: Wir setzen dann $t = \max\{x_0, 1\}$ und erhalten für $x > t$ dann auch $y = f(x_0) < f(x)$, d.h.

$$|f(x) - d| = d - f(x) < d - y < \varepsilon.$$

In beiden Fällen ist damit $\lim_{x \rightarrow b} f(x) = d$ gezeigt.

2. Fall: $d = \infty$: Zu $s > 0$ gibt es dann ein $y \in \text{Im}(f)$ mit $y > s$ und wieder gibt es ein $x_0 \in (a, b)$ mit $f(x_0) = y$.

Fall 1.1: $b \in \mathbb{R}$: Wir setzen nun $\delta_\varepsilon := b - x_0$ und erhalten für $x \in (a, b)$ mit $b - x = |x - b| < \delta_\varepsilon = b - x_0$ notwendigerweise $x_0 < x$ und somit auch

$$f(x) > f(x_0) = y > s.$$

Fall 1.2: $b = \infty$: Wir setzen dann $t = \max\{x_0, 1\}$ und erhalten für $x > t$ dann auch

$$f(x) > f(x_0) = y > s.$$

In beiden Fällen ist damit wieder $\lim_{x \rightarrow b} f(x) = d$ gezeigt.

Läßt sich nun zudem f in $\mathbf{b} \in \mathbb{R}$ stetig fortsetzen, so heißt dies, daß der Grenzwert

$$\mathbf{d} := \lim_{x \rightarrow \mathbf{b}} f(x) \in \mathbb{R}$$

in \mathbb{R} liegt. Da f^{-1} stetig und streng monoton wachsend auf (\mathbf{c}, \mathbf{d}) ist, gilt zudem

$$\mathbf{b} = \lim_{x \rightarrow \mathbf{d}} f^{-1}(x),$$

und somit läßt sich f^{-1} in \mathbf{d} durch $f^{-1}(\mathbf{d}) = \mathbf{b}$ stetig fortsetzen. \square

Beispiel 14.23 (Wurzelfunktion)

Für $n \geq 2$ ist die Funktion

$$f : (0, \infty) \longrightarrow \mathbb{R} : x \mapsto x^n$$

nach Beispiel 14.19 streng monoton wachsend und nach Beispiel 14.4 stetig. Zudem gilt

$$\inf(\text{Im}(f)) = 0 \quad \text{und} \quad \sup(\text{Im}(f)) = \infty.$$

Nach dem Umkehrsatz 14.21 gibt es also eine Umkehrfunktion

$$\sqrt[n]{\cdot} : (0, \infty) \longrightarrow (0, \infty) : x \mapsto \sqrt[n]{x}$$

und diese ist streng monoton wachsend und stetig.

Dies ist unter anderem ein alternativer Beweis zu Satz 9.10 für die Existenz von n -ten Wurzeln!

Man beachte zudem, daß wegen Bemerkung 14.22

$$\lim_{x \rightarrow 0} \sqrt[n]{x} = 0$$

gilt, so daß die Wurzelfunktion stetig nach 0 fortgesetzt werden kann:

$$\sqrt[n]{\cdot} : [0, \infty) \longrightarrow [0, \infty) : x \mapsto \sqrt[n]{x}.$$

Insbesondere ist auch die Funktion $\sqrt{\cdot} : [0, \infty) \longrightarrow [0, \infty) : x \mapsto \sqrt{x}$ stetig.

Korollar 14.24

$$\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1.$$

Beweis: Wir müssen zeigen, daß die Folge $(\mathbf{a}_n)_{n \geq 2}$ mit $\mathbf{a}_n := \sqrt[n]{n} - 1$ eine Nullfolge ist. Da die Funktion $\sqrt[n]{\cdot}$ streng monoton wachsend ist, folgt aus $n > 1$ auch $\sqrt[n]{n} > \sqrt[n]{1} = 1$, und somit $\mathbf{a}_n > 0$. Aus dem Binomischen Lehrsatz 7.15 folgt damit

$$n = (1 + \mathbf{a}_n)^n = \sum_{k=0}^n \binom{n}{k} \cdot 1^k \cdot \mathbf{a}_n^{n-k} \geq 1 + \frac{n \cdot (n-1)}{2} \cdot \mathbf{a}_n^2 > 1,$$

oder alternativ

$$0 \leftarrow \frac{2}{n} \geq \mathbf{a}_n^2 > 0.$$

Der Einschachtelungssatz 11.17 bedingt dann, daß

$$\mathbf{a}_n^2 \longrightarrow 0,$$

und da die Wurzelfunktion stetig in 0 ist, folgt damit

$$\mathbf{a}_n = \sqrt{\mathbf{a}_n^2} \longrightarrow \sqrt{0} = 0.$$

□

D) Gleichmäßige Stetigkeit

Bemerkung 14.25 (Stetigkeit auf \mathbf{U})

Wir erinnern uns, eine Funktion $f : \mathbf{U} \longrightarrow \mathbb{R}$ heißt *stetig auf \mathbf{U}* , wenn sie in jedem Punkt $\mathbf{a} \in \mathbf{U}$ stetig ist, d.h.

$$\forall \mathbf{a} \in \mathbf{U} \forall \varepsilon > 0 \exists \delta_{\varepsilon, \mathbf{a}} > 0 : \forall \mathbf{x} \in \mathbf{U} \text{ mit } |\mathbf{x} - \mathbf{a}| < \delta_{\varepsilon, \mathbf{a}} \text{ gilt } |f(\mathbf{x}) - f(\mathbf{a})| < \varepsilon.$$

Wir schreiben diesmal $\delta_{\varepsilon, \mathbf{a}}$ statt δ_ε , um zu verdeutlichen, daß wir bei gegebenem $\varepsilon > 0$ zwar in jedem Punkt \mathbf{a} ein geeignetes δ finden müssen, daß dieses δ sich mit dem Punkt \mathbf{a} aber ändern kann! Es hängt also vom Punkt \mathbf{a} ab. In der nächsten Definition wollen wir einen stärkeren Begriff der Stetigkeit einführen, bei dem genau das nicht mehr der Fall ist.

Definition 14.26 (Gleichmäßige Stetigkeit)

Eine Funktion $f : \mathbf{U} \longrightarrow \mathbb{R}$ heißt *gleichmäßig stetig auf \mathbf{U}* , wenn

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : \forall \mathbf{x}, \mathbf{y} \in \mathbf{U} \text{ mit } |\mathbf{x} - \mathbf{y}| < \delta_\varepsilon \text{ gilt } |f(\mathbf{x}) - f(\mathbf{y})| < \varepsilon.$$

Bemerkung 14.27

Offenbar ist jede auf \mathbf{U} gleichmäßig stetige Funktion $f : \mathbf{U} \longrightarrow \mathbb{R}$ auch stetig auf \mathbf{U} .

Satz 14.28

Eine stetige Funktion $f : [\mathbf{a}, \mathbf{b}] \longrightarrow \mathbb{R}$ ist gleichmäßig stetig auf $[\mathbf{a}, \mathbf{b}]$.

Beweis: Angenommen, f wäre nicht gleichmäßig stetig auf $[\mathbf{a}, \mathbf{b}]$. Dann gilt:

$$\exists \varepsilon > 0 : \forall \delta_\varepsilon > 0 : \exists \mathbf{x}_{\delta_\varepsilon}, \mathbf{y}_{\delta_\varepsilon} \in [\mathbf{a}, \mathbf{b}] \text{ mit } |\mathbf{x}_{\delta_\varepsilon} - \mathbf{y}_{\delta_\varepsilon}| < \delta_\varepsilon, \text{ aber } |f(\mathbf{x}_{\delta_\varepsilon}) - f(\mathbf{y}_{\delta_\varepsilon})| \geq \varepsilon.$$

Für $n \geq 1$ und $\delta_\varepsilon := \frac{1}{n}$ setzen wir $\mathbf{a}_n := \mathbf{x}_{\delta_\varepsilon} = \mathbf{x}_{\frac{1}{n}}$ und $\mathbf{b}_n := \mathbf{y}_{\delta_\varepsilon} = \mathbf{y}_{\frac{1}{n}}$. Damit erhalten wir zwei beschränkte Folgen $(\mathbf{a}_n)_{n \geq 1}$ und $(\mathbf{b}_n)_{n \geq 1}$ in $[\mathbf{a}, \mathbf{b}]$. Nach dem Satz von Bolzano-Weierstraß 11.26 besitzt $(\mathbf{a}_n)_{n \geq 1}$ eine konvergente Teilfolge $(\mathbf{a}_{n_k})_{k \in \mathbb{N}}$, und ebenso besitzt dann $(\mathbf{b}_{n_k})_{k \in \mathbb{N}}$ eine konvergente Teilfolge $(\mathbf{b}_{n_{k_1}})_{k_1 \in \mathbb{N}}$. Nach Konstruktion gilt

$$0 \leq |\mathbf{a}_{n_{k_1}} - \mathbf{b}_{n_{k_1}}| \leq \frac{1}{n_{k_1}} \longrightarrow 0,$$

so daß die Grenzwerte von $(\mathbf{a}_{n_{k_1}})_{k_1 \in \mathbb{N}}$ und $(\mathbf{b}_{n_{k_1}})_{k_1 \in \mathbb{N}}$ wegen des Einschachtelungssatzes 11.17 übereinstimmen müssen, d.h.

$$\mathbf{a}_{n_{k_1}} \longrightarrow \mathbf{y} \quad \text{und} \quad \mathbf{b}_{n_{k_1}} \longrightarrow \mathbf{y}.$$

Da das Intervall $[\mathbf{a}, \mathbf{b}]$ abgeschlossen ist, gilt nach Satz 11.28 zudem

$$\mathbf{y} \in [\mathbf{a}, \mathbf{b}].$$

Da f und die Betragsfunktion stetig sind, folgt damit

$$0 = |f(\mathbf{y}) - f(\mathbf{y})| \leftarrow |f(\mathbf{a}_{n_{k_l}}) - f(\mathbf{b}_{n_{k_l}})| \geq \varepsilon,$$

was ein offensichtlicher Widerspruch ist. \square

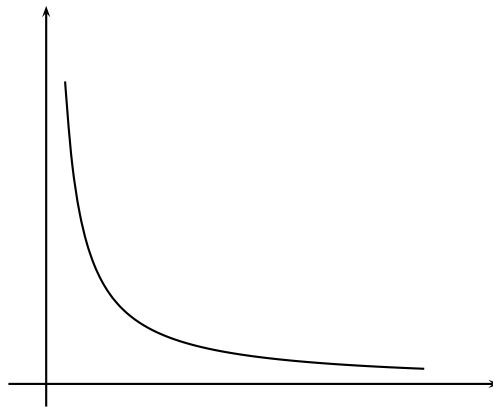
Beispiel 14.29

a. Die Funktion $f : [0, 1] \rightarrow \mathbb{R} : x \mapsto x^2$ ist gleichmäßig stetig auf $[0, 1]$.

Dies folgt aus Satz 14.28. Will man es aus der Definition selbst herleiten, so kann man $\delta_\varepsilon := \frac{\varepsilon}{2}$ zu gegebenem $\varepsilon > 0$ wählen, denn aus $|x - y| < \delta_\varepsilon$ folgt dann

$$|f(x) - f(y)| = |x^2 - y^2| = |x - y| \cdot |x + y| \leq |x - y| \cdot 2 < 2 \cdot \delta_\varepsilon = \varepsilon.$$

b. Die Funktion $f : (0, \infty) \rightarrow \mathbb{R} : x \mapsto \frac{1}{x}$ ist *nicht* gleichmäßig stetig auf $(0, \infty)$.



Dazu setzen wir $\varepsilon := 1$ und wählen $\delta > 0$ beliebig. Dann setzen wir $x := \delta$ und $y := \frac{\delta}{1+\delta}$, also $x, y \in (0, \infty)$ mit

$$|x - y| = \delta - \frac{\delta}{1+\delta} < \delta,$$

aber

$$|f(x) - f(y)| = \left| \frac{1}{\delta} - \frac{1+\delta}{\delta} \right| = 1 \geq \varepsilon.$$

Also ist f nicht gleichmäßig stetig auf $(0, \infty)$.

Das Problem liegt darin, daß bei fest vorgegebenem $\varepsilon > 0$ das $\delta_{\varepsilon, a}$, das man für die Stetigkeit in a wählen muß, immer kleiner werden muß, je näher a an 0 liegt, da die Steigung des Graphen von f nahe bei Null immer steiler wird.

Aufgaben

Aufgabe 14.30

Sei $f : \mathbf{U} \rightarrow \mathbb{R}$ stetig, $a \in \mathbf{U}$ und $b \in \mathbb{R}$.

a. Zeige, ist $f(a) > b$, so gibt, es ein $\delta > 0$, so dass $f(x) > b$ für alle $x \in \mathbf{U} \cap (a - \delta, a + \delta)$.

- b. Zeige, ist $f(\mathbf{a}) \neq \mathbf{b}$, so gibt es ein $\delta > 0$, so dass $f(x) \neq \mathbf{b}$ für alle $x \in U \cap (\mathbf{a} - \delta, \mathbf{a} + \delta)$.

Aufgabe 14.31

Die Wurzelfunktion $\sqrt{\cdot} : [0, \infty) \rightarrow \mathbb{R} : x \mapsto \sqrt{x}$ ist gleichmäßig stetig auf $[0, \infty)$.

Aufgabe 14.32

Verwende die ϵ - δ -Definition der Stetigkeit, um zu zeigen, dass die Funktion $f : [0, 1] \rightarrow \mathbb{R}, x \mapsto \sqrt{1-x^3}$ stetig in $[0, 1]$ ist.

Aufgabe 14.33 (Lipschitz-Stetigkeit)

Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine Funktion und $L \in \mathbb{R}_{>0}$. Zeige, wenn $|f(x) - f(y)| \leq L \cdot |x - y|$ für alle $x, y \in \mathbb{R}$ gilt, so ist f stetig in \mathbb{R} .

Aufgabe 14.34

Sei $f : [0, 1] \rightarrow \mathbb{R}$ eine Funktion definiert durch

$$f(x) := \begin{cases} \frac{2^n}{n!} & \text{für } x = \frac{1}{n} \text{ mit } n \geq 1 \\ 0 & \text{für } x \in [0, 1] \setminus \{\frac{1}{n} \mid n \geq 1\} \end{cases} .$$

Bestimme (mit Beweis) sämtliche Punkte auf $[0, 1]$, in denen f stetig ist.

Aufgabe 14.35

Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ stetig und $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ mit $f(\mathbf{a}) \neq \mathbf{b}$. Zeige, es gibt ein $\delta > 0$, so dass $f(x) \neq \mathbf{b}$ für alle $x \in (\mathbf{a} - \delta, \mathbf{a} + \delta)$.

Aufgabe 14.36 (Fixpunktsatz von Banach)

Sei $f : [\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ eine stetige Abbildung mit $\text{Im}(f) \subseteq [\mathbf{a}, \mathbf{b}]$. Zeige, daß f einen Fixpunkt hat, d.h. es gibt ein $\mathbf{c} \in [\mathbf{a}, \mathbf{b}]$ mit $f(\mathbf{c}) = \mathbf{c}$.

Aufgabe 14.37

Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine stetige Abbildung und $\mathbf{a} \in \mathbb{R}_{>0}$ mit $f(x) = f(x + \mathbf{a})$ für alle $x \in \mathbb{R}$. Zeige, daß es ein $\mathbf{b} \in (0, \mathbf{a})$ gibt mit $f(\mathbf{b} + \frac{\mathbf{a}}{2}) = f(\mathbf{b})$.

Aufgabe 14.38 (Stetige Fortsetzbarkeit)

- Sei $f : (\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ eine stetige Funktion. Zeige, dass f genau dann stetig in \mathbf{a} fortsetzbar ist, wenn f gleichmäßig stetig ist.
- Gibt es eine beschränkte Funktion $f : (0, 1] \rightarrow \mathbb{R}$, die sich nicht stetig in 0 fortsetzen lässt?

§ 15 Konvergenz von Funktionenfolgen

Definition 15.1 (Konvergenz von Funktionenfolgen)

- a. Für jedes $n \in \mathbb{N}$ sei $f_n : U \rightarrow \mathbb{R}$ eine Funktion, so nennen wir $(f_n)_{n \in \mathbb{N}}$ eine *Folge von Funktionen* auf U .
- b. Wir nennen die Folge $(f_n)_{n \in \mathbb{N}}$ von Funktionen *punktweise konvergent auf U* , wenn für jedes $x \in U$ der Grenzwert $\lim_{n \rightarrow \infty} f_n(x)$ existiert, d.h. In diesem Fall nennen wir die Funktion

$$f : U \rightarrow \mathbb{R} : x \mapsto \lim_{n \rightarrow \infty} f_n(x)$$

den *Grenzwert* oder die *Grenzfunktion* der Funktionenfolge $(f_n)_{n \in \mathbb{N}}$, und wir sagen auch, daß $(f_n)_{n \in \mathbb{N}}$ *punktweise gegen f konvergiert*. Wir schreiben dann

$$f = \lim_{n \rightarrow \infty} f_n.$$

Man beachte, $(f_n)_{n \in \mathbb{N}}$ konvergiert auf U genau dann punktweise gegen f , wenn

$$\forall x \in U \forall \varepsilon > 0 \exists n_{\varepsilon, x} : \forall n \geq n_{\varepsilon, x} \text{ gilt } |f_n(x) - f(x)| < \varepsilon.$$

- c. Wir sagen $(f_n)_{n \in \mathbb{N}}$ *konvergiert gleichmäßig auf U gegen f* , wenn

$$\forall \varepsilon > 0 \exists n_\varepsilon : \forall n \geq n_\varepsilon \text{ und } \forall x \in U \text{ gilt } |f_n(x) - f(x)| < \varepsilon.$$

Bemerkung 15.2

Konvergiert $(f_n)_{n \in \mathbb{N}}$ auf U gleichmäßig gegen f , so konvergiert die Folge auch punktweise gegen f .

Beispiel 15.3

Die Folge $f_n : [0, 1] \rightarrow \mathbb{R} : x \mapsto x^n$ konvergiert auf $[0, 1]$ punktweise gegen die Funktion

$$f : [0, 1] \rightarrow \mathbb{R} : x \mapsto \lim_{n \rightarrow \infty} x^n = \begin{cases} 0, & \text{falls } x < 1, \\ 1, & \text{falls } x = 1. \end{cases}$$

Aber, $(f_n)_{n \in \mathbb{N}}$ konvergiert auf $[0, 1]$ *nicht* gleichmäßig gegen f .

Beachte auch, daß die Grenzfunktion nicht stetig in 1 ist, obwohl alle f_n stetig waren!

Um zu sehen, daß die Konvergenz nicht gleichmäßig ist, betrachten wir $\varepsilon := \frac{1}{4} > 0$ und ein beliebiges $n_\varepsilon \in \mathbb{N}$. Setze $n := \max\{n_\varepsilon, 2\} \geq n_\varepsilon$ und $x = \frac{1}{\sqrt[3]{2}} \in [0, 1)$, dann gilt

$$|f_n(x) - f(x)| = \frac{1}{2} > \frac{1}{4} = \varepsilon.$$

Satz 15.4 (Gleichmäßige Konvergenz von Potenzreihen)

Es sei $\sum_{k=0}^{\infty} a_k \cdot t^k$ eine Potenzreihe über \mathbb{R} mit Konvergenzradius r und für $n \in \mathbb{N}$ sei

$$f_n : (-r, r) \rightarrow \mathbb{R} : x \mapsto \sum_{k=0}^n a_k \cdot x^k.$$

Dann konvergiert die Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ auf $(-r, r)$ punktweise gegen

$$f : (-r, r) \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{\infty} a_k \cdot x^k,$$

und für jedes $0 \leq R < r$ konvergiert $(f_n)_{n \in \mathbb{N}}$ auf $[-R, R]$ gleichmäßig gegen f .

Beweis: Daß die f_n auf $(-r, r)$ punktweise gegen f konvergieren, folgt unmittelbar aus der Definition von f_n und f . Es bleibt also nur, für $0 \leq R < r$ zu zeigen, daß die f_n auf $[-R, R]$ gleichmäßig konvergieren.

Sei dazu $\varepsilon > 0$ vorgegeben. Aus Satz 12.32 wissen wir, daß die Reihe $\sum_{k=0}^{\infty} |a_k| \cdot R^k$ konvergiert, und wegen Lemma 12.6 ist die Folge der Restglieder dann eine Nullfolge, so daß es ein $n_\varepsilon \in \mathbb{N}$ gibt mit

$$\sum_{k=n}^{\infty} |a_k| \cdot R^k = \left| \sum_{k=n}^{\infty} |a_k| \cdot R^k \right| < \varepsilon$$

für alle $n \geq n_\varepsilon$. Sei nun $n \geq n_\varepsilon$ und $x \in [-R, R]$, so gilt

$$|f_n(x) - f(x)| = \left| - \sum_{k=n+1}^{\infty} a_k \cdot x^k \right| \leq \sum_{k=n+1}^{\infty} |a_k| \cdot |x|^k \leq \sum_{k=n+1}^{\infty} |a_k| \cdot R^k < \varepsilon.$$

Man beachte hierbei, daß wir hier mehrfach die Proposition 11.17 a. für die betrachteten Folgen der Partialsummen verwenden. \square

Beispiel 15.5

Die Folge $f_n : (-1, 1) \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^n x^k$ konvergiert auf $(-1, 1)$ *nicht* gleichmäßig gegen die geometrische Reihe $f : (-1, 1) \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{\infty} x^k$.

Um dies zu sehen, seien $\varepsilon := 1$ gegeben und $n_\varepsilon \in \mathbb{N}$ beliebig. Wir betrachten zunächst die stetige Funktion

$$g : (-1, 1) \longrightarrow \mathbb{R} : x \mapsto \frac{x^{n_\varepsilon+1}}{1-x}.$$

Man sieht leicht, daß $\lim_{x \rightarrow 1} g(x) = \infty$, so daß es sicher ein $x \in (0, 1)$ mit

$$\frac{x^{n_\varepsilon+1}}{1-x} = g(x) \geq 1 = \varepsilon$$

geben muß. Für dieses x gilt nun

$$|f_{n_\varepsilon}(x) - f(x)| = \sum_{k=n_\varepsilon+1}^{\infty} x^k = x^{n_\varepsilon+1} \cdot \sum_{k=0}^{\infty} x^k = \frac{x^{n_\varepsilon+1}}{1-x} \geq \varepsilon.$$

Mithin konvergiert f_n auf $(-1, 1)$ nicht gleichmäßig gegen f .

Satz 15.6 (Der gleichmäßige Grenzwert stetiger Funktionen ist stetig.)

Ist $f_n : \mathcal{U} \longrightarrow \mathbb{R}$ stetig auf \mathcal{U} für $n \in \mathbb{N}$ und konvergiert $(f_n)_{n \in \mathbb{N}}$ auf \mathcal{U} gleichmäßig gegen f , so ist f stetig auf \mathcal{U} .

Beweis: Seien $\mathbf{a} \in \mathbf{U}$ und $\varepsilon > 0$ gegeben. Da die f_n gleichmäßig gegen f konvergieren, gilt:

$$\exists n_\varepsilon \in \mathbb{N} : \forall n \geq n_\varepsilon \text{ und } \forall x \in \mathbf{U} : |f_n(x) - f(x)| < \frac{\varepsilon}{3}.$$

Da zudem f_{n_ε} stetig in \mathbf{a} ist, gilt:

$$\exists \delta_\varepsilon > 0 \quad \forall x \in \mathbf{U} \text{ mit } |x - \mathbf{a}| < \delta_\varepsilon \text{ gilt } |f_{n_\varepsilon}(x) - f_{n_\varepsilon}(\mathbf{a})| < \frac{\varepsilon}{3}.$$

Sei nun $x \in \mathbf{U}$ mit $|x - \mathbf{a}| < \delta_\varepsilon$ gegeben, so gilt

$$|f(x) - f(\mathbf{a})| \leq |f(x) - f_{n_\varepsilon}(x)| + |f_{n_\varepsilon}(x) - f_{n_\varepsilon}(\mathbf{a})| + |f_{n_\varepsilon}(\mathbf{a}) - f(\mathbf{a})| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Mithin ist f stetig in \mathbf{a} . □

Korollar 15.7 (Potenzreihen sind stetig.)

Ist $\sum_{n=0}^{\infty} a_n \cdot t^n$ eine Potenzreihe mit Konvergenzradius r , dann ist

$$f : (-r, r) \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} a_n \cdot x^n$$

stetig auf $(-r, r)$.

Beweis: Sei $x \in (-r, r)$ beliebig, so ist $0 \leq R < r$ für $R := |x|$. Nach Satz 15.4 konvergiert die Folge stetiger Funktionen

$$f_n : [-R, R] \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^n a_k \cdot x^k$$

auf $[-R, R]$ gleichmäßig gegen f , und nach Satz 15.6 ist f mithin stetig auf $[-R, R]$ und damit insbesondere in $x \in [-R, R]$. □

Beispiel 15.8

Die Exponentialfunktion, der Sinus und der Cosinus sind stetig auf \mathbb{R} .

Dies folgt aus Korollar 15.7 zusammen mit den Sätzen 12.36 und 12.38.

Aufgaben

Aufgabe 15.9

Für $n \geq 2$ sei $f_n = \sqrt[n]{\cdot} : [0, \infty) \longrightarrow \mathbb{R}$. Beweise oder widerlege die folgenden Aussagen.

- a. $(f_n)_{n \in \mathbb{N}}$ konvergiert gleichmäßig auf $[0, \infty)$.
- b. $(f_n)_{n \in \mathbb{N}}$ konvergiert gleichmäßig auf $[1, 100]$.

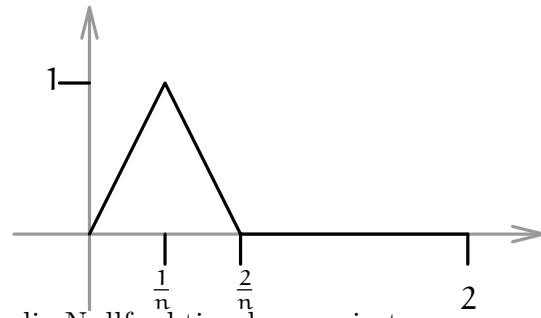
Aufgabe 15.10

Finde eine Folge $(f_n)_{n \in \mathbb{N}}$ von stetigen Funktionen $f_n : [0, 1] \longrightarrow \mathbb{R}$, die punktweise gegen die Nullfunktion konvergiert, aber unbeschränkt ist, d.h., so daß zu jedem $c \in \mathbb{R}$ ein $n \in \mathbb{N}$ und ein $x \in [0, 1]$ existiert mit $|f_n(x)| > c$.

Aufgabe 15.11

Zeige, daß die Folge $(f_n)_{n \geq 1}$ von Funktionen $f_n : [0, 2] \rightarrow \mathbb{R}$, gegeben durch

$$f_n(x) = \begin{cases} nx, & x \in [0, \frac{1}{n}) \\ 2 - nx, & x \in [\frac{1}{n}, \frac{2}{n}) \\ 0, & x \in [\frac{2}{n}, 2] \end{cases} ,$$



punktweise, aber nicht gleichmäßig gegen die Nullfunktion konvergiert.

§ 16 Exponentialfunktion, Logarithmus, trigonometrische Funktionen

A) Exponentialfunktionen und Logarithmusfunktionen

Satz 16.1 (Die Exponentialfunktion)

Die Exponentialfunktion

$$\exp : (-\infty, \infty) \longrightarrow (0, \infty)$$

ist stetig, streng monoton wachsend und bijektiv. Insbesondere gelten

$$\lim_{x \rightarrow -\infty} \exp(x) = 0 \quad \text{und} \quad \lim_{x \rightarrow \infty} \exp(x) = \infty.$$

Beweis: Für $z \in \mathbb{R}$ mit $z > 0$ gilt offenbar

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \geq \frac{z^1}{1!} + \frac{z^0}{0!} = z + 1 > 1, \quad (20)$$

und für $x, y \in \mathbb{R}$ mit $x < y$ folgt mit Hilfe der Funktionalgleichung in Satz 12.36 dann

$$\exp(y) = \exp(y - x + x) = \exp(y - x) \cdot \exp(x) \stackrel{(20)}{>} 1 \cdot \exp(x) = \exp(x),$$

so daß \exp streng monoton wachsend ist. Da \exp nach Beispiel 15.8 zudem stetig ist, können wir den Umkehrsatz für streng monotone Funktionen 14.21 anwenden und erhalten, daß

$$\exp : (-\infty, \infty) \longrightarrow (c, d)$$

auch bijektiv ist, wobei

$$c = \inf(\text{Im}(f)) \stackrel{14.22}{=} \lim_{x \rightarrow -\infty} \exp(x)$$

und

$$d = \sup(\text{Im}(f)) \stackrel{14.22}{=} \lim_{x \rightarrow \infty} \exp(x).$$

Nun gilt für $x > 0$ aber

$$\exp(x) \stackrel{(20)}{\geq} x + 1 \xrightarrow{x \rightarrow \infty} \infty,$$

so daß $d = \lim_{x \rightarrow \infty} \exp(x) = \infty$ folgt.

Außerdem folgt aus der Funktionalgleichung der Exponentialfunktion

$$\exp(x) \cdot \exp(-x) = \exp(x - x) = 1,$$

und aus den Grenzwertsätzen für uneigentliche Grenzwerte von Funktionen folgt dann

$$\exp(-x) = \frac{1}{\exp(x)} \xrightarrow{x \rightarrow \infty} \frac{1}{\infty} = 0,$$

d.h. $c = \lim_{x \rightarrow -\infty} \exp(x) = 0$. □

Bemerkung 16.2

Die Zahl $e = \exp(1)$ ist irrational und es gilt $2 < e < 3$ (siehe auch Beispiel 18.32).

Beweis: Aus (20) wissen wir

$$\exp(1) > 1 + 1 = 2.$$

Wegen

$$\frac{1}{k!} = \frac{1}{1 \cdot 2 \cdot 3 \cdot \dots \cdot k} < \frac{1}{1 \cdot 2 \cdot 2 \cdot \dots \cdot 2} = \frac{1}{2^{k-1}}$$

für $k \geq 4$ folgt unter Berücksichtigung der endlichen geometrischen Reihe 7.12

$$\begin{aligned} s_n &:= \sum_{k=0}^n \frac{1}{k!} \leq 1 + 1 + \frac{1}{2} + \frac{1}{6} + \sum_{k=4}^n \frac{1}{2^{k-1}} \\ &= \frac{8}{3} + \frac{1}{8} \cdot \sum_{k=0}^{n-4} \frac{1}{2^k} = \frac{8}{3} + \frac{1}{8} \cdot \frac{1 - (\frac{1}{2})^{n-3}}{1 - \frac{1}{2}} \\ &\leq \frac{8}{3} + \frac{1}{8} \cdot \frac{1}{1 - \frac{1}{2}} = \frac{8}{3} + \frac{1}{4} = \frac{35}{12} \end{aligned}$$

für alle $n \in \mathbb{N}$. Grenzwertbildung liefert deshalb $\exp(1) = \lim_{n \rightarrow \infty} s_n \leq \frac{35}{12} < 3$.

Wir müssen nun noch zeigen, daß e keine rationale Zahl sein kann. Nehmen wir dazu an, es gelte $e = \frac{p}{q} \in \mathbb{Q}$ mit $p, q \in \mathbb{N}$. Wegen $2 < e < 3$ muß $q \geq 2$ sein. Wir betrachten nun die Zahlen

$$a := \frac{q!}{0!} + \frac{q!}{1!} + \frac{q!}{2!} + \dots + \frac{q!}{q!} \in \mathbb{Z}$$

und

$$b := \sum_{n=q+1}^{\infty} \frac{q!}{n!} = q! \cdot e - a = (q-1)! \cdot p - a \in \mathbb{Z}. \quad (21)$$

Da $q \geq 2$ ist, folgt für $n > q$

$$\frac{q!}{n!} = \frac{1}{(q+1) \cdot (q+2) \cdot \dots \cdot n} \leq \frac{1}{3 \cdot 3 \cdot \dots \cdot 3} = \frac{1}{3^{n-q}}$$

und deshalb

$$b = \sum_{n=q+1}^{\infty} \frac{q!}{n!} \leq \sum_{n=q+1}^{\infty} \frac{1}{3^{n-q}} = \sum_{n=1}^{\infty} \frac{1}{3^n} = \frac{1}{1 - \frac{1}{3}} - 1 = \frac{1}{2}$$

Da aber aufgrund der Definition von b auch

$$b = \sum_{n=q+1}^{\infty} \frac{q!}{n!} > 0$$

gilt, kann b keine ganze Zahl sein, im Widerspruch zu (21). Der Widerspruch kommt von unserer Annahme, daß e eine rationale Zahl wäre. \square

Definition 16.3 (Natürlicher Logarithmus)

Die Umkehrabbildung der Exponentialfunktion wird mit

$$\ln : (0, \infty) \longrightarrow (-\infty, \infty)$$

bezeichnet und (*natürlicher*) *Logarithmus* genannt.

Satz 16.4 (Natürlicher Logarithmus)*Der natürliche Logarithmus*

$$\ln : (0, \infty) \longrightarrow (-\infty, \infty)$$

ist stetig, streng monoton wachsend und bijektiv. Insbesondere gelten

$$\lim_{x \rightarrow 0} \ln(x) = -\infty \quad \text{und} \quad \lim_{x \rightarrow \infty} \ln(x) = \infty.$$

Beweis: Die Aussagen folgen aus dem Umkehrsatz für streng monotone Funktionen 14.21 und Satz 16.1 unter Berücksichtigung von Bemerkung 14.22. \square

Bemerkung 16.5

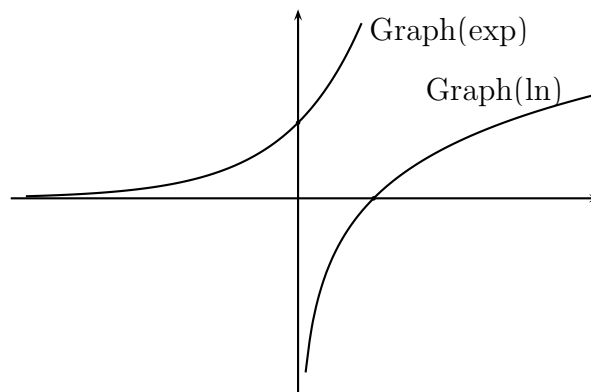
Man beachte, daß aus

$$\exp(0) = 1 \quad \text{und} \quad \exp(1) = e$$

unmittelbar

$$\ln(1) = 0 \quad \text{und} \quad \ln(e) = 1$$

folgt. Die Graphen der Exponentialfunktion und des natürlichen Logarithmus sind in der folgenden Abbildung dargestellt.

**Definition 16.6** (Exponentialfunktion zur Basis a)

Für $a \in \mathbb{R}$ mit $a > 0$ und $x \in \mathbb{R}$ definieren wir

$$a^x := \exp(x \cdot \ln(a)).$$

Man beachte, daß damit $e^x = \exp(x \cdot \ln(e)) = \exp(x \cdot 1) = \exp(x)$ gilt, so daß die neue Definition im Fall $a = e$ mit der Definition aus Bemerkung 12.37 übereinstimmt.

Satz 16.7 (Exponential- und Logarithmusfunktion zur Basis a)

Es sei $a \in \mathbb{R}$ mit $a > 0$ und $a \neq 1$.

a. Die Abbildung

$$\exp_a : \mathbb{R} \longrightarrow (0, \infty) : x \mapsto a^x$$

heißt Exponentialfunktion zur Basis a , ist stetig, bijektiv und

- streng monoton wachsend, falls $a > 1$, und
- streng monoton fallend, falls $a < 1$.

b. *Die Umkehrabbildung*

$$\log_a : (0, \infty) \longrightarrow \mathbb{R}$$

von \exp_a heißt Logarithmus zur Basis a , ist stetig, bijektiv und

- streng monoton wachsend, falls $a > 1$, und
- streng monoton fallend, falls $a < 1$.

Beweis: Für $a > 1$ ist $\ln(a) > 0$, da $\ln(1) = 0$ und \ln streng monoton wachsend, so daß aus $x < y$ auch

$$\ln(a) \cdot x < \ln(a) \cdot y$$

und damit

$$\exp_a(x) = \exp(x \cdot \ln(a)) < \exp(y \cdot \ln(a)) = \exp_a(y)$$

folgt. \exp_a ist dann also streng monoton wachsend. Außerdem gilt

$$\lim_{x \rightarrow \infty} x \cdot \ln(a) = \infty \quad \text{sowie} \quad \lim_{x \rightarrow -\infty} x \cdot \ln(a) = -\infty,$$

und da \exp stetig ist folgt dann auch

$$\lim_{x \rightarrow \infty} \exp_a(x) = \lim_{x \rightarrow \infty} \exp(x \cdot \ln(a)) = \infty$$

sowie

$$\lim_{x \rightarrow -\infty} \exp_a(x) = \lim_{x \rightarrow -\infty} \exp(x \cdot \ln(a)) = 0.$$

Aus dem Umkehrsatz für streng monotone Funktionen 14.21 folgt dann, daß \exp_a stetig, bijektiv und streng monoton wachsend ist. Zudem folgen die entsprechenden Aussagen über die Umkehrfunktion \log_a für $a > 1$.

Den Fall $a < 1$ beweist man analog, da dann $\ln(a) < 0$ gilt. □

Korollar 16.8 (Potenzgesetze)

Seien $a, b, x, y \in \mathbb{R}$ mit $a, b > 0$.

- a. $a^{x+y} = a^x \cdot a^y$.
- b. $a^{x \cdot y} = (a^x)^y$.
- c. $(a \cdot b)^x = a^x \cdot b^x$.
- d. $a^{-x} = \frac{1}{a^x}$.
- e. Für $n \in \mathbb{Z}$ stimmen die Definitionen von a^n in 7.9 und 16.6 überein.
- f. Für $p, q \in \mathbb{Z}$ mit $q \geq 2$ gilt $a^{\frac{p}{q}} = \sqrt[q]{a^p}$.
Insbesondere stimmen die Definitionen von $a^{\frac{1}{q}}$ in 9.10 und 16.6 überein.

Beweis:

- a. Mit Hilfe der Funktionalgleichung für die Exponentialfunktion sieht man:

$$\begin{aligned} a^{x+y} &= \exp((x+y) \cdot \ln(a)) = \exp(x \cdot \ln(a) + y \cdot \ln(a)) \\ &= \exp(x \cdot \ln(a)) \cdot \exp(y \cdot \ln(a)) = a^x \cdot a^y. \end{aligned}$$

b. Wegen $\mathbf{a}^x = \exp(x \cdot \ln(\mathbf{a}))$ gilt auch

$$\ln(\mathbf{a}^x) = \ln(\exp(x \cdot \ln(\mathbf{a}))) = x \cdot \ln(\mathbf{a})$$

und damit

$$(\mathbf{a}^x)^y = \exp(y \cdot \ln(\mathbf{a}^x)) = \exp(x \cdot y \cdot \ln(\mathbf{a})) = \mathbf{a}^{x \cdot y}.$$

c. Wir verwenden in der folgenden Gleichung bereits ein Logarithmusgesetz 16.9, dessen Beweis unabhängig von diesem Potenzgesetz ist:

$$\begin{aligned} (\mathbf{a} \cdot \mathbf{b})^x &= \exp(x \cdot \ln(\mathbf{a} \cdot \mathbf{b})) \stackrel{16.9b.}{=} \exp(x \cdot (\ln(\mathbf{a}) + \ln(\mathbf{b}))) \\ &= \exp(x \cdot \ln(\mathbf{a})) \cdot \exp(x \cdot \ln(\mathbf{b})) = \mathbf{a}^x \cdot \mathbf{b}^x. \end{aligned}$$

d. Wegen

$$\mathbf{a}^x \cdot \mathbf{a}^{-x} = \mathbf{a}^{x-x} = \mathbf{a}^0 = \exp(0 \cdot \ln(\mathbf{a})) = \exp(0) = 1$$

gilt auch $\mathbf{a}^{-x} = \frac{1}{\mathbf{a}^x}$.

e. Mit Induktion nach $n \geq 1$ und a. sieht man, daß $\mathbf{a}^n = \prod_{k=1}^n \mathbf{a}$. Wir haben bereits gesehen, daß zudem $\mathbf{a}^0 = 1$ gilt, und aus d. folgt dann

$$\mathbf{a}^{-n} = \frac{1}{\mathbf{a}^n} = \frac{1}{\prod_{k=1}^n \mathbf{a}} = \prod_{k=1}^n \frac{1}{\mathbf{a}}.$$

Die Definitionen von \mathbf{a}^n in 7.9 und 16.6 stimmen also überein.

f. Es gilt

$$\left(\mathbf{a}^{\frac{p}{q}}\right)^q \stackrel{b.}{=} \mathbf{a}^{\frac{p}{q} \cdot q} = \mathbf{a}^p,$$

so daß $\mathbf{a}^{\frac{p}{q}} = \sqrt[q]{\mathbf{a}^p}$ nach Satz 9.10 gilt.

□

Korollar 16.9 (Logarithmusgesetze)

Seien $\mathbf{a}, x, y \in \mathbb{R}_{>0}$ mit $\mathbf{a} \neq 1$ und $z \in \mathbb{R}$.

- $\log_a(x) = \frac{\ln(x)}{\ln(\mathbf{a})}$.
- $\log_a(x \cdot y) = \log_a(x) + \log_a(y)$.
- $\log_a(x^z) = z \cdot \log_a(x)$.
- $\log_a\left(\frac{x}{y}\right) = \log_a(x) - \log_a(y)$.

Beweis:

a. Falls $\mathbf{a} \neq 1$, so gilt

$$\exp_a\left(\frac{\ln(x)}{\ln(\mathbf{a})}\right) = \exp\left(\frac{\ln(x)}{\ln(\mathbf{a})} \cdot \ln(\mathbf{a})\right) = \exp(\ln(x)) = x = \exp_a(\log_a(x)),$$

und da \exp_a injektiv ist, gilt dann auch

$$\frac{\ln(x)}{\ln(\mathbf{a})} = \log_a(x).$$

b. Es gilt

$$\begin{aligned} \exp_a(\log_a(x \cdot y)) &= x \cdot y = \exp_a(\log_a(x)) \cdot \exp_a(\log_a(y)) \\ &\stackrel{16.8a.}{=} \exp_a(\log_a(x) + \log_a(y)), \end{aligned}$$

und da \exp_a injektiv ist, folgt somit

$$\log_a(x \cdot y) = \log_a(x) + \log_a(y).$$

c. Falls $a \neq 1$ und $x > 0$, so ist

$$x^z = \exp(z \cdot \ln(x)) \stackrel{a.}{=} \exp(z \cdot \log_a(x) \cdot \ln(a)) = \exp_a(z \cdot \log_a(x))$$

definiert. Wenden wir auf beiden Seiten die Funktion \log_a an, so erhalten wir

$$\log_a(x^z) = \log_a(\exp_a(z \cdot \log_a(x))) = z \cdot \log_a(x).$$

d. Es gilt

$$\log_a\left(\frac{x}{y}\right) \stackrel{b.}{=} \log_a(x) + \log_a(y^{-1}) \stackrel{c.}{=} \log_a(x) - \log_a(y).$$

□

B) Trigonometrische Funktionen

Wir wollen uns nun den trigonometrischen Funktionen zuwenden. Dazu führen wir zunächst die Zahl π als kleinste positive Nullstelle des Sinus ein.

Satz 16.10 (Definition der Zahl π .)

Der Sinus besitzt eine kleinste positive Nullstelle, die wir π nennen, und für alle $x \in (0, \pi)$ gilt $\sin(x) > 0$.

Beweis: Wir wählen ein $x \in (0, 4]$ und setzen $a_n := \frac{x^{2n+1}}{(2n+1)!}$ für $n \geq 1$. Die Folge $(a_n)_{n \geq 1}$ ist monoton fallend, denn

$$a_{n+1} = \frac{x^{2n+3}}{(2n+3)!} = \frac{x^{2n+1}}{(2n+1)!} \cdot \frac{x \cdot x}{(2n+3) \cdot (2n+2)} \leq \frac{x^{2n+1}}{(2n+1)!} \cdot \frac{16}{20} < a_n,$$

und da die Reihe $\sum_{n=1}^{\infty} (-1)^n \cdot a_n = \sin(x) - x$ absolut konvergiert, muß die Folge $(a_n)_{n \geq 1}$ auch eine Nullfolge sein. Aus dem Beweis des Leibniz-Kriteriums erfüllen die Partialsummen der Reihe $\sum_{n=1}^{\infty} (-1)^n \cdot a_n$ dann insbesondere

$$s_1 \leq \sum_{n=1}^{\infty} (-1)^n \cdot a_n \leq s_4,$$

und damit

$$x - \frac{x^3}{6} = x + s_1 \leq \sin(x) \leq x + s_4. \tag{22}$$

Wenden wir dies für $x = 1$ an, so erhalten wir

$$\sin(1) \geq 1 - \frac{1}{6} > 0,$$

und wenden wir die Aussage für $x = 4$ an, so erhalten wir

$$\sin(4) \leq 4 + s_4 = -\frac{268}{405} < 0.$$

Da der Sinus auf dem abgeschlossenen Intervall $[1, 4]$ stetig ist mit $\sin(1) > 0$ und $\sin(4) < 0$, muß er nach dem Zwischenwertsatz 14.12 eine Nullstelle besitzen, das heißt, die Menge

$$A := \{x \in [1, 4] \mid \sin(x) = 0\}$$

ist nicht leer und nach unten beschränkt. Dann existiert aber ihr Infimum

$$\pi := \inf(A).$$

Für jedes $n \geq 1$ ist $\pi + \frac{1}{n}$ keine untere Schranke von A , so daß es ein $b_n \in A$ gibt mit

$$\pi \leq b_n < \pi + \frac{1}{n}.$$

Folglich gilt nach dem Einschachtelungssatz 11.17 $b_n \rightarrow \pi$. Da der Sinus stetig ist, gilt dann auch

$$0 = \sin(b_n) \rightarrow \sin(\pi),$$

also $\sin(\pi) = 0$, d.h. π ist eine Nullstelle des Sinus.

Wir müssen nun noch zeigen, daß $\sin(x) > 0$ für alle $x \in (0, \pi)$. Für $x \in [1, \pi)$ ist dies der Fall, da entweder aus $\sin(x) = 0$ oder aus $\sin(x) < 0$ und $\sin(1) > 0$ mit Hilfe des Zwischenwertsatzes 14.12 die Existenz einer kleineren Nullstelle des Sinus als π im Intervall $[1, 4]$ folgen würde. Für $x \in (0, 1)$ folgt aber aus (22)

$$\sin(x) \geq x - \frac{x^3}{6} > x - \frac{x}{6} \geq \frac{5x}{6} > 0,$$

da $x^3 < x$. Also haben wir $\sin(x) > 0$ für alle $x \in (0, \pi)$ gezeigt, so daß π die kleinste positive Nullstelle des Sinus ist. \square

Bemerkung 16.11 (Approximation von π)

Aus dem Beweis von Satz 16.10 wissen wir bislang nur, daß $1 < \pi < 4$ gilt. Wir werden später sehen (siehe Aufgabe 18.39), daß man die Zahl π approximieren kann durch

$$3,14159\dots$$

Satz 16.12 (Monotonie des Cosinus)

Der Cosinus $\cos : [0, \pi] \rightarrow [-1, 1]$ ist streng monoton fallend und bijektiv.

Beweis: Es seien $x, y \in [0, \pi]$ mit $x < y$. Aus dem Additionstheorem für den Cosinus sowie der Tatsache, daß der Cosinus eine gerade und der Sinus eine ungerade Funktion ist (siehe Satz 12.38), folgen

$$\begin{aligned} \cos(y) &= \cos\left(\frac{y+x}{2} + \frac{y-x}{2}\right) \\ &= \cos\left(\frac{y+x}{2}\right) \cdot \cos\left(\frac{y-x}{2}\right) - \sin\left(\frac{y+x}{2}\right) \cdot \sin\left(\frac{y-x}{2}\right) \end{aligned}$$

und

$$\begin{aligned}\cos(x) &= \cos\left(\frac{y+x}{2} - \frac{y-x}{2}\right) \\ &= \cos\left(\frac{y+x}{2}\right) \cdot \cos\left(-\frac{y-x}{2}\right) - \sin\left(\frac{y+x}{2}\right) \cdot \sin\left(-\frac{y-x}{2}\right) \\ &= \cos\left(\frac{y+x}{2}\right) \cdot \cos\left(\frac{y-x}{2}\right) + \sin\left(\frac{y+x}{2}\right) \cdot \sin\left(\frac{y-x}{2}\right).\end{aligned}$$

Subtrahieren wir die beiden Gleichungen voneinander, so erhalten wir

$$\cos(x) - \cos(y) = 2 \cdot \sin\left(\frac{y+x}{2}\right) \cdot \sin\left(\frac{y-x}{2}\right) \stackrel{16.10}{>} 0,$$

da mit $x, y \in [0, \pi]$ und $x < y$ auch

$$0 < \frac{y+x}{2} < \pi$$

und

$$0 < \frac{y-x}{2} < \pi$$

gelten muß.

Somit haben wir gezeigt, daß der Cosinus auf dem Intervall $[0, \pi]$ streng monoton fallend ist. Aus dem Umkehrsatz 14.21 folgt damit, daß

$$\cos : [0, \pi] \longrightarrow \left[\lim_{x \rightarrow \pi} \cos(x), \lim_{x \rightarrow 0} \cos(x) \right]$$

bijektiv ist. Da der Cosinus stetig ist, gilt nun

$$\lim_{x \rightarrow 0} \cos(x) = \cos(0) = \sum_{n=0}^{\infty} (-1)^n \cdot \frac{0^{2n}}{(2n)!} = 1$$

und

$$\lim_{x \rightarrow \pi} \cos(x) = \cos(\pi).$$

Aus Satz 12.38 wissen wir zudem, daß

$$\cos(\pi)^2 = 1 - \sin(\pi)^2 = 1 - 0 = 1$$

gilt, so daß $\cos(\pi) \in \{1, -1\}$. Da der Cosinus auf dem Intervall $[0, \pi]$ streng monoton fallend mit $\cos(0) = 1$ ist, muß somit $\cos(\pi) = -1$ gelten. \square

Satz 16.13 (Eigenschaften des Sinus und des Cosinus)

a. Für $x \in \mathbb{R}$ gelten

$$\sin(x + \pi) = -\sin(x) \quad \text{und} \quad \cos(x + \pi) = -\cos(x).$$

b. Für $x \in \mathbb{R}$ gelten zudem

$$\sin(x) \in [-1, 1] \quad \text{und} \quad \cos(x) \in [-1, 1].$$

c. Wir können folgende Werte des Sinus und Cosinus explizit angeben:

x	0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π
$\sin(x)$	0	$\frac{1}{\sqrt{2}}$	1	0	-1	0
$\cos(x)$	1	$\frac{1}{\sqrt{2}}$	0	-1	0	1

d. Sinus und Cosinus sind 2π -periodisch, d.h. für $x \in \mathbb{R}$ gelten

$$\sin(x + 2\pi) = \sin(x) \quad \text{und} \quad \cos(x + 2\pi) = \cos(x).$$

e. Die Perioden von Sinus und Cosinus sind um $\frac{\pi}{2}$ verschoben, d.h. für $x \in \mathbb{R}$ gelten

$$\sin\left(x + \frac{\pi}{2}\right) = \cos(x)$$

und

$$\cos\left(x - \frac{\pi}{2}\right) = \sin(x).$$

f. Der Sinus

$$\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \longrightarrow [-1, 1]$$

ist streng monoton wachsend und bijektiv.

g. Die Nullstellen des Sinus sind genau die ganzzahligen Vielfachen von π , d.h.

$$\sin(x) = 0 \iff x \in \{k \cdot \pi \mid k \in \mathbb{Z}\},$$

und für den Cosinus gilt mithin

$$\cos(x) = 0 \iff x \in \left\{ \frac{\pi}{2} + k \cdot \pi \mid k \in \mathbb{Z} \right\}.$$

Beweis:

a. Aus den Additionstheoremen 12.38 erhalten wir

$$\begin{aligned} \sin(x + \pi) &= \sin(x) \cdot \cos(\pi) + \sin(\pi) \cdot \cos(x) \\ &= \sin(x) \cdot (-1) + 0 \cdot \cos(x) = -\sin(x) \end{aligned}$$

und

$$\begin{aligned} \cos(x + \pi) &= \cos(x) \cdot \cos(\pi) - \sin(x) \cdot \sin(\pi) \\ &= \cos(x) \cdot (-1) - \sin(x) \cdot 0 = -\cos(x). \end{aligned}$$

b. Nach Satz 12.38 gilt

$$|\sin(x)| \leq \sqrt{\cos(x)^2 + \sin(x)^2} = 1$$

und

$$|\cos(x)| \leq \sqrt{\cos(x)^2 + \sin(x)^2} = 1.$$

- c. Die Werte für $x = 0$ folgen unmittelbar aus der Definition von Sinus und Cosinus als Potenzreihen

$$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!} = 0$$

und

$$\cos(x) = \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n}}{(2n)!} = 1.$$

Die Werte für $x = \pi$ folgen aus Satz 16.10 und Satz 16.12 oder alternativ aus Teil a.. Mit Hilfe der Additionstheoreme 12.38 folgt dann

$$-1 = \cos(\pi) = \cos\left(\frac{\pi}{2} + \frac{\pi}{2}\right) = \cos\left(\frac{\pi}{2}\right)^2 - \sin\left(\frac{\pi}{2}\right)^2$$

und somit

$$0 \leq \cos\left(\frac{\pi}{2}\right)^2 = \sin\left(\frac{\pi}{2}\right)^2 - 1 \stackrel{\text{b.}}{\leq} 0.$$

Damit müssen notwendigerweise

$$\cos\left(\frac{\pi}{2}\right) = 0$$

und

$$\sin\left(\frac{\pi}{2}\right) \in \{-1, 1\}$$

gelten. Da wir aber aus Satz 16.10 wissen, daß der Sinus auf dem Intervall $(0, \pi)$ strikt positiv ist, folgt

$$\sin\left(\frac{\pi}{2}\right) = 1.$$

Aus dem Additionstheorem für den Cosinus erhalten wir dann

$$0 = \cos\left(\frac{\pi}{2}\right) = \cos\left(\frac{\pi}{4} + \frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right)^2 - \sin\left(\frac{\pi}{4}\right)^2$$

und damit

$$\cos\left(\frac{\pi}{4}\right)^2 = \sin\left(\frac{\pi}{4}\right)^2.$$

Aus Satz 12.38 wissen wir zudem

$$1 = \cos\left(\frac{\pi}{4}\right)^2 + \sin\left(\frac{\pi}{4}\right)^2 = 2 \cdot \cos\left(\frac{\pi}{4}\right)^2,$$

und damit

$$\cos\left(\frac{\pi}{4}\right), \sin\left(\frac{\pi}{4}\right) \in \left\{-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right\}.$$

Nach Satz 16.12 ist der Cosinus auf $[0, \pi]$ streng monoton fallend mit Nullstelle bei $\frac{\pi}{2}$, also muß $\cos\left(\frac{\pi}{4}\right)$ positiv sein, und aus Satz 16.10 wissen wir, daß auch $\sin\left(\frac{\pi}{4}\right)$ positiv ist. Mithin gilt

$$\sin\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}.$$

Die übrigen Werte folgen, indem wir Teil a. auf die bisherigen Ergebnisse anwenden.

d. Durch Anwenden der Additionstheoreme 12.38 erhalten wir wie in Teil a.

$$\begin{aligned}\sin(x + 2\pi) &= \sin(x) \cdot \cos(2\pi) + \sin(2\pi) \cdot \cos(x) \\ &= \sin(x) \cdot 1 + 0 \cdot \cos(x) = \sin(x)\end{aligned}$$

und

$$\begin{aligned}\cos(x + 2\pi) &= \cos(x) \cdot \cos(2\pi) - \sin(x) \cdot \sin(2\pi) \\ &= \cos(x) \cdot 1 - \sin(x) \cdot 0 = \cos(x).\end{aligned}$$

e. Auch diese Aussage folgt aus den Additionstheoremen 12.38 wie in Teil a.:

$$\begin{aligned}\sin\left(x + \frac{\pi}{2}\right) &= \sin(x) \cdot \cos\left(\frac{\pi}{2}\right) + \sin\left(\frac{\pi}{2}\right) \cdot \cos(x) \\ &= \sin(x) \cdot 0 + 1 \cdot \cos(x) = \cos(x)\end{aligned}$$

und

$$\begin{aligned}\cos\left(x - \frac{\pi}{2}\right) &= \cos(x) \cdot \cos\left(-\frac{\pi}{2}\right) - \sin(x) \cdot \sin\left(-\frac{\pi}{2}\right) \\ &= \cos(x) \cdot \cos\left(\frac{\pi}{2}\right) + \sin(x) \cdot \sin\left(\frac{\pi}{2}\right) \\ &= \cos(x) \cdot 0 + \sin(x) \cdot 1 = \sin(x).\end{aligned}$$

f. Aus Teil a. und e. folgt

$$\sin(x) = \cos\left(x - \frac{\pi}{2}\right) = -\cos\left(x + \frac{\pi}{2}\right),$$

so daß die Aussage aus Satz 16.12 folgt.

g. Aus Teil a. und $\sin(\pi) = 0$ folgt mit Induktion, daß $\sin(k \cdot \pi) = 0$ für alle $k \in \mathbb{Z}$. Ist umgekehrt $x \in \mathbb{R}$ eine Nullstelle des Sinus, so gibt es eine ganze Zahl $k \in \mathbb{Z}$, so daß

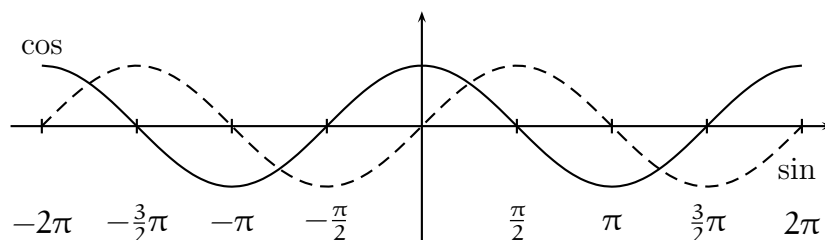
$$0 \leq x - k \cdot \pi < \pi,$$

und da der Sinus im Intervall $(0, \pi)$ keine Nullstelle besitzt, muß mithin $x = k \cdot \pi$ gelten. Aus Teil e. folgt dann die Aussage für die Nullstellen des Cosinus.

□

Bemerkung 16.14

Aus Satz 16.13 können wir den Verlauf der Graphen des Sinus und des Cosinus im wesentlichen herleiten:



Bemerkung 16.15 (Polarkoordinaten)

Ist $x \in \mathbb{R}$ und betrachten wir ein rechtwinkliges Dreieck mit den Kathetenlängen $\sin(x)$ und $\cos(x)$, so folgt wegen

$$\cos(x)^2 + \sin(x)^2 = 1 = 1^2$$

aus dem Satz von Pythagoras, daß die Hypotenuse die Seitenlänge 1 besitzt. D.h. der Punkt

$$(\cos(x), \sin(x)) = \cos(x) + i \cdot \sin(x) = \exp(i \cdot x) \in \mathbb{C}$$

liegt auf dem Einheitskreis und wir nennen x den Winkel im Bogenmaß, den der Strahl vom Ursprung durch diesen Punkt mit der x -Achse einschließt (siehe Abbildung 7).

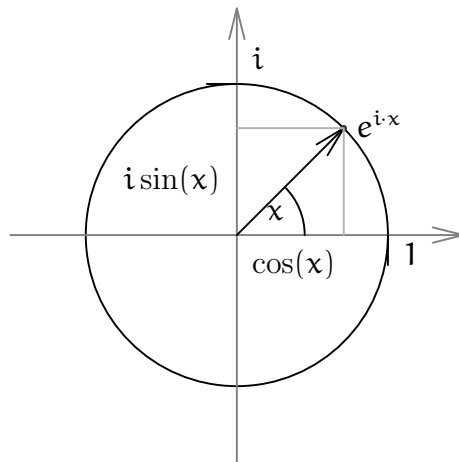


ABBILDUNG 7. Polarkoordinaten von $e^{i \cdot x} = \cos(x) + i \cdot \sin(x)$

Ist umgekehrt $z = a + ib = (a, b)$ ein Punkt auf dem Einheitskreis, so folgt aus $1 = a^2 + b^2$ sofort, daß $a \in [-1, 1]$ liegt. Da der Cosinus bijektiv auf dem Intervall $[0, \pi]$ mit Bild $[-1, 1]$ ist, gibt es genau ein $x \in [0, \pi]$ mit $a = \cos(x)$, und es gilt

$$\sin(x)^2 = 1 - \cos(x)^2 = 1 - a^2 = b^2,$$

d.h. $b \in \{-\sin(x), \sin(x)\} = \{\sin(-x), \sin(x)\}$. Wegen $a = \cos(x) = \cos(-x)$ finden wir also ein $y \in [-\pi, \pi]$ mit

$$z = a + ib = \cos(y) + i \cdot \sin(y) = e^{i \cdot y},$$

d.h. jeder Punkt auf dem Einheitskreis hat die Gestalt $z = e^{i \cdot y}$ mit $y \in \mathbb{R}$. Genauer kann man sogar sagen, daß es genau ein solches $y \in [-\pi, \pi)$ gibt.

Wir haben damit die Behauptung aus Bemerkung 10.9 gezeigt, daß jede komplexe Zahl z sich schreiben läßt als

$$z = |z| \cdot e^{i \cdot \arg z},$$

und wir können $\arg(z)$ im Intervall $[-\pi, \pi)$ eindeutig wählen. Wir nennen diese Darstellung die *Polarkoordinatendarstellung* von z .

Außerdem haben wir damit auch Bemerkung 10.10 gezeigt, daß nämlich jede komplexe Zahl eine n -te Wurzel besitzt, da wir dazu nur die Polarkoordinatendarstellung von z benötigen haben.

Man beachte, daß für $n \geq 2$ die Zahlen

$$e^{\frac{2 \cdot k \cdot \pi \cdot i}{n}} \quad \text{mit} \quad k = 0, \dots, n-1$$

genau die n -ten Wurzeln aus 1 sind. Man nennt sie auch die n -ten *Einheitswurzeln*.

Daß sie in der Tat n -te Wurzeln von 1 sind, folgt unmittelbar aus

$$\left(e^{\frac{2 \cdot k \cdot \pi \cdot i}{n}} \right)^n = e^{2 \cdot k \cdot \pi \cdot i} = \cos(2 \cdot k \cdot \pi) + i \cdot \sin(2 \cdot k \cdot \pi) = 1.$$

Und daß es keine weiteren n -ten Wurzeln geben kann, folgt aus der Tatsache, daß jede n -te Wurzel eine Nullstelle des Polynoms $t^n - 1$ ist und dieses nach Bemerkung 13.12 höchstens n verschiedene Nullstellen besitzen kann.

Definition 16.16 (Tangens und Cotangens)

Die Funktion

$$\tan : \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k \cdot \pi \mid k \in \mathbb{Z} \right\} \longrightarrow \mathbb{R} : x \mapsto \frac{\sin(x)}{\cos(x)}$$

heißt *Tangens* und die Funktion

$$\cot : \mathbb{R} \setminus \{k \cdot \pi \mid k \in \mathbb{Z}\} \longrightarrow \mathbb{R} : x \mapsto \frac{\cos(x)}{\sin(x)}$$

heißt *Cotangens*.

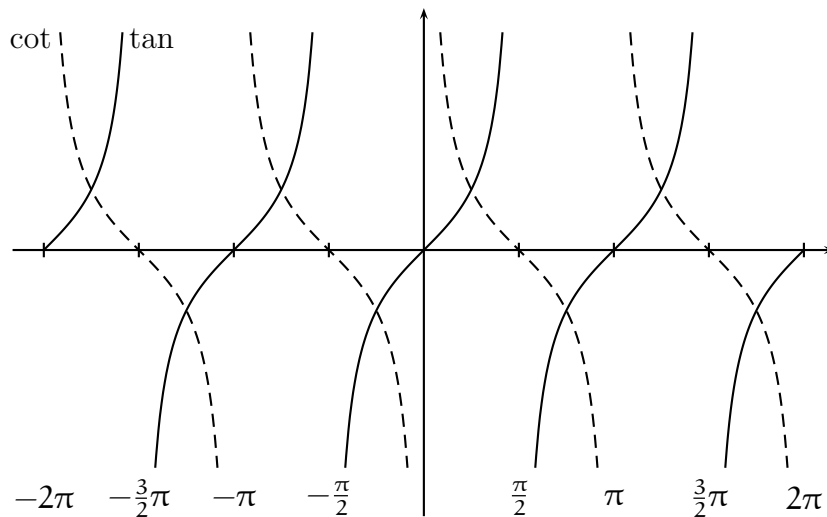


ABBILDUNG 8. Tangens und Cotangens

Satz 16.17 (Tangens und Cotangens)

a. Für $x \in \mathbb{R}$ gelten

$$\tan(-x) = -\tan(x) \quad \text{und} \quad \cot(-x) = -\cot(x)$$

und

$$\tan(x + \pi) = \tan(x) \quad \text{und} \quad \cot(x + \pi) = \cot(x).$$

- b. Der Tangens ist auf jedem der Intervalle $(-\frac{\pi}{2} + k \cdot \pi, \frac{\pi}{2} + k \cdot \pi)$, $k \in \mathbb{Z}$, streng monoton wachsend, stetig, bijektiv mit Bild \mathbb{R} und punktsymmetrisch zu seiner Nullstelle $k \cdot \pi$.
- c. Der Cotangens ist auf jedem der Intervalle $(k \cdot \pi, (k+1) \cdot \pi)$, $k \in \mathbb{Z}$, streng monoton fallend, stetig, bijektiv mit Bild \mathbb{R} und punktsymmetrisch zu seiner Nullstelle $\frac{2k+1}{2} \cdot \pi$.

Beweis:

- a. Die Aussagen folgen unmittelbar aus den entsprechenden Aussagen für den Sinus und Cosinus in Satz 12.38 und Satz 16.13.
- b. Für $0 \leq x < y < \frac{\pi}{2}$ folgt aus der Monotonie des Cosinus (Satz 16.12) und des Sinus (Satz 16.13)

$$\cos(x) > \cos(y) > 0$$

und

$$0 < \sin(x) < \sin(y),$$

so daß mithin

$$\tan(x) = \frac{\sin(x)}{\cos(x)} < \frac{\sin(y)}{\cos(y)} = \tan(y).$$

Der Tangens ist auf dem Intervall $[0, \frac{\pi}{2})$ also streng monoton wachsend. Wegen $\tan(-x) = -\tan(x)$ ist der Tangens aber punktsymmetrisch zum Ursprung und somit auch streng monoton wachsend auf $(\frac{\pi}{2}, 0]$, also streng monoton wachsend auf $(\frac{\pi}{2}, \frac{\pi}{2})$.

Wegen der Stetigkeit von Sinus und Cosinus erhalten wir zudem

$$\lim_{x \rightarrow \frac{\pi}{2}} \tan(x) = \frac{\lim_{x \rightarrow \frac{\pi}{2}} \sin(x)}{\lim_{x \rightarrow \frac{\pi}{2}} \cos(x)} = \frac{1}{0} = \infty$$

und

$$\lim_{x \rightarrow -\frac{\pi}{2}} \tan(x) = \frac{\lim_{x \rightarrow -\frac{\pi}{2}} \sin(x)}{\lim_{x \rightarrow -\frac{\pi}{2}} \cos(x)} = \frac{-1}{0} = -\infty.$$

Aus dem Umkehrsatz 14.21 folgt dann, daß die Funktion

$$\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \longrightarrow (-\infty, \infty) = \mathbb{R}$$

stetig und bijektiv ist. Aus Teil a. folgt die Aussage zur Punktsymmetrie, und zudem die Aussage für die verschobenen Intervalle.

- c. Die Aussage wird analog zur Aussage in Teil b. bewiesen.

□

Satz 16.18 (Arcusfunktionen)

- a. Die Funktion $\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \longrightarrow [-1, 1]$ besitzt eine stetige, streng monoton wachsende Umkehrabbildung, die wir Arcussinus nennen,

$$\arcsin : [-1, 1] \longrightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

- b. Die Funktion $\cos : [0, \pi] \longrightarrow [-1, 1]$ besitzt eine stetige, streng monoton fallende Umkehrabbildung, die wir Arcuscosinus nennen,

$$\arccos : [-1, 1] \longrightarrow [0, \pi].$$

- c. Die Funktion $\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \longrightarrow \mathbb{R}$ besitzt eine stetige, streng monoton wachsende Umkehrabbildung, die wir Arcustangens nennen,

$$\arctan : \mathbb{R} \longrightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

- d. Die Funktion $\cot : (0, \pi) \longrightarrow \mathbb{R}$ besitzt eine stetige, streng monoton fallende Umkehrabbildung, die wir Arcuscotangens nennen,

$$\operatorname{arccot} : \mathbb{R} \longrightarrow (0, \pi).$$

Beweis: Die Aussagen folgen aus dem Umkehrsatz 14.21 zusammen mit den Monotonieaussagen in den Sätzen 16.12, 16.13 und 16.17. \square

Aufgaben**Aufgabe 16.19**

Sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ stetig mit $f(x + y) = f(x) \cdot f(y)$ für alle $x, y \in \mathbb{R}$ und $f(1) = a > 0$. Zeige, dann ist $f = \exp_a$.

Aufgabe 16.20 (Additionstheoreme für Tangens und Arcustangens)

- a. Zeige das folgende Additionstheorem für den Tangens:

$$\tan(x + y) = \frac{\tan(x) + \tan(y)}{1 - \tan(x) \cdot \tan(y)},$$

wobei $x, y, x + y \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ gelten soll.

- b. Folgere daraus das folgende Additionstheorem für den Arcustangens:

$$\arctan(x) + \arctan(y) = \arctan\left(\frac{x + y}{1 - xy}\right).$$

- c. Zeige die Gleichung

$$4 \cdot \arctan\left(\frac{1}{5}\right) - \arctan\left(\frac{1}{239}\right) = \frac{\pi}{4}.$$

§ 17 Differenzierbarkeit

A) Differenzierbarkeit

Definition 17.1 (Differenzenquotient)

Es sei $U \subseteq \mathbb{R}$, $f : U \rightarrow \mathbb{R}$ und $a \in U$. Die Funktion

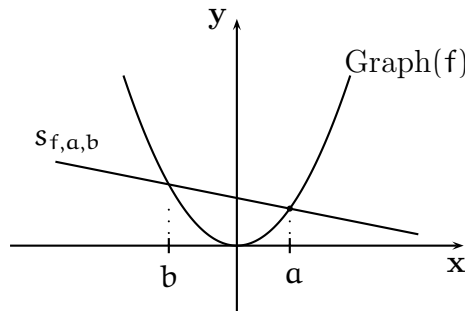
$$\text{Diff}_{f,a} : U \setminus \{a\} \rightarrow \mathbb{R} : x \mapsto \frac{f(x) - f(a)}{x - a}$$

heißt der *Differenzenquotient* von f an der Stelle a .

Für ein festes b ist der Wert des Differenzenquotienten $\text{Diff}_{f,a}(b)$ die *Steigung* der *Sekante* $s_{f,a,b}$ an den Graphen von f durch die Punkte $(b, f(b))$ und $(a, f(a))$, deren Geradengleichung durch

$$y = \frac{f(b) - f(a)}{b - a} \cdot x + \frac{f(a) \cdot b - f(b) \cdot a}{b - a}$$

gegeben ist.

**Beispiel 17.2**

Ist $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^n$ mit $n \geq 1$, so ist

$$\text{Diff}_{f,a}(x) = \frac{x^n - a^n}{x - a} = x^{n-1} + a \cdot x^{n-2} + a^2 \cdot x^{n-3} + \dots + a^{n-2} \cdot x + a^{n-1}$$

für $x \in \mathbb{R} \setminus \{a\}$.

Definition 17.3

Es sei $U \subseteq \mathbb{R}$, $f : U \rightarrow \mathbb{R}$ eine Funktion und $a \in U$. Wir nennen f *differenzierbar in a* , wenn a ein Häufungspunkt von U ist und der Grenzwert

$$\lim_{x \rightarrow a} \text{Diff}_{f,a}(x) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \in \mathbb{R}$$

des Differenzenquotienten in a existiert. In diesem Fall schreiben wir

$$f'(a) := \frac{\partial f}{\partial x}(a) := \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a},$$

und nennen diesen Grenzwert die *Ableitung* von f an der Stelle a .

Wir nennen die Funktion f *differenzierbar (auf U)*, wenn f in jedem Punkt von U differenzierbar ist. In diesem Fall nennen wir die Funktion

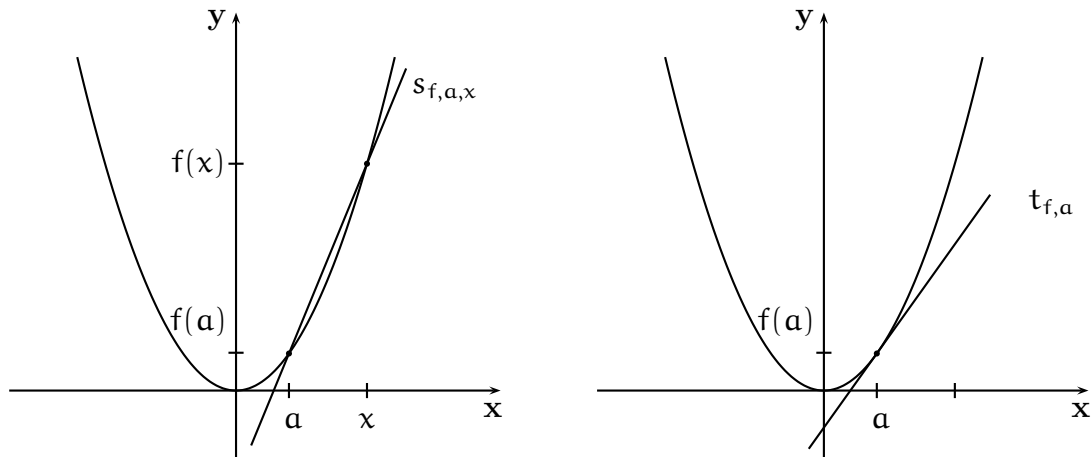
$$f' : U \rightarrow \mathbb{R} : x \mapsto f'(x)$$

die *Ableitung* von f . Beachte auch, daß dann insbesondere jeder Punkt von U ein Häufungspunkt von U sein muß!

Bemerkung 17.4

Der Definition liegt die Idee zugrunde, daß sich die Sekante $s_{f,a,x}$ für $x \rightarrow a$ einer Geraden annähert, die im Punkt $(a, f(a))$ den Graphen von f berührt und ihn optimal *linear approximiert*. Diese Gerade wollen wir die *Tangente* $t_{f,a}$ von f in a nennen, und der Grenzwert des Differenzenquotienten, d.h. die Steigung von $s_{f,a,x}$ konvergiert dann für $x \rightarrow a$ gegen die Steigung der Tangenten. D.h. die Tangente an den Graphen von f im Punkt $(a, f(a))$ hat die Geradengleichung

$$y = f'(a) \cdot x + (f(a) - a \cdot f'(a)) = f(a) + f'(a) \cdot (x - a).$$



Beispiel 17.5

Die folgenden Funktionen sind alle auf ganz \mathbb{R} definiert, und dort ist jeder Punkt ein Häufungspunkt!

- a. Eine konstante Funktion $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto c$ ist differenzierbar auf \mathbb{R} und die Ableitung ist die Nullfunktion

$$f' : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto 0,$$

da für jedes $a \in \mathbb{R}$ der Differenzenquotient $\text{Diff}_{f,a}$ die Nullfunktion ist und somit der Grenzwert $f'(a) = 0$ existiert.

- b. Für $n \in \mathbb{N}$ ist die Funktion

$$f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^n$$

differenzierbar auf \mathbb{R} mit Ableitung

$$f' : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto n \cdot x^{n-1},$$

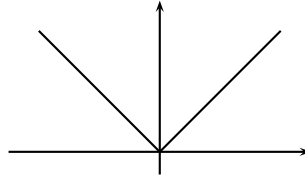
da sich für $a \in \mathbb{R}$ aus Beispiel 17.2 folgendes ergibt:

$$f'(a) = \lim_{x \rightarrow a} (x^{n-1} + a \cdot x^{n-2} + a^2 \cdot x^{n-3} + \dots + a^{n-2} \cdot x + a^{n-1}) = n \cdot a^{n-1}.$$

c. Die Betragsfunktion

$$|\cdot| : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto |x|$$

ist in $\mathbf{a} = 0$ *nicht* differenzierbar. In jedem anderen Punkt \mathbf{a} ist sie jedoch differenzierbar mit $f'(\mathbf{a}) = -1$ falls $\mathbf{a} < 0$ und $f'(\mathbf{a}) = 1$ falls $\mathbf{a} > 0$.



Anschaulich bedeutet die Nicht-Differenzierbarkeit im Punkt $\mathbf{a} = 0$, daß man am Graphen im Ursprung keine klare Tangente findet.

Um die Nicht-Differenzierbarkeit in $\mathbf{a} = 0$ zu sehen, betrachten wir die Nullfolge $\left(\frac{(-1)^n}{n}\right)_{n \geq 1}$. Die zugehörige Folge der Werte des Differenzenquotienten

$$\left(\text{Diff}_{f,0}\left(\frac{(-1)^n}{n}\right)\right)_{n \geq 1} = \left(\frac{\frac{1}{n}}{\frac{(-1)^n}{n}}\right)_{n \geq 1} = ((-1)^n)_{n \geq 1}$$

ist nicht konvergent. Mithin existiert der Grenzwert des Differenzenquotienten in $\mathbf{a} = 0$ nicht, und somit ist die Funktion in $\mathbf{a} = 0$ nicht differenzierbar.

Außerdem, ist $\mathbf{a} < 0$ und x nahe bei \mathbf{a} , so ist auch $x < 0$ und mithin

$$\text{Diff}_{f,\mathbf{a}}(x) = \frac{|x| - |\mathbf{a}|}{x - \mathbf{a}} = \frac{-x + \mathbf{a}}{x - \mathbf{a}} = -1 \xrightarrow{x \rightarrow \mathbf{a}} -1,$$

und analog ist für $\mathbf{a} > 0$ und x nahe bei \mathbf{a} auch $x > 0$, so daß

$$\text{Diff}_{f,\mathbf{a}}(x) = \frac{|x| - |\mathbf{a}|}{x - \mathbf{a}} = \frac{x - \mathbf{a}}{x - \mathbf{a}} = 1 \xrightarrow{x \rightarrow \mathbf{a}} 1.$$

Damit ist auch gezeigt, daß die Ableitung in allen Punkten $\mathbf{a} \neq 0$ existiert.

Bemerkung 17.6

- Wie bei der Stetigkeit wollen wir auch bei der Differenzierbarkeit anmerken, daß es sich um eine *lokale* Eigenschaft der Funktion handelt. D.h. sie ist punktweise definiert und hängt nur vom Verhalten der Funktion in einer sehr kleinen ε -Umgebung des betrachteten Punktes \mathbf{a} ab!
- Ist \mathbf{a} ein Häufungspunkt von \mathbf{U} und $f : \mathbf{U} \longrightarrow \mathbb{R}$, so ist f genau dann in \mathbf{a} differenzierbar, wenn der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(\mathbf{a} + h) - f(\mathbf{a})}{h} \in \mathbb{R}$$

existiert.

Um dies zu sehen ersetzt man im Differenzenquotienten einfach $x - \mathbf{a}$ durch h .

- Ist \mathbf{a} ein Häufungspunkt von \mathbf{U} und $f : \mathbf{U} \longrightarrow \mathbb{R}$, so ist f genau dann in \mathbf{a} differenzierbar, wenn es eine Zahl $c \in \mathbb{R}$ und eine Funktion $\rho : \mathbf{U} \longrightarrow \mathbb{R}$ gibt, so daß

$$f(x) = f(\mathbf{a}) + c \cdot (x - \mathbf{a}) + \rho(x) \quad \text{und} \quad \lim_{x \rightarrow \mathbf{a}} \frac{\rho(x)}{|x - \mathbf{a}|} = 0$$

gilt.

Man beachte, ist f differenzierbar in \mathbf{a} , so wählt man $\mathbf{c} = f'(\mathbf{a})$ und $\rho(\mathbf{x}) = (\text{Diff}_{f,\mathbf{a}}(\mathbf{x}) - f'(\mathbf{a})) \cdot (\mathbf{x} - \mathbf{a})$. Umgekehrt, wenn \mathbf{c} und ρ existieren, so ist $\text{Diff}_{f,\mathbf{a}}(\mathbf{x}) = \mathbf{c} + \frac{\rho(\mathbf{x})}{\mathbf{x} - \mathbf{a}}$ und der Grenzwert des Differenzenquotienten existiert nach Voraussetzung.

Wir erwähnen diese äquivalente Formulierung der Differenzierbarkeit an dieser Stelle, da sie für die Verallgemeinerung des Begriffes in der mehrdimensionalen Analysis von Vorteil ist. Die Bedeutung der Bedingung $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho(\mathbf{x})}{|\mathbf{x} - \mathbf{a}|} = 0$ ist, daß die Funktion ρ , die den Unterschied des Differenzenquotienten und der Ableitung beschreibt, sehr schnell gegen Null konvergiert für $\mathbf{x} \rightarrow \mathbf{a}$, jedenfalls schneller als die lineare Funktion $\mathbf{x} - \mathbf{a}$.

Satz 17.7 (Differenzierbar impliziert stetig.)

Ist $f : \mathbf{U} \rightarrow \mathbb{R}$ differenzierbar in \mathbf{a} , so ist f stetig in \mathbf{a} .

Beweis: Da nach Voraussetzung \mathbf{a} ein Häufungspunkt von \mathbf{U} ist, müssen wir nach Lemma 14.3 nur zeigen, daß $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) = f(\mathbf{a})$ oder alternativ $\lim_{\mathbf{x} \rightarrow \mathbf{a}} (f(\mathbf{x}) - f(\mathbf{a})) = 0$ gilt. Nach Voraussetzung existiert der Grenzwert

$$f'(\mathbf{a}) = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a})}{\mathbf{x} - \mathbf{a}} \in \mathbb{R},$$

und da die Identität stetig ist, gilt zudem $\lim_{\mathbf{x} \rightarrow \mathbf{a}} (\mathbf{x} - \mathbf{a}) = 0$. Mithin erhalten wir aus den Grenzwertsätzen für Funktionen 13.10

$$\begin{aligned} \lim_{\mathbf{x} \rightarrow \mathbf{a}} (f(\mathbf{x}) - f(\mathbf{a})) &= \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a})}{\mathbf{x} - \mathbf{a}} \cdot (\mathbf{x} - \mathbf{a}) \\ &= \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a})}{\mathbf{x} - \mathbf{a}} \cdot \lim_{\mathbf{x} \rightarrow \mathbf{a}} (\mathbf{x} - \mathbf{a}) = f'(\mathbf{a}) \cdot 0 = 0. \end{aligned}$$

Also ist f stetig in \mathbf{a} . □

Beispiel 17.8

Die Umkehrung von Satz 17.7 gilt nicht, wie das Beispiel der Betragsfunktion zeigt, die stetig in $\mathbf{a} = 0$ ist (siehe Beispiel 14.6), ohne dort differenzierbar zu sein (siehe Beispiel 17.5).

B) Ableitungsregeln

Proposition 17.9 (Linearität der Ableitung)

Seien $f : \mathbf{U} \rightarrow \mathbb{R}$ und $g : \mathbf{U} \rightarrow \mathbb{R}$ in $\mathbf{a} \in \mathbf{U}$ differenzierbar und sei $\mathbf{c}, \mathbf{d} \in \mathbb{R}$.

Dann ist $\mathbf{c} \cdot f + \mathbf{d} \cdot g$ differenzierbar in \mathbf{a} mit $(\mathbf{c} \cdot f + \mathbf{d} \cdot g)'(\mathbf{a}) = \mathbf{c} \cdot f'(\mathbf{a}) + \mathbf{d} \cdot g'(\mathbf{a})$.

Beweis: Wir beachten zunächst, daß nach Voraussetzung \mathbf{a} ein Häufungspunkt von \mathbf{U} ist und daß \mathbf{U} jeweils der Definitionsbereich der Funktionen ist. Dann folgt aus

den Grenzwertsätzen für Funktionen 13.10, daß der Grenzwert

$$\begin{aligned} (c \cdot f + d \cdot g)'(\mathbf{a}) &= \lim_{x \rightarrow \mathbf{a}} \frac{c \cdot f(x) + d \cdot g(x) - c \cdot f(\mathbf{a}) - d \cdot g(\mathbf{a})}{x - \mathbf{a}} \\ &= c \cdot \lim_{x \rightarrow \mathbf{a}} \frac{f(x) - f(\mathbf{a})}{x - \mathbf{a}} + d \cdot \lim_{x \rightarrow \mathbf{a}} \frac{g(x) - g(\mathbf{a})}{x - \mathbf{a}} \\ &= c \cdot f'(\mathbf{a}) + d \cdot g'(\mathbf{a}) \end{aligned}$$

existiert. □

Beispiel 17.10 (Polynomfunktionen sind differenzierbar.)

Ist $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \sum_{k=0}^n a_k \cdot x^k$ eine Polynomfunktion, so ist f differenzierbar auf \mathbb{R} mit

$$f' : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \sum_{k=1}^n k \cdot a_k \cdot x^{k-1}.$$

Dies folgt unmittelbar aus Proposition 17.9 und Beispiel 17.5.

Proposition 17.11 (Produktregel)

Seien $f : \mathbf{U} \rightarrow \mathbb{R}$ und $g : \mathbf{U} \rightarrow \mathbb{R}$ in $\mathbf{a} \in \mathbf{U}$ differenzierbar, so ist $f \cdot g$ differenzierbar in \mathbf{a} mit

$$(f \cdot g)'(\mathbf{a}) = f'(\mathbf{a}) \cdot g(\mathbf{a}) + f(\mathbf{a}) \cdot g'(\mathbf{a}).$$

Beweis: Wir beachten, daß nach Voraussetzung \mathbf{a} ein Häufungspunkt von \mathbf{U} ist und daß \mathbf{U} der Definitionsbereich von $f \cdot g$ ist. Der Differenzenquotient von $f \cdot g$ an der Stelle \mathbf{a} genügt der Gleichung

$$\begin{aligned} \text{Diff}_{f \cdot g, \mathbf{a}}(x) &= \frac{f(x) \cdot g(x) - f(\mathbf{a}) \cdot g(\mathbf{a})}{x - \mathbf{a}} \\ &= \frac{f(x) \cdot g(x) - f(\mathbf{a}) \cdot g(x) + f(\mathbf{a}) \cdot g(x) - f(\mathbf{a}) \cdot g(\mathbf{a})}{x - \mathbf{a}} \\ &= \frac{f(x) \cdot g(x) - f(\mathbf{a}) \cdot g(x)}{x - \mathbf{a}} + \frac{f(\mathbf{a}) \cdot g(x) - f(\mathbf{a}) \cdot g(\mathbf{a})}{x - \mathbf{a}} \\ &= \frac{f(x) - f(\mathbf{a})}{x - \mathbf{a}} \cdot g(x) + f(\mathbf{a}) \cdot \frac{g(x) - g(\mathbf{a})}{x - \mathbf{a}}. \end{aligned}$$

Da f und g in \mathbf{a} differenzierbar sind und da g nach Satz 17.7 zudem stetig in \mathbf{a} ist, existiert damit der Grenzwert

$$(f \cdot g)'(\mathbf{a}) = \lim_{x \rightarrow \mathbf{a}} \text{Diff}_{f \cdot g, \mathbf{a}}(x) = f'(\mathbf{a}) \cdot g(\mathbf{a}) + f(\mathbf{a}) \cdot g'(\mathbf{a})$$

aufgrund der Grenzwertsätze für Funktionen 13.10. □

Proposition 17.12 (Quotientenregel)

Seien $f : \mathbf{U} \rightarrow \mathbb{R}$ und $g : \mathbf{U} \rightarrow \mathbb{R}$ in $\mathbf{a} \in \mathbf{U}$ differenzierbar mit $g(\mathbf{a}) \neq 0$, so ist auch $\frac{f}{g} : \{x \in \mathbf{U} \mid g(x) \neq 0\} \rightarrow \mathbb{R}$ differenzierbar in \mathbf{a} mit

$$\left(\frac{f}{g}\right)'(\mathbf{a}) = \frac{f'(\mathbf{a}) \cdot g(\mathbf{a}) - f(\mathbf{a}) \cdot g'(\mathbf{a})}{g(\mathbf{a})^2}.$$

Beweis: Wir müssen zunächst einmal zeigen, daß \mathbf{a} ein Häufungspunkt der Menge

$$V := \{x \in U \mid g(x) \neq 0\}$$

ist. Wegen Satz 17.7 ist g stetig in \mathbf{a} , und daher folgt aus Aufgabe 14.30 mit $g(\mathbf{a}) \neq 0$, daß es ein $\delta > 0$ gibt mit

$$g(x) \neq 0 \quad \forall x \in U \cap (\mathbf{a} - \delta, \mathbf{a} + \delta),$$

d.h.

$$U \cap (\mathbf{a} - \delta, \mathbf{a} + \delta) \subseteq V.$$

Da \mathbf{a} ein Häufungspunkt von U ist, gibt es eine Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in $U \setminus \{\mathbf{a}\}$, die gegen \mathbf{a} konvergiert, und somit gibt es ein $n_\delta \in \mathbb{N}$, so daß für $n \geq n_\delta$

$$\mathbf{a}_n \in (\mathbf{a} - \delta, \mathbf{a} + \delta) \cap U \subseteq V$$

Damit haben wir eine Folge in $V \setminus \{\mathbf{a}\}$ gefunden, die gegen \mathbf{a} konvergiert, und somit ist \mathbf{a} ein Häufungspunkt der Menge V .

Ferner gilt für den Differenzenquotienten

$$\text{Diff}_{\frac{1}{g}, \mathbf{a}}(x) = \frac{\frac{1}{g(x)} - \frac{1}{g(\mathbf{a})}}{x - \mathbf{a}} = -\frac{g(x) - g(\mathbf{a})}{x - \mathbf{a}} \cdot \frac{1}{g(x) \cdot g(\mathbf{a})}.$$

Da g in \mathbf{a} differenzierbar und stetig ist, existiert damit der Grenzwert

$$\left(\frac{1}{g}\right)'(\mathbf{a}) = \lim_{x \rightarrow \mathbf{a}} \text{Diff}_{\frac{1}{g}, \mathbf{a}}(x) = -\lim_{x \rightarrow \mathbf{a}} \frac{g(x) - g(\mathbf{a})}{x - \mathbf{a}} \cdot \lim_{x \rightarrow \mathbf{a}} \frac{1}{g(x) \cdot g(\mathbf{a})} = -\frac{g'(\mathbf{a})}{g(\mathbf{a})^2}.$$

Wenden wir nun die Produktformel 17.11 auf $f \cdot \frac{1}{g}$ an, so folgt das Ergebnis. \square

Aus der Quotientenregel und Beispiel 17.10 folgt die folgende Aussage.

Beispiel 17.13

Jede rationale Funktion $\frac{f}{g} : \mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) = 0\} \rightarrow \mathbb{R}$ ist differenzierbar.

Z.B. gilt für $h : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R} : x \mapsto \frac{1}{x^n}$ mit $n \geq 1$ für die Ableitung

$$h' : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R} : x \mapsto -\frac{n \cdot x^{n-1}}{x^{2n}} = -\frac{n}{x^{n+1}}.$$

\square

Aufgrund des Umkehrsatzes für streng monotone Funktionen 14.21 sowie Bemerkung 14.22 wissen wir, daß eine stetige und streng monotone Funktion auf einem Intervall eine stetige Umkehrfunktion besitzt. Dabei kann das Intervall offen, halboffen oder abgeschlossen sein und es kann auch ein uneigentliches Intervall sein. Wir wenden uns nun der Frage zu, ob die Umkehrfunktion differenzierbar ist, wenn f differenzierbar ist.

Satz 17.14 (Ableitung der Umkehrfunktion)

Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ sei stetig und streng monoton (wachsend oder fallend). Ist f differenzierbar in \mathbf{a} und ist $f'(\mathbf{a}) \neq 0$, so ist die Umkehrfunktion

$$f^{-1} : f(I) \rightarrow I$$

differenzierbar in $\mathbf{b} := f(\mathbf{a})$ und es gilt

$$(f^{-1})'(\mathbf{b}) = \frac{1}{f'(\mathbf{a})} = \frac{1}{f'(f^{-1}(\mathbf{b}))}.$$

Beweis: Aus dem Umkehrsatz 14.21 sowie Bemerkung 14.22 wissen wir, daß die Umkehrfunktion

$$f^{-1} : f(I) \longrightarrow I$$

existiert und daß sie stetig und bijektiv ist.

Wir betrachten nun eine Folge $(\mathbf{y}_n)_{n \in \mathbb{N}}$ in $f(I) \setminus \{\mathbf{b}\}$ mit $\mathbf{y}_n \longrightarrow \mathbf{b}$. Da f^{-1} stetig ist, gilt dann

$$\mathbf{x}_n := f^{-1}(\mathbf{y}_n) \longrightarrow f^{-1}(\mathbf{b}) = \mathbf{a}.$$

Da f^{-1} bijektiv ist, ist $(\mathbf{x}_n)_{n \in \mathbb{N}}$ somit eine Folge in $I \setminus \{\mathbf{a}\}$, die gegen \mathbf{a} konvergiert. Aufgrund der Grenzwertsätze für Folgen 11.15 erhalten wir dann

$$\frac{f^{-1}(\mathbf{y}_n) - f^{-1}(\mathbf{b})}{\mathbf{y}_n - \mathbf{b}} = \frac{\mathbf{x}_n - \mathbf{a}}{f(\mathbf{x}_n) - f(\mathbf{a})} = \frac{1}{\text{Diff}_{f,\mathbf{a}}(\mathbf{x}_n)} \longrightarrow \frac{1}{f'(\mathbf{a})},$$

und wegen des Folgenkriteriums für Grenzwerte 13.7 existiert somit der Grenzwert

$$(f^{-1})'(\mathbf{b}) = \lim_{\mathbf{y} \rightarrow \mathbf{b}} \text{Diff}_{f^{-1},\mathbf{b}}(\mathbf{y}) = \lim_{\mathbf{y} \rightarrow \mathbf{b}} \frac{f^{-1}(\mathbf{y}) - f^{-1}(\mathbf{b})}{\mathbf{y} - \mathbf{b}} = \frac{1}{f'(\mathbf{a})}.$$

□

Beispiel 17.15

Für $n \geq 2$ ist die Funktion

$$f : [0, \infty) \longrightarrow [0, \infty) : x \mapsto x^n$$

streng monoton wachsend und stetig nach Beispiel 14.23 mit der Wurzelfunktion als Umkehrfunktion und mit der Ableitung

$$f' : [0, \infty) \longrightarrow [0, \infty) : x \mapsto n \cdot x^{n-1}.$$

Da $f'(x) \neq 0$ für $x \neq 0$, folgt aus Satz 17.14, daß die Wurzelfunktion

$$\sqrt[n]{\cdot} : [0, \infty) \longrightarrow [0, \infty) : y \mapsto y^{\frac{1}{n}}$$

auf dem Intervall $(0, \infty)$ differenzierbar ist mit Ableitung

$$(\sqrt[n]{\cdot})' : (0, \infty) \rightarrow \mathbb{R} : y \mapsto \frac{1}{n \cdot (\sqrt[n]{y})^{n-1}} = \frac{1}{n} \cdot y^{\frac{1}{n}-1}.$$

Im Falle von $n = 2$ erhalten wir insbesondere

$$(\sqrt{y})' = \frac{1}{2 \cdot \sqrt{y}}.$$

Wir wollen nun noch zeigen, daß die Wurzelfunktion $\sqrt[n]{\cdot}$ in $\mathbf{a} = 0$ in der Tat *nicht* differenzierbar ist!

Dazu betrachten wir die Nullfolge $\left(\frac{1}{k^n}\right)_{k \in \mathbb{N}}$ und die zugehörigen Werte des Differenzenquotienten

$$\text{Diff}_{\mathbb{V},0} \left(\frac{1}{k^n} \right) = \frac{\frac{1}{k} - 0}{\frac{1}{k^n} - 0} = k^{n-1} \longrightarrow \infty$$

für $k \longrightarrow \infty$, da $n \geq 2$. Also existiert der Grenzwert des Differenzenquotienten in $\mathbf{a} = 0$ nicht, und somit ist die Wurzelfunktion dort auch nicht differenzierbar.

Proposition 17.16 (Kettenregel – äußere Ableitung \times innere Ableitung)

Es seien $f : \mathbf{U} \longrightarrow \mathbb{R}$ und $g : \mathbf{V} \longrightarrow \mathbb{R}$ Funktionen mit $\text{Im}(f) \subseteq \mathbf{V}$ und es sei $\mathbf{a} \in \mathbf{U}$. Ist f differenzierbar in \mathbf{a} und g differenzierbar in $f(\mathbf{a})$, so ist $g \circ f$ differenzierbar in \mathbf{a} mit

$$(g \circ f)'(\mathbf{a}) = g'(f(\mathbf{a})) \cdot f'(\mathbf{a}).$$

Beweis: Wir definieren auf \mathbf{V} eine Funktion durch

$$h : \mathbf{V} \longrightarrow \mathbb{R} : \mathbf{y} \mapsto \begin{cases} \text{Diff}_{g,f(\mathbf{a})}(\mathbf{y}) = \frac{g(\mathbf{y}) - g(f(\mathbf{a}))}{\mathbf{y} - f(\mathbf{a})}, & \text{falls } \mathbf{y} \neq f(\mathbf{a}), \\ g'(f(\mathbf{a})), & \text{falls } \mathbf{y} = f(\mathbf{a}). \end{cases}$$

Da g in $f(\mathbf{a})$ differenzierbar ist, gilt dann $\lim_{\mathbf{y} \rightarrow f(\mathbf{a})} h(\mathbf{y}) = h(f(\mathbf{a}))$, d.h. h ist stetig in $f(\mathbf{a})$. Außerdem gilt für alle $\mathbf{y} \in \mathbf{V}$

$$h(\mathbf{y}) \cdot (\mathbf{y} - f(\mathbf{a})) = g(\mathbf{y}) - g(f(\mathbf{a})). \quad (23)$$

Wir beachten nun noch, daß nach Proposition 14.8 $\lim_{x \rightarrow \mathbf{a}} h(f(x)) = h(f(\mathbf{a})) = g'(f(\mathbf{a}))$ gilt, da die Funktion h stetig in $f(\mathbf{a})$ und die Funktion f nach Satz 17.7 stetig in \mathbf{a} ist. Damit erhalten wir dann, daß der Grenzwert

$$\begin{aligned} (g \circ f)'(\mathbf{a}) &= \lim_{x \rightarrow \mathbf{a}} \frac{g(f(x)) - g(f(\mathbf{a}))}{x - \mathbf{a}} \stackrel{(23)}{=} \lim_{x \rightarrow \mathbf{a}} \frac{h(f(x)) \cdot (f(x) - f(\mathbf{a}))}{x - \mathbf{a}} \\ &= \lim_{x \rightarrow \mathbf{a}} h(f(x)) \cdot \lim_{x \rightarrow \mathbf{a}} \frac{f(x) - f(\mathbf{a})}{x - \mathbf{a}} = g'(f(\mathbf{a})) \cdot f'(\mathbf{a}) \end{aligned}$$

existiert. □

Beispiel 17.17

Die Funktion

$$h : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sqrt{x^2 + 1}$$

läßt sich schreiben als $g \circ f$ mit

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2 + 1$$

und $g = \sqrt{\cdot}$. Also ist h differenzierbar auf \mathbb{R} mit Ableitung

$$h'(x) = g'(f(x)) \cdot f'(x) = \frac{1}{2 \cdot \sqrt{x^2 + 1}} \cdot 2x = \frac{x}{\sqrt{x^2 + 1}}$$

in x .

C) Stetige Differenzierbarkeit

Abschließend wollen wir in diesem Abschnitt noch einige Begriffe einführen, die im folgenden nützlich sein werden.

Definition 17.18

Es sei $U \subseteq \mathbb{R}$ und $f : U \rightarrow \mathbb{R}$ eine Funktion.

- Wir nennen f *stetig differenzierbar*, wenn f differenzierbar auf U und f' stetig auf U ist.
- Wir definieren die *k-fache Differenzierbarkeit* und die *k-te Ableitung* von f rekursiv. f heißt *1-fach differenzierbar* auf U , wenn f auf U differenzierbar ist, und $f^{(1)} := f'$ heißt die *erste Ableitung* von f . Für $k > 1$ heißt f *k-fach differenzierbar*, wenn $f^{(k-1)}$ differenzierbar ist, und $f^{(k)} := (f^{(k-1)})'$ heißt dann die *k-te Ableitung* von f . Wir schreiben auch $f^{(0)} := f$, $f'' := f^{(2)}$ und $f''' := f^{(3)}$.
- f heißt *k-fach stetig differenzierbar*, wenn f k-fach differenzierbar auf U und zudem $f^{(k)}$ stetig auf U ist. Mit

$$\mathcal{C}^k(U, \mathbb{R}) := \{f : U \rightarrow \mathbb{R} \mid f \text{ ist } k\text{-fach stetig differenzierbar}\}$$

bezeichnen wir die Menge der k-fach stetig differenzierbaren Funktionen auf U .

- f heißt *unendlich oft differenzierbar* auf U , wenn $f \in \mathcal{C}^k(U, \mathbb{R})$ für alle $k \geq 1$. Wir bezeichnen mit

$$\mathcal{C}^\infty(U, \mathbb{R}) := \{f : U \rightarrow \mathbb{R} \mid f \text{ ist unendlich oft differenzierbar}\}$$

die Menge der unendlich oft differenzierbaren Funktionen auf U .

Beispiel 17.19

- Die Funktion

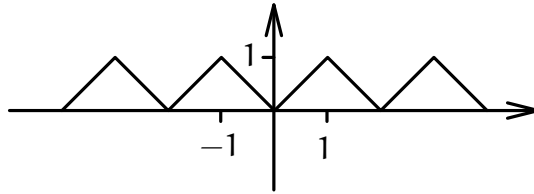
$$f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \begin{cases} x^2 \cdot \sin\left(\frac{1}{x}\right), & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0, \end{cases}$$

ist differenzierbar auf \mathbb{R} , die Ableitung ist aber nicht stetig in $a = 0$. Der Beweis ist eine Übungsaufgabe, für die man unter anderem Korollar 18.20 benötigt.

- Leitet man eine Polynomfunktion oder eine rationale Funktion ab, so erhält man wieder eine Polynomfunktion oder eine rationale Funktion mit dem jeweils gleichen Definitionsbereich. Da diese wieder differenzierbar sind, sehen wir, daß Polynomfunktionen und rationale Funktionen unendlich oft differenzierbar sind.

Bemerkung 17.20 (Überall stetig, nirgendwo differenzierbar)

Betrachte die periodische Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$, deren Graph in folgendem Bild dargestellt ist.



Die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} \frac{g(2^n \cdot x)}{2^n}$$

ist ein Beispiel für eine Funktion, die in jedem Punkt stetig und in keinem Punkt differenzierbar ist!

Aufgaben

Aufgabe 17.21

Sei $f : (0, 1] \longrightarrow \mathbb{R}$ eine differenzierbare Funktion mit beschränkter Ableitung. Zeige, dass f dann stetig in 0 fortsetzbar ist.

Aufgabe 17.22

Berechne die Ableitungen der folgenden Funktionen $f : (a, \infty) \longrightarrow \mathbb{R}$.

- $f(x) = \ln\left(\frac{\ln(x)}{x}\right)$ mit $a = 1$.
- $f(x) = \frac{x^2+4}{x-4}$ mit $a = 4$.
- $f(x) = \sqrt{e^{\cos(\sqrt{x})}}$ mit $a = 0$.

Aufgabe 17.23

Für $n \in \{0, 1, 2\}$ sei

$$f_n : [0, \infty) \longrightarrow \mathbb{R}, x \mapsto \begin{cases} x^n \cdot \sin\left(\frac{1}{x}\right), & \text{für } x > 0, \\ 0, & \text{für } x = 0. \end{cases}$$

Welche der Funktionen sind stetig in $a = 0$, differenzierbar in $a = 0$, stetig differenzierbar auf $[0, \infty)$?

Aufgabe 17.24 (Leibnitz-Regel)

Sei $n \geq 1$, $U \subseteq \mathbb{R}$ und $f, g \in \mathcal{C}^n(U, \mathbb{R})$ zwei n -fach differenzierbare Funktionen mit gleichem Definitionsbereich. Zeige:

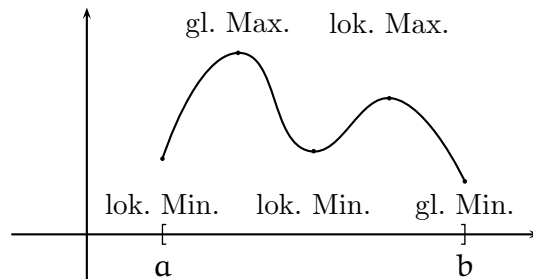
$$(f \cdot g)^{(n)} = \sum_{k=0}^n \binom{n}{k} \cdot f^{(k)} \cdot g^{(n-k)}.$$

§ 18 Der Mittelwertsatz und seine Anwendungen

Definition 18.1 (Extremstellen)

Es sei $f : U \rightarrow \mathbb{R}$ eine Funktion und $a \in U$.

- a. f hat in a ein *globales Maximum*, wenn $f(a) \geq f(x)$ für alle $x \in U$.
- b. f hat in a ein *lokales Maximum*, wenn es ein $\delta > 0$ gibt, so daß $f(a) \geq f(x)$ für alle $x \in U \cap U_\delta(a)$.
- c. f hat in a ein *globales Minimum*, wenn $f(a) \leq f(x)$ für alle $x \in U$.
- d. f hat in a ein *lokales Minimum*, wenn es ein $\delta > 0$ gibt, so daß $f(a) \leq f(x)$ für alle $x \in U \cap U_\delta(a)$.
- e. a heißt *Extremstelle* und $f(a)$ *Extremum* von f , wenn f in a ein lokales Minimum oder ein lokales Maximum hat.



Proposition 18.2 (Notwendige Bedingung für eine Extremstelle: $f'(c) = 0$)

Ist $f : (a, b) \rightarrow \mathbb{R}$ in einer Extremstelle c differenzierbar, so ist $f'(c) = 0$.

Beweis: Wir können ohne Einschränkung annehmen, daß c ein lokales Maximum ist, da der Beweis für ein lokales Minimum dann durch Übergang von f zu $-f$ folgt.

Nach Definition gibt es ein $\delta > 0$, so daß $f(c) \geq f(x)$ für alle $x \in (a, b) \cap (c - \delta, c + \delta)$. Ersetzen wir δ durch $\min\{\delta, b - c, c - a\}$ so können wir annehmen, daß $(c - \delta, c + \delta) \subseteq (a, b)$. Wir betrachten nun die Folgen $(a_n)_{n \geq 2}$ und $(b_n)_{n \geq 2}$ mit

$$a_n := c - \frac{\delta}{n} < c$$

und

$$b_n := c + \frac{\delta}{n} > c.$$

Dann konvergiert die Folge $(a_n)_{n \geq 2}$ von links gegen c , und die Folge $(b_n)_{n \geq 2}$ konvergiert von rechts gegen c . Nun betrachten wir den Grenzwert des Differenzenquotienten von f in c für die beiden Folgen und berücksichtigen, daß stets $f(a_n) - f(c) \leq 0$ und $f(b_n) - f(c) \leq 0$ gilt und daß außerdem $a_n - c < 0$ und $b_n - c > 0$ gilt:

$$0 \leq \frac{f(a_n) - f(c)}{a_n - c} \rightarrow f'(c)$$

und

$$0 \geq \frac{f(b_n) - f(c)}{b_n - c} \rightarrow f'(c).$$

Für den Grenzwert $f'(c)$ gilt also

$$0 \leq f'(c) \leq 0,$$

und mithin

$$f'(c) = 0.$$

□

Beispiel 18.3 a. Ist $n \geq 2$ gerade, so nimmt die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^n$$

in $a = 0$ ein globales Minimum an, und es gilt auch

$$f'(0) = n \cdot 0^{n-1} = 0.$$

- b. Die Funktion $f : (-1, 1) \longrightarrow \mathbb{R} : x \mapsto x^3$ hat in Null *keine* Extremstelle, da $f(x) < 0$ für $x < 0$ und $f(x) > 0$ für $x > 0$, dennoch gilt $f'(0) = 3 \cdot 0^2 = 0$. Die Bedingung $f'(c) = 0$ für eine Extremstelle c ist also *notwendig*, aber sie ist *nicht hinreichend*.

Bemerkung 18.4

Selbst wenn f auf dem abgeschlossenen Intervall $[a, b]$ definiert und dort überall differenzierbar ist, macht Proposition 18.2 *keine* Aussagen über die Ableitung in den *Randpunkten* a und b , falls diese Extremstellen sind!

Die Funktion $f : [-1, 1] \longrightarrow \mathbb{R} : x \mapsto x^3$ nimmt in $a = -1$ ihr globales Minimum und in $a = 1$ ihr globales Maximum an, aber die Ableitungen $f'(-1) = 3 = f'(1)$ sind beide nicht Null.

Satz 18.5 (Satz von Rolle)

Ist $a < b$ und ist $f : [a, b] \longrightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf (a, b) mit $f(a) = f(b)$, so gibt es ein $c \in (a, b)$ mit $f'(c) = 0$.

Beweis: Ist f konstant auf dem Intervall $[a, b]$ so ist $f'(c) = 0$ für jedes $c \in (a, b)$. Wir können also annehmen, daß es ein $y \in (a, b)$ mit $f(y) \neq f(a) = f(b)$ gibt.

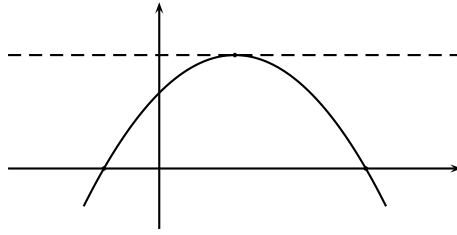
Wir betrachten zunächst den Fall, daß $f(y) > f(a) = f(b)$ gilt. Da f stetig auf dem abgeschlossenen Intervall $[a, b]$ ist, nimmt f dort nach Satz 14.16 sein Maximum an, d.h. es gibt ein $c \in [a, b]$ mit $f(c) \geq f(x)$ für alle $x \in [a, b]$. c ist also eine Extremstelle, und wegen $f(y) > f(a) = f(b)$, muß $c \in (a, b)$ gelten, so daß wir Proposition 18.2 anwenden können und $f'(c) = 0$ erhalten.

Der Fall $f(y) < f(a) = f(b)$ geht analog, da dann ein globales Minimum von f in (a, b) existiert. □

Bemerkung 18.6

Der Satz von Rolle besagt insbesondere, daß die Ableitung zwischen zwei Nullstellen einer differenzierbaren Funktion mindestens einmal Null werden muß, und im Beweis

haben wir gesehen, daß das daran liegt, daß die Funktion dort eine Extremstelle besitzt.



Satz 18.7 (Mittelwertsatz)

Ist $a < b$ und ist $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf (a, b) , so gibt es ein $c \in (a, b)$ mit $f'(c) = \frac{f(b)-f(a)}{b-a}$.

Beweis: Die Funktion

$$g : [a, b] \rightarrow \mathbb{R} : x \mapsto f(x) - \frac{f(b) - f(a)}{b - a} \cdot (x - a)$$

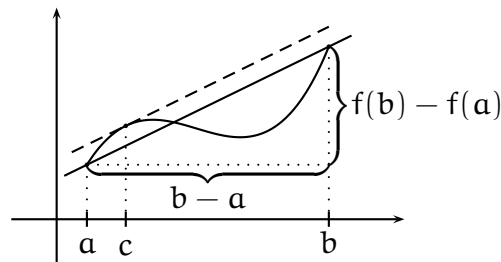
ist stetig auf $[a, b]$ und differenzierbar auf (a, b) . Außerdem gilt $g(a) = f(a) = g(b)$. Aus dem Satz von Rolle 18.5 folgt die Existenz eines $c \in (a, b)$ mit

$$0 = g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}.$$

□

Bemerkung 18.8

Der Mittelwertsatz besagt, daß zwischen a und b ein c liegt, in dem die Steigung der Tangente $t_{f,c}$ an den Graphen von f mit der Steigung der Sekante $s_{f,a,b}$ durch a und b übereinstimmt.



Beispiel 18.9

Betrachten wir die Funktion

$$f : [-1, 1] \rightarrow \mathbb{R} : x \mapsto x^3.$$

Aus dem Mittelwertsatz folgt, daß es ein $c \in (-1, 1)$ geben muß, so daß die Tangente an den Graphen von f im Punkt (c, c^3) die Steigung

$$\frac{f(1) - f(-1)}{1 - (-1)} = \frac{2}{2} = 1$$

hat. Da wir die Ableitungsfunktion kennen, können wir versuchen, c zu bestimmen. Es muß gelten

$$1 = f'(c) = 3 \cdot c^2.$$

Wir finden also zwei solcher Stellen:

$$c = \frac{1}{\sqrt{3}} \quad \text{und} \quad c = -\frac{1}{\sqrt{3}}.$$

Korollar 18.10 (Allgemeiner Mittelwertsatz)

Ist $a < b$ und sind $f, g : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf (a, b) , so gibt es ein $c \in (a, b)$ mit $f'(c) \cdot (g(b) - g(a)) = g'(c) \cdot (f(b) - f(a))$.

Beweis: Wir betrachten die Funktion

$$h : [a, b] \rightarrow \mathbb{R} : x \mapsto f(x) \cdot (g(b) - g(a)) - g(x) \cdot (f(b) - f(a)).$$

h ist auf $[a, b]$ stetig und auf (a, b) differenzierbar mit $h(a) = f(a) \cdot g(b) - g(a) \cdot f(b) = h(b)$. Aus dem Satz von Rolle folgt, daß es ein $c \in (a, b)$ gibt mit

$$0 = h'(c) = f'(c) \cdot (g(b) - g(a)) - g'(c) \cdot (f(b) - f(a)).$$

□

Wir wollen uns nun den Anwendungen des Mittelwertsatzes zuwenden.

A) Monotonie und Ableitung

Mit Hilfe der Ableitung läßt sich bei differenzierbaren Funktionen ein hinreichendes Kriterium für Monotonie angeben.

Proposition 18.11 (Hinreichendes Kriterium für Monotonie)

Es sei $a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ sei stetig auf $[a, b]$ und differenzierbar auf (a, b) .

- Ist $f'(x) > 0$ für alle $x \in (a, b)$, so ist f streng monoton wachsend auf $[a, b]$.
- Ist $f'(x) < 0$ für alle $x \in (a, b)$, so ist f streng monoton fallend auf $[a, b]$.

Beweis:

- Es seien $x, y \in [a, b]$ mit $x < y$ gegeben. Dann ist f stetig auf $[x, y]$ und differenzierbar auf (x, y) . Aus dem Mittelwertsatz folgt deshalb, daß es ein $c \in (x, y)$ gibt mit

$$f(y) - f(x) = f'(c) \cdot (y - x) > 0.$$

Also ist f streng monoton wachsend.

- Der Beweis geht analog zum ersten Teil.

□

Beispiel 18.12

Betrachte für $n \geq 1$ die Funktion

$$f : [0, \infty) \rightarrow \mathbb{R} : x \mapsto x^n.$$

Für die Ableitung gilt

$$f'(x) = n \cdot x^{n-1} > 0$$

für alle $x \in (0, \infty)$. Mithin ist die Funktion streng monoton wachsend auf jedem Intervall $[0, b] \subseteq [0, \infty)$ und mithin auf $[0, \infty)$. Dies ist ein alternativer Beweis der Aussage in Beispiel 14.19.

B) Hinreichende Bedingung für Extremstellen

Proposition 18.13 (Hinreichende Bedingung für eine Extremstelle)

Es sei $f : (a, b) \rightarrow \mathbb{R}$ eine zweifach differenzierbare Funktion und $c \in (a, b)$.

- a. Falls $f'(c) = 0$ und $f''(c) < 0$, so ist c ein lokales Maximum.
- b. Falls $f'(c) = 0$ und $f''(c) > 0$, so ist c ein lokales Minimum.

Beweis:

b. Nach Voraussetzung ist

$$\lim_{x \rightarrow c} \frac{f'(x)}{x - c} = \lim_{x \rightarrow c} \frac{f'(x) - f'(c)}{x - c} = f''(c) > 0.$$

Zu $\varepsilon := \frac{f''(c)}{2} > 0$ gibt es dann ein $\delta_\varepsilon > 0$, so daß

$$-\frac{f''(c)}{2} = -\varepsilon < \frac{f'(x)}{x - c} - f''(c) < \varepsilon = \frac{f''(c)}{2}$$

für alle $x \in (a, b)$ mit $|x - c| < \delta_\varepsilon$. Insbesondere folgt für diese x dann

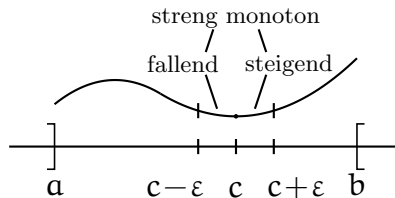
$$\frac{f'(x)}{x - c} > -\frac{f''(c)}{2} + f''(c) = \frac{f''(c)}{2} > 0. \tag{24}$$

Wir können ohne Einschränkung annehmen, daß δ_ε so klein ist, daß $a < c - \delta_\varepsilon < c + \delta_\varepsilon < b$ gilt.

Für $x \in (c - \delta_\varepsilon, c)$ folgt aus (24) dann $f'(x) < 0$, und nach Proposition 18.11 ist f dann streng monoton fallend auf dem Intervall $[c - \delta_\varepsilon, c]$.

Analog folgt für $x \in (c, c + \delta_\varepsilon)$ aus (24) $f'(x) > 0$ und aus Proposition 18.11 folgt, daß f streng monoton wachsend auf dem Intervall $[c, c + \delta_\varepsilon]$ ist.

Insbesondere heißt das, daß $f(c) \leq f(x)$ für alle $x \in [c - \delta_\varepsilon, c + \delta_\varepsilon]$, so daß f in c ein Minimum besitzt.



a. Die Aussage beweist man analog.

□

Beispiel 18.14

Wir betrachten die Funktion

$$f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^3 - 3x^2 - 1.$$

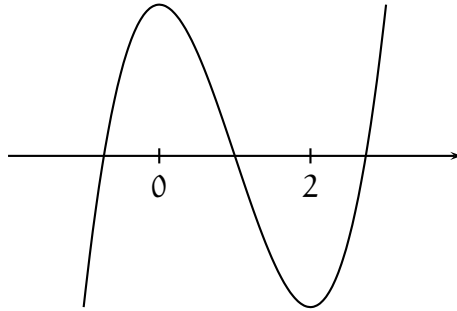
Um mögliche Extremstellen zu finden, müssen wir die Nullstellen der ersten Ableitung

$$f'(x) = 3x^2 - 6x$$

finden. Das ist für $x = 0$ und $x = 2$ der Fall. In diesen Punkten schauen wir uns die zweite Ableitung

$$f''(x) = 6x - 6$$

an. Aus $f''(0) = -6 < 0$ folgt, daß in $x = 0$ ein Maximum vorliegt, und aus $f''(2) = 6 > 0$ folgt, daß in $x = 2$ ein Minimum vorliegt.



Bemerkung 18.15 (Hinreichende Bedingung für Extremstellen)

Anstatt zweifacher Differenzierbarkeit und der Bedingung an die zweite Ableitung kann man auch einfach fordern, daß die erste Ableitung in c einen Vorzeichenwechsel hat, wie wir ihn im Beweis von Proposition 18.13 aus den Bedingungen an $f''(c)$ ableiten.

C) Vertauschbarkeit von Grenzwert und Ableitung

Satz 18.16 (Vertauschbarkeit von Grenzwert und Ableitung)

Ist $(f_n)_{n \in \mathbb{N}}$ eine Funktionenfolge stetig differenzierbarer Funktionen auf dem Intervall $[a, b]$, so daß $(f_n)_{n \in \mathbb{N}}$ punktweise gegen f und $(f'_n)_{n \in \mathbb{N}}$ gleichmäßig gegen g konvergiert, dann ist f stetig differenzierbar auf $[a, b]$ mit Ableitung $f' = g$.

Beweis: Nach Voraussetzung sind die f'_n stetig auf $[a, b]$, so daß die Grenzfunktion g als gleichmäßiger Grenzwert stetiger Funktionen nach Satz 15.6 ebenfalls stetig ist.

Sei nun $\varepsilon > 0$ und $c \in [a, b]$ gegeben, so müssen wir ein $\delta_\varepsilon > 0$ finden, so daß für alle $c \neq x \in [a, b]$ mit $|x - c| < \delta_\varepsilon$ auch

$$|\text{Diff}_{f,c}(x) - g(c)| = \left| \frac{f(x) - f(c)}{x - c} - g(c) \right| < \varepsilon$$

gilt, d.h. $g(c)$ ist der Grenzwert des Differenzenquotienten von f in c .

Da $(f'_n)_{n \in \mathbb{N}}$ gleichmäßig gegen g konvergiert, gibt es ein $n_\varepsilon \in \mathbb{N}$, so daß

$$|f'_n(x) - g(x)| < \frac{\varepsilon}{3} \tag{25}$$

für alle $n \geq n_\varepsilon$ und alle $x \in [a, b]$ gilt.

Da g stetig in c ist, gibt es zudem ein $\delta_\varepsilon > 0$, so daß für alle $x \in [a, b]$ mit $|x - c| < \delta_\varepsilon$ auch

$$|g(x) - g(c)| < \frac{\varepsilon}{3} \quad (26)$$

gilt.

Sei nun $c \neq x \in [a, b]$ mit $|x - c| < \delta_\varepsilon$ gegeben. Für $n \geq n_\varepsilon$ können wir den Mittelwertsatz 18.7 auf die differenzierbare Funktion f_n anwenden und finden somit ein y zwischen x und c mit

$$\frac{f_n(x) - f_n(c)}{x - c} = f'_n(y) \quad (27)$$

und da y zwischen x und c liegt, gilt auch $|y - c| \leq |x - c| < \delta_\varepsilon$.

Setzen wir die obigen Ergebnisse nun zusammen, so erhalten wir

$$\begin{aligned} \left| \frac{f_n(x) - f_n(c)}{x - c} - g(c) \right| &\stackrel{(27)}{=} |f'_n(y) - g(c)| \\ &\leq |f'_n(y) - g(y)| + |g(y) - g(c)| \\ &\stackrel{(25)(26)}{<} \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2 \cdot \varepsilon}{3} \end{aligned}$$

für alle $n \geq n_\varepsilon$. Lassen wir nun n gegen unendlich laufen, so erhalten wir für den Grenzwert

$$\left| \frac{f(x) - f(c)}{x - c} - g(c) \right| = \lim_{n \rightarrow \infty} \left| \frac{f_n(x) - f_n(c)}{x - c} - g(c) \right| \leq \frac{2 \cdot \varepsilon}{3} < \varepsilon.$$

Damit haben wir gezeigt, daß f in einem beliebigen Punkt c des Intervalls $[a, b]$ differenzierbar ist und daß $f' = g$ gilt. Da wir bereits wissen, daß g stetig ist, ist f mithin stetig differenzierbar auf $[a, b]$. \square

Bemerkung 18.17

- a. Die Aussage in Satz 18.16 besagt, daß

$$\lim_{n \rightarrow \infty} (f'_n) = \left(\lim_{n \rightarrow \infty} f_n \right)',$$

d.h. die Grenzwertbildung für die Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ vertauscht mit der Ableitung!

Auf die Differenzenquotienten zurückgeführt, bedeutet dies

$$\lim_{n \rightarrow \infty} \lim_{x \rightarrow a} \frac{f_n(x) - f_n(a)}{x - a} = \lim_{x \rightarrow a} \lim_{n \rightarrow \infty} \frac{f_n(x) - f_n(a)}{x - a}.$$

Hier vertauschen zwei Grenzwertprozesse! Das ist eine Besonderheit!

- b. Man kann in Satz 18.16 auf die Bedingung, daß die Ableitungen f'_n stetig sind, verzichten. Der Beweis wird dann aber etwas technischer.
- c. Auch wenn wir in Satz 18.16 nur die punktweise Konvergenz für die Folge $(f_n)_{n \in \mathbb{N}}$ gefordert haben, erzwingt die gleichmäßige Konvergenz der Folge $(f'_n)_{n \in \mathbb{N}}$ letztlich die gleichmäßige Konvergenz der Folge $(f_n)_{n \in \mathbb{N}}$.

Beweis der Aussagen b. und c. Wir verwenden die Notation und die Voraussetzungen von Satz 18.16, verzichten aber auf die Bedingung, daß die f'_n stetig sind!

Zu Teil b.: f ist differenzierbar mit $f' = g$: Dazu definieren wir uns für ein fest gegebenes $c \in [a, b]$ die Funktionenfolge

$$h_n : [a, b] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \frac{f_n(x) - f_n(c)}{x - c} - f'_n(c), & \text{falls } x \neq c, \\ 0, & \text{falls } x = c. \end{cases}$$

Da die f_n in c differenzierbar sind, ist die Funktion h_n stetig in c . In den übrigen Punkten von $[a, b]$ ist die Funktion aber ohnehin stetig, da die f_n als differenzierbare Funktionen auch stetig sind nach Satz 17.7.

Wir wollen zeigen, daß die Funktionenfolge $(h_n)_{n \in \mathbb{N}}$ gleichmäßig auf $[a, b]$ gegen die Funktion

$$h : [a, b] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \frac{f(x) - f(c)}{x - c} - g(c), & \text{falls } x \neq c, \\ 0, & \text{falls } x = c \end{cases}$$

konvergiert. Die punktweise Konvergenz von $(h_n)_{n \in \mathbb{N}}$ gegen h ist klar, da $(f_n)_{n \in \mathbb{N}}$ punktweise gegen f konvergiert und $(f'_n)_{n \in \mathbb{N}}$ auch punktweise gegen g konvergiert.

Sei also $\varepsilon > 0$ gegeben. Wir müssen ein $n_\varepsilon \in \mathbb{N}$ finden, so daß

$$|h_n(x) - h(x)| < \varepsilon$$

für alle $n \geq n_\varepsilon$ und $x \in [a, b]$ gilt. Wegen $h_n(c) = 0 = h(c)$ können wir uns dazu ein $c \neq x \in [a, b]$ vorgeben.

Da die Folge $(f'_n)_{n \in \mathbb{N}}$ gleichmäßig gegen g konvergiert, gibt es ein $n_\varepsilon \in \mathbb{N}$, so daß

$$|f'_n(y) - g(y)| < \frac{\varepsilon}{6}$$

für alle $n \geq n_\varepsilon$ und $y \in [a, b]$. Aus der Dreiecksungleichung erhalten wir dann

$$|f'_m(y) - f'_n(y)| \leq |f'_m(y) - g(y)| + |g(y) - f'_n(y)| < \frac{\varepsilon}{3} \quad (28)$$

für alle $m > n \geq n_\varepsilon$ und $y \in [a, b]$.

Wenden wir nun den Mittelwertsatz 18.7 auf die differenzierbare Funktion $f_m - f_n$ für $m > n \geq n_\varepsilon$ an, so finden wir ein y zwischen x und c mit

$$\frac{(f_m - f_n)(x) - (f_m - f_n)(c)}{x - c} = (f'_m - f'_n)(y) \quad (29)$$

Für $m > n \geq n_\varepsilon$ erhalten wir damit

$$\begin{aligned}
 |h_m(x) - h_n(x)| &= \left| \frac{f_m(x) - f_m(c)}{x - c} - f'_m(c) - \frac{f_n(x) - f_n(c)}{x - c} + f'_n(c) \right| \\
 &= \left| \frac{(f_m - f_n)(x) - (f_m - f_n)(c)}{x - c} - (f'_m(c) - f'_n(c)) \right| \\
 &\stackrel{(29)}{=} |(f'_m(y) - f'_n(y)) - (f'_m(c) - f'_n(c))| \\
 &\leq |f'_m(y) - f'_n(y)| + |f'_m(c) - f'_n(c)| \\
 &\stackrel{(28)}{<} \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2 \cdot \varepsilon}{3}.
 \end{aligned}$$

Lassen wir nun m gegen unendlich gehen, erhalten wir für den Grenzwert

$$|h(x) - h_n(x)| = \lim_{m \rightarrow \infty} |h_m(x) - h_n(x)| \leq \frac{2 \cdot \varepsilon}{3} < \varepsilon.$$

Damit haben wir gezeigt, daß $(h_n)_{n \in \mathbb{N}}$ auf $[a, b]$ gleichmäßig gegen h konvergiert. Als gleichmäßige Grenzfunktion stetiger Funktionen ist h damit nach Satz 15.6 stetig auf $[a, b]$, und insbesondere stetig in c . D.h.

$$\lim_{x \rightarrow c} \text{Diff}_{f,c}(x) - g(c) = \lim_{x \rightarrow c} h(x) = h(c) = 0,$$

also ist f in c differenzierbar mit $f'(c) = g(c)$.

Zu Teil c.: $(f_n)_{n \in \mathbb{N}}$ konvergiert gleichmäßig gegen f : Sei $\varepsilon > 0$ gegeben. Wir müssen ein $n_\varepsilon \in \mathbb{N}$ finden, so daß

$$|f_n(x) - f(x)| < \varepsilon$$

für alle $n \geq n_\varepsilon$ und alle $x \in [a, b]$ gilt.

Da $(f_n)_{n \in \mathbb{N}}$ punktweise gegen f konvergiert, ist insbesondere die Folge $(f_n(a))_{n \in \mathbb{N}}$ eine Cauchy-Folge und wir finden ein n'_ε , so daß für alle $m > n \geq n'_\varepsilon$

$$|f_m(a) - f_n(a)| < \frac{\varepsilon}{4}. \quad (30)$$

Da die Folge $(f'_n)_{n \in \mathbb{N}}$ gleichmäßig gegen g konvergiert, finden wir zudem ein n''_ε , so daß

$$|f'_n(y) - g(y)| < \frac{\varepsilon}{4 \cdot (b - a)} \quad (31)$$

für alle $y \in [a, b]$ und $n \geq n''_\varepsilon$ gilt.

Wir betrachten nun ein beliebiges $x \in [a, b]$ und beliebige $m > n \geq n_\varepsilon := \max\{n'_\varepsilon, n''_\varepsilon\}$. Wenden wir den Mittelwertsatz 18.7 auf die differenzierbare Funktion $f_m - f_n$ an, so finden wir ein $y \in [a, x] \subseteq [a, b]$ mit

$$(f_m - f_n)(x) - (f_m - f_n)(a) = (f'_m - f'_n)(y) \cdot (x - a). \quad (32)$$

Diese Gleichung wollen wir nun ausnutzen, um die Differenz $f_m(x) - f_n(x)$ für $x \in [a, b]$ und $m > n \geq n_\varepsilon$ abzuschätzen:

$$\begin{aligned} |f_m(x) - f_n(x)| &\leq |(f_m - f_n)(x) - (f_m - f_n)(a)| + |f_m(a) - f_n(a)| \\ &\stackrel{(32)}{=} |f'_m(y) - f'_n(y)| \cdot |x - a| + |f_m(a) - f_n(a)| \\ &\stackrel{(30)}{<} (|f'_m(y) - g(y)| + |g(y) - f'_n(y)|) \cdot |b - a| + \frac{\varepsilon}{4} \\ &\stackrel{(31)}{<} \left(\frac{\varepsilon}{4 \cdot (b - a)} + \frac{\varepsilon}{4 \cdot (b - a)} \right) \cdot (b - a) + \frac{\varepsilon}{4} = \frac{3 \cdot \varepsilon}{4}. \end{aligned}$$

Halten wir n fest und betrachten $m \rightarrow \infty$, so erhalten wir für den Grenzwert

$$|f(x) - f_n(x)| = \lim_{m \rightarrow \infty} |f_m(x) - f_n(x)| \leq \frac{3 \cdot \varepsilon}{4} < \varepsilon,$$

und dies gilt für jedes $n \geq n_\varepsilon$ und jedes $x \in [a, b]$. □

D) Ableitung von Potenzreihen

Korollar 18.18 (Ableitung von Potenzreihen)

Ist $\sum_{n=0}^{\infty} a_n \cdot t^n$ eine Potenzreihe über \mathbb{R} mit Konvergenzradius $r > 0$, dann ist die Funktion

$$f : (-r, r) \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} a_n \cdot x^n$$

differenzierbar auf $(-r, r)$ und die Ableitung in $x \in (-r, r)$ ist gegeben durch

$$f'(x) = \sum_{n=1}^{\infty} n \cdot a_n \cdot x^{n-1},$$

d.h. durch die formale Ableitung $\sum_{n=1}^{\infty} n \cdot a_n \cdot t^{n-1}$ der Potenzreihe.

Beweis: Aus Aufgabe 12.44 wissen wir, daß die beiden Reihen $\sum_{n=0}^{\infty} a_n \cdot t^n$ und ihre formale Ableitung $\sum_{n=1}^{\infty} n \cdot a_n \cdot t^{n-1}$ den gleichen Konvergenzradius r besitzen. Insbesondere definiert letztere Reihe eine Funktion

$$g : (-r, r) \longrightarrow \mathbb{R} : x \mapsto \sum_{n=1}^{\infty} n \cdot a_n \cdot x^{n-1},$$

die nach Korollar 15.7 stetig ist.

Sei nun $a \in (-r, r)$ gegeben. Wir wollen zeigen, daß f in a differenzierbar ist mit

$$f'(a) = g(a).$$

Dazu setzen wir $R := \frac{r+|a|}{2} < r$, so daß $a \in (-R, R)$ liegt. Die Folge $(f_n)_{n \in \mathbb{N}}$ mit

$$f_n : [-R, R] \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^n a_k \cdot x^k$$

konvergiert nach Satz 15.4 auf dem abgeschlossenen Intervall $[-R, R]$ gleichmäßig gegen die Funktion f . Nach Beispiel 17.10 sind die f_n differenzierbar mit stetiger Ableitung

$$f'_n : [-R, R] \longrightarrow \mathbb{R} : x \mapsto \sum_{k=1}^n k \cdot a_k \cdot x^{k-1}.$$

Die Folge $(f'_n)_{n \in \mathbb{N}}$ der Ableitungen konvergiert dann wieder nach Satz 15.4 auf $[-R, R]$ gleichmäßig gegen g . Da die Voraussetzungen von Satz 18.16 erfüllt sind, ist f auf $[-R, R]$ differenzierbar mit $f' = g$. Insbesondere ist f also in \mathbf{a} differenzierbar mit $f'(\mathbf{a}) = g(\mathbf{a})$. \square

Da wir die Aussage des Korollars auch auf die formale Ableitung der Potenzreihe anwenden können, erhalten wir durch Induktion die folgende Aussage.

Korollar 18.19 (Differenzierbarkeit von Potenzreihen)

Eine durch eine Potenzreihe definierte Funktion ist auf ihrem Konvergenzbereich unendlich oft differenzierbar.

Die Exponentialfunktion, der Sinus und der Cosinus sind also differenzierbar.

Korollar 18.20 (Ableitungen wichtiger Funktionen)

a. *Die Exponentialfunktion*

$$\exp : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

ist unendlich oft differenzierbar auf \mathbb{R} mit Ableitung

$$\exp'(x) = \exp(x).$$

b. *Der Sinus*

$$\sin : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n+1}}{(2n+1)!}$$

ist unendlich oft differenzierbar auf \mathbb{R} mit Ableitung

$$\sin'(x) = \cos(x).$$

c. *Der Cosinus*

$$\cos : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} (-1)^n \cdot \frac{x^{2n}}{(2n)!}$$

ist unendlich oft differenzierbar auf \mathbb{R} mit Ableitung

$$\cos'(x) = -\sin(x).$$

d. *Für $\mathbf{a} \in \mathbb{R}_{>0}$ ist die Exponentialfunktion zur Basis \mathbf{a}*

$$\exp_{\mathbf{a}} : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \exp(x \cdot \ln(\mathbf{a}))$$

stetig differenzierbar auf \mathbb{R} mit Ableitung

$$\exp'_{\mathbf{a}}(x) = \ln(\mathbf{a}) \cdot \exp_{\mathbf{a}}(x).$$

- e. Für $a \in \mathbb{R}_{>0}$ mit $a \neq 1$ ist die Logarithmusfunktion zur Basis a

$$\log_a : (0, \infty) \longrightarrow \mathbb{R}$$

stetig differenzierbar auf $(0, \infty)$ mit Ableitung

$$\log'_a(x) = \frac{1}{x \cdot \ln(a)}.$$

Insbesondere gilt für die Ableitung des natürlichen Logarithmus

$$\ln'(x) = \frac{1}{x}.$$

- f. Der Tangens

$$\tan : \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \longrightarrow \mathbb{R} : x \mapsto \frac{\sin(x)}{\cos(x)}$$

ist auf $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ stetig differenzierbar mit Ableitung

$$\tan'(x) = \frac{1}{\cos^2(x)}.$$

- g. Der Cotangens

$$\cot : (0, \pi) \longrightarrow \mathbb{R} : x \mapsto \frac{\cos(x)}{\sin(x)}$$

ist stetig differenzierbar auf $(0, \pi)$ mit Ableitung

$$\cot'(x) = -\frac{1}{\sin^2(x)}.$$

- h. Der Arcustangens ist auf \mathbb{R} stetig differenzierbar mit

$$\arctan'(x) = \frac{1}{1+x^2}.$$

- i. Der Arcuscotangens ist auf \mathbb{R} stetig differenzierbar mit

$$\operatorname{arccot}'(x) = -\frac{1}{1+x^2}.$$

- j. Der Arcussinus ist auf $(-1, 1)$ differenzierbar mit

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}}.$$

- k. Der Arcuscosinus ist auf $(-1, 1)$ differenzierbar mit

$$\arccos'(x) = -\frac{1}{\sqrt{1-x^2}}.$$

Beweis:

- a. \exp ist nach Korollar 18.18 und 18.19 unendlich oft differenzierbar auf \mathbb{R} mit Ableitung

$$\exp'(x) = \sum_{n=1}^{\infty} n \cdot \frac{x^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty} \frac{x^n}{n!} = \exp(x).$$

- b. \sin ist nach Korollar 18.18 und 18.19 unendlich oft differenzierbar auf \mathbb{R} mit Ableitung

$$\sin'(x) = \sum_{n=0}^{\infty} (-1)^n (2n+1) \cdot \frac{x^{2n}}{(2n+1)!} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = \cos(x).$$

- c. \cos ist nach Korollar 18.18 und 18.19 unendlich oft differenzierbar auf \mathbb{R} mit Ableitung

$$\begin{aligned} \cos'(x) &= \sum_{n=1}^{\infty} (-1)^n (2n) \cdot \frac{x^{2n-1}}{(2n)!} = - \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^{2n-1}}{(2n-1)!} \\ &= - \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = -\sin(x). \end{aligned}$$

- d. Aus der Kettenregel erhalten wir, daß die Exponentialfunktion zur Basis a differenzierbar ist mit Ableitung

$$\exp'_a(x) = \ln(a) \cdot \exp_a(x),$$

und diese Funktion ist offenbar wieder stetig.

- e. Aus dem Satz zur Ableitung von Umkehrfunktionen 17.14 folgt, daß \log_a auf $(0, \infty)$ differenzierbar ist, da die Ableitung der Exponentialfunktion \exp_a nie Null wird. Für die Ableitung erhalten wir zudem

$$\log'_a(x) = \frac{1}{\exp'_a(\log_a(x))} = \frac{1}{\ln(a) \cdot \exp_a(\log_a(x))} = \frac{1}{\ln(a) \cdot x}.$$

Zudem ist die Ableitung offenbar stetig.

- f. Aus der Quotientenregel erhalten wir, daß der Tangens in $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ differenzierbar ist mit

$$\tan'(x) = \frac{\sin'(x) \cdot \cos(x) - \sin(x) \cdot \cos'(x)}{\cos(x)^2} = \frac{\sin(x)^2 + \cos(x)^2}{\cos(x)^2} = \frac{1}{\cos(x)^2}.$$

Als Quotient stetiger Funktionen ist die Ableitung insbesondere stetig.

- g. Aus der Quotientenregel erhalten wir, daß der Cotangens in $x \in (0, \pi)$ differenzierbar ist mit

$$\cot'(x) = \frac{\cos'(x) \cdot \sin(x) - \cos(x) \cdot \sin'(x)}{\sin(x)^2} = \frac{-\cos(x)^2 - \sin(x)^2}{\sin(x)^2} = -\frac{1}{\sin(x)^2}.$$

Als Quotient stetiger Funktionen ist die Ableitung insbesondere stetig.

- h. Aus dem Satz zur Ableitung von Umkehrfunktionen 17.14 zusammen mit Satz 16.18 folgt, daß \arctan auf \mathbb{R} differenzierbar ist, da die Ableitung des Tangens

nie Null wird auf $(-\frac{\pi}{2}, \frac{\pi}{2})$. Für die Ableitung erhalten wir zudem

$$\begin{aligned} \arctan'(x) &= \frac{1}{\tan'(\arctan(x))} = \frac{1}{\frac{1}{\cos^2(\arctan(x))}} \\ &= \frac{1}{\frac{\sin^2(\arctan(x)) + \cos^2(\arctan(x))}{\cos^2(\arctan(x))}} = \frac{1}{\frac{\sin^2(\arctan(x))}{\cos^2(\arctan(x))} + 1} \\ &= \frac{1}{\tan(\arctan(x))^2 + 1} = \frac{1}{x^2 + 1}. \end{aligned}$$

Die Ableitung ist zudem offenbar stetig.

- i. Aus dem Satz zur Ableitung von Umkehrfunktionen 17.14 zusammen mit Satz 16.18 folgt, daß arccot auf \mathbb{R} differenzierbar ist, da die Ableitung des Cotangens nie Null wird auf $(0, \pi)$. Für die Ableitung erhalten wir zudem

$$\begin{aligned} \operatorname{arccot}'(x) &= \frac{1}{\cot'(\operatorname{arccot}(x))} = \frac{1}{-\frac{1}{\sin^2(\operatorname{arccot}(x))}} \\ &= -\frac{1}{\frac{\sin^2(\operatorname{arccot}(x)) + \cos^2(\operatorname{arccot}(x))}{\sin^2(\operatorname{arccot}(x))}} = -\frac{1}{1 + \frac{\cos^2(\operatorname{arccot}(x))}{\sin^2(\operatorname{arccot}(x))}} \\ &= -\frac{1}{1 + \cot(\operatorname{arccot}(x))^2} = -\frac{1}{1 + x^2}. \end{aligned}$$

Die Ableitung ist zudem offenbar stetig.

- j. Aus dem Satz zur Ableitung von Umkehrfunktionen 17.14 zusammen mit Satz 16.18 folgt, daß \arcsin auf $(-1, 1)$ differenzierbar ist, da die Ableitung des Sinus nie Null wird auf $(-\frac{\pi}{2}, \frac{\pi}{2})$. Für die Ableitung erhalten wir zudem unter Berücksichtigung der Tatsache, daß der Cosinus auf $(-\frac{\pi}{2}, \frac{\pi}{2})$ positiv ist:

$$\begin{aligned} \arcsin'(x) &= \frac{1}{\sin'(\arcsin(x))} = \frac{1}{\cos(\arcsin(x))} \\ &= \frac{1}{\sqrt{\cos^2(\arcsin(x))}} = \frac{1}{\sqrt{1 - \sin^2(\arcsin(x))}} \\ &= \frac{1}{\sqrt{1 - x^2}}. \end{aligned}$$

Die Ableitung ist zudem offenbar stetig.

- k. Aus dem Satz zur Ableitung von Umkehrfunktionen 17.14 zusammen mit Satz 16.18 folgt, daß \arccos auf $(-1, 1)$ differenzierbar ist, da die Ableitung des Cosinus nie Null wird auf $(0, \pi)$. Für die Ableitung erhalten wir zudem unter

Berücksichtigung der Tatsache, daß der Sinus auf $(0, \pi)$ positiv ist:

$$\begin{aligned} \arccos'(x) &= \frac{1}{\cos'(\arccos(x))} = \frac{1}{-\sin(\arccos(x))} \\ &= -\frac{1}{\sqrt{\sin^2(\arccos(x))}} = -\frac{1}{\sqrt{1 - \cos^2(\arccos(x))}} \\ &= -\frac{1}{\sqrt{1 - x^2}}. \end{aligned}$$

Die Ableitung ist zudem offenbar stetig.

□

Bemerkung 18.21

Schaut man sich die Ableitungen der Funktionen in Korollar 18.20 d.-k. an, so kann man leicht durch Induktion zeigen, daß jede der Funktionen auf ihrem Definitionsbereich unendlich oft differenzierbar ist.

Beispiel 18.22

Für $a \in \mathbb{R}$ ist die Funktion

$$f : (0, \infty) \longrightarrow \mathbb{R} : x \mapsto x^a$$

unendlich oft differenzierbar auf $(0, \infty)$ mit

$$f'(x) = a \cdot x^{a-1}.$$

Um dies zu sehen, beachten wir, daß $f(x) = \exp(a \cdot \ln(x))$ gilt, so daß f nach Korollar 18.20 die Verkettung zweier differenzierbarer Funktionen ist. Aus der Kettenregel 17.16 folgt dann

$$f'(x) = \exp'(a \cdot \ln(x)) \cdot \frac{a}{x} = \exp(a \cdot \ln(x)) \cdot \frac{a}{x} = x^a \cdot \frac{a}{x} = a \cdot x^{a-1}.$$

Daß f sogar unendlich oft differenzierbar ist, folgt dann mit Induktion aus der Tatsache, daß f' eine Funktion der gleichen Gestalt ist.

E) Die Regeln von de l'Hôpital

Im folgenden Satz soll $[-\infty, \infty] := \mathbb{R} \cup \{-\infty, \infty\}$ bezeichnen.

Satz 18.23 (Regeln von de l'Hôpital)

Seien $a, b \in [-\infty, \infty]$ mit $a < b$, $f, g : (a, b) \longrightarrow \mathbb{R}$ differenzierbar und $c \in [a, b]$. Ferner gelte $g'(x) \neq 0$ für alle $x \in (a, b)$ und $\lim_{x \rightarrow c} \frac{f(x)}{g(x)}$ existiere eigentlich oder uneigentlich.

- Falls $\lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} g(x) = 0$, so gilt $\lim_{x \rightarrow c} \frac{f(x)}{g(x)} = \lim_{x \rightarrow c} \frac{f'(x)}{g'(x)}$.
- Falls $\lim_{x \rightarrow c} g(x) \in \{\infty, -\infty\}$, so gilt $\lim_{x \rightarrow c} \frac{f(x)}{g(x)} = \lim_{x \rightarrow c} \frac{f'(x)}{g'(x)}$.

Beweis: Wir beschränken uns im Beweis auf den Fall $c \in \mathbb{R}$ und $k := \lim_{x \rightarrow c} \frac{f'(x)}{g'(x)} \in \mathbb{R}$. Die Fälle $c \in \{-\infty, \infty\}$ oder $k \in \{-\infty, \infty\}$ beweist man analog.

Sei $\varepsilon > 0$ gegeben, so müssen wir ein $\delta_\varepsilon > 0$ finden, so daß

$$\left| \frac{f(x)}{g(x)} - k \right| < \varepsilon \quad (33)$$

für alle $c \neq x \in (a, b)$ mit $|x - c| < \delta_\varepsilon$.

Da $\frac{f'(x)}{g'(x)}$ gegen k konvergiert für x gegen c , gibt es ein $\delta'_\varepsilon > 0$, so daß für alle $z \in (a, b)$ mit $|z - c| < \delta'_\varepsilon$ auch

$$\left| \frac{f'(z)}{g'(z)} - k \right| < \frac{\varepsilon}{2} \quad (34)$$

gilt.

Wir betrachten nun $c \neq x, y \in (a, b) \cap (c - \delta'_\varepsilon, c + \delta'_\varepsilon)$ mit $x \neq y$ und wenden den allgemeinen Mittelwertsatz 18.10 an. Dann gibt es ein z zwischen x und y mit

$$f'(z) \cdot (g(x) - g(y)) = g'(z) \cdot (f(x) - f(y)). \quad (35)$$

Da z zwischen x und y liegt, gilt auch

$$z \in (a, b) \cap (c - \delta'_\varepsilon, c + \delta'_\varepsilon). \quad (36)$$

Nach Voraussetzung ist $g'(z) \neq 0$, und wir behaupten, daß auch $g(x) - g(y) \neq 0$ gilt, da es nach dem Satz von Rolle 18.5 sonst ein w zwischen x und y geben würde mit $g'(w) = 0$, was aber nach Voraussetzung nicht möglich ist. Damit können wir Gleichung (35) auch in der folgenden Form schreiben:

$$\frac{f(x) - f(y)}{g(x) - g(y)} = \frac{f'(z)}{g'(z)}. \quad (37)$$

- a. Wir betrachten nun den Fall, daß $\lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} g(x) = 0$. Ist $c \in (a, b)$, so folgt aus der Stetigkeit von f und g automatisch $f(c) = 0 = g(c)$. Ist $c \notin (a, b)$, so können wir f und g in c stetig fortsetzen durch $f(c) = 0 = g(c)$.

Aus (34), (36) und (37) folgt

$$\left| \frac{f(x) - f(y)}{g(x) - g(y)} - k \right| = \left| \frac{f'(z)}{g'(z)} - k \right| < \frac{\varepsilon}{2},$$

und dies gilt für alle $c \neq x, y \in (a, b) \cap (c - \delta'_\varepsilon, c + \delta'_\varepsilon)$ mit $x \neq y$. Da die Funktionen f und g nun stetig in c mit Funktionswert 0 sind, können wir y gegen c gehen lassen und erhalten im Grenzwert

$$\left| \frac{f(x)}{g(x)} - k \right| = \left| \frac{f(x) - f(c)}{g(x) - g(c)} - k \right| = \lim_{y \rightarrow c} \left| \frac{f(x) - f(y)}{g(x) - g(y)} - k \right| \leq \frac{\varepsilon}{2} < \varepsilon.$$

Dies gilt für alle $c \neq x \in (a, b) \cap (c - \delta'_\varepsilon, c + \delta'_\varepsilon)$, so daß wir mit $\delta_\varepsilon := \delta'_\varepsilon$ unsere Aussage in diesem Fall bewiesen haben.

b. Wir können annehmen, daß f nicht konstant 0 in einer kleinen Umgebung von c ist, da sonst auch $k = 0$ gilt und (33) sicher erfüllt ist. Deshalb können wir ein $c \neq y \in (a, b) \cap (c - \delta'_\varepsilon, c + \delta'_\varepsilon)$ festhalten mit $f(y) \neq 0$, und wegen $\lim_{x \rightarrow c} g(x) = \pm\infty$ können wir auch $g(y) \neq 0$ annehmen.

Aus $\lim_{x \rightarrow c} g(x) = \pm\infty$ folgt $\lim_{x \rightarrow c} \frac{1}{g(x)} = 0$, und deshalb gibt es ein $\delta''_\varepsilon > 0$, so daß

$$\left| \frac{1}{g(x)} \right| < \frac{\varepsilon}{4 \cdot |f(y)|} \tag{38}$$

für alle $x \in (a, b) \cap (c - \delta''_\varepsilon, c + \delta''_\varepsilon)$ mit $x \neq c, y$.

Aus (34), (36) und (37) folgt, daß

$$\left| \frac{f(x) - f(y)}{g(x) - g(y)} \right| \leq \left| \frac{f(x) - f(y)}{g(x) - g(y)} - k \right| + |k| < \frac{\varepsilon}{2} + |k| =: s \tag{39}$$

für alle $x \in (a, b) \cap (c - \delta'_\varepsilon, c + \delta'_\varepsilon)$ mit $x \neq c, y$, d.h. der Ausdruck ist auf dem angegebenen Intervall nach oben beschränkt.

Wegen $\lim_{x \rightarrow c} \frac{1}{g(x)} = 0$ gibt es ein $\delta'''_\varepsilon > 0$ mit

$$\left| \frac{1}{g(x)} \right| < \frac{\varepsilon}{4 \cdot |g(y)| \cdot s} \tag{40}$$

für alle $x \in (a, b) \cap (c - \delta'''_\varepsilon, c + \delta'''_\varepsilon)$ mit $x \neq c, y$.

Nun setzen wir $\delta_\varepsilon := \min\{\delta'_\varepsilon, \delta''_\varepsilon, \delta'''_\varepsilon, |y - c|\}$ und betrachten ein beliebiges $c \neq x \in (a, b) \cap (c - \delta_\varepsilon, c + \delta_\varepsilon)$. (33) gilt dann auch in diesem Fall wegen

$$\begin{aligned} \left| \frac{f(x)}{g(x)} - k \right| &= \left| \frac{f(y)}{g(x)} + \frac{f(x) - f(y)}{g(x)} - k \right| \\ &= \left| \frac{f(y)}{g(x)} + \frac{f(x) - f(y)}{g(x) - g(y)} \cdot \frac{g(x) - g(y)}{g(x)} - k \right| \\ &= \left| \frac{f(y)}{g(x)} - \frac{f(x) - f(y)}{g(x) - g(y)} \cdot \frac{g(y)}{g(x)} + \frac{f(x) - f(y)}{g(x) - g(y)} - k \right| \\ &\leq \underbrace{\left| \frac{f(y)}{g(x)} \right|}_{(38) \frac{\varepsilon}{4}} + \underbrace{\left| \frac{f(x) - f(y)}{g(x) - g(y)} \cdot \frac{g(y)}{g(x)} \right|}_{(39) \leq s} + \underbrace{\left| \frac{f(x) - f(y)}{g(x) - g(y)} - k \right|}_{(40) \frac{\varepsilon}{4 \cdot s}} \\ &= \underbrace{\left| \frac{f(y)}{g(x)} \right|}_{(38) \frac{\varepsilon}{4}} + \underbrace{\left| \frac{f(x) - f(y)}{g(x) - g(y)} \right|}_{(39) \leq s} \cdot \underbrace{\left| \frac{g(y)}{g(x)} \right|}_{(40) \frac{\varepsilon}{4 \cdot s}} + \underbrace{\left| \frac{f(x) - f(y)}{g(x) - g(y)} - k \right|}_{(34)(36)(37) \frac{\varepsilon}{2}} < \varepsilon. \end{aligned}$$

□

Bemerkung 18.24 (Die Regeln von de l'Hôpital)

a. Im Beweis von Satz 18.23 haben wir gesehen, daß die Funktion g auf dem Intervall $[a, b]$ keinen Wert *zweimal* annehmen kann, da wegen des Satzes von Rolle 18.5 die Ableitung ansonsten auch einmal Null würde. Insbesondere zeigt das, daß g höchstens eine Nullstelle haben kann! Die Bedingung $g'(x) \neq 0$ für alle $x \in (a, b)$ erzwingt also, daß auch $g(x)$ im wesentlichen ungleich Null ist.

- b. Ist g' stetig, so muß g' auf (a, b) entweder stets positiv oder stets negativ sein. Aus Proposition 18.11 folgt dann, daß g *streng monoton* auf dem Intervall (a, b) sein muß. Das zeigt, für welchen Typ von Funktionen g man die Regeln von de l'Hôpital überhaupt nur anwenden kann!
- c. Man beachte, daß die zweite Regel von de l'Hôpital 18.23 nur in der Situation $\lim_{x \rightarrow c} f(x) = \pm\infty$, d.h.

$$\frac{\lim_{x \rightarrow c} f(x)}{\lim_{x \rightarrow c} g(x)} = \frac{\pm\infty}{\pm\infty},$$

interessant ist, um den Grenzwert $\lim_{x \rightarrow c} \frac{f(x)}{g(x)}$ zu bestimmen, da für $\lim_{x \rightarrow c} f(x) = k \in \mathbb{R}$ schon aus den normalen Grenzwertsätzen

$$\lim_{x \rightarrow c} \frac{f(x)}{g(x)} = \frac{\lim_{x \rightarrow c} f(x)}{\lim_{x \rightarrow c} g(x)} = \frac{k}{\pm\infty} = 0$$

folgen würde!

Beispiel 18.25

- a. Wir betrachten die Funktionen $f = \sin$ und $g = \sqrt{\cdot}$ auf dem Intervall $(0, \infty)$. Dort sind beide differenzierbar mit

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

und

$$g'(x) = \frac{1}{2 \cdot \sqrt{x}} \neq 0$$

für alle $x \in (0, \infty)$. Aus der ersten Regel von de l'Hôpital 18.23 folgt dann

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{\sqrt{x}} = \lim_{x \rightarrow 0} \frac{f'(x)}{g'(x)} = \lim_{x \rightarrow 0} \cos(x) \cdot 2 \cdot \sqrt{x} = \cos(0) \cdot 2 \cdot \sqrt{0} = 0.$$

- b. Wir betrachten die Funktionen

$$\ln : (0, \infty) \longrightarrow \mathbb{R}$$

und

$$g : (0, \infty) \longrightarrow (0, \infty) : x \mapsto x^a = \exp(a \cdot \ln(x))$$

für ein festes $a \in \mathbb{R}_{>0}$. Nach Korollar 18.20 und Beispiel 18.22 sind beide Funktionen differenzierbar auf $(0, \infty)$.

Da sowohl $\exp(x)$, als auch $\ln(x)$ für $x \rightarrow \infty$ gegen ∞ divergieren und da a positiv ist, folgt aus den Grenzwertsätzen für uneigentliche Grenzwerte 13.17

$$\lim_{x \rightarrow \infty} g(x) = \lim_{x \rightarrow \infty} \exp(a \cdot \ln(x)) = \infty.$$

Außerdem gilt nach Beispiel 18.22

$$g'(x) = a \cdot x^{a-1} \neq 0$$

für alle $x \in (0, \infty)$. Aus der zweiten Regel von de l'Hôpital 18.23 folgt dann

$$\lim_{x \rightarrow \infty} \frac{\ln(x)}{x^a} = \lim_{x \rightarrow \infty} \frac{\ln'(x)}{g'(x)} = \lim_{x \rightarrow \infty} \frac{\frac{1}{x}}{a \cdot x^{a-1}} = \lim_{x \rightarrow \infty} \frac{1}{a \cdot g(x)} = \frac{1}{\infty} = 0.$$

Korollar 18.26 (Wachstum der Exponentialfunktion)

Die Exponentialfunktion wächst schneller als jede Polynomfunktion, d.h. ist $f = \sum_{k=0}^n a_k \cdot t^k$ ein Polynom über \mathbb{R} , so gilt

$$\lim_{x \rightarrow \infty} \frac{f(x)}{\exp(x)} = 0.$$

Beweis: Wir bezeichnen die zu f gehörige Polynomfunktion

$$f: \mathbb{R} \longrightarrow \mathbb{R}: x \mapsto \sum_{k=0}^n a_k \cdot x^k$$

wieder mit f . Dann ist f differenzierbar und die Ableitung von f ist die Polynomfunktion zum Polynom

$$f' = \sum_{k=1}^n k \cdot a_k \cdot t^{k-1}.$$

Ist $f = a_0$ konstant, so folgt die Aussage aus den Grenzwertsätzen,

$$\lim_{x \rightarrow \infty} \frac{f(x)}{\exp(x)} = \frac{a_0}{\lim_{x \rightarrow \infty} \exp(x)} = 0,$$

da $\lim_{x \rightarrow \infty} \exp(x) = \infty$.

Für ein allgemeines Polynom $f \neq 0$ führt man den Beweis am besten durch Induktion nach dem Grad n des Polynoms. Den Fall $n = 0$ haben wir bereits betrachtet. Ist $n \neq 0$, so können wir die zweite Regel von de l'Hôpital 18.23 anwenden, da $\exp'(x) = \exp(x) \neq 0$ für alle $x \in \mathbb{R}$ und da $\lim_{x \rightarrow \infty} \exp(x) = \infty$ gilt. Damit erhalten wir

$$\lim_{x \rightarrow \infty} \frac{f(x)}{\exp(x)} = \lim_{x \rightarrow \infty} \frac{f'(x)}{\exp'(x)} = \lim_{x \rightarrow \infty} \frac{f'(x)}{\exp(x)},$$

aber die rechte Seite ist dann Null nach Induktion. \square

F) Der Satz von Taylor

Definition 18.27 (Taylorpolynome)

Es sei $f: U \longrightarrow \mathbb{R}$ und $a \in U$.

Ist f n -fach differenzierbar, so nennen wir das Polynom

$$T_{f,a}^n := \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} \cdot (t - a)^k$$

das n -te Taylorpolynom von f mit Entwicklungspunkt a .

Ist f unendlich oft differenzierbar, so nennen wir die Potenzreihe

$$T_{f,a} := \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} \cdot (t - a)^k$$

die *Taylorreihe* von f mit Entwicklungspunkt \mathbf{a} oder die *Taylor-Entwicklung* von f im Punkt \mathbf{a} . Beachte, stets gilt $T_{f,\mathbf{a}}^n(\mathbf{a}) = T_{f,\mathbf{a}}(\mathbf{a}) = f(\mathbf{a})$.

Bemerkung 18.28 (Tangenten und das 1. Taylorpolynome)

Ist $f: \mathbb{R} \rightarrow \mathbb{R}$ differenzierbar, so ist die Gleichung der Tangente an den Graphen von f im Punkt $(\mathbf{a}, f(\mathbf{a}))$ gegeben durch

$$y = f'(\mathbf{a}) \cdot (x - \mathbf{a}) + f(\mathbf{a}) = T_{f,\mathbf{a}}^1(x).$$

D.h. das erste Taylorpolynom von f mit Entwicklungspunkt \mathbf{a} ist die optimale lineare Approximation der Funktion f lokal in \mathbf{a} .

Die Idee ist nun, daß mit steigendem n die Taylorpolynome $T_{f,\mathbf{a}}^n$ immer bessere Approximationen von f lokal in \mathbf{a} sein werden, und daß im Grenzwert dann die Taylorreihe vielleicht sogar mit f übereinstimmt. Das wird nicht immer aber doch oft der Fall sein – siehe Beispiel 18.30! Funktionen, für die das gilt, nennt man *analytisch* in \mathbf{a} .

Man kann die Theorie der Differenzierbarkeit statt für Funktionen auf \mathbb{R} auch für Funktionen auf \mathbb{C} einführen. In der Vorlesung Einführung in die Funktionentheorie wird das getan, und dort zeigt man, daß über \mathbb{C} jede einmal auf \mathbb{C} differenzierbare Funktion schon analytisch ist, d.h. durch eine Potenzreihe gegeben und damit unendlich oft differenzierbar ist! Die komplexen Zahlen verhalten sich also weit unkomplizierter als die reellen Zahlen.

Beispiel 18.29 (Potenzreihen als Taylorreihen)

Ist $\sum_{n=0}^{\infty} a_n \cdot t^n$ eine Potenzreihe auf \mathbb{R} mit Konvergenzradius $r > 0$, so ist die Funktion

$$f: (-r, r) \rightarrow \mathbb{R}: x \mapsto \sum_{n=0}^{\infty} a_n \cdot x^n$$

nach Korollar 18.19 unendlich oft differenzierbar, und mittels Induktion nach n zeigt man, daß

$$f^{(n)}(0) = n! \cdot a_n.$$

Damit stimmt f also mit seiner Taylorreihe

$$T_{f,0} = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} \cdot t^n = \sum_{n=0}^{\infty} a_n \cdot t^n$$

auf dem Konvergenzbereich $(-r, r)$ überein, und die Taylorpolynome

$$T_{f,0}^n = \sum_{k=0}^n a_k \cdot t^k$$

definieren eine Folge von Funktionen, die auf jedem abgeschlossenen Intervall $[-R, R] \subseteq (-r, r)$ gleichmäßig gegen f konvergieren.

Beispiel 18.30

Die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \exp\left(-\frac{1}{x^2}\right), & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0, \end{cases}$$

ist unendlich oft differenzierbar mit $f^{(n)}(0) = 0$ für alle $n \in \mathbb{N}$. Insbesondere ist die Taylorreihe von f also Null,

$$T_{f,0} = 0.$$

In diesem Fall stimmt die Taylorreihe also nur im Punkt $x = 0$ mit der Funktion überein, da $f(x) \neq 0$ für alle $x \neq 0$. Der Beweis ist Aufgabe 18.38.

Der Satz von Taylor sagt etwas darüber aus, wie gut das n -te Taylorpolynom f approximiert.

Satz 18.31 (Satz von Taylor – Restglied nach Lagrange)

Sei I ein Intervall, $f : I \longrightarrow \mathbb{R}$ eine $n+1$ -fach differenzierbare Funktion und $x, a \in I$. Dann gibt es ein c zwischen x und a , so daß

$$f(x) - T_{f,a}^n(x) = \frac{f^{(n+1)}(c)}{(n+1)!} \cdot (x-a)^{n+1}.$$

Wir nennen die rechte Seite auch das Restglied des n -ten Taylorpolynoms.

Beweis: Wir können ohne Einschränkung $x > a$ annehmen.

Dann definieren wir eine reelle Zahl

$$z := \frac{(f(x) - T_{f,a}^n(x)) \cdot (n+1)!}{(x-a)^{n+1}} \in \mathbb{R}$$

und eine Funktion $g : [a, x] \longrightarrow \mathbb{R}$ durch

$$\begin{aligned} g(y) &:= f(x) - T_{f,y}^n(x) - \frac{z}{(n+1)!} \cdot (x-y)^{n+1} \\ &= f(x) - \sum_{k=0}^n \frac{f^{(k)}(y)}{k!} \cdot (x-y)^k - \frac{z}{(n+1)!} \cdot (x-y)^{n+1} \\ &= f(x) - f(y) - \sum_{k=1}^n \frac{f^{(k)}(y)}{k!} \cdot (x-y)^k - \frac{z}{(n+1)!} \cdot (x-y)^{n+1} \end{aligned}$$

für $y \in [a, x]$ — man beachte hier, daß g eine Funktion in der Veränderlichen y ist, während x konstant ist!

Nach Voraussetzung ist f $n+1$ -fach differenzierbar auf I , so daß die Funktion g differenzierbar auf $[a, x]$ ist, und mit Hilfe der Produktregel erhalten wir für die

Ableitung

$$\begin{aligned}
 g'(y) &= -f'(y) - \sum_{k=1}^n \left(\frac{f^{(k+1)}(y)}{k!} \cdot (x-y)^k - \frac{f^{(k)}(y)}{k!} \cdot k \cdot (x-y)^{k-1} \right) + \frac{z \cdot (n+1)}{(n+1)!} \cdot (x-y)^n \\
 &= -f'(y) - \sum_{k=1}^n \left(\frac{f^{(k+1)}(y)}{k!} \cdot (x-y)^k - \frac{f^{(k)}(y)}{(k-1)!} \cdot (x-y)^{k-1} \right) + \frac{z}{n!} \cdot (x-y)^n \\
 &= -f'(y) - \left(\frac{f^{(n+1)}(y)}{n!} \cdot (x-y)^n - f'(y) \right) + \frac{z}{n!} \cdot (x-y)^n \\
 &= \frac{z}{n!} \cdot (x-y)^n + \frac{f^{(n+1)}(y)}{n!} \cdot (x-y)^n = \frac{z - f^{(n+1)}(y)}{n!} \cdot (x-y)^n,
 \end{aligned}$$

wobei wir beachten, daß die Summe in der zweiten Zeile eine Teleskopsumme ist, so daß nur die Randsummanden übrig bleiben.

Zudem folgt aus der Definition von z

$$g(a) = f(x) - T_{f,a}^n(x) - \frac{z}{(n+1)!} \cdot (x-a)^{n+1} = 0,$$

und aus der Definition des Taylorpolynoms folgt

$$g(x) = f(x) - \sum_{k=0}^n \frac{f^{(k)}(x)}{k!} \cdot (x-x)^k - \frac{z}{(n+1)!} \cdot (x-x)^{n+1} = f(x) - f(x) = 0.$$

Wir können also den Satz von Rolle 18.5 anwenden und finden ein $c \in (a, x)$ mit

$$0 = g'(c) = \frac{z - f^{(n+1)}(c)}{n!} \cdot (x-c)^n.$$

Da $x - c \neq 0$, muß

$$f^{(n+1)}(c) = z = \frac{(f(x) - T_{f,a}^n(x)) \cdot (n+1)!}{(x-a)^{n+1}}$$

gelten, und damit

$$f(x) - T_{f,a}^n(x) = \frac{f^{(n+1)}(c)}{(n+1)!} \cdot (x-a)^{n+1}.$$

□

Beispiel 18.32 (Näherungswert für die Eulersche Zahl e)

Wir betrachten die Funktion $f = \exp$, $x = 1$ und $a = 0$. Dann ist $f^{(n+1)} = \exp$ und das n -te Taylorpolynom erfüllt

$$T_{\exp,0}^n(1) = T_{f,a}^n(1) = \sum_{k=0}^n \frac{1}{k!}.$$

Mit Hilfe des Satzes von Taylor finden wir ein $c \in (0, 1)$ mit

$$\left| e - \sum_{k=0}^n \frac{1}{k!} \right| = |\exp(1) - T_{\exp,0}^n(1)| = \frac{|\exp(c)|}{(n+1)!} \cdot |1-0|^{n+1} < \frac{e}{(n+1)!} < \frac{3}{(n+1)!},$$

wenn wir ausnutzen, daß die Exponentialfunktion streng monoton wachsend und positiv auf $[0, 1]$ ist. Wenden wir diese Abschätzung mit $n = 6$ an, so erhalten wir

$$\left| e - \frac{1957}{720} \right| < \frac{1}{1680} < \frac{1}{1000}.$$

Die Dezimalzahldarstellung von $\frac{1957}{720}$ stimmt also bis zur dritten Nachkommastelle mit der Zahl e überein, und daraus ersehen wir:

$$e = 2,718\dots$$

Beispiel 18.33 (Taylor-Entwicklung des natürlichen Logarithmus)

Wir wissen, daß der natürliche Logarithmus

$$\ln : (0, \infty) \longrightarrow \mathbb{R}$$

unendlich oft differenzierbar mit Ableitung

$$\ln' : (0, \infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x}$$

ist. Eine einfache Induktion zeigt, daß für $n \geq 1$ dann

$$\ln^{(n)} : (0, \infty) \longrightarrow \mathbb{R} : x \mapsto (-1)^{n-1} \cdot \frac{(n-1)!}{x^n}$$

gilt. Das n -te Taylorpolynom mit Entwicklungspunkt 1 ist mithin

$$T_{\ln,1}^n(x) = \sum_{k=0}^n \frac{\ln^{(k)}(1)}{k!} \cdot (x-1)^k = \sum_{k=1}^n (-1)^{k-1} \cdot \frac{(x-1)^k}{k}.$$

Der Betrag aller Ableitungen von \ln ist auf dem Intervall $[1, 2]$ streng monoton fallend, so daß insbesondere

$$|\ln^{(n+1)}(c)| \leq |\ln^{(n+1)}(1)| = n!$$

für jedes $c \in [1, 2]$ gilt. Mit dem Satz von Taylor finden wir zu $x \in [1, 2]$ ein c zwischen 1 und $x \leq 2$, so daß

$$|\ln(x) - T_{\ln,x}^n(x)| = \frac{|\ln^{(n+1)}(c)|}{(n+1)!} \cdot |x-1|^{n+1} \leq \frac{1}{n+1}.$$

Auf dem Intervall $[1, 2]$ konvergiert die durch die Taylorpolynome definierte Funktionenfolge mithin gleichmäßig gegen die Funktion \ln , und zugleich konvergiert sie dort gleichmäßig gegen die durch die Taylorreihe definierte Funktion, d.h. für $x \in [1, 2]$ gilt

$$\ln(x) = \sum_{n=1}^{\infty} (-1)^{n-1} \cdot \frac{(x-1)^n}{n}.$$

Werten wir diese Gleichheit in $x = 2$ aus, so erhalten wir den Wert für die alternierende harmonische Reihe als

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -\ln(2).$$

G) Allgemeinere Bedingung für Extremstellen

Satz 18.34 (Allgemeinere Bedingung für Extremstellen)

Es sei $f : (a, b) \rightarrow \mathbb{R}$ eine $n + 1$ -fach stetig differenzierbare Funktion mit $n \in \mathbb{N}$ und es sei $c \in (a, b)$ mit $f'(c) = f''(c) = \dots = f^{(n)}(c) = 0$ und $f^{(n+1)}(c) \neq 0$.

- Falls n gerade ist, so ist c keine Extremstelle von f .
- Falls n ungerade und $f^{(n+1)}(c) < 0$, so ist c ein lokales Maximum.
- Falls n ungerade und $f^{(n+1)}(c) > 0$, so ist c ein lokales Minimum.

Beweis: Wir können uns ohne Einschränkung auf den Fall $f^{(n+1)}(c) > 0$ beschränken.

Weil $f^{(n+1)}$ stetig ist und in c nicht den Wert Null annimmt, muß es nach Aufgabe 14.30 ein $\varepsilon > 0$ geben, so daß

$$f^{(n+1)}(x) \neq 0$$

für alle $x \in (c - \varepsilon, c + \varepsilon)$. Da $f^{(n+1)}$ stetig ist, garantiert der Zwischenwertsatz 14.12 dann, daß $f^{(n+1)}$ in der Tat strikt positiv auf diesem Intervall ist.

Wir betrachten das n -te Taylorpolynom

$$T_{f,c}^n = \sum_{k=0}^n \frac{f^{(k)}(c)}{k!} \cdot (t - c)^k = f(c) + \sum_{k=1}^n \frac{0}{k!} \cdot (t - c)^k = f(c)$$

von f mit Entwicklungspunkt c sowie ein $c \neq x \in (c - \varepsilon, c + \varepsilon)$. Aus dem Satz von Taylor erhalten wir ein d_x zwischen c und x , so daß

$$f(x) - f(c) = f(x) - T_{f,c}^n(x) = \frac{f^{(n+1)}(d_x)}{(n+1)!} \cdot (x - c)^{n+1}.$$

Man beachte, daß $d_x \in (c - \varepsilon, c + \varepsilon)$ liegen muß, so daß $f^{(n+1)}(d_x) > 0$ ist.

Ist n ungerade, so ist auch $(x - c)^{n+1}$ stets positiv, und wir erhalten

$$f(x) - f(c) \geq 0$$

für alle $x \in (c - \varepsilon, c + \varepsilon)$, d.h. c ist ein Minimum.

Ist n gerade, so wechselt $(x - c)^{n+1}$ in c das Vorzeichen, d.h.

$$f(x) - f(c) \begin{cases} < 0 & \text{falls } x < c, \\ > 0 & \text{falls } x > c. \end{cases}$$

In c liegt also keine Extremstelle vor. □

Im Fall $n = 1$ lautet die Bedingung für eine Extremstelle einfach $f'(c) = 0$ und $f''(c) \neq 0$ und stimmt mit der Bedingung in Proposition 18.13 überein. Aber wir haben diesmal vorausgesetzt, daß die zweite Ableitung *stetig* ist. Darauf konnten wir in Proposition 18.13 verzichten.

Beispiel 18.35

Betrachten wir die Funktion $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^4$, dann ist $\mathbf{a} = 0$ offenbar ein globales Minimum, da $f(x) > 0$ für alle $x \neq 0$ gilt. Proposition 18.13 hätte uns dieses Minimum nicht offenbart, da

$$f'(0) = 4 \cdot 0^3 = 0, \quad f''(0) = 12 \cdot 0^2 = 0 \quad \text{und} \quad f'''(0) = 24 \cdot 0 = 0$$

gilt. Aber wegen $f^{(4)}(0) = 24 > 0$ weist Satz 18.34 $\mathbf{a} = 0$ als ein wenn auch nur lokales Minimum aus.

Bemerkung 18.36

Man beachte, daß die Funktion in Beispiel 18.30 im Punkt $\mathbf{a} = 0$ ein globales Minimum besitzt und unendlich oft differenzierbar ist, daß aber alle Ableitungen im Punkt $\mathbf{a} = 0$ verschwinden! Satz 18.34 ist also nicht immer anwendbar, um Extremstellen zu finden.

Aufgaben**Aufgabe 18.37**

Es sei $f : [\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ stetig auf $[\mathbf{a}, \mathbf{b}]$ und differenzierbar auf (\mathbf{a}, \mathbf{b}) mit $f'(x) = 0$ für alle $x \in (\mathbf{a}, \mathbf{b})$, dann ist f eine konstante Funktion.

Aufgabe 18.38

Zeige, daß für die Funktion

$$f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \begin{cases} \exp\left(-\frac{1}{x^2}\right) & \text{für } x \neq 0, \\ 0 & \text{für } x = 0. \end{cases}$$

die folgenden Aussagen gelten:

- a. Für alle $n \geq 1$ gibt es ein Polynom $p_n \in \mathbb{R}[t]$, so daß für $x \neq 0$ gilt:

$$f^{(n)}(x) = \frac{p_n(x)}{x^{3 \cdot 2^{n-1}}} \cdot \exp\left(-\frac{1}{x^2}\right).$$

- b. Für alle $k \in \mathbb{N}$ gilt $\lim_{x \rightarrow 0} \frac{\exp(-\frac{1}{x^2})}{x^k} = 0$.

- c. Für alle $n \in \mathbb{N}$ gilt $f^{(n)}(0) = 0$.

- d. $f \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ und $T_{f,0} = 0$.

Aufgabe 18.39 (Näherungsweise Berechnung von π)

Betrachte die Funktion $f = \arctan$ auf \mathbb{R} .

- a. Berechne das dritte Taylorpolynom $T_{f,0}^3$ von \arctan mit Entwicklungspunkt 0.
 b. Benutze $T_{f,0}^3$ und Aufgabe 16.20 c., um $\frac{\pi}{4}$ und damit π näherungsweise zu bestimmen. Zeige dabei, daß die in der Näherung bis auf zwei Nachkommastellen exakt ist mit

$$\pi = 3,14 \dots$$

Aufgabe 18.40

Berechne die folgenden Grenzwerte:

- $\lim_{x \rightarrow \pi} \frac{\cos(x)+1}{x^2-\pi^2}$ mit $x < \pi$.
- $\lim_{x \rightarrow 1} \left(\frac{1}{x-1} - \frac{1}{\ln(x)} \right)$ mit $x > 1$.
- $\lim_{x \rightarrow 0} x^x$ mit $x > 0$.

Aufgabe 18.41

Bestimme alle Extrema der Funktion $f : [0, 1] \rightarrow \mathbb{R}$, $x \mapsto (1-x) \cdot \sqrt{1+9x^2}$.

Aufgabe 18.42

Berechne für die Funktion

$$f : \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), x \mapsto \arctan(x)$$

das zweite Taylor-Polynom $T_{f,0}^2$ um 0 und gib eine Abschätzung für das Restglied $|f(x) - T_{f,0}^2(x)|$ auf dem Intervall $[-1, 1]$ an.

Aufgabe 18.43

Berechne für die Funktion

$$f : \left(-\frac{1}{2}, \frac{1}{2}\right) \rightarrow \mathbb{R}, x \mapsto \frac{\cos(x)}{1-(2x)^4}$$

das vierte Taylor-Polynom $T_{f,0}^4$ um 0.

Hinweis: mit etwas Überlegung kann man die Berechnung aller vier Ableitungen von f vermeiden.

Aufgabe 18.44

Beweise oder widerlege durch ein Gegenbeispiel die folgenden Aussagen:

- Sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige und auf (a, b) differenzierbare Funktion, so daß $\lim_{x \rightarrow a} f'(x)$ existiert. Dann ist f differenzierbar in a .
- Sei $f \in \mathcal{C}^1([a, b], \mathbb{R})$ eine stetig differenzierbare streng monoton wachsende Funktion. Dann gilt $f'(x) > 0$ für alle $x \in [a, b]$.

§ 19 Das Riemann-Integral

Wir werden in diesem Abschnitt im wesentlichen nur Funktionen betrachten, die auf einem *abgeschlossenen Intervall* $[a, b]$ definiert und die dort *beschränkt* sind.

A) Riemann-integrierbare Funktionen

Definition 19.1 (Zerlegungen eines Intervalls)

Es seien $a, b \in \mathbb{R}$ mit $a < b$. Ein Tupel $Z = (x_0, \dots, x_n)$ mit $n \geq 1$ heißt eine *Zerlegung* des Intervalls $[a, b]$, falls

$$a = x_0 < x_1 < \dots < x_{n-1} < x_n = b.$$

Die Zahl $l(Z) := \max\{x_i - x_{i-1} \mid i = 1, \dots, n\}$ heißt die *Länge* oder *Feinheit* der Zerlegung, die Menge $\text{supp}(Z) := \{x_0, \dots, x_n\}$ ihr *Träger*, die Zahl $|Z| := n$ ihre *Mächtigkeit* und die x_i ihre *Stützpunkte*.

Eine zweite Zerlegung $Z' = (y_0, \dots, y_m)$ von $[a, b]$ heißt *Verfeinerung* von Z , falls

$$\{x_0, \dots, x_n\} \subseteq \{y_0, \dots, y_m\}.$$

Zu zwei Zerlegungen $Z = (x_0, \dots, x_n)$ und $Z' = (y_0, \dots, y_m)$ definieren wir

$$Z * Z' := (z_0, \dots, z_k),$$

indem wir die Elemente der Vereinigung $\text{supp}(Z) \cup \text{supp}(Z') = \{z_0, \dots, z_k\}$ der Größe nach ordnen. Sind Z und Z' Zerlegungen des gleichen Intervalls, so nennen wir $Z * Z'$ ihre *gemeinsame Verfeinerung*.

Beispiel 19.2

Die Tupel $Z = (0, 1, 3, 5)$ und $Z' = (0, 2, 5)$ sind Zerlegungen von $[0, 5]$ der Länge 2 bzw. 3, und die gemeinsame Verfeinerung von Z und Z' ist $Z * Z' = (0, 1, 2, 3, 5)$.

Definition 19.3 (Obersummen und Untersummen)

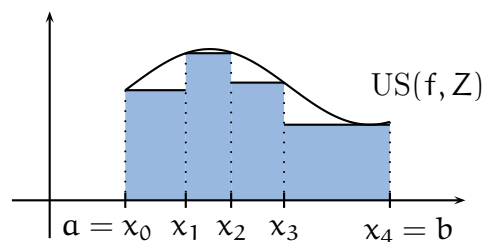
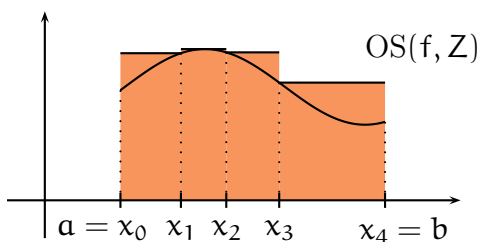
Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt, $a < b$, und $Z = (x_0, \dots, x_n)$ eine Zerlegung von $[a, b]$.

Wir definieren die *Obersumme* von f bezüglich Z als

$$OS(f, Z) := \sum_{i=1}^n (x_i - x_{i-1}) \cdot \sup\{f(x) \mid x \in [x_{i-1}, x_i]\},$$

und die *Untersumme* von f bezüglich Z als

$$US(f, Z) := \sum_{i=1}^n (x_i - x_{i-1}) \cdot \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}.$$



Beispiel 19.4

Wir betrachten die Identität $\text{id} : [0, 1] \rightarrow \mathbb{R} : x \mapsto x$ auf dem Intervall $[0, 1]$ sowie die folgende äquidistante Zerlegung der Länge $\frac{1}{n}$

$$Z^n = (x_0, \dots, x_n) = \left(0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\right).$$

Auf einem Teilintervall $[x_{i-1}, x_i] = \left[\frac{i-1}{n}, \frac{i}{n}\right]$ gilt dann

$$m_i := \inf \{ \text{id}(x) \mid x \in [x_{i-1}, x_i] \} = x_{i-1} = \frac{i-1}{n}$$

und

$$M_i := \sup \{ \text{id}(x) \mid x \in [x_{i-1}, x_i] \} = x_i = \frac{i}{n}.$$

Für die Unter- und Obersumme von f bezüglich Z ergibt sich unter Berücksichtigung von Beispiel 7.11 damit

$$\text{US}(f, Z^n) = \sum_{i=1}^n (x_i - x_{i-1}) \cdot m_i = \sum_{i=1}^n \frac{1}{n} \cdot \frac{i-1}{n} = \frac{n \cdot (n-1)}{2 \cdot n^2} = \frac{1}{2} - \frac{1}{2n}$$

und

$$\text{OS}(f, Z^n) = \sum_{i=1}^n (x_i - x_{i-1}) \cdot M_i = \sum_{i=1}^n \frac{1}{n} \cdot \frac{i}{n} = \frac{n \cdot (n+1)}{2 \cdot n^2} = \frac{1}{2} + \frac{1}{2n}.$$

Lemma 19.5

Es sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt mit $|f(x)| \leq M$ für alle $x \in [a, b]$, $a < b$.

- a. Ist Z' eine Verfeinerung der Zerlegung Z von $[a, b]$, so gelten

$$0 \leq \text{US}(f, Z') - \text{US}(f, Z) \leq 2 \cdot M \cdot l(Z) \cdot (|Z'| - |Z|)$$

und

$$0 \leq \text{OS}(f, Z) - \text{OS}(f, Z') \leq 2 \cdot M \cdot l(Z) \cdot (|Z'| - |Z|).$$

Insbesondere gilt also

$$\text{US}(f, Z) \leq \text{US}(f, Z') \leq \text{OS}(f, Z') \leq \text{OS}(f, Z).$$

- b. Für je zwei Zerlegungen Z und Z' von $[a, b]$ gilt

$$\text{US}(f, Z) \leq \text{OS}(f, Z').$$

- c. Es gelten

$$-M \cdot (b - a) \leq \text{US}(f, Z) \leq \text{OS}(f, Z) \leq (b - a) \cdot M.$$

Beweis:

- a. Es sei $Z = (x_0, \dots, x_n)$, und wir setzen für $i = 1, \dots, n$ wieder

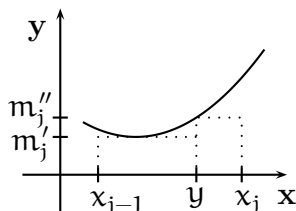
$$m_i := \inf \{ f(x) \mid x \in [x_{i-1}, x_i] \}.$$

Wir zeigen die Aussage zu den Untersummen für den Fall, daß Z' einen Punkt mehr enthält als Z . Sei also $Z' = (x_0, \dots, x_{j-1}, y, x_j, \dots, x_n)$. Dann gilt

$$m'_j := \inf \{f(x) \mid x \in [x_{j-1}, y]\} \geq m_j$$

und

$$m''_j := \inf \{f(x) \mid x \in [y, x_j]\} \geq m_j.$$



Daraus ergibt sich

$$\begin{aligned} \text{US}(f, Z) &= \sum_{i \neq j} (x_i - x_{i-1}) \cdot m_i + (x_j - x_{j-1}) \cdot m_j \\ &= \sum_{i \neq j} (x_i - x_{i-1}) \cdot m_i + (y - x_{j-1}) \cdot m_j + (x_j - y) \cdot m_j \\ &\leq \sum_{i \neq j} (x_i - x_{i-1}) \cdot m_i + (y - x_{j-1}) \cdot m'_j + (x_j - y) \cdot m''_j \\ &= \text{US}(f, Z'). \end{aligned}$$

Für die Differenz der beiden Terme erhalten wir

$$\begin{aligned} 0 &\leq \text{US}(f, Z') - \text{US}(f, Z) \\ &= (y - x_{j-1}) \cdot (m'_j - m_j) + (x_j - y) \cdot (m''_j - m_j) \\ &\leq (y - x_{j-1}) \cdot (M + M) + (x_j - y) \cdot (M + M) \\ &= (x_j - x_{j-1}) \cdot 2 \cdot M \leq 2 \cdot M \cdot l(Z). \end{aligned}$$

Für eine beliebige Verfeinerung Z' von Z wenden wir dann Induktion an und erhalten die Formel

$$0 \leq \text{US}(f, Z') - \text{US}(f, Z) \leq 2 \cdot M \cdot l(Z) \cdot (|Z'| - |Z|)$$

Die Aussage für Obersummen zeigt man analog.

- b. Wir betrachten die gemeinsame Verfeinerung $Z * Z' = (y_0, \dots, y_k)$. Wegen

$$m_i := \inf \{f(x) \mid x \in [y_{i-1}, y_i]\} \leq \sup \{f(x) \mid x \in [y_{i-1}, y_i]\} =: M_i$$

folgt dann

$$\begin{aligned} \text{US}(f, Z) &\stackrel{\text{a.}}{\leq} \text{US}(f, Z * Z') = \sum_{i=1}^k (y_i - y_{i-1}) \cdot m_i \\ &\leq \sum_{i=1}^k (y_i - y_{i-1}) \cdot M_i = \text{OS}(f, Z * Z') \stackrel{\text{a.}}{\leq} \text{OS}(f, Z'). \end{aligned}$$

- c. Die Aussage folgt aus a., da Z eine Verfeinerung der Zerlegung (\mathbf{a}, \mathbf{b}) ist und da $M \geq \sup \{f(x) \mid x \in [\mathbf{a}, \mathbf{b}]\} \geq \inf \{f(x) \mid x \in [\mathbf{a}, \mathbf{b}]\} \geq -M$.

□

Beispiel 19.6

In Beispiel 19.4 gilt $US(\text{id}, Z^n) = \frac{1}{2} - \frac{1}{2n} \leq \frac{1}{2} + \frac{1}{2n} = OS(\text{id}, Z^n)$.

Da die Menge der Obersummen und die Menge der Untersummen nach Lemma 19.5 c. beschränkt sind, können wir ihr Infimum und ihr Supremum betrachten.

Definition 19.7 (Riemann-integrierbar)

Es sei $f : [\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ beschränkt, $\mathbf{a} < \mathbf{b}$. Wir definieren das *Oberintegral*

$$OI(f) := \inf \{ OS(f, Z) \mid Z \text{ Zerlegung von } [\mathbf{a}, \mathbf{b}] \}$$

von f und das *Unterintegral*

$$UI(f) := \sup \{ US(f, Z) \mid Z \text{ Zerlegung von } [\mathbf{a}, \mathbf{b}] \}$$

von f . Wegen Lemma 19.5 b. und Lemma 8.19 gilt

$$UI(f) \leq OI(f).$$

Wir nennen f (Riemann-)integrierbar auf $[\mathbf{a}, \mathbf{b}]$, falls $UI(f) = OI(f)$. Dann heißt

$$\int_{\mathbf{a}}^{\mathbf{b}} f(x) \, dx := OI(f) \in \mathbb{R}$$

das *Integral* von f auf $[\mathbf{a}, \mathbf{b}]$.

Beispiel 19.8

Aus Beispiel 19.4 wissen wir für $\text{id} : [0, 1] \rightarrow \mathbb{R}$

$$\frac{1}{2} - \frac{1}{2n} = US(\text{id}, Z^n) \leq UI(\text{id}) \leq OI(\text{id}) \leq OS(\text{id}, Z^n) = \frac{1}{2} + \frac{1}{2n}$$

für alle $n \in \mathbb{N}$. Bilden wir nun den Grenzwert für n gegen unendlich, so erhalten wir

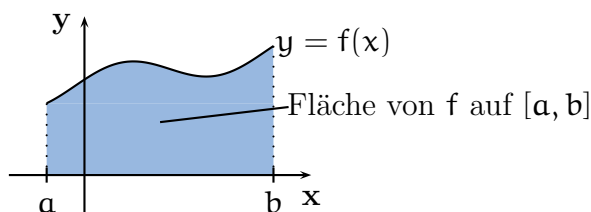
$$\frac{1}{2} = \lim_{n \rightarrow \infty} US(\text{id}, Z^n) \leq UI(\text{id}) \leq OI(\text{id}) \leq \lim_{n \rightarrow \infty} OS(\text{id}, Z^n) = \frac{1}{2},$$

d.h. id ist integrierbar auf $[0, 1]$ mit

$$\int_0^1 x \, dx = OI(\text{id}) = UI(\text{id}) = \frac{1}{2}.$$

Bemerkung 19.9 (Das Riemann-Integral als Flächeninhalt)

Wenn die Funktion nur nicht-negative Werte annimmt, dann sind die Untersummen von f nach oben beschränkt durch den Flächeninhalt I der Fläche, die der Graph von f mit der x -Achse einschließt, und die Obersummen von f sind durch diesen nach unten beschränkt. Aufgrund der Definition von $OI(f)$ als Infimum und $UI(f)$ als Supremum gilt also stets $UI(f) \leq I \leq OI(f)$. Daß f integrierbar ist, bedeutet mithin nichts anderes, als daß das Integral $\int_{\mathbf{a}}^{\mathbf{b}} f(x) \, dx$ den Flächeninhalt der Fläche beschreibt, die der Graph von f auf dem Intervall $[\mathbf{a}, \mathbf{b}]$ mit der x -Achse einschließt.



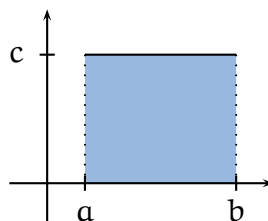
Beispiel 19.10

a. Jede konstante Funktion $f : [a, b] \rightarrow \mathbb{R} : x \mapsto c$ ist integrierbar mit

$$\int_a^b f(x) \, dx = (b - a) \cdot c.$$

Denn dann gilt für jede Zerlegung $Z = (x_0, \dots, x_n)$ von $[a, b]$ bereits

$$OS(f, Z) = \sum_{i=1}^n (x_i - x_{i-1}) \cdot c = (b - a) \cdot c = \sum_{i=1}^n (x_i - x_{i-1}) \cdot c = US(f, Z).$$



b. Die *Dirichletsche Sprungfunktion*

$$f : [0, 1] \rightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x \in \mathbb{Q}, \\ 1, & \text{falls } x \notin \mathbb{Q} \end{cases}$$

ist *nicht* integrierbar auf $[0, 1]$. Denn ist $Z = (x_0, \dots, x_n)$ eine beliebige Zerlegung von $[0, 1]$, so gibt es im Intervall $[x_{i-1}, x_i]$ sowohl eine rationale Zahl, als auch eine irrationale. Mithin gilt

$$US(f, Z) = \sum_{i=1}^n (x_i - x_{i-1}) \cdot 0 = 0$$

und

$$OS(f, Z) = \sum_{i=1}^n (x_i - x_{i-1}) \cdot 1 = 1$$

für jede Zerlegung Z , so daß

$$UI(f) = 0 < 1 = OI(f).$$

B) Das Riemannsches Integrierbarkeitskriterium

Satz 19.11 (Riemannsches Integrierbarkeitskriterium)

Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt, $a < b$. Genau dann ist f integrierbar auf $[a, b]$, wenn

$$\forall \varepsilon > 0 \exists Z \text{ Zerlegung von } [a, b] : OS(f, Z) - US(f, Z) < \varepsilon.$$

Beweis: \implies : Sei zunächst f integrierbar auf $[a, b]$ und sei $\varepsilon > 0$ gegeben. Aufgrund der Definition von $UI(f)$ als Supremum und $OI(f)$ als Infimum und wegen Proposition 8.18 gibt es Zerlegungen Z' und Z'' von $[a, b]$ mit

$$OI(f) + \frac{\varepsilon}{2} > OS(f, Z') \stackrel{19.5}{\geq} OS(f, Z' * Z'')$$

und

$$UI(f) - \frac{\varepsilon}{2} < US(f, Z'') \stackrel{19.5}{\leq} US(f, Z' * Z'').$$

Damit erhalten wir mit $Z = Z' * Z''$ und wegen $UI(f) = OI(f)$

$$OS(f, Z) - US(f, Z) < \left(OI(f) + \frac{\varepsilon}{2}\right) - \left(UI(f) - \frac{\varepsilon}{2}\right) = \varepsilon.$$

\impliedby : Für $\varepsilon := \frac{1}{n}$ mit $n \geq 1$ gibt es eine Zerlegung Z^n von $[a, b]$ mit

$$\frac{1}{n} > OS(f, Z^n) - US(f, Z^n) \geq OI(f) - UI(f) \geq 0.$$

Da die linke Seite der Ungleichung für $n \rightarrow \infty$ gegen Null konvergiert, folgt im Grenzwert

$$0 = \lim_{n \rightarrow \infty} \frac{1}{n} \geq OI(f) - UI(f) \geq 0,$$

also $OI(f) = UI(f)$. Mithin ist f integrierbar auf $[a, b]$. \square

Satz 19.12 (Stetige Funktionen sind integrierbar.)

Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig, $a < b$, so ist f integrierbar auf $[a, b]$.

Beweis: Da f stetig auf dem abgeschlossenen Intervall $[a, b]$ ist, ist f dort beschränkt nach Proposition 14.15 und gleichmäßig stetig nach Satz 14.28.

Sei nun $\varepsilon > 0$ gegeben. Da f gleichmäßig stetig auf $[a, b]$ ist, gibt es $\delta_\varepsilon > 0$, so daß

$$|f(x) - f(y)| < \frac{\varepsilon}{b - a} \tag{41}$$

für alle $x, y \in [a, b]$ mit $|x - y| < \delta_\varepsilon$. Wir wählen nun eine Zerlegung $Z = (x_0, \dots, x_n)$ mit Länge $l(Z) < \delta_\varepsilon$. Da f stetig auf $[x_{i-1}, x_i]$ ist, existieren $y_i, z_i \in [x_{i-1}, x_i]$ mit

$$f(y_i) = \sup \{f(y) \mid y \in [x_{i-1}, x_i]\}$$

und

$$f(z_i) = \inf \{f(y) \mid y \in [x_{i-1}, x_i]\},$$

und wegen $|y_i - z_i| \leq |x_i - x_{i-1}| < \delta_\varepsilon$ folgt aus (41) zudem

$$0 \leq f(y_i) - f(z_i) < \frac{\varepsilon}{b - a}.$$

Damit erhalten wir insbesondere

$$\begin{aligned} OS(f, Z) - US(f, Z) &= \sum_{i=1}^n (x_i - x_{i-1}) \cdot (f(y_i) - f(z_i)) \\ &< \sum_{i=1}^n (x_i - x_{i-1}) \cdot \frac{\varepsilon}{b - a} = (b - a) \cdot \frac{\varepsilon}{b - a} = \varepsilon. \end{aligned}$$

Somit ist f integrierbar nach dem Riemannsches Integrierbarkeitskriterium 19.11. \square

Beispiel 19.13

Wir betrachten die Funktion

$$f : [0, 1] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x = 0, \\ 1, & \text{falls } x \neq 0 \end{cases}$$

und die Zerlegung $Z^n = (0, \frac{1}{n}, 1)$ für $n \geq 1$. Dann gilt

$$\text{US}(f, Z^n) = \left(\frac{1}{n} - 0\right) \cdot 0 + \left(1 - \frac{1}{n}\right) \cdot 1 = 1 - \frac{1}{n}$$

und

$$\text{OS}(f, Z^n) = \left(\frac{1}{n} - 0\right) \cdot 1 + \left(1 - \frac{1}{n}\right) \cdot 1 = 1.$$

Wir erhalten also

$$1 \longleftarrow 1 - \frac{1}{n} = \text{US}(f, Z^n) \leq \text{UI}(f) \leq \text{OI}(f) \leq \text{OS}(f, Z^n) = 1.$$

Mithin ist f auf $[0, 1]$ integrierbar mit

$$\int_0^1 f(x) \, dx = 1.$$

Dies zeigt, daß eine Funktion nicht stetig sein muß, um integrierbar zu sein.

Proposition 19.14 (Monotone Funktionen sind integrierbar.)

Ist $f : [a, b] \longrightarrow \mathbb{R}$ monoton wachsend oder fallend, $a < b$, so ist f integrierbar.

Beweis: Wir können ohne Einschränkung annehmen, daß f monoton wachsend und nicht konstant ist. Insbesondere ist $f(b) > f(a)$. Außerdem ist f beschränkt, da $f(a) \leq f(x) \leq f(b)$ für alle $x \in [a, b]$.

Sei $\varepsilon > 0$ gegeben. Wir wählen eine natürliche Zahl n so, daß

$$\frac{1}{n} < \frac{\varepsilon}{(b-a) \cdot (f(b) - f(a))}, \quad (42)$$

und betrachten die Zerlegung $Z = (x_0, \dots, x_n)$ mit

$$x_i := a + i \cdot \frac{(b-a)}{n}.$$

Da f monoton wachsend ist, ist

$$\sup \{f(x) \mid x \in [x_{i-1}, x_i]\} = f(x_i)$$

und

$$\inf \{f(x) \mid x \in [x_{i-1}, x_i]\} = f(x_{i-1}).$$

Für die Ober- und Untersumme von f bezüglich Z folgt damit

$$\begin{aligned}
 \text{OS}(f, Z) - \text{US}(f, Z) &= \sum_{i=1}^n (x_i - x_{i-1}) \cdot (f(x_i) - f(x_{i-1})) \\
 &= \sum_{i=1}^n \frac{b-a}{n} \cdot (f(x_i) - f(x_{i-1})) \\
 &= \frac{b-a}{n} \cdot \sum_{i=1}^n (f(x_i) - f(x_{i-1})) \\
 &= \frac{b-a}{n} \cdot (f(x_n) - f(x_0)) \\
 &= \frac{b-a}{n} \cdot (f(b) - f(a)) \stackrel{(42)}{<} \varepsilon.
 \end{aligned}$$

Somit ist f integrierbar nach dem Riemannsches Integrierbarkeitskriterium 19.11. \square

Beispiel 19.15

Die Funktion in Beispiel 19.13 ist monoton wachsend und deshalb nach Proposition 19.14 auch integrierbar. 19.14 sagt aber nichts über den Wert des Integrals aus!

C) Riemannsches Folgenkriterium für Integrierbarkeit

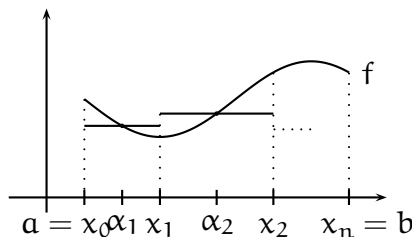
Definition 19.16 (Riemannsches Zwischenpunktkriterium)

Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt, $a < b$, und $Z = (x_0, \dots, x_n)$ eine Zerlegung von $[a, b]$.

Erfüllt $\alpha = (\alpha_1, \dots, \alpha_n)$ die Bedingung $\alpha_i \in [x_{i-1}, x_i]$ für $i = 1, \dots, n$, so nennen wir

$$\text{ZS}(f, Z, \alpha) := \sum_{i=1}^n (x_i - x_{i-1}) \cdot f(\alpha_i)$$

die *Riemannsches Zwischenpunktsomme* von f bezüglich der Zerlegung Z und den *Zwischenpunkten* α .



Das nächste Lemma sagt, daß man Obersummen und Untersummen beliebig gut approximieren kann durch Zwischenpunktsommen.

Lemma 19.17 (Approximation von Ober- und Unter- durch Zwischenpunktsommen)

Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt, $a < b$, Z eine Zerlegung von $[a, b]$ und $\varepsilon > 0$.

- Dann gibt es Zwischenpunkte α mit $0 \leq \text{OS}(f, Z) - \text{ZS}(f, Z, \alpha) < \varepsilon$.
- Dann gibt es Zwischenpunkte β mit $0 \leq \text{ZS}(f, Z, \beta) - \text{US}(f, Z) < \varepsilon$.

Beweis: Sei $Z = (x_0, \dots, x_n)$ und sei

$$M_i := \sup \{ f(x) \mid x \in [x_{i-1}, x_i] \}.$$

Aufgrund der Definition von M_i als Supremum der Funktionswerte auf dem Intervall $[x_{i-1}, x_i]$ gibt es ein $\alpha_i \in [x_{i-1}, x_i]$, so daß

$$f(\alpha_i) > M_i - \frac{\varepsilon}{b-a}.$$

Damit erhalten wir für $\alpha = (\alpha_1, \dots, \alpha_n)$

$$\begin{aligned} \text{OS}(f, Z) - \text{ZS}(f, Z, \alpha) &= \sum_{i=1}^n (x_i - x_{i-1}) \cdot (M_i - f(\alpha_i)) \\ &< \sum_{i=1}^n (x_i - x_{i-1}) \cdot \frac{\varepsilon}{b-a} \\ &= (x_n - x_0) \cdot \frac{\varepsilon}{b-a} = \varepsilon. \end{aligned}$$

Damit ist a. gezeigt, und b. zeigt man analog. □

Das folgende Lemma sagt, daß für integrierbare Funktionen Untersummen und Obersummen beliebig nahe beieinander und damit beim Wert des Integrals liegen, wenn nur die Länge der Zerlegung hinreichend klein gewählt ist.

Lemma 19.18 (Verschärfung des Riemannsches Integritätskriteriums)

Ist $f : [a, b] \rightarrow \mathbb{R}$ integrierbar, $a < b$, so gilt:

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : \forall Z \text{ Zerlegung mit } l(Z) < \delta_\varepsilon \text{ gilt } \text{OS}(f, Z) - \text{US}(f, Z) < \varepsilon.$$

Beweis: Sei $\varepsilon > 0$ gegeben. Aus dem Riemannsches Integritätskriterium erhalten wir eine Zerlegung Z' von $[a, b]$, so daß

$$\text{OS}(f, Z') - \text{US}(f, Z') < \frac{\varepsilon}{2}. \quad (43)$$

Wir setzen nun

$$\delta_\varepsilon := \frac{\varepsilon}{8 \cdot |Z'| \cdot M} > 0,$$

wobei $M := \sup \{ |f(x)| \mid x \in [a, b] \}$. Ist Z eine Zerlegung von $[a, b]$ mit $l(Z) < \delta_\varepsilon$, so folgt aus Lemma 19.5 und $|Z * Z'| - |Z| \leq |Z'|$

$$\text{OS}(f, Z) - \text{OS}(f, Z * Z') \leq 2 \cdot M \cdot l(Z) \cdot (|Z * Z'| - |Z|) < 2 \cdot M \cdot \delta_\varepsilon \cdot |Z'| = \frac{\varepsilon}{4} \quad (44)$$

und

$$\text{US}(f, Z * Z') - \text{US}(f, Z) \leq 2 \cdot M \cdot l(Z) \cdot (|Z * Z'| - |Z|) < 2 \cdot M \cdot \delta_\varepsilon \cdot |Z'| = \frac{\varepsilon}{4}. \quad (45)$$

Da $Z * Z'$ eine Verfeinerung von Z' ist, folgt aus (43) zusammen mit Lemma 19.5

$$\text{OS}(f, Z * Z') - \text{US}(f, Z * Z') \leq \text{OS}(f, Z') - \text{US}(f, Z') < \frac{\varepsilon}{2}. \quad (46)$$

Insgesamt erhalten wir damit

$$\begin{aligned} \text{OS}(f, Z) - \text{US}(f, Z) &= \text{OS}(f, Z) - \text{OS}(f, Z * Z') + \text{OS}(f, Z * Z') - \text{US}(f, Z * Z') \\ &\quad + \text{US}(f, Z * Z') - \text{US}(f, Z) \stackrel{(44)(46)(45)}{<} \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon. \end{aligned}$$

□

Satz 19.19 (Riemannsches Folgenkriterium für Integrierbarkeit)

Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion, $a < b$, und $I \in \mathbb{R}$.

Genau dann ist f auf $[a, b]$ integrierbar mit $I = \int_a^b f(x) dx$, wenn für jede Folge $(Z^n, \alpha^n)_{n \in \mathbb{N}}$ von Zerlegungen von $[a, b]$ und Zwischenpunkten mit $l(Z^n) \rightarrow 0$ gilt

$$\text{ZS}(f, Z^n, \alpha^n) \rightarrow I.$$

Beweis:

\Rightarrow : Es sei $(Z^n, \alpha^n)_{n \in \mathbb{N}}$ eine Folge von Zerlegungen von $[a, b]$ mit Zwischenpunkten, so daß $\lim_{n \rightarrow \infty} l(Z^n) = 0$, und sei $I = \int_a^b f(x) dx$.

Sei $\varepsilon > 0$ gegeben. Wir müssen ein $n_\varepsilon \in \mathbb{N}$ finden, so daß

$$|\text{ZS}(f, Z^n, \alpha^n) - I| < \varepsilon \tag{47}$$

für alle $n \geq n_\varepsilon$.

Da f integrierbar ist, gibt es nach Lemma 19.18 ein $\delta_\varepsilon > 0$, so daß für eine Zerlegung Z von $[a, b]$ aus $l(Z) < \delta_\varepsilon$ auch

$$\text{OS}(f, Z) - \text{US}(f, Z) < \varepsilon \tag{48}$$

gilt. Wegen $\lim_{n \rightarrow \infty} l(Z^n) = 0$ gibt es ein $n_\varepsilon \in \mathbb{N}$, so daß $l(Z^n) < \delta_\varepsilon$ für $n \geq n_\varepsilon$. Für $n \geq n_\varepsilon$ leiten wir dann aus (48)

$$\text{ZS}(f, Z^n, \alpha^n) - I \leq \text{OS}(f, Z^n) - I \leq \text{OS}(f, Z^n) - \text{US}(f, Z^n) < \varepsilon$$

her, sowie

$$\text{ZS}(f, Z^n, \alpha^n) - I \geq \text{US}(f, Z^n) - I \geq \text{US}(f, Z^n) - \text{OS}(f, Z^n) > -\varepsilon.$$

Damit ist (47) für $n \geq n_\varepsilon$ erfüllt, und das heißt $\text{ZS}(f, Z^n, \alpha^n) \rightarrow I$.

\Leftarrow : Wir wollen das Riemannsches Integrierbarkeitskriterium anwenden.

Sei dazu $\varepsilon > 0$ gegeben. Wir betrachten die Zerlegung $Z^n = (x_0^n, \dots, x_n^n)$, $n \geq 1$, mit

$$x_i^n := a + i \cdot \frac{(b-a)}{n}$$

für $i = 0, \dots, n$. Dann gilt

$$\lim_{n \rightarrow \infty} l(Z^n) = \lim_{n \rightarrow \infty} \frac{b-a}{n} = 0.$$

Mit Lemma 19.17 finden wir zu $n \in \mathbb{N}$ Zwischenpunkte α^n und β^n , so daß

$$\text{OS}(f, Z^n) - \text{ZS}(f, Z^n, \alpha^n) < \frac{\varepsilon}{4}.$$

und

$$ZS(f, Z^n, \beta^n) - US(f, Z^n) < \frac{\varepsilon}{4}.$$

Außerdem gelten nach Voraussetzung

$$ZS(f, Z^n, \alpha^n) \longrightarrow I$$

und

$$ZS(f, Z^n, \beta^n) \longrightarrow I.$$

Wir finden also ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $n \geq n_\varepsilon$ gilt

$$|ZS(f, Z^n, \alpha^n) - I| < \frac{\varepsilon}{4}$$

und

$$|ZS(f, Z^n, \beta^n) - I| < \frac{\varepsilon}{4}.$$

Dann gilt aber auch

$$\begin{aligned} OS(f, Z_{n_\varepsilon}) - US(f, Z_{n_\varepsilon}) &= |OS(f, Z_{n_\varepsilon}) - US(f, Z_{n_\varepsilon})| \\ &\leq |OS(f, Z_{n_\varepsilon}) - ZS(f, Z_{n_\varepsilon}, \alpha^{n_\varepsilon})| + |ZS(f, Z_{n_\varepsilon}, \alpha^{n_\varepsilon}) - I| \\ &\quad + |I - ZS(f, Z_{n_\varepsilon}, \beta^{n_\varepsilon})| + |ZS(f, Z_{n_\varepsilon}, \beta^{n_\varepsilon}) - US(f, Z_{n_\varepsilon})| \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon. \end{aligned}$$

Also ist f integrierbar nach dem Riemannsches Integrierbarkeitskriterium 19.11.

Außerdem haben wir für $n \geq n_\varepsilon$

$$\begin{aligned} |OS(f, Z^n) - I| &\leq |OS(f, Z^n) - ZS(f, Z^n, \alpha^n)| + |ZS(f, Z^n, \alpha^n) - I| \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < \varepsilon, \end{aligned}$$

woraus

$$OS(f, Z^n) \longrightarrow I$$

folgt. Analog sehen wir

$$US(f, Z^n) \longrightarrow I.$$

Damit erhalten wir dann

$$I \longleftarrow US(f, Z^n) \leq \int_a^b f(x) \, dx \leq OS(f, Z^n) \longrightarrow I,$$

so daß $I = \int_a^b f(x) \, dx$ aus dem Einschachtelungssatz 11.17 folgt.

□

Beispiel 19.20

Die Funktion $f : [0, b] \longrightarrow \mathbb{R} : x \mapsto x^2$, $b > 0$, ist stetig und mithin integrierbar.

Setzen wir

$$x_i := \frac{i \cdot b}{n},$$

so ist $Z^n = (x_0, \dots, x_n)$ eine Zerlegung von $[0, b]$ mit Zwischenpunkten $\alpha^n = (x_1, \dots, x_n)$, und es gilt $l(Z^n) \rightarrow 0$. Um die Zwischensumme berechnen zu können, verwenden wir die Formel

$$\sum_{i=1}^n i^2 = \frac{n \cdot (n+1) \cdot (2n+1)}{6}, \quad (49)$$

die man mit Hilfe von Induktion leicht zeigen kann. Damit ergibt sich

$$\begin{aligned} \text{ZS}(f, Z^n, \alpha^n) &= \sum_{i=1}^n (x_i - x_{i-1}) \cdot x_i^2 = \sum_{i=1}^n \frac{b}{n} \cdot \frac{i^2 \cdot b^2}{n^2} \\ &= \frac{b^3}{n^3} \cdot \sum_{i=1}^n i^2 \stackrel{(49)}{=} \frac{b^3}{n^3} \cdot \frac{n \cdot (n+1) \cdot (2n+1)}{6} \\ &= \frac{b^3}{3} \cdot \left(1 + \frac{1}{n}\right) \cdot \left(1 + \frac{1}{2n}\right) \rightarrow \frac{b^3}{3}. \end{aligned}$$

Aus dem Riemannschen Folgenkriterium für Integrierbarkeit 19.19 folgt dann

$$\int_0^b x^2 dx = \frac{b^3}{3}.$$

D) Rechenregeln für Integrale

Korollar 19.21 (Linearität und Monotonie des Integrals)

Seien $f, g : [a, b] \rightarrow \mathbb{R}$ integrierbar, $a < b$, und $c, d \in \mathbb{R}$.

a. Dann ist $c \cdot f + d \cdot g$ integrierbar auf $[a, b]$ mit

$$\int_a^b (c \cdot f + d \cdot g)(x) dx = c \cdot \int_a^b f(x) dx + d \cdot \int_a^b g(x) dx.$$

b. Ist $f(x) \leq g(x)$ für alle $x \in [a, b]$, so ist auch

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

Beweis:

a. Wir beachten zunächst, daß für jede Zerlegung $Z = (x_0, \dots, x_n)$ von $[a, b]$ mit Zwischenpunkten $\alpha = (\alpha_1, \dots, \alpha_n)$ offenbar gilt:

$$\begin{aligned} \text{ZS}(cf + dg, Z, \alpha) &= \sum_{i=1}^n (x_i - x_{i-1}) \cdot (cf + dg)(\alpha_i) \\ &= c \cdot \sum_{i=1}^n (x_i - x_{i-1}) \cdot f(\alpha_i) + d \cdot \sum_{i=1}^n (x_i - x_{i-1}) \cdot g(\alpha_i) \\ &= c \cdot \text{ZS}(f, Z, \alpha) + d \cdot \text{ZS}(g, Z, \alpha). \end{aligned}$$

Es sei nun $(Z^n, \alpha^n)_{n \in \mathbb{N}}$ eine Folge von Zerlegungen von $[a, b]$ und Zwischenpunkten mit $l(Z^n) \rightarrow 0$. Aus den Grenzwertsätzen für Folgen 11.15 und Satz 19.19 folgt dann

$$ZS(cf + dg, Z^n, \alpha^n) = c \cdot ZS(f, Z^n, \alpha^n) + d \cdot ZS(g, Z^n, \alpha^n) \rightarrow c \cdot \int_a^b f(x) dx + d \cdot \int_a^b g(x) dx.$$

Das Riemannsche Folgenkriterium für Integrierbarkeit 19.19 liefert dann die Behauptung.

- b. Es sei $(Z^n, \alpha^n)_{n \in \mathbb{N}}$ eine Folge von Zerlegungen von $[a, b]$ und Zwischenpunkten mit $l(Z^n) \rightarrow 0$. Wegen $f(x) \leq g(x)$ für alle $x \in [a, b]$ gilt dann offenbar

$$\int_a^b f(x) dx \leftarrow ZS(f, Z^n, \alpha^n) \leq ZS(g, Z^n, \alpha^n) \rightarrow \int_a^b g(x) dx,$$

wobei die Grenzwerte aus dem Riemannschen Folgenkriterium für Integrierbarkeit folgen. Damit gilt dann aber auch für die Grenzwerte

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

□

Beispiel 19.22

Aus Beispiel 19.10 und 19.20 erhalten wir aus der Linearität des Integrals

$$\int_0^b 3x^2 + 5 dx = 3 \cdot \int_0^b x^2 dx + \int_0^b 5 dx = b^3 + 5b.$$

Bemerkung 19.23 (Aneinanderhängen von Zerlegungen)

Ist $Z' = (x_0, \dots, x_n)$ eine Zerlegung von $[a, c]$ und $Z'' = (y_0, \dots, y_m)$ eine Zerlegung von $[c, b]$, so ist $Z' * Z'' = (x_0, \dots, x_n, y_1, \dots, y_m)$ eine Zerlegung von $[a, b]$ und sie entsteht durch aneinanderhängen der beiden Zerlegungen. Ist $\alpha = (\alpha_1, \dots, \alpha_n)$ ein Tupel von Zwischenpunkten von Z' und $\beta = (\beta_1, \dots, \beta_m)$ ein Tupel von Zwischenpunkten von Z'' , so definieren wir $\alpha \sqcup \beta = (\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_m)$ und erhalten damit ein Tupel von Zwischenpunkten von $Z' * Z''$.

Außerdem gelten offenbar

$$OS(f, Z' * Z'') = OS(f, Z') + OS(f, Z''),$$

$$US(f, Z' * Z'') = US(f, Z') + US(f, Z''),$$

$$ZS(f, Z' * Z'', \alpha \sqcup \beta) = ZS(f, Z', \alpha) + ZS(f, Z'', \beta).$$

Proposition 19.24 (Additivität des Integrals)

Es sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt, $a < b$ und $c \in (a, b)$.

Genau dann ist f integrierbar auf $[a, b]$, wenn f integrierbar auf $[a, c]$ und auf $[c, b]$ ist. Zudem gilt dann

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Beweis: Es sei $\varepsilon > 0$ gegeben.

Ist f integrierbar auf $[a, c]$ und auf $[c, b]$, so gibt es wegen des Riemannsches Integritätskriteriums 19.11 Zerlegungen $Z' = (x_0, \dots, x_n)$ von $[a, c]$ und $Z'' = (y_0, \dots, y_m)$ von $[c, b]$, so daß

$$\text{OS}(f, Z') - \text{US}(f, Z') < \frac{\varepsilon}{2}$$

und

$$\text{OS}(f, Z'') - \text{US}(f, Z'') < \frac{\varepsilon}{2}.$$

Dann ist aber $Z = Z' * Z'' = (x_0, \dots, x_n, y_1, \dots, y_m)$ eine Zerlegung von $[a, b]$ und

$$\begin{aligned} \text{OS}(f, Z) - \text{US}(f, Z) &= (\text{OS}(f, Z') + \text{OS}(f, Z'')) - (\text{US}(f, Z') + \text{US}(f, Z'')) \\ &= (\text{OS}(f, Z') - \text{US}(f, Z')) + (\text{OS}(f, Z'') - \text{US}(f, Z'')) \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Das Riemannsches Integritätskriterium 19.11 impliziert also, daß f auf $[a, b]$ integrierbar ist.

Ist umgekehrt f auf $[a, b]$ integrierbar, so gibt es wegen des Riemannsches Integritätskriteriums eine Zerlegung $Z = (x_1, \dots, x_n)$ von $[a, b]$ mit

$$\text{OS}(f, Z) - \text{US}(f, Z) < \varepsilon.$$

Nach eventueller Verfeinerung können wir ohne Einschränkung annehmen, daß $c = x_j \in \text{supp}(Z)$ ein Stützpunkt von Z ist. Dann ist $Z' := (x_0, \dots, x_j)$ eine Zerlegung von $[a, c]$ und $Z'' := (x_j, \dots, x_n)$ eine Zerlegung von $[c, b]$. Außerdem gilt $Z = Z' * Z''$ und

$$\begin{aligned} &(\text{OS}(f, Z') - \text{US}(f, Z')) + (\text{OS}(f, Z'') - \text{US}(f, Z'')) \\ &= (\text{OS}(f, Z') + \text{OS}(f, Z'')) - (\text{US}(f, Z') + \text{US}(f, Z'')) \\ &= \text{OS}(f, Z) - \text{US}(f, Z) < \varepsilon. \end{aligned}$$

Mithin gilt auch

$$\text{OS}(f, Z') - \text{US}(f, Z') < \varepsilon \quad \text{und} \quad \text{OS}(f, Z'') - \text{US}(f, Z'') < \varepsilon,$$

so daß aus dem Riemannsches Integritätskriterium 19.11 wieder folgt, daß f auf $[a, c]$ und auf $[c, b]$ integrierbar ist.

Wir wählen nun zwei Folgen $(Z^{n'}, \alpha^n)_{n \in \mathbb{N}}$ und $(Z^{n''}, \beta^n)_{n \in \mathbb{N}}$ von Zerlegungen von $[a, c]$ bzw. von $[c, b]$ mit Zwischenpunkten, so daß $\lim_{n \rightarrow \infty} l(Z^{n'}) = \lim_{n \rightarrow \infty} l(Z^{n''}) = 0$.

Wie oben können wir die Zerlegungen $Z^{n'}$ und $Z^{n''}$ zu einer Zerlegung $Z^n := Z^{n'} * Z^{n''}$ von $[a, b]$ zusammenfügen und ebenfalls die Zwischenpunkte α^n und β^n zu Zwischenpunkten $\gamma^n := \alpha^n \sqcup \beta^n$ von Z^n . Dann gilt $l(Z^n) = \max\{l(Z^{n'}), l(Z^{n''})\} \rightarrow 0$, und somit folgt aus dem Folgenkriterium für Integrierbarkeit 19.19

$$\int_a^b f(x) dx \leftarrow \text{ZS}(f, Z^n, \gamma^n) = \text{ZS}(f, Z^{n'}, \alpha^n) + \text{ZS}(f, Z^{n''}, \beta^n) \rightarrow \int_a^c f(x) dx + \int_c^b f(x) dx.$$

□

Beispiel 19.25

Aus Proposition 19.24 und Beispiel 19.20 erhalten wir für $0 < a < b$

$$\int_a^b x^2 dx = \int_0^b x^2 dx - \int_0^a x^2 dx = \frac{b^3}{3} - \frac{a^3}{3}.$$

Die Dreiecksungleichung für Summen liefert mit Induktion, daß

$$|a_1 + \dots + a_n| \leq |a_1| + \dots + |a_n|$$

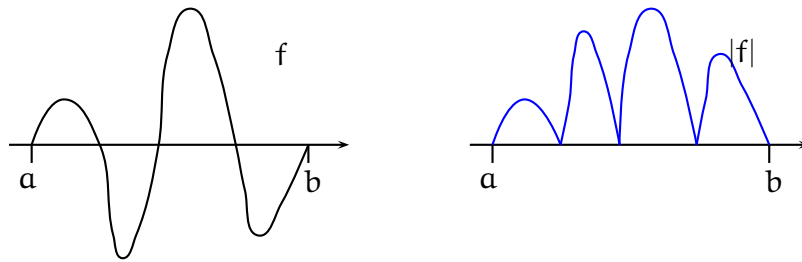
gilt. Integrale sind verallgemeinerte Summen, und die Dreiecksungleichung nimmt dann die folgende Gestalt an.

Proposition 19.26 (Dreiecksungleichung für Integrale)

Ist $f : [a, b] \rightarrow \mathbb{R}$ integrierbar auf $[a, b]$, $a < b$, so ist $|f|$ integrierbar auf $[a, b]$, und es gilt

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

Wir nennen das Integral über $|f|$ auch den Flächeninhalt, den der Graph von f auf dem Intervall $[a, b]$ mit der x -Achse einschließt.

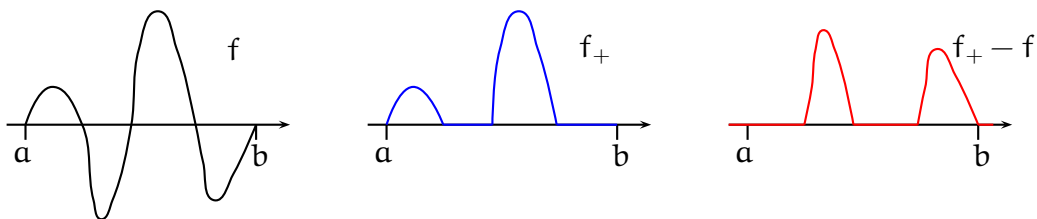


Beweis: Wir betrachten die Funktion

$$f_+ : [a, b] \rightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } f(x) \geq 0, \\ 0, & \text{falls } f(x) < 0. \end{cases}$$

Dann gilt

$$|f| = 2 \cdot f_+ - f.$$



Wir wollen nun mit Hilfe des Riemannsches Integrierbarkeitskriteriums zeigen, daß f_+ auf $[a, b]$ integrierbar ist. Sei $\varepsilon > 0$ gegeben. Da f auf $[a, b]$ integrierbar ist, gibt es eine Zerlegung $Z = (x_0, \dots, x_n)$ von $[a, b]$, so daß

$$OS(f, Z) - US(f, Z) < \varepsilon.$$

Wir behaupten, daß für jede Teilmenge $I \subseteq [a, b]$ die Ungleichung

$$\sup\{f_+(x) \mid x \in I\} - \inf\{f_+(x) \mid x \in I\} \leq \sup\{f(x) \mid x \in I\} - \inf\{f(x) \mid x \in I\}, \quad (50)$$

gilt. Dazu betrachten wir verschiedene Fälle.

- 1. Fall:** $f(x) < 0$ für alle $x \in I$: Dann gilt $f_+ \equiv 0$ auf I , so daß die linke Seite in (50) Null ist. Zugleich gilt

$$\inf\{f(x) \mid x \in I\} \leq \sup\{f(x) \mid x \in I\} \leq 0,$$

so daß die rechte Seite von (50) nicht-negativ ist. In diesem Fall gilt (50).

- 2. Fall:** $\exists y, z \in I$ mit $f(y) < 0 \leq f(z)$: Also $\sup\{f_+(x) \mid x \in I\} = \sup\{f(x) \mid x \in I\}$ und $\inf\{f_+(x) \mid x \in I\} = 0 > \inf\{f(x) \mid x \in I\}$. Damit gilt die Ungleichung (50).

- 3. Fall:** $f(x) \geq 0$ für alle $x \in I$: Dann ist $f = f_+$ auf I und (50) gilt.

Damit haben wir gezeigt, daß (50) stets erfüllt ist. Für die Differenz der Ober- und Untersumme von f_+ ergibt sich mit $I_i := [x_{i-1}, x_i]$ dann

$$\begin{aligned} \text{OS}(f_+, Z) - \text{US}(f_+, Z) &= \sum_{i=1}^n (x_i - x_{i-1}) \cdot (\sup\{f_+(x) \mid x \in I_i\} - \inf\{f_+(x) \mid x \in I_i\}) \\ &\leq \sum_{i=1}^n (x_i - x_{i-1}) \cdot (\sup\{f(x) \mid x \in I_i\} - \inf\{f(x) \mid x \in I_i\}) \\ &= \text{OS}(f, Z) - \text{US}(f, Z) < \varepsilon. \end{aligned}$$

Mit Hilfe des Riemannsches Integrabilitätskriteriums 19.11 folgt dann, daß f_+ auf $[a, b]$ integrierbar ist. Aus der Linearität des Integrals 19.21 folgt dann, daß auch

$$|f| = 2 \cdot f_+ - f$$

auf $[a, b]$ integrierbar ist.

Für eine Zerlegung $Z = (x_0, \dots, x_n)$ mit Zwischenpunkten $\alpha = (\alpha_1, \dots, \alpha_n)$ gilt

$$|\text{ZS}(f, Z, \alpha)| = \left| \sum_{i=1}^n (x_i - x_{i-1}) \cdot f(\alpha_i) \right| \leq \sum_{i=1}^n (x_i - x_{i-1}) \cdot |f(\alpha_i)| = \text{ZS}(|f|, Z, \alpha).$$

Sei nun $(Z^n, \alpha^n)_{n \in \mathbb{N}}$ eine Folge von Zerlegungen von $[a, b]$ und Zwischenpunkten mit $\lim_{n \rightarrow \infty} l(Z^n) = 0$, dann folgt

$$\left| \int_a^b f(x) \, dx \right| \longleftarrow |\text{ZS}(f, Z^n, \alpha^n)| \leq \text{ZS}(|f|, Z^n, \alpha^n) \longrightarrow \int_a^b |f(x)| \, dx.$$

Die Ungleichung bleibt für die Grenzwerte erhalten. □

Bemerkung 19.27

Es sei $f : [a, b] \rightarrow \mathbb{R}$ integrierbar auf $[a, b]$. Wenden wir Proposition 19.24 zweimal an, so sehen wir, daß f auf jedem Teilintervall $[c, d]$ von $[a, b]$ mit $c < d$ ebenfalls integrierbar ist. Wir definieren nun

$$\int_c^c f(x) \, dx := 0$$

und

$$\int_d^c f(x) dx := - \int_c^d f(x) dx.$$

Damit müssen die Integrationsgrenzen also nicht mehr verschieden sein, und die untere Integrationsgrenze muß auch nicht mehr die kleinere sein. Die Linearität und Additivität des Integrals verallgemeinern sich dann in naheliegender Weise.

Aufgaben

Aufgabe 19.28

Wir nennen eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ *stückweise stetig*, wenn es eine Zerlegung $Z = (x_0, \dots, x_n)$ von $[a, b]$ gibt, so daß die Funktionen $f_i : [x_{i-1}, x_i] \rightarrow \mathbb{R} : x \mapsto f(x)$ für $i = 1, \dots, n$ auf (x_{i-1}, x_i) stetig sind und so daß die Grenzwerte $\lim_{x \rightarrow x_{i-1}} f_i(x)$ und $\lim_{x \rightarrow x_i} f_i(x)$ in \mathbb{R} existieren.

Zeige, eine stückweise stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist integrierbar auf $[a, b]$.

Aufgabe 19.29

Bestimme die folgenden Integrale.

- $\int_0^{\frac{\pi}{3}} \frac{3}{\cos^2(x)} dx.$
- $\int_0^1 \frac{1}{1+x} dx.$
- $\int_{-1}^2 \left(8 \cdot (x-2)^3 + \frac{1}{\sqrt{x+2}} \right) dx.$
- $\int_0^{\frac{\pi}{4}} \sin(x) \cos(x) dx.$
- $\int_0^{\frac{1}{4}} x^2 e^{4x} dx.$

Aufgabe 19.30

Sei $f : [0, 1] \rightarrow \mathbb{R}, x \mapsto e^x$, $n \in \mathbb{N}$, $Z_n = (0, \frac{1}{2^n}, \frac{2}{2^n}, \dots, \frac{2^n-1}{2^n}, 1)$ eine Zerlegung und $\alpha^n = (\frac{1}{2^n}, \frac{2}{2^n}, \dots, \frac{2^n-1}{2^n}, 1)$ Zwischenpunkte. Zeige die folgenden Aussagen.

- $ZS(f, Z_n, \alpha^n) = (e-1) \cdot e^y \cdot \frac{1}{\frac{e^y-1}{y}}$ für $y = \frac{1}{2^n}$.
- $\lim_{y \rightarrow 0} \frac{e^y-1}{y} = 1.$
- Berechne $\int_0^1 e^x dx$ mit Hilfe der Zwischensumme aus Aufgabenteil a..

Aufgabe 19.31

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion mit $f(x) \geq 0$ für alle $x \in [a, b]$. Desweiteren existiere ein $c \in [a, b]$ mit $f(c) > 0$. Zeige, daß $\int_a^b f(x) dx > 0$ ist.

§ 20 Hauptsatz der Differential- und Integralrechnung mit Anwendungen

Definition 20.1 (Stammfunktion)

Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ eine Funktion.

Eine differenzierbare Funktion $F : I \rightarrow \mathbb{R}$ mit $F' = f$ heißt *Stammfunktion* von f .

Proposition 20.2 (Stammfunktionen sind eindeutig bis auf eine Konstante.)

Sei I ein Intervall, $f : I \rightarrow \mathbb{R}$ und $F, G : I \rightarrow \mathbb{R}$ zwei Stammfunktionen von f . Dann gibt es ein $c \in \mathbb{R}$, so daß $F(x) = G(x) + c$ für alle $x \in I$.

Beweis: Wähle einen Punkt $a \in I$ und setze $c := F(a) - G(a)$. Sei nun $a \neq b \in I$ gegeben, so müssen wir

$$F(b) = G(b) + c$$

zeigen. Wir können ohne Einschränkung annehmen, daß $a < b$ gilt. Nach Voraussetzung ist $F - G$ auf dem Intervall $[a, b] \subseteq I$ differenzierbar, also ist $F - G$ dort auch stetig. Wegen

$$(F - G)'(x) = F'(x) - G'(x) = f(x) - f(x) = 0$$

für alle $x \in [a, b]$ folgt aus Aufgabe 18.37, daß $F - G$ auf $[a, b]$ konstant ist. Es gilt also insbesondere, daß

$$F(b) - G(b) = F(a) - G(a) = c.$$

□

Beispiel 20.3

Die Funktion $F : [0, \infty) \rightarrow \mathbb{R} : x \mapsto \frac{x^3}{3}$ ist eine Stammfunktion von $f : [0, \infty) \rightarrow \mathbb{R} : x \mapsto x^2$, da $F' = f$. Man beachte, daß wir aus Beispiel 19.20 wissen, daß

$$F(y) = \frac{y^3}{3} = \int_0^y f(x) \, dx.$$

Diese Beobachtung werden wir im folgenden Satz verallgemeinern.

Satz 20.4 (Hauptsatz der Differential- und Integralrechnung)

Es sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \rightarrow \mathbb{R}$ stetig und $a \in I$.

Dann ist die Funktion $F : I \rightarrow \mathbb{R} : y \mapsto \int_a^y f(x) \, dx$ eine Stammfunktion von f .

Beweis: Sei $c \in I$ gegeben. Wir müssen zeigen, daß F in c differenzierbar ist mit $F'(c) = f(c)$. Sei dazu wiederum $\varepsilon > 0$ gegeben. Dann müssen wir ein $\delta_\varepsilon > 0$ finden, so daß

$$\left| \frac{F(y) - F(c)}{y - c} - f(c) \right| < \varepsilon$$

für alle $y \in I$ mit $0 < |y - c| < \delta_\varepsilon$ gilt.

Da f stetig in c ist, gibt es ein $\delta_\varepsilon > 0$, so daß

$$|f(x) - f(c)| < \frac{\varepsilon}{2} \tag{51}$$

für alle $x \in I$ mit $|x - c| < \delta_\varepsilon$. Sei nun $c \neq y \in I$ mit $|y - c| < \delta_\varepsilon$, dann gilt

$$\begin{aligned}
 \left| \frac{F(y) - F(c)}{y - c} - f(c) \right| &= \left| \frac{1}{y - c} \cdot \left(\int_a^y f(x) \, dx - \int_a^c f(x) \, dx \right) - f(c) \right| \\
 &\stackrel{19.24}{=} \left| \frac{1}{y - c} \cdot \int_c^y f(x) \, dx - f(c) \right| \\
 &= \left| \frac{1}{y - c} \cdot \int_c^y f(x) \, dx - \frac{f(c) \cdot (y - c)}{y - c} \right| \\
 &= \left| \frac{1}{y - c} \cdot \int_c^y f(x) \, dx - \frac{\int_c^y f(c) \, dx}{y - c} \right| \\
 &\stackrel{19.21}{=} \left| \frac{1}{y - c} \cdot \int_c^y (f(x) - f(c)) \, dx \right| \\
 &\stackrel{19.26}{\leq} \frac{1}{|y - c|} \cdot \left| \int_c^y |f(x) - f(c)| \, dx \right| \\
 &\stackrel{(51), 19.21}{\leq} \frac{1}{|y - c|} \cdot \left| \int_c^y \frac{\varepsilon}{2} \, dx \right| \\
 &= \frac{|y - c|}{|y - c|} \cdot \frac{\varepsilon}{2} < \varepsilon.
 \end{aligned}$$

□

Korollar 20.5 (Hauptsatz der Differential- und Integralrechnung)

Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und F sei eine Stammfunktion von f . Dann gilt

$$\int_a^b f(x) \, dx = F(b) - F(a).$$

Beweis: Wegen Satz 20.4 und 20.2 gibt eine Konstante $c \in \mathbb{R}$, so daß

$$F(y) = \int_a^y f(x) \, dx + c$$

für alle $y \in I$ gilt. Setzen wir $y = a$ ein, so erhalten wir

$$F(a) = \int_a^a f(x) \, dx + c = 0 + c = c,$$

und mithin gilt insbesondere

$$\int_a^b f(x) \, dx = F(b) - c = F(b) - F(a).$$

□

Bemerkung 20.6

Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ stetig.

- a. Der Hauptsatz der Differential- und Integralrechnung 20.4 besagt im wesentlichen, daß die Differentiation die Umkehrung der Integration ist.

b. Ist $F : I \rightarrow \mathbb{R}$ eine Stammfunktion von f und $a, b \in I$, so schreiben wir auch

$$F(x) \Big|_a^b := F(b) - F(a) = \int_a^b f(x) \, dx.$$

c. Wir nennen den Ausdruck

$$\int f(x) \, dx$$

ein *unbestimmtes Integral*. Man verwendet ihn gemeinhin, um eine beliebige Stammfunktion F zu bezeichnen, und schreibt dann $F(y) = \int^y f(x) \, dx$.

A) Stammfunktionen aus Ableitungen ablesen

Beispiel 20.7 (Einige ausgewählte Stammfunktionen)

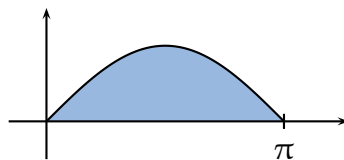
In den Abschnitten 17 und 18 haben wir für eine Vielzahl von stetig differenzierbarer Abbildungen die Ableitungen kennengelernt. Im Umkehrschluß haben wir damit für die Ableitungsfunktionen auch Stammfunktionen gefunden. Wir wollen für einige wichtige Beispiele von Funktionen f hier die Stammfunktionen F in tabellarischer Form zusammenstellen.

f	$F = \int f(x) \, dx$	f	$F = \int f(x) \, dx$
exp	exp	$\exp_a, a \neq 1$	$\frac{1}{\ln(a)} \cdot \exp_a$
cos	sin	sin	$-\cos$
$x \mapsto \frac{1}{x}$	ln	$x \mapsto x^a, -1 \neq a \in \mathbb{R}$	$x \mapsto \frac{1}{a+1} \cdot x^{a+1}$
$x \mapsto \frac{1}{x^2+1}$	arctan	$\frac{1}{\cos^2}$	tan

Beispiel 20.8 (Flächeninhalt eines Sinusbogens)

Wir können den Flächeninhalt unter einem der Bögen der Sinusfunktion berechnen als

$$\int_0^\pi \sin(x) \, dx = -\cos(x) \Big|_0^\pi = -\cos(\pi) + \cos(0) = 1 + 1 = 2.$$



B) Der Mittelwertsatz der Integralrechnung

Korollar 20.9 (Mittelwertsatz der Integralrechnung)

Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, $a < b$. Dann gibt es ein $c \in (a, b)$ mit

$$\int_a^b f(x) \, dx = f(c) \cdot (b - a).$$

Beweis: Die Funktion $F : [a, b] \rightarrow \mathbb{R} : y \mapsto \int_a^y f(x) dx$ ist nach dem Hauptsatz der Differential- und Integralrechnung 20.4 eine Stammfunktion von f und damit differenzierbar. Aus dem Mittelwertsatz der Differentialrechnung erhalten wir deshalb ein $c \in (a, b)$ mit

$$\int_a^b f(x) dx = F(b) - F(a) = F'(c) \cdot (b - a) = f(c) \cdot (b - a).$$

□

Wir geben jetzt noch einen alternativen Beweis des Mittelwertsatzes der Integralrechnung, der den Zwischenwertsatz verwendet sowie die Tatsache, daß stetige Funktionen auf abgeschlossenen Intervallen ihr Maximum und Minimum haben. Er hat den Vorteil, daß er sich direkt ins Mehrdimensionale verallgemeinern läßt (siehe Satz 44.27).

Alternativer Beweis des Mittelwertsatzes 20.9: Da f stetig auf $[a, b]$ ist, nimmt f auf $[a, b]$ sein Minimum und sein Maximum an (siehe Satz 14.16), d.h. es gibt $y, z \in [a, b]$ mit

$$f(y) \leq f(x) \leq f(z)$$

für alle $x \in [a, b]$. Für die Zerlegung $Z = (a, b)$ gilt dann

$$(b - a) \cdot f(y) = \text{US}(f, Z) \leq \int_a^b f(x) dx \leq \text{OS}(f, Z) = (b - a) \cdot f(z),$$

und damit

$$f(y) \leq \frac{1}{b - a} \cdot \int_a^b f(x) dx \leq f(z).$$

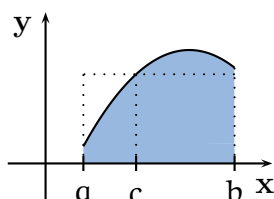
Nach dem Zwischenwertsatz 14.12 nimmt f jeden Wert zwischen $f(y)$ und $f(z)$ für ein geeignetes c zwischen y und z an. Insbesondere gibt es also ein $c \in (a, b)$ mit

$$f(c) = \frac{1}{b - a} \cdot \int_a^b f(x) dx.$$

Daraus folgt die Behauptung. □

Bemerkung 20.10 (Geometrische Interpretation des Mittelwertsatzes)

Der Mittelwertsatz besagt, daß das Rechteck mit den Seitenlängen $b - a$ und $f(c)$ den gleichen Flächeninhalt hat, wie die Fläche, die der Graph von f mit der x -Achse einschließt.



C) Partielle Integration

Satz 20.11 (Partielle Integration)

Sind $u, v : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar, dann gilt

$$\int_a^b u(x) \cdot v'(x) \, dx = u(x) \cdot v(x) \Big|_a^b - \int_a^b u'(x) \cdot v(x) \, dx.$$

Beweis: Aufgrund der Produktregel gilt $(u \cdot v)'(x) = u(x) \cdot v'(x) + u'(x) \cdot v(x)$ für $x \in [a, b]$, und mithin ist $u \cdot v$ eine Stammfunktion von $x \mapsto u(x) \cdot v'(x) + u'(x) \cdot v(x)$. Da letztere Funktion stetig ist auf $[a, b]$ folgt aus dem Hauptsatz der Differential- und Integralrechnung 20.4 und wegen der Linearität des Integrals 19.21

$$\int_a^b u(x) \cdot v'(x) \, dx + \int_a^b u'(x) \cdot v(x) \, dx = \int_a^b u(x) \cdot v'(x) + u'(x) \cdot v(x) \, dx = u(x) \cdot v(x) \Big|_a^b.$$

Damit ist die Aussage bewiesen. \square

Bemerkung 20.12 (Partielle Integration als Umkehrung der Produktregel)

Die partielle Integration ist die Umkehrung der Produktregel. Man wendet sie an, wenn man hofft, das Integral über $u' \cdot v$ leichter berechnen zu können als das über $u \cdot v'$. Auch mit partieller Integration kann man Stammfunktionen berechnen, indem man b durch die Variable y ersetzt und a ignoriert.

Beispiel 20.13 (Stammfunktion von \cos^2)

Wir wollen eine Stammfunktion von \cos^2 mit Hilfe partieller Integration berechnen. Dazu betrachten wir $u(x) = \cos(x)$ und $v'(x) = \cos(x)$. Dann ist $v(x) = \sin(x)$, und es gilt

$$\begin{aligned} \int^y \cos^2(x) \, dx &= \int^y u(x) \cdot v'(x) \, dx = u(x) \cdot v(x) \Big|_a^y - \int^y u'(x) \cdot v(x) \, dx \\ &= \cos(x) \cdot \sin(x) \Big|_a^y - \int^y -\sin^2(x) \, dx \\ &= \cos(x) \cdot \sin(x) \Big|_a^y - \int^y \cos^2(x) - 1 \, dx \\ &= \cos(x) \cdot \sin(x) \Big|_a^y - \int^y \cos^2(x) \, dx + \int^y 1 \, dx \\ &= \cos(x) \cdot \sin(x) \Big|_a^y - \int^y \cos^2(x) \, dx + x \Big|_a^y. \end{aligned}$$

Addieren wir auf beiden Seiten $\int^y \cos^2(x) \, dx$ und teilen durch 2, so erhalten wir

$$\int^y \cos^2(x) \, dx = \frac{1}{2} \cdot (y + \cos(y) \cdot \sin(y)).$$

D) Der Satz von Taylor

Aus dem Hauptsatz der Differential- und Integralrechnung und unter Anwendung der Methode der partiellen Integration ergibt sich eine Integralform für das Restglied im Satz von Taylor.

Korollar 20.14 (Satz von Taylor – Restglied in Integralform)

Sei I ein Intervall, $f : I \rightarrow \mathbb{R}$ eine $n+1$ -fach differenzierbare Funktion und $x, a \in I$. Dann gilt

$$f(x) - T_{f,a}^n(x) = \int_a^x \frac{f^{(n+1)}(y)}{n!} \cdot (x-y)^n dy.$$

Beweis: Wir führen den Beweis durch Induktion nach n . Für $n = 0$ gilt

$$f(x) - T_{f,a}^0(x) = f(x) - f(a) \stackrel{20.5}{=} \int_a^x f'(y) dy = \int_a^x \frac{f'(y)}{0!} \cdot (x-y)^0 dy$$

nach dem Hauptsatz der Differential- und Integralrechnung 20.4.

Nun setzen wir voraus, daß $n \geq 1$ ist und daß die Aussage für $n-1$ bereits gezeigt ist und wir wollen sie für n zeigen. Aufgrund der Induktionsvoraussetzung gilt dann

$$f(x) - T_{f,a}^{n-1}(x) = \int_a^x \frac{f^{(n)}(y)}{(n-1)!} \cdot (x-y)^{n-1} dy.$$

Wir setzen nun $u(y) := \frac{f^{(n)}(y)}{(n-1)!}$ und $v'(y) := (x-y)^{n-1}$ und wenden partielle Integration an. Die Stammfunktion von v' ist durch $v(y) = \frac{-(x-y)^n}{n}$ gegeben, so daß wir

$$\begin{aligned} f(x) - T_{f,a}^{n-1}(x) &= \int_a^x \frac{f^{(n)}(y)}{(n-1)!} \cdot (x-y)^{n-1} dy = \int_a^x u(y) \cdot v'(y) dy \\ &= u(y) \cdot v(y) \Big|_a^x - \int_a^x u'(y) \cdot v(y) dy \\ &= \frac{f^{(n)}(y)}{(n-1)!} \cdot \frac{-(x-y)^n}{n} \Big|_a^x - \int_a^x \frac{f^{(n+1)}(y)}{(n-1)!} \cdot \frac{-(x-y)^n}{n} dy \\ &= \frac{f^{(n)}(a)}{n!} \cdot (x-a)^n + \int_a^x \frac{f^{(n+1)}(y)}{n!} \cdot (x-y)^n dy. \end{aligned}$$

Bringen wir den Summanden $\frac{f^{(n)}(a)}{n!} \cdot (x-a)^n$ auf die linke Seite, so erhalten wir

$$f(x) - T_{f,a}^n(x) = f(x) - T_{f,a}^{n-1}(x) - \frac{f^{(n)}(a)}{n!} \cdot (x-a)^n = \int_a^x \frac{f^{(n+1)}(y)}{n!} \cdot (x-y)^n dy,$$

und damit folgt die Behauptung mittels des Prinzips der Induktion. \square

E) Die Substitutionsregel**Satz 20.15** (Substitutionsregel)

Es sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \rightarrow \mathbb{R}$ stetig und $\varphi : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar mit $\text{Im}(\varphi) \subseteq I$. Dann gilt

$$\int_{\varphi(a)}^{\varphi(b)} f(x) dx = \int_a^b f(\varphi(x)) \cdot \varphi'(x) dx.$$

Beweis: Da φ stetig ist, nimmt φ sein Minimum und sein Maximum an, d.h. es gibt $\mathbf{y}, \mathbf{z} \in [\mathbf{a}, \mathbf{b}]$ mit $\varphi(\mathbf{y}) \leq \varphi(\mathbf{x}) \leq \varphi(\mathbf{z})$ für alle $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$, und mithin ist

$$\text{Im}(\varphi) = [\varphi(\mathbf{y}), \varphi(\mathbf{z})] \subseteq I$$

ein Intervall. Als stetige Funktion besitzt f nach dem Hauptsatz der Differential- und Integralrechnung 20.4 auf diesem Intervall eine Stammfunktion F . Aus der Kettenregel 17.16 folgt dann

$$(F \circ \varphi)'(\mathbf{x}) = F'(\varphi(\mathbf{x})) \cdot \varphi'(\mathbf{x}) = f(\varphi(\mathbf{x})) \cdot \varphi'(\mathbf{x}),$$

so daß $F \circ \varphi$ eine Stammfunktion von $\mathbf{x} \mapsto f(\varphi(\mathbf{x})) \cdot \varphi'(\mathbf{x})$ auf $[\mathbf{a}, \mathbf{b}]$ ist. Dann können wir Korollar 20.5 anwenden und erhalten

$$\int_{\mathbf{a}}^{\mathbf{b}} f(\varphi(\mathbf{x})) \cdot \varphi'(\mathbf{x}) \, d\mathbf{x} \stackrel{20.5}{=} (F \circ \varphi)(\mathbf{b}) - (F \circ \varphi)(\mathbf{a}) = F(\varphi(\mathbf{b})) - F(\varphi(\mathbf{a})) \stackrel{20.5}{=} \int_{\varphi(\mathbf{a})}^{\varphi(\mathbf{b})} f(\mathbf{x}) \, d\mathbf{x}.$$

□

Bemerkung 20.16 (Die Substitutionsregel als Umkehrung der Kettenregel)

- a. Die Substitutionsregel ist die Umkehrung der Kettenregel.
- b. Es ist üblich, bei der Formel für die Substitutionsregel auf der linken Seite statt der Variablen \mathbf{x} die Variable \mathbf{z} zu verwenden, so daß die Formel folgende Gestalt hat:

$$\int_{\varphi(\mathbf{a})}^{\varphi(\mathbf{b})} f(\mathbf{z}) \, d\mathbf{z} = \int_{\mathbf{a}}^{\mathbf{b}} f(\varphi(\mathbf{x})) \cdot \varphi'(\mathbf{x}) \, d\mathbf{x}.$$

Man sagt dann, daß man $\varphi(\mathbf{x})$ durch \mathbf{z} substituiert oder umgekehrt, je nachdem ob man die linke durch die rechte Seite ausrechnen will oder umgekehrt. Man schreibt $\mathbf{z} = \varphi(\mathbf{x})$.

Diese Schreibweise kann man nutzen, um sich für die Substitution eine Eselsbrücke zu bauen. In Anlehnung an die Schreibweise $\varphi' = \frac{\partial \varphi}{\partial \mathbf{x}}$ kann man mit $\mathbf{z} = \varphi(\mathbf{x})$ dann auch

$$\varphi'(\mathbf{x}) \, d\mathbf{x} = \frac{d\mathbf{z}}{d\mathbf{x}} \, d\mathbf{x} = d\mathbf{z}$$

schreiben. Damit wird aus der Substitutionsformel ohne Integralgrenzen dann

$$\int f(\varphi(\mathbf{x})) \cdot \varphi'(\mathbf{x}) \, d\mathbf{x} = \int f(\mathbf{z}) \cdot \frac{d\mathbf{z}}{d\mathbf{x}} \, d\mathbf{x} = \int f(\mathbf{z}) \, d\mathbf{z}.$$

- c. Man kann mit Hilfe der Substitutionsregel auch Stammfunktionen ausrechnen, indem man die Integrationsgrenze \mathbf{b} durch die Variable \mathbf{y} ersetzt und \mathbf{a} ignoriert.

Beispiel 20.17 (Stammfunktion von $\mathbf{x} \mapsto \mathbf{x} \cdot \exp(\mathbf{x}^2)$)

Wir wollen das Integral $\int_{\mathbf{a}}^{\mathbf{b}} \mathbf{x} \cdot \exp(\mathbf{x}^2) \, d\mathbf{x}$ für $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ berechnen. Dazu substituieren wir $\mathbf{z} = \mathbf{x}^2$, d.h. wir betrachten $\varphi : \mathbb{R} \rightarrow \mathbb{R} : \mathbf{x} \mapsto \mathbf{x}^2$, $\varphi'(\mathbf{x}) = 2\mathbf{x}$ und $f : \mathbb{R} \rightarrow \mathbb{R} :$

$z \mapsto \frac{\exp(z)}{2}$. Da zudem f eine Stammfunktion von f ist, folgt damit

$$\begin{aligned} \int_a^b x \cdot \exp(x^2) dx &= \int_a^b f(\varphi(x)) \cdot \varphi'(x) dx = \int_{\varphi(a)}^{\varphi(b)} f(z) dz \\ &= \int_{a^2}^{b^2} \frac{\exp(z)}{2} dz = \frac{\exp(b^2)}{2} - \frac{\exp(a^2)}{2}. \end{aligned}$$

Beispiel 20.18 (Stammfunktion von \tan)

Wir wollen eine Stammfunktion für den Tangens auf dem Intervall $(-\frac{\pi}{2}, \frac{\pi}{2})$ bestimmen. Dazu substituieren wir $z = \cos(x)$, d.h. $\varphi(x) = \cos(x)$, $\varphi'(x) = -\sin(x)$ und $f(z) = -\frac{1}{z}$. Dann erhalten wir

$$\begin{aligned} \int^y \tan(x) dx &= \int^y -\frac{1}{\cos(x)} \cdot (-\sin(x)) dx = \int^y f(\varphi(x)) \cdot \varphi'(x) dx \\ &= \int^{\cos(y)} f(z) dz = \int^{\cos(y)} -\frac{1}{z} dz \\ &= -\ln(z) \Big|_{\cos(y)} = -\ln(\cos(y)). \end{aligned}$$

Also ist $\ln \circ \cos$ eine Stammfunktion von \tan auf $(-\frac{\pi}{2}, \frac{\pi}{2})$.

Beispiel 20.19 (Stammfunktion von $x \mapsto \sqrt{1-x^2}$)

Wir wollen mit Hilfe von Substitution eine Stammfunktion für die stetige Funktion

$$f: [-1, 1] \longrightarrow \mathbb{R} : z \mapsto \sqrt{1-z^2}$$

bestimmen. Dazu substituieren wir $z = \sin(x)$, d.h. $\varphi(x) = \sin(x)$, $\varphi'(x) = \cos(x)$ und $b = \arcsin(y)$. Dann definiert

$$\begin{aligned} F(y) &= \int^y \sqrt{1-z^2} dz = \int^{\varphi(b)} f(z) dz = \int^b f(\varphi(x)) \cdot \varphi'(x) dx \\ &= \int^{\arcsin(y)} \sqrt{1-\sin^2(x)} \cdot \cos(x) dx \\ &= \int^{\arcsin(y)} \sqrt{\cos^2(x)} \cdot \cos(x) dx \\ &= \int^{\arcsin(y)} \cos^2(x) dx \\ &\stackrel{20.13}{=} \frac{1}{2} \cdot (x + \cos(x) \cdot \sin(x)) \Big|_{\arcsin(y)} \\ &= \frac{1}{2} \cdot (x + \sqrt{\cos^2(x)} \cdot \sin(x)) \Big|_{\arcsin(y)} \\ &= \frac{1}{2} \cdot (x + \sqrt{1-\sin^2(x)} \cdot \sin(x)) \Big|_{\arcsin(y)} \\ &= \frac{\arcsin(y) + y \cdot \sqrt{1-y^2}}{2}. \end{aligned}$$

eine Stammfunktion von f auf $[-1, 1]$.

Beispiel 20.20 (Flächeninhalt eines Kreises)

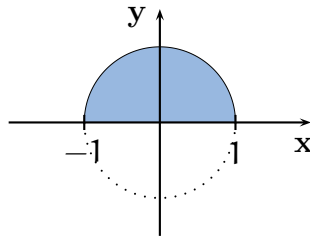
Wir wollen nun den Flächeninhalt eines Kreises berechnen. Die obere Hälfte des Einheitskreises mit dem Ursprung als Mittelpunkt

$$K_1(0) := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$$

ist die Fläche, die der Graph der Funktion

$$f : [-1, 1] \longrightarrow \mathbb{R} : x \mapsto \sqrt{1 - x^2}$$

mit der x -Achse einschließt.



Mithin ist das Integral

$$\begin{aligned} \int_{-1}^1 \sqrt{1 - x^2} \, dx &\stackrel{20.19}{=} \left. \frac{\arcsin(x) + x \cdot \sqrt{1 - x^2}}{2} \right|_{-1}^1 \\ &= \frac{\arcsin(1)}{2} - \frac{\arcsin(-1)}{2} = \frac{\pi}{4} - \left(-\frac{\pi}{4}\right) = \frac{\pi}{2} \end{aligned}$$

die Hälfte des Flächeninhaltes des Einheitskreises.

Der Kreis mit Radius r und dem Ursprung als Mittelpunkt ist

$$K_r(0) := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2\},$$

und sein Flächeninhalt ist entsprechend das Integral

$$2 \cdot \int_{-r}^r \sqrt{r^2 - x^2} \, dx.$$

Um dieses zu berechnen, substituieren wir $z = \frac{x}{r}$, d.h. $\varphi(x) = \frac{x}{r}$, $\varphi'(x) = \frac{1}{r}$, $\varphi(r) = 1$, $\varphi(-r) = -1$ und $f(z) = \sqrt{1 - z^2}$. Wir erhalten dann

$$\begin{aligned} 2 \cdot \int_{-r}^r \sqrt{r^2 - x^2} \, dx &= 2 \cdot r^2 \cdot \int_{-r}^r \sqrt{1 - \frac{x^2}{r^2}} \cdot \frac{1}{r} \, dx \\ &= 2 \cdot r^2 \cdot \int_{-1}^1 \sqrt{1 - z^2} \, dz = 2 \cdot \frac{\pi}{2} \cdot r^2 = \pi \cdot r^2 \end{aligned}$$

als Flächeninhalt von $K_r(0)$.

F) Partialbruchzerlegung

Bemerkung 20.21 (Partialbruchzerlegung)

Jede rationale Funktion $r = \frac{f}{g}$ läßt sich schreiben als

$$r = \frac{f}{g} = h + \frac{p}{q}, \quad (52)$$

wobei h , p und q Polynome sind mit $\deg(p) < \deg(q)$. Dies folgt sofort mittels einer einfachen Polynomdivision, wie sie in der Vorlesung Algebraische Strukturen eingeführt wird.

Nicht offensichtlich ist, daß sich der Bruch $\frac{p}{q}$ als Summe von Ausdrücken der Form

$$\frac{A}{(t-a)^k} \quad \text{und} \quad \frac{Bt+C}{(t^2+bt+c)^k} \quad (53)$$

für geeignete $A, B, C, a, b, c \in \mathbb{R}$ mit $4b - a^2 > 0$ schreiben läßt. Genauer kann man zeigen, daß wenn $(t-a)^n$ bzw. $(t^2+bt+c)^n$ die höchste Potenz von $t-a$ bzw. von t^2+bt+c ist, die das Polynom q teilt, so kommen in der Summe Ausdrücke der Form (53) für $k = 1, \dots, n$ vor. Eine solche Darstellung nennt man dann die *Partialbruchzerlegung* von $\frac{p}{q}$. Wir werden unten in einem Beispiel sehen, wie man diese unter Umständen finden kann.

Eine rationale Funktion wie in (52) ist stetig auf ihrem Definitionsbereich, der eine Vereinigung von endlich vielen Intervallen ist. Mithin ist sie auf allen abgeschlossenen Teilintervallen ihres Definitionsbereiches integrierbar. Um nun eine Stammfunktion von r zu bestimmen, reicht es im wesentlichen, Funktionen der Form (53) zu integrieren. Dies ist mit Hilfe unserer bisherigen Methoden vergleichsweise einfach, sei in der allgemeinen Form aber dem Leser als Übungsaufgabe überlassen.

Beispiel 20.22 (Integration mit Partialbruchzerlegung)

a. Wir wollen das folgende Integral berechnen:

$$\int_1^2 \frac{3x^5 + 3x^4 + 6x^2 + x - 2}{x^3 + x^2} dx.$$

Polynomdivision von $3t^5 + 3t^4 + 6t^2 + t - 2$ durch $t^3 + t^2$ liefert

$$r = \frac{3t^5 + 3t^4 + 6t^2 + t - 2}{t^3 + t^2} = 3t^2 + \frac{6t^2 + t - 2}{t^3 + t^2} = f + \frac{p}{q}.$$

Dabei faktorisiert q als

$$q = t^2 \cdot (t + 1).$$

Das Prinzip der Partialbruchzerlegung läßt uns nun nach Zahlen $A, B, C \in \mathbb{R}$ suchen, so daß

$$\frac{6t^2 + t - 2}{t^3 + t^2} = \frac{A}{t} + \frac{B}{t^2} + \frac{C}{t+1}$$

gilt. Bringen wir die rechte Seite auf den Hauptnenner, so erhalten wir

$$\begin{aligned} \frac{6t^2 + t - 2}{t^3 + t^2} &= \frac{A \cdot t \cdot (t + 1)}{t^3 + t^2} + \frac{B \cdot (t + 1)}{t^3 + t^2} + \frac{C \cdot t^2}{t^3 + t^2} \\ &= \frac{(A + C) \cdot t^2 + (A + B) \cdot t + B}{t^3 + t^2}. \end{aligned}$$

Ein Koeffizientenvergleich der Polynome im Zähler der beiden Seiten führt zu den Gleichungen:

$$A + C = 6, \quad A + B = 1 \quad \text{und} \quad B = -2.$$

Daraus lesen wir ohne Schwierigkeiten

$$A = 3, \quad B = -2 \quad \text{und} \quad C = 3$$

ab. Für unser Integral ergibt sich daraus

$$\begin{aligned} \int_1^2 \frac{3x^5 + 3x^4 + 6x^2 + x - 2}{x^3 + x^2} dx &= \int_1^2 3x^2 dx + \int_1^2 \frac{3}{x} dx + \int_1^2 \frac{-2}{x^2} dx + \int_1^2 \frac{3}{x+1} dx \\ &= x^3 + 3 \cdot \ln(x) + \frac{2}{x} + 3 \cdot \ln(x+1) \Big|_1^2 \\ &= 8 + 3 \cdot \ln(2) + 1 + 3 \cdot \ln(3) - 1 - 3 \cdot \ln(1) - 2 - 3 \cdot \ln(2) \\ &= 3 \cdot \ln(3) + 6. \end{aligned}$$

b. Wie sieht eine Stammfunktion zu folgender Funktion aus

$$x \mapsto \frac{2x^3 + x^2 + 1}{x^2 \cdot (1 + x^2)^2}?$$

Das Prinzip der Partialbruchzerlegung läßt uns nach reellen Zahlen $A, B, C, D, E, F \in \mathbb{R}$ suchen, so daß

$$\frac{2t^3 + t^2 + 1}{t^2 \cdot (1 + t^2)^2} = \frac{A}{t} + \frac{B}{t^2} + \frac{Ct + D}{1 + t^2} + \frac{Et + F}{(1 + t^2)^2}$$

gilt. Bringen wir die rechte Seite auf den Hauptnenner, so erhalten wir

$$\begin{aligned} \frac{2t^3 + t^2 + 1}{t^2 \cdot (1 + t^2)^2} &= \\ &= \frac{(A + C) \cdot t^5 + (B + D) \cdot t^4 + (2A + C + E) \cdot t^3 + (2B + D + F) \cdot t^2 + A \cdot t + B}{t^2 \cdot (1 + t^2)^2}. \end{aligned}$$

Durch Koeffizientenvergleich erhalten wir dann

$$A = 0, \quad B = 1, \quad C = 0, \quad D = -1, \quad E = 2 \quad \text{und} \quad F = 0,$$

d.h.

$$\frac{2t^3 + t^2 + 1}{t^2 \cdot (1 + t^2)^2} = \frac{1}{t^2} - \frac{1}{1 + t^2} + \frac{2t}{(1 + t^2)^2}.$$

Damit erhalten wir als Stammfunktion

$$\begin{aligned} F(y) &= \int^y \frac{2x^3 + x^2 + 1}{x^2 \cdot (1 + x^2)^2} dx = \int^y \frac{1}{x^2} dx - \int^y \frac{1}{1 + x^2} dx + \int^y \frac{2x}{(1 + x^2)^2} dx \\ &= -\frac{1}{y} - \arctan y - \frac{1}{1 + y^2}, \end{aligned}$$

wobei wir zur Berechnung des letzten Integrals die Substitution $z = 1 + x^2$ vornehmen und so

$$\int^y \frac{2x}{(1+x^2)^2} dx = \int^{1+y^2} \frac{1}{z^2} dz = -\frac{1}{z} \Big|^{1+y^2} = -\frac{1}{1+y^2}$$

erhalten.

G) Vertauschbarkeit von Grenzwert und Integration

Satz 20.23 (Vertauschbarkeit von Grenzwert und Integration)

Die Folge $(f_n)_{n \in \mathbb{N}}$ stetiger Funktionen $f_n : [a, b] \rightarrow \mathbb{R}$ konvergiere auf $[a, b]$, $a < b$, gleichmäßig gegen $f : [a, b] \rightarrow \mathbb{R}$.

Dann ist auch die Grenzfunktion f stetig auf $[a, b]$ und für alle $y \in [a, b]$ gilt

$$\int_a^y f(x) dx = \lim_{n \rightarrow \infty} \int_a^y f_n(x) dx. \quad (54)$$

D.h. der Grenzwert der Stammfunktionen der f_n ist eine Stammfunktion von f .

Beweis: Nach Satz 15.6 ist f als gleichmäßiger Grenzwert stetiger Funktionen stetig. Da stetige Funktionen nach Satz 19.12 integrierbar sind, existieren die Integrale in (54). Es bleibt, für $y \in [a, b]$ zu zeigen, daß

$$\lim_{n \rightarrow \infty} \int_a^y f_n(x) dx = \int_a^y f(x) dx.$$

Sei dazu $\varepsilon > 0$ gegeben. Da $(f_n)_{n \in \mathbb{N}}$ auf $[a, b]$ gleichmäßig gegen f konvergiert, gibt es ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $n \geq n_\varepsilon$ und für alle $x \in [a, b]$

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2 \cdot (b - a)}. \quad (55)$$

Dann gilt für $n \geq n_\varepsilon$ auch

$$\left| \int_a^y f_n(x) dx - \int_a^y f(x) dx \right| \stackrel{19.26}{\leq} \int_a^y |f_n(x) - f(x)| dx \stackrel{19.21, (55)}{\leq} \int_a^y \frac{\varepsilon}{2 \cdot (b - a)} dx \leq \frac{\varepsilon}{2} < \varepsilon.$$

Damit ist die Behauptung gezeigt. \square

Bemerkung 20.24

- a. Wie in Bemerkung 18.17 wollen wir wieder darauf hinweisen, daß wir in Satz 20.23 gezeigt haben, daß zwei Grenzwertprozesse vertauschen. Es gilt nämlich

$$\lim_{n \rightarrow \infty} \lim_{m \rightarrow \infty} \text{ZS}(f_n, Z_m, \alpha^m) = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx = \lim_{m \rightarrow \infty} \lim_{n \rightarrow \infty} \text{ZS}(f_n, Z_m, \alpha^m).$$

- b. Ersetzen wir in Satz 20.23 die Voraussetzung *stetig* durch *integrierbar*, so wird auch die Grenzfunktion f nur noch *integrierbar* sein, es gilt aber nach wie vor

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

Wir geben unten einen Beweis für diese Aussage.

- c. Wir können mit Satz 20.23 und dem Hauptsatz der Differential- und Integralrechnung 20.4 einen wesentlich kürzeren Beweis der Vertauschbarkeit von Grenzwert und Ableitung 18.16 geben:

Die Funktionen f_n in Satz 18.16 sind jeweils Stammfunktion von f'_n , und da die f'_n stetig sind, folgt aus dem Hauptsatz der Differential- und Integralrechnung 20.5

$$f_n(\mathbf{y}) - f_n(\mathbf{a}) = \int_{\mathbf{a}}^{\mathbf{y}} f'_n(x) \, dx.$$

Bilden wir auf beiden Seiten den Grenzwert, so folgt mit Satz 20.23

$$f(\mathbf{y}) - f(\mathbf{a}) \longleftarrow f_n(\mathbf{y}) - f_n(\mathbf{a}) = \int_{\mathbf{a}}^{\mathbf{y}} f'_n(x) \, dx \longrightarrow \int_{\mathbf{a}}^{\mathbf{y}} g(x) \, dx.$$

Also ist

$$f(\mathbf{y}) = f(\mathbf{a}) + \int_{\mathbf{a}}^{\mathbf{y}} g(x) \, dx$$

für $\mathbf{y} \in [\mathbf{a}, \mathbf{b}]$, so daß f nach dem Hauptsatz der Differential- und Integralrechnung 20.4 differenzierbar ist auf $[\mathbf{a}, \mathbf{b}]$ mit

$$f'(\mathbf{y}) = 0 + g(\mathbf{y}) = g(\mathbf{y})$$

für alle $\mathbf{y} \in [\mathbf{a}, \mathbf{b}]$.

Beweis von Teil b. Wir zeigen zunächst, daß die Folge $(\int_{\mathbf{a}}^{\mathbf{b}} f_n(x) \, dx)_{n \in \mathbb{N}}$ konvergiert.

Sei dazu $\varepsilon > 0$ gegeben. Da $(f_n)_{n \in \mathbb{N}}$ auf $[\mathbf{a}, \mathbf{b}]$ gleichmäßig gegen f konvergiert, gibt es ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $n \geq n_\varepsilon$ und für alle $x \in [\mathbf{a}, \mathbf{b}]$

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3 \cdot (\mathbf{b} - \mathbf{a})}. \quad (56)$$

Dann gilt für $m > n \geq n_\varepsilon$

$$|f_m(x) - f_n(x)| \leq |f_m(x) - f(x)| + |f(x) - f_n(x)| < \frac{2\varepsilon}{3 \cdot (\mathbf{b} - \mathbf{a})} \quad (57)$$

für alle $x \in [\mathbf{a}, \mathbf{b}]$ und mithin

$$\left| \int_{\mathbf{a}}^{\mathbf{b}} f_m(x) \, dx - \int_{\mathbf{a}}^{\mathbf{b}} f_n(x) \, dx \right| \stackrel{19.26}{\leq} \int_{\mathbf{a}}^{\mathbf{b}} |f_m(x) - f_n(x)| \, dx \stackrel{19.21, (57)}{\leq} \int_{\mathbf{a}}^{\mathbf{b}} \frac{2\varepsilon}{3 \cdot (\mathbf{b} - \mathbf{a})} \, dx = \frac{2\varepsilon}{3} < \varepsilon.$$

Also ist die Folge $(\int_{\mathbf{a}}^{\mathbf{b}} f_n(x) \, dx)_{n \in \mathbb{N}}$ eine Cauchy-Folge und damit konvergent. Wir setzen

$$I := \lim_{n \rightarrow \infty} \int_{\mathbf{a}}^{\mathbf{b}} f_n(x) \, dx.$$

Nun wollen wir das Folgenkriterium für Integrierbarkeit 19.19 anwenden, um zu zeigen, daß f integrierbar ist und daß $I = \int_{\mathbf{a}}^{\mathbf{b}} f(x) \, dx$. Sei dazu $(Z_m, \alpha^m)_{m \in \mathbb{N}}$ eine Folge von Zerlegungen von $[\mathbf{a}, \mathbf{b}]$ und Zwischenpunkten mit $\iota(Z_m) \rightarrow 0$, und sei $\varepsilon > 0$ gegeben. Wir wählen nun $n_\varepsilon \in \mathbb{N}$ wie oben.

Außerdem, da $\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = I$, gibt es ein $N_\varepsilon \in \mathbb{N}$, so daß

$$\left| \int_a^b f_n(x) dx - I \right| < \frac{\varepsilon}{3} \quad (58)$$

für alle $n \geq N_\varepsilon$.

Wir setzen $n := \max\{n_\varepsilon, N_\varepsilon\}$. Für eine beliebige Zerlegung $Z = (x_0, \dots, x_k)$ von $[a, b]$ mit Zwischenpunkten $\alpha = (\alpha_1, \dots, \alpha_k)$ gilt dann

$$\begin{aligned} |ZS(Z, f_n, \alpha) - ZS(Z, f, \alpha)| &= \left| \sum_{i=1}^k (x_i - x_{i-1}) \cdot (f_n(\alpha_i) - f(\alpha_i)) \right| \\ &\leq \sum_{i=1}^k (x_i - x_{i-1}) \cdot |f_n(\alpha_i) - f(\alpha_i)| \\ &\stackrel{(56)}{<} \sum_{i=1}^k (x_i - x_{i-1}) \cdot \frac{\varepsilon}{3 \cdot (b-a)} = \frac{\varepsilon}{3}. \end{aligned} \quad (59)$$

Da f_n integrierbar ist, gibt es wegen des Folgenkriteriums für Integrierbarkeit 19.19 zudem ein $m_\varepsilon \in \mathbb{N}$, so daß

$$\left| ZS(f_n, Z_m, \alpha^m) - \int_a^b f_n(x) dx \right| < \frac{\varepsilon}{3} \quad (60)$$

für alle $m \geq m_\varepsilon$.

Damit erhalten wir für $m \geq m_\varepsilon$ insgesamt

$$\begin{aligned} |ZS(f, Z_m, \alpha^m) - I| &\leq \\ &\leq |ZS(f, Z_m, \alpha^m) - ZS(f_n, Z_m, \alpha^m)| + \left| ZS(f_n, Z_m, \alpha^m) - \int_a^b f_n(x) dx \right| + \left| \int_a^b f_n(x) dx - I \right| \\ &\stackrel{(58), (59), (60)}{<} \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Damit haben wir gezeigt, daß $ZS(f, Z_m, \alpha^m) \rightarrow I$ und die Behauptung folgt dann aus dem Folgenkriterium für Integrierbarkeit 19.19. \square

Korollar 20.25 (Integration von Potenzreihen)

Es sei $\sum_{n=0}^{\infty} a_n \cdot t^n$ ein Potenzreihe über \mathbb{R} mit Konvergenzradius $r > 0$.

Dann ist die Funktion $f : (-r, r) \rightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} a_n \cdot x^n$ auf jedem Intervall $[a, b] \subset (-r, r)$ integrierbar und

$$F : (-r, r) \rightarrow \mathbb{R} : y \mapsto \sum_{n=0}^{\infty} \frac{a_n}{n+1} \cdot y^{n+1}$$

ist eine Stammfunktion von f . Sie entsteht durch gliedweises Integrieren.

Beweis: Die Folge stetiger Funktionen $(f_n)_{n \in \mathbb{N}}$ mit

$$f_n : (-r, r) \rightarrow \mathbb{R} : x \mapsto \sum_{k=0}^n a_k \cdot x^k$$

konvergiert auf $[a, b]$ gleichmäßig gegen f (siehe Satz 15.4). Also ist f nach Satz 15.6 stetig auf $[a, b]$. Für $y \in (-r, r)$ gilt nach Satz 20.23 zudem

$$\int_0^y f(x) dx = \lim_{n \rightarrow \infty} \int_0^y f_n(x) dx \stackrel{19.21}{=} \lim_{n \rightarrow \infty} \sum_{k=0}^n a_k \cdot \int_0^y x^k dx \stackrel{20.7}{=} \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{a_k}{k+1} \cdot y^{k+1}.$$

Damit ist die Behauptung gezeigt. \square

Man kann durch gliedweises Integrieren Stammfunktionen berechnen oder auch Potenzreihendarstellungen von Funktionen aus der Potenzreihendarstellung ihrer Ableitungen herleiten, wie wir im folgenden Beispiel sehen werden.

Beispiel 20.26 (Reihenentwicklung durch gliedweise Integration)

a. Die Potenzreihe zur Exponentialfunktion ist

$$\sum_{n=0}^{\infty} \frac{t^n}{n!}.$$

Durch gliedweises Integrieren erhalten wir die Potenzreihe

$$\sum_{n=0}^{\infty} \frac{1}{n+1} \cdot \frac{t^{n+1}}{n!} = \sum_{n=1}^{\infty} \frac{t^n}{n!},$$

und diese definiert eine Stammfunktion von \exp . Sie unterscheidet sich von der bereits bekannten Stammfunktion \exp von \exp um die Konstante 1.

b. Die Funktion $f : (-1, \infty) \rightarrow \mathbb{R} : x \mapsto \ln(1+x)$ ist differenzierbar mit Ableitung

$$f' : (-1, \infty) \rightarrow \mathbb{R} : x \mapsto \frac{1}{1+x}.$$

Mit Hilfe der geometrischen Reihe sehen wir, daß

$$f'(x) = \frac{1}{1-(-x)} = \sum_{n=0}^{\infty} (-x)^n = \sum_{n=0}^{\infty} (-1)^n \cdot x^n$$

für alle $x \in (-1, 1)$ gilt. f' ist dort also durch die Potenzreihe $\sum_{n=0}^{\infty} (-1)^n \cdot t^n$ gegeben. Durch gliedweises Integrieren finden wir eine Potenzreihendarstellung einer Stammfunktion von f' :

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} \cdot t^{n+1} = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \cdot t^n$$

auf dem Intervall $(-1, 1)$. Da auch f auf $(-1, 1)$ eine Stammfunktion von f' ist und zwei Stammfunktionen sich nur um eine Konstante c unterscheiden, werten wir f und diese Potenzreihe in $a = 0$ aus, um c zu bestimmen. Wir erhalten damit

$$c = f(0) - \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \cdot 0^n = \ln(1) - 0 = 0.$$

Wir haben also eine Potenzreihendarstellung für f auf $(-1, 1)$ gefunden; für $x \in (-1, 1)$ gilt

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \cdot x^n.$$

Daraus ergibt sich dann die Potenzreihendarstellung für den natürlichen Logarithmus

$$\ln(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \cdot (x-1)^n$$

für alle $x \in (0, 2)$, d.h. die Taylorreihenentwicklung des natürlichen Logarithmus aus Beispiel 18.33 gilt auf ganz $(0, 2)$, und sie gilt auch im Punkt $a = 2$, wie wir dort bereits gesehen haben.

Aufgaben

Aufgabe 20.27

Sei $f : (-1, 1) \rightarrow \mathbb{R}$, $x \mapsto \arctan(x)$.

- Finde die Taylorreihenentwicklung von f mit Hilfe gliedweiser Integration.
- Bestimme eine Reihendarstellung für $\frac{\pi}{4}$ und π .

Aufgabe 20.28

Es sei I ein Intervall und $f : I \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion ohne Nullstelle. Zeige, daß $\ln(|f|)$ eine Stammfunktion von $\frac{f'}{f}$ auf I ist.

Aufgabe 20.29

Berechne mittels Partialbruchzerlegung eine Stammfunktion von

$$x \mapsto \frac{x^3}{x^4 - 4x^3 + 8x^2 - 16x + 16}.$$

§ 21 Uneigentliche Integrale

Definition 21.1 (Uneigentliche Integrale)

- a. Es seien $a \in \mathbb{R}$ und $b \in \mathbb{R} \cup \{\infty\}$ mit $a < b$, und $f : [a, b) \rightarrow \mathbb{R}$ sei auf $[a, y]$ integrierbar für alle $y \in (a, b)$. Falls der Grenzwert

$$\int_a^b f(x) dx := \lim_{y \rightarrow b} \int_a^y f(x) dx \in \mathbb{R} \cup \{\infty, -\infty\}$$

existiert, so nennen wir ihn ein *uneigentliches Integral*. Ist $\int_a^b f(x) dx \in \mathbb{R}$, so nennen wir das uneigentliche Integral *konvergent*.

- b. Es seien $b \in \mathbb{R}$ und $a \in \mathbb{R} \cup \{-\infty\}$ mit $a < b$, und $f : (a, b] \rightarrow \mathbb{R}$ sei auf $[y, b]$ integrierbar für alle $y \in (a, b)$. Falls der Grenzwert

$$\int_a^b f(x) dx := \lim_{y \rightarrow a} \int_y^b f(x) dx \in \mathbb{R} \cup \{\infty, -\infty\}$$

existiert, so nennen wir ihn ein *uneigentliches Integral*. Ist $\int_a^b f(x) dx \in \mathbb{R}$, so nennen wir das uneigentliche Integral *konvergent*.

- c. Es seien $a \in \mathbb{R} \cup \{-\infty\}$ und $b \in \mathbb{R} \cup \{\infty\}$ mit $a < b$ und für $f : (a, b) \rightarrow \mathbb{R}$ gebe es ein $c \in (a, b)$, so daß die uneigentlichen Integrale $\int_c^b f(x) dx$ und $\int_a^c f(x) dx$ existieren, dann nennen wir auch

$$\int_a^b f(x) dx := \int_a^c f(x) dx + \int_c^b f(x) dx$$

ein *uneigentliches Integral*. Aus der Additivität des Integrals folgt, daß die Definition der linken Seite unabhängig von der Wahl von c ist.

Beispiel 21.2

- a. Das folgende uneigentliche Integral ist konvergent mit Grenzwert 1:

$$\int_0^{\infty} \exp(-x) dx = \lim_{y \rightarrow \infty} \int_0^y \exp(-x) dx = \lim_{y \rightarrow \infty} -\exp(-x) \Big|_0^y = \lim_{y \rightarrow \infty} 1 - \exp(-y) = 1.$$

- b. Es sei $a \in \mathbb{R}$ mit $a \neq 1$. Dann gilt für das uneigentliche Integral

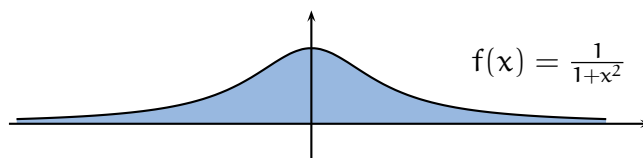
$$\begin{aligned} \int_1^{\infty} \frac{1}{x^a} dx &= \lim_{y \rightarrow \infty} \int_1^y x^{-a} dx = \lim_{y \rightarrow \infty} \frac{x^{-a+1}}{-a+1} \Big|_1^y \\ &= \lim_{y \rightarrow \infty} \frac{y^{1-a}}{1-a} - \frac{1}{1-a} = \begin{cases} \infty, & \text{falls } a < 1, \\ \frac{1}{a-1}, & \text{falls } a > 1. \end{cases} \end{aligned}$$

- c. Das folgende uneigentliche Integral ist konvergent mit Grenzwert 2:

$$\int_0^1 \frac{1}{\sqrt{1-x}} dx = \lim_{y \rightarrow 1} -2 \cdot \sqrt{1-x} \Big|_0^y = 2 - \lim_{y \rightarrow 1} 2 \cdot \sqrt{1-y} = 2.$$

d. Der Flächeninhalt unter dem Graphen von $f: \mathbb{R} \rightarrow \mathbb{R}: x \mapsto \frac{1}{1+x^2}$ ist π :

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{1}{1+x^2} dx &= \int_0^{\infty} \frac{1}{1+x^2} dx + \int_{-\infty}^0 \frac{1}{1+x^2} dx \\ &= \lim_{y \rightarrow \infty} \arctan(x) \Big|_0^y + \lim_{y \rightarrow -\infty} \arctan(x) \Big|_y^0 \\ &= \lim_{y \rightarrow \infty} \arctan(y) - \lim_{y \rightarrow -\infty} \arctan(y) = \frac{\pi}{2} + \frac{\pi}{2} = \pi. \end{aligned}$$



Bemerkung 21.3

Die *Linearität* und die *Monotonie* des Integrals (siehe Korollar 19.21) sowie die *Additivität* des Integrals (siehe Proposition 19.24) und die *Dreiecksungleichung* für Integrale (siehe Proposition 19.26) gelten analog auch für uneigentliche Integrale. Der Beweis folgt unmittelbar aus den entsprechenden Aussagen für Integrale zusammen mit den Grenzwertsätzen 13.10.

Lemma 21.4 (Monotoniekriterium für uneigentliche Integrale)

Es sei $a \in \mathbb{R}$ und $b \in \mathbb{R} \cup \{\infty\}$ mit $a < b$, und es sei $f: [a, b) \rightarrow \mathbb{R}_{\geq 0}$ integrierbar auf $[a, y]$ für alle $y \in (a, b)$. Gibt es ein $s \in \mathbb{R}$ mit $\int_a^y f(x) dx < s$ für alle $y \in (a, b)$, so ist $\int_a^b f(x) dx$ konvergent.

Beweis: Die Funktion

$$F: [a, b) \rightarrow \mathbb{R}: y \mapsto \int_a^y f(x) dx$$

ist monoton wachsend, da f nur nicht-negative Werte annimmt.

Wir betrachten nun eine monoton wachsende Folge $(a_n)_{n \in \mathbb{N}}$ und die zugehörige Folge $(F(a_n))_{n \in \mathbb{N}}$ von Funktionswerten. Diese ist monoton wachsend und beschränkt durch s , mithin ist sie konvergent, d.h. es gibt ein $I \in \mathbb{R}$ mit

$$\int_a^{a_n} f(x) dx = F(a_n) \rightarrow I.$$

Zu gegebenem $\varepsilon > 0$ gibt es also ein $n_\varepsilon \in \mathbb{N}$, so daß

$$I - F(a_{n_\varepsilon}) = |F(a_{n_\varepsilon}) - I| < \varepsilon.$$

Wir wollen nun zeigen, daß $\int_a^b f(x) dx$ gegen I konvergiert und unterscheiden dazu die beiden Fälle $b = \infty$ und $b \in \mathbb{R}$.

1. Fall: $b = \infty$: Für $\varepsilon > 0$ setzen wir nun $s_\varepsilon := a_{n_\varepsilon}$, so daß für $y \in [a, \infty)$ mit $y > a_{n_\varepsilon}$ dann

$$|F(y) - I| = I - F(y) \leq I - F(a_{n_\varepsilon}) < \varepsilon,$$

gilt, da F monoton wachsend ist. Mithin gilt

$$\int_a^\infty f(x) dx = \lim_{y \rightarrow \infty} F(y) = I.$$

2. Fall: $b \in \mathbb{R}$: Für $\varepsilon > 0$ setzen wir $\delta_\varepsilon := b - a_{n_\varepsilon} > 0$, so daß für $y \in [a, b]$ mit $b - y = |y - b| < \delta_\varepsilon = b - a_{n_\varepsilon}$ auch $y > a_{n_\varepsilon}$ und damit

$$|F(y) - I| = I - F(y) \leq I - F(a_{n_\varepsilon}) < \varepsilon,$$

gilt, da F monoton wachsend ist. Mithin gilt wieder

$$\int_a^\infty f(x) dx = \lim_{y \rightarrow \infty} F(y) = I.$$

□

Satz 21.5 (Integralkriterium für Reihen)

Es sei $a \in \mathbb{N}$ und $f: [a, \infty) \rightarrow \mathbb{R}_{\geq 0}$ sei monoton fallend und auf $[a, c]$ integrierbar für alle $c \in [a, \infty)$. Dann gilt:

$$\int_a^\infty f(x) dx \text{ ist konvergent} \iff \sum_{n=a}^\infty f(n) \text{ ist konvergent.}$$

In dieser Situation gilt zudem

$$\sum_{n=a+1}^\infty f(n) \leq \int_a^\infty f(x) dx \leq \sum_{n=a}^\infty f(n).$$

Beweis: Für $m \in \mathbb{N}$ mit $m > a$ betrachten wir die Zerlegung

$$Z_m := (a, a+1, a+2, \dots, m)$$

des Intervalls $[a, m]$. Da f auf $[a, m]$ monoton fallend ist, erhalten wir

$$\sum_{n=a+1}^m f(n) = \text{US}(f, Z_m) \leq \int_a^m f(x) dx \leq \text{OS}(f, Z_m) = \sum_{n=a}^{m-1} f(n). \quad (61)$$

Ist die Reihe $\sum_{n=a}^\infty f(n)$ konvergent und $y \in (a, \infty)$, so wählen wir ein $m \in \mathbb{N}$ mit $y \leq m$ und aus (61) folgt dann

$$\int_a^y f(x) dx \leq \int_a^m f(x) dx \leq \sum_{n=a}^{m-1} f(n) \leq \sum_{n=a}^\infty f(n),$$

so daß das Integral $\int_a^\infty f(x) dx$ nach dem Monotoniekriterium 21.4 konvergent ist.

Ist umgekehrt das Integral $\int_a^\infty f(x) dx$ konvergent, so ist die Folge der Partialsummen

$$s_m := \sum_{n=a}^m f(n) \stackrel{(61)}{\leq} f(a) + \int_a^m f(x) dx \leq f(a) + \int_a^\infty f(x) dx$$

monoton wachsend und beschränkt, mithin konvergent.

Die Abschätzung für die Grenzwerte der Reihen und des Integrals folgt unmittelbar aus (61), indem man m gegen unendlich gehen läßt. □

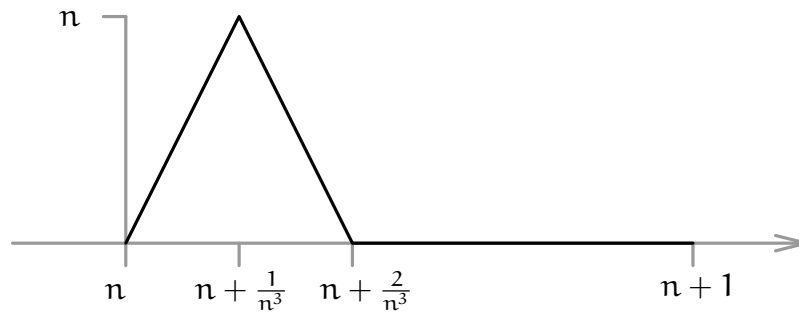
Beispiel 21.6

Ist $a \in \mathbb{R}$ mit $a > 1$, so ist die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^a}$ konvergent, nach Satz 21.5 und Teil b. von Beispiel 21.2 .

Bemerkung 21.7

Wenn $\int_a^{\infty} f(x) dx$ konvergiert, so muß *nicht* $\lim_{x \rightarrow \infty} f(x) = 0$ gelten!

Wir betrachten eine Funktion $f : [1, \infty) \rightarrow \mathbb{R}$, die auf dem Intervall $[n, n+1]$ den folgenden Graphen besitzt:



Dann ist

$$\int_n^{n+1} f(x) dx = \frac{1}{2} \cdot \frac{2}{n^3} \cdot n = \frac{1}{n^2}$$

der Flächeninhalt des obigen Dreiecks. Also ist das uneigentliche Integral

$$\int_1^{\infty} f(x) dx = \lim_{n \rightarrow \infty} \int_1^{n+1} f(x) dx = \lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{1}{k^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

konvergent, aber $\lim_{x \rightarrow \infty} f(x) \neq 0$, da

$$f\left(n + \frac{1}{n^3}\right) = n \rightarrow \infty.$$

Aufgaben**Aufgabe 21.8**

- a. Zeige, für $y \in (0, \infty)$ ist das uneigentliche Integral

$$\int_0^{\infty} x^{y-1} \cdot \exp(-x) dx$$

konvergent.

- b. Die Funktion $\Gamma : (0, \infty) \rightarrow \mathbb{R} : y \mapsto \int_0^{\infty} x^{y-1} \cdot \exp(-x) dx$ erfüllt die Funktionalgleichung

$$\Gamma(y+1) = y \cdot \Gamma(y)$$

für $y \in (0, \infty)$.

- c. Für $n \in \mathbb{N}$ gilt $\Gamma(n+1) = n!$.

KAPITEL III

Lineare Algebra

Im folgenden wollen wir die Theorie der Vektorräume und der linearen Abbildungen studieren, unter anderem mit dem Ziel, die Lösungsmengen von linearen Gleichungssystemen zu verstehen und berechnen zu können. Der Analysis lagen die Körper der reellen und komplexen Zahlen zugrunde. Wesentlichster Baustein neben der Addition und Multiplikation waren der Absolutbetrag mit Werten in \mathbb{R} und die Ordnungsrelation auf \mathbb{R} , die \mathbb{R} zu einem vollständigen angeordneten Körper machte. Für die lineare Algebra spielen der Absolutbetrag und die Ordnungsrelation keine Rolle mehr. Wir kommen ohne ε 's und δ 's und komplizierte Abschätzungen aus. Deshalb können wir unser Arsenal an Grundstrukturen, mit denen wir arbeiten wollen, auch erweitern.

K wird im folgenden einen *beliebigen Körper* bezeichnen,

etwa \mathbb{Q} , \mathbb{R} oder \mathbb{C} oder auch einen endlichen Körper wie etwa \mathbb{F}_2 in Beispiel 7.6. Manchmal können wir sogar auf die Division verzichten und statt einem Körper eine Struktur wie die ganzen Zahlen \mathbb{Z} zugrunde legen.

§ 22 Rechnen mit Matrizen

Definition 22.1 (Matrizen und der K^n)

Es seien $m, n \geq 1$ zwei positive ganze Zahlen.

- a. Eine $m \times n$ -*Matrix* über K ist ein rechteckiges Schema A mit Einträgen aus K der Form

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

Wenn keine Unklarheiten zu befürchten sind, schreiben wir verkürzt auch

$$A = (a_{ij})_{i=1, \dots, m; j=1, \dots, n} = (a_{ij}).$$

- b. Die Menge aller $m \times n$ -Matrizen über K wird mit

$$\text{Mat}(m \times n, K)$$

bezeichnet, und falls $m = n$, dann auch kurz mit $\text{Mat}_n(K) = \text{Mat}(n, K)$ und man spricht von *quadratischen Matrizen*.

- c. Ist $A = (a_{ij})$ eine $m \times n$ -Matrix, dann bezeichnen wir

$$\mathbf{a}_i := (a_{i1}, \dots, a_{in})$$

als den i -ten *Zeilenvektor* von A und

$$\mathbf{a}^j := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

als den j -ten *Spaltenvektor* von A .

- d. Ist $A = (a_{ij}) \in \text{Mat}(m \times n, K)$, so heißt die $n \times m$ -Matrix

$$A^t := \begin{pmatrix} a_{11} & a_{21} & \dots & a_{m1} \\ a_{12} & a_{22} & \dots & a_{m2} \\ \vdots & \vdots & & \vdots \\ a_{1n} & a_{2n} & \dots & a_{mn} \end{pmatrix},$$

d. h. für $A^t = (a'_{ij})$ gilt $a'_{ij} = a_{ji}$, die *Transponierte* von A .

- e. Schließlich definieren wir

$$K^n := \text{Mat}(n \times 1, K) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_1, \dots, x_n \in K \right\}.$$

Die Elemente von K^n heißen *Vektoren* oder *Punkte* im K^n . x_i heißt die i -te *Komponente* des Vektors x .

Definition 22.2 (Operationen mit Matrizen)

- a. Es seien $A = (a_{ij}), B = (b_{ij}) \in \text{Mat}(m \times n, K)$ und $\lambda \in K$. Dann definiert man

$$A + B := (a_{ij} + b_{ij}) = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \dots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \dots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \dots & a_{mn} + b_{mn} \end{pmatrix},$$

sowie

$$\lambda \cdot A := (\lambda a_{ij}) = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \dots & \lambda a_{mn} \end{pmatrix}.$$

- b. Sind $A = (a_{ij}) \in \text{Mat}(m \times n, K)$ und $B = (b_{jk}) \in \text{Mat}(n \times p, K)$ zwei Matrizen, wobei A genauso viele Spalten wie B Zeilen hat. Dann definieren wir das *Matrixprodukt* durch

$$A \circ B := C, \quad \text{mit } C = (c_{ik}) \in \text{Mat}(m \times p, K) \quad \text{und} \quad c_{ik} := \sum_{j=1}^n a_{ij} b_{jk}.$$

Beispiel 22.3

Folgende Matrizen $A, B \in \text{Mat}(2 \times 3, \mathbb{K})$ und $C \in \text{Mat}(3 \times 2, \mathbb{K})$ seien gegeben:

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 3 & 1 & 2 \end{pmatrix}, B = \begin{pmatrix} 3 & 2 & 1 \\ 0 & 4 & 5 \end{pmatrix}, C = \begin{pmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 1 \end{pmatrix}.$$

Dann gilt:

$$A + B = \begin{pmatrix} 4 & 2 & 3 \\ 3 & 5 & 7 \end{pmatrix}, 3 \cdot A = \begin{pmatrix} 3 & 0 & 6 \\ 9 & 3 & 6 \end{pmatrix} \quad \text{und} \quad A \circ C = \begin{pmatrix} 7 & 2 \\ 11 & 3 \end{pmatrix}.$$

Bemerkung 22.4

- Die in Definition 22.2 a. definierte Addition zweier Matrizen definiert auf $\text{Mat}(m \times n, \mathbb{K})$ offensichtlich eine zweistellige Operation, bezüglich derer $\text{Mat}(m \times n, \mathbb{K})$ eine abelsche Gruppe $(\text{Mat}(m \times n, \mathbb{K}), +)$ wird, wie man leicht nachprüft.
- Wir werden meist kurz λA bzw. λx schreiben, statt $\lambda \cdot A$ bzw. $\lambda \cdot x$, wenn $\lambda \in \mathbb{K}$, $A \in \text{Mat}(m \times n, \mathbb{K})$ und $x \in \mathbb{K}^n$.
- Wir schreiben statt $A \circ B$ häufig kurz AB , insbesondere auch Ax statt $A \circ x$.
- Spaltenvektoren nehmen im Skript sehr viel Raum ein. Um platzsparend arbeiten zu können, werden wir deshalb statt den Spaltenvektor $x \in \mathbb{K}^n$ als

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

anzugeben, meist den *transponierten* Zeilenvektor

$$x = (x_1 \ \dots \ x_n)^t$$

betrachten, und um Mißverständnissen vorzubeugen, fügen wir zudem meist Kommata als Trennsymbole ein

$$x = (x_1, \dots, x_n)^t.$$

- Man beachte, daß das Produkt nur dann definiert ist, wenn A soviele Spalten wie B Zeilen hat. Das Produkt $A \circ B$ hat dann soviele Zeilen wie A und soviele Spalten wie B .

Jede Matrix definiert wie folgt eine Abbildung.

Definition 22.5 (Die Abbildung f_A)

Ist $A \in \text{Mat}(m \times n, \mathbb{K})$, so definieren wir

$$f_A : \mathbb{K}^n \rightarrow \mathbb{K}^m : x \mapsto Ax = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}.$$

f_A heißt die zu A assoziierte oder zu A gehörige Abbildung.

Beispiel 22.6

Die Matrix

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

definiert die Abbildung

$$f_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x_1, x_2)^t \mapsto (x_1 + 2x_2, 3x_2)^t.$$

Bemerkung 22.7 (Einheitsvektoren)

Um den Zusammenhang zwischen A und f_A besser zu verstehen, betrachten wir für $i = 1, \dots, n$ den i -ten *Einheitsvektor* $e_i = (\delta_{1i}, \dots, \delta_{ni})^t \in \mathbb{K}^n$, wobei

$$\delta_{ji} := \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases}$$

das *Kronecker Symbol* ist, d. h.

$$e_i := \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

wobei die Eins in der i -ten Komponente steht.

Es ist dann

$$f_A(e_i) = \begin{pmatrix} \sum_{j=1}^n a_{1j} \delta_{ji} \\ \vdots \\ \sum_{j=1}^n a_{mj} \delta_{ji} \end{pmatrix} = \begin{pmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{pmatrix} = a^i,$$

d. h. die i -te Spalte von A ist das Bild des i -ten Einheitsvektors unter f_A .

Hieraus folgt insbesondere, daß A durch f_A eindeutig bestimmt ist.

Lemma 22.8 (Einfache Rechenregeln für Matrizen)

Für $x, y \in \mathbb{K}^n$, $A, B \in \text{Mat}(m \times n, \mathbb{K})$, $C \in \text{Mat}(n \times p, \mathbb{K})$ und $\lambda \in \mathbb{K}$ gelten:

- $A(x + y) = Ax + Ay$ und $A(\lambda x) = \lambda Ax$,
- $\lambda \cdot (A \circ C) = (\lambda \cdot A) \circ C = A \circ (\lambda \cdot C)$,
- $f_{A+B} = f_A + f_B$, und
- $f_{\lambda A} = \lambda f_A$.

Beweis: Der Beweis ist eine Übungsaufgabe. □

Wir wollen jetzt sehen, wie sich die Multiplikation von Matrizen mit den zugehörigen Abbildungen verträgt.

Satz 22.9 (Matrixmultiplikation und Komposition)

Für $A \in \text{Mat}(m \times n, K)$ und $B \in \text{Mat}(n \times p, K)$ gilt:

$$f_{A \circ B} = f_A \circ f_B.$$

Beweis: Da Definitionsbereich und Wertebereich von beiden Abbildungen übereinstimmen, reicht es zu zeigen:

$$(f_{A \circ B})(x) = (f_A \circ f_B)(x), \quad \text{für alle } x \in K^p.$$

Seien $A = (a_{ij})$ und $B = (b_{jk})$, und sei $x = (x_1, \dots, x_p)^t \in K^p$ gegeben.

$$\begin{aligned} (f_{A \circ B})(x) &= (A \circ B)x = \begin{pmatrix} \sum_{j=1}^n a_{1j}b_{j1} & \cdots & \sum_{j=1}^n a_{1j}b_{jp} \\ \vdots & & \vdots \\ \sum_{j=1}^n a_{mj}b_{j1} & \cdots & \sum_{j=1}^n a_{mj}b_{jp} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix} \\ &= \begin{pmatrix} \sum_{k=1}^p \sum_{j=1}^n a_{1j}b_{jk}x_k \\ \vdots \\ \sum_{k=1}^p \sum_{j=1}^n a_{mj}b_{jk}x_k \end{pmatrix}. \end{aligned}$$

Ferner gilt:

$$\begin{aligned} (f_A \circ f_B)(x) &= f_A(Bx) = A(Bx) = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} \sum_{k=1}^p b_{1k}x_k \\ \vdots \\ \sum_{k=1}^p b_{nk}x_k \end{pmatrix} \\ &= \begin{pmatrix} \sum_{j=1}^n a_{1j} \sum_{k=1}^p b_{jk}x_k \\ \vdots \\ \sum_{j=1}^n a_{mj} \sum_{k=1}^p b_{jk}x_k \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n \sum_{k=1}^p a_{1j}b_{jk}x_k \\ \vdots \\ \sum_{j=1}^n \sum_{k=1}^p a_{mj}b_{jk}x_k \end{pmatrix}. \end{aligned}$$

Beide Ausdrücke stimmen (bis auf die Reihenfolge der Summation) überein, was zu zeigen war. \square

Korollar 22.10 (Die Matrixmultiplikation ist assoziativ.)

Für $A \in \text{Mat}(m \times n, K)$, $B \in \text{Mat}(n \times p, K)$ und $C \in \text{Mat}(p \times q, K)$ gilt

$$A \circ (B \circ C) = (A \circ B) \circ C.$$

Beweis: Dies folgt aus Satz 22.9, da die Komposition von Abbildungen assoziativ ist und da eine Matrix A durch die Abbildung f_A eindeutig bestimmt ist. \square

Man kann die Aussage des Korollars natürlich auch direkt nachweisen, was auf die gleiche Rechnung wie in 22.9 führt - statt des einen Vektors x hat man die q Spaltenvektoren von C zu multiplizieren, was etwas mehr Schreibarbeit bedeutet.

Lemma 22.11 (Distributivgesetze)

Sind $A, B \in \text{Mat}(m \times n, K)$ und $C, D \in \text{Mat}(n \times p, K)$, so gelten die Distributivgesetze:

$$A \circ (C + D) = A \circ C + A \circ D,$$

sowie

$$(A + B) \circ C = A \circ C + B \circ C.$$

Beweis: Die Aussage kann wie Korollar 22.10 aus Lemma 22.8 und Satz 22.9 abgeleitet werden und sei dem Leser als Übung anempfohlen. \square

Definition 22.12 (Invertierbare Matrizen)

Eine Matrix $A \in \text{Mat}_n(\mathbb{K})$ heißt *invertierbar*, falls es eine Matrix $A^{-1} \in \text{Mat}_n(\mathbb{K})$ gibt, so daß

$$A \circ A^{-1} = A^{-1} \circ A = \mathbf{1}_n,$$

wobei die Matrix $\mathbf{1}_n = (\delta_{ij})_{i,j=1,\dots,n} \in \text{Mat}_n(\mathbb{K})$ die *Einheitsmatrix* ist, die auf der Diagonalen Einsen und außerhalb der Diagonalen Nullen als Einträge hat. Eine Matrix mit der Eigenschaft von A^{-1} nennt man eine *Inverse* zu A .

Satz 22.13 (Die allgemeine lineare Gruppe $\text{Gl}_n(\mathbb{K})$)

Die Menge der invertierbaren $n \times n$ -Matrizen

$$\text{Gl}_n(\mathbb{K}) = \{A \in \text{Mat}_n(\mathbb{K}) \mid A \text{ ist invertierbar}\}$$

ist eine Gruppe mit neutralem Element $\mathbf{1}_n$, die für $n > 1$ nicht kommutativ ist.

Insbesondere ist die Inverse zu A eindeutig bestimmt und es gelten für $A, B \in \text{Gl}_n(\mathbb{K})$

$$(AB)^{-1} = B^{-1}A^{-1} \quad \text{und} \quad (A^{-1})^{-1} = A.$$

Beweis: Der Beweis ist eine Übungsaufgabe. \square

Aufgaben

Aufgabe 22.14

Zeige, daß die Matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{Mat}_2(\mathbb{K})$$

genau dann invertierbar ist, wenn $ad - bc \neq 0$ ist. Die Inverse ist dann

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Aufgabe 22.15

Es sei $A \in \text{Mat}(m \times n, \mathbb{K})$ und $B \in \text{Mat}(n \times p, \mathbb{K})$. Zeige, $(AB)^t = B^t A^t$.

Aufgabe 22.16 (Nilpotente Matrizen)

Es sei $N = (n_{ij}) \in \text{Mat}_n(\mathbb{K})$ eine Matrix, für die die Einträge auf der oberen Nebendiagonale alle Eins sind und für die alle anderen Einträge Null sind, d.h. $n_{ij} = \delta_{j-i,1}$.

Zeige für $k = 1, \dots, n$, daß die Einträge der Matrix $N^k = (n_{ij}^{(k)})$ auf der k -ten oberen Nebendiagonale alle eins und alle anderen Einträge null sind, d.h. $n_{ij}^{(k)} = \delta_{j-i,k}$. Insbesondere ist $N^n = 0$ und $N^k \neq 0$ für $k < n$.

§ 23 Vektorräume und lineare Abbildungen

A) Vektorräume

Definition 23.1 (Vektorräume)

Ein K -Vektorraum (oder Vektorraum über K) besteht aus einer nicht-leeren Menge V sowie einer zweistelligen Operation

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

die *Vektoraddition* genannt wird, und einer zweistelligen Operation

$$\cdot : K \times V \rightarrow V : (\lambda, x) \mapsto \lambda \cdot x = \lambda x,$$

die *Skalarmultiplikation* genannt wird, so daß die folgenden Gesetze gelten:

- a. $(V, +)$ ist eine abelsche Gruppe,
- b. für $\lambda, \mu \in K$ und $x, y \in V$ gelten:
 - (i) $(\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x$, (“verallgemeinertes Distributivgesetz”)
 - (ii) $\lambda \cdot (x + y) = \lambda \cdot x + \lambda \cdot y$, und (“verallgemeinertes Distributivgesetz”)
 - (iii) $(\lambda \cdot \mu) \cdot x = \lambda \cdot (\mu \cdot x)$. (“verallgemeinertes Assoziativgesetz”)
 - (iv) $1 \cdot x = x$.

Die Elemente aus V nennt man *Vektoren* und die aus K *Skalare*. Der *Nullvektor*, d. h. das neutrale Element aus V bezüglich der Addition, wird mit 0 bzw. mit 0_V bezeichnet und das neutrale Element von $(K, +)$ ebenfalls mit 0 bzw. mit 0_K .

Beispiel 23.2

- a. Der *Nullraum* $V = \{0\}$ mit $\lambda \cdot 0 = 0$ für alle $\lambda \in K$ ist für jeden Körper K ein K -Vektorraum. Man bezeichnet den Nullraum auch mit K^0 .
- b. Der Körper K selbst mit der Körperaddition als Vektoraddition und der Körpermultiplikation als Skalarmultiplikation ist ein K -Vektorraum.
- c. \mathbb{R} ist ein \mathbb{Q} -Vektorraum und \mathbb{C} ist ein \mathbb{R} -Vektorraum, jeweils mit der üblichen Addition und Multiplikation.
- d. Die Menge $\text{Mat}(m \times n, K)$ der $m \times n$ -Matrizen über K mittels der in Definition 22.2 definierten Addition und Skalarmultiplikation ist ein K -Vektorraum mit der *Nullmatrix*

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

als Nullvektor.

- e. Damit ist insbesondere K^n mit der komponentenweisen Addition und Skalarmultiplikation ein K -Vektorraum mit $0_{K^n} = (0, \dots, 0)^t$.
Speziell sind \mathbb{R}^n , \mathbb{C}^n und \mathbb{F}_2^n Vektorräume über \mathbb{R} , \mathbb{C} bzw. \mathbb{F}_2 .

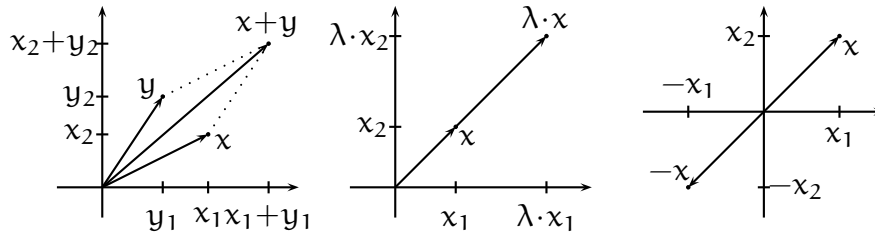


ABBILDUNG 1. Addition und Skalarmultiplikation in \mathbb{R}^2

- f. Ist M eine Menge und V ein K -Vektorraum, so wird die Menge

$$V^M = \{f : M \rightarrow V \mid f \text{ ist eine Abbildung}\}$$

durch die Operationen

$$+ : V^M \times V^M \rightarrow V^M : (f, g) \mapsto (f + g : M \rightarrow V : x \mapsto f(x) + g(x))$$

und

$$\cdot : K \times V^M \rightarrow V^M : (\lambda, f) \mapsto (\lambda \cdot f : M \rightarrow V : x \mapsto \lambda \cdot f(x))$$

zu einem K -Vektorraum, wie man leicht nachrechnet.

Ist z.B. $M = \mathbb{N}$ und $K = V = \mathbb{R}$, so ist

$$\mathbb{R}^{\mathbb{N}} = \{\mathbf{a} : \mathbb{N} \rightarrow \mathbb{R} \mid \mathbf{a} \text{ ist eine Abbildung}\} = \{(\mathbf{a}_n)_{n \in \mathbb{N}} \mid \mathbf{a}_n \in \mathbb{R}\}$$

der Vektorraum der Folgen in \mathbb{R} . Unsere Definitionen sorgen dafür, daß Folgen komponentenweise addiert und mit Skalaren multipliziert werden.

- g. Die Menge

$$K[t] := \left\{ \sum_{k=0}^n \mathbf{a}_k \cdot t^k \mid n \in \mathbb{N}, \mathbf{a}_k \in K \right\}$$

der Polynome mit Koeffizienten in K (vgl. Definition 13.11) wird durch die Addition

$$\sum_{k=0}^m \mathbf{a}_k \cdot t^k + \sum_{k=0}^n \mathbf{b}_k \cdot t^k := \sum_{k=0}^{\max\{m,n\}} (\mathbf{a}_k + \mathbf{b}_k) \cdot t^k,$$

wobei $\mathbf{a}_k = 0$ für $k > m$ und $\mathbf{b}_k = 0$ für $k > n$ gelten soll, und durch die Skalarmultiplikation

$$\lambda \cdot \sum_{k=0}^n \mathbf{a}_k \cdot t^k := \sum_{k=0}^n (\lambda \cdot \mathbf{a}_k) \cdot t^k$$

zu einem K -Vektorraum, wie man leicht nachprüft.

- h. Da man für $M = \{1, \dots, n\}$ eine Abbildung $f : M \rightarrow K$ in eindeutiger Weise durch das Tupel der Bildelemente $(f(1), \dots, f(n))$ beschreiben kann, sieht man leicht, daß die Zuordnung

$$K^M \rightarrow K^n : f \mapsto (f(1), \dots, f(n))^t$$

in diesem Falle eine Bijektion ist. Man prüft überdies leicht nach, daß diese Abbildung ein Vektorraumhomomorphismus im Sinne von Definition 23.19 ist. K^M und K^n sind dann also isomorph.

Lemma 23.3 (Einfache Rechenregeln für Vektoren)

In einem K -Vektorraum gelten folgende Rechenregeln:

- a. $0_K \cdot x = 0_V$ und $\lambda \cdot 0_V = 0_V$ für alle $x \in V$, $\lambda \in K$.
 b. Für $\lambda \in K$ und $x \in V$ gilt:

$$\lambda \cdot x = 0_V \Rightarrow \lambda = 0 \quad \text{oder} \quad x = 0.$$

- c. $(-1) \cdot x = -x$ für alle $x \in V$.

Beweis: Es seien $x \in V$ und $\lambda \in K$ gegeben.

- a. Es gilt:

$$0_V + 0_K \cdot x = 0_K \cdot x = (0_K + 0_K) \cdot x = 0_K \cdot x + 0_K \cdot x,$$

also $0_V = 0_K \cdot x$, wie aus den Kürzungsregeln für $(V, +)$ folgt. Analog gilt:

$$0_V + \lambda \cdot 0_V = \lambda \cdot 0_V = \lambda \cdot (0_V + 0_V) = \lambda \cdot 0_V + \lambda \cdot 0_V,$$

und damit $0_V = \lambda \cdot 0_V$.

- b. Ist $\lambda \in K$ mit $\lambda \neq 0$, dann gibt es ein Inverses $\lambda^{-1} \in K$. Aus $\lambda \cdot x = 0$ folgt dann aber wegen a. und den Vektorraumaxiomen

$$0_V = \lambda^{-1} \cdot 0_V = \lambda^{-1} \cdot (\lambda \cdot x) = (\lambda^{-1} \cdot \lambda) \cdot x = 1 \cdot x = x.$$

- c. Für $x \in V$ gilt:

$$x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = (1 + (-1)) \cdot x = 0_K \cdot x = 0_V.$$

Also ist $(-1) \cdot x$ das (eindeutig bestimmte) additive Inverse zu x .

□

B) Unterräume

Definition 23.4 (Unterräume)

Es sei V ein Vektorraum über K . Eine nicht-leere Teilmenge $U \subseteq V$ von V heißt *Unterraum*, wenn für alle $\lambda \in K$ und $x, y \in U$ gilt

$$\lambda \cdot x \in U \quad \text{und} \quad x + y \in U. \tag{62}$$

Man sagt, U sei *abgeschlossen* bezüglich der Addition und der Skalarmultiplikation. Wir schreiben $U \leq V$, um auszudrücken, daß U ein Unterraum von V ist.

Proposition 23.5 (Unterräume sind Vektorräume.)

Jeder Unterraum eines K -Vektorraums ist selbst ein K -Vektorraum.

Beweis: Als Unterraum eines K -Vektorraums ist U eine nicht-leere Menge. Für die Addition und Skalarmultiplikation, die U von V erbt, gilt nach Voraussetzung

$$U \times U \longrightarrow U : (x, y) \mapsto x + y$$

und

$$K \times U \longrightarrow U : (\lambda, x) \mapsto \lambda \cdot x,$$

sie sind also zweistellige Operationen mit Werten in U !

Da U nicht leer ist, gibt es ein $y \in U$ und mithin folgt aus (62)

$$0_V = 0_K \cdot y \in U.$$

Damit besitzt U offenbar ein neutrales Element der Addition, da für alle $x \in U$ auch

$$0_V + x = x$$

gilt. Ist $x \in U$ beliebig, so ist zudem

$$-x = (-1) \cdot x \in U,$$

so daß U wegen

$$-x + x = 0_V = 0_U$$

auch das Inverse von x enthält. Da das Assoziativgesetz für alle Elemente in V gilt und die Elemente aus U auch in V sind, gilt es automatisch für alle Elemente in U . Wir haben also gezeigt, daß U eine abelsche Gruppe bezüglich $+$ ist.

Die verbleibenden Axiome aus Teil b. in Definition 23.1 vererben sich analog zum Assoziativgesetz unmittelbar von V auf U . Damit ist U also ein K -Vektorraum. \square

Beispiel 23.6

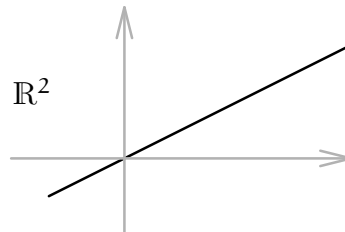
- Ist V ein K -Vektorraum, so ist $\{0_V\}$ stets ein Unterraum von V . Ferner ist V selbst ein Unterraum. Man nennt diese beiden auch die *trivialen Unterräume*.
- Eine Gerade G durch den Ursprung in der Ebene \mathbb{R}^2 mit Steigung m genügt der Gleichung $y = m \cdot x$ und ist deshalb

$$G = \{(x, mx)^t \mid x \in \mathbb{R}\}.$$

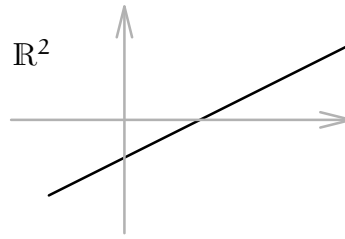
Für $v = (x, mx)^t, w = (x', mx')^t \in G$ und $\lambda \in \mathbb{R}$ gilt dann

$$v + w = (x + x', m \cdot (x + x'))^t, \lambda \cdot v = (\lambda x, m\lambda x)^t \in G.$$

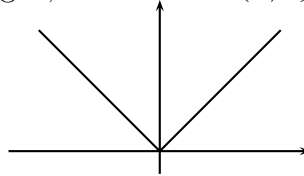
Mithin ist G ein Unterraum von V .



- Eine Gerade in \mathbb{R}^2 , die nicht durch den Ursprung geht, kann kein Unterraum von \mathbb{R}^2 sein, da ein Unterraum den Nullvektor enthalten muß.



- d. Der Graph der Betragsfunktion ist *kein* Unterraum von \mathbb{R}^2 , da $(-1, 1)^t$ und $(1, 1)^t$ auf dem Graphen liegen, ihre Summe $(0, 2)^t$ aber nicht.



- e. Die Menge

$$\mathcal{U} := \{(\mathbf{a}_n)_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid (\mathbf{a}_n)_{n \in \mathbb{N}} \text{ ist konvergent}\}$$

der *konvergenten Folgen* in \mathbb{R} ist ein Unterraum des \mathbb{R} -Vektorraums $\mathbb{R}^{\mathbb{N}}$ aller Folgen. Dies folgt aus den Grenzwertsätzen für Folgen 11.15.

- f. Ist I ein Intervall in \mathbb{R} , so sind die Menge $\mathcal{C}(I, \mathbb{R})$ aller auf I stetigen Abbildungen sowie die Menge $\mathcal{C}^k(I, \mathbb{R})$ aller auf I k -fach stetig differenzierbaren Abbildungen Unterräume des Vektorraums \mathbb{R}^I aller Abbildungen von I nach \mathbb{R} .

Solche Funktionenräume spielen in der Analysis eine große Rolle. Sie sind für kein n isomorph zu \mathbb{R}^n , und sie sind ein gutes Beispiel für den Wert der abstrakten Theorie der Vektorräume.

- g. Ist $n \in \mathbb{N}$ fest vorgegeben, so bilden die Polynome vom Grad höchstens n

$$\mathcal{P}_n := \left\{ \sum_{k=0}^n \mathbf{a}_k \cdot t^k \mid \mathbf{a}_k \in K \right\}$$

einen Unterraum vom Vektorraum der Polynome $K[t]$.

Lemma 23.7 (Durchschnitt von Unterräumen)

Der Durchschnitt beliebig vieler Unterräume ist wieder ein Unterraum.

Beweis: Es seien \mathcal{U}_i für $i \in I$ Unterräume eines K -Vektorraums V . Da $0_V \in \mathcal{U}_i$ für alle $i \in I$, ist $\mathcal{U} := \bigcap_{i \in I} \mathcal{U}_i$ nicht die leere Menge. Es bleibt also zu zeigen, daß für $x, y \in \mathcal{U}$ und $\lambda \in K$ gilt:

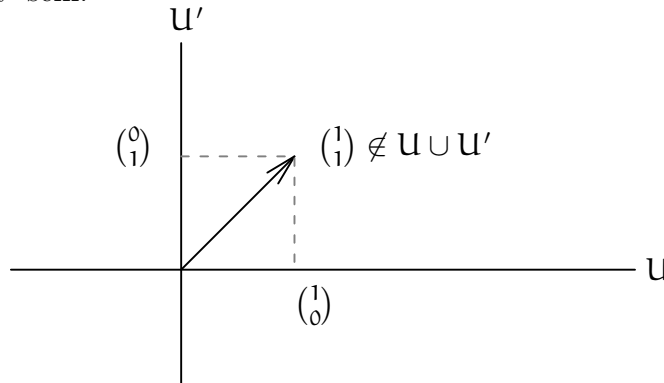
$$x + y \in \mathcal{U} \quad \text{und} \quad \lambda x \in \mathcal{U}.$$

Für ein beliebiges $i \in I$ gilt, da \mathcal{U}_i ein Unterraum von V ist und da $x, y \in \mathcal{U} \subseteq \mathcal{U}_i$, daß $x + y \in \mathcal{U}_i$ und $\lambda x \in \mathcal{U}_i$. Also liegen die Vektoren im Durchschnitt \mathcal{U} . \square

Bemerkung 23.8

Die Vereinigung von zwei Unterräumen ist i. a. kein Unterraum mehr!

Sei etwa U die x -Achse und U' die y -Achse im \mathbb{R}^2 . Beides sind offenbar Unterräume von \mathbb{R}^2 . Dann liegt $(1, 1)^t = e_1 + e_2$ nicht in $U \cup U'$, und mithin kann $U \cup U'$ kein Unterraum von \mathbb{R}^2 sein.



Definition 23.9 (Linearkombination und lineare Hülle)

Es sei V ein K -Vektorraum.

- a. Wir nennen $x \in V$ eine *Linearkombination* von $x_1, \dots, x_r \in V$, falls es $\lambda_1, \dots, \lambda_r \in K$ gibt mit

$$x = \lambda_1 x_1 + \dots + \lambda_r x_r.$$

Ist eines der λ_i ungleich Null, so nennen wir die Linearkombination *nicht-trivial*.

- b. Ist $M \subseteq V$, so nennen wir den Durchschnitt

$$\text{Lin}(M) := \langle M \rangle := \bigcap_{M \subseteq U \leq V} U$$

aller Unterräume von V , die M enthalten, die *lineare Hülle* oder das *Erzeugnis* von M .

Bemerkung 23.10

- a. Man beachte, daß die lineare Hülle von M wegen Lemma 23.7 ein Unterraum von V ist. Aufgrund der Definition ist es *der kleinste* Unterraum, der M enthält.
- b. Ist $M = \emptyset$, so ist $\text{Lin}(M) = \text{Lin}(\emptyset) = \{0\}$.
- c. Eine Linearkombination ist immer eine *endliche* Summe von Vielfachen von Vektoren aus V . In der linearen Algebra wird es *nie* unendliche Summen geben.
- d. Mit Induktion nach der Anzahl der Summanden folgt aus (62) unmittelbar, daß ein Unterraum U abgeschlossen bezüglich endlicher Linearkombinationen von Vektoren aus U ist.

Proposition 23.11 (Lineare Hülle = Menge der Linearkombinationen)

Ist V ein K -Vektorraum und $\emptyset \neq M \subseteq V$, so ist die lineare Hülle von M

$$\text{Lin}(M) = \{\lambda_1 \cdot x_1 + \dots + \lambda_r \cdot x_r \mid r \in \mathbb{N}, x_1, \dots, x_r \in M, \lambda_1, \dots, \lambda_r \in K\} \leq V$$

die Menge aller Linearkombinationen von Elementen in M .

Beweis: Wir setzen $\mathbf{U} := \{\lambda_1 \cdot \mathbf{x}_1 + \dots + \lambda_r \cdot \mathbf{x}_r \mid r \in \mathbb{N}, \mathbf{x}_1, \dots, \mathbf{x}_r \in \mathbf{M}, \lambda_1, \dots, \lambda_r \in \mathbf{K}\}$.

Als Unterraum, der \mathbf{M} enthält, enthält $\text{Lin}(\mathbf{M})$ auch jede endliche Linearkombination von Elementen in \mathbf{M} , also auch \mathbf{U} .

Wir wollen nun zeigen, daß \mathbf{U} ein Unterraum von \mathbf{V} ist, der \mathbf{M} enthält, da er aufgrund der Definition des Erzeugnisses dann auch das Erzeugnis von \mathbf{M} enthält. Dazu beachten wir zunächst, daß für $\mathbf{x} \in \mathbf{M}$ auch $\mathbf{x} = 1 \cdot \mathbf{x} \in \mathbf{U}$ gilt. Also ist $\mathbf{M} \subseteq \mathbf{U}$ und somit ist \mathbf{U} auch nicht leer. Seien nun $\mathbf{x} = \sum_{i=1}^r \lambda_i \cdot \mathbf{x}_i$ und $\mathbf{y} = \sum_{j=1}^s \mu_j \cdot \mathbf{y}_j$ mit $\mathbf{x}_i, \mathbf{y}_j \in \mathbf{M}$ und $\lambda_i, \mu_j \in \mathbf{K}$ sowie $\lambda \in \mathbf{K}$ gegeben. Dann ist

$$\mathbf{x} + \mathbf{y} = \lambda_1 \cdot \mathbf{x}_1 + \dots + \lambda_r \cdot \mathbf{x}_r + \mu_1 \cdot \mathbf{y}_1 + \dots + \mu_s \cdot \mathbf{y}_s \in \mathbf{U},$$

weil es eine endliche Linearkombination von Elementen in \mathbf{M} ist, und ebenso ist

$$\lambda \cdot \mathbf{x} = \sum_{i=1}^r (\lambda \cdot \lambda_i) \cdot \mathbf{x}_i \in \mathbf{U}.$$

Also ist \mathbf{U} ein Unterraum von \mathbf{V} . □

Beispiel 23.12

- a. Ist $\mathbf{M} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subseteq \mathbf{V}$ endlich, so ist das Erzeugnis von \mathbf{M}

$$\text{Lin}(\mathbf{x}_1, \dots, \mathbf{x}_n) := \text{Lin}(\mathbf{M}) = \{\lambda_1 \cdot \mathbf{x}_1 + \dots + \lambda_n \cdot \mathbf{x}_n \mid \lambda_1, \dots, \lambda_n \in \mathbf{K}\}.$$

Insbesondere gilt $\text{Lin}(\mathbf{x}) := \text{Lin}(\{\mathbf{x}\}) = \{\lambda \cdot \mathbf{x} \mid \lambda \in \mathbf{K}\}$.

- b. Die Lineare Hülle der Vektoren $\mathbf{x}_1 = (1, 0)^t$ und $\mathbf{x}_2 = (0, 1)^t$ in \mathbb{R}^2 ist

$$\text{Lin}(\mathbf{x}_1, \mathbf{x}_2) = \{\lambda_1 \cdot \mathbf{x}_1 + \lambda_2 \cdot \mathbf{x}_2 = (\lambda_1, \lambda_2) \mid \lambda_1, \lambda_2 \in \mathbb{R}\} = \mathbb{R}^2.$$

- c. Die Lineare Hülle von $\mathbf{x} = (1, m)^t \in \mathbb{R}^2$ ist die Gerade

$$\text{Lin}(\mathbf{x}) = \{(\lambda, \lambda m)^t \mid \lambda \in \mathbb{R}\}.$$

- d. Es gilt offenbar $\text{Lin}(\mathbf{t}^0, \mathbf{t}^1, \dots, \mathbf{t}^n) = \mathbf{P}_n$.

Proposition 23.13 (Summe zweier Unterräume)

Es seien \mathbf{U} und \mathbf{U}' Unterräume des \mathbf{K} -Vektorraums \mathbf{V} . Dann gilt

$$\mathbf{U} + \mathbf{U}' := \{\mathbf{u} + \mathbf{u}' \mid \mathbf{u} \in \mathbf{U}, \mathbf{u}' \in \mathbf{U}'\} = \text{Lin}(\mathbf{U} \cup \mathbf{U}') \leq \mathbf{V},$$

und wir nennen diesen Unterraum von \mathbf{V} die Summe von \mathbf{U} und \mathbf{U}' .

Beweis: Wegen Proposition 23.11 ist $\mathbf{U} + \mathbf{U}'$ in $\text{Lin}(\mathbf{U} \cup \mathbf{U}')$ enthalten, da die Elemente von $\mathbf{U} + \mathbf{U}'$ Linearkombinationen von Elementen in $\mathbf{U} \cup \mathbf{U}'$ sind.

Umgekehrt ist jede Linearkombination von Elementen in $\mathbf{U} \cup \mathbf{U}'$ von der Form $\sum_{i=1}^r \lambda_i \cdot \mathbf{u}_i + \sum_{j=1}^s \mu_j \cdot \mathbf{u}'_j$ mit $\mathbf{u}_i \in \mathbf{U}$, $\mathbf{u}'_j \in \mathbf{U}'$ und $\lambda_i, \mu_j \in \mathbf{K}$. Da \mathbf{U} und \mathbf{U}' aber Unterräume sind, ist

$$\mathbf{u} := \sum_{i=1}^r \lambda_i \cdot \mathbf{u}_i \in \mathbf{U}$$

und

$$\mathbf{u}' := \sum_{j=1}^s \mu_j \cdot \mathbf{u}'_j \in \mathbf{U}'.$$

Deshalb ist die Linearkombination

$$\sum_{i=1}^r \lambda_i \cdot \mathbf{u}_i + \sum_{j=1}^s \mu_j \cdot \mathbf{u}'_j = \mathbf{u} + \mathbf{u}' \in \mathbf{U} + \mathbf{U}'$$

in $\mathbf{U} + \mathbf{U}'$, und mit Proposition 23.11 enthält $\mathbf{U} + \mathbf{U}'$ auch $\text{Lin}(\mathbf{U} \cup \mathbf{U}')$. \square

Analog zeigt man die allgemeinere Aussage.

Bemerkung 23.14 (Summen von Unterräumen)

- Die Summe zweier Unterräume ersetzt ihre Vereinigung in der Theorie der Vektorräume. Sie ist der kleinste Unterraum, der beide enthält. Im Beispiel aus Bemerkung 23.8 ist die Summe der beiden Unterräume ganz \mathbb{R}^2 .
- Analog zu Proposition 23.13 zeigt man allgemeiner: Sind $\mathbf{U}_1, \dots, \mathbf{U}_n$ Unterräume des K -Vektorraums V , so gilt

$$\mathbf{U}_1 + \dots + \mathbf{U}_n := \{\mathbf{u}_1 + \dots + \mathbf{u}_n \mid \mathbf{u}_i \in \mathbf{U}_i\} = \text{Lin}(\mathbf{U}_1 \cup \dots \cup \mathbf{U}_n).$$

Beispiel 23.15

Jeder Vektor \mathbf{x} in $\mathbf{U} + \mathbf{U}'$ läßt sich schreiben als $\mathbf{x} = \mathbf{u} + \mathbf{u}'$ mit $\mathbf{u} \in \mathbf{U}$ und $\mathbf{u}' \in \mathbf{U}'$, diese Darstellung muß aber nicht eindeutig sein.

Sind z.B. $\mathbf{U} = \text{Lin}((1, 0, 0)^t, (0, 1, 1)^t)$ und $\mathbf{U}' = \text{Lin}((1, 1, 0)^t, (1, 0, 1)^t)$ als Unterräume von \mathbb{R}^3 gegeben, so können wir den Vektor $\mathbf{x} = (1, 0, -1)^t$ auf folgende beiden Weisen als Summe zweier Vektoren in \mathbf{U} und \mathbf{U}' schreiben:

$$\mathbf{x} = (0, -1, -1)^t + (1, 1, 0)^t = (2, 0, 0)^t + (-1, 0, -1)^t.$$

Definition 23.16 (Direkte Summe)

Es seien $\mathbf{U}_1, \dots, \mathbf{U}_n$ Unterräume des K -Vektorraums V . Wir nennen die Summe $\mathbf{U} = \mathbf{U}_1 + \dots + \mathbf{U}_n$ eine *direkte Summe*, wenn sich jeder Vektor $\mathbf{x} \in \mathbf{U}_1 + \dots + \mathbf{U}_n$ auf eindeutige Weise als Summe $\mathbf{x} = \mathbf{u}_1 + \dots + \mathbf{u}_n$ mit $\mathbf{u}_i \in \mathbf{U}_i$ schreiben läßt. Wir schreiben dann $\mathbf{U} = \mathbf{U}_1 \oplus \dots \oplus \mathbf{U}_n$.

Proposition 23.17 (Direkte Summe zweier Unterräume)

Es seien \mathbf{U} , \mathbf{U}' und \mathbf{W} Unterräume des K -Vektorraums V .

Genau dann gilt $\mathbf{W} = \mathbf{U} \oplus \mathbf{U}'$, wenn $\mathbf{W} = \mathbf{U} + \mathbf{U}'$ und $\mathbf{U} \cap \mathbf{U}' = \{0\}$.

Beweis: Ist die Summe $\mathbf{W} = \mathbf{U} \oplus \mathbf{U}'$, so gilt insbesondere $\mathbf{W} = \mathbf{U} + \mathbf{U}'$. Für $\mathbf{x} \in \mathbf{U} \cap \mathbf{U}'$, gilt zudem

$$\mathbf{x} = \mathbf{x} + 0 = 0 + \mathbf{x} \in \mathbf{U} + \mathbf{U}',$$

und wegen der Eindeutigkeit der Darstellung in $\mathbf{U} + \mathbf{U}'$ muß $\mathbf{x} = 0$ sein.

Ist umgekehrt $W = U + U'$ und $U \cap U' = \{0\}$ und sind $x_1 + x'_1 = x_2 + x'_2 \in U + U' = W$ mit $x_i \in U$ und $x'_i \in U'$, $i = 1, 2$, so gilt:

$$x_1 - x_2 = x'_2 - x'_1 \in U \cap U' = \{0\}.$$

Also ist $x_1 = x_2$ und $x'_1 = x'_2$, d. h. die Darstellung ist eindeutig. \square

Beispiel 23.18

Betrachte die Unterräume $U = \text{Lin}((1, 1, 1)^t)$ und $U' = \text{Lin}((1, 0, 1)^t)$ von \mathbb{R}^3 . Ein Vektor x liegt genau dann im Durchschnitt $U \cap U'$, wenn es $\lambda, \mu \in K$ gibt mit

$$x = \lambda \cdot (1, 1, 1)^t = (\lambda, \lambda, \lambda)^t$$

und

$$x = \mu \cdot (1, 0, 1)^t = (\mu, 0, \mu)^t.$$

Gleichsetzen der beiden Ausdrücke liefert die Bedingungen $\lambda = \mu$ und $\lambda = 0$, also gilt $x = (0, 0, 0)^t$, d.h.

$$U \cap U' = \{(0, 0, 0)^t\}.$$

Damit ist die Summe $U + U'$ eine direkte Summe.

C) Lineare Abbildungen

Zu jeder Struktur gehören die strukturerhaltenden Abbildungen.

Definition 23.19 (Lineare Abbildungen)

Es seien V und W zwei K -Vektorräume.

- a. Eine Abbildung $f : V \rightarrow W$ heißt *K-lineare Abbildung* oder *Vektorraumhomomorphismus*, wenn für alle $\lambda \in K$ und $x, y \in V$ gilt

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

- b. Eine injektive (bzw. surjektive bzw. bijektive) K -lineare Abbildung heißt auch *Monomorphismus* (bzw. *Epimorphismus* bzw. *Isomorphismus*). Gilt $V = W$, so nennen wir eine K -lineare Abbildung auch einen *Endomorphismus*, und ist sie zudem bijektiv, so sprechen wir von einem *Automorphismus*.
- c. Existiert ein Isomorphismus von V nach W , so nennen wir V und W *isomorph* und schreiben $V \cong W$.
- d. Die Menge aller K -linearen Abbildungen von V nach W bezeichnen wir mit $\text{Hom}_K(V, W)$ und die Menge aller Endomorphismen von V mit $\text{End}_K(V)$.

Bemerkung 23.20

Die beiden Bedingungen in Definition 23.19 a. lassen sich zusammenfassen zu der Bedingung $f(\lambda x + \mu y) = \lambda f(x) + \mu f(y)$ für alle $\lambda, \mu \in K$ und $x, y \in V$.

Beispiel 23.21 a. Die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R} : (x_1, x_2)^t \mapsto x_1 - x_2$ ist \mathbb{R} -linear.

Denn für $x = (x_1, x_2)$ und $y = (y_1, y_2)$ sowie $\lambda \in \mathbb{R}$ gilt

$$\begin{aligned} f(x + y) &= f((x_1 + y_1, x_2 + y_2)) = x_1 + y_1 - x_2 - y_2 \\ &= x_1 - x_2 + y_1 - y_2 = f(x) + f(y) \end{aligned}$$

und

$$f(\lambda x) = f((\lambda x_1, \lambda x_2)) = \lambda x_1 - \lambda x_2 = \lambda \cdot (x_1 - x_2) = \lambda \cdot f(x).$$

b. Ist I ein Intervall, so ist die Abbildung

$$D : \mathcal{C}^1(I, \mathbb{R}) \rightarrow \mathcal{C}(I, \mathbb{R}) : f \mapsto f'$$

\mathbb{R} -linear, da aus der Linearität der Ableitung 17.9 folgt

$$D(\lambda f + \mu g) = (\lambda f + \mu g)' = \lambda \cdot f' + \mu \cdot g' = \lambda \cdot D(f) + \mu \cdot D(g).$$

c. Die Abbildung

$$f : \mathbb{R}[t] \rightarrow \mathbb{R}^{\mathbb{N}} : \sum_{k=0}^n a_k \cdot t^k \mapsto (a_k)_{k \in \mathbb{N}},$$

wobei $a_k = 0$ für $k > n$ gelten soll, ist eine \mathbb{R} -lineare Abbildung. Ihr Bild ist der Unterraum

$$\text{Im}(f) = \{(a_k)_{k \in \mathbb{N}} \mid \text{nur endlich viele } a_k \text{ sind nicht } 0\}$$

der *abbrechenden Folgen* in \mathbb{R} .

d. Die formale Ableitung

$$d : K[t] \rightarrow K[t] : \sum_{k=0}^n a_k \cdot t^k \mapsto \sum_{k=0}^n k \cdot a_k \cdot t^{k-1}$$

ist eine K -lineare Abbildung, wie man leicht nachrechnet.

Lemma 23.22 (Einfache Eigenschaften linearer Abbildungen)

Seien U, V und W K -Vektorräume und $f : U \rightarrow V$ und $g : V \rightarrow W$ seien K -linear.

Ferner seien $x, x_1, \dots, x_n \in U$ und $\lambda_1, \dots, \lambda_n \in K$. Dann gelten:

- $f(0_U) = 0_V$ und $f(-x) = -f(x)$.
- $f(\lambda_1 x_1 + \dots + \lambda_n x_n) = \lambda_1 f(x_1) + \dots + \lambda_n f(x_n)$.
- Ist f bijektiv, so ist $f^{-1} : V \rightarrow U$ K -linear.
- $g \circ f : U \rightarrow W$ ist K -linear.
- $\text{Hom}_K(U, V)$ ist ein Unterraum von V^U .

Beweis: a. Aus der Verträglichkeit mit der Skalarmultiplikation folgen

$$f(0_U) = f(0_K \cdot 0_U) = 0_K \cdot f(0_U) = 0_V$$

und

$$f(-x) = f((-1) \cdot x) = (-1) \cdot f(x) = -f(x).$$

- b. Die Aussage folgt mittels Induktion aus den beiden Bedingungen für Linearität.
- c. Seien $\mathbf{y}, \mathbf{y}' \in V$ und $\lambda, \lambda' \in K$ sowie $\mathbf{x} = f^{-1}(\mathbf{y})$ und $\mathbf{x}' = f^{-1}(\mathbf{y}')$. Wegen der Linearität von f gilt

$$f(\lambda \cdot \mathbf{x} + \lambda' \cdot \mathbf{x}') = \lambda \cdot f(\mathbf{x}) + \lambda' \cdot f(\mathbf{x}').$$

Wenden wir auf beiden Seiten f^{-1} an, so erhalten wir

$$\begin{aligned} \lambda \cdot f^{-1}(\mathbf{y}) + \lambda' \cdot f^{-1}(\mathbf{y}') &= \lambda \cdot \mathbf{x} + \lambda' \cdot \mathbf{x}' = f^{-1}(f(\lambda \cdot \mathbf{x} + \lambda' \cdot \mathbf{x}')) \\ &= f^{-1}(\lambda \cdot f(\mathbf{x}) + \lambda' \cdot f(\mathbf{x}')) = f^{-1}(\lambda \cdot \mathbf{y} + \lambda' \cdot \mathbf{y}'). \end{aligned}$$

Mithin ist f^{-1} eine lineare Abbildung.

- d. Seien $\lambda, \mu \in K$ und $\mathbf{x}, \mathbf{y} \in U$, so gelten

$$\begin{aligned} (g \circ f)(\lambda \mathbf{x} + \mu \mathbf{y}) &= g(f(\lambda \mathbf{x} + \mu \mathbf{y})) = g(\lambda f(\mathbf{x}) + \mu f(\mathbf{y})) \\ &= \lambda g(f(\mathbf{x})) + \mu g(f(\mathbf{y})) = \lambda(g \circ f)(\mathbf{x}) + \mu(g \circ f)(\mathbf{y}). \end{aligned}$$

- e. Dies folgt aus Aufgabe 23.44. □

Proposition 23.23 (f_A ist linear.)

Für $A \in \text{Mat}(m \times n, K)$ ist $f_A : K^n \rightarrow K^m$ eine K -lineare Abbildung.

Beweis: Aus Lemma 22.8 folgt für $\mathbf{x}, \mathbf{y} \in K^n$ und $\lambda \in K$

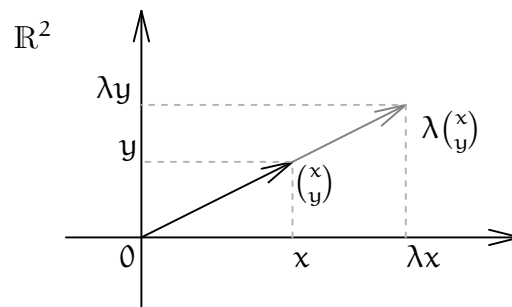
$$f_A(\mathbf{x} + \mathbf{y}) = A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y} = f_A(\mathbf{x}) + f_A(\mathbf{y})$$

und

$$f_A(\lambda \mathbf{x}) = A(\lambda \mathbf{x}) = \lambda(A\mathbf{x}) = \lambda f_A(\mathbf{x}).$$
□

Beispiel 23.24

- a. Im Fall $n = 1$ und $A = (\mathbf{a})$ ist die K -lineare Abbildung $f_A : K \rightarrow K : x \mapsto \mathbf{a} \cdot x$ gerade die Multiplikation mit \mathbf{a} .
- b. Die lineare Abbildung $f_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2 : (x, y)^t \mapsto (\lambda x, \lambda y)$ zu $A = \lambda \mathbb{1}_2$ ist eine *Streckung* um den Faktor λ .



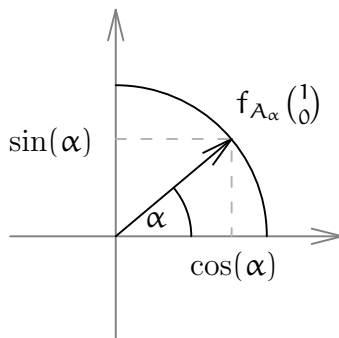
c. Für $\alpha \in \mathbb{R}$ setzen wir

$$A_\alpha := \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Dann ist die lineare Abbildung $f_{A_\alpha} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ eine Drehung um den Winkel α . Beachte dazu, daß

$$A_\alpha e_1 = (\cos(\alpha), \sin(\alpha))^t \quad \text{und} \quad A_\alpha e_2 = (-\sin(\alpha), \cos(\alpha))^t,$$

woraus die Aussage für die *Einheitsvektoren* e_1 und e_2 unmittelbar folgt.



Daraus leitet sich die Aussage für einen beliebigen Vektor $(x, y)^t$ mittels der Linearität von f_{A_α} ab: $f_{A_\alpha}((x, y)^t) = x f_{A_\alpha}(e_1) + y f_{A_\alpha}(e_2)$.

d. Ist $n \geq m$, so ist die Abbildung

$$\text{pr} : \mathbb{K}^n \rightarrow \mathbb{K}^m : (x_1, \dots, x_n)^t \mapsto (x_1, \dots, x_m)^t$$

eine \mathbb{K} -lineare Abbildung, genannt die kanonische *Projektion*.

Ist $m \geq n$, dann ist die kanonische *Inklusion*

$$i_{\mathbb{K}^n} : \mathbb{K}^n \rightarrow \mathbb{K}^m : (x_1, \dots, x_n)^t \mapsto (x_1, \dots, x_n, 0, \dots, 0)^t$$

ebenfalls \mathbb{K} -linear. Beides prüft man leicht nach.

Proposition 23.25 (Kern und Bild sind Unterräume)

Es seien V und W \mathbb{K} -Vektorräume und $f : V \rightarrow W$ sei \mathbb{K} -linear.

- Ist U ein Unterraum von V , so ist $f(U)$ ein Unterraum von W .
- Ist U ein Unterraum von W , so ist $f^{-1}(U)$ ein Unterraum von V .
- Das Bild $\text{Im}(f) = f(V)$ von f ist ein Unterraum von W .
- Der Kern von f , $\text{Ker}(f) = \{x \in V \mid f(x) = 0\}$, ist ein Unterraum von V .

Beweis:

- Es sei U ein Unterraum von V . Dann ist $0_V \in U$ und somit $0_W = f(0_V) \in f(U)$, so daß $f(U)$ nicht leer ist. Sind $\lambda \in \mathbb{K}$ und $u = f(x), v = f(y) \in f(U)$ mit $x, y \in U$, so gilt

$$u + v = f(x) + f(y) = f(x + y) \in f(U)$$

und

$$\lambda \mathbf{u} = \lambda f(\mathbf{x}) = f(\lambda \mathbf{x}) \in f(\mathbf{U}).$$

Also ist $f(\mathbf{U})$ ein Unterraum von \mathbf{W} .

- b. Es sei \mathbf{U} ein Unterraum von \mathbf{W} . Dann ist $0_{\mathbf{W}} \in \mathbf{U}$ und wegen $f(0_{\mathbf{V}}) = 0_{\mathbf{W}}$ ist dann $0_{\mathbf{V}} \in f^{-1}(\mathbf{U})$, so daß $f^{-1}(\mathbf{U})$ nicht leer ist. Sind $\lambda \in \mathbf{K}$ und $\mathbf{x}, \mathbf{y} \in f^{-1}(\mathbf{U})$, so gilt $f(\mathbf{x}), f(\mathbf{y}) \in \mathbf{U}$ und somit

$$f(\mathbf{x} + \mathbf{y}) = f(\mathbf{x}) + f(\mathbf{y}) \in \mathbf{U}$$

und

$$f(\lambda \mathbf{x}) = \lambda f(\mathbf{x}) \in \mathbf{U}.$$

Also auch $\mathbf{x} + \mathbf{y} \in f^{-1}(\mathbf{U})$ und $\lambda \mathbf{x} \in f^{-1}(\mathbf{U})$, so daß $f^{-1}(\mathbf{U})$ ein Unterraum von \mathbf{V} ist.

- c. Dies folgt aus a. mit $\mathbf{U} = \mathbf{V}$.
 d. Dies folgt aus b. mit $\mathbf{U} = \{0_{\mathbf{W}}\}$.

□

Beispiel 23.26

Die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R} : (\mathbf{x}_1, \mathbf{x}_2) \mapsto \mathbf{x}_1 - \mathbf{x}_2$ aus Beispiel 23.21 hat den Kern

$$\text{Ker}(f) = \{\mathbf{x} \in \mathbb{R}^2 \mid f(\mathbf{x}) = \mathbf{x}_1 - \mathbf{x}_2 = 0\} = \{(\lambda, \lambda)^t \mid \lambda \in \mathbf{K}\}$$

und das Bild ist $\text{Im}(f) = \mathbb{R}$.

Proposition 23.27 (Injektivität linearer Abbildungen)

Eine lineare Abbildung $f : \mathbf{V} \rightarrow \mathbf{W}$ ist genau dann injektiv, wenn $\text{Ker}(f) = \{0_{\mathbf{V}}\}$.

Beweis: Ist f injektiv und $\mathbf{x} \in \text{Ker}(f)$, so folgt aus $f(\mathbf{x}) = 0_{\mathbf{W}} = f(0_{\mathbf{V}})$ auch $\mathbf{x} = 0_{\mathbf{V}}$. Also ist $\text{Ker}(f) = \{0_{\mathbf{V}}\}$.

Sei nun umgekehrt $\text{Ker}(f) = \{0_{\mathbf{V}}\}$ und seien $\mathbf{x}, \mathbf{y} \in \mathbf{V}$ mit $f(\mathbf{x}) = f(\mathbf{y})$. Dann folgt $f(\mathbf{x} - \mathbf{y}) = f(\mathbf{x}) - f(\mathbf{y}) = 0_{\mathbf{W}}$ und damit $\mathbf{x} - \mathbf{y} \in \text{Ker}(f) = \{0_{\mathbf{V}}\}$. Also ist $\mathbf{x} = \mathbf{y}$, und somit ist f injektiv. □

D) Faktorräume

Definition 23.28 (Faktorraum)

Es sei \mathbf{V} ein \mathbf{K} -Vektorraum und \mathbf{U} ein Unterraum von \mathbf{V} .

- a. Für $\mathbf{x} \in \mathbf{V}$ nennen wir

$$\bar{\mathbf{x}} := \mathbf{x} + \mathbf{U} := \{\mathbf{x} + \mathbf{u} \mid \mathbf{u} \in \mathbf{U}\}$$

die *Restklasse* oder *Nebenklasse* von \mathbf{x} modulo \mathbf{U} und \mathbf{x} einen *Vertreter* der Restklasse. Man nennt $\mathbf{x} + \mathbf{U}$ auch einen *affinen Raum* parallel zum Unterraum \mathbf{U} mit Aufpunkt \mathbf{x} .

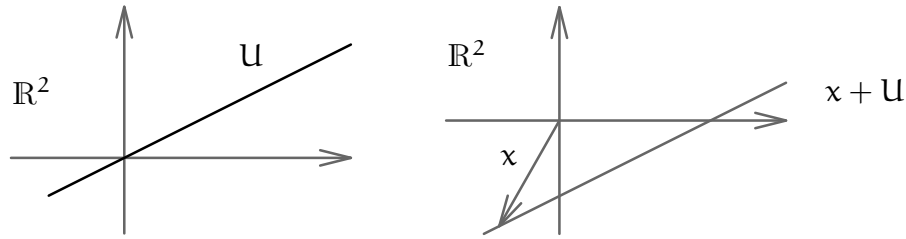


ABBILDUNG 2. Ein affiner Raum $x + \mathbf{U}$ zum Unterraum \mathbf{U} .

Man beachte, daß aus der Notation \bar{x} nicht mehr abzulesen ist, modulo welchem Unterraum man rechnet. Die Notation ist deshalb mit Vorsicht zu verwenden.

- b. Wir nennen die Menge der Restklassen modulo \mathbf{U}

$$V/\mathbf{U} := \{x + \mathbf{U} \mid x \in V\} = \{\bar{x} \mid x \in V\}$$

auch den *Faktorraum* von V modulo \mathbf{U} .

Bemerkung 23.29 (Restklassen als Äquivalenzklassen)

Ist \mathbf{U} ein Unterraum des K -Vektorraums V , so wird durch

$$x \sim y :\iff x - y \in \mathbf{U}$$

für $x, y \in V$ eine *Äquivalenzrelation* auf V definiert, wie man leicht nachprüft. Die Äquivalenzklasse von x ist dann gerade $x + \mathbf{U}$, und V/\mathbf{U} ist die Menge aller Äquivalenzklassen der Äquivalenzrelation. Dies ist konzeptionell der richtige Weg, die Restklassen von V modulo \mathbf{U} sowie den Faktorraum einzuführen.

Lemma 23.30 (Rechnen mit Restklassen)

Es sei V ein K -Vektorraum, \mathbf{U} ein Unterraum, $x, x', y, y' \in V$ und $\lambda \in K$. In V/\mathbf{U} gelten dann die folgenden Aussagen:

- a. Entweder $\bar{x} = \bar{y}$ oder $\bar{x} \cap \bar{y} = \emptyset$.
 b. Es gilt:

$$\bar{x} = \bar{y} \iff x - y \in \mathbf{U}.$$

Insbesondere, $\bar{x} = \bar{0} = \mathbf{U}$ genau dann, wenn $x \in \mathbf{U}$.

- c. Gilt $\bar{x} = \overline{x'}$ und $\bar{y} = \overline{y'}$, so gelten auch

$$\overline{x + y} = \overline{x' + y'} \quad \text{und} \quad \overline{\lambda x} = \overline{\lambda x'}.$$

Beweis: Im Beweis wollen wir $x + \mathbf{U}$ und $y + \mathbf{U}$ statt \bar{x} und \bar{y} schreiben.

- a. Ist der Durchschnitt nicht leer, so finden wir $u, v \in \mathbf{U}$ mit

$$x + u = y + v \in (x + \mathbf{U}) \cap (y + \mathbf{U}).$$

Ist nun $z = x + w \in x + \mathbf{U}$ beliebig mit $w \in \mathbf{U}$, so ist

$$z = x + w = y + (v - u + w) \in y + \mathbf{U}.$$

Das zeigt, daß $x + \mathcal{U} \subseteq y + \mathcal{U}$. Die umgekehrte Inklusion zeigt man analog. Also gilt $x + \mathcal{U} = y + \mathcal{U}$, sobald der Durchschnitt der beiden nicht leer ist.

- b. Ist $x + \mathcal{U} = y + \mathcal{U}$, so finden wir wieder $u, v \in \mathcal{U}$ mit $x + u = y + v$ und somit gilt $x - y = v - u \in \mathcal{U}$. Ist umgekehrt $x - y \in \mathcal{U}$, so ist

$$y + (x - y) = x + 0 \in (y + \mathcal{U}) \cap (x + \mathcal{U}),$$

und mit Teil a. sind die beiden Restklassen also gleich.

- c. Aus $x + \mathcal{U} = x' + \mathcal{U}$ sowie $y + \mathcal{U} = y' + \mathcal{U}$ folgt nach b.

$$x - x', y - y' \in \mathcal{U}.$$

Damit gilt dann auch

$$(x + y) - (x' + y') = (x - x') + (y - y') \in \mathcal{U}$$

und

$$(\lambda x - \lambda x') = \lambda \cdot (x - x') \in \mathcal{U}.$$

Wegen b. gilt dann wieder $(x + y) + \mathcal{U} = (x' + y') + \mathcal{U}$ und $\lambda x + \mathcal{U} = \lambda x' + \mathcal{U}$.

□

Beispiel 23.31

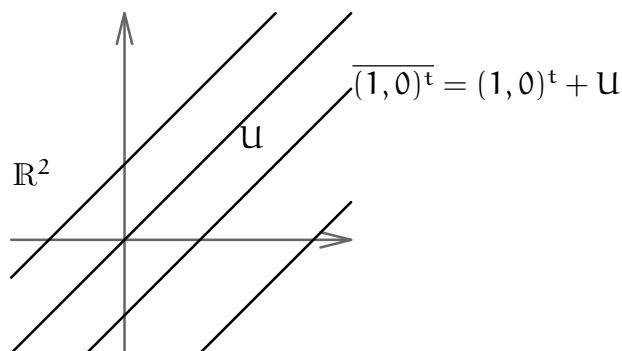
Sei $V = \mathbb{R}^2$ und $\mathcal{U} = \text{Lin}((1, 1)^t)$. Dann gilt

$$\overline{(1, 0)^t} = (1, 0)^t + \mathcal{U} = (2, 1)^t + \mathcal{U} = \overline{(2, 1)^t}$$

und

$$\overline{(1, 0)^t} = (1, 0)^t + \mathcal{U} \neq (0, 1)^t + \mathcal{U} = \overline{(0, 1)^t}.$$

Die Restklassen in V/\mathcal{U} sind in diesem Fall genau die Geraden, die parallel zur Geraden \mathcal{U} sind.



Satz 23.32 (Der Faktorraum ist ein Vektorraum.)

Es sei V ein K -Vektorraum und \mathcal{U} ein Unterraum. Dann definiert

$$\bar{x} + \bar{y} := \overline{x + y} \tag{63}$$

und

$$\lambda \cdot \bar{x} := \overline{\lambda x} \tag{64}$$

für $\bar{x}, \bar{y} \in V/U$ und $\lambda \in K$ eine Addition und eine Skalarmultiplikation auf V/U bezüglich derer der Faktorraum V/U ein K -Vektorraum ist.

Zudem ist die Abbildung

$$\pi: V \longrightarrow V/U: x \mapsto \bar{x}$$

eine surjektive K -lineare Abbildung mit $\text{Ker}(\pi) = U$, die wir die Restklassenabbildung nennen.

Beweis: Bevor wir uns den Vektorraumaxiomen widmen können, müssen wir zeigen, daß durch (63) und (64) überhaupt Operationen auf V/U definiert werden. Das Problem dabei ist, daß wir zur Definition der Summe und der Skalarmultiplikation Vertreter der Restklassen verwendet haben. Diese sind aber nicht eindeutig bestimmt. Wir müssen also sicherstellen, daß wir das gleiche Ergebnis erhalten, wenn wir andere Vertreter wählen. Man nennt dies die *Wohldefiniertheit* der Operationen. Dazu ist zu zeigen, daß aus für $\bar{x} = \bar{x}'$ und $\bar{y} = \bar{y}'$ auch $\overline{\bar{x} + \bar{y}} = \overline{\bar{x}' + \bar{y}'}$ und $\overline{\lambda \bar{x}} = \overline{\lambda \bar{x}'}$ gelten. Dies folgt aber aus Lemma 23.30.

Wir wollen nun zeigen, daß V/U den Vektorraumaxiomen genügt. Dazu seien $\bar{x}, \bar{y}, \bar{z} \in V/U$ und $\lambda, \mu \in K$ gegeben. Dann gilt

$$(\bar{x} + \bar{y}) + \bar{z} = \overline{\bar{x} + \bar{y}} + \bar{z} = \overline{(\bar{x} + \bar{y}) + \bar{z}} = \overline{\bar{x} + (\bar{y} + \bar{z})} = \bar{x} + \overline{\bar{y} + \bar{z}} = \bar{x} + (\bar{y} + \bar{z}).$$

Die Addition ist also assoziativ. Zudem gilt

$$\bar{0} + \bar{x} = \overline{\bar{0} + \bar{x}} = \bar{x}$$

und

$$\overline{-\bar{x}} + \bar{x} = \overline{-\bar{x} + \bar{x}} = \bar{0},$$

so daß $\bar{0}$ der Nullvektor ist und \bar{x} ein Inverses besitzt. Da zudem

$$\bar{x} + \bar{y} = \overline{\bar{x} + \bar{y}} = \overline{\bar{y} + \bar{x}} = \bar{y} + \bar{x}$$

gilt, ist V/U eine abelsche Gruppe bezüglich $+$. Ähnlich wie sich die Axiome für die Addition von V auf V/U vererbt haben, vererben sich auch die Axiome für die Skalarmultiplikation.

$$(\lambda + \mu) \cdot \bar{x} = \overline{(\lambda + \mu) \cdot \bar{x}} = \overline{\lambda \bar{x} + \mu \bar{x}} = \overline{\lambda \bar{x}} + \overline{\mu \bar{x}} = \lambda \cdot \bar{x} + \mu \cdot \bar{x}$$

und

$$\lambda \cdot \overline{\bar{x} + \bar{y}} = \overline{\lambda \cdot (\bar{x} + \bar{y})} = \overline{\lambda \bar{x} + \lambda \bar{y}} = \overline{\lambda \bar{x}} + \overline{\lambda \bar{y}} = \lambda \cdot \bar{x} + \lambda \cdot \bar{y}$$

und

$$(\lambda \cdot \mu) \cdot \bar{x} = \overline{(\lambda \cdot \mu) \cdot \bar{x}} = \overline{\lambda \cdot (\mu \cdot \bar{x})} = \overline{\lambda \mu \cdot \bar{x}} = \lambda \cdot (\mu \cdot \bar{x})$$

und

$$1 \cdot \bar{x} = \overline{1 \cdot \bar{x}} = \bar{x}.$$

Also ist V/U ein K -Vektorraum.

Es bleibt, die Aussagen zur Restklassenabbildung π zu zeigen. Die Linearität von π folgt unmittelbar aus der Definition der Operationen auf V/U . Sind $x, y \in V$ und $\lambda \in K$, dann gilt

$$\pi(x + y) = \overline{x + y} = \bar{x} + \bar{y} = \pi(x) + \pi(y)$$

und

$$\pi(\lambda x) = \overline{\lambda x} = \lambda \cdot \bar{x} = \lambda \cdot \pi(x).$$

Außerdem ist π surjektiv, da jedes $\bar{x} \in V/U$ sich schreiben läßt als $\bar{x} = \pi(x)$. Und es gilt

$$x \in \text{Ker}(\pi) \iff \bar{x} = \pi(x) = \bar{0} \iff x \in U.$$

Damit sind alle Aussagen des Satzes bewiesen. □

Bemerkung 23.33 (Die vier Rechenregeln für den Faktorraum)

Um mit dem Faktorraum rechnen zu können, braucht man nur die Rechenregeln:

- a. $\bar{0} = 0 + U = U$ ist der Nullvektor.
- b. $\bar{x} + \bar{y} = \overline{x + y}$.
- c. $\lambda \cdot \bar{x} = \overline{\lambda x}$.
- d. $\bar{x} = \bar{y} \iff x - y \in U$.

Satz 23.34 (Homomorphiesatz)

Ist $f : V \rightarrow W$ eine K -lineare Abbildung, so ist

$$\bar{f} : V/\text{Ker}(f) \rightarrow \text{Im}(f) : \bar{x} \mapsto f(x)$$

ein Isomorphismus. Insbesondere gilt also $V/\text{Ker}(f) \cong \text{Im}(f)$.

Beweis: Da wir für die Definition von $\bar{f}(\bar{x})$ wieder den Restklassenvertreter x verwendet haben, müssen wir wieder zeigen, daß unsere Definition nicht von der speziellen Wahl des Vertreters abhängt. Man sagt wieder, wir müssen die *Wohldefiniertheit* von \bar{f} zeigen.

Seien dazu $\bar{x} = x + \text{Ker}(f) = y + \text{Ker}(f) = \bar{y}$ gegeben. Dann gilt

$$x - y \in \text{Ker}(f),$$

und mithin $0 = f(x - y) = f(x) - f(y)$, oder alternativ $f(x) = f(y)$. Die Abbildung \bar{f} ist also wohldefiniert.

Die Linearität von \bar{f} folgt dann aus der Linearität von f . Seien dazu $\bar{x}, \bar{y} \in V/\text{Ker}(f)$ und $\lambda \in K$, dann gilt

$$\bar{f}(\bar{x} + \bar{y}) = \bar{f}(\overline{x + y}) = f(x + y) = f(x) + f(y) = \bar{f}(\bar{x}) + \bar{f}(\bar{y})$$

und

$$\bar{f}(\lambda \cdot \bar{x}) = \bar{f}(\overline{\lambda \cdot x}) = f(\lambda \cdot x) = \lambda \cdot f(x) = \lambda \cdot \bar{f}(\bar{x}).$$

Es bleibt noch zu zeigen, daß \bar{f} surjektiv und injektiv ist.

Ist $\mathbf{y} \in \text{Im}(f)$, so gibt es ein $\mathbf{x} \in V$ mit $f(\mathbf{x}) = \mathbf{y}$, und somit gilt

$$\mathbf{y} = f(\mathbf{x}) = \bar{f}(\bar{\mathbf{x}}).$$

Also ist \bar{f} surjektiv.

Für die Injektivität nutzen wir Proposition 23.27. Es gilt

$$\bar{\mathbf{x}} \in \text{Ker}(\bar{f}) \iff 0 = \bar{f}(\bar{\mathbf{x}}) = f(\mathbf{x}) \iff \mathbf{x} \in \text{Ker}(f) \iff \bar{\mathbf{x}} = \bar{0}.$$

Also enthält der Kern von \bar{f} nur den Nullvektor, und somit ist \bar{f} injektiv. \square

Definition 23.35 (Direkte Komplemente)

Es sei V ein K -Vektorraum und U und U' seien Unterräume von V . Dann heißt U' ein (direktes) *Komplement* von U , falls $V = U \oplus U'$.

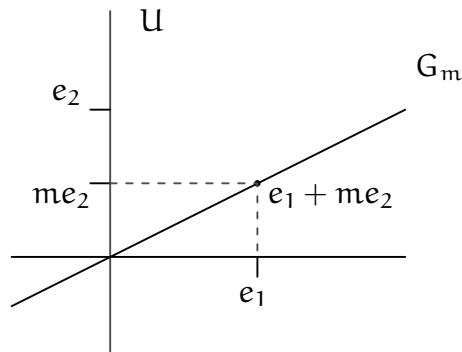
Beispiel 23.36 (Komplemente sind nicht eindeutig.)

Ist $V = \mathbb{R}^2$ und $U = \text{Lin}(e_2)$ die y -Achse, dann ist die Ursprungsgerade mit Steigung m

$$G_m := \text{Lin}(e_1 + me_2)$$

für jedes $m \in \mathbb{R}$ ein Komplement von U . Beachte dazu nur, daß sich die Geraden U und G_m nur im Ursprung schneiden, d.h. $U \cap G_m = \{0\}$, und daß ein beliebiger Vektor $(x, y)^t \in \mathbb{R}^2$ sich schreiben läßt als

$$(x, y)^t = x \cdot (e_1 + me_2) + (-xm + y) \cdot e_2 \in G_m + U.$$



Proposition 23.37 (Der Faktorraum als Komplementersatz)

Sei V ein K -Vektorraum, U ein Unterraum von V und U' ein Komplement von U . Dann ist die Einschränkung der Restklassenabbildung

$$\pi|_{U'} : U' \rightarrow V/U : x \mapsto \bar{x}$$

auf U' ein Isomorphismus. Insbesondere sind je zwei Komplemente von U isomorph.

Beweis: Es ist klar, daß $\pi|_{U'}$ als Einschränkung einer K -linearen Abbildung wieder K -linear ist.

Wir zeigen zunächst, daß $\pi|_{U'}$ surjektiv ist. Sei dazu $\bar{x} \in V/U$ gegeben. Wegen $V = U \oplus U'$ läßt sich x als $x = y + z$ mit $y \in U$ und $z \in U'$ schreiben. Damit gilt:

$$\bar{x} = \bar{z} = \pi|_{U'}(z) \in \text{Im}(\pi|_{U'}).$$

Also ist $\pi|$ surjektiv.

Es bleibt zu zeigen, daß $\pi|$ injektiv ist, d. h. $\text{Ker}(\pi|) = \{0\}$. Sei dazu $z \in \text{Ker}(\pi|)$, dann gilt

$$\bar{0} = \pi|(z) = \bar{z}.$$

D. h. $z \in \mathbf{U}$. Damit gilt aber $z \in \mathbf{U} \cap \mathbf{U}' = \{0\}$, also $z = 0$.

Seien schließlich \mathbf{U}' und \mathbf{U}'' zwei Komplemente von \mathbf{U} , dann ist die Komposition

$$\mathbf{U}' \xrightarrow{\pi|_{\mathbf{U}'}} \mathbf{V}/\mathbf{U} \xrightarrow{\pi|_{\mathbf{U}''}^{-1}} \mathbf{U}''$$

ein Isomorphismus von \mathbf{U}' nach \mathbf{U}'' . Die beiden Komplemente sind also isomorph zueinander. \square

Bemerkung 23.38

Daß \mathbf{V}/\mathbf{U} isomorph zu jedem Komplement von \mathbf{U} ist, heißt im wesentlichen, daß man bei Betrachtungen, bei denen man ein Komplement von \mathbf{U} benötigt, stattdessen auch mit \mathbf{V}/\mathbf{U} arbeiten kann. Während es sehr viele Komplemente von \mathbf{U} geben kann, gibt es nur *einen* Faktorraum. Dieser ist durch \mathbf{U} eindeutig bestimmt. Das ist unter Umständen ein großer Vorteil!

E) Ringe und Moduln

Bemerkung 23.39 (Kommutative Ringe mit Eins)

Eine Menge \mathbf{K} mit zwei zweistelligen Operationen $+$ und \cdot , die allen Axiomen eines Körpers genügt, außer eventuell der Existenz von multiplikativen Inversen, nennt man einen *kommutativen Ring mit Eins*.

Ein typisches Beispiel dafür sind die ganzen Zahlen \mathbb{Z} , und in der Vorlesung algebraische Strukturen lernt man den Polynomring, z.B.

$$\mathbb{R}[t] = \left\{ \sum_{k=0}^n a_k \cdot t^k \mid a_k \in \mathbb{R}, n \in \mathbb{N} \right\},$$

als ein weiteres typisches Beispiel kennen.

Der Nachteil von Ringen im Vergleich zu Körpern ist, daß man im allgemeinen nicht durch die Elemente teilen darf. Deshalb sind nicht alle Aussagen, die für Körper gelten auch gleichermaßen für Ringe richtig. Aber immer dann, wenn man ohne Division auskommt, geht alles gut. Das trifft auf die meisten Definitionen und Aussagen in diesem Kapitel zu!

Bemerkung 23.40 (Matrizen über Ringen)

Setzen wir in Abschnitt 22 nur voraus, daß \mathbf{K} ein kommutativer Ring mit Eins ist, so bleiben alle Definitionen und Aussagen korrekt.

Bemerkung 23.41 (Moduln und lineare Abbildungen)

Setzen wir in Definition 23.1 nur voraus, daß \mathbf{K} ein kommutativer Ring mit Eins ist, so nennen wir \mathbf{V} einen *K-Modul*. Der Name ist dabei aber auch das einzige, was sich

ändert. Entsprechend wird man in Definition 23.4 dann von einem *Untermodul* statt von einem Unterraum reden. Alle anderen Begriffe, Beispiele und Aussagen dieses Abschnitts bleiben ohne Änderung und mit dem jeweils gleichen Beweis korrekt, bis auf eine einzige Ausnahme:

$$\lambda \cdot v = 0 \not\Rightarrow \lambda = 0 \text{ oder } v = 0.$$

Im Beweis dieser Aussage mußten wir durch λ teilen können, wenn λ nicht Null war!

Wir werden an einigen Stellen der Vorlesung lineare Abbildungen oder Matrizen über dem Polynomring benötigen. Deshalb werde ich immer mal wieder anmerken, welche Aussagen auch für Ringe und Moduln gelten. Wenn ohnehin vieles ohne Änderung im Beweis korrekt bleibt, hätte man die Aussagen natürlich auch gleich für Ringe und Moduln formulieren können. Die Erfahrung zeigt aber, daß die Mehrzahl der Studenten sich mit Körpern und Vektorräumen als Begriffen wohler fühlt.

Aufgaben

Aufgabe 23.42

Welche der folgenden Teilmengen von K^3 sind Unterräume des K^3 ? Begründe Deine Antworten.

- $\{(x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid x_1 \cdot x_2 = 2x_3\}$ für $K = \mathbb{R}$.
- $\{(x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid ax_1 + x_2 + x_3 = a + 1\}$ für ein festes $a \in \mathbb{R}$ für $K = \mathbb{R}$.
- $\{(x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid x_1 \leq 0\}$ für $K = \mathbb{R}$.
- $\{(1, 0, 0)^t, (0, 1, 0)^t, (1, 1, 0)^t, (0, 0, 0)^t\}$ für $K = \mathbb{R}$ oder $K = \mathbb{F}_2$.

Aufgabe 23.43

Gegeben seien die folgenden Teilmengen des \mathbb{Q} -Vektorraums $V = \mathbb{Q}^4$:

- $U_1 = \{(x, x + 1, x + 2, x + 4)^t \mid x \in \mathbb{Q}\}$,
- $U_2 = \{(x, 2x, 3x, 4x)^t \mid x \in \mathbb{Q}\}$,
- $U_3 = \{(x_1, x_2, x_3, x_4)^t \mid x_1, x_2, x_3, x_4 \in \mathbb{Q}, x_3 = x_1 + x_2, x_4 = x_2 + x_3\}$,
- $U_4 = \{(x_1, x_2, x_3, x_4)^t \mid x_1, x_2, x_3, x_4 \in \mathbb{Q}, x_2 \geq 0\}$,
- $U_5 = \{(x_1, x_2, x_3, x_4)^t \mid x_1, x_2, x_3, x_4 \in \mathbb{Z}\}$.

Welche dieser Mengen sind Unterräume von V ? Begründe Deine Aussage.

Aufgabe 23.44

Seien U, V und W K -Vektorräume, $\lambda, \lambda' \in K$ und $f, f' \in \text{Hom}_K(U, V)$ und $g, g' \in \text{Hom}_K(V, W)$. Dann gelten:

- $f + f', \lambda \cdot f \in \text{Hom}_K(U, V)$, d.h. $\text{Hom}_K(U, V)$ ist ein Unterraum von V^U .
- $g \circ (\lambda f + \lambda' f') = \lambda(g \circ f) + \lambda'(g \circ f')$ und $(\lambda g + \lambda' g') \circ f = \lambda(g \circ f) + \lambda'(g' \circ f)$.

$$c. \lambda(g \circ f) = (\lambda g) \circ f = g \circ (\lambda f).$$

Aufgabe 23.45

Seien V, W K -Vektorräume und $f, g \in \text{Hom}_K(V, W)$. Zeige

$$\text{Ker}(f + g) \supseteq \text{Ker}(f) \cap \text{Ker}(g)$$

und

$$\text{Im}(f + g) \subseteq \text{Im}(f) + \text{Im}(g).$$

Finde außerdem Beispiele, so dass die Inklusionen strikt sind.

Aufgabe 23.46 (f -invariante Unterräume)

Ist V ein K -Vektorraum, $f : V \rightarrow V$ K -linear und $U \leq V$ ein Unterraum von V mit $f(U) \subseteq U$, so nennen wir U einen *f-invarianten* Unterraum von V .

Zeige, daß durch

$$f_U : U \rightarrow U : x \mapsto f(x)$$

und

$$f_{V/U} : V/U \rightarrow V/U : \bar{x} \mapsto \overline{f(x)}$$

K -lineare Abbildungen definiert werden.

Aufgabe 23.47

Es sei $V = \mathbb{R}^{\mathbb{R}}$ der \mathbb{R} -Vektorraum aller Abbildungen von \mathbb{R} nach \mathbb{R} und

$$U = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f(-x) = -f(x) \forall x \in \mathbb{R}\}$$

sowie

$$U' = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f(-x) = f(x) \forall x \in \mathbb{R}\}.$$

Zeige, daß U und U' Unterräume von V sind mit $V = U \oplus U'$.

Aufgabe 23.48 (Erster Isomorphiesatz)

Sei V ein K -Vektorraum und $U, U' \leq V$. Zeige

$$U/(U \cap U') \cong (U + U')/U'.$$

Aufgabe 23.49 (Projektionen)

Es sei V ein K -Vektorraum. $f \in \text{End}_K(V)$ heißt *Projektion*, falls $f^2 = f$ gilt. Zeige, die folgenden Aussagen sind äquivalent:

- f ist eine Projektion,
- $\text{id}_V - f$ ist eine Projektion,
- $\text{Im}(\text{id}_V - f) = \text{Ker}(f)$,
- $\text{Ker}(\text{id}_V - f) = \text{Im}(f)$.

Zeige auch, sind obige Bedingungen erfüllt, so gilt zudem $V = \text{Ker}(f) \oplus \text{Im}(f)$.

§ 24 Basen von Vektorräumen

In diesem Abschnitt ist V stets ein K -Vektorraum.

Das wesentlichste Konzept im Zusammenhang mit Vektorräumen ist das der Basis. Mit Hilfe einer Basis können die Elemente eines Vektorraums effizient auf eindeutige Weise dargestellt werden. Wir führen in diesem Kapitel Basen als linear unabhängige Erzeugendensysteme ein.

A) Linear unabhängige Familien von Vektoren**Definition 24.1** (Familien)

Es seien I und X zwei Mengen.

- a. Wir nennen ein Tupel der Form $(x_i)_{i \in I}$ mit $x_i \in X$ eine *Familie* von Elementen in X . Ist I endlich, so nennen wir die Familie *endlich* und setzen $|F| := |I|$.
- b. Ist $F = (x_i)_{i \in I}$ eine Familie von Elementen in X , schreiben wir $x \in F$, wenn es ein $i \in I$ gibt mit $x = x_i$.
- c. Sind $F = (x_i)_{i \in I}$ und $G = (y_i)_{i \in I}$ zwei Familien von Elementen in X , so schreiben wir $F = G$, falls $x_i = y_i$ für alle $i \in I$.
- d. Ist $F = (x_i)_{i \in I}$ eine Familie von Elementen in X und $J \subseteq I$, so nennen wir $F' = (x_j)_{j \in J}$ eine *Teilfamilie* von F und F eine *Oberfamilie* von F' , und wir schreiben $F' \subseteq F$.
- e. Wir schreiben kurz $\text{Lin}(F)$ für die lineare Hülle $\text{Lin}(\{x_i \mid i \in I\})$ und nennen $\text{Lin}(F)$ die *lineare Hülle* von F .

Beispiel 24.2

Ist $I = \{1, 2, 3\}$ und $X = \mathbb{R}^2$, so wird durch $x_1 = (1, 0)^t$, $x_2 = (1, 1)^t$, $x_3 = (1, 0)^t$ eine endliche Familie $(x_1, x_2, x_3) = ((1, 0)^t, (1, 1)^t, (1, 0)^t)$ definiert. $(x_1, x_3) = ((1, 0)^t, (1, 0)^t)$ ist eine Teilfamilie.

Bemerkung 24.3 (Familien von Vektoren)

- a. In einer Familie können Elemente auch *mehrfach* auftreten, in einer Menge geht das nicht. Z.B. $F = ((1, 0)^t, (1, 1)^t, (1, 0)^t)$.
- b. Ist die Menge I geordnet, so ordnen wir die Mitglieder der Familie F in der gleichen Weise. Z.B. $((1, 0)^t, (1, 1)^t, (1, 0)^t) \neq ((1, 0)^t, (1, 0)^t, (1, 1)^t)$.
- c. In unseren Anwendungen wird die Menge I meist $\{1, \dots, n\}$ für eine positive natürliche Zahl n sein, und wir ordnen die Elemente dann in der naheliegenden Weise.
- d. Formal korrekt sollte man die Familie $F = (x_i)_{i \in I}$ als Abbildung $F : I \rightarrow X : i \mapsto x_i$ angeben. Die Tupelschreibweise ist aber suggestiver als die Schreibweise als Abbildung.

Definition 24.4 (Lineare Unabhängigkeit)

Es sei V ein K -Vektorraum und F eine Familie von Vektoren in V .

- a. Eine endliche Familie (x_1, \dots, x_n) von Vektoren in V heißt *linear unabhängig*, wenn aus

$$\lambda_1 \cdot x_1 + \dots + \lambda_n \cdot x_n = 0$$

stets

$$\lambda_1 = \dots = \lambda_n = 0$$

folgt, d.h. wenn nur die triviale Linearkombination der x_i Null ergibt.

Wir sagen dann oft einfach, die Vektoren x_1, \dots, x_n seien linear unabhängig.

- b. Eine endliche Familie (x_1, \dots, x_n) von Vektoren in V heißt *linear abhängig*, wenn es Skalare $\lambda_1, \dots, \lambda_n \in K$ gibt, so daß

$$\lambda_1 \cdot x_1 + \dots + \lambda_n \cdot x_n = 0,$$

aber nicht alle λ_i sind Null, d.h. wenn eine nicht-triviale Linearkombination der x_i Null ergibt.

Wir nennen oft einfach die Vektoren x_1, \dots, x_n linear abhängig.

- c. F heißt *linear abhängig*, wenn es eine endliche linear abhängige Teilfamilie gibt.
 d. F heißt *linear unabhängig*, wenn jede endliche Teilfamilie linear unabhängig ist.

Beispiel 24.5 (Lineare Unabhängigkeit)

- a. Die Einheitsvektoren $e_1, \dots, e_n \in K^n$ sind linear unabhängig. Denn aus

$$(0, \dots, 0)^t = \lambda_1 \cdot e_1 + \dots + \lambda_n \cdot e_n = (\lambda_1, \dots, \lambda_n)$$

folgt unmittelbar $\lambda_1 = \dots = \lambda_n = 0$.

- b. Die Familie $((1, 0)^t, (0, 1)^t, (1, 1)^t)$ ist linear abhängig, da

$$(1, 0)^t + (0, 1)^t - (1, 1)^t = (0, 0)^t.$$

- c. Wir betrachten die Folge $e_k = (\delta_{n,k})_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, die als k -ten Eintrag eine Eins hat und ansonsten konstant Null ist. Dann ist die Familie $F = (e_k)_{k \in \mathbb{N}}$ linear unabhängig in $\mathbb{R}^{\mathbb{N}}$.

Um das zu sehen, betrachten wir die endliche Teilfamilie $(e_{k_1}, \dots, e_{k_l})$ für $0 \leq k_1 < \dots < k_l$. Dann folgt aus

$$\lambda_{k_1} \cdot e_{k_1} + \dots + \lambda_{k_l} \cdot e_{k_l} = (0)_{n \in \mathbb{N}}$$

unmittelbar $\lambda_{k_1} = \dots = \lambda_{k_l} = 0$, da die linke Folge als Folgenglied k_i den Wert λ_{k_i} hat. Also ist jede endliche Teilfamilie von F linear unabhängig, und somit ist auch F linear unabhängig.

Lemma 24.6 (Kriterien für lineare Abhängigkeit)

Es sei $F = (x_i)_{i \in I}$ eine Familie von Vektoren im K -Vektorraum V .

- a. Ist $0 \in F$, so ist F linear abhängig.

- b. *Gibt es ein $i \neq j$ mit $x_i = x_j$, so ist F linear abhängig.*
- c. *F ist genau dann linear abhängig, wenn es ein $x \in F$ gibt, das Linearkombination anderer Vektoren in F ist.*

Beweis: Im ersten Fall ist $1 \cdot 0_V = 0_V$ eine nicht-triviale Linearkombination, im zweiten Fall ist $1 \cdot x_i - 1 \cdot x_j = 0_V$ eine solche. In jedem Fall ist F also linear abhängig, weil F eine endliche linear abhängige Teilfamilie enthält. Damit sind a. und b. gezeigt.

Ist F linear abhängig, so gibt es eine nicht-triviale Linearkombination

$$\sum_{j \in J} \lambda_j \cdot x_j = 0$$

mit $J \subseteq I$ endlich und nicht alle λ_j sind Null. Sei also $i \in J$ mit $\lambda_i \neq 0$, dann ist

$$x_i = \sum_{i \neq j \in J} -\frac{\lambda_j}{\lambda_i} \cdot x_j$$

Linearkombination anderer Vektoren in F .

Ist umgekehrt $x_i = \sum_{j \in J} \lambda_j \cdot x_j$ mit $J \subseteq I$ endlich und $i \in I \setminus J$, so ist

$$-x_i + \sum_{j \in J} \lambda_j \cdot x_j = 0$$

eine nicht-triviale Linearkombination, die Null ergibt. Mithin ist F linear abhängig. \square

Beispiel 24.7

In Beispiel 24.5 b. gilt

$$(1, 0)^t = -(0, 1)^t + (1, 1)^t,$$

woraus ebenfalls die lineare Abhängigkeit der Familie folgt.

Notation 24.8 (Linearkombination)

Sei $F = (x_i)_{i \in I}$ eine Familie von Vektoren in V und I sei nicht notwendigerweise endlich. Wir werden des öfteren

$$x = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot x_i \tag{65}$$

schreiben, wenn wir sagen wollen, daß x eine Linearkombination von Vektoren in F ist. Formal korrekt müßte es lauten: es gibt eine endliche Teilfamilie $(x_j)_{j \in J}$ von F und Skalare $\lambda_j \in K$ für $j \in J$, so daß

$$x = \sum_{j \in J} \lambda_j \cdot x_j.$$

Wir interpretieren dies so, daß in (65) nur endlich viele der λ_i nicht Null sind, und daß somit die Summe auf der rechten Seite doch eine endliche Summe ist.

Mit dieser neuen Notation ist F genau dann linear unabhängig, wenn aus

$$\sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot x_i = 0$$

stets $\lambda_i = 0$ für alle $i \in I$ folgt; und analog ist F linear abhängig, wenn es eine Linearkombination

$$\sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot x_i = 0$$

gibt, bei der nicht alle λ_i Null sind.

Lemma 24.9 (Ergänzung linear unabhängiger Familien)

Ist $B = (x_i)_{i \in I}$ eine linear unabhängige Familie in V mit $\text{Lin}(B) \subsetneq V$, so ist die Familie $(x, x_i \mid i \in I)$ für jedes $x \in V \setminus \text{Lin}(B)$ linear unabhängig.

Beweis: Seien dazu $\lambda, \lambda_i \in K$, $i \in I$, mit

$$\lambda x + \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i x_i = 0.$$

Wäre $\lambda \neq 0$, so wäre

$$x = \sum_{\substack{i \in I \\ \text{endlich}}} -\frac{\lambda_i}{\lambda} \cdot x_i \in \text{Lin}(B)$$

im Widerspruch zur Wahl von x . Also ist $\lambda = 0$, und somit folgt aus

$$\sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i x_i = \lambda x + \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i x_i = 0$$

und der linearen Unabhängigkeit von B , daß auch alle anderen λ_i Null sind. Also ist $(x, x_i \mid i \in I)$ linear unabhängig. \square

B) Erzeugendensysteme und Basen

Definition 24.10 (Erzeugendensystem und Basis)

Es sei V ein K -Vektorraum und F eine Familie von Vektoren in V .

- F heißt ein *Erzeugendensystem* von V , wenn $V = \text{Lin}(F)$, d.h. wenn jeder Vektor in V eine Linearkombination von Vektoren in F ist.
- F heißt eine *Basis* von V , wenn F ein linear unabhängiges Erzeugendensystem von V ist.
- V heißt *endlich erzeugt*, wenn V ein endliches Erzeugendensystem besitzt.

Beispiel 24.11 (Erzeugendensystem und Basis)

- Die Familie $B = (e_1, \dots, e_n)$ der Einheitsvektoren im K^n ist eine Basis des K^n , die wir auch die *Standardbasis* oder die *kanonische Basis* des K^n nennen. Denn nach Beispiel 24.5 ist B linear unabhängig und zudem ist ein beliebiger Vektor $x = (x_1, \dots, x_n)^t \in K^n$ eine Linearkombination

$$x = x_1 \cdot e_1 + \dots + x_n \cdot e_n$$

der Vektoren in B .

- b. Analog sieht man, daß für $n, m \geq 1$ die Familie

$$(\mathbb{E}_i^j \mid i = 1, \dots, m; j = 1, \dots, n),$$

wobei $\mathbb{E}_i^j = (e_{lk})_{l=1, \dots, m; k=1, \dots, n}$ mit

$$e_{lk} = \delta_{il} \cdot \delta_{jk} = \begin{cases} 1, & \text{falls } l = i \text{ und } k = j, \\ 0, & \text{sonst} \end{cases}$$

die Matrix ist, die in Zeile i und Spalte j eine Eins als Eintrag hat und sonst nur Nullen, eine Basis des K -Vektorraums $\text{Mat}(m \times n, K)$ ist.

- c. Die Familie $((1, 0)^t, (0, 1)^t, (1, 1)^t)$ in Beispiel 24.5 b. ist ein Erzeugendensystem von \mathbb{R}^2 , aber keine Basis, da sie linear abhängig ist.
- d. Die Familie $(e_k)_{k \in \mathbb{N}}$ im Vektorraum der Folgen $\mathbb{R}^{\mathbb{N}}$ aus Beispiel 24.5 c. ist *kein* Erzeugendensystem von $\mathbb{R}^{\mathbb{N}}$. Es scheint zwar, als gelte für eine beliebige Folge $(a_n)_{n \in \mathbb{N}}$

$$(a_n)_{n \in \mathbb{N}} = (a_0, a_1, a_2, \dots) = \sum_{n \in \mathbb{N}} a_n \cdot e_n,$$

aber diese Summe ist *nicht* endlich und mithin keine zulässige Linearkombination! Die konstante Folge $(1)_{n \in \mathbb{N}}$ ist sicher keine endliche Linearkombination der e_k , da eine solche nur endlich viele Folgenglieder ungleich Null haben kann.

- e. Die Familie $(1, i)$ ist eine Basis von \mathbb{C} als \mathbb{R} -Vektorraum, da jede komplexe Zahl von der Gestalt $x + iy$ mit $x, y \in \mathbb{R}$ ist und da eine solche Zahl nur dann Null ist, wenn x und y beide Null sind.
- f. Die Familie $B = (t^0, t^1, t^2, \dots)$ ist eine Basis von $K[t]$.

Proposition 24.12 (Eindeutige Darstellbarkeit bezüglich einer Basis)

Eine Familie B von Vektoren in V ist genau dann eine Basis von V , wenn jeder Vektor in V in eindeutiger Weise als Linearkombination von Elementen in B geschrieben werden kann.

Beweis: Sei zunächst $B = (x_i)_{i \in I}$ eine Basis von V und $x \in V$. Nach Voraussetzung ist B ein Erzeugendensystem von V und mithin ist

$$x = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot x_i$$

eine Linearkombination von Vektoren in B . Ist nun

$$x = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda'_i \cdot x_i$$

eine zweite Linearkombination von Vektoren in B , die x ergibt, so ist

$$0 = x - x = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot x_i - \sum_{\substack{i \in I \\ \text{endlich}}} \lambda'_i \cdot x_i = \sum_{\substack{i \in I \\ \text{endlich}}} (\lambda_i - \lambda'_i) \cdot x_i$$

eine Linearkombination von Vektoren in B , die Null ergibt. Da B linear unabhängig ist, muß dann aber stets

$$\lambda_i - \lambda'_i = 0$$

gelten. Die Darstellung ist also eindeutig.

Sie nun umgekehrt jeder Vektor x in V auf eindeutige Weise als Linearkombination der Vektoren in B darstellbar. Dann ist offenbar B ein Erzeugendensystem von V , und 0_V kann nur auf die triviale Weise als Linearkombination von Vektoren in B dargestellt werden, so daß B auch linear unabhängig ist. \square

Beispiel 24.13

$B = ((1, 1)^t, (1, -1)^t)$ ist eine Basis von \mathbb{R}^2 , da sich ein beliebiger Vektor $(\lambda_1, \lambda_2)^t \in \mathbb{R}^2$ in eindeutiger Weise als

$$(\lambda_1, \lambda_2)^t = \frac{\lambda_1 + \lambda_2}{2} \cdot (1, 1)^t + \frac{\lambda_1 - \lambda_2}{2} \cdot (1, -1)^t$$

schreiben läßt, wie man leicht sieht.

Satz 24.14 (Existenz- und Eindeutigkeitsatz für lineare Abbildungen)

Es seien V und W zwei K -Vektorräume, $B = (x_i)_{i \in I}$ eine Basis von V und $F = (y_i)_{i \in I}$ eine Familie von Vektoren in W .

Dann gibt es genau eine K -lineare Abbildung $f: V \rightarrow W$, so daß für alle $i \in I$

$$f(x_i) = y_i.$$

Insbesondere, zwei lineare Abbildungen sind gleich, sobald sie auf einer Basis übereinstimmen.

Beweis: Jeder Vektor $x \in V$ läßt sich nach Proposition 24.12 in eindeutiger Weise als Linearkombination

$$x = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot x_i$$

schreiben. Wir definieren die Abbildung f dann durch

$$f(x) := \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot y_i. \quad (66)$$

Wir wollen nun zeigen, daß $f: V \rightarrow W$ dann K -linear ist. Seien dazu

$$x = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot x_i, \quad x' = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda'_i \cdot x_i \in V$$

und $\lambda, \lambda' \in K$ gegeben, dann gilt für die eindeutige Darstellung von $\lambda x + \lambda' x'$ offenbar

$$\lambda x + \lambda' x' = \sum_{\substack{i \in I \\ \text{endlich}}} (\lambda \cdot \lambda_i + \lambda' \cdot \lambda'_i) \cdot x_i,$$

und mithin erhalten wir

$$\begin{aligned}
 f(\lambda x + \lambda' x') &= f\left(\sum_{\substack{i \in I \\ \text{endlich}}} (\lambda \cdot \lambda_i + \lambda' \cdot \lambda'_i) \cdot x_i\right) \\
 &= \sum_{\substack{i \in I \\ \text{endlich}}} (\lambda \cdot \lambda_i + \lambda' \cdot \lambda'_i) \cdot y_i \\
 &= \lambda \cdot \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot y_i + \lambda' \cdot \sum_{\substack{i \in I \\ \text{endlich}}} \lambda'_i \cdot y_i \\
 &= \lambda \cdot f(x) + \lambda' \cdot f(x').
 \end{aligned}$$

Die Abbildung f ist also K -linear, und nach Definition gilt auch $f(x_i) = y_i$.

Es bleibt zu zeigen, daß es keine zweite K -lineare Abbildung geben kann, die diese Eigenschaft hat. Sei dazu $g : V \rightarrow W$ eine K -lineare Abbildung mit $g(x_i) = y_i$ für alle $i \in I$. Ein beliebiges $x \in V$ läßt sich wieder schreiben als

$$x = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot x_i$$

und dann gilt

$$f(x) = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot f(x_i) = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot y_i = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot g(x_i) = g(x).$$

Mithin stimmt f mit g überein. □

Bemerkung 24.15 (Existenz- und Eindeutigkeitssatz für lineare Abbildungen)

Satz 24.14 besagt, daß man die Werte einer linearen Abbildung auf einer Basis beliebig vorschreiben kann. Egal welche Vektoren im Zielbereich man als Bilder wählt, es gibt eine und nur eine lineare Abbildung, die den Basiselementen genau diese Vektoren zuordnet!

Wegen der Formel für $f(x)$ in (66) sagt man auch, daß sich f aus der Vorschrift $f(x_i) = y_i$, $i \in I$, durch *lineare Fortsetzung* ergibt.

Beispiel 24.16

Setzen wir $x_1 = (1, 1)^t$ und $x_2 = (1, -1)^t$, so ist $B = (x_1, x_2)$ eine Basis von \mathbb{R}^2 . Wählen wir nun zudem $y_1 = (1, 1)^t$ und $y_2 = (3, 1)^t$, so muß es genau eine \mathbb{R} -lineare Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ geben mit

$$f((1, 1)^t) = f(x_1) = y_1 = (1, 1)^t \quad \text{und} \quad f((1, -1)^t) = f(x_2) = y_2 = (3, 1)^t.$$

Diese besitzt die Abbildungsvorschrift

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2 : (x, y)^t \mapsto (2x - y, x)^t.$$

Wir werden später sehen, wie man die Abbildungsvorschrift systematisch bestimmen kann.

Korollar 24.17 (Alle linearen Abbildungen $K^n \rightarrow K^m$ sind von der Form f_A .)
 Jede lineare Abbildung $f : K^n \rightarrow K^m$ ist von der Form $f = f_A$ für eine eindeutig bestimmte Matrix $A \in \text{Mat}(m \times n, K)$.

Beweis: Ist $f : K^n \rightarrow K^m$ eine lineare Abbildung, so setzen wir $\mathbf{a}^i := f(\mathbf{e}_i) \in K^m$ für $i = 1, \dots, n$ und bilden eine Matrix A mit den \mathbf{a}^i als Spaltenvektoren. Dann ist f_A eine lineare Abbildung, mit

$$f_A(\mathbf{e}_i) = A\mathbf{e}_i = \mathbf{a}^i = f(\mathbf{e}_i),$$

so daß aus der Eindeutigkeitsaussage in 24.14 unmittelbar $f_A = f$ folgt. Die Eindeutigkeit der Matrix A folgt aus der Tatsache, daß A die Abbildung f_A eindeutig festlegt. \square

Proposition 24.18 (Charakterisierung von Basen)

Für eine Familie B von Vektoren in V sind die folgenden Aussagen gleichwertig:

- B ist eine Basis von V .
- B ist ein minimales Erzeugendensystem von V .
- B ist eine maximale linear unabhängige Familie in V .

Bemerkung 24.19

Ein Erzeugendensystem B von V heißt *minimal*, wenn keine echte Teilfamilie von B ein Erzeugendensystem ist. Dies heißt *nicht*, daß sie in jedem anderen Erzeugendensystem enthalten ist! Es gibt nicht *das* minimale Erzeugendensystem.

Eine linear unabhängige Familie B in V heißt *maximal*, wenn keine echte Oberfamilie linear unabhängig ist. Dies heißt *nicht*, daß sie jede andere linear unabhängige Familie enthält! Es gibt nicht *die* maximale linear unabhängige Familie.

Beweis von Proposition 24.18: Es sei $B = (\mathbf{x}_i)_{i \in I}$.

a. \Rightarrow b.: Ist B eine Basis, so erzeugt B den Vektorraum V per definitionem. Ist $(\mathbf{x}_j \mid j \in J)$ eine echte Teilfamilie von B und ist $i \in I \setminus J$, so gibt es wegen der linearen Unabhängigkeit von B keine Darstellung

$$\mathbf{x}_i - \sum_{\substack{j \in J \\ \text{endlich}}} \lambda_j \mathbf{x}_j = 0$$

also ist $\mathbf{x}_i \notin \text{Lin}(\mathbf{x}_j \mid j \in J)$.

b. \Rightarrow c.: Wir zeigen zunächst, daß B linear unabhängig ist. Angenommen, dies sei nicht der Fall, dann läßt sich nach Lemma 24.6 ein \mathbf{x}_i als Linearkombination

$$\mathbf{x}_i = \sum_{\substack{i \neq j \in I \\ \text{endlich}}} \lambda_j \mathbf{x}_j$$

der übrigen Vektoren in B darstellen. Damit gilt dann aber

$$\text{Lin}(\mathbf{x}_j \mid j \in I \setminus \{i\}) = \text{Lin}(\mathbf{x}_j \mid j \in I) = V,$$

im Widerspruch zur Minimalität des Erzeugendensystems B .

Sei nun $(x_j \mid j \in J)$ mit $I \subsetneq J$ eine echte Oberfamilie von B und $j \in J \setminus I$, so ist x_j eine Linearkombination

$$x_j = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i x_i$$

der Elemente in B , da B ein Erzeugendensystem ist. Folglich ist $(x_j \mid j \in J)$ linear abhängig nach Lemma 24.6.

c. \Rightarrow a.: Da B linear unabhängig ist, bleibt zu zeigen, daß $\text{Lin}(B) = V$. Gäbe es ein $x \in V \setminus \text{Lin}(B)$, so wäre wegen Lemma 24.9 auch $(x, x_i \mid i \in I)$ linear unabhängig, im Widerspruch zur Maximalität von B .

□

Beispiel 24.20

Kommen wir zu unserem Beispiel $B = ((1, 1)^t, (1, -1)^t)$ zurück. Da sich ein beliebiger Vektor $(\lambda_1, \lambda_2)^t \in \mathbb{R}^2$ als

$$(\lambda_1, \lambda_2)^t = \frac{\lambda_1 + \lambda_2}{2} \cdot (1, 1)^t + \frac{\lambda_1 - \lambda_2}{2} \cdot (1, -1)^t$$

schreiben läßt, ist B ein Erzeugendensystem von \mathbb{R}^2 , und offenbar kann man weder $(1, 1)^t$ noch $(1, -1)^t$ weglassen. B ist also ein minimales Erzeugendensystem und mithin eine Basis von \mathbb{R}^2 . Damit sparen wir uns, die Eindeutigkeit obiger Darstellung von $(\lambda_1, \lambda_2)^t$ zu zeigen, von der wir in Beispiel 24.13 nur gesagt haben, daß man sie leicht nachweisen könne!

C) Existenz von Basen

Wir wollen jetzt zeigen, daß jeder Vektorraum eine Basis besitzt. Da wir dies nicht nur für Vektorräume mit einem endlichen Erzeugendensystem beweisen wollen, benötigen wir das folgende *Zornsche Lemma*, eine Aussage, die zu den logischen Grundlagen der Mathematik gehört. Grob gesprochen gehört es zu den (im Rahmen einer formalen Mengenlehre) nicht aus anderen Axiomen herleitbaren Axiomen. Man kann aber zeigen, daß das Zornsche Lemma äquivalent zum *Wohlordnungssatz* und zum *Auswahlaxiom* ist, vgl. [Moo82, Sze50]. Ohne diese Axiome läßt sich der Existenzsatz über Basen nicht für beliebige Vektorräume beweisen.

Für endlich erzeugbare Vektorräume, die für die Praxis natürlich ausreichen, werden wir im nächsten Abschnitt einen konstruktiven Beweis geben. Nicht nur in der Theorie der Vektorräume ist das Zornsche Lemma ein wichtiges Beweismittel für unendliche Objekte.

Definition 24.21 (Ketten)

Es sei (M, \leq) eine teilgeordnete Menge. Eine Familie $K = (K_n)_{n \in \mathbb{N}}$ heißt eine *Kette* in M , falls $K_0 \leq K_1 \leq K_2 \leq \dots$

Axiom 24.22 (Lemma von Zorn)

Es sei (M, \leq) eine nicht-leere, teilgeordnete Menge. Wenn jede Kette in M eine obere Schranke in M besitzt, so besitzt M ein maximales Element.

Die Hauptarbeit der Existenz von Basen wird in der folgenden Proposition geleistet.

Proposition 24.23 (Basisexistenzsatz)

Sei E ein Erzeugendensystem des K -Vektorraums V und sei F eine linear unabhängige Teilfamilie von E , dann gibt es eine Basis B , die Teilfamilie von E ist und F als Teilfamilie enthält, d.h. $F \subseteq B \subseteq E$.

Beweis: Wir betrachten das Mengensystem

$$M = \{G \mid G \text{ ist linear unabhängige Familie in } V, F \subseteq G \subseteq E\}.$$

Da $F \in M$, ist M nicht leer. Außerdem ist M bezüglich der Inklusion " \subseteq " teilgeordnet (vgl. Beispiel 8.4).

Behauptung: Jede Kette K in M besitzt eine obere Schranke.

Sei also $K = (G_n)_{n \in \mathbb{N}}$ eine Kette in M . Wir setzen $X := \bigcup_{n \in \mathbb{N}} G_n$. Dann gilt sicher $G_n \subseteq X$ für alle $n \in \mathbb{N}$. Sobald wir gezeigt haben, daß $X \in M$, ist also X eine obere Schranke von K . Wegen $F \subseteq G_n \subseteq E$ für alle $n \in \mathbb{N}$, gilt auch

$$F \subseteq X \subseteq E.$$

Es bleibt also noch zu zeigen, daß X linear unabhängig ist. Seien dazu (x_1, \dots, x_r) eine endliche Teilfamilie von X . Dann gibt es zu jedem $i \in \{1, \dots, r\}$ ein G_{n_i} mit $x_i \in G_{n_i}$. Da K eine Kette ist, gibt es aber ein n mit $G_{n_i} \subseteq G_n$ für alle $i = 1, \dots, r$, so daß (x_1, \dots, x_r) eine Teilfamilie der linear unabhängigen Familie G_n ist. Damit ist auch (x_1, \dots, x_r) linear unabhängig. Also ist X linear unabhängig, d. h. in M , und damit eine obere Schranke von K .

Wir können also das Zornsche Lemma auf M anwenden und erhalten, daß M ein maximales Element B besitzt.

Behauptung: B ist eine Basis von V .

Da $B \in M$, ist also B linear unabhängig, und es bleibt zu zeigen, daß B ein Erzeugendensystem von V ist. Nehmen wir an, dies sei nicht der Fall. Dann gilt insbesondere $E \not\subseteq \text{Lin}(B)$ und es gibt ein Element $x \in E \setminus \text{Lin}(B)$. Nach Lemma 24.9 gilt dann aber, daß $(x, y \mid y \in B)$ linear unabhängig ist und somit, wegen $F \subseteq (x, y \mid y \in B) \subseteq E$,

$$(x, y \mid y \in B) \in M,$$

im Widerspruch zur Maximalität von B . Also war die Annahme falsch und B ist eine Basis. □

Korollar 24.24 (Basisergänzungssatz)

- a. *Jede linear unabhängige Familie in V kann zu einer Basis ergänzt werden.*
- b. *Jedes Erzeugendensystem von V enthält eine Basis.*

Beweis: Für a. wähle zu gegebenem F' in Proposition 24.23 $F = F'$ und $E = V$, und für b. wähle zu gegebenem E' dort $F = \emptyset$ und $E = E'$. \square

Beispiel 24.25

Das Erzeugendensystem $E = ((1, 0)^t, (1, 1)^t, (0, 1)^t)$ von \mathbb{R}^2 enthält kanonische Basis, und die linear unabhängige Familie $F = ((1, 1)^t)$ kann zur Basis $B = ((1, 1)^t, (1, -1)^t)$ von \mathbb{R}^2 ergänzt werden.

Korollar 24.26 (Existenz von Basen)

Jeder Vektorraum besitzt eine Basis.

Korollar 24.27 (Existenz von Komplementen)

Jeder Unterraum U von V besitzt ein direktes Komplement.

Beweis: Wähle eine Basis B von U und ergänze sie durch eine linear unabhängige Menge B' zu einer Basis $B \cup B'$ von V gemäß dem Basisergänzungssatz 24.24. Dann ist $U' := \text{Lin}(B')$ ein Komplement von U , denn

$$U + U' = \text{Lin}(U \cup U') \supseteq \text{Lin}(B \cup B') = V$$

und aus

$$x = \sum_{\substack{y \in B \\ \text{endlich}}} \lambda_y \cdot y = \sum_{\substack{z \in B' \\ \text{endlich}}} \lambda_z \cdot z \in U \cap U'$$

folgt wegen der linearen Unabhängigkeit von $B \cup B'$ mit

$$\sum_{\substack{y \in B \\ \text{endlich}}} \lambda_y \cdot y + \sum_{\substack{z \in B' \\ \text{endlich}}} -\lambda_z \cdot z = 0,$$

daß alle λ_y und λ_z Null sein müssen, so daß auch $U \cap U' = \{0\}$. \square

Beispiel 24.28 (\mathbb{R} als \mathbb{Q} -Vektorraum)

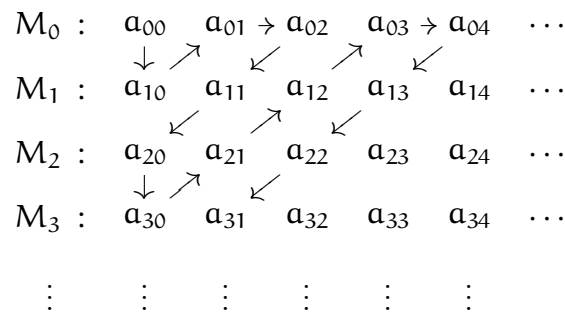
Auch wenn jeder Vektorraum eine Basis besitzt, kann nicht notwendigerweise für jeden Vektorraum eine Basis angegeben werden. \mathbb{R} als \mathbb{Q} -Vektorraum ist ein gutes Beispiel für einen Vektorraum, bei dem man keine Basis angeben kann.

Behauptung: Eine Basis von \mathbb{R} als \mathbb{Q} -Vektorraum ist überabzählbar.

Hierzu argumentieren wir wie folgt, wobei wir eine Menge *höchstens abzählbar* nennen, wenn sie endlich oder abzählbar ist.

- \mathbb{R} ist überabzählbar nach Proposition 5.9.
- Die abzählbare Vereinigung höchstens abzählbarer Mengen ist wieder höchstens abzählbar. Seien dazu $M_i = \{\mathbf{a}_{i0}, \mathbf{a}_{i1}, \mathbf{a}_{i2}, \dots\}$, $i \in \mathbb{N}$, (ohne Einschränkung)

abzählbare Mengen, dann schreiben wir sie wie folgt auf:



Abzählen der Elemente wie angedeutet, wobei man Elemente, die mehrfach vorkommen, nur beim ersten Mal berücksichtigt, liefert eine Bijektion von $\mathbb{N} \rightarrow \bigcup_{i \in \mathbb{N}} M_i$, mithin ist die Vereinigung abzählbar.

c. Es gilt also $\mathbb{Z} = \mathbb{N} \cup (-\mathbb{N})$ und

$$\mathbb{Q} = \bigcup_{0 \neq q \in \mathbb{N}} \frac{1}{q} \mathbb{Z} = \bigcup_{0 \neq q \in \mathbb{N}} \left\{ \frac{p}{q} \mid p \in \mathbb{Z} \right\}$$

sind abzählbar.

d. Das kartesische Produkt zweier höchstens abzählbarer Mengen ist wieder höchstens abzählbar. Seien dazu M und N zwei höchstens abzählbare Mengen, dann gilt

$$M \times N = \bigcup_{m \in M} \{m\} \times N,$$

wobei $N \rightarrow \{m\} \times N : n \mapsto (m, n)$ eine Bijektion ist, $\{m\} \times N$ also höchstens abzählbar ist.

e. Ein Vektorraum V über einem höchstens abzählbaren Körper K mit höchstens abzählbarer Basis ist höchstens abzählbar. Sei dazu (ohne Einschränkung) $B = (x_i \mid i \in \mathbb{N})$, eine abzählbare Basis von V . Für $n \in \mathbb{N}$ setzen wir

$$V_n := \text{Lin}(x_1, \dots, x_n).$$

Dann gilt $V_n \cong K^n$, also ist V_n nach d. mit Induktion über n abzählbar. Aber dann ist $V = \bigcup_{n \in \mathbb{N}} V_n$ nach b. abzählbar.

f. Da \mathbb{Q} abzählbar ist, \mathbb{R} aber überabzählbar, folgt aus e. die Behauptung.

Bemerkung 24.29 (Ringe und Moduln)

In diesem Abschnitt haben wir in vielen Beweisen verwendet, daß man durch Körperelemente teilen darf, sobald sie nicht Null sind. Deshalb gelten viele Aussagen nicht mehr im allgemeinen für Moduln über Ringen. Die Definitionen lassen sich aber dennoch in der gleichen Weise geben, und wir wollen hier zusammenstellen, welche Aussagen für Moduln über Ringen letztlich wahr bleiben.

Die Definitionen 24.1, 24.4, 24.8 und 24.10 können für Moduln über Ringen in der gleichen Weise gegeben werden. Beispiele 24.5 und 24.11 bleiben dann ebenso richtig

wie die Aussagen in Lemma 24.6 a. und c. und in den wichtigen Sätzen 24.12, 24.14 und 24.17.

Die Aussagen in Lemma 24.6 und 24.9, Proposition 24.18 sowie in den Korollaren 24.24 und 24.26 gelten für Moduln über Ringen im allgemeinen *nicht* mehr. In ihren Beweisen wird durch Skalare geteilt, von denen nur bekannt ist, daß sie nicht Null sind.

Aufgaben

Aufgabe 24.30

Welche der folgenden Familien sind linear unabhängig / Erzeugendensysteme / Basen von \mathbb{R}^2 ?

- a. $((1, 0)^t, (0, 1)^t, (1, 1)^t)$.
- b. $((1, 1)^t, (2, 2)^t)$.
- c. $((1, 3)^t)$.
- d. $((1, 1)^t, (1, -2)^t)$.
- e. $((1, 1)^t, (0, 0)^t)$.
- f. $((1, 1)^t, (0, 0)^t, (1, -2)^t)$.
- g. $((1, 2)^t, (2, 1)^t)$.

Aufgabe 24.31

Es sei V ein K -Vektorraum, $U \subset V$ ein Unterraum, $0 \neq x \in U$ und $y \in V \setminus U$.
Zeige, daß (x, y) linear unabhängig ist.

Aufgabe 24.32

Ist $f: V \rightarrow W$ eine K -lineare Abbildung, F eine Familie von Vektoren in V , so ist

$$f(\text{Lin}(F)) = \text{Lin}(f(x) \mid x \in F).$$

Aufgabe 24.33

Es seien V und W zwei K -Vektorräume, $f: V \rightarrow W$ eine K -lineare Abbildung und B eine Basis von V .

- a. Genau dann ist f surjektiv, wenn $f(B)$ ein Erzeugendensystem von W ist.
- b. Genau dann ist f injektiv, wenn $f(B)$ linear unabhängig ist.
- c. Genau dann ist f bijektiv, wenn $f(B)$ eine Basis von W ist.

Aufgabe 24.34

Es seien U_1, \dots, U_k Unterräume eines K -Vektorraums V mit Basen B_1, \dots, B_k .
Zeige, genau dann ist $V = U_1 \oplus \dots \oplus U_k$ die direkte Summe der U_i , wenn $B = B_1 \cup \dots \cup B_k$ eine Basis von V ist.

Aufgabe 24.35

Sei V ein \mathbb{C} -Vektorraum, dann ist V offensichtlich auch ein \mathbb{R} -Vektorraum, und seien $x_1, \dots, x_n \in V$. Zeige, daß (x_1, \dots, x_n) genau dann linear unabhängig über \mathbb{C} ist, wenn $(x_1, ix_1, \dots, x_n, ix_n)$ linear unabhängig über \mathbb{R} ist.

Aufgabe 24.36

Es sei $(V, +, \cdot)$ ein K -Vektorraum, und $x_1, \dots, x_n \in V$ seien linear abhängige Vektoren mit der Eigenschaft, daß je $n - 1$ der Vektoren linear unabhängig sind. Zeige:

- a. Es gibt $\lambda_1, \dots, \lambda_n \in K \setminus \{0\}$ mit der Eigenschaft

$$\sum_{i=1}^n \lambda_i x_i = 0.$$

- b. Gilt für $\mu_1, \dots, \mu_n \in K$ ebenfalls $\sum_{i=1}^n \mu_i x_i = 0$, so gibt es ein $v \in K$ mit $\mu_i = \lambda_i \cdot v$ für alle $i = 1, \dots, n$.

Aufgabe 24.37

Es sei $A \in \text{Mat}_n(K)$.

- a. Genau dann ist f_A bijektiv, wenn $A \in \text{Gl}_n(K)$.
- b. Ist $A \in \text{Gl}_n(K)$, so gilt $(f_A)^{-1} = f_{A^{-1}}$.

§ 25 Endlich-dimensionale Vektorräume

Wir betrachten jetzt *endlich erzeugte Vektorräume* V , d. h. Vektorräume, die ein endliches Erzeugendensystem besitzen. Nach Korollar 24.24 besitzt V dann auch eine endliche Basis. Für solche Vektorräume kann man die Sätze des vorigen Abschnitts teilweise verschärfen und vor allem kann man in diesen Vektorräumen mit Hilfe von Basen und Matrizen effizient rechnen.

In diesem Abschnitt ist V stets ein endlich erzeugter K -Vektorraum.

A) Austauschatz von Steinitz

Lemma 25.1 (Austauschlemma)

Sei $B = (x_1, \dots, x_n)$ eine Basis von V , $y = \sum_{i=1}^n \lambda_i x_i$ und $\lambda_j \neq 0$ für ein j .

Dann ist $(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_n)$ eine Basis von V , d.h. man kann in der Basis B den Vektor x_j gegen y austauschen.

Beweis: Wegen $\lambda_j \neq 0$ gilt

$$x_j = \frac{1}{\lambda_j} \cdot y - \sum_{i \neq j} \frac{\lambda_i}{\lambda_j} \cdot x_i,$$

und somit

$$V = \text{Lin}(x_1, \dots, x_n) = \text{Lin}(y, x_1, \dots, x_n) = \text{Lin}(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_n).$$

Bleibt also zu zeigen, daß $(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_n)$ linear unabhängig ist. Seien dazu $\mu_i \in K$, $i = 1, \dots, n$, gegeben mit

$$\begin{aligned} 0 &= \mu_j y + \sum_{i \neq j} \mu_i x_i = \sum_{i=1}^n \mu_j \lambda_i x_i + \sum_{i \neq j} \mu_i x_i \\ &= \mu_j \lambda_j x_j + \sum_{i \neq j} (\mu_j \lambda_i + \mu_i) x_i. \end{aligned}$$

Dann folgt aus der linearen Unabhängigkeit von x_1, \dots, x_n

$$\mu_j \lambda_j = 0 \quad \text{und} \quad \mu_i = -\mu_j \lambda_i, \quad \text{für } i \neq j.$$

Wegen $\lambda_j \neq 0$, ist also $\mu_j = 0$ und damit auch

$$\mu_i = 0 \quad \text{für } i \neq j.$$

Damit ist die lineare Unabhängigkeit von $(x_1, \dots, x_{j-1}, y, x_{j+1}, \dots, x_n)$ gezeigt. \square

Beispiel 25.2

Ist zum Beispiel $E = (e_1, \dots, e_n)$ die kanonische Basis des K^n und $x = (\lambda_1, \dots, \lambda_n)^t \in K^n$ mit $\lambda_j \neq 0$, so können wir e_j gegen x austauschen und erhalten wieder eine Basis.

Konkret kann man in der Basis $E = (e_1, e_2, e_3)$ von \mathbb{R}^3 den Vektor $(1, 2, 0)^t$ gegen e_1 oder e_2 austauschen, nicht aber gegen e_3 .

Das Austauschlemma wird benutzt, um den wichtigen Steinitzschen Austauschsatz zu beweisen.

Satz 25.3 (Austauschsatz von Steinitz)

Sei (x_1, \dots, x_n) eine Basis von V und (y_1, \dots, y_r) sei linear unabhängig in V .

Dann lassen sich die x_1, \dots, x_n so umnumerieren, daß $(y_1, \dots, y_r, x_{r+1}, \dots, x_n)$ eine Basis von V ist. Insbesondere gilt $r \leq n$.

Beweis von Satz 25.3: Wir führen den Beweis mittels Induktion über r .

Für $r = 0$ ist die Behauptung offensichtlich richtig. Nehmen wir also an, daß $r > 0$ und daß die Behauptung bereits richtig ist für $r - 1$. D. h. nach evt. Umnumerieren ist $(y_1, \dots, y_{r-1}, x_r, \dots, x_n)$ eine Basis von V . Dann besitzt y_r eine Darstellung der Form

$$y_r = \lambda_1 y_1 + \dots + \lambda_{r-1} y_{r-1} + \lambda_r x_r + \dots + \lambda_n x_n,$$

mit $\lambda_i \in K$. Angenommen, $\lambda_r = \dots = \lambda_n = 0$, dann wäre (y_1, \dots, y_r) linear abhängig, im Widerspruch zur Voraussetzung. Also gibt es ein $j \in \{r, \dots, n\}$ mit $\lambda_j \neq 0$. Durch Umnumerieren können wir annehmen, daß $j = r$ gilt. Dann können wir aber nach dem Austauschlemma 25.1 y_r gegen x_r austauschen, und die Behauptung ist bewiesen. \square

Bemerkung 25.4

- Der Austauschsatz von Steinitz besagt also, daß man - nach eventuellem Umnumerieren - die linear unabhängigen Vektoren x_1, \dots, x_r durch y_1, \dots, y_r ersetzen kann.
- Im Austauschsatz tauschen wir nacheinander x_{i_1} durch y_1 , x_{i_2} durch y_2 , etc. und schließlich x_{i_r} durch y_r für geeignete i_1, \dots, i_r aus. Im j -ten Schritt wissen wir, daß wir eine Darstellung

$$y_j = \sum_{l=1}^{j-1} \lambda_l y_l + \sum_{l \notin \{i_1, \dots, i_{j-1}\}} \lambda_l x_l$$

haben mit $\lambda_l \neq 0$ für ein $l \notin \{i_1, \dots, i_{j-1}\}$, und setzen wir dann $i_j := l$, so können wir x_{i_j} durch y_j ersetzen.

Wie wir eine solche Darstellung von y_j mit Hilfe des Gauß'schen Algorithmus berechnen können, werden wir später sehen. Damit haben wir dann ein konstruktives Verfahren für die Anwendung des Steinitz'schen Austauschsatzes.

B) Die Dimension eines endlich-erzeugten Vektorraums

Als Folgerung des Steinitz'schen Austauschsatzes erhalten wir den folgenden Satz.

Satz 25.5 (Alle Basen sind gleichmächtig.)

- Ist V endlich erzeugt, so ist jede Basis von V endlich und alle Basen haben gleich viele Elemente.
- Ist V nicht endlich erzeugt, so hat jede Basis unendlich viele Elemente.

- Beweis:** a. Nach Voraussetzung besitzt V ein endliches Erzeugendensystem E und nach Proposition 24.23 folgt dann auch, daß V eine endliche Basis $B = (x_1, \dots, x_n)$ besitzt. Dabei können wir o. E. annehmen, daß n die minimale Mächtigkeit einer Basis ist. Sei nun B' eine weitere Basis von V . Angenommen, $|B'| > n$. Dann gibt es eine linear unabhängige Teilfamilie (y_1, \dots, y_{n+1}) in B' , im Widerspruch zum Austauschsatz von Steinitz, der verlangt $n + 1 \leq n$.
- b. Dies ist offensichtlich, da jede Basis V erzeugt.

□

Satz 25.5 rechtfertigt die folgende Definition.

Definition 25.6 (Dimension eines Vektorraums)

Für einen (nicht notwendig endlich erzeugten) K -Vektorraum V definieren wir die *Dimension* von V durch

$$\dim_K(V) := \begin{cases} n, & \text{falls } V \text{ eine Basis mit } n < \infty \text{ Elementen besitzt,} \\ \infty, & \text{falls } V \text{ nicht endlich erzeugt ist.} \end{cases}$$

Ist $\dim_K(V) < \infty$, so nennen wir V einen *endlich-dimensionalen* K -Vektorraum.

Aus Satz 25.5 und Definition 25.6 folgt unmittelbar das folgende Korollar.

Korollar 25.7

Sei $\dim_K(V) = n$, E ein Erzeugendensystem von V und F linear unabhängig in V . Dann gelten

$$|E| \geq n \quad \text{und} \quad |F| \leq n.$$

Zudem gilt die Gleichheit genau dann, wenn die jeweilige Familie eine Basis ist.

Beweis: Nach Proposition 24.23 ist F in einer Basis von V enthalten und E enthält eine Basis von V . Die Ungleichungen folgen dann aus Satz 25.5, und derselbe Satz liefert Gleichheit, wenn die Familien Basen sind. Gilt umgekehrt die Gleichheit, so muß E bzw. F ein minimales Erzeugendensystem bzw. eine maximal linear unabhängige Familie sein und somit nach Proposition 24.18 eine Basis. □

Beispiel 25.8

- a. Es gilt:

$$\dim_K(V) = 0 \Leftrightarrow V = \text{Lin}(\emptyset) \Leftrightarrow V = \{0\}.$$

- b. $\dim_K(K^n) = n$, da die kanonische Basis $E = (e_1, \dots, e_n)$ genau n Elemente enthält.
- c. $\dim_{\mathbb{Q}}(\mathbb{Q}) = \dim_{\mathbb{R}}(\mathbb{R}) = \dim_{\mathbb{C}}(\mathbb{C}) = 1$, aber $\dim_{\mathbb{Q}}(\mathbb{R}) = \infty$ und $\dim_{\mathbb{R}}(\mathbb{C}) = 2$. Für letzteres beachte man, daß die Familie $(1, i)$ eine \mathbb{R} -Basis von \mathbb{C} ist.
- d. Die Dimension des Vektorraums P_n der Polynome vom Grad höchstens n ist $\dim_K(P_n) = n + 1$, da $B = (t^0, t^1, \dots, t^n)$ eine Basis ist.

Satz 25.9 (Karten eines Vektorraums)

Es sei $B = (x_1, \dots, x_n)$ eine Basis von V und $E = (e_1, \dots, e_n)$ die kanonische Basis des K^n . Dann bestimmt B einen Isomorphismus

$$\phi_B : V \rightarrow K^n : x_i \mapsto e_i, \text{ für } i = 1, \dots, n,$$

durch lineare Fortsetzung. Man nennt ϕ_B die Karte von V zur Basis B .

Beweis: Nach Satz 24.14 bestimmen die Zuordnungen

$$x_i \mapsto e_i, \quad i = 1, \dots, n, \quad \text{und} \quad e_i \mapsto x_i, \quad i = 1, \dots, n,$$

zwei lineare Abbildungen $\phi_B : V \rightarrow K^n$ und $\phi^B : K^n \rightarrow V$. Es bleibt zu zeigen, daß

$$\phi_B \circ \phi^B = \text{id}_{K^n} \quad \text{und} \quad \phi^B \circ \phi_B = \text{id}_V.$$

Dazu reicht es wegen Satz 24.14, daß die beiden Seiten jeweils auf einer Basis übereinstimmen, und das tun sie offenbar. \square

Insbesondere haben wir das folgende Korollar gezeigt.

Korollar 25.10

Ist $\dim_K(V) = n$, so gilt $V \cong K^n$.

Korollar 25.11 (Die Dimension ist die einzige Invariante eines Vektorraums.)

Für zwei endlich-dimensionale K -Vektorräume V und W sind gleichwertig:

- a. $V \cong W$.
- b. $\dim_K(V) = \dim_K(W)$.

Beweis: Es seien $n = \dim_K(V)$ und $m = \dim_K(W)$.

Ist $f : V \rightarrow W$ ein Isomorphismus, so überführt er laut Aufgabe 24.33 eine Basis von V , die n Elemente enthält, in eine Basis von W , die m Elemente enthält. Mithin gilt $n = m$.

Ist umgekehrt $n = m$, so gibt es nach Korollar 25.10 Isomorphismen $f : V \rightarrow K^n$ und $g : K^n \rightarrow W$. Dann ist $g \circ f : V \rightarrow W$ ebenfalls ein Isomorphismus. \square

Beispiel 25.12

Die Abbildungen $\sin, \cos \in \mathbb{R}^{\mathbb{R}}$ sind linear unabhängig, da aus

$$\lambda \cdot \sin + \mu \cdot \cos = 0$$

insbesondere

$$0 = \lambda \cdot \sin(0) + \mu \cdot \cos(0) = \mu$$

und

$$0 = \lambda \cdot \sin\left(\frac{\pi}{2}\right) + \mu \cdot \cos\left(\frac{\pi}{2}\right) = \lambda$$

folgt. Aber dann hat der Vektorraum $\text{Lin}(\sin, \cos)$ die Dimension zwei, da (\sin, \cos) eine Basis ist, und deshalb gilt

$$\text{Lin}(\sin, \cos) \cong \mathbb{R}^2.$$

C) Dimensionsformeln

Lemma 25.13

Ist $\dim_{\mathbb{K}}(V) < \infty$ und $U \subseteq V$ ein Unterraum. Dann gelten

$$\dim_{\mathbb{K}}(U) \leq \dim_{\mathbb{K}}(V),$$

und

$$U = V \iff \dim_{\mathbb{K}}(U) = \dim_{\mathbb{K}}(V).$$

Beweis: Ist U ein Unterraum, so kann eine Basis B von U zu einer Basis B' von V ergänzt werden, so daß $\dim_{\mathbb{K}}(U) = |B| \leq |B'| = \dim_{\mathbb{K}}(V)$ gelten muß.

Ist $U = V$, so ist offenbar auch $\dim_{\mathbb{K}}(U) = \dim_{\mathbb{K}}(V)$. Gilt umgekehrt $\dim_{\mathbb{K}}(U) = \dim_{\mathbb{K}}(V)$ und ist $B = (x_1, \dots, x_n)$ eine Basis von U , so können wir sie zu einer Basis B' von V ergänzen nach dem Basisergänzungssatz 24.24. Wegen $B \subseteq B'$ und $|B| = n = |B'|$ folgt dann aber notwendigerweise $B = B'$, und somit $U = \text{Lin}(B) = \text{Lin}(B') = V$. \square

Satz 25.14 (Dimensionsformel für Unterräume)

Ist $\dim_{\mathbb{K}}(V) < \infty$ und sind U und U' Unterräume von V , dann gilt:

$$\dim_{\mathbb{K}}(U + U') = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(U') - \dim_{\mathbb{K}}(U \cap U').$$

Beweis: Wir beweisen mehr, nämlich wie wir geeignete Basen von U , U' und $U \cap U'$ wählen können. Sei $B_{U \cap U'} := (x_1, \dots, x_r)$ eine Basis von $U \cap U'$. Wir ergänzen $B_{U \cap U'}$ zu einer Basis $B_U := (x_1, \dots, x_r, y_1, \dots, y_s)$ von U , und zu einer Basis $B_{U'} := (x_1, \dots, x_r, z_1, \dots, z_t)$ von U' . Das geht nach dem Basisergänzungssatz 24.24.

Behauptung: $B_{U+U'} := (x_1, \dots, x_r, y_1, \dots, y_s, z_1, \dots, z_t)$ ist Basis von $U + U'$.

Dazu zeigen wir zunächst, daß jedes Element von $U + U'$ eine Linearkombination von Elementen aus $B_{U+U'}$ ist. Sei also $x + x' \in U + U'$ mit $x \in U$ und $x' \in U'$. Dann gilt:

$$x = \sum_{i=1}^r \lambda_i x_i + \sum_{j=1}^s \mu_j y_j \quad \text{und} \quad x' = \sum_{i=1}^r \lambda'_i x_i + \sum_{k=1}^t \mu'_k z_k,$$

mit $\lambda_i, \lambda'_i, \mu_j, \mu'_k \in \mathbb{K}$, $i = 1, \dots, r$, $j = 1, \dots, s$, $k = 1, \dots, t$. Daraus folgt:

$$x + x' = \sum_{i=1}^r (\lambda_i + \lambda'_i) x_i + \sum_{j=1}^s \mu_j y_j + \sum_{k=1}^t \mu'_k z_k \in \text{Lin}(B_{U+U'}).$$

Dann müssen wir noch zeigen, daß $B_{U+U'}$ linear unabhängig ist. Sei dazu

$$\sum_{i=1}^r \lambda_i x_i + \sum_{j=1}^s \mu_j y_j + \sum_{k=1}^t \nu_k z_k = 0 \tag{67}$$

eine Linearkombination der Null. Dann ist

$$\underbrace{\sum_{i=1}^r \lambda_i x_i}_{\in \mathbf{U}} + \underbrace{\sum_{j=1}^s \mu_j y_j}_{\in \mathbf{U}'} = \underbrace{\sum_{k=1}^t -\nu_k z_k}_{\in \mathbf{U}'} \in \mathbf{U} \cap \mathbf{U}'.$$

Da $B_{\mathbf{U} \cap \mathbf{U}'}$ eine Basis von $\mathbf{U} \cap \mathbf{U}'$ ist, gibt es also λ'_i , so daß

$$\sum_{i=1}^r \lambda_i x_i + \sum_{j=1}^s \mu_j y_j = \sum_{i=1}^r \lambda'_i x_i = \sum_{i=1}^r \lambda'_i x_i + \sum_{j=1}^s 0 \cdot y_j$$

gilt. Da $B_{\mathbf{U}}$ linear unabhängig ist, ergibt ein Koeffizientenvergleich auf beiden Seiten insbesondere $\mu_j = 0$ für alle $j = 1, \dots, s$. Damit erhalten wir aus (67) dann

$$\sum_{i=1}^r \lambda_i x_i + \sum_{k=1}^t \nu_k z_k = 0,$$

und da $B_{\mathbf{U}'}$ linear unabhängig ist, müssen notwendigerweise alle λ_i und ν_k Null sein. Damit haben wir dann auch gezeigt, daß $B_{\mathbf{U} + \mathbf{U}'}$ linear unabhängig ist.

Aus der Behauptung folgt,

$$\dim_{\mathbb{K}}(\mathbf{U} + \mathbf{U}') = r + s + t = (r + s) + (r + t) - r = \dim_{\mathbb{K}}(\mathbf{U}) + \dim_{\mathbb{K}}(\mathbf{U}') - \dim_{\mathbb{K}}(\mathbf{U} \cap \mathbf{U}').$$

□

Beispiel 25.15

Für die Unterräume $\mathbf{U} = \text{Lin}((1, 0, 0)^t, (1, 1, 1)^t)$ und $\mathbf{U}' = \text{Lin}((1, 1, 1)^t, (0, 0, 1)^t)$ von \mathbb{R}^3 sieht man leicht, daß

$$\mathbf{U} \cap \mathbf{U}' = \text{Lin}((1, 1, 1)^t)$$

ein Vektorraum von Dimension eins ist, während \mathbf{U} und \mathbf{U}' jeweils Dimension zwei haben, da die angegebenen Erzeugendensysteme auch linear unabhängig sind. Mit hin erhalten wir

$$\dim_{\mathbb{R}}(\mathbf{U} + \mathbf{U}') = 2 + 2 - 1 = 3 = \dim_{\mathbb{R}}(\mathbb{R}^3),$$

so daß

$$\mathbf{U} + \mathbf{U}' = \mathbb{R}^3$$

gelten muß. Da zudem

$$\mathbf{U} + \mathbf{U}' = \text{Lin}((1, 0, 0)^t, (1, 1, 1)^t, (0, 0, 1)^t)$$

gilt, sehen wir, daß dieses Erzeugendensystem eine Basis des \mathbb{R}^3 ist.

Korollar 25.16 (Dimensionsformel für Komplemente)

Ist $\dim_{\mathbb{K}}(V) < \infty$, dann sind für Unterräume \mathbf{U} und \mathbf{U}' von V die folgenden Aussagen äquivalent:

- $V = \mathbf{U} \oplus \mathbf{U}'$.
- $V = \mathbf{U} + \mathbf{U}'$ und $\mathbf{U} \cap \mathbf{U}' = \{0\}$.

- c. $V = U + U'$ und $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(U')$.
 d. $U \cap U' = \{0\}$ und $\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(U')$.

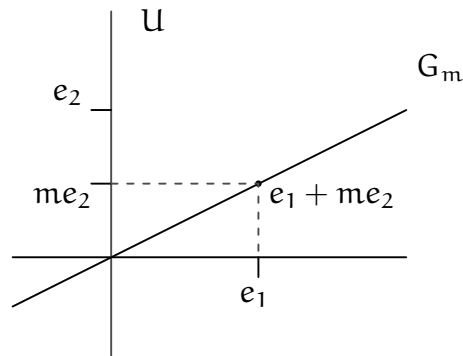
Beweis: Dies ist eine direkte Folgerung aus Lemma 23.17, Lemma 25.13 und Satz 25.14. \square

Beispiel 25.17

Wir erinnern uns an Beispiel 23.36. Dort haben wir in $V = \mathbb{R}^2$ den Unterraum $U = \text{Lin}(e_2)$, die y -Achse, betrachtet und gezeigt, daß jede Ursprungsgerade mit Steigung m

$$G_m := \text{Lin}(e_1 + me_2)$$

ein Komplement von U ist. Dies können wir nun mit weniger Aufwand begründen, denn die beiden Geraden schneiden sich offenbar nur im Ursprung und ihre Dimensionen addieren sich zu zwei.



Korollar 25.18 (Dimensionsformel für Faktorräume)

Es sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum und U ein Unterraum von V . Dann gilt

$$\dim_{\mathbb{K}}(V/U) = \dim_{\mathbb{K}}(V) - \dim_{\mathbb{K}}(U).$$

Beweis: Nach Korollar 24.27 besitzt U ein Komplement U' , und nach Proposition 23.37 gilt $U' \cong V/U$. Aus Korollar 25.11 und Korollar 25.16 folgt dann

$$\dim_{\mathbb{K}}(V) = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(U') = \dim_{\mathbb{K}}(U) + \dim_{\mathbb{K}}(V/U).$$

\square

Bemerkung 25.19 (Basis von V/U)

Es sei V ein \mathbb{K} -Vektorraum und U ein Unterraum von V mit Basis (x_1, \dots, x_r) . Dann sind die folgenden Aussagen für $y_1, \dots, y_s \in V$ gleichwertig:

- $(x_1, \dots, x_r, y_1, \dots, y_s)$ ist eine Basis von V .
- (y_1, \dots, y_s) ist Basis eines Komplementes von U .
- $(\overline{y}_1, \dots, \overline{y}_s)$ ist eine Basis von V/U .

Beweis: Die Äquivalenz von b. und c. folgt aus Proposition 23.37 und Aufgabe 24.33. Die Äquivalenz von a. und b. folgt aus dem Beweis von 24.27 und Korollar 25.16. \square

Satz 25.20 (Dimensionsformel für lineare Abbildungen)

Es sei $f: V \rightarrow W$ eine K -lineare Abbildung und $\dim_K(V) < \infty$. Dann gilt

$$\dim_K(V) = \dim_K(\text{Ker}(f)) + \dim_K(\text{Im}(f)).$$

Beweis: Aus dem Homomorphiesatz 23.34 erhalten wir den Isomorphismus

$$V/\text{Ker}(f) \cong \text{Im}(f),$$

so daß die Formel dann aus Korollar 25.18 folgt. \square

Beispiel 25.21

Die lineare Abbildung $f: \mathbb{R}^2 \rightarrow \mathbb{R}: (x_1, x_2)^t \mapsto x_1 - x_2$ hat den Kern $\text{Lin}((1, 1)^t)$ von Dimension eins und ist surjektiv. Wir erhalten also die Formel

$$\dim_{\mathbb{R}}(\mathbb{R}^2) = 2 = 1 + 1 = \dim_{\mathbb{R}}(\text{Ker}(f)) + \dim_{\mathbb{R}}(\text{Im}(f)).$$

D) Bijektivität linearer Abbildungen

Korollar 25.22 (Injektiv = surjektiv = bijektiv)

Es seien V und W zwei endlich-dimensionale K -Vektorräume gleicher Dimension und $f: V \rightarrow W$ sei K -linear. Dann sind gleichwertig:

- f ist bijektiv,
- f ist injektiv,
- f ist surjektiv.

Beweis: Ist f injektiv, so ist $\text{Ker}(f) = \{0\}$, und wir erhalten aus der Dimensionsformel für lineare Abbildungen 25.20

$$\dim_K(W) = \dim_K(V) = \dim_K(\text{Ker}(f)) + \dim_K(\text{Im}(f)) = \dim_K(\text{Im}(f)).$$

Wegen Lemma 25.13 gilt dann $W = \text{Im}(f)$ und f ist surjektiv.

Ist f surjektiv, so ist $W = \text{Im}(f)$ und wir erhalten aus der Dimensionsformel für lineare Abbildungen 25.20

$$\dim_K(\text{Ker}(f)) = \dim_K(V) - \dim_K(\text{Im}(f)) = \dim_K(V) - \dim_K(W) = 0.$$

Dann ist aber $\text{Ker}(f) = \{0\}$ und somit ist f injektiv. \square

Korollar 25.23 (Invertierbare Matrizen)

Sind $A, B \in \text{Mat}_n(K)$ mit $AB = \mathbf{1}_n$, so gilt auch $BA = \mathbf{1}_n$ und $A \in \text{Gl}_n(K)$.

Beweis: Aus $AB = \mathbb{1}_n$ folgt

$$f_A \circ f_B = f_{AB} = f_{\mathbb{1}_n} = \text{id}_{K^n},$$

so daß f_B injektiv mit Linksinverser f_A ist. Nach Korollar 25.22 ist f_B dann aber schon bijektiv, und die Linksinverse ist die Inverse von f_B . Damit folgt dann auch

$$f_{\mathbb{1}_n} = \text{id}_{K^n} = f_B \circ f_A = f_{BA},$$

und damit $BA = \mathbb{1}_n$. □

Bemerkung 25.24 (Ringe und Moduln)

In den vorhergehenden Kapiteln haben wir stets angemerkt, welche Aussagen auch für Moduln und lineare Abbildungen über kommutativen Ringen mit Eins wahr bleiben. In diesem Kapitel gilt das im wesentlichen für keine Aussage. Die Beweise beruhen sämtlich auf dem Basisergänzungssatz und dem Austauschlemma, und beide Aussagen sind über beliebigen Ringen falsch, ihre Beweise benötigen Division. Allein Korollar 25.23 bleibt wahr, allerdings braucht man einen neuen Beweis.

Aufgaben

Aufgabe 25.25

Es sei V ein K -Vektorraum mit $\dim_K(V) = 5$, und U und U' Unterräume mit $\dim_K(U) = 3$ und $\dim_K(U') = 4$.

- a. Welche Werte kann $\dim_K(U \cap U')$ annehmen?
- b. Gib für jeden der Werte von $\dim_K(U \cap U')$ ein Beispiel (K, V, U, U') an.

Aufgabe 25.26

Finde einen K -Vektorraum V sowie zwei K -lineare Abbildungen $f, g : V \rightarrow V$, so daß folgendes gilt:

- a. f ist injektiv, aber nicht surjektiv.
- b. g ist surjektiv, aber nicht injektiv.

Aufgabe 25.27

Es sei V ein K -Vektorraum mit $1 \leq \dim_K(V) = n < \infty$ und $g \in \text{End}_K(V)$.

Zeige, es gibt eine Zahl $0 \leq k \leq n$ mit

$$\text{Ker}(g^0) \subsetneq \text{Ker}(g^1) \subsetneq \dots \subsetneq \text{Ker}(g^k) = \text{Ker}(g^{k+i})$$

für alle $i \geq 1$.

Aufgabe 25.28

Es sei $B := ((3, 5, 2)^t, (1, 1, -1)^t, (2, 4, 1)^t)$.

- a. Zeige, B ist eine Basis von \mathbb{R}^3 .
- b. Ersetze mit Hilfe des Austauschsatzes von Steinitz zwei Vektoren in B durch die Vektoren $(1, 3, 2)^t$ und $(-2, 1, 2)^t$.

Aufgabe 25.29

Sei V ein K -Vektorraum und $F = (v_1, \dots, v_5)$ eine linear unabhängige Familie in V . Welchen der Vektoren v_1, \dots, v_5 kann man durch $v := v_2 - v_3 + v_4 - v_5$ ersetzen, so dass die daraus resultierende Familie wieder linear unabhängig ist? Begründe Deine Aussage.

Aufgabe 25.30

Sei K ein Körper.

- Begründe, weshalb die Mengen $\mathcal{U} := \{(a_1, \dots, a_n)^t \in K^n \mid a_1 = \dots = a_n\}$ und $\mathcal{U}' := \{(a_1, \dots, a_n)^t \in K^n \mid a_1 + \dots + a_n = 0\}$ Unterräume des K^n sind.
- Bestimme $\dim_K(\mathcal{U})$, $\dim_K(\mathcal{U}')$, $\dim_K(\mathcal{U} \cap \mathcal{U}')$ und $\dim_K(\mathcal{U} + \mathcal{U}')$.

§ 26 Lineare Abbildungen und Matrizen

In diesem Abschnitt sind V und W zwei endlich-dimensionale K -Vektorräume mit Basen $B = (b_1, \dots, b_n)$ und $D = (d_1, \dots, d_m)$. Ferner bezeichnen wir mit $E = (e_1, \dots, e_n)$ die kanonische Basis von K^n und mit $F = (f_1, \dots, f_m)$ die kanonische Basis von K^m .

A) Matrixdarstellung linearer Abbildungen

Satz 26.1 ($\text{Mat}(m \times n, K) \cong \text{Hom}_K(K^n, K^m)$)

Die Vektorräume $\text{Mat}(m \times n, K)$ und $\text{Hom}_K(K^n, K^m)$ sind zueinander isomorph und die Abbildungen Φ und Ψ sind zueinander inverse Isomorphismen:

$$\begin{aligned} \Phi : \text{Mat}(m \times n, K) &\rightarrow \text{Hom}_K(K^n, K^m) \\ A &\mapsto f_A, \end{aligned}$$

und

$$\begin{aligned} \Psi : \text{Hom}_K(K^n, K^m) &\rightarrow \text{Mat}(m \times n, K) \\ f &\mapsto A_f, \end{aligned}$$

wobei $f(e_i)$ der i -te Spaltenvektor der Matrix A_f ist.

Beweis: Nach Proposition 23.23 ist das Bild von Φ in $\text{Hom}_K(K^n, K^m)$ enthalten, und aus Lemma 22.8 wissen wir, daß die Abbildung Φ K -linear ist. Ferner folgt aus Korollar 24.17, daß es zu jeder linearen Abbildung $f: K^n \rightarrow K^m$ eine *eindeutig bestimmte* $m \times n$ -Matrix A mit $f = f_A$ gibt, d. h. Φ ist bijektiv und damit ein Isomorphismus.

Außerdem gilt für die Umkehrabbildung Φ^{-1} , ebenfalls nach Korollar 24.17, daß der i -te Spaltenvektor von $\Phi^{-1}(f)$ für $f \in \text{Hom}_K(K^n, K^m)$ gerade $f(e_i)$ ist. Also ist $\Phi^{-1} = \Psi$, und damit ist Ψ linear und ein Isomorphismus nach Lemma 23.22. \square

Bemerkung 26.2 (K -Algebren)

a. Es seien $f \in \text{Hom}_K(K^n, K^m)$ und $g \in \text{Hom}_K(K^m, K^p)$. Dann gilt:

$$A_{g \circ f} = A_g \circ A_f.$$

Beachte dazu, daß wegen $f_{A_f} = f$ und $f_{A_g} = g$ für $i = 1, \dots, n$ gilt:

$$A_{g \circ f} e_i = (g \circ f)(e_i) = g(f(e_i)) = g(A_f e_i) = A_g(A_f e_i) = (A_g \circ A_f) e_i.$$

Das heißt aber, daß die n Spalten der beiden Matrizen übereinstimmen.

b. Ein K -Vektorraum $(B, +, \cdot)$, auf dem zusätzlich eine Multiplikation

$$\circ : B \times B \rightarrow B : (x, y) \mapsto x \circ y$$

definiert ist, so daß $(\mathbf{B}, +, \circ)$ ein (nicht unbedingt kommutativer) Ring mit Eins $1_{\mathbf{B}}$ ist, heißt eine *K-Algebra*, falls die Skalarmultiplikation mit der Ringmultiplikation verträglich ist, d. h. für $\lambda \in \mathbf{K}$ und $\mathbf{x}, \mathbf{y} \in \mathbf{B}$ gelten:

$$\lambda \cdot (\mathbf{x} \circ \mathbf{y}) = (\lambda \cdot \mathbf{x}) \circ \mathbf{y} = \mathbf{x} \circ (\lambda \cdot \mathbf{y}).$$

Ein *K-Algebrenisomorphismus* ist eine bijektive Abbildung $\varphi : \mathbf{A} \rightarrow \mathbf{B}$ zwischen zwei K-Algebren \mathbf{A} und \mathbf{B} , die mit allen drei Operationen verträglich ist und die $1_{\mathbf{A}}$ auf $1_{\mathbf{B}}$ abbildet.

- c. $(\text{End}_{\mathbf{K}}(\mathbf{V}), +, \cdot, \circ)$ ist eine K-Algebra.
- d. $(\text{Mat}(\mathbf{n}, \mathbf{K}), +, \cdot, \circ)$ ist ebenfalls eine K-Algebra.
- e. Die in Satz 26.1 definierten Abbildungen Φ und Ψ sind für $\mathbf{m} = \mathbf{n}$ zueinander inverse K-Algebrenisomorphismen der K-Algebren $\text{Mat}(\mathbf{n}, \mathbf{K})$ und $\text{End}_{\mathbf{K}}(\mathbf{K}^{\mathbf{n}})$.

Wir kommen jetzt zum Zusammenhang zwischen Matrizen und linearen Abbildungen zwischen Vektorräumen mit Basen.

Definition 26.3 (Koordinaten- und Matrixdarstellung)

Es sei $f \in \text{Hom}_{\mathbf{K}}(\mathbf{V}, \mathbf{W})$ und $\phi_{\mathbf{B}}$ und $\phi_{\mathbf{D}}$ seien die Karten von \mathbf{V} zur Basis \mathbf{B} bzw. von \mathbf{W} zur Basis \mathbf{D} .

- a. Die Abbildung

$$\phi_{\mathbf{D}} \circ f \circ \phi_{\mathbf{B}}^{-1} : \mathbf{K}^{\mathbf{n}} \rightarrow \mathbf{K}^{\mathbf{m}}$$

heißt die *Koordinatendarstellung* von f bezüglich der Basen \mathbf{B} und \mathbf{D} oder kurz bezüglich (\mathbf{B}, \mathbf{D}) .

- b. Die zur Abbildung $\phi_{\mathbf{D}} \circ f \circ \phi_{\mathbf{B}}^{-1}$ gehörende Matrix

$$\mathbf{M}_{\mathbf{D}}^{\mathbf{B}}(f) := \Psi(\phi_{\mathbf{D}} \circ f \circ \phi_{\mathbf{B}}^{-1}) = \mathbf{A}_{\phi_{\mathbf{D}} \circ f \circ \phi_{\mathbf{B}}^{-1}}$$

nennen wir die *Matrixdarstellung* von f bezüglich (\mathbf{B}, \mathbf{D}) .

- c. Für einen Vektor $\mathbf{x} = \sum_{i=1}^{\mathbf{n}} \lambda_i \mathbf{b}_i$ nennt man den Vektor

$$\mathbf{M}_{\mathbf{B}}(\mathbf{x}) := \phi_{\mathbf{B}}(\mathbf{x}) = (\lambda_1, \dots, \lambda_{\mathbf{n}})^t \in \mathbf{K}^{\mathbf{n}}$$

den *Koordinatenvektor* oder die *Koordinaten* von \mathbf{x} bezüglich \mathbf{B} .

Bemerkung 26.4 (Koordinaten- und Matrixdarstellung)

- a. Die Koordinatendarstellung von f schreibt man am besten in Form eines *kommutativen Diagramms*:

$$\begin{array}{ccc} \mathbf{V} & \xrightarrow{f} & \mathbf{W} \\ \phi_{\mathbf{B}} \downarrow \cong & \circlearrowleft & \cong \downarrow \phi_{\mathbf{D}} \\ \mathbf{K}^{\mathbf{n}} & \xrightarrow{\phi_{\mathbf{D}} \circ f \circ \phi_{\mathbf{B}}^{-1}} & \mathbf{K}^{\mathbf{m}} \end{array}$$

Die *Kommutativität eines Diagramms* von Abbildungen bedeutet, daß die Komposition von Abbildungen mit gleichem Anfangs- und Endpunkt längs Pfeilen (in Pfeilrichtung) gleich sind. Hier bedeutet dies einfach

$$\phi_D \circ f = (\phi_D \circ f \circ \phi_B^{-1}) \circ \phi_B.$$

- b. Wie wir in Satz 25.9 gesehen haben, ist die Wahl eines Isomorphismus $V \rightarrow K^n$ äquivalent zur Wahl einer Basis von V . Die Wahl einer Basis B von V bestimmt eindeutig einen Isomorphismus $\phi_B : V \rightarrow K^n$, mit dessen Hilfe wir V und K^n identifizieren können.
- c. Durch Wahl von Basen B in V und D in W können wir lineare Abbildungen $f : V \rightarrow W$ mit ihrer Koordinatendarstellung $\phi_D \circ f \circ \phi_B^{-1} : K^n \rightarrow K^m$ identifizieren und damit auch mit der zugehörigen Matrix $M_D^B(f)$, vermittels Satz 26.1. Wir werden häufig nicht zwischen der Koordinatendarstellung $\phi_D \circ f \circ \phi_B^{-1}$ und der Matrixdarstellung $M_D^B(f)$ unterscheiden.
- d. Für einen Vektor $x = \sum_{i=1}^n \lambda_i b_i$ gilt offenbar

$$M_D(f(x)) = M_D^B(f) \circ M_B(x),$$

d. h. der Koordinatenvektor $M_D(f(x))$ von $f(x)$ bezüglich der Basis D ist das Matrixprodukt der Matrixdarstellung $M_D^B(f)$ von f bezüglich B und D mit dem Koordinatenvektor $M_B(x)$ von x bezüglich B . Insbesondere gilt:

Die i -te Spalte von $M_D^B(f)$ ist der Koordinatenvektor von $f(b_i)$ bez. D .

Beispiel 26.5

- a. Ist $V = K^n$ und $W = K^m$, dann gilt $\phi_E = \text{id}_{K^n}$ und $\phi_F = \text{id}_{K^m}$. Mithin folgt $M_F^E(f) = A_f$ für $f \in \text{Hom}_K(K^n, K^m)$.
- b. Sei $V = K^2$ mit Basis $B = (b_1, b_2) = ((1, 2)^t, (1, 1)^t)$ und $W = K^3$ mit Basis $D = (d_1, d_2, d_3) = ((1, 1, 0)^t, (0, 1, 1)^t, (0, 0, 1)^t)$, und sei $f : V \rightarrow W$ die lineare Abbildung, die definiert wird durch

$$\begin{aligned} b_1 &\mapsto 3d_1 - 4d_2 + 6d_3, \\ b_2 &\mapsto 3d_1 - 3d_2 + 4d_3. \end{aligned}$$

Dann gilt:

$$M_D^B(f) = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix}.$$

Unser Ziel ist es nun, in obigem Beispiel aus $M_D^B(f)$ die Matrix $A_f = M_F^E(f)$ zu bestimmen. Dazu führen wir folgende allgemeine Begriffsbildung ein.

Definition 26.6 (Basiswechsel)

Sind $B = (b_1, \dots, b_n)$ und $B' = (b'_1, \dots, b'_n)$ zwei Basen von V , so heißen der Isomorphismus

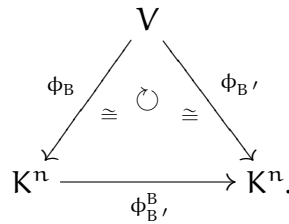
$$\phi_B^B := \phi_{B'} \circ \phi_B^{-1} : K^n \rightarrow K^n$$

und die zugehörige Matrixdarstellung

$$T_{B'}^B := M_E^E(\phi_{B'} \circ \phi_B^{-1}) = A_{\phi_{B'} \circ \phi_B^{-1}}$$

Basiswechsel oder *Koordinatentransformation* bezüglich (B, B') .

Wir haben also folgendes kommutatives Diagramm:



Bemerkung 26.7 (Basiswechsel als Matrixdarstellung)

Offensichtlich ist der Basiswechsel ein Spezialfall der Koordinatendarstellung, denn es gilt

$$\phi_B^B = \phi_{B'} \circ \phi_B^{-1} = \phi_{B'} \circ \text{id}_V \circ \phi_B^{-1},$$

und entsprechend für die Matrixdarstellung

$$T_{B'}^B = M_{B'}^B(\text{id}_V).$$

Damit gilt ferner:

Die i -te Spalte von $T_{B'}^B$ ist der Koordinatenvektor $M_{B'}(b_i)$ von b_i bez. B' .

Außerdem folgt unmittelbar aus der Definition von $T_{B'}^B$, daß $T_{B'}^B$ invertierbar ist mit

$$(T_{B'}^B)^{-1} = T_B^{B'}.$$

Satz 26.8 (Basiswechsel)

Seien B und B' Basen von V , D und D' Basen von W und $f \in \text{Hom}_K(V, W)$.

Dann gilt:

$$M_{D'}^{B'}(f) = T_{D'}^D \circ M_D^B(f) \circ T_B^{B'}.$$

Beweis: Es gilt:

$$(\phi_{D'} \circ \phi_D^{-1}) \circ (\phi_D \circ f \circ \phi_B^{-1}) \circ (\phi_B \circ \phi_{B'}^{-1}) = \phi_{D'} \circ f \circ \phi_{B'}^{-1}.$$

Mittels Bemerkung 26.2 gilt dann aber:

$$\begin{aligned} T_{D'}^D \circ M_D^B(f) \circ T_B^{B'} &= A_{\phi_{D'} \circ \phi_D^{-1}} \circ A_{\phi_D \circ f \circ \phi_B^{-1}} \circ A_{\phi_B \circ \phi_{B'}^{-1}} \\ &= A_{(\phi_{D'} \circ \phi_D^{-1}) \circ (\phi_D \circ f \circ \phi_B^{-1}) \circ (\phi_B \circ \phi_{B'}^{-1})} = A_{\phi_{D'} \circ f \circ \phi_{B'}^{-1}} = M_{D'}^{B'}(f). \end{aligned}$$

□

Korollar 26.9 (Basiswechsel bei Endomorphismen)

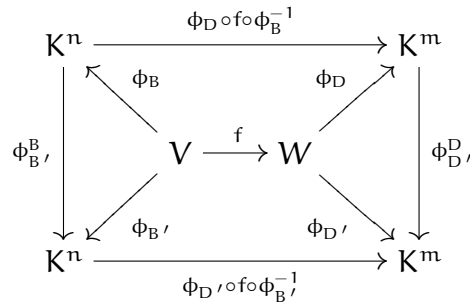
Sind B und B' Basen von V , ist $T = T_B^{B'}$ und ist $f \in \text{End}_K(V)$, so gilt

$$M_{B'}^{B'}(f) = T^{-1} \circ M_B^B(f) \circ T.$$

Beweis: Dies folgt aus Satz 26.8, weil $(T_B^{B'})^{-1} = T_{B'}^B$. □

Bemerkung 26.10

Man kann den Beweis von Satz 26.8 als *Diagrammjagd* in folgendem kommutativen Diagramm bezeichnen:



Dabei impliziert die Kommutativität der inneren Diagramme die des äußeren. Die Aussage des Satzes zusammen mit $(T_B^{B'})^{-1} = T_{B'}^B$ ist dann (unter Berücksichtigung von Bemerkung 26.2) gerade die Kommutativität des Diagramms.

Beispiel 26.11

Wir wollen nun für die Abbildung in Beispiel 26.5 b. die Matrixdarstellung $M_F^E(f)$ bezüglich der kanonischen Basen berechnen. Nach Satz 26.8 gilt:

$$M_F^E(f) = T_F^D \circ M_D^B(f) \circ T_B^E.$$

Um T_F^D auszurechnen, müssen wir d_1, d_2 und d_3 in der kanonischen Basis ausdrücken und die Koeffizienten als Spaltenvektoren in die Matrix T_F^D übertragen:

$$T_F^D = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Um T_B^E zu ermitteln, müßten wir die Einheitsvektoren e_1 und e_2 als Linearkombination der Basis B darstellen, was auf das Lösen zweier Gleichungssysteme hinaus liefe. Stattdessen können wir aber auch T_E^B bestimmen und anschließend invertieren, was sich im Falle einer (2×2) -Matrix anbietet, da das Invertieren sehr einfach ist (vgl. Aufgabe 22.14),

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

sofern die Matrix invertierbar ist.

Analog zum Fall von T_F^D erhalten wir

$$T_E^B = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix},$$

und somit

$$T_B^E = (T_E^B)^{-1} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}.$$

Also gilt:

$$M_F^E(f) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \circ \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \circ \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Wir können jetzt für Vektorräume mit Basen, ganz analog zu Satz 26.1 für K^n , den Vektorraum $\text{Hom}_K(V, W)$ mit dem Vektorraum $\text{Mat}(m \times n, K)$ identifizieren.

Satz 26.12 (Matrixdarstellung linearer Abbildungen)

Die Abbildung

$$M_D^B : \text{Hom}_K(V, W) \longrightarrow \text{Mat}(m \times n, K) : f \mapsto M_D^B(f)$$

ist ein Isomorphismus von K -Vektorräumen.

Beweis: Wegen Aufgabe 23.44 b. ist die Abbildung

$$\Psi_D^B : \text{Hom}_K(V, W) \rightarrow \text{Hom}_K(K^n, K^m) : f \mapsto \phi_D \circ f \circ \phi_B^{-1}$$

eine K -lineare Abbildung. Ferner ist die Abbildung

$$\text{Hom}_K(K^n, K^m) \rightarrow \text{Hom}_K(V, W) : g \mapsto \phi_D^{-1} \circ g \circ \phi_B$$

offensichtlich die inverse Abbildung, so daß Ψ_D^B ein Isomorphismus ist. Ferner gilt, mit den Notationen aus Satz 26.1,

$$M_D^B = \Psi \circ \Psi_D^B$$

so daß M_D^B als Komposition von zwei Isomorphismen ein Isomorphismus ist. \square

Bemerkung 26.13

Für $V = W$ ist in Satz 26.12

$$M_B^B : \text{End}_K(V) \rightarrow \text{Mat}(n, K)$$

ein K -Algebrenisomorphismus.

B) f -invariante Unterräume

Definition 26.14 (f -invarianter Unterraum)

Ist $f : V \rightarrow V$ eine K -lineare Abbildung und $U \leq V$ ein Unterraum mit $f(U) \subseteq U$, so nennen wir U einen f -invarianten Unterraum.

Bemerkung 26.15 (f -invariante Unterräume)

Aus Aufgabe 23.46 wissen wir, daß jeder f -invariante Unterraum U zwei K -lineare Abbildungen

$$f_U : U \rightarrow U : x \mapsto f(x)$$

und

$$f_{V/U} : V/U \rightarrow V/U : \bar{x} \mapsto \overline{f(x)}$$

induziert. Mit Hilfe dieser Abbildungen erhalten wir eine vorteilhafte Blockgestalt bei der Matrixdarstellung.

Proposition 26.16 (Matrixdarstellung bei in Blockform)

Es sei $f : V \rightarrow V$ eine K -lineare Abbildung und $U \leq V$ ein f -invarianter Unterraum. Ferner sei $B' = (x_1, \dots, x_k)$ eine Basis von U und $B = (x_1, \dots, x_n)$ eine Ergänzung von B' zu einer Basis von V .

Dann ist $B'' = (\overline{x_{k+1}}, \dots, \overline{x_n})$ eine Basis von V/U und es gilt

$$M_B^B(f) = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array} \right),$$

wobei $0 \in \text{Mat}((n-k) \times k, K)$ die Nullmatrix ist und $* \in \text{Mat}(k \times (n-k), K)$ eine geeignete Matrix ist.

Beweis: Daß B'' eine Basis von V/U ist, wissen wir bereits aus Bemerkung 25.19. Sei nun $M_B^B(f) = (a_{ij}) \in \text{Mat}_n(K)$, $M_{B'}^{B'}(f_U) = (b_{ij}) \in \text{Mat}_k(K)$ und $M_{B''}^{B''}(f_{V/U}) = (c_{ij}) \in \text{Mat}_{n-k}(K)$.

Für $j = 1, \dots, k$ gilt dann

$$\sum_{i=1}^k b_{ij} \cdot x_i + \sum_{i=k+1}^n 0 \cdot x_i = \sum_{i=1}^k b_{ij} \cdot x_i = f_U(x_j) = f(x_j) = \sum_{i=1}^n a_{ij} \cdot x_i.$$

Da die Darstellung als Linearkombination von Basisvektoren eindeutig ist, folgt somit

$$(a_{ij})_{i,j=1,\dots,k} = (b_{ij})_{i,j=1,\dots,k} = M_{B'}^{B'}(f_U)$$

und

$$(a_{ij})_{i=k+1,\dots,n, j=1,\dots,k} = 0 \in \text{Mat}((n-k) \times k, K).$$

Für $j = k+1, \dots, n$ erhalten wir analog

$$\sum_{i=k+1}^n c_{ij} \overline{x_i} = f_{V/U}(\overline{x_j}) = \overline{f(x_j)} = \overline{\sum_{i=1}^n a_{ij} \cdot x_i} = \sum_{i=1}^n a_{ij} \cdot \overline{x_i} = \sum_{i=k+1}^n a_{ij} \cdot \overline{x_i},$$

da $\bar{x}_i = \bar{0}$ für $i = 1, \dots, k$. Aus der Eindeutigkeit der Darstellung als Linearkombination von Basisvektoren erhalten wir also wieder

$$(\mathbf{a}_{ij})_{i,j=k+1,\dots,n} = (\mathbf{c}_{ij})_{i,j=k+1,\dots,n} = M_{B''}^{B''}(f_{V/U}).$$

Insgesamt haben wir damit die Behauptung

$$M_B^B(f) = (\mathbf{a}_{ij})_{i,j=1,\dots,n} = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array} \right)$$

gezeigt. □

Proposition 26.17 (Matrixdarstellung in Blockdiagonalgestalt)

Sei $f : V \rightarrow V$ eine K -lineare Abbildung und $V = U_1 \oplus \dots \oplus U_k$ die direkte Summe nicht-trivialer f -invarianter Unterräume U_1, \dots, U_k mit Basen B_1, \dots, B_k .

Dann ist $B = B_1 \cup \dots \cup B_k$ eine Basis von V und es gilt

$$M_B^B(f) = \left(\begin{array}{c|ccc} M_{B_1}^{B_1}(f_{U_1}) & 0 & \dots & 0 \\ \hline 0 & M_{B_2}^{B_2}(f_{U_2}) & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & M_{B_k}^{B_k}(f_{U_k}) \end{array} \right).$$

Beweis: Daß B eine Basis von V ist, wissen wir aus Aufgabe 24.34.

Es sei $B = (x_1, \dots, x_n)$ und $M_B^B(f) = (\mathbf{a}_{ij}) \in \text{Mat}_n(K)$. Halten wir ein $1 \leq l \leq k$ fest, so ist $B_l = (x_r, \dots, x_s)$ für geeignete $1 \leq r < s \leq n$ und $M_{B_l}^{B_l}(f_{U_l}) = (\mathbf{b}_{ij})_{i,j=r,\dots,s} \in \text{Mat}_{s-r+1}(K)$. Wie im Beweis von Proposition 26.16 sehen wir für $j = r, \dots, s$ dann

$$\sum_{i=1}^{r-1} 0 \cdot x_i + \sum_{i=r}^s \mathbf{b}_{ij} \cdot x_i + \sum_{i=s+1}^n 0 \cdot x_i = \sum_{i=1}^k \mathbf{b}_{ij} \cdot x_i = f_{U_l}(x_j) = f(x_j) = \sum_{i=1}^n \mathbf{a}_{ij} \cdot x_i,$$

woraus wieder

$$(\mathbf{a}_{ij})_{i,j=r,\dots,s} = (\mathbf{b}_{ij})_{i,j=r,\dots,s} = M_{B_l}^{B_l}(f_{U_l})$$

sowie $\mathbf{a}_{ij} = 0$ für alle $i < r$ und $i > s$ folgt. Damit ist die Behauptung gezeigt. □

C) Äquivalenz von Matrizen und der Rang

Die Koordinatentransformationen in Vektorräumen mit Basen führen auf folgende Äquivalenzbegriffe für Matrizen.

Definition 26.18 (Äquivalenz von Matrizen)

Zwei Matrizen A und A' in $\text{Mat}(m \times n, K)$ heißen *äquivalent*, falls es invertierbare Matrizen $S \in \text{Gl}_m(K)$ und $T \in \text{Gl}_n(K)$ gibt mit

$$A' = S \circ A \circ T.$$

Bemerkung 26.19 (Äquivalenz von Matrizen als Äquivalenzrelation)

Die Äquivalenz von Matrizen eine Äquivalenzrelation auf $\text{Mat}(\mathfrak{m} \times \mathfrak{n}, \mathbb{K})$ ist.

Beispiel 26.20

Die Matrizen

$$A = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \quad \text{und} \quad A' = \begin{pmatrix} 3 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}$$

sind äquivalent, da wir in Beispiel 26.11

$$A' = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \circ \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \circ \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}$$

gezeigt haben.

Definition 26.21 (Rang)

- a. Ist $f \in \text{Hom}_{\mathbb{K}}(V, W)$, so definieren wir den *Rang* von f als

$$\text{rang}(f) := \dim_{\mathbb{K}}(\text{Im}(f)).$$

- b. Ferner definieren wir für eine Matrix $A \in \text{Mat}(\mathfrak{m} \times \mathfrak{n}, \mathbb{K})$ den *Rang* von A durch:

$$\text{rang}(A) := \text{rang}(f_A).$$

Bemerkung 26.22 (Rangabschätzung)

- a. Für $f \in \text{Hom}_{\mathbb{K}}(V, W)$ gilt wegen der Dimensionsformel für lineare Abbildungen

$$\text{rang}(f) = \dim_{\mathbb{K}}(V) - \dim_{\mathbb{K}}(\text{Ker}(f)) \leq \dim_{\mathbb{K}}(V)$$

und da $\text{Im}(f)$ ein Unterraum von W ist, gilt auch $\text{rang}(f) \leq \dim_{\mathbb{K}}(W)$.

- b. Man beachte, daß das Bild von f_A von den Spalten von A erzeugt wird, so daß der Rang von A die Anzahl linear unabhängiger Spalten von A ist.

Zudem folgt aus a. für $A \in \text{Mat}(\mathfrak{m} \times \mathfrak{n}, \mathbb{K})$

$$\text{rang}(A) \leq \min\{\mathfrak{m}, \mathfrak{n}\}.$$

Beispiel 26.23

Da der Rang einer Matrix die Anzahl linear unabhängiger Spalten ist, gelten

$$\text{rang} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0, \quad \text{rang} \begin{pmatrix} 1 & 0 & 2 \\ 1 & 0 & 2 \end{pmatrix} = 1, \quad \text{rang} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix} = 2.$$

Satz 26.24 (Invertierbare Matrizen haben vollen Rang.)

Eine Matrix $A \in \text{Mat}_{\mathfrak{n}}(\mathbb{K})$ ist genau dann invertierbar, wenn $\text{rang}(A) = \mathfrak{n}$.

Inbesondere sind die Spalten einer invertierbaren $\mathfrak{n} \times \mathfrak{n}$ -Matrix eine Basis des $\mathbb{K}^{\mathfrak{n}}$.

Beweis: Nach Aufgabe 24.37 ist A genau dann invertierbar, wenn $f_A : K^n \rightarrow K^n$ bijektiv ist. Wegen Korollar 25.22 ist dies genau dann der Fall, wenn f_A surjektiv ist, d.h. wenn $\text{Im}(f_A) = K^n$. Wegen $\text{Im}(f_A) \subseteq K^n$ und Lemma 25.13 ist dies wiederum genau dann der Fall, wenn

$$n = \dim_K(\text{Im}(f_A)) = \text{rang}(A).$$

Also ist A genau dann invertierbar, wenn $\text{rang}(A) = n$ gilt. In diesem Fall sind die Spalten von A linear unabhängig in K^n und bilden mithin eine Basis des K^n . \square

Bemerkung 26.25

Sind die Matrizen $A, A' \in \text{Mat}(m \times n, K)$ äquivalent, so gibt es Basen B von K^n und D von K^m , so daß

$$A' = M_D^B(f_A),$$

d.h. A und A' sind Matrixdarstellungen derselben linearen Abbildung f_A bezüglich verschiedener Basen!

Dazu betrachten wir einfach die Matrizen $S \in \text{Gl}_m(K)$ und $T \in \text{Gl}_n(K)$ mit $A' = S \circ A \circ T$. Die Spalten von S^{-1} sind linear unabhängig und bilden eine Basis D von K^m nach Satz 26.24. Ebenso bilden die Spalten von T eine Basis B von K^n . Für diese Basen gilt aber nach Konstruktion

$$T_F^D = S^{-1} \quad \text{und} \quad T_E^B = T.$$

Insgesamt erhalten wir also

$$M_D^B(f_A) = T_F^D \circ M_F^E(f_A) \circ T_E^B = S \circ A \circ T = A'.$$

Wir werden nun zeigen, daß der Rang der Matrixdarstellung einer linearen Abbildung *nicht* von der Wahl der Basen abhängt, bezüglich derer man die Matrixdarstellung bildet.

Proposition 26.26 (Rang einer Matrixdarstellung)

Ist $f \in \text{Hom}_K(V, W)$, so gilt

$$\text{rang}(f) = \text{rang}(M_D^B(f)).$$

Insbesondere haben äquivalente Matrizen den gleichen Rang.

Beweis: Wir betrachten die Karten $\phi_B : V \rightarrow K^n$ und $\phi_D : W \rightarrow K^m$. Da ϕ_D ein Isomorphismus ist, erhält ϕ_D die Dimension eines Vektorraumes und es folgt

$$\begin{aligned} \text{rang}(M_D^B(f)) &= \text{rang}(f_{M_D^B(f)}) = \text{rang}(\phi_D \circ f \circ \phi_B^{-1}) \\ &= \dim_K(\phi_D(f(\phi_B^{-1}(K^n)))) = \dim_K(\phi_D(f(V))) \\ &= \dim_K(f(V)) = \text{rang}(f) \end{aligned}$$

Sind A und A' äquivalent, so sind sie nach Bemerkung 26.25 Matrixdarstellungen der gleichen Abbildung f_A bezüglich verschiedener Basen und wir haben gerade gesehen, daß der Rang der Matrixdarstellung nicht von der Wahl der Basen abhängt. \square

Beispiel 26.27

Die Abbildung f in Beispiel 26.5 hat den Rang zwei, wie man an ihrer Matrixdarstellung sieht:

$$M_D^B(f) = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix}.$$

Satz 26.28 (Normalform einer Matrixdarstellung bezüglich Äquivalenz)

Es sei $f \in \text{Hom}_K(V, W)$ mit $\text{rang}(f) = r$. Dann gibt es Basen B von V und D von W mit

$$M_D^B(f) = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array} \right),$$

wobei hier 0 jeweils die Nullmatrix der entsprechenden Größe meint.¹ Wir bezeichnen die rechte Seite der obigen Gleichung auch als die Normalform von f bezüglich Äquivalenz.

Beweis: Wähle vermöge Lemma 24.27 ein Komplement U von $\text{Ker}(f)$. Nach Satz 23.34 und Lemma 23.37 ist die folgende Abbildung ein Isomorphismus

$$f|_U : U \rightarrow \text{Im}(f) : x \mapsto f(x).$$

Wähle eine Basis (d_1, \dots, d_r) von $\text{Im}(f)$. Dann ist (b_1, \dots, b_r) mit $b_i := (f|_U)^{-1}(d_i)$ eine Basis von U , nach Aufgabe 24.33. Wähle nun eine Basis (b_{r+1}, \dots, b_n) von $\text{Ker}(f)$, dann ist $B = (b_1, \dots, b_n)$ eine Basis von $V = U \oplus \text{Ker}(f)$. Ergänze ferner (d_1, \dots, d_r) zu einer Basis $D = (d_1, \dots, d_m)$ von W vermöge Korollar 24.24. Dann:

$$f(b_i) = \begin{cases} d_i, & i = 1, \dots, r, \\ 0, & i = r+1, \dots, n. \end{cases}$$

Also hat $M_D^B(f)$ die gewünschte Gestalt. □

Korollar 26.29 (Normalform einer Matrix bezüglich Äquivalenz)

Zu $A \in \text{Mat}(m \times n, K)$ mit $r = \text{rang}(A)$ existieren Matrizen $S \in \text{Gl}_m(K)$ und $T \in \text{Gl}_n(K)$ mit

$$S \circ A \circ T = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array} \right). \quad (68)$$

Die rechte Seite heißt die Normalform von A bezüglich Äquivalenz.

Beweis: Anwendung des Satzes 26.28 auf $f_A : K^n \rightarrow K^m$ liefert B und D von K^n bzw. K^m mit

$$\left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array} \right) = M_D^B(f_A) = T_D^F \circ M_F^E(f_A) \circ T_E^B = T_D^F \circ A \circ T_E^B.$$

¹Man bezeichnet die vier Matrizen $\mathbb{1}_r \in \text{Mat}_r(K)$, $0 \in \text{Mat}(r \times (n-r), K)$, $0 \in \text{Mat}((m-r) \times r, K)$ und $0 \in \text{Mat}((m-r) \times (n-r), K)$ auch als *Blöcke* von $M_D^B(f)$ und die Matrix $M_D^B(f)$ als eine *Blockmatrix*.

Die Behauptung folgt also, da $S := T_D^F$ und $T := T_E^B$ invertierbar sind. □

Beispiel 26.30

Die folgende Matrix $A \in \text{Mat}(3 \times 4, \mathbb{R})$ hat Rang 2 und hat somit die Matrix B als Normalform:

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 2 & 4 & 6 & 8 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Bemerkung 26.31 (Normalform als guter Repräsentant einer Äquivalenzklasse)
 Aus Korollar 26.29 folgt, daß zwei Matrizen genau dann äquivalent sind, wenn sie den gleichen Rang haben. $\text{Mat}(m \times n, K)$ zerfällt also in $\min\{m, n\} + 1$ Äquivalenzklassen und jede Äquivalenzklasse ist durch den Rang einer ihrer Matrizen eindeutig bestimmt. Darüber hinaus besitzt jede Äquivalenzklasse $[A]$, $A \in \text{Mat}(m \times n, K)$, einen besonders schönen Repräsentanten, nämlich

$$\left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array} \right).$$

Diesen Repräsentanten der Äquivalenzklasse von A nennt man die *Normalform von A bezüglich Äquivalenz*.

Korollar 26.32 (Zeilen- und Spaltenrang)

- a. $A \in \text{Mat}_n(K)$ ist genau dann invertierbar, wenn A^t invertierbar ist.
 In dem Fall gilt

$$(A^t)^{-1} = (A^{-1})^t.$$

- b. Für eine Matrix $A \in \text{Mat}(m \times n, K)$ gilt

$$\text{rang}(A) = \text{rang}(A^t).$$

Insbesondere ist die Anzahl linear unabhängiger Spalten in A gleich der Anzahl linear unabhängiger Zeilen!

Beweis: a. Es sei A invertierbar und $B = A^{-1}$. Dann gilt

$$B^t A^t = (AB)^t = \mathbb{1}_n^t = \mathbb{1}_n,$$

so daß A^t nach Korollar 25.23 invertierbar ist mit Inverser $(A^t)^{-1} = (A^{-1})^t$. Ist umgekehrt A^t invertierbar, so ist nach dem eben gezeigten auch $A = (A^t)^t$ invertierbar.

- b. Nach Korollar 26.29 finden wir invertierbare Matrizen $S \in \text{Gl}_m(K)$ und $T \in \text{Gl}_n(K)$, so daß

$$S \circ A \circ T = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array} \right)$$

in Normalform mit $r = \text{rang}(A)$ gegeben ist. Dann ist aber

$$T^t \circ A^t \circ S^t = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array} \right)^t \in \text{Mat}(n \times m, K)$$

ebenfalls eine Matrix in Normalform. Es gilt also

$$\text{rang}(T^t \circ A^t \circ S^t) = r$$

und wegen Teil a. ist die Matrix $T^t \circ A^t \circ S^t$ äquivalent zu A^t , so daß sie nach Proposition 26.26 den gleichen Rang hat wie A^t .

□

Beispiel 26.33

Die Matrix

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

hat offenbar den Rang 3, da die ersten drei Spalten schon linear unabhängig sind. Mithin hat auch die transponierte Matrix

$$A^t = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \\ 3 & 2 & 3 \\ 5 & 2 & 1 \end{pmatrix}$$

den Rang 3, d.h. die drei Spalten sind linear unabhängig.

Bemerkung 26.34 (Ringe und Moduln)

Die Identifikation von linearen Abbildungen und Matrizen funktioniert auch für Moduln über Ringen, wenn sie endliche Basen besitzen. Die Beweise ändern sich nicht. Man erhält also die Aussagen der Sätze und Bemerkungen 26.1, 26.2, 26.8 und 26.9 ohne Änderung für Moduln, die endliche Basen besitzen — die zugehörigen Definitionen kann man ebenfalls ohne Änderung übernehmen. Die weiteren Aussagen des Abschnitts zur Äquivalenz von Matrizen und zu deren Rang gelten in dieser Form nicht allgemein für lineare Abbildungen von Moduln. Selbst wenn zwei Moduln V und W Basen besitzen, muß das Bild einer linearen Abbildung von V nach W keine Basis haben, so daß man den Rang der Abbildung dann gar nicht definieren kann.

Aufgaben

Aufgabe 26.35 (Zyklische Unterräume)

Es sei $f \in \text{End}_K(V)$, $0 \neq x \in V$ und $m \in \mathbb{N}$ mit $f^{m-1}(x) \neq 0$ und $f^m(x) = 0$.

- Zeige, $B = (f^{m-1}(x), f^{m-2}(x), \dots, f(x), x)$ ist eine Basis von $U = \text{Lin}(B)$.
- Zeige, U ist f -invariant.
- Bestimme $M_B^B(f|_U)$.

Wir nennen U einen *zyklischen Unterraum* von V .

Aufgabe 26.36

Für Matrizen $A \in \text{Mat}(\mathbf{n} \times \mathbf{p}, \mathbf{K})$ und $B \in \text{Mat}(\mathbf{m} \times \mathbf{n}, \mathbf{K})$ gilt:

$$\text{rang}(B \circ A) \leq \min \{ \text{rang}(A), \text{rang}(B) \}.$$

Aufgabe 26.37

Für eine Matrix $A \in \text{Mat}(\mathbf{m} \times \mathbf{n}, \mathbf{K})$ bezeichne $\text{ZR}(A)$ die lineare Hülle der Zeilen von A und $\text{SR}(A)$ die lineare Hülle der Spalten von A .

Zeige für $A \in \text{Mat}(\mathbf{m} \times \mathbf{n}, \mathbf{K})$, $S \in \text{Gl}_m(\mathbf{K})$ und $T \in \text{Gl}_n(\mathbf{K})$

$$\text{ZR}(A) = \text{ZR}(SA) \quad \text{und} \quad \text{SR}(A) = \text{SR}(AT).$$

Aufgabe 26.38

Betrachte den Vektorraum P_n der Polynome vom Grad höchstens n (siehe Beispiel 23.6) mit Basis $B = (t^0, t^1, \dots, t^n)$ und die formale Ableitung

$$d : P_n \longrightarrow P_n : \sum_{k=0}^n a_k \cdot t^k \mapsto \sum_{k=1}^n k \cdot a_k \cdot t^{k-1},$$

von der wir aus Beispiel 23.21 bereits wissen, daß sie \mathbf{K} -linear ist.

- Berechne die Matrixdarstellung $M_B^B(d)$ und den Rang von d .
- Zeige, daß im Fall $n = 3$ auch $D = (t^0, t^0 + t^1, t^1 + t^2, t^2 + t^3)$ eine Basis von P_3 ist und berechne die Basiswechsel T_B^D und T_D^B sowie die Matrixdarstellung $M_D^D(d)$.

§ 27 Der Gauß-Algorithmus

In diesem Abschnitt wollen wir zeigen, daß man jede Matrix durch elementare Zeilenoperationen in Zeilen-Stufen-Form transformieren kann, und einen Algorithmus angeben, der dies tut, den *Gauß-Algorithmus*.

Definition 27.1 (Zeilen-Stufen-Form)

Es sei $A = (a_{ij}) \in \text{Mat}(m \times n, K)$.

a. A ist in *Zeilen-Stufen-Form*, kurz ZSF, falls es ein r , mit $0 \leq r \leq m$ und Indizes j_1, \dots, j_r mit $1 \leq j_1 < j_2 < \dots < j_r \leq n$ gibt, so daß folgendes gilt:

- (i) $a_{ij} = 0$ für $1 \leq i \leq r$ und $1 \leq j < j_i$,
- (ii) $a_{ij} = 0$ für $r < i \leq m$ und $j = 1, \dots, n$, und
- (iii) $a_{ij_i} \neq 0$ für $i = 1, \dots, r$.

Die Körperelemente a_{ij_i} heißen die *Pivots* der Zeilen-Stufen-Form. Man beachte, daß A genau r linear unabhängige Zeilen hat und daß somit $r = \text{rang}(A)$!

b. Eine Zeilen-Stufen-Form von A heißt *reduziert*, falls zusätzlich gilt:

- (iv) $a_{ij_i} = 1$ für $i = 1, \dots, r$, und
- (v) $a_{ki} = 0$ für $k < i$ und $i = 1, \dots, r$.

Bemerkung 27.2

Eine Matrix A in Zeilen-Stufen-Form ist also von der folgenden Gestalt:

$$A = \begin{pmatrix} 0 & \dots & 0 & \boxed{a_{1j_1}} & * & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & * \\ 0 & \dots & \dots & 0 & \dots & 0 & \boxed{a_{2j_2}} & * & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & * \\ 0 & \dots & \dots & \dots & \dots & \dots & 0 & \dots & 0 & \boxed{a_{3j_3}} & * & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & * \\ \vdots & & & & & & & & & \ddots & & & & & & & & & & & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \boxed{a_{rj_r}} & * & \dots & * & & & & & * \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ \vdots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix}$$

Hat A reduzierte Zeilen-Stufen-Form, so sind die Pivots alle Eins und die Einträge in der Spalte oberhalb der Pivots sind alle Null.

Beispiel 27.3

Betrachte die Matrizen $A, B, C \in \text{Mat}(4 \times 5, K)$ mit

$$A = \begin{pmatrix} 0 & \boxed{1} & 0 & 3 & 0 \\ 0 & 0 & \boxed{1} & 2 & 0 \\ 0 & 0 & 0 & 0 & \boxed{1} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} \boxed{1} & 2 & 3 & 4 & 5 \\ 0 & \boxed{2} & 3 & 4 & 5 \\ 0 & 0 & \boxed{3} & 0 & 0 \\ 0 & 0 & 0 & \boxed{4} & 1 \end{pmatrix} \text{ und } C = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 3 & 4 & 0 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Die Matrix A ist in reduzierter ZSF mit $\text{rang}(A) = r = 3$, $j_1 = 2$, $j_2 = 3$ und $j_3 = 5$. Die Matrix B ist in ZSF mit $\text{rang}(B) = r = 4$ und $j_1 = 1$, $j_2 = 2$, $j_3 = 3$ und $j_4 = 4$.

Die ZSF ist aber nicht reduziert.

Die Matrix C ist nicht in ZSF. Aber durch Vertauschen der beiden ersten Zeilen entsteht eine Matrix, die ZSF hat.

Für die folgende Definition erinnern wir uns (siehe Beispiel 24.11) an die Matrizen

$$E_i^j = (e_{lk})_{l,k=1,\dots,n} = (\delta_{il} \cdot \delta_{jk})_{l,k=1,\dots,n} \in \text{Mat}_n(\mathbb{K}),$$

die an der Stelle (i, j) den Eintrag 1 und sonst nur Nullen als Einträge haben.

Definition 27.4 (Elementarmatrizen)

Es seien $0 \neq \lambda \in \mathbb{K}$, $n > 0$ und $1 \leq i, j \leq n$ mit $i \neq j$. Wir definieren die folgenden quadratischen Matrizen in $\text{Mat}(n, \mathbb{K})$, die auch als *Elementarmatrizen* bezeichnet werden:

- a. $S_i(\lambda) := \mathbb{1}_n + (\lambda - 1) \cdot E_i^i$,
- b. $Q_i^j(\lambda) := \mathbb{1}_n + \lambda \cdot E_i^j$, und
- c. $P_i^j := \mathbb{1}_n - E_i^i - E_j^j + E_i^j + E_j^i$.

Die Matrizen P_i^j heißen zudem *Permutationsmatrizen*.

Bemerkung 27.5 (Elementare Zeilen- und Spaltenoperationen)

Es seien $0 \neq \lambda \in \mathbb{K}$, $1 \leq i, j \leq n$ mit $i \neq j$ und $A \in \text{Mat}(n \times m, \mathbb{K})$.

- I $S_i(\lambda) \circ A$ geht aus A hervor, indem man die i -te Zeile von A mit λ multipliziert.
- II $Q_i^j(\lambda) \circ A$ geht aus A hervor, indem man zur i -ten Zeile das λ -fache der j -ten Zeile addiert.
- III $P_i^j \circ A$ geht aus A hervor, indem man die i -te und j -te Zeile vertauscht.

Man nennt die Multiplikation von links mit diesen Matrizen auch *elementare Zeilenoperationen*. Analog erhält man *elementare Spaltenoperationen*, indem man mit den Matrizen von rechts multipliziert.

- I' $A \circ S_j(\lambda)$ geht aus A hervor, indem man die j -te Spalte von A mit λ multipliziert.
- II' $A \circ Q_i^j(\lambda)$ geht aus A hervor, indem man zur j -ten Spalte das λ -fache der i -ten Spalte addiert.
- III' $A \circ P_i^j$ geht aus A hervor, indem man die i -te und j -te Spalte vertauscht.

Proposition 27.6 (Elementarmatrizen sind invertierbar.)

Es seien $0 \neq \lambda \in \mathbb{K}$, $1 \leq i, j \leq n$ mit $i \neq j$ und $A \in \text{Mat}(n \times m, \mathbb{K})$. Dann gelten:

- a. $S_i(\lambda^{-1}) \circ S_i(\lambda) = \mathbb{1}_n$,
- b. $Q_i^j(-\lambda) \circ Q_i^j(\lambda) = \mathbb{1}_n$, und
- c. $P_i^j \circ P_i^j = \mathbb{1}_n$.

Insbesondere sind die Elementarmatrizen invertierbar und die Inversen sind wiederum Elementarmatrizen vom gleichen Typ.

Beweis: Wir führen den Beweis für b. vor. Die übrigen Teile lassen sich dann analog zeigen. Für $0 \neq \lambda \in \mathbb{K}$ gilt, mittels der Distributivität der Matrixmultiplikation:

$$Q_i^j(-\lambda) \circ Q_i^j(\lambda) = (\mathbb{1}_n - \lambda \cdot E_i^j) \circ (\mathbb{1}_n + \lambda \cdot E_i^j) = \mathbb{1}_n - \lambda^2 \cdot E_i^j \circ E_i^j = \mathbb{1}_n,$$

da $E_i^j \circ E_i^j = 0$ wegen $i \neq j$. Beachte dazu, daß für $E_i^j \circ E_i^j = (c_{ik})$ gilt:

$$c_{ik} = \sum_{p=1}^n \delta_{ip} \delta_{jp} \delta_{ip} \delta_{jk},$$

und daß für $i \neq j$ und p beliebig gilt $\delta_{jp} \delta_{ip} = 0$. □

Satz 27.7 (Existenz der reduzierten Zeilen-Stufen-Form)

Jede Matrix $A \in \text{Mat}(m \times n, \mathbb{K})$ läßt sich mittels endlich vieler elementarer Zeilenoperationen in reduzierte Zeilen-Stufen-Form $rZSF(A)$ überführen, d.h. es gibt Elementarmatrizen T_1, \dots, T_k , so daß

$$rZSF(A) = T_1 \circ \dots \circ T_k \circ A.$$

Beweis: Sei $A \in \text{Mat}(m \times n, \mathbb{K})$. Ist $A = 0$, so hat A bereits ZSF mit $r = 0$ und wir sind fertig.

Ist $A \neq 0$, so führe folgende Schritte durch:

- Schritt:** Durchlaufe die Spalten von oben nach unten, mit der ersten Spalte beginnend, bis der erste Eintrag $a_{i_1 j_1} \neq 0$ gefunden ist:

$$\begin{pmatrix} 0 & \dots & 0 & 0 & * & \dots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & 0 & * & \dots & * \\ \vdots & & \vdots & a_{i_1 j_1} & * & \dots & * \\ \vdots & & \vdots & * & * & \dots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & * & * & \dots & * \end{pmatrix}$$

- Schritt:** Steht $a_{i_1 j_1}$ nicht in der ersten Zeile, d. h. $i_1 \neq 1$, dann vertausche die Zeilen a_1 und a_{i_1} - Zeilenoperation vom Typ III. Die so entstandene Matrix heie $\tilde{A}_1 = (\tilde{a}_{ij})$. Dann ist \tilde{a}_{1j_1} unser erstes Pivot.

- Schritt:** Erzeuge in der Spalte \tilde{a}^{j_1} von \tilde{A}_1 unterhalb von \tilde{a}_{1j_1} Nullen durch elementare Operationen vom Typ II, d. h. addiere für $k = 2, \dots, m$ zur k -ten Zeile das $-\frac{\tilde{a}_{kj_1}}{\tilde{a}_{1j_1}}$ -fache der ersten Zeile. Die Spalten mit Index kleiner als j_1 werden dadurch nicht geändert. Das Ergebnis ist dann eine Matrix von der Form:

$$A^{(1)} := \left(\begin{array}{cccc|ccc} 0 & \dots & 0 & a_{1j_1}^{(1)} & * & \dots & * \\ 0 & \dots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & & \\ 0 & \dots & 0 & 0 & & A_2 & \end{array} \right),$$

wobei A_2 eine $(m - 1) \times (n - j_1)$ -Matrix ist, sofern $j_1 \neq n$.

Ist $n - j_1 = 0$ oder $m - 1 = 0$ oder $A^{(2)} = 0$, so sind wir fertig.

Andernfalls ist $A_2 \neq 0$, und wir führen Schritt 1-3 mit A_2 durch. Dabei kann man alle Zeilenoperationen auf die Matrix $A^{(1)}$ ausdehnen, ohne daß sich in den ersten j_1 Spalten etwas ändert, da dort nur Nullen stehen. Ist A_2 umgeformt, so erhält man eine Matrix $A^{(2)}$ der Form:

$$A^{(2)} = \left(\begin{array}{ccccccc|ccc} 0 & \dots & 0 & a_{1j_1}^{(2)} & * & \dots & \dots & \dots & \dots & * \\ 0 & \dots & 0 & 0 & \dots & 0 & a_{2j_2}^{(2)} & * & \dots & * \\ \hline 0 & \dots & 0 & 0 & \dots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & & \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & & & \end{array} \right) \begin{array}{c} \\ \\ A_3 \\ \\ \end{array}$$

mit einem Pivot $a_{2j_2}^{(2)}$ und, sofern nicht $m - 2 = 0$ oder $n - j_2 = 0$, einer Matrix A_3 , die eine Zeile und mindestens eine Spalte weniger als A_2 hat.

Ist $A_3 = 0$, so sind wir fertig. ansonsten fahren wir fort wie bisher und erhalten Matrizen $A^{(3)}, A_4, A^{(4)}, \dots$. Das Verfahren stoppt, falls nach r -maligem Durchlaufen der Schritte 1-3 entweder $r = m$ oder $r = n$ oder $A_{r+1} = 0$. In jedem der drei Fälle ist die Matrix $A^{(r)}$ in ZSF.

Um die Matrix $A^{(r)} = (a_{ij}^{(r)})$ in reduzierte ZSF zu bringen, multiplizieren wir zunächst die Zeilen $a_i^{(r)}$, für $i = 1, \dots, r$, mit $\frac{1}{a_{ij_i}^{(r)}}$, was einer elementaren Zeilenoperation vom Typ I entspricht. Die so entstehende Matrix heie $A' = (a'_{ij})$. Sodann addiert man für $i = 1, \dots, r$ und $k = 1, \dots, i - 1$ zur k -ten Zeile das $-a'_{ki}$ -fache der i -ten Zeile – elementare Operationen vom Typ II – und nennt in jedem Schritt i die neue Matrix wieder A' . Man sieht unmittelbar, daß die entstehende Matrix $A'' = (a''_{ij})$ reduzierte ZSF hat, da in Spalte j_i die Elemente a'_{ki} in $a''_{ki} = 0$, für $k < i$, übergegangen sind. □

Bemerkung 27.8 (Eindeutigkeit der reduzierten Zeilen-Stufen-Form)

- a. Der Beweis des Satzes ist konstruktiv, das heißt, aus dem Beweis läßt sich ein Algorithmus zur Berechnung einer ZSF von A herleiten, der sogenannte *Gauß-Algorithmus*.
- b. Die reduzierte Zeilenstufenform einer Matrix A ist eindeutig bestimmt, was die Bezeichnung $rZSF(A)$ rechtfertigt.

Beweis der Eindeutigkeit der Zeilenstufenform: Es sei also $A \in \text{Mat}(m \times n, K)$ eine $m \times n$ -Matrix.

Da elementare Zeilenoperationen durch Multiplikation mit Elementarmatrizen von links realisiert werden, gilt für eine ZSF B von A , daß es eine invertierbare Matrix $S \in \text{Gl}_m(K)$ gibt mit $B = S \circ A$ (vgl. auch Satz 27.7). Mit Aufgabe 26.37 folgt dann

$ZR(A) = ZR(B)$, insbesondere gilt mit Korollar 25.7 also, daß die Nicht-Null-Zeilen von B eine Basis von $ZR(A)$ bilden, da

$$r := \dim_K (ZR(A)) = \text{rang}(A) = \text{rang}(B). \quad (69)$$

Seien nun $B = (b_{ij})$ und $B' = (b'_{ij})$ zwei reduzierte ZSF von A mit Zeilenvektoren b_1, \dots, b_m bzw. b'_1, \dots, b'_m und Pivotspalten $\{j_1, \dots, j_r\}$ bzw. $\{j'_1, \dots, j'_r\}$ - beachte, daß die Anzahl $r = \text{rang}(A)$ nach (69) für beide gleich ist. Wir zeigen nun per Induktion, daß die Zeilen der Matrizen B und B' übereinstimmen.

Induktionsbehauptung: Für $i \in \mathbb{N}$ gilt entweder $i \geq r$ oder $b_{r-i} = b'_{r-i}$, insbesondere also $j_{r-i} = j'_{r-i}$.

Induktionsanfang: $i = 0$. O. E. gelte $j_r \geq j'_r$. Da $b_r \in ZR(A) = \text{Lin}(b'_1, \dots, b'_r)$, gibt es $\lambda_1, \dots, \lambda_r \in K$ mit

$$b_r = \sum_{i=1}^r \lambda_i b'_i.$$

Insbesondere gilt für $i = 1, \dots, r-1$

$$0 = b_{rj'_i} = \lambda_i \quad \text{und} \quad b_{rj'_r} = \lambda_r,$$

nach (iv) und (v) in Definition 27.1 angewandt auf die reduzierte ZSF B' mit Pivotspalten j'_1, \dots, j'_r sowie (i) angewandt auf die ZSF B . Also folgt $b_r = \lambda_r \cdot b'_r$. Da $b_r \neq 0$, muß $\lambda_r \neq 0$ gelten und somit $j'_r = j_r$ wegen (i) in 27.1. Aber dann gilt nach (iv) in 27.1 $1 = b_{rj_r} = \lambda_r$ und somit $b_r = b'_r$.

Induktionsschritt: $0 < i < r-1$ und die Behauptung gelte schon für $0, \dots, i-1$. O. E. gelte $j_{r-i} \geq j'_{r-i}$. Nach Induktionsvoraussetzung gilt nun $b_{r-i} \in ZR(A) = \text{Lin}(b'_1, \dots, b'_{r-i}, b_{r-i+1}, \dots, b_r)$ also gibt es $\lambda_1, \dots, \lambda_r \in K$ mit

$$b_{r-i} = \sum_{k=1}^{r-i} \lambda_k b'_k + \sum_{k=r-i+1}^r \lambda_k b_k.$$

Insbesondere gilt nach (v) in Definition 27.1, angewandt auf die reduzierte ZSF B , für $k = r-i+1, \dots, r$

$$0 = b_{r-i j_k} = \lambda_k,$$

da $r-i < k$, und (i) angewandt auf B sowie (v) auf B' liefert für $k = 1, \dots, r-i-1$

$$0 = b_{r-i j'_k} = \lambda_k,$$

da $j'_k < j'_{r-i} \leq j_{r-i}$. Insgesamt erhalten wir also wieder

$$b_{r-i} = \lambda_{r-i} b'_{r-i}. \quad (70)$$

Wäre $j_{r-i} > j'_{r-i}$, dann wäre wieder mit (i) $0 = b_{r-i j'_{r-i}} = \lambda_{r-i}$ im Widerspruch zu (70) und $b_{r-i} \neq 0$. Also ist $j_{r-i} = j'_{r-i}$ und dann folgt mit (iv) aus 27.1, daß $\lambda_{r-i} = b_{r-i j_{r-i}} = 1$, und damit aus (70)) $b_{r-i} = b'_{r-i}$.

Also haben wir mit Induktion gezeigt, daß die Zeilen von B und B' übereinstimmen, d. h. daß die reduzierte Zeilenstufenform von A eindeutig bestimmt ist. \square

Beispiel 27.9

Wir überführen nun die folgende Matrix in reduzierte ZSF.

$$\begin{array}{ccc}
 \begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} & \xrightarrow{\text{I} \leftrightarrow \text{II}} & \begin{pmatrix} -1 & 1 & -3 & 0 & 2 \\ 0 & 0 & -1 & 2 & 3 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} & \xrightarrow{\text{III} \rightarrow \text{III} + \text{I}} \\
 \begin{pmatrix} -1 & 1 & -3 & 0 & 2 \\ 0 & 0 & -1 & 2 & 3 \\ 0 & 0 & -2 & 4 & 5 \end{pmatrix} & \xrightarrow{\text{III} \rightarrow \text{III} - 2 \cdot \text{II}} & \begin{pmatrix} -1 & 1 & -3 & 0 & 2 \\ 0 & 0 & -1 & 2 & 3 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix} & \xrightarrow{\begin{array}{l} \text{I} \rightarrow -\text{I}, \text{II} \rightarrow -\text{II} \\ \text{III} \rightarrow -\text{III} \end{array}} \\
 \begin{pmatrix} 1 & -1 & 3 & 0 & -2 \\ 0 & 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} & \xrightarrow{\begin{array}{l} \text{I} \rightarrow \text{I} + 2 \cdot \text{III} \\ \text{II} \rightarrow \text{II} + 3 \cdot \text{III} \end{array}} & \begin{pmatrix} 1 & -1 & 3 & 0 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} & \xrightarrow{\text{I} \rightarrow \text{I} - 3 \cdot \text{II}} \\
 \begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} & & & &
 \end{array}$$

Die vierte Matrix besitzt bereits ZSF mit unterstrichenen Pivots, die letzte ist in reduzierter ZSF.

Wir bemerken, daß wir auch auf anderem Weg zum Ziel gekommen wären, und zwar durch andere Wahl der Pivots.

$$\begin{array}{ccc}
 \begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} & \xrightarrow{\text{I} \leftrightarrow \text{III}} & \begin{pmatrix} 1 & -1 & 1 & 4 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ 0 & 0 & -1 & 2 & 3 \end{pmatrix} & \xrightarrow{\text{II} \rightarrow \text{II} + \text{I}} \\
 \begin{pmatrix} 1 & -1 & 1 & 4 & 3 \\ 0 & 0 & -2 & 4 & 5 \\ 0 & 0 & -1 & 2 & 3 \end{pmatrix} & \xrightarrow{\text{III} \rightarrow \text{III} - \frac{1}{2} \cdot \text{II}} & \begin{pmatrix} 1 & -1 & 1 & 4 & 3 \\ 0 & 0 & -2 & 4 & 5 \\ 0 & 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} & \xrightarrow{\begin{array}{l} \text{II} \rightarrow -\frac{1}{2} \cdot \text{II} \\ \text{III} \rightarrow 2 \cdot \text{III} \end{array}} \\
 \begin{pmatrix} 1 & -1 & 1 & 4 & 3 \\ 0 & 0 & 1 & -2 & -\frac{5}{2} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} & \xrightarrow{\begin{array}{l} \text{I} \rightarrow \text{I} - 3 \cdot \text{III} \\ \text{II} \rightarrow \text{II} + \frac{5}{2} \cdot \text{III} \end{array}} & \begin{pmatrix} 1 & -1 & 1 & 4 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} & \xrightarrow{\text{I} \rightarrow \text{I} - \text{II}} \\
 \begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} & & & &
 \end{array}$$

In der Praxis sind 1000×1000 -Matrizen keine Seltenheit. Dort wird mit einer festen Stellenzahl gerechnet und deshalb treten bei großen Matrizen unter Umständen

erhebliche Rundungsfehler auf. Es kommt der Wahl der richtigen Pivots eine große Bedeutung zu. Ist das gewählte Pivot zu klein, so kann bei Division durch dieses Pivot im dritten Schritt der Rundungsfehler riesig werden - für den Computer bedeutet dies in etwa, als ob man durch Null zu dividieren versuche. Deshalb wählt man in der Praxis das betragsmäßig größte Element als Pivot.

Rechnet man allerdings in Computeralgebrasystemen mit exakter Arithmetik, so spielt die Auslöschung durch Rundungsfehler keine Rolle. Dort muß man eher dafür sorgen, daß die Zahlen, d. h. die Zähler und Nenner, nicht zu groß werden, da dies zu erheblichen Geschwindigkeitsverlusten führen würde.

Wir wollen abschließend den Gauß-Algorithmus in leicht abgewandelter Form als rekursiven Algorithmus zur Bestimmung der reduzierten ZSF einer Matrix formulieren.

Algorithmus 27.10 (Gauß-Algorithmus)

INPUT: $A \in \text{Mat}(m \times n, K)$.

OUTPUT: rZSF(A), die reduzierte Zeilen-Stufen-Form von A.

1. **Schritt:** Falls $A = 0$ oder $m = 1$, gehe zu Schritt 8.
2. **Schritt:** Durchlaufe die erste Spalte von oben nach unten, bis ein Element ungleich Null a_{i1} gefunden wurde oder das Ende der Spalte erreicht ist.
3. **Schritt:** Wurde kein $a_{i1} \neq 0$ gefunden, bilde eine Untermatrix B von A durch Streichen der ersten Spalte von A und gehe zu Schritt 6. Andernfalls, vertausche die Zeilen a_1 und a_i .
4. **Schritt:** Für $k = 2, \dots, m$ addiere zur k-ten Zeile das $-\frac{a_{k1}}{a_{i1}}$ -fache der ersten.
5. **Schritt:** Falls $n = 1$, gehe zu Schritt 7. Andernfalls bilde eine Untermatrix B von A, durch Streichen der ersten Zeile und der ersten Spalte von A.
6. **Schritt:** Wende den Algorithmus auf die Untermatrix B an.²
7. **Schritt:** Die Matrix A ist nun in ZSF. Für $i = m$ bis $i = 1$, d. h. rückwärts zählend, durchlaufe die Zeile a_i , beginnend mit der ersten Spalte, bis ein Element $a_{ij} \neq 0$ gefunden wurde oder das Ende der Zeile erreicht ist. In letzterem Fall tue nichts, in ersterem multipliziere die Zeile a_i mit $\frac{1}{a_{ij}}$ und addiere für $k = 1, \dots, i - 1$ zur k-ten Zeile das $-a_{kj}$ -fache der i-ten Zeile.
8. **Schritt:** Gib die (veränderte) Matrix A zurück.

A) Algorithmus zur Berechnung des Rangs einer Matrix

Lemma 27.11

Elementare Zeilen- oder Spaltenoperationen ändern den Rang einer Matrix nicht.

Beweis: Multipliziert man eine Matrix A mit einer invertierbaren Matrix, so erhält man eine äquivalente Matrix. Wegen Proposition 26.26 ändert dies den Rang der

²Dies ist der Rekursionsschritt, indem der Algorithmus mit einer kleineren Untermatrix aufgerufen wird. Das Ergebnis, das man dabei zurück erhält, wird wieder in die Matrix A eingefügt.

Matrix nicht. Da elementare Zeilen- und Spaltenoperationen durch Multiplikation mit invertierbaren Matrizen entstehen, ändern auch diese den Rang der Matrix nicht. \square

Algorithmus 27.12 (zur Bestimmung des Rangs)

INPUT: $A \in \text{Mat}(m \times n, K)$.

OUTPUT: $\text{rang}(A)$

1. **Schritt:** Überführe A in ZSF.
2. **Schritt:** Zähle die Anzahl r der Nicht-Nullzeilen in der ZSF.
3. **Schritt:** Gib r zurück.

Beispiel 27.13

In Beispiel 27.9 haben wir eine ZSF berechnet:

$$A := \begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} \longmapsto \dots \longmapsto \begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Die Matrix A hat also Rang 3.

B) Algorithmus zur Berechnung der Inversen einer Matrix

Satz 27.14 (Kriterium für die Invertierbarkeit einer Matrix)

Es sei $A \in \text{Mat}(n, K)$. Dann sind gleichwertig:

- a. A ist invertierbar.
- b. $r\text{ZSF}(A) = \mathbb{1}_n$.
- c. Es gibt Elementarmatrizen $T_1, \dots, T_k \in \text{Mat}(n, K)$ mit:

$$T_k \circ \dots \circ T_1 \circ A = \mathbb{1}_n.$$

- d. Es gibt Elementarmatrizen $T'_1, \dots, T'_k \in \text{Mat}(n, K)$ mit:

$$A = T'_1 \circ \dots \circ T'_k.$$

Insbesondere wird die Gruppe $\text{Gl}_n(K)$ also von den Elementarmatrizen erzeugt.

Beweis: Nach Korollar 26.24 gilt, daß A genau dann invertierbar ist, wenn $\text{rang}(A) = n$. Also folgt die Äquivalenz von a.-d. aus Satz 27.7 unter Berücksichtigung von Proposition 27.6. \square

Aus Satz 27.14 leitet sich folgendes Verfahren zur Bestimmung der Inversen einer invertierbaren Matrix ab. Hierzu beachte man, daß für Elementarmatrizen T_1, \dots, T_k , für die gilt, daß $T_k \circ \dots \circ T_1 \circ A = \mathbb{1}_n$, auch gilt, daß

$$T_k \circ \dots \circ T_1 \circ (A, \mathbb{1}_n) = (\mathbb{1}_n, T_k \circ \dots \circ T_1) = (\mathbb{1}_n, A^{-1}).$$

Algorithmus 27.15 (zur Bestimmung der Inversen)

INPUT: $A \in \text{Mat}(n, K)$.

OUTPUT: Inverse von A , falls sie existiert, eine Fehlermeldung sonst.

1. **Schritt:** Erweitere die Matrix A um $\mathbb{1}_n$ zu $C = (A, \mathbb{1}_n) \in \text{Mat}(n \times 2n, K)$.
2. **Schritt:** Überführe C in reduzierte ZSF $C' = (A', B)$.
3. **Schritt:** Falls $\text{rang}(A') = n$, dann gib B zurück, sonst eine Fehlermeldung.

Beispiel 27.16

Wir betrachten die 3×3 -Matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in \text{Mat}(3 \times 3, K)$$

und versuchen die Inverse mittels des Algorithmus 27.15 zu bestimmen.

A	$\mathbb{1}_n$	
1 1 1	1 0 0	
0 1 1	0 1 0	
1 0 1	0 0 1	III \mapsto III - I
1 1 1	1 0 0	
0 1 1	0 1 0	
0 -1 0	-1 0 1	III \mapsto III + II
1 1 1	1 0 0	I \mapsto I - III
0 1 1	0 1 0	II \mapsto II - III
0 0 1	-1 1 1	
1 1 0	2 -1 -1	I \mapsto I - II
0 1 0	1 0 -1	
0 0 1	-1 1 1	
1 0 0	1 -1 0	
0 1 0	1 0 -1	
0 0 1	-1 1 1	

Hieraus ergibt sich gemäß obigem Algorithmus zunächst, daß A invertierbar ist, und ferner, daß

$$A^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

C) Algorithmus zur Berechnung der Normalform einer Matrix

Korollar 27.17 (Normalform einer Matrix)

Sei $A \in \text{Mat}(m \times n, K)$ mit $r = \text{rang}(A)$, so lässt sich A durch endlich viele elementare Zeilen- und Spaltenoperationen auf die folgende Form bringen:

$$\left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array} \right). \tag{71}$$

Beweis: Die Aussage folgt aus Korollar 26.29 und Korollar 27.14, da elementare Operationen nach Bemerkung 27.5 durch Multiplikation mit Elementarmatrizen realisierbar sind. □

Wir wollen nun noch an einem Beispiel zeigen, wie man eine Matrix mittels des gaußschen Verfahrens auf Normalform (71) bringt.

Algorithmus 27.18 (Normalform-Algorithmus)

INPUT: $A \in \text{Mat}(m \times n, K)$.

OUTPUT: Normalform $\text{NF}(A)$ von A bezüglich Äquivalenz sowie die zugehörigen Transformationsmatrizen $S \in \text{Gl}_m(K)$ und $T \in \text{Gl}_n(K)$

1. **Schritt:** Überführe A durch elementare Zeilenoperationen in (reduzierte) ZSF und überführe $\mathbb{1}_m$ durch die selben Zeilenoperationen in eine Matrix S .
2. **Schritt:** Überführe A durch elementare Spaltenoperationen in Normalform und überführe $\mathbb{1}_n$ durch die selben Spaltenoperationen in eine Matrix T .
3. **Schritt:** Gib die Normalform von A sowie die Matrizen S und T zurück.

Beispiel 27.19

Durch elementare Zeilen und Spaltenoperationen überführt man $A_\lambda, \lambda \in K$, in Normalform:

$\mathbb{1}_m$	A_λ	$\mathbb{1}_n$	
1 0 0	1 0 λ	1 0 0	
0 1 0	0 1 0	0 1 0	ZIII \mapsto ZIII $- \lambda \cdot$ ZI
0 0 1	λ 0 1	0 0 1	
1 0 0	1 0 λ	1 0 0	
0 1 0	0 1 0	0 1 0	SIII \mapsto SIII $- \lambda \cdot$ SI
$-\lambda$ 0 1	0 0 $1 - \lambda^2$	0 0 1	
1 0 0	1 0 0	1 0 $-\lambda$	falls $\lambda = \pm 1$ fertig,
0 1 0	0 1 0	0 1 0	sonst SIII $\mapsto \frac{1}{1-\lambda^2} \cdot$ SIII
$-\lambda$ 0 1	0 0 $1 - \lambda^2$	0 0 1	
1 0 0	1 0 0	1 0 $-\frac{\lambda}{1-\lambda^2}$	
0 1 0	0 1 0	0 1 0	
$-\lambda$ 0 1	0 0 1	0 0 $\frac{1}{1-\lambda^2}$	
S	NF(A_λ)	T	

Für die Normalform $NF(A) = SAT$ erhalten wir also

$$NF(A_\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\lambda & 0 & 1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 & -\frac{\lambda}{1-\lambda^2} \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{1-\lambda^2} \end{pmatrix},$$

falls $\lambda \neq \pm 1$, und sonst

$$NF(A_\lambda) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -\lambda & 0 & 1 \end{pmatrix}, \quad T = \begin{pmatrix} 1 & 0 & -\lambda \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Insbesondere gilt, $\text{rang}(A) = 3$ für $\lambda^2 \neq 1$ und $\text{rang}(A) = 2$ sonst.

D) Algorithmus zur Berechnung einer Basis

Der folgende Algorithmus zur Bestimmung einer Basis aus gegebenem Erzeugendensystem beruht auf der Tatsache, daß elementare Zeilenoperationen den Zeilenraum nicht verändern - vgl. Aufgabe 26.37.

Algorithmus 27.20 (Basisberechnung)

INPUT: Ein Erzeugendensystem F des Unterraums $U \subseteq K^n$.

OUTPUT: Eine Basis von U .

1. **Schritt:** Schreibe die Vektoren von F als Zeilen in eine Matrix A und überführe A in Zeilen-Stufen-Form.
2. **Schritt:** Gib die ersten $\text{rang}(A)$ Zeilen als Vektoren zurück.

Beispiel 27.21

Betrachte $U = \text{Lin}((1, 0, -1, 2, 3)^t, (1, -1, 1, 4, 3)^t, (0, 2, -4, -4, 0)^t) \leq \mathbb{R}^5$:

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

Also ist $B = ((1, 0, -1, 2, 3)^t, (0, -1, 2, 2, 0)^t)$ eine Basis von U .

E) Algorithmus zum Test auf Injektivität / Surjektivität / Bijektivität

Bemerkung 27.22

Sei $A \in \text{Mat}(m \times n, K)$. Aus der Dimensionsformel für lineare Abbildungen 25.20

$$\dim_K(\text{Ker}(f_A)) = n - \text{rang}(A)$$

folgt unmittelbar:

- f_A ist injektiv $\iff \text{rang}(A) = n$.
- f_A ist surjektiv $\iff \text{rang}(A) = m$.
- f_A ist bijektiv $\iff \text{rang}(A) = n = m$.

Algorithmus 27.23 (Test auf Injektivität / Surjektivität / Bijektivität)

INPUT: $A \in \text{Mat}(\mathfrak{m} \times \mathfrak{n}, \mathbb{K})$.

OUTPUT: Meldung, ob f_A injektiv, surjektiv oder bijektiv ist.

1. **Schritt:** Bestimme den Rang r von A .
2. **Schritt:** Ist $r = \mathfrak{m} = \mathfrak{n}$, gib “ f_A ist bijektiv” zurück. Ist $r = \mathfrak{m} < \mathfrak{n}$, gib “ f_A ist surjektiv” zurück. Ist $r = \mathfrak{n} < \mathfrak{m}$, gib “ f_A ist injektiv” zurück.

Beispiel 27.24

Die zur folgenden Matrix $A \in \text{Mat}(3 \times 5, \mathbb{R})$ gehörende Abbildung $f_A : \mathbb{R}^5 \rightarrow \mathbb{R}^3$ ist weder injektiv noch surjektiv, da $\text{rang}(A) = 2 < 3 = \mathfrak{m}$ und $\text{rang}(A) = 2 < 5 = \mathfrak{n}$:

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

F) Algorithmus zur Berechnung der Summe zweier Unterräume

Die Berechnung der Summe zweier Unterräume, die durch Erzeuger gegeben sind, ist einfach, da man nur die Erzeuger der beiden Unterräume vereinigen muß.

Algorithmus 27.25 (Summe zweier Unterräume)

INPUT: Erzeugendensysteme F und G von zwei Unterräumen U und U' des K^n .

OUTPUT: Eine Basis von $U + U'$.

1. **Schritt:** Bilde aus F und G ein Erzeugendensystem und berechne mittels 27.20 eine Basis von $U + U' = \langle F \cup G \rangle$.
2. **Schritt:** Gib diese Basis zurück.

G) Algorithmus zum Testen auf lineare Unabhängigkeit

Da eine endliche Familie von Vektoren genau dann linear unabhängig ist, wenn sie eine Basis ihres Erzeugnisses ist, und da die Dimension des Erzeugnisses einer solchen Familie gerade der Rang der Matrix ist, deren Spalten die Erzeuger sind, liefert Korollar 25.7 den folgenden Algorithmus.

Algorithmus 27.26 (Test auf lineare Unabhängigkeit)

INPUT: Eine Familie F von \mathfrak{m} Vektoren in K^n .

OUTPUT: Eins, falls F linear unabhängig ist, Null sonst.

1. **Schritt:** Ist F leer, gib Eins zurück, sonst schreibe die Vektoren in F als Spalten in eine Matrix A .
2. **Schritt:** Ist $\text{rang}(A) = \mathfrak{m}$, so gib Eins zurück, sonst Null.

Ist $f = f_A$ für eine $\mathfrak{m} \times \mathfrak{n}$ -Matrix A , dann wird das Bild von f von den Spalten von A erzeugt. Wir können eine Basis des Bildes also wie folgt bestimmen.

H) Algorithmus zur Berechnung des Bildes einer linearen Abbildung**Algorithmus 27.27** (Bild von f_A)INPUT: $A \in \text{Mat}(m \times n, K)$.OUTPUT: Eine Basis von $\text{Im}(f_A)$.

- 1. Schritt:** Transponiere A und überführe die Transponierte in ZSF.
- 2. Schritt:** Transponiere das Ergebnis wieder und gib die ersten $\text{rang}(A)$ Spaltenvektoren zurück.

Bemerkung 27.28 (Ringe und Moduln)

Der Gauß-Algorithmus zur Berechnung einer reduzierten ZSF funktioniert über beliebigen kommutativen Ringen mit Eins im allgemeinen nicht mehr, da man dazu teilen muß. Man kann eine Matrix aber auch durch elementare Zeilenoperationen in nicht-reduzierte ZSF überführen, ohne zu teilen. Für manche Fragen ist eine solche ZSF hinreichend, die dann über Ringen berechnet werden kann. Im Übrigen funktioniert auch der Algorithmus zur Berechnung der reduzierten Zeilenstufenform, wenn man zwischendurch nur durch Elemente teilen muß, die ein Inverses im Ring besitzen. Damit kann man bei invertierbaren Matrizen dann z.B. die Inverse berechnen.

Aufgaben**Aufgabe 27.29**Es seien $0 \neq \lambda \in K$, $1 \leq i, j \leq n$ mit $i \neq j$. Dann gelten:

$$Q_i^j(\lambda) = S_j(\lambda^{-1}) \circ Q_i^j(1) \circ S_j(\lambda),$$

und

$$P_i^j = Q_j^i(1) \circ Q_i^j(-1) \circ Q_j^i(1) \circ S_j(-1).$$

Aufgabe 27.30Berechne den Rang der folgenden Matrix in Abhängigkeit von \mathbf{a} und \mathbf{b} :

$$\begin{pmatrix} 0 & \mathbf{b} & \mathbf{b} & \mathbf{b} \\ \mathbf{a} & 0 & \mathbf{b} & \mathbf{b} \\ \mathbf{a} & \mathbf{a} & 0 & \mathbf{b} \end{pmatrix} \in \text{Mat}_4(\mathbb{R}).$$

Aufgabe 27.31

Bestimme die Inverse der Matrix

$$\begin{pmatrix} 1 & 3 & -1 & 4 \\ 2 & 5 & -1 & 3 \\ 0 & 4 & -3 & 1 \\ -3 & 1 & -5 & -2 \end{pmatrix} \in \text{Mat}_4(\mathbb{R}).$$

Aufgabe 27.32

Transformiere die folgende Matrix A in Normalform bezüglich Äquivalenz und gib auch die Transformationsmatrizen S und T^{-1} an:

$$A = \begin{pmatrix} 1 & -2 & 3 & 0 \\ 2 & -7 & 10 & -1 \\ -2 & 4 & -7 & 2 \\ 3 & -5 & 7 & 1 \end{pmatrix} \in \text{Mat}_4(\mathbb{R}).$$

Aufgabe 27.33

Überprüfe die folgende Abbildung auf Injektivität und Surjektivität:

$$g: \mathbb{R}^4 \rightarrow \mathbb{R}^4 : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} 2x_1 + x_2 + x_4 \\ 3x_1 + x_2 + x_3 + x_4 \\ 2x_1 + x_3 + x_4 \\ 2x_1 + x_2 + x_3 \end{pmatrix}.$$

Aufgabe 27.34

Es sei $U = \langle (1, 2, 3, 4)^t, (1, 1, 1, 1)^t \rangle \leq \mathbb{R}^4$. Bestimme mit Hilfe des Austauschsatzes von Steinitz eine Basis von \mathbb{R}^4/U .

Aufgabe 27.35

Es sei $U = \{(\mathbf{a}_1, \dots, \mathbf{a}_5)^t \in \mathbb{R}^5 \mid \mathbf{a}_1 - 2\mathbf{a}_2 = \mathbf{0} = 2\mathbf{a}_4 + \mathbf{a}_5\} \leq \mathbb{R}^5$. Bestimme die Dimension von U sowie eine Basis von U , die den Vektor $(2, 1, 1, -1, 2)^t$ enthält.

Aufgabe 27.36

Seien $U = \langle (1, 0, 1, 1)^t, (-1, 1, 0, 0)^t \rangle \leq \mathbb{R}^4$ und $U' = \langle (1, 0, 1, 0)^t, (1, 1, 1, 1)^t \rangle \leq \mathbb{R}^4$. Zeige, $\mathbb{R}^4 = U \oplus U'$.

§ 28 Lineare Gleichungssysteme

Definition 28.1 (Lineare Gleichungssysteme)

- a. Ein *lineares Gleichungssystem* über K

$$\begin{array}{r}
 \text{(LGS)} \quad \begin{array}{l}
 \mathbf{a}_{11}x_1 + \mathbf{a}_{12}x_2 + \dots + \mathbf{a}_{1n}x_n = \mathbf{b}_1 \\
 \mathbf{a}_{21}x_1 + \mathbf{a}_{22}x_2 + \dots + \mathbf{a}_{2n}x_n = \mathbf{b}_2 \\
 \vdots \\
 \mathbf{a}_{m1}x_1 + \mathbf{a}_{m2}x_2 + \dots + \mathbf{a}_{mn}x_n = \mathbf{b}_m
 \end{array}
 \end{array}$$

besteht aus m Gleichungen in n *Unbestimmten* oder *Variablen* x_1, \dots, x_n mit $\mathbf{a}_{ij}, \mathbf{b}_i \in K$ für $1 \leq i \leq m$ und $1 \leq j \leq n$.

Da sich (LGS) mit $A = (\mathbf{a}_{ij})$, $\mathbf{b} = (\mathbf{b}_1, \dots, \mathbf{b}_m)^t$ und $\mathbf{x} = (x_1, \dots, x_n)^t$ mittels Matrixmultiplikation auch kurz schreiben läßt als

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

sprechen wir meist von dem *linearen Gleichungssystem* $A\mathbf{x} = \mathbf{b}$.

- b. Die Matrix

$$A = (\mathbf{a}_{ij}) = \begin{pmatrix} \mathbf{a}_{11} & \dots & \mathbf{a}_{1n} \\ \vdots & & \vdots \\ \mathbf{a}_{m1} & \dots & \mathbf{a}_{mn} \end{pmatrix} \in \text{Mat}(m \times n, K)$$

heißt *Koeffizientenmatrix* und der Vektor $\mathbf{b} = (\mathbf{b}_1, \dots, \mathbf{b}_m)^t \in K^m$ die *Inhomogenität* des Gleichungssystems (LGS). Ferner heißt die Matrix

$$(A | \mathbf{b}) := \left(\begin{array}{ccc|c} \mathbf{a}_{11} & \dots & \mathbf{a}_{1n} & \mathbf{b}_1 \\ \vdots & & \vdots & \vdots \\ \mathbf{a}_{m1} & \dots & \mathbf{a}_{mn} & \mathbf{b}_m \end{array} \right) \in \text{Mat}(m \times (n+1), K)$$

die *erweiterte Koeffizientenmatrix* von (LGS).

- c. Das lineare Gleichungssystem (LGS) heißt *homogen*, falls $\mathbf{b} = \mathbf{0}_{K^m}$ der Nullvektor in K^m ist. Ansonsten heißt das System *inhomogen*.
- d. Ist ein lineares Gleichungssystem $A\mathbf{x} = \mathbf{b}$ gegeben, so heißt das Gleichungssystem $A\mathbf{x} = \mathbf{0}$ (mit $\mathbf{0} = \mathbf{0}_{K^m}$) das *zugehörige homogene Gleichungssystem*.
- e. Ein Vektor $\mathbf{c} = (c_1, \dots, c_n)^t \in K^n$ heißt *Lösung* von (LGS), wenn die Gleichung $A\mathbf{c} = \mathbf{b}$ erfüllt ist. Die Menge aller Lösungen von (LGS) wird mit

$$\text{Lös}(A, \mathbf{b}) := \{\mathbf{c} \in K^n \mid A\mathbf{c} = \mathbf{b}\}.$$

bezeichnet.

Bei einem linearen Gleichungssystem sind also Körperelemente \mathbf{a}_{ij} und \mathbf{b}_i fest vorgegeben, während für die Unbestimmten x_j Körperelemente c_j gesucht werden, die das Gleichungssystem lösen.

Falls $K = \mathbb{R}$, so kann ein lineares Gleichungssystem entweder gar keine, genau eine oder unendlich viele Lösungen haben. Wir werden im Folgenden mehrere Verfahren zur Lösung kennenlernen und uns, im Fall von mehr als einer Lösung, mit der Struktur der Lösungsmenge $\text{Lös}(A, \mathbf{b})$ beschäftigen. Eine wichtige Rolle spielt dabei die lineare Abbildung $f_A : K^n \rightarrow K^m$.

Bemerkung 28.2 (Struktur des Lösungsraums)

Es sei $A \in \text{Mat}(m \times n, K)$ und $\mathbf{b} \in K^m$.

- a. Aus den Definitionen folgt unmittelbar

$$\text{Lös}(A, \mathbf{0}) = \{\mathbf{c} \in K^n \mid A\mathbf{c} = \mathbf{0}\} = \text{Ker}(f_A),$$

so daß $\text{Lös}(A, \mathbf{0})$ ein Unterraum des K^n ist mit Dimension

$$\dim_K(\text{Lös}(A, \mathbf{0})) = \dim_K(\text{Ker}(f_A)) = n - \text{rang}(A).$$

Insbesondere ist $A\mathbf{x} = \mathbf{0}$ genau dann eindeutig lösbar, wenn $\text{rang}(A) = n$.

- b. Ebenfalls anhand der Definitionen sieht man, daß das lineare Gleichungssystem $A\mathbf{x} = \mathbf{b}$ genau dann eine Lösung besitzt, wenn $\mathbf{b} \in \text{Im}(f_A) = \{A\mathbf{c} \mid \mathbf{c} \in K^n\}$.

Beispiel 28.3

Das lineare Gleichungssystem

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ 2x_1 + 3x_2 &= 1 \\ x_2 - x_3 &= 0 \end{aligned}$$

ist inhomogen, hat als Koeffizientenmatrix

$$A = \begin{pmatrix} 1 & 2 & 1 \\ 2 & 3 & 0 \\ 0 & 1 & -1 \end{pmatrix} \in \text{Mat}(3 \times 3, \mathbb{R}),$$

und als erweiterte Koeffizientenmatrix

$$(A \mid \mathbf{b}) = \left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 3 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{array} \right) \in \text{Mat}(3 \times 4, \mathbb{R}).$$

Die Lösung ist in diesem Fall ganz einfach. Wir erhalten $x_3 = x_2$ aus der 3. Gleichung, $3x_2 = 1 - 2x_1$ aus der 2. und, wenn wir das in die erste Gleichung einsetzen, $x_1 + (1 - 2x_1) = 1$, also $x_1 = 0$. Einsetzen von $x_1 = 0$ in die 2. und 3. Gleichung liefert, daß $(0, \frac{1}{3}, \frac{1}{3})^t$ die einzige Lösung ist.

Wir geben zunächst ein Kriterium für die Lösbarkeit eines Gleichungssystems.

Satz 28.4 (Kriterium für die Lösbarkeit eines LGS)

Ein Gleichungssystem $A\mathbf{x} = \mathbf{b}$ ist genau dann lösbar, wenn $\text{rang}(A) = \text{rang}(A \mid \mathbf{b})$.

Beweis: Wir beachten, daß $\text{Im}(f_A)$ von den Spaltenvektoren $\mathbf{a}^1, \dots, \mathbf{a}^n$ von A erzeugt wird, und erhalten deshalb:

$$\begin{aligned} Ax = \mathbf{b} \text{ lösbar} &\iff \mathbf{b} \in \text{Im}(f_A) = \text{Lin}(\mathbf{a}^1, \dots, \mathbf{a}^n) \\ &\iff \mathbf{b} \text{ ist Linearkombination von } \mathbf{a}^1, \dots, \mathbf{a}^n \\ &\iff \text{Im}(f_A) = \text{Lin}(\mathbf{a}^1, \dots, \mathbf{a}^n) = \text{Lin}(\mathbf{a}^1, \dots, \mathbf{a}^n, \mathbf{b}) = \text{Im}(f_{(A|\mathbf{b})}) \\ &\iff \text{rang}(A) = \dim_{\mathbb{K}}(\text{Im}(f_A)) = \dim_{\mathbb{K}}(\text{Im}(f_{(A|\mathbf{b})})) = \text{rang}(A | \mathbf{b}), \end{aligned}$$

wobei wir für die letzte Äquivalenz berücksichtigen, daß $\text{Im}(f_A) \subseteq \text{Im}(f_{(A|\mathbf{b})})$ gilt. \square

Der folgende Satz gibt Auskunft über die Struktur der Lösungsmenge eines linearen Gleichungssystems. Wir haben bereits gesehen, daß diese ein Unterraum ist, wenn das Gleichungssystem homogen ist, und wir werden nun zeigen, daß sie ein affiner Unterraum ist, wenn das Gleichungssystem inhomogen ist.

Satz 28.5 (Struktur von $\text{Lös}(A, \mathbf{b})$ als affiner Raum)

Seien $A \in \text{Mat}(m \times n, \mathbb{K})$, $\mathbf{b} \in \mathbb{K}^m$ und sei $\mathbf{c} \in \mathbb{K}^n$ eine Lösung des linearen Gleichungssystems $Ax = \mathbf{b}$. Dann gilt:

$$\text{Lös}(A, \mathbf{b}) = \mathbf{c} + \text{Lös}(A, \mathbf{0}).$$

Beweis: Sei zunächst $\mathbf{y} \in \text{Lös}(A, \mathbf{0})$. Dann gilt:

$$A(\mathbf{c} + \mathbf{y}) = A\mathbf{c} + A\mathbf{y} = \mathbf{b} + \mathbf{0} = \mathbf{b},$$

also ist $\mathbf{c} + \mathbf{y} \in \text{Lös}(A, \mathbf{b})$.

Ist umgekehrt $\mathbf{x} \in \text{Lös}(A, \mathbf{b})$. Dann gilt für $\mathbf{y} := \mathbf{x} - \mathbf{c}$

$$A\mathbf{y} = A(\mathbf{x} - \mathbf{c}) = A\mathbf{x} - A\mathbf{c} = \mathbf{b} - \mathbf{b} = \mathbf{0},$$

also ist $\mathbf{y} \in \text{Lös}(A, \mathbf{0})$. Aber damit ist $\mathbf{x} = \mathbf{c} + \mathbf{y} \in \mathbf{c} + \text{Lös}(A, \mathbf{0})$. \square

Wir wollen nun einen Algorithmus kennenlernen, der es uns erlaubt, die Lösung eines linearen Gleichungssystems $Ax = \mathbf{b}$ in parametrisierter Form zu bestimmen, d. h. eine spezielle Lösung und eine Basis des Lösungsraumes des zugehörigen homogenen Gleichungssystems zu berechnen. Der wichtigste Schritt ist hierbei die Überführung der erweiterten Koeffizientenmatrix $(A | \mathbf{b})$ in reduzierte Zeilen-Stufen-Form.

Lemma 28.6 (Elementare Zeilenoperationen ändern den Lösungsraum nicht.)

Sind $A, A' \in \text{Mat}(m \times n, \mathbb{K})$ und $\mathbf{b}, \mathbf{b}' \in \mathbb{K}^m$ und entsteht die Matrix $(A' | \mathbf{b}')$ aus $(A | \mathbf{b})$ durch elementare Zeilenoperationen, so gilt

$$\text{Lös}(A, \mathbf{b}) = \text{Lös}(A', \mathbf{b}').$$

Beweis: Daß $(A' | \mathbf{b}')$ aus $(A | \mathbf{b})$ durch elementare Zeilenoperationen hervorgeht, bedeutet, daß es eine invertierbare Matrix $S \in \text{Gl}_m(\mathbb{K})$ gibt mit $A' = SA$ und $\mathbf{b}' = S\mathbf{b}$.

Ist nun $\mathbf{c} \in \text{Lös}(A, \mathbf{b})$, dann gilt $A\mathbf{c} = \mathbf{b}$ und damit

$$\mathbf{b}' = S\mathbf{b} = SA\mathbf{c} = A'\mathbf{c}.$$

Also ist $\mathbf{c} \in \text{Lös}(A', \mathbf{b}')$.

Ist andererseits $\mathbf{c} \in \text{Lös}(A', \mathbf{b}')$, dann gilt $A'\mathbf{c} = \mathbf{b}'$ und damit

$$\mathbf{b} = S^{-1}\mathbf{b}' = S^{-1}A'\mathbf{c} = A\mathbf{c}.$$

Also ist $\mathbf{c} \in \text{Lös}(A, \mathbf{b})$. □

Bemerkung 28.7

Aus Lemma 28.6 und Satz 27.7 folgt, daß wir die erweiterte Koeffizientenmatrix eines Gleichungssystems $A\mathbf{x} = \mathbf{b}$ mit $A \in \text{Mat}(\mathbf{m} \times \mathbf{n}, K)$ und $\mathbf{b} \in K^{\mathbf{m}}$ mittels Gauß-Algorithmus in (reduzierte) ZSF überführen können, ohne daß sich die Lösungsmenge ändert.

Wir betrachten deshalb den Fall, daß die Matrix A in ZSF gegeben ist, näher.

Satz 28.8 (Lösbarkeitskriterium für ein LGS mittels Gauß-Algorithmus)

Sei $A \in \text{Mat}(\mathbf{m} \times \mathbf{n}, K)$ eine Matrix in Zeilen-Stufen-Form und $\mathbf{b} \in K^{\mathbf{m}}$. Die erweiterte Koeffizientenmatrix habe die Gestalt

$$(A | \mathbf{b}) = \left(\begin{array}{cccccccccccc|c} 0 & \dots & 0 & \mathbf{a}_{1j_1} & * & \dots & \dots & \dots & \dots & \dots & \dots & * & \mathbf{b}_1 \\ 0 & \dots & \dots & 0 & \dots & 0 & \mathbf{a}_{2j_2} & * & \dots & \dots & \dots & * & \mathbf{b}_2 \\ \vdots & & & & & & & \ddots & & & & & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \mathbf{a}_{rj_r} & * & \dots & * & \mathbf{b}_r \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \mathbf{b}_{r+1} \\ \vdots & & & & & & & & & & & & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & \mathbf{b}_m \end{array} \right) \tag{72}$$

mit Pivots $\mathbf{a}_{ij_i} \neq 0$ für $i = 1, \dots, r$. Dann gilt:

- a. $A\mathbf{x} = \mathbf{b}$ ist genau dann lösbar, wenn $\mathbf{b}_{r+1} = \dots = \mathbf{b}_m = 0$.
- b. Sind $\mathbf{b}_{r+1} = \dots = \mathbf{b}_m = 0$ und gilt $r = n$, so besitzt $A\mathbf{x} = \mathbf{b}$ genau eine Lösung.
- c. Sind $\mathbf{b}_{r+1} = \dots = \mathbf{b}_m = 0$ und ist $r < n$, so hat $A\mathbf{x} = \mathbf{b}$ mehr als eine Lösung. Genauer $\text{Lös}(A, \mathbf{b}) = \mathbf{c} + \text{Lös}(A, 0)$, wobei \mathbf{c} eine spezielle Lösung ist und $\text{Lös}(A, 0)$ die Dimension $n - r$ hat.

Beweis: Die Aussagen folgen aus Satz 28.4, Satz 28.5 und Bemerkung 28.2. □

Bemerkung 28.9 (Parametrisierung von $\text{Lös}(A, \mathbf{b})$)

Wir wollen nun angeben, wie man im Fall c. aus Satz 28.8 aus der Zeilen-Stufen-Form (72) von A die Parametrisierung von $\text{Lös}(A, \mathbf{b})$ als sogenannte affine Abbildung

$$\phi : K^{n-r} \rightarrow \text{Lös}(A, \mathbf{b})$$

herleitet. Sei hierzu $A = r\text{ZSF}(A)$ in reduzierter ZSF gegeben.

Die Parameter x_{j_1}, \dots, x_{j_r} nennen wir die *gebundenen Parameter* und die x_j mit $j \in I := \{1, \dots, n\} \setminus \{j_1, \dots, j_r\}$ die *freien Parameter*. Dies rührt daher, daß sich aus (72) für eine Lösung \mathbf{x} ergibt

$$x_{j_i} = b_i - \sum_{j \in I} a_{ij} x_j, \quad i = 1, \dots, r. \quad (73)$$

D. h. die gebundenen Parameter hängen von den freien Parametern ab.

Identifizieren wir K^{n-r} nun mit K^I und schreiben somit $\mathbf{y} = (y_j \mid j \in I)$ für einen Vektor $\mathbf{y} \in K^{n-r}$, dann ergibt sich die Parametrisierung hieraus als

$$\phi : K^{n-r} \rightarrow \text{Lös}(A, \mathbf{b}) : \mathbf{y} \mapsto \mathbf{c} + f(\mathbf{y}), \quad (74)$$

wobei

$$c_j = \begin{cases} 0, & \text{falls } j \in I, \\ b_i, & \text{falls } j = j_i, \end{cases} \quad (75)$$

und

$$f : K^{n-r} \rightarrow K^n : \mathbf{y} \mapsto (z_1, \dots, z_n)^t, \quad (76)$$

mit

$$z_j = \begin{cases} y_j, & \text{falls } j \in I, \\ -\sum_{k \in I} a_{ik} y_k, & \text{falls } j = j_i. \end{cases} \quad (77)$$

Damit ist f eine lineare Abbildung und deshalb nennt man ϕ affin.

Man beachte, daß \mathbf{c} in diesem Fall eine spezielle Lösung von $A\mathbf{x} = \mathbf{b}$ ist, während $\text{Im}(f) = \text{Lös}(A, 0)$.

A) Der Gauß-Algorithmus zur Lösung eines (LGS)

Algorithmus 28.10 (Algorithmus zur Lösung eines LGS)

INPUT: Die erweiterte Koeffizientenmatrix $(A \mid \mathbf{b})$ eines LGS $A\mathbf{x} = \mathbf{b}$.

OUTPUT: Eine spezielle Lösung \mathbf{c} von $A\mathbf{x} = \mathbf{b}$ und eine Basis B von $\text{Lös}(A, 0)$, sofern das Gleichungssystem lösbar ist.

1. **Schritt:** Berechne eine reduzierte Zeilen-Stufen-Form $(A' \mid \mathbf{b}')$ von $(A \mid \mathbf{b})$ mit $r = \text{rang}(A')$.
2. **Schritt:** Ist $b'_{r+1} \neq 0$, dann ist das LGS nicht lösbar.
3. **Schritt:** Überführe $(A' \mid \mathbf{b}')$ in eine $n \times (n+1)$ -Matrix $(A'' \mid \mathbf{b}'')$ durch Einfügen und Streichen von Nullzeilen, so daß die Pivotelemente anschließend auf der Diagonale der Matrix A'' stehen.
4. **Schritt:** Ersetze jede Null auf der Diagonale von A'' durch -1 .
5. **Schritt:** Die spezielle Lösung ist $\mathbf{c} := \mathbf{b}''$ und die Spalten von A'' , die eine -1 auf der Diagonale haben, sind eine Basis von $\text{Lös}(A, 0)$.

Beispiel 28.11

Wir betrachten das Gleichungssystem:

$$\begin{aligned}x_1 + x_2 + x_3 - 2x_4 &= 1 \\x_1 + x_2 - x_3 &= -1 \\3x_1 + 3x_2 + x_3 - 4x_4 &= 1\end{aligned}\tag{78}$$

In Matrixschreibweise lautet das Gleichungssystem:

$$\begin{pmatrix} 1 & 1 & 1 & -2 \\ 1 & 1 & -1 & 0 \\ 3 & 3 & 1 & -4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

Durch den Gauß-Algorithmus überführen wir die erweiterte Koeffizientenmatrix in reduzierte Zeilen-Stufen-Form:

$$\left(\begin{array}{cccc|c} 1 & 1 & 1 & -2 & 1 \\ 1 & 1 & -1 & 0 & -1 \\ 3 & 3 & 1 & -4 & 1 \end{array} \right) \mapsto \left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right).$$

Wir sehen, daß $\text{rang}(A) = \text{rang}(A | b) = 2$, so daß das Gleichungssystem lösbar ist.

Um die Lösung zu berechnen, fügen wir als zweite Zeile eine Nullzeile ein, um eine 4×5 -Matrix zu erzeugen und die Pivotelemente auf der Diagonalen zu haben, und ersetzen die Nullen auf der Diagonalen anschließend durch -1 :

$$\left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \mapsto \left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \mapsto \left(\begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & -1 & 0 \end{array} \right).$$

Damit erhalten wir die letzte Spalte

$$c = (0, 0, 1, 0)^t$$

als spezielle Lösung von (78) und die Spalten 2 und 4 als Basis

$$B = ((1, -1, 0, 0)^t, (-1, 0, -1, -1)^t)$$

des Lösungsraums $\text{Lös}(A, 0)$ des homogenen Gleichungssystems $Ax = 0$. Insgesamt gilt damit

$$\text{Lös}(A, b) = c + \text{Lös}(A, 0) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + s \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} -1 \\ 0 \\ -1 \\ -1 \end{pmatrix} \mid s, t \in K \right\}.$$

Wollte man eine Parametrisierung wie in Bemerkung 28.9 angeben, so erhält man

$$\phi: \mathbb{K}^2 \rightarrow \text{Lös}(A, \mathbf{b}) \subset \mathbb{K}^4: \begin{pmatrix} x_2 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} x_2 - x_4 \\ -x_2 \\ -x_4 \\ -x_4 \end{pmatrix}.$$

Wir wollen nun einige Algorithmen angeben, denen der Algorithmus zur Lösung eines linearen Gleichungssystems zugrunde liegt.

B) Algorithmus zur Berechnung des Kerns einer linearen Abbildung

Ist $f = f_A$ für eine $m \times n$ -Matrix A , dann ist der Kern von f gerade die Lösungsmenge $\text{Lös}(A, 0)$ des homogenen Gleichungssystems $Ax = 0$.

Algorithmus 28.12 (Kern von f_A)

INPUT: $A \in \text{Mat}(m \times n, \mathbb{K})$.

OUTPUT: Eine Basis von $\text{Ker}(f_A)$.

1. **Schritt:** Bestimme eine Lösung (c, B) von $Ax = 0$ gemäß 28.10.
2. **Schritt:** Gib B als Basis zurück.

Beispiel 28.13

Wir wollen den Kern der Linearen Abbildung $f_A: \mathbb{K}^4 \rightarrow \mathbb{K}^3$ berechnen, die durch die Koeffizientenmatrix

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

in Beispiel 28.11 gegeben ist. Dann gehen wir wie in Beispiel 28.11 vor, wobei wir die Inhomogenität durch den Nullvektor ersetzen oder einfach gänzlich ignorieren können. Die Rechnungen ändern sich nicht und wir erhalten wie dort

$$B = ((1, -1, 0, 0)^t, (-1, 0, -1, -1)^t)$$

als Basis von $\text{Ker}(f_A) = \text{Lös}(A, 0)$.

C) Algorithmus zur Berechnung einer Transformationsmatrix $T_{B'}^B$

Sind $B = (\mathbf{b}_1, \dots, \mathbf{b}_n)$ und $B' = (\mathbf{b}'_1, \dots, \mathbf{b}'_n)$ zwei Basen des \mathbb{K}^n und wollen wir die Transformationsmatrix $T_{B'}^B$ bestimmen, so müssen wir die Basisvektoren in B als Linearkombination der Basisvektoren in B' darstellen und die so erhaltenen Koeffizienten liefern die Spalten von $T_{B'}^B$. Wir müssen also n Gleichungssysteme lösen, bei denen die Koeffizientenmatrix stets $\mathbf{b}'_1, \dots, \mathbf{b}'_n$ als Spaltenvektoren hat und bei denen die Inhomogenitäten durch die Vektoren $\mathbf{b}_1, \dots, \mathbf{b}_n$ gegeben werden. Da die Koeffizientenmatrix sich nicht ändert, können wir die n Gleichungssysteme simultan lösen, indem wir der erweiterten Koeffizientenmatrix gleich alle Vektoren $\mathbf{b}_1, \dots, \mathbf{b}_n$ als zusätzliche Spalten anhängen.

Algorithmus 28.14 (Transformationsmatrix $T_{B'}^B$)

INPUT: Zwei Basen $B = (b_1, \dots, b_n)$ und $B' = (b'_1, \dots, b'_n)$ im K^n .

OUTPUT: Die Transformationsmatrix $T_{B'}^B$.

1. **Schritt:** Schreibe die Vektoren $b'_1, \dots, b'_n, b_1, \dots, b_n$ in dieser Reihenfolge als Spalten in eine Matrix A .
2. **Schritt:** Bringe A auf reduzierte ZSF.
3. **Schritt:** Die letzten n Spalten von $rZSF(A)$ sind $T_{B'}^B$.

Beispiel 28.15

Seien die zwei Basen $B = ((1, 1)^t, (1, -1)^t)$ und $B' = ((1, 2)^t, (-1, 0)^t)$ des \mathbb{R}^2 gegeben.

B'		B	
1	-1	1	1
2	0	1	-1
1	-1	1	1
0	2	-1	-3
1	-1	1	1
0	1	$-\frac{1}{2}$	$-\frac{3}{2}$
1	0	$\frac{1}{2}$	$-\frac{1}{2}$
0	1	$-\frac{1}{2}$	$-\frac{3}{2}$
$\mathbb{1}_2$		$T_{B'}^B$	

D) Algorithmus zur Berechnung einer Matrixdarstellung $M_D^B(f)$

Wir wollen hier angeben, wie man die Matrixdarstellung einer linearen Abbildung $f : K^n \rightarrow K^m$ bezüglich zweier Basen $B = (b_1, \dots, b_n)$ von K^n und $D = (d_1, \dots, d_m)$ von K^m berechnet. Die Grundidee ist ähnlich wie beim Algorithmus zur Berechnung der Transformationsmatrix.

Algorithmus 28.16 (Matrixdarstellung $M_D^B(f)$)

INPUT: Eine lineare Abbildung $f : K^n \rightarrow K^m$, eine Basis $B = (b_1, \dots, b_n)$ von K^n und eine Basis $D = (d_1, \dots, d_m)$ im K^m .

OUTPUT: Die Matrixdarstellung $M_D^B(f)$.

1. **Schritt:** Schreibe die Vektoren $d_1, \dots, d_m, f(b_1), \dots, f(b_n)$ in dieser Reihenfolge als Spalten in eine Matrix A .
2. **Schritt:** Bringe A auf reduzierte ZSF.
3. **Schritt:** Die letzten n Spalten von $rZSF(A)$ sind $M_D^B(f)$.

Beispiel 28.17

Für die Basen $B = ((1, 0, 1)^t, (1, 1, 0)^t, (0, 0, 1)^t)$ des K^3 und $D = ((1, 1)^t, (1, -1)^t)$ des K^2 sowie die lineare Abbildung

$$f : K^3 \rightarrow K^2 : (x, y, z)^t \mapsto (x + y + z, x - z)^t$$

wollen wir die Matrixdarstellung $M_D^B(f)$ berechnen.

D		f(B)		
1	1	2	2	1
1	-1	0	1	-1
<hr/>				
1	1	2	2	1
0	-2	-2	-1	-2
<hr/>				
1	1	2	2	1
0	1	1	$\frac{1}{2}$	1
<hr/>				
1	0	1	$\frac{3}{2}$	0
0	1	1	$\frac{1}{2}$	1
<hr/>				
$\mathbb{1}_2$		$M_D^B(f)$		

Bemerkung 28.18

Natürlich könnte man auch zunächst die Matrixdarstellung $M_E^E(f)$ bezüglich der kanonischen Basen berechnen, da man dazu einfach die Vektoren $f(e_i)$ in die Spalten der Matrix schreiben muß. Analog erhält man T_E^B , indem man die Vektoren von B in die Spalten der Matrix schreibt. Dann muß man nur noch T_D^E mit Hilfe des Algorithmus' zur Berechnung einer Transformationsmatrix bestimmen und kann die Matrizen multiplizieren, um $M_D^B(f)$ zu erhalten.

E) Algorithmus zum Austauschverfahren von Steinitz

Beim Austauschsatz von Steinitz müssen wir die Vektoren in $F = (y_1, \dots, y_r)$, die wir in die Basis $B = (x_1, \dots, x_n)$ hineintauschen wollen, sukzessive als Linearkombination der Basisvektoren in (dem veränderten) B darstellen, d.h. wir müssen immer wieder lineare Gleichungssysteme lösen.

Algorithmus 28.19 (Austauschverfahren von Steinitz)

INPUT: Eine Basis $B = (x_1, \dots, x_n)$ und eine linear unabhängige Familie $F = (y_1, \dots, y_r)$ von Vektoren in $V = \text{Lin}(B) \subseteq K^n$.

OUTPUT: Eine Basis B' von V , die F enthält.

1. Schritt: Für $i = 1, \dots, r$ tue:

- Schreibe die Vektoren in B als Spalten in eine Matrix A.
- Bilde die erweiterte Matrix (A, y_i) .
- Überführe (A, y_i) in reduzierte Zeilen-Stufen-Form und suche in der letzten Spalte den ersten Eintrag ungleich Null.
- Streiche den entsprechenden Vektor aus B und füge y_i als letzten Vektor in B ein.

2. Schritt: Gib B zurück.

Beispiel 28.20

Betrachte die linear unabhängige Familie $F = (y_1, y_2) = ((1, 2, 1)^t, (1, 2, 2)^t)$ und

die Basis $B = (x_1, x_2, x_3) = ((1, 1, 0)^t, (1, 0, 1)^t, (0, 0, 1)^t)$. Wir wollen nun F in B hineintauschen.

Wir bilden die erweiterte Matrix (A, y_1) und überführen sie in reduzierte ZSF:

$$(A, y_1) = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 2 \\ 0 & 1 & 1 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{pmatrix}$$

Da der erste Eintrag in der letzten Spalte nicht Null ist, streichen wir aus B den Vektor x_1 und fügen y_1 als letzten Vektor ein. Wir erhalten die neue Basis

$$B = (x_2, x_3, y_1).$$

Dann bilden wir wieder die erweiterte Matrix (A, y_2) und überführen sie in reduzierte ZSF:

$$(A, y_2) = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 1 & 1 & 1 & 2 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

Der erste Eintrag der letzten Spalte, der nicht Null ist, ist der zweite, mithin müssen wir den zweiten Vektor in B streichen, das ist x_3 , und fügen y_2 am Ende ein. Wir erhalten die Basis

$$B = (x_2, y_1, y_2).$$

Bemerkung 28.21 (Berechnung eines Komplementes oder einer Basis für K^n/U)
Will man ein Komplement eines Unterraumes U in K^n berechnen, so berechnet man zunächst eine Basis von U und tauscht diese anschließend mit Steinitz in die kanonische Basis von K^n . Die verbleibenden Vektoren der kanonischen Basis sind dann eine Basis für ein Komplement und zugleich sind deren Restklassen eine Basis für den Faktorraum K^n/U (siehe Bemerkung 25.19). Der obige Algorithmus erlaubt also auch die Berechnung eines Komplementes und einer Basis eines Faktorraums.

F) Algorithmus zur Berechnung der Gleichungen eines Unterraumes

Wir haben gesehen, daß Unterräume des K^n als Lösungsmengen von homogenen linearen Gleichungssystemen auftauchen. Um etwa den Schnitt zweier Unterräume des K^n zu bestimmen, ist es nützlich, aus dem Erzeugendensystem eines Unterraumes ein Gleichungssystem bestimmen zu können, das den Unterraum beschreibt.

Algorithmus 28.22 (Gleichungen eines Unterraumes)

INPUT: Eine Familie $F = (x_1, \dots, x_m)$ von Vektoren im K^n .

OUTPUT: Eine Matrix $A \in \text{Mat}(k \times n, K)$ mit $\text{Lös}(A, 0) = \text{Lin}(F)$.

1. **Schritt:** Schreibe die Vektoren aus F als Zeilen in eine Matrix $B \in \text{Mat}(m \times n, K)$ und bestimme eine Basis (y_1, \dots, y_k) von $\text{Ker}(f_B) = \text{Lös}(B, 0)$.
2. **Schritt:** Schreibe die y_1, \dots, y_k als Zeilenvektoren in eine Matrix A .
3. **Schritt:** Gib A zurück.

Beispiel 28.23

Finde ein lineares Gleichungssystem $Ax = 0$ mit Lösungsmenge

$$\text{Lös}(A, 0) = \text{Lin}((1, 2, 1)^t, (0, 1, 0)^t) \leq \mathbb{R}^3.$$

Dazu bilden wir die 2×3 -Matrix

$$B = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

und berechnen ihren Kern:

$$\text{rZSF}(B, 0) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Somit ist der Vektor in der dritten Spalte eine Basis von $\text{Lös}(B, 0)$ und wir erhalten

$$A = (1 \ 0 \ -1).$$

G) Algorithmus zur Berechnung des Durchschnitts zweier Unterräume

Abschließend sind wir nun in der Lage, einen Algorithmus anzugeben, mittels dessen sich eine Basis des Schnitts zweier Unterräume des K^n ermitteln läßt.

Algorithmus 28.24 (Durchschnitt zweier Unterräume)

INPUT: Zwei Familien F und G von Vektoren in K^n .

OUTPUT: Eine Basis des Schnitts von $\text{Lin}(F)$ und $\text{Lin}(G)$.

1. **Schritt:** Bestimme Matrizen A und A' gemäß 28.22, so daß $\text{Lin}(F) = \text{Lös}(A, 0)$ und $\text{Lin}(G) = \text{Lös}(A', 0)$.
2. **Schritt:** Bilde aus den Zeilen von A und A' eine gemeinsame Matrix A'' .
3. **Schritt:** Bestimme eine Basis B von $\text{Ker}(f_{A''}) = \text{Lös}(A'', 0)$ gemäß 28.12 und gib B zurück.

Beispiel 28.25

Wir wollen den Durchschnitt der Unterräume

$$U = \text{Lin}((1, 2, 1)^t, (0, 1, 0)^t)$$

und

$$U' = \{(x, y, z)^t \in \mathbb{R}^3 \mid x + y + z = 0\}$$

berechnen. Der zweite Unterraum ist bereits als Lösungsmenge eines Gleichungssystems mit Koeffizientenmatrix $A' = (1 \ 1 \ 1)$ gegeben. Für den ersten Unterraum haben wir eine solche Darstellung $\text{Lös}(A, 0)$ bereits in Beispiel 28.23 berechnet. Wir bilden eine neue Matrix A''

$$A'' = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 1 & 1 \end{pmatrix}$$

aus A und A' und lösen das zugehörige homogene Gleichungssystem

$$(A'', 0) = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Damit ist die dritte Spalte

$$B = ((-1, 2, -1)^t)$$

eine Basis von $U \cap U'$.

H) Beispiele linearer Gleichungssysteme in Anwendung

Wir geben jetzt einige Beispiele von Gleichungssystemen, die zum Teil aus Anwendungen kommen. Wir werden diese nicht in der Vorlesung besprechen. Sie sollen dem interessierten Leser die große praktische Bedeutung linearer Gleichungssysteme illustrieren.

Beispiel 28.26 (Wie alt ist der Vater?)

Ein Vater hat einen Sohn und eine Tochter. Der Vater ist viermal so alt wie sein Sohn und der Sohn ist fünf Jahre älter als seine Schwester. In fünf Jahren sind Vater und Sohn zusammen sechsmal so alt wie die Tochter.

Wie alt sind Vater, Sohn und Tochter?

Das lineare Gleichungssystem mit v = Alter des Vaters, s = Alter des Sohnes, und t = Alter der Tochter lautet:

$$v = 4s, \quad s = t + 5, \quad (v + 5) + (s + 5) = 6(t + 5).$$

Das Gleichungssystem schreiben wir systematisch folgendermaßen auf:

$$\begin{aligned} v - 4s + 0 \cdot t &= 0, \\ 0 \cdot v + s - t &= 5, \\ v + s - 6t &= 20. \end{aligned}$$

Dies ist ein lineares Gleichungssystem in den Unbestimmten v , s , t .

Die Lösung mit Hilfe des Gaußschen Algorithmus geht wie folgt:

$$\begin{aligned} \left(\begin{array}{ccc|c} 1 & -4 & 0 & 0 \\ 0 & 1 & -1 & 5 \\ 1 & 1 & -6 & 20 \end{array} \right) &\mapsto \left(\begin{array}{ccc|c} 1 & -4 & 0 & 0 \\ 0 & 1 & -1 & 5 \\ 0 & 5 & -6 & 20 \end{array} \right) \mapsto \left(\begin{array}{ccc|c} 1 & -4 & 0 & 0 \\ 0 & 1 & -1 & 5 \\ 0 & 0 & 1 & 5 \end{array} \right) \\ &\mapsto \left(\begin{array}{ccc|c} 1 & -4 & 0 & 0 \\ 0 & 1 & 0 & 10 \\ 0 & 0 & 1 & 5 \end{array} \right) \mapsto \left(\begin{array}{ccc|c} 1 & 0 & 0 & 40 \\ 0 & 1 & 0 & 10 \\ 0 & 0 & 1 & 5 \end{array} \right) \end{aligned}$$

Als Lösung erhalten wir also: $t = 5$, $s = 10$, $v = 40$, d. h. der Vater ist 40 Jahre alt, sein Sohn zehn und seine Tochter fünf.

Beispiel 28.27 (Schnitt zweier Ebenen)

Wir definieren eine Ebene im \mathbb{R}^3 als Lösungsmenge einer linearen Gleichung

$$E : \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \mathbf{a}_3 x_3 = \mathbf{b}$$

mit $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{b} \in \mathbb{R}$ und $\mathbf{a}_i \neq 0$ für mindestens ein i .

Dies stimmt mit der Anschauung überein (sind alle \mathbf{a}_i und \mathbf{b} gleich 0, so erhalten wir als Lösungsmenge den ganzen \mathbb{R}^3 , sind alle $\mathbf{a}_i = 0$ und $\mathbf{b} \neq 0$, so ist die Lösungsmenge leer).

Um den Schnitt der beiden Ebenen, die durch die Gleichungen $E_1 : x_1 + x_2 + 2x_3 = 2$ und $E_2 : x_1 + x_3 = 4$ gegeben sind, zu bestimmen, müssen wir also das Gleichungssystem aus diesen beiden Gleichungen lösen, wobei wir wie in Abschnitt A) beschrieben vorgehen:

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

Wir erhalten als Lösungsmenge

$$E_1 \cap E_2 = \begin{pmatrix} 4 \\ -2 \\ 0 \end{pmatrix} + \mathbb{R} \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}.$$

Dies ist offensichtlich die Parameterdarstellung einer Geraden im \mathbb{R}^3 durch die Punkte $(4, -2, 0)^t$ und $(5, -1, -1)^t$.

Beispiel 28.28 (Schnitt zweier Ebenen)

Im allgemeinen werden sich zwei Ebenen, E_1, E_2 , im \mathbb{R}^3 in einer Geraden schneiden, in Spezialfällen können die Ebenen aber parallel sein ($E_1 \cap E_2 = \emptyset$) oder übereinstimmen ($E_1 = E_2$).

Sei E_1 die Ebene

$$E_1 : x_1 + x_2 + 2x_3 = 3$$

und E_2 eine beliebige Ebene

$$E_2 : \mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \mathbf{a}_3 x_3 = \mathbf{b}.$$

Wir wollen feststellen für welche $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{b}$ entweder $E_1 \cap E_2$ eine Gerade, leer oder E_1 ist:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{b} \end{array} \right) \mapsto \left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & \mathbf{a}_2 - \mathbf{a}_1 & \mathbf{a}_3 - 2\mathbf{a}_1 & \mathbf{b} - 3\mathbf{a}_1 \end{array} \right).$$

Die letzte Gleichung lautet

$$(\mathbf{a}_2 - \mathbf{a}_1)x_2 + (\mathbf{a}_3 - 2\mathbf{a}_1)x_3 = \mathbf{b} - 3\mathbf{a}_1.$$

Ein wenig Überlegung liefert (da die Lösungsmenge der ersten Gleichung E_1 ist, und da die Lösungsmenge der zweiten Gleichung unabhängig von x_1 ist):

$$E_1 \cap E_2 = \emptyset \Leftrightarrow a_2 - a_1 = a_3 - 2a_1 = 0, (b - 3a_1) \neq 0, \quad (79)$$

$$E_1 = E_2 \Leftrightarrow a_2 - a_1 = a_3 - 2a_1 = b - 3a_1 = 0. \quad (80)$$

In allen anderen Fällen ist $E_1 \cap E_2$ eine Gerade.

Im Fall $E_1 = E_2$ haben wir wieder ein Gleichungssystem (80) mit drei Gleichungen in den vier Unbestimmten a_1, a_2, a_3, b zu lösen:

$$\begin{aligned} \left(\begin{array}{cccc|c} -1 & 1 & 0 & 0 & 0 \\ -2 & 0 & 1 & 0 & 0 \\ -3 & 0 & 0 & 1 & 0 \end{array} \right) &\mapsto \left(\begin{array}{cccc|c} -1 & 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & -3 & 0 & 1 & 0 \end{array} \right) \mapsto \left(\begin{array}{cccc|c} -1 & 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & -\frac{3}{2} & 1 & 0 \end{array} \right) \\ &\mapsto \left(\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & -\frac{2}{3} & 0 \end{array} \right) \mapsto \left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 1 & 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{2}{3} & 0 \end{array} \right) \end{aligned}$$

Als Lösung ergibt sich $a_1 = -\frac{b}{3}$, $a_2 = \frac{b}{3}$ und $a_3 = \frac{2b}{3}$, oder kurz

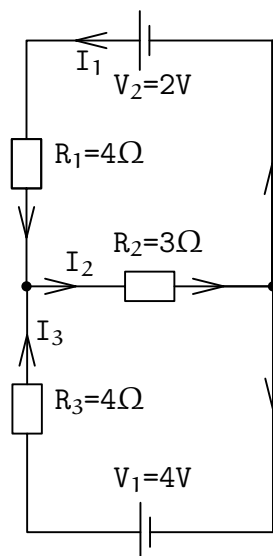
$$(a_1, a_2, a_3, b) = t \cdot (-1, 1, 2, 3)$$

mit $t \in \mathbb{R}$ beliebig. Daraus können wir aber alle drei Fälle ablesen:

$E_1 = E_2$ genau dann, wenn die Gleichung von E_2 ein Vielfaches $\neq 0$ der Gleichung von E_1 ist; $E_1 \cap E_2 = \emptyset$ genau dann, wenn der Koeffizientenvektor (a_1, a_2, a_3) ein Vielfaches $\neq 0$ des Koeffizientenvektors von E_1 ist, aber die rechte Seite b von E_2 nicht das gleiche Vielfache der rechten Seite von E_1 ist; und $E_1 \cap E_2$ ist eine Gerade in allen anderen Fällen.

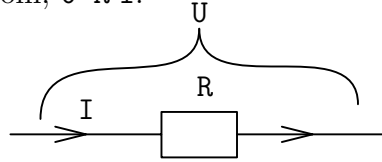
Beispiel 28.29 (Elektrische Netzwerke)

In einem einfachen elektrischen Netzwerk, wie z. B.

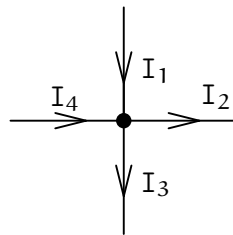


bezeichnet man mit U die Spannung, mit I den Strom und mit R den Widerstand, gemessen in Volt (V), Ampere (A) und Ohm (Ω) respektive. Dabei gelten folgende Gesetze:

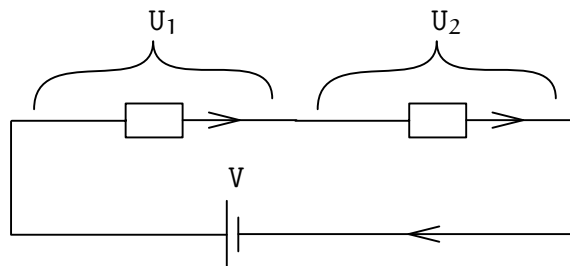
- *Ohmsches Gesetz*: Der Spannungsabfall über einen Widerstand ist das Produkt von Widerstand und Strom, $U=R \cdot I$.



- *1. Kirchhoffsches Gesetz (Knotengleichung)*: Die Summe der in einen Knoten hineinfließenden Ströme ist gleich der Summe der hinausfließenden Ströme. Beispiel: $I_1+I_4=I_2+I_3$



- *2. Kirchhoffsches Gesetz (Maschengleichung)*: Die Summe der Spannungsverluste in einem geschlossenen Kreis ist gleich der Gesamtspannung in einem Kreis. Beispiel: $V=U_1+U_2$



Im obigen Beispiel stellt man mit Hilfe der drei Gesetze das folgende lineare Gleichungssystem auf:

$$\begin{aligned} I_1 + I_3 &= I_2, & (\text{Knotengleichung}) \\ 4I_1 + 3I_2 &= 2, & (1. \text{ Maschengleichung}) \\ 4I_3 + 3I_2 &= 4. & (2. \text{ Maschengleichung}) \end{aligned}$$

Wir erhalten das folgende Gleichungssystem:

$$\begin{aligned} \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 4 & 3 & 0 & 2 \\ 0 & 3 & 4 & 4 \end{array} \right) &\mapsto \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 7 & -4 & 2 \\ 0 & 3 & 4 & 4 \end{array} \right) &\mapsto \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 7 & -4 & 2 \\ 0 & 0 & 40 & 22 \end{array} \right) \\ &\mapsto \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & -\frac{4}{7} & \frac{2}{7} \\ 0 & 0 & 1 & \frac{11}{20} \end{array} \right) &\mapsto \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{20} \\ 0 & 1 & 0 & \frac{3}{5} \\ 0 & 0 & 1 & \frac{11}{20} \end{array} \right) \end{aligned}$$

woraus sich die folgende Lösung ergibt:

$$I_3 = \frac{11}{20}, I_2 = \frac{3}{5} \text{ und } I_1 = \frac{1}{20}.$$

Beispiel 28.30 (Kubische Splines)

Im “Computer aided geometric design” (CAGD) werden zum Design von Flächen und Kurven (z. B. im Automobil- oder Flugzeugbau) Flächen- und Kurvenstücke verwendet (meist durch sogenannte kubische Splines realisiert), die dann an den Endpunkten oder Randkurven glatt zusammenpassen müssen. Am bekanntesten sind die Bézier-Kubiken, die von dem französischen Auto-Designer bei Renault, P. Bézier, eingeführt wurden (diese werden heute z. B. auch in der Text-Beschreibungssprache PostScript verwendet).

Ein typisches Problem ist z.B. die Bestimmung einer kubischen Parabel

$$f(x) = ax^3 + bx^2 + cx + d$$

durch zwei Punkte (x_1, y_1) , (x_2, y_2) in der Ebene mit vorgegebener Steigung m_1 in (x_1, y_1) und m_2 in (x_2, y_2) .

Für $(x_1, y_1) = (0, 2)$, $(x_2, y_2) = (4, 0)$, $m_1 = -3$, $m_2 = -3$ ergibt sich aus

$$f'(x) = 3ax^2 + 2bx + c$$

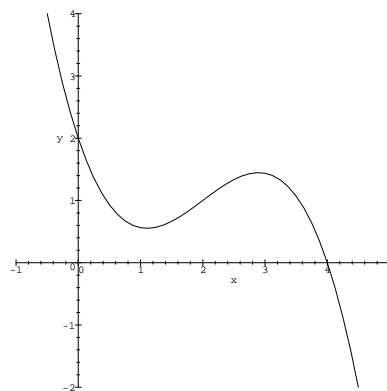
und

$$f(0) = 2, f(4) = 0, f'(0) = -3 \text{ und } f'(4) = -3$$

das lineare Gleichungssystem

$$\begin{aligned} d &= 2, \\ 64a + 16b + 4c + d &= 0, \\ c &= -3, \\ 48a + 8b + c &= -3, \end{aligned}$$

also $d = 2$, $c = -3$, $6a + b = 0$, $32a + 8b = 5$, und damit $a = -\frac{5}{16}$ und $b = \frac{15}{8}$. Die Kurve $y = -\frac{5}{16}x^3 + \frac{15}{8}x^2 - 3x + 2$ hat etwa die folgende Gestalt



Die Aufgabe ist, wie leicht zu sehen ist, stets lösbar und daher können kubische Splines stückweise definiert und glatt aneinander gesetzt werden.

Beispiel 28.31 (Leontieff-Modell)

Die folgende Planungsaufgabe zeigt, daß durchaus Gleichungen mit vielen Veränderlichen in der Praxis auftauchen.

Ein Konzern besitzt n Fabriken F_1, \dots, F_n , in der Fabrik F_i wird das Produkt P_i hergestellt.

Zur Produktion einer Einheit von P_k werden a_{jk} Einheiten von P_j benötigt; wir nehmen an $a_{ii} = 0$.

Am Ende eines Produktionszyklus sind x_k Einheiten von P_k hergestellt, $k = 1, \dots, n$; wir haben also einen Produktionsvektor $x = (x_1, \dots, x_n)$.

Zur Herstellung von $x = (x_1, \dots, x_n)$ werden

$$\sum_{k=1}^n a_{jk} x_k = a_{j1} x_1 + \dots + a_{jn} x_n$$

Einheiten von P_j verbraucht.

Für den Markt verbleiben damit

$$y_j = x_j - \sum_{k=1}^n a_{jk} x_k$$

Einheiten von P_j .

Die Planungsaufgabe lautet nun:

Der Mehrbedarf $y = (y_1, \dots, y_n)$ ist vorgegeben. Gesucht ist ein Produktionsvektor $x = (x_1, \dots, x_n)$ mit

$$\begin{aligned} x_1 - (a_{11}x_1 + \dots + a_{1n}x_n) &= y_1 \\ \vdots & \\ x_n - (a_{n1}x_1 + \dots + a_{nn}x_n) &= y_n. \end{aligned}$$

Also ist ein lineares Gleichungssystem zu lösen. Allerdings, und das macht das Problem schwerer, ist zu beachten, daß alle $x_i \geq 0$ sein müssen (natürlich sind auch die y_j und die $a_{jk} \geq 0$).

(Das Modell heißt Leontieff-Modell und ist nach Vassili Leontieff benannt, der 1973 den Nobelpreis für Wirtschaftswissenschaften erhielt.)

Ein einfaches Beispiel mit zwei Fabriken, Verbrauchsmatrix

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

und zunächst unbestimmtem Mehrbedarf (y_1, y_2) liefert das Gleichungssystem

$$\begin{aligned} \left(\begin{array}{cc|c} 1 & -\frac{1}{2} & y_1 \\ -\frac{1}{3} & 1 & y_2 \end{array} \right) &\mapsto \left(\begin{array}{cc|c} 1 & -\frac{1}{2} & y_1 \\ 0 & \frac{5}{6} & \frac{1}{3}y_1 + y_2 \end{array} \right) \mapsto \left(\begin{array}{cc|c} 1 & -\frac{1}{2} & y_1 \\ 0 & 1 & \frac{2}{5}y_1 + \frac{6}{5}y_2 \end{array} \right) \\ &\mapsto \left(\begin{array}{cc|c} 1 & 0 & \frac{6}{5}y_1 + \frac{3}{5}y_2 \\ 0 & 1 & \frac{2}{5}y_1 + \frac{6}{5}y_2 \end{array} \right) \end{aligned}$$

und damit $x_1 = \frac{6}{5}y_1 + \frac{3}{5}y_2$, $x_2 = \frac{2}{5}y_1 + \frac{6}{5}y_2$.

Beispiel 28.32 (Finde ein Gleichungssystem zu gegebener Lösung.)

Ein Gleichungssystem besitze die spezielle Lösung $(1, 0, 1)^t$ und das zugehörige homogene System besitze $(1, 1, 1)^t$ als Lösung und habe den Rang zwei. Finde ein Gleichungssystem, das diese Bedingungen erfüllt.

Da die Lösungen Vektoren im \mathbb{R}^3 sind, ist es ein System in drei Variablen.

Da der Rang zwei ist, hat die Zeilen-Stufen-Form zwei Zeilen. Da die Lösungsmenge nicht von der Form abhängt, können wir das System in Zeilen-Stufen-Form annehmen:

Problem: Finde eine Gerade im \mathbb{R}^3 , die selbst durch $(1, 0, 1)^t$ geht und für die die in den Nullpunkt verschobene Gerade durch $(1, 1, 1)^t$ geht.

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 &= b_1, \\ a_{22}x_2 + a_{23}x_3 &= b_2. \end{aligned}$$

$(1, 0, 1)^t$ ist Lösung:

$$a_{11} + a_{13} = b_1, \tag{1}$$

$$a_{23} = b_2. \tag{2}$$

$(1, 1, 1)^t$ ist Lösung des homogenen Systems:

$$a_{11} + a_{12} + a_{13} = 0, \tag{3}$$

$$a_{22} + a_{23} = 0. \tag{4}$$

Das zugehörige lineare Gleichungssystem in $a_{11}, a_{12}, a_{13}, a_{22}, a_{23}, b_1, b_2$ lautet:

$$\begin{array}{cccccc} & a_{11} & a_{12} & a_{13} & a_{22} & a_{23} & b_1 & b_2 & & a_{11} & a_{12} & a_{13} & a_{22} & a_{23} & b_1 & b_2 \\ (1) & \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 0 & 0 \end{array} \right) & & & & & & & \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 0 \end{array} \right) \\ (3) & \left(\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & 0 & 0 \end{array} \right) & & & & & & & \left(\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \end{array} \right) \\ (4) & \left(\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right) & & & & & & & \left(\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right) \\ (2) & \left(\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right) & & & & & & & \left(\begin{array}{cc|c} 0 & 0 & 0 \\ 0 & 1 & 0 \end{array} \right) \end{array} \mapsto \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{array} \right)$$

Das System hat unendlich viele Lösungen, und da der Rang 2 sein soll, muß $a_{22} \neq 0$ und damit auch $a_{23} = -a_{22} \neq 0$ sein.

Wir wählen

$$a_{22} = 1 \Rightarrow a_{23} = b_2 = -1,$$

$$a_{12} = 1 \Rightarrow b_1 = -1,$$

$$a_{11} = 1 \Rightarrow a_{13} = -2.$$

Also ist

$$x_1 + x_2 - 2x_3 = -1,$$

$$x_2 - x_3 = -1$$

ein geeignetes Gleichungssystem.

I) Ringe und Moduln

Bemerkung 28.33 (Ringe und Moduln)

Man kann lineare Gleichungssysteme in analoger Weise über kommutativen Ringen mit Eins betrachten. Es bleibt richtig, daß die Lösungsmenge eines homogenen LGS ein Modul $\text{Lös}(A, 0) = \text{Ker}(f_A)$, und daß $Ax = b$ genau dann lösbar ist, wenn $b \in \text{Im}(f_A)$. Auch die Strukturaussage $\text{Lös}(A, b) = c + \text{Lös}(A, 0)$ in Satz 28.5 bleibt wahr. Alle Aussagen, die den Rang einer Matrix verwenden, sind jedoch nicht mehr richtig. Außerdem kann sich die Lösungsmenge eines Gleichungssystems ändern, wenn man Zeilen der Matrix mit einer Konstanten multipliziert, die kein Inverses im Ring besitzt. Es ist also Vorsicht geboten, wenn man den abgewandelten Gauß-Algorithmus, der ohne Division auskommt, verwenden will, um die erweiterte Matrix auf ZSF zu bringen. Ist der Ring ein sogenannter Integritätsbereich, wie etwa die ganzen Zahlen \mathbb{Z} , so entstehen dadurch keine wirklichen Probleme, da aus $\lambda \cdot x = 0$ mit $\lambda \neq 0$ immer noch $x = 0$ folgt.

Aufgaben

Aufgabe 28.34

Berechne die Lösungsmenge des folgenden Gleichungssystems, sofern es lösbar ist:

$$-x + 6y + 2z = 4$$

$$2x - 2y - z = 2$$

$$3x - 4y - 2z = 1$$

Aufgabe 28.35

Für welche $a, b \in \mathbb{R}$ besitzt das lineare Gleichungssystem

$$ax + z = ab$$

$$-2x + by + az = -b$$

$$by + (a+1)z = b$$

außer $(b, 1, 0)$ noch weitere Lösungen. Bestimme sie.

Aufgabe 28.36

Bestimme die Lösungsmenge des folgenden linearen Gleichungssystems über \mathbb{R} in Abhängigkeit vom Parameter $t \in \mathbb{R}$:

$$\begin{aligned}x + y + z &= 1 \\ty + z &= 1 \\tx + ty + z &= 1 + t\end{aligned}$$

Aufgabe 28.37

Bestimme eine Basis des Kerns und des Bildes von f_A mit

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ -1 & -1 & 0 & 1 & 1 \\ -1 & -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & -1 & 0 \end{pmatrix} \in \text{Mat}_5(\mathbb{R}).$$

Aufgabe 28.38

Es sei $U = \{(x + y, y, y - x)^t \mid x, y \in \mathbb{R}\}$ und $U' = \{(x, y, z)^t \in \mathbb{R}^3 \mid z = 2x + y\}$. Bestimme Basen von $U + U'$, $U \cap U'$, \mathbb{R}^3/U und \mathbb{R}^3/U' .

Aufgabe 28.39

Bestimme eine Basis für $U \cap U'$ mit

$$U = \langle (2, -1, 1, -1)^t, (1, -2, 2, 1)^t, (3, -1, 0, 2)^t \rangle \leq \mathbb{R}^4$$

und

$$U' = \langle (3, -2, 3, 8)^t, (2, 1, -5, 3)^t \rangle \leq \mathbb{R}^4.$$

Aufgabe 28.40

Wir betrachten

$$B = \langle (1, 1, 1, 1)^t, (-1, 0, 0, 1)^t, (0, -1, 0, 1)^t, (0, 0, -1, 1)^t \rangle$$

und

$$D = \langle (1, 1, 0)^t, (0, 1, 1)^t, (0, 0, 1)^t \rangle.$$

- Zeige, dass B eine Basis des \mathbb{R}^4 und D eine Basis des \mathbb{R}^3 ist.
- Bestimme $M_D^B(f)$ für $f: \mathbb{R}^4 \rightarrow \mathbb{R}^3: (x_1, x_2, x_3, x_4)^t \mapsto (x_1 - x_2, x_3, x_2 + x_4)^t$.
- Bestimme umgekehrt die Funktionsvorschrift für $g \in \text{Hom}_{\mathbb{R}}(\mathbb{R}^4, \mathbb{R}^3)$ mit

$$M_D^B(g) = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 2 & 1 & 3 & 1 \\ 0 & -1 & 2 & 0 \end{pmatrix}.$$

Aufgabe 28.41

Es sei V ein \mathbb{R} -Vektorraum, $B = (x_1, x_2, x_3)$ eine Basis von V und $B' = (y_1, y_2, y_3)$ mit $y_1 = x_1 + x_3$, $y_2 = x_1 + x_2$ und $y_3 = x_1 + x_2 + x_3$.

- a. Zeige, dass B' eine Basis von V ist.
- b. Bestimme $M_{B'}^{B'}(f)$, wobei $f \in \text{End}_{\mathbb{R}}(V)$ gegeben ist durch

$$M_{B'}^B(f) = \begin{pmatrix} a & 0 & b \\ -b & a & a \\ a & b & b \end{pmatrix} \text{ mit } a, b \in \mathbb{R}.$$

Aufgabe 28.42

Seien $B = ((1, 1, 1)^t, (1, 1, 0)^t, (1, 0, -1)^t)$ und $B' = ((2, 1)^t, (1, 1)^t)$. E bzw. E' seien die kanonischen Basen des \mathbb{R}^3 bzw. des \mathbb{R}^2 . Ferner sei $f \in \text{Hom}_{\mathbb{K}}(\mathbb{R}^3, \mathbb{R}^2)$ gegeben durch $f((x, y, z)^t) = (x - y + z, 2x + y)^t$.

- a. Zeige, dass B und B' Basen des \mathbb{R}^3 bzw. des \mathbb{R}^2 sind.
- b. Bestimme $M_{E'}^E(f)$.
- c. Bestimme $M_{B'}^B(f)$ sowie die Transformationsmatrizen T_E^B und $T_{B'}^{E'}$ mit $T_{B'}^{E'} \cdot M_{E'}^E(f) \cdot T_E^B = M_{B'}^B(f)$.

§ 29 Die symmetrische Gruppe

Definition 29.1 (Permutationen)

Eine bijektive Abbildung $\sigma : \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$ nennen wir eine *Permutation* der Menge $\{1, \dots, n\}$, und wir bezeichnen mit

$$\mathbb{S}_n = \text{Sym}(\{1, \dots, n\}) = \{ \sigma : \{1, \dots, n\} \longrightarrow \{1, \dots, n\} \mid \sigma \text{ ist bijektiv} \}$$

die Menge aller Permutation der Menge $\{1, \dots, n\}$.

Eine Permutation $\sigma \in \mathbb{S}_n$ kann durch eine *Wertetabelle* der folgenden Form beschrieben werden:

$$\begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix}$$

bzw.

$$\begin{pmatrix} a_1 & a_2 & \dots & a_n \\ \sigma(a_1) & \sigma(a_2) & \dots & \sigma(a_n) \end{pmatrix},$$

falls a_1, \dots, a_n irgendeine Anordnung der Zahlen $1, \dots, n$ ist.

Bemerkung 29.2 (Die symmetrische Gruppe \mathbb{S}_n)

In Beispiel 7.2 haben wir gezeigt, daß \mathbb{S}_n mit der Komposition von Abbildungen eine Gruppe ist. Wir nennen (\mathbb{S}_n, \circ) die *symmetrische Gruppe* vom Grad n . Die \mathbb{S}_n enthält genau $n!$ Elemente.

Beispiel 29.3

Die Gruppe \mathbb{S}_n ist für $n \geq 3$ nicht abelsch. In \mathbb{S}_3 gilt für die Permutationen

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in \mathbb{S}_3$$

nämlich

$$\begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \circ \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

Beachte, daß es bei dem Schema nicht darauf ankommt, in welcher Reihenfolge die Zahlen von 1 bis n in der ersten Zeile stehen. Es gilt etwa:

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

Es empfiehlt sich aber der Übersichtlichkeit halber für gewöhnlich, die Ziffern in aufsteigender Reihenfolge anzuordnen.

Bemerkung 29.4 (Invertieren einer Permutation)

Die oben eingeführte Darstellung einer Permutation hat den angenehmen Nebeneffekt, daß man das Inverse der Permutation leicht angeben kann, indem man einfach die beiden Zeilen vertauscht. Sprich, für eine Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & \dots & n \\ \sigma(1) & \sigma(2) & \dots & \sigma(n) \end{pmatrix} \in \mathbb{S}_n$$

ist das Inverse σ^{-1} gegeben durch

$$\sigma^{-1} = \begin{pmatrix} \sigma(1) & \sigma(2) & \dots & \sigma(n) \\ 1 & 2 & \dots & n \end{pmatrix}.$$

Definition 29.5 (Zyklen und Transpositionen)

- a. Sei $\{1, \dots, n\} = \{a_1, \dots, a_k\} \cup \{b_1, \dots, b_{n-k}\}$, $k \geq 2$, und

$$\sigma = \begin{pmatrix} a_1 & a_2 & \dots & a_{k-1} & a_k & b_1 & \dots & b_{n-k} \\ a_2 & a_3 & \dots & a_k & a_1 & b_1 & \dots & b_{n-k} \end{pmatrix} \in \mathbb{S}_n,$$

so heißt σ ein **k-Zyklus**, und wir sagen, daß sie die Zahlen a_1, \dots, a_k *zyklisch vertauscht*. Die Abbildungsvorschrift eines solchen k-Zyklus läßt sich deutlich kompakter durch das folgende einzeilige Schema repräsentieren:

$$\sigma = (a_1 \dots a_k). \quad (81)$$

- b. Ein 2-Zyklus wird auch eine **Transposition** genannt. Eine Transposition $\tau = (i \ j)$ ist mithin eine Permutation, die nur die zwei Zahlen i und j miteinander vertauscht, alle anderen aber fest läßt.
- c. Das neutrale Element von \mathbb{S}_n , per definitionem $\text{id}_{\{1, \dots, n\}}$, wollen wir der Einfachheit halber mit id bezeichnen.

Bemerkung 29.6

- a. Die Interpretation der Schreibweise in Gleichung (81) ist offensichtlich, das erste Element a_1 wird auf das zweite a_2 abgebildet, das zweite auf das dritte, und so weiter, bis schließlich das letzte, nämlich a_k , auf das erste, das heißt auf a_1 , abgebildet wird – der *Kreis* schließt sich. Beachte hierbei, daß die Zyklen $(a_1 \dots a_k)$, $(a_k a_1 \dots a_{k-1})$, etc. übereinstimmen! Um diese Mehrdeutigkeit zu vermeiden, empfiehlt es sich, einen Zyklus stets mit der kleinsten der Zahlen a_1, \dots, a_k zu beginnen.

Bisher haben wir k-Zyklen nur für $k \geq 2$ definiert. Wir können nun auch 1-Zyklen, etwa (1) oder (3), zulassen und definieren diese in natürlicher Weise als die Identität.

- b. Die Permutationen

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix} \in \mathbb{S}_4 \quad \text{und} \quad \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 1 & 3 & 2 & 5 \end{pmatrix} \in \mathbb{S}_5$$

sind jeweils 3-Zyklen, die die Zahlen 1, 4, 2 zyklisch vertauschen. In der oben eingeführten Zykelschreibweise gilt

$$\sigma = (1 \ 4 \ 2) \quad \text{und} \quad \pi = (1 \ 4 \ 2).$$

Damit wird der Nachteil dieser Schreibweise gegenüber dem zweizeiligen Schema deutlich – weder der Definitionsbereich noch der Wertebereich lassen sich aus der Zykelschreibweise eindeutig ablesen. Aber diesen Preis sind wir für die gewonnene *Übersichtlichkeit* gerne bereit zu zahlen. Denn einerseits ist in Anwendungen meist zweifelsfrei bekannt, was n ist, und andererseits ist die

wesentliche Information für uns letztlich, welche Zahlen durch die Permutation vertauscht werden, und nicht, welche unbewegt bleiben.

- c. Für eine Transposition $\tau \in \mathbb{S}_n$ gilt $\tau^{-1} = \tau$, also $\tau^2 = \text{id}$.
- d. Für kleine Werte n ist \mathbb{S}_n sehr übersichtlich, für große Werte n wird \mathbb{S}_n jedoch riesig. $\mathbb{S}_1 = \{\text{id}\}$ und $\mathbb{S}_2 = \{\text{id}, (1\ 2)\}$. $\mathbb{S}_3 = \{\text{id}, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$ hat schon sechs Elemente, \mathbb{S}_4 gar 24 und \mathbb{S}_{60} ungefähr 10^{82} . Letztere Zahl entspricht in etwa der angenommenen Anzahl der Nukleone des Universums.

Satz 29.7 (Zyklenzerlegung und Signum)

- a. Jede Permutation $\sigma \in \mathbb{S}_n$ läßt sich als Produkt von disjunkten Zyklen schreiben.
- b. Jede Permutation $\sigma \in \mathbb{S}_n$ läßt sich als Produkt von Transpositionen schreiben.
- c. Jede Permutation $\sigma \in \mathbb{S}_n$ läßt sich als Produkt von Transpositionen benachbarter Zahlen schreiben.
- d. Es gibt genau einen Gruppenhomomorphismus, das Signum genannt,

$$\text{sgn} : (\mathbb{S}_n, \circ) \longrightarrow (\{1, -1\}, \cdot)$$

mit $\text{sgn}(\tau) = -1$ für jede Transposition $\tau \in \mathbb{S}_n$. Insbesondere gilt, ist $\sigma \in \mathbb{S}_n$ ein Produkt von k Transpositionen, dann gilt mithin

$$\text{sgn}(\sigma) = (-1)^k.$$

- e. Für $\sigma \in \mathbb{S}_n$ gilt $\text{sgn}(\sigma) = \text{sgn}(\sigma^{-1})$.
- f. Ist $\mathbb{A}_n = \{\sigma \in \mathbb{S}_n \mid \text{sgn}(\sigma) = 1\}$ und ist $\tau = (i\ j)$ eine Transposition, so gilt

$$\mathbb{S}_n = \mathbb{A}_n \cup \mathbb{A}_n \tau$$

wobei $\mathbb{A}_n \tau = \{\sigma \circ \tau \mid \sigma \in \mathbb{A}_n\}$.

Bemerkung 29.8

- a. Daß die Abbildung sgn ein Gruppenhomomorphismus ist, heißt

$$\text{sgn}(\sigma \cdot \pi) = \text{sgn}(\sigma) \cdot \text{sgn}(\pi)$$

für alle $\sigma, \pi \in \mathbb{S}_n$. Ist also $\sigma = \tau_1 \circ \dots \circ \tau_k \in \mathbb{S}_n$ ein Produkt von k Transpositionen, dann gilt

$$\text{sgn}(\sigma) = \text{sgn}(\tau_1) \cdot \dots \cdot \text{sgn}(\tau_k) = (-1)^k.$$

- b. Bewiesen werden die Aussagen in der Vorlesung Algebraische Strukturen (siehe [Mar08, §3]). Wir wollen uns hier damit begnügen, an einem Beispiel zu zeigen, was die Aussagen bedeuten und wie man die Zerlegungen bzw. das Signum berechnen kann.

Beispiel 29.9 (Zyklenzerlegung)

Die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 5 & 4 & 3 & 1 \end{pmatrix} \in \mathbb{S}_5$$

hat die Zyklenzerlegung

$$\sigma = (1\ 2\ 5) \circ (3\ 4) = (3\ 4) \circ (1\ 2\ 5). \quad (82)$$

Eine berechtigte Frage ist, wie wir die Zyklenzerlegung in (82) gefunden haben. Wir wollen versuchen, dies so in Worte zu fassen, daß dem Leser daraus die allgemeine Vorgehensweise ersichtlich wird. Man starte mit der kleinsten Zahl, 1, und suche ihr Bild unter σ , also $\sigma(1) = 2$. Das liefert den Startteil des ersten Zyklus:

$$(1\ 2$$

Sodann betrachte man das Bild von 2 unter σ , also $\sigma(2) = 5$, und erhält:

$$(1\ 2\ 5$$

Man fährt mit dem Bild von 5 unter σ , also $\sigma(5) = 1$, fort. Da dieses das erste Element des ersten Zyklus war, schließen wir den Zyklus,

$$(1\ 2\ 5),$$

und beginnen den zweiten Zyklus mit der kleinsten Zahl in $\{1, \dots, 5\}$, die noch nicht in dem ersten Zyklus vorkommt, also mit 3:

$$(1\ 2\ 5) \circ (3$$

Dann betrachten wir deren Bild unter σ , also $\sigma(3) = 4$, und setzen so unseren zweiten Zyklus fort:

$$(1\ 2\ 5) \circ (3\ 4$$

Da bereits alle fünf Elemente von $\{1, \dots, 5\}$ aufgebraucht sind, muß notwendig $\sigma(4) = 3$ gelten, was es auch tut, und wir können damit auch den zweiten Zyklus schließen:

$$\sigma = (1\ 2\ 5) \circ (3\ 4).$$

Wie gesagt, da in $\{1, \dots, 5\}$ keine Zahl mehr übrig ist, sind wir fertig und haben die Zyklenzerlegung von σ gefunden. \square

Beispiel 29.10 (Zerlegung in Transpositionen)

Wir wollen nun zeigen, wie man die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 8 & 7 & 1 & 4 & 5 & 6 & 9 & 2 \end{pmatrix} \in \mathbb{S}_9$$

als Produkt von Transpositionen schreiben kann und wie man ihr Signum berechnet.

Dazu zerlegen wir sie zunächst in ein Produkt disjunkter Zyklen, und mit Hilfe des Verfahrens aus Beispiel 29.9 erhalten wir

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 3 & 8 & 7 & 1 & 4 & 5 & 6 & 9 & 2 \end{pmatrix} = (1\ 3\ 7\ 6\ 5\ 4) \circ (2\ 8\ 9).$$

Dann schreiben wir die Zyklen als Produkte von Transpositionen:

$$(1\ 3\ 7\ 6\ 5\ 4) = (1\ 3) \circ (3\ 7) \circ (7\ 6) \circ (6\ 5) \circ (5\ 4)$$

und

$$(2\ 8\ 9) = (2\ 8) \circ (8\ 9).$$

Die Ergebnisse können wir dann zusammensetzen und erhalten

$$\sigma = (1\ 3) \circ (3\ 7) \circ (7\ 6) \circ (6\ 5) \circ (5\ 4) \circ (2\ 8) \circ (8\ 9).$$

Aus dem Beispiel läßt sich leicht ein allgemeines Verfahren ableiten, um *eine* solche Zerlegung zu berechnen. Man sollte beachten, daß die Zerlegung in ein Produkt nicht eindeutig ist. Sie läßt sich auf viele Arten variieren. Wichtig ist sie allein, um das Signum zu berechnen, denn es gilt

$$\operatorname{sgn}(\sigma) = (-1)^7 = -1,$$

da σ Produkt von sieben Transpositionen ist.

Will man die Permutation gar als Produkt von Transpositionen benachbarter Zahlen schreiben, so reicht es, zu zeigen, wie man eine beliebige Transposition als Produkt solcher Transpositionen schreiben kann. Dann kann man das Verfahren auf jede Transposition in der obigen Zerlegung anwenden. Wir führen dies hier nur am Beispiel der Transposition $(3\ 7)$ vor. Das allgemeine Verfahren kann man daraus leicht ablesen:

$$(3\ 7) = (3\ 4) \circ (4\ 5) \circ (5\ 6) \circ (6\ 7) \circ (5\ 6) \circ (4\ 5) \circ (3\ 4).$$

Aufgaben

Aufgabe 29.11

Betrachte die Permutationen

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 6 & 3 & 7 & 5 & 1 & 4 \end{pmatrix}, \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 2 & 7 & 1 & 4 & 3 & 6 \end{pmatrix} \in \mathbb{S}_7.$$

- Berechne $\sigma \circ \pi$, $\pi \circ \sigma$, σ^{-1} , π^{-1} .
- Bestimme für jede der Permutationen in a. die Zyklenzerlegung.
- Schreibe $\sigma \circ \pi$ als ein Produkt von Transpositionen.
- Schreibe π^{-1} als ein Produkt von Transpositionen aufeinander folgender Zahlen.
- Berechne für jede der Permutationen in a. das Signum.

§ 30 Die Determinante

Wir werden jetzt eine ganz neue Möglichkeit kennenlernen, um quadratische lineare Gleichungssysteme zu lösen, nämlich mit Hilfe von Determinanten. Die Determinante ordnet einer quadratischen Matrix über einem Körper ein Element des Körpers zu, das genau dann ungleich Null ist, wenn die Matrix invertierbar ist. Die Determinante liefert aber nicht nur ein nützliches Kriterium für die Invertierbarkeit, sie ist vor allem aus theoretischen Gründen von unschätzbarem Wert. Z. B. liefert die Cramersche Regel mit Hilfe der Determinante eine geschlossene Formel für die Lösung eines linearen Gleichungssystems. Aus dieser Formel lassen sich Eigenschaften der Lösungen als Funktionen der Koeffizienten der Matrix bestimmen.

Die Determinante einer Matrix ist eine *polynomiale Funktion* in den Einträgen der Matrix. Sind diese Einträge etwa reelle oder komplexe Zahlen, so hängt die Determinante stetig von den Einträgen ab. Daraus folgt z. B. die wichtige Tatsache, daß eine invertierbare Matrix bei kleiner Störung der Einträge invertierbar bleibt. Damit wird eine Verbindung zur Analysis hergestellt (siehe Beispiel 40.21). Eine weitere wichtige Bedeutung in der Analysis hat die Determinante für die Volumenberechnung (siehe auch Bemerkung 37.32) und Beispiel 47.11).

Wir werden die Eigenschaften der Determinante soweit entwickeln, wie sie in der linearen Algebra wichtig sind. Allerdings führt uns die Determinante auch hier schon auf eine höhere Stufe: die Determinante ist nicht nur linear, sie ist *multilinear*, wie wir gleich sehen werden.

A) Die Leibnitz-Formel für die Determinante

Definition 30.1 (Determinante)

Wir definieren für $A \in \text{Mat}_n(\mathbb{K})$ die *Determinante* von A durch die *Leibniz-Formel*

$$\det(A) := |A| := \sum_{\sigma \in \mathbb{S}_n} \text{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)}. \quad (83)$$

Beispiel 30.2 (Determinanten für $n = 1, 2, 3$) a. Ist $n = 1$, dann ist $A = (a) \in \text{Mat}(1, \mathbb{K})$ und $\det(A) = a$.

b. Ist $n = 2$, dann ist $\mathbb{S}_2 = \{\text{id}, (1\ 2)\}$ und damit folgt:

$$\det(A) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

d. h. $\det(A)$ ist das Produkt der Elemente der Hauptdiagonalen minus dem Produkt der Elemente der Gegendiagonalen. Z.B.

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

- c. Für $n = 3$ hat S_n bereits sechs Elemente. Man berechnet in diesem Fall die Determinante mit der *Regel von Sarrus*:

$$\begin{array}{ccccc}
 a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\
 a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\
 a_{31} & a_{32} & a_{33} & a_{31} & a_{32}
 \end{array}$$

Die Produkte der Elemente längs der gestrichelten Linien tauchen bei der Berechnung der Determinante als positive Summanden auf, die Produkte der Elemente längs der gepunkteten Linien als negative Summanden. D. h., wir erhalten:

$$\det(A) = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Wenden wir das obige Schema auf die Matrix

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

an, so erhalten wir

$$\begin{array}{ccccc}
 2 & 0 & 1 & 2 & 0 \\
 0 & 5 & 0 & 0 & 5 \\
 -1 & 0 & 1 & -1 & 0
 \end{array}$$

und damit

$$\det(A) = 2 \cdot 5 \cdot 1 - (-1) \cdot 5 \cdot 1 = 15.$$

- d. Für $n = 4$ ergeben sich schon $4! = 24$ Summanden und für $n = 10$ gar $10! = 3628800$. In numerischen Anwendungen sind 1000×1000 -Matrizen keine Seltenheit, so daß es sich von selbst versteht, daß dabei nicht die Definition, bei der dann für die Determinante über 10^{2567} Produkte berechnet werden müßten, zur Berechnung verwendet werden kann. In der Tat wird zur Berechnung von Determinanten über Körpern wieder der Gauß-Algorithmus eine wichtige Rolle spielen.

Proposition 30.3 (Determinanten von Dreiecksmatrizen)

Ist $A = (a_{ij}) \in \text{Mat}_n(\mathbb{K})$ eine obere (bzw. untere) Dreiecksmatrix, d. h. $a_{ij} = 0$ für $i > j$ (bzw. $i < j$), dann ist

$$\det(A) = a_{11} \cdots a_{nn}$$

das Produkt der Diagonalelemente.

Beweis: Ist $\text{id} \neq \sigma \in S_n$, so gilt $i > \sigma(i)$ (bzw. $i < \sigma(i)$) für mindestens ein i . Wegen der Voraussetzung $a_{i\sigma(i)} = 0$ für $i > \sigma(i)$ (bzw. $i < \sigma(i)$) bleibt von den Summanden in (83) also nur der für id übrig. \square

Beispiel 30.4

$$\det \begin{pmatrix} 2 & 3 & 12 & -3 \\ 0 & -1 & 5 & -11 \\ 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 2 \cdot (-1) \cdot 3 \cdot 1 = -6.$$

Lemma 30.5 (Alternative Leibnitz-Formel)

Für die Determinante von $A \in \text{Mat}_n(\mathbb{K})$ gilt

$$\det(A) = \sum_{\sigma \in \mathbb{S}_n} \text{sgn}(\sigma) \cdot a_{\sigma(1)1} \cdot \dots \cdot a_{\sigma(n)n}. \quad (84)$$

Beweis: Man beachte, daß für $\sigma \in \mathbb{S}_n$ auch σ^{-1} eine Permutation der Zahlen $1, \dots, n$ ist, d. h. $\{1, \dots, n\} = \{\sigma^{-1}(1), \dots, \sigma^{-1}(n)\}$. Zudem wissen wir aus Satz 29.7 e., daß $\text{sgn}(\sigma) = \text{sgn}(\sigma^{-1})$, und es ist gleich, ob wir über $\sigma \in \mathbb{S}_n$ summieren oder über $\sigma^{-1} \in \mathbb{S}_n$, da auf beide Weisen alle Elemente von \mathbb{S}_n je einmal erreicht werden. Aus diesen Vorbetrachtungen ergibt sich:

$$\begin{aligned} \det(A) &= \sum_{\sigma \in \mathbb{S}_n} \text{sgn}(\sigma) \cdot a_{\sigma^{-1}(1)\sigma(\sigma^{-1}(1))} \cdot \dots \cdot a_{\sigma^{-1}(n)\sigma(\sigma^{-1}(n))} \\ &= \sum_{\sigma \in \mathbb{S}_n} \text{sgn}(\sigma) \cdot a_{\sigma^{-1}(1)1} \cdot \dots \cdot a_{\sigma^{-1}(n)n} \\ &\stackrel{29.7e.}{=} \sum_{\sigma \in \mathbb{S}_n} \text{sgn}(\sigma^{-1}) \cdot a_{\sigma^{-1}(1)1} \cdot \dots \cdot a_{\sigma^{-1}(n)n} \\ &= \sum_{\sigma^{-1} \in \mathbb{S}_n} \text{sgn}(\sigma^{-1}) \cdot a_{\sigma^{-1}(1)1} \cdot \dots \cdot a_{\sigma^{-1}(n)n} \\ &= \sum_{\pi \in \mathbb{S}_n} \text{sgn}(\pi) \cdot a_{\pi(1)1} \cdot \dots \cdot a_{\pi(n)n}. \end{aligned}$$

□

Proposition 30.6 (Die Determinante der Transponierten)

Für $A \in \text{Mat}_n(\mathbb{K})$ gilt:

$$\det(A) = \det(A^t).$$

Beweis: Sei $A = (a_{ij})$ und $A^t = (a'_{ij})$, dann gilt $a'_{ij} = a_{ji}$. Mithin erhalten wir mit Hilfe von Lemma 30.5

$$\begin{aligned} \det(A) &= \sum_{\sigma \in \mathbb{S}_n} \text{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot \dots \cdot a_{n\sigma(n)} \\ &= \sum_{\sigma \in \mathbb{S}_n} \text{sgn}(\sigma) \cdot a'_{\sigma(1)1} \cdot \dots \cdot a'_{\sigma(n)n} = \det(A^t). \end{aligned}$$

□

Beispiel 30.7

Beispiel 30.2 b. aufgreifend gilt

$$\det \begin{pmatrix} 5 & 6 \\ 2 & 3 \end{pmatrix} = \det \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix} = 3.$$

B) Die Determinante als Volumenform

Definition 30.8 (Multilineare Abbildungen)

Es seien V und W zwei K -Vektorräume.

a. Eine Abbildung

$$f : V^n = V \times \dots \times V \rightarrow W$$

heißt *multilinear*, falls f in jedem Argument linear ist, d. h. es gelten

$$f(x_1, \dots, x_i + y_i, \dots, x_n) = f(x_1, \dots, x_i, \dots, x_n) + f(x_1, \dots, y_i, \dots, x_n)$$

und

$$f(x_1, \dots, \lambda x_i, \dots, x_n) = \lambda \cdot f(x_1, \dots, x_i, \dots, x_n)$$

für jedes $i \in \{1, \dots, n\}$ und für alle $x_1, \dots, x_n, y_i \in V$ und $\lambda \in K$.

b. Eine multilineare Abbildung $f : V^n \rightarrow W$ heißt *alternierend*, falls für $(x_1, \dots, x_n) \in V^n$ mit $x_i = x_j$ für ein $i \neq j$, gilt:

$$f(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = 0.$$

Lemma 30.9

Ist $f : V^n \rightarrow W$ eine alternierende multilineare Abbildung, dann gilt für $\sigma \in \mathfrak{S}_n$

$$f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) = \operatorname{sgn}(\sigma) \cdot f(x_1, \dots, x_n).$$

Inbesondere gilt $f(x_1, \dots, x_i, \dots, x_j, \dots, x_n) = -f(x_1, \dots, x_j, \dots, x_i, \dots, x_n)$.

Beweis: Wir betrachten zunächst den Fall, daß $\sigma = (i j)$ eine Transposition ist. Da f alternierend und multilinear ist, folgt die Behauptung für σ aus

$$\begin{aligned} 0 &= f(x_1, \dots, x_i + x_j, \dots, x_i + x_j, \dots, x_n) \\ &= f(x_1, \dots, x_i, \dots, x_i, \dots, x_n) + f(x_1, \dots, x_i, \dots, x_j, \dots, x_n) \\ &\quad + f(x_1, \dots, x_j, \dots, x_i, \dots, x_n) + f(x_1, \dots, x_j, \dots, x_j, \dots, x_n) \\ &= f(x_1, \dots, x_i, \dots, x_j, \dots, x_n) + f(x_1, \dots, x_j, \dots, x_i, \dots, x_n). \end{aligned}$$

Ist $\sigma \in \mathfrak{S}_n$ beliebig, so können wir $\sigma = \tau_1 \circ \dots \circ \tau_k$ als Produkt von Transpositionen schreiben und die Behauptung folgt mittels Induktion nach der Anzahl k der Transpositionen. Den Induktionsanfang $k = 1$ haben wir bereits gezeigt. Ist $k \geq 2$ und setzen wir $\pi = \tau_2 \circ \dots \circ \tau_k$, so folgt mit der Vorüberlegung

$$\begin{aligned} f(x_{\sigma(1)}, \dots, x_{\sigma(n)}) &= f(x_{\tau_1(\pi(1))}, \dots, x_{\tau_1(\pi(n))}) = -f(x_{\pi(1)}, \dots, x_{\pi(n)}) \\ &\stackrel{\text{Ind.}}{=} -\operatorname{sgn}(\pi) \cdot f(x_1, \dots, x_n) \stackrel{\text{Satz 29.7}}{=} \operatorname{sgn}(\tau_1 \circ \pi) \cdot f(x_1, \dots, x_n) \\ &= \operatorname{sgn}(\sigma) \cdot f(x_1, \dots, x_n). \end{aligned}$$

□

Bemerkung 30.10

Wir können $\operatorname{Mat}_n(K)$ auf recht natürliche Weise mit $K^n \times \dots \times K^n$ identifizieren, indem wir eine Matrix $A = (a_{ij})$ mit dem n -Tupel ihrer Spaltenvektoren (a^1, \dots, a^n) gleichsetzen. Das wollen wir im folgenden tun.

Satz 30.11 (Die Determinante als Volumenform)a. *Die Determinante*

$$\det : \text{Mat}_n(\mathbb{K}) \rightarrow \mathbb{K} : \mathbf{A} \mapsto \det(\mathbf{A})$$

ist eine alternierende multilineare Abbildung mit $\det(\mathbf{1}_n) = 1$.

b. Ist $f : \text{Mat}_n(\mathbb{K}) \rightarrow \mathbb{K}$ eine alternierende multilineare Abbildung und $\mathbf{A} \in \text{Mat}_n(\mathbb{K})$, so gilt

$$f(\mathbf{A}) = f(\mathbf{1}_n) \cdot \det(\mathbf{A}).$$

Beweis:

a. Wir werden im Beweis die Formel (84) aus Lemma 30.5 zur Berechnung der Determinante verwenden, da sie auf die Bedürfnisse der Determinante als multilineare Abbildung bezüglich der Spalten zugeschnitten ist.

Es seien $\mathbf{a}^j = (\mathbf{a}_{1j}, \dots, \mathbf{a}_{nj})^t$, $j = 1, \dots, n$, und $\mathbf{b}^i = (\mathbf{b}_{1i}, \dots, \mathbf{b}_{ni})^t$. Wir setzen $\mathbf{A} := (\mathbf{a}^1 \dots \mathbf{a}^i \dots \mathbf{a}^n)$, $\mathbf{B} := (\mathbf{a}^1 \dots \mathbf{b}^i \dots \mathbf{a}^n)$ und $\mathbf{C} := (\mathbf{a}^1 \dots \lambda \mathbf{a}^i + \mu \mathbf{b}^i \dots \mathbf{a}^n)$. Dann gilt

$$\begin{aligned} \det(\mathbf{C}) &= \sum_{\sigma \in \mathbb{S}_n} \text{sgn}(\sigma) \cdot \mathbf{a}_{\sigma(1)1} \cdot \dots \cdot (\lambda \mathbf{a}_{\sigma(i)i} + \mu \mathbf{b}_{\sigma(i)i}) \cdot \dots \cdot \mathbf{a}_{\sigma(n)n} \\ &= \lambda \cdot \sum_{\sigma \in \mathbb{S}_n} \text{sgn}(\sigma) \cdot \mathbf{a}_{\sigma(1)1} \cdot \dots \cdot \mathbf{a}_{\sigma(i)i} \cdot \dots \cdot \mathbf{a}_{\sigma(n)n} \\ &\quad + \mu \cdot \sum_{\sigma \in \mathbb{S}_n} \text{sgn}(\sigma) \cdot \mathbf{a}_{\sigma(1)1} \cdot \dots \cdot \mathbf{b}_{\sigma(i)i} \cdot \dots \cdot \mathbf{a}_{\sigma(n)n} \\ &= \lambda \cdot \det(\mathbf{A}) + \mu \cdot \det(\mathbf{B}), \end{aligned}$$

so daß \det multilinear ist.

Sei nun $\mathbf{a}^i = \mathbf{a}^j$, für ein $i \neq j$. Ist $\tau = (i \ j)$, die Transposition, die i und j vertauscht, dann besitzt \mathbb{S}_n nach Satz 29.7 die Zerlegung $\mathbb{S}_n = \mathbb{A}_n \cup \mathbb{A}_n \tau$. Ferner gilt für $\sigma \in \mathbb{A}_n$

$$\text{sgn}(\sigma) = 1 \quad \text{und} \quad \text{sgn}(\sigma\tau) = -1.$$

Wir erhalten also

$$\begin{aligned} \det(\mathbf{A}) &= \sum_{\sigma \in \mathbb{A}_n} \mathbf{a}_{\sigma(1)1} \cdot \dots \cdot \mathbf{a}_{\sigma(i)i} \cdot \dots \cdot \mathbf{a}_{\sigma(j)j} \cdot \dots \cdot \mathbf{a}_{\sigma(n)n} \\ &\quad - \sum_{\sigma \in \mathbb{A}_n} \mathbf{a}_{\sigma\tau(1)1} \cdot \dots \cdot \mathbf{a}_{\sigma\tau(i)i} \cdot \dots \cdot \mathbf{a}_{\sigma\tau(j)j} \cdot \dots \cdot \mathbf{a}_{\sigma\tau(n)n} \\ &= \sum_{\sigma \in \mathbb{A}_n} \mathbf{a}_{\sigma(1)1} \cdot \dots \cdot \mathbf{a}_{\sigma(i)i} \cdot \dots \cdot \mathbf{a}_{\sigma(j)j} \cdot \dots \cdot \mathbf{a}_{\sigma(n)n} \\ &\quad - \sum_{\sigma \in \mathbb{A}_n} \mathbf{a}_{\sigma(1)1} \cdot \dots \cdot \mathbf{a}_{\sigma(j)j} \cdot \dots \cdot \mathbf{a}_{\sigma(i)i} \cdot \dots \cdot \mathbf{a}_{\sigma(n)n} = 0, \end{aligned}$$

und somit ist \det alternierend.

Außerdem folgt $\det(\mathbf{1}_n) = 1 \cdot \dots \cdot 1 = 1$ aus Proposition 30.3.

b. Mit den Notationen von a. gilt $\mathbf{a}^i = \sum_{j=1}^n \mathbf{a}_{ji} \mathbf{e}^j$, wenn \mathbf{e}^j der j -te Einheitsvektor ist. Aus der Multilinearität von f folgt:

$$f(\mathbf{A}) = \sum_{j_1=1}^n \mathbf{a}_{j_1 1} f(\mathbf{e}^{j_1} \mathbf{a}^2 \dots \mathbf{a}^n) = \sum_{j_1=1}^n \mathbf{a}_{j_1 1} \sum_{j_2=1}^n \mathbf{a}_{j_2 2} f(\mathbf{e}^{j_1} \mathbf{e}^{j_2} \mathbf{a}^3 \dots \mathbf{a}^n)$$

$$= \dots = \sum_{j_1, \dots, j_n=1}^n a_{j_1 1} \cdot \dots \cdot a_{j_n n} f(e^{j_1} \dots e^{j_n}).$$

Genau dann, wenn die j_1, \dots, j_n paarweise verschieden sind, existiert eine Permutation $\sigma \in S_n$ mit $(e^{j_1} \dots e^{j_n}) = (e^{\sigma(1)} \dots e^{\sigma(n)})$, und wegen Lemma 30.9 gilt dann

$$f(e^{\sigma(1)} \dots e^{\sigma(n)}) = \text{sgn}(\sigma) \cdot f(e^1 \dots e^n) = \text{sgn}(\sigma) \cdot f(\mathbf{1}_n).$$

Andernfalls stimmen zwei der j_i überein und $f(e^{j_1} \dots e^{j_n}) = 0$, da f alternierend ist. Insgesamt haben wir damit gezeigt:

$$f(A) = \sum_{\sigma \in S_n} a_{\sigma(1)1} \cdot \dots \cdot a_{\sigma(n)n} \cdot \text{sgn}(\sigma) \cdot f(\mathbf{1}_n) = \det(A) \cdot f(\mathbf{1}_n).$$

□

Bemerkung 30.12 (Das Volumen des Parallelotops)

Eine alternierende multilineare Abbildung $f : \text{Mat}_n(\mathbb{K}) \rightarrow \mathbb{K}$ wird auch eine *Volumenform* genannt. Aus Satz 30.11 b. folgt, daß die Determinante die einzige Volumenform f mit $f(\mathbf{1}_n) = 1$ ist, d.h. \det ist durch die Eigenschaften in Satz 30.11 a. eindeutig bestimmt.

Die Determinante hat eine wichtige geometrische Interpretation, die den Begriff *Volumenform* rechtfertigt. Seien $x_1, \dots, x_n \in \mathbb{R}^n$ und sei

$$P(x_1, \dots, x_n) := \{ \lambda_1 x_1 + \dots + \lambda_n x_n \in \mathbb{R}^n \mid 0 \leq \lambda_i \leq 1, i = 1, \dots, n \}$$

das von den Vektoren x_1, \dots, x_n aufgespannte *Parallelotop* (siehe Abbildung 3).

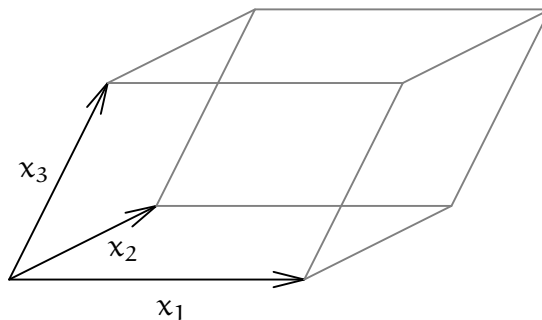


ABBILDUNG 3. Das Parallelotop $P(x_1, x_2, x_3)$ im \mathbb{R}^3

Dann definiert man das n -dimensionale Volumen von $P(x_1, \dots, x_n)$ mit Hilfe der Determinante als

$$\text{Volumen}(P(x_1, \dots, x_n)) = |\det(x_1 \dots x_n)|.$$

In Dimension $n = 1$ ist $|\det(x_1)| = |x_1|$ in der Tat die Länge der Strecke von 0 nach x_1 , und diese ist gerade $P(x_1)$. Wir werden in Bemerkung 37.32 zeigen, daß auch in Dimension $n = 2$ und $n = 3$ das so definierte Volumen mit dem euklidischen Flächeninhalt bzw. mit dem euklidischen Volumen übereinstimmt, daß die Definition

also sinnvoll ist. Sie wird im Rahmen der mehrdimensionalen Integrationstheorie und der Verallgemeinerung der Substitutionsregel eine wichtige Rolle spielen. \square

C) Der Gauß-Algorithmus zur Berechnung der Determinante

Korollar 30.13 (Spaltenoperationen und die Determinante)

Es sei $A \in \text{Mat}_n(\mathbb{K})$ und $\lambda \in \mathbb{K}$.

- Bei Vertauschung zweier Spalten von A ändert sich das Vorzeichen von $\det(A)$.
- Bei Multiplikation einer Spalte von A mit λ multipliziert sich $\det(A)$ mit λ .
- Bei Addition des λ -fachen einer Spalte zu einer anderen Spalte ändert sich $\det(A)$ nicht.
- Enthält A eine Nullspalte, so ist $\det(A) = 0$.
- Sind zwei Spalten von A gleich, so ist $\det(A) = 0$.

Beweis:

- Das ist ein Spezialfall von Lemma 30.9, da \det nach Satz 30.11 alternierend ist.
- Dies folgt aus der Multilinearität von \det , siehe Satz 30.11.
- Für $A = (\mathbf{a}^1 \dots \mathbf{a}^n)$ und $A' = (\mathbf{a}^1 \dots \mathbf{a}^j + \lambda \mathbf{a}^i \dots \mathbf{a}^n)$ folgt aus der Multilinearität und da \det alternierend ist:

$$\det(A') = \det(A) + \lambda \cdot \det(\mathbf{a}^1 \dots \mathbf{a}^i \dots \mathbf{a}^i \dots \mathbf{a}^n) = \det(A) + \lambda \cdot 0 = \det(A).$$
- Ist eine Spalte von A Null, so folgt $\det(A) = 0$ aus b. mit $\lambda = 0$.
- Das folgt, da \det alternierend ist.

\square

Da die Determinante einer Matrix gleich der Determinante der Transponierten ist, sind die Begriffe Spalte und Zeile austauschbar. Eine exaktere Formulierung bietet das folgende Korollar.

Korollar 30.14 (Zeilenoperationen und die Determinante)

Wir können $\det : \text{Mat}_n(\mathbb{K}) \rightarrow \mathbb{K}$ auch als multilineare Abbildung auf den Zeilen einer Matrix A auffassen. Entsprechend gilt Korollar 30.13 auch für Zeilen statt Spalten.

Da sich die Determinante bei der Addition eines Vielfachen einer Zeile zu einer anderen nicht ändert, können wir den Gauß-Algorithmus zur Berechnung von Determinanten einsetzen.

Algorithmus 30.15 (Algorithmus zur Berechnung der Determinante über \mathbb{K})

INPUT: $A \in \text{Mat}_n(\mathbb{K})$.

OUTPUT: $\det(A)$.

1. Schritt: Setze $d = 1$.

2. Schritt: Überführe A mittels Gauß-Algorithmus in nicht-reduzierte ZSF, d. h. führe im Gauß-Algorithmus 27.10 Schritt sieben nicht aus. Jedesmal, wenn dabei zwei Zeilen vertauscht werden, ersetze d durch $-d$. - Wird bei der Gaußreduktion ein Pivotelement zu Null, gib Null zurück und brich ab.

3. Schritt: Gib das Produkt von d mit den Diagonalelementen der ZSF zurück.

Beispiel 30.16

Wir wollen die Determinante der Matrix

$$A = \begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 0 \end{pmatrix} \in \text{Mat}_3(\mathbb{R})$$

berechnen. Dazu überführen wir sie mittels des Gauß-Algorithmus in ZSF und merken uns die Zeilenvertauschungen.

$$\begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 0 \end{pmatrix} \xrightarrow[\text{d}=-1]{\text{I} \leftrightarrow \text{II}} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{pmatrix} \xrightarrow[\text{III} \rightarrow \text{III} - 7\text{I}]{\text{II} \rightarrow \text{II} - 4\text{I}} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -21 \end{pmatrix}$$

$$\xrightarrow{\text{III} \rightarrow \text{III} - 2\text{II}} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & -9 \end{pmatrix}$$

Damit gilt dann

$$\det(A) = d \cdot 1 \cdot (-3) \cdot (-9) = -27.$$

Beispiel 30.17

Sei $A \in \text{Mat}(n+1, K)$ gegeben durch

$$A = \begin{pmatrix} 0 & 1 & 2 & 3 & \dots & n \\ 1 & 0 & 1 & 2 & \dots & n-1 \\ 2 & 1 & 0 & 1 & \dots & n-2 \\ 3 & 2 & 1 & 0 & \dots & n-3 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ n & n-1 & n-2 & n-3 & \dots & 0 \end{pmatrix}.$$

Ziehe für $i = 1, \dots, n$ von der i -ten Zeile die $(i+1)$ -te Zeile ab. Wir erhalten:

$$A = \begin{pmatrix} -1 & 1 & 1 & 1 & \dots & 1 & 1 \\ -1 & -1 & 1 & 1 & \dots & 1 & 1 \\ -1 & -1 & -1 & 1 & \dots & 1 & 1 \\ -1 & -1 & -1 & -1 & \dots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & -1 & \dots & -1 & 1 \\ n & n-1 & n-2 & n-3 & \dots & 1 & 0 \end{pmatrix}.$$

Addiere nun für $i = 2, \dots, n + 1$ die erste Spalte zur i -ten Spalte. Dann erhalten wir:

$$A = \begin{pmatrix} -1 & 0 & 0 & 0 & \dots & 0 & 0 \\ * & -2 & 0 & 0 & \dots & 0 & 0 \\ * & * & -2 & 0 & \dots & 0 & 0 \\ * & * & * & -2 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ & * & * & * & \dots & -2 & 0 \\ * & * & * & * & \dots & * & n \end{pmatrix}.$$

Es folgt:

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

Bemerkung 30.18

In Beispiel 30.17 haben wir durch ganz wenige Zeilen- und Spaltenoperationen die Matrix in Dreiecksgestalt überführt. Das lag aber an der speziellen Struktur der Matrix. Im allgemeinen Fall braucht der oben beschriebene Algorithmus zur Berechnung der Determinante mit Hilfe des Gauß-Algorithmus $\sim \frac{n^3}{3}$ Multiplikationen für eine $n \times n$ -Matrix. In der Definition der Determinante tauchen dagegen $n!$ Summanden von je n Produkten auf, mit $n! \sim \left(\frac{n}{e}\right)^n$, wobei e die Eulersche Zahl ist. Man sagt, daß der Gauß-Algorithmus *polynomial*, die Definition aber *exponentiell* in der Größe der Matrix ist. Grundsätzlich gelten polynomiale Algorithmen als effizient, exponentielle dagegen als unakzeptabel ineffizient. Allerdings gibt es Fälle, wo keine polynomialen Algorithmen bekannt sind.

D) Der Determinantenmultiplikationssatz und der Kästchensatz

Satz 30.19 (Determinantenmultiplikationssatz)

Für Matrizen $A, B \in \text{Mat}_n(\mathbb{K})$ gilt

$$\det(A \circ B) = \det(A) \cdot \det(B).$$

Beweis: Wähle $A \in \text{Mat}_n(\mathbb{K})$ fest und betrachte die Abbildung

$$f : \text{Mat}_n(\mathbb{K}) \rightarrow \mathbb{K} : B \mapsto \det(A \circ B).$$

f ist multilinear bezüglich der Spalten von B , da A auf jede Spalte von B linear wirkt. Außerdem ist f alternierend, da mit B auch $A \circ B$ zwei gleiche Spalten hat. Damit folgt aus Satz 30.11:

$$\det(A \circ B) = f(B) = f(\mathbf{1}_n) \cdot \det(B) = \det(A) \cdot \det(B).$$

□

Beispiel 30.20

In Beispiel 30.2 b. gilt

$$A = \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \in \text{Mat}_2(\mathbb{R}),$$

so folgt aus dem Determinantenmultiplikationssatz 30.19 und weil die beiden Matrizen auf der rechten Seite Dreiecksmatrizen sind:

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

Das folgende Korollar ist eine Verallgemeinerung der Aussage in Aufgabe 22.14.

Korollar 30.21 (Determinante und Invertierbarkeit)

Genau dann ist $A \in \text{Mat}_n(\mathbb{K})$ invertierbar, wenn $\det(A) \neq 0$. In diesem Fall gilt

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

Beweis: Ist A invertierbar, so gilt

$$1 = \det(\mathbb{1}_n) = \det(A \circ A^{-1}) = \det(A) \cdot \det(A^{-1}).$$

Dies zeigt, daß $\det(A)$ nicht Null sein kann, und zudem ist damit die obige Formel bewiesen.

Ist A nicht invertierbar, so sind die Spalten von A linear abhängig und durch mehrfache Addition von Vielfachen bestimmter Spalten zu einer anderen können wir eine Nullspalte erzeugen. Nach Korollar 30.13 c. ändert sich dabei der Wert der Determinante nicht, und nach Korollar 30.13 d. muß er somit 0 sein. \square

Beispiel 30.22

Die Matrix A in Beispiel 30.20 ist invertierbar und ihre Inverse hat Determinante $\frac{1}{3}$. Dies wissen wir, ohne die Inverse auszurechnen. Diese können wir mit Hilfe von Aufgabe 22.14 berechnen. Für eine invertierbare 2×2 -Matrix gilt

$$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \text{Mat}_n(\mathbb{K})$$

gilt

$$B^{-1} = \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \frac{1}{\det(B)} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

so daß wir im Beispiel

$$A^{-1} = \frac{1}{3} \cdot \begin{pmatrix} 3 & -2 \\ -6 & 5 \end{pmatrix}$$

erhalten.

Bemerkung 30.23 (\det ist ein Gruppenepimorphismus.)

In der Sprache der Vorlesung Algebraische Strukturen folgt aus Satz 30.19 und Korollar 30.21, daß

$$\det : (\text{Gl}_n(\mathbb{K}), \circ) \rightarrow (\mathbb{K}^*, \cdot)$$

ein Gruppenepimorphismus ist. Dazu beachte man, daß \det surjektiv ist wegen

$$\det \left(\begin{array}{c|ccc} \lambda & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & & \\ 0 & & \mathbb{1}_{n-1} & \end{array} \right) = \lambda.$$

Satz 30.24 (Kästchensatz)

Es sei $A \in \text{Mat}_n(\mathbb{K})$ eine Blockmatrix der Form

$$A = \left(\begin{array}{c|c} \mathbb{B} & \mathbb{C} \\ \hline 0 & \mathbb{D} \end{array} \right)$$

mit $\mathbb{B} \in \text{Mat}_k(\mathbb{K})$, $\mathbb{C} \in \text{Mat}(k \times l, \mathbb{K})$, $\mathbb{D} \in \text{Mat}_l(\mathbb{K})$, $0 \in \text{Mat}(l \times k, \mathbb{K})$ und $n = k + l$.

Dann gilt:

$$\det(A) = \det(\mathbb{B}) \cdot \det(\mathbb{D}).$$

Beweis: Man beachte, daß

$$A = \left(\begin{array}{c|c} \mathbb{B} & \mathbb{C} \\ \hline 0 & \mathbb{D} \end{array} \right) = \left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & \mathbb{D} \end{array} \right) \circ \left(\begin{array}{c|c} \mathbb{B} & \mathbb{C} \\ \hline 0 & \mathbb{1}_l \end{array} \right).$$

Wegen des Determinantenmultiplikationssatzes 30.19 reicht es mithin zu zeigen:

$$\det \left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & \mathbb{D} \end{array} \right) = \det(\mathbb{D}) \quad (85)$$

und

$$\det \left(\begin{array}{c|c} \mathbb{B} & \mathbb{C} \\ \hline 0 & \mathbb{1}_l \end{array} \right) = \det(\mathbb{B}). \quad (86)$$

Die Abbildung

$$f : \text{Mat}_l(\mathbb{K}) \rightarrow \mathbb{K} : D' \mapsto \det \left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D' \end{array} \right)$$

ist offensichtlich multilinear und alternierend, und wegen Satz 30.11 b. gilt mithin

$$\det \left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & \mathbb{D} \end{array} \right) = f(\mathbb{D}) = f(\mathbb{1}_l) \cdot \det(\mathbb{D}) = \det(\mathbb{1}_n) \cdot \det(\mathbb{D}) = \det(\mathbb{D}),$$

d. h. (85) ist erfüllt.

Analog ist die Abbildung

$$g : \text{Mat}_k(\mathbb{K}) \rightarrow \mathbb{K} : B' \mapsto \det \left(\begin{array}{c|c} B' & \mathbb{C} \\ \hline 0 & \mathbb{1}_l \end{array} \right)$$

alternierend und multilinear in den Spalten von B' , also eine Volumenform. Wieder folgt aus Satz 30.11 mit Hilfe von Proposition 30.3, daß

$$\det \left(\begin{array}{c|c} \mathbb{B} & \mathbb{C} \\ \hline 0 & \mathbb{1}_l \end{array} \right) = g(\mathbb{B}) = g(\mathbb{1}_k) \cdot \det(\mathbb{B}) = \det \left(\begin{array}{c|c} \mathbb{1}_k & \mathbb{C} \\ \hline 0 & \mathbb{1}_l \end{array} \right) \cdot \det(\mathbb{B}) = \det(\mathbb{B}),$$

womit auch (86) gezeigt ist. □

E) Laplacescher Entwicklungssatz und Cramersche Regel

Wir kommen jetzt zu einer alternativen Berechnung der Determinante. Im Gegensatz zum Gaußalgorithmus kommt sie ohne Division aus und funktioniert deshalb über jedem kommutativen Ring mit Eins (siehe Bemerkung 30.36). Zu ihrer Herleitung führen wir zunächst verschiedene Hilfsmatrizen ein.

Definition 30.25

Es sei $A = (a_{ij}) = (a^1 \dots a^n) \in \text{Mat}_n(K)$, $n \geq 2$, und $b = (b_1, \dots, b_n)^t \in K^n$.

Wir definieren die *Ersetzungsmatrix*

$$A_i(b) := (a^1 \dots a^{i-1} \ b \ a^{i+1} \dots a^n),$$

in der die i -te Spalte von A durch b ersetzt wurde.

Ist $b = e_j$ der j -te Einheitsvektor, so gilt:

$$A_i(e_j) = \begin{pmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{j1} & \dots & 1 & \dots & a_{jn} \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{pmatrix}.$$

Ersetzen wir in $A_i(e_j)$ zusätzlich noch die j -te Zeile durch den i -ten Einheitsvektor, dann erhält man die Matrix

$$S_{ji}(A) = \begin{pmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ 0 & \dots & 1 & \dots & 0 \\ \vdots & & \vdots & & \vdots \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{pmatrix}.$$

Streicht man in der Matrix A die j -te Zeile und die i -te Spalte, so erhält man die *Streichungsmatrix*

$$A_{ji} = \left(\begin{array}{ccc|ccc} a_{11} & \dots & a_{1\ i-1} & a_{1\ i+1} & \dots & a_{1n} \\ \vdots & & \vdots & & & \vdots \\ a_{j-1\ 1} & \dots & a_{j-1\ i-1} & a_{j-1\ i+1} & \dots & a_{j-1\ n} \\ \hline a_{j+1\ 1} & \dots & a_{j+1\ i-1} & a_{j+1\ i+1} & \dots & a_{j+1\ n} \\ \vdots & & \vdots & & & \vdots \\ a_{n\ 1} & \dots & a_{n\ i-1} & a_{n\ i+1} & \dots & a_{n\ n} \end{array} \right).$$

Lemma 30.26

Für $A \in \text{Mat}_n(K)$, $n \geq 2$, $1 \leq i, j \leq n$, gilt:

$$\det(A_i(e_j)) = \det(S_{ji}(A)) = (-1)^{i+j} \det(A_{ji}).$$

Beweis: $S_{ji}(A)$ entsteht aus $A_i(e_j)$ durch Subtraktion des a_{jk} -fachen der i -ten Spalte von der k -ten Spalte, $k \in \{1, \dots, n\} \setminus \{i\}$. Also gilt nach Korollar 30.13:

$$\det(A_i(e_j)) = \det(S_{ji}(A)).$$

Durch $i - 1$ Spaltenvertauschungen und $j - 1$ Zeilenvertauschungen entsteht aus $S_{ji}(A)$ die Matrix

$$\left(\begin{array}{c|ccc} 1 & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & A_{ji} & \\ 0 & & & \end{array} \right).$$

Also folgt aus dem Kästchensatz 30.24 unter Beachtung der Korollare 30.13 und 30.14

$$\det(S_{ji}(A)) = (-1)^{i+j} \det(A_{ji}).$$

□

In der folgenden Definition beachte man die Vertauschung der Indizes!

Definition 30.27

Für $A \in \text{Mat}_n(\mathbb{K})$, $n \geq 2$, $1 \leq i, j \leq n$ heißt

$$a_{ij}^\# := (-1)^{i+j} \det(A_{ji})$$

ein *Kofaktor* von A . Die Matrix der Kofaktoren

$$A^\# := (a_{ij}^\#) \in \text{Mat}_n(\mathbb{K})$$

heißt die *Adjunkte* oder *Komplementärmatrix* von A .

Satz 30.28 (Satz über die Adjunkte)

Für $A \in \text{Mat}_n(\mathbb{K})$, $n \geq 2$, gilt:

$$A^\# \circ A = A \circ A^\# = \det(A) \cdot \mathbf{1}_n.$$

Beweis: Sei $A^\# \circ A = (c_{ik})$. Dann gilt mit Lemma 30.26:

$$\begin{aligned} c_{ik} &= \sum_{j=1}^n a_{ij}^\# \cdot a_{jk} = \sum_{j=1}^n a_{jk} \cdot \det(a^1 \dots a^{i-1} e_j a^{i+1} \dots a^n) \\ &= \det \left(a^1 \dots a^{i-1} \sum_{j=1}^n a_{jk} e_j a^{i+1} \dots a^n \right) \\ &= \det(a^1 \dots a^{i-1} a^k a^{i+1} \dots a^n) = \delta_{ik} \cdot \det(A), \end{aligned}$$

wobei δ_{ik} das Kronecker-Symbol ist. Das dritte Gleichheitszeichen folgt aus der Multilinearität von \det , das letzte, da \det alternierend ist.

Der Beweis, daß $A \circ A^\# = \det(A) \cdot \mathbf{1}_n$ geht analog. □

Korollar 30.29

Es sei $A \in \text{Mat}_n(\mathbb{K})$ invertierbar, so ist

$$A^{-1} = \frac{1}{\det(A)} \cdot A^\#.$$

Wir wollen an dieser Stelle einmal die vielen Aussagen, die zur Invertierbarkeit einer quadratischen Matrix über einem Körper äquivalent sind, sammeln.

Korollar 30.30

Für eine Matrix $A \in \text{Mat}(n, \mathbb{K})$ sind gleichwertig:

- a. A ist invertierbar.
- b. $\text{rang}(A) = n$.
- c. $\det(A) \neq 0$.
- d. f_A ist bijektiv.
- e. f_A ist injektiv.
- f. f_A ist surjektiv.
- g. $\text{rZSF}(A) = \mathbb{1}_n$.
- h. A ist das Produkt endlich vieler Elementarmatrizen.
- i. Es gibt eine Matrix $B \in \text{Mat}(n, \mathbb{K})$ mit $A \circ B = \mathbb{1}_n$.

Beweis: Die unterschiedlichen Äquivalenzen sind in den Sätzen 24.37, 25.22, 25.23, 26.24, 27.14, 28.2 und 30.21 gezeigt worden. \square

Beispiel 30.31

Für eine 2×2 -Matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

gilt $\det(A) = ad - bc$ und

$$A^\# = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Ist also $ad - bc \neq 0$, so gilt:

$$A^{-1} = \frac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Damit ist Aufgabe 22.14 bewiesen.

Sei nun konkret $\mathbb{K} = \mathbb{Q}$ und

$$A = \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix}.$$

Dann ist $\det(\mathbf{A}) = 1$ und somit gilt

$$\mathbf{A}^{-1} = \begin{pmatrix} 4 & -1 \\ -3 & 1 \end{pmatrix}.$$

Der Satz über die Adjunkte führt zu einer rekursiven Berechnungsformel für die Determinante, die für theoretische Überlegungen sehr nützlich ist. Sie ist auch als rekursive Prozedur sehr einfach zu programmieren, aber nicht sehr effizient. Sie hat die gleiche Komplexität, wie die Leibnizsche Formel (83) zur Definition der Determinante.

Satz 30.32 (Laplacescher Entwicklungssatz)

Es sei $\mathbf{A} \in \text{Mat}_n(\mathbf{K})$.

a. Wir nennen die folgende Formel, die Entwicklung nach der i -ten Zeile:

$$\det(\mathbf{A}) = \sum_{j=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det(\mathbf{A}_{ij}). \quad (87)$$

b. Entsprechend nennen wir die folgende Formel, die Entwicklung nach der j -ten Spalte:

$$\det(\mathbf{A}) = \sum_{i=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det(\mathbf{A}_{ij}). \quad (88)$$

Beweis: Nach Satz 30.28 gilt für $\mathbf{A} \circ \mathbf{A}^\# = (c_{ik})$

$$\det(\mathbf{A}) = c_{ii} = \sum_{j=1}^n a_{ij} \cdot a_{ji}^\# = \sum_{j=1}^n (-1)^{i+j} \cdot a_{ij} \cdot \det(\mathbf{A}_{ij}).$$

Damit folgt (87), und (88) zeigt man analog durch die Betrachtung von $\mathbf{A}^\# \circ \mathbf{A}$. \square

Bemerkung 30.33

Entwickelt man $\mathbf{A} = (a_{ij})$ nach der ersten Zeile, so gilt:

$$\begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} = a_{11} \cdot \begin{vmatrix} a_{22} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n2} & \dots & a_{nn} \end{vmatrix} - a_{12} \cdot \begin{vmatrix} a_{21} & a_{23} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n3} & \dots & a_{nn} \end{vmatrix} \\ + \dots + (-1)^{n+1} a_{1n} \cdot \begin{vmatrix} a_{21} & \dots & a_{2\ n-1} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{n\ n-1} \end{vmatrix}.$$

Benutzt man dieses Verfahren, so entwickelt man am Besten nach Zeilen bzw. Spalten, die möglichst viele Nullen enthalten. Die Vorzeichen merkt man sich am

Günstigsten mit der sogenannten *Schachbrettregel*:

$$\begin{vmatrix} + & - & + & - & \dots \\ - & + & - & + & \dots \\ + & - & + & - & \dots \\ - & + & - & + & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{vmatrix}$$

Für kleine Matrizen, insbesondere wenn die Matrix dünn besetzt ist, ist dieses Verfahren zur Berechnung der Determinante (und zur Berechnung der Inversen) durchaus anwendbar. Für größere Matrizen ist auf jeden Fall der Gaußsche Eliminationsalgorithmus vorzuziehen.

Wir berechnen nun die Determinante der Matrix

$$A = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 3 & 4 \\ 2 & 5 & 3 \end{pmatrix} \in \text{Mat}_2(\mathbb{R})$$

mit Hilfe der Entwicklung nach der ersten Zeile. Dann gilt

$$\det(A) = 0 \cdot \det(A_{11}) - 2 \cdot \det(A_{12}) + 0 \cdot \det(A_{13}) = -2 \cdot \det \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} = (-2) \cdot (-5) = 10.$$

Algorithmus 30.34 (Laplace-Entwicklung)

INPUT: $A \in \text{Mat}_n(\mathbb{K})$.

OUTPUT: $\det(A)$.

- 1. Schritt:** Initialisiere \det auf Null.
- 2. Schritt:** Falls $n = 1$, setze $\det = a_{11}$ und gehe zu Schritt 3. Sonst tue für $i = 1, \dots, n$:
 - Bilde eine Hilfsmatrix B durch Streichen der ersten Spalte und der i -ten Zeile von A .
 - Rufe den Algorithmus mit B auf und merke Dir das Ergebnis in einer Hilfsvariablen x .
 - Addiere zu \det die Zahl $(-1)^{i+1} \cdot a_{i1} \cdot x$.
- 3. Schritt:** Gib \det zurück.

Der Satz über die Adjunkte liefert auch eine für theoretische Überlegungen sehr wichtige geschlossene Formel für die Lösungen eines linearen Gleichungssystems. Dies ist die berühmte *Cramersche Regel*. Wir werden sie in der mehrdimensionalen Analysis nutzen, um zu sehen, daß die Lösung eines eindeutig lösbaren linearen Gleichungssystems *stetig* von den Koeffizienten der erweiterten Koeffizientenmatrix (A, b) abhängt und somit kleine Störungen der Einträge nur zu kleinen Störungen in der Lösung führen. (siehe Beispiel 40.21).

Satz 30.35 (Cramersche Regel)

Es sei $A \in \text{Mat}_n(\mathbb{K})$ invertierbar und $\mathbf{b} \in \mathbb{K}^n$.

Für die eindeutig bestimmte Lösung $\mathbf{x} = (x_1, \dots, x_n)^t \in \mathbb{K}^n$ von $A\mathbf{x} = \mathbf{b}$ gilt dann

$$\begin{aligned} x_i &= \frac{1}{\det(A)} \cdot \det(A_i(\mathbf{b})) \\ &= \frac{1}{\det(A)} \cdot \det \begin{pmatrix} a_{11} & \dots & a_{1\ i-1} & b_1 & a_{1\ i+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n\ i-1} & b_n & a_{n\ i+1} & \dots & a_{nn} \end{pmatrix}. \end{aligned}$$

Beweis: Wegen Korollar 30.29 ist

$$\mathbf{x} = A^{-1}\mathbf{b} = \frac{1}{\det(A)} \cdot A^\# \mathbf{b}$$

die eindeutig bestimmte Lösung des linearen Gleichungssystems. Also folgt mit Lemma 30.26 und der Multilinearität der Determinante

$$\begin{aligned} x_i &= \frac{1}{\det(A)} \cdot \sum_{j=1}^n a_{ij}^\# \cdot b_j = \frac{1}{\det(A)} \cdot \sum_{j=1}^n \det(A_i(\mathbf{e}_j)) \cdot b_j \\ &= \frac{1}{\det(A)} \cdot \sum_{j=1}^n \det(a^1 \dots a^{i-1} \mathbf{e}_j a^{i+1} \dots a^n) \cdot b_j \\ &= \frac{1}{\det(A)} \cdot \det(a^1 \dots a^{i-1} \mathbf{b} a^{i+1} \dots a^n) \\ &= \frac{1}{\det(A)} \cdot \det(A_i(\mathbf{b})). \end{aligned}$$

□

Bemerkung 30.36 (Determinanten über kommutativen Ringen mit Eins)

Ist \mathbb{K} nur ein kommutativer Ring mit Eins, so können wir die Determinante einer Matrix in $\text{Mat}_n(\mathbb{K})$ ebenfalls durch die Leibniz-Formel definieren, und alle Aussagen dieses Abschnitts, die *ohne Division* auskommen, gelten mit dem gleichen Beweis.

Wir können den Gauß-Algorithmus 30.15 über beliebigen Ringen in der angegebenen Form *nicht* mehr anwenden, da dabei Divisionen nötig sind. Außerdem gilt Korollar 30.21 *nicht* mehr in der angegebenen Form. Dort muß man die Bedingung $\det(A) \neq 0$ durch $\det(A)$ invertierbar ersetzen. Ebenso gilt Korollar 30.30 in *nicht* in vollem Umfang.

Alle anderen Aussagen gelten jedoch ohne jede Änderung. Dies trifft insbesondere auf den Satz zur Adjunkten 30.28 zu, den wir später für Matrizen mit Koeffizienten in einem Polynomring anwenden wollen. Außerdem können wir den Laplaceschen Entwicklungssatz im Gegensatz zum Gaußschen Algorithmus über jedem kommutativen Ring mit Eins anwenden, um die Determinante auszurechnen. Es gibt aber auch hier geschicktere Verfahren, indem man den Gaußschen Algorithmus abwandelt zum sogenannten Bareiss Algorithmus (siehe [Coh96]).

Für die Aussage in Korollar 30.29 beachte man, daß aus dem Determinantenmultiplikationssatz 30.19 und dem Satz zur Adjunkten 30.28 unmittelbar folgt, daß

eine quadratische Matrix über einem kommutativen Ring genau dann invertierbar ist, wenn $\det(A)$ invertierbar ist. In diesem Fall darf man dann auch in dem Ring durch $\det(A)$ teilen. Das trifft auf Korollar 30.29 ebenso zu wie auf die Cramersche Regel 30.35.

Betrachten wir konkret den Ring \mathbb{Z} der ganzen Zahlen, dann sind nur 1 und -1 invertierbar. Mithin sind nur ganzzahlige Matrizen mit Determinante 1 oder -1 über \mathbb{Z} invertierbar, d.h. nur für solche enthält die Inverse wieder nur ganze Zahlen. Ein Beispiel dafür haben wir in Beispiel 30.31 gesehen. Betrachten wir stattdessen die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix},$$

so gilt $\det(A) = -2 \notin \{1, -1\}$ und die Einträge von

$$A^{-1} = -\frac{1}{2} \cdot \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}$$

sind nicht mehr alle ganzzahlig, obwohl A nur ganzzahlige Einträge hatte. A ist als Matrix in $\text{Mat}_2(\mathbb{Q})$ also invertierbar, als Matrix in $\text{Mat}_2(\mathbb{Z})$ aber nicht.

Aufgaben

Aufgabe 30.37 (Vandermonde-Determinante)

Zeige, für $a_0, \dots, a_n \in K$ und $n \geq 1$ gilt

$$\det \begin{pmatrix} 1 & a_0 & a_0^2 & a_0^3 & \dots & a_0^n \\ 1 & a_1 & a_1^2 & a_1^3 & \dots & a_1^n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & a_n^3 & \dots & a_n^n \end{pmatrix} = \prod_{1 \leq i < j \leq n} (a_j - a_i).$$

Aufgabe 30.38

Berechne die Determinanten der folgenden Matrizen:

a. $\begin{pmatrix} 1 & -4 \\ -3 & 8 \end{pmatrix} \in \text{Mat}_2(\mathbb{R}).$

b. $\begin{pmatrix} -1 & 0 & -1 \\ 0 & 5 & 0 \\ 3 & 0 & 3 \end{pmatrix} \in \text{Mat}_3(\mathbb{R}).$

c. $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \in \text{Mat}_3(\mathbb{R}).$

Aufgabe 30.39

Sei K ein Körper und $\lambda \in K$. Bestimme die Determinante der Matrix

$$\begin{pmatrix} 1 & \lambda & \lambda^2 & \dots & \lambda^{n-1} \\ \lambda^{n-1} & 1 & \lambda & \dots & \lambda^{n-2} \\ \lambda^{n-2} & \lambda^{n-1} & 1 & \dots & \lambda^{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda & \lambda^2 & \lambda^3 & \dots & 1 \end{pmatrix} \in \text{Mat}(n \times n, K).$$

Aufgabe 30.40

Sei V ein n -dimensionaler \mathbb{C} -Vektorraum und $f : V \rightarrow V$ \mathbb{C} -linear. Mittels Einschränkung der Skalarmultiplikation können wir V als \mathbb{R} -Vektorraum und f als \mathbb{R} -lineare Abbildung auffassen. Des Weiteren bezeichnen wir mit $\det_{\mathbb{C}}(f)$ die Determinante von f als \mathbb{C} -lineare Abbildung und $\det_{\mathbb{R}}(f)$ die Determinante von f als \mathbb{R} -lineare Abbildung. Zeige:

$$\det_{\mathbb{R}}(f) = |\det_{\mathbb{C}}(f)|^2.$$

Hinweis: Für eine \mathbb{C} -Basis (v_1, \dots, v_n) von V betrachte man die zugehörige \mathbb{R} -Basis $(v_1, \dots, v_n, iv_1, \dots, iv_n)$ sowie jeweils die zugehörige Matrixdarstellung von f . Wem der allgemeine Fall zu schwer ist, der beschränke sich auf die Abbildung $f : \mathbb{C} \rightarrow \mathbb{C} : z \mapsto (a + ib) \cdot z$ mit $a, b \in \mathbb{R}$ fest vorgegeben. Was ist eine Basis von \mathbb{C} als \mathbb{R} -Vektorraum?

Aufgabe 30.41

Berechne die Determinante folgender Matrix mit Hilfe des Gauß-Algorithmus':

$$A = \begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ -2 & 0 & 3 & 1 & 3 \\ 0 & 2 & 5 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 3 & 4 & 1 \end{pmatrix} \in \text{Mat}_5(\mathbb{R}).$$

Aufgabe 30.42

Berechne die folgende Determinante mit Hilfe des Determinantenentwicklungssatzes:

$$B = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 3 & 1 & 3 \\ 2 & 5 & 0 & -1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \in \text{Mat}_4(\mathbb{Q}).$$

Aufgabe 30.43

Bestimme für welche $s \in \mathbb{R}$ die Abbildung

$$f_s : \mathbb{R}^3 \rightarrow \mathbb{R}^3 : (x, y, z)^t \mapsto (x + z, x + 2y + z, sx + y - z)^t$$

invertierbar ist und berechne für diese die Inverse mit Hilfe der Adjunkten.

Aufgabe 30.44

Es sei $A \in \text{Mat}_n(\mathbb{R})$ mit ungeradem n und $A^t = -A$. Zeige, A ist nicht invertierbar. Bleibt die Aussage wahr, wenn wir \mathbb{R} durch einen anderen Körper ersetzen?

Aufgabe 30.45

Zeige, ist $A \in \text{Mat}_n(\mathbb{K})$ eine invertierbare obere Dreiecksmatrix, so ist auch A^{-1} eine obere Dreiecksmatrix.

Aufgabe 30.46

Löse das folgende lineare Gleichungssystem mit Hilfe der Cramerschen Regel:

$$x + 2z = 3, \quad 3x + y = 5 \quad \text{und} \quad -x + y = 1.$$

§ 31 Der Polynomring $K[t]$

Wir verallgemeinern hier zunächst den Begriff des Polynoms, den wir in Definition 13.11 über den reellen Zahlen eingeführt haben. Auf Beweise werden wir weitgehend verzichten, da diese Bestandteil der Vorlesung Algebraische Strukturen oder weiterführender Vorlesungen sind.

Definition 31.1 (Der Polynomring)

Wir nennen einen Ausdruck der Form

$$\sum_{k=0}^n a_k \cdot t^k = a_n \cdot t^n + a_{n-1} \cdot t^{n-1} + \dots + a_1 \cdot t^1 + a_0 \cdot t^0$$

mit $a_0, \dots, a_n \in K$ und $n \in \mathbb{N}$ ein *Polynom* in der Unbestimmten t mit Koeffizienten in K .

Für zwei Polynome $f = \sum_{k=0}^n a_k \cdot t^k$ und $g = \sum_{k=0}^m b_k \cdot t^k$ soll gelten

$$f = g \iff a_k = b_k \quad \forall k = 0, \dots, \max\{m, n\}, \quad (89)$$

wobei $a_k = 0$ für $k > n$ und $b_k = 0$ für $k > m$. Wir sagen, zwei Polynome sind gleich, wenn ihre Koeffizienten übereinstimmen und sprechen dabei vom *Koeffizientenvergleich*.

Die Menge

$$K[t] = \left\{ \sum_{k=0}^n a_k \cdot t^k \mid a_k \in K, n \in \mathbb{N} \right\}$$

aller Polynome in der Unbestimmten t mit Koeffizienten in K heißt der *Polynomring* in der Unbestimmten t über dem Körper K .

Für zwei Polynome $f = \sum_{k=0}^n a_k \cdot t^k \in K[t]$ und $g = \sum_{k=0}^m b_k \cdot t^k \in K[t]$ sowie ein Skalar $\lambda \in K$ definieren wir

$$\lambda \cdot f = \sum_{k=0}^n (\lambda \cdot a_k) \cdot t^k$$

und

$$f + g = \sum_{k=0}^{\max\{m, n\}} (a_k + b_k) \cdot t^k$$

mit $a_k = 0$ für $k > n$ und $b_k = 0$ für $k > m$ sowie

$$f \cdot g = \sum_{k=0}^{m+n} \left(\sum_{i=0}^k a_i \cdot b_{k-i} \right) \cdot t^k = \sum_{k=0}^{m+n} \left(\sum_{i+j=k} a_i \cdot b_j \right) \cdot t^k$$

wobei $a_i = 0$ für $i > n$ und $b_j = 0$ für $j > m$.

Wir erinnern uns, daß eine K -Algebra ein K -Vektorraum mit einer Multiplikation ist, bezüglich derer der Vektorraum ein Ring mit Eins ist und mit der die Skalarmultiplikation verträglich ist (siehe Bemerkung 26.2).

Proposition 31.2 (Der Polynomring als K -Algebra.)

Der Polynomring $K[t]$ ist eine kommutative K -Algebra mit Basis $B = (t^k \mid k \in \mathbb{N})$.

Beweis: In der Vorlesung Algebraische Strukturen wird gezeigt, daß $K[t]$ ein kommutativer Ring mit Eins ist (siehe [Mar08, Satz 6.15]). Beachtet man, daß die Skalarmultiplikation $\lambda \cdot f = (\lambda \cdot t^0) \cdot f$ erfüllt, so folgen damit automatisch auch die fehlenden Vektorraumgesetze (siehe auch Bemerkung 31.22) sowie die Verträglichkeit von Multiplikation und Skalarmultiplikation. $K[t]$ ist also eine K -Algebra.

Da jedes Polynom eine endliche Linearkombination der t^k ist, ist die Familie $(t^k \mid k \in \mathbb{N})$ ein Erzeugendensystem des K -Vektorraumes $K[t]$, und wegen (89) ist sie zudem linear unabhängig, da $\sum_{k=0}^n a_k \cdot t^k = 0$ genau dann gilt wenn $a_k = 0$ für alle $k = 0, \dots, n$. \square

Bemerkung 31.3

Die Abbildung $i : K \hookrightarrow K[t] : a \mapsto a \cdot t^0$ ist ein K -Algebrenmonomorphismus, und wir identifizieren die *konstanten Polynome* deshalb mit den Elementen aus K . Das paßt damit zusammen, daß das Polynom t^0 die Eins der K -Algebra $K[t]$ ist. Wir schreiben deshalb

$$a_n \cdot t^n + \dots + a_1 \cdot t^1 + a_0 \cdot t^0 = a_n \cdot t^n + \dots + a_1 \cdot t + a_0.$$

Definition 31.4 (Der Grad eines Polynoms)

Sei $f = \sum_{k=0}^n a_k t^k \in K[t]$ mit $a_n \neq 0$, dann heißt $\deg(f) := n$ der *Grad* von f und $\text{lc}(f) := a_n$ der *Leitkoeffizient* von f . Zudem setzen wir $\deg(0) := -\infty$ und $\text{lc}(0) := 0$. Ist $\text{lc}(f) = 1$ oder $f = 0$, so nennen wir f *normiert*.

Beachte, ein Polynom f ist genau dann konstant, wenn $\deg(f) \leq 0$.

Lemma 31.5 (Gradformeln)

Seien $f, g \in K[t] \setminus \{0\}$. Dann gelten:

- $\deg(f + g) \leq \max \{ \deg(f), \deg(g) \}$.
- $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Beweis: Der Beweis wird in der Vorlesung Algebraische Strukturen geführt, siehe [Mar08, Proposition 6.16]. Er ergibt sich durch Einsetzen der Definition. \square

Beispiel 31.6

Sei $f = 2t + 1, g = -2t + 1 \in \mathbb{Q}[t]$, dann gilt $f + g = 2$, also $\deg(f + g) < \max \{ \deg(f), \deg(g) \}$, aber $f \cdot g = -4t^2 + 1$ und somit $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Definition 31.7 (Nullstellen von Polynomen)

Es sei L eine K -Algebra, $b \in L$ und $f = \sum_{k=0}^n a_k \cdot t^k \in K[t]$. Wir setzen

$$f(b) := \sum_{k=0}^n a_k \cdot b^k \in L.$$

Gilt $f(b) = 0$, so heißt b eine *Nullstelle* von f in L .

Beispiel 31.8

Sei $f = t^2 - 4 = t^2 - 4 \cdot t^0 \in \mathbb{R}[t]$, $L = \text{Mat}_2(\mathbb{R})$ und $\mathbf{b} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$, dann gilt

$$f(\mathbf{b}) = \mathbf{b}^2 - 4 \cdot \mathbf{b}^0 = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}^2 - 4 \cdot \mathbb{1}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Also ist \mathbf{b} eine Nullstelle von f in $\text{Mat}_2(\mathbb{R})$.

Satz 31.9 (Division mit Rest)

Es seien $f, g \in K[t] \setminus \{0\}$.

- Es gibt eindeutige Polynome $q, r \in K[t]$ mit $f = q \cdot g + r$ und $\deg(r) < \deg(g)$.
- Ist $\lambda \in K$ eine Nullstelle von f in K , dann gibt es ein Polynom $q \in K[t]$ mit $f = q \cdot (t - \lambda)$, d.h. wir können $t - \lambda$ als Linearfaktor abspalten.
- Ist $\deg(f) = n$, so hat f höchstens n Nullstellen in K .

Beweis: Der Beweis wird in der Vorlesung Algebraische Strukturen geführt, siehe [Mar08, Satz 7.27, Proposition 7.39, Satz 7.42]. Teil a. beweist man dabei mit Induktion nach dem Grad von f und kürzt dabei den Leitterm von f mit Hilfe des Leitterms von g . Teil b. folgt unmittelbar aus Teil a. indem man f durch den Linearfaktor $t - \lambda$ teilt, und Teil c. folgt aus Teil b. mit Induktion nach n . \square

Beispiel 31.10

Sei $f = t^3 - 1 \in \mathbb{R}[t]$, dann gilt offenbar $f(1) = 1^3 - 1 = 0$. Polynomdivision liefert:

$$\begin{array}{r} (t^3 - 1) : (t - 1) = t^2 + t + 1. \\ \underline{t^3 - t^2} \\ t^2 - t \\ \underline{t^2 - t} \\ t - 1 \\ \underline{t - 1} \\ - \end{array}$$

Also gilt $f = (t^2 + t + 1) \cdot (t - 1)$.

Proposition 31.11 (Das Minimalpolynom)

Es sei L eine K -Algebra und $\mathbf{b} \in L$.

- Der Einsetzhomomorphismus

$$\phi_{\mathbf{b}} : K[t] \longrightarrow L : f \mapsto f(\mathbf{b})$$

ist ein K -Algebrenhomomorphismus, d.h. $(f + g)(\mathbf{b}) = f(\mathbf{b}) + g(\mathbf{b})$, $(f \cdot g)(\mathbf{b}) = f(\mathbf{b}) \cdot g(\mathbf{b})$, $(\lambda \cdot f)(\mathbf{b}) = \lambda \cdot f(\mathbf{b})$ und $1(\mathbf{b}) = 1$.

- Es gibt ein eindeutig bestimmtes normiertes Polynom $\mu_{\mathbf{b}} \in K[t]$, so daß

$$\mu_{\mathbf{b}} K[t] := \{\mu_{\mathbf{b}} \cdot g \mid g \in K[t]\} \stackrel{!}{=} \{h \in K[t] \mid h(\mathbf{b}) = 0\} =: \text{Ker}(\phi_{\mathbf{b}}).$$

$\mu_{\mathbf{b}}$ heißt das Minimalpolynom von \mathbf{b} .

- c. Gibt es ein $0 \neq h \in K[t]$ mit $h(b) = 0$, so ist μ_b das normierte Nicht-Null-Polynom kleinsten Grades, das b als Nullstelle hat.

Beweis: Daß der Einsetzhomomorphismus ein K -Algebrenhomomorphismus ist, folgt unmittelbar aus den Definitionen (siehe auch [Mar08, Lemma 7.36]). Mithin ist der Kern

$$\text{Ker}(\phi_b) = \{h \in K[t] \mid h(b) = 0\}$$

von ϕ_b ein Ideal im Ring $K[t]$, siehe [Mar08, Satz 6.43]. Aus der Vorlesung Algebraische Strukturen wissen wir, daß $K[t]$ ein Hauptidealring ist, siehe [Mar08, Satz 7.51]. Es gibt also ein normiertes Polynom $\mu_b \in K[t]$ mit

$$\text{Ker}(\phi_b) = \mu_b K[t] = \{\mu_b \cdot g \mid g \in K[t]\}.$$

Die Eindeutigkeit von μ_b folgt dann leicht aus den Kürzungsregeln in $K[t]$, siehe [Mar08, Beispiel 7.2]. Zudem folgt aus der Gradformel unmittelbar, daß jedes Nicht-Null-Polynom in $\mu_b K[t]$ mindestens den Grad von μ_b hat. \square

Bemerkung 31.12 (Polynome versus Polynomfunktionen)

- a. Auch die Menge K^K aller Abbildungen von K nach K ist eine K -Algebra und die Abbildung

$$\psi : K[t] \longrightarrow K^K : f \mapsto f$$

die einem Polynom die zugehörige *Polynomfunktion* zuordnet, ist ein K -Algebrenhomomorphismus.

- b. Zwei verschiedene Polynome können dieselbe Polynomfunktion liefern! Die beiden Polynome $f = t^2 - t \in \mathbb{F}_2[t]$ und $g = 0 \in \mathbb{F}_2[t]$ induzieren beide als Polynomfunktion die Funktion konstant Null, da

$$f(0) = 0^2 - 0 = 0 \quad \text{und} \quad f(1) = 1^2 - 1 = 0.$$

Die Polynome f und g sind aber verschieden. In diesem Fall ist die Abbildung ψ aus Teil a. nicht injektiv.

- c. Enthält K *unendlich* viele Elemente, so ist die Abbildung ψ aus Teil a. injektiv, d.h. zwei Polynome sind genau dann verschieden, wenn die zugehörigen Polynomfunktionen verschieden sind. Liefern nämlich zwei Polynome f und g die gleichen Polynomfunktionen, so hat die Differenz $f - g$ unendlich viele Nullstellen und muß wegen Satz 31.9 somit das Nullpolynom sein.

Definition 31.13 (Vielfachheiten von Nullstellen)

- a. Ist $\lambda \in K$ und $f = (t - \lambda)^m \cdot g \in K[t]$ mit $m \geq 1$ und $g(\lambda) \neq 0$, so nennen wir λ eine Nullstelle von f mit *Vielfachheit* $\text{mult}(f, \lambda) = m$.
- b. Es sei $K \subseteq L$ ein Teilkörper des Körpers L und $f \in K[t]$ mit $n = \deg(f) > 0$. Gibt es $b_1, \dots, b_n \in L$ mit $f = (t - b_1) \cdot \dots \cdot (t - b_n)$ so sagen wir, daß f über L in *Linearfaktoren* zerfällt.

- c. Wir nennen einen Körper K *algebraisch abgeschlossen*, wenn jedes nicht-konstante Polynom in $K[t]$ über K in Linearfaktoren zerfällt.

Beispiel 31.14

Betrachte das Polynom $f = t^4 - 2t^3 + 2t^2 - 2t + 1 = (t - 1)^2 \cdot (t^2 + 1) \in \mathbb{R}[t]$. Dann ist $\lambda = 1$ keine Nullstelle von $t^2 + 1$. Mithin ist $\lambda = 1$ eine Nullstelle von f mit Vielfachheit $\text{mult}(f, 1) = 2$. Man beachte, daß f über \mathbb{R} nicht in Linearfaktoren zerfällt, da $t^2 + 1$ keine Nullstelle besitzt. Über \mathbb{C} zerfällt f hingegen in Linearfaktoren

$$f = (t - 1)^2 \cdot (t - i) \cdot (t + i).$$

Satz 31.15 (Fundamentalsatz der Algebra)

- a. Der Körper \mathbb{C} der komplexen Zahlen ist algebraisch abgeschlossen.
 b. Jeder Körper K ist Teilkörper eines algebraisch abgeschlossenen Körpers. Der kleinste solche Oberkörper von K ist bis auf Isomorphie eindeutig bestimmt und wird der algebraische Abschluß \bar{K} von K genannt.

Beweis: Die als Fundamentalsatz der Algebra bekannte Aussage in Teil a. wird in der Vorlesung Einführung in die Funktionentheorie mit Mitteln der Analysis bewiesen und unter Umständen auch in der Vorlesung Einführung in die Algebra mit algebraischen Mitteln. Die Aussage in Teil b. ist Bestandteil der Vorlesung Einführung in die Algebra. \square

Ist der Körper K nicht algebraisch abgeschlossen, so zerfällt nicht jedes Polynom in Linearfaktoren. Zumindest aber läßt sich jedes Polynom als Produkt von nicht mehr weiter zerlegbaren Polynomen schreiben.

Definition 31.16 (Irreduzible Polynome)

Ein nicht-konstantes Polynom $f \in K[t] \setminus K$ heißt *irreduzibel*, wenn aus $f = g \cdot h$ mit $g, h \in K[t]$ stets $\deg(g) = 0$ oder $\deg(h) = 0$ folgt.

Beispiel 31.17

Aus $f = g \cdot h$ folgt mit der Gradformel

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

Ist $\deg(f) = 1$, so folgt unmittelbar, daß f irreduzibel ist. Ist $\deg(f) \in \{2, 3\}$, so ist f genau dann irreduzibel, wenn man von f keinen Faktor vom Grad 1 abspalten kann, d.h. wenn f keine Nullstelle in K hat.

Satz 31.18 (Primfaktorzerlegung im Polynomring)

Jedes nicht-konstante normierte Polynom in $K[t]$ läßt sich als Produkt von endlich vielen normierten irreduziblen Polynomen schreiben, und diese Faktoren sind bis auf die Reihenfolge eindeutig bestimmt.

Beweis: Die Aussage ist eines der Highlights der Vorlesung Algebraische Strukturen, siehe [Mar08, Satz 7.65]. \square

Beispiel 31.19

Das Polynom $f = t^4 - 2t^3 + 2t^2 - 2t + 1$ aus Beispiel 31.14 hat in $\mathbb{R}[t]$ die Primfaktorzerlegung

$$f = (t - 1)^2 \cdot (t^2 + 1)$$

und in $\mathbb{C}[t]$ die Primfaktorzerlegung

$$f = (t - 1)^2 \cdot (t - i) \cdot (t + i).$$

Satz 31.20 (Bézout-Identität)

Seien $f, g \in K[t]$ zwei normierte teilerfremde Polynome, d.h. sie haben keinen Primfaktor gemeinsam, so gibt es Polynome $p, q \in K[t]$ mit

$$1 = p \cdot f + q \cdot g.$$

Allgemeiner gilt, sind $q_1, \dots, q_r \in K[t]$ normierte Polynome und gibt es keinen Primfaktor, den alle gemeinsam haben, dann gibt es Polynome $p_1, \dots, p_r \in K[t]$ mit

$$1 = p_1 \cdot q_1 + \dots + p_r \cdot q_r.$$

Beweis: Die Aussage wird in der Vorlesung Algebraische Strukturen bewiesen, siehe [Mar08, Korollar 7.51, Satz 7.54].

Wir wollen die Aussage hier im Spezialfall $g = t - \lambda$ beweisen. Teilen wir f durch $t - \lambda$ mit Rest, so finden wir Polynome $q, r \in K[t]$ mit

$$f = q \cdot (t - \lambda) + r$$

und $\deg(r) < \deg(t - \lambda) = 1$. Damit ist r eine Konstante. Wäre $r = 0$, so wäre $t - \lambda$ ein gemeinsamer Primfaktor von f und g , also ist $r \neq 0$. Dann ist

$$1 = \frac{1}{r} \cdot f - \frac{q}{r} \cdot g$$

die gesuchte Darstellung. □

Bemerkung 31.21 (Rationale Funktionen)

Der Polynomring $K[t]$ über einem Körper hat sehr viele Eigenschaften mit dem Ring \mathbb{Z} der ganzen Zahlen gemeinsam – in beiden Ringen gibt es eine Division mit Rest (d.h. sie sind Euklidische Ringe im Sinne der Vorlesung Algebraische Strukturen, siehe [Mar08, Beispiel 7.26, Korollar 7.28]) und in beiden Ringen hat man eine eindeutige Primfaktorzerlegung (siehe [Mar08, Korollar 7.63, Korollar 7.65]). Deshalb kann man bei den Polynomen das Problem der fehlenden multiplikativen Inversen genauso lösen wie im Fall der ganzen Zahlen, man führt Brüche ein. Dies führt zum Körper

$$K(t) = \left\{ \frac{f}{g} \mid f, g \in K[t], g \neq 0 \right\}$$

der *rationalen Funktionen*. Das Kürzen von Brüchen funktioniert wie in den rationalen Zahlen, und gleiches gilt für die Addition und die Multiplikation.

Bemerkung 31.22 (Abbrechende Folgen)

Wir bezeichnen mit $e_k = (\delta_{ik} \mid i \in \mathbb{N}) \in K^{\mathbb{N}}$ die Folge in K , die an der Stelle k den Eintrag 1 und sonst stets den Eintrag 0 hat, und wir betrachten den K -Vektorraum

$$V = \text{Lin}(e_k \mid k \in \mathbb{N}) = \text{Lin}(e_0, e_1, e_2, \dots)$$

der sogenannten *abbrechenden Folgen* — beachte, daß eine Folge in V nur endlich viele Glieder ungleich Null haben kann!

$B = (e_k \mid k \in \mathbb{N})$ ist eine Basis von V , und es gibt mithin genau eine K -lineare Abbildung

$$\varphi : V \longrightarrow K[t]$$

mit

$$\varphi(e_k) = t^k$$

für alle $k \in \mathbb{N}$. Da φ eine Basis auf eine Basis abbildet, ist φ ein Isomorphismus. Wir können die Vektorräume V und $K[t]$ mittels φ miteinander identifizieren.

Zudem existiert wegen Aufgabe 31.23 auf V genau eine Multiplikation, die bilinear ist und für die $e_i \cdot e_j = e_{i+j}$ gilt. Mit dieser Multiplikation wird V eine K -Algebra und φ wird zum K -Algebrenisomorphismus.

In vielen Büchern zur Linearen Algebra ist es üblich, den Polynomring als den Vektorraum der abbrechenden Folgen einzuführen und auf diesem dann wie oben eine Multiplikation einzuführen. Dies hat den Vorteil, daß Polynome klar definierte Objekte, nämlich abbrechende Folgen, sind und man Polynome nicht als *Ausdrücke* einer bestimmten Form einführen muß, ohne genau zu sagen, was das eigentlich heißen soll. Ich habe auf diesen Zugang verzichtet, da es Studienanfänger erfahrungsgemäß eher verwirrt, wenn Polynome Folgen sein sollen, während sie meist ohne Probleme hinnehmen, nicht wirklich gesagt zu bekommen, was ein Ausdruck der Form $\sum_{k=0}^n a_k \cdot t^k$ eigentlich sein soll. Die Identifikation von V mit $K[t]$ rechtfertigt unser Vorgehen nun im Nachhinein und uns reicht im weiteren Verlauf, daß wir wissen, wie wir mit Polynomen zu rechnen haben.

Aufgaben**Aufgabe 31.23** (Existenz- und Eindeutigkeitssatz für bilineare Abbildungen)

Sei V ein K -Vektorraum mit Basis $B = (x_i \mid i \in I)$ und sei $F = (y_{ij} \mid (i, j) \in I \times I)$ eine Familie von Vektoren im K -Vektorraum W . Dann gibt es genau eine multilineare Abbildung $f : V \times V \rightarrow W$ mit $f(x_i, x_j) = y_{ij}$ für alle $(i, j) \in I \times I$.

Sind $x = \sum_{\substack{i \in I \\ \text{endlich}}} a_i x_i, y = \sum_{\substack{i \in I \\ \text{endlich}}} b_i x_i \in V$, so gilt

$$f(x, y) = \sum_{\substack{i \in I \\ \text{endlich}}} \sum_{\substack{j \in I \\ \text{endlich}}} a_i b_j y_{ij}. \quad (90)$$

Aufgabe 31.24 (Polynominterpolation)

Es seien $b_0, \dots, b_n \in K$ paarweise verschieden und $c_0, \dots, c_n \in K$ beliebig. Zeige, es gibt genau ein Polynom $f \in K[t]$ vom Grad $\deg(f) \leq n$ mit $f(b_i) = c_i$ für alle $i = 0, \dots, n$.

Aufgabe 31.25

Zerlege das Polynom $f = t^4 + t^3 + 2t - 4 \in \mathbb{C}[t]$ in Linearfaktoren.

Aufgabe 31.26

Bestimme das Minimalpolynom $\mu_b \in \mathbb{Q}[t]$ von $b = \sqrt{2} \in \mathbb{R}$.

Aufgabe 31.27

Zeige, ist $f \in \mathbb{R}[t]$ irreduzibel, so ist $\deg(f) \in \{1, 2\}$.

Hinweis: Betrachte für eine komplexe Nullstelle λ von f die Fälle $\lambda \in \mathbb{R}$ und $\lambda \in \mathbb{C} \setminus \mathbb{R}$. In letzterem Fall zeige, daß auch das konjugiert Komplexe von λ eine Nullstelle von f ist und betrachte dann das Polynom $g = (t - \lambda) \cdot (t - \bar{\lambda})$.

§ 32 Endomorphismen und ihre Eigenwerte

In diesem Abschnitt sei V ein K -Vektorraum mit $1 \leq \dim_K(V) = n < \infty$.

A) Invarianten von Endomorphismen unter Konjugation

Bemerkung 32.1 (Endomorphismen)

Wir erinnern uns, daß K -lineare Abbildungen

$$f : V \longrightarrow V$$

auch *Endomorphismen* des K -Vektorraums V genannt werden (siehe Definition 23.19) und daß

$$\text{End}_K(V) = \{f : V \longrightarrow V \mid f \text{ ist } K\text{-linear}\}$$

die K -*Algebra* der Endomorphismen von V ist (siehe Bemerkung 26.2).

Zudem wissen wir, wie sich die Matrixdarstellungen von Endomorphismen unter Basiswechsel verhalten. Sind B und D zwei Basen des Vektorraums V und ist $T = T_B^D$, so gilt (siehe Korollar 26.9)

$$M_D^D(f) = T^{-1} \circ M_B^B(f) \circ T.$$

Dabei ist es von großer Wichtigkeit, daß wir jeweils im Definitions- und Zielbereich von f *dieselbe* Basis verwenden, und das wollen wir von nun an stets tun, wenn wir Matrixdarstellungen von Endomorphismen betrachten!

Wir können deshalb Eigenschaften von Matrizen, die unter Transformationen der Form

$$A \mapsto T^{-1} \circ A \circ T$$

erhalten bleiben, auch für Endomorphismen definieren, indem wir dazu ihre Matrixdarstellungen bezüglich einer beliebigen Basis verwenden. In diesem Abschnitt wollen wir einige Beispiele hierfür kennen lernen.

Definition 32.2 (Konjugiert oder ähnlich)

Zwei quadratische Matrizen $A, B \in \text{Mat}_n(K)$ heißen *konjugiert* oder *ähnlich*, wenn es eine invertierbare Matrix $T \in \text{Gl}_n(K)$ gibt, so daß $B = T^{-1} \circ A \circ T$ ist.

Bemerkung 32.3 (Konjugation ist eine Äquivalenzrelation)

Konjugation von Matrizen ist ein Beispiel für eine Äquivalenzrelation auf der Menge $\text{Mat}_n(K)$ der quadratischen $n \times n$ -Matrizen über K . D.h.

- jede Matrix ist zu sich selbst konjugiert, denn $A = \mathbb{1}_n^{-1} \cdot A \cdot \mathbb{1}_n$;
- ist A zu B konjugiert, so ist auch B zu A konjugiert, da aus $B = T^{-1} \circ A \circ T$ auch $A = (T^{-1})^{-1} \circ B \circ T^{-1}$ folgt;
- ist A zu B und B zu C konjugiert, so ist auch A zu C konjugiert, da aus $B = T^{-1} \circ A \circ T$ und $C = S^{-1} \circ B \circ S$ auch $C = (T \circ S)^{-1} \circ A \circ (T \circ S)$ folgt.

Satz 32.4 (Konjugierte Matrizen haben die gleiche Determinante.)

Sind $A, B \in \text{Mat}_n(\mathbb{K})$ konjugiert, dann gilt

$$\det(A) = \det(B).$$

Beweis: Aus Satz 30.19 und Korollar 30.21 folgen für $B = T^{-1} \circ A \circ T$:

$$\det(B) = \det(T^{-1} \circ A \circ T) = \det(T)^{-1} \cdot \det(A) \cdot \det(T) = \det(A).$$

□

Damit können wir die Determinante eines Endomorphismus definieren.

Definition 32.5 (Die Determinante eines Endomorphismus)

Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum der Dimension mindestens Eins mit Basis B und $f \in \text{End}_{\mathbb{K}}(V)$. Wir definieren die *Determinante* von f durch

$$\det(f) := \det(M_B^B(f)).$$

Da die Matrixdarstellungen eines Endomorphismus f zu verschiedenen Basen nach Korollar 26.9 konjugiert sind, ist diese Definition unter Berücksichtigung von Satz 32.4 unabhängig von der Wahl der Basis B .

Beispiel 32.6

Wir betrachten den Endomorphismus

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}^2: (x, y)^t \mapsto (5x + 2y, 6x + 3y)^t.$$

Ist E die Standardbasis des \mathbb{R}^2 , so gilt

$$M_E^E(f) = \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix},$$

und mithin gilt

$$\det(f) = \det(M_E^E(f)) = 5 \cdot 3 - 2 \cdot 6 = 3.$$

Alternativ könnte man die Basis $B = ((1, 1)^t, (0, 1)^t)$ betrachten und erhält dann

$$M_B^B(f) = \begin{pmatrix} 7 & 2 \\ 2 & 1 \end{pmatrix},$$

und in der Tat gilt auch

$$\det(M_B^B(f)) = 7 \cdot 1 - 2 \cdot 2 = 3.$$

Definition 32.7 (Die Spur einer Matrix)

Wir definieren die *Spur* einer quadratischen Matrix $A = (a_{ij}) \in \text{Mat}_n(\mathbb{K})$ als Summe der Diagonalelemente, d.h.

$$\text{Spur}(A) := \sum_{i=1}^n a_{ii} = a_{11} + a_{22} + a_{33} + \dots + a_{nn}.$$

Proposition 32.8 (Konjugierte Matrizen haben die gleiche Spur.)

Es seien $A, B \in \text{Mat}_n(\mathbb{K})$ zwei $n \times n$ -Matrizen. Dann gilt

$$\text{Spur}(A \circ B) = \text{Spur}(B \circ A).$$

Insbesondere, zwei konjugierte Matrizen haben die gleiche Spur.

Beweis: Es seien $A = (a_{ij})$, $B = (b_{ij})$, $A \circ B = (c_{ij})$ und $B \circ A = (d_{ij})$. Dann gilt

$$c_{ii} = \sum_{k=1}^n a_{ik} \cdot b_{ki}$$

und

$$d_{kk} = \sum_{i=1}^n b_{ki} \cdot a_{ik}.$$

Also gilt

$$\text{Spur}(A \circ B) = \sum_{i=1}^n c_{ii} = \sum_{i=1}^n \sum_{k=1}^n a_{ik} b_{ki} = \sum_{k=1}^n \sum_{i=1}^n b_{ki} a_{ik} = \sum_{k=1}^n d_{kk} = \text{Spur}(B \circ A).$$

Ist nun $T \in \text{Gl}_n(\mathbb{K})$, so folgt

$$\text{Spur}(T^{-1} \circ A \circ T) = \text{Spur}(A \circ T \circ T^{-1}) = \text{Spur}(A),$$

so daß konjugierte Matrizen die gleiche Spur haben. \square

Damit können wir die Spur eines Endomorphismus definieren.

Definition 32.9 (Die Spur eines Endomorphismus)

Sei V ein endlich-dimensionaler \mathbb{K} -Vektorraum der Dimension mindestens Eins mit Basis B und $f \in \text{End}_{\mathbb{K}}(V)$. Wir definieren die *Spur* von f durch

$$\text{Spur}(f) := \text{Spur}(M_B^B(f)).$$

Da die Matrixdarstellungen eines Endomorphismus f zu verschiedenen Basen nach Korollar 26.9 konjugiert sind, ist diese Definition unter Berücksichtigung von Proposition 32.8 unabhängig von der Wahl der Basis B .

Beispiel 32.10

Mit den Bezeichnungen aus Beispiel 32.6 gilt

$$\text{Spur}(f) = \text{Spur}(M_E^E(f)) = \text{Spur} \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix} = 5 + 3 = 8$$

oder alternativ

$$\text{Spur}(f) = \text{Spur}(M_B^B(f)) = \text{Spur} \begin{pmatrix} 7 & 2 \\ 2 & 1 \end{pmatrix} = 7 + 1 = 8.$$

Definition 32.11 (Das Charakteristische Polynom einer Matrix)

Für eine quadratische Matrix $A \in \text{Mat}_n(\mathbb{K})$ nennen wir

$$\chi_A := \det(t \cdot \mathbb{1}_n - A) \in \mathbb{K}[t]$$

das *charakteristische Polynom* von A , wobei

$$t \cdot \mathbb{1}_n - A = \begin{pmatrix} t - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & t - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & t - a_{nn} \end{pmatrix} \in \text{Mat}_n(\mathbb{K}[t])$$

eine quadratische Matrix mit Polynomen als Einträgen ist.

Beispiel 32.12

Für die Matrix

$$A = \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix}$$

gilt

$$\chi_A = \det \begin{pmatrix} t - 5 & -2 \\ -6 & t - 3 \end{pmatrix} = (t - 5) \cdot (t - 3) - (-2) \cdot (-6) = t^2 - 8t + 3.$$

Man beachte, daß der konstante Term von χ_A gerade $\det(A) = 3$ und daß der Koeffizient von t gerade $-\text{Spur}(A) = -8$ ist.

Proposition 32.13 (Charakteristisches Polynom)

Es sei $A \in \text{Mat}_n(\mathbb{K})$ eine quadratische Matrix. Dann ist

$$\chi_A = t^n + \alpha_{n-1} \cdot t^{n-1} + \alpha_{n-2} \cdot t^{n-2} + \dots + \alpha_1 \cdot t + \alpha_0 \in \mathbb{K}[t]$$

ein normiertes Polynom vom Grad n mit $\alpha_{n-1} = -\text{Spur}(A)$ und $\alpha_0 = (-1)^n \cdot \det(A)$.

Beweis: Ist $A = (a_{ij})$ und $t \cdot \mathbb{1}_n - A = (p_{ij})$, dann folgt aus der Leibnizschen Formel für die Determinante

$$\chi_A = \det(t \cdot \mathbb{1}_n - A) = (t - a_{11}) \cdots (t - a_{nn}) + \sum_{\text{id} \neq \sigma \in \mathbb{S}_n} \text{sgn}(\sigma) \cdot p_{1\sigma(1)} \cdots p_{n\sigma(n)}.$$

Da für $\sigma \neq \text{id}$ mindestens zwei Faktoren in $p_{1\sigma(1)} \cdots p_{n\sigma(n)}$ konstante Polynome sind, ergibt $\sum_{\text{id} \neq \sigma \in \mathbb{S}_n} \text{sgn}(\sigma) \cdot p_{1\sigma(1)} \cdots p_{n\sigma(n)}$ ein Polynom vom Grad kleiner gleich $n - 2$. Damit lassen sich die Koeffizienten α_n und α_{n-1} von t^n und t^{n-1} in χ_A aus $(t - a_{11}) \cdots (t - a_{nn})$ herleiten und sind wie oben angegeben $\alpha_n = 1$ und

$$\alpha_{n-1} = -a_{11} - a_{22} - \dots - a_{nn} = -\text{Spur}(A).$$

Ferner ist

$$\alpha_0 = \chi_A(0) = \det(-A) = (-1)^n \cdot \det(A)$$

der konstante Term im charakteristischen Polynom. □

Bemerkung 32.14

Man beachte, daß es bei der Berechnung von $\chi_A(\lambda)$ für $\lambda \in K$ keinen Unterschied macht, ob wir zuerst t durch λ ersetzen und dann die Leibnitzformel zum Berechnen der Determinante anwenden oder ob wir zuerst die Determinante berechnen und dann t durch λ ersetzen. Das liegt daran, daß der Einsetzhomomorphismus mit der Multiplikation und Addition verträglich ist, vgl. Proposition 31.11. Diese Tatsache haben wir im obigen Beweis bei der Berechnung des konstanten Terms des charakteristischen Polynoms verwendet.

Proposition 32.15 (Konjugierte Matrizen haben das gleiche charakt. Polynom.)
Sind $A, B \in \text{Mat}_n(K)$ konjugiert, so gilt $\chi_A = \chi_B$.

Beweis: Sei $T \in \text{Gl}_n(K)$ mit $B = T^{-1} \circ A \circ T$, dann gilt auch

$$T^{-1} \circ (t \cdot \mathbb{1}_n - A) \circ T = t \cdot T^{-1} \circ \mathbb{1}_n \circ T - T^{-1} \circ A \circ T = t \cdot \mathbb{1}_n - B.$$

Der Determinantenmultiplikationssatz 30.19 gilt auch für kommutative Ringe mit Eins (siehe Bemerkung 30.36), also insbesondere für Matrizen mit Einträgen im Polynomring, und somit folgt

$$\begin{aligned} \chi_B &= \det(t \cdot \mathbb{1}_n - B) = \det(T^{-1} \circ (t \cdot \mathbb{1}_n - A) \circ T) \\ &= \det(T^{-1}) \cdot \det(t \cdot \mathbb{1}_n - A) \cdot \det(T) \\ &= \frac{1}{\det(T)} \cdot \det(t \cdot \mathbb{1}_n - A) \cdot \det(T) = \chi_A. \end{aligned}$$

□

Bemerkung 32.16

Aus Proposition 32.15 folgt wegen Proposition 32.13 insbesondere, daß konjugierte Matrizen die gleiche Spur und die gleiche Determinante haben. Dies ist also ein alternativer Beweis für die wesentlichen Aussagen von Satz 32.4 und Proposition 32.8.

Damit können wir das charakteristische Polynom eines Endomorphismus definieren.

Definition 32.17 (Das charakteristische Polynom eines Endomorphismus)

Sei V ein endlich-dimensionaler K -Vektorraum der Dimension mindestens Eins mit Basis B und $f \in \text{End}_K(V)$. Wir definieren das *charakteristische Polynom* von f durch

$$\chi_f := \chi_{M_B^B(f)}.$$

Da die Matrixdarstellungen eines Endomorphismus f zu verschiedenen Basen nach Korollar 26.9 konjugiert sind, ist diese Definition unter Berücksichtigung von Proposition 32.15 unabhängig von der Wahl der Basis B .

Beispiel 32.18

Mit den Bezeichnungen aus Beispiel 32.6 gilt

$$\chi_f = \chi_{M_E^E(f)} = \det \begin{pmatrix} t-5 & -2 \\ -6 & t-3 \end{pmatrix} = t^2 - 8t + 3$$

oder alternativ

$$\chi_f = \chi_{M_B^B(f)} = \det \begin{pmatrix} t-7 & -2 \\ -2 & t-1 \end{pmatrix} = (t-7) \cdot (t-1) - 4 = t^2 - 8t + 3.$$

Bemerkung 32.19 (Normalformen bezüglich Konjugation als Ziel)

Die restlichen Abschnitte zur linearen Algebra sind folgender Aufgabe gewidmet:

Finde eine Basis B so, daß $M_B^B(f)$ eine besonders einfache Gestalt hat und wichtige Eigenschaften von f direkt aus $M_B^B(f)$ ersichtlich sind!

Alternativ kann man die Frage auch für quadratische Matrizen formulieren:

Finde ein invertierbares $T \in \text{Gl}_n(\mathbb{K})$ so, daß $T^{-1} \circ A \circ T$ eine besonders einfache Gestalt hat und wichtige Eigenschaften von A sofort sichtbar sind!

Solche *einfachen Repräsentanten* der Äquivalenzklassen bezüglich Konjugation nennt man dann *Normalformen bezüglich Konjugation*.

Ich möchte an dieser Stelle daran erinnern, daß wir uns schon mal eine ähnliche Aufgabe gestellt haben. Wir wollten Basen B und D finden, so daß die Matrixdarstellung $M_D^B(f)$ möglichst einfache Gestalt hat, oder alternativ invertierbare Matrizen S und T , so daß $S \circ A \circ T$ möglichst einfach ist. Die Aufgabe haben wir in Satz 26.28 und Korollar 26.29 gelöst und festgestellt, daß wir stets eine Matrix der Form

$$\begin{pmatrix} \mathbb{1}_r & 0 \\ 0 & 0 \end{pmatrix}$$

erhalten können, wobei r der Rang von f bzw. von A ist. Aus dieser Form kann man über die Abbildung bzw. die Matrix außer dem Rang keine interessante Information mehr ablesen. Das ist der Grund, weshalb es wichtig ist, daß wir uns von nun an auf die Situation $B = D$ bei Matrixdarstellungen bzw. $S = T^{-1}$ bei Matrizen beschränken! Und wir haben oben schon gesehen, daß bei solchen Transformationen interessante Eigenschaften wie die Determinante, die Spur und das charakteristische Polynom erhalten bleiben.

B) Eigenwerte

Der Begriff des Eigenwertes ist von zentraler Bedeutung für die in Bemerkung 32.19 angestrebte Klassifikation.

Definition 32.20 (Eigenwerte und Eigenvektoren)

Sei $f \in \text{End}_{\mathbb{K}}(V)$ und $A \in \text{Mat}_n(\mathbb{K})$.

- a. $\lambda \in \mathbb{K}$ heißt *Eigenwert* von f , falls es ein $0 \neq x \in V$ mit $f(x) = \lambda x$ gibt.

Der Vektor x heißt dann *Eigenvektor* zum Eigenwert λ von f .

$\text{Eig}(f, \lambda) := \{x \in V \mid f(x) = \lambda x\}$ heißt der *Eigenraum* von f zum Eigenwert λ .

Die Menge $\sigma(f) := \{\lambda \in \mathbb{K} \mid \lambda \text{ ist Eigenwert von } f\}$ der Eigenwerte von f heißt das *Spektrum* von f .

- b. $\lambda \in \mathbb{K}$ heißt *Eigenwert* von A , falls es ein $0 \neq x \in \mathbb{K}^n$ mit $Ax = \lambda x$ gibt.

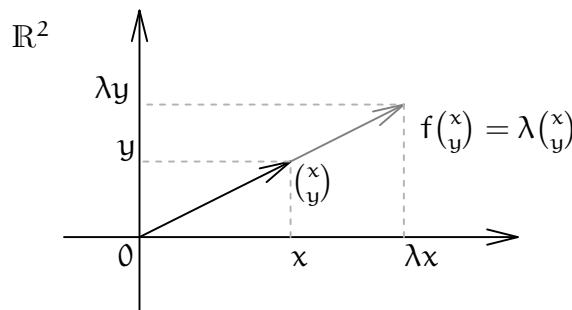
Der Vektor x heißt dann *Eigenvektor* zum Eigenwert λ von A .

$\text{Eig}(A, \lambda) := \{x \in V \mid Ax = \lambda x\}$ heißt der *Eigenraum* von A zum Eigenwert λ .

Die Menge $\sigma(A) := \{\lambda \in \mathbb{K} \mid \lambda \text{ ist Eigenwert von } A\}$ der Eigenwerte von A heißt das *Spektrum* von A .

Bemerkung 32.21 (Geometrische Interpretation von Eigenvektoren)

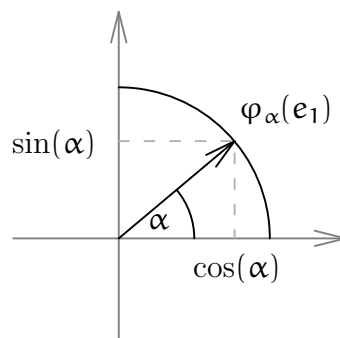
Ist λ Eigenwert von f mit Eigenvektor x , so bedeutet das anschaulich, daß f in *Richtung* von x durch Multiplikation mit λ wirkt. Diese Anschauung liefert im Fall $V = \mathbb{R}^n$ und $\lambda > 0$, daß f den Vektor x um den Faktor λ streckt, falls $\lambda > 1$, und um den Faktor λ staucht, falls $0 < \lambda < 1$.



Beispiel 32.22

- a. Ist $\dim_{\mathbb{K}}(V) = 1$, so ist jeder Vektor ungleich Null ein Eigenvektor von f , da f schlicht die Multiplikation mit einer Konstanten ist.
- b. Ist $\dim_{\mathbb{K}}(V) \geq 2$, so braucht f hingegen keine Eigenwerte und Eigenvektoren zu besitzen. Dabei hängt die Frage der Existenz wesentlich vom Grundkörper \mathbb{K} ab. Betrachte etwa die Drehung $\varphi_{\alpha} = f_{A_{\alpha}} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ um den Winkel $\alpha \in \mathbb{R}$ aus Beispiel 23.24. Die Matrixdarstellung bezüglich der kanonischen Basis $E = (e_1, e_2)$ ist

$$A_{\alpha} = M_E^E(\varphi_{\alpha}) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$



Aus einer rein geometrischen Betrachtung folgt unmittelbar, daß φ_α bzw. A_α nur dann einen Eigenvektor besitzen können, wenn α ein ganzzahliges Vielfaches von π ist.

Bemerkung 32.23 (Eigenräume)

Es seien $f \in \text{End}_K(V)$ und $A \in \text{Mat}_n(K)$.

- a. Da $f(x) = \lambda x$ für $x \in V$ und $\lambda \in K$ genau dann erfüllt ist, wenn x im Kern der linearen Abbildung $f - \lambda \text{id}_V \in \text{End}_K(V)$ liegt, gilt also

$$\text{Eig}(f, \lambda) = \text{Ker}(f - \lambda \text{id}_V) = \text{Ker}(\lambda \text{id}_V - f).$$

Analog erhält man:

$$\text{Eig}(A, \lambda) = \text{Ker}(f_A - \lambda \text{id}_V) = \text{Lös}(A - \lambda \mathbf{1}_n, 0) = \text{Lös}(\lambda \mathbf{1}_n - A, 0).$$

- b. Aus der Definition folgt unmittelbar, daß $\sigma(A) = \sigma(f_A)$ und $\sigma(f) = \sigma(M_B^B(f))$.
 c. Ebenso folgt unmittelbar, daß der Eigenraum $\text{Eig}(f, \lambda)$ von f zum Eigenwert λ f -invariant ist.
 d. Kennt man einen Eigenwert $\lambda \in K$ von A , so kann man das lineare Gleichungssystem

$$(A - \lambda \mathbf{1}_n)x = 0 \quad \text{oder} \quad (\lambda \mathbf{1}_n - A)x = 0$$

lösen und damit eine Basis des Eigenraumes $\text{Eig}(A, \lambda) = \text{Lös}(A - \lambda \mathbf{1}_n, 0)$ bestimmen. D. h., bei Kenntnis des Eigenwertes λ lassen sich die Eigenvektoren von A zu λ durch Lösen eines linearen Gleichungssystems bestimmen. Aber wie kommt man zu den Eigenwerten von A ?

Satz 32.24 (Eigenwerte und das charakteristische Polynome)

Es sei $f \in \text{End}_K(V)$ und $A \in \text{Mat}_n(K)$.

- a. Die Eigenwerte von f sind genau die Nullstellen von χ_f in K .
 b. Die Eigenwerte von A sind genau die Nullstellen von χ_A in K .

Inbesondere, f und A haben höchstens n paarweise verschiedene Eigenwerte.

Beweis: Für $\lambda \in K$ gilt unter Berücksichtigung von Korollar 25.22:

$$\begin{aligned} \lambda \text{ ist Eigenwert von } f &\iff \text{Ker}(\lambda \text{id}_V - f) = \text{Eig}(f, \lambda) \neq \{0\} \\ &\iff \lambda \text{id}_V - f \text{ ist nicht injektiv} \\ &\stackrel{25.22}{\iff} \lambda \text{id}_V - f \text{ ist nicht bijektiv} \\ &\iff \chi_f(\lambda) = \det(\lambda \text{id}_V - f) = 0. \end{aligned}$$

Der Beweis für die Matrizen geht analog. □

Bevor wir das charakteristische Polynom weiter untersuchen, wollen wir zunächst einige Beispiele betrachten.

Beispiel 32.25

a. Betrachten wir zunächst die folgende Matrix:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix} \in \text{Mat}(3, \mathbb{Q}).$$

Mit Hilfe der Regel von Sarrus oder durch den Laplaceschen Entwicklungssatz bestimmen wir das charakteristische Polynom von A als

$$\chi_A = \det \begin{pmatrix} t & -1 & -1 \\ 1 & t-2 & -1 \\ 1 & -1 & t-2 \end{pmatrix} = t^3 - 4t^2 + 5t - 2 = (t-1)^2 \cdot (t-2).$$

Alternativ kann man allgemein die Matrix $t\mathbb{1}_n - A \in \text{Mat}_n(\mathbb{Q}(t))$ auch als Matrix über dem Körper $\mathbb{Q}(t)$ auffassen (siehe Bemerkung 31.21). Da $\mathbb{Q}(t)$ ein Körper ist, dürfen wir die Determinante mittels des Gaußschen Algorithmus 30.15 bestimmen. Insbesondere dürfen wir dabei durch Polynome dividieren!

$$\begin{aligned} \begin{pmatrix} t & -1 & -1 \\ 1 & t-2 & -1 \\ 1 & -1 & t-2 \end{pmatrix} &\xrightarrow{\substack{\text{II} \rightarrow \text{II} - \frac{1}{t} \text{I} \\ \text{III} \rightarrow \text{III} - \frac{1}{t} \text{I}}} \begin{pmatrix} t & -1 & -1 \\ 0 & t-2 + \frac{1}{t} & \frac{1}{t} - 1 \\ 0 & \frac{1}{t} - 1 & t-2 + \frac{1}{t} \end{pmatrix} = \\ &\xrightarrow{\text{III} \rightarrow \text{III} + \frac{1}{t-1} \text{II}} \begin{pmatrix} t & -1 & -1 \\ 0 & \frac{(t-1)^2}{t} & -\frac{t-1}{t} \\ 0 & 0 & t-2 \end{pmatrix}. \end{aligned} \quad (91)$$

Entsprechend erhalten wir für das charakteristische Polynom

$$\chi_A = t \cdot \frac{(t-1)^2}{t} \cdot (t-2) = (t-1)^2 \cdot (t-2).$$

Das charakteristische Polynom hat also die Nullstellen $\lambda = 1$ und $\lambda = 2$, wobei $\lambda = 1$ eine zweifache Nullstelle ist. Insbesondere ist also $\sigma(A) = \{1, 2\}$.

Wir können jetzt für $\lambda = 1$ und für $\lambda = 2$ jeweils den Eigenraum $\text{Eig}(A, \lambda) = \text{Lös}(\lambda\mathbb{1}_n - A, 0)$ mit Hilfe des Gauß-Algorithmus bestimmen.³

Der Algorithmus zur Bestimmung von $\text{Eig}(A, 1) = \text{Lös}(\mathbb{1}_n - A, 0)$ sieht vor, daß wir die Matrix zunächst auf reduzierte ZSF bringen und dann in den Nullzeilen die Diagonalelemente durch -1 ersetzen:

$$\begin{pmatrix} -1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

³Man beachte, daß es zur Berechnung der reduzierten Zeilen-Stufen-Form von $\lambda\mathbb{1}_n - A$ für $\lambda = 1$ *nicht* erlaubt ist, in (91) in der letzten Matrix t etwa durch $\lambda = 1$ zu ersetzen, um die ZSF zu erhalten, da wir bei den vorgenommenen Umformungen zur Ermittlung obiger Matrix durch das Polynom $t-1$ dividiert haben. Dies ist über $\mathbb{Q}(t)$ eine erlaubte Operation gewesen. Ersetzen wir jedoch t durch 1 , so ist die Operation nicht mehr erlaubt!

Die letzten beiden Spalten, d.h. die, bei denen eine -1 auf der Diagonalen steht, bilden dann eine Basis des Eigenraumes zum Eigenwert 1 :

$$\text{Eig}(A, 1) = \text{Lin}((-1, -1, 0)^t, (-1, 0, -1)^t).$$

$\text{Eig}(A, 1)$ ist also zweidimensional.

Analog ergibt sich $\text{Eig}(A, 2)$ aus

$$\begin{pmatrix} -2 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & 1 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \end{pmatrix},$$

und damit gilt $\text{Eig}(A, 2) = \text{Lin}((-1, -1, -1)^t)$.

- b. Wir hatten schon durch eine geometrische Argumentation gesehen, daß die Drehung um einen Winkel α im allgemeinen keinen reellen Eigenwert besitzt. Den gleichen Sachverhalt prüfen wir nun noch einmal mit algebraischen Methoden. Die Matrixdarstellung der Drehung bezüglich der kanonischen Basis von \mathbb{R}^2 ist

$$A_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

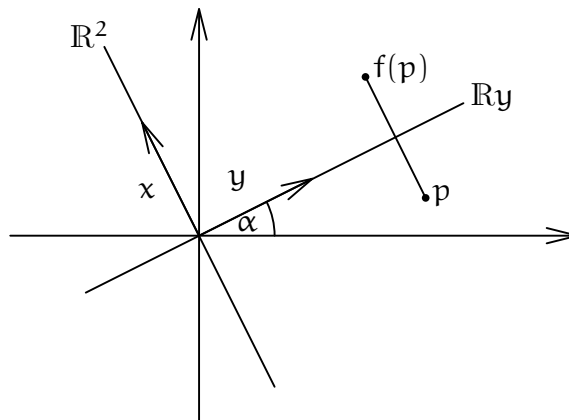
Aber

$$\chi_{A_\alpha} = (t - \cos(\alpha))^2 + \sin^2(\alpha) = t^2 - 2\cos(\alpha)t + 1.$$

Die Nullstellen von χ_{A_α} sind $\cos(\alpha) + \sqrt{\cos^2(\alpha) - 1}$ und $\cos(\alpha) - \sqrt{\cos^2(\alpha) - 1}$. Für beide Terme gilt, sie sind genau dann reell, wenn α ein ganzzahliges Vielfaches von π ist.

Insbesondere hat die Drehung also nur dann reelle Eigenwerte, wenn α ein ganzzahliges Vielfaches von π ist, d. h. $A_\alpha = \mathbb{1}_2$ oder $A_\alpha = -\mathbb{1}_2$.

- c. Es sei $f \in \text{End}_{\mathbb{R}}(\mathbb{R}^2)$ die Spiegelung an einer Geraden $\text{Lin}(\mathbf{y}) = \mathbb{R} \cdot \mathbf{y} \subset \mathbb{R}^2$ mit $0 \neq \mathbf{y} = (y_1, y_2) \in \mathbb{R}^2$.



Wir setzen $\mathbf{x} = (-y_2, y_1) \in \mathbb{R}^2$. Dann steht \mathbf{x} senkrecht auf \mathbf{y} und $B = (\mathbf{y}, \mathbf{x})$ ist eine Basis von \mathbb{R}^2 . Die Spiegelung f bildet mithin \mathbf{y} auf sich selbst und \mathbf{x} auf $-\mathbf{x}$ ab, da \mathbf{x} senkrecht auf $\text{Lin}(\mathbf{y})$ steht. Damit hat f die folgende Matrixdarstellung

bezüglich B

$$M_B^B(f) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

und das charakteristische Polynom von f ist gerade

$$\chi_f = (t - 1) \cdot (t + 1).$$

Die Spiegelung von f hat also Spektrum $\sigma(f) = \{-1, 1\}$.

Beschreiben wir f in den Standardkoordinaten $E = (e_1, e_2)$ von \mathbb{R}^2 , so ist f die Spiegelung an $\text{Lin}(e_1) = \mathbb{R} \cdot e_1$ gefolgt von der Drehung um den Winkel 2α , wenn α der Winkel ist, den $\text{Lin}(y)$ mit $\text{Lin}(e_1)$ einschließt. Wir erhalten also

$$M_E^E(f) = \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix} = \begin{pmatrix} \cos(2\alpha) & -\sin(2\alpha) \\ \sin(2\alpha) & \cos(2\alpha) \end{pmatrix} \circ \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Das charakteristische Polynom errechnet sich aus dieser Matrixdarstellung als

$$(t - \cos(2\alpha)) \cdot (t + \cos(2\alpha)) - \sin^2(2\alpha) = t^2 - 1 = (t - 1) \cdot (t + 1).$$

Korollar 32.26 (Eigenwerte einer Dreiecksmatrix)

Ist $A = (a_{ij}) \in \text{Mat}_n(\mathbb{K})$ eine obere oder untere Dreiecksmatrix, dann ist

$$\chi_A = (t - a_{11}) \cdot \dots \cdot (t - a_{nn})$$

und die Einträge auf der Diagonalen sind genau die Eigenwerte von A .

Beweis: Für eine obere Dreiecksmatrix $A = (a_{ij})$ gilt

$$\chi_A = \det \begin{pmatrix} t - a_{11} & * & \dots & \dots & * \\ 0 & t - a_{22} & * & \dots & * \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & 0 & t - a_{nn} \end{pmatrix} = \prod_{i=1}^n (t - a_{ii}),$$

und diese Polynom hat genau die Nullstellen a_{11}, \dots, a_{nn} . Der Beweis für untere Dreiecksmatrizen geht analog. \square

Definition 32.27 (Vielfachheit von Eigenwerten)

Es sei $f \in \text{End}_{\mathbb{K}}(V)$, $A \in \text{Mat}_n(\mathbb{K})$ und $\lambda \in \mathbb{K}$.

- a. $\text{mult}(\chi_f, \lambda)$ heißt *algebraische Vielfachheit* von λ als Eigenwert von f .
 $\dim_{\mathbb{K}} \text{Eig}(f, \lambda)$ heißt *geometrische Vielfachheit* von λ als Eigenwert von f .
- b. $\text{mult}(\chi_A, \lambda)$ heißt *algebraische Vielfachheit* von λ als Eigenwert von A .
 $\dim_{\mathbb{K}} \text{Eig}(A, \lambda)$ heißt *geometrische Vielfachheit* von λ als Eigenwert von A .

Die algebraischen Vielfachheiten nennt man auch *arithmetische Vielfachheiten*.

Beispiel 32.28

Die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \text{Mat}_2(\mathbb{C})$$

aus Beispiel 33.4 hat nur den Eigenwert 0, da $\chi_A = t^2$. Die algebraische Vielfachheit von 0 als Eigenwert von A ist

$$\text{mult}(\chi_A, 0) = \text{mult}(t^2, 0) = 2,$$

während die geometrische Vielfachheit

$$\dim_{\mathbb{C}} \text{Eig}(A, 0) = \dim_{\mathbb{C}} \text{Lös}(A, 0) = \dim_{\mathbb{C}} \text{Lin}((1, 0)^t) = 1$$

ist.

Lemma 32.29 (Geometrische und algebraische Vielfachheit)

Es sei $f \in \text{End}_{\mathbb{K}}(V)$, $A \in \text{Mat}_n(\mathbb{K})$ und $\lambda \in \mathbb{K}$. Dann gilt stets

$$\dim_{\mathbb{K}} \text{Eig}(f, \lambda) \leq \text{mult}(\chi_f, \lambda) \quad \text{und} \quad \dim_{\mathbb{K}} \text{Eig}(A, \lambda) \leq \text{mult}(\chi_A, \lambda),$$

d.h. die geometrische Vielfachheit eines Eigenwertes ist stets nach oben durch die algebraische Vielfachheit beschränkt.

Beweis: Man beachte, daß $U := \text{Eig}(f, \lambda)$ ein f -invarianter Unterraum ist und daß $f|_U = \lambda \cdot \text{id}_U$ gilt. Mithin ist

$$\chi_{f|_U} = \chi_{\lambda \cdot \text{id}_U} = \det(t \cdot \text{id}_U - \lambda \cdot \text{id}_U) = \det((t - \lambda) \cdot \text{id}_U) = (t - \lambda)^s$$

wobei $s = \dim_{\mathbb{K}}(U) = \dim_{\mathbb{K}} \text{Eig}(f, \lambda)$. Außerdem gilt nach Aufgabe 32.32

$$\chi_f = \chi_{f|_U} \cdot \chi_{f|_{V/U}} = (t - \lambda)^s \cdot \chi_{f|_{V/U}}.$$

Daraus folgt unmittelbar

$$\text{mult}(f, \lambda) \geq s = \dim_{\mathbb{K}} \text{Eig}(f, \lambda).$$

Die analoge Aussage für A folgt hieraus mit $f = f_A$. □

Lemma 32.30 (Eigenwerte bei konjugierten Matrizen)

Für $A, B \in \text{Mat}_n(\mathbb{K})$ und $T \in \text{GL}_n(\mathbb{K})$ mit $B = T^{-1} \circ A \circ T$ sowie $\lambda \in \mathbb{K}$ gelten:

- a. $\sigma(A) = \sigma(B)$.
- b. $\text{mult}(\chi_A, \lambda) = \text{mult}(\chi_B, \lambda)$.
- c. $\dim_{\mathbb{K}} \text{Eig}(A, \lambda) = \dim_{\mathbb{K}} \text{Eig}(B, \lambda)$.
- d. $x \in \text{Eig}(A, \lambda) \iff T^{-1}x \in \text{Eig}(B, \lambda)$.

D.h. konjugierte Matrizen haben die gleichen Eigenwerte und für jeden Eigenwert stimmen ihre geometrischen Vielfachheiten ebenso überein wie ihre algebraischen Vielfachheiten.

Beweis: Nach Proposition 32.15 haben A und B die gleichen charakteristischen Polynome. Mithin stimmen wegen Satz 32.24 die Eigenwerte von A und B sowie deren algebraische Vielfachheiten überein. Damit sind a. und b. gezeigt. Ferner gilt

$$\begin{aligned} x \in \text{Eig}(A, \lambda) &\iff \lambda x = Ax = A T T^{-1} x \\ &\iff \lambda T^{-1} x = T^{-1} A T T^{-1} x = B T^{-1} x \iff T^{-1} x \in \text{Eig}(B, \lambda). \end{aligned}$$

Damit ist d. gezeigt und außerdem folgt, daß der Isomorphismus f_T den Eigenraum $\text{Eig}(A, \lambda)$ isomorph auf den Eigenraum $\text{Eig}(B, \lambda)$ abbildet. Die beiden müssen also die gleiche Dimension haben, womit auch c. gezeigt ist. \square

Aufgaben

Aufgabe 32.31 (Nilpotente Endomorphismen und Matrizen)

Es sei $A \in \text{Mat}_n(\mathbb{K})$ und $f \in \text{End}_{\mathbb{K}}(V)$ mit $1 \leq \dim_{\mathbb{K}}(V) < \infty$.

- Zeige, gibt es ein $r \in \mathbb{N}$ mit $f^r = 0$, so gilt $\text{Spur}(f) = 0$.
- Zeige, gibt es ein $r \in \mathbb{N}$ mit $A^r = 0$, so gilt $\text{Spur}(A) = 0$.
- Finde ein Beispiel für eine Matrix wie in Teil b., bei der nicht alle Diagonalelemente Null sind.

Hinweis zum Beweis von a.: Führe Induktion über $n = \dim_{\mathbb{K}}(V)$. Dazu zeige man, daß $M_B^B(f)$ für eine geeignete Wahl von B Blockgestalt mit einem Nullblock in der oberen linken Ecke hat. Proposition 26.16b. mit $U = \text{Ker}(f)$ ist dabei hilfreich.

Aufgabe 32.32

Es sei $1 \leq \dim_{\mathbb{K}}(V) < \infty$, $f \in \text{End}_{\mathbb{K}}(V)$.

- Ist $U \subseteq V$ ein f -invarianter Unterraum, dann gilt

$$\det(f) = \det(f|_U) \cdot \det(f|_{V/U})$$

und

$$\chi_f = \chi_{f|_U} \cdot \chi_{f|_{V/U}}.$$

- Ist $V = U_1 \oplus \dots \oplus U_k$, wobei die U_i f -invariant seien, dann gilt

$$\det(f) = \det(f|_{U_1}) \cdot \dots \cdot \det(f|_{U_k})$$

und

$$\chi_f = \chi_{f|_{U_1}} \cdot \dots \cdot \chi_{f|_{U_k}}.$$

Aufgabe 32.33

Für ein Polynom $p \in \mathbb{K}[t]$ und zwei konjugierte Matrizen $A, B \in \text{Mat}_n(\mathbb{K})$ gilt

$$p(A) = 0 \iff p(B) = 0.$$

Aufgabe 32.34 (Zyklische Unterräume)

Zeige, $\chi_{f|_U} = \mu_{f|_U} = t^m$ für den Endomorphismus $f|_U$ aus Aufgabe 26.35.

Aufgabe 32.35

Bestimme die Eigenwerte und die Eigenräume der folgenden Matrix A und entscheide, ob sie diagonalisierbar bzw. trigonalisierbar ist:

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 3 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix} \in \text{Mat}_4(\mathbb{Q}).$$

Aufgabe 32.36 (Die Eigenräume bilden eine direkte Summe.)

Es sei $f \in \text{End}_{\mathbb{K}}(\mathbb{V})$ und $A \in \text{Mat}_n(\mathbb{K})$.

- a. Sind $x_1, \dots, x_r \in \mathbb{V}$ Eigenvektoren von f zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r \in \mathbb{K}$, dann ist die Familie (x_1, \dots, x_r) linear unabhängig. Insbesondere gilt

$$\text{Eig}(f, \lambda_1) + \dots + \text{Eig}(f, \lambda_r) = \text{Eig}(f, \lambda_1) \oplus \dots \oplus \text{Eig}(f, \lambda_r).$$

- b. Sind $x_1, \dots, x_r \in \mathbb{K}^n$ Eigenvektoren von A zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r \in \mathbb{K}$, dann ist die Familie (x_1, \dots, x_r) linear unabhängig. Insbesondere gilt

$$\text{Eig}(A, \lambda_1) + \dots + \text{Eig}(A, \lambda_r) = \text{Eig}(A, \lambda_1) \oplus \dots \oplus \text{Eig}(A, \lambda_r).$$

Aufgabe 32.37

Sind $f, g \in \text{End}_{\mathbb{K}}(\mathbb{V})$, so gilt $\sigma(f \circ g) = \sigma(g \circ f)$.

Aufgabe 32.38

Sind $f, g \in \text{End}_{\mathbb{K}}(\mathbb{V})$, so gilt $\sigma(f \circ g) = \sigma(g \circ f)$.

§ 33 Diagonalisierbarkeit und Trigonalisierbarkeit

In diesem Abschnitt sei V ein K -Vektorraum mit $1 \leq \dim_K(V) = n < \infty$.

In Bemerkung 32.19 haben wir erläutert, daß unser zentrales Anliegen darin besteht, eine Basis B bzw. eine invertierbare Matrix $T \in \text{Gl}_n(K)$ zu finden, so daß $M_B^B(f)$ bzw. $T^{-1} \circ A \circ T$ eine möglichst einfache Gestalt hat. Dazu zählt sicher, daß die Matrix möglichst viele Nullen enthält.

Definition 33.1 (Diagonalisierbar und trigonalisierbar)

Es sei $f \in \text{End}_K(V)$ und $A \in \text{Mat}_n(K)$.

- a. f heißt *diagonalisierbar* (bzw. *trigonalisierbar*), falls es eine Basis B von V gibt, so daß $M_B^B(f)$ eine Diagonalmatrix (bzw. eine obere Dreiecksmatrix) ist.
- b. A heißt *diagonalisierbar* (bzw. *trigonalisierbar*), falls es eine Matrix $T \in \text{Gl}_n(K)$ gibt, so daß $T^{-1} \circ A \circ T$ eine Diagonalmatrix (bzw. eine obere Dreiecksmatrix) ist.

A) Trigonalisierbarkeit

Satz 33.2 (Trigonalisierbarkeit)

Es sei $f \in \text{End}_K(V)$ und $A \in \text{Mat}_n(K)$.

- a. Genau dann ist f trigonalisierbar, wenn χ_f über K in Linearfaktoren zerfällt.
- b. Genau dann ist A trigonalisierbar, wenn χ_A über K in Linearfaktoren zerfällt.

Beweis: Ist f trigonalisierbar, so gibt es eine Basis B mit

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & * & \dots & \dots & * \\ 0 & \lambda_2 & * & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & \lambda_{n-1} & * \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix}.$$

Damit folgt, daß das charakteristische Polynom

$$\chi_f = (t - \lambda_1) \cdots (t - \lambda_n)$$

von f über K in Linearfaktoren zerfällt.

Zerfälle nun umgekehrt das charakteristische Polynom von f in Linearfaktoren $\chi_f = (t - \lambda_1) \cdots (t - \lambda_n)$. Wir beweisen mit Induktion über $n = \dim_K(V)$, daß dann f trigonalisierbar ist. Im Fall $n = 1$ ist f nach Beispiel 32.22 sogar diagonalisierbar.

Sei also $n > 1$ und sei $0 \neq x_1 \in V$ ein Eigenvektor von f zum Eigenwert λ_1 . Wir setzen $U := \text{Lin}(x_1) \leq V$. Wegen $f(x_1) = \lambda_1 x_1 \in U$ ist U ein f -invarianter Unterraum

von V und $\chi_{f_U} = t - \lambda_1$. Mithin folgt aus Aufgabe 32.32

$$\chi_{f_{V/U}} = (t - \lambda_2) \cdots (t - \lambda_n),$$

d. h. das charakteristische Polynom von $f_{V/U}$ zerfällt über K in Linearfaktoren. Da $\dim_K(V/U) = n - 1 < n$, existiert per Induktion eine Basis $B'' = (\bar{x}_2, \dots, \bar{x}_n)$ von V/U , so daß $M_{B''}^{B''}(f_{V/U})$ eine obere Dreiecksmatrix ist. Dann ist aber $B = (x_1, \dots, x_n)$ eine Basis von V und mit $B' = (x_1)$ gilt wegen Proposition 26.16

$$M_B^B(f) = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array} \right) = \left(\begin{array}{c|c} \lambda_1 & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array} \right).$$

Damit ist $M_B^B(f)$ eine obere Dreiecksmatrix und f ist trigonalisierbar.

Die Aussage für A erhalten wir aus der entsprechenden Aussage für f_A . \square

Bemerkung 33.3

Ist K ein algebraisch abgeschlossener Körper, etwa $K = \mathbb{C}$, so sind somit jede Matrix A und jeder Endomorphismus f trigonalisierbar. Eine vergleichbare Aussage für die Diagonalisierbarkeit gilt nicht.

Beispiel 33.4

a. Die Drehmatrix

$$A_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

hat das charakteristische Polynom $\chi_{A_\alpha} = t^2 - 2\cos(\alpha)t + 1 = (t - \lambda) \cdot (t - \bar{\lambda})$ mit $\lambda = \cos(\alpha) + i\sin(\alpha) \in \mathbb{C}$, $\alpha \in \mathbb{R}$. Damit hat χ_{A_α} also keine reellen Nullstellen, wenn α kein ganzzahliges Vielfaches von π ist, und somit ist A_α über \mathbb{R} nicht trigonalisierbar.

Hingegen zerfällt χ_{A_α} über \mathbb{C} in Linearfaktoren, so daß A_α über \mathbb{C} trigonalisierbar sein muß. In der Tat ist A_α sogar diagonalisierbar mit Diagonalgestalt

$$\begin{pmatrix} \lambda & 0 \\ 0 & \bar{\lambda} \end{pmatrix}.$$

Ist α kein ganzzahliges Vielfaches von π , so besitzt A_α zwei verschiedene Eigenwerte, so daß zugehörige Eigenvektoren nach Aufgabe 32.36 eine Basis von \mathbb{C}^2 bilden müssen und diese transformieren A_α in obige Diagonalmatrix. Ist α hingegen ein ganzzahliges Vielfaches von π , so ist $A_\alpha = \mathbf{1}_2$ oder $A_\alpha = -\mathbf{1}_2$ und hat bereits Diagonalgestalt.

b. Die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \text{Mat}_2(\mathbb{C})$$

ist hingegen auch über \mathbb{C} nicht diagonalisierbar. Denn, gäbe es eine Matrix $T \in \text{Gl}_2(\mathbb{C})$ mit

$$T^{-1} \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \in \text{Mat}_2(\mathbb{C}),$$

dann wäre

$$\begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda_2^2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^2 = T^{-1} \circ A^2 \circ T = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

also wären $\lambda_1 = \lambda_2 = 0$. Aber damit würde gelten:

$$0 = \text{rang} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \text{rang}(T^{-1} \circ A \circ T) = \text{rang}(A) = 1,$$

da $T \in \text{Gl}_2(\mathbb{C})$. Dies ist jedoch ein Widerspruch.

B) Diagonalblockmatrizen

Definition 33.5 (Diagonalblockmatrizen)

Wir werden im Folgenden sehr häufig mit Blockmatrizen der folgenden Form arbeiten:

$$A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & A_r \end{pmatrix} \in \text{Mat}_n(\mathbb{K}),$$

wobei $A_i \in \text{Mat}_{n_i}(\mathbb{K})$, $i = 1, \dots, r$ mit $n = n_1 + \dots + n_r$. Es empfiehlt sich deshalb, eine Kurzschreibweise für solche *Diagonalblockmatrizen* einzuführen. Wir schreiben kurz:

$$A = A_1 \oplus \dots \oplus A_r = \bigoplus_{i=1}^r A_i.$$

Bemerkung 33.6 (Diagonalblockmatrizen)

- Man beachte, daß es bei der obigen Schreibweise für Diagonalblockmatrizen auf die Reihenfolge der Summation ankommt, daß aber Matrizen, die durch Änderung der Summationsreihenfolge entstehen, zueinander konjugiert sind!
- Mit Hilfe dieser Notation gilt beispielsweise, daß eine Matrix A genau dann diagonalisierbar ist, wenn es Körperelemente $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ und positive natürliche Zahlen $n_1, \dots, n_r \in \mathbb{N}$ gibt sowie eine invertierbare Matrix $T \in \text{Gl}_n(\mathbb{K})$ mit

$$T^{-1} \circ A \circ T = \bigoplus_{i=1}^r \lambda_i \mathbb{1}_{n_i}.$$

- c. Ist $A = \bigoplus_{i=1}^r A_i$ eine Diagonalblockmatrix, so verifiziert man leicht, daß für $k \in \mathbb{N}$ gilt $A^k = \bigoplus_{i=1}^r A_i^k$, und damit, daß für ein Polynom $p \in K[t]$ gilt

$$p(A) = \bigoplus_{i=1}^r p(A_i).$$

Insbesondere gilt also für eine Diagonalmatrix $D = \bigoplus_{i=1}^n \lambda_i \mathbb{1}_1$, daß

$$p(D) = \bigoplus_{i=1}^n p(\lambda_i) \mathbb{1}_1 = \begin{pmatrix} p(\lambda_1) & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & p(\lambda_n) \end{pmatrix}$$

In der Tat kann man sogar zeigen, daß für eine Blockmatrix der Form

$$A = \begin{pmatrix} A_1 & * & \dots & * \\ 0 & A_2 & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & A_r \end{pmatrix} \in \text{Mat}_n(K),$$

gilt, daß

$$p(A) = \begin{pmatrix} p(A_1) & * & \dots & * \\ 0 & p(A_2) & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & p(A_r) \end{pmatrix} \in \text{Mat}_n(K),$$

wobei sich die Sterne oberhalb der Blöcke verändert haben.

Damit gilt insbesondere, daß $p(A)$ eine obere Dreiecksmatrix ist, falls A eine solche war.

C) Der Satz von Cayley-Hamilton

Da $\dim_K(\text{Mat}_n(K)) = n^2$ gilt, sind die $n^2 + 1$ Matrizen

$$\mathbb{1}_n = A^0, A^1, A^2, \dots, A^{n^2}$$

in $\text{Mat}_n(K)$ linear abhängig. D. h. es existieren $\lambda_0, \dots, \lambda_{n^2} \in K$, nicht alle null, mit

$$\lambda_0 A^0 + \lambda_1 A^1 + \dots + \lambda_{n^2} A^{n^2} = 0 \in \text{Mat}_n(K).$$

Ein einfaches Dimensionsargument zeigt also, es gibt ein Polynom $0 \neq p = \lambda_{n^2} t^{n^2} + \dots + \lambda_0 \in K[t]$ vom Grad kleiner gleich n^2 mit $p(A) = 0$. Der folgende wichtige Satz von Cayley-Hamilton besagt nun, daß es sogar ein Polynom vom Grad n gibt, das A annulliert.

Satz 33.7 (Cayley-Hamilton)

Für $f \in \text{End}_K(V)$ und $A \in \text{Mat}_n(K)$ gilt $\chi_f(f) = 0$ und $\chi_A(A) = 0$.

Beweis: Da für eine Basis D von V die Abbildung $M_D^D: \text{End}_K(V) \rightarrow \text{Mat}_n(K)$ ein K -Algebrenhomomorphismus ist, gilt

$$M_D^D(\chi_f(f)) = \chi_f(M_D^D(f)).$$

Dann gilt aber $\chi_f(f) = 0$ genau dann, wenn

$$0 = M_D^D(\chi_f(f)) = \chi_f(M_D^D(f)) = \chi_{M_{B_D^D}(f)}(M_D^D(f)).$$

Es reicht deshalb, die Aussage für Matrizen zu beweisen.

Betrachte dazu die Matrix

$$B_t := t \cdot \mathbb{1}_n - A \in \text{Mat}_n(K[t])$$

sowie die Adjunkte $B_t^\# = (p_{ij}) \in \text{Mat}_n(K[t])$ von B_t , die auch *Busadjunkte* von A genannt wird. Nach dem Satz über die Adjunkte 30.28 in $\text{Mat}_n(K[t])$ gilt die Adjunktengleichung

$$B_t \circ B_t^\# = (t\mathbb{1}_n - A) \circ (t\mathbb{1}_n - A)^\# = \det(t\mathbb{1}_n - A) \cdot \mathbb{1}_n = \chi_\lambda \cdot \mathbb{1}_n. \quad (92)$$

Man beachte nun noch, daß die Einträge von $B_t^\#$ Determinanten von gewissen $(n-1) \times (n-1)$ -Matrizen von B_t sind, also Polynome vom Grad höchstens $n-1$. Wir können nun $B_t^\#$ auch als Polynom schreiben, dessen Koeffizienten Matrizen sind, und dieses Polynom hat dann höchstens den Grad $n-1$, d. h. es gibt Matrizen $B_0, \dots, B_{n-1} \in \text{Mat}_n(K)$ mit

$$B_t^\# = B_{n-1}t^{n-1} + \dots + B_1t + B_0.$$

Ist $\chi_\lambda = t^n + \alpha_{n-1}t^{n-1} + \dots + \alpha_0$, so folgt aus der Adjunktengleichung (92)

$$(\mathbb{1}_n t - A) \circ (B_{n-1}t^{n-1} + \dots + B_1t + B_0) \stackrel{!}{=} \mathbb{1}_n t^n + \alpha_{n-1}\mathbb{1}_n t^{n-1} + \dots + \alpha_0\mathbb{1}_n \quad (93)$$

durch Koeffizientenvergleich für die t^i , $i = 0, \dots, n$:

$$\begin{aligned} B_{n-1} &= \mathbb{1}_n \\ -AB_{n-1} + B_{n-2} &= \alpha_{n-1}\mathbb{1}_n \\ -AB_{n-2} + B_{n-3} &= \alpha_{n-2}\mathbb{1}_n \\ &\vdots \\ -AB_1 + B_0 &= \alpha_1\mathbb{1}_n \\ -AB_0 &= \alpha_0\mathbb{1}_n \end{aligned} \quad (94)$$

Multipliziert man die i -te Zeile in (94) mit A^{n-i+1} und summiert die beiden Seiten auf, so erhält man die Behauptung:

$$\begin{array}{rcl}
 A^n B_{n-1} & = & A^n \\
 -A^n B_{n-1} + A^{n-1} B_{n-2} & = & \alpha_{n-1} A^{n-1} \\
 -A^{n-1} B_{n-2} + A^{n-2} B_{n-3} & = & \alpha_{n-2} A^{n-2} \\
 & \vdots & \\
 -A^2 B_1 + A B_0 & = & \alpha_1 A \\
 -A B_0 & = & \alpha_0 \mathbf{1}_n \\
 \hline
 0 & = & \chi_A(A).
 \end{array}$$

□

Bemerkung 33.8

- a. Man beachte, daß der folgende *offensichtliche* Beweis für $\chi_A(A) = 0$, nämlich

$$\chi_A(A) = \det(A * \mathbf{1}_n - A) = \det(0) = 0$$

falsch ist, da “*” beim Einsetzen von A in $\det(t\mathbf{1}_n - A) \in K[t]$ eben *nicht* die Matrixmultiplikation ist! Man beachte ferner, daß die Gleichung auch schon deshalb keinen Sinn ergeben kann, da $\chi_A(A)$ die Nullmatrix ist, während $\det(0)$ die Null in K ist.

- b. Kennt man das charakteristische Polynom $\chi_A = t^n + \alpha_{n-1}t^{n-1} + \dots + \alpha_0$, so läßt sich daraus mittels (93) und der Rekursionsformel (94) die Busadjunkte

$$(t\mathbf{1}_n - A)^\# = B_{n-1}t^{n-1} + \dots + B_1t + B_0$$

von A bestimmen. Für die B_{n-k} , $k = 1, \dots, n$, gilt dabei explizit:

$$B_{n-k} = A^{k-1} + \alpha_{n-1}A^{k-2} + \dots + \alpha_{n-k+1}A^0,$$

und speziell gilt für die Adjunkte von A

$$A^\# = B_0 = A^{n-1} + \alpha_{n-1}A^{n-2} + \dots + \alpha_1A^0.$$

Diese Formel zur Berechnung der Adjunkten von A ist weit effizienter, als die Determinanten sämtlicher Streichungsmatrizen zu berechnen.

D) Das Minimalpolynom

Satz 33.9 (Das Minimalpolynom)

Es sei $f \in \text{End}_K(V)$ und $A \in \text{Mat}_n(K)$.

- a. Es gibt ein eindeutiges normiertes Polynom $0 \neq \mu_f \in K[t]$ mit $\mu_f(f) = 0$, so daß

$$\mu_f \cdot K[t] = \{p \in K[t] \mid p(f) = 0\}.$$

Insbesondere ist μ_f das Nicht-Null-Polynom kleinsten Grades mit $\mu_f(f) = 0$. Wir nennen μ_f das Minimalpolynom von f .

- b. *Es gibt ein eindeutiges normiertes Polynom $0 \neq \mu_A \in K[t]$ mit $\mu_A(A) = 0$, so daß*

$$\mu_A \cdot K[t] = \{p \in K[t] \mid p(A) = 0\}.$$

Insbesondere ist μ_A das Nicht-Null-Polynom kleinsten Grades mit $\mu_A(A) = 0$. Wir nennen μ_A das Minimalpolynom von A .

Beweis: Die Aussage folgt unmittelbar aus Proposition 31.11, da $\text{End}_K(V)$ sowie $\text{Mat}_n(K)$ beides K -Algebren sind. Man beachte dabei, daß das Minimalpolynom nicht Null ist, da nach dem Satz von Cayley-Hamilton der Kern des Einsetzhomomorphismus nicht Null ist. \square

Bemerkung 33.10 (Minimalpolynome konjugierter Matrizen)

- a. Sei $f \in \text{End}_K(V)$, B eine Basis von V und $p \in K[t]$, dann gilt

$$M_B^B(p(f)) = p(M_B^B(f)),$$

da M_B^B ein K -Algebrenhomomorphismus ist.

Insbesondere gilt daher $p(f) = 0$ genau dann, wenn $p(M_B^B(f)) = 0$, und deshalb

$$\mu_f = \mu_{M_B^B(f)}.$$

Entsprechend gilt dann auch $\mu_{f_A} = \mu_{M_E^E(f_A)} = \mu_A$, wobei E die kanonische Basis von K^n bezeichnet.

- b. Konjugierte Matrizen haben dasselbe Minimalpolynom, denn wegen Aufgabe 32.33 und Satz 33.9 gilt für konjugierte Matrizen $A, B \in \text{Mat}_n(K)$

$$\mu_A \cdot K[t] = \{p \in K[t] \mid p(A) = 0\} = \{p \in K[t] \mid p(B) = 0\} = \mu_B \cdot K[t].$$

Korollar 33.11 (Primfaktorzerlegung von χ und μ)

Es sei $f \in \text{End}_K(V)$ und $A \in \text{Mat}_n(K)$.

- a. *Hat das charakteristische Polynom von f die Primfaktorzerlegung*

$$\chi_f = p_1^{n_1} \cdot \dots \cdot p_r^{n_r},$$

so hat das Minimalpolynom von f eine Primfaktorzerlegung der Form

$$\mu_f = p_1^{m_1} \cdot \dots \cdot p_r^{m_r}$$

mit $1 \leq m_i \leq n_i$ für $i = 1, \dots, r$.

- b. *Hat das charakteristische Polynom von A die Primfaktorzerlegung*

$$\chi_A = p_1^{n_1} \cdot \dots \cdot p_r^{n_r},$$

so hat das Minimalpolynom von A eine Primfaktorzerlegung der Form

$$\mu_A = p_1^{m_1} \cdot \dots \cdot p_r^{m_r}$$

mit $1 \leq m_i \leq n_i$ für $i = 1, \dots, r$.

Beweis: Der Satz von Cayley-Hamilton besagt, daß $\chi_A \in \text{Ker}(\phi_A) = \mu_A \mathbb{K}[t]$ gilt. Also gibt es ein $h \in \mathbb{K}[t]$ mit $p_1^{n_1} \cdot \dots \cdot p_r^{n_r} = \chi_A = \mu_A \cdot h$.

Daraus folgt $\mu_A = p_1^{m_1} \cdot \dots \cdot p_r^{m_r}$ für $0 \leq m_i \leq n_i$, $i = 1, \dots, r$, geeignet. Wir müssen zeigen, daß jedes p_i in μ_A auch vorkommt, d. h., daß $m_i \geq 1$ für alle $i = 1, \dots, r$.

Nehmen wir an, daß es ein i mit $m_i = 0$ gibt. Dann sind μ_A und p_i teilerfremde Polynome, also gibt es wegen der Bézout-Identität 31.20 Polynome $p, q \in \mathbb{K}[t]$ mit

$$1 = p \cdot \mu_A + q \cdot p_i.$$

Wir führen die Annahme zum Widerspruch, indem wir zum algebraischen Abschluß $\bar{\mathbb{K}}$ von \mathbb{K} übergehen. Da p_i vom Grad $\deg(p_i) \geq 1$ ist, besitzt es eine Nullstelle $\lambda \in \bar{\mathbb{K}}$. Setzen wir λ in der obigen Gleichung für t ein, so erhalten wir

$$1 = p(\lambda) \cdot \mu_A(\lambda) + q(\lambda) \cdot p_i(\lambda) = p(\lambda) \cdot \mu_A(\lambda).$$

Also muß $\mu_A(\lambda) \neq 0$ gelten.

Wegen $p_i(\lambda) = 0$, ist dann aber auch $\chi_A(\lambda) = 0$ und somit ist λ ein Eigenwert von $A \in \text{Mat}_n(\bar{\mathbb{K}})$. Sei nun $0 \neq x \in \bar{\mathbb{K}}^n$ ein Eigenvektor von A zum Eigenwert λ . Dann gilt für das Polynom $\mu_A = \sum_{i=0}^m a_i t^i \in \mathbb{K}[t] \subseteq \bar{\mathbb{K}}[t]$

$$\mu_A(A)x = \sum_{i=0}^m a_i (A^i x) = \sum_{i=0}^m a_i (\lambda^i x) = \mu_A(\lambda) \cdot x \neq 0,$$

im Widerspruch zu $\mu_A(A) = 0$. Also muß $m_i \geq 1$ gelten.

Die analoge Aussage für f folgt aus der entsprechenden Aussage für eine Matrixdarstellung $M_B^B(f)$. \square

Korollar 33.12 (Die Eigenwerte sind die Nullstellen des Minimalpolynoms.)

Es sei $f \in \text{End}_{\mathbb{K}}(V)$ und $A \in \text{Mat}_n(\mathbb{K})$.

- Die Eigenwerte von f sind genau die Nullstellen des Minimalpolynoms μ_f , und χ_f zerfällt genau dann über \mathbb{K} in Linearfaktoren, wenn μ_f zerfällt.
- Die Eigenwerte von A sind genau die Nullstellen des Minimalpolynoms μ_A , und χ_A zerfällt genau dann über \mathbb{K} in Linearfaktoren, wenn μ_A zerfällt.

Beweis: Genau dann ist $\lambda \in \mathbb{K}$ ein Eigenwert von f , wenn $t - \lambda$ ein Primfaktor von χ_f ist, und dies ist wegen Korollar 33.11 genau dann der Fall, wenn $t - \lambda$ ein Primfaktor von μ_f ist, d.h. wenn λ eine Nullstelle von μ_f ist. Wegen Korollar 33.11 zerfällt χ_f zudem genau dann über \mathbb{K} in Linearfaktoren, wenn μ_f in Linearfaktoren zerfällt. Die entsprechenden Aussagen für eine Matrix A folgen aus a. mit $f = f_A$. \square

Beispiel 33.13

- Ist $A = \lambda \mathbb{1}_n \in \text{Mat}_n(\mathbb{K})$ eine Diagonalmatrix mit gleichen Diagonalelementen, so gilt wegen $\lambda \mathbb{1}_n - A = 0$ offenbar $\chi_A = (t - \lambda)^n$ und $\mu_A = t - \lambda$.

b. Sei $\lambda \in \mathbf{K}$ und

$$J := J_n(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix} \in \text{Mat}_n(\mathbf{K}),$$

d. h. $J_n(\lambda)$ hat auf der Hauptdiagonalen den Wert λ und auf der oberen Nebendiagonalen Einsen stehen, ansonsten nur Nullen. Wir nennen $J_n(\lambda)$ einen *Jordanblock* (oder eine *Jordanzelle* oder ein *Jordankästchen*) der Größe n zum Eigenwert λ .

Offenbar gilt wieder

$$\chi_J = (t - \lambda)^n.$$

Nach Korollar 33.11 ist mithin $\mu_J = (t - \lambda)^m$ für ein $1 \leq m \leq n$. Dabei ist m die kleinste natürliche Zahl mit $(J - \lambda \mathbf{1}_n)^m = 0$. Nun ist aber

$$J - \lambda \mathbf{1}_n = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix} =: N$$

und man sieht mittels einer einfachen Induktion, daß $N^k \neq 0$ für $k < n$, aber $N^n = 0$ (vgl. Aufgabe 22.16). Also gilt

$$\mu_J = (t - \lambda)^n.$$

c. Ist $A = A_1 \oplus \dots \oplus A_r \in \text{Mat}_n(\mathbf{K})$ eine Diagonalblockmatrix mit $A_i \in \text{Mat}_{n_i}(\mathbf{K})$, so folgt aus der Definition des charakteristischen Polynoms unmittelbar (vgl. Aufgabe 32.32)

$$\chi_A = \prod_{i=1}^r \chi_{A_i}.$$

Die entsprechende Formel für das Minimalpolynom gilt nicht. Sei etwa $A_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \text{Mat}_2(\mathbf{K})$ und $A_2 = (1) \in \text{Mat}_1(\mathbf{K})$, dann gilt für $A = A_1 \oplus A_2$

$$\mu_A = (t - 1)^2 \neq (t - 1)^3 = \mu_{A_1} \cdot \mu_{A_2}.$$

Man kann zeigen, daß für eine Diagonalblockmatrix wie oben μ_A ein kleinstes gemeinsames Vielfaches von $\mu_{A_1}, \dots, \mu_{A_r}$ im Sinne der Vorlesung Algebraische Strukturen ist. Darauf wollen wir hier nicht näher eingehen.

Bemerkung 33.14 (Berechnung des Minimalpolynoms)

Zur praktischen Berechnung des Minimalpolynoms von $A \in \text{Mat}_n(\mathbf{K})$ kann man wie folgt vorgehen. Aufgrund des Satzes von Cayley-Hamilton wissen wir, daß die Matrizen A^0, \dots, A^n linear abhängig sind. Fassen wir die Matrix A^i als einen *langen*

Spaltenvektor in K^{n^2} auf und bezeichnen wir diesen mit x_i , dann suchen wir das minimale m , so daß x_0, \dots, x_m linear abhängig sind, und wir suchen ferner geeignete $\beta_0, \dots, \beta_{m-1}$ mit

$$x_m + \beta_{m-1}x_{m-1} + \dots + \beta_0x_0 = 0.$$

Dies ist dann gleichbedeutend damit, daß

$$t^m + \beta_{m-1}t^{m-1} + \dots + \beta_0 \in K[t]$$

das gesuchte Minimalpolynom von A ist.

Bezeichne $X = (x_0 \dots x_n) \in \text{Mat}(n^2 \times (n+1), K)$ die Matrix, deren Spalten x_0, \dots, x_n sind, dann suchen wir eine Lösung des linearen Gleichungssystems

$$X \cdot \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix} = 0 \in K^{n^2} \quad (95)$$

mit $\beta_{m+1} = \dots = \beta_n = 0$ und $\beta_m = 1$ und so, daß m minimal mit dieser Eigenschaft ist. Da (x_0, \dots, x_{m-1}) nach Definition von m linear unabhängig, (x_0, \dots, x_m) aber linear abhängig ist, bedeutet dies, daß in einer ZSF von X die Zahlen $1, \dots, m$ Pivotindizes sind, während $m+1$ kein Pivotindex mehr ist.

Berechnet man eine Basis des Lösungsraums von (95) mittels des Algorithmus 28.10, so erhalten wir den gesuchten Koeffizientenvektor β als das Negative des ersten Basisvektors, d.h. des ersten Vektors mit einer -1 auf der Diagonalen.

Dies führt zu folgendem Algorithmus zur Berechnung des Minimalpolynoms einer Matrix $A \in \text{Mat}_n(K)$.

Algorithmus 33.15 (Algorithmus zur Berechnung des Minimalpolynoms)

INPUT: $A \in \text{Mat}_n(K)$

OUTPUT: μ_A

1. **Schritt:** Falls A nicht quadratisch ist, gib 0 zurück.
2. **Schritt:** Bilde die Potenzen A^0, \dots, A^n und schreibe die Matrizen in Form von Spaltenvektoren der Länge n^2 in eine Matrix $X \in \text{Mat}(n^2 \times (n+1), K)$.
3. **Schritt:** Berechne eine Basis von $\text{Lös}(X, 0)$.
4. **Schritt:** Verwende die Negativen der Koeffizienten des ersten Basisvektors als Koeffizienten eines Polynoms und gib dieses zurück.

E) Die Hauptraumzerlegung

Für unsere weiteren Betrachtungen brauchen wir einen neuen Begriff, der auch im folgenden Abschnitt für die Jordansche Normalform von Bedeutung sein wird. Für $\lambda \in K$ haben wir aufsteigende Ketten von Unterräumen von V (vgl. Aufgabe 25.27)

$$\text{Ker}(f - \lambda \text{id}_V) \subseteq \text{Ker}((f - \lambda \text{id}_V)^2) \subseteq \text{Ker}((f - \lambda \text{id}_V)^3) \subseteq \dots \subseteq V$$

und

$$\text{Lös}(A - \lambda \mathbf{1}_n, 0) \subseteq \text{Lös}((A - \lambda \mathbf{1}_n)^2, 0) \subseteq \text{Lös}((A - \lambda \mathbf{1}_n)^3, 0) \subseteq \dots \subseteq V$$

Die Vereinigung all dieser Unterräume ist offenbar wieder ein Unterraum und führt zu folgender Definition.

Definition 33.16 (Hauptraum)

Es sei $f \in \text{End}_K(V)$, $A \in \text{Mat}_n(K)$ und $\lambda \in K$. Dann heißen

$$\text{Hau}(f, \lambda) = \bigcup_{k \in \mathbb{N}} \text{Ker}((f - \lambda \text{id}_V)^k) \quad \text{und} \quad \text{Hau}(A, \lambda) = \bigcup_{k \in \mathbb{N}} \text{Lös}((A - \lambda \mathbf{1}_n)^k, 0)$$

der *Hauptraum* oder *verallgemeinerte Eigenraum* von f bzw. A zu λ .

Lemma 33.17 (Nilpotenzindex und Hauptraum)

Es sei $\lambda \in K$ gegeben.

a. Es gibt eine natürliche Zahl $0 \leq m \leq n$ mit

$$\text{Ker}((f - \lambda \text{id}_V)^0) \subsetneq \text{Ker}((f - \lambda \text{id}_V)^1) \subsetneq \dots \subsetneq \text{Ker}((f - \lambda \text{id}_V)^m)$$

und für $k > m$

$$\text{Hau}(f, \lambda) = \text{Ker}((f - \lambda \text{id}_V)^m) = \text{Ker}((f - \lambda \text{id}_V)^k).$$

Die Zahl m heißt Nilpotenzindex von λ und erfüllt $m \leq \text{mult}(\mu_f, \lambda)$.

b. Für jedes $k \in \mathbb{N}$ ist $\text{Ker}((f - \lambda \text{id}_V)^k)$ ein f -invarianter Unterraum von V . Insbesondere sind also Eigenräume und Haupträume von f auch f -invariant.

Die entsprechenden Aussagen für eine Matrix $A \in \text{Mat}_n(K)$ gelten analog.

Beweis: Durch Betrachtung von f_A ergibt sich die Aussage für eine Matrix A unmittelbar aus der entsprechenden Aussage für Endomorphismen.

a. Aus Aufgabe 25.27 wissen wir, daß es eine natürliche Zahl $0 \leq m \leq n$ gibt mit

$$\text{Ker}((f - \lambda \text{id}_V)^0) \subsetneq \text{Ker}((f - \lambda \text{id}_V)^1) \subsetneq \dots \subsetneq \text{Ker}((f - \lambda \text{id}_V)^m)$$

und

$$\text{Ker}((f - \lambda \text{id}_V)^m) = \text{Ker}((f - \lambda \text{id}_V)^k)$$

für $k > m$. Aus der Definition des Hauptraumes folgt dann unmittelbar

$$\text{Hau}(f, \lambda) = \bigcup_{k \in \mathbb{N}} \text{Ker}((f - \lambda \text{id}_V)^k) = \text{Ker}((f - \lambda \text{id}_V)^m).$$

Es bleibt zu zeigen, daß

$$m \leq \text{mult}(\mu_f, \lambda).$$

Angenommen, $l := \text{mult}(\mu_f, \lambda) < m$. Dann gibt es ein $x \in \text{Ker}((f - \lambda \text{id}_V)^m)$ mit $y := (f - \lambda \text{id}_V)^l(x) \neq 0$. Da λ eine l -fache Nullstelle von μ_f ist, gibt es ein $h \in K[t]$ mit

$$\mu_f = h \cdot (t - \lambda)^l,$$

wobei λ keine Nullstelle von h ist. Deshalb sind h und $(t - \lambda)^{m-l}$ teilerfremd und nach der Bézout-Identität 31.20 gibt es Polynome $p, q \in K[t]$ mit

$$p \cdot (t - \lambda)^{m-l} + q \cdot h = 1.$$

Es folgt

$$(p(f) \circ (f - \lambda \operatorname{id}_V)^{m-l} + q(f) \circ h(f))(y) = \operatorname{id}_V(y) = y. \quad (96)$$

Andererseits gilt aber $(f - \lambda \operatorname{id}_V)^{m-l}(y) = (f - \lambda \operatorname{id}_V)^m(x) = 0$ sowie

$$h(f)(y) = (h(f) \circ (f - \lambda \operatorname{id}_V)^l)(x) = \mu_f(f)(x) = 0.$$

Aus (96) folgt damit $y = 0$, im Widerspruch zur Voraussetzung. Damit haben wir $l \geq m$ gezeigt.

- b. Da f mit Potenzen von f und mit der Identität vertauschbar ist, gilt für $k \in \mathbb{N}$ und $x \in \operatorname{Ker}((f - \lambda \operatorname{id}_V)^k)$

$$(f - \lambda \operatorname{id}_V)^k(f(x)) = f((f - \lambda \operatorname{id}_V)^k(x)) = f(0) = 0,$$

woraus die Behauptung folgt. □

Satz 33.18 (Hauptraumzerlegung)

Es sei $f \in \operatorname{End}_K(V)$ so, daß χ_f über K in Linearfaktoren zerfällt, d. h. es gibt paarweise verschiedene $\lambda_1, \dots, \lambda_r \in K$ und $0 < m_i \leq n_i$, so daß

$$\chi_f = (t - \lambda_1)^{n_1} \cdots (t - \lambda_r)^{n_r} \quad \text{und} \quad \mu_f = (t - \lambda_1)^{m_1} \cdots (t - \lambda_r)^{m_r}.$$

Dann gelten:

- $V = \operatorname{Hau}(f, \lambda_1) \oplus \dots \oplus \operatorname{Hau}(f, \lambda_r)$,
- $n_i = \operatorname{mult}(\chi_f, \lambda_i) = \dim_K(\operatorname{Hau}(f, \lambda_i))$ und
- $m_i = \operatorname{mult}(\mu_f, \lambda_i)$ ist der Nilpotenzindex von $f - \lambda_i \operatorname{id}_V$.

Die analogen Aussagen für eine Matrix $A \in \operatorname{Mat}_n(K)$, deren charakteristisches Polynom zerfällt, gelten analog.

Beweis: Wir setzen $V_i := \operatorname{Hau}(f, \lambda_i)$, wählen eine Basis B_i von V_i und setzen

$$q_i := \frac{\mu_f}{(t - \lambda_i)^{m_i}} = \prod_{j \neq i} (t - \lambda_j)^{m_j}.$$

Zeige: $V = V_1 + \dots + V_r$ und $B = B_1 \cup \dots \cup B_r$ erzeugt V .

Offenbar gibt es keinen Primfaktor, den alle q_1, \dots, q_r gemeinsam haben, so daß aus der Bézout-Identität 31.20 die Existenz von Polynomen $p_1, \dots, p_r \in K[t]$ folgt mit

$$p_1 q_1 + \dots + p_r q_r = 1.$$

Setzt man $Q_i := q_i p_i$, dann folgt

$$Q_1(f) + \dots + Q_r(f) = \operatorname{id}_V.$$

Sei $x \in V$ beliebig. Wegen $(f - \lambda_i \text{id}_V)^{m_i} \circ Q_i(f) = p_i(f) \circ \mu_f(f) = 0$ gilt

$$Q_i(f)(x) \in \text{Ker}(f - \lambda_i \text{id}_V)^{m_i} = V_i,$$

und somit

$$x = Q_1(f)(x) + \dots + Q_r(f)(x) \in V_1 + \dots + V_r.$$

Damit ist $V = V_1 + \dots + V_r$ gezeigt, und insbesondere ist $B = B_1 \cup \dots \cup B_r$ ein Erzeugendensystem von V , da die B_i die V_i erzeugen.

Zeige: $\dim_{\mathbb{K}}(V_i) = |B_i| \leq n_i$.

Wir wissen aus Lemma 33.17, daß V_i ein f -invarianter Unterraum ist, und aus der Definition von V_i folgt unmittelbar

$$(f_{V_i} - \lambda_i \text{id}_{V_i})^{m_i} = 0.$$

Also muß nach Satz 33.9 $(t - \lambda_i)^{m_i}$ ein Vielfaches des Minimalpolynoms von f_{V_i} sein, d.h. $\mu_{f_{V_i}} = (t - \lambda_i)^m$ für ein $0 \leq m \leq m_i$. Wegen Korollar 33.11 gilt dann aber auch

$$\chi_{f_{V_i}} = (t - \lambda_i)^k$$

und dabei muß dann $k = \dim_{\mathbb{K}}(V_i) = |B_i|$ gelten. Aus Aufgabe 32.32 wissen wir, daß χ_f ein Vielfaches von $\chi_{f_{V_i}}$ ist, und mithin muß $k \leq n_i$ gelten.

Zeige: $V = V_1 \oplus \dots \oplus V_r$ und $n_i = \dim_{\mathbb{K}}(V_i)$ (d.h. Teil a. und b.).

Da B ein Erzeugendensystem von V ist und da $|B_i| \leq n_i$ gilt, erhalten wir

$$n \leq |B| = |B_1| + \dots + |B_r| \leq n_1 + \dots + n_r = \deg(\chi_f) = n.$$

Dies zeigt, daß alle beteiligten Ungleichheitszeichen in der Tat Gleichheitszeichen waren, d.h. $\dim_{\mathbb{K}}(V_i) = |B_i| = n_i$ für $i = 1, \dots, r$. Aber als Erzeugendensystem mit genau $n = \dim_{\mathbb{K}}(V)$ Elementen ist B dann eine Basis von V und aus Aufgabe 24.34 folgt dann, daß V die direkte Summe der V_i ist.

Zeige: m_i ist der Nilpotenzindex von f (d.h. Teil c.).

Für den Nilpotenzindex m von $f - \lambda_i \text{id}_V$ gilt nach Lemma 33.17 $m_i \geq m$.

Angenommen, $m_i > m$. Sei nun $x = x_1 + \dots + x_r \in V$ beliebig mit $x_j \in V_j$ für $j = 1, \dots, r$. Dann gilt $(f - \lambda_i \text{id}_V)^m(x_i) = 0$ und für $j \neq i$ gilt $q_i(f)(x_j) = 0$, also folgt für $p := (t - \lambda_i)^m q_i \in \mathbb{K}[t]$

$$p(f)(x) = 0.$$

Also ist $p(f) = 0$ die Nullabbildung, aber wegen $0 \leq \deg(p) < \deg(\mu_f)$ ist dies ein Widerspruch zur Definition des Minimalpolynoms in Satz 33.9.

Die entsprechende Aussage für eine Matrix A läßt sich unmittelbar auf die Aussage für f_A zurückführen. □

Bemerkung 33.19 (Hauptraumzerlegung)

Man kann im Beweis von Satz 33.18 auch ohne Rückgriff auf Aufgabe 24.34 zeigen, daß die Summe $V_1 + \dots + V_r$ eine direkte Summe ist:

Mit der Notation aus dem Beweis gilt für $x \in V_j$ und $i \neq j$

$$Q_i(f)(x) = p_i(f) \circ \prod_{k \neq i, j} (f - \lambda_k \text{id}_V)^{m_k} \circ (f - \lambda_j \text{id}_V)^{m_j}(x) = 0$$

und für $x \in V_i$ gilt deshalb

$$Q_i(f)(x) = \sum_{j=1}^r Q_j(f)(x) = \text{id}_V(x) = x.$$

Sind nun also $x = x_1 + \dots + x_r = y_1 + \dots + y_r$ zwei Darstellungen des Vektors x mit $x_i, y_i \in V_i$, $i = 1, \dots, r$, dann gilt

$$0 = Q_i(f)(x - x) = Q_i(f)((x_1 - y_1) + \dots + (x_r - y_r)) = Q_i(f)(x_i - y_i) = x_i - y_i,$$

d.h. $x_i = y_i$ für alle $i = 1, \dots, r$. Die Summe ist also direkt.

Man beachte, daß wir hier $Q_i(f)^2 = Q_i(f)$ gezeigt haben, und daß damit $Q_i(f)$ die Projektion auf V_i mit Kern $\text{Ker}(Q_i(f)) = \bigoplus_{j \neq i} V_j$ ist - vgl. Aufgabe 23.49.

Aus Satz 33.18 Teil b. und c. folgt, da die Haupträume von f f -invariant sind, unmittelbar das folgende Korollar.

Korollar 33.20

Sei f wie in Satz 33.18, dann gilt

$$\chi_{f_{\text{Hau}(f, \lambda_i)}} = (t - \lambda_i)^{n_i}$$

und

$$\mu_{f_{\text{Hau}(f, \lambda_i)}} = (t - \lambda_i)^{m_i}.$$

F) Diagonalisierbarkeit

Satz 33.21 (Diagonalisierbarkeit von Endomorphismen)

Für einen Endomorphismus $f \in \text{End}_K(V)$ sind die folgenden Aussagen äquivalent:

- f ist diagonalisierbar.
- V hat eine Basis aus Eigenvektoren von f .
- Sind $\lambda_1, \dots, \lambda_r$ die paarweise verschiedenen Eigenwerte von f in K , dann gilt

$$V = \bigoplus_{i=1}^r \text{Eig}(f, \lambda_i).$$

- Das charakteristische Polynom von f zerfällt über K in Linearfaktoren und für jeden Eigenwert λ stimmen algebraische und geometrische Vielfachheit überein.
- Das Minimalpolynom von f zerfällt über K in paarweise verschiedene Linearfaktoren.

Beweis:

a. \Rightarrow c.: Ist f diagonalisierbar, dann gibt es eine Basis B von V , so daß

$$M_B^B(f) = \bigoplus_{i=1}^r \lambda_i \mathbb{1}_{n_i},$$

mit $\lambda_i \neq \lambda_j$ für $i \neq j$. Setzen wir $p = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_r) \in K[t]$, so ist

$$p(M_B^B(f)) = p(\lambda_1 \mathbb{1}_{n_1}) \oplus \dots \oplus p(\lambda_r \mathbb{1}_{n_r})$$

wegen Bemerkung 33.6 eine Diagonalmatrix, und für die Blöcke gilt

$$p(\lambda_i \mathbb{1}_{n_i}) = p(\lambda_i) \cdot \mathbb{1}_{n_i} = 0.$$

Also ist schon $p(f) = 0$ erfüllt und p ist ein Vielfaches von μ_f . Dann zerfällt μ_f aber in paarweise verschiedene Linearfaktoren.

e. \Rightarrow d.: Zerfällt μ_f über K in paarweise verschiedene Linearfaktoren

$$\mu_f = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_r),$$

so zerfällt wegen Korollar 33.12 auch χ_f in Linearfaktoren

$$\chi_f = (t - \lambda_1)^{n_1} \cdot \dots \cdot (t - \lambda_r)^{n_r}.$$

Aus dem Satz zur Hauptraumzerlegung 33.18 folgt zudem $\text{Hau}(f, \lambda_i) = \text{Eig}(f, \lambda_i)$, da der Nilpotenzindex von λ_i eins ist, und

$$\text{mult}(\chi_f, \lambda_i) = n_i = \dim_K \text{Hau}(f, \lambda_i) = \dim_K \text{Eig}(f, \lambda_i),$$

d.h. die geometrische und die algebraische Vielfachheit jedes Eigenwertes stimmen überein.

d. \Rightarrow c.: Das charakteristische Polynom habe die Primfaktorzerlegung

$$\chi_f = (t - \lambda_1)^{n_1} \cdot \dots \cdot (t - \lambda_r)^{n_r}.$$

Aus dem Satz über die Hauptraumzerlegung und der Voraussetzung folgt dann

$$\dim_K \text{Hau}(f, \lambda_i) = n_i = \text{mult}(\chi_f, \lambda_i) = \dim_K \text{Eig}(f, \lambda_i),$$

und da stets $\text{Eig}(f, \lambda_i) \subseteq \text{Hau}(f, \lambda_i)$ gilt, folgt

$$\text{Eig}(f, \lambda_i) = \text{Hau}(f, \lambda_i).$$

Aus dem Satz über die Hauptraumzerlegung folgt dann aber wiederum

$$V = \text{Hau}(f, \lambda_1) \oplus \dots \oplus \text{Hau}(f, \lambda_r) = \text{Eig}(f, \lambda_1) \oplus \dots \oplus \text{Eig}(f, \lambda_r).$$

c. \Rightarrow b.: Es sei B_i eine Basis von $\text{Eig}(f, \lambda_i)$, dann ist nach Aufgabe 24.34 $B = B_1 \cup \dots \cup B_r$ eine Basis von V , die aus Eigenvektoren besteht.

b. \Rightarrow a.: Ist $B = (x_1, \dots, x_n)$ eine Basis von V aus Eigenvektoren, so ist $f(x_i) = \lambda_i \cdot x_i$ für ein geeignetes $\lambda_i \in K$. Damit ist dann aber $M_B^B(f)$ eine Diagonalmatrix mit den Diagonaleinträgen $\lambda_1, \dots, \lambda_n$. \square

Korollar 33.22 (Diagonalisierbarkeit von Matrizen)

Für eine quadratische Matrix $A \in \text{Mat}_n(\mathbb{K})$ sind die folgenden Aussagen äquivalent:

- A ist diagonalisierbar.
- \mathbb{K}^n hat eine Basis aus Eigenvektoren von A .
- Sind $\lambda_1, \dots, \lambda_r$ die paarweise verschiedenen Eigenwerte von A , dann gilt

$$\mathbb{K}^n = \bigoplus_{i=1}^r \text{Eig}(A, \lambda_i).$$

- Das charakteristische Polynom von A zerfällt über \mathbb{K} in Linearfaktoren und für jeden Eigenwert λ stimmen algebraische und geometrische Vielfachheit überein.
- Das Minimalpolynom von A zerfällt über \mathbb{K} in paarweise verschiedene Linearfaktoren.

Insbesondere, genau dann ist $T \in \text{GL}_n(\mathbb{K})$ so, daß $T^{-1} \circ A \circ T$ eine Diagonalmatrix ist, wenn die Spalten von T eine Basis des \mathbb{K}^n aus Eigenvektoren von A sind.

Beweis: Wende Satz 33.21 auf die Abbildung f_A an. □

Falls ein Endomorphismus oder eine Matrix hinreichend viele verschiedene Eigenwerte hat, so folgt aus den obigen Überlegungen unmittelbar deren Diagonalisierbarkeit.

Korollar 33.23 (Diagonalisierbarkeit)

Es sei $f \in \text{End}_{\mathbb{K}}(V)$ und $A \in \text{Mat}_n(\mathbb{K})$.

- Hat f genau n paarweise verschiedene Eigenwerte, so ist f diagonalisierbar.
- Hat A genau n paarweise verschiedene Eigenwerte, so ist A diagonalisierbar.

Beweis: Hat f genau n paarweise verschiedene Eigenwerte $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, so muß

$$\chi_f = \mu_f = (t - \lambda_1) \cdot \dots \cdot (t - \lambda_n)$$

gelten. D.h. μ_f zerfällt in paarweise verschiedene Linearfaktoren und f ist diagonalisierbar. Der Beweis für A geht analog. □

Aus Korollar 33.22 können wir ein Verfahren ableiten, das es uns erlaubt, eine Matrix zu diagonalisieren und die Transformationsmatrix T zu berechnen.

Algorithmus 33.24 (Algorithmus zur Diagonalisierung)

INPUT: $A \in \text{Mat}_n(\mathbb{K})$.

OUTPUT: 0, falls A über \mathbb{K} nicht diagonalisierbar ist,
 1, D, T , falls A diagonalisierbar ist, wobei D eine zu A konjugierte Diagonalmatrix ist, und T die zugehörige Transformationsmatrix mit $T^{-1} \circ A \circ T = D$.

1. Schritt: Berechne das charakteristische Polynom von A .

- 2. Schritt:** Faktorisiere das charakteristische Polynom über K . Ist einer der Faktoren nicht linear, ist A nicht diagonalisierbar (nicht einmal trigonalisierbar) und man gebe 0 zurück. Sind alle Faktoren linear, so liefert die Faktorisierung die Eigenwerte $\lambda_1, \dots, \lambda_r$ sowie ihre algebraischen Vielfachheiten n_1, \dots, n_r .
- 3. Schritt:** Bestimme für jeden Eigenwert λ_i eine Basis des Eigenraums $\text{Eig}(A, \lambda_i)$ als $\text{Lös}(A - \lambda_i \mathbb{1}_n, 0)$ - vgl. Algorithmus 28.10 - sowie seine Dimension - vgl. Algorithmus 27.20 -, d. h. die geometrische Vielfachheit von λ_i .
- 4. Schritt:** Stimmt für jeden Eigenwert die algebraische Vielfachheit mit der geometrischen überein, so schreibe man die im 3. Schritt bestimmten Basen als Spalten in eine Matrix und erhält so T . Ferner erhält man D , indem man die Eigenwerte $\lambda_1, \dots, \lambda_r$ entsprechend ihren algebraischen Vielfachheiten in der Diagonalen einer Nullmatrix einträgt.

Beispiel 33.25

Gegeben sei die Matrix

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & -1 & -1 & 1 \end{pmatrix} \in \text{Mat}_4(\mathbb{Q}).$$

Das charakteristische Polynom von A berechnet man mit Hilfe zweifacher Laplace-Entwicklung nach der jeweils letzten Spalte als

$$\chi_A = \begin{vmatrix} t-2 & 1 & 0 & 0 \\ 0 & t-1 & 0 & 0 \\ 0 & 0 & t-2 & 0 \\ -1 & 1 & 1 & t-1 \end{vmatrix} = (t-1) \cdot (t-2) \cdot \begin{vmatrix} t-2 & 1 \\ 0 & t-1 \end{vmatrix} = (t-1)^2 \cdot (t-2)^2.$$

Damit ist also $\sigma(A) = \{1, 2\}$ mit $\text{mult}(\chi_A, 1) = \text{mult}(\chi_A, 2) = 2$.

Als nächstes berechnen wir den Eigenraum $\text{Lös}(2\mathbb{1}_4 - A, 0)$ zum Eigenwert $\lambda = 2$:

$$2 \cdot \mathbb{1}_4 - A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 1 \end{pmatrix} \xrightarrow[\text{IV} \rightarrow \text{IV} - \text{II}]{\begin{matrix} \text{I} \leftrightarrow \text{IV} \\ \text{I} \rightarrow -\text{I} \end{matrix}} \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{-1\text{'en einfügen}} \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Mithin ist

$$\text{Eig}(A, 2) = \text{Lin}((-1, 0, -1, 0)^t, (-1, 0, 0, -1)^t)$$

und

$$\dim_{\mathbb{Q}} \text{Eig}(A, 2) = 2 = \text{mult}(\chi_A, 2).$$

Dann berechnen wir den Eigenraum $\text{Lös}(\mathbb{1}_4 - A, 0)$ zum Eigenwert $\lambda = 1$:

$$1 \cdot \mathbb{1}_4 - A = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 1 & 1 & 0 \end{pmatrix} \xrightarrow[\text{III} \rightarrow -\text{III}]{\begin{matrix} \text{IV} \rightarrow \text{IV} - \text{I} + \text{III} \\ \text{I} \rightarrow -\text{I} \end{matrix}} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{-1\text{'en einfügen}} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Mithin ist

$$\text{Eig}(A, 1) = \text{Lin}((-1, -1, 0, 0)^t, (0, 0, 0, -1)^t)$$

und

$$\dim_{\mathbb{Q}} \text{Eig}(A, 1) = 2 = \text{mult}(\chi_A, 1).$$

Also zerfällt χ_A über \mathbb{Q} in Linearfaktoren und die geometrischen Vielfachheiten der Eigenwerte stimmen mit den algebraischen überein, so daß A diagonalisierbar ist. Zudem gilt für

$$T = \begin{pmatrix} -1 & -1 & -1 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & -1 \end{pmatrix},$$

daß

$$T^{-1} \circ A \circ T = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Aufgaben

Aufgabe 33.26

Es sei $E = (E_{11}, E_{12}, E_{21}, E_{22})$ die kanonische Basis von $V = \text{Mat}_2(\mathbb{K})$ und $T = E_{11} + E_{12} + E_{22} \in \text{Gl}_2(\mathbb{K})$. Zeige, daß der Endomorphismus $f: V \rightarrow V: A \mapsto T \circ A \circ T^{-1}$ trigonalisierbar, aber nicht diagonalisierbar ist, und bestimme eine Basis B von V , so daß $M_B^B(f)$ eine obere Dreiecksmatrix ist.

Aufgabe 33.27

Zeige, ist $A \in \text{Gl}_n(\mathbb{K})$, so gibt es ein Polynom $p \in \mathbb{K}[t]$ mit $A^{-1} = p(A)$.

Aufgabe 33.28

Zeige, ist $1 \leq \dim_{\mathbb{K}}(V) = n < \infty$, so sind für $\mathcal{A} \subseteq \text{End}_{\mathbb{K}}(V)$ die folgenden beiden Aussagen gleichwertig:

- \mathcal{A} ist simultan diagonalisierbar, d. h. es gibt eine Basis B von V , so daß für alle $f \in \mathcal{A}$ gilt $M_B^B(f)$ ist eine Diagonalmatrix.
- Für alle $f \in \mathcal{A}$ gilt, f ist diagonalisierbar, und für alle $f, g \in \mathcal{A}$ gilt, $f \circ g = g \circ f$.

Hinweis: Führe für "b. \Rightarrow a." Induktion über n und zerlege dazu V in zwei invariante Unterräume kleinerer Dimension.

§ 34 Die Jordansche Normalform

In diesem Abschnitt sei V ein K -Vektorraum mit $1 \leq \dim_K(V) = n < \infty$.

Eine Matrix $A \in \text{Mat}_n(K)$, deren charakteristisches Polynom in Linearfaktoren zerfällt, was etwa für einen algebraisch abgeschlossenen Körper wie \mathbb{C} stets der Fall ist, ist zu einer Matrix konjugiert, die besonders einfach gebaut ist, der sog. Jordanschen Normalform von A . Aus der Jordanschen Normalform lassen sich Invarianten von A einfach ablesen und diese Invarianten bestimmen die Matrix A bis auf Konjugation eindeutig.

Satz 34.1 (Jordansche Normalform eines Endomorphismus)

Es sei $f \in \text{End}_K(V)$ ein Endomorphismus, dessen charakteristisches Polynom über K zerfällt, $\chi_f = (t - \lambda_1)^{m_1} \cdot \dots \cdot (t - \lambda_r)^{m_r}$, und es sei $\mu_f = (t - \lambda_1)^{m_1} \cdot \dots \cdot (t - \lambda_r)^{m_r}$.

Dann gibt es für jedes $i = 1, \dots, r$ und $1 \leq j \leq m_i$, je eine natürliche Zahl t_{ij} und es gibt eine Basis B so, daß

- (1) $\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K \text{Hau}(f, \lambda_i)$,
- (2) $\sum_{j=1}^{m_i} t_{ij} = \dim_K \text{Eig}(f, \lambda_i)$,
- (3) $t_{im_i} \geq 1$ und

$$J_f := M_B^B(f) = \bigoplus_{i=1}^r \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i).$$

J_f heißt Jordansche Normalform von f , und die t_{ij} werden Elementarteiler von f zum Eigenwert λ_i genannt.

Korollar 34.2 (Jordansche Normalform einer quadratischen Matrix)

Es sei $A \in \text{Mat}_n(K)$ ein Endomorphismus, dessen charakteristisches Polynom über K zerfällt, $\chi_A = (t - \lambda_1)^{m_1} \cdot \dots \cdot (t - \lambda_r)^{m_r}$, und es sei $\mu_A = (t - \lambda_1)^{m_1} \cdot \dots \cdot (t - \lambda_r)^{m_r}$.

Dann gibt es für jedes $i = 1, \dots, r$ und $1 \leq j \leq m_i$, je eine natürliche Zahl t_{ij} und es gibt $T \in \text{GL}_n(K)$ so, daß

- (1) $\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K \text{Hau}(A, \lambda_i)$,
- (2) $\sum_{j=1}^{m_i} t_{ij} = \dim_K \text{Eig}(A, \lambda_i)$,
- (3) $t_{im_i} \geq 1$ und

$$J_A := T^{-1} \circ A \circ T = \bigoplus_{i=1}^r \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i).$$

J_A heißt Jordansche Normalform von A , und die t_{ij} werden Elementarteiler von A zum Eigenwert λ_i genannt.

Beweis: Der Beweis folgt aus Satz 34.1 mit $f = f_A$. \square

Es scheint angebracht, den Satz zunächst etwas zu erläutern, um ihn verständlicher zu machen.

Bemerkung 34.3 (Jordansche Normalform)

- a. Ziel des Abschnittes ist es, zu zeigen, daß eine Matrix A , deren charakteristisches Polynom zerfällt, konjugiert zu einer Matrix von besonders einfacher Gestalt ist. Der obige Satz sagt nun, daß in der Tat A konjugiert ist zu einer Diagonalblockmatrix, deren Diagonalblöcke, die $J_j(\lambda_i)$, alle Jordanblöcke sind, also obere Dreiecksmatrizen, die auf der Diagonalen stets den gleichen Wert λ_i stehen haben, auf der oberen Nebendiagonalen nur Einsen und ansonsten nur Nullen (vgl. Beispiel 33.13).

Dabei gelten:

- Die natürlichen Zahlen t_{ij} geben an, wieviele Jordanblöcke der Größe $j \times j$ zum Eigenwert λ_i denn vorkommen.
 - $j \leq m_i$ bedeutet, die maximale Größe eines Jordanblockes ist $m_i \times m_i$.
 - $t_{im_i} \geq 1$ besagt, daß auch mindestens ein Block der maximalen Größe $m_i \times m_i$ vorkommt. D. h. die Vielfachheit von λ_i als Nullstelle von μ_A gibt die maximale Größe eines vorkommenden Jordanblockes in J_A zum Eigenwert λ_i an.
 - Die Summe $\sum_{j=1}^{m_i} j \cdot t_{ij}$ gibt gerade an, wie oft der Eigenwert λ_i auf der Diagonalen der Diagonalblockmatrix vorkommt, und da diese das gleiche charakteristische Polynom wie A besitzt, muß die Summe mithin n_i , also die algebraische Vielfachheit von λ_i als Eigenwert von A , sein.
 - Und $\sum_{j=1}^{m_i} t_{ij} = \dim_{\mathbb{K}} \text{Eig}(A, \lambda_i)$ bedeutet schließlich, daß die Anzahl der Jordanblöcke zum Eigenwert λ_i , die in J_A vorkommen, der Dimension des Eigenraumes von A zum Eigenwert λ_i entspricht, d.h. seiner geometrischen Vielfachheit.
- b. Schon die direkte Summenschreibweise der Jordanschen Normalform bringt zum Ausdruck, daß die Jordansche Normalform nur bis auf die Reihenfolge der Jordanblöcke eindeutig bestimmt sein kann, und in der Tat ist sie es auch, d. h.:

Zwei Jordansche Normalformen sind genau dann konjugiert, wenn ihre Eigenwerte und die zugehörigen Elementarteiler übereinstimmen.

Es ist leicht einsichtig, daß eine Vertauschung der Blöcke durch Konjugation mit einer Reihe von Permutationsmatrizen erreicht werden kann, daß mithin zwei Jordansche Normalformen, deren Eigenwerte mit zugehörigen Elementarteilern übereinstimmen, zueinander konjugiert sind.

Seien umgekehrt zwei Jordansche Normalformen zueinander konjugiert, dann stimmen zunächst die charakteristischen Polynome und damit die Eigenwerte

überein. Ferner folgt aus Aufgabe 34.16, daß die Elementarteiler übereinstimmen, da für eine invertierbare Matrix $T \in \text{GL}_n(\mathbb{K})$ und ein $k \in \mathbb{N}$ gilt

$$\begin{aligned} \text{rang} \left((T^{-1} \circ A \circ T - \lambda \mathbb{1}_n)^k \right) &= \text{rang} \left(T^{-1} \circ (A - \lambda \mathbb{1}_n)^k \circ T \right) \\ &= \text{rang} \left((A - \lambda \mathbb{1}_n)^k \right). \end{aligned}$$

Damit ist natürlich auch die Jordansche Normalform eines Endomorphismus bis auf die Reihenfolge der Jordanblöcke eindeutig bestimmt.

- c. Wir wollen folgende Notation einführen, die die Jordanblöcke von A (bzw. f) zu einem Eigenwert λ_i zusammenfaßt:

$$J_A(\lambda_i) := \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i) \quad \text{bzw.} \quad J_f(\lambda_i) := \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i)$$

Dann gilt

$$J_A = \bigoplus_{i=1}^r J_A(\lambda_i) \quad \text{bzw.} \quad J_f = \bigoplus_{i=1}^r J_f(\lambda_i).$$

Beispiel 34.4 (Jordansche Normalform)

Wir wollen nun in einigen einfachen Fällen die Jordansche Normalform bestimmen.

- a. Die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \text{Mat}_4(\mathbb{Q})$$

ist eine obere Dreiecksmatrix und ihr charakteristisches Polynom

$$\chi_A = (t - 1) \cdot (t - 5) \cdot (t - 8) \cdot (t - 2)$$

zerfällt in paarweise verschiedene Linearfaktoren. Da zu jedem der Eigenwerte ein Jordanblock gehören muß und da die Matrix J_A nicht mehr als vier Jordanblöcke aufnehmen kann, gilt also

$$J_A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 5 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}.$$

Über die Transformationsmatrix T ist damit noch nichts gesagt. Da die Matrix J_A aber eine Diagonalmatrix ist, wissen wir aus Korollar 33.22 bereits, daß die Spalten von T Eigenvektoren zu den vier Eigenwerten sein müssen. Wir könnten T also leicht berechnen.

b. Die Matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \in \text{Mat}_4(\mathbb{Q})$$

hat offenbar den Rang eins. Deshalb gilt für die geometrische Vielfachheit von 0 als Eigenwert von A

$$\dim_{\mathbb{K}} \text{Eig}(A, 0) = \dim_{\mathbb{K}} \text{Lös}(A, 0) = 4 - \text{rang}(A) = 3. \quad (97)$$

Da die algebraische Vielfachheit von 0 als Eigenwert von A mindestens so groß sein muß wie die geometrische, besitzt im charakteristischen Polynom χ_A von A der Linearfaktor t also mindestens Vielfachheit 3. Deshalb gibt es ein $\lambda \in \mathbb{Q}$ mit

$$\chi_A = t^3 \cdot (t - \lambda) = t^4 - \lambda \cdot t^3.$$

Aus Lemma 32.13 wissen wir aber, daß der zweithöchste Koeffizient des charakteristischen Polynoms das Negative der Spur der Matrix ist, d.h. $\lambda = \text{Spur}(A) = 4$. Wir haben also

$$\chi_A = t^3 \cdot (t - 4).$$

Aus (97) folgt, daß es drei Jordanblöcke zum Eigenwert 0 geben muß, und außerdem muß es einen Jordanblock zum Eigenwert 4 geben. Da aber wieder höchstens vier Jordanblöcke in J_A passen, gilt

$$J_A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

Die Transformationsmatrix T enthält als Spalten also auch wieder Eigenvektoren und läßt sich so leicht berechnen.

c. Wir betrachten wieder die Matrix A aus dem vorherigen Teil, fassen sie nun aber als Matrix über dem Körper \mathbb{F}_2 auf, d.h.

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \in \text{Mat}_4(\mathbb{F}_2).$$

Wie oben sieht man, daß die Matrix Rang eins hat und somit 0 die geometrische Vielfachheit 3 besitzt. Und mit den gleichen Argumenten erhalten wir

$$\chi_A = t^3 \cdot (t - \text{Spur}(A)).$$

Allerdings ist die Spur diesmal

$$\text{Spur}(A) = 1 + 1 + 1 + 1 = 0 \in \mathbb{F}_2,$$

so daß wir

$$\chi_A = t^4$$

erhalten. 0 hat die geometrische Vielfachheit 3 und hat somit exakt drei Jordanblöcke zum Eigenwert 0, und da A keine anderen Eigenwerte besitzt, muß einer dieser Jordanblöcke diesmal die Größe zwei haben! Wir erhalten also

$$J_A = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right).$$

Die Matrix A ist diesmal also *nicht* diagonalisierbar und wir wissen deshalb auch noch nicht, wie wir die Transformationsmatrix T bestimmen sollten!

- d. Es sei $A \in \text{Mat}_3(\mathbb{Q})$ mit der Eigenschaft $A^3 - A^2 = 0$. Was können wir über die Jordansche Normalform von A sagen?

A ist eine Nullstelle des Polynoms

$$p = t^3 - t^2 = t^2 \cdot (t - 1).$$

Das Minimalpolynom von A muß nach Satz 33.9 ein Teiler von p sein, so daß für μ_A nur folgende Möglichkeiten in Betracht kommen:

$$\mu_A \in \{t, t - 1, t \cdot (t - 1), t^2, t^2 \cdot (t - 1)\}.$$

Daraus ergeben sich für die Jordansche Normalform bis auf die Reihenfolge der Jordanblöcke folgenden Möglichkeiten:

μ_A	t	$t - 1$	$t \cdot (t - 1)$	t^2	$t^2 \cdot (t - 1)$
J_A	$\left(\begin{array}{ccc cc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array} \right)$	$\left(\begin{array}{ccc cc} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \end{array} \right)$	$\left(\begin{array}{ccc cc} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \end{array} \right)$	$\left(\begin{array}{ccc cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array} \right)$	$\left(\begin{array}{ccc cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \end{array} \right)$
			$\left(\begin{array}{ccc cc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 1 & 0 & 0 \end{array} \right)$		

Dabei ist die Situation für $\mu_A = t$ oder $\mu_A = t - 1$ klar, da dann schon A selbst die angegebene Jordansche Normalform sein muß, wie man durch einsetzen von A in die Gleichung sieht.

Ist $\mu_A = t \cdot (t - 1)$, so zerfällt das Minimalpolynom in paarweise verschiedene Linearfaktoren und A ist nach Korollar 33.22 diagonalisierbar. Zudem muß für jeden Eigenwert mindestens ein Jordanblock vorkommen, so daß genau die beiden angegebenen Matrizen in Frage kommen.

Wenn $\mu_A = t^2$ ist, so muß ein Jordanblock der Größe zwei zum Eigenwert 0 vorkommen und da nur Blöcke zum Eigenwert 0 vorkommen können, sind wir dann auch schon fertig. $\mu_A = t^2 \cdot (t - 1)$ geht analog.

Definition 34.5 (Nilpotent)

Wir nennen einen Endomorphismus $f \in \text{End}_K(V)$ bzw. eine Matrix $A \in \text{Mat}_n(K)$ *nilpotent*, wenn es ein $r \in \mathbb{N}$ gibt, so daß $f^r = 0$ bzw. $A^r = 0$. Offenbar gilt dann $\mu_f = \mathfrak{t}^m$ bzw. $\mu_A = \mathfrak{t}^m$ für ein $1 \leq m \leq r$.

Wir werden Satz 34.1 zunächst für nilpotente Endomorphismen zeigen, d. h. für Endomorphismen, die nur einen Eigenwert, nämlich $\lambda = 0$, besitzen, und den allgemeinen Fall dann auf diesen zurückführen.

Lemma 34.6 (Jordansche Normalform nilpotenter Endomorphismen)

Es sei $f \in \text{End}_K(V)$ ein nilpotenter Endomorphismus mit $\mu_f = \mathfrak{t}^m$. Dann gibt es für jedes $1 \leq j \leq m$ je eine natürliche Zahl t_j und es gibt eine Basis B so, daß

- (1) $\sum_{j=1}^m j \cdot t_j = n = \dim_K \text{Hau}(f, 0) = \dim_K(V)$,
- (2) $\sum_{j=1}^m t_j = \dim_K \text{Eig}(f, 0)$,
- (3) $t_m \geq 1$ und

$$J_f := M_B^B(f) = \bigoplus_{j=1}^m \bigoplus_{k=1}^{t_j} J_j(0).$$

Bemerkung 34.7 (Jordanbasis eines nilpotenten Endomorphismus)

Bevor wir uns dem Beweis der Aussage zuwenden, wollen wir etwas besser verstehen, welche Eigenschaften die Basis B von V haben muß, damit $M_B^B(f)$ eine Diagonalblockmatrix aus Jordanblöcken ist, wie in Lemma 34.6 angegeben. Wir schauen uns dazu folgende exemplarische Situation an.

Es sei $f : V \rightarrow V$ ein nilpotenter Endomorphismus des achtdimensionalen Vektorraums V mit einer Basis $B = (x_1, \dots, x_8)$, so daß

$$M_B^B(f) = \left(\begin{array}{cccc|cccc} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) = J_4(0) \oplus J_2(0) \oplus J_2(0) \in \text{Mat}_8(K)$$

Da $M_B^B(f)$ eine Diagonalblockmatrix ist und sich beim Potenzieren die Blöcke auf der Diagonalen potenzieren (siehe Bemerkung 33.6) erhalten wir aus Aufgabe 32.34, daß der Eigenwert 0 von f den Nilpotenzindex $m = 4$ hat, d.h.

$$\mu_f = \mathfrak{t}^4.$$

Außerdem gilt mit der Notation von Lemma 34.6

$$t_1 = 0, \quad t_2 = 2, \quad t_3 = 0, \quad t_m = t_4 = 1 \geq 1$$

und damit

$$1 \cdot t_1 + 2 \cdot t_2 + 3 \cdot t_3 + 4 \cdot t_4 = 0 + 4 + 0 + 4 = 8 = n.$$

Wir setzen nun $U_j := \text{Ker}(f^j)$ für $j \in \mathbb{N}$ und erhalten somit

$$\{0\} = U_0 \subsetneq U_1 = \text{Eig}(f, 0) \subsetneq U_2 \subsetneq U_3 \subsetneq U_4 = \text{Hau}(f, 0) = V.$$

Dann betrachten wir zunächst den ersten Jordanblock, der zu den Basisvektoren x_1, x_2, x_3, x_4 gehört. Aus der Matrixdarstellung lesen wir sofort ab, daß

$$x_3 = f(x_4), \quad x_2 = f(x_3) = f^2(x_4) \quad \text{und} \quad x_1 = f(x_2) = f^3(x_4).$$

Da zudem $f^4(x_4) = 0$ gilt, ist damit

$$\text{Lin}(x_1, x_2, x_3, x_4) = \text{Lin}(f^3(x_4), f^2(x_4), f(x_4), x_4)$$

ein *zyklischer Unterraum* von V im Sinne von Aufgabe 26.35. Analog sind auch

$$\text{Lin}(x_5, x_6) = \text{Lin}(f(x_6), x_6) \quad \text{und} \quad \text{Lin}(x_7, x_8) = \text{Lin}(f(x_8), x_8).$$

zyklische Unterräume, und wir stellen fest, daß es reicht, die Vektoren

$$x_4, \quad x_6 \quad \text{und} \quad x_8$$

zu kennen, um B und damit $M_B^B(f)$ zu bestimmen! Was zeichnet diese Vektoren aus?

Wir stellen zunächst fest, daß

$$\begin{aligned} x_4 &\in U_4 \setminus U_3, & x_3 &\in U_3 \setminus U_2, & x_2 &\in U_2 \setminus U_1, & x_1 &\in U_1 \setminus U_0, \\ x_5 &\in U_1 \setminus U_0, & x_6 &\in U_2 \setminus U_1, & x_7 &\in U_1 \setminus U_0, & x_8 &\in U_2 \setminus U_1. \end{aligned}$$

Daraus lassen sich leicht Basen für die Unterräume U_j ablesen:

j	U_j	Basis von U_j	$\dim_{\mathbb{K}}(U_j)$
0	$\{0\}$	$B_0 = \emptyset$	0
1	$\text{Eig}(f, 0)$	$B_1 = (x_1, x_5, x_7)$	3
2	$\text{Ker}(f^2)$	$B_2 = B_1 \cup (x_2, x_6, x_8)$	6
3	$\text{Ker}(f^3)$	$B_3 = B_2 \cup (x_3)$	7
4	V	$B_4 = B_3 \cup (x_4) = B$	8

Wir betrachten nun für $j \geq 1$ die Räume

$$\bar{U}_j = U_j / (U_{j-1} + f(U_{j+1})).$$

Dann erhalten wir

j	Basis von $f(U_{j+1})$	Basis von U_{j-1}	Basis von U_j	Basis von \bar{U}_j	$\dim_{\mathbb{K}} \bar{U}_j$	t_j
4	$(x_1, x_2, x_3, x_5, x_7)$	$(x_1, x_2, x_3, x_5, \dots, x_8)$	(x_1, \dots, x_8)	(\bar{x}_4)	1	1
3	$(x_1, x_2, x_3, x_5, x_7)$	$(x_1, x_2, x_5, \dots, x_8)$	$(x_1, x_2, x_3, x_5, \dots, x_8)$	\emptyset	0	0
2	(x_1, x_2, x_5, x_7)	(x_1, x_5, x_7)	$(x_1, x_2, x_5, \dots, x_8)$	(\bar{x}_6, \bar{x}_8)	2	2
1	(x_1, x_5, x_7)	\emptyset	(x_1, x_5, x_7)	\emptyset	0	0

Die Vektoren x_4 , x_6 und x_8 tauchen also als Repräsentanten von Basen der Unterräume \bar{U}_j auf. Dabei taucht ein Vektor genau dann als Basisvektor von \bar{U}_j auf, wenn der zugehörige Jordanblock die Größe j hat. Insbesondere ist also die Anzahl t_j der Jordanblöcke der Größe j genau die Dimension von \bar{U}_j .

Wieso sollten denn die Basen der \bar{U}_j gerade das richtige liefern? Wir haben oben gesehen, daß der Teil einer Jordanbasis zu einem Block der Größe j die Form $f^{j-1}(x), f^{j-2}(x), \dots, f(x), x$ für einen Vektor x hat, der in $U_j \setminus U_{j-1}$ liegt. Zudem sollte der Vektor x nicht schon in $f(U_{j+1})$ liegen, da er sonst bereits in einem größeren Jordanblock in der Basis vorkommen könnte — in unserem Beispiel sind $x_2, x_6, x_8 \in U_2$, aber $x_2 = f(x_3) = f^2(x_4) \in f(U_3)$ und ist damit schon im ersten Jordanblock verbraucht. Insgesamt legt dies nahe, daß die Restklasse von x in $U_j / (U_{j-1} + f(U_{j+1}))$ nicht-null sein sollte, und somit, daß wir eine Basis von \bar{U}_j betrachten sollten.

Aus der gewonnenen Erkenntnis wollen wir nun einen allgemeinen Beweis von Lemma 34.6 ableiten.

Beweis von Lemma 34.6: Wir beachten zunächst, daß nach dem Satz zur Hauptraumzerlegung 33.18

$$V = \text{Hau}(f, 0) = \text{Ker}(f^m)$$

gilt, und mithin insbesondere $f^m = 0$.

Wir führen den Beweis konstruktiv, so daß sich daraus auch ein Algorithmus zur Bestimmung der Basis B herleiten läßt. Dabei gehen wir schrittweise vor und konstruieren in jedem Schritt alle Jordanblöcke gleicher Größe, beginnend mit dem größten.

Für $j = 0, \dots, m+1$ definieren wir $U_j := \text{Ker}(f^j)$. Da m der Nilpotenzindex ist, gilt mithin

$$0 = U_0 \subsetneq U_1 \subsetneq \dots \subsetneq U_m = U_{m+1} = V.$$

Außerdem gilt $f(U_{j+1}) \subseteq U_j$ für $j = 0, \dots, m$, was folgende Definition ermöglicht:

$$\bar{U}_j := U_j / (U_{j-1} + f(U_{j+1})).$$

Schließlich setzen wir für $j = 1, \dots, m$

$$t_j := \dim_{\mathbb{K}}(\bar{U}_j).$$

Unter Berücksichtigung des Austauschsatzes von Steinitz wählen wir rekursiv für $j = m, \dots, 1$ Vektoren $x_{j,1}, \dots, x_{j,t_j} \in U_j$, deren Restklassen eine Basis von \bar{U}_j bilden.

Sodann bilden wir für $j = 1, \dots, m$, $k = 1, \dots, t_j$ und $r \in \mathbb{N}$ die Familien von Vektoren

$$B_{j,k}^r = (f^{j-1}(x_{j,k}), f^{j-2}(x_{j,k}), \dots, f^{r+1}(x_{j,k}), f^r(x_{j,k})).$$

Man beachte, daß

$$\mathbf{B}_{jk}^{r+1} \subseteq \mathbf{B}_{jk}^r \quad \text{und} \quad \text{Lin}(\mathbf{B}_{j,k}^r) \subseteq \mathbf{U}_{j-r}.$$

Wir zeigen zunächst die folgenden Aussagen:

- (a) $\mathbf{B} = \bigcup_{j=1}^m \bigcup_{k=1}^{t_j} \mathbf{B}_{j,k}^0$ ist linear unabhängig.
- (b) $\mathbf{U}_j = \mathbf{U}_{j-1} + \text{Lin}(f^l(x_{j+l,k}) \mid l = 0, \dots, m-j; k = 1, \dots, t_{j+l})$ für $j = 1, \dots, m+1$.
- (c) $\mathbf{B}_j = \left(\bigcup_{l=1}^j \bigcup_{k=1}^{t_l} \mathbf{B}_{l,k}^0 \right) \cup \left(\bigcup_{l=j+1}^m \bigcup_{k=1}^{t_l} \mathbf{B}_{l,k}^{l-j} \right)$ ist Basis von \mathbf{U}_j für $j = 0, \dots, m$.

Zu (a): Dazu zeigen wir zunächst eine Hilfsaussage. Sei $1 \leq d \leq m$ beliebig aber fest gegeben und es gelte

$$\sum_{j=d}^m \sum_{k=1}^{t_j} \lambda_{j,k,j-d} f^{j-1}(x_{j,k}) = 0. \quad (98)$$

Dann wollen wir mit Induktion nach j zeigen, daß für $j = d, \dots, m$ gilt

$$\lambda_{j,k,j-d} = 0 \quad \forall k = 1, \dots, t_j. \quad (99)$$

Da der Induktionsanfang und der Induktionsschritt exakt die gleichen Argumente verwenden, beschränken wir uns auf den Induktionsschritt. Dazu nehmen wir an, daß wir schon

$$\lambda_{j,k,j-d} = 0 \quad \forall j = d, \dots, e-1, \quad k = 1, \dots, t_j \quad (100)$$

gezeigt haben, und wir wollen dies nun auch für $j = e$ tun. Wegen (100) hat (98) die Form

$$0 = \sum_{j=e}^m \sum_{k=1}^{t_j} \lambda_{j,k,j-d} f^{j-1}(x_{j,k}) = f^{e-1} \left(\sum_{j=e}^m \sum_{k=1}^{t_j} \lambda_{j,k,j-d} f^{j-e}(x_{j,k}) \right).$$

Dann gilt aber

$$\sum_{k=1}^{t_e} \lambda_{e,k,e-d} x_{e,k} + f \left(\sum_{j=e+1}^m \sum_{k=1}^{t_j} \lambda_{j,k,j-d} f^{j-e-1}(x_{j,k}) \right) = \sum_{j=e}^m \sum_{k=1}^{t_j} \lambda_{j,k,j-d} f^{j-e}(x_{j,k}) \in \mathbf{U}_{e-1}.$$

Nun ist aber $x_{j,k} \in \mathbf{U}_j = \text{Ker}(f^j)$ und somit $f^{j-e-1}(x_{j,k}) \in \text{Ker}(f^{e+1}) = \mathbf{U}_{e+1}$ für alle $j = e+1, \dots, m$ und alle $k = 1, \dots, t_j$. Damit erhalten wir

$$\sum_{k=1}^{t_e} \lambda_{e,k,e-d} x_{e,k} \in \mathbf{U}_{e-1} + f(\mathbf{U}_{e+1}),$$

aber die $x_{e,k}$, $k = 1, \dots, t_e$, sind nach Definition linear unabhängig modulo $\mathbf{U}_{e-1} + f(\mathbf{U}_{e+1})$ und es folgt

$$\lambda_{e,k,e-d} = 0 \quad \forall k = 1, \dots, t_e.$$

Damit ist der Induktionsschritt gezeigt, und den Induktionsanfang erhält man, indem man im obigen Argument überall e durch d ersetzt. Also haben wir die Hilfsaussage bewiesen.

Nun wollen wir zeigen, daß B linear unabhängig ist. Dazu betrachten wir eine Linearkombination der Elemente von B

$$\sum_{j=1}^m \sum_{k=1}^{t_j} \sum_{l=0}^{j-1} \lambda_{j,k,l} f^l(x_{j,k}) = 0 \quad (101)$$

und müssen zeigen, daß die Koeffizienten alle null sind, d.h.

$$\lambda_{j,k,l} = 0 \quad \forall j = 1, \dots, m, \quad k = 1, \dots, t_j, \quad l = 0, \dots, j-1.$$

Das wollen wir durch Induktion nach $p = m - j + l$ tun, und da der Induktionsanfang die gleichen Argumente verwendet, wie der Induktionsschritt, werden wir wieder nur den Induktionsschritt ausführen. Dazu nehmen wir an, daß wir für alle $1 \leq j \leq m$ und $0 \leq l \leq j-1$ mit $m - j + l \in \{0, \dots, p-1\}$ bereits gezeigt haben, daß

$$\lambda_{j,k,l} = 0 \quad \forall k = 1, \dots, t_j,$$

und wir müssen nur zeigen, daß dies auch für alle j und l mit

$$m - j + l = p, \quad \text{d.h.} \quad l = j + p - m,$$

gilt. Man beachte, daß dann notwendigerweise $m - p \leq j \leq m$ gilt, da $l \geq 0$. Wenden wir nun f^{m-p-1} auf die Gleichung (101) an, so erhalten wir

$$0 = f^{m-p-1}(0) = \sum_{j=1}^m \sum_{k=1}^{t_j} \sum_{l=0}^{j-1} \lambda_{j,k,l} f^{m-p-1+l}(x_{j,k}). \quad (102)$$

Ist $l \leq j + p - m - 1$, so ist $m - j + l \leq p - 1$ und somit ist $\lambda_{j,k,l} = 0$.

Ist $l \geq j + p - m + 1$, so ist $m - p - 1 + l \geq j$ und somit ist $f^{m-p-1+l}(x_{j,k}) = 0$. Damit bleiben in (102) nur Summanden mit $l = j + p - m \geq 0$ übrig, für die notwendigerweise $j \geq m - p$ gelten muß. (102) hat deshalb die Form

$$\sum_{j=m-p}^m \sum_{k=1}^{t_j} \lambda_{j,k,j+p-m} f^{j-1}(x_{j,k}) = 0. \quad (103)$$

Aus der Hilfsaussage in (98) und (99) folgt dann mit $d = m - p$, daß

$$\lambda_{j,k,j+p-m} = 0 \quad \forall j = p - m, \dots, m, \quad k = 1, \dots, t_j.$$

Damit ist der Induktionsschritt gezeigt. Den Induktionsanfang zeigt man analog, indem man überall $p = 0$ setzt, was $j = m$ und $l = 0$ zur Folge hat und weshalb oben der Fall $l \leq j + p - m - 1$ nicht auftritt. Insgesamt haben wir mittels Induktion gezeigt, daß B linear unabhängig ist.

Zu (b): Wir führen den Beweis durch absteigende Induktion nach j . Wegen $U_{m+1} = V = U_m$ ist für $j = m + 1$ nichts zu zeigen. Sei also $1 \leq j \leq m$.

Nach Definition der $x_{j,k}$ und Induktionsvoraussetzung gilt unter Berücksichtigung von $f(U_j) \subseteq U_{j-1}$ und der Linearität von f

$$\begin{aligned}
U_j &= \text{Lin}(x_{j,k} \mid k = 1, \dots, t_j) + U_{j-1} + f(U_{j+1}) \\
&\stackrel{\text{Ind.}}{=} \text{Lin}(x_{j,k} \mid k = 1, \dots, t_j) + U_{j-1} + f(U_j) \\
&\quad + \text{Lin}(f^{l+1}(x_{j+1+l,k}) \mid l = 0, \dots, m - (j+1); k = 1, \dots, t_{j+1+l}) \\
&= \text{Lin}(x_{j,k} \mid k = 1, \dots, t_j) + U_{j-1} + \\
&\quad + \text{Lin}(f^l(x_{j+l,k}) \mid l = 1, \dots, m - j; k = 1, \dots, t_{j+l}) \\
&= U_{j-1} + \text{Lin}(f^l(x_{j+l,k}) \mid l = 0, \dots, m - j; k = 1, \dots, t_{j+l}).
\end{aligned} \tag{104}$$

Zu (c): Nach (a) ist $B_j \subseteq B$ linear unabhängig für $j = 0, \dots, m$. Es reicht also, zu zeigen, daß $U_j = \text{Lin}(B_j)$. Dies tun wir durch Induktion nach j . Dabei ist für den Induktionsanfang mit $j = 0$ nichts zu zeigen, da $U_0 = \{0\} = \text{Lin}(\emptyset)$ gilt.

Sei also $j \geq 1$. Dann gilt nach Teil (b)

$$U_j = U_{j-1} + \text{Lin}(f^l(x_{j+l,k}) \mid l = 0, \dots, m - j; k = 1, \dots, t_{j+l})$$

Da $f^l(x_{j+l,k}) \in B_{j+l,k}^l$ für $l = 0, \dots, m - j$ und $k = 1, \dots, t_{j+l}$ gilt und wegen $B_{j,k}^1 \subseteq B_{j,k}^0$ für $k = 1, \dots, t_j$, folgt mit Induktion

$$\begin{aligned}
U_j &\subseteq \text{Lin} \left(\left(\bigcup_{l=1}^{j-1} \bigcup_{k=1}^{t_l} B_{l,k}^0 \right) \cup \left(\bigcup_{l=j}^m \bigcup_{k=1}^{t_l} B_{l,k}^{l-j+1} \right) \right) + \text{Lin} \left(\bigcup_{l=0}^{m-j} \bigcup_{k=1}^{t_{j+l}} B_{j+l,k}^l \right) \\
&= \text{Lin} \left(\left(\bigcup_{l=1}^{j-1} \bigcup_{k=1}^{t_l} B_{l,k}^0 \right) \cup \left(\bigcup_{l=j}^m \bigcup_{k=1}^{t_l} B_{l,k}^{l-j+1} \right) \right) + \text{Lin} \left(\bigcup_{l=j}^m \bigcup_{k=1}^{t_l} B_{l,k}^{l-j} \right) \\
&\subseteq \text{Lin} \left(\left(\bigcup_{l=1}^j \bigcup_{k=1}^{t_l} B_{l,k}^0 \right) \cup \left(\bigcup_{l=j+1}^m \bigcup_{k=1}^{t_l} B_{l,k}^{l-j} \right) \right).
\end{aligned}$$

Die umgekehrte Inklusion “ \supseteq ” ist hingegen klar.

Damit sind die Aussagen (a)-(c) bewiesen. Insbesondere folgt aus (c) damit, daß $B = B_m$ eine Basis von V ist, und somit gilt die Aussage in (1):

$$n = \dim_{\mathbb{K}}(V) = |B| = \sum_{j=1}^m j \cdot t_j.$$

Ferner ist $U_1 = \text{Ker}(f) = \text{Eig}(f, 0)$, so daß mit (c) auch

$$\dim_{\mathbb{K}} \text{Eig}(f, 0) = |B_1| = \sum_{j=1}^m \sum_{k=1}^{t_j} |B_{j,k}^{j-1}| = \sum_{j=1}^m t_j$$

folgt, wegen $|B_{j,k}^{j-1}| = 1$ für $j = 1, \dots, m$ und $k = 1, \dots, t_j$. Damit ist auch die Aussage (2) gezeigt.

Für $j = 1, \dots, m$ und $k = 1, \dots, t_j$ ist zudem $W_{j,k} = \text{Lin}(B_{j,k}^0)$ aufgrund der Definition von $B_{j,k}^0$ ein f -invarianter, zyklischer Unterraum mit

$$M_{B_{j,k}^0}^{B_{j,k}^0}(f_{W_{j,k}}) = J_j(0).$$

Wegen $B = B_m$ gilt zudem

$$V = \bigoplus_{j=1}^m \bigoplus_{k=1}^{t_j} W_{j,k}$$

und mithin

$$J_f := M_B^B(f) = \bigoplus_{j=1}^m \bigoplus_{k=1}^{t_j} J_j(0).$$

Es bleibt also nur, $t_m \geq 1$ zu zeigen. Aber wäre $t_m = 0$, so wäre $J_f^{m-1} = 0$ im Widerspruch zu $\mu_{J_f} = \mu_f = t^m$. Damit ist das Lemma bewiesen. \square

Der Beweis von Lemma 34.6 zeigt mehr, als wir in der Aussage des Lemmas festgehalten haben. Er beschreibt die Elementarteiler t_j in natürlicher Weise als Dimensionen eines Vektorraums und zeigt genau, wie wir die *Jordanbasis* B , für die $M_B^B(f)$ in Jordanscher Normalform ist, finden können. Das Ergebnis halten wir in folgendem Korollar fest. Es ist die Grundlage für die algorithmische Bestimmung der Jordanbasis in Algorithmus 34.11.

Korollar 34.8 (Jordanbasis eines nilpotenten Endomorphismus)

Es sei $f \in \text{End}_K(V)$ ein nilpotenter Endomorphismus mit $\mu_f = t^m$ und Elementarteilern t_j für $j = 1, \dots, m$. Ferner sei $U_j = \text{Ker}(f^j)$ und

$$\bar{U}_j = U_j / (U_{j-1} + f(U_{j+1})).$$

a. *Dann gilt*

$$t_j = \dim_K(\bar{U}_j)$$

b. *Wählen wir für jedes $j = 1, \dots, m$ eine Basis $\bar{x}_{j,1}, \dots, \bar{x}_{j,t_j}$ von \bar{U}_j und setzen*

$$B_{j,k} = (f^{j-1}(x_{j,k}), f^{j-2}(x_{j,k}), \dots, f(x_{j,k}), x_{j,k}),$$

dann ist $M_B^B(f)$ in Jordanscher Normalform für $B = \bigcup_{j=1}^m \bigcup_{k=1}^{t_j} B_{j,k}$.

c. *Zu jedem $B_{j,k}$ gehört ein Jordanblock der Größe j , denn $W_{j,k} = \text{Lin}(B_{j,k})$ ist f -invariant und $M_{B_{j,k}}^{B_{j,k}}(f_{W_{j,k}}) = J_j(0)$ ist der zugehörige Jordanblock.*

d. *Für die Berechnung der $B_{j,k}$ ist es zudem wichtig, zu wissen, daß*

$$U_{j-1} + f(U_{j+1}) = U_{j-1} + \text{Lin}(f^l(x_{j+l,k}) \mid l = 1, \dots, m - j; k = 1, \dots, t_{j+l}).$$

Insbesondere gilt also

$$U_{m-1} + f(U_{m+1}) = U_{m-1}$$

und

$$U_{m-2} + f(U_m) = U_{m-2} + \text{Lin}(f(x_{m,1}), \dots, f(x_{m,t_m})).$$

Beweis: Die Aussagen a.-c. wurden so bereits in Beweis von Lemma 34.6 bewiesen. Es bleibt also nur Aussage d. zu zeigen. Dazu verwenden wir die Hilfsaussage (b) im Beweis von Lemma 34.6. Ersetzen wir dort j durch $j+1$, so erhalten wir

$$\mathbf{U}_{j+1} = \mathbf{U}_j + \text{Lin} (f^l(x_{j+l+1,k}) \mid l = 0, \dots, m-j-1; k = 1, \dots, t_{j+1+l}).$$

Wenden wir auf beiden Seiten die Abbildung f an, so ergibt sich daraus

$$\begin{aligned} f(\mathbf{U}_{j+1}) &= f(\mathbf{U}_j) + \text{Lin} (f^{l+1}(x_{j+l+1,k}) \mid l = 0, \dots, m-j-1; k = 1, \dots, t_{j+1+l}) \\ &= f(\mathbf{U}_j) + \text{Lin} (f^l(x_{j+l,k}) \mid l = 1, \dots, m-j; k = 1, \dots, t_{j+l}), \end{aligned}$$

wobei wir für die letzte Gleichheit eine Indexverschiebung $l+1 \mapsto l$ vornehmen. Da $f(\mathbf{U}_j) \subseteq \mathbf{U}_{j-1}$ gilt, ergibt sich daraus wie gewünscht

$$\mathbf{U}_{j-1} + f(\mathbf{U}_{j+1}) = \mathbf{U}_{j-1} + \text{Lin} (f^l(x_{j+l,k}) \mid l = 1, \dots, m-j; k = 1, \dots, t_{j+l}).$$

□

Bemerkung 34.9 (Jordanbasis einer nilpotenten Matrix)

Ist A eine nilpotente Matrix mit $\mu_A = t^m$ und bestimmt man wie in Korollar 34.8 linear unabhängige Familien

$$B_{j,k} = (A^{j-1}(x_{j,k}), A^{j-2}(x_{j,k}), \dots, A(x_{j,k}), x_{j,k}) \subset K^n$$

in $\text{Lös}(A^j, 0)$ für $j = 1, \dots, m$ und $k = 1, \dots, t_j$, dann ist die Matrix $T \in \text{Gl}_n(K)$, deren Spalten gerade die Vektoren der Basis B von K^n sind, eine Transformationsmatrix, die A in Jordansche Normalform überführt.

Beispiel 34.10 (Jordansche Normalform einer nilpotenten Matrix)

Wir wollen nun für die folgende nilpotente Matrix

$$A = \begin{pmatrix} -2 & 0 & 0 & 1 & 0 \\ -2 & -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ -4 & 0 & -1 & 2 & 0 \\ -2 & -1 & 0 & 1 & 1 \end{pmatrix} \in \text{Mat}_5(\mathbb{Q})$$

die Jordansche Normalform sowie die Transformationsmatrix $T \in \text{Gl}_5(\mathbb{Q})$ bestimmen.

Dazu berechnen wir zunächst den Nilpotenzindex von A und merken uns die Potenzen A^k von A , da wir sie anschließend benötigen:

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

und

$$A^3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Der Nilpotenzindex von A ist also $m = 3$, und damit $\mu_A = t^3$ und

$$\text{Hau}(A, 0) = \mathbb{Q}^5.$$

Damit muß in der Jordanschen Normalform von A also ein Jordanblock $J_3(0)$ der Größe $m = 3$ vorkommen, und aufgrund der geringen Größe der Matrix A bleiben damit nur die beiden folgenden Möglichkeiten für die Jordansche Normalform übrig:

$$J_A = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \quad \text{oder} \quad J_A = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right).$$

Gemäß Korollar 34.8 reicht es im wesentlichen, Basen der Vektorräume \bar{U}_j für $j = 3, 2, 1$ zu bestimmen, um die Jordanbasis B zusammensetzen zu können.

Wir beginnen mit $j = 3$ und

$$\bar{U}_3 = \mathcal{U}_3 / (\mathcal{U}_2 + f_A(\mathcal{U}_4)) = \mathcal{U}_3 / \mathcal{U}_2 = \mathbb{Q}^5 / \mathcal{U}_2,$$

wobei wir $\mathcal{U}_3 = \mathbb{Q}^5$ und mit Korollar 34.8 auch $\mathcal{U}_2 + f(\mathcal{U}_4) = \mathcal{U}_2$ beachten. Um eine Basis von \bar{U}_3 zu finden, muß man einerseits eine Basis von \mathcal{U}_2 berechnen und diese dann zu einer Basis von \mathcal{U}_3 ergänzen, indem man sie mit Steinitz in eine Basis von \mathcal{U}_3 hineintauscht. Bestimmen wir also zunächst eine Basis B'_2 von $\mathcal{U}_2 = \text{Lös}(A^2, 0)$. Aufgrund der einfachen Form von A^2 mit Rang 1 geschieht dies durch einfaches Draufschaun — vier der Einheitsvektoren tun es offenbar:

$$B'_2 = ((1, 0, 0, 0, 0)^t, (0, 1, 0, 0, 0)^t, (0, 0, 0, 1, 0)^t, (0, 0, 0, 0, 1)^t).$$

Für $\mathcal{U}_3 = \text{Lös}(A^3, 0) = \mathbb{Q}^5$ ist es noch einfacher, eine Basis zu bestimmen, die kanonische Basis tut's:

$$B'_3 = ((1, 0, 0, 0, 0)^t, (0, 1, 0, 0, 0)^t, (0, 0, 1, 0, 0)^t, (0, 0, 0, 1, 0)^t, (0, 0, 0, 0, 1)^t).$$

Damit ist es in dem vorliegenden Beispiel auch denkbar einfach, die Basis B'_2 in die Basis B'_3 hineinzutauschen, es fehlt nämlich einfach der Vektor e_3 , und wir setzen deshalb

$$x_{3,1} = e_3 = (0, 0, 1, 0, 0)^t.$$

Damit erhalten wir $t_3 = \dim_{\mathbb{Q}}(\bar{U}_3) = 1$ und die erste Teilbasis der Jordanbasis:

$$B_{3,1} = (A^2 x_{3,1}, A x_{3,1}, x_{3,1}) = ((-1, 0, 0, -2, 0)^t, (0, -1, 0, -1, 0)^t, (0, 0, 1, 0, 0)^t).$$

Als nächstes betrachten wir $j = 2$ und

$$\bar{\mathbf{U}}_2 = \mathbf{U}_2 / (\mathbf{U}_1 + f_A(\mathbf{U}_3)).$$

Um eine Basis von $\bar{\mathbf{U}}_2$ zu bestimmen, müssen wir eine Basis von $\mathbf{U}_1 + f_A(\mathbf{U}_3)$ berechnen und diese zu einer Basis von \mathbf{U}_2 ergänzen. Dazu berechnen wir zunächst eine Basis \mathbf{B}'_1 von $\mathbf{U}_1 = \text{Lös}(A, 0)$ mit Hilfe von Algorithmus 28.10:

$$A = \begin{pmatrix} -2 & 0 & 0 & 1 & 0 \\ -2 & -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ -4 & 0 & -1 & 2 & 0 \\ -2 & -1 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{\text{Gauß}} \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{einfügen}]{-1 \text{ 'en}} \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

Die letzten beiden Spaltenvektoren bilden also eine Basis von $\text{Lös}(A, 0)$. Wir dürfen die Vektoren aber auch mit einem Skalar multiplizieren, um schönere Vektoren zu erhalten, und tun dies. Unsere Basis von $\mathbf{U}_1 = \text{Eig}(A, 0) = \text{Lös}(A, 0)$ ist dann

$$\mathbf{B}'_1 = ((1, 0, 0, 2, 0)^t, (0, 1, 0, 0, 1)^t).$$

Damit wissen wir unter anderem, daß J_A nur $\dim_{\mathbb{K}} \text{Eig}(A, 0) = 2$ Jordanblöcke enthält, d.h. die erste der oberen beiden Möglichkeiten tritt ein, und es gilt

$$t_1 = 0, \quad t_2 = 1, \quad t_3 = 1.$$

Nachdem wir eine Basis für \mathbf{U}_1 berechnet haben, erhalten wir mit Korollar 34.8

$$\mathbf{U}_1 + f_A(\mathbf{U}_3) = \mathbf{U}_1 + \text{Lin}(A\mathbf{x}_{3,1}) = \text{Lin}(\mathbf{B}'_1 \cup (A\mathbf{x}_{3,1})). \quad (105)$$

Dann ist die offenbar linear unabhängige Familie

$$\mathbf{B}''_2 = \mathbf{B}'_1 \cup (A\mathbf{e}_3) = ((1, 0, 0, 2, 0)^t, (0, 1, 0, 0, 1)^t, (0, -1, 0, -1, 0)^t),$$

aber eine Basis von $\mathbf{U}_1 + f_A(\mathbf{U}_3)$. Man kann die lineare Unabhängigkeit auch mit Algorithmus 27.26 nachprüfen.

Diese Basis müssen wir nun zu einer Basis von \mathbf{U}_2 ergänzen. Dazu können wir sie mit Hilfe des Austauschsatzes von Steinitz in die Basis \mathbf{B}'_2 von \mathbf{U}_2 hineintauschen, oder alternativ kann man auch einfach genau hinschauen. Man sieht nämlich leicht, daß der Vektor

$$\mathbf{x}_{2,1} = (0, 0, 0, 0, 1)^t$$

von den drei Vektoren in \mathbf{B}''_2 linear unabhängig ist, und somit ergänzt er \mathbf{B}''_2 zu einer Basis von \mathbf{U}_2 . Von der linearen Unabhängigkeit der vier Vektoren kann man sich auch überzeugen, indem man die Vektoren in eine Matrix schreibt und den Rang bestimmt, was schneller ist als dreimal Steinitz und trotzdem ausreicht. Wir überlassen die Rechnung dem Leser. Nachdem wir nun $\mathbf{x}_{2,1}$ bestimmt haben, erhalten wir die zweite Teilbasis der Jordanbasis:

$$\mathbf{B}_{2,1} = (A\mathbf{x}_{2,1}, \mathbf{x}_{2,1}) = ((0, 1, 0, 0, 1)^t, (0, 0, 0, 0, 1)^t).$$

Im Prinzip bliebe noch der Fall $\underline{j = 1}$ zu untersuchen, aber wegen $\dim_{\mathbb{K}}(\overline{U}_1) = t_1 = 0$ muß \overline{U}_1 der Nullraum sein und liefert somit keinen Beitrag zur Jordanbasis.

Wir haben also die Jordanbasis $B = B_{3,1} \cup B_{2,1}$ bestimmt und damit auch die Transformationsmatrix T , deren Spalten die Vektoren in B sind. Wir haben

$$T = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \in \text{Gl}_5(\mathbb{Q})$$

mit

$$T^{-1} = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 2 & 1 & 0 & -1 & 0 \\ -2 & -1 & 0 & 1 & 1 \end{pmatrix},$$

und es gilt

$$J_A = T^{-1} \circ A \circ T = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right).$$

Beweis von Satz 34.1: Nach Satz 33.18 zerfällt V in die direkte Summe der Haupträume $V_i := \text{Hau}(f, \lambda_i)$, $i = 1, \dots, r$, und diese sind nach Lemma 33.17 invariant unter f und $f - \lambda_i \text{id}_V$. Betrachten wir nun die Abbildungen

$$(f - \lambda_i \text{id}_V)|_{V_i} : V_i \rightarrow V_i$$

für $i = 1, \dots, r$, so sind diese nilpotent mit $\chi_{(f - \lambda_i \text{id}_V)|_{V_i}} = t^{n_i}$ und $\mu_{(f - \lambda_i \text{id}_V)|_{V_i}} = t^{m_i}$ (vgl. Korollar 33.20). Nach Lemma 34.6 gibt es dann aber für jedes $i = 1, \dots, r$ Basen B_i von V_i und natürliche Zahlen t_{ij} , $j = 1, \dots, m_i$, so daß gilt

- (1) $\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_{\mathbb{K}} \text{Hau}(f, \lambda_i)$,
- (2) $\sum_{j=1}^{m_i} t_{ij} = \dim_{\mathbb{K}} \text{Eig}((f - \lambda_i \text{id}_V)|_{V_i}, 0) = \dim_{\mathbb{K}} \text{Eig}(f, \lambda_i)$,
- (3) $t_{im_i} \geq 1$ und

$$\begin{aligned} M_{B_i}^{B_i}(f|_{V_i}) &= \lambda_i \mathbb{1}_{n_i} + M_{B_i}^{B_i}((f - \lambda_i \text{id}_V)|_{V_i}) \\ &= \lambda_i \mathbb{1}_{n_i} + \left(\bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(0) \right) = \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i). \end{aligned}$$

Damit folgt die Behauptung, da für $B = B_1 \cup \dots \cup B_r$ gilt

$$M_B^B(f) = \bigoplus_{i=1}^r M_{B_i}^{B_i}(f|_{V_i}).$$

□

Wie wir schon gesehen haben, ist der Beweis zur Berechnung der Jordanschen Normalform algorithmisch. Wir wollen nun den Algorithmus beschreiben, mit Hilfe dessen man die Jordansche Normalform einer Matrix A inklusive der zugehörigen Transformationsmatrix bestimmen kann.

Algorithmus 34.11 (Jordansche Normalform - I)

INPUT: $A \in \text{Mat}_n(\mathbb{Q})$ mit μ_A zerfällt in Linearfaktoren.

OUTPUT: J_A und eine Transformationsmatrix $T \in \text{Gl}_n(\mathbb{K})$ mit $T^{-1} \circ A \circ T = J_A$.

1. **Schritt:** Bestimme das Minimalpolynom μ_A von A und faktorisierere es.
2. **Schritt:** Wenn μ_A nicht in Linearfaktoren zerfällt, gebe man eine Fehlermeldung zurück, andernfalls gilt $\mu_A = \prod_{i=1}^r (t - \lambda_i)^{m_i}$.
3. **Schritt:** Für $i = 1, \dots, r$ bilde man die Matrix $B_i = A - \lambda_i \mathbb{1}_n$ und führe folgende Schritte aus:

Schritt a.: Bestimme eine Basis D_{m_i} von Lös $(B_i^{m_i}, 0)$ sowie eine Basis D_{m_i-1} von Lös $(B_i^{m_i-1}, 0)$.

Schritt b.: Tausche D_{m_i-1} mittels des Satzes von Steinitz in D_{m_i} hinein und bestimme die in D_{m_i} verbliebenen Vektoren $x_{m_i,1}^i, \dots, x_{m_i,t_{m_i}}^i$.

Schritt c.: Für $k = 1, \dots, t_{m_i}$ schreibe man $B_i^{m_i-1} x_{m_i,k}^i, \dots, B_i^0 x_{m_i,k}^i$ als Spaltenvektoren in eine Matrix T .

Schritt d.: Für $j = m_i - 1, \dots, 1$ führe man folgendes aus:

- bestimme eine Basis D_{j-1} von Lös $(B_i^{j-1}, 0)$;
- tausche $D_{j-1} \cup \{B_i^{l-j} x_{l,k}^i \mid l = j+1, \dots, m_i, k = 1, \dots, t_l\}$ mittels des Satzes von Steinitz in D_j hinein;
- bestimme die in D_j verbliebenen Vektoren $x_{j,1}^i, \dots, x_{j,t_{ij}}^i$;
- für $k = 1, \dots, t_{ij}$ schreibe die Vektoren $B_i^{j-1} x_{j,k}^i, \dots, B_i^0 x_{j,k}^i$ als Spaltenvektoren in die Matrix T .

4. **Schritt:** Gib $T^{-1} \circ A \circ T$ und T zurück.

Beispiel 34.12 (Jordansche Normalform)

Wir wollen nun die Jordansche Normalform und die Transformationsmatrix von

$$A = \begin{pmatrix} 2 & 1 & 1 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \text{Mat}_4(\mathbb{Q})$$

bestimmen.

Das charakteristische Polynom berechnet man mit Hilfe des Kästchensatzes als

$$\chi_A = \begin{vmatrix} t-2 & -1 & -1 & -2 \\ 0 & t-1 & 0 & 0 \\ 0 & -1 & t-1 & -1 \\ 0 & 0 & 0 & t-2 \end{vmatrix} = (t-2) \cdot \begin{vmatrix} t-1 & 0 \\ -1 & t-1 \end{vmatrix} \cdot (t-2) = (t-2)^2 \cdot (t-1)^2.$$

Dann berechnen wir eine Basis von $\text{Eig}(A, 1) = \text{Lös}(A - \mathbb{1}_4, 0)$:

$$A - \mathbb{1}_4 = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\text{Gauß}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{einfügen}]{-1 \cdot \text{en}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Also ist $\text{Eig}(A, 1) = \text{Lin}((1, 0, -1, 0)^t)$, und damit stimmen die geometrische und die algebraische Vielfachheit von 1 als Eigenwert von A nicht überein. Wir müssen auch noch $\text{Hau}(A, 1) = \text{Lös}((A - \mathbb{1}_4)^2, 0)$ bestimmen:

$$(A - \mathbb{1}_4)^2 = \begin{pmatrix} 1 & 2 & 1 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{\text{Gauß}} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{einfügen}]{-1 \cdot \text{en}} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Als Basis für $\text{Hau}(A, 1)$ erhalten wir also

$$B_1 = ((2, -1, 0, 0)^t, (1, 0, -1, 0)^t).$$

Der erste der beiden Vektoren ist nicht in $\text{Eig}(A, 1)$, so daß wir ihn als x_2 wählen können. Damit erhalten wir

$$x_1 = (A - \mathbb{1}_4)x_2 = (1, 0, -1, 0)^t, \quad x_2 = (2, -1, 0, 0)^t$$

als die ersten beiden Spalten von T .

Nun wenden wir uns der Berechnung von $\text{Eig}(A, 2)$ zu:

$$A - 2 \cdot \mathbb{1}_4 = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{\text{Gauß}} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{einfügen}]{-1 \cdot \text{en}} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Also ist $\text{Eig}(A, 2) = \text{Lin}((-1, 0, 0, 0)^t)$ und somit stimmen wieder die geometrische und die algebraische Vielfachheit des Eigenwertes nicht überein. Wir müssen also wieder $\text{Hau}(A, 2) = \text{Lös}((A - \mathbb{1}_4)^2, 0)$ berechnen:

$$(A - 2 \cdot \mathbb{1}_4)^2 = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{\text{Gauß}} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow[\text{einfügen}]{-1 \cdot \text{en}} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Wir erhalten als Basis für $\text{Hau}(A, 2)$ also

$$B_2((-1, 0, 0, 0)^t, (0, 0, -1, -1)^t),$$

und somit ist $x_4 = (0, 0, 1, 1)^t$ im Hauptraum, aber nicht im Eigenraum von 2. Wir erhalten deshalb

$$x_3 = (A - 2 \cdot \mathbb{1}_4)x_4 = (3, 0, 0, 0)^t, \quad x_4 = (0, 0, 1, 1)^t$$

als die Spalten 3 und 4 der Matrix T .

Insgesamt haben wir also

$$T = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \text{Gl}_4(\mathbb{Q})$$

mit

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

und für die Jordansche Normalform erhalten wir

$$T^{-1} \circ A \circ T = \left(\begin{array}{cc|cc} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 2 \end{array} \right)$$

Will man nur die Normalform von A , aber nicht die Transformationsmatrix wissen, dann reicht es, die Elementarteiler zu bestimmen, was mit Hilfe von Aufgabe 34.16 sehr viel einfacher zu bewerkstelligen ist. Dies führt auf folgenden Algorithmus zur Bestimmung der Jordanschen Normalform einer Matrix A , deren charakteristisches Polynom zerfällt.

Algorithmus 34.13 (Jordansche Normalform - II)

INPUT: $A \in \text{Mat}_n(\mathbb{Q})$ mit μ_A zerfällt in Linearfaktoren

OUTPUT: Liste mit den Eigenwerten von A und den Elementarteilern

1. **Schritt:** Bestimme das Minimalpolynom μ_A von A und faktorisiere es.
2. **Schritt:** Wenn μ_A nicht in Linearfaktoren zerfällt, gib eine Fehlermeldung zurück.
3. **Schritt:** Für jeden Eigenwert λ_i mit $\text{mult}(\mu_A, \lambda_i) = m_i$ bestimme man für $j = 0, \dots, m_i + 1$ die Zahlen $\text{rang}((A - \lambda_i \mathbb{1}_n)^j)$ und berechne daraus den Vektor der Elementarteiler $(t_{i1}, \dots, t_{im_i})$. Den Eigenwert und den Vektor der Elementarteiler speichere man als i -ten Eintrag in einer Liste nf .
4. **Schritt:** Man gebe die Liste nf zurück.

Bemerkung 34.14 (Jordanzerlegung einer Matrix)

Es sei $J = (a_{ij})$ eine Matrix in Jordanscher Normalform. $S = (s_{ij})$ bezeichne die Diagonalmatrix, die entsteht, wenn man in J alle Nicht-Diagonalelemente zu Null setzt, d. h. $s_{ii} = a_{ii}$ und $s_{ij} = 0$ für $i \neq j$. Ferner setzen wir $N = J - S$, d. h. N ist eine Matrix, die nur auf der oberen Nebendiagonalen Elemente ungleich Null besitzen kann.

Dann ist N *nilpotent*, und es gelten

$$J = S + N \quad \text{mit} \quad N \circ S = S \circ N.$$

Man nennt dies auch die *Jordan-Zerlegung* von J .

Um die Aussage einzusehen, beachte man, daß für $i = 1, \dots, r$ und $1 \leq j \leq m_i$ gilt

$$J_j(\lambda_i) = \lambda_i \mathbb{1}_j + J_j(0).$$

Damit gilt

$$S = \bigoplus_{i=1}^r \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} \lambda_i \mathbb{1}_j$$

und

$$N = \bigoplus_{i=1}^r \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(0).$$

Aber damit folgt unmittelbar

$$N \circ S = \bigoplus_{i=1}^r \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} \lambda_i J_j(0) = S \circ N.$$

Allgemeiner nennt man die Darstellung einer Matrix $A \in \text{Mat}_n(K)$ als $A = S + N$ mit N nilpotent und S diagonalisierbar (auch *halbeinfach* genannt, engl. semi-simple, daher das S) und $S \circ N = N \circ S$ eine Jordan-Zerlegung von A . Solche Zerlegungen von Objekten in einen halbeinfachen und einen nilpotenten Anteil spielen auch in anderen Bereichen der Mathematik eine Rolle - siehe etwa Lie-Algebren oder Jordan-Algebren.

Bemerkung 34.15 (Anwendungsmöglichkeit der Jordanschen Normalform)

Anwendung findet die Jordansche Normalform zum Beispiel in der Theorie der linearen Differentialgleichungen, wo ein Fundamentalsystem mit Hilfe der Exponentialabbildung einer Matrix beschrieben wird. Diese kann mit Hilfe der Jordanschen Normalform von A berechnet werden.

Aufgaben

Aufgabe 34.16 (Berechnung der Elementarteiler)

Mit den Bezeichnungen aus Satz 34.1 zeige man, für $i = 1, \dots, r$ und $1 \leq j \leq m_i$ gilt:

$$t_{ij} = \text{rang} \left((f - \lambda_i \text{id}_V)^{j-1} \right) - 2 \cdot \text{rang} \left((f - \lambda_i \text{id}_V)^j \right) + \text{rang} \left((f - \lambda_i \text{id}_V)^{j+1} \right)$$

bzw.

$$t_{ij} = \text{rang} \left((A - \lambda_i \mathbb{1}_n)^{j-1} \right) - 2 \cdot \text{rang} \left((A - \lambda_i \mathbb{1}_n)^j \right) + \text{rang} \left((A - \lambda_i \mathbb{1}_n)^{j+1} \right).$$

Hinweise: 1. Zeige, $J_j(0)^l = (\delta_{\mu+l,\nu})_{\mu,\nu=1,\dots,j}$ und $\text{rang} (J_j(0)^l) = \max\{0, j-l\}$ für $l \in \mathbb{N}$. 2. Man betrachte zunächst den Fall $r = 1$ und $\lambda_1 = 0$. 3. Den allgemeinen Fall führe man auf die Abbildungen $g_i := (f - \lambda_i \text{id}_V)_{\text{Hau}(f, \lambda_i)}$ zurück.

Aufgabe 34.17

Bestimme die Jordansche Normalform und die zugehörige Transformationsmatrix für die Matrix A

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 & 0 \\ -2 & 0 & 1 & 0 & 0 \\ 0 & -4 & 0 & 0 & -2 \end{pmatrix} \in \text{Mat}_5(\mathbb{Q}).$$

Aufgabe 34.18

Bestimme die Jordansche Normalform und die zugehörige Transformationsmatrix $T^{-1} \in \text{Gl}_4(\mathbb{Q})$ für

$$A = \begin{pmatrix} 3 & 6 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 9 & 3 & 0 \\ 0 & 12 & 0 & 3 \end{pmatrix} \in \text{Mat}_4(\mathbb{Q}).$$

Aufgabe 34.19

Es sei $A \in \text{Mat}(5, K)$ mit $\chi_A = t(t-1)^4$, $\mu_A = t(t-1)^2$ und $\text{rang}(A - \mathbb{1}_5) = 2$. Bestimme die Jordansche Normalform von A .

Aufgabe 34.20

Zeige, ist $A \in \text{Mat}_n(K)$ so, daß χ_A über K in Linearfaktoren zerfällt, so sind A und A^t konjugiert.

Aufgabe 34.21

Beweise oder widerlege die folgende Aussage für zwei Matrizen $A, B \in \text{Mat}_n(K)$:

$$A \text{ ist konjugiert zu } B \iff \chi_A = \chi_B, \mu_A = \mu_B \text{ und } \text{rang}(A) = \text{rang}(B).$$

§ 35 Lineare Algebra mit Singular

Im vorliegenden Abschnitt wollen wir zeigen, wie ein Computeralgebrasystem eingesetzt werden kann, um Rechnungen in der linearen Algebra durchzuführen. Wir verwenden hierzu das am Fachbereich entwickelte System SINGULAR. Es ist frei erhältlich für die Betriebssysteme Linux, Windows und MacOS von der Webseite:

<http://www.singular.uni-kl.de>

Auf den Linuxrechnern des Fachbereichs startet man SINGULAR einfach durch den Befehl `Singular` von einer einfachen Textkonsole aus. Man erhält dann zunächst einige Informationen zum Programm sowie einen Eingabeprompt `>`:

```

                SINGULAR                               /
A Computer Algebra System for Polynomial Computations / version 3-1-1
                                                    0<
    by: G.-M. Greuel, G. Pfister, H. Schoenemann      \ Feb 2010
FB Mathematik der Universitaet, D-67653 Kaiserslautern \
>

```

Der Eingabeprompt `>` fordert zur Eingabe von SINGULAR-Befehlen auf. Wir wollen hier nur einige kurze Anmerkungen zur allgemeinen Syntax machen und hoffen, daß sich alles weitere aus den im folgenden besprochenen Beispielen erschließt. Unsere Konvention dabei ist, daß SINGULAR-Ein- und Ausgaben im Gegensatz zu begleitenden Erläuterungen stets im `Typewriter`-Stil geschrieben werden.

- a. Jede SINGULAR-Sitzung sollte mit dem Befehl

```
ring r=0,t,lp;
```

beginnen. Dadurch wird der Polynomring $\mathbb{Q}[t]$ als Grundring festgelegt und erhält den Namen `r`. Selbst, wenn man nicht vor hat, Polynome zu verwenden, ist dies nötig, um mit den rationalen Zahlen rechnen zu können. Ersetzt man die Zahl `0` in der Definition von `r` durch eine Primzahl `p`, so verwendet man statt der rationalen Zahlen den Körper $\mathbb{Z}/p\mathbb{Z}$; ersetzt man sie durch `real` oder `complex`, so rechnet man mit reellen oder komplexen Dezimalzahlen, was aber tunlichst vermieden werden sollte, da dann Rundungsfehler auftreten können.

- b. Man kann Ergebnisse von Rechnungen sowie Eingaben auch in Variablen speichern. Ein Beispiel dafür ist die Variable `r` in Teil a., in der der Polynomring $\mathbb{Q}[t]$ abgespeichert wurde. Jede Variable in SINGULAR hat einen Namen und einen festgelegten Typen, der sagt, ob es sich um einen Ring (`ring`), ein Polynom (`poly`), ein Körperelement (`number`), eine ganze Zahl (`int`), eine Matrix (`matrix`) oder eine Liste (`list`) von Objekten handelt.
- c. Nicht alle in SINGULAR im Prinzip verfügbaren Befehle sind schon unmittelbar mit dem Programmstart geladen, viele liegen in sogenannten Bibliotheken vor.

Sie sind erst verfügbar, wenn man die entsprechende Bibliothek mit dem Befehl LIB eingebunden hat. Wie dies geschieht, werden wir in Beispielen sehen.

- d. Jede SINGULAR-Eingabe schließt mit einem Semikolon ; und dem anschließenden Drücken der Return-Taste ab. Das Semikolon fordert den SINGULAR-Interpreter dazu auf, die Eingabe zu übersetzen und auszuführen. Will man eine Eingabe über mehrere Zeilen strecken, so läßt man das Semikolon am Zeilenende weg und drückt die Return-Taste. Man erhält statt des üblichen Promptzeichens > dann einen Punkt . als Prompt. Dieser zeigt an, daß die Eingabe noch nicht beendet ist und sich über mehrere Zeilen erstreckt.
- e. In den folgenden Beispielen ist alles, was auf einen Prompt > oder . am Zeilenanfang folgt, eine Eingabe, und jede Zeile, die ohne eines dieser Zeichen beginnt, enthält SINGULAR-Ausgaben. Text, der auf // folgt, enthält Kommentare, die beim Ausführen des Kommandos nicht beachtet werden. Will man das Beispiel selbst in SINGULAR nachprüfen, kann man sie getrost weglassen. Sie dienen nur der Erläuterung für den Leser. Ausgaben, die beim Laden von Bibliotheken auftreten, werden wir in den Beispielen weglassen.
- f. Man beendet SINGULAR mit dem Befehl exit. Hilfe zur Syntax von SINGULAR findet man im Manual auf der SINGULAR-Webseite oder durch den Befehl help.

Beispiel 35.1 (Reduzierte Zeilen-Stufen-Form und Rang einer Matrix)

Wir wollen eine reduzierte Zeilen-Stufenform und damit den Rang der folgenden Matrix berechnen:

$$A = \begin{pmatrix} 1 & 2 & 1 & 3 & 1 \\ 2 & 4 & 7 & 3 & 0 \\ 4 & 8 & 9 & 9 & 2 \\ 3 & 6 & 0 & 2 & 1 \end{pmatrix} \in \text{Mat}(4 \times 5, \mathbb{Q})$$

Dazu benutzen wir die SINGULAR-Befehle rowred.

```
> LIB "matrix.lib";
> ring r=0,t,dp;
> matrix A[4][5]=1,2,1,3,1,2,4,7,3,0,4,8,9,9,2,3,6,0,2,1;
> print(rowred(A)); // zeigt die rZSF von A
1,2,0,0,1/11,
0,0,1,0,-2/11,
0,0,0,1,4/11,
0,0,0,0,0
```

Beispiel 35.2 (Kern einer Matrix)

Mit dem Befehl syz können wir eine Basis des Kerns der Matrix in Beispiel 35.1 berechnen. Man bezeichnet die Relationen zwischen den Spalten der Matrix, die durch die Vektoren im Kern beschrieben werden, auch als *Syzygien*, und syz ist die Abkürzung dieses Begriffs.

```

> LIB "matrix.lib";
> ring r=0,t,dp;
> matrix A[4][5]=1,2,1,3,1,2,4,7,3,0,4,8,9,9,2,3,6,0,2,1;
> matrix B=syz(A);
> print(B);
-2,0,
1, -1,
0, 4,
0, -8,
0, 22

```

Der Kern von A hat also die Basisvektoren $(-2, 1, 0, 0, 0)^t$ und $(0, -1, 4, -8, 22)^t$.

Beispiel 35.3 (Lösung eines linearen Gleichungssystems)

Wir setzen nun $\mathbf{b} = (2, 2, 6, 4)^t$ und wollen das lineare Gleichungssystem $Ax = \mathbf{b}$ lösen. Der Befehl `concat` hängt zwei Matrizen hintereinander.

```

> matrix b[4][1]=2,2,6,4;
> matrix Ab=concat(A,b); // Bilde die erweiterte Koeffizientenmatrix.
> print(Ab);
1,2,1,3,1,2,
2,4,7,3,0,2,
4,8,9,9,2,6,
3,6,0,2,1,4
> print(syz(Ab)); // Berechne eine Basis des Kerns von Ab.
-2,0, 0,
1, -1,0,
0, 4, -2,
0, -8,4,
0, 22,-12,
0, 0, 1

```

Wir haben nun eine Basis des Kerns der erweiterten Koeffizientenmatrix berechnet. Der Algorithmus stellt sicher, daß es genau dann einen Vektor mit letzter Komponente ungleich null gibt, wenn das Gleichungssystem lösbar ist. Es gibt dann auch nur einen solchen Vektor und das ist der letzte Basisvektor. Dividiert man die ersten fünf Einträge des Vektors durch das Negative des letzten Eintrags, so erhält man eine spezielle Lösung, hier

$$\mathbf{c} = (0, 0, 2, -4, 12)^t.$$

Vergißt man bei den übrigen Vektoren in der berechneten Basis die letzte Komponente, so erhält man eine Basis des homogenen Lösungsraums, hier

$$\text{Lös}(A, 0) = \text{Lin} \left((-2, 1, 0, 0, 0)^t, (0, -1, 4, -8, 22)^t \right),$$

wie wir schon aus Beispiel 35.2 wissen.

Beispiel 35.4 (Eigenwerte einer Matrix)

Wir wollen die Eigenwerte der folgenden Matrix bestimmen

$$A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix} \in \text{Mat}_3(\mathbb{Q}).$$

Dazu verwenden wir unter anderem die SINGULAR-Befehle `det` zum Berechnen der Determinante, `unitmat` für die Einheitsmatrix und `factorize` zum Berechnen der Primfaktorzerlegung eines Polynoms.

```
> LIB "matrix.lib";
> ring r=0,t,dp;
> matrix A[3][3]=0,1,1,-1,2,1,-1,1,2;
> poly p=det(t*unitmat(3)-A);
> p;
t3-4t2+5t-2
> short=0;
> p;
t^3-4*t^2+5*t-2
> factorize(p);
[1]:
  _[1]=1
  _[2]=t-1
  _[3]=t-2
[2]:
  1,2,1
```

Das charakteristische Polynom von A ist also

$$\chi_A = t^3 - 4t^2 + 5t - 2 = (t - 1)^2 \cdot (t - 2),$$

so daß die Eigenwerte $\lambda = 1$ mit algebraischer Vielfachheit 2 sowie $\lambda = 2$ mit algebraischer Vielfachheit 1 sind.

Beispiel 35.5 (Minimalpolynom einer Matrix)

Als nächstes wollen wir das Minimalpolynom der Matrix A in Beispiel 35.4 berechnen. Dazu verwenden wir den Algorithmus 33.15 sowie einige SINGULAR-Befehle. `transpose` transponiert eine Matrix, `flatten` schreibt die Einträge einer Matrix in einen Zeilenvektor und `power` potenziert eine Matrix.

```
> matrix C=transpose(flatten(power(A,0)));
> for (int i=1;i<=3;i++)
. {
.   C=concat(C,transpose(flatten(power(A,i))));
. }
```

```

> matrix D=syz(C);
> print(D);
2, 0,
-3,2,
1, -3,
0, 1
> poly mu;
> for (i=1;i<=4;i++){mu=mu+D[1][i]*t^(i-1);}
> mu;
t^2-3*t+2
> factorize(mu);
[1]:
  _[1]=1
  _[2]=t-1
  _[3]=t-2
[2]:
  1,1,1

```

Das Minimalpolynom von A ist also

$$\mu_A = t^2 - 3t + 2 = (t - 1) \cdot (t - 2),$$

und die Matrix A ist somit diagonalisierbar, da das Minimalpolynom in paarweise verschiedene Linearfaktoren zerfällt.

Die obige Befehlssequenz ist recht lang. Falls man bereits weiß, daß das Minimalpolynom über dem Grundkörper in Linearfaktoren zerfällt, so kann man auch den SINGULAR-Befehl `minipoly` aus der Bibliothek `linalg.lib` verwenden, aber nur dann! Um sicherzustellen, daß das Minimalpolynom zerfällt, kann man zunächst das charakteristische Polynom berechnen und faktorisieren, denn nur wenn dieses zerfällt, zerfällt auch das Minimalpolynom. Für die Matrix A aus unserem Beispiel wissen wir bereits, daß es zerfällt. Wir können also den Befehl `minipoly` anwenden.

```

> LIB "linalg.lib";
> minipoly(A);
[1]:          // das Minimalpolynom hat die zwei Nullstellen 1 und 2
  _[1]=1
  _[2]=2
[2]:          // beide kommen mit Vielfachheit 1 vor
  1,1

```

Beispiel 35.6 (Diagonalisierung einer Matrix)

Wir haben in Beispiel 35.5 gesehen, daß die Matrix A aus Beispiel 35.4 diagonalisierbar ist. Nun wollen wir die zugehörige Transformationsmatrix T bestimmen. Dazu

erinnern wir uns, daß A genau die Eigenwerte 1 und 2 besitzt. Zu diesen müssen wir Basen der Eigenräume bestimmen.

```
> matrix T1=syz(A-unitmat(3));
> print(T1);
1,0,
1,-1,
0,1
> matrix T2=syz(A-2*unitmat(3));
> print(T2);
1,
1,
1
> matrix T=concat(T1,T2);
> print(T);
1,0, 1,
1,-1,1,
0,1, 1
> print(inverse(T)*A*T);
1,0,0,
0,1,0,
0,0,2
```

Beispiel 35.7 (Jordansche Normalform)

In diesem Beispiel wollen wir die Jordansche Normalform und die Transformationsmatrix T für die Matrix

$$A = \begin{pmatrix} 21 & 5 & 1 & 21 & 5 \\ -23 & 4 & 8 & -31 & 1 \\ -2 & -1 & -2 & -1 & -1 \\ -17 & -4 & -1 & -17 & -4 \\ 22 & -2 & -8 & 30 & 1 \end{pmatrix} \in \text{Mat}_5(\mathbb{Q})$$

berechnen. Dazu bestimmen wir zunächst das charakteristische Polynom und das Minimalpolynom von A und faktorisieren diese.

```
> LIB "matrix.lib";
> LIB "linalg.lib";
> ring R=0,t,dp;
> matrix A[5][5]=21, 5, 1, 21, 5,
. -23,4, 8, -31,1,
. -2, -1,-2,-1, -1,
. -17,-4,-1,-17,-4,
. 22, -2,-8,30, 1;
```

```

> print(A);
21, 5, 1, 21, 5,
-23,4, 8, -31,1,
-2, -1,-2,-1, -1,
-17,-4,-1,-17,-4,
22, -2,-8,30, 1
> short=0;
> poly chi=det(t*unitmat(5)-A);
> chi;
t^5-7*t^4+10*t^3+18*t^2-27*t-27
> factorize(chi);
[1]:
  _[1]=1
  _[2]=t-3
  _[3]=t+1
[2]:
  1,3,2
> minipoly(A);
[1]:
  _[1]=-1
  _[2]=3
[2]:
  2,2

```

Wir sehen also, daß

$$\chi_A = (t - 3)^3 \cdot (t + 1)^2$$

und

$$\mu_A = (t - 3)^2 \cdot (t + 1)^2.$$

Damit ist die Jordansche Normalform von A festgelegt. Sie muß zu den beiden Eigenwerten 3 und -1 je mindestens einen Jordanblock der Größe 2 enthalten, weil sie im Minimalpolynom beide mit Vielfachheit zwei vorkommen. Zudem muß sie den Eigenwert 3 noch ein drittes Mal auf der Diagonalen haben, so daß ein weiterer Jordanblock der Größe eins zum Eigenwert 3 nötig ist. Also gilt

$$J_A = \left(\begin{array}{cc|cc} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ \hline 0 & 0 & 0 & 0 & -1 \end{array} \right).$$

Da die Jordansche Normalform drei Jordanblöcke besitzt, müssen wir letztlich drei Basisvektoren finden, die die zyklischen A -invarianten Unterräume definieren, zu denen die Blöcke gehören. Dabei wenden wir den Algorithmus 34.11 an.

```

> matrix B=syz(A-3*unitmat(5)); // Basis von Ker(A-3*id)
> print(B);
-5,0,
1, -1,
1, 0,
4, 0,
0, 1
> matrix C=syz(power(A-3*unitmat(5),2)); // Basis von Ker((A-3*id)^2)
> print(C);
0,-5,0,
1,0, 0,
0,1, 0,
0,4, 0,
0,0, 1

```

Ein kurzer Blick genügt, um zu sehen, daß der erste und der dritte Basisvektor von $\text{Ker}((A - 3 \cdot \mathbf{1}_5)^2)$ nicht im $\text{Ker}(A - 3 \cdot \mathbf{1}_5)$ liegt. Wir können also jeden der beiden wählen, um den zyklischen Unterraum der Größe zwei zum Eigenwert 3 zu bilden. Wählen wir

$$x_1 = (0, 1, 0, 0, 0)^t.$$

```

> // berechne (A-3*unitmat(5)) * erste Spalte von C
. matrix X1[5][1]=C[1..5,1];
> print((A-3*unitmat(5))*X1);
5,
1,
-1,
-4,
-2

```

Damit hat der zyklische Unterraum der Größe zwei zum Eigenwert 3 die Basisvektoren

$$(A - 3 \cdot \mathbf{1}_5)x_1 = (5, 1, -1, -4, -2)^t \quad \text{und} \quad x_1 = (0, 1, 0, 0, 0)^t,$$

und diese sind die ersten beiden Spalten der Matrix T .

Nun müssen wir noch den Vektor $(A - 3 \cdot \mathbf{1}_5)x_1$ zu einer Basis von $\text{Ker}(A - 3 \cdot \mathbf{1}_5)$ ergänzen. Ein Blick auf die Basis von $\text{Ker}(A - 3 \cdot \mathbf{1}_5)$ zeigt, daß jeder der beiden Vektoren es tut.

```

> matrix X2[5][1]=B[1..5,2]; // waehle X2

```

Wir wählen deshalb

$$x_2 = (0, -1, 0, 0, 1)^t,$$

und dieser ist die dritte Spalte von T .

```

> print(syz(A+unitmat(5))); // Basis von Ker(A+id)
-1,
0,
1,
1,
0
> matrix D=syz(power(A+unitmat(5),2)); // Basis von Ker((A+id)^2)
> print(D);
-1,0,
0, -2,
1, 1,
1, 0,
0, 2
> matrix X3[5][1]=D[1..5,2];
> print((A+unitmat(5))*X3); // (A+unitmat(5)) * 2. Spalte von D
1,
0,
-1,
-1,
0

```

Daraus folgt, daß der Vektor

$$x_3 = (0, -2, 1, 0, 2)^t \in \text{Ker}(A + \mathbb{1}_5^2) \setminus \text{Ker}(A + \mathbb{1}_5)$$

liegt, und daß die letzten beiden Spalten von T die Vektoren

$$(A + \mathbb{1}_5)x_3 = (1, 0, -1, 1, 0)^t \quad \text{und} \quad x_3 = (0, -2, 1, 0, 2)^t$$

sind.

```

> // bestuecke die Matrix T
. matrix T=(A-3*unitmat(5))*X1;
> T=concat(T,X1);
> T=concat(T,X2);
> T=concat(T,(A+unitmat(5))*X3);
> T=concat(T,X3);
> print(T);
5, 0,0, 1, 0,
1, 1,-1,0, -2,
-1,0,0, -1,1,
-4,0,0, -1,0,
-2,0,1, 0, 2
> // invertiere die Matrix T
. matrix S=inverse(T);

```



```

> print(S);
1, 0,0, 1, 0,
1, 1,0, 1, 1,
8, 0,-2,10,1,
-4,0,0, -5,0,
-3,0,1, -4,0
> print(inverse(T)*A*T);
3,1,0,0, 0,
0,3,0,0, 0,
0,0,3,0, 0,
0,0,0,-1,1,
0,0,0,0, -1

```

Wir erhalten also

$$T = \begin{pmatrix} 5 & 0 & 0 & 1 & 0 \\ 1 & 1 & -1 & 0 & -2 \\ -1 & 0 & 0 & -1 & 1 \\ -4 & 0 & 0 & -1 & 0 \\ -2 & 0 & 1 & 0 & 2 \end{pmatrix} \quad \text{und} \quad T^{-1} = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 8 & 0 & -2 & 10 & 1 \\ -4 & 0 & 0 & -5 & 0 \\ -3 & 0 & 1 & -4 & 0 \end{pmatrix}$$

sowie

$$J_A = T^{-1} \circ A \circ T = \left(\begin{array}{cc|cc} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{array} \right).$$

Es gibt in SINGULAR auch einen schnelleren Weg zur Jordanschen Normalform und der Transformationsmatrix, wenn man den Befehl `jordanbasis` verwendet, was in den SINGULAR-Übungsaufgaben aber nicht gemacht werden soll!

```

> matrix E=jordanbasis(A)[1];
> matrix Z[5][5];
> for (int j=1;j<=5;j++) { Z[1..5,j]=E[1..5,6-j]; }
> print(Z);
-5,5, 0,1, 0,
1, 1, 1,0, -2,
1, -1,0,-1,1,
4, -4,0,-1,0,
0, -2,0,0, 2
> print(inverse(Z)*A*Z);
3,0,0,0, 0,
0,3,1,0, 0,
0,0,3,0, 0,
0,0,0,-1,1,
0,0,0,0, -1

```

Die `for`-Schleife oben kehrt die Reihenfolge der Spalten in der Matrix E um. Die neue Matrix Z ist dann eine zulässige Transformationsmatrix T , wobei die Reihenfolge der Jordanblöcke sich geändert hat. Die Vertauschung der Spalten ist nötig, da die Konvention der Jordanschen Normalform in SINGULAR nicht mit unserer Konvention übereinstimmt. Darauf möchte ich hier aber nicht näher eingehen.

Beispiel 35.8 (Näherungsweise Bestimmung von Eigenwerten)

Eine zufällig ausgewählte Matrix in $\text{Mat}_n(\mathbb{Q})$ wird keine rationalen Eigenwerte haben. Betrachten wir sie aber als Matrix in $\text{Mat}_n(\mathbb{C})$, so zerfällt sie in Linearfaktoren und mit Wahrscheinlichkeit 1 sind diese paarweise verschieden. Exakt berechnen können wir sie aber nicht, da die Zerlegung eines Polynoms in $\mathbb{C}[t]$ in seine Primfaktoren im allgemeinen nicht möglich ist. Wir können die Eigenwerte aber näherungsweise berechnen, und dies reicht häufig aus, um zu sehen, daß sie paarweise verschieden sind.

```
> LIB "matrix.lib";
> ring S=complex,t,lp;
> matrix M[3][3];
> int i,j;
> for (i=1;i<=3;i++){for (j=1;j<=3;j++){M[i,j]=random(-9,9);}}
> print(M);
8, 5, 0,
-6,-2,3,
9, -9,7
> poly f=det(t*unitmat(3)-M);
> short=0;
> f;
t^3-13*t^2+83*t-449
> LIB "solve.lib";
> solve(f);
[1]:
  9.27119961
[2]:
  (1.86440019-i*6.70474155)
[3]:
  (1.86440019+i*6.70474155)
```

Das charakteristische Polynom der 3×3 -Matrix ist ein Polynom vom Grad drei mit reellen Koeffizienten. Wegen des Zwischenwertsatzes muß es eine reelle Nullstelle haben. Wenn es keine weitere reelle Nullstelle besitzt, so müssen die übrigen beiden Nullstellen komplex konjugiert zueinander sein. Unsere Rechnung oben approximiert die Nullstellen mit dem Befehl `solve` aus der Bibliothek `solve.lib`, und wir sehen an den approximierten Nullstellen das geschilderte Phänomen.

Aufgaben**Aufgabe 35.9**

Bestimme mit Hilfe von SINGULAR eine Basis B von \mathbb{Q}^5 , bezüglich derer die Matrixdarstellung der Abbildung $f: \mathbb{Q}^5 \rightarrow \mathbb{Q}^5$ Jordansche Normalform hat, wo:

$$f(x_1, x_2, x_3, x_4, x_5) = (x_1 - x_2, x_1 + 2x_2 - x_3, -x_1 + 3x_3, -x_1 - 2x_2 - 2x_3 + 2x_4 - x_5, x_1 - x_3 + 2x_5)^t.$$

§ 36 Bilinearformen und Sesquilinearformen

A) Bilinearformen

In Definition 30.8 haben wir den Begriff einer multilinearen Abbildung auf Vektorräumen eingeführt. In diesem Abschnitt wollen wir einen besonders wichtigen Spezialfall dieser Begriffsbildung untersuchen, die Bilinearformen.

Definition 36.1 (Bilinearformen)

Es sei V ein K -Vektorraum. Eine Abbildung

$$\mathbf{b} : V \times V \rightarrow K,$$

die linear in beiden Argumenten ist, nennen wir *bilinear* oder eine *Bilinearform*, d. h. für $x, y, z \in V$ und $\lambda, \mu \in K$ gilt (vgl. Definition 30.8):

$$\mathbf{b}(\lambda x + \mu y, z) = \lambda \mathbf{b}(x, z) + \mu \mathbf{b}(y, z)$$

und

$$\mathbf{b}(z, \lambda x + \mu y) = \lambda \mathbf{b}(z, x) + \mu \mathbf{b}(z, y).$$

Die Menge aller Bilinearformen auf V bezeichnen wir mit

$$\text{Bil}_K(V) = \{\mathbf{b} : V \times V \rightarrow K \mid \mathbf{b} \text{ ist bilinear}\}.$$

Beispiel 36.2 (Bilinearformen)

- a. Die Determinante definiert eine Bilinearform

$$\det : K^2 \times K^2 \longrightarrow K : \left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right) \mapsto \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} = x_1 y_2 - x_2 y_1.$$

- b. Ist $A \in \text{Mat}_n(K)$ eine quadratische Matrix, dann wird durch

$$\mathbf{b}_A : K^n \times K^n \longrightarrow K : (x, y) \mapsto \mathbf{b}_A(x, y) = x^t \circ A \circ y = x^t A y$$

eine Bilinearform auf K^n definiert, wie unmittelbar aus der Distributivität des Matrixproduktes folgt.

- c. Wählen wir in Teil b. die Matrix $A = \mathbb{1}_n$, so erhalten wir die Bilinearform

$$\mathbf{b}_A : K^n \times K^n \longrightarrow K : (x, y) \mapsto \sum_{i=1}^n x_i \cdot y_i.$$

- d. Ist $n = 2$ und ist

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \text{Mat}_2(K),$$

so ist $\mathbf{b}_A = \det$.

Bemerkung 36.3

Die Menge $\text{Bil}_K(V)$ ist offenbar ein Unterraum des K -Vektorraums $(V \times V)^K$ aller Abbildungen von $V \times V$ nach K , d.h. die Summe zweier Bilinearformen sowie das skalare Vielfache einer Bilinearform sind wieder bilinear.

Definition 36.4 (Matrixdarstellung von Bilinearformen)

Es sei V ein K -Vektorraum mit Basis $B = (x_1, \dots, x_n)$ und $b : V \times V \rightarrow K$ eine Bilinearform auf V . Wir nennen die Matrix

$$M_B(b) = (b(x_i, x_j))_{1 \leq i, j \leq n} \in \text{Mat}_n(K)$$

die *Matrixdarstellung* von b bezüglich der Basis B .

Beispiel 36.5

Wir betrachten die Bilinearform $b = b_{\mathbb{1}_2}$ auf \mathbb{R}^2 sowie die Basis $B = (x_1, x_2) = ((1, 1)^t, (1, 2)^t)$. Dann ist

$$M_B(b) = \begin{pmatrix} b(x_1, x_1) & b(x_1, x_2) \\ b(x_2, x_1) & b(x_2, x_2) \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix}.$$

Bemerkung 36.6 (Matrixdarstellung von Vektoren)

Wir sollten darauf hinweisen, daß wir die Bezeichnung M_B bereits einmal verwendet haben, nämlich bei der Matrixdarstellung von Vektoren bezüglich einer Basis $B = (x_1, \dots, x_n)$. Ist $x = \sum_{i=1}^n \lambda_i x_i$, so bezeichnet

$$M_B(x) = (\lambda_1, \dots, \lambda_n)^t \in K^n$$

die Matrixdarstellung des Vektors $x \in V$ bezüglich der Basis B . Wir werden diese Bezeichnung auch im folgenden wieder benötigen, es wird aus dem Kontext aber stets unmittelbar ersichtlich sein, in welcher Bedeutung M_B verwendet wird.

Proposition 36.7 (Matrixdarstellung von Bilinearformen)

Es sei V ein K -Vektorraum mit Basis $B = (x_1, \dots, x_n)$. Dann ist die Abbildung

$$M_B : \text{Bil}_K(V) \longrightarrow \text{Mat}_n(K) : b \mapsto M_B(b)$$

ein Isomorphismus von K -Vektorräumen.

Insbesondere ist eine Bilinearform durch ihre Matrixdarstellung eindeutig bestimmt, und für $x, y \in V$ gilt

$$b(x, y) = M_B(x)^t \circ M_B(b) \circ M_B(y).$$

Beweis: Sind $b, b' \in \text{Bil}_K(V)$ und $\lambda, \lambda' \in K$, so gilt

$$\begin{aligned} M_B(\lambda \cdot b + \lambda' \cdot b') &= ((\lambda b + \lambda' b')(x_i, x_j))_{i,j} = (\lambda \cdot b(x_i, x_j) + \lambda' \cdot b'(x_i, x_j))_{i,j} \\ &= \lambda \cdot (b(x_i, x_j))_{i,j} + \lambda' \cdot (b'(x_i, x_j))_{i,j} = \lambda \cdot M_B(b) + \lambda' \cdot M_B(b'). \end{aligned}$$

Also ist die Abbildung M_B linear. Zudem folgt aus Aufgabe 31.23, daß es zu jeder Matrix $Y = (y_{ij}) \in \text{Mat}_n(K)$ genau eine bilineare Abbildung $b : V \times V \rightarrow K$ gibt mit $b(x_i, x_j) = y_{ij}$, d.h.

$$M_B(b) = (b(x_i, x_j))_{i,j} = (y_{ij})_{i,j} = Y.$$

Damit ist die Abbildung M_B aber bijektiv.

Ferner gilt für $\mathbf{x} = \sum_{i=1}^n \lambda_i \mathbf{x}_i$ und $\mathbf{y} = \sum_{j=1}^n \mu_j \mathbf{x}_j$

$$\mathbf{b}(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \lambda_i \mathbf{b}(\mathbf{x}_i, \mathbf{y}) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \mu_j \mathbf{b}(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{M}_B(\mathbf{x})^t \cdot \mathbf{M}_B(\mathbf{b}) \cdot \mathbf{M}_B(\mathbf{y}).$$

□

Beispiel 36.8

Wir betrachten die Bilinearform aus Beispiel 36.5 sowie die Vektoren $\mathbf{x} = (0, 1)^t = -\mathbf{x}_1 + \mathbf{x}_2$ und $\mathbf{y} = (3, 4)^t = 2\mathbf{x}_1 + \mathbf{x}_2$. Dann gilt

$$\mathbf{M}_B(\mathbf{x}) = (-1, 1)^t \quad \text{und} \quad \mathbf{M}_B(\mathbf{y}) = (2, 1)^t$$

sowie

$$4 = \mathbf{b}(\mathbf{x}, \mathbf{y}) = \mathbf{M}_B(\mathbf{x})^t \circ \mathbf{M}_B(\mathbf{b}) \circ \mathbf{M}_B(\mathbf{y}) = \begin{pmatrix} -1 & 1 \end{pmatrix} \circ \begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} \circ \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$$

Korollar 36.9 (Alle Bilinearformen auf \mathbb{K}^n sind von der Form \mathbf{b}_A .)

Die Abbildung $\text{Mat}_n(\mathbb{K}) \rightarrow \text{Bil}_{\mathbb{K}}(\mathbb{K}^n) : A \mapsto \mathbf{b}_A$ ist ein Isomorphismus.

Insbesondere ist jede Bilinearform auf \mathbb{K}^n von der Gestalt \mathbf{b}_A für eine eindeutig bestimmte Matrix $A \in \text{Mat}_n(\mathbb{K})$.

Beweis: Die Abbildung $A \mapsto \mathbf{b}_A$ ist die Umkehrabbildung der Matrixdarstellung

$$\mathbf{M}_E : \text{Bil}_{\mathbb{K}}(\mathbb{K}^n) \rightarrow \text{Mat}_n(\mathbb{K})$$

bezüglich der kanonischen Basis $E = (\mathbf{e}_1, \dots, \mathbf{e}_n)$ von \mathbb{K}^n , da

$$\mathbf{b}_A(\mathbf{e}_i, \mathbf{e}_j) = \mathbf{e}_i^t \circ A \circ \mathbf{e}_j = a_{ij}$$

und somit $\mathbf{M}_E(\mathbf{b}_A) = A$. □

Bemerkung 36.10

Die Aussage in Korollar 36.9 bedeutet insbesondere, daß

$$\mathbf{b} = \mathbf{b}_{\mathbf{M}_E(\mathbf{b})} \quad \text{und} \quad A = \mathbf{M}_E(\mathbf{b}_A).$$

für $\mathbf{b} \in \text{Bil}_{\mathbb{K}}(\mathbb{K}^n)$ und $A \in \text{Mat}_n(\mathbb{K})$, wenn E die kanonische Basis von \mathbb{K}^n ist.

Satz 36.11 (Basiswechsel bei Bilinearformen)

Sei V ein \mathbb{K} -Vektorraum mit Basen $B = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ und $D = (\mathbf{y}_1, \dots, \mathbf{y}_n)$ und sei $\mathbf{b} \in \text{Bil}_{\mathbb{K}}(V)$, dann gilt

$$\mathbf{M}_B(\mathbf{b}) = (\mathbf{T}_D^B)^t \circ \mathbf{M}_D(\mathbf{b}) \circ \mathbf{T}_D^B.$$

Beweis: Wir setzen $\mathbf{M}_B(\mathbf{b}) = (a_{ij})$ und $(\mathbf{T}_D^B)^t \circ \mathbf{M}_D(\mathbf{b}) \circ \mathbf{T}_D^B = C = (c_{ij}) \in \text{Mat}_n(\mathbb{K})$. Man beachte, daß für einen beliebigen Vektor $\mathbf{z} \in V$ nach Bemerkung 26.4 die Gleichung

$$\mathbf{M}_D(\mathbf{z}) = \mathbf{M}_D(\text{id}_V(\mathbf{z})) = \mathbf{M}_D^B(\text{id}_V) \circ \mathbf{M}_B(\mathbf{z}) = \mathbf{T}_D^B \circ \mathbf{M}_B(\mathbf{z}) \quad (106)$$

gilt. Damit erhalten wir für $1 \leq i, j \leq n$ dann

$$\begin{aligned} \mathbf{a}_{ij} &\stackrel{\text{Def.}}{=} \mathbf{b}(x_i, x_j) \stackrel{36.7}{=} M_D(x_i)^t \circ M_D(\mathbf{b}) \circ M_D(x_j) \\ &\stackrel{(106)}{=} (T_D^B \circ M_B(x_i))^t \circ M_D(\mathbf{b}) \circ (T_D^B \circ M_B(x_j)) \\ &= M_B(x_i)^t \circ (T_D^B)^t \circ M_D(\mathbf{b}) \circ T_D^B \circ M_B(x_j) = \mathbf{e}_i^t \circ C \circ \mathbf{e}_j = \mathbf{c}_{ij}. \end{aligned}$$

Also gilt $M_B(\mathbf{b}) = (\mathbf{a}_{ij}) = (\mathbf{c}_{ij}) = (T_D^B)^t \circ M_D(\mathbf{b}) \circ T_D^B$, was zu zeigen war. □

Beispiel 36.12

Die Matrixdarstellung der Bilinearform $\mathbf{b} = \mathbf{b}_{\mathbb{1}_2}$ aus Beispiel 36.5 bezüglich der kanonischen Basis E ist die Einheitsmatrix $\mathbb{1}_2$. Zugleich ist

$$T_E^B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

die Transformationsmatrix für die Basis $B = ((1, 1)^t, (1, 2)^t)$. Dann gilt aber

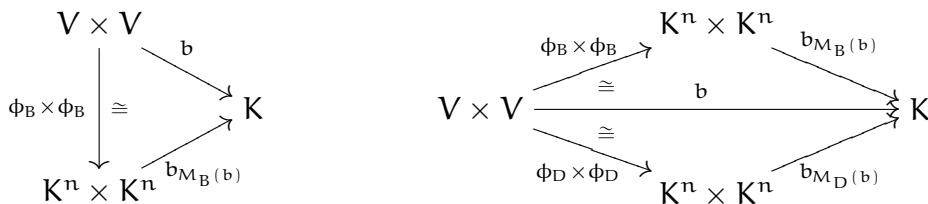
$$\begin{pmatrix} 2 & 3 \\ 3 & 5 \end{pmatrix} = M_B(\mathbf{b}) = (T_E^B)^t \circ M_E(\mathbf{b}) \circ T_E^B = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \circ \mathbb{1}_2 \circ \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Bemerkung 36.13

- a. Es bleibt festzuhalten, daß sowohl Endomorphismen $f : V \rightarrow V$ als auch Bilinearformen $\mathbf{b} : V \times V \rightarrow K$ sich nach Wahl einer Basis B durch Matrizen $M_B^B(f)$ bzw. $M_B(\mathbf{b})$ beschreiben lassen. Bei Basiswechsel, der durch die Matrix $T = T_D^B$ beschrieben wird, haben Endomorphismen und Bilinearformen aber ein unterschiedliches Transformationsverhalten. Es gilt:

$$M_B^B(f) = T^{-1} \circ M_D^D(f) \circ T \quad \text{und} \quad M_B(\mathbf{b}) = T^t \circ M_D(\mathbf{b}) \circ T.$$

- b. Die Aussagen von Proposition 36.7 und Satz 36.11 lassen sich mit Hilfe der Karten $\phi_B : V \rightarrow K^n : x_i \mapsto e_i$ und $\phi_D : V \rightarrow K^n : y_i \mapsto e_i$ auch durch folgende kommutative Diagramme veranschaulichen:



B) Normalform symmetrischer Bilinearformen

Definition 36.14 (Symmetrische Bilinearformen)

Es sei V ein K -Vektorraum.

- a. Eine Bilinearform $\mathbf{b} \in \text{Bil}_K(V)$ heißt *symmetrisch*, falls für $x, y \in V$ stets gilt

$$\mathbf{b}(x, y) = \mathbf{b}(y, x).$$

- b. Eine Matrix $A \in \text{Mat}_n(K)$ heißt *symmetrisch*, falls $A = A^t$.

Beispiel 36.15

a. Die Bilinearform \det aus Beispiel 36.2 ist für $K = \mathbb{R}$ nicht symmetrisch, da

$$\det(\mathbf{e}_1, \mathbf{e}_2) = 1 \neq -1 = \det(\mathbf{e}_2, \mathbf{e}_1).$$

b. Die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \in \text{Mat}_2(\mathbb{Q})$$

ist symmetrisch.

Proposition 36.16 (Symmetrische Bilinearformen)

Sei V ein K -Vektorraum mit Basis $B = (\mathbf{x}_1, \dots, \mathbf{x}_n)$, $\mathbf{b} \in \text{Bil}_K(V)$ und $A \in \text{Mat}_n(K)$.

a. Die Bilinearform \mathbf{b} ist genau dann symmetrisch, wenn $M_B(\mathbf{b})$ symmetrisch ist.

b. Die Bilinearform \mathbf{b}_A ist genau dann symmetrisch, wenn A symmetrisch ist.

Beweis: Es sei $M_B(\mathbf{b}) = A = (a_{ij}) \in \text{Mat}_n(K)$.

Ist A symmetrisch, so ist auch \mathbf{b} symmetrisch, da für $\mathbf{x}, \mathbf{y} \in V$ gilt

$$\begin{aligned} \mathbf{b}(\mathbf{x}, \mathbf{y}) &= M_B(\mathbf{x})^t \circ A \circ M_B(\mathbf{y}) = (M_B(\mathbf{x})^t \circ A \circ M_B(\mathbf{y}))^t \\ &= M_B(\mathbf{y})^t \circ A^t \circ M_B(\mathbf{x}) = M_B(\mathbf{y})^t \circ A \circ M_B(\mathbf{x}) = \mathbf{b}(\mathbf{y}, \mathbf{x}). \end{aligned}$$

Ist umgekehrt \mathbf{b} symmetrisch, dann folgt für $1 \leq i, j \leq n$

$$a_{ij} = \mathbf{b}(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{b}(\mathbf{x}_j, \mathbf{x}_i) = a_{ji},$$

so daß auch $A = A^t$ symmetrisch ist. Damit ist a. gezeigt, und b. folgt aus a., da $A = M_E(\mathbf{b}_A)$ für die kanonische Basis E von K^n gilt. \square

Definition 36.17 (Quadratische Form)

Ist $\mathbf{b} \in \text{Bil}_K(V)$ eine symmetrische Bilinearform auf V , dann nennen wir

$$q_{\mathbf{b}} : V \rightarrow K : \mathbf{x} \mapsto \mathbf{b}(\mathbf{x}, \mathbf{x})$$

die *quadratische Form* zu \mathbf{b} . Für $A \in \text{Mat}_n(K)$ schreiben wir auch q_A statt $q_{\mathbf{b}_A}$.

Beispiel 36.18

Ist $A = (a_{ij}) \in \text{Mat}_n(K)$ symmetrisch, dann gilt für $\mathbf{x} = (t_1, \dots, t_n)^t \in K^n$

$$q_A(\mathbf{x}) = \mathbf{x}^t A \mathbf{x} = \sum_{i,j} a_{ij} t_i t_j = \sum_{i=1}^n a_{ii} t_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij} t_i t_j.$$

Damit können wir q_A als ein homogenes quadratisches Polynom in den Unbestimmten t_1, \dots, t_n auffassen.

Z.B., ist

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 5 \end{pmatrix} \in \text{Mat}_2(\mathbb{R}),$$

so gilt

$$q_A(t_1, t_2) = t_1^2 + 4t_1 t_2 + 5t_2^2.$$

Auf den ersten Blick scheint es, daß die quadratische Form q_b weit weniger Information enthält, als die symmetrische Bilinearform b . Erstaunlicherweise kann man b jedoch aus q_b zurückgewinnen, wenn $\text{char}(\mathbb{K}) \neq 2$ (siehe Definition 8.9) ist.

Proposition 36.19 (Polarisierung einer Bilinearform)

Sei \mathbb{K} ein Körper mit $\text{char}(\mathbb{K}) \neq 2$ und $b \in \text{Bil}_{\mathbb{K}}(\mathbb{V})$ eine symmetrische Bilinearform.

Dann gilt für $x, y \in \mathbb{V}$

$$b(x, y) = \frac{1}{2}(q_b(x + y) - q_b(x) - q_b(y)).$$

Beweis: Die Aussage folgt durch einfaches Einsetzen der Definition von q_b in die rechte Seite. \square

Satz 36.20 (Existenz einer Orthogonalbasis)

Sei \mathbb{K} ein Körper mit $\text{char}(\mathbb{K}) \neq 2$ und \mathbb{V} ein \mathbb{K} -Vektorraum mit $1 \leq \dim_{\mathbb{K}}(\mathbb{V}) < \infty$.

Ist $b \in \text{Bil}_{\mathbb{K}}(\mathbb{V})$ eine symmetrische Bilinearform, dann besitzt \mathbb{V} eine Basis $B = (x_1, \dots, x_n)$, so daß $M_B(b)$ eine Diagonalmatrix ist, d.h.

$$b(x_i, x_j) = 0 \quad \forall i \neq j.$$

Wir nennen eine solche Basis B eine Orthogonalbasis bezüglich der Bilinearform b .

Beweis: Wir führen den Beweis durch Induktion über $n = \dim_{\mathbb{K}}(\mathbb{V})$, wobei im Fall $n = 1$ nichts zu zeigen ist.

Sei also $n > 1$. Wir bezeichnen mit $q_b : \mathbb{V} \rightarrow \mathbb{K} : x \mapsto b(x, x)$ die zu b gehörende quadratische Form. Ist q_b identisch Null, so ist nach Proposition 36.19 auch b identisch Null, da $\text{char}(\mathbb{K}) \neq 2$, und jede Basis ist eine Orthogonalbasis bezüglich b .

Wir können also annehmen, daß es ein $x \in \mathbb{V}$ gibt mit $b(x, x) = q_b(x) \neq 0$. Setze $U := \text{Lin}(x)$ und

$$U^\perp := \{y \in \mathbb{V} \mid b(x, y) = 0\}.$$

Aus der Bilinearität von b folgt, daß U^\perp ein Unterraum von \mathbb{V} ist. Wir wollen nun zeigen, daß in der Tat $\mathbb{V} = U \oplus U^\perp$ gilt.

Sei dazu zunächst $y \in \mathbb{V}$ beliebig. Dann setzen wir

$$x' := \frac{b(x, y)}{b(x, x)} \cdot x \in U$$

und erhalten

$$b(y - x', x) = b(y, x) - \frac{b(x, y)}{b(x, x)} \cdot b(x, x) = 0$$

Deshalb ist $y - x' \in U^\perp$ und

$$y = x' + (y - x') \in U + U^\perp,$$

womit $\mathbb{V} = U + U^\perp$ gezeigt ist.

Sei nun $y \in U \cap U^\perp$, dann gibt es ein $\lambda \in \mathbb{K}$ mit $y = \lambda x$ und damit

$$\lambda \cdot q_b(x) = b(x, \lambda x) = b(x, y) = 0.$$

Da aber $q_b(x) \neq 0$ gilt, ist $\lambda = 0$ und damit $y = 0$. Also gilt $U \cap U^\perp = \{0\}$. Insgesamt haben wir damit $V = U \oplus U^\perp$ gezeigt.

Schränken wir nun b auf U^\perp ein, so erhalten wir per Induktion eine Orthogonalbasis (x_2, \dots, x_n) von U^\perp bezüglich b und $B = (x, x_2, \dots, x_n)$ ist dann die gesuchte Orthogonalbasis von V . □

Korollar 36.21 (Diagonalisierbarkeit symmetrischer Bilinearformen)

Es sei K ein Körper mit $\text{char}(K) \neq 2$ und $A \in \text{Mat}_n(K)$ eine symmetrische Matrix. Dann gibt es eine Matrix $T \in \text{Gl}_n(K)$ mit

$$T^t \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix},$$

wobei $\lambda_1, \dots, \lambda_r \in K \setminus \{0\}$, $\lambda_{r+1} = \dots = \lambda_n = 0$ und $r = \text{rang}(A)$.

Beweis: Nach Satz 36.20 besitzt K^n eine Orthogonalbasis $B = (x_1, \dots, x_n)$ bezüglich der Bilinearform b_A . Dabei können wir durch Ummumerieren o. E. annehmen, daß $b(x_i, x_i) \neq 0$ für $i = 1, \dots, r$ und $b(x_i, x_i) = 0$ für $i = r + 1, \dots, n$ für ein geeignetes r gilt. Wähle nun T als die Matrix, deren Spalten die Vektoren in B sind, dann ist $M_B(b) = T^t \circ A \circ T$ und hat die gewünschte Gestalt.

Es bleibt zu zeigen, daß $r = \text{rang}(A)$. Aber, da T invertierbar ist, gilt

$$r = \text{rang}(T^t \circ A \circ T) = \text{rang}(A).$$

□

Bemerkung 36.22

- a. Man beachte, daß die λ_i i. a. nicht nur von A abhängen und auch *nicht* die Eigenwerte von A sind (siehe Beispiel 36.24). Die Anzahl der Diagonalelemente ungleich Null hängt jedoch stets nur von A ab.
- b. Korollar 36.21 führt zu folgender Überlegung. Da $T \in \text{Gl}_n(K)$ ist, ist T das Produkt von Elementarmatrizen $T = P_1 \circ \dots \circ P_k$ und somit gilt

$$D := T^t \circ A \circ T = P_k^t \circ \dots \circ P_1^t \circ A \circ P_1 \circ \dots \circ P_k.$$

Das heißt, daß die Diagonalmatrix D aus A durch gleichzeitiges Durchführen von elementaren Zeilenoperationen und den zugehörigen Spaltenoperationen entsteht. Dabei ist es wegen $P^t \circ (A \circ P) = (P^t \circ A) \circ P$ egal, ob zuerst die Zeilenoperation oder die Spaltenoperation durchgeführt wird.

Die Überführung einer symmetrischen Matrix A in Diagonalgestalt mittels gleichzeitiger Zeilen- und Spaltenoperationen nennt man das *symmetrische Gaußsche Eliminationsverfahren* oder den *symmetrischen Gaußalgorithmus*.

Es ist klar, daß man diesen Algorithmus ebenso einfach implementieren kann, wie den Gaußschen Algorithmus. Will man zusätzlich die Transformationsmatrix T bestimmen, so startet man wie bei der Invertierung einer Matrix mit $(A \mid \mathbb{1}_n)$, führt bei A die Zeilen- und Spaltenoperationen durch, bei $\mathbb{1}_n$ aber nur die Spaltenoperationen. Ist dann A diagonalisiert, so ist $\mathbb{1}_n$ in die Transformationsmatrix überführt.

Wir formulieren den Algorithmus nun in rekursiver Form. Die Eingabe muß dann ein Schema der Form $(A \mid \mathbb{1}_n)$ sein, damit die Ausgabe den gewünschten Erfolg hat.

Algorithmus 36.23 (Symmetrischer Gaußalgorithmus)

INPUT: A, T mit $A \in \text{Mat}_n(\mathbb{K})$ symmetrisch und $T \in \text{GL}_r(\mathbb{K})$, $r \geq n$.

OUTPUT: $T \in \text{GL}_r(\mathbb{K})$ so, daß $\tilde{T}^t \circ A \circ \tilde{T}$ eine Diagonalmatrix ist, wobei \tilde{T} durch Streichen der ersten $r - n$ Spalten und Zeilen aus T entsteht.

1. **Schritt:** Setze $m = r - n$.
2. **Schritt:** Man suche in der ersten Spalte von A den ersten Eintrag, der nicht Null ist. Existiert ein solcher, merke man sich die Zeilennummer z , sonst gehe man zu Schritt 5.
3. **Schritt:** Ist $z \neq 1$, so addiere die z -te Zeile von A zur ersten und die z -te Spalte zur ersten. Addiere ferner die $z + m$ -te Spalte von T zur $m + 1$ -ten Spalte.
4. **Schritt:** Für $k = 2, \dots, n$ addiere man das $-A[1, k]/A[1, 1]$ -fache der ersten Zeile von A zur k -ten und das $-A[1, k]/A[1, 1]$ -fache der ersten Spalte zur k -ten. Sodann addiere man das $-A[1, k]/A[1, 1]$ -fache der $1 + m$ -ten Spalte von T zur $k + m$ -ten.
5. **Schritt:** Falls $n > 1$, dann erzeuge man eine Matrix B , indem man aus A die erste Zeile und die erste Spalte streicht. Sodann rufe man die Prozedur mit den Parametern B und T auf und speichere das Ergebnis in T .
6. **Schritt:** Man gebe T zurück.

Beispiel 36.24

Wir betrachten die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \text{Mat}_n(\mathbb{R}).$$

Sodann bilden wir das Schema $(A \mid \mathbb{1}_2)$ und wenden den symmetrischen Gaußalgorithmus an:

$$\left(\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array} \right) \xrightarrow{I=I+II} \left(\begin{array}{cc|cc} 3 & 2 & 1 & 0 \\ 2 & 1 & 1 & 1 \end{array} \right) \xrightarrow{II=II-\frac{2}{3}I} \left(\begin{array}{cc|cc} 3 & 0 & 1 & -\frac{2}{3} \\ 0 & -\frac{1}{3} & 1 & \frac{1}{3} \end{array} \right).$$

Für $T = \begin{pmatrix} 1 & -\frac{2}{3} \\ 1 & \frac{1}{3} \end{pmatrix} \in \text{Gl}_2(\mathbb{R})$ gilt also

$$T^t \circ A \circ T = \begin{pmatrix} 3 & 0 \\ 0 & -\frac{1}{3} \end{pmatrix}.$$

Alternativ können wir auch wie folgt vorgehen:

$$\left(\begin{array}{cc|cc} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array} \right) \xrightarrow{I \leftrightarrow II} \left(\begin{array}{cc|cc} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{array} \right) \xrightarrow{II=II-I} \left(\begin{array}{cc|cc} 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & -1 \end{array} \right).$$

Dann gilt für $S = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \in \text{Gl}_2(\mathbb{R})$

$$S^t \circ A \circ S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Man beachte, daß die beiden Diagonalmatrizen nicht die gleichen Diagonaleinträge besitzen!

C) Der Sylvestersche Trägheitssatz

Über dem Körper $K = \mathbb{R}$ der reellen Zahlen können wir noch eine etwas schönere Normalform für Bilinearformen herleiten, bei der nicht nur die Zahl der Nicht-Null-Einträge auf der Diagonalen invariant ist.

Korollar 36.25 (Sylvesterscher Trägheitssatz für Bilinearformen über \mathbb{R})

Sei V ein \mathbb{R} -Vektorraum der Dimension $n > 0$ und $b \in \text{Bil}_{\mathbb{R}}(V)$ eine symmetrische Bilinearform. Dann besitzt V eine Orthogonalbasis B , so daß

$$M_B(b) = \left(\begin{array}{c|c|c} \mathbf{1}_k & 0 & 0 \\ 0 & -\mathbf{1}_l & 0 \\ \hline 0 & 0 & 0 \end{array} \right).$$

Dabei hängen die Zahlen k und l nur von b ab, nicht von der Orthogonalbasis B . Wir nennen k den Trägheitsindex, l den Morseindex und $k - l$ die Signatur von b .

Beweis: Wir wählen zunächst eine Basis $D = (y_1, \dots, y_n)$ wie in Satz 36.20, d.h. $M_D(b)$ ist eine Diagonalmatrix mit Diagonaleinträgen $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Dabei können wir nach Ummumerieren der y_i ohne Einschränkung annehmen, daß $\lambda_1, \dots, \lambda_k > 0$, $\lambda_{k+1}, \dots, \lambda_{k+l} < 0$ und $\lambda_{k+l+1}, \dots, \lambda_n = 0$ für geeignete $k, l \in \mathbb{N}$ gilt. Setzen wir nun

$$x_i := \begin{cases} \frac{1}{\sqrt{|\lambda_i|}} \cdot y_i, & i = 1, \dots, k+l, \\ y_i, & i = k+l+1, \dots, n \end{cases}$$

und $B = (x_1, \dots, x_n)$, dann hat $M_B(b)$ die gewünschte Gestalt, da nach wie vor $b(x_i, x_j) = 0$ für $i \neq j$ und

$$b(x_i, x_i) = \begin{cases} 1, & i = 1, \dots, k, \\ -1, & i = k+1, \dots, k+l, \\ 0, & i = k+l+1, \dots, n. \end{cases}$$

Es bleibt, zu zeigen, daß die Zahlen k und l unabhängig von der Wahl von B sind. Dazu zeigen wir zunächst, daß

$$k = \max \{ \dim_{\mathbb{K}}(\mathbf{U}) \mid \mathbf{U} \leq \mathbf{V}, q_b(x) > 0 \forall 0 \neq x \in \mathbf{U} \}. \quad (107)$$

Ist $0 \neq x = \sum_{i=1}^k \lambda_i x_i \in \text{Lin}(x_1, \dots, x_k)$, so gilt

$$q_b(x) = b(x, x) = \sum_{i,j=1}^k \lambda_i \cdot \lambda_j \cdot b(x_i, x_j) = \sum_{i=1}^k \lambda_i^2 > 0. \quad (108)$$

Mithin ist $\text{Lin}(x_1, \dots, x_k)$ einer der Unterräume von \mathbf{V} , die auf der rechten Seite betrachtet werden, und mithin ist das Maximum mindestens k .

Sei nun $\mathbf{U} \leq \mathbf{V}$ irgendein Unterraum von \mathbf{V} mit $q_b(x) > 0$ für alle $0 \neq x \in \mathbf{U}$. Für ein beliebiges $x \in \mathbf{W} := \text{Lin}(x_{k+1}, \dots, x_n)$ zeigt man wie in (108), daß $q_b(x) \leq 0$ gilt. Daraus folgt unmittelbar, daß

$$\mathbf{U} \cap \mathbf{W} = \{0\}$$

gelten muß, und dann folgt

$$\dim_{\mathbb{R}}(\mathbf{U}) = \dim_{\mathbb{R}}(\mathbf{U} + \mathbf{W}) - \dim_{\mathbb{R}}(\mathbf{W}) + \dim_{\mathbb{R}}(\mathbf{U} \cap \mathbf{W}) \leq n - (n - k) = k.$$

Damit ist (107) gezeigt und k hängt allein von b ab.

Zudem gilt für jede Basis C von \mathbf{V}

$$k + l = \text{rang}(M_B(b)) = \text{rang}\left((T_C^B)^t \circ M_C(b) \circ T_C^B\right) = \text{rang}(M_C(b)),$$

so daß auch $k + l$ nicht von der Wahl der Orthogonalbasis B abhängt, aber dann trifft dies auch auf die Differenz $l = (k + l) - k$ zu. \square

Korollar 36.26 (Sylvesterscher Trägheitssatz für symmetrische Matrizen über \mathbb{R})
Sei $A \in \text{Mat}_n(\mathbb{R})$ eine symmetrische Matrix. Dann existiert eine invertierbare Matrix $T \in \text{Gl}_n(\mathbb{R})$, so daß

$$T^t \circ A \circ T = \left(\begin{array}{c|c|c} \mathbb{1}_k & 0 & 0 \\ \hline 0 & -\mathbb{1}_l & 0 \\ \hline 0 & 0 & 0 \end{array} \right).$$

Dabei hängen die Zahlen k und l nur von A ab und nicht von T .

Wir nennen k den Trägheitsindex, l den Morseindex und $k - l$ die Signatur von A .

Beweis: Dies folgt, indem wir Korollar 36.25 auf $b = b_A$ anwenden und $T = T_E^B$ setzen, wobei B die Orthogonalbasis aus Korollar 36.25 ist und E die kanonische Basis von \mathbb{R}^n . \square

Beispiel 36.27

In Beispiel 36.24 haben wir die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \in \text{Mat}_n(\mathbb{R})$$

auf zwei mögliche Weisen als Bilinearform zu einer Diagonalmatrix transformiert. Bei der ersten Möglichkeit haben wir die Matrix

$$T = \begin{pmatrix} 1 & -\frac{2}{3} \\ 1 & \frac{1}{3} \end{pmatrix} \in \text{Gl}_2(\mathbb{R})$$

verwendet und

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

erhalten. Unser Beweis legt nahe, die erste Spalte von T durch $\sqrt{3}$ zu dividieren und die zweite mit $\sqrt{3}$ zu multiplizieren. Wir erhalten dann

$$\tilde{T} = \begin{pmatrix} \frac{1}{\sqrt{3}} & -\frac{2}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{pmatrix}$$

und

$$\tilde{T}^t \circ A \circ \tilde{T} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

A hat also den Trägheitsindex 1, den Morseindex 1 und die Signatur 0.

Die zweite Transformation von A mittels der Matrix

$$S = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \in \text{Gl}_2(\mathbb{R})$$

hat gleich zur Normalform

$$S^t \circ A \circ S = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

geführt. Wir sehen, daß die Transformationsmatrix nicht eindeutig bestimmt ist.

D) Sesquilinearformen

Wenn der Grundkörper der Körper der komplexen Zahlen ist, dann kann man die Bedingung der Linearität einer Bilinearform in der zweiten Komponente verändern und kommt zum Begriff der Sesquilinearform. Diese sind im Zusammenhang mit geometrischen Begriffen wie Länge und Winkel wichtig.

Definition 36.28 (Sesquilinearformen)

Es sei V ein \mathbb{C} -Vektorraum und $\bar{\cdot} : \mathbb{C} \rightarrow \mathbb{C}$ bezeichne die komplexe Konjugation.

- a. Eine Abbildung

$$\mathbf{b} : V \times V \rightarrow \mathbb{C}$$

heißt *sesquilinear*⁴ oder eine *Sesquilinearform*, falls für $x, y, z \in V$ und $\lambda, \mu \in \mathbb{C}$

$$\mathbf{b}(\lambda x + \mu y, z) = \bar{\lambda} \cdot \mathbf{b}(x, z) + \bar{\mu} \cdot \mathbf{b}(y, z)$$

gilt sowie

$$\mathbf{b}(z, \lambda x + \mu y) = \lambda \cdot \mathbf{b}(z, x) + \mu \cdot \mathbf{b}(z, y).$$

$\text{Sesq}(V)$ bezeichnet den \mathbb{C} -Vektorraum aller Sesquilinearformen auf V .

- b. Ist $B = (x_1, \dots, x_n)$ eine Basis und \mathbf{b} eine Sesquilinearform von V , so heißt

$$M_B(\mathbf{b}) = (\mathbf{b}(x_i, x_j))_{1 \leq i, j \leq n} \in \text{Mat}_n(\mathbb{C})$$

die *Matrixdarstellung* von \mathbf{b} bezüglich der Basis B .

- c. Für eine Matrix $A = (a_{ij})_{i,j} \in \text{Mat}_n(\mathbb{C})$ heißt die Matrix

$$A^* = \bar{A}^t = (\bar{a}_{ji})_{i,j} \in \text{Mat}_n(\mathbb{C})$$

die *adjungierte Matrix* zu A .

Beispiel 36.29 (Sesquilinearformen)

Ist $A \in \text{Mat}_n(\mathbb{C})$, so wird durch

$$\mathbf{b}_A^s : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C} : (x, y) \mapsto \bar{x}^t \circ A \circ y = x^* \circ A \circ y$$

eine Sesquilinearform auf \mathbb{C}^n definiert.

Proposition 36.30 (Matrixdarstellung einer Sesquilinearform)

Es sei V ein \mathbb{C} -Vektorraum mit Basen $B = (x_1, \dots, x_n)$ und $D = (y_1, \dots, y_n)$.

- a. Zu jedem $A \in \text{Mat}_n(\mathbb{C})$ gibt es genau ein $\mathbf{b} \in \text{Sesq}(V)$ mit $M_B(\mathbf{b}) = A$.
 b. Ist \mathbf{b} eine Sesquilinearform auf V und sind $x, y \in V$, so gilt

$$\mathbf{b}(x, y) = M_B(x)^* \circ M_B(\mathbf{b}) \circ M_B(y).$$

- c. Ist \mathbf{b} eine Sesquilinearform auf V , so gilt

$$M_B(\mathbf{b}) = (T_D^B)^* \circ M_D(\mathbf{b}) \circ T_D^B.$$

Beweis: Der Beweis geht analog zu den entsprechenden Beweisen für Bilinearformen, siehe Proposition 36.7 und Satz 36.11. □

Definition 36.31 (Hermitesche Sesquilinearformen)

- a. Eine Sesquilinearform \mathbf{b} auf einem \mathbb{C} -Vektorraum V heißt *hermitesch*, falls

$$\mathbf{b}(x, y) = \overline{\mathbf{b}(y, x)}$$

für alle $x, y \in V$ gilt.

- b. Eine Matrix $A \in \text{Mat}_n(\mathbb{C})$ heißt *hermitesch* oder *selbstadjungiert*, falls $A = A^*$.

⁴Sesquilinear bedeutet ein-einhalb-fach linear. Dies bezieht sich darauf, daß die Abbildung in der ersten Komponente nur die eine Hälfte der Linearitätsbedingung erfüllt.

Beispiel 36.32

Die Matrix

$$A = \begin{pmatrix} 1 & i \\ -i & 0 \end{pmatrix} \in \text{Mat}_2(\mathbb{C})$$

ist hermitesch, da

$$A^* = \overline{A}^t = \begin{pmatrix} \overline{1} & \overline{i} \\ \overline{-i} & \overline{0} \end{pmatrix}^t = \begin{pmatrix} 1 & -i \\ i & 0 \end{pmatrix}^t = A.$$

Proposition 36.33 (Hermitesche Sesquilinearformen)

Sei V ein \mathbb{C} -Vektorraum mit Basis $B = (x_1, \dots, x_n)$, $b \in \text{Sesq}(V)$ und $A \in \text{Mat}_n(\mathbb{C})$.

- b ist genau dann hermitesch, wenn $M_B(b)$ hermitesch ist.
- Die Sesquilinearform b_A^s ist genau dann hermitesch, wenn A hermitesch ist.

Beweis: Der Beweis geht analog zur entsprechenden Aussage für Bilinearformen, siehe Proposition 36.16. □

Bemerkung 36.34

- Ist $A \in \text{Mat}_n(\mathbb{R}) \subseteq \text{Mat}_n(\mathbb{C})$, dann gilt

$$A^* = \overline{A}^t = A^t.$$

Insbesondere, A ist genau dann hermitesch, wenn A symmetrisch ist.

- Ist b eine hermitesche Sesquilinearform auf dem \mathbb{C} -Vektorraum V , so gilt

$$b(x, x) = \overline{b(x, x)}$$

für alle $x \in V$. Das geht aber nur, wenn

$$b(x, x) \in \mathbb{R}$$

stets eine reelle Zahl ist!

Man bezeichnet die Abbildung

$$q_b : V \rightarrow \mathbb{R} : x \mapsto b(x, x)$$

dann als die zu b gehörende *quadratische Form*, und man prüft leicht nach, daß für $x, y \in V$ stets

$$b(x, y) = \frac{1}{4}(q_b(x+y) - q_b(x-y) + iq_b(x+iy) - iq_b(x-iy))$$

gilt, so daß die quadratische Form die Sesquilinearform b eindeutig bestimmt.

E) Definite Bilinearformen und Sesquilinearformen

Im folgenden beschränken wir uns auf \mathbb{R} - und \mathbb{C} -Vektorräume, da wir für die Werte $b(x, x)$ einer Bilinearform bzw. einer Sesquilinearform entscheiden müssen, ob sie positiv oder negativ sind.

Definition 36.35 (Definitheit)

- a. Eine symmetrische Bilinearform $b \in \text{Bil}_{\mathbb{R}}(V)$ heißt *positiv definit*, falls

$$b(x, x) > 0$$

für alle $0 \neq x \in V$ gilt. Sie heißt *negativ definit*, falls stattdessen

$$b(x, x) < 0$$

für alle $0 \neq x \in V$ gilt. Und sie heißt schließlich *indefinit*, falls es $x, y \in V$ gibt mit

$$b(x, x) > 0 > b(y, y).$$

- b. Eine hermitesche Sesquilinearform $b \in \text{Sesq}(V)$ heißt *positiv definit*, falls

$$b(x, x) > 0$$

für alle $0 \neq x \in V$ gilt. Sie heißt *negativ definit*, falls stattdessen

$$b(x, x) < 0$$

für alle $0 \neq x \in V$ gilt. Und sie heißt schließlich *indefinit*, falls es $x, y \in V$ gibt mit

$$b(x, x) > 0 > b(y, y).$$

Beispiel 36.36

- a. Die hermitesche Sesquilinearform

$$b_{\mathbb{1}_n}^s : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C}^n : (x, y) \mapsto \bar{x}^t \circ y = \sum_{i=1}^n \bar{x}_i \cdot y_i$$

ist positiv definit, da

$$b_{\mathbb{1}_n}^s(x, x) = \sum_{i=1}^n \bar{x}_i \cdot x_i = \sum_{i=1}^n |x_i|^2 > 0$$

für alle $0 \neq x = (x_1, \dots, x_n) \in \mathbb{C}^n$.

- b. Für

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \in \text{Mat}_2(\mathbb{R})$$

ist die Bilinearform b_A auf \mathbb{R}^2 symmetrisch. Da für $x = (x_1, x_2)^t \in \mathbb{R}^2$ mit $x \neq (0, 0)$ ferner gilt

$$b_A(x, x) = x_1^2 + 2x_1x_2 + 4x_2^2 = (x_1 + x_2)^2 + 3x_2^2 > 0$$

ist b_A zudem positiv definit.

Bemerkung 36.37

Im allgemeinen ist die Bedingung der positiven Definitheit durchaus nicht einfach nachzuprüfen, da man meist nicht alle Vektoren $0 \neq x \in V$ überprüfen kann. Man beachte auch, daß es nicht reicht, etwa für eine Basis $B = (x_1, \dots, x_n)$ von V nachzuprüfen, daß $\langle x_i, x_i \rangle > 0$ für alle $i = 1, \dots, n$ gilt.

Betrachte dazu die folgende symmetrische Bilinearform auf \mathbb{R}^2

$$\langle \cdot, \cdot \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R} : ((a_1, a_2)^t, (b_1, b_2)^t) \mapsto a_1 b_1 - a_2 b_2$$

sowie die Basis $(x_1, x_2) = ((1, 0)^t, (2, 1)^t)$. Dann gilt $\langle x_1, x_1 \rangle = 1 > 0$ und $\langle x_2, x_2 \rangle = 3 > 0$, aber $\langle e_2, e_2 \rangle = -1 < 0$.

Wir werden in Satz 38.39 Kriterien kennenlernen, die es uns erlauben, positive Definitheit zu entscheiden. Diese formuliert man dann für symmetrische und hermitesche Matrizen (siehe auch Definition 38.36).

Aufgaben**Aufgabe 36.38**

Es sei $b : K^2 \times K^2 \rightarrow K : ((x_1, x_2)^t, (y_1, y_2)^t) \mapsto 2 \cdot x_1 \cdot y_1 + x_1 \cdot y_2 + y_1 \cdot x_2 - x_2 \cdot y_2$. Ferner bezeichne $E = (e_1, e_2)$ die kanonische Basis des K^2 und $B = (x_1, x_2)$ mit $x_1 = (1, 1)^t$ und $x_2 = (1, -1)^t$ sei eine weitere Basis.

Zeige, daß b eine symmetrische Bilinearform ist und berechne die Matrixdarstellungen $M_E(b)$ und $M_B(b)$ sowie die Transformationsmatrix T_E^B mit

$$M_B(b) = (T_E^B)^t \cdot M_E(b) \cdot T_E^B.$$

Aufgabe 36.39

Bestimme für die folgende symmetrische Matrix $A \in \text{Mat}_4(\mathbb{R})$ eine Transformationsmatrix $T \in \text{Gl}_4(\mathbb{R})$, so daß $T^t \circ A \circ T$ eine Diagonalmatrix wie im Sylvesterschen Trägheitssatz ist, und bestimme den Trägheitsindex, den Morseindex, die Signatur und den Rang von A :

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

Aufgabe 36.40

Sei K ein Körper mit $\text{char}(K) \neq 2$ und $b \in \text{Bil}_K(V)$.

- Zeige, es gibt eine symmetrische Bilinearform $b' \in \text{Bil}_K(V)$ und eine schiefsymmetrische Bilinearform $b'' \in \text{Bil}_K(V)$, so daß $b = b' + b''$.
- Zeige, die Bilinearformen b' und b'' in a. sind eindeutig bestimmt.
- Gelten die Aussagen in a. und b. auch noch, wenn $\text{char}(K) = 2$?

Anmerkung, b heißt *schiefsymmetrisch*, wenn $b(x, y) = -b(y, x)$ für alle $x, y \in V$.

§ 37 Euklidische und unitäre Räume

Im folgenden sei stets $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ einer der beiden Körper \mathbb{R} oder \mathbb{C} .

Zur Motivation beginnen wir den Abschnitt mit einigen Überlegungen zur euklidischen Geometrie in der reellen Ebene \mathbb{R}^2 .

Wir definieren uns zunächst zwei Abbildungen

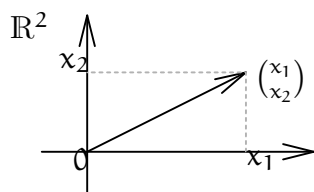
$$\|\cdot\| : \mathbb{R}^2 \rightarrow \mathbb{R}_{\geq 0},$$

die einem Vektor $\mathbf{x} = (x_1, x_2)^t \in \mathbb{R}^2$ seine Länge $\|\mathbf{x}\|$ zuordnet, sowie

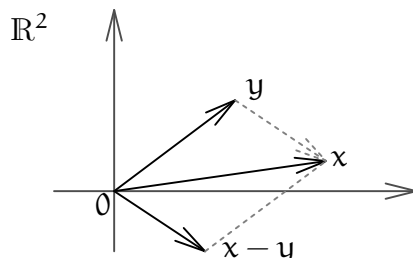
$$d : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}_{\geq 0},$$

die zwei Punkten $\mathbf{x} \in \mathbb{R}^2$ und $\mathbf{y} \in \mathbb{R}^2$ ihren Abstand $d(\mathbf{x}, \mathbf{y})$ zuweist.

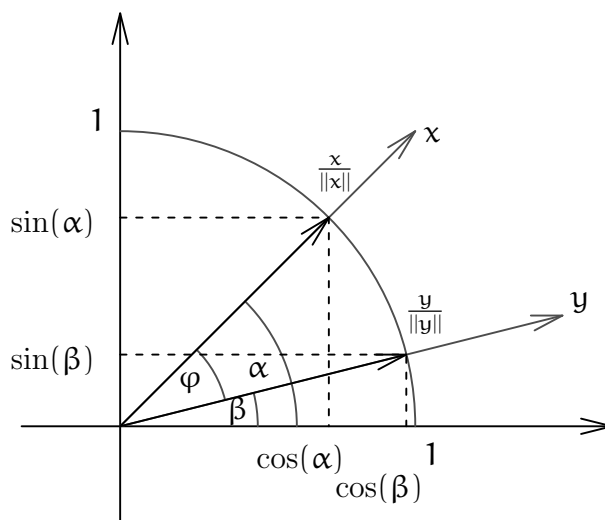
Der Satz von Pythagoras liefert dann $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2}$.



Wir nennen $\|\mathbf{x}\|$ auch die *Norm* des Vektors \mathbf{x} . Da der Abstand der Punkte $\mathbf{x} = (x_1, x_2)^t$ und $\mathbf{y} = (y_1, y_2)^t$ gerade die Länge des Vektors $\mathbf{x} - \mathbf{y}$ ist, folgt somit $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$.



Mit Hilfe der Norm können wir - nach einigen geometrischen Überlegungen - auch den Winkel $\angle(\mathbf{x}, \mathbf{y})$, den zwei Vektoren \mathbf{x} und \mathbf{y} miteinander einschließen, bestimmen.



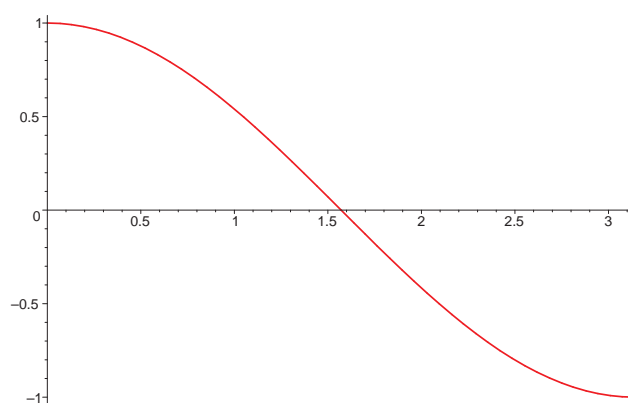
Dazu gehen wir zunächst zu den normierten Vektoren $\frac{x}{\|x\|}$ und $\frac{y}{\|y\|}$ über, die beide die Länge eins haben, wobei wir $x \neq 0$ und $y \neq 0$ voraussetzen. Mit den Bezeichnungen in der Skizze gilt dann

$$\angle(x, y) = \angle\left(\frac{x}{\|x\|}, \frac{y}{\|y\|}\right) = \alpha - \beta = \varphi.$$

Um φ selbst (im Bogenmaß) auszudrücken, müßte man die Länge des Kreisbogens zwischen $\frac{x}{\|x\|}$ und $\frac{y}{\|y\|}$ messen, also einer gekrümmten Linie. Dazu greifen wir auf unsere Analysiskenntnisse zurück.

Zur anschaulichen Herleitung des Winkels φ mit $0 \leq \varphi \leq \pi$, benötigen wir nur, daß die Funktion

$$\cos : [0, \pi] \rightarrow \mathbb{R} : \varphi \mapsto \cos(\varphi)$$



injektiv ist. Also reicht es, $\cos(\varphi)$ zu kennen, um den Winkel φ eindeutig beschrieben zu haben. Unter Zuhilfenahme der obigen Skizze und des Additionstheorems 12.38 für den Cosinus erhalten wir

$$\cos(\varphi) = \cos(\alpha - \beta) = \cos(\alpha) \cos(\beta) + \sin(\alpha) \sin(\beta) = \frac{x_1 y_1 + x_2 y_2}{\|x\| \cdot \|y\|}.$$

Dies führt zur Definition einer weiteren Abbildung

$$\langle \cdot, \cdot \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R} : (\mathbf{x}, \mathbf{y}) = ((x_1, x_2)^t, (y_1, y_2)^t) \mapsto \langle \mathbf{x}, \mathbf{y} \rangle := x_1 y_1 + x_2 y_2,$$

welche wir *Skalarprodukt* nennen. Mit deren Hilfe erhalten wir

$$\cos(\varphi) = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}$$

oder alternativ

$$\angle(\mathbf{x}, \mathbf{y}) = \varphi = \arccos \left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} \right).$$

Wir sind also mittels recht einfacher Abbildungen in der Lage, Längen und Winkel auszudrücken. Dieses Beispiel motiviert die folgenden Begriffsbildungen.

A) Skalarprodukte

Definition 37.1 (Skalarprodukt)

Es sei V ein \mathbb{K} -Vektorraum. Eine Abbildung

$$\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{K}$$

heißt ein *Skalarprodukt* auf V , falls für alle $\lambda, \mu \in \mathbb{K}$ und $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ gilt:

- (1) $\langle \mathbf{x}, \lambda \mathbf{y} + \mu \mathbf{z} \rangle = \lambda \cdot \langle \mathbf{x}, \mathbf{y} \rangle + \mu \cdot \langle \mathbf{x}, \mathbf{z} \rangle$,
- (2) $\langle \mathbf{x}, \mathbf{y} \rangle = \overline{\langle \mathbf{y}, \mathbf{x} \rangle}$,
- (3) $\langle \mathbf{x}, \mathbf{x} \rangle \in \mathbb{R}$ und $\langle \mathbf{x}, \mathbf{x} \rangle > 0$ für $\mathbf{x} \neq 0$.

Bemerkung 37.2 (Skalarprodukte als Bilinear- bzw. Sesquilinearformen)

Es sei V ein \mathbb{K} -Vektorraum.

- a. Ist $\mathbb{K} = \mathbb{R}$, so ist $\langle \cdot, \cdot \rangle$ genau dann ein Skalarprodukt, wenn $\langle \cdot, \cdot \rangle$ eine positiv-definite, symmetrische Bilinearform auf V ist.
- b. Ist $\mathbb{K} = \mathbb{C}$, so ist $\langle \cdot, \cdot \rangle$ genau dann ein Skalarprodukt, wenn $\langle \cdot, \cdot \rangle$ eine positiv-definite, hermitesche Sesquilinearform auf V ist.

Definition 37.3 (Euklidische und unitäre Räume)

Es sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf den \mathbb{K} -Vektorraum V .

- a. Ist $\mathbb{K} = \mathbb{R}$, so nennen wir das Quadrupel $(V, +, \cdot, \langle \cdot, \cdot \rangle)$ einen *euklidischen Raum*.
- b. Ist $\mathbb{K} = \mathbb{C}$, so nennen wir das Quadrupel $(V, +, \cdot, \langle \cdot, \cdot \rangle)$ einen *unitären Raum*.

Wir werden meist nur V statt $(V, +, \cdot, \langle \cdot, \cdot \rangle)$ schreiben.

Beispiel 37.4 (Skalarprodukte)

- a. Wir nennen

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} : (\mathbf{x}, \mathbf{y}) \mapsto \mathbf{x}^t \circ \mathbf{y} = \sum_{i=1}^n x_i \cdot y_i$$

das *kanonische Skalarprodukt* oder *Standardskalarprodukt* auf \mathbb{R}^n . $(\mathbb{R}^n, +, \cdot, \langle \cdot, \cdot \rangle)$ ist ein euklidischer Raum.

b. Analog definieren wir das *kanonische Skalarprodukt* auf \mathbb{C}^n durch

$$\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C} : (\mathbf{x}, \mathbf{y}) \mapsto \bar{\mathbf{x}}^t \circ \mathbf{y} = \mathbf{x}^* \circ \mathbf{y} = \sum_{i=1}^n \bar{x}_i \cdot y_i,$$

und $(\mathbb{C}^n, +, \cdot, \langle \cdot, \cdot \rangle)$ ist ein unitärer Raum.

c. Die positiv-definite symmetrische Bilinearform in Beispiel 36.36 ist ein Skalarprodukt auf \mathbb{R}^2 .

d. Sei $V = \mathcal{C}([0, 1], \mathbb{R})$ der \mathbb{R} -Vektorraum der auf dem Intervall $[0, 1]$ stetigen Funktionen. Dann wird durch

$$\langle f, g \rangle := \int_0^1 f(x)g(x)dx \in \mathbb{R}$$

für $f, g \in V$ ein Skalarprodukt definiert (siehe Aufgabe 37.33) und $(\mathcal{C}([0, 1], \mathbb{R}), +, \cdot, \langle \cdot, \cdot \rangle)$ ist ein euklidischer Raum.

Bemerkung 37.5

Wenn wir im Folgenden den Fall eines euklidischen und eines unitären Raumes parallel behandeln wollen, dann werden wir uns häufig zunutze machen, daß für eine reelle Zahl λ gilt $\lambda = \bar{\lambda}$. Mithin sind auf einem *reellen* Vektorraum V die Bedingungen

$$\langle \lambda \mathbf{x} + \mu \mathbf{y}, \mathbf{z} \rangle = \lambda \langle \mathbf{x}, \mathbf{z} \rangle + \mu \langle \mathbf{y}, \mathbf{z} \rangle$$

oder

$$\langle \lambda \mathbf{x} + \mu \mathbf{y}, \mathbf{z} \rangle = \bar{\lambda} \langle \mathbf{x}, \mathbf{z} \rangle + \bar{\mu} \langle \mathbf{y}, \mathbf{z} \rangle.$$

gleichwertig, und für eine reelle Matrix $A \in \text{Mat}(n, \mathbb{R})$ gilt genau dann $A = A^t$, wenn $A = \bar{A}^t$ erfüllt ist. Dies erspart uns viele Fallunterscheidungen!

B) Die euklidische Norm zu einem Skalarprodukt

Definition 37.6 (Normierter Raum)

Es sei V ein \mathbb{K} -Vektorraum. Eine Abbildung

$$\|\cdot\| : V \rightarrow \mathbb{R}_{\geq 0},$$

heißt eine *Norm* auf V , falls für alle $\mathbf{x}, \mathbf{y} \in V$ und $\lambda \in \mathbb{K}$

$$(1) \quad \|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}, \quad (\text{“Positive Definitheit”})$$

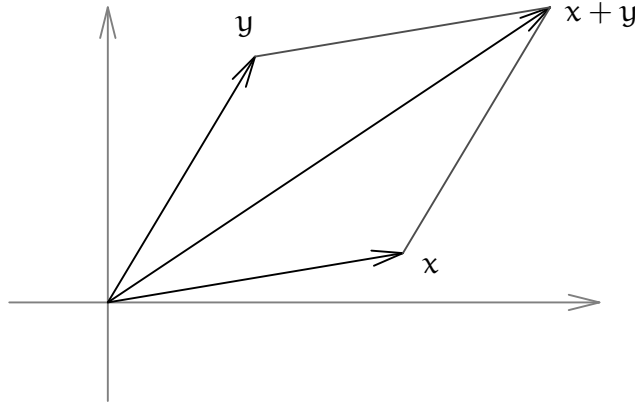
$$(2) \quad \|\lambda \mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|, \text{ und} \quad (\text{“Homogenität”})$$

$$(3) \quad \|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|. \quad (\text{“Dreiecksungleichung”})$$

Das Quadrupel $(V, +, \cdot, \|\cdot\|)$ heißt dann ein *normierter Raum*.

Bemerkung 37.7 (Norm als Längenmaß)

Wir erinnern uns, daß eine Norm die Länge von Vektoren sowie Abstände messen soll. Bedingung (1) kann dann so interpretiert werden, daß jeder Vektor eine nicht-negative Länge hat und daß nur der Nullvektor die Länge null hat. Bedingung (2) bedeutet, daß die Streckung eines Vektors um den Faktor λ seine Länge um $|\lambda|$ strecken möge. Und Bedingung (3) kann dahingehend interpretiert werden, daß der Weg vom Ursprung über den Punkt x hin zum Punkt $x + y$ unter gar keinen Umständen kürzer ist, als der direkte Weg vom Ursprung zum Punkt $x + y$.



Diese Forderungen scheinen allesamt für eine Funktion, die die Länge von Vektoren beziehungsweise Abstände von Punkten messen soll, nicht unbillig. Und in der Tat reichen diese Forderungen auch bereits aus, um einen vernünftigen Längenbegriff zu erhalten.

Satz 37.8 (Cauchy-Schwarzsche Ungleichung)

Ist V ein euklidischer oder unitärer Raum, dann gilt für alle $x, y \in V$

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle} \cdot \sqrt{\langle y, y \rangle}, \quad (109)$$

zudem gilt die Gleichheit genau dann, wenn x und y linear abhängig sind.

Beweis: Für $x = 0$ oder $y = 0$ ist die Aussage offensichtlich richtig. Wir können also $x, y \neq 0$ annehmen. Dann gilt für $\lambda \in \mathbb{K}$

$$0 \leq \langle x - \lambda y, x - \lambda y \rangle = \langle x, x \rangle - \overline{\lambda} \overline{\langle x, y \rangle} - \lambda \langle x, y \rangle + \overline{\lambda} \lambda \langle y, y \rangle. \quad (110)$$

Wählen wir nun speziell $\lambda = \frac{\langle x, y \rangle}{\langle y, y \rangle} \in \mathbb{K}$, dann folgt

$$\begin{aligned} 0 &\leq \langle x, x \rangle - \frac{\langle x, y \rangle}{\langle y, y \rangle} \cdot \overline{\langle x, y \rangle} - \frac{\overline{\langle x, y \rangle}}{\langle y, y \rangle} \cdot \langle x, y \rangle + \frac{\overline{\langle x, y \rangle}}{\langle y, y \rangle} \frac{\langle x, y \rangle}{\langle y, y \rangle} \cdot \langle y, y \rangle \\ &= \langle x, x \rangle - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} - \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle} + \frac{|\langle x, y \rangle|^2}{\langle y, y \rangle}, \end{aligned}$$

also

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \cdot \langle y, y \rangle. \quad (111)$$

Durch Ziehen der positiven Wurzel folgt die gesuchte Ungleichung (109).

Nun sind x und y genau dann linear abhängig, wenn es ein $\lambda \in \mathbb{K}$ gibt, für das $x = \lambda y$ gilt. Das wiederum ist wegen der positiven Definitheit von $\langle \cdot, \cdot \rangle$ gleichbedeutend

dazu, daß es ein $\lambda \in \mathbb{K}$ gibt, für das in (110) das Gleichheitszeichen gilt. Dieses λ ist eindeutig bestimmt, und erfüllt

$$\lambda = \lambda \cdot \frac{\langle \mathbf{y}, \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle} = \frac{\langle \mathbf{y}, \lambda \mathbf{y} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle} = \frac{\langle \mathbf{y}, \mathbf{x} \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle} = \frac{\overline{\langle \mathbf{x}, \mathbf{y} \rangle}}{\langle \mathbf{y}, \mathbf{y} \rangle}.$$

Damit ist die Gleichheit in (110) gleichwertig zur Gleichheit in (111). □

Satz 37.9 (Die euklidische Norm zu einem Skalarprodukt)

Es sei V ein euklidischer oder unitärer Raum. Dann wird durch

$$\|\cdot\| : V \rightarrow \mathbb{R}_{\geq 0} : \mathbf{x} \mapsto \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

eine Norm auf V definiert. Wir werden euklidische und unitäre Räume im folgenden stets mit dieser zugehörigen euklidischen Norm als normierte Räume betrachten, so daß die Cauchy-Schwarzsche Ungleichung die folgende Form hat:

$$|\langle \mathbf{x}, \mathbf{y} \rangle| \leq \|\mathbf{x}\| \cdot \|\mathbf{y}\|.$$

Beweis: Seien $\mathbf{x} \in V$ und $\lambda \in \mathbb{K}$. Aus der positiven Definitheit von $\langle \cdot, \cdot \rangle$ folgt, daß $\langle \mathbf{x}, \mathbf{x} \rangle \geq 0$ und somit $\|\mathbf{x}\|$ definiert und stets nicht-negativ ist. Ferner folgt, daß $\|\mathbf{x}\| = 0$ genau dann gilt, wenn \mathbf{x} der Nullvektor ist. Aus der Bilinearität bzw. Sesquilinearität von $\langle \cdot, \cdot \rangle$ leiten wir her, daß

$$\langle \lambda \mathbf{x}, \lambda \mathbf{x} \rangle = \bar{\lambda} \lambda \langle \mathbf{x}, \mathbf{x} \rangle = |\lambda|^2 \langle \mathbf{x}, \mathbf{x} \rangle,$$

und somit $\|\lambda \mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|$.

Allein, die Dreiecksungleichung ist etwas schwieriger zu zeigen. Wir verwenden hierfür die Cauchy-Schwarzsche Ungleichung aus Satz 37.8. Beachten wir noch, daß für eine komplexe Zahl $c \in \mathbb{C}$ stets

$$\operatorname{Re}(c) = \frac{1}{2} \cdot (c + \bar{c})$$

gilt, so erhalten wir für $\mathbf{x}, \mathbf{y} \in V$

$$\begin{aligned} \|\mathbf{x} + \mathbf{y}\|^2 &= \langle \mathbf{x} + \mathbf{y}, \mathbf{x} + \mathbf{y} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + \langle \mathbf{x}, \mathbf{y} \rangle + \overline{\langle \mathbf{x}, \mathbf{y} \rangle} + \langle \mathbf{y}, \mathbf{y} \rangle \\ &= \|\mathbf{x}\|^2 + 2 \cdot \operatorname{Re}(\langle \mathbf{x}, \mathbf{y} \rangle) + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2 \cdot |\langle \mathbf{x}, \mathbf{y} \rangle| + \|\mathbf{y}\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2 \cdot \|\mathbf{x}\| \cdot \|\mathbf{y}\| + \|\mathbf{y}\|^2 \\ &= (\|\mathbf{x}\| + \|\mathbf{y}\|)^2. \end{aligned}$$

Hieraus folgt dann die Dreiecksungleichung. □

Bemerkung 37.10 (Winkel in euklidischen Räumen)

Die Cauchy-Schwarzsche Ungleichung erlaubt es uns nun, in einem beliebigen *euklidischen Raum* V Winkel zwischen zwei Vektoren zu definieren. Denn aus der

Ungleichung (109) folgt für $0 \neq x, y \in V$

$$-1 \leq \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|} \leq 1. \quad (112)$$

Aus Satz 16.12 wissen wir aber, daß es zu jeder reellen Zahl $-1 \leq r \leq 1$ genau einen Winkel $\alpha \in [0, \pi]$ gibt mit $r = \cos(\alpha)$, nämlich $\alpha = \arccos(r)$. Man definiert deshalb

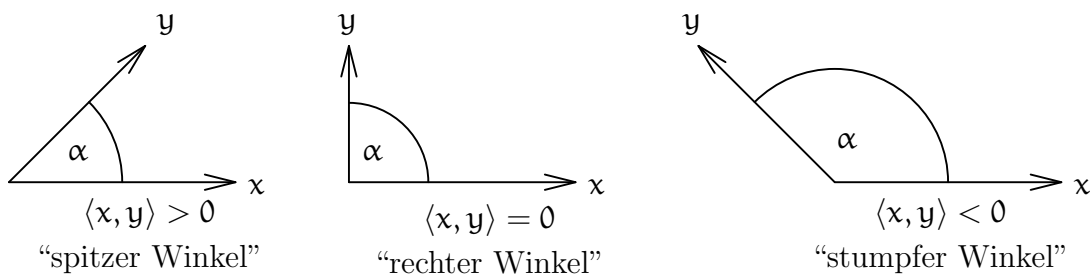
$$\angle(x, y) = \arccos\left(\frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}\right) \in [0, \pi]$$

als den *Winkel* zwischen x und y .

Ist $\langle x, y \rangle > 0$, also $\angle(x, y) \in [0, \frac{\pi}{2}[$, so spricht man von einem *spitzen Winkel*.

Ist $\langle x, y \rangle < 0$, also $\angle(x, y) \in]\frac{\pi}{2}, \pi]$, so spricht man von einem *stumpfen Winkel*.

Ist $\langle x, y \rangle = 0$, also $\angle(x, y) = \frac{\pi}{2}$, so spricht man von einem *rechten Winkel*.



C) Orthonormalbasen und Parsevalsche Gleichung

Definition 37.11 (Orthogonal)

Sei V ein euklidischer oder unitärer Raum, $x, y \in V$, $M, N \subseteq V$ und $U \leq V$.

- x heißt *orthogonal* zu y , falls $\langle x, y \rangle = 0$. Wir schreiben dann $x \perp y$.
- M heißt *orthogonal* zu N , falls $m \perp n$ für alle $m \in M$ und $n \in N$.
Wir schreiben dann $M \perp N$.
- Wir nennen $U^\perp := \{z \in V \mid z \perp U\}$ das *orthogonale Komplement* von U .

Lemma 37.12 (Orthogonales Komplement)

Ist V ein euklidischer oder unitärer Raum und $U \leq V$, dann ist $U^\perp \leq V$.

Beweis: Wegen $0 \in U^\perp$ ist $U^\perp \neq \emptyset$. Sind $x, y \in U^\perp$ und $\lambda, \mu \in \mathbb{K}$, so gilt für $z \in U$

$$\langle \lambda x + \mu y, z \rangle = \bar{\lambda} \langle x, z \rangle + \bar{\mu} \langle y, z \rangle = 0,$$

Also $\lambda x + \mu y \in U^\perp$. Damit ist U^\perp ein Unterraum von V . □

Definition 37.13 (Orthonormalbasis)

Sei V ein euklidischer oder unitärer Raum und $B = (x_i \mid i \in I)$ eine Familie in V .

- B heißt *orthogonal*, falls $x_i \perp x_j$ für alle $i, j \in I$ mit $i \neq j$ gilt.
- B heißt *orthonormal*, falls B orthogonal ist und $\|x_i\| = 1$ für alle $i \in I$ gilt.
- Ist B eine Basis und orthonormal, so heißt B eine *Orthonormalbasis*, kurz *ONB*.

Beispiel 37.14 (ONB)

Betrachten wir \mathbb{K}^n mit dem kanonischen Skalarprodukt, dann ist die kanonische Basis $E = (e_1, \dots, e_n)$ offenbar eine Orthonormalbasis von \mathbb{K}^n , da $\langle e_i, e_j \rangle = \delta_{ij}$ für $i, j \in \{1, \dots, n\}$.

Lemma 37.15 (Orthogonal impliziert linear unabhängig.)

Sei V ein euklidischer oder unitärer Raum und $B = (x_i \mid i \in I)$ eine orthogonale Familie in $V \setminus \{0\}$. Dann ist B linear unabhängig.

Beweis: Aus $\sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i x_i = 0$ folgt für jedes $j \in I$

$$0 = \langle x_j, 0 \rangle = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \langle x_j, x_i \rangle = \lambda_j \langle x_j, x_j \rangle.$$

Da $\langle x_j, x_j \rangle \neq 0$, muß also $\lambda_j = 0$ gelten. □

Proposition 37.16 (Parsevalsche Gleichung)

Sei V ein euklidischer oder unitärer Raum, $B = (x_i \mid i \in I)$ eine ONB und $x \in V$, dann gilt

$$x = \sum_{\substack{i \in I \\ \text{endlich}}} \langle x_i, x \rangle \cdot x_i. \quad (113)$$

Insbesondere sind nur endlich viele $\langle x_i, x \rangle$, $i \in I$, ungleich null.

Beweis: Da die Darstellung $x = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i x_i$ von x als endliche Linearkombination von B eindeutig ist, folgt die Behauptung aus

$$\langle x_j, x \rangle = \left\langle x_j, \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i x_i \right\rangle = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \langle x_j, x_i \rangle = \lambda_j \langle x_j, x_j \rangle = \lambda_j.$$

□

Bemerkung 37.17

Ist B eine ONB von V , so erlaubt es die Gleichung (113), einen Vektor aus V als Linearkombination von B darzustellen, ohne hierzu eigens ein LGS lösen zu müssen, durch simples Einsetzen der Vektoren in das Skalarprodukt. Dieses Verfahren ist sehr effizient und von hoher praktischer Bedeutung. Die Tatsache, daß sich die Koordinaten eines Vektors bezüglich einer ONB mit Hilfe des Skalarproduktes so einfach ausdrücken lassen, spielt aber auch in vielen Beweisen eine Rolle, und ist somit ebenfalls für die Theorie von Bedeutung.

D) Das Orthonormalisierungsverfahren von Gram-Schmidt

Wir beweisen jetzt, daß jeder endlich-dimensionale euklidische bzw. unitäre Raum eine ONB besitzt. Etwas allgemeiner gilt der folgende Satz.

Satz 37.18 (Gram-Schmidt)

Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und B eine orthogonale Familie in V , dann läßt sich B zu einer ONB von V ergänzen.

Beweis: Ist B schon eine Basis von V , so sind wir fertig. Wir dürfen also annehmen, daß B keine Basis und wegen Lemma 37.15 dann auch kein Erzeugendensystem von V ist. Wir zeigen nun konstruktiv, wie wir die orthonormale Familie $B = (z_1, \dots, z_r)$ zu einer orthonormalen Familie (z_1, \dots, z_{r+1}) ergänzen können. Wenden wir dieses Verfahren dann $\dim_{\mathbb{K}}(V) - r$ mal an, so haben wir die Aussage bewiesen.

Dazu wählen wir zunächst einen Vektor x_{r+1} , der linear unabhängig von B ist. Dann setzen wir

$$y_{r+1} := x_{r+1} - \sum_{i=1}^r \langle z_i, x_{r+1} \rangle \cdot z_i. \quad (114)$$

Da x_{r+1} linear unabhängig von B ist, ist $y_{r+1} \neq 0$, und wir können deshalb

$$z_{r+1} := \frac{1}{\|y_{r+1}\|} \cdot y_{r+1} \quad (115)$$

setzen. Dann ist $\|z_{r+1}\| = 1$ und außerdem gilt für $i = 1, \dots, r$

$$\begin{aligned} \langle z_i, z_{r+1} \rangle &= \frac{1}{\|y_{r+1}\|} \cdot \langle z_i, y_{r+1} \rangle \\ &= \frac{1}{\|y_{r+1}\|} \cdot \left(\langle z_i, x_{r+1} \rangle - \sum_{j=1}^r \langle z_j, x_{r+1} \rangle \cdot \langle z_i, z_j \rangle \right) \\ &= \frac{1}{\|y_{r+1}\|} \cdot (\langle z_i, x_{r+1} \rangle - \langle z_i, x_{r+1} \rangle) = 0. \end{aligned}$$

Dann ist aber (z_1, \dots, z_{r+1}) orthonormal und wir sind fertig. \square

Korollar 37.19 (Existenz einer ONB)

Jeder endlich-dimensionale euklidische bzw. unitäre Raum besitzt eine ONB.

Beweis: Wende Satz 37.18 mit $B = \emptyset$ an. \square

Der Beweis von Satz 37.18 ist konstruktiv und wird auch das *Gram-Schmidtsche Orthonormalisierungsverfahren* genannt. Es erlaubt, aus einem gegebenen Erzeugendensystem eine ONB zu konstruieren.

Algorithmus 37.20 (Gram-Schmidtsches Orthonormalisierungsverfahren)

INPUT: $M \subseteq \mathbb{K}^n$ und ein Skalarprodukt $\langle \cdot, \cdot \rangle$ auf \mathbb{K}^n

OUTPUT: ONB B von $\langle M \rangle$

1. **Schritt:** Bestimme eine Basis $B = (x_1, \dots, x_r)$ von $\langle M \rangle$, z. B. mittels Algorithmus 27.20.
2. **Schritt:** Für $i = 1, \dots, r$ führe man folgende Schritte aus:
 - Schritt a.:** berechne die Summe $y_i = x_i - \sum_{j=1}^{i-1} \langle z_j, x_i \rangle \cdot z_j$;
 - Schritt b.:** berechne $z_i = \frac{1}{\|y_i\|} \cdot y_i$;
3. **Schritt:** Gib die veränderte Basis (z_1, \dots, z_r) zurück.

Bemerkung 37.21

- a. Will man in der Praxis ein Skalarprodukt übergeben, so wird man im reellen Fall eine symmetrische Matrix übergeben und im komplexen Fall eine hermitesche. Das Skalarprodukt wird dann gemäß Beispiel 36.2 bzw. Beispiel 36.29 gebildet.
- b. Um zu normieren, ist in Algorithmus 37.20 das Ziehen von Wurzeln notwendig. Verzichtet man jedoch auf die Normierung der Vektoren, so kommt man ohne Wurzelziehen aus. Läßt man im obigen Algorithmus Schritt 2.b. weg und ersetzt dafür in Schritt 2.a. die rechte Seite der Gleichung durch

$$\mathbf{y}_i = \mathbf{x}_i - \sum_{j=1}^{i-1} \frac{\langle \mathbf{y}_j, \mathbf{x}_i \rangle}{\langle \mathbf{y}_j, \mathbf{y}_j \rangle} \cdot \mathbf{y}_j,$$

dann liefert Algorithmus 37.20 eine orthogonale Basis $(\mathbf{y}_1, \dots, \mathbf{y}_r)$ von $\langle M \rangle$. Das hat den Vorteil, daß man exakt rechnen kann - etwa in SINGULAR, wenn die Eingabedaten rationale Zahlen waren.

Beispiel 37.22 (Gram-Schmidtsches Orthonormalisierungsverfahren)

Es sei $B = (\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) = \{(1, 0, 1)^t, (1, 1, 1)^t, (0, 0, 4)^t\} \subseteq \mathbb{R}^3$, wobei wir \mathbb{R}^3 mit dem kanonischen Skalarprodukt versehen betrachten. Man sieht leicht, daß B bereits eine Basis von \mathbb{R}^3 ist. Wir wollen hier B in eine ONB von \mathbb{R}^3 überführen.

Wir setzen nun $\mathbf{y}_1 := (1, 0, 1)^t$, dann ist $\langle \mathbf{y}_1, \mathbf{y}_1 \rangle = 2$ und somit ersetzen wir \mathbf{x}_1 in B durch

$$\mathbf{z}_1 = \frac{1}{\|\mathbf{y}_1\|} \cdot \mathbf{y}_1 = \frac{1}{\sqrt{2}} \cdot \mathbf{y}_1 = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right)^t.$$

Im nächsten Schritt setzen wir

$$\mathbf{y}_2 = \mathbf{x}_2 - \langle \mathbf{z}_1, \mathbf{x}_2 \rangle \cdot \mathbf{z}_1 = (1, 1, 1)^t - \frac{2}{\sqrt{2}} \cdot \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right)^t = (0, 1, 0)^t.$$

Dann ist $\langle \mathbf{y}_2, \mathbf{y}_2 \rangle = 1$ und somit ersetzen wir \mathbf{x}_2 in B durch $\mathbf{z}_2 = \mathbf{y}_2$.

Schließlich bilden wir

$$\begin{aligned} \mathbf{y}_3 &= \mathbf{x}_3 - \langle \mathbf{z}_1, \mathbf{x}_3 \rangle \cdot \mathbf{z}_1 - \langle \mathbf{z}_2, \mathbf{x}_3 \rangle \cdot \mathbf{z}_2 \\ &= (0, 0, 4)^t - \frac{4}{\sqrt{2}} \cdot \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right)^t - 0 \cdot (0, 1, 0)^t \\ &= (-2, 0, 2)^t, \end{aligned}$$

und erhalten $\langle \mathbf{y}_3, \mathbf{y}_3 \rangle = 8$. Somit müssen wir \mathbf{x}_3 durch den Vektor

$$\mathbf{z}_3 = \frac{1}{\|\mathbf{y}_3\|} \cdot \mathbf{y}_3 = \frac{1}{\sqrt{8}} \cdot (-2, 0, 2)^t = \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right)^t$$

ersetzen. Damit ergibt sich aus dem Gram-Schmidtschen Orthonormalisierungsverfahren die ONB

$$\left(\left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right)^t, (0, 1, 0)^t, \left(-\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right)^t \right).$$

E) Orthogonale und unitäre Matrizen

Definition 37.23 (Orthogonale / unitäre Matrizen)

- Eine Matrix $A \in \text{Mat}_n(\mathbb{R})$ heißt *orthogonal*, wenn $A^t \circ A = \mathbb{1}_n$ gilt. Wir nennen $O(n) := \{A \in \text{Mat}_n(\mathbb{R}) \mid A \text{ ist orthogonal}\}$ *orthogonale Gruppe* vom Grad n .
- Eine Matrix $A \in \text{Mat}_n(\mathbb{C})$ heißt *unitär*, wenn $A^* \circ A = \mathbb{1}_n$ gilt, und wir nennen $U(n) := \{A \in \text{Mat}_n(\mathbb{C}) \mid A \text{ ist unitär}\}$ die *unitäre Gruppe* vom Grad n .

Proposition 37.24 (Die Determinante orthogonaler / unitärer Matrizen)

Ist $A \in \text{Mat}_n(\mathbb{K})$ *orthogonal oder unitär*, so gilt $|\det(A)| = 1$.

Beweis: Wegen $A^* \circ A = \mathbb{1}_n$ gilt

$$\begin{aligned} 1 = \det(\mathbb{1}_n) &= \det(A^* \circ A) = \det(A^*) \cdot \det(A) \\ &= \overline{\det(A^t)} \cdot \det(A) = \overline{\det(A)} \cdot \det(A) = |\det(A)|^2. \end{aligned}$$

Daraus folgt $|\det(A)| = 1$. □

Proposition 37.25 (Orthogonale / unitäre Matrizen)

Für eine quadratische Matrix $A \in \text{Mat}_n(\mathbb{K})$ sind äquivalent:

- A ist *orthogonal bzw. unitär*.
- A ist *invertierbar mit $A^* = A^{-1}$* .
- Die Spalten von A sind eine *ONB von \mathbb{K}^n mit kanonischem Skalarprodukt*.
- Die Zeilen von A sind eine *ONB von \mathbb{K}^n mit kanonischem Skalarprodukt*.

Beweis: Die Äquivalenz von a. und b. folgt unmittelbar aus der Definition, denn $A^* \circ A = \mathbb{1}_n$ heißt, daß A^* die Inverse von A ist.

Ist \mathbf{a}^i der i -te Spaltenvektor von A , so ist $\overline{\mathbf{a}^i}^t$ der i -te Zeilenvektor von A^* und deshalb ist

$$\langle \mathbf{a}^i, \mathbf{a}^j \rangle = \overline{\mathbf{a}^i}^t \circ \mathbf{a}^j$$

der Eintrag an der Stelle (i, j) von $A^* \circ A$. Deshalb sind die Spalten von A genau dann eine ONB von \mathbb{K}^n , wenn $A^* = A^{-1}$ die Inverse von A ist. Dies zeigt die Äquivalenz von b. und c..

Ist \mathbf{a}_i der i -te Zeilenvektor von A , so ist $\overline{\mathbf{a}_i}^t$ der i -te Spaltenvektor von A^* . Also ist

$$\overline{\langle \mathbf{a}_i^t, \mathbf{a}_j^t \rangle} = \mathbf{a}_i \circ \overline{\mathbf{a}_j}^t$$

der Eintrag von $A \circ A^*$ an der Stelle (i, j) . Dies zeigt schließlich, daß b. und d. äquivalent sind. □

Beispiel 37.26 (Orthogonale Matrix)

Die Matrix

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \in \text{Mat}_3(\mathbb{R})$$

ist orthogonal, da ihre Spalten nach Beispiel 37.22 eine ONB von \mathbb{R}^3 bezüglich des kanonischen Skalarproduktes sind.

Korollar 37.27 (Die orthogonale und die unitäre Gruppe)
 $(O(\mathfrak{n}), \circ)$ und $(U(\mathfrak{n}), \circ)$ sind Gruppen.

Beweis: Es reicht, zu zeigen, daß sie Untergruppen von $GL_n(\mathbb{C})$ sind. Offenbar sind $O(\mathfrak{n})$ und $U(\mathfrak{n})$ nicht-leere Teilmengen von $GL_n(\mathbb{C})$. Sind nun A und B in $O(\mathfrak{n})$ bzw. in $U(\mathfrak{n})$, so gilt

$$(A \circ B)^* = B^* \circ A^* = B^{-1} \circ A^{-1} = (A \circ B)^{-1},$$

und

$$(A^{-1})^* = A^{**} = A = (A^{-1})^{-1}.$$

Mithin liegen auch $A \circ B$ und A^{-1} in $O(\mathfrak{n})$ bzw. in $U(\mathfrak{n})$. Damit ist gezeigt, daß $O(\mathfrak{n})$ und $U(\mathfrak{n})$ Untergruppen von $GL_n(\mathbb{C})$ sind. \square

Bemerkung 37.28 (Die orthogonale Gruppe $O(2)$)

Die Determinante

$$\det : O(2) \longrightarrow \{1, -1\}$$

ist wegen des Determinantenmultiplikationssatzes ein Gruppenepimorphismus. Der Kern von \det ist der Normalteiler

$$SO(2) := \{A \in O(2) \mid \det(A) = 1\}$$

von $O(2)$ und wird die *spezielle orthogonale Gruppe* vom Grad 2 genannt. Man kann zeigen, daß

$$SO(2) = \{T(\alpha) \mid \alpha \in [0, 2\pi)\}$$

und

$$O(2) \setminus SO(2) = \{S(\alpha) \mid \alpha \in [0, 2\pi)\},$$

wobei

$$T(\alpha) := \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

eine Drehung um den Winkel α ist und

$$S(\alpha) := \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{pmatrix}$$

eine Spiegelung an der Geraden $\text{Lin} \left(\left(\cos \left(\frac{\alpha}{2} \right), \sin \left(\frac{\alpha}{2} \right) \right)^t \right)$. Insbesondere ist im Fall $\mathfrak{n} = 2$ also jede orthogonale Matrix eine Drehung oder eine Spiegelung.

Man beachte auch daß $S(\alpha) = T(\alpha) \circ S(0)$ d. h. die von $S(\alpha)$ induzierte Spiegelung ist Komposition der Spiegelung an der x -Achse gefolgt von einer Drehung um den Winkel α . Damit gilt zugleich, daß jede Drehung im \mathbb{R}^2 Komposition von zwei Spiegelungen ist.

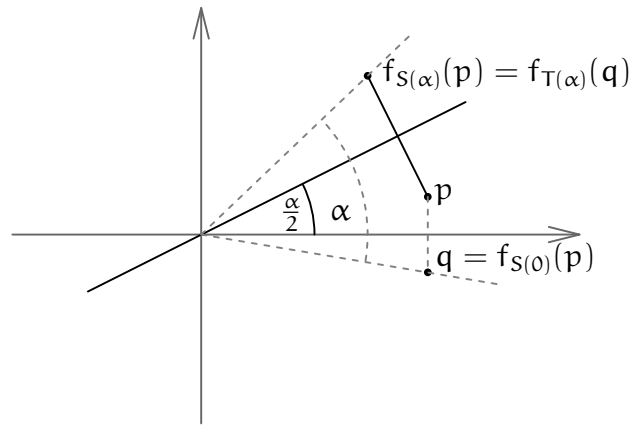


ABBILDUNG 4. Die Spielung $S(\alpha) = T(\alpha) \circ S(0)$

F) Orthogonale Summe und orthogonale Projektion

Definition 37.29 (Orthogonale Summe)

Sei V ein euklidischer oder unitärer Raum. Wir nennen V die *orthogonale Summe* der Unterräume U_1, \dots, U_r , falls $V = U_1 \oplus \dots \oplus U_r$ und $U_i \perp U_j$ für $i \neq j$. In diesem Fall schreiben wir $V = U_1 \perp \dots \perp U_r$.

Proposition 37.30 (Orthogonales Komplement)

Ist V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$, so gilt

$$V = U \perp U^\perp.$$

Insbesondere, U^\perp ist ein Komplement von U .

Beweis: Da nach Voraussetzung $U \perp U^\perp$ gilt, bleibt $U \cap U^\perp = \{0\}$ und $V = U + U^\perp$ zu zeigen, wobei für letzteres auch $V \subseteq U + U^\perp$ reicht.

Ist $x \in U \cap U^\perp$, so gilt $\langle x, x \rangle = 0$ und damit $x = 0$, da das Skalarprodukt positiv definit ist. Also ist $U \cap U^\perp = \{0\}$.

Zudem können wir wegen Satz 37.18 eine ONB (x_1, \dots, x_r) von U wählen und diese zu einer ONB (x_1, \dots, x_n) von V ergänzen. Dann gilt aber

$$V = \text{Lin}(x_1, \dots, x_r) + \text{Lin}(x_{r+1}, \dots, x_n) \subseteq U + U^\perp,$$

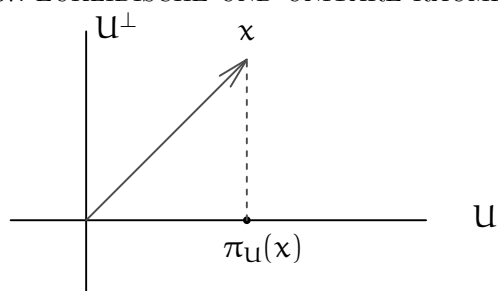
da nach Wahl $x_{r+1}, \dots, x_n \in U^\perp$. Hierbei beachte man, daß ein Vektor, der orthogonal zu einer Basis von U ist, automatisch orthogonal zu jedem Vektor in U ist. \square

Bemerkung 37.31

Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$. Da sich jeder Vektor $x \in V$ in eindeutiger Weise darstellen läßt als $x = u + u'$ mit $u \in U$ und $u' \in U^\perp$, können wir die *orthogonale Projektion* von V auf U

$$\pi_U : V \rightarrow V$$

definieren durch $\pi(u + u') = u$ für $u \in U$ und $u' \in U^\perp$.



In Aufgabe 37.34 wird gezeigt, daß π_U in der Tat eine Projektion ist, d.h. π_U ist linear mit $\pi_U^2 = \pi_U$. Außerdem ist $\text{Im}(\pi_U) = U$ das Bild von π_U und $\text{Ker}(\pi_U) = U^\perp$ der Kern.

Bemerkung 37.32 (Determinante als Volumenform)

In Bemerkung 30.12 haben wir das Parallelotop

$$P(x, y, z) := \{\lambda x + \mu y + \nu z \in \mathbb{R}^3 \mid 0 \leq \lambda, \mu, \nu \leq 1\}$$

betrachtet, das von den Vektoren $0 \neq x, y, z \in \mathbb{R}^3$ aufgespannt wird (siehe Abbildung 5).

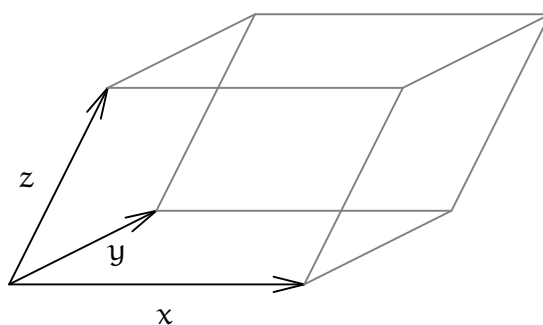


ABBILDUNG 5. Das Parallelotop $P(x, y, z)$ im \mathbb{R}^3

Wir wollen zeigen, daß das Volumen

$$\text{Volumen}(P(x, y, z)) = |\det(x \ y \ z)|$$

die Determinante der Matrix ist, deren Spalten die Vektoren x , y und z sind.

Elementargeometrisch berechnet sich das Volumen von $P(x, y, z)$ als Grundfläche mal Höhe, d.h. als Fläche A des von x und y aufgespannten Parallelogramms multipliziert mit der Höhe h des Parallelogramms. Dabei berechnet sich A als Länge von x mal der Höhe h' des Parallelogramms.

Wenden wir uns zunächst letzterer Berechnung zu. Es sei $U = \text{Lin}(x)$ und π_U sei die orthogonale Projektion auf U . Dann ist die Höhe h' des Parallelogramms genau die Länge des Vektors $y - \pi_U(y)$ (siehe Abbildung 6).

Für die Fläche A des Parallelogramms gilt deshalb

$$A = \|x\| \cdot \|y - \pi_U(y)\|.$$

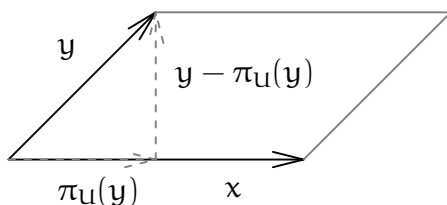


ABBILDUNG 6. Die Höhe im Parallelogramm zu x und y .

Auf ähnliche Weise kann man die Höhe h des Parallelotops $P(x, y, z)$ bestimmen. Hierzu betrachten wir den Unterraum $W = \text{Lin}(x, y)$ und die orthogonale Projektion π_W auf W . Dann ist h die Länge des Vektors $z - \pi_W(z)$ (siehe Abbildung 7).

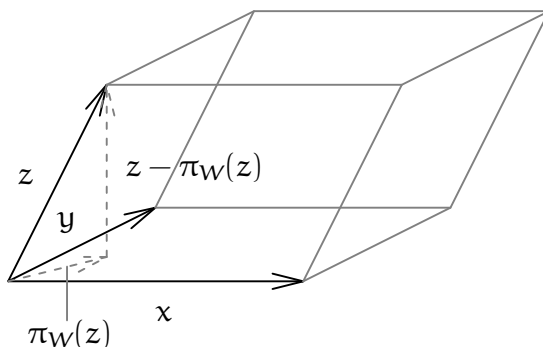


ABBILDUNG 7. Die Höhe in $P(x, y, z)$

Für das Volumen von $P(x, y, z)$ erhalten wir deshalb

$$\text{Volumen}(P(x, y, z)) = A \cdot h = \|x\| \cdot \|y - \pi_U(y)\| \cdot \|z - \pi_W(z)\|.$$

Wegen $\pi_U(y) \in U = \text{Lin}(x)$ $\pi_W(z) \in W = \text{Lin}(x, y)$, gibt es Zahlen $\lambda, \mu, \nu \in \mathbb{R}$ mit

$$\pi_U(y) = \lambda x \quad \text{und} \quad \pi_W(z) = \mu x + \nu y.$$

Dann gilt aber

$$\det(x \ y \ z) = \det(x \ y - \lambda x \ z - \mu x - \nu y) = \det(x \ y - \pi_U(y) \ z - \pi_W(z)),$$

da sich die Determinante einer Matrix nicht ändert, wenn wir Vielfache einer Spalte zu einer anderen addieren. Nun beachten wir, daß nach Konstruktion die Spalten der rechten Matrix orthogonal zueinander sind (siehe Abbildung 6 und 7). Normieren wir sie, so bilden sie eine ONB von \mathbb{R}^3 und die Matrix wird orthogonal. Da die Determinante einer orthogonalen Matrix Betrag 1 hat, erhalten wir also

$$\begin{aligned} |\det(x \ y \ z)| &= \|x\| \cdot \|y - \pi_U(y)\| \cdot \|z - \pi_W(z)\| \cdot \left| \det \begin{pmatrix} \frac{x}{\|x\|} & \frac{y - \pi_U(y)}{\|y - \pi_U(y)\|} & \frac{z - \pi_W(z)}{\|z - \pi_W(z)\|} \end{pmatrix} \right| \\ &= \|x\| \cdot \|y - \pi_U(y)\| \cdot \|z - \pi_W(z)\|. \end{aligned}$$

Dies beweist die Aussage und begründet den Begriff *Volumenform* im Zusammenhang mit Determinanten. Man beachte auch, daß die entsprechende Aussage für Parallelogramme analog gezeigt werden kann.

Aufgaben

Aufgabe 37.33

Es sei $V = \mathcal{C}([0, 1], \mathbb{R})$ der Vektorraum der auf $[0, 1]$ stetigen Funktionen. Zeige, daß durch

$$\langle f, g \rangle := \int_0^1 f(x) \cdot g(x) \, dx$$

ein Skalarprodukt auf V definiert wird.

Aufgabe 37.34 (Orthogonale Projektion)

Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$.

- Zeige, $\pi_U \in \text{End}_{\mathbb{K}}(V)$ ist eine Projektion mit $\text{Ker}(\pi_U) = U^\perp$ und $\text{Im}(\pi_U) = U$.
- Zeige, ist $\pi \in \text{End}_{\mathbb{K}}(V)$ eine Projektion mit $\text{Ker}(\pi) = U^\perp$ und $\text{Im}(\pi) = U$, dann ist $\pi = \pi_U$.
- Ist (x_1, \dots, x_r) eine ONB von U und $x \in V$, dann gilt

$$\pi_U(x) = \sum_{i=1}^r \langle x_i, x \rangle \cdot x_i.$$

Aufgabe 37.35

Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$. Dann gilt $(U^\perp)^\perp = U$.

Aufgabe 37.36

Zeige, durch $\langle (x_1, x_2, x_3)^t, (y_1, y_2, y_3)^t \rangle := x_1 y_1 + x_1 y_2 + x_2 y_1 + 2x_2 y_2 + x_2 y_3 + x_3 y_2 + 2x_3 y_3$ für $(x_1, x_2, x_3)^t, (y_1, y_2, y_3)^t \in \mathbb{R}^3$ wird ein Skalarprodukt auf \mathbb{R}^3 definiert und bestimme eine Orthonormalbasis des \mathbb{R}^3 bezüglich dieses Skalarproduktes.

Aufgabe 37.37

Betrachte den Vektorraum $U = \{f : [0, 1] \rightarrow \mathbb{R} : t \mapsto at^2 + bt + c \mid a, b, c \in \mathbb{R}\}$ und bestimme eine ONB bezüglich des Skalarproduktes aus Aufgabe 37.33.

Aufgabe 37.38

Für $V = \text{Mat}_n(\mathbb{R})$ definieren wir $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R} : (A, B) \mapsto \text{Spur}(A^t \circ B)$.

- Zeige, $\langle \cdot, \cdot \rangle$ ist ein Skalarprodukt auf V .
- Zeige, für $U = \{A \in V \mid A^t = A\}$ gilt $U^\perp = \{A \in V \mid A^t = -A\}$.

Aufgabe 37.39

Sei V ein euklidischer oder unitärer Raum und $\|\cdot\| : V \rightarrow \mathbb{R} : x \mapsto \sqrt{\langle x, x \rangle}$ die durch das Skalarprodukt definierte Norm. Zeige, für $x, y \in V$ gelten:

- Die Parallelogrammgleichung: $\|x + y\|^2 + \|x - y\|^2 = 2 \cdot (\|x\|^2 + \|y\|^2)$.
- Der Satz des Pythagoras: $x \perp y \implies \|x\|^2 + \|y\|^2 = \|x + y\|^2$.

§ 38 Spektralsatz und Hauptachsentransformation

In diesem Abschnitt sei V ein euklidischer oder unitärer Raum der Dimension $1 \leq n < \infty$ mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und euklidischer Norm $\| \cdot \|$.

A) Die adjungierte Abbildung

Satz 38.1 (Die adjungierte Abbildung)

Zu jedem Endomorphismus $f \in \text{End}_{\mathbb{K}}(V)$ gibt es genau ein $f^* \in \text{End}_{\mathbb{K}}(V)$, so daß

$$\langle f(x), y \rangle = \langle x, f^*(y) \rangle \quad (116)$$

für alle $x, y \in V$ gilt. Ist $B = (x_1, \dots, x_n)$ eine ONB von V , so gilt für $y \in V$

$$f^*(y) = \sum_{i=1}^n \langle f(x_i), y \rangle \cdot x_i. \quad (117)$$

Die Abbildung f^* heißt die adjungierte Abbildung zu f .

Beweis: Wir wollen zunächst zeigen, daß es einen Endomorphismus $f^* \in \text{End}_{\mathbb{K}}(V)$ mit der Eigenschaft (116) gibt. Dazu wählen wir eine ONB $B = (x_1, \dots, x_n)$ von V und definieren

$$f^* : V \longrightarrow V : y \mapsto \sum_{i=1}^n \langle f(x_i), y \rangle \cdot x_i,$$

d.h. wir definieren $f^*(y)$ durch die Formel in (117). Da das Skalarprodukt in der zweiten Komponente linear ist, ist f^* in der Tat eine lineare Abbildung, also ein Endomorphismus von V .

Seien nun $x, y \in V$ gegeben. Unter Anwendung der Parsevalschen Gleichung 37.16 gilt dann

$$\begin{aligned} \langle f(x), y \rangle &\stackrel{37.16}{=} \left\langle f \left(\sum_{i=1}^n \langle x_i, x \rangle \cdot x_i \right), y \right\rangle = \left\langle \sum_{i=1}^n \langle x_i, x \rangle \cdot f(x_i), y \right\rangle \\ &= \sum_{i=1}^n \overline{\langle x_i, x \rangle} \cdot \langle f(x_i), y \rangle = \sum_{i=1}^n \langle x, x_i \rangle \cdot \langle f(x_i), y \rangle \\ &= \sum_{i=1}^n \langle x, \langle f(x_i), y \rangle \cdot x_i \rangle = \left\langle x, \sum_{i=1}^n \langle f(x_i), y \rangle \cdot x_i \right\rangle = \langle x, f^*(y) \rangle. \end{aligned}$$

Damit ist gezeigt, daß der durch (117) definierte Endomorphismus die Gleichung (116) erfüllt.

Es bleibt noch die Eindeutigkeit zu zeigen. Sei dazu $h \in \text{End}_{\mathbb{K}}(V)$ mit

$$\langle f(x), y \rangle = \langle x, h(y) \rangle \quad (118)$$

für alle $x, y \in V$. Wegen der Parsevalschen Gleichung 37.16 gilt für $y \in V$ dann

$$h(y) \stackrel{37.16}{=} \sum_{i=1}^n \langle x_i, h(y) \rangle \cdot x_i \stackrel{(118)}{=} \sum_{i=1}^n \langle f(x_i), y \rangle \cdot x_i \stackrel{(117)}{=} f^*(y).$$

Mithin stimmen f^* und h überein, so daß f^* eindeutig bestimmt ist. □

Korollar 38.2 ($f^{**} = f$)

Ist $f \in \text{End}_{\mathbb{K}}(V)$, so gilt $f^{**} = f$, d.h. $\langle f^*(x), y \rangle = \langle x, f(y) \rangle$ für $x, y \in V$.

Beweis: Für $x, y \in V$ gilt

$$\langle f^*(x), y \rangle = \overline{\langle y, f^*(x) \rangle} \stackrel{(116)}{=} \overline{\langle f(y), x \rangle} = \langle x, f(y) \rangle.$$

Damit erfüllt f die Bedingung, durch die die Abbildung f^{**} eindeutig festgelegt ist. Also muß $f = f^{**}$ gelten. □

Korollar 38.3 (Matrixdarstellung der adjungierten Abbildung)

Sei B eine ONB von V und $f \in \text{End}_{\mathbb{K}}(V)$, so gilt

$$M_B^B(f^*) = M_B^B(f)^*,$$

d.h. die Matrixdarstellung der adjungierten Abbildung ist die Adjungierte der Matrixdarstellung.

Beweis: Seien $M_B^B(f) = (a_{ij})_{i,j}$ und $M_B^B(f^*) = (b_{ji})_{j,i}$. Unter Berücksichtigung der Parsevalschen Gleichung gilt 37.16 gilt

$$\sum_{i=1}^n a_{ij} \cdot x_i = f(x_j) \stackrel{37.16}{=} \sum_{i=1}^n \langle x_i, f(x_j) \rangle \cdot x_i \tag{119}$$

und

$$\sum_{j=1}^n b_{ji} \cdot x_j = f^*(x_i) \stackrel{37.16}{=} \sum_{j=1}^n \langle x_j, f^*(x_i) \rangle \cdot x_j. \tag{120}$$

Da die Darstellung als Linearkombination einer Basis eindeutig ist, erhalten wir

$$\overline{a_{ij}} \stackrel{(119)}{=} \overline{\langle x_i, f(x_j) \rangle} = \langle f(x_j), x_i \rangle \stackrel{(116)}{=} \langle x_j, f^*(x_i) \rangle \stackrel{(120)}{=} b_{ji}.$$

Daraus folgt

$$M_B^B(f^*) = (b_{ji})_{j,i} = (\overline{a_{ij}})_{j,i} = (\overline{a_{ij}})_{i,j}^t = M_B^B(f)^*.$$

□

Beispiel 38.4 (Adjungierte)

Wir betrachten $V = \mathbb{C}^2$ als unitären Raum mit dem kanonischen Skalarprodukt sowie die Abbildung

$$f : \mathbb{C}^2 \longrightarrow \mathbb{C}^2 : (x, y)^t \mapsto (2x + 4y, 2y - 4x)^t.$$

Bezüglich der kanonischen Basis E hat f die Matrixdarstellung

$$M_E^E(f) = \begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix}$$

und somit gilt

$$M_E^E(f)^* = \overline{\begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix}}^t = \begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix}^t = \begin{pmatrix} 2 & -4 \\ 4 & 2 \end{pmatrix}.$$

Da E eine ONB bezüglich des kanonischen Skalarproduktes ist, ist somit

$$f^* : \mathbb{C}^2 \longrightarrow \mathbb{C}^2 : (x, y)^t \mapsto (2x - 4y, 2y + 4x)^t$$

nach Korollar 38.3 die Adjungierte von f .

B) Der Spektralsatz für normale Abbildungen

Definition 38.5

Es sei $f \in \text{End}_{\mathbb{K}}(V)$ und $A \in \text{Mat}_n(\mathbb{K})$.

- f heißt *normal*, falls $f^* \circ f = f \circ f^*$.
- A heißt *normal*, falls $A^* \circ A = A \circ A^*$.

Bemerkung 38.6

- Jede symmetrische Matrix $A \in \text{Mat}_n(\mathbb{R})$ und jede hermitesche Matrix $A \in \text{Mat}_n(\mathbb{C})$ ist normal, denn wegen $A = A^*$ gilt auch

$$A \circ A^* = A \circ A = A^* \circ A.$$

- Jede orthogonale Matrix $A \in O(n)$ und jede unitäre Matrix $A \in U(n)$ ist normal, denn wegen $A^* = A^{-1}$ gilt auch

$$A \circ A^* = A \circ A^{-1} = \mathbb{1}_n = A^{-1} \circ A = A^* \circ A.$$

Lemma 38.7 (Matrixdarstellung normaler Endomorphismen)

Sei $f \in \text{End}_{\mathbb{K}}(V)$ und B eine ONB von V .

Genau dann ist f normal, wenn $M_B^B(f)$ normal ist.

Beweis: Ist f normal, so gilt

$$\begin{aligned} M_B^B(f)^* \circ M_B^B(f) &\stackrel{38.3}{=} M_B^B(f^*) \circ M_B^B(f) = M_B^B(f^* \circ f) \\ &= M_B^B(f \circ f^*) = M_B^B(f) \circ M_B^B(f^*) \stackrel{38.3}{=} M_B^B(f) \circ M_B^B(f)^* \end{aligned}$$

und somit ist $M_B^B(f)$ normal. Ist umgekehrt $M_B^B(f)$ normal, so gilt

$$\begin{aligned} M_B^B(f^* \circ f) &= M_B^B(f^*) \circ M_B^B(f) \stackrel{38.3}{=} M_B^B(f)^* \circ M_B^B(f) \\ &= M_B^B(f) \circ M_B^B(f)^* \stackrel{38.3}{=} M_B^B(f) \circ M_B^B(f^*) = M_B^B(f \circ f^*). \end{aligned}$$

Dann stimmen aber die Abbildungen $f^* \circ f$ und $f \circ f^*$ überein, und somit ist f normal. \square

Beispiel 38.8 (Normale Abbildung)

In Beispiel 38.4 gilt

$$\begin{aligned} M_E^E(f) \circ M_E^E(f)^* &= \begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix} \circ \begin{pmatrix} 2 & -4 \\ 4 & 2 \end{pmatrix} \\ &= \begin{pmatrix} 20 & 0 \\ 0 & 20 \end{pmatrix} = \begin{pmatrix} 2 & -4 \\ 4 & 2 \end{pmatrix} \circ \begin{pmatrix} 2 & 4 \\ -4 & 2 \end{pmatrix} = M_E^E(f)^* \circ M_E^E(f). \end{aligned}$$

Somit ist $M_E^E(f)$ normal und da E eine ONB ist, ist dann auch f normal.

Lemma 38.9 (Normale Abbildungen)

Sei $f \in \text{End}_{\mathbb{K}}(V)$ normal. Genau dann gilt $x \in \text{Eig}(f, \lambda)$, wenn $x \in \text{Eig}(f^*, \bar{\lambda})$.

Beweis: Für ein beliebiges $x \in V$ gilt

$$\begin{aligned} \langle f^*(x) - \bar{\lambda}x, f^*(x) - \bar{\lambda}x \rangle &= \langle f^*(x), f^*(x) \rangle - \langle f^*(x), \bar{\lambda}x \rangle - \langle \bar{\lambda}x, f^*(x) \rangle + \langle \bar{\lambda}x, \bar{\lambda}x \rangle \\ &= \langle f \circ f^*(x), x \rangle - \bar{\lambda} \langle f^*(x), x \rangle - \lambda \langle x, f^*(x) \rangle + \lambda \bar{\lambda} \langle x, x \rangle \\ &= \langle f^* \circ f(x), x \rangle - \bar{\lambda} \langle x, f(x) \rangle - \lambda \langle f(x), x \rangle + \lambda \bar{\lambda} \langle x, x \rangle \\ &= \langle f(x), f(x) \rangle - \langle \lambda x, f(x) \rangle - \langle f(x), \lambda x \rangle + \langle \lambda x, \lambda x \rangle \\ &= \langle f(x) - \lambda x, f(x) - \lambda x \rangle. \end{aligned}$$

Da das Skalarprodukt positiv definit ist, gilt somit

$$\begin{aligned} x \in \text{Eig}(f, \lambda) &\iff f(x) - \lambda x = 0 \iff \langle f(x) - \lambda x, f(x) - \lambda x \rangle = 0 \\ &\iff \langle f^*(x) - \bar{\lambda}x, f^*(x) - \bar{\lambda}x \rangle = 0 \iff f^*(x) - \bar{\lambda}x = 0 \\ &\iff x \in \text{Eig}(f^*, \bar{\lambda}). \end{aligned}$$

□

Satz 38.10 (Spektralsatz für normale Abbildungen)

Für $f \in \text{End}_{\mathbb{K}}(V)$ sind die folgenden beiden Aussagen gleichwertig:

- f ist normal und χ_f zerfällt über \mathbb{K} in Linearfaktoren.
- V besitzt eine ONB aus Eigenvektoren von f .

Inbesondere ist f dann bezüglich einer ONB diagonalisierbar.

Beweis:

b. \implies a.: Besitzt V eine Basis B aus Eigenvektoren, so zerfällt χ_f nach Satz 33.21 über \mathbb{K} in Linearfaktoren. Zudem ist dann $M_B^B(f)$ eine Diagonalmatrix, und wegen Korollar 38.3 ist dann auch

$$M_B^B(f)^* = \overline{M_B^B(f)}^t$$

eine Diagonalmatrix. Da zwei Diagonalmatrizen stets kommutieren, gilt also

$$M_B^B(f)^* \circ M_B^B(f) = M_B^B(f) \circ M_B^B(f)^*,$$

d.h. $M_B^B(f)$ ist normal. Nach Lemma 38.7 ist dann aber auch f normal.

a. \implies b.: Wir führen den Beweis mit Induktion nach $n = \dim_{\mathbb{K}}(V)$, wobei für $n = 1$ der Endomorphismus f für jede ONB $B = (x_1)$ diagonalisierbar ist und zudem x_1 ein Eigenvektor von f ist. Sei also $n > 1$.

Da χ_f über \mathbb{K} in Linearfaktoren zerfällt, besitzt f einen Eigenwert $\lambda \in \mathbb{K}$ und einen zugehörigen Eigenvektor $0 \neq x \in V$. Dann ist $U := \text{Lin}(x)$ ein f -invarianter Unterraum der Dimension 1 und es gilt

$$V = U \perp U^\perp = U \oplus U^\perp$$

nach Proposition 37.30.

Wir zeigen nun zunächst, daß auch U^\perp ein f -invarianter Unterraum ist, der dann die Dimension $n - 1$ hat. Sei dazu $y \in U^\perp$, dann gilt

$$\langle f(y), x \rangle = \langle y, f^*(x) \rangle \stackrel{38.9}{=} \langle y, \bar{\lambda}x \rangle = \bar{\lambda} \cdot \langle y, x \rangle = 0,$$

da $y \perp x$. Damit gilt dann aber auch $f(y) \perp x$, und somit $f(y) \in U^\perp$.

Als nächstes wollen wir zeigen, daß f_{U^\perp} normal ist. Dazu beachten wir zunächst, daß U^\perp auch f^* -invariant ist, da für $y \in U^\perp$ wie oben

$$\langle f^*(y), x \rangle = \langle y, f(x) \rangle = \langle y, \lambda x \rangle = \lambda \cdot \langle y, x \rangle = 0$$

und somit $f^*(y) \perp x$ und $f^*(y) \in U^\perp$ gilt. Aus der definierenden Eigenschaft der adjungierten Abbildung folgt dann aber, daß die adjungierte Abbildung $(f_{U^\perp})^*$ der Einschränkung f_{U^\perp} genau die Einschränkung $(f^*)_{U^\perp}$ der adjungierten Abbildung f^* auf U^\perp ist. Die Normalität von f überträgt sich also direkt auf f_{U^\perp} durch

$$f_{U^\perp} \circ (f_{U^\perp})^* = (f \circ f^*)_{U^\perp} = (f^* \circ f)_{U^\perp} = (f_{U^\perp})^* \circ f_{U^\perp}.$$

Außerdem gilt

$$\chi_f = \chi_{f_U} \cdot \chi_{f_{U^\perp}},$$

da V die direkte Summe der beiden f -invarianten Unterräume U und U^\perp ist, und deshalb zerfällt $\chi_{f_{U^\perp}}$ über \mathbb{K} in Linearfaktoren. Nach Induktion besitzt U^\perp deshalb eine ONB (x_2, \dots, x_n) aus Eigenvektoren von f_{U^\perp} . Dann ist aber $B = (x, x_2, \dots, x_n)$ eine ONB aus Eigenvektoren von f . \square

Korollar 38.11 (Spektralsatz für normale Matrizen)

Für eine Matrix $A \in \text{Mat}_n(\mathbb{K})$ sind die folgenden beiden Aussagen gleichwertig:

- A ist normal und χ_A zerfällt über \mathbb{K} in Linearfaktoren.
- Es gibt ein T in $O(n)$ bzw. $U(n)$, so daß $T^{-1} \circ A \circ T$ eine Diagonalmatrix ist.

Beweis: Wenden wir den Spektralsatz 38.10 auf f_A und \mathbb{K}^n mit dem kanonischen Skalarprodukt an, so enthält die Basistransformationsmatrix $T = T_E^B$ genau die Vektoren der ONB B als Spalten und ist nach Proposition 37.25 daher orthogonal bzw. unitär. \square

Der Beweis ist konstruktiv, sofern man die Eigenwerte von A exakt kennt. Man leitet daraus folgenden prinzipiellen Algorithmus zur Bestimmung von T her.

Algorithmus 38.12

INPUT: $A \in \text{Mat}_n(\mathbb{K})$ normal mit χ_A zerfällt über \mathbb{K} .

OUTPUT: T in $O(n)$ bzw. $U(n)$, so daß $T^{-1} \circ A \circ T$ Diagonalgestalt hat.

1. **Schritt:** Bestimme die Eigenwerte von A .
2. **Schritt:** Bestimme für jeden Eigenwert von A eine Basis des zugehörigen Eigenraumes.
3. **Schritt:** Orthonormalisiere die Basen der Eigenräume mit dem Orthonormalisierungsverfahren von Gram-Schmidt und schreibe die Vektoren als Spalten in eine Matrix T .
4. **Schritt:** Gib schließlich T zurück.

Beispiel 38.13 (Diagonalisierung einer normalen Abbildung)

Die Abbildung f in Beispiel 38.4 ist nach Beispiel 38.8 normal.

$$\chi_f = \det(t \cdot \mathbb{1}_2 - M_E^E(f)) = \begin{vmatrix} t-2 & -4 \\ 4 & t-2 \end{vmatrix} = t^2 - 4t + 20 = (t - (2+4i)) \cdot (t - (2-4i))$$

zerfällt über \mathbb{C} in Linearfaktoren. Mithin gibt es wegen des Spektralsatzes eine ONB aus Eigenvektoren von f , so daß

$$M_B^B(f) = \begin{pmatrix} 2+4i & 0 \\ 0 & 2-4i \end{pmatrix}.$$

Um B zu berechnen, berechnen wir zunächst die beiden Eigenräume $\text{Eig}(f, 2+4i)$ und $\text{Eig}(f, 2-4i)$ ausgehend von der Matrixdarstellung $M_E^E(f)$ in Beispiel 38.4.

Für $\text{Eig}(f, 2+4i) = \text{Lös}(M_E^E(f) - (2+4i) \cdot \mathbb{1}_2, 0)$ liefert unser Algorithmus

$$\begin{pmatrix} -4i & 4 \\ -4 & -4i \end{pmatrix} \xrightarrow[\text{I} \rightarrow \frac{1}{-4i} \cdot \text{I}]{\text{II} \rightarrow \text{II} + i \cdot \text{I}} \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} \xrightarrow[\text{erg\u00e4nzen}]{-1 \cdot \text{en}} \begin{pmatrix} 1 & i \\ 0 & -1 \end{pmatrix},$$

so daß $x_1 = (i, -1)^t$ eine Basis von $\text{Eig}(f, 2+4i)$ ist. Analog erhalten wir für $\text{Eig}(f, 2-4i) = \text{Lös}(M_E^E(f) - (2-4i) \cdot \mathbb{1}_2, 0)$

$$\begin{pmatrix} 4i & 4 \\ -4 & 4i \end{pmatrix} \xrightarrow[\text{I} \rightarrow \frac{1}{4i} \cdot \text{I}]{\text{II} \rightarrow \text{II} - i \cdot \text{I}} \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix} \xrightarrow[\text{erg\u00e4nzen}]{-1 \cdot \text{en}} \begin{pmatrix} 1 & -i \\ 0 & -1 \end{pmatrix},$$

so daß $x_2 = (-i, -1)^t$ eine Basis von $\text{Eig}(f, 2-4i)$ ist. Die Vektoren x_1 und x_2 sind bereits orthogonal zueinander, da

$$\langle x_1, x_2 \rangle = \bar{i} \cdot (-i) + \overline{-1} \cdot (-1) = -1 + 1 = 0.$$

Mithin reicht es, sie zu normieren, und wir erhalten die gewünschte ONB

$$B = \left(\frac{1}{\|x_1\|} \cdot x_1, \frac{1}{\|x_2\|} \cdot x_2 \right) = \left(\left(\frac{i}{\sqrt{2}}, \frac{-1}{\sqrt{2}} \right)^t, \left(\frac{-i}{\sqrt{2}}, \frac{-1}{\sqrt{2}} \right)^t \right).$$

Es ist übrigens kein Zufall, daß die beiden Eigenvektoren im letzten Beispiel orthogonal zueinander standen.

Lemma 38.14

Ist $f \in \text{End}_{\mathbb{K}}(\mathbf{V})$ normal, $\mathbf{x} \in \text{Eig}(f, \lambda)$, $\mathbf{y} \in \text{Eig}(f, \mu)$ und $\lambda \neq \mu$, dann gilt $\mathbf{x} \perp \mathbf{y}$.

Beweis: Nach Voraussetzung gilt

$$\lambda \cdot \langle \mathbf{x}, \mathbf{y} \rangle = \langle \bar{\lambda} \mathbf{x}, \mathbf{y} \rangle \stackrel{38.9}{=} \langle f^*(\mathbf{x}), \mathbf{y} \rangle \stackrel{38.2}{=} \langle \mathbf{x}, f(\mathbf{y}) \rangle = \langle \mathbf{x}, \mu \mathbf{y} \rangle = \mu \cdot \langle \mathbf{x}, \mathbf{y} \rangle.$$

Für die Differenz der beiden Seiten erhalten wir dann $(\lambda - \mu) \cdot \langle \mathbf{x}, \mathbf{y} \rangle = 0$, wobei nach Voraussetzung $\lambda - \mu \neq 0$ gilt. Also ist $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ und somit $\mathbf{x} \perp \mathbf{y}$. \square

Satz 38.15 (Spektralzerlegung für normale Endomorphismen)

Sei $f \in \text{End}_{\mathbb{K}}(\mathbf{V})$ normal mit zerfallendem charakteristischem Polynom und seien $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ die paarweise verschiedenen Eigenwerte von f . Ferner bezeichne

$$\pi_i : \mathbf{V} \longrightarrow \text{Eig}(f, \lambda_i)$$

die orthogonale Projektion von \mathbf{V} auf $\text{Eig}(f, \lambda_i)$.

Dann ist

$$\mathbf{V} = \text{Eig}(f, \lambda_1) \perp \dots \perp \text{Eig}(f, \lambda_r)$$

die orthogonale Summe der Eigenräume von f und es gilt

$$f = \lambda_1 \cdot \pi_1 + \dots + \lambda_r \cdot \pi_r.$$

Man nennt dies die Spektralzerlegung von f .

Beweis: Nach dem Spektralsatz für normale Abbildungen ist f diagonalisierbar. Deshalb folgt aus Satz 33.21, daß \mathbf{V} die direkte Summe der Eigenräume von f ist. Wegen Lemma 38.14 ist diese Summe dann eine orthogonale Summe.

Ist nun $\mathbf{x} = \mathbf{x}_1 + \dots + \mathbf{x}_r \in \mathbf{V}$ mit $\mathbf{x}_i \in \text{Eig}(f, \lambda_i)$ gegeben, so gilt

$$\pi_j(\mathbf{x}_i) = \delta_{ij} \cdot \mathbf{x}_i,$$

da $\mathbf{x}_j \in \text{Eig}(f, \lambda_j)$ und $\mathbf{x}_i \perp \text{Eig}(f, \lambda_j)$ für $i \neq j$, und somit

$$\pi_j(\mathbf{x}) = \pi_j(\mathbf{x}_1) + \dots + \pi_j(\mathbf{x}_r) = \mathbf{x}_j.$$

Damit erhalten wir dann

$$\begin{aligned} (\lambda_1 \cdot \pi_1 + \dots + \lambda_r \cdot \pi_r)(\mathbf{x}) &= \lambda_1 \cdot \pi_1(\mathbf{x}) + \dots + \lambda_r \cdot \pi_r(\mathbf{x}) \\ &= \lambda_1 \cdot \mathbf{x}_1 + \dots + \lambda_r \cdot \mathbf{x}_r = f(\mathbf{x}_1) + \dots + f(\mathbf{x}_r) = f(\mathbf{x}). \end{aligned}$$

\square

Bemerkung 38.16 (Spektralzerlegung und Hauptraumzerlegung)

Die Projektionen π_i stimmen übrigens mit den Projektionen $Q_i(f)$ aus Bemerkung 33.19 überein.

C) Orthogonale und unitäre Abbildungen

Wir kommen jetzt zu den strukturerhaltenden Abbildungen, d. h. zu solchen, die mit dem Skalarprodukt verträglich sind. Diese haben einen speziellen Namen.

Definition 38.17 (Orthogonale / unitäre Abbildungen)

- Ein Endomorphismus $f \in \text{End}_{\mathbb{R}}(V)$ heißt *orthogonal*, falls $f^* \circ f = \text{id}_V$ gilt.
 $O(V) := \{f \in \text{End}_{\mathbb{R}}(V) \mid f \text{ ist orthogonal}\}$ heißt die *orthogonale Gruppe* von V .
- Ein Endomorphismus $f \in \text{End}_{\mathbb{C}}(V)$ heißt *unitär*, falls $f^* \circ f = \text{id}_V$ gilt.
Wir nennen $U(V) := \{f \in \text{End}_{\mathbb{C}}(V) \mid f \text{ ist unitär}\}$ die *unitäre Gruppe* von V .

Proposition 38.18 (Matrixdarstellung orthogonaler / unitärer Abbildungen)

Es sei $f \in \text{End}_{\mathbb{K}}(V)$ und B eine ONB von V .

Genau dann ist f orthogonal bzw. unitär, wenn $M_B^B(f)$ orthogonal bzw. unitär ist.

Beweis: Es gilt:

$$\begin{aligned} f \text{ ist orthogonal bzw. unitär} &\iff f^* \circ f = \text{id}_V \\ &\iff M_B^B(f^*) \circ M_B^B(f) = M_B^B(f^* \circ f) = \mathbb{1}_n \\ &\stackrel{38.3}{\iff} M_B^B(f)^* \circ M_B^B(f) = \mathbb{1}_n \\ &\iff M_B^B(f) \text{ ist orthogonal bzw. unitär.} \end{aligned}$$

□

Korollar 38.19 (Orthogonal / unitär \implies normal)

Orthogonale und unitäre Abbildungen sind normal.

Beweis: Ist $f \in \text{End}_{\mathbb{K}}(V)$ orthogonal oder unitär und B eine ONB von V , so ist $M_B^B(f)$ nach Proposition 38.18 orthogonal oder unitär. Nach Bemerkung 38.6 ist dann $M_B^B(f)$ auch normal, und mit Lemma 38.7 ist f deshalb normal. □

Proposition 38.20 (Charakterisierung orthogonaler / unitärer Endomorphismen)

Ein Endomorphismus $f \in \text{End}_{\mathbb{K}}(V)$ ist genau dann orthogonal bzw. unitär, wenn

$$\langle f(x), f(y) \rangle = \langle x, y \rangle$$

für alle $x, y \in V$ gilt.

Beweis: Ist f orthogonal bzw. unitär und sind $x, y \in V$, so gilt

$$\langle x, y \rangle = \langle x, \text{id}_V(y) \rangle = \langle x, f^* \circ f(y) \rangle = \langle f(x), f(y) \rangle.$$

Ist umgekehrt f mit dem Skalarprodukt verträglich, so gilt

$$\langle x, y \rangle = \langle f(x), f(y) \rangle = \langle (f^* \circ f)(x), y \rangle$$

und somit

$$\langle (f^* \circ f)(x) - \text{id}_V(x), y \rangle = \langle (f^* \circ f)(x), y \rangle - \langle x, y \rangle = 0$$

für alle $\mathbf{x}, \mathbf{y} \in V$. Für $\mathbf{x} \in V$ beliebig setzen wir $\mathbf{y} = (f^* \circ f)(\mathbf{x}) - \text{id}_V(\mathbf{x})$ und erhalten

$$\langle (f^* \circ f)(\mathbf{x}) - \text{id}_V(\mathbf{x}), (f^* \circ f)(\mathbf{x}) - \text{id}_V(\mathbf{x}) \rangle = 0,$$

was $(f^* \circ f)(\mathbf{x}) - \text{id}_V(\mathbf{x}) = \mathbf{0}$ und $(f^* \circ f)(\mathbf{x}) = \text{id}_V(\mathbf{x})$ zur Folge hat. Mithin ist $f^* \circ f = \text{id}_V$ und f ist orthogonal bzw. unitär. \square

Proposition 38.21 (Eigenschaften orthogonaler / unitärer Abbildungen)

Es seien $f, g \in \text{End}_{\mathbb{K}}(V)$ orthogonal bzw. unitär und $\mathbf{x}, \mathbf{y} \in V$

- $\|f(\mathbf{x})\| = \|\mathbf{x}\|$.
- $\mathbf{x} \perp \mathbf{y}$ genau dann, wenn $f(\mathbf{x}) \perp f(\mathbf{y})$.
- Jeder Eigenwert von f hat Betrag 1.
- f ist bijektiv.
- f^{-1} und $f \circ g$ sind orthogonal bzw. unitär, d.h. $O(V)$ und $U(V)$ sind Gruppen.

Insbesondere, orthogonale und unitäre Abbildungen erhalten Längen und Abstände.

Beweis:

- $\|f(\mathbf{x})\|^2 = \langle f(\mathbf{x}), f(\mathbf{x}) \rangle = \langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2$.
- $\mathbf{x} \perp \mathbf{y} \iff \langle f(\mathbf{x}), f(\mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle = 0 \iff f(\mathbf{x}) \perp f(\mathbf{y})$.
- Ist $\mathbf{0} \neq \mathbf{x} \in V$ ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{K}$, so gilt nach a.

$$\|\mathbf{x}\| = \|f(\mathbf{x})\| = \|\lambda \mathbf{x}\| = |\lambda| \cdot \|\mathbf{x}\|,$$

also $|\lambda| = 1$, da $\mathbf{x} \neq \mathbf{0}$.

- Ist $\mathbf{x} \in \text{Ker}(f)$, so gilt nach a. $\mathbf{0} = \|f(\mathbf{x})\| = \|\mathbf{x}\|$, und somit $\mathbf{x} = \mathbf{0}$. Also ist f injektiv, und da V endlich-dimensional ist, ist f somit auch bijektiv.
- f^{-1} und $f \circ g$ sind orthogonal bzw. unitär, da für $\mathbf{x}, \mathbf{y} \in V$ gelten

$$\langle f^{-1}(\mathbf{x}), f^{-1}(\mathbf{y}) \rangle = \langle f(f^{-1}(\mathbf{x})), f(f^{-1}(\mathbf{y})) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle$$

und

$$\langle (f \circ g)(\mathbf{x}), (f \circ g)(\mathbf{y}) \rangle = \langle f(g(\mathbf{x})), f(g(\mathbf{y})) \rangle = \langle g(\mathbf{x}), g(\mathbf{y}) \rangle = \langle \mathbf{x}, \mathbf{y} \rangle.$$

\square

Bemerkung 38.22 (Orthogonale Abbildungen sind winkeltreu.)

Orthogonale Abbildungen erhalten Winkel, d.h. ist $f \in O(V)$ und sind $\mathbf{0} \neq \mathbf{x}, \mathbf{y} \in V$, so gilt

$$\angle(f(\mathbf{x}), f(\mathbf{y})) = \angle(\mathbf{x}, \mathbf{y}),$$

da f das Skalarprodukt und die euklidische Norm erhält. Es gilt nämlich

$$\angle(f(\mathbf{x}), f(\mathbf{y})) = \arccos \left(\frac{\langle f(\mathbf{x}), f(\mathbf{y}) \rangle}{\|f(\mathbf{x})\| \cdot \|f(\mathbf{y})\|} \right) = \arccos \left(\frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|} \right) = \angle(\mathbf{x}, \mathbf{y}).$$

Satz 38.23 (Spektralsatz für unitäre Abbildungen)

Ein Endomorphismus $f \in \text{End}_{\mathbb{C}}(V)$ ist genau dann unitär, wenn V eine ONB aus Eigenvektoren von f besitzt und alle Eigenwerte von f Betrag 1 haben.

Beweis: Ist f unitär, so ist f nach Korollar 38.19 normal und χ_f zerfällt nach dem Fundamentalsatz der Algebra über \mathbb{C} in Linearfaktoren. Aufgrund des Spektralsatzes für normale Abbildungen 38.10 besitzt V dann eine ONB aus Eigenvektoren von f . Zudem haben die Eigenwerte nach Proposition 38.21 alle Betrag 1.

Besitzt umgekehrt V eine ONB $B = (x_1, \dots, x_n)$ aus Eigenvektoren von f zu Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ vom Betrag 1, so gilt

$$\begin{aligned} M_B^B(f) \circ M_B^B(f)^* &= \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \circ \begin{pmatrix} \overline{\lambda_1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \overline{\lambda_n} \end{pmatrix} \\ &= \begin{pmatrix} \lambda_1 \overline{\lambda_1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \overline{\lambda_n} \end{pmatrix} = \begin{pmatrix} |\lambda_1|^2 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & |\lambda_n|^2 \end{pmatrix} = \mathbb{1}_n. \end{aligned}$$

Mithin ist $M_B^B(f)$ unitär, und damit ist auch f unitär nach Proposition 38.18. □

Korollar 38.24 (Spektralsatz für unitäre Matrizen)

Ist $A \in U(n)$, dann gibt es ein $T \in U(n)$ mit

$$T^{-1} \circ A \circ T = T^* \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix}$$

wobei $\lambda_i \in \mathbb{C}$ mit $|\lambda_i| = 1$, $i = 1, \dots, n$, die Eigenwerte von A sind. Insbesondere ist jede unitäre Matrix diagonalisierbar.

Beweis: Ist A unitär, dann ist f_A unitär und wir finden eine ONB von \mathbb{C}^n aus Eigenvektoren von f_A , also von A , und alle Eigenwerte haben Betrag 1. Schreiben wir die Eigenvektoren als Spalten in eine Matrix T , so ist $T \in U(n)$ und T transformiert A in eine Diagonalmatrix. □

Beispiel 38.25 (Unitäre Matrix)

Betrachten wir \mathbb{C}^3 mit dem kanonischen Skalarprodukt sowie die Matrix

$$A = \frac{1}{9} \cdot \begin{pmatrix} 1 & 8 & -4 \\ -4 & 4 & 7 \\ 8 & 1 & 4 \end{pmatrix} \in \text{Mat}_3(\mathbb{C}).$$

Man rechnet sofort nach, daß $A \circ A^* = \mathbf{1}_3$, daß A also orthogonal bzw. unitär ist, mit charakteristischem Polynom

$$\chi_A = t^3 - t^2 + t - 1 = (t - 1) \cdot (t - i) \cdot (t + i).$$

Es gibt also eine unitäre Matrix $T \in U(3)$ mit

$$T^* \circ A \circ T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}.$$

Um T zu bestimmen, berechnen wir zunächst den Eigenraum $\text{Eig}(A, 1)$ und finden $(\frac{1}{3}, \frac{2}{3}, \frac{2}{3})^t$ als ONB. Durch Einsetzen in das Gleichungssystem überzeugt man sich, daß $(4, -1 + 3i, -1 - 3i)^t$ eine Lösung von $(A - i\mathbf{1}_3)x = 0$ ist, und durch Normierung erhalten wir dann $(\frac{2}{3}, \frac{-1+3i}{6}, \frac{-1-3i}{6})^t$ als ONB von $\text{Eig}(A, i)$. Da A eine reelle Matrix ist, muß somit $-i$ gerade den konjugiert komplexen Vektor als Eigenvektor haben, d. h. $(\frac{2}{3}, \frac{-1-3i}{6}, \frac{-1+3i}{6})^t$ ist eine ONB von $\text{Eig}(A, -i)$.

Wir erhalten also

$$T = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{-1+3i}{6} & \frac{-1-3i}{6} \\ \frac{2}{3} & \frac{-1-3i}{6} & \frac{-1+3i}{6} \end{pmatrix} \in U(3)$$

als Transformationsmatrix mit

$$T^{-1} \circ A \circ T = T^* \circ A \circ T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{pmatrix}.$$

Bemerkung 38.26 (Spektralsatz für orthogonale Abbildungen)

Mit dem gleichen Beweis wie in Satz 38.23 kann man zeigen, daß eine orthogonale Abbildung $f \in O(V)$, deren charakteristisches Polynom über \mathbb{R} in Linearfaktoren zerfällt, ebenfalls bezüglich einer ONB diagonalisierbar ist.

Orthogonale Abbildungen lassen sich im allgemeinen aber nicht diagonalisieren, insbesondere nicht durch eine ONB. Wir haben in Beispiel 32.22 gesehen, daß die Matrix

$$T(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix},$$

die eine Drehung um den Ursprung um den Winkel α beschreibt, i.a. nicht diagonalisierbar ist. Man kann zeigen, daß dieses Beispiel im wesentlichen auch das einzige ist. Es gilt nämlich:

Ist $f \in \text{End}_{\mathbb{R}}(V)$ orthogonal, dann gibt es eindeutig bestimmte Zahlen $r, s, t \in \mathbb{N}$ sowie Winkel $\alpha_1, \dots, \alpha_t \in (0, 2\pi) \setminus \{\pi\}$ und eine ONB B von V , so daß

$$M_B^B(f) = \mathbf{1}_r \oplus -\mathbf{1}_s \oplus T(\alpha_1) \oplus \dots \oplus T(\alpha_t).$$

D) Selbstadjungierte Endomorphismen und Hauptachsentransformation

Definition 38.27 (Selbstadjungierter Endomorphismus)

$f \in \text{End}_{\mathbb{K}}(V)$ heißt *selbstadjungiert* oder *hermitesch*, wenn $f = f^*$ gilt.

Proposition 38.28 (Matrixdarstellung selbstadjungierter Endomorphismen)

Es sei $f \in \text{End}_{\mathbb{K}}(V)$ und B eine ONB von V .

Genau dann ist f selbstadjungiert, wenn $M_B^B(f)$ symmetrisch bzw. hermitesch ist.

Beweis: Aus Korollar 38.3 wissen wir, daß $M_B^B(f^*) = M_B^B(f)^*$ gilt. Deshalb gilt

$$\begin{aligned} f \text{ selbstadjungiert} &\iff f = f^* \iff M_B^B(f) = M_B^B(f^*) = M_B^B(f)^* \\ &\iff M_B^B(f) \text{ symmetrisch bzw. hermitesch.} \end{aligned}$$

□

Korollar 38.29 (Selbstadjungiert impliziert normal.)

Ist $f \in \text{End}_{\mathbb{K}}(V)$ selbstadjungiert, so ist f normal.

Beweis: Ist f selbstadjungiert, so gilt $f \circ f^* = f \circ f = f^* \circ f$, also ist f normal. □

Lemma 38.30 (Selbstadjungierte Endomorphismen haben nur reelle Eigenwerte.)

Ist $f \in \text{End}_{\mathbb{K}}(V)$ selbstadjungiert, dann ist $\chi_f \in \mathbb{R}[t]$ und χ_f zerfällt über \mathbb{R} .

Inbesondere gilt, ist $\lambda \in \sigma(f)$ ein Eigenwert von f , dann ist $\lambda \in \mathbb{R}$.

Beweis: Ist B eine ONB, dann ist $A = M_B^B(f)$ symmetrisch bzw. hermitesch und es reicht zu zeigen, daß χ_A in $\mathbb{R}[t]$ liegt und über \mathbb{R} zerfällt.

Hierfür machen wir uns zunutze, daß wir A auf alle Fälle als eine Matrix in $\text{Mat}_n(\mathbb{C})$ auffassen können und daß $A = A^*$ gilt. Über \mathbb{C} zerfällt das charakteristische Polynom von A , d. h. es gibt $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ mit

$$\chi_A = (t - \lambda_1) \cdots (t - \lambda_n).$$

Es reicht, zu zeigen, daß $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ gilt. Nun gibt es zu jedem λ_i aber einen Vektor $0 \neq x_i \in \mathbb{C}^n$ mit $Ax_i = \lambda_i x_i$. Wegen $A = A^* = \overline{A}^t$ gilt für diesen Vektor

$$\begin{aligned} \lambda_i \cdot (\overline{x_i}^t \circ x_i) &= \overline{x_i}^t \circ (\lambda_i x_i) = \overline{x_i}^t \circ (Ax_i) = \overline{x_i}^t \circ A \circ x_i \\ &= \overline{x_i}^t \circ \overline{A}^t \circ x_i = \overline{Ax_i}^t \circ x_i = \overline{\lambda_i x_i}^t \circ x_i = \overline{\lambda_i} \cdot (\overline{x_i}^t \circ x_i). \end{aligned}$$

Aus $x_i^t \circ \overline{x_i} \neq 0$ folgt dann $\lambda_i = \overline{\lambda_i}$, d. h. $\lambda_i \in \mathbb{R}$. □

Satz 38.31 (Spektralsatz für selbstadjungierte Endomorphismen)

Ein Endomorphismus $f \in \text{End}_{\mathbb{K}}(V)$ ist genau dann selbstadjungiert, wenn V eine ONB aus Eigenvektoren von f besitzt und alle Eigenwerte von f reell sind.

Beweis: Ist f selbstadjungiert, so ist f nach Korollar 38.29 normal und nach Lemma 38.30 zerfällt χ_f über \mathbb{R} in Linearfaktoren. Aus dem Spektralsatz für normale

Abbildungen folgt dann, daß V eine ONB aus Eigenvektoren von V besitzt und die Eigenwerte sind alle reell.

Besitzt umgekehrt V eine ONB $B = (x_1, \dots, x_n)$ aus Eigenvektoren von V mit reellen Eigenwerten $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, so gilt

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} = \begin{pmatrix} \overline{\lambda_1} & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \overline{\lambda_n} \end{pmatrix} = M_B^B(f)^*.$$

Mithin ist $M_B^B(f)$ symmetrisch bzw. hermitesch, und somit ist f selbstadjungiert. \square

Korollar 38.32 (Spektralsatz für symmetrische und hermitesche Matrizen)

Zu jeder symmetrischen bzw. hermiteschen Matrix $A \in \text{Mat}_n(\mathbb{K})$ gibt es eine Matrix $T \in O(n)$ bzw. $T \in U(n)$ mit

$$T^* \circ A \circ T = T^{-1} \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & 0 & \lambda_n \end{pmatrix}$$

und $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Insbesondere ist jede symmetrische bzw. hermitesche Matrix diagonalisierbar und hat nur reelle Eigenwerte.

Beweis: Die Aussage folgt aus dem Spektralsatz für selbstadjungierte Endomorphismen über \mathbb{K}^n mit kanonischem Skalarprodukt angewendet auf f_A , wenn wir $T = T_E^B$ für die dortige ONB B wählen. \square

Dies ist eine wichtige Ergänzung des Satzes über die Jordansche Normalform.

Korollar 38.33 (Hauptachsentransformationssatz für Bilinearformen)

Ist $A \in \text{Mat}_n(\mathbb{R})$ eine symmetrische Matrix, so gibt es eine ONB B von \mathbb{R}^n zum kanonischen Skalarprodukt, so daß $M_B(b_A)$ eine Diagonalmatrix ist mit den Eigenwerten von A auf der Diagonalen.

Beweis: Wählen wir die Spalten der Matrix T in Korollar 38.32 als Basis B , so ist B eine ONB von \mathbb{R}^n mit $T = T_E^B$ und die Matrix

$$M_B(b_A) = (T_E^B)^t \circ M_E(b_A) \circ T_E^B = T^{-1} \circ A \circ T$$

ist eine Diagonalmatrix, auf deren Diagonalen die Eigenwerte von A stehen. \square

Bemerkung 38.34 (Normalform symmetrischer Matrizen)

Die beiden Korollare besagen, daß reelle symmetrische Matrizen sowohl als Bilinearformen, als auch als Endomorphismen mit derselben orthogonalen Koordinatentransformation diagonalisiert werden können!

Man beachte, daß die Eigenwerte $\lambda_1, \dots, \lambda_n$ nur Invarianten des Endomorphismus sind, aber *nicht* der Bilinearform. Das ergibt sich schon aus dem Sylvesterschen Trägheitssatz 36.25.

Beispiel 38.35

Wir betrachten \mathbb{C}^n mit dem kanonischen Skalarprodukt sowie die Matrix

$$A = \begin{pmatrix} 0 & -1 & i \\ -1 & 0 & -i \\ -i & i & 0 \end{pmatrix} \in \text{Mat}_3(\mathbb{C}).$$

Da $A = A^*$ gilt, ist A hermitesch, und wir wollen mit Hilfe des Algorithmus 38.12 eine unitäre Transformationsmatrix berechnen, die A auf Diagonalgestalt transformiert.

Dazu bestimmen wir zunächst das charakteristische Polynom von A als

$$\chi_A = \begin{vmatrix} t & 1 & -i \\ 1 & t & i \\ i & -i & t \end{vmatrix} = t^3 - 3t + 2 = (t - 2) \cdot (t + 1)^2.$$

Da A diagonalisierbar ist, wissen wir nun schon, daß

$$\dim_{\mathbb{C}} \text{Eig}(A, 2) = \text{mult}(\chi_A, 2) = 1$$

und

$$\dim_{\mathbb{C}} \text{Eig}(A, -1) = \text{mult}(\chi_A, -1) = 2$$

gilt. Mit Hilfe des Gauß-Algorithmus können wir dann Basen der Eigenräume von A zu den Eigenwerten 2 und -1 berechnen. Die Rechnung wollen wir hier nicht vorführen, sondern nur das Ergebnis angeben:

$$B' = ((-i, i, -1)^t)$$

ist eine Basis von $\text{Eig}(A, 2)$ und

$$B'' = ((1, 1, 0)^t, (0, i, 1)^t)$$

ist eine Basis von $\text{Eig}(A, -1)$.

Dann müssen wir B' und B'' mittels des Gram-Schmidtschen-Orthonormalisierungsverfahrens in Orthonormalbasen der jeweiligen Eigenräume überführen. Bei B' ist das sehr einfach, da wir den einzigen Vektor in B' nur normieren müssen. Wir erhalten als einzigen Vektor in der ONB von $\text{Eig}(A, 2)$ deshalb

$$z_1 = \left(\frac{-i}{\sqrt{3}}, \frac{i}{\sqrt{3}}, \frac{-1}{\sqrt{3}} \right)^t.$$

Für $B'' = (x_2, x_3)$ ist es etwas mehr Aufwand. Wir setzen zunächst

$$z_2 = \frac{1}{\|x_2\|} \cdot x_2 = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right)^t.$$

Als nächstes setzen wir

$$y_3 = x_3 - \langle z_2, x_3 \rangle \cdot z_2 = (0, i, 1)^t - \frac{i}{\sqrt{2}} \cdot \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0 \right)^t = \left(\frac{-i}{2}, \frac{i}{2}, 1 \right)^t$$

und normieren diesen Vektor anschließend zu

$$z_3 = \frac{1}{\|y_3\|} \cdot y_3 = \left(\frac{-i}{\sqrt{6}}, \frac{i}{\sqrt{6}}, \frac{2}{\sqrt{6}} \right)^t.$$

Die Vektoren z_2 und z_3 bilden eine ONB von $\text{Eig}(A, -1)$, und $B = (z_1, z_2, z_3)$ ist somit eine ONB von \mathbb{C}^3 . Schreiben wir die Vektoren als Spalten in die Matrix T , so erhalten wir die gesuchte Transformationsmatrix

$$T = \begin{pmatrix} \frac{-i}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{6}} \\ \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} \in U(3).$$

Man rechnet folgendes leicht nach:

$$T^* \circ A \circ T = \begin{pmatrix} \frac{i}{\sqrt{3}} & \frac{-i}{\sqrt{3}} & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{i}{\sqrt{6}} & \frac{-i}{\sqrt{6}} & \frac{2}{\sqrt{6}} \end{pmatrix} \circ \begin{pmatrix} 0 & -1 & i \\ -1 & 0 & -i \\ -i & i & 0 \end{pmatrix} \circ \begin{pmatrix} \frac{-i}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{6}} \\ \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} = \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

E) Positiv definite symmetrische und hermitesche Matrizen

Eine symmetrische Bilinearform auf \mathbb{R}^n bzw. eine hermitesche Sesquilinearform auf \mathbb{C}^n ist genau dann ein Skalarprodukt, wenn sie positiv definit ist. Es ist daher außerordentlich nützlich, Kriterien für die positive Definitheit zur Hand zu haben.

Definition 38.36 (Definite Matrizen)

- Eine symmetrische Matrix $A \in \text{Mat}_n(\mathbb{R})$ heißt *positiv definit* bzw. *negativ definit* bzw. *indefinit*, wenn die Bilinearform b_A positiv bzw. negativ definit bzw. indefinit ist.
- Eine hermitesche Matrix $A \in \text{Mat}_n(\mathbb{C})$ heißt *positiv* bzw. *negativ definit* bzw. *indefinit*, wenn die Sesquilinearform b_A^s positiv bzw. negativ definit bzw. indefinit ist.
- Sei $A \in \text{Mat}_n(\mathbb{K})$ und entsteht die $k \times k$ -Untermatrix $A(k)$ von A durch Streichen der letzten $n - k$ Zeilen und Spalten, so nennen wir $A(k)$ die k -te *Hauptmatrix* von A und $\det(A(k))$ den k -ten *Hauptminor* von A .

Bemerkung 38.37 (Negativ definite Matrizen)

Man beachte, daß eine symmetrische oder hermitesche Matrix A genau dann positiv definit ist, wenn für alle $0 \neq x \in \mathbb{K}^n$

$$\bar{x}^t \circ A \circ x > 0$$

gilt, und daß A genau dann negativ definit ist, wenn

$$\bar{x}^t \circ A \circ x < 0$$

für alle $0 \neq x \in \mathbb{K}^n$ gilt.

Daraus folgt aber unmittelbar, daß A genau dann negativ definit ist, wenn $-A$ positiv definit ist. Es reicht deshalb, ein Kriterium für positive Definitheit zu finden,

um zugleich ein Kriterium für negative Definitheit zu erhalten, indem man A durch $-A$ ersetzt.

Lemma 38.38 (Positiv definite Matrizen)

Sei $A \in \text{Mat}_n(\mathbb{K})$ symmetrisch bzw. hermitesch und $T \in \text{Gl}_n(\mathbb{K})$ invertierbar. Genau dann ist A positiv definit, wenn $T^* \circ A \circ T$ positiv definit ist.

Beweis: Wir beachten, daß

$$\mathbb{K}^n = \{T \circ x \mid x \in \mathbb{K}^n\} \quad (121)$$

gilt, da T invertierbar ist. Wir erhalten deshalb

$$\begin{aligned} A \text{ ist positiv definit} &\iff \bar{x}^t \circ A \circ x > 0 \quad \forall x \in \mathbb{K}^n \\ &\stackrel{(121)}{\iff} \overline{T \circ x}^t \circ A \circ (T \circ x) > 0 \quad \forall x \in \mathbb{K}^n \\ &\iff \bar{x}^t \circ T^* \circ A \circ T \circ x > 0 \quad \forall x \in \mathbb{K}^n \\ &\iff T^* \circ A \circ T \text{ ist positiv definit} \end{aligned}$$

□

Satz 38.39 (Hurwitz-Kriterium für positive Definitheit)

Für eine symmetrische bzw. hermitesche Matrix $A \in \text{Mat}_n(\mathbb{K})$ sind die folgenden Aussagen gleichwertig:

- A ist positiv definit.
- Alle Eigenwerte von A sind positiv.
- Alle Hauptminoren von A sind positiv.

Beweis:

a. \iff b.: Wegen des Spektralsatzes für symmetrische und hermitesche Matrizen 38.32 gibt es eine orthogonale bzw. unitäre Matrix $T \in \text{Mat}_n(\mathbb{K})$, so daß

$$T^* \circ A \circ T = T^{-1} \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \quad (122)$$

gilt, wobei $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ genau die Eigenwerte von A sind.

Ist A positiv definit und ist x_i die i -te Spalte von T , so folgt aus (122)

$$\lambda_i = \bar{x}_i^t \circ A \circ x_i > 0.$$

Seien nun umgekehrt alle Eigenwerte positiv und sei $0 \neq x \in \mathbb{K}^n$ gegeben. Die Spaltenvektoren x_1, \dots, x_n von T sind eine ONB $B = (x_1, \dots, x_n)$ von \mathbb{K}^n , da T

orthogonal bzw. unitär ist. Ist $M_B(x) = (\mu_1, \dots, \mu_n)^t$ der Koordinatenvektor von x bezüglich B , so gilt

$$T \cdot M_B(x) = \sum_{i=1}^n \mu_i \cdot x_i = x$$

und

$$\begin{aligned} \bar{x}^t \cdot A \cdot x &= \overline{T \circ M_B(x)}^t \circ A \circ (T \circ M_B(x)) \\ &= \overline{M_B(x)}^t \circ (T^* \circ A \circ T) \circ M_B(x) \\ &\stackrel{(122)}{=} \sum_{i=1}^n \bar{\mu}_i \cdot \mu_i \cdot \lambda_i = \sum_{i=1}^n |\mu_i|^2 \cdot \lambda_i > 0, \end{aligned}$$

da nicht alle μ_i null sind. Damit ist die Äquivalenz von a. und b. gezeigt.

a. \implies c.: Da wir die Äquivalenz von a. und b. bereits gezeigt haben, können wir hier beide Bedingungen voraussetzen. Wegen des Spektralsatzes für symmetrische und hermitesche Matrizen 38.32 gibt es eine Matrix $T \in GL_n(\mathbb{K})$, so daß (122) erfüllt ist, und deshalb gilt

$$\det(A) = \det(T^{-1} \circ A \circ T) = \lambda_1 \cdot \dots \cdot \lambda_n > 0.$$

Jede positiv definite symmetrische oder hermitesche Matrix hat also eine positive Determinante.

Die k -te Hauptmatrix $A(k)$ beschreibt die Einschränkung der durch A definierten Bilinearform b_A bzw. Sesquilinearform b_A^s auf den Unterraum $\text{Lin}(e_1, \dots, e_k)$ von \mathbb{K}^n . Da die Einschränkung einer positiv definiten Bilinearform bzw. Sesquilinearform offenbar wieder positiv definit ist, ist mithin auch die definierende Matrix $A(k)$ positiv definit. Dann ist aber ihre Determinante, der k -te Hauptminor von A , positiv.

c. \implies a.: Wir führen den Beweis durch Induktion über n unter Ausnutzung der bereits gezeigten Äquivalenzen, wobei für $n = 1$ die Determinante $\det(A) > 0$ der einzige Eigenwert ist.

Sei also $n > 1$. Wegen des Spektralsatzes für symmetrische und hermitesche Matrizen 38.32 existiert für die symmetrische bzw. hermitesche Matrix $A(n-1)$ eine orthogonale bzw. unitäre Matrix $S \in GL_{n-1}(\mathbb{K})$, die $A(n-1)$ auf Diagonalgestalt transformiert:

$$S^{-1} \circ A(n-1) \circ S = S^* \circ A(n-1) \circ S = \bigoplus_{i=1}^{n-1} \lambda_i \mathbb{1}_1 =: D.$$

Da $A(n-1)$ die Induktionsvoraussetzung erfüllt, muß $A(n-1)$ dann positiv definit sein und somit sind die Eigenwerte $\lambda_1, \dots, \lambda_{n-1}$ von $A(n-1)$ positiv.

Wir setzen nun $T = S \oplus \mathbb{1}_1 \in \text{Gl}_n(\mathbb{K})$. Dann gilt

$$T^* \circ A \circ T = \left(\begin{array}{ccc|c} & & & \mathbf{a}_1 \\ & & & \vdots \\ & & D & \mathbf{a}_{n-1} \\ \hline \overline{\mathbf{a}_1} & \dots & \overline{\mathbf{a}_{n-1}} & \mathbf{a}_n \end{array} \right) =: B$$

für geeignete $\mathbf{a}_1, \dots, \mathbf{a}_n \in \mathbb{K}$. Setzen wir ferner $\mathbf{c}_j = -\frac{\mathbf{a}_j}{\lambda_j}$, $j = 1, \dots, n-1$, und

$$C = \left(\begin{array}{ccc|c} & & & \mathbf{c}_1 \\ & & & \vdots \\ & & \mathbb{1}_{n-1} & \mathbf{c}_{n-1} \\ \hline 0 & \dots & 0 & 1 \end{array} \right) \in \text{Gl}_n(\mathbb{K}),$$

dann folgt

$$E := (T \circ C)^* \circ A \circ (T \circ C) = C^* \circ T^* \circ A \circ T \circ C = C^* \circ B \circ C = \bigoplus_{i=1}^n \lambda_i \mathbb{1}_1,$$

wobei $\lambda_n \in \mathbb{R}$ geeignet ist. Man beachte, daß damit $\lambda_1, \dots, \lambda_n$ die Eigenwerte von E sind, und daß

$$\lambda_1 \cdot \dots \cdot \lambda_n = \det((T \circ C)^* \circ A \circ (T \circ C)) = \det(A) \cdot |\det(C \circ T)|^2 > 0,$$

da $\det(A) > 0$ der n -te Hauptminor von A ist. Da aber $\lambda_1, \dots, \lambda_{n-1}$ nach Voraussetzung positiv waren, ist dann auch λ_n positiv. E hat somit nur positive Eigenwerte und ist deshalb positiv definit. Aber mit Lemma 38.38 ist dann auch A positiv definit. □

Bemerkung 38.40 (Negativ definite und indefinite Matrizen)

Wie im Beweis von “a. \iff b.” im Beweis von Satz 38.39 sieht man:

$$A \text{ ist negativ definit} \iff A \text{ hat nur negative Eigenwerte}$$

und

$$A \text{ ist indefinit} \iff A \text{ hat einen positiven und einen negativen Eigenwert.}$$

Bemerkung 38.41 (Geometrische Interpretation der Hauptachsentransformation)

Ist $A \in \text{Mat}_n(\mathbb{R})$ eine positiv definite symmetrische Matrix mit quadratischer Form q_A , so interessieren wir uns für die Einheitssphäre zu q_A

$$S_{q_A} = \{x \in \mathbb{R}^n \mid q_A(x) = x^t \circ A \circ x = 1\}.$$

Der Satz zur Hauptachsentransformation liefert die Existenz einer orthogonalen Matrix $T \in O(n)$, so daß

$$T^t \circ A \circ T = \lambda_1 \cdot \mathbb{1}_1 \oplus \dots \oplus \lambda_n \cdot \mathbb{1}_1$$

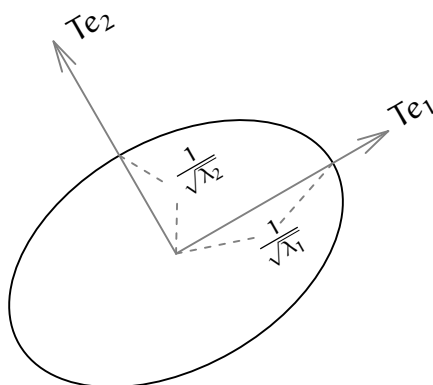


ABBILDUNG 8. Ellipse S_{q_A} mit Hauptachsen Te_1 und Te_2

eine Diagonalmatrix ist, bei der die $\lambda_i > 0$ die Eigenwerte von A sind. Die Spaltenvektoren von T sind dann ein neues orthonormales Koordinatensystem, in dem die quadratische Form die Gestalt

$$q = \lambda_1 \cdot x_1^2 + \dots + \lambda_n \cdot x_n^2$$

hat. Die Einheitssphäre zu q ist dann ein n -dimensionales Ellipsoid

$$S_q = \{x \in \mathbb{R}^n \mid q(x) = \lambda_1 \cdot x_1^2 + \dots + \lambda_n \cdot x_n^2 = 1\}.$$

Man kann sich S_{q_A} deshalb als eine Einheitskugel vorstellen, die in Richtung Te_i um den Faktor $\frac{1}{\sqrt{\lambda_i}}$ gestreckt wurde. Man nennt die Koordinatenvektoren Te_i auch die *Hauptachsen* des Ellipsoids (siehe Abbildung 8). Im Fall $n = 2$ besagt der Satz der Hauptachsentransformation dann, daß wir allein durch Drehen und Spiegeln die Ellipse S_{q_A} so bewegen können, daß ihre Hauptachsen mit den Koordinatenachsen zusammenfallen. Daher rührt der Begriff der *Hauptachsentransformation*.

F) Klassifikation der Kegelschnitte in der euklidischen Ebene

Dieser Abschnitt ist dem interessierten Leser zur Lektüre überlassen. Er wird in der Vorlesung nicht besprochen und ist nicht prüfungsrelevant.

Man kann die Überlegungen in 38.41 verallgemeinern, was wir hier im Fall $n = 2$ tun wollen. Dazu betrachten wir die Lösungsmenge einer allgemeinen quadratischen Gleichung in zwei Unbekannten. Dies führt zur Klassifikation der Kegelschnitte in der euklidischen Ebene.

Definition 38.42

Es sei $(V, \langle \cdot, \cdot \rangle)$ ein euklidischer Raum.

- Eine Abbildung $f: V \rightarrow V$ heißt eine *affine Abbildung* auf V , falls es ein $y \in V$ gibt und ein $g \in \text{End}_{\mathbb{R}}(V)$ mit $f(x) = y + g(x)$ für alle $x \in V$.
- Für $y \in V$ nennen wir die affine Abbildung

$$\tau_y: V \rightarrow V: x \mapsto x + y$$

die *Translation* um den Vektor \mathbf{y} .

- c. Eine Abbildung $f: V \rightarrow V$ heißt eine *Ähnlichkeit*, wenn es einen Vektor $\mathbf{y} \in V$ gibt und eine orthogonale Abbildung $g \in O(V)$ mit $f = \tau_{\mathbf{y}} \circ g$, d. h.

$$f(x) = \tau_{\mathbf{y}}(g(x)) = \mathbf{y} + g(x) \quad \forall x \in V.$$

- d. Ist $V = \mathbb{R}^n$ und sei $f = \tau_{\mathbf{y}} \circ g$ mit $g \in \text{End}_{\mathbb{R}}(V)$ eine bijektive affine Abbildung auf V , dann nennen wir die induzierte Abbildung

$$\mathbb{R}[t_1, \dots, t_n] \rightarrow \mathbb{R}[t_1, \dots, t_n] : p \mapsto p(f(t_1, \dots, t_n))$$

auf der Menge der Polynome in den Veränderlichen t_1, \dots, t_n einen *affinen Koordinatenwechsel* von $\mathbb{R}[t_1, \dots, t_n]$. Für die Definition von $\mathbb{R}[t_1, \dots, t_n]$ siehe auch Beispiel 40.14.

Bemerkung 38.43

- a. Jede affine Abbildung $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ läßt sich offenbar in eindeutiger Weise schreiben, als $f = \tau_{\mathbf{y}} \circ g$ mit $\mathbf{y} = f(0) \in V$ und $g \in \text{End}_{\mathbb{R}}(V)$.
- b. Ist $f = \tau_{\mathbf{y}} \circ g: \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine affine Abbildung mit $\mathbf{y} \in \mathbb{R}^n$ und $g \in \text{End}_{\mathbb{R}}(\mathbb{R}^n)$ bijektiv, dann gibt es eine eindeutig bestimmte Matrix $T \in \text{Gl}_n(\mathbb{R})$ mit $g = f_T$. Damit gilt für $p \in \mathbb{R}[t_1, \dots, t_n]$ und $\mathbf{t} = (t_1, \dots, t_n)^t$

$$p(f(t_1, \dots, t_n)) = p(T\mathbf{t} + \mathbf{y}).$$

Ist beispielsweise $p = t_1^2 + 3t_2 - 1 \in \mathbb{R}[t_1, t_2]$, $T = T\left(\frac{\pi}{2}\right)$ die Drehung um 90° und $\mathbf{y} = (2, -2)$, dann ist für $f = \tau_{\mathbf{y}} \circ f_T$

$$p(f(t_1, t_2)) = p(-t_2 + 2, t_1 - 2) = (-t_2 + 2)^2 + 3(t_1 - 2) - 1.$$

Definition 38.44

Es sei $p \in \mathbb{R}[t_1, \dots, t_n]$ dann nennen wir die Menge

$$N(p) = \{\lambda = (\lambda_1, \dots, \lambda_n)^t \in \mathbb{R}^n \mid p(\lambda) = 0\}$$

eine *algebraische Hyperfläche* von \mathbb{R}^n . Ist $\deg(p) = d$, so nennen wir d auch den *Grad* der Hyperfläche. Ist $n = 2$, so sprechen wir auch von *algebraischen Kurven* statt von algebraischen Hyperflächen.

Definition 38.45

Wir definieren auf $\mathbb{R}[t_1, \dots, t_n]$ eine Relation durch

$$p \equiv q \quad :\Leftrightarrow \quad \exists c \in \mathbb{R} \setminus \{0\} : p = c \cdot q$$

für $p, q \in \mathbb{R}[t_1, \dots, t_n]$. Wir nennen p und q mit $p \equiv q$ auch *äquivalent*.

Bemerkung 38.46

Man sieht sofort, daß \equiv eine Äquivalenzrelation auf $\mathbb{R}[t_1, \dots, t_n]$ definiert.

Ferner gilt offensichtlich, daß für zwei äquivalente Polynome $p, q \in \mathbb{R}[t_1, \dots, t_n]$ auch $N(p) = N(q)$ gilt. Interessiert man sich also nur für das Nullstellengebilde

von p , so kann man p getrost durch ein äquivalentes Polynom ersetzen und somit erreichen, daß der konstante Anteil von p entweder 0 oder -1 ist.

Im Folgenden interessieren wir uns nur noch für algebraische Kurven vom Grad zwei.

Bemerkung 38.47

Ist $p \in \mathbb{R}[t_1, t_2]$ ein allgemeines Polynom zweiten Grades, dann gibt es reelle Zahlen $\alpha_{11}, \alpha_{12} = \alpha_{21}, \alpha_{22}, \alpha_1, \alpha_2, \alpha \in \mathbb{R}$ so, daß

$$p = \alpha_{11}t_1^2 + 2\alpha_{12}t_1t_2 + \alpha_{22}t_2^2 + \alpha_1t_1 + \alpha_2t_2 + \alpha = \langle t, St \rangle + \langle a, t \rangle + \alpha,$$

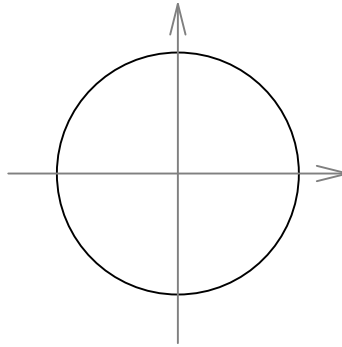
wobei $t = (t_1, t_2)^t$, $0 \neq S = (\alpha_{ij})_{i,j \in \{1,2\}} \in \text{Mat}_2(\mathbb{R})$ und $a = (\alpha_1, \alpha_2)^t$.

Beispiel 38.48

Für $S = \mathbb{1}_2$, $a = (0, 0)^t$ und $\alpha = -1$ erhalten wir $p = t_1^2 + t_2^2 - 1$, und die Nullstellenmenge davon,

$$N(t_1^2 + t_2^2 - 1) = \{\lambda = (\lambda_1, \lambda_2)^t \in \mathbb{R}^2 \mid \lambda_1^2 + \lambda_2^2 = 1\},$$

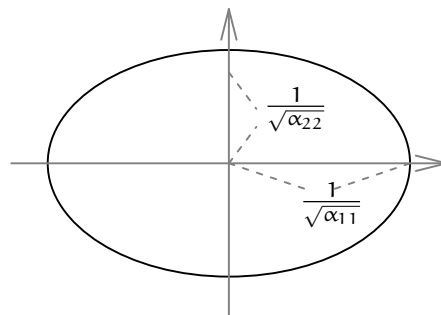
ist offenbar der Einheitskreis.



Ist S eine Diagonalmatrix mit positiven Diagonaleinträgen, d. h. $\alpha_{11}, \alpha_{22} > 0$ und $\alpha_{12} = \alpha_{21} = 0$, und ist ferner $a = (0, 0)^t$ und $\alpha = -1$, dann erhalten wir als Nullstellengebilde von p

$$N\left(\left(\sqrt{\alpha_{11}}t_1\right)^2 + \left(\sqrt{\alpha_{22}}t_2\right)^2 - 1\right) = \left\{(\lambda_1, \lambda_2)^t \in \mathbb{R}^2 \mid \left(\sqrt{\alpha_{11}}\lambda_1\right)^2 + \left(\sqrt{\alpha_{22}}\lambda_2\right)^2 = 1\right\}$$

eine Ellipse.



Satz 38.49

Es sei

$$p = \langle t, St \rangle + \langle a, t \rangle + \alpha \in \mathbb{R}[t_1, t_2] \tag{123}$$

ein Polynom zweiten Grades mit symmetrischer Matrix $0 \neq S = (\alpha_{ij}) \in \text{Mat}_2(\mathbb{R})$. Dann gibt es eine affine Koordinatentransformation mittels einer Ähnlichkeit $f = \tau_y \circ f_T$ von \mathbb{R}^2 mit $T \in \text{SO}(2)$, so daß $q := p(f(t_1, t_2))$ äquivalent zu einer der folgenden Normalformen ist:

I: $\det(S) > 0$.

I.1: $\alpha \neq 0$ und $\alpha_{11} > 0$. Dann ist $q \equiv (\lambda_1 t_1)^2 + (\lambda_2 t_2)^2 - 1$ und $N(q)$ ist eine Ellipse.

I.2: $\alpha \neq 0$ und $\alpha_{11} < 0$. Dann ist $q \equiv (\lambda_1 t_1)^2 + (\lambda_2 t_2)^2 + 1$ und $N(q)$ ist die leere Menge.

I.3: $\alpha = 0$. Dann ist $q \equiv (\lambda_1 t_1)^2 + (\lambda_2 t_2)^2$ und $N(q)$ ist ein Punkt.

II: $\det(S) < 0$.

II.1: $\alpha \neq 0$. Dann ist $q \equiv (\lambda_1 t_1)^2 - (\lambda_2 t_2)^2 - 1$ und $N(q)$ ist eine Hyperbel.

II.2: $\alpha = 0$. Dann ist $q \equiv (\lambda_1 t_1)^2 - (\lambda_2 t_2)^2$ und $N(q)$ besteht aus zwei verschiedenen Geraden durch den Ursprung.

III: $\det(S) = 0$, $\alpha \neq (0, 0)^t$. Dann ist $q \equiv t_1^2 - \lambda t_2$ und $N(q)$ ist eine Parabel.

IV: $\det(S) = 0$, $\alpha = (0, 0)^t$.

IV.1: $\alpha \neq 0$ und S hat einen positiven Eigenwert. Dann ist $q \equiv t_1^2 - \lambda$, $\lambda > 0$, und $N(q)$ besteht aus zwei parallelen Geraden.

IV.2: $\alpha \neq 0$ und S hat einen negativen Eigenwert. Dann ist $q \equiv t_1^2 + \lambda$, $\lambda > 0$, und $N(q)$ ist die leere Menge.

IV.3: $\alpha = 0$. Dann ist $q \equiv t_1^2$ und $N(q)$ besteht aus einer Doppelgeraden, d. h. einer Geraden, die man doppelt zählt.

Bemerkung 38.50

Dies ist die vollständige Klassifikation der Kurven zweiten Grades. Sie heißen auch *Kegelschnitte*, da alle, bis auf die Fälle I.2, IV.1 und IV.2 als Schnitt des Kreiskegels

$$N(t_1^2 + t_2^2 - t_3^2) \subset \mathbb{R}^3$$

mit einer geeigneten Ebene im \mathbb{R}^3 realisierbar sind (siehe Abbildung 9).

I.1 besagt, daß sich jede Ellipse durch Translation und Drehung so bewegen läßt, daß die Hauptachsen der Ellipse mit den Koordinatenachsen übereinstimmen. Daher kommt der Name Hauptachsentransformation.

Beweis von Satz 38.49:

1. Fall: $\alpha = (0, 0)^t$: Wir betrachten zunächst den Fall $\alpha = (0, 0)^t$.

Nach dem Satz über die Hauptachsentransformation 38.32 existiert ein $T \in \text{SO}(2)$, so daß

$$T^t \circ S \circ T = T^{-1} \circ S \circ T = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}.$$

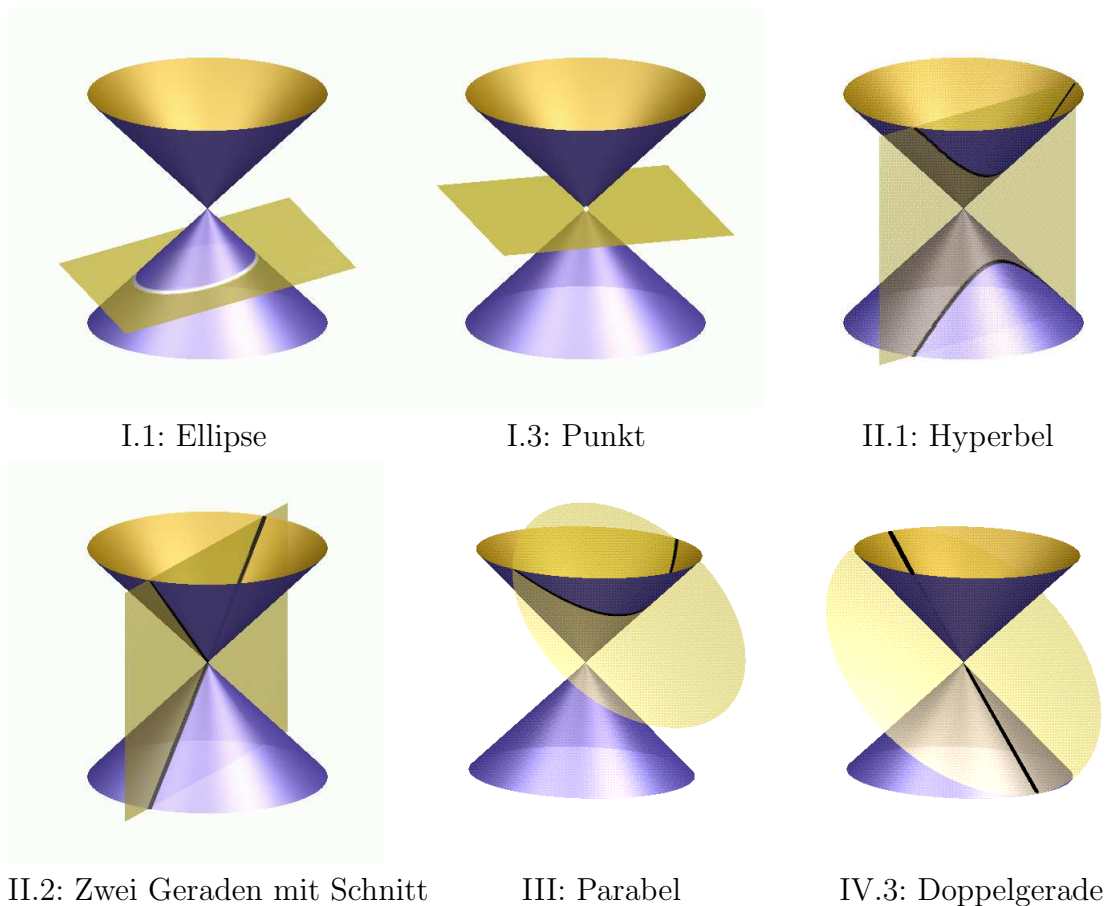


ABBILDUNG 9. Kegelschnitte

Man beachte noch, daß nicht beide Eigenwerte μ_1 und μ_2 null sein können, da $S \neq 0$. Also können wir o. E. annehmen, daß $\mu_1 \neq 0$ und daß $\mu_1 \geq \mu_2$ gilt, falls $\mu_2 \neq 0$.

Die lineare Abbildung $f_T : \mathbb{R}^2 \rightarrow \mathbb{R}^2 : x \mapsto Tx$ ist eine Drehung und es gilt

$$\begin{aligned} p(Tt) &= \langle Tt, (S \circ T)t \rangle + \alpha \\ &= \langle t, (T^t \circ S \circ T)t \rangle + \alpha \\ &= \mu_1 t_1^2 + \mu_2 t_2^2 + \alpha. \end{aligned}$$

Da wir p ohnehin nur bis auf Äquivalenz klassifizieren wollen, können wir o. E. annehmen, daß $\alpha = 0$ oder $\alpha = -1$ gilt. Setzen wir nun noch $\lambda_i = \sqrt{|\mu_i|}$, dann erhalten wir folgende Fälle.

Fall 1.1: $\mu_1, \mu_2 > 0$: Dies ist gleichbedeutend dazu, daß S positiv definit ist, und nach dem Hauptminorenkriterium dazu, daß $\det(S) > 0$ und $\alpha_{11} > 0$. Ist $\alpha = -1$, so sind wir im Fall I.1, und ist $\alpha = 0$, so sind wir Fall I.3.

Fall 1.2: $\mu_1, \mu_2 < 0$: Dies ist gleichbedeutend dazu, daß $-S$ positiv definit ist, daß also $\det(S) = \det(-S) > 0$ und $-\alpha_{11} > 0$. Ist $\alpha = -1$, so sind wir im Fall I.2, und für

$\alpha = 0$ wieder im Fall I.3, da wir dann das Polynom nochmals mit -1 multiplizieren können, um ein äquivalentes der gesuchten Form zu erhalten.

Fall 1.3: $\mu_1 > 0, \mu_2 < 0$: Dies ist gleichbedeutend dazu, daß $\mu_1 \cdot \mu_2 = \det(S) < 0$ ist. Im Fall $\alpha = -1$ führt dies zu Fall II.1, und im Fall $\alpha = 0$ führt es zu Fall II.2.

Fall 1.4: $\mu_1 > 0, \mu_2 = 0$ oder $\mu_1 < 0, \mu_2 = 0$: Das ist dann gleichbedeutend dazu, daß $\det(S) = 0$ ist. Für $\mu_1 > 0$ und $\alpha = -1$ erhalten wir Fall IV.1, für $\mu_1 < 0$ und $\alpha = -1$ den Fall IV.2, und für $\alpha = 0$ in den Fall IV.3.

2. Fall: $\mathbf{a} \neq (0, 0)^t$: Sind wir im Fall $\mathbf{a} = (0, 0)^t$ noch ohne Translation ausgekommen, so werden wir jetzt doch Translationen betrachten müssen.

Für $\mathbf{c} \in \mathbb{R}^2$ bewirkt die Translation $\tau_{\mathbf{c}} : \mathbb{R}^2 \rightarrow \mathbb{R}^2 : \mathbf{x} \mapsto \mathbf{x} + \mathbf{c}$ folgende Koordinatentransformation für \mathfrak{p}

$$\begin{aligned} \mathfrak{p}(\mathbf{t} + \mathbf{c}) &= \langle \mathbf{t} + \mathbf{c}, S\mathbf{t} + S\mathbf{c} \rangle + 2\langle \mathbf{a}, \mathbf{t} + \mathbf{c} \rangle + \alpha \\ &= \langle \mathbf{t}, S\mathbf{t} \rangle + 2\langle \mathbf{a} + S\mathbf{c}, \mathbf{t} \rangle + \langle \mathbf{c}, S\mathbf{c} \rangle + 2\langle \mathbf{a}, \mathbf{c} \rangle + \alpha \\ &= \langle \mathbf{t}, S\mathbf{t} \rangle + 2\langle \mathbf{b}, \mathbf{t} \rangle + \beta, \end{aligned} \quad (124)$$

wenn wir $\mathbf{b} = \mathbf{a} + S\mathbf{c}$ und $\beta = \langle \mathbf{c}, S\mathbf{c} \rangle + 2\langle \mathbf{a}, \mathbf{c} \rangle + \alpha$ setzen.

Fall 2.1: $\exists \mathbf{c} \in \mathbb{R}^2 : \mathbf{b} = \mathbf{a} + S\mathbf{c} = (0, 0)^t$: Dann haben wir \mathfrak{p} durch $\mathfrak{p}(\tau_{\mathbf{c}}(\mathbf{t}))$ auf den ersten Fall " $\mathbf{a} = (0, 0)^t$ " zurückgeführt. Es gibt also ein $T \in SO(2)$, so daß $\mathfrak{q} = \mathfrak{p}((\tau_{\mathbf{c}} \circ f_T)(\mathbf{t}))$ äquivalent zu einem der Fälle I, II oder IV ist.

Fall 2.2: $\forall \mathbf{c} \in \mathbb{R}^2 : \mathbf{b} = \mathbf{a} + S\mathbf{c} \neq (0, 0)^t$: Aus Lemma 38.51 folgt, daß es ein $\mathbf{c} \in \mathbb{R}^2$ gibt mit $S\mathbf{b} = S^2\mathbf{c} + S\mathbf{a} = 0$. Setzen wir nun noch $\delta := -\frac{\beta}{2\langle \mathbf{b}, \mathbf{b} \rangle}$, dann gilt für die Translation $\tau_{\mathbf{c} + \delta\mathbf{b}}$ ⁵

$$\begin{aligned} \mathfrak{p}(\mathbf{t} + \mathbf{c} + \delta\mathbf{b}) &= \langle \mathbf{t}, S\mathbf{t} \rangle + 2\langle \mathbf{a} + S(\mathbf{c} + \delta\mathbf{b}), \mathbf{t} \rangle + \langle \mathbf{c} + \delta\mathbf{b}, S(\mathbf{c} + \delta\mathbf{b}) \rangle + 2\langle \mathbf{a}, \mathbf{c} + \delta\mathbf{b} \rangle + \alpha \\ &= \langle \mathbf{t}, S\mathbf{t} \rangle + 2\langle \mathbf{b} + \delta S\mathbf{b}, \mathbf{t} \rangle + \delta^2\langle \mathbf{b}, S\mathbf{b} \rangle + 2\delta\langle \mathbf{b}, \mathbf{b} \rangle + \beta \\ &= \langle \mathbf{t}, S\mathbf{t} \rangle + 2\langle \mathbf{b}, \mathbf{t} \rangle + 2\delta\langle \mathbf{b}, \mathbf{b} \rangle + \beta \\ &= \langle \mathbf{t}, S\mathbf{t} \rangle + 2\langle \mathbf{b}, \mathbf{t} \rangle. \end{aligned}$$

Beachtet man, daß, wegen $S\mathbf{b} = 0$, Null auf alle Fälle ein Eigenwert von S ist und daß $S \neq 0$, so folgt aus dem Satz über Hauptachsentransformation 38.32 die Existenz eines $T \in SO(2)$, so daß

$$D := T^t \circ S \circ T = T^{-1} \circ S \circ T = \begin{pmatrix} \mu_1 & 0 \\ 0 & 0 \end{pmatrix},$$

wobei $\mu_1 \neq 0$. Insbesondere sind wir also in dem Fall $\det(S) = 0$.

⁵Man setze zunächst in der Gleichung (124) für \mathbf{c} den Wert $\mathbf{c} + \delta\mathbf{b}$ ein. Dann ziehe man die Skalarprodukte auseinander und gruppieren sie neu, so daß man $\mathbf{b} = \mathbf{a} + S\mathbf{c}$, $S\mathbf{b} = 0$ sowie die Definition von β verwenden kann. Man beachte auch, daß S symmetrisch, also selbstadjungiert, ist.

Ferner gilt für $T^t \mathbf{b} =: (\mu, \lambda)^t$ unter Berücksichtigung, daß $T^t = T^{-1}$,

$$(\mu_1 \mu, 0) = (T^t \circ S \circ T) \circ (T^t \mathbf{b}) = T^t \circ (S \mathbf{b}) = 0,$$

und mithin ist $T^t \mathbf{b} = (0, \lambda)^t$, wobei $\lambda \neq 0$, da T^t invertierbar und $\mathbf{b} \neq (0, 0)^t$. Aber dann überführt $t \mapsto Tt$ das Polynom $\langle t, St \rangle + 2\langle \mathbf{b}, t \rangle$ in das Polynom

$$\langle Tt, (S \circ T)t \rangle + 2\langle \mathbf{b}, Tt \rangle = \langle t^t, Dt \rangle + 2\langle T^t \mathbf{b}, t \rangle = \mu_1 t_1^2 + 2\lambda t_2.$$

D. h. dann aber, daß

$$q := p((\tau_{c+\delta b} \circ f_T)(t)) = \mu_1 t_1^2 + 2\lambda t_2,$$

und damit sind wir genau im Fall III. □

Lemma 38.51

Ist $S \in \text{Mat}_n(\mathbb{R})$ symmetrisch, so gilt für die lineare Abbildung $f_S: \mathbb{R}^n \rightarrow \mathbb{R}^n$.

- a. $\text{Ker}(f_S^2) = \text{Ker}(f_S)$ und $\text{Im}(f_S^2) = \text{Im}(f_S)$.
- b. Zu jedem $\mathbf{a} \in \mathbb{R}^n$ existiert ein $\mathbf{c} \in \mathbb{R}^n$, so daß $S^2 \mathbf{c} + S \mathbf{a} = 0$.

Beweis: a. Für $\mathbf{x} \in \text{Ker}(f_S^2)$ ergibt sich aus

$$0 = \langle \mathbf{x}, S^2 \mathbf{x} \rangle = \langle S \mathbf{x}, S \mathbf{x} \rangle,$$

also $f_S(\mathbf{x}) = S \mathbf{x} = 0$ und $\mathbf{x} \in \text{Ker}(f_S)$. Die umgekehrte Inklusion ist klar.

Wir wissen bereits, daß $\text{Im}(f_S) \supseteq \text{Im}(f_S^2)$ gilt. Da nun ferner

$$\begin{aligned} \dim_{\mathbb{R}}(\text{Im}(f_S)) &= n - \dim_{\mathbb{R}}(\text{Ker}(f_S)) \\ &= n - \dim_{\mathbb{R}}(\text{Ker}(f_S^2)) = \dim_{\mathbb{R}}(\text{Im}(f_S^2)) \end{aligned}$$

gilt, folgt also die Gleichheit.

- b. Es gilt für $\mathbf{a} \in \mathbb{R}^n$, daß $S(-\mathbf{a}) = f_S(-\mathbf{a}) \in \text{Im}(f_S) = \text{Im}(f_S^2)$, also gibt es nach a. ein $\mathbf{c} \in \mathbb{R}^n$ mit $S^2 \mathbf{c} + S \mathbf{a} = f_S^2(\mathbf{c}) - f_S(-\mathbf{a}) = 0$. □

Aufgaben

Aufgabe 38.52

Zeige, wenn für $f \in \text{End}_{\mathbb{K}}(V)$ stets $\|f(\mathbf{x})\| = \|\mathbf{x}\|$ gilt, so ist f orthogonal bzw. unitär.

Aufgabe 38.53 (Lineare Funktionale)

Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum.

Dann gibt es für jedes $g \in \text{Hom}_{\mathbb{K}}(V, \mathbb{K})$ genau ein $\mathbf{y} \in V$, so daß für alle $\mathbf{x} \in V$ gilt

$$g(\mathbf{x}) = \langle \mathbf{y}, \mathbf{x} \rangle.$$

Aufgabe 38.54 (Die adjungierte Abbildung)

Seien V und W zwei endlich-dimensionale euklidische oder unitäre Räume mit Skalarprodukten $\langle \cdot, \cdot \rangle_V$ und $\langle \cdot, \cdot \rangle_W$. Dann gibt es zu jeder linearen Abbildung $f : V \rightarrow W$ genau eine lineare Abbildung $f^* : W \rightarrow V$, so daß

$$\langle f(x), y \rangle_W = \langle x, f^*(y) \rangle_V \quad (125)$$

für alle $x \in V$ und $y \in W$. Die Abbildung f^* heißt die *adjungierte Abbildung* zu f .

Aufgabe 38.55 (Matrixdarstellung der adjungierten Abbildung)

Seien V und W zwei endlich-dimensionale euklidische oder unitäre Räume mit Orthonormalbasen B bzw. D . Dann gilt für jede \mathbb{K} -lineare Abbildung $f : V \rightarrow W$

$$M_B^D(f^*) = (M_D^B(f))^*,$$

d.h. die Matrixdarstellung der adjungierten Abbildung ist die Adjungierte der Matrixdarstellung.

Aufgabe 38.56

Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum.

Zeige, ist $f \in \text{End}_{\mathbb{K}}(V)$ normal, so gelten

$$\text{Ker}(f) = \text{Ker}(f^*)$$

und

$$V = \text{Ker}(f) \perp \text{Im}(f).$$

Aufgabe 38.57

Sei V ein endlich-dimensionaler unitärer Raum und $f \in \text{End}_{\mathbb{C}}(V)$ normal.

Zeige, es gibt ein Polynom $p \in \mathbb{C}[t]$ mit $f^* = p(f)$.

Aufgabe 38.58

Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $f \in \text{End}_{\mathbb{K}}(V)$ bijektiv. Zeige, die folgenden Aussagen sind äquivalent:

- Für $x, y \in V$ mit $x \perp y$ gilt $f(x) \perp f(y)$.
- Für $x, y \in V$ mit $\|x\| = \|y\|$ gilt $\|f(x)\| = \|f(y)\|$.
- Es gibt ein $\lambda \in \mathbb{R}_{>0}$ und ein $g \in O(V)$ bzw. $g \in U(V)$ mit $f = \lambda g$.

Aufgabe 38.59

Es sei $V \neq 0$ ein endlich-dimensionaler unitärer Raum und $f \in \text{End}_{\mathbb{C}}(V)$. Zeige, die folgenden Aussagen sind gleichwertig:

- $f^* = -f$.
- Für alle $x \in V$ gilt: $\langle f(x), x \rangle \in i\mathbb{R}$.
- Es gibt eine Orthonormalbasis von V aus Eigenvektoren von f und der Realteil aller Eigenwerte ist Null.

Aufgabe 38.60

Überprüfe, ob die folgende symmetrische Matrix $A \in \text{Mat}_3(\mathbb{R})$ positiv definit ist:

$$A = \begin{pmatrix} 9 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

Aufgabe 38.61

Bestimme eine orthogonale Matrix $T \in O(3)$, die die folgende symmetrische Matrix A diagonalisiert:

$$A = \begin{pmatrix} 5 & 2 & 2 \\ 2 & 2 & -4 \\ 2 & -4 & 2 \end{pmatrix}.$$

Ist die Matrix A positiv definit?

Mehrdimensionale Analysis

Mehrdimensionale Analysis ist in dieser Vorlesung im wesentlichen Analysis auf dem \mathbb{R}^n . Wir werden die Konzepte aber etwas allgemeiner auf beliebigen (nicht notwendigerweise endlich-dimensionalen) \mathbb{R} -Vektorräumen einführen, und uns immer dann auf den endlich-dimensionalen Fall zurückziehen, wenn dies geboten scheint.

§ 39 Topologische Grundbegriffe in metrischen und in normierten Räumen

Wir wollen in diesem Abschnitt die Grundlagen dafür legen, die Analysis über \mathbb{R} zu erweitern auf den \mathbb{R}^n . Dafür ist es wesentlich, den Begriff des Abstands $|x - y|$ zweier Zahlen x und y zu erweitern und allgemein den Abstand zweier Punkte x und y im n -dimensionalen Raum \mathbb{R}^n zu definieren. Eigentlich haben wir das schon in Abschnitt 37 im Zusammenhang mit den euklidischen Räumen getan, aber ein wenig Überlegung zeigt, daß es im \mathbb{R}^n oder allgemeiner noch in einem beliebigen \mathbb{R} -Vektorraum mehr als eine Möglichkeit gibt, den Abstandsbegriff $|x - y|$ zu verallgemeinern. Dies ist Gegenstand der folgenden Überlegungen.

A) Metrische Räume

Definition 39.1 (Metrischer Raum)

Eine Abbildung

$$d : M \times M \longrightarrow \mathbb{R}_{\geq 0},$$

die den folgenden Axiomen genügt, heißt eine *Metrik* oder ein *Abstand* auf der Menge M :

- (1) $d(x, y) = 0 \iff x = y$, (“Definitheit”)
- (2) $d(x, y) = d(y, x)$, (“Symmetrie”)
- (3) $d(x, z) \leq d(x, y) + d(y, z)$, (“Dreiecksungleichung”)

wobei $x, y, z \in M$ gelten soll. Das Tupel (M, d) heißt dann ein *metrischer Raum*.

Die anschauliche Bedeutung der Axiome ist offensichtlich. Zwei verschiedene Punkte sollen einen positiven Abstand haben, der Abstand von x zu y soll sich nicht von dem von y zu x unterscheiden, und schließlich soll der Weg von x zu z über y nicht kürzer als der direkte Weg von x zu z sein.

Beispiel 39.2 (Eisenbahnmetrik)

Ist M eine beliebige Menge, so wird durch

$$d : M \times M \longrightarrow \mathbb{R}_{\geq 0} : (x, y) \mapsto \begin{cases} 0, & \text{falls } x = y, \\ 1, & \text{falls } x \neq y \end{cases}$$

eine Metrik auf M definiert, die man auch die *Eisenbahnmetrik* nennt.

Metriken wie die in Beispiel 39.2 sind pathologisch und sollen uns im folgenden nicht weiter interessieren. Wir werden ausschließlich Teilmengen von Vektorräumen betrachten und Metriken, die sich vom Begriff der Länge, d.h. der Norm, eines Vektors ableiten.

Bemerkung 39.3 (Normierte Räume)

Wir erinnern uns (siehe Definition 37.6), daß eine Abbildung

$$\|\cdot\| : V \longrightarrow \mathbb{R}_{\geq 0},$$

eine *Norm* auf V heißt, falls für alle $x, y \in V$ und $\lambda \in K$ folgendes gilt:

- (1) $\|x\| = 0 \iff x = 0$, (“Positive Definitheit”)
- (2) $\|\lambda x\| = |\lambda| \cdot \|x\|$ und (“Homogenität”)
- (3) $\|x + y\| \leq \|x\| + \|y\|$. (“Dreiecksungleichung”)

Man beachte, daß jede Teilmenge eines metrischen Raumes mit der Einschränkung der Metrik wieder ein metrischer Raum ist. Für normierte Räume gilt dies nicht, da die meisten Teilmengen keine Vektorräume sind. Das ist ein wesentlicher Vorzug des Begriffs metrischer Raum, auch wenn wir uns im folgenden auf Metriken beschränken wollen, die von Normen herrühren.

Lemma 39.4 (Normierte Räume sind metrische Räume.)

Ist $(V, \cdot, +, \|\cdot\|)$ ein normierter Raum $M \subseteq V$, so wird durch

$$d : M \times M \longrightarrow \mathbb{R}_{\geq 0} : (x, y) \mapsto \|x - y\|$$

eine Metrik auf M definiert.

Beweis: Seien $x, y, z \in M$, dann gilt

- (1) $0 = d(x, y) = \|x - y\| \iff x - y = 0 \iff x = y$,
- (2) $d(x, y) = \|x - y\| = \|(-1) \cdot (y - x)\| = |-1| \cdot \|y - x\| = d(y, x)$, und
- (3) $d(x, z) = \|x - z\| = \|(x - y) + (y - z)\| \leq \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$.

□

Wir werden im folgenden auf einer Teilmenge eines normierten Raums stets die in Lemma 39.4 definierte Metrik zur Abstandsberechnung verwenden. Außerdem werden wir die Norm immer mit $\|\cdot\|$ bezeichnen, sofern nichts anderes angegeben ist, und wir sprechen dann einfach vom normierten Raum V . Jede Teilmenge eines metrischen Raumes betrachten wir mit der Einschränkung der Metrik selbst als metrischen Raum.

Beispiel 39.5 (Normierte Räume)

- a. Die *euklidische Norm* auf dem \mathbb{R}^n

$$\|\cdot\|_2 : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geq 0} : (x_1, \dots, x_n)^t \mapsto \sqrt{x_1^2 + \dots + x_n^2}$$

haben wir bereits in Satz 37.9 betrachtet.

- b. Die *Summennorm* auf dem \mathbb{R}^n wird definiert durch

$$\|\cdot\|_1 : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geq 0} : (x_1, \dots, x_n)^t \mapsto |x_1| + \dots + |x_n|.$$

Die Definitheit und die Homogenität der Summennorm sind dabei offensichtlich, und die Dreiecksungleichung ergibt sich aus der Dreiecksungleichung des Betrages.

- c. Die *Maximumsnorm* auf dem \mathbb{R}^n wird definiert durch

$$\|\cdot\|_\infty : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geq 0} : (x_1, \dots, x_n)^t \mapsto \max\{|x_1|, \dots, |x_n|\}.$$

- d. Auf $\mathbb{R} = \mathbb{R}^1$ stimmen die euklidische Norm, die Summennorm und die Maximumsnorm alle mit dem Absolutbetrag überein. Wir werden \mathbb{R} stets mit dem Absolutbetrag als normierten Raum betrachten.

- e. Auch auf $V = \mathcal{C}([0, 1], \mathbb{R})$ erhalten wir aus Satz 37.9 eine euklidische Norm

$$\|\cdot\|_{L_2} : \mathcal{C}([0, 1], \mathbb{R}) \longrightarrow \mathbb{R}_{\geq 0} : f \mapsto \sqrt{\int_0^1 f^2(x) \, dx}.$$

- f. Analog zu Teil c. kann man auf $V = \mathcal{C}([0, 1], \mathbb{R})$ auch eine *Maximumsnorm* definieren durch

$$\|\cdot\|_\infty : \mathcal{C}([0, 1], \mathbb{R}) \longrightarrow \mathbb{R}_{\geq 0} : f \mapsto \max\{|f(x)| \mid x \in [0, 1]\}.$$

Man beachte, daß das Maximum nach Satz 14.16 auch angenommen wird.

B) Topologische Grundbegriffe

Definition 39.6 (ε -Umgebung)

Sei M ein metrischer Raum, $a \in M$ und $\varepsilon > 0$. Wir nennen

$$U_\varepsilon(a) := \{x \in M \mid d(x, a) < \varepsilon\}$$

die ε -Umgebung von a in M .

Beispiel 39.7

Ist V ein normierter Raum und $\mathbf{a} \in V$, so ist

$$U_\varepsilon(\mathbf{a}) = \{\mathbf{x} \in V \mid \|\mathbf{x} - \mathbf{a}\| < \varepsilon\}.$$

Für $V = \mathbb{R}$ stimmt dies mit der Definition in Bemerkung 13.2 überein.

In Abbildung 1 ist die ε -Umgebung von $\mathbf{a} = (0, 0)$ im \mathbb{R}^2 für die Summennorm, die euklidische Norm und die Maximumsnorm gezeigt, wobei der schwarze Rand der grauen Gebiete nicht zur ε -Umgebung gehört.

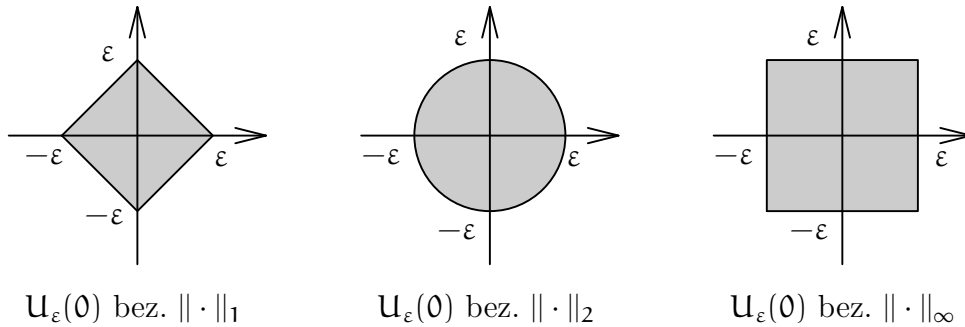


ABBILDUNG 1. Die ε -Umgebung des Ursprungs bez. $\|\cdot\|_1$, $\|\cdot\|_2$ und $\|\cdot\|_\infty$

Proposition 39.8 (Metrische Räume sind hausdorffsch.)

Ist M ein metrischer Raum und sind $x, y \in M$ mit $x \neq y$, so gibt es ein $\varepsilon > 0$ mit

$$U_\varepsilon(x) \cap U_\varepsilon(y) = \emptyset.$$

Beweis: Wähle $\varepsilon = \frac{d(x,y)}{2}$, so würde für $z \in U_\varepsilon(x) \cap U_\varepsilon(y)$

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

gelten, was ein offensichtlicher Widerspruch ist. □

Definition 39.9 (Offene und abgeschlossene Mengen)

Es sei M ein metrischer Raum, $U, O, A \subseteq M$ und $\mathbf{a} \in M$.

- a. \mathbf{a} heißt ein *innerer Punkt* von U , falls es ein $\varepsilon > 0$ gibt mit $U_\varepsilon(\mathbf{a}) \subseteq U$.
- b. O heißt *offen* in M , falls jeder Punkt von O ein innerer Punkt ist.
- c. A heißt *abgeschlossen* in M , falls $M \setminus A$ offen in M ist.

Beispiel 39.10

- a. Offene Intervalle (\mathbf{a}, \mathbf{b}) für $\mathbf{a}, \mathbf{b} \in \mathbb{R} \cup \{-\infty, \infty\}$ sind offen in \mathbb{R} .
- b. Abgeschlossene Intervalle $[\mathbf{a}, \mathbf{b}]$ für $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ sind abgeschlossen in \mathbb{R} .
- c. Ein halboffenes Intervall $[\mathbf{a}, \mathbf{b})$ mit $\mathbf{a}, \mathbf{b} \in \mathbb{R}$ ist weder offen noch abgeschlossen, da \mathbf{a} kein innerer Punkt von $[\mathbf{a}, \mathbf{b})$ und \mathbf{b} kein innerer Punkt von $\mathbb{R} \setminus [\mathbf{a}, \mathbf{b})$ ist. Analog ist auch $(\mathbf{a}, \mathbf{b}]$ weder offen noch abgeschlossen.
- d. Intervalle der Form $(-\infty, \mathbf{a}]$ und $[\mathbf{a}, \infty)$ für $\mathbf{a} \in \mathbb{R}$ sind abgeschlossen.

- e. Ist M ein metrischer Raum, so sind M und \emptyset offen und abgeschlossen.
- f. Ist M ein metrischer Raum, $U \subseteq W \subseteq M$ und a ein innerer Punkt von U , dann ist a auch ein innerer Punkt von W , denn aus $U_\varepsilon(a) \subseteq U$ folgt $U_\varepsilon(a) \subseteq W$.
- g. \mathbb{Q} ist weder offen, noch abgeschlossen in \mathbb{R} , da jede ε -Umgebung einer rationalen Zahl sowohl weitere rationale Zahlen, als auch irrationale Zahlen enthält.

Lemma 39.11 (ε -Umgebungen sind offen.)

Sei M ein metrischer Raum, $a \in M$ und $\varepsilon > 0$. Dann ist $U_\varepsilon(a)$ offen in M .

Beweis: Für $b \in U_\varepsilon(a)$ setzen wir $\delta := \varepsilon - d(a, b) > 0$. Ist nun $x \in U_\delta(b)$, so gilt

$$d(x, a) \leq d(x, b) + d(b, a) < \delta + d(a, b) = \varepsilon.$$

Mithin ist $x \in U_\varepsilon(a)$ und somit $U_\delta(b) \subseteq U_\varepsilon(a)$, so daß b ein innerer Punkt von $U_\varepsilon(a)$ ist. \square

Proposition 39.12 (Vereinigungen und Schnitte offener / abgeschlossener Mengen)

Sei M ein metrischer Raum.

- Endliche Durchschnitte offener Mengen sind offen in M .
- Beliebige Vereinigungen offener Mengen sind offen in M .
- Endliche Vereinigungen abgeschlossener Mengen sind abgeschlossen in M .
- Beliebige Durchschnitte abgeschlossener Mengen sind abgeschlossen in M .

Beweis:

- a. Seien $O, O' \subseteq M$ zwei offene Mengen und sei $a \in O \cap O'$. Da O und O' offen in M sind, gibt es ein $\varepsilon > 0$ und ein $\delta > 0$ mit $U_\varepsilon(a) \subseteq O$ und $U_\delta(a) \subseteq O'$. Für $\gamma := \min\{\varepsilon, \delta\}$ gilt dann

$$U_\gamma(a) = U_\varepsilon(a) \cap U_\delta(a) \subseteq O \cap O'.$$

Mithin ist a ein innerer Punkt von $O \cap O'$, und $O \cap O'$ ist offen in M . Die Aussage ist also für den Durchschnitt zweier offener Mengen gezeigt und folgt für endliche Durchschnitte offener Mengen mittels Induktion.

- b. Sei O_i offen in M für $i \in I$, und sei $a \in \bigcup_{i \in I} O_i$. Dann gibt es ein $j \in I$ mit $a \in O_j$. Da O_j offen in M ist, gibt es ein $\varepsilon > 0$ mit

$$U_\varepsilon(a) \subseteq O_j \subseteq \bigcup_{i \in I} O_i.$$

Mithin ist a ein innerer Punkt von $\bigcup_{i \in I} O_i$, und $\bigcup_{i \in I} O_i$ ist offen in M .

- c. Seien A_1, \dots, A_n abgeschlossen in M , so ist $M \setminus A_i$ offen in M und deshalb ist nach a. auch

$$\bigcap_{i=1}^n (M \setminus A_i) = M \setminus \bigcup_{i=1}^n A_i$$

offen in M . Dann ist aber $A_1 \cup \dots \cup A_n$ abgeschlossen in M .

- d. Sei A_i abgeschlossen in M für $i \in I$, so ist $M \setminus A_i$ offen in M . Aus Teil b. folgt dann, daß

$$\bigcup_{i \in I} (M \setminus A_i) = M \setminus \bigcap_{i \in I} A_i$$

offen in M ist. Also ist $\bigcap_{i \in I} A_i$ abgeschlossen in M .

□

Beispiel 39.13

Die Intervalle $U_n = (-\frac{1}{n}, \frac{1}{n})$ sind offen in \mathbb{R} für $n \geq 1$. Aber ihr Durchschnitt

$$U := \bigcap_{n \geq 1} U_n = \{0\}$$

ist nicht offen in \mathbb{R} , da 0 kein innerer Punkt der Menge U ist.

Korollar 39.14 (Das Innere und der Abschluß)

Sei M ein metrischer Raum und $U \subseteq M$.

- Der Durchschnitt \bar{U} aller abgeschlossenen Teilmengen von M , die U enthalten, ist die kleinste abgeschlossene Teilmenge von M , die U enthält, und heißt der Abschluß von U .
- Die Vereinigung $\overset{\circ}{U}$ aller offenen Teilmengen von M , die in U enthalten sind, ist die größte offene Teilmenge von M , die in U enthalten ist, und heißt das Innere von U .

Beweis: Nach Proposition 39.12 ist \bar{U} als Durchschnitt abgeschlossener Mengen wieder abgeschlossen, und nach Definition ist \bar{U} in jeder abgeschlossenen Teilmenge von M enthalten, die U enthält, so daß \bar{U} die kleinste solche Menge ist. Die Aussage für $\overset{\circ}{U}$ sieht man analog. □

Definition 39.15 (Häufungspunkte und Randpunkte)

Sei M ein metrischer Raum, $U \subseteq M$ und $a \in M$.

- a heißt ein *Häufungspunkt* von U , wenn jede ε -Umgebung von a einen von a verschiedenen Punkt aus U enthält, d.h.

$$\forall \varepsilon > 0 \exists x \in U \setminus \{a\} : 0 < d(x, a) < \varepsilon,$$

d.h.

$$\forall \varepsilon > 0 : U_\varepsilon(a) \cap (U \setminus \{a\}) \neq \emptyset.$$

$HP(U)$ bezeichnet die Menge aller Häufungspunkte von U .

- a heißt *Randpunkt* von U , wenn jede ε -Umgebung von a mindestens einen Punkt aus U und einen aus $M \setminus U$ enthält, d.h.

$$\forall \varepsilon > 0 : U_\varepsilon(a) \cap U \neq \emptyset \text{ und } U_\varepsilon(a) \cap (M \setminus U) \neq \emptyset.$$

Die Menge ∂U aller Randpunkte von U heißt der *Rand* von U .

Beispiel 39.16 (Der Rand einer ε -Umgebung)

Ist V ein *normierter* Raum, $\mathbf{a} \in V$ und $\varepsilon > 0$, so ist

$$\partial U_\varepsilon(\mathbf{a}) = \{\mathbf{x} \in V \mid d(\mathbf{x}, \mathbf{a}) = \varepsilon\}$$

der Rand der ε -Umgebung von \mathbf{a} . In Abbildung 1 ist der Rand der ε -Umgebung also genau der schwarze Rand der grauen Fläche.

Um dies zu sehen, setzen wir $A = \{\mathbf{x} \in V \mid d(\mathbf{x}, \mathbf{a}) = \varepsilon\}$, wählen ein $\mathbf{b} \in A$ und ein $\delta > 0$, wobei wir ohne Einschränkung $\delta < \varepsilon$ annehmen können. Da \mathbf{b} nicht in $U_\varepsilon(\mathbf{a})$ liegt, gilt zunächst

$$\mathbf{b} \in U_\delta(\mathbf{b}) \cap (V \setminus U_\varepsilon(\mathbf{a})),$$

so daß dieser Durchschnitt nicht-leer ist. Setzen wir nun noch

$$\mathbf{x} := \mathbf{b} + \frac{\delta}{2 \cdot \varepsilon} \cdot (\mathbf{a} - \mathbf{b}),$$

so gilt

$$d(\mathbf{x}, \mathbf{b}) = \left\| \mathbf{b} + \frac{\delta}{2 \cdot \varepsilon} \cdot (\mathbf{a} - \mathbf{b}) - \mathbf{b} \right\| = \frac{\delta}{2 \cdot \varepsilon} \cdot \|\mathbf{a} - \mathbf{b}\| = \frac{\delta}{2 \cdot \varepsilon} \cdot \varepsilon < \delta$$

und

$$d(\mathbf{x}, \mathbf{a}) = \left\| \mathbf{b} + \frac{\delta}{2 \cdot \varepsilon} \cdot (\mathbf{a} - \mathbf{b}) - \mathbf{a} \right\| = \left| 1 - \frac{\delta}{2 \cdot \varepsilon} \right| \cdot \|\mathbf{b} - \mathbf{a}\| = \left| 1 - \frac{\delta}{2 \cdot \varepsilon} \right| \cdot \varepsilon < \varepsilon.$$

Damit liegt \mathbf{x} also im Durchschnitt von $U_\delta(\mathbf{b})$ und $U_\varepsilon(\mathbf{a})$, und dieser Durchschnitt ist nicht leer. Das zeigt, daß \mathbf{b} im Rand von $U_\varepsilon(\mathbf{a})$ liegt, d.h. $A \subseteq \partial U_\varepsilon(\mathbf{a})$.

Umgekehrt liegt kein Punkt $\mathbf{b} \in V$ mit $d(\mathbf{b}, \mathbf{a}) > \varepsilon$ im Rand von $U_\varepsilon(\mathbf{a})$, da die δ -Umgebung von \mathbf{b} für $\delta = \frac{d(\mathbf{b}, \mathbf{a}) - \varepsilon}{2}$ keinen Punkt mit $U_\varepsilon(\mathbf{a})$ gemeinsam hat. Für einen Punkt $\mathbf{x} \in U_\delta(\mathbf{b})$ gilt nämlich

$$d(\mathbf{a}, \mathbf{b}) \leq d(\mathbf{a}, \mathbf{x}) + d(\mathbf{x}, \mathbf{b}) < d(\mathbf{a}, \mathbf{x}) + \delta = d(\mathbf{a}, \mathbf{x}) + \frac{d(\mathbf{b}, \mathbf{a}) - \varepsilon}{2}$$

und somit

$$\varepsilon < \frac{d(\mathbf{b}, \mathbf{a})}{2} + \frac{\varepsilon}{2} < d(\mathbf{a}, \mathbf{x}).$$

Analog sieht man auch, daß kein Punkt $\mathbf{b} \in V$ mit $d(\mathbf{b}, \mathbf{a}) < \varepsilon$ im Rand von $U_\varepsilon(\mathbf{a})$ liegen kann. Dies zeigt, schließlich $\partial U_\varepsilon(\mathbf{a}) \subseteq A$.

Im Beweis haben wir verwendet, daß wir in V Punkte addieren können, d.h. wir haben die Vektorraumstruktur von V ausgenutzt, um zu zeigen, daß $U_\varepsilon(\mathbf{a}) \cap U_\delta(\mathbf{a})$ nicht leer ist. In einem beliebigen *metrischen* Raum muß dies *nicht* mehr gelten – dazu betrachte man die 1-Umgebung eines Punktes bezüglich der Eisenbahnmetrik.

Proposition 39.17 (Das Innere und der Abschluß von U)

Sei M ein *metrischer Raum* und $U \subseteq M$.

- $\overset{\circ}{U}$ ist die Menge aller inneren Punkte von U .
- $\bar{U} = U \cup \partial U$.
- $\bar{U} = U \cup \text{HP}(U)$.

Beweis:

a. Sei O die Menge aller inneren Punkte von U , dann ist $O \subseteq U$. Ist nun $a \in O$ ein innerer Punkt von U , so gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(a) \subseteq U$. Da $U_\varepsilon(a)$ offen ist, ist jeder Punkt $b \in U_\varepsilon(a)$ somit ein innerer Punkt von $U_\varepsilon(a)$ und damit auch ein innerer Punkt von U . Also ist $U_\varepsilon(a) \subseteq O$ und deshalb ist O offen. Dann ist O aber eine offene Teilmenge von U , und somit ist O in der größten offenen Teilmenge $\overset{\circ}{U}$ von U enthalten.

Andererseits ist $\overset{\circ}{U}$ offen, so daß jeder Punkt in $\overset{\circ}{U}$ ein innerer Punkt von $\overset{\circ}{U}$ und somit auch ein innerer Punkt der größeren Menge U ist. Deshalb ist $\overset{\circ}{U}$ in O enthalten.

b. Wir setzen $A = U \cup \partial U$ und wollen zunächst zeigen, daß A abgeschlossen ist, d.h., daß $M \setminus A$ offen ist. Sei dazu $a \in M \setminus A$. Dann ist $a \notin \partial U$ kein Randpunkt von U . Es gibt also ein $\varepsilon > 0$, so daß

$$U_\varepsilon(a) \cap U = \emptyset \quad \text{oder} \quad U_\varepsilon(a) \cap (M \setminus U) = \emptyset.$$

Da nach Voraussetzung $a \in U_\varepsilon(a) \cap (M \setminus U)$ gilt, muß also notwendigerweise

$$U_\varepsilon(a) \cap U = \emptyset$$

und damit

$$U_\varepsilon(a) \subseteq M \setminus U$$

gelten. Für $\delta = \frac{\varepsilon}{2}$ und $b \in U_\delta(a)$ gilt

$$U_\delta(b) \subseteq U_{2\delta}(a) = U_\varepsilon(a) \subseteq M \setminus U$$

und somit

$$U_\delta(b) \cap U = \emptyset.$$

Dann ist aber b weder in U noch im Rand ∂U von U enthalten, d.h. $b \notin U \cup \partial U = A$. Mithin liegt b in $M \setminus A$, und somit gilt

$$U_\delta(a) \subseteq M \setminus A.$$

Dies zeigt, daß a ein innerer Punkt von $M \setminus A$ ist. Mithin ist $M \setminus A$ offen in M und A abgeschlossen.

Damit ist aber A eine abgeschlossene Teilmenge von M , die U enthält, und enthält somit auch die kleinste solche abgeschlossene Teilmenge \bar{U} von M , d.h. $\bar{U} \subseteq A$.

Sei nun umgekehrt B irgendeine abgeschlossene Teilmenge von M , die U enthält. Wir wollen zeigen, daß sie dann auch den Rand ∂U enthält und daß somit $A = U \cup \partial U \subseteq B$ gilt. Sei dazu $a \in M \setminus B$ gegeben. Da $M \setminus B$ offen ist, gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(a) \subseteq M \setminus B$. Dann ist aber

$$U_\varepsilon(a) \cap U \subseteq U_\varepsilon(a) \cap B = \emptyset,$$

so daß a kein Randpunkt von U sein kann. Mithin gilt $\partial U \cap (M \setminus B) = \emptyset$, oder alternativ $\partial U \subseteq B$. Damit enthält also jede abgeschlossene Obermenge von U

auch $A = U \cup \partial U$, und somit enthält der Durchschnitt \bar{U} aller solcher Mengen auch die Menge A , d.h. $A \subseteq \bar{U}$.

- c. Für einen Punkt $a \in M \setminus U$ gilt:

$$a \in \text{HP}(U) \iff \forall \varepsilon > 0 : U_\varepsilon(a) \cap U \neq \emptyset \iff a \in \partial U.$$

Mithin gilt $\text{HP}(U) \setminus U = \partial U \setminus U$, und somit folgt aus b.

$$\bar{U} = U \cup (\partial U \setminus U) = U \cup (\text{HP}(U) \setminus U) = U \cup \text{HP}(U).$$

□

Beispiel 39.18 (Der Rand und der Abschluß)

- a. Die Endpunkte a und b des Intervalls (a, b) für $a < b$ sind genau die Randpunkte der Menge (a, b) , und mithin ist das abgeschlossene Intervall $[a, b]$ der Abschluß des Intervalls (a, b) in \mathbb{R} .
- b. Ist V ein *normierter Raum*, $a \in V$ und $\varepsilon > 0$. Dann ist

$$\overline{U_\varepsilon(a)} = \{x \in V \mid d(x, a) \leq \varepsilon\}$$

der Abschluß von $U_\varepsilon(a)$ in V , wegen Beispiel 39.16 und Proposition 39.17. In Abbildung 1 ist der Abschluß der ε -Umgebung also genau die schwarzumrandete Fläche inklusive des Randes.

Durch Komplementbildung sieht man zudem, daß die Menge

$$V \setminus U_\varepsilon(a) = \{x \in V \mid d(x, a) \geq \varepsilon\}$$

abgeschlossen ist in V und daß die folgende Menge offen ist

$$V \setminus \overline{U_\varepsilon(a)} = \{x \in V \mid d(x, a) > \varepsilon\}.$$

- c. Der Rand von \mathbb{Q} in \mathbb{R} ist ganz \mathbb{R} , da in jeder ε -Umgebung jeder reellen Zahl sowohl rationale, als auch irrationale Zahlen liegen. Damit ist \mathbb{R} auch der Abschluß von \mathbb{Q} in \mathbb{R} . Das Innere von \mathbb{Q} in \mathbb{R} ist hingegen leer, da \mathbb{Q} keine ε -Umgebung einer rationalen Zahl ganz enthalten kann.
- d. Ist $U = (0, 1] \cup \{2\} \subseteq \mathbb{R}$, so ist 2 kein Häufungspunkt von $(0, 1]$, da die $\frac{1}{2}$ -Umgebung von 2 keinen weiteren Punkt aus U enthält. Hingegen ist $\{2\}$ ein Randpunkt von U , da jede ε -Umgebung von 2 Punkte aus U und aus dem Komplement enthält. Dies zeigt, daß im allgemeinen nicht $\text{HP}(U) = \partial U$ gilt, sondern lediglich

$$\text{HP}(U) \setminus U = \partial U \setminus U.$$

Der Abschluß der Menge ist $\bar{U} = [0, 1] \cup \{2\}$.

C) Kompakte Mengen in normierten Räumen

Definition 39.19 (Beschränktheit und Kompaktheit)

Sei M ein metrischer Raum und $U \subseteq M$.

- U heißt *beschränkt*, wenn es ein $r > 0$ und ein $a \in M$ gibt mit $d(x, a) \leq r$ für alle $x \in U$.
- Eine Familie $(U_i)_{i \in I}$ offener Mengen in M heißt eine *offene Überdeckung* von U , falls $U \subseteq \bigcup_{i \in I} U_i$.
- U heißt *kompakt*, falls jede offene Überdeckung $(U_i)_{i \in I}$ von U eine endliche Teilüberdeckung U_{i_1}, \dots, U_{i_n} , $i_1, \dots, i_n \in I$, besitzt.

Beispiel 39.20

- Ein Intervall in \mathbb{R} mit endlichen Intervallgrenzen ist beschränkt.
- In jedem metrischen Raum sind ε -Umgebungen beschränkt.
- Endliche Teilmengen eines metrischen Raumes sind offenbar kompakt.
- Das offene Intervall $(0, 1)$ ist nicht kompakt, da die offene Überdeckung $(0, 1) = \bigcup_{n \geq 2} (\frac{1}{n}, 1)$ keine endliche Teilüberdeckung besitzt.

Satz 39.21 (Kompakte Mengen)

In einem metrischen Raum ist jede kompakte Menge abgeschlossen und beschränkt.

Beweis: Sei M ein metrischer Raum und $A \subseteq M$ eine kompakte Menge. Wir wollen zunächst zeigen, daß A beschränkt ist. Dazu betrachten wir die Überdeckung

$$A \subseteq \bigcup_{x \in A} U_1(x)$$

von A . Da A kompakt ist, gibt es eine endliche Teilüberdeckung, d.h. es gibt $x_1, \dots, x_n \in A$ mit

$$A \subseteq \bigcup_{i=1}^n U_1(x_i).$$

Wir wählen $a \in A$ beliebig, aber fest und setzen

$$r := \max\{d(x_1, a), \dots, d(x_n, a)\} + 1 > 0.$$

Für ein beliebiges $x \in A$ gibt es nun ein $i \in \{1, \dots, n\}$, so daß $x \in U_1(x_i)$, und somit gilt

$$d(x, a) \leq d(x, x_i) + d(x_i, a) \leq 1 + d(x_i, a) \leq r.$$

Also ist A beschränkt.

Um zu zeigen, daß A abgeschlossen ist, reicht es, zu zeigen, daß jedes $a \in M \setminus A$ eine ε -Umgebung besitzt, die ganz in $M \setminus A$ liegt. Sei also $a \in M \setminus A$ gegeben. Nach Proposition 39.8 gibt es zu jedem $x \in A$ ein $\varepsilon_x > 0$ mit

$$U_{\varepsilon_x}(a) \cap U_{\varepsilon_x}(x) = \emptyset. \quad (126)$$

Außerdem gilt

$$A \subseteq \bigcup_{x \in A} U_{\varepsilon_x}(x).$$

Da A kompakt ist, gibt es eine endliche Teilüberdeckung, d.h. es gibt $x_1, \dots, x_n \in A$ mit

$$A \subseteq \bigcup_{i=1}^n U_{\varepsilon_{x_i}}(x_i).$$

Wir setzen nun

$$\varepsilon := \min\{\varepsilon_{x_1}, \dots, \varepsilon_{x_n}\} > 0$$

und erhalten somit aus (126)

$$U_\varepsilon(a) \cap U_{\varepsilon_{x_i}}(x_i) = \emptyset$$

für alle $i = 1, \dots, n$ und somit

$$U_\varepsilon(a) \cap A \subseteq U_\varepsilon(a) \cap \bigcup_{i=1}^n U_{\varepsilon_{x_i}}(x_i) = \bigcup_{i=1}^n U_\varepsilon(a) \cap U_{\varepsilon_{x_i}}(x_i) = \emptyset.$$

Mithin liegt die ε -Umgebung $U_\varepsilon(a)$ ganz in $M \setminus A$. Dies zeigt, daß A auch abgeschlossen ist. \square

D) Folgen in metrischen und normierten Räumen

Um Kriterien formulieren und beweisen zu können, die es uns erlauben, festzustellen, ob eine Menge kompakt ist, benötigen wir den Begriff der konvergenten Folge.

Definition 39.22 (Folgen in metrischen Räumen)

Sei M ein metrischer Raum.

- Eine *Folge* in M ist eine Abbildung $\alpha : \mathbb{N} \rightarrow M$. Ist $\mathbf{a}_n := \alpha(n)$ für $n \in \mathbb{N}$, so schreiben wir meist $(\mathbf{a}_n)_{n \in \mathbb{N}}$ statt α .
- Wir nennen $\mathbf{a} \in M$ einen *Grenzwert* der Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in M , wenn

$$\forall 0 < \varepsilon \in \mathbb{R} \quad \exists n_\varepsilon \in \mathbb{N} : \forall n \geq n_\varepsilon : d(\mathbf{a}_n, \mathbf{a}) < \varepsilon.$$

In diesem Fall sagen wir auch, daß $(\mathbf{a}_n)_{n \in \mathbb{N}}$ *gegen \mathbf{a} konvergiert* und schreiben

$$\lim_{n \rightarrow \infty} \mathbf{a}_n = \mathbf{a}$$

oder

$$\mathbf{a}_n \rightarrow \mathbf{a}.$$

- Wir nennen eine Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in M genau dann *konvergent*, wenn $(\mathbf{a}_n)_{n \in \mathbb{N}}$ einen Grenzwert $\mathbf{a} \in M$ besitzt. Ist $(\mathbf{a}_n)_{n \in \mathbb{N}}$ nicht konvergent, so nennen wir $(\mathbf{a}_n)_{n \in \mathbb{N}}$ *divergent*. Ist M ein normierter Raum und ist der Grenzwert $\mathbf{0}$, so nennen wir $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine *Nullfolge*.
- Eine Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in M heißt *Cauchy-Folge*, falls

$$\forall 0 < \varepsilon \in \mathbb{R} \quad \exists n_\varepsilon \in \mathbb{N} : \forall m > n \geq n_\varepsilon : d(\mathbf{a}_m, \mathbf{a}_n) < \varepsilon.$$

e. Ist $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in M und ist zudem

$$\mathbf{n}_0 < \mathbf{n}_1 < \mathbf{n}_2 < \mathbf{n}_3 < \dots$$

eine aufsteigende Folge natürlicher Zahlen, so nennen wir die Folge

$$(\mathbf{a}_{\mathbf{n}_k})_{k \in \mathbb{N}} = (\mathbf{a}_{\mathbf{n}_0}, \mathbf{a}_{\mathbf{n}_1}, \mathbf{a}_{\mathbf{n}_2}, \mathbf{a}_{\mathbf{n}_3}, \dots)$$

eine *Teilfolge* von $(\mathbf{a}_n)_{n \in \mathbb{N}}$.

f. Eine Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in M heißt *beschränkt*, wenn $\{\mathbf{a}_n \mid n \in \mathbb{N}\}$ beschränkt ist, d.h. wenn es ein $r > 0$ und ein $\mathbf{a} \in M$ gibt mit $d(\mathbf{a}_n, \mathbf{a}) \leq r$ für alle $n \in \mathbb{N}$.

Proposition 39.23 (Einfache Eigenschaften von Folgen)

Sei M ein metrischer Raum, $\lambda \in \mathbb{R}$ und $(\mathbf{a}_n)_{n \in \mathbb{N}}$ und $(\mathbf{b}_n)_{n \in \mathbb{N}}$ Folgen in M .

- Genau dann gilt $\mathbf{a}_n \rightarrow \mathbf{a}$, wenn $d(\mathbf{a}_n, \mathbf{a}) \rightarrow 0$.
- Der Grenzwert einer konvergenten Folge in M ist eindeutig bestimmt.
- Jede konvergente Folge in M ist eine Cauchy-Folge.
- Cauchy-Folgen und konvergente Folgen in M sind beschränkt.
- Falls $\mathbf{a}_n \rightarrow \mathbf{a}$ gilt, so konvergiert jede Teilfolge von $(\mathbf{a}_n)_{n \in \mathbb{N}}$ gegen \mathbf{a} .
- Ist M ein normierter Raum und gelten $\mathbf{a}_n \rightarrow \mathbf{a}$ und $\mathbf{b}_n \rightarrow \mathbf{b}$, dann gelten

$$\mathbf{a}_n \pm \mathbf{b}_n \rightarrow \mathbf{a} \pm \mathbf{b} \quad \text{und} \quad \lambda \cdot \mathbf{a}_n \rightarrow \lambda \cdot \mathbf{a}.$$

- Wenn $\mathbf{a}_n \rightarrow \mathbf{a}$ und $\mathbf{b}_n \rightarrow \mathbf{b}$ gilt, dann gilt auch $d(\mathbf{a}_n, \mathbf{b}_n) \rightarrow d(\mathbf{a}, \mathbf{b})$.

Beweis: Ersetzt man in den entsprechenden Beweisen in Abschnitt 11 die Betragsdifferenz durch den Abstand d und den Körper \mathbb{K} durch M , so übertragen sich die Beweise von b.-f. wörtlich. Wir führen sie hier der Vollständigkeit halber noch einmal vor.

- Dies folgt unmittelbar aus der Definition der Konvergenz in M und in \mathbb{R} .
- Nehmen wir an, eine Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in M besitze zwei verschiedene Grenzwerte $\mathbf{a}, \mathbf{b} \in M$. Dann ist die reelle Zahl

$$\varepsilon := \frac{d(\mathbf{a}, \mathbf{b})}{2} > 0$$

positiv. Mithin gibt es zwei natürliche Zahlen $\mathbf{n}_\varepsilon, \mathbf{n}'_\varepsilon \in \mathbb{N}$, so daß

$$d(\mathbf{a}_n, \mathbf{a}) < \varepsilon$$

für $n \geq \mathbf{n}_\varepsilon$ und

$$d(\mathbf{a}_n, \mathbf{b}) < \varepsilon$$

für $n \geq \mathbf{n}'_\varepsilon$. Setzen wir nun $\mathbf{N} := \max\{\mathbf{n}_\varepsilon, \mathbf{n}'_\varepsilon\}$, so gilt

$$d(\mathbf{a}, \mathbf{b}) \leq d(\mathbf{a}, \mathbf{a}_\mathbf{N}) + d(\mathbf{a}_\mathbf{N}, \mathbf{b}) < \varepsilon + \varepsilon = d(\mathbf{a}, \mathbf{b}),$$

was ein offensichtlicher Widerspruch ist.

- c. Wir setzen voraus, daß $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine konvergente Folge ist mit Grenzwert \mathbf{a} . Sei nun $\varepsilon > 0$ gegeben, dann gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$d(\mathbf{a}_n, \mathbf{a}) < \frac{\varepsilon}{2}$$

für alle $n \geq n_\varepsilon$. Für zwei natürliche Zahlen $m > n \geq n_\varepsilon$ folgt dann

$$d(\mathbf{a}_m, \mathbf{a}_n) \leq d(\mathbf{a}_m, \mathbf{a}) + d(\mathbf{a}, \mathbf{a}_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Also ist $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge.

- d. Wegen c. reicht es, die Aussage für Cauchy-Folgen $(\mathbf{a}_n)_{n \in \mathbb{N}}$ zu zeigen. Zu $\varepsilon = 1$ gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$d(\mathbf{a}_m, \mathbf{a}_n) < \varepsilon = 1$$

für alle $m > n \geq n_\varepsilon$. Wähle $\mathbf{a} \in M$ fest und setze

$$r := \max\{1 + d(\mathbf{a}_{n_\varepsilon}, \mathbf{a}), d(\mathbf{a}_0, \mathbf{a}), d(\mathbf{a}_1, \mathbf{a}), \dots, d(\mathbf{a}_{n_\varepsilon-1}, \mathbf{a})\}.$$

Damit erhalten wir dann

$$d(\mathbf{a}_n, \mathbf{a}) \leq \begin{cases} r, & \text{falls } n < n_\varepsilon, \\ d(\mathbf{a}_n, \mathbf{a}_{n_\varepsilon}) + d(\mathbf{a}_{n_\varepsilon}, \mathbf{a}) \leq 1 + d(\mathbf{a}_{n_\varepsilon}, \mathbf{a}) \leq r, & \text{falls } n \geq n_\varepsilon. \end{cases}$$

Mithin ist $(\mathbf{a}_n)_{n \in \mathbb{N}}$ beschränkt.

- e. Die Aussage für Teilfolgen ist klar.
f. Sei $\varepsilon > 0$ gegeben. Dann gibt es natürliche Zahlen $n'_\varepsilon, n''_\varepsilon \in \mathbb{N}$, so daß

$$\|\mathbf{a}_n - \mathbf{a}\| = d(\mathbf{a}_n, \mathbf{a}) < \frac{\varepsilon}{2}$$

für $n \geq n'_\varepsilon$ und

$$\|\mathbf{b}_n - \mathbf{b}\| = d(\mathbf{b}_n, \mathbf{b}) < \frac{\varepsilon}{2}$$

für $n \geq n''_\varepsilon$. Mit $n_\varepsilon := \max\{n'_\varepsilon, n''_\varepsilon\}$ gilt dann

$$d(\mathbf{a}_n + \mathbf{b}_n, \mathbf{a} + \mathbf{b}) = \|\mathbf{a}_n - \mathbf{a} + \mathbf{b}_n - \mathbf{b}\| \leq \|\mathbf{a}_n - \mathbf{a}\| + \|\mathbf{b}_n - \mathbf{b}\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

für $n \geq n_\varepsilon$. Mithin konvergiert $(\mathbf{a}_n + \mathbf{b}_n)_{n \in \mathbb{N}}$ gegen $\mathbf{a} + \mathbf{b}$. Analog sieht man $\mathbf{a}_n - \mathbf{b}_n \rightarrow \mathbf{a} - \mathbf{b}$ und $\lambda \cdot \mathbf{a}_n \rightarrow \lambda \cdot \mathbf{a}$.

- g. Aus der Dreiecksungleichung leiten wir

$$d(\mathbf{x}_n, \mathbf{y}_n) \leq d(\mathbf{x}_n, \mathbf{x}) + d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{y}_n) = d(\mathbf{x}_n, \mathbf{x}) + d(\mathbf{y}_n, \mathbf{y}) + d(\mathbf{x}, \mathbf{y})$$

sowie

$$d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{x}_n) + d(\mathbf{x}_n, \mathbf{y}_n) + d(\mathbf{y}_n, \mathbf{y}) = d(\mathbf{x}_n, \mathbf{x}) + d(\mathbf{y}_n, \mathbf{y}) + d(\mathbf{x}_n, \mathbf{y}_n)$$

her. Daraus folgt dann aber unmittelbar

$$0 \leq |d(\mathbf{x}_n, \mathbf{y}_n) - d(\mathbf{x}, \mathbf{y})| \leq d(\mathbf{x}_n, \mathbf{x}) + d(\mathbf{y}_n, \mathbf{y}) \rightarrow 0 + 0 = 0.$$

Also konvergiert die reelle Zahlenfolge $(d(\mathbf{x}_n, \mathbf{y}_n))_{n \in \mathbb{N}}$ gegen $d(\mathbf{x}, \mathbf{y})$.

□

Beispiel 39.24

Die Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ mit

$$\mathbf{a}_n = \left(\frac{\cos(n)}{2^n}, \frac{\sin(n)}{2^n} \right)^t$$

konvergiert in \mathbb{R}^2 bezüglich der euklidischen Norm gegen $\mathbf{a} = (0, 0)^t$, da

$$d(\mathbf{a}_n, \mathbf{a}) = \sqrt{\frac{\cos^2(n) + \sin^2(n)}{2^{2n}}} = \frac{1}{2^n} \longrightarrow 0.$$

Lemma 39.25 (Folgenkriterium für Häufungspunkte)

Sei M ein metrischer Raum, $U \subseteq M$ und $\mathbf{a} \in M$. Genau dann ist \mathbf{a} ein Häufungspunkt von U , wenn es eine Folge in $U \setminus \{\mathbf{a}\}$ gibt, die gegen \mathbf{a} konvergiert.

Beweis: Der Beweis geht wörtlich wie der Beweis für das Folgenkriterium für Häufungspunkte in \mathbb{R} , Proposition 13.4, wenn wir die Betragsdifferenzen durch den Abstand d ersetzen. Wir führen ihn der Vollständigkeit halber noch mal aus.

Ist \mathbf{a} ein Häufungspunkt von U , so gibt es für jedes $n \geq 1$ ein

$$\mathbf{a}_n \in U_{\frac{1}{n}}(\mathbf{a}) \cap (U \setminus \{\mathbf{a}\}).$$

Dann gilt aber

$$0 \leq d(\mathbf{a}_n, \mathbf{a}) \leq \frac{1}{n} \longrightarrow 0,$$

und $d(\mathbf{a}_n, \mathbf{a})$ ist eine Nullfolge. Also ist $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in U , die gegen \mathbf{a} konvergiert.

Ist umgekehrt $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in $U \setminus \{\mathbf{a}\}$, die gegen \mathbf{a} konvergiert, und ist $\varepsilon > 0$, so gibt es ein $n_\varepsilon \in \mathbb{N}$ mit $d(\mathbf{a}_n, \mathbf{a}) < \varepsilon$ für alle $n \geq n_\varepsilon$. Insbesondere ist also $U_\varepsilon(\mathbf{a}) \cap (U \setminus \{\mathbf{a}\})$ nicht leer. Das zeigt, daß \mathbf{a} ein Häufungspunkt von U ist. \square

Beispiel 39.26

Die Häufungspunkte in $U = (0, 1] \cup \{2\}$ sind genau die Zahlen \mathbf{a} in $[0, 1]$, da nur diese Grenzwerte einer Folge in $U \setminus \{\mathbf{a}\}$ sind. Die Randpunkte von U sind genau die Zahlen $0, 1, 2$.

Satz 39.27 (Der Abschluß ist die Menge der Grenzwerte.)

Ist M ein metrischer Raum und $U \subseteq M$, so gilt

$$\bar{U} = \{\mathbf{a} \in M \mid \exists (\mathbf{a}_n)_{n \in \mathbb{N}} \subseteq U \text{ mit } \mathbf{a}_n \rightarrow \mathbf{a}\},$$

d.h. der Abschluß von U ist die Menge der Grenzwerte von Folgen in U .

Beweis: Sei zunächst \mathbf{a} der Grenzwert einer Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in U . Ist $\mathbf{a} \in U$, so ist $\mathbf{a} \in \bar{U}$. Ist $\mathbf{a} \notin U$, so ist \mathbf{a} nach Lemma 39.25 ein Häufungspunkt von U und somit gilt $\mathbf{a} \in \text{HP}(U) \subseteq \bar{U}$ nach Proposition 39.17.

Sei nun umgekehrt $\mathbf{a} \in \bar{U} = U \cup \text{HP}(U)$. Ist $\mathbf{a} \in U$, so ist \mathbf{a} der Grenzwert der konstanten Folge $(\mathbf{a})_{n \in \mathbb{N}}$. Ist $\mathbf{a} \notin U$, so muß \mathbf{a} ein Häufungspunkt von U sein und ist dann nach Lemma 39.25 Grenzwert einer Folge in $U \setminus \{\mathbf{a}\}$. \square

Beispiel 39.28

Der Abschluß der Menge $U = (0, 1] \cup \{2\}$ aus Beispiel 39.26 ist $[0, 1] \cup \{2\}$, da dies genau die reellen Zahlen sind, die als Grenzwerte von Folgen in U auftreten.

Korollar 39.29 (Folgenkriterium für Abgeschlossenheit)

Genau dann ist eine Teilmenge U eines metrischen Raumes M abgeschlossen in M , wenn jede in M konvergente Folge $(a_n)_{n \in \mathbb{N}} \subseteq U$ ihren Grenzwert schon in U hat.

Beweis: Nach Aufgabe 39.50 ist U genau dann abgeschlossen, wenn $U = \bar{U}$ gilt. Sei also U abgeschlossen und $(a_n)_{n \in \mathbb{N}}$ eine Folge in U , die in M konvergiert, so liegt der Grenzwert nach Satz 39.27 in $\bar{U} = U$. Ist umgekehrt jeder Grenzwert einer in M konvergenten Folge aus U schon in U , so gilt mit Satz 39.27 $U = \bar{U}$ und U ist abgeschlossen. \square

Satz 39.30 (Bolzano-Weierstraß)

Sei M ein metrischer Raum und $A \subseteq M$. Folgende Aussagen sind gleichwertig:

- A ist kompakt.
- Jede Folge in A besitzt eine konvergente Teilfolge mit Grenzwert in A .

Beweis: ¹

a. \implies b.: Angenommen, $(a_n)_{n \in \mathbb{N}}$ ist eine Folge in A , die keine konvergente Teilfolge besitzt. Dann hat $(a_n)_{n \in \mathbb{N}}$ insbesondere keine konstante Teilfolge und somit enthält die Menge

$$U = \{a_n \mid n \in \mathbb{N}\}$$

unendlich viele verschiedene Elemente. Da $(a_n)_{n \in \mathbb{N}}$ keine konvergente Teilfolge besitzt, hat U wegen Lemma 39.25 auch keinen Häufungspunkt. Also gibt es für jedes $x \in A$ ein $\varepsilon_x > 0$ mit $U_{\varepsilon_x}(x) \cap U \subseteq \{x\}$. Die $U_{\varepsilon_x}(x)$ für $x \in A$ überdecken A und A ist kompakt, so daß endlich viele davon ausreichen. D.h. es gibt $x_1, \dots, x_n \in A$ mit

$$A \subseteq \bigcup_{i=1}^n U_{\varepsilon_{x_i}}(x_i).$$

Es folgt also

$$U = U \cap A = \bigcup_{i=1}^n U \cap U_{\varepsilon_{x_i}}(x_i) \subseteq \{x_1, \dots, x_n\},$$

im Widerspruch dazu, daß die Menge U unendlich ist. Also muß $(a_n)_{n \in \mathbb{N}}$ eine konvergente Teilfolge besitzen.

Da die kompakte Menge A nach Satz 39.21 aber abgeschlossen ist, muß der Grenzwert der konvergenten Teilfolge wegen des Folgenkriteriums für Abgeschlossenheit 39.29 dann auch in A liegen.

b. \implies a.: Sei $(U_i)_{i \in I}$ eine offene Überdeckung von A . Wir müssen zeigen, daß es bereits eine endliche Teilüberdeckung gibt.

¹Den Beweis von "b. \implies a." habe ich von der Webseite von Prof. Franz Merkl, München.

Wir wollen dazu zunächst folgende Aussage zeigen:

$$\exists N \geq 1 : \forall x \in A \exists i \in I \text{ mit } U_{\frac{1}{N}}(x) \subseteq U_i. \quad (127)$$

Dazu definieren wir für $n \geq 1$ die Menge

$$X_n = \{x \in A \mid \exists i \in I : U_{\frac{1}{n}}(x) \subseteq U_i\}.$$

Aus der Definition folgt unmittelbar

$$X_1 \subseteq X_2 \subseteq X_3 \subseteq \dots,$$

da mit $U_{\frac{1}{n}}(x)$ auch $U_{\frac{1}{n+1}}(x)$ in U_i enthalten ist.

Ist $x \in A$ beliebig, so gibt es ein $i \in I$ mit $x \in U_i$, und da U_i offen ist, enthält U_i auch eine kleine Umgebung $U_{\frac{1}{n}}(x)$ für ein hinreichend großes n , d.h. für n hinreichend groß gilt $x \in X_n$. Wir erhalten also

$$A = \bigcup_{n \geq 1} X_n.$$

Um (127) zu zeigen, müssen wir nur zeigen, daß es in der Tat ein $N \geq 1$ gibt mit

$$A = X_N.$$

Wir nehmen das Gegenteil an, dann gibt es zu jedem $n \geq 1$ ein

$$a_n \in A \setminus X_n,$$

und nach Voraussetzung besitzt die Folge $(a_n)_{n \geq 1}$ eine konvergente Teilfolge mit Grenzwert $a \in A = \bigcup_{n \geq 1} X_n$. Es gibt also eine natürliche Zahl m , so daß $a \in X_m$ und somit

$$U_{\frac{1}{m}}(a) \subseteq U_i$$

für ein geeignetes $i \in I$. Aufgrund der Dreiecksungleichung gilt für $x \in U_{\frac{1}{2m}}(a)$ auch

$$U_{\frac{1}{2m}}(x) \subseteq U_{\frac{1}{m}}(a) \subseteq U_i.$$

Dies zeigt, $x \in X_{2m}$ und somit

$$U_{\frac{1}{2m}}(a) \subseteq X_{2m} \subseteq X_n$$

für alle $n \geq 2m$. Da $a_n \notin X_n$ gilt, muß notwendigerweise

$$d(a_n, a) \geq \frac{1}{2m}$$

für alle $n \geq 2m$ gelten, im Widerspruch dazu, daß eine Teilfolge von $(a_n)_{n \geq 1}$ gegen a konvergiert. Damit ist (127) gezeigt.

Wir nehmen nun an, daß $(U_i)_{i \in I}$ keine endliche Teilüberdeckung von A besitzt, und wollen dies zum Widerspruch führen. Wir konstruieren uns dazu rekursiv eine Folge $(i_n)_{n \in \mathbb{N}}$ in I sowie eine Folge $(x_n)_{n \in \mathbb{N}}$ in A mit

$$x_n \notin U_{i_0} \cup \dots \cup U_{i_n}.$$

i_0 wählen wir beliebig und dann wählen wir $x_0 \in A \setminus U_{i_0}$, wobei letzteres möglich ist, da sonst U_{i_0} die Menge A überdecken würde.

Seien nun $i_0, \dots, i_{n-1} \in I$ und $x_0, \dots, x_{n-1} \in A$ schon konstruiert. Wegen (127) können wir ein $i_n \in I$ wählen mit

$$U_{\frac{1}{N}}(x_{n-1}) \subseteq U_{i_n}$$

und da nach Annahme U_{i_0}, \dots, U_{i_n} die Menge A nicht überdecken, gibt es auch ein

$$x_n \in A \setminus (U_{i_0} \cup \dots \cup U_{i_n}).$$

Seien nun natürliche Zahlen $n > m \geq 0$ gegeben. Dann gilt nach Konstruktion

$$x_n \notin U_{i_0} \cup \dots \cup U_{i_n} \supseteq U_{i_{m+1}} \supseteq U_{\frac{1}{N}}(x_m)$$

und mithin

$$d(x_n, x_m) \geq \frac{1}{N}.$$

Dies zeigt, daß keine Teilfolge von $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge sein kann, und mithin kann nach Proposition 39.23 erst recht keine Teilfolge von $(x_n)_{n \in \mathbb{N}}$ konvergent sein, was im Widerspruch zur Voraussetzung steht. \square

Beispiel 39.31 (Abgeschlossene Intervalle sind kompakt.)

Ein abgeschlossenes Intervall $[a, b]$ in \mathbb{R} ist kompakt.

Um dies zu sehen, nutzen wir aus, daß das Intervall $[a, b]$ beschränkt ist, so daß nach dem Satz von Bolzano-Weierstraß 11.26 in \mathbb{R} jede Folge in $[a, b]$ eine konvergente Teilfolge mit Grenzwert c in \mathbb{R} besitzt, und daß nach dem Folgenkriterium für Abgeschlossenheit 39.29 dieser Grenzwert c dann schon in $[a, b]$ liegen muß, da $[a, b]$ abgeschlossen ist.

Korollar 39.32 (Abgeschlossene Teilmengen von Kompakta sind kompakt.)

Ist M kompakt und $A \subseteq M$ abgeschlossen in M , dann ist A kompakt.

Beweis: Ist $(a_n)_{n \in \mathbb{N}}$ eine Folge in A , so ist sie auch eine Folge in M . Nach dem Satz von Bolzano-Weierstraß 39.30 besitzt $(a_n)_{n \in \mathbb{N}}$ also eine konvergente Teilfolge mit Grenzwert in M . Da aber A abgeschlossen ist, ist dieser Grenzwert dann schon in A . Mithin ist die Menge A nach dem Satz von Bolzano-Weierstraß 39.30 kompakt. \square

E) Äquivalente Normen

Definition 39.33 (Äquivalente Normen)

Zwei Normen $|\cdot| : V \rightarrow \mathbb{R}_{\geq 0}$ und $\|\cdot\| : V \rightarrow \mathbb{R}_{\geq 0}$ auf einem \mathbb{R} -Vektorraum V heißen äquivalent, wenn es positive reelle Zahlen $s, r \in \mathbb{R}_{> 0}$ gibt mit

$$s \cdot |x| \leq \|x\| \leq r \cdot |x|$$

für alle $x \in V$.

Bemerkung 39.34 (Die Äquivalenz von Normen ist eine Äquivalenzrelation.)

Die Äquivalenz von Normen ist eine Äquivalenzrelation auf der Menge aller Normen eines Vektorraums V .

Beweis: Für die Reflexivität können wir $s = r = 1$ wählen. Für die Symmetrie leiten wir aus

$$s \cdot |x| \leq \|x\| \leq r \cdot |x|$$

für alle $x \in V$ her, daß auch

$$\frac{1}{r} \cdot \|x\| \leq |x| \leq \frac{1}{s} \cdot \|x\|$$

für alle $x \in V$ gilt. Und für die Transitivität beachten wir schließlich, daß aus

$$s \cdot |x| \leq \|x\| \leq r \cdot |x|$$

und

$$S \cdot \|x\| \leq ||x|| \leq R \cdot \|x\|$$

auch

$$S \cdot s \cdot |x| \leq S \cdot \|x\| \leq ||x|| \leq R \cdot \|x\| \leq R \cdot r \cdot |x|$$

für alle $x \in V$ gilt. □

Bemerkung 39.35 (Äquivalente Normen erzeugen die gleiche Topologie!)

Topologische Begriffe in einem normierten Raum wie die Offenheit oder Abgeschlossenheit einer Menge, der Rand einer Menge oder ihre Häufungspunkte, die Kompaktheit einer Menge oder die Konvergenz einer Folge hängen alle von der gewählten Norm ab, da diese den Abstandsbegriff festlegt. In aller Regel muß man davon ausgehen, daß unterschiedliche Normen zu unterschiedlichen Ergebnissen führen, d.h., daß eine Folge bezüglich einer Norm konvergiert, aber bezüglich einer anderen nicht, oder daß eine Menge bezüglich einer Norm offen ist, aber bezüglich einer anderen nicht. Sind Normen je *äquivalent*, so passiert das *nicht*.

Denn, gilt für zwei Normen $|\cdot|$ und $\|\cdot\|$ auf V

$$s \cdot |x| \leq \|x\| \leq r \cdot |x|$$

für alle $x \in V$ mit $s, r > 0$, so gilt (siehe auch Abbildung 2)

$$U_{s \cdot \varepsilon}^{|\cdot|}(\mathbf{a}) \subseteq U_{\varepsilon}^{\|\cdot\|}(\mathbf{a})$$

und

$$U_{\frac{\varepsilon}{r}}^{|\cdot|}(\mathbf{a}) \subseteq U_{\varepsilon}^{\|\cdot\|}(\mathbf{a}),$$

wobei $U_{\delta}^{|\cdot|}(\mathbf{a})$ die δ -Umgebung von \mathbf{a} bezüglich der Norm $|\cdot|$ bezeichnen soll und analog $U_{\delta}^{\|\cdot\|}(\mathbf{a})$ die δ -Umgebung von \mathbf{a} bezüglich der Norm $\|\cdot\|$. D.h. jede kleine Umgebung eines Punktes bezüglich der einen Norm enthält eine kleine Umgebung des Punktes bezüglich der anderen Norm. Damit erhalten wir für die beiden Äquivalenten Normen $|\cdot|$ und $\|\cdot\|$ auf V sowie $U, A \subseteq V$, $\mathbf{a} \in V$ und $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in V :

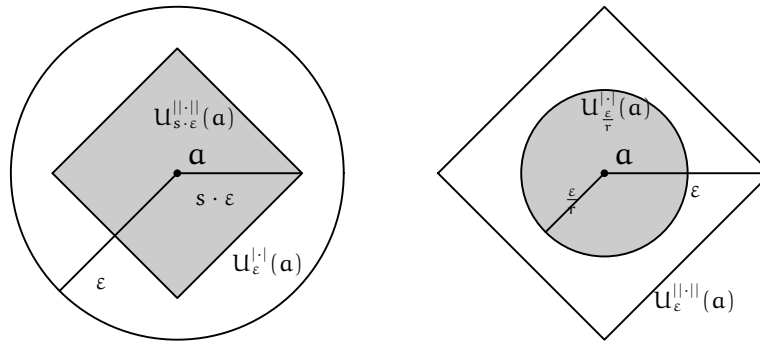


ABBILDUNG 2. Umgebungen bei äquivalenten Normen

Eigenschaften bezüglich $ \cdot $		Eigenschaften bezüglich $\ \cdot\ $
\mathbf{a} ist ein innerer Punkt von \mathbf{U} .	\Leftrightarrow	\mathbf{a} ist ein innerer Punkt von \mathbf{U} .
\mathbf{a} ist ein Randpunkt von \mathbf{U} .	\Leftrightarrow	\mathbf{a} ist ein Randpunkt von \mathbf{U} .
\mathbf{a} ist ein Häufungspunkt von \mathbf{U} .	\Leftrightarrow	\mathbf{a} ist ein Häufungspunkt von \mathbf{U} .
\mathbf{U} ist beschränkt.	\Leftrightarrow	\mathbf{U} ist beschränkt.
\mathbf{U} ist offen in \mathbf{V} .	\Leftrightarrow	\mathbf{U} ist offen in \mathbf{V} .
\mathbf{A} ist abgeschlossen in \mathbf{V} .	\Leftrightarrow	\mathbf{A} ist abgeschlossen in \mathbf{V} .
\mathbf{A} ist kompakt.	\Leftrightarrow	\mathbf{A} ist kompakt.
$(\mathbf{a}_n)_{n \in \mathbb{N}}$ konvergiert gegen \mathbf{a} .	\Leftrightarrow	$(\mathbf{a}_n)_{n \in \mathbb{N}}$ konvergiert gegen \mathbf{a} .
$(\mathbf{a}_n)_{n \in \mathbb{N}}$ ist eine Cauchy-Folge.	\Leftrightarrow	$(\mathbf{a}_n)_{n \in \mathbb{N}}$ ist eine Cauchy-Folge.

Bemerkung 39.36

Wir werden in Lemma 39.41 zeigen, daß jede Norm auf \mathbb{R}^n zur Maximumsnorm äquivalent ist, und daß somit alle Normen auf \mathbb{R}^n äquivalent sind. Wegen Bemerkung 39.35 ist es deshalb egal, welche Norm wir auf dem \mathbb{R}^n wählen, wenn wir topologische Begriffe oder Abstände verwenden wollen. Bis zum Beweis dieser Aussage werden wir uns aber auf die Verwendung der Maximumsnorm beschränken:

Bis zum Beweis von Lemma 39.41 werden wir \mathbb{R}^n mit der *Maximumsnorm*

$$\|\cdot\|_\infty : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geq 0} : \mathbf{x} = (x_1, \dots, x_n)^t \mapsto \|\mathbf{x}\|_\infty = \max\{|x_1|, \dots, |x_n|\}$$

als normierten Raum und damit auch als metrischen Raum betrachten.

F) Folgen und Kompaktheit im \mathbb{R}^n

Definition 39.37

Ist $(\mathbf{a}_k)_{k \in \mathbb{N}}$ eine Folge in \mathbb{R}^n mit $\mathbf{a}_k = (a_{k1}, \dots, a_{kn})^t$, so nennen wir die Folgen $(a_{ki})_{k \in \mathbb{N}}$ in \mathbb{R} die *Komponentenfolgen* von $(\mathbf{a}_k)_{k \in \mathbb{N}}$.

Proposition 39.38 (Komponentenweise Konvergenz)

Sei $(\mathbf{a}_k)_{k \in \mathbb{N}}$ eine Folge in \mathbb{R}^n und $\mathbf{c} \in \mathbb{R}^n$.

Genau dann gilt $\mathbf{a}_k \rightarrow \mathbf{c}$, wenn $a_{ki} \rightarrow c_i$ für $i = 1, \dots, n$.

Beweis: Aus $\mathbf{a}_k \rightarrow \mathbf{c}$ folgt

$$0 \leq |\mathbf{a}_{ki} - \mathbf{c}_i| \leq \max\{|\mathbf{a}_{k1} - \mathbf{c}_1|, \dots, |\mathbf{a}_{kn} - \mathbf{c}_n|\} = \|\mathbf{a}_k - \mathbf{c}\|_\infty = d(\mathbf{a}_k, \mathbf{c}) \longrightarrow 0$$

und somit $\mathbf{a}_{ki} \rightarrow \mathbf{c}_i$. Umgekehrt folgt aus $\mathbf{a}_{ki} \rightarrow \mathbf{c}_i$ für alle $i = 1, \dots, n$ auch $|\mathbf{a}_{ki} - \mathbf{c}_i| \rightarrow 0$ für alle $i = 1, \dots, n$ und somit

$$\|\mathbf{a}_k - \mathbf{c}\|_\infty = \max\{|\mathbf{a}_{k1} - \mathbf{c}_1|, \dots, |\mathbf{a}_{kn} - \mathbf{c}_n|\} \longrightarrow 0$$

und somit $\mathbf{a}_k \rightarrow \mathbf{c}$. □

Beispiel 39.39

Die Folge $(\mathbf{a}_n)_{n \geq 1}$ mit $\mathbf{a}_n = \left(\frac{2n^3+1}{n^3+n}, \left(1 + \frac{1}{n}\right)^n\right)$ konvergiert in \mathbb{R}^2 gegen $(2, e)$.

Satz 39.40 (Bolzano-Weierstraß)

Jede beschränkte Folge im \mathbb{R}^n besitzt eine konvergente Teilfolge.

Beweis: Es sei $(\mathbf{a}_k)_{k \in \mathbb{N}}$ eine beschränkte Folge in \mathbb{R}^n , dann gibt es ein $r > 0$, so daß

$$|\mathbf{a}_{ki}| \leq \|\mathbf{a}_k\|_\infty \leq r \tag{128}$$

für alle $k \in \mathbb{N}$ und $i = 1, \dots, n$. Mithin ist die Komponentenfolge $(\mathbf{a}_{k1})_{k \in \mathbb{N}}$ beschränkt und besitzt nach dem Satz von Bolzano-Weierstraß 11.26 in \mathbb{R} eine konvergente Teilfolge $(\mathbf{a}_{k_j 1})_{j \in \mathbb{N}}$. Die Komponentenfolge $(\mathbf{a}_{k_j 2})_{j \in \mathbb{N}}$ ist wegen (128) aber ebenfalls beschränkt und besitzt eine konvergente Teilfolge $(\mathbf{a}_{k_{j_l} 2})_{l \in \mathbb{N}}$. Die entsprechende Teilfolge in der ersten Komponente, $(\mathbf{a}_{k_{j_l} 1})_{l \in \mathbb{N}}$ ist aber ebenfalls konvergent als Teilfolge einer konvergenten Folge. Fahren wir so fort, erhalten wir schließlich eine Teilfolge von $(\mathbf{a}_k)_{k \in \mathbb{N}}$, so daß alle Komponentenfolgen konvergieren. Wegen des Satzes zur komponentenweisen Konvergenz 39.38 konvergiert dann aber auch die Teilfolge selbst. □

Lemma 39.41 (Jede Norm auf \mathbb{R}^n ist äquivalent zu $\|\cdot\|_\infty$.)

Sei $\|\cdot\|$ eine beliebige Norm auf \mathbb{R}^n , so gibt es reelle Zahlen $s, r > 0$ mit

$$s \cdot \|\mathbf{x}\|_\infty \leq \|\mathbf{x}\| \leq r \cdot \|\mathbf{x}\|_\infty$$

für alle $\mathbf{x} \in \mathbb{R}^n$.

Beweis: Mit $(\mathbf{e}_1, \dots, \mathbf{e}_n)$ bezeichnen wir die kanonische Basis von \mathbb{R}^n , und wir setzen

$$r := \|\mathbf{e}_1\| + \dots + \|\mathbf{e}_n\| > 0.$$

Für $\mathbf{x} = x_1 \cdot \mathbf{e}_1 + \dots + x_n \cdot \mathbf{e}_n \in \mathbb{R}^n$ gilt dann wegen der Dreiecksungleichung

$$\|\mathbf{x}\| = \left\| \sum_{i=1}^n x_i \cdot \mathbf{e}_i \right\| \leq \sum_{i=1}^n |x_i| \cdot \|\mathbf{e}_i\| \leq \sum_{i=1}^n \|\mathbf{x}\|_\infty \cdot \|\mathbf{e}_i\| = r \cdot \|\mathbf{x}\|_\infty. \tag{129}$$

Nehmen wir nun an, es gebe kein $s > 0$ mit $s \cdot \|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|$ für alle $\mathbf{x} \in \mathbb{R}^n$. Dann gibt es für jedes $k > 0$ ein $\mathbf{x}_k \in \mathbb{R}^n$ mit

$$\frac{1}{k} \cdot \|\mathbf{x}_k\|_\infty > \|\mathbf{x}_k\|. \tag{130}$$

Für $\mathbf{y}_k = \frac{1}{\|\mathbf{x}_k\|_\infty} \cdot \mathbf{x}_k$ gilt dann einerseits

$$0 \leq \|\mathbf{y}_k\| = \left\| \frac{\mathbf{x}_k}{\|\mathbf{x}_k\|_\infty} \right\| \stackrel{(130)}{<} \frac{1}{k} \longrightarrow 0 \quad (131)$$

und andererseits

$$\|\mathbf{y}_k\|_\infty = \frac{\|\mathbf{x}_k\|_\infty}{\|\mathbf{x}_k\|_\infty} = 1.$$

Bezüglich der Maximumsnorm ist die Folge $(\mathbf{y}_k)_{k \in \mathbb{N}}$ also beschränkt und besitzt nach dem Satz von Bolzano-Weierstraß 39.40 somit eine konvergente Teilfolge $(\mathbf{y}_{k_j})_{j \in \mathbb{N}}$ mit Grenzwert \mathbf{a} , d.h.

$$\|\mathbf{y}_{k_j} - \mathbf{a}\|_\infty \longrightarrow 0. \quad (132)$$

Für den Grenzwert gilt nun unter Berücksichtigung von (131) und (132)

$$0 \leq \|\mathbf{a}\| \leq \|\mathbf{a} - \mathbf{y}_{k_j}\| + \|\mathbf{y}_{k_j}\| \stackrel{(129)}{\leq} r \cdot \|\mathbf{a} - \mathbf{y}_{k_j}\|_\infty + \|\mathbf{y}_{k_j}\| \longrightarrow 0 + 0 = 0,$$

was $\|\mathbf{a}\| = 0$ bedingt. Dann ist aber $\mathbf{a} = \mathbf{0}$ der Nullvektor, da die Norm $\|\cdot\|$ definit ist. Daraus ergibt sich dann der Widerspruch

$$1 = \|\mathbf{y}_{k_j}\|_\infty \leq \|\mathbf{y}_{k_j} - \mathbf{a}\|_\infty + \|\mathbf{a}\|_\infty = \|\mathbf{y}_{k_j} - \mathbf{a}\|_\infty \longrightarrow 0.$$

□

Beispiel 39.42

Für $\mathbf{x} \in \mathbb{R}^n$ gilt offenbar $\|\mathbf{x}\|_\infty \leq \|\mathbf{x}\|_2 \leq \sqrt{n} \cdot \|\mathbf{x}\|_\infty$.

Korollar 39.43

Je zwei Normen auf \mathbb{R}^n sind äquivalent.

Beweis: Dies folgt aus Lemma 39.41, da die Äquivalenz von Normen nach Bemerkung 39.34 eine Äquivalenzrelation ist. □

Von jetzt an ist also egal, welche Norm wir auf \mathbb{R}^n verwenden!

Satz 39.44 (Heine-Borel)

$A \subseteq \mathbb{R}^n$ ist genau dann kompakt, wenn A abgeschlossen und beschränkt ist.

Beweis: Ist A kompakt in \mathbb{R}^n , so ist A nach Satz 39.21 abgeschlossen und beschränkt in \mathbb{R}^n .

Sei nun umgekehrt A abgeschlossen und beschränkt in \mathbb{R}^n und sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in A . Nach dem Satz von Bolzano-Weierstraß 39.40 im \mathbb{R}^n besitzt $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine konvergente Teilfolge mit Grenzwert $\mathbf{a} \in \mathbb{R}^n$. Da A abgeschlossen ist, liegt der Grenzwert \mathbf{a} in A , wegen des Folgenkriteriums für Abgeschlossenheit 39.29. Dann ist die Menge A aber kompakt wegen des Satzes von Bolzano-Weierstraß 39.30 in metrischen Räumen. □

Beispiel 39.45

a. Der Abschluß der ε -Umgebung $\{x \in \mathbb{R}^n \mid \|x\|_2 \leq \varepsilon\}$ ist kompakt in \mathbb{R}^n , da er abgeschlossen und beschränkt ist.

b. Der n -dimensionale Quader

$$[a_1, b_1] \times \dots \times [a_n, b_n] = \{(x_1, \dots, x_n)^t \mid a_i \leq x_i \leq b_i \forall i = 1, \dots, n\}$$

ist kompakt, da er abgeschlossen und beschränkt ist.

Um die Abgeschlossenheit zu sehen, beachtet man, daß jede konvergente Folge $(a_k)_{k \in \mathbb{N}}$ im Quader konvergente Komponentenfolgen in den Intervallen $[a_i, b_i]$ hat, deren Grenzwert dann in diesem Intervall liegen muß, so daß der Grenzwert von $(a_k)_{k \in \mathbb{N}}$ auch im Quader liegt.

c. Die abgeschlossenen Intervalle $[a, b]$ sind *genau* die kompakten Intervalle in \mathbb{R} , da alle anderen Intervalle nicht abgeschlossen oder nicht beschränkt sind. Die zentrale Rolle, die abgeschlossenen Intervallen bei vielen Aussagen der eindimensionalen Analysis zugekommen ist, kommt in höheren Dimensionen allgemeiner kompakten Mengen zu.

d. Die $n - 1$ -dimensionale Einheitssphäre

$$S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}$$

ist abgeschlossen und beschränkt, also kompakt.

Korollar 39.46 (Kompakta enthalten ihre Maxima und Minima.)

Ist A eine kompakte Menge in \mathbb{R} , so gilt $\inf(A) = \min(A)$ und $\sup(A) = \max(A)$.

Beweis: Da A kompakt ist, ist A beschränkt und $\sup(A)$ sowie $\inf(A)$ existieren. Angenommen, $a = \sup(A) \notin A$, dann gilt wegen Proposition 8.18

$$a \notin U_\varepsilon(a) \cap A \neq \emptyset.$$

Mithin ist a ein Häufungspunkt von A , und da A abgeschlossen ist, ist dann $a \in A$ im Widerspruch zur Annahme. Diese ist also falsch und es gilt $\sup(A) = a \in A$. Dann ist das Supremum aber ein Maximum. Analog sieht man $\inf(A) = \min(A)$. \square

G) Vollständige metrische Räume**Definition 39.47** (Banachraum und Hilbertraum)

- Ein metrischer Raum heißt *vollständig*, wenn jede Cauchy-Folge konvergiert.
- Ein normierter Raum heißt *Banachraum*, wenn er als metrischer Raum vollständig ist.
- Ein euklidischer Raum heißt *Hilbertraum*, wenn er mit der euklidischen Norm ein Banachraum ist.

Satz 39.48 (Cauchy-Kriterium für Folgenkonvergenz)

Jede Cauchy-Folge in \mathbb{R}^n ist konvergent.

Beweis: Es sei $(\mathbf{a}_k)_{k \in \mathbb{N}}$ eine Cauchy-Folge in \mathbb{R}^n und es sei $\varepsilon > 0$ gegeben. Dann gibt es ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $k > l \geq n_\varepsilon$

$$\varepsilon > d(\mathbf{a}_k, \mathbf{a}_l) = \|\mathbf{a}_k - \mathbf{a}_l\|_2 \geq |\mathbf{a}_{ki} - \mathbf{a}_{li}|$$

gilt, wobei $i \in \{1, \dots, n\}$ beliebig ist. Daraus folgt, daß die i -te Komponentenfolge von $(\mathbf{a}_k)_{k \in \mathbb{N}}$ eine Cauchy-Folge in \mathbb{R} ist und mithin nach dem Cauchy-Kriterium 11.30 auch konvergent. Dann ist $(\mathbf{a}_k)_{k \in \mathbb{N}}$ nach Proposition 39.38 aber konvergent. \square

Korollar 39.49 (\mathbb{R}^n ist ein Banachraum.)

\mathbb{R}^n ist bezüglich jeder Norm ein Banachraum, und bezüglich der euklidischen Norm ist \mathbb{R}^n ein Hilbertraum.

Aufgaben

Aufgabe 39.50

Sei M ein metrischer Raum und $U \subseteq M$.

- a. U ist genau dann offen in M , wenn $U = \overset{\circ}{U}$.
- b. U ist genau dann abgeschlossen in M , wenn $U = \bar{U}$.
- c. $\bar{\bar{U}} = \bar{U}$ und $\overset{\circ}{\overset{\circ}{U}} = \overset{\circ}{U}$.
- d. $\bar{U} = \overset{\circ}{U} \cup \partial U$.
- e. ∂U ist abgeschlossen in M .

Aufgabe 39.51

Bestimme für die Summennorm und die euklidische Norm jeweils Konstanten r und s wie in Lemma 39.41.

Aufgabe 39.52 (Relativ offen und abgeschlossen)

Betrachte $U \subseteq M$ als metrischen Raum mit der Einschränkung der Metrik von M .

- a. $X \subseteq U$ ist offen in $U \iff \exists O$ offen in M mit $X = U \cap O$.
- b. $X \subseteq U$ ist abgeschlossen in $U \iff \exists A$ abgeschlossen in M mit $X = U \cap A$.

Man nennt die Menge X dann auch *relativ* offen oder abgeschlossen in M .

Aufgabe 39.53 (Normen auf Funktionenräumen)

Sei $V = \mathcal{C}([0, 1], \mathbb{R})$ der Vektorraum der auf dem Intervall $[0, 1]$ stetigen Funktionen und $U = \mathcal{C}^1([0, 1], \mathbb{R})$ der Vektorraum der dort stetig-differenzierbaren Funktionen. Zeige die folgenden Aussagen:

- a. V ist nicht vollständig bezüglich der euklidischen L_2 -Norm aus Beispiel 39.5.
- b. V ist ein Banachraum bezüglich der Maximumsnorm aus Beispiel 39.5.
- c. Der Differentialoperator $D : U \rightarrow V : f \mapsto f'$ ist ein linearer Operator, der *nicht* stetig ist, wenn U und V normiert sind bzgl. der Maximumsnorm.

Aufgabe 39.54 (Abgeschlossene Teilräume vollständiger Räume sind vollständig.)
Ist (M, d) ein vollständiger metrischer Raum und $A \subseteq M$ eine abgeschlossene Teilmenge, so ist A mit der Einschränkung von d auf A ebenfalls vollständig.

Aufgabe 39.55 (Submultiplikativität der euklidischen Norm)

Identifizieren wir den \mathbb{R} -Vektorraum $\text{Mat}(\mathbf{m} \times \mathbf{n}, \mathbb{R})$ der $\mathbf{m} \times \mathbf{n}$ -Matrizen mit $\mathbb{R}^{\mathbf{m} \times \mathbf{n}}$, so definiert

$$\|(\mathbf{a}_{ij})\|_2 := \sqrt{\sum_{i=1}^{\mathbf{m}} \sum_{j=1}^{\mathbf{n}} \mathbf{a}_{ij}^2}$$

gerade die euklidische Norm auf $\text{Mat}(\mathbf{m} \times \mathbf{n}, \mathbb{R})$.

Zeige, für eine Matrix $A \in \text{Mat}(\mathbf{m} \times \mathbf{n}, \mathbb{R})$ und einen Vektor $\mathbf{x} \in \mathbb{R}^{\mathbf{n}}$ gilt stets

$$\|A\mathbf{x}\|_2 \leq \|A\|_2 \cdot \|\mathbf{x}\|_2.$$

Aufgabe 39.56

Zeige, $\|f\|_\infty := \max \{|f(x)| \mid x \in [0, 1]\}$ definiert eine Norm auf $V = \mathcal{C}([0, 1], \mathbb{R})$.

Aufgabe 39.57

Sei M ein metrischer Raum und $f : M \rightarrow \mathbb{R}$ eine stetige Abbildung. Finde geeignete Beispiele für M , f und Teilmengen $A, O \subseteq M$ bzw. $K \subseteq \mathbb{R}$ mit den folgenden Eigenschaften:

- $A \subseteq M$ ist abgeschlossen und $f(A) \subseteq \mathbb{R}$ ist *nicht* abgeschlossen.
- $O \subseteq M$ ist offen und $f(O) \subseteq \mathbb{R}$ ist *nicht* offen.
- $K \subseteq \mathbb{R}$ ist kompakt und $f^{-1}(K) \subseteq M$ ist *nicht* kompakt.

§ 40 Stetigkeit in metrischen und in normierten Räumen

In diesem Abschnitt seien (M, d) und (M', d') stets metrische Räume. Außerdem sei V stets ein normierter \mathbb{R} -Vektorraum mit Norm $\|\cdot\|_V$ und W ein normierter \mathbb{R} -Vektorraum mit Norm $\|\cdot\|_W$. Der Einfachheit halber bezeichnen wir die Metrik, die durch die Normen definiert wird, ebenfalls mit d . \mathbb{R}^n und \mathbb{R}^m betrachten wir als normierte Räume bezüglich einer beliebigen Norm.

A) Grenzwerte in metrischen und in normierten Räumen

Definition 40.1 (ε - δ -Kriterium für Grenzwerte von Abbildungen)

Sei $U \subseteq M$, $f: U \rightarrow M'$ eine Abbildung und $a \in M$ ein Häufungspunkt von U . Wir nennen $y \in M'$ den *Grenzwert* von f in a , falls

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : \forall x \in U \text{ mit } 0 < d(x, a) < \delta_\varepsilon \text{ gilt } d'(f(x), y) < \varepsilon,$$

d.h.

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : f((U_{\delta_\varepsilon}(a) \setminus \{a\}) \cap U) \subseteq U_\varepsilon(y).$$

Wir schreiben dann

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

oder " $f(x) \rightarrow y$ für $x \rightarrow a$ " und sagen, $f(x)$ *konvergiert* gegen y für x gegen a .

Proposition 40.2 (Folgenkriterium für Grenzwerte von Abbildungen)

Es sei $U \subseteq M$, $f: U \rightarrow M'$ eine Abbildung und a ein Häufungspunkt von U .

Dann sind die beiden folgenden Aussagen gleichwertig:

- a. $\lim_{x \rightarrow a} f(x) = y$.
- b. $\forall (a_n)_{n \in \mathbb{N}}$ mit $a_n \in U \setminus \{a\}$ und $\lim_{n \rightarrow \infty} a_n = a$ gilt $\lim_{n \rightarrow \infty} f(a_n) = y$.

Beweis: Der Beweis geht wörtlich wie der Beweis des Folgenkriteriums für Grenzwerte von Funktionen in \mathbb{R} , Proposition 13.7, wenn wir die Betragsdifferenz durch den Abstand d bzw. d' ersetzen.

a. \implies b.: Es sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in $U \setminus \{a\}$ mit $\lim_{n \rightarrow \infty} a_n = a$. Wir müssen

$\lim_{n \rightarrow \infty} f(a_n) = y$ zeigen. Dazu sei $\varepsilon > 0$ gegeben.

Wegen $\lim_{x \rightarrow a} f(x) = y$ gibt es ein $\delta_\varepsilon > 0$, so daß aus $x \in U$ mit $0 < d(x, a) < \delta_\varepsilon$ auch $d'(f(x), y) < \varepsilon$ folgt.

Wegen $\lim_{n \rightarrow \infty} a_n = a$ gibt es zu δ_ε nun ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $n \geq n_\varepsilon$ auch $d(a_n, a) < \delta_\varepsilon$ gilt.

Sei nun $n \geq n_\varepsilon$ dann erfüllt $a_n \in U$ die Bedingung $0 < d(a_n, a) < \delta_\varepsilon$ und somit ist auch $d'(f(a_n), y) < \varepsilon$. Damit ist $f(a_n) \rightarrow y$ gezeigt.

b. \implies a.: Wir nehmen an, \mathbf{y} wäre nicht der Grenzwert von f in \mathbf{a} . Dann gilt:

$$\exists \varepsilon > 0 : \forall \delta_\varepsilon > 0 \exists \mathbf{x}_{\delta_\varepsilon} \in \mathbf{U} \text{ mit } 0 < d(\mathbf{x}_{\delta_\varepsilon}, \mathbf{a}) < \delta_\varepsilon, \text{ aber } d'(f(\mathbf{x}_{\delta_\varepsilon}), \mathbf{y}) \geq \varepsilon.$$

Für $n \geq 1$ und $\delta_\varepsilon = \frac{1}{n}$ setzen wir $\mathbf{a}_n := \mathbf{x}_{\delta_\varepsilon} = \mathbf{x}_{\frac{1}{n}} \in \mathbf{U} \setminus \{\mathbf{a}\}$. Dann gilt

$$0 < d(\mathbf{a}_n, \mathbf{a}) < \frac{1}{n} \longrightarrow 0,$$

so daß $\mathbf{a}_n \longrightarrow \mathbf{a}$, und zugleich gilt

$$d'(f(\mathbf{a}_n), \mathbf{y}) \geq \varepsilon$$

für alle $n \in \mathbb{N}$. Dies ist ein Widerspruch dazu, daß $f(\mathbf{a}_n)$ gegen \mathbf{y} konvergieren muß.

□

Bemerkung 40.3 (Unabhängigkeit des Grenzwertes von der Norm)

Betrachten wir \mathbb{R}^n und \mathbb{R}^m als metrische Räume mittels einer (beliebigen) Norm und $\mathbf{U} \subseteq \mathbb{R}^n$. Dann ist der Grenzwert $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x})$ wegen Proposition 40.2 unabhängig von den gewählten Normen, da die Konvergenz von Folgen unabhängig von diesen ist.

Lemma 40.4

Es sei $\mathbf{U} \subseteq \mathbf{M}$, $f : \mathbf{U} \longrightarrow \mathbf{M}'$ eine Abbildung und \mathbf{a} ein Häufungspunkt von \mathbf{U} .

Existiert der Grenzwert $\lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x})$, so ist er eindeutig bestimmt.

Beweis: Dies folgt wie in Proposition 13.10 aus dem Folgenkriterium für Grenzwerte von Funktionen 40.2 und der Eindeutigkeit des Grenzwertes bei Folgen 39.23. Genauer, da \mathbf{a} ein Häufungspunkt von \mathbf{U} ist, gibt es nach Lemma 39.25 eine Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in $\mathbf{U} \setminus \{\mathbf{a}\}$ mit $\mathbf{a}_n \longrightarrow \mathbf{a}$. Falls nun \mathbf{y} und \mathbf{z} zwei Grenzwerte von $f(\mathbf{x})$ für $\mathbf{x} \rightarrow \mathbf{a}$ sind, so folgt mit den eben erwähnten Sätzen

$$\mathbf{y} = \lim_{n \rightarrow \infty} f(\mathbf{a}_n) = \mathbf{z}.$$

□

Beispiel 40.5 (Grenzwerte)

Sei $\mathbf{M} = \mathbb{R}^2$ und $\mathbf{U} = \mathbb{R}^2 \setminus \{(0, 0)^t\}$, dann ist $\mathbf{a} = (0, 0)^t$ ein Häufungspunkt von \mathbf{U} , da es Folgen in \mathbf{U} gibt, die gegen \mathbf{a} konvergieren.

a. Wir betrachten die Funktion

$$f : \mathbf{U} \longrightarrow \mathbb{R} : (\mathbf{x}, \mathbf{y})^t \mapsto \frac{\mathbf{x}\mathbf{y}}{\mathbf{x}^2 + \mathbf{y}^2}.$$

Wir betrachten die Folgen $(\mathbf{a}_n)_{n \geq 1}$ und $(\mathbf{b}_n)_{n \geq 1}$ mit

$$\mathbf{a}_n = \left(\frac{1}{n}, \frac{1}{n} \right)^t$$

und

$$\mathbf{b}_n = \left(\frac{1}{n}, 0 \right)^t.$$

Beide konvergieren gegen \mathbf{a} , aber es gilt

$$f(\mathbf{a}_n) = \frac{\frac{1}{n^2}}{\frac{1}{n^2} + \frac{1}{n^2}} = \frac{1}{2} \rightarrow \frac{1}{2}$$

und

$$f(\mathbf{b}_n) = 0 \rightarrow 0.$$

Mithin existiert der Grenzwert $\lim_{z \rightarrow \mathbf{a}} f(z)$ *nicht*!

Man beachte übrigens, daß die beiden Grenzwerte

$$\lim_{x \rightarrow 0} f(x, 0) = 0 \quad \text{und} \quad \lim_{y \rightarrow 0} f(0, y) = 0$$

existieren und übereinstimmen. Es reicht für die Existenz eines Grenzwertes in \mathbb{R}^n also *nicht*, die Existenz entlang der Koordinatenrichtungen zu testen!

b. Betrachten wir statt dessen die Funktion

$$g : \mathbf{U} \rightarrow \mathbb{R} : (x, y)^t \mapsto \frac{xy}{\sqrt{x^2 + y^2}},$$

so gilt für $(x, y) \in \mathbf{U}$

$$|g(x, y) - 0| = \frac{|x| \cdot |y|}{\sqrt{x^2 + y^2}} \leq \frac{\sqrt{x^2 + y^2} \cdot \sqrt{x^2 + y^2}}{\sqrt{x^2 + y^2}} = \sqrt{x^2 + y^2} = \|(x, y)^t\|_2.$$

Haben wir nun eine beliebige Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in \mathbf{U} , die gegen \mathbf{a} konvergiert, so gilt $\|\mathbf{a}_n\|_2 = \|\mathbf{a}_n - \mathbf{a}\|_2 \rightarrow 0$ und mithin

$$|g(\mathbf{a}_n) - 0| \rightarrow 0.$$

Wir erhalten deshalb

$$\lim_{z \rightarrow \mathbf{a}} g(z) = 0.$$

Proposition 40.6 (Grenzwertsätze für Abbildungen in normierten Räumen)

Es seien $f : \mathbf{U} \rightarrow \mathbf{V}$ und $g : \mathbf{U} \rightarrow \mathbf{V}$ zwei Funktionen mit $\mathbf{U} \subseteq \mathbf{M}$, \mathbf{a} ein Häufungspunkt von \mathbf{U} und $c \in \mathbb{R}$. Wenn $\lim_{x \rightarrow \mathbf{a}} f(x)$ und $\lim_{x \rightarrow \mathbf{a}} g(x)$ existieren, so gelten:

- $\lim_{x \rightarrow \mathbf{a}} (c \cdot f)(x) = c \cdot \lim_{x \rightarrow \mathbf{a}} f(x)$.
- $\lim_{x \rightarrow \mathbf{a}} (f + g)(x) = \lim_{x \rightarrow \mathbf{a}} f(x) + \lim_{x \rightarrow \mathbf{a}} g(x)$.
- $\lim_{x \rightarrow \mathbf{a}} (f - g)(x) = \lim_{x \rightarrow \mathbf{a}} f(x) - \lim_{x \rightarrow \mathbf{a}} g(x)$.
- Ist $\mathbf{V} = \mathbb{R}$, so ist auch $\lim_{x \rightarrow \mathbf{a}} (f \cdot g)(x) = \lim_{x \rightarrow \mathbf{a}} f(x) \cdot \lim_{x \rightarrow \mathbf{a}} g(x)$.
- Falls zudem $\mathbf{V} = \mathbb{R}$ und $\lim_{x \rightarrow \mathbf{a}} f(x) \neq 0$, so ist \mathbf{a} ein Häufungspunkt der Menge $X = \{x \in \mathbf{U} \mid f(x) \neq 0\}$ und es gilt

$$\lim_{x \rightarrow \mathbf{a}} \frac{1}{f}(x) = \frac{1}{\lim_{x \rightarrow \mathbf{a}} f(x)}.$$

Beweis: Die Aussagen a.-c. folgen aus dem Folgenkriterium für Grenzwerte von Funktionen 40.2 und den Grenzwertsätzen für Folgen 39.23 unter Berücksichtigung von Lemma 39.25. Die Aussage d. folgt analog, allerdings ersetzt man Proposition 39.23 durch Proposition 11.15.

Der Beweis für e. ist schließlich wörtlich der gleiche wie in Proposition 13.10. Wir führen ihn hier der Vollständigkeit halber noch mal auf. Nach Lemma 39.25 gibt es eine Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in $U \setminus \{\mathbf{a}\}$, die gegen \mathbf{a} konvergiert, und nach dem Folgenkriterium 40.2 gilt dann

$$f(\mathbf{a}_n) \longrightarrow \lim_{x \rightarrow \mathbf{a}} f(x) =: \mathbf{y}.$$

Wegen $\mathbf{y} \neq 0$ gibt es wegen der Grenzwertsätze für Folgen 11.15 ein n_0 , so daß $f(\mathbf{a}_n) \neq 0$ für alle $n \geq n_0$, so daß $(\mathbf{a}_n)_{n \geq n_0}$ eine Folge in X ist mit $\mathbf{a}_n \longrightarrow \mathbf{a}$. Nach Proposition 13.4 ist dann \mathbf{a} ein Häufungspunkt von X . Die Aussage zum Grenzwert folgt dann wieder aus den Grenzwertsätzen für Folgen 11.15 und dem Folgenkriterium 13.7. \square

B) Stetigkeit in metrischen und in normierten Räumen

Definition 40.7 (ε - δ -Kriterium für Stetigkeit)

Sei $f : M \longrightarrow M'$ eine Abbildung und $\mathbf{a} \in M$. Wir nennen f *stetig in \mathbf{a}* , wenn

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : \forall \mathbf{x} \in M \text{ mit } d(\mathbf{x}, \mathbf{a}) < \delta_\varepsilon \text{ gilt } d'(f(\mathbf{x}), f(\mathbf{a})) < \varepsilon,$$

d.h., wenn

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : f(U_{\delta_\varepsilon}(\mathbf{a})) \subseteq U_\varepsilon(f(\mathbf{a})).$$

Die Funktion f heißt *stetig* (auf M), wenn sie stetig in jedem Punkt in M ist.

Lemma 40.8

Es sei $f : M \longrightarrow M'$ und $\mathbf{a} \in M$ ein Häufungspunkt.

Genau dann ist f stetig in \mathbf{a} , wenn $\lim_{x \rightarrow \mathbf{a}} f(x) = f(\mathbf{a})$.

Beweis: Dies folgt unmittelbar aus den Definitionen 40.1 und 40.7. \square

Beispiel 40.9 (Metriken und Normen sind stetig.)

- a. Ist M ein metrischer Raum und $\mathbf{c} \in M$ fest gegeben. Dann ist die Abbildung

$$d(\cdot, \mathbf{c}) : M \longrightarrow \mathbb{R} : \mathbf{x} \mapsto d(\mathbf{x}, \mathbf{c})$$

stetig auf M . Um dies zu sehen, setzen wir $\delta_\varepsilon = \varepsilon$ für $\mathbf{a} \in M$ und $\varepsilon > 0$ und erhalten wegen der Dreiecksungleichung, daß

$$|d(\mathbf{x}, \mathbf{c}) - d(\mathbf{a}, \mathbf{c})| \leq d(\mathbf{x}, \mathbf{a}) < \varepsilon$$

gilt, sobald nur $d(\mathbf{x}, \mathbf{a}) < \delta_\varepsilon$ erfüllt ist.

- b. Ist V ein normierter Raum, so ist die Norm

$$\|\cdot\|_V : V \longrightarrow \mathbb{R}$$

stetig auf V . Dies folgt aus Teil a. mit $\mathbf{c} = 0$.

- c. Ist $\|\cdot\| : \mathbb{R}^n \rightarrow \mathbb{R}$ eine Norm auf \mathbb{R}^n , so ist sie stetig bezüglich der Maximumnorm. Um dies zu sehen, beachten wir, daß es ein $r > 0$ mit

$$\|\mathbf{x}\| \leq r \cdot \|\mathbf{x}\|_\infty$$

gibt, und betrachten zu $\mathbf{a} \in \mathbb{R}^n$ und $\varepsilon > 0$ dann $\delta_\varepsilon = \frac{\varepsilon}{r}$. Dann gilt für alle $\mathbf{x} \in \mathbb{R}^n$ mit $\|\mathbf{x} - \mathbf{a}\|_\infty < \delta_\varepsilon$ auch

$$|\|\mathbf{x}\| - \|\mathbf{a}\|| \leq \|\mathbf{x} - \mathbf{a}\| \leq r \cdot \|\mathbf{x} - \mathbf{a}\|_\infty < r \cdot \delta_\varepsilon = \varepsilon.$$

Mit Bemerkung 40.11 folgt die Aussage auch unmittelbar aus Teil b..

Satz 40.10 (Folgenkriterium für Stetigkeit)

Genau dann ist $f : M \rightarrow M'$ stetig in $\mathbf{a} \in M$, wenn

$$\forall (\mathbf{a}_n)_{n \in \mathbb{N}} \text{ mit } \mathbf{a}_n \in M \text{ und } \lim_{n \rightarrow \infty} \mathbf{a}_n = \mathbf{a} \text{ gilt } \lim_{n \rightarrow \infty} f(\mathbf{a}_n) = f(\mathbf{a}). \quad (133)$$

Beweis: Der Beweis geht genau wie der Beweis des Folgenkriteriums für Stetigkeit in \mathbb{R} , Satz 14.5, wenn man die Betragsdifferenz durch den Abstand d bzw. d' ersetzt.

“ \implies ”: Sei zunächst f stetig in \mathbf{a} und sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in M mit $\lim_{n \rightarrow \infty} \mathbf{a}_n = \mathbf{a}$. Wir müssen $\lim_{n \rightarrow \infty} f(\mathbf{a}_n) = f(\mathbf{a})$ zeigen. Dazu sei $\varepsilon > 0$ gegeben.

Da f stetig in \mathbf{a} ist, gibt es ein $\delta_\varepsilon > 0$, so daß aus $\mathbf{x} \in M$ mit $d(\mathbf{x}, \mathbf{a}) < \delta_\varepsilon$ auch $d'(f(\mathbf{x}), f(\mathbf{a})) < \varepsilon$ folgt.

Wegen $\lim_{n \rightarrow \infty} \mathbf{a}_n = \mathbf{a}$ gibt es zu δ_ε nun ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $n \geq n_\varepsilon$ auch $d(\mathbf{a}_n, \mathbf{a}) < \delta_\varepsilon$ gilt.

Sei nun $n \geq n_\varepsilon$ dann erfüllt $\mathbf{a}_n \in M$ die Bedingung $d(\mathbf{a}_n, \mathbf{a}) < \delta_\varepsilon$ und somit ist auch $d'(f(\mathbf{a}_n), f(\mathbf{a})) < \varepsilon$. Damit ist $f(\mathbf{a}_n) \rightarrow f(\mathbf{a})$ gezeigt.

“ \impliedby ”: Wir nehmen an, f wäre nicht stetig in \mathbf{a} . Dann gilt:

$$\exists \varepsilon > 0 : \forall \delta_\varepsilon > 0 \exists \mathbf{x}_{\delta_\varepsilon} \in M \text{ mit } d(\mathbf{x}_{\delta_\varepsilon}, \mathbf{a}) < \delta_\varepsilon, \text{ aber } d'(f(\mathbf{x}_{\delta_\varepsilon}), f(\mathbf{a})) \geq \varepsilon.$$

Für $n \geq 1$ und $\delta_\varepsilon = \frac{1}{n}$ setzen wir $\mathbf{a}_n := \mathbf{x}_{\delta_\varepsilon} = \mathbf{x}_{\frac{1}{n}} \in M \setminus \{\mathbf{a}\}$. Dann gilt

$$0 \leq d(\mathbf{a}_n, \mathbf{a}) < \frac{1}{n} \rightarrow 0,$$

so daß $\mathbf{a}_n \rightarrow \mathbf{a}$, und zugleich gilt

$$d'(f(\mathbf{a}_n), f(\mathbf{a})) \geq \varepsilon$$

für alle $n \in \mathbb{N}$. Dies ist ein Widerspruch dazu, daß $f(\mathbf{a}_n)$ gegen $f(\mathbf{a})$ konvergieren muß. □

Bemerkung 40.11 (Stetigkeit unabhängig von gewählter Norm)

Betrachten wir \mathbb{R}^n und \mathbb{R}^m als metrische Räume mittels einer (beliebigen) Norm und $\mathbf{U} \subseteq \mathbb{R}^n$ als metrischen Raum mittels der Einschränkung der Metrik. Dann ist es für die Frage, ob eine Abbildung $f : \mathbf{U} \rightarrow \mathbb{R}^m$ stetig ist, wegen des Folgenkriteriums der Stetigkeit 40.10 unerheblich welche Normen wir auf \mathbb{R}^n und \mathbb{R}^m gewählt haben, da Konvergenz einer Folge nicht von der gewählten Norm abhängt.

Beispiel 40.12 (Stetigkeit von $x \mapsto x_i$)

Die Funktion $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R} : (x_1, \dots, x_n)^t \mapsto x_i$ ist stetig auf \mathbb{R}^n .

Um dies zu sehen, betrachten wir eine Folge $(\mathbf{a}_k)_{k \in \mathbb{N}}$ in \mathbb{R}^n , die in der Maximumsnorm gegen $\mathbf{c} = (c_1, \dots, c_n)^t$ konvergiert. Dann gilt

$$|\pi_i(\mathbf{a}_k) - \pi_i(\mathbf{c})| = |a_{ki} - c_i| \leq \|\mathbf{a}_k - \mathbf{c}\|_\infty \rightarrow 0$$

Mithin ist π_i stetig in \mathbf{c} .

Proposition 40.13 (Grenzwertsätze für stetige Abbildungen)

Seien $f : M \rightarrow V$ und $g : M \rightarrow V$ stetig in $\mathbf{a} \in M$ und sei $\mathbf{c} \in \mathbb{R}$.

- $\mathbf{c} \cdot f$, $f + g$ und $f - g$ sind stetig in \mathbf{a} .
- Ist $V = \mathbb{R}$, so ist auch $f \cdot g$ stetig in \mathbf{a} .
- Ist $V = \mathbb{R}$ und $g(\mathbf{a}) \neq 0$, so ist auch $\frac{f}{g} : M \setminus g^{-1}(0) \rightarrow \mathbb{R}$ stetig in \mathbf{a} .

Beweis: Der Beweis folgt aus dem Folgenkriterium für Stetigkeit 40.10 und den Grenzwertsätzen für Folgen in V , Proposition 39.23, bzw. in \mathbb{R} , Proposition 11.15.

Z.B. sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in M mit $\mathbf{a}_n \rightarrow \mathbf{a}$, dann gilt

$$(f + g)(\mathbf{a}_n) = f(\mathbf{a}_n) + g(\mathbf{a}_n) \rightarrow f(\mathbf{a}) + g(\mathbf{a}) = (f + g)(\mathbf{a}),$$

da f und g in \mathbf{a} stetig sind. Also ist auch $f + g$ stetig in \mathbf{a} . □

Beispiel 40.14 (Polynomfunktionen)

Sind t_1, \dots, t_n Veränderliche und ist $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$, so setzen wir

$$\mathbf{t} = (t_1, \dots, t_n)$$

und

$$\mathbf{t}^\alpha = t_1^{\alpha_1} \cdot \dots \cdot t_n^{\alpha_n}.$$

Ferner nennen wir

$$|\alpha| = \alpha_1 + \dots + \alpha_n$$

den *Grad* des *Monoms* \mathbf{t}^α . Für reelle Zahlen $\mathbf{a}_\alpha \in \mathbb{R}$ mit $\alpha \in \mathbb{N}^n$ und $0 \leq |\alpha| \leq d$ nennen wir

$$p = \sum_{|\alpha|=0}^d \mathbf{a}_\alpha \cdot \mathbf{t}^\alpha = \sum_{\substack{\alpha \in \mathbb{N}^n \\ 0 \leq |\alpha| \leq d}} \mathbf{a}_\alpha \cdot t_1^{\alpha_1} \cdot \dots \cdot t_n^{\alpha_n}$$

ein *Polynom* in den Veränderlichen t_1, \dots, t_n . Wir definieren den *Grad* des Polynoms p als

$$\deg(p) := \sup \{ |\alpha| \mid \mathbf{a}_\alpha \neq 0 \} \in \mathbb{N} \cup \{-\infty\},$$

wobei automatisch $\deg(0) := -\infty$ gilt. Mit $\mathbb{R}[t] = \mathbb{R}[t_1, \dots, t_n]$ bezeichnen wir die Menge aller Polynome in den Veränderlichen t_1, \dots, t_n , so daß der Grad eine Abbildung $\deg : \mathbb{R}[t] \rightarrow \mathbb{N} \cup \{-\infty\}$ ist.

Für ein Polynom $p = \sum_{|\alpha|=0}^d a_\alpha \cdot t^\alpha$ und ein $x = (x_1, \dots, x_n)^t \in \mathbb{R}^n$ setzen wir schließlich

$$p(x) := \sum_{|\alpha|=0}^d a_\alpha \cdot x^\alpha = \sum_{\substack{\alpha \in \mathbb{N}^n \\ 0 \leq |\alpha| \leq d}} a_\alpha \cdot x_1^{\alpha_1} \cdot \dots \cdot x_n^{\alpha_n},$$

wobei wir uns wieder die *Multiindexnotation* $x^\alpha = x_1^{\alpha_1} \cdot \dots \cdot x_n^{\alpha_n}$ zu eigen machen.

Sind $p, q \in \mathbb{R}[t]$ zwei Polynome, $q \neq 0$ nicht das Nullpolynom, so nennen wir die Funktion

$$p : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto p(x)$$

eine *Polynomfunktion* und die Funktion

$$\frac{p}{q} : \mathbb{R}^n \setminus \{x \in \mathbb{R}^n \mid q(x) = 0\} \longrightarrow \mathbb{R} : x \mapsto \frac{p(x)}{q(x)}$$

nennen wir eine *rationale Funktion*.

Polynomfunktionen und rationale Funktionen sind *stetig* auf ihrem Definitionsbereich. Dies folgt unmittelbar aus den Grenzwertsätzen für stetige Funktionen 40.13 und Beispiel 40.12.

Proposition 40.15 (Komposition stetiger Funktionen)

Seien (M, d) , (M', d') und (M'', d'') metrische Räume, $f : M \rightarrow M'$, $g : M' \rightarrow M''$. Ist f stetig in $a \in M$ und g stetig in $f(a)$, so ist $g \circ f$ stetig in a .

Beweis: Der Beweis geht genau wie im eindimensionalen Fall, wenn man die Betragsdifferenz durch die Abstände ersetzt.

Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in M mit $a_n \rightarrow a$, dann ist $(f(a_n))_{n \in \mathbb{N}}$ eine Folge in M' und, da f stetig in a ist, gilt zudem $f(a_n) \rightarrow f(a)$. Nun ist auch g stetig in $f(a)$, so daß daraus

$$(g \circ f)(a_n) = g(f(a_n)) \longrightarrow g(f(a)) = (g \circ f)(a)$$

folgt. Aufgrund des Folgenkriteriums für Stetigkeit 40.10 ist dann $g \circ f$ stetig in a . \square

Beispiel 40.16

Die Funktion

$$\mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y)^t \mapsto \sqrt{x^2 + y^2}$$

ist als Komposition der stetigen Funktionen $\mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y)^t \mapsto x^2 + y^2$ und $\mathbb{R}_{\geq 0} \longrightarrow \mathbb{R} : x \mapsto \sqrt{x}$ stetig auf \mathbb{R}^2 .

Proposition 40.17 (Stetige Urbilder offener Mengen sind offen.)

Sei $f : M \rightarrow M'$ eine Abbildung. Dann sind die folgenden Aussagen gleichwertig:

- f ist stetig auf M .
- Das Urbild jeder offenen Menge in M' ist offen in M .
- Das Urbild jeder abgeschlossenen Menge in M' ist abgeschlossen in M .

Beweis:

a. \implies b.: Sei zunächst f stetig und $O' \subseteq M'$ sei offen in M' . Da O' offen in M' ist, gibt es zu jedem $\mathbf{a} \in f^{-1}(O')$ ein $\varepsilon_{\mathbf{a}} > 0$ mit $U_{\varepsilon_{\mathbf{a}}}(f(\mathbf{a})) \subseteq O'$, und da f stetig in \mathbf{a} ist, gibt es ein $\delta_{\mathbf{a}} > 0$ mit

$$f(U_{\delta_{\mathbf{a}}}(\mathbf{a})) \subseteq U_{\varepsilon_{\mathbf{a}}}(f(\mathbf{a})) \subseteq O'.$$

Die Menge

$$O = \bigcup_{\mathbf{a} \in f^{-1}(O')} U_{\delta_{\mathbf{a}}}(\mathbf{a})$$

ist als Vereinigung offener Mengen offen in M . Außerdem gilt

$$f(O) = \bigcup_{\mathbf{a} \in f^{-1}(O')} f(U_{\delta_{\mathbf{a}}}(\mathbf{a})) \subseteq \bigcup_{\mathbf{a} \in f^{-1}(O')} U_{\varepsilon_{\mathbf{a}}}(f(\mathbf{a})) \subseteq O'.$$

Wenn wir nun noch berücksichtigen, daß O jedes $\mathbf{a} \in f^{-1}(O')$ enthält, so erhalten wir

$$f^{-1}(O') = O$$

ist offen in M .

b. \implies a.: Sei umgekehrt das Urbild jeder offenen Menge offen in M , und seien $\mathbf{a} \in M$ und $\varepsilon > 0$ beliebig. Dann ist $f^{-1}(U_{\varepsilon}(f(\mathbf{a})))$ offen in M . Deshalb gibt es ein $\delta > 0$ mit

$$U_{\delta}(\mathbf{a}) \subseteq f^{-1}(U_{\varepsilon}(f(\mathbf{a})))$$

und mithin

$$f(U_{\delta}(\mathbf{a})) \subseteq f\left(f^{-1}(U_{\varepsilon}(f(\mathbf{a})))\right) \subseteq U_{\varepsilon}(f(\mathbf{a})).$$

f ist also stetig in \mathbf{a} , und somit stetig auf M .

b. \implies c.: Ist $A' \subseteq M'$ abgeschlossen in M' , dann ist $M' \setminus A'$ offen in M' . Nach Voraussetzung ist somit

$$f^{-1}(M' \setminus A') = M \setminus f^{-1}(A')$$

offen in M , so daß $f^{-1}(A')$ abgeschlossen in M ist.

c. \implies b.: Ist $O' \subseteq M'$ offen in M' , dann ist $M' \setminus O'$ abgeschlossen in M' . Nach Voraussetzung ist somit

$$f^{-1}(M' \setminus O') = M \setminus f^{-1}(O')$$

abgeschlossen in M , so daß $f^{-1}(O')$ offen in M ist.

□

C) Stetigkeit für Abbildungen nach \mathbb{R}^m

Definition 40.18

Ist $f : M \rightarrow \mathbb{R}^m : x \mapsto (f_1(x), \dots, f_m(x))^t$ eine Abbildung, so heißen die Funktionen $f_i : M \rightarrow \mathbb{R}$ die *Komponentenfunktionen* von f .

Satz 40.19 (Komponentenweise Stetigkeit)

Es sei $f : M \rightarrow \mathbb{R}^m$ eine Abbildung und $\mathbf{a} \in M$.

- Genau dann existiert der Grenzwert $\lim_{x \rightarrow \mathbf{a}} f(x)$, wenn die Grenzwerte $\lim_{x \rightarrow \mathbf{a}} f_i(x)$ aller Komponentenfunktionen existieren.
- Genau dann ist f stetig in \mathbf{a} , wenn alle Komponentenfunktionen stetig in \mathbf{a} sind.

Beweis: Sei $(\mathbf{a}_n)_{n \in \mathbb{N}}$ eine Folge in M , die gegen \mathbf{a} konvergiert. Wegen des Satzes zur komponentenweisen Konvergenz von Folgen in \mathbb{R}^m , Proposition 39.38, gilt $f(\mathbf{a}_n) \rightarrow f(\mathbf{a})$ genau dann, wenn $f_i(\mathbf{a}_n) \rightarrow f_i(\mathbf{a})$ für alle $i = 1, \dots, m$. Die Aussage folgt also aus dem Folgenkriterium für Grenzwerte 40.2 bzw. für Stetigkeit 40.10. \square

Beispiel 40.20

- Die Abbildung

$$\mathbb{R}^2 \setminus \{(0, 0)^t\} \rightarrow \mathbb{R}^2 : (x, y)^t \mapsto \left(\frac{xy}{x^2 + y^2}, x^3 + 5xy \right)^t$$

ist stetig auf $\mathbb{R}^2 \setminus \{(0, 0)^t\}$, da die Komponentenfunktionen nach Beispiel 40.14 stetig sind.

- Ist $A = (a_{ij}) \in \text{Mat}(m \times n, \mathbb{R})$, so hat die lineare Abbildung

$$f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m : x \mapsto Ax$$

die Komponentenfunktionen

$$f_i : \mathbb{R}^n \rightarrow \mathbb{R} : (x_1, \dots, x_n)^t \mapsto a_{i1} \cdot x_1 + \dots + a_{in} \cdot x_n.$$

Als Polynomfunktionen sind diese stetig, so daß auch f_A stetig ist.

Beispiel 40.21 (Determinante und Invertieren als stetige Abbildungen)

Wir können $\text{Mat}_n(\mathbb{R})$ mit dem Vektorraum \mathbb{R}^{n^2} identifizieren, so daß $\text{Mat}_n(\mathbb{R})$ ein normierter Raum ist. Die Determinante

$$\det : \text{Mat}_n(\mathbb{R}) \rightarrow \mathbb{R}$$

ist aufgrund der Leibnitz-Formel eine Polynomfunktion und somit stetig. Dann ist das Urbild

$$\det^{-1}(0) = \{A \in \text{Mat}_n(\mathbb{R}) \mid \det(A) = 0\}$$

abgeschlossen in $\text{Mat}_n(\mathbb{R})$ als Urbild der abgeschlossenen Menge $\{0\}$ in \mathbb{R} . Ihr Komplement

$$\text{Gl}_n(\mathbb{R}) = \{A \in \text{Mat}_n(\mathbb{R}) \mid \det(A) \neq 0\},$$

die Menge der invertierbaren $\mathbf{n} \times \mathbf{n}$ -Matrizen, ist somit offen in $\text{Mat}_{\mathbf{n}}(\mathbb{R})$.

Betrachten wir nun die Abbildung

$$\text{Gl}_{\mathbf{n}}(\mathbb{R}) \longrightarrow \text{Mat}_{\mathbf{n}}(\mathbb{R}) : \mathbf{A} \mapsto \mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})} \cdot \mathbf{A}^{\#},$$

die einer invertierbaren Matrix ihre Inverse zuordnet, so sind ihre Komponentenfunktionen wegen des Satzes über die Adjunkte, Korollar 30.29, rationale Funktionen in den Einträgen von \mathbf{A} , und mithin ist die Abbildung stetig! D.h. eine Matrix zu invertieren ist ein stetiger Prozeß, was wiederum bedeutet, daß eine kleine Störung in den Einträgen von \mathbf{A} nur zu einer kleinen Störung in den Einträgen der Inversen führt! Gleiches gilt damit für die Lösung eines eindeutig lösbaren Gleichungssystems.

Bemerkung 40.22 (Stetigkeit ist stärker als getrennte Stetigkeit)

Ist $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ eine Abbildung und $\mathbf{y} = (\mathbf{y}_1, \dots, \mathbf{y}_n)^t \in \mathbb{R}^n$, so können wir die Funktionen

$$F_i : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto f(\mathbf{y}_1, \dots, \mathbf{y}_{i-1}, x, \mathbf{y}_{i+1}, \dots, \mathbf{y}_n)$$

definieren.

- a. Ist f stetig in \mathbf{y} , so sind die F_i alle stetig in \mathbf{y}_i .

Denn ist $(\mathbf{a}_k)_{k \in \mathbb{N}}$ eine Folge in \mathbb{R} , die gegen \mathbf{y}_i konvergiert, so konvergiert auch die Folge der $(\mathbf{y}_1, \dots, \mathbf{y}_{i-1}, \mathbf{a}_k, \mathbf{y}_{i+1}, \dots, \mathbf{y}_n)$ gegen \mathbf{y} und somit gilt

$$F_i(\mathbf{a}_k) = f(\mathbf{y}_1, \dots, \mathbf{y}_{i-1}, \mathbf{a}_k, \mathbf{y}_{i+1}, \dots, \mathbf{y}_n) \longrightarrow f(\mathbf{y}) = F_i(\mathbf{y}_i).$$

Man sagt, daß Stetigkeit die getrennte Stetigkeit in Richtung der Koordinatenachsen impliziert.

- b. Die Umkehrung gilt nicht, wie im wesentlichen aus Beispiel 40.5 folgt. Setzen wir

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (\mathbf{x}, \mathbf{y})^t \mapsto \begin{cases} \frac{\mathbf{x}\mathbf{y}}{\mathbf{x}^2 + \mathbf{y}^2}, & \text{falls } (\mathbf{x}, \mathbf{y})^t \neq (0, 0)^t, \\ 0, & \text{falls } (\mathbf{x}, \mathbf{y})^t = (0, 0)^t, \end{cases}$$

so ist die Funktion nicht stetig in $\mathbf{a} = (0, 0)^t$, da der Grenzwert von $f(\mathbf{x}, \mathbf{y})$ in \mathbf{a} nach Beispiel 40.5 nicht existiert. Hingegen sind die Funktionen

$$F_1 : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto f(x, 0)$$

und

$$F_2 : \mathbb{R} \longrightarrow \mathbb{R} : y \mapsto f(0, y)$$

jeweils stetig in 0, wie wir ebenfalls aus Beispiel 40.5 ablesen können.

D) Stetige Funktionen auf kompakten Mengen

Satz 40.23 (Stetige Bilder von Kompakta sind kompakt.)

Ist M kompakt und $f : M \longrightarrow M'$ stetig, so ist $f(M)$ kompakt in M' .

Beweis: Ist $(\mathbf{b}_n)_{n \in \mathbb{N}}$ eine Folge in $f(M)$, so besitzt jedes \mathbf{b}_n ein Urbild $\mathbf{a}_n \in M$ unter f , d.h. wir haben eine Folge $(\mathbf{a}_n)_{n \in \mathbb{N}}$ in M mit $f(\mathbf{a}_n) = \mathbf{b}_n$. Da M kompakt ist, besitzt $(\mathbf{a}_n)_{n \in \mathbb{N}}$ dann eine konvergente Teilfolge $(\mathbf{a}_{n_k})_{k \in \mathbb{N}}$ mit Grenzwert $\mathbf{a} \in M$. Da f stetig ist, gilt dann aber auch

$$\mathbf{b}_{n_k} = f(\mathbf{a}_{n_k}) \longrightarrow f(\mathbf{a}) \in f(M),$$

so daß auch $(\mathbf{b}_n)_{n \in \mathbb{N}}$ eine konvergente Teilfolge mit Grenzwert in $f(M)$ besitzt. Aus dem Satz von Bolzano-Weierstraß 39.30 folgt dann, daß $f(M)$ kompakt ist. \square

Der obige Beweis von Satz 40.23 funktioniert nur in metrischen Räumen, da ansonsten der Satz von Bolzano-Weierstraß nicht gelten muß. Mit Hilfe der Charakterisierung von Stetigkeit in Proposition 40.17 kann man aber alternativ auch einen Beweis geben.

Alternativer Beweis für Satz 40.23: Sei $(O_i)_{i \in I}$ eine offene Überdeckung von $f(A)$ in M' . Da f stetig ist, ist $f^{-1}(O_i)$ wegen Proposition 40.17 offen in M und

$$A \subseteq f^{-1}(f(A)) \subseteq \bigcup_{i \in I} f^{-1}(O_i).$$

Also ist $(f^{-1}(O_i))_{i \in I}$ eine offene Überdeckung von A in M , und da A kompakt ist, gibt es eine endliche Teilüberdeckung. D.h. es gibt $i_1, \dots, i_n \in I$ mit

$$A \subseteq f^{-1}(O_{i_1}) \cup \dots \cup f^{-1}(O_{i_n}).$$

Wenden wir f auf beide Seiten an, so erhalten wir

$$f(A) \subseteq f(f^{-1}(O_{i_1}) \cup \dots \cup f^{-1}(O_{i_n})) \subseteq O_{i_1} \cup \dots \cup O_{i_n}.$$

Mithin besitzt die offene Überdeckung $(O_i)_{i \in I}$ eine endliche Teilüberdeckung, und somit ist $f(A)$ kompakt. \square

Das folgende Korollar ist eine direkte Verallgemeinerung von Satz 14.16.

Korollar 40.24 (Maxima & Minima stetiger Funktionen auf Kompakta)

Sei $f : M \longrightarrow \mathbb{R}$ stetig und M kompakt, dann gibt es $\mathbf{c}, \mathbf{d} \in M$ mit

$$f(\mathbf{c}) \leq f(\mathbf{x}) \leq f(\mathbf{d})$$

für alle $\mathbf{x} \in M$.

Beweis: Wegen Satz 40.23 ist $f(M)$ kompakt in \mathbb{R} und wegen Korollar 39.46 enthält $f(M)$ dann sein Maximum und sein Minimum. \square

Korollar 40.25 (Umkehrsatz für injektive Abbildungen auf Kompakta)

Ist M kompakt und $f : M \longrightarrow M'$ stetig und injektiv, so ist $f^{-1} : f(M) \longrightarrow M$ stetig.

Beweis: Da f injektiv ist, bildet f bijektiv auf sein Bild ab und die Abbildung

$$f^{-1} : f(M) \longrightarrow M$$

ist definiert. Wegen Proposition 40.17 reicht es, zu zeigen, daß das Urbild jeder abgeschlossenen Menge in M unter f^{-1} abgeschlossen in $f(M)$ ist, d.h. für $A \subseteq M$ abgeschlossen, müssen wir zeigen, daß

$$f(A) = (f^{-1})^{-1}(A) \subseteq f(M)$$

abgeschlossen in $f(M)$ ist.

Da A abgeschlossen im kompakten metrischen Raum M ist, ist A nach Korollar 39.32 ebenfalls kompakt. Dann ist aber $f(A)$ kompakt in M' nach Satz 40.23. Als kompakte Menge ist $f(A)$ mithin abgeschlossen in M' nach Satz 39.21. Aber dann ist die Menge

$$f(A) = f(A) \cap f(M)$$

abgeschlossen in $f(M)$ nach Aufgabe 39.52. □

Beispiel 40.26

Die Kompaktheit des Definitionsbereiches im Umkehrsatz 40.25 ist eine notwendige Bedingung. Um dies zu sehen, betrachten wir die Abbildung

$$f : [0, 2\pi) \longrightarrow \mathbb{R}^2 : x \mapsto (\cos(x), \sin(x))^t.$$

f bildet bijektiv auf sein Bild ab, den Einheitskreis mit dem Ursprung als Mittelpunkt

$$\text{Im}(f) = S^1 = \{(x, y)^t \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}.$$

Die Umkehrabbildung

$$f^{-1} : S^1 \longrightarrow [0, 2\pi)$$

ist aber nicht stetig, da es keine δ -Umgebung des Punktes $\mathbf{a} = (1, 0)^t \in S^1$ gibt, die ganz in die 1-Umgebung des Punktes 0 in $[0, 2\pi)$ abgebildet wird; das Bild jeder noch so kleinen Umgebung von \mathbf{a} enthält Zahlen nahe bei 2π .

E) Gleichmäßige Stetigkeit

Definition 40.27 (Gleichmäßige Stetigkeit)

Eine Abbildung $f : M \longrightarrow M'$ heißt *gleichmäßig stetig* auf M , wenn

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : \forall x, y \in M \text{ mit } d(x, y) < \delta_\varepsilon \text{ gilt } d'(f(x), f(y)) < \varepsilon,$$

d.h.

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : \forall x \in M \text{ gilt } f(\mathbf{U}_{\delta_\varepsilon}(x)) \subseteq \mathbf{U}_\varepsilon(f(x)).$$

Bemerkung 40.28 (Gleichmäßige Stetigkeit impliziert Stetigkeit.)

Jede auf M gleichmäßig stetige Funktion $f : M \longrightarrow M'$ ist auch stetig auf M .

Der folgende Satz ist eine direkte Verallgemeinerung von Satz 14.28

Satz 40.29 (Stetigkeit auf Kompakta impliziert gleichmäßige Stetigkeit)

Ist M kompakt und $f : M \rightarrow M'$ stetig, so ist f gleichmäßig stetig auf M .

Beweis: Der Beweis geht wörtlich wie der Beweis von Satz 14.28, wenn wir $[a, b]$ durch M und die Betragsdifferenzen durch die Abstände ersetzen.

Angenommen, f wäre nicht gleichmäßig stetig auf M . Dann gilt:

$$\exists \varepsilon > 0 : \forall \delta_\varepsilon > 0 \exists x_{\delta_\varepsilon}, y_{\delta_\varepsilon} \in M : d(x_{\delta_\varepsilon}, y_{\delta_\varepsilon}) < \delta_\varepsilon, \text{ aber } d'(f(x_{\delta_\varepsilon}), f(y_{\delta_\varepsilon})) \geq \varepsilon.$$

Für $n \geq 1$ und $\delta_\varepsilon := \frac{1}{n}$ setzen wir $\mathbf{a}_n := x_{\delta_\varepsilon} = x_{\frac{1}{n}}$ und $\mathbf{b}_n := y_{\delta_\varepsilon} = y_{\frac{1}{n}}$. Damit erhalten wir zwei Folgen $(\mathbf{a}_n)_{n \geq 1}$ und $(\mathbf{b}_n)_{n \geq 1}$ in der kompakten Menge M . Nach dem Satz von Bolzano-Weierstraß 39.30 besitzt $(\mathbf{a}_n)_{n \geq 1}$ eine konvergente Teilfolge $(\mathbf{a}_{n_k})_{k \in \mathbb{N}}$, und ebenso besitzt dann $(\mathbf{b}_{n_k})_{k \in \mathbb{N}}$ eine konvergente Teilfolge $(\mathbf{b}_{n_{k_l}})_{l \in \mathbb{N}}$. Nach Konstruktion gilt

$$0 \leq d(\mathbf{a}_{n_{k_l}}, \mathbf{b}_{n_{k_l}}) \leq \frac{1}{n_{k_l}} \rightarrow 0,$$

so daß die Grenzwerte von $(\mathbf{a}_{n_{k_l}})_{l \in \mathbb{N}}$ und $(\mathbf{b}_{n_{k_l}})_{l \in \mathbb{N}}$ wegen Proposition 39.23 g. übereinstimmen müssen, d.h.

$$\mathbf{a}_{n_{k_l}} \rightarrow \mathbf{y} \quad \text{und} \quad \mathbf{b}_{n_{k_l}} \rightarrow \mathbf{y}.$$

Da f stetig ist, folgt unter Berücksichtigung von Proposition 39.23 g.

$$0 = d'(f(\mathbf{y}), f(\mathbf{y})) \leftarrow d'(f(\mathbf{a}_{n_{k_l}}), f(\mathbf{b}_{n_{k_l}})) \geq \varepsilon,$$

was ein offensichtlicher Widerspruch ist. □

Beispiel 40.30

Es sei

$$M = \{(x_1, x_2)^t \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$$

die Kreisscheibe vom Radius 1 in der Ebene \mathbb{R}^2 . Wir betrachten M als metrischen Raum mittels der Maximumsnorm auf \mathbb{R}^2 . Die Abbildung

$$f : M \rightarrow \mathbb{R} : \mathbf{x} = (x_1, x_2)^t \mapsto x_1^2 + x_2^2$$

ist dann als Polynomfunktion stetig auf M und M ist kompakt.

Mithin ist f gleichmäßig stetig auf M , und für alle $\varepsilon > 0$ gibt es also ein $\delta_\varepsilon > 0$, so daß für $\max\{|x_1 - y_1|, |x_2 - y_2|\} = \|\mathbf{x} - \mathbf{y}\|_\infty < \delta_\varepsilon$ schon

$$|x_1^2 + x_2^2 - y_1^2 - y_2^2| = |f(\mathbf{x}) - f(\mathbf{y})| < \varepsilon$$

gilt. Wir können hierzu $\delta_\varepsilon = \frac{\varepsilon}{4}$ wählen, denn dann gilt

$$\begin{aligned} |f(\mathbf{x}) - f(\mathbf{y})| &\leq |x_1^2 - y_1^2| + |x_2^2 - y_2^2| = |x_1 + y_1| \cdot |x_1 - y_1| + |x_2 + y_2| \cdot |x_2 - y_2| \\ &\leq (|x_1| + |y_1|) \cdot |x_1 - y_1| + (|x_2| + |y_2|) \cdot |x_2 - y_2| \leq 2 \cdot \delta_\varepsilon + 2 \cdot \delta_\varepsilon = \varepsilon, \end{aligned}$$

da $|x_i| \leq \sqrt{x_1^2 + x_2^2} \leq 1$ und $|y_i| \leq \sqrt{y_1^2 + y_2^2} \leq 1$. Wir sehen bei der Abschätzung auch, daß der geringste Spielraum bei den Punkten auf dem Rand von M ist, was

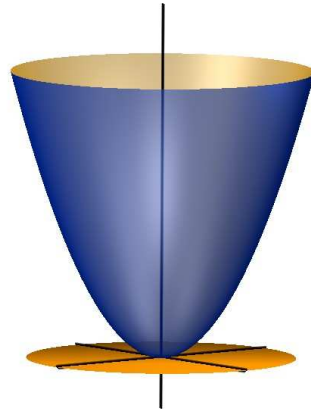


ABBILDUNG 3. Der Graph von $f : M \rightarrow \mathbb{R} : (x_1, x_2)^t \mapsto x_1^2 + x_2^2$

daran liegt, daß in diesen die Steigung des Graphen am größten ist. Mit solchen Betrachtungen beschäftigen wir uns im Rahmen der Differentialrechnung in \mathbb{R}^n nochmal ausführlicher.

F) Lipschitz-Stetigkeit

Eine besonders starke Form der Stetigkeit ist die Lipschitz-Stetigkeit, die wir im Zusammenhang mit dem Banachschen Fixpunktsatz 43.4 und dem mehrdimensionalen Riemann-Integral benötigen.

Definition 40.31 (Lipschitz-stetig)

Eine Abbildung $f : M \rightarrow M'$ heißt *Lipschitz-stetig* auf M mit Lipschitzkonstante $q \in \mathbb{R}_{>0}$, wenn

$$d'(f(x), f(y)) \leq q \cdot d(x, y)$$

für alle $x, y \in M$ gilt (siehe auch Aufgabe 14.33).

Beispiel 40.32

Ist $a \in \mathbb{R}$, so ist die lineare Abbildung $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto a \cdot x$ Lipschitz-stetig mit Lipschitz-Konstante $q = |a|$, da

$$|f(x) - f(y)| = |ax - ay| = |a| \cdot |x - y|.$$

Proposition 40.33 (Lipschitz stetige Abbildungen sind gleichmäßig stetig.)

Jede Lipschitz stetige Abbildung $f : M \rightarrow M'$ ist gleichmäßig stetig auf M .

Beweis: Sei $\varepsilon > 0$ vorgegeben und sei $q > 0$ die Lipschitzkonstante von f , so setzen wir $\delta_\varepsilon := \frac{\varepsilon}{q} > 0$ und erhalten für $x, y \in M$ mit $d(x, y) < \delta_\varepsilon$ auch

$$d'(f(x), f(y)) \leq q \cdot d(x, y) < q \cdot \delta_\varepsilon = \varepsilon.$$

□

Beispiel 40.34 (Gleichmäßig stetig impliziert nicht Lipschitz-stetig)

Die Wurzelfunktion $\sqrt{\cdot} : [0, \infty) \rightarrow \mathbb{R} : x \mapsto \sqrt{x}$ ist nach Aufgabe 14.31 gleichmäßig

stetig auf $[0, \infty)$. Nehmen wir an, $\sqrt{\cdot}$ sei auch Lipschitz-stetig mit Lipschitz-Konstante $q > 0$. Wir wählen nun eine positive reelle Zahl $0 \neq x \in [0, \infty)$ mit

$$\sqrt{x} < \frac{1}{q}$$

und erhalten dann

$$|\sqrt{x} - \sqrt{0}| = \frac{x}{\sqrt{x}} > q \cdot |x - 0|$$

im Widerspruch zur Annahme der Lipschitz-Stetigkeit. Die Wurzelfunktion ist also nicht Lipschitz-stetig auf $[0, \infty)$.

Bemerkung 40.35 (Lipschitz-Stetigkeit auf \mathbb{R}^n)

Lipschitz-Stetigkeit für Abbildungen $f : U \rightarrow \mathbb{R}^m$ mit $U \subseteq \mathbb{R}^n$ ist unabhängig von den gewählten Normen. Beim Wechsel der Norm ändert sich die Lipschitz-Konstante durch die Äquivalenzkonstanten der Norm (siehe Lemma 39.41).

G) Konvergenz von Folgen von Abbildungen

Definition 40.36 (Punktweise und gleichmäßige Konvergenz)

- Für jedes $n \in \mathbb{N}$ sei $f_n : M \rightarrow M'$ eine Abbildung, so nennen wir $(f_n)_{n \in \mathbb{N}}$ eine *Folge von Abbildungen* auf M .
- Wir nennen die Folge $(f_n)_{n \in \mathbb{N}}$ von Abbildungen *punktweise konvergent auf M* , wenn für jedes $x \in M$ der Grenzwert $\lim_{n \rightarrow \infty} f_n(x)$ existiert. In diesem Fall nennen wir die Funktion

$$f : M \rightarrow M' : x \mapsto \lim_{n \rightarrow \infty} f_n(x)$$

den *Grenzwert* oder die *Grenzfunktion* der Funktionenfolge $(f_n)_{n \in \mathbb{N}}$, und wir sagen auch, daß $(f_n)_{n \in \mathbb{N}}$ *punktweise gegen f konvergiert*. Wir schreiben dann

$$f = \lim_{n \rightarrow \infty} f_n.$$

Man beachte, $(f_n)_{n \in \mathbb{N}}$ konvergiert auf M genau dann punktweise gegen f , wenn

$$\forall x \in M \forall \varepsilon > 0 \exists n_{\varepsilon, x} : \forall n \geq n_{\varepsilon, x} \text{ gilt } d'(f_n(x), f(x)) < \varepsilon.$$

- Wir sagen $(f_n)_{n \in \mathbb{N}}$ *konvergiert gleichmäßig auf M gegen f* , wenn

$$\forall \varepsilon > 0 \exists n_\varepsilon : \forall n \geq n_\varepsilon \text{ und } \forall x \in M \text{ gilt } d'(f_n(x), f(x)) < \varepsilon.$$

Bemerkung 40.37

Konvergiert $(f_n)_{n \in \mathbb{N}}$ auf M gleichmäßig gegen f , so konvergiert die Folge auch punktweise gegen f .

Satz 40.38 (Der gleichmäßige Grenzwert stetiger Funktionen ist stetig.)

Ist $f_n : M \rightarrow M'$ stetig auf M für $n \in \mathbb{N}$ und konvergiert $(f_n)_{n \in \mathbb{N}}$ auf M gleichmäßig gegen f , so ist f stetig auf M .

Beweis: Der Beweis ist identisch mit dem Beweis von Satz 15.6, wenn man die Betragsdifferenzen durch die Abstände ersetzt.

Seien $\mathbf{a} \in M$ und $\varepsilon > 0$ gegeben. Da die f_n gleichmäßig gegen f konvergieren, gilt:

$$\exists n_\varepsilon \in \mathbb{N} : \forall n \geq n_\varepsilon \text{ und } \forall \mathbf{x} \in M : d'(f_n(\mathbf{x}), f(\mathbf{x})) < \frac{\varepsilon}{3}.$$

Da zudem f_{n_ε} stetig in \mathbf{a} ist, gilt:

$$\exists \delta_\varepsilon > 0 \quad \forall \mathbf{x} \in M \text{ mit } d(\mathbf{x}, \mathbf{a}) < \delta_\varepsilon \text{ gilt } d'(f_{n_\varepsilon}(\mathbf{x}), f_{n_\varepsilon}(\mathbf{a})) < \frac{\varepsilon}{3}.$$

Sei nun $\mathbf{x} \in M$ mit $d(\mathbf{x}, \mathbf{a}) < \delta_\varepsilon$ gegeben, so gilt

$$d'(f(\mathbf{x}), f(\mathbf{a})) \leq d'(f(\mathbf{x}), f_{n_\varepsilon}(\mathbf{x})) + d'(f_{n_\varepsilon}(\mathbf{x}), f_{n_\varepsilon}(\mathbf{a})) + d'(f_{n_\varepsilon}(\mathbf{a}), f(\mathbf{a})) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Mithin ist f stetig in \mathbf{a} . □

Bemerkung 40.39 (Gleichmäßige Konvergenz ist Konvergenz bezüglich $\|\cdot\|_\infty$.)
Ist M kompakt, so ist die Menge der stetigen Funktionen

$$\mathcal{C}(M, \mathbb{R}) = \{f : M \longrightarrow \mathbb{R} \mid f \text{ ist stetig auf } M\}$$

ein normierter Raum vermittelt der Maximumsnorm

$$\|f\|_\infty = \max \{|f(\mathbf{x})| \mid \mathbf{x} \in M\},$$

wie man leicht nachprüft. Aus der Definition der gleichmäßigen Konvergenz folgt dann, daß eine Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ in $\mathcal{C}(M, \mathbb{R})$ genau dann in der Maximumsnorm gegen eine Funktion f konvergiert, wenn $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen f konvergiert.

Bemerkung 40.40 (Die Peano-Kurve)

Stetige Abbildungen haben viele sehr schöne topologische Eigenschaften wie die Verträglichkeit mit offenen und abgeschlossenen Mengen unter Urbildern in Proposition 40.17 oder mit kompakten Mengen in Satz 40.23. Aber wie sieht es mit der Verträglichkeit mit geometrischen Begriffen wie der Dimension aus. Denkt Ihr, daß es eine *surjektive stetige* Abbildung

$$f : [0, 1] \longrightarrow [0, 1] \times [0, 1]$$

geben kann?

Man kann sich das Bild von f als *Kurve* vorstellen, auf der man entlang läuft, indem man zum Zeitpunkt $t = 0$ bei $f(0)$ startet und schließlich zum Zeitpunkt $t = 1$ im Punkt $f(1)$ endet, und die Stetigkeit von f legt nahe, daß man zwischendurch nicht springen muß. Die Existenz einer solchen Abbildung würde also bedeuten, daß es eine stetige Kurve gibt, die jeden Punkt des zwei-dimensionalen Quaders durchläuft. Das widerspricht jeglichem Gefühl für Anstand, eine ein-dimensionale Kurve sollte kein zwei-dimensionales Gebiet ausfüllen können.

Aber, so viel zur Warnung, das geht! Ein Beispiel für eine solche Abbildung ist die sogenannte *Peano-Kurve*. Sie wird als gleichmäßiger Grenzwert einer Funktionenfolge definiert. Die Konstruktion ist sehr anschaulich, der Beweis der Konvergenz und der gewünschten Eigenschaften etwas technisch. Wer sich dafür interessiert, sei auf das Vorlesungsskript von Andreas Gathmann [**Gat08**, Satz 22.30] verwiesen.

H) Stetige lineare Operatoren

Satz 40.41 (Stetige lineare Operatoren)

Für eine lineare Abbildung $f : V \rightarrow W$ sind die folgenden Aussagen gleichwertig:

- f ist Lipschitz-stetig auf V .
- f ist gleichmäßig stetig auf V .
- f ist stetig auf V .
- f ist stetig in 0 .
- $\exists r > 0 : \forall x \in V : \|f(x)\|_W \leq r \cdot \|x\|_V$.

Man nennt f dann auch einen stetigen oder beschränkten linearen Operator.

Beweis:

a. \implies b.: Klar, siehe Proposition 40.33

b. \implies c.: Klar, siehe Bemerkung 40.28.

c. \implies d.: Klar.

d. \implies e.: Ist f stetig in 0 , so gibt es für $\varepsilon = 1$ ein $\delta_\varepsilon > 0$ mit

$$\|f(x)\|_W = \|f(x) - f(0)\|_W < 1 \quad (134)$$

für alle $x \in V$ mit $\|x\|_V = \|x - 0\|_V < \delta_\varepsilon$. Wir setzen $r = \frac{2}{\delta_\varepsilon} > 0$. Für $x \in V$ beliebig ist die Norm von $\frac{\delta_\varepsilon \cdot x}{2 \cdot \|x\|_V}$ kleiner als δ_ε und somit gilt

$$\begin{aligned} \|f(x)\|_W &= \left\| f \left(\frac{2 \cdot \|x\|_V}{\delta_\varepsilon} \cdot \frac{\delta_\varepsilon \cdot x}{2 \cdot \|x\|_V} \right) \right\|_W = \left\| \frac{2 \cdot \|x\|_V}{\delta_\varepsilon} \cdot f \left(\frac{\delta_\varepsilon \cdot x}{2 \cdot \|x\|_V} \right) \right\|_W \\ &= \frac{2 \cdot \|x\|_V}{\delta_\varepsilon} \cdot \left\| f \left(\frac{\delta_\varepsilon \cdot x}{2 \cdot \|x\|_V} \right) \right\|_W \stackrel{(134)}{<} \frac{2 \cdot \|x\|_V}{\delta_\varepsilon} \cdot 1 = r \cdot \|x\|_V. \end{aligned}$$

e. \implies a.: Für $x, y \in V$ gilt dann

$$\|f(x) - f(y)\|_W = \|f(x - y)\|_W \leq r \cdot \|x - y\|_V.$$

Mithin ist f Lipschitz-stetig auf V mit Lipschitz-Konstante r .

□

Beispiel 40.42 (Der Integraloperator als stetiger Operator)

Wir betrachten den Vektorraum $V = \mathcal{C}([0, 1], \mathbb{R})$ mit der euklidischen L_2 -Norm aus Beispiel 39.5 sowie die lineare Abbildung

$$I : V \rightarrow \mathbb{R} : f \mapsto \int_0^1 f(x) \, dx.$$

Für $f \in V$ gilt dann

$$|I(f)| = \left| \int_0^1 f(x) \, dx \right| \leq \int_0^1 |f(x)| \, dx = \int_0^1 (1 \cdot |f(x)|) \, dx \stackrel{37.8}{\leq} \|1\|_{L_2} \cdot \|f\|_{L_2},$$

wobei die letzte Ungleichung von der Ungleichung von Cauchy-Schwartz 37.8 für das Skalarprodukt auf V herrührt. Also ist der lineare Operator I beschränkt mit Schranke $r = \|1\|_{L_2} = 1$ und deshalb stetig auf V .

Satz 40.43

$L(V, W) = \{f : V \rightarrow W \mid f \text{ ist linear und stetig}\}$ ist ein normierter Raum vermittelt

$$\|\cdot\| : L(V, W) \rightarrow \mathbb{R}_{\geq 0} : f \mapsto \|f\| := \sup_{0 \neq x \in V} \frac{\|f(x)\|_W}{\|x\|_V}.$$

Wir nennen $\|f\|$ die Operatornorm von f .

Beweis: Nach Proposition 40.13 ist die Summe zweier stetiger Abbildungen sowie das skalare Vielfache einer stetigen Abbildung wieder stetig. Waren die Abbildungen zudem linear, sind auch die Summe oder das skalare Vielfache wieder linear. Mithin ist $L(V, W)$ ein Unterraum des Vektorraums aller Abbildungen von V nach W .

Es bleibt also nur zu zeigen, daß die Operatornorm eine Norm ist.

Gilt $\|f\| = 0$, so muß notwendigerweise $\|f(x)\|_W = 0$ für alle $x \in V$ gelten und mithin $f(x) = 0$ für alle $x \in V$. Dann ist f aber die Nullabbildung. Dies zeigt, daß $\|\cdot\|$ definit ist.

Ist $f \in L(V, W)$ und $\lambda \in \mathbb{R}$, so gilt

$$\|\lambda \cdot f\| = \sup_{0 \neq x \in V} \frac{\|\lambda \cdot f(x)\|_W}{\|x\|_V} = \sup_{0 \neq x \in V} \frac{|\lambda| \cdot \|f(x)\|_W}{\|x\|_V} = |\lambda| \cdot \sup_{0 \neq x \in V} \frac{\|f(x)\|_W}{\|x\|_V} = |\lambda| \cdot \|f\|,$$

so daß $\|\cdot\|$ auch homogen ist.

Sind $f, g \in L(V, W)$, so gilt zudem

$$\begin{aligned} \|f + g\| &= \sup_{0 \neq x \in V} \frac{\|f(x) + g(x)\|_W}{\|x\|_V} \leq \sup_{0 \neq x \in V} \left(\frac{\|f(x)\|_W}{\|x\|_V} + \frac{\|g(x)\|_W}{\|x\|_V} \right) \\ &\leq \sup_{0 \neq x \in V} \frac{\|f(x)\|_W}{\|x\|_V} + \sup_{0 \neq x \in V} \frac{\|g(x)\|_W}{\|x\|_V} = \|f\| + \|g\|. \end{aligned}$$

Die Dreiecksungleichung gilt also auch, und $\|\cdot\|$ ist eine Norm. □

Die Aussage des folgenden Korollars folgt unmittelbar aus der Definition der Operatornorm.

Korollar 40.44 (Submultiplikativität der Operatornorm)

Für $f \in L(V, W)$ und $x \in V$ gilt $\|f(x)\|_W \leq \|f\| \cdot \|x\|_V$.

Beispiel 40.45 (Operatornorm des Integraloperators)

In Beispiel 40.42 haben wir gesehen, daß die Operatornorm von I höchstens 1 sein kann, da

$$\frac{|I(f)|}{\|f\|_{L_2}} \leq \|1\|_{L_2} = 1$$

gilt und mithin auch

$$\|I\| = \sup_{0 \neq f \in V} \frac{|I(f)|}{\|f\|_{L_2}} \leq 1$$

gelten muß. Da außerdem $|I(1)| = \int_0^1 1 \, dx = 1$ gilt, folgt auch

$$\|I\| = \sup_{0 \neq f \in V} \frac{|I(f)|}{\|f\|_{L_2}} \geq \frac{|I(1)|}{\|1\|_{L_2}} = 1.$$

Der Integraloperator I hat in diesem Fall also die Operatornorm $\|I\| = 1$.

Beispiel 40.46

Da alle linearen Abbildungen von \mathbb{R}^n nach \mathbb{R}^m von der Form f_A für eine Matrix $A \in \text{Mat}(m \times n, \mathbb{R})$ sind und da diese alle stetig sind, wissen wir

$$L(\mathbb{R}^n, \mathbb{R}^m) = \text{Hom}_{\mathbb{R}}(\mathbb{R}^n, \mathbb{R}^m).$$

Bemerkung 40.47 (Operatornorm auf $\text{Mat}(m \times n, \mathbb{R})$)

Definieren wir die Norm $\|A\| := \|f_A\|$ einer Matrix $A \in \text{Mat}(m \times n, \mathbb{R})$ als Operatornorm der zugehörigen linearen Abbildung f_A bezüglich fest vorgegebener Normen auf \mathbb{R}^n und \mathbb{R}^m , so erhalten wir in der Tat eine Norm auf $\text{Mat}(m \times n, \mathbb{R})$ und diese ist submultiplikativ im Sinne von Korollar 40.44.

Aufgaben**Aufgabe 40.48** (Topologische Isomorphismen)

Zeige, für eine bijektive lineare Abbildung $f: V \rightarrow W$ sind gleichwertig:

- f und f^{-1} sind stetig.
- $\exists r, s > 0 : \forall x \in V : s \cdot \|x\|_V \leq \|f(x)\|_W \leq r \cdot \|x\|_V$.

Man nennt den Isomorphismus f dann einen *topologischen Isomorphismus*

Aufgabe 40.49

Zeige, ist $\dim_{\mathbb{R}}(V) = n$ und betrachten wir \mathbb{R}^n mit der euklidischen Norm, so ist jeder Isomorphismus $f: \mathbb{R}^n \rightarrow V$ ein topologischer Isomorphismus.

Aufgabe 40.50 (Endlich-dimensionale Banachräume)

Zeige, jeder endlich-dimensionale normierte \mathbb{R} -Vektorraum ist ein Banachraum.

Aufgabe 40.51

Zeige, auf einem endlich-dimensionalen \mathbb{R} -Vektorraum sind alle Normen äquivalent.

Aufgabe 40.52

Sind V und W endlich-dimensional, so gilt $L(V, W) = \text{Hom}_{\mathbb{R}}(V, W)$.

Aufgabe 40.53 (Der Differentialoperator als nicht stetiger Operator)

Es sei $V = C^1([0, 1], \mathbb{R})$ der Vektorraum der stetig-differenzierbaren Funktionen auf $[0, 1]$ und $W = C([0, 1], \mathbb{R})$. Aus Aufgabe 39.53 wissen wir, daß W ein normierter Raum bezüglich der Maximumsnorm aus Beispiel 39.5 ist, und somit ist auch V als Unterraum von W normiert bezüglich der Maximumsnorm. Zeige, der Differentialoperator

$$D : V \longrightarrow W : f \mapsto f'$$

ist ein linearer Operator, der nicht stetig ist.

Aufgabe 40.54

Betrachte $V = \text{Mat}_n(\mathbb{R}) \cong \mathbb{R}^{n^2}$ als normierten Raum mit der euklidischen Norm

$$\|(\mathbf{a}_{ij})\|_2 = \sqrt{\sum_{i=1}^n \sum_{j=1}^n \mathbf{a}_{ij}^2},$$

(siehe auch Aufgabe 39.55) und betrachte $U = \{A \in V \mid A \text{ ist symmetrisch}\}$ als metrischen Raum mit der Einschränkung der Metrik von V auf U .

- Zeige, die Menge $P = \{A \in U \mid A \text{ ist positiv definit}\}$ ist offen in U .
- Zeige, die Menge $N = \{A \in U \mid A \text{ ist negativ definit}\}$ ist offen in U .
- Zeige, die Menge $I = \{A \in U \mid A \text{ ist indefinit}\}$ ist offen in U .

Aufgabe 40.55

Es sei $\mathbf{a} \in M$ und $f : M \longrightarrow \mathbb{R}$ sei stetig in \mathbf{a} mit $f(\mathbf{a}) \neq 0$. Zeige, es gibt ein $\delta > 0$ mit $f(\mathbf{x}) \neq 0$ für alle $\mathbf{x} \in U_\delta(\mathbf{a})$.

Aufgabe 40.56

Sei $X := [0, \infty) \times [0, \infty)$ und $f : X \longrightarrow \mathbb{R}$ die Potenzfunktion

$$f(x, y) = \begin{cases} 0 & \text{für } x > 0 \text{ und } y = 0 \\ 1 & \text{für } x = y = 0 \\ y^x & \text{sonst} \end{cases}.$$

Untersuche, an welchen Stellen f stetig ist.

Aufgabe 40.57

Zeige, daß die Menge $U = \{(x, y)^t \in \mathbb{R}^2 \mid x < 2y^2 - 1\}$ offen in \mathbb{R}^2 ist.

Aufgabe 40.58

Zeige, daß die Funktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}, (x, y)^t \longmapsto \begin{cases} \frac{xy^2}{x^2+y^4} & \text{falls } (x, y) \neq (0, 0) \\ 0 & \text{falls } (x, y) = (0, 0) \end{cases}$$

im Nullpunkt das Folgenkriterium für Stetigkeit für jede Folge erfüllt, die sich auf einer Geraden dem Nullpunkt nähert. Ist f in $(0, 0)$ stetig?

§ 41 Differenzierbarkeit im \mathbb{R}^n

In diesem Abschnitt betrachten wir \mathbb{R}^n und \mathbb{R}^m stets als normierte Räume bezüglich der euklidischen Norm $\|\cdot\|_2$. Außerdem sei $U \subseteq \mathbb{R}^n$ stets eine *offene* Teilmenge des \mathbb{R}^n , so daß jeder Punkt in U ein Häufungspunkt ist.

Motivation 41.1

Eine Funktion $f : U \rightarrow \mathbb{R}$ mit $U \subseteq \mathbb{R}^n$ haben wir *differenzierbar* im Häufungspunkt a von U genannt, wenn der Grenzwert des Differenzenquotienten

$$\lim_{x \rightarrow a} \text{Diff}_{f,a}(x) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = \lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \in \mathbb{R}$$

in a existiert (siehe Definition 17.3). Diesen Grenzwert haben wir dann die *Ableitung* von f im Punkt a genannt.

Eine dazu äquivalente Bedingung haben wir in Bemerkung 17.6 gesehen. f ist genau dann *differenzierbar* in a , wenn es eine Zahl $c \in \mathbb{R}$ und eine Funktion $\rho : U \rightarrow \mathbb{R}$ gibt, so daß

$$f(x) = f(a) + c \cdot (x - a) + \rho(x)$$

und

$$\lim_{x \rightarrow a} \frac{|\rho(x)|}{|x - a|} = \lim_{h \rightarrow 0} \frac{|\rho(a+h)|}{|h|} = 0.$$

Hierbei war dann $c = f'(a)$ die *Ableitung* von f in a .

Wieso beide Beschreibungen äquivalent waren, interessiert uns im Augenblick nicht so sehr. Vielmehr freuen wir uns, daß wir zwei Ansätze haben, die wir eventuell nutzen können, um den Begriff der Differenzierbarkeit vom Eindimensionalen ins Mehrdimensionale zu verallgemeinern.

In beiden Ansätzen sehen wir, daß die Differenz oder Summe von Punkten im Start- und Zielraum gebildet werden muß, und daß in unseren Definitionen dabei nicht der Abstand der Punkte ausreicht. Dies legt nahe, daß wir für den Begriff der Differenzierbarkeit keine allgemeinen metrischen Räume mehr betrachten sollten, sondern nur noch solche, wo diese Operationen erlaubt sind, d.h. metrische Räume, die Teilmengen von normierten Räumen sind. In der Tat wollen wir uns in diesem Abschnitt dabei im wesentlichen auf Teilmengen von \mathbb{R}^n und \mathbb{R}^m beschränken.

Betrachten wir nun die erste mögliche Definition oben, so sehen wir, daß wir durch $x - a$ bzw. durch h dividieren müssen. Ist der Definitionsbereich U aber eine Teilmenge von \mathbb{R}^n mit $n \geq 1$, so sind $x - a$ und h Vektoren. Diese Operation ist dann nicht erlaubt, und damit ist die Definition schlecht für eine Verallgemeinerung der Differenzierbarkeit geeignet. Die zweite Definition kommt hingegen ohne Division durch Vektoren aus. Dort wird lediglich durch die Länge von Vektoren dividiert, was erlaubt ist. Insofern bietet sie einen kanonischen Ansatz zur Verallgemeinerung.

Allerdings birgt sie ein anderes Problem! Bei der ersten Definition ergab sich die Ableitung einfach als der Grenzwert, dessen Existenz die Differenzierbarkeit definierte. In der zweiten Definition mußte man die Ableitung c a priori kennen, um die Differenzierbarkeit prüfen zu können. Wenn der Grenzwert, den wir im Eindimensionalen zur Berechnung der Ableitung benutzt haben, im Mehrdimensionalen aber keinen Sinn mehr ergibt, dann stellt sich die Frage, wie man die Ableitung denn a priori ausrechnen soll, und schlimmer noch, was sie eigentlich ist? Eine einzige Zahl wird wohl kaum noch ausreichen, um die *Steigung* im Mehrdimensionalen zu beschreiben, wo man sich in so viele verschiedene Richtungen bewegen kann.

Die folgende Überlegung soll dabei weiterhelfen. Im Eindimensionalen war die Ableitung von f in \mathbf{a} , wie schon erwähnt, eine Zahl $c = f'(\mathbf{a})$. Nun kann man eine Zahl c auch als Abbildung auffassen, indem wir einfach mit dieser Zahl multiplizieren:

$$\mathbb{R} \longrightarrow \mathbb{R} : x \mapsto c \cdot x.$$

Dadurch wird aus c eine lineare Abbildung. Dies ist ein Konzept, das sich verallgemeinern läßt. Wir werden die Ableitung einer Abbildung $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ an einer Stelle \mathbf{a} als lineare Abbildung $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ ansehen oder auch einfach als Matrix in $\text{Mat}(m \times n, \mathbb{R})$, da jede lineare Abbildung von \mathbb{R}^n nach \mathbb{R}^m ja als Multiplikation mit einer solchen gegeben ist.

Es bleibt das Problem, wie man die Ableitung a priori berechnen soll, aber dabei werden uns sogenannte partielle Ableitungen helfen, das sind Ableitungen in Richtung der Koordinatenachsen, bei denen wir dann auch unsere alte Formel wieder teilweise nutzen können.

A) Totale, partielle und stetige Differenzierbarkeit

Definition 41.2 (Total differenzierbar)

Eine Abbildung $f : \mathbf{U} \longrightarrow \mathbb{R}^m$ heißt *total differenzierbar* in $\mathbf{a} \in \mathbf{U}$, wenn es eine Matrix $\mathbf{A} \in \text{Mat}(m \times n, \mathbb{R})$ und eine Funktion $\rho : \mathbf{U} \longrightarrow \mathbb{R}^m$ gibt mit

$$f(\mathbf{x}) = f(\mathbf{a}) + \mathbf{A} \cdot (\mathbf{x} - \mathbf{a}) + \rho(\mathbf{x})$$

und

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} = \mathbf{0}.$$

Wir nennen die Matrix $Df(\mathbf{a}) := \mathbf{A}$ dann die *Ableitung* von f in \mathbf{a} .

f heißt *total differenzierbar* auf \mathbf{U} , wenn f für jedes $\mathbf{a} \in \mathbf{U}$ total differenzierbar ist.

Bemerkung 41.3 (Totale Differenzierbarkeit in \mathbb{R})

Aufgrund von Bemerkung 17.6 ist eine Funktion $f : \mathbf{U} \longrightarrow \mathbb{R}$ mit $\mathbf{U} \subseteq \mathbb{R}$ offen genau dann in $\mathbf{a} \in \mathbf{U}$ differenzierbar im Sinne von Definition 17.3, wenn f in \mathbf{a} total differenzierbar ist.

Lemma 41.4 (Eindeutigkeit der Ableitung)

Ist $f: \mathbf{U} \rightarrow \mathbb{R}^m$ total differenzierbar in $\mathbf{a} \in \mathbf{U}$ und $0 \neq \mathbf{x} \in \mathbb{R}^n$, so gilt²

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{x}) - f(\mathbf{a})}{t} = Df(\mathbf{a}) \circ \mathbf{x}.$$

Inbesondere ist die Ableitung $Df(\mathbf{a})$ von f in \mathbf{a} eindeutig bestimmt.

Beweis: Es seien A und ρ wie in Definition 41.2. Da \mathbf{a} ein innerer Punkt der offenen Menge \mathbf{U} ist, gibt es ein $\varepsilon > 0$, so daß $\mathbf{a} + t\mathbf{x} \in \mathbf{U}$ für alle $t \in \mathbf{U}_\varepsilon(0)$. Wir werden im folgenden nur solche t betrachten. Wir erhalten dann

$$\frac{f(\mathbf{a} + t\mathbf{x}) - f(\mathbf{a})}{t} = \frac{A t\mathbf{x} + \rho(\mathbf{a} + t\mathbf{x})}{t} = A\mathbf{x} \pm \frac{\rho(\mathbf{a} + t\mathbf{x})}{\|\mathbf{a} + t\mathbf{x} - \mathbf{a}\|_2} \cdot \|\mathbf{x}\|_2 \xrightarrow{t \rightarrow 0} A\mathbf{x}.$$

Das zeigt die erste Aussage des Lemmas. Wenden wir die Formel auf $\mathbf{x} = \mathbf{e}_i$ an, erhalten wir, daß die i -te Spalte $A\mathbf{e}_i$ von A gerade

$$A\mathbf{e}_i = \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{e}_i) - f(\mathbf{a})}{t}$$

ist. Die rechte Seite hängt aber nur von f und nicht von A ab. Dies zeigt, daß A durch f eindeutig bestimmt ist. \square

Beispiel 41.5 (Totale Differenzierbarkeit)

Wir wollen zeigen, daß die Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}: (x_1, x_2)^t \mapsto 3x_1 + x_1x_2 - 2x_2 + 1$ in $\mathbf{a} = (0, 0)^t$ total differenzierbar ist. Aus Lemma 41.4 wissen wir, daß die Ableitung $Df(\mathbf{a}) = A$ dann eine 1×2 -Matrix sein muß, deren erste Spalte

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{e}_1) - f(\mathbf{a})}{t} = \lim_{t \rightarrow 0} \frac{3t}{t} = 3$$

und deren zweite Spalte

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{e}_2) - f(\mathbf{a})}{t} = \lim_{t \rightarrow 0} \frac{-2t}{t} = -2$$

ist, d.h.

$$A = (3 \quad -2)$$

Die Abbildung ρ muß dann die Abbildungsvorschrift

$$\rho: \mathbb{R}^2 \rightarrow \mathbb{R}: (x_1, x_2)^t \mapsto f(x_1, x_2) - f(\mathbf{a}) - A \circ (x_1, x_2)^t = x_1x_2$$

haben. Wir erhalten deshalb

$$\lim_{(x_1, x_2)^t \rightarrow (0, 0)^t} \frac{\rho(x_1, x_2)}{\|(x_1, x_2)^t - \mathbf{a}\|_2} = \lim_{(x_1, x_2)^t \rightarrow (0, 0)^t} \frac{x_1x_2}{\sqrt{x_1^2 + x_2^2}} = 0,$$

wie wir aus Beispiel 40.5 wissen. Also ist f im Ursprung total differenzierbar.

Die Beschreibung der Ableitung von f im Punkt \mathbf{a} in Lemma 41.4 motiviert die folgende Definition.

²In der Definition dieses Grenzwertes und aller folgenden betrachten wir nur solche t , für die $\mathbf{a} + t\mathbf{x}$ in \mathbf{U} liegt. Da \mathbf{U} offen ist, gilt dies für alle t , die hinreichend nahe an 0 liegen, d.h. es gibt ein $\varepsilon > 0$, so daß für alle $t \in \mathbf{U}_\varepsilon(0)$ auch $\mathbf{a} + t\mathbf{x} \in \mathbf{U}$ gilt.

Definition 41.6 (Partielle Ableitungen)

- a. Für eine Funktion $f : \mathcal{U} \rightarrow \mathbb{R}$, einen Punkt $\mathbf{a} \in \mathcal{U}$ und $\mathbf{v} \in \mathbb{R}^n$ mit $\|\mathbf{v}\|_2 = 1$ heißt der Grenzwert

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a}) := \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a})}{t}$$

die *Richtungsableitung* von f in \mathbf{a} in Richtung \mathbf{v} , sofern er existiert.

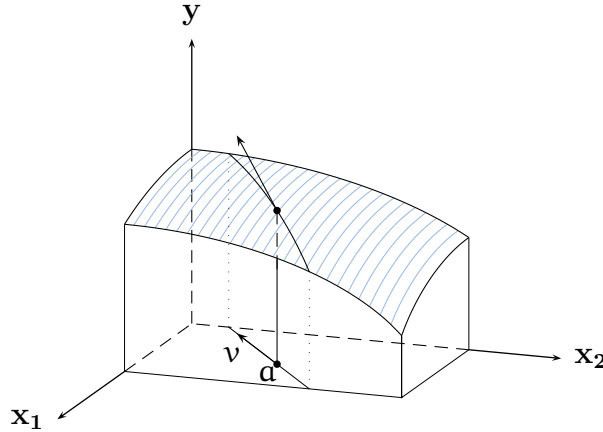


ABBILDUNG 4. Richtungsableitung $\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a})$ in Richtung \mathbf{v}

Die Richtungsableitung in Richtung der i -ten Koordinate $\mathbf{v} = \mathbf{e}_i$

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{e}_i) - f(\mathbf{a})}{t}$$

heißt auch die *partielle Ableitung* von f nach der i -ten Variablen in \mathbf{a} . Folgende Bezeichnungen sind für diese partielle Ableitung gebräuchlich

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) = \frac{\partial f}{\partial \mathbf{e}_i}(\mathbf{a}) = f_{x_i}(\mathbf{a}) = \partial_i f(\mathbf{a}) = D_i f(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{e}_i) - f(\mathbf{a})}{t}.$$

- b. Eine Abbildung $f : \mathcal{U} \rightarrow \mathbb{R}^m$ heißt *partiell differenzierbar* in $\mathbf{a} \in \mathcal{U}$, falls alle partiellen Ableitungen aller Komponentenfunktionen f_1, \dots, f_m von f in \mathbf{a} existieren. Wir nennen die Matrix

$$Jf(\mathbf{a}) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \frac{\partial f_1}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{a}) & \frac{\partial f_2}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_2}{\partial x_n}(\mathbf{a}) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \frac{\partial f_m}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{pmatrix}$$

dann die *Jacobi-Matrix* von f in \mathbf{a} .

f heißt *partiell differenzierbar* auf \mathcal{U} , wenn f in jedem $\mathbf{a} \in \mathcal{U}$ partiell differenzierbar ist. In diesem Fall sind die partiellen Ableitungen

$$\frac{\partial f_i}{\partial x_j} : \mathcal{U} \rightarrow \mathbb{R}$$

selbst wieder Funktionen auf \mathcal{U} .

c. Eine Abbildung $f : \mathbf{U} \rightarrow \mathbb{R}^m$ heißt *stetig differenzierbar* auf \mathbf{U} , wenn f partiell differenzierbar auf \mathbf{U} ist und alle partiellen Ableitungen stetig auf \mathbf{U} sind.

Wir bezeichnen mit $\mathcal{C}^1(\mathbf{U}, \mathbb{R}^m)$ die Menge der stetig differenzierbaren Funktionen von \mathbf{U} nach \mathbb{R}^m .

Bemerkung 41.7 (Richtungsableitung)

Ist $f : \mathbf{U} \rightarrow \mathbb{R}$, $\mathbf{U}_\varepsilon(\mathbf{a}) \subseteq \mathbf{U}$ und $\mathbf{v} \in \mathbb{R}^n$ mit $\|\mathbf{v}\|_2 = 1$, so betrachten wir die Funktion

$$g : (-\varepsilon^2, \varepsilon^2) \rightarrow \mathbb{R} : t \mapsto f(\mathbf{a} + t\mathbf{v}).$$

Die Richtungsableitung von f in \mathbf{a} in Richtung \mathbf{v} existiert genau dann, wenn g in 0 differenzierbar ist, und in diesem Fall gilt

$$g'(0) = \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a})}{t} = \frac{\partial f}{\partial \mathbf{v}}(\mathbf{a}).$$

Diese Aussage läßt sich insbesondere auf die partiellen Ableitungen, d.h. den Fall $\mathbf{v} = \mathbf{e}_i$, anwenden. Man führt die partielle Differenzierbarkeit also auf die Differenzierbarkeit im Eindimensionalen zurück.

Korollar 41.8 (Total differenzierbar impliziert partiell differenzierbar.)

Ist $f : \mathbf{U} \rightarrow \mathbb{R}^m$ total differenzierbar in $\mathbf{a} \in \mathbf{U}$, so ist f partiell differenzierbar in \mathbf{a} und die Ableitung von f in \mathbf{a} ist die Jacobi-Matrix von f in \mathbf{a}

$$Df(\mathbf{a}) = Jf(\mathbf{a}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \frac{\partial f_1}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_1}{\partial x_n}(\mathbf{a}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{a}) & \frac{\partial f_2}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_2}{\partial x_n}(\mathbf{a}) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{a}) & \frac{\partial f_m}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_m}{\partial x_n}(\mathbf{a}) \end{pmatrix}.$$

Beweis: In Lemma 41.4 haben wir gezeigt, daß die Grenzwerte

$$\lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{e}_j) - f(\mathbf{a})}{t}$$

existieren, und Satz 40.19 existieren dann die Grenzwerte der Komponentenfunktionen

$$\frac{\partial f_i}{\partial x_j}(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{f_i(\mathbf{a} + t\mathbf{e}_j) - f_i(\mathbf{a})}{t}.$$

f ist also partiell differenzierbar in \mathbf{a} , und zudem folgt aus Lemma 41.4, daß die j -te Spalte von $Df(\mathbf{a})$ genau die j -te Spalte von $Jf(\mathbf{a})$ ist. \square

Beispiel 41.9 (Partiell differenzierbar impliziert nicht total differenzierbar.)

Eine partiell differenzierbare Abbildung muß *nicht* total differenzierbar sein. Selbst die Existenz aller Richtungsableitungen reicht nicht aus, um die totale Differenzierbarkeit zu garantieren.

Um dies zu sehen, betrachten wir die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} : (x_1, x_2)^t \mapsto \begin{cases} \frac{x_1 \cdot x_2^3}{x_1^2 + x_2^6}, & \text{falls } (x_1, x_2)^t \neq (0, 0), \\ 0 & \text{sonst.} \end{cases}$$

Wir wollen zeigen, daß für f in $\mathbf{a} = (0, 0)^t$ sogar alle Richtungsableitungen existieren, aber f trotzdem in \mathbf{a} nicht total differenzierbar ist.

Sei dazu $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^2$ mit $\|\mathbf{v}\|_2 = 1$. Ist $v_1 = 0$, so ist

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a})}{t} = \lim_{t \rightarrow 0} \frac{0}{t^7} = 0.$$

Ist $v_1 \neq 0$, so ist

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a})}{t} = \lim_{t \rightarrow 0} \frac{t^4 v_1 v_2^3}{t^3 v_1^2 + t^7 v_2^6} = \lim_{t \rightarrow 0} \frac{t v_1 v_2^3}{v_1^2 + t^4 v_2^4} = 0.$$

Mithin existiert die Richtungsableitung in Richtung \mathbf{v} im Punkt \mathbf{a} und ist 0.

f ist damit in \mathbf{a} insbesondere partiell differenzierbar mit Jacobi-Matrix $Jf(\mathbf{a}) = (0 \ 0)$.

Für die Funktion ρ ergibt sich damit

$$\rho(\mathbf{x}_1, \mathbf{x}_2) = f(\mathbf{x}_1, \mathbf{x}_2) - f(0, 0) - Jf(\mathbf{a}) \cdot (\mathbf{x}_1, \mathbf{x}_2)^t = f(\mathbf{x}_1, \mathbf{x}_2),$$

und wir wollen nun zeigen, daß der Grenzwert

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} \tag{135}$$

nicht existiert. Dazu betrachten wir die Folge $(\mathbf{a}_n)_{n \geq 1}$ mit

$$\mathbf{a}_n = \left(\frac{1}{n^3}, \frac{1}{n}\right)^t \longrightarrow (0, 0)^t.$$

Dann gilt

$$\frac{\rho(\mathbf{a}_n)}{\|\mathbf{a}_n - \mathbf{a}\|_2} = \frac{\frac{1}{n^6}}{\frac{2}{n^6} \cdot \sqrt{\frac{1}{n^6} + \frac{1}{n^2}}} = \frac{n}{2 \cdot \sqrt{\frac{1}{n^4} + 1}} \longrightarrow \infty.$$

Aus dem Folgenkriterium für Grenzwerte erhalten wir deshalb, daß der Grenzwert in (135) nicht existiert. Mithin ist die Funktion f nicht total differenzierbar im Ursprung.

Man beachte auch, daß die partiellen Ableitungen von f im Punkt $\mathbf{a} = (0, 0)^t$ nicht stetig sein können, wie aus dem folgenden Satz folgt.

Satz 41.10 (Stetig differenzierbar impliziert total differenzierbar.)

Ist $f : \mathbf{U} \longrightarrow \mathbb{R}^m$ stetig differenzierbar auf \mathbf{U} , so ist f total differenzierbar auf \mathbf{U} .

Beweis: Es seien wieder f_1, \dots, f_m die Komponentenfunktionen von f und $\mathbf{a} \in \mathbf{U}$. Wir wählen ein $\varepsilon > 0$, so daß die ε -Umgebung von \mathbf{a} bezüglich der Maximumsnorm ganz in \mathbf{U} liegt, und wir betrachten nur noch Punkte \mathbf{x} in dieser ε -Umgebung. Außerdem halten wir ein $i \in \{1, \dots, m\}$ fest. Dann gilt

$$f_i(\mathbf{x}) - f_i(\mathbf{a}) = \sum_{j=1}^n f_i(x_1, \dots, x_j, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n) - f_i(x_1, \dots, x_{j-1}, \mathbf{a}_j, \dots, \mathbf{a}_n).$$

Da die Funktion

$$\mathbf{U}_\varepsilon(\mathbf{a}_i) = (\mathbf{a}_i - \varepsilon, \mathbf{a}_i + \varepsilon) \longrightarrow \mathbb{R} : z \mapsto f_i(x_1, \dots, x_{j-1}, z, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n)$$

nach Voraussetzung differenzierbar in \mathbf{a}_j ist, kann man den Mittelwertsatz der Differentialrechnung 18.7 im Eindimensionalen anwenden und erhält

$$\begin{aligned} \sum_{j=1}^n f_i(x_1, \dots, x_j, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n) - f_i(x_1, \dots, x_{j-1}, \mathbf{a}_j, \dots, \mathbf{a}_n) \\ \stackrel{\text{MWS 18.7}}{=} \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(x_1, \dots, x_{j-1}, \mathbf{c}_j, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n) \cdot (x_j - \mathbf{a}_j). \end{aligned}$$

für $\mathbf{c}_1, \dots, \mathbf{c}_n \in \mathbb{R}$ mit $|\mathbf{c}_j - \mathbf{a}_j| < |x_j - \mathbf{a}_j| < \varepsilon$ geeignet. Wir setzen nun

$$\begin{aligned} \rho_i(\mathbf{x}) &:= f_i(\mathbf{x}) - f_i(\mathbf{a}) - \mathbf{J}f_i(\mathbf{a}) \circ (\mathbf{x} - \mathbf{a}) \\ &= \sum_{j=1}^n \left(\frac{\partial f_i}{\partial x_j}(x_1, \dots, x_{j-1}, \mathbf{c}_j, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n) - \frac{\partial f_i}{\partial x_j}(\mathbf{a}) \right) \cdot (x_j - \mathbf{a}_j) \\ &= \langle \mathbf{r}(\mathbf{x}), \mathbf{x} - \mathbf{a} \rangle, \end{aligned}$$

wobei $\mathbf{r}(\mathbf{x})$ ein Vektor ist, dessen j -te Komponente gerade

$$r_j(\mathbf{x}) = \frac{\partial f_i}{\partial x_j}(x_1, \dots, x_{j-1}, \mathbf{c}_j, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n) - \frac{\partial f_i}{\partial x_j}(\mathbf{a})$$

ist. Die Schreibweise mit Hilfe des Skalarproduktes hat den Vorteil, daß wir den Satz von Cauchy-Schwarz 37.8 anwenden können. Wir erhalten dann

$$\frac{|\rho_i(\mathbf{x})|}{\|\mathbf{x} - \mathbf{a}\|_2} = \frac{|\langle \mathbf{r}(\mathbf{x}), \mathbf{x} - \mathbf{a} \rangle|}{\|\mathbf{x} - \mathbf{a}\|_2} \stackrel{\text{CS 37.8}}{\leq} \frac{\|\mathbf{r}(\mathbf{x})\|_2 \cdot \|\mathbf{x} - \mathbf{a}\|_2}{\|\mathbf{x} - \mathbf{a}\|_2} = \|\mathbf{r}(\mathbf{x})\|_2$$

Die Stetigkeit der partiellen Ableitungen in \mathbf{a} impliziert nun aber, daß

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} r_j(\mathbf{x}) = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\partial f_i}{\partial x_j}(x_1, \dots, x_{j-1}, \mathbf{c}_j, \mathbf{a}_{j+1}, \dots, \mathbf{a}_n) - \frac{\partial f_i}{\partial x_j}(\mathbf{a}) = 0$$

für alle $j = 1, \dots, n$, und mithin folgen $\lim_{\mathbf{x} \rightarrow \mathbf{a}} \|\mathbf{r}(\mathbf{x})\|_2 = 0$ und deshalb

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_i(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} = 0.$$

Insgesamt erhalten wir für

$$\rho(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{a}) - \mathbf{J}f(\mathbf{a}) \circ (\mathbf{x} - \mathbf{a}) = (\rho_1(\mathbf{x}), \dots, \rho_m(\mathbf{x}))^t$$

dann

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} = 0.$$

Also ist f in \mathbf{a} total differenzierbar. □

Beispiel 41.11 (Total differenzierbar impliziert nicht stetig differenzierbar.)

Die Funktion

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}: \mathbf{x} \mapsto \begin{cases} \|\mathbf{x}\|_2^2 \cdot \sin\left(\frac{1}{\|\mathbf{x}\|_2}\right), & \text{falls } \mathbf{x} \neq \mathbf{0}, \\ 0 & \text{sonst} \end{cases}$$

ist total differenzierbar auf \mathbb{R}^2 , aber nicht stetig differenzierbar (siehe Aufgabe 41.29).

Bemerkung 41.12 (Wie prüft man totale Differenzierbarkeit nach?)

a. Satz 41.10 ist für konkrete Anwendungen ungemein wichtig. Wollen wir überprüfen, ob eine Abbildung $f : \mathbf{U} \rightarrow \mathbb{R}^m$ total differenzierbar ist, so können wir also wie folgt vorgehen:

- 1) Überprüfe, ob alle partiellen Ableitungen auf \mathbf{U} existieren. Falls nicht, so ist f nicht total differenzierbar auf \mathbf{U} .
- 2) Überprüfe, ob alle partiellen Ableitungen stetig auf \mathbf{U} sind. Falls ja, so ist f total differenzierbar auf \mathbf{U} .

Nur, wenn die partiellen Ableitungen existieren, aber nicht stetig sind, müssen wir auf die Definition zurück greifen und den Grenzwert

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f(\mathbf{x}) - f(\mathbf{a}) - \mathbf{J}f(\mathbf{a}) \circ (\mathbf{x} - \mathbf{a})}{\|\mathbf{x} - \mathbf{a}\|_2}$$

für $\mathbf{a} \in \mathbf{U}$ untersuchen. Ist er null, so ist f in \mathbf{a} total differenzierbar.

b. Wenn man den Beweis von Satz 41.10 genau betrachtet, sieht man, daß es reicht, zu fordern, daß f auf \mathbf{U} partiell differenzierbar ist und daß die partiellen Ableitungen in \mathbf{a} stetig sind, um die totale Differenzierbarkeit von f in \mathbf{a} zu folgern. Die Stetigkeit der partiellen Ableitungen außerhalb von \mathbf{a} ist nicht erforderlich, in der Praxis aber in aller Regel gegeben.

Proposition 41.13 (Total differenzierbar impliziert stetig.)

Ist $f : \mathbf{U} \rightarrow \mathbb{R}^m$ total differenzierbar in $\mathbf{a} \in \mathbf{U}$, so ist f stetig in \mathbf{a} .

Beweis: Wir setzen $\mathbf{A} = \mathbf{D}f(\mathbf{a})$ und erhalten dann

$$\begin{aligned} \lim_{\mathbf{x} \rightarrow \mathbf{a}} f(\mathbf{x}) &= \lim_{\mathbf{x} \rightarrow \mathbf{a}} (f(\mathbf{a}) + \mathbf{A} \circ (\mathbf{x} - \mathbf{a}) + \rho(\mathbf{x})) \\ &= f(\mathbf{a}) + \lim_{\mathbf{x} \rightarrow \mathbf{a}} \mathbf{A} \circ (\mathbf{x} - \mathbf{a}) + \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} \cdot \lim_{\mathbf{x} \rightarrow \mathbf{a}} \|\mathbf{x} - \mathbf{a}\|_2 = f(\mathbf{a}), \end{aligned}$$

da die lineare Abbildung, die durch die Multiplikation mit \mathbf{A} definiert wird, stetig ist. Also ist f stetig im Häufungspunkt \mathbf{a} von \mathbf{U} . \square

Beispiel 41.14 (Stetig impliziert nicht differenzierbar.)

Die Abbildung

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} : \mathbf{x} = (x_1, x_2)^t \mapsto \begin{cases} \frac{x_1 x_2}{\sqrt{x_1^2 + x_2^2}}, & \text{falls } (x_1, x_2)^t \neq (0, 0)^t, \\ 0 & \text{sonst} \end{cases}$$

ist wegen Beispiel 40.5 b. stetig in $\mathbf{a} = (0, 0)^t$. Wir wollen nun zeigen, daß sie nicht total differenzierbar in \mathbf{a} ist, wohl aber partiell differenzierbar. Die partiellen Ableitungen von f sind

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{e}_i) - f(\mathbf{a})}{t} = \lim_{t \rightarrow 0} \frac{0}{t^2} = 0$$

und die Jacobi-Matrix von f im Punkt $\mathbf{a} = (0, 0)^t$ ist mithin

$$\mathbf{J}f(\mathbf{a}) = (0 \ 0).$$

Für die Funktion ρ ergibt sich daraus

$$\rho(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{a}) - Jf(\mathbf{a}) \circ \mathbf{x} = f(\mathbf{x}).$$

Wir müssen den Grenzwert von

$$\frac{\rho(\mathbf{x})}{\|\mathbf{x}\|_2} = \frac{x_1 x_2}{x_1^2 + x_2^2}$$

im Punkt $\mathbf{a} = (0, 0)^t$ betrachten, aber dieser existiert nach Beispiel 40.5 a. nicht. Mithin ist die Funktion f nicht total differenzierbar in $\mathbf{a} = (0, 0)^t$.

Bemerkung 41.15 (Die Ableitung als Abbildung)

In der eindimensionalen Analysis haben wir die Ableitung einer auf \mathbf{U} differenzierbaren Funktion als eine Funktion auf \mathbf{U} aufgefaßt, die Werte in \mathbb{R} annimmt. Das war möglich, weil die Ableitung von f in einem Punkt \mathbf{a} eine reelle Zahl war. Im Mehrdimensionalen ist die Ableitung von $f : \mathbf{U} \rightarrow \mathbb{R}^m$ in einem Punkt aber eine $m \times n$ -Matrix. In diesem Sinne nimmt die Ableitung

$$Df : \mathbf{U} \rightarrow \text{Mat}(m \times n, \mathbb{R}) : \mathbf{a} \mapsto Df(\mathbf{a}) = Jf(\mathbf{a})$$

Werte in $\text{Mat}(m \times n, \mathbb{R})$ an. $\text{Mat}(m \times n, \mathbb{R})$ können wir mit \mathbb{R}^{mn} identifizieren und mit der euklidischen Norm versehen,

$$\|(\mathbf{a}_{ij})\|_2 = \sqrt{\sum_{i=1}^m \sum_{j=1}^n a_{ij}^2},$$

als normierten Raum betrachten. Das folgende Korollar zeigt dann, daß der Begriff der stetigen Differenzierbarkeit auch im Mehrdimensionalen sinnvoll gewählt ist, da eine Abbildung genau dann stetig differenzierbar ist, wenn sie (total) differenzierbar und ihre Ableitung stetig ist. Aus der Definition ist dies nicht ohne weiteres ersichtlich, da diese lediglich auf die partiellen Ableitungen Bezug nimmt.

Korollar 41.16 (Stetige Differenzierbarkeit)

Eine Abbildung $f : \mathbf{U} \rightarrow \mathbb{R}^m$ ist genau dann stetig differenzierbar auf \mathbf{U} , wenn f auf \mathbf{U} total differenzierbar ist und $Df : \mathbf{U} \rightarrow \text{Mat}(m \times n, \mathbb{R})$ auf \mathbf{U} stetig ist.

Beweis: Wir identifizieren bei dieser Aussage $\text{Mat}(m \times n, \mathbb{R})$ mit \mathbb{R}^{mn} und betrachten diesen Raum mit einer beliebigen Norm als normierten Raum.

Ist f stetig differenzierbar auf \mathbf{U} , so ist f nach Satz 41.10 auch differenzierbar auf \mathbf{U} . Mithin existiert die Abbildung Df , und ihre Komponentenfunktionen sind die partiellen Ableitungen von f , die nach Voraussetzung stetig auf \mathbf{U} sind. Dann ist aber nach Satz 40.19 auch Df stetig.

Ist umgekehrt f differenzierbar auf \mathbf{U} , so existieren nach Korollar 41.8 auch alle partiellen Ableitungen von f auf \mathbf{U} und diese sind die Komponentenfunktionen von Df . Wenn nun die Abbildung Df als stetig auf \mathbf{U} voraus gesetzt wird, so sind nach Satz 40.19 die partiellen Ableitungen von f also stetig auf \mathbf{U} , und f ist mithin stetig differenzierbar auf \mathbf{U} . \square

B) Rechenregeln für Ableitungen

Proposition 41.17 (Lineare Abbildungen sind total differenzierbar.)

Für $A \in \text{Mat}(m \times n, \mathbb{R})$ ist $f_A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ total differenzierbar in allen $\mathbf{a} \in \mathbb{R}^n$ mit

$$Df_A(\mathbf{a}) = A.$$

Beweis: Dies folgt unmittelbar aus der Definition der totalen Differenzierbarkeit mit der Nullabbildung als ρ . \square

Beispiel 41.18

Die Funktion $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R} : \mathbf{x} \mapsto x_i$ ist linear und somit total differenzierbar.

Proposition 41.19 (Linearität der Ableitung)

Sind $f, g : \mathcal{U} \rightarrow \mathbb{R}^m$ total differenzierbar in $\mathbf{a} \in \mathcal{U}$ und $\lambda, \mu \in \mathbb{R}$, so ist $\lambda \cdot f + \mu \cdot g$ total differenzierbar in \mathbf{a} mit

$$D(\lambda \cdot f + \mu \cdot g)(\mathbf{a}) = \lambda \cdot Df(\mathbf{a}) + \mu \cdot Dg(\mathbf{a}).$$

Beweis: Nach Voraussetzung gilt für

$$\rho_f(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{a}) - Df(\mathbf{a}) \circ (\mathbf{x} - \mathbf{a})$$

und

$$\rho_g(\mathbf{x}) = g(\mathbf{x}) - g(\mathbf{a}) - Dg(\mathbf{a}) \circ (\mathbf{x} - \mathbf{a}),$$

daß

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_f(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_g(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} = 0.$$

Dann gilt aber auch

$$(\lambda \cdot f + \mu \cdot g)(\mathbf{x}) - (\lambda \cdot f + \mu \cdot g)(\mathbf{a}) - (\lambda \cdot Df(\mathbf{a}) + \mu \cdot Dg(\mathbf{a})) \circ (\mathbf{x} - \mathbf{a}) = \lambda \cdot \rho_f(\mathbf{x}) + \mu \cdot \rho_g(\mathbf{x})$$

und

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\lambda \cdot \rho_f(\mathbf{x}) + \mu \cdot \rho_g(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} = \lambda \cdot \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_f(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} + \mu \cdot \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_g(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} = 0.$$

Damit folgt die Behauptung. \square

Proposition 41.20 (Produktregel)

Sind $f, g : \mathcal{U} \rightarrow \mathbb{R}$ total differenzierbar in $\mathbf{a} \in \mathcal{U}$, so ist auch $f \cdot g$ total differenzierbar in \mathbf{a} mit $D(f \cdot g)(\mathbf{a}) = f(\mathbf{a}) \cdot Dg(\mathbf{a}) + Df(\mathbf{a}) \cdot g(\mathbf{a})$.

Beweis: Nach Voraussetzung existieren $A = Df(\mathbf{a})$ und $B = Dg(\mathbf{a})$ und für die Funktionen $\rho_f : \mathcal{U} \rightarrow \mathbb{R}$ und $\rho_g : \mathcal{U} \rightarrow \mathbb{R}$ mit

$$\rho_f(\mathbf{x}) = f(\mathbf{x}) - f(\mathbf{a}) - A \circ (\mathbf{x} - \mathbf{a}) \tag{136}$$

und

$$\rho_g(\mathbf{x}) = g(\mathbf{x}) - g(\mathbf{a}) - B \circ (\mathbf{x} - \mathbf{a}) \tag{137}$$

gilt

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_f(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} = \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_g(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} = 0.$$

Der Übersichtlichkeit halber setzen wir zwischendurch $\mathbf{y} = \mathbf{x} - \mathbf{a}$ und erhalten dann unter Berücksichtigung von (136) und (137)

$$\begin{aligned} \rho_{f \cdot g}(\mathbf{x}) &:= (f \cdot g)(\mathbf{x}) - (f \cdot g)(\mathbf{a}) - (f(\mathbf{a}) \cdot \mathbf{B} + \mathbf{A} \cdot g(\mathbf{a})) \circ (\mathbf{x} - \mathbf{a}) \\ &= (f(\mathbf{x}) - f(\mathbf{a})) \cdot g(\mathbf{x}) + f(\mathbf{a}) \cdot (g(\mathbf{x}) - g(\mathbf{a})) - f(\mathbf{a}) \cdot \mathbf{B}\mathbf{y} - \mathbf{A}\mathbf{y} \cdot g(\mathbf{a}) \\ &= (\rho_f(\mathbf{x}) + \mathbf{A}\mathbf{y}) \cdot g(\mathbf{x}) + f(\mathbf{a}) \cdot (\rho_g(\mathbf{x}) + \mathbf{B}\mathbf{y}) - f(\mathbf{a}) \cdot \mathbf{B}\mathbf{y} - \mathbf{A}\mathbf{y} \cdot g(\mathbf{a}) \\ &= \rho_f(\mathbf{x}) \cdot g(\mathbf{x}) + f(\mathbf{a}) \cdot \rho_g(\mathbf{x}) + \mathbf{A}\mathbf{y} \cdot (g(\mathbf{x}) - g(\mathbf{a})) \\ &= \rho_f(\mathbf{x}) \cdot g(\mathbf{x}) + f(\mathbf{a}) \cdot \rho_g(\mathbf{x}) + \mathbf{A}\mathbf{y} \cdot (\rho_g(\mathbf{x}) + \mathbf{B}\mathbf{y}) \\ &= \rho_f(\mathbf{x}) \cdot g(\mathbf{x}) + f(\mathbf{a}) \cdot \rho_g(\mathbf{x}) + f_{\mathbf{A}}(\mathbf{x} - \mathbf{a}) \cdot \rho_g(\mathbf{x}) + f_{\mathbf{A}}(\mathbf{x} - \mathbf{a}) \cdot f_{\mathbf{B}}(\mathbf{x} - \mathbf{a}). \end{aligned}$$

Wir beachten nun, daß die Multiplikation mit einer Matrix eine stetige lineare Abbildung ist, und mit Korollar 40.44 gilt deshalb

$$\frac{|f_{\mathbf{A}}(\mathbf{x} - \mathbf{a})| \cdot |f_{\mathbf{B}}(\mathbf{x} - \mathbf{a})|}{\|\mathbf{x} - \mathbf{a}\|_2} \leq \frac{\|f_{\mathbf{A}}\| \cdot \|\mathbf{x} - \mathbf{a}\|_2 \cdot \|f_{\mathbf{B}}\| \cdot \|\mathbf{x} - \mathbf{a}\|_2}{\|\mathbf{x} - \mathbf{a}\|_2} = \|f_{\mathbf{A}}\| \cdot \|f_{\mathbf{B}}\| \cdot \|\mathbf{x} - \mathbf{a}\|_2,$$

wobei $\|f_{\mathbf{A}}\|$ und $\|f_{\mathbf{B}}\|$ die Operatornormen von $f_{\mathbf{A}}$ und $f_{\mathbf{B}}$ sind. Da die rechte Seite für $\mathbf{x} \rightarrow \mathbf{a}$ gegen Null konvergiert, folgt

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f_{\mathbf{A}}(\mathbf{x} - \mathbf{a}) \cdot f_{\mathbf{B}}(\mathbf{x} - \mathbf{a})}{\|\mathbf{x} - \mathbf{a}\|_2} = 0.$$

Nutzen wir nun noch aus, daß die in \mathbf{a} total differenzierbare Funktion g nach Proposition 41.13 auch stetig in \mathbf{a} ist und daß die Multiplikation mit einer Matrix eine stetige Abbildung ist, so erhalten wir

$$\begin{aligned} \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_{f \cdot g}(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} &= \\ &= \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_f(\mathbf{x}) \cdot g(\mathbf{x}) + f(\mathbf{a}) \cdot \rho_g(\mathbf{x}) + f_{\mathbf{A}}(\mathbf{x} - \mathbf{a}) \cdot \rho_g(\mathbf{x}) + f_{\mathbf{A}}(\mathbf{x} - \mathbf{a}) \cdot f_{\mathbf{B}}(\mathbf{x} - \mathbf{a})}{\|\mathbf{x} - \mathbf{a}\|_2} \\ &= \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_f(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} \cdot \lim_{\mathbf{x} \rightarrow \mathbf{a}} g(\mathbf{x}) + f(\mathbf{a}) \cdot \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_g(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} \\ &\quad + \lim_{\mathbf{x} \rightarrow \mathbf{a}} f_{\mathbf{A}}(\mathbf{x} - \mathbf{a}) \cdot \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_g(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2} + \lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f_{\mathbf{A}}(\mathbf{x} - \mathbf{a}) \cdot f_{\mathbf{B}}(\mathbf{x} - \mathbf{a})}{\|\mathbf{x} - \mathbf{a}\|_2} \\ &= 0 \cdot g(\mathbf{a}) + f(\mathbf{a}) \cdot 0 + 0 \cdot 0 + 0 = 0. \end{aligned}$$

Also ist $f \cdot g$ total differenzierbar in \mathbf{a} mit der Ableitung

$$D(f \cdot g)(\mathbf{a}) = f(\mathbf{a}) \cdot \mathbf{B} + \mathbf{A} \cdot g(\mathbf{a}) = f(\mathbf{a}) \cdot Dg(\mathbf{a}) + Df(\mathbf{a}) \cdot g(\mathbf{a}).$$

□

Beispiel 41.21 (Polynomfunktionen sind total differenzierbar.)

- a. Aus der Linearität der Ableitung 41.19, der Produktregel 41.20 und Beispiel 41.18 folgt, daß jede Polynomfunktion

$$\mathbf{p} : \mathbb{R}^n \longrightarrow \mathbb{R} : \mathbf{x} \mapsto \sum_{|\alpha|=0}^d \mathbf{a}_{\alpha} \cdot \mathbf{x}^{\alpha}$$

total differenzierbar auf \mathbb{R}^n ist mit

$$\frac{\partial p}{\partial x_j}(x) = \sum_{|\alpha|=1}^d a_\alpha \cdot \alpha_j \cdot x^{\alpha - e_j}.$$

b. Das Normquadrat

$$N^2 : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto \|x\|_2^2 = x_1^2 + \dots + x_n^2$$

ist eine Polynomfunktion und ist mithin total differenzierbar mit

$$DN^2(a) = (2a_1 \ \dots \ 2a_n).$$

Proposition 41.22 (Kettenregel)

Ist $f : U \longrightarrow \mathbb{R}^m$ total differenzierbar in $a \in U$ mit $f(U) \subseteq V$ und V offen in \mathbb{R}^m und ist $g : V \longrightarrow \mathbb{R}^k$ total differenzierbar in $f(a)$, so ist auch $h = g \circ f : U \longrightarrow \mathbb{R}^k$ total differenzierbar in a mit

$$Dh(a) = Dg(f(a)) \circ Df(a).$$

Beweis: Nach Voraussetzung existieren $A = Df(a) \in \text{Mat}(m \times n, \mathbb{R})$ und $B = Dg(f(a)) \in \text{Mat}(k \times m, \mathbb{R})$ und für die Funktionen $\rho_f : U \longrightarrow \mathbb{R}$ und $\rho_g : U \longrightarrow \mathbb{R}$ mit

$$\rho_f(x) = f(x) - f(a) - A \circ (x - a) \quad (138)$$

und

$$\rho_g(y) = g(y) - g(f(a)) - B \circ (y - f(a)) \quad (139)$$

gilt

$$\lim_{x \rightarrow a} \frac{\rho_f(x)}{\|x - a\|_2} = \lim_{y \rightarrow f(a)} \frac{\rho_g(y)}{\|y - f(a)\|_2} = 0. \quad (140)$$

Wir erhalten damit

$$\begin{aligned} \rho_h(x) &:= h(x) - h(a) - B \circ A \circ (x - a) \\ &= g(f(x)) - g(f(a)) - B \circ A \circ (x - a) \\ &\stackrel{(139)}{=} \rho_g(f(x)) + B \circ (f(x) - f(a)) - B \circ A \circ (x - a) \\ &= \rho_g(f(x)) + B \circ (f(x) - f(a) - A \circ (x - a)) \\ &= \rho_g(f(x)) + B \circ \rho_f(x) = \rho_g(f(x)) + f_B(\rho_f(x)). \end{aligned}$$

Die Kettenregel ist nun bewiesen, sobald wir zeigen können, daß

$$\lim_{x \rightarrow a} \frac{\rho_h(x)}{\|x - a\|_2} = \lim_{x \rightarrow a} \frac{\rho_g(f(x)) + f_B(\rho_f(x))}{\|x - a\|_2} = 0$$

gilt. Dazu beachten wir zunächst, daß

$$\frac{\|f_B(\rho_f(x))\|_2}{\|x - a\|_2} \stackrel{40.44}{\leq} \|f_B\| \cdot \left\| \frac{\rho_f(x)}{\|x - a\|_2} \right\|_2 \longrightarrow 0$$

für $x \rightarrow a$, wobei $\|f_B\|$ die Operatornorm von f_B ist. Also gilt schon mal

$$\lim_{x \rightarrow a} \frac{f_B(\rho_f(x))}{\|x - a\|_2} = 0,$$

und es reicht mithin noch

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_g(f(\mathbf{x}))}{\|\mathbf{x} - \mathbf{a}\|_2} = 0$$

zu zeigen. Dies wollen wir im folgenden tun.

Wegen (140) ist die Abbildung

$$\sigma : V \longrightarrow \mathbb{R}^k : \mathbf{y} \mapsto \begin{cases} \frac{\rho_g(\mathbf{y})}{\|\mathbf{y} - f(\mathbf{a})\|_2}, & \text{falls } \mathbf{y} \neq f(\mathbf{a}), \\ 0, & \text{sonst} \end{cases}$$

stetig in $\mathbf{y} = f(\mathbf{a})$. Da außerdem $\rho_g(f(\mathbf{a})) = 0$ wegen (139) gilt, erhalten wir

$$\rho_g(\mathbf{y}) = \sigma(\mathbf{y}) \cdot \|\mathbf{y} - f(\mathbf{a})\|_2$$

für $\mathbf{y} \in V$ und somit mit $\mathbf{y} = f(\mathbf{x})$

$$\begin{aligned} \left\| \frac{\rho_g(f(\mathbf{x}))}{\|\mathbf{x} - \mathbf{a}\|_2} \right\|_2 &= \|\sigma(f(\mathbf{x}))\|_2 \cdot \frac{\|f(\mathbf{x}) - f(\mathbf{a})\|_2}{\|\mathbf{x} - \mathbf{a}\|_2} \\ &\stackrel{(138)}{=} \|\sigma(f(\mathbf{x}))\|_2 \cdot \frac{\|A \circ (\mathbf{x} - \mathbf{a}) + \rho_f(\mathbf{x})\|_2}{\|\mathbf{x} - \mathbf{a}\|_2} \\ &\leq \|\sigma(f(\mathbf{x}))\|_2 \cdot \left(\frac{\|f_A(\mathbf{x} - \mathbf{a})\|_2}{\|\mathbf{x} - \mathbf{a}\|_2} + \frac{\|\rho_f(\mathbf{x})\|_2}{\|\mathbf{x} - \mathbf{a}\|_2} \right) \\ &\leq \|\sigma(f(\mathbf{x}))\|_2 \cdot \left(\frac{\|f_A\| \cdot \|\mathbf{x} - \mathbf{a}\|_2}{\|\mathbf{x} - \mathbf{a}\|_2} + \frac{\|\rho_f(\mathbf{x})\|_2}{\|\mathbf{x} - \mathbf{a}\|_2} \right) \\ &\leq \|\sigma(f(\mathbf{x}))\|_2 \cdot \left(\|f_A\| + \frac{\|\rho_f(\mathbf{x})\|_2}{\|\mathbf{x} - \mathbf{a}\|_2} \right), \end{aligned}$$

wobei $\|f_A\|$ die Operatornorm von f_A ist. Nach Proposition 41.13 ist f stetig in \mathbf{a} und zudem ist σ stetig in $f(\mathbf{a})$, so daß nach Proposition 40.15 auch $\sigma \circ f$ stetig in \mathbf{a} ist. Unter Berücksichtigung von (140) konvergiert die rechte Seite der Ungleichung für $\mathbf{x} \rightarrow \mathbf{a}$ deshalb gegen

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \left(\|\sigma(f(\mathbf{x}))\|_2 \cdot \left(\|f_A\| + \frac{\|\rho_f(\mathbf{x})\|_2}{\|\mathbf{x} - \mathbf{a}\|_2} \right) \right) = \|\sigma(f(\mathbf{a}))\|_2 \cdot (\|f_A\| + 0) = 0,$$

so daß auch die linke Seite der Ungleichung gegen 0 konvergieren muß. Wir erhalten deshalb

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho_g(f(\mathbf{x}))}{\|\mathbf{x} - \mathbf{a}\|_2} = 0,$$

und damit ist die Behauptung bewiesen. \square

Beispiel 41.23 (Kettenregel)

a. Aus der Kettenregel folgt, daß die euklidische Norm

$$N : \mathbb{R}^n \longrightarrow \mathbb{R} : \mathbf{x} \mapsto \|\mathbf{x}\|_2 = \sqrt{x_1^2 + \dots + x_n^2}$$

als Verkettung einer Polynomfunktion

$$p : \mathbb{R}^n \longrightarrow \mathbb{R} : \mathbf{x} \mapsto x_1^2 + \dots + x_n^2$$

mit der Wurzelfunktion

$$w : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R} : y \mapsto \sqrt{y}$$

differenzierbar auf $\mathbb{R}^n \setminus \{0\}$ ist mit

$$DN(\mathbf{a}) = Dw(p(\mathbf{a})) \cdot Dp(\mathbf{a}) = \frac{1}{2\sqrt{p(\mathbf{a})}} \cdot (2\mathbf{a}_1, \dots, 2\mathbf{a}_n) = \left(\frac{\mathbf{a}_1}{\|\mathbf{a}\|_2}, \dots, \frac{\mathbf{a}_n}{\|\mathbf{a}\|_2} \right).$$

Man beachte hierbei, daß die Wurzelfunktion in 0 nicht differenzierbar ist.

b. Die Funktion

$$g : \mathbb{R}^n \setminus \{0\} \longrightarrow \mathbb{R} : \mathbf{x} \mapsto \frac{1}{\|\mathbf{x}\|_2}$$

ist dann als Verkettung $g = \text{inv} \circ N$ der Norm $N : \mathbb{R}^2 \longrightarrow \mathbb{R}$ mit der Inversion $\text{inv} : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} \setminus \{0\} : x \mapsto \frac{1}{x}$ ebenfalls total differenzierbar auf $\mathbb{R}^n \setminus \{0\}$ mit

$$\begin{aligned} Dg(\mathbf{a}) &= D \text{inv}(N(\mathbf{a})) \circ DN(\mathbf{a}) \\ &= -\frac{1}{N(\mathbf{a})^2} \cdot \left(\frac{x_1}{\|\mathbf{x}\|_2}, \dots, \frac{x_n}{\|\mathbf{x}\|_2} \right) = \left(\frac{-x_1}{\|\mathbf{x}\|_2^3}, \dots, \frac{-x_n}{\|\mathbf{x}\|_2^3} \right). \end{aligned}$$

Proposition 41.24 (Quotientenregel)

Seien $f, g : U \longrightarrow \mathbb{R}$ total differenzierbar in $\mathbf{a} \in U$ mit $g(\mathbf{a}) \neq 0$ und g stetig auf U . Dann ist $U \setminus g^{-1}(0)$ offen in \mathbb{R}^n und $\frac{f}{g} : U \setminus g^{-1}(0) \longrightarrow \mathbb{R}$ ist total differenzierbar in \mathbf{a} mit

$$D \left(\frac{f}{g} \right) (\mathbf{a}) = \frac{Df(\mathbf{a}) \cdot g(\mathbf{a}) - f(\mathbf{a}) \cdot Dg(\mathbf{a})}{g^2(\mathbf{a})}.$$

Beweis: Da g stetig auf U ist, ist $g^{-1}(0)$ abgeschlossen in U nach Proposition 40.17 und $U \setminus g^{-1}(0)$ ist offen in U . Nach Aufgabe 39.52 ist $U \setminus g^{-1}(0)$ dann der Durchschnitt von U mit einer in \mathbb{R}^n offenen Menge O , und ist damit als Durchschnitt zweier in \mathbb{R}^n offenen Mengen offen in \mathbb{R}^n .

Die Abbildung

$$\text{inv} : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} \setminus \{0\} : x \mapsto \frac{1}{x},$$

ist total differenzierbar auf $\mathbb{R} \setminus \{0\}$ mit

$$D \text{inv}(x) = -\frac{1}{x^2}.$$

Somit ist auch

$$\frac{f}{g} = f \cdot (\text{inv} \circ g)$$

total differenzierbar in \mathbf{a} wegen der Produktregel 41.20 und der Kettenregel 41.22.

Für die Ableitung folgt zudem

$$\begin{aligned} D \left(\frac{f}{g} \right) (\mathbf{a}) &= Df(\mathbf{a}) \cdot (\text{inv} \circ g)(\mathbf{a}) + f(\mathbf{a}) \cdot D(\text{inv} \circ g)(\mathbf{a}) \\ &= Df(\mathbf{a}) \cdot (\text{inv} \circ g)(\mathbf{a}) + f(\mathbf{a}) \cdot D \text{inv}(g(\mathbf{a})) \cdot Dg(\mathbf{a}) \\ &= \frac{Df(\mathbf{a})}{g(\mathbf{a})} - \frac{f(\mathbf{a}) \cdot Dg(\mathbf{a})}{g^2(\mathbf{a})} \\ &= \frac{Df(\mathbf{a}) \cdot g(\mathbf{a}) - f(\mathbf{a}) \cdot Dg(\mathbf{a})}{g^2(\mathbf{a})}. \end{aligned}$$

□

Beispiel 41.25 (Rationale Funktionen sind total differenzierbar.)

Da Polynomfunktionen nach Beispiel 41.21 total differenzierbar und stetig sind, folgt aus der Quotientenregel 41.24, daß auch jede rationale Funktion auf ihrem Definitionsbereich total differenzierbar ist.

C) Geometrische Interpretation des Gradienten

Definition 41.26 (Gradient)

Ist $f : \mathbf{U} \rightarrow \mathbb{R}$ eine in $\mathbf{a} \in \mathbf{U}$ partiell differenzierbare Funktion, so nennen wir

$$\text{grad}(f(\mathbf{a})) = \nabla f(\mathbf{a}) = \mathbf{J}f(\mathbf{a})^t = \left(\frac{\partial f}{\partial x_1}(\mathbf{a}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{a}) \right)^t$$

auch den *Gradienten* von f im Punkt \mathbf{a} .³

Bemerkung 41.27 (Geometrische Interpretation des Gradienten)

Ist $f : \mathbf{U} \rightarrow \mathbb{R}$ total differenzierbar in \mathbf{a} , so ist

$$\mathbf{x} \mapsto f(\mathbf{a}) + \mathbf{J}f(\mathbf{a}) \circ (\mathbf{x} - \mathbf{a}) = f(\mathbf{a}) + \langle \nabla f(\mathbf{a}), \mathbf{x} - \mathbf{a} \rangle$$

die lineare Approximation von f im Punkt \mathbf{a} . Ihr Graph ist eine Hyperebene, d.h. ein affiner Raum der Dimension n , im \mathbb{R}^{n+1} , die den Graphen von f im Punkt $(\mathbf{a}, f(\mathbf{a}))$ approximiert. Man nennt diese Hyperebene auch den Tangentialraum an den Graphen von f im Punkt $(\mathbf{a}, f(\mathbf{a}))$ (siehe Abbildung 5).

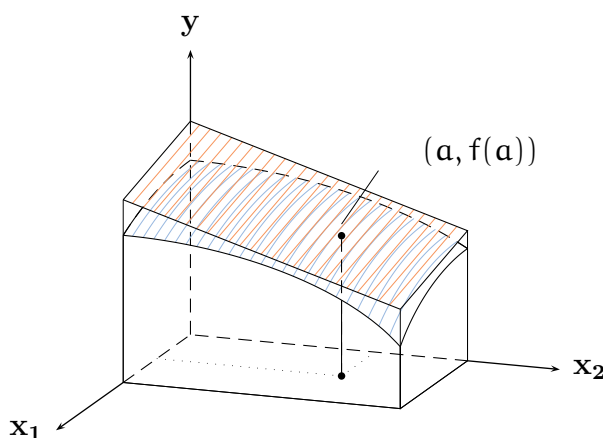


ABBILDUNG 5. Tangentialraum an den Graphen von f in $(\mathbf{a}, f(\mathbf{a}))$

Der Gradient von f in \mathbf{a} , als Vektor aufgefaßt, gibt dabei die Richtung des *größten Anstiegs* von f an. Um dies zu sehen, beachten wir, daß für den Winkel α zwischen $\mathbf{x} - \mathbf{a}$ und $\nabla f(\mathbf{a})$

$$\langle \nabla f(\mathbf{a}), \mathbf{x} - \mathbf{a} \rangle = \|\nabla f(\mathbf{a})\|_2 \cdot \|\mathbf{x} - \mathbf{a}\|_2 \cdot \cos(\alpha)$$

³Wir verwenden die Kommata in der Darstellung der $1 \times n$ -Matrix $\mathbf{J}f(\mathbf{a})$, um die einzelnen Spalten besser unterscheiden zu können. Dies haben wir auch weiter oben in den Beispielen schon so getan. Das Symbol ∇ wird Nabla genannt.

gilt. Mithin ist der Zuwachs von $f(\mathbf{a})$ zu

$$f(\mathbf{x}) \approx f(\mathbf{a}) + \langle \nabla f(\mathbf{a}), \mathbf{x} - \mathbf{a} \rangle = f(\mathbf{a}) + \|\nabla f(\mathbf{a})\|_2 \cdot \|\mathbf{x} - \mathbf{a}\|_2 \cdot \cos(\alpha)$$

maximal, wenn $\cos(\alpha) = 1$ ist. Das ist aber genau dann der Fall, wenn $\alpha = 0$ gilt, wenn also $\nabla f(\mathbf{a})$ und $\mathbf{x} - \mathbf{a}$ die gleiche Richtung haben.

Man kann diesen Sachverhalt auch mit Hilfe der Richtungsableitung ausdrücken. Wegen Lemma 41.4 ist die Richtungsableitung von f im Punkt \mathbf{a} in Richtung \mathbf{v} mit $\|\mathbf{v}\|_2 = 1$ durch

$$\frac{\partial f}{\partial \mathbf{v}}(\mathbf{a}) = Df(\mathbf{a}) \circ \mathbf{v} = \langle \nabla f(\mathbf{a}), \mathbf{v} \rangle$$

gegeben (siehe auch Abbildung 4). Wegen der Gleichung von Cauchy-Schwarz 37.8 gilt

$$\left| \frac{\partial f}{\partial \mathbf{v}}(\mathbf{a}) \right| = |\langle \nabla f(\mathbf{a}), \mathbf{v} \rangle| \leq \|\nabla f(\mathbf{a})\|_2 \cdot \|\mathbf{v}\|_2 = \|\nabla f(\mathbf{a})\|_2,$$

und es gilt genau dann Gleichheit, wenn $\nabla f(\mathbf{a})$ und \mathbf{v} linear abhängig voneinander sind. Die Richtungsableitung in Richtung \mathbf{v} wird also maximal, in Richtung

$$\mathbf{v} = \frac{\nabla f(\mathbf{a})}{\|\nabla f(\mathbf{a})\|_2}.$$

Bei all diesen Betrachtungen sollte der Gradient $\nabla f(\mathbf{a})$ natürlich nicht der Nullvektor sein.

D) Differenzierbarkeit in normierten Räumen

Bemerkung 41.28 (Fréchet-Differenzierbarkeit)

Es seien V und W zwei normierte Räume, $U \subseteq V$ sei offen und $f : U \rightarrow W$ sei eine Abbildung. Dann heißt f *Fréchet-differenzierbar* in $\mathbf{a} \in U$, wenn es eine stetige lineare Abbildung $A \in L(V, W)$ und eine Abbildung $\rho : U \rightarrow W$ gibt, mit

$$f(\mathbf{x}) = f(\mathbf{a}) + A(\mathbf{x} - \mathbf{a}) + \rho(\mathbf{x})$$

und

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{\rho(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_V} = 0.$$

Die Ableitung $Df(\mathbf{a}) := A$ ist dann wieder eindeutig bestimmt. Außerdem gelten die Linearität der Ableitung, die Produktregel (wenn $W = \mathbb{R}$), die Kettenregel und die Quotientenregel (wenn $W = \mathbb{R}$) ohne Änderung des Beweises. Ebenso folgt aus der Differenzierbarkeit in \mathbf{a} die Stetigkeit in \mathbf{a} und stetige lineare Abbildungen sind differenzierbar, auch diese Aussagen gelten mit demselben Beweis.

Ist die Abbildung f nun in jedem Punkt von U Fréchet-differenzierbar, so ist ihre Ableitung eine Abbildung

$$Df : U \rightarrow L(V, W) : \mathbf{a} \mapsto Df(\mathbf{a})$$

die Werte in $L(V, W)$ annimmt.

Aufgaben

Aufgabe 41.29 (Total differenzierbar impliziert nicht stetig differenzierbar.)

Die Funktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \|x\|_2^2 \cdot \sin\left(\frac{1}{\|x\|_2}\right), & \text{falls } x \neq 0, \\ 0 & \text{sonst} \end{cases}$$

ist total differenzierbar auf \mathbb{R}^2 , aber nicht stetig differenzierbar.

Aufgabe 41.30

Begründe, weshalb $f : \mathbb{R}_{>0}^3 \longrightarrow \mathbb{R}^2 : (x, y, z)^t \mapsto (z\sqrt{y} + \sqrt{z}, xyz + \ln(x + y))^t$ total differenzierbar ist, und berechne die Ableitung von f .

Aufgabe 41.31

Zeige, für $A \in \text{Mat}_n(\mathbb{R})$ ist $f : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto x^t \circ A \circ x$ total differenzierbar auf \mathbb{R}^n und berechne die Ableitung.

Aufgabe 41.32

Sei $\Delta = D_1^2 + \dots + D_n^2$ der *Laplace-Operator* und $f \in \mathcal{C}^2((0, 1), \mathbb{R})$.

Zeige, daß für die Funktion $\varphi : \mathbb{R}^n \setminus \{(0, \dots, 0)^t\} \longrightarrow \mathbb{R}, x \mapsto f(\|x\|_2)$ die Gleichung

$$\Delta\varphi(x) = f''(\|x\|_2) + \frac{n-1}{\|x\|_2} \cdot f'(\|x\|_2) \quad \text{für alle } x \in \mathbb{R}^n \setminus \{(0, \dots, 0)^t\}$$

gilt und berechne $\Delta\varphi$ für den Fall $f(t) = \frac{1}{t^{n-2}}$ mit $n > 2$.

Aufgabe 41.33

Zeige, ist $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ eine in $(0, \dots, 0)^t \in \mathbb{R}^n$ total differenzierbare Abbildung mit der Eigenschaft $f(\lambda x) = \lambda f(x)$ für alle $\lambda \in \mathbb{R}, x \in \mathbb{R}^n$, so ist f \mathbb{R} -linear.

§ 42 Der Satz von Taylor und seine Anwendungen

In diesem Abschnitt betrachten wir \mathbb{R}^n und \mathbb{R}^m stets als normierten Raum bezüglich der euklidischen Norm $\|\cdot\|_2$. Außerdem sei $U \subseteq \mathbb{R}^n$ stets eine *offene* Teilmenge des \mathbb{R}^n , so daß jeder Punkt in U ein Häufungspunkt ist.

A) Der Satz von Schwarz

Beispiel 42.1

Die Polynomfunktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : \mathbf{x} = (x_1, x_2)^t \mapsto 3x_1^5 + 2x_1^3x_2 + 5x_2$$

ist partiell differenzierbar auf \mathbb{R}^2 . Ihre partiellen Ableitungen

$$D_1f : \mathbb{R}^2 \longrightarrow \mathbb{R} : \mathbf{x} \mapsto D_1f(\mathbf{x}) = \frac{\partial f}{\partial x_1}(\mathbf{x}) = 15x_1^4 + 6x_1^2x_2$$

und

$$D_2f : \mathbb{R}^2 \longrightarrow \mathbb{R} : \mathbf{x} \mapsto D_2f(\mathbf{x}) = \frac{\partial f}{\partial x_2}(\mathbf{x}) = 2x_1^3 + 5$$

sind ebenfalls partiell differenzierbar auf \mathbb{R}^2 . Wir können also auch von diesen wieder die partiellen Ableitungen bilden und erhalten:

$$D_1D_1f(\mathbf{x}) = 60x_1^3 + 12x_1x_2, \quad D_2D_1f(\mathbf{x}) = 6x_1^2, \quad D_1D_2f(\mathbf{x}) = 6x_1^2, \quad D_2D_2f(\mathbf{x}) = 0.$$

Dabei fällt uns auf, daß

$$D_2D_1f(\mathbf{x}) = 6x_1^2 = D_1D_2f(\mathbf{x})$$

gilt. In unserem Beispiel war es also egal, ob wir zuerst nach x_1 und dann nach x_2 abgeleitet haben oder umgekehrt. Dies ist ein allgemeines Prinzip, wie wir im Satz von Schwarz 42.5 sehen werden. Zunächst wollen wir aber allgemein höhere partielle Ableitungen einführen.

Definition 42.2 (Höhere partielle Ableitungen)

Eine Funktion $f : U \longrightarrow \mathbb{R}$ heißt *k-fach stetig differenzierbar* auf U , falls für alle $j_1, \dots, j_k \in \{1, \dots, n\}$ die *k-fachen* partiellen Ableitungen

$$D_{j_k} \dots D_{j_1} f : U \longrightarrow \mathbb{R}$$

existieren und stetig sind, d.h. für $l = 1, \dots, k$ ist die Funktion $D_{j_{l-1}} \dots D_{j_1} f$ auf U partiell differenzierbar nach x_{j_l} und die Funktion $D_{j_l} \dots D_{j_1} f$ ist stetig auf U . Man beachte, daß es insgesamt n^k solche *k-fachen* partiellen Ableitungen von f gibt.

Ein Abbildung $f : U \longrightarrow \mathbb{R}^m$ heißt dann *k-fach stetig differenzierbar* auf U , wenn alle Komponentenfunktionen *k-fach* stetig differenzierbar auf U sind.

Wir bezeichnen mit

$$\mathcal{C}^k(U, \mathbb{R}^m) = \{f : U \longrightarrow \mathbb{R}^m \mid f \text{ ist } k\text{-fach stetig differenzierbar auf } U\}$$

die Menge der auf U *k-fach* stetig differenzierbaren Abbildungen nach \mathbb{R}^m .

Bemerkung 42.3 ($\mathcal{C}^k(\mathbf{U}, \mathbb{R}^m)$)

Da die Differenzierbarkeit und die Stetigkeit verträglich mit Linearkombinationen sind (siehe Proposition 41.19 und Proposition 40.13) ist $\mathcal{C}^k(\mathbf{U}, \mathbb{R}^m)$ ein \mathbb{R} -Vektorraum. Zudem folgt aus der Definition und den Ergebnissen des letzten Abschnittes unmittelbar, daß

$$\mathcal{C}(\mathbf{U}, \mathbb{R}^m) \supseteq \mathcal{C}^1(\mathbf{U}, \mathbb{R}^m) \supseteq \mathcal{C}^2(\mathbf{U}, \mathbb{R}^m) \supseteq \mathcal{C}^3(\mathbf{U}, \mathbb{R}^m) \supseteq \dots$$

Man nennt dann den \mathbb{R} -Vektorraum

$$\mathcal{C}^\infty(\mathbf{U}, \mathbb{R}^m) = \bigcap_{k \geq 0} \mathcal{C}^k(\mathbf{U}, \mathbb{R}^m)$$

den Vektorraum der *unendlich oft differenzierbaren Abbildungen* auf \mathbf{U} .

Beispiel 42.4 (Die Glockenfunktion)

Die Funktion

$$f: \mathbb{R}^n \longrightarrow \mathbb{R}: \mathbf{x} \mapsto \begin{cases} \exp\left(\frac{1}{(1-\|\mathbf{x}\|_2^2)^2}\right), & \text{falls } \|\mathbf{x}\|_2 < 1, \\ 0, & \text{falls } \|\mathbf{x}\|_2 \geq 1 \end{cases}$$

ist unendlich oft differenzierbar auf \mathbb{R}^n , da sie die Verkettung der unendlich oft differenzierbaren Funktion $\mathbb{R} \rightarrow \mathbb{R}: t \mapsto \exp(-\frac{1}{t^2})$ aus Beispiel 18.30 mit der unendlich oft differenzierbaren Funktion $\mathbb{R}^n \rightarrow \mathbb{R}: \mathbf{x} \mapsto 1 - \|\mathbf{x}\|_2^2$ ist. f nimmt im Inneren der Einheitskugel nur positive Werte an und ist ansonsten stets null.

Satz 42.5 (Schwarz)

Ist $f: \mathbf{U} \rightarrow \mathbb{R}^m$ zweifach stetig differenzierbar, so gilt für alle $\mathbf{a} \in \mathbf{U}$ und $1 \leq i, j \leq n$

$$D_i D_j f(\mathbf{a}) = D_j D_i f(\mathbf{a}).$$

Beweis: Es reicht, die Aussage für die Komponentenfunktionen von f zu zeigen, so daß wir ohne Einschränkung annehmen können, daß $m = 1$ gilt.

Da \mathbf{U} offen ist, gibt es ein $\varepsilon > 0$ mit $\mathbf{U}_\varepsilon(\mathbf{a}) \subseteq \mathbf{U}$, und da die euklidische Norm nach Lemma 39.41 äquivalent zur Maximumnorm ist, gibt es zudem ein $\delta > 0$, so daß

$$\{\mathbf{a} + s \cdot \mathbf{e}_i + t \cdot \mathbf{e}_j \mid -\delta \leq s, t \leq \delta\} \subset \mathbf{U}_\varepsilon(\mathbf{a}) \subseteq \mathbf{U}.$$

Wir interessieren uns im folgenden für die Funktion $H: \mathbf{O} \rightarrow \mathbb{R}$ mit

$$H(s, t) = \frac{f(\mathbf{a} + s\mathbf{e}_i + t\mathbf{e}_j) - f(\mathbf{a} + t\mathbf{e}_j) - f(\mathbf{a} + s\mathbf{e}_i) + f(\mathbf{a})}{s \cdot t},$$

mit $\mathbf{O} = \{(s, t)^t \in \mathbb{R}^2 \mid -\delta < s, t < \delta \text{ und } s, t \neq 0\} \subset \mathbb{R}^2$ offen, da

$$D_j D_i f(\mathbf{a}) = \lim_{t \rightarrow 0} \frac{D_i f(\mathbf{a} + t\mathbf{e}_j) - D_i f(\mathbf{a})}{t} = \lim_{t \rightarrow 0} \lim_{s \rightarrow 0} H(s, t)$$

und

$$D_i D_j f(\mathbf{a}) = \lim_{s \rightarrow 0} \frac{D_j f(\mathbf{a} + s\mathbf{e}_i) - D_j f(\mathbf{a})}{s} = \lim_{s \rightarrow 0} \lim_{t \rightarrow 0} H(s, t).$$

In der Tat werden wir zeigen, daß

$$D_i D_j f(\mathbf{a}) = \lim_{(s,t) \rightarrow (0,0)} H(s, t) = D_j D_i f(\mathbf{a})$$

gilt, wenn wir hier die Konvergenz von $(s, t)^t$ gegen $(0, 0)^t$ im \mathbb{R}^2 bezüglich der euklidischen Norm betrachten.

Halten wir $s \in (-\delta, \delta)$ fest, so ist die Funktion

$$F_s : (-\delta, \delta) \longrightarrow \mathbb{R} : t \mapsto f(\mathbf{a} + se_i + te_j) - f(\mathbf{a} + te_j)$$

differenzierbar und die Ableitung von F_s ergibt sich aus der partiellen Ableitung von f nach x_j wie folgt:

$$\begin{aligned} F'_s(t) &= \lim_{h \rightarrow 0} \frac{F_s(t+h) - F_s(t)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(\mathbf{a} + se_i + (t+h)e_j) - f(\mathbf{a} + (t+h)e_j) - f(\mathbf{a} + se_i + te_j) + f(\mathbf{a} + te_j)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(\mathbf{a} + se_i + te_j + he_j) - f(\mathbf{a} + se_i + te_j)}{h} - \lim_{h \rightarrow 0} \frac{f(\mathbf{a} + te_j + he_j) - f(\mathbf{a} + te_j)}{h} \\ &= D_j f(\mathbf{a} + se_i + te_j) - D_j f(\mathbf{a} + te_j). \end{aligned}$$

Wenden wir nun den Mittelwertsatz 18.7 an, so finden wir zu jedem $t \in (-\delta, \delta)$ ein $\theta_{s,t} \in \mathbb{R}$ mit $|\theta_{s,t}| < |t|$ und

$$F_s(t) - F_s(0) = t \cdot F'_s(\theta_{s,t}) = t \cdot (D_j f(\mathbf{a} + se_i + \theta_{s,t}e_j) - D_j f(\mathbf{a} + \theta_{s,t}e_j)). \quad (141)$$

Auch die Funktion

$$G_{s,t} : (-\delta, \delta) \longrightarrow \mathbb{R} : r \mapsto D_j f(\mathbf{a} + re_i + \theta_{s,t}e_j)$$

ist differenzierbar und ihre Ableitung ergibt sich aus der partiellen Ableitung von $D_j f$ nach x_i als

$$\begin{aligned} G'_{s,t}(r) &= \lim_{h \rightarrow 0} \frac{G_{s,t}(r+h) - G_{s,t}(r)}{h} \\ &= \lim_{h \rightarrow 0} \frac{D_j f(\mathbf{a} + (r+h)e_i + \theta_{s,t}e_j) - D_j f(\mathbf{a} + re_i + \theta_{s,t}e_j)}{h} \\ &= \lim_{h \rightarrow 0} \frac{D_j f(\mathbf{a} + r \cdot e_i + \theta_{s,t}e_j + he_i) - D_j f(\mathbf{a} + re_i + \theta_{s,t}e_j)}{h} \\ &= D_i D_j f(\mathbf{a} + re_i + \theta_{s,t}e_j). \end{aligned}$$

Wenden wir nun wieder den Mittelwertsatz an, so finden wir zu jedem $s = r \in (-\delta, \delta)$ ein $\sigma_{s,t} \in \mathbb{R}$ mit $|\sigma_{s,t}| < |s|$ und

$$G_{s,t}(s) - G_{s,t}(0) = s \cdot G'_{s,t}(\sigma_{s,t}) = s \cdot D_i D_j f(\mathbf{a} + \sigma_{s,t}e_i + \theta_{s,t}e_j). \quad (142)$$

Für $0 \neq s, t \in (-\delta, \delta)$ ergibt sich dann aber

$$\begin{aligned} H(s, t) &= \frac{F_s(t) - F_s(0)}{s \cdot t} \\ &\stackrel{(141)}{=} \frac{D_j f(\mathbf{a} + se_i + \theta_{s,t}e_j) - D_j f(\mathbf{a} + \theta_{s,t}e_j)}{s} \\ &\stackrel{(142)}{=} D_i D_j f(\mathbf{a} + \sigma_{s,t}e_i + \theta_{s,t}e_j). \end{aligned}$$

Lassen wir nun $(s, t)^t$ in der euklidischen Norm gegen $(0, 0)^t$ konvergieren, konvergiert auch $(\sigma_{s,t}, \theta_{s,t})^t$ gegen $(0, 0)^t$, da $|\sigma_{s,t}| < |s|$ und $|\theta_{s,t}| < |t|$. Aus der Stetigkeit der zweiten partiellen Ableitung $D_i D_j f$ erhalten wir deshalb

$$\lim_{(s,t) \rightarrow (0,0)} H(s, t) = D_i D_j f(\mathbf{a}).$$

Bei den obigen Überlegungen können wir die Rollen von e_i und e_j aber auch austauschen und erhalten dann ein $|\sigma'_{s,t}| < |s|$ und ein $|\theta'_{s,t}| < |t|$ mit

$$H(s, t) = D_j D_i f(\mathbf{a} + \sigma'_{s,t} e_i + \theta'_{s,t} e_j),$$

so daß mit dem gleichen Argument auch

$$\lim_{(s,t) \rightarrow (0,0)} H(s, t) = D_j D_i f(\mathbf{a})$$

folgt. Damit ist insbesondere die Gleichheit

$$D_i D_j f(\mathbf{a}) = D_j D_i f(\mathbf{a})$$

gezeigt. □

Beispiel 42.6 (Satz von Schwarz)

Im Satz von Schwarz ist die Voraussetzung, daß die zweiten partiellen Ableitungen stetig sind, wesentlich. Für die Funktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : \mathbf{x} \mapsto \begin{cases} \frac{x_1^3 x_2 - x_1 x_2^3}{x_1^2 + x_2^2}, & \text{falls } (x_1, x_2)^t \neq (0, 0)^t, \\ 0, & \text{sonst} \end{cases}$$

existieren die beiden partiellen Ableitungen $D_1 D_2 f(0, 0)$ und $D_2 D_1 f(0, 0)$, stimmen aber nicht überein (siehe Aufgabe 42.29).

Korollar 42.7 (Schwarz)

Ist $f \in C^k(\mathbf{U}, \mathbb{R}^m)$, $\sigma \in \mathbb{S}_k$ und $j_1, \dots, j_k \in \{1, \dots, n\}$, so gilt für alle $\mathbf{a} \in \mathbf{U}$

$$D_{j_1} \dots D_{j_k} f(\mathbf{a}) = D_{j_{\sigma(1)}} \dots D_{j_{\sigma(k)}} f(\mathbf{a}),$$

d.h. beim partiellen Ableiten kommt es nicht auf die Reihenfolge an.

Beweis: Man beachte zunächst, daß aus der Anwendung des Satzes von Schwarz 42.5 auf $D_{j_{l+2}} \dots D_{j_k} f$ die Gleichheit

$$D_{j_1} \dots D_{j_l} D_{j_{l+1}} D_{j_{l+2}} \dots D_{j_k} f(\mathbf{a}) = D_{j_1} \dots D_{j_{l+1}} D_{j_l} D_{j_{l+2}} \dots D_{j_k} f(\mathbf{a})$$

folgt, d.h. man kann zwei beliebige benachbarte partielle Ableitungen vertauschen — hierbei beachte man, daß die Abbildung $D_{j_{l+2}} \dots D_{j_k} f$ mindestens zweifach stetig differenzierbar ist.

Da sich die Permutation σ nach Satz 29.7 als Produkt von Transpositionen benachbarter Zahlen schreiben läßt, folgt die Aussage also durch wiederholte Anwendung des Satzes von Schwarz 42.5. □

B) Der Satz von Taylor

Notation 42.8

Wir haben bei mehrfachen partiellen Ableitungen oben die Notation D_j für die Ableitung nach x_j verwendet. Wir hätten statt dessen in unserer bisherigen Notation auch

$$D_i D_j f(\mathbf{a}) = \frac{\partial}{\partial x_i} \frac{\partial f}{\partial x_j}(\mathbf{a})$$

schreiben können. Das ist zu umständlich. Gebräuchlicher ist dann die Notation

$$D_i D_j f(\mathbf{a}) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{a}),$$

wobei man bei doppelter Ableitung nach der gleichen Variablen dann auch

$$D_i^2 f(\mathbf{a}) = \frac{\partial^2 f}{\partial x_i^2}(\mathbf{a})$$

schreibt. Diese Notation verallgemeinert sich dann in naheliegender Weise zu

$$D^\alpha f(\mathbf{a}) = D_1^{\alpha_1} \dots D_n^{\alpha_n} f(\mathbf{a}) = \frac{\partial^{\alpha_1 + \dots + \alpha_n} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}(\mathbf{a})$$

für $\alpha = (\alpha_1, \dots, \alpha_n)^t \in \mathbb{N}^n$. Hierbei machen wir uns schon zunutze, daß wir beim partiellen Differenzieren die Reihenfolge vertauschen können, so daß wir die Ableitungen nach den Variablen x_1, \dots, x_n in aufsteigender Reihenfolge sortieren können, ohne dabei an Allgemeinheit einzubüßen, sofern f nur hinreichend oft stetig differenzierbar ist.

Offenbar benötigt man im Zusammenhang mit höheren Ableitungen auf natürliche Weise wieder *Multiindizes*. Wir wollen deshalb die Multiindexnotation, die wir im Zusammenhang mit Polynomen in Beispiel 40.14 bereits eingeführt haben, nochmals in Erinnerung rufen und erweitern. Für $\alpha = (\alpha_1, \dots, \alpha_n)^t \in \mathbb{N}^n$, $\mathbf{t} = (t_1, \dots, t_n)^t$ und $\mathbf{a} = (a_1, \dots, a_n)^t \in \mathbb{R}^n$ haben wir bereits

$$|\alpha| = \alpha_1 + \dots + \alpha_n$$

und

$$(\mathbf{t} - \mathbf{a})^\alpha = (t_1 - a_1)^{\alpha_1} \cdot \dots \cdot (t_n - a_n)^{\alpha_n}$$

eingeführt. Außerdem definieren wir nun die *Fakultät*

$$\alpha! = \alpha_1! \cdot \dots \cdot \alpha_n!$$

als Produkt der Fakultäten der Komponenten α_i und den *Binomialkoeffizienten*

$$\binom{\alpha}{\beta} = \frac{\alpha!}{\beta! \cdot (\alpha - \beta)!} = \frac{\alpha_1! \cdot \dots \cdot \alpha_n!}{\beta_1! \cdot \dots \cdot \beta_n! \cdot (\alpha_1 - \beta_1)! \cdot \dots \cdot (\alpha_n - \beta_n)!} = \binom{\alpha_1}{\beta_1} \cdot \dots \cdot \binom{\alpha_n}{\beta_n}$$

für $\alpha, \beta \in \mathbb{N}^n$ mit $\beta \leq \alpha$, d.h. $0 \leq \beta_j \leq \alpha_j$ für alle $j = 1, \dots, n$.

Bemerkung 42.9 (Das Taylorpolynom einer Polynomfunktion)

Man sieht dann leicht, daß

$$D^\beta x^\alpha = \beta! \cdot \binom{\alpha}{\beta} \cdot x^{\alpha-\beta}$$

gilt für $\alpha, \beta \in \mathbb{N}^n$ mit $\beta \leq \alpha$. Insbesondere gilt also

$$D^\alpha x^\alpha = \alpha!.$$

Ist nun

$$p : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto \sum_{|\alpha|=0}^d a_\alpha \cdot x^\alpha$$

eine Polynomfunktion, so gilt

$$D^\beta p(0) = \sum_{|\alpha|=0}^d a_\alpha \cdot D^\beta x^\alpha|_{x=0} = a_\beta \cdot \beta!$$

und mithin gilt

$$p(x) = \sum_{|\alpha|=0}^d \frac{D^\alpha p(0)}{\alpha!} \cdot x^\alpha = \sum_{|\alpha|=0}^d \frac{D^\alpha p(0)}{\alpha!} \cdot (x - 0)^\alpha.$$

Dies motiviert die folgende Definition des Taylorpolynoms im Mehrdimensionalen, dessen Ziel es ist, eine gegebene Funktion möglichst gut durch eine Polynomfunktion zu approximieren.

Definition 42.10 (Taylorpolynome)

Für eine k -fach stetig differenzierbare Funktion $f : U \longrightarrow \mathbb{R}$ und $a \in U$ heißt

$$T_{f,a}^k := \sum_{|\alpha|=0}^k \frac{D^\alpha f(a)}{\alpha!} \cdot (t - a)^\alpha = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| \leq k}} \frac{D_1^{\alpha_1} \dots D_n^{\alpha_n} f(a)}{\alpha_1! \cdot \dots \cdot \alpha_n!} \cdot (t_1 - a_1)^{\alpha_1} \cdot \dots \cdot (t_n - a_n)^{\alpha_n}$$

das k -te Taylorpolynom von f mit Entwicklungspunkt a .

Gilt sogar $f \in C^\infty(U, \mathbb{R})$, so nennen wir

$$T_{f,a} := \sum_{|\alpha|=0}^\infty \frac{D^\alpha f(a)}{\alpha!} \cdot (t - a)^\alpha = \sum_{\alpha \in \mathbb{N}^n} \frac{D_1^{\alpha_1} \dots D_n^{\alpha_n} f(a)}{\alpha_1! \cdot \dots \cdot \alpha_n!} \cdot (t_1 - a_1)^{\alpha_1} \cdot \dots \cdot (t_n - a_n)^{\alpha_n}$$

die Taylorreihe von f mit Entwicklungspunkt a .

Wir führen außerdem für zwei Punkte $a, x \in \mathbb{R}^n$ die Bezeichnung

$$\overline{ax} := \{a + t \cdot (x - a) \mid 0 \leq t \leq 1\} \subset \mathbb{R}^n$$

für die Strecke im \mathbb{R}^n zwischen a und x ein.

Bemerkung 42.11

Ist $f : U \longrightarrow \mathbb{R}$ stetig differenzierbar und $a \in U$, so ist das erste Taylorpolynom

$$T_{f,a}^1(x) = f(a) + \sum_{i=1}^n D_i f(a) \cdot (x_i - a_i) = f(a) + \langle \nabla f(a), x - a \rangle$$

die lineare Approximation von f in a , die wir in Bemerkung 41.27 betrachtet haben.

Beispiel 42.12 (Taylorpolynom)

Die Funktion

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R}: x \mapsto x_2 \cdot \cos(x_1)$$

ist unendlich oft stetig differenzierbar und das 2-te Taylorpolynom im Entwicklungspunkt $\mathbf{a} = (0, 0)^t$ ist

$$\begin{aligned} T_{f,\mathbf{a}}^2 &= f(\mathbf{a}) + \frac{D_1 f(\mathbf{a})}{1! \cdot 0!} \cdot (t_1 - a_1) + \frac{D_2 f(\mathbf{a})}{0! \cdot 1!} \cdot (t_2 - a_2) \\ &\quad + \frac{D_1^2 f(\mathbf{a})}{2! \cdot 0!} \cdot (t_1 - a_1)^2 + \frac{D_1 D_2 f(\mathbf{a})}{1! \cdot 1!} \cdot (t_1 - a_1) \cdot (t_2 - a_2) + \frac{D_2^2 f(\mathbf{a})}{0! \cdot 2!} \cdot (t_2 - a_2)^2 \\ &= 0 \cdot \cos(0) - 0 \cdot \sin(0) \cdot t_1 + \cos(0) \cdot t_2 - \frac{0 \cdot \cos(0)}{2} \cdot t_1^2 - \sin(0) \cdot t_1 \cdot t_2 + \frac{0}{2} \cdot t_2^2 = t_2. \end{aligned}$$

Will man die Taylorreihe von f mit Entwicklungspunkt $\mathbf{a} = (0, 0)^t$ bestimmen, so kann man sich in diesem konkreten Beispiel auch die Potenzreihenentwicklung des Cosinus zunutze machen und erhält

$$T_{f,\mathbf{a}} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \cdot t_1^{2n} \cdot t_2.$$

Die Kettenregel 41.22 erlaubt es, den Satz von Taylor 18.31 bzw. 20.14 ins Mehrdimensionale zu verallgemeinern.

Satz 42.13 (Satz von Taylor)

Ist $f \in \mathcal{C}^{k+1}(\mathbf{U}, \mathbb{R})$ und $\mathbf{a}, \mathbf{x} \in \mathbf{U}$ mit $\overline{\mathbf{a}\mathbf{x}} \subseteq \mathbf{U}$, dann gibt es ein $\mathbf{c} \in \overline{\mathbf{a}\mathbf{x}}$ mit

$$f(\mathbf{x}) - T_{f,\mathbf{a}}^k(\mathbf{x}) = \sum_{|\alpha|=k+1} \frac{D^\alpha f(\mathbf{c})}{\alpha!} \cdot (\mathbf{x} - \mathbf{a})^\alpha.$$

Die rechte Seite heißt auch Restglied des k -ten Taylorpolynoms nach Lagrange.

In Integralform gilt für das Restglied zudem

$$f(\mathbf{x}) - T_{f,\mathbf{a}}^k(\mathbf{x}) = (k+1) \cdot \int_0^1 (1-t)^k \cdot \sum_{|\alpha|=k+1} \frac{D^\alpha f(\mathbf{a} + t \cdot (\mathbf{x} - \mathbf{a}))}{\alpha!} \cdot (\mathbf{x} - \mathbf{a})^\alpha dt.$$

Der wesentliche Schritt im Beweis der mehrdimensionalen Taylorformel ist das folgende Lemma, das eine einfache Anwendung der Kettenregel darstellt.

Lemma 42.14

Es sei $f \in \mathcal{C}^{k+1}(\mathbf{U}, \mathbb{R})$ und $\mathbf{a}, \mathbf{x} \in \mathbf{U}$ mit $\overline{\mathbf{a}\mathbf{x}} \subseteq \mathbf{U}$. Dann ist die Funktion

$$\mathbf{h}: [0, 1] \longrightarrow \mathbb{R}: t \mapsto f(\mathbf{a} + t \cdot (\mathbf{x} - \mathbf{a}))$$

$k+1$ -fach stetig differenzierbar auf $[0, 1]$ und für $0 \leq l \leq k+1$ gilt

$$\frac{\mathbf{h}^{(l)}(t)}{l!} = \sum_{|\alpha|=l} \frac{D^\alpha f(\mathbf{a} + t \cdot (\mathbf{x} - \mathbf{a}))}{\alpha!} \cdot (\mathbf{x} - \mathbf{a})^\alpha. \quad (143)$$

Beweis: Die Funktion $\mathbf{h} = f \circ g$ ist als Verkettung der $k+1$ -fach stetig differenzierbaren Funktion f mit der unendlich oft differenzierbaren Funktion

$$g: [0, 1] \longrightarrow \mathbf{U}: t \mapsto \mathbf{a} + t \cdot (\mathbf{x} - \mathbf{a})$$

$k + 1$ -fach stetig differenzierbar. Wir wollen nun mit abbrechender Induktion nach l zeigen, daß die l -te Ableitung von h die Gleichung (143) für $0 \leq l \leq k + 1$ erfüllt.

Für $l = 0$ gilt dabei

$$\frac{h^{(0)}(t)}{0!} = h(t) = f(a + t \cdot (x - a)) = \sum_{|\alpha|=0} \frac{D^{\alpha}f(a + t \cdot (x - a))}{\alpha!} \cdot (x - a)^{\alpha},$$

da $\alpha = (0, \dots, 0)$ die einzige Möglichkeit für $|\alpha| = 0$ ist.

Sei also $l > 0$. Die Kettenregel liefert für die Funktion

$$F_{\alpha} : [0, 1] \longrightarrow \mathbb{R} : t \mapsto (D^{\alpha}f \circ g)(t)$$

die Ableitung

$$\begin{aligned} \frac{\partial F_{\alpha}}{\partial t}(t) &= D(D^{\alpha}f)(g(t)) \circ Dg(t) = D(D^{\alpha}f)(g(t)) \circ (x - a) \\ &= \sum_{j=1}^n D_j D^{\alpha}f(g(t)) \cdot (x_j - a_j) = \sum_{j=1}^n D^{\alpha+e_j} f(a + t \cdot (x - a)) \cdot (x_j - a_j). \end{aligned} \tag{144}$$

Dann folgt mit Hilfe von Induktion

$$\begin{aligned} \frac{h^{(l)}(t)}{l!} &= \frac{1}{l} \cdot \frac{\partial \frac{h^{(l-1)}}{(l-1)!}}{\partial t}(t) \\ &\stackrel{\text{Ind.}}{=} \frac{1}{l} \cdot \frac{\partial \sum_{|\alpha|=l-1} \frac{D^{\alpha}f(a+t \cdot (x-a))}{\alpha!} \cdot (x-a)^{\alpha}}{\partial t} \\ &= \frac{1}{l} \cdot \sum_{|\alpha|=l-1} \frac{\partial D^{\alpha}f(a+t \cdot (x-a))}{\partial t} \cdot \frac{(x-a)^{\alpha}}{\alpha!} \\ &= \frac{1}{l} \cdot \sum_{|\alpha|=l-1} \frac{\partial F_{\alpha}(t)}{\partial t} \cdot \frac{(x-a)^{\alpha}}{\alpha!} \\ &\stackrel{(144)}{=} \frac{1}{l} \cdot \sum_{|\alpha|=l-1} \sum_{j=1}^n D^{\alpha+e_j} f(a+t \cdot (x-a)) \cdot \frac{(x-a)^{\alpha+e_j}}{\alpha!} \\ &= \sum_{|\alpha|=l-1} \sum_{j=1}^n D^{\alpha+e_j} f(a+t \cdot (x-a)) \cdot \frac{(x-a)^{\alpha+e_j}}{(\alpha+e_j)!} \cdot \frac{\alpha_j+1}{l} \\ &\stackrel{!}{=} \sum_{|\beta|=l} \frac{D^{\beta}f(a+t \cdot (x-a))}{\beta!} \cdot (x-a)^{\beta}. \end{aligned}$$

Um die letzte Gleichheit zu sehen, beachten wir folgendes. Ist $\beta \in \mathbb{N}^n$ mit $|\beta| = l$ gegeben und setzen wir für $j \in \{1, \dots, n\}$

$$\alpha^j = (\beta_1, \dots, \beta_{j-1}, \beta_j - 1, \beta_{j+1}, \dots, \beta_n),$$

so gilt

$$\beta = \alpha^j + e_j$$

und entweder $\alpha^j \in \mathbb{N}^n$ mit $|\alpha^j| = l - 1$ oder $\beta_j = 0$. Dies sind die einzigen Möglichkeiten, β als Summe $\beta = \alpha + \gamma$ mit $|\alpha| = l - 1$ zu schreiben, und für alle diese Fälle ist

$$D^{\alpha^j + e_j} f(\mathbf{a} + \mathbf{t} \cdot (\mathbf{x} - \mathbf{a})) \cdot \frac{(\mathbf{x} - \mathbf{a})^{\alpha^j + e_j}}{(\alpha^j + e_j)!} = D^\beta f(\mathbf{a} + \mathbf{t} \cdot (\mathbf{x} - \mathbf{a})) \cdot \frac{(\mathbf{x} - \mathbf{a})^\beta}{\beta!}$$

gleich, und summieren wir sie auf, so erhalten wir

$$\begin{aligned} \sum_{j=1}^n D^{\alpha^j + e_j} f(\mathbf{a} + \mathbf{t} \cdot (\mathbf{x} - \mathbf{a})) \cdot \frac{(\mathbf{x} - \mathbf{a})^{\alpha^j + e_j}}{(\alpha^j + e_j)!} \cdot \frac{\alpha_j^j + 1}{l} &= D^\beta f(\mathbf{a} + \mathbf{t} \cdot (\mathbf{x} - \mathbf{a})) \cdot \frac{(\mathbf{x} - \mathbf{a})^\beta}{\beta!} \cdot \sum_{j=1}^n \frac{\beta_j}{l} \\ &= D^\beta f(\mathbf{a} + \mathbf{t} \cdot (\mathbf{x} - \mathbf{a})) \cdot \frac{(\mathbf{x} - \mathbf{a})^\beta}{\beta!} \cdot \frac{\sum_{j=1}^n \beta_j}{l} = D^\beta f(\mathbf{a} + \mathbf{t} \cdot (\mathbf{x} - \mathbf{a})) \cdot \frac{(\mathbf{x} - \mathbf{a})^\beta}{\beta!}. \end{aligned}$$

Damit ist (143) gezeigt. \square

Beweis von Satz 42.13: Wenden wir nun den Satz von Taylor 18.31 auf die Funktion

$$h : [0, 1] \longrightarrow \mathbb{R} : t \mapsto f(\mathbf{a} + \mathbf{t} \cdot (\mathbf{x} - \mathbf{a}))$$

an, so erhalten wir

$$\begin{aligned} f(\mathbf{x}) - T_{f,\mathbf{a}}^k(\mathbf{x}) &= f(\mathbf{x}) - \sum_{|\alpha|=0}^k \frac{D^\alpha f(\mathbf{a})}{\alpha!} \cdot (\mathbf{x} - \mathbf{a})^\alpha \\ &\stackrel{(143)}{=} h(1) - \sum_{l=0}^k \frac{h^{(l)}(0)}{l!} \cdot (1 - 0)^l \\ &\stackrel{18.31}{=} \frac{h^{(k+1)}(\theta)}{(k+1)!} \cdot (1 - 0)^{k+1} \\ &\stackrel{(143)}{=} \sum_{|\alpha|=k+1} \frac{D^\alpha f(\mathbf{c})}{\alpha!} \cdot (\mathbf{x} - \mathbf{a})^k \end{aligned}$$

für ein geeignetes $\theta \in (0, 1)$ und $\mathbf{c} = \mathbf{a} + \theta \cdot (\mathbf{x} - \mathbf{a}) \in \overline{\mathbf{a}\mathbf{x}}$.

Die Integralform des Restgliedes folgt mit dem gleichen Argument aus Korollar 20.14:

$$\begin{aligned} f(\mathbf{x}) - T_{f,\mathbf{a}}^k(\mathbf{x}) &= f(\mathbf{x}) - \sum_{|\alpha|=0}^k \frac{D^\alpha f(\mathbf{a})}{\alpha!} \cdot (\mathbf{x} - \mathbf{a})^\alpha \\ &\stackrel{(143)}{=} h(1) - \sum_{l=0}^k \frac{h^{(l)}(0)}{l!} \cdot (1 - 0)^l \\ &\stackrel{20.14}{=} \int_0^1 \frac{h^{(k+1)}(t)}{k!} \cdot (1 - t)^k dt \\ &\stackrel{(143)}{=} (k+1) \cdot \int_0^1 (1 - t)^k \cdot \sum_{|\alpha|=k+1} \frac{D^\alpha f(\mathbf{a} + \mathbf{t} \cdot (\mathbf{x} - \mathbf{a}))}{\alpha!} \cdot (\mathbf{x} - \mathbf{a})^\alpha dt. \end{aligned}$$

\square

Bemerkung 42.15 (Fehlerabschätzung mittels Taylorentwicklung)

- a. Ist $f \in \mathcal{C}(\mathbf{U}, \mathbb{R})$ und ist $\mathbf{x} \in \overline{\mathbf{U}_\varepsilon(\mathbf{a})} \subset \mathbf{U}$, so ergibt sich aus der Taylorformel 42.13 und der Dreiecksungleichung folgende Fehlerabschätzung für das Restglied

$$\|f(\mathbf{x}) - \mathbf{T}_{f,\mathbf{a}}^k(\mathbf{x})\|_2 \leq \sum_{|\alpha|=k+1} \frac{\max\{|D^\alpha f(\mathbf{c})| \mid \mathbf{c} \in \overline{\mathbf{U}_\varepsilon(\mathbf{a})}\}}{\alpha!} \cdot \|\mathbf{x} - \mathbf{a}\|_2^{k+1},$$

da $|(x - \mathbf{a})^\alpha| = |x_1 - a_1|^{\alpha_1} \cdots |x_n - a_n|^{\alpha_n} \leq \|x - \mathbf{a}\|_2^{\alpha_1} \cdots \|x - \mathbf{a}\|_2^{\alpha_n} = \|x - \mathbf{a}\|_2^{|\alpha|}$.
Damit folgt dann aber unmittelbar

$$\lim_{\mathbf{x} \rightarrow \mathbf{a}} \frac{f(\mathbf{x}) - \mathbf{T}_{f,\mathbf{a}}^k(\mathbf{x})}{\|\mathbf{x} - \mathbf{a}\|_2^k} = 0.$$

Man sagt auch, daß das Restglied für $\mathbf{x} \rightarrow \mathbf{a}$ mit Ordnung $k + 1$ gegen 0 konvergiert.

- b. Die Integralform des Restgliedes hat den Vorteil, daß sie ohne Änderung für Funktionen $f : \mathbf{U} \rightarrow \mathbb{R}^m$ übernommen werden kann, wenn das Integral dort den Vektor in \mathbb{R}^m bezeichnet, dessen Einträge die entsprechenden Integrale der Komponentenfunktionen sind. Die Lagrange-Form des Restgliedes benötigt hingegen für jede Komponente einen eigenen Zwischenpunkt \mathbf{c}_i .

Der Mittelwertsatz ist ein Spezialfall des Satzes von Taylor mit $k = 0$.

Korollar 42.16 (Mittelwertsatz)

Ist $f \in \mathcal{C}^1(\mathbf{U}, \mathbb{R})$ und $\mathbf{x}, \mathbf{a} \in \mathbf{U}$ mit $\overline{\mathbf{a}\mathbf{x}} \subseteq \mathbf{U}$, so gibt es ein $\mathbf{c} \in \overline{\mathbf{a}\mathbf{x}}$ mit

$$f(\mathbf{x}) - f(\mathbf{a}) = \langle \nabla f(\mathbf{c}), \mathbf{x} - \mathbf{a} \rangle = Df(\mathbf{c}) \circ (\mathbf{x} - \mathbf{a}).$$

Ferner gilt

$$f(\mathbf{x}) - f(\mathbf{a}) = \int_0^1 Df(\mathbf{a} + t \cdot (\mathbf{x} - \mathbf{a})) \circ (\mathbf{x} - \mathbf{a}) dt$$

Wir wollen nun den Mittelwertsatz anwenden, um zu zeigen, daß stetig differenzierbare Abbildungen auf kompakten konvexen Mengen Lipschitz-stetig sind.

Definition 42.17 (Konvexe Mengen)

Eine Teilmenge $C \subseteq \mathbb{R}^n$ heißt *konvex*, falls für je zwei Punkte $\mathbf{x}, \mathbf{y} \in C$ auch die Verbindungsstrecke $\overline{\mathbf{x}\mathbf{y}}$ in C enthalten ist.

Beispiel 42.18 (Eine kompakte konvexe Menge)

Eine abgeschlossene ε -Umgebung

$$\overline{\mathbf{U}_\varepsilon(\mathbf{a})} = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\| \leq \varepsilon\}$$

in \mathbb{R}^n bezüglich einer beliebigen Norm $\|\cdot\|$ ist stets konvex und kompakt.

Für das folgende Korollar identifizieren wir den \mathbb{R} -Vektorraum $\text{Mat}(\mathbf{m} \times \mathbf{n}, \mathbb{R})$ der $\mathbf{m} \times \mathbf{n}$ -Matrizen mit $\mathbb{R}^{\mathbf{m} \cdot \mathbf{n}}$, so daß wir für $\mathbf{A} = (a_{ij}) \in \text{Mat}(\mathbf{m} \times \mathbf{n}, \mathbb{R})$ die euklidische Norm

$$\|\mathbf{A}\|_2 = \sqrt{\sum_{i=1}^{\mathbf{m}} \sum_{j=1}^{\mathbf{n}} a_{ij}^2}$$

haben und auch die Maximumsnorm

$$\|A\|_\infty = \max \{ |a_{ij}| \mid i = 1, \dots, m, j = 1, \dots, n \}.$$

Korollar 42.19 (Lipschitz-Stetigkeit auf kompakten konvexen Mengen)

Sei $U \subseteq \mathbb{R}^n$ offen, $C \subseteq U$ konvex und kompakt, $f : U \rightarrow \mathbb{R}^m$ stetig differenzierbar.

Für $x, y \in C$ gilt dann stets

$$\|f(x) - f(y)\|_2 \leq \|x - y\|_2 \cdot n^2 \cdot \sup_{z \in C} \|Df(z)\|_2$$

und

$$\|f(x) - f(y)\|_\infty \leq \|x - y\|_\infty \cdot n \cdot \sup_{z \in C} \|Df(z)\|_\infty.$$

Insbesondere ist eine stetig-differenzierbare Abbildung auf einer kompakten konvexen Menge im \mathbb{R}^n also Lipschitz-stetig.

Beweis: Aus dem Mittelwertsatz 42.16 angewendet auf die Komponentenfunktion f_i von f folgt

$$f_i(x) - f_i(y) = Df_i(c_i) \circ (x - y) = \sum_{j=1}^n D_j f_i(c_i) \cdot (x_j - y_j)$$

für ein geeignetes $c_i \in \overline{xy}$. Mit der Dreiecksungleichung erhalten wir dann

$$\begin{aligned} |f_i(x) - f_i(y)| &\leq \sum_{j=1}^n |D_j f_i(c_i)| \cdot |x_j - y_j| \leq \sum_{j=1}^n \|Df(c_i)\|_\infty \cdot \|x - y\|_\infty \\ &\leq n \cdot \max_{z \in C} \|Df(z)\|_\infty \cdot \|x - y\|_\infty, \end{aligned}$$

wobei wir beachten, daß die stetige Funktion $x \mapsto \|Df(x)\|_\infty$ auf der kompakten Menge C ihr Maximum annimmt. Wir erhalten daraus zunächst

$$\|f(x) - f(y)\|_\infty = \max_{i=1, \dots, n} |f_i(x) - f_i(y)| \leq n \cdot \max_{z \in C} \|Df(z)\|_\infty \cdot \|x - y\|_\infty.$$

Aus Beispiel 39.42 wissen wir zudem, daß die euklidische Norm durch das \sqrt{n} -Fache der Maximumsnorm beschränkt ist und daß umgekehrt die Maximumsnorm stets durch die euklidische Norm beschränkt ist. Damit erhalten wir

$$\begin{aligned} \|f(x) - f(y)\|_2 &\leq \sqrt{n} \cdot \|f(x) - f(y)\|_\infty \leq \sqrt{n} \cdot n \cdot \max_{z \in C} \|Df(z)\|_\infty \cdot \|x - y\|_\infty \\ &\leq n^2 \cdot \max_{z \in C} \|Df(z)\|_2 \cdot \|x - y\|_2. \end{aligned}$$

□

Auch der Spezialfall des Satzes von Taylor für $k = 1$ ist von Interesse. Dazu führen wir folgenden Begriff ein.

Definition 42.20 (Hesse-Matrix)

Ist $f \in \mathcal{C}^2(\mathbf{U}, \mathbb{R})$ und $\mathbf{x} \in \mathbf{U}$, so nennen wir

$$H_f(\mathbf{x}) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}) \right)_{i,j=1,\dots,n} = \begin{pmatrix} D_1 D_1 f(\mathbf{x}) & D_1 D_2 f(\mathbf{x}) & \dots & D_1 D_n f(\mathbf{x}) \\ D_2 D_1 f(\mathbf{x}) & D_2 D_2 f(\mathbf{x}) & \dots & D_2 D_n f(\mathbf{x}) \\ \vdots & \vdots & & \vdots \\ D_n D_1 f(\mathbf{x}) & D_n D_2 f(\mathbf{x}) & \dots & D_n D_n f(\mathbf{x}) \end{pmatrix}$$

die *Hesse-Matrix* von f in \mathbf{x} . Man beachte, daß $H_f(\mathbf{x})$ aufgrund des Satzes von Schwarz 42.5 eine *symmetrische* Matrix ist.

Beispiel 42.21

Für die Funktion f aus Beispiel 42.1 gilt

$$H_f(\mathbf{x}) = \begin{pmatrix} 60x_1^3 + 12x_1x_2 & 6x_1^2 \\ 6x_1^2 & 0 \end{pmatrix}.$$

Korollar 42.22 (Satz von Taylor für $k = 1$)

Ist $f \in \mathcal{C}^2(\mathbf{U}, \mathbb{R})$ und $\mathbf{x}, \mathbf{a} \in \mathbf{U}$ mit $\overline{\mathbf{a}\mathbf{x}} \subseteq \mathbf{U}$, so gibt es ein $\mathbf{c} \in \overline{\mathbf{a}\mathbf{x}}$ mit

$$f(\mathbf{x}) = f(\mathbf{a}) + Df(\mathbf{a}) \circ (\mathbf{x} - \mathbf{a}) + \frac{(\mathbf{x} - \mathbf{a})^t \circ H_f(\mathbf{c}) \circ (\mathbf{x} - \mathbf{a})}{2}.$$

Beweis: Es gilt

$$\begin{aligned} \sum_{|\alpha|=2} \frac{D^\alpha f(\mathbf{c})}{\alpha!} \cdot (\mathbf{x} - \mathbf{a})^\alpha &= \sum_{i=1}^n \frac{D_i^2 f(\mathbf{c})}{2} \cdot (x_i - a_i)^2 + \sum_{i=1}^n \sum_{j < i} \frac{D_i D_j f(\mathbf{c})}{1 \cdot 1} \cdot (x_i - a_i) \cdot (x_j - a_j) \\ &= \sum_{i=1}^n \frac{D_i^2 f(\mathbf{c})}{2} \cdot (x_i - a_i)^2 + \sum_{i=1}^n \sum_{j \neq i} \frac{D_i D_j f(\mathbf{c})}{2} \cdot (x_i - a_i) \cdot (x_j - a_j) \\ &= \frac{\sum_{i=1}^n \sum_{j=1}^n (x_j - a_j) \cdot D_i D_j f(\mathbf{c}) \cdot (x_i - a_i)}{2} \\ &= \frac{(\mathbf{x} - \mathbf{a})^t \circ H_f(\mathbf{c}) \circ (\mathbf{x} - \mathbf{a})}{2}. \end{aligned}$$

Der Rest folgt dann aus dem Satz von Taylor 42.13. □

C) Lokale Extrema im Mehrdimensionalen

Wir wollen nun die Bedingungen 18.2 und 18.13 für Extremstellen verallgemeinern.

Definition 42.23 (Lokale Extrema)

Sei $D \subseteq \mathbb{R}^n$, $f : D \rightarrow \mathbb{R}$ und $\mathbf{a} \in D$.

- f hat in \mathbf{a} ein *globales Maximum*, wenn $f(\mathbf{a}) \geq f(\mathbf{x})$ für alle $\mathbf{x} \in D$.
- f hat in \mathbf{a} ein *lokales Maximum*, wenn es ein $\delta > 0$ gibt, so daß $f(\mathbf{a}) \geq f(\mathbf{x})$ für alle $\mathbf{x} \in D \cap \mathbf{U}_\delta(\mathbf{a})$.
- f hat in \mathbf{a} ein *isoliertes lokales Maximum*, wenn es ein $\delta > 0$ gibt mit $f(\mathbf{a}) > f(\mathbf{x})$ für alle $\mathbf{a} \neq \mathbf{x} \in D \cap \mathbf{U}_\delta(\mathbf{a})$.
- f hat in \mathbf{a} ein *globales Minimum*, wenn $f(\mathbf{a}) \leq f(\mathbf{x})$ für alle $\mathbf{x} \in D$.

- e. f hat in \mathbf{a} ein *lokales Minimum*, wenn es ein $\delta > 0$ gibt, so daß $f(\mathbf{a}) \leq f(\mathbf{x})$ für alle $\mathbf{x} \in \mathbf{D} \cap \mathbf{U}_\delta(\mathbf{a})$.
- f. f hat in \mathbf{a} ein *isoliertes lokales Minimum*, wenn es ein $\delta > 0$ gibt mit $f(\mathbf{a}) < f(\mathbf{x})$ für alle $\mathbf{a} \neq \mathbf{x} \in \mathbf{D} \cap \mathbf{U}_\delta(\mathbf{a})$.
- g. \mathbf{a} heißt *Extremstelle* und $f(\mathbf{a})$ *Extremum* von f , wenn f in \mathbf{a} ein lokales Minimum oder ein lokales Maximum hat.
- h. Ist f in \mathbf{a} differenzierbar mit $Df(\mathbf{a}) = (0, \dots, 0)$, so heißt \mathbf{a} ein *kritischer Punkt* von f .
- i. Ist \mathbf{a} ein kritischer Punkt von f und gibt es in jeder ε -Umgebung von \mathbf{a} Punkte \mathbf{x} und \mathbf{y} mit $f(\mathbf{x}) > f(\mathbf{a}) > f(\mathbf{y})$, so heißt \mathbf{a} ein *Sattelpunkt* von f (siehe Abbildung 7 auf Seite 560).

Proposition 42.24 (Extremstellen sind kritische Punkte)

Ist $f \in \mathcal{C}^1(\mathbf{U}, \mathbb{R})$ und ist $\mathbf{a} \in \mathbf{U}$ eine Extremstelle von f , so ist $Df(\mathbf{a}) = (0, \dots, 0)$.

Beweis: Da \mathbf{U} offen ist, gibt es ein $\varepsilon > 0$ mit $\mathbf{U}_\varepsilon(\mathbf{a}) \subseteq \mathbf{U}$. Für $j \in \{1, \dots, n\}$ betrachten wir die Funktion

$$g_j : (-\varepsilon, \varepsilon) \longrightarrow \mathbb{R} : t \mapsto f(\mathbf{a} + t \cdot \mathbf{e}_j).$$

Besitzt f ein lokales Maximum bzw. Minimum in \mathbf{a} , so besitzt g_j ein lokales Maximum bzw. Minimum in 0 . Nach der notwendigen Bedingung für Extrema im Eindimensionalen, Proposition 18.2, ist dann

$$0 = g_j'(0) = \lim_{t \rightarrow 0} \frac{g_j(t) - g_j(0)}{t - 0} = \lim_{t \rightarrow 0} \frac{f(\mathbf{a} + t\mathbf{e}_j) - f(\mathbf{a})}{t} = D_j f(\mathbf{a}).$$

Dies zeigt, daß der Gradient von f der Nullvektor ist. □

Satz 42.25 (Hinreichende Bedingung für Extremstellen)

Es sei $f \in \mathcal{C}^2(\mathbf{U}, \mathbb{R})$ und $\mathbf{a} \in \mathbf{U}$ mit $Df(\mathbf{a}) = (0, \dots, 0)$.

- a. Ist die Hesse-Matrix $H_f(\mathbf{a})$ positiv definit, so hat f in \mathbf{a} ein isoliertes lokales Minimum.
- b. Ist die Hesse-Matrix $H_f(\mathbf{a})$ negativ definit, so hat f in \mathbf{a} ein isoliertes lokales Maximum.
- c. Ist die Hesse-Matrix $H_f(\mathbf{a})$ indefinit, so ist \mathbf{a} ein Sattelpunkt von f und keine Extremstelle.

Beweis:

- a. Für $k = 1, \dots, n$ ist die Funktion

$$g_k : \mathbf{U} \longrightarrow \mathbb{R} : \mathbf{x} \mapsto \det \begin{pmatrix} D_1 D_1 f(\mathbf{x}) & D_1 D_2 f(\mathbf{x}) & \dots & D_1 D_k f(\mathbf{x}) \\ D_2 D_1 f(\mathbf{x}) & D_2 D_2 f(\mathbf{x}) & \dots & D_2 D_k f(\mathbf{x}) \\ \vdots & \vdots & & \vdots \\ D_k D_1 f(\mathbf{x}) & D_k D_2 f(\mathbf{x}) & \dots & D_k D_k f(\mathbf{x}) \end{pmatrix}$$

stetig, da die Funktion f nach Voraussetzung zweifach stetig differenzierbar ist. Mithin sind die Mengen

$$g_k^{-1}((0, \infty))$$

als Urbilder offener Mengen unter stetigen Abbildungen nach Proposition 40.17 offen in \mathbf{U} für $k = 1, \dots, n$, und damit ist auch

$$O = \bigcup_{k=1}^n g_k^{-1}((0, \infty))$$

als Vereinigung offener Mengen wieder offen.

Man beachte nun, daß $g_k(\mathbf{x})$ gerade der k -te Hauptminor der Hesse-Matrix $H_f(\mathbf{x})$ ist. Aus dem Hurwitz-Kriterium für positive Definitheit 38.39 wissen wir, daß $g_k(\mathbf{a}) > 0$ für alle $k = 1, \dots, n$ gilt, da die symmetrische Matrix $H_f(\mathbf{a})$ positiv definit ist, und damit gilt also

$$\mathbf{a} \in O.$$

Da O offen ist, gibt es mithin ein $\delta > 0$, so daß

$$U_\delta(\mathbf{a}) \subseteq O.$$

Sei nun $\mathbf{a} \neq \mathbf{x} \in U_\delta(\mathbf{a})$ beliebig, so gilt

$$\overline{\mathbf{ax}} \subset U_\delta(\mathbf{a}) \subseteq O \subseteq \mathbf{U},$$

so daß wir den Satz von Taylor in der Fassung von Korollar 42.22 anwenden können. Wir erhalten

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{a}) + Df(\mathbf{a}) \circ (\mathbf{x} - \mathbf{a}) + \frac{(\mathbf{x} - \mathbf{a})^t \circ H_f(\mathbf{c}) \circ (\mathbf{x} - \mathbf{a})}{2} \\ &= f(\mathbf{a}) + \frac{(\mathbf{x} - \mathbf{a})^t \circ H_f(\mathbf{c}) \circ (\mathbf{x} - \mathbf{a})}{2} > f(\mathbf{a}), \end{aligned}$$

denn $\mathbf{c} \in \overline{\mathbf{ax}} \subset O$ und mithin hat $H_f(\mathbf{c})$ nur positive Hauptminoren, so daß $H_f(\mathbf{c})$ wegen des Hurwitz-Kriteriums 38.39 positiv definit ist.

- b. Ist $H_f(\mathbf{a})$ negativ definit, dann ist $H_{-f}(\mathbf{a}) = -H_f(\mathbf{a})$ positiv definit. Aus Teil a. folgt also, daß $-f$ ein isoliertes lokales Minimum in \mathbf{a} hat. Dann hat f dort aber ein isoliertes lokales Maximum.
- c. Ist $H_f(\mathbf{a})$ indefinit, so gibt es $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ mit

$$\mathbf{x}^t \circ H_f(\mathbf{a}) \circ \mathbf{x} > 0 > \mathbf{y}^t \circ H_f(\mathbf{a}) \circ \mathbf{y}. \quad (145)$$

Wir müssen zeigen, daß f in jeder ε -Umgebung von \mathbf{a} Werte größer und kleiner als $f(\mathbf{a})$ annimmt.

Sei also $\varepsilon > 0$ gegeben. Da \mathbf{U} offen ist und wir nur kleine ε betrachten müssen, können wir ohne Einschränkung annehmen, daß

$$U_\varepsilon(\mathbf{a}) \subseteq \mathbf{U}$$

gilt. Für $\delta := \min \left\{ \frac{\varepsilon}{\|\mathbf{x}\|_2}, \frac{\varepsilon}{\|\mathbf{y}\|_2} \right\} > 0$ und $\mathbf{t} \in (-\delta, \delta)$ gilt $\mathbf{a} + \mathbf{tx}, \mathbf{a} + \mathbf{ty} \in U_\varepsilon(\mathbf{a})$.

Wir betrachten nun zunächst die Funktion

$$g : (-\delta, \delta) \longrightarrow \mathbb{R} : t \mapsto x^t \circ H_f(\mathbf{a} + t\mathbf{x}) \circ x.$$

Sie ist stetig, da f zweifach stetig differenzierbar auf \mathbf{U} ist. Außerdem gilt $g(0) > 0$ wegen (145), so daß mit Aufgabe 14.30 ein $\gamma > 0$ existiert, so daß

$$0 < g(t) = x^t \circ H_f(\mathbf{a} + t\mathbf{x}) \circ x$$

für alle $t \in (-\gamma, \gamma)$. Wenden wir nun den Satz von Taylor in der Fassung von Korollar 42.22 auf \mathbf{a} und $\mathbf{a} + t\mathbf{x}$ für $t \in (-\gamma, \gamma)$ an, so finden wir ein \mathbf{c}_t auf der Strecke von \mathbf{a} nach $\mathbf{a} + t\mathbf{x}$, d.h. wir finden ein θ_t mit $\theta_t \in (-\gamma, \gamma)$ und $\mathbf{c}_t = \mathbf{a} + \theta_t \mathbf{x}$, mit

$$\begin{aligned} f(\mathbf{a} + t\mathbf{x}) &= f(\mathbf{a}) + Df(\mathbf{a}) \circ (\mathbf{a} + t\mathbf{x} - \mathbf{a}) + \frac{(\mathbf{a} + t\mathbf{x} - \mathbf{a})^t \circ H_f(\mathbf{c}_t) \circ (\mathbf{a} + t\mathbf{x} - \mathbf{a})}{2} \\ &= f(\mathbf{a}) + \frac{t^2}{2} \cdot x^t \circ H_f(\mathbf{a} + \theta_t \mathbf{x}) \circ x = f(\mathbf{a}) + \frac{t^2}{2} \cdot g(\theta_t) > f(\mathbf{a}), \end{aligned}$$

da $\theta_t \in (-\gamma, \gamma)$. Also gibt es in der ε -Umgebung von \mathbf{a} einen Punkt $\mathbf{a} + t\mathbf{x}$, so daß

$$f(\mathbf{a} + t\mathbf{x}) > f(\mathbf{a}).$$

Analog findet man ein Intervall $(-\beta, \beta)$, so daß $\mathbf{a} + t\mathbf{y} \in \mathbf{U}_\varepsilon(\mathbf{a})$ und

$$f(\mathbf{a} + t\mathbf{y}) < f(\mathbf{a})$$

für alle $t \in (-\beta, \beta)$. Damit ist gezeigt, daß \mathbf{a} ein Sattelpunkt und keine Extremstelle von f ist.

□

Bemerkung 42.26

- a. Man beachte, daß die drei in Satz 42.25 betrachteten Bedingungen *nicht* alle möglichen Fälle abdecken. Z.B. hat die Funktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto x_1^2$$

den kritischen Punkt $\mathbf{x} = (0, 0)^t$, da $Df(\mathbf{x}) = (2x_1, 0)$ ist, und sie hat in $\mathbf{x} = (0, 0)^t$ die Hesse-Matrix

$$H_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Diese ist weder positiv definit, noch negativ definit, noch indefinit. Der Satz erlaubt also keine Aussage in diesem Fall. Die Funktion hat in $(0, 0)^t$ aber ein globales Minimum, das allerdings nicht isoliert ist.

- b. Allgemein gilt festzuhalten, daß Satz 42.25 dann keine Aussage zu den Extremstellen von f macht, wenn $H_f(\mathbf{a})$ für den kritischen Punkt \mathbf{a} den Eigenwert 0 hat.

c. Die Definitheit einer symmetrischen Matrix überprüft man am Besten mit dem Hurwitz-Kriterium 38.39 durch Berechnen der Hauptminoren. Sei $A \in \text{Mat}_n(\mathbb{R})$, dann gilt:

- (i) A ist positiv definit $\iff \det(A(k)) > 0$ für alle $k = 1, \dots, n$.
- (ii) A ist negativ definit $\iff (-1)^k \cdot \det(A(k)) > 0$ für alle $k = 1, \dots, n$.
- (iii) Falls $\det(A) \neq 0$ gilt, dann ist A genau dann indefinit, wenn A weder positiv definit noch negativ definit ist.

Beispiel 42.27

a. Die Funktion

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto x_1^2 + x_2^2$$

hat den Gradienten

$$Df(x) = (2x_1, 2x_2),$$

so daß $x = (0, 0)^t$ der einzige kritische Punkt ist. In diesem Punkt ist die Hesse-Matrix

$$H_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$

positiv definit. Also ist der Punkt $(0, 0)^t$ ein isoliertes lokales Minimum von f (siehe Abbildung 6). Es ist in der Tat sogar ein globales Minimum, da $f(x) = x_1^2 + x_2^2 > 0 = f(0, 0)$ für alle $x \neq (0, 0)^t$.

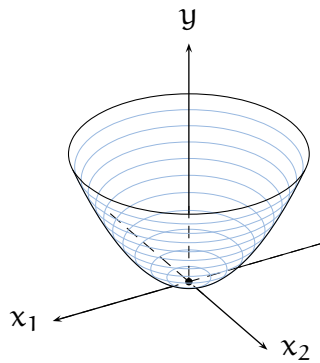


ABBILDUNG 6. Das isolierte Minimum von $x \mapsto x_1^2 + x_2^2$

b. Die Funktion

$$f: \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto x_1^2 - x_2^2$$

hat den Gradienten

$$Df(x) = (2x_1, -2x_2),$$

so daß wieder $x = (0, 0)^t$ der einzige kritische Punkt ist. In diesem Punkt ist die Hesse-Matrix

$$H_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

indefinit. Satz 42.25 sagt also, daß $(0, 0)^t$ keine Extremstelle ist, und Bemerkung 42.26 sagt etwas genauer, daß f in $(0, 0)^t$ einen Sattelpunkt hat (siehe Abbildung 7).

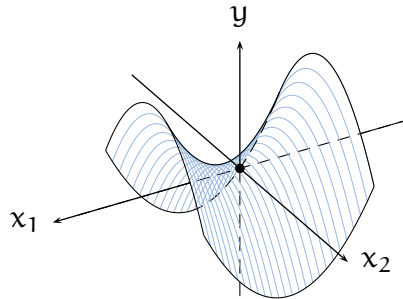


ABBILDUNG 7. Der Sattelpunkt von $x \mapsto x_1^2 - x_2^2$

Aufgaben

Aufgabe 42.28 (Vertauschbarkeit der Integrationsreihenfolge)

Es sei $U \subseteq \mathbb{R}^2$ offen, $f : U \rightarrow \mathbb{R}$ stetig und $[a, b] \times [c, d] \subset U$. Zeige die folgenden Aussagen:

- Die Funktion $H : [a, b] \times [c, d] \rightarrow \mathbb{R} : (x, y)^t \mapsto \int_a^x f(s, y) ds$ ist stetig.
- Falls $\frac{\partial f}{\partial y} : U \rightarrow \mathbb{R}$ existiert und stetig ist, so die Abbildung

$$F : [c, d] \rightarrow \mathbb{R} : y \mapsto \int_a^b f(s, y) ds$$

stetig differenzierbar auf $[c, d]$ mit

$$F'(y) = \int_a^b \frac{\partial f}{\partial y}(s, y) ds.$$

- Es gilt

$$\int_a^b \int_c^d f(s, t) dt ds = \int_c^d \int_a^b f(s, t) ds dt.$$

- Man leite aus c. und dem Hauptsatz der Differenzial- und Integralrechnung 20.4 den Satz von Schwarz her:

$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y).$$

Aufgabe 42.29 (Beispiel zum Satz von Schwarz)

Zeige, für die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} : x \mapsto \begin{cases} \frac{x_1^3 x_2 - x_1 x_2^3}{x_1^2 + x_2^2}, & \text{falls } (x_1, x_2)^t \neq (0, 0)^t, \\ 0, & \text{sonst} \end{cases}$$

existieren die beiden partiellen Ableitungen $D_1 D_2 f(0, 0)$ und $D_2 D_1 f(0, 0)$, stimmen aber nicht überein

Aufgabe 42.30 (Leibniz-Regel)

Sind $f, g \in \mathcal{C}^k(\mathcal{U}, \mathbb{R})$ und ist $|\alpha| \leq k$, so gilt für $x \in \mathcal{U}$

$$D^\alpha(f \cdot g)(x) = \sum_{\substack{\beta \in \mathbb{N}^n \\ \beta \leq \alpha}} \binom{\alpha}{\beta} \cdot D^\beta f(x) \cdot D^{\alpha-\beta} g(x).$$

Aufgabe 42.31 (Verbesserte Abschätzung für Propositiono 40.33)

a. Ist $g : [a, b] \rightarrow \mathbb{R}^m$ eine stetige Funktion und definieren wir das Integral

$$\int_a^b g(t) dt = \left(\int_a^b g_1(t) dt, \dots, \int_a^b g_n(t) dt \right)^t \in \mathbb{R}^n$$

als den Vektor der Integrale über die Koordinatenfunktionen, so gilt

$$\left\| \int_a^b g(t) dt \right\|_2 \leq \int_a^b \|g(t)\|_2 dt.$$

b. Ist $\mathcal{U} \subseteq \mathbb{R}^n$ offen, $\overline{\mathcal{U}_r(\mathbf{a})} \subseteq \mathcal{U}$ und $h : \mathcal{U} \rightarrow \mathbb{R}^m$ stetig differenzierbar auf \mathcal{U} , dann gilt für $x, y \in \overline{\mathcal{U}_r(\mathbf{a})}$ stets

$$\|h(x) - h(y)\|_2 \leq \|x - y\|_2 \cdot \sup_{z \in \overline{\mathcal{U}_r(\mathbf{a})}} \|Dh(z)\|_2.$$

Aufgabe 42.32

Bestimme alle lokalen Extrema und Sattelpunkte der Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y)^t \mapsto x^4 + 2x^2y^2 + y^4 - 2x^3 - 2xy^2 + x^2 + y^2.$$

Aufgabe 42.33

Bestimme das zweite Taylor-Polynom $T_{f, \mathbf{a}}^2$ im Entwicklungspunkt $\mathbf{a} = (0, \pi)^t$ für

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y)^t \mapsto \sin(y) - x(x + \sin(y) + 1).$$

Aufgabe 42.34

Sei $\mathbf{a} \in \mathbb{R}^n$, $\epsilon > 0$ und $f : \mathcal{U}_\epsilon(\mathbf{a}) \rightarrow \mathbb{R}^m$ eine differenzierbare Abbildung mit $Df(x) = 0$ für alle $x \in \mathcal{U}_\epsilon(\mathbf{a})$. Zeige, daß f konstant ist.

Aufgabe 42.35

Bestimme das sechste Taylor-Polynom $T_{f, \mathbf{a}}^6$ für $\mathbf{a} = (0, 0)^t$ und

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}, (x, y, z)^t \mapsto (x + y^3) \cdot \cos\left(\frac{z}{1 + y^2}\right).$$

§ 43 Der Satz über implizite Funktionen

A) Der Banachsche Fixpunktsatz

Definition 43.1 (Strikte Kontraktion)

Sei (M, d) ein metrischer Raum.

- Eine Lipschitz stetige Selbstabbildung $f : M \rightarrow M$ mit Lipschitzkonstante $q < 1$ heißt eine *strikte Kontraktion* oder kurz eine *q-Kontraktion*.
- $x \in M$ heißt ein *Fixpunkt* der Selbstabbildung $f : M \rightarrow M$, wenn $f(x) = x$.

Bemerkung 43.2 (Strikte Kontraktionen über \mathbb{R})

Eine differenzierbare Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ ist genau dann eine q-Kontraktion, wenn

$$|f'(x)| \leq q$$

für alle $x \in \mathbb{R}$, d.h. die Steigung in jedem Punkt ist durch $q < 1$ beschränkt.

Beweis: Ist f eine q-Kontraktion, so gilt

$$|f(x) - f(y)| \leq q \cdot |x - y|$$

für alle $x, y \in \mathbb{R}$. Dann ist aber

$$|f'(x)| = \lim_{y \rightarrow x} \frac{|f(y) - f(x)|}{|y - x|} \leq q$$

für alle $x \in \mathbb{R}$.

Gilt umgekehrt $|f'(c)| \leq q$ für alle $c \in \mathbb{R}$, so folgt für $x, y \in \mathbb{R}$ aus dem Mittelwertsatz 18.7

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

für ein geeignetes c zwischen x und y . Also ist f eine q-Kontraktion. \square

Beispiel 43.3 (Strikte Kontraktion)

Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto \frac{\sin(x)}{2}$ ist eine $\frac{1}{2}$ -Kontraktion, da

$$|f'(x)| = \frac{|\cos(x)|}{2} \leq \frac{1}{2}$$

für alle $x \in \mathbb{R}$. x ist genau dann ein Fixpunkt von f , wenn der Graph von f in $(x, f(x))$ die Winkelhalbierende schneidet (siehe Abbildung 8). Deshalb hat f genau einen Fixpunkt, nämlich $x = 0$.

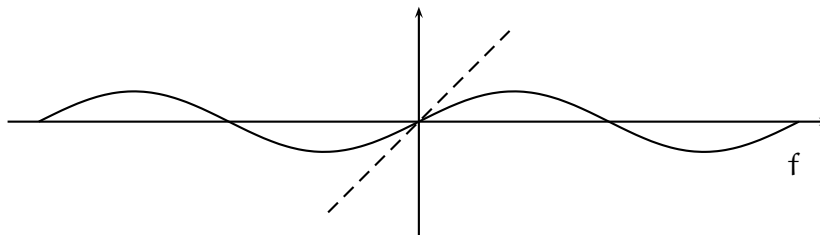


ABBILDUNG 8. Der Graph von $x \mapsto \frac{\cos(x)}{2}$

Der folgende Satz zeigt, daß es kein Zufall war, daß die Funktion f in Beispiel 43.3 genau einen Fixpunkt hat (siehe auch Aufgabe 14.36).

Satz 43.4 (Fixpunktsatz von Banach)

Sei (M, d) ein vollständiger metrischer Raum, $f : M \rightarrow M$ eine strikte Kontraktion. Dann hat f genau einen Fixpunkt x und für jedes $y \in M$ gilt $\lim_{n \rightarrow \infty} f^n(y) = x$.

Beweis: Wir zeigen zunächst die Existenz des Fixpunktes x von f . Dazu wählen wir ein beliebiges $y \in M$ und setzen

$$y_n := f^n(y).$$

Nach Voraussetzung gibt es ein $0 \leq q < 1$, so daß

$$d(f(u), f(v)) \leq q \cdot d(u, v) \quad (146)$$

für alle $u, v \in M$ gilt. Mit Hilfe von Induktion nach n leiten wir daraus her, daß

$$d(y_{n+1}, y_n) \leq q^n \cdot d(y_1, y_0). \quad (147)$$

Für den Induktionsanfang $n = 0$ ist nichts zu zeigen, und der Induktionsschritt $n - 1 \mapsto n$ folgt dann sofort:

$$\begin{aligned} d(y_{n+1}, y_n) &= d(f(y_n), f(y_{n-1})) \stackrel{(146)}{\leq} q \cdot d(y_n, y_{n-1}) \\ &\stackrel{\text{Ind.}}{\leq} q \cdot q^{n-1} \cdot d(y_1, y_0) = q^n \cdot d(y_1, y_0). \end{aligned}$$

Damit ist (147) gezeigt. Mit Induktion nach m folgt für ein festes $n \in \mathbb{N}$ nun daraus wiederum

$$d(y_{n+m}, y_n) \leq \sum_{k=n}^{n+m-1} q^k \cdot d(y_1, y_0). \quad (148)$$

Für den Induktionsanfang $m = 1$ ist (148) gerade (147), und für den Induktionsschritt $m - 1 \mapsto m$ folgt mit der Dreiecksungleichung sofort

$$\begin{aligned} d(y_{n+m}, y_n) &\stackrel{\text{DUGL}}{\leq} d(y_{n+m}, y_{n+m-1}) + d(y_{n+m-1}, y_n) \\ &\stackrel{\text{Ind.}/(147)}{\leq} q^{n+m-1} \cdot d(y_1, y_0) + \sum_{k=n}^{n+m-2} q^k \cdot d(y_1, y_0) \\ &= \sum_{k=n}^{n+m-1} q^k \cdot d(y_1, y_0) \end{aligned}$$

Die Ungleichung (148) führt zu folgender Abschätzung

$$d(y_{n+m}, y_n) \leq \sum_{k=n}^{n+m-1} q^k \cdot d(y_1, y_0) \leq q^n \cdot \sum_{k=0}^{\infty} q^k \cdot d(y_1, y_0) = \frac{q^n}{1-q} \cdot d(y_1, y_0)$$

Da die rechte Seite für $n \rightarrow \infty$ gegen Null konvergiert, sehen wir, daß die Folge $(y_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge ist, und da wir zudem vorausgesetzt haben, daß der metrische Raum M vollständig ist, ist die Folge also konvergent, d.h. der Grenzwert

$$x := \lim_{n \rightarrow \infty} y_n \in M$$

existiert. Nach Proposition 40.33 ist die Kontraktion f stetig, so daß aus der Eindeutigkeit des Grenzwertes

$$f(x) \longleftarrow f(y_n) = y_{n+1} \longrightarrow x$$

für $n \rightarrow \infty$ unmittelbar

$$f(x) = x$$

folgt, d.h. x ist ein Fixpunkt von f .

Es bleibt noch, die *Eindeutigkeit* des Fixpunktes zu zeigen. Nehmen wir dazu an, daß $z \in M$ ebenfalls ein Fixpunkt ist, so erhalten wir

$$d(x, z) = d(f(x), f(z)) \leq q \cdot d(x, z),$$

so daß $d(x, z) = 0$ wegen $q < 1$ gelten muß. Mithin ist $x = z$ und die Eindeutigkeit ist gezeigt. \square

B) Der Satz über implizite Funktionen

Motivation 43.5

Das wesentliche Ziel dieses Abschnitts besteht darin, einzelne Gleichungen oder ganze Gleichungssysteme nach einzelnen Variablen aufzulösen.

- a. Eine einzelne lineare Gleichung, z.B.

$$f(x) = 2x_1 + 4x_3 - x_4 = 0$$

kann man leicht nach jeder Variablen auflösen, die in der Gleichung vorkommt, z.B.

$$x_4 = 2x_1 + 4x_3.$$

Dies bedeutet, daß man die Lösungsmenge der Gleichung parametrisiert, z.B.

$$\begin{aligned} V(f) &:= \{x \in \mathbb{R}^4 \mid f(x) = 0\} = \{(x_1, x_2, x_3, 2x_1 + 4x_3)^t \mid x_1, x_2, x_3 \in \mathbb{R}\} \\ &= \text{Graph}(\varphi) = \text{Im}(\psi) \end{aligned}$$

mit

$$\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R} : (x_1, x_2, x_3)^t \mapsto 2x_1 + 4x_3$$

und

$$\psi : \mathbb{R}^3 \longrightarrow \mathbb{R}^4 : (x_1, x_2, x_3) \mapsto (x_1, x_2, x_3, 2x_1 + 4x_3)^t.$$

Man unterscheidet deshalb zwischen der *impliziten Darstellung* von $V(f)$ als Lösungsmenge einer Gleichung $f(x) = 0$ und der *expliziten Darstellung* von $V(f)$ als Graph einer Abbildung φ oder allgemeiner als Bild einer Parametrisierung ψ . Dabei wird die Abbildung φ durch die folgende Eigenschaft charakterisiert:

$$f(x_1, x_2, x_3, \varphi(x_1, x_2, x_3)) = 0.$$

- b. Ist $D \subseteq \mathbb{R}^n$ und $f : D \rightarrow \mathbb{R}^m$, so nennen wir die Lösungsmenge

$$V(f) = \{x \in D \mid f(x) = 0\}$$

der Gleichung $f(x) = 0$ *Verschwindungsmenge* oder *Nullstellenmenge* von f .

Ist $U \subseteq \mathbb{R}^p$ offen und $\psi : U \rightarrow \mathbb{R}^n$ mit $\text{Im}(\psi) = V(f)$, so nennen wir ψ eine *Parametrisierung* von $V(f)$.

Der *Graph* einer Funktion $\varphi : U \rightarrow \mathbb{R}^m$ mit $U \subseteq \mathbb{R}^n$ besitzt stets eine Parametrisierung, nämlich

$$\text{Graph}(\varphi) = \{(y, \varphi(y))^t \mid y \in U\} = \text{Im}(\psi)$$

für

$$\psi : U \rightarrow \mathbb{R}^{n+m} : y \mapsto (y, \varphi(y))^t.$$

Mit diesen Begriffsbildungen heißt, das Gleichungssystem $f(x) = 0$, d.h. $f_i(x_1, \dots, x_n) = 0$ für $i = 1, \dots, m$, nach den letzten m Variablen aufzulösen,

$$V(f) = \{x \in \mathbb{R}^n \mid x_{n-m+j} = \varphi_j(x_1, \dots, x_{n-m}) \quad j = 1, \dots, m\} = \text{Graph}(\varphi)$$

als den Graphen einer Abbildung φ zu schreiben und damit zu parametrisieren.

- c. Die Überlegungen in Teil a. funktionieren auch für ein ganzes lineares Gleichungssystem, wie wir in Bemerkung 28.9 gesehen haben. Wir können die Lösungsmengen eines lösbaren linearen Gleichungssystems

$$Ax = b$$

stets nach bestimmten Variablen auflösen und so parametrisieren. Um dies algorithmisch durchzuführen, bringt man die erweiterte Koeffizientenmatrix (A, b) auf reduzierte Zeilenstufenform und die Pivot-Spalten gehören dann zu Variablen, nach denen man das Gleichungssystem auflösen kann.

Nehmen wir der Einfachheit halber an, daß unser Gleichungssystem genau m Gleichungen hat, von denen keine redundant ist, d.h. $\text{rang}(A) = \text{rang}(A, b) = m$. Dann können wir nach genau m Variablen auflösen. Wie kann man nun für x_{i_1}, \dots, x_{i_m} feststellen, ob man nach diesen Variablen auflösen kann? Man schaut sich die Matrix $A' \in \text{Mat}_m(\mathbb{R})$ an, die genau die Spalten i_1, \dots, i_m von A enthält, und benötigt, daß diese den vollen Rang m hat, oder alternativ, daß ihre Determinante $\det(A')$ ungleich null ist.

Unser Ziel ist es, die implizit gegebene Lösungsmenge der Gleichung

$$f(x) = Ax - b = 0$$

als Graph einer Abbildung zu parametrisieren. Nehmen wir an, daß keine Gleichung redundant ist und daß die letzten m Spalten von $A \in \text{Mat}(m \times n + m, \mathbb{R})$ mögliche Pivotspalten sind, d.h.

$$\det((D_j f_i(0))_{i=1, \dots, m, j=n+1, \dots, n+m}) = \det((a_{ij})_{i=1, \dots, m, j=n+1, \dots, n+m}) \neq 0 \quad (149)$$

dann erhalten wir eine Parametrisierung von

$$\text{Lös}(A, \mathbf{b}) = \{\mathbf{x} \in \mathbb{R}^{n+m} \mid f(\mathbf{x}) = 0\}$$

durch den Graphen einer Abbildung

$$\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

mit

$$f(z, \varphi(z)) = 0$$

für alle $z \in \mathbb{R}^n$.

- d. Ist eine Gleichung oder ein Gleichungssystem nicht mehr linear, so können wir nicht mehr so ohne weiteres nach allen vorkommenden Variablen auflösen.

$$f(x_1, x_2) = x_1 \cdot \exp(x_2) - \sin(x_2) = 0$$

mit $x_1, x_2 \in \mathbb{R}$ läßt sich z.B. leicht nach x_1 auflösen als

$$x_1 = \frac{\sin(x_2)}{\exp(x_2)},$$

es ist aber vollkommen unklar ob und ggf. wie man die Gleichung nach der Variablen x_2 auflösen sollte.

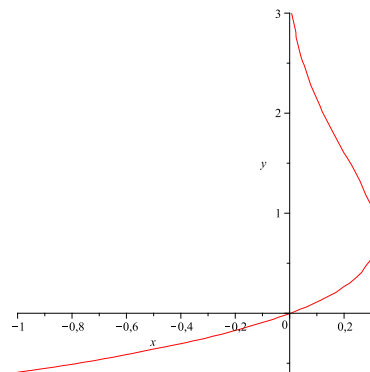


ABBILDUNG 9. Die Lösungsmenge $x_1 \cdot \exp(x_2) - \sin(x_2) = 0$

- e. Bei der nicht-linearen Gleichung

$$f(\mathbf{x}) = x_2^2 - x_1^2 - x_1^3 = 0$$

ist es nicht so schwer, die Gleichung nach x_2 aufzulösen als

$$x_2 = \pm \sqrt{x_1^2 + x_1^3},$$

aber man erhält auf dem Weg keine global gültige *eindeutige* Lösung mehr. Zieht man sich aber auf eine kleine Umgebung eines Punktes

$$(\mathbf{a}, \mathbf{b})^t \in V(f) = \{\mathbf{x} \in \mathbb{R}^2 \mid f(\mathbf{x}) = 0\}$$

zurück, so kann man die *implizit* gegebene Menge $V(f)$ in dieser Umgebung meistens wieder *explizit* als Graph einer Funktion parametrisieren — siehe den linken kleinen Kasten in Abbildung 10. Dabei heißt “meistens”, daß $\mathbf{b} \neq 0$ gilt,

da man sich dann in einer kleinen Umgebung von $(\mathbf{a}, \mathbf{b})^t$ für einen der beiden Zweige von $\pm\sqrt{x_1^2 + x_1^3}$ entscheiden kann. Genauer gesagt, man kann dann die folgende Vorschrift für die Abbildung φ wählen

$$\varphi(x_1) = \begin{cases} \sqrt{x_1^2 + x_1^3}, & \text{falls } \mathbf{b} > 0, \\ -\sqrt{x_1^2 + x_1^3}, & \text{falls } \mathbf{b} < 0. \end{cases}$$

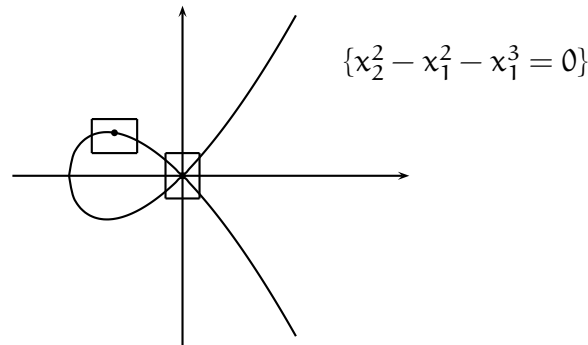


ABBILDUNG 10. Die Lösungsmenge von $x_2^2 - x_1^2 - x_1^3 = 0$

Ist hingegen $\mathbf{b} = 0$, wie etwa im Punkt $(\mathbf{a}, \mathbf{b})^t = (0, 0)^t$, so kann man sich nicht für einen der beiden Zweige entscheiden und es gibt dann lokal keine Möglichkeit, die Gleichung nach x_2 aufzulösen, sprich lokal eine eindeutige Parametrisierung mit $x_2 = \varphi(x_1)$ zu finden, wie man sehr schön in Abbildung 10 für den Punkt $(\mathbf{a}, \mathbf{b})^t = (0, 0)^t$ sieht.

Man beachte auch, daß die Punkte $(\mathbf{a}, \mathbf{b})^t$ mit lokaler Parametrisierung als Graph einer Funktion genau die Punkte mit

$$D_2f(\mathbf{a}, \mathbf{b}) = 2\mathbf{b} \neq 0$$

sind. Das ist kein Zufall!

Nach diesen Betrachtungen wollen wir unser oben angegebenes Ziel, Gleichungen und Gleichungssystem nach einzelnen Variablen aufzulösen wie folgt konkretisieren:

Ziel der folgenden Überlegungen ist es, ein Kriterium dafür zu finden, daß eine *implizit* gegebene Menge $\{x \mid f(x) = 0\}$ lokal in einem Punkt *explizit* als Graph einer Abbildung φ beschrieben werden kann. Interessanterweise werden wir dabei die Ableitung von φ angeben können, ohne φ selbst kennen zu müssen.

Bemerkung 43.6

Mit $(x, y)^t = (x_1, \dots, x_n, y_1, \dots, y_m)^t$ wollen wir sowohl die Koordinaten auf $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$ bezeichnen, als auch Punkte in \mathbb{R}^{n+m} .

- a. Sei $U \subseteq \mathbb{R}^{n+m}$ offen und $f : U \rightarrow \mathbb{R}^m$ stetig differenzierbar. Für $z \in U$ setze

$$D_x f(z) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(z) & \cdots & \frac{\partial f_1}{\partial x_n}(z) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(z) & \cdots & \frac{\partial f_m}{\partial x_n}(z) \end{pmatrix} \in \text{Mat}(m \times n, \mathbb{R})$$

und

$$D_y f(z) = \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(z) & \cdots & \frac{\partial f_1}{\partial y_m}(z) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(z) & \cdots & \frac{\partial f_m}{\partial y_m}(z) \end{pmatrix} \in \text{Mat}_m(\mathbb{R}).$$

Dann gilt

$$Df(z) = (D_x f(z), D_y f(z)) \in \text{Mat}(m \times n + m, \mathbb{R})$$

und

$$Df(z) \circ \begin{pmatrix} x \\ y \end{pmatrix} = D_x f(z) \circ x + D_y f(z) \circ y. \quad (150)$$

- b. Man sieht leicht, daß durch

$$\|(x, y)^t\|_p := \max\{\|x\|_2, \|y\|_2\}$$

für $(x, y)^t \in \mathbb{R}^{n+m}$ eine Norm auf \mathbb{R}^{n+m} definiert wird, die wir die *Produktnorm* nennen wollen und die wir im Beweis des Satzes über implizite Funktionen 43.7 an einer Stelle verwenden werden. In dieser Norm ist die ε -Umgebung von (a, b)

$$U_\varepsilon^{\|\cdot\|_p}(a, b) = U_\varepsilon^{\|\cdot\|_2}(a) \times U_\varepsilon^{\|\cdot\|_2}(b)$$

das Produkt der ε -Umgebungen von a und b .

- c. Wenn wir im folgenden Normen von Matrizen $A = (a_{ij}) \in \text{Mat}(p \times q, \mathbb{R})$ betrachten, so identifizieren wir $\text{Mat}(p \times q, \mathbb{R})$ stets mit $\mathbb{R}^{p \cdot q}$ und verwenden die euklidische Norm (siehe Aufgabe 39.55).

Satz 43.7 (Satz über implizite Funktionen)

Es seien $U \subseteq \mathbb{R}^{n+m}$ offen, $f : U \rightarrow \mathbb{R}^m$ stetig differenzierbar und $(a, b) = (a_1, \dots, a_n, b_1, \dots, b_m) \in U$ mit $f(a, b) = 0$ und $\det(D_y f(a, b)) \neq 0$.

Dann gibt es $\varepsilon, r > 0$ mit $U_\varepsilon(a) \times U_r(b) \subseteq U$ so, daß es genau eine Abbildung

$$\varphi : U_\varepsilon(a) \rightarrow U_r(b)$$

gibt mit

$$f(x, \varphi(x)) = 0$$

für alle $x \in U_\varepsilon(a)$. Insbesondere gilt $\varphi(a) = b$.

Zudem ist die Abbildung φ stetig differenzierbar auf $U_\varepsilon(a)$ mit

$$D\varphi(x) = - (D_y f(x, \varphi(x)))^{-1} \circ D_x f(x, \varphi(x))$$

für $x \in U_\varepsilon(a)$. Insbesondere gilt also

$$D\varphi(a) = - (D_y f(a, b))^{-1} \circ D_x f(a, b)$$

Bevor wir uns dem Beweis des Satzes zuwenden, wollen wir seine Aussage an zwei Beispielen verdeutlichen.

Beispiel 43.8 (Die Kreislinie)

Die Lösungsmenge $V(f)$ der Gleichung

$$f(x, y) = x^2 + y^2 - 1 = 0$$

für $(x, y)^t \in \mathbb{R}^2$ ist die Kreislinie (siehe Abbildung 11). Für $(a, b)^t \in V(f)$ mit $b \neq 0$

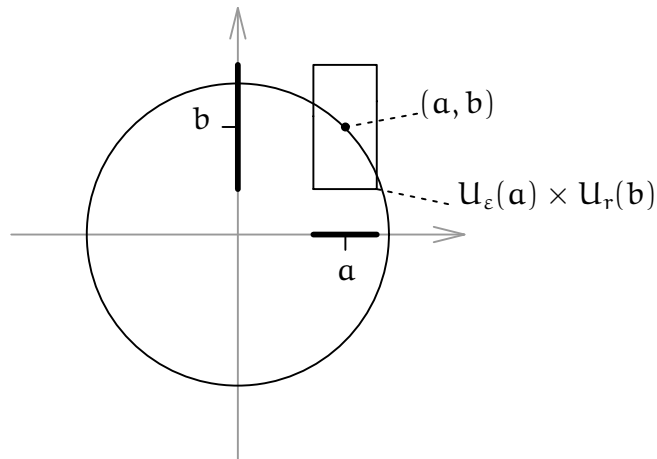


ABBILDUNG 11. Die Kreislinie $\{x^2 + y^2 - 1 = 0\}$

gilt

$$D_y f(a, b) = 2b \neq 0,$$

so daß wir $V(f)$ lokal in (a, b) als Graph einer Funktion φ beschreiben können, nämlich

$$\varphi : U_\varepsilon(a) \longrightarrow U_r(b) : x \mapsto \begin{cases} \sqrt{1 - x^2}, & \text{falls } b > 0, \\ -\sqrt{1 - x^2}, & \text{falls } b < 0. \end{cases}$$

Aus dem Satz über implizite Funktionen wissen wir, daß

$$D\varphi(x) = \varphi'(x) = - (D_y f(x, \varphi(x)))^{-1} \circ D_x f(x, \varphi(x)) = -\frac{1}{2\varphi(x)} \cdot 2x = -\frac{x}{\varphi(x)}$$

gilt, was wir aufgrund der konkreten Beschreibung von φ auch direkt nachrechnen können.

Man beachte, daß in den Punkten $(1, 0)$ und $(-1, 0)$ die Menge $V(f)$ offensichtlich *nicht* lokal als Graph einer Funktion geschrieben werden kann! Dies sind genau die Punkte, in denen $D_y f$ den Wert 0 annimmt, so daß der Satz über implizite Funktionen dort nicht anwendbar ist.

Beispiel 43.9

In unserer Motivation 43.5 haben wir die Gleichung

$$f(x, y) = x \cdot \exp(y) - \sin(y) = 0$$

betrachtet. Wollen wir diese nach \mathbf{y} auflösen, so müssen wir uns die Ableitung

$$\frac{\partial f}{\partial \mathbf{y}}(\mathbf{x}) = \mathbf{x} \cdot \exp(\mathbf{y}) - \cos(\mathbf{y})$$

anschauen. Dort, wo diese nicht 0 ist, ist die Lösungsmenge der Gleichung lokal parametrisierbar als Graph einer Funktion. Die schlechten Punkte sind also die, in denen

$$\mathbf{x} \cdot \exp(\mathbf{y}) - \sin(\mathbf{y}) = 0$$

und

$$\mathbf{x} \cdot \exp(\mathbf{y}) - \cos(\mathbf{y}) = 0$$

gilt, und damit muß

$$\sin(\mathbf{y}) = \cos(\mathbf{y})$$

gelten, was genau für

$$\mathbf{y} = \frac{\pi}{4} + k \cdot \pi \quad \text{und} \quad \mathbf{x} = \frac{\sin(\mathbf{y})}{\exp(\mathbf{y})} = \frac{\pm 1}{\sqrt{2} \cdot \exp(\mathbf{y})}$$

mit $k \in \mathbb{Z}$ der Fall ist. Außer in diesen Punkten ist die Gleichung also lokal parametrisierbar als Graph einer Funktion φ , und obwohl wir für φ keine geschlossene Formel angeben können, folgt aus dem Satz über implizite Funktionen 43.7 für die Ableitung von φ in \mathbf{a} für den Punkt $(\mathbf{a}, \mathbf{b})^t \in V(f) = \{(\mathbf{x}, \mathbf{y})^t \in \mathbb{R}^2 \mid f(\mathbf{x}, \mathbf{y}) = 0\}$ dann

$$D\varphi(\mathbf{a}) = -\frac{1}{\frac{\partial f}{\partial \mathbf{y}}(\mathbf{a}, \mathbf{b})} \cdot \frac{\partial f}{\partial \mathbf{x}}(\mathbf{a}, \mathbf{b}) = \frac{\exp(\mathbf{b})}{\cos(\mathbf{b}) - \mathbf{a} \cdot \exp(\mathbf{b})}.$$

Wir können die Ableitung der unbekannteten Funktion φ im Punkt \mathbf{a} also angeben.

Beweis des Satzes über implizite Funktionen 43.7: Wir teilen den recht langen und komplexen Beweis in mehrere Schritte auf. Im Zentrum steht eine geschickte Anwendung des Banachschen Fixpunktsatzes 43.4, um die Existenz und Eindeutigkeit der Abbildung φ herzuleiten.

1) Ohne Einschränkung gilt $\det(D_{\mathbf{y}}f(\mathbf{x}, \mathbf{y})) \neq 0$ für alle $(\mathbf{x}, \mathbf{y})^t \in \mathbf{U}$:

Die Abbildung

$$\mathbf{U} \longrightarrow \mathbb{R} : (\mathbf{x}, \mathbf{y})^t \mapsto \det(D_{\mathbf{y}}f(\mathbf{x}, \mathbf{y}))$$

ist stetig auf \mathbf{U} als Verknüpfung der stetigen Polynomfunktion \det mit der stetigen Ableitung $D_{\mathbf{y}}f$. Zudem ist $\det(D_{\mathbf{y}}f(\mathbf{a}, \mathbf{b})) \neq 0$ und somit gibt es wegen der Stetigkeit eine ganze Umgebung von $(\mathbf{a}, \mathbf{b})^t$ auf der die Funktion nicht null ist (siehe Aufgabe 40.55). Wir können nun \mathbf{U} einfach durch diese Umgebung ersetzen.

2) Definition einer Hilfsfunktion g :

Wir setzen

$$\mathbf{A} := D_{\mathbf{y}}f(\mathbf{a}, \mathbf{b}) \in \text{Gl}_m(\mathbb{R})$$

und definieren

$$g : \mathbf{U} \longrightarrow \mathbb{R}^m : (\mathbf{x}, \mathbf{y})^t \mapsto \mathbf{y} - \mathbf{A}^{-1} \circ f(\mathbf{x}, \mathbf{y}).$$

Dann ist g stetig differenzierbar auf \mathbf{U} mit

$$D_x g(x, y) = D_x y - DA^{-1}(f(x, y)) \circ D_x f(x, y) = -A^{-1} \circ D_x f(x, y)$$

und

$$D_y g(x, y) = D_y y - DA^{-1}(f(x, y)) \circ D_y f(x, y) = \mathbb{1}_m - A^{-1} \circ D_y f(x, y).$$

Außerdem gilt

$$g(\mathbf{a}, \mathbf{b}) = \mathbf{b} - A^{-1} \circ f(\mathbf{a}, \mathbf{b}) = \mathbf{b} - A^{-1} \circ 0 = \mathbf{b} \in \mathbb{R}^m \quad (151)$$

und

$$D_y g(\mathbf{a}, \mathbf{b}) = \mathbb{1}_m - A^{-1} \circ D_y f(\mathbf{a}, \mathbf{b}) = 0 \in \text{Mat}_m(\mathbb{R}) \quad (152)$$

ist die Nullmatrix.

3) Konstruktion von φ mit Hilfe des Banachschen Fixpunktsatzes:

Idee: Finde $\varepsilon, r > 0$, so daß für $x \in \mathbf{U}_\varepsilon(\mathbf{a})$ die Abbildung

$$g_x : \overline{\mathbf{U}_r(\mathbf{b})} \longrightarrow \overline{\mathbf{U}_r(\mathbf{b})} : y \mapsto g(x, y)$$

eine strikte Kontraktion ist und somit genau einen Fixpunkt y_x besitzt, den wir dann als $\varphi(x)$ definieren werden.

Nach Voraussetzung ist die Abbildung

$$\mathbf{U} \longrightarrow \text{Mat}_m(\mathbb{R}) : (x, y)^t \mapsto D_y g(x, y)$$

stetig auf \mathbf{U} , und Stetigkeit ist unabhängig von der gewählten Norm. Wir können auf \mathbf{U} also die Produktnorm aus Bemerkung 43.6 und auf $\text{Mat}_m(\mathbb{R})$ die euklidische Norm wählen. Wegen $D_y g(\mathbf{a}, \mathbf{b}) = 0$ (siehe (152)) gibt es dann ein $r > 0$ mit

$$\|D_y g(x, y)\|_2 = \|D_y g(x, y) - D_y g(\mathbf{a}, \mathbf{b})\|_2 \leq \frac{1}{2m^2} \quad (153)$$

für alle $(x, y)^t \in \overline{\mathbf{U}_r(\mathbf{a})} \times \overline{\mathbf{U}_r(\mathbf{b})}$. Aus Korollar 42.19 erhalten wir damit

$$\|g(x, y) - g(x, y')\|_2 \stackrel{42.19}{\leq} \|y - y'\|_2 \cdot m^2 \cdot \max_{z \in \overline{\mathbf{U}_r(\mathbf{b})}} \|D_y g(x, z)\|_2 \stackrel{(153)}{\leq} \frac{1}{2} \cdot \|y - y'\|_2 \quad (154)$$

für alle $x \in \overline{\mathbf{U}_r(\mathbf{a})}$ und $y, y' \in \overline{\mathbf{U}_r(\mathbf{b})}$. Außerdem ist die Abbildung g stetig in (\mathbf{a}, \mathbf{b}) mit $g(\mathbf{a}, \mathbf{b}) = \mathbf{b}$ (siehe (151)), so daß es sicher ein $\varepsilon > 0$ gibt mit

$$\|g(x, \mathbf{b}) - \mathbf{b}\|_2 = \|g(x, \mathbf{b}) - g(\mathbf{a}, \mathbf{b})\|_2 < \frac{r}{2} \quad (155)$$

für alle $x \in \mathbf{U}_\varepsilon(\mathbf{a})$. Wir können dabei ohne Einschränkung $\varepsilon \leq r$ wählen.

Für $x \in \mathbf{U}_\varepsilon(\mathbf{a})$ betrachten wir nun die Abbildung

$$g_x : \overline{\mathbf{U}_r(\mathbf{b})} \longrightarrow \mathbb{R}^m : y \mapsto g(x, y).$$

Ist $y \in \overline{\mathbf{U}_r(\mathbf{b})}$, so gilt

$$\|g_x(y) - \mathbf{b}\|_2 \leq \|g_x(y) - g_x(\mathbf{b})\|_2 + \|g_x(\mathbf{b}) - \mathbf{b}\|_2 \stackrel{(154), (155)}{<} \frac{1}{2} \cdot \|y - \mathbf{b}\|_2 + \frac{r}{2} \leq r, \quad (156)$$

so daß in der Tat

$$g_x : \overline{U_r(\mathbf{b})} \longrightarrow \overline{U_r(\mathbf{b})}$$

eine Selbstabbildung von $\overline{U_r(\mathbf{b})}$ und wegen (154) sogar eine $\frac{1}{2}$ -Kontraktion ist. Als abgeschlossene Teilmenge eines vollständigen Raumes ist $\overline{U_r(\mathbf{b})}$ zudem selbst vollständig, da jede Cauchy-Folge in $\overline{U_r(\mathbf{b})}$ in \mathbb{R}^m konvergiert und wegen der Abgeschlossenheit ihren Grenzwert in $\overline{U_r(\mathbf{b})}$ hat (siehe auch Aufgabe 39.54). Wir können deshalb den Banachschen Fixpunktsatz 43.4 anwenden und finden genau ein $\mathbf{y}_x \in \overline{U_r(\mathbf{b})}$ mit

$$\mathbf{y}_x = g_x(\mathbf{y}_x) = \mathbf{y}_x - \mathbf{A}^{-1} \circ f(\mathbf{x}, \mathbf{y}_x), \quad (157)$$

was äquivalent zu

$$0 = \mathbf{A} \circ 0 = \mathbf{A} \circ (\mathbf{y}_x - \mathbf{y}_x + \mathbf{A}^{-1} \circ f(\mathbf{x}, \mathbf{y}_x)) = \mathbf{A} \circ \mathbf{A}^{-1} \circ f(\mathbf{x}, \mathbf{y}_x) = f(\mathbf{x}, \mathbf{y}_x)$$

ist, und wegen (156) gilt sogar $\mathbf{y}_x \in U_r(\mathbf{b})$. Wir definieren deshalb

$$\varphi : U_\varepsilon(\mathbf{a}) \longrightarrow U_r(\mathbf{b}) : \mathbf{x} \mapsto \mathbf{y}_x.$$

Dann gilt zumindest schon mal

$$f(\mathbf{x}, \varphi(\mathbf{x})) = 0$$

und φ ist auf $U_\varepsilon(\mathbf{a})$ eindeutig mit dieser Eigenschaft.

4) Zeige, daß φ auf $U_\varepsilon(\mathbf{a})$ stetig ist:

Seien $\mathbf{x}, \mathbf{x}' \in U_\varepsilon(\mathbf{a})$. Dann gilt

$$\varphi(\mathbf{x}) - g(\mathbf{x}, \varphi(\mathbf{x})) \stackrel{(157)}{=} 0 \stackrel{(157)}{=} \varphi(\mathbf{x}') - g(\mathbf{x}', \varphi(\mathbf{x}')),$$

und wir erhalten durch Umstellen der Gleichung

$$\begin{aligned} \|\varphi(\mathbf{x}) - \varphi(\mathbf{x}')\|_2 &= \|g(\mathbf{x}, \varphi(\mathbf{x})) - g(\mathbf{x}', \varphi(\mathbf{x}'))\|_2 \\ &\leq \|g(\mathbf{x}, \varphi(\mathbf{x})) - g(\mathbf{x}, \varphi(\mathbf{x}'))\|_2 + \|g(\mathbf{x}, \varphi(\mathbf{x}')) - g(\mathbf{x}', \varphi(\mathbf{x}'))\|_2 \end{aligned} \quad (158)$$

Wenden wir Korollar 42.19 auf die Funktion $z \mapsto g(z, \varphi(\mathbf{x}'))$ an, so erhalten wir die Abschätzung

$$\begin{aligned} \|g(\mathbf{x}, \varphi(\mathbf{x}')) - g(\mathbf{x}', \varphi(\mathbf{x}'))\|_2 &\stackrel{42.19}{\leq} \|\mathbf{x} - \mathbf{x}'\|_2 \cdot n^2 \cdot \max_{z \in U_\varepsilon(\mathbf{a})} \|D_x g(z, \varphi(\mathbf{x}'))\|_2 \\ &\leq \|\mathbf{x} - \mathbf{x}'\|_2 \cdot L, \end{aligned} \quad (159)$$

wobei

$$L = n^2 \cdot \max \{ \|D_x g(z, \mathbf{y})\|_2 \mid z \in \overline{U_\varepsilon(\mathbf{a})}, \mathbf{y} \in \overline{U_r(\mathbf{b})} \} \in \mathbb{R}$$

eine von \mathbf{x} und \mathbf{x}' unabhängige Konstante ist. Setzen wir (159) in (158) ein und berücksichtigen zudem noch (154), so erhalten wir

$$\|\varphi(\mathbf{x}) - \varphi(\mathbf{x}')\|_2 \leq \frac{1}{2} \cdot \|\varphi(\mathbf{x}) - \varphi(\mathbf{x}')\|_2 + \|\mathbf{x} - \mathbf{x}'\|_2 \cdot L$$

und damit

$$\|\varphi(\mathbf{x}) - \varphi(\mathbf{x}')\|_2 \leq 2 \cdot L \cdot \|\mathbf{x} - \mathbf{x}'\|_2. \quad (160)$$

Mithin ist φ Lipschitz stetig und nach Proposition 40.33 somit auch stetig auf $U_\varepsilon(\mathbf{a})$.

5) Zeige, daß φ auf $U_\varepsilon(\mathbf{a})$ total differenzierbar ist mit

$$D\varphi(\mathbf{x}) = - (D_y f(\mathbf{x}, \varphi(\mathbf{x})))^{-1} \circ D_x f(\mathbf{x}, \varphi(\mathbf{x})) :$$

Sei $(\mathbf{x}, \mathbf{y})^t \in U_\varepsilon(\mathbf{a}) \times U_r(\mathbf{b})$ gegeben. Wir definieren

$$\rho : U \longrightarrow \mathbb{R}^m$$

durch

$$\begin{aligned} \rho(\mathbf{x}', \mathbf{y}') &= f(\mathbf{x}', \mathbf{y}') - f(\mathbf{x}, \mathbf{y}) - Df(\mathbf{x}, \mathbf{y}) \circ ((\mathbf{x}', \mathbf{y}')^t - (\mathbf{x}, \mathbf{y})^t) \\ &\stackrel{(150)}{=} f(\mathbf{x}', \mathbf{y}') - f(\mathbf{x}, \mathbf{y}) - D_x f(\mathbf{x}, \mathbf{y}) \circ (\mathbf{x}' - \mathbf{x}) - D_y f(\mathbf{x}, \mathbf{y}) \circ (\mathbf{y}' - \mathbf{y}). \end{aligned} \quad (161)$$

Da f in $(\mathbf{x}, \mathbf{y})^t$ total differenzierbar ist, gilt dann

$$\lim_{(\mathbf{x}', \mathbf{y}') \rightarrow (\mathbf{x}, \mathbf{y})} \frac{\rho(\mathbf{x}', \mathbf{y}')}{\|(\mathbf{x}', \mathbf{y}')^t - (\mathbf{x}, \mathbf{y})^t\|_2} = 0. \quad (162)$$

Aufgrund der Definition von φ erhalten wir aus (161) mit $\mathbf{y} = \varphi(\mathbf{x})$ und $\mathbf{y}' = \varphi(\mathbf{x}')$

$$\begin{aligned} 0 &= f(\mathbf{x}', \varphi(\mathbf{x}')) - f(\mathbf{x}, \varphi(\mathbf{x})) \\ &= D_x f(\mathbf{x}, \varphi(\mathbf{x})) \circ (\mathbf{x}' - \mathbf{x}) + D_y f(\mathbf{x}, \varphi(\mathbf{x})) \circ (\varphi(\mathbf{x}') - \varphi(\mathbf{x})) + \rho(\mathbf{x}', \varphi(\mathbf{x}')). \end{aligned}$$

Lösen wir die Gleichung nach $\varphi(\mathbf{x}') - \varphi(\mathbf{x})$ auf, so ergibt sich

$$\begin{aligned} \varphi(\mathbf{x}') - \varphi(\mathbf{x}) &= - (D_y f(\mathbf{x}, \varphi(\mathbf{x})))^{-1} \circ D_x f(\mathbf{x}, \varphi(\mathbf{x})) \circ (\mathbf{x}' - \mathbf{x}) \\ &\quad - (D_y f(\mathbf{x}, \varphi(\mathbf{x})))^{-1} \circ \rho(\mathbf{x}', \varphi(\mathbf{x}')), \end{aligned}$$

wobei wir berücksichtigen, daß $D_y f(\mathbf{x}, \varphi(\mathbf{x}))$ wegen 1) invertierbar ist.

Für die totale Differenzierbarkeit von φ in $(\mathbf{x}, \mathbf{y})^t$ bleibt zu zeigen, daß

$$\lim_{\mathbf{x}' \rightarrow \mathbf{x}} \frac{- (D_y f(\mathbf{x}, \varphi(\mathbf{x})))^{-1} \circ \rho(\mathbf{x}', \varphi(\mathbf{x}'))}{\|\mathbf{x}' - \mathbf{x}\|_2} = 0,$$

und da die Multiplikation mit der Matrix $- (D_y f(\mathbf{x}, \varphi(\mathbf{x})))^{-1}$ eine stetige lineare Abbildung ist, reicht es dazu, zu zeigen, daß

$$\lim_{\mathbf{x}' \rightarrow \mathbf{x}} \frac{\rho(\mathbf{x}', \varphi(\mathbf{x}'))}{\|\mathbf{x}' - \mathbf{x}\|_2} = 0 \quad (163)$$

gilt. Dazu beachten wir, daß man für $\mathbf{u} \in \mathbb{R}^n$ und $\mathbf{v} \in \mathbb{R}^m$ stets

$$\|(\mathbf{u}, \mathbf{v})\|_2 = \sqrt{\sum_{i=1}^n u_i^2 + \sum_{j=1}^m v_j^2} \leq \sqrt{\sum_{i=1}^n u_i^2} + \sqrt{\sum_{j=1}^m v_j^2} \leq \|\mathbf{u}\|_2 + \|\mathbf{v}\|_2 \quad (164)$$

erhält und damit auch

$$\frac{\rho(\mathbf{x}', \varphi(\mathbf{x}'))}{\|(\mathbf{x}', \varphi(\mathbf{x}')) - (\mathbf{x}, \varphi(\mathbf{x}))\|_2} \stackrel{(164)}{\geq} \frac{\rho(\mathbf{x}', \varphi(\mathbf{x}'))}{\|\mathbf{x}' - \mathbf{x}\|_2 + \|\varphi(\mathbf{x}') - \varphi(\mathbf{x})\|_2} \stackrel{(160)}{\geq} \frac{\rho(\mathbf{x}', \varphi(\mathbf{x}'))}{(1 + 2L) \cdot \|\mathbf{x}' - \mathbf{x}\|_2}.$$

Wegen (162) und da φ nach 4) stetig ist, konvergiert die linke Seite für $\mathbf{x}' \rightarrow \mathbf{x}$ gegen 0, also muß auch die rechte Seite gegen 0 konvergieren und (163) ist gezeigt.

6) Zeige, daß $D\varphi$ auf $U_\varepsilon(\mathbf{a})$ stetig ist:

Wegen Teil 4) ist φ stetig und nach Voraussetzung ist zudem $D_x f$ stetig, so daß auch die Verkettung

$$U_\varepsilon(\mathbf{a}) \longrightarrow \text{Mat}(\mathbf{m} \times \mathbf{n}, \mathbb{R}) : \mathbf{x} \mapsto D_x f(\mathbf{x}, \varphi(\mathbf{x}))$$

stetig ist. Analog ist auch die Abbildung

$$U_\varepsilon(\mathbf{a}) \longrightarrow \text{Mat}_m(\mathbb{R}) : \mathbf{x} \mapsto D_y f(\mathbf{x}, \varphi(\mathbf{x}))$$

stetig und da auch das Invertieren einer Matrix nach Beispiel 40.21 stetig ist, ist auch die Abbildung

$$U_\varepsilon(\mathbf{a}) \longrightarrow \text{Mat}_m(\mathbb{R}) : \mathbf{x} \mapsto (D_y f(\mathbf{x}, \varphi(\mathbf{x})))^{-1}$$

stetig. Aber damit ist dann auch

$$U_\varepsilon(\mathbf{a}) \longrightarrow \text{Mat}(\mathbf{m} \times \mathbf{n}, \mathbb{R}) : \mathbf{x} \mapsto D\varphi(\mathbf{x}) = - (D_y f(\mathbf{x}, \varphi(\mathbf{x})))^{-1} \circ D_x f(\mathbf{x}, \varphi(\mathbf{x}))$$

stetig, d.h. die Komponentenfunktionen sind stetig. \square

Bemerkung 43.10 (Satz über implizite Funktionen)

- Der Satz über implizite Funktionen 43.7 gibt leider *keinerlei* Hinweis darauf, wie man die implizit gegebene Abbildung φ finden kann. Es ist eine reine Existenzaussage. Aber, obwohl man φ in aller Regel nicht näher bestimmen kann, kann man die Ableitung von φ im Punkt \mathbf{a} berechnen!
- Der Satz über implizite Funktionen spielt eine wichtige Rolle bei der Betrachtung von Extrema mit Nebenbedingungen und im Beweis der Transformationsformel für Integrale.
- Man kann den Satz über implizite Funktionen ohne wesentliche Änderungen auch allgemeiner für Banachräume zeigen (siehe [Heu04, S. 330-345]).

C) Der Satz über die Umkehrfunktion**Bemerkung 43.11** (Umkehrsatz für stetig differenzierbare Funktionen)

Ist eine Funktion $f : (\mathbf{a}, \mathbf{b}) \longrightarrow \mathbb{R}$ stetig differenzierbar mit $f'(x) \neq 0$ für alle $x \in (\mathbf{a}, \mathbf{b})$, so ist die stetige Funktion $f' : (\mathbf{a}, \mathbf{b}) \longrightarrow \mathbb{R}$ wegen des Zwischenwertsatzes 14.12 entweder stets positiv oder stets negativ. Aus Proposition 18.11 ist die Funktion f dann aber streng monoton und wegen des Umkehrsatzes für streng monotone Funktionen 14.21 ist f deshalb bijektiv mit $\text{Im}(f) = (\mathbf{c}, \mathbf{d})$ für geeignete $\mathbf{c}, \mathbf{d} \in \mathbb{R} \cup \{-\infty, \infty\}$, d.h. es existiert die Umkehrfunktion

$$f^{-1} : (\mathbf{c}, \mathbf{d}) \longrightarrow (\mathbf{a}, \mathbf{b}).$$

Wegen des Satzes über die Ableitung der Umkehrfunktion 17.14 ist diese sogar differenzierbar und wir können zudem die Ableitung der Umkehrfunktion berechnen,

ohne diese wirklich zu kennen; für $\mathbf{y} = f(\mathbf{x})$ gilt nämlich

$$(f^{-1})'(\mathbf{y}) = \frac{1}{f'(\mathbf{x})} = (f'(\mathbf{x}))^{-1}. \quad (165)$$

Diesen Sachverhalt wollen wir nun ins Mehrdimensionale verallgemeinern. Der Satz über implizite Funktionen legt dabei nahe, die Bedingung $f'(\mathbf{x}) \neq 0$ durch die Bedingung $\det(Df(\mathbf{x})) \neq 0$ zu ersetzen. Das folgende Beispiel zeigt jedoch, daß wir dabei vorsichtig sein und gewisse Abstriche machen müssen.

Beispiel 43.12 (Polarkoordinaten)

Wir betrachten die Polarkoordinatenabbildung

$$f : (0, \infty) \times \mathbb{R} \longrightarrow \mathbb{R}^2 : (r, \theta)^t \mapsto (r \cdot \cos(\theta), r \cdot \sin(\theta))^t.$$

Diese ist stetig differenzierbar auf ihrem Definitionsbereich und die Ableitungsdeterminante

$$\det(Df(r, \theta)) = \begin{vmatrix} \cos(\theta) & -r \cdot \sin(\theta) \\ \sin(\theta) & r \cdot \cos(\theta) \end{vmatrix} = r \cdot (\sin^2(\theta) + \cos^2(\theta)) = r > 0$$

ist stets positiv. Die Abbildung ist aber nicht injektiv, da

$$f(r, \theta) = f(r, \theta + 2\pi)$$

für alle (r, θ) im Definitionsbereich. f besitzt deshalb *global* keine Umkehrfunktion.

Beschränkt man sich zu gegebenem (r, θ) jedoch auf die $\frac{\pi}{2}$ -Umgebung $\mathbf{U}_{\frac{\pi}{2}}(r, \theta)$ von (r, θ) , so ist die Abbildung injektiv, wie man leicht sieht, und der folgende Satz über die Umkehrfunktion zeigt, daß auf dieser Umgebung dann auch die Umkehrfunktion stetig differenzierbar ist und wie wir ihre Ableitung berechnen können, indem wir die Formel (165) verallgemeinern.

Definition 43.13 (Diffeomorphismen)

Es sei $\mathbf{U} \subseteq \mathbb{R}^n$ offen, $f : \mathbf{U} \longrightarrow \mathbb{R}^n$ und $\mathbf{a} \in \mathbf{U}$.

- f ist ein *Diffeomorphismus*, wenn $f(\mathbf{U})$ offen in \mathbb{R}^n ist und $f : \mathbf{U} \longrightarrow f(\mathbf{U})$ bijektiv und stetig differenzierbar mit stetig differenzierbarer Umkehrabbildung f^{-1} ist.
- f ist *lokaler Diffeomorphismus in \mathbf{a}* , wenn es eine offene Umgebung $\mathbf{V}_{\mathbf{a}}$ von \mathbf{a} gibt, so daß $f|_{\mathbf{V}_{\mathbf{a}}}$ ein Diffeomorphismus ist.
- f ist ein *lokaler Diffeomorphismus*, wenn f in jedem Punkt von \mathbf{U} ein lokaler Diffeomorphismus ist.

Das Bestimmen einer Umkehrfunktion von f kann man auch als Lösen der impliziten Gleichung $\mathbf{y} = f(\mathbf{x})$ nach \mathbf{x} auffassen. In diesem Sinn ist der folgende Satz über die Umkehrfunktion ein Spezialfall des Satzes über implizite Funktionen, und die angegebene Idee sagt auch schon, wie der Satz bewiesen wird.

Satz 43.14 (Satz über die Umkehrfunktion)

Sei $\mathbf{c} \in \mathbf{U} \subseteq \mathbb{R}^n$ offen und $f : \mathbf{U} \rightarrow \mathbb{R}^n$ stetig differenzierbar mit $\det(Df(\mathbf{c})) \neq 0$.

Dann gibt es $\mathbf{V} \subseteq \mathbf{U}$ offen und $\mathbf{W} \subseteq \mathbb{R}^n$ offen, so daß

$$f|_{\mathbf{V}} : \mathbf{V} \rightarrow \mathbf{W}$$

bijektiv ist. Zudem ist die Umkehrabbildung $f|_{\mathbf{V}}^{-1} : \mathbf{W} \rightarrow \mathbf{V}$ stetig differenzierbar mit

$$Df|_{\mathbf{V}}^{-1}(\mathbf{y}) = \left(Df(f|_{\mathbf{V}}^{-1}(\mathbf{y})) \right)^{-1}.$$

für $\mathbf{y} \in \mathbf{W}$. Insbesondere ist f ein lokaler Diffeomorphismus in \mathbf{c} mit

$$Df|_{\mathbf{V}}^{-1}(f(\mathbf{c})) = (Df(\mathbf{c}))^{-1}.$$

Beweis: Wir setzen $\mathbf{a} = f(\mathbf{c})$ und $\mathbf{b} = \mathbf{c}$. Die Abbildung

$$\mathbf{h} : \mathbb{R}^n \times \mathbf{U} \rightarrow \mathbb{R}^n : (\mathbf{x}, \mathbf{y}) \mapsto f(\mathbf{y}) - \mathbf{x}$$

ist stetig differenzierbar auf $\mathbb{R}^n \times \mathbf{U}$ mit $\mathbf{h}(\mathbf{a}, \mathbf{b}) = f(\mathbf{b}) - \mathbf{a} = 0$ und

$$\det(D_{\mathbf{y}}\mathbf{h}(\mathbf{a}, \mathbf{b})) = \det(Df(\mathbf{c})) \neq 0.$$

Damit erfüllt \mathbf{h} die Voraussetzungen des Satzes über implizite Funktionen und wir finden eine *eindeutig* bestimmte Abbildung

$$\varphi : \mathbf{U}_{\varepsilon}(\mathbf{a}) \rightarrow \mathbf{U}_r(\mathbf{b}) = \mathbf{U}_r(\mathbf{c}) \subseteq \mathbf{U}$$

mit

$$0 = \mathbf{h}(\mathbf{x}, \varphi(\mathbf{x})) = f(\varphi(\mathbf{x})) - \mathbf{x}, \quad (166)$$

für alle $\mathbf{x} \in \mathbf{U}_{\varepsilon}(\mathbf{a})$. Zudem ist die Funktion φ stetig differenzierbar auf $\mathbf{U}_{\varepsilon}(\mathbf{a})$ mit

$$\begin{aligned} D\varphi(\mathbf{x}) &= - (D_{\mathbf{y}}\mathbf{h}(\mathbf{x}, \varphi(\mathbf{x})))^{-1} \circ D_{\mathbf{x}}\mathbf{h}(\mathbf{x}, \varphi(\mathbf{x})) \\ &= - (Df(\varphi(\mathbf{x})))^{-1} \circ (-\mathbf{1}_n) = (Df(\varphi(\mathbf{x})))^{-1}. \end{aligned} \quad (167)$$

Setzen wir nun $\mathbf{W} = \mathbf{U}_{\varepsilon}(\mathbf{a}) = \mathbf{U}_{\varepsilon}(f(\mathbf{c}))$ und $\mathbf{V} = \varphi(\mathbf{W})$, so ist die Abbildung

$$\varphi : \mathbf{W} \rightarrow \mathbf{V}$$

aufgrund der Definition von \mathbf{V} surjektiv und wegen (166) gilt

$$f|_{\mathbf{V}} \circ \varphi = \text{id}_{\mathbf{W}}.$$

Damit ist φ aber auch injektiv, da aus $\varphi(\mathbf{x}) = \varphi(\mathbf{x}')$ automatisch

$$\mathbf{x} = f(\varphi(\mathbf{x})) = f(\varphi(\mathbf{x}')) = \mathbf{x}'$$

folgt, und $f|_{\mathbf{V}}$ ist die Inverse von φ , d.h. auch

$$f|_{\mathbf{V}} : \mathbf{V} \rightarrow \mathbf{W}$$

ist bijektiv mit $\varphi = f|_{\mathbf{V}}^{-1}$ als Umkehrabbildung. Außerdem gilt auch

$$\mathbf{c} = \mathbf{b} = \varphi(\mathbf{a}) \in \mathbf{V}.$$

Unter Berücksichtigung von (167) ist damit alles bewiesen außer, daß \mathbf{V} *offen* ist.

Sei dazu $x \in V = \varphi(W) \subseteq U_r(\mathbf{b}) \subseteq U$ beliebig gegeben. Dann gilt

$$f(x) = f|_V(x) \in W,$$

und da W offen in \mathbb{R}^n ist, gibt es ein $\gamma > 0$ mit

$$U_\gamma(f(x)) \subseteq W.$$

Nach Voraussetzung ist die Abbildung f stetig auf $U_r(\mathbf{b})$ und deshalb gibt es ein $\delta > 0$ mit $U_\delta(x) \subseteq U_r(\mathbf{b})$ und

$$f(U_\delta(x)) \subseteq U_\gamma(f(x)) \subseteq W.$$

Wir wollen nun zeigen, daß in der Tat schon

$$U_\delta(x) \subseteq V$$

gilt, da dann der beliebige Punkt x in V ein innerer Punkt von V und V offen ist.

Sei also $z \in U_\delta(x) \subseteq U_r(\mathbf{b})$ beliebig, dann ist $f(z) \in W$ und es gilt zudem

$$h(f(z), z) = f(z) - f(z) = 0.$$

Wegen der Eindeutigkeit der Abbildung φ auf $U_r(\mathbf{b})$ muß dann notwendigerweise

$$z = \varphi(f(z)) \in V$$

gelten. Damit ist der Satz dann bewiesen. \square

Bemerkung 43.15 (Satz über die Umkehrfunktion)

- a. Der Satz über die Umkehrfunktion ist die naheliegende lokale Verallgemeinerung des eindimensionalen Umkehrsatzes in Bemerkung 43.11. Die Bedingung $f'(x) \neq 0$ wird dabei zu $\det(Df(x)) \neq 0$, und aus dem Kehrwert der Ableitung zur Bestimmung der Ableitung der Umkehrfunktion wird die Inverse der Ableitung als Matrix.
- b. Man nennt $\det(Df(x))$ auch die *Funktionaldeterminante* von f in x .
- c. Die Bedingung $\det(Df(c)) \neq 0$ ist hinreichend, um lokal die Existenz einer stetig differenzierbaren Umkehrfunktion von f zu sichern. Sie ist aber wegen der Kettenregel zugleich auch notwendig, denn aus

$$\mathbb{1}_n = D \operatorname{id}(c) = D(f^{-1} \circ f)(c) = Df^{-1}(f(c)) \circ Df(c)$$

folgt unmittelbar, daß $Df(c)$ invertierbar sein muß.

- d. Die letzte Anmerkung besagt, daß eine stetig differenzierbare Abbildung f genau dann ein *lokaler Diffeomorphismus in \mathbf{a}* ist, wenn $\det(Df(\mathbf{a})) \neq 0$ gilt.
- e. Der Satz über die Umkehrfunktion 43.14 gibt leider wieder *keinerlei* Hinweis darauf, wie man die Umkehrfunktion bestimmen kann. Es ist eine reine Existenzaussage. Aber, obwohl man f^{-1} in aller Regel nicht näher kennt, kann man wieder die Ableitung von f^{-1} im Punkt \mathbf{c} berechnen!

- f. Wir haben den Satz über implizite Funktionen verwendet, um den Satz über die Umkehrfunktion zu beweisen. Man hätte auch umgekehrt den Satz über die Umkehrfunktion zuerst beweisen und daraus den Satz über implizite Funktionen herleiten können (siehe [Gat08]).
- g. Auch der Satz über die Umkehrfunktion läßt sich wieder ohne wesentliche Änderungen auf Banachräume verallgemeinern (siehe [Heu04, S. 330-345]).

Beispiel 43.16

- a. Die Polarkoordinatenabbildung aus Beispiel 43.12

$$f : (0, \infty) \times \mathbb{R} \longrightarrow \mathbb{R}^2 : (r, \theta)^t \mapsto (r \cdot \cos(\theta), r \cdot \sin(\theta))^t.$$

ist lokal in jedem Punkt (r, θ) ein Diffeomorphismus, da

$$\det(Df(r, \theta)) = r > 0.$$

Und wir können nun auch die Ableitung der Umkehrfunktion in $f(r, \theta)$ berechnen als

$$Df^{-1}(f(r, \theta)) = \begin{pmatrix} \cos(\theta) & -r \cdot \sin(\theta) \\ \sin(\theta) & r \cdot \cos(\theta) \end{pmatrix}^{-1} = \frac{1}{r} \cdot \begin{pmatrix} r \cdot \cos(\theta) & r \cdot \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}.$$

In der Tat könnte man in diesem Beispiel die Umkehrfunktion auch wieder direkt angeben. Z.B. kann man einen Punkt $(x, y)^t \in \mathbb{R}^2$ mit $x > 0$ in Polarkoordinaten (r, θ) angeben mit

$$r = \|(x, y)^t\|_2 = \sqrt{x^2 + y^2}$$

und

$$\tan(\theta) = \frac{\sin(\theta)}{\cos(\theta)} = \frac{y}{x}.$$

Die Funktion

$$f_1 : (0, \infty) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \longrightarrow (0, \infty) \times \mathbb{R}$$

ist also bijektiv mit der Umkehrabbildung

$$f_1^{-1} : (0, \infty) \times \mathbb{R} \longrightarrow (0, \infty) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) : (x, y)^t \mapsto \left(\sqrt{x^2 + y^2}, \arctan\left(\frac{y}{x}\right)\right).$$

Wir überlassen es dem Leser, die Ableitung im Punkt $f(r, \theta)$ mit Hilfe dieser Formel nachzurechnen.

- b. Die Abbildung

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (x_1, x_2, x_3)^t \mapsto (x_1 + x_2 + x_3, x_1x_2 + x_2x_3 + x_3x_1, x_1x_2x_3)^t$$

hat die sogenannten dritten elementarsymmetrischen Polynome als Koordinatenfunktionen. f im Punkt $\mathbf{x} = (x_1, x_2, x_3)^t$ hat die Funktionaldeterminante

$$\begin{aligned} \det(Df(\mathbf{x})) &= \begin{vmatrix} 1 & 1 & 1 \\ x_2 + x_3 & x_1 + x_3 & x_1 + x_2 \\ x_2x_3 & x_1x_3 & x_1x_2 \end{vmatrix} \\ &= \begin{vmatrix} 1 & 0 & 0 \\ x_2 + x_3 & x_1 - x_2 & x_1 - x_3 \\ x_2x_3 & (x_1 - x_2) \cdot x_3 & (x_1 - x_3) \cdot x_2 \end{vmatrix} \\ &= (x_1 - x_2) \cdot (x_1 - x_3) \cdot \begin{vmatrix} 1 & 0 & 0 \\ x_2 + x_3 & 1 & 1 \\ x_2x_3 & x_3 & x_2 \end{vmatrix} \\ &= (x_1 - x_2) \cdot (x_1 - x_3) \cdot (x_2 - x_3), \end{aligned}$$

wie man leicht nachrechnet. Damit gilt, daß f immer dann lokal in $\mathbf{x} = (x_1, x_2, x_3)^t$ invertierbar ist, wenn die x_i paarweise verschieden sind.

Betrachtet man das Polynom

$$\mathbf{p} = (t - x_1) \cdot (t - x_2) \cdot (t - x_3) = t^3 + f_1(\mathbf{x}) \cdot t^2 + f_2(\mathbf{x}) \cdot t + f_3(\mathbf{x}),$$

so sind die Komponentenfunktionen von f gerade die Koeffizienten von \mathbf{p} . Die Inverse f^{-1} von f lokal in $\mathbf{y} = (y_1, y_2, y_3)^t$ zu bestimmen, bedeutet also, die Nullstellen des Polynoms

$$\mathbf{p}_{\mathbf{y}} = t^3 + y_1 t^2 + y_2 t + y_3$$

in Abhängigkeit von \mathbf{y} anzugeben. Der Satz über die Umkehrfunktion besagt nun, daß die Nullstellen von $\mathbf{p}_{\mathbf{y}}$ stetig differenzierbar von \mathbf{y} abhängen, sofern man \mathbf{y} in einer kleinen Umgebung eines Punktes \mathbf{a} bewegt, für den $\mathbf{p}_{\mathbf{a}}$ drei paarweise verschiedene Nullstellen hat.

Korollar 43.17 (Satz von der offenen Abbildung)

Seien $\mathbf{U} \subseteq \mathbb{R}^n$ offen und $f \in C^1(\mathbf{U}, \mathbb{R}^n)$ mit $\det(Df(\mathbf{x})) \neq 0$ für alle $\mathbf{x} \in \mathbf{U}$.

- Ist $\mathbf{O} \subseteq \mathbf{U}$ offen, so ist $f(\mathbf{O})$ offen in \mathbb{R}^n .
- Ist $\mathbf{O} \subseteq \mathbf{U}$ offen und f injektiv, so ist $f^{-1} : f(\mathbf{O}) \rightarrow \mathbf{O}$ stetig differenzierbar.

Beweis: Sei $\mathbf{O} \subseteq \mathbf{U}$ offen und $\mathbf{c} \in \mathbf{O}$ beliebig. Nach dem Satz über die Umkehrfunktion 43.14 angewendet auf

$$f|_{\mathbf{O}} : \mathbf{O} \rightarrow \mathbb{R}^n$$

gibt es dann offene Mengen $V_{\mathbf{c}} \subseteq \mathbf{O}$ und $W_{\mathbf{c}} \subseteq \mathbb{R}^n$, so daß

$$f|_{V_{\mathbf{c}}} : V_{\mathbf{c}} \rightarrow W_{\mathbf{c}}$$

bijektiv ist. Damit gilt dann aber insbesondere

$$W_{\mathbf{c}} = f(V_{\mathbf{c}}) \subseteq f(\mathbf{O}),$$

und somit

$$f(\mathcal{O}) = \bigcup_{c \in \mathcal{O}} W_c \quad (168)$$

als Vereinigung offener Mengen offen.

Ist nun zudem f injektiv, so ist die Abbildung $f : \mathcal{O} \rightarrow f(\mathcal{O})$ bijektiv, und wegen des Satzes über die Umkehrfunktion ist die Umkehrabbildung f^{-1} für jedes $c \in \mathcal{O}$ auf W_c stetig differenzierbar. Damit ist sie wegen (168) aber in jedem Punkt von $f(\mathcal{O})$ differenzierbar und ihre Ableitung ist in jedem Punkt von $f(\mathcal{O})$ stetig, d.h. f^{-1} ist stetig differenzierbar auf ganz $f(\mathcal{O})$. \square

Beispiel 43.18

Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R} : x \mapsto x^2$ ist stetig differenzierbar, aber es gibt einen Punkt $x = 0$ mit

$$\det(Df(x)) = f'(x) = 2x = 0.$$

Die Voraussetzungen des Satzes von der offenen Abbildung 43.17 sind also nicht erfüllt, und in der Tat ist das Bild auch nicht offen in \mathbb{R} ,

$$f(\mathbb{R}) = [0, \infty)$$

Bemerkung 43.19 (Diffeomorphismen)

Kombinieren wir die Aussagen im Satz über die Umkehrfunktion 43.14 mit dem Satz der offenen Abbildung 43.17 und Bemerkung 43.15, so erhalten wir die folgenden Charakterisierungen von (lokalen) Diffeomorphismen $f : \mathcal{U} \rightarrow \mathbb{R}^n$:

- f ist ein *lokaler Diffeomorphismus in \mathbf{a}* $\iff \det(Df(\mathbf{a})) \neq 0$.
- f ist ein *lokaler Diffeomorphismus* $\iff \det(Df(\mathbf{x})) \neq 0$ für alle $\mathbf{x} \in \mathcal{U}$.
- f ist ein *Diffeomorphismus* $\iff f$ ist ein injektiver lokaler Diffeomorphismus.

D) Extrema unter Nebenbedingungen

In Abschnitt 42 haben wir Bedingungen dafür untersucht, daß eine Funktion ein lokales Extremum in einer offenen Menge besitzt. In diesem Abschnitt wollen wir uns mit der gleichen Frage beschäftigen, wenn der Definitionsbereich die Lösungsmenge einer oder mehrerer Gleichungen und damit nicht offen ist. Der Rand vieler interessanter Mengen ist ein Spezialfall solcher Lösungsmengen. Man kann auf dem Weg also auch Extremwertbetrachtungen im Rand des Definitionsbereiches anstellen.

Definition 43.20

Sei $D \subseteq \mathbb{R}^n$, $f : D \rightarrow \mathbb{R}$, $g : D \rightarrow \mathbb{R}^m$ und $\mathbf{a} \in D$.

- f hat in \mathbf{a} ein *globales Maximum unter der Nebenbedingung $g = 0$* , wenn $\mathbf{a} \in V(g)$ und $f(\mathbf{a}) \geq f(\mathbf{x})$ für alle $\mathbf{x} \in V(g)$.
- f hat in \mathbf{a} ein *lokales Maximum unter der Nebenbedingung $g = 0$* , wenn $\mathbf{a} \in V(g)$ und wenn es ein $\delta > 0$ gibt, so daß $f(\mathbf{a}) \geq f(\mathbf{x})$ für alle $\mathbf{x} \in V(g) \cap \mathcal{U}_\delta(\mathbf{a})$.

- c. f hat in \mathbf{a} ein *isoliertes lokales Maximum unter der Nebenbedingung $\mathbf{g} = 0$* , wenn $\mathbf{a} \in V(\mathbf{g})$ und wenn es ein $\delta > 0$ gibt mit $f(\mathbf{a}) > f(\mathbf{x})$ für alle $\mathbf{a} \neq \mathbf{x} \in V(\mathbf{g}) \cap U_\delta(\mathbf{a})$.
- d. f hat in \mathbf{a} ein *globales Minimum unter der Nebenbedingung $\mathbf{g} = 0$* , wenn $\mathbf{a} \in V(\mathbf{g})$ und wenn $f(\mathbf{a}) \leq f(\mathbf{x})$ für alle $\mathbf{x} \in V(\mathbf{g})$.
- e. f hat in \mathbf{a} ein *lokales Minimum unter der Nebenbedingung $\mathbf{g} = 0$* , wenn $\mathbf{a} \in V(\mathbf{g})$ und wenn es ein $\delta > 0$ gibt, so daß $f(\mathbf{a}) \leq f(\mathbf{x})$ für alle $\mathbf{x} \in V(\mathbf{g}) \cap U_\delta(\mathbf{a})$.
- f. f hat in \mathbf{a} ein *isoliertes lokales Minimum unter der Nebenbedingung $\mathbf{g} = 0$* , wenn $\mathbf{a} \in V(\mathbf{g})$ und wenn es ein $\delta > 0$ gibt mit $f(\mathbf{a}) < f(\mathbf{x})$ für alle $\mathbf{a} \neq \mathbf{x} \in V(\mathbf{g}) \cap U_\delta(\mathbf{a})$.
- g. \mathbf{a} heißt *Extremstelle von f unter der Nebenbedingung $\mathbf{g} = 0$* , wenn $\mathbf{a} \in V(\mathbf{g})$ und wenn f in \mathbf{a} ein lokales Minimum oder ein lokales Maximum unter der Nebenbedingung $\mathbf{g} = 0$ hat.

Beispiel 43.21

Manchmal ist es möglich, eine Verschwindungsmenge $V(\mathbf{g})$ global zu parametrisieren. Dann kann man die Untersuchung von Extremstellen mit Nebenbedingungen auf die Untersuchung von Extremstellen ohne Nebenbedingungen zurück führen.

- a. Wir betrachten die Abbildung

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : \mathbf{x} \mapsto \|\mathbf{x}\|_2^2 = x_1^2 + x_2^2$$

und die Nebenbedingung $\mathbf{g} = 0$ mit

$$\mathbf{g} : \mathbb{R}^2 \longrightarrow \mathbb{R} : \mathbf{x} \mapsto x_1 + x_2 - 1.$$

Dann ist $\mathbf{g}(\mathbf{x}) = x_1 + x_2 - 1 = 0$ gleichwertig zu $x_2 = 1 - x_1$, und wir erhalten die Parametrisierung

$$\psi : \mathbb{R} \longrightarrow \mathbb{R}^2 : t \mapsto (t, 1 - t)^t$$

von $V(\mathbf{g})$, d.h.

$$V(\mathbf{g}) = \{(t, 1 - t)^t \mid t \in \mathbb{R}\} = \text{Im}(\psi).$$

Um die Extremstellen von f mit Nebenbedingung $\mathbf{g} = 0$ zu finden, untersuchen wir nun die Extremstellen von

$$\mathbf{h} = f \circ \psi : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto f(t, 1 - t) = 2t^2 - 2t + 1.$$

Die Funktion ist unendlich oft differenzierbar mit

$$\mathbf{h}'(t) = 4t - 2 = 0$$

genau dann, wenn $t = \frac{1}{2}$. Sie hat also nur einen kritischen Punkt, und in diesem gilt

$$\mathbf{h}''\left(\frac{1}{2}\right) = 4 > 0.$$

Mithin hat die Funktion h in $t = \frac{1}{2}$ ein lokales Minimum, und deshalb hat f in $\psi(\frac{1}{2}) = (\frac{1}{2}, \frac{1}{2})^t$ ein lokales Minimum unter der Nebenbedingung $g = 0$.

- b. In Teil a. wurde die globale Parametrisierung durch den Graphen einer Funktion gegeben. Für unsere Zwecke kann die Parametrisierung aber auch allgemeinerer Natur sein.

Betrachten wir etwa die Funktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto \exp(x_1 + x_2)$$

und die Nebenbedingung $g = 0$ für

$$g : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto x_1^2 - x_2^3.$$

Man sieht leicht, daß die Abbildung

$$\psi : \mathbb{R} \longrightarrow \mathbb{R}^2 : t \mapsto (t^3, t^2)^t$$

eine injektive Parametrisierung von $V(g)$ ist. Um die Extremstellen von f unter der Nebenbedingung $g = 0$ zu finden, betrachten wir wieder die Funktion

$$h = f \circ \psi : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto \exp(t^3 + t^2).$$

Sie ist unendlich oft differenzierbar mit

$$h'(t) = (3t^2 + 2t) \cdot \exp(t^3 + t^2) = 0$$

genau dann, wenn $t = 0$ oder $t = -\frac{2}{3}$ gilt. Für die zweite Ableitung

$$h''(t) = (6t + 2) \cdot \exp(t^3 + t^2) + (3t^2 + 2t)^2 \cdot \exp(t^3 + t^2)$$

von h gilt in den kritischen Punkten

$$h''(0) = 2 > 0$$

und

$$h''\left(-\frac{2}{3}\right) = -2 \cdot \exp\left(\frac{4}{27}\right) < 0,$$

so daß h in $t = 0$ ein lokales Minimum und in $t = -\frac{2}{3}$ ein lokales Maximum hat. Entsprechend hat f in $\psi(0) = (0, 0)^t$ ein lokales Minimum unter der Nebenbedingung $g = 0$ und in $\psi(-\frac{2}{3}) = (-\frac{8}{27}, \frac{4}{9})^t$ ein lokales Maximum unter der Nebenbedingung $g = 0$.

Wenn $V(g)$ keine globale Parametrisierung besitzt oder wir sie nicht ausrechnen können, dann hilft uns unter Umständen eine Anwendung des Satzes über implizite Funktionen, der Satz zu den Lagrange-Multiplikatoren, dabei, Kandidaten für Extremstellen unter Nebenbedingungen zu finden. In Proposition 42.24 haben wir ein notwendiges Kriterium für die Existenz einer Extremstelle kennen gelernt. Dieses wollen wir nun auf die Situation einer Extremstelle unter Nebenbedingung verallgemeinern.

Satz 43.22 (Lagrange-Multiplikatoren)

Es seien $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}$ und $g : U \rightarrow \mathbb{R}^m$ stetig differenzierbar auf U und $\mathbf{a} \in V(g)$ mit $\text{rang}(Dg(\mathbf{a})) = m < n$.

Ist \mathbf{a} Extremstelle von f unter der Nebenbedingung $g = 0$, so gibt es ein $\mathbf{b} \in \mathbb{R}^m$ mit

$$Df(\mathbf{a}) = \sum_{j=1}^m b_j \cdot Dg_j(\mathbf{a}).$$

Man bezeichnet b_1, \dots, b_m auch als Lagrange-Multiplikatoren.

Beweis: Die Grundidee des Beweises besteht darin, eine Hilfsfunktion

$$q : U' \rightarrow \mathbb{R}$$

auf einer offenen Menge $U' \subset \mathbb{R}^{n-m}$ zu konstruieren, die eine Extremstelle hat, und dann Proposition 42.24 auf diese Funktion q anzuwenden. Der verschwindende Gradient von q wird die Behauptung des Satzes liefern. Zur Konstruktion von q parametrisiert man mit Hilfe des Satzes über implizite Funktionen 43.7 die Verschwindungsmenge $V(g)$ lokal in \mathbf{a} .

Nach Voraussetzung besitzt $Dg(\mathbf{a})$ m linear unabhängige Spalten. Um die Notation einfach zu halten, zerlegen wir im Beweis \mathbb{R}^n als $\mathbb{R}^n = \mathbb{R}^{n-m} \times \mathbb{R}^m$ mit Koordinaten $\mathbf{x} = (\mathbf{u}, \mathbf{v}) = (\mathbf{u}_1, \dots, \mathbf{u}_{n-m}, \mathbf{v}_1, \dots, \mathbf{v}_m)$ und nehmen ohne Einschränkung an, daß

$$\text{rang}(D_v g(\mathbf{a})) = \text{rang}(Dg(\mathbf{a})) = m$$

gilt, d.h. die letzten m Spalten von $Dg(\mathbf{a})$ sind linear unabhängig. Dann gilt aber

$$\det(D_v g(\mathbf{a})) \neq 0.$$

Wir können den Satz über implizite Funktionen 43.7 also auf g anwenden und damit $V(g)$ lokal in \mathbf{a} parametrisieren. D.h. Satz 43.7 liefert uns eine stetig differenzierbare Funktion

$$\varphi : U_\varepsilon(\mathbf{u}_a) \rightarrow U_r(\mathbf{v}_a)$$

mit $\varphi(\mathbf{u}_a) = \mathbf{v}_a$,

$$D\varphi(\mathbf{u}_a) = -(D_v g(\mathbf{a}))^{-1} \circ D_u g(\mathbf{a}) \quad (169)$$

und

$$g(\mathbf{u}, \varphi(\mathbf{u})) = 0 \quad (170)$$

für alle $\mathbf{u} \in U_\varepsilon(\mathbf{u}_a)$, wobei $\mathbf{u}_a = (\mathbf{a}_1, \dots, \mathbf{a}_{n-m})$ und $\mathbf{v}_a = (\mathbf{a}_{n-m+1}, \dots, \mathbf{a}_n)$.

Nach Voraussetzung hat f in \mathbf{a} ein lokales Extremum unter der Nebenbedingung $g = 0$, und wir nehmen ohne Einschränkung an, daß es ein lokales Maximum ist. Dann gibt es ein $\delta > 0$ mit

$$f(\mathbf{x}) \leq f(\mathbf{a}) \quad (171)$$

für alle $\mathbf{x} \in V(g) \cap U_\delta(\mathbf{a})$. Da die Abbildung

$$p : U_\varepsilon(\mathbf{u}_a) \rightarrow \mathbb{R}^n : \mathbf{u} \mapsto (\mathbf{u}, \varphi(\mathbf{u}))$$

stetig ist, gibt es eine kleine Umgebung von \mathbf{u}_a , die ganz in die δ -Umgebung $\mathbf{U}_\delta(\mathbf{a})$ von $\mathbf{p}(\mathbf{u}_a) = \mathbf{a}$ abgebildet wird. Indem wir ε verkleinern können wir ohne Einschränkung annehmen, daß dies schon für ganz $\mathbf{U}_\varepsilon(\mathbf{u}_a)$ gilt. Unter Berücksichtigung von (170) heißt das

$$(\mathbf{u}, \varphi(\mathbf{u})) = \mathbf{p}(\mathbf{u}) \in \mathbf{U}_\delta(\mathbf{a}) \cap \mathbf{V}(\mathbf{g}) \quad (172)$$

für alle $\mathbf{u} \in \mathbf{U}_\varepsilon(\mathbf{u}_a)$. Wir betrachten nun die stetig differenzierbare Funktion

$$\mathbf{q} : \mathbf{U}_\varepsilon(\mathbf{u}_a) \longrightarrow \mathbb{R} : \mathbf{u} \mapsto f(\mathbf{u}, \varphi(\mathbf{u})).$$

Wegen (171) und (172) gilt

$$\mathbf{q}(\mathbf{u}) \leq \mathbf{q}(\mathbf{u}_a)$$

für alle $\mathbf{u} \in \mathbf{U}_\varepsilon(\mathbf{u}_a)$, so daß \mathbf{q} eine Extremstelle in \mathbf{u}_a hat. Nach Proposition 42.24 ist \mathbf{u}_a somit ein kritischer Punkt von \mathbf{q} , und wir erhalten mit der Kettenregel

$$\begin{aligned} 0 &= D\mathbf{q}(\mathbf{u}_a) = Df(\mathbf{p}(\mathbf{u}_a)) \circ D\mathbf{p}(\mathbf{u}_a) \\ &= D_{\mathbf{u}}f(\mathbf{a}) \circ \mathbf{1}_{n-m} + D_{\mathbf{v}}f(\mathbf{a}) \circ D\varphi(\mathbf{u}_a) \\ &\stackrel{(169)}{=} D_{\mathbf{u}}f(\mathbf{a}) - D_{\mathbf{v}}f(\mathbf{a}) \circ (D_{\mathbf{v}}\mathbf{g}(\mathbf{a}))^{-1} \circ D_{\mathbf{u}}\mathbf{g}(\mathbf{a}). \end{aligned} \quad (173)$$

Wir setzen nun

$$\mathbf{b}^t := D_{\mathbf{v}}f(\mathbf{a}) \circ (D_{\mathbf{v}}\mathbf{g}(\mathbf{a}))^{-1} \quad (174)$$

und erhalten damit einerseits aus (174)

$$D_{\mathbf{v}}f(\mathbf{a}) = \mathbf{b}^t \circ D_{\mathbf{v}}\mathbf{g}(\mathbf{a}) = \sum_{j=1}^m \mathbf{b}_j \cdot D_{\mathbf{v}}\mathbf{g}_j(\mathbf{a})$$

und andererseits aus (173) und (174)

$$D_{\mathbf{u}}f(\mathbf{a}) = \mathbf{b}^t \circ D_{\mathbf{u}}\mathbf{g}(\mathbf{a}) = \sum_{j=1}^m \mathbf{b}_j \cdot D_{\mathbf{u}}\mathbf{g}_j(\mathbf{a}).$$

Insgesamt ist damit

$$Df(\mathbf{a}) = \sum_{j=1}^m \mathbf{b}_j \cdot D\mathbf{g}_j(\mathbf{a})$$

gezeigt. □

Bemerkung 43.23 (Lagrange-Multiplikatoren)

Wir setzen die Bezeichnungen und Bedingungen aus Satz 43.22 voraus.

- a. Damit f in \mathbf{a} eine Extremstelle unter der Nebenbedingung $\mathbf{g} = 0$ hat, muß es wegen Satz 43.22 ein $\mathbf{b} \in \mathbb{R}^m$ geben, so daß $(\mathbf{a}, \mathbf{b})^t \in \mathbb{R}^{n+m}$ eine Lösung der $n + m$ Gleichungen

$$\begin{aligned} D_i f(\mathbf{a}) - \sum_{j=1}^m \mathbf{b}_j \cdot D_i \mathbf{g}_j(\mathbf{a}) &= 0, \quad i = 1, \dots, n \\ \mathbf{g}_j(\mathbf{a}) &= 0, \quad j = 1, \dots, m \end{aligned} \quad (175)$$

ist. Sucht man also Kandidaten für den Punkt \mathbf{a} , so sollte man versuchen, dieses Gleichungssystem zu lösen. Man beachte, daß es in aller Regel ein nicht-lineares Gleichungssystem ist.

- b. Daß $(\mathbf{a}, \mathbf{b})^t$ das Gleichungssystem (175) löst, ist eine *notwendige* Bedingung dafür, daß f in \mathbf{a} eine Extremstelle unter der Nebenbedingung $\mathbf{g} = \mathbf{0}$ hat. Es ist aber weder garantiert, daß f in \mathbf{a} eine Extremstelle hat, noch läßt der Satz Rückschlüsse darauf zu, ob es sich um ein Maximum oder ein Minimum handelt. Hierzu sind in aller Regel weitere Untersuchungen von f nötig.
- c. Der Beweis von Satz 43.22 gibt zu gegebenem $\mathbf{a} \in \mathbf{U}$ konkret den Vektor \mathbf{b} als

$$\mathbf{b} = \left(D_{\mathbf{v}}f(\mathbf{a}) \circ (D_{\mathbf{v}}\mathbf{g}(\mathbf{a}))^{-1} \right)^t$$

an, wenn $\mathbf{v} = (x_{i_1}, \dots, x_{i_m})$ die Koordinaten bezeichnet, die zu m linear unabhängigen Spalten von $D\mathbf{g}(\mathbf{a})$ gehören. Für die Anwendungen des Satzes ist dies aber irrelevant, da wir in aller Regel \mathbf{a} nicht kennen, sondern suchen, indem wir die Gleichungen (175) für \mathbf{a} und \mathbf{b} lösen.

- d. Man kann die Aussage von Satz 43.22 auch etwas anders formulieren, wenn man die Hilfsfunktion

$$\mathbf{h} : \mathbf{U} \times \mathbb{R}^m \longrightarrow \mathbb{R} : (\mathbf{x}, \mathbf{y})^t \mapsto f(\mathbf{x}) - \sum_{j=1}^m y_j \cdot g_j(\mathbf{x})$$

betrachtet. Bis auf das Vorzeichen sind die linken Seiten von (175) genau die partiellen Ableitungen von \mathbf{h} .

Hat also f in \mathbf{a} eine Extremstelle unter der Nebenbedingung $\mathbf{g} = \mathbf{0}$, so gibt es ein $\mathbf{b} \in \mathbb{R}^m$, so daß $(\mathbf{a}, \mathbf{b})^t$ ein kritischer Punkt von \mathbf{h} ist. Dies ist eine in der Literatur gebräuchliche Formulierung des Satzes.

Beispiel 43.24 (Lagrange-Multiplikatoren)

Wir betrachten die lineare Funktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : \mathbf{x} \mapsto x_1 - 2x_2$$

und die Nebenbedingung

$$\mathbf{g} : \mathbb{R}^2 \longrightarrow \mathbb{R} : \mathbf{x} \mapsto \|\mathbf{x}\|_2^2 - 1 = x_1^2 + x_2^2 - 1,$$

so daß

$$V(\mathbf{g}) = \{\mathbf{x} \in \mathbb{R}^2 \mid \mathbf{g}(\mathbf{x}) = 0\} = \overline{\partial \mathbf{U}_1(\mathbf{0})}$$

der Einheitskreis um den Ursprung ist. Unser Ziel ist es, die Extremstellen von f unter der Nebenbedingung $\mathbf{g} = \mathbf{0}$ zu finden, d.h. die Extremstellen von f eingeschränkt auf den Einheitskreis.

Wegen (175) müssen wir dazu das Gleichungssystem

$$\begin{aligned} 0 &= D_1 f(x) - b \cdot D_1 g(x) = 1 - b \cdot 2x_1 \\ 0 &= D_2 f(x) - b \cdot D_2 g(x) = -2 - b \cdot 2x_2 \\ 0 &= g(x) = x_1^2 + x_2^2 - 1 \end{aligned}$$

lösen. Dabei müssen wir die Nebenbedingung

$$\text{rang}(Dg(x)) = \text{rang} \begin{pmatrix} 2x_1 & 2x_2 \end{pmatrix} = 1$$

berücksichtigen, die aber für alle $x \neq (0,0)^t$ erfüllt ist, und $x = (0,0)^t$ erfüllt die letzte der Gleichungen schon nicht, wird also keine Lösung sein.

Aus der ersten Gleichung sehen wir sofort, daß $b \neq 0$ gelten muß, so daß wir die ersten beiden Gleichungen nach x_1 und x_2 auflösen können,

$$x_1 = \frac{1}{2b}$$

und

$$x_2 = -\frac{1}{b}.$$

Um nun b zu bestimmen, setzen wir x_1 und x_2 in die dritte Gleichung ein und erhalten

$$\frac{1}{4b^2} + \frac{1}{b^2} - 1 = 0$$

oder alternativ

$$b^2 = \frac{5}{4}.$$

Wir erhalten also zwei Kandidaten für Extremstellen von f unter der Nebenbedingung $g = 0$, nämlich für $b = -\frac{\sqrt{5}}{2}$ den Punkt

$$\mathbf{a}^{(1)} = \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}} \right)$$

und für $b = \frac{\sqrt{5}}{2}$ den Punkt

$$\mathbf{a}^{(2)} = \left(\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}} \right).$$

Um nun festzustellen, ob die beiden Kandidaten für Extremstellen auch wirklich welche sind, benötigen wir eine Zusatzüberlegung. Die Funktion

$$f|_V : V(g) \longrightarrow \mathbb{R}$$

ist stetig und die Menge

$$V(g) = g^{-1}(0)$$

ist kompakt, mithin nimmt die Funktion $f|_V$ auf $V(g)$ ihr (globales) Maximum und ihr (globales) Minimum an, siehe Korollar 40.24. Als Kandidaten dafür kommen nur die beiden Punkte $\mathbf{a}^{(1)}$ und $\mathbf{a}^{(2)}$ in Frage, und wegen

$$f(\mathbf{a}^{(1)}) = -\sqrt{5} < \sqrt{5} = f(\mathbf{a}^{(2)})$$

folgt dann, daß $\mathbf{a}^{(1)}$ ein globales Minimum von f unter der Nebenbedingung $g = 0$ ist, während $\mathbf{a}^{(2)}$ ein globales Maximum von f unter der Nebenbedingung $g = 0$ ist.

Bemerkung 43.25

In Beispiel 43.24 hätte man in der Tat auch auf die Lagrange-Multiplikatoren verzichten können, da man die Kreislinie durch

$$\psi : \mathbb{R} \longrightarrow \mathbb{R}^2 : t \mapsto (\cos(t), \sin(t))^t$$

auch global parametrisieren kann. Wenn wir dann die Funktion

$$h = f \circ \psi : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto \cos(t) - 2 \cdot \sin(t)$$

auf Extremstellen untersuchen, müssen wir

$$h'(t) = -\sin(t) - 2 \cdot \cos(t) = 0$$

lösen, was zur Gleichung

$$\cot(t) = -\frac{1}{2}$$

führt. Da wir die Parametrisierung ψ ohnehin nur auf dem Intervall $[0, 2\pi)$ zu betrachten brauchen, um die Kreislinie abzudecken, führt diese Gleichung zu zwei zulässigen Werten für t , nämlich

$$t = \operatorname{arccot}\left(-\frac{1}{2}\right)$$

und

$$t = \operatorname{arccot}\left(-\frac{1}{2}\right) + \pi.$$

Man sieht diesen Werten aber nicht mehr so leicht an, daß sie unter ψ auf die Punkte $\left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$ und $\left(\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}}\right)$ abgebildet werden. Insofern liefern die Lagrange-Multiplikatoren hier ein befriedigenderes Ergebnis.

Die Kreislinie besitzt übrigens noch eine weitere Parametrisierung, wenn man den Punkt $(0, 1)^t$ wegläßt, nämlich

$$\psi : \mathbb{R} \longrightarrow \mathbb{R}^2 : t \mapsto \left(\frac{t^2-1}{t^2+1}, \frac{2t}{t^2+1}\right)^t.$$

Wir überlassen es dem Leser, mit dieser Parametrisierung die lokalen Extremstellen von f unter der Nebenbedingung $g = 0$ zu finden.

Aufgaben**Aufgabe 43.26** (Kugelkoordinaten)

Zeige, daß die Abbildung (siehe auch Beispiel 47.14)

$$\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (r, \theta, \vartheta) \mapsto (r \cdot \cos(\theta) \cdot \cos(\vartheta), r \cdot \sin(\theta) \cdot \cos(\vartheta), r \cdot \sin(\vartheta))$$

auf der offenen Menge

$$U = (0, \infty) \times (0, 2\pi) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

ein Diffeomorphismus ist mit Funktionaldeterminante

$$\det(D\varphi(r, \theta, z)) = r^2 \cdot \cos(\vartheta) > 0.$$

Aufgabe 43.27 (Zylinderkoordinaten)

Zeige, daß die Abbildung (siehe auch Beispiel 47.13)

$$\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (r, \theta, z) \mapsto (r \cdot \cos(\theta), r \cdot \sin(\theta), z)$$

auf der offenen Menge

$$U = (0, \infty) \times (0, 2\pi) \times \mathbb{R}$$

ein Diffeomorphismus ist mit Funktionaldeterminante

$$\det(D\varphi(r, \theta, z)) = r > 0.$$

Aufgabe 43.28

Finde zu $a > 0$ ein Intervall $[c, d]$, so daß die Funktion $f : [c, d] \longrightarrow [c, d] : x \mapsto \frac{x}{2} + \frac{a}{2x}$ eine strikte Kontraktion ist und bestimme den Fixpunkt von f .

Aufgabe 43.29

Sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ eine q -Kontraktion. Zeige mit Hilfe des Banachschen Fixpunktsatzes, daß dann die folgende Funktion surjektiv ist:

$$g : \mathbb{R}^2 \longrightarrow \mathbb{R}^2, (x_1, x_2)^t \longmapsto (x_1 + f(x_2), x_2 + f(x_1))^t.$$

Aufgabe 43.30

Zeige, daß die Funktion $f : \mathbb{R}^2 \longrightarrow \mathbb{R}, (x, y)^t \longmapsto (y - x^2) \cdot (y - 2x^2)$ keine Extremstelle hat, daß aber für jede Gerade $G \subseteq \mathbb{R}^2$ durch $(0, 0)^t$ die Funktion $f|_G$ ein isoliertes lokales Minimum in $(0, 0)^t$ besitzt.

Aufgabe 43.31

Zeige, daß die Verschwindungsmenge $V(f)$ für

$$f : \mathbb{R}^4 \longrightarrow \mathbb{R}^2 : (x_1, x_2, y_1, y_2)^t \mapsto (x_1^2 + x_2^2 - 2y_1y_2, x_1^3 + x_2^3 + y_1^3 - y_2^3)^t$$

lokal in $(-1, 1, 1, 1)$ als Graph einer Abbildung $\varphi : U_\varepsilon(-1, 1) \longrightarrow U_r(1, 1)$ darstellbar ist und berechne $D\varphi(-1, 1)$.

Aufgabe 43.32

Berechne die lokalen Extrema der Funktion $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto x_1^2 + x_2^2 + x_2$ unter der Nebenbedingung $(x_1 - 1)^2 + x_2^2 = 5$.

Aufgabe 43.33

Es seien $p, q \in \mathbb{R}_{>1}$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Bestimme die Extremstellen der Funktion

$$f : \mathbb{R}_{>0}^2 \longrightarrow \mathbb{R} : x \mapsto \frac{x_1^p}{p} + \frac{x_2^q}{q}$$

unter der Nebenbedingung $x_1x_2 = 1$. Folgere daraus für $u, v > 0$ die Höldersche Ungleichung $\frac{u^p}{p} + \frac{v^q}{q} \geq uv$.

Aufgabe 43.34

Es sei $U \subseteq \mathbb{R}^n$ offen und beschränkt und $f : \bar{U} \longrightarrow \mathbb{R}^n$ sei stetig auf \bar{U} und stetig differenzierbar auf U . Ferner sei $y \in \mathbb{R}^n$ so, daß $f^{-1}(y) \subseteq U$ und $\det(Df(x)) \neq 0$ für alle $x \in f^{-1}(y)$. Zeige, daß $f^{-1}(y)$ nur endlich viele Punkte enthält.

Aufgabe 43.35 (Spektralsatz für symmetrische Matrizen)

Sei $A \in \text{Mat}_n(\mathbb{R})$ eine symmetrische Matrix und $f : \mathbb{R}^n \rightarrow \mathbb{R} : x \mapsto x^t \circ A \circ x = \langle Ax, x \rangle$. Für $Y \subseteq \mathbb{R}^n$ setzen wir

$$M(Y) = \{x \in \mathbb{R}^n \mid \|x\|_2 = 1, \langle x, y \rangle = 0 \forall y \in Y\}.$$

- a. Zeige, für $i = 1, \dots, n$ gibt es $y_i \in M(y_1, \dots, y_{i-1})$ mit

$$\lambda_i := f(y_i) \stackrel{!}{=} \max_{x \in M(y_1, \dots, y_{i-1})} f(x).$$

- b. Zeige, es gilt $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$.
 c. Zeige $Ay_i = \lambda_i y_i$ für $i = 1, \dots, n$ mit Hilfe der Lagrange Multiplikatoren.
 d. Bestimme alle lokalen Extrema von f unter der Nebenbedingung $\|x\|_2 = 1$.

Aufgabe 43.36

Bestimme für die Abbildung

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}^2 : x \mapsto (x_1, x_2^2 - 2x_1x_2)^t$$

alle Punkte $y \in \mathbb{R}^2$, so daß $\det(Df(x)) \neq 0$ für alle $x \in f^{-1}(y)$.

Aufgabe 43.37

Welcher Punkt in $V(x_1^2 + x_2^2 - x_3)$ hat den kleinsten Abstand vom Punkt $(1, 1, \frac{1}{2})^t$?

§ 44 Das Riemann-Integral über n -dimensionalen Quadern

Wir wollen in diesem Abschnitt das eindimensionale Riemann-Integral auf *beschränkte* Funktionen in mehreren Veränderlichen über n -dimensionalen Quadern verallgemeinern. Die Theorie funktioniert im wesentlichen eins-zu-eins wie im eindimensionalen Fall.

A) Zerlegungen von n -dimensionalen Quadern

Definition 44.1

Es seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ mit $\mathbf{a} < \mathbf{b}$, d.h. $a_i < b_i$ für alle $i = 1, \dots, n$.

Wir nennen $[\mathbf{a}, \mathbf{b}] := [a_1, b_1] \times \dots \times [a_n, b_n]$ einen n -dimensionalen *Quader* oder ein n -dimensionales *Intervall* (siehe auch Beispiel 39.45), und wir nennen

$$V([\mathbf{a}, \mathbf{b}]) = (b_1 - a_1) \cdot \dots \cdot (b_n - a_n)$$

das *Volumen* des Quaders $[\mathbf{a}, \mathbf{b}]$.

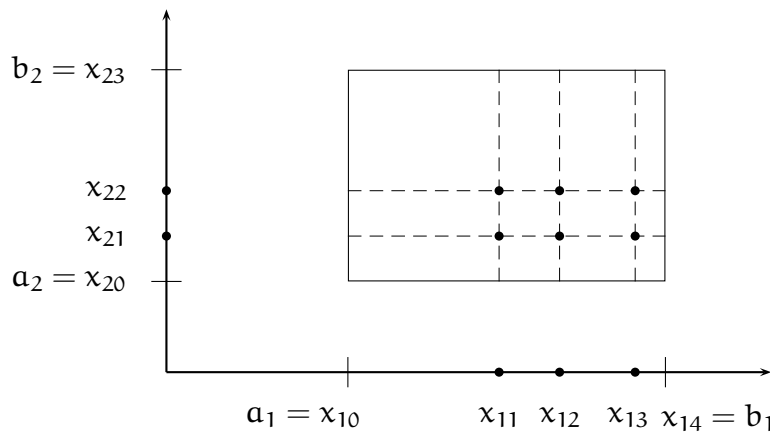


ABBILDUNG 12. Zerlegung eines 2-dimensionalen Quaders

Ist $Z_i = (x_{i0}, \dots, x_{im_i})$ eine Zerlegung des Intervalls $[a_i, b_i]$ für $i = 1, \dots, n$, so nennen wir

$$Z = ((x_{1j_1}, \dots, x_{nj_n}) \mid j_i = 0, \dots, m_i, i = 1, \dots, n)$$

eine *Zerlegung* von $[\mathbf{a}, \mathbf{b}]$, und der Einfachheit halber schreiben wir auch

$$Z = Z_1 \times \dots \times Z_n.$$

Die Punkte in Z heißen die *Stützpunkte* der Zerlegung, und

$$\text{supp}(Z) = \text{supp}(Z_1) \times \dots \times \text{supp}(Z_n)$$

heißt ihr *Träger*. Wir nennen

$$l(Z) := \max_{i=1, \dots, n} l(Z_i)$$

die *Länge* oder *Feinheit* von Z und

$$|Z| := |Z_1| \cdot \dots \cdot |Z_n| = m_1 \cdot \dots \cdot m_n$$

ihre *Mächtigkeit*. Die $|Z|$ Quader

$$[x_{1j_1-1}, x_{1j_1}] \times \dots \times [x_{nj_n-1}, x_{nj_n}],$$

$j_i = 1, \dots, m_i$, $i = 1, \dots, n$, heißen die *Teilquader* oder *Teilintervalle* der Zerlegung Z , und mit $\text{TQ}(Z)$ bezeichnen wir die *Menge aller Teilquader* von Z .

Eine zweite Zerlegung Z' von $[a, b]$ heißt *Verfeinerung* von Z , falls $\text{supp}(Z) \subseteq \text{supp}(Z')$ gilt.

Zu zwei Zerlegungen $Z = Z_1 \times \dots \times Z_n$ und $Z' = Z'_1 \times \dots \times Z'_n$ definieren wir

$$Z * Z' := (Z_1 * Z'_1) \times \dots \times (Z_n * Z'_n).$$

Sind Z und Z' Zerlegungen des gleichen Quaders, so nennen wir $Z * Z'$ ihre *gemeinsame Verfeinerung*.

Beispiel 44.2

Für $a = (0, 0)$ und $b = (4, 3)$ ist $[a, b] = \{(x, y) \mid 0 \leq x \leq 4, 0 \leq y \leq 3\}$. Betrachten wir nun die Zerlegung $Z_1 = (0, 2, 3, 4)$ und $Z_2 = (0, 1, 3)$, so hat die Zerlegung

$$Z = Z_1 \times Z_2 = ((0, 0), (0, 1), (0, 3), (2, 0), (2, 1), (2, 3), (3, 0), (3, 1), (3, 3), (4, 0), (4, 1), (4, 3))$$

den Träger

$$\text{supp}(Z) = \{(0, 0), (0, 1), (0, 3), (2, 0), (2, 1), (2, 3), (3, 0), (3, 1), (3, 3), (4, 0), (4, 1), (4, 3)\},$$

die Länge

$$l(Z) = \max\{l(Z_1), l(Z_2)\} = \max\{2, 2\} = 2$$

und die Teilquader

$$\text{TQ}(Z) = \{[0, 2] \times [0, 1], [0, 2] \times [1, 3], [2, 3] \times [0, 1], [2, 3] \times [1, 3], [3, 4] \times [0, 1], [3, 4] \times [1, 3]\}.$$

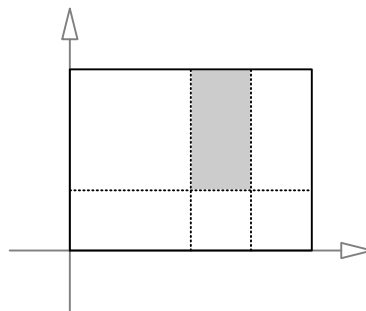


ABBILDUNG 13. Zerlegung von $[0, 4] \times [0, 3]$ mit markiertem Teilquader

Beispiel 44.3 (Äquidistante Zerlegung)

Seien $a, b \in \mathbb{R}^n$ mit $a < b$ und sei zudem $m \geq 1$. Wir setzen dann

$$x_{ij} := a_i + j \cdot \frac{b_i - a_i}{m}$$

für $i = 1, \dots, n$ und $j = 0, \dots, m$. Dann ist

$$Z_i^m := (x_{i0}, \dots, x_{im})$$

eine Zerlegung von $[a_i, b_i]$ und

$$Z^m := Z_1^m \times \dots \times Z_n^m$$

ist eine Zerlegung von $[a, b]$, die wir die *m-te äquidistante Zerlegung* von $[a, b]$ nennen wollen. Z^m hat genau

$$|Z^m| = m^n$$

Teilquader und jeder Teilquader $Q \in \text{TQ}(Z^m)$ hat das gleiche Volumen

$$V(Q) = \frac{b_1 - a_1}{m} \cdot \dots \cdot \frac{b_n - a_n}{m} = \frac{V([a, b])}{m^n}.$$

Die Länge der Zerlegung Z^m ist

$$l(Z^m) = \max_{i=1, \dots, n} \frac{b_i - a_i}{m}.$$

Ist m ein Teiler von m' , so ist $Z^{m'}$ eine Verfeinerung von Z^m .

Bemerkung 44.4 (Verfeinerung von Zerlegungen)

- a. Fügt man in der Zerlegung eines eindimensionalen Intervalls *einen* Stützpunkt ein, so zerteilt man damit nur *ein einziges* Teilintervall und erhöht damit die Anzahl der Teilintervalle der Zerlegung um *eins*.

Im mehrdimensionalen wird die Situation komplizierter. Gehen wir von einer Zerlegung $Z = Z_1 \times \dots \times Z_n$ eines Quaders $[a, b] \subset \mathbb{R}^n$ wie in der Definition 44.1 aus, bei der die Zerlegung Z_i von $[a_i, b_i]$ aus m_i Teilintervallen besteht. Fügt man nun in einer Zerlegungen Z_i einen Stützpunkt y ein, etwa zwischen x_{ij-1} und x_{ij} , so zerteilt man damit automatisch die

$$\frac{m_1 \cdot \dots \cdot m_n}{m_i} = \frac{|Z|}{|Z_i|}$$

Teilquader, an denen das Intervall $[x_{ij-1}, x_{ij}]$ beteiligt war, in zwei Teile (siehe Abbildung 14) und die neue Zerlegung

$$Z' = Z_1 \times \dots \times Z_{i-1} \times (x_{i0}, \dots, x_{ij-1}, y, x_{ij}, \dots, x_{im_i}) \times Z_{i+1} \times \dots \times Z_n$$

besitzt deshalb genau diese Anzahl Teilquader mehr als Z , d.h.

$$|Z'| - |Z| = \frac{|Z|}{|Z_i|}. \quad (176)$$

- b. Im Eindimensionalen ist zudem die Feinheit $l(Z)$ einer Zerlegung Z eine obere Schranke für die Länge und damit das Volumen der Teilintervalle der Zerlegung. Im mehrdimensionalen ist $l(Z)$ immer noch eine obere Schranke für die Seitenlängen der Teilquader, deren Volumen ist dann aber durch $l(Z)^n$ beschränkt, d.h. für alle $Q \in \text{TQ}(Z)$ gilt

$$V(Q) \leq l(Z)^n. \quad (177)$$

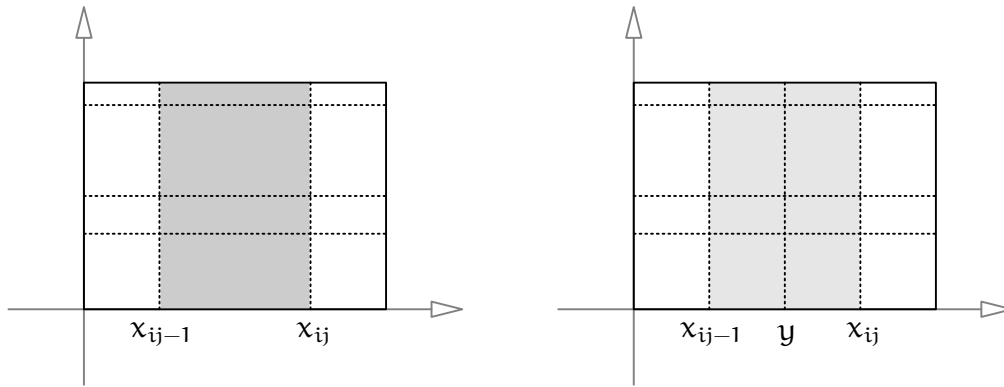


ABBILDUNG 14. Verfeinerung von Z um einen Stützpunkt in Z_i

- c. Die Zerlegung Z zerlegt den Quader $[a, b]$ in Teilquader, und das Volumen genügt offenbar der einfachen, aber sehr wichtigen Formel

$$V([a, b]) = \sum_{Q \in \text{TQ}(Z)} V(Q). \quad (178)$$

Will man diese Formel formal beweisen, sollte man dies am Besten durch Induktion nach

$$m_Z := |Z_1| + \dots + |Z_n|$$

tun, wobei die folgende einfache Beobachtung der Kern des Induktionsschrittes ist. Fügen wir in das Intervall $[a_i, b_i]$ einen Stützpunkt y ein, so zerteilen wir $Q = [a, b]$ in zwei neue Quader $Q = Q_1 \cup Q_2$ und für die Volumina gilt

$$\begin{aligned} V(Q_1) + V(Q_2) &= \prod_{j \neq i} (b_j - a_j) \cdot (b_i - y) + \prod_{j \neq i} (b_j - a_j) \cdot (y - a_i) \\ &= \prod_{j \neq i} (b_j - a_j) \cdot (b_i - a_i) = V(Q). \end{aligned}$$

B) Untersummen, Obersummen und das Riemann-Integral

Definition 44.5

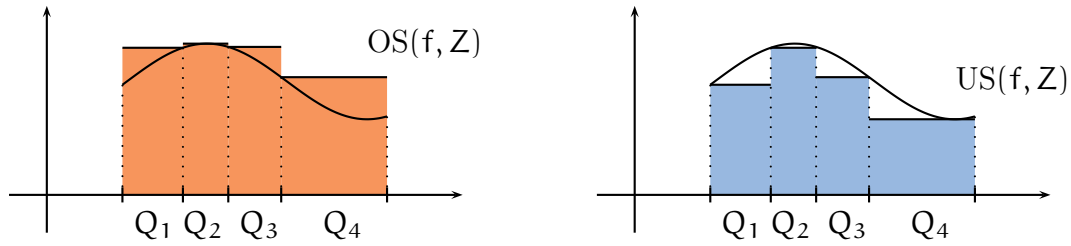
Seien $a, b \in \mathbb{R}^n$, $a < b$, $f : [a, b] \rightarrow \mathbb{R}$ beschränkt und Z eine Zerlegung von $[a, b]$.

Wir definieren die *Obersumme* von f bezüglich Z als

$$\text{OS}(f, Z) := \sum_{Q \in \text{TQ}(Z)} V(Q) \cdot \sup\{f(x) \mid x \in Q\},$$

und die *Untersumme* von f bezüglich Z als

$$\text{US}(f, Z) := \sum_{Q \in \text{TQ}(Z)} V(Q) \cdot \inf\{f(x) \mid x \in Q\}.$$

**Beispiel 44.6**

Wir betrachten die Funktion $f: \mathbb{R}^2 \rightarrow \mathbb{R}: x \mapsto x_1 + x_2$ auf dem Quader $[0, 2] \times [0, 2]$ mit der Zerlegung

$$Z = (0, 1, 2) \times (0, 1, 2).$$

Dann gilt

$$\text{TQ}(Z) = \{Q_1 = [0, 1] \times [0, 1], Q_2 = [0, 1] \times [1, 2], Q_3 = [1, 2] \times [0, 1], Q_4 = [1, 2] \times [1, 2]\}$$

und

$$\begin{aligned} \text{OS}(f, Z) &= V(Q_1) \cdot \sup_{x \in Q_1} f(x) + V(Q_2) \cdot \sup_{x \in Q_2} f(x) + V(Q_3) \cdot \sup_{x \in Q_3} f(x) + V(Q_4) \cdot \sup_{x \in Q_4} f(x) \\ &= 1 \cdot 2 + 1 \cdot 3 + 1 \cdot 3 + 1 \cdot 4 = 12 \end{aligned}$$

sowie

$$\begin{aligned} \text{US}(f, Z) &= V(Q_1) \cdot \inf_{x \in Q_1} f(x) + V(Q_2) \cdot \inf_{x \in Q_2} f(x) + V(Q_3) \cdot \inf_{x \in Q_3} f(x) + V(Q_4) \cdot \inf_{x \in Q_4} f(x) \\ &= 1 \cdot 0 + 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 2 = 4. \end{aligned}$$

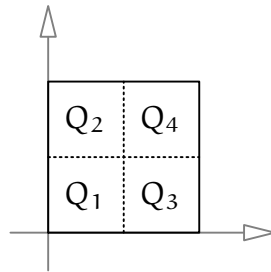


ABBILDUNG 15. Zerlegung von $[(0, 0)^t, (2, 2)^t] = [0, 2] \times [0, 2]$

Bemerkung 44.7

Fast alle Definitionen und Aussagen aus Abschnitt 19 übertragen sich nahezu wörtlich für Riemann-Integrale über n -dimensionalen Quadern. Dabei wird man Summen der Form

$$\sum_{i=1}^n (x_i - x_{i-1}) \cdot \dots$$

durch

$$\sum_{Q \in \text{TQ}(Z)} V(Q) \cdot \dots$$

ersetzen, d.h. insbesondere wird die Länge von Intervallen durch das Volumen der Quader ersetzt. In den meisten Beweisen ist dies die einzige Änderung die nötig

ist. An ein oder zwei Stellen muß man noch die Länge $l(Z)$ der Zerlegung durch ihre n -te Potenz $l(Z)^n$ ersetzen, und hin und wieder braucht man andere Zitate als im Eindimensionalen Fall. Aber das war's im wesentlichen dann auch schon. Der Vollständigkeit halber werden wir die Beweise mit den notwendigen kosmetischen Änderungen aber alle ausführen. Wir empfehlen dem Leser, sie mit den Beweisen im Eindimensionalen Fall zu vergleichen.

Lemma 44.8

Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ mit $\mathbf{a} < \mathbf{b}$ und $f : [\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ mit $|f(\mathbf{x})| \leq M$ für alle $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$.

a. Ist Z' eine Verfeinerung der Zerlegung Z von $[\mathbf{a}, \mathbf{b}]$, so gelten

$$0 \leq \text{US}(f, Z') - \text{US}(f, Z) \leq 2 \cdot M \cdot l(Z)^n \cdot (|Z'| - |Z|)$$

und

$$0 \leq \text{OS}(f, Z) - \text{OS}(f, Z') \leq 2 \cdot M \cdot l(Z)^n \cdot (|Z'| - |Z|).$$

Insbesondere gilt also

$$\text{US}(f, Z) \leq \text{US}(f, Z') \leq \text{OS}(f, Z') \leq \text{OS}(f, Z).$$

b. Für je zwei Zerlegungen Z und Z' von $[\mathbf{a}, \mathbf{b}]$ gilt

$$\text{US}(f, Z) \leq \text{OS}(f, Z').$$

c. Es gelten

$$-M \cdot V([\mathbf{a}, \mathbf{b}]) \leq \text{US}(f, Z) \leq \text{OS}(f, Z) \leq V([\mathbf{a}, \mathbf{b}]) \cdot M.$$

Beweis: Der Beweis geht wie im eindimensionalen Fall in Lemma 19.5.

a. Wir stellen zunächst folgende Vorüberlegung an. Ist $Q \subseteq [\mathbf{a}, \mathbf{b}]$ ein Quader und zerlegen wir ihn in zwei Quader $Q = Q_1 \cup Q_2$ indem wir eine Seite zweiteilen, so gilt

$$\inf_{\mathbf{x} \in Q} f(\mathbf{x}) \leq \inf_{\mathbf{x} \in Q_i} f(\mathbf{x})$$

und mithin

$$\begin{aligned} V(Q) \cdot \inf_{\mathbf{x} \in Q} f(\mathbf{x}) &= V(Q_1) \cdot \inf_{\mathbf{x} \in Q} f(\mathbf{x}) + V(Q_2) \cdot \inf_{\mathbf{x} \in Q} f(\mathbf{x}) \\ &\leq V(Q_1) \cdot \inf_{\mathbf{x} \in Q_1} f(\mathbf{x}) + V(Q_2) \cdot \inf_{\mathbf{x} \in Q_2} f(\mathbf{x}). \end{aligned}$$

Für die Differenz der beiden Seiten folgt damit

$$\begin{aligned} 0 &\leq V(Q_1) \cdot \inf_{\mathbf{x} \in Q_1} f(\mathbf{x}) + V(Q_2) \cdot \inf_{\mathbf{x} \in Q_2} f(\mathbf{x}) - V(Q) \cdot \inf_{\mathbf{x} \in Q} f(\mathbf{x}) \\ &= V(Q_1) \cdot \left(\inf_{\mathbf{x} \in Q_1} f(\mathbf{x}) - \inf_{\mathbf{x} \in Q} f(\mathbf{x}) \right) + V(Q_2) \cdot \left(\inf_{\mathbf{x} \in Q_2} f(\mathbf{x}) - \inf_{\mathbf{x} \in Q} f(\mathbf{x}) \right) \quad (179) \\ &\leq V(Q_1) \cdot (M + M) + V(Q_2) \cdot (M + M) = 2 \cdot M \cdot V(Q). \end{aligned}$$

Wir betrachten dann, ähnlich wie im Eindimensionalen, zunächst den Fall, daß die Zerlegung $Z = Z_1 \times \dots \times Z_n$ zu $Z' = Z_1 \times \dots \times Z'_i \times \dots \times Z_n$ verfeinert

wird, indem wir zu Z_i einen Stützpunkt hinzufügen. Aus (176) in Bemerkung 44.4 wissen wir, daß sich die Anzahl der Teilquader dadurch um

$$\frac{|Z|}{|Z_i|} = |Z'| - |Z|$$

erhöht. Zudem gilt $V(Q) \leq \mathfrak{l}(Z)^n$ wegen (177), so daß aus (179)

$$0 \leq US(f, Z') - US(f, Z) \leq 2 \cdot M \cdot \mathfrak{l}(Z)^n \cdot (|Z'| - |Z|) \quad (180)$$

folgt, da die Summanden in den Untersummen, die nicht zu den geteilten Teilquadern gehören, in beiden Summen gleich sind. Damit ist die erste Aussage zu Untersummen für diesen Spezialfall gezeigt. Den allgemeinen Fall führen wir darauf zurück.

Hat man eine beliebige Verfeinerung $Z' = Z'_1 \times \dots \times Z'_n$ von Z , so kann man von Z zu Z' gelangen, indem man sukzessive je einen Stützpunkt in einer der Komponenten hinzufügt. Seien Z^k für $k = 0, \dots, m$ die so entstehenden Zwischenzerlegungen mit $Z^0 = Z$ und $Z^m = Z'$. Wir erhalten aus (180) und wegen $\mathfrak{l}(Z^i) \leq \mathfrak{l}(Z)$ dann

$$\begin{aligned} 0 \leq US(f, Z') - US(f, Z) &= \sum_{i=1}^m US(f, Z^i) - US(f, Z^{i-1}) \\ &\leq \sum_{i=1}^m 2 \cdot M \cdot \mathfrak{l}(Z^{i-1})^n \cdot (|Z^i| - |Z^{i-1}|) \\ &\leq \sum_{i=1}^m 2 \cdot M \cdot \mathfrak{l}(Z)^n \cdot (|Z^i| - |Z^{i-1}|) \\ &= 2 \cdot M \cdot \mathfrak{l}(Z)^n \cdot \sum_{i=1}^m (|Z^i| - |Z^{i-1}|) \\ &= 2 \cdot M \cdot \mathfrak{l}(Z)^n \cdot (|Z'| - |Z|). \end{aligned}$$

Die Aussage für Obersummen zeigt man analog.

- b. Wir betrachten die gemeinsame Verfeinerung $Z * Z'$ und erhalten

$$\begin{aligned} US(f, Z) \stackrel{a.}{\leq} US(f, Z * Z') &= \sum_{Q \in \text{TQ}(Z * Z')} V(Q) \cdot \inf_{x \in Q} f(x) \\ &\leq \sum_{Q \in \text{TQ}(Z * Z')} V(Q) \cdot \sup_{x \in Q} f(x) = OS(f, Z * Z') \stackrel{a.}{\leq} OS(f, Z'). \end{aligned}$$

- c. Dies folgt aus a., da Z eine Verfeinerung der Zerlegung $(a_1, b_1) \times \dots \times (a_n, b_n)$ ist und da $M \geq \sup \{f(x) \mid x \in [a, b]\} \geq \inf \{f(x) \mid x \in [a, b]\} \geq -M$.

□

Da die Menge der Obersummen und die Menge der Untersummen nach Lemma 44.8 c. beschränkt sind, können wir ihr Infimum und ihr Supremum betrachten.

Definition 44.9 (Riemann-integrierbar)

Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ mit $\mathbf{a} < \mathbf{b}$ und sei $f : [\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ beschränkt.

Wir definieren das *Oberintegral*

$$OI(f) := \inf \{ OS(f, Z) \mid Z \text{ Zerlegung von } [\mathbf{a}, \mathbf{b}] \}$$

von f und das *Unterintegral*

$$UI(f) := \sup \{ US(f, Z) \mid Z \text{ Zerlegung von } [\mathbf{a}, \mathbf{b}] \}$$

von f . Wegen Lemma 44.8 b. und Lemma 8.19 gilt

$$UI(f) \leq OI(f).$$

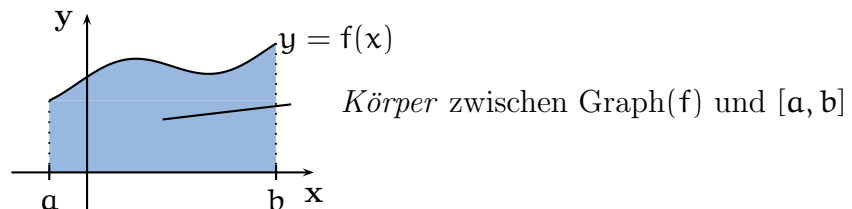
Wir nennen f (*Riemann-*)*integrierbar* auf $[\mathbf{a}, \mathbf{b}]$, falls $UI(f) = OI(f)$. Dann heißt

$$\int_{[\mathbf{a}, \mathbf{b}]} f(x) \, dx := OI(f) \in \mathbb{R}$$

das *Integral* von f auf $[\mathbf{a}, \mathbf{b}]$.

Bemerkung 44.10

Wenn die Funktion nur nicht-negative Werte annimmt, dann sind die Untersummen von f nach oben beschränkt durch das Volumen V des n -dimensionalen Körpers, den der Graph von f mit dem Quader $[\mathbf{a}, \mathbf{b}]$ einschließt, und die Obersummen von f sind durch dieses nach unten beschränkt. Aufgrund der Definition von $OI(f)$ als Infimum und $UI(f)$ als Supremum gilt also stets $UI(f) \leq V \leq OI(f)$. Daß f integrierbar ist, bedeutet mithin nichts anderes, als daß das Integral $\int_{[\mathbf{a}, \mathbf{b}]} f(x) \, dx$ das Volumen des Körpers beschreibt, den der Graph von f mit dem Quader $[\mathbf{a}, \mathbf{b}]$ einschließt.



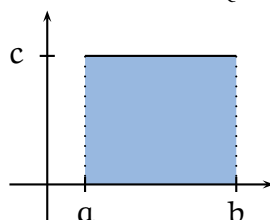
Beispiel 44.11

Jede konstante Funktion $f : Q \rightarrow \mathbb{R} : x \mapsto c$ auf einem Quader $Q \subseteq \mathbb{R}^n$ ist integrierbar mit

$$\int_Q f(x) \, dx = V(Q) \cdot c.$$

Denn dann gilt für jede Zerlegung Z von Q bereits

$$OS(f, Z) = \sum_{Q' \in \text{TQ}(Z)} V(Q') \cdot c = V(Q) \cdot c = \sum_{Q' \in \text{TQ}(Z)} V(Q') \cdot c = US(f, Z).$$



Insbesondere ist das Volumen

$$V(Q) = \int_Q 1 \, dx$$

von Q also der Wert des Integrals der Funktion 1 über Q .

C) Das Riemannsches Integrabilitätskriterium

Satz 44.12 (Riemannsches Integrabilitätskriterium)

Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ mit $\mathbf{a} < \mathbf{b}$ und sei $f : [\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ beschränkt.

Genau dann ist f integrierbar auf $[\mathbf{a}, \mathbf{b}]$, wenn

$$\forall \varepsilon > 0 \exists Z \text{ Zerlegung von } [\mathbf{a}, \mathbf{b}] : OS(f, Z) - US(f, Z) < \varepsilon.$$

Beweis: Der Beweis geht wörtlich wie im eindimensionalen Fall, siehe Satz 19.11.

\implies : Sei zunächst f integrierbar auf $[\mathbf{a}, \mathbf{b}]$ und sei $\varepsilon > 0$ gegeben. Aufgrund der Definition von $UI(f)$ als Supremum und $OI(f)$ als Infimum und wegen Proposition 8.18 gibt es Zerlegungen Z' und Z'' von $[\mathbf{a}, \mathbf{b}]$ mit

$$OI(f) + \frac{\varepsilon}{2} > OS(f, Z') \stackrel{44.8}{\geq} OS(f, Z' * Z'')$$

und

$$UI(f) - \frac{\varepsilon}{2} < US(f, Z'') \stackrel{44.8}{\leq} US(f, Z' * Z'').$$

Damit erhalten wir mit $Z = Z' * Z''$ und wegen $UI(f) = OI(f)$

$$OS(f, Z) - US(f, Z) < \left(OI(f) + \frac{\varepsilon}{2}\right) - \left(UI(f) - \frac{\varepsilon}{2}\right) = \varepsilon.$$

\impliedby : Für $\varepsilon := \frac{1}{n}$ mit $n \geq 1$ gibt es eine Zerlegung Z^n von $[\mathbf{a}, \mathbf{b}]$ mit

$$\frac{1}{n} > OS(f, Z^n) - US(f, Z^n) \geq OI(f) - UI(f) \geq 0.$$

Da die linke Seite der Ungleichung für $n \rightarrow \infty$ gegen Null konvergiert, folgt im Grenzwert

$$0 = \lim_{n \rightarrow \infty} \frac{1}{n} \geq OI(f) - UI(f) \geq 0,$$

also $OI(f) = UI(f)$. Mithin ist f integrierbar auf $[\mathbf{a}, \mathbf{b}]$. □

Satz 44.13 (Stetige Funktionen sind integrierbar.)

Ist $f : [\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ stetig für $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ mit $\mathbf{a} < \mathbf{b}$, so ist f integrierbar auf $[\mathbf{a}, \mathbf{b}]$.

Beweis: Der Beweis geht genau wie im eindimensionalen Fall, siehe Satz 19.12.

Der Quader $[\mathbf{a}, \mathbf{b}]$ ist nach Beispiel 39.45 kompakt, und die stetige Funktion f ist dann nach Satz 40.29 gleichmäßig stetig auf $[\mathbf{a}, \mathbf{b}]$ und nach Korollar 40.24 auch beschränkt.

Sei nun $\varepsilon > 0$ gegeben. Da f gleichmäßig stetig auf $[\mathbf{a}, \mathbf{b}]$ ist, gibt es $\delta_\varepsilon > 0$, so daß

$$|f(\mathbf{x}) - f(\mathbf{y})| < \frac{\varepsilon}{V([\mathbf{a}, \mathbf{b}])} \tag{181}$$

für alle $\mathbf{x}, \mathbf{y} \in [\mathbf{a}, \mathbf{b}]$ mit $\|\mathbf{x} - \mathbf{y}\|_\infty < \delta_\varepsilon$. Wir wählen nun eine Zerlegung Z von $[\mathbf{a}, \mathbf{b}]$ der Länge $l(Z) < \delta_\varepsilon$. Sei

$$TQ(Z) = \{Q_1, \dots, Q_m\}$$

die Menge der Teilquader von Z . Da f stetig auf Q_i ist, existieren nach Korollar 40.24 $\mathbf{y}_i, \mathbf{z}_i \in Q_i$ mit

$$f(\mathbf{y}_i) = \sup \{f(\mathbf{y}) \mid \mathbf{y} \in Q_i\}$$

und

$$f(\mathbf{z}_i) = \inf \{f(\mathbf{y}) \mid \mathbf{y} \in Q_i\},$$

und wegen

$$\|\mathbf{y}_i - \mathbf{z}_i\|_\infty = \max \{|\mathbf{y}_{ij} - \mathbf{z}_{ij}| \mid j = 1, \dots, n\} \leq l(Z) < \delta_\varepsilon$$

folgt aus (181) zudem

$$0 \leq f(\mathbf{y}_i) - f(\mathbf{z}_i) < \frac{\varepsilon}{V([\mathbf{a}, \mathbf{b}])}.$$

Damit erhalten wir insbesondere

$$\begin{aligned} OS(f, Z) - US(f, Z) &= \sum_{i=1}^m V(Q_i) \cdot (f(\mathbf{y}_i) - f(\mathbf{z}_i)) \\ &< \sum_{i=1}^m V(Q_i) \cdot \frac{\varepsilon}{V([\mathbf{a}, \mathbf{b}])} \stackrel{(178)}{=} V([\mathbf{a}, \mathbf{b}]) \cdot \frac{\varepsilon}{V([\mathbf{a}, \mathbf{b}])} = \varepsilon. \end{aligned}$$

Somit ist f integrierbar nach dem Riemannschem Integrierbarkeitskriterium 44.12. \square

D) Riemannsche Zwischensummen und das Folgenkriterium

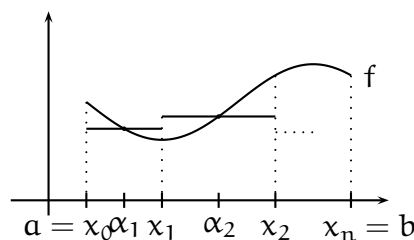
Definition 44.14

Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$, $\mathbf{a} < \mathbf{b}$, $f : [\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ beschränkt und Z eine Zerlegung von $[\mathbf{a}, \mathbf{b}]$.

Erfüllt $\alpha = (\alpha_Q \mid Q \in TQ(Z))$ die Bedingung $\alpha_Q \in Q$ für $Q \in TQ(Z)$, so heißt

$$ZS(f, Z, \alpha) := \sum_{Q \in TQ(Z)} V(Q) \cdot f(\alpha_Q)$$

die *Riemannsche Zwischensumme* von f bezüglich der Zerlegung Z und den *Zwischenpunkten* α . Wir bezeichnen mit $\mathcal{Z}([\mathbf{a}, \mathbf{b}])$ die Menge aller Paare (Z, α) mit Z Zerlegung von $[\mathbf{a}, \mathbf{b}]$ mit Zwischenpunkten α .



Das nächste Lemma sagt, daß man Obersummen und Untersummen beliebig gut approximieren kann durch Zwischensummen.

Lemma 44.15

Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt, $a, b \in \mathbb{R}^n$, $a < b$, Z eine Zerlegung von $[a, b]$, $\varepsilon > 0$.

- a. Dann gibt es Zwischenpunkte α von Z mit $0 \leq OS(f, Z) - ZS(f, Z, \alpha) < \varepsilon$.
- b. Dann gibt es Zwischenpunkte β von Z mit $0 \leq ZS(f, Z, \beta) - US(f, Z) < \varepsilon$.

Beweis: Der Beweis geht genau wie im eindimensionalen Fall.

Für $Q \in \mathcal{TQ}(Z)$ sei

$$M_Q := \sup \{ f(x) \mid x \in Q \}.$$

Aufgrund der Definition von M_Q als Supremum der Funktionswerte auf dem Quader Q gibt es ein $\alpha_Q \in Q$, so daß

$$f(\alpha_Q) > M_Q - \frac{\varepsilon}{V([a, b])}.$$

Damit erhalten wir für $\alpha = (\alpha_Q \mid Q \in \mathcal{TQ}(Z))$

$$\begin{aligned} OS(f, Z) - ZS(f, Z, \alpha) &= \sum_{Q \in \mathcal{TQ}(Z)} V(Q) \cdot (M_Q - f(\alpha_Q)) \\ &< \sum_{Q \in \mathcal{TQ}(Z)} V(Q) \cdot \frac{\varepsilon}{V([a, b])} \\ &\stackrel{(178)}{=} V([a, b]) \cdot \frac{\varepsilon}{V([a, b])} = \varepsilon. \end{aligned}$$

Damit ist a. gezeigt, und b. zeigt man analog. □

Das folgende Lemma sagt, daß für integrierbare Funktionen Untersummen und Obersummen beliebig nahe beieinander und damit beim Wert des Integrals liegen, wenn nur die Länge der Zerlegung hinreichend klein gewählt ist.

Lemma 44.16

Ist $f : [a, b] \rightarrow \mathbb{R}$ integrierbar für $a, b \in \mathbb{R}^n$ mit $a < b$, so gilt:

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : \forall Z \text{ Zerlegung mit } l(Z) < \delta_\varepsilon \text{ gilt } OS(f, Z) - US(f, Z) < \varepsilon.$$

Beweis: Der Beweis geht genau wie im Eindimensionalen, siehe Lemma 19.18.

Sei $\varepsilon > 0$ gegeben. Aus dem Riemannschen Integrabilitätskriterium erhalten wir eine Zerlegung Z' von $[a, b]$, so daß

$$OS(f, Z') - US(f, Z') < \frac{\varepsilon}{2}. \tag{182}$$

Wir setzen nun

$$\delta_\varepsilon := \min \left\{ \frac{\varepsilon}{8 \cdot |Z'| \cdot M}, 1 \right\} > 0,$$

wobei $M := \sup \{ |f(\mathbf{x})| \mid \mathbf{x} \in [\mathbf{a}, \mathbf{b}] \}$. Ist Z eine Zerlegung von $[\mathbf{a}, \mathbf{b}]$ mit $l(Z) < \delta_\varepsilon \leq 1$, so ist auch $l(Z)^n \leq l(Z) < \delta_\varepsilon$ und aus Lemma 44.8 und $|Z * Z'| - |Z| \leq |Z'|$ folgt

$$OS(f, Z) - OS(f, Z * Z') \leq 2 \cdot M \cdot l(Z)^n \cdot (|Z * Z'| - |Z|) < 2 \cdot M \cdot \delta_\varepsilon \cdot |Z'| = \frac{\varepsilon}{4} \quad (183)$$

und

$$US(f, Z * Z') - US(f, Z) \leq 2 \cdot M \cdot l(Z)^n \cdot (|Z * Z'| - |Z|) < 2 \cdot M \cdot \delta_\varepsilon \cdot |Z'| = \frac{\varepsilon}{4}. \quad (184)$$

Da $Z * Z'$ eine Verfeinerung von Z' ist, folgt aus (182) zusammen mit Lemma 19.5

$$OS(f, Z * Z') - US(f, Z * Z') \leq OS(f, Z') - US(f, Z') < \frac{\varepsilon}{2}. \quad (185)$$

Insgesamt erhalten wir damit

$$\begin{aligned} OS(f, Z) - US(f, Z) &= OS(f, Z) - OS(f, Z * Z') + OS(f, Z * Z') - US(f, Z * Z') \\ &\quad + US(f, Z * Z') - US(f, Z) \stackrel{(183)(185)(184)}{<} \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon. \end{aligned}$$

□

Satz 44.17 (Riemannsches Folgenkriterium für Integrierbarkeit)

Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ mit $\mathbf{a} < \mathbf{b}$, $f: [\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ sei beschränkt und $I \in \mathbb{R}$.

Genau dann ist f auf $[\mathbf{a}, \mathbf{b}]$ integrierbar mit $I = \int_{[\mathbf{a}, \mathbf{b}]} f(\mathbf{x}) \, d\mathbf{x}$, wenn für jede Folge $(Z^m, \alpha^m)_{m \in \mathbb{N}}$ von Zerlegungen von $[\mathbf{a}, \mathbf{b}]$ und Zwischenpunkten mit $l(Z^m) \rightarrow 0$ gilt

$$ZS(f, Z^m, \alpha^m) \rightarrow I.$$

Beweis: Der Beweis geht genau wie im Eindimensionalen, siehe Satz 19.19.

\implies : Es sei $(Z^m, \alpha^m)_{m \in \mathbb{N}}$ eine Folge von Zerlegungen von $[\mathbf{a}, \mathbf{b}]$ mit Zwischenpunkten, so daß $\lim_{m \rightarrow \infty} l(Z^m) = 0$, und sei $I = \int_{[\mathbf{a}, \mathbf{b}]} f(\mathbf{x}) \, d\mathbf{x}$.

Sei $\varepsilon > 0$ gegeben. Wir müssen ein $n_\varepsilon \in \mathbb{N}$ finden, so daß

$$|ZS(f, Z^m, \alpha^m) - I| < \varepsilon \quad (186)$$

für alle $m \geq n_\varepsilon$.

Da f integrierbar ist, gibt es nach Lemma 44.16 ein $\delta_\varepsilon > 0$, so daß für eine Zerlegung Z von $[\mathbf{a}, \mathbf{b}]$ aus $l(Z) < \delta_\varepsilon$ auch

$$OS(f, Z) - US(f, Z) < \varepsilon \quad (187)$$

gilt. Wegen $\lim_{m \rightarrow \infty} l(Z^m) = 0$ gibt es ein $n_\varepsilon \in \mathbb{N}$, so daß $l(Z^m) < \delta_\varepsilon$ für $m \geq n_\varepsilon$.

Für $m \geq n_\varepsilon$ leiten wir dann aus (187)

$$ZS(f, Z^m, \alpha^m) - I \leq OS(f, Z^m) - I \leq OS(f, Z^m) - US(f, Z^m) < \varepsilon$$

her, sowie

$$ZS(f, Z^m, \alpha^m) - I \geq US(f, Z^m) - I \geq US(f, Z^m) - OS(f, Z^m) > -\varepsilon.$$

Damit ist (186) für $m \geq n_\varepsilon$ erfüllt, und das heißt $ZS(f, Z^m, \alpha^m) \rightarrow I$.

⇐: Wir wollen das Riemannsche Integrabilitätskriterium anwenden.

Sei dazu $\varepsilon > 0$ gegeben. Wir betrachten die äquidistante Zerlegung Z^m aus Beispiel 44.3 für $m \geq 1$. Dann gilt

$$\lim_{m \rightarrow \infty} l(Z^m) = \lim_{m \rightarrow \infty} \frac{\max\{b_i - a_i \mid i = 1, \dots, n\}}{m} = 0.$$

Mit Lemma 44.15 finden wir zu $m \in \mathbb{N}$ Zwischenpunkte α^m und β^m , so daß

$$OS(f, Z^m) - ZS(f, Z^m, \alpha^m) < \frac{\varepsilon}{4}.$$

und

$$ZS(f, Z^m, \beta^m) - US(f, Z^m) < \frac{\varepsilon}{4}.$$

Außerdem gelten nach Voraussetzung

$$ZS(f, Z^m, \alpha^m) \rightarrow I$$

und

$$ZS(f, Z^m, \beta^m) \rightarrow I.$$

Wir finden also ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $m \geq n_\varepsilon$ gilt

$$|ZS(f, Z^m, \alpha^m) - I| < \frac{\varepsilon}{4}$$

und

$$|ZS(f, Z^m, \beta^m) - I| < \frac{\varepsilon}{4}.$$

Dann gilt aber auch

$$\begin{aligned} OS(f, Z^{n_\varepsilon}) - US(f, Z^{n_\varepsilon}) &= |OS(f, Z^{n_\varepsilon}) - US(f, Z^{n_\varepsilon})| \\ &\leq |OS(f, Z^{n_\varepsilon}) - ZS(f, Z^{n_\varepsilon}, \alpha^{n_\varepsilon})| + |ZS(f, Z^{n_\varepsilon}, \alpha^{n_\varepsilon}) - I| \\ &\quad + |I - ZS(f, Z^{n_\varepsilon}, \beta^{n_\varepsilon})| + |ZS(f, Z^{n_\varepsilon}, \beta^{n_\varepsilon}) - US(f, Z^{n_\varepsilon})| \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon. \end{aligned}$$

Also ist f integrierbar nach dem Riemannschen Integrabilitätskriterium 44.12.

Außerdem haben wir für $m \geq n_\varepsilon$

$$\begin{aligned} |OS(f, Z^m) - I| &\leq |OS(f, Z^m) - ZS(f, Z^m, \alpha^m)| + |ZS(f, Z^m, \alpha^m) - I| \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < \varepsilon, \end{aligned}$$

woraus

$$OS(f, Z^m) \rightarrow I$$

folgt. Analog sehen wir

$$US(f, Z^m) \rightarrow I.$$

Damit erhalten wir dann

$$I \leftarrow US(f, Z^m) \leq \int_{[a,b]} f(x) dx \leq OS(f, Z^m) \rightarrow I,$$

so daß $I = \int_{[a,b]} f(x) dx$ aus dem Einschachtelungssatz 11.17 folgt.

□

Korollar 44.18 (Riemannsches Zwischensummenkriterium)

Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ mit $\mathbf{a} < \mathbf{b}$, $f : [\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ sei beschränkt und $I \in \mathbb{R}$.

Genau dann ist f auf $[\mathbf{a}, \mathbf{b}]$ integrierbar mit $I = \int_{[\mathbf{a}, \mathbf{b}]} f(\mathbf{x}) \, d\mathbf{x}$, wenn

$$\forall \varepsilon > 0 \exists \delta_\varepsilon > 0 : \forall (Z, \alpha) \in \mathcal{Z}([\mathbf{a}, \mathbf{b}]) \text{ mit } \mathfrak{l}(Z) < \delta_\varepsilon : |\text{ZS}(f, Z, \alpha) - I| < \varepsilon.$$

Beweis:

\Rightarrow : Ist f integrierbar mit $I = \int_{[\mathbf{a}, \mathbf{b}]} f(\mathbf{x}) \, d\mathbf{x}$, so gibt es zu $\varepsilon > 0$ nach Lemma 44.16 ein $\delta_\varepsilon > 0$, so daß aus $\mathfrak{l}(Z) < \delta_\varepsilon$ stets

$$\text{OS}(f, Z) - \text{US}(f, Z) < \varepsilon$$

folgt. Für beliebige Zwischenpunkte von Z gilt dann aber

$$|\text{ZS}(f, Z, \alpha) - I| \leq \text{OS}(f, Z) - \text{US}(f, Z) < \varepsilon,$$

da sowohl die Zwischensumme, als auch das Integral zwischen der Ober- und der Untersumme liegen.

\Leftarrow : Sei $(Z^n, \alpha^n) \in \mathcal{Z}([\mathbf{a}, \mathbf{b}])$ mit $\mathfrak{l}(Z^n) \rightarrow 0$. Zu $\varepsilon > 0$ sei $\delta_\varepsilon > 0$ wie in der Voraussetzung gegeben. Dann gibt es zu $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$, so daß $\mathfrak{l}(Z^n) < \delta_\varepsilon$ für alle $n \geq n_\varepsilon$, und nach Voraussetzung gilt deshalb

$$|\text{ZS}(f, Z^n, \alpha^n) - I| < \varepsilon$$

für alle $n \geq n_\varepsilon$. Mithin konvergiert $\text{ZS}(f, Z^n, \alpha^n)$ gegen I , und wegen des Riemannschen Folgenkriteriums für Integrierbarkeit 44.17 ist f also integrierbar mit $I = \int_{[\mathbf{a}, \mathbf{b}]} f(\mathbf{x}) \, d\mathbf{x}$. \square

Korollar 44.19 (Linearität und Monotonie des Integrals)

Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ mit $\mathbf{a} < \mathbf{b}$, $f, g : [\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ seien integrierbar und $\mathbf{c}, \mathbf{d} \in \mathbb{R}$.

a. Dann ist $\mathbf{c} \cdot f + \mathbf{d} \cdot g$ integrierbar auf $[\mathbf{a}, \mathbf{b}]$ mit

$$\int_{[\mathbf{a}, \mathbf{b}]} (\mathbf{c} \cdot f + \mathbf{d} \cdot g)(\mathbf{x}) \, d\mathbf{x} = \mathbf{c} \cdot \int_{[\mathbf{a}, \mathbf{b}]} f(\mathbf{x}) \, d\mathbf{x} + \mathbf{d} \cdot \int_{[\mathbf{a}, \mathbf{b}]} g(\mathbf{x}) \, d\mathbf{x}.$$

b. Ist $f(\mathbf{x}) \leq g(\mathbf{x})$ für alle $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$, so ist auch

$$\int_{[\mathbf{a}, \mathbf{b}]} f(\mathbf{x}) \, d\mathbf{x} \leq \int_{[\mathbf{a}, \mathbf{b}]} g(\mathbf{x}) \, d\mathbf{x}.$$

Beweis: Der Beweis geht genau wie im Eindimensionalen, siehe Korollar 19.21.

a. Wir beachten zunächst, daß für jede Zerlegung Z von $[\mathbf{a}, \mathbf{b}]$ mit Zwischenpunkten α offenbar gilt:

$$\begin{aligned} \text{ZS}(\mathbf{c}f + \mathbf{d}g, Z, \alpha) &= \sum_{Q \in \text{TQ}(Z)} V(Q) \cdot (\mathbf{c}f + \mathbf{d}g)(\alpha_Q) \\ &= \mathbf{c} \cdot \sum_{Q \in \text{TQ}(Z)} V(Q) \cdot f(\alpha_Q) + \mathbf{d} \cdot \sum_{Q \in \text{TQ}(Z)} V(Q) \cdot g(\alpha_Q) \\ &= \mathbf{c} \cdot \text{ZS}(f, Z, \alpha) + \mathbf{d} \cdot \text{ZS}(g, Z, \alpha). \end{aligned}$$

Es sei nun $(Z^m, \alpha^m)_{m \in \mathbb{N}}$ eine Folge von Zerlegungen von $[a, b]$ und Zwischenpunkten mit $l(Z^m) \rightarrow 0$. Aus den Grenzwertsätzen für Folgen 11.15 und Satz 44.17 folgt dann

$$\begin{aligned} \lim_{m \rightarrow \infty} \text{ZS}(cf + dg, Z^m, \alpha^m) &= c \cdot \lim_{m \rightarrow \infty} \text{ZS}(f, Z^m, \alpha^m) + d \cdot \lim_{m \rightarrow \infty} \text{ZS}(g, Z^m, \alpha^m) \\ &= c \cdot \int_{[a,b]} f(x) \, dx + d \cdot \int_{[a,b]} g(x) \, dx. \end{aligned}$$

Das Riemannsche Folgenkriterium für Integrierbarkeit 44.17 liefert dann die Behauptung.

- b. Es sei $(Z^m, \alpha^m)_{m \in \mathbb{N}}$ eine Folge von Zerlegungen von $[a, b]$ und Zwischenpunkten mit $l(Z^m) \rightarrow 0$. Wegen $f(x) \leq g(x)$ für alle $x \in [a, b]$ gilt dann offenbar

$$\int_{[a,b]} f(x) \, dx \leftarrow \text{ZS}(f, Z^m, \alpha^m) \leq \text{ZS}(g, Z^m, \alpha^m) \rightarrow \int_{[a,b]} g(x) \, dx,$$

wobei die Grenzwerte aus dem Riemannschen Folgenkriterium für Integrierbarkeit folgen. Damit gilt dann aber auch für die Grenzwerte

$$\int_{[a,b]} f(x) \, dx \leq \int_{[a,b]} g(x) \, dx.$$

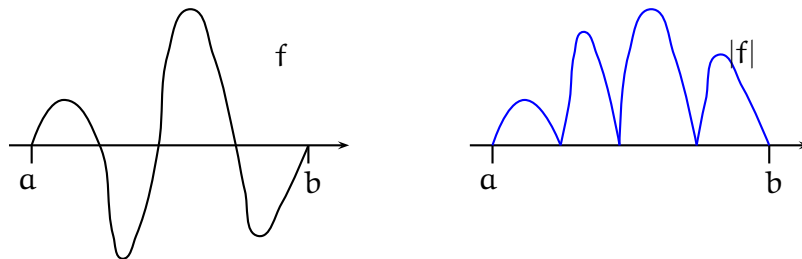
□

Proposition 44.20 (Dreiecksungleichung für Integrale)

Ist $f : [a, b] \rightarrow \mathbb{R}$ integrierbar auf $[a, b]$ für $a, b \in \mathbb{R}^n$ mit $a < b$, so ist $|f|$ integrierbar auf $[a, b]$, und es gilt

$$\left| \int_{[a,b]} f(x) \, dx \right| \leq \int_{[a,b]} |f(x)| \, dx.$$

Wir nennen das Integral über $|f|$ auch das Volumen, das der Graph von f mit dem Quader $[a, b]$ einschließt.

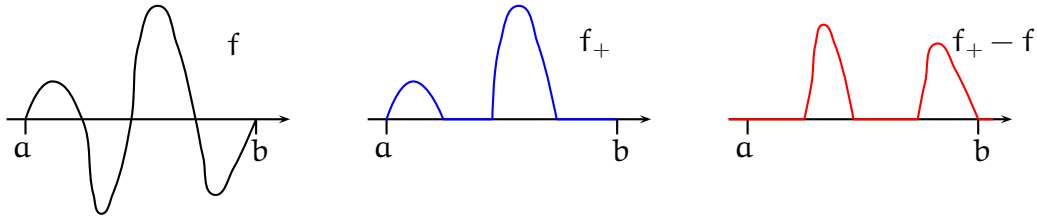


Beweis: Wir betrachten die Funktion

$$f_+ : [a, b] \rightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } f(x) \geq 0, \\ 0, & \text{falls } f(x) < 0. \end{cases}$$

Dann gilt

$$|f| = 2 \cdot f_+ - f.$$



Wir wollen nun mit Hilfe des Riemannsches Integrierbarkeitskriteriums zeigen, daß f_+ auf $[a, b]$ integrierbar ist. Sei $\varepsilon > 0$ gegeben. Da f auf $[a, b]$ integrierbar ist, gibt es eine Zerlegung Z von $[a, b]$, so daß

$$OS(f, Z) - US(f, Z) < \varepsilon.$$

Wir behaupten, daß für jede Teilmenge $Q \subseteq [a, b]$ die Ungleichung

$$\sup\{f_+(x) \mid x \in Q\} - \inf\{f_+(x) \mid x \in Q\} \leq \sup\{f(x) \mid x \in Q\} - \inf\{f(x) \mid x \in Q\}, \quad (188)$$

gilt. Dazu betrachten wir verschiedene Fälle.

1. Fall: $f(x) < 0$ für alle $x \in Q$: Dann gilt $f_+ \equiv 0$ auf Q , so daß die linke Seite in (188) Null ist. Zugleich gilt

$$\inf\{f(x) \mid x \in Q\} \leq \sup\{f(x) \mid x \in Q\} \leq 0,$$

so daß die rechte Seite von (188) nicht-negativ ist. In diesem Fall gilt (188).

2. Fall: $\exists y, z \in Q$ mit $f(y) < 0 \leq f(z)$: Also $\sup\{f_+(x) \mid x \in Q\} = \sup\{f(x) \mid x \in Q\}$ und $\inf\{f_+(x) \mid x \in Q\} = 0 > \inf\{f(x) \mid x \in Q\}$. Damit gilt die Ungleichung (188).

3. Fall: $f(x) \geq 0$ für alle $x \in Q$: Dann ist $f = f_+$ auf Q und (188) gilt.

Damit haben wir gezeigt, daß (188) stets erfüllt ist. Für die Differenz der Ober- und Untersumme von f_+ ergibt sich dann

$$\begin{aligned} OS(f_+, Z) - US(f_+, Z) &= \sum_{Q \in TQ(Z)} V(Q) \cdot (\sup\{f_+(x) \mid x \in Q\} - \inf\{f_+(x) \mid x \in Q\}) \\ &\leq \sum_{Q \in TQ(Z)} V(Q) \cdot (\sup\{f(x) \mid x \in Q\} - \inf\{f(x) \mid x \in Q\}) \\ &= OS(f, Z) - US(f, Z) < \varepsilon. \end{aligned}$$

Mit Hilfe des Riemannsches Integrierbarkeitskriteriums 44.12 folgt dann, daß f_+ auf $[a, b]$ integrierbar ist. Aus der Linearität des Integrals 44.19 folgt damit, daß auch

$$|f| = 2 \cdot f_+ - f$$

auf $[a, b]$ integrierbar ist.

Für eine Zerlegung Z mit Zwischenpunkten α gilt

$$|ZS(f, Z, \alpha)| = \left| \sum_{Q \in TQ(Z)} V(Q) \cdot f(\alpha_Q) \right| \leq \sum_{Q \in TQ(Z)} V(Q) \cdot |f(\alpha_Q)| = ZS(|f|, Z, \alpha).$$

Sei nun $(Z^m, \alpha^m)_{m \in \mathbb{N}}$ eine Folge von Zerlegungen von $[a, b]$ und Zwischenpunkten mit $\lim_{m \rightarrow \infty} l(Z^m) = 0$, dann folgt

$$\left| \int_{[a,b]} f(x) dx \right| \leftarrow |ZS(f, Z^m, \alpha^m)| \leq ZS(|f|, Z^m, \alpha^m) \longrightarrow \int_{[a,b]} |f(x)| dx.$$

Die Ungleichung bleibt für die Grenzwerte erhalten. \square

E) Der Satz von Fubini für Quader

Bislang haben wir noch kein gutes Verfahren kennen gelernt, das es uns wirklich erlauben würde, mehrdimensionale Riemann-Integrale auszurechnen. Wir wollen nun den Satz von Fubini beweisen, der es uns erlauben wird, ein mehrdimensionales Riemann-Integral auf die Berechnung mehrerer eindimensionaler Riemann-Integrale zurückzuführen.

Notation 44.21 (Vorbereitung des Satzes von Fubini)

Wir teilen die Koordinaten auf \mathbb{R}^{p+q} auf in $x = (x_1, \dots, x_p)$ und $y = (y_1, \dots, y_q)$.

Sind $X = [a, b] \subset \mathbb{R}^p$ und $Y = [c, d] \subset \mathbb{R}^q$ zwei Quader, so ist auch

$$X \times Y = [a, b] \times [c, d] = [(a, c), (b, d)] \subset \mathbb{R}^{p+q}$$

ein Quader, und es gilt

$$V(X \times Y) = \prod_{i=1}^p (b_i - a_i) \cdot \prod_{j=1}^q (d_j - c_j) = V(X) \cdot V(Y). \quad (189)$$

Ist $Z_x = Z_1 \times \dots \times Z_p$ eine Zerlegung von X und $Z_y = Z_{p+1} \times \dots \times Z_{p+q}$ eine Zerlegung von Y , so ist

$$Z_x \times Z_y = Z_1 \times \dots \times Z_{p+q}$$

eine Zerlegung von $X \times Y$, und es gilt

$$TQ(Z_x \times Z_y) = \{P \times Q \mid P \in TQ(Z_x), Q \in TQ(Z_y)\}. \quad (190)$$

Sind ferner $\alpha = (\alpha_P \mid P \in TQ(Z_x))$ Zwischenpunkte von Z_x und $\beta = (\beta_Q \mid Q \in TQ(Z_y))$ Zwischenpunkte von Z_y , so sind

$$\alpha \times \beta := ((\alpha_P, \beta_Q)^t \mid P \in TQ(Z_x), Q \in TQ(Z_y))$$

Zwischenpunkte von $Z_x \times Z_y$.

Satz 44.22 (Satz von Fubini für Quader)

Es seien $X \subset \mathbb{R}^p$ und $Y \subset \mathbb{R}^q$ zwei Quader, $f : X \times Y \rightarrow \mathbb{R}$ sei integrierbar und die Funktion

$$g : Y \rightarrow \mathbb{R} : y \mapsto \int_X f(x, y) dx$$

sei definiert, d.h. das Integral auf der rechten Seite existiere für alle $y \in Y$.

Dann ist g integrierbar auf Y und es gilt

$$\int_{X \times Y} f(x, y) d(x, y) = \int_Y \int_X f(x, y) dx dy.$$

Beweis: Wir müssen zeigen, daß g integrierbar ist mit

$$\int_Y g(y) \, dy = \int_{X \times Y} f(x, y) \, d(x, y) =: I.$$

Dazu wählen wir eine beliebige Folge von Zerlegungen Z_y^m von Y mit Zwischenpunkten β^m , so daß $\lim_{m \rightarrow \infty} l(Z_y^m) = 0$, und wegen des Riemannsches Folgenkriteriums für Integrierbarkeit 44.17 reicht es, zu zeigen, daß

$$\lim_{m \rightarrow \infty} ZS(g, Z_y^m, \beta^m) = I,$$

d.h. für alle $\varepsilon > 0$ gibt es ein $n_\varepsilon \in \mathbb{N}$, so daß

$$|ZS(g, Z_y^m, \beta^m) - I| < \varepsilon$$

für alle $m \geq n_\varepsilon$ gilt.

Um dies zu zeigen, wählen wir noch eine Folge von Zerlegungen Z_x^k von X mit Zwischenpunkten α^k und $\lim_{k \rightarrow \infty} l(Z_x^k) = 0$.

Sei nun $\varepsilon > 0$ gegeben. Aufgrund des Riemannsches Zwischensummenkriteriums 44.18 und weil f auf $X \times Y$ integrierbar ist, gibt es dann ein $\delta_\varepsilon > 0$, so daß für alle $(Z, \gamma) \in \mathcal{Z}(X \times Y)$ mit $l(Z) < \delta_\varepsilon$

$$|ZS(f, Z, \gamma) - I| < \frac{\varepsilon}{2} \tag{191}$$

gilt. Wegen $\lim_{m \rightarrow \infty} l(Z_y^m) = 0$ und $\lim_{k \rightarrow \infty} l(Z_x^k) = 0$ gibt es ein $n_\varepsilon > 0$, so daß

$$l(Z_x^k \times Z_y^m) = \max \{l(Z_x^k), l(Z_y^m)\} < \delta_\varepsilon$$

für alle $k, m \geq n_\varepsilon$, und wegen (191) folgt dann

$$|ZS(f, Z_x^k \times Z_y^m, \alpha^k \times \beta^m) - I| < \frac{\varepsilon}{2} \tag{192}$$

für alle $k, m \geq n_\varepsilon$. Für die Zwischensumme auf der linken Seite der Ungleichung gilt

$$\begin{aligned} ZS(f, Z_x^k \times Z_y^m, \alpha^k \times \beta^m) &\stackrel{(190)}{=} \sum_{P \times Q \in \text{TQ}(Z_x^k \times Z_y^m)} V(P \times Q) \cdot f(\alpha_P^k, \beta_Q^m) \\ &\stackrel{(189)}{=} \sum_{Q \in \text{TQ}(Z_y^m)} V(Q) \cdot \sum_{P \in \text{TQ}(Z_x^k)} V(P) \cdot f(\alpha_P^k, \beta_Q^m). \end{aligned}$$

Nach Voraussetzung ist die Funktion $x \mapsto f(x, \beta_Q^m)$ auf X integrierbar und aus dem Riemannsches Folgenkriterium für Integrierbarkeit 44.17 folgt mit $\lim_{k \rightarrow \infty} l(Z_x^k) = 0$ dann

$$\lim_{k \rightarrow \infty} \sum_{P \in \text{TQ}(Z_x^k)} V(P) \cdot f(\alpha_P^k, \beta_Q^m) = \int_X f(x, \beta_Q^m) \, dx = g(\beta_Q^m)$$

und somit

$$\lim_{k \rightarrow \infty} ZS(f, Z_x^k \times Z_y^m, \alpha^k \times \beta^m) = \sum_{Q \in \text{TQ}(Z_y^m)} V(Q) \cdot g(\beta_Q^m) = ZS(g, Z_y^m, \beta^m).$$

Mit (192) erhalten wir deshalb

$$|ZS(g, Z_y^m, \beta^m) - I| = \lim_{k \rightarrow \infty} |ZS(f, Z_x^k \times Z_y^m, \alpha^k \times \beta^m) - I| \leq \frac{\varepsilon}{2} < \varepsilon$$

für alle $m \geq n_\varepsilon$. Damit ist der Satz von Fubini für Quader bewiesen. \square

Beispiel 44.23 (Die Existenz des iterierten Integrals reicht nicht.)

Die Funktion

$$f: [0, 1] \times [0, 1] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 1, & \text{falls } x_2 \in \mathbb{Q}, \\ 2x_1, & \text{falls } x_2 \notin \mathbb{Q}, \end{cases}$$

ist auf den Quader $[0, 1] \times [0, 1]$ nicht integrierbar, aber das zugehörige iterierte Integral

$$\int_0^1 \int_0^1 f(x_1, x_2) \, dx_1 \, dx_2 = 1$$

existiert.

Beweis: Wir zeigen die Behauptung in drei Schritten.

1. Schritt: Sei $x_2 \in [0, 1]$ beliebig und

$$h: [0, 1] \longrightarrow \mathbb{R} : x_1 \mapsto f(x_1, x_2).$$

Ist $x_2 \in \mathbb{Q}$, so ist die Funktion h konstant 1 und mithin integrierbar mit

$$\int_0^1 h(x_1) \, dx_1 = \int_0^1 1 \, dx_1 = 1.$$

Ist $x_2 \notin \mathbb{Q}$, so ist $h(x_1) = 2x_1$ eine Polynomfunktion und mithin ist h ebenfalls integrierbar mit

$$\int_0^1 h(x_1) \, dx_1 = \int_0^1 2x_1 \, dx_1 = x_1^2 \Big|_0^1 = 1.$$

2. Schritt: Die Funktion

$$g: [0, 1] \longrightarrow \mathbb{R} : x_2 \mapsto \int_0^1 f(x_1, x_2) \, dx_1 = 1$$

ist konstant und mithin integrierbar, und für das iterierte Integral erhalten wir deshalb

$$\int_0^1 \int_0^1 f(x_1, x_2) \, dx_1 \, dx_2 = \int_0^1 g(x_2) \, dx_2 = \int_0^1 1 \, dx_2 = 1.$$

3. Schritt: Zeige, daß f nicht integrierbar auf $[0, 1] \times [0, 1]$ ist.

Für $m \geq 1$ betrachten wir die 2^m -te äquidistante Zerlegung Z^{2^m} des Quaders $[0, 1] \times [0, 1]$. Dann gilt

$$\text{TQ}(Z^{2^m}) = \left\{ Q_{ij} = \left[\frac{i-1}{2^m}, \frac{i}{2^m} \right] \times \left[\frac{j-1}{2^m}, \frac{j}{2^m} \right] \mid i, j = 1, \dots, 2^m \right\}$$

sowie

$$\sup_{x \in Q_{ij}} f(x) = \begin{cases} 2 \cdot \frac{i}{2^m} = \frac{i}{2^{m-1}}, & \text{falls } i > 2^{m-1}, \\ 1, & \text{falls } i \leq 2^{m-1}, \end{cases}$$

und

$$\inf_{x \in Q_{ij}} f(x) = \begin{cases} 1, & \text{falls } i > 2^{m-1}, \\ 2 \cdot \frac{i-1}{2^m} = \frac{i-1}{2^{m-1}}, & \text{falls } i \leq 2^{m-1}. \end{cases}$$

Außerdem ist das Volumen jedes der 2^{2m} Quader Teilquader Q_{ij} gerade

$$V(Q_{ij}) = \frac{1}{2^{2m}}.$$

Für die Obersumme von f bezüglich Z^{2^m} gilt deshalb

$$\begin{aligned} \text{OS}(f, Z^{2^m}) &= \sum_{j=1}^{2^m} \left(\sum_{i \leq 2^{m-1}} \frac{1}{2^{2m}} \cdot 1 + \sum_{i > 2^{m-1}} \frac{1}{2^{2m}} \cdot \frac{i}{2^{m-1}} \right) \\ &= 2^m \cdot \frac{1}{2^{2m}} \cdot \left(\sum_{i \leq 2^{m-1}} 1 + \frac{1}{2^{m-1}} \cdot \sum_{i=2^{m-1}+1}^{2^m} i \right) \\ &= \frac{1}{2^m} \cdot \left(2^{m-1} + \frac{2^m \cdot (2^m + 1) - 2^{m-1} \cdot (2^{m-1} + 1)}{2^m} \right) \\ &= \frac{1}{2} + \frac{3}{4} + \frac{1}{2^{m+1}} = \frac{5}{4} + \frac{1}{2^{m+1}} \geq \frac{5}{4} \end{aligned}$$

und für die Untersumme gilt analog

$$\begin{aligned} \text{US}(f, Z^{2^m}) &= \sum_{j=1}^{2^m} \left(\sum_{i \leq 2^{m-1}} \frac{1}{2^{2m}} \cdot \frac{i-1}{2^{m-1}} + \sum_{i > 2^{m-1}} \frac{1}{2^{2m}} \cdot 1 \right) \\ &= \frac{1}{2^m} \cdot \left(\frac{2^{m-1} \cdot (2^{m-1} - 1)}{2^m} + 2^{m-1} \right) \\ &= \frac{1}{4} + \frac{1}{2^{m+1}} + \frac{1}{2} = \frac{3}{4} - \frac{1}{2^{m+1}} \leq \frac{3}{4}. \end{aligned}$$

Wäre f integrierbar, so müßte es zu $\varepsilon = \frac{1}{4}$ ein $\delta_\varepsilon > 0$ geben, so daß für alle Zerlegungen von $[0, 1] \times [0, 1]$ mit $l(Z) < \delta_\varepsilon$ auch

$$\text{OS}(f, Z) - \text{US}(f, Z) < \frac{1}{4}.$$

Aber $l(Z^{2^m}) \rightarrow 0$ für $m \rightarrow \infty$, so daß für m hinreichend groß auch $l(Z^{2^m}) < \delta_\varepsilon$ gesichert ist, im Widerspruch zu

$$\text{OS}(f, Z^{2^m}) - \text{US}(f, Z^{2^m}) \geq \frac{1}{2}.$$

Also ist f nicht integrierbar auf $[0, 1] \times [0, 1]$. □

Man kann im Satz von Fubini die Rollen von X und Y auch vertauschen, sofern die notwendigen Voraussetzungen erfüllt sind, und dann erhält man, daß es bei der Aufteilung des Integrals nicht auf die Integrationsreihenfolge ankommt.

Korollar 44.24 (Satz über die Vertauschbarkeit der Integrationsreihenfolge)
Es seien $X \subset \mathbb{R}^p$ und $Y \subset \mathbb{R}^q$ zwei Quader, $f : X \times Y \rightarrow \mathbb{R}$ sei integrierbar und die Funktionen

$$g : Y \rightarrow \mathbb{R} : y \mapsto \int_X f(x, y) \, dx$$

und

$$h : X \rightarrow \mathbb{R} : x \mapsto \int_Y f(x, y) \, dy$$

seien definiert, d.h. die Integrale existieren für alle $y \in Y$ bzw. $x \in X$.

Dann sind g auf Y und h auf X integrierbar, und es gilt

$$\int_X \int_Y f(x, y) \, dy \, dx = \int_{X \times Y} f(x, y) \, d(x, y) = \int_Y \int_X f(x, y) \, dx \, dy.$$

Beweis: Die Aussage zu X und die zweite Gleichheit der Integrale ist die Aussage des Satzes von Fubini für Quader 44.22. Die Aussage für Y und die erste Gleichheit der Integrale folgt analog, indem man die Rollen von X und Y im Satz von Fubini 44.22 vertauscht. \square

Die Funktionen g und h in Korollar 44.24 sind sicher definiert, wenn die Funktion f stetig ist. Der Satz von Fubini erlaubt also, die Berechnung eines mehrdimensionalen Integrals auf das wiederholte, d.h. iterierte, Berechnen eindimensionaler Integrale zurückzuführen.

Korollar 44.25 (Iterierte Integration von mehrdimensionalen Integralen)

Es seien $a, b \in \mathbb{R}^n$ mit $a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann ist

$$\int_{[a, b]} f(x) \, dx = \int_{a_n}^{b_n} \dots \int_{a_1}^{b_1} f(x_1, \dots, x_n) \, dx_1 \dots dx_n.$$

Zudem darf die Reihenfolge der Integration auf der rechten Seite beliebig vertauscht werden.

Beweis: Dies folgt durch n -faches Anwenden des Satzes über die Vertauschbarkeit der Integration 44.24. \square

Beispiel 44.26

Wir wollen als einfaches Beispiel das Integral der Funktion

$$f : [0, 1] \times [0, 1] \rightarrow \mathbb{R} : x \mapsto x_1 + x_2$$

auf dem Quader $Q = [0, 1] \times [0, 1]$ berechnen:

$$\begin{aligned} \int_Q f(x) \, dx &= \int_0^1 \int_0^1 x_1 + x_2 \, dx_1 \, dx_2 = \int_0^1 \left. \frac{x_1^2}{2} + x_1 x_2 \right|_0^1 dx_2 \\ &= \int_0^1 \left. \frac{1}{2} + x_2 \right|_0^1 dx_2 = \left. \frac{x_2}{2} + \frac{x_2^2}{2} \right|_0^1 = 1. \end{aligned}$$

F) Der Mittelwertsatz für Riemann-Integrale auf Quadern**Satz 44.27** (Mittelwertsatz der Integralrechnung für Quader)Seien $\mathbf{a}, \mathbf{b} \in \mathbb{R}^n$ mit $\mathbf{a} < \mathbf{b}$, $f : [\mathbf{a}, \mathbf{b}] \rightarrow \mathbb{R}$ stetig. Dann gibt es ein $\mathbf{c} \in [\mathbf{a}, \mathbf{b}]$ mit

$$\int_{[\mathbf{a}, \mathbf{b}]} f(\mathbf{x}) \, d\mathbf{x} = f(\mathbf{c}) \cdot V([\mathbf{a}, \mathbf{b}]).$$

Beweis: Der Beweis geht genau wie im Eindimensionalen, siehe den alternativen Beweis von Korollar 20.9.Da f stetig auf dem Kompaktum $[\mathbf{a}, \mathbf{b}]$ ist, gibt es nach Korollar 40.24 $\mathbf{y}, \mathbf{z} \in [\mathbf{a}, \mathbf{b}]$ mit

$$f(\mathbf{y}) \leq f(\mathbf{x}) \leq f(\mathbf{z})$$

für alle $\mathbf{x} \in [\mathbf{a}, \mathbf{b}]$. Für die triviale Zerlegung Z von $[\mathbf{a}, \mathbf{b}]$ mit nur einem Teilquader gilt dann

$$f(\mathbf{y}) \cdot V([\mathbf{a}, \mathbf{b}]) = \text{US}(f, Z) \leq \int_{[\mathbf{a}, \mathbf{b}]} f(\mathbf{x}) \, d\mathbf{x} \leq \text{OS}(f, Z) = f(\mathbf{z}) \cdot V([\mathbf{a}, \mathbf{b}]). \quad (193)$$

Die Funktion

$$g : [0, 1] \rightarrow \mathbb{R} : t \mapsto f(\mathbf{y} + t \cdot (\mathbf{z} - \mathbf{y}))$$

ist stetig, da f stetig auf $\overline{\mathbf{y}\mathbf{z}} \subset [\mathbf{a}, \mathbf{b}]$ ist, und aus (193) folgt

$$g(0) = f(\mathbf{y}) \leq \frac{1}{V([\mathbf{a}, \mathbf{b}])} \cdot \int_{[\mathbf{a}, \mathbf{b}]} f(\mathbf{x}) \, d\mathbf{x} \leq f(\mathbf{z}) = g(1).$$

Wegen des Zwischenwertsatzes 14.12 nimmt g jeden Wert zwischen $g(0)$ und $g(1)$ an, so daß es ein $t \in [0, 1]$ gibt mit

$$f(\mathbf{c}) = g(t) = \frac{1}{V([\mathbf{a}, \mathbf{b}])} \cdot \int_{[\mathbf{a}, \mathbf{b}]} f(\mathbf{x}) \, d\mathbf{x}$$

für $\mathbf{c} = \mathbf{y} + t \cdot (\mathbf{z} - \mathbf{y}) \in [\mathbf{a}, \mathbf{b}]$. □**Aufgaben****Aufgabe 44.28**

Berechne das folgende Integral:

$$\int_{[0, \frac{\pi}{2}] \times [0, \frac{\pi}{2}]} \sin(\mathbf{x} + \mathbf{y}) \, d(\mathbf{x}, \mathbf{y})$$

Aufgabe 44.29

Berechne das folgende Integral:

$$\int_{[0, 1] \times [0, 1] \times [0, 1]} \frac{x^3 y^2}{1 + z^2} \, d(\mathbf{x}, \mathbf{y}, \mathbf{z}).$$

Aufgabe 44.30

Für eine rationale Zahl $x = \frac{a}{b}$ mit $a, b > 0$ setzen wir $N(x) = \frac{b}{\text{ggT}(a,b)}$, d.h. $N(x)$ ist der Nenner von x in gekürzter Form und wir setzen $N(0) = 1$. Untersuche die Funktionen

$$f : [0, 1] \times [0, 1] \longrightarrow \mathbb{R} : (x, y)^t \mapsto \begin{cases} \frac{1}{N(x)}, & \text{falls } x, y \in \mathbb{Q}, \\ 0, & \text{sonst,} \end{cases}$$

und

$$g : [0, 1] \times [0, 1] \longrightarrow \mathbb{R} : (x, y)^t \mapsto \begin{cases} 1, & \text{falls } x, y \in \mathbb{Q} \text{ mit } N(x) = N(y), \\ 0, & \text{sonst,} \end{cases}$$

bezüglich ihrer Integrierbarkeit auf $[0, 1] \times [0, 1]$ und bestimme ggf. den Wert des Integrals über diesem Quader.

Aufgabe 44.31

Zeige, sind $f, g : [a, b] \longrightarrow \mathbb{R}$ auf dem Intervall $[a, b] \subseteq \mathbb{R}$ integrierbar, so ist $f \cdot g$ auf $[a, b] \times [a, b]$ integrierbar und

$$\int_{[a,b] \times [a,b]} f(x) \cdot g(y) \, d(x, y) = \int_a^b f(x) \, dx \cdot \int_a^b g(y) \, dy.$$

Aufgabe 44.32

Es seien $a, b \in \mathbb{R}^n$ mit $a < b$ und $f : [a, b] \longrightarrow \mathbb{R}_{\geq 0}$ sei integrierbar auf $[a, b]$ und stetig in $c \in [a, b]$ mit $f(c) > 0$. Zeige, $\int_{[a,b]} f(x) \, dx > 0$.

Aufgabe 44.33

Es sei $[a, b]$ ein Intervall in \mathbb{R} und $f : [a, b] \longrightarrow \mathbb{R}_{>0}$ sei stetig. Zeige,

$$\int_a^b f(x) \, dx \cdot \int_a^b \frac{1}{f(x)} \, dx \geq (b - a)^2.$$

Hinweis: Setze die linke Seite der Ungleichung zu $\frac{1}{2} \cdot \int_{[a,b] \times [a,b]} \frac{f(x)}{f(y)} + \frac{f(y)}{f(x)} \, d(x, y)$ in Beziehung.

§ 45 Das Lebesguesche Integrabilitätskriterium

Wir verzichten in dieser Vorlesung darauf, das Lebesgue-Integral als Alternative zum Riemann-Integral einzuführen, da es dazu unerlässlich ist, den Begriff des Maßes einer Teilmenge einzuführen. Wir wollen aber das Kriterium von Lebesgue behandeln, das eine sehr prägnante Charakterisierung der Riemann-integrierbaren Funktionen gibt. Hierzu reicht es aus, einzuführen, wann eine Menge das Maß null hat, was recht einfach und anschaulich ist. Kann man eine Menge mit Quadern überdecken, deren Gesamtvolumen beliebig klein gewählt werden kann, so sollte man der Menge wohl kein positives Volumen, d.h. kein positives Maß zuweisen.

A) Nullmengen

Wir wollen nun eine mögliche Definition dafür geben, daß eine Menge Volumen null hat. Da wir bislang nur in der Lage sind, das Volumen von Quadern wirklich leicht auszurechnen, ist die Grundidee, die Menge mit Quadern zu überdecken, deren Gesamtvolumen beliebig klein gewählt werden kann.

Definition 45.1 (Nullmenge)

Eine Teilmenge $N \subseteq \mathbb{R}^n$ heißt eine *Nullmenge*, wenn es zu jedem $\varepsilon > 0$ eine Folge von Quadern $(Q_m)_{m \in \mathbb{N}}$ im \mathbb{R}^n gibt, die N überdecken, d.h. $N \subseteq \bigcup_{m \in \mathbb{N}} Q_m$, und für die die Reihe der Volumina absolut konvergent ist mit

$$\sum_{m=0}^{\infty} V(Q_m) < \varepsilon.$$

Von einer Teilmenge des \mathbb{R}^n , deren Dimension echt kleiner als n ist, was auch immer das im konkreten Fall heißen mag, würden wir erwarten, daß ihr n -dimensionales Volumen null ist. Für diesen Sachverhalt geben wir jetzt ein erstes Beispiel.

Beispiel 45.2 (Nullmengen)

Eine Hyperebene der Form

$$V(x_j - c) = \{x \in \mathbb{R}^n \mid x_j = c\}$$

für ein $j \in \{1, \dots, n\}$ und $c \in \mathbb{R}$ ist eine Nullmenge.

Um dies zu sehen, geben wir uns ein $\varepsilon > 0$ vor und definieren für $m \geq 1$ den Quader

$$Q_m = [a^m, b^m]$$

durch

$$a_i^m := -m \quad \text{und} \quad b_i^m := m$$

für $i \neq j$ und

$$a_j^m := c - \frac{\varepsilon}{2^{m+1} \cdot 2^{n-1} \cdot m^{n-1}}$$

sowie

$$b_j^m := c + \frac{\varepsilon}{2^{m+1} \cdot 2^{n-1} \cdot m^{n-1}}.$$

Dann überdecken die Q_m die Hyperebene $V(x_j - c)$, d.h. $V(x_j - c) \subseteq \bigcup_{m=1}^{\infty} Q_m$, da es für einen beliebigen Punkt $x \in V(x_j - c)$ ein $m \geq 1$ gibt mit

$$\max_{i \neq j} |x_i| < m$$

und damit $x \in Q_m$. Außerdem gilt für das Volumen von Q_m

$$V(Q_m) = (2m)^{n-1} \cdot \frac{2 \cdot \varepsilon}{2^{m+1} \cdot 2^{n-1} \cdot m^{n-1}} = \frac{\varepsilon}{2^m},$$

und damit ist die Reihe der Volumina konvergent mit

$$\sum_{m=1}^{\infty} V(Q_m) = \varepsilon \cdot \sum_{m=1}^{\infty} \frac{1}{2^m} = \varepsilon.$$

Bemerkung 45.3 (Umgebungen in der Maximumsnorm sind Würfel.)

Wenn man mit Quadern arbeiten möchte, ist die Maximumsnorm besonders gut geeignet. Ist etwa $x \in \mathbb{R}^n$ und $\delta > 0$, so ist die δ -Umgebung von x in der Maximumsnorm

$$U_{\delta}(x) = (x_1 - \delta, x_1 + \delta) \times \dots \times (x_n - \delta, x_n + \delta)$$

das Innere des Quaders

$$\overline{U_{\delta}(x)} = [x_1 - \delta, x_1 + \delta] \times \dots \times [x_n - \delta, x_n + \delta].$$

Quader dieser Form nennt man auch *Würfel*, da sie dadurch gekennzeichnet sind, daß die Längen aller beteiligten Intervalle gleich sind.

Proposition 45.4 (Nullmengen)

- Teilmengen von Nullmengen sind Nullmengen.
- Jede endliche oder abzählbare Vereinigung von Nullmengen ist eine Nullmenge.
- $N \subsetneq \mathbb{R}^n$ ist genau dann eine Nullmenge in \mathbb{R}^n , wenn es zu jedem $\varepsilon > 0$ eine Folge von Quadern $(Q_m)_{m \in \mathbb{N}}$ in \mathbb{R}^n gibt mit

$$N \subsetneq \bigcup_{m=0}^{\infty} \overset{\circ}{Q}_m$$

und so, daß die Reihe der Volumina absolut konvergent ist mit

$$\sum_{m=0}^{\infty} V(Q_m) < \varepsilon,$$

d.h. man kann in der Definition von Nullmenge für die Überdeckung die abgeschlossenen Quader auch durch offene ersetzen.

- Ist $N \subseteq \mathbb{R}^n$ eine kompakte Nullmenge, so besitzt N zu jedem $\varepsilon > 0$ eine endliche Überdeckung durch das Innere von Quadern Q_1, \dots, Q_k mit $\sum_{i=1}^k V(Q_i) < \varepsilon$.
- In der Definition des Begriffs Nullmenge kann man Quader auch durch Würfel ersetzen.

Beweis:

- a. Ist N eine Nullmenge, so gibt es zu $\varepsilon > 0$ eine abzählbare Überdeckung $N \subseteq \bigcup_{m=0}^{\infty} Q_m$ von N durch Quader mit

$$\sum_{m=0}^{\infty} V(Q_m) < \varepsilon,$$

und ist $M \subseteq N$, so überdecken die Q_m 's die Menge M ebenfalls. Also ist M auch eine Nullmenge.

- b. Seien N_k für $k \in \mathbb{N}$ Nullmengen in \mathbb{R}^n . Zu jedem $\varepsilon > 0$ gibt es dann Quader Q_{ki} , $i \in \mathbb{N}$, mit

$$N_k \subseteq \bigcup_{i=0}^{\infty} Q_{ki}$$

und

$$\sum_{i=0}^{\infty} V(Q_{ki}) < \frac{\varepsilon}{2^{k+2}}.$$

Mit Hilfe des Cantorschen Diagonalverfahrens können wir die Q_{ki} dann zu einer Folge $(Q'_m)_{m \in \mathbb{N}}$ anordnen und dabei gilt

$$\bigcup_{k=0}^{\infty} N_k \subseteq \bigcup_{k=0}^{\infty} \bigcup_{i=0}^{\infty} Q_{ki} = \bigcup_{m=0}^{\infty} Q'_m.$$

Zudem gilt

$$\sum_{m=0}^p V(Q'_m) \leq \sum_{k=0}^p \sum_{i=0}^p V(Q_{ki}) \leq \sum_{k=0}^{\infty} \sum_{i=0}^{\infty} V(Q_{ki}) = \sum_{k=0}^{\infty} \frac{\varepsilon}{2^{k+2}} = \frac{\varepsilon}{2}$$

und damit auch

$$\sum_{m=0}^{\infty} V(Q'_m) = \lim_{p \rightarrow \infty} \sum_{m=0}^p V(Q'_m) \leq \frac{\varepsilon}{2} < \varepsilon.$$

Also ist die Vereinigung abzählbar vieler Nullmengen eine Nullmenge. Die Aussage schließt den Fall der Vereinigung endlich vieler Nullmengen ein, da man jede endliche Familie von Nullmengen zu einer abzählbaren ergänzen kann, indem man entsprechend oft die leere Menge hinzunimmt.

- c. Wir wollen zunächst zeigen, daß es zu jedem Quader $P = [a, b] \subseteq \mathbb{R}^n$ auch einen Quader $Q \subseteq \mathbb{R}^n$ mit $P \subseteq \overset{\circ}{Q}$ und $V(Q) < 2 \cdot V(P)$ gibt. Für $t \geq 0$ definieren wir $Q_t = [a_1 - t, b_1 + t] \times \dots \times [a_n - t, b_n + t]$ und erhalten dann eine Funktion

$$v: \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R}: t \mapsto V(Q_t) = \prod_{i=1}^n (d_i - c_i) = \prod_{i=1}^n (b_i - a_i + 2t),$$

die als Polynomfunktion stetig ist. Mithin gibt es zu $\varepsilon = V(P) > 0$ ein $\delta_\varepsilon > 0$, so daß für alle $0 \leq t < \delta_\varepsilon$ auch

$$V(Q_t) - V(P) = |v(t) - v(0)| < \varepsilon = V(P)$$

gilt. Für $t = \frac{\delta_\varepsilon}{2}$ und $Q = Q_t$ erhalten wir dann

$$P = [a, b] \subseteq (a_1 - t, b_1 + t) \times \dots \times (a_n - t, b_n + t) = \mathring{Q}$$

und

$$V(Q) < 2 \cdot V(P).$$

Ist nun N eine Nullmenge und $\varepsilon > 0$ gegeben, so finden wir eine Folge von Quadern P_m , $m \in \mathbb{N}$, mit

$$N \subsetneq \bigcup_{m=0}^{\infty} P_m$$

und

$$\sum_{m=0}^{\infty} V(P_m) < \frac{\varepsilon}{2}.$$

Aufgrund der Vorüberlegung können wir zu jedem P_m einen Quader Q_m mit $P_m \subsetneq \mathring{Q}_m$ und $V(Q_m) < 2 \cdot V(P_m)$ wählen. Für diese Quader gilt dann

$$N \subsetneq \bigcup_{m=0}^{\infty} P_m \subsetneq \bigcup_{m=0}^{\infty} \mathring{Q}_m$$

und

$$\sum_{m=0}^{\infty} V(Q_m) \leq 2 \cdot \sum_{m=0}^{\infty} V(P_m) < \varepsilon.$$

- d. Es sei N eine kompakte Nullmenge in \mathbb{R}^n und $\varepsilon > 0$. Dann gibt es wegen c. eine offene Überdeckung von N ,

$$N \subsetneq \bigcup_{m=0}^{\infty} \mathring{Q}_m,$$

mit Quadern Q_m , $m \in \mathbb{N}$, und

$$\sum_{m=0}^{\infty} V(Q_m) < \varepsilon.$$

Da N kompakt ist reichen aber endlich viele dieser Quader aus, und die Summe der Volumina von diesen ist dann erst recht kleiner als ε .

- e. Wir zeigen zunächst, daß man jeden Quader $Q = [a, b] \in \mathbb{R}^n$ durch endlich viele Würfel W_1, \dots, W_k überdecken kann, so daß

$$V(W_1) + \dots + V(W_k) \leq 2^n \cdot V(Q). \quad (194)$$

Dazu können wir ohne Einschränkung annehmen, daß $a = 0$ der Ursprung ist. Für jedes $i = 1, \dots, n$ gibt es eine rationale Zahl $\frac{p_i}{q_i} \in [b_i, 2b_i]$. Indem wir diese Zahlen auf ihren Hauptnenner bringen, können wir $q_i = q$ für alle $i = 1, \dots, n$ annehmen. Den Quader

$$P = \left[0, \frac{p_1}{q}\right] \times \dots \times \left[0, \frac{p_n}{q}\right]$$

kann man offenbar in $k = p_1 \cdot \dots \cdot p_n$ Würfeln W_1, \dots, W_k der Seitenlänge $\frac{1}{q}$ zerlegen, und damit gilt

$$Q \subseteq P = \bigcup_{i=1}^k W_i \subseteq [0, 2b_1] \times \dots \times [0, 2b_n] =: Q'$$

und

$$V(P) \leq V(Q') = \prod_{i=1}^n 2b_i = 2^n \cdot V(Q).$$

Damit ist die Hilfsaussage gezeigt.

Ist nun N eine Nullmenge, so besitzt N für jedes $\varepsilon > 0$ eine abzählbare Überdeckung durch Quader Q_m mit

$$\sum_{m=0}^{\infty} V(Q_m) < \frac{\varepsilon}{2^n}.$$

Jeden Quader Q_m überdecken wir mit endlich vielen Würfeln $W_{m,1}, \dots, W_{m,k_m}$, die (194) für Q_m genügen. Dann überdecken die so gewonnenen abzählbar vielen Würfel N und für die Reihe ihrer Volumina gilt

$$\sum_{m=0}^{\infty} \sum_{i=1}^{k_m} V(W_{m,i}) \leq \sum_{m=0}^{\infty} 2^n \cdot V(Q_m) < \varepsilon.$$

Jede Nullmenge läßt sich also schon durch Würfel mit beliebig kleinem Gesamtvolumen überdecken. Umgekehrt ist jede solche Überdeckung natürlich auch eine Überdeckung durch Quader. Dies zeigt, daß man in der Definition von Nullmenge auch Würfel statt Quader hätte nehmen können.

□

Beispiel 45.5 (Ränder von Quadern sind kompakte Nullmengen.)

Die Seiten eines Quaders in \mathbb{R}^n sowie sein Rand sind kompakte Nullmengen.

Um dies zu sehen, beachte man folgendes: ist $Q = [\mathbf{a}, \mathbf{b}] \subsetneq \mathbb{R}^n$ ein Quader, so sind die Seiten von Q in den Hyperebenen $V(\mathbf{x}_j - \mathbf{a}_j)$ bzw. $V(\mathbf{x}_j - \mathbf{b}_j)$ für $j \in \{1, \dots, n\}$ enthalten, so daß sie wegen Proposition 45.4 a. und Beispiel 45.2 Nullmengen sind. Der Rand des Quaders ist dann als Vereinigung der endlich vielen Seiten nach Proposition 45.4 b. ebenfalls eine Nullmenge. Als abgeschlossene Teilmengen der kompakten Menge Q sind die Seiten und auch der Rand des Quaders zudem kompakt (siehe Korollar 39.32).

Beispiel 45.6 (\mathbb{Q}^n ist eine Nullmenge.)

\mathbb{Q}^n ist eine Nullmenge in \mathbb{R}^n und ebenso $[\mathbf{a}, \mathbf{b}] \cap \mathbb{Q}^n$ in \mathbb{R}^n .

Um dies zu sehen, beachten wir, daß eine einpunktige Menge als Teilmenge der Seite eines Quaders sicher eine Nullmenge ist und daß mit Q auch \mathbb{Q}^n abzählbar unendlich ist. Also ist \mathbb{Q}^n als abzählbare Vereinigung von Nullmengen eine Nullmenge nach Proposition 45.4, und dann ist auch $[\mathbf{a}, \mathbf{b}] \cap \mathbb{Q}^n$ als Teilmenge einer Nullmenge eine Nullmenge.

Lemma 45.7 (Lipschitz stetige Bilder von Nullmengen sind Nullmengen.)

Ist $f : D \rightarrow \mathbb{R}^n$ mit $D \subseteq \mathbb{R}^n$ Lipschitz stetig auf D und N eine Nullmenge in \mathbb{R}^n , so ist $f(N \cap D)$ ebenfalls eine Nullmenge in \mathbb{R}^n .

Beweis: Egal welche Norm wir auf \mathbb{R}^n für die Lipschitz-Stetigkeit vorgegeben haben, f ist *auch* Lipschitz stetig bezüglich der Maximumsnorm. Sei nun $q > 0$ eine Lipschitz-Konstante für f und sei zudem $\varepsilon > 0$ gegeben.

Nach Proposition 45.4 gibt es eine Überdeckung von N durch Würfel $W_i = \overline{U_{\delta_i}(x_i)}$, $i \in \mathbb{N}$, mit

$$\sum_{i=0}^{\infty} 2^n \cdot \delta_i^n = \sum_{i=0}^{\infty} V(W_i) < \frac{\varepsilon}{q^n}. \quad (195)$$

Da f Lipschitz-stetig mit Lipschitz-Konstante q ist, gilt

$$\|f(x) - f(y)\|_{\infty} \leq q \cdot \|x - y\|_{\infty}$$

für alle $x, y \in D$ und mithin auch

$$f(W_i \cap D) = f(\overline{U_{\delta_i}(x_i)} \cap D) \subseteq \overline{U_{q \cdot \delta_i}(f(x_i))}.$$

Es folgt, daß

$$f(N \cap D) \subseteq \bigcup_{i=0}^{\infty} f(W_i \cap D) \subseteq \bigcup_{i=0}^{\infty} \overline{U_{q \cdot \delta_i}(f(x_i))}$$

eine abzählbare Überdeckung mit Würfeln besitzt, wobei

$$\sum_{i=0}^{\infty} V\left(\overline{U_{q \cdot \delta_i}(f(x_i))}\right) = \sum_{i=0}^{\infty} (2 \cdot q \cdot \delta_i)^n \stackrel{(195)}{<} \varepsilon$$

gilt. Mithin ist $f(N)$ eine Nullmenge. □

Proposition 45.8 (Stetig differenzierbare Bilder von kompakten Nullmengen)

Ist $U \subseteq \mathbb{R}^n$ offen, $f : U \rightarrow \mathbb{R}^n$ stetig differenzierbar und $N \subsetneq U$ eine kompakte Nullmenge in \mathbb{R}^n , so ist $f(N)$ eine kompakte Nullmenge in \mathbb{R}^n .

Beweis: Im Beweis wollen wir Lemma 45.7 anwenden und überdecken dazu N mit endlich vielen Mengen, auf denen f Lipschitz stetig ist.

Da U offen ist, besitzt jeder Punkt $x \in N$ eine offene Umgebung, deren Abschluß

$$\overline{U_{\delta_x}(x)} \subseteq U$$

noch ganz in U liegt. Da N kompakt ist, besitzt die offene Überdeckung

$$N \subseteq \bigcup_{x \in N} U_{\delta_x}(x)$$

von N zudem eine endliche Teilüberdeckung, d.h. es gibt $x_1, \dots, x_k \in N$ mit

$$N \subseteq \bigcup_{i=1}^k U_{\delta_{x_i}}(x_i).$$

Wegen Proposition 40.33 ist f Lipschitz stetig auf $\overline{U_{\delta_{x_i}}(x_i)}$, und somit ist $f(N \cap U_{\delta_{x_i}}(x_i))$ nach Lemma 45.7 dann eine Nullmenge in \mathbb{R}^n . Als endliche Vereinigung von Nullmengen ist dann aber auch

$$f(N) = \bigcup_{i=1}^k f(N \cap U_{\delta_{x_i}}(x_i))$$

eine Nullmenge in \mathbb{R}^n . Außerdem ist $f(N)$ als Bild einer kompakten Menge unter einer stetigen Abbildung nach Satz 40.23 kompakt. \square

Beispiel 45.9 (Die Kreislinie ist eine Nullmenge.)

Wir wollen zeigen, daß die Kreislinie $S^1 = V(x_1^2 + x_2^2 - 1)$ eine Nullmenge ist.

Dazu betrachten wir die stetig differenzierbaren Polarkoordinaten

$$f : \mathbb{R}_{>0} \times \mathbb{R} \longrightarrow \mathbb{R}^2 : (r, \theta) \mapsto (r \cdot \cos(\theta), r \cdot \sin(\theta))$$

und die kompakte Nullmenge $\{1\} \times [0, 2\pi]$. Deren Bild unter f ist dann mit Proposition 45.8 die kompakte Nullmenge

$$f(\{1\} \times [0, 2\pi]) = S^1.$$

B) Das Lebesguesche Integrabilitätskriterium

Definition 45.10 (Fast überall stetig)

Eine Funktion $f : D \longrightarrow \mathbb{R}$ mit $D \subseteq \mathbb{R}^n$ heißt *fast überall stetig*, falls es eine Nullmenge $N \subseteq \mathbb{R}^n$ gibt, so daß f stetig in x ist für alle $x \in D \setminus N$.

Beispiel 45.11

Die Dirichletsche Sprungfunktion

$$g : [0, 1] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x \in \mathbb{Q}, \\ 1, & \text{falls } x \notin \mathbb{Q} \end{cases}$$

ist zwar außerhalb der Nullmenge $[0, 1] \cap \mathbb{Q}$ konstant, aber in keinem Punkt stetig!

Lemma 45.12

Ist ein Quader $[a, b] \subseteq \mathbb{R}^n$ überdeckt

$$[a, b] \subseteq \bigcup_{m=1}^k \overset{\circ}{Q}_m$$

von Inneren endlich vieler Quader Q_m , so können wir eine Zerlegung Z von $[a, b]$ wählen, so daß jeder Teilquader $Q \in TQ(Z)$ ganz in einem der Q_m enthalten ist.

Beweis: Um dies zu sehen, setzen wir $Q_m = [a^m, b^m]$, $m = 1, \dots, k$, und betrachten für $i \in \{1, \dots, n\}$ die *endliche* Menge

$$\{a_i, b_i, a_i^m, b_i^m \mid m = 1, \dots, k, a_i < a_i^m, b_i^m < b_i\}.$$

Nehmen wir die Punkte dieser Menge als Stützpunkt einer Zerlegung Z_i von $[a_i, b_i]$, dann ist

$$Z = Z_1 \times \dots \times Z_n$$

eine Zerlegung von $[a, b]$. Sei nun $Q = [c, d] \in \text{TQ}(Z)$ ein Teilquader von Z , so gibt es nach Voraussetzung ein m mit $Q \cap \overset{\circ}{Q}_m \neq \emptyset$. Wir wählen ein $x \in Q \cap \overset{\circ}{Q}_m$ und erhalten für $i \in \{1, \dots, n\}$

$$c_i \leq x_i \leq d_i$$

und zugleich

$$a_i^m < x_i < b_i^m.$$

Aufgrund der Definition von Z_i können weder a_i^m noch b_i^m echt zwischen c_i und d_i liegen, so daß wir mithin notwendigerweise auch

$$a_i^m \leq c_i \leq x_i \leq d_i \leq b_i^m$$

erhalten. Damit folgt aber

$$Q = [c, d] \subseteq [a^m, b^m] = Q_m.$$

□

Das folgende Lebesguesche Integrabilitätskriterium ist eine Verallgemeinerung der Aussage von Aufgabe 19.28.

Satz 45.13 (Lebesguesches Integrabilitätskriterium)

Seien $a, b \in \mathbb{R}^n$ mit $a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ beschränkt.

Genau dann ist f integrierbar auf $[a, b]$, wenn f fast überall stetig ist.

Beweis: In diesem Beweis verwenden wir auf \mathbb{R}^n die Maximumsnorm $\|\cdot\|_\infty$, da dann

$$U_\delta(x) = (x_1 - \delta, x_1 + \delta) \times \dots \times (x_n - \delta, x_n + \delta)$$

das Innere eines Quaders ist. Für die Frage der Stetigkeit von f ist die gewählte Norm aber unerheblich.

\implies : Wir nehmen zunächst an, daß f auf $[a, b]$ integrierbar ist.

Unser Ziel ist es, zu zeigen, daß die Menge

$$N = \{x \in [a, b] \mid f \text{ ist nicht stetig in } x\}$$

eine Nullmenge in \mathbb{R}^n ist.

Nach Definition ist die Funktion f genau dann stetig in $x \in [a, b]$, wenn

$$\forall \varepsilon > 0 \exists \delta > 0 : \forall y \in U_\delta(x) \cap [a, b] : |f(y) - f(x)| < \varepsilon.$$

Mithin ist f genau dann nicht stetig in x , wenn

$$\exists \varepsilon > 0 : \forall \delta > 0 \exists y \in U_\delta(x) \cap [a, b] : |f(y) - f(x)| \geq \varepsilon.$$

Wir definieren für $m \geq 1$ deshalb die Menge

$$N_m := \left\{ x \in [a, b] \mid \forall \delta > 0 \exists y \in U_\delta(x) \cap [a, b] : |f(y) - f(x)| \geq \frac{1}{m} \right\}.$$

Dann gilt

$$N_1 \subseteq N_2 \subseteq N_3 \subseteq \dots$$

und

$$N = \bigcup_{m=1}^{\infty} N_m.$$

Wegen Proposition 45.4 b. reicht es, zu zeigen, daß N_m für jedes $m \geq 1$ eine Nullmenge ist, und wir können ohne Einschränkung $N_m \neq \emptyset$ annehmen.

Dazu geben wir uns zunächst ein $\varepsilon > 0$ vor. Da f auf $[a, b]$ integrierbar ist, gibt es wegen des Riemannsches Integrabilitätskriteriums 44.12 eine Zerlegung Z von $[a, b]$ mit

$$OS(f, Z) - US(f, Z) < \frac{\varepsilon}{2m}. \tag{196}$$

Um zu zeigen, daß N_m durch endlich viele Quader mit Volumensumme höchstens ε überdeckt werden kann, zerlegen wir N_m in die Menge N_m° der Punkte, die im Inneren von Teilquadern von Z liegen, und die Menge N_m^∂ der Punkte, die auf dem Rand solcher Teilquader liegen, und überdecken jede der beiden Teilmengen von N_m mit endlich vielen Quadern der Volumensumme höchstens $\frac{\varepsilon}{2}$.

Wir betrachten deshalb zunächst die Menge

$$T = \{ Q \in \text{TQ}(Z) \mid N_m \cap \overset{\circ}{Q} \neq \emptyset \}$$

aller Teilquader von Z , die einen Punkt aus N_m im Inneren enthalten. Ist $Q \in T$ und $x \in N_m \cap \overset{\circ}{Q}$, so gibt es ein $\delta > 0$ mit $U_\delta(x) \subsetneq \overset{\circ}{Q}$, da $\overset{\circ}{Q}$ offen ist, und da $x \in N_m$ ist, gibt es dann ein $y \in U_\delta(x) \subsetneq Q$ mit

$$\frac{1}{m} \leq |f(y) - f(x)| \leq \sup_{z \in Q} f(z) - \inf_{z \in Q} f(z). \tag{197}$$

Für die Summe der Volumina aller Quader in T ergibt sich dann aber

$$\begin{aligned} \sum_{Q \in T} V(Q) &= m \cdot \sum_{Q \in T} V(Q) \cdot \frac{1}{m} \\ &\stackrel{(197)}{\leq} m \cdot \sum_{Q \in T} V(Q) \cdot \left(\sup_{z \in Q} f(z) - \inf_{z \in Q} f(z) \right) \\ &\leq m \cdot \sum_{Q \in \text{TQ}(Z)} V(Q) \cdot \left(\sup_{z \in Q} f(z) - \inf_{z \in Q} f(z) \right) \\ &\leq m \cdot (OS(f, Z) - US(f, Z)) \stackrel{(196)}{<} \frac{\varepsilon}{2}. \end{aligned}$$

Damit ist die Menge

$$N_m^\circ \subseteq \bigcup_{Q \in T} Q$$

aller Punkte von N_m , die im Inneren von Teilquadern von Z liegen, durch die endlich vielen Quader in T überdeckt und deren Volumensumme ist höchstens $\frac{\varepsilon}{2}$.

Die Vereinigung

$$M = \bigcup_{Q \in TQ(Z)} \partial Q$$

der Ränder aller Teilquader in Z ist nach Beispiel 45.5 eine endliche Vereinigung von Nullmengen und ist nach Proposition 45.4 b. als solche wieder eine Nullmenge. Zudem ist sie als endliche Vereinigung abgeschlossener Mengen abgeschlossen in der kompakten Menge Q , also nach Korollar 39.32 auch kompakt. Dann gibt es nach Proposition 45.4 d. aber eine endliche Überdeckung durch Quader P_1, \dots, P_k mit

$$\sum_{i=1}^k V(P_i) < \frac{\varepsilon}{2}.$$

Diese Quader überdecken dann aber erst recht die Menge N_m^∂ der Punkte in N_m , die im Rand eines Teilquaders von Z liegen.

Insgesamt haben wir also gezeigt, daß

$$N_m = N_m^\circ \cup N_m^\partial \subsetneq \bigcup_{Q \in T} Q \cup \bigcup_{i=1}^k P_i$$

eine endliche Überdeckung durch Quader besitzt mit

$$\sum_{Q \in T} V(Q) + \sum_{i=1}^k V(P_i) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Also ist N_m eine Nullmenge, und somit ist auch N eine Nullmenge.

⇐: Wir wollen das Riemannsches Integrierbarkeitskriterium anwenden und geben uns deshalb ein $\varepsilon > 0$ vor. Wir müssen nun eine Zerlegung Z von $[a, b]$ finden, so daß

$$OS(f, Z) - US(f, Z) < \varepsilon.$$

Die Zerlegung Z wollen wir mit Lemma 45.12 finden, und dazu konstruieren wir uns eine geeignete endliche Überdeckung von $[a, b]$ durch offene Quader.

Da die Funktion f auf $[a, b]$ beschränkt ist, gibt es eine Konstante $c > 0$ mit

$$|f(x)| \leq c \tag{198}$$

für alle $x \in [a, b]$. Die Menge N aller Unstetigkeitsstellen von f ist nach Voraussetzung eine Nullmenge in \mathbb{R}^n , so daß es wegen Proposition 45.4 c. eine abzählbare Überdeckung

$$N \subsetneq \bigcup_{m=0}^{\infty} \overset{\circ}{Q}_m$$

von N durch das Innere von Quadern Q_m gibt mit

$$\sum_{m=0}^{\infty} V(Q_m) < \frac{\varepsilon}{2 \cdot c + V([a, b])}. \tag{199}$$

Ist $x \in [a, b] \setminus N$, so ist f stetig in x und deshalb gibt es ein $\delta > 0$ mit

$$|f(y) - f(x)| < \frac{\varepsilon}{8 \cdot c + 4 \cdot V([a, b])}$$

für alle $y \in U_\delta(x) \cap [a, b]$. Man beachte nun, daß der Abschluß

$$P_x := \overline{U_{\frac{\delta}{2}}(x)}$$

ein Quader in \mathbb{R}^n ist, und für je zwei Punkte $y, z \in P_x \cap [a, b]$ gilt

$$|f(y) - f(z)| \leq |f(y) - f(x)| + |f(x) - f(z)| < \frac{\varepsilon}{4 \cdot c + 2 \cdot V([a, b])}.$$

Durch Übergang zum Supremum und Infimum erhalten wir dann

$$\sup_{y \in P_x \cap [a, b]} f(y) - \inf_{z \in P_x \cap [a, b]} f(z) \leq \frac{\varepsilon}{4 \cdot c + 2 \cdot V([a, b])} < \frac{\varepsilon}{2 \cdot c + V([a, b])}. \quad (200)$$

Wir haben nun eine offene Überdeckung

$$[a, b] = N \cup ([a, b] \setminus N) \subseteq \bigcup_{m=0}^{\infty} \overset{\circ}{Q}_m \cup \bigcup_{x \in [a, b] \setminus N} \overset{\circ}{P}_x$$

der kompakten Menge $[a, b]$ konstruiert, und diese muß eine endliche Teilüberdeckung besitzen, d.h. es gibt $m_1, \dots, m_k \in \mathbb{N}$ und $x_1, \dots, x_l \in [a, b] \setminus N$ mit

$$[a, b] \subseteq \bigcup_{i=1}^k \overset{\circ}{Q}_{m_i} \cup \bigcup_{j=1}^l \overset{\circ}{P}_{x_j}.$$

Nach Lemma 45.12 gibt es eine Zerlegung Z von $[a, b]$, so daß jeder Teilquader von Z ganz in einem der Q_{m_i} oder der P_{x_j} liegt. Dann gilt

$$\begin{aligned} OS(f, Z) - US(f, Z) &= \sum_{Q \in TQ(Z)} V(Q) \cdot \left(\sup_{x \in Q} f(x) - \inf_{x \in Q} f(x) \right) \\ &\leq \sum_{\substack{Q \in TQ(Z) \\ \exists i: Q \subseteq Q_{m_i}}} V(Q) \cdot \left(\sup_{x \in Q} f(x) - \inf_{x \in Q} f(x) \right) \\ &\quad + \sum_{\substack{Q \in TQ(Z) \\ \exists j: Q \subseteq P_{x_j}}} V(Q) \cdot \left(\sup_{x \in Q} f(x) - \inf_{x \in Q} f(x) \right) \\ &\stackrel{(198), (200)}{\leq} \sum_{\substack{Q \in TQ(Z) \\ \exists i: Q \subseteq Q_{m_i}}} V(Q) \cdot 2 \cdot c + \sum_{\substack{Q \in TQ(Z) \\ \exists j: Q \subseteq P_{x_j}}} V(Q) \cdot \frac{\varepsilon}{2 \cdot c + V([a, b])} \\ &\leq \sum_{m=0}^{\infty} V(Q_m) \cdot 2 \cdot c + \sum_{Q \in TQ(Z)} V(Q) \cdot \frac{\varepsilon}{2 \cdot c + V([a, b])} \\ &\stackrel{(199)}{\leq} \frac{\varepsilon}{2 \cdot c + V([a, b])} \cdot 2 \cdot c + V([a, b]) \cdot \frac{\varepsilon}{2 \cdot c + V([a, b])} = \varepsilon. \end{aligned}$$

Aus dem Riemannschem Integrabilitätskriterium 44.12 folgt nun, daß f auf $[a, b]$ integrierbar ist. □

Beispiel 45.14

Es sei $Q \subseteq \mathbb{R}^n$ ein Quader und $f : \overset{\circ}{Q} \rightarrow \mathbb{R}$ sei stetig. Dann ist die Funktion

$$g : Q \rightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } x \in \overset{\circ}{Q}, \\ 0, & \text{falls } x \in \partial Q, \end{cases}$$

integrierbar auf Q , da g höchstens in der Nullmenge ∂Q unstetig ist.

Bemerkung 45.15

Mit Hilfe des Lebesgueschen Integrabilitätskriteriums 45.13 können wir den Beweis, daß mit $f : [a, b] \rightarrow \mathbb{R}$ auch $|f|$ auf $[a, b]$ integrierbar ist, erheblich verkürzen. Mit f ist nämlich auch $|f|$ beschränkt und $|f|$ ist in jedem Punkt stetig, in dem auch f stetig ist, so daß $|f|$ fast überall stetig und damit integrierbar auf $[a, b]$ ist.

Korollar 45.16 (Produkte integrierbarer Funktionen)

Es seien $a, b \in \mathbb{R}^n$ mit $a < b$ und $f, g : [a, b] \rightarrow \mathbb{R}$ seien integrierbar auf $[a, b]$.

- a. Dann ist $f \cdot g$ integrierbar auf $[a, b]$.
- b. $\exists q > 0$ mit $|g(x)| \geq q$ für alle $x \in [a, b]$, so ist auch $\frac{f}{g}$ integrierbar auf $[a, b]$.

Beweis: Die Funktionen $f \cdot g$ und $\frac{f}{g}$ sind unter den gegebenen Voraussetzungen jeweils beschränkt, und die Menge der Unstetigkeitsstellen von $f \cdot g$ und $\frac{f}{g}$ ist in der Vereinigung der Menge der Unstetigkeitsstellen von f und von g enthalten, und ist somit eine Nullmenge. Damit sind $f \cdot g$ und $\frac{f}{g}$ dann aber auch fast überall stetig und mit dem Lebesgueschen Integrabilitätskriterium auch integrierbar auf $[a, b]$. \square

Bemerkung 45.17

Das Lebesguesche Integrabilitätskriterium erlaubt uns zwar zu zeigen, daß das Produkt von zwei integrierbaren Funktionen integrierbar ist, es gibt aber keinen Hinweis auf den Wert des Integrals, und diesen zu berechnen ist im allgemeinen ein Problem.

Aufgaben**Aufgabe 45.18** (Integrale fast überall gleicher Funktionen)

Es seien $a, b \in \mathbb{R}^n$ mit $a < b$ und $f, g : [a, b] \rightarrow \mathbb{R}$ seien fast überall gleich, d.h. es gibt eine Nullmenge N , so daß $f(x) = g(x)$ für alle $x \in [a, b] \setminus N$.

- a. Ist N kompakt und f auf $[a, b]$ integrierbar, so ist auch g auf $[a, b]$ integrierbar.
- b. Sind f und g auf $[a, b]$ integrierbar, so gilt

$$\int_{[a,b]} f(x) \, dx = \int_{[a,b]} g(x) \, dx.$$

Aufgabe 45.19

Zeige, daß jede Hyperebene eine Nullmenge ist.

§ 46 Das Riemann-Integral über Jordan-meßbaren Mengen

Bei der eindimensionalen Integration war die Beschränkung auf Intervalle als Integrationsbereich sehr naheliegend und ausreichend. In der mehrdimensionalen Analysis kann man sich viele schöne Definitionsbereiche, etwa Kreisscheiben, vorstellen, die keine Quader sind und auch nicht einfach als endliche Vereinigung solcher dargestellt werden können. Wir wollen unsere Integrationstheorie deshalb erweitern, um auch allgemeinere Teilmengen von \mathbb{R}^n als Integrationsbereiche zulassen zu können.

A) Jordan-meßbare Mengen

Definition 46.1 (Charakteristische Funktionen)

Es sei $B \subseteq \mathbb{R}^n$ und $f : B \rightarrow \mathbb{R}$. Wir nennen die Funktion

$$1_B : \mathbb{R}^n \rightarrow \mathbb{R} : x \mapsto \begin{cases} 1, & \text{falls } x \in B, \\ 0, & \text{falls } x \notin B, \end{cases}$$

die *charakteristische Funktion von B*, und wir nennen die Funktion

$$f_B : \mathbb{R}^n \rightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } x \in B, \\ 0, & \text{falls } x \notin B, \end{cases}$$

die *charakteristische Funktion von f auf B*.

Lemma 46.2 (Integration charakteristischer Funktionen)

Es sei $B \subseteq \mathbb{R}^n$, $f : B \rightarrow \mathbb{R}$ und $P, Q \subset \mathbb{R}^n$ seien Quader die B enthalten.

Genau dann ist f_B auf P integrierbar, wenn f_B auf Q integrierbar ist, und in diesem Fall gilt zudem

$$\int_P f_B(x) \, dx = \int_Q f_B(x) \, dx.$$

Beweis: Wir betrachten zunächst den Fall, daß P in Q enthalten ist und wenden das Lebesguesche Integrabilitätskriterium 45.13 an.

Ist nun f_B auf P integrierbar, so ist f_B auf P und somit auch auf Q beschränkt und es gibt eine Nullmenge N in \mathbb{R}^n , so daß die Einschränkung von f_B auf P stetig in jedem $x \in P \setminus N$ ist. Dann ist die Einschränkung von f_B auf Q aber schlimmstenfalls in der Nullmenge $N \cup \partial P$ unstetig, so daß f_B auf Q integrierbar ist.

Ist umgekehrt f_B auf Q integrierbar, so ist die Einschränkung von f_B auf Q beschränkt und fast überall stetig. Die Eigenschaften übertragen sich offenbar auf die Einschränkung von f_B auf den kleineren Quader P, so daß f_B auf P integrierbar ist.

Sei nun also f_B auf $P = [a, b]$ und auf $Q = [c, d]$ integrierbar, so betrachten wir zunächst die Zerlegung Z, die durch P auf Q induziert wird, indem wir auf dem Intervall $[c_i, d_i]$ die Punkte in $\{c_i, a_i, b_i, d_i\}$ als Stützpunkte wählen. Man beachte, daß dann P ein Teilquader von Z ist. Dann wählen wir eine Folge von Verfeinerungen Z^m , $m \in \mathbb{N}$, von Z mit $l(Z^m) \rightarrow 0$, und für jeden Teilquader $Q \in \text{TQ}(Z^m)$ wählen wir einen Punkt $\alpha_Q^m \in \overset{\circ}{Q}$ im Inneren von Q. Auf dem Weg erhalten wir auch eine

Folge von Zwischenpunkten α^m zu den Z^m . Außerdem enthält jede Zerlegung Z^m eine Zerlegung X^m von P und α^m enthält eine Familie von Zwischenpunkten β^m von X^m , da P ein Teilquader von Z war und die Z^m die Zerlegung Z verfeinern. Schließlich gilt aufgrund der Konstruktion

$$ZS(f_B, Z^m, \alpha^m) = \sum_{Q \in \text{TQ}(Z^m)} V(Q) \cdot f(\alpha_Q^m) = \sum_{Q \in \text{TQ}(X^m)} V(Q) \cdot f(\beta_Q^m) = ZS(f_B, X^m, \beta^m),$$

da $f(\alpha_Q^m) = 0$ für alle $Q \in \text{TQ}(Z^m) \setminus \text{TQ}(X^m)$. Da die linke Seite nach dem Folgenkriterium für Integrierbarkeit 44.17 gegen $\int_Q f_B(x) \, dx$ konvergiert und die rechte Seite gegen $\int_P f_B(x) \, dx$, folgt die Gleichheit der beiden Integrale.

Damit ist die Aussage für den Fall bewiesen, daß P in Q enthalten ist. Sind nun P und Q beliebig, so wählen wir einen Quader X , der beide enthält. Dann gilt mit dem bereits Bewiesenen:

$$f_B \text{ integrierbar auf } P \iff f_B \text{ integrierbar auf } X \iff f_B \text{ integrierbar auf } Q,$$

und in diesem Fall gilt zudem

$$\int_P f_B(x) \, dx = \int_X f_B(x) \, dx = \int_Q f_B(x) \, dx.$$

□

Die Aussage dieses Lemmas erlaubt uns die folgende Definition des Integralbegriffs über beschränkten Mengen, die keine Quader sind.

Definition 46.3 (Riemann-integrierbar und Jordan-meßbar)

Es sei $B \subseteq \mathbb{R}^n$ eine beschränkte Menge.

- a. Eine Funktion $f : B \rightarrow \mathbb{R}$ heißt (*Riemann-*)*integrierbar* auf B , falls es einen Quader $Q \subseteq \mathbb{R}^n$ gibt, der B enthält und auf dem die charakteristische Funktion f_B integrierbar ist. In diesem Fall nennen wir

$$\int_B f(x) \, dx := \int_Q f_B(x) \, dx$$

das *Integral* von f auf B .

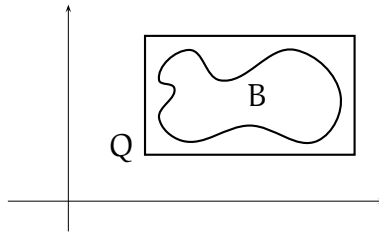


ABBILDUNG 16. Ein Quader Q , der B enthält.

b. Wir nennen B *Jordan-meßbar*, falls das Integral

$$V(B) := \int_B 1 \, dx$$

existiert, und wir nennen $V(B)$ dann das *Volumen* von B .

Bemerkung 46.4 (Riemann-integrierbar und Jordan-meßbar)

Es sei $B \subseteq \mathbb{R}^n$ beschränkt und $f : B \rightarrow \mathbb{R}$.

- Wegen Lemma 46.2 ist f_B genau dann über einem Quader integrierbar, der B enthält, wenn f_B über jedem Quader integrierbar ist, der B enthält, und in diesem Fall hängt der Wert des Integrals von f_B nicht vom gewählten Quader ab. Die Definition der Integrierbarkeit und des Integrals sind also unabhängig von Q .
- Ist B ein Quader, so können wir $Q = B$ wählen und erhalten genau die Definition der Integrierbarkeit und des Integrals aus Definition 44.9.
- Jeder Quader B ist wegen Beispiel 44.11 Jordan-meßbar und die Definition des Volumens in Definition 44.1 stimmt mit der in Definition 46.3 überein.

Beispiel 46.5 (Die leere Menge hat das Volumen 0.)

Die leere Menge ist Jordan-meßbar mit

$$V(\emptyset) = \int_{\emptyset} 1 \, dx = 0.$$

Dies sieht man, indem man einen beliebigen Quader Q wählt und

$$\int_Q 1_{\emptyset}(x) \, dx = \int_Q 0 \, dx = 0$$

berechnet.

Satz 46.6 (Kriterium für Jordan-Meßbarkeit)

$B \subset \mathbb{R}^n$ beschränkt ist genau dann Jordan-meßbar, wenn ∂B eine Nullmenge ist.

Beweis: Wir wählen zunächst einen Quader Q in \mathbb{R}^n , dessen Inneres B enthält, und wir beachten, daß die Funktion 1_B auf der offenen Menge $Q \setminus \partial B$ stetig ist, und daß ihre Unstetigkeitsstellen genau die Randpunkte ∂B sind.

B ist genau dann Jordan-meßbar, wenn die Funktion 1_B auf Q integrierbar ist, und dies ist wegen des Lebesgueschen Integrabilitätskriteriums 45.13 genau dann der Fall wenn die Menge ∂B ihrer Unstetigkeitsstellen eine Nullmenge ist. \square

Beispiel 46.7 (Die Kreisscheibe ist Jordan-meßbar.)

Die Kreisscheibe

$$D = \{x \in \mathbb{R}^2 \mid \|x\|_2 \leq 1\}$$

hat die Kreislinie S^1 als Rand, die nach Beispiel 45.9 eine Nullmenge ist. Mithin ist D Jordan-meßbar mit

$$V(D) = \int_D 1 \, dx.$$

Wir sind im Augenblick aber noch nicht in der Lage, dieses Integral auszurechnen.

Korollar 46.8 (Rechenregeln für Jordan-Meßbarkeit)

Sind $A, B \subseteq \mathbb{R}^n$ Jordan-meßbar, so sind auch $A \cup B$, $A \cap B$ und $A \setminus B$ Jordan-meßbar.

Beweis: A und B sind als Jordan-meßbare Mengen beschränkt, und der Rand der somit ebenfalls beschränkten Mengen $A \cup B$, $A \cap B$ und $A \setminus B$ ist jeweils enthalten in der Nullmenge $\partial A \cup \partial B$ und ist mithin eine Nullmenge (siehe auch Abbildung 17). Die Aussage folgt dann aus Satz 46.6. \square

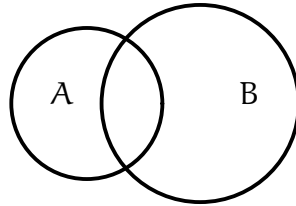


ABBILDUNG 17. $\partial(A \cup B), \partial(A \cap B), \partial(A \setminus B) \subseteq \partial A \cup \partial B$

B) Riemann-Integrale über Jordan-meßbaren Mengen

Wir haben in Definition 46.3 zwar den Begriff des Riemann-Integrals über beliebigen Teilmengen des \mathbb{R}^n eingeführt, eine wirklich gute Theorie ergibt sich aber nur über Jordan-meßbaren Mengen. Deshalb beschränken wir uns im folgenden auf diese.

Satz 46.9 (Lebesguesches Integrabilitätskriterium)

Es sei $B \subset \mathbb{R}^n$ Jordan-meßbar und $f : B \rightarrow \mathbb{R}$ sei beschränkt.

Genau dann ist f integrierbar auf B , wenn f fast überall stetig auf B ist.

Beweis: Wir wählen einen Quader Q , der B enthält.

Ist f auf B integrierbar, so ist f_B nach dem Lebesgueschen Integrabilitätskriterium für Quader 45.13 fast überall stetig auf Q . Da die Unstetigkeitsstellen von f auf B aber eine Teilmenge der Unstetigkeitsstellen von f_B auf Q sind, ist dann f erst recht fast überall stetig auf B .

Ist umgekehrt f fast überall stetig auf B , so gibt es eine Nullmenge N in \mathbb{R}^n , die die Unstetigkeitsstellen von f auf B enthält. Da B Jordan-meßbar ist, ist der Rand von B nach Satz 46.23 eine Nullmenge. Dann ist aber f_B auf Q stetig außerhalb der Nullmenge $N \cup \partial B$. Mithin ist f_B mit Satz 45.13 integrierbar auf Q und somit ist f integrierbar auf B . \square

Korollar 46.10 (Stetige Funktionen auf Kompakta sind integrierbar.)

Ist B kompakt und Jordan-meßbar und $f : B \rightarrow \mathbb{R}$ stetig, so ist f auf B integrierbar.

Beweis: Ist B in \mathbb{R}^n kompakt, so nimmt f auf B sein Minimum und sein Maximum an und ist somit insbesondere beschränkt auf B . Die Aussage folgt dann aus dem Lebesgueschen Integrabilitätskriterium für Jordan-meßbare Mengen 46.9. \square

Korollar 46.11 (Rechenregeln für Integrale)

Es sei $B \subset \mathbb{R}^n$ Jordan-meßbar und $f, g : B \rightarrow \mathbb{R}$ seien integrierbar auf B .

- a. Für $c, d \in \mathbb{R}$ ist die Funktion $c \cdot f + d \cdot g$ integrierbar auf B mit

$$\int_B (c \cdot f + d \cdot g)(x) \, dx = c \cdot \int_B f(x) \, dx + d \cdot \int_B g(x) \, dx.$$

- b. $|f|$ ist auf B integrierbar und es gilt die Dreiecksungleichung

$$\left| \int_B f(x) \, dx \right| \leq \int_B |f(x)| \, dx.$$

- c. Gilt $f(x) \leq g(x)$ für alle $x \in B$, so gilt auch $\int_B f(x) \, dx \leq \int_B g(x) \, dx$.

- d. Die Funktion $f \cdot g$ ist integrierbar auf B .

- e. $\exists q > 0$ mit $|g(x)| \geq q$ für alle $x \in B$, so ist auch $\frac{f}{g}$ integrierbar auf B .

Beweis: Wir wählen einen Quader Q im \mathbb{R}^n , der B enthält.

- a. Mit f und g ist auch $c \cdot f + d \cdot g$ beschränkt und fast überall stetig auf B . Mithin ist $c \cdot f + d \cdot g$ dann nach dem Lebesgueschen Integrierbarkeitskriterium für Jordan-meßbare Mengen 46.9 integrierbar auf B . Aus Korollar 44.19 folgt dann

$$\begin{aligned} \int_B (c \cdot f + d \cdot g)(x) \, dx &= \int_Q (c \cdot f + d \cdot g)_B(x) \, dx \\ &= \int_Q (c \cdot f_B + d \cdot g_B)(x) \, dx \\ &\stackrel{44.19}{=} c \cdot \int_Q f_B(x) \, dx + d \cdot \int_Q g_B(x) \, dx \\ &= c \cdot \int_B f(x) \, dx + d \cdot \int_B g(x) \, dx. \end{aligned}$$

- b. Mit f ist auch $|f|$ beschränkt und fast überall stetig auf B und deshalb integrierbar nach dem Lebesgueschen Integrierbarkeitskriterium 46.9. Aus Proposition 44.20 folgt dann

$$\begin{aligned} \left| \int_B f(x) \, dx \right| &\leq \left| \int_Q f_B(x) \, dx \right| \\ &\stackrel{44.20}{\leq} \int_Q |f_B(x)| \, dx = \int_Q |f|_B(x) \, dx \\ &= \int_B |f(x)| \, dx. \end{aligned}$$

- c. Für alle $x \in Q$ gilt $f_B(x) \leq g_B(x)$, so daß die Aussage aus der Monotonie des Integrals auf Quadern 44.19 folgt:

$$\int_B f(x) \, dx = \int_Q f_B(x) \, dx \leq \int_Q g_B(x) \, dx = \int_B g(x) \, dx.$$

- d. Mit f und g ist auch $f \cdot g$ beschränkt und fast überall stetig auf B . Die Behauptung folgt dann aus dem Lebesgueschen Integrierbarkeitskriterium 46.9.
- e. Aufgrund der Voraussetzung an g ist mit f und g auch $\frac{f}{g}$ beschränkt und fast überall stetig auf B . Die Behauptung folgt dann aus dem Lebesgueschen Integrierbarkeitskriterium 46.9.

□

Korollar 46.12 (Mittelwertsatz der Integralrechnung)

Sei $B \subset \mathbb{R}^n$ Jordan-meßbar und $f : B \rightarrow \mathbb{R}$ integrierbar auf B . Dann gilt

$$\inf_{x \in B} f(x) \cdot V(B) \leq \int_B f(x) \, dx \leq \sup_{x \in B} f(x) \cdot V(B).$$

Beweis: Da f als integrierbare Funktion auf B beschränkt ist, gilt

$$m := \inf_{x \in B} f(x) \in \mathbb{R}$$

und

$$M := \sup_{x \in B} f(x) \in \mathbb{R}.$$

Zudem folgt aus den Rechenregeln für Integrale 46.11

$$m \cdot V(B) = m \cdot \int_B 1 \, dx = \int_B m \, dx \leq \int_B f(x) \, dx \leq \int_B M \, dx = M \cdot \int_B 1 \, dx = M \cdot V(B).$$

□

Korollar 46.13 (Integration auf Jordan-meßbaren Teilmengen)

Sind $A \subseteq B \subset \mathbb{R}^n$ Jordan-meßbar und ist $f : B \rightarrow \mathbb{R}$ auf B integrierbar, so ist f auch auf A integrierbar.

Beweis: Ist f auf B beschränkt und fast überall stetig, so ist f auch auf A beschränkt und fast überall stetig. Die Aussage folgt also aus dem Lebesgueschen Integrierbarkeitskriterium für Jordan-meßbare Mengen 46.9. □

Satz 46.14 (Additivität des Integrals)

Seien $A, B \subset \mathbb{R}^n$ Jordan-meßbar und $f : A \cup B \rightarrow \mathbb{R}$ sei integrierbar auf A und B . Dann ist f auch integrierbar auf $A \cup B$ und auf $A \cap B$ und es gilt

$$\int_{A \cup B} f(x) \, dx = \int_A f(x) \, dx + \int_B f(x) \, dx - \int_{A \cap B} f(x) \, dx.$$

Beweis: Nach Korollar 46.8 sind die Menge $A \cup B$ und $A \cap B$ Jordan-meßbar. Zudem ist die Funktion f wegen des Lebesgueschen Integrierbarkeitskriteriums 46.9 beschränkt und fast überall stetig auf A und auf B , also auch auf $A \cup B$ und auf $A \cap B$. Wiederum mit Satz 46.9 ist f also auf $A \cup B$ und $A \cap B$ integrierbar.

Wir betrachten zunächst den Fall $A \cap B = \emptyset$. Dann gilt offenbar

$$f_{A \cup B} = f_A + f_B,$$

und wählen wir nun einen Quader Q in \mathbb{R}^n mit $A \cup B \subseteq Q$, so erhalten wir

$$\begin{aligned} \int_A f(x) \, dx + \int_B f(x) \, dx &= \int_Q f_A(x) \, dx + \int_Q f_B(x) \, dx \\ &= \int_Q (f_A + f_B)(x) \, dx = \int_Q f_{A \cup B}(x) \, dx = \int_{A \cup B} f(x) \, dx. \end{aligned}$$

Damit ist die Aussage für diesen Fall gezeigt, da

$$\int_{A \cap B} f(x) \, dx = \int_{\emptyset} f(x) \, dx = \int_Q f_{\emptyset}(x) \, dx = \int_Q 0 \, dx = 0.$$

Für den allgemeinen Fall beachten wir die Zerlegungen

$$A = (A \setminus B) \cup (A \cap B) \quad (201)$$

und

$$B = (B \setminus A) \cup (A \cap B) \quad (202)$$

sowie

$$A \cup B = (A \setminus B) \cup (B \setminus A) \cup (A \cap B). \quad (203)$$

Wenden wir den Spezialfall nun mehrfach an, so erhalten wir in Kurzform

$$\int_A f \stackrel{(201)}{=} \int_{A \setminus B} f + \int_{A \cap B} f \quad (204)$$

und

$$\int_B f \stackrel{(202)}{=} \int_{B \setminus A} f + \int_{A \cap B} f \quad (205)$$

sowie

$$\begin{aligned} \int_{A \cup B} f &\stackrel{(203)}{=} \int_{(A \setminus B) \cup (B \setminus A)} f + \int_{A \cap B} f \stackrel{(203)}{=} \int_{A \setminus B} f + \int_{B \setminus A} f + \int_{A \cap B} f \\ &\stackrel{(204), (205)}{=} \int_A f - \int_{A \cap B} f + \int_B f - \int_{A \cap B} f + \int_{A \cap B} f = \int_A f + \int_B f - \int_{A \cap B} f. \end{aligned}$$

□

Korollar 46.15 (Additivität des Volumens)

Sind $A, B \subset \mathbb{R}^n$ Jordan-meßbar, so gilt

$$V(A \cup B) = V(A) + V(B) - V(A \cap B).$$

Insbesondere folgt aus $A \subseteq B$ auch $V(A) \leq V(B)$.

Beweis: Man wende Satz 46.14 auf die Funktion $f : A \cup B \rightarrow \mathbb{R} : x \mapsto 1$ an und erhält

$$V(A \cup B) = V(A) + V(B) - V(A \cap B).$$

Ist nun $A \subseteq B$, so ist $B = A \cup (B \setminus A)$ und wir erhalten mithin

$$V(B) = V(A) + V(B \setminus A) - V(\emptyset) = V(A) + V(B \setminus A) \geq V(A).$$

□

C) Jordan-Nullmengen

Definition 46.16 (Jordan-Nullmenge)

Eine Jordan-meßbare Menge $B \subseteq \mathbb{R}^n$ heißt *Jordan-Nullmenge*, falls $V(B) = 0$.

Wir haben jetzt zwei verschiedene Nullmengen-Begriffe kennen gelernt, und um des besseren Verständnisses willen sollten wir zunächst klären, welche Beziehungen zwischen den beiden Begriffen gelten. Das folgende Lemma zeigt, daß Jordan-Nullmengen durch eine ähnliche Überdeckungseigenschaft charakterisiert werden wie die Nullmengen. Die Bedingung ist jedoch restriktiver, da stets endlich viele Quader ausreichen müssen.

Lemma 46.17 (Überdeckungseigenschaft von Jordan-Nullmengen)

Eine Teilmenge B in \mathbb{R}^n ist genau dann eine Jordan-Nullmenge, wenn B für jedes $\varepsilon > 0$ eine endliche Überdeckung aus Quadern mit Volumensumme kleiner ε besitzt.

Beweis: Sei zunächst B eine Jordan-Nullmenge und sei $\varepsilon > 0$ gegeben. Dann können wir einen Quader Q wählen, der B enthält, und die Funktion 1_B ist integrierbar auf Q . Da das Integral das Infimum aller Obersummen von 1_B ist, gibt es eine Zerlegung Z von Q mit

$$\varepsilon > OS(1_B, Z) - \int_Q 1_B(x) \, dx = \sum_{P \in TQ(Z)} V(P) \cdot \sup_{x \in P} 1_B(x) = \sum_{\substack{P \in TQ(Z) \\ P \cap B \neq \emptyset}} V(P).$$

Damit haben wir die gesuchte Überdeckung gefunden, nämlich

$$B \subseteq \bigcup_{\substack{P \in TQ(Z) \\ P \cap B \neq \emptyset}} P.$$

Gebe es nun umgekehrt zu jedem $\varepsilon > 0$ eine endliche Überdeckung von B durch Quader $Q_1^\varepsilon, \dots, Q_{m_\varepsilon}^\varepsilon$ mit

$$\sum_{i=1}^{m_\varepsilon} V(Q_i^\varepsilon) < \varepsilon.$$

Wir halten zunächst fest, daß B dann notwendigerweise beschränkt sein muß. Da die Vereinigung der Quader $Q_1^\varepsilon, \dots, Q_{m_\varepsilon}^\varepsilon$ abgeschlossen ist, enthält sie mit B auch den Abschluß von B und damit den Rand von B ,

$$\partial B \subseteq \bar{B} \subseteq \bigcup_{i=1}^{m_\varepsilon} Q_i^\varepsilon.$$

Da $\varepsilon > 0$ beliebig war, muß mithin ∂B eine Nullmenge sein. Nach Satz 46.6 ist B dann auch Jordan-meßbar. Zudem gilt wegen der Additivität des Volumens, Korollar 46.15, für das Volumen von B

$$V(B) \stackrel{46.15}{\leq} v\left(\bigcup_{i=1}^{m_\varepsilon} Q_i^\varepsilon\right) \stackrel{46.15}{\leq} \sum_{i=1}^{m_\varepsilon} V(Q_i^\varepsilon) < \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, muß das Volumen $V(B)$ notwendigerweise null sein. \square

Bemerkung 46.18 (Überdeckungseigenschaft von Jordan-Nullmengen)

Wie in Proposition 45.4 sieht man, daß man zur Überdeckung einer Jordan-Nullmenge genauso gut auch offene Quader oder Würfel nehmen kann.

Beispiel 46.19 (Eine Nullmenge muß keine Jordan-Nullmenge sein.)

Wegen Beispiel 45.6 ist $B = [0, 1] \cap \mathbb{Q}$ eine Nullmenge in \mathbb{R} , aber B ist keine Jordan-Nullmenge, da die charakteristische Funktion 1_B die Differenz der integrierbaren Funktion konstant 1 und der nicht integrierbaren Dirichletschen Sprungfunktion ist.

Korollar 46.20 (Nullmengen und Jordan-Nullmengen)

- Jede Jordan-Nullmenge in \mathbb{R}^n ist eine Nullmenge in \mathbb{R}^n .
- Jede kompakte Nullmenge in \mathbb{R}^n ist eine Jordan-Nullmenge in \mathbb{R}^n .
- Jede Jordan-meßbare Nullmenge in \mathbb{R}^n ist ein Jordan-Nullmenge in \mathbb{R}^n .

Beweis: Aus Lemma 46.17 folgt unmittelbar, daß jede Jordan-Nullmenge auch eine Nullmenge ist, und aus Proposition 45.4 d. zusammen mit Lemma 46.17 folgt, daß kompakte Nullmengen auch Jordan-Nullmengen sind. Schließlich folgt aus Aufgabe 45.18, daß eine Jordan-meßbare Nullmenge N auch Volumen null hat, da die charakteristische Funktion 1_N auf einem N enthaltenden Quader Q fast überall mit der Nullfunktion übereinstimmt. \square

Beispiel 46.21 (Eine Jordan-Nullmenge muß nicht kompakt sein.)

Die Menge

$$[0, 1] \times \{0\}$$

ist als Seite eines Quaders eine kompakte Nullmenge in \mathbb{R}^2 und mithin eine Jordan-Nullmenge. Ihre Teilmenge

$$(0, 1) \times \{0\} \subseteq \mathbb{R}^2$$

ist dann ebenfalls eine Jordan-Nullmenge, die aber nicht kompakt ist.

Beispiel 46.22 (Die Kreislinie ist eine Jordan-Nullmenge.)

Die Kreislinie $S^1 = V(x_1^2 + x_2^2 - 1)$ ist nach Beispiel 45.9 eine Nullmenge und sie ist kompakt. Mithin ist S^1 eine Jordan-Nullmenge.

Satz 46.23 (Kriterium für Jordan-Meßbarkeit)

Für eine beschränkte Menge $B \subseteq \mathbb{R}^n$ sind die folgenden Aussagen gleichwertig:

- B ist Jordan-meßbar.
- ∂B ist eine Nullmenge.
- ∂B ist eine Jordan-Nullmenge.

Beweis: Die Äquivalenz von a. und b. wurde in Satz 46.6 gezeigt, und die Äquivalenz von b. und c. folgt aus Korollar 46.20, da die abgeschlossene und beschränkte Menge ∂B nach dem Satz von Heine-Borel 39.44 kompakt ist. \square

Proposition 46.24 (Integrale über Jordan-Nullmengen sind null.)

Ist B eine Jordan-Nullmenge und $f : B \rightarrow \mathbb{R}$ beschränkt, so ist f integrierbar mit

$$\int_B f(x) \, dx = 0.$$

Beweis: Da die Jordan-Nullmenge eine Nullmenge ist, ist f auf B automatisch fast überall stetig. Nach dem Lebesgueschen Integrabilitätskriterium 46.9 ist die beschränkte Funktion f also integrierbar auf B . Aus dem Mittelwertsatz der Integraltheorie 46.12 folgt zudem

$$0 = \inf_{x \in B} f(x) \cdot V(B) \leq \int_B f(x) \, dx \leq \sup_{x \in B} f(x) \cdot V(B) = 0.$$

□

Bemerkung 46.25

Man beachte, daß die analoge Aussage für Nullmengen statt Jordan-Nullmengen im allgemeinen nicht gilt. Daß das Riemann-Integral es nicht erlaubt, jeder Nullmenge durch Integration das Volumen null zu zuordnen, ist einer der wesentlichen Gründe, das Riemann-Integral durch das Lebesgue-Integral zu ersetzen.

Wir betrachten die Nullmenge $N = [0, 1] \cap \mathbb{Q}$ und die konstante Funktion 1 auf N . Wäre diese dort integrierbar, so wäre 1_N auf $[0, 1]$ integrierbar, im Widerspruch dazu, daß die Dirichletsche Sprungfunktion $1 - 1_N$ auf $[0, 1]$ nicht integrierbar ist.

Korollar 46.26 (Integrale Jordan-fast überall gleicher Funktionen sind gleich.)

Sei $B \subset \mathbb{R}^n$ Jordan-meßbar, $f : B \rightarrow \mathbb{R}$ sei integrierbar auf B und $g : B \rightarrow \mathbb{R}$ sei beschränkt und stimme mit f außerhalb einer Jordan-Nullmenge überein. Dann ist auch g integrierbar auf B mit

$$\int_B g(x) \, dx = \int_B f(x) \, dx.$$

Beweis: Sei N eine Jordan-Nullmenge außerhalb derer f und g identisch sind, dann ist die Funktion $g - f$ auf $B \setminus N$ die Nullfunktion und $B \setminus N$ ist nach Korollar 46.8 Jordan-meßbar. Aus der Additivität des Integrals 46.14 und Proposition 46.24 folgt dann daß $g - f$ auf $B = (B \setminus N) \cup N$ integrierbar ist mit

$$\int_B (g - f)(x) \, dx = \int_{B \setminus N} (g - f)(x) \, dx + \int_N (g - f)(x) \, dx \stackrel{46.24}{=} \int_{B \setminus N} 0 \, dx + 0 = 0.$$

Dann ist aber auch g als Summe der integrierbaren Funktionen f und $g - f$ auf B integrierbar und aus der Linearität des Integrals folgt

$$\int_B g(x) \, dx = \int_B (g - f)(x) \, dx + \int_B f(x) \, dx = \int_B f(x) \, dx.$$

□

Beispiel 46.27

Die Aussage in Korollar 46.26 gilt im allgemeinen nicht, wenn sich f und g auf einer

Nullmenge unterscheiden, die keine Jordan-Nullmenge ist, wie man wieder an der nicht-integrierbaren Dirichletschen Sprungfunktion

$$g : [0, 1] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x \in \mathbb{Q}, \\ 1, & \text{falls } x \notin \mathbb{Q} \end{cases}$$

sieht, die auf sich nur auf der Nullmenge $[0, 1] \cap \mathbb{Q}$ (siehe Beispiel 45.6) von der integrierbaren konstanten Funktion

$$f : [0, 1] \longrightarrow \mathbb{R} : x \mapsto 1$$

unterscheidet.

Beispiel 46.28

Ist $Q \subset \mathbb{R}^n$ ein Quader und $f : Q \longrightarrow \mathbb{R}$ ist integrierbar auf Q , dann ist auch

$$g : Q \longrightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } x \in \overset{\circ}{Q}, \\ 0, & \text{falls } x \in \partial Q, \end{cases}$$

integrierbar auf Q und es gilt

$$\int_Q f(x) \, dx = \int_Q g(x) \, dx,$$

weil sich f und g nur auf ∂Q unterscheiden, die als kompakte Nullmenge eine Jordan-Nullmenge ist.

Korollar 46.29 (Spezielle Additivität des Integrals)

Seien $A, B \subset \mathbb{R}^n$ Jordan-meßbar, $A \cap B$ eine Jordan-Nullmenge und $f : A \cup B \longrightarrow \mathbb{R}$ sei integrierbar auf $A \cup B$. Dann gilt

$$\int_{A \cup B} f(x) \, dx = \int_A f(x) \, dx + \int_B f(x) \, dx.$$

Beweis: Dies folgt aus der Additivität des Integral 46.14 und Proposition 46.24. \square

Beispiel 46.30

Ist Z eine Zerlegung des Quaders Q und $f : Q \longrightarrow \mathbb{R}$ ist integrierbar, so gilt

$$\int_Q f(x) \, dx = \sum_{P \in \text{TQ}(Z)} \int_P f(x) \, dx.$$

Dies geht, da die Schnitte der Quader in $\text{TQ}(Z)$ als kompakte Nullmengen stets Jordan-Nullmengen sind.

D) Der Satz von Fubini für Normalbereiche

Lemma 46.31 (Graphen integrierbarer Funktionen sind Jordan-Nullmengen.)

Sei $B \subseteq \mathbb{R}^n$ Jordan-meßbar und $f : B \longrightarrow \mathbb{R}$ integrierbar, so ist der Graph

$$\text{Graph}(f) = \{(x, f(x))^t \mid x \in B\}$$

von f eine Jordan-Nullmenge in \mathbb{R}^{n+1} .

Beweis: Sei Q ein Quader in \mathbb{R}^n , der B enthält. Nach Voraussetzung ist f_B auf Q integrierbar, so daß wir für jedes $\varepsilon > 0$ eine Zerlegung Z von Q finden mit

$$\varepsilon > OS(f, Z) - US(f, Z) = \sum_{P \in TQ(Z)} V(P) \cdot \left(\sup_{x \in P} f(x) - \inf_{x \in P} f(x) \right).$$

Definieren wir nun

$$Q_P := P \times \left[\inf_{x \in P} f(x), \sup_{x \in P} f(x) \right]$$

für $P \in TQ(Z)$, so überdecken die Quader Q_P mit $P \in TQ(Z)$ den Graphen von f_B auf Q und damit erst recht den Graphen von f auf B ,

$$\text{Graph}(f) \subseteq \text{Graph}(f_B) \subseteq \bigcup_{P \in TQ(Z)} Q_P.$$

Zudem ist ihr Gesamtvolumen

$$\sum_{P \in TQ(Z)} V(Q_P) = OS(f, Z) - US(f, Z) < \varepsilon.$$

Nach Lemma 46.17 ist der Graph von f dann eine Jordan-Nullmenge. \square

Beispiel 46.32 (Die Oberfläche einer Halbkugel ist eine Jordan-Nullmenge.)

Die Oberfläche der oberen Halbkugel

$$M = \{(x, y, z)^t \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1, z \geq 0\}$$

ist der Graph der stetigen Funktion

$$f: D \longrightarrow \mathbb{R}: (x, y)^t \mapsto \sqrt{1 - x^2 - y^2}$$

auf der Jordan-meßbaren Kreisscheibe

$$D = \{(x, y)^t \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}.$$

Mithin ist M eine Jordan-Nullmenge.

Satz 46.33 (Das Prinzip von Cavalieri)

Sei $B \subset \mathbb{R}^n$ Jordan-meßbar, $f, g: B \rightarrow \mathbb{R}$ integrierbar auf B mit $f(x) \leq g(x) \forall x \in B$.

Dann ist die Menge

$$B_{f,g} := \{(x, y)^t \in \mathbb{R}^{n+1} \mid x \in B, f(x) \leq y \leq g(x)\}$$

Jordan-meßbar in \mathbb{R}^{n+1} mit

$$V(B_{f,g}) = \int_B g(x) - f(x) \, dx.$$

Beweis: Um zu zeigen, daß $B_{f,g}$ Jordan-meßbar ist, reicht es wegen Satz 46.23, zu zeigen, daß der Rand von $B_{f,g}$ eine Nullmenge ist.

Wir zeigen zunächst, daß ein Punkt $(x, y)^t \in B_{f,g}$ mit $x \in \overset{\circ}{B}$, $f(x) < y < g(x)$ und f und g stetig in x stets ein innerer Punkt von $B_{f,g}$ ist. Wegen der Stetigkeit in x gibt es zu

$$\varepsilon := \frac{\min\{y - f(x), g(x) - y\}}{4} > 0$$

ein $\delta_\varepsilon > 0$, so daß $|f(z) - f(x)|, |g(z) - g(x)| < \varepsilon$ für alle $z \in B$ mit $\|z - x\|_\infty < \delta_\varepsilon$. Da x im Inneren von B liegt, können wir ohne Einschränkung annehmen, daß $U_{\delta_\varepsilon}(x) \subseteq B$. Für $\gamma = \min\{\delta_\varepsilon, \varepsilon\} > 0$ ist dann aber die Umgebung

$$U_\gamma((x, y)^t) = \{(z, u)^t \mid \|z - x\|_\infty < \gamma, |u - y| < \gamma\} \subseteq B_{f,g}$$

ganz in $B_{f,g}$ enthalten, da $z \in B$ wegen $\|z - x\|_\infty < \gamma \leq \delta_\varepsilon$ und

$$u \in (y - \gamma, y + \gamma) \subseteq (y - \varepsilon, y + \varepsilon) \subseteq (f(z), g(z)).$$

Der Punkt $(x, y)^t$ ist unter obigen Voraussetzungen also innerer Punkt von $B_{f,g}$, und für den Rand $B_{f,g}$ gilt mithin

$$\partial B_{f,g} \subseteq \text{Graph}(f) \cup \text{Graph}(g) \cup X$$

mit

$$X = \{(x, y)^t \in \mathbb{R}^{n+1} \mid x \in N, -M \leq y \leq M\}$$

wobei

$$N = \partial B \cup \{x \in B \mid f \text{ oder } g \text{ unstetig in } x\}$$

und

$$M := \max \left\{ \sup_{x \in B} |f(x)|, \sup_{x \in B} |g(x)| \right\}.$$

Da $\text{Graph}(f)$ und $\text{Graph}(g)$ nach Lemma 46.31 Nullmengen sind, ist $\partial B_{f,g}$ eine Nullmenge, wenn X eine solche ist.

Da B Jordan-meßbar ist, ist ∂B eine Nullmenge (siehe Satz 46.23), und da f und g integrierbar sind, ist auch die Menge ihrer Unstetigkeitsstellen eine Nullmenge (siehe Satz 45.13). Als Vereinigung dieser Nullmengen ist N mithin ebenfalls eine Nullmenge und besitzt deshalb für $\varepsilon > 0$ eine abzählbare Überdeckung durch Quader Q_m , $m \in \mathbb{N}$, mit

$$\sum_{m=0}^{\infty} V(Q_m) < \frac{\varepsilon}{2M}.$$

Die Quader

$$P_m := Q_m \times [-M, M] \subset \mathbb{R}^{n+1}$$

für $m \in \mathbb{N}$ überdecken dann X und für das $n + 1$ -dimensionale Gesamtvolumen der Überdeckung gilt dann

$$\sum_{m=0}^{\infty} V(P_m) = \sum_{m=0}^{\infty} V(Q_m) \cdot 2M < \varepsilon.$$

Dies zeigt, daß X eine Nullmenge ist.

Damit ist gezeigt, daß $B_{f,g}$ Jordan-meßbar ist, und es bleibt die Formel für das Volumen von $B_{f,g}$ zu zeigen. Dazu wählen wir einen Quader $Q \subset \mathbb{R}^n$, der B enthält, und betrachten dazu den Quader

$$P = Q \times [-M, M] \subset \mathbb{R}^{n+1}.$$

Für $\mathbf{x} \in B$ fest, ist die stückweise stetige Funktion

$$[-M, M] \longrightarrow \mathbb{R} : \mathbf{y} \mapsto 1_{B_{f,g}}(\mathbf{x}, \mathbf{y})$$

auf dem Intervall $[-M, M]$ integrierbar, und aus der Jordan-Meßbarkeit von $B_{f,g}$ folgt, daß $1_{B_{f,g}}$ zudem integrierbar auf P ist. Damit sind die Voraussetzungen des Satzes von Fubini für Quader 44.22 erfüllt und wir erhalten

$$\begin{aligned} V(B_{f,g}) &= \int_P 1_{B_{f,g}}(\mathbf{x}, \mathbf{y}) \, d(\mathbf{x}, \mathbf{y}) = \int_Q \int_{-M}^M 1_{B_{f,g}}(\mathbf{x}, \mathbf{y}) \, d\mathbf{y} \, d\mathbf{x} \\ &= \int_Q \int_{f_B(\mathbf{x})}^{g_B(\mathbf{x})} 1 \, d\mathbf{y} \, d\mathbf{x} = \int_Q g_B(\mathbf{x}) - f_B(\mathbf{x}) \, d\mathbf{x} = \int_B g(\mathbf{x}) - f(\mathbf{x}) \, d\mathbf{x}. \end{aligned}$$

□

Beispiel 46.34 (Der Flächeninhalt des Kreises)

Wir haben bereits in Beispiel 46.7 gesehen, daß die Kreisscheibe

$$D = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_2 \leq 1\}$$

Jordan-meßbar ist. Dies folgt auch aus Satz 46.33, da

$$D = B_{f,g}$$

für

$$f : [-1, 1] \longrightarrow \mathbb{R} : t \mapsto -\sqrt{1-t^2}$$

und

$$g : [-1, 1] \longrightarrow \mathbb{R} : t \mapsto \sqrt{1-t^2}.$$

Dadurch sind wir nun aber in der Lage das Volumen von D auszurechnen:

$$V(D) = \int_{-1}^1 2 \cdot \sqrt{1-t^2} \, dt = t \cdot \sqrt{1-t^2} + \arcsin(t) \Big|_{-1}^1 = \pi.$$

Das Volumen von D ist also genau der Flächeninhalt des Kreises, was ein Hinweis dafür ist, daß unsere Begriffsbildungen sinnvoll gewählt sind.

Definition 46.35 (Normalbereich)

Sei i_1, \dots, i_n eine Permutation der Zahlen von $1, \dots, n$.

Ein *Normalbereich* bezüglich $(x_{i_1}, \dots, x_{i_n})$ in \mathbb{R}^n ist eine Menge der Form

$$B = \{\mathbf{x} \in \mathbb{R}^n \mid \varphi_j(x_{i_1}, \dots, x_{i_{j-1}}) \leq x_{i_j} \leq \psi_j(x_{i_1}, \dots, x_{i_{j-1}}) \, \forall j = 1, \dots, n\},$$

wobei $\varphi_1, \psi_1 \in \mathbb{R}$ Konstanten sind, φ_j und ψ_j für $j = 2, \dots, n$ geeignete stetige reellwertige Funktionen auf Kompakta sind und zudem

$$\varphi_j \leq \psi_j$$

für alle $j = 1, \dots, n$ gilt. Wir bezeichnen den Normalbereich B auch mit

$$B = NB(\varphi_1, \psi_1, \dots, \varphi_n, \psi_n).$$

Proposition 46.36 (Normalbereiche)

Normalbereiche sind kompakt und Jordan-meßbar.

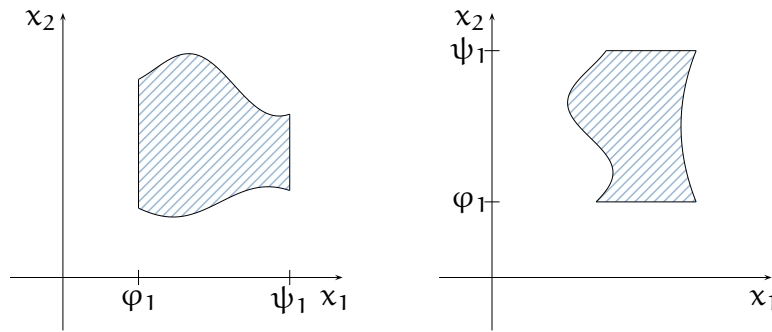


ABBILDUNG 18. Normalbereiche bez. (x_1, x_2) und bez. (x_2, x_1)

Beweis: Ohne Einschränkung können wir $i_j = j$ für $j = 1, \dots, n$ annehmen. Mit der Notation in Satz 46.33 und Definition 46.35 gilt für $k \in \{2, \dots, n\}$ und

$$B = \text{NB}(\varphi_1, \psi_1, \dots, \varphi_{k-1}, \psi_{k-1})$$

dann

$$\text{NB}(\varphi_1, \psi_1, \dots, \varphi_k, \psi_k) = B_{\varphi_k, \psi_k}.$$

Wenden wir Satz 46.33 nun $n - 1$ -fach an, so sehen wir, daß Normalbereiche Jordan-meßbar sind. Zudem sind sie beschränkt und abgeschlossen, da die Funktionen φ_j und ψ_j stetig sind. Nach dem Satz von Heine-Borel 39.44 sind Normalbereiche also auch kompakt. □

Satz 46.37 (Satz von Fubini für Normalbereiche)

Ist B ein Normalbereich wie in Definition 46.35 und ist $f : B \rightarrow \mathbb{R}$ stetig, so gilt

$$\int_B f(x) \, dx = \int_{\varphi_1}^{\psi_1} \int_{\varphi_2(x_{i_1})}^{\psi_2(x_{i_1})} \dots \int_{\varphi_n(x_{i_1}, \dots, x_{i_{n-1}})}^{\psi_n(x_{i_1}, \dots, x_{i_{n-1}})} f(x) \, dx_{i_n} \dots dx_{i_2} \, dx_{i_1}.$$

Beweis: Nach Proposition 46.36 ist B Jordan-meßbar und kompakt, so daß die stetige Funktion f auf B auch integrierbar ist. Um die Notationen zu vereinfachen, nehmen wir wieder ohne Einschränkung an, daß $i_j = j$ für $j = 1, \dots, n$, und wir führen den Beweis durch Induktion nach n , wobei für den Induktionsanfang $n = 1$ nichts zu zeigen ist.

Wir wählen nun einen Quader $Q = [a, b]$, der den Normalbereich B ganz enthält und für den $a_1 = \varphi_1$ und $b_1 = \psi_1$ gilt. Für jeden festen Wert $x_1 \in [a_1, b_1]$ ist die Menge

$$B' = \{(x_2, \dots, x_n) \in \mathbb{R}^{n-1} \mid \varphi_j(x_1, \dots, x_{j-1}) \leq x_j \leq \psi_j(x_1, \dots, x_{j-1}), j = 2, \dots, n\}$$

ein Normalbereich in \mathbb{R}^{n-1} und somit Jordan-meßbar. Zudem ist

$$Q' = [a_2, b_2] \times \dots \times [a_n, b_n] \subset \mathbb{R}^{n-1}$$

ein Quader in \mathbb{R}^{n-1} , der B' ganz enthält. Die Funktion

$$B' \rightarrow \mathbb{R} : x' \mapsto f(x_1, x')$$

ist stetig auf der kompakten Jordan-meßbaren Menge B' und deshalb integrierbar auf B' , d.h. die Funktion

$$Q' \longrightarrow \mathbb{R} : x' \mapsto f_B(x_1, x')$$

ist integrierbar auf Q' für jedes $x_1 \in [a_1, b_1]$ und mit Induktion erhalten wir

$$\int_{Q'} f(x_1, x') \, dx' = \int_{\varphi_2(x_1)}^{\psi_2(x_1)} \dots \int_{\varphi_n(x_1, \dots, x_{n-1})}^{\psi_n(x_1, \dots, x_{n-1})} f(x) \, dx_n \dots dx_2. \quad (206)$$

Außerdem ist die Funktion f_B integrierbar auf Q , so daß die Voraussetzungen des Satzes von Fubini für Quader 44.22 erfüllt sind. Wir erhalten deshalb

$$\begin{aligned} \int_B f(x) \, dx &= \int_Q f_B(x) \, dx = \int_{\varphi_1}^{\psi_1} \int_{Q'} f(x_1, x') \, dx' \, dx_1 \\ &\stackrel{(206)}{=} \int_{\varphi_1}^{\psi_1} \int_{\varphi_2(x_1)}^{\psi_2(x_1)} \dots \int_{\varphi_n(x_1, \dots, x_{n-1})}^{\psi_n(x_1, \dots, x_{n-1})} f(x) \, dx_n \dots dx_2 \, dx_1, \end{aligned}$$

womit die Behauptung gezeigt ist. □

Beispiel 46.38

Die Menge

$$B = \{(x, y)^t \mid -1 \leq x \leq 1, x^2 \leq y \leq 2 - x^2\}$$

ist der in Abbildung 19 skizzierte Normalbereich. Wir können nun mit Hilfe des

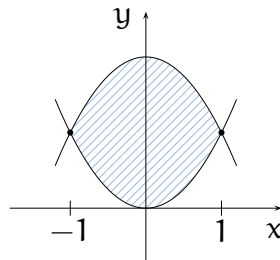


ABBILDUNG 19. Der Normalbereich aus Beispiel 46.38

Satzes von Fubini 46.37 die stetige Funktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y)^t \mapsto y$$

über B integrieren:

$$\begin{aligned} \int_B y \, d(x, y) &= \int_{-1}^1 \int_{x^2}^{2-x^2} y \, dy \, dx = \int_{-1}^1 \frac{y^2}{2} \Big|_{x^2}^{2-x^2} dx \\ &= \int_{-1}^1 2 - 2x^2 \, dx = 2x - \frac{2x^3}{3} \Big|_{-1}^1 = 4 - \frac{4}{3} = \frac{8}{3}. \end{aligned}$$

Beispiel 46.39

Die Menge

$$B = \{(x, y, z)^t \mid 0 \leq x \leq 1, 0 \leq y \leq 2, 0 \leq z \leq 2 - xy\}$$

ist ebenfalls ein Normalbereich (siehe Abbildung 20), und sein Volumen berechnet sich mit dem Satz von Fubini als

$$\begin{aligned} V(B) &= \int_B 1 \, d(x, y, z) = \int_0^1 \int_0^2 \int_0^{2-x \cdot y} 1 \, dz \, dy \, dx \\ &= \int_0^1 \int_0^2 (2 - x \cdot y) \, dy \, dx = \int_0^2 \int_0^1 (2 - x \cdot y) \, dx \, dy \\ &= \int_0^2 2 \cdot x - \frac{x^2 \cdot y}{2} \Big|_0^1 \, dy = \int_0^2 2 - \frac{y}{2} \, dy = 3. \end{aligned}$$

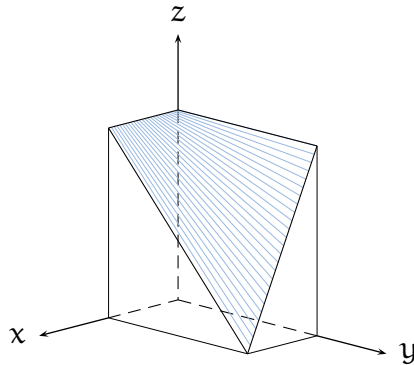


ABBILDUNG 20. Der Normalbereich aus Beispiel 46.39

Aufgaben

Aufgabe 46.40 (Abschlüsse von Jordan-Nullmengen sind Jordan-Nullmengen)

Ist $B \subset \mathbb{R}^n$ eine Jordan-Nullmengen, dann ist auch der Abschluß \bar{B} eine Jordan-Nullmenge.

Aufgabe 46.41

Zeige, ist W ein Würfel in \mathbb{R}^n und $f : W \rightarrow \mathbb{R}^n$ Lipschitz-stetig mit Lipschitz-Konstante q bezüglich der Maximumsnorm, dann gilt

$$V(f(W)) \leq q^n \cdot V(W).$$

Aufgabe 46.42 (Das Prinzip von Cavalieri)

Sei $B \subseteq [a, b] \subset \mathbb{R}^n$ Jordan-meßbar und für jedes $t \in [a_1, b_1]$ sei der Hyperebenenschnitt $H_t = B \cap V(x_1 - t)$ Jordan-meßbar mit Volumen $v(t)$, dann ist

$$V(B) = \int_{a_1}^{b_1} v(t) \, dt.$$

Ist zudem B kompakt und $f : B \rightarrow \mathbb{R}$ stetig, so gilt auch

$$\int_B f(x) \, dx = \int_{a_1}^{b_1} \int_{H_t} f(t, y) \, dy \, dt.$$

Aufgabe 46.43

Sei $j \in \{1, \dots, n\}$, $c \in \mathbb{R}$ und $M \subseteq V(x_j - c) \subseteq \mathbb{R}^n$ beschränkt. Zeige, daß M eine Jordan-Nullmenge in \mathbb{R}^n ist.

Aufgabe 46.44

Zeige, daß jede beschränkte Teilmenge des \mathbb{R}^n , die höchstens endlich viele Häufungspunkte besitzt, eine Jordan-Nullmenge ist.

Aufgabe 46.45

Zeige, daß die folgenden Mengen Normalbereiche bezüglich (x_1, x_2) und (x_2, x_1) sind.

- a. $B_1 = \{(x_1, x_2)^t \in \mathbb{R}^2 \mid 0 \leq x_1 \leq 1, 0 \leq x_2 \leq x_1\}$.
- b. $B_2 = \{(x_1, x_2)^t \in \mathbb{R}^2 \mid 0 \leq x_1 \leq \sin(x_2), 0 \leq x_2 \leq \frac{\pi}{2}\}$.

Aufgabe 46.46

Sei $B \subseteq \mathbb{R}^2$ der Normalbereich im ersten Quadranten zwischen der Geraden $x_2 = x_1$ und der Parabel $x_2 = x_1^2$. Berechne $\int_B x_1 x_2 \, d(x_1, x_2)$.

Aufgabe 46.47

Berechne das Volumen des Tetraeders, der von den drei Koordinatenachsen und der Ebene $x_3 = 2 - 2x_1 - x_2$ begrenzt wird.

§ 47 Der Transformationssatz für Integrale

A) Diffeomorphismen

Bemerkung 47.1

Ist $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ eine stetig differenzierbare Abbildung mit $\varphi'(x) \neq 0$ für alle $x \in \mathbb{R}$, dann folgt aus dem Zwischenwertsatz schon, daß die Ableitung stets positiv oder stets negativ ist. Die Funktion φ ist dann streng monoton wachsend oder streng monoton fallend, und für ein Intervall $[a, b] \subseteq \mathbb{R}$ gilt deshalb in ersterem Fall

$$\varphi([a, b]) = [\varphi(a), \varphi(b)]$$

und in letzterem Fall

$$\varphi([a, b]) = [\varphi(b), \varphi(a)].$$

Zudem wird der Rand $\partial[a, b] = \{a, b\}$ auf den Rand des Bildes abgebildet,

$$\varphi(\partial[a, b]) = \varphi(\{a, b\}) = \{\varphi(a), \varphi(b)\},$$

und das Innere $\overset{\circ}{[a, b]} = (a, b)$ wird auf das Innere des Bildes abgebildet. Man beachte auch, daß φ als streng monotone Funktion injektiv ist und damit bijektiv auf ihr Bild abbildet.

Wir haben schon im Zusammenhang mit dem Satz über implizite Funktionen gesehen, daß im Mehrdimensionalen die Bedingung $\varphi'(x) \neq 0$ durch $\det(D\varphi(x)) \neq 0$ ersetzt werden muß, daß dies aber nicht mehr ausreicht, um die Injektivität folgern zu können. Diese muß zusätzlich vorausgesetzt werden. Solche Abbildungen nennt man dann Diffeomorphismen (siehe Bemerkung 43.19). Die folgende Proposition ist also die naheliegende Verallgemeinerung des eben beschriebenen Sachverhalts.

Proposition 47.2 (Bilder Jordan-meßbarer Mengen unter Diffeomorphismen)

Es sei $U \subseteq \mathbb{R}^n$ offen, $\varphi : U \rightarrow \mathbb{R}^n$ sei ein Diffeomorphismus und $B \subseteq U$ sei Jordan-meßbar und kompakt. Dann ist $\varphi(B)$ Jordan-meßbar mit $\partial\varphi(B) = \varphi(\partial B)$.

Wir beweisen zunächst eine etwas schwächere Aussage unter einer ebenfalls abgeschwächten und technischeren Voraussetzung.

Lemma 47.3 (Bilder Jordan-meßbarer Mengen unter Diffeomorphismen)

Es sei $U \subseteq \mathbb{R}^n$ offen, N sei eine abgeschlossene Nullmenge in \mathbb{R}^n , $\varphi : U \rightarrow \mathbb{R}^n$ sei stetig differenzierbar auf U , und $B \subseteq U$ sei Jordan-meßbar und kompakt, so daß φ ein lokaler Diffeomorphismus auf $\overset{\circ}{B} \setminus N$ ist.

Dann ist $\varphi(B)$ kompakt und Jordan-meßbar mit $\partial\varphi(B) \subseteq \varphi(\partial B) \cup \varphi(N \cap B)$.

Beweis: Wir halten zunächst fest, daß φ stetig auf der kompakten Menge B ist, so daß $\varphi(B)$ nach Satz 40.23 wieder kompakt ist. Nach Korollar 39.32 ist die in B abgeschlossene Teilmenge ∂B ebenfalls kompakt. Zudem ist ∂B als Rand der Jordan-meßbaren Menge B nach Satz 46.23 eine nun kompakte Nullmenge, und da φ stetig differenzierbar auf U ist, ist dann nach Proposition 45.8 $\varphi(\partial B)$ ebenfalls eine kompakte Nullmenge.

Wir beachten ferner, daß $N \cap B$ als Teilmenge von N eine Nullmenge und als abgeschlossene Teilmenge von B kompakt ist. Demnach ist auch das stetig differenzierbare Bild $\varphi(N \cap B)$ eine Nullmenge.

Um zu zeigen, daß $\varphi(B)$ Jordan-meßbar ist, reicht es nun wegen Satz 46.23

$$\partial\varphi(B) \subseteq \varphi(\partial B) \cup \varphi(N \cap B) \quad (207)$$

zu zeigen, da der Rand von $\varphi(B)$ dann eine Nullmenge ist.

Dazu beachten wir nun zunächst, daß die Menge

$$\mathring{B} \setminus N = \mathring{B} \cap (\mathbb{R}^n \setminus N)$$

offen in \mathbb{R}^n ist, da N abgeschlossen ist. Außerdem ist φ nach Voraussetzung ein lokaler Diffeomorphismus auf $\mathring{B} \setminus N$, so daß nach dem Satz der offenen Abbildung 43.17 auch $\varphi(\mathring{B} \setminus N)$ offen in \mathbb{R}^n ist. Dann liegt diese Menge aber ganz im Inneren von $\varphi(B)$ und es gilt

$$\varphi(\mathring{B} \setminus N) \cap \partial\varphi(B) = \emptyset. \quad (208)$$

Wir beachten nun noch, daß die kompakte Menge $\varphi(B)$ abgeschlossen ist und mithin ihren Rand enthält. Ist also $\mathbf{y} \in \partial\varphi(B)$, so gibt es ein $\mathbf{x} \in B$ mit $\varphi(\mathbf{x}) = \mathbf{y}$. Wegen (208) gilt dabei $\mathbf{x} \notin \mathring{B} \setminus N$, und somit

$$\mathbf{x} \in \partial B \cup (N \cap B)$$

und

$$\mathbf{y} \in \varphi(\partial B) \cup \varphi(N \cap B).$$

Damit ist (207) gezeigt, und wir sind fertig. \square

Beweis von Proposition 47.2: Ist φ ein Diffeomorphismus auf U , so sind die Voraussetzungen von Lemma 47.3 mit $N = \emptyset$ erfüllt und $\varphi(B)$ ist kompakt und Jordan-meßbar mit

$$\partial\varphi(B) \subseteq \varphi(\partial B). \quad (209)$$

Es bleibt also nur, die umgekehrte Inklusion zu zeigen.

Aus dem Satz der offenen Abbildung 43.17 wissen wir aber, daß φ bijektiv auf sein Bild $\varphi(U)$ abbildet und daß

$$\varphi^{-1} : \varphi(U) \longrightarrow U$$

ebenfalls ein Diffeomorphismus ist, da

$$\det(D\varphi^{-1}(\mathbf{y})) = \det\left(\left(D\varphi(\varphi^{-1}(\mathbf{y}))\right)^{-1}\right) = \frac{1}{\det(D\varphi(\varphi^{-1}(\mathbf{y})))} \neq 0$$

für alle $\mathbf{y} \in \varphi(U)$. Wenden wir (209) also auf φ^{-1} und die kompakte Jordan-meßbare Menge $\varphi(B)$ an, so erhalten wir

$$\partial\varphi^{-1}(\varphi(B)) \subseteq \varphi^{-1}(\partial\varphi(B)).$$

Kombinieren wir dies mit (209), so sehen wir

$$\partial B = \partial \varphi^{-1}(\varphi(B)) \subseteq \varphi^{-1}(\partial \varphi(B)) \subseteq \varphi^{-1}(\varphi(\partial B)) = \partial B$$

und mithin

$$\partial B = \varphi^{-1}(\partial \varphi(B)).$$

Wenden wir auf beiden Seiten φ an, so sind wir fertig. □

Beispiel 47.4 (Parallelotope als Bilder eines Würfels)

Sei $A \in \text{Gl}_n(\mathbb{R})$ eine invertierbare Matrix, dann ist die zugehörige lineare Abbildung

$$f_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n : x \mapsto Ax$$

ein Diffeomorphismus, da f_A bijektiv mit $Df_A(x) = A$ für alle $x \in \mathbb{R}^n$ ist.

Betrachten wir nun den Würfel

$$W = [0, 1] \times \dots \times [0, 1],$$

so ist W eine kompakte, Jordan-meßbare Menge, und somit ist auch das Parallelotop

$$P(\mathbf{a}_1, \dots, \mathbf{a}_n) = \{\lambda_1 \mathbf{a}_1 + \dots + \lambda_n \mathbf{a}_n \in \mathbb{R}^n \mid 0 \leq \lambda_i \leq 1, i = 1, \dots, n\} = f_A(W)$$

Jordan-meßbar, wobei $\mathbf{a}_1, \dots, \mathbf{a}_n$ die Spalten von A sind.

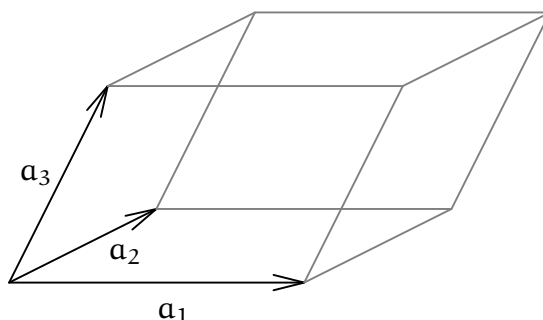


ABBILDUNG 21. Das Parallelotop $P(\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3)$ im \mathbb{R}^3

Bemerkung 47.5

Die Aussage von Proposition 47.2 zum Rand ist im allgemeinen nicht mehr richtig, wenn f nicht injektiv ist, also nur ein lokaler Diffeomorphismus ist. Die Polarkoordinatenabbildung

$$\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (r, \theta) \mapsto (r \cdot \cos(\theta), r \cdot \sin(\theta))$$

ist auf $(0, \infty) \times \mathbb{R}$ ein nicht injektiver lokaler Diffeomorphismus und der kompakte, Jordan-meßbare Quader

$$B = [r_1, r_2] \times [0, 2\pi]$$

hat den Kreisring

$$\varphi(B) = \{(x, y)^t \in \mathbb{R}^2 \mid r_1^2 \leq x^2 + y^2 \leq r_2^2\}$$

als Bild. Die Randstücke $[r_1, r_2] \times \{0\}$ und $[r_1, r_2] \times \{2\pi\}$ werden jedoch aufeinander abgebildet und liegen im Innern des Kreisrings. Es gilt aber nach wie vor, daß der Rand des Kreisrings im Bild des Randes des Quaders enthalten ist.

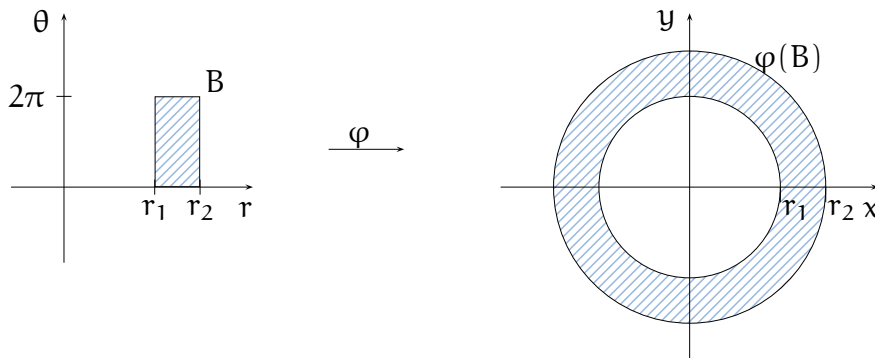


ABBILDUNG 22. Transformation eines Rechtecks auf einen Kreisringausschnitt

Setzen wir nun

$$N = \{0\} \times \mathbb{R},$$

so ist N eine abgeschlossene Nullmenge in \mathbb{R}^2 und f ist auf $\mathbb{R}^2 \setminus N$ ein lokaler Diffeomorphismus, so daß wir Lemma 47.3 auf φ anwenden können. Das Bild des kompakten Quaders

$$B = [0, r_0] \times [0, 2\pi]$$

ist also Jordan-meßbar, was wir ohnehin wußten, da es ein Kreis vom Radius r_0 um den Ursprung ist,

$$\varphi(B) = \{(x, y)^t \in \mathbb{R}^2 \mid x^2 + y^2 \leq r_0^2\}.$$

Man beachte, daß hier auch die Jordan-Nullmenge $\{0\} \times [0, 2\pi]$ im Inneren des Bildes verschwunden ist.

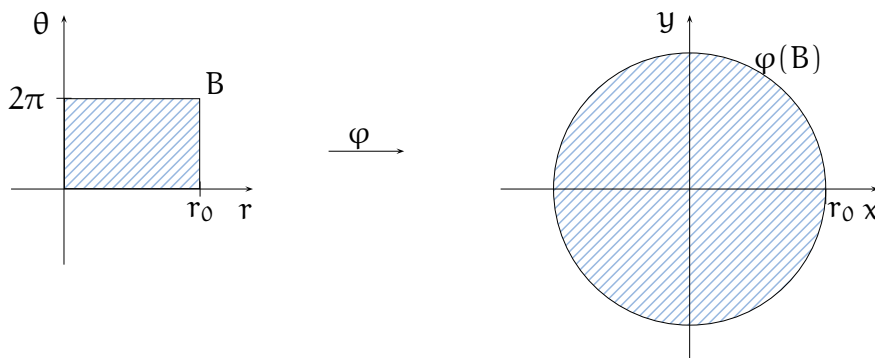


ABBILDUNG 23. Transformation eines Rechtecks auf einen Kreisringausschnitt

Situationen dieser Art sind für die Anwendungen im Zusammenhang mit dem Transformationsatz aber wichtig. Man sollte sie als Grenzprozesse der regulären Situation in Proposition 47.2 verstehen.

B) Der Transformationssatz für Integrale

Bemerkung 47.6 (Die Substitutionsregel)

Im Eindimensionalen haben wir die Substitutionsregel für Integrale kennengelernt. Ist $\varphi : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar und $f : \mathbb{R} \rightarrow \mathbb{R}$ stetig, so gilt

$$\int_{\varphi(a)}^{\varphi(b)} f(x) \, dx = \int_a^b f(\varphi(x)) \cdot \varphi'(x) \, dx.$$

Wir wollen diese Aussage gerne ins Mehrdimensionale verallgemeinern. Dazu würden wir die Integrationsbereiche bei den Integralen schon im Eindimensionalen gerne etwas anders schreiben, nämlich als

$$\int_{\varphi([a,b])} f(x) \, dx = \int_{[a,b]} f(\varphi(x)) \cdot \varphi'(x) \, dx.$$

Aber so ist die Formel schon im Eindimensionalen nicht mehr richtig, da die stetige Differenzierbarkeit von φ nicht ausreicht, um

$$\varphi([a, b]) = [\varphi(a), \varphi(b)] \quad \text{oder} \quad \varphi([a, b]) = [\varphi(b), \varphi(a)]$$

zu garantieren. Z.B. $[a, b] = [0, \pi]$ und $\varphi = \sin$, so ist $\varphi(a) = \varphi(b) = 0$, aber $\varphi([a, b]) = [0, 1]$.

Hier hilft die Forderung $\varphi'(x) \neq 0$ für alle x weiter, da sie die Monotonie von φ sichert. Falls φ' stets positiv ist, so gilt

$$\varphi([a, b]) = [\varphi(a), \varphi(b)]$$

und somit

$$\begin{aligned} \int_{\varphi([a,b])} f(x) \, dx &= \int_{\varphi(a)}^{\varphi(b)} f(x) \, dx = \int_a^b f(\varphi(x)) \cdot \varphi'(x) \, dx \\ &= \int_{[a,b]} f(\varphi(x)) \cdot \varphi'(x) \, dx = \int_{[a,b]} f(\varphi(x)) \cdot |\varphi'(x)| \, dx. \end{aligned}$$

Falls φ' stets negativ ist, so gilt

$$\varphi([a, b]) = [\varphi(b), \varphi(a)]$$

und somit

$$\begin{aligned} \int_{\varphi([a,b])} f(x) \, dx &= \int_{\varphi(b)}^{\varphi(a)} f(x) \, dx = - \int_{\varphi(a)}^{\varphi(b)} f(x) \, dx \\ &= - \int_a^b f(\varphi(x)) \cdot \varphi'(x) \, dx = \int_a^b f(\varphi(x)) \cdot |\varphi'(x)| \, dx \\ &= \int_{[a,b]} f(\varphi(x)) \cdot |\varphi'(x)| \, dx. \end{aligned}$$

Zusammen mit den Betrachtungen in Bemerkung 47.1 und Proposition 47.2 legt dies nahe, die Formel

$$\int_{\varphi(B)} f(x) \, dx = \int_B f(\varphi(x)) \cdot |\det D\varphi(x)| \, dx$$

zu erwarten, wenn f stetig, φ ein injektiver Diffeomorphismus und B kompakt und Jordan-meßbar ist.

Man muß aber noch ein klein wenig Vorsicht walten lassen, denn im Eindimensionalen haben wir den Zwischenwertsatz auf einem Intervall angewendet, um aus $\varphi' \neq 0$ zu folgern, daß φ' schon stets positiv oder negativ ist. Das können wir bei einem beliebigen B nicht mehr, wenn nicht die Verbindungsstrecke zwischen je zwei Punkten ganz in B liegt. Wir werden deshalb zusätzlich fordern, daß die Funktionaldeterminante ihr Vorzeichen nicht wechselt.

Satz 47.7 (Transformationssatz für Integrale)

Sei $U \subseteq \mathbb{R}^n$ offen, $\varphi : U \rightarrow \mathbb{R}^n$ ein Diffeomorphismus mit $\det(D\varphi(x))$ stets positiv oder stets negativ, $B \subseteq U$ kompakt und Jordan-meßbar und $f : \varphi(B) \rightarrow \mathbb{R}$ stetig.

Dann ist f auf der Jordan-meßbaren Menge $\varphi(B)$ integrierbar mit

$$\int_{\varphi(B)} f(x) \, dx = \int_B f(\varphi(x)) \cdot |\det D\varphi(x)| \, dx. \quad (210)$$

Bemerkung 47.8 (Allgemeiner Transformationssatz für Integrale)

Man kann die Voraussetzungen des Transformationssatzes etwas abschwächen.

Ist φ zwar auf ganz U stetig differenzierbar, aber nur auf $U \setminus N$ soll φ injektiv mit stets positiver oder stets negativer Funktionaldeterminante sein, während die Funktionaldeterminante auf der *Jordan-Nullmenge* N auch den Wert null annehmen darf, so ist $\varphi(B)$ immer noch Jordan-meßbar und die Transformationsformel (210) gilt auch noch. Wir beweisen diesen Sachverhalt in Lemma 47.22. \square

Ist die Menge B konvex, dann reicht es im Übrigen auch, daß φ ein lokaler Diffeomorphismus ist, da dann die Funktionaldeterminante ihr Vorzeichen nicht wechseln kann.

Lemma 47.9 (Funktionaldeterminante)

Ist $U \subseteq \mathbb{R}^n$ offen, $\varphi : U \rightarrow \mathbb{R}^n$ ein lokaler Diffeomorphismus und $C \subseteq U$ konvex, dann ändert die Funktionaldeterminante von φ auf C ihr Vorzeichen nicht.

Beweis: Sind $x, y \in C$ zwei beliebige Punkte, so liegt \overline{xy} in C , da C konvex ist. Die Funktion

$$h : [0, 1] \rightarrow \mathbb{R} : t \mapsto \det(D\varphi(x + t \cdot (y - x)))$$

ist stetig, da φ stetig differenzierbar ist und die Determinante ebenfalls eine stetige Abbildung ist. Zudem gilt

$$h(0) = \det(D\varphi(x))$$

und

$$h(1) = \det(D\varphi(y)).$$

Hätten die beiden Funktionaldeterminanten unterschiedliches Vorzeichen, so würde es wegen des Zwischenwertsatzes einen Punkt t geben mit

$$\det(D\varphi(x + t \cdot (y - x))) = h(t) = 0$$

im Widerspruch dazu, daß φ ein lokaler Diffeomorphismus ist. \square

In den meisten interessanten Fällen ist B ein Quader, aber φ nur im Inneren des Quaders ein Diffeomorphismus und wir verlieren die Injektivität nicht nur auf einer Nullmenge, wenn wir eine offene Umgebung von B als Definitionsbereich zulassen. Durch eine Grenzwertbetrachtung können wir die Transformationsformel aber auch in diesen Fällen retten, wie das folgende Korollar zeigt.

Korollar 47.10 (Transformationssatz für Integrale auf Quadern)

Sei $U \subseteq \mathbb{R}^n$ offen, $\varphi : U \rightarrow \mathbb{R}^n$ stetig differenzierbar, $Q \subseteq U$ ein Quader, φ auf dem Inneren \dot{Q} von Q ein Diffeomorphismus und $f : \varphi(Q) \rightarrow \mathbb{R}$ stetig.

Dann ist f auf der Jordan-meßbaren Menge $\varphi(Q)$ integrierbar mit

$$\int_{\varphi(Q)} f(x) \, dx = \int_Q f(\varphi(x)) \cdot |\det D\varphi(x)| \, dx.$$

Beweis: In diesem Beweis verwenden wir auf \mathbb{R}^n die Maximumsnorm.

Nach Voraussetzung ist die Funktion

$$h : Q \rightarrow \mathbb{R} : x \mapsto f(\varphi(x)) \cdot |\det(D\varphi(x))|$$

stetig auf dem Quader Q und somit integrierbar, was die Existenz des rechten Integrals sichert.

Wir betrachten nun für $Q = [a, b]$ und ein sehr kleines $\varepsilon > 0$ den Quader

$$Q_\varepsilon = [a_1 + \varepsilon, b_1 - \varepsilon] \times \dots \times [a_n + \varepsilon, b_n - \varepsilon] \subseteq Q$$

mit Volumen

$$V(Q_\varepsilon) = \prod_{i=1}^n (b_i - a_i - 2\varepsilon) \xrightarrow{\varepsilon \rightarrow 0} \prod_{i=1}^n (b_i - a_i) = V(Q). \quad (211)$$

Damit erhalten wir aus dem Mittelwertsatz der Integration 46.12 dann

$$\inf_{x \in Q} h(x) \cdot (V(Q) - V(Q_\varepsilon)) \leq \int_{Q \setminus Q_\varepsilon} h(x) \, dx \leq \sup_{x \in Q} h(x) \cdot (V(Q) - V(Q_\varepsilon)),$$

wobei die linke und die rechte Seite für $\varepsilon \rightarrow 0$ gegen null konvergieren. Also gilt dies auch für das Integral in der Mitte, und aus der speziellen Additivität des Integrals 46.29 erhalten wir dann

$$\int_{Q_\varepsilon} h(x) \, dx = \int_Q h(x) \, dx - \int_{Q \setminus Q_\varepsilon} h(x) \, dx \longrightarrow \int_Q h(x) \, dx$$

für $\varepsilon \rightarrow 0$.

Wir wollen nun eine analoge Betrachtung für das linke Integral anstellen, was aber komplizierter ist, da $\varphi(Q)$ kein Quader mehr ist.

Wenden wir Lemma 47.3 mit $N = \emptyset$ auf φ und Q an, so erhalten wir, daß $\varphi(Q)$ eine kompakte Jordan-meßbare Menge ist, auf der f zudem stetig und damit integrierbar ist. Das linke Integral existiert also ebenfalls.

Nach Korollar 42.19 ist die stetig differenzierbare Abbildung φ auf dem kompakten konvexen Quader Q zudem Lipschitz-stetig mit einer Lipschitz-Konstanten $q > 0$. Für einen Würfel

$$W = \overline{U_r(\mathbf{a})}$$

gilt dann (siehe auch den Beweis von Lemma 45.7)

$$\varphi(W) \subseteq \overline{U_{qr}(\varphi(\mathbf{a}))}$$

und damit (siehe auch Aufgabe 46.41)

$$V(\varphi(W)) \leq (2qr)^n = q^n \cdot V(W). \quad (212)$$

Das Komplement $Q \setminus Q_\varepsilon$ ist offensichtlich eine Vereinigung von Quadern, und wir haben im Beweis von Proposition 45.4 e. (siehe (194)) gesehen, daß wir jeden Quader durch endlich viele Würfel überdecken können, wobei sich das Volumen höchstens um den Faktor 2^n vergrößert. Zu ε gibt es also Würfel W_1, \dots, W_k mit

$$Q \setminus Q_\varepsilon \subseteq W_1 \cup \dots \cup W_k$$

und

$$\sum_{i=1}^k V(W_i) \leq 2^n \cdot V(Q \setminus Q_\varepsilon) = 2^n \cdot (V(Q) - V(Q_\varepsilon)).$$

Damit erhalten wir dann aber auch

$$\varphi(Q) \setminus \varphi(Q_\varepsilon) \subseteq \varphi(Q \setminus Q_\varepsilon) \subseteq \bigcup_{i=1}^k \varphi(W_i)$$

und somit

$$V(\varphi(Q) \setminus \varphi(Q_\varepsilon)) \leq \sum_{i=1}^k V(\varphi(W_i)) \stackrel{(212)}{\leq} q^n \cdot \sum_{i=1}^k V(W_i) \leq 2^n q^n \cdot (V(Q) - V(Q_\varepsilon)).$$

Für $\varepsilon \rightarrow 0$ konvergiert die rechte Seite wegen (211) gegen null, so daß auch

$$\lim_{\varepsilon \rightarrow 0} V(\varphi(Q) \setminus \varphi(Q_\varepsilon)) = 0$$

gilt. Aus dem Mittelwertsatz der Integration 46.12 erhalten wir dann wieder

$$\begin{aligned} \inf_{x \in \varphi(Q)} f(x) \cdot V(\varphi(Q) \setminus \varphi(Q_\varepsilon)) &\leq \int_{\varphi(Q) \setminus \varphi(Q_\varepsilon)} f(x) \, dx \\ &\leq \sup_{x \in \varphi(Q)} f(x) \cdot V(\varphi(Q) \setminus \varphi(Q_\varepsilon)), \end{aligned}$$

und für $\varepsilon \rightarrow 0$ konvergieren die linke und rechte Seite gegen null, so daß dies auch für das Integral in der Mitte gilt. Dann wenden wir wieder die spezielle Additivität des Integrals 46.29 an und erhalten für $\varepsilon \rightarrow 0$

$$\int_{\varphi(Q_\varepsilon)} f(x) \, dx = \int_{\varphi(Q)} f(x) \, dx - \int_{\varphi(Q) \setminus \varphi(Q_\varepsilon)} f(x) \, dx \longrightarrow \int_{\varphi(Q)} f(x) \, dx.$$

Nun können wir auf dem Quader Q_ε aber den Transformationssatz 47.7 anwenden, wobei wir beachten, daß die Funktionaldeterminante von φ auf der konvexen Menge $\overset{\circ}{B}$ nach Lemma 47.9 ihr Vorzeichen nicht wechselt, und erhalten somit insgesamt

$$\int_{\varphi(Q)} f(x) \, dx \longleftarrow \int_{\varphi(Q_\varepsilon)} f(x) \, dx \stackrel{47.7}{=} \int_{Q_\varepsilon} h(x) \, dx \longrightarrow \int_Q h(x) \, dx,$$

was zu beweisen war. \square

Bevor wir uns dem Beweis des Transformationssatzes 47.7 zuwenden wollen, betrachten wir Beispiele, die seine Bedeutung veranschaulichen sollen.

C) Anwendungen des Transformationssatzes für Integrale

Beispiel 47.11 (Das Volumen des Parallelotops)

In Bemerkung 37.32 haben wir das Volumen eines Parallelotops im \mathbb{R}^3 bestimmt. Diese Rechnung können wir nun mit Hilfe der Transformationsformel und von Beispiel 47.4 vereinfachen. Ist nämlich P das Parallelotop, das von den Spalten der Matrix A aufgespannt wird, so gilt

$$V(P) = \int_P 1 \, dx = \int_{f_A(W)} 1 \, dx = \int_W |\det(Df_A)| \, dx = \int_W |\det(A)| \, dx = |\det(A)|,$$

wenn wie in Beispiel 47.4 W der Würfel mit Kantenlänge Eins ist. Diese Formel gilt in dieser Form in jeder Dimension, so daß unsere Definition des Volumens in Bemerkung 30.12 sinnvoll war.

Beispiel 47.12 (Polarkoordinaten)

Die Polarkoordinatenabbildung

$$\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (r, \theta)^t \mapsto (r \cdot \cos(\theta), r \cdot \sin(\theta))$$

eingeschränkt auf die offene Menge

$$U = (0, \infty) \times (0, 2\pi)$$

in \mathbb{R}^2 ist ein Diffeomorphismus mit

$$\det(D\varphi(r, \theta)) = r > 0$$

für alle $r \in (0, \infty)$. Wählen wir nun ein kompaktes, achsenparalleles Rechteck

$$B = [r_1, r_2] \times [\theta_1, \theta_2] \subseteq U$$

in U , so ist B kompakt und Jordan-meßbar. Das Bild von B unter φ ist dann ein Ausschnitt aus einem Kreisring, wie in Abbildung 24 dargestellt. Für eine auf $\varphi(B)$ stetige Funktion gilt dann

$$\begin{aligned} \int_{\varphi(B)} f(x, y) \, d(x, y) &= \int_B f(\varphi(r, \theta)) \cdot |D\varphi(r, \theta)| \, d(r, \theta) \\ &= \int_{\theta_1}^{\theta_2} \int_{r_1}^{r_2} f(r \cdot \cos(\theta), r \cdot \sin(\theta)) \cdot r \, dr \, d\theta. \end{aligned} \tag{213}$$

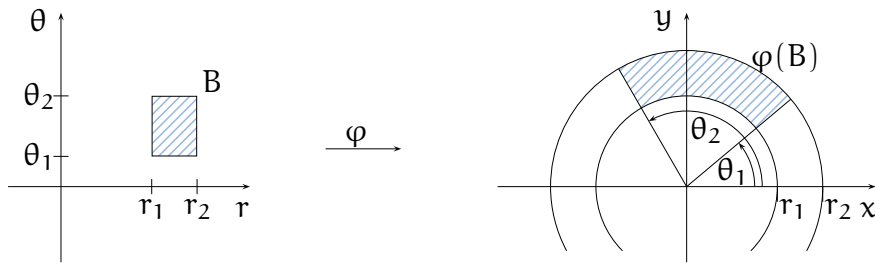


ABBILDUNG 24. Transformation eines Rechtecks auf einen Kreisringausschnitt

Insbesondere können wir damit den Inhalt des Kreisbogenschnitts berechnen:

$$V(\varphi(B)) = \int_{\theta_1}^{\theta_2} \int_{r_1}^{r_2} r \, dr \, d\theta = \frac{r_2^2 - r_1^2}{2} \cdot (\theta_2 - \theta_1).$$

Wegen des Transformationssatzes für Integrale auf Quadern 47.10 gilt die Transformationsformel (213) auch auf dem Quader

$$B = [0, r_0] \times [0, 2\pi],$$

und wir können somit den Inhalt eines Kreises mit Radius r_0

$$\varphi(B) = \{(x, y)^t \mid x^2 + y^2 \leq r_0^2\}$$

ausrechnen als

$$V(\varphi(B)) = \int_0^{r_0} \int_0^{2\pi} r \, d\theta \, dr = \pi \cdot r_0^2.$$

Beispiel 47.13 (Zylinderkoordinaten)

Nach Aufgabe 43.27 ist die Zylinderkoordinatenabbildung

$$\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (r, \theta, z) \mapsto (r \cdot \cos(\theta), r \cdot \sin(\theta), z)$$

auf der offenen Menge

$$U = (0, \infty) \times (0, 2\pi) \times \mathbb{R}$$

ein Diffeomorphismus mit

$$\det(D\varphi(r, \theta, z)) = \begin{vmatrix} \cos(\theta) & -r \cdot \sin(\theta) & 0 \\ \sin(\theta) & r \cdot \cos(\theta) & 0 \\ 0 & 0 & 1 \end{vmatrix} = r > 0.$$

Für jeden Quader

$$Q = [r_1, r_2] \times [\theta_1, \theta_2] \times [z_1, z_2] \subseteq U$$

und jede stetige Funktion f auf $\varphi(Q)$ gilt deshalb

$$\int_{\varphi(Q)} f(x, y, z) \, d(x, y, z) = \int_{r_1}^{r_2} \int_{\theta_1}^{\theta_2} \int_{z_1}^{z_2} f(r \cdot \cos(\theta), r \cdot \sin(\theta), z) \cdot r \, dz \, d\theta \, dr.$$

Aufgrund von Korollar 47.10 gilt die Transformationsformel auch noch falls $r_1 = 0$, $\theta_1 = 0$ und $\theta_2 = 2\pi$. Insbesondere erhält man für das Volumen des Zylinders

$$Z = \{(x, y, z)^t \in \mathbb{R}^3 \mid x^2 + y^2 \leq r_0^2, 0 \leq z \leq z_0\}$$

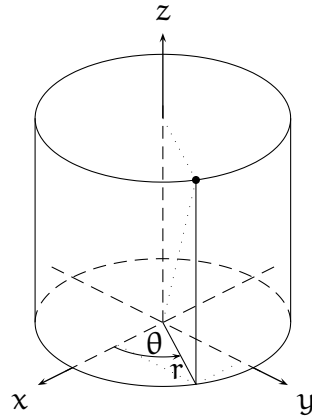


ABBILDUNG 25. Zylinderkoordinaten

dann

$$V(Z) = \int_0^{r_0} \int_0^{2\pi} \int_0^{z_0} r \, dz \, d\theta \, dr = z_0 \cdot r_0^2 \cdot \pi.$$

Beispiel 47.14 (Kugelkoordinaten)

Nach Aufgabe 43.26 ist die Kugelkoordinatenabbildung

$$\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (r, \theta, \vartheta) \mapsto (r \cdot \cos(\theta) \cdot \cos(\vartheta), r \cdot \sin(\theta) \cdot \cos(\vartheta), r \cdot \sin(\vartheta))$$

ist auf der offenen Menge

$$U = (0, \infty) \times (0, 2\pi) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$$

ein Diffeomorphismus mit

$$\det(D\varphi(r, \theta, z)) = r^2 \cdot \cos(\vartheta) > 0.$$

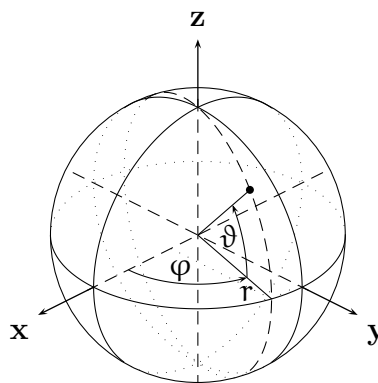


ABBILDUNG 26. Kugelkoordinaten

Für jeden Quader

$$Q = [r_1, r_2] \times [\theta_1, \theta_2] \times [\vartheta_1, \vartheta_2] \subseteq U$$

und jede stetige Funktion f auf $\varphi(Q)$ gilt deshalb

$$\int_{\varphi(Q)} f(x, y, z) \, d(x, y, z) = \int_{r_1}^{r_2} \int_{\theta_1}^{\theta_2} \int_{\vartheta_1}^{\vartheta_2} f(r \cos(\theta) \cos(\vartheta), r \sin(\theta) \cos(\vartheta), r \sin(\vartheta)) \cdot r^2 \cdot \cos(\vartheta) \, d\vartheta \, d\theta \, dr.$$

Wieder erlaubt uns Korollar 47.10 den Fall $r_1 = 0$, $\theta_1 = 0$, $\theta_2 = 2\pi$, $\vartheta_1 = -\frac{\pi}{2}$ und $\vartheta_2 = \frac{\pi}{2}$ zuzulassen. Insbesondere erhält man für das Volumen der Kugel

$$K = \{(x, y, z)^t \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq r_0^2\}$$

dann

$$\begin{aligned} V(K) &= \int_0^{r_0} \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} r^2 \cdot \cos(\vartheta) \, d\vartheta \, d\theta \, dr = \int_0^{r_0} \int_0^{2\pi} r^2 \cdot \sin(\vartheta) \Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \, d\theta \, dr \\ &= \int_0^{r_0} \int_0^{2\pi} 2r^2 \, d\theta \, dr = \int_0^{r_0} 4r^2 \pi \, dr = \frac{4 \cdot r_0^3 \cdot \pi}{3}. \end{aligned}$$

D) Der Beweis des Transformationsatzes 47.7

In diesem Abschnitt beschäftigen wir uns ausschließlich mit dem Beweis des Transformationsatzes für Integrale 47.7 sowie seiner Verallgemeinerung in Bemerkung 47.8. Wir zerlegen den Beweis in viele einzelne Schritte, die wir als Lemmata formulieren und beweisen. Dabei setzen wir stets die Bezeichnungen aus Satz 47.7 voraus.

Die *Grundidee* des Beweises besteht darin, sich auf den Spezialfall zurückzuziehen, daß B ein Quader ist und dort den Beweis mit Induktion und Fubini für Normalbereiche 46.37 auf die Substitutionsregel im Eindimensionalen 20.15 zurückzuführen.

Wir verwenden stets die Maximumsnorm, so daß Umgebungen Würfel sind.

Lemma 47.15

$\varphi(B)$ ist kompakt und Jordan-meßbar, und die Integrale in Satz 47.7 existieren, und zwar auch noch unter den abgeschwächten Voraussetzungen von Bemerkung 47.8.

Beweis: Wir beachten zunächst, daß der Abschluß \overline{N} der Jordan-Nullmenge N wegen Aufgabe 46.40 eine abgeschlossene Nullmenge ist. Aus $\det(D\varphi(x)) \neq 0$ für $x \in \overset{\circ}{B} \setminus \overline{N} \subseteq U \setminus N$ folgt dann mit dem Satz über die Umkehrfunktion 43.14, daß φ ein lokaler Diffeomorphismus auf $\overset{\circ}{B} \setminus \overline{N}$ ist. Die Voraussetzungen von Lemma 47.3 sind also erfüllt und $\varphi(B)$ ist somit Jordan-meßbar und kompakt.

Als stetige Funktion auf einer kompakten und Jordan-meßbaren Menge sind f auf $\varphi(B)$ und

$$g : B \longrightarrow \mathbb{R} : x \mapsto f(\varphi(x)) \cdot |\det(D\varphi(x))|$$

auf B dann integrierbar, so daß die beiden Integrale in Satz 47.7 existieren. \square

In Satz 47.7 bleibt also nur noch, die Transformationsformel (210) zu beweisen. Wir reduzieren das Problem nun darauf, die Transformationsformel für Quader zu zeigen.

Lemma 47.16

Gilt (210) für Quader, so gilt (210) für alle kompakten Jordan-meßbaren B .

Beweis: In der Aussage des Lemmas sind die offene Menge U und der Diffeomorphismus φ fest vorgegeben. Wir setzen nun voraus, daß die Transformationsformel (210) für jeden Quader Q in U gilt, und wir wollen zeigen, daß sie dann auch für die kompakte Jordan-meßbare Menge B gilt.

Die *Grundidee* des Beweises besteht darin, B zu zerlegen in eine Vereinigung Y von Würfeln, auf denen (210) dann nach Voraussetzung und der speziellen Additivität des Integrals gilt, sowie einen Restanteil $B \setminus Y$, auf dem die Differenz der beiden zugehörigen Integrale in (210) beliebig klein wird. Dies geht, da $B \setminus Y$ und $\varphi(B) \setminus \varphi(Y)$ jeweils durch eine Vereinigung von Würfeln sehr kleinen Volumens überdeckt werden, wobei für $\varphi(B) \setminus \varphi(Y)$ die Lipschitz-Stetigkeit von φ auf den Würfeln wesentlich ist.

1. Schritt: Wir zeigen zunächst, daß B eine Überdeckung durch endlich viele Würfel $\overline{U_{\varepsilon_i}(x_i)}$, $i = 1, \dots, k$, besitzt mit

$$U_{3\varepsilon_i}(x_i) \subseteq U. \quad (214)$$

Da U offen ist, besitzt jeder Punkt x von B eine offene Umgebung $U_{\delta_x}(x)$, die ganz in U enthalten ist. Dann überdecken die $U_{\frac{\delta_x}{3}}(x)$ die kompakte Menge B ebenfalls, und diese Überdeckung besitzt mithin eine endliche Teilüberdeckung, d.h. es reichen endlich viele Punkte x_1, \dots, x_k aus und wir können $\varepsilon_i = \frac{\delta_{x_i}}{3}$ wählen.

2. Schritt: Als nächstes überzeugen wir uns davon, daß wir ohne Einschränkung annehmen können, daß $D\varphi$ auf U beschränkt ist mit

$$L := \sup_{x \in U} \|D\varphi(x)\|_{\infty}.$$

Dazu beachten wir nur, daß alle $\overline{U_{2\varepsilon_i}(x_i)}$ ganz in U liegen und $D\varphi$ auf diesen kompakten Mengen stetig und somit beschränkt ist. Dann ist $D\varphi$ aber auf der Vereinigung der offenen Mengen $U_{2\varepsilon_i}(x_i)$ beschränkt und indem wir U durch diese ersetzen, können wir ohne Einschränkung annehmen, daß φ auf U beschränkt ist.

3. Schritt: Wir fixieren nun einen Würfel P , der B enthält, und setzen

$$l := \min\{\varepsilon_1, \dots, \varepsilon_k\}.$$

Ist Z eine äquidistante Zerlegung des Würfels P in *Würfel* mit $l(Z) \leq l$, so gilt für jeden Teilwürfel $W \in TQ(Z)$ mit $W \cap B \neq \emptyset$ schon

$$W \subseteq U.$$

Um dies zu sehen, beachten wir, daß der Würfel W die Form $W = \overline{U_{\frac{l(Z)}{2}}(z)}$ besitzt und daß für $x \in W \cap B$ ein i mit $x \in \overline{U_{\varepsilon_i}(x_i)}$ existiert. Somit gilt für $y \in W$

$$\|y - x_i\|_\infty \leq \|y - z\|_\infty + \|z - x\|_\infty + \|x - x_i\| \leq \frac{l(Z)}{2} + \frac{l(Z)}{2} + \varepsilon_i < 3 \cdot \varepsilon_i,$$

so daß mit (214) in der Tat $y \in U$ und somit $W \subseteq U$ folgt.

4. Schritt: Wir geben uns nun eine positive natürliche Zahl m vor.

Da $V(\partial B) = \int_{\partial B} 1 \, dx = 0$ gilt, gibt es nach Lemma 44.16 ein $\delta_m > 0$, so daß

$$OS(1_{\partial B}, Z) = OS(1_{\partial B}, Z) - \int_P 1_{\partial B}(x) \, dx = OS(1_{\partial B}, Z) - US(1_{\partial B}, Z) < \frac{1}{m} \quad (215)$$

für jede Zerlegung Z von P mit $l(Z) < \delta_m$. Wir wählen nun eine äquidistante Zerlegung Z^m von P in Würfel mit

$$l(Z^m) < \min\{\delta_m, l\}.$$

Dann setzen wir

$$X := \bigcup_{\substack{W \in \mathcal{TQ}(Z^m) \\ W \cap \partial B \neq \emptyset}} W$$

und

$$Y := \bigcup_{\substack{W \in \mathcal{TQ}(Z^m) \\ W \subseteq B}} W \subseteq B.$$

Da $l(Z)$ kleiner als l ist, sind die Würfel, die an X und Y beteiligt sind, zudem alle in U enthalten, und deshalb gilt

$$B \subseteq X \cup Y \subseteq U$$

und somit

$$B \setminus Y \subseteq X.$$

Wir können nun φ anwenden und erhalten

$$\varphi(B) \setminus \varphi(Y) \subseteq \varphi(B \setminus Y) \subseteq \varphi(X). \quad (216)$$

Aus (215) erhalten wir

$$V(X) = \sum_{\substack{W \in \mathcal{TQ}(Z^m) \\ W \cap \partial B \neq \emptyset}} V(W) = OS(1_{\partial B}, Z^m) < \frac{1}{m}, \quad (217)$$

und mit

$$M := \sup_{y \in \varphi(B)} |f(y)|$$

dann auch

$$\begin{aligned} \left| \int_{B \setminus Y} f(\varphi(x)) \cdot |\det(D\varphi(x))| \, dx \right| &\leq \int_{B \setminus Y} |f(\varphi(x))| \cdot |\det(D\varphi(x))| \, dx \\ &\leq \int_{B \setminus Y} M \cdot L \, dx = M \cdot L \cdot V(B \setminus Y) \quad (218) \\ &\leq M \cdot L \cdot V(X) \stackrel{(217)}{<} \frac{M \cdot L}{m}. \end{aligned}$$

Zudem ist die Abbildung φ nach Korollar 42.19 auf jedem Würfel $W \in \text{TQ}(Z^m)$ mit $W \cap \partial B \neq \emptyset$ Lipschitz-stetig und

$$L = \sup_{x \in U} \|D\varphi(x)\|_\infty$$

ist eine mögliche Lipschitz-Konstante für φ , die unabhängig von der gewählten Zerlegung ist. Für einen solchen Würfel $W = \overline{U_r(\mathbf{a})}$ gilt dann aber $\varphi(W) \subseteq \overline{U_{Lr}(\varphi(\mathbf{a}))}$ und somit (siehe auch Aufgabe 46.41)

$$V(\varphi(W)) \leq V(\overline{U_{Lr}(\varphi(\mathbf{a}))}) = (2Lr)^n = L^n \cdot V(W).$$

Da die Abschätzung für jeden an X beteiligten Würfel gilt, erhalten wir auch

$$V(\varphi(X)) \leq L^n \cdot V(X) \stackrel{(217)}{<} \frac{L^n}{m}. \quad (219)$$

Damit erhalten wir analog zu obigen Betrachtungen

$$\begin{aligned} \left| \int_{\varphi(B) \setminus \varphi(Y)} f(x) \, dx \right| &\leq \int_{\varphi(B) \setminus \varphi(Y)} |f(x)| \, dx \\ &\leq \int_{\varphi(B) \setminus \varphi(Y)} M \, dx = M \cdot V(\varphi(B) \setminus \varphi(Y)) \quad (220) \\ &\stackrel{(216)}{\leq} M \cdot V(\varphi(X)) \stackrel{(217)}{<} \frac{M \cdot L^n}{m}. \end{aligned}$$

Nach Voraussetzung ist Y eine Vereinigung von Würfeln, die ganz in B enthalten sind und deren Schnitt jeweils nur eine Jordan-Nullmenge ist. Da φ ein Diffeomorphismus auf einer Umgebung der Würfel ist, schneiden sich die Bilder der Würfel unter φ dann auch nur im Bild der Ränder und diese sind Jordan-Nullmengen (siehe Lemma 47.3). Wir können die spezielle Additivität für Integrale 46.29 also bei Y und bei $\varphi(Y)$ anwenden, und da wir zudem vorausgesetzt haben, daß die Transformationsformel auf Quadern gilt, erhalten wir folgende Gleichung:

$$\begin{aligned} \int_{\varphi(Y)} f(x) \, dx &= \sum_{\substack{W \in \text{TQ}(Z^m) \\ W \subseteq \overline{B}}} \int_{\varphi(W)} f(x) \, dx \\ &= \sum_{\substack{W \in \text{TQ}(Z^m) \\ W \subseteq \overline{B}}} \int_W f(\varphi(x)) \cdot |\det(D\varphi(x))| \, dx \quad (221) \\ &= \int_Y f(\varphi(x)) \cdot |\det(D\varphi(x))| \, dx. \end{aligned}$$

Der Übersichtlichkeit halber setzen wir nun $g(x) = f(\varphi(x)) \cdot |\det(D\varphi(x))|$ und tragen die Ergebnisse aus (218), (220) und (221) in Kurzform zusammen:

$$\begin{aligned} \left| \int_{\varphi(B)} f - \int_B g \right| &= \left| \int_{\varphi(Y)} f + \int_{\varphi(B) \setminus \varphi(Y)} f - \int_Y g - \int_{B \setminus Y} g \right| \\ &\stackrel{(221)}{=} \left| \int_{\varphi(B) \setminus \varphi(Y)} f - \int_{B \setminus Y} g \right| \leq \left| \int_{\varphi(B) \setminus \varphi(Y)} f \right| + \left| \int_{B \setminus Y} g \right| \\ &\stackrel{(218), (220)}{<} \frac{M \cdot L^n + M \cdot L}{m}. \end{aligned}$$

Dabei ist die linke Seite unabhängig von m und die rechte Seite konvergiert für $m \rightarrow \infty$ gegen null, so daß wir

$$\left| \int_{\varphi(B)} f(x) \, dx - \int_B f(\varphi(x)) \cdot |\det(D\varphi(x))| \, dx \right| = 0$$

erhalten. Damit ist (210) dann für B gezeigt. \square

Lemma 47.17

(210) gilt, wenn

$$\varphi(x) = (x_1, \dots, x_{n-1}, \varphi_n(x))^t$$

für eine notwendigerweise stetig differenzierbare Funktion $\varphi_n : U \rightarrow \mathbb{R}$ ist.

Beweis: Wegen Lemma 47.16 können wir ohne Einschränkung annehmen, daß B ein Quader ist. Die Funktionaldeterminante von φ ist dann

$$\det(D\varphi(x)) = \begin{vmatrix} 1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & \ddots & 0 & 0 \\ 0 & \dots & \dots & 0 & 1 & 0 \\ D_1\varphi_n(x) & D_2\varphi_n(x) & \dots & \dots & D_{n-1}\varphi_n(x) & D_n\varphi_n(x) \end{vmatrix} = D_n\varphi_n(x), \quad (222)$$

so daß nach Voraussetzung ohne Einschränkung

$$D_n\varphi_n(x) > 0 \quad (223)$$

für alle $x \in U$ gilt. Ist nun $B = [a, b]$ und $x_i \in [a_i, b_i]$ fest für $i = 1, \dots, n-1$, so hat die Funktion

$$[a_n, b_n] \rightarrow \mathbb{R} : t \mapsto \varphi_n(x_1, \dots, x_{n-1}, t)$$

eine stets positive Ableitung und ist mithin streng monoton wachsend. Dann ist aber

$$\varphi(B) = \{x \in \mathbb{R}^n \mid a_i \leq x_i \leq b_i \text{ für } i = 1, \dots, n-1, \\ \varphi_n(x_1, \dots, x_{n-1}, a_n) \leq x_n \leq \varphi_n(x_1, \dots, x_{n-1}, b_n)\}$$

ein Normalbereich, so daß uns der Satz von Fubini für Normalbereiche 46.37 und die Substitutionsregel 20.15 erlauben, das linke Integral in (210) auszurechnen:

$$\begin{aligned} \int_{\varphi(B)} f(x) \, dx &\stackrel{46.37}{=} \int_{a_1}^{b_1} \dots \int_{a_{n-1}}^{b_{n-1}} \int_{\varphi_n(x_1, \dots, x_{n-1}, a_n)}^{\varphi_n(x_1, \dots, x_{n-1}, b_n)} f(x_1, \dots, x_n) \, dx_n \, dx_{n-1} \dots dx_1 \\ &\stackrel{20.15}{=} \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} f(x_1, \dots, x_{n-1}, \varphi_n(x)) \cdot D_n \varphi_n(x) \, dx_n \, dx_{n-1} \dots dx_1 \\ &\stackrel{(222), (223)}{=} \int_B f(\varphi(x)) \cdot |\det(D\varphi(x))| \, dx \end{aligned}$$

Damit ist dann (210) in diesem Fall gezeigt. □

Wir wollen den Transformationssatz nun für eine weitere Klasse von Diffeomorphismen zeigen, die eine sehr spezielle Zerlegung besitzen. Wir werden dies aber nur unter der Annahme tun können, daß der Transformationssatz für Diffeomorphismen 47.7 in Dimension $n - 1$ bereits bewiesen ist.

Lemma 47.18

Der Transformationssatz 47.7 sei in \mathbb{R}^{n-1} schon bewiesen. Zudem sei U ein offener Quader, B ein Quader und $\varphi = \beta \circ \gamma$ für Diffeomorphismen γ auf U und β auf $\gamma(U)$ mit

$$\beta(x) = (x_1, \dots, x_{n-1}, \beta_n(x))^t$$

und

$$\gamma(x) = (\varphi_1(x), \dots, \varphi_{n-1}(x), x_n)^t,$$

dann gilt (210).

Beweis: Wir wollen den Beweis der Transformationsformel für φ nun auf die Transformationsformel im \mathbb{R}^{n-1} zurückführen. Dazu zerlegen wir den Quader $B = [a, b]$ in

$$B = [a_1, b_1] \times \dots \times [a_n, b_n] = Q \times [a_n, b_n]$$

mit

$$Q = [a_1, b_1] \times \dots \times [a_{n-1}, b_{n-1}] \subseteq \mathbb{R}^{n-1}.$$

Halten wir ein $x_n \in [a_n, b_n]$ fest, so ist die Menge

$$U_{x_n} = \{y \in \mathbb{R}^{n-1} \mid (y, x_n)^t \in U\}$$

ein offener Quader in \mathbb{R}^{n-1} . Ferner ist die Abbildung

$$\phi_{x_n} : U_{x_n} \longrightarrow \mathbb{R}^{n-1} : y \mapsto (\varphi_1(y, x_n), \dots, \varphi_{n-1}(y, x_n))^t$$

stetig differenzierbar mit

$$\det(D\phi_{x_n}(y)) = \det(D\gamma(y, x_n)) \neq 0 \tag{224}$$

für alle $y \in U_{x_n}$, da γ ein Diffeomorphismus ist. Zudem folgt aus der Injektivität von γ unmittelbar die Injektivität von ϕ_{x_n} , so daß ϕ_{x_n} ein Diffeomorphismus ist,

und da U_{x_n} als offener Quader konvex ist, ändert die Funktionaldeterminante von ϕ_{x_n} nach Lemma 47.9 auch ihr Vorzeichen nicht.

Beachtet man noch, daß

$$\gamma(B) = \bigcup_{x_n \in [a_n, b_n]} \phi_{x_n}(Q) \times \{x_n\} \subseteq \gamma(U) \quad (225)$$

gilt, so ist die Funktion

$$g : \gamma(B) \longrightarrow \mathbb{R} : (y, x_n) \mapsto f(\beta(y, x_n)) \cdot |\det(D\beta(y, x_n))|$$

nach Voraussetzung stetig und wir können den Transformationssatz 47.7 im \mathbb{R}^{n-1} für ein festes $x_n \in [a_n, b_n]$ anwenden:

$$\begin{aligned} \int_{\phi_{x_n}(Q)} g(y, x_n) \, dy &= \int_Q g(\phi_{x_n}(y), x_n) \cdot |\det(D\phi_{x_n}(y))| \, dy \\ &\stackrel{(224)}{=} \int_Q g(\gamma(y, x_n)) \cdot |\det(D\gamma(y, x_n))| \, dy. \end{aligned} \quad (226)$$

Nun beachten wir noch die Gleichung

$$\det(D\varphi(x)) = \det(D\beta(\gamma(x))) \cdot \det(D\gamma(x)). \quad (227)$$

und mithin

$$\begin{aligned} g(\gamma(x)) \cdot |\det(D\gamma(x))| &= f(\beta(\gamma(x))) \cdot |\det(D\beta(\gamma(x)))| \cdot |\det(D\gamma(x))| \\ &\stackrel{(227)}{=} f(\varphi(x)) \cdot |\det(D\varphi(x))|. \end{aligned} \quad (228)$$

Damit sind wir nun in der Lage (210) in dieser Situation zu beweisen, wobei wir beachten, daß wir den Satz von Fubini in der Fassung von Aufgabe 46.42 auf $\gamma(B)$ anwenden können:

$$\begin{aligned} \int_{\varphi(B)} f(x) \, dx &= \int_{\beta(\gamma(B))} f(x) \, dx \stackrel{47.17}{=} \int_{\gamma(B)} f(\beta(x)) \cdot |\det(D\beta(x))| \, dx \\ &= \int_{\gamma(B)} g(x) \, dx = \int_{\gamma(B)} g(y, x_n) \, d(y, x_n) \\ &\stackrel{46.42}{=} \int_{a_n}^{b_n} \int_{\phi_{x_n}(Q)} g(y, x_n) \, dy \, dx_n \\ &\stackrel{(226)}{=} \int_{a_n}^{b_n} \int_Q g(\gamma(y, x_n)) \cdot |\det(D\gamma(y, x_n))| \, dy \, dx_n \\ &\stackrel{(228)}{=} \int_{a_n}^{b_n} \int_Q f(\varphi(y, x_n)) \cdot |\det(D\varphi(y, x_n))| \, dy \, dx_n \\ &= \int_B f(\varphi(x)) \cdot |\det(D\varphi(x))| \, dx. \end{aligned}$$

□

Wir wollen nun ein einfaches Beispiel für Diffeomorphismen mit einer globalen Zerlegung wie in Lemma 47.18 geben.

Beispiel 47.19

Wir betrachten eine invertierbare Matrix

$$A = (a_{ij})_{i,j=1,\dots,n} = \left(\begin{array}{c|c} D & \mathbf{b} \\ \hline \mathbf{c}^t & a_{nn} \end{array} \right) \in \text{Gl}_n(\mathbb{R})$$

mit $D = (a_{ij})_{i,j=1,\dots,n-1}$, $\mathbf{b} = (a_{1n}, \dots, a_{n-1n})^t$ und $\mathbf{c} = (a_{n1}, \dots, a_{nn-1})^t$. Außerdem setzen wir noch voraus, daß auch $D \in \text{Gl}_{n-1}(\mathbb{R})$ invertierbar ist.

Dann setzen wir

$$B = \left(\begin{array}{c|c} \mathbf{1}_{n-1} & 0 \\ \hline \mathbf{c}^t \circ D^{-1} & a_{nn} - \mathbf{c}^t \circ D^{-1} \circ \mathbf{b} \end{array} \right)$$

und

$$C = \left(\begin{array}{c|c} D & \mathbf{b} \\ \hline 0 & 1 \end{array} \right).$$

Die Abbildungen $\varphi = f_A$, $\beta = f_B$ und $\gamma = f_C$ erfüllen dann die Voraussetzungen von Lemma 47.18. Insbesondere gilt

$$\varphi = f_A = f_B \circ f_C = \beta \circ \gamma,$$

da man leicht $A = B \circ C$ nachrechnet.

Lemma 47.20 (Zerlegungssatz)

Für jeden Punkt in $x \in U$ gibt es (nach eventuellem Umsortieren der Koordinaten) eine offene Umgebung V sowie Diffeomorphismen

$$\gamma : V \longrightarrow \mathbb{R}^n : x \mapsto (\varphi_1(x), \dots, \varphi_{n-1}(x), x_n)^t$$

und

$$\beta : \gamma(V) \longrightarrow \mathbb{R}^n : x \mapsto (x_1, \dots, x_{n-1}, \beta_n(x))^t$$

mit

$$\varphi|_V = \beta \circ \gamma.$$

Beweis: Da φ ein Diffeomorphismus ist, ist $\det(D\varphi(x)) \neq 0$. Dann muß aber auch einer der $(n-1) \times (n-1)$ -Minoren ungleich null sein, und wir können nach eventuellem Ummumerieren der Koordinaten annehmen, daß es der $n-1$ -te Hauptminor ist, d.h.

$$0 \neq \begin{vmatrix} D_1\varphi_1(x) & \dots & D_{n-1}\varphi_1(x) \\ \vdots & & \vdots \\ D_1\varphi_{n-1}(x) & \dots & D_{n-1}\varphi_{n-1}(x) \end{vmatrix} = \det(D\gamma(x)).$$

Nach dem Satz über die Umkehrfunktion 43.14 gibt es dann eine offene Umgebung V von x , so daß

$$\gamma : V \longrightarrow \mathbb{R}^n$$

ein Diffeomorphismus ist, und insbesondere ist $\gamma(V)$ offen. Die Umkehrabbildung γ^{-1} von γ habe nun die Komponentenfunktionen $\alpha_1, \dots, \alpha_n$, d.h.

$$\gamma^{-1} : \gamma(V) \longrightarrow \mathbb{R}^n : y \mapsto (\alpha_1(y), \dots, \alpha_n(y))^t.$$

Dann definieren wir

$$\beta_n : \gamma(V) \longrightarrow \mathbb{R} : \mathbf{y} \mapsto \varphi_n(\alpha_1(\mathbf{y}), \dots, \alpha_{n-1}(\mathbf{y}), \mathbf{y}_n)^t$$

und erhalten damit

$$\begin{aligned} \beta(\gamma(\mathbf{z})) &= (\varphi_1(\mathbf{z}), \dots, \varphi_{n-1}(\mathbf{z}), \varphi_n(\alpha_1(\gamma(\mathbf{z})), \dots, \alpha_{n-1}(\gamma(\mathbf{z})), z_n))^t \\ &= (\varphi_1(\mathbf{z}), \dots, \varphi_{n-1}(\mathbf{z}), \varphi_n(z_1, \dots, z_{n-1}, z_n))^t = \varphi(\mathbf{z}). \end{aligned}$$

Zudem ist $\beta = \varphi \circ \gamma^{-1}$ als Komposition zweier Diffeomorphismen offenbar ein Diffeomorphismus. \square

Lemma 47.21

(210) gilt für ein beliebiges B und φ .

Beweis: Wegen Lemma 47.16 können wir uns auf den Fall zurück ziehen, daß B ein Quader ist.

Wir führen den Beweis nun durch Induktion nach n , wobei der Induktionsanfang $n = 1$ aus der Substitutionsformel 20.15 folgt (siehe auch Bemerkung 47.6). Für den Induktionsschritt können wir also annehmen, daß (210) in \mathbb{R}^{n-1} bereits gilt.

Aus dem Zerlegungssatz 47.20 wissen wir, daß jeder Punkt $\mathbf{x} \in B$ eine Umgebung $V_{\mathbf{x}}$ besitzt, auf der φ sich (bis auf die Reihenfolge der Variablen) wie in Lemma 47.18 zerlegen läßt, und durch Verkleinern können wir ohne Einschränkung annehmen, daß $V_{\mathbf{x}}$ ein offener Quader ist. Die $V_{\mathbf{x}}$ überdecken B und da B kompakt ist, reichen dazu schon endlich viele $V_{\mathbf{x}_i}$, $i = 1, \dots, k$, aus. Wir können nun nach Lemma 45.12 eine Zerlegung Z des Quaders B wählen, die so fein ist, daß jeder Teilquader von Z ganz in einem der $V_{\mathbf{x}_i}$ enthalten ist. Dann können wir Lemma 47.18 auf jeden dieser Teilquader anwenden, wobei die Rolle von U durch den zugehörigen offenen Quader $V_{\mathbf{x}_i}$ übernommen wird, und wir erhalten, daß (210) auf jedem Teilquader gilt. Dann folgt aber (210) auf B mittels der speziellen Additivität des Integrals 46.29:

$$\begin{aligned} \int_{\varphi(B)} f(\mathbf{x}) \, d\mathbf{x} &= \sum_{Q \in \text{TQ}(Z)} \int_{\varphi(Q)} f(\mathbf{x}) \, d\mathbf{x} \\ &= \sum_{Q \in \text{TQ}(Z)} \int_Q f(\varphi(\mathbf{x})) \cdot |\det(D\varphi(\mathbf{x}))| \, d\mathbf{x} \\ &= \int_B f(\varphi(\mathbf{x})) \cdot |\det(D\varphi(\mathbf{x}))| \, d\mathbf{x}. \end{aligned}$$

\square

Damit ist der Transformationssatz für Integrale 47.7 gezeigt. Es bleibt noch, die allgemeine Fassung des Satzes mit den abgeschwächten Voraussetzungen in Bemerkung 47.8 zu zeigen. Wegen Lemma 47.15 reicht es dazu, die Transformationsformel (210) für diesen Fall herzuleiten.

Lemma 47.22

(210) *gilt auch unter den abgeschwächten Voraussetzungen von Bemerkung 47.8.*

Beweis: Wir geben uns eine positive natürliche Zahl m vor. Die Jordan-Nullmenge \bar{N} (siehe Aufgabe 46.40) läßt sich nach Bemerkung 46.18 durch das Innere endlich vieler Quader Q_1, \dots, Q_k in U überdecken mit

$$\sum_{i=1}^k V(Q_i) < \frac{1}{m}.$$

Wir setzen nun

$$X := \bigcup_{i=1}^k Q_i \subseteq U$$

und

$$Y := B \setminus X.$$

Dann ist Y abgeschlossen und beschränkt und nach dem Satz von Heine-Borel 39.44 deshalb kompakt. Zudem ist φ auf der offenen Umgebung $U \setminus \bar{N}$ von Y ein Diffeomorphismus, so daß die Transformationsformel nach Satz 47.7 auf Y gilt.

Zudem ist $V(X) < \frac{1}{m}$ und wie im Beweis des 4. Schritts von Lemma 47.16 zeigt man, daß

$$V(\varphi(X)) \leq L^n \cdot V(X) < \frac{L^n}{m}$$

gilt für eine geeignete Konstante L , die unabhängig von m ist.

Dann kann man wie im Beweis des 4. Schrittes von Lemma 47.16 fortfahren und erhält schließlich die Transformationsformel (210) auch in diesem Fall. \square

Literaturverzeichnis

- [BF87] Martin Barner and Friedrich Flohr, *Analysis I*, 3. Auflage ed., Walter de Gruyter, 1987.
- [Coh96] Henri Cohen, *A course in computational algebraic number theory*, 3 ed., Graduate Texts in Mathematics, no. 138, Springer, 1996.
- [Dec10] Wolfram Decker, *Grundlagen der Mathematik I*, Vorlesungsskript Wintersemester 2009/10, 2010.
- [Ebb92] Heinz-Dieter Ebbinghaus (ed.), *Zahlen*, 3 ed., Springer, 1992.
- [Gat08] Andreas Gathmann, *Grundlagen der Mathematik*, Vorlesungsskript WS2007/08, 2008.
- [GK00] Gert-Martin Greuel and Thomas Keilen, *Lineare Algebra I & II*, Vorlesungsskript, FB Mathematik, Universität Kaiserslautern, 2000.
- [Heu03] Harro Heuser, *Lehrbuch der Analysis, Teil 1*, 15 ed., Teubner, 2003.
- [Heu04] _____, *Lehrbuch der Analysis, Teil 2*, 13 ed., Teubner, 2004.
- [Mar08] Thomas Markwig, *Algebraische Strukturen*, Vorlesungsskript, 2008.
- [Moo82] Gregory H. Moore, *Zermelo's axiom of choice: Its origins, development and influence*, Studies in the History of Mathematics and Physical Sciences, no. 8, Springer, 1982.
- [Mül07] Rainer Müller, *Aufgaben zur vollständigen induktion*, <http://www.emath.de>, 2007.
- [Sze50] Tibor Szele, *On Zorn's lemma*, *Publicationes Mathematicae Debrecen* **1** (1949/50), 254–57.