Analysis

Vorlesungsskript

Thomas Markwig
Fachbereich Mathematik
Universität Tübingen

Sommersemester 2025

Inhaltsverzeichnis

Einleitung 1

A Analysis 1 3

Kapitel I Grundlegende Begriffsbildungen 5

§ 1 Etwas Logik 5

§ 2 Mengen 18

§ 3 Abbildungen 22

§ 4 Vollständige Induktion 31

§ 5 Mächtigkeit von Mengen 34

§ 6 Äquivalenzrelationen 40

§ 7 Gruppen und Körper 48

§ 8 Ordnungsrelationen 60

§ 9 Eigenschaften der reellen Zahlen $\mathbb{R}$ 68

§ 10 Der Körper der komplexen Zahlen 76

Kapitel II Eindimensionale Analysis 89

§ 11 Folgen und ihre Grenzwerte 89

§ 12 Unendliche Reihen 109

§ 13 Grenzwerte von Funktionen 144

§ 14 Stetigkeit 157

§ 15 Konvergenz von Funktionenfolgen 173

§ 16 Exponentialfunktion, Logarithmus, trigonometrische Funktionen 178

§ 17 Differenzierbarkeit 196

§ 18 Der Mittelwertsatz und seine Anwendungen 209

§ 19 Das Riemann-Integral 239
§ 20 Hauptsatz der Differential- und Integralrechnung mit Anwendungen 258
§ 21 Uneigentliche Integrale 275

B Analysis 2 281

Kapitel III Mehrdimensionale Analysis 283
§ 22 Topologische Grundbegriffe in metrischen und in normierten Räumen 283
§ 23 Stetigkeit in metrischen und in normierten Räumen 314
§ 24 Differenzierbarkeit im $\mathbb{R}^n$ 349
§ 25 Der Satz von Taylor und seine Anwendungen 369
§ 26 Der Satz über implizite Funktionen 394
§ 27 Das Riemann-Integral über n-dimensionalen Quadern 424
§ 28 Das Lebesguesche Integrabilitätskriterium 450
§ 29 Riemann-Integral über Jordan-meßbaren Mengen 464
§ 30 Der Transformationssatz für Integrale 487

Kapitel IV Gewöhnliche Differentialgleichungen 511
§ 31 Gewöhnliche Differentialgleichungen 511
§ 32 Lineare gewöhnliche Differentialgleichungen 530

C Anhang 549

Anhang A Lineare Algebra 551
§ A1 Grundlegendes zu Gruppen und Körpern 551
§ A2 Rechnen mit Matrizen und der Gauß-Algorithmus 577
§ A3 Vektorräume und lineare Abbildungen 592
§ A4 Basen von Vektorräumen 619
§ A5 Endlich-dimensionale Vektorräume 637
§ A6 Lineare Abbildungen und Matrizen 650
§ A7 Lineare Gleichungssysteme 673
§ A8 Die symmetrische Gruppe 695
§ A9 Die Determinante 702

§ A10 Der Polynomring $K[t]$ 728
§ A11 Endomorphismen und ihre Eigenwerte 741
§ A12 Diagonalisierbarkeit und Trigonalisierbarkeit 758
§ A13 Die Jordansche Normalform 779
§ A14 Euklidische und unitäre Räume 808
§ A15 Spektralsatz und Hauptachsentransformation 829
§ A16 Klassifikation der Kegelschnitte in der euklidischen Ebene 854
§ A17 Singulärwertzerlegung 861
§ A18 Lineare Algebra mit SINGULAR 870

Anhang B Die Integralsätze von Green, Gauß und Stokes 883
§ B1 Kurvenintegrale im $\mathbb{R}^n$ und der Integralsatz von Green 883
§ B2 Die Integralsätze von Gauß und Stokes 905

Literaturverzeichnis 919

Einleitung

Die vorliegende Ausarbeitung zu den Vorlesungen Analysis 1 und 2 im Sommersemester 2017 und Wintersemester 2017/18 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.

Teil A

Analysis 1

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: $A, B, C, \ldots$.

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 **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:

(1)

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

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 Wahrheitstafeln 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 sinnvoll ist. 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

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

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

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

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

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

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 nicht viel.

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 Wahrheitstafeln 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.

A. Jede gerade Zahl ist Summe zweier ungerader Zahlen.

B. Es gibt unendlich viele Primzahlen.

C. Jede gerade Zahl größer zwei ist Summe zweier Primzahlen.

D. Zu jedem Kreis läßt sich, nur mit Zirkel und Lineal, ein Quadrat konstruieren, das den gleichen Flächeninhalt hat.

E. 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.

Aussage A ist offensichtlich wahr, und auch Aussage B ist richtig, allerdings ist dies keine triviale Aussage. Sie muß bewiesen werden. 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 jahrhundertelang 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. □

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

Bemerkung 1.3

a. Eine *Aussage* ist eine Äußerung, der eindeutig ein Wahrheitswert wahr (**w**) oder falsch (**f**) zugeordnet ist.

b. Aus Aussagen X und Y können wir durch Anwenden *logischer Operatoren* neue Aussagen bilden:

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

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.

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

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".

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

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

Eine Aussageform, die für jeden möglichen Wert der Variablen eine wahre Aussage liefert, nennt man auch eine *Tautologie*. Die in Proposition 1.5 gezeigten Äquivalenzen aufgefaßt als Aussageformen sind in diesem Sinne Tautologien.

Beispiel 1.7 (Nand-Gatter)

Die Digitalelektronik basiert auf den Prinzipien der Aussagenlogik. Die Bauelemente elektronischer Schaltungen realisieren im Grunde Aussageformen, die je nach Wert der boolschen Variablen zwei Zustände annehmen können. Dabei kann man zeigen, daß alle logischen Aussagen, die wir aus den in diesem Abschnitt eingeführten Operationen erstellen können, sich auch mittels einer einzigen Operation gewinnen lassen, dem sogenannten Nand $\uparrow$, die durch folgende Wahrheitstafel definiert wird:

X	Y	$X \uparrow Y$
w	**w**	**f**
w	**f**	**w**
f	**w**	**w**
f	**f**	**w**

Wir überlassen es dem Leser, mittels Wahrheitstafeln die folgenden Äquivalenzen zu zeigen:

$$\begin{aligned}\neg(X \wedge Y) &\iff (X \uparrow Y)\\ \neg X &\iff (X \uparrow X)\\ (X \vee Y) &\iff ((X \uparrow X) \uparrow (Y \uparrow Y))\\ (X \wedge Y) &\iff ((X \uparrow Y) \uparrow (X \uparrow Y))\end{aligned}$$

Da sich die anderen Operatoren alle auf die Negation, die Disjunktion und die Konjunktion zurück führen lassen, ist damit auch gezeigt, daß sich alle auch allein durch das Nand ausdrücken lassen.

Bemerkung 1.8 (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:

$A\ \alpha$ Alpha	$B\ \beta$ Beta	$\Gamma\ \gamma$ Gamma	$\Delta\ \delta$ Delta	$E\ \epsilon\ \varepsilon$ Epsilon	$Z\ \zeta$ Zeta	$H\ \eta$ Eta	$\Theta\ \theta\ \vartheta$ Theta
$I\ \iota$ Iota	$K\ \kappa$ Kappa	$\Lambda\ \lambda$ Lambda	$M\ \mu$ My	$N\ \nu$ Ny	$\Xi\ \xi$ Xi	$O\ o$ Omikron	$\Pi\ \pi$ Pi
$P\ \rho$ Rho	$\Sigma\ \sigma$ Sigma	$T\ \tau$ Tau	$Y\ \upsilon$ Ypsilon	$\Phi\ \phi\ \varphi$ Phi	$X\ \chi$ Chi	$\Psi\ \psi$ Psi	$\Omega\ \omega$ Omega

Aufgaben

Aufgabe 1.9
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. $(\neg X \Longrightarrow f) \iff X$.

Aufgabe 1.10 (Nand)
Es seien X und Y zwei Aussagen.

Aufgabe 1.11

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 1.12

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 1.13

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 1.14

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 1.15 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.

Aufgabe 1.16

Seien X, Y und Z Aussagen. Beweise die folgende Aussage:

$$(X \vee Y) \iff ((X \uparrow X) \uparrow (Y \uparrow Y)).$$

§ 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*:

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

b. *Mengen angeben durch Vorschreiben einer Eigenschaft*:

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) $M \subseteq N \;:\Longleftrightarrow\; (x \in M \;\Rightarrow x \in N)$ "M ist *Teilmenge* von N"

2) $M = N \;:\Longleftrightarrow\; (M \subseteq N \;\wedge\; N \subseteq M)$
$\Longleftrightarrow\; (x \in M \;\Leftrightarrow\; x \in N)$

3) $M \neq N \;:\Longleftrightarrow\; \neg(M = N)$
$\Longleftrightarrow\; ((\exists\, x \in M \;:\; x \notin N) \;\vee\; (\exists\, x \in N \;:\; x \notin M))$

4) $M \subsetneqq N \;:\Longleftrightarrow\; (M \subseteq N \;\wedge\; M \neq N)$ "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. $\{1, 3\} \subsetneqq \{1, 2, 3\}$.

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:

- $\mathbb{N} = \{0, 1, 2, 3, 4, \ldots\}$ die Menge der *natürlichen Zahlen*,
- $\mathbb{Z} = \{\ldots, -3, -2, -1, 0, 1, 2, 3, \ldots\}$ 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. $P = M \dot\cup N \ :\iff \ (P = M \cup N \ \wedge \ M \cap N = \emptyset)$.
Wir sagen dann, P ist die *disjunkte Vereinigung* von M und N.

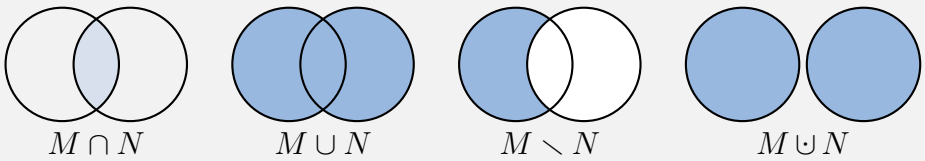


Abbildung 1: Durchschnitt, Vereinigung, Differenzmenge, disjunkte Vereinigung

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. $P = \dot\bigcup_{i \in I} M_i \ :\iff \ (P = \bigcup_{i \in I} M_i \ \wedge \ M_i \cap M_j = \emptyset \ \forall\, i, j \in I \text{ mit } i \neq j)$.
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:

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

und

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

c. Es gilt:

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

und

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

□

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:

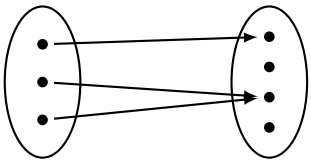
$$f : M \longrightarrow N : x \mapsto f(x).$$

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

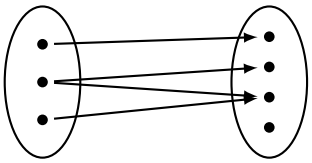
$$f = g \iff \big(M = X \ \wedge \ N = Y \ \wedge \ \forall\, x \in M \ : \ f(x) = g(x)\big).$$

Beispiel 3.2

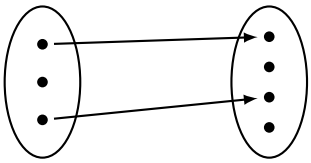
a. Die folgenden Bilder sollen den Begriff der Abbildung graphisch veranschaulichen:



Abbildung



keine Abbildung



keine Abbildung

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

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

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. $\mathrm{Graph}(f) := \{(x, f(x)) \mid x \in M\} \subseteq M \times N$ heißt der *Graph* von f.

b. $f(A) := \{f(x) \mid x \in A\} \subseteq N$ heißt das *Bild* von A unter f.

c. $\mathrm{Im}(f) := f(M) \subseteq N$ 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 $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$.

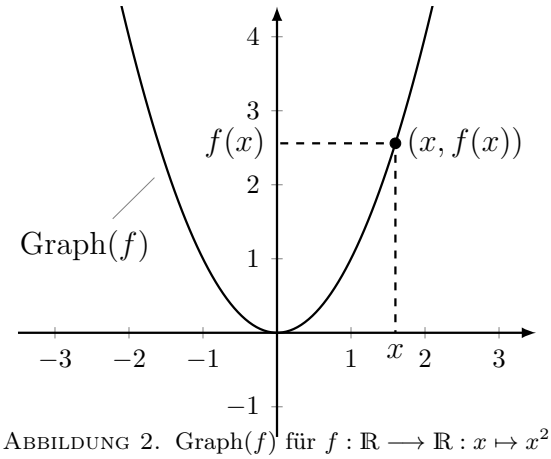


Abbildung 2. Graph(f) für $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$

- 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 $f^{-1}(B') = \emptyset$.
- $\mathrm{Im}(f) = \{x \in \mathbb{R} \mid x \geq 0\}$.

b. Die Abbildung $\mathrm{nf} : \mathbb{N} \longrightarrow \mathbb{N} : x \mapsto x + 1$ nennen wir die *Nachfolgerfunktion*. Es gelten

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

und

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

Bemerkung 3.5 (Abbildungen und ihre Graphen)

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

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

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 $f : M \longrightarrow N$ mit $\Gamma = \mathrm{Graph}(f)$.

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 $f : \emptyset \longrightarrow M$ gibt, und daß es für eine nicht-leere Menge M keine Abbildung $f : M \longrightarrow \emptyset$ geben kann.

Definition 3.6 (Injektiv, surjektiv, bijektiv)
Es sei $f : M \longrightarrow N$ 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

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

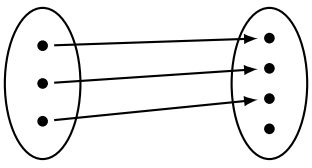
d.h. wenn $\mathrm{Im}(f) = N$.

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

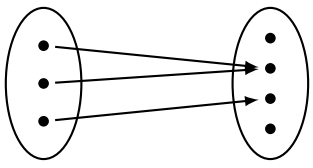
Bemerkung 3.7 (Injektiv, surjektiv, bijektiv)
Es sei $f : M \longrightarrow N$ eine Abbildung.

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

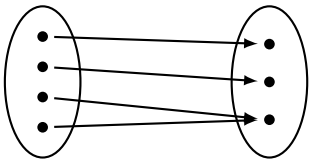
b. Es gelten:
- f ist injektiv $\iff$ jedes $y \in N$ hat *höchstens* ein Urbild.
- f ist surjektiv $\iff$ jedes $y \in N$ hat *mindestens* ein Urbild.
- f ist bijektiv $\iff$ jedes $y \in N$ hat *genau* ein Urbild.



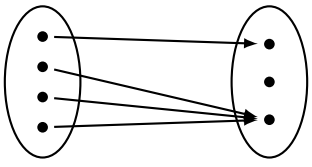
injektiv



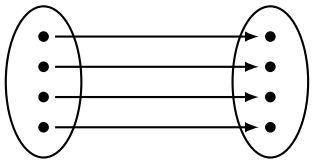
nicht injektiv



surjektiv



nicht surjektiv



bijektiv

Beispiel 3.8

a. Die Nachfolgerfunktion $\mathrm{nf} : \mathbb{N} \longrightarrow \mathbb{N} : x \mapsto x+1$ ist injektiv, aber nicht surjektiv. Denn, $x+1 = \mathrm{nf}(x) = \mathrm{nf}(y) = y+1$ für $x, y \in \mathbb{N}$ impliziert $x = y$, und $0 \notin \mathrm{Im}(f)$.

b. $g : \mathbb{Z} \longrightarrow \mathbb{N} : x \mapsto x^2$ 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 $\mathrm{id}_M(y) = y$, so daß id_M surjektiv ist, und für $x, x' \in M$ mit $\mathrm{id}_M(x) = \mathrm{id}_M(x')$ gilt $x = x'$, so daß id_M injektiv ist.

d. Ist $f : M \longrightarrow N$ injektiv, so ist die Abbildung $M \longrightarrow \mathrm{Im}(f) : x \mapsto f(x)$ offenbar bijektiv.

Definition 3.9 (Komposition von Abbildungen)
Seien $f : M \longrightarrow N$ und $g : N \longrightarrow P$ zwei Abbildungen. Die Abbildung

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

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

Beispiel 3.10

Seien $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$ und $g : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x+1$. 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 $f : M \longrightarrow N$, $g : N \longrightarrow P$ und $h : P \longrightarrow Q$ 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

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

Dies zeigt die Behauptung. □

Satz 3.12 (Bijektivität = Existenz einer Umkehrabbildung)
Es sei $f : M \longrightarrow N$ eine Abbildung.

a. f ist genau dann bijektiv, wenn eine Abbildung $g : N \longrightarrow M$ existiert, so daß $g \circ f = \mathrm{id}_M$ und $f \circ g = \mathrm{id}_N$.

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

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

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

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

Also ist f injektiv.
"$\Longrightarrow$": Da f bijektiv ist, gibt es für jedes $y \in N$ genau ein Urbild $x_y \in M$ von y unter f, d.h. $f(x_y) = y$. Wir definieren nun eine Abbildung

$$g : N \longrightarrow M : y \mapsto x_y.$$

Dann gilt zunächst für $y \in N$

$$(f \circ g)(y) = f\big(g(y)\big) = f(x_y) = y = \mathrm{id}_N(y).$$

Also ist $f \circ g = \mathrm{id}_N$.
Zudem gilt für $x \in M$ und $y := f(x) \in N$

$$f(x_y) = y = f(x).$$

Da f injektiv ist, folgt daraus $x = x_y$, und wir erhalten

$$(g \circ f)(x) = g\big(f(x)\big) = g(y) = x_y = x = \mathrm{id}_M(x).$$

Damit ist auch $g \circ f = \mathrm{id}_M$ gezeigt.

b. Sei $h : N \longrightarrow M$ eine zweite Abbildung mit $h \circ f = \mathrm{id}_M$ und $f \circ h = \mathrm{id}_N$. Dann gilt für $y \in N$

$$f\big(g(y)\big) = (f \circ g)(y) = \mathrm{id}_N(y) = (f \circ h)(y) = f\big(h(y)\big).$$

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

□

Beispiel 3.13

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

Denn für $y \in \mathbb{R}$ gilt

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

und für $x \in \mathbb{R}$ gilt

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

Die Behauptung folgt also aus Satz 3.12.

Proposition 3.14 (Injektivität, Surjektivität, Bijektivität unter Komposition)

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

a. Sind f und g injektiv, so ist $g \circ f$ injektiv.

b. Sind f und g surjektiv, so ist $g \circ f$ surjektiv.

c. Sind f und g bijektiv, so ist $g \circ f$ 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.

□

Aufgaben

Aufgabe 3.15

Überprüfe die folgende Abbildung auf Injektivität und Surjektivität:

$$f : \mathbb{N} \longrightarrow \mathbb{N} : n \mapsto 2^n.$$

Aufgabe 3.16
Ist $f : M \longrightarrow N$ eine surjektive Abbildung und $y \in N$, so ist

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

eine surjektive Abbildung.

Aufgabe 3.17
Es sei $f : M \longrightarrow N$ eine injektive Abbildung, $x' \in M$ und $y' = f(x') \in N$.

a. Dann ist $g : M \setminus \{x'\} \longrightarrow N \setminus \{y'\} : x \mapsto f(x)$ eine injektive Abbildung.

b. Ist f bijektiv, so ist g auch bijektiv.

Aufgabe 3.18
Für eine Abbildung $f : M \longrightarrow N$, $M \neq \emptyset$, beweise man die folgenden Aussagen:

a. f ist injektiv $\iff \exists\, g : N \longrightarrow M$, so dass $g \circ f = \mathrm{id}_M$.

b. f ist surjektiv $\iff \exists\, g : N \longrightarrow M$, so dass $f \circ g = \mathrm{id}_N$.

Aufgabe 3.19
Untersuche ob die folgenden Abbildungen injektiv, surjektiv oder bijektiv sind:

a. $f_1 : \mathbb{R} \longrightarrow \mathbb{R} : \; x \longmapsto 3x + 2$

b. $f_2 : \mathbb{Z} \longrightarrow \mathbb{Z} : \; x \longmapsto 3x + 2$

c. $f_3 : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \times \mathbb{R} : \; (x, y) \longmapsto (xy, x + 1)$

d. $f_4 : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \times \mathbb{R} : \; (x, y) \longmapsto (x - 2y, 2x + y)$

Aufgabe 3.20
Bestimme die Menge $(h \circ g \circ f)^{-1}(B)$ für $B = \{x \in \mathbb{R} \mid x \geq 3\}$ und

$$f : \mathbb{R} \longrightarrow \mathbb{R} \times \mathbb{R} : x \mapsto (x, x - 2),$$

$$g : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{Z} \times \mathbb{R} : (x, y) \mapsto \begin{cases} (1, x \cdot y), & \text{wenn } x > y \\ (-1, x \cdot y), & \text{wenn } x \leq y, \end{cases}$$

$$h : \mathbb{Z} \times \mathbb{R} \longrightarrow \mathbb{R} : (z, x) \mapsto x^z.$$

Aufgabe 3.21
Seien M, N zwei nicht-leere Mengen und $f : M \longrightarrow N$ 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 $g : N \longrightarrow X$ und $h : N \longrightarrow X$ aus $g \circ f = h \circ f$ schon $g = h$ folgt.

Aufgabe 3.22

Seien L, M, N Mengen und $f : L \longrightarrow M$, $g : M \longrightarrow N$ 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.

c. Ist $g \circ f$ surjektiv, so ist g surjektiv.

d. Ist $g \circ f$ surjektiv, so ist f surjektiv.

Aufgabe 3.23

Seien M, N Mengen, $A_1, A_2 \subseteq M$ und $B, B_1, B_2 \subseteq N$ Teilmengen und $f : M \longrightarrow N$ 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.

Aufgabe 3.24

Gib für die folgenden Abbildungsvorschriften die maximale Teilmenge von $\mathbb{R}$ an, die als Definitionsbereich in Frage kommt, und berechne jeweils auch das Bild der Abbildung:

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

b. $g(x) = x^2 - 4$,

c. $h(x) = \sqrt{(g \circ f)(x)}$,

d. $k(x) = \frac{1}{g(x)}$.

Aufgabe 3.25

Überprüfe die folgenden Abbildungen auf Injektivität, Surjektivität und Bijektivität:

a. $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x^2+1}$,

b. $g : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} \setminus \{0\} : x \mapsto \frac{1}{x}$,

c. $h : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} : (x, y) \mapsto 2x + y$,

d. $k : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \times \mathbb{R} : (x, y) \mapsto (x + y, x - y)$.

Zeige oder widerlege jeweils, daß die Abbildung die jeweilige Eigenschaft besitzt, und gib im Falle der Bijektivität auch die Umkehrabbildung an.

§ 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 $\mathrm{nf} : \mathbb{N} \longrightarrow \mathbb{N} : n \mapsto n+1$ können wir die Eigenschaft auch wie folgt formulieren:

> Ist $M \subseteq \mathbb{N}$ mit $0 \in M$ und $\forall\, n \in M \;:\; n+1 = \mathrm{nf}(n) \in M$, 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 $n \in \mathbb{N}$. □

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. □

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 $k = n_0, \ldots, n$ als richtig voraussetzen und auf $\mathcal{A}(n+1)$ schließen (wobei $\mathcal{A}(n_0)$ der Induktionsanfang sein soll). Das ist manchmal hilfreich.

Bemerkung 4.6 (Vollständige Induktion ist ohne Induktionsanfang wertlos.)

Bei einem Induktionsbeweis ist es meist weit leichter, den Induktionsanfang zu zeigen als den Induktionsschluß. Man sollte diesen aber auf keinen Fall vernachlässigen, da ein Beweis mittels Induktion ohne Induktionsanfang schlicht kein Beweis ist.

Wir wollen dies an folgendem Beispiel demonstrieren. Wir betrachten dazu die Aussageform

$$\mathcal{A}(n) : \exists\, k \in \mathbb{N} \;:\; 10^n = 3 \cdot k,$$

d.h. jede Potenz von 10 ist durch 3 teilbar. Wollte man dies durch vollständige Induktion beweisen, wäre der Induktionsschluß leicht zu führen. Man nimmt dazu an, die Aussage sei für eine natürliche Zahl n schon gezeigt, d.h. es gibt ein $k \in \mathbb{N}$ mit

$$10^n = 3 \cdot k$$

und folgert dann, daß auch

$$10^{n+1} = 10^n \cdot 10 = (3 \cdot k) \cdot 10 = 3 \cdot (10k)$$

ein Vielfaches von 3 ist. Die Korrektheit des Schlusses allein reicht aber nicht aus, um die Aussage auch nur für eine einzige natürliche Zahl n zu zeigen. Solange wir nicht mindestens eine natürliche Zahl gefunden haben, für die die Aussage korrekt ist und die uns dann als Induktionsanfang dient, ist der Beweis unvollständig. Einen solchen Induktionsanfang werden wir im vorliegenden Beispiel aber nicht finden, weil schlicht keine Potenz von 10 durch 3 teilbar ist!

Das Beispiel sollte auch noch mal verdeutlichen, daß der Beweis der Korrektheit des Schlusses "$\mathcal{A}(n) \Rightarrow \mathcal{A}(n+1)$" *nicht* bedeutet, daß $\mathcal{A}(n+1)$ wahr ist! Dies folgt nur, *wenn* $\mathcal{A}(n)$ wirklich wahr ist, nicht aber, wenn $\mathcal{A}(n)$ falsch ist, was der Korrektheit der Schlußfolgerung an sich keinen Abbruch tut!

Aufgaben

Aufgabe 4.7
Zeige, daß $3^{n+1} - 3$ für jede natürliche Zahl $n \in \mathbb{N}$ durch 6 teilbar ist.

Aufgabe 4.8
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.

Aufgabe 4.9
Beweise mittels vollständiger Induktion $\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \cdots + \frac{1}{\sqrt{n-1}} + \frac{1}{\sqrt{n}} > \sqrt{n}$ für alle $n \geq 2$.

Aufgabe 4.10
Beweise die folgenden Aussagen mittels vollständiger Induktion:

a. Für $n \geq 2$ gilt $\prod_{k=2}^{n} \left(1 - \frac{1}{k}\right) = \frac{1}{n}$.

b. Für $n \in \mathbb{N}$ gilt $\sum_{k=0}^{n} (k+1) \cdot \binom{n}{k} = 2^{n-1} \cdot (n+2)$.

c. Für $n \geq 2$ gilt $\frac{4^n}{n+1} < \frac{(2n)!}{(n!)^2}$.

d. Für $n \in \mathbb{N}$ gilt $\sum_{k=1}^{n} (-1)^{n-k} \cdot k^2 = \binom{n+1}{2}$.

§ 5 Mächtigkeit von Mengen

A) Endliche Mengen

Definition 5.1 (Die Mächtigkeit von Mengen)

a. Wir nennen eine Menge M *endlich*, wenn sie nur endlich viele Elemente enthält. In diesem Fall bezeichnen wir mit $\#M = |M|$ die Anzahl an Elementen in M und nennen die Zahl die *Mächtigkeit* von M. Enthält M unendlich viele Elemente, so nennen wir M *unendlich* und setzen $\#M := |M| := \infty$.

b. Zwei Mengen M und N heißen *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.

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

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

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

Bemerkung 5.2 (Einfache Eigenschaften der Mächtigkeit endlicher Mengen)

a. Ist eine Menge endlich und enthält genau n Elemente, so können wir die Elemente in M abzählen, etwa $x_1, x_2, x_3, \ldots, x_n$ und wir erhalten so eine bijektive Abbildung

$$f : \{1, \ldots, n\} \longrightarrow M : i \mapsto x_i.$$

Umgekehrt erlaubt eine solche Abbildung, die Elemente von M abzuzählen und wir erhalten $|M| = n$. Damit sehen wir, daß eine Menge genau dann endlich von Mächtigkeit n ist, wenn es eine Bijektion von $\{1, \ldots, n\}$ nach M gibt.

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 $|M| = |A| + |B|$.

Wir wollen den in Bemerkung 5.2 angedeuteten Zusammenhang zwischen der Mächtigkeit einer endlichen Menge und der Existenz von Abbildungen mit bestimmten Eigenschaften im folgenden Satz vertiefen.

Satz 5.3
Es seien M und N zwei nicht-leere endliche Mengen.

a. Genau dann gilt $|M| \leq |N|$, wenn es eine injektive Abbildung $f : M \to N$ gibt.

b. Genau dann gilt $|M| \geq |N|$, wenn es eine surjektive Abbildung $f : M \to N$ gibt.

c. Genau dann gilt $|M| = |N|$, wenn es eine bijektive Abbildung $f : M \to N$ gibt.

Beweis: Es seien $M = \{x_1, \ldots, x_m\}$ und $N = \{y_1, \ldots, y_n\}$ mit paarweise verschiedenen Elementen $x_i \neq x_j$ für $i \neq j$ und $y_i \neq y_j$ für $i \neq j$. Es gilt $|M| = m > 0$ und $|N| = n > 0$.

a. Ist $m \leq n$, so definiere $f : M \to N$ durch $f(x_i) = y_i$ für $i = 1, \ldots, m$. Dann gilt für $i, j \in \{1, \ldots, m\}$ mit $i \neq j$

$$f(x_i) = y_i \neq y_j = f(x_j).$$

Mithin ist f injektiv.
Ist umgekehrt $f : M \to N$ eine injektive Abbildung, so gilt $f(M) = \{f(x_1), \ldots, f(x_m)\} \subseteq N$ eine Teilmenge von paarweise verschiedenen Elementen. Mithin enthält N mindestens m Elemente, und folglich gilt $m \leq n$.

b. Ist $m \geq n$, so definiere $f : M \to N$ durch $f(x_i) = y_i$ für $i = 1, \ldots, n$ und $f(x_i) = y_1$ für $i = n+1, \ldots, m$. Dann gilt offenbar $f(M) = \{y_1, \ldots, y_n\} = N$ und f ist surjektiv.
Ist umgekehrt $f : M \to N$ eine surjektive Abbildung, so gilt $\{y_1, \ldots, y_n\} = N = f(M) = \{f(x_1), \ldots, f(x_m)\}$. Mithin enthält die Menge $\{f(x_1), \ldots, f(x_m)\}$ auch n verschiedene Elemente, und folglich ist $m \geq n$.

c. Die Rückrichtung folgt unmittelbar aus den ersten beiden Teilen. Für die Hinrichtung beachte man, daß die in a. und b. definierten Abbildungen im Fall $|M| = |N|$ übereinstimmen und somit bijektiv sind.

□

Bemerkung 5.4 (Schubfachprinzip)
Die Kontraposition von Teil a. in Satz 5.3 nennt man auch das *Schubfachprinzip*: wenn $f : M \longrightarrow N$ eine Abbildung ist und $|M| > |N|$ gilt, dann ist f nicht injektiv.

Übersetzt, wenn man $m > n$ Gegenstände auf n Schubfächer verteilen möchte, dann muß man in mindestens ein Schubfach zwei legen.

Beispiel 5.5
In einem Raum mit 13 Personen haben mindestens zwei im selben Monat Geburtstag.

Aus Satz 5.3 leitet sich zudem unmittelbar ab, daß für Selbstabbildungen endlicher Mengen die Begriffe injektiv, surjektiv und bijektiv zusammen fallen.

Korollar 5.6 (Injektiv = surjektiv = bijektiv für gleichmächtige endliche 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. $\Longrightarrow$ **b.:** Angenommen, f wäre nicht surjektiv, dann gibt es ein

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

und mithin ist

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

Da f injektiv ist, ist $g : M \longrightarrow \operatorname{Im}(f) : x \mapsto f(x)$ nach Beispiel 3.8 bijektiv, so daß mit Satz 5.3

$$|M| \overset{5.3}{=} |\operatorname{Im}(f)| \leq |N| - 1 < |N| = |M|$$

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

b. $\Longrightarrow$ **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 3.16 surjektiv. Mithin gilt nach Satz 5.3

$$|M| - 1 \overset{\text{Vor.}}{=} |N| - 1 = |N \setminus \{y\}| \overset{5.3}{\leq} |M \setminus f^{-1}(\{y\})| \leq |M \setminus \{x, x'\}| = |M| - 2,$$

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

c. $\Longrightarrow$ **a.:** Jede bijektive Abbildung ist auch injektiv, also ist f injektiv.

Damit haben wir die Aussage durch einen *Ringschluß* gezeigt. □

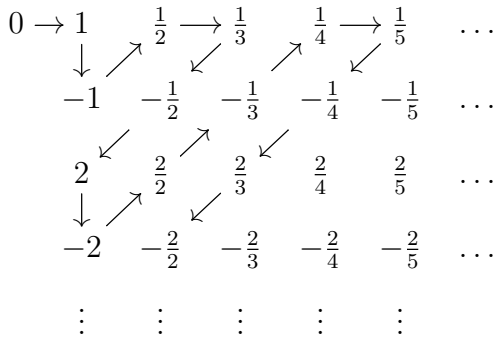
B) Das Cantorsche Diagonalisierungsverfahren

Nachdem wir uns bislang im wesentlichen mit endlichen Mengen beschäftigt haben, wollen wir uns nun unendlichen Mengen zuwenden und dabei zeigen, daß es unterschiedliche Qualitäten der Unendlichkeit gibt.

Proposition 5.7 (Cantorsches Diagonalverfahren)
Die Menge $\mathbb{Q}$ der rationalen Zahlen ist abzählbar unendlich.

Beweis: Wir zeigen, wie man mit Hilfe des Cantorschen Diagonalverfahrens 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 5.8 ($\mathbb{R}$ 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 $\varphi : \mathbb{N} \to \mathbb{R}$, und wir schreiben dann $\varphi(i)$, $i \in \mathbb{N}$, in Dezimaldarstellung (siehe auch Satz 12.45):

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

Dann setzen wir $a := a_{00}, a_{11}a_{22}a_{33}\cdots \in \mathbb{R}$, d. h. a ist diejenige Zahl, die in obiger Aufzählung durch die unterstrichenen Diagonalelemente gegeben ist. Nun ändern wir jede der Ziffern von a ab (etwa $b_{ii} = 2$, falls $a_{ii} = 0$ und $b_{ii} = 0$ sonst) und erhalten eine Zahl

$$b = b_{00}, b_{11}b_{22}b_{33}\cdots \in \mathbb{R},$$

mit $a_{ii} \neq b_{ii}$ für alle $i \in \mathbb{N}$. Da φ bijektiv ist, gibt es ein $i \in \mathbb{N}$ mit $\varphi(i) = b$, also $a_{ii} = b_{ii}$, im Widerspruch zur Konstruktion von b. (Wir müssen noch berücksichtigen,

daß $0,9999\cdots = 1$, was aber die einzige Zweideutigkeit der Dezimaldarstellung ist, und dieser weichen wir durch unsere Wahl der b_{ii} aus.) Also ist $\mathbb{R}$ überabzählbar. □

Bemerkung 5.9 (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 $\mathbb{Q} \subsetneqq M \subsetneqq \mathbb{R}$ 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.

C) Potenzmengen

Definition 5.10 (Potenzmenge)
Es sei M ein eine Menge. Wir nennen die Menge

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

aller Teilmengen von M die *Potenzmenge* von M.

Beispiel 5.11
$\mathcal{P}(\emptyset) = \{\emptyset\}, \quad \mathcal{P}(\{1\}) = \{\emptyset, \{1\}\}, \quad \mathcal{P}(\{1,2\}) = \{\emptyset, \{1\}, \{2\}, \{1,2\}\}.$

Proposition 5.12 (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:

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

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 5.13
Die Menge $\mathbb{Z}$ der ganzen Zahlen ist abzählbar unendlich.

Aufgabe 5.14
In einer Gruppe mit mindestens zwei Personen gibt es mindestens zwei, die dieselbe Anzahl an Personen in der Gruppe kennen.

Aufgabe 5.15
Sei M eine Menge und $N := \{g \mid g : M \to \{0,1\} \text{ Abbildung}\}$. Zeige, die Menge N ist gleichmächtig zur Potenzmenge $\mathcal{P}(M)$.

§ 6 Äquivalenzrelationen

Äquivalenzrelationen stellen ein sehr wichtiges *Ordnungs-* und *Konstruktionsprinzip* innerhalb der Mathematik dar, das im Verlauf der ersten Semester an einigen zentralen Stellen benötigt wird, etwa im Zusammenhang mit Faktorräumen, der Äquivalenz von Matrizen oder der Konjugation von Matrizen und der Jordanschen Normalform.

A) Äquivalenzrelationen und Äquivalenzklassen

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)

a. Der Graph einer Abbildung $f : M \longrightarrow N$ ist ein Beispiel einer Relation, bei der jedes $x \in M$ zu genau einem $y \in N$ in Relation steht.

b. Ist M die Menge der Hörer der Vorlesung und N die Menge der in Tübingen 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, $i = 1, \ldots, k$, ein, so daß wir anschließend eine *disjunkte Zerlegung*

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

der Menge S in die Schulklassen $K_1, \ldots, K_k$ 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 derselben Schulklasse ist wie Ben, dann ist Ben auch in derselben Schulklasse wie Alfred.
3) Wenn Alfred in derselben Schulklasse ist wie Ben und wenn zugleich Ben in derselben Schulklasse ist wie Christoph, dann ist auch Alfred in derselben 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 selben 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

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

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 derselben 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: $(x, y) \in R \implies (y, x) \in R$, ("Symmetrie")
R3: $(x, y), (y, z) \in R \implies (x, z) \in R$. ("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)

Äquivalenzrelationen werden oft eher mit Symbolen wie $\sim$ statt Buchstaben wie R bezeichnet. Mit der Schreibweise aus Bemerkung 6.2 gilt dann

$$x \sim y \iff (x, y) \in \sim$$

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

R1: $x \sim x$, ("Reflexivität")
R2: $x \sim y \implies y \sim x$, ("Symmetrie")
R3: $x \sim y, y \sim z \implies x \sim z$. ("Transitivität")

Definition 6.7 (Äquivalenzklassen)

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

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

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

$$M/\sim := \left\{\overline{x} \mid x \in M\right\}$$

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

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

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

B) Äquivalenzrelationen und disjunkte Zerlegungen

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. □

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 $\overline{x}$ und $\overline{y}$ entweder $\overline{x} = \overline{y}$ oder $\overline{x} \cap \overline{y} = \emptyset$.

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

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

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

Seien $\overline{x}, \overline{y} \in M/\sim$ mit $\overline{x} \cap \overline{y} \neq \emptyset$. Dann gibt es ein $z \in \overline{x} \cap \overline{y}$, 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 $u \in \overline{x}$ beliebig, dann gilt $u \sim x$ und wieder wegen der Transitivität $u \sim y$. Also $u \in \overline{y}$ und damit $\overline{x} \subseteq \overline{y}$. Vertauschung der Rollen von x und y in der Argumentation liefert schließlich $\overline{x} = \overline{y}$. □

Korollar 6.11 (Äquivalenzrelationen auf endlichen Mengen)
Sei M eine endliche Menge, $\sim$ eine Äquivalenzrelation auf M und $M_1, \ldots, M_s$ 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 5.2. □

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 $(p, q), (p', q') \in M := \mathbb{Z} \times \big(\mathbb{Z} \setminus \{0\}\big)$ definiere

$$(p, q) \sim (p', q') \quad :\Longleftrightarrow \quad pq' = p'q.$$

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 $x' = (p', q') \sim (p, q) = x$.

[1] Man 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

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

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

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

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

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 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:

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

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 Klassen 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 $(x, y), (x', y') \in \mathbb{R}^2$

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

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

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

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

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

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

a. 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 $\mathbb{P}^1_{\mathbb{R}}$.

b. 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 Abbilung

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

surjektiv ist.

c. 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 $\overline{0}, \overline{1}, \ldots, \overline{n-1}$.
Man nennt zwei äquivalente Zahlen x und y dann auch *kongruent modulo* n. Diese Äquivalenzrelation wird in der Vorlesung algebraische Strukturen genauer untersucht.

Aufgabe 6.17
Für zwei reelle Zahlen $x, y \in \mathbb{R}$ definieren wir

$$x \sim y \iff x^2 - y^2 = 2x - 2y.$$

Zeige, daß $\sim$ eine Äquivalenzrelation auf $\mathbb{R}$ ist und bestimme die Äquivalenzklassen von 0 und 1.

§ 7 Gruppen und Körper

Gruppen und Körper sind Themen, die ausführlich in Vorlesungen zur Algebra behandelt werden. Wir wollen die Begriffe hier nur so weit einführen, wie sie für die Analysis erforderlich sind.

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

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

so daß die folgenden *Gruppenaxiome* gelten:

G1: $(g * h) * k = g * (h * k) \quad \forall\, g, h, k \in G,$ ("Assoziativgesetz")

G2: $\exists\, e \in G \;:\; \forall\, g \in G \;:\; e * g = g,$ ("Existenz eines Neutralen")

G3: $\forall\, g \in G\, \exists\, g^{-1} \in G \;:\; g^{-1} * g = e.$ ("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:

G4: $g * h = h * g \quad \forall\, g, h \in G$ ("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 nicht-leere Menge, so ist die Menge

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

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 $(\mathrm{Sym}(M), \circ)$ die *symmetrische Gruppe* auf M. Enthält M mehr als zwei Elemente, so ist $\mathrm{Sym}(M)$ 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 $\left(g^{-1}\right)^{-1} = g$ 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 Vorlesungen zur (Linearen) Algebra 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

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

Also folgt:

$$\begin{aligned} g * g^{-1} &\overset{G2}{=} e * \left(g * g^{-1}\right) \overset{(2)}{=} \left((g^{-1})^{-1} * g^{-1}\right) * \left(g * g^{-1}\right) \overset{G1}{=} (g^{-1})^{-1} * \left(g^{-1} * (g * g^{-1})\right) \\ &\overset{G1}{=} (g^{-1})^{-1} * \left((g^{-1} * g) * g^{-1}\right) \overset{G3}{=} (g^{-1})^{-1} * \left(e * g^{-1}\right) \overset{G2}{=} (g^{-1})^{-1} * g^{-1} \overset{(2)}{=} e. \end{aligned} \tag{3}$$

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

$$g * e \overset{G3}{=} g * \left(g^{-1} * g\right) \overset{G1}{=} \left(g * g^{-1}\right) * g \overset{(3)}{=} e * g \overset{G2}{=} g. \tag{4}$$

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.

$$\tilde{e} * h = h \tag{5}$$

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:

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

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

$$\tilde{g}^{-1} * g = e \tag{6}$$

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

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

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:

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

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

$$\left(g^{-1}\right)^{-1} = g.$$

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:

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

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. $g * a = g * b \implies a = b$, und

b. $a * g = b * g \implies a = b$.

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

$$a \overset{G2}{=} e * a \overset{G3}{=} \left(g^{-1} * g\right) * a \overset{G1}{=} g^{-1} * (g * a)$$
$$\overset{Vor.}{=} g^{-1} * (g * b) \overset{G1}{=} \left(g^{-1} * g\right) * b \overset{G3}{=} e * b \overset{G2}{=} b.$$

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

$$+ : K \times K \to K : (x, y) \mapsto x + y, \qquad (\text{“}\textit{Addition}\text{”})$$

und

$$\cdot : K \times K \to K : (x, y) \mapsto x \cdot y, \qquad (\text{“}\textit{Multiplikation}\text{”})$$

so daß folgende Axiome erfüllt sind:

a. $(K, +)$ ist eine abelsche Gruppe mit neutralem Element 0.

b. $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe mit neutralem Element 1.

c. Es gelten die *Distributivgesetze* $x \cdot (y+z) = x \cdot y + x \cdot z$ und $(y+z) \cdot x = y \cdot x + z \cdot x$ 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 $\mathbb{F}_p$)

a. 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 $\mathbb{R}$.

b. Die ganzen Zahlen $(\mathbb{Z}, +, \cdot)$ sind kein Körper, da z.B. der Zahl 2 ein multiplikatives Inverses fehlt.

c. Auf der Menge $\mathbb{F}_2 := \{0, 1\}$ definieren wir zwei Operationen durch folgende Additions- und Multiplikationstafeln:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

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 Verknüpfungstafeln besteht, wenn man einen Körper mit genau zwei Elementen haben möchte. — Beachte auch, daß $\mathbb{F}_2$ kein Teilkörper von $\mathbb{R}$ ist, da das Ergebnis von $1 + 1$ in den beiden Körpern nicht übereinstimmt.

d. Allgemeiner zeigt man in Vorlesungen zur (Linearen) Algebra, daß man für eine Primzahl p die Menge

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

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 "$\cdot$" auf der rechten Seite jeweils die Operation in den ganzen Zahlen bezeichnet, während das "$+$" und das "$\cdot$" 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

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

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 der (Linearen) Algebra, 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 der (Linearen) Algebra schreibt man übrigens oft $\mathbb{Z}/p\mathbb{Z}$ oder $\mathbb{Z}_p$ anstatt $\mathbb{F}_p$, und die Zahl a wird dort meist mit $\overline{a}$ 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. $x + y = z \iff x = z - y$,

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. $\left(x^{-1}\right)^{-1} = x$, für $x \neq 0$,

i. $x \cdot y = 0 \iff x = 0$ oder $y = 0$,

j. $z \cdot x = z \cdot y,\ z \neq 0 \implies x = y$,

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.:

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

g. Für $x, y, z \in K$ impliziert e.:

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

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

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

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

□

C) Summen, Produkte und der Binomische Lehrsatz

Notation 7.9 (Produkte und Summen)

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

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

für das *Produkt* der Zahlen $x_0, \ldots, x_n$ und

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

für die *Summe* der Zahlen $x_0, \ldots, x_n$. Wir einigen uns dabei darauf, daß das leere Produkt (d.h. ein Produkt, bei dem der obere Index kleiner als der untere ist) den Wert 1 hat und die leere Summe den Wert 0.

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

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

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

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

Analog dazu setzen wir

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

und

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

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

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

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. □

Definition 7.13 (Fakultät)
Für eine natürliche Zahl $n \in \mathbb{N}$ definieren wir die *Fakultät* durch

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

falls $n \geq 1$, und durch $0! := 1$.
Für eine natürliche Zahl $n \in \mathbb{N}$ und $k \in \mathbb{Z}$ erklären wir den *Binomialkoeffizienten* von n über k durch

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

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

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

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:

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

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

0. Zeile:				1				
1. Zeile:			1		1			
2. Zeile:		1		2		1		
3. Zeile:		1		3		3		1
4. Zeile:	1		4		6		4		1

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:

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

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 $(x+y)^n$ ausschreiben. Z.B.

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

Aufgaben

Aufgabe 7.17

Sei M eine Menge. Zeige, die Potenzmenge $\mathcal{P}(M)$ wird mittels der symmetrischen Differenz

$$A + B := (A \cup B) \setminus (A \cap B)$$

für $A, B \in \mathcal{P}(M)$ eine abelsche Gruppe.

Aufgabe 7.18
Es sei K ein Körper und $x \in K$. Zeige, $x^2 = 1$ genau dann, wenn $x \in \{1, -1\}$.

Aufgabe 7.19

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

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

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

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

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 "·" wie in a. und b. definiert sind.

Aufgabe 7.20
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 7.21 (Die projektive Gerade als Gruppe)
Wir haben in Aufgabe 6.15 die Projektive Gerade $\mathbb{P}^1_{\mathbb{R}}$ 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 $\mathbb{P}^1_{\mathbb{R}}$ mit dieser Operation eine Gruppe ist.

Aufgabe 7.22
Für zwei natürliche Zahlen $1 \leq k \leq n$ gilt

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

§ 8 Ordnungsrelationen

In diesem Abschnitt wollen wir Körper mit einer zusätzlichen Struktur einführen und betrachten, sogenannte angeordnete Körper. Dazu führen wir zunächst den Begriff der Ordnungsrelation ein.

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")

Wir nennen dann (M, R) auch eine *partiell* oder *teilgeordneten Menge*.

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 8.18).

Notation 8.3 (Schreibweise $\leq$ für Ordnungsrelationen)
Ordnungsrelationen bezeichnet man meist eher mit Symbolen wie $\leq$ anstatt mit R. Sei also M eine Menge und $\leq$ eine Ordnungsrelation auf M. Dann gilt in unserer Notation aus Bemerkung 6.2 für $x, y \in M$

$$x \leq y \iff (x, y) \in \leq$$

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

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

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.

a. 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 $y \leq x$.

b. 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 $x = y$.

c. 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

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

b. Gleiches trifft auf $(\mathbb{Z}, \leq)$ mit der üblichen Kleiner-Gleich-Relation

$$\ldots -2 < -1 < 0 < 1 < 2 < \ldots$$

zu. Allerdings definiert die "unübliche" Anordnung

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

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*.

B) Das Supremumsaxiom

Definition 8.9 (Supremum und Infimum)
Es sei "$\leq$" eine Totalordnung auf einer Menge M und $\emptyset \neq A \subseteq M$ eine nicht-leere Teilmenge von M.

a. Wir nennen $s \in M$ eine *obere Schranke* von A, falls $s \geq x$ für alle $x \in A$.

b. Wir nennen A *nach oben beschränkt*, falls A eine obere Schranke besitzt.

c. 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 $\sup(A)$.

d. Wir nennen $s \in M$ eine *untere Schranke* von A, falls $s \leq x$ für alle $x \in A$.

e. Wir nennen A *nach unten beschränkt*, falls A eine untere Schranke besitzt.

f. 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 $\inf(A)$.

g. Wir nennen A *beschränkt*, wenn A nach oben und nach unten beschränkt ist.

Beispiel 8.10

a. 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.

b. 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.

c. 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 9.10).

Bemerkung 8.11 (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 8.10 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.

C) Angeordnete Körper

Definition 8.12 (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 $x, y, z \in K$

$$x < y \quad \Longrightarrow \quad x + z < y + z$$

und

$$x < y, \;\; 0 < z \quad \Longrightarrow \quad x \cdot z < y \cdot z$$

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

Beispiel 8.13

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 8.10), $\mathbb{R}$ 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 8.14 (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 8.15 (Charakterisierung des Supremums und Infimums)
Ist $(K, +, \cdot, \leq)$ ein angeordneter Körper, $A \subseteq K$ und $s \in K$, dann gelten

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

sowie

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

Beweis: Ist $s = \sup(A)$, so ist s ein 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 8.16
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 8.17

Ist M eine endliche Menge, so gilt

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

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

Aufgabe 8.18

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 8.19

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

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

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 8.20

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 8.21

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

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

und

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

Aufgabe 8.22 (Jeder angeordnete Körper enthält $\mathbb{N}$.)
Sei $(K, +, \cdot, \leq)$ ein angeordneter Körper mit Eins 1_K. Zeige, die Abbildung

$$\mathbb{N} \longrightarrow K : n \mapsto \sum_{i=1}^{n} 1_K$$

ist injektiv und verträglich mit den Operationen $+$ und $\cdot$.

Aufgabe 8.23
Gib alle möglichen Ordnungsrelationen auf der 2-elementigen Menge $M = \{a, b\}$ als Teilmengen von $M \times M$ an. Begründe, weshalb dies alle sind.

Aufgabe 8.24
Für zwei Tupel $(a, b), (c, d) \in \mathbb{N} \times \mathbb{N}$ definieren wir

$$(a, b) \ \leq \ (c, d) \quad :\Longleftrightarrow \quad a + b < c + d \ \text{ oder } \ (a + b = c + d \ \text{ und } \ a \leq c).$$

Zeige, $\leq$ ist eine Wohlordnung auf der Menge $M = \mathbb{N} \times \mathbb{N}$ und stelle graphisch (in der Ebene) dar, wie die Tupel (a, b) mit $a + b \leq 4$ angeordnet sind.

§ 9 Eigenschaften der reellen Zahlen $\mathbb{R}$

In diesem Abschnitt wollen wir einige wichtige Eigenschaften der reellen Zahlen betrachten.

A) Axiomatische Charakterisierung der reellen Zahlen

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

$$x \leq y \quad \Longleftrightarrow \quad f(x) \leq f(y)$$

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 Zahlsystemen 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 $y < n \cdot x$. □

Korollar 9.4 (Konsequenzen der archimedischen Anordnung)

a. Für alle $x \in \mathbb{R}$ gibt es eine ganze Zahl n, so daß $n \leq x < n + 1$.

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

Beweis:

a. 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 $n := -m - 1$.

b. Wegen $\varepsilon > 0$ ist nach Lemma 8.14 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 8.14 folgt dann

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

□

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

$$0 < \frac{1}{n} < b - a. \tag{7}$$

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

$$m \leq n \cdot a < m + 1. \tag{8}$$

Damit gilt dann

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

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

B) Weitere Eigenschaften der reellen Zahlen

Satz 9.7 (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 8.14 b. ist $x^2 \geq 0$ und nach Voraussetzung gilt zudem $1 + x \geq 0$. Damit erhalten wir dann:

$$(1+x)^{n+1} = (1+x)^n \cdot (1+x) \overset{\text{Ind}}{\geq}$$

$$(1 + n \cdot x) \cdot (1+x) = 1 + (n+1) \cdot x + n \cdot x^2 \overset{8.14b.}{\geq} 1 + (n+1) \cdot x.$$

Die Aussage ist damit also mittels Induktion gezeigt. □

Mit Hilfe der Bernoulli-Ungleichung kann man die Existenz von n-ten Wurzeln nicht-negativer reeller Zahlen für alle $n \geq 2$ herleiten. Wir führen den Beweis hier auf, werden in der Vorlesung in Beispiel 14.23 aber mit Hilfe des Umkehrsatzes für streng monotone stetige Funktionen einen viel kürzeren und eleganteren Beweis geben.

Satz 9.8 (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 8.14 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

$$y^n \overset{8.14g.}{\geq} (1+x)^n \overset{9.7}{\geq} 1+n\cdot x > x,$$

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* $a+\varepsilon \in A$*.* $-\lightning$

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 8.14. 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

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

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* $y^n > (a-\varepsilon)^n \geq x$*.* $-\lightning$

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

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

Aus der Definition von ε folgt zum einen

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

und zum anderen unter Anwendung der Bernoullischen Ungleichung

$$x \leq a^n\cdot\left(1+n\cdot\frac{-\varepsilon}{a}\right) \overset{9.7}{\leq} a^n\cdot\left(1-\frac{\varepsilon}{a}\right)^n = (a-\varepsilon)^n\,; \tag{11}$$

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

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

$$y > a-\varepsilon > 0.$$

Dann gilt nach Lemma 8.14 auch

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

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

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

Bemerkung 9.9

In $\mathbb{R}$ besitzt also insbesondere jede nicht-negative Zahl eine Quadratwurzel. Dies gilt in den rationalen Zahlen nicht (siehe Satz 9.10), 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 9.10 ($\sqrt{2}$ ist irrational.)

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

Beweis: Nehmen wir an, es wäre $a = \frac{p}{q}\in\mathbb{Q}$ eine solche Zahl. Wir können ohne weiteres annehmen, daß der Bruch in gekürzter Form vorliegt. Aus

$$\frac{p^2}{q^2} = a^2 = 2$$

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 $b \in \mathbb{Z}$ mit $p = 2 \cdot b$. Also ist

$$4 \cdot b^2 = p^2 = 2 \cdot q^2,$$

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. □

Aufgaben

Aufgabe 9.11

Zeige durch vollständige Induktion, daß

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} + \ldots + \frac{1}{\sqrt{n-1}} + \frac{1}{\sqrt{n}} > \sqrt{n}$$

für alle $n \geq 2$ gilt.

Aufgabe 9.12

Zeige, für je drei reelle Zahlen $a, b, c \in \mathbb{R}$ gilt

$$\frac{|a-c|}{1+|a-c|} \leq \frac{|a-b|}{1+|a-b|} + \frac{|b-c|}{1+|b-c|}$$

Aufgabe 9.13

a. Schreibe die folgenden Mengen als Vereinigung von Intervallen:

(i) $A_1 = \{x \in \mathbb{R} \mid |x-1| < 1 \ \vee \ |x+2| \leq 2\}$.

(ii) $A_2 = \{x \in \mathbb{R} \mid |x-1| > 1 \ \wedge \ |x+2| \leq 2\}$.

(iii) $A_3 = \{x \in \mathbb{R} \mid x^2 - 2 \leq 2 \ \vee \ \sqrt{|x-10|} > 5\}$.

b. Überprüfe zudem jede der Mengen in Teil (a) sowie die folgende Menge, ob sie nach oben oder unten beschränkt ist, gib ggf. das Infimum oder Supremum der Menge an und überprüfe auch, ob es ein Minimim oder Maximum ist:

$$M = \left\{ \frac{m-n}{m+n} \mid 0 \neq m, n \in \mathbb{N} \right\}.$$

Aufgabe 9.14

Für je zwei reelle Zahlen $x, y \in \mathbb{R}$ gilt

$$\max\{x, y\} = \frac{1}{2} \cdot (x + y + |x - y|).$$

§ 10 Der Körper der komplexen Zahlen

Wir betachten in diesem Abschnitt den 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 Zahlsystem 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.

[2] Erstmals tauchte $\sqrt{-1}$ wohl um 1540 bei Cardano auf. Wirklich als Zahlsystem 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.

A) Die Arithmetik der komplexen Zahlen

Bemerkung 10.1 (Konstruktion der komplexen Zahlen)
Es ist unser erklärtes Ziel, auf der reellen Zahlenebene $\mathbb{R}^2$ mit der Vektoraddtion

$$(x, y) + (u, v) := (x + u, y + v)$$

eine *Multiplikation* zu definieren, so daß einerseits die üblichen Rechenregeln (Assoziativgesetze, Kommutativgesetze und Distributivgesetze) gelten und daß außerdem der Vektor

$$i := (0, 1)$$

eine Lösung der Gleichung

$$z^2 = -1$$

ist. Um letzteres richtig zu interpretieren, denken wir uns die reelle Zahlengerade $\mathbb{R}$ als Teilmenge von $\mathbb{R}^2$, indem wir sie mit der x-Achse identifizieren, d.h.

$$\mathbb{R} \mathrel{\hat{=}} \{(a, 0) \mid a \in \mathbb{R}\} = x\text{-Achse}.$$

Die Multiplikation soll also der Bedingung

$$i^2 = (0, 1) \cdot (0, 1) = (-1, 0) \mathrel{\hat{=}} -1$$

genügen. Außerdem würden wir uns sicher wünschen, daß die Multiplikation eines Vektors mit der reellen Zahl

$$a \mathrel{\hat{=}} (a, 0)$$

wie die Streckung des Vektors um den Faktor a funktioniert, d.h.

$$(a, 0) \cdot (x, y) \mathrel{\hat{=}} a \cdot (x, y) = (ax, ay).$$

Wenn eine Multiplikation diese Wunschliste erfüllt, so gilt offenbar:

$$\begin{aligned}
(x, y) \cdot (u, v) &= \big((x, 0) + (0, y)\big) \cdot \big((u, 0) + (0, v)\big) \\
&= \big((x, 0) + (y, 0) \cdot (0, 1)\big) \cdot \big((u, 0) + (v, 0) \cdot (0, 1)\big) \\
&= (x, 0) \cdot (u, 0) + (y, 0) \cdot (0, 1) \cdot (u, 0) + (x, 0) \cdot (v, 0) \cdot (0, 1) \\
&\qquad + (y, 0) \cdot (0, 1) \cdot (v, 0) \cdot (0, 1) \\
&= (xu, 0) + (yu, 0) \cdot (0, 1) + (xv, 0) \cdot (0, 1) + (yv, 0) \cdot (0, 1) \cdot (0, 1) \\
&= (xu, 0) + (yu, 0) \cdot (0, 1) + (xv, 0) \cdot (0, 1) + (yv, 0) \cdot (-1, 0) \\
&= (xu, 0) + (0, yu) + (0, xv) + (-yv, 0) \\
&= (xu - yv, xv + yu).
\end{aligned}$$

Wir haben für die Definition der Multiplikation also nur *eine einzige* Möglichkeit, und die funktioniert zum Glück auch.

Satz 10.2 (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

$$(x, y) \cdot (u, v) := (xu - yv, xv + yu), \quad \text{für } (x, y), (u, v) \in \mathbb{C},$$

definierten Addition und Multiplikation ist ein Körper, den wir den Körper der *komplexen Zahlen* nennen. .

Beweis: Dies folgt aus Aufgabe 7.19. □

Bemerkung 10.3

a. Daß $\mathbb{C}$ mit den beiden Operationen ein *Körper* ist, bedeutet, daß die oben erwähnten üblichen Rechenregeln bezüglich der Addition, Subtraktion, Multiplikation und Division gelten, so wie wir sie von den reellen Zahlen her kennen. Man beachte dabei, daß die reelle Zahl $0 \hat{=} (0, 0)$ bei der Addtion nichts tut und die reelle Zahl $1 \hat{=} (1, 0)$ bei der Multiplikation ohne Wirkung ist:

$$(x, y) + (0, 0) = (x + 0, y + 0) = (x, y)$$

und

$$(x, y) \cdot (1, 0) = (x \cdot 1 - y \cdot 0, x \cdot 0 + y \cdot 1) = (x, y).$$

Das multiplikative Inverse der Zahl $(0, 0) \neq (x, y) \in \mathbb{C}$ ist

$$(x, y)^{-1} = \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right).$$

b. Die Abbildung

$$\iota : \mathbb{R} \longrightarrow \mathbb{C} : x \mapsto (x, 0)$$

ist mit der Addition und der Multiplikation verträglich und identifiziert den Körper der reellen Zahlen $\mathbb{R}$ mit dem Teilkörper $\mathbb{R} \times \{0\}$ von $\mathbb{C}$. Wir fassen $\mathbb{R}$ in diesem Sinne als Teilmenge von $\mathbb{C}$ auf.

c. 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 $z = (x, y) \in \mathbb{C}$

$$z = (x, y) = (x, 0) + (0, y) = (x, 0) + (0, 1) \cdot (y, 0) \hat{=} x + iy.$$

Diese Schreibweise wollen wir künftig für komplexe Zahlen verwenden. Damit gilt dann:

$$i^2 = (0, 1) \cdot (0, 1) = -1.$$

Ferner ergibt sich die etwas willkürlich anmutende Definition der Multiplikation ganz "natürlich" aus

$$(x+iy)(u+iv) = \big(xu + i^2yv\big) + i(xv+yu) = (xu-yv) + i(xv+yu).$$

Lemma 10.4 ($\mathbb{C}$ 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. Nach Lemma 8.14 muß dann $0 < i^2 = -1$ gelten, was im Widerspruch zu $0 < 1$ steht. □

Definition 10.5 (Der Betrag und die komplexe Konjugation)

a. Wir definieren die *Betragsfunktion* auf $\mathbb{C}$ durch

$$|\cdot| : \mathbb{C} \longrightarrow \mathbb{R}_{\geq 0} : x+iy \mapsto \sqrt{x^2+y^2}$$

und nennen $|x|$ auch den *Absolutbetrag* von x. Wegen Satz 9.8 ist der Betrag einer komplexen Zahl definiert und ist stets eine nicht-negative reelle Zahl. Beachte zudem, für $x \in \mathbb{R}$ gilt

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

b. Wir definieren die *komplexe Konjugation* als

$$\bar{\cdot} : \mathbb{C} \longrightarrow \mathbb{C} : z = x+iy \mapsto \overline{z} := x - iy.$$

Für $z \in \mathbb{C}$ heißt $\overline{z}$ die zu z *konjugiert komplexe Zahl*.

c. Wir definieren die Abbildungen *Realteil*

$$\mathrm{Re} : \mathbb{C} \longrightarrow \mathbb{R} : x+iy \mapsto x$$

und *Imaginärteil*

$$\mathrm{Im} : \mathbb{C} \longrightarrow \mathbb{R} : x+iy \mapsto y$$

und nennen $\mathrm{Re}(x+iy) = x$ den *Realteil* von z und $\mathrm{Im}(x+iy) = y$ den *Imaginärteil* von z.

Beispiel 10.6
Wir betrachten die komplexe Zahl

$$z = i - 1 = -1 + i.$$

Dann gilt $\mathrm{Re}(z) = -1$, $\mathrm{Im}(z) = 1$ und

$$\overline{z} = -1 - i = -(1+i).$$

Für den Betrag von z rechnen wir

$$|z| = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2} = \sqrt{1+1} = \sqrt{2}$$

und damit erhalten wir die Gleichung

$$z \cdot \overline{z} = (-1+i) \cdot (-1-i) = 2 = |z|^2.$$

Lemma 10.7 (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.
$$|z+w| \leq |z| + |w|,$$
und es gilt stets
$$\big||z| - |w|\big| \leq |z-w|.$$

c. $z = 0 \iff |z| = 0$.

d. $z \cdot \overline{z} = |z|^2$.

e. Wenn $z \neq 0$, dann ist $z^{-1} = \frac{1}{z} = \frac{\overline{z}}{|z|^2}$.

f. Die komplexe Konjugation ist additiv, d.h.
$$\overline{z} + \overline{w} = \overline{z+w}.$$

g. Die komplexe Konjugation ist multiplikativ, d.h.
$$\overline{z} \cdot \overline{w} = \overline{z \cdot w}.$$

h. $\overline{\overline{z}} = z$.

i. $\overline{z} = z \iff z \in \mathbb{R}$.

j. $\mathrm{Re}(z) = \frac{z+\overline{z}}{2} \leq |z|$.

k. $\mathrm{Im}(z) = \frac{z-\overline{z}}{2i} \leq |z|$.

l. $|z| = \big|\overline{z}\big| = |-z|$.

Beweis: Die Aussagen in den Teilen c.-l. überlassen wir dem Leser als Übungsaufgabe.

a. Seien $z = x + iy$ und $w = u + iv$ mit $x, y, u, v \in \mathbb{R}$. Dann gilt

$$\begin{aligned}|zw|^2 =& |(xu - yv) + i \cdot (xv + yu)|^2 = (xu - yv)^2 + (xv + yu)^2\\ =& x^2u^2 - 2xuyv + y^2v^2 + x^2v^2 + 2xvyu + y^2u^2\\ =& x^2u^2 + y^2v^2 + x^2v^2 + y^2u^2 = (x^2 + y^2) \cdot (u^2 + v^2)\\ =& |z|^2 \cdot |w|^2 = \big(|z| \cdot |w|\big)^2.\end{aligned}$$

Aus der Eindeutigkeit der nicht-negativen Quadratwurzel (Satz 9.8) folgt dann

$$|zw| = |z| \cdot |w|.$$

b. Wir wollen nun die Dreiecksungleichung unter Verwendung der übrigen Aussagen zeigen. Es gilt

$$\begin{aligned}|z + w|^2 &\overset{d.}{=} (z + w) \cdot \overline{(z + w)}\\ &\overset{f.}{=} z \cdot \overline{z} + (z \cdot \overline{w} + \overline{z} \cdot w) + w \cdot \overline{w}\\ &\overset{d.,h.}{=} |z|^2 + \big(z \cdot \overline{w} + \overline{z \cdot \overline{w}}\big) + |w|^2\\ &\overset{j.}{=} |z|^2 + 2 \cdot \mathrm{Re}(z \cdot \overline{w}) + |w|^2\\ &\overset{j.}{\leq} |z|^2 + 2 \cdot |z \cdot \overline{w}| + |w|^2\\ &\overset{a.}{=} |z|^2 + 2 \cdot |z| \cdot |\overline{w}| + |w|^2\\ &\overset{l.}{=} |z|^2 + 2 \cdot |z| \cdot |w| + |w|^2\\ &= \big(|z| + |w|\big)^2.\end{aligned}$$

Da dies eine Ungleichung von nicht-negativen Zahlen in dem angeordneten Körper $\mathbb{R}$ ist, folgt aus Lemma 8.14, daß

$$|z + w| \leq |z| + |w|.$$

Es bleibt, die zweite Aussage in Teil b. zu zeigen. Aus der Dreiecksungleichung erhalten wir

$$|z| = |(z - w) + w| \leq |z - w| + |w|,$$

und somit

$$|z| - |w| \leq |z - w|.$$

Analog folgt

$$-(|z| - |w|) = |w| - |z| \leq |w - z| = |-(w - z)| = |z - w|.$$

Wegen

$$\big||z| - |w|\big| = \begin{cases} |z| - |w|, & \text{falls } |z| - |w| \geq 0,\\ -(|z| - |w|) & \text{falls } |z| - |w| < 0,\end{cases}$$

folgt dann $\big||z| - |w|\big| \leq |z - w|$.

□

Beispiel 10.8

a. Gegeben seien $z = 3 + 2i$ und $w = 5 - i$. Dann gelten

$$z \cdot w = (3 \cdot 5 - 2 \cdot (-1)) + (3 \cdot (-1) + 2 \cdot 5) \cdot i = 17 + 7i$$

sowie

$$|w| = \sqrt{5^2 + (-1)^2} = \sqrt{26}$$

und

$$\begin{aligned} \frac{z}{w} =& z \cdot \frac{\overline{w}}{|w|^2} = (3 + 2i) \cdot \left(\frac{5}{26} + \frac{1}{26} \cdot i \right) \\ =& \left(3 \cdot \frac{5}{26} - 2 \cdot \frac{1}{26} \right) + \left(3 \cdot \frac{1}{26} + 2 \cdot \frac{5}{26} \right) \cdot i \\ =& \frac{13}{26} + \frac{13}{26} \cdot i = \frac{1}{2} + \frac{1}{2} \cdot i. \end{aligned}$$

b. Für die komplexen Zahlen $z = 3 + 4i$ und $w = 5 - 12i$ gilt

$$z + w = (3 + 5) + (4 - 12) \cdot i = 8 - 8i$$

und somit

$$\begin{aligned} |z + w| =& \sqrt{8^2 + 8^2} = \sqrt{2} \cdot 8 < 16 < 18 = 5 + 13 \\ =& \sqrt{25} + \sqrt{169} = \sqrt{3^2 + 4^2} + \sqrt{5^2 + 12^2} = |z| + |w|. \end{aligned}$$

Außerdem gilt

$$\frac{z + \overline{z}}{2} = \frac{(3 + 4i) + (3 - 4i)}{2} = \frac{6}{2} = 3 = \mathrm{Re}(z).$$

B) Geometrische Interpretation der Arithmetik der komplexen Zahlen

Bemerkung 10.9 (Geometrische Deutung und Polarkoordinaten)

Wir wollen hier einige der bisher eingeführten Operationen auf den komplexen Zahlen und der angeführten Eigenschaften derselben geometrisch interpretieren.

- Die Addition ist einfach die komponentenweise Addition, also die Addition der Vektoren (siehe Abbildung 3).

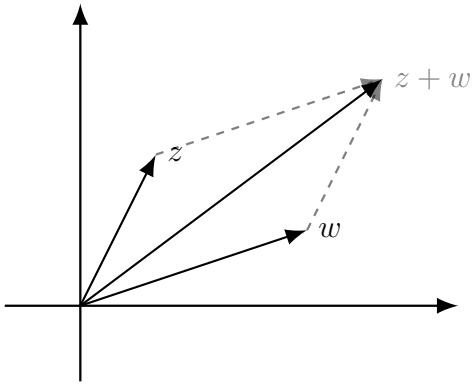


Abbildung 3: Addition in $\mathbb{C}$ als Vektoraddition

- Die komplexe Konjugation ist die Spiegelung an der x-Achse.
- Der Realteil ist die orthogonale Projektion auf die x-Achse und der Imaginärteil die orthogonale Projektion auf die y-Achse.
- Der Betrag $|z| = \sqrt{x^2 + y^2}$ einer komplexen Zahl $z = x + iy$ ist die euklidische Länge des Vektors z, d.h. der Abstand von z zum Ursprung. Dies ergibt sich unmittelbar aus dem Satz von Pythagoras (siehe Abbildung 4).

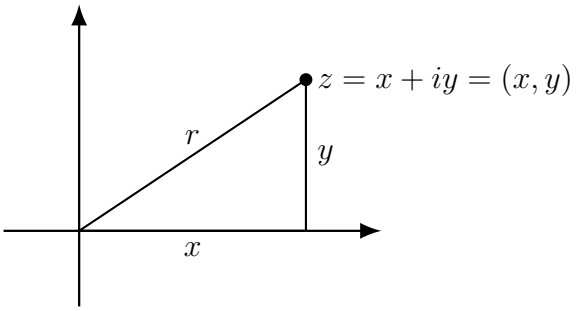


Abbildung 4: Pythagoras: $x^2 + y^2 = r^2$

- 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.
- Die Menge

$$K := \{z \in \mathbb{C} \mid |z| = 1\}$$

der Punkte in der Ebene, deren Abstand zum Ursprung genau 1 ist, ist der Einheitskreis um den Ursprung. Man beachte, daß bei einem Punkt

$$z = x + iy,$$

der auf dem Einheitskreis liegt, die kartesichen Koordinaten x und y schon vollständig durch den Winkel $\alpha \in [0, 2\pi)$ bestimmt sind, den der Vektor z mit der x-Achse einschließt. Es gilt nämlich (siehe Abbildung 5)

$$x = \cos(\alpha)$$

und

$$y = \sin(\alpha)$$

und somit

$$z = \cos(\alpha) + i \cdot \sin(\alpha).$$

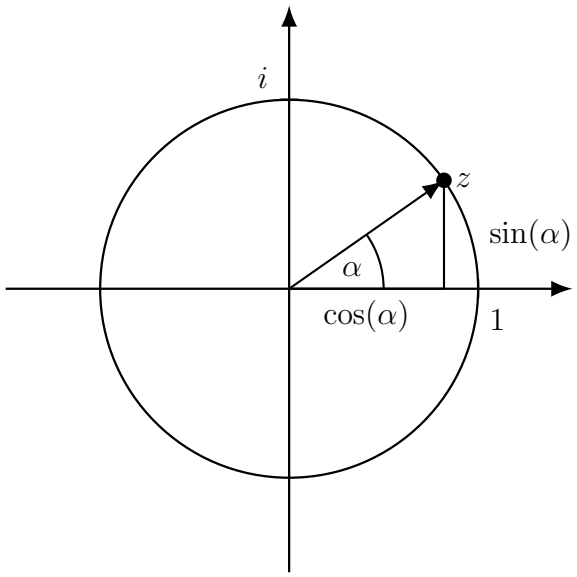


Abbildung 5: Koordinaten eines Punktes $z = \cos(\alpha) + i \cdot \sin(\alpha)$ auf dem Einheitskreis

- Es bleibt, die Multiplikation zweier komplexer Zahlen $0 \neq z, w \in \mathbb{C}$ geometrisch zu deuten. Dazu schreiben wir die Zahl z als

$$z = |z| \cdot \frac{z}{|z|} = r \cdot z'$$

mit $r = |z|$ und $z' = \frac{z}{|z|}$. Man beachte, daß die Zahl z' den Betrag 1 hat, so daß es genau einen Winkel $\alpha \in [0, 2\pi)$ gibt mit

$$z' = (\cos(\alpha), \sin(\alpha)) = \cos(\alpha) + i \cdot \sin(\alpha).$$

Die komplexe Zahl $z \neq 0$ ist also eindeutig durch ihren Betrag und den Winkel α bestimmt. Wir nennen

$$\arg(z) := \alpha$$

das *Argument* von z und das Paar

$$(r, \alpha) = \big(|z|, \arg(z)\big)$$

die *Polarkoordinaten* von z.

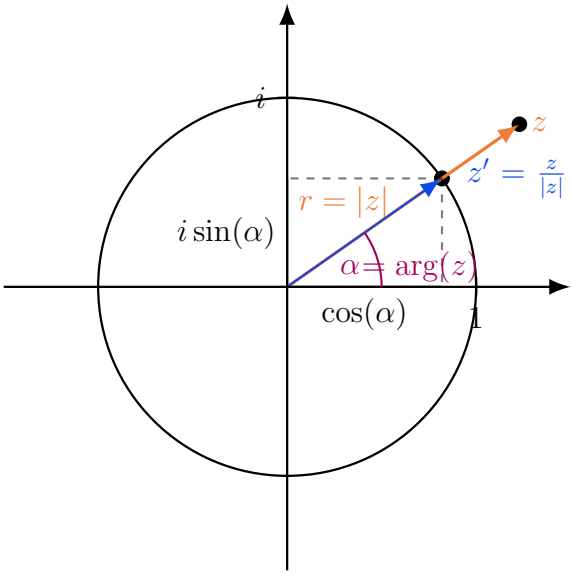


Abbildung 6: Polarkoordinaten von $z = r \cdot (\cos(\alpha) + i \cdot \sin(\alpha))$

Wir erinnern hier an die beiden Additionstheoreme für den Sinus und den Cosinus (siehe auch Satz 12.38):

$$\cos(\alpha + \beta) = \cos(\alpha) \cdot \cos(\beta) - \sin(\alpha) \cdot \sin(\beta) \tag{12}$$

und

$$\sin(\alpha + \beta) = \cos(\alpha) \cdot \sin(\beta) + \sin(\alpha) \cdot \cos(\beta). \tag{13}$$

Betrachten wir zunächst die Multiplikation von zwei komplexen Zahlen $z = |z| \cdot \big(\cos(\alpha) + i \cdot \sin(\alpha)\big)$ und $w = |w| \cdot \big(\cos(\beta) + i \cdot \sin(\beta)\big)$:

$$\begin{aligned} z \cdot w &= |z| \cdot |w| \cdot \big(\cos(\alpha) + i \cdot \sin(\alpha)\big) \cdot \big(\cos(\beta) + i \cdot \sin(\beta)\big) \\ &= |z| \cdot |w| \cdot \big(\cos(\alpha) \cdot \cos(\beta) - \sin(\alpha) \cdot \sin(\beta)\big) + i \cdot \big(\cos(\alpha) \cdot \sin(\beta) + \sin(\alpha) \cdot \cos(\beta)\big) \\ &\overset{(12),(13)}{=} |z| \cdot |w| \cdot \big(\cos(\alpha + \beta) + i \cdot \sin(\alpha + \beta)\big). \end{aligned}$$

Die beiden Zahlen werden also multipliziert, indem man die Argumente addiert und die Beträge multipliziert (siehe Abbildung 7).

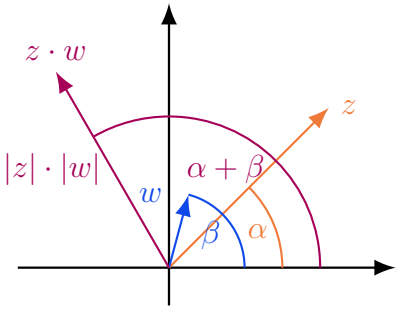


Abbildung 7: Multiplikation zweier komplexer Zahlen

In Polarkoordinaten könnte man dies schreiben als

$$(r, \alpha) \cdot (s, \beta) = (r \cdot s, \alpha + \beta).$$

Beispiel 10.10

Zur Ermittlung von $\alpha = \arg(z)$ für $z = i - 1$ betrachten wir die Zahl

$$\frac{z}{|z|} = -\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2}i$$

vom Betrag 1, für die gilt

$$\frac{z}{|z|} = \cos(\alpha) + i\sin(\alpha),$$

d.h. $\cos(\alpha) = -\frac{\sqrt{2}}{2}$ und $\sin(\alpha) = \frac{\sqrt{2}}{2}$, also $\alpha = \frac{3}{4}\pi$.

Bemerkung 10.11 (n-te Wurzeln)

Aus der Polarkoordinatendarstellung einer komplexen Zahl

$$w = r \cdot \big(\cos(\alpha) + i \cdot \sin(\alpha)\big)$$

läßt sich leicht ableiten, daß die Zahl

$$a = \sqrt[n]{r} \cdot \left(\cos\left(\tfrac{\alpha}{n}\right) + i \cdot \sin\left(\tfrac{\alpha}{n}\right)\right)$$

eine n-te Wurzel aus w ist, d.h.

$$a^n = w.$$

Dabei ist $\sqrt[n]{r}$ die eindeutig bestimmte nicht-negative n-te Wurzel der nicht-negativen Zahl r.

Die obige Zahl a ist aber nicht die einzige Lösung der Gleichung

$$z^n = w$$

in $\mathbb{C}$. Denn addiert man zum Argument einen der folgenden Winkel

$$\frac{2\pi k}{n}, \quad \text{mit } k = 1, \ldots, n-1,$$

so erhalten wir

$$\begin{aligned}\left(\sqrt[n]{r} \cdot \left(\cos\left(\tfrac{\alpha+2\pi k}{n}\right) + i \cdot \sin\left(\tfrac{\alpha+2\pi k}{n}\right)\right)\right)^n &= \sqrt[n]{r}^n \cdot \left(\cos\left(\alpha + 2\pi k\right) + i \cdot \sin\left(\alpha + 2\pi k\right)\right) \\ &= \sqrt[n]{r}^n \cdot \left(\cos\left(\alpha\right) + i \cdot \sin\left(\alpha\right)\right) = w.\end{aligned}$$

Wir haben also in der Tat n verschiedene n-te Wurzeln von w gefunden:

$$a_k = \sqrt[n]{r} \cdot \left(\cos\left(\tfrac{\alpha+2\pi\cdot k}{n}\right) + i \cdot \sin\left(\tfrac{\alpha+2\pi\cdot k}{n}\right)\right), \quad k = 0, \ldots, n-1.$$

Damit sehen wir, daß die Polynomgleichung

$$z^n = 1$$

in $\mathbb{C}$ genau n Lösungen hat, wobei n der Grad der Gleichung ist. Das ist ein Spezialfall des Fundamentalsatzes der Algebra.

Aufgaben

Aufgabe 10.12

Bestimme für die folgenden komplexen Zahlen $\operatorname{Re} z$, $\operatorname{Im} z$, $\arg z$, $|z|$, $\overline{z}$ und z^{-1}:

a. $z = i - 1$.

b. $z = \frac{4i}{1+i}$.

c. $z = \frac{(2+2i)^7}{(1-i)^3}$.

d. $z = \frac{4+2i}{2-2i}$.

Aufgabe 10.13

Berechne die komplexe Zahl $\left(\frac{1}{\sqrt{2}} + \frac{i}{\sqrt{2}}\right)^{4n}$ für jede natürliche Zahl $n \in \mathbb{N}$.

Aufgabe 10.14

Es sei $z \in \mathbb{C}$ eine komplexe Zahl mit $\operatorname{Im}(z) \neq 0$ und

$$\frac{2 + 3z + 4z^2}{2 - 3z + 4z^2} \in \mathbb{R}.$$

Bestimme $|z|^2$.

Aufgabe 10.15

Wie viele Lösungen $z \in \mathbb{C}$ hat die Gleichung

$$\overline{z} - z^2 = i \cdot \left(\overline{z} + z^2\right)?$$

Bestimme alle Lösungen.

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 $a_n := \alpha(n)$ mit $n \in \mathbb{N}$. Wir schreiben deshalb statt $\alpha : \mathbb{N} \longrightarrow \mathbb{K}$ gemeinhin nur $(a_n)_{n\in\mathbb{N}}$ oder $(a_0, a_1, a_2, \ldots)$.

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 $(a_n)_{n\geq k}$. Formal würde dem dann die Abbildung

$$\mathbb{N} \longrightarrow \mathbb{K} : n \mapsto a_{n+k}$$

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 $a_n = c$ für $n \in \mathbb{N}$, und mithin $(a_n)_{n\in\mathbb{N}} = (c)_{n\in\mathbb{N}}$. Z.B. $(a_n)_{n\in\mathbb{N}} = (2, 2, 2, 2, 2, \ldots)$.

b. Für $q \in \mathbb{K}$ ist auch $\alpha : \mathbb{N} \longrightarrow \mathbb{K} : n \mapsto q^n$ eine Folge mit $a_n = q^n$, also $(a_n)_{n\in\mathbb{N}} = (q^n)_{n\in\mathbb{N}}$. Z.B. $(a_n)_{n\in\mathbb{N}} = (1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots)$.

c. $\left(\frac{1}{n}\right)_{n\geq 1}$ 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 $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $a \in \mathbb{K}$.

a. Wir nennen a genau dann einen *Grenzwert* von $(a_n)_{n\in\mathbb{N}}$, wenn

$$\forall\, 0 < \varepsilon \in \mathbb{R} \quad \exists\, n_\varepsilon \in \mathbb{N} \;:\; \forall\, n \geq n_\varepsilon \;:\; |a_n - a| < \varepsilon.$$

In diesem Fall sagen wir auch, daß $(a_n)_{n\in\mathbb{N}}$ *gegen a konvergiert* und schreiben

$$\lim_{n\to\infty} a_n = a$$

oder

$$a_n \longrightarrow a.$$

b. Wir nennen $(a_n)_{n\in\mathbb{N}}$ genau dann *konvergent*, wenn es ein $a \in \mathbb{K}$ gibt, so daß $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert. Andernfalls nennen wir $(a_n)_{n\in\mathbb{N}}$ *divergent*.

c. Wir nennen eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{K}$ eine *Nullfolge*, wenn $(a_n)_{n\in\mathbb{N}}$ gegen 0 konvergiert, d.h. $a_n \longrightarrow 0$.

Beispiel 11.5

a. Die konstante Folge $(a_n)_{n\in\mathbb{N}} = (c)_{n\in\mathbb{N}}$ konvergiert gegen c, d.h. $\lim\limits_{n\to\infty} c = c$.
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

$$|a_n - c| = |c - c| = 0 < \varepsilon.$$

b. Die Folge $\left(\frac{1}{n}\right)_{n\geq 1}$ konvergiert gegen 0, d.h. $\lim\limits_{n\to\infty} \frac{1}{n} = 0$.
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 $(a_n)_{n\in\mathbb{N}} = \left((-1)^n\right)_{n\in\mathbb{N}}$ ist divergent.
Denn: nehmen wir an, $(a_n)_{n\in\mathbb{N}}$ konvergiert gegen a. Dann gibt es zu $\varepsilon = \frac{1}{2}$ ein n_ε, so

daß $|a_n - a| < \varepsilon$ für $n \geq n_\varepsilon$. Insbesondere gilt dann wegen der Dreiecksungleichung

$$2 = |(-1)^{n_\varepsilon} - (-1)^{n_\varepsilon+1}| = |a_{n_\varepsilon} - a_{n_\varepsilon+1}| \leq |a_{n_\varepsilon} - a| + |a - a_{n_\varepsilon+1}| < \varepsilon + \varepsilon = 1,$$

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

$$\frac{1}{n_\varepsilon} < x \cdot \varepsilon \tag{14}$$

Ist nun $n \geq n_\varepsilon$, so gilt wegen $|q| = \frac{1}{1+x}$ und der Bernoullischen Ungleichung auch

$$|q^n - 0| = |q|^n = \frac{1}{(1+x)^n} \overset{9.7}{\leq} \frac{1}{1+n\cdot x} < \frac{1}{n\cdot x} \leq \frac{1}{n_\varepsilon \cdot x} \overset{(14)}{<} \frac{x\cdot\varepsilon}{x} = \varepsilon.$$

□

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 $\left(\frac{1}{n}\right)_{n\geq 1}$ ist beschränkt, da $\left|\frac{1}{n}\right| \leq 1$ für alle $n \geq 1$.

b. Die divergente Folge $(a_n)_{n\in\mathbb{N}} = \left((-1)^n\right)_{n\in\mathbb{N}}$ ist ebenfalls beschränkt, da $\{|a_n| \in \mathbb{R} \mid n \in \mathbb{N}\} = \{1\}$.

Satz 11.11 (Konvergente Folgen sind beschränkt.)
Jede konvergente Folge in $\mathbb{K}$ ist beschränkt.

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $\lim\limits_{n\to\infty} a_n = a$. Dann gibt es zu $\varepsilon = 1$ eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$|a_n - a| < \varepsilon = 1$$

für $n \geq n_\varepsilon$. Setze

$$s := \max\{1 + |a|, |a_0|, |a_1|, \ldots, |a_{n_\varepsilon - 1}|\},$$

wobei man beachte, daß das Maximum existiert, weil die Menge endlich ist.

Damit erhalten wir dann

$$|a_n| \leq \begin{cases} s, & \text{falls } n < n_\varepsilon, \\ |a_n - a| + |a| < 1 + |a| \leq s, & \text{falls } n \geq n_\varepsilon. \end{cases}$$

Mithin ist s eine Schranke für $(a_n)_{n\in\mathbb{N}}$. □

Beispiel 11.12

a. Beispiel 11.10 zeigt, daß die Umkehrung von Satz 11.11 nicht gilt.

b. Für $k \in \mathbb{N}$ mit $k \geq 1$ ist die Folge $(a_n)_{n\in\mathbb{N}} = (n^k)_{n\in\mathbb{N}}$ nicht beschränkt, also auch nicht konvergent.

c. Für $q \in \mathbb{K}$ mit $|q| > 1$ ist die Folge $(a_n)_{n\in\mathbb{N}} = (q^n)_{n\in\mathbb{N}}$ nicht beschränkt und somit divergent.
Um dies zu sehen, nehmen wir an, $s > 0$ sei eine Schranke für die Folge $(a_n)_{n\in\mathbb{N}}$ und setzen $x := |q| - 1 > 0$. Da $\mathbb{R}$ archimedisch angeordnet ist (siehe Satz 9.3), gibt es eine natürliche Zahl $n \in \mathbb{N}$, so daß

$$s < n \cdot x.$$

Aus der Bernoullischen Ungleichung erhalten wir damit

$$|q|^n = (1 + x)^n \geq 1 + n \cdot x > s,$$

was ein Widerspruch zur Wahl von s als Schranke von $(a_n)_{n\in\mathbb{N}}$ ist. Dies zeigt, daß $(a_n)_{n\in\mathbb{N}}$ nicht beschränkt ist.

Lemma 11.13
Ist $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge in $\mathbb{K}$ und $(b_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{K}$, so ist $(a_n \cdot b_n)_{n\in\mathbb{N}}$ eine Nullfolge.

Beweis: Da $(b_n)_{n\in\mathbb{N}}$ beschränkt ist, gibt es eine positive reelle Zahl $s \in \mathbb{R}_{>0}$ mit

$$|b_n| \leq s$$

für alle $n \in \mathbb{N}$.

Sei nun $\varepsilon > 0$ gegeben. Wegen $a_n \longrightarrow 0$ gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$|a_n - 0| < \frac{\varepsilon}{s}$$

für $n \geq n_\varepsilon$. Für $n \geq n_\varepsilon$ erhalten wir damit

$$|a_n \cdot b_n - 0| = |a_n \cdot b_n - 0 \cdot b_n| = |a_n - 0| \cdot |b_n| \leq |a_n - 0| \cdot s < \frac{\varepsilon}{s} \cdot s = \varepsilon.$$

Mithin konvergiert $(a_n \cdot b_n)_{n\in\mathbb{N}}$ gegen 0. □

Beispiel 11.14
Da die Folge $\left(\frac{1}{n}\right)_{n\geq 1}$ eine Nullfolge ist und da zudem die Folge $\left((-1)^n\right)_{n\geq 1}$ beschränkt ist, gilt

$$\frac{(-1)^n}{n} \longrightarrow 0.$$

C) Grenzwertsätze

Proposition 11.15 (Grenzwertsätze)
Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergente Folgen in $\mathbb{K}$ mit $a_n \longrightarrow a$ und $b_n \longrightarrow b$.

a. $a_n + b_n \longrightarrow a + b$ und $a_n - b_n \longrightarrow a - b$.

b. $a_n \cdot b_n \longrightarrow a \cdot b$.

c. $|a_n| \longrightarrow |a|$.

d. Ist zudem $b \neq 0$, so gibt es ein $n_0 \in \mathbb{N}$ mit $b_n \neq 0$ für alle $n \geq n_0$ und die Folge $\left(\frac{a_n}{b_n}\right)_{n\geq n_0}$ ist konvergent mit

$$\frac{a_n}{b_n} \longrightarrow \frac{a}{b}.$$

Beweis:

a. Sei $\varepsilon > 0$ gegeben. Dann gibt es natürliche Zahlen $n'_\varepsilon, n''_\varepsilon \in \mathbb{N}$, so daß

$$|a_n - a| < \frac{\varepsilon}{2}$$

für $n \geq n'_\varepsilon$ und

$$|b_n - b| < \frac{\varepsilon}{2}$$

für $n \geq n''_\varepsilon$. Mit $n_\varepsilon := \max\{n'_\varepsilon, n''_\varepsilon\}$ gilt dann

$$|(a_n + b_n) - (a + b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

für $n \geq n_\varepsilon$. Mithin konvergiert $(a_n + b_n)_{n\in\mathbb{N}}$ gegen $a + b$. Analog sieht man $a_n - b_n \longrightarrow a - b$.

b. Nach Satz 11.11 ist $(a_n)_{n\in\mathbb{N}}$ als konvergente Folge beschränkt und nach Voraussetzung ist $(b_n - b)_{n\in\mathbb{N}}$ eine Nullfolge, so daß

$$a_n \cdot (b_n - b) \longrightarrow 0$$

nach Lemma 11.13. Analog ist nach Voraussetzung $(a_n - a)_{n\in\mathbb{N}}$ eine Nullfolge und die konstante Folge $(b)_{n\in\mathbb{N}}$ ist als konvergente Folge beschränkt, so daß

$$(a_n - a) \cdot b \longrightarrow 0.$$

Aus a. folgt dann, daß die Summe der beiden Nullfolgen eine Nullfolge ist, d.h.

$$a_n \cdot b_n - a \cdot b = a_n \cdot (b_n - b) + (a_n - a) \cdot b \longrightarrow 0 + 0 = 0.$$

Also gilt auch $a_n \cdot b_n \longrightarrow a \cdot b$.

c. Ist $\varepsilon > 0$ gegeben, so gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$ mit

$$|a_n - a| < \varepsilon$$

für alle $n \geq n_\varepsilon$. Aber dann gilt nach Lemma 10.7 auch

$$\big||a_n| - |a|\big| \leq |a_n - a| < \varepsilon$$

für alle $n \geq n_\varepsilon$. Es folgt die Behauptung.

d. Wegen $b \neq 0$ ist

$$\epsilon := \frac{|b|}{2} > 0$$

und es gibt ein $n_0 \in \mathbb{N}$ mit

$$|b_n - b| < \epsilon = \frac{|b|}{2}$$

für alle $n \geq n_0$. Mithin ist

$$|b| = |b_n + (b - b_n)| \leq |b_n| + |b - b_n| = |b_n| + |b_n - b| < |b_n| + \frac{|b|}{2},$$

so daß $0 < \frac{|b|}{2} \leq |b_n|$ für $n \geq n_0$. Insbesondere ist $b_n \neq 0$ in diesen Fällen.
Aus Lemma 8.14 folgt zudem

$$0 < \frac{1}{|b_n|} \leq \frac{2}{|b|} \tag{15}$$

für $n \geq n_0$.
Ist nun $\varepsilon > 0$ beliebig gegeben, so gibt es eine natürliche Zahl $n_\varepsilon \geq n_0$ mit

$$|b_n - b| < \frac{\varepsilon \cdot |b|^2}{2} \tag{16}$$

für alle $n \geq n_\varepsilon$. Für diese n erhalten wir damit

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b - b_n|}{|b_n \cdot b|} = \frac{1}{|b_n|} \cdot \frac{1}{|b|} \cdot |b_n - b| \overset{(15),(16)}{<} \frac{2}{|b|^2} \cdot \frac{\varepsilon \cdot |b|^2}{2} = \varepsilon.$$

Also gilt

$$\frac{1}{b_n} \longrightarrow \frac{1}{b},$$

und mit Teil b. folgt dann die Behauptung $\frac{a_n}{b_n} \longrightarrow \frac{a}{b}$.

□

Beispiel 11.16

a. Die Folge $\left(\frac{1}{n^2}\right)_{n\geq 1}$ ist eine Nullfolge, da $\lim\limits_{n\to\infty} \frac{1}{n^2} = \lim\limits_{n\to\infty} \frac{1}{n} \cdot \lim\limits_{n\to\infty} \frac{1}{n} = 0 \cdot 0 = 0$.

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

$$a_n = \frac{7 + \frac{3}{n^2}}{4 + \frac{1}{n} + \frac{1}{n^2}} \longrightarrow \frac{7 + 0}{4 + 0 + 0} = \frac{7}{4}.$$

D) Konvergenzkriterien für reelle Zahlenfolgen

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.

a. Ist $a_n \leq b_n$ für alle $n \geq n_0$, so ist $a \leq b$.

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

$$a - b < 2 \cdot \varepsilon = a - b. \quad \lightning$$

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 $0 \longrightarrow 0$ und $\frac{1}{n} \longrightarrow 0$ auch

$$\frac{1}{n^k} \longrightarrow 0.$$

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 $\left(\frac{1}{n}\right)_{n\geq 1}$ 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 $a_n \longrightarrow a$ gezeigt. Der Fall einer monoton fallenden Folge wird analog mit Hilfe des Infimums bewiesen. □

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. □

Beispiel 11.23 (Rekursive Folgen — das Heron-Verfahren)
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: $a_{n+1}^2 \geq c$ 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 9.8.

Beachte, daß man die Folge $(a_n)_{n\in\mathbb{N}}$ nutzen kann, um die Wurzel $\sqrt{c}$ näherungsweise zu berechnen — man nennt dieses rekursive Verfahren auch das *Heron-Verfahren*. Versucht dies einmal für $c = 2$ oder $c = 4$.

E) 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

$$n_0 < n_1 < n_2 < n_3 < \ldots$$

eine aufsteigende Folge natürlicher Zahlen, so nennen wir die Folge

$$(a_{n_k})_{k\in\mathbb{N}} = (a_{n_0}, a_{n_1}, a_{n_2}, a_{n_3}, \ldots)$$

eine *Teilfolge* von $(a_n)_{n\in\mathbb{N}}$.

Beispiel 11.25
Die Folge $\left(\frac{1}{n^2}\right)_{n\geq 1}$ ist eine Teilfolge von $\left(\frac{1}{n}\right)_{n\geq 1}$.

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

$$[b_0, c_0] \supsetneqq [b_1, c_1] \supsetneqq [b_2, c_2] \supsetneqq [b_3, c_3] \supsetneqq \dots,$$

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

$$b_j \le a_{n_j} \le c_j$$

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

$$c - b \longleftarrow c_n - b_n = \frac{2s}{2^n} = 2s \cdot \left(\frac{1}{2}\right)^n \longrightarrow 0,$$

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

$$\lim_{j\to\infty} a_{n_j} = b.$$

2. Fall: $\mathbb{K} = \mathbb{C}$**:** Aus Lemma 10.7 wissen wir, daß

$$|\operatorname{Re}(a_n)| \le |a_n|,$$

so daß die Folge $\big(\operatorname{Re}(a_n)\big)_{n\in\mathbb{N}}$ 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 10.7 folgt

$$|\operatorname{Im}(a_{n_k})| \leq |a_{n_k}|,$$

so daß auch die Folge $\big(\operatorname{Im}(a_{n_k})\big)_{k\in\mathbb{N}}$ 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 $\big(\operatorname{Re}(a_{n_{k_j}})\big)_{j\in\mathbb{N}}$ von $\big(\operatorname{Re}(a_{n_k})\big)_{k\in\mathbb{N}}$ 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 $\big((-1)^n\big)_{n\in\mathbb{N}}$ besitzt als konvergente Teilfolge die konstante Folge $\big((-1)^{2k}\big)_{k\in\mathbb{N}} = (1)_{k\in\mathbb{N}}$.

Satz 11.28 (Abgeschlossene Intervalle sind abgeschlossen.)

Ist $(a_n)_{n\in\mathbb{N}}$ eine konvergente Folge im abgeschlossenen Intervall $[a, b]$, so gilt

$$\lim_{n\to\infty} a_n \in [a, b].$$

Beweis: Nehmen wir an, $c := \lim\limits_{n\to\infty} a_n \in \mathbb{R} \setminus [a, b]$. 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. □

F) 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:

$\Longrightarrow$**:** 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.

$\Longleftarrow$**:** 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

   $$|a_m - a_n| < \varepsilon = 1$$

   für alle $m > n \geq n_\varepsilon$. Setze

   $$s := \max\{1 + |a_{n_\varepsilon}|, |a_0|, |a_1|, \ldots, |a_{n_\varepsilon - 1}|\}.$$

   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

   $$a := \lim_{k\to\infty} a_{n_k}.$$

3. **Schritt:** Zeige, $a_n \longrightarrow a$.

   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 $a_{n_k} \longrightarrow a$ existiert auch ein $k_\varepsilon \in \mathbb{N}$ mit

   $$|a_{n_k} - a| < \frac{\varepsilon}{2}$$

[1] Der Beweis geht wie der Beweis von Satz 11.11, wenn man dort den Grenzwert a durch a_{n_ε} ersetzt.

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

$$|q^m - q^n| = |q|^n \cdot |q-1| = 1^n \cdot \varepsilon = \varepsilon.$$

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 9.10) 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. 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).

G) 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 $a_n \longrightarrow \infty$ oder $\lim\limits_{n\to\infty} a_n = \infty$, 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 $a_n \longrightarrow -\infty$ oder $\lim\limits_{n\to\infty} a_n = -\infty$, 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 $\big((-1)^n \cdot n\big)_{n\in\mathbb{N}}$ 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 $a - \infty := -\infty$.
- $a \cdot \infty := \infty$ und $a \cdot -\infty := -\infty$, falls $a > 0$.
- $a \cdot \infty := -\infty$ und $a \cdot -\infty := \infty$, falls $a < 0$.
- $\frac{a}{\infty} := 0$ und $\frac{a}{-\infty} := 0$.

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 $a_n \longrightarrow a$ und $b_n \longrightarrow \infty$, so gilt $a_n + b_n \longrightarrow a + \infty = \infty$.

b. Wenn $a_n \longrightarrow a$ und $b_n \longrightarrow \infty$, so gilt $\frac{a_n}{b_n} \longrightarrow \frac{a}{\infty} = 0$.

Zudem kann man die Grenzwertsätze auch für Brüche von Folgen formulieren, wenn im Nenner eine Nullfolge steht. Allerdings ist dabei etwas Vorsicht geboten:

a. Wenn $a_n \longrightarrow a \neq 0$ und $b_n \longrightarrow 0$ mit $b_n > 0$ für $n \geq n_0$, so gilt $\frac{a_n}{b_n} \longrightarrow \infty \cdot a$.

b. Wenn $a_n \longrightarrow a \neq 0$ und $b_n \longrightarrow 0$ mit $b_n < 0$ für $n \geq n_0$, so gilt $\frac{a_n}{b_n} \longrightarrow -\infty \cdot a$.

c. Ist das Vorzeichen der b_n nicht ab einer gewissen Stelle fest, so existiert $\lim\limits_{n\to\infty} \frac{a_n}{b_n}$ nicht.

Aufgaben

Aufgabe 11.36

Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge komplexer Zahlen. Zeige, daß die folgenden Aussagen äquivalent sind:

a. $a_n \longrightarrow a$.

b. $\mathrm{Re}(a_n) \longrightarrow \mathrm{Re}(a)$ und $\mathrm{Im}(a_n) \longrightarrow \mathrm{Im}(a)$.

Aufgabe 11.37

Ist $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$ bijektiv, so nennen wir die Folge

$$(a_{\sigma(n)})_{n\in\mathbb{N}} = (a_{\sigma(0)}, a_{\sigma(1)}, a_{\sigma(2)}, a_{\sigma(3)}, \ldots)$$

eine *Umordnung* von $(a_n)_{n\in\mathbb{N}}$. Beweise die folgenden beiden Aussagen.

a. Wenn $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert, so konvergiert jede Teilfolge von $(a_n)_{n\in\mathbb{N}}$ gegen a.

b. Wenn $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert, so konvergiert jede Umordnung von $(a_n)_{n\in\mathbb{N}}$ gegen a.

Aufgabe 11.38

a. Zeige, daß die Folge $(s_n)_{n\in\mathbb{N}}$ mit

$$s_n := \sum_{k=0}^{n} \frac{1}{k!}$$

konvergent ist.

b. Zeige, daß die Folge $(t_n)_{n\in\mathbb{N}}$ mit

$$t_n := \left(1 + \frac{1}{n}\right)^n$$

konvergent ist.

c. 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 $(a_n)_{n\geq 1}$ auf Konvergenz und berechne gegebenenfalls den Grenzwert:

a. $a_n = \frac{n^4-3n+5}{3n^5+6n^3+11}$.

b. $a_n = \frac{(3n+1)\cdot(n+1)^2-5(n-1)}{1+n+n^2}$.

c. $a_n = \sum_{i=1}^{n} \frac{i}{n^2}$.

d. $a_n = \frac{n}{n^2+1} + \ldots + \frac{n}{n^2+n}$.

e. $a_n = \frac{n}{2^n}$.

Aufgabe 11.40

Zeige, dass die rekursiv definierte Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_0 = 1$ und

$$a_{n+1} := \sqrt{1 + a_n}$$

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 $(a_n)_{n\in\mathbb{N}}$ mit $a_0 = 1$ und

$$a_{n+1} := 1 + \frac{1}{a_n}$$

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 $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen, so daß $|a_{n+1} - a_n| < \frac{1}{2^n}$ 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 $|a_{n+1} - a_n| < \frac{1}{n}$ voraussetzen? Gib einen Beweis oder ein Gegenbeispiel an.

Aufgabe 11.43

Bestimme für der in der folgenden Tabelle angegebenen Nullfolgen $(a_n)_{n\geq 1}$ und für jedes der angegebenen $\varepsilon > 0$ eine natürliche Zahl n_ε, so daß $|a_n - 0| < \varepsilon$ für alle $n \geq n_\varepsilon$:

	$\varepsilon = \frac{1}{4}$	$\varepsilon = \frac{1}{16}$	$\varepsilon = \frac{1}{32}$
$a_n = \frac{1}{n^2}$			
$a_n = \frac{n^2}{2^n}$			
$a_n = \frac{n+1}{n} - 1$			

Aufgabe 11.44

Zeige mit Hilfe der Definition des Grenzwertes, daß $\lim\limits_{n\to\infty} \frac{n^2+1}{4n^2} = \frac{1}{4}$.

Aufgabe 11.45

Gib zwei divergente Folgen an, deren Produkt konvergiert. Begründe Deine Antwort.

Aufgabe 11.46

Untersuche die fogenden Folgen $(a_n)_{n\geq 1}$, ob sie konvergent oder divergent sind und bestimme ggf. ihren Grenzwert:

a. $a_n = \frac{n+(-1)^n}{n+1}$

b. $a_n = \frac{(n-1)^3}{n^3+1}$

c. $a_n = \frac{3^{n+1}+2^{n+1}}{3^n+2^n}$

d. $a_n = (-1)^n \cdot \frac{1-n}{2+n}$

e. $a_n = \sqrt{n^2+3n} - \sqrt{n^2-n}$

f. $a_n = \frac{n^3-2n+1}{n^2+1}$

g. $a_n = \frac{1+\sqrt{n-1}}{n+1}$

Aufgabe 11.47

Welche der folgenden Folgen $(a_n)_{n\geq 1}$ sind monoton, beschränkt, konvergent oder bestimmt divergent? Es ist keine Begründung erforderlich, es darf aber eine Begründung gegeben werden.

a. $a_n = 5$

b. $a_n = \frac{1}{(-1)^n}$

c. $a_n = \frac{(-1)^n}{n^2}$

d. $a_n = \frac{3}{n}$

e. $a_n = \frac{n}{n+1}$

f. $a_n = \frac{n+1}{n}$

g. $a_n = -n$

h. $a_n = n^2 - n$

i. $a_n = 2^n$

j. $a_n = (-2)^n$

k. $a_1 = 1$ und $a_{n+1} = a_n + \frac{1}{n}$ für $n \geq 1$.

l. $a_1 = 1$ und $a_{n+1} = \frac{n}{n+1} \cdot a_n$ für $n \geq 1$.

Aufgabe 11.48

Die Folge $(a_n)_{n\geq 1}$ sei rekursiv definiert durch $a_1 = 5$ und

$$a_{n+1} = \frac{a_n}{2} + \frac{2}{a_n}.$$

a. Zeige mit Induktion nach n, daß $a_n \geq 2$ für alle $n \geq 1$ gilt.

b. Zeige, die Folge ist monoton fallend.

c. Zeige, die Folge ist konvergent und bestimme den Grenzwert.

§ 12 Unendliche Reihen

A) Konvergenz unendlicher Reihen

Definition 12.1 (Unendliche Reihen)
Ist $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$, so nennen wir die Folge $(s_n)_{n\in\mathbb{N}}$ mit

$$s_n := \sum_{k=0}^{n} a_k$$

der *Partialsummen* von $(a_n)_{n\in\mathbb{N}}$ auch die durch $(a_n)_{n\in\mathbb{N}}$ 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

$$\sum_{n=0}^{\infty} a_n.$$

Beachte, wie stets bei Folgen müssen weder $(a_n)_{n\geq n_0}$ noch $(s_m)_{m\geq n_0} = \sum_{n=n_0}^{\infty} a_n$ (wobei $s_m = \sum_{k=n_0}^{m} a_k$) 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

$$\begin{aligned} s_n &= \sum_{k=1}^{n} \frac{1}{k\cdot(k+1)} = \sum_{k=1}^{n} \frac{1}{k} - \frac{1}{k+1} \\ &= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \ldots + \left(\frac{1}{n-1} - \frac{1}{n}\right) + \left(\frac{1}{n} - \frac{1}{n+1}\right) \\ &= 1 - \frac{1}{n+1} \longrightarrow 1. \end{aligned}$$

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

$$\begin{aligned} s_{n_k} &= \sum_{i=1}^{2^k} \frac{1}{i} \\ &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \ldots + \left(\frac{1}{2^{k-1}+1} + \ldots + \frac{1}{2^k}\right) \\ &\geq 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \ldots + \left(\frac{1}{2^k} + \ldots + \frac{1}{2^k}\right) \\ &= 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \ldots + \frac{1}{2} \\ &= 1 + \frac{k}{2} \longrightarrow \infty, \end{aligned}$$

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 $a \in \mathbb{K}$.

a. $\sum\limits_{n=0}^{\infty} (a_n + b_n) = \sum\limits_{n=0}^{\infty} a_n + \sum\limits_{n=0}^{\infty} b_n.$

b. $\sum\limits_{n=0}^{\infty} (a_n - b_n) = \sum\limits_{n=0}^{\infty} a_n - \sum\limits_{n=0}^{\infty} b_n.$

c. $\sum\limits_{n=0}^{\infty} a \cdot a_n = a \cdot \sum\limits_{n=0}^{\infty} a_n.$

d. $\mathbb{K} = \mathbb{R}$ und $a_n \leq b_n$ für alle $n \in \mathbb{N}$, so ist $\sum\limits_{n=0}^{\infty} a_n \leq \sum\limits_{n=0}^{\infty} b_n.$

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. □

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

$$s_m - s_n = \sum_{k=n+1}^{m} a_k.$$

□

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.

$$\lim_{n\to\infty} \sum_{k=n}^{\infty} a_k = 0.$$

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

$$\left| \sum_{k=n+1}^{\infty} a_k \right| = \lim_{m\to\infty} \left| \sum_{k=n+1}^{m} a_k \right| \leq \varepsilon$$

für alle $n \geq n_\varepsilon$. Also ist die Folge der Restglieder eine Nullfolge. □

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

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} = \sum_{n=0}^{\infty} a_n - \sum_{n=0}^{\infty} a_n = 0.$$

□

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 $\sum\limits_{n=1}^{\infty} \frac{(-1)^n}{n}$ 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 18.37).

Lemma 12.12 (Umklammern in Reihen)

Es sei $\sum_{n=0}^{\infty} a_n$ eine konvergente Reihe in $\mathbb{K}$ und $0 = k_0 < k_1 < k_2 < \ldots$ eine aufsteigende Folge natürlicher Zahlen. Setzen wir

$$b_n := \sum_{k=k_n}^{k_{n+1}-1} a_k = a_{k_n} + a_{k_n+1} + \ldots + a_{k_{n+1}-1}$$

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 $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$ und $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$ bijektiv. Wir nennen die Folge

$$(a_{\sigma(n)})_{n\in\mathbb{N}} = (a_{\sigma(0)}, a_{\sigma(1)}, a_{\sigma(2)}, a_{\sigma(3)}, \ldots)$$

eine *Umordnung* von $(a_n)_{n\in\mathbb{N}}$ und die Reihe

$$\sum_{n=0}^{\infty} a_{\sigma(n)} = a_{\sigma(0)} + a_{\sigma(1)} + a_{\sigma(2)} + a_{\sigma(3)} + \ldots$$

eine *Umordnung* der Reihe $\sum_{n=0}^{\infty} a_n$.

Beispiel 12.14
Betrachten wir folgende Umordnung der alternierenden harmonischen Reihe

$$\left(-\frac{1}{1}+\frac{1}{2}\right)+\frac{1}{4}+\left(-\frac{1}{3}+\frac{1}{6}\right)+\frac{1}{8}+\left(-\frac{1}{5}+\frac{1}{10}\right)+\frac{1}{12}+\left(-\frac{1}{7}+\frac{1}{14}\right)+\frac{1}{16}+\ldots,$$

d.h. in den Klammern sind sukzessive die ungeraden Folgenglieder a_n zusammen jeweils mit dem zugehörigen Folgenglied a_{2n} aufgeführt, und zwischen den Klammerausdrücken 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

$$-\frac{1}{2}+\frac{1}{4}-\frac{1}{6}+\frac{1}{8}-\frac{1}{10}+\frac{1}{12}-\frac{1}{14}+\frac{1}{16}-\ldots = \frac{1}{2}\cdot\sum_{n=1}^{\infty}\frac{(-1)^n}{n}$$

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 $\sum_{n=0}^{\infty} a_n$ in $\mathbb{K}$ heißt *absolut konvergent*, wenn die Reihe ihrer *Absolut*beträge $\sum_{n=0}^{\infty} |a_n|$ konvergiert. Da die Folge der Parialsummen $t_n := \sum_{k=0}^{n} |a_k|$ monoton wächst, ist dies gleichwertig dazu, daß die Folge $(t_n)_{n\in\mathbb{N}}$ 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 $\sum_{n=0}^{\infty} a_n$ 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. □

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 $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$ sei bijektiv.

Wir wollen zunächst zeigen, daß die Reihe $\sum_{n=0}^{\infty} \left(a_n - a_{\sigma(n)}\right)$ 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 $\{0, 1, \ldots, n'_\varepsilon\}$ endlich ist, gibt es eine Zahl $n_\varepsilon \geq n'_\varepsilon$ mit

$$\{0, 1, \ldots, n'_\varepsilon\} \subseteq \{\sigma(0), \sigma(1), \ldots, \sigma(n_\varepsilon)\}.$$

Für $n \geq n_\varepsilon$ heben sich deshalb in der Partialsumme

$$\sum_{k=0}^{n} \left(a_k - a_{\sigma(k)}\right)$$

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

$$m := \max\{\sigma(0), \sigma(1), \ldots, \sigma(n), 0, 1, \ldots, n\} + 1 > n'_\varepsilon,$$

so erhalten wir insgesamt

$$\left| \sum_{k=0}^{n} \left(a_k - a_{\sigma(k)}\right) - 0 \right| \leq \sum_{k=n'_\varepsilon+1}^{m} |a_k| < \varepsilon.$$

Also konvergiert die Reihe $\sum_{n=0}^{\infty} \left(a_n - a_{\sigma(n)}\right)$ gegen Null.

Aus den Grenzwertsätzen für Reihen 12.4 erhalten wir deshalb, daß die Reihe

$$\sum_{n=0}^{\infty} a_{\sigma(n)} = \sum_{n=0}^{\infty} a_n - \sum_{n=0}^{\infty} \left(a_n - a_{\sigma(n)}\right) = \sum_{n=0}^{\infty} a_n$$

konvergent ist mit dem Grenzwert $\sum_{n=0}^{\infty} a_n$.

Wenden wir das Ergebnis auf die konvergente Reihe $\sum_{n=0}^{\infty} |a_n|$ und ihre Umordnung $\sum_{n=0}^{\infty} \left|a_{\sigma(n)}\right|$ an, so folgt auch, daß die Umordnung absolut konvergent ist. □

D) Konvergenzkriterien für absolute Konvergenz

Satz 12.19 (Majorantenkriterium)
Es seien $\sum\limits_{n=0}^{\infty} a_n$ und $\sum\limits_{n=0}^{\infty} b_n$ zwei Reihen in $\mathbb{K}$. Ist $\sum\limits_{n=0}^{\infty} b_n$ absolut konvergent und $|a_n| \leq |b_n|$ für alle $n \geq n_0$, so ist auch $\sum\limits_{n=0}^{\infty} a_n$ absolut konvergent.
Wir nennen $\sum\limits_{n=0}^{\infty} b_n$ dann eine konvergente *Majorante* von $\sum\limits_{n=0}^{\infty} a_n$.

Beweis: Die Folge von Partialsummen $(s_n)_{n\in\mathbb{N}}$ mit

$$s_n := \sum_{k=0}^{n} |a_k|$$

ist beschränkt durch $s_{n_0} + \sum_{n=0}^{\infty} |b_n|$. Also ist die Reihe absolut konvergent. □

Proposition 12.20 (Minorantenkriterium)
Es seien $\sum\limits_{n=0}^{\infty} a_n$ und $\sum\limits_{n=0}^{\infty} b_n$ zwei Reihen in $\mathbb{R}$.
Ist $\sum\limits_{n=0}^{\infty} b_n$ divergent und $a_n \geq b_n \geq 0$ für alle $n \in \mathbb{N}$, so ist $\sum\limits_{n=0}^{\infty} a_n$ divergent.

Wir nennen $\sum_{n=0}^{\infty} b_n$ dann eine divergente *Minorante* von $\sum_{n=0}^{\infty} a_n$.

Beweis: Wegen $b_n \geq 0$ ist die Folge der Partialsummen $(t_n)_{n\in\mathbb{N}}$ mit

$$t_n := \sum_{k=0}^{n} b_k$$

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

$$s_n := \sum_{k=0}^{n} a_k$$

unbeschränkt, wegen $s_n \geq t_n$, und mithin ist sie divergent nach Satz 11.11. □

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 $\mathbb{K}$.

a. 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.

b. 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. □

Satz 12.23 (Quotientenkriterium)
Es sei $\sum_{n=0}^{\infty} a_n$ eine Reihe in $\mathbb{K}$ mit $a_n \neq 0$ für alle $n \geq n_0$.

a. Existiert ein $q < 1$ mit $\left|\frac{a_{n+1}}{a_n}\right| \leq q$ für $n \geq n_0$, so ist $\sum\limits_{n=0}^{\infty} a_n$ absolut konvergent.

b. Ist $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ für alle $n \geq n_0$, so ist $\sum\limits_{n=0}^{\infty} a_n$ 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

$$|a_n| \leq q \cdot |a_{n-1}| \leq q^2 \cdot |a_{n-2}| \leq \ldots \leq q^{n-n_0} \cdot |a_{n_0}|.$$

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 $\sum_{n=0}^{\infty} a_n$.

Ist $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ für alle $n \geq n_0$, so ist $|a_{n+1}| \geq |a_n| \neq 0$ für alle $n \geq n_0$. Mithin ist $(a_n)_{n\in\mathbb{N}}$ keine Nullfolge und die Reihe $\sum_{n=0}^{\infty} a_n$ ist wegen des Nullfolgekriteriums dann divergent. □

Korollar 12.24 (Praktikables Quotienten-/Wurzelkriterium)
Sei $\sum_{n=0}^{\infty} a_n$ eine Reihe in $\mathbb{K}$ mit $a_n \neq 0$ für alle $n \geq n_0$.

a. Falls $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$ oder $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} < 1$, so ist $\sum\limits_{n=0}^{\infty} a_n$ absolut konvergent.

b. Falls $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| > 1$ oder $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} > 1$, so ist $\sum\limits_{n=0}^{\infty} a_n$ divergent.

c. Im Fall $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = 1$ oder $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} = 1$ wird keine Aussage getroffen!

Beweis: Falls $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$ bzw. $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} < 1$, so kann man Satz 12.22 bzw. Satz 12.23 mit

$$q := \frac{1 + \lim\limits_{n\to\infty} \sqrt[n]{|a_n|}}{2} \quad \text{bzw.} \quad q := \frac{1 + \lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|}{2}$$

anwenden.

Falls $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} > 1$ bzw. $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| > 1$, so ist sicher $\sqrt[n]{|a_n|} > 1$ bzw. $\left|\frac{a_{n+1}}{a_n}\right| > 1$ für n hinreichend groß, so daß die Aussage ebenfalls aus Satz 12.22 bzw. Satz 12.23 folgt. □

Bemerkung 12.25

Man beachte, daß die harmonische Reihe $\sum_{n\geq 1} a_n = \sum_{n\geq 1} \frac{1}{n}$ divergent ist, obwohl stets

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{\frac{1}{n+1}}{\frac{1}{n}} = 1 - \frac{1}{n+1} < 1$$

gilt. Aber, es gibt kein $q < 1$ mit

$$\left|\frac{a_{n+1}}{a_n}\right| = 1 - \frac{1}{n+1} < q$$

für alle hinreichend großen n. Das Quotientenkriterium ist deshalb nicht anwendbar. Beachte auch, daß in diesem Fall

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty} 1 - \frac{1}{n+1} = 1$$

gilt.

Beispiel 12.26

Die Reihe $\sum_{n=1}^{\infty} \frac{n^2}{n!}$ ist absolut konvergent, da

$$\left|\frac{\frac{(n+1)^2}{(n+1)!}}{\frac{n^2}{n!}}\right| = \frac{(n+1)^2 \cdot n!}{n^2 \cdot (n+1)!} = \frac{n+1}{n^2} = \frac{1}{n} + \frac{1}{n^2} \longrightarrow 0 + 0 = 0.$$

E) Das Cauchy-Produkt absolut konvergenter Reihen

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 $\sigma : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$, indem wir die Elemente in $\mathbb{N} \times \mathbb{N}$ in der in Abbildung 1 angegebenen Weise durchlaufen.

Aufgrund der Definition von σ gilt für $m \in \mathbb{N}$ offenbar

$$\{\sigma(0), \sigma(1), \ldots, \sigma((m+1)^2 - 1)\} = \{(k,l) \mid 0 \leq k, l \leq m\} \tag{17}$$

Dann definieren wir uns eine Folge $(d_n)_{n\in\mathbb{N}}$ durch

$$d_n := a_k \cdot b_l, \qquad \text{wenn } (k,l) = \sigma(n).$$

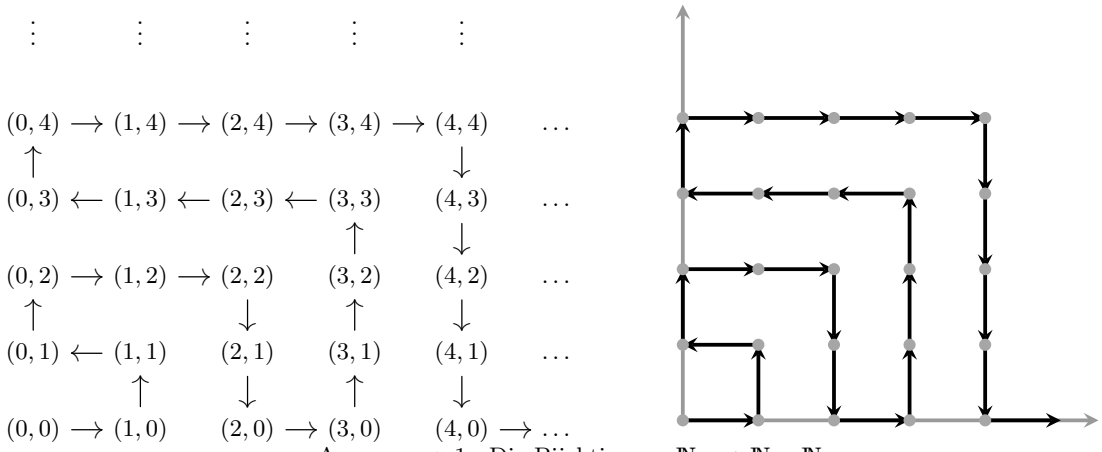


ABBILDUNG 1. Die Bijektion $\sigma : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$

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

$$\sum_{n=0}^{m} |d_n| \leq \sum_{n=0}^{(m+1)^2-1} |d_n| \overset{(17)}{=} \sum_{k=0}^{m} \sum_{l=0}^{m} |a_k \cdot b_l| = \sum_{k=0}^{m} |a_k| \cdot \sum_{l=0}^{m} |b_l| \leq \sum_{k=0}^{\infty} |a_k| \cdot \sum_{l=0}^{\infty} |b_l|$$

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

$$\sum_{i=0}^{\infty} d_i \longleftarrow \sum_{i=0}^{(n+1)^2-1} d_i \overset{(17)}{=} \sum_{k=0}^{n} \sum_{l=0}^{n} a_k \cdot b_l = \sum_{k=0}^{n} a_k \cdot \sum_{l=0}^{n} b_l \longrightarrow \sum_{k=0}^{\infty} a_k \cdot \sum_{l=0}^{\infty} b_l,$$

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 $\pi : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$ wie in Abbildung 2 angedeutet.

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

$$\sum_{n=0}^{\infty} e_n = \sum_{n=0}^{\infty} d_n.$$

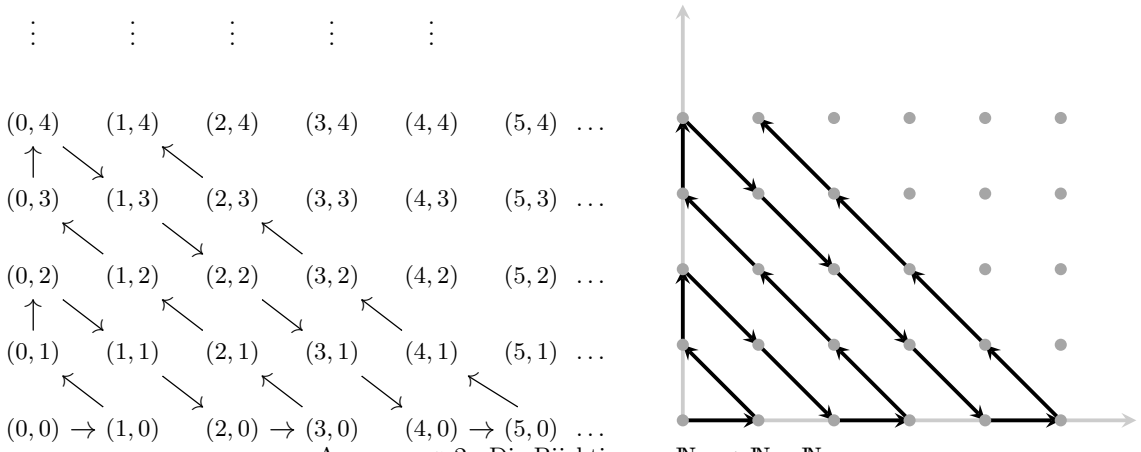


ABBILDUNG 2. Die Bijektion $\pi : \mathbb{N} \longrightarrow \mathbb{N} \times \mathbb{N}$

Nun entsteht die Reihe $\sum_{n=0}^{\infty} c_n$ offenbar aus der Reihe $\sum_{n=0}^{\infty} e_n$ durch Einfügen von Klammern[2] im Sinne von Lemma 12.12. Mithin ist die Reihe nach eben diesem Lemma ebenfalls konvergent mit dem gleichen Grenzwert, d.h.

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} e_n = \sum_{n=0}^{\infty} d_n = \sum_{n=0}^{\infty} a_n \cdot \sum_{n=0}^{\infty} b_n.$$

Aufgrund der Dreiecksungleichung erhalten wir für $m \in \mathbb{N}$ zudem

$$\sum_{n=0}^{m} |c_n| = \sum_{n=0}^{m} \left| \sum_{k+l=n} a_k \cdot b_l \right| \leq \sum_{n=0}^{m} \sum_{k+l=n} |a_k \cdot b_l| \leq \sum_{k=0}^{m} |a_k| \cdot \sum_{l=0}^{m} |b_l| \leq \sum_{k=0}^{\infty} |a_k| \cdot \sum_{l=0}^{\infty} |b_l|,$$

so daß auch die Reihe $\sum_{n=0}^{\infty} |c_n|$ beschränkt und monoton wachsend, also konvergent ist, d.h. $\sum_{n=0}^{\infty} c_n$ ist absolut konvergent. □

[2]Wir 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

$$\{(k,l) \mid k+l=n\} = \left\{ \pi(i) \;\middle|\; \frac{n\cdot(n+1)}{2} \leq i \leq \frac{(n+2)\cdot(n+1)}{2} - 1 \right\}. \tag{18}$$

Für c_n ergibt sich daraus

$$c_n = \sum_{k+l=n} a_k \cdot b_l = \sum_{i=\frac{n\cdot(n+1)}{2}}^{\frac{(n+2)\cdot(n+1)}{2}-1} e_i = e_{\frac{n\cdot(n+1)}{2}} + e_{\frac{n\cdot(n+1)}{2}+1} + \ldots + e_{\frac{(n+2)\cdot(n+1)}{2}-1},$$

d.h. $\sum_{n=0}^{\infty} c_n$ entsteht aus $\sum_{n=0}^{\infty} e_n$ durch Zusammenfassung von Summanden mittels Einfügen von Klammern.

F) 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 $\sum\limits_{n=0}^{\infty} a_n \cdot x^n$ 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. □

Notation 12.30
Wir wollen den Begriff des Supremums etwas erweitern, indem wir $\sup(\emptyset) := -\infty$ 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 für $x \in \mathbb{R}$ bereits die Konvention $\frac{x}{\infty} := \frac{x}{-\infty} := 0$ eingeführt haben. Wir vereinbaren nun zudem $\frac{x}{0} := \infty$ für $x > 0$ sowie $\frac{x}{0} := -\infty$ für $x < 0$.

Definition 12.31

Wir definieren den *Konvergenzradius* der Potenzreihe $\sum_{n=0}^{\infty} a_n \cdot t^n$ über $\mathbb{K}$ als

$$r := \sup \left\{ |y| \;\middle|\; y \in \mathbb{K}, \sum_{n=0}^{\infty} a_n \cdot y^n \text{ ist konvergent} \right\} \in \mathbb{R}_{\geq 0} \cup \{\infty\}.$$

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.

a. Ist $x \in \mathbb{K}$ mit $|x| < r$, so ist $\sum_{n=0}^{\infty} a_n \cdot x^n$ absolut konvergent.

b. 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

a. 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.

b. Konvergenzradius $r = \infty$ heißt, daß $\sum_{n=0}^{\infty} a_n \cdot x^n$ für alle $x \in \mathbb{K}$ absolut konvergent ist.

c. 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.

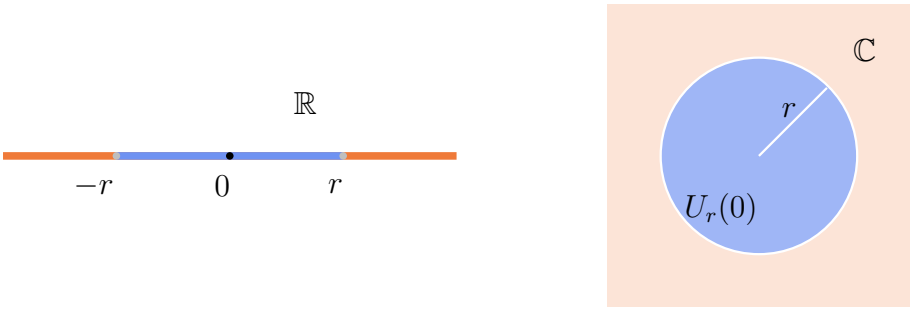


Abbildung 3: Konvergenzbereich $U_r(0)$

Beweis von Satz 12.32: a. Wir betrachten die Menge

$$A := \left\{ |y| \;\middle|\; y \in \mathbb{K}, \sum_{n=0}^{\infty} a_n \cdot y^n \text{ ist konvergent} \right\},$$

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 $r - \varepsilon = \sup(A) - \varepsilon$ 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 $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ existiert, so ist der Konvergenzradius von $\sum_{n=0}^{\infty} a_n \cdot t^n$ gegeben durch

$$r = \frac{1}{\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|}.$$

b. Falls der eigentliche oder uneigentliche Grenzwert $\lim\limits_{n\to\infty} \sqrt[n]{|a_n|} \in \mathbb{R}_{\geq 0} \cup \{\infty\}$ existiert, so ist der Konvergenzradius von $\sum_{n=0}^{\infty} a_n \cdot t^n$ gegeben durch

$$r = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{|a_n|}}.$$

Beweis:

a. Es sei $r = \frac{1}{\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|}$ 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

$$\lim_{n\to\infty} \left|\frac{a_{n+1} \cdot x^{n+1}}{a_n \cdot x^n}\right| = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| \cdot |x| < 1.$$

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 $r = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{1}} = 1$.

Damit wissen wir, daß die Reihe $\sum_{n=0}^{\infty} x^n$ 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

$$r = \frac{1}{\lim\limits_{n\to\infty} \frac{n}{n+1}} = \frac{1}{\lim\limits_{n\to\infty} \left(1 - \frac{1}{n+1}\right)} = 1.$$

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.

G) Exponentialfunktion, Sinus und Cosinus als Potenzreihen

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

$$r = \frac{1}{\lim\limits_{n\to\infty} \frac{n!}{(n+1)!}} = \frac{1}{\lim\limits_{n\to\infty} \frac{1}{n+1}} = \frac{1}{0} = \infty.$$

Aus dem Cauchy-Produkt für Reihen 12.27 und dem Binomischen Lehrsatz 7.15 folgt

$$\begin{aligned} \exp(x) \cdot \exp(y) &= \sum_{n=0}^{\infty} \frac{x^n}{n!} \cdot \sum_{n=0}^{\infty} \frac{y^n}{n!} \overset{12.27}{=} \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{x^k}{k!} \cdot \frac{y^{n-k}}{(n-k)!} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} \cdot x^k \cdot y^{n-k} \overset{7.15}{=} \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} = \exp(x+y). \end{aligned}$$

□

Bemerkung 12.37
Nach Aufgabe 11.38 gilt $\exp(1) = e$ und mit Induktion folgt aus der Funktionalgleichung leicht, daß $\exp(n) = e^n$ und $\exp(\frac{1}{n}) = e^{\frac{1}{n}} = \sqrt[n]{e}$ für $n \geq 2$. Wir setzen deshalb

$$e^x := \exp(x)$$

für $x \in \mathbb{K}$, so daß die neue Notation mit der üblichen Potenzschreibweise und mit der Notation in Satz 9.8 übereinstimmt, und das Potenzgesetz $e^{x+y} = e^x \cdot e^y$ gilt.

Satz 12.38 (Sinus und Cosinus)

a. Die Potenzreihe $\sum_{n=0}^{\infty}(-1)^n \cdot \frac{t^{2n}}{(2n)!}$ 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 $\sum_{n=0}^{\infty}(-1)^n \cdot \frac{t^{2n+1}}{(2n+1)!}$ 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

$$e^{i \cdot x} = \exp(i \cdot x) = \cos(x) + i \cdot \sin(x).$$

e. Für $x \in \mathbb{K}$ gilt

$$\cos(x)^2 + \sin(x)^2 = 1.$$

f. Für $x \in \mathbb{K}$ gilt

$$\cos(x) = \frac{1}{2} \cdot \left(e^{ix} + e^{-ix}\right)$$

und

$$\sin(x) = \frac{1}{2i} \cdot \left(e^{ix} - e^{-ix}\right).$$

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

$$r = \sup\left\{|x| \;\middle|\; x \in \mathbb{K}\right\} = \sup(\mathbb{R}_{\geq 0}) = \infty.$$

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

$$r = \sup\left\{|x| \;\middle|\; x \in \mathbb{K}\right\} = \sup(\mathbb{R}_{\geq 0}) = \infty.$$

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 Lemmas 12.12 zum Umklammern konvergenter Reihen.

e. Für $x \in \mathbb{K}$ gilt

$$\begin{aligned}\cos(x)^2 + \sin(x)^2 &= \big(\cos(x) + i \cdot \sin(x)\big) \cdot \big(\cos(x) - i \cdot \sin(x)\big) \\ &\overset{c.}{=} \big(\cos(x) + i \cdot \sin(x)\big) \cdot \big(\cos(-x) + i \cdot \sin(-x)\big) \\ &\overset{d.}{=} \exp(ix) \cdot \exp(-ix) \\ &\overset{12.36}{=} \exp(ix - ix) = \exp(0) = 1.\end{aligned}$$

f. Für $x \in \mathbb{K}$ gilt

$$e^{ix} + e^{-ix} \overset{d.}{=} \cos(x) + i \cdot \sin(x) + \cos(-x) + i \cdot \sin(-x) \overset{c.}{=} 2 \cdot \cos(x)$$

und

$$e^{ix} - e^{-ix} \overset{d.}{=} \cos(x) + i \cdot \sin(x) - \cos(-x) - i \cdot \sin(-x) \overset{c.}{=} 2 \cdot i \cdot \sin(x).$$

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

$$\begin{aligned}\cos(x+y) + i \cdot \sin(x+y) &\overset{d.}{=} \exp(i \cdot (x+y)) \overset{12.36}{=} \exp(ix) \cdot \exp(iy) \\ &\overset{d.}{=} \big(\cos(x) + i \cdot \sin(x)\big) \cdot \big(\cos(y) + i \cdot \sin(y)\big) \\ &= \big(\cos(x) \cdot \cos(y) - \sin(x) \cdot \sin(y)\big) \\ &\qquad + i \cdot \big(\cos(x) \cdot \sin(y) + \sin(x) \cdot \cos(y)\big).\end{aligned}$$

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

$$\begin{aligned}
\cos(x)^2 + \sin(x)^2 &= 1 + \sum_{n=0}^{\infty}(-1)^{n+1} \cdot \frac{x^{2n+2}}{(2n+2)!} \cdot \left(\sum_{k=0}^{n+1}\binom{2n+2}{2k} - \sum_{k=0}^{n}\binom{2n+2}{2k+1}\right) \\
&= 1 + \sum_{n=0}^{\infty}(-1)^{n+1} \cdot \frac{x^{2n+2}}{(2n+2)!} \cdot \left(\sum_{l=0}^{2n+2}\binom{2n+2}{l} \cdot (-1)^l \cdot 1^{2n-l}\right) \\
&\overset{7.15}{=} 1 + \sum_{n=0}^{\infty}(-1)^{n+1} \cdot \frac{x^{2n+2}}{(2n+2)!} \cdot (1-1)^{2n} \\
&= 1
\end{aligned}$$

unter Berücksichtigung der Grenzwertsätze für Reihen 12.4 sowie des Umordnungssatzes 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 der obigen Bemerkung 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

$$r := \sup \left\{ |y| \;\middle|\; y \in \mathbb{K}, \sum_{n=0}^{\infty} a_n \cdot y^n \text{ ist konvergent} \right\} \in \mathbb{R}_{\geq 0} \cup \{\infty\},$$

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

$$r = \frac{1}{\lim\limits_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right|}$$

bzw. als

$$r = \frac{1}{\lim\limits_{n\to\infty} \sqrt[n]{|a_n|}}.$$

H) Darstellung von Zahlen in Zahlsystemen

Motivation 12.42 (Zahldarstellung)
Wir sind es gewohnt, reelle Zahlen als Dezimalzahlen darzustellen, d.h. eine Zahldarstellung in einem Stellenwertsystem zur Basis $b = 10$ zu verwenden. Die Wahl der Basis ist dabei eigentlich willkürlich. In der Antike waren auch andere Basen üblich, bei den Babyloniern etwa die Basis $b = 60$. In der Informatik nutzt man vornehmlich die Basen $b = 2$, das Dualsystem, und $b = 16$, das Hexadezimalsystem.

Definition und Satz 12.43 (b-adische Zahldarstellung)
Es sei $b \in \mathbb{N}$ eine natürliche Zahl mit $b \geq 2$ und es sei $z \in \mathbb{Z}$. Dann ist die Reihe

$$s = \sum_{n=z}^{\infty} a_n \cdot b^{-n}$$

konvergent in $\mathbb{R}$ für jede Wahl einer Folge $(a_n)_{n\geq z}$ mit

$$a_n \in \{0, 1, \ldots, b-1\}$$

für alle $n \geq z$. Wir nennen die Reihe dann eine Darstellung der reellen Zahl s als *b-adischen Bruch* und schreiben diesen auch in der Form

$$s = (a_z, a_{z+1}a_{z+2}a_{z+3}\ldots \; E-z)_b.$$

Wenn $a_n = 0$ für $n \geq N$ gilt, dann schreiben wir auch

$$s = (a_z, a_{z+1}a_{z+2}a_{z+3}\ldots a_N \; E-z)_b$$

und sprechen von einem *endlichen b-adischen Bruch*. Wenn sich die Ziffernfolge ab einer Stelle N periodisch wiederholt, d.h.

$$a_{N+i} = a_{N+i+k\cdot m}$$

für alle $i = 1, \ldots, m$ und alle $k \in \mathbb{N}$, dann schreiben wir

$$s = (a_z, a_{z+1} \ldots a_N \, \overline{a_{N+1} \ldots a_{N+m}} \, E - z)_b$$

und sprechen von einem *periodischen b-adischen Bruch*.

Beweis: Da die a_n alle nicht negativ und durch $b-1$ beschränkt sind, gilt

$$|a_n \cdot b^{-n}| = a_n \cdot b^{-n} \leq (b-1) \cdot b^{-n},$$

und wegen $\frac{1}{b} < 1$ ist dann die geometrische Reihe

$$\sum_{n=z}^{\infty} (b-1) \cdot b^{-n} = (b-1) \cdot b^{-z} \cdot \sum_{n=0}^{\infty} \frac{1}{b^n} = \frac{(b-1) \cdot b^{-z}}{1 - \frac{1}{b}} = b^{1-z} \tag{19}$$

ist eine konvergente Majorante von s. Mithin ist die Reihe absolut konvergent und ihr Grenzwert s ist eine reelle Zahl. □

Beispiel 12.44

a. Die natürliche Zahl 1274 hat die 10-adische Darstellung

$$1274 = 1 \cdot 10^3 + 2 \cdot 10^2 + 7 \cdot 10^1 + 4 \cdot 10^0 = (1, 274 \; E \; 3)_{10}.$$

b. Die rationale Zahl $\frac{1}{3}$ hat die 10-adische Darstellung

$$\frac{1}{3} = \sum_{n=1}^{\infty} 3 \cdot 10^{-n} = (3, \overline{3} \; E \; -1)_{10} = (0, \overline{3} \; E \; 0)_{10}.$$

c. Die rationale Zahl $\frac{1685}{49}$ hat die 7-adische Darstellung

$$\frac{1685}{49} = 4 \cdot 7^1 + 6 \cdot 7^0 + \frac{2}{7} + \frac{5}{7^2} = (4, 625 \; E \; 1)_7.$$

d. Die natürliche Zahl 2 besitzt die 2-adische Darstellungen $2 = (1, 0 \; E1)_2$ und

$$2 = \frac{1}{1 - \frac{1}{2}} = \sum_{n=0}^{\infty} 2^{-n} = (0, \overline{1} \; E \; 1)_2.$$

Satz 12.45 (*b*-adische Zahldarstellung)

Sei $b \in \mathbb{N}$ eine natürliche Zahl mit $b \geq 2$.

a. Jede nicht-negative reelle Zahl besitzt eine Darstellung als b-adischer Bruch.

b. Die Darstellung einer reellen Zahl als b-adischer Bruch ist bis auf die folgende Uneindeutigkeit eindeutig: Ist $N \geq z$ eine feste ganze Zahl mit $a_n = c_n$ für

$z \leq n \leq N-1$, $a_N < c_N$ und

$$(a_z, a_{z+1}a_{z+2}a_{z+3}\dots\ E-z)_b = (c_z, c_{z+1}c_{z+2}c_{z+3}\dots\ E-z)_b,$$

so gilt $a_N + 1 = c_N$ sowie $a_n = b-1$ und $c_n = 0$ für alle $n > N$.

c. Jede natürliche Zahl $s \in \mathbb{N}$ besitzt eine Darstellung als endlicher b-adischer Bruch der Form

$$s = (a_z, a_{z+1} \dots a_0\ E-z)_b.$$

d. Ein b-adischer Bruch ist genau dann in $\mathbb{Q}$, wenn er endlich oder periodisch ist.

Beweis:

a. Es reicht, die Aussage für positive reelle Zahlen $s \in \mathbb{R}_{>0}$ zu zeigen, da sie für $s = 0$ offenbar gilt. Weil $(b^{-n})_{n\in\mathbb{N}}$ streng monoton fallend mit

$$\lim_{n\to\infty} b^{-n} = 0$$

und $(b^n)_{n\in\mathbb{N}}$ streng monoton wachsend mit

$$\lim_{n\to\infty} b^n = \infty$$

ist, können wir eine ganze Zahl $z \in \mathbb{Z}$ wählen, so daß

$$b^{-z} \leq s < b^{-z+1}.$$

Wir wollen die Koeffizienten a_n für $n \geq z$ nun rekursiv so definieren, daß stets

$$a_n := \max(A_n)$$

für

$$A_n := \left\{ a \in \mathbb{N} \;\middle|\; s - \sum_{i=z}^{n-1} a_i \cdot b^{-i} - a \cdot b^{-n} \geq 0 \right\}$$

gilt. Wir nehmen dazu an, daß wir für ein festes $n \geq z$ solche Koeffizienten $a_z, \dots, a_{n-1}$ bereits gefunden haben. Dann ist die Menge A_n nicht leer, da $a = 0$ die gesuchte Bedingung sicher erfüllt. Wir zeigen nun, daß

$$A_n \subseteq \{0, 1, \dots, b-1\} \tag{20}$$

gilt. Wäre dies nicht der Fall, dann wäre $b \in A_n$ und es würde

$$0 \leq s - \sum_{i=z}^{n-1} a_i \cdot b^{-i} - b \cdot b^{-n} = s - \sum_{i=z}^{n-2} a_i \cdot b^{-i} - (a_{n-1}+1) \cdot b^{-(n-1)}$$

gelten und somit wäre $a_{n-1} + 1 \in A_{n-1}$ im Widerspruch dazu, daß a_{n-1} das Maximum dieser Menge ist. Also ist (20) erfüllt und es folgt

$$a_n \in \{0, 1, \dots, b-1\}.$$

Wir erhalten auf diese Weise also einen b-adischen Bruch und für die Folge der zugehörigen Partialsummen gilt nach Konstruktion

$$s - \sum_{i=z}^{n} a_i \cdot b^{-i} - b^{-n} = s - \sum_{i=z}^{n-1} a_i \cdot b^{-i} - (a_n + 1) \cdot b^{-n} < 0$$

und somit

$$0 \leq s - \sum_{i=z}^{n} a_i \cdot b^{-i} \leq b^{-n}.$$

Die Folge in der Mitte ist nun durch zwei Nullfolgen eingeschachtelt und ist somit auch eine Nullfolge, woraus

$$s = \sum_{n=z}^{\infty} a_n \cdot b^{-n}$$

folgt.

b. Wegen $\sum_{n=z}^{\infty} a_n \cdot b^{-n} = \sum_{n=z}^{\infty} c_n \cdot b^{-n}$ und $a_n = c_n$ für $n = z, \ldots, N-1$ gilt auch

$$\sum_{n=N}^{\infty} a_n \cdot b^{-n} = \sum_{n=N}^{\infty} c_n \cdot b^{-n}. \tag{21}$$

Zudem folgt aus $a_N < c_N$ auch

$$c_N - a_N \in \{1, \ldots, b-1\}.$$

Insgesamt erhalten wir damit

$$\begin{aligned} 0 < (c_N - a_N) \cdot b^{-N} &= \sum_{N+1}^{\infty} a_n \cdot b^{-n} - \sum_{n=N+1}^{\infty} c_n \cdot b^{-n} = \sum_{N+1}^{\infty} (a_n - c_n) \cdot b^{-n} \\ &\leq \sum_{n=N+1}^{\infty} (b-1) \cdot b^{-n} \overset{(19)}{=} b^{-N}. \end{aligned}$$

Dies bedingt aber $c_N - a_N = 1$ und mithin

$$a_N + 1 = c_N.$$

Setzen wir dies in (21) ein und ziehen auf beiden Seiten $a_N \cdot b^{-N}$ ab, so erhalten wir

$$0 \leq b^{-N} + \sum_{n=N+1}^{\infty} c_n \cdot b^{-n} = \sum_{n=N+1}^{\infty} a_n \cdot b^{-n} \leq \sum_{n=N+1}^{\infty} (b-1) \cdot b^{-n} \overset{(19)}{=} b^{-N} \tag{22}$$

und damit

$$\sum_{n=N+1}^{\infty} c_n \cdot b^{-n} = 0.$$

Wäre einer der Koeffizienten c_n positiv, so wäre der Grenzwert der Reihe durch dieses c_n nach unten beschränkt und nicht null. Also gilt $c_n = 0$ für alle $n \geq N+1$. Damit folgt aus (22) dann

$$\sum_{n=N+1}^{\infty} a_n \cdot b^{-n} = b^{-N} \overset{(19)}{=} \sum_{n=N+1}^{\infty} (b-1) \cdot b^{-n}$$

und somit

$$\sum_{n=N+1} (b - 1 - a_n) \cdot b^{-n} = 0.$$

Da die Koeffizienten der Reihe alle nicht-negativ sind, folgt wie oben auch in diesem Fall, daß sie alle null sind, d.h.

$$a_n = b - 1$$

für alle $n \geq N + 1$.

c. Wir zeigen mittels Induktion nach s, daß s eine Darstellung als b-adischer Bruch der Form

$$s = \sum_{n=z}^{0} a_n \cdot b^{-n} = (a_z, a_{z+1} \dots a_0 \; E - z)_b$$

besitzt. Für $s = 0$ ist dies offensichtlich der Fall. Sei also $s > 0$. Teilen wir s durch b mit Rest, so erhalten wir natürliche Zahlen $q, r \in \mathbb{N}$, so daß

$$s = q \cdot b + r$$

und $0 \leq r < b$. Die Zahl $q = \frac{s-r}{b}$ ist kleiner als s und besitzt mittels Induktion eine Darstellung als b-adischer Bruch der Form

$$q = \sum_{n=z}^{0} a_n \cdot b^{-n}.$$

Dann ist aber

$$s = q \cdot b + r = \sum_{n=z+1}^{1} a_{n-1} \cdot b^{-n} + r = (a_z, a_{z+1} \dots a_0 r \; E - (z+1))_b$$

ebenfalls eine solche.

d. Ist $s = (a_z, a_{z+1} a_{z+2} \dots a_N \; E - z)_b$ ein endlicher b-adischer Bruch, dann ist

$$s = \sum_{n=z}^{N} a_n \cdot b^{-n} \in \mathbb{Q}$$

als Summe endlich vieler rationaler Zahlen eine rationale Zahl.

Ist $s = (a_z, a_{z+1} \dots a_N \ \overline{a_{N+1} \dots a_{N+m}} \ E - z)_b$ ein periodischer b-adischer Bruch, dann gilt

$$\begin{aligned} s - b^m \cdot s &= \sum_{n=z}^{\infty} a_n \cdot b^{-n} - \sum_{n=z-m}^{\infty} a_{n+m} \cdot b^{-n} \\ &= \sum_{n=z}^{N} a_n \cdot b^{-n} - \sum_{n=z-m}^{N} a_{n+m} \cdot b^{-n} =: q, \end{aligned}$$

wobei dann q als Differenz zweier endlicher b-adischer Brüche eine rationale Zahl ist. Aber damit gilt dann auch

$$s = \frac{q}{1 - b^m} \in \mathbb{Q}.$$

Es bleibt noch zu zeigen, daß eine b-adische Darstellung einer rationalen Zahl

$$s = \sum_{n=z}^{\infty} a_n \cdot b^{-n} \in \mathbb{Q}$$

stets endlich oder periodisch ist. Es reicht dabei, dies für

$$s - \sum_{n=z}^{0} a_n \cdot b^{-n} = \sum_{n=1}^{\infty} a_n \cdot b^n$$

zu zeigen, da der endliche Anfangsteil des b-adischen Bruches nichts an der Eigenschaft ändert endlich oder periodisch zu sein. Wir können also ohne Einschränkung $z > 0$ annehmen.
Zunächst schreiben wir $s = \frac{p}{q}$ als Bruch mit $p, q \in \mathbb{N}$ und setzen

$$r_n := q \cdot b^n \cdot \left(s - \sum_{i=z}^{n} a_i \cdot b^{-i} \right) \in \mathbb{N}$$

für $n \geq z - 1 \geq 0$. Wir stellen zunächst fest, daß

$$r_n = q \cdot b^n \cdot \sum_{i=n+1}^{\infty} a_i \cdot b^{-i} \leq q \cdot b^n \sum_{i=n+1}^{\infty} (b-1) \cdot b^{-i} \overset{(19)}{=} q$$

gilt und daß r_n nicht negativ ist.
Gilt $r_n = 0$ für ein n, dann folgt $a_i = 0$ für alle $i \geq n + 1$ und s ist ein endlicher b-adischer Bruch.
Gilt $r_n = q$ für ein n, dann gilt $a_i = b-1$ für alle $i \geq n+1$ und s ist ein periodischer Bruch.

Wir können also $0 < r_n < q$ für alle $n \geq z$ annehmen. Aus der Definition von r_n folgt dann unmittelbar

$$\begin{aligned} b \cdot r_{n-1} &= q \cdot b^n \cdot \left(s - \sum_{i=z}^{n-1} a_i \cdot b^{-i} \right) \\ &= q \cdot b^n \cdot \left(s - \sum_{i=z}^{n} a_i \cdot b^{-i} \right) + q \cdot a_n \\ &= a_n \cdot q + r_n, \end{aligned}$$

d.h. r_n ist der Rest bei Division mit Rest von $b \cdot r_{n-1}$ durch q und a_n ist der ganzzahlige Anteil dabei. Beide sind durch den Wert $b \cdot r_{n-1}$ eindeutig bestimmt. Wegen $0 < r_n < q$ kommen für den Rest nur endlich viele Werte in Frage und somit kommen auch für die Zahlen $b \cdot r_{n-1}$, die geteilt werden sollen, nur endlich viele Werte in Frage, so daß einer dieser Werte irgendwann ein zweites Mal auftaucht. Damit wiederholen sich die Reste und die ganzzahligen Anteile von diesem Index an aber periodisch und s ist ein periodischer b-adischer Bruch.

□

Bemerkung 12.46 (Algorithmus zur Berechnung b-adischer Darstellungen)
Aus dem Beweis von Satz 12.45 läßt sich ein Algorithmus zur Berechnung der b-adischen Darstellung einer rationalen Zahl s herleiten.

Zunächst zerlegen wir

$$s = x + y$$

in seinen ganzzahligen Anteil $x \in \mathbb{N}$ und seinen gebrochenen Anteil $y = \frac{p}{q} \in [0, 1)$.

Für x berechnen wir eine b-adische Darstellung wie im Beweis von Beweis von Satz 12.45 c.: Wir setzen $x_0 = x$ und berechnen dann rekursiv mittels Division mit Rest natürliche Zahlen $x_n, a_{1-n} \in \mathbb{N}$ mit

$$x_{n-1} = x_n \cdot b + a_{1-n}$$

mit

$$0 \leq a_{1-n} < b$$

für $n \geq 1$. Wegen $0 \leq x_n < x_{n-1}$ bricht der Prozess nach endlich vielen Schritten mit einem $x_N = 0$ ab. Wir erhalten damit die b-adische Darstellung

$$x = \sum_{n=-N}^{0} a_n \cdot b^{-n}.$$

Für $y = \frac{p}{q}$ setzen wir zunächst $r_0 = q \cdot y = p$ und bestimmen rekursiv mittels Division mit Rest ganze Zahlen a_n, r_n mit

$$b \cdot r_{n-1} = a_n \cdot q + r_n$$

und

$$0 \leq r_n < q$$

für $n \geq z$.[3]Wir stoppen den Prozess, wenn der Rest $r_N = 0$ null wird oder erstmals ein Rest r_{N+1} zum zweiten Mal als r_{N+m+1} auftaucht. Im ersten Fall ist

$$y = \sum_{n=1}^{N} a_n \cdot b^{-n} = (a_1, a_{z+1} \dots a_N \; E - 1)_b$$

ein endlicher b-adischer Bruch, im zweiten Fall ist

$$y = (a_1, a_{z+1} \dots a_N \; \overline{a_{N+1} \dots a_{N+m}} \; E - 1)_b$$

ein periodischer b-adischer Bruch. Um die Darstellung für s zu erhalten, brauchen wir dann nur die Darstellungen für x und y zusammenzusetzen.

Beispiel 12.47 (Berechnung einer 2-adischen Darstellung)
Wir wollen mit dem in Bemerkung 12.46 beschriebenen Algorithmus für die rationale Zahl

$$s = \frac{37}{5} = 7 + \frac{2}{5}$$

die Darstellung als 2-adischer Bruch berechnen.

Für den ganzzahligen Anteil $x = x_0 = 7$ erhalten wir durch wiederholte Division mit Rest durch $b = 2$:

$$\begin{aligned} x_0 &= 7 = 3 \cdot 2 + 1 = x_1 \cdot b + a_0 \\ x_1 &= 3 = 1 \cdot 2 + 1 = x_2 \cdot b + a_{-1} \\ x_2 &= 1 = 0 \cdot 2 + 1 = x_3 \cdot b + a_{-2} \end{aligned}$$

Der Algorithmus für x bricht hier wegen $x_3 = 0$ ab und wir erhalten die 2-adische Darstellung

$$x = 7 = 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = (1, 11 \; E2)_2.$$

[3]Woher wissen wir, daß $a_n \in \{0, \dots, b-1\}$ liegt? Einerseits folgt dies aus dem Beweis von Satz 12.45 d., wo gezeigt wurde, daß die Koeffizienten der b-adischen Darstellung dem ganzzahligen Anteil in obiger Division mit Rest entsprechen, und dieser ist eindeutig bestimmt. Also muß der ganzzahlige Anteil, den wir hier bei der Division mit Rest ausrechnen auch kleiner als b sein. Andererseits kann man dies auch direkt nachrechnen. Wir wissen, daß stets $r_{n-1} < q$ gilt und damit erhalten wir

$$b = \frac{b \cdot q}{q} > \frac{b \cdot r_{n-1}}{q} = \frac{a_n \cdot q + r_n}{q} \geq \frac{a_n \cdot q}{q} = a_n.$$

Für $y = \frac{p}{q} = \frac{2}{5}$ und $r_0 = p = 2$ erhalten wir durch wiederholte Division mit Rest durch $q = 5$:

$$\begin{aligned} b \cdot r_0 &= 2 \cdot 2 = 0 \cdot 5 + 4 = a_1 \cdot 5 + r_1 \\ b \cdot r_1 &= 2 \cdot 4 = 1 \cdot 5 + 3 = a_2 \cdot 5 + r_2 \\ b \cdot r_2 &= 2 \cdot 3 = 1 \cdot 5 + 1 = a_3 \cdot 5 + r_3 \\ b \cdot r_3 &= 2 \cdot 1 = 0 \cdot 5 + 2 = a_4 \cdot 5 + r_4 \end{aligned}$$

Der Algorithmus bricht wegen $r_4 = r_0$ hier ab und wir erhalten die periodische b-adische Darstellung

$$y = (\overline{0,110}\ E\ -1)_2.$$

Setzen wir diese zusammen, so erhalten wir

$$s = \frac{37}{5} = (1, 11\overline{0110}\ E2)_2.$$

Bemerkung 12.48 (Binärzahlen)

Für einige Basen b haben die b-adischen Zahlsysteme eigene Namen, die sich aus dem Wert der Basis ableiten. Bei $b = 10$ spricht man von *Dezimalzahlen*, bei $b = 2$ von *Binärzahlen* und bei $b = 16$ von *Hexadezimalzahlen*.

Bemerkung 12.49 (Binärzahlen in der Informatik)

Reelle Zahlen werden in Rechnern als Binärzahlen gespeichert. Im Standard *IEEE754*, double-precision binary floating point, stehen für eine reelle Zahl 8 Byte (64 Bit) als Speicher zur Verfügung. Dabei wird ein Bit für das Vorzeichen verwendet, damit auch negative Zahlen gespeichert werden können. Für den Exponenten (den Wert nach dem E) stehen 11 Bit zur Verfügung, von denen wiederum eines für das Vorzeichen des Exponenten verwendet wird. Die Zahlenfolge vor dem E in der Binärdarstellung nennt man die *Mantisse* und für diese stehen 52 Bit zur Verfügung womit 53 Stellen gespeichert werden können, indem man die führende Zahl weg läßt, da sie stets eine 1 sein muß. Eine solche sieht dann in etwa wie folgt aus:

$$s = \underbrace{-}_{1\text{ Bit}} \cdot 1, \underbrace{01100011\ldots001}_{52\text{ Bit}} \cdot 2^{\underbrace{-(1,001111001\ E\ 10)_2}_{11\text{ Bit}}}.$$

Der maximale positve Wert, der so dargestellt werden kann, ist

$$\left(2 - 2^{-52}\right) \cdot 2^{1023} \sim 1,8 \cdot 10^{308},$$

und der minimale positive Wert ist

$$2^{-1022} \sim 2,3 \cdot 10^{-308}.$$

Man beachte aber, daß auch bei weitem nicht alle positiven reellen Zahlen zwischen diesen beiden Werten im Rechner dargestellt werden können, so daß bei Rechnungen of schon beim Abspeichern der Eingabedaten Fehler entstehen (müssen), weil die binäre

Darstellung der eingegebenen Zahlen nicht in das Raster paßt. Wie sich solche Fehler auf nachfolgende Rechnungen auswirken und was man tun sollte, um sie im Griff zu behalten, untersucht die Numerische Mathematik.

Aufgaben

Aufgabe 12.50

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 12.51 (Nicht-konvergentes Cauchy-Produkt konvergenter Reihen)

Betrachte die Folgen $(a_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ mit

$$a_n = (-1)^n \cdot \frac{1}{\sqrt{n+1}}$$

und

$$c_n = \sum_{k=0}^{n} a_k \cdot a_{n-k} = \sum_{k=0}^{n} \frac{(-1)^k}{\sqrt{k+1}} \cdot \frac{(-1)^{n-k}}{\sqrt{n-k+1}}.$$

Zeige, die Reihe $\sum_{n=0}^{\infty} a_n$ ist konvergent, aber das Cauchy-Produkt $\sum_{n=0}^{\infty} c_n$ ist divergent.

Aufgabe 12.52

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 12.53

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 12.52.

Aufgabe 12.54

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 12.55

Beweise für $x, y \in \mathbb{K}$ die Gleichung $\sin(x+y) = \sin(x) \cdot \cos(y) + \cos(x) \cdot \sin(y)$.

Aufgabe 12.56

Zeige, daß die beiden Reihen

$$\sum_{n=0}^{\infty} \frac{\sin(n)}{n!}$$

und

$$\sum_{n=0}^{\infty} \frac{\cos(n)}{n!}$$

absolut konvergent sind und berechne ihren Grenzwert.

Aufgabe 12.57

Zeige für alle $x, y \in \mathbb{R}$ die Gleichung

$$\cos(x) - \cos(y) = -2 \cdot \sin\left(\frac{x+y}{2}\right) \cdot \sin\left(\frac{x-y}{2}\right).$$

Aufgabe 12.58

Für $x \in \mathbb{R}$ mit $e^{ix} \neq 1$ und $n \in \mathbb{N}$ zeige

$$\sum_{k=1}^{n} \cos(k \cdot x) = \frac{\sin\left(\frac{2n+1}{2} \cdot x\right)}{2 \cdot \sin\left(\frac{x}{2}\right)} - \frac{1}{2}.$$

Aufgabe 12.59 (Binomialreihe)

Für eine beliebige reelle Zahl $a \in \mathbb{R}$ und eine natürliche Zahl $k \in \mathbb{N}$ definieren wir den *Binomialkoeffizienten*

$$\binom{a}{k} = \frac{a \cdot (a-1) \cdot \ldots (a-k+1)}{k!} \in \mathbb{R}.$$

a. Zeige, für $a \in \mathbb{R} \setminus \mathbb{N}$ hat die Potenzreihe $\sum_{n=0}^{\infty} \binom{a}{n} \cdot t^n$ den Konvergenzradius 1, (siehe auch Aufgabe 18.50).

b. Berechne die Binomialkoeffizienten $\binom{a}{k}$ für den Fall $a = 3$ für alle k und berechne damit auch die Potenzreihe in (a) exakt. Welchen Konvergenzradius hat die Binomialreihe für $a \in \mathbb{N}$?

Aufgabe 12.60

a. Zeige, daß die geometrische Reihe $\sum_{n=0}^{\infty} q^n$ mit $q = \frac{2}{3}$ eine konvergente Majorante der Reihe $\sum_{n=1}^{\infty} \frac{1}{n!}$ ist und leite daraus eine obere Schranke für den Grenzwert der Reihe ab.

b. Zeige, die Folge $(a_n)_{n\geq 1}$ mit $a_n = \left(1 + \frac{1}{n}\right)^n$ ist monton wachsend und konvergent.

Hinweis zu Teil b., man kann die Bernoulli-Ungleichung an geeigneter Stelle verwenden.

Aufgabe 12.61

Bestimme für die folgenden Reihen, ob sie konvergent, absolut konvergent oder divergent sind:

a. $\sum_{n=1}^{\infty} (-1)^n \cdot \left(1 - \frac{1}{n}\right)^n$.

b. $\sum_{n=0}^{\infty} \frac{3^n}{3 \cdot n!}$.

c. $\sum_{n=1}^{\infty} (-1)^n \cdot \frac{1}{\sqrt{n}}$.

d. $\sum_{n=1}^{\infty} \left(\frac{n+2}{n}\right)^n$.

e. $\sum_{n=1}^{\infty} \frac{1}{n+\sqrt{n}}$.

f. $\sum_{n=1}^{\infty} \frac{n!}{n^n}$.

Aufgabe 12.62
Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ zwei Folgen in $\mathbb{C}$, so daß die Reihen $\sum_{n=0}^{\infty} a_n^2$ und $\sum_{n=0}^{\infty} b_n^2$ absolut konvergent sind. Zeige, dann ist auch die Reihe $\sum_{n=0}^{\infty} a_n \cdot b_n$ absolut konvergent.

Aufgabe 12.63
Gib ein Beispiel für eine absolut konvergente Reihe $\sum_{n=0}^{\infty} a_n$ mit $\left|\frac{a_{n+1}}{a_n}\right| > 1$ für unendlich viele $n \in \mathbb{N}$. (Wie immer mit Begründung!)

Aufgabe 12.64
Bestimme den Konvergenzradius der Reihe $\sum_{n=1}^{\infty} t^n$ über $\mathbb{R}$ und untersuche die Reihe auf Konvergenz in den Randpunkten des Konvergenzintervalls.

Aufgabe 12.65
Bestimme für die folgenden Potenzreihen über $\mathbb{R}$ den Konvergenzradius r und untersuche das Konvergenzverhalten in den Randpunkten des Konvergenzintervalls $(-r, r)$:

a. $\sum_{n=1}^{\infty} \frac{1}{n\cdot 4^n} \cdot t^n$.
b. $\sum_{n=1}^{\infty} \frac{n}{2^n} \cdot t^{n+1}$.
c. $\sum_{n=0}^{\infty} \sqrt{n} \cdot t^n$.
d. $\sum_{n=0}^{\infty} n! \cdot t^n$.
e. $\sum_{n=1}^{\infty} \left(\frac{n+1}{n}\right)^n \cdot t^n$.
f. $\sum_{n=0}^{\infty} (-1)^n \cdot t^n$.

Aufgabe 12.66
Zeige, wenn $\sum_{n=0}^{\infty} a_n \cdot t^n$ den Konvergenzradius r hat und $c \in \mathbb{R} \setminus \{0\}$ ist, dann hat $\sum_{n=0}^{\infty} \frac{a_n}{c^n} \cdot t^n$ den Konvergenzradius $|c| \cdot r$.

Aufgabe 12.67
Bestimme die 5-adische Darstellung der rationalen Zahl $\frac{111336}{1000}$ mittels des Algorithmus' aus der Vorlesung.

Aufgabe 12.68

a. Berechne die 7-adische Darstellung der Zahl 4973.

b. Berechne das Produkt der Zahlen $(2,33\ E2)_5$ und $(1,01\ E1)_5$ im Dezimalsystem.

c. Berechne die 3-adische Zahldarstellung der rationalen Zahl $\frac{123}{75}$.

d. Berechne die Dezimalzahldarstellung des 3-adischen Bruches $(2,12121\ E2)_3$.

e. Berechne die b-adische Zahldarstellung von $\frac{1}{7}$ für $b = 2, 7, 10, 16$.

§ 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

$$\forall\, \varepsilon > 0 \;\; \exists\, x \in U \setminus \{a\} \;:\; 0 < |x - a| < \varepsilon.$$

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.

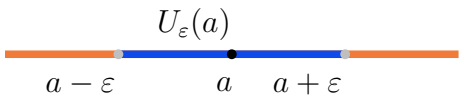


Abbildung 4: Die ε-Umgebung $U_\varepsilon(a)$ von a.

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 $\lim\limits_{n\to\infty} a_n = a$ gibt, d.h. a ist Grenzwert einer Folge in $U \setminus \{a\}$.

Beweis: Ist a ein Häufungspunkt von U und $n \in \mathbb{N}$, so gibt es zu $\varepsilon = \frac{1}{n+1} > 0$ ein $x \in U$ mit $0 < |x - a| < \varepsilon = \frac{1}{n+1}$. Wir wählen a_n als dieses x. Dann konvergiert die Folge $(a_n)_{n\in\mathbb{N}}$ offenbar gegen a nach dem Einschließungssatz, da

$$0 < |a_n - a| < \frac{1}{n+1} \longrightarrow 0.$$

Gibt es umgekehrt eine Folge $(a_n)_{n\in\mathbb{N}}$ in $U \setminus \{a\}$, die gegen a konvergiert, so gibt es für jedes $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$ mit $0 < |a_n - a| < \varepsilon$ für alle $n \geq n_\varepsilon$. Setzen wir nun $x = a_{n_\varepsilon} \in U$, so folgt $0 < |x - a| < \varepsilon$ und somit ist a ein Häufungspunkt von U. □

Das folgende Beispiel zeigt, welche Art von Häufungspunkten einer Menge U, die nicht bereits in U liegen, wir typischerweise erwarten.

Beispiel 13.5
Sind $a, b \in \mathbb{R}$ mit $a < b$, so enthält $[a, b]$ genau die Häufungspunkte von (a, 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 $c \in \mathbb{R}$ ein Häufungspunkt von (a, b), so gibt es nach dem Folgenkriterium 13.4 eine Folge $(a_n)_{n\in\mathbb{N}}$ in $(a, b) \subset [a, b]$, die gegen c konvergiert, und nach Satz 11.28 ist dann $c \in [a, b]$.
Ist $a \leq c < b$, so gilt für $n \geq 1$

$$(a, b) \ni c + \frac{b - c}{2 \cdot n} \longrightarrow c,$$

also ist c ein Häufungspunkt von (a, b). Analog gilt $(a, b) \ni b - \frac{b-a}{2\cdot n} \longrightarrow b$, so daß auch b ein Häufungspunkt von (a, b) ist. □

B) Grenzwerte von Funktionen

Definition 13.6 (ε-δ-Kriterium für Grenzwerte von Funktionen)
Sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ eine Funktion und a ein Häufungspunkt von U.
Wir nennen $y \in \mathbb{R}$ den *Grenzwert* von f in a, falls

$$\forall\, \varepsilon > 0 \;\; \exists\, \delta_\varepsilon > 0 \;:\; \forall\, x \in U \text{ mit } 0 < |x - a| < \delta_\varepsilon \;\text{ gilt }\; |f(x) - y| < \varepsilon.$$

Wir schreiben dann

$$\lim_{x\to a} f(x) = y$$

oder "$f(x) \longrightarrow y$ für $x \to a$" und sagen, $f(x)$ *konvergiert* gegen y für x gegen a.

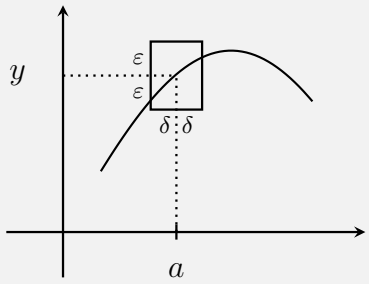


Abbildung 5: ε-δ-Kriterium für Grenzwerte

Proposition 13.7 (Folgenkriterium für Grenzwerte von Funktionen)
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ eine Funktion und a ein Häufungspunkt von U.
Dann sind die beiden folgenden Aussagen gleichwertig:

a. $\lim\limits_{x\to a} f(x) = y$.

b. $\forall\, (a_n)_{n\in\mathbb{N}}$ mit $a_n \in U \setminus \{a\}$ und $\lim\limits_{n\to\infty} a_n = a$ gilt $\lim\limits_{n\to\infty} f(a_n) = y$.

Beweis: **a. $\Longrightarrow$ b.:** Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $U\setminus\{a\}$ mit $\lim\limits_{n\to\infty} a_n = a$. Wir müssen $\lim\limits_{n\to\infty} f(a_n) = y$ zeigen. Dazu sei $\varepsilon > 0$ gegeben.
Wegen $\lim\limits_{x\to a} f(x) = y$ 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 $\lim\limits_{n\to\infty} a_n = a$ 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 $f(a_n) \longrightarrow y$ gezeigt.

b. $\Longrightarrow$ a.: Wir nehmen an, y wäre nicht der Grenzwert von f in a. Dann gilt:

$$\exists\, \varepsilon > 0 \;:\; \forall\, \delta_\varepsilon > 0\; \exists\, x_{\delta_\varepsilon} \in U \;\text{ mit }\; 0 < |x_{\delta_\varepsilon} - a| < \delta_\varepsilon, \;\text{ aber } |f(x_{\delta_\varepsilon}) - y| \geq \varepsilon.$$

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

$$0 < |a_n - a| < \frac{1}{n} \longrightarrow 0,$$

so daß $a_n \longrightarrow a$, 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 $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$ und $a = 3$. Für eine Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n \longrightarrow 3$ gilt dann wegen der Grenzwertsätze für Folgen 11.15

$$f(a_n) = a_n^2 = a_n \cdot a_n \longrightarrow 3 \cdot 3 = 9.$$

Mithin ist 9 der Grenzwert von f in 3, d.h.

$$\lim_{x\to 3} x^2 = 9 = f(3).$$

b. Betrachte die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 1, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0 \end{cases}$$

und $a = 0$. Ist nun $(a_n)_{n\in\mathbb{N}}$ eine Folge mit $a_n \longrightarrow 0$ und $a_n \neq 0$, dann gilt

$$f(a_n) = 1 \longrightarrow 1.$$

Mithin ist 1 der Grenzwert von f in 0, d.h.

$$\lim_{x\to 0} f(x) = 1 \neq 0 = f(0).$$

c. Betrachte die Funktion

$$f : \mathbb{R} \setminus \{1\} \longrightarrow \mathbb{R} : x \mapsto \frac{x^2 - 1}{x - 1}$$

und $a = 1$. Da es in $\mathbb{R} \setminus \{1\}$ offenbar eine Folge gibt, die gegen 1 konvergiert, ist a ein Häufungspunkt von $\mathbb{R} \setminus \{1\}$. Sei nun $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R} \setminus \{1\}$ mit $a_n \longrightarrow 1$, so gilt

$$f(a_n) = \frac{a_n^2 - 1}{a_n - 1} = a_n + 1 \longrightarrow 2.$$

Mithin ist 2 der Grenzwert von f in 1, d.h.

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

Man beachte, daß in diesem Fall $a = 1$ gar nicht im Definitionsbereich von f liegt.

d. Betrachte die Funktion

$$f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x < 0, \\ 1, & \text{falls } x > 0 \end{cases}$$

und $a = 0$. Dann gilt für $a_n := -\frac{1}{n} \longrightarrow 0$ und $f(a_n) = 0 \longrightarrow 0$ sowie $b_n := \frac{1}{n} \longrightarrow 0$ und $f(b_n) = 1 \longrightarrow 1$. Mithin existiert der Grenzwert von f in $a = 0$ nicht.

!!! Warnung !!!

In manchen Büchern und von manchen Dozenten wird der Begriff des Grenzwertes leicht anders definiert (siehe z.B. [Dec10] oder [Gat08])! Wenn f im Punkt a definiert ist, muß bei uns **nicht** notwendig $\lim_{x\to a} f(x) = f(a)$ gelten (siehe Beispiel 13.8 b.), was nach deren Definition gelten muß!

C) Die Grenzwertsätze für Funktionen

Definition 13.9

Für zwei Funktionen $f : U \longrightarrow \mathbb{R}$ und $g : V \longrightarrow \mathbb{R}$ sowie $c \in \mathbb{R}$ definieren wir

$$c \cdot f : U \longrightarrow \mathbb{R} : x \mapsto c \cdot f(x),$$

$$f + g : U \cap V \longrightarrow \mathbb{R} : x \mapsto f(x) + g(x),$$

$$f - g : U \cap V \longrightarrow \mathbb{R} : x \mapsto f(x) - g(x)$$

und

$$f \cdot g : U \cap V \longrightarrow \mathbb{R} : x \mapsto f(x) \cdot g(x).$$

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 $\lim_{x\to a} f(x) = y$ und $\lim_{x\to a} f(x) = z$, so ist $y = z$.

b. Wenn $\lim_{x\to a} f(x)$ und $\lim_{x\to a} g(x)$ existieren, so gelten:

   (i) $\lim_{x\to a}(c \cdot f)(x) = c \cdot \lim_{x\to a} f(x)$.

   (ii) $\lim_{x\to a}(f + g)(x) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x)$.

   (iii) $\lim_{x\to a}(f - g)(x) = \lim_{x\to a} f(x) - \lim_{x\to a} g(x)$.

   (iv) $\lim_{x\to a}(f \cdot g)(x) = \lim_{x\to a} f(x) \cdot \lim_{x\to a} g(x)$.

c. Falls zudem $\lim_{x\to a} f(x) \neq 0$, so ist a ein Häufungspunkt der Menge $V = \{x \in U \mid f(x) \neq 0\}$ und es gilt

$$\lim_{x\to a} \frac{1}{f}(x) = \frac{1}{\lim_{x\to a} f(x)}.$$

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

$$y = \lim_{n\to\infty} f(a_n) = z.$$

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

$$f(a_n) \longrightarrow \lim_{x\to a} f(x) =: y.$$

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 $a_0, \ldots, a_n \in \mathbb{R}$, so nennen wir einen Ausdruck der Form

$$\sum_{k=0}^{n} a_k \cdot t^k = a_n \cdot t^n + a_{n-1} \cdot t^{n-1} + \ldots + a_1 \cdot t + a_0$$

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

$$\mathbb{R}[t] := \left\{ \sum_{k=0}^{n} a_k \cdot t^k \;\middle|\; n \in \mathbb{N}, a_0, \ldots, a_n \in \mathbb{R} \right\}$$

bezeichnen wir die Menge aller Polynome in der Veränderlichen t mit Koeffizienten in $\mathbb{R}$, so daß der Grad eine Abbildung $\deg : \mathbb{R}[t] \longrightarrow \mathbb{N} \cup \{-\infty\}$ 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

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto f(x)$$

eine *Polynomfunktion* und die Funktion

$$\frac{f}{g} : \mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) = 0\} \longrightarrow \mathbb{R} : x \mapsto \frac{f(x)}{g(x)}$$

nennen wir eine *rationale Funktion*.
Ist $h : \mathbb{R} \longrightarrow \mathbb{R}$ 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:

$$\big|\{x \in \mathbb{R} \mid g(x) = 0\}\big| \leq \deg(g) < \infty.$$

Beispiel 13.13

a. Ist $f = \sum_{k=0}^{n} a_k \cdot t^k$ ein Polynom und $a \in \mathbb{R}$, so gilt

$$\lim_{x \to a} f(x) = f(a).$$

Dies folgt aus den Grenzwertsätzen für Funktionen 13.10, da offenbar $\lim_{x \to a} \mathrm{id}(x) = \lim_{x \to a} x = a$ 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

$$\lim_{x \to a} \frac{f}{g}(x) = \frac{f(a)}{g(a)} = \frac{f}{g}(a).$$

D) Uneigentliche Grenzwerte

Definition 13.14 (Grenzwerte für $x \longrightarrow \pm\infty$)
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ 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 $\lim\limits_{x\to\infty} f(x) = y$.

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 $\lim\limits_{x\to-\infty} f(x) = y$.

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.

$$\lim_{x\to\infty} f(x) = y \iff \forall\, (a_n)_{n\in\mathbb{N}} \text{ mit } a_n \in U \text{ und } a_n \to \infty \text{ gilt } f(a_n) \longrightarrow y$$

und

$$\lim_{x\to-\infty} f(x) = y \iff \forall\, (a_n)_{n\in\mathbb{N}} \text{ mit } a_n \in U \text{ und } a_n \to -\infty \text{ gilt } f(a_n) \longrightarrow y.$$

Zudem gelten auch die Grenzwertsätze für Funktionen 13.10 für Grenzwerte in $\pm\infty$.

Definition 13.16 (Uneigentliche Grenzwerte)

Sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ 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 $\lim\limits_{x\to a} f(x) = \infty$.

b. Wir nennen $-\infty$ 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 $\lim\limits_{x\to a} f(x) = -\infty$.

c. Ist U nach oben unbeschränkt, so nennen wir ∞ den *uneigentlichen Grenzwert* von f in ∞, wenn

$$\forall\, s > 0\ \exists t > 0\ :\ \forall\, x \in U \text{ mit } x > t \text{ gilt } f(x) > s.$$

Wir schreiben dann $\lim\limits_{x\to\infty} f(x) = \infty$.

d. Ist U nach oben unbeschränkt, so nennen wir $-\infty$ den *uneigentlichen Grenzwert* von f in ∞, wenn

$$\forall\, s < 0 \;\; \exists t > 0 \;:\; \forall\, x \in U \text{ mit } x > t \text{ gilt } f(x) < s.$$

Wir schreiben dann $\lim\limits_{x\to\infty} f(x) = -\infty$.

e. Ist U nach unten unbeschränkt, so nennen wir ∞ den *uneigentlichen Grenzwert* von f in $-\infty$, wenn

$$\forall\, s > 0 \;\; \exists t < 0 \;:\; \forall\, x \in U \text{ mit } x < t \text{ gilt } f(x) > s.$$

Wir schreiben dann $\lim\limits_{x\to-\infty} f(x) = \infty$.

f. Ist U nach unten unbeschränkt, so nennen wir $-\infty$ den *uneigentlichen Grenzwert* von f in $-\infty$, wenn

$$\forall\, s < 0 \;\; \exists t < 0 \;:\; \forall\, x \in U \text{ mit } x < t \text{ gilt } f(x) < s.$$

Wir schreiben dann $\lim\limits_{x\to-\infty} f(x) = -\infty$.

Bemerkung 13.17 (Folgenkriterium und Grenzwertsätze für uneigentliche GWe)

Auch für uneigentliche Grenzwerte gelten naheliegende Folgenkriterien:

a. $\lim\limits_{x\to a} f(x) = \infty \iff \forall\, (a_n)_{n\in\mathbb{N}}$ mit $a_n \in U \setminus \{a\}$ und $a_n \to a$ gilt $f(a_n) \to \infty$.

b. $\lim\limits_{x\to\infty} f(x) = \infty \iff \forall\, (a_n)_{n\in\mathbb{N}}$ mit $a_n \in U$ und $a_n \to \infty$ gilt $f(a_n) \to \infty$.

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.

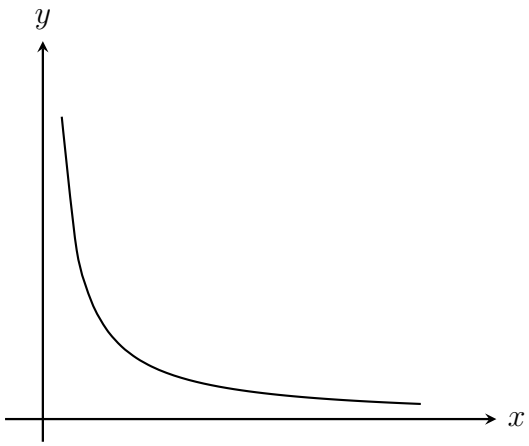
Beispiel 13.18

Für die Funktion $f : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x}$ ist 0 ein Häufungspunkt des Definitionsbereiches, und es gilt

$$\lim_{x\to 0} f(x) = \infty.$$

Zudem ist $(0,\infty)$ nach oben unbeschränkt und

$$\lim_{x\to\infty} f(x) = 0.$$



Beweis: Wenn $(a_n)_{n\in\mathbb{N}}$ eine Folge in $(0,\infty)$ mit $a_n \longrightarrow 0$ 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. $f(a_n) \longrightarrow \infty$.

Wenn $(a_n)_{n\in\mathbb{N}}$ eine Folge in $(0,\infty)$ mit $a_n \longrightarrow \infty$ 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. $f(a_n) \longrightarrow 0$. □

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

$$\lim_{x\to\infty} f(x) = \begin{cases} \infty, & \text{falls } a_n > 0, \\ -\infty, & \text{falls } a_n < 0, \end{cases}$$

und

$$\lim_{x\to-\infty} f(x) = \begin{cases} \infty, & \text{falls } (a_n > 0 \text{ und } n \text{ gerade }) \text{ oder } (a_n < 0 \text{ und } n \text{ ungerade}), \\ -\infty, & \text{falls } (a_n < 0 \text{ und } n \text{ gerade }) \text{ oder } (a_n > 0 \text{ und } n \text{ ungerade}). \end{cases}$$

Wir beweisen die Aussage nur für $\lim\limits_{x\to\infty} f(x)$ und $a_n > 0$, da der Rest sich analog zeigen läßt. Hierzu betrachten wir ein beliebiges $x \in \mathbb{R}$ mit

$$x \geq \max\left\{\frac{-2\cdot n\cdot a_0}{a_n}, \frac{-2\cdot n\cdot a_1}{a_n}, \ldots, \frac{-2\cdot n\cdot a_{n-1}}{a_n}, 1\right\}.$$

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 $\lim\limits_{x \to \infty} \frac{a_n \cdot x^n}{2} = \infty$, muß auch $\lim\limits_{x \to \infty} f(x) = \infty$ gelten. □

E) Rechts- und linksseitige Grenzwerte

Bemerkung 13.20 (Rechts- und linksseitge Grenzwerte)

Sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ und $a \in \mathbb{R}$. Wir definieren

$$U_{<a} := \{x \in U \mid x < a\}$$

und

$$U_{>a} := \{x \in U \mid x > a\}.$$

Ist a ein Häufungspunkt von $U_{<a}$ und existiert der Grenzwert der Einschränkung $f_{|U_{<a}}$ in a, so bezeichnen wir diesen mit

$$\lim_{x \to a-} f(x) := \lim_{x \to a} f_{|U_{<a}}(x)$$

und nennen ihn den *linksseitigen Grenzwert* von f in a.

Analog, ist a ein Häufungspunkt von $U_{>a}$ und existiert der Grenzwert der Einschränkung $f_{|U_{>a}}$ in a, so bezeichnen wir diesen mit

$$\lim_{x \to a+} f(x) := \lim_{x \to a} f_{|U_{>a}}(x)$$

und nennen ihn den *rechtsseitigen Grenzwert* von f in a.

Sei nun a ein Häufungspunkt $U_{<a}$ und von $U_{>a}$. Genau dann existieren links- und rechtsseitiger Grenzwert von f in a und stimmen überein, wenn auch der Grenzwert von f in a existiert und

$$\lim_{x \to a} f(x) = \lim_{x \to a-} f(x) = \lim_{x \to a+} f(x)$$

gilt. In manchen Situationen, insbesondere wenn Funktionen abschnittsweise definiert sind, ist es leichter, die links- und rechtsseitigen Grenzwerte getrennt zu berechnen.

Beweis: Wenn a ein Häufungspunkt einer Teilmenge von U ist, ist es automatisch auch ein Häufungspunkt von U. Es bleibt also zu zeigen, dass der Grenzwert von f in a existiert, wenn links- und rechtsseitiger Grenzwert übereinstimmen.

Sei $y = \lim_{x\to a-} f(x) = \lim_{x\to a+} f(x)$ und sei $\varepsilon > 0$ gegeben. Dann existiert ein $\delta'_\varepsilon > 0$, so daß für alle $x \in U_{<a}$ mit $|x - a| < \delta'_\varepsilon$ auch

$$|f(x) - y| < \varepsilon$$

gilt. Zudem existiert ein $\delta''_\varepsilon > 0$, so daß für alle $x \in U_{>a}$ mit $|x - a| < \delta''_\varepsilon$ auch

$$|f(x) - y| < \varepsilon$$

gilt. Wir setzen nun $\delta_\varepsilon := \min\{\delta'_\varepsilon, \delta''_\varepsilon\} > 0$. Dann gilt für alle $x \in U$ mit $|x - a| < \delta_\varepsilon$ auch

$$|f(x) - y| < \varepsilon.$$

Also ist $y = \lim_{x\to a} f(x)$. □

Beispiel 13.21 (Links- und rechtsseitiger Grenzwert)
Wir betrachten die Funktion

$$f : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} x + 1, & \text{wenn } x < 0, \\ x^3 + 2x^2 + 1, & \text{wenn } x > 0. \end{cases}$$

Mit $U = \mathbb{R}\setminus\{0\}$ ist dann $U_{<0} = (-\infty, 0)$ und $U_{>0} = (0, \infty)$, so daß 0 ein Häufungspunkt von beiden Mengen ist. Ferner gilt

$$\lim_{x\to 0-} f(x) = \lim_{x\to 0-} x + 1 = 1$$

und

$$\lim_{x\to 0+} f(x) = \lim_{x\to 0+} 3x^3 + 2x^2 + 1 = 1.$$

Links- und rechtsseitger Grenzwert von f in 0 existieren also und stimmen überein. Damit existiert dann auch der Grenzwert von f in 0 und es gilt

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

Aufgaben

Aufgabe 13.22
Bestimme die Häufungspunkte der folgenden Mengen:

a. $A = \left\{(-1)^n + \frac{n+1}{n} \;\middle|\; n \in \mathbb{N}\right\}$.

b. $B = \left\{x \in \mathbb{R} \;\middle|\; 3 < |x + 1|\right\}$.

Aufgabe 13.23
Bestimme für die nachfolgenden Mengen jeweils die Menge aller ihrer Häufungspunkte:

a. $A = \left\{ (-1)^n + \left(\frac{-1}{n}\right)^{n+1} \;\middle|\; n \in \mathbb{Z}_{\geq 1} \right\}$.

b. $B = \mathbb{N}$.

c. $C = \{x \in \mathbb{Q} \mid x^2 < 2\}$.

Aufgabe 13.24

Gib ein Beispiel für eine Menge, die genau drei Häufungspunkte hat.

Aufgabe 13.25

Bestimme die folgenden Grenzwerte oder zeige, daß sie nicht existieren:

a. $\lim\limits_{x\to 3} \frac{x^3-9x}{x-3}$.

b. $\lim\limits_{x\to 5} \frac{2x^4+x^3-3x^2+4}{x^2-5}$.

c. $\lim\limits_{x\to \infty} \sqrt{\frac{2x^3+5x}{(x+1)^3}}$.

d. $\lim\limits_{x\to \infty} \sqrt{x+\sqrt{x}} - \sqrt{x}$.

e. $\lim\limits_{x\to 2} \frac{x^3-2x^2-x+2}{x^2-x-2}$.

f. $\lim\limits_{x\to x_0} \frac{x^n-x_0^n}{x-x_0}$, wobei $n \in \mathbb{N}$ und $x_0 \in \mathbb{R}$ beliebig, aber fest vorgegeben sind.

g. $\lim\limits_{x\to \infty} \left(\sqrt{x+5} - \sqrt{x}\right)$.

Aufgabe 13.26 (Cauchy-Kriterium für Grenzwerte)

Es sei $U \subseteq \mathbb{R}$, $a \in \mathbb{R}$ ein Häugfungspunkt von U und $f : U \longrightarrow \mathbb{R}$ eine Funktion. Zeige, der Grenzwert $\lim\limits_{x\to a} f(x)$ existiert genau dann, wenn

$$\forall\, \varepsilon > 0\ \exists\, \delta_\varepsilon > 0\ :\ \forall\, x, y \in \big(U \cap U_{\delta_\varepsilon}(a)\big) \setminus \{a\} \ \text{ gilt } |f(x) - f(y)| < \varepsilon.$$

§ 14 Stetigkeit

A) Stetige Funktionen

Definition 14.1 (ε-δ-Kriterium für Stetigkeit)
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ 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.
$\mathcal{C}(U, \mathbb{R}) := \{f : U \longrightarrow \mathbb{R} \mid f \text{ stetig}\}$ ist die Menge der auf U stetigen Funktionen.

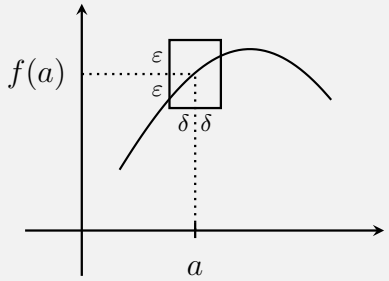


Abbildung 6: ε-δ-Kriterium für Stetigkeit

Bemerkung 14.2
Für die Stetigkeit einer Funktion in einem Punkt a ist nur das Verhalten von f in einer kleinen ϵ-Umgebung $U_\epsilon(a) = (a-\epsilon, a+\epsilon)$ von a maßgeblich. Wir sagen deshalb auch, daß die Stetigkeit eine *lokale* Eigenschaft ist!

Lemma 14.3 (Stetigkeit in Häufungspunkten)
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ und $a \in U$ ein Häufungspunkt.
Genau dann ist f stetig in a, wenn $\lim\limits_{x \to a} f(x) = f(a)$.

Beweis: Dies folgt unmittelbar aus den Definitionen 13.6 und 14.1. □

Beispiel 14.4 (Polynomfunktionen sind stetig.)

a. Jede Polynomfunktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ ist stetig.
Denn nach Beispiel 13.13 gilt für $a \in \mathbb{R}$ auch $\lim\limits_{x \to a} f(x) = f(a)$.

b. Jede rationale Funktion $\frac{f}{g} : \mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) = 0\} \longrightarrow \mathbb{R}$ 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 $\lim\limits_{x \to a} \frac{f}{g}(x) = \frac{f}{g}(a)$.

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 nicht stetig in 0, da $\lim_{x\to 0} f(x) = 1 \neq 0 = f(0)$. Aber, f ist stetig in jedem $a \neq 0$, wie man leicht sieht.

d. Ist $f : U \longrightarrow \mathbb{R}$ stetig und $V \subseteq U$, so ist die Einschränkung $f_| : V \longrightarrow \mathbb{R}$ von f auf V offenbar ebenfalls stetig.

e. Ist $f : \mathbb{Z} \longrightarrow \mathbb{R}$ irgendeine Funktion, so ist f stetig! (Kein nützliches Konzept!) Denn, ist $a \in \mathbb{Z}$ und $\varepsilon > 0$ wählen wir $\delta_\varepsilon := \frac{1}{2}$. Für $x \in \mathbb{Z}$ mit $|x - a| < \delta_\varepsilon = \frac{1}{2}$ muß dann $x = a$ gelten und somit auch $|f(x) - f(a)| = 0 < \varepsilon$.

Satz 14.5 (Folgenkriterium für Stetigkeit)
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ und $a \in U$.
Genau dann ist f stetig in a, wenn

(23) $\quad \forall\, (a_n)_{n\in\mathbb{N}}$ mit $a_n \in U$ und $\lim_{n\to\infty} a_n = a$ gilt $\lim_{n\to\infty} f(a_n) = f(a)$.

Beweis: Der Beweis geht genau wie der Beweis des Folgenkriteriums für Grenzwerte von Funktionen 13.7.

"$\Longrightarrow$": Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in U mit $\lim_{n\to\infty} a_n = a$. Wir müssen $\lim_{n\to\infty} f(a_n) = f(a)$ zeigen. Dazu sei $\varepsilon > 0$ gegeben.
Da f stetig in a ist, gibt es ein $\delta_\varepsilon > 0$, so daß aus $x \in U$ mit $|x - a| < \delta_\varepsilon$ auch $|f(x) - f(a)| < \varepsilon$ folgt.
Wegen $\lim_{n\to\infty} a_n = a$ 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 $|a_n - a| < \delta_\varepsilon$ und somit ist auch $|f(a_n) - f(a)| < \varepsilon$. Damit ist $f(a_n) \longrightarrow f(a)$ gezeigt.

"$\Longleftarrow$": Wir nehmen an, f wäre nicht stetig in a. Dann gilt:

$$\exists\, \varepsilon > 0 \;:\; \forall\, \delta_\varepsilon > 0 \;\; \exists\, x_{\delta_\varepsilon} \in U \;\text{ mit }\; |x_{\delta_\varepsilon} - a| < \delta_\varepsilon, \;\text{ aber } |f(x_{\delta_\varepsilon}) - f(a)| \geq \varepsilon.$$

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

$$0 \leq |a_n - a| < \frac{1}{n} \longrightarrow 0,$$

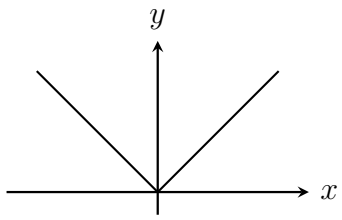
so daß $a_n \longrightarrow a$, und zugleich gilt

$$|f(a_n) - f(a)| \geq \varepsilon$$

für alle $n \in \mathbb{N}$. Dies ist ein Widerspruch dazu, daß $f(a_n)$ gegen $f(a)$ konvergieren muß.

□

Beispiel 14.6 (Die Betragsfunktion ist stetig)
Die Betragsfunktion $|\cdot| : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto |x|$ ist stetig.



Denn für $a \in \mathbb{R}$ und $(a_n)_{n\in\mathbb{N}}$ mit $a_n \longrightarrow a$ gilt aufgrund der Grenzwertsätze für Folgen 11.15 auch $|a_n| \longrightarrow |a|$.

B) Rechnen mit stetigen Funktionen

Proposition 14.7 (Rechenregeln für stetige Funktionen)
Seien $f : U \longrightarrow \mathbb{R}$ und $g : U \longrightarrow \mathbb{R}$ Funktionen, die in $a \in U$ stetig sind, und $c \in \mathbb{R}$.

a. $c \cdot f$, $f + g$, $f - g$ und $f \cdot g$ sind stetig in a.

b. Ist $g(a) \neq 0$, so ist auch $\frac{f}{g} : U \setminus \{x \in U \mid g(x) = 0\} \longrightarrow \mathbb{R}$ stetig in a.

Beweis: Der Beweis folgt aus dem Folgenkriterium für Stetigkeit 14.5 und den Grenzwertsätzen für Folgen 11.15.

Z.B. sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in U mit $a_n \longrightarrow a$, dann gilt

$$(f + g)(a_n) = f(a_n) + g(a_n) \longrightarrow f(a) + g(a) = (f + g)(a),$$

da f und g in a stetig sind. Also ist auch $f + g$ stetig in a. □

Proposition 14.8 (Komposition stetiger Funktionen)
Es seien $f : U \longrightarrow \mathbb{R}$ und $g : V \longrightarrow \mathbb{R}$ Funktionen mit $\mathrm{Im}(f) \subseteq V$ und es sei $a \in U$. Ist f stetig in a und g stetig in $f(a)$, so ist $g \circ f$ stetig in a.

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in U mit $a_n \longrightarrow a$, dann ist $(f(a_n))_{n\in\mathbb{N}}$ eine Folge in V und, da f stetig in a ist, gilt zudem $f(a_n) \longrightarrow f(a)$. 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 14.5 ist dann $g \circ f$ stetig in a. □

Beispiel 14.9

Ist $f : U \longrightarrow \mathbb{R}$ stetig in $a \in U$, so ist auch $|f| : U \longrightarrow \mathbb{R} : x \mapsto |f(x)|$ als Komposition stetiger Funktionen stetig in a.

Definition 14.10 (Stetig fortsetzbar)

Es sei $f : U \longrightarrow \mathbb{R}$ eine stetige Funktion und $a \in \mathbb{R} \setminus U$ ein Häufungspunkt von U. Wir nennen f in a *stetig fortsetzbar*, wenn $\lim\limits_{x \to a} f(x)$ existiert.
In dieser Situation nennen wir

$$g : U \cup \{a\} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } x \neq a, \\ \lim\limits_{z \to a} f(z), & \text{falls } x = a, \end{cases}$$

die *stetige Fortsetzung* von f, und g ist nach Lemma 14.3 stetig in a und damit stetig auf $U \cup \{a\}$.

Beispiel 14.11

a. Die Funktion

$$f : \mathbb{R} \setminus \{1\} \longrightarrow \mathbb{R} : x \mapsto \frac{x^2 - 1}{x - 1}$$

aus Beispiel 13.8 c. ist in $a = 1$ 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 $a = 0$ 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 $a = 0$ nicht stetig fortsetzbar, obwohl der Grenzwert von f in 0 existiert, da 0 bereits zum Definitionsbereich der Funktion gehört!

C) Der Zwischenwertsatz

Satz 14.12 (Zwischenwertsatz)

Eine stetige Funktion $f : [a, b] \longrightarrow \mathbb{R}$ nimmt jeden Wert zwischen $f(a)$ und $f(b)$ an.

Beweis: Für den Beweis können wir $f(a) \leq f(b)$ annehmen. Für $c \in [f(a), f(b)]$ definieren wir eine Funktion

$$g : [a, b] \longrightarrow \mathbb{R} : x \mapsto f(x) - c,$$

und diese ist aufgrund der Proposition 14.7 stetig auf $[a, b]$.

Wir müssen zeigen, daß g eine Nullstelle in $[a, b]$ besitzt.

Dazu wenden wir wie im Beweis des Satzes von Bolzano-Weierstraß 11.26 ein Intervallschachtelungsverfahren an. Wir setzen

$$[a_0, b_0] := [a, b]$$

und betrachten den Punkt

$$x_0 = \frac{a_0 + b_0}{2} \in [a, b].$$

Ist $g(x_0) = 0$, so sind wir fertig. Andernfalls gilt entweder $g(x_0) > 0$ und wir setzen $[a_1, b_1] := [a_0, x_0]$, oder es gilt $g(x_0) < 0$ und wir setzen $[a_1, b_1] := [x_0, b_0]$.

Mit dem neuen Intervall verfahren wir wie mit dem vorherigen. Auf dem Weg finden wir entweder nach endlich vielen Schritten einen Punkt $x_n \in [a, b]$ mit $g(x_n) = 0$, oder wir konstruieren rekursiv eine monoton steigende, beschränkte Folge $(a_n)_{n\in\mathbb{N}}$ in $[a, b]$ und eine monoton fallende, beschränkte $(b_n)_{n\in\mathbb{N}}$ in $[a, b]$ mit

$$b_n - a_n = \frac{b - a}{2^n} \longrightarrow 0. \tag{24}$$

Aufgrund des Monotoniekriteriums für Folgen konvergiert $(a_n)_{n\in\mathbb{N}}$ gegen einen Wert x und $(b_n)_{n\in\mathbb{N}}$ gegen einen Wert y, und wegen (24) gilt dann

$$x = y.$$

Da das Intervall $[a, b]$ abgeschlossen ist, gilt zudem nach Satz 11.28

$$x \in [a, b].$$

Man beachte auch, daß aufgrund der Konstruktion von $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ stets

$$g(a_n) < 0 \quad \text{und} \quad g(b_n) > 0.$$

Für die stetige Funktion g folgt dann aus dem Folgenkriterium 14.5 und Satz 11.17

$$g(x) = \lim_{n\to\infty} g(a_n) \leq 0 \leq \lim_{n\to\infty} g(b_n) = g(x),$$

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 $\lim_{x\to\infty} f(x)$ und $\lim_{x\to-\infty} f(x)$ verschiedene Vorzeichen haben, so daß es $a, b \in \mathbb{R}$ mit $f(a) > 0$ und $f(b) < 0$ geben muß. Wenden wir dann den Zwischenwertsatz auf $f_{|[a,b]}$ bzw. $f_{|[b,a]}$ an, so folgt die Behauptung.

D) Beschränktheit stetiger Funktionen

Definition 14.14 (Beschränkte Funktionen)
Eine Funktion $f : U \longrightarrow \mathbb{R}$ heißt *beschränkt*, wenn $\mathrm{Im}(f)$ beschränkt ist.

Proposition 14.15 (Beschränktheit stetiger Funktionen)
Eine stetige Funktion $f : [a, b] \longrightarrow \mathbb{R}$ ist beschränkt.

Beweis: Nehmen wir an, f wäre nicht beschränkt. Dann gibt es für jedes $n \in \mathbb{N}$ ein $a_n \in [a, b]$ mit

$$|f(a_n)| > n.$$

Die Folge $(a_n)_{n\in\mathbb{N}}$ ist beschränkt, da sie im abgeschlossenen Intervall $[a, b]$ liegt, und nach dem Satz von Bolzano-Weierstraß 11.26 gibt es also eine konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ mit Grenzwert c, d.h.

$$a_{n_k} \longrightarrow c.$$

Da das Intervall $[a, b]$ abgeschlossen ist, gilt nach Satz 11.28

$$c \in [a, b].$$

Da f und somit nach Beispiel 14.9 auch $|f|$ stetig auf $[a, b]$ ist, folgt

$$|f(c)| \longleftarrow |f(a_{n_k})| \geq n_k \longrightarrow \infty,$$

was ein offensichtlicher Widerspruch ist. □

Satz 14.16 (Maximum / Minimum stetiger Funktionen)
Eine stetige Funktion $f : [a, b] \longrightarrow \mathbb{R}$ 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

$$A := \mathrm{Im}(f) = \{f(x) \mid x \in [a, b]\}$$

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

$$y \longleftarrow y - \frac{1}{n_k} < f(a_{n_k}) \leq y \longrightarrow y,$$

so daß aufgrund des Einschachtelungssatzes 11.17 auch

$$f(a_{n_k}) \longrightarrow y$$

gilt. Da f aber stetig ist, folgt dann

$$f(d) = \lim_{k\to\infty} f(a_{n_k}) = y.$$

Die Existenz von c zeigt man analog mit Hilfe von $\inf(A)$. □

Beispiel 14.17

a. Die Funktion $f : [-1,1] \longrightarrow \mathbb{R} : x \mapsto x^2$ ist beschränkt, und es gilt $f(0) = 0$ ist das Minimum und $f(1) = f(-1) = 1$ ist das Maximum von $\mathrm{Im}(f)$.

b. Die Funktion $f : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x}$ ist nicht beschränkt und nimmt weder ihr Minimum noch ihr Maximum an.

E) Umkehrsatz für streng monotone stetige Funktionen

Definition 14.18 (Monotone Funktionen)

Es sei $f : U \longrightarrow \mathbb{R}$ 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 $f : [0,\infty) \longrightarrow \mathbb{R} : x \mapsto x^n$ ist für jedes $n \geq 1$ streng monoton wachsend, da nach Lemma 8.14 aus $0 \leq x < y$ stets $x^n < y^n$ folgt.

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 monoton wachsend, aber nicht streng monoton wachsend.

Bemerkung 14.20

Ist $f : U \longrightarrow \mathbb{R}$ 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 $a < b$, $f : (a, b) \longrightarrow \mathbb{R}$ sei eine Funktion und es seien $c := \inf(\mathrm{Im}(f)) \in \mathbb{R} \cup \{-\infty\}$ und $d := \sup(\mathrm{Im}(f)) \in \mathbb{R} \cup \{\infty\}$.

a. Ist f streng monoton wachsend und stetig, so gelten:
   (i) $f : (a, b) \longrightarrow (c, d)$ ist bijektiv.
   (ii) $f^{-1} : (c, d) \longrightarrow (a, b)$ ist streng monoton wachsend und stetig.

b. Ist f streng monoton fallend und stetig, so gelten:
   (i) $f : (a, b) \longrightarrow (c, d)$ ist bijektiv.
   (ii) $f^{-1} : (c, d) \longrightarrow (a, b)$ 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: $c, d \notin \mathrm{Im}(f)$**:** Wäre $d \in \mathrm{Im}(f)$, so würde es ein $x \in (a, b)$ geben mit $f(x) = d$. Wegen $x < b$ gibt es ein $x' \in (a, b)$ mit $x < x'$ und somit

$$d = f(x) < f(x') \in \mathrm{Im}(f),$$

im Widerspruch dazu, daß d das Supremum von $\mathrm{Im}(f)$ ist. Analog sieht man, daß $c \notin \mathrm{Im}(f)$.

Zeige: $\mathrm{Im}(f) = (c, d)$**:** Nach Definition von $c = \inf(\mathrm{Im}(f))$ und $d = \sup(\mathrm{Im}(f))$ sowie nach der obigen Vorüberlegung folgt für $y \in \mathrm{Im}(f)$ sofort $c < y < d$, d.h.

$$\mathrm{Im}(f) \subseteq (c, d).$$

Sei nun $y \in (c, d)$. Wegen $y > c = \inf(\mathrm{Im}(f))$ gibt es ein $x_1 \in (a, b)$ mit $y > f(x_1)$, und wegen $y < d = \sup(\mathrm{Im}(f))$ 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.

$$(c,d) \subseteq \mathrm{Im}(f).$$

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 $f : (a,b) \longrightarrow \mathbb{R}$ im Umkehrsatz 14.21 streng monoton wachsend, so ist

$$c = \inf(\mathrm{Im}(f)) = \lim_{x \to a} f(x) \quad \text{und} \quad d = \sup(\mathrm{Im}(f)) = \lim_{x \to b} f(x),$$

und ist f streng monoton fallend, so ist

$$c = \inf(\mathrm{Im}(f)) = \lim_{x\to b} f(x) \quad \text{und} \quad d = \sup(\mathrm{Im}(f)) = \lim_{x\to a} f(x).$$

Außerdem, falls f stetig in $a \in \mathbb{R}$ bzw. in $b \in \mathbb{R}$ fortgesetzt werden kann, so ist $\lim\limits_{x\to a} f(x) \in \mathbb{R}$ bzw. $\lim\limits_{x\to b} f(x) \in \mathbb{R}$ und f^{-1} wird durch

$$f^{-1}\left(\lim_{x\to a} f(x)\right) = a \quad \text{bzw.} \quad f^{-1}\left(\lim_{x\to b} f(x)\right) = b$$

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

$$d := \sup(\mathrm{Im}(f)) \in \mathbb{R} \cup \{\infty\}.$$

1. **Fall:** $d \in \mathbb{R}$**:** Zu $\varepsilon > 0$ gibt es ein $y \in \mathrm{Im}(f)$ 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 $\lim\limits_{x\to b} f(x) = d$ gezeigt.
2. **Fall:** $d = \infty$**:** Zu $s > 0$ gibt es dann ein $y \in \mathrm{Im}(f)$ 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 $\lim\limits_{x\to b} f(x) = d$ gezeigt.

Läßt sich nun zudem f in $b \in \mathbb{R}$ stetig fortsetzen, so heißt dies, daß der Grenzwert

$$d := \lim_{x\to b} f(x) \in \mathbb{R}$$

in $\mathbb{R}$ liegt. Da f^{-1} stetig und streng monton wachsend auf (c, d) ist, gilt zudem

$$b = \lim_{x\to d} f^{-1}(x),$$

und somit läßt sich f^{-1} in d durch $f^{-1}(d) = b$ stetig fortsetzen. □

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

$$\inf(\mathrm{Im}(f)) = 0 \quad \text{und} \quad \sup(\mathrm{Im}(f)) = \infty.$$

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 9.8 für die Existenz von n-ten Wurzeln!

Man beachte zudem, daß wegen Bemerkung 14.22

$$\lim_{x \to 0} \sqrt[n]{x} = 0$$

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
$\lim_{n \to \infty} \sqrt[n]{n} = 1.$

Beweis: Wir müssen zeigen, daß die Folge $(a_n)_{n \geq 2}$ mit $a_n := \sqrt[n]{n} - 1$ 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 $a_n > 0$. Aus dem Binomischen Lehrsatz 7.15 folgt damit

$$n = (a_n + 1)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot a_n^k \cdot 1^{n-k} \geq 1 + \frac{n \cdot (n-1)}{2} \cdot a_n^2 > 1,$$

oder alternativ

$$0 \longleftarrow \frac{2}{n} \geq a_n^2 > 0.$$

Der Einschachtelungssatz 11.17 bedingt dann, daß

$$a_n^2 \longrightarrow 0,$$

und da die Wurzelfunktion stetig in 0 ist, folgt damit

$$a_n = \sqrt{a_n^2} \longrightarrow \sqrt{0} = 0.$$

□

F) Gleichmäßige Stetigkeit

Bemerkung 14.25 (Stetigkeit auf U)
Wir erinnern uns, eine Funktion $f : U \longrightarrow \mathbb{R}$ heißt *stetig auf* U, wenn sie in jedem Punkt $a \in U$ stetig ist, d.h.

$$\forall\, a \in U\ \forall\, \varepsilon > 0\ \exists\, \delta_{\varepsilon,a} > 0\ :\ \forall\, x \in U \text{ mit } |x-a| < \delta_{\varepsilon,a} \text{ gilt } |f(x)-f(a)| < \varepsilon.$$

Wir schreiben diesmal $\delta_{\varepsilon,a}$ statt δ_ε, um zu verdeutlichen, daß wir bei gegebenem $\varepsilon > 0$ zwar in jedem Punkt a ein geeignetes δ finden müssen, daß dieses δ sich mit dem Punkt a aber ändern kann! Es hängt also vom Punkt 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 $f : U \longrightarrow \mathbb{R}$ heißt *gleichmäßig stetig* auf U, wenn

$$\forall\, \varepsilon > 0\ \exists\, \delta_\varepsilon > 0\ :\ \forall\, x, y \in U \text{ mit } |x-y| < \delta_\varepsilon \text{ gilt } |f(x)-f(y)| < \varepsilon.$$

Bemerkung 14.27
Offenbar ist jede auf U gleichmäßig stetige Funktion $f : U \longrightarrow \mathbb{R}$ auch stetig auf U.

Satz 14.28
Eine stetige Funktion $f : [a,b] \longrightarrow \mathbb{R}$ ist gleichmäßig stetig auf $[a,b]$.

Beweis: Angenommen, f wäre nicht gleichmäßig stetig auf $[a,b]$. Dann gilt:

$$\exists\, \varepsilon > 0\ :\ \forall\, \delta_\varepsilon > 0\ :\ \exists\, x_{\delta_\varepsilon}, y_{\delta_\varepsilon} \in [a,b] \text{ mit } |x_{\delta_\varepsilon} - y_{\delta_\varepsilon}| < \delta_\varepsilon, \text{ aber } |f(x_{\delta_\varepsilon}) - f(y_{\delta_\varepsilon})| \geq \varepsilon.$$

Für $n \geq 1$ und $\delta_\varepsilon := \frac{1}{n}$ setzen wir $a_n := x_{\delta_\varepsilon} = x_{\frac{1}{n}}$ und $b_n := y_{\delta_\varepsilon} = y_{\frac{1}{n}}$. Damit erhalten wir zwei beschränkte Folgen $(a_n)_{n\geq 1}$ und $(b_n)_{n\geq 1}$ in $[a,b]$. Nach dem Satz von Bolzano-Weierstraß 11.26 besitzt $(a_n)_{n\geq 1}$ eine konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$, und ebenso besitzt dann $(b_{n_k})_{k\in\mathbb{N}}$ eine konvergente Teilfolge $(b_{n_{k_l}})_{l\in\mathbb{N}}$. Nach Konstruktion gilt

$$0 \leq |a_{n_{k_l}} - b_{n_{k_l}}| \leq \frac{1}{n_{k_l}} \longrightarrow 0,$$

so daß die Grenzwerte von $(a_{n_{k_l}})_{l\in\mathbb{N}}$ und $(b_{n_{k_l}})_{l\in\mathbb{N}}$ wegen des Einschachtelungssatzes 11.17 übereinstimmen müssen, d.h.

$$a_{n_{k_l}} \longrightarrow y \quad \text{und} \quad b_{n_{k_l}} \longrightarrow y.$$

Da das Intervall $[a,b]$ abgeschlossen ist, gilt nach Satz 11.28 zudem

$$y \in [a,b].$$

Da f und die Betragsfunktion stetig sind, folgt damit

$$0 = |f(y) - f(y)| \longleftarrow |f(a_{n_{k_l}}) - f(b_{n_{k_l}})| \geq \varepsilon,$$

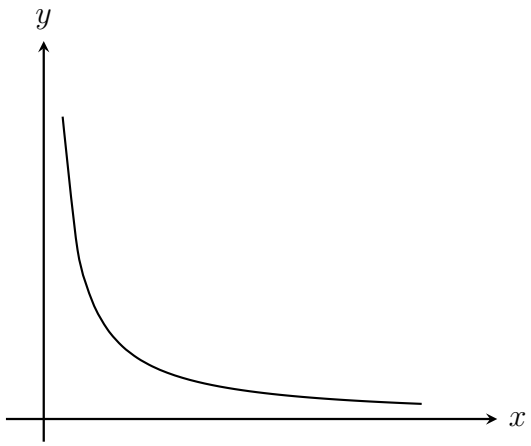
was ein offensichtlicher Widerspruch ist. □

Beispiel 14.29

a. Die Funktion $f : [0,1] \longrightarrow \mathbb{R} : x \mapsto x^2$ 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 $f : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x}$ 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

Entscheide, ob die folgende Funktion in $a = -2$ stetig ist

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \frac{4x^2+x+1}{1-x}, & \text{falls } x < -2, \\ |x-1|, & \text{falls } x \geq -2. \end{cases}$$

Aufgabe 14.31

Bestimme eine reelle Zahl a, so daß die folgende Funktion stetig ist:

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 5x - 1, & \text{falls } x < a, \\ x + 7, & \text{falls } x \geq a. \end{cases}$$

Aufgabe 14.32

Bestimme eine reelle Zahl b, so daß die folgende Funktion stetig ist:

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} x^2 - b, & \text{falls } x < 0, \\ \sqrt{x} + b, & \text{falls } x \geq 0. \end{cases}$$

Aufgabe 14.33

Sei $f : U \longrightarrow \mathbb{R}$ stetig, $a \in U$ und $b \in \mathbb{R}$.

a. Zeige, ist $f(a) > b$, so gibt, es ein $\delta > 0$, so dass $f(x) > b$ für alle $x \in U \cap (a - \delta, a + \delta)$.

b. Zeige, ist $f(a) \neq b$, so gibt es ein $\delta > 0$, so dass $f(x) \neq b$ für alle $x \in U \cap (a - \delta, a + \delta)$.

Aufgabe 14.34

Die Wurzelfunktion $\sqrt{\cdot} : [0, \infty) \longrightarrow \mathbb{R} : x \mapsto \sqrt{x}$ ist gleichmäßig stetig auf $[0, \infty)$.

Aufgabe 14.35

Verwende die ϵ-δ-Definition der Stetigkeit, um zu zeigen, dass die Funktion $f : [0, 1] \longrightarrow \mathbb{R}$, $x \longmapsto \sqrt{1 - x^3}$ stetig in $[0, 1]$ ist.

Aufgabe 14.36 (Lipschitz-Stetigkeit)
Sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ 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 14.37
Sei $f : [0,1] \longrightarrow \mathbb{R}$ eine Funktion definiert durch

$$f(x) := \begin{cases} \frac{2^n}{n!} & \text{für } x = \frac{1}{n} \text{ mit } n \geq 1 \\ 0 & \text{für } x \in [0,1] \backslash \{\frac{1}{n} \mid n \geq 1\} \end{cases}.$$

Bestimme (mit Beweis) sämtliche Punkte auf $[0,1]$, in denen f stetig ist.

Aufgabe 14.38
Sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ stetig und $a, b \in \mathbb{R}$ mit $f(a) \neq b$. Zeige, es gibt ein $\delta > 0$, so dass $f(x) \neq b$ für alle $x \in (a-\delta, a+\delta)$.

Aufgabe 14.39 (Fixpunktsatz von Banach)
Sei $f : [a,b] \longrightarrow \mathbb{R}$ eine stetige Abbildung mit $\mathrm{Im}(f) \subseteq [a,b]$. Zeige, daß f einen Fixpunkt hat, d.h. es gibt ein $c \in [a,b]$ mit $f(c) = c$.

Aufgabe 14.40
Zeige, ist $f : [0,1] \longrightarrow \mathbb{R}$ eine stetige Funktion mit $f(0) = f(1)$. Zeige, dann gibt es ein $a \in [0, \frac{1}{2}]$ mit $f(a) = f(a + \frac{1}{2})$.

Aufgabe 14.41
Sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ eine stetige Abbildung und $a \in \mathbb{R}_{>0}$ mit $f(x) = f(x+a)$ für alle $x \in \mathbb{R}$. Zeige, daß es ein $b \in (0,a)$ gibt mit $f(b + \frac{a}{2}) = f(b)$.

Aufgabe 14.42 (Stetige Fortsetzbarkeit)

a. Sei $f : (a,b] \longrightarrow \mathbb{R}$ eine stetige Funktion. Zeige, dass f genau dann stetig in a fortsetzbar ist, wenn f gleichmäßig stetig ist.

b. Gibt es eine beschränkte Funktion $f : (0,1] \longrightarrow \mathbb{R}$, die sich nicht stetig in 0 fortsetzen lässt?

Aufgabe 14.43
Es seien $f : [a,b] \longrightarrow \mathbb{R}$ stetig auf $[a,b]$ und $g : [b,c] \longrightarrow \mathbb{R}$ stetig auf $[b,c]$ mit

$f(b) = g(b)$, so ist auch die Funktion

$$h : [a, c] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} g(x), & \text{falls } x \leq b, \\ h(x), & \text{falls } x > b \end{cases}$$

stetig auf $[a, c]$.

Aufgabe 14.44

Es sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ eine Funktion und $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}$, so daß die Folge $(f(x_n))_{n\in\mathbb{N}}$ konvergiert.

a. Zeige durch ein Beispiel, daß die Folge $(x_n)_{n\in\mathbb{N}}$ nicht konvergieren muß.

b. Zeige, wenn f stetig und streng monoton ist, dann ist $(x_n)_{n\in\mathbb{N}}$ konvergent.

Aufgabe 14.45

Betrachte die Funktion

$$f : (0, \infty) \longrightarrow \mathbb{R} : x \mapsto \sin\left(\frac{1}{x}\right).$$

a. Ist f stetig auf $(0, \infty)$?

b. Ist f beschränkt auf $(0, \infty)$?

c. Ist f gleichmäßig stetig auf $(0, \infty)$?

Aufgabe 14.46

Zeige, die Funktion $f : [0, \infty) \longrightarrow \mathbb{R} : x \mapsto \sqrt{x}$ ist gleichmäßig stetig.

Aufgabe 14.47

Zeige, ist $f : [0, \infty) \longrightarrow [0, \infty)$ gleichmäßig stetig mit $f(0) = 0$, so gibt es eine Konstante $K > 0$ mit

$$f(x) \leq 1 + K \cdot x$$

für alle $x \in [0, \infty)$.

Aufgabe 14.48 (Thomaesche Funktion, die nur an $R \setminus \mathbb{Q}$ stetig ist.)

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 $N(0) = 1$. Zeige,

$$f : [0, 1] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \frac{1}{N(x)}, & \text{falls } x \in \mathbb{Q} \text{ und } x \neq 0, \\ 0, & \text{sonst,} \end{cases}$$

ist genau dann stetig in $x \in [0, 1]$, wenn x irrational ist.

§ 15 Konvergenz von Funktionenfolgen

A) Punktweise & gleichmäßige Konvergenz von Funktionenfolgen

Definition 15.1 (Konvergenz von Funktionenfolgen)

a. Für jedes $n \in \mathbb{N}$ sei $f_n : U \longrightarrow \mathbb{R}$ 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 $\lim\limits_{n\to\infty} f_n(x)$ existiert, d.h. In diesem Fall nennen wir die Funktion

$$f : U \longrightarrow \mathbb{R} : x \mapsto \lim_{n\to\infty} f_n(x)$$

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

$$f = \lim_{n\to\infty} f_n.$$

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 $f_n : [0,1] \longrightarrow \mathbb{R} : x \mapsto x^n$ konvergiert auf $[0,1]$ punktweise gegen die Funktion

$$f : [0,1] \longrightarrow \mathbb{R} : x \mapsto \lim_{n\to\infty} x^n = \begin{cases} 0, & \text{falls } x < 1, \\ 1, & \text{falls } x = 1. \end{cases}$$

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 $x = \frac{1}{\sqrt[n]{2}} \in [0,1)$, 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

$$f_n : (-r, r) \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{n} a_k \cdot x^k.$$

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. □

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 $\lim\limits_{x\to 1} g(x) = \infty$, 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.

B) Gleichmäßiger Grenzwert stetiger Funktionen

Satz 15.6 (Der gleichmäßige Grenzwert stetiger Funktionen ist stetig.)
Ist $f_n : U \longrightarrow \mathbb{R}$ stetig auf U für $n \in \mathbb{N}$ und konvergiert $(f_n)_{n\in\mathbb{N}}$ auf U gleichmäßig gegen f, so ist f stetig auf U.

Beweis: Seien $a \in U$ und $\varepsilon > 0$ gegeben. Da die f_n gleichmäßig gegen f konvergieren, gilt:

$$\exists\, n_\varepsilon \in \mathbb{N} \;:\; \forall\, n \geq n_\varepsilon \text{ und } \forall\, x \in U \;:\; |f_n(x) - f(x)| < \frac{\varepsilon}{3}.$$

Da zudem f_{n_ε} stetig in a ist, gilt:

$$\exists\, \delta_\varepsilon > 0 \quad \forall\, x \in U \text{ mit } |x-a| < \delta_\varepsilon \text{ gilt } |f_{n_\varepsilon}(x) - f_{n_\varepsilon}(a)| < \frac{\varepsilon}{3}.$$

Sei nun $x \in U$ mit $|x-a| < \delta_\varepsilon$ gegeben, so gilt

$$|f(x)-f(a)| \leq |f(x)-f_{n_\varepsilon}(x)| + |f_{n_\varepsilon}(x) - f_{n_\varepsilon}(a)| + |f_{n_\varepsilon}(a) - f(a)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Mithin ist f stetig in 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 $R := \frac{|x|+r}{2}$. 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

Wir betrachten die Funktionenfolge $(f_n)_{n\geq 1}$ mit

$$f_n : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \frac{1}{1 + |x|^n}.$$

a. Zeige, die Funktionenfolge konvergiert punktweise gegen eine Grenzfunktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ und gib die Werte $f(1)$, $f(\frac{1}{2})$ und $f(2)$ an.

b. Skizziere den Graphen der Grenzfunktion f.

c. Welche der folgenden Grenzwerte existieren?

   (i) $\lim_{x\to 1} f(x)$ für $x \in (1, \infty)$.

   (ii) $\lim_{x\to 1} f(x)$ für $x \in (-1, 1)$.

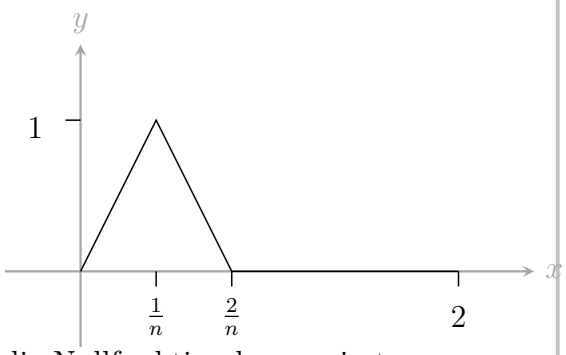
   (iii) $\lim_{x\to 1} f(x)$ für $x \in \mathbb{R}$.

d. An welchen Stellen $a \in \mathbb{R}$ ist die Funktion f stetig.

Aufgabe 15.12

Zeige, daß die Folge $(f_n)_{n\geq 1}$ von Funktionen $f_n : [0, 2] \longrightarrow \mathbb{R}$, 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.

Aufgabe 15.13

Es sei $g : \mathbb{R} \longrightarrow \mathbb{R}_{\geq 0}$ eine beschränkte, nicht-negative Funktion. Zeige, dass die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{\infty} \frac{g(2^k \cdot x)}{2^k}$$

wohldefiniert und stetig ist.

Aufgabe 15.14 a. Zeige, eine Folge von beschränkten Funktionen $f_n : [a, b] \longrightarrow \mathbb{R}$ konvergiert genau dann gleichmäßig gegen eine Grenzfunktion $f : [a, b] \longrightarrow \mathbb{R}$, wenn

$$\lim_{n\to\infty} \sup_{x\in[a,b]} |f_n(x) - f(x)| = 0.$$

b. Untersuche die Funktionenfolge $(f_n)_{n\geq 1}$ mit $f_n : [0, 1] \longrightarrow \mathbb{R} : x \mapsto e^{\frac{x}{n}}$ auf punktweise und gleichmäßige Konvergenz

Aufgabe 15.15

Es sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von monoton wachsenden Funktionen $f_n : [a, b] \longrightarrow \mathbb{R}$, die punktweise gegen eine stetige Funktion $f : [a, b] \longrightarrow \mathbb{R}$ konvergieren. Zeige, dann konvergiert die Folge schon gleichmäßig gegen f.

§ 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

$$\lim_{x \to -\infty} \exp(x) = 0 \quad \text{und} \quad \lim_{x \to \infty} \exp(x) = \infty.$$

Beweis: Für $z \in \mathbb{R}$ mit $z > 0$ gilt offenbar

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!} \geq \frac{z^1}{1!} + \frac{z^0}{0!} = z + 1 > 1, \tag{25}$$

und mit Hilfe der Funktionalgleichung in Satz 12.36 folgt dann

$$\exp(-z) \cdot \exp(z) = \exp(-z + z) = \exp(0) = 1$$

sowie

$$\exp(-z) = \frac{1}{\exp(z)} > 0, \tag{26}$$

die Exponentialfunktion nimmt also nur positive Werte an. Wenden wir die Funktionalgleichung noch mal für $x, y \in \mathbb{R}$ mit $x < y$ wie folgt an

$$\exp(y) = \exp(y - x + x) = \exp(y - x) \cdot \exp(x) \overset{(25),(26)}{>} 1 \cdot \exp(x) = \exp(x),$$

so erhalten wir, 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

$$c = \inf(\mathrm{Im}(f)) \overset{14.22}{=} \lim_{x \to -\infty} \exp(x)$$

und

$$d = \sup(\mathrm{Im}(f)) \overset{14.22}{=} \lim_{x \to \infty} \exp(x).$$

Nun gilt für $x > 0$ aber

$$\exp(x) \overset{(25)}{\geq} x + 1 \overset{x \to \infty}{\longrightarrow} \infty,$$

so daß $d = \lim_{x \to \infty} \exp(x) = \infty$ folgt.

Mit (26) und aus den Grenzwertsätzen für uneigentliche Grenzwerte von Funktionen folgt zudem

$$\exp(-x) = \frac{1}{\exp(x)} \overset{x\to\infty}{\longrightarrow} \frac{1}{\infty} = 0,$$

d.h. $c = \lim\limits_{x\to-\infty} \exp(x) = 0$. □

Bemerkung 16.2

Die Zahl $e = \exp(1)$ ist irrational und es gilt $2 < e < 3$ (siehe auch Beispiel 18.36).

Beweis: Aus (25) wissen wir

$$\exp(1) > 1 + 1 = 2.$$

Wegen

$$\frac{1}{k!} = \frac{1}{1\cdot 2\cdot 3\cdot \ldots \cdot k} \le \frac{1}{1\cdot 2\cdot 2\cdot \ldots \cdot 2} = \frac{1}{2^{k-1}}$$

für $k \ge 1$ folgt unter Berücksichtigung der endlichen geometrischen Reihe 7.12

$$\begin{aligned} s_n &:= \sum_{k=0}^{n} \frac{1}{k!} \le 1 + 1 + \frac{1}{2} + \frac{1}{6} + \sum_{k=4}^{n} \frac{1}{2^{k-1}} \\ &= \frac{8}{3} + \frac{1}{8}\cdot \sum_{k=0}^{n-4} \frac{1}{2^k} = \frac{8}{3} + \frac{1}{8}\cdot \frac{1-\left(\frac{1}{2}\right)^{n-3}}{1-\frac{1}{2}} \\ &\le \frac{8}{3} + \frac{1}{8}\cdot \frac{1}{1-\frac{1}{2}} = \frac{8}{3} + \frac{1}{4} = \frac{35}{12} \end{aligned}$$

für alle $n \in \mathbb{N}$. Grenzwertbildung liefert deshalb $\exp(1) = \lim\limits_{n\to\infty} s_n \le \frac{35}{12} < 3$.

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 $q \ge 2$ sein. Wir betrachten nun die Zahlen

$$a := \frac{q!}{0!} + \frac{q!}{1!} + \frac{q!}{2!} + \ldots + \frac{q!}{q!} \in \mathbb{Z}$$

und

$$b := \sum_{n=q+1}^{\infty} \frac{q!}{n!} = q!\cdot e - a = (q-1)!\cdot p - a \in \mathbb{Z}. \tag{27}$$

Da $q \ge 2$ ist, folgt für $n > q$

$$\frac{q!}{n!} = \frac{1}{(q+1)\cdot(q+2)\cdots n} \le \frac{1}{3\cdot 3\cdots 3} = \frac{1}{3^{n-q}}$$

und deshalb

$$b = \sum_{n=q+1}^{\infty} \frac{q!}{n!} \le \sum_{n=q+1}^{\infty} \frac{1}{3^{n-q}} = \sum_{n=1}^{\infty} \frac{1}{3^n} = \frac{1}{1-\frac{1}{3}} - 1 = \frac{1}{2}$$

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 (27). Der Widerspruch kommt von unserer Annahme, daß e eine rationale Zahl wäre. □

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

$$\lim_{x \to 0} \ln(x) = -\infty \quad \text{und} \quad \lim_{x \to \infty} \ln(x) = \infty.$$

Beweis: Die Aussagen folgen aus dem Umkehrsatz für streng monotone Funktionen 14.21 und Satz 16.1 unter Berücksichtigung von Bemerkung 14.22. □

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.

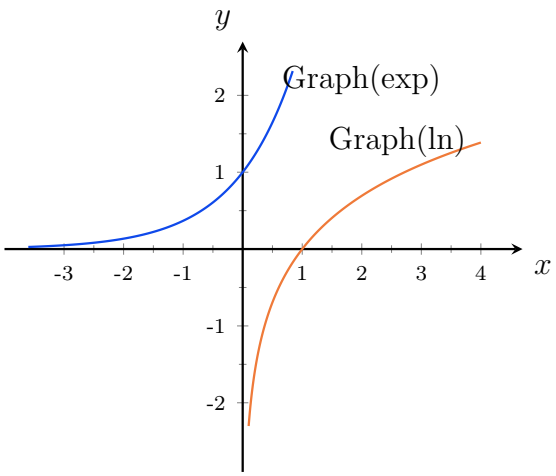


Abbildung 7: Die Graphen von exp und ln

Definition 16.6 (Exponentialfunktion zur Basis a)
Für $a \in \mathbb{R}$ mit $a > 0$ und $x \in \mathbb{R}$ definieren wir

$$a^x := \exp\big(x \cdot \ln(a)\big).$$

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

$$\exp_a(x) = \exp\big(x \cdot \ln(a)\big) < \exp\big(y \cdot \ln(a)\big) = \exp_a(y)$$

folgt. $\exp_a$ ist dann also streng monoton wachsend. Außerdem gilt

$$\lim_{x \to \infty} x \cdot \ln(a) = \infty \quad \text{sowie} \quad \lim_{x \to -\infty} x \cdot \ln(a) = -\infty,$$

und da exp stetig ist folgt dann auch

$$\lim_{x \to \infty} \exp_a(x) = \lim_{x \to \infty} \exp\big(x \cdot \ln(a)\big) = \infty$$

sowie

$$\lim_{x \to -\infty} \exp_a(x) = \lim_{x \to -\infty} \exp\big(x \cdot \ln(a)\big) = 0.$$

Aus dem Umkehrsatz für streng monotone Funktionen 14.21 folgt dann, weil $\exp_a$ stetig, und streng monoton wachsend ist, daß $\exp_a$ auch bijektiv 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. □

B) Potenz- und Logarithmusgesetze

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 9.8 und 16.6 überein.

Beweis:

a. Mit Hilfe der Funktionalgleichung für die Exponentialfunktion sieht man:

$$\begin{aligned} a^{x+y} &= \exp\big((x+y) \cdot \ln(a)\big) = \exp\big(x \cdot \ln(a) + y \cdot \ln(a)\big) \\ &= \exp\big(x \cdot \ln(a)\big) \cdot \exp\big(y \cdot \ln(a)\big) = a^x \cdot a^y. \end{aligned}$$

b. Wegen $a^x = \exp(x \cdot \ln(a))$ gilt auch

$$\ln(a^x) = \ln\big(\exp(x \cdot \ln(a))\big) = x \cdot \ln(a)$$

und damit

$$(a^x)^y = \exp\big(y \cdot \ln(a^x)\big) = \exp\big(x \cdot y \cdot \ln(a)\big) = a^{x \cdot y}.$$

c. Wir verwenden in der folgenden Gleichung bereits ein Logarithmusgesetz 16.9, dessen Beweis unabhängig von diesem Potenzgesetz ist:

$$\begin{aligned} (a \cdot b)^x &= \exp\big(x \cdot \ln(a \cdot b)\big) \overset{16.9b.}{=} \exp\big(x \cdot (\ln(a) + \ln(b))\big) \\ &= \exp\big(x \cdot \ln(a)\big) \cdot \exp\big(x \cdot \ln(b)\big) = a^x \cdot b^x. \end{aligned}$$

d. Die Gleichung $a^{-x} = \frac{1}{a^x}$ folgt unmittelbar aus

$$a^x \cdot a^{-x} = a^{x-x} = a^0 = \exp(0 \cdot \ln(a)) = \exp(0) = 1.$$

e. Mit Induktion nach $n \geq 1$ und a. sieht man, daß $a^n = \prod_{k=1}^{n} a$. Wir haben bereits gesehen, daß zudem $a^0 = 1$ gilt, und aus d. folgt dann

$$a^{-n} = \frac{1}{a^n} = \frac{1}{\prod_{k=1}^{n} a} = \prod_{k=1}^{n} \frac{1}{a}.$$

Die Definitionen von a^n in 7.9 und 16.6 stimmen also überein.

f. Mit Satz 9.8 folgt $a^{\frac{p}{q}} = \sqrt[q]{a^p}$ aus

$$\left(a^{\frac{p}{q}}\right)^q \overset{b.}{=} a^{\frac{p}{q}\cdot q} = a^p.$$

□

Korollar 16.9 (Logarithmusgesetze)
Seien $a, x, y \in \mathbb{R}_{>0}$ mit $a \neq 1$ und $z \in \mathbb{R}$.

a. $\log_a(x) = \frac{\ln(x)}{\ln(a)}$.

b. $\log_a(x \cdot y) = \log_a(x) + \log_a(y)$.

c. $\log_a(x^z) = z \cdot \log_a(x)$.

d. $\log_a(\frac{x}{y}) = \log_a(x) - \log_a(y)$.

Beweis:

a. Falls $a \neq 1$, so gilt

$$\exp_a\left(\frac{\ln(x)}{\ln(a)}\right) = \exp\left(\frac{\ln(x)}{\ln(a)} \cdot \ln(a)\right) = \exp\big(\ln(x)\big) = x = \exp_a\big(\log_a(x)\big),$$

und da $\exp_a$ injektiv ist, gilt dann auch

$$\frac{\ln(x)}{\ln(a)} = \log_a(x).$$

b. Es gilt

$$\begin{aligned}\exp_a\big(\log_a(x \cdot y)\big) &= x \cdot y = \exp_a\big(\log_a(x)\big) \cdot \exp_a\big(\log_a(y)\big) \\ &\overset{16.8a.}{=} \exp_a\big(\log_a(x) + \log_a(y)\big),\end{aligned}$$

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

$$x^z = \exp\big(z \cdot \ln(x)\big) \overset{a.}{=} \exp\big(z \cdot \log_a(x) \cdot \ln(a)\big) = \exp_a\big(z \cdot \log_a(x)\big)$$

definiert. Wenden wir auf beiden Seiten die Funktion $\log_a$ an, so erhalten wir

$$\log_a\big(x^z\big) = \log_a\big(\exp_a\big(z \cdot \log_a(x)\big)\big) = z \cdot \log_a(x).$$

d. Es gilt

$$\log_a\left(\frac{x}{y}\right) \overset{b.}{=} \log_a(x) + \log_a(y^{-1}) \overset{c.}{=} \log_a(x) - \log_a(y).$$

□

C) Die Zahl π

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

$$x - \frac{x^3}{6} = x + s_1 \leq \sin(x) \leq x + s_4. \tag{28}$$

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

$$\sin(4) \leq 4 + s_4(4) = -\frac{268}{405} < 0.$$

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

$$\pi := \inf(A).$$

Nach Bemerkung 11.22 gibt es eine monoton fallende Folge $(b_n)_{n\in\mathbb{N}}$ in A, die gegen das Infimum $\pi = \inf(A)$ konvergiert. Da der Sinus stetig ist, gilt dann auch

$$0 = \sin(b_n) \longrightarrow \sin(\pi),$$

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 (28)

$$\sin(x) \geq x - \frac{x^3}{6} > x - \frac{x}{6} \geq \frac{5x}{6} > 0,$$

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. □

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 18.44), daß man die Zahl π approximieren kann durch

$$3,14159\ldots.$$

D) Der Cosinus

Satz 16.12 (Monotonie des Cosinus)
Der Cosinus $\cos : [0, \pi] \longrightarrow [-1, 1]$ 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

$$\cos(x) - \cos(y) = 2 \cdot \sin\left(\frac{y+x}{2}\right) \cdot \sin\left(\frac{y-x}{2}\right) \overset{16.10}{>} 0,$$

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

$$\cos : [0,\pi] \longrightarrow [\lim_{x\to\pi}\cos(x), \lim_{x\to 0}\cos(x)]$$

bijektiv ist. Da der Cosinus stetig ist, gilt nun

$$\lim_{x\to 0}\cos(x) = \cos(0) = \sum_{n=0}^{\infty}(-1)^n \cdot \frac{0^{2n}}{(2n)!} = 1$$

und

$$\lim_{x\to\pi}\cos(x) = \cos(\pi).$$

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. □

E) Eigenschaften des Sinus' und des Cosinus'

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	$\frac{3\cdot\pi}{2}$	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.

$$\sin(x)=0\quad\Longleftrightarrow\quad x\in\{k\cdot\pi\mid k\in\mathbb{Z}\},$$

und für den Cosinus gilt mithin

$$\cos(x)=0\quad\Longleftrightarrow\quad x\in\left\{\frac{\pi}{2}+k\cdot\pi\;\middle|\;k\in\mathbb{Z}\right\}.$$

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

$$|\sin(x)|\le\sqrt{\cos(x)^2+\sin(x)^2}=1$$

und

$$|\cos(x)|\le\sqrt{\cos(x)^2+\sin(x)^2}=1.$$

c. Die Werte für $x=0$ folgen unmittelbar aus der Definition von Sinus und Cosinus als Potenzreihen

$$\sin(0)=\sum_{n=0}^{\infty}(-1)^n\cdot\frac{0^{2n+1}}{(2n+1)!}=0$$

und

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

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

$$0 \leq \cos\left(\frac{\pi}{2}\right)^2 = \sin\left(\frac{\pi}{2}\right)^2 - 1 \overset{b.}{\leq} 0.$$

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. Den Beweis überlassen wir dem Leser als Übungsaufgabe.

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

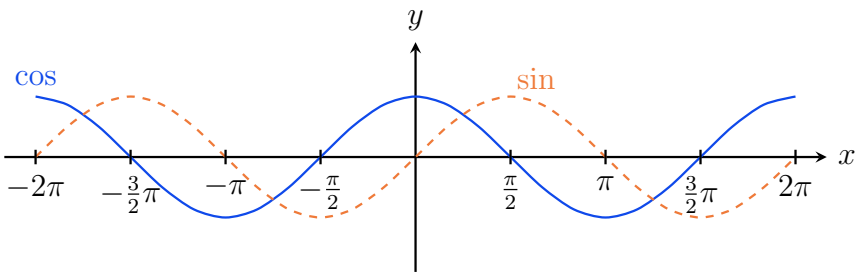
$$0 \leq x - k \cdot \pi < \pi,$$

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:



F) Polarkoordinaten

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 Hypothenuse 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 8).

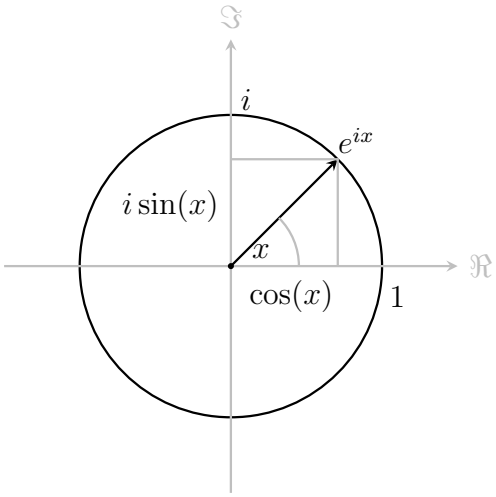


Abbildung 8: Polarkoordinaten von $e^{i\cdot x} = \cos(x) + i \cdot \sin(x)$

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 10.11 gezeigt, daß nämlich jede komplexe Zahl eine n-te Wurzel besitzt, da wir dazu nur die Polarkoordinatendarstellung von z benötigt haben.

Man beachte, daß für $n \geq 2$ die Zahlen

$$e^{\frac{2\cdot k\cdot \pi\cdot i}{n}} \quad \text{mit} \quad k = 0, \ldots, n-1$$

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.

G) Weitere trigonometrische Funktionen

Definition 16.16 (Tangens und Cotangens)

Die Funktion

$$\tan : \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k \cdot \pi \;\middle|\; k \in \mathbb{Z} \right\} \longrightarrow \mathbb{R} : x \mapsto \frac{\sin(x)}{\cos(x)}$$

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*.

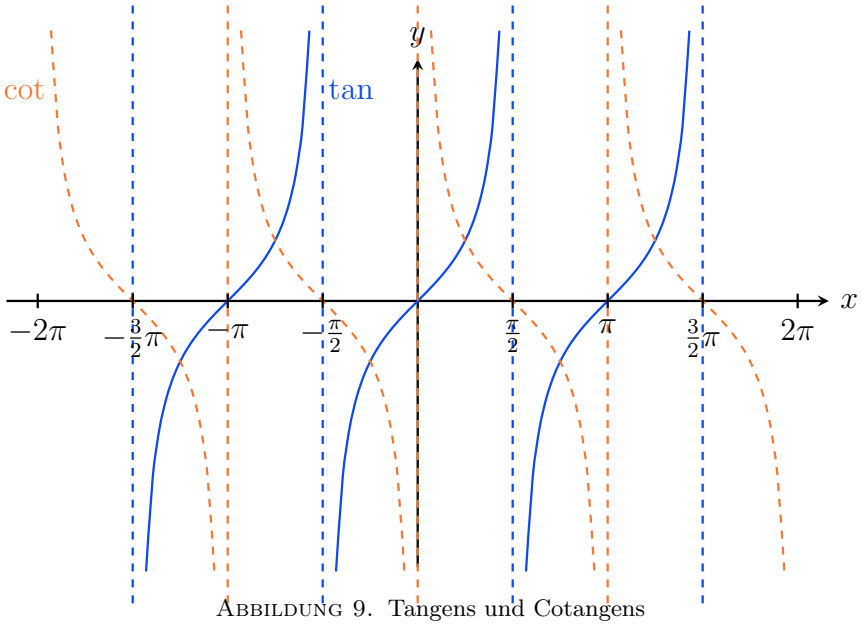


Abbildung 9. Tangens und Cotangens

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 $\left(-\frac{\pi}{2}+k\cdot\pi, \frac{\pi}{2}+k\cdot\pi\right)$, $k\in\mathbb{Z}$, 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 $0\le x<y<\frac{\pi}{2}$ 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 $\left[0,\frac{\pi}{2}\right)$ also streng monoton wachsend. Wegen $\tan(-x)=-\tan(x)$ ist der Tangens aber punktsymmetrisch zum Ursprung und somit auch streng monoton wachsend auf $\left(-\frac{\pi}{2},0\right]$, also streng monoton wachsend auf $\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$.
Wegen der Stetigkeit von Sinus und Cosinus erhalten wir zudem

$$\lim_{x\to\frac{\pi}{2}}\tan(x)=\frac{\lim\limits_{x\to\frac{\pi}{2}}\sin(x)}{\lim\limits_{x\to\frac{\pi}{2}}\cos(x)}=\frac{1}{0}=\infty$$

und

$$\lim_{x\to-\frac{\pi}{2}}\tan(x)=\frac{\lim\limits_{x\to-\frac{\pi}{2}}\sin(x)}{\lim\limits_{x\to-\frac{\pi}{2}}\cos(x)}=\frac{-1}{0}=-\infty.$$

Aus dem Umkehrsatz 14.21 folgt dann, daß die stetige Funktion

$$\tan:\left(-\frac{\pi}{2},\frac{\pi}{2}\right)\longrightarrow(-\infty,\infty)=\mathbb{R}$$

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. □

Aufgaben

Aufgabe 16.19

Löse die folgenden Gleichungen für $x \in \mathbb{R}$:

a. $\log_2(x) - \log_2(x-6) = 3$.

b. $4^x + 4 = 2^{x+2} + 2^x$.

c. $\log_4(x+2) - \log_4(x-2) = \frac{1}{2}$.

d. $2^{3-x} \cdot 3^{x-1} = 6^{2x-3}$.

Aufgabe 16.20

a. Zeige, ist $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge mit $0 \neq a_n \in \mathbb{R}_{>-1}$ für alle $n \in \mathbb{N}$, so gilt

$$\lim_{n\to\infty} (1+a_n)^{\frac{1}{a_n}} = e.$$

b. Für $x \in \mathbb{R}$ zeige $\lim_{n\to\infty} \left(1+\frac{x}{n}\right)^n = e^x$.

c. Für $a \in \mathbb{R}_{>0}$ zeige $\lim_{n\to\infty} n \cdot (\sqrt[n]{a} - 1) = \ln(a)$.

Aufgabe 16.21

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 16.22

Bestimme das Bild der folgenden Funktionen f und entscheide, ob diese injektiv sind. Falls die Umkehrfunktion $f^{-1} : \mathrm{Im}(f) \longrightarrow \mathbb{R}$ existiert, ist sie dann stetig?

a. $f : [-1, 1] \longrightarrow \mathbb{R} : x \mapsto \log_2(x^2 + 1)$.

b. $f : [0, \infty) \longrightarrow \mathbb{R} : x \mapsto \sum_{n=1}^{20} \frac{x}{n} \cdot e^{nx^2+n}$.

Aufgabe 16.23 (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

$$4 \cdot \arctan\left(\frac{1}{5}\right) - \arctan\left(\frac{1}{239}\right) = \frac{\pi}{4}.$$

Aufgabe 16.24 (Rechenregeln für die Arcusfunktionen)

Beweise die folgenden Gleichungen:

a. $\sin\big(\arccos(x)\big) = \sqrt{1 - x^2}$ für $x \in [-1, 1]$.

b. $\cos\big(\arcsin(x)\big) = \sqrt{1 - x^2}$ für $x \in [-1, 1]$.

c. $\sin\big(\arctan(x)\big) = \frac{x}{\sqrt{1+x^2}}$ für $x \in \mathbb{R}$.

d. $\cos\big(\arctan(x)\big) = \frac{1}{\sqrt{1+x^2}}$ für $x \in \mathbb{R}$.

e. $\arcsin(x) = \arctan\left(\frac{x}{\sqrt{1-x^2}}\right)$ für $x \in (-1, 1)$.

f. $\arccos(x) = \frac{\pi}{2} - \arctan\left(\frac{x}{\sqrt{1-x^2}}\right)$ für $x \in (-1, 1)$.

g. $\arcsin(x) = 2 \cdot \arctan\left(\frac{x}{1+\sqrt{1-x^2}}\right)$ für $x \in (-1, 1)$.

Aufgabe 16.25

a. Zeige, für alle $x \in \mathbb{R}$ gilt die Gleichung

$$\sin(3x) = 3 \cdot \sin(x) - 4 \cdot \sin^3(x).$$

b. Zeige mit Hilfe der Formel in Teil a. die Gleichung

$$\sin\left(\frac{\pi}{6}\right) = \frac{1}{2}.$$

Aufgabe 16.26 (Eingangsfrage aus dem JEE advanced 2022)

Berechne die folgenden Ausdruck

$$\frac{3}{2} \cdot \arccos\left(\sqrt{\frac{2}{2+\pi^2}}\right) + \frac{1}{4} \cdot \arcsin\left(\frac{2 \cdot \sqrt{2} \cdot \pi}{2+\pi^2}\right) + \arctan\left(\frac{\sqrt{2}}{\pi}\right).$$

Aufgabe 16.27

Computer speichern Daten in Form von Bits, d.h. in Variablen, die nur die Wert 0 oder 1 annehmen können.

a. Wie viele Bits benötigt ein Computer, um die Zahl 65^{365} zu speichern?

b. Ein Algorithmus benötige bei einem Eingabedatum mit n Bits $n \log_2(n)$ Rechenschritte. Berechne die Laufzeit in Rechenschritten für die Wert $n = 10$, $n = 256$ und $n = 1000$.

c. Was ist die kleinste Eingabe in Bits, für die der Algorithmus mindestens 100 Rechenschritte benötigt?

§ 17 Differenzierbarkeit

A) Der Differenzenquotient

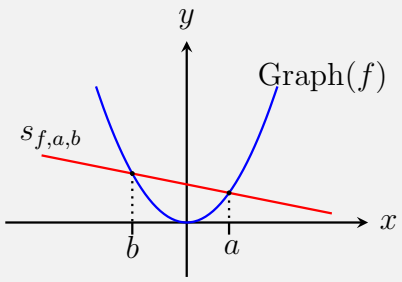
Definition 17.1 (Differenzenquotient)
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ und $a \in U$. Die Funktion

$$\mathrm{Diff}_{f,a} : U \setminus \{a\} \longrightarrow \mathbb{R} : x \mapsto \frac{f(x) - f(a)}{x - a}$$

heißt der *Differenzenquotient* von f an der Stelle a.
Für ein festes b ist der Wert des Differenzenquotienten $\mathrm{Diff}_{f,a}(b)$ 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

$$\begin{aligned} y =& \frac{f(b) - f(a)}{b - a} \cdot x + \frac{f(a) \cdot b - f(b) \cdot a}{b - a} \\ =& \frac{f(b) - f(a)}{b - a} \cdot x + f(a) - \frac{f(b) - f(a)}{b - a} \cdot a \end{aligned} \tag{29}$$

gegeben ist.



Beispiel 17.2
Ist $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^n$ mit $n \geq 1$, so ist

$$\mathrm{Diff}_{f,a}(x) = \frac{x^n - a^n}{x - a} = x^{n-1} + a \cdot x^{n-2} + a^2 \cdot x^{n-3} + \ldots + a^{n-2} \cdot x + a^{n-1}$$

für $x \in \mathbb{R} \setminus \{a\}$.

B) Differenzierbarkeit und die Ableitung

Definition 17.3
Es sei $U \subseteq \mathbb{R}$, $f : U \longrightarrow \mathbb{R}$ eine Funktion und $a \in U$. Wir nennen f *differenzierbar in* a, wenn a ein Häufungspunkt von U ist und der Grenzwert

$$\lim_{x \to a} \mathrm{Diff}_{f,a}(x) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \in \mathbb{R}$$

des Differenzenquotienten in a existiert. In diesem Fall schreiben wir

$$f'(a) := \frac{\partial f}{\partial x}(a) := \lim_{x \to a} \frac{f(x) - f(a)}{x - a},$$

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

$$f' : U \longrightarrow \mathbb{R} : x \mapsto f'(x)$$

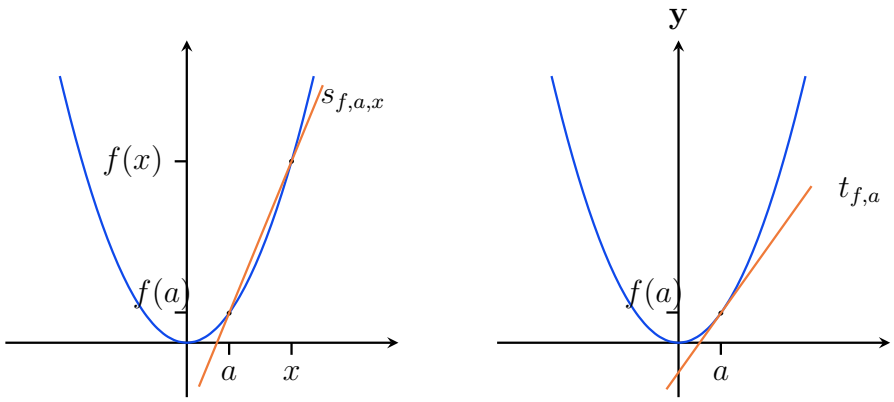
die *Ableitung* von f. Beachte auch, daß dann insbesondere jeder Punkt von U ein Häufungspunkt von U sein muß!

Bemerkung 17.4

Der Definition liegt die Idee zugrunde, daß sich die Sekante $s_{f,a,x}$ für $x \to a$ 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 $x \to a$ gegen die Steigung der Tangenten. D.h. die Tangente an den Graphen von f im Punkt $(a, f(a))$ hat die Geradengleichung

$$y = f'(a) \cdot x + \big(f(a) - a \cdot f'(a)\big) = f(a) + f'(a) \cdot (x - a),$$

die sich aus (29) ergibt, indem man den Grenzwert für $b = x$ gegen a betrachtet.



Beispiel 17.5

Die folgenden Funktionen sind alle auf ganz $\mathbb{R}$ definiert, und dort ist jeder Punkt ein Häufungspunkt!

a. Eine konstante Funktion $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto c$ ist differenzierbar auf $\mathbb{R}$ und die Ableitung ist die Nullfunktion

$$f' : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto 0,$$

da für jedes $a \in \mathbb{R}$ der Differenzenquotient $\mathrm{Diff}_{f,a}$ die Nullfunktion ist und somit der Grenzwert $f'(a) = 0$ existiert.

b. Für $n \in \mathbb{N}$ ist die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^n$$

differenzierbar auf $\mathbb{R}$ mit Ableitung

$$f' : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto n \cdot x^{n-1},$$

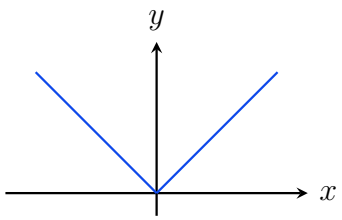
da sich für $a \in \mathbb{R}$ aus Beispiel 17.2 folgendes ergibt:

$$f'(a) = \lim_{x \to a} \left(x^{n-1} + a \cdot x^{n-2} + a^2 \cdot x^{n-3} + \ldots + a^{n-2} \cdot x + a^{n-1} \right) = n \cdot a^{n-1}.$$

c. Die Betragsfunktion

$$|\cdot| : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto |x|$$

ist in $a = 0$ *nicht* differenzierbar. In jedem anderen Punkt a ist sie jedoch differenzierbar mit $f'(a) = -1$ falls $a < 0$ und $f'(a) = 1$ falls $a > 0$.



Anschaulich bedeutet die Nicht-Differenzierbarkeit im Punkt $a = 0$, daß man am Graphen im Ursprung keine klare Tangente findet.
Um die Nicht-Differenzierbarkeit in $a = 0$ zu sehen, betrachten wir die Nullfolge $\left(\frac{(-1)^n}{n}\right)_{n\geq 1}$. Die zugehörige Folge der Werte des Differenzenquotienten

$$\left(\mathrm{Diff}_{f,0} \left(\frac{(-1)^n}{n} \right) \right)_{n\geq 1} = \left(\frac{\frac{1}{n}}{\frac{(-1)^n}{n}} \right)_{n\geq 1} = \left((-1)^n\right)_{n\geq 1}$$

ist nicht konvergent. Mithin existiert der Grenzwert des Differenzenquotienten in $a = 0$ nicht, und somit ist die Funktion in $a = 0$ nicht differenzierbar.
Außerdem, ist $a < 0$ und x nahe bei a, so ist auch $x < 0$ und mithin

$$\mathrm{Diff}_{f,a}(x) = \frac{|x| - |a|}{x - a} = \frac{-x + a}{x - a} = -1 \xrightarrow{x \to a} -1,$$

und analog ist für $a > 0$ und x nahe bei a auch $x > 0$, so daß

$$\mathrm{Diff}_{f,a}(x) = \frac{|x| - |a|}{x - a} = \frac{x - a}{x - a} = 1 \xrightarrow{x \to a} 1.$$

Damit ist auch gezeigt, daß die Ableitung in allen Punkten $a \neq 0$ existiert.

Bemerkung 17.6

a. 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 a ab!

b. Ist a ein Häufungspunkt von U und $f : U \longrightarrow \mathbb{R}$, so ist f genau dann in a differenzierbar, wenn der Grenzwert

$$\lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \in \mathbb{R}$$

existiert.
Um dies zu sehen ersetzt man im Differenzenquotienten einfach $x-a$ durch h.

c. Ist a ein Häufungspunkt von U und $f : U \longrightarrow \mathbb{R}$, so ist f genau dann in a differenzierbar, wenn es eine Zahl $c \in \mathbb{R}$ und eine Funktion $\rho : U \longrightarrow \mathbb{R}$ gibt, so daß

$$f(x) = f(a) + c\cdot(x-a) + \rho(x) \quad \text{und} \quad \lim_{x\to a} \frac{\rho(x)}{|x-a|} = 0$$

gilt.
Man beachte, ist f differenzierbar in a, so wählt man $c = f'(a)$ und $\rho(x) = \big(\operatorname{Diff}_{f,a}(x) - f'(a)\big)\cdot(x-a)$. Umgekehrt, wenn c und ρ existieren, so ist $\operatorname{Diff}_{f,a}(x) = c + \frac{\rho(x)}{x-a}$ 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 $\lim_{x\to a} \frac{\rho(x)}{|x-a|} = 0$ ist, daß die Funktion ρ, die den Unterschied des Differenzenquotienten und der Ableitung beschreibt, sehr schnell gegen Null konvergiert für $x \to a$, jedenfalls schneller als die lineare Funktion $x-a$.

C) Differenzierbarkeit und Stetigkeit

Satz 17.7 (Differenzierbar impliziert stetig.)
Ist $f : U \longrightarrow \mathbb{R}$ differenzierbar in a, so ist f stetig in a.

Beweis: Da nach Voraussetzung a ein Häufungspunkt von U ist, müssen wir nach Lemma 14.3 nur zeigen, daß $\lim_{x\to a} f(x) = f(a)$ oder alternativ $\lim_{x\to a}(f(x)-f(a)) = 0$ gilt. Nach Voraussetzung existiert der Grenzwert

$$f'(a) = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} \in \mathbb{R},$$

und da die Identität stetig ist, gilt zudem $\lim_{x\to a}(x-a) = 0$. Mithin erhalten wir aus den Grenzwertsätzen für Funktionen 13.10

$$\begin{aligned}\lim_{x\to a}\big(f(x)-f(a)\big) &= \lim_{x\to a} \frac{f(x)-f(a)}{x-a}\cdot(x-a)\\ &= \lim_{x\to a} \frac{f(x)-f(a)}{x-a}\cdot\lim_{x\to a}(x-a) = f'(a)\cdot 0 = 0.\end{aligned}$$

Also ist f stetig in a. □

Beispiel 17.8

a. Die Umkehrung von Satz 17.7 gilt nicht, wie das Beispiel der Betragsfunktion zeigt, die stetig in $a = 0$ ist (siehe Beispiel 14.6), ohne dort differenzierbar zu sein (siehe Beispiel 17.5).

b. Die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} x^2, & \text{wenn } x \notin \mathbb{Q}, \\ 0, & \text{wenn } x \in \mathbb{Q} \end{cases}$$

ist im Punkt $a = 0$ differenzierbar, weil

$$\mathrm{Diff}_{f,0}(x) = \begin{cases} \frac{x^2-0}{x-0} = x, & \text{wenn } x \notin \mathbb{Q}, \\ \frac{0-0}{x-0} = 0, & \text{wenn } x \in \mathbb{Q} \end{cases}$$

und deshalb

$$\lim_{x\to 0} \mathrm{Diff}_{f,0}(x) = 0$$

existiert. Die Funktion ist aber in keinem anderen Punkt $0 \neq a \in \mathbb{R}$ differenzierbar, weil sie in den anderen Punkten nicht mal stetig ist; denn für $a \neq 0$ enthält jede noch so kleine Umgebung von a Punkte, an denen der Funktionswert 0 ist und Punkte, an denen der Funktionswert größer als a^2 ist. f ist somit ein Beispiel für einen Funktion, die in genau einem Punkt differenzierbar ist.

D) Ableitungsregeln – Linearität, Produkt- und Quotientenregel

Proposition 17.9 (Linearität der Ableitung)
Seien $f : U \longrightarrow \mathbb{R}$ und $g : U \longrightarrow \mathbb{R}$ in $a \in U$ differenzierbar und sei $c, d \in \mathbb{R}$. Dann ist $c \cdot f + d \cdot g$ differenzierbar in a mit $(c \cdot f + d \cdot g)'(a) = c \cdot f'(a) + d \cdot g'(a)$.

Beweis: Wir beachten zunächst, daß nach Voraussetzung a ein Häufungspunkt von U ist und daß U jeweils der Definitionsbereich der Funktionen ist. Dann folgt aus den Grenzwertsätzen für Funktionen 13.10, daß der Grenzwert

$$\begin{aligned}(c \cdot f + d \cdot g)'(a) =& \lim_{x\to a} \frac{c \cdot f(x) + d \cdot g(x) - c \cdot f(a) - d \cdot g(a)}{x - a} \\ =& c \cdot \lim_{x\to a} \frac{f(x) - f(a)}{x - a} + d \cdot \lim_{x\to a} \frac{g(x) - g(a)}{x - a} \\ =& c \cdot f'(a) + d \cdot g'(a)\end{aligned}$$

existiert. □

Beispiel 17.10 (Polynomfunktionen sind differenzierbar.)

Ist $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{n} a_k \cdot x^k$ eine Polynomfunktion, so ist f differenzierbar auf $\mathbb{R}$

mit

$$f' : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{k=1}^{n} k \cdot a_k \cdot x^{k-1}.$$

Dies folgt unmittelbar aus Proposition 17.9 und Beispiel 17.5.

Proposition 17.11 (Produktregel)
Seien $f : U \longrightarrow \mathbb{R}$ und $g : U \longrightarrow \mathbb{R}$ in $a \in U$ differenzierbar, so ist $f \cdot g$ differenzierbar in a mit

$$(f \cdot g)'(a) = f'(a) \cdot g(a) + f(a) \cdot g'(a).$$

Beweis: Wir beachten, daß nach Voraussetzung a ein Häufungspunkt von U ist und daß U der Definitionsbereich von $f \cdot g$ ist. Der Differenzenquotient von $f \cdot g$ an der Stelle a genügt der Gleichung

$$\begin{aligned}
\mathrm{Diff}_{f\cdot g,a}(x) =& \frac{f(x) \cdot g(x) - f(a) \cdot g(a)}{x - a} \\
=& \frac{f(x) \cdot g(x) - f(a) \cdot g(x) + f(a) \cdot g(x) - f(a) \cdot g(a)}{x - a} \\
=& \frac{f(x) \cdot g(x) - f(a) \cdot g(x)}{x - a} + \frac{f(a) \cdot g(x) - f(a) \cdot g(a)}{x - a} \\
=& \frac{f(x) - f(a)}{x - a} \cdot g(x) + f(a) \cdot \frac{g(x) - g(a)}{x - a}.
\end{aligned}$$

Da f und g in a differenzierbar sind und da g nach Satz 17.7 zudem stetig in a ist, existiert damit der Grenzwert

$$(f \cdot g)'(a) = \lim_{x \to a} \mathrm{Diff}_{f\cdot g,a}(x) = f'(a) \cdot g(a) + f(a) \cdot g'(a)$$

aufgrund der Grenzwertsätze für Funktionen 13.10. □

Proposition 17.12 (Quotientenregel)
Seien $f : U \longrightarrow \mathbb{R}$ und $g : U \longrightarrow \mathbb{R}$ in $a \in U$ differenzierbar mit $g(a) \neq 0$, so ist auch $\frac{f}{g} : \{x \in U \mid g(x) \neq 0\} \longrightarrow \mathbb{R}$ differenzierbar in a mit

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a) \cdot g(a) - f(a) \cdot g'(a)}{g(a)^2}.$$

Beweis: Wir müssen zunächst einmal zeigen, daß a ein Häufungspunkt der Menge

$$V := \{x \in U \mid g(x) \neq 0\}$$

ist. Wegen Satz 17.7 ist g stetig in a, und da a ein Häufungspunkt von U ist gilt somit

$$\lim_{x \to a} g(x) = g(a) \neq 0.$$

Dann folgt aus Proposition 13.10 c. aber bereits, daß a auch ein Häufungspunkt von V ist.

Ferner gilt für den Differenzenquotienten

$$\mathrm{Diff}_{\frac{1}{g},a}(x) = \frac{\frac{1}{g(x)} - \frac{1}{g(a)}}{x-a} = -\frac{g(x)-g(a)}{x-a} \cdot \frac{1}{g(x)\cdot g(a)}.$$

Da g in a differenzierbar und stetig ist, existiert damit der Grenzwert

$$\left(\frac{1}{g}\right)'(a) = \lim_{x\to a} \mathrm{Diff}_{\frac{1}{g},a}(x) = -\lim_{x\to a} \frac{g(x)-g(a)}{x-a} \cdot \lim_{x\to a} \frac{1}{g(x)\cdot g(a)} = -\frac{g'(a)}{g(a)^2}.$$

Wenden wir nun die Produktformel 17.11 auf $f \cdot \frac{1}{g}$ an, so folgt das Ergebnis. □

Aus der Quotientenregel und Beispiel 17.10 folgt die folgende Aussage.

Beispiel 17.13
Jede rationale Funktion $\frac{f}{g} : \mathbb{R} \setminus \{x \in \mathbb{R} \mid g(x) = 0\} \longrightarrow \mathbb{R}$ ist differenzierbar.
Z.B. gilt für $h : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x^n}$ mit $n \geq 1$ für die Ableitung

$$h' : \mathbb{R} \setminus \{0\} \longrightarrow \mathbb{R} : x \mapsto -\frac{n \cdot x^{n-1}}{x^{2n}} = -\frac{n}{x^{n+1}}.$$

□

E) Ableitungsregeln – Ableitung der Umkehrfunktion und Kettenregel

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 $f : I \longrightarrow \mathbb{R}$ sei stetig und streng monoton (wachsend oder fallend). Ist f differenzierbar in a und ist $f'(a) \neq 0$, so ist die Umkehrfunktion

$$f^{-1} : f(I) \longrightarrow I$$

differenzierbar in $b := f(a)$ und es gilt

$$\left(f^{-1}\right)'(b) = \frac{1}{f'(a)} = \frac{1}{f'(f^{-1}(b))}.$$

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 $(y_n)_{n\in\mathbb{N}}$ in $f(I) \setminus \{b\}$ mit $y_n \longrightarrow b$. Da f^{-1} stetig ist, gilt dann

$$x_n := f^{-1}(y_n) \longrightarrow f^{-1}(b) = a.$$

Da f^{-1} bijektiv ist, ist $(x_n)_{n\in\mathbb{N}}$ somit eine Folge in $I \setminus \{a\}$, die gegen a konvergiert. Aufgrund der Grenzwertsätze für Folgen 11.15 erhalten wir dann

$$\frac{f^{-1}(y_n) - f^{-1}(b)}{y_n - b} = \frac{x_n - a}{f(x_n) - f(a)} = \frac{1}{\mathrm{Diff}_{f,a}(x_n)} \longrightarrow \frac{1}{f'(a)},$$

und wegen des Folgenkriteriums für Grenzwerte 13.7 existiert somit der Grenzwert

$$\left(f^{-1}\right)'(b) = \lim_{y\to b} \mathrm{Diff}_{f^{-1},b}(y) = \lim_{y\to b} \frac{f^{-1}(y) - f^{-1}(b)}{y - b} = \frac{1}{f'(a)}.$$

□

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

$$\left(\sqrt[n]{\cdot}\right)' : (0,\infty) \to \mathbb{R} : y \mapsto \frac{1}{n \cdot (\sqrt[n]{y})^{n-1}} = \frac{1}{n} \cdot y^{\frac{1}{n}-1}.$$

Im Falle von $n = 2$ erhalten wir insbesondere

$$\left(\sqrt{y}\right)' = \frac{1}{2 \cdot \sqrt{y}}.$$

Wir wollen nun noch zeigen, daß die Wurzelfunktion $\sqrt[n]{\cdot}$ in $a = 0$ 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

$$\mathrm{Diff}_{\sqrt[n]{\cdot},0}\left(\frac{1}{k^n}\right) = \frac{\frac{1}{k} - 0}{\frac{1}{k^n} - 0} = k^{n-1} \longrightarrow \infty$$

für $k \longrightarrow \infty$, da $n \geq 2$. Also existiert der Grenzwert des Differenzenquotienten in $a = 0$ nicht, und somit ist die Wurzelfunktion dort auch nicht differenzierbar.

Proposition 17.16 (Kettenregel – äußere Ableitung × innere Ableitung)
Es seien $f : U \longrightarrow \mathbb{R}$ und $g : V \longrightarrow \mathbb{R}$ Funktionen mit $\mathrm{Im}(f) \subseteq V$ und es sei $a \in U$. Ist f differenzierbar in a und g differenzierbar in $f(a)$, so ist $g \circ f$ differenzierbar in a mit

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a).$$

Beweis: Wir definieren auf V eine Funktion durch

$$h : V \longrightarrow \mathbb{R} : y \mapsto \begin{cases} \mathrm{Diff}_{g,f(a)}(y) = \frac{g(y)-g(f(a))}{y-f(a)}, & \text{falls } y \neq f(a), \\ g'(f(a)), & \text{falls } y = f(a). \end{cases}$$

Da g in $f(a)$ differenzierbar ist, gilt dann $\lim\limits_{y \to f(a)} h(y) = h(f(a))$, d.h. h ist stetig in $f(a)$. Außerdem gilt für alle $y \in V$

$$h(y) \cdot \big(y - f(a)\big) = g(y) - g\big(f(a)\big). \tag{30}$$

Wir beachten nun noch, daß nach Proposition 14.8 $\lim\limits_{x \to a} h(f(x)) = h(f(a)) = g'(f(a))$ gilt, da die Funktion h stetig in $f(a)$ und die Funktion f nach Satz 17.7 stetig in a ist. Damit erhalten wir dann, daß der Grenzwert

$$\begin{aligned}(g \circ f)'(a) &= \lim_{x \to a} \frac{g(f(x)) - g(f(a))}{x - a} \overset{(30)}{=} \lim_{x \to a} \frac{h(f(x)) \cdot (f(x) - f(a))}{x - a} \\ &= \lim_{x \to a} h(f(x)) \cdot \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = g'(f(a)) \cdot f'(a)\end{aligned}$$

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.

F) 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 $f : U \longrightarrow \mathbb{R}$ eine Funktion.

a. Wir nennen f *stetig differenzierbar*, wenn f differenzierbar auf U und f' stetig auf U ist.

b. 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 $f^{(k)} := \left(f^{(k-1)}\right)'$ heißt dann die *k-te Ableitung* von f. Wir schreiben auch $f^{(0)} := f$, $f'' := f^{(2)}$ und $f''' := f^{(3)}$.

c. f heißt *k-fach stetig differenzierbar*, wenn f k-fach differenzierbar auf U und zudem $f^{(k)}$ stetig auf U ist. Mit

$$\mathcal{C}^k(U, \mathbb{R}) := \{f : U \longrightarrow \mathbb{R} \mid f \text{ ist } k\text{-fach stetig differenzierbar}\}$$

bezeichnen wir die Menge der k-fach stetig differenzierbaren Funktionen auf U.

d. f heißt *unendlich oft differenzierbar* auf U, wenn $f \in \mathcal{C}^k(U, \mathbb{R})$ für alle $k \geq 1$. Wir bezeichnen mit

$$\mathcal{C}^\infty(U, \mathbb{R}) := \{f : U \longrightarrow \mathbb{R} \mid f \text{ ist unendlich oft differenzierbar}\}$$

die Menge der unendlich oft differenzierbaren Funktionen auf U.

Beispiel 17.19 (Nicht stetig differenzierbare Funktion)

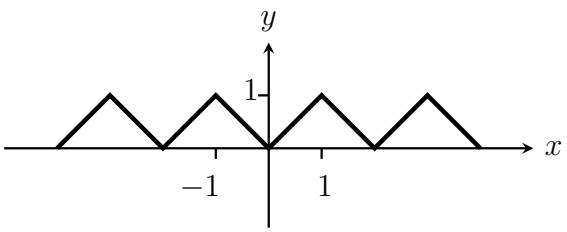
a. Die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} x^2 \cdot \sin\left(\frac{1}{x}\right), & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0, \end{cases}$$

ist differenzierbar auf $\mathbb{R}$, die Ableitung ist aber nicht stetig in $a = 0$. Der Beweis ist ein Übungsaufgabe, für die man unter anderem Korollar 18.22 benötigt.

b. Leitet man eine Polynomfunktion oder eine rationale Funktion ab, so erhält man wieder eine Polynomfunktion oder eine rationale Funktion mit dem jeweils gleichen Definitionsbreich. 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 $g : \mathbb{R} \longrightarrow \mathbb{R}$, 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 (siehe Aufgabe 15.13) und in keinem Punkt differenzierbar ist! Funktionen, die wie f gebildet werden, nennt man auch *Weiterstraß-Funktionen*.

Aufgaben

Aufgabe 17.21
Bestimme mittels der Ableitungsregeln die Ableitung der folgenden Funktionen:

a. $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto 10 + x \cdot \cos(x)$.

b. $g : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto e^{-3x} \cdot \left(x^3 - 2x\right)^2$.

Aufgabe 17.22
Für $n \in \{0, 1, 2\}$ sei

$$f_n : [0, \infty) \longrightarrow \mathbb{R},\ x \longmapsto \begin{cases} x^n \cdot \sin\left(\frac{1}{x}\right), & \text{für } x > 0, \\ 0, & \text{für } x = 0. \end{cases}$$

Welche der Funktionen sind stetig in $a = 0$, differenzierbar in $a = 0$, stetig differenzierbar auf $[0, \infty)$?

Aufgabe 17.23 (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:

$$(f \cdot g)^{(n)} = \sum_{k=0}^{n} \binom{n}{k} \cdot f^{(k)} \cdot g^{(n-k)}.$$

Aufgabe 17.24

Mit $\lfloor x \rfloor = \max\{z \in \mathbb{Z} \mid z \leq x\}$ sei die Abrundung der reellen Zahl $x \in \mathbb{R}$ bezeichnet. Skizziere den Graphen der Funktion

$$f : (-1,1) \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \frac{1}{\lfloor \frac{1}{x} \rfloor}, & \text{falls } x \neq 0, \\ 0, & \text{falls } x = 0, \end{cases}$$

schematisch, überprüfe, an welchen Stellen die Funktion differenzierbar ist, und bestimme dort die Ableitung.

Aufgabe 17.25

Bestimme für die folgenden Funktionen $f : I \longrightarrow \mathbb{R}$ ein maximales Intervall I, auf dem sie differenzierbar sind, und berechne ihre Ableitungen:

a. $f(x) = x^{\sqrt{x}}$.

b. $f(x) = \log_3(e^{2x+1})$.

c. $f(x) = x \cdot \ln(x) - x$.

d. $f(x) = (\cos(x))^{\sin(x)}$.

e. $f(x) = e^{3x} \cdot \big(\tan(x) - \sin(3x^2) + 3\big) \cdot x$.

f. $f(x) = e^{e^x}$.

Aufgabe 17.26

Für eine auf U differenzierbare Funktion $f : U \longrightarrow \mathbb{R}$ ohne Nullstellen nennen wir $L(f) := \frac{f'}{f}$ die logarithmische Ableitung von f.
Zeige die folgenden Rechenregeln für die logarithmische Ableitung:

a. $L(f_1 \cdot \ldots \cdot f_n) = L(f_1) + \ldots + L(f_n)$.

b. $L\left(\frac{f}{g}\right) = L(f) - L(g)$.

c. $L\big(\exp(f(x))\big) = f'(x)$.

Aufgabe 17.27

Es sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ eine Funktion, die in a differenzierbar ist und es seien $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$ zwei Folgen mit $x_n < a < y_n$ und $\lim_{n\to\infty} x_n = \lim_{n\to\infty} y_n = a$.
Dann gilt

$$\lim_{n\to\infty} \frac{f(y_n) - f(x_n)}{y_n - x_n} = f'(a).$$

Man zeige zudem, daß die Aussage i.a. nicht mehr richtig ist, wenn man auf die Voraussetzung $x_n < a < y_n$ verzichtet.

Aufgabe 17.28

Zeige, daß die Funktion f in Bemerkung 17.20 in keinem Punkt ihres Definitionsbereiches differenzierbar ist.

Aufgabe 17.29

Es sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ die Funktion aus Bemerkung 17.20 und

$$h : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x \cdot f(x).$$

Zeige, h ist stetig auf ganz $\mathbb{R}$, aber nur im Punkt $a = 0$ differenzierbar.

§ 18 Der Mittelwertsatz und seine Anwendungen

A) Notwendige Bedingung für Extremstellen

Definition 18.1 (Extremstellen)
Es sei $f : U \longrightarrow \mathbb{R}$ eine Funktion und $a \in U$.

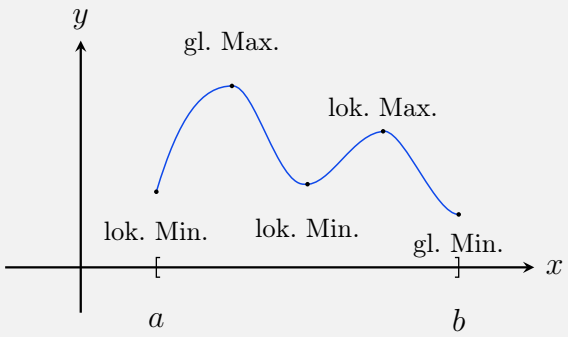
a. f hat in a ein *globales Maximum*, wenn $f(a) \geq f(x)$ für alle $x \in U$.

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 $f : (a, b) \longrightarrow \mathbb{R}$ 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:

$$0 \leq \frac{f(a_n) - f(c)}{a_n - c} \longrightarrow f'(c)$$

und

$$0 \geq \frac{f(b_n) - f(c)}{b_n - c} \longrightarrow f'(c).$$

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.

B) Der Satz von Rolle und der Mittelwertsatz

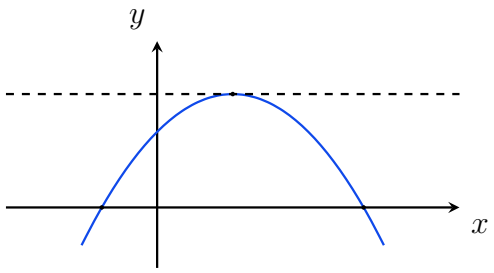
Satz 18.5 (Satz von Rolle)
Ist $a < b$ und ist $f : [a, b] \longrightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf (a, b) mit $f(a) = f(b)$, so gibt es ein $c \in (a, b)$ mit $f'(c) = 0$.

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 $f : [a, b] \longrightarrow \mathbb{R}$ 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

$$g : [a, b] \longrightarrow \mathbb{R} : x \mapsto f(x) - \frac{f(b) - f(a)}{b - a} \cdot (x - a)$$

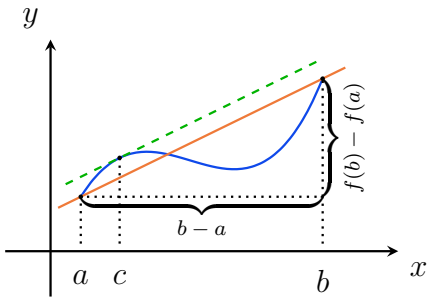
ist stetig auf $[a, b]$ und differenzierbar auf (a, b). Außerdem gilt $g(a) = f(a) = g(b)$. Aus dem Satz von Rolle 18.5 folgt die Existenz eines $c \in (a, b)$ mit

$$0 = g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}.$$

□

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

$$f : [-1, 1] \longrightarrow \mathbb{R} : x \mapsto x^3.$$

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 $f, g : [a, b] \longrightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf (a, b), so gibt es ein $c \in (a, b)$ mit $f'(c) \cdot \big(g(b) - g(a)\big) = g'(c) \cdot \big(f(b) - f(a)\big)$.

Beweis: Wir betrachten die Funktion

$$h : [a, b] \longrightarrow \mathbb{R} : x \mapsto f(x) \cdot \big(g(b) - g(a)\big) - g(x) \cdot \big(f(b) - f(a)\big).$$

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

$$0 = h'(c) = f'(c) \cdot \big(g(b) - g(a)\big) - g'(c) \cdot \big(f(b) - f(a)\big).$$

□

Bemerkung 18.11 (Geometrische Interpretation)
Wollen wir den allgemeinen Mittelwertsatz geometrisch interpretieren, brauchen wir einen Vorgriff auf die mehrdimensionale Analysis. Sind f und g zwei Funktionen wie in Korollar 18.10, dann definiert

$$\varphi : [a,b] \longrightarrow \mathbb{R}^2 : t \mapsto \big(f(t), g(t)\big)^t$$

eine differenzierbare Abbildung vom Intervall $[a,b]$ in die Ebene $\mathbb{R}^2$. Das Bild der Abbildung φ ist eine ebene Kurve mit den Endpunkten $(f(a), g(a))^t$ und $(f(b), g(b))^t$ (siehe Abbildung 10).

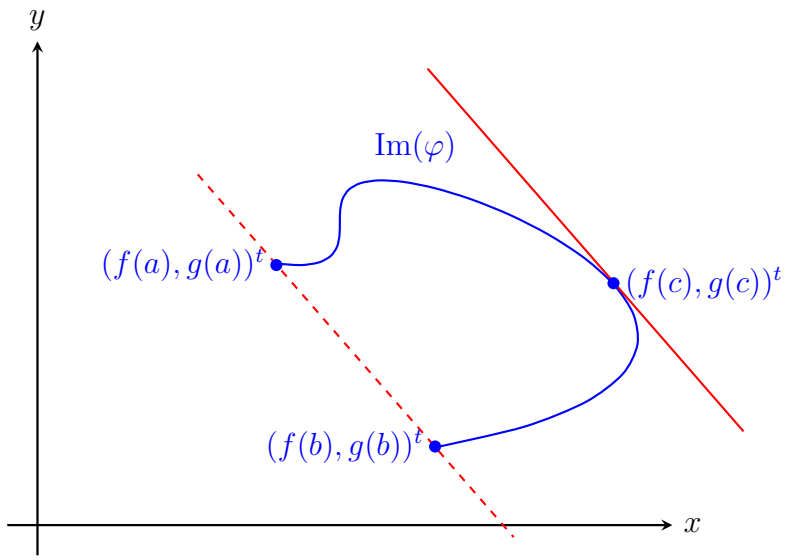


Abbildung 10: Geometrische Interpretation des Allgemeinen Mittelwertsatzes

Die Richtung der Tangente an die Kurve im Punkt $(f(c), g(c))^t$ kann man mit Hilfe des Gradienten bestimmen, denn wenn der Vektor

$$v(c) = \nabla\varphi(c) = \big(f'(c), g'(c)\big)^t$$

nicht der Nullvektor ist, ist er ein Richtungsvektor der Tagente. Schauen wir uns nun die Gleichung

$$\frac{g'(c)}{f'(c)} = \frac{g(b) - g(a)}{f(b) - f(a)}$$

an, die wir aus dem Allgemeinen Mittelwertsatz erhalten, so ist die linke Seite die Steigung der Tangente an die Kurve im Punkt $(f(c), g(c))^t$ und die rechte Seite ist die Steigung der Geraden durch die beiden Endpunkte der Geraden. Der Allgemeine Mittelwertsatz sagt mithin, dass es einen Punkt auf der Kurve gibt, in dem die Tangete parallel zur Geraden durch die beiden Endpunkte ist.

C) Anwendungen des Mittelwertsatzes

Wir wollen uns nun den Anwendungen des Mittelwertsatzes zuwenden.

C.1) Konstante Funktionen

Proposition 18.12
Es sei $f : [a, b] \longrightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf (a, b) mit $f'(x) = 0$ für alle $x \in (a, b)$, dann ist f eine konstante Funktion.

Beweis: Sei $x \in (a, b]$, so ist f stetig auf $[a, x]$ und differenzierbar auf (a, x). Aus dem Mittelwertsatz folgt dann, daß es ein $c \in (a, x)$ gibt mit

$$\frac{f(x) - f(a)}{x - a} = f'(c) = 0.$$

Also gilt $f(x) = f(a)$, und dies gilt für alle $x \in (a, b]$. □

C.2) Monotonie und Ableitung

Mit Hilfe der Ableitung läßt sich bei differenzierbaren Funktionen ein hinreichendes Kriterium für Monotonie angeben.

Proposition 18.13 (Hinreichendes Kriterium für Monotonie)
Es sei $a < b$ und $f : [a, b] \longrightarrow \mathbb{R}$ sei stetig auf $[a, b]$ und differenzierbar auf (a, b).

a. Ist $f'(x) > 0$ für alle $x \in (a, b)$, so ist f streng monoton wachsend auf $[a, b]$.

b. Ist $f'(x) < 0$ für alle $x \in (a, b)$, so ist f streng monoton fallend auf $[a, b]$.

Beweis: a. 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.

b. Der Beweis geht analog zum ersten Teil.

□

Beispiel 18.14
Betrachte für $n \geq 1$ die Funktion

$$f : [0, \infty) \longrightarrow \mathbb{R} : x \mapsto x^n.$$

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.

C.3) Hinreichende Bedingung für Extremstellen

Proposition 18.15 (Hinreichende Bedingung für eine Extremstelle)
Es sei $f : (a, b) \longrightarrow \mathbb{R}$ 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

$$\lim_{x \to c} \frac{f'(x)}{x - c} = \lim_{x \to c} \frac{f'(x) - f'(c)}{x - c} = f''(c) > 0.$$

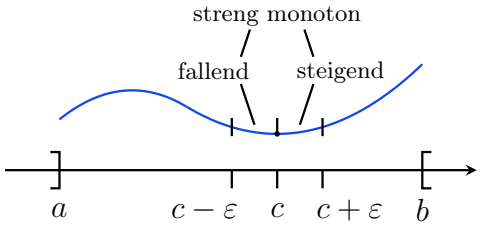
Zu $\varepsilon := \frac{f''(c)}{2} > 0$ gibt es dann ein $\delta_\varepsilon > 0$, so daß

$$-\frac{f''(c)}{2} = -\varepsilon < \frac{f'(x)}{x - c} - f''(c) < \varepsilon = \frac{f''(c)}{2}$$

für alle $x \in (a, b)$ mit $|x - c| < \delta_\varepsilon$. Insbesondere folgt für diese x dann

$$\frac{f'(x)}{x - c} > -\frac{f''(c)}{2} + f''(c) = \frac{f''(c)}{2} > 0. \tag{31}$$

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 (31) dann $f'(x) < 0$, und nach Proposition 18.13 ist f dann streng monoton fallend auf dem Intervall $[c - \delta_\varepsilon, c]$.
Analog folgt für $x \in (c, c + \delta_\varepsilon)$ aus (31) $f'(x) > 0$ und aus Proposition 18.13 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 18.16
Wir betrachten die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^3 - 3x^2 + 2.$$

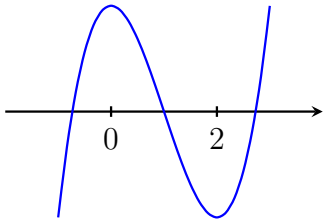
Um mögliche Extremstellen zu finden, müssen wir die Nullstellen der ersten Ableitung

$$f'(x) = 3x^2 - 6x$$

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 18.17 (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 18.15 aus den Bedingungen an $f''(c)$ ableiten.

C.4) Vertauschbarkeit von Grenzwert und Ableitung

Satz 18.18 (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

$$|\operatorname{Diff}_{f,c}(x) - g(c)| = \left| \frac{f(x) - f(c)}{x - c} - g(c) \right| < \varepsilon$$

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

$$|f_n'(x) - g(x)| < \frac{\varepsilon}{3} \tag{32}$$

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

$$|g(x) - g(c)| < \frac{\varepsilon}{3} \tag{33}$$

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

$$\frac{f_n(x) - f_n(c)}{x-c} = f_n'(y) \tag{34}$$

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

$$\begin{aligned}\left|\frac{f_n(x) - f_n(c)}{x-c} - g(c)\right| &\overset{(34)}{=} |f_n'(y) - g(c)| \\ &\leq |f_n'(y) - g(y)| + |g(y) - g(c)| \\ &\overset{(32)(33)}{<} \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2\cdot\varepsilon}{3}\end{aligned}$$

für alle $n \geq n_\varepsilon$. Lassen wir nun n gegen unendlich laufen, so erhalten wir für den Grenzwert

$$\left|\frac{f(x) - f(c)}{x-c} - g(c)\right| = \lim_{n\to\infty}\left|\frac{f_n(x) - f_n(c)}{x-c} - g(c)\right| \leq \frac{2\cdot\varepsilon}{3} < \varepsilon.$$

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 $[a,b]$. □

Bemerkung 18.19

a. Die Aussage in Satz 18.18 besagt, daß

$$\lim_{n\to\infty}(f_n') = \Big(\lim_{n\to\infty} f_n\Big)',$$

d.h. die Grenzwertbildung für die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ vertauscht mit der Ableitung!

Auf die Differenzenquotienten zurückgeführt, bedeutet dies

$$\lim_{n\to\infty}\lim_{x\to a}\frac{f_n(x) - f_n(a)}{x-a} = \lim_{x\to a}\lim_{n\to\infty}\frac{f_n(x) - f_n(a)}{x-a}.$$

Hier vertauschen zwei Grenzwertprozesse! Das ist eine Besonderheit!

b. Man kann in Satz 18.18 auf die Bedingung, daß die Ableitungen f_n' stetig sind, verzichten. Der Beweis wird dann aber etwas technischer.

c. Auch wenn wir in Satz 18.18 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 $(f_n)_{n\in\mathbb{N}}$.

C.5) Ableitung von Potenzreihen

Korollar 18.20 (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 12.53 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 18.18 erfüllt sind, ist f auf

$[-R, R]$ differenzierbar mit $f' = g$. Insbesondere ist f also in a differenzierbar mit $f'(a) = g(a)$. □

Da wir die Aussage des Korollars auch auf die formale Ableitung der Potenzreihe anwenden können, erhalten wir durch Induktion die folgende Aussage.

Korollar 18.21 (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.

C.6) Ableitungen einiger spezieller Funktionen

Korollar 18.22 (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 $a \in \mathbb{R}_{>0}$ ist die Exponentialfunktion zur Basis a

$$\exp_a : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \exp\big(x \cdot \ln(a)\big)$$

stetig differenzierbar auf $\mathbb{R}$ mit Ableitung

$$\exp_a'(x) = \ln(a) \cdot \exp_a(x).$$

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 18.20 und 18.21 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 18.20 und 18.21 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 18.20 und 18.21 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 $x\in\left(-\frac{\pi}{2},\frac{\pi}{2}\right)$ 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 $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. 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 $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Für die Ableitung erhalten wir zudem unter Berücksichtigung der Tatsache, daß der Cosinus auf $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ 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 18.23

Schaut man sich die Ableitungen der Funktionen in Korollar 18.22 d.-k. an, so kann man leicht durch Induktion zeigen, daß jede der Funktionen auf ihrem Definitionsbereich unendlich oft differenzierbar ist.

Beispiel 18.24

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 18.22 die Verkettung zweier differenzierbarer Funktionen ist. Aus der Kettenregel 17.16 folgt dann

$$f'(x) = \exp'\big(a \cdot \ln(x)\big) \cdot \frac{a}{x} = \exp\big(a \cdot \ln(x)\big) \cdot \frac{a}{x} = x^a \cdot \frac{a}{x} = a \cdot x^{a-1}.$$

Daß f sogar unendlich oft differenzierbar ist, folgt dann mit Induktion aus der Tatsache, daß f' eine Funktion der gleichen Gestalt ist.

C.7) Die Regeln von de l'Hôpital

Bemerkung 18.25 (Anwendung der Grenzwertsätze für Funktionen)

Es gibt Situationen, in denen man mit Hilfe der Grenzwersätze Grenzwerte der Form

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x\to c} f(x)}{\lim\limits_{x\to c} g(x)}$$

leicht bestimmen kann, egal ob es sich um eigentliche oder uneigentliche Grenzwerte handelt:

- $\lim\limits_{x\to 0} \frac{\cos(x)}{x^2+1} = \frac{\lim\limits_{x\to 0}\cos(x)}{\lim\limits_{x\to 0} x^2+1} = \frac{1}{1} = 1.$
- $\lim\limits_{x\to 0} \frac{\exp(x)}{\sin^2(x)} = \frac{\lim\limits_{x\to 0}\exp(x)}{\lim\limits_{x\to 0}\sin^2(x)} = \frac{1}{0} = \infty.$
- $\lim\limits_{x\to 0} \frac{x^2+1}{\ln(x)} = \frac{\lim\limits_{x\to 0} x^2+1}{\lim\limits_{x\to 0}\ln(x)} = \frac{1}{-\infty} = 0.$
- $\lim\limits_{x\to\infty} \frac{\arctan(x)}{1+\frac{1}{x}} = \frac{\lim\limits_{x\to\infty}\arctan(x)}{\lim\limits_{x\to\infty} 1+\frac{1}{x}} = \frac{\frac{\pi}{2}}{1} = \frac{\pi}{2}.$
- $\lim\limits_{x\to\infty} \frac{1+\frac{1}{x}}{\exp(x)} = \frac{\lim\limits_{x\to\infty} 1+\frac{1}{x}}{\lim\limits_{x\to\infty}\exp(x)} = \frac{1}{\infty} = 0.$

Aber es gibt auch Situationen, in denen die Grenzwertsätze nicht weiterhelfen:

$$\lim_{x\to 0}\frac{\sin(x)}{\sqrt{x}} \stackrel{?}{=} \frac{\lim\limits_{x\to 0}\sin(x)}{\lim\limits_{x\to 0}\sqrt{x}} \stackrel{?}{=} \frac{0}{0} \qquad \text{oder} \qquad \lim_{x\to\infty}\frac{\ln(x)}{x} \stackrel{?}{=} \frac{\lim\limits_{x\to\infty}\ln(x)}{\lim\limits_{x\to\infty} x} \stackrel{?}{=} \frac{\infty}{\infty}.$$

In solchen Situationen könnnen u.U. die Regeln von de l'Hôpital helfen.

Im folgenden Satz soll $[-\infty,\infty] := \mathbb{R}\cup\{-\infty,\infty\}$ bezeichnen.

Satz 18.26 (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 $\lim\limits_{x\to c}\frac{f'(x)}{g'(x)}$ existiere eigentlich oder uneigentlich.

a. Falls $\lim\limits_{x\to c} f(x) = \lim\limits_{x\to c} g(x) = 0$, so gilt $\lim\limits_{x\to c}\frac{f(x)}{g(x)} = \lim\limits_{x\to c}\frac{f'(x)}{g'(x)}$.

b. Falls $\lim\limits_{x\to c} g(x)\in\{\infty,-\infty\}$, so gilt $\lim\limits_{x\to c}\frac{f(x)}{g(x)} = \lim\limits_{x\to c}\frac{f'(x)}{g'(x)}$.

Beweis: Wir beschränken uns im Beweis auf den Fall $c\in\mathbb{R}$ und $k := \lim\limits_{x\to c}\frac{f'(x)}{g'(x)}\in\mathbb{R}$. 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

$$\left|\frac{f(x)}{g(x)} - k\right| < \varepsilon \tag{35}$$

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 $c\neq z\in(a,b)$ mit $|z-c|<\delta'_\varepsilon$ auch

$$\left|\frac{f'(z)}{g'(z)} - k\right| < \frac{\varepsilon}{2} \tag{36}$$

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

$$f'(z)\cdot\big(g(x)-g(y)\big) = g'(z)\cdot\big(f(x)-f(y)\big). \tag{37}$$

Da z zwischen x und y liegt, gilt auch

$$z \in (a,b) \cap (c-\delta'_\varepsilon, c+\delta'_\varepsilon). \tag{38}$$

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 (37) auch in der folgenden Form schreiben:

$$\frac{f(x)-f(y)}{g(x)-g(y)} = \frac{f'(z)}{g'(z)}. \tag{39}$$

a. Wir betrachten nun den Fall, daß $\lim\limits_{x\to c} f(x) = \lim\limits_{x\to c} g(x) = 0$. 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 (36), (38) und (39) 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

$$\left|\frac{f(x)}{g(x)} - k\right| = \left|\frac{f(x)-f(c)}{g(x)-g(c)} - k\right| = \lim_{y\to c}\left|\frac{f(x)-f(y)}{g(x)-g(y)} - k\right| \leq \frac{\varepsilon}{2} < \varepsilon.$$

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 (35) 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 $\lim\limits_{x\to c} g(x) = \pm\infty$ können wir auch $g(y) \neq 0$ annehmen.
Aus $\lim\limits_{x\to c} g(x) = \pm\infty$ folgt $\lim\limits_{x\to c} \frac{1}{g(x)} = 0$, und deshalb gibt es ein $\delta''_\varepsilon > 0$, so daß

$$\left|\frac{1}{g(x)}\right| < \frac{\varepsilon}{4\cdot|f(y)|} \tag{40}$$

für alle $x \in (a,b) \cap (c-\delta''_\varepsilon, c+\delta''_\varepsilon)$ mit $x \neq c, y$.
Aus (36), (38) und (39) folgt, daß

$$\left|\frac{f(x)-f(y)}{g(x)-g(y)}\right| \leq \left|\frac{f(x)-f(y)}{g(x)-g(y)} - k\right| + |k| < \frac{\varepsilon}{2} + |k| =: s \tag{41}$$

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 $\lim\limits_{x\to c} \frac{1}{g(x)} = 0$ gibt es ein $\delta'''_\varepsilon > 0$ mit

$$\left|\frac{1}{g(x)}\right| < \frac{\varepsilon}{4\cdot|g(y)|\cdot s} \tag{42}$$

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 $c \neq x \in (a,b) \cap (c-\delta_\varepsilon, c+\delta_\varepsilon)$. (35) gilt dann auch in diesem Fall wegen

$$\begin{aligned}
\left|\frac{f(x)}{g(x)} - k\right| &= \left|\frac{f(y)}{g(x)} + \frac{f(x)-f(y)}{g(x)} - k\right| \\
&= \left|\frac{f(y)}{g(x)} + \frac{f(x)-f(y)}{g(x)-g(y)} \cdot \frac{g(x)-g(y)}{g(x)} - k\right| \\
&= \left|\frac{f(y)}{g(x)} - \frac{f(x)-f(y)}{g(x)-g(y)} \cdot \frac{g(y)}{g(x)} + \frac{f(x)-f(y)}{g(x)-g(y)} - k\right| \\
&\leq \left|\frac{f(y)}{g(x)}\right| + \left|\frac{f(x)-f(y)}{g(x)-g(y)} \cdot \frac{g(y)}{g(x)}\right| + \left|\frac{f(x)-f(y)}{g(x)-g(y)} - k\right| \\
&= \underbrace{\left|\frac{f(y)}{g(x)}\right|}_{\substack{(40)\\ <\frac{\varepsilon}{4}}} + \underbrace{\left|\frac{f(x)-f(y)}{g(x)-g(y)}\right|}_{\substack{(41)\\ \leq s}} \cdot \underbrace{\left|\frac{g(y)}{g(x)}\right|}_{\substack{(42)\\ <\frac{\varepsilon}{4\cdot s}}} + \underbrace{\left|\frac{f(x)-f(y)}{g(x)-g(y)} - k\right|}_{\substack{(36)(38)(39)\\ <\frac{\varepsilon}{2}}} < \varepsilon.
\end{aligned}$$

□

Bemerkung 18.27 (Die Regeln von de l'Hôpital)

a. Wenn c in Satz 18.26 eine Nullstelle der Funktionen f und g ist und diese lokal in c beide stetig differenzierbar sind, dann kann man einen sehr einfachen Beweis für die Aussage des Satzes geben:

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{\frac{f(x)-f(c)}{x-c}}{\frac{g(x)-g(c)}{x-c}} = \frac{\lim\limits_{x\to c} \frac{f(x)-f(c)}{x-c}}{\lim\limits_{x\to c} \frac{g(x)-g(c)}{x-c}} = \frac{f'(c)}{g'(c)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}.$$

Man beachte auch, daß es hier nicht reicht, daß f und g in c differenzierbar sind, da daraus noch nicht einmal die Existenz des Grenzwertes auf der rechten Seite folgt. Dazu kann man sich z.B.

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} x^2 \cdot \sin\left(\frac{1}{x}\right), & \text{für } x \neq 0, \\ 0, & \text{für } x = 0 \end{cases}$$

aus Beispiel 17.19 anschauen. Wir haben dort gezeigt, daß diese Funktion in $c = 0$ differenzierbar ist, daß die Ableitung aber nicht stetig ist. Als zweite Funktion wählen wir den Sinus $g = \sin$. Dann haben f und g in c eine Nullstelle und sind auf ganz $\mathbb{R}$ differenzierbar, aber der Grenzwert

$$\lim_{x\to 0} \frac{f'(x)}{g'(x)} = \lim_{x\to 0} \frac{2\cdot x \cdot \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right)}{\cos(x)}$$

existiert nicht, da $\lim_{x\to 0} \cos\left(\frac{1}{x}\right)$ nicht existiert, wie man leicht sieht.

b. Im Beweis von Satz 18.26 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.

c. Ist g' stetig, so muß g' auf (a,b) entweder stets positiv oder stets negativ sein. Aus Proposition 18.13 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!

d. Man beachte, daß die zweite Regel von de l'Hôpital 18.26 nur in der Situation $\lim\limits_{x\to c} f(x) = \pm\infty$, d.h.

$$\frac{\lim\limits_{x\to c} f(x)}{\lim\limits_{x\to c} g(x)} = \frac{\pm\infty}{\pm\infty},$$

interessant ist, um den Grenzwert $\lim\limits_{x\to c} \frac{f(x)}{g(x)}$ zu bestimmen, da für $\lim\limits_{x\to c} f(x) = k \in \mathbb{R}$ schon aus den normalen Grenzwertsätzen

$$\lim_{x\to c} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x\to c} f(x)}{\lim\limits_{x\to c} g(x)} = \frac{k}{\pm\infty} = 0$$

folgen würde!

Beispiel 18.28

a. Wir betrachten die Funktionen $f = \sin$ und $g = \sqrt{\cdot}$ auf dem Intervall $(0,\infty)$. Dort sind beide differenzierbar mit

$$\lim_{x\to 0} f(x) = \lim_{x\to 0} g(x) = 0$$

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 18.26 folgt dann

$$\lim_{x\to 0} \frac{\sin(x)}{\sqrt{x}} = \lim_{x\to 0} \frac{f'(x)}{g'(x)} = \lim_{x\to 0} \cos(x)\cdot 2\cdot\sqrt{x} = \cos(0)\cdot 2\cdot\sqrt{0} = 0.$$

b. Wir betrachten die Funktionen

$$\ln : (0,\infty) \longrightarrow \mathbb{R}$$

und

$$g : (0,\infty) \longrightarrow (0,\infty) : x \mapsto x^a = \exp\big(a\cdot\ln(x)\big)$$

für ein festes $a \in \mathbb{R}_{>0}$. Nach Korollar 18.22 und Beispiel 18.24 sind beide Funktionen differenzierbar auf $(0,\infty)$.
Da sowohl $\exp(x)$, als auch $\ln(x)$ für $x \to \infty$ gegen ∞ divergieren und da a positiv ist, folgt aus den Grenzwertsätzen für uneigentliche Grenzwerte 13.17

$$\lim_{x\to\infty} g(x) = \lim_{x\to\infty} \exp\big(a\cdot\ln(x)\big) = \infty.$$

Außerdem gilt nach Beispiel 18.24

$$g'(x) = a \cdot x^{a-1} \neq 0$$

für alle $x \in (0, \infty)$. Aus der zweiten Regel von de l'Hôpital 18.26 folgt dann

$$\lim_{x\to\infty} \frac{\ln(x)}{x^a} = \lim_{x\to\infty} \frac{\ln'(x)}{g'(x)} = \lim_{x\to\infty} \frac{\frac{1}{x}}{a \cdot x^{a-1}} = \lim_{x\to\infty} \frac{1}{a \cdot g(x)} = \frac{1}{\infty} = 0.$$

C.8) Wachstum der Exponentialfunktion

Korollar 18.29 (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

$$\lim_{x\to\infty} \frac{f(x)}{\exp(x)} = 0.$$

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,

$$\lim_{x\to\infty} \frac{f(x)}{\exp(x)} = \frac{a_0}{\lim\limits_{x\to\infty} \exp(x)} = 0,$$

da $\lim\limits_{x\to\infty} \exp(x) = \infty$.

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 18.26 anwenden, da $\exp'(x) = \exp(x) \neq 0$ für alle $x \in \mathbb{R}$ und da $\lim\limits_{x\to\infty} \exp(x) = \infty$ gilt. Damit erhalten wir

$$\lim_{x\to\infty} \frac{f(x)}{\exp(x)} = \lim_{x\to\infty} \frac{f'(x)}{\exp'(x)} = \lim_{x\to\infty} \frac{f'(x)}{\exp(x)},$$

aber die rechte Seite ist dann Null nach Induktion. □

Bemerkung 18.30
Korollar 18.29 besagt, daß die Exponentialfunktion asymptotisch für x gegen ∞ erheblich viel schneller wächst als jede Polynomfunktion. Das ist der Grund, weshalb

Algorithmen mit exponentiellem Laufzeitverhalten wesentlich schlechter sind als Algorithmen mit polynomialem Laufzeitverhalten, zumindest wenn man Eingabedaten beliebiger Größe zulassen möchte. Für konkrete, hinreichend kleine Eingabedaten muß das nicht der Fall sein. Zudem gibt es Probleme, wie etwa die Berechnung von Gröbnerbasen, für die nur Algorithmen mit exponentiellem Laufzeitverhalten bekannt sind. Damit muß man dann ggf. leben.

C.9) Der Satz von Taylor

Definition 18.31 (Taylorpolynome)
Es sei $f : U \longrightarrow \mathbb{R}$ und $a \in U$.
Ist f n-fach differenzierbar, so nennen wir das Polynom

$$T^n_{f,a} := \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!} \cdot (t-a)^k$$

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 a oder die *Taylor-Entwicklung* von f im Punkt a. Beachte, stets gilt $T^n_{f,a}(a) = T_{f,a}(a) = f(a)$.

Bemerkung 18.32 (Tangenten und das 1. Taylorpolynome)
Ist $f : \mathbb{R} \longrightarrow \mathbb{R}$ differenzierbar, so ist die Gleichung der Tangente an den Graphen von f im Punkt $\big(a, f(a)\big)$ gegeben durch

$$y = f'(a) \cdot (x-a) + f(a) = T^1_{f,a}(x).$$

D.h. das erste Taylorpolynom von f mit Entwicklungspunkt a ist die optimale lineare Approximation der Funktion f lokal in a.

Die Idee ist nun, daß mit steigendem n die Taylorpolynome $T^n_{f,a}$ immer bessere Approximationen von f lokal in 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 18.34! Funktionen, für die das gilt, nennt man *analytisch* in 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 18.33 (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

$$f : (-r, r) \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} a_n \cdot x^n$$

nach Korollar 18.21 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 18.34
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 18.43.

Der Satz von Taylor sagt etwas darüber aus, wie gut das n-te Taylorpolynom f approximiert.

Satz 18.35 (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

$$z := \frac{\big(f(x) - T^n_{f,a}(x)\big) \cdot (n+1)!}{(x-a)^{n+1}} \in \mathbb{R}$$

und eine Funktion $g : [a, x] \longrightarrow \mathbb{R}$ durch

$$\begin{aligned} g(y) :=& f(x) - T^n_{f,y}(x) - \frac{z}{(n+1)!} \cdot (x-y)^{n+1} \\ =& f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(y)}{k!} \cdot (x-y)^k - \frac{z}{(n+1)!} \cdot (x-y)^{n+1} \\ =& f(x) - f(y) - \sum_{k=1}^{n} \frac{f^{(k)}(y)}{k!} \cdot (x-y)^k - \frac{z}{(n+1)!} \cdot (x-y)^{n+1} \end{aligned}$$

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

$$g(a) = f(x) - T^n_{f,a}(x) - \frac{z}{(n+1)!} \cdot (x-a)^{n+1} = 0,$$

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

$$f^{(n+1)}(c) = z = \frac{\big(f(x) - T^n_{f,a}(x)\big) \cdot (n+1)!}{(x-a)^{n+1}}$$

gelten, und damit

$$f(x) - T^n_{f,a}(x) = \frac{f^{(n+1)}(c)}{(n+1)!} \cdot (x-a)^{n+1}.$$

□

Beispiel 18.36 (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

$$T^n_{\exp,0}(1) = T^n_{f,a}(1) = \sum_{k=0}^{n} \frac{1}{k!}.$$

Mit Hilfe des Satzes von Taylor finden wir ein $c \in (0,1)$ mit

$$\left| e - \sum_{k=0}^{n} \frac{1}{k!} \right| = \left| \exp(1) - T^n_{\exp,0}(1) \right| = \frac{|\exp(c)|}{(n+1)!} \cdot |1-0|^{n+1} < \frac{e}{(n+1)!} < \frac{3}{(n+1)!},$$

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:

$$e = 2,718\ldots.$$

Beispiel 18.37 (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

$$T^n_{\ln,1}(x) = \sum_{k=0}^{n} \frac{\ln^{(k)}(1)}{k!} \cdot (x-1)^k = \sum_{k=1}^{n} (-1)^{k-1} \cdot \frac{(x-1)^k}{k}.$$

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

$$|\ln(x) - T^n_{\ln,1}(x)| = \frac{|\ln^{(n+1)}(c)|}{(n+1)!} \cdot |(x-1)|^{n+1} \leq \frac{1}{n+1}.$$

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

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -\ln(2).$$

C.10) Allgemeinere Bedingung für Extremstellen

Satz 18.38 (Allgemeinere Bedingung für Extremstellen)
Es sei $f : (a,b) \longrightarrow \mathbb{R}$ eine $n+1$-fach stetig differenzierbare Funktion mit $n \in \mathbb{N}$ und es sei $c \in (a,b)$ mit $f'(c) = f''(c) = \ldots = f^{(n)}(c) = 0$ und $f^{(n+1)}(c) \neq 0$.

a. Falls n gerade ist, so ist c keine Extremstelle von f.

b. Falls n ungerade und $f^{(n+1)}(c) < 0$, so ist c ein lokales Maximum.

c. 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 annimt, muß es nach Aufgabe 14.33 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

$$T^n_{f,c} = \sum_{k=0}^{n} \frac{f^{(k)}(c)}{k!} \cdot (t-c)^k = f(c) + \sum_{k=1}^{n} \frac{0}{k!} \cdot (t-c)^k = f(c)$$

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

$$f(x) - f(c) = f(x) - T^n_{f,c}(x) = \frac{f^{(n+1)}(d_x)}{(n+1)!} \cdot (x-c)^{n+1}.$$

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 18.15 überein. Aber wir haben diesmal vorausgesetzt, daß die zweite Ableitung *stetig* ist. Darauf konnten wir in Proposition 18.15 verzichten.

Beispiel 18.39

Betrachten wir die Funktion $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^4$, dann ist $a = 0$ offenbar ein globales Minimum, da $f(x) > 0$ für alle $x \neq 0$ gilt. Proposition 18.15 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 18.38 $a = 0$ als ein wenn auch nur lokales Minimum aus.

Bemerkung 18.40

Man beachte, daß die Funktion in Beispiel 18.34 im Punkt $a = 0$ ein globales Minimum besitzt und unendlich oft differenzierbar ist, daß aber alle Ableitungen im Punkt $a = 0$ verschwinden! Satz 18.38 ist also nicht immer anwendbar, um Extremstellen zu finden.

Aufgaben

Aufgabe 18.41

Berechne die Ableitungen der folgenden Funktionen $f : (a, \infty) \longrightarrow \mathbb{R}$.

a. $f(x) = \ln\left(\frac{\ln(x)}{x}\right)$ mit $a = 1$.

b. $f(x) = \frac{x^2+4}{x-4}$ mit $a = 4$.

c. $f(x) = \sqrt{e^{\cos(\sqrt{x})}}$ mit $a = 0$.

Aufgabe 18.42

Sei $f : (0,1] \longrightarrow \mathbb{R}$ eine differenzierbare Funktion mit beschränkter Ableitung. Zeige, dass f dann stetig in 0 fortsetzbar ist.

Aufgabe 18.43

Zeige, daß für die Funktion

$$f : \mathbb{R} \longmapsto \mathbb{R},\ x \longmapsto \begin{cases} \exp\left(-\frac{1}{x^2}\right) & \text{für } x \neq 0, \\ 0 & \text{für } x = 0. \end{cases}$$

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 $\lim\limits_{x\to 0} \frac{\exp(-\frac{1}{x^2})}{x^k} = 0$.

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 18.44 (Näherungsweise Berechnung von π)

Betrachte die Funktion $f = \arctan$ auf $\mathbb{R}$.

a. Berechne das dritte Taylorpolynom $T^3_{f,0}$ von arctan mit Entwicklungspunkt 0.

b. Benutze $T^3_{f,0}$ und Aufgabe 16.23 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

$$\pi = 3,14\ldots.$$

Aufgabe 18.45

Berechne die folgenden Grenzwerte:

a. $\lim_{x\to\pi} \frac{\cos(x)+1}{x^2-\pi^2}$ mit $x < \pi$.

b. $\lim_{x\to 1} \left(\frac{1}{x-1} - \frac{1}{\ln(x)}\right)$ mit $x > 1$.

c. $\lim_{x\to 0} x^x$ mit $x > 0$.

Aufgabe 18.46

Bestimme alle Extrema der Funktion $f : [0,1] \longrightarrow \mathbb{R},\ x \longmapsto (1-x)\cdot\sqrt{1+9x^2}$.

Aufgabe 18.47

Berechne für die Funktion

$$f : \mathbb{R} \longrightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right), \; x \longmapsto \arctan(x)$$

das zweite Taylor-Polynom $T^2_{f,0}$ um 0 und gib eine Abschätzung für das Restglied $|f(x) - T^2_{f,0}(x)|$ auf dem Intervall $[-1, 1]$ an.

Aufgabe 18.48

Berechne für die Funktion

$$f : \left(-\frac{1}{2}, \frac{1}{2}\right) \longrightarrow \mathbb{R}, \; x \longmapsto \frac{\cos(x)}{1 - (2x)^4}$$

das vierte Taylor-Polynom $T^4_{f,0}$ um 0.

Hinweis: mit etwas Überlegung kann man die Berechnung aller vier Ableitungen von f vermeiden.

Aufgabe 18.49

Beweise oder widerlege durch eine Gegenbeispiel die folgenden Aussagen:

a. Sei $f : [a, b] \longrightarrow \mathbb{R}$ eine stetige und auf (a, b) differenzierbare Funktion, so daß $\lim\limits_{x \to a} f'(x)$ existiert. Dann ist f differenzierbar in a.

b. Sei $f \in \mathcal{C}^1([a, b], \mathbb{R})$ eine stetig differenzierbare streng monoton wachsende Funktion. Dann gilt $f'(x) > 0$ für alle $x \in [a, b]$.

Aufgabe 18.50 (Binomialreihe)

Zeige mit Hilfe der Taylorentwicklung des natürlichen Logarithmus in Beispiel 18.37 für alle $a \in \mathbb{R} \setminus \mathbb{N}$ und $x \in (-1, 1)$ die Gleichheit

$$(1 + x)^a = \sum_{n=0}^{\infty} \binom{a}{n} \cdot x^n.$$

Siehe auch Aufgabe 12.59.

Aufgabe 18.51

Zeige, daß die Funktion $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto 2e^x + e^{2x} - 1$ streng monoton wachsend ist, bestimme das Bild von f sowie die Umkehrfunktion und deren Ableitung.

Aufgabe 18.52

Bestimme alle Extremstellen von $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^3 - x^2 - 8x + 1$.

Aufgabe 18.53 (Vorsicht bei den Regeln von de l'Hôpital)
Wir betrachten die beiden Funktionen

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x + \sin(x) \cdot \cos(x)$$

und

$$g : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto f(x) \cdot \exp(\sin(x)).$$

Man zeige, daß $\lim_{x\to\infty} g(x) = \infty$ gilt und daß auch der Grenzwert $\lim_{x\to\infty} \frac{f'(x)}{g'(x)}$ existiert, daß der Grenzwert $\lim_{x\to\infty} \frac{f(x)}{g(x)}$ aber trotzdem nicht mit der Regel von de l'Hôpital bestimmt werden kann. Welche Voraussetzung von Satz 18.26 ist verletzt?

Aufgabe 18.54
Zeige, daß die Funktion

$$f_a : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^3 - 3a^2x + 7$$

für keinen Parameterwert $a \in \mathbb{R}_{>0}$ drei nicht-negative Nullstellen hat.

Aufgabe 18.55
Bestimme eine monoton steigende, differenzierbare Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ mit

$$f(x) = \begin{cases} 0, & \text{für } x < -1, \\ 4, & \text{für } x > 1. \end{cases}$$

Aufgabe 18.56
Berechne die folgenden Grenzwerte mittels der Regeln von de l'Hôpital:

a. $\lim_{x\to 0} \frac{1-\cos(x)}{x}$.

b. $\lim_{x\to -1} \frac{x^2-1}{x^3+2x+3}$.

c. $\lim_{x\to\infty} \frac{x^2+\ln(x)}{x^2-\ln(x)}$.

Aufgabe 18.57
Berechne die folgenden Grenzwerte mittels der Regeln von de l'Hôpital:

a. $\lim_{x\to 0} \frac{20\cdot x^{1300}}{\tan(x^{1300})}$.

b. $\lim_{x\to 0} \frac{e^x-x-1}{1-\cos(x)}$.

c. $\lim_{x\to 0} \frac{a^x-1}{x}$ mit $1 \neq a \in \mathbb{R}_{>0}$.

Aufgabe 18.58

Bestimme für die Funktion $f : [a, b] \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x}$ mit $0 < a < b$ einen Wert $c \in (a, b)$ mit

$$f'(c) = \frac{f(b) - f(a)}{b - a}.$$

Aufgabe 18.59

Es sei $a \in \mathbb{R}$ mit $a > 1$ und

$$f : (0, \infty) \longrightarrow \mathbb{R} : x \mapsto x \cdot \log_a(x).$$

Untersuche, auf welchen Teilintervallen des Definitionsbereichs die Funktion streng monton wächst bzw. fällt, und untersuche das Grenzverhalten für $x \to 0$.

Aufgabe 18.60

Gegeben seien reelle Zahlen $0 \leq a_1 < a_2 < \ldots < a_n$ sowie die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sum_{i=1}^{n} (x - a_i)^2.$$

Bestimme den minimalen Wert von f und die Stelle $x \in \mathbb{R}$, an der dieser Wert angenommen wird.

Aufgabe 18.61

Berechne das dritte Taylorpolynom der folgenden Funktionen $f : \mathbb{R} \longrightarrow \mathbb{R}$ im angegebenen Entwicklungspunkt a:

a. $f(x) = \sin(x)$, $a = \frac{\pi}{2}$.

b. $f(x) = e^{4x^2+x}$, $a = 0$.

Hinweis: Mit etwas Überlegung braucht man in Teil b. keine Ableitung zu berechnen!

§ 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) Obersummen und Untersummen

Definition 19.1 (Zerlegungen eines Intervalls)
Es seien $a, b \in \mathbb{R}$ mit $a < b$. Ein Tupel $Z = (x_0, \ldots, x_n)$ mit $n \geq 1$ heißt eine *Zerlegung* des Intervalls $[a, b]$, falls

$$a = x_0 < x_1 < \ldots < x_{n-1} < x_n = b.$$

Die Zahl $l(Z) := \max\{x_i - x_{i-1} \mid i = 1, \ldots, n\}$ heißt die *Länge* oder *Feinheit* der Zerlegung, die Menge $\operatorname{supp}(Z) := \{x_0, \ldots, x_n\}$ ihr *Träger*, die Zahl $|Z| := n$ ihre *Mächtigkeit* und die x_i ihre *Stützpunkte*.
Eine zweite Zerlegung $Z' = (y_0, \ldots, y_m)$ von $[a, b]$ heißt *Verfeinerung* von Z, falls

$$\{x_0, \ldots, x_n\} \subseteq \{y_0, \ldots, y_m\}.$$

Zu zwei Zerlegungen $Z = (x_0, \ldots, x_n)$ und $Z' = (y_0, \ldots, y_m)$ definieren wir

$$Z * Z' := (z_0, \ldots, z_k),$$

indem wir die Elemente der Vereinigung $\operatorname{supp}(Z) \cup \operatorname{supp}(Z') = \{z_0, \ldots, z_k\}$ der Größe nach ordnen. Sind Z und Z' Zerlegungen des gleichen Intervalls, so nennen wir $Z * Z'$ ihre *gemeinsame Verfeinerung*.

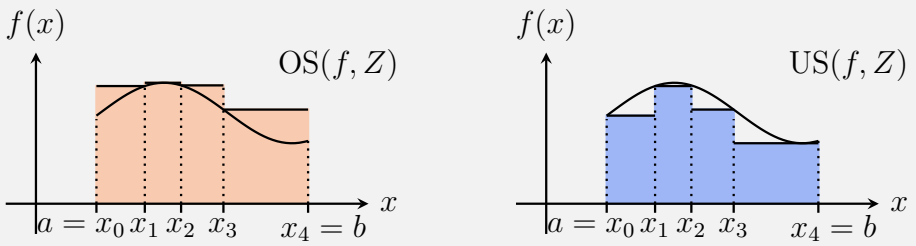
Beispiel 19.2
Die Tupel $Z = (0, 1, 3, 5)$ und $Z' = (0, 2, 5)$ sind Zerlegungen von $[0, 5]$ der Länge 2 bzw. 3, und die gemeinsame Verfeinerung von Z und Z' ist $Z * Z' = (0, 1, 2, 3, 5)$.

Definition 19.3 (Obersummen und Untersummen)
Sei $f : [a, b] \to \mathbb{R}$ beschränkt, $a < b$, und $Z = (x_0, \ldots, x_n)$ eine Zerlegung von $[a, b]$.
Wir definieren die *Obersumme* von f bezüglich Z als

$$\operatorname{OS}(f, Z) := \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \sup\{f(x) \mid x \in [x_{i-1}, x_i]\},$$

und die *Untersumme* von f bezüglich Z als

$$\operatorname{US}(f, Z) := \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}.$$



Beispiel 19.4

Wir betrachten die Identität $\mathrm{id} : [0,1] \longrightarrow \mathbb{R} : x \mapsto x$ auf dem Intervall $[0,1]$ sowie die folgende äquidistante Zerlegung der Länge $\frac{1}{n}$

$$Z^n = (x_0, \ldots, x_n) = \left(0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\right).$$

Auf einem Teilintervall $[x_{i-1}, x_i] = \left[\frac{i-1}{n}, \frac{i}{n}\right]$ gilt dann

$$m_i := \inf\left\{ \mathrm{id}(x) \;\middle|\; x \in [x_{i-1}, x_i]\right\} = x_{i-1} = \frac{i-1}{n}$$

und

$$M_i := \sup\left\{ \mathrm{id}(x) \;\middle|\; x \in [x_{i-1}, x_i]\right\} = x_i = \frac{i}{n}.$$

Für die Unter- und Obersumme von id bezüglich Z ergibt sich unter Berücksichtigung von Beispiel 7.11 damit

$$\mathrm{US}(\mathrm{id}, Z^n) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot m_i = \sum_{i=1}^{n} \frac{1}{n} \cdot \frac{i-1}{n} = \frac{n \cdot (n-1)}{2 \cdot n^2} = \frac{1}{2} - \frac{1}{2n}$$

und

$$\mathrm{OS}(\mathrm{id}, Z^n) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot M_i = \sum_{i=1}^{n} \frac{1}{n} \cdot \frac{i}{n} = \frac{n \cdot (n+1)}{2 \cdot n^2} = \frac{1}{2} + \frac{1}{2n}.$$

Lemma 19.5

Es sei $f : [a,b] \longrightarrow \mathbb{R}$ 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

$$0 \leq \mathrm{US}(f, Z') - \mathrm{US}(f, Z) \leq 2 \cdot M \cdot l(Z) \cdot \big(|Z'| - |Z|\big)$$

und

$$0 \leq \mathrm{OS}(f, Z) - \mathrm{OS}(f, Z') \leq 2 \cdot M \cdot l(Z) \cdot \big(|Z'| - |Z|\big).$$

Insbesondere gilt also

$$\mathrm{US}(f, Z) \leq \mathrm{US}(f, Z') \leq \mathrm{OS}(f, Z') \leq \mathrm{OS}(f, Z).$$

b. Für je zwei Zerlegungen Z und Z' von $[a,b]$ gilt

$$\mathrm{US}(f,Z) \le \mathrm{OS}(f,Z').$$

c. Es gelten

$$-M\cdot(b-a) \le \mathrm{US}(f,Z) \le \mathrm{OS}(f,Z) \le (b-a)\cdot M.$$

Beweis:

a. Es sei $Z=(x_0,\dots,x_n)$, und wir setzen für $i=1,\dots,n$ wieder

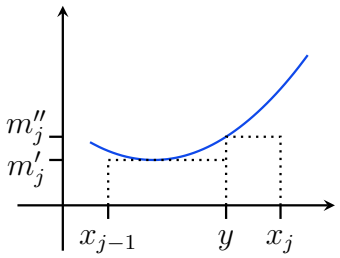
$$m_i := \inf\big\{f(x) \;\big|\; x\in[x_{i-1},x_i]\big\}.$$

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

$$m'_j := \inf\big\{f(x) \;\big|\; x\in[x_{j-1},y]\big\} \ge m_j$$

und

$$m''_j := \inf\big\{f(x) \;\big|\; x\in[y,x_j]\big\} \ge m_j.$$



Daraus ergibt sich

$$\begin{aligned}
\mathrm{US}(f,Z) &= \sum_{i\neq j}(x_i-x_{i-1})\cdot m_i + (x_j-x_{j-1})\cdot m_j \\
&= \sum_{i\neq j}(x_i-x_{i-1})\cdot m_i + (y-x_{j-1})\cdot m_j + (x_i-y)\cdot m_j \\
&\le \sum_{i\neq j}(x_i-x_{i-1})\cdot m_i + (y-x_{j-1})\cdot m'_j + (x_i-y)\cdot m''_j \\
&= \mathrm{US}(f,Z').
\end{aligned}$$

Für die Differenz der beiden Terme erhalten wir

$$\begin{aligned}
0 \le{}& \mathrm{US}(f,Z') - \mathrm{US}(f,Z) \\
={}& (y-x_{j-1})\cdot(m'_j-m_j) + (x_j-y)\cdot(m''_j-m_j) \\
\le{}& (y-x_{j-1})\cdot(M+M) + (x_j-y)\cdot(M+M) \\
={}& (x_j-x_{j-1})\cdot 2\cdot M \le 2\cdot M\cdot l(Z).
\end{aligned}$$

Für eine beliebige Verfeinerung Z' von Z wenden wir dann Induktion an und erhalten die Formel

$$0 \leq \mathrm{US}(f, Z') - \mathrm{US}(f, Z) \leq 2 \cdot M \cdot l(Z) \cdot \big(|Z'| - |Z|\big)$$

Die Aussage für Obersummen zeigt man analog.

b. Wir betrachten die gemeinsame Verfeinerung $Z * Z' = (y_0, \ldots, y_k)$. Wegen

$$m_i := \inf\big\{f(x) \;\big|\; x \in [y_{i-1}, y_i]\big\} \leq \sup\big\{f(x) \;\big|\; x \in [y_{i-1}, y_i]\big\} =: M_i$$

folgt dann

$$\begin{aligned}\mathrm{US}(f, Z) &\overset{a.}{\leq} \mathrm{US}(f, Z * Z') = \sum_{i=1}^{k} (y_i - y_{i-1}) \cdot m_i \\ &\leq \sum_{i=1}^{k} (y_i - y_{i-1}) \cdot M_i = \mathrm{OS}(f, Z * Z') \overset{a.}{\leq} \mathrm{OS}(f, Z').\end{aligned}$$

c. Die Aussage folgt aus a., da Z eine Verfeinerung der Zerlegung (a, b) ist und da $M \geq \sup\big\{f(x) \;\big|\; x \in [a, b]\big\} \geq \inf\big\{f(x) \;\big|\; x \in [a, b]\big\} \geq -M$.

□

Beispiel 19.6
In Beispiel 19.4 gilt $\mathrm{US}(\mathrm{id}, Z^n) = \frac{1}{2} - \frac{1}{2n} \leq \frac{1}{2} + \frac{1}{2n} = \mathrm{OS}(\mathrm{id}, Z^n)$.

B) Riemann-integrierbare Funktionen

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 $f : [a, b] \longrightarrow \mathbb{R}$ beschränkt, $a < b$. Wir definieren das *Oberintegral*

$$\mathrm{OI}(f) := \inf\big\{\, \mathrm{OS}(f, Z) \;\big|\; Z \text{ Zerlegung von } [a, b]\big\}$$

von f und das *Unterintegral*

$$\mathrm{UI}(f) := \sup\big\{\, \mathrm{US}(f, Z) \;\big|\; Z \text{ Zerlegung von } [a, b]\big\}$$

von f. Wegen Lemma 19.5 b. und Lemma 8.16 gilt

$$\mathrm{UI}(f) \leq \mathrm{OI}(f).$$

Wir nennen f (Riemann-)*integrierbar* auf $[a, b]$, falls $\mathrm{UI}(f) = \mathrm{OI}(f)$. Dann heißt

$$\int_a^b f(x)\, dx := \mathrm{OI}(f) \in \mathbb{R}$$

das *Integral* von f auf $[a, b]$.

Beispiel 19.8

Aus Beispiel 19.4 wissen wir für $\mathrm{id} : [0,1] \longrightarrow \mathbb{R}$

$$\frac{1}{2} - \frac{1}{2n} = \mathrm{US}(\mathrm{id}, Z^n) \leq \mathrm{UI}(\mathrm{id}) \leq \mathrm{OI}(\mathrm{id}) \leq \mathrm{OS}(\mathrm{id}, Z^n) = \frac{1}{2} + \frac{1}{2n}$$

für alle $n \in \mathbb{N}$. Bilden wir nun den Grenzwert für n gegen unendlich, so erhalten wir

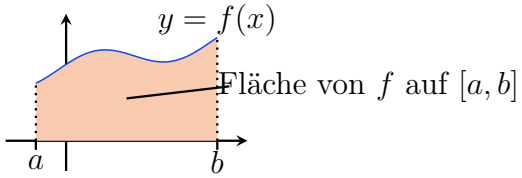
$$\frac{1}{2} = \lim_{n\to\infty} \mathrm{US}(\mathrm{id}, Z^n) \leq \mathrm{UI}(\mathrm{id}) \leq \mathrm{OI}(\mathrm{id}) \leq \lim_{n\to\infty} \mathrm{OS}(\mathrm{id}, Z^n) = \frac{1}{2},$$

d.h. id ist integrierbar auf $[0,1]$ mit

$$\int_0^1 x\,dx = \mathrm{OI}(\mathrm{id}) = \mathrm{UI}(\mathrm{id}) = \frac{1}{2}.$$

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 $\mathrm{OI}(f)$ als Infimum und $\mathrm{UI}(f)$ als Supremum gilt also stets $\mathrm{UI}(f) \leq I \leq \mathrm{OI}(f)$. Daß f integrierbar ist, bedeutet mithin nichts anderes, als daß das Integral $\int_a^b f(x)\,dx$ den Flächeninhalt der Fläche beschreibt, die der Graph von f auf dem Intervall $[a,b]$ mit der x-Achse einschließt.



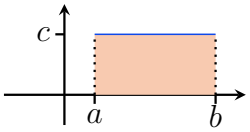
Beispiel 19.10

a. Jede konstante Funktion $f : [a,b] \longrightarrow \mathbb{R} : x \mapsto c$ ist integrierbar mit

$$\int_a^b f(x)\,dx = (b-a)\cdot c.$$

Denn dann gilt für jede Zerlegung $Z = (x_0, \ldots, x_n)$ von $[a,b]$ bereits

$$\mathrm{OS}(f,Z) = \sum_{i=1}^{n}(x_i - x_{i-1})\cdot c = (b-a)\cdot c = \sum_{i=1}^{n}(x_i - x_{i-1})\cdot c = \mathrm{US}(f,Z).$$



b. Die *Dirichletsche Sprungfunktion*

$$f : [0,1] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{falls } x \in \mathbb{Q}, \\ 1, & \text{falls } x \notin \mathbb{Q} \end{cases}$$

ist *nicht* integrierbar auf $[0,1]$. Denn ist $Z = (x_0, \ldots, x_n)$ 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

$$\mathrm{US}(f,Z) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot 0 = 0$$

und

$$\mathrm{OS}(f,Z) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot 1 = 1$$

für jede Zerlegung Z, so daß

$$\mathrm{UI}(f) = 0 < 1 = \mathrm{OI}(f).$$

C) Das Riemannsche Integrabilitätskriterium

Satz 19.11 (Riemannsches Integrabilitätskriterium)
Sei $f : [a,b] \to \mathbb{R}$ beschränkt, $a < b$. Genau dann ist f integrierbar auf $[a,b]$, wenn

$$\forall\, \varepsilon > 0 \;\; \exists\, Z \text{ Zerlegung von } [a,b] \;:\; \mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon.$$

Beweis: $\Longrightarrow$: Sei zunächst f integrierbar auf $[a,b]$ und sei $\varepsilon > 0$ gegeben. Aufgrund der Definition von $\mathrm{UI}(f)$ als Supremum und $\mathrm{OI}(f)$ als Infimum und wegen Proposition 8.15 gibt es Zerlegungen Z' und Z'' von $[a,b]$ mit

$$\mathrm{OI}(f) + \frac{\varepsilon}{2} > \mathrm{OS}(f,Z') \overset{19.5}{\geq} \mathrm{OS}(f, Z' * Z'')$$

und

$$\mathrm{UI}(f) - \frac{\varepsilon}{2} < \mathrm{US}(f,Z'') \overset{19.5}{\leq} \mathrm{US}(f, Z' * Z'').$$

Damit erhalten wir mit $Z = Z' * Z''$ und wegen $\mathrm{UI}(f) = \mathrm{OI}(f)$

$$\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \left(\mathrm{OI}(f) + \frac{\varepsilon}{2}\right) - \left(\mathrm{UI}(f) - \frac{\varepsilon}{2}\right) = \varepsilon.$$

$\Longleftarrow$: Für $\varepsilon := \frac{1}{n}$ mit $n \geq 1$ gibt es eine Zerlegung Z^n von $[a,b]$ mit

$$\frac{1}{n} > \mathrm{OS}(f,Z^n) - \mathrm{US}(f,Z^n) \geq \mathrm{OI}(f) - \mathrm{UI}(f) \geq 0.$$

Da die linke Seite der Ungleichung für $n \to \infty$ gegen Null konvergiert, folgt im Grenzwert

$$0 = \lim_{n\to\infty} \frac{1}{n} \geq \mathrm{OI}(f) - \mathrm{UI}(f) \geq 0,$$

also $\mathrm{OI}(f) = \mathrm{UI}(f)$. Mithin ist f integrierbar auf $[a,b]$. □

Das folgende Verschärfung des Riemannschen Integrabilitätskriteriums sagt, daß für integrierbare Funktionen Untersummen und Obersummen schon beliebig nahe beieinander und damit beim Wert des Integrals liegen, wenn nur die Länge der Zerlegung hinreichend klein gewählt ist.

Lemma 19.12 (Verschärfung des Riemannschen Integrabilitätskriteriums)
Ist $f : [a,b] \longrightarrow \mathbb{R}$ integrierbar, $a < b$, so gilt:

$\forall\, \varepsilon > 0\ \exists\, \delta_\varepsilon > 0\ :\ \forall\ Z$ Zerlegung mit $l(Z) < \delta_\varepsilon$ gilt $\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon$.

Beweis: Sei $\varepsilon > 0$ gegeben. Aus dem Riemannschen Integrabilitätskriterium erhalten wir eine Zerlegung Z' von $[a,b]$, so daß

$$\mathrm{OS}(f,Z') - \mathrm{US}(f,Z') < \frac{\varepsilon}{2}. \tag{43}$$

Wir setzen nun

$$\delta_\varepsilon := \frac{\varepsilon}{8 \cdot |Z'| \cdot M} > 0,$$

wobei $M := \sup\big\{|f(x)| \;\big|\; x \in [a,b]\big\}$. Ist Z eine Zerlegung von $[a,b]$ mit $l(Z) < \delta_\varepsilon$, so folgt aus aus Lemma 19.5 und $|Z * Z'| - |Z| \le |Z'|$

$$\mathrm{OS}(f,Z) - \mathrm{OS}(f,Z*Z') \le 2 \cdot M \cdot l(Z) \cdot (|Z*Z'| - |Z|) < 2 \cdot M \cdot \delta_\varepsilon \cdot |Z'| = \frac{\varepsilon}{4} \tag{44}$$

und

$$\mathrm{US}(f,Z*Z') - \mathrm{US}(f,Z) \le 2 \cdot M \cdot l(Z) \cdot (|Z*Z'| - |Z|) < 2 \cdot M \cdot \delta_\varepsilon \cdot |Z'| = \frac{\varepsilon}{4}. \tag{45}$$

Da $Z * Z'$ eine Verfeinerung von Z' ist, folgt aus (43) zusammen mit Lemma 19.5

$$\mathrm{OS}(f,Z*Z') - \mathrm{US}(f,Z*Z') \le \mathrm{OS}(f,Z') - \mathrm{US}(f,Z') < \frac{\varepsilon}{2}. \tag{46}$$

Insgesamt erhalten wir damit

$$\begin{aligned}\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) &= \mathrm{OS}(f,Z) - \mathrm{OS}(f,Z*Z') + \mathrm{OS}(f,Z*Z') - \mathrm{US}(f,Z*Z') \\ &\quad + \mathrm{US}(f,Z*Z') - \mathrm{US}(f,Z) \overset{(44)(46)(45)}{<} \frac{\varepsilon}{4} + \frac{\varepsilon}{2} + \frac{\varepsilon}{4} = \varepsilon.\end{aligned}$$

□

D) Zwei Klassen integrierbarer Funktionen

Im folgenden wollen wir das Riemannsche Integrabilitätskriterium anwenden, um von zwei wichtigen Klassen von Funktionen zu zeigen, daß sie integrierbar sind.

Satz 19.13 (Stetige Funktionen sind integrierbar.)
Ist $f : [a,b] \longrightarrow \mathbb{R}$ 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

$$|f(x) - f(y)| < \frac{\varepsilon}{b-a} \tag{47}$$

für alle $x, y \in [a,b]$ mit $|x-y| < \delta_\varepsilon$. Wir wählen nun eine Zerlegung $Z = (x_0, \ldots, x_n)$ 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

$$f(y_i) = \sup\big\{f(y) \;\big|\; y \in [x_{i-1}, x_i]\big\}$$

und

$$f(z_i) = \inf\big\{f(y) \;\big|\; y \in [x_{i-1}, x_i]\big\},$$

und wegen $|y_i - z_i| \leq |x_i - x_{i-1}| < \delta_\varepsilon$ folgt aus (47) zudem

$$0 \leq f(y_i) - f(z_i) < \frac{\varepsilon}{b-a}.$$

Damit erhalten wir insbesondere

$$\begin{aligned} \mathrm{OS}(f,Z) - \mathrm{US}(f,Z) &= \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \big(f(y_i) - f(z_i)\big) \\ &< \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \frac{\varepsilon}{b-a} = (b-a) \cdot \frac{\varepsilon}{b-a} = \varepsilon. \end{aligned}$$

Somit ist f integrierbar nach dem Riemannschen Integrabilitätskriterium 19.11. □

Beispiel 19.14
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 $Z^n = \big(0, \frac{1}{n}, 1\big)$ für $n \geq 1$. Dann gilt

$$\mathrm{US}(f, Z^n) = \left(\frac{1}{n} - 0\right) \cdot 0 + \left(1 - \frac{1}{n}\right) \cdot 1 = 1 - \frac{1}{n}$$

und

$$\mathrm{OS}(f, Z^n) = \left(\frac{1}{n} - 0\right) \cdot 1 + \left(1 - \frac{1}{n}\right) \cdot 1 = 1.$$

Wir erhalten also

$$1 \longleftarrow 1 - \frac{1}{n} = \mathrm{US}(f, Z^n) \leq \mathrm{UI}(f) \leq \mathrm{OI}(f) \leq \mathrm{OS}(f, Z^n) = 1.$$

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 19.15 (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 $f(a) \le f(x) \le f(b)$ für alle $x \in [a,b]$.

Sei $\varepsilon > 0$ gegeben. Wir wählen eine natürliche Zahl n so, daß

$$\frac{1}{n} < \frac{\varepsilon}{(b-a)\cdot(f(b)-f(a))}, \tag{48}$$

und betrachten die Zerlegung $Z = (x_0, \ldots, x_n)$ mit

$$x_i := a + i \cdot \frac{(b-a)}{n}.$$

Da f monoton wachsend ist, ist

$$\sup\big\{f(x) \;\big|\; x \in [x_{i-1}, x_i]\big\} = f(x_i)$$

und

$$\inf\big\{f(x) \;\big|\; x \in [x_{i-1}, x_i]\big\} = f(x_{i-1}).$$

Für die Ober- und Untersumme von f bezüglich Z folgt damit

$$\begin{aligned}
\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) &= \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \big(f(x_i) - f(x_{i-1})\big) \\
&= \sum_{i=1}^{n} \frac{b-a}{n} \cdot \big(f(x_i) - f(x_{i-1})\big) \\
&= \frac{b-a}{n} \cdot \sum_{i=1}^{n} \big(f(x_i) - f(x_{i-1})\big) \\
&= \frac{b-a}{n} \cdot \big(f(x_n) - f(x_0)\big) \\
&= \frac{b-a}{n} \cdot \big(f(b) - f(a)\big) \overset{(48)}{<} \varepsilon.
\end{aligned}$$

Somit ist f integrierbar nach dem Riemannschen Integrabilitätskriterium 19.11. □

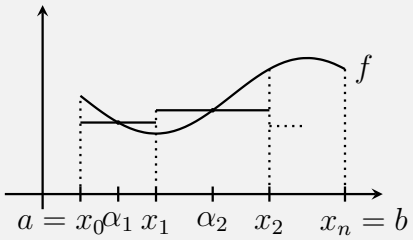
Beispiel 19.16
Die Funktion in Beispiel 19.14 ist monoton wachsend und deshalb nach Proposition 19.15 auch integrierbar. 19.15 sagt aber nichts über den Wert des Integrals aus!

E) Riemannsche Zwischensummen

Definition 19.17 (Riemannsche Zwischensummen)
Sei $f : [a,b] \to \mathbb{R}$ beschränkt, $a < b$, und $Z = (x_0, \ldots, x_n)$ eine Zerlegung von $[a,b]$.
Erfüllt $\alpha = (\alpha_1, \ldots, \alpha_n)$ die Bedingung $\alpha_i \in [x_{i-1}, x_i]$ für $i = 1, \ldots, n$, so nennen wir

$$\mathrm{ZS}(f, Z, \alpha) := \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot f(\alpha_i)$$

die *Riemannsche Zwischensumme* von f bezüglich der Zerlegung Z und den *Zwischenpunkten* α.



Das nächste Lemma sagt, daß man Obersummen und Untersummen beliebig gut approximieren kann durch Zwischensummen.

Lemma 19.18 (Approximation von Ober- und Unter- durch Zwischensummen)
Sei $f : [a,b] \longrightarrow \mathbb{R}$ beschränkt, $a < b$, Z eine Zerlegung von $[a,b]$ und $\varepsilon > 0$.

a. Dann gibt es Zwischenpunkte α mit $0 \le \mathrm{OS}(f, Z) - \mathrm{ZS}(f, Z, \alpha) < \varepsilon$.

b. Dann gibt es Zwischenpunkte β mit $0 \le \mathrm{ZS}(f, Z, \beta) - \mathrm{US}(f, Z) < \varepsilon$.

Beweis: Sei $Z = (x_0, \ldots, x_n)$ und sei

$$M_i := \sup \big\{ f(x) \;\big|\; x \in [x_{i-1}, x_i] \big\}.$$

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 $\alpha = (\alpha_1, \ldots, \alpha_n)$

$$\begin{aligned} \mathrm{OS}(f, Z) - \mathrm{ZS}(f, Z, \alpha) &= \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \big(M_i - f(\alpha_i)\big) \\ &< \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \frac{\varepsilon}{b-a} \quad = \quad (x_n - x_0) \cdot \frac{\varepsilon}{b-a} = \varepsilon. \end{aligned}$$

Damit ist a. gezeigt, und b. zeigt man analog. □

F) Riemannsches Folgenkriterium für Integrierbarkeit

Satz 19.19 (Riemannsches Folgenkriterium für Integrierbarkeit)
Es sei $f : [a,b] \longrightarrow \mathbb{R}$ 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 $l(Z^n) \to 0$ gilt

$$\mathrm{ZS}(f, Z^n, \alpha^n) \longrightarrow I.$$

Beweis:

$\Longrightarrow$**:** Es sei $(Z^n, \alpha^n)_{n\in\mathbb{N}}$ eine Folge von Zerlegungen von $[a,b]$ mit Zwischenpunkten, so daß $\lim\limits_{n\to\infty} l(Z^n) = 0$, und sei $I = \int_a^b f(x)\,dx$.
Sei $\varepsilon > 0$ gegeben. Wir müssen ein $n_\varepsilon \in \mathbb{N}$ finden, so daß

$$|\mathrm{ZS}(f, Z^n, \alpha^n) - I| < \varepsilon \tag{49}$$

für alle $n \geq n_\varepsilon$.
Da f integrierbar ist, gibt es nach Lemma 19.12 ein $\delta_\varepsilon > 0$, so daß für eine Zerlegung Z von $[a,b]$ aus $l(Z) < \delta_\varepsilon$ auch

$$\mathrm{OS}(f, Z) - \mathrm{US}(f, Z) < \varepsilon \tag{50}$$

gilt. Wegen $\lim\limits_{n\to\infty} l(Z^n) = 0$ 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 (50)

$$\mathrm{ZS}(f, Z^n, \alpha^n) - I \leq \mathrm{OS}(f, Z^n) - I \leq \mathrm{OS}(f, Z^n) - \mathrm{US}(f, Z^n) < \varepsilon$$

her, sowie

$$\mathrm{ZS}(f, Z^n, \alpha^n) - I \geq \mathrm{US}(f, Z^n) - I \geq \mathrm{US}(f, Z^n) - \mathrm{OS}(f, Z^n) > -\varepsilon.$$

Damit ist (49) für $n \geq n_\varepsilon$ erfüllt, und das heißt $\mathrm{ZS}(f, Z^n, \alpha^n) \longrightarrow I$.

$\Longleftarrow$**:** Wir wollen das Riemannsche Integrabilitätskriterium anwenden.
Sei dazu $\varepsilon > 0$ gegeben. Wir betrachten die Zerlegung $Z^n = (x_0^n, \ldots, x_n^n)$, $n \geq 1$, mit

$$x_i^n := a + i \cdot \frac{(b-a)}{n}$$

für $i = 0, \ldots, n$. Dann gilt

$$\lim_{n\to\infty} l(Z^n) = \lim_{n\to\infty} \frac{b-a}{n} = 0.$$

Wir wollen zunächst

$$\lim_{n\to\infty} \mathrm{OS}(f, Z^n) = I \tag{51}$$

zeigen. Sei dazu $\varepsilon > 0$ gegeben, dann müssen wir ein n_ε finden, so daß

$$|\operatorname{OS}(f, Z^n) - I| < \varepsilon$$

für alle $n \geq n_\varepsilon$ gilt.
Mit Lemma 19.18 finden wir zu $n \in \mathbb{N}$ Zwischenpunkte α^n, so daß

$$\operatorname{OS}(f, Z^n) - \operatorname{ZS}(f, Z^n, \alpha^n) < \frac{\varepsilon}{2}.$$

Außerdem gilt nach Voraussetzung

$$\lim_{n\to\infty} \operatorname{ZS}(f, Z^n, \alpha^n) = I,$$

so daß es ein n_ε gibt mit

$$|\operatorname{ZS}(f, Z^n, \alpha^n) - I| < \frac{\varepsilon}{2}$$

für alle $n \geq n_\varepsilon$. Sei nun $n \geq n_\varepsilon$ gegeben, dann folgt

$$\begin{aligned} |\operatorname{OS}(f, Z^n) - I| \leq & |\operatorname{OS}(f, Z^n) - \operatorname{ZS}(f, Z^n, \alpha^n)| + |\operatorname{ZS}(f, Z^n, \alpha^n) - I| \\ < & \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Damit ist (51) gezeigt, und analog sieht man

$$\lim_{n\to\infty} \operatorname{US}(f, Z^n) = I$$

Damit erhalten wir dann insgesamt

$$I \longleftarrow \operatorname{US}(f, Z^n) \leq \operatorname{UI}(f) \leq \operatorname{OI}(f) \leq \operatorname{OS}(f, Z^n) \longrightarrow I,$$

so daß aus dem Einschachtelungssatz 11.17

$$I = \operatorname{OI}(f) = \operatorname{UI}(f) = \int_a^b f(x)\,dx$$

und insbesondere die Integrierbarkeit von f 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 $Z^n = (x_0, \ldots, x_n)$ eine Zerlegung von $[0, b]$ mit Zwischenpunkten $\alpha^n = (x_1, \ldots, x_n)$, und es gilt $l(Z^n) \longrightarrow 0$. Um die Zwischensumme berechnen zu können, verwenden wir die Formel

$$\sum_{i=1}^{n} i^2 = \frac{n \cdot (n+1) \cdot (2n+1)}{6}, \tag{52}$$

die man mit Hilfe von Induktion leicht zeigen kann. Damit ergibt sich

$$\begin{aligned}\mathrm{ZS}(f, Z^n, \alpha^n) &= \sum_{i=1}^{n}(x_i - x_{i-1}) \cdot x_i^2 = \sum_{i=1}^{n} \frac{b}{n} \cdot \frac{i^2 \cdot b^2}{n^2} \\ &= \frac{b^3}{n^3} \cdot \sum_{i=1}^{n} i^2 \overset{(52)}{=} \frac{b^3}{n^3} \cdot \frac{n \cdot (n+1) \cdot (2n+1)}{6} \\ &= \frac{b^3}{3} \cdot \left(1 + \frac{1}{n}\right) \cdot \left(1 + \frac{1}{2n}\right) \longrightarrow \frac{b^3}{3}.\end{aligned}$$

Aus dem Riemannschen Folgenkriterium für Integrierbarkeit 19.19 folgt dann

$$\int_0^b x^2\, dx = \frac{b^3}{3}.$$

G) Rechenregeln für Integrale – Linearität und Monotonie

Korollar 19.21 (Linearität und Monotonie des Integrals)
Seien $f, g : [a, b] \longrightarrow \mathbb{R}$ 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 $Z = (x_0, \ldots, x_n)$ von $[a, b]$ mit Zwischenpunkten $\alpha = (\alpha_1, \ldots, \alpha_n)$ offenbar gilt:

$$\begin{aligned}\mathrm{ZS}(cf + dg, Z, \alpha) &= \sum_{i=1}^{n}(x_i - x_{i-1}) \cdot (cf + dg)(\alpha_i) \\ &= c \cdot \sum_{i=1}^{n}(x_i - x_{i-1}) \cdot f(\alpha_i) + d \cdot \sum_{i=1}^{n}(x_i - x_{i-1}) \cdot g(\alpha_i) \\ &= c \cdot \mathrm{ZS}(f, Z, \alpha) + d \cdot \mathrm{ZS}(g, Z, \alpha).\end{aligned}$$

Es sei nun $(Z^n, \alpha^n)_{n \in \mathbb{N}}$ eine Folge von Zerlegungen von $[a, b]$ und Zwischenpunkten mit $l(Z^n) \longrightarrow 0$. Aus den Grenzwertsätzen für Folgen 11.15 und Satz 19.19 folgt dann

$$\mathrm{ZS}(cf + dg, Z^n, \alpha^n) = c \cdot \mathrm{ZS}(f, Z^n, \alpha^n) + d \cdot \mathrm{ZS}(g, Z^n, \alpha^n) \longrightarrow c \cdot \int_a^b f(x)\, dx + d \cdot \int_a^b g(x)\, dx.$$

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 $l(Z^n) \longrightarrow 0$. Wegen $f(x) \le g(x)$ für alle $x \in [a,b]$ gilt dann offenbar

$$\int_a^b f(x)\,dx \longleftarrow \mathrm{ZS}(f, Z^n, \alpha^n) \le \mathrm{ZS}(g, Z^n, \alpha^n) \longrightarrow \int_a^b g(x)\,dx,$$

wobei die Grenzwerte aus dem Riemannschen Folgenkriterium für Integrierbarkeit folgen. Damit gilt dann aber auch für die Grenzwerte

$$\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx.$$

□

Beispiel 19.22

Aus Beispiel 19.10 und 19.20 erhalten wir aus der Linearität des Integrals

$$\int_0^b 3x^2 + 5\,dx = 3 \cdot \int_0^b x^2\,dx + \int_0^b 5\,dx = b^3 + 5b.$$

H) Rechenregeln für Integrale – Additivität

Bemerkung 19.23 (Aneinanderhängen von Zerlegungen)

Ist $Z' = (x_0, \ldots, x_n)$ eine Zerlegung von $[a,c]$ und $Z'' = (y_0, \ldots, y_m)$ eine Zerlegung von $[c,b]$, so ist $Z' * Z'' = (x_0, \ldots, x_n, y_1, \ldots, y_m)$ eine Zerlegung von $[a,b]$ und sie entsteht durch aneinanderhängen der beiden Zerlegungen. Ist $\alpha = (\alpha_1, \ldots, \alpha_n)$ ein Tupel von Zwischenpunkten von Z' und $\beta = (\beta_1, \ldots, \beta_m)$ ein Tupel von Zwischenpunkten von Z'', so definieren wir $\alpha \sqcup \beta = (\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m)$ und erhalten damit ein Tupel von Zwischenpunkten von $Z' * Z''$.

Außerdem gelten offenbar

$$\mathrm{OS}(f, Z' * Z'') = \mathrm{OS}(f, Z') + \mathrm{OS}(f, Z''),$$

$$\mathrm{US}(f, Z' * Z'') = \mathrm{US}(f, Z') + \mathrm{US}(f, Z''),$$

$$\mathrm{ZS}(f, Z' * Z'', \alpha \sqcup \beta) = \mathrm{ZS}(f, Z', \alpha) + \mathrm{ZS}(f, Z'', \beta).$$

Proposition 19.24 (Additivität des Integrals)

Es sei $f : [a,b] \longrightarrow \mathbb{R}$ 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 Riemannschen Integrabilitätskriteriums 19.11 Zerlegungen $Z' = (x_0, \dots, x_n)$ von $[a,c]$ und $Z'' = (y_0, \dots, y_m)$ von $[c,b]$, so daß

$$\mathrm{OS}(f, Z') - \mathrm{US}(f, Z') < \frac{\varepsilon}{2}$$

und

$$\mathrm{OS}(f, Z'') - \mathrm{US}(f, Z'') < \frac{\varepsilon}{2}.$$

Dann ist aber $Z = Z' * Z'' = (x_0, \dots, x_n, y_1, \dots, y_m)$ eine Zerlegung von $[a,b]$ und

$$\begin{aligned}\mathrm{OS}(f, Z) - \mathrm{US}(f, Z) =& \big(\mathrm{OS}(f, Z') + \mathrm{OS}(f, Z'')\big) - \big(\mathrm{US}(f, Z') + \mathrm{US}(f, Z'')\big)\\ =& \big(\mathrm{OS}(f, Z') - \mathrm{US}(f, Z')\big) + \big(\mathrm{OS}(f, Z'') - \mathrm{US}(f, Z'')\big)\\ <& \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

Das Riemannsche Integrabilitätskriterium 19.11 impliziert also, daß f auf $[a,b]$ integrierbar ist.

Ist umgekehrt f auf $[a,b]$ integrierbar, so gibt es wegen des Riemannschen Integrabilitätskriteriums eine Zerlegung $Z = (x_1, \dots, x_n)$ von $[a,b]$ mit

$$\mathrm{OS}(f, Z) - \mathrm{US}(f, Z) < \varepsilon.$$

Nach eventueller Verfeinerung können wir ohne Einschränkung annehmen, daß $c = x_j \in \mathrm{supp}(Z)$ 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

$$\begin{aligned}&\big(\mathrm{OS}(f, Z') - \mathrm{US}(f, Z')\big) + \big(\mathrm{OS}(f, Z'') - \mathrm{US}(f, Z'')\big)\\ &= \big(\mathrm{OS}(f, Z') + \mathrm{OS}(f, Z'')\big) - \big(\mathrm{US}(f, Z') + \mathrm{US}(f, Z'')\big)\\ &= \mathrm{OS}(f, Z) - \mathrm{US}(f, Z) < \varepsilon.\end{aligned}$$

Mithin gilt auch

$$\mathrm{OS}(f, Z') - \mathrm{US}(f, Z') < \varepsilon \quad \text{und} \quad \mathrm{OS}(f, Z'') - \mathrm{US}(f, Z'') < \varepsilon,$$

so daß aus dem Riemannschen Integrabilitätskriterium 19.11 wieder folgt, daß f auf $[a,c]$ und auf $[c,b]$ integrierbar ist.

Wir wählen nun zwei Folgen $(Z^{n\prime}, \alpha^n)_{n\in\mathbb{N}}$ und $(Z^{n\prime\prime}, \beta^n)_{n\in\mathbb{N}}$ von Zerlegungen von $[a,c]$ bzw. von $[c,b]$ mit Zwischenpunkten, so daß $\lim\limits_{n\to\infty} l(Z^{n\prime}) = \lim\limits_{n\to\infty} l(Z^{n\prime\prime}) = 0$. Wie oben können wir die Zerlegungen $Z^{n\prime}$ und $Z^{n\prime\prime}$ zu einer Zerlegung $Z^n := Z^{n\prime} * Z^{n\prime\prime}$ 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 $l(Z^n) = \max\{l(Z^{n\prime}), l(Z^{n\prime\prime})\} \longrightarrow 0$, und somit folgt aus dem Folgenkriterium für Integrierbarkeit 19.19

$$\int_a^b f(x)\,dx \longleftarrow \mathrm{ZS}(f, Z^n, \gamma^n) = \mathrm{ZS}(f, Z^{n\prime}, \alpha^n) + \mathrm{ZS}(f, Z^{n\prime\prime}, \beta^n) \longrightarrow \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

□

Beispiel 19.25

Aus Proposition 19.24 und Beispiel 19.20 erhalten wir für $0 < a < b$

$$\int_a^b x^2\,dx = \int_0^b x^2\,dx - \int_0^a x^2\,dx = \frac{b^3}{3} - \frac{a^3}{3}.$$

I) Rechenregeln für Integrale – Dreiecksungleichung

Die Dreiecksungleichung für Summen liefert mit Induktion, daß

$$|a_1 + \ldots + a_n| \leq |a_1| + \ldots + |a_n|$$

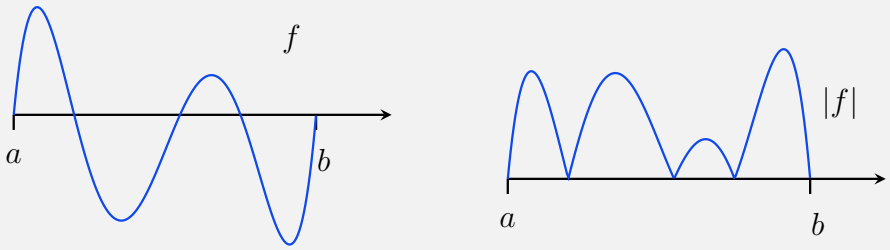
gilt. Integrale sind verallgemeinerte Summen, und die Dreiecksungleichung nimmt dann die folgende Gestalt an.

Proposition 19.26 (Dreiecksungleichung für Integrale)

Ist $f : [a,b] \longrightarrow \mathbb{R}$ 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

$$f_+ : [a,b] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } f(x) \geq 0, \\ 0, & \text{falls } f(x) < 0. \end{cases}$$

Dann gilt (siehe Abbildung 11)

$$|f| = 2 \cdot f_+ - f.$$

Wir wollen nun mit Hilfe des Riemannschen Integrabilitätskriteriums zeigen, daß f_+ auf $[a,b]$ integrierbar ist. Sei $\varepsilon > 0$ gegeben. Da f auf $[a,b]$ integrierbar ist, gibt es eine Zerlegung $Z = (x_0, \ldots, x_n)$ von $[a,b]$, so daß

$$\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon.$$

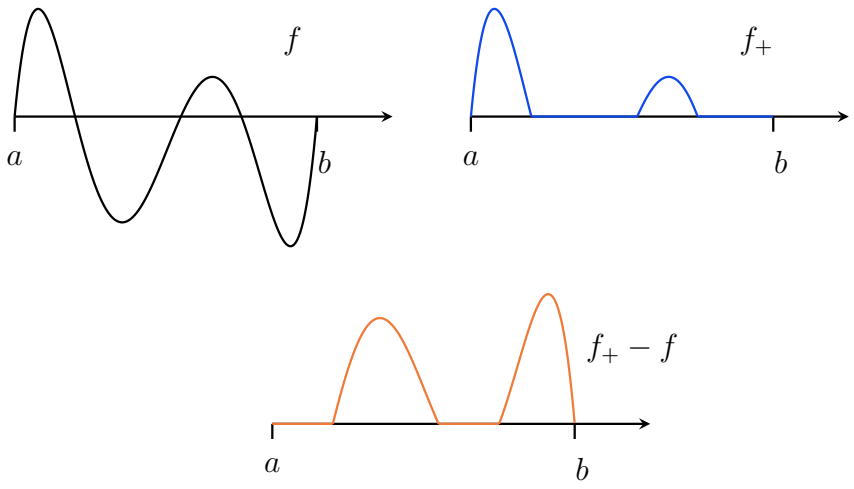


Abbildung 11. Die Graphen von f_+ und f_-

Wir behaupten, daß für jede Teilmenge $I \subseteq [a,b]$ die Ungleichung

$$\sup\{f_+(x) \mid x \in I\} - \inf\{f_+(x) \mid x \in I\} \leq \sup\{f(x) \mid x \in I\} - \inf\{f(x) \mid x \in I\}, \tag{53}$$

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 (53) 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 (53) nicht-negativ ist. In diesem Fall gilt (53).
2. **Fall:** $\exists\, y, z \in I$ **mit** $f(y) < 0 \leq f(z)$**:** Also $\sup\{f_+(x) |\ x \in I\} = \sup\{f(x) |\ x \in I\}$ und $\inf\{f_+(x) |\ x \in I\} = 0 > \inf\{f(x) \mid x \in I\}$. Damit gilt die Ungleichung (53).
3. **Fall:** $f(x) \geq 0$ **für alle** $x \in I$**:** Dann ist $f = f_+$ auf I und (53) gilt.

Damit haben wir gezeigt, daß (53) stets erfüllt ist. Für die Differenz der Ober- und Untersumme von f_+ ergibt sich mit $I_i := [x_{i-1}, x_i]$ dann

$$\begin{aligned}
\mathrm{OS}(f_+, Z) - \mathrm{US}(f_+, Z) &= \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \big(\sup\{f_+(x) \mid x \in I_i\} - \inf\{f_+(x) \mid x \in I_i\}\big) \\
&\leq \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \big(\sup\{f(x) \mid x \in I_i\} - \inf\{f(x) \mid x \in I_i\}\big) \\
&= \mathrm{OS}(f, Z) - \mathrm{US}(f, Z) < \varepsilon.
\end{aligned}$$

Mit Hilfe des Riemannschen 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

$$|\operatorname{ZS}(f,Z,\alpha)| = \left|\sum_{i=1}^{n}(x_i-x_{i-1})\cdot f(\alpha_i)\right| \le \sum_{i=1}^{n}(x_i-x_{i-1})\cdot|f(\alpha_i)| = \operatorname{ZS}(|f|,Z,\alpha).$$

Sei nun $(Z^n,\alpha^n)_{n\in\mathbb{N}}$ eine Folge von Zerlegungen von $[a,b]$ und Zwischenpunkten mit $\lim\limits_{n\to\infty} l(Z^n)=0$, dann folgt

$$\left|\int_a^b f(x)\,dx\right| \longleftarrow |\operatorname{ZS}(f,Z^n,\alpha^n)| \le \operatorname{ZS}(|f|,Z^n,\alpha^n) \longrightarrow \int_a^b |f(x)|\,dx.$$

Die Ungleichung bleibt für die Grenzwerte erhalten. □

Bemerkung 19.27

Es sei $f:[a,b]\longrightarrow\mathbb{R}$ 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 $f:[a,b]\longrightarrow\mathbb{R}$ *stückweise stetig*, wenn es eine Zerlegung $Z=(x_0,\dots,x_n)$ von $[a,b]$ gibt, so daß die Funktionen $f_i:[x_{i-1},x_i]\longrightarrow\mathbb{R}: x\mapsto f(x)$ für $i=1,\dots,n$ auf (x_{i-1},x_i) stetig sind und so daß die Grenzwerte $\lim\limits_{x\to x_{i-1}} f_i(x)$ und $\lim\limits_{x\to x_i} f_i(x)$ in $\mathbb{R}$ existieren.
Zeige, eine stückweise stetige Funktion $f:[a,b]\longrightarrow\mathbb{R}$ ist integrierbar auf $[a,b]$.

Aufgabe 19.29

Sei $f:[a,b]\longrightarrow\mathbb{R}$ eine stetige Funktion mit $f(x)\ge 0$ 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.

Aufgabe 19.30

Bestimme die folgenden Integrale.

a. $\int_0^{\frac{\pi}{3}} \frac{3}{\cos^2(x)} dx$.

b. $\int_0^1 \frac{1}{1+x} dx$.

c. $\int_{-1}^2 \left(8 \cdot (x-2)^3 + \frac{1}{\sqrt{x+2}}\right) dx$.

d. $\int_0^{\frac{\pi}{4}} \sin(x)\cos(x) dx$.

e. $\int_0^{\frac{1}{4}} x^2 e^{4x} dx$.

Aufgabe 19.31

Betrachte für $n \in \mathbb{N}$ die Zerlegung $Z^n = (0, \frac{1}{2^n}, \frac{2}{2^n}, \ldots, \frac{2^n-1}{2^n}, 1)$ des Intervalls $[0,1]$ mit den Zwischenpunkten $\alpha^n = (\frac{1}{2^n}, \frac{2}{2^n}, \ldots, \frac{2^n-1}{2^n}, 1)$. Zeige die folgenden Aussagen.

a. $\mathrm{ZS}(\exp, Z^n, \alpha^n) = (e-1) \cdot e^y \cdot \frac{1}{\frac{e^y-1}{y}}$ für $y = \frac{1}{2^n}$.

b. $\lim\limits_{y \to 0} \frac{e^y-1}{y} = 1$.

c. Berechne $\int_0^1 e^x dx$ mit Hilfe der Zwischensumme aus Aufgabenteil a..

Aufgabe 19.32

Es sei $f : [0,1] \longrightarrow \mathbb{R}_{\geq 0}$ eine Funktion, so daß es für jedes $\varepsilon > 0$ nur endlich viele Werte $x \in [0,1]$ gibt mit $f(x) > \varepsilon$. Zeige, f ist integrierbar auf $[0,1]$ mit $\int_0^1 f(x)\, dx = 0$.

§ 20 Hauptsatz der Differential- und Integralrechnung mit Anwendungen

A) Hauptsatz der Differential- und Integralrechnung

Definition 20.1 (Stammfunktion)
Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \longrightarrow \mathbb{R}$ eine Funktion.
Eine differenzierbare Funktion $F : I \longrightarrow \mathbb{R}$ mit $F' = f$ heißt *Stammfunktion* von f.

Proposition 20.2 (Stammfunktionen sind eindeutig bis auf eine Konstante.)
Sei I ein Intervall, $f : I \longrightarrow \mathbb{R}$ und $F, G : I \longrightarrow \mathbb{R}$ 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 Proposition 18.12, 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 $F : [0, \infty) \longrightarrow \mathbb{R} : x \mapsto \frac{x^3}{3}$ ist eine Stammfunktion von $f : [0, \infty) \longrightarrow \mathbb{R} : x \mapsto x^2$, 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, $f : I \longrightarrow \mathbb{R}$ stetig und $a \in I$.
Dann ist die Funktion $F : I \longrightarrow \mathbb{R} : y \mapsto \int_a^y f(x)\, dx$ 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

$$|f(x) - f(c)| < \frac{\varepsilon}{2} \tag{54}$$

für alle $x \in I$ mit $|x - c| < \delta_\varepsilon$. Sei nun $c \neq y \in I$ mit $|y - c| < \delta_\varepsilon$, dann gilt

$$\begin{aligned}
\left| \frac{F(y) - F(c)}{y - c} - f(c) \right| &= \left| \frac{1}{y - c} \cdot \left(\int_a^y f(x)\,dx - \int_a^c f(x)\,dx \right) - f(c) \right| \\
&\overset{19.24}{=} \left| \frac{1}{y - c} \cdot \int_c^y f(x)\,dx - f(c) \right| \\
&= \left| \frac{1}{y - c} \cdot \int_c^y f(x)\,dx - \frac{f(c) \cdot (y - c)}{y - c} \right| \\
&= \left| \frac{1}{y - c} \cdot \int_c^y f(x)\,dx - \frac{\int_c^y f(c)\,dx}{y - c} \right| \\
&\overset{19.21}{=} \left| \frac{1}{y - c} \cdot \int_c^y \big(f(x) - f(c)\big)\,dx \right| \\
&\overset{19.26}{\leq} \frac{1}{|y - c|} \cdot \left| \int_c^y \big|f(x) - f(c)\big|\,dx \right| \\
&\overset{(54),19.21}{\leq} \frac{1}{|y - c|} \cdot \left| \int_c^y \frac{\varepsilon}{2}\,dx \right| \\
&= \frac{|y - c|}{|y - c|} \cdot \frac{\varepsilon}{2} < \varepsilon.
\end{aligned}$$

□

Korollar 20.5 (Hauptsatz der Differential- und Integralrechnung)
Es sei $f : [a, b] \longrightarrow \mathbb{R}$ 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 $f : I \longrightarrow \mathbb{R}$ stetig.

a. Der Hauptsatz der Differential- und Integralrechnung 20.4 besagt im wesentlichen, daß die Differentiation die Umkehrung der Integration ist.

b. Ist $F : I \longrightarrow \mathbb{R}$ 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 $F(y) = \int^y f(x)\,dx$.

B) Stammfunktionen aus Ableitungen ablesen

Beispiel 20.7 (Einige ausgewählte Stammfunktionen)

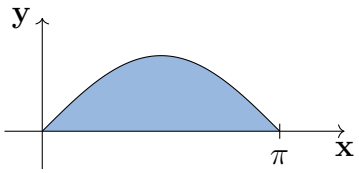
In den Abschnitten 17 und 18 haben wir für eine Vielzahl stetig differenzierbarer Abbildungen die Ableitungen kennengelernt. Im Umkehrschluß haben wir damit für die Ableitungsfunktionen auch Stammfunktionen gefunden. Wir wollen für einige wichtige Beispiele von Funktionen f hier die Stammfunktionen F in tabellarischer Form zusammenstellen.

f	$F = \int f(x)\,dx$	f	$F = \int f(x)\,dx$
$\exp$	$\exp$	$\exp_a,\quad a \neq 1$	$\frac{1}{\ln(a)} \cdot \exp_a$
$\cos$	$\sin$	$\sin$	$-\cos$
$x \mapsto \frac{1}{x}$	$\ln$	$x \mapsto x^a,\quad -1 \neq a \in \mathbb{R}$	$x \mapsto \frac{1}{a+1} \cdot x^{a+1}$
$x \mapsto \frac{1}{x^2+1}$	$\arctan$	$\frac{1}{\cos^2}$	$\tan$

Beispiel 20.8 (Flächeninhalt eines Sinusbogens)

Wir können den Flächeninhalt unter einem der Bögen der Sinusfunktion berechnen als

$$\int_0^\pi \sin(x)\,dx = -\cos(x)\big|_0^\pi = -\cos(\pi) + \cos(0) = 1 + 1 = 2.$$



C) Der Mittelwertsatz der Integralrechnung

Korollar 20.9 (Mittelwertsatz der Integralrechnung)
Es sei $f : [a, b] \longrightarrow \mathbb{R}$ 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 $F : [a, b] \longrightarrow \mathbb{R} : y \mapsto \int_a^y f(x)\,dx$ 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.

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

$$f(y) \le f(x) \le f(z)$$

für alle $x \in [a, b]$. Für die Zerlegung $Z = (a, b)$ gilt dann

$$(b - a) \cdot f(y) = \mathrm{US}(f, Z) \le \int_a^b f(x)\,dx \le \mathrm{OS}(f, Z) = (b - a) \cdot f(z),$$

und damit

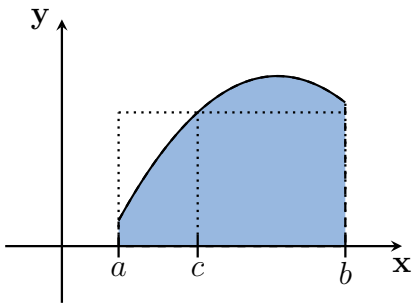
$$f(y) \le \frac{1}{b - a} \cdot \int_a^b f(x)\,dx \le f(z).$$

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.



D) Partielle Integration

Satz 20.11 (Partielle Integration)
Sind $u, v : [a,b] \longrightarrow \mathbb{R}$ 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. □

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

$$\begin{aligned}\int^y \cos^2(x)\,dx &= \int^y u(x)\cdot v'(x)\,dx = u(x)\cdot v(x)\Big|^y - \int^y u'(x)\cdot v(x)\,dx\\ &= \cos(x)\cdot\sin(x)\Big|^y - \int^y -\sin^2(x)\,dx\\ &= \cos(x)\cdot\sin(x)\Big|^y - \int^y \cos^2(x) - 1\,dx\\ &= \cos(x)\cdot\sin(x)\Big|^y - \int^y \cos^2(x)\,dx + \int^y 1\,dx\\ &= \cos(x)\cdot\sin(x)\Big|^y - \int^y \cos^2(x)\,dx + x\Big|^y.\end{aligned}$$

Addieren wir auf beiden Seiten $\int^y \cos^2(x)\,dx$ und teilen durch 2, so erhalten wir

$$\int^y \cos^2(x)\,dx = \frac{1}{2}\cdot\big(y + \cos(y)\cdot\sin(y)\big).$$

E) 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 $f : I \longrightarrow \mathbb{R}$ $n+1$-fach stetig differenzierbar auf dem Intervall I und $x, a \in I$. Dann gilt

$$f(x) - T^n_{f,a}(x) = \int_a^x \frac{f^{(n+1)}(y)}{n!}\cdot(x-y)^n\,dy.$$

Beweis: Wir führen den Beweis durch Induktion nach n. Für $n = 0$ gilt

$$f(x) - T^0_{f,a}(x) = f(x) - f(a) \overset{20.5}{=} \int_a^x f'(y)\,dy = \int_a^x \frac{f'(y)}{0!}\cdot(x-y)^0\,dy$$

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

$$f(x) - T^{n-1}_{f,a}(x) = \int_a^x \frac{f^{(n)}(y)}{(n-1)!}\cdot(x-y)^{n-1}\,dy.$$

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. □

F) Die Substitutionsregel

Satz 20.15 (Substitutionsregel)
Es sei $I \subseteq \mathbb{R}$ ein Intervall, $f : I \longrightarrow \mathbb{R}$ stetig und $\varphi : [a,b] \longrightarrow \mathbb{R}$ stetig differenzierbar mit $\mathrm{Im}(\varphi) \subseteq I$. Dann gilt

$$\int_{\varphi(a)}^{\varphi(b)} f(x)\, dx = \int_a^b f\big(\varphi(x)\big) \cdot \varphi'(x)\, dx.$$

Beweis: Da φ stetig ist, nimmt φ sein Minimum und sein Maximum an, d.h. es gibt $y, z \in [a,b]$ mit $\varphi(y) \le \varphi(x) \le \varphi(z)$ für alle $x \in [a,b]$, und mithin ist

$$\mathrm{Im}(\varphi) = \big[\varphi(y), \varphi(z)\big] \subseteq I$$

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

$$(F \circ \varphi)'(x) = F'\big(\varphi(x)\big) \cdot \varphi'(x) = f\big(\varphi(x)\big) \cdot \varphi'(x),$$

so daß $F \circ \varphi$ eine Stammfunktion von $x \mapsto f\big(\varphi(x)\big) \cdot \varphi'(x)$ auf $[a,b]$ ist. Dann können wir Korollar 20.5 anwenden und erhalten

$$\int_a^b f\big(\varphi(x)\big) \cdot \varphi'(x)\, dx \overset{20.5}{=} (F\circ\varphi)(b) - (F\circ\varphi)(a) = F\big(\varphi(b)\big) - F\big(\varphi(a)\big) \overset{20.5}{=} \int_{\varphi(a)}^{\varphi(b)} f(x)\, dx.$$

□

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 x die Variable z zu verwenden, so daß die Formel folgende Gestalt hat:

$$\int_{\varphi(a)}^{\varphi(b)} f(z)\,dz = \int_a^b f\big(\varphi(x)\big)\cdot\varphi'(x)\,dx.$$

Man sagt dann, daß man $\varphi(x)$ durch z substituiert oder umgekehrt, je nachdem ob man die linke durch die rechte Seite ausrechnen will oder umgekehrt. Man schreibt $z = \varphi(x)$.

Diese Schreibweise kann man nutzen, um sich für die Substitution eine Eselsbrücke zu bauen. In Anlehnung an die Schreibweise $\varphi' = \frac{\partial\varphi}{\partial x}$ kann man mit $z = \varphi(x)$ dann auch

$$\varphi'(x)\,dx = \frac{dz}{dx}\,dx = dz$$

schreiben. Damit wird aus der Substitutionsformel ohne Integralgrenzen dann

$$\int f(\varphi(x))\cdot\varphi'(x)\,dx = \int f(z)\cdot\frac{dz}{dx}\,dx = \int f(z)\,dz.$$

c. Man kann mit Hilfe der Substitutionsregel auch Stammfunktionen ausrechnen, indem man die Integrationsgrenze b durch die Variable y ersetzt und a ignoriert.

Beispiel 20.17 (Stammfunktion von $x \mapsto x\cdot\exp(x^2)$)

Wir wollen das Integral $\int_a^b x\cdot\exp\big(x^2\big)\,dx$ für $a, b \in \mathbb{R}$ berechnen. Dazu substituieren wir $z = x^2$, d.h. wir betrachten $\varphi : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$, $\varphi'(x) = 2x$ und $f : \mathbb{R} \longrightarrow \mathbb{R} : z \mapsto \frac{\exp(z)}{2}$. Da zudem f eine Stammfunktion von f ist, folgt damit

$$\begin{aligned}\int_a^b x\cdot\exp\big(x^2\big)\,dx &= \int_a^b f\big(\varphi(x)\big)\cdot\varphi'(x)\,dx = \int_{\varphi(a)}^{\varphi(b)} f(z)\,dz\\ &= \int_{a^2}^{b^2}\frac{\exp(z)}{2}\,dz = \frac{\exp\big(b^2\big)}{2} - \frac{\exp\big(a^2\big)}{2}.\end{aligned}$$

Beispiel 20.18 (Stammfunktion von tan)

Wir wollen eine Stammfunktion für den Tangens auf dem Intervall $\big(-\frac{\pi}{2}, \frac{\pi}{2}\big)$ 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

$$\begin{aligned}\int^y \tan(x)\,dx &= \int^y -\frac{1}{\cos(x)}\cdot\big(-\sin(x)\big)\,dx = \int^y f\big(\varphi(x)\big)\cdot\varphi'(x)\,dx\\ &= \int^{\cos(y)} f(z)\,dz = \int^{\cos(y)} -\frac{1}{z}\,dz\\ &= -\ln(z)\Big|^{\cos(y)} = -\ln\big(\cos(y)\big).\end{aligned}$$

Also ist $-\ln\circ\cos$ eine Stammfunktion von tan auf $\big(-\frac{\pi}{2}, \frac{\pi}{2}\big)$.

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

$$\begin{aligned}
F(y) = \int^{y} \sqrt{1-z^2}\,dz &= \int^{\varphi(b)} f(z)\,dz = \int^{b} f\big(\varphi(x)\big)\cdot\varphi'(x)\,dx \\
&= \int^{\arcsin(y)} \sqrt{1-\sin^2(x)}\cdot\cos(x)\,dx \\
&= \int^{\arcsin(y)} \sqrt{\cos^2(x)}\cdot\cos(x)\,dx \\
&= \int^{\arcsin(y)} \cos^2(x)\,dx \\
&\overset{20.13}{=} \frac{1}{2}\cdot\big(x+\cos(x)\cdot\sin(x)\big)\Big|^{\arcsin(y)} \\
&= \frac{1}{2}\cdot\big(x+\sqrt{\cos^2(x)}\cdot\sin(x)\big)\Big|^{\arcsin(y)} \\
&= \frac{1}{2}\cdot\big(x+\sqrt{1-\sin^2(x)}\cdot\sin(x)\big)\Big|^{\arcsin(y)} \\
&= \frac{\arcsin(y)+y\cdot\sqrt{1-y^2}}{2}.
\end{aligned}$$

eine Stammfunktion von f auf $[-1,1]$.

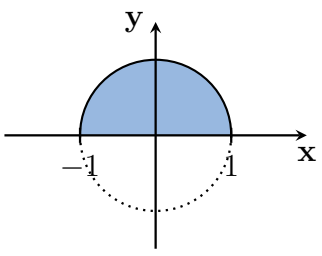
Beispiel 20.20 (Flächeninhalt eines Kreises)
Wir wollen nun den Flächeninhalt eines Kreises berechnen. Die obere Hälfte des Einheitskreises mit dem Ursprung als Mittelpunkt

$$K_1(0) := \big\{(x,y) \in \mathbb{R}^2 \;\big|\; x^2+y^2 = 1\big\}$$

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

$$\int_{-1}^{1} \sqrt{1-x^2}\,dx \overset{20.19}{=} \left.\frac{\arcsin(x) + x\cdot\sqrt{1-x^2}}{2}\right|_{-1}^{1}$$
$$= \frac{\arcsin(1)}{2} - \frac{\arcsin(-1)}{2} = \frac{\pi}{4} - \left(-\frac{\pi}{4}\right) = \frac{\pi}{2}$$

die Hälfte des Flächeninhaltes des Einheitskreises.

Der Kreis mit Radius r und dem Ursprung als Mittelpunkt ist

$$K_r(0) := \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; x^2+y^2 = r^2\right\},$$

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

$$2\cdot\int_{-r}^{r} \sqrt{r^2-x^2}\,dx = 2\cdot r^2\cdot\int_{-r}^{r} \sqrt{1-\frac{x^2}{r^2}}\cdot\frac{1}{r}\,dx$$
$$= 2\cdot r^2\cdot\int_{-1}^{1} \sqrt{1-z^2}\,dz = 2\cdot\frac{\pi}{2}\cdot r^2 = \pi\cdot r^2$$

als Flächeninhalt von $K_r(0)$.

G) Partialbruchzerlegung

Bemerkung 20.21 (Partialbruchzerlegung)

Jede rationale Funktion $r = \frac{f}{g}$ läßt sich schreiben als

$$r = \frac{f}{g} = h + \frac{p}{q}, \tag{55}$$

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

$$\frac{A}{(t-a)^k} \quad \text{und} \quad \frac{Bt+C}{(t^2+bt+c)^k} \tag{56}$$

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 (56) für $k = 1, \ldots, n$ 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 (55) 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 (56) 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

$$A = 0,\ B = 1,\ C = 0,\ D = -1,\ E = 2 \quad \text{und} \quad F = 0,$$

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.

H) Vertauschbarkeit von Grenzwert und Integration

Satz 20.23 (Vertauschbarkeit von Grenzwert und Integration)
Die Folge $(f_n)_{n\in\mathbb{N}}$ stetiger Funktionen $f_n : [a,b] \longrightarrow \mathbb{R}$ konvergiere auf $[a,b]$, $a<b$, gleichmäßig gegen $f : [a,b] \longrightarrow \mathbb{R}$.
Dann ist auch die Grenzfunktion f stetig auf $[a,b]$ und für alle $y \in [a,b]$ gilt

$$\int_a^y f(x)\,dx = \lim_{n\to\infty} \int_a^y f_n(x)\,dx. \tag{57}$$

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 19.13 integrierbar sind, existieren die Integrale in (57). Es bleibt, für $y \in [a,b]$ zu zeigen, daß

$$\lim_{n\to\infty} \int_a^y f_n(x)\,dx = \int_a^y f(x)\,dx.$$

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 $x \in [a,b]$

$$|f_n(x) - f(x)| < \frac{\varepsilon}{2\cdot(b-a)}. \tag{58}$$

Dann gilt für $n \geq n_\varepsilon$ auch

$$\left|\int_a^y f_n(x)\,dx - \int_a^y f(x)\,dx\right| \overset{19.26}{\leq} \int_a^y |f_n(x) - f(x)|\,dx \overset{19.21,(58)}{\leq} \int_a^y \frac{\varepsilon}{2\cdot(b-a)}\,dx \leq \frac{\varepsilon}{2} < \varepsilon.$$

Damit ist die Behauptung gezeigt. □

Bemerkung 20.24

a. Wie in Bemerkung 18.19 wollen wir wieder darauf hinweisen, daß wir in Satz 20.23 gezeigt haben, daß zwei Grenzwertprozesse vertauschen. Es gilt nämlich

$$\lim_{n\to\infty}\lim_{m\to\infty} \mathrm{ZS}(f_n, Z_m, \alpha^m) = \lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx = \lim_{m\to\infty}\lim_{n\to\infty} \mathrm{ZS}(f_n, Z_m, \alpha^m).$$

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

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx.$$

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 18.18 geben:

Die Funktionen f_n in Satz 18.18 sind jeweils Stammfunktion von f_n', und da die f_n' stetig sind, folgt aus dem Hauptsatz der Differential- und Integralrechnung 20.5

$$f_n(y) - f_n(a) = \int_a^y f_n'(x)\,dx.$$

Bilden wir auf beiden Seiten den Grenzwert, so folgt mit Satz 20.23

$$f(y) - f(a) \longleftarrow f_n(y) - f_n(a) = \int_a^y f_n'(x)\,dx \longrightarrow \int_a^y g(x)\,dx.$$

Also ist

$$f(y) = f(a) + \int_a^y g(x)\,dx$$

für $y \in [a,b]$, so daß f nach dem Hauptsatz der Differential- und Integralrechnung 20.4 differenzierbar ist auf $[a,b]$ mit

$$f'(y) = 0 + g(y) = g(y)$$

für alle $y \in [a,b]$.

I) Integration von Potenzreihen

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 $f : (-r,r) \longrightarrow \mathbb{R} : x \mapsto \sum_{n=0}^{\infty} a_n \cdot x^n$ auf jedem Intervall $[a,b] \subset (-r,r)$ integrierbar und

$$F : (-r,r) \longrightarrow \mathbb{R} : y \mapsto \sum_{n=0}^{\infty} \frac{a_n}{n+1} \cdot y^{n+1}$$

ist eine Stammfunktion von f. Sie entsteht durch gliedweises Integrieren.

Beweis: Die Folge stetiger Funktionen $(f_n)_{n\in\mathbb{N}}$ mit

$$f_n : (-r,r) \longrightarrow \mathbb{R} : x \mapsto \sum_{k=0}^{n} a_k \cdot x^k$$

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

$$\int_0^y f(x)\,dx = \lim_{n\to\infty} \int_0^y f_n(x)\,dx \overset{19.21}{=} \lim_{n\to\infty} \sum_{k=0}^{n} a_k \cdot \int_0^y x^k\,dx \overset{20.7}{=} \lim_{n\to\infty} \sum_{k=0}^{n} \frac{a_k}{k+1} \cdot y^{k+1}.$$

Damit ist die Behauptung gezeigt. □

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 $f : (-1, \infty) \longrightarrow \mathbb{R} : x \mapsto \ln(1+x)$ ist differenzierbar mit Ableitung

$$f' : (-1, \infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{1+x}.$$

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 18.37 gilt auf ganz $(0, 2)$, und sie gilt auch im Punkt $a = 2$, wie wir dort bereits gesehen haben.

Aufgaben

Aufgabe 20.27 (Nicht Riemann-integrierbare Funktion mit Stammfunktion)

Zeige, die Funktion

$$F : [0,1] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} x^{\frac{3}{2}} \cdot \sin\left(\frac{1}{x}\right), & \text{wenn } x > 0, \\ 0, & \text{wenn } x = 0 \end{cases}$$

ist Stammfunktion einer nicht Riemann-integrierbaren Funktion.

Aufgabe 20.28

Sei $f : (-1,1) \longrightarrow \mathbb{R}$, $x \longmapsto \arctan(x)$.

a. Finde die Taylorreihenentwicklung von f mit Hilfe gliedweiser Integration.

b. Bestimme eine Reihendarstellung für $\frac{\pi}{4}$ und π.

Aufgabe 20.29

Es sei I ein Intervall und $f : I \to \mathbb{R}$ eine stetig differenzierbare Funktion ohne Nullstelle. Zeige, daß $\ln(|f|)$ eine Stammfunktion von $\frac{f'}{f}$ auf I ist.

Aufgabe 20.30

Berechne mittels Partialbruchzerlegung eine Stammfunktion von

$$x \mapsto \frac{x^3}{x^4 - 4x^3 + 8x^2 - 16x + 16}.$$

Aufgabe 20.31

Berechne die folgenden bestimmten und unbestimmten Integrale:

a. $\int_{1\frac{1}{4}}^{0} \sqrt{8x+2}\, dx$.

b. $\int_{1}^{2} \frac{x^4+2x}{x^5+5x^2-2}\, dx$.

c. $\int x^2 - \sin(x) + e^{3x}\, dx$.

d. $\int \sin(x) \cdot e^x\, dx$.

e. $\int_{0}^{\frac{\pi}{2}} x^2 \cdot \cos(x)\, dx$.

f. $\int_{2}^{5} \frac{3x}{x^2+1}\, dx$.

g. $\int_{\frac{\pi}{4}}^{\frac{\pi}{2}} \ln\left(\sqrt{\sin(x)}\right) \cdot \cos(x)\, dx$.

h. $\int e^{\sqrt{x}}\, dx$.

Aufgabe 20.32

Ist $f : [0,1] \longrightarrow \mathbb{R}$ stetig, so gibt es ein $c \in (0,1)$ mit $\int_0^1 f(x) \cdot x^2 \, dx = \frac{f(c)}{3}$.

Aufgabe 20.33

Berechne die folgenden bestimmten und unbestimmten Integrale:

a. $\int_0^{2\pi} \cos(kx) \, dx$ für $k \in \mathbb{N}$.

b. $\int_0^{\frac{\pi}{4}} \sin(x)\cos(x) dx$.

c. $\int \frac{1}{\sin(x)} \, dx$, substituiere $z = \tan\left(\frac{x}{2}\right)$.

Aufgabe 20.34

Berechne mittels Partialbruchzerlegung eine Stammfunktion von

$$x \mapsto \frac{2x^5 + 7x^4 + 9x^3 + 9x^2 + 6x - 5}{x^4 + 4x^3 + 5x^2 + 4x + 4}.$$

Aufgabe 20.35

Es sei $f : [a,b] \longrightarrow \mathbb{R}$ eine stetig differenzierbare Funktion und

$$g : \mathbb{R} \longrightarrow \mathbb{R} : y \mapsto \int_a^b f(x) \cdot \sin(y \cdot x) \, dx.$$

Zeige mit Hilfe partieller Integration $\lim\limits_{y \to \infty} g(y) = 0$ und $\lim\limits_{y \to -\infty} g(y) = 0$.

§ 21 Uneigentliche Integrale

A) 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 $f : [a, b) \longrightarrow \mathbb{R}$ sei auf $[a, y]$ integrierbar für alle $y \in (a, b)$. Falls der Grenzwert

$$\int_a^b f(x)\,dx := \lim_{y \to b} \int_a^y f(x)\,dx \in \mathbb{R} \cup \{\infty, -\infty\}$$

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 $f : (a, b] \longrightarrow \mathbb{R}$ sei auf $[y, b]$ integrierbar für alle $y \in (a, b)$. Falls der Grenzwert

$$\int_a^b f(x)\,dx := \lim_{y \to a} \int_y^b f(x)\,dx \in \mathbb{R} \cup \{\infty, -\infty\}$$

existiert, so nennen wir ihn ein *uneigentliches Integral*. Ist $\int_a^b f(x)\,dx \in \mathbb{R}$, so nennen wir das uneigentliche Intgral *konvergent*.

c. Es seien $a \in \mathbb{R} \cup \{-\infty\}$ und $b \in \mathbb{R} \cup \{\infty\}$ mit $a < b$ und für $f : (a, b) \longrightarrow \mathbb{R}$ gebe es ein $c \in (a, b)$, so daß die uneigentlichen Integrale $\int_c^b f(x)\,dx$ und $\int_a^c f(x)\,dx$ konvergent sind, dann nennen wir auch

$$\int_a^b f(x)\,dx := \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

ein *konvergentes 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:

$$\int_0^\infty \exp(-x)\,dx = \lim_{y \to \infty} \int_0^y \exp(-x) = \lim_{y \to \infty} -\exp(-x)\Big|_0^y = \lim_{y \to \infty} 1 - \exp(-y) = 1.$$

b. Es sei $a \in \mathbb{R}$ mit $a \neq 1$. Dann gilt für das uneigentliche Integral

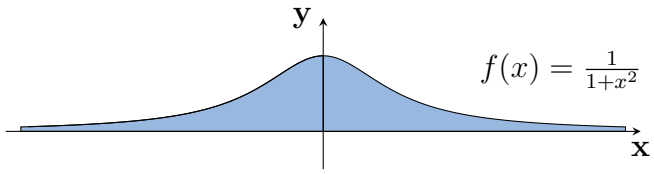
$$\begin{aligned} \int_1^\infty \frac{1}{x^a}\,dx &= \lim_{y \to \infty} \int_1^y x^{-a}\,dx = \lim_{y \to \infty} \frac{x^{-a+1}}{-a+1}\Big|_1^y \\ &= \lim_{y \to \infty} \frac{y^{1-a}}{1-a} - \frac{1}{1-a} = \begin{cases} \infty, & \text{falls } a < 1, \\ \frac{1}{a-1}, & \text{falls } a > 1. \end{cases} \end{aligned}$$

c. Das folgende uneigentliche Integral ist konvergent mit Grenzwert 2:

$$\int_0^1 \frac{1}{\sqrt{1-x}}\,dx = \lim_{y\to 1} -2\cdot\sqrt{1-x}\Big|_0^y = 2 - \lim_{y\to 1} 2\cdot\sqrt{1-y} = 2.$$

d. Der Flächeninhalt unter dem Graphen von $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \frac{1}{1+x^2}$ ist π:

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{1}{1+x^2}\,dx &= \int_0^{\infty} \frac{1}{1+x^2}\,dx + \int_{-\infty}^{0} \frac{1}{1+x^2}\,dx \\ &= \lim_{y\to\infty} \arctan(x)\Big|_0^y + \lim_{y\to-\infty} \arctan(x)\Big|_y^0 \\ &= \lim_{y\to\infty} \arctan(y) - \lim_{y\to-\infty} \arctan(y) = \frac{\pi}{2} + \frac{\pi}{2} = \pi.\end{aligned}$$



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.

B) Integralkriterium für Reihen

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 $f : [a, b) \longrightarrow \mathbb{R}_{\geq 0}$ 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

$$F : [a, b) \longrightarrow \mathbb{R} : y \mapsto \int_a^y f(x)\,dx$$

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 $\big(F(a_n)\big)_{n\in\mathbb{N}}$ von Funktionswerten. Diese ist monoton wachsend und beschränkt durch s, mithin ist sie konvergent, d.h. es gibt ein $I \in \mathbb{R}$ mit

$$\int_a^{a_n} f(x)\,dx = F(a_n) \longrightarrow I.$$

Zu gegebenem $\varepsilon > 0$ gibt es also ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $n \geq n_\varepsilon$ gilt

$$I - F(a_n) = |F(a_n) - I| < \varepsilon.$$

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

   $$\int_a^\infty f(x)\,dx = \lim_{y\to\infty} F(y) = I.$$

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

   $$\int_a^\infty f(x)\,dx = \lim_{y\to\infty} F(y) = I.$$

□

Satz 21.5 (Integralkriterium für Reihen)
Es sei $a \in \mathbb{N}$ und $f : [a, \infty) \longrightarrow \mathbb{R}_{\geq 0}$ sei monoton fallend und auf $[a, c]$ integrierbar für alle $c \in [a, \infty)$. Dann gilt:

$$\int_a^\infty f(x)\,dx \text{ ist konvergent} \quad \Longleftrightarrow \quad \sum_{n=a}^{\infty} f(n) \text{ ist konvergent.}$$

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

$$Z_m := (a, a+1, a+2, \ldots, m)$$

des Intervalls $[a, m]$. Da f auf $[a, m]$ monoton fallend ist, erhalten wir

$$\sum_{n=a+1}^{m} f(n) = \mathrm{US}(f, Z_m) \leq \int_a^m f(x)\,dx \leq \mathrm{OS}(f, Z_m) = \sum_{n=a}^{m-1} f(n). \tag{59}$$

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 (59) 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

$$s_m := \sum_{n=a}^{m} f(n) \overset{(59)}{\leq} f(a) + \int_a^m f(x)\,dx \leq f(a) + \int_a^\infty f(x)\,dx$$

monoton wachsend und beschränkt, mithin konvergent.

Die Abschätzung für die Grenzwerte der Reihen und des Integrals folgt unmittelbar aus (59), indem man m gegen unendlich gehen läßt. □

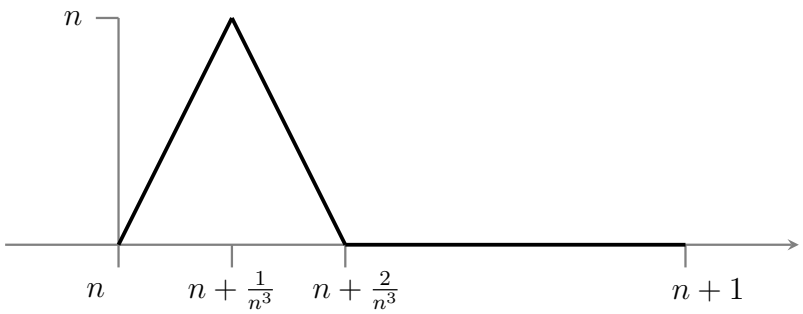
Beispiel 21.6

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* $\lim\limits_{x\to\infty} f(x) = 0$ gelten!

Wir betrachten eine Funktion $f : [1, \infty) \longrightarrow \mathbb{R}$, 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

$$\int_1^\infty f(x)\,dx = \lim_{n\to\infty} \int_1^{n+1} f(x)\,dx = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k^2} = \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty$$

konvergent, aber $\lim\limits_{x\to\infty} f(x) \neq 0$, da

$$f\left(n + \frac{1}{n^3}\right) = n \longrightarrow \infty.$$

Aufgaben

Aufgabe 21.8
Berechne den Wert des uneigentlichen Integrals $\int_{-\infty}^{0} \frac{e^x}{1+e^{2x}}\,dx$.

Aufgabe 21.9
Untersuche, für welche $t \in \mathbb{R}$ das uneigentliche Integral $\int_0^\infty x \cdot e^{-tx}\,dx$ konvergiert und bestimme für diese den Wert des Integrals.

Aufgabe 21.10

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 $\Gamma : (0,\infty) \longrightarrow \mathbb{R} : y \mapsto \int_0^\infty x^{y-1} \cdot \exp(-x)\,dx$ erfüllt die Funktionalgleichung
$$\Gamma(y+1) = y \cdot \Gamma(y)$$
für $y \in (0,\infty)$.

c. Für $n \in \mathbb{N}$ gilt $\Gamma(n+1) = n!$.

Aufgabe 21.11
Zeige, $\int_0^1 x^t \cdot \ln(x)^n\,dx = \frac{(-1)^n \cdot n!}{(1+t)^{n+1}}$ für alle $t > -1$ und für alle $n \in \mathbb{N}$.

Aufgabe 21.12 (Cauchy-Kriterium für uneigentliche Integrale)
Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a,b) \longrightarrow \mathbb{R}$ sei auf $[a,y]$ integrierbar für alle $y \in (a,b)$. Zeige, genau dann existiert das uneigentliche Integral $\int_a^b f(x)\,dx$, wenn
$$\forall\, \varepsilon > 0\ \exists\, c_\varepsilon \in [a,b)\ :\ \forall\, c_\varepsilon < s < t < b \ \text{ gilt } \left| \int_s^t f(x)\,dx \right| < \varepsilon.$$

Aufgabe 21.13 (Uneigentliche Integrale beschränkter Funktionen)
Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a,b) \longrightarrow \mathbb{R}$ sei eine beschränkte Funktion, die für alle $y \in (a,b)$ auf $[a,y)$ integrierbar ist. Zeige, dann ist das uneigentliche Integral $\int_a^b f(x)\,dx$ konvergent.

Aufgabe 21.14

a. Zeige, für jedes $y \in (0,\infty)$ ist die Funktion

$$g : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x} - \frac{1}{x+y}$$

streng monoton fallend.

b. Zeige mit Hilfe des Integralkriteriums für Reihen, daß die Reihe

$$\sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{k+y} \right)$$

für jedes $y \in (0,\infty)$ konvergiert.

c. Zeige, die Folge $(f_n)_{n\in\mathbb{N}}$ mit

$$f_n : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x^2} + \sum_{k=1}^{n} \frac{1}{(x+k)^2}$$

konvergiert auf jedem Intervall $[\delta,\infty)$ mit $\delta > 0$ gleichmäßig gegen

$$f : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \frac{1}{x^2} + \sum_{k=1}^{\infty} \frac{1}{(x+k)^2}.$$

d. Zeige, die Funktion

$$F : (0,\infty) \longrightarrow \mathbb{R} : y \mapsto -\frac{1}{y} + \sum_{k=1}^{\infty} \left(\frac{1}{k} - \frac{1}{k+y} \right)$$

ist eine Stammfunktion von f auf $(0,\infty)$ und erfüllt für alle $y \in (0,\infty)$ die Bedingungen $F'(y) > 0$ und $F(y+1) - F(y) = \frac{1}{y}$.

e. Zeige, ist $G : (0,\infty) \longrightarrow \mathbb{R}$ eine differenzierbare Funktion mit $G'(y) \geq 0$ und $G(y+1) - G(y) = \frac{1}{y}$ für alle $y \in (0,\infty)$, so unterscheiden sich F und G nur um eine Konstante.

f. Zeige, die differenzierbare Funktion

$$H : (0,\infty) \longrightarrow \mathbb{R} : x \mapsto \big(\ln\big(\Gamma(x)\big)\big)' = \frac{\Gamma'(x)}{\Gamma(x)}$$

erfüllt die Bedingungen $H'(y) \geq 0$ und $H(y+1) - H(y) = \frac{1}{y}$ für alle $y \in (0,\infty)$. Man nennt H auch die *Digammafunktion*.

Hinweis zu Teil e.: setze $F_n(y) = -\frac{1}{y} + \sum_{k=1}^{n} \left(\frac{1}{k} - \frac{1}{k+y} \right)$ und zeige $G(y) - F_n(y) = G(y+n+1) - \sum_{k=1}^{n} \frac{1}{k}$; nutze dies, um $(G-F)'(y) \geq 0$ für alle y zu zeigen; zeige ferner, daß $G - F$ 1-periodisch ist und folgere dann, daß $G - F$ konstant ist.

Teil B

Analysis 2

Kapitel III

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 oder noch allgemeiner auf metrischen Räumen einführen, und uns immer dann auf den endlich-dimensionalen Fall zurückziehen, wenn dies geboten scheint.

In der eindimensionalen Analysis haben wir uns zunächst mit der Konvergenz von Folgen beschäftigt und darauf aufbauend Eigenschaften von Funktionen wie die Stetigkeit, die Differenzierbarkeit und die Integrierbarkeit und untersucht. Alle diese Begriffe wollen wir nun für Abbildungen einführen, die auf geeigneten Teilmengen des $\mathbb{R}^n$ definiert sind, und eine erste Hürde wird darin bestehen, festzulegen, was im jeweiligen Kontext eine zielführende Verallgemeinerung eines offenen oder abgeschlossenen Intevalls ist. Danach wird sich manches ohne wesentliche Änderung verallgemeinern lassen und für manches wird man neue Ideen und Ansätze brauchen. Einige zentrale Sätze der eindimensionalen Analysis wie der Mittelwertsatz der Differentialrechnung oder der Umkehrsatz für stetige streng monotone Funktionen werden in neuem Gewand auftauchen und wichtige Ergebnisse im zweiten Teil der Vorlesung sein.

§ 22 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 wurde das schon in der Linearen Algebra im Zusammenhang mit den euklidischen Räumen getan (siehe Abschnitt A14), 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 22.1 (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 22.2 (Diskrete Metrik)
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 *diskrete Metrik* nennt.

Metriken wie die in Beispiel 22.2 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 22.3 (Normierte Räume)
Wir erinnern uns (siehe Definition A14.6), daß eine Abbildung

$$||\cdot|| : V \longrightarrow \mathbb{R}_{\geq 0},$$

eine *Norm* auf V heißt, falls für alle $x, y \in V$ und $\lambda \in \mathbb{R}$ folgendes gilt:

(1) $||x|| = 0 \quad \Leftrightarrow \quad x = 0,$ ("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 22.4 (Normierte Räume sind metrische Räume.)
Ist $(V, \cdot, +, ||\cdot||)$ ein normierter Raum und $M \subseteq V$, 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 22.4 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 22.5 (Normierte Räume)

a. Die *euklidische Norm* auf dem $\mathbb{R}^n$

$$||\cdot||_2 : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geq 0} : (x_1, \ldots, x_n)^t \mapsto \sqrt{x_1^2 + \ldots + x_n^2}$$

haben wir bereits in Satz A14.9 betrachtet.

b. Die *Summennorm* auf dem $\mathbb{R}^n$ wird definiert durch

$$||\cdot||_1 : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geq 0} : (x_1, \ldots, x_n)^t \mapsto |x_1| + \ldots + |x_n|.$$

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

$$||\cdot||_\infty : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geq 0} : (x_1, \ldots, x_n)^t \mapsto \max\{|x_1|, \ldots, |x_n|\}.$$

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 A14.9 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

$$||\cdot||_\infty : \mathcal{C}([0,1],\mathbb{R}) \longrightarrow \mathbb{R}_{\geq 0} : f \mapsto \max\big\{|f(x)| \;\big|\; x \in [0,1]\big\}.$$

Man beachte, daß das Maximum nach Satz 14.16 auch angenommen wird.

B) Metrische Räume sind haussdorffsch

Definition 22.6 (ε-Umgebung)
Sei M ein metrischer Raum, $a \in M$ und $\varepsilon > 0$. Wir nennen

$$U_\varepsilon(a) := \{x \in M \mid d(x,a) < \varepsilon\}$$

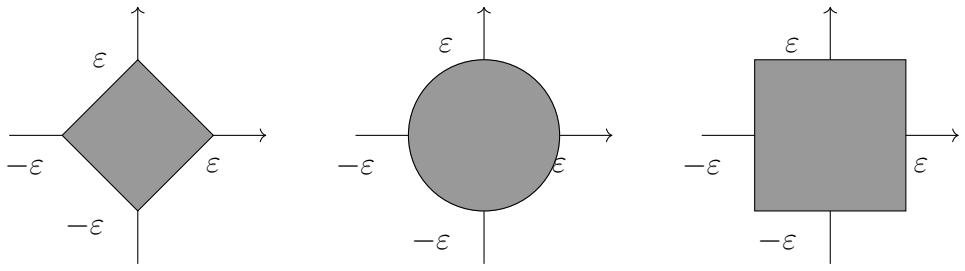
die *ε-Umgebung* von a in M.

Beispiel 22.7
Ist V ein normierter Raum und $a \in V$, so ist

$$U_\varepsilon(a) = \{x \in V \mid ||x-a|| < \varepsilon\}.$$

Für $V = \mathbb{R}$ stimmt dies mit der Definition in Bemerkung 13.2 überein.

In Abbildung 1 ist die ε-Umgebung von $a = (0,0)$ 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.



$U_\varepsilon(0)$ bez. $\|\cdot\|_1$ $U_\varepsilon(0)$ bez. $\|\cdot\|_2$ $U_\varepsilon(0)$ bez. $\|\cdot\|_\infty$

Abbildung 1: Die ε-Umgebung des Ursprungs bez. $||\cdot||_1$, $||\cdot||_2$ und $||\cdot||_\infty$

Proposition 22.8 (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 $\varepsilon = \frac{d(x,y)}{2} > 0$. Für ein $z \in U_\varepsilon(x) \cap U_\varepsilon(y)$ würde dann

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

gelten, was ein offensichtlicher Widerspruch ist. □

C) Offene und abgeschlossene Mengen

Definition 22.9 (Offene und abgeschlossene Mengen)
Es sei M ein metrischer Raum, $U, O, A \subseteq M$ und $a \in U$.

a. a heißt ein *innerer Punkt* von U, falls es ein $\varepsilon > 0$ gibt mit $U_\varepsilon(a) \subseteq U$.

b. O heißt *offen* in M, falls jeder Punkt von O ein innerer Punkt von O ist.

c. A heißt *abgeschlossen* in M, falls $M \setminus A$ offen in M ist.

Beispiel 22.10

a. Offene Intervalle (a,b) für $a, b \in \mathbb{R} \cup \{-\infty, \infty\}$ sind offen in $\mathbb{R}$.

b. Abgeschlossene Intervalle $[a,b]$ für $a, b \in \mathbb{R}$ sind abgeschlossen in $\mathbb{R}$.

c. Ein halboffenes Intervall $[a,b)$ mit $a, b \in \mathbb{R}$ ist weder offen noch abgeschlossen, da a kein innerer Punkt von $[a,b)$ und b kein innerer Punkt von $\mathbb{R} \setminus [a,b)$ ist. Analog ist auch $(a,b]$ weder offen noch abgeschlossen.

d. Intervalle der Form $(-\infty, a]$ und $[a, \infty)$ für $a \in \mathbb{R}$ 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 reellen Zahl sowohl rationale Zahlen, als auch irrationale Zahlen enthält.

Lemma 22.11 (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. □

Proposition 22.12 (Vereinigungen und Schnitte offener / abgeschlossener Mengen)
Sei M ein metrischer Raum.

a. Endliche Durchschnitte offener Mengen sind offen in M.

b. Beliebige Vereinigungen offener Mengen sind offen in M.

c. Endliche Vereinigungen abgeschlossener Mengen sind abgeschlossen in M.

d. 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 $A_1, \ldots, A_n$ 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 $A_1 \cup \ldots \cup A_n$ 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 22.13
Die Intervalle $U_n = \left(-\frac{1}{n}, \frac{1}{n}\right)$ 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 22.14 (Das Innere und der Abschluß)
Sei M ein metrischer Raum und $U \subseteq M$.

a. Der Durchschnitt $\overline{U}$ aller abgeschlossenen Teilmengen von M, die U enthalten, ist die kleinste abgeschlossene Teilmenge von M, die U enthält, und heißt der *Abschluß* von U.

b. Die Vereinigung $\mathring{U}$ 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 22.12 ist $\overline{U}$ als Durchschnitt abgeschlossener Mengen wieder abgeschlossen, und nach Definition ist $\overline{U}$ in jeder abgeschlossenen Teilmenge von M enthalten, die U enthält, so daß $\overline{U}$ die kleinste solche Menge ist. Die Aussage für $\mathring{U}$ sieht man analog. □

D) Häufungspunkte und Randpunkte

Definition 22.15 (Häufungspunkte und Randpunkte)
Sei M ein metrischer Raum, $U \subseteq M$ und $a \in M$.

a. 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.

$$\forall\, \varepsilon > 0 \;:\; U_\varepsilon(a) \cap \big(U \setminus \{a\}\big) \neq \emptyset.$$

$\mathrm{HP}(U)$ bezeichnet die Menge aller Häufungspunkte von U.

b. 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 22.16 (Der Rand einer ε-Umgebung)

a. Ist V ein *normierter* Raum, $a \in V$ und $\varepsilon > 0$, so ist

$$\partial U_\varepsilon(a) = \{x \in V \mid d(x,a) = \varepsilon\}$$

der Rand der ε-Umgebung von a. In Abbildung 1 ist der Rand der ε-Umgebung also genau der schwarze Rand der grauen Fläche.
Um dies zu sehen, setzen wir $A = \{x \in V \mid d(x,a) = \varepsilon\}$, wählen ein $b \in A$ und ein $\delta > 0$, wobei wir ohne Einschränkung $\delta < \varepsilon$ annehmen können. Da b nicht in $U_\varepsilon(a)$ liegt, gilt zunächst

$$b \in U_\delta(b) \cap \big(V \setminus U_\varepsilon(a)\big),$$

so daß dieser Durchschnitt nicht-leer ist. Setzen wir nun noch

$$x := b + \frac{\delta}{2 \cdot \varepsilon} \cdot (a - b),$$

so gilt

$$d(x,b) = \left|\left| b + \frac{\delta}{2 \cdot \varepsilon} \cdot (a-b) - b \right|\right| = \frac{\delta}{2 \cdot \varepsilon} \cdot ||a - b|| = \frac{\delta}{2 \cdot \varepsilon} \cdot \varepsilon < \delta$$

und

$$d(x,a) = \left|\left| b + \frac{\delta}{2 \cdot \varepsilon} \cdot (a-b) - a \right|\right| = \left| 1 - \frac{\delta}{2 \cdot \varepsilon} \right| \cdot ||b - a|| = \left| 1 - \frac{\delta}{2 \cdot \varepsilon} \right| \cdot \varepsilon < \varepsilon.$$

Damit liegt x also im Durchschnitt von $U_\delta(b)$ und $U_\varepsilon(a)$, und dieser Durchschnitt ist nicht leer. Das zeigt, daß b im Rand von $U_\varepsilon(a)$ liegt, d.h. $A \subseteq \partial U_\varepsilon(a)$.
Umgekehrt liegt kein Punkt $b \in V$ mit $d(b,a) > \varepsilon$ im Rand von $U_\varepsilon(a)$, da die δ-Umgebung von b für $\delta = \frac{d(b,a)-\varepsilon}{2}$ keinen Punkt mit $U_\varepsilon(a)$ gemeinsam hat. Für einen Punkt $x \in U_\delta(b)$ gilt nämlich

$$d(a,b) \leq d(a,x) + d(x,b) < d(a,x) + \delta = d(a,x) + \frac{d(b,a) - \varepsilon}{2}$$

und somit

$$\varepsilon < \frac{d(b,a)}{2} + \frac{\varepsilon}{2} < d(a,x).$$

Analog sieht man auch, daß kein Punkt $b \in V$ mit $d(b,a) < \varepsilon$ im Rand von $U_\varepsilon(a)$ liegen kann. Dies zeigt, schließlich $\partial U_\varepsilon(a) \subseteq A$.

b. Im Beweis von Teil a. haben wir verwendet, daß wir in V Punkte addieren können, d.h. wir haben die Vektorraumstruktur von V ausgenutzt, um zu zeigen, daß $U_\varepsilon(a) \cap U_\delta(b)$ nicht leer ist. In einem beliebigen *metrischen* Raum muß die Aussage *nicht* mehr gelten.
Betrachten wir etwa die reelle Ebene mit der diskreten Metrik aus Beispiel 22.2, dann enthält die 1-Umgebung

$$U_1(p) = \{p\}$$

eines Punkte p nur den Punkt p selbst; die Menge der Punkte mit Abstand 1 ist

$$\{q \in \mathbb{R}^2 \mid d(q,p) = 1\} = \mathbb{R}^2 \setminus \{p\};$$

der Rand der Umgebung ist aber leer, da z.B. die $\frac{1}{2}$-Umgebung $U_{\frac{1}{2}}(q) = \{q\}$ eines Punktes q entweder mit $\{p\}$ oder mit $\mathbb{R}^2 \setminus \{p\}$ einen leeren Schnitt hat. Also gilt die Aussage in Teil a. hier nicht.

Proposition 22.17 (Das Innere und der Abschluß von U)
Sei M ein metrischer Raum und $U \subseteq M$.

a. $\mathring{U}$ ist die Menge aller inneren Punkte von U.

b. $\overline{U} = U \cup \partial U$.

c. $\overline{U} = U \cup \mathrm{HP}(U)$.

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 $\mathring{U}$ von U enthalten.
Andererseits ist $\mathring{U}$ offen, so daß jeder Punkt in $\mathring{U}$ ein innerer Punkt von $\mathring{U}$ und somit auch ein innerer Punkt der größeren Menge U ist. Deshalb ist $\mathring{U}$ 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 $\overline{U}$ von M, d.h. $\overline{U} \subseteq A$.
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 $\overline{U}$ aller solcher Mengen auch die Menge A, d.h. $A \subseteq \overline{U}$.

c. Für einen Punkt $a \in M \setminus U$ gilt:

$$a \in \mathrm{HP}(U) \iff \forall\, \varepsilon > 0 \,:\, U_\varepsilon(a) \cap U \neq \emptyset \iff a \in \partial U.$$

Mithin gilt $\mathrm{HP}(U) \setminus U = \partial U \setminus U$, und somit folgt aus b.

$$\overline{U} = U \cup \big(\partial U \setminus U\big) = U \cup \big(\mathrm{HP}(U) \setminus U\big) = U \cup \mathrm{HP}(U).$$

□

Beispiel 22.18 (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 22.16 und Proposition 22.17. 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 U, 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 $\mathrm{HP}(U) = \partial U$ gilt, sondern lediglich

$$\mathrm{HP}(U) \setminus U = \partial U \setminus U.$$

Der Abschluß der Menge ist $\overline{U} = [0,1] \cup \{2\}$.

E) Kompakte Mengen in metrischen Räumen

Definition 22.19 (Beschränktheit und Kompaktheit)
Sei M ein metrischer Raum und $U \subseteq M$.

a. U heißt *beschränkt*, wenn es ein $r > 0$ und ein $a \in M$ gibt mit $d(x,a) \leq r$ für alle $x \in U$.

b. Eine Familie $(U_i)_{i\in I}$ offener Mengen in M heißt eine *offene Überdeckung* von U, falls $U \subseteq \bigcup_{i\in I} U_i$.

c. U heißt *kompakt*, falls jede offene Überdeckung $(U_i)_{i\in I}$ von U eine endliche Teilüberdeckung $(U_{i_1}, \ldots, U_{i_n})$, mit $i_1, \ldots, i_n \in I$, besitzt.

Beispiel 22.20

a. Ein Intervall in $\mathbb{R}$ mit endlichen Intervallgrenzen ist beschränkt.

b. In jedem metrischen Raum sind ε-Umgebungen beschränkt.

c. Endliche Teilmengen eines metrischen Raumes sind offenbar kompakt.

d. Das offene Intervall $(0,1)$ ist nicht kompakt, da die offene Überdeckung $(0,1) = \bigcup_{n\geq 2} \left(\frac{1}{n}, 1\right)$ keine endliche Teilüberdeckung besitzt.

Satz 22.21 (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 $x_1, \ldots, x_n \in A$ mit

$$A \subseteq \bigcup_{i=1}^{n} U_1(x_i).$$

Wir wählen $a \in A$ beliebig, aber fest und setzen

$$r := \max\{d(x_1, a), \ldots, d(x_n, a)\} + 1 > 0.$$

Für ein beliebiges $x \in A$ gibt es nun ein $i \in \{1, \ldots, n\}$, 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 22.8 gibt es zu jedem $x \in A$ ein $\varepsilon_x > 0$ mit

$$U_{\varepsilon_x}(a) \cap U_{\varepsilon_x}(x) = \emptyset. \tag{60}$$

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 $x_1, \ldots, x_n \in A$ mit

$$A \subseteq \bigcup_{i=1}^{n} U_{\varepsilon_{x_i}}(x_i).$$

Wir setzen nun

$$\varepsilon := \min\{\varepsilon_{x_1}, \ldots, \varepsilon_{x_n}\} > 0$$

und erhalten somit aus (60)

$$U_\varepsilon(a) \cap U_{\varepsilon_{x_i}}(x_i) = \emptyset$$

für alle $i = 1, \ldots, n$ 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. □

F) 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. Ersetzen wir den Betragsabstand durch die Metrik d, so verallgemeinern sich die Definitionen zu Folgen aus der eindimensionalen Analysis ohne Änderungen auf metrische Räume.

Definition 22.22 (Folgen in metrischen Räumen)
Sei M ein metrischer Raum.

a. Eine *Folge* in M ist eine Abbildung $\alpha : \mathbb{N} \longrightarrow M$. Ist $a_n := \alpha(n)$ für $n \in \mathbb{N}$, so schreiben wir meist $(a_n)_{n\in\mathbb{N}}$ statt α.

b. Wir nennen $a \in M$ einen *Grenzwert* der Folge $(a_n)_{n\in\mathbb{N}}$ in M, wenn

$$\forall\, 0 < \varepsilon \in \mathbb{R} \;\; \exists\, n_\varepsilon \in \mathbb{N} \;:\; \forall\, n \geq n_\varepsilon \;:\; d(a_n, a) < \varepsilon.$$

In diesem Fall sagen wir auch, daß $(a_n)_{n\in\mathbb{N}}$ *gegen* a *konvergiert* und schreiben

$$\lim_{n\to\infty} a_n = a$$

oder

$$a_n \longrightarrow a.$$

c. Wir nennen eine Folge $(a_n)_{n\in\mathbb{N}}$ in M genau dann *konvergent*, wenn $(a_n)_{n\in\mathbb{N}}$ einen Grenzwert $a \in M$ besitzt. Ist $(a_n)_{n\in\mathbb{N}}$ nicht konvergent, so nennen wir $(a_n)_{n\in\mathbb{N}}$ *divergent*. Ist M ein normierter Raum und ist der Grenzwert 0, so nennen wir $(a_n)_{n\in\mathbb{N}}$ eine *Nullfolge*.

d. Eine Folge $(a_n)_{n\in\mathbb{N}}$ in M heißt *Cauchy-Folge*, falls

$$\forall\, 0 < \varepsilon \in \mathbb{R} \;\; \exists\, n_\varepsilon \in \mathbb{N} \;:\; \forall\, m > n \geq n_\varepsilon \;:\; d(a_m, a_n) < \varepsilon.$$

e. Ist $(a_n)_{n\in\mathbb{N}}$ eine Folge in M und ist zudem

$$n_0 < n_1 < n_2 < n_3 < \ldots$$

eine aufsteigende Folge natürlicher Zahlen, so nennen wir die Folge

$$(a_{n_k})_{k\in\mathbb{N}} = (a_{n_0}, a_{n_1}, a_{n_2}, a_{n_3}, \ldots)$$

eine *Teilfolge* von $(a_n)_{n\in\mathbb{N}}$.

f. Eine Folge $(a_n)_{n\in\mathbb{N}}$ in M heißt *beschränkt*, wenn $\{a_n \mid n \in \mathbb{N}\}$ beschränkt ist, d.h. wenn es ein $r > 0$ und ein $a \in M$ gibt mit $d(a_n, a) \leq r$ für alle $n \in \mathbb{N}$.

Ebenso wie die Definitionen verallgemeinern sich auch die elementaren Eigenschaften zur Konvergenz von Folgen aus der eindimensionalen Analysis auf metrische Räume.

Proposition 22.23 (Einfache Eigenschaften von Folgen)
Sei M ein metrischer Raum, $\lambda \in \mathbb{R}$ und $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ Folgen in M.

a. Genau dann gilt $a_n \longrightarrow a$, wenn $d(a_n, a) \longrightarrow 0$.

b. Der Grenzwert einer konvergenten Folge in M ist eindeutig bestimmt.

c. Jede konvergente Folge in M ist eine Cauchy-Folge.

d. Cauchy-Folgen und konvergente Folgen in M sind beschränkt.

e. Falls $a_n \to a$ gilt, so konvergiert jede Teilfolge von $(a_n)_{n\in\mathbb{N}}$ gegen a.

f. Ist M ein normierter Raum und gelten $a_n \to a$ und $b_n \to b$, dann gelten

$$a_n \pm b_n \to a \pm b \quad \text{und} \quad \lambda \cdot a_n \to \lambda \cdot a.$$

g. Wenn $a_n \to a$ und $b_n \to b$ gilt, dann gilt auch $d(a_n, b_n) \to d(a, b)$.

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.

a. Dies folgt unmittelbar aus der Definition der Konvergenz in M und in $\mathbb{R}$.

b. Nehmen wir an, eine Folge $(a_n)_{n\in\mathbb{N}}$ in M besitze zwei verschiedene Grenzwerte $a, b \in M$. Dann ist die reelle Zahl

$$\varepsilon := \frac{d(a,b)}{2} > 0$$

positiv. Mithin gibt es zwei natürliche Zahlen $n_\varepsilon, n'_\varepsilon \in \mathbb{N}$, so daß

$$d(a_n, a) < \varepsilon$$

für $n \geq n_\varepsilon$ und

$$d(a_n, b) < \varepsilon$$

für $n \geq n'_\varepsilon$. Setzen wir nun $N := \max\{n_\varepsilon, n'_\varepsilon\}$, so gilt

$$d(a,b) \leq d(a, a_N) + d(a_N, b) < \varepsilon + \varepsilon = d(a,b),$$

was ein offensichtlicher Widerspruch ist.

c. 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

$$d(a_n, a) < \frac{\varepsilon}{2}$$

für alle $n \geq n_\varepsilon$. Für zwei natürliche Zahlen $m > n \geq n_\varepsilon$ folgt dann

$$d(a_m, a_n) \leq d(a_m, a) + d(a, a_n) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Also ist $(a_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge.

d. Wegen c. reicht es, die Aussage für Cauchy-Folgen $(a_n)_{n\in\mathbb{N}}$ zu zeigen. Zu $\varepsilon = 1$ gibt es eine natürliche Zahl $n_\varepsilon \in \mathbb{N}$, so daß

$$d(a_m, a_n) < \varepsilon = 1$$

für alle $m > n \geq n_\varepsilon$. Wähle $a \in M$ fest und setze

$$r := \max\{1 + d(a_{n_\varepsilon}, a), d(a_0, a), d(a_1, a), \ldots, d(a_{n_\varepsilon - 1}, a)\}.$$

Damit erhalten wir dann

$$d(a_n, a) \leq \begin{cases} r, & \text{falls } n < n_\varepsilon, \\ d(a_n, a_{n_\varepsilon}) + d(a_{n_\varepsilon}, a) \leq 1 + d(a_{n_\varepsilon}, a) \leq r, & \text{falls } n \geq n_\varepsilon. \end{cases}$$

Mithin ist $(a_n)_{n\in\mathbb{N}}$ 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

$$||a_n - a|| = d(a_n, a) < \frac{\varepsilon}{2}$$

für $n \geq n'_\varepsilon$ und

$$||b_n - b|| = d(b_n, b) < \frac{\varepsilon}{2}$$

für $n \geq n''_\varepsilon$. Mit $n_\varepsilon := \max\{n'_\varepsilon, n''_\varepsilon\}$ gilt dann

$$d(a_n + b_n, a + b) = ||a_n - a + b_n - b|| \leq ||a_n - a|| + ||b_n - b|| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

für $n \geq n_\varepsilon$. Mithin konvergiert $(a_n + b_n)_{n\in\mathbb{N}}$ gegen $a + b$. Analog sieht man $a_n - b_n \longrightarrow a - b$ und $\lambda \cdot a_n \longrightarrow \lambda \cdot a$.

g. Aus der Dreiecksungleichung leiten wir

$$d(a_n, b_n) \leq d(a_n, a) + d(a, b) + d(b, b_n) = d(a_n, a) + d(b_n, b) + d(a, b)$$

sowie

$$d(a, b) \leq d(a, a_n) + d(a_n, b_n) + d(b_n, b) = d(a_n, a) + d(b_n, b) + d(a_n, b_n)$$

her. Daraus folgt dann aber unmittelbar

$$0 \leq |d(a_n, b_n) - d(a, b)| \leq d(a_n, a) + d(b_n, b) \longrightarrow 0 + 0 = 0.$$

Also konvergiert die reelle Zahlenfolge $\big(d(a_n, b_n)\big)_{n\in\mathbb{N}}$ gegen $d(a, b)$.

□

Beispiel 22.24

Die Folge $(a_n)_{n\in\mathbb{N}}$ mit

$$a_n = \left(\frac{\cos(n)}{2^n}, \frac{\sin(n)}{2^n}\right)^t$$

konvergiert in $\mathbb{R}^2$ bezüglich der euklidischen Norm gegen $a = (0,0)^t$, da

$$d(a_n, a) = \sqrt{\frac{\cos^2(n) + \sin^2(n)}{2^{2n}}} = \frac{1}{2^n} \longrightarrow 0.$$

Lemma 22.25 (Folgenkriterium für Häufungspunkte)
Sei M ein metrischer Raum, $U \subseteq M$ und $a \in M$. Genau dann ist a ein Häufungspunkt von U, wenn es eine Folge in $U \setminus \{a\}$ gibt, die gegen 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 a ein Häufungspunkt von U, so gibt es für jedes $n \geq 1$ ein

$$a_n \in U_{\frac{1}{n}}(a) \cap \big(U \setminus \{a\}\big).$$

Dann gilt aber

$$0 \leq d(a_n, a) \leq \frac{1}{n} \longrightarrow 0,$$

und $d(a_n, a)$ ist eine Nullfolge. Also ist $(a_n)_{n\in\mathbb{N}}$ eine Folge in U, die gegen a konvergiert.

Ist umgekehrt $(a_n)_{n\in\mathbb{N}}$ eine Folge in $U \setminus \{a\}$, die gegen a konvergiert, und ist $\varepsilon > 0$, so gibt es ein $n_\varepsilon \in \mathbb{N}$ mit $d(a_n, a) < \varepsilon$ für alle $n \geq n_\varepsilon$. Insbesondere ist also $U_\varepsilon(a) \cap \big(U \setminus \{a\}\big)$ nicht leer. Das zeigt, daß a ein Häufungspunkt von U ist. □

Beispiel 22.26
Die Häufungspunkte in $U = (0,1] \cup \{2\}$ sind genau die Zahlen a in $[0,1]$, da nur diese Grenzwerte einer Folge in $U \setminus \{a\}$ sind. Die Randpunkte von U sind genau die Zahlen $0, 1, 2$.

Satz 22.27 (Der Abschluß ist die Menge der Grenzwerte.)
Ist M ein metrischer Raum und $U \subseteq M$, so gilt

$$\overline{U} = \{a \in M \mid \exists\, (a_n)_{n\in\mathbb{N}} \subseteq U \text{ mit } a_n \to a\},$$

d.h. der Abschluß von U ist die Menge der Grenzwerte von Folgen in U.

Beweis: Sei zunächst a der Grenzwert einer Folge $(a_n)_{n\in\mathbb{N}}$ in U. Ist $a \in U$, so ist $a \in \overline{U}$. Ist $a \notin U$, so ist a nach Lemma 22.25 ein Häufungspunkt von U und somit gilt $a \in \mathrm{HP}(U) \subseteq \overline{U}$ nach Proposition 22.17.

Sei nun umgekehrt $a \in \overline{U} = U \cup \mathrm{HP}(U)$. Ist $a \in U$, so ist a der Grenzwert der konstanten Folge $(a)_{n\in\mathbb{N}}$. Ist $a \notin U$, so muß a ein Häufungspunkt von U sein und ist dann nach Lemma 22.25 Grenzwert einer Folge in $U \setminus \{a\}$. □

Beispiel 22.28
Der Abschluß der Menge $U = (0,1] \cup \{2\}$ aus Beispiel 22.26 ist $[0,1] \cup \{2\}$, da dies genau die reellen Zahlen sind, die als Grenzwerte von Folgen in U auftreten.

Korollar 22.29 (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 22.52 ist U genau dann abgeschlossen, wenn $U = \overline{U}$ 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 22.27 in $\overline{U} = U$. Ist umgekehrt jeder Grenzwert einer in M konvergenten Folge aus U schon in U, so gilt mit Satz 22.27 $U = \overline{U}$ und U ist abgeschlossen. □

G) Der Satz von Bolzano-Weierstraß

Satz 22.30 (Bolzano-Weierstraß)
Sei M ein metrischer Raum und $A \subseteq M$. Folgende Aussagen sind gleichwertig:

a. A ist kompakt.

b. Jede Folge in A besitzt eine konvergente Teilfolge mit Grenzwert in A.

Beweis: [1]

a. $\Longrightarrow$ 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 22.25 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 $x_1, \ldots, x_n \in A$ mit

$$A \subseteq \bigcup_{i=1}^{n} U_{\varepsilon_{x_i}}(x_i).$$

Es folgt also

$$U = U \cap A = \bigcup_{i=1}^{n} U \cap U_{\varepsilon_{x_i}}(x_i) \subseteq \{x_1, \ldots, x_n\},$$

[1] Den Beweis von "b. ⇒ a." habe ich von der Webseite von Prof. Franz Merkl, München.

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 22.21 aber abgeschlossen ist, muß der Grenzwert der konvergenten Teilfolge wegen des Folgenkriteriums für Abgeschlossenheit 22.29 dann auch in A liegen.

b. $\Longrightarrow$ a.: Sei $(U_i)_{i\in I}$ eine offene Überdeckung von A. Wir müssen zeigen, daß es bereits eine endliche Teilüberdeckung gibt.

Wir wollen dazu zunächst folgende Aussage zeigen:

$$\exists\, N \geq 1 \;:\; \forall\, x \in A \;\; \exists\, i \in I \;\text{ mit }\; U_{\frac{1}{N}}(x) \subseteq U_i. \tag{61}$$

Dazu definieren wir für $n \geq 1$ die Menge

$$X_n = \left\{ x \in A \;\middle|\; \exists\, i \in I \;:\; U_{\frac{1}{n}}(x) \subseteq U_i \right\}.$$

Aus der Definition folgt unmittelbar

$$X_1 \subseteq X_2 \subseteq X_3 \subseteq \ldots,$$

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 (61) 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

$$d(a_n, a) \geq \tfrac{1}{2m}$$

für alle $n \geq 2m$ gelten, im Widerspruch dazu, daß eine Teilfolge von $(a_n)_{n\geq 1}$ gegen a konvergiert. Damit ist (61) 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, so daß

$$x_n \notin U_{i_0} \cup \ldots \cup U_{i_n}$$

gilt sowie für alle $m < n$ stets auch

$$U_{\frac{1}{N}}(x_m) \subseteq U_{i_0} \cup \ldots \cup U_{i_n}.$$

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 $i_0, \ldots, i_{n-1} \in I$ und $x_0, \ldots, x_{n-1} \in A$ schon konstruiert. Wegen (61) können wir ein $i_n \in I$ wählen mit

$$U_{\frac{1}{N}}(x_{n-1}) \subseteq U_{i_n} \subseteq U_{i_0} \cup \ldots \cup U_{i_n}$$

und da nach Annahme $U_{i_0}, \ldots, U_{i_n}$ die Menge A nicht überdecken, gibt es auch ein

$$x_n \in A \setminus \big(U_{i_0} \cup \ldots \cup U_{i_n}\big).$$

Seien nun natürliche Zahlen $n > m \geq 0$ gegeben. Dann gilt nach Konstruktion

$$x_n \notin U_{i_0} \cup \ldots \cup U_{i_n} \supseteq U_{\frac{1}{N}}(x_m)$$

und mithin

$$d(x_n, x_m) \geq \tfrac{1}{N}.$$

Dies zeigt, daß keine Teilfolge von $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge sein kann, und mithin kann nach Proposition 22.23 erst recht keine Teilfolge von $(x_n)_{n\in\mathbb{N}}$ konvergent sein, was im Widerspruch zur Voraussetzung steht. □

Beispiel 22.31 (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 22.29 dieser Grenzwert c dann schon in $[a, b]$ liegen muß, da $[a, b]$ abgeschlossen ist.

Korollar 22.32 (Abgeschloss. 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 22.30 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 22.30 kompakt. □

H) Äquivalente Normen

Definition 22.33 (Äquivalente Normen)
Zwei Normen $|\cdot| : V \longrightarrow \mathbb{R}_{\geq 0}$ und $||\cdot|| : V \longrightarrow \mathbb{R}_{\geq 0}$ 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 22.34 (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 22.35 (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 *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)

$$U_{s\cdot\varepsilon}^{||\cdot||}(a) \subseteq U_{\varepsilon}^{|\cdot|}(a)$$

und

$$U_{\frac{\varepsilon}{r}}^{|\cdot|}(a) \subseteq U_{\varepsilon}^{||\cdot||}(a),$$

wobei $U_{\delta}^{|\cdot|}(a)$ die δ-Umgebung von a bezüglich der Norm $|\cdot|$ bezeichnen soll und analog $U_{\delta}^{||\cdot||}(a)$ die δ-Umgebung von a bezüglich der Norm $||\cdot||$. D.h. jede kleine Umgebung

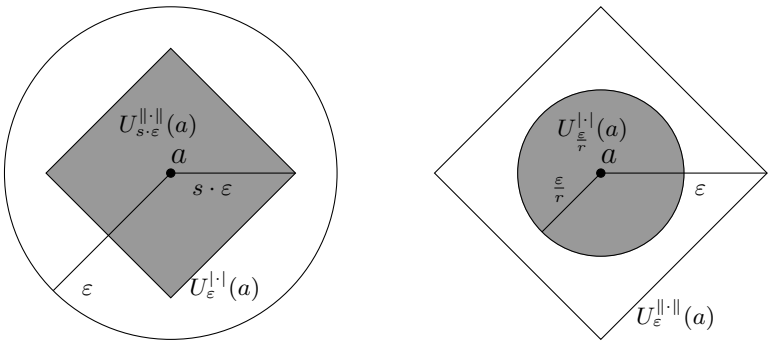


Abbildung 2. Umgebungen bei äquivalenten Normen

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 $U, A \subseteq V$, $a \in V$ und $(a_n)_{n\in\mathbb{N}}$ eine Folge in V:

Eigenschaften bezüglich $\lvert\cdot\rvert$		Eigenschaften bezüglich $\lVert\cdot\rVert$
a ist ein innerer Punkt von U.	$\Longleftrightarrow$	a ist ein innerer Punkt von U.
a ist ein Randpunkt von U.	$\Longleftrightarrow$	a ist ein Randpunkt von U.
a ist ein Häufungspunkt von U.	$\Longleftrightarrow$	a ist ein Häufungspunkt von U.
U ist beschränkt.	$\Longleftrightarrow$	U ist beschränkt.
U ist offen in V.	$\Longleftrightarrow$	U ist offen in V.
A ist abgeschlossen in V.	$\Longleftrightarrow$	A ist abgeschlossen in V.
A ist kompakt.	$\Longleftrightarrow$	A ist kompakt.
$(a_n)_{n\in\mathbb{N}}$ konvergiert gegen a.	$\Longleftrightarrow$	$(a_n)_{n\in\mathbb{N}}$ konvergiert gegen a.
$(a_n)_{n\in\mathbb{N}}$ ist eine Cauchy-Folge.	$\Longleftrightarrow$	$(a_n)_{n\in\mathbb{N}}$ ist eine Cauchy-Folge.

Bemerkung 22.36

Wir werden in Lemma 22.41 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 22.35 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 22.41 werden wir $\mathbb{R}^n$ mit der *Maximumsnorm*

$$||\cdot||_\infty : \mathbb{R}^n \longrightarrow \mathbb{R}_{\geq 0} : x = (x_1, \ldots, x_n)^t \mapsto ||x||_\infty = \max\{|x_1|, \ldots, |x_n|\}$$

als normierten Raum und damit auch als metrischen Raum betrachten.

I) Der Satz von Bolzano-Weierstraß im $\mathbb{R}^n$

Definition 22.37
Ist $(a_k)_{k\in\mathbb{N}}$ eine Folge in $\mathbb{R}^n$ mit $a_k = (a_{k1}, \ldots, a_{kn})^t$, so nennen wir die Folgen $(a_{ki})_{k\in\mathbb{N}}$ in $\mathbb{R}$ die *Komponentenfolgen* von $(a_k)_{k\in\mathbb{N}}$.

Proposition 22.38 (Komponentenweise Konvergenz)
Sei $(a_k)_{k\in\mathbb{N}}$ eine Folge in $\mathbb{R}^n$ und $c \in \mathbb{R}^n$.
Genau dann gilt $a_k \to c$, wenn $a_{ki} \to c_i$ für $i = 1, \ldots, n$.

Beweis: Aus $a_k \to c$ folgt

$$0 \leq |a_{ki} - c_i| \leq \max\{|a_{k1} - c_1|, \ldots, |a_{kn} - c_n|\} = ||a_k - c||_\infty = d(a_k, c) \longrightarrow 0$$

und somit $a_{ki} \to c_i$. Umgekehrt folgt aus $a_{ki} \to c_i$ für alle $i = 1, \ldots, n$ auch $|a_{ki} - c_i| \to 0$ für alle $i = 1, \ldots, n$ und somit

$$||a_k - c||_\infty = \max\left\{|a_{k1} - c_1|, \ldots, |a_{kn} - c_n|\right\} \longrightarrow 0$$

und somit $a_k \to c$. □

Beispiel 22.39
Die Folge $(a_n)_{n\geq 1}$ mit $a_n = \left(\frac{2n^3+1}{n^3+n}, \left(1 + \frac{1}{n}\right)^n\right)$ konvergiert in $\mathbb{R}^2$ gegen $(2, e)$.

Satz 22.40 (Bolzano-Weierstraß)
Jede beschränkte Folge im $\mathbb{R}^n$ besitzt eine konvergente Teilfolge.

Beweis: Es sei $(a_k)_{k\in\mathbb{N}}$ eine beschränkte Folge in $\mathbb{R}^n$, dann gibt es ein $r > 0$, so daß

$$|a_{ki}| \leq ||a_k||_\infty \leq r \tag{62}$$

für alle $k \in \mathbb{N}$ und $i = 1, \ldots, n$. Mithin ist die Komponentenfolge $(a_{k1})_{k\in\mathbb{N}}$ beschränkt und besitzt nach dem Satz von Bolzano-Weierstraß 11.26 in $\mathbb{R}$ eine konvergente Teilfolge $(a_{k_j 1})_{j\in\mathbb{N}}$. Die Komponentenfolge $(a_{k_j 2})_{j\in\mathbb{N}}$ ist wegen (62) aber ebenfalls beschränkt und besitzt eine konvergente Teilfolge $(a_{k_{j_l} 2})_{l\in\mathbb{N}}$. Die entsprechende Teilfolge in der ersten Komponente, $(a_{k_{j_l} 1})_{l\in\mathbb{N}}$ ist aber ebenfalls konvergent als Teilfolge einer konvergenten Folge. Fahren wir so fort, erhalten wir schließlich eine Teilfolge von $(a_k)_{k\in\mathbb{N}}$, so daß alle

Komponentenfolgen konvergieren. Wegen des Satzes zur komponentenweisen Konvergenz 22.38 konvergiert dann aber auch die Teilfolge selbst. □

J) Im $\mathbb{R}^n$ sind alle Normen äquivalent.

Lemma 22.41 (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

$$s \cdot ||x||_\infty \leq ||x|| \leq r \cdot ||x||_\infty$$

für alle $x \in \mathbb{R}^n$.

Beweis: Mit $(e_1, \ldots, e_n)$ bezeichen wir die kanonische Basis von $\mathbb{R}^n$, und wir setzen

$$r := ||e_1|| + \ldots + ||e_n|| > 0.$$

Für $x = x_1 \cdot e_1 + \ldots + x_n \cdot e_n \in \mathbb{R}^n$ gilt dann wegen der Dreiecksungleichung

$$||x|| = \left|\left| \sum_{i=1}^{n} x_i \cdot e_i \right|\right| \leq \sum_{i=1}^{n} |x_i| \cdot ||e_i|| \leq \sum_{i=1}^{n} ||x||_\infty \cdot ||e_i|| = r \cdot ||x||_\infty. \tag{63}$$

Nehmen wir nun an, es gebe kein $s > 0$ mit $s \cdot ||x||_\infty \leq ||x||$ für alle $x \in \mathbb{R}^n$. Dann gibt es für jedes $k > 0$ ein $x_k \in \mathbb{R}^n$ mit

$$\frac{1}{k} \cdot ||x_k||_\infty > ||x_k||. \tag{64}$$

Für $y_k = \frac{1}{||x_k||_\infty} \cdot x_k$ gilt dann einerseits

$$0 \leq ||y_k|| = \left|\left| \frac{x_k}{||x_k||_\infty} \right|\right| \overset{(64)}{<} \frac{1}{k} \longrightarrow 0 \tag{65}$$

und andererseits

$$||y_k||_\infty = \frac{||x_k||_\infty}{||x_k||_\infty} = 1.$$

Bezüglich der Maximumsnorm ist die Folge $(y_k)_{k \in \mathbb{N}}$ also beschränkt und besitzt nach dem Satz von Bolzano-Weierstraß 22.40 somit eine konvergente Teilfolge $(y_{k_j})_{j \in \mathbb{N}}$ mit Grenzwert a, d.h.

$$||y_{k_j} - a||_\infty \longrightarrow 0. \tag{66}$$

Für den Grenzwert gilt nun unter Berücksichtigung von (65) und (66)

$$0 \leq ||a|| \leq ||a - y_{k_j}|| + ||y_{k_j}|| \overset{(63)}{\leq} r \cdot ||a - y_{k_j}||_\infty + ||y_{k_j}|| \longrightarrow 0 + 0 = 0,$$

was $||a|| = 0$ bedingt. Dann ist aber $a = 0$ der Nullvektor, da die Norm $||\cdot||$ definit ist. Daraus ergibt sich dann der Widerspruch

$$1 = ||y_{k_j}||_\infty \leq ||y_{k_j} - a||_\infty + ||a||_\infty = ||y_{k_j} - a||_\infty \longrightarrow 0.$$

□

Beispiel 22.42
Für $x \in \mathbb{R}^n$ gilt offenbar $||x||_\infty \leq ||x||_2 \leq \sqrt{n} \cdot ||x||_\infty$.

Korollar 22.43
Je zwei Normen auf $\mathbb{R}^n$ sind äquivalent.

Beweis: Dies folgt aus Lemma 22.41, da die Äquivalenz von Normen nach Bemerkung 22.34 eine Äquivalenzrelation ist. □

Von jetzt an ist also egal, welche Norm wir auf $\mathbb{R}^n$ verwenden!

K) Der Satz von Heine-Borel

Satz 22.44 (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 22.21 abgeschlossen und beschränkt in $\mathbb{R}^n$.

Sei nun umgekehrt A abgeschlossen und beschränkt in $\mathbb{R}^n$ und sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in A. Nach dem Satz von Bolzano-Weierstraß 22.40 im $\mathbb{R}^n$ besitzt $(a_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge mit Grenzwert $a \in \mathbb{R}^n$. Da A abgeschlossen ist, liegt der Grenzwert a in A, wegen des Folgenkriteriums für Abgeschlossenheit 22.29. Dann ist die Menge A aber kompakt wegen des Satzes von Bolzano-Weierstraß 22.30 in metrischen Räumen. □

Beispiel 22.45

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

$$[a_1, b_1] \times \ldots \times [a_n, b_n] = \left\{(x_1, \ldots, x_n)^t \;\middle|\; a_i \leq x_i \leq b_i \;\forall\, i = 1, \ldots, n\right\}$$

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 22.46 (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 wenn A nicht-leer ist, so existieren $\sup(A)$ sowie $\inf(A)$. Wegen Bemerkung 11.22 gibt es Folgen in A, die gegen $\sup(A)$ bzw. $\inf(A)$ konvergieren, und da die kompakte Menge A auch abgeschlossen ist, liegen die Grenzwerte der Folgen deshalb in A. □

L) Vollständige metrische Räume

Definition 22.47 (Banachraum und Hilbertraum)

a. Ein metrischer Raum heißt *vollständig*, wenn jede Cauchy-Folge konvergiert.

b. Ein normierter Raum heißt *Banachraum*, wenn er als metrischer Raum vollständig ist.

c. Ein euklidischer Raum heißt *Hilbertraum*, wenn er mit der euklidischen Norm ein Banachraum ist.

Satz 22.48 (Cauchy-Kriterium für Folgenkonvergenz)
Jede Cauchy-Folge in $\mathbb{R}^n$ ist konvergent.

Beweis: Es sei $(a_k)_{k\in\mathbb{N}}$ 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 $k > l \geq n_\varepsilon$

$$\varepsilon > d(a_k, a_l) = ||a_k - a_l||_2 \geq |a_{ki} - a_{li}|$$

gilt, wobei $i \in \{1, \ldots, n\}$ beliebig ist. Daraus folgt, daß die i-te Komponentenfolge von $(a_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$ ist und mithin nach dem Cauchy-Kriterium 11.30 auch konvergent. Dann ist $(a_k)_{k\in\mathbb{N}}$ nach Proposition 22.38 aber konvergent. □

Korollar 22.49 ($\mathbb{R}^n$ ist ein Banachraum.)
$\mathbb{R}^n$ ist bezüglich jeder Norm ein Banachraum, und bezüglich der euklidischen Norm ist $\mathbb{R}^n$ ein Hilbertraum.

M) Reihen in Banachräumen

In Banachräumen kann man die Konvergenz von Reihen genauso definieren wie in $\mathbb{R}$ und die wesentlichen Ergebnisse übertragen sich unmittelbar. Wir werden uns hier auf einige wenige Aussagen beschränken, die wir im weiteren Verlauf benötigen.

Definition 22.50 (Reihen in Banachräumen)
Sei $(V, ||\cdot||)$ ein Banachraum und $(a_n)_{n\in\mathbb{N}}$ sei eine Folge in V.

a. Wir nennen die Folge $(s_n)_{n\in\mathbb{N}}$ mit
$$s_n := \sum_{k=0}^{n} a_k \in V$$
der *Partialsummen* von $(a_n)_{n\in\mathbb{N}}$ auch die durch $(a_n)_{n\in\mathbb{N}}$ 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
$$\sum_{n=0}^{\infty} a_n.$$

b. Die Reihe $\sum_{n=0}^{\infty} a_n$ heißt *absolut konvergent*, wenn die Reihe $\sum_{n=0}^{\infty} ||a_n||$ konvergiert.

Proposition 22.51 (Absolute Konvergenz impliziert Konvergenz)
Eine absolut konvergente Reihe $\sum_{n=0}^{\infty} a_n$ in einem Banachraum ist konvergent, und es gilt
$$\left|\left|\sum_{n=k}^{\infty} a_n\right|\right| \leq \sum_{n=k}^{\infty} ||a_n||$$
für alle $k \in \mathbb{N}$.

Beweis: Der Beweis geht wörtlich wie der Beweis von Lemma 12.17 in der eindimensionalen Analysis, wenn man den Betrag durch die Norm ersetzt.

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

$$\left|\left| \sum_{k=n+1}^{m} a_k \right|\right| \leq \sum_{k=n+1}^{m} ||a_k|| = \left| \sum_{k=n+1}^{m} ||a_k|| \right| < \varepsilon$$

für alle $m > n \geq n_\varepsilon$ gilt. Mithin ist die Folge der Partialsummen eine Cauchy-Folge und mithin im Banachraum V konvergent. □

Aufgaben

Aufgabe 22.52

Sei M ein metrischer Raum und $U \subseteq M$.

a. U ist genau dann offen in M, wenn $U = \mathring{U}$.

b. U ist genau dann abgeschlossen in M, wenn $U = \overline{U}$.

c. $\overline{\overline{U}} = \overline{U}$ und $\mathring{\mathring{U}} = \mathring{U}$.

d. $\overline{U} = \mathring{U} \cup \partial U$.

e. ∂U ist abgeschlossen in M.

Aufgabe 22.53 (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 22.54

Zeige, $||f||_\infty := \max\{|f(x) \mid x \in [0,1]\}$ definiert eine Norm auf $V = \mathcal{C}([0,1], \mathbb{R})$.

Aufgabe 22.55 (Normen auf Funktionenenräumen)

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 22.5.

b. V ist ein Banachraum bezüglich der Maximumsnorm aus Beispiel 22.5.

c. Der Differentialoperator $D : U \longrightarrow V : f \longmapsto f'$ ist ein linearer Operator, der *nicht* stetig ist, wenn U und V normiert sind bzgl. der Maximumsnorm.

Aufgabe 22.56

Sei $V = \mathcal{C}([0,1],\mathbb{R})$ und für $f \in V$ definieren wir

$$||f|| := \int_0^1 |f(x)|\, dx.$$

a. Zeige, daß $||\cdot||$ ein Norm auf V definiert.

b. Untersuche den Grenzwert der Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ in V bezüglich $||\cdot||$, wobei

$$f_n : [0,1] \longrightarrow \mathbb{R} : x \mapsto (x-1)^n.$$

Aufgabe 22.57 (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 22.58

Bestimme für die Summennorm und die euklidische Norm jeweils Konstanten r und s wie in Lemma 22.41.

Aufgabe 22.59 (Submultiplikativität der euklidischen Norm)

Identifizieren wir den $\mathbb{R}$-Vektorraum $\mathrm{Mat}(m\times n,\mathbb{R})$ der $m\times n$-Matrizen mit $\mathbb{R}^{m\times n}$, so definiert

$$||(a_{ij})||_2 := \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2}$$

gerade die euklidische Norm auf $\mathrm{Mat}(m\times n,\mathbb{R})$.

Zeige, für eine Matrix $A \in \mathrm{Mat}(m\times n,\mathbb{R})$ und einen Vektor $x \in \mathbb{R}^n$ gilt stets

$$||Ax||_2 \leq ||A||_2 \cdot ||x||_2.$$

Aufgabe 22.60 (Intervalle sind zusammenhängend)
Ein metrischer Raum M heißt *zusammenhängend*, wenn M und $\emptyset$ die einzigen Teilmengen sind, die zugleich offen und abgeschlossen sind.
Zeige, daß jedes abgeschlossene Intervall $[a, b]$ in $\mathbb{R}$ zusammenhängend ist.

Aufgabe 22.61
Es sei (M, d) ein metrischer Raum.

a. Begründe, weshalb das Minimum $d(x, A) := \min\{d(x, a) \mid a \in A\}$ existiert, wenn $x \in M$ beliebig und $A \subseteq M$ kompakt ist.

b. Für $A \subseteq M$ kompakt und $x \in M$ zeige, $d(x, A) = 0$ genau dann, wenn $x \in A$.

c. Sei $M' = \{A \subseteq M \mid A \text{ kompakt}\}$ und

$$d' : M' \times M' \longrightarrow \mathbb{R} : (A, B) \mapsto \max\{d(a, B) \mid a \in A\} + \max\{d(A, b) \mid b \in B\}.$$

Zeige, daß d' eine Metrik auf M' ist.

Aufgabe 22.62 a. Es sei $M = \mathbb{R}^{\mathbb{N}} = \{(a_n)_{n\in\mathbb{N}} \mid a_n \in \mathbb{R}\}$ der $\mathbb{R}$-Vektorraum der reellen Zahlenfolgen und für $A = (a_n)_{n\in\mathbb{N}}, B = (b_n)_{n\in\mathbb{N}} \in M$ sei

$$d(A, B) = \sum_{k=0}^{\infty} \frac{1}{2^k} \cdot \frac{|a_k - b_k|}{1 + |a_k - b_k|}.$$

Zeige, d ist eine Metrik auf M.

b. Zeige, eine Folge $(A_n)_{n\in\mathbb{N}}$ in M mit $A_n = (a_{nk})_{k\in\mathbb{N}}$ ist genau dann konvergent, wenn alle Komponentenfolgen $(a_{nk})_{n\in\mathbb{N}}$ konvergieren.

Aufgabe 22.63 (Kompakte Räume haben abzählbare dichte Teilmengen.)
Es sei M ein kompakter metrischer Raum. Zeige, daß es eine abzählbare Teilmenge A von M gibt, die dicht in M liegt, d.h. $\overline{A} = M$.

Aufgabe 22.64 (Abgeschlossen und beschränkt, aber nicht kompakt)
Wir betrachten die Abbildung

$$d : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} : (x, y) \mapsto \min\{1, |x - y|\}.$$

a. Zeige, d ist eine Metrik auf $\mathbb{R}$.

b. Zeige, $\mathbb{R}$ ist abgeschlossen und beschränkt bezüglich d.

c. Zeige, die Intervalle der Form (a, b) für $a < b$ sind offen in $\mathbb{R}$.

d. Zeige, $\mathbb{R}$ ist nicht kompakt bezüglich d.

Kann d als die Metrik zu einer Norm auf $\mathbb{R}$ geschrieben werden?

Aufgabe 22.65 (Abgeschlossen und beschränkt, nicht kompakt)
Wir betrachten die rationalen Zahlen $\mathbb{Q}$ mit der Betragsdifferenz als metrischen Raum. Zeige, ist $\delta > 0$ eine irrationale Zahl. dann ist

$$U_\delta(0) = \{x \in \mathbb{Q} \mid |x| < \delta\}$$

beschränkt sowie offen und abgeschlossen in $\mathbb{Q}$, aber nicht kompakt.

Aufgabe 22.66 (Zusammenhang: abgeschlossen, beschränkt, kompakt)
Sei (M, d) ein metrischer Raum. Zeige, daß die folgenden Aussagen gleichwertig sind:

a. Jede abgeschlossene und beschränkte Teilmenge von M ist kompakt.

b. Für jedes $\varepsilon > 0$ und jedes $x \in M$ ist $\overline{U_\varepsilon(x)}$ kompakt.

Aufgabe 22.67
Sei V ein normierter Raum, $W \leq V$ ein Unterraum von V.

a. Zeige, ist W endlich-dimensional, dann ist W abgeschlossen in V.

b. Zeige, ist W abgeschlossen und $x \in V \setminus W$, dann gilt

$$d(x, W) := \inf_{y \in W} ||x - y|| > 0.$$

Aufgabe 22.68 (Heine-Borel gilt nur im Endlich-Dimensionalen.)
Sei V ein normierter Raum, so daß $\overline{U_1(0)}$ kompakt ist, dann gilt $\dim_{\mathbb{R}}(V) < \infty$.

Aufgabe 22.69 (Die Vielfachen irrationaler Zahlen liegen dicht.)
Zeige, ist $q \in \mathbb{R} \setminus \mathbb{Q}$ eine irrationale Zahl und bezeichnet

$$x \bmod 1 := x - \lfloor x \rfloor$$

den gebrochenen Anteil von x, dann ist die Menge

$$A := \{m \cdot q \bmod 1 \mid m \in \mathbb{N}\}$$

dicht im Intervall $[0, 1)$.

Aufgabe 22.70

Zeige, jeder Punkt des Einheitskreises

$$S^1(\mathbb{R}) = \{(x,y)^t \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$$

ist ein Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}}$ mit

$$a_n = \big(\cos(n), \sin(n)\big)^t.$$

Insbesondere sind die Mengen $\{\cos(n) \mid n \in \mathbb{N}\}$ und $\{\sin(n) \mid n \in \mathbb{N}\}$ dicht im Intervall $[-1,1]$.

Hinweis: beachte Aufgabe 22.69 und verwende ohne Beweis, daß π eine irrationale Zahl ist.

§ 23 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

Wie bei Folgen verallgemeinern sich auch bei Funktionen die Definitionen für Grenzwerte von der eindimensionalen Analysis auf metrische Räume, indem man die Betragsdifferenz durch die Metrik d ersetzt.

Definition 23.1 (ε-δ-Kriterium für Grenzwerte von Abbildungen)
Sei $U \subseteq M$, $f : U \longrightarrow M'$ eine Abbildung und $a \in M$ ein Häufungspunkt von U. Wir nennen $y \in M'$ den *Grenzwert* von f in a, falls

$$\forall\, \varepsilon > 0 \;\; \exists\, \delta_\varepsilon > 0 \;:\; \forall\, x \in U \text{ mit } 0 < d(x,a) < \delta_\varepsilon \;\text{ gilt }\; d'\big(f(x), y\big) < \varepsilon,$$

d.h.

$$\forall\, \varepsilon > 0 \;\; \exists\, \delta_\varepsilon > 0 \;:\; f\big(U_{\delta_\varepsilon}(a) \cap \big(U \setminus \{a\}\big)\big) \subseteq U_\varepsilon(y).$$

Wir schreiben dann

$$\lim_{x \to a} f(x) = y$$

oder "$f(x) \longrightarrow y$ für $x \to a$" und sagen, $f(x)$ *konvergiert* gegen y für x gegen a.

Auch in metrischen Räumen gilt das Folgenkriterium für Grenzwerte von Abbildungen.

Proposition 23.2 (Folgenkriterium für Grenzwerte von Abbildungen)
Es sei $U \subseteq M$, $f : U \longrightarrow M'$ eine Abbildung und a ein Häufungspunkt von U. Dann sind die beiden folgenden Aussagen gleichwertig:

a. $\lim\limits_{x \to a} f(x) = y$.

b. $\forall\, (a_n)_{n \in \mathbb{N}}$ mit $a_n \in U \setminus \{a\}$ und $\lim\limits_{n \to \infty} a_n = a$ gilt $\lim\limits_{n \to \infty} f(a_n) = y$.

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. $\Longrightarrow$ b.: Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $U\setminus\{a\}$ mit $\lim\limits_{n\to\infty} a_n = a$. Wir müssen $\lim\limits_{n\to\infty} f(a_n) = y$ zeigen. Dazu sei $\varepsilon > 0$ gegeben.
Wegen $\lim\limits_{x\to a} f(x) = y$ gibt es ein $\delta_\varepsilon > 0$, so daß aus $x \in U$ mit $0 < d(x,a) < \delta_\varepsilon$ auch $d'\big(f(x), y\big) < \varepsilon$ folgt.
Wegen $\lim\limits_{n\to\infty} a_n = a$ 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 $d'\big(f(a_n), y\big) < \varepsilon$. Damit ist $f(a_n) \longrightarrow y$ gezeigt.

b. $\Longrightarrow$ a.: Wir nehmen an, y wäre nicht der Grenzwert von f in a. Dann gilt:

$$\exists\, \varepsilon > 0 \;:\; \forall\, \delta_\varepsilon > 0\; \exists\, x_{\delta_\varepsilon} \in U \;\text{ mit }\; 0 < d(x_{\delta_\varepsilon}, a) < \delta_\varepsilon, \;\text{ aber } d'\big(f(x_{\delta_\varepsilon}), y\big) \geq \varepsilon.$$

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

$$0 < d(a_n, a) < \frac{1}{n} \longrightarrow 0,$$

so daß $a_n \longrightarrow a$, und zugleich gilt

$$d'(f(a_n), y) \geq \varepsilon$$

für alle $n \in \mathbb{N}$. Dies ist ein Widerspruch dazu, daß $f(a_n)$ gegen y konvergieren muß.

□

Bemerkung 23.3 (Unabhängigkeit des Grenzwertes von der Norm)
Betrachten wir $\mathbb{R}^n$ und $\mathbb{R}^m$ als metrische Räume mittels einer (beliebigen) Norm und $U \subseteq \mathbb{R}^n$. Dann ist der Grenzwert $\lim\limits_{x\to a} f(x)$ wegen Proposition 23.2 unabhängig von den gewählten Normen, da die Konvergenz von Folgen unabhängig von diesen ist.

Lemma 23.4
Es sei $U \subseteq M$, $f : U \longrightarrow M'$ eine Abbildung und a ein Häufungspunkt von U. Existiert der Grenzwert $\lim\limits_{x\to a} f(x)$, so ist er eindeutig bestimmt.

Beweis: Dies folgt wie in Proposition 13.10 aus dem Folgenkriterium für Grenzwerte von Funktionen 23.2 und der Eindeutigkeit des Grenzwertes bei Folgen 22.23. Genauer, da a ein Häufungspunkt von U ist, gibt es nach Lemma 22.25 eine Folge $(a_n)_{n\in\mathbb{N}}$ in $U\setminus\{a\}$ mit $a_n \longrightarrow a$. Falls nun y und z zwei Grenzwerte von $f(x)$ für $x \to a$ sind, so folgt mit den eben erwähnten Sätzen

$$y = \lim_{n\to\infty} f(a_n) = z.$$

□

Beispiel 23.5 (Grenzwerte)
Sei $M = \mathbb{R}^2$ und $U = \mathbb{R}^2 \setminus \{(0,0)^t\}$, dann ist $a = (0,0)^t$ ein Häufungspunkt von U, da es Folgen in U gibt, die gegen a konvergieren.

a. Wir betrachten die Funktion

$$f : U \longrightarrow \mathbb{R} : (x,y)^t \mapsto \frac{xy}{x^2+y^2}.$$

Wir betrachten die Folgen $(a_n)_{n\geq 1}$ und $(b_n)_{n\geq 1}$ mit

$$a_n = \left(\frac{1}{n}, \frac{1}{n}\right)^t$$

und

$$b_n = \left(\frac{1}{n}, 0\right)^t.$$

Beide konvergieren gegen a, aber es gilt

$$f(a_n) = \frac{\frac{1}{n^2}}{\frac{1}{n^2}+\frac{1}{n^2}} = \frac{1}{2} \to \frac{1}{2}$$

und

$$f(b_n) = 0 \to 0.$$

Mithin existiert der Grenzwert $\lim_{z\to a} f(z)$ *nicht*!
Man beachte übrigens, daß die beiden Grenzwerte

$$\lim_{x\to 0} f(x,0) = 0 \quad \text{und} \quad \lim_{y\to 0} f(0,y) = 0$$

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

$$g : U \longrightarrow \mathbb{R} : (x,y)^t \mapsto \frac{xy}{\sqrt{x^2+y^2}},$$

so gilt für $(x,y) \in U$

$$|g(x,y) - 0| = \frac{|x|\cdot|y|}{\sqrt{x^2+y^2}} \leq \frac{\sqrt{x^2+y^2}\cdot\sqrt{x^2+y^2}}{\sqrt{x^2+y^2}} = \sqrt{x^2+y^2} = ||(x,y)^t||_2.$$

Haben wir nun eine beliebige Folge $(a_n)_{n\in\mathbb{N}}$ in U, die gegen $a = (0,0)^t$ konvergiert, so gilt $||a_n||_2 = ||a_n - a||_2 \to 0$ und mithin

$$|g(a_n) - 0| \to 0.$$

Wir erhalten deshalb

$$\lim_{z\to a} g(z) = 0.$$

Proposition 23.6 (Grenzwertsätze in normierten Räumen)
Es seien $f : U \longrightarrow V$ und $g : U \longrightarrow V$ zwei Funktionen mit $U \subseteq M$, a ein Häufungspunkt von U und $c \in \mathbb{R}$. Wenn $\lim\limits_{x\to a} f(x)$ und $\lim\limits_{x\to a} g(x)$ existieren, so gelten:

a. $\lim\limits_{x\to a}(c \cdot f)(x) = c \cdot \lim\limits_{x\to a} f(x)$.

b. $\lim\limits_{x\to a}(f+g)(x) = \lim\limits_{x\to a} f(x) + \lim\limits_{x\to a} g(x)$.

c. $\lim\limits_{x\to a}(f-g)(x) = \lim\limits_{x\to a} f(x) - \lim\limits_{x\to a} g(x)$.

d. Ist $V = \mathbb{R}$, so ist auch $\lim\limits_{x\to a}(f \cdot g)(x) = \lim\limits_{x\to a} f(x) \cdot \lim\limits_{x\to a} g(x)$.

e. Falls zudem $V = \mathbb{R}$ und $\lim\limits_{x\to a} f(x) \neq 0$, so ist a ein Häufungspunkt der Menge $X = \{x \in U \mid f(x) \neq 0\}$ und es gilt

$$\lim_{x\to a} \frac{1}{f}(x) = \frac{1}{\lim\limits_{x\to a} f(x)}.$$

Beweis: Die Aussagen a.-c. folgen aus dem Folgenkriterium für Grenzwerte von Funktionen 23.2 und den Grenzwertsätzen für Folgen 22.23 unter Berücksichtigung von Lemma 22.25. Die Aussage d. folgt analog, allerdings ersetzt man Proposition 22.23 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 22.25 gibt es eine Folge $(a_n)_{n\in\mathbb{N}}$ in $U \setminus \{a\}$, die gegen a konvergiert, und nach dem Folgenkriterium 23.2 gilt

$$f(a_n) \longrightarrow \lim_{x\to a} f(x) =: y.$$

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 X ist mit $a_n \longrightarrow a$. Nach Lemma 22.25 ist dann a ein Häufungspunkt von X. Die Aussage zum Grenzwert folgt dann wieder aus den Grenzwertsätzen für Folgen 11.15 und dem Folgenkriterium 23.2. □

B) Stetigkeit in metrischen und in normierten Räumen

Definition 23.7 (ε-δ-Kriterium für Stetigkeit)
Sei $f : M \longrightarrow M'$ eine Abbildung und $a \in M$. Wir nennen f *stetig in* a, wenn

$$\forall\, \varepsilon > 0\ \ \exists\, \delta_\varepsilon > 0\ :\ \forall\, x \in M \text{ mit } d(x,a) < \delta_\varepsilon \text{ gilt } d'\big(f(x), f(a)\big) < \varepsilon,$$

d.h., wenn

$$\forall\, \varepsilon > 0\ \ \exists\, \delta_\varepsilon > 0\ :\ f\big(U_{\delta_\varepsilon}(a)\big) \subseteq U_\varepsilon\big(f(a)\big).$$

Die Funktion f heißt *stetig* (auf M), wenn sie stetig in jedem Punkt in M ist.

Lemma 23.8
Es sei $f : M \longrightarrow M'$ und $a \in M$ ein Häufungspunkt.
Genau dann ist f stetig in a, wenn $\lim\limits_{x\to a} f(x) = f(a)$.

Beweis: Dies folgt unmittelbar aus den Definitionen 23.1 und 23.7. □

Beispiel 23.9 (Metriken und Normen sind stetig.)

a. Ist M ein metrischer Raum und $c \in M$ fest gegeben. Dann ist die Abbildung

$$d(\cdot, c) : M \longrightarrow \mathbb{R} : x \mapsto d(x, c)$$

stetig auf M. Um dies zu sehen, setzen wir $\delta_\varepsilon = \varepsilon$ für $a \in M$ und $\varepsilon > 0$ und erhalten wegen der Dreiecksungleichung, daß

$$|d(x, c) - d(a, c)| \leq d(x, a) < \varepsilon$$

gilt, sobald nur $d(x, a) < \delta_\varepsilon$ erfüllt ist. Man beachte, daß der Beweis sogar zeigt, daß die Abbildung gleichmäßig stetig im Sinne von Definiton 23.27 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 $c = 0$.

c. Ist $||\cdot|| : \mathbb{R}^n \longrightarrow \mathbb{R}$ eine Norm auf $\mathbb{R}^n$, so ist sie stetig bezüglich der Maximumsnorm. Um dies zu sehen, beachten wir, daß es ein $r > 0$ mit

$$||x|| \leq r \cdot ||x||_\infty$$

gibt, und betrachten zu $a \in \mathbb{R}^n$ und $\varepsilon > 0$ dann $\delta_\varepsilon = \frac{\varepsilon}{r}$. Dann gilt für alle $x \in \mathbb{R}^n$ mit $||x - a||_\infty < \delta_\varepsilon$ auch

$$\big|\, ||x|| - ||a|| \,\big| \leq ||x - a|| \leq r \cdot ||x - a||_\infty < r \cdot \delta_\varepsilon = \varepsilon.$$

Mit Bemerkung 23.11 folgt die Aussage auch unmittelbar aus Teil b..

Satz 23.10 (Folgenkriterium für Stetigkeit)
Genau dann ist $f : M \longrightarrow M'$ stetig in $a \in M$, wenn

(67) $\quad \forall\, (a_n)_{n\in\mathbb{N}}$ mit $a_n \in M$ und $\lim\limits_{n\to\infty} a_n = a$ gilt $\lim\limits_{n\to\infty} f(a_n) = f(a)$.

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.

"$\Longrightarrow$": Sei zunächst f stetig in a und sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in M mit $\lim\limits_{n\to\infty} a_n = a$.
Wir müssen $\lim\limits_{n\to\infty} f(a_n) = f(a)$ zeigen. Dazu sei $\varepsilon > 0$ gegeben.
Da f stetig in a ist, gibt es ein $\delta_\varepsilon > 0$, so daß aus $x \in M$ mit $d(x,a) < \delta_\varepsilon$ auch $d'\big(f(x), f(a)\big) < \varepsilon$ folgt.
Wegen $\lim\limits_{n\to\infty} a_n = a$ 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 $a_n \in M$ die Bedingung $d(a_n, a) < \delta_\varepsilon$ und somit ist auch $d'\big(f(a_n), f(a)\big) < \varepsilon$. Damit ist $f(a_n) \longrightarrow f(a)$ gezeigt.
"$\Longleftarrow$": Wir nehmen an, f wäre nicht stetig in a. Dann gilt:

$$\exists\, \varepsilon > 0 \;:\; \forall\, \delta_\varepsilon > 0 \;\; \exists\, x_{\delta_\varepsilon} \in M \;\text{ mit }\; d(x_{\delta_\varepsilon}, a) < \delta_\varepsilon, \;\text{ aber } d'\big(f(x_{\delta_\varepsilon}), f(a)\big) \geq \varepsilon.$$

Für $n \geq 1$ und $\delta_\varepsilon = \frac{1}{n}$ setzen wir $a_n := x_{\delta_\varepsilon} = x_{\frac{1}{n}} \in M \setminus \{a\}$. Dann gilt

$$0 \leq d(a_n, a) < \frac{1}{n} \longrightarrow 0,$$

so daß $a_n \longrightarrow a$, und zugleich gilt

$$d'\big(f(a_n), f(a)\big) \geq \varepsilon$$

für alle $n \in \mathbb{N}$. Dies ist ein Widerspruch dazu, daß $f(a_n)$ gegen $f(a)$ konvergieren muß.

□

Bemerkung 23.11 (Stetigkeit unabhängig von gewählter Norm)
Betrachten wir $\mathbb{R}^n$ und $\mathbb{R}^m$ als metrische Räume mittels einer (beliebigen) Norm und $U \subseteq \mathbb{R}^n$ als metrischen Raum mittels der Einschränkung der Metrik. Dann ist es für die Frage, ob eine Abbildung $f : U \longrightarrow \mathbb{R}^m$ stetig ist, wegen des Folgenkriteriums der Stetigkeit 23.10 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 23.12 (Stetigkeit von $x \mapsto x_i$)
Die Funktion $\pi_i : \mathbb{R}^n \longrightarrow \mathbb{R} : (x_1, \ldots, x_n)^t \mapsto x_i$ ist stetig auf $\mathbb{R}^n$.
Um dies zu sehen, betrachten wir eine Folge $(a_k)_{k\in\mathbb{N}}$ in $\mathbb{R}^n$, die gegen c konvergiert. Dann ist $\big(\pi_i(a_k)\big)_{k\in\mathbb{N}}$ die i-te Komponentenfolge von $(a_k)_{k\in\mathbb{N}}$ und konvergiert somit gegen $c_i = \pi_i(c)$. Mithin ist π_i stetig in c.

C) Einfache Eigenschaften stetiger Abbildungen

Proposition 23.13 (Grenzwertsätze für stetige Abbildungen)
Seien $f : M \longrightarrow V$ und $g : M \longrightarrow V$ stetig in $a \in M$ und sei $c \in \mathbb{R}$.

a. $c \cdot f$, $f + g$ und $f - g$ sind stetig in a.

b. Ist $V = \mathbb{R}$, so ist auch $f \cdot g$ stetig in a.

c. Ist $V = \mathbb{R}$ und $g(a) \neq 0$, so ist auch $\frac{f}{g} : M \setminus g^{-1}(0) \longrightarrow \mathbb{R}$ stetig in a.

Beweis: Der Beweis folgt aus dem Folgenkriterium für Stetigkeit 23.10 und den Grenzwertsätzen für Folgen in V, Proposition 22.23, bzw. in $\mathbb{R}$, Proposition 11.15.

Z.B. sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in M mit $a_n \longrightarrow a$, dann gilt

$$(f+g)(a_n) = f(a_n) + g(a_n) \longrightarrow f(a) + g(a) = (f+g)(a),$$

da f und g in a stetig sind. Also ist auch $f + g$ stetig in a. □

Beispiel 23.14 (Polynomfunktionen)
Sind $t_1, \ldots, t_n$ Veränderliche und ist $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, so setzen wir

$$t = (t_1, \ldots, t_n)$$

und

$$t^\alpha = t_1^{\alpha_1} \cdot \ldots \cdot t_n^{\alpha_n}.$$

Ferner nennen wir

$$|\alpha| = \alpha_1 + \ldots + \alpha_n$$

den *Grad* des *Monoms* t^α. Für reelle Zahlen $a_\alpha \in \mathbb{R}$ mit $\alpha \in \mathbb{N}^n$ und $0 \leq |\alpha| \leq d$ nennen wir einen Ausdruck der folgenden Form ein *Polynom* in den Veränderlichen $t_1, \ldots, t_n$:

$$p = \sum_{|\alpha|=0}^{d} a_\alpha \cdot t^\alpha = \sum_{\substack{\alpha\in\mathbb{N}^n \\ 0\leq|\alpha|\leq d}} a_\alpha \cdot t_1^{\alpha_1} \cdot \ldots \cdot t_n^{\alpha_n}$$

Wir definieren den *Grad* des Polynoms p als

$$\deg(p) := \sup\left\{|\alpha| \;\middle|\; a_\alpha \neq 0\right\} \in \mathbb{N} \cup \{-\infty\},$$

wobei automatisch $\deg(0) := -\infty$ gilt. Mit $\mathbb{R}[t] = \mathbb{R}[t_1, \ldots, t_n]$ bezeichnen wir die Menge aller Polynome in den Veränderlichen $t_1, \ldots, t_n$, so daß der Grad eine Abbildung $\deg : \mathbb{R}[t] \longrightarrow \mathbb{N} \cup \{-\infty\}$ ist.

Für ein Polynom $p = \sum_{|\alpha|=0}^{d} a_\alpha \cdot t^\alpha$ und ein $x = (x_1, \ldots, x_n)^t \in \mathbb{R}^n$ setzen wir

$$p(x) := \sum_{|\alpha|=0}^{d} a_\alpha \cdot x^\alpha = \sum_{\substack{\alpha\in\mathbb{N}^n \\ 0\leq|\alpha|\leq d}} a_\alpha \cdot x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n},$$

wobei wir uns wieder die *Multiindexnotation* $x^\alpha = x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}$ zu eigen machen.

Sind $p, q \in \mathbb{R}[t]$ zwei Polynome, $q \neq 0$ nicht das Nullpolynom, so nennen wir

$$p : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto p(x)$$

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 23.13 und Beispiel 23.12.

Proposition 23.15 (Komposition stetiger Funktionen)
Seien (M, d), (M', d') und (M'', d'') metrische Räume, $f : M \to M'$, $g : M' \to M''$. 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, Proposition 14.8.

Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in M mit $a_n \longrightarrow a$, dann ist $(f(a_n))_{n\in\mathbb{N}}$ eine Folge in M' und, da f stetig in a ist, gilt zudem $f(a_n) \longrightarrow f(a)$. Nun ist auch g stetig in $f(a)$, so daß

$$(g \circ f)(a_n) = g(f(a_n)) \longrightarrow g(f(a)) = (g \circ f)(a)$$

folgt. Wegen des Folgenkriteriums für Stetigkeit 23.10 ist $g \circ f$ dann stetig in a. □

Beispiel 23.16
Die Funktion $||\cdot||_2 : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y)^t \mapsto \sqrt{x^2 + y^2}$ 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 23.17 (Stetige Urbilder offener Mengen sind offen.)
Sei $f : M \longrightarrow M'$ eine Abbildung. Dann sind die folgenden Aussagen gleichwertig:

a. f ist stetig auf M.

b. Das Urbild jeder offenen Menge in M' ist offen in M.

c. Das Urbild jeder abgeschlossenen Menge in M' ist abgeschlossen in M.

Beweis:

a. $\Longrightarrow$ b.: Sei zunächst f stetig, $O' \subseteq M'$ sei offen in M' und $a \in f^{-1}(O')$. Wir müssen zeigen, daß eine ganze Umgebung von a in $f^{-1}(O')$ liegt.

Da O' offen in M' ist und $f(a) \in O'$, gibt es ein $\varepsilon > 0$ mit $U_\varepsilon\big(f(a)\big) \subseteq O'$, und da f stetig in a ist, gibt es ein $\delta > 0$ mit

$$f\big(U_\delta(a)\big) \subseteq U_\varepsilon\big(f(a)\big) \subseteq O'.$$

Dann gilt aber auch

$$U_\delta(a) \subseteq f^{-1}\big(f\big(U_\delta(a)\big)\big) \subseteq f^{-1}(O').$$

Also ist $f^{-1}(O')$ offen in M.

b. $\Longrightarrow$ a.: Sei umgekehrt das Urbild jeder offenen Menge offen in M, und seien $a \in M$ und $\varepsilon > 0$ beliebig. Dann ist $f^{-1}\big(U_\varepsilon(f(a))\big)$ offen in M. Deshalb gibt es ein $\delta > 0$ mit

$$U_\delta(a) \subseteq f^{-1}\big(U_\varepsilon(f(a))\big)$$

und mithin

$$f\big(U_\delta(a)\big) \subseteq f\Big(f^{-1}\big(U_\varepsilon(f(a))\big)\Big) \subseteq U_\varepsilon\big(f(a)\big).$$

f ist also stetig in a, und somit stetig auf M.

b. $\Longrightarrow$ 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. $\Longrightarrow$ 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.

□

D) Stetigkeit für Abbildungen nach $\mathbb{R}^m$

Abbildungen in den $\mathbb{R}^m$ zeichnen sich dadurch aus, daß ihre Verknüpfung mit der Projektion auf die i-te Koordinate eine Abbildung nach $\mathbb{R}$ liefert, die i-te Komponentenfunktion.

Definition 23.18
Ist $f : M \longrightarrow \mathbb{R}^m : x \mapsto \big(f_1(x), \ldots, f_m(x)\big)^t$ eine Abbildung, so heißen die Funktionen $f_i : M \longrightarrow \mathbb{R}$ die *Komponentenfunktionen* von f.

Wir können nun die Untersuchung bestimmter Eigenschaften von Abbildungen in den $\mathbb{R}^m$ auf die Untersuchung von Eigenschaften der Komponentenfunktionen zurückführen.

Satz 23.19 (Komponentenweise Stetigkeit)
Es sei $f : M \longrightarrow \mathbb{R}^m$ eine Abbildung und $a \in M$.

a. Genau dann existiert der Grenzwert $\lim\limits_{x\to a} f(x)$, wenn die Grenzwerte $\lim\limits_{x\to a} f_i(x)$ aller Komponentenfunktionen existieren.

b. Genau dann ist f stetig in a, wenn dies auch alle Komponentenfunktionen sind.

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in M, die gegen a konvergiert. Wegen des Satzes zur komponentenweisen Konvergenz von Folgen in $\mathbb{R}^m$, Proposition 22.38, gilt $f(a_n) \to f(a)$ genau dann, wenn $f_i(a_n) \to f_i(a)$ für alle $i = 1, \ldots, m$. Die Aussage folgt also aus dem Folgenkriterium für Grenzwerte 23.2 bzw. für Stetigkeit 23.10. □

Beispiel 23.20

a. Die Abbildung

$$\mathbb{R}^2 \setminus \{(0,0)^t\} \longrightarrow \mathbb{R}^2 : (x,y)^t \mapsto \left(\frac{xy}{x^2+y^2}, x^3+5xy\right)^t$$

ist stetig auf $\mathbb{R}^2 \setminus \{(0,0)^t\}$, da die Komponentenfunktionen nach Beispiel 23.14 stetig sind.

b. Ist $A = (a_{ij}) \in \mathrm{Mat}(m \times n, \mathbb{R})$, so hat die lineare Abbildung

$$f_A : \mathbb{R}^n \longrightarrow \mathbb{R}^m : x \mapsto Ax$$

die Komponentenfunktionen

$$f_i : \mathbb{R}^n \longrightarrow \mathbb{R} : (x_1, \ldots, x_n)^t \mapsto a_{i1} \cdot x_1 + \ldots + a_{in} \cdot x_n.$$

Als Polynomfunktionen sind diese stetig, so daß auch f_A stetig ist.

Beispiel 23.21 (Determinante und Invertieren als stetige Abbildungen)
Wir können $\mathrm{Mat}_n(\mathbb{R})$ mit dem Vektorraum $\mathbb{R}^{n^2}$ identifizieren, so daß $\mathrm{Mat}_n(\mathbb{R})$ ein normierter Raum ist. Die Determinante

$$\det : \mathrm{Mat}_n(\mathbb{R}) \longrightarrow \mathbb{R}$$

ist aufgrund der Leibnitz-Formel eine Polynomfunktion und somit stetig. Dann ist das Urbild

$$\det^{-1}(0) = \{A \in \mathrm{Mat}_n(\mathbb{R}) \mid \det(A) = 0\}$$

abgeschlossen in $\mathrm{Mat}_n(\mathbb{R})$ als Urbild der abgeschlossenen Menge $\{0\}$ in $\mathbb{R}$. Ihr Komplement

$$\mathrm{Gl}_n(\mathbb{R}) = \{A \in \mathrm{Mat}_n(\mathbb{R}) \mid \det(A) \neq 0\},$$

die Menge der invertierbaren $n \times n$-Matrizen, ist somit offen in $\mathrm{Mat}_n(\mathbb{R})$.

Betrachten wir nun die Abbildung

$$\mathrm{Gl}_n(\mathbb{R}) \longrightarrow \mathrm{Mat}_n(\mathbb{R}) : A \mapsto A^{-1} = \frac{1}{\det(A)} \cdot A^{\#},$$

die einer invertierbaren Matrix ihre Inverse zuordnet, so sind ihre Komponentenfunktionen wegen des Satzes über die Adjunkte, Korollar A9.30, rationale Funktionen in den Einträgen von 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 A nur zu einer kleinen Störung in den Einträgen der Inversen führt! Gleiches gilt damit für die Lösung eines eindeutig lösbaren Gleichungssystems.

Bemerkung 23.22 (Stetigkeit ist stärker als getrennte Stetigkeit)
Ist $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ eine Abbildung und $y = (y_1, \ldots, y_n)^t \in \mathbb{R}^n$, so können wir die Funktionen

$$F_i : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto f(y_1, \ldots, y_{i-1}, x, y_{i+1}, \ldots, y_n)$$

definieren.

a. Ist f stetig in y, so sind die F_i alle stetig in y_i.
Denn ist $(a_k)_{k\in\mathbb{N}}$ eine Folge in $\mathbb{R}$, die gegen y_i konvergiert, so konvergiert auch die Folge der $(y_1, \ldots, y_{i-1}, a_k, y_{i+1}, \ldots, y_n)$ gegen y und somit gilt

$$F_i(a_k) = f(y_1, \ldots, y_{i-1}, a_k, y_{i+1}, \ldots, y_n) \longrightarrow f(y) = F_i(y_i).$$

Man sagt, daß Stetigkeit die getrennte Stetigkeit in Richtung der Koordinatenachsen impliziert.

b. Die Umkehrung gilt nicht, wie im wesentlichen aus Beispiel 23.5 folgt. Setzen wir

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y)^t \mapsto \begin{cases} \frac{xy}{x^2+y^2}, & \text{falls } (x, y)^t \neq (0, 0)^t, \\ 0, & \text{falls } (x, y)^t = (0, 0)^t, \end{cases}$$

so ist die Funktion nicht stetig in $a = (0, 0)^t$, da der Grenzwert von $f(x, y)$ in a nach Beispiel 23.5 nicht existiert. Hingegen sind die Funktionen

$$F_1 : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto f(x, 0) = 0$$

und

$$F_2 : \mathbb{R} \longrightarrow \mathbb{R} : y \mapsto f(0, y) = 0$$

jeweils stetig in 0.

E) Stetige Funktionen auf kompakten Mengen

Satz 23.23 (Stetige Bilder von Kompakta sind kompakt.)
Ist M kompakt und $f : M \longrightarrow M'$ stetig, so ist $f(M)$ kompakt in M'.

Beweis: Ist $(b_n)_{n\in\mathbb{N}}$ eine Folge in $f(M)$, so besitzt jedes b_n ein Urbild $a_n \in M$ unter f, d.h. wir haben eine Folge $(a_n)_{n\in\mathbb{N}}$ in M mit $f(a_n) = b_n$. Da M kompakt ist, besitzt $(a_n)_{n\in\mathbb{N}}$ dann eine konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ mit Grenzwert $a \in M$. Da f stetig ist, gilt dann aber auch

$$b_{n_k} = f(a_{n_k}) \longrightarrow f(a) \in f(M),$$

so daß auch $(b_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge mit Grenzwert in $f(M)$ besitzt. Aus dem Satz von Bolzano-Weierstraß 22.30 folgt dann, daß $f(M)$ kompakt ist. □

Der obige Beweis von Satz 23.23 funktioniert nur in metrischen Räumen, da ansonsten der Satz von Bolzano-Weierstraß nicht gelten muß. Mit Hilfe der Charakterisierung von Stetigkeit in Proposition 23.17 kann man aber auch einen alternativen Beweis geben.

Alternativer Beweis für Satz 23.23: 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 23.17 offen in M und

$$M = f^{-1}\big(f(M)\big) \subseteq \bigcup_{i\in I} f^{-1}(O_i).$$

Also ist $\big(f^{-1}(O_i)\big)_{i\in I}$ eine offene Überdeckung von M, und da M kompakt ist, gibt es eine endliche Teilüberdeckung. D.h. es gibt $i_1, \ldots, i_n \in I$ mit

$$M \subseteq f^{-1}(O_{i_1}) \cup \ldots \cup f^{-1}(O_{i_n}).$$

Wenden wir f auf beide Seiten an, so erhalten wir

$$f(A) \subseteq f\big(f^{-1}(O_{i_1}) \cup \ldots \cup f^{-1}(O_{i_n})\big) \subseteq O_{i_1} \cup \ldots \cup O_{i_n}.$$

Mithin besitzt die offene Überdeckung $(O_i)_{i\in I}$ eine endliche Teilüberdeckung, und somit ist $f(M)$ kompakt. □

Das folgende Korollar ist eine direkte Verallgemeinerung von Satz 14.16.

Korollar 23.24 (Maxima & Minima stetiger Funktionen auf Kompakta)
Sei $f : M \longrightarrow \mathbb{R}$ stetig und M kompakt, dann gibt es $c, d \in M$ mit

$$f(c) \leq f(x) \leq f(d)$$

für alle $x \in M$.

Beweis: Wegen Satz 23.23 ist $f(M)$ kompakt in $\mathbb{R}$ und wegen Korollar 22.46 enthält $f(M)$ dann sein Maximum und sein Minimum. □

Korollar 23.25 (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 23.17 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

$$f(A) = \big(f^{-1}\big)^{-1}(A) \subseteq f(M)$$

abgeschlossen in $f(M)$ ist.

Da A abgeschlossen im kompakten metrischen Raum M ist, ist A nach Korollar 22.32 ebenfalls kompakt. Dann ist aber $f(A)$ kompakt in M' nach Satz 23.23. Als kompakte Menge ist $f(A)$ mithin abgeschlossen in M' nach Satz 22.21. Aber dann ist die Menge

$$f(A) = f(A) \cap f(M)$$

abgeschlossen in $f(M)$ nach Aufgabe 22.53. □

Beispiel 23.26
Die Kompaktheit des Definitionsbereiches im Umkehrsatz 23.25 ist eine notwendige Bedingung. Um dies zu sehen, betrachten wir die Abbildung

$$f : [0, 2\pi) \longrightarrow \mathbb{R}^2 : x \mapsto \big(\cos(x), \sin(x)\big)^t.$$

f bildet bijektiv auf sein Bild ab, den Einheitskreis mit dem Ursprung als Mittelpunkt

$$\mathrm{Im}(f) = S^1 = \{(x, y)^t \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}.$$

Die Umkehrabbildung

$$f^{-1} : S^1 \longrightarrow [0, 2\pi)$$

ist aber nicht stetig, da es keine δ-Umgebung des Punktes $a = (1, 0)^t \in S^1$ gibt, die ganz in die 1-Umgebung des Punktes 0 in $[0, 2\pi)$ abgebildet wird; das Bild jeder noch so kleinen Umgebung von a enthält Zahlen nahe bei 2π.

F) Gleichmäßige Stetigkeit

Definition 23.27 (Gleichmäßige Stetigkeit)
Eine Abbildung $f : M \longrightarrow M'$ heißt *gleichmäßig stetig* auf M, wenn

$$\forall\, \varepsilon > 0\ \exists\, \delta_\varepsilon > 0\ :\ \forall\, x, y \in M \text{ mit } d(x,y) < \delta_\varepsilon \text{ gilt } d'\big(f(x), f(y)\big) < \varepsilon,$$

d.h.

$$\forall\, \varepsilon > 0\ \exists\, \delta_\varepsilon > 0\ :\ \forall\, x \in M \text{ gilt } f\big(U_{\delta_\varepsilon}(x)\big) \subseteq U_\varepsilon\big(f(x)\big).$$

Bemerkung 23.28 (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 23.29 (Stetigkeit auf Kompakta impliziert gleichmäßige Stetigkeit)
Ist M kompakt und $f : M \longrightarrow M'$ 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:

$$\exists\, \varepsilon > 0\ :\ \forall\, \delta_\varepsilon > 0\ \exists\, x_{\delta_\varepsilon}, y_{\delta_\varepsilon} \in M\ :\ d(x_{\delta_\varepsilon}, y_{\delta_\varepsilon}) < \delta_\varepsilon, \text{ aber } d'\big(f(x_{\delta_\varepsilon}), f(y_{\delta_\varepsilon})\big) \geq \varepsilon.$$

Für $n \geq 1$ und $\delta_\varepsilon := \frac{1}{n}$ setzen wir $a_n := x_{\delta_\varepsilon} = x_{\frac{1}{n}}$ und $b_n := y_{\delta_\varepsilon} = y_{\frac{1}{n}}$. Damit erhalten wir zwei Folgen $(a_n)_{n\geq 1}$ und $(b_n)_{n\geq 1}$ in der kompakten Menge M. Nach dem Satz von Bolzano-Weierstraß 22.30 besitzt $(a_n)_{n\geq 1}$ eine konvergente Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$, und ebenso besitzt dann $(b_{n_k})_{k\in\mathbb{N}}$ eine konvergente Teilfolge $(b_{n_{k_l}})_{l\in\mathbb{N}}$. Nach Konstruktion gilt

$$0 \leq d(a_{n_{k_l}}, b_{n_{k_l}}) \leq \frac{1}{n_{k_l}} \longrightarrow 0,$$

so daß die Grenzwerte von $(a_{n_{k_l}})_{l\in\mathbb{N}}$ und $(b_{n_{k_l}})_{l\in\mathbb{N}}$ wegen Proposition 22.23 g. übereinstimmen müssen, d.h.

$$a_{n_{k_l}} \longrightarrow y \quad \text{und} \quad b_{n_{k_l}} \longrightarrow y.$$

Da f stetig ist, folgt unter Berücksichtigung von Proposition 22.23 g.

$$0 = d'\big(f(y), f(y)\big) \longleftarrow d'\big(f(a_{n_{k_l}}), f(b_{n_{k_l}})\big) \geq \varepsilon,$$

was ein offensichtlicher Widerspruch ist. □

Beispiel 23.30
Es sei

$$M = \big\{(x_1, x_2)^t \in \mathbb{R}^2 \;\big|\; x_1^2 + x_2^2 \leq 1\big\}$$

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

$$f : M \longrightarrow \mathbb{R} : x = (x_1, x_2)^t \mapsto x_1^2 + x_2^2.$$

ist dann als Polynomfunktion stetig auf M und M ist kompakt.

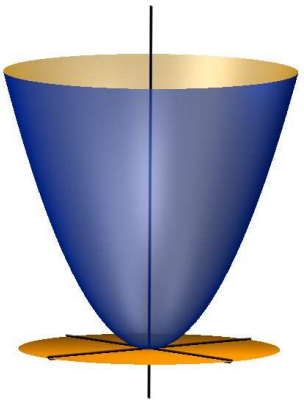


Abbildung 3. Der Graph von $f : M \longrightarrow \mathbb{R} : (x_1, x_2)^t \mapsto x_1^2 + x_2^2$

Mithin ist f gleichmäßig stetig auf M, und für alle $\varepsilon > 0$ gibt es also ein $\delta_\varepsilon > 0$, so daß für $\max\{|x_1 - y_1|, |x_2 - y_2|\} = ||x - y||_\infty < \delta_\varepsilon$ schon

$$|x_1^2 + x_2^2 - y_1^2 - y_2^2| = |f(x) - f(y)| < \varepsilon$$

gilt. Wir können hierzu $\delta_\varepsilon = \frac{\varepsilon}{4}$ wählen, denn dann gilt

$$\begin{aligned} |f(x) - f(y)| &\leq |x_1^2 - y_1^2| + |x_2^2 - y_2^2| = |x_1 + y_1| \cdot |x_1 - y_1| + |x_2 + y_2| \cdot |x_2 - y_2| \\ &\leq \big(|x_1| + |y_1|\big) \cdot |x_1 - y_1| + \big(|x_2| + |y_2|\big) \cdot |x_2 - y_2| \leq 2 \cdot \delta_\varepsilon + 2 \cdot \delta_\varepsilon = \varepsilon, \end{aligned}$$

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 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.

G) Lipschitz-Stetigkeit

Eine besonders starke Form der Stetigkeit ist die Lipschitz-Stetigkeit, die wir im Zusammenhang mit dem Banachschen Fixpunktsatz 26.4 und dem mehrdimensionalen Riemann-Integral benötigen.

Definition 23.31 (Lipschitz-stetig)
Eine Abbildung $f : M \longrightarrow M'$ heißt *Lipschitz-stetig* auf M mit Lipschitzkonstante $q \in \mathbb{R}_{>0}$, wenn

$$d'\big(f(x), f(y)\big) \leq q \cdot d(x, y)$$

für alle $x, y \in M$ gilt (siehe auch Aufgabe 14.36).

Beispiel 23.32
Ist $a \in \mathbb{R}$, so ist die lineare Abbildung $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto a \cdot x$ Lipschitz-stetig mit Lipschitz-Konstante $q = |a|$, da

$$|f(x) - f(y)| = |ax - ay| = |a| \cdot |x - y|.$$

Proposition 23.33 (Lipschitz stetige Abbildungen sind gleichmäßig stetig.)
Jede Lipschitz stetige Abbildung $f : M \longrightarrow M'$ 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

$$d'\big(f(x), f(y)\big) \leq q \cdot d(x, y) < q \cdot \delta_\varepsilon = \varepsilon.$$

□

Beispiel 23.34 (Gleichmäßig stetig impliziert nicht Lipschitz-stetig)
Die Wurzelfunktion $\sqrt{\cdot} : [0, \infty) \longrightarrow \mathbb{R} : x \mapsto \sqrt{x}$ ist nach Aufgabe 14.34 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 23.35 (Lipschitz-Stetigkeit auf $\mathbb{R}^n$)
Lipschitz-Stetigkeit für Abbildungen $f : U \longrightarrow \mathbb{R}^m$ 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 22.41).

H) Lokal Lipschitz-stetige Abbildungen

Definition 23.36 (Lokal Lipschitz-stetig)
$f : M \longrightarrow M'$ heißt *lokal Lipschitz-stetig*, wenn jeder Punkt in M eine offene Umgebung besitzt, so daß f auf diese eingeschränkt Lipschitz-stetig ist.

Beispiel 23.37
Die Wurzelfunktion $\sqrt{}:(0,\infty) \longrightarrow \mathbb{R}$ ist zwar nicht Lipschitz-stetig, wohl aber lokal Lipschitz-stetig. Denn für $a \in (0,\infty)$ und $\varepsilon = \frac{a}{2} > 0$ gilt für alle $x, y \in U_\varepsilon(a) = \left(\frac{a}{2}, \frac{3a}{2}\right)$ wegen des Mittelwertsatzes der Differentialrechnung

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

für ein geeignetes c zwischen x und y, und somit

$$|f(x) - f(y)| \leq q \cdot |x - y|$$

mit

$$q = \max\left\{ |f'(c)| \;\middle|\; \frac{a}{2} \leq c \leq \frac{3a}{2} \right\} \in \mathbb{R}.$$

Proposition 23.38 (Lokal Lipschitz auf Kompaktum = Lipschitz)
Ist $f : M \longrightarrow M'$ lokal Lipschitz-stetig und M kompakt, dann ist f Lipschitz-stetig auf M.

Beweis: Wir nehmen an, daß f nicht Lipschitz-stetig ist. Dann gibt es für jedes $n \geq 1$ Punkte $x_n, y_n \in M$ mit

$$d'\big(f(x_n), f(y_n)\big) > n \cdot d(x_n, y_n).$$

Da M kompakt ist, können wir den Satz von Bolzano-Weierstraß 22.30 anwenden und finden eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ von $(x_n)_{n\geq 1}$ mit

$$x_{n_k} \longrightarrow x \in M$$

und eine konvergente Teilfolge $(y_{n_{k_l}})_{l\in\mathbb{N}}$ von $(y_{n_k})_{k\in\mathbb{N}}$ mit

$$y_{n_{k_l}} \longrightarrow y \in M.$$

Zunächst wollen wir

$$x = y$$

zeigen. Dazu beachten wir, daß die stetige Abbildung

$$M \longrightarrow \mathbb{R} : z \mapsto d'\big(f(z), f(x)\big)$$

auf der kompakten Menge M nach Korollar 23.24 beschränkt ist. Es gibt also eine Konstante $C \geq 0$ mit

$$d'\big(f(z), f(x)\big) \leq C$$

für alle $z \in M$ und damit

$$d'\big(f(z), f(z')\big) \leq d'\big(f(z), f(x)\big) + d'\big(f(x), f(z')\big) \leq 2C$$

für alle $z, z' \in M$. Damit erhalten wir dann

$$\begin{aligned} 0 \leq d(x,y) \leq & d(x, x_{n_{k_l}}) + d(x_{n_{k_l}}, y_{n_{k_l}}) + d(y_{n_{k_l}}, y) \\ \leq & d(x, x_{n_{k_l}}) + \tfrac{1}{n_{k_l}} \cdot d'\big(f(x_{n_{k_l}}), f(y_{n_{k_l}})\big) + d(y_{n_{k_l}}, y) \\ \leq & d(x, x_{n_{k_l}}) + \tfrac{2C}{n_{k_l}} + d(y_{n_{k_l}}, y) \longrightarrow 0 + 0 + 0 = 0 \end{aligned}$$

für $l \to \infty$, woraus $x = y$ folgt.

Da f lokal Lipschitz-stetig ist, gibt es ein $\varepsilon > 0$ und eine Lipschitz-Konstante L mit

$$d'(f(z), f(z')) \leq q \cdot d(z, z')$$

für alle $z, z' \in U_\varepsilon(x)$. Für $l \in \mathbb{N}$ sehr groß gilt nun $x_{n_{k_l}}, y_{n_{k_l}} \in U_\varepsilon(x)$ und zudem $n_{k_l} > q$, womit wir dann den Widerspruch

$$q \cdot d(x_{n_{k_l}}, y_{n_{k_l}}) \leq n_{k_l} \cdot d(x_{n_{k_l}}, y_{n_{k_l}}) < d'\big(f(x_{n_{k_l}}), f(y_{n_{k_l}})\big)$$

erhalten. □

I) Konvergenz von Folgen von Abbildungen

Definition 23.39 (Punktweise und gleichmäßige Konvergenz)

a. Für jedes $n \in \mathbb{N}$ sei $f_n : M \longrightarrow M'$ eine Abbildung, so nennen wir $(f_n)_{n\in\mathbb{N}}$ eine *Folge von Abbildungen* auf M.

b. Wir nennen die Folge $(f_n)_{n\in\mathbb{N}}$ von Abbildungen *punktweise konvergent auf* M, wenn für jedes $x \in M$ der Grenzwert $\lim\limits_{n\to\infty} f_n(x)$ existiert. In diesem Fall nennen wir die Funktion

$$f : M \longrightarrow M' : x \mapsto \lim_{n\to\infty} f_n(x)$$

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

$$f = \lim_{n\to\infty} f_n.$$

Man beachte, $(f_n)_{n\in\mathbb{N}}$ konvergiert auf M genau dann punktweise gegen f, wenn

$$\forall\, x \in M\ \forall\, \varepsilon > 0\ \ \exists\, n_{\varepsilon,x}\ :\ \forall\, n \geq n_{\varepsilon,x} \text{ gilt } d'\big(f_n(x), f(x)\big) < \varepsilon.$$

c. Wir sagen $(f_n)_{n\in\mathbb{N}}$ *konvergiert gleichmäßig auf* M gegen f, wenn

$$\forall\, \varepsilon > 0\ \ \exists\, n_\varepsilon\ :\ \forall\, n \geq n_\varepsilon \text{ und } \forall\, x \in M \text{ gilt } d'\big(f_n(x), f(x)\big) < \varepsilon.$$

Bemerkung 23.40
Konvergiert $(f_n)_{n\in\mathbb{N}}$ auf M gleichmäßig gegen f, so konvergiert die Folge auch punktweise gegen f.

Satz 23.41 (Der gleichmäßige Grenzwert stetiger Funktionen ist stetig.)
Ist $f_n : M \longrightarrow M'$ 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 $a \in M$ und $\varepsilon > 0$ gegeben. Da die f_n gleichmäßig gegen f konvergieren, gilt:

$$\exists\, n_\varepsilon \in \mathbb{N} \;:\; \forall\, n \geq n_\varepsilon \text{ und } \forall\, x \in M \;:\; d'\big(f_n(x), f(x)\big) < \frac{\varepsilon}{3}.$$

Da zudem f_{n_ε} stetig in a ist, gilt:

$$\exists\, \delta_\varepsilon > 0 \quad \forall\, x \in M \text{ mit } d(x,a) < \delta_\varepsilon \text{ gilt } d'\big(f_{n_\varepsilon}(x), f_{n_\varepsilon}(a)\big) < \frac{\varepsilon}{3}.$$

Sei nun $x \in M$ mit $d(x,a) < \delta_\varepsilon$ gegeben, so gilt

$$d'\big(f(x), f(a)\big) \leq d'\big(f(x), f_{n_\varepsilon}(x)\big) + d'\big(f_{n_\varepsilon}(x), f_{n_\varepsilon}(a)\big) + d'\big(f_{n_\varepsilon}(a), f(a)\big) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Mithin ist f stetig in a. □

Bemerkung 23.42 (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 vermittels der Maximumsnorm

$$||f||_\infty = \max\big\{|f(x)| \;\big|\; x \in M\big\},$$

der sogar vollständig ist, wie man leicht nachprüft. Aus der Definition der gleichmäßigen Konvergenz folgt, 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.

Beweis: Wir müssen zeigen, daß jede Cauchy-Folge $(f_n)_{n\in\mathbb{N}}$ in $(\mathcal{C}(M,\mathbb{R}), ||\cdot||_\infty)$ sogar eine konvergente Folge ist. Sei dazu zunächst $x \in M$ gegeben. Aus der Voraussetzung folgt, daß es zu jedem $\varepsilon > 0$ ein $n_\varepsilon \in \mathbb{N}$ gibt, so daß für alle $n, m \geq n_\varepsilon$

$$||f_n - f_m||_\infty < \varepsilon$$

gilt und mithin auch

$$|f_n(x) - f_m(x)| \leq ||f_n - f_m||_\infty < \varepsilon. \tag{68}$$

Damit ist aber $(f_n(x))_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$, und da $\mathbb{R}$ vollständig ist, ist diese Folge konvergent. Damit ist die Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ punktweise konvergent und wir setzen

$$f := \lim_{n\to\infty} f_n.$$

Lassen wir in Gleichung (68) m gegen unendlich laufen, so erhalten wir

$$|f_n(x) - f(x)| \leq \varepsilon$$

für alle $x \in M$ und für alle $n \geq n_\varepsilon$. Daraus folgt, daß die Folge $(f_n)_{n\in\mathbb{N}}$ sogar gleichmäßig gegen f konvergiert und damit auch in der Maximumsnorm. □

Bemerkung 23.43 (Die Peano-Kurve)
Stetige Abbildungen haben viele sehr schöne topologische Eigenschaften wie die Verträglichkeit mit offenen und abgeschlossenen Mengen unter Urbildern in Proposition 23.17 oder mit kompakten Mengen in Satz 23.23. 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.

J) Der Satz von Arzelà-Ascoli

Der Begriff der Gleichstetigkeit ist eine Verschärfung der gleichmäßigen Stetigkeit. Wir betrachten dabei eine Menge von gleichmäßig stetigen Funktionen und fordern, dass bei gegebenem $\varepsilon > 0$ das zugehörige δ_ε gleichermaßen für alle Funktionen in der Menge funktionieren soll.

Definition 23.44 (Gleichstetig)
Eine Menge $K \subseteq \mathcal{C}(M, \mathbb{R})$ stetiger Funktionen heißt *gleichstetig* oder *gleichgradig stetig* auf M, wenn

$$\forall\, \varepsilon > 0\ \exists\, \delta_\varepsilon > 0\ \forall\, f \in K\ \wedge\ \forall\, x, y \in M \text{ mit } d(x,y) < \delta_\varepsilon \text{ gilt } |f(x) - f(y)| < \varepsilon.$$

Die Funktionen in K sind also gleichmäßig stetig und für jede der Funktionen f in K funktioniert zu jedem ε sogar dasselbe δ_ε.

Wir haben in Bemerkung 23.42 festgestellt, daß $\mathcal{C}(M, \mathbb{R})$ für ein kompaktes M bezüglich der Maximumsnorm ein normierter Raum ist. Wir wollen nun ein Kriterium dafür angeben, wann eine Teilmenge von $\mathcal{C}(M, \mathbb{R})$ kompakt ist, vergleichbar dem Satz von Heine-Borel.

Satz 23.45 (Satz von Arzelà-Ascoli)
Es sei M kompakt und $K \subseteq \mathcal{C}(M, \mathbb{R})$ eine Menge stetiger Funktionen auf M. Dann sind die folgenden Aussagen gleichwertig:

a. K ist kompakt bezüglich der Maximumsnorm $||\cdot||_\infty$.

b. K ist abgeschlossen in $\mathcal{C}(M, \mathbb{R})$, beschränkt und gleichstetig auf M.

Beweis: Wir wollen zunächst voraussetzen, daß K kompakt ist. Aus Satz 22.21 wissen wir, daß K dann auch abgeschlossen und beschränkt ist. Es bleibt also die Gleichstetigkeit von K auf M zu zeigen. Sei dazu $\varepsilon > 0$ gegeben.

Wir erhalten dann eine Überdeckung

$$K \subseteq \bigcup_{f \in K} U_{\frac{\varepsilon}{3}}(f)$$

von K in $\mathcal{C}(M, \mathbb{R})$, und da K als kompakt vorausgesetzt wird, gibt es dazu eine endliche Teilüberdeckung

$$K \subseteq \bigcup_{i=1}^{k} U_{\frac{\varepsilon}{3}}(f_i) \tag{69}$$

für geeignete Funktionen $f_1, \ldots, f_k \in K$. Da auch M kompakt ist, ist f_i nach Satz 23.29 sogar gleichmäßig stetig auf M und mithin gibt es ein $\delta_i > 0$, so daß für alle $x, y \in M$ mit $d(x,y) < \delta_i$ schon

$$|f_i(x) - f_i(y)| < \frac{\varepsilon}{3} \tag{70}$$

gilt. Wir setzen nun $\delta := \min\{\delta_1, \ldots, \delta_k\} > 0$ und betrachten $f \in K$ beliebig sowie $x, y \in M$ mit $d(x,y) < \delta$. Wegen Gleichung (69) gibt es dann ein i mit $f \in U_{\frac{\varepsilon}{3}}(f_i)$ und

wir erhalten

$$\begin{aligned}|f(x)-f(y)| \leq & |f(x)-f_i(x)|+|f_i(x)-f_i(y)|+|f_i(y)-f(y)| \\ \leq & ||f-f_i||_\infty+|f_i(x)-f_i(y)|+||f_i-f||_\infty \overset{(70)}{<} \frac{\varepsilon}{3}+\frac{\varepsilon}{3}+\frac{\varepsilon}{3}=\varepsilon.\end{aligned}$$

Damit ist die Gleichstetigkeit von K auf M gezeigt.

Nun wollen wir umgekehrt voraussetzen, daß K abgeschlossen in $\mathcal{C}(M,\mathbb{R})$, beschränkt und gleichstetig auf M ist, und wir wollen daraus die Kompaktheit von K folgern. Sei dazu $(f_n)_{n\in\mathbb{N}}$ eine Folge in K. Wegen des Satzes von Bolzano-Weierstrass 22.30 ist zu zeigen, daß $(f_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge mit Grenzwert in K besitzt.

Da K beschränkt und gleichstetig auf M ist, erfüllt $(f_n)_{n\in\mathbb{N}}$ die Voraussetzungen von Lemma 23.46 und besitzt somit eine konvergente Teilfolge mit Grenzwert in $\mathcal{C}(M,\mathbb{R})$. Nun ist die Teilfolge aber in K enthalten und K ist abgeschlossen in $\mathcal{C}(M,\mathbb{R})$, weshalb der Grenzwert nach Korollar 22.29 schon in K liegt und K kompakt ist. □

Den technischen Teil der Rückrichtung im obigen Satz von Arzelà-Ascoli haben wir als eigenes Lemma ausgegliedert, das seinerseits als Satz von Arzelà-Ascoli bezeichnet wird.

Lemma 23.46 (Satz von Arzelà-Ascoli)
Sei M kompakt und $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen in $\mathcal{C}(M,\mathbb{R})$. Wenn die Folgen $(f_n(x))_{n\in\mathbb{N}}$ für alle $x\in M$ beschränkt sind und die Menge $\{f_n \mid n\in\mathbb{N}\}$ gleichstetig auf M ist, dann besitzt $(f_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge mit Grenzwert in $\mathcal{C}(M,\mathbb{R})$.

Beweis: Nach Voraussetzung ist M kompakt und aus Aufgabe 22.63 erhalten wir eine abzählbare Teilmenge $A=\{x_n \mid n\in\mathbb{N}\}$, die dicht in M liegt. Diese wollen wir nun verwenden, um eine geeignete Teilfolge von $(f_n)_{n\in\mathbb{N}}$ zu konstruieren.

Da nach Voraussetzung die Folge $(f_n(x_0))_{n\in\mathbb{N}}$ beschränkt ist, liefert der Satz von Bolzano-Weierstrass 11.26 uns eine konvergente Teilfolge $(f_{n_k}(x_0))_{k\in\mathbb{N}}$. Wir setzen nun

$$g_{0,k} := f_{n_k}$$

und erhalten somit eine Teilfolge $(g_{0,k})_{k\in\mathbb{N}}$ von $(f_n)_{n\in\mathbb{N}}$ mit der Eigenschaft, daß der Grenzwert

$$\lim_{k\to\infty} g_{0,k}(x_0)\in\mathbb{R}$$

existiert. Wir betrachten dann die Folge $(g_{0,n}(x_1))_{n\in\mathbb{N}}$, die als Teilfolge von $(f_n(x_1))_{n\in\mathbb{N}}$ beschränkt ist und als beschränkte Folge wieder eine konvergente Teilfolge $(g_{0,n_k}(x_1))_{k\in\mathbb{N}}$ besitzt, und setzen

$$g_{1,k} := g_{0,n_k}.$$

Die Folge $(g_{1,k})_{k\in\mathbb{N}}$ ist damit eine Teilfolge von $(g_{0,n})_{n\in\mathbb{N}}$ und somit auch von $(f_n)_{n\in\mathbb{N}}$ mit der Eigenschaft, daß die Grenzwerte

$$\lim_{k\to\infty} g_{1,k}(x_0), \lim_{k\to\infty} g_{1,k}(x_1) \in \mathbb{R}$$

existieren. Wir fahren nun so fort und bilden damit eine Folge von Teilfolgen $(g_{n,k})_{k\in\mathbb{N}}$ von $(f_n)_{n\in\mathbb{N}}$ mit der Eigenschaft, daß für ein festes $n \in \mathbb{N}$ die Grenzwerte

$$\lim_{k\to\infty} g_{n,k}(x_l) \in \mathbb{R}$$

für $l = 0, \ldots, n$ existieren. Denken wir uns die Folgen wie folgt aufgelistet

$$\begin{array}{ccccc} g_{0,0} & g_{0,1} & g_{0,2} & g_{0,3} & \cdots \\ g_{1,0} & g_{1,1} & g_{1,2} & g_{1,3} & \cdots \\ g_{2,0} & g_{2,1} & g_{2,2} & g_{2,3} & \cdots \\ g_{3,0} & g_{3,1} & g_{3,2} & g_{3,3} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots, \end{array}$$

so können wir die Diagonalfolge $(h_n)_{n\in\mathbb{N}}$ mit den Folgengliedern

$$h_n := g_{n,n}$$

betrachten. Nach Konstruktion ist diese eine Teilfolge von $(f_n)_{n\in\mathbb{N}}$ und für alle $k \in \mathbb{N}$ existiert der Grenzwert

$$\lim_{n\to\infty} h_n(x_k) \in \mathbb{R},$$

weil $(h_n(x_k))_{n\geq k}$ eine Teilfolge von $(g_{k,n}(x_k))_{n\in\mathbb{N}}$ ist.

Wir wollen nun zeigen, daß die Folge $(h_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathcal{C}(M, \mathbb{R})$ ist. Sei dazu $\varepsilon > 0$ gegeben.

Da $\{f_n \mid n \in \mathbb{N}\}$ gleichstetig auf M ist, gibt es ein $\delta > 0$, so daß für alle $h \in \{f_n \mid n \in \mathbb{N}\}$ und für alle $x, y \in M$ mit $d(x, y) < \delta$ schon

$$|h(x) - h(y)| < \frac{\varepsilon}{3} \tag{71}$$

gilt. Da M kompakt ist, besitzt die offene Überdeckung

$$M = \bigcup_{x\in M} U_{\frac{\delta}{2}}(x)$$

eine endliche Teilüberdeckung

$$M = \bigcup_{i=1}^{s} U_{\frac{\delta}{2}}(y_i)$$

für geeignete $y_1, \ldots, y_s \in M$.

Da A dicht in M liegt, gibt es für jedes $i = 1, \ldots, s$ ein $k_i \in \mathbb{N}$ mit

$$x_{k_i} \in U_{\frac{\delta}{2}}(y_i). \tag{72}$$

Nach Konstruktion ist jede der endliche vielen Folgen $(h_n(x_{k_i}))_{n\in\mathbb{N}}$ konvergent und damit eine Cauchy-Folge. Für $i \in \{1,\dots,s\}$ gibt es also ein n_i, so daß für alle $n, m \geq n_i$ schon

$$|h_n(x_{k_i}) - h_m(x_{k_i})| < \frac{\varepsilon}{3} \tag{73}$$

gilt, und setzen wir $n_\varepsilon = \max\{n_1,\dots,n_s\}$, dann gilt dies für alle $n, m \geq n_\varepsilon$ und alle $i = 1,\dots,s$ gleichzeitig.

Sei nun $x \in M$ beliebig gegeben. Dann gibt es ein i, so daß

$$x \in U_{\frac{\delta}{2}}(y_i)$$

gilt. Mit (72) folgt dann aber

$$d(x, x_{k_i}) \leq d(x, y_i) + d(y_i, x_{k_i}) < \frac{\delta}{2} + \frac{\delta}{2} = \delta. \tag{74}$$

Wegen (71) zusammen mit (73) sowie (74) erhalten wir damit

$$\begin{aligned} |h_n(x) - h_m(x)| \leq & |h_n(x) - h_n(x_{k_i})| + |h_n(x_{k_i}) - h_m(x_{k_i})| + |h_m(x_{k_i}) - h_m(x)| \\ < & \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

für alle $m, n \geq n_\varepsilon$. Da die Ungleichung für ein beliebiges $x \in M$ gilt und n_ε nicht von x abhängt, folgt für alle $m, n \geq n_\varepsilon$ schon

$$||h_n - h_m||_\infty = \max\left\{|h_n(x) - h_m(x)| \;\middle|\; x \in M\right\} < \varepsilon$$

und somit ist $(h_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathcal{C}(M,\mathbb{R})$.

Da $\mathcal{C}(M,\mathbb{R})$ nach Bemerkung 23.42 vollständig ist, ist $(h_n)_{n\in\mathbb{N}}$ dann aber auch konvergent in $\mathcal{C}(M,\mathbb{R})$. □

Bemerkung 23.47 (Satz von Arzelà-Ascoli)
Es ist leicht, zu sehen, daß jede gleichmäßig konvergente Folge stetiger Funkionen auf einer kompakten Menge M schon gleichstetig ist. Der Satz von Arzelà-Ascoli in der Form von Lemma 23.46 liefert eine teilweise Umkehrung, indem er zeigt, daß eine gleichstetige Folge auf einer kompakten Menge, die punktweise beschränkt ist, zumindest eine gleichmäßig konvergente Teilfolge besitzt.

K) Stetige lineare Operatoren

Stetigkeit ist eine lokale Eigenschaft und gemeinhin kann eine Abbildung in manchen Punkten stetig und in anderen unstetig sein. Für lineare Abbildungen geht das nicht. Im endlich-dimensionalen Fall sind sie stets stetig, im unendlich-dimensionalen Fall folgt auch der Stetigkeit in einem Punkt schon die globale Lipschitz-Stetigkeit.

Satz 23.48 (Stetige lineare Operatoren)
Für eine lineare Abbildung $f : V \longrightarrow W$ sind die folgenden Aussagen gleichwertig:

a. f ist Lipschitz-stetig auf V.

b. f ist gleichmäßig stetig auf V.

c. f ist stetig auf V.

d. f ist stetig in 0.

e. $\exists\, r > 0 \;:\; \forall\, x \in V \;:\; ||f(x)||_W \leq r \cdot ||x||_V$.

Man nennt f dann auch einen stetigen oder beschränkten *linearen Operator*.

Beweis:

a. $\Longrightarrow$ b.: Klar, siehe Proposition 23.33.
b. $\Longrightarrow$ c.: Klar, siehe Bemerkung 23.28.
c. $\Longrightarrow$ d.: Klar.
d. $\Longrightarrow$ e.: Ist f stetig in 0, so gibt es für $\varepsilon = 1$ ein $\delta_\varepsilon > 0$ mit

$$||f(x)||_W = ||f(x) - f(0)||_W < 1 \tag{75}$$

für alle $x \in V$ mit $||x||_V = ||x - 0||_V < \delta_\varepsilon$. Wir setzen $r = \frac{2}{\delta_\varepsilon} > 0$. Für $0 \neq x \in V$ beliebig ist die Norm von $\frac{\delta_\varepsilon \cdot x}{2 \cdot ||x||_V}$ kleiner als δ_ε und somit gilt

$$||f(x)||_W = \left|\left| f\left(\tfrac{2\cdot||x||_V}{\delta_\varepsilon} \cdot \tfrac{\delta_\varepsilon \cdot x}{2\cdot||x||_V}\right)\right|\right|_W = \left|\left| \tfrac{2\cdot||x||_V}{\delta_\varepsilon} \cdot f\left(\tfrac{\delta_\varepsilon \cdot x}{2\cdot||x||_V}\right)\right|\right|_W$$

$$= \tfrac{2\cdot||x||_V}{\delta_\varepsilon} \cdot \left|\left| f\left(\tfrac{\delta_\varepsilon \cdot x}{2\cdot||x||_V}\right)\right|\right|_W \overset{(75)}{<} \tfrac{2\cdot||x||_V}{\delta_\varepsilon} \cdot 1 = r \cdot ||x||_V.$$

Für $x = 0$ gilt ohnehin $||f(x)||_W = 0 = r \cdot ||x||_V$.
e. $\Longrightarrow$ 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 23.49 (Der Integraloperator als stetiger Operator)
Wir betrachten den Vektorraum $V = \mathcal{C}\big([0,1], \mathbb{R}\big)$ mit der euklidischen L_2-Norm aus Beispiel 22.5 sowie die lineare Abbildung

$$I : V \longrightarrow \mathbb{R} : f \mapsto \int_0^1 f(x)\, dx.$$

Für $f \in V$ gilt dann

$$|I(f)| = \left|\int_0^1 f(x)\, dx\right| = \left|\int_0^1 \big(1 \cdot f(x)\big)\, dx\right| \overset{A14.8}{\leq} ||1||_{L_2} \cdot ||f||_{L_2},$$

wobei die letzte Ungleichung von der Ungleichung von Cauchy-Schwartz A14.8 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.

L) Die Operatornorm

Sind V und W normierte Räume, so ist der Vektorraum $L(V,W)$ der linearen Operatoren von V nach W in kanonischer Weise selbst ein normierter Raum.

Satz 23.50 (Operatornorm)
$L(V,W) = \{f : V \longrightarrow W \mid f \text{ ist linear und stetig}\}$ ist ein normierter Raum vermittels

$$||\cdot|| : L(V,W) \longrightarrow \mathbb{R}_{\geq 0} : f \mapsto ||f|| := \sup_{0 \neq x \in V} \frac{||f(x)||_W}{||x||_V}.$$

Wir nennen $||f||$ die *Operatornorm* von f.

Beweis: Nach Proposition 23.13 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 Aussagen des folgenden Korollars folgen aus der Definition der Operatornorm.

Korollar 23.51 (Submultiplikativität der Operatornorm)

a. Für $f \in L(V,W)$ und $x \in V$ gilt $||f(x)||_W \leq ||f|| \cdot ||x||_V$.

b. Für $f, g \in L(V,V)$ gilt zudem $||f \circ g|| \leq ||f|| \cdot ||g||$.

Beweis: Die Aussage in a. folgt unmittelbar aus der Definition der Operatornorm. Für die Aussage in b. betrachten wir einen Vektor $x \in V$ und erhalten aus a. dann

$$||(f \circ g)(x)||_V = ||f(g(x))||_V \leq ||f|| \cdot ||g(x)||_V \leq ||f|| \cdot ||g|| \cdot ||x||_V$$

und damit

$$||f \circ g|| = \sup_{0 \neq x \in V} \frac{||(f \circ g)(x)||}{||x||} \leq \sup_{0 \neq x \in V} \frac{||f|| \cdot ||g|| \cdot ||x||}{||x||} = ||f|| \cdot ||g||.$$

□

Beispiel 23.52 (Operatornorm des Integraloperators)

In Beispiel 23.49 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 23.53

Da alle linearen Abbildungen von $\mathbb{R}^n$ nach $\mathbb{R}^m$ von der Form f_A für eine Matrix $A \in \mathrm{Mat}(m \times n, \mathbb{R})$ sind und da diese alle stetig sind, wissen wir

$$L\big(\mathbb{R}^n, \mathbb{R}^m\big) = \mathrm{Hom}_{\mathbb{R}}\big(\mathbb{R}^n, \mathbb{R}^m\big).$$

Bemerkung 23.54 (Operatornorm auf $\mathrm{Mat}(m \times n, \mathbb{R})$)

Definieren wir die Norm $||A|| := ||f_A||$ einer Matrix $A \in \mathrm{Mat}(m \times n, \mathbb{R})$ 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 $\mathrm{Mat}(m \times n, \mathbb{R})$ und diese ist submultiplikativ im Sinne von Korollar 23.51.

M) $L(V,W)$ als Banachraum

Abschließend beschäftigen wir uns mit der Frage, wann der Raum $L(V,W)$ mit der Operatornorm ein Banachraum wird.

Satz 23.55 ($L(V,W)$ als Banachraum)
Ist W ein Banachraum, so ist $L(V,W)$ mit der Operatornorm $||\cdot||$ ein Banachraum.

Beweis: Aus Satz 23.50 wissen wir, daß $L(V,W)$ mit der Operatornorm ein normierter Raum ist. Sei nun $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $L(V,W)$.

Wir zeigen zunächst, daß dann $(f_n(x))_{n\in\mathbb{N}}$ für jedes $x \in V$ eine Cauchy-Folge in W und somit konvergent ist. Sei dazu $x \in V$ gegeben. Dann existiert ein $n_\varepsilon > 0$, so daß

$$||f_n - f_m|| < \frac{\varepsilon}{||x||_V}$$

für alle $n > m \geq n_\varepsilon$. Mithin folgt nach Korollar 23.51 aus der Submultiplikativität der Operatornorm

$$||f_n(x) - f_m(x)||_W = ||(f_n - f_m)(x)||_W \leq ||f_n - f_m|| \cdot ||x||_V < \varepsilon$$

für alle $n > m \geq n_\varepsilon$. Also ist $(f_n(x))_{n\in\mathbb{N}}$ eine Cauchy-Folge im Banachraum W und somit konvergent.

Wir definieren nun eine Abbildung

$$f : V \longrightarrow W : x \mapsto \lim_{n\to\infty} f_n(x).$$

Für $x, y \in V$ und $\lambda, \mu \in \mathbb{R}$ folgt aus der Eindeutigkeit des Grenzwertes, den Grenzwertsätzen und weil die f_n linear sind

$$\begin{aligned} f(\lambda \cdot x + \mu \cdot y) &= \lim_{n\to\infty} f_n(\lambda \cdot x + \mu \cdot y) = \lim_{n\to\infty} \lambda \cdot f_n(x) + \mu \cdot f_n(y) \\ &= \lambda \cdot \lim_{n\to\infty} f_n(x) + \mu \cdot \lim_{n\to\infty} f_n(y) = \lambda \cdot f(x) + \mu \cdot f(y). \end{aligned}$$

Also ist f eine lineare Abbildung.

Wir wollen nun noch zeigen, daß f auch beschränkt und damit stetig ist. Da $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist, ist $(f_n)_{n\in\mathbb{N}}$ nach Proposition 22.23 auch beschränkt. Es gibt also eine Konstante $C > 0$ mit

$$||f_n|| \leq C$$

für alle $n \in \mathbb{N}$. Sei nun $x \in V$ mit $||x|| = 1$. Dann konvergiert $f_n(x)$ gegen $f(x)$ und mithin gibt es ein $n \in \mathbb{N}$ mit

$$||f(x) - f_n(x)||_W < 1.$$

Dann gilt aber

$$||f(x)||_W \leq ||f(x) - f_n(x)||_W + ||f_n(x)||_W < 1 + ||f_n|| \cdot ||x||_V \leq 1 + C,$$

und somit

$$||f|| = \sup_{0\neq x\in V} \frac{||f(x)||_W}{||x||_V} = \sup_{0\neq x\in V} \left|\left| f\left(\frac{x}{||x||_V}\right)\right|\right|_W \leq 1 + C.$$

Nach Satz 23.48 ist f dann stetig, also $f \in L(V,W)$.

Als nächstes wollen wir

$$\lim_{n\to\infty} ||f_n - f|| = 0 \tag{76}$$

zeigen, woraus dann folgt, daß $L(V,W)$ ein Banachraum ist. Sei dazu $\varepsilon > 0$ gegeben und sei $n_\varepsilon > 0$ so, daß

$$||f_n - f_m|| < \varepsilon$$

für alle $n > m \geq n_\varepsilon$. Für $0 \neq x \in V$ gilt dann

$$||(f_n - f_m)(x)||_W \leq ||f_n - f_m|| \cdot ||x||_V < \varepsilon \cdot ||x||_V,$$

und damit folgt

$$||f_n(x) - f(x)||_W = \lim_{m\to\infty} ||(f_n - f_m)(x)||_W \leq \lim_{m\to\infty} ||f_n - f_m|| \cdot ||x||_V \leq \varepsilon \cdot ||x||_V.$$

Wir leiten daraus

$$\frac{||(f_n - f)(x)||_W}{||x||_V} \leq \varepsilon$$

für alle $n \geq n_\varepsilon$ ab oder alternativ

$$||f_n - f|| = \sup_{0\neq x\in V} \frac{||(f_n - f)(x)||_W}{||x||_V} \leq \varepsilon$$

für alle $n \geq n_\varepsilon$. Damit ist (76) gezeigt. □

N) Die Neummansche Reihe

Proposition 23.56 (Neumannsche Reihe)
Sei V ein Banachraum und $f \in L(V,V)$, so daß die Reihe $\sum_{k=0}^{\infty} ||f^k||$ konvergent ist, wobei $||\cdot||$ die Operatornorm auf $L(V,V)$ bezeichnet.
Dann ist $\mathrm{id}_V - f$ bijektiv mit

$$(\mathrm{id}_V - f)^{-1} = \sum_{k=0}^{\infty} f^k \in L(V,V),$$

und wenn $||f|| < 1$ gilt, dann ist zudem

$$||(\mathrm{id}_V - f)^{-1}|| \leq \frac{1}{1 - ||f||}.$$

Beweis: Aus Satz 23.55 wissen wir, daß $L(V,V)$ mit der Operatornorm ein Banachraum ist, und Proposition 22.51 liefert uns dann, daß die absolut konvergente Reihe

$$g = \sum_{k=0}^{\infty} f^k$$

in $L(V,V)$ konvergent ist. Aus der absoluten Konvergenz der Reihe erhalten wir

$$\lim_{k\to\infty} ||f^k|| = 0$$

und damit bezüglich der Operatornorm

$$\lim_{k\to\infty} f^k = 0 \in L(V,V).$$

Beachten wir nun, daß die Komposition mit der stetigen Abbildung $\mathrm{id}_V - f$ stetig ist, so erhalten wir

$$(\mathrm{id}_V - f) \circ \sum_{k=0}^{n} f^k = \mathrm{id}_V - f^{n+1}$$

und damit

$$(\mathrm{id}_V - f) \circ g = (\mathrm{id}_V - f) \circ \lim_{n\to\infty} \sum_{k=0}^{n} f^k = \mathrm{id}_V - \lim_{n\to\infty} f^{n+1} = \mathrm{id}_V .$$

Analog zeigt man

$$g \circ (\mathrm{id}_V - f) = \mathrm{id}_V,$$

so daß $\mathrm{id}_V - f$ bijektiv mit $(\mathrm{id}_V - f)^{-1} = g$ ist.

Es bleibt, die Abschätzung zu zeigen. Unter Beachtung der Submultiplikativität der Operatornorm 23.51 und der geometrischen Reihe folgt diese unmittelbar aus der Ungleichung

$$\left|\left| \sum_{k=0}^{n} f^k \right|\right| \le \sum_{k=0}^{n} ||f^k|| \le \sum_{k=0}^{n} ||f||^k$$

durch Grenzwertbildung

$$||(\mathrm{id}_V - f)^{-1}|| = \lim_{n\to\infty} \left|\left| \sum_{k=0}^{n} f^k \right|\right| \le \lim_{n\to\infty} \sum_{k=0}^{n} ||f||^k = \sum_{k=0}^{\infty} ||f||^k = \frac{1}{1-||f||}.$$

□

Aufgaben

Aufgabe 23.57

Es sei $a \in M$ und $f : M \longrightarrow \mathbb{R}$ sei stetig in a mit $f(a) \neq 0$. Zeige, es gibt ein $\delta > 0$ mit $f(x) \neq 0$ für alle $x \in U_\delta(a)$.

Aufgabe 23.58
Es sei $M := \{(x,y)^t \in \mathbb{R}^2 \mid x > y\}$ und $f : M \longrightarrow \mathbb{R}$ sei definiert durch

$$f(x,y) = \arctan\left(\frac{xy}{x-y}\right)$$

Für welche $a \in \mathbb{R}$ existiert der Grenzwert $\lim_{(x,y)\to(a,a)} f(x,y)$?

Aufgabe 23.59
Sei M ein metrischer Raum und $f : M \longrightarrow \mathbb{R}$ 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. $A \subseteq M$ ist abgeschlossen und $f(A) \subseteq \mathbb{R}$ ist *nicht* abgeschlossen.

b. $O \subseteq M$ ist offen und $f(O) \subseteq \mathbb{R}$ ist *nicht* offen.

c. $K \subseteq \mathbb{R}$ ist kompakt und $f^{-1}(K) \subseteq M$ ist *nicht* kompakt.

Aufgabe 23.60 (Topologische Isomorphismen)
Zeige, für eine bijektive lineare Abbildung $f : V \longrightarrow W$ sind gleichwertig:

a. f und f^{-1} sind stetig.

b. $\exists\, r, s > 0 \;:\; \forall\, x \in V \;:\; s \cdot ||x||_V \leq ||f(x)||_W \leq r \cdot ||x||_V$.

Man nennt den Isomorphismus f dann einen *topologischen Isomorphismus*

Aufgabe 23.61
Zeige, ist $\dim_{\mathbb{R}}(V) = n$ und betrachten wir $\mathbb{R}^n$ mit der euklidischen Norm, so ist jeder Isomorphismus $f : \mathbb{R}^n \longrightarrow V$ ein topologischer Isomorphismus.

Aufgabe 23.62 (Endlich-dimensionale Banachräume)
Zeige, jeder endlich-dimensionale normierte $\mathbb{R}$-Vektorraum ist ein Banachraum.

Aufgabe 23.63
Zeige, auf einem endlich-dimensionalen $\mathbb{R}$-Vektorraum sind alle Normen äquivalent.

Aufgabe 23.64
Sind V und W endlich-dimensional, so gilt $L(V,W) = \mathrm{Hom}_{\mathbb{R}}(V,W)$.

Aufgabe 23.65 (Der Differentialoperator als nicht stetiger Operator)
Es sei $V = \mathcal{C}^1\big([0,1],\mathbb{R}\big)$ der Vektorraum der stetig-differenzierbaren Funktionen auf $[0,1]$ und $W = \mathcal{C}\big([0,1],\mathbb{R}\big)$. Aus Aufgabe 22.55 wissen wir, daß W ein normierter Raum bezüglich der Maximumsnorm aus Beispiel 22.5 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 23.66
Wir betrachten den Vektorraum $V = \mathcal{C}\big([0,1],\mathbb{R}\big)$ als normierten Raum mit der L^2-Norm

$$||f|| = \sqrt{\int_0^1 f^2(x)\,dx}$$

und die Auswertungsabbildung

$$P : V \longrightarrow \mathbb{R} : f \mapsto f(0).$$

Zeige, P ist ein nicht-stetiger lineare Operator.

Aufgabe 23.67
Zeige, der Integraloperator $I : \mathcal{C}([0,1],\mathbb{R}) \longrightarrow \mathbb{R} : f \mapsto \int_0^1 f(x)\,dx$ ist bezüglich der Maximumsnorm auf $\mathcal{C}([0,1],\mathbb{R})$ gleichmäßig stetig.

Aufgabe 23.68
Betrachte $V = \mathrm{Mat}_n(\mathbb{R}) \hat{=} \mathbb{R}^{n^2}$ als normierten Raum mit der euklidischen Norm

$$||(a_{ij})||_2 = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}^2},$$

(siehe auch Aufgabe 22.59) und betrachte $U = \{A \in V \mid A \text{ ist symmetrisch}\}$ als metrischen Raum mit der Einschränkung der Metrik von V auf U.

a. Zeige, die Menge $P = \{A \in U \mid A \text{ ist positiv definit}\}$ ist offen in U.

b. Zeige, die Menge $N = \{A \in U \mid A \text{ ist negativ definit}\}$ ist offen in U.

c. Zeige, die Menge $I = \{A \in U \mid A \text{ ist indefinit}\}$ ist offen in U.

Aufgabe 23.69

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 23.70

Zeige, daß die Menge $U = \{(x,y)^t \in \mathbb{R}^2 \mid x < 2y^2 - 1\}$ offen in $\mathbb{R}^2$ ist.

Aufgabe 23.71

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 $(0,0)^t$ stetig?

Aufgabe 23.72

Finde eine Funktion $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$, die in $(0,0)^t$ nicht stetig ist, so daß aber dennoch das Folgenkriterium für Stetigkeit für jede Folge erfüllt ist, die sich dem Nullpunkt entlang einer Parabel der Form $x = ay^2$ für $0 \neq a \in \mathbb{R}$ nähert.

Aufgabe 23.73

Betrachte die Funktion

$$f : \mathbb{R}^2 \setminus \{(0,0)^t\} \longrightarrow \mathbb{R} : (x,y)^t \mapsto \frac{x^2y^2}{x^2y^2 + (x-y)^2}.$$

Zeige, $\lim\limits_{x\to 0}\lim\limits_{y\to 0} f(x,y) = \lim\limits_{y\to 0}\lim\limits_{x\to 0} f(x,y)$. Ist f stetig in $(0,0)^t$ fortsetzbar?

Aufgabe 23.74

Betrachte die Funktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x,y)^t \mapsto \begin{cases} \frac{x^3-x^2y+9xy^2}{5x^2y+7y^3}, & \text{wenn } y \neq 0, \\ 0, & \text{wenn } y = 0. \end{cases}$$

Untersuche, wo die Funktion f stetig ist.

Aufgabe 23.75
Sei $f : (M, d) \longrightarrow (M', d')$ eine gleichmäßig stetige Abbildung metrischer Räume. Zeige, ist $(a_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge, so ist auch $\big(f(a_n)\big)_{n\in\mathbb{N}}$ eine Cauchy-Folge. Gilt die Aussage auch noch, wenn f nur stetig ist?

Aufgabe 23.76
Zeige mit Hilfe einer stetigen Abbildung $f : [0, 1] \longrightarrow \mathbb{R}^n$, daß $\mathbb{R}^n$ zusammenhängend im Sinne von Aufgabe 22.60 ist.

Aufgabe 23.77
Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \longrightarrow \mathbb{R}^n$ sei gleichmäßig stetig und bijektiv.
Zeige, ist f^{-1} stetig, so ist $U = \mathbb{R}^n$.
Gilt die Aussage auch noch, wenn f nur stetig aber nicht gleichmäßig stetig ist?
Hinweis, wenn $U \neq \mathbb{R}^n$, so kann man zeigen, daß U einen Häufungspunkt besitzt und diesen ausnutzen.

Aufgabe 23.78
Zeige, ist $A \in \mathrm{Mat}_n(\mathbb{R})$ symmetrisch, so ist

$$||f_A|| = \max\big\{|\lambda| \;\big|\; \lambda \text{ ist Eigenwert von } A\big\}$$

die Operatornorm von $f_A : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ bezüglich der euklidischen Norm auf $\mathbb{R}^n$.

Aufgabe 23.79 (Arzelà-Ascoli)
Sei (M, d) ein kompakter metrischer Raum und $(f_n)_{n\in\mathbb{N}}$ eine gleichmäßig konvergente Folge stetiger Funktionen. Zeige, daß die Menge $\{f_n \mid n \in \mathbb{N}\}$ dann auch gleichstetig auf M ist.

Aufgabe 23.80 (Abgeschlossen und beschränkt, aber nicht kompakt)
Wir betrachten $\mathcal{C}([0, 1], \mathbb{R})$ mit der Maximumsnorm als normierten Raum.

a. Zeige, die Folge $(f_n)_{n\in\mathbb{N}}$

$$f_n : [0, 1) \longrightarrow \mathbb{R} : x \mapsto x^n$$

besitzt keine konvergente Teilfolge.

b. Zeige, der Abschluss $\overline{U_1(0)}$ der 1-Umgebung der Nullfunktion in $\mathcal{C}([0, 1], \mathbb{R})$ ist nicht gleichstetig auf $[0, 1]$.

Damit ist $\overline{U_1(0)}$ dann abgeschlossen und beschränkt, aber nicht kompakt.

Aufgabe 23.81
Finde eine Abbildung, die gleichmäßig stetig, aber nicht lokal Lipschitz-stetig ist, und eine die lokal Lipschitz-stetig, aber nicht gleichmäßig stetig ist.

Aufgabe 23.82
Sei

$$V = B(\mathbb{R}, \mathbb{R}) = \{f : \mathbb{R} \longrightarrow \mathbb{R} \mid f \text{ ist beschränkt}\}$$

der Vektorraum der beschränkten Funktionen auf $\mathbb{R}$.

a. Zeige, durch

$$||f||_\infty := \sup_{x \in \mathbb{R}} |f(x)|$$

wird eine Norm auf V definiert.

b. Finde eine Teilmenge von V, die abgeschlossen, beschränkt und gleichgradig stetig, aber nicht kompakt ist.

Aufgabe 23.83
Zeige mit Hilfe der Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ mit

$$f_n : [0, 1] \longrightarrow \mathbb{R} : x \mapsto e^{-nx},$$

daß die im Satz von Arzéla-Ascoli (Lemma 23.46) vorausgesetzte gleichgradige Stetigkeit eine notwendige Voraussetzung ist.

Aufgabe 23.84
Finde ein Gegenbeispiel zu Lemma 23.46, wenn man dort die punktweise Beschränktheit wegläßt.

Aufgabe 23.85
Betrachte die Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ mit

$$f_n : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 0, & \text{wenn } x < n, \\ n \cdot x - n^2, & \text{wenn } x \in \left[n, n + \frac{1}{n}\right], \\ 1, & \text{wenn } x > n + \frac{1}{n}, \end{cases}$$

im Vektorraum der beschränkten Funktionen $B(\mathbb{R}, \mathbb{R})$, den wir mit der Supremumsnorm aus Aufgabe 23.82 als normierten Raum betrachten.
Überprüfe, ob die Funktionenfolge beschränkt und gleichstetig ist und ob sie eine konvergente Teilfolge besitzt.

§ 24 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.

A) Totale Differenzierbarkeit

Motivation 24.1

Eine Funktion $f : U \longrightarrow \mathbb{R}$ mit $U \subseteq \mathbb{R}$ haben wir *differenzierbar* im Häufungspunkt a von U genannt, wenn der Grenzwert des Differenzenquotienten

$$\lim_{x\to a} \mathrm{Diff}_{f,a}(x) = \lim_{x\to a} \frac{f(x)-f(a)}{x-a} = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h} \in \mathbb{R}$$

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 $\rho : U \longrightarrow \mathbb{R}$ gibt, so daß

$$f(x) = f(a) + c \cdot (x-a) + \rho(x)$$

und

$$\lim_{x\to a} \frac{\rho(x)}{|x-a|} = \lim_{h\to 0} \frac{\rho(a+h)}{|h|} = 0.$$

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 $n \geq 2$, 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 a, wie schon erwähnt, eine Zahl $c = f'(a)$. 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 a als lineare Abbildung $\mathbb{R}^n \longrightarrow \mathbb{R}^m$ ansehen oder auch einfach als Matrix in $\mathrm{Mat}(m \times n, \mathbb{R})$, 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.

Definition 24.2 (Total differenzierbar)
Eine Abbildung $f : U \longrightarrow \mathbb{R}^m$ heißt *total differenzierbar* in $a \in U$, wenn es eine Matrix $A \in \mathrm{Mat}(m \times n, \mathbb{R})$ und eine Abbildung $\rho : U \longrightarrow \mathbb{R}^m$ gibt mit

$$f(x) = f(a) + A \cdot (x - a) + \rho(x)$$

und

$$\lim_{x \to a} \frac{\rho(x)}{||x - a||_2} = 0.$$

Wir nennen die Matrix $Df(a) := A$ dann die *Ableitung* von f in a.
f heißt *total differenzierbar* auf U, wenn f für jedes $a \in U$ total differenzierbar ist.

Bemerkung 24.3 (Totale Differenzierbarkeit in $\mathbb{R}$)

Aufgrund von Bemerkung 17.6 ist eine Funktion $f : U \longrightarrow \mathbb{R}$ mit $U \subseteq \mathbb{R}$ offen genau dann in $a \in U$ differenzierbar im Sinne von Definition 17.3, wenn f in a total differenzierbar ist.

Lemma 24.4 (Eindeutigkeit der Ableitung)

Ist $f : U \longrightarrow \mathbb{R}^m$ *total differenzierbar* in $a \in U$ und $0 \neq x \in \mathbb{R}^n$, so gilt[2]

$$\lim_{t\to 0} \frac{f(a+tx) - f(a)}{t} = Df(a) \circ x.$$

Insbesondere ist die Ableitung $Df(a)$ von f in a eindeutig bestimmt.

Beweis: Es seien A und ρ wie in Definition 24.2. Da a ein innerer Punkt der offenen Menge U ist, gibt es ein $\varepsilon > 0$, so daß $a + tx \in U$ für alle $t \in U_\varepsilon(0)$. Wir werden im folgenden nur solche t betrachten. Wir erhalten dann

$$\frac{f(a+tx) - f(a)}{t} = \frac{Atx + \rho(a+tx)}{t} = Ax \pm \frac{\rho(a+tx)}{||a+tx-a||_2} \cdot ||x||_2 \xrightarrow{t\to 0} Ax.$$

Das zeigt die erste Aussage des Lemmas. Wenden wir die Formel auf $x = e_j$ an, erhalten wir, daß die j-te Spalte Ae_j von A gerade

$$Ae_j = \lim_{t\to 0} \frac{f(a+te_j) - f(a)}{t}$$

ist. Die rechte Seite hängt aber nur von f und nicht von A ab. Dies zeigt, daß A durch f eindeutig bestimmt ist. □

Beispiel 24.5 (Totale Differenzierbarkeit)

Wir wollen zeigen, daß die Funktion $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x_1, x_2)^t \mapsto 3x_1 + x_1x_2 - 2x_2 + 1$ in $a = (0,0)^t$ total differenzierbar ist. Aus Lemma 24.4 wissen wir, daß die Ableitung $Df(a) = A$ dann eine 1×2-Matrix sein muß, deren erste Spalte

$$\lim_{t\to 0} \frac{f(a+te_1) - f(a)}{t} = \lim_{t\to 0} \frac{3t}{t} = 3$$

und deren zweite Spalte

$$\lim_{t\to 0} \frac{f(a+te_2) - f(a)}{t} = \lim_{t\to 0} \frac{-2t}{t} = -2$$

ist, d.h.

$$A = (3 \quad -2)$$

Die Abbildung ρ muß dann die Abbildungsvorschrift

$$\rho : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x_1, x_2)^t \mapsto f(x_1, x_2) - f(a) - A \circ (x_1, x_2)^t = x_1x_2$$

[2] In der Definition dieses Grenzwertes und aller folgenden betrachten wir nur solche t, für die $a+tx$ in U liegt. Da 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 $t \in U_\varepsilon(0)$ auch $a + tx \in U$ gilt.

haben. Wir erhalten deshalb

$$\lim_{(x_1,x_2)^t\to(0,0)^t} \frac{\rho(x_1,x_2)}{||(x_1,x_2)^t-a||_2} = \lim_{(x_1,x_2)^t\to(0,0)^t} \frac{x_1x_2}{\sqrt{x_1^2+x_2^2}} = 0,$$

wie wir aus Beispiel 23.5 wissen. Also ist f im Ursprung total differenzierbar.

B) Partielle Differenzierbarkeit

Die Beschreibung der Ableitung von f im Punkt a in Lemma 24.4 motiviert die folgende Definition.

Definition 24.6 (Partielle Ableitungen)

a. Für eine Funktion $f : U \longrightarrow \mathbb{R}$, einen Punkt $a \in U$ und $v \in \mathbb{R}^n$ mit $||v||_2 = 1$ heißt der Grenzwert

$$\frac{\partial f}{\partial v}(a) := \lim_{t\to 0} \frac{f(a+tv)-f(a)}{t}$$

die *Richtungsableitung* von f in a in Richtung v, sofern er existiert.

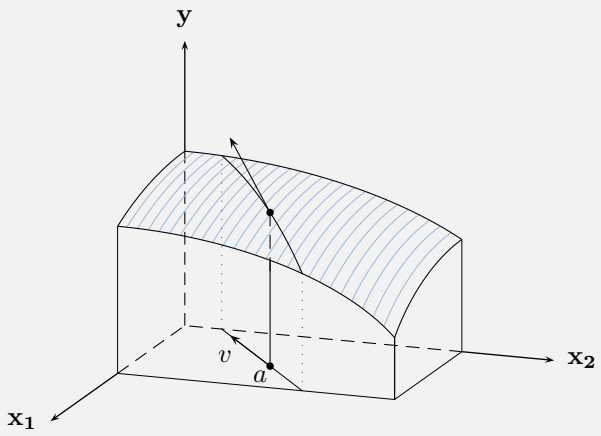


Abbildung 4: Richtungsableitung $\frac{\partial f}{\partial v}(a)$ in Richtung v

Die Richtungsableitung in Richtung der i-ten Koordinate $v = e_i$

$$\lim_{t\to 0} \frac{f(a+te_i)-f(a)}{t}$$

heißt auch die *partielle Ableitung* von f nach der i-ten Variablen in a. Folgende Bezeichnungen sind für diese partielle Ableitung gebräuchlich

$$\frac{\partial f}{\partial x_i}(a) = \frac{\partial f}{\partial e_i}(a) = f_{x_i}(a) = \partial_i f(a) = D_i f(a) = \lim_{t\to 0} \frac{f(a+te_i)-f(a)}{t}.$$

b. Eine Abbildung $f : U \longrightarrow \mathbb{R}^m$ heißt *partiell differenzierbar* in $a \in U$, falls alle partiellen Ableitungen aller Komponentenfunktionen $f_1, \ldots, f_m$ von f in a existieren. Wir nennen die Matrix

$$Jf(a) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \ldots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \ldots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \ldots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}$$

dann die *Jacobi-Matrix* von f in a.
f heißt *partiell differenzierbar* auf U, wenn f in jedem $a \in U$ partiell differenzierbar ist. In diesem Fall sind die partiellen Ableitungen

$$\frac{\partial f_i}{\partial x_j} : U \longrightarrow \mathbb{R}$$

selbst wieder Funktionen auf U.

Bemerkung 24.7 (Richtungsableitung)
Ist $f : U \longrightarrow \mathbb{R}$, $U_\varepsilon(a) \subseteq U$ und $v \in \mathbb{R}^n$ mit $||v||_2 = 1$, so betrachten wir die Funktion

$$g : (-\varepsilon, \varepsilon) \longrightarrow \mathbb{R} : t \mapsto f(a + tv).$$

Die Richtungsableitung von f in a in Richtung v existiert genau dann, wenn g in 0 differenzierbar ist, und in diesem Fall gilt

$$g'(0) = \lim_{t \to 0} \frac{f(a+tv) - f(a)}{t} = \frac{\partial f}{\partial v}(a).$$

Diese Aussage läßt sich insbesondere auf die partiellen Ableitungen, d.h. den Fall $v = e_i$, anwenden. Man führt die partielle Differenzierbarkeit also auf die Differenzierbarkeit im Eindimensionalen zurück.

Korollar 24.8 (Total differenzierbar impliziert partiell differenzierbar.)
Ist $f : U \longrightarrow \mathbb{R}^m$ total differenzierbar in $a \in U$, so ist f partiell differenzierbar in a und die Ableitung von f in a ist die Jacobi-Matrix von f in a

$$Df(a) = Jf(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \frac{\partial f_1}{\partial x_2}(a) & \ldots & \frac{\partial f_1}{\partial x_n}(a) \\ \frac{\partial f_2}{\partial x_1}(a) & \frac{\partial f_2}{\partial x_2}(a) & \ldots & \frac{\partial f_2}{\partial x_n}(a) \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \frac{\partial f_m}{\partial x_2}(a) & \ldots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}.$$

Beweis: In Lemma 24.4 haben wir gezeigt, daß die Grenzwerte

$$\lim_{t \to 0} \frac{f(a + te_j) - f(a)}{t}$$

existieren, und nach Satz 23.19 existieren dann die Grenzwerte der Komponentenfunktionen

$$\frac{\partial f_i}{\partial x_j}(a) = \lim_{t\to 0} \frac{f_i(a+te_j) - f_i(a)}{t}.$$

f ist also partiell differenzierbar in a, und zudem folgt aus Lemma 24.4, daß die j-te Spalte von $Df(a)$ genau die j-te Spalte von $Jf(a)$ ist. □

Beispiel 24.9 (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

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x_1, x_2)^t \mapsto \begin{cases} \frac{x_1 \cdot x_2^3}{x_1^2 + x_2^6}, & \text{falls } (x_1, x_2)^t \neq (0,0), \\ 0 & \text{sonst.} \end{cases}$$

Wir wollen zeigen, daß für f in $a = (0,0)^t$ sogar alle Richtungsableitungen existieren, aber f trotzdem in a nicht total differenzierbar ist.

Sei dazu $v = (v_1, v_2) \in \mathbb{R}^2$ mit $||v||_2 = 1$. Ist $v_1 = 0$, so ist

$$\frac{\partial f}{\partial v}(a) = \lim_{t\to 0} \frac{f(a+tv) - f(a)}{t} = \lim_{t\to 0} \frac{0}{t^7} = 0.$$

Ist $v_1 \neq 0$, so ist

$$\frac{\partial f}{\partial v}(a) = \lim_{t\to 0} \frac{f(a+tv) - f(a)}{t} = \lim_{t\to 0} \frac{t^4 v_1 v_2^3}{t^3 v_1^2 + t^7 v_2^6} = \lim_{t\to 0} \frac{t v_1 v_2^3}{v_1^2 + t^4 v_2^4} = 0.$$

Mithin existiert die Richtungsableitung in Richtung v im Punkt a und ist 0.

f ist damit in a insbesondere partiell differenzierbar mit Jacobi-Matrix $Jf(a) = (0\ 0)$. Für die Funktion ρ ergibt sich damit

$$\rho(x_1, x_2) = f(x_1, x_2) - f(0,0) - Jf(a) \cdot (x_1, x_2)^t = f(x_1, x_2),$$

und wir wollen nun zeigen, daß der Grenzwert

$$\lim_{x\to a} \frac{\rho(x)}{||x-a||_2} \tag{77}$$

nicht existiert. Dazu betrachten wir die Folge $(a_n)_{n\geq 1}$ mit

$$a_n = \left(\tfrac{1}{n^3}, \tfrac{1}{n}\right)^t \longrightarrow (0,0)^t.$$

Dann gilt

$$\frac{\rho(a_n)}{||a_n - a||_2} = \frac{\frac{1}{n^6}}{\frac{2}{n^6} \cdot \sqrt{\frac{1}{n^6} + \frac{1}{n^2}}} = \frac{n}{2 \cdot \sqrt{\frac{1}{n^4} + 1}} \longrightarrow \infty.$$

Aus dem Folgenkriterium für Grenzwerte erhalten wir deshalb, daß der Grenzwert in (77) nicht existiert. Mithin ist die Funktion f nicht total differenzierbar im Ursprung.

Die Funktion f ist im Ursprung *nicht mal stetig*, da $\lim_{n\to\infty} f(a_n) = \frac{1}{2} \neq 0 = f(0,0)$.

Man beachte auch, daß die partiellen Ableitungen von f im Punkt $a = (0,0)^t$ nicht stetig sein können, wie aus dem folgenden Satz folgt.

C) Stetige Differenzierbarkeit

Definition 24.10 (Stetige Differenzierbarkeit)
Eine Abbildung $f : U \longrightarrow \mathbb{R}^m$ heißt *stetig differenzierbar* auf U, wenn f partiell differenzierbar auf U ist und alle partiellen Ableitungen stetig auf U sind.
Wir bezeichnen mit $\mathcal{C}^1(U, \mathbb{R}^m)$ die Menge der stetig differenzierbaren Funktionen von U nach $\mathbb{R}^m$.

Satz 24.11 (Stetig differenzierbar impliziert total differenzierbar.)
Ist $f : U \longrightarrow \mathbb{R}^m$ stetig differenzierbar auf U, so ist f total differenzierbar auf U.

Beweis: Es seien wieder $f_1, \ldots, f_m$ die Komponentenfunktionen von f und $a \in U$. Wir wählen ein $\varepsilon > 0$, so daß die ε-Umgebung von a bezüglich der Maximumsnorm ganz in U liegt, und wir betrachten nur noch Punkte x in dieser ε-Umgebung. Außerdem halten wir ein $i \in \{1, \ldots, m\}$ fest. Dann gilt

$$f_i(x) - f_i(a) = \sum_{j=1}^{n} f_i(x_1, \ldots, x_j, a_{j+1}, \ldots, a_n) - f_i(x_1, \ldots, x_{j-1}, a_j, \ldots, a_n).$$

Da die Funktion

$$U_\varepsilon(a_j) = (a_j - \varepsilon, a_j + \varepsilon) \longrightarrow \mathbb{R} : z \mapsto f_i(x_1, \ldots, x_{j-1}, z, a_{j+1}, \ldots, a_n)$$

nach Voraussetzung differenzierbar in a_j ist, kann man den Mittelwertsatz der Differentialrechnung 18.7 im Eindimensionalen anwenden und erhält

$$\sum_{j=1}^{n} f_i(x_1, \ldots, x_j, a_{j+1}, \ldots, a_n) - f_i(x_1, \ldots, x_{j-1}, a_j, \ldots, a_n)$$
$$\overset{MWS\ 18.7}{=} \sum_{j=1}^{n} \frac{\partial f_i}{\partial x_j}(x_1, \ldots, x_{j-1}, c_j, a_{j+1}, \ldots, a_n) \cdot (x_j - a_j).$$

für $c_1, \ldots, c_n \in \mathbb{R}$ mit $|c_j - a_j| < |x_j - a_j| < \varepsilon$ geeignet. Wir setzen nun

$$\begin{aligned}
\rho_i(x) &:= f_i(x) - f_i(a) - Jf_i(a) \circ (x - a) \\
&= \sum_{j=1}^{n} \left(\frac{\partial f_i}{\partial x_j}(x_1, \ldots, x_{j-1}, c_j, a_{j+1}, \ldots, a_n) - \frac{\partial f_i}{\partial x_j}(a) \right) \cdot (x_j - a_j) \\
&= \langle r(x), x - a \rangle,
\end{aligned}$$

wobei $r(x)$ ein Vektor ist, dessen j-te Komponente gerade

$$r_j(x) = \frac{\partial f_i}{\partial x_j}(x_1, \ldots, x_{j-1}, c_j, a_{j+1}, \ldots, a_n) - \frac{\partial f_i}{\partial x_j}(a)$$

ist. Die Schreibweise mit Hilfe des Skalarproduktes hat den Vorteil, daß wir den Satz von Cauchy-Schwarz A14.8 anwenden können. Wir erhalten dann

$$\frac{|\rho_i(x)|}{||x-a||_2} = \frac{|\langle r(x), x-a\rangle|}{||x-a||_2} \overset{CS\ A14.8}{\leq} \frac{||r(x)||_2 \cdot ||x-a||_2}{||x-a||_2} = ||r(x)||_2$$

Die Stetigkeit der partiellen Ableitungen in a impliziert nun aber, daß

$$\lim_{x\to a} r_j(x) = \lim_{x\to a} \frac{\partial f_i}{\partial x_j}(x_1, \ldots, x_{j-1}, c_j, a_{j+1}, \ldots, a_n) - \frac{\partial f_i}{\partial x_j}(a) = 0$$

für alle $j = 1, \ldots, n$, und mithin folgen $\lim\limits_{x\to a} ||r(x)||_2 = 0$ und deshalb

$$\lim_{x\to a} \frac{\rho_i(x)}{||x-a||_2} = 0.$$

Insgesamt erhalten wir für

$$\rho(x) = f(x) - f(a) - Jf(a) \circ (x-a) = \big(\rho_1(x), \ldots, \rho_m(x)\big)^t$$

dann

$$\lim_{x\to a} \frac{\rho(x)}{||x-a||_2} = 0.$$

Also ist f in a total differenzierbar. □

Beispiel 24.12 (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 (siehe Aufgabe 24.30).

Bemerkung 24.13 (Wie prüft man totale Differenzierbarkeit nach?)

a. Satz 24.11 ist für konkrete Anwendungen ungemein wichtig. Wollen wir überprüfen, ob eine Abbildung $f : U \longrightarrow \mathbb{R}^m$ total differenzierbar ist, so können wir also wie folgt vorgehen:

   1) Überprüfe, ob alle partiellen Ableitungen auf U existieren. Falls nicht, so ist f nicht total differenzierbar auf U.

   2) Überprüfe, ob alle partiellen Ableitungen stetig auf U sind. Falls ja, so ist f total differenzierbar auf U.

   Nur, wenn die partiellen Ableitungen existieren, aber nicht stetig sind, müssen wir auf die Definition zurück greifen und den Grenzwert

$$\lim_{x\to a} \frac{f(x) - f(a) - Jf(a) \circ (x-a)}{||x-a||_2}$$

für $a \in U$ untersuchen. Ist er null, so ist f in a total differenzierbar.

b. Wenn man den Beweis von Satz 24.11 genau betrachtet, sieht man, daß es reicht, zu fordern, daß f auf U partiell differenzierbar ist und daß die partiellen Ableitungen in a stetig sind, um die totale Differenzierbarkeit von f in a zu folgern. Die Stetigkeit der partiellen Ableitungen außerhalb von a ist nicht erforderlich, in der Praxis aber in aller Regel gegeben.

D) Erste einfache Eigenschaften differenzierbarer Abbildungen

Proposition 24.14 (Total differenzierbar impliziert stetig.)
Ist $f : U \longrightarrow \mathbb{R}^m$ total differenzierbar in $a \in U$, so ist f stetig in a.

Beweis: Wir setzen $A = Df(a)$ und erhalten dann

$$\begin{aligned}\lim_{x\to a} f(x) &= \lim_{x\to a}\left(f(a) + A\circ(x-a) + \rho(x)\right)\\ &= f(a) + \lim_{x\to a} A\circ(x-a) + \lim_{x\to a}\frac{\rho(x)}{||x-a||_2}\cdot\lim_{x\to a}||x-a||_2 = f(a),\end{aligned}$$

da die lineare Abbildung, die durch die Multiplikation mit A definiert wird, stetig ist. Also ist f stetig im Häufungspunkt a von U. □

Beispiel 24.15 (Stetig impliziert nicht differenzierbar.)
Die Abbildung

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : x = (x_1, x_2)^t \mapsto \begin{cases} \frac{x_1x_2}{\sqrt{x_1^2+x_2^2}}, & \text{falls } (x_1,x_2)^t \neq (0,0)^t,\\ 0 & \text{sonst}\end{cases}$$

ist wegen Beispiel 23.5 b. stetig in $a = (0,0)^t$. Wir wollen nun zeigen, daß sie nicht total differenzierbar in a ist, wohl aber partiell differenzierbar. Die partiellen Ableitungen von f sind

$$\frac{\partial f}{x_i}(a) = \lim_{t\to 0}\frac{f(a+te_i)-f(a)}{t} = \lim_{t\to 0}\frac{0}{t^2} = 0$$

und die Jacobi-Matrix von f im Punkt $a = (0,0)^t$ ist mithin

$$Jf(a) = (0\ 0).$$

Für die Funktion ρ ergibt sich daraus

$$\rho(x) = f(x) - f(a) - Jf(a)\circ x = f(x).$$

Wir müssen den Grenzwert von

$$\frac{\rho(x)}{||x||_2} = \frac{x_1x_2}{x_1^2+x_2^2}$$

im Punkt $a = (0,0)^t$ betrachten, aber dieser existiert nach Beispiel 23.5 a. nicht. Mithin ist die Funktion f nicht total differenzierbar in $a = (0,0)^t$.

Bemerkung 24.16 (Die Ableitung als Abbildung)
In der eindimensionalen Analysis haben wir die Ableitung einer auf U differenzierbaren Funktion als eine Funktion auf U aufgefaßt, die Werte in $\mathbb{R}$ annimmt. Das war möglich, weil die Ableitung von f in einem Punkt a eine reelle Zahl war. Im Mehrdimensionalen ist die Ableitung von $f : U \longrightarrow \mathbb{R}^m$ in einem Punkt aber eine $m \times n$-Matrix. In diesem Sinne nimmt die Ableitung

$$Df : U \longrightarrow \mathrm{Mat}(m \times n, \mathbb{R}) : a \mapsto Df(a) = Jf(a)$$

Werte in $\mathrm{Mat}(m \times n, \mathbb{R})$ an. $\mathrm{Mat}(m \times n, \mathbb{R})$ können wir mit $\mathbb{R}^{mn}$ identifizieren und mit der euklidischen Norm versehen,

$$||(a_{ij})||_2 = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2},$$

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 24.17 (Stetige Differenzierbarkeit)
Eine Abbildung $f : U \longrightarrow \mathbb{R}^m$ ist genau dann stetig differenzierbar auf U, wenn f auf U total differenzierbar ist und $Df : U \longrightarrow \mathrm{Mat}(m \times n, \mathbb{R})$ auf U stetig ist.

Beweis: Wir identifizieren bei dieser Aussage $\mathrm{Mat}(m \times n, \mathbb{R})$ mit $\mathbb{R}^{mn}$ und betrachten diesen Raum mit einer beliebigen Norm als normierten Raum.

Ist f stetig differenzierbar auf U, so ist f nach Satz 24.11 auch total differenzierbar auf U. Mithin existiert die Abbildung Df, und ihre Komponentenfunktionen sind die partiellen Ableitungen von f, die nach Voraussetzung stetig auf U sind. Dann ist aber nach Satz 23.19 auch Df stetig.

Ist umgekehrt f total differenzierbar auf U, so existieren nach Korollar 24.8 auch alle partiellen Ableitungen von f auf U und diese sind die Komponentenfunktionen von Df. Wenn nun die Abbildung Df als stetig auf U voraus gesetzt wird, so sind nach Satz 23.19 die partiellen Ableitungen von f also stetig auf U, und f ist mithin stetig differenzierbar auf U. □

E) Die Linearität der Ableitung

Proposition 24.18 (Lineare Abbildungen sind total differenzierbar.)
Für $A \in \mathrm{Mat}(m \times n, \mathbb{R})$ ist $f_A : \mathbb{R}^n \to \mathbb{R}^m$ total differenzierbar in allen $a \in \mathbb{R}^n$ mit

$$Df_A(a) = A.$$

Beweis: Dies folgt unmittelbar aus der Definition der totalen Differenzierbarkeit mit der Nullabbildung als ρ. □

Beispiel 24.19
Die Funktion $\pi_i : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto x_i$ ist linear und somit total differenzierbar.

Proposition 24.20 (Linearität der Ableitung)
Sind $f, g : U \longrightarrow \mathbb{R}^m$ total differenzierbar in $a \in U$ und $\lambda, \mu \in \mathbb{R}$, so ist $\lambda \cdot f + \mu \cdot g$ total differenzierbar in a mit

$$D(\lambda \cdot f + \mu \cdot g)(a) = \lambda \cdot Df(a) + \mu \cdot Dg(a).$$

Beweis: Nach Voraussetzung gilt für

$$\rho_f(x) = f(x) - f(a) - Df(a) \circ (x - a)$$

und

$$\rho_g(x) = g(x) - g(a) - Dg(a) \circ (x - a),$$

daß

$$\lim_{x \to a} \frac{\rho_f(x)}{||x - a||_2} = \lim_{x \to a} \frac{\rho_g(x)}{||x - a||_2} = 0.$$

Dann gilt aber auch

$$(\lambda \cdot f + \mu \cdot g)(x) - (\lambda \cdot f + \mu \cdot g)(a) - (\lambda \cdot Df(a) + \mu \cdot Dg(a)) \circ (x - a) \;=\; \lambda \cdot \rho_f(x) + \mu \cdot \rho_g(x)$$

und

$$\lim_{x \to a} \frac{\lambda \cdot \rho_f(x) + \mu \cdot \rho_g(x)}{||x - a||_2} = \lambda \cdot \lim_{x \to a} \frac{\rho_f(x)}{||x - a||_2} + \mu \cdot \lim_{x \to a} \frac{\rho_g(x)}{||x - a||_2} = 0.$$

Damit folgt die Behauptung. □

F) Die Produktregel

Proposition 24.21 (Produktregel)
Sind $f, g : U \longrightarrow \mathbb{R}$ total differenzierbar in $a \in U$, so ist auch $f \cdot g$ total differenzierbar in a mit $D(f \cdot g)(a) = f(a) \cdot Dg(a) + Df(a) \cdot g(a)$.

Beweis: Nach Voraussetzung existieren $A = Df(a)$ und $B = Dg(a)$ und für die Funktionen $\rho_f : U \longrightarrow \mathbb{R}$ und $\rho_g : U \longrightarrow \mathbb{R}$ mit

$$\rho_f(x) = f(x) - f(a) - A \circ (x - a) \tag{78}$$

und

$$\rho_g(x) = g(x) - g(a) - B \circ (x - a) \tag{79}$$

gilt

$$\lim_{x \to a} \frac{\rho_f(x)}{||x - a||_2} = \lim_{x \to a} \frac{\rho_g(x)}{||x - a||_2} = 0.$$

Der Übersichtlichkeit halber setzen wir zwischendurch $y = x - a$ und erhalten dann unter Berücksichtigung von (78) und (79)

$$\begin{aligned}
\rho_{f\cdot g}(x) &:= (f \cdot g)(x) - (f \cdot g)(a) - \big(f(a) \cdot B + A \cdot g(a)\big) \circ (x - a) \\
&= \big(f(x) - f(a)\big) \cdot g(x) + f(a) \cdot \big(g(x) - g(a)\big) - f(a) \cdot By - Ay \cdot g(a) \\
&= \big(\rho_f(x) + Ay\big) \cdot g(x) + f(a) \cdot \big(\rho_g(x) + By\big) - f(a) \cdot By - Ay \cdot g(a) \\
&= \rho_f(x) \cdot g(x) + f(a) \cdot \rho_g(x) + Ay \cdot \big(g(x) - g(a)\big) \\
&= \rho_f(x) \cdot g(x) + f(a) \cdot \rho_g(x) + Ay \cdot \big(\rho_g(x) + By\big) \\
&= \rho_f(x) \cdot g(x) + f(a) \cdot \rho_g(x) + f_A(x - a) \cdot \rho_g(x) + f_A(x - a) \cdot f_B(x - a).
\end{aligned}$$

Wir beachten nun, daß die Multiplikation mit einer Matrix eine stetige lineare Abbildung ist, und mit Korollar 23.51 gilt deshalb

$$\frac{|f_A(x - a)| \cdot |f_B(x - a)|}{||x - a||_2} \leq \frac{||f_A|| \cdot ||x - a||_2 \cdot ||f_B|| \cdot ||x - a||_2}{||x - a||_2} = ||f_A|| \cdot ||f_B|| \cdot ||x - a||_2,$$

wobei $||f_A||$ und $||f_B||$ die Operatornormen von f_A und f_B sind. Da die rechte Seite für $x \to a$ gegen Null konvergiert, folgt

$$\lim_{x \to a} \frac{f_A(x - a) \cdot f_B(x - a)}{||x - a||_2} = 0.$$

Nutzen wir nun noch aus, daß die in a total differenzierbare Funktion g nach Proposition 24.14 auch stetig in a ist und daß die Multiplikation mit einer Matrix eine stetige

Abbildung ist, so erhalten wir

$$\begin{aligned}
&\lim_{x\to a}\frac{\rho_{f\cdot g}(x)}{||x-a||_2} = \\
&= \lim_{x\to a}\frac{\rho_f(x)\cdot g(x) + f(a)\cdot\rho_g(x) + f_A(x-a)\cdot\rho_g(x) + f_A(x-a)\cdot f_B(x-a)}{||x-a||_2} \\
&= \lim_{x\to a}\frac{\rho_f(x)}{||x-a||_2}\cdot\lim_{x\to a} g(x) + f(a)\cdot\lim_{x\to a}\frac{\rho_g(x)}{||x-a||_2} \\
&\quad + \lim_{x\to a} f_A(x-a)\cdot\lim_{x\to a}\frac{\rho_g(x)}{||x-a||_2} + \lim_{x\to a}\frac{f_A(x-a)\cdot f_B(x-a)}{||x-a||_2} \\
&= 0\cdot g(a) + f(a)\cdot 0 + 0\cdot 0 + 0 = 0.
\end{aligned}$$

Also ist $f\cdot g$ total differenzierbar in a mit der Ableitung

$$D(f\cdot g)(a) = f(a)\cdot B + A\cdot g(a) = f(a)\cdot Dg(a) + Df(a)\cdot g(a).$$

□

Beispiel 24.22 (Polynomfunktionen sind total differenzierbar.)

a. Aus der Linearität der Ableitung 24.20, der Produktregel 24.21 und Beispiel 24.19 folgt, daß jede Polynomfunktion

$$p : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto \sum_{|\alpha|=0}^{d} a_\alpha\cdot x^\alpha$$

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

$$N^2 : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto ||x||_2^2 = x_1^2 + \ldots + x_n^2$$

ist eine Polynomfunktion und ist mithin total differenzierbar mit

$$DN^2(a) = \begin{pmatrix}2a_1 & \ldots & 2a_n\end{pmatrix}.$$

G) Die Kettenregel

Proposition 24.23 (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

$$Dh(a) = Dg\big(f(a)\big)\circ Df(a).$$

Beweis: Nach Voraussetzung existieren $A = Df(a) \in \mathrm{Mat}(m \times n, \mathbb{R})$ und $B = Dg\big(f(a)\big) \in \mathrm{Mat}(k \times m, \mathbb{R})$ und für die Funktionen $\rho_f : U \longrightarrow \mathbb{R}^m$ und $\rho_g : V \longrightarrow \mathbb{R}^k$ mit

$$\rho_f(x) = f(x) - f(a) - A \circ (x - a) \tag{80}$$

und

$$\rho_g(y) = g(y) - g\big(f(a)\big) - B \circ \big(y - f(a)\big) \tag{81}$$

gilt

$$\lim_{x \to a} \frac{\rho_f(x)}{||x - a||_2} = \lim_{y \to f(a)} \frac{\rho_g(y)}{||y - f(a)||_2} = 0. \tag{82}$$

Wir erhalten damit

$$\begin{aligned}
\rho_h(x) &:= h(x) - h(a) - B \circ A \circ (x - a) \\
&= g\big(f(x)\big) - g\big(f(a)\big) - B \circ A \circ (x - a) \\
&\overset{(81)}{=} \rho_g\big(f(x)\big) + B \circ \big(f(x) - f(a)\big) - B \circ A \circ (x - a) \\
&= \rho_g\big(f(x)\big) + B \circ \big(f(x) - f(a) - A \circ (x - a)\big) \\
&= \rho_g\big(f(x)\big) + B \circ \rho_f(x) = \rho_g\big(f(x)\big) + f_B\big(\rho_f(x)\big).
\end{aligned}$$

Die Kettenregel ist nun bewiesen, sobald wir zeigen können, daß

$$\lim_{x \to a} \frac{\rho_h(x)}{||x - a||_2} = \lim_{x \to a} \frac{\rho_g\big(f(x)\big) + f_B(\rho_f(x))}{||x - a||_2} = 0$$

gilt. Dazu beachten wir zunächst, daß

$$\frac{||f_B(\rho_f(x))||_2}{||x - a||_2} \overset{23.51}{\leq} ||f_B|| \cdot \left|\left| \frac{\rho_f(x)}{||x - a||_2} \right|\right|_2 \longrightarrow 0$$

für $x \to a$, wobei $||f_B||$ die Operatornorm von f_B ist. Also gilt schon mal

$$\lim_{x \to a} \frac{f_B(\rho_f(x))}{||x - a||_2} = 0,$$

und es reicht mithin noch

$$\lim_{x \to a} \frac{\rho_g(f(x))}{||x - a||_2} = 0$$

zu zeigen. Dies wollen wir im folgenden tun.

Wegen (82) ist die Abbildung

$$\sigma : V \longrightarrow \mathbb{R}^k : y \mapsto \begin{cases} \frac{\rho_g(y)}{||y - f(a)||_2}, & \text{falls } y \neq f(a), \\ 0, & \text{sonst} \end{cases}$$

stetig in $y = f(a)$. Da außerdem $\rho_g\big(f(a)\big) = 0$ wegen (81) gilt, erhalten wir

$$\rho_g(y) = \sigma(y) \cdot ||y - f(a)||_2$$

für $y \in V$ und somit mit $y = f(x)$

$$\begin{aligned}
\left|\left|\frac{\rho_g(f(x))}{||x-a||_2}\right|\right|_2 &= ||\sigma(f(x))||_2 \cdot \frac{||f(x)-f(a)||_2}{||x-a||_2} \\
&\overset{(80)}{=} ||\sigma(f(x))||_2 \cdot \frac{||A \circ (x-a) + \rho_f(x)||_2}{||x-a||_2} \\
&\leq ||\sigma(f(x))||_2 \cdot \left(\frac{||f_A(x-a)||_2}{||x-a||_2} + \frac{||\rho_f(x)||_2}{||x-a||_2}\right) \\
&\leq ||\sigma(f(x))||_2 \cdot \left(\frac{||f_A|| \cdot ||(x-a)||_2}{||x-a||_2} + \frac{||\rho_f(x)||_2}{||x-a||_2}\right) \\
&\leq ||\sigma(f(x))||_2 \cdot \left(||f_A|| + \frac{||\rho_f(x)||_2}{||x-a||_2}\right),
\end{aligned}$$

wobei $||f_A||$ die Operatornorm von f_A ist. Nach Proposition 24.14 ist f stetig in a und zudem ist σ stetig in $f(a)$, so daß nach Proposition 23.15 auch $\sigma \circ f$ stetig in a ist. Unter Berücksichtigung von (82) konvergiert die rechte Seite der Ungleichung für $x \to a$ deshalb gegen

$$\lim_{x \to a} \left(||\sigma(f(x))||_2 \cdot \left(||f_A|| + \frac{||\rho_f(x)||_2}{||x-a||_2}\right)\right) = ||\sigma(f(a))||_2 \cdot (||f_A|| + 0) = 0,$$

so daß auch die linke Seite der Ungleichung gegen 0 konvergieren muß. Wir erhalten deshalb

$$\lim_{x \to a} \frac{\rho_g(f(x))}{||x-a||_2} = 0,$$

und damit ist die Behauptung bewiesen. □

Beispiel 24.24 (Kettenregel)

a. Aus der Kettenregel folgt, daß die euklidische Norm

$$N : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto ||x||_2 = \sqrt{x_1^2 + \ldots + x_n^2}$$

als Verkettung einer Polynomfunktion

$$p : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto x_1^2 + \ldots + x_n^2$$

mit der Wurzelfunktion

$$w : \mathbb{R}_{\geq 0} \longrightarrow \mathbb{R} : y \mapsto \sqrt{y}$$

differenzierbar auf $\mathbb{R}^n \setminus \{0\}$ ist mit

$$DN(a) = Dw(p(a)) \cdot Dp(a) = \frac{1}{2\sqrt{p(a)}} \cdot (2a_1, \ldots, 2a_n) = \left(\frac{a_1}{||a||_2}, \ldots, \frac{a_n}{||a||_2}\right).$$

Man beachte hierbei, daß die Wurzelfunktion in 0 nicht differenzierbar ist.

b. Die Funktion

$$g : \mathbb{R}^n \setminus \{0\} \longrightarrow \mathbb{R} : x \mapsto \frac{1}{||x||_2}$$

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

$$\begin{aligned} Dg(a) &= D\,\text{inv}\big(N(a)\big) \circ DN(a) \\ &= -\frac{1}{N(a)^2} \cdot \left(\frac{x_1}{||x||_2}, \ldots, \frac{x_n}{||x||_2}\right) = \left(\frac{-x_1}{||x||_2^3}, \ldots, \frac{-x_n}{||x||_2^3}\right). \end{aligned}$$

H) Die Quotientenregel

Proposition 24.25 (Quotientenregel)
Seien $f, g : U \longrightarrow \mathbb{R}$ total differenzierbar in $a \in U$ mit $g(a) \neq 0$ 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 a mit

$$D\left(\frac{f}{g}\right)(a) = \frac{Df(a) \cdot g(a) - f(a) \cdot Dg(a)}{g^2(a)}.$$

Beweis: Da g stetig auf U ist, ist $g^{-1}(0)$ abgeschlossen in U nach Proposition 23.17 und $U \setminus g^{-1}(0)$ ist offen in U. Nach Aufgabe 22.53 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 a wegen der Produktregel 24.21 und der Kettenregel 24.23. Für die Ableitung folgt zudem

$$\begin{aligned} D\left(\frac{f}{g}\right)(a) &= Df(a) \cdot (\text{inv} \circ g)(a) + f(a) \cdot D(\text{inv} \circ g)(a) \\ &= Df(a) \cdot (\text{inv} \circ g)(a) + f(a) \cdot D\,\text{inv}\big(g(a)\big) \cdot Dg(a) \\ &= \frac{Df(a)}{g(a)} - \frac{f(a) \cdot Dg(a)}{g^2(a)} \\ &= \frac{Df(a) \cdot g(a) - f(a) \cdot Dg(a)}{g^2(a)}. \end{aligned}$$

□

Beispiel 24.26 (Rationale Funktionen sind total differenzierbar.)
Da Polynomfunktionen nach Beispiel 24.22 total differenzierbar und stetig sind, folgt aus der Quotientenregel 24.25, daß auch jede rationale Funktion auf ihrem Definitionsbereich total differenzierbar ist.

I) Geometrische Interpretation des Gradienten

Definition 24.27 (Gradient)
Ist $f : U \longrightarrow \mathbb{R}$ eine in $a \in U$ partiell differenzierbare Funktion, so nennen wir

$$\operatorname{grad}\big(f(a)\big) = \nabla f(a) = Jf(a)^t = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right)^t$$

auch den *Gradienten* von f im Punkt a.[3]

Bemerkung 24.28 (Geometrische Interpretation des Gradienten)
Ist $f : U \longrightarrow \mathbb{R}$ total differenzierbar in a, so ist

$$x \mapsto f(a) + Jf(a) \circ (x - a) = f(a) + \langle \nabla f(a), x - a \rangle$$

die lineare Approximation von f im Punkt 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 $(a, f(a))$ approximiert. Man nennt diese Hyperebene auch den Tangentialraum an den Graphen von f im Punkt $(a, f(a))$ (siehe Abbildung 5).

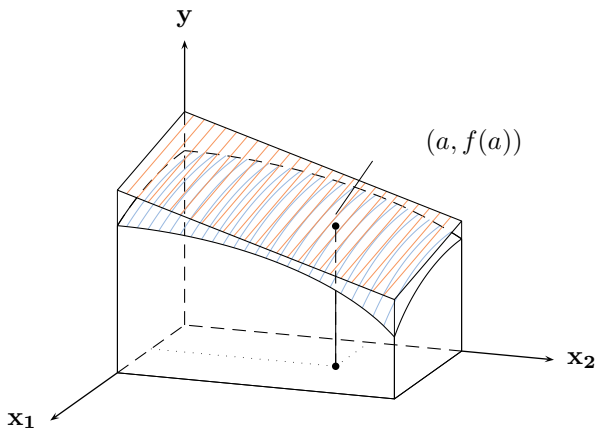


Abbildung 5: Tangentialraum an den Graphen von f in $(a, f(a))$

[3]Wir verwenden die Kommata in der Darstellung der $1 \times n$-Matrix $Jf(a)$, 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.

Der Gradient von f in 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 $x-a$ und $\nabla f(a)$

$$\langle \nabla f(a), x-a\rangle = ||\nabla f(a)||_2 \cdot ||x-a||_2 \cdot \cos(\alpha)$$

gilt. Mithin ist der Zuwachs von $f(a)$ zu

$$f(x) \approx f(a) + \langle \nabla f(a), x-a\rangle = f(a) + ||\nabla f(a)||_2 \cdot ||x-a||_2 \cdot \cos(\alpha)$$

maximal, wenn $\cos(\alpha) = 1$ ist. Das ist aber genau dann der Fall, wenn $\alpha = 0$ gilt, wenn also $\nabla f(a)$ und $x-a$ die gleiche Richtung haben.

Man kann diesen Sachverhalt auch mit Hilfe der Richtungsableitung ausdrücken. Wegen Lemma 24.4 ist die Richtungsableitung von f im Punkt a in Richtung v mit $||v||_2 = 1$ durch

$$\frac{\partial f}{\partial v}(a) = Df(a) \circ v = \langle \nabla f(a), v\rangle$$

gegeben (siehe auch Abbildung 4). Wegen der Gleichung von Cauchy-Schwarz A14.8 gilt

$$\left|\frac{\partial f}{\partial v}(a)\right| = |\langle \nabla f(a), v\rangle| \leq ||\nabla f(a)||_2 \cdot ||v||_2 = ||\nabla f(a)||_2,$$

und es gilt genau dann Gleichheit, wenn $\nabla f(a)$ und v linear abhängig voneinander sind. Die Richtungsableitung in Richtung v wird also maximal, in Richtung

$$v = \frac{\nabla f(a)}{||\nabla f(a)||_2}.$$

Bei all diesen Betrachtungen sollte der Gradient $\nabla f(a)$ natürlich nicht der Nullvektor sein.

J) Differenzierbarkeit in normierten Räumen

Bemerkung 24.29 (Fréchet-Differenzierbarkeit)
Es seien V und W zwei normierte Räume, $U \subseteq V$ sei offen und $f : U \longrightarrow W$ sei eine Abbildung. Dann heißt f *Fréchet-differenzierbar* in $a \in U$, wenn es eine stetige lineare Abbildung $A \in L(V,W)$ und eine Abbildung $\rho : U \longrightarrow W$ gibt, mit

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

und

$$\lim_{x\to a} \frac{\rho(x)}{||x-a||_V} = 0.$$

Die Ableitung $Df(a) := A$ 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 a die Stetigkeit in 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

$$Df : U \longrightarrow L(V,W) : a \mapsto Df(a)$$

die Werte in $L(V,W)$ annimmt.

Aufgaben

Aufgabe 24.30 (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 24.31
Begründe, weshalb $f : \mathbb{R}^3_{>0} \longrightarrow \mathbb{R}^2 : (x,y,z)^t \mapsto \left(z\sqrt{y} + \sqrt{z}, xyz + \ln(x+y)\right)^t$ total differenzierbar ist, und berechne die Ableitung von f.

Aufgabe 24.32
Zeige, für $A \in \mathrm{Mat}_n(\mathbb{R})$ 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 24.33
Sei $\Delta = D_1^2 + \ldots + D_n^2$ der *Laplace-Operator* und $f \in \mathcal{C}^2((0,1),\mathbb{R})$.
Zeige, daß für die Funktion $\varphi : \mathbb{R}^n \setminus \{(0,\ldots,0)^t\} \longrightarrow \mathbb{R},\ x \longmapsto f(||x||_2)$ die Gleichung

$$\Delta\varphi(x) = f''(||x||_2) + \frac{n-1}{||x||_2} \cdot f'(||x||_2) \quad \text{für alle } x \in \mathbb{R}^n \setminus \{(0,\ldots,0)^t\}$$

gilt und berechne $\Delta\varphi$ für den Fall $f(t) = \frac{1}{t^{n-2}}$ mit $n > 2$.

Aufgabe 24.34
Zeige, ist $f : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ eine in $(0,\ldots,0)^t \in \mathbb{R}^n$ 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.

Aufgabe 24.35 a. Begründe bitte, weshalb die Abbildung $\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ gegeben durch

$$(r, \theta, \vartheta) \mapsto \big(r \cdot \cos(\theta) \cdot \cos(\vartheta), r \cdot \sin(\theta) \cdot \cos(\vartheta), r \cdot \sin(\vartheta)\big)$$

total differenzierbar auf $\mathbb{R}^3$ ist und berechne die Ableitung $D\varphi(r, \theta, \vartheta)$. Für welche $(r, \theta, \vartheta)^t \in \mathbb{R}^3$ ist die Matrix $D\varphi(r, \theta, \vartheta)$ invertierbar?

b. Berechne die Ableitung von $f \circ \varphi$ im Punkt $a = (1, 0, \frac{\pi}{2})^t$ für $f : \mathbb{R}^3 \longrightarrow \mathbb{R} : (x, y, z)^t \mapsto x^2 + y^2 + z^2$ zunächst mit Hilfe der Kettenregel und dann ohne diese.

Aufgabe 24.36
Sei $U \subseteq \mathbb{R}^n$ offen, $a \in U$, $f : U \longrightarrow \mathbb{R}$ sei stetig in a und $g : U \longrightarrow \mathbb{R}$ sei total differenzierbar in a mit $g(a) = 0$.
Zeige, dann ist $f \cdot g$ total differenzierbar in a mit $D(f \cdot g)(a) = f(a) \cdot Dg(a)$.

§ 25 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 25.1

Die Polynomfunktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : x = (x_1, x_2)^t \mapsto 3x_1^5 + 2x_1^3x_2 + 5x_2$$

ist partiell differenzierbar auf $\mathbb{R}^2$. Ihre partiellen Ableitungen

$$D_1 f : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto D_1 f(x) = \frac{\partial f}{\partial x_1}(x) = 15x_1^4 + 6x_1^2x_2$$

und

$$D_2 f : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto D_2 f(x) = \frac{\partial f}{\partial x_2}(x) = 2x_1^3 + 5$$

sind ebenfalls partiell differenzierbar auf $\mathbb{R}^2$. Wir können also auch von diesen wieder die partiellen Ableitungen bilden und erhalten:

$$D_1D_1f(x) = 60x_1^3 + 12x_1x_2, \quad D_2D_1f(x) = 6x_1^2, \quad D_1D_2f(x) = 6x_1^2, \quad D_2D_2f(x) = 0.$$

Dabei fällt uns auf, daß

$$D_2D_1f(x) = 6x_1^2 = D_1D_2f(x)$$

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 25.5 sehen werden. Zunächst wollen wir aber allgemein höhere partielle Ableitungen einführen.

Definition 25.2 (Höhere partielle Ableitungen)

Eine Funktion $f : U \longrightarrow \mathbb{R}$ heißt *k-fach stetig differenzierbar* auf U, falls für alle $j_1, \ldots, j_k \in \{1, \ldots, n\}$ die k-fachen partiellen Ableitungen

$$D_{j_k} \ldots D_{j_1} f : U \longrightarrow \mathbb{R}$$

existieren und stetig sind, d.h. für $l = 1, \ldots, k$ ist die Funktion $D_{j_{l-1}} \ldots D_{j_1} f$ auf U partiell differenzierbar nach x_{j_l} und die Funktion $D_{j_l} \ldots D_{j_1} f$ 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 25.3 ($\mathcal{C}^k(U, \mathbb{R}^m)$)
Da die Differenzierbarkeit und die Stetigkeit verträglich mit Linearkombinationen sind (siehe Proposition 24.20 und Proposition 23.13) ist $\mathcal{C}^k(U, \mathbb{R}^m)$ ein $\mathbb{R}$-Vektorraum. Zudem folgt aus der Definition und den Ergebnissen des letzten Abschnittes unmittelbar, daß

$$\mathcal{C}(U, \mathbb{R}^m) \supseteq \mathcal{C}^1(U, \mathbb{R}^m) \supseteq \mathcal{C}^2(U, \mathbb{R}^m) \supseteq \mathcal{C}^3(U, \mathbb{R}^m) \supseteq \dots.$$

Man nennt dann den $\mathbb{R}$-Vektorraum

$$\mathcal{C}^\infty(U, \mathbb{R}^m) = \bigcap_{k \geq 0} \mathcal{C}^k(U, \mathbb{R}^m)$$

den Vektorraum der *unendlich oft differenzierbaren Abbildungen* auf U.

Beispiel 25.4 (Die Glockenfunktion)
Die Funktion

$$f : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \exp\left(-\frac{1}{(1-||x||_2^2)^2}\right), & \text{falls } ||x||_2 < 1, \\ 0, & \text{falls } ||x||_2 \geq 1 \end{cases}$$

ist unendlich oft differenzierbar auf $\mathbb{R}^n$, da sie die Verkettung der unendlich oft differenzierbaren Funktion $\mathbb{R} \longrightarrow \mathbb{R} : t \mapsto \exp\left(-\frac{1}{t^2}\right)$ aus Beispiel 18.34 mit der unendlich oft differenzierbaren Funktion $\mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto 1 - ||x||_2^2$ ist. f nimmt im Inneren der Einheitskugel nur positive Werte an und ist ansonsten stets null.

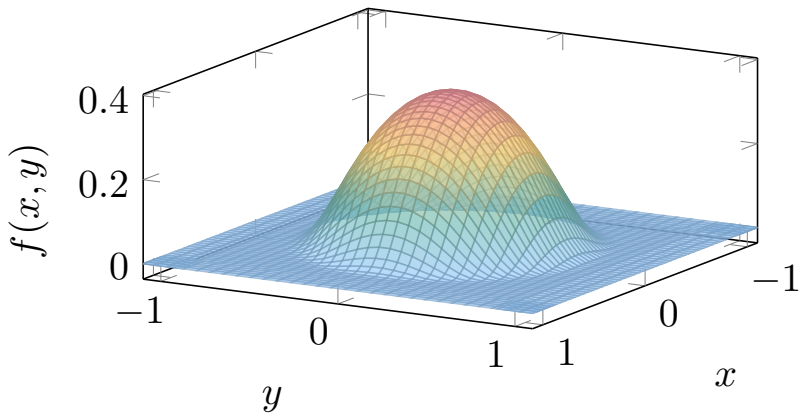


Abbildung 6: Graph der Glockenfunktion für $n = 2$

Satz 25.5 (Schwarz)
Ist $f : U \to \mathbb{R}^m$ zweifach stetig differenzierbar, so gilt für alle $a \in U$ und $1 \le i, j \le n$

$$D_i D_j f(a) = D_j D_i f(a).$$

Beweis: Es reicht, die Aussage für die Komponentenfunktionen von f zu zeigen, so daß wir ohne Einschränkung annehmen können, daß $m = 1$ gilt.

Da U offen ist, gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(a) \subseteq U$, und da die euklidische Norm nach Lemma 22.41 äquivalent zur Maximumsnorm ist, gibt es zudem ein $\delta > 0$, so daß

$$\left\{a + s \cdot e_i + t \cdot e_j \mid -\delta \le s, t \le \delta\right\} \subset U_\varepsilon(a) \subseteq U.$$

Wir interessieren uns im folgenden für die Funktion $H : O \to \mathbb{R}$ mit

$$H(s,t) = \frac{f(a + se_i + te_j) - f(a + te_j) - f(a + se_i) + f(a)}{s \cdot t},$$

mit $O = \{(s,t)^t \in \mathbb{R}^2 \mid -\delta < s, t < \delta \text{ und } s, t \neq 0\} \subset \mathbb{R}^2$ offen, da

$$D_j D_i f(a) = \lim_{t \to 0} \frac{D_i f(a + te_j) - D_i f(a)}{t} = \lim_{t \to 0} \lim_{s \to 0} H(s,t)$$

und

$$D_i D_j f(a) = \lim_{s \to 0} \frac{D_j f(a + se_i) - D_j f(a)}{s} = \lim_{s \to 0} \lim_{t \to 0} H(s,t).$$

In der Tat werden wir zeigen, daß

$$D_i D_j f(a) = \lim_{(s,t) \to (0,0)} H(s,t) = D_j D_i f(a)$$

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

$$F_s : (-\delta, \delta) \longrightarrow \mathbb{R} : t \mapsto f(a + se_i + te_j) - f(a + te_j)$$

differenzierbar und die Ableitung von F_s ergibt sich aus der partiellen Ableitung von f nach x_j wie folgt:

$$\begin{aligned}
F_s'(t) &= \lim_{h \to 0} \frac{F_s(t+h) - F_s(t)}{h} \\
&= \lim_{h \to 0} \frac{f(a + se_i + (t+h)e_j) - f(a + (t+h)e_j) - f(a + se_i + te_j) + f(a + te_j)}{h} \\
&= \lim_{h \to 0} \frac{f(a + se_i + te_j + he_j) - f(a + se_i + te_j)}{h} - \lim_{h \to 0} \frac{f(a + te_j + he_j) - f(a + te_j)}{h} \\
&= D_j f(a + se_i + te_j) - D_j f(a + te_j).
\end{aligned}$$

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

$$F_s(t) - F_s(0) = t \cdot F_s'(\theta_{s,t}) = t \cdot \left(D_j f(a + se_i + \theta_{s,t} e_j) - D_j f(a + \theta_{s,t} e_j)\right). \tag{83}$$

Auch die Funktion

$$G_{s,t} : (-\delta, \delta) \longrightarrow \mathbb{R} : r \mapsto D_j f(a + re_i + \theta_{s,t} e_j)$$

ist differenzierbar und ihre Ableitung ergibt sich aus der partiellen Ableitung von $D_j f$ nach x_i als

$$\begin{aligned} G'_{s,t}(r) =& \lim_{h \to 0} \frac{G_{s,t}(r+h) - G_{s,t}(r)}{h} \\ =& \lim_{h \to 0} \frac{D_j f(a + (r+h)e_i + \theta_{s,t} e_j) - D_j f(a + re_i + \theta_{s,t} e_j)}{h} \\ =& \lim_{h \to 0} \frac{D_j f(a + r \cdot e_i + \theta_{s,t} e_j + he_i) - D_j f(a + re_i + \theta_{s,t} e_j)}{h} \\ =& D_i D_j f(a + re_i + \theta_{s,t} e_j). \end{aligned}$$

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

$$G_{s,t}(s) - G_{s,t}(0) = s \cdot G'_{s,t}(\sigma_{s,t}) = s \cdot D_i D_j f(a + \sigma_{s,t} e_i + \theta_{s,t} e_j). \tag{84}$$

Für $0 \neq s, t \in (-\delta, \delta)$ ergibt sich dann aber

$$\begin{aligned} H(s,t) \;=\; & \frac{F_s(t) - F_s(0)}{s \cdot t} \\ \overset{(83)}{=} & \frac{D_j f(a + se_i + \theta_{s,t} e_j) - D_j f(a + \theta_{s,t} e_j)}{s} \\ \overset{(84)}{=} & D_i D_j f(a + \sigma_{s,t} e_i + \theta_{s,t} e_j). \end{aligned}$$

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

$$\lim_{(s,t) \to (0,0)} H(s,t) = D_i D_j f(a).$$

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

$$H(s,t) = D_j D_i f(a + \sigma'_{s,t} e_i + \theta'_{s,t} e_j),$$

so daß mit dem gleichen Argument auch

$$\lim_{(s,t) \to (0,0)} H(s,t) = D_j D_i f(a)$$

folgt. Damit ist insbesondere die Gleichheit

$$D_i D_j f(a) = D_j D_i f(a)$$

gezeigt. □

Beispiel 25.6 (Satz von Schwarz)
Im Satz von Schwarz ist die Voraussetzung, daß die zweiten partiellen Ableitungen stetig sind, wesentlich. Für die Funktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \frac{x_1^3 x_2 - x_1 x_2^3}{x_1^2 + x_2^2}, & \text{falls } (x_1, x_2)^t \neq (0,0)^t, \\ 0, & \text{sonst} \end{cases}$$

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 25.34).

Korollar 25.7 (Schwarz)
Ist $f \in \mathcal{C}^k(U, \mathbb{R}^m)$, $\sigma \in \mathbb{S}_k$ und $j_1, \ldots, j_k \in \{1, \ldots, n\}$, so gilt für alle $a \in U$

$$D_{j_1} \ldots D_{j_k} f(a) = D_{j_{\sigma(1)}} \ldots D_{j_{\sigma(k)}} f(a),$$

d.h. beim partiellen Ableiten kommt es nicht auf die Reihenfolge an.

Beweis: Man beachte zunächst, daß aus der Anwendung des Satzes von Schwarz 25.5 auf $D_{j_{l+2}} \ldots D_{j_k} f$ die Gleichheit

$$D_{j_1} \ldots D_{j_l} D_{j_{l+1}} D_{j_{l+2}} \ldots D_{j_k} f(a) = D_{j_1} \ldots D_{j_{l+1}} D_{j_l} D_{j_{l+2}} \ldots D_{j_k} f(a)$$

folgt, d.h. man kann zwei beliebige benachbarte partielle Ableitungen vertauschen — hierbei beachte man, daß die Abbildung $D_{j_{l+2}} \ldots D_{j_k} f$ mindestens zweifach stetig differenzierbar ist.

Da sich die Permutation σ nach Satz A8.7 als Produkt von Transpositionen benachbarter Zahlen schreiben läßt, folgt die Aussage also durch wiederholte Anwendung des Satzes von Schwarz 25.5. □

B) Taylorpolynome

Notation 25.8
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

$$D_i D_j f(a) = \frac{\partial \frac{\partial f}{\partial x_j}}{\partial x_i}(a)$$

schreiben können. Das ist zu umständlich. Gebräuchlicher ist dann die Notation

$$D_i D_j f(a) = \frac{\partial^2 f}{\partial x_i \partial x_j}(a),$$

wobei man bei doppelter Ableitung nach der gleichen Variablen dann auch

$$D_i^2 f(a) = \frac{\partial^2 f}{\partial x_i^2}(a)$$

schreibt. Diese Notation verallgemeinert sich dann in naheliegender Weise zu

$$D^\alpha f(a) = D_1^{\alpha_1} \dots D_n^{\alpha_n} f(a) = \frac{\partial^{\alpha_1+\dots+\alpha_n} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}(a)$$

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 23.14 bereits eingeführt haben, nochmals in Erinnerung rufen und erweitern. Für $\alpha = (\alpha_1, \dots, \alpha_n)^t \in \mathbb{N}^n$, $t = (t_1, \dots, t_n)^t$ und $a = (a_1, \dots, a_n)^t \in \mathbb{R}^n$ haben wir bereits

$$|\alpha| = \alpha_1 + \dots + \alpha_n$$

und

$$(t-a)^\alpha = (t_1 - a_1)^{\alpha_1} \cdot \dots \cdot (t_n - a_n)^{\alpha_n}$$

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 $\beta \le \alpha$, d.h. $0 \le \beta_j \le \alpha_j$ für alle $j = 1, \dots, n$.

Bemerkung 25.9 (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 $\beta \le \alpha$. 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 25.10 (Taylorpolynome)
Für eine k-fach stetig differenzierbare Funktion $f : U \longrightarrow \mathbb{R}$ und $a \in U$ heißt

$$T^k_{f,a} := \sum_{|\alpha|=0}^{k} \frac{D^\alpha f(a)}{\alpha!} \cdot (t-a)^\alpha = \sum_{\substack{\alpha \in \mathbb{N}^n \\ |\alpha| \le k}} \frac{D_1^{\alpha_1} \dots D_n^{\alpha_n} f(a)}{\alpha_1! \cdot \ldots \cdot \alpha_n!} \cdot (t_1 - a_1)^{\alpha_1} \cdot \ldots \cdot (t_n - a_n)^{\alpha_n}$$

das *k-te Taylorpolynom* von f mit Entwicklungspunkt a.
Gilt sogar $f \in \mathcal{C}^\infty(U, \mathbb{R})$, so nennen wir

$$T_{f,a} := \sum_{|\alpha|=0}^{\infty} \frac{D^\alpha f(a)}{\alpha!} \cdot (t-a)^\alpha = \sum_{\alpha \in \mathbb{N}^n} \frac{D_1^{\alpha_1} \dots D_n^{\alpha_n} f(a)}{\alpha_1! \cdot \ldots \cdot \alpha_n!} \cdot (t_1 - a_1)^{\alpha_1} \cdot \ldots \cdot (t_n - a_n)^{\alpha_n}$$

die *Taylorreihe* von f mit Entwicklungspunkt a.
Wir führen außerdem für zwei Punkte $a, x \in \mathbb{R}^n$ die Bezeichnung

$$\overline{ax} := \{a + t \cdot (x-a) \mid 0 \le t \le 1\} \subset \mathbb{R}^n$$

für die *Strecke* im $\mathbb{R}^n$ zwischen a und x ein.

Bemerkung 25.11
Ist $f : U \longrightarrow \mathbb{R}$ stetig differenzierbar und $a \in U$, so ist das erste Taylorpolynom

$$T^1_{f,a}(x) = f(a) + \sum_{i=1}^{n} D_i f(a) \cdot (x_i - a_i) = f(a) + \langle \nabla f(a), x - a \rangle$$

die lineare Approximation von f in a, die wir in Bemerkung 24.28 betrachtet haben.

Beispiel 25.12 (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 $a = (0,0)^t$ ist

$$\begin{aligned} T^2_{f,a} =& f(a) + \frac{D_1 f(a)}{1! \cdot 0!} \cdot (t_1 - a_1) + \frac{D_2 f(a)}{0! \cdot 1!} \cdot (t_2 - a_2) \\ &+ \frac{D_1^2 f(a)}{2! \cdot 0!} \cdot (t_1 - a_1)^2 + \frac{D_1 D_2 f(a)}{1! \cdot 1!} \cdot (t_1 - a_1) \cdot (t_2 - a_2) + \frac{D_2^2 f(a)}{0! \cdot 2!} \cdot (t_2 - a_2)^2 \\ =& 0 \cdot \cos(0) - 0 \cdot \sin(0) \cdot t_1 + \cos(0) \cdot t_2 - \frac{0 \cdot \cos(0)}{2} \cdot t_1^2 - \sin(0) \cdot t_1 \cdot t_2 + \frac{0}{2} \cdot t_2^2 = t_2. \end{aligned}$$

Will man die Taylorreihe von f mit Entwicklungspunkt $a = (0,0)^t$ bestimmen, so kann man sich in diesem konkreten Beispiel auch die Potenzreihenentwicklung des Cosinus zunutze machen und erhält

$$T_{f,a} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \cdot t_1^{2n} \cdot t_2.$$

C) Der Satz von Taylor

Die Kettenregel 24.23 erlaubt es, den Satz von Taylor 18.35 bzw. 20.14 ins Mehrdimensionale zu verallgemeinern.

Satz 25.13 (Satz von Taylor)
Ist $f \in \mathcal{C}^{k+1}(U, \mathbb{R})$ und $a, x \in U$ mit $\overline{ax} \subseteq U$, dann gibt es ein $c \in \overline{ax}$ mit

$$f(x) - T_{f,a}^k(x) = \sum_{|\alpha|=k+1} \frac{D^\alpha f(c)}{\alpha!} \cdot (x-a)^\alpha.$$

Die rechte Seite heißt auch *Restglied* des k-ten Taylorpolynoms nach Lagrange. In Integralform gilt für das Restglied zudem

$$f(x) - T_{f,a}^k(x) = (k+1) \cdot \int_0^1 (1-t)^k \cdot \sum_{|\alpha|=k+1} \frac{D^\alpha f(a + t\cdot(x-a))}{\alpha!} \cdot (x-a)^\alpha \, dt.$$

Der wesentliche Schritt im Beweis der mehrdimensionalen Taylorformel ist das folgende Lemma, das eine einfache Anwendung der Kettenregel darstellt.

Lemma 25.14
Es sei $f \in \mathcal{C}^{k+1}(U, \mathbb{R})$ und $a, x \in U$ mit $\overline{ax} \subseteq U$. Dann ist die Funktion

$$h : [0,1] \longrightarrow \mathbb{R} : t \mapsto f(a + t\cdot(x-a))$$

$k+1$-fach stetig differenzierbar auf $[0,1]$ und für $0 \le l \le k+1$ gilt

$$\frac{h^{(l)}(t)}{l!} = \sum_{|\alpha|=l} \frac{D^\alpha f(a + t\cdot(x-a))}{\alpha!} \cdot (x-a)^\alpha. \tag{85}$$

Beweis: Die Funktion $h = f \circ g$ ist als Verkettung der $k+1$-fach stetig differenzierbaren Funktion f mit der unendlich oft differenzierbaren Funktion

$$g : [0,1] \longrightarrow U : t \mapsto a + t\cdot(x-a)$$

$k+1$-fach stetig differenzierbar. Wir wollen nun mit abbrechender Induktion nach l zeigen, daß die l-te Ableitung von h die Gleichung (85) für $0 \le l \le k+1$ 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 $\alpha = (0, \ldots, 0)$ die einzige Möglichkeit für $|\alpha| = 0$ ist.

Sei also $l > 0$. Die Kettenregel liefert für die Funktion

$$F_\alpha : [0, 1] \longrightarrow \mathbb{R} : t \mapsto \big(D^\alpha f \circ g\big)(t)$$

die Ableitung

$$\begin{aligned} \frac{\partial F_\alpha}{\partial t}(t) =& D\big(D^\alpha f\big)(g(t)) \circ Dg(t) = D\big(D^\alpha f\big)(g(t)) \circ (x - a) \\ =& \sum_{j=1}^{n} D_j D^\alpha f(g(t)) \cdot (x_j - a_j) = \sum_{j=1}^{n} D^{\alpha + e_j} f(a + t \cdot (x - a)) \cdot (x_j - a_j). \end{aligned} \tag{86}$$

Dann folgt mit Hilfe von Induktion

$$\begin{aligned} \frac{h^{(l)}(t)}{l!} &= \frac{1}{l} \cdot \frac{\partial \frac{h^{(l-1)}}{(l-1)!}}{\partial t}(t) \\ &\overset{Ind.}{=} \frac{1}{l} \cdot \frac{\partial \sum_{|\alpha|=l-1} \frac{D^\alpha f(a+t\cdot(x-a))}{\alpha!} \cdot (x-a)^\alpha}{\partial t} \\ &= \frac{1}{l} \cdot \sum_{|\alpha|=l-1} \frac{\partial D^\alpha f(a + t \cdot (x - a))}{\partial t} \cdot \frac{(x - a)^\alpha}{\alpha!} \\ &= \frac{1}{l} \cdot \sum_{|\alpha|=l-1} \frac{\partial F_\alpha(t)}{\partial t} \cdot \frac{(x - a)^\alpha}{\alpha!} \\ &\overset{(86)}{=} \frac{1}{l} \cdot \sum_{|\alpha|=l-1} \sum_{j=1}^{n} D^{\alpha + e_j} f(a + t \cdot (x - a)) \cdot \frac{(x - a)^{\alpha + e_j}}{\alpha!} \\ &= \sum_{|\alpha|=l-1} \sum_{j=1}^{n} D^{\alpha + e_j} f(a + t \cdot (x - a)) \cdot \frac{(x - a)^{\alpha + e_j}}{(\alpha + e_j)!} \cdot \frac{\alpha_j + 1}{l} \\ &\overset{!}{=} \sum_{|\beta|=l} \frac{D^\beta f(a + t \cdot (x - a))}{\beta!} \cdot (x - a)^\beta. \end{aligned}$$

Um die letzte Gleichheit zu sehen, beachten wir folgendes. Ist $\beta \in \mathbb{N}^n$ mit $|\beta| = l$ gegeben und setzen wir für $j \in \{1, \ldots, n\}$

$$\alpha^j = (\beta_1, \ldots, \beta_{j-1}, \beta_j - 1, \beta_{j+1}, \ldots, \beta_n),$$

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

$$D^{\alpha^j+e_j} f(a+t\cdot(x-a)) \cdot \frac{(x-a)^{\alpha^j+e_j}}{(\alpha^j+e_j)!} = D^\beta f(a+t\cdot(x-a)) \cdot \frac{(x-a)^\beta}{\beta!}$$

gleich, und summieren wir sie auf, so erhalten wir

$$\sum_{j=1}^n D^{\alpha^j+e_j} f(a+t\cdot(x-a)) \cdot \frac{(x-a)^{\alpha^j+e_j}}{(\alpha^j+e_j)!} \cdot \frac{\alpha_j^j+1}{l} = D^\beta f(a+t\cdot(x-a)) \cdot \frac{(x-a)^\beta}{\beta!} \cdot \sum_{j=1}^n \frac{\beta_j}{l}$$

$$= D^\beta f(a+t\cdot(x-a)) \cdot \frac{(x-a)^\beta}{\beta!} \cdot \frac{\sum_{j=1}^n \beta_j}{l} = D^\beta f(a+t\cdot(x-a)) \cdot \frac{(x-a)^\beta}{\beta!}.$$

Damit ist (85) gezeigt. □

Beweis von Satz 25.13: Wenden wir nun den Satz von Taylor 18.35 auf die Funktion

$$h : [0,1] \longrightarrow \mathbb{R} : t \mapsto f(a+t\cdot(x-a))$$

an, so erhalten wir

$$\begin{aligned} f(x) - T_{f,a}^k(x) = \; & f(x) - \sum_{|\alpha|=0}^k \frac{D^\alpha f(a)}{\alpha!} \cdot (x-a)^\alpha \\ \overset{(85)}{=} \; & h(1) - \sum_{l=0}^k \frac{h^{(l)}(0)}{l!} \cdot (1-0)^l \\ \overset{18.35}{=} \; & \frac{h^{(k+1)}(\theta)}{(k+1)!} \cdot (1-0)^{k+1} \\ \overset{(85)}{=} \; & \sum_{|\alpha|=k+1} \frac{D^\alpha f(c)}{\alpha!} \cdot (x-a)^\alpha \end{aligned}$$

für ein geeignetes $\theta \in (0,1)$ und $c = a + \theta\cdot(x-a) \in \overline{ax}$.

Die Integralform des Restgliedes folgt mit dem gleichen Argument aus Korollar 20.14:

$$\begin{aligned} f(x) - T_{f,a}^k(x) = \; & f(x) - \sum_{|\alpha|=0}^k \frac{D^\alpha f(a)}{\alpha!} \cdot (x-a)^k \\ \overset{(85)}{=} \; & h(1) - \sum_{l=0}^k \frac{h^{(l)}(0)}{l!} \cdot (1-0)^l \\ \overset{20.14}{=} \; & \int_0^1 \frac{h^{(k+1)}(t)}{k!} \cdot (1-t)^k \, dt \\ \overset{(85)}{=} \; & (k+1) \cdot \int_0^1 (1-t)^k \cdot \sum_{|\alpha|=k+1} \frac{D^\alpha f(a+t\cdot(x-a))}{\alpha!} \cdot (x-a)^\alpha \, dt. \end{aligned}$$

□

Bemerkung 25.15 (Fehlerabschätzung mittels Taylorentwicklung)

a. Ist $f \in \mathcal{C}^{k+1}(U, \mathbb{R})$ und ist $x \in \overline{U_\varepsilon(a)} \subset U$, so ergibt sich aus der Taylorformel 25.13 und der Dreiecksungleichung folgende Fehlerabschätzung für das Restglied

$$\left|f(x) - T^k_{f,a}(x)\right| \leq \sum_{|\alpha|=k+1} \frac{\max\left\{|D^\alpha f(c)| \;\middle|\; c \in \overline{U_\varepsilon(a)}\right\}}{\alpha!} \cdot ||x-a||_2^{k+1},$$

da $|(x-a)^\alpha| = |x_1-a_1|^{\alpha_1} \cdot \ldots \cdot |x_n-a_n|^{\alpha_n} \leq ||x-a||_2^{\alpha_1} \cdot \ldots \cdot ||x-a||_2^{\alpha_n} = ||x-a||_2^{|\alpha|}$. Damit folgt dann aber unmittelbar

$$\lim_{x \to a} \frac{f(x) - T^k_{f,a}(x)}{||x-a||_2^k} = 0.$$

Man sagt auch, daß das Restglied für $x \to a$ mit Ordnung $k+1$ gegen 0 konvergiert.

b. Die Integralform des Restgliedes hat den Vorteil, daß sie ohne Änderung für Funktionen $f : U \longrightarrow \mathbb{R}^m$ 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 c_i.

Der Mittelwertsatz ist ein Spezialfall des Satzes von Taylor mit $k = 0$.

Korollar 25.16 (Mittelwertsatz)
Ist $f \in \mathcal{C}^1(U, \mathbb{R})$ und $x, a \in U$ mit $\overline{ax} \subseteq U$, so gibt es ein $c \in \overline{ax}$ mit

$$f(x) - f(a) = \langle \nabla f(c), x-a \rangle = Df(c) \circ (x-a).$$

Ferner gilt

$$f(x) - f(a) = \int_0^1 Df(a + t \cdot (x-a)) \circ (x-a)\, dt$$

D) Differenzierbare Abbildungen auf konvexen Mengen

Wir wollen nun den Mittelwertsatz anwenden, um zu zeigen, daß stetig differenzierbare Abbildungen auf kompakten konvexen Mengen Lipschitz-stetig sind.

Definition 25.17 (Konvexe Mengen)
Eine Teilmenge $C \subseteq \mathbb{R}^n$ heißt *konvex*, falls für je zwei Punkte $x, y \in C$ auch die Verbindungsstrecke $\overline{xy}$ in C enthalten ist.

Beispiel 25.18 (Eine kompakte konvexe Menge)
Eine abgeschlossene ε-Umgebung

$$\overline{U_\varepsilon(a)} = \left\{x \in \mathbb{R}^n \;\middle|\; ||x-a|| \leq \varepsilon\right\}$$

in $\mathbb{R}^n$ bezüglich einer beliebigen Norm $||\cdot||$ ist stets konvex und kompakt.

Für das folgende Korollar identifizieren wir den $\mathbb{R}$-Vektorraum $\mathrm{Mat}(m \times n, \mathbb{R})$ der $m \times n$-Matrizen mit $\mathbb{R}^{m \cdot n}$, so daß wir für $A = (a_{ij}) \in \mathrm{Mat}(m \times n, \mathbb{R})$ die euklidische Norm

$$||A||_2 = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij}^2}$$

haben und auch die Maximumsnorm

$$||A||_\infty = \max\left\{|a_{ij}| \;\middle|\; i = 1, \ldots, m, j = 1, \ldots, n\right\}.$$

Korollar 25.19 (Lipschitz-Stetigkeit auf kompakten konvexen Mengen)
Sei $U \subseteq \mathbb{R}^n$ offen, $C \subseteq U$ konvex und kompakt, $f : U \longrightarrow \mathbb{R}^m$ stetig differenzierbar.
Für $x, y \in C$ gilt dann stets

$$||f(x) - f(y)||_2 \leq ||x - y||_2 \cdot n^2 \cdot \max_{z \in C} ||Df(z)||_2$$

und

$$||f(x) - f(y)||_\infty \leq ||x - y||_\infty \cdot n \cdot \max_{z \in C} ||Df(z)||_\infty.$$

Insbesondere ist eine stetig-differenzierbare Abbildung auf einer kompakten konvexen Menge im $\mathbb{R}^n$ also Lipschitz-stetig.

Beweis: Aus dem Mittelwertsatz 25.16 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

$$||f(x) - f(y)||_\infty = \max_{i=1,\ldots,n} |f_i(x) - f_i(y)| \leq n \cdot \max_{z \in C} ||Df(z)||_\infty \cdot ||x - y||_\infty.$$

Aus Beispiel 22.42 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

$$\begin{aligned} ||f(x)-f(y)||_2 &\leq \sqrt{n}\cdot||f(x)-f(y)||_\infty \leq \sqrt{n}\cdot n\cdot \max_{z\in c}||Df(z)||_\infty \cdot ||x-y||_\infty \\ &\leq n^2\cdot \max_{z\in C}||Df(z)||_2\cdot||x-y||_2. \end{aligned}$$

□

Korollar 25.20 (Stetig differenzierbar impliziert lokal Lipschitz-stetig)
Eine stetig differenzierbare Abbildung $f: U \longrightarrow \mathbb{R}^m$ ist lokal Lipschitz-stetig.

Beweis: Jeder Punkt $x \in U$ hat eine kompakte konvexe Umgebung $\overline{U_\varepsilon(x)}$ für $\varepsilon > 0$ hinreichend klein, auf der f nach Korollar 25.19 Lipschitz-stetig ist. □

E) Der Satz von Taylor und die Hesse-Matrix

Auch der Spezialfall des Satzes von Taylor für $k = 1$ ist von Interesse. Dazu führen wir folgenden Begriff ein.

Definition 25.21 (Hesse-Matrix)
Ist $f \in \mathcal{C}^2(U, \mathbb{R})$ und $x \in U$, so nennen wir

$$H_f(x) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)_{i,j=1,\ldots,n} = \begin{pmatrix} D_1D_1f(x) & D_1D_2f(x) & \ldots & D_1D_nf(x) \\ D_2D_1f(x) & D_2D_2f(x) & \ldots & D_2D_nf(x) \\ \vdots & \vdots & & \vdots \\ D_nD_1f(x) & D_nD_2f(x) & \ldots & D_nD_nf(x) \end{pmatrix}$$

die *Hesse-Matrix* von f in x. Man beachte, daß $H_f(x)$ aufgrund des Satzes von Schwarz 25.5 eine *symmetrische* Matrix ist.

Beispiel 25.22
Für die Funktion f aus Beispiel 25.1 gilt

$$H_f(x) = \begin{pmatrix} 60x_1^3 + 12x_1x_2 & 6x_1^2 \\ 6x_1^2 & 0 \end{pmatrix}.$$

Korollar 25.23 (Satz von Taylor für $k = 1$)
Ist $f \in \mathcal{C}^2(U, \mathbb{R})$ und $x, a \in U$ mit $\overline{ax} \subseteq U$, so gibt es ein $c \in \overline{ax}$ mit

$$f(x) = f(a) + Df(a)\circ(x-a) + \frac{(x-a)^t \circ H_f(c) \circ (x-a)}{2}.$$

Beweis: Es gilt

$$\begin{aligned}\sum_{|\alpha|=2} \frac{D^\alpha f(c)}{\alpha!} \cdot (x-a)^\alpha &= \sum_{i=1}^n \frac{D_i^2 f(c)}{2} \cdot (x_i - a_i)^2 + \sum_{i=1}^n \sum_{j<i} \frac{D_i D_j f(c)}{1 \cdot 1} \cdot (x_i - a_i) \cdot (x_j - a_j) \\ &= \sum_{i=1}^n \frac{D_i^2 f(c)}{2} \cdot (x_i - a_i)^2 + \sum_{i=1}^n \sum_{j\neq i} \frac{D_i D_j f(c)}{2} \cdot (x_i - a_i) \cdot (x_j - a_j) \\ &= \frac{\sum_{i=1}^n \sum_{j=1}^n (x_j - a_j) \cdot D_i D_j f(c) \cdot (x_i - a_i)}{2} \\ &= \frac{(x-a)^t \circ H_f(c) \circ (x-a)}{2}.\end{aligned}$$

Der Rest folgt dann aus dem Satz von Taylor 25.13. □

F) Lokale Extrema im Mehrdimensionalen

Wir wollen nun die Bedingungen 18.2 und 18.15 für Extremstellen verallgemeinern.

Definition 25.24 (Lokale Extrema)
Sei $D \subseteq \mathbb{R}^n$, $f : D \longrightarrow \mathbb{R}$ und $a \in D$.

a. f hat in a ein *globales Maximum*, wenn $f(a) \geq f(x)$ für alle $x \in D$.

b. f hat in a ein *lokales Maximum*, wenn es ein $\delta > 0$ gibt, so daß $f(a) \geq f(x)$ für alle $x \in D \cap U_\delta(a)$.

c. f hat in a ein *isoliertes lokales Maximum*, wenn es ein $\delta > 0$ gibt mit $f(a) > f(x)$ für alle $a \neq x \in D \cap U_\delta(a)$.

d. f hat in a ein *globales Minimum*, wenn $f(a) \leq f(x)$ für alle $x \in D$.

e. f hat in a ein *lokales Minimum*, wenn es ein $\delta > 0$ gibt, so daß $f(a) \leq f(x)$ für alle $x \in D \cap U_\delta(a)$.

f. f hat in a ein *isoliertes lokales Minimum*, wenn es ein $\delta > 0$ gibt mit $f(a) < f(x)$ für alle $a \neq x \in D \cap U_\delta(a)$.

g. a heißt *Extremstelle* und $f(a)$ *Extremum* von f, wenn f in a ein lokales Minimum oder ein lokales Maximum hat.

h. Ist f in a differenzierbar mit $Df(a) = (0, \ldots, 0)$, so heißt a ein *kritischer Punkt* von f.

i. Ist a ein kritischer Punkt von f und gibt es in jeder ε-Umgebung von a Punkte x und y mit $f(x) > f(a) > f(y)$, so heißt a ein *Sattelpunkt* von f (siehe Abbildung 8 auf Seite 387) .

Proposition 25.25 (Extremstellen sind kritische Punkte)
Ist $f \in \mathcal{C}^1(U, \mathbb{R})$ und ist $a \in U$ eine Extremstelle von f, so ist $Df(a) = (0, \ldots, 0)$.

Beweis: Da U offen ist, gibt es ein $\varepsilon > 0$ mit $U_\varepsilon(a) \subseteq U$. Für $j \in \{1, \ldots, n\}$ betrachten wir die Funktion

$$g_j : (-\varepsilon, \varepsilon) \longrightarrow \mathbb{R} : t \mapsto f(a + t \cdot e_j).$$

Besitzt f ein lokales Maximum bzw. Minimum in 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

$$0 = g_j'(0) = \lim_{t \to 0} \frac{g_j(t) - g_j(0)}{t - 0} = \lim_{t \to 0} \frac{f(a + te_j) - f(a)}{t} = D_j f(a).$$

Dies zeigt, daß der Gradient von f der Nullvektor ist. □

Satz 25.26 (Hinreichende Bedingung für Extremstellen)
Es sei $f \in \mathcal{C}^2(U, \mathbb{R})$ und $a \in U$ mit $Df(a) = (0, \ldots, 0)$.

a. Ist die Hesse-Matrix $H_f(a)$ positiv definit, so hat f in a ein isoliertes lokales Minimum.

b. Ist die Hesse-Matrix $H_f(a)$ negativ definit, so hat f in a ein isoliertes lokales Maximum.

c. Ist die Hesse-Matrix $H_f(a)$ indefinit, so ist a ein Sattelpunkt von f und keine Extremstelle.

Beweis:

a. Für $k = 1, \ldots, n$ ist die Funktion

$$g_k : U \longrightarrow \mathbb{R} : x \mapsto \det \begin{pmatrix} D_1D_1f(x) & D_1D_2f(x) & \ldots & D_1D_kf(x) \\ D_2D_1f(x) & D_2D_2f(x) & \ldots & D_2D_kf(x) \\ \vdots & \vdots & & \vdots \\ D_kD_1f(x) & D_kD_2f(x) & \ldots & D_kD_kf(x) \end{pmatrix}$$

stetig, da die Funktion f nach Voraussetzung zweifach stetig differenzierbar ist. Mithin sind die Mengen

$$g_k^{-1}\big((0, \infty)\big)$$

als Urbilder offener Mengen unter stetigen Abbildungen nach Proposition 23.17 offen in U für $k = 1, \ldots, n$, und damit ist auch

$$O = \bigcap_{k=1}^{n} g_k^{-1}\big((0, \infty)\big)$$

als Schnitt offener Mengen wieder offen.

Man beachte nun, daß $g_k(x)$ gerade der k-te Hauptminor der Hesse-Matrix $H_f(x)$ ist. Aus dem Hurwitz-Kriterium für positive Definitheit A15.37 wissen wir, daß $g_k(a) > 0$ für alle $k = 1, \ldots, n$ gilt, da die symmetrische Matrix $H_f(a)$ positiv definit ist, und damit gilt also

$$a \in O.$$

Da O offen ist, gibt es mithin ein $\delta > 0$, so daß

$$U_\delta(a) \subseteq O.$$

Sei nun $a \neq x \in U_\delta(a)$ beliebig, so gilt

$$\overline{ax} \subset U_\delta(a) \subseteq O \subseteq U,$$

so daß wir den Satz von Taylor in der Fassung von Korollar 25.23 anwenden können. Wir erhalten

$$\begin{aligned} f(x) =& f(a) + Df(a) \circ (x-a) + \frac{(x-a)^t \circ H_f(c) \circ (x-a)}{2} \\ =& f(a) + \frac{(x-a)^t \circ H_f(c) \circ (x-a)}{2} > f(a), \end{aligned}$$

denn $c \in \overline{ax} \subset O$ und mithin hat $H_f(c)$ nur positive Hauptminoren, so daß $H_f(c)$ wegen des Hurwitz-Kriteriums A15.37 positiv definit ist.

b. Ist $H_f(a)$ negativ definit, dann ist $H_{-f}(a) = -H_f(a)$ positiv definit. Aus Teil a. folgt also, daß $-f$ ein isoliertes lokales Minimum in a hat. Dann hat f dort aber ein isoliertes lokales Maximum.

c. Ist $H_f(a)$ indefinit, so gibt es $x, y \in \mathbb{R}^n$ mit

$$x^t \circ H_f(a) \circ x > 0 > y^t \circ H_f(a) \circ y. \tag{87}$$

Wir müssen zeigen, daß f in jeder ε-Umgebung von a Werte größer und kleiner als $f(a)$ annimmt.
Sei also $\varepsilon > 0$ gegeben. Da U offen ist und wir nur kleine ε betrachten müssen, können wir ohne Einschränkung annehmen, daß

$$U_\varepsilon(a) \subseteq U$$

gilt. Für $\delta := \min\left\{\frac{\varepsilon}{||x||_2}, \frac{\varepsilon}{||y||_2}\right\} > 0$ und $t \in (-\delta, \delta)$ gilt $a + tx, a + ty \in U_\varepsilon(a)$.
Wir betrachten nun zunächst die Funktion

$$g : (-\delta, \delta) \longrightarrow \mathbb{R} : t \mapsto x^t \circ H_f(a + tx) \circ x.$$

Sie ist stetig, da f zweifach stetig differenzierbar auf U ist. Außerdem gilt $g(0) > 0$ wegen (87), so daß mit Aufgabe 14.33 ein $\gamma > 0$ existiert, so daß

$$0 < g(t) = x^t \circ H_f(a + tx) \circ x$$

für alle $t \in (-\gamma, \gamma)$. Wenden wir nun den Satz von Taylor in der Fassung von Korollar 25.23 auf a und $a + tx$ für $t \in (-\gamma, \gamma)$ an, so finden wir ein c_t auf der

Strecke von a nach $a+tx$, d.h. wir finden ein θ_t mit $\theta_t \in (-\gamma,\gamma)$ und $c_t = a+\theta_t x$, mit

$$\begin{aligned} f(a+tx) =& f(a) + Df(a)\circ(a+tx-a) + \frac{(a+tx-a)^t \circ H_f(c_t)\circ(a+tx-a)}{2} \\ =& f(a) + \frac{t^2}{2}\cdot x^t \circ H_f(a+\theta_t x)\circ x = f(a) + \frac{t^2}{2}\cdot g(\theta_t) > f(a), \end{aligned}$$

da $\theta_t \in (-\gamma,\gamma)$. Also gibt es in der ε-Umgebung von a einen Punkt $a+tx$, so daß

$$f(a+tx) > f(a).$$

Analog findet man ein Intervall $(-\beta,\beta)$, so daß $a+ty \in U_\varepsilon(a)$ und

$$f(a+ty) < f(a)$$

für alle $t \in (-\beta,\beta)$. Damit ist gezeigt, daß a ein Sattelpunkt und keine Extremstelle von f ist.

□

Bemerkung 25.27

a. Man beachte, daß die drei in Satz 25.26 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 $x = (0,0)^t$, da $Df(x) = (2x_1, 0)$ ist, und sie hat in $x = (0,0)^t$ 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 25.26 dann keine Aussage zu den Extremstellen von f macht, wenn $H_f(a)$ für den kritischen Punkt a den Eigenwert 0 hat.

c. Die Definitheit einer symmetrischen Matrix überprüft man am Besten mit dem Hurwitz-Kriterium A15.37 durch Berechnen der Hauptminoren. Sei $A \in \mathrm{Mat}_n(\mathbb{R})$, 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 25.28

Wir wollen nun zwei Funktionen betrachten, bei denen wir Satz 25.26 für die Untersuchung der kritischen Punkte anwenden können.

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 7). Es ist in der Tat sogar ein globales Minimum, da $f(x) = x_1^2 + x_2^2 > 0 = f(0,0)$ für alle $x \neq (0,0)^t$.

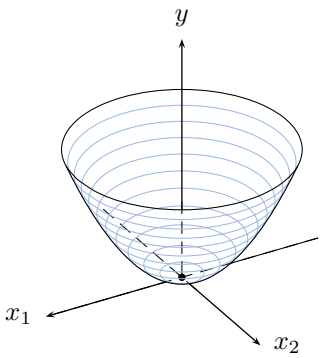


Abbildung 7: Das isolierte Minimum von $x \mapsto x_1^2 + x_2^2$

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 25.26 sagt also, daß $(0,0)^t$ keine Extremstelle ist, und Bemerkung 25.27 sagt etwas genauer, daß f in $(0,0)^t$ einen Sattelpunkt hat (siehe Abbildung 8).

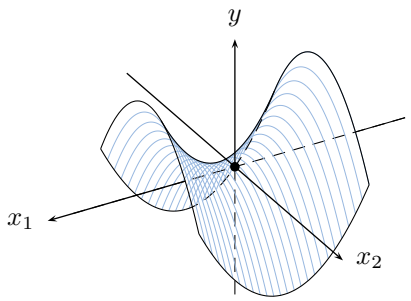


Abbildung 8: Der Sattelpunkt von $x \mapsto x_1^2 - x_2^2$

G) Potenzreihen für Matrizen

Mit Blick auf Anwendungen bei der Lösung von Systemen linearer Differentialgleichungen in Abschnitt 31 wollen wir Ausdrücken der Form e^{At} für eine quadratische Matrix A eine Bedeutung geben und deren Eigenschaften in Abhängigkeit von t untersuchen. Als erster Schritt in die Richtung verallgemeinern wir zunächst Aussagen zur Vertauschbarkeit von Grenzwert und Integration bzw. Ableitung bei Funktionenfolgen in den $\mathbb{R}^N$ durch Rückführung auf die Komponentenfunktionen. Dabei formulieren wir die Aussagen nicht in der Allgemeinheit, wie wir sie im Eindimensionalen kennengelernt haben, sondern mit spezielleren Voraussetzungen, die für unsere Anwendungen gut genug sind.

Definition 25.29 (Komponentenweises Integral)
Für eine stetige Abbildung $h : I \longrightarrow \mathbb{R}^n$ für ein Intervall I und $a, b \in I$ definieren wir das Integral

$$\int_a^b h(s)\, ds := \left(\int_a^b h_1(s)\, ds, \ldots, \int_a^b h_n(s)\, ds \right)^t \in \mathbb{R}^n$$

als den Vektor der Integrale über die Komponentenfunktionen.

Proposition 25.30 (Vertauschbarkeit von Grenzwert und Integration / Ableitung)
Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge stetiger Funktionen $f_n : [a, b] \longrightarrow \mathbb{R}^N$ die auf $[a, b]$, $a < b$, gleichmäßig gegen $f : [a, b] \longrightarrow \mathbb{R}^N$ konvergiere.

a. Dann ist f stetig und es gilt

$$\int_a^b f(t)\, dt = \lim_{n\to\infty} \int_a^b f_n(t)\, dt.$$

b. Sind die f_n stetig differenzierbar und ist $(f'_n)_{n\in\mathbb{N}}$ auf $[a,b]$ gleichmäßig konvergent, dann ist f stetig differenzierbar

$$f' = \lim_{n\to\infty} f'_n.$$

Beweis: Man beachte, daß Konvergenz im $\mathbb{R}^N$ komponentenweise Konvergenz ist (siehe Korollar 22.38); zudem beachte man, daß die gleichmäßige Konvergenz der Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ auch die gleichmäßige Konvergenz der Komponentenfunktionen impliziert, weil der Betrag der Komponenten eines Vektors stets durch die euklidische Norm des Vektors nach oben beschränkt ist. Mithin folgt die Aussagen aus den entsprechenden Aussage im Eindimensionalen (siehe Satz 18.18 und Satz 20.23) angewandt auf die Komponentenfunktionen. □

Proposition 25.31 (Potenzreihen von Matrizen)
Sei $g = \sum_{k=0}^{\infty} a_k \cdot t^k$ eine Potenzreihe in $\mathbb{R}$ mit Konvergenzradius r und sei

$$g : (-r,r) \longrightarrow \mathbb{R} : s \mapsto \sum_{k=0}^{\infty} a_k \cdot s^k$$

die durch die Potenzreihe definierte Funktion ist. Ferner sei $||\cdot||$ eine submultiplikative Norm auf $\mathrm{Mat}_n(\mathbb{R})$ und $A \in \mathrm{Mat}_n(\mathbb{R})$.
Dann ist die Reihe

$$g(s\cdot A) = \sum_{k=0}^{\infty} a_k \cdot s^k \cdot A^k$$

für alle $s \in \mathbb{R}$ mit $||s\cdot A|| < r$ absolut konvergent in $\mathrm{Mat}_n(\mathbb{R})$ bezüglich $||\cdot||$ und die Abbildung

$$f : \left(-\tfrac{r}{||A||}, \tfrac{r}{||A||}\right) \longrightarrow \mathrm{Mat}_n(\mathbb{R}) : s \mapsto g(s\cdot A) = \sum_{k=0}^{\infty} a_k \cdot s^k \cdot A^k$$

ist unendlich oft differenzierbar mit Ableitung

$$f'(s) = A \circ \sum_{k=1}^{\infty} k \cdot a_k \cdot s^{k-1} \cdot A^{k-1} = A \circ g'(s\cdot A) = g'(s\cdot A) \circ A.$$

Beweis: Da die Norm auf $\mathrm{Mat}_n(\mathbb{R})$ submultiplikativ ist, gilt

$$\sum_{k=0}^{\infty} ||a_k \cdot s^k \cdot A^k|| \le \sum_{k=0}^{\infty} |a_k| \cdot ||s\cdot A||^k$$

und die rechte Reihe ist aufgrund der Definition des Konvergenzradius für $s \in \mathbb{R}$ mit $||s\cdot A|| < r$ konvergent. Also ist auch die linke Reihe konvergent, da die rechte eine konvergente Majorante ist. Damit ist die Reihe $g(s\cdot A)$ dann aber absolut konvergent und damit konvergent für diese s.

Sei nun $0 < R < \frac{r}{||A||}$ beliebig gegeben und betrachten wir die Abbildungen

$$f_m : [-R, R] \longrightarrow \mathrm{Mat}_n(\mathbb{R}) : s \mapsto \sum_{k=0}^{m} a_k \cdot s^k \cdot A^k.$$

Da die Komponentenfunktionen Polynomfunktionen in s sind, sind sie stetig differenzierbar mit Ableitung

$$f'_m : [-R, R] \longrightarrow \mathrm{Mat}_n(\mathbb{R}) : s \mapsto \sum_{k=1}^{m} k \cdot a_k \cdot s^{k-1} \cdot A^k.$$

Wir wissen aus der eindimensionalen Analysis, daß die Folge

$$h_m : [-R, R] \longrightarrow \mathbb{R} : s \mapsto \sum_{k=0}^{m} |a_k| \cdot ||s \cdot A||^k$$

gleichmäßig auf $[-R, R]$ gegen

$$h : [-R, R] \longrightarrow \mathbb{R} : s \mapsto \sum_{k=0}^{\infty} |a_k| \cdot ||s \cdot A||^k$$

konvergiert (siehe Satz 15.4). Für $s \in [-R, R]$ erhalten wir nun aus Proposition 22.51 und unter Beachtung der Submultiplikativität der Norm

$$||f_m(s) - f(s)|| = \left|\left| \sum_{k=m+1}^{\infty} a_k \cdot s^k \cdot A^k \right|\right| \leq \sum_{k=m+1}^{\infty} ||a_k \cdot s^k \cdot A^k|| \leq \sum_{k=m+1}^{\infty} |a_k| \cdot ||s \cdot A||^k.$$

Da die Folge $(h_m)_{m\in\mathbb{N}}$ gleichmäßig gegen h konvergiert, konvergiert die rechte Seite der Ungleichung auf $[-R, R]$ gleichmäßig in s gegen 0, und somit tut dies auch die linke Seite. Analog sieht man, daß f'_m auf $[-R, R]$ gleichmäßig gegen

$$[-R, R] \longrightarrow \mathrm{Mat}_n(\mathbb{R}) : s \mapsto \sum_{k=1}^{\infty} k \cdot a_k \cdot s^{k-1} \cdot A^k$$

konvergiert. Aus Proposition 25.30 folgt dann, daß f stetig differenzierbar mit

$$f'(s) = \sum_{k=1}^{\infty} k \cdot a_k \cdot s^{k-1} \cdot A^k = A \circ \sum_{k=1}^{\infty} k \cdot a_k \cdot s^{k-1} \cdot A^{k-1}$$

ist, wobei wir für die letzte Gleichheit beachten, daß die Multiplikation mit einer Matrix eine stetige Operation ist und deshalb mit der Grenzwertbildung vertauscht. Da die Aussage für jedes $0 < R < \frac{r}{||A||}$ gilt, ist f auf $\left(-\frac{r}{||A||}, \frac{r}{||A||}\right)$ stetig differenzierbar mit der angegebenen Ableitung.

Per Induktion zeigt man dann, daß f in der Tat unendlich oft differenzierbar ist. □

H) Die Exponentialfunktion für Matrizen

Wir wollen nun Proposition 25.31 anwenden, um die Exponentialfunktion für Matrizen einzuführen.

Korollar 25.32 (Die Exponentialfunktion)
Sei $A \in \mathrm{Mat}_n(\mathbb{R})$.

a. Die Abbildung

$$f : \mathbb{R} \longrightarrow \mathrm{Mat}_n(\mathbb{R}) : t \mapsto e^{At} := \sum_{k=0}^{\infty} \frac{A^k \cdot t^k}{k!}$$

ist unendlich oft differenzierbar mit

$$f'(t) = A \circ e^{At} = e^{At} \circ A.$$

b. Für $T \in \mathrm{Gl}_n(\mathbb{R})$ gilt

$$e^{T^{-1} \circ A \circ T} = T^{-1} \circ e^A \circ T.$$

c. Für $A, B \in \mathrm{Mat}_n(\mathbb{R})$ mit $A \circ B = B \circ A$ gelten zudem

$$A \circ e^B = e^B \circ A$$

und

$$e^{A+B} = e^A \circ e^B = e^B \circ e^A.$$

d. Ferner gelten für $s, t \in \mathbb{R}$ und $0 \in \mathrm{Mat}_n(\mathbb{R})$

- $||e^{At}|| \leq e^{||A|| \cdot |t|}$,
- $e^0 = \mathbb{1}_n$,
- $e^{A \cdot (s+t)} = e^{As} \circ e^{At}$,
- $e^{-A} = (e^A)^{-1}$,
- $e^{\mathbb{1}_n \cdot t} = e^t \cdot \mathbb{1}_n$.

Beweis: a. Die Aussage folgt aus Proposition 25.31 mit der Potenzreihe der Exponentialfunktion.

b. Man sieht leicht, daß

$$(T^{-1} \circ A \circ T)^k = T^{-1} \circ A^k \circ T$$

gilt. Da die Multiplikation mit einer Matrix eine stetige Abbildung ist und deshalb mit Grenzwertbildung vertauscht, folgt dann

$$T^{-1} \circ e^{At} \circ T = \sum_{k=0}^{\infty} \frac{T^{-1} \circ A^k \circ T}{k!} = \sum_{k=0}^{\infty} \frac{(T^{-1} \circ A \circ T)^k}{k!} = e^{T^{-1} \circ A \circ T}.$$

c. Da die Multiplikation mit einer Matrix stetig ist und deshalb mit der Grenzwertbildung vertauscht, erhalten wir

$$A \circ e^B = A \circ \sum_{k=0}^{\infty} \frac{B^k}{k!} = \sum_{k=0}^{\infty} \frac{A \circ B^k}{k!} = \sum_{k=0}^{\infty} \frac{B^k \circ A}{k!} = \sum_{k=0}^{\infty} \frac{B^k}{k!} \circ A = e^B \circ A.$$

Die Formel $e^{A+B} = e^A \circ e^B$ folgt wörtlich wie die analoge Aussage in den reellen Zahlen (siehe Satz 12.36).

d. Die Ungleichung folgt aus Proposition 22.51, die übrigen Formeln folgen unmittelbar aus c. sowie der Definition der Exponentialfunktion.

□

Aufgaben

Aufgabe 25.33 (Vertauschbarkeit der Integrationsreihenfolge)
Es sei $U \subseteq \mathbb{R}^2$ offen, $f : U \longrightarrow \mathbb{R}$ stetig und $[a,b] \times [c,d] \subset U$. Zeige die folgenden Aussagen:

a. Die Funktion $H : [a,b] \times [c,d] \longrightarrow \mathbb{R} : (x,y)^t \mapsto \int_a^x f(s,y)\, ds$ ist stetig.

b. Falls $\frac{\partial f}{\partial y} : U \longrightarrow \mathbb{R}$ existiert und stetig ist, so die Abbildung

$$F : [c,d] \longrightarrow \mathbb{R} : y \mapsto \int_a^b f(s,y)\, ds$$

stetig differenzierbar auf $[c,d]$ mit

$$F'(y) = \int_a^b \frac{\partial f}{\partial y}(s,y)\, ds.$$

c. Es gilt

$$\int_a^b \int_c^d f(s,t)\, dt\, ds = \int_c^d \int_a^b f(s,t)\, ds\, dt.$$

d. 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 25.34 (Beispiel zum Satz von Schwarz)
Zeige, für die Funktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto \begin{cases} \frac{x_1^3 x_2 - x_1 x_2^3}{x_1^2 + x_2^2}, & \text{falls } (x_1,x_2)^t \neq (0,0)^t, \\ 0, & \text{sonst} \end{cases}$$

existieren die beiden partiellen Ableitungen $D_1D_2f(0,0)$ und $D_2D_1f(0,0)$, stimmen aber nicht überein.

Aufgabe 25.35 (Leibniz-Regel)

Sind $f, g \in \mathcal{C}^k(U, \mathbb{R})$ und ist $|\alpha| \leq k$, so gilt für $x \in U$

$$D^\alpha(f \cdot g)(x) = \sum_{\substack{\beta \in \mathbb{N}^n \\ \beta \leq \alpha}} \binom{\alpha}{\beta} \cdot D^\beta f(x) \cdot D^{\alpha-\beta} g(x).$$

Aufgabe 25.36 (Verbesserte Abschätzung für Korollar 25.19)

a. Ist $g : [a, b] \longrightarrow \mathbb{R}^m$ eine stetige Funktion und definieren wir das Integral

$$\int_a^b g(t)\, dt = \left(\int_a^b g_1(t)\, dt, \ldots, \int_a^b g_n(t)\, dt \right)^t \in \mathbb{R}^n$$

als den Vektor der Integrale über die Koordinatenfunktionen, so gilt

$$\left|\left| \int_a^b g(t)\, dt \right|\right|_2 \leq \int_a^b ||g(t)||_2\, dt.$$

b. Ist $U \subseteq \mathbb{R}^n$ offen, $\overline{U_r(a)} \subseteq U$ und $h : U \longrightarrow \mathbb{R}^m$ stetig differenzierbar auf U, dann gilt für $x, y \in \overline{U_r(a)}$ stets

$$||h(x) - h(y)||_2 \leq ||x - y||_2 \cdot \sup_{z \in \overline{U_r(a)}} ||Dh(z)||_2.$$

Aufgabe 25.37

a. Es sei $U \subseteq \mathbb{R}^n$ offen, $f \in \mathcal{C}^2(U, \mathbb{R})$ und $a \in U$. Zeige, für das Restglied $r_a(h) := f(a+h) - f(a) + Df(a) \circ h$ im Satz von Taylor gilt die Abschätzung

$$|r_a(h)| \leq \frac{||h||_\infty^2 \cdot \sup\left\{ ||H_f(x)||_1 \;\middle|\; x \in \overline{a\ a+h} \right\}}{2}.$$

b. Wende die Abschätzung in Teil a. auf die Funktion $f : (\mathbb{R}_{>0})^2 \longrightarrow \mathbb{R} : (x, y)^t \mapsto x^y$ mit $a = (1, 1)^t$ an, um die Zahl $(1,05)^{1,02}$ näherungsweise zu berechnen. Bis auf wieviele Nachkommastellen genau ist die Abschätzung?

Aufgabe 25.38

Bestimme alle lokalen Extrema und Sattelpunkte der Funktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y)^t \longmapsto x^4 + 2x^2y^2 + y^4 - 2x^3 - 2xy^2 + x^2 + y^2.$$

Aufgabe 25.39

Bestimme das zweite Taylor-Polynom $T^2_{f,a}$ im Entwicklungspunkt $a = (0, \pi)^t$ für

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R},\ (x, y)^t \longmapsto \sin(y) - x(x + \sin(y) + 1).$$

Aufgabe 25.40

Sei $a \in \mathbb{R}^n$, $\epsilon > 0$ und $f : U_\epsilon(a) \longrightarrow \mathbb{R}^m$ eine differenzierbare Abbildung mit $Df(x) = 0$ für alle $x \in U_\epsilon(a)$. Zeige, daß f konstant ist.

Aufgabe 25.41

Bestimme das sechste Taylor-Polynom $T^6_{f,a}$ für $a = (0, 0, 0)^t$ und

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R} : (x, y, z)^t \longmapsto (x + y^3) \cdot \cos\left(\frac{z}{1 + y^2}\right).$$

§ 26 Der Satz über implizite Funktionen

A) Der Banachsche Fixpunktsatz

Definition 26.1 (Strikte Kontraktion)
Sei (M, d) ein metrischer Raum.

a. Eine Lipschitz stetige Selbstabbildung $f : M \longrightarrow M$ mit Lipschitzkonstante $q < 1$ heißt eine *strikte Kontraktion* oder kurz eine *q-Kontraktion*.

b. $x \in M$ heißt ein *Fixpunkt* der Selbstabbildung $f : M \longrightarrow M$, wenn $f(x) = x$.

Bemerkung 26.2 (Strikte Kontraktionen über $\mathbb{R}$)
Eine differenzierbare Funktion $f : \mathbb{R} \longrightarrow \mathbb{R}$ 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

$$|f'(x)| = \lim_{y \to x} \frac{|f(y) - f(x)|}{|y - x|} \leq q$$

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. □

Beispiel 26.3 (Strikte Kontraktion)
Die Funktion $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \frac{\sin(x)}{2}$ 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 9). Deshalb hat f genau einen Fixpunkt, nämlich $x = 0$.

Der folgende Satz zeigt, daß es kein Zufall war, daß die Funktion f in Beispiel 26.3 genau einen Fixpunkt hat (siehe auch Aufgabe 14.39).

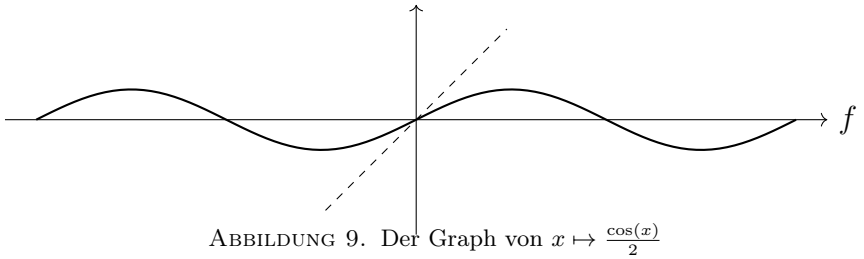


ABBILDUNG 9. Der Graph von $x \mapsto \frac{\cos(x)}{2}$

Satz 26.4 (Fixpunktsatz von Banach)
Sei (M, d) ein vollständiger metrischer Raum, $f : M \to M$ eine strikte Kontraktion.
Dann hat f genau einen Fixpunkt x und für jedes $y \in M$ gilt $\lim\limits_{n\to\infty} f^n(y) = x$.

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 $0 \le q < 1$, so daß

$$d\big(f(u), f(v)\big) \le q \cdot d(u, v) \tag{88}$$

für alle $u, v \in M$ gilt. Mit Hilfe von Induktion nach n leiten wir daraus her, daß

$$d(y_{n+1}, y_n) \le q^n \cdot d(y_1, y_0). \tag{89}$$

Für den Induktionsanfang $n = 0$ ist nichts zu zeigen, und der Induktionsschritt $n-1 \mapsto n$ folgt dann sofort:

$$\begin{aligned} d(y_{n+1}, y_n) &= d\big(f(y_n), f(y_{n-1})\big) \overset{(88)}{\le} q \cdot d(y_n, y_{n-1}) \\ &\overset{Ind.}{\le} q \cdot q^{n-1} \cdot d(y_1, y_0) = q^n \cdot d(y_1, y_0). \end{aligned}$$

Damit ist (89) gezeigt. Mit Induktion nach m folgt für ein festes $n \in \mathbb{N}$ nun daraus wiederum

$$d(y_{n+m}, y_n) \le \sum_{k=n}^{n+m-1} q^k \cdot d(y_1, y_0). \tag{90}$$

Für den Induktionsanfang $m = 1$ ist (90) gerade (89), und für den Induktionsschritt $m - 1 \mapsto m$ folgt mit der Dreiecksungleichung sofort

$$\begin{aligned} d(y_{n+m}, y_n) &\overset{DUGL}{\le} d(y_{n+m}, y_{n+m-1}) + d(y_{n+m-1}, y_n) \\ &\overset{Ind./(89)}{\le} q^{n+m-1} \cdot d(y_1, y_0) + \sum_{k=n}^{n+m-2} q^k \cdot d(y_1, y_0) \\ &= \sum_{k=n}^{n+m-1} q^k \cdot d(y_1, y_0) \end{aligned}$$

Die Ungleichung (90) führt zu folgender Abschätzung

$$d(y_{n+m}, y_n) \leq \sum_{k=n}^{n+m-1} q^k \cdot d(y_1, y_0) \leq q^n \cdot \sum_{k=0}^{\infty} q^k \cdot d(y_1, y_0) = \frac{q^n}{1-q} \cdot d(y_1, y_0). \tag{91}$$

Da die rechte Seite für $n \to \infty$ 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

$$x := \lim_{n\to\infty} y_n \in M$$

existiert. Nach Proposition 23.33 ist die Kontraktion f stetig, so daß aus der Eindeutigkeit des Grenzwertes

$$f(x) \longleftarrow f(y_n) = y_{n+1} \longrightarrow x$$

für $n \to \infty$ 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

$$d(x, z) = d\big(f(x), f(z)\big) \leq q \cdot d(x, z),$$

so daß $d(x, z) = 0$ wegen $q < 1$ gelten muß. Mithin ist $x = z$ und die Eindeutigkeit ist gezeigt. □

Bemerkung 26.5

Aus dem Beweis des Banachschen Fixpunktsatzes ergeben sich unmittelbar Abschätzungen, die über das Konvergenzverhalten der Folge $(y_n)_{n\in\mathbb{N}}$ mit $y_n = f^n(y)$ Auskunft geben und die dazu verwendet werden können, den Grenzwert der Folge und damit den eindeutigen Fixpunkt näherungsweise zu berechnen unter Angabe einer Schranke für den gemachten Fehler.

Bildet man in Gleichung (91) auf der linken Seite den Grenzwert für $m \to \infty$, so erhält man die *a priori Abschätzung*

$$d(x, y_n) \leq \frac{q^n}{1-q} \cdot d(y_1, y_0),$$

die nur die ersten beiden Iterationsschritte für die Fehlerabschätzung benötigt. Man kann den maximalen Fehler nach n Iterationsschritten also angeben, ohne dazu den Wert $y_n = f^n(y)$ zu berechnen.

Analog kann man auch die sogenannte *a posteriori Abschätzung*

$$d(x, y_n) \leq \frac{q}{1-q} \cdot d(y_n, y_{n-1})$$

herleiten, die im allgemeinen eine bessere Fehlerschranke liefert, für die aber die Berechnung von y_n erforderlich ist.

Beispiel 26.6 (Das Heron-Verfahren)
In Beispiel 11.23 haben wir das Heron-Verfahren zur näherungsweisen Berechnung der Wurzel $\sqrt{c}$ einer positiven reellen Zahl c kennen gelernt, da $\sqrt{c}$ der Grenzwert einer Folge mit der Rekursionsvorschrift $a_{n+1} = \frac{1}{2} \cdot \left(a_n + \frac{c}{a_n}\right)$ ist. Wir wollen dieses Beispiel nun noch einmal mit Hilfe des Banachschen Fixpunktsatzes untersuchen.

Offenbar gilt stets

$$x^2 - 2 \cdot \sqrt{\frac{c}{2}} \cdot x + \frac{c}{2} = \left(x - \sqrt{\frac{c}{2}}\right)^2 \geq -\frac{c}{2}$$

und mithin

$$x^2 + c \geq 2 \cdot \sqrt{\frac{c}{2}} \cdot x$$

sowie für positive x also

$$\frac{1}{2} \cdot \left(x + \frac{c}{x}\right) \geq \sqrt{\frac{c}{2}}.$$

Damit erhalten wir insbesondere, daß die differenzierbare Funktion

$$f : \left[\sqrt{\frac{c}{2}}, \infty\right) \longrightarrow \left[\sqrt{\frac{c}{2}}, \infty\right) : x \mapsto \frac{1}{2} \cdot \left(x + \frac{c}{x}\right)$$

eine Selbstabbildung ist. Zudem ist das Intervall als abgeschlossene Teilmenge von $\mathbb{R}$ ein vollständiger metrischer Raum. Außerdem ist

$$f'(x) = \frac{1}{2} \cdot \left(1 - \frac{c}{x^2}\right)$$

die Ableitung von f und für x im obigen Intervall gilt

$$-\frac{1}{2} \leq f'(x) \leq \frac{1}{2},$$

wie man leicht nachrechnet. Mit Hilfe des Mittelwertsatzes sieht man dann wie in Bemerkung 26.2, daß f eine $\frac{1}{2}$-Kontraktion ist, und aus dem Banachschen Fixpunktsatz folgt deshalb, daß die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n = f^n(a_0)$ für jedes $a_0 \geq \sqrt{\frac{c}{2}}$ gegen den eindeutig bestimmten Fixpunkt $\sqrt{c}$ der Funktion f konvergiert. Man beachte dabei, daß die Folge der Rekusionsvorschrift

$$a_{n+1} = f(a_n) = \frac{1}{2} \cdot \left(a_n + \frac{c}{a_n}\right)$$

des Heron-Verfahrens genügt.

Ist nun $c \leq 2$, so gilt $1 \geq \sqrt{\frac{c}{2}}$ und $a_0 = 1$ wie in Beispiel 11.23 ist ein zulässiger Startwert gemäß unserer obigen Betrachtungen, so daß wir für diese Fälle hier einen alternativen Beweis von Beispiel 11.23 gegeben haben. Die obigen Betrachtungen zeigen aber, daß man auch mit einem beliebigen anderen hinreichend großen Startwert hätte beginnen können.

Mit dem Startwert $a_0 = 1$ erhalten wir aus der a priori Abschätzung für $c = 2$ übrigens, daß a_n den Wert $\sqrt{2}$ mit einem Fehler von höchstens $\frac{1}{2^n}$ approximiert.

B) Gleichungen auflösen nach Variablen

Das Ziel dieses Abschnitts besteht darin, einzelne Gleichungen und ganze Gleichungssysteme (zumindest lokal) nach einzelnen Variablen aufzulösen. Wir werden zunächst einige Begriffe einführen und dann motivierende Beispiele betrachten, bevor wir uns mit dem Satz über implizite Funktionen einem der schwierigsten Sätze der Vorlesung zuwenden.

Definition 26.7 (Eine Gleichung nach bestimmten Variablen auflösen.)
Es sei $D \subseteq \mathbb{R}^N$ und $f : D \longrightarrow \mathbb{R}^m$.

a. Wir nennen
$$V(f) = \{x \in D \mid f(x) = 0\}$$
die *Verschwindungsmenge* oder *Nullstellenmenge* von f oder auch die Lösungsmenge des Gleichungssystems

$$\begin{aligned} f_1(x) &= 0 \\ &\vdots \\ f_m(x) &= 0. \end{aligned} \tag{92}$$

b. Ist $N = n + m$, so heißt das Gleichungssystem (92) *nach den Variablen* $x_{n+1}, \ldots, x_{n+m}$ *auflösbar*, wenn es eine Abbildung $\varphi : U \longrightarrow \mathbb{R}^m$ mit $U \subseteq \mathbb{R}^n$ gibt, so daß
$$V(f) = \operatorname{Graph}(\varphi) = \big\{\big(y, \varphi(y)\big)^t \mid y \in U\big\},$$
d.h. wenn $V(f)$ der *Graph* einer Abbildung φ ist.

Bemerkung 26.8 (Eine Gleichung nach bestimmten Variablen auflösen.)
Ist das Gleichungssystem (92) *nach den Variablen* $x_{n+1}, \ldots, x_{n+m}$ *auflösbar* durch die Abbildung $\varphi : U \longrightarrow \mathbb{R}^m$, so gilt insbesondere

$$f(y, \varphi(y)) = 0$$

für alle $y \in U$ und für einen Punkt $(x_1, \ldots, x_{n+m})^t \in V(f)$ in der Lösungsmenge des Gleichungssystems gilt dann

$$x_{n+i} = \varphi_i(x_1, \ldots, x_n)$$

für $i = 1, \ldots, m$. In diesem Sinne ist es also gelungen, das Gleichungssystem nach den letzten m Variablen aufzulösen.

Man kann in analoger Weise definieren, was es heißt, das Gleichungssystem nach einem anderen Satz von m Variablen $x_{i_1}, \ldots, x_{i_m}$ aufzulösen. Dies hinzuschreiben ist nur etwas technischer und deshalb verzichten wir hier darauf.

Motivation 26.9 (Eine Gleichung nach bestimmten Variablen auflösen.) Wir betrachten die Gleichung des Newtonschen Knotens

$$f(x) = x_2^2 - x_1^2 - x_1^3 = 0.$$

Ein Blick auf die Lösungsmenge $V(f)$ (siehe Abbildung 10) zeigt, daß man sie nicht

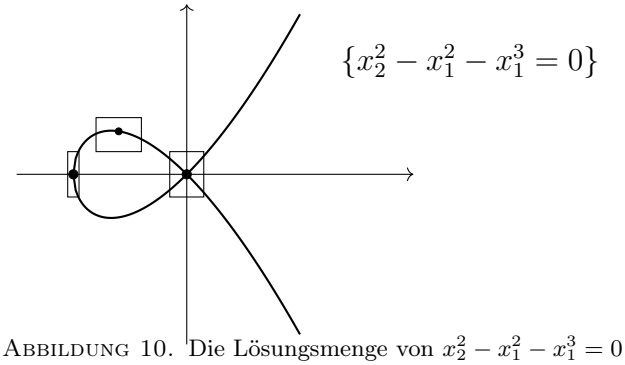


Abbildung 10. Die Lösungsmenge von $x_2^2 - x_1^2 - x_1^3 = 0$

global als den Graphen einer Funktion in der Variablen x_1 schreiben kann, daß man sie also nicht global nach x_2 auflösen kann. Dies sieht man auch analytisch, da der Versuch, die Gleichung nach x_2 aufzulösen zu der nicht eindeutigen Lösung

$$x_2 = \pm\sqrt{x_1^2 + x_1^3}, \tag{93}$$

führt. Wir können also bestenfalls hoffen, $V(f)$ lokal in einer kleinen Umgebung eines Punktes in $V(f)$ als den Graphen einer Funktion schreiben zu können (siehe die kleinen Kästen in Abbildung 10). Ein erneuter Blick auf die Lösungsmenge $V(f)$ zeigt auch, in welchen Punkten $p = (a, b)$ wir das wohl hinbekommen werden und in welchen nicht, nämlich nur dann, wenn der $b \neq 0$ gilt. Für diese Punkte liefert jeweils einer der beiden Wurzelzweige in Gleichung (93) eine lokale Auflösung nach x_2.

Die schlechten Punkte sind also die beiden Punkte $p = (-1, 0)$ und $q = (0, 0)$. Betrachtet man die Ableitung von f in einem Punkt $(a, b) \in V(f)$,

$$Df(a, b) = (D_1 f(a, b), D_2 f(a, b)) = (-2a - 3a^2, 2b),$$

so sieht man, daß die beiden Punkte dadurch charakterisiert sind, daß

$$D_2 f(a, b) = 2b = 0$$

gilt. Das ist kein Zufall!

Nach diesen Betrachtungen wollen wir unser oben angegebenes Ziel, Gleichungen und Gleichungssysteme nach einzelnen Variablen aufzulösen wie folgt konkretisieren:

Ziel der folgenden Überlegungen ist es, ein Kriterium dafür zu finden, daß eine *implizit* gegebene Menge $V(f) = \{x \mid f(x) = 0\}$ lokal in einem Punkt *explizit* als den Graphen einer Abbildung φ beschrieben werden kann. Interessanterweise werden wir dabei die Ableitung von φ angeben können, ohne φ selbst zu kennen.

C) Der Satz über implizite Funktionen

Bemerkung 26.10

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 $f : U \longrightarrow \mathbb{R}^m$ stetig differenzierbar. Für $z \in U$ setze

$$D_x f(z) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(z) & \dots & \frac{\partial f_1}{\partial x_n}(z) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1}(z) & \dots & \frac{\partial f_m}{\partial x_n}(z) \end{pmatrix} \in \mathrm{Mat}(m \times n, \mathbb{R})$$

und

$$D_y f(z) = \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(z) & \dots & \frac{\partial f_1}{\partial y_m}(z) \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(z) & \dots & \frac{\partial f_m}{\partial y_m}(z) \end{pmatrix} \in \mathrm{Mat}_m(\mathbb{R}).$$

Dann gilt

$$Df(z) = \big(D_x f(z), D_y f(z)\big) \in \mathrm{Mat}(m \times n+m, \mathbb{R})$$

und

$$Df(z) \circ \begin{pmatrix} x \\ y \end{pmatrix} = D_x f(z) \circ x + D_y f(z) \circ y. \tag{94}$$

b. Man sieht leicht, daß durch

$$||(x,y)^t||_P := \max\big\{||x||_2, ||y||_2\big\}$$

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 26.11 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 $A = (a_{ij}) \in \mathrm{Mat}(p \times q, \mathbb{R})$ betrachten, so identifizieren wir $\mathrm{Mat}(p \times q, \mathbb{R})$ stets mit $\mathbb{R}^{p \cdot q}$ und verwenden die euklidische Norm (siehe Aufgabe 22.59).

Satz 26.11 (Satz über implizite Funktionen)
Es seien $U \subseteq \mathbb{R}^{n+m}$ offen, $f : U \longrightarrow \mathbb{R}^m$ stetig differenzierbar und $(a,b) = (a_1,\dots,a_n,b_1,\dots,b_m) \in U$ mit $f(a,b) = 0$ und $\det\big(D_y f(a,b)\big) \neq 0$.
Dann gibt es $\varepsilon, r > 0$ mit $U_\varepsilon(a) \times U_r(b) \subseteq U$ so, daß es genau eine Abbildung

$$\varphi : U_\varepsilon(a) \longrightarrow U_r(b)$$

gibt mit

$$f\big(x,\varphi(x)\big) = 0$$

für alle $x \in U_\varepsilon(a)$. Insbesondere gilt $\varphi(a) = b$.
Zudem ist die Abbildung φ stetig differenzierbar auf $U_\varepsilon(a)$ mit

$$D\varphi(x) = -\big(D_y f\big(x,\varphi(x)\big)\big)^{-1} \circ D_x f\big(x,\varphi(x)\big)$$

für $x \in U_\varepsilon(a)$. Insbesondere gilt also

$$D\varphi(a) = -\left(D_y f(a,b)\right)^{-1} \circ D_x f(a,b)$$

Bemerkung 26.12 (Interpretation des Satzes über implizite Funktionen)
Der Satz über implizite Funktionen besagt im wesentlichen, daß die Verschwindungsmenge einer stetig differenzierbaren Abbildgung f, deren Funktionaldeterminante $\det\big(D_y f(a,b)\big)$ nicht null ist, lokal in $(a,b)^t$ als Graph einer eindeutig bestimmten Abbildung φ parametrisierbar ist.

Bevor wir uns dem Beweis des Satzes zuwenden, wollen wir seine Aussage an zwei Beispielen verdeutlichen.

Beispiel 26.13 (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 $b \neq 0$

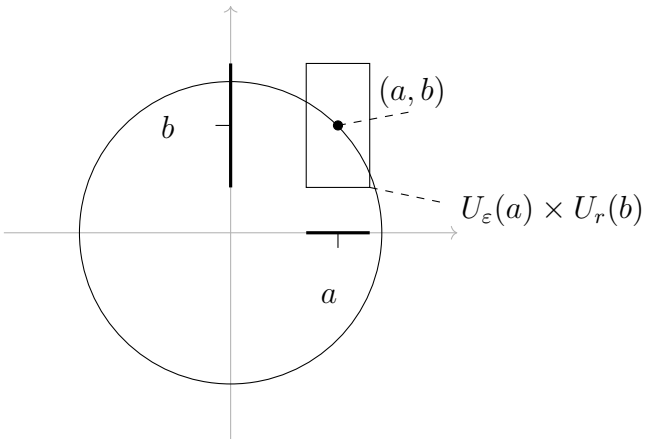


Abbildung 11. Die Kreislinie $\{x^2 + y^2 - 1 = 0\}$

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

$$D\varphi(x) = \varphi'(x) = -\big(D_y f\big(x,\varphi(x)\big)\big)^{-1} \circ D_x f\big(x,\varphi(x)\big) = -\frac{1}{2\varphi(x)} \cdot 2x = -\frac{x}{\varphi(x)}$$

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 26.14

Bei der folgenden Gleichung

$$f(x,y) = x \cdot \exp(y) - \sin(y) = 0$$

können wir nicht mehr so leicht entscheiden, in welchen Punkten die Gleichung nach y auflösbar ist, und vor allem können wir, auch lokal, die Funktion φ nicht mehr bestimmen, die die Gleichung auflöst. Um den Satz über implizite Funktionen anwenden

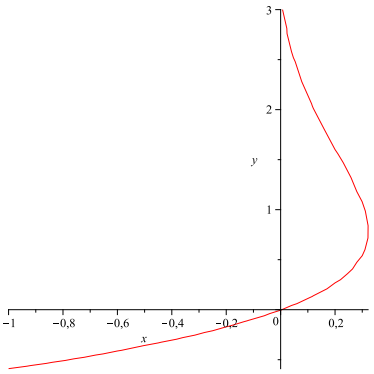


Abbildung 12. Die Lösungsmenge $x \cdot \exp(y) - \sin(y) = 0$

zu können, bestimmen wir zunächst die die Ableitung

$$\frac{\partial f}{\partial y}(x) = x \cdot \exp(y) - \cos(y).$$

Dort, wo diese nicht 0 ist, ist die Lösungsmenge der Gleichung lokal darstellbar als Graph einer Funktion. Die schlechten Punkte sind also die, in denen

$$x \cdot \exp(y) - \sin(y) = 0$$

und

$$x \cdot \exp(y) - \cos(y) = 0$$

gilt, und damit muß

$$\sin(y) = \cos(y)$$

gelten, was genau für

$$y = \frac{\pi}{4} + k \cdot \pi \quad \text{und} \quad x = \frac{\sin(y)}{\exp(y)} = \frac{\pm 1}{\sqrt{2} \cdot \exp(y)}$$

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 26.11 für die Ableitung von φ in a für den Punkt $(a,b)^t \in V(f) = \{(x,y)^t \in \mathbb{R}^2 \mid f(x,y) = 0\}$ dann

$$D\varphi(a) = -\frac{1}{\frac{\partial f}{\partial y}(a,b)} \cdot \frac{\partial f}{\partial x}(a,b) = \frac{\exp(b)}{\cos(b) - a \cdot \exp(b)}.$$

Wir können die Ableitung der unbekannten Funktion φ im Punkt a also angeben.

D) Der Beweis des Satzes über implizite Funktionen

Beweis des Satzes über implizite Funktionen 26.11: Wir teilen den recht langen und komplexen Beweis in mehrere Schritte auf. Im Zentrum steht eine geschickte Anwendung des Banachschen Fixpunktsatzes 26.4, um die Existenz und Eindeutigkeit der Abbildung φ herzuleiten.

1) Ohne Einschränkung gilt $\det\big(D_y f(x,y)\big) \neq 0$ für alle $(x,y)^t \in U$:

Die Abbildung

$$U \longrightarrow \mathbb{R} : (x,y)^t \mapsto \det\big(D_y f(x,y)\big)$$

ist stetig auf U als Verknüpfung der stetigen Polynomfunktion det mit der stetigen Ableitung $D_y f$. Zudem ist $\det\big(D_y f(a,b)\big) \neq 0$ und somit gibt es wegen der Stetigkeit eine ganze Umgebung von $(a,b)^t$ auf der die Funktion nicht null ist (siehe Aufgabe 23.57). Wir können nun U einfach durch diese Umgebung ersetzen.

2) Definition einer Hilfsfunktion g:

Wir setzen

$$A := D_y f(a,b) \in \mathrm{Gl}_m(\mathbb{R})$$

und definieren

$$g : U \longrightarrow \mathbb{R}^m : (x,y)^t \mapsto y - A^{-1} \circ f(x,y).$$

Dann ist g stetig differenzierbar auf 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

$$g(a,b) = b - A^{-1} \circ f(a,b) = b - A^{-1} \circ 0 = b \in \mathbb{R}^m \tag{95}$$

und

$$D_y g(a,b) = \mathbb{1}_m - A^{-1} \circ D_y f(a,b) = 0 \in \mathrm{Mat}_m(\mathbb{R}) \tag{96}$$

ist die Nullmatrix.

3) Konstruktion von φ mit Hilfe des Banachschen Fixpunktsatzes:

Idee: Finde $\varepsilon, r > 0$, so daß für $x \in U_\varepsilon(a)$ die Abbildung

$$g_x : \overline{U_r(b)} \longrightarrow \overline{U_r(b)} : y \mapsto g(x,y)$$

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

$$U \longrightarrow \mathrm{Mat}_m(\mathbb{R}) : (x,y)^t \mapsto D_y g(x,y)$$

stetig auf U, und Stetigkeit ist unabhängig von der gewählten Norm. Wir können auf U also die Produktnorm aus Bemerkung 26.10 und auf $\mathrm{Mat}_m(\mathbb{R})$ die euklidische Norm wählen. Wegen $D_y g(a,b) = 0$ (siehe (96)) gibt es dann ein $r > 0$ mit

$$||D_y g(x,y)||_2 = ||D_y g(x,y) - D_y g(a,b)||_2 \leq \frac{1}{2m^2} \tag{97}$$

für alle $(x,y)^t \in \overline{U_r(a)} \times \overline{U_r(b)}$. Aus Korollar 25.19 erhalten wir damit

$$||g(x,y) - g(x,y')||_2 \overset{25.19}{\leq} ||y - y'||_2 \cdot m^2 \cdot \max_{z \in \overline{U_r(b)}} ||D_y g(x,z)||_2 \overset{(97)}{\leq} \frac{1}{2} \cdot ||y - y'||_2 \tag{98}$$

für alle $x \in \overline{U_r(a)}$ und $y, y' \in \overline{U_r(b)}$. Außerdem ist die Abbildung g stetig in (a,b) mit $g(a,b) = b$ (siehe (95)), so daß es sicher ein $\varepsilon > 0$ gibt mit

$$||g(x,b) - b||_2 = ||g(x,b) - g(a,b)||_2 < \frac{r}{2} \tag{99}$$

für alle $x \in U_\varepsilon(a)$. Wir können dabei ohne Einschränkung $\varepsilon \leq r$ wählen.

Für $x \in U_\varepsilon(a)$ betrachten wir nun die Abbildung

$$g_x : \overline{U_r(b)} \longrightarrow \mathbb{R}^m : y \mapsto g(x,y).$$

Ist $y \in \overline{U_r(b)}$, so gilt

$$||g_x(y) - b||_2 \leq ||g_x(y) - g_x(b)||_2 + ||g_x(b) - b||_2 \overset{(98),(99)}{<} \frac{1}{2} \cdot ||y - b||_2 + \frac{r}{2} \leq r, \tag{100}$$

so daß in der Tat

$$g_x : \overline{U_r(b)} \longrightarrow \overline{U_r(b)}$$

eine Selbstabbildung von $\overline{U_r(b)}$ und wegen (98) sogar eine $\frac{1}{2}$-Kontraktion ist. Als abgeschlossene Teilmenge eines vollständigen Raumes ist $\overline{U_r(b)}$ zudem selbst vollständig, da jede Cauchy-Folge in $\overline{U_r(b)}$ in $\mathbb{R}^m$ konvergiert und wegen der Abgeschlossenheit ihren Grenzwert in $\overline{U_r(b)}$ hat (siehe auch Aufgabe 22.57). Wir können deshalb den Banachschen Fixpunktsatz 26.4 anwenden und finden genau ein $y_x \in \overline{U_r(b)}$ mit

$$y_x = g_x(y_x) = y_x - A^{-1} \circ f(x, y_x), \tag{101}$$

was äquivalent zu

$$0 = A \circ 0 = A \circ \big(y_x - y_x + A^{-1} \circ f(x, y_x)\big) = A \circ A^{-1} \circ f(x, y_x) = f(x, y_x)$$

ist, und wegen (100) gilt sogar $y_x \in U_r(b)$. Wir definieren deshalb

$$\varphi : U_\varepsilon(a) \longrightarrow U_r(b) : x \mapsto y_x.$$

Dann gilt zumindest schon mal

$$f\big(x, \varphi(x)\big) = 0$$

und φ ist auf $U_\varepsilon(a)$ eindeutig mit dieser Eigenschaft.

4) Zeige, daß φ auf $U_\varepsilon(a)$ stetig ist:

Seien $x, x' \in U_\varepsilon(a)$. Dann gilt

$$\varphi(x) - g\big(x, \varphi(x)\big) \overset{(101)}{=} 0 \overset{(101)}{=} \varphi(x') - g\big(x', \varphi(x')\big),$$

und wir erhalten durch Umstellen der Gleichung

$$\begin{aligned} ||\varphi(x) - \varphi(x')||_2 \;&=\; ||g\big(x, \varphi(x)\big) - g\big(x', \varphi(x')\big)||_2 \\ &\leq\; ||g\big(x, \varphi(x)\big) - g\big(x, \varphi(x')\big)||_2 + ||g\big(x, \varphi(x')\big) - g\big(x', \varphi(x')\big)||_2 \end{aligned} \tag{102}$$

Wenden wir Korollar 25.19 auf die Funktion $z \mapsto g\big(z, \varphi(x')\big)$ an, so erhalten wir die Abschätzung

$$\begin{aligned} ||g\big(x, \varphi(x')\big) - g\big(x', \varphi(x')\big)||_2 &\overset{25.19}{\leq} ||x - x'||_2 \cdot n^2 \cdot \max_{z \in \overline{U_\varepsilon(a)}} ||D_x g(z, \varphi(x'))||_2 \\ &\leq ||x - x'||_2 \cdot L, \end{aligned} \tag{103}$$

wobei

$$L = n^2 \cdot \max\big\{||D_x g(z, y)||_2 \;\big|\; z \in \overline{U_\varepsilon(a)}, y \in \overline{U_r(b)}\big\} \in \mathbb{R}$$

eine von x und x' unabhängige Konstante ist. Setzen wir (103) in (102) ein und berücksichtigen zudem noch (98), so erhalten wir

$$||\varphi(x) - \varphi(x')||_2 \leq \frac{1}{2} \cdot ||\varphi(x) - \varphi(x')||_2 + ||x - x'||_2 \cdot L$$

und damit

$$||\varphi(x) - \varphi(x')||_2 \leq 2 \cdot L \cdot ||x - x'||_2. \tag{104}$$

Mithin ist φ Lipschitz stetig und nach Proposition 23.33 somit auch stetig auf $U_\varepsilon(a)$.

5) Zeige, daß φ auf $U_\varepsilon(a)$ total differenzierbar ist mit

$$D\varphi(x) = -\big(D_y f\big(x,\varphi(x)\big)\big)^{-1} \circ D_x f\big(x,\varphi(x)\big) :$$

Sei $(x,y)^t \in U_\varepsilon(a) \times U_r(b)$ gegeben. Wir definieren

$$\rho : U \longrightarrow \mathbb{R}^m$$

durch

$$\begin{aligned} \rho(x',y') &= f(x',y') - f(x,y) - Df(x,y) \circ \big((x',y')^t - (x,y)^t\big) \\ &\overset{(94)}{=} f(x',y') - f(x,y) - D_x f(x,y) \circ (x'-x) - D_y f(x,y) \circ (y'-y). \end{aligned} \tag{105}$$

Da f in $(x,y)^t$ total differenzierbar ist, gilt dann

$$\lim_{(x',y')\to(x,y)} \frac{\rho(x',y')}{||(x',y')^t - (x,y)^t||_2} = 0. \tag{106}$$

Aufgrund der Definition von φ erhalten wir aus (105) mit $y = \varphi(x)$ und $y' = \varphi(x')$

$$\begin{aligned} 0 &= f\big(x',\varphi(x')\big) - f\big(x,\varphi(x)\big) \\ &= D_x f\big(x,\varphi(x)\big) \circ (x'-x) + D_y f\big(x,\varphi(x)\big) \circ \big(\varphi(x') - \varphi(x)\big) + \rho\big(x',\varphi(x')\big). \end{aligned}$$

Lösen wir die Gleichung nach $\varphi(x') - \varphi(x)$ auf, so ergibt sich

$$\begin{aligned} \varphi(x') - \varphi(x) = &-\big(D_y f\big(x,\varphi(x)\big)\big)^{-1} \circ D_x f\big(x,\varphi(x)\big) \circ (x'-x) \\ &-\big(D_y f\big(x,\varphi(x)\big)\big)^{-1} \circ \rho\big(x',\varphi(x')\big), \end{aligned}$$

wobei wir berücksichtigen, daß $D_y f\big(x,\varphi(x)\big)$ wegen 1) invertierbar ist.

Für die totale Differenzierbarkeit von φ in $(x,y)^t$ bleibt zu zeigen, daß

$$\lim_{x'\to x} \frac{-\big(D_y f\big(x,\varphi(x)\big)\big)^{-1} \circ \rho\big(x',\varphi(x')\big)}{||x'-x||_2} = 0,$$

und da die Multiplikation mit der Matrix $-\big(D_y f\big(x,\varphi(x)\big)\big)^{-1}$ eine stetige lineare Abbildung ist, reicht es dazu, zu zeigen, daß

$$\lim_{x'\to x} \frac{\rho\big(x',\varphi(x')\big)}{||x'-x||_2} = 0 \tag{107}$$

gilt. Dazu beachten wir, daß man für $u \in \mathbb{R}^n$ und $v \in \mathbb{R}^m$ stets

$$||(u,v)||_2 = \sqrt{\sum_{i=1}^n u_i^2 + \sum_{j=1}^m v_j^2} \le \sqrt{\sum_{i=1}^n u_i^2} + \sqrt{\sum_{j=1}^m v_j^2} \le ||u||_2 + ||v||_2 \tag{108}$$

erhält und damit auch

$$\frac{||\rho\big(x',\varphi(x')\big)||_2}{||(x',\varphi(x')) - (x,\varphi(x))||_2} \overset{(108)}{\ge} \frac{||\rho\big(x',\varphi(x')\big)||_2}{||x'-x||_2 + ||\varphi(x')) - \varphi(x)||_2} \overset{(104)}{\ge} \frac{||\rho\big(x',\varphi(x')\big)||_2}{(1+2L)\cdot||x'-x||_2}.$$

Wegen (106) und da φ nach 4) stetig ist, konvergiert die linke Seite für $x' \to x$ gegen 0, also muß auch die rechte Seite gegen 0 konvergieren und (107) ist gezeigt.

6) Zeige, daß $D\varphi$ auf $U_\varepsilon(a)$ stetig ist:

Wegen Teil 4) ist φ stetig und nach Voraussetzung ist zudem $D_x f$ stetig, so daß auch die Verkettung

$$U_\varepsilon(a) \longrightarrow \mathrm{Mat}(m \times n, \mathbb{R}) : x \mapsto D_x f\big(x, \varphi(x)\big)$$

stetig ist. Analog ist auch die Abbildung

$$U_\varepsilon(a) \longrightarrow \mathrm{Mat}_m(\mathbb{R}) : x \mapsto D_y f\big(x, \varphi(x)\big)$$

stetig und da auch das Invertieren einer Matrix nach Beispiel 23.21 stetig ist, ist auch die Abbildung

$$U_\varepsilon(a) \longrightarrow \mathrm{Mat}_m(\mathbb{R}) : x \mapsto \big(D_y f\big(x, \varphi(x)\big)\big)^{-1}$$

stetig. Aber damit ist dann auch

$$U_\varepsilon(a) \longrightarrow \mathrm{Mat}(m \times n, \mathbb{R}) : x \mapsto D\varphi(x) = -\big(D_y f\big(x, \varphi(x)\big)\big)^{-1} \circ D_x f\big(x, \varphi(x)\big)$$

stetig, d.h. die Komponentenfunktionen sind stetig. □

Bemerkung 26.15 (Satz über implizite Funktionen)

a. Der Satz über implizite Funktionen 26.11 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 a berechnen!

b. Der Satz über implizite Funktionen spielt eine wichtige Rolle bei der Betrachtung von Extrema mit Nebenbedingungen und im Beweis der Transformationsformel für Integrale.

c. Man kann den Satz über implizite Funktionen ohne wesentliche Änderungen auch allgemeiner für Banachräume zeigen (siehe [Heu04, S. 330-345]).

E) Der Satz über die Umkehrfunktion

Bemerkung 26.16 (Umkehrsatz für stetig differenzierbare Funktionen)

Ist eine Funktion $f : (a, b) \longrightarrow \mathbb{R}$ stetig differenzierbar mit $f'(x) \neq 0$ für alle $x \in (a, b)$, so ist die stetige Funktion $f' : (a, b) \longrightarrow \mathbb{R}$ wegen des Zwischenwertsatzes 14.12 entweder stets positiv oder stets negativ. Wegen Proposition 18.13 ist die Funktion f dann aber streng monoton und wegen des Umkehrsatzes für streng monotone Funktionen 14.21 ist f deshalb bijektiv mit $\mathrm{Im}(f) = (c, d)$ für geeignete $c, d \in \mathbb{R} \cup \{-\infty, \infty\}$, d.h. es existiert die Umkehrfunktion

$$f^{-1} : (c, d) \longrightarrow (a, b).$$

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 $y = f(x)$ gilt nämlich

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)} = \left(f'(x)\right)^{-1}. \tag{109}$$

Diesen Sachverhalt wollen wir nun ins Mehrdimensionale verallgemeinern. Der Satz über implizite Funktionen legt dabei nahe, die Bedingung $f'(x) \neq 0$ durch die Bedingung $\det\big(Df(x)\big) \neq 0$ zu ersetzen. Das folgende Beispiel zeigt jedoch, daß wir dabei vorsichtig sein und gewisse Abstriche machen müssen.

Beispiel 26.17 (Polarkoordinaten)
Wir betrachten die Polarkoordinatenabbildung

$$f : (0,\infty) \times \mathbb{R} \longrightarrow \mathbb{R}^2 : (r,\theta)^t \mapsto \big(r \cdot \cos(\theta), r \cdot \sin(\theta)\big)^t.$$

Diese ist stetig differenzierbar auf ihrem Definitionsbereich und die Ableitungsdeterminante

$$\det\big(Df(r,\theta)\big) = \begin{vmatrix} \cos(\theta) & -r \cdot \sin(\theta) \\ \sin(\theta) & r \cdot \cos(\theta) \end{vmatrix} = r \cdot \big(\sin^2(\theta) + \cos^2(\theta)\big) = r > 0$$

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 $U_{\frac{\pi}{2}}(r,\theta)$ 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 (109) verallgemeinern.

Definition 26.18 (Diffeomorphismen)
Es sei $U \subseteq \mathbb{R}^n$ offen, $f : U \longrightarrow \mathbb{R}^n$ und $a \in U$.

a. f ist ein *Diffeomorphismus*, wenn $f(U)$ offen in $\mathbb{R}^n$ ist und $f : U \longrightarrow f(U)$ bijektiv und stetig differenzierbar mit stetig differenzierbarer Umkehrabbildung f^{-1} ist.

b. f ist ein *lokaler Diffeomorphismus in* a, wenn es eine offene Umgebung V_a von a gibt, so daß $f_{|V_a}$ ein Diffeomorphismus ist.

c. f ist ein *lokaler Diffeomorphismus*, wenn f in jedem Punkt von U ein lokaler Diffeomorphismus ist.

Das Bestimmen einer Umkehrfunktion von f kann man auch als Lösen der impliziten Gleichung $y = f(x)$ nach 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 26.19 (Satz über die Umkehrfunktion)
Sei $c \in U \subseteq \mathbb{R}^n$ offen und $f : U \longrightarrow \mathbb{R}^n$ stetig differenzierbar mit $\det\big(Df(c)\big) \neq 0$. Dann gibt es $c \in V \subseteq U$ offen und $W \subseteq \mathbb{R}^n$ offen, so daß

$$f_| : V \longrightarrow W$$

bijektiv ist. Zudem ist die Umkehrabbildung $f_|^{-1} : W \longrightarrow V$ stetig differenzierbar mit

$$Df_|^{-1}(y) = \Big(Df\big(f_|^{-1}(y)\big)\Big)^{-1}$$

für $y \in W$. Insbesondere ist f ein lokaler Diffeomorphismus in c mit

$$Df_|^{-1}\big(f(c)\big) = \big(Df(c)\big)^{-1}.$$

Beweis: Wir setzen $a = f(c)$ und $b = c$. Die Abbildung

$$h : \mathbb{R}^n \times U \longrightarrow \mathbb{R}^n : (x, y)^t \mapsto f(y) - x$$

ist stetig differenzierbar auf $\mathbb{R}^n \times U$ mit $h(a, b) = f(c) - f(c) = 0$ und

$$\det\big(D_y h(a, b)\big) = \det\big(Df(c)\big) \neq 0.$$

Damit erfüllt h die Voraussetzungen des Satzes über implizite Funktionen und wir finden eine *eindeutig* bestimmte Abbildung

$$\varphi : U_\varepsilon(a) \longrightarrow U_r(b) = U_r(c) \subseteq U$$

mit

$$0 = h\big(x, \varphi(x)\big) = f\big(\varphi(x)\big) - x, \tag{110}$$

für alle $x \in U_\varepsilon(a)$. Zudem ist die Funktion φ stetig differenzierbar auf $U_\varepsilon(a)$ mit

$$\begin{aligned} D\varphi(x) &= -\big(D_y h\big(x, \varphi(x)\big)\big)^{-1} \circ D_x h\big(x, \varphi(x)\big) \\ &= -\big(Df\big(\varphi(x)\big)\big)^{-1} \circ (-\mathbb{1}_n) = \big(Df\big(\varphi(x)\big)\big)^{-1}. \end{aligned} \tag{111}$$

Setzen wir nun $W = U_\varepsilon(a) = U_\varepsilon\big(f(c)\big)$ und $V = \varphi(W)$, so ist die Abbildung

$$\varphi : W \longrightarrow V$$

aufgrund der Definition von V surjektiv und wegen (110) gilt

$$f_{|V} \circ \varphi = \mathrm{id}_W.$$

Damit ist φ aber auch injektiv, da aus $\varphi(x) = \varphi(x')$ automatisch

$$x = f\big(\varphi(x)\big) = f\big(\varphi(x')\big) = x'$$

folgt, und $f_{|V}$ ist die Inverse von φ, d.h. auch

$$f_| : V \longrightarrow W$$

ist bijektiv mit $\varphi = f_|^{-1}$ als Umkehrabbildung. Außerdem gilt auch

$$c = b = \varphi(a) \in V.$$

Unter Berücksichtigung von (111) ist damit alles bewiesen außer, daß V *offen* ist.

Sei dazu $x \in V = \varphi(W) \subseteq U_r(b) \subseteq U$ 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

$$U_\gamma\big(f(x)\big) \subseteq W.$$

Nach Voraussetzung ist die Abbildung f stetig auf $U_r(b)$ und deshalb gibt es ein $\delta > 0$ mit $U_\delta(x) \subseteq U_r(b)$ und

$$f\big(U_\delta(x)\big) \subseteq U_\gamma\big(f(x)\big) \subseteq W.$$

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 $z \in U_\delta(x) \subseteq U_r(b)$ 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 $U_r(b)$ muß dann notwendigerweise

$$z = \varphi\big(f(z)\big) \in V$$

gelten. Damit ist der Satz dann bewiesen. □

Bemerkung 26.20 (Satz über die Umkehrfunktion)

a. Der Satz über die Umkehrfunktion ist die naheliegende lokale Verallgemeinerung des eindimensionalen Umkehrsatzes in Bemerkung 26.16. 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

$$\mathbb{1}_n = D\,\mathrm{id}(c) = D\big(f^{-1} \circ f\big)(c) = Df^{-1}\big(f(c)\big) \circ Df(c)$$

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* a ist, wenn $\det\big(Df(a)\big) \neq 0$ gilt.

e. Der Satz über die Umkehrfunktion 26.19 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 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 26.21

a. Die Polarkoordinatenabbildung aus Beispiel 26.17

$$f : (0,\infty) \times \mathbb{R} \longrightarrow \mathbb{R}^2 : (r,\theta)^t \mapsto \big(r \cdot \cos(\theta), r \cdot \sin(\theta)\big)^t.$$

ist lokal in jedem Punkt (r,θ) ein Diffeomorphismus, da

$$\det\big(Df(r,\theta)\big) = r > 0.$$

Und wir können nun auch die Ableitung der Umkehrfunktion in $f(r,\theta)$ berechnen als

$$Df^{-1}\big(f(r,\theta)\big) = \begin{pmatrix} \cos(\theta) & -r \cdot \sin(\theta) \\ \sin(\theta) & r \cdot \cos(\theta) \end{pmatrix}^{-1} = \frac{1}{r} \cdot \begin{pmatrix} r \cdot \cos(\theta) & r \cdot \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{pmatrix}.$$

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

$$f_| : (0,\infty) \times \big(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\big) \longrightarrow (0,\infty) \times \mathbb{R}$$

ist also bijektiv mit der Umkehrabbildung

$$f_|^{-1} : (0,\infty) \times \mathbb{R} \longrightarrow (0,\infty) \times \big(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\big) : (x,y)^t \mapsto \left(\sqrt{x^2+y^2}, \arctan\left(\tfrac{y}{x}\right)\right).$$

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 $x = (x_1, x_2, x_3)^t$ hat die Funktionaldeterminante

$$\begin{aligned}
\det\big(Df(x)\big) &= \begin{vmatrix} -1 & -1 & -1 \\ x_2 + x_3 & x_1 + x_3 & x_1 + x_2 \\ -x_2x_3 & -x_1x_3 & -x_1x_2 \end{vmatrix} \\
&= \begin{vmatrix} -1 & 0 & 0 \\ x_2 + x_3 & x_1 - x_2 & x_1 - x_3 \\ -x_2x_3 & -(x_1 - x_2)\cdot x_3 & -(x_1 - x_3)\cdot x_2 \end{vmatrix} \\
&= (x_1 - x_2)\cdot(x_1 - x_3)\cdot \begin{vmatrix} -1 & 0 & 0 \\ x_2 + x_3 & 1 & 1 \\ -x_2x_3 & -x_3 & -x_2 \end{vmatrix} \\
&= (x_1 - x_2)\cdot(x_1 - x_3)\cdot(x_2 - x_3),
\end{aligned}$$

wie man leicht nachrechnet. Damit gilt, daß f immer dann lokal in $x = (x_1, x_2, x_3)^t$ invertierbar ist, wenn die x_i paarweise verschieden sind.
Betrachtet man das Polynom

$$p = (t - x_1)\cdot(t - x_2)\cdot(t - x_3) = t^3 + f_1(x)\cdot t^2 + f_2(x)\cdot t + f_3(x),$$

so sind die Komponentenfunktionen von f gerade die Koeffizienten von p. Die Inverse f^{-1} von f lokal in $y = (y_1, y_2, y_3)^t$ zu bestimmen, bedeutet also, die Nullstellen des Polynoms

$$p_y = t^3 + y_1t^2 + y_2t + y_3$$

in Abhängigkeit von y anzugeben. Der Satz über die Umkehrfunktion besagt nun, daß die Nullstellen von p_y stetig differenzierbar von y abhängen, sofern man y in einer kleinen Umgebung eines Punktes a bewegt, für den p_a drei paarweise verschiedene Nullstellen hat.

F) Der Satz von der offenen Abbildung

Korollar 26.22 (Satz von der offenen Abbildung)
Seien $U \subseteq \mathbb{R}^n$ offen und $f \in \mathcal{C}^1(U, \mathbb{R}^n)$ mit $\det\big(Df(x)\big) \neq 0$ für alle $x \in U$.

a. Ist $O \subseteq U$ offen, so ist $f(O)$ offen in $\mathbb{R}^n$.

b. Ist $O \subseteq U$ offen und f injektiv, so ist $f^{-1} : f(O) \longrightarrow O$ stetig differenzierbar.

Beweis: Sei $O \subseteq U$ offen und $c \in O$ beliebig. Nach dem Satz über die Umkehrfunktion 26.19 angewendet auf

$$f_{|O} : O \longrightarrow \mathbb{R}^n$$

gibt es dann offene Mengen $V_c \subseteq O$ mit $c \in V_c$ und $W_c \subseteq \mathbb{R}^n$, so daß

$$f_{|V_c} : V_c \longrightarrow W_c$$

bijektiv ist. Damit gilt dann aber insbesondere

$$W_c = f(V_c) \subseteq f(O),$$

und somit ist

$$f(O) = \bigcup_{c \in O} W_c \tag{112}$$

als Vereinigung offener Mengen offen.

Ist nun zudem f injektiv, so ist die Abbildung $f : O \longrightarrow f(O)$ bijektiv, und wegen des Satzes über die Umkehrfunktion ist die Umkehrabbildung f^{-1} für jedes $c \in O$ auf W_c stetig differenzierbar. Damit ist sie wegen (112) aber in jedem Punkt von $f(O)$ differenzierbar und ihre Ableitung ist in jedem Punkt von $f(O)$ stetig, d.h. f^{-1} ist stetig differenzierbar auf ganz $f(O)$. □

Beispiel 26.23

Die Funktion $f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$ ist stetig differenzierbar, aber es gibt einen Punkt $x = 0$ mit

$$\det\big(Df(x)\big) = f'(x) = 2x = 0.$$

Die Voraussetzungen des Satzes von der offenen Abbildung 26.22 sind also nicht erfüllt, und in der Tat ist das Bild auch nicht offen in $\mathbb{R}$,

$$f(\mathbb{R}) = [0, \infty)$$

Bemerkung 26.24 (Diffeomorphismen)

Kombinieren wir die Aussagen im Satz über die Umkehrfunktion 26.19 mit dem Satz der offenen Abbildung 26.22 und Bemerkung 26.20, so erhalten wir die folgenden Charakterisierungen von (lokalen) Diffeomorphismen $f : U \longrightarrow \mathbb{R}^n$:

a. f ist ein *lokaler Diffeomorphismus in a* $\iff$ $\det\big(Df(a)\big) \neq 0$.

b. f ist ein *lokaler Diffeomorphismus* $\iff$ $\det\big(Df(x)\big) \neq 0$ für alle $x \in U$.

c. f ist ein *Diffeomorphismus* $\iff$ f ist ein injektiver lokaler Diffeomorphismus.

G) Extrema unter Nebenbedingungen

In Abschnitt 25 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 26.25

Sei $D \subseteq \mathbb{R}^n$, $f : D \longrightarrow \mathbb{R}$, $g : D \longrightarrow \mathbb{R}^m$ und $a \in D$.

a. f hat in a ein *globales Maximum unter der Nebenbedingung* $g = 0$, wenn $a \in V(g)$ und $f(a) \geq f(x)$ für alle $x \in V(g)$.

b. f hat in a ein *lokales Maximum unter der Nebenbedingung* $g = 0$, wenn $a \in V(g)$ und wenn es ein $\delta > 0$ gibt, so daß $f(a) \geq f(x)$ für alle $x \in V(g) \cap U_\delta(a)$.

c. f hat in a ein *isoliertes lokales Maximum unter der Nebenbedingung* $g = 0$, wenn $a \in V(g)$ und wenn es ein $\delta > 0$ gibt mit $f(a) > f(x)$ für alle $a \neq x \in V(g) \cap U_\delta(a)$.

d. f hat in a ein *globales Minimum unter der Nebenbedingung* $g = 0$, wenn $a \in V(g)$ und wenn $f(a) \leq f(x)$ für alle $x \in V(g)$.

e. f hat in a ein *lokales Minimum unter der Nebenbedingung* $g = 0$, wenn $a \in V(g)$ und wenn es ein $\delta > 0$ gibt, so daß $f(a) \leq f(x)$ für alle $x \in V(g) \cap U_\delta(a)$.

f. f hat in a ein *isoliertes lokales Minimum unter der Nebenbedingung* $g = 0$, wenn $a \in V(g)$ und wenn es ein $\delta > 0$ gibt mit $f(a) < f(x)$ für alle $a \neq x \in V(g) \cap U_\delta(a)$.

g. a heißt *Extremstelle von f unter der Nebenbedingung* $g = 0$, wenn $a \in V(g)$ und wenn f in a ein lokales Minimum oder ein lokales Maximum unter der Nebenbedingung $g = 0$ hat.

Beispiel 26.26 (Extremwerte unter Nebenbedingungen mittels Parametrisierung)

Manchmal ist es möglich, eine Verschwindungsmenge $V(g)$ 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

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto ||x||_2^2 = x_1^2 + x_2^2$$

und die Nebenbedingung $g = 0$ mit

$$g : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto x_1 + x_2 - 1.$$

Dann ist $g(x) = x_1 + x_2 - 1 = 0$ 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 $V(g)$, d.h.

$$V(g) = \left\{ (t, 1-t)^t \;\middle|\; t \in \mathbb{R} \right\} = \mathrm{Im}(\psi).$$

Um die Extremstellen von f mit Nebenbedingung $g = 0$ zu finden, untersuchen wir nun die Extremstellen von

$$h = f \circ \psi : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto f(t, 1-t) = 2t^2 - 2t + 1.$$

Die Funktion ist unendlich oft differenzierbar mit

$$h'(t) = 4t - 2 = 0$$

genau dann, wenn $t = \frac{1}{2}$. Sie hat also nur einen kritischen Punkt, und in diesem gilt

$$h''\left(\tfrac{1}{2}\right) = 4 > 0.$$

Mithin hat die Funktion h in $t = \frac{1}{2}$ ein lokales Minimum, und deshalb hat f in $\psi\left(\frac{1}{2}\right) = \left(\frac{1}{2}, \frac{1}{2}\right)^t$ 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

$$h'(t) = \left(3t^2 + 2t\right) \cdot \exp\left(t^3 + t^2\right) = 0$$

genau dann, wenn $t = 0$ oder $t = -\frac{2}{3}$ gilt. Für die zweite Ableitung

$$h''(t) = (6t+2) \cdot \exp\left(t^3+t^2\right) + \left(3t^2+2t\right)^2 \cdot \exp\left(t^3+t^2\right)$$

von h gilt in den kritischen Punkten

$$h''(0) = 2 > 0$$

und

$$h''\left(-\tfrac{2}{3}\right) = -2 \cdot \exp\left(\tfrac{4}{27}\right) < 0,$$

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 $\psi\left(-\frac{2}{3}\right) = \left(-\frac{8}{27}, \frac{4}{9}\right)^t$ ein lokales Maximum unter der Nebenbedingung $g = 0$.

H) Lagrange-Multiplikatoren

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 25.25 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 26.27 (Lagrange-Multiplikatoren)
Es seien $U \subseteq \mathbb{R}^n$ offen, $f : U \longrightarrow \mathbb{R}$ und $g : U \longrightarrow \mathbb{R}^m$ stetig differenzierbar auf U und $a \in V(g)$ mit $\operatorname{rang}\left(Dg(a)\right) = m < n$.
Ist a Extremstelle von f unter der Nebenbedingung $g = 0$, so gibt es ein $b \in \mathbb{R}^m$ mit

$$Df(a) = \sum_{j=1}^{m} b_j \cdot Dg_j(a).$$

Man bezeichnet $b_1, \ldots, b_m$ auch als *Lagrange-Multiplikatoren*.

Beweis: Die Grundidee des Beweises besteht darin, eine Hilfsfunktion

$$q : U' \longrightarrow \mathbb{R}$$

auf einer offenen Menge $U' \subset \mathbb{R}^{n-m}$ zu konstruieren, die eine Extremstelle hat, und dann Proposition 25.25 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 26.11 die Verschwindungsmenge $V(g)$ lokal in a.

Nach Voraussetzung besitzt $Dg(a)$ 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 $x = (u, v) = (u_1, \dots, u_{n-m}, v_1, \dots, v_m)$ und nehmen ohne Einschränkung an, daß

$$\operatorname{rang}\big(D_v g(a)\big) = \operatorname{rang}\big(Dg(a)\big) = m$$

gilt, d.h. die letzten m Spalten von $Dg(a)$ sind linear unabhängig. Dann gilt aber

$$\det\big(D_v g(a)\big) \neq 0.$$

Wir können den Satz über implizite Funktionen 26.11 also auf g anwenden und damit $V(g)$ lokal in a parametrisieren. D.h. Satz 26.11 liefert uns eine stetig differenzierbare Funktion

$$\varphi : U_\varepsilon(u_a) \longrightarrow U_r(v_a)$$

mit $\varphi(u_a) = v_a$,

$$D\varphi(u_a) = -\left(D_v g(a)\right)^{-1} \circ D_u g(a) \tag{113}$$

und

$$g\big(u, \varphi(u)\big) = 0 \tag{114}$$

für alle $u \in U_\varepsilon(u_a)$, wobei $u_a = (a_1, \dots, a_{n-m})$ und $v_a = (a_{n-m+1}, \dots, a_n)$.

Nach Voraussetzung hat f in 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

$$f(x) \leq f(a) \tag{115}$$

für alle $x \in V(g) \cap U_\delta(a)$. Da die Abbildung

$$p : U_\varepsilon(u_a) \longrightarrow \mathbb{R}^n : u \mapsto \big(u, \varphi(u)\big)$$

stetig ist, gibt es eine kleine Umgebung von u_a, die ganz in die δ-Umgebung $U_\delta(a)$ von $p(u_a) = a$ abgebildet wird. Indem wir ε verkleinern können wir ohne Einschränkung annehmen, daß dies schon für ganz $U_\varepsilon(u_a)$ gilt. Unter Berücksichtigung von (114) heißt das

$$\big(u, \varphi(u)\big) = p(u) \in U_\delta(a) \cap V(g) \tag{116}$$

für alle $u \in U_\varepsilon(a)$. Wir betrachten nun die stetig differenzierbare Funktion

$$q : U_\varepsilon(u_a) \longrightarrow \mathbb{R} : u \mapsto f\big(u, \varphi(u)\big).$$

Wegen (115) und (116) gilt

$$q(u) \leq q(u_a)$$

für alle $u \in U_\varepsilon(u_a)$, so daß q eine Extremstelle in u_a hat. Nach Proposition 25.25 ist u_a somit ein kritischer Punkt von q, und wir erhalten mit der Kettenregel

$$\begin{aligned} 0 &= Dq(u_a) = Df\big(p(u_a)\big) \circ Dp(u_a) \\ &= D_u f(a) \circ \mathbb{1}_{n-m} + D_v f(a) \circ D\varphi(u_a) \\ &\overset{(113)}{=} D_u f(a) - D_v f(a) \circ \big(D_v g(a)\big)^{-1} \circ D_u g(a). \end{aligned} \tag{117}$$

Wir setzen nun

$$b^t := D_v f(a) \circ \big(D_v g(a)\big)^{-1} \tag{118}$$

und erhalten damit einerseits aus (118)

$$D_v f(a) = b^t \circ D_v g(a) = \sum_{j=1}^{m} b_j \cdot D_v g_j(a)$$

und andererseits aus (117) und (118)

$$D_u f(a) = b^t \circ D_u g(a) = \sum_{j=1}^{m} b_j \cdot D_u g_j(a).$$

Insgesamt ist damit

$$Df(a) = \sum_{j=1}^{m} b_j \cdot Dg_j(a)$$

gezeigt. □

Bemerkung 26.28 (Lagrange-Multiplikatoren)
Wir setzen die Bezeichnungen und Bedingungen aus Satz 26.27 voraus.

a. Damit f in a eine Extremstelle unter der Nebenbedingung $g = 0$ hat, muß es wegen Satz 26.27 ein $b \in \mathbb{R}^m$ geben, so daß $(a, b)^t \in \mathbb{R}^{n+m}$ eine Lösung der $n + m$ Gleichungen

$$\begin{aligned} D_i f(a) - \sum_{j=1}^{m} b_j \cdot D_i g_j(a) = 0, \quad i = 1, \dots, n \\ g_j(a) = 0, \quad j = 1, \dots, m \end{aligned} \tag{119}$$

ist. Sucht man also Kandidaten für den Punkt a, so sollte man versuchen, dieses Gleichungssystem zu lösen. Man beachte, daß es in aller Regel ein nicht-lineares Gleichungssystem ist.

b. Daß $(a, b)^t$ das Gleichungssystem (119) löst, ist eine *notwendige* Bedingung dafür, daß f in a eine Extremstelle unter der Nebenbedingung $g = 0$ hat. Es ist aber weder garantiert, daß f in 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 26.27 gibt zu gegebenem $a \in U$ konkret den Vektor b als

$$b = \Big(D_v f(a) \circ \big(D_v g(a)\big)^{-1}\Big)^t$$

an, wenn $v = (x_{i_1}, \ldots, x_{i_m})$ die Koordinaten bezeichnet, die zu m linear unabhängigen Spalten von $Dg(a)$ gehören. Für die Anwendungen des Satzes ist dies aber irrelevant, da wir in aller Regel a nicht kennen, sondern suchen, indem wir die Gleichungen (119) für a und b lösen.

d. Man kann die Aussage von Satz 26.27 auch etwas anders formulieren, wenn man die Hilfsfunktion

$$h : U \times \mathbb{R}^m \longrightarrow \mathbb{R} : (x, y)^t \mapsto f(x) - \sum_{j=1}^{m} y_j \cdot g_j(x)$$

betrachtet. Bis auf das Vorzeichen sind die linken Seiten von (119) genau die partiellen Ableitungen von h.
Hat also f in a eine Extremstelle unter der Nebenbedingung $g = 0$, so gibt es ein $b \in \mathbb{R}^m$, so daß $(a, b)^t$ ein kritischer Punkt von h ist. Dies ist eine in der Literatur gebräuchliche Formulierung des Satzes.

e. Das Verfahren der Lagrange-Multiplikatoren 26.27 ist in Beispiel 26.26 b. im Fall $t = 0$, d.h. im Punkt $(0, 0)^t$, nicht anwendbar, da dort die Funktionaldeterminante den Rang 0 hat. Die globale Parametrisierung, die wir in dem Beispiel verwendet haben, kennt der Satz über implizite Funktionen nicht!

Beispiel 26.29 (Lagrange-Multiplikatoren)
Wir betrachten die lineare Funktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto x_1 - 2x_2$$

und die Nebenbedingung

$$g : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto ||x||_2^2 - 1 = x_1^2 + x_2^2 - 1,$$

so daß

$$V(g) = \{x \in \mathbb{R}^2 \mid g(x) = 0\} = \partial\overline{U_1(0)}$$

der Einheitskreis um den Ursprung ist. Unser Ziel ist es, die Extremstellen von f unter der Nebenbedingung $g = 0$ zu finden, d.h. die Extremstellen von f eingeschränkt auf den Einheitskreis.

Wegen (119) müssen wir dazu das Gleichungssystem

$$\begin{array}{rcccl} 0 & = & D_1 f(x) - b \cdot D_1 g(x) & = & 1 - b \cdot 2x_1 \\ 0 & = & D_2 f(x) - b \cdot D_2 g(x) & = & -2 - b \cdot 2x_2 \\ 0 & = & g(x) & = & x_1^2 + x_2^2 - 1 \end{array}$$

lösen. Dabei müssen wir die Nebenbedingung

$$\operatorname{rang}\big(Dg(x)\big) = \operatorname{rang}\begin{pmatrix} 2x_1 & 2x_2 \end{pmatrix} = 1$$

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,

$$x_1 = \tfrac{1}{2b}$$

und

$$x_2 = -\tfrac{1}{b}.$$

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

$$a^{(1)} = \left(-\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$$

und für $b = \frac{\sqrt{5}}{2}$ den Punkt

$$a^{(2)} = \left(\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}}\right).$$

Um nun festzustellen, ob die beiden Kandidaten für Extremstellen auch wirklich welche sind, benötigen wir eine Zusatzüberlegung. Die Funktion

$$f_| : V(g) \longrightarrow \mathbb{R}$$

ist stetig und die Menge

$$V(g) = g^{-1}(0)$$

ist kompakt, mithin nimmt die Funktion $f_|$ auf $V(g)$ ihr (globales) Maximum und ihr (globales) Minimum an, siehe Korollar 23.24. Als Kandidaten dafür kommen nur die beiden Punkte $a^{(1)}$ und $a^{(2)}$ in Frage, und wegen

$$f\big(a^{(1)}\big) = -\sqrt{5} < \sqrt{5} = f\big(a^{(2)}\big)$$

folgt dann, daß $a^{(1)}$ ein globales Minimum von f unter der Nebenbedingung $g = 0$ ist, während $a^{(2)}$ ein globales Maximum von f unter der Nebenbedingung $g = 0$ ist.

Bemerkung 26.30

In Beispiel 26.29 hätte man in der Tat auch auf die Lagrange-Multiplikatoren verzichten können, da man die Kreislinie durch

$$\psi : \mathbb{R} \longrightarrow \mathbb{R}^2 : t \mapsto \big(\cos(t), \sin(t)\big)^t$$

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

$$t = \operatorname{arccot}\left(-\tfrac{1}{2}\right)$$

und

$$t = \operatorname{arccot}\left(-\tfrac{1}{2}\right) + \pi.$$

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

$$\psi : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto \left(\tfrac{t^2-1}{t^2+1}, \tfrac{2t}{t^2+1}\right)^t.$$

Wir überlassen es dem Leser, mit dieser Parametrisierung die lokalen Extremstellen von f unter der Nebenbedingung $g = 0$ zu finden.

Aufgaben

Aufgabe 26.31

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 26.32 (Kugelkoordinaten)
Zeige, daß die Abbildung (siehe auch Beispiel 30.14)

$$\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (r, \theta, \vartheta) \mapsto \big(r \cdot \cos(\theta) \cdot \cos(\vartheta), r \cdot \sin(\theta) \cdot \cos(\vartheta), r \cdot \sin(\vartheta)\big)$$

auf der offenen Menge

$$U = (0, \infty) \times (0, 2\pi) \times \big(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\big)$$

ein Diffeomorphismus ist mit Funktionaldeterminante

$$\det(D\varphi(r, \theta, \vartheta)) = r^2 \cdot \cos(\vartheta) > 0.$$

Aufgabe 26.33 (Zylinderkoordinaten)
Zeige, daß die Abbildung (siehe auch Beispiel 30.13)

$$\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (r, \theta, z) \mapsto \big(r \cdot \cos(\theta), r \cdot \sin(\theta), z\big)$$

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 26.34
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 26.35
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 26.36
Zeige, daß die Verschwindungsmenge $V(f)$ für

$$f : \mathbb{R}^4 \longrightarrow \mathbb{R}^2 : (x_1, x_2, y_1, y_2)^t \mapsto \big(x_1^2 + x_2^2 - 2y_1y_2, x_1^3 + x_2^3 + y_1^3 - y_2^3\big)^t$$

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 26.37
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 26.38
Es seien $p, q \in \mathbb{R}_{>1}$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Bestimme die Extremstellen der Funktion

$$f : \mathbb{R}^2_{>0} \longrightarrow \mathbb{R} : x \mapsto \frac{x_1^p}{p} + \frac{x_2^q}{q}$$

unter der Nebenbedingung $x_1 x_2 = 1$. Folgere daraus für $u, v > 0$ die Höldersche Ungleichung $\frac{u^p}{p} + \frac{v^q}{q} \geq uv$.

Aufgabe 26.39
Es sei $U \subseteq \mathbb{R}^n$ offen und beschränkt und $f : \overline{U} \longrightarrow \mathbb{R}^n$ sei stetig auf $\overline{U}$ und stetig differenzierbar auf U. Ferner sei $y \in \mathbb{R}^n$ so, daß $f^{-1}(y) \subseteq U$ und $\det\big(Df(x)\big) \neq 0$ für alle $x \in f^{-1}(y)$. Zeige, daß $f^{-1}(y)$ nur endlich viele Punkte enthält.

Aufgabe 26.40 (Spektralsatz für symmetrische Matrizen)
Sei $A \in \mathrm{Mat}_n(\mathbb{R})$ eine symmetrische Matrix und $f : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto x^t \circ A \circ x = \langle Ax, x\rangle$. Für $Y \subseteq \mathbb{R}^n$ setzen wir

$$M(Y) = \big\{x \in \mathbb{R}^n \;\big|\; ||x||_2 = 1, \langle x, y\rangle = 0 \;\forall\, y \in Y\big\}.$$

a. Zeige, für $i = 1, \ldots, n$ gibt es $y_i \in M(y_1, \ldots, y_{i-1})$ mit

$$\lambda_i := f(y_i) \stackrel{!}{=} \max_{x \in M(y_1, \ldots, y_{i-1})} f(x).$$

b. Zeige, es gilt $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$.

c. Zeige $Ay_i = \lambda_i y_i$ für $i = 1, \ldots, n$ mit Hilfe der Lagrange Multiplikatoren.

d. Bestimme alle lokalen Extrema von f unter der Nebenbedingung $||x||_2 = 1$.

Aufgabe 26.41
Bestimme für die Abbildung

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : x \mapsto (x_1, x_2^2 - 2x_1x_2)^t$$

alle Punkte $y \in \mathbb{R}^2$, so daß $\det(Df(x)) \neq 0$ für alle $x \in f^{-1}(y)$.

Aufgabe 26.42
Welcher Punkt in $V(x_1^2 + x_2^2 - x_3)$ hat den kleinsten Abstand vom Punkt $\big(1, 1, \frac{1}{2}\big)^t$?

§ 27 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 27.1

Es seien $a, b \in \mathbb{R}^n$ mit $a < b$, d.h. $a_i < b_i$ für alle $i = 1, \ldots, n$.
Wir nennen $[a, b] := [a_1, b_1] \times \ldots \times [a_n, b_n]$ einen n-dimensionalen *Quader* oder ein n-dimensionales *Intervall* (siehe auch Beispiel 22.45), und wir nennen

$$V\big([a,b]\big) = (b_1 - a_1) \cdot \ldots \cdot (b_n - a_n)$$

das *Volumen* des Quaders $[a, b]$.

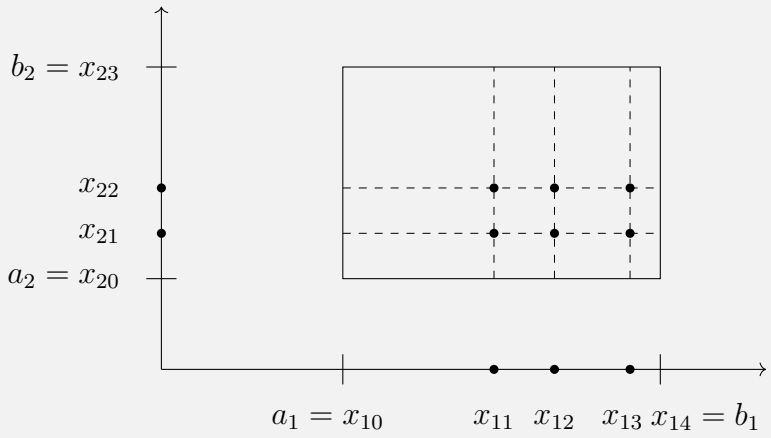


Abbildung 13: Zerlegung eines 2-dimensionalen Quaders

Ist $Z_i = (x_{i0}, \ldots, x_{im_i})$ eine Zerlegung des Intervalls $[a_i, b_i]$ für $i = 1, \ldots, n$, so nennen wir

$$Z = \big((x_{1j_1}, \ldots, x_{nj_n}) \mid j_i = 0, \ldots, m_i, i = 1, \ldots, n\big)$$

eine *Zerlegung* von $[a, b]$, und der Einfachheit halber schreiben wir auch

$$Z = Z_1 \times \ldots \times Z_n.$$

Die Punkte in Z heißen die *Stützpunkte* der Zerlegung, und

$$\operatorname{supp}(Z) = \operatorname{supp}(Z_1) \times \ldots \times \operatorname{supp}(Z_n)$$

heißt ihr *Träger*. Wir nennen

$$l(Z) := \max_{i=1,\dots,n} l(Z_i)$$

die *Länge* oder *Feinheit* von Z und

$$|Z| := |Z_1| \cdot \ldots \cdot |Z_n| = m_1 \cdot \ldots \cdot m_n$$

ihre *Mächtigkeit*. Die $|Z|$ Quader

$$[x_{1j_1-1}, x_{1j_1}] \times \ldots \times [x_{nj_n-1}, x_{nj_n}],$$

$j_i = 1, \dots, m_i$, $i = 1, \dots, n$, heißen die *Teilquader* oder *Teilintervalle* der Zerlegung Z, und mit $\mathrm{TQ}(Z)$ bezeichnen wir die *Menge aller Teilquader* von Z.
Eine zweite Zerlegung Z' von $[a, b]$ heißt *Verfeinerung* von Z, falls $\mathrm{supp}(Z) \subseteq \mathrm{supp}(Z')$ gilt.
Zu zwei Zerlegungen $Z = Z_1 \times \ldots \times Z_n$ und $Z' = Z'_1 \times \ldots \times Z'_n$ definieren wir

$$Z * Z' := (Z_1 * Z'_1) \times \ldots \times (Z_n * Z'_n).$$

Sind Z und Z' Zerlegungen des gleichen Quaders, so nennen wir $Z * Z'$ ihre *gemeinsame Verfeinerung*.

Beispiel 27.2
Für $a = (0, 0)$ und $b = (4, 3)$ ist $[a, b] = \{(x, y) \mid 0 \le x \le 4, 0 \le y \le 3\}$. Betrachten wir nun die Zerlegung $Z_1 = (0, 2, 3, 4)$ und $Z_2 = (0, 1, 3)$, so hat die Zerlegung

$$Z = Z_1 \times Z_2 = \big((0,0), (0,1), (0,3), (2,0), (2,1), (2,3), (3,0), (3,1), (3,3), (4,0), (4,1), (4,3)\big)$$

den Träger

$$\mathrm{supp}(Z) = \big\{(0,0), (0,1), (0,3), (2,0), (2,1), (2,3), (3,0), (3,1), (3,3), (4,0), (4,1), (4,3)\big\},$$

die Länge

$$l(Z) = \max\{l(Z_1), l(Z_2)\} = \max\{2, 2\} = 2$$

und die Teilquader

$$\mathrm{TQ}(Z) = \big\{[0,2]\times[0,1],\ [0,2]\times[1,3],\ [2,3]\times[0,1], [2,3]\times[1,3],\ [3,4]\times[0,1],\ [3,4]\times[1,3]\big\}.$$

Beispiel 27.3 (Äquidistante Zerlegung)
Seien $a, b \in \mathbb{R}^n$ mit $a < b$ und sei zudem $m \ge 1$. 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

$$Z_i^m := (x_{i0}, \dots, x_{im})$$

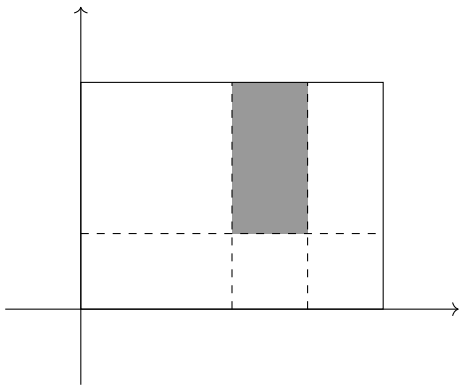


Abbildung 14. Zerlegung von $[0,4] \times [0,3]$ mit markiertem Teilquader

eine Zerlegung von $[a_i, b_i]$ und

$$Z^m := Z_1^m \times \ldots \times Z_n^m$$

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 $Q \in \mathrm{TQ}(Z^m)$ hat das gleiche Volumen

$$V(Q) = \frac{b_1 - a_1}{m} \cdot \ldots \cdot \frac{b_n - a_n}{m} = \frac{V\big([a,b]\big)}{m^n}.$$

Die Länge der Zerlegung Z^m ist

$$l(Z^m) = \max_{i=1,\ldots,n} \frac{b_i - a_i}{m}.$$

Ist m ein Teiler von m', so ist $Z^{m'}$ eine Verfeinerung von Z^m.

Bemerkung 27.4 (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 $Z = Z_1 \times \ldots \times Z_n$ eines Quaders $[a,b] \subset \mathbb{R}^n$ wie in der Definition 27.1 aus, bei der die Zerlegung Z_i von $[a_i, b_i]$ aus m_i Teilintervallen besteht. Fügt man nun in einer der Zerlegungen Z_i einen Stützpunkt y ein, etwa zwischen x_{ij-1} und x_{ij}, so zerteilt man damit automatisch die

$$\frac{m_1 \cdot \ldots \cdot m_n}{m_i} = \frac{|Z|}{|Z_i|}$$

Teilquader, an denen das Intervall $[x_{ij-1}, x_{ij}]$ beteiligt war, in zwei Teile (siehe Abbildung 15) und die neue Zerlegung

$$Z' = Z_1 \times \ldots \times Z_{i-1} \times (x_{i0}, \ldots, x_{ij-1}, y, x_{ij}, \ldots, x_{im_i}) \times Z_{i+1} \times \ldots \times Z_n$$

besitzt deshalb genau diese Anzahl Teilquader mehr als Z, d.h.

$$|Z'| - |Z| = \frac{|Z|}{|Z_i|}.$$

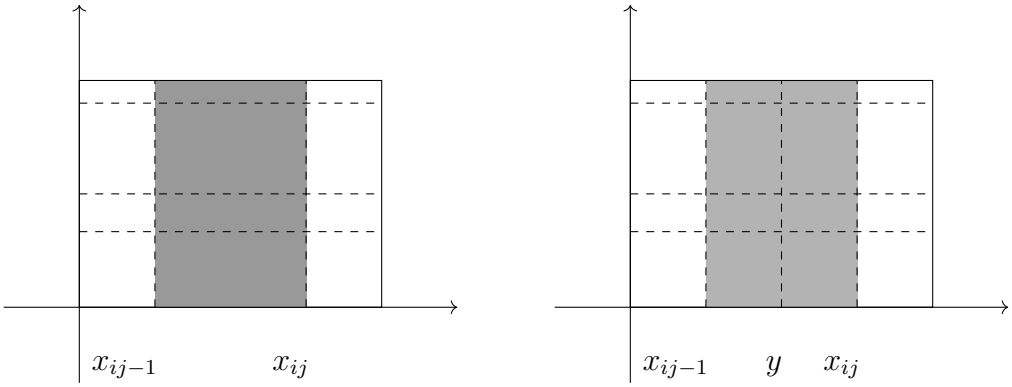


Abbildung 15. Verfeinerung von Z um einen Stützpunkt in Z_i

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 $Q \in \mathrm{TQ}(Z)$ gilt

$$V(Q) \leq l(Z)^n.$$

c. Die Zerlegung Z zerlegt den Quader $[a, b]$ in Teilquader, und das Volumen genügt offenbar der einfachen, aber sehr wichtigen Formel

$$V\big([a, b]\big) = \sum_{Q \in \mathrm{TQ}(Z)} V(Q). \tag{120}$$

Will man diese Formel formal beweisen, sollte man dies am Besten durch Induktion nach

$$m_Z := |Z_1| + \ldots + |Z_n|$$

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

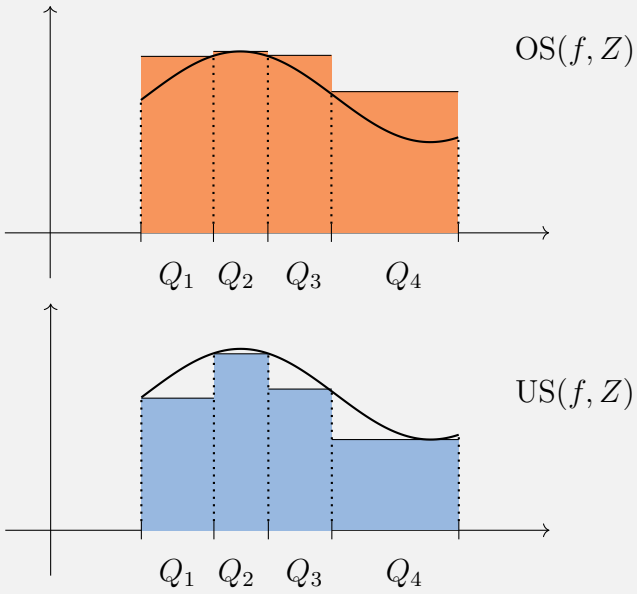
Definition 27.5

Seien $a, b \in \mathbb{R}^n$, $a < b$, $f : [a, b] \to \mathbb{R}$ beschränkt und Z eine Zerlegung von $[a, b]$. Wir definieren die *Obersumme* von f bezüglich Z als

$$\mathrm{OS}(f, Z) := \sum_{Q \in \mathrm{TQ}(Z)} V(Q) \cdot \sup\{f(x) \mid x \in Q\},$$

und die *Untersumme* von f bezüglich Z als

$$\mathrm{US}(f, Z) := \sum_{Q \in \mathrm{TQ}(Z)} V(Q) \cdot \inf\{f(x) \mid x \in Q\}.$$



Beispiel 27.6

Wir betrachten die Funktion $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : x \mapsto x_1 + x_2$ auf dem Quader $[0, 2] \times [0, 2]$ mit der Zerlegung

$$Z = (0, 1, 2) \times (0, 1, 2).$$

Dann gilt

$$\mathrm{TQ}(Z) = \{Q_1 = [0, 1] \times [0, 1],\ Q_2 = [0, 1] \times [1, 2],\ Q_3 = [1, 2] \times [0, 1],\ Q_4 = [1, 2] \times [1, 2]\}$$

und

$$\begin{aligned}\mathrm{OS}(f, Z) =& V(Q_1) \cdot \sup_{x \in Q_1} f(x) + V(Q_2) \cdot \sup_{x \in Q_2} f(x) + V(Q_3) \cdot \sup_{x \in Q_3} f(x) + V(Q_4) \cdot \sup_{x \in Q_4} f(x) \\ =& 1 \cdot 2 + 1 \cdot 3 + 1 \cdot 3 + 1 \cdot 4 = 12\end{aligned}$$

sowie

$$\begin{aligned}\mathrm{US}(f, Z) =& V(Q_1) \cdot \inf_{x \in Q_1} f(x) + V(Q_2) \cdot \inf_{x \in Q_2} f(x) + V(Q_3) \cdot \inf_{x \in Q_3} f(x) + V(Q_4) \cdot \inf_{x \in Q_4} f(x) \\ =& 1 \cdot 0 + 1 \cdot 1 + 1 \cdot 1 + 1 \cdot 2 = 4.\end{aligned}$$

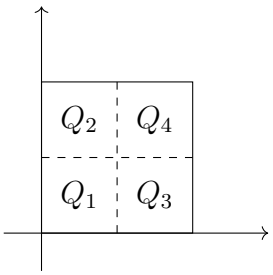


Abbildung 16. Zerlegung von $[(0,0)^t,(2,2)^t] = [0,2] \times [0,2]$

Bemerkung 27.7

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

$$\sum_{i=1}^{n}(x_i - x_{i-1}) \cdot \ldots$$

durch

$$\sum_{Q \in \mathrm{TQ}(Z)} V(Q) \cdot \ldots$$

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 27.8

Seien $a, b \in \mathbb{R}^n$ mit $a < b$ und $f : [a,b] \longrightarrow \mathbb{R}$ mit $|f(x)| \leq M$ für alle $x \in [a,b]$.

a. Ist $Z' = Z'_1 \times \ldots \times Z'_n$ eine Verfeinerung der Zerlegung $Z = Z_1 \times \ldots \times Z_n$ von $[a,b]$ und ist $V_i = \frac{V([a,b])}{b_i - a_i} = \prod_{j \neq i}(b_j - a_j)$, so gelten

$$0 \leq \mathrm{US}(f,Z') - \mathrm{US}(f,Z) \leq 2 \cdot M \cdot l(Z) \cdot \sum_{i=1}^{n} \left(|Z'_i| - |Z_i|\right) \cdot V_i$$

und

$$0 \leq \mathrm{OS}(f,Z) - \mathrm{OS}(f,Z') \leq 2 \cdot M \cdot l(Z) \cdot \sum_{i=1}^{n} \left(|Z'_i| - |Z_i|\right) \cdot V_i.$$

Insbesondere gilt also

$$\mathrm{US}(f,Z) \leq \mathrm{US}(f,Z') \leq \mathrm{OS}(f,Z') \leq \mathrm{OS}(f,Z).$$

b. Für je zwei Zerlegungen Z und Z' von $[a,b]$ gilt

$$\mathrm{US}(f,Z) \le \mathrm{OS}(f,Z').$$

c. Es gelten

$$-M \cdot V\big([a,b]\big) \le \mathrm{US}(f,Z) \le \mathrm{OS}(f,Z) \le V\big([a,b]\big) \cdot M.$$

Beweis: Der Beweis geht wie im eindimensionalen Fall in Lemma 19.5.

a. Wir stellen zunächst folgende Vorüberlegung an. Ist $Q \subseteq [a,b]$ ein Quader und zerlegen wir ihn in zwei Quader $Q = Q_1 \cup Q_2$ indem wir eine Seite zweiteilen, so gilt

$$\inf_{x\in Q} f(x) \le \inf_{x\in Q_i} f(x)$$

und mithin

$$\begin{aligned} V(Q)\cdot \inf_{x\in Q} f(x) =& V(Q_1)\cdot \inf_{x\in Q} f(x) + V(Q_2)\cdot \inf_{x\in Q} f(x) \\ \le& V(Q_1)\cdot \inf_{x\in Q_1} f(x) + V(Q_2)\cdot \inf_{x\in Q_2} f(x). \end{aligned}$$

Für die Differenz der beiden Seiten folgt damit

$$\begin{aligned} 0 &\le V(Q_1)\cdot \inf_{x\in Q_1} f(x) + V(Q_2)\cdot \inf_{x\in Q_2} f(x) - V(Q)\cdot \inf_{x\in Q} f(x) \\ &= V(Q_1)\cdot \left(\inf_{x\in Q_1} f(x) - \inf_{x\in Q} f(x)\right) + V(Q_2)\cdot \left(\inf_{x\in Q_2} f(x) - \inf_{x\in Q} f(x)\right) \\ &\le V(Q_1)\cdot (M+M) + V(Q_2)\cdot (M+M) = 2\cdot M\cdot V(Q). \end{aligned} \tag{121}$$

Wir betrachten dann, ähnlich wie im Eindimensionalen, zunächst den Fall, daß die Zerlegung $Z = Z_1 \times \ldots \times Z_n$ zu $Z' = Z_1 \times \ldots \times Z_i' \times \ldots \times Z_n$ verfeinert wird, indem wir zu $Z_i = (x_{i0}, \ldots, x_{im_i})$ einen Stützpunkt y zwischen x_{ij-1} und x_{ij} hinzufügen (siehe Abbildung 15). Wir zerteilen dadurch genau die Teilquader in $\mathrm{TQ}(Z)$, die in Komponente i den zusätzlichen Stützpunkt y enthalten und deren Vereinigung der Quader

$$P = [a_1,b_1] \times \ldots \times [a_{i-1},b_{i-1}] \times [x_{ij-1},x_{ij}] \times [a_{i+1},b_{i+1}] \times \ldots \times [a_n,b_n]$$

ist (in Abbildung 15 entspricht das dem grau schattierten Bereich). Für das Gesamtvolumen dieser Teilquader gilt dann

$$V(P) = V_i \cdot (x_{ij} - x_{ij-1}) \le V_i \cdot l(Z).$$

Für die Differenz der beiden Untersummen erhalten wir wegen (121) mithin die Abschätzung

$$0 \le \mathrm{US}(f,Z') - \mathrm{US}(f,Z) \le 2\cdot M \cdot l(Z) \cdot V_i, \tag{122}$$

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 $Z' = Z'_1 \times \ldots \times Z'_n$ 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 $k = 0, \ldots, m$ die so entstehenden Zwischenzerlegungen mit $Z^0 = Z$ und $Z^m = Z'$ und sei $i_k \in \{1, \ldots, n\}$ die Komponente, in der im Schritt k ein Stützpunkt eingefügt wird. Wir erhalten aus (122) und wegen $l(Z^k) \leq l(Z)$ dann

$$\begin{aligned} 0 \leq \mathrm{US}(f, Z') - \mathrm{US}(f, Z) &= \sum_{k=1}^{m} \mathrm{US}(f, Z^k) - \mathrm{US}(f, Z^{k-1}) \\ &\leq \sum_{k=1}^{m} 2 \cdot M \cdot l(Z^{k-1}) \cdot V_{i_k} \\ &\leq 2 \cdot M \cdot l(Z) \cdot \sum_{k=1}^{m} V_{i_k} \\ &= 2 \cdot M \cdot l(Z) \cdot \sum_{i=1}^{n} V_i \cdot \big(|Z'_i| - |Z_i|\big), \end{aligned}$$

wobei man für den letzten Schritt beachte, daß $|Z'_i| - |Z_i|$ genau die Anzahl der Stützpunkte ist, die in der i-ten Komponente hinzugefügt worden sind. Die Aussage für Obersummen zeigt man analog.

b. Wir betrachten die gemeinsame Verfeinerung $Z * Z'$ und erhalten

$$\begin{aligned} \mathrm{US}(f, Z) \overset{a.}{\leq} \mathrm{US}(f, Z * Z') &= \sum_{Q \in \mathrm{TQ}(Z * Z')} V(Q) \cdot \inf_{x \in Q} f(x) \\ &\leq \sum_{Q \in \mathrm{TQ}(Z * Z')} V(Q) \cdot \sup_{x \in Q} f(x) = \mathrm{OS}(f, Z * Z') \overset{a.}{\leq} \mathrm{OS}(f, Z'). \end{aligned}$$

c. Dies folgt aus a., da Z eine Verfeinerung der Zerlegung $(a_1, b_1) \times \ldots \times (a_n, b_n)$ ist und da $M \geq \sup\big\{f(x) \;\big|\; x \in [a, b]\big\} \geq \inf\big\{f(x) \;\big|\; x \in [a, b]\big\} \geq -M$.

□

Wie beim Riemann-Integral in der eindimensionalen Analysis können wir uns auch jetzt wieder zunutze machen, daß die Menge der Obersummen und die Menge der Untersummen nach Lemma 27.8 c. beschränkt sind. Deshalb existieren ihr Infimum und ihr Supremum, und wenn diese übereinstimmen, dann wollen wir die Funktion Riemann-integrierbar nennen.

Definition 27.9 (Riemann-integrierbar)
Seien $a, b \in \mathbb{R}^n$ mit $a < b$ und sei $f : [a, b] \longrightarrow \mathbb{R}$ beschränkt.
Wir definieren das *Oberintegral*

$$\mathrm{OI}(f) := \inf \left\{ \mathrm{OS}(f, Z) \;\middle|\; Z \text{ Zerlegung von } [a, b] \right\}$$

von f und das *Unterintegral*

$$\mathrm{UI}(f) := \sup \left\{ \mathrm{US}(f, Z) \;\middle|\; Z \text{ Zerlegung von } [a, b] \right\}$$

von f. Wegen Lemma 27.8 b. und Lemma 8.16 gilt

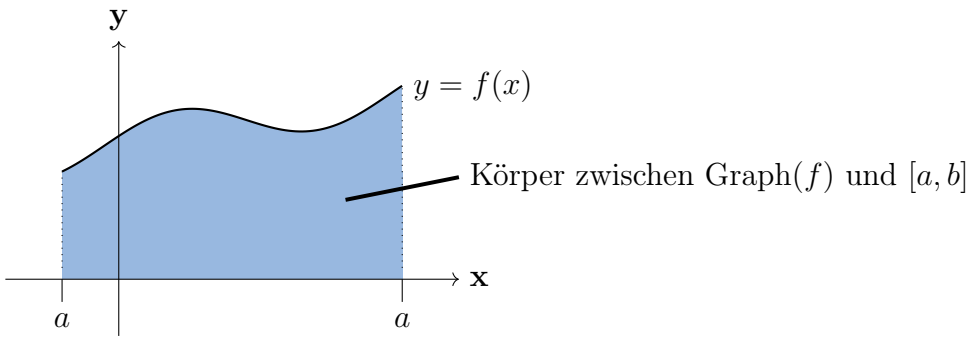
$$\mathrm{UI}(f) \leq \mathrm{OI}(f).$$

Wir nennen f *(Riemann-)integrierbar* auf $[a, b]$, falls $\mathrm{UI}(f) = \mathrm{OI}(f)$. Dann heißt

$$\int_{[a,b]} f(x)\, dx := \mathrm{OI}(f) \in \mathbb{R}$$

das *Integral* von f auf $[a, b]$.

Bemerkung 27.10
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 $[a, b]$ einschließt, und die Obersummen von f sind durch dieses nach unten beschränkt. Aufgrund der Definition von $\mathrm{OI}(f)$ als Infimum und $\mathrm{UI}(f)$ als Supremum gilt also stets $\mathrm{UI}(f) \leq V \leq \mathrm{OI}(f)$. Daß f integrierbar ist, bedeutet mithin nichts anderes, als daß das Integral $\int_{[a,b]} f(x)\, dx$ das Volumen des Körpers beschreibt, den der Graph von f mit dem Quader $[a, b]$ einschließt.

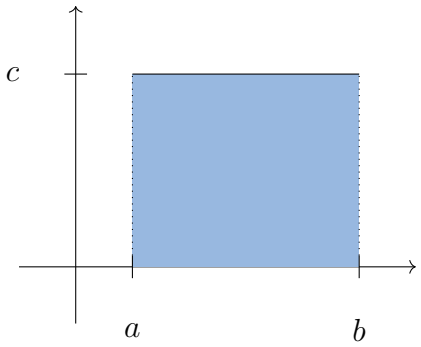


Beispiel 27.11
Jede konstante Funktion $f : Q \longrightarrow \mathbb{R} : x \mapsto c$ auf einem Quader $Q \subseteq \mathbb{R}^n$ ist integrierbar mit

$$\int_Q f(x)\, dx = V(Q) \cdot c.$$

Denn dann gilt für jede Zerlegung Z von Q bereits

$$\mathrm{OS}(f,Z) = \sum_{Q'\in \mathrm{TQ}(Z)} V(Q')\cdot c = V(Q)\cdot c = \sum_{Q'\in \mathrm{TQ}(Z)} V(Q')\cdot c = \mathrm{US}(f,Z).$$



Insbesondere ist das Volumen

$$V(Q) = \int_Q 1\,dx$$

von Q also der Wert des Integrals der Funktion 1 über Q.

C) Das Riemannsche Integrabilitätskriterium

Satz 27.12 (Riemannsches Integrabilitätskriterium)
Seien $a,b\in\mathbb{R}^n$ mit $a<b$ und sei $f:[a,b]\to\mathbb{R}$ beschränkt.
Genau dann ist f integrierbar auf $[a,b]$, wenn

$$\forall\,\varepsilon>0\ \ \exists\ Z \text{ Zerlegung von } [a,b]\ :\ \mathrm{OS}(f,Z)-\mathrm{US}(f,Z)<\varepsilon.$$

Beweis: Der Beweis geht wörtlich wie im eindimensionalen Fall, siehe Satz 19.11.

$\Longrightarrow$: Sei zunächst f integrierbar auf $[a,b]$ und sei $\varepsilon>0$ gegeben. Aufgrund der Definition von $\mathrm{UI}(f)$ als Supremum und $\mathrm{OI}(f)$ als Infimum und wegen Proposition 8.15 gibt es Zerlegungen Z' und Z'' von $[a,b]$ mit

$$\mathrm{OI}(f)+\frac{\varepsilon}{2} > \mathrm{OS}(f,Z') \overset{27.8}{\geq} \mathrm{OS}(f,Z'*Z'')$$

und

$$\mathrm{UI}(f)-\frac{\varepsilon}{2} < \mathrm{US}(f,Z'') \overset{27.8}{\leq} \mathrm{US}(f,Z'*Z'').$$

Damit erhalten wir mit $Z=Z'*Z''$ und wegen $\mathrm{UI}(f)=\mathrm{OI}(f)$

$$\mathrm{OS}(f,Z)-\mathrm{US}(f,Z) < \Big(\mathrm{OI}(f)+\frac{\varepsilon}{2}\Big)-\Big(\mathrm{UI}(f)-\frac{\varepsilon}{2}\Big)=\varepsilon.$$

$\Longleftarrow$: Für $\varepsilon:=\frac{1}{n}$ mit $n\geq 1$ gibt es eine Zerlegung Z^n von $[a,b]$ mit

$$\frac{1}{n} > \mathrm{OS}(f,Z^n)-\mathrm{US}(f,Z^n) \geq \mathrm{OI}(f)-\mathrm{UI}(f)\geq 0.$$

Da die linke Seite der Ungleichung für $n \to \infty$ gegen Null konvergiert, folgt im Grenzwert

$$0 = \lim_{n\to\infty} \frac{1}{n} \geq \mathrm{OI}(f) - \mathrm{UI}(f) \geq 0,$$

also $\mathrm{OI}(f) = \mathrm{UI}(f)$. Mithin ist f integrierbar auf $[a, b]$.

□

Satz 27.13 (Stetige Funktionen sind integrierbar.)
Ist $f : [a, b] \longrightarrow \mathbb{R}$ stetig für $a, b \in \mathbb{R}^n$ mit $a < b$, so ist f integrierbar auf $[a, b]$.

Beweis: Der Beweis geht genau wie im eindimensionalen Fall, siehe Satz 19.13.

Der Quader $[a, b]$ ist nach Beispiel 22.45 kompakt, und die stetige Funktion f ist dann nach Satz 23.29 gleichmäßig stetig auf $[a, b]$ und nach Korollar 23.24 auch beschränkt.

Sei nun $\varepsilon > 0$ gegeben. Da f gleichmäßig stetig auf $[a, b]$ ist, gibt es $\delta_\varepsilon > 0$, so daß

$$|f(x) - f(y)| < \frac{\varepsilon}{V([a, b])} \tag{123}$$

für alle $x, y \in [a, b]$ mit $||x - y||_\infty < \delta_\varepsilon$. Wir wählen nun eine Zerlegung Z von $[a, b]$ der Länge $l(Z) < \delta_\varepsilon$. Sei

$$\mathrm{TQ}(Z) = \{Q_1, \ldots, Q_m\}$$

die Menge der Teilquader von Z. Da f stetig auf Q_i ist, existieren nach Korollar 23.24 $y_i, z_i \in Q_i$ mit

$$f(y_i) = \sup\big\{f(y) \;\big|\; y \in Q_i\big\}$$

und

$$f(z_i) = \inf\big\{f(y) \;\big|\; y \in Q_i\big\},$$

und wegen

$$||y_i - z_i||_\infty = \max\big\{|y_{ij} - z_{ij}| \;\big|\; j = 1, \ldots, n\big\} \leq l(Z) < \delta_\varepsilon$$

folgt aus (123) zudem

$$0 \leq f(y_i) - f(z_i) < \frac{\varepsilon}{V([a, b])}.$$

Damit erhalten wir insbesondere

$$\begin{aligned}\mathrm{OS}(f, Z) - \mathrm{US}(f, Z) &= \sum_{i=1}^{m} V(Q_i) \cdot \big(f(y_i) - f(z_i)\big) \\ &< \sum_{i=1}^{m} V(Q_i) \cdot \frac{\varepsilon}{V([a, b])} \overset{(120)}{=} V([a, b]) \cdot \frac{\varepsilon}{V([a, b])} = \varepsilon.\end{aligned}$$

Somit ist f integrierbar nach dem Riemannschen Integrabilitätskriterium 27.12. □

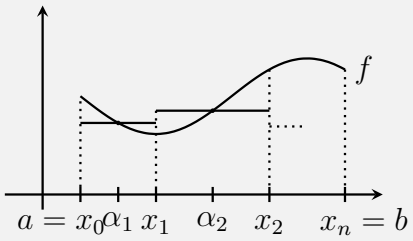
D) Riemannsche Zwischensummen und das Folgenkriterium

Definition 27.14

Seien $a, b \in \mathbb{R}^n$, $a < b$, $f : [a, b] \to \mathbb{R}$ beschränkt und Z eine Zerlegung von $[a, b]$. Erfüllt $\alpha = (\alpha_Q \mid Q \in \mathrm{TQ}(Z))$ die Bedingung $\alpha_Q \in Q$ für $Q \in \mathrm{TQ}(Z)$, so heißt

$$\mathrm{ZS}(f, Z, \alpha) := \sum_{Q \in \mathrm{TQ}(Z)} V(Q) \cdot f(\alpha_Q)$$

die *Riemannsche Zwischensumme* von f bezüglich der Zerlegung Z und den *Zwischenpunkten* α. Wir bezeichnen mit $\mathcal{ZP}\big([a, b]\big)$ die Menge aller Paare (Z, α) mit Z Zerlegung von $[a, b]$ mit Zwischenpunkten α.



Das nächste Lemma sagt, daß man Obersummen und Untersummen beliebig gut approximieren kann durch Zwischensummen.

Lemma 27.15

Sei $f : [a, b] \to \mathbb{R}$ 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 $0 \leq \mathrm{OS}(f, Z) - \mathrm{ZS}(f, Z, \alpha) < \varepsilon$.

b. Dann gibt es Zwischenpunkte β von Z mit $0 \leq \mathrm{ZS}(f, Z, \beta) - \mathrm{US}(f, Z) < \varepsilon$.

Beweis: Der Beweis geht genau wie im eindimensionalen Fall.

Für $Q \in \mathrm{TQ}(Z)$ sei

$$M_Q := \sup\big\{f(x) \;\big|\; x \in Q\big\}.$$

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 $\alpha = (\alpha_Q \mid Q \in \mathrm{TQ}(Z))$

$$\begin{aligned}\mathrm{OS}(f,Z) - \mathrm{ZS}(f,Z,\alpha) &= \sum_{Q\in\mathrm{TQ}(Z)} V(Q)\cdot\big(M_Q - f(\alpha_Q)\big)\\ &< \sum_{Q\in\mathrm{TQ}(Z)} V(Q)\cdot\frac{\varepsilon}{V([a,b])}\\ &\overset{(120)}{=} V([a,b])\cdot\frac{\varepsilon}{V([a,b])} = \varepsilon.\end{aligned}$$

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 27.16
Ist $f : [a,b] \longrightarrow \mathbb{R}$ integrierbar für $a, b \in \mathbb{R}^n$ mit $a < b$, so gilt:

$\forall\, \varepsilon > 0\ \exists\, \delta_\varepsilon > 0\ :\ \forall\, Z$ Zerlegung mit $l(Z) < \delta_\varepsilon$ gilt $\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon$.

Beweis: Der Beweis geht genau wie im Eindimensionalen, siehe Lemma 19.12.

Sei $\varepsilon > 0$ gegeben. Aus dem Riemannschen Integrabilitätskriterium erhalten wir eine Zerlegung Z' von $[a,b]$, so daß

$$\mathrm{OS}(f,Z') - \mathrm{US}(f,Z') < \frac{\varepsilon}{2}. \tag{124}$$

Wir setzen nun

$$\delta_\varepsilon := \frac{\varepsilon}{8\cdot M\cdot\sum_{i=1}^n V_i\cdot|Z_i'|} > 0,$$

wobei $M := \sup\big\{|f(x)| \;\big|\; x\in[a,b]\big\}$ und $V_i = \prod_{j\neq i}(b_j - a_j)$. Ist Z eine Zerlegung von $[a,b]$ mit $l(Z) < \delta_\varepsilon$, so folgt aus Lemma 27.8 wegen $|Z_i * Z_i'| - |Z_i| \le |Z_i'|$

$$\begin{aligned}\mathrm{OS}(f,Z) - \mathrm{OS}(f,Z*Z') &\le 2\cdot M\cdot l(Z)\cdot\sum_{i=1}^n V_i\cdot(|Z_i*Z_i'| - |Z_i|)\\ &< 2\cdot M\cdot\delta_\varepsilon\cdot\sum_{i=1}^n V_i\cdot|Z_i'| = \frac{\varepsilon}{4}\end{aligned} \tag{125}$$

und analog

$$\mathrm{US}(f,Z*Z') - \mathrm{US}(f,Z) < \frac{\varepsilon}{4}. \tag{126}$$

Da $Z * Z'$ eine Verfeinerung von Z' ist, folgt aus (124) zusammen mit Lemma 19.5

$$\mathrm{OS}(f,Z*Z') - \mathrm{US}(f,Z*Z') \le \mathrm{OS}(f,Z') - \mathrm{US}(f,Z') < \frac{\varepsilon}{2}. \tag{127}$$

Insgesamt erhalten wir damit

$$\begin{aligned}\mathrm{OS}(f,Z)-\mathrm{US}(f,Z) &= \mathrm{OS}(f,Z)-\mathrm{OS}(f,Z*Z')+\mathrm{OS}(f,Z*Z')-\mathrm{US}(f,Z*Z')\\ &\quad+\mathrm{US}(f,Z*Z')-\mathrm{US}(f,Z) \overset{(125)(127)(126)}{<} \frac{\varepsilon}{4}+\frac{\varepsilon}{2}+\frac{\varepsilon}{4}=\varepsilon.\end{aligned}$$

□

Satz 27.17 (Riemannsches Folgenkriterium für Integrierbarkeit)
Seien $a,b\in\mathbb{R}^n$ mit $a<b$, $f:[a,b]\longrightarrow\mathbb{R}$ sei beschränkt 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 $(Z^m,\alpha^m)_{m\in\mathbb{N}}$ von Zerlegungen von $[a,b]$ und Zwischenpunkten mit $l(Z^m)\to 0$ gilt

$$\mathrm{ZS}(f,Z^m,\alpha^m)\longrightarrow I.$$

Beweis: Der Beweis geht genau wie im Eindimensionalen, siehe Satz 19.19.

$\Longrightarrow$**:** Es sei $(Z^m,\alpha^m)_{m\in\mathbb{N}}$ eine Folge von Zerlegungen von $[a,b]$ mit Zwischenpunkten, so daß $\lim_{m\to\infty} l(Z^m)=0$, und sei $I=\int_{[a,b]} f(x)\,dx$.
Sei $\varepsilon>0$ gegeben. Wir müssen ein $n_\varepsilon\in\mathbb{N}$ finden, so daß

$$|\mathrm{ZS}(f,Z^m,\alpha^m)-I|<\varepsilon \tag{128}$$

für alle $m\geq n_\varepsilon$.
Da f integrierbar ist, gibt es nach Lemma 27.16 ein $\delta_\varepsilon>0$, so daß für eine Zerlegung Z von $[a,b]$ aus $l(Z)<\delta_\varepsilon$ auch

$$\mathrm{OS}(f,Z)-\mathrm{US}(f,Z)<\varepsilon \tag{129}$$

gilt. Wegen $\lim_{m\to\infty} l(Z^m)=0$ 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 (129)

$$\mathrm{ZS}(f,Z^m,\alpha^m)-I\leq \mathrm{OS}(f,Z^m)-I\leq \mathrm{OS}(f,Z^m)-\mathrm{US}(f,Z^m)<\varepsilon$$

her, sowie

$$\mathrm{ZS}(f,Z^m,\alpha^m)-I\geq \mathrm{US}(f,Z^m)-I\geq \mathrm{US}(f,Z^m)-\mathrm{OS}(f,Z^m)>-\varepsilon.$$

Damit ist (128) für $m\geq n_\varepsilon$ erfüllt, und das heißt $\mathrm{ZS}(f,Z^m,\alpha^m)\longrightarrow I$.

$\Longleftarrow$**:** Wir wollen das Riemannsche Integrabilitätskriterium anwenden.
Sei dazu $\varepsilon>0$ gegeben. Wir betrachten die äquidistante Zerlegung Z^m aus Beispiel 27.3 für $m\geq 1$. Dann gilt

$$\lim_{m\to\infty} l(Z^m)=\lim_{m\to\infty}\frac{\max\{b_i-a_i\mid i=1,\ldots,n\}}{m}=0.$$

Mit Lemma 27.15 finden wir zu $m\in\mathbb{N}$ Zwischenpunkte α^m und β^m, so daß

$$\mathrm{OS}(f,Z^m)-\mathrm{ZS}(f,Z^m,\alpha^m)<\frac{\varepsilon}{4}$$

und

$$\mathrm{ZS}(f, Z^m, \beta^m) - \mathrm{US}(f, Z^m) < \frac{\varepsilon}{4}.$$

Außerdem gelten nach Voraussetzung

$$\mathrm{ZS}(f, Z^m, \alpha^m) \longrightarrow I$$

und

$$\mathrm{ZS}(f, Z^m, \beta^m) \longrightarrow I.$$

Wir finden also ein $n_\varepsilon \in \mathbb{N}$, so daß für alle $m \geq n_\varepsilon$ gilt

$$|\,\mathrm{ZS}(f, Z^m, \alpha^m) - I| < \frac{\varepsilon}{4}$$

und

$$|\,\mathrm{ZS}(f, Z^m, \beta^m) - I| < \frac{\varepsilon}{4}.$$

Dann gilt aber auch

$$\begin{aligned}
\mathrm{OS}(f, Z^{n_\varepsilon}) - \mathrm{US}(f, Z^{n_\varepsilon}) &= |\,\mathrm{OS}(f, Z^{n_\varepsilon}) - \mathrm{US}(f, Z^{n_\varepsilon})| \\
&\leq |\,\mathrm{OS}(f, Z^{n_\varepsilon}) - \mathrm{ZS}(f, Z^{n_\varepsilon}, \alpha^{n_\varepsilon})| + |\,\mathrm{ZS}(f, Z^{n_\varepsilon}, \alpha^{n_\varepsilon}) - I| \\
&\qquad + |I - \mathrm{ZS}(f, Z^{n_\varepsilon}, \beta^{n_\varepsilon})| + |\,\mathrm{ZS}(f, Z^{n_\varepsilon}, \beta^{n_\varepsilon}) - \mathrm{US}(f, Z^{n_\varepsilon})| \\
&< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon.
\end{aligned}$$

Also ist f integrierbar nach dem Riemannschen Integrabilitätskriterium 27.12. Außerdem haben wir für $m \geq n_\varepsilon$

$$\begin{aligned}
|\,\mathrm{OS}(f, Z^m) - I| \leq& |\,\mathrm{OS}(f, Z^m) - \mathrm{ZS}(f, Z^m, \alpha^m)| + |\,\mathrm{ZS}(f, Z^m, \alpha^m) - I| \\
<& \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < \varepsilon,
\end{aligned}$$

woraus

$$\mathrm{OS}(f, Z^m) \longrightarrow I$$

folgt. Analog sehen wir

$$\mathrm{US}(f, Z^m) \longrightarrow I.$$

Damit erhalten wir dann

$$I \longleftarrow \mathrm{US}(f, Z^m) \leq \int_{[a,b]} f(x)\,dx \leq \mathrm{OS}(f, Z^m) \longrightarrow I,$$

so daß $I = \int_{[a,b]} f(x)\,dx$ aus dem Einschachtelungssatz 11.17 folgt.

□

Korollar 27.18 (Riemannsches Zwischensummenkriterium)
Seien $a, b \in \mathbb{R}^n$ mit $a < b$, $f : [a, b] \longrightarrow \mathbb{R}$ sei beschränkt und $I \in \mathbb{R}$.
Genau dann ist f auf $[a, b]$ integrierbar mit $I = \int_{[a,b]} f(x)\,dx$, wenn

$$\forall\, \varepsilon > 0\ \exists\, \delta_\varepsilon > 0\ :\ \forall\, (Z, \alpha) \in \mathcal{ZP}\big([a, b]\big) \text{ mit } l(Z) < \delta_\varepsilon\ :\ |\,\mathrm{ZS}(f, Z, \alpha) - I| < \varepsilon.$$

Beweis:
$\Longrightarrow$: Ist f integrierbar mit $I = \int_{[a,b]} f(x)\,dx$, so gibt es zu $\varepsilon > 0$ nach Lemma 27.16 ein $\delta_\varepsilon > 0$, so daß aus $l(Z) < \delta_\varepsilon$ stets

$$\mathrm{OS}(f, Z) - \mathrm{US}(f, Z) < \varepsilon$$

folgt. Für beliebige Zwischenpunkte von Z gilt dann aber

$$|\,\mathrm{ZS}(f, Z, \alpha) - I| \le \mathrm{OS}(f, Z) - \mathrm{US}(f, Z) < \varepsilon,$$

da sowohl die Zwischensumme, als auch das Integral zwischen der Ober- und der Untersumme liegen.

$\Longleftarrow$: Sei $(Z^m, \alpha^m) \in \mathcal{ZP}\big([a,b]\big)$ mit $l(Z^m) \longrightarrow 0$. 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 $l(Z^m) < \delta_\varepsilon$ für alle $m \ge n_\varepsilon$, und nach Voraussetzung gilt deshalb

$$|\,\mathrm{ZS}(f, Z^m, \alpha^m) - I| < \varepsilon$$

für alle $m \ge n_\varepsilon$. Mithin konvergiert $\mathrm{ZS}(f, Z^m, \alpha^m)$ gegen I, und wegen des Riemannschen Folgenkriteriums für Integrierbarkeit 27.17 ist f also integrierbar mit $I = \int_{[a,b]} f(x)\,dx$. □

E) Erste einfache Eigenschaften des Integrals

Korollar 27.19 (Linearität und Monotonie des Integrals)
Seien $a, b \in \mathbb{R}^n$ mit $a < b$, $f, g : [a,b] \longrightarrow \mathbb{R}$ seien integrierbar 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 $f(x) \le g(x)$ für alle $x \in [a,b]$, so ist auch

$$\int_{[a,b]} f(x)\,dx \le \int_{[a,b]} g(x)\,dx.$$

Beweis: Der Beweis geht genau wie im Eindimensionalen, siehe Korollar 19.21.

a. Wir beachten zunächst, daß für jede Zerlegung Z von $[a,b]$ mit Zwischenpunkten α offenbar gilt:

$$\begin{aligned}\mathrm{ZS}(cf + dg, Z, \alpha) &= \sum_{Q \in \mathrm{TQ}(Z)} V(Q) \cdot (cf + dg)(\alpha_Q) \\ &= c \cdot \sum_{Q \in \mathrm{TQ}(Z)} V(Q) \cdot f(\alpha_Q) + d \cdot \sum_{Q \in \mathrm{TQ}(Z)} V(Q) \cdot g(\alpha_Q) \\ &= c \cdot \mathrm{ZS}(f, Z, \alpha) + d \cdot \mathrm{ZS}(g, Z, \alpha).\end{aligned}$$

Es sei nun $(Z^m, \alpha^m)_{m\in\mathbb{N}}$ eine Folge von Zerlegungen von $[a,b]$ und Zwischenpunkten mit $l(Z^m) \longrightarrow 0$. Aus den Grenzwertsätzen für Folgen 11.15 und Satz 27.17 folgt dann

$$\begin{aligned}\lim_{m\to\infty} \mathrm{ZS}(cf+dg, Z^m, \alpha^m) &= c \cdot \lim_{m\to\infty} \mathrm{ZS}(f, Z^m, \alpha^m) + d \cdot \lim_{m\to\infty} \mathrm{ZS}(g, Z^m, \alpha^m) \\ &= c \cdot \int_{[a,b]} f(x)\,dx + d \cdot \int_{[a,b]} g(x)\,dx.\end{aligned}$$

Das Riemannsche Folgenkriterium für Integrierbarkeit 27.17 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 $l(Z^m) \longrightarrow 0$. Wegen $f(x) \le g(x)$ für alle $x \in [a,b]$ gilt dann offenbar

$$\int_{[a,b]} f(x)\,dx \longleftarrow \mathrm{ZS}(f, Z^m, \alpha^m) \le \mathrm{ZS}(g, Z^m, \alpha^m) \longrightarrow \int_{[a,b]} g(x)\,dx,$$

wobei die Grenzwerte aus dem Riemannschen Folgenkriterium für Integrierbarkeit folgen. Damit gilt dann aber auch für die Grenzwerte

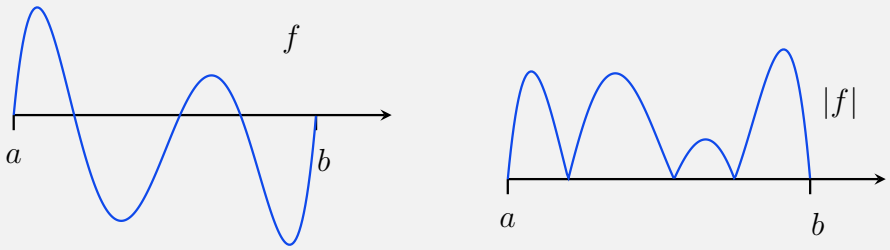
$$\int_{[a,b]} f(x)\,dx \le \int_{[a,b]} g(x)\,dx.$$

□

Proposition 27.20 (Dreiecksungleichung für Integrale)
Ist $f : [a,b] \longrightarrow \mathbb{R}$ integrierbar auf $[a,b]$ für $a, b \in \mathbb{R}^n$ mit $a < b$, so ist $|f|$ integrierbar auf $[a,b]$, und es gilt

$$\left|\int_{[a,b]} f(x)\,dx\right| \le \int_{[a,b]} |f(x)|\,dx.$$

Wir nennen das Integral über $|f|$ auch das *Volumen*, das der Graph von f mit dem Quader $[a,b]$ einschließt.



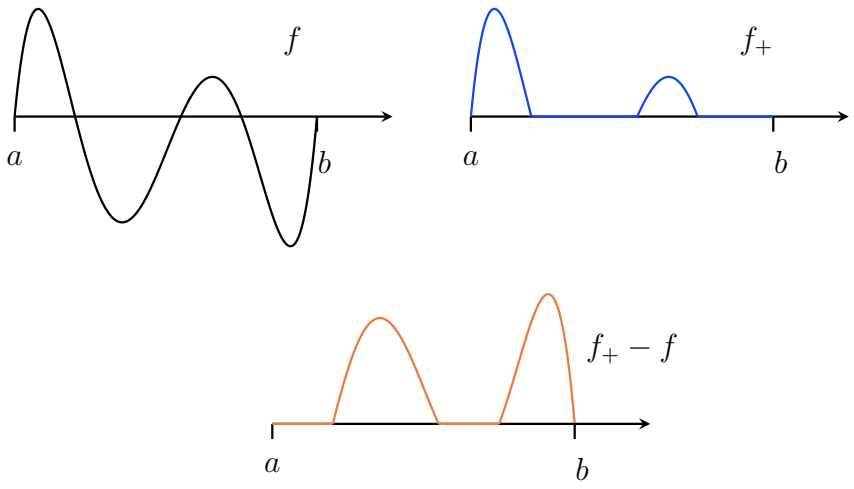
Beweis: Der Beweis geht wörtlich wie im Eindimensionalen, Proposition 19.26.

Wir betrachten die Funktion

$$f_+ : [a,b] \longrightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } f(x) \ge 0, \\ 0, & \text{falls } f(x) < 0. \end{cases}$$

Dann gilt

$$|f| = 2 \cdot f_+ - f.$$



Wir wollen nun mit Hilfe des Riemannschen Integrabilitätskriteriums 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

$$\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon.$$

Wir behaupten, daß für jede Teilmenge $Q \subseteq [a,b]$ die Ungleichung

$$\sup\{f_+(x) \mid x \in Q\} - \inf\{f_+(x) \mid x \in Q\} \leq \sup\{f(x) \mid x \in Q\} - \inf\{f(x) \mid x \in Q\}, \tag{130}$$

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 (130) 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 (130) nicht-negativ ist. In diesem Fall gilt (130).
2. **Fall:** $\exists\, y, z \in Q$ **mit** $f(y) < 0 \leq f(z)$**:** Also $\sup\{f_+(x) |\ x \in Q\} = \sup\{f(x) |\ x \in Q\}$ und $\inf\{f_+(x) |\ x \in Q\} = 0 > \inf\{f(x) \mid x \in Q\}$. Damit gilt die Ungleichung (130).
3. **Fall:** $f(x) \geq 0$ **für alle** $x \in Q$**:** Dann ist $f = f_+$ auf Q und (130) gilt.

Damit haben wir gezeigt, daß (130) stets erfüllt ist. Für die Differenz der Ober- und Untersumme von f_+ ergibt sich dann

$$\begin{aligned}
\mathrm{OS}(f_+,Z) - \mathrm{US}(f_+,Z) &= \sum_{Q \in \mathrm{TQ}(Z)} V(Q) \cdot \big(\sup\{f_+(x) \mid x \in Q\} - \inf\{f_+(x) \mid x \in Q\}\big) \\
&\leq \sum_{Q \in \mathrm{TQ}(Z)} V(Q) \cdot \big(\sup\{f(x) \mid x \in Q\} - \inf\{f(x) \mid x \in Q\}\big) \\
&= \mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon.
\end{aligned}$$

Mit Hilfe des Riemannschen Integrabilitätskriteriums 27.12 folgt dann, daß f_+ auf $[a,b]$ integrierbar ist. Aus der Linearität des Integrals 27.19 folgt damit, daß auch

$$|f| = 2 \cdot f_+ - f$$

auf $[a,b]$ integrierbar ist.

Für eine Zerlegung Z mit Zwischenpunkten α gilt

$$|\operatorname{ZS}(f,Z,\alpha)| = \left|\sum_{Q\in\operatorname{TQ}(Z)} V(Q)\cdot f(\alpha_Q)\right| \le \sum_{Q\in\operatorname{TQ}(Z)} V(Q)\cdot|f(\alpha_Q)| = \operatorname{ZS}(|f|,Z,\alpha).$$

Sei nun $(Z^m,\alpha^m)_{m\in\mathbb{N}}$ eine Folge von Zerlegungen von $[a,b]$ und Zwischenpunkten mit $\lim_{m\to\infty} l(Z^m) = 0$, dann folgt

$$\left|\int_{[a,b]} f(x)\,dx\right| \longleftarrow |\operatorname{ZS}(f,Z^m,\alpha^m)| \le \operatorname{ZS}(|f|,Z^m,\alpha^m) \longrightarrow \int_{[a,b]} |f(x)|\,dx.$$

Die Ungleichung bleibt für die Grenzwerte erhalten. □

F) 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 27.21 (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

$$X \times Y = [a,b]\times[c,d] = \big[(a,c),(b,d)\big] \subset \mathbb{R}^{p+q}$$

ein Quader, und es gilt

$$V(X\times Y) = \prod_{i=1}^{p}(b_i-a_i)\cdot\prod_{j=1}^{q}(d_j-c_j) = V(X)\cdot V(Y). \tag{131}$$

Ist $Z_x = Z_1\times\ldots\times Z_p$ eine Zerlegung von X und $Z_y = Z_{p+1}\times\ldots\times Z_{p+q}$ eine Zerlegung von Y, so ist

$$Z_x \times Z_y = Z_1\times\ldots\times Z_{p+q}$$

eine Zerlegung von $X\times Y$, und es gilt

$$\operatorname{TQ}(Z_x\times Z_y) = \big\{P\times Q \;\big|\; P\in\operatorname{TQ}(Z_x), Q\in\operatorname{TQ}(Z_y)\big\}. \tag{132}$$

Sind ferner $\alpha = \big(\alpha_P \mid P\in\operatorname{TQ}(Z_x)\big)$ Zwischenpunkte von Z_x und $\beta = \big(\beta_Q \mid Q\in\operatorname{TQ}(Z_y)\big)$ Zwischenpunkte von Z_y, so erhalten wir Zwischenpunkte von $Z_x\times Z_y$ als

$$\alpha\times\beta := \big((\alpha_P,\beta_Q)^t \mid P\in\operatorname{TQ}(Z_x), Q\in\operatorname{TQ}(Z_y)\big).$$

Satz 27.22 (Satz von Fubini für Quader)
Es seien $X \subset \mathbb{R}^p$ und $Y \subset \mathbb{R}^q$ zwei Quader, $f : X \times Y \longrightarrow \mathbb{R}$ sei integrierbar und die Funktion

$$g : Y \longrightarrow \mathbb{R} : y \mapsto \int_X f(x,y)\, dx$$

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 $\lim\limits_{m\to\infty} l(Z_y^m) = 0$, und wegen des Riemannschen Folgenkriteriums für Integrierbarkeit 27.17 reicht es, zu zeigen, daß

$$\lim_{m\to\infty} \mathrm{ZS}(g, Z_y^m, \beta^m) = I,$$

d.h. für alle $\varepsilon > 0$ gibt es ein $n_\varepsilon \in \mathbb{N}$, so daß

$$\left|\mathrm{ZS}(g, Z_y^m, \beta^m) - I\right| < \varepsilon$$

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 $\lim\limits_{k\to\infty} l(Z_x^k) = 0$.
Sei nun $\varepsilon > 0$ gegeben. Aufgrund des Riemannschen Zwischensummenkriteriums 27.18 und weil f auf $X \times Y$ integrierbar ist, gibt es dann ein $\delta_\varepsilon > 0$, so daß für alle $(Z,\gamma) \in \mathcal{ZP}(X \times Y)$ mit $l(Z) < \delta_\varepsilon$

$$|\mathrm{ZS}(f, Z, \gamma) - I| < \frac{\varepsilon}{2} \tag{133}$$

gilt. Wegen $\lim\limits_{m\to\infty} l(Z_y^m) = 0$ und $\lim\limits_{k\to\infty} l(Z_x^k) = 0$ gibt es ein $n_\varepsilon > 0$, so daß

$$l(Z_x^k \times Z_y^m) = \max\left\{l(Z_x^k), l(Z_y^m)\right\} < \delta_\varepsilon$$

für alle $k, m \geq n_\varepsilon$, und wegen (133) folgt dann

$$\left|\mathrm{ZS}(f, Z_x^k \times Z_y^m, \alpha^k \times \beta^m) - I\right| < \frac{\varepsilon}{2} \tag{134}$$

für alle $k, m \geq n_\varepsilon$. Für die Zwischensumme auf der linken Seite der Ungleichung gilt

$$\begin{aligned}\mathrm{ZS}(f, Z_x^k \times Z_y^m, \alpha^k \times \beta^m) &\overset{(132)}{=} \sum_{P\times Q \in \mathrm{TQ}(Z_x^k \times Z_y^m)} V(P\times Q)\cdot f(\alpha_P^k, \beta_Q^m)\\ &\overset{(131)}{=} \sum_{Q\in \mathrm{TQ}(Z_y^m)} V(Q)\cdot \sum_{P\in \mathrm{TQ}(Z_x^k)} V(P)\cdot f(\alpha_P^k, \beta_Q^m).\end{aligned}$$

Nach Voraussetzung ist die Funktion $x \mapsto f(x, \beta_Q^m)$ auf X integrierbar und aus dem Riemannschen Folgenkriterium für Integrierbarkeit 27.17 folgt mit $\lim\limits_{k\to\infty} l(Z_x^k) = 0$ dann

$$\lim_{k\to\infty} \sum_{P\in \mathrm{TQ}(Z_x^k)} V(P)\cdot f(\alpha_P^k, \beta_Q^m) = \int_X f(x, \beta_Q^m)\, dx = g(\beta_Q^m)$$

und somit

$$\lim_{k\to\infty} \mathrm{ZS}(f, Z_x^k \times Z_y^m, \alpha^k \times \beta^m) = \sum_{Q\in \mathrm{TQ}(Z_y^m)} V(Q)\cdot g(\beta_Q^m) = \mathrm{ZS}(g, Z_y^m, \beta^m).$$

Mit (134) erhalten wir deshalb

$$\left|\mathrm{ZS}(g, Z_y^m, \beta^m) - I\right| = \lim_{k\to\infty} \left|\mathrm{ZS}(f, Z_x^k \times Z_y^m, \alpha^k \times \beta^m) - I\right| \leq \frac{\varepsilon}{2} < \varepsilon$$

für alle $m \geq n_\varepsilon$. Damit ist der Satz von Fubini für Quader bewiesen. □

Beispiel 27.23 (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

$$\mathrm{TQ}\left(Z^{2^m}\right) = \left\{ Q_{ij} = \left[\tfrac{i-1}{2^m}, \tfrac{i}{2^m}\right] \times \left[\tfrac{j-1}{2^m}, \tfrac{j}{2^m}\right] \;\middle|\; i,j = 1, \ldots, 2^m \right\}$$

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

$$\begin{aligned} \mathrm{OS}\left(f, Z^{2^m}\right) &= \sum_{j=1}^{2^m} \left(\sum_{i \leq 2^{m-1}} \frac{1}{2^{2m}} \cdot 1 + \sum_{i > 2^{m-1}} \frac{1}{2^{2m}} \cdot \frac{i}{2^{m-1}} \right) \\ &= 2^m \cdot \frac{1}{2^{2m}} \cdot \left(\sum_{i \leq 2^{m-1}} 1 + \frac{1}{2^{m-1}} \cdot \sum_{i=2^{m-1}+1}^{2^m} i \right) \\ &= \frac{1}{2^m} \cdot \left(2^{m-1} + \frac{2^m \cdot (2^m+1) - 2^{m-1} \cdot (2^{m-1}+1)}{2^m} \right) \\ &= \frac{1}{2} + \frac{3}{4} + \frac{1}{2^{m+1}} = \frac{5}{4} + \frac{1}{2^{m+1}} \geq \frac{5}{4} \end{aligned}$$

und für die Untersumme gilt analog

$$\begin{aligned} \mathrm{US}\left(f, Z^{2^m}\right) &= \sum_{j=1}^{2^m} \left(\sum_{i \leq 2^{m-1}} \frac{1}{2^{2m}} \cdot \frac{i-1}{2^{m-1}} + \sum_{i > 2^{m-1}} \frac{1}{2^{2m}} \cdot 1 \right) \\ &= \frac{1}{2^m} \cdot \left(\frac{2^{m-1} \cdot (2^{m-1}-1)}{2^m} + 2^{m-1} \right) \\ &= \frac{1}{4} - \frac{1}{2^{m+1}} + \frac{1}{2} = \frac{3}{4} - \frac{1}{2^{m+1}} \leq \frac{3}{4}. \end{aligned}$$

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

$$\mathrm{OS}(f, Z) - \mathrm{US}(f, Z) < \frac{1}{4}.$$

Aber $l(Z^{2^m}) \longrightarrow 0$ für $m \to \infty$, so daß für m hinreichend groß auch $l(Z^{2^m}) < \delta_\varepsilon$ gesichert ist, im Widerspruch zu

$$\mathrm{OS}\left(f, Z^{2^m}\right) - \mathrm{US}\left(f, Z^{2^m}\right) \geq \frac{1}{2}.$$

Also ist f nicht integrierbar auf $[0,1] \times [0,1]$. □

G) Satz über die Vertauschbarkeit der Integrationsreihenfolge

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 27.24 (Vertauschbarkeit der Integrationsreihenfolge)
Es seien $X \subset \mathbb{R}^p$ und $Y \subset \mathbb{R}^q$ zwei Quader, $f : X \times Y \longrightarrow \mathbb{R}$ sei integrierbar und die Funktionen

$$g : Y \longrightarrow \mathbb{R} : y \mapsto \int_X f(x,y)\, dx$$

und

$$h : X \longrightarrow \mathbb{R} : x \mapsto \int_Y f(x,y)\, dy$$

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 27.22. Die Aussage für Y und die erste Gleichheit der Integrale folgt analog, indem man die Rollen von X und Y im Satz von Fubini 27.22 vertauscht. □

Die Funktionen g und h in Korollar 27.24 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 27.25 (Iterierte Integration mehrdimensionaler Integrale)
Es seien $a, b \in \mathbb{R}^n$ mit $a < b$ und $f : [a, b] \longrightarrow \mathbb{R}$ 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 27.24. □

Beispiel 27.26
Wir wollen als einfaches Beispiel den Wert des Integrals der Funktion

$$f : [0,1] \times [0,1] \longrightarrow \mathbb{R} : x \mapsto x_1 + x_2$$

auf dem Quader $Q = [0,2] \times [0,2]$ aus Beispiel 27.6 berechnen

$$\begin{aligned} \int_Q f(x)\, dx &= \int_0^2 \int_0^2 x_1 + x_2\, dx_1\, dx_2 = \int_0^2 \frac{x_1^2}{2} + x_1 x_2 \Big|_0^2 dx_2 \\ &= \int_0^2 2 + 2x_2\, dx_2 = 2x_2 + x_2^2 \Big|_0^2 = 8. \end{aligned}$$

H) Der Mittelwertsatz für Riemann-Integrale auf Quadern

Satz 27.27 (Mittelwertsatz der Integralrechnung für Quader)
Seien $a, b \in \mathbb{R}^n$ mit $a < b$, $f : [a, b] \longrightarrow \mathbb{R}$ stetig. Dann gibt es ein $c \in [a, b]$ mit

$$\int_{[a,b]} f(x)\, dx = f(c) \cdot V\big([a, b]\big).$$

Beweis: Der Beweis geht genau wie im Eindimensionalen, siehe den alternativen Beweis von Korollar 20.9.

Da f stetig auf dem Kompaktum $[a, b]$ ist, gibt es nach Korollar 23.24 $y, z \in [a, b]$ mit

$$f(y) \le f(x) \le f(z)$$

für alle $x \in [a, b]$. Für die triviale Zerlegung Z von $[a, b]$ mit nur einem Teilquader gilt dann

$$f(y) \cdot V\big([a, b]\big) = \mathrm{US}(f, Z) \le \int_{[a,b]} f(x)\, dx \le \mathrm{OS}(f, Z) = f(z) \cdot V\big([a, b]\big). \tag{135}$$

Die Funktion

$$g : [0,1] \longrightarrow \mathbb{R} : t \mapsto f\big(y + t \cdot (z - y)\big)$$

ist stetig, da f stetig auf $\overline{yz} \subset [a,b]$ ist, und aus (135) folgt

$$g(0) = f(y) \leq \frac{1}{V\big([a,b]\big)} \cdot \int_{[a,b]} f(x)\, dx \leq f(z) = g(1).$$

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

$$f(c) = g(t) = \frac{1}{V\big([a,b]\big)} \cdot \int_{[a,b]} f(x)\, dx$$

für $c = y + t \cdot (z - y) \in [a,b]$. □

Aufgaben

Aufgabe 27.28

Berechne das folgende Integral:

$$\int\limits_{[0,\frac{\pi}{2}]\times[0,\frac{\pi}{2}]} \sin(x+y)\, d(x,y)$$

Aufgabe 27.29

Berechne das folgende Integral:

$$\int\limits_{[0,1]\times[0,1]\times[0,1]} \frac{x^3y^2}{1+z^2}\, d(x,y,z).$$

Aufgabe 27.30

Für eine rationale Zahl $x = \frac{a}{b}$ mit $a, b > 0$ setzen wir $N(x) = \frac{b}{\operatorname{ggt}(a,b)}$, 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 27.31

Zeige, sind $f, g : [a,b] \longrightarrow \mathbb{R}$ auf dem Intervall $[a,b] \subseteq \mathbb{R}$ integrierbar, so ist

$$[a,b] \times [a,b] \longrightarrow \mathbb{R} : (x,y) \mapsto f(x) \cdot g(y)$$

auf $[a,b] \times [a,b]$ integrierbar und

$$\int\limits_{[a,b]\times[a,b]} f(x) \cdot g(y)\, d(x,y) = \int_a^b f(x)\, dx \cdot \int_a^b g(y)\, dy.$$

Aufgabe 27.32

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 27.33

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.

Aufgabe 27.34

Berechne das Volumen des Tetraeders, der von den drei Koordinatenachsen und der Ebene $x_3 = 2 - 2x_1 - x_2$ begrenzt wird.

Aufgabe 27.35 (Gegenbeispiel zum Satz über die Vertauschbarkeit der Integrationsreihenfolge)

Zeige, daß die beiden Doppelintegrale in der folgenden Ungleichung existieren und daß ihr Wert nicht übereinstimmt:

$$\int_0^1 \int_0^1 \frac{x-y}{(x+y)^3}\, dx\, dy \neq \int_0^1 \int_0^1 \frac{x-y}{(x+y)^3}\, dy\, dx.$$

Wie ist dies mit dem Satz von Fubini über die Vertauschbarkeit der Integrationsreihenfolge verträglich? Was gilt für den Grenzwert

$$\lim_{\substack{(a,b)\to(0,0)\\ a,b>0}} \int_{[a,1]\times[b,1]} \frac{x-y}{(x+y)^3}\, dx\, dy.$$

§ 28 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 28.1 (Nullmenge)
Eine Teilmenge $N \subsetneqq \mathbb{R}^n$ 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. $N \subsetneqq \bigcup_{m\in\mathbb{N}} Q_m$, 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 28.2 (Nullmengen)
Eine Hyperebene der Form

$$V(x_j - c) = \{x \in \mathbb{R}^n \mid x_j = c\}$$

für ein $j \in \{1, \ldots, n\}$ 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

$$Q_m = \left[a^m, b^m\right]$$

durch

$$a_i^m := -m \quad \text{und} \quad b_i^m := m$$

für $i \neq j$ und

$$a_j^m := c - \frac{\varepsilon}{2^{m+1} \cdot 2^n \cdot m^{n-1}}$$

sowie

$$b_j^m := c + \frac{\varepsilon}{2^{m+1} \cdot 2^n \cdot m^{n-1}}.$$

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

$$V(Q_m) = (2m)^{n-1} \cdot \frac{2 \cdot \varepsilon}{2^{m+1} \cdot 2^n \cdot m^{n-1}} = \frac{\varepsilon}{2^{m+1}},$$

und damit ist die Reihe der Volumina konvergent mit

$$\sum_{m=1}^{\infty} V(Q_m) = \frac{\varepsilon}{2} \cdot \sum_{m=1}^{\infty} \frac{1}{2^m} = \frac{\varepsilon}{2} < \varepsilon.$$

Bemerkung 28.3 (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

$$U_\delta(x) = (x_1 - \delta, x_1 + \delta) \times \ldots \times (x_n - \delta, x_n + \delta)$$

das Innere des Quaders

$$\overline{U_\delta(x)} = [x_1 - \delta, x_1 + \delta] \times \ldots \times [x_n - \delta, x_n + \delta].$$

Quader dieser Form nennt man auch *Würfel*, da sie dadurch gekennzeichnet sind, daß die Längen aller beteiligten Intervalle gleich sind.

B) Eigenschaften von Nullmengen

Proposition 28.4 (Nullmengen)

a. Teilmengen von Nullmengen sind Nullmengen.

b. Jede endliche oder abzählbare Vereinigung von Nullmengen ist eine Nullmenge.

c. $N \subsetneqq \mathbb{R}^n$ 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

$$N \subsetneqq \bigcup_{m=0}^{\infty} \mathring{Q}_m$$

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.

d. Ist $N \subseteq \mathbb{R}^n$ eine kompakte Nullmenge, so besitzt N zu jedem $\varepsilon > 0$ eine *endliche* Überdeckung durch das Innere von Quadern $Q_1, \ldots, Q_k$ mit $\sum_{i=1}^{k} V(Q_i) < \varepsilon$.

e. 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 $N \subsetneqq \bigcup_{m=0}^{\infty} Q_m$ 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

$$N_k \subsetneqq \bigcup_{i=0}^{\infty} Q_{ki}$$

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

$$\bigcup_{k=0}^{\infty} N_k \subsetneqq \bigcup_{k=0}^{\infty}\bigcup_{i=0}^{\infty} Q_{ki} = \bigcup_{m=0}^{\infty} Q'_m.$$

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

$$\sum_{m=0}^{\infty} V(Q'_m) = \lim_{p\to\infty} \sum_{m=0}^{p} V(Q'_m) \leq \frac{\varepsilon}{2} < \varepsilon.$$

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 $P = [a, b] \subsetneqq \mathbb{R}^n$ auch einen Quader $Q \subsetneqq \mathbb{R}^n$ mit $P \subsetneqq \mathring{Q}$ und $V(Q) < 2 \cdot V(P)$ gibt. Für $t \geq 0$ definieren wir $Q_t = [a_1 - t, b_1 + t] \times \ldots \times [a_n - t, b_n + t]$ 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

$$P = [a, b] \subseteq (a_1 - t, b_1 + t) \times \ldots \times (a_n - t, b_n + t) = \mathring{Q}$$

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

$$N \subsetneqq \bigcup_{m=0}^{\infty} P_m$$

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 $P_m \subsetneqq \mathring{Q}_m$ und $V(Q_m) < 2 \cdot V(P_m)$ wählen. Für diese Quader gilt dann

$$N \subsetneqq \bigcup_{m=0}^{\infty} P_m \subsetneqq \bigcup_{m=0}^{\infty} \mathring{Q}_m$$

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,

$$N \subsetneqq \bigcup_{m=0}^{\infty} \mathring{Q}_m,$$

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 $W_1, \ldots, W_k$ überdecken kann, so daß

$$V(W_1) + \ldots + V(W_k) \leq 2^n \cdot V(Q). \tag{136}$$

Dazu können wir ohne Einschränkung annehmen, daß $a = 0$ der Ursprung ist. Für jedes $i = 1, \ldots, n$ 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 $i = 1, \ldots, n$ annehmen. Den Quader

$$P = \left[0, \tfrac{p_1}{q}\right] \times \ldots \times \left[0, \tfrac{p_n}{q}\right]$$

kann man offenbar in $k = p_1 \cdot \ldots \cdot p_n$ Würfel $W_1, \ldots, W_k$ der Seitenlänge $\frac{1}{q}$ zerlegen, und damit gilt

$$Q \subseteq P = \bigcup_{i=1}^{k} W_i \subseteq [0, 2b_1] \times \ldots \times [0, 2b_n] =: Q'$$

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 $W_{m,1}, \ldots, W_{m,k_m}$, die (136) 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 28.5 (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 $Q = [a, b] \subsetneqq \mathbb{R}^n$ ein Quader, so sind die Seiten von Q in den Hyperebenen $V(x_j - a_j)$ bzw. $V(x_j - b_j)$ für $j \in \{1, \ldots, n\}$ enthalten, so daß sie wegen Proposition 28.4 a. und Beispiel 28.2 Nullmengen sind. Der Rand des Quaders ist dann als Vereinigung der endlich vielen Seiten nach Proposition 28.4 b. ebenfalls eine Nullmenge. Als abgeschlossene Teilmengen der kompakten Menge Q sind die Seiten und auch der Rand des Quaders zudem kompakt (siehe Korollar 22.32).

Beispiel 28.6 ($\mathbb{Q}^n$ ist eine Nullmenge.)
$\mathbb{Q}^n$ ist eine Nullmenge in $\mathbb{R}^n$ und ebenso $[a, b] \cap \mathbb{Q}^n$ 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 $\mathbb{Q}$ auch $\mathbb{Q}^n$ abzählbar unendlich ist. Also ist $\mathbb{Q}^n$ als abzählbare Vereinigung von Nullmengen eine Nullmenge nach Proposition 28.4, und dann ist auch $[a, b] \cap \mathbb{Q}^n$ als Teilmenge einer Nullmenge eine Nullmenge.

C) Bilder von Nullmengen

Lemma 28.7 (Lipschitz stetige Bilder von Nullmengen sind Nullmengen.)
Ist $f : D \longrightarrow \mathbb{R}^n$ 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 28.4 gibt es eine Überdeckung von N durch Würfel $W_i = \overline{U_{\delta_i}(x_i)}$, $i \in \mathbb{N}$, mit

$$\sum_{i=0}^{\infty} 2^n \cdot \delta_i^n = \sum_{i=0}^{\infty} V(W_i) < \frac{\varepsilon}{q^n}. \tag{137}$$

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

$$f(W_i \cap D) = f\big(\overline{U_{\delta_i}(x_i)} \cap D\big) \subseteq \overline{U_{q\cdot\delta_i}(f(x_i))}.$$

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

$$\sum_{i=0}^{\infty} V\left(\overline{U_{q\cdot\delta_i}(f(x_i))}\right) = \sum_{i=0}^{\infty} (2\cdot q\cdot\delta_i)^n \overset{(137)}{<} \varepsilon$$

gilt. Mithin ist $f(N)$ eine Nullmenge. □

Beispiel 28.8 (Peano Kurve)

Die stetige Abbildung $f : [0,1] \longrightarrow [0,1] \times [0,1]$ aus Bemerkung 23.43 liefert eine stetige Abbildung

$$g : [0,1] \times [0,1] \longrightarrow [0,1] \times [0,1] : (x,y) \mapsto f(x)$$

des Quaders $Q = [0,1] \times [0,1]$ in sich selbst. Dabei wird die Nullmenge $N = [0,1] \times \{0\}$ surjektiv auf die Nicht-Nullmenge Q abgebildet. Dies zeigt, daß Stetigkeit für die Aussage in Lemma 28.7 nicht ausreicht.

Proposition 28.9 (Stetig differenzierbare Bilder von kompakten Nullmengen)

Ist $U \subseteq \mathbb{R}^n$ offen, $f : U \longrightarrow \mathbb{R}^n$ 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: In Korollar 25.20 haben wir gezeigt, daß eine stetig differenzierbare Abbildung lokal Lipschitz-stetig ist. Damit ist f lokal Lipschitz-stetig auf U und erst recht auf der Teilmenge N. Da N aber kompakt ist, können wir Proposition 23.38 anwenden und erhalten, daß f sogar global Lipschitz stetig auf N ist. Dann ist $f(N)$ aber wegen Lemma 28.7 eine Nullmenge. □

Beispiel 28.10 (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

$$f : \mathbb{R}_{>0} \times \mathbb{R} \longrightarrow \mathbb{R}^2 : (r,\theta) \mapsto \big(r\cdot\cos(\theta), r\cdot\sin(\theta)\big)$$

und die kompakte Nullmenge $\{1\} \times [0, 2\pi]$. Deren Bild unter f ist dann mit Proposition 28.9 die kompakte Nullmenge

$$f\big(\{1\} \times [0,2\pi]\big) = S^1.$$

D) Das Lebesguesche Integrabilitätskriterium

Definition 28.11 (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 28.12
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 28.13
Ist ein Quader $[a,b] \subseteq \mathbb{R}^n$ überdeckt

$$[a,b] \subseteq \bigcup_{m=1}^{k} \mathring{Q}_m$$

vom Inneren endlich vieler Quader Q_m, so können wir eine Zerlegung Z von $[a,b]$ wählen, so daß jeder Teilquader $Q \in \mathrm{TQ}(Z)$ ganz in einem der Q_m enthalten ist.

Beweis: Um dies zu sehen, setzen wir $Q_m = [a^m, b^m]$, $m = 1, \ldots, k$, und betrachten für $i \in \{1, \ldots, n\}$ die *endliche* Menge

$$\left\{ a_i, b_i, a_i^m, b_i^m \;\middle|\; m = 1, \ldots, k, a_i < a_i^m, b_i^m < b_i \right\}.$$

Nehmen wir die Punkte dieser Menge als Stützpunkte einer Zerlegung Z_i von $[a_i, b_i]$, dann ist

$$Z = Z_1 \times \ldots \times Z_n$$

eine Zerlegung von $[a,b]$. Sei nun $Q = [c,d] \in \mathrm{TQ}(Z)$ ein Teilquader von Z, so gibt es nach Voraussetzung ein m mit $Q \cap \mathring{Q}_m \neq \emptyset$. Wir wählen ein $x \in Q \cap \mathring{Q}_m$ und erhalten für $i \in \{1, \ldots, n\}$

$$c_i \leq x_i \leq d_i$$

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 28.14 (Lebesguesches Integrabilitätskriterium)
Seien $a, b \in \mathbb{R}^n$ mit $a < b$ und $f : [a, b] \longrightarrow \mathbb{R}$ 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

$$U_\delta(x) = (x_1 - \delta, x_1 + \delta) \times \ldots \times (x_n - \delta, x_n + \delta)$$

das Innere eines Quaders ist. Für die Frage der Stetigkeit von f ist die gewählte Norm aber unerheblich.

$\underline{\Longrightarrow:}$ Wir nehmen zunächst an, daß f auf $[a, b]$ integrierbar ist.

Unser Ziel ist es, zu zeigen, daß die Menge

$$N = \big\{x \in [a, b] \;\big|\; f \text{ ist nicht stetig in } x\big\}$$

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

$$N_m := \big\{x \in [a, b] \;\big|\; \forall \; \delta > 0 \;\; \exists \; y \in U_\delta(x) \cap [a, b] \;:\; |f(y) - f(x)| \geq \tfrac{1}{m}\big\}.$$

Dann gilt

$$N_1 \subseteq N_2 \subseteq N_3 \subseteq \ldots$$

und

$$N = \bigcup_{m=1}^{\infty} N_m.$$

Wegen Proposition 28.4 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 Riemannschen Integrabilitätskriteriums 27.12 eine Zerlegung Z von $[a, b]$ mit

$$\mathrm{OS}(f, Z) - \mathrm{US}(f, Z) < \frac{\varepsilon}{2m}. \tag{138}$$

Um zu zeigen, daß N_m durch endlich viele Quader mit Volumensumme kleiner als 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

$$T = \left\{ Q \in \mathrm{TQ}(Z) \;\middle|\; N_m \cap \mathring{Q} \neq \emptyset \right\}$$

aller Teilquader von Z, die einen Punkt aus N_m im Inneren enthalten. Ist $Q \in T$ und $x \in N_m \cap \mathring{Q}$, so gibt es ein $\delta > 0$ mit $U_\delta(x) \subsetneq \mathring{Q}$, da $\mathring{Q}$ offen ist, und da $x \in N_m$ ist, gibt es dann ein $y \in U_\delta(x) \subsetneq Q$ mit

$$\frac{1}{m} \leq |f(y) - f(x)| \leq \sup_{z \in Q} f(z) - \inf_{z \in Q} f(z). \tag{139}$$

Für die Summe der Volumina aller Quader in T ergibt sich dann aber

$$\begin{aligned}
\sum_{Q \in T} V(Q) \;&=\; m \cdot \sum_{Q \in T} V(Q) \cdot \frac{1}{m} \\
&\overset{(139)}{\leq}\; m \cdot \sum_{Q \in T} V(Q) \cdot \left(\sup_{z \in Q} f(z) - \inf_{z \in Q} f(z) \right) \\
&\leq\; m \cdot \sum_{Q \in \mathrm{TQ}(Z)} V(Q) \cdot \left(\sup_{z \in Q} f(z) - \inf_{z \in Q} f(z) \right) \\
&\leq\; m \cdot \big(\mathrm{OS}(f, Z) - \mathrm{US}(f, Z) \big) \overset{(138)}{<} \frac{\varepsilon}{2}.
\end{aligned}$$

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 kleiner als $\frac{\varepsilon}{2}$.

Die Vereinigung

$$M = \bigcup_{Q \in \mathrm{TQ}(Z)} \partial Q$$

der Ränder aller Teilquader in Z ist nach Beispiel 28.5 eine endliche Vereinigung von Nullmengen und ist nach Proposition 28.4 b. als solche wieder eine Nullmenge. Zudem ist sie als endliche Vereinigung abgeschlossener Mengen abgeschlossen in der kompakten Menge Q, also nach Korollar 22.32 auch kompakt. Dann gibt es nach Proposition 28.4 d. aber eine endliche Überdeckung durch Quader $P_1, \ldots, P_k$ 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

$$N_m = N_m^\circ \cup N_m^\partial \subsetneqq \bigcup_{Q\in T} Q \cup \bigcup_{i=1}^{k} P_i$$

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.

$\underline{\Longleftarrow}$: Wir wollen das Riemannsche Integrabilitätskriterium anwenden und geben uns deshalb ein $\varepsilon > 0$ vor. Wir müssen nun eine Zerlegung Z von $[a,b]$ finden, so daß

$$\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon.$$

Die Zerlegung Z wollen wir mit Lemma 28.13 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

$$|f(x)| \le c \tag{140}$$

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 28.4 c. eine abzählbare Überdeckung

$$N \subsetneqq \bigcup_{m=0}^{\infty} \mathring{Q}_m$$

von N durch das Innere von Quadern Q_m gibt mit

$$\sum_{m=0}^{\infty} V(Q_m) < \frac{\varepsilon}{2\cdot c + V([a,b])}. \tag{141}$$

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

$$|f(y) - f(z)| \le |f(y) - f(x)| + |f(x) - f(z)| < \frac{\varepsilon}{4\cdot c + 2\cdot V([a,b])}.$$

Durch Übergang zum Supremum und Infimum erhalten wir dann

$$\sup_{y\in P_x\cap[a,b]} f(y) - \inf_{z\in P_x\cap[a,b]} f(z) \le \frac{\varepsilon}{4\cdot c + 2\cdot V([a,b])} < \frac{\varepsilon}{2\cdot c + V([a,b])}. \tag{142}$$

Wir haben nun eine offene Überdeckung

$$[a,b] = N \cup \big([a,b] \setminus N\big) \subseteq \bigcup_{m=0}^{\infty} \mathring{Q}_m \cup \bigcup_{x\in[a,b]\setminus N} \mathring{P}_x$$

der kompakten Menge $[a,b]$ konstruiert, und diese muß eine endliche Teilüberdeckung besitzen, d.h. es gibt $m_1,\ldots,m_k \in \mathbb{N}$ und $x_1,\ldots,x_l \in [a,b]\setminus N$ mit

$$[a,b] \subseteq \bigcup_{i=1}^{k} \mathring{Q}_{m_i} \cup \bigcup_{j=1}^{l} \mathring{P}_{x_j}.$$

Nach Lemma 28.13 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

$$\begin{aligned}
\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) &= \sum_{Q\in\mathrm{TQ}(Z)} V(Q)\cdot\Big(\sup_{x\in Q} f(x) - \inf_{x\in Q} f(x)\Big)\\
&\le \sum_{\substack{Q\in\mathrm{TQ}(Z)\\ \exists\, i:Q\subseteq Q_{m_i}}} V(Q)\cdot\Big(\sup_{x\in Q} f(x) - \inf_{x\in Q} f(x)\Big)\\
&\quad + \sum_{\substack{Q\in\mathrm{TQ}(Z)\\ \exists\, j:Q\subseteq P_{x_j}}} V(Q)\cdot\Big(\sup_{x\in Q} f(x) - \inf_{x\in Q} f(x)\Big)\\
&\overset{(140),(142)}{\le} \sum_{\substack{Q\in\mathrm{TQ}(Z)\\ \exists\, i:Q\subseteq Q_{m_i}}} V(Q)\cdot 2\cdot c + \sum_{\substack{Q\in\mathrm{TQ}(Z)\\ \exists\, j:Q\subseteq P_{x_j}}} V(Q)\cdot\frac{\varepsilon}{2\cdot c + V([a,b])}\\
&\le \sum_{m=0}^{\infty} V(Q_m)\cdot 2\cdot c + \sum_{Q\in\mathrm{TQ}(Z)} V(Q)\cdot\frac{\varepsilon}{2\cdot c + V([a,b])}\\
&\overset{(141)}{<} \frac{\varepsilon}{2\cdot c + V([a,b])}\cdot 2\cdot c + V\big([a,b]\big)\cdot\frac{\varepsilon}{2\cdot c + V([a,b])} = \varepsilon.
\end{aligned}$$

Aus dem Riemannschen Integrabilitätskriterium 27.12 folgt nun, daß f auf $[a,b]$ integrierbar ist. □

Beispiel 28.15

Es sei $Q \subseteq \mathbb{R}^n$ ein Quader und $f : \mathring{Q} \longrightarrow \mathbb{R}$ sei stetig und beschränkt. Dann ist die Funktion

$$g : Q \longrightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } x \in \mathring{Q},\\ 0, & \text{falls } x \in \partial Q, \end{cases}$$

integrierbar auf Q, da g höchstens in der Nullmenge ∂Q unstetig ist.

Bemerkung 28.16

Mit Hilfe des Lebesgueschen Integrabilitätskriteriums 28.14 können wir den Beweis, daß mit $f : [a,b] \longrightarrow \mathbb{R}$ 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 28.17 (Produkte integrierbarer Funktionen)
Es seien $a,b \in \mathbb{R}^n$ mit $a < b$ und $f,g : [a,b] \longrightarrow \mathbb{R}$ 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 $[a,b]$. □

Bemerkung 28.18
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.

Bemerkung 28.19 (Beschränkt, nicht-integrierbar mit Stammfunktion)
In Aufgabe 20.27 haben wir ein Beispiel für eine differenzierbare Funktion gesehen, deren Ableitung nicht Riemann-integrierbar ist. Dies scheiterte daran, dass die Ableitung nicht beschränkt war. Es gibt auch Beispiele für differenzierbare Funktionen mit beschränkter Ableitung, deren Ableitung nicht Riemann-integrierbar ist. Ein Beispiel dafür ist die *Volterra-Funktion*. Die Menge der Unstetigkeitsstellen ihrer Ableitung ist die *Smith-Volterra-Cantor-Menge*, die keine Nullmenge ist.

Aufgaben

Aufgabe 28.20 (Integrale fast überall gleicher Funktionen)
Es seien $a,b \in \mathbb{R}^n$ mit $a < b$ und $f,g : [a,b] \longrightarrow \mathbb{R}$ seien beschränkt und 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 28.21
Zeige, daß jede Hyperebene eine Nullmenge ist.

Aufgabe 28.22
Sei $Q \subseteq \mathbb{R}^n$ ein Quader. Wir nennen eine beschränkte Funktion $f : Q \longrightarrow \mathbb{R}$ eine Treppenfunktion, wenn es eine Zerlegung Z von Q gibt, so dass f im Inneren jedes Teilquaders von Z konstant ist, und wir nennen die Zerlegung Z dann zu f passend.

a. Zeige, jede Treppenfunktion $f : Q \longrightarrow \mathbb{R}$ ist Riemann-integrierbar mit

$$\int_Q f(x)\,dx = \sum_{P \in \mathrm{TQ}(Z)} c_P \cdot V(P),$$

wenn Z eine zu f passende Zerlegung ist und c_P der konstante Wert ist, den f auf dem Quader P annimmt.

b. Zeige, ist $(f_m)_{m\in\mathbb{N}}$ eine Folge von Treppenfunktionen auf Q, die auf Q gleichmäßig gegen eine Funktion $f : Q \longrightarrow \mathbb{R}$ konvergiert, dann ist auch f Riemann-integrierbar und es gilt

$$\int_Q f(x)\,dx = \lim_{m\to\infty} \int_Q f_m(x)\,dx.$$

Aufgabe 28.23 (Gleichmäßiger Grenzwert fast überall stetiger Funktionen)
Sei $D \subseteq \mathbb{R}^n$ und $(f_m)_{m\in\mathbb{N}}$ eine Folge von Funktionen $f_m : D \longrightarrow \mathbb{R}$, die fast überall stetig sind. Ferner sei f der gleichmäßige Grenzwert der Folge. Zeige, dann ist auch f fast überall stetig auf D.

§ 29 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) Charakteristische Funktionen

Definition 29.1 (Charakteristische Funktionen)
Es sei $B \subseteq \mathbb{R}^n$ und $f : B \longrightarrow \mathbb{R}$. Wir nennen die Funktion

$$1_B : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto \begin{cases} 1, & \text{falls } x \in B, \\ 0, & \text{falls } x \notin B, \end{cases}$$

die *charakteristische Funktion von* B, und wir nennen die Funktion

$$f_B : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } x \in B, \\ 0, & \text{falls } x \notin B, \end{cases}$$

die *charakteristische Funktion von* f auf B.

Lemma 29.2 (Integration charakteristischer Funktionen)
Es sei $B \subseteq \mathbb{R}^n$, $f : B \longrightarrow \mathbb{R}$ 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 28.14 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 $l(Z^m) \longrightarrow 0$, und für jeden Teilquader $Y \in \mathrm{TQ}(Z^m)$ wählen wir einen Punkt $\alpha_Y^m \in \mathring{Y}$ im Inneren von Y. 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

$$\mathrm{ZS}(f_B, Z^m, \alpha^m) = \sum_{Y \in \mathrm{TQ}(Z^m)} V(Y) \cdot f(\alpha_Y^m) = \sum_{Y \in \mathrm{TQ}(X^m)} V(Y) \cdot f(\beta_Y^m) = \mathrm{ZS}(f_B, X^m, \beta^m),$$

da $f(\alpha_Y^m) = 0$ für alle $Y \in \mathrm{TQ}(Z^m) \setminus \mathrm{TQ}(X^m)$. Da die linke Seite nach dem Folgenkriterium für Integrabilität 27.17 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

$$\int_P f_B(x)\,dx = \int_X f_B(x)\,dx = \int_Q f_B(x)\,dx.$$

□

B) Jordan-meßbare Mengen

Die Aussage dieses Lemmas erlaubt uns die folgende Definition des Integralbegriffs über beschränkten Mengen, die keine Quader sind.

Definition 29.3 (Riemann-integrierbar und Jordan-meßbar)
Es sei $B \subseteq \mathbb{R}^n$ eine beschränkte Menge.

a. Eine Funktion $f : B \longrightarrow \mathbb{R}$ 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.

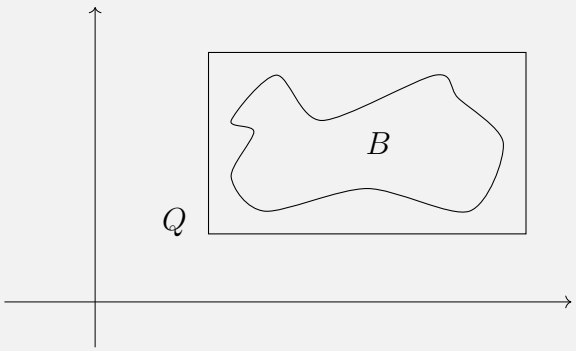


Abbildung 17: Ein Quader Q, der B enthält.

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 29.4 (Riemann-integrierbar und Jordan-meßbar)
Es sei $B \subseteq \mathbb{R}^n$ beschränkt und $f : B \longrightarrow \mathbb{R}$.

a. Wegen Lemma 29.2 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.

b. Ist B ein Quader, so können wir $Q = B$ wählen und erhalten genau die Definition der Integrierbarkeit und des Integrals aus Definition 27.9.

c. Jeder Quader B ist wegen Beispiel 27.11 Jordan-meßbar und die Definition des Volumens in Definition 27.1 stimmt mit der in Definition 29.3 überein.

Beispiel 29.5 (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 29.6 (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 28.14 genau dann der Fall wenn die Menge ∂B ihrer Unstetigkeitsstellen eine Nullmenge ist. □

Beispiel 29.7 (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 28.10 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 29.8 (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 18). Die Aussage folgt dann aus Satz 29.6. □

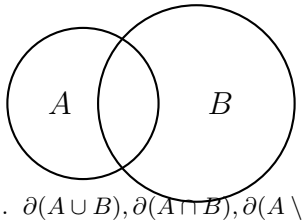


Abbildung 18. $\partial(A \cup B), \partial(A \cap B), \partial(A \setminus B) \subseteq \partial A \cup \partial B$

C) Lebesguesches Integrabilitätskriterium

Wir haben in Definition 29.3 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 29.9 (Lebesguesches Integrabilitätskriterium)
Es sei $B \subset \mathbb{R}^n$ Jordan-meßbar und $f : B \longrightarrow \mathbb{R}$ 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 28.14 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 29.23 eine Nullmenge. Dann ist aber f_B auf Q stetig außerhalb der Nullmenge $N \cup \partial B$. Mithin ist f_B mit Satz 28.14 integrierbar auf Q und somit ist f integrierbar auf B. □

Korollar 29.10 (Stetige Funktionen auf Kompakta sind integrierbar.)
Ist B kompakt und Jordan-meßbar und $f : B \longrightarrow \mathbb{R}$ 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 29.9. □

D) Folgerungen aus dem Lebesgueschen Integrabilitätskriterium

Korollar 29.11 (Rechenregeln für Integrale)
Es sei $B \subset \mathbb{R}^n$ Jordan-meßbar und $f, g : B \longrightarrow \mathbb{R}$ 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 Integrabilitätskriterium 29.9 integrierbar auf B. Aus Korollar 27.19 folgt dann

$$\begin{aligned}\int_B (c \cdot f + d \cdot g)(x)\,dx &= \int_Q (c \cdot f + d \cdot g)_B(x)\,dx \\ &= \int_Q (c \cdot f_B + d \cdot g_B)(x)\,dx \\ &\overset{27.19}{=} c \cdot \int_Q f_B(x)\,dx + d \cdot \int_Q g_B(x)\,dx \\ &= c \cdot \int_B f(x)\,dx + d \cdot \int_B g(x)\,dx.\end{aligned}$$

b. Mit f ist auch $|f|$ beschränkt und fast überall stetig auf B und deshalb integrierbar nach dem Lebesgueschen Integrabilitätskriterium 29.9. Aus Proposition 27.20 folgt dann

$$\begin{aligned}\left|\int_B f(x)\,dx\right| &= \left|\int_Q f_B(x)\,dx\right| \\ &\overset{27.20}{\leq} \int_Q |f_B(x)|\,dx = \int_Q |f|_B(x)\,dx \\ &= \int_B |f(x)|\,dx.\end{aligned}$$

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 27.19 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 Integrabilitätskriterium 29.9.

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 Integrabilitätskriterium 29.9.

□

Korollar 29.12 (Mittelwertsatz der Integralrechnung)
Sei $B \subset \mathbb{R}^n$ Jordan-meßbar und $f : B \longrightarrow \mathbb{R}$ 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 29.11

$$m\cdot V(B) = m\cdot \int_B 1\,dx = \int_B m\,dx \le \int_B f(x)\,dx \le \int_B M\,dx = M\cdot \int_B 1\,dx = M\cdot V(B).$$

□

Korollar 29.13 (Integration auf Jordan-meßbaren Teilmengen)
Sind $A \subseteq B \subset \mathbb{R}^n$ Jordan-meßbar und ist $f : B \longrightarrow \mathbb{R}$ 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 Integrabilitätskriterium für Jordan-meßbare Mengen 29.9. □

E) Die Additivität des Integrals und des Volumens

Satz 29.14 (Additivität des Integrals)
Seien $A, B \subset \mathbb{R}^n$ Jordan-meßbar und $f : A\cup B \longrightarrow \mathbb{R}$ 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 29.8 sind die Mengen $A\cup B$ und $A\cap B$ Jordan-meßbar. Zudem ist die Funktion f wegen des Lebesgueschen Integrabilitätskriteriums 29.9 beschränkt und fast überall stetig auf A und auf B, also auch auf $A\cup B$ und auf $A\cap B$. Wiederum mit Satz 29.9 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

$$A = (A \setminus B) \dot\cup (A \cap B) \tag{143}$$

und

$$B = (B \setminus A) \dot\cup (A \cap B) \tag{144}$$

sowie

$$A \cup B = (A \setminus B) \dot\cup (B \setminus A) \dot\cup (A \cap B). \tag{145}$$

Wenden wir den Spezialfall nun mehrfach an, so erhalten wir in Kurzform

$$\int_A f \overset{(143)}{=} \int_{A\setminus B} f + \int_{A\cap B} f \tag{146}$$

und

$$\int_B f \overset{(144)}{=} \int_{B\setminus A} f + \int_{A\cap B} f \tag{147}$$

sowie

$$\int_{A\cup B} f \overset{(145)}{=} \int_{(A\setminus B)\,\dot\cup\,(B\setminus A)} f + \int_{A\cap B} f \overset{(145)}{=} \int_{A\setminus B} f + \int_{B\setminus A} f + \int_{A\cap B} f$$
$$\overset{(146),(147)}{=} \int_A f - \int_{A\cap B} f + \int_B f - \int_{A\cap B} f + \int_{A\cap B} f = \int_A f + \int_B f - \int_{A\cap B} f.$$

□

Korollar 29.15 (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 29.14 auf die Funktion $f : A \cup B \longrightarrow \mathbb{R} : x \mapsto 1$ an und erhält

$$V(A \cup B) = V(A) + V(B) - V(A \cap B).$$

Ist nun $A \subseteq B$, so ist $B = A \dot\cup (B \setminus A)$ und wir erhalten mithin

$$V(B) = V(A) + V(B \setminus A) - V(\emptyset) = V(A) + V(B \setminus A) \geq V(A).$$

□

F) Jordan-Nullmengen

Definition 29.16 (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 29.17 (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

$$\varepsilon > \mathrm{OS}(1_B, Z) - \int_Q 1_B(x)\,dx = \sum_{P \in \mathrm{TQ}(Z)} V(P) \cdot \sup_{x \in P} 1_B(x) = \sum_{\substack{P \in \mathrm{TQ}(Z) \\ P \cap B \neq \emptyset}} V(P).$$

Damit haben wir die gesuchte Überdeckung gefunden, nämlich

$$B \subseteq \bigcup_{\substack{P \in \mathrm{TQ}(Z) \\ P \cap B \neq \emptyset}} P.$$

Gebe es nun umgekehrt zu jedem $\varepsilon > 0$ eine endliche Überdeckung von B durch Quader $Q_1^\varepsilon, \ldots, Q_{m_\varepsilon}^\varepsilon$ 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 $Q_1^\varepsilon, \ldots, Q_{m_\varepsilon}^\varepsilon$ abgeschlossen ist, enthält sie mit B auch den Abschluß von B und damit den Rand von B,

$$\partial B \subseteq \overline{B} \subseteq \bigcup_{i=1}^{m_\varepsilon} Q_i^\varepsilon.$$

Da $\varepsilon > 0$ beliebig war, muß mithin ∂B eine Nullmenge sein. Nach Satz 29.6 ist B dann auch Jordan-meßbar. Zudem gilt wegen der Additivität des Volumens, Korollar 29.15,

für das Volumen von B

$$V(B) \overset{29.15}{\leq} V\left(\bigcup_{i=1}^{m_\varepsilon} Q_i^\varepsilon\right) \overset{29.15}{\leq} \sum_{i=1}^{m_\varepsilon} V(Q_i^\varepsilon) < \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, muß das Volumen $V(B)$ notwendigerweise null sein. □

Bemerkung 29.18 (Überdeckungseigenschaft von Jordan-Nullmengen)
Wie in Proposition 28.4 sieht man, daß man zur Überdeckung einer Jordan-Nullmenge genauso gut auch offene Quader oder Würfel nehmen kann.

Beispiel 29.19 (Eine Nullmenge muß keine Jordan-Nullmenge sein.)
Wegen Beispiel 28.6 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 29.20 (Nullmengen und Jordan-Nullmengen)

a. Jede Jordan-Nullmenge in $\mathbb{R}^n$ ist eine Nullmenge in $\mathbb{R}^n$.

b. Jede kompakte Nullmenge in $\mathbb{R}^n$ ist eine Jordan-Nullmenge in $\mathbb{R}^n$.

c. Jede Jordan-meßbare Nullmenge in $\mathbb{R}^n$ ist eine Jordan-Nullmenge in $\mathbb{R}^n$.

Beweis: Aus Lemma 29.17 folgt unmittelbar, daß jede Jordan-Nullmenge auch eine Nullmenge ist, und aus Proposition 28.4 d. zusammen mit Lemma 29.17 folgt, daß kompakte Nullmengen auch Jordan-Nullmengen sind. Schließlich folgt aus Aufgabe 28.20, 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. □

Beispiel 29.21 (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 29.22 (Die Kreislinie ist eine Jordan-Nullmenge.)
Die Kreislinie $S^1 = V(x_1^2 + x_2^2 - 1)$ ist nach Beispiel 28.10 eine Nullmenge und sie ist kompakt. Mithin ist S^1 eine Jordan-Nullmenge.

Satz 29.23 (Kriterium für Jordan-Meßbarkeit)
Für eine beschränkte Menge $B \subseteq \mathbb{R}^n$ sind die folgenden Aussagen gleichwertig:

a. B ist Jordan-meßbar.

b. ∂B ist eine Nullmenge.

c. ∂B ist eine Jordan-Nullmenge.

Beweis: Die Äquivalenz von a. und b. wurde in Satz 29.6 gezeigt, und die Äquivalenz von b. und c. folgt aus Korollar 29.20, da die abgeschlossene und beschränkte Menge ∂B nach dem Satz von Heine-Borel 22.44 kompakt ist. □

G) Integrale und Jordan-Nullmengen

Proposition 29.24 (Integrale über Jordan-Nullmengen sind null.)
Ist B eine Jordan-Nullmenge und $f : B \to \mathbb{R}$ 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 29.9 ist die beschränkte Funktion f also integrierbar auf B. Aus dem Mittelwertsatz der Integraltheorie 29.12 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 29.25
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 29.26 (Integrale Jordan-fast überall gleicher Funktionen)
Sei $B \subset \mathbb{R}^n$ Jordan-meßbar, $f : B \longrightarrow \mathbb{R}$ sei integrierbar auf B und $g : B \longrightarrow \mathbb{R}$ 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 29.8 Jordan-meßbar. Aus der Additivität des Integrals 29.14 und Proposition 29.24 folgt dann daß $g - f$ auf $B = (B \setminus N) \cup N$ integrierbar ist mit

$$\int_B (g-f)(x)\,dx = \int_{B\setminus N} (g-f)(x)\,dx + \int_N (g-f)(x)\,dx \overset{29.24}{=} \int_{B\setminus N} 0\,dx + 0 = 0.$$

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 29.27
Die Aussage in Korollar 29.26 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 sich nur auf der Nullmenge $[0,1] \cap \mathbb{Q}$ (siehe Beispiel 28.6) von der integrierbaren konstanten Funktion

$$f : [0,1] \longrightarrow \mathbb{R} : x \mapsto 1$$

unterscheidet.

Beispiel 29.28
Ist $Q \subset \mathbb{R}^n$ ein Quader und $f : Q \longrightarrow \mathbb{R}$ ist integrierbar auf Q, dann ist auch

$$g : Q \longrightarrow \mathbb{R} : x \mapsto \begin{cases} f(x), & \text{falls } x \in \mathring{Q}, \\ 0, & \text{falls } x \in \partial Q, \end{cases}$$

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 29.29 (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 29.14 und Proposition 29.24. □

Beispiel 29.30
Ist Z eine Zerlegung des Quaders Q und $f : Q \longrightarrow \mathbb{R}$ ist integrierbar, so gilt

$$\int_Q f(x)\,dx = \sum_{P\in \mathrm{TQ}(Z)} \int_P f(x)\,dx.$$

Dies geht, da die Schnitte der Quader in $\mathrm{TQ}(Z)$ als kompakte Nullmengen stets Jordan-Nullmengen sind.

H) Das Prinzip von Cavalieri

Lemma 29.31 (Graphen integrierbarer Funk. sind Jordan-0-Mengen.)
Sei $B \subseteq \mathbb{R}^n$ Jordan-meßbar und $f : B \longrightarrow \mathbb{R}$ integrierbar, so ist der Graph

$$\mathrm{Graph}(f) = \{(x, f(x))^t \mid x \in B\}$$

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

$$\varepsilon > \mathrm{OS}(f, Z) - \mathrm{US}(f, Z) = \sum_{P\in \mathrm{TQ}(Z)} V(P) \cdot \left(\sup_{x\in P} f(x) - \inf_{x\in P} f(x) \right).$$

Definieren wir nun

$$Q_P := P \times \left[\inf_{x\in P} f(x), \sup_{x\in P} f(x) \right]$$

für $P \in \mathrm{TQ}(Z)$, so überdecken die Quader Q_P mit $P \in \mathrm{TQ}(Z)$ den Graphen von f_B auf Q und damit erst recht den Graphen von f auf B,

$$\mathrm{Graph}(f) \subseteq \mathrm{Graph}(f_B) \subseteq \bigcup_{P\in \mathrm{TQ}(P)} Q_P.$$

Zudem ist ihr Gesamtvolumen

$$\sum_{P\in \mathrm{TQ}(Z)} V(Q_P) = \mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon.$$

Nach Lemma 29.17 ist der Graph von f dann eine Jordan-Nullmenge. □

Beispiel 29.32 (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 29.33 (Das Prinzip von Cavalieri)
Sei $B \subset \mathbb{R}^n$ Jordan-meßbar, $f, g : B \to \mathbb{R}$ integrierbar auf B mit $f(x) \leq g(x)$ $\forall\, x \in B$. Dann ist die Menge

$$B_{f,g} := \big\{(x,y)^t \in \mathbb{R}^{n+1} \;\big|\; x \in B, f(x) \leq y \leq g(x)\big\}$$

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 29.23, 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 $x \in \mathring{B}$, $f(x) < y < g(x)$ 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

$$U_\gamma\big((x,y)^t\big) = \{(z,u)^t \mid ||z-x||_\infty < \gamma, |u-y| < \gamma\} \subseteq B_{f,g}$$

ganz in $B_{f,g}$ enthalten, da $z \in B$ wegen $||z-x||_\infty < \gamma \leq \delta_\varepsilon$ und

$$u \in (y-\gamma, y+\gamma) \subseteq (y-\varepsilon, y+\varepsilon) \subseteq \big(f(z), g(z)\big).$$

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

$$\partial B_{f,g} \subseteq \operatorname{Graph}(f) \cup \operatorname{Graph}(g) \cup X$$

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 $\operatorname{Graph}(f)$ und $\operatorname{Graph}(g)$ nach Lemma 29.31 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 29.23), und da f und g integrierbar sind, ist auch die Menge ihrer Unstetigkeitsstellen eine Nullmenge (siehe Satz 28.14). 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 $x \in Q$ fest, ist die stückweise stetige Funktion

$$[-M, M] \longrightarrow \mathbb{R} : y \mapsto 1_{B_{f,g}}(x,y)$$

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 27.22 erfüllt und wir erhalten

$$\begin{aligned} V(B_{f,g}) &= \int_P 1_{B_{f,g}}(x,y)\, d(x,y) = \int_Q \int_{-M}^{M} 1_{B_{f,g}}(x,y)\, dy\, dx \\ &= \int_Q \int_{f_B(x)}^{g_B(x)} 1\, dy\, dx = \int_Q g_B(x) - f_B(x)\, dx = \int_B g(x) - f(x)\, dx. \end{aligned}$$

□

Beispiel 29.34 (Der Flächeninhalt des Kreises)
Wir haben bereits in Beispiel 29.7 gesehen, daß die Kreisscheibe

$$D = \{x \in \mathbb{R}^2 \mid ||x||_2 \leq 1\}$$

Jordan-meßbar ist. Dies folgt auch aus Satz 29.33, 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.

I) Der Satz von Fubini für Normalbereiche

Definition 29.35 (Normalbereich)
Sei $i_1, \ldots, i_n$ eine Permutation der Zahlen von $1, \ldots, n$.
Ein *Normalbereich* bezüglich $(x_{i_1}, \ldots, x_{i_n})$ in $\mathbb{R}^n$ ist eine Menge der Form

$$B = \left\{(x_1, \ldots, x_n)^t \;\middle|\; \varphi_j(x_{i_1}, \ldots, x_{i_{j-1}}) \leq x_{i_j} \leq \psi_j(x_{i_1}, \ldots, x_{i_{j-1}})\ \forall\, j = 1, \ldots, n\right\},$$

wobei $\varphi_1, \psi_1 \in \mathbb{R}$ Konstanten sind, φ_j und ψ_j für $j = 2, \ldots, n$ geeignete stetige reellwertige Funktionen auf Kompakta sind und zudem

$$\varphi_j \leq \psi_j$$

für alle $j = 1, \ldots, n$ gilt. Wir bezeichnen den Normalbereich B auch mit

$$B = NB(\varphi_1, \psi_1, \ldots, \varphi_n, \psi_n).$$

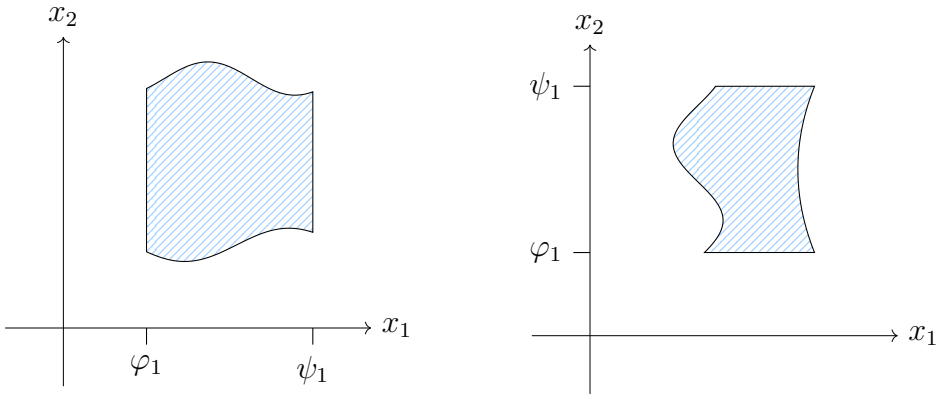


Abbildung 19: Normalbereiche bez. (x_1, x_2) und bez. (x_2, x_1)

Proposition 29.36 (Normalbereiche)
Normalbereiche sind kompakt und Jordan-meßbar.

Beweis: Ohne Einschränkung können wir $i_j = j$ für $j = 1, \ldots, n$ annehmen. Mit der Notation in Satz 29.33 und Definition 29.35 gilt für $k \in \{2, \ldots, n\}$ und

$$B = NB(\varphi_1, \psi_1, \ldots, \varphi_{k-1}, \psi_{k-1})$$

dann

$$NB(\varphi_1, \psi_1, \ldots, \varphi_k, \psi_k) = B_{\varphi_k, \psi_k}.$$

Wenden wir Satz 29.33 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 22.44 sind Normalbereiche also auch kompakt. □

Satz 29.37 (Satz von Fubini für Normalbereiche)
Ist B ein Normalbereich wie in Definition 29.35 und ist $f : B \longrightarrow \mathbb{R}$ stetig, so gilt

$$\int_B f(x)\, dx = \int_{\varphi_1}^{\psi_1} \int_{\varphi_2(x_{i_1})}^{\psi_2(x_{i_1})} \ldots \int_{\varphi_n(x_{i_1}, \ldots, x_{i_{n-1}})}^{\psi_n(x_{i_1}, \ldots, x_{i_{n-1}})} f(x)\, dx_{i_n} \ \ldots\ dx_{i_2}\, dx_{i_1}.$$

Beweis: Nach Proposition 29.36 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 $j = 1, \ldots, n$, 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

$$B' = \{(x_2, \ldots, x_n) \in \mathbb{R}^{n-1} \mid \varphi_j(x_1, \ldots, x_{j-1}) \leq x_j \leq \psi_j(x_1, \ldots, x_{j-1}), j = 2, \ldots, n\}$$

ein Normalbereich in $\mathbb{R}^{n-1}$ und somit Jordan-meßbar. Zudem ist

$$Q' = [a_2, b_2] \times \ldots \times [a_n, b_n] \subset \mathbb{R}^{n-1}$$

ein Quader in $\mathbb{R}^{n-1}$, der B' ganz enthält. Die Funktion

$$B' \longrightarrow \mathbb{R} : x' \mapsto f(x_1, x')$$

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

$$\int_{Q'} f_B(x_1, x')\, dx' = \int_{\varphi_2(x_1)}^{\psi_2(x_1)} \ldots \int_{\varphi_n(x_1,\ldots,x_{n-1})}^{\psi_n(x_1,\ldots,x_{n-1})} f(x)\, dx_n \ \ldots\ dx_2. \tag{148}$$

Außerdem ist die Funktion f_B integrierbar auf Q, so daß die Voraussetzungen des Satzes von Fubini für Quader 27.22 erfüllt sind. Wir erhalten deshalb

$$\begin{aligned}\int_B f(x)\, dx \ &= \ \int_Q f_B(x)\, dx \ = \ \int_{\varphi_1}^{\psi_1} \int_{Q'} f_B(x_1, x')\, dx'\, dx_1 \\ &\overset{(148)}{=} \int_{\varphi_1}^{\psi_1} \int_{\varphi_2(x_1)}^{\psi_2(x_1)} \ldots \int_{\varphi_n(x_1,\ldots,x_{n-1})}^{\psi_n(x_1,\ldots,x_{n-1})} f(x)\, dx_n \ \ldots\ dx_2\, dx_1,\end{aligned}$$

womit die Behauptung gezeigt ist. □

Beispiel 29.38

Die Menge

$$B = \left\{(x, y)^t \;\middle|\; -1 \le x \le 1, x^2 \le y \le 2 - x^2\right\}$$

ist der in Abbildung 20 skizzierte Normalbereich. Wir können nun mit Hilfe des Satzes

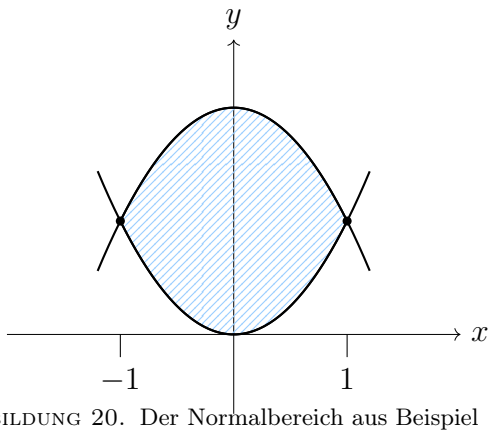


Abbildung 20. Der Normalbereich aus Beispiel 29.38

von Fubini 29.37 die stetige Funktion

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y)^t \mapsto y$$

über B integrieren:

$$\begin{aligned}\int_B y\, d(x,y) &= \int_{-1}^{1}\int_{x^2}^{2-x^2} y\, dy\, dx = \int_{-1}^{1} \tfrac{y^2}{2}\Big|_{x^2}^{2-x^2} dx \\ &= \int_{-1}^{1} 2-2x^2\, dx = 2x - \frac{2x^3}{3}\Big|_{-1}^{1} = 4 - \frac{4}{3} = \frac{8}{3}.\end{aligned}$$

Beispiel 29.39

Die Menge

$$B = \left\{(x,y,z)^t \;\middle|\; 0 \le x \le 1, 0 \le y \le 2, 0 \le z \le 2 - xy\right\}$$

ist ebenfalls ein Normalbereich (siehe Abbildung 21), und sein Volumen berechnet sich mit dem Satz von Fubini als

$$\begin{aligned}V(B) &= \int_B 1\, d(x,y,z) = \int_0^1\int_0^2\int_0^{2-x\cdot y} 1\, dz\, dy\, dx \\ &= \int_0^1\int_0^2 (2 - x\cdot y)\, dy\, dx = \int_0^2\int_0^1 (2-x\cdot y)\, dx\, dy \\ &= \int_0^2 2\cdot x - \frac{x^2\cdot y}{2}\Big|_0^1 dy = \int_0^2 2 - \frac{y}{2}\, dy = 3.\end{aligned}$$

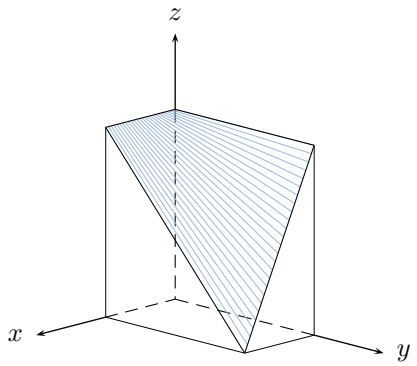


Abbildung 21: Der Normalbereich aus Beispiel 29.39

Aufgaben

Aufgabe 29.40 (Abschlüsse von Jordan-Nullmengen sind Jordan-Nullmengen)

Ist $B \subset \mathbb{R}^n$ eine Jordan-Nullmengen, dann ist auch der Abschluß $\overline{B}$ eine Jordan-Nullmenge.

Aufgabe 29.41

Zeige, ist W ein Würfel in $\mathbb{R}^n$ und $f : W \longrightarrow \mathbb{R}^n$ Lipschitz-stetig mit Lipschitz-Konstante q bezüglich der Maximumsnorm, dann gilt

$$V\big(f(W)\big) \leq q^n \cdot V(W).$$

Aufgabe 29.42 (Die Cantor-Menge)

Wir definieren rekursiv eine Folge von Mengen durch $U_1 := \left(\frac{1}{3}, \frac{2}{3}\right)$ und

$$U_n := \left\{ \frac{x}{3}, \frac{x+2}{3} \;\middle|\; x \in U_{n-1} \right\}$$

für $n \geq 2$, und nennen

$$C := [0,1] \setminus \bigcup_{n \geq 1} U_n$$

dann die Cantor-Menge.[4] Zeige, daß C eine überabzählbare, kompakte Jordan-Nullmenge ist.

Aufgabe 29.43

Sei $B \subseteq \mathbb{R}^n$ Jordan-messbar und $(f_m)_{m \in \mathbb{N}}$ mit $f_m : B \longrightarrow \mathbb{R}$ sei eine Folge integrierbarer Funktionen, die auf B gleichmäßig gegen eine Funktion f konvergiere. Zeige, dann ist f integrierbar auf B und es gilt

$$\lim_{m \to \infty} \int_B f_m(x)\, dx = \int_B f(x)\, dx.$$

Aufgabe 29.44 (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 $f : B \longrightarrow \mathbb{R}$ stetig, so gilt auch

$$\int_B f(x)\, dx = \int_{a_1}^{b_1} \int_{H_t} f(t,y)\, dy\, dt.$$

[4] Stellen wir uns die Konstruktion von C als Prozess vor, bei dem wir sukkzessive die Mengen U_n entfernen, so wird im ersten Schritt mit U_1 das mittlere Drittel des Intervalls $[0,1]$ entfernt; im zweiten Schritt wird mit U_2 dann das mittlere Drittel der verbleibenden beiden Intervalle entfernt; und so fahren wir dann fort.

Aufgabe 29.45
Sei $j \in \{1, \ldots, n\}$, $c \in \mathbb{R}$ und $M \subseteq V(x_j - c) \subseteq \mathbb{R}^n$ beschränkt. Zeige, daß M eine Jordan-Nullmenge in $\mathbb{R}^n$ ist.

Aufgabe 29.46
Zeige, daß jede beschränkte Teilmenge des $\mathbb{R}^n$, die höchstens endlich viele Häufungspunkte besitzt, eine Jordan-Nullmenge ist.

Aufgabe 29.47
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 29.48
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 29.49
Sei $B \subseteq \mathbb{R}^2$ der Normalbereich, der von den beiden Funktionen

$$\varphi : [-1, 1] \longrightarrow \mathbb{R} : x \mapsto \sqrt{1 + |x|} + \sqrt{1 - x^2}$$

und

$$\psi : [-1, 1] \longrightarrow \mathbb{R} : x \mapsto \sqrt{1 + |x|} - \sqrt{1 - x^2}$$

eingeschlossen wird. Berechne den Flächeninhalt von B und fertige eine Skizze von B an.

Aufgabe 29.50
Berechne das Volumen des Tetraeders, der von den drei Koordinatenachsen und der Ebene $x_3 = 2 - 2x_1 - x_2$ begrenzt wird.

Aufgabe 29.51
Berechne das Integral $\int_{[0,1]\times[0,1]} \max\{x, y\} \, d(x, y)$.

Aufgabe 29.52

a. Beweise durch Induktion nach n die Gleichung

$$\int_0^1 x^n(1-x)^p\,dx = \frac{n!p!}{(n+p+1)!} \text{ für alle } n,p \in \mathbb{N}.$$

b. Sei $B = \{(x_1,x_2)^t \in \mathbb{R}^2 \mid x_1 \geq 0, x_2 \geq 0, x_1+x_2 \leq 1\}$. Zeige: für alle $n,m \in \mathbb{N}$ ist

$$\int_B x_1^n x_2^m\,dx = \frac{n!m!}{(n+m+2)!}.$$

Aufgabe 29.53 (Inneres und äußeres Jordan-Maß)

Wir bezeichnen mit

$$\mathcal{P} := \left\{ \bigcup_{i=1}^k Q_i \;\middle|\; k \in \mathbb{N}, Q_1,\ldots,Q_k \text{ Quader in } \mathbb{R}^n \text{ mit } \mathring{Q}_i \cap \mathring{Q}_j = \emptyset \;\forall\, i \neq j \right\}$$

Menge der endlichen Vereinigungen von Quaderen, die sich nur im Rand schneiden.[5] Für eine beschränkte Teilmenge $B \subset \mathbb{R}^n$ nennen wir

$$V_*(B) := \sup\{V(P) \mid P \in \mathcal{P} \text{ mit } P \subseteq B\}$$

das *innere Jordan-Maß* von B und

$$V^*(B) := \inf\{V(P) \mid P \in \mathcal{P} \text{ mit } B \subseteq P\}$$

das *äußere Jordan-Maß* von B.

a. Zeige, ist $Q \subset \mathbb{R}^n$ ein Quader mit $B \subseteq Q$ und ist $1_B : Q \longrightarrow \mathbb{R}$ die charakteristische Funktion von B auf Q, dann gilt

$$\mathrm{UI}(1_B) \leq V_*(B) \leq V^*(B) \leq \mathrm{OI}(1_B).$$

b. Zeige, B ist genau dann Jordan-meßbar, wenn $V_*(B) = V^*(B)$ gilt. In diesem Fall gilt dann zudem

$$V(B) = V_*(B) = V^*(B).$$

Aufgabe 29.54

Zeige, ist $P \in \mathcal{P}$, dann gilt:

$$\forall\, \varepsilon > 0\; \exists\, Q \in \mathcal{P} \;:\; P \subseteq \mathring{Q} \text{ und } V(Q) - V(P) = V(Q \setminus P) < \varepsilon.$$

[5] Beachte, für $k = 0$ ist die Vereinigung die leere Menge, die somit in $\mathcal{P}$ enthalten ist.

Aufgabe 29.55 (σ-Additvität des Jordan-Maßes)

a. Zeige, ist $P \in \mathcal{P}$ und $(P_n)_{n\in\mathbb{N}}$ ein Folge in $\mathcal{P}$ mit $P \subseteq \bigcup_{n\in\mathbb{N}} P_n$, dann gilt

$$V(P) \leq \sum_{n=0}^{\infty} V(P_n).$$

b. Zeige, ist B Jordan-meßbar und $(B_n)_{n\in\mathbb{N}}$ eine Familie Jordan-meßbarer Mengen mit $B = \bigcup_{n=0}^{\infty} B_n$, dann gilt

$$V(B) \leq \sum_{n=0}^{\infty} V(B_n).$$

Sind die B_n paarweise disjunkt, gilt sogar die Gleichheit.

Aufgabe 29.56 (Smith-Cantor-Volterra-Menge)

Wir definieren die Smith-Cantor-Volterra-Menge rekursiv wie folgt: wir starten mit dem Intervall $[0,1]$ und nehmen in der Mitte ein offenes Intervall der Länge $\frac{1}{4}$ heraus. In jedem weiteren Schritt entfernen wir aus allen verbleibenden Intervallen jeweils in der Mitte ein offenes Intervall von einem viertel der Länge des betreffenden Intervalls. Bezeichnen wir nun mit C_n die Vereinigung der 2^{n-1} Intervalle der Länge $\frac{1}{4^n}$, die im n-ten Schritt entfernt werden, dann nennen wir

$$C := [0,1] \setminus \bigcup_{n=1}^{\infty} C_n = \bigcap_{n=1}^{\infty} [0,1] \setminus C_n$$

die *Smith-Cantor-Volterra-Menge*.

a. Zeige, C ist kompakt und hat keine inneren Punkte, d.h. $C = \partial C$.

b. Zeige, C ist nicht Jordan-meßbar und keine Nullmenge.[6]

[6] In der Tat ist C ein Beispiel für eine Lebesgue-meßbare Menge vom Lebesgue-Maß $\frac{1}{2}$ ohne inneren Punkt. Man kann die Konstruktion statt mit dem Wert $\frac{1}{4}$ auch mit jedem anderen Wert a zwischen 0 und $\frac{1}{3}$ machen. Nur im Fall $a = \frac{1}{3}$ kommt eine Jordan-meßbare Menge vom Volumen 0 heraus. In allen anderen Fälle ist die Menge nur Lebesgue-meßbar mit Lebesgue-Maß $\frac{1-3a}{1-2a}$.

§ 30 Der Transformationssatz für Integrale

A) Diffeomorphismen

Bemerkung 30.1

Ist $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ 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

$$\varphi\big([a,b]\big) = [\varphi(a), \varphi(b)]$$

und in letzterem Fall

$$\varphi\big([a,b]\big) = [\varphi(b), \varphi(a)].$$

Zudem wird der Rand $\partial[a,b] = \{a,b\}$ auf den Rand des Bildes abgebildet,

$$\varphi\big(\partial[a,b]\big) = \varphi\big(\{a,b\}\big) = \{\varphi(a), \varphi(b)\},$$

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 26.24). Die folgende Proposition ist also die naheliegende Verallgemeinerung des eben beschriebenen Sachverhalts.

Proposition 30.2 (Bilder Jordan-meßbarer Mengen unter Diffeomorph.)
Es sei $U \subseteq \mathbb{R}^n$ offen, $\varphi : U \longrightarrow \mathbb{R}^n$ sei ein Diffeomorphismus und $B \subseteq U$ sei Jordan-meßbar und kompakt. Dann ist $\varphi(B)$ Jordan-meßbar mit $\partial\varphi(B) = \varphi\big(\partial B\big)$.

Wir beweisen zunächst eine etwas schwächere Aussage unter einer ebenfalls abgeschwächten und technischeren Voraussetzung.

Lemma 30.3 (Bilder Jordan-meßbarer Mengen unter Diffeomorph.)
Es sei $U \subseteq \mathbb{R}^n$ offen, N sei eine abgeschlossene Nullmenge in $\mathbb{R}^n$, $\varphi : U \longrightarrow \mathbb{R}^n$ 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 $\partial\varphi(B) \subseteq \varphi\big(\partial B\big) \cup \varphi(N \cap B)$.

Beweis: Wir halten zunächst fest, daß φ stetig auf der kompakten Menge B ist, so daß $\varphi(B)$ nach Satz 23.23 wieder kompakt ist. Nach Korollar 22.32 ist die in B abgeschlossene Teilmenge ∂B ebenfalls kompakt. Zudem ist ∂B als Rand der Jordan-meßbaren Menge B nach Satz 29.23 eine nun kompakte Nullmenge, und da φ stetig differenzierbar auf U ist, ist dann nach Proposition 28.9 $\varphi(\partial B)$ 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 29.23

$$\partial\varphi(B) \subseteq \varphi\big(\partial B\big) \cup \varphi(N \cap B) \tag{149}$$

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 26.22 auch $\varphi\big(\mathring{B} \setminus N\big)$ offen in $\mathbb{R}^n$ ist. Dann liegt diese Menge aber ganz im Inneren von $\varphi(B)$ und es gilt

$$\varphi\big(\mathring{B} \setminus N\big) \cap \partial\varphi(B) = \emptyset. \tag{150}$$

Wir beachten nun noch, daß die kompakte Menge $\varphi(B)$ abgeschlossen ist und mithin ihren Rand enthält. Ist also $y \in \partial\varphi(B)$, so gibt es ein $x \in B$ mit $\varphi(x) = y$. Wegen (150) gilt dabei $x \notin \mathring{B} \setminus N$, und somit

$$x \in \partial B \cup (N \cap B)$$

und

$$y \in \varphi(\partial B) \cup \varphi(N \cap B).$$

Damit ist (149) gezeigt, und wir sind fertig. □

Beweis von Proposition 30.2: Ist φ ein Diffeomorphismus auf U, so sind die Voraussetzungen von Lemma 30.3 mit $N = \emptyset$ erfüllt und $\varphi(B)$ ist kompakt und Jordan-meßbar mit

$$\partial\varphi(B) \subseteq \varphi\big(\partial B\big). \tag{151}$$

Es bleibt also nur, die umgekehrte Inklusion zu zeigen.

Aus dem Satz der offenen Abbildung 26.22 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

$$\det\big(D\varphi^{-1}(y)\big) = \det\Big(\big(D\varphi\big(\varphi^{-1}(y)\big)\big)^{-1}\Big) = \frac{1}{\det\big(D\varphi\big(\varphi^{-1}(y)\big)\big)} \neq 0$$

für alle $y \in \varphi(U)$. Wenden wir (151) also auf φ^{-1} und die kompakte Jordan-meßbare Menge $\varphi(B)$ an, so erhalten wir

$$\partial\varphi^{-1}\big(\varphi(B)\big) \subseteq \varphi^{-1}\big(\partial\varphi(B)\big).$$

Kombinieren wir dies mit (151), so sehen wir

$$\partial B = \partial\varphi^{-1}\big(\varphi(B)\big) \subseteq \varphi^{-1}\big(\partial\varphi(B)\big) \subseteq \varphi^{-1}\big(\varphi\big(\partial B\big)\big) = \partial B$$

und mithin

$$\partial B = \varphi^{-1}\big(\partial\varphi(B)\big).$$

Wenden wir auf beiden Seiten φ an, so sind wir fertig. □

Beispiel 30.4 (Parallelotope als Bilder eines Würfels)

Sei $A \in \mathrm{Gl}_n(\mathbb{R})$ 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

$$W = [0,1] \times \ldots \times [0,1],$$

so ist W eine kompakte, Jordan-meßbare Menge, und somit ist auch das Parallelotop

$$P(a_1, \ldots, a_n) = \big\{\lambda_1 a_1 + \ldots + \lambda_n a_n \in \mathbb{R}^n \mid 0 \leq \lambda_i \leq 1, i = 1, \ldots, n\big\} = f_A(W)$$

Jordan-meßbar, wobei $a_1, \ldots, a_n$ die Spalten von A sind.

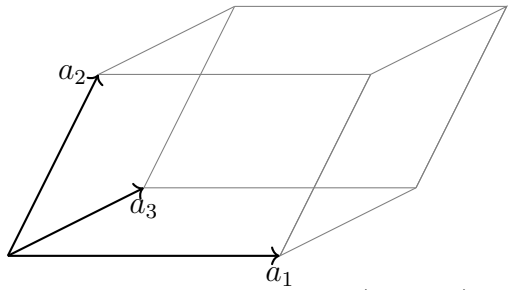


Abbildung 22. Das Parallelotop $P(a_1, a_2, a_3)$ im $\mathbb{R}^3$

Bemerkung 30.5

Die Aussage von Proposition 30.2 zum Rand ist im allgemeinen nicht mehr richtig, wenn f nicht injektiv ist, also nur ein lokaler Diffeomorphismus ist. Die Polarkoordinatenabbildung

$$\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (r, \theta) \mapsto \big(r \cdot \cos(\theta), r \cdot \sin(\theta)\big)$$

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

$$\varphi(B)=\left\{(x,y)^t\in\mathbb{R}^2\;\middle|\;r_1^2\le x^2+y^2\le r_2^2\right\}$$

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.

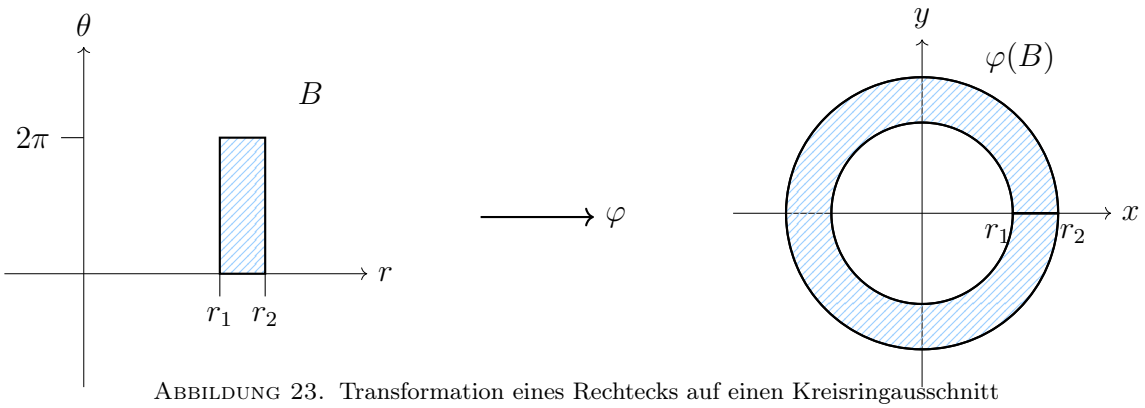


Abbildung 23. Transformation eines Rechtecks auf einen Kreisringausschnitt

Setzen wir nun

$$N=\{0\}\times\mathbb{R},$$

so ist N eine abgeschlossene Nullmenge in $\mathbb{R}^2$ und φ ist auf $\mathbb{R}^2\setminus N$ ein lokaler Diffeomorphismus, so daß wir Lemma 30.3 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,

$$\varphi(B)=\{(x,y)^t\in\mathbb{R}^2\;|\;x^2+y^2\le r_0^2\}.$$

Man beachte, daß hier auch die Jordan-Nullmenge $\{0\}\times[0,2\pi]$ im Inneren des Bildes verschwunden ist.

Situationen dieser Art sind für die Anwendungen im Zusammenhang mit dem Transformationssatz aber wichtig. Man sollte sie als Grenzprozesse der regulären Situation in Proposition 30.2 verstehen.

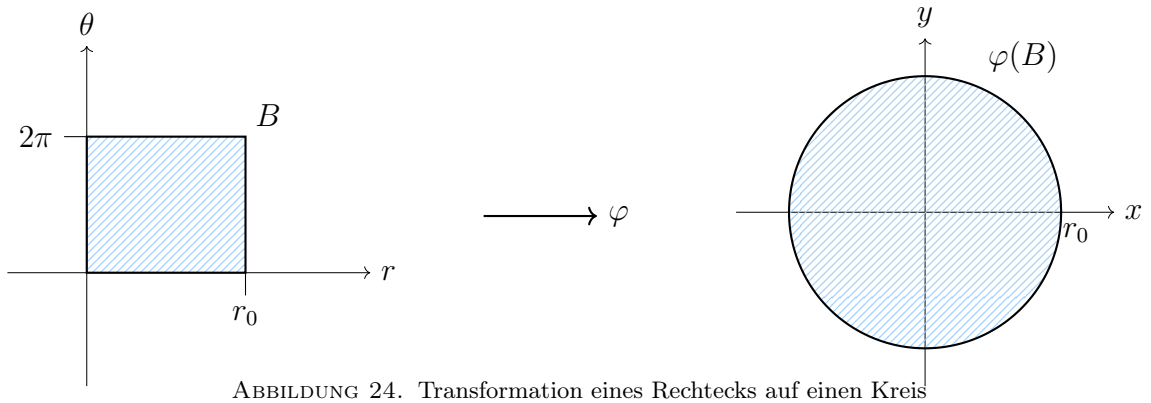


ABBILDUNG 24. Transformation eines Rechtecks auf einen Kreis

B) Der Transformationssatz für Integrale

Bemerkung 30.6 (Die Substitutionsregel)

Im Eindimensionalen haben wir die Substitutionsregel für Integrale kennengelernt. Ist $\varphi : [a,b] \longrightarrow \mathbb{R}$ stetig differenzierbar und $f : \mathbb{R} \longrightarrow \mathbb{R}$ 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

$$\varphi\big([a,b]\big) = \big[\varphi(a), \varphi(b)\big] \quad \text{oder} \quad \varphi\big([a,b]\big) = \big[\varphi(b), \varphi(a)\big]$$

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

$$\varphi\big([a,b]\big) = \big[\varphi(a), \varphi(b)\big]$$

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

$$\varphi\big([a,b]\big) = \big[\varphi(b), \varphi(a)\big]$$

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 30.1 und Proposition 30.2 legt dies nahe, die Formel

$$\int_{\varphi(B)} f(x)\,dx = \int_B f\big(\varphi(x)\big)\cdot|\det D\varphi(x)|\,dx$$

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 30.7 (Transformationssatz für Integrale)
Sei $U \subseteq \mathbb{R}^n$ offen, $\varphi : U \longrightarrow \mathbb{R}^n$ ein Diffeomorphismus mit $\det(D\varphi(x))$ stets positiv oder stets negativ, $B \subseteq U$ kompakt und Jordan-meßbar und $f : \varphi(B) \longrightarrow \mathbb{R}$ stetig.
Dann ist f auf der Jordan-meßbaren Menge $\varphi(B)$ integrierbar mit

$$\int_{\varphi(B)} f(x)\,dx = \int_B f(\varphi(x))\cdot|\det D\varphi(x)|\,dx. \tag{152}$$

Bemerkung 30.8 (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 (152) gilt auch noch. Wir beweisen diesen Sachverhalt in Lemma 30.22. □

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 30.9 (Funktionaldeterminate)
Ist $U \subseteq \mathbb{R}^n$ offen, $\varphi : U \longrightarrow \mathbb{R}$ 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

$$h : [0,1] \longrightarrow \mathbb{R} : t \mapsto \det\big(D\varphi(x + t \cdot (y - x))\big)$$

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

$$\det\big(D\varphi(x + t \cdot (y - x))\big) = h(t) = 0$$

im Widerspruch dazu, daß φ ein lokaler Diffeomorphismus ist. □

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 30.10 (Transformationssatz für Integrale auf Quadern)
Sei $U \subseteq \mathbb{R}^n$ offen, $\varphi : U \longrightarrow \mathbb{R}^n$ stetig differenzierbar, $Q \subseteq U$ ein Quader, φ auf dem Inneren $\mathring{Q}$ von Q ein Diffeomorphismus und $f : \varphi(Q) \longrightarrow \mathbb{R}$ 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

$$h : Q \longrightarrow \mathbb{R} : x \mapsto f(\varphi(x)) \cdot |\det(D\varphi(x))|$$

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

$$Q_\varepsilon = [a_1 + \varepsilon, b_1 - \varepsilon] \times \ldots \times [a_n + \varepsilon, b_n - \varepsilon] \subseteq Q$$

mit Volumen

$$V(Q_\varepsilon) = \prod_{i=1}^{n}(b_i - a_i - 2\varepsilon) \xrightarrow{\varepsilon\to 0} \prod_{i=1}^{n}(b_i - a_i) = V(Q). \tag{153}$$

Damit erhalten wir aus dem Mittelwertsatz der Integration 29.12 dann

$$\inf_{x\in Q} h(x)\cdot\big(V(Q) - V(Q_\varepsilon)\big) \le \int_{Q\setminus Q_\varepsilon} h(x)\,dx \le \sup_{x\in Q} h(x)\cdot\big(V(Q) - V(Q_\varepsilon)\big),$$

wobei die linke und die rechte Seite für $\varepsilon \to 0$ gegen null konvergieren. Also gilt dies auch für das Integral in der Mitte, und aus der speziellen Additivität des Integrals 29.29 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 $\varepsilon \to 0$.

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 30.3 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 25.19 ist die stetig differenzierbare Abbildung φ auf dem kompakten konvexen Quader Q zudem Lipschitz-stetig mit einer Lipschitz-Konstanten $q > 0$. Für einen Würfel

$$W = \overline{U_r(a)}$$

gilt dann (siehe auch den Beweis von Lemma 28.7)

$$\varphi(W) \subseteq \overline{U_{qr}(\varphi(a))}$$

und damit (siehe auch Aufgabe 29.41)

$$V(\varphi(W)) \le (2qr)^n = q^n \cdot V(W). \tag{154}$$

Das Komplement $Q \setminus Q_\varepsilon$ ist offensichtlich eine Vereinigung von Quadern, und wir haben im Beweis von Proposition 28.4 e. (siehe (136)) 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 $W_1, \ldots, W_k$ mit

$$Q \setminus Q_\varepsilon \subseteq W_1 \cup \ldots \cup W_k$$

und

$$\sum_{i=1}^{k} V(W_i) \le 2^n \cdot V(Q \setminus Q_\varepsilon) = 2^n \cdot \big(V(Q) - V(Q_\varepsilon)\big).$$

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

$$V\big(\varphi(Q) \setminus \varphi(Q_\varepsilon)\big) \leq \sum_{i=1}^{k} V\big(\varphi(W_i)\big) \overset{(154)}{\leq} q^n \cdot \sum_{i=1}^{k} V(W_i) \leq 2^n q^n \cdot \big(V(Q) - V(Q_\varepsilon)\big).$$

Für $\varepsilon \to 0$ konvergiert die rechte Seite wegen (153) gegen null, so daß auch

$$\lim_{\varepsilon \to 0} V\big(\varphi(Q) \setminus \varphi(Q_\varepsilon)\big) = 0$$

gilt. Aus dem Mittelwertsatz der Integration 29.12 erhalten wir dann wieder

$$\begin{aligned} \inf_{x \in \varphi(Q)} f(x) \cdot V\big(\varphi(Q) \setminus \varphi(Q_\varepsilon)\big) &\leq \int_{\varphi(Q) \setminus \varphi(Q_\varepsilon)} f(x)\, dx \\ &\leq \sup_{x \in \varphi(Q)} f(x) \cdot V\big(\varphi(Q) \setminus \varphi(Q_\varepsilon)\big), \end{aligned}$$

und für $\varepsilon \to 0$ 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 29.29 an und erhalten für $\varepsilon \to 0$

$$\int_{\varphi(Q_\varepsilon)} f(x)\, dx = \int_{\varphi(Q)} f(x)\, dx - \int_{\varphi(Q) \setminus \varphi(Q_\varepsilon)} f(x)\, dx \longrightarrow \int_{\varphi(Q)} f(x)\, dx.$$

Nun können wir auf dem Quader Q_ε aber den Transformationssatz 30.7 anwenden, wobei wir beachten, daß die Funktionaldeterminante von φ auf der konvexen Menge $\mathring{B}$ nach Lemma 30.9 ihr Vorzeichen nicht wechselt, und erhalten somit insgesamt

$$\int_{\varphi(Q)} f(x)\, dx \longleftarrow \int_{\varphi(Q_\varepsilon)} f(x)\, dx \overset{30.7}{=} \int_{Q_\varepsilon} h(x)\, dx \longrightarrow \int_{Q} h(x)\, dx,$$

was zu beweisen war. □

Bevor wir uns dem Beweis des Transformationssatzes 30.7 zuwenden wollen, betrachten wir Beispiele, die seine Bedeutung veranschaulichen sollen.

C) Anwendungsbeispiele für den Transformationssatz

Beispiel 30.11 (Das Volumen des Parallelotops)

In Bemerkung A14.31 haben wir das Volumen eines Parallelotops im $\mathbb{R}^3$ bestimmt. Diese Rechnung können wir nun mit Hilfe der Transformationsformel und von Beispiel 30.4 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 30.4 W der Würfel mit Kantenlänge Eins ist. Diese Formel gilt in dieser Form in jeder Dimension, so daß unsere Definition des Volumens in Bemerkung A9.12 sinnvoll war.

Beispiel 30.12 (Polarkoordinaten)
Die Polarkoordinatenabbildung

$$\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (r,\theta)^t \mapsto \big(r\cdot\cos(\theta), r\cdot\sin(\theta)\big)$$

eingeschränkt auf die offene Menge

$$U = (0,\infty)\times(0,2\pi)$$

in $\mathbb{R}^2$ ist ein Diffeomorphismus mit

$$\det\big(D\varphi(r,\theta)\big) = r > 0$$

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 25 dargestellt. Für eine auf $\varphi(B)$

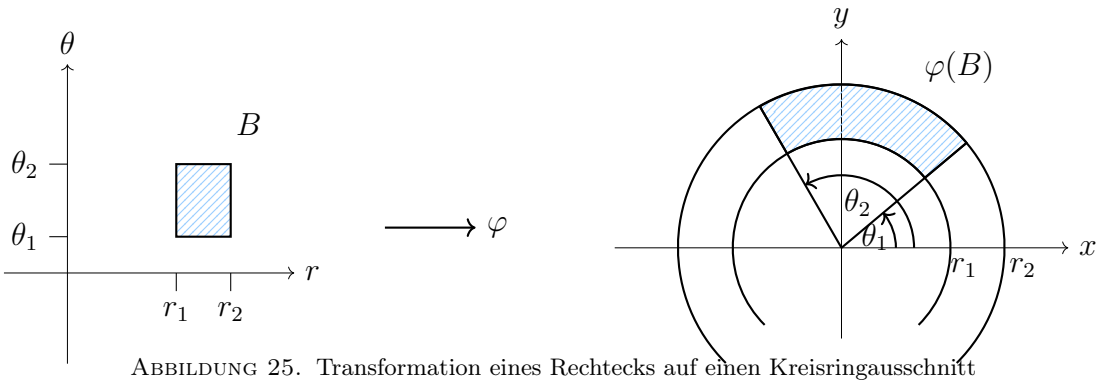


Abbildung 25. Transformation eines Rechtecks auf einen Kreisringausschnitt

stetige Funktion gilt dann

$$\begin{aligned}\int_{\varphi(B)} f(x,y)\,d(x,y) &= \int_B f\big(\varphi(r,\theta)\big)\cdot|det(D\varphi(r,\theta))|\,d(r,\theta)\\ &= \int_{\theta_1}^{\theta_2}\int_{r_1}^{r_2} f(r\cdot\cos(\theta), r\cdot\sin(\theta))\cdot r\,dr\,d\theta.\end{aligned} \tag{155}$$

Insbesondere können wir damit den Inhalt des Kreisbogenausschnitts berechnen:

$$V\big(\varphi(B)\big) = \int_{\theta_1}^{\theta_2}\int_{r_1}^{r_2} r\,dr\,d\theta = \frac{r_2^2 - r_1^2}{2}\cdot(\theta_2-\theta_1).$$

Wegen des Transformationssatzes für Integrale auf Quadern 30.10 gilt die Transformationsformel (155) 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 30.13 (Zylinderkoordinaten)
Nach Aufgabe 26.33 ist die Zylinderkoordinatenabbildung

$$\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (r,\theta,z) \mapsto \big(r \cdot \cos(\theta), r \cdot \sin(\theta), z\big)$$

auf der offenen Menge

$$U = (0,\infty) \times (0,2\pi) \times \mathbb{R}$$

ein Diffeomorphismus mit

$$\det(D\varphi(r,\theta,z)) = \begin{vmatrix} \cos(\theta) & -r \cdot \sin(\theta) & 0 \\ \sin(\theta) & r \cdot \cos(\theta) & 0 \\ 0 & 0 & 1 \end{vmatrix} = r > 0.$$

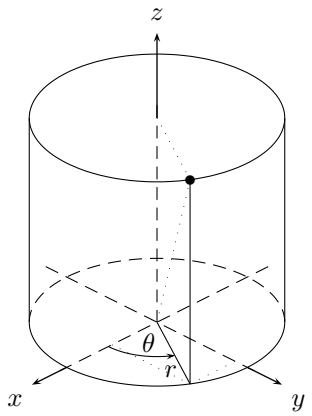


Abbildung 26: Zylinderkoordinaten

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

$$\int_{\varphi(Q)} f(x,y,z) \; d(x,y,z) = \int_{r_1}^{r_2} \int_{\theta_1}^{\theta_2} \int_{z_1}^{z_2} f\big(r \cdot \cos(\theta), r \cdot \sin(\theta), z\big) \cdot r \; dz \; d\theta \; dr.$$

Aufgrund von Korollar 30.10 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

$$Z = \{(x,y,z)^t \in \mathbb{R}^3 \mid x^2 + y^2 \leq r_0^2, 0 \leq z \leq z_0\}$$

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 30.14 (Kugelkoordinaten)
Nach Aufgabe 26.32 ist die Kugelkoordinatenabbildung

$$\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (r, \theta, \vartheta) \mapsto \big(r \cdot \cos(\theta) \cdot \cos(\vartheta), r \cdot \sin(\theta) \cdot \cos(\vartheta), r \cdot \sin(\vartheta)\big)$$

ist auf der offenen Menge

$$U = (0, \infty) \times (0, 2\pi) \times \big(-\tfrac{\pi}{2}, \tfrac{\pi}{2}\big)$$

ein Diffeomorphismus mit

$$\det(D\varphi(r, \theta, \vartheta)) = r^2 \cdot \cos(\vartheta) > 0.$$

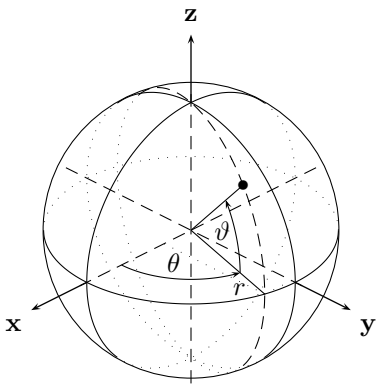


Abbildung 27: Kugelkoordinaten

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

$$\int_{\varphi(Q)} f(x, y, z) \; d(x, y, z) =$$
$$\int_{r_1}^{r_2} \int_{\theta_1}^{\theta_2} \int_{\vartheta_1}^{\vartheta_2} f\big(r\cos(\theta)\cos(\vartheta), r\sin(\theta)\cos(\vartheta), r\sin(\vartheta)\big) \cdot r^2 \cdot \cos(\vartheta) \; d\vartheta \; d\theta \; dr.$$

Wieder erlaubt uns Korollar 30.10 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

$$V(K) = \int_0^{r_0}\int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} r^2\cdot\cos(\vartheta)\,d\vartheta\,d\theta\,dr = \int_0^{r_0}\int_0^{2\pi} r^2\cdot\sin(\vartheta)\Big|_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\,d\theta\,dr$$
$$= \int_0^{r_0}\int_0^{2\pi} 2r^2\,d\theta\,dr = \int_0^{r_0} 4r^2\pi\,dr = \frac{4\cdot r_0^3\cdot\pi}{3}.$$

D) Der Beweis des Transformationssatzes 30.7

In diesem Abschnitt beschäftigen wir uns ausschließlich mit dem Beweis des Transformationssatzes für Integrale 30.7 sowie seiner Verallgemeinerung in Bemerkung 30.8. Wir zerlegen den Beweis in viele einzelne Schritte, die wir als Lemmata formulieren und beweisen. Dabei setzen wir stets die Bezeichnungen aus Satz 30.7 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 29.37 auf die Substitutionsregel im Eindimensionalen 20.15 zurückzuführen.

Wir verwenden stets die Maximumsnorm, so daß Umgebungen Würfel sind.

Lemma 30.15
$\varphi(B)$ ist kompakt und Jordan-meßbar, und die Integrale in Satz 30.7 existieren, und zwar auch noch unter den abgeschwächten Voraussetzungen von Bemerkung 30.8.

Beweis: Wir beachten zunächst, daß der Abschluß $\overline{N}$ der Jordan-Nullmenge N wegen Aufgabe 29.40 eine abgeschlossene Nullmenge ist. Aus $\det(D\varphi(x)) \neq 0$ für $x \in \mathring{B}\setminus\overline{N} \subseteq U\setminus N$ folgt dann mit dem Satz über die Umkehrfunktion 26.19, daß φ ein lokaler Diffeomorphismus auf $\mathring{B}\setminus\overline{N}$ ist. Die Voraussetzungen von Lemma 30.3 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 30.7 existieren. □

In Satz 30.7 bleibt also nur noch, die Transformationsformel (152) zu beweisen. Wir reduzieren das Problem nun darauf, die Transformationsformel für Quader zu zeigen.

Lemma 30.16
Gilt (152) für Quader, so gilt (152) 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 (152) 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 (152) 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 (152) 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 $\overline{U_{\varepsilon_i}(x_i)}$, $i = 1, \ldots, k$, besitzt mit

$$U_{4\varepsilon_i}(x_i) \subseteq U. \tag{156}$$

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 $U_{\frac{\delta_x}{4}}(x)$ die kompakte Menge B ebenfalls, und diese Überdeckung besitzt mithin eine endliche Teilüberdeckung, d.h. es reichen endlich viele Punkte $x_1, \ldots, x_k$ aus und wir können $\varepsilon_i = \frac{\delta_{x_i}}{4}$ 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 $\overline{U_{3\varepsilon_i}(x_i)}$ 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 $U_{3\varepsilon_i}(x_i)$ 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

$$l := \min\{\varepsilon_1, \ldots, \varepsilon_k\}.$$

Ist Z eine äquidistante Zerlegung des Würfels P in *Würfel* mit $l(Z) \leq l$, so gilt für jeden Teilwürfel $W \in \mathrm{TQ}(Z)$ 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 (156) 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 27.16 ein $\delta_m > 0$, so daß

$$\text{(157)}\qquad \mathrm{OS}(1_{\partial B}, Z) = \mathrm{OS}(1_{\partial B}, Z) - \int_P 1_{\partial B}(x)\,dx = \mathrm{OS}(1_{\partial B}, Z) - \mathrm{US}(1_{\partial B}, Z) < \frac{1}{m}$$

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

$$X := \bigcup_{\substack{W \in \mathrm{TQ}(Z^m) \\ W \cap \partial B \neq \emptyset}} W$$

und

$$Y := \bigcup_{\substack{W \in \mathrm{TQ}(Z^m) \\ W \subseteq \mathring{B}}} W \subseteq B.$$

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

$$\text{(158)}\qquad \varphi(B) \setminus \varphi(Y) \subseteq \varphi(B \setminus Y) \subseteq \varphi(X).$$

Aus (157) erhalten wir

$$\text{(159)}\qquad V(X) = \sum_{\substack{W \in \mathrm{TQ}(Z^m) \\ W \cap \partial B \neq \emptyset}} V(W) = \mathrm{OS}(1_{\partial B}, Z^m) < \frac{1}{m},$$

und mit

$$M := \sup_{y \in \varphi(B)} |f(y)|$$

dann auch

$$|\det(D\varphi(x))| \leq \sum_{\sigma \in \mathfrak{S}_n} |D_1\varphi_{\sigma(1)}(x)| \cdot \ldots \cdot |D_n\varphi_{\sigma(n)}(x)| \leq L^n \cdot n!$$

und

$$\text{(160)}\qquad \begin{aligned} \left| \int_{B \setminus Y} f(\varphi(x)) \cdot |\det(D\varphi(x))|\,dx \right| &\leq \int_{B \setminus Y} |f(\varphi(x))| \cdot |\det(D\varphi(x))|\,dx \\ &\leq \int_{B \setminus Y} M \cdot L^n \cdot n!\,dx = M \cdot L^n \cdot n! \cdot V(B \setminus Y) \\ &\leq M \cdot L^n \cdot n! \cdot V(X) \overset{(159)}{<} \frac{M \cdot L^n \cdot n!}{m}. \end{aligned}$$

Zudem ist die Abbildung φ nach Korollar 25.19 auf jedem Würfel $W \in \mathrm{TQ}(Z^m)$ 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 $W = \overline{U_r(a)}$ gilt dann aber $\varphi(W) \subseteq \overline{U_{Lr}(\varphi(a))}$ und somit (siehe auch Aufgabe 29.41)

$$V(\varphi(W)) \leq V\big(\overline{U_{Lr}(\varphi(a))}\big) = (2Lr)^n = L^n \cdot V(W).$$

Da die Abschätzung für jeden an X beteiligten Würfel gilt, erhalten wir auch

$$V(\varphi(X)) \leq L^n \cdot V(X) \overset{(159)}{<} \frac{L^n}{m}. \tag{161}$$

Damit erhalten wir analog zu obigen Betrachtungen

$$\begin{aligned} \left| \int_{\varphi(B)\setminus\varphi(Y)} f(x)\, dx \right| &\leq \int_{\varphi(B)\setminus\varphi(Y)} |f(x)|\, dx \\ &\leq \int_{\varphi(B)\setminus\varphi(Y)} M\, dx = M \cdot V\big(\varphi(B) \setminus \varphi(Y)\big) \\ &\overset{(158)}{\leq} M \cdot V(\varphi(X)) \overset{(159)}{<} \frac{M \cdot L^n}{m}. \end{aligned} \tag{162}$$

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 30.3). Wir können die spezielle Additivität für Integrale 29.29 also bei Y und bei $\varphi(Y)$ anwenden, und da wir zudem vorausgesetzt haben, daß die Transformationsformel auf Quadern gilt, erhalten wir folgende Gleichung:

$$\begin{aligned} \int_{\varphi(Y)} f(x)\, dx &= \sum_{\substack{W \in \mathrm{TQ}(Z^m) \\ W \subseteq \mathring{B}}} \int_{\varphi(W)} f(x)\, dx \\ &= \sum_{\substack{W \in \mathrm{TQ}(Z^m) \\ W \subseteq \mathring{B}}} \int_W f(\varphi(x)) \cdot |\det(D\varphi(x))|\, dx \\ &= \int_Y f(\varphi(x)) \cdot |\det(D\varphi(x))|\, dx. \end{aligned} \tag{163}$$

Der Übersichtlichkeit halber setzen wir nun $g(x) = f(\varphi(x)) \cdot |\det(D\varphi(x))|$ und tragen die Ergebnisse aus (160), (162) und (163) in Kurzform zusammen:

$$\begin{aligned}\left|\int_{\varphi(B)} f - \int_B g\right| &= \left|\int_{\varphi(Y)} f + \int_{\varphi(B)\setminus\varphi(Y)} f - \int_Y g - \int_{B\setminus Y} g\right| \\ &\overset{(163)}{=} \left|\int_{\varphi(B)\setminus\varphi(Y)} f - \int_{B\setminus Y} g\right| \le \left|\int_{\varphi(B)\setminus\varphi(Y)} f\right| + \left|\int_{B\setminus Y} g\right| \\ &\overset{(160),(162)}{<} \frac{M\cdot L^n + M\cdot L^n\cdot n!}{m}.\end{aligned}$$

Dabei ist die linke Seite unabhängig von m und die rechte Seite konvergiert für $m \to \infty$ 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 (152) dann für B gezeigt. □

Lemma 30.17
(152) gilt, wenn

$$\varphi(x) = \big(x_1, \ldots, x_{n-1}, \varphi_n(x)\big)^t$$

für eine notwendigerweise stetig differenzierbare Funktion $\varphi_n : U \longrightarrow \mathbb{R}$ ist.

Beweis: Wegen Lemma 30.16 können wir ohne Einschränkung annehmen, daß B ein Quader ist. Die Funktionaldeterminante von φ ist dann

$$(164)\quad \det(D\varphi(x)) = \begin{vmatrix} 1 & 0 & 0 & \ldots & 0 & 0 \\ 0 & 1 & 0 & \ldots & 0 & 0 \\ \vdots & \ddots & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & \ddots & 0 & 0 \\ 0 & \ldots & \ldots & 0 & 1 & 0 \\ D_1\varphi_n(x) & D_2\varphi_n(x) & \ldots & \ldots & D_{n-1}\varphi_n(x) & D_n\varphi_n(x) \end{vmatrix} = D_n\varphi_n(x),$$

so daß nach Voraussetzung ohne Einschränkung

$$(165)\qquad D_n\varphi_n(x) > 0$$

für alle $x \in U$ gilt. Ist nun $B = [a, b]$ und $x_i \in [a_i, b_i]$ fest für $i = 1, \ldots, n-1$, so hat die Funktion

$$[a_n, b_n] \longrightarrow \mathbb{R} : t \mapsto \varphi_n(x_1, \ldots, x_{n-1}, t)$$

eine stets positive Ableitung und ist mithin streng monoton wachsend. Dann ist aber

$$\begin{aligned}\varphi(B) = \big\{x \in \mathbb{R}^n \;\big|\; & a_i \le x_i \le b_i \text{ für } i = 1, \ldots, n-1, \\ & \varphi_n(x_1, \ldots, x_{n-1}, a_n) \le x_n \le \varphi_n(x_1, \ldots, x_{n-1}, b_n)\big\}\end{aligned}$$

ein Normalbereich, so daß uns der Satz von Fubini für Normalbereiche 29.37 und die Substitutionsregel 20.15 erlauben, das linke Integral in (152) auszurechnen:

$$\begin{aligned}\int_{\varphi(B)} f(x)\,dx &\overset{29.37}{=} \int_{a_1}^{b_1}\dots\int_{a_{n-1}}^{b_{n-1}}\int_{\varphi_n(x_1,\dots,x_{n-1},a_n)}^{\varphi_n(x_1,\dots,x_{n-1},b_n)} f(x_1,\dots,x_n)\,dx_n\,dx_{n-1}\,\dots\,dx_1\\ &\overset{20.15}{=} \int_{a_1}^{b_1}\dots\int_{a_n}^{b_n} f(x_1,\dots,x_{n-1},\varphi_n(x))\cdot D_n\varphi_n(x)\,dx_n\,dx_{n-1}\,\dots\,dx_1\\ &\overset{(164),(165)}{=} \int_B f(\varphi(x))\cdot|\det(D\varphi(x))|\,dx\end{aligned}$$

Damit ist dann (152) 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 30.7 in Dimension $n-1$ bereits bewiesen ist.

Lemma 30.18
Der Transformationssatz 30.7 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

$$\beta(x)=\big(x_1,\dots,x_{n-1},\beta_n(x)\big)^t$$

und

$$\gamma(x)=\big(\varphi_1(x),\dots,\varphi_{n-1}(x),x_n\big)^t,$$

dann gilt (152).

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

$$B=[a_1,b_1]\times\ldots\times[a_n,b_n]=Q\times[a_n,b_n]$$

mit

$$Q=[a_1,b_1]\times\ldots\times[a_{n-1},b_{n-1}]\subseteq\mathbb{R}^{n-1}.$$

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

$$\phi_{x_n}:U_{x_n}\longrightarrow\mathbb{R}^{n-1}:y\mapsto\big(\varphi_1(y,x_n),\dots,\varphi_{n-1}(y,x_n)\big)^t$$

stetig differenzierbar mit

$$\det\big(D\phi_{x_n}(y)\big)=\det(D\gamma(y,x_n))\neq0 \tag{166}$$

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 30.9 auch ihr Vorzeichen nicht.

Beachtet man noch, daß

$$\gamma(B) = \bigcup_{x_n \in [a_n, b_n]} \phi_{x_n}(Q) \times \{x_n\} \subseteq \gamma(U) \tag{167}$$

gilt, so ist die Funktion

$$g : \gamma(B) \longrightarrow \mathbb{R} : (y, x_n) \mapsto f\big(\beta(y, x_n)\big) \cdot |\det(D\beta(y, x_n))|$$

nach Voraussetzung stetig und wir können den Transformationssatz 30.7 im $\mathbb{R}^{n-1}$ für ein festes $x_n \in [a_n, b_n]$ anwenden:

$$\begin{aligned} \int_{\phi_{x_n}(Q)} g(y, x_n)\, dy &= \int_Q g\big(\phi_{x_n}(y), x_n\big) \cdot |\det(D\phi_{x_n}(y))|\, dy \\ &\overset{(166)}{=} \int_Q g\big(\gamma(y, x_n)\big) \cdot |\det(D\gamma(y, x_n))|\, dy. \end{aligned} \tag{168}$$

Nun beachten wir noch die Gleichung

$$\det(D\varphi(x)) = \det\big(D\beta(\gamma(x))\big) \cdot \det(D\gamma(x)). \tag{169}$$

und mithin

$$\begin{aligned} g(\gamma(x)) \cdot |\det(D\gamma(x))| &= f\big(\beta(\gamma(x))\big) \cdot |\det(D\beta(\gamma(x)))| \cdot |\det(D\gamma(x))| \\ &\overset{(169)}{=} f\big(\varphi(x)\big) \cdot |\det(D\varphi(x))|. \end{aligned} \tag{170}$$

Damit sind wir nun in der Lage (152) in dieser Situation zu beweisen, wobei wir beachten, daß wir den Satz von Fubini in der Fassung von Aufgabe 29.44 auf $\gamma(B)$ anwenden können:

$$\begin{aligned} \int_{\varphi(B)} f(x)\, dx &= \int_{\beta(\gamma(B))} f(x)\, dx \overset{30.17}{=} \int_{\gamma(B)} f(\beta(x)) \cdot |\det(D\beta(x))|\, dx \\ &= \int_{\gamma(B)} g(x)\, dx = \int_{\gamma(B)} g(y, x_n)\, d(y, x_n) \\ &\overset{29.44}{=} \int_{a_n}^{b_n} \int_{\phi_{x_n}(Q)} g(y, x_n)\, dy\, dx_n \\ &\overset{(168)}{=} \int_{a_n}^{b_n} \int_Q g\big(\gamma(y, x_n)\big) \cdot |\det(D\gamma(y, x_n))|\, dy\, dx_n \\ &\overset{(170)}{=} \int_{a_n}^{b_n} \int_Q f\big(\varphi(y, x_n)\big) \cdot |\det(D\varphi(y, x_n))|\, dy\, dx_n \\ &= \int_B f(\varphi(x)) \cdot |\det(D\varphi(x))|\, dx. \end{aligned}$$

□

Wir wollen nun ein einfaches Beispiel für Diffeomorphismen mit einer globalen Zerlegung wie in Lemma 30.18 geben.

Beispiel 30.19
Wir betrachten eine invertierbare Matrix

$$A = (a_{ij})_{i,j=1,\dots,n} = \left(\begin{array}{c|c} D & b \\ \hline c^t & a_{nn} \end{array}\right) \in \mathrm{Gl}_n(\mathbb{R})$$

mit $D = (a_{ij})_{i,j=1,\dots,n-1}$, $b = (a_{1n}, \dots, a_{n-1n})^t$ und $c = (a_{n1}, \dots, a_{nn-1})^t$. Außerdem setzen wir noch voraus, daß auch $D \in \mathrm{Gl}_{n-1}(\mathbb{R})$ invertierbar ist.

Dann setzen wir

$$B = \left(\begin{array}{c|c} \mathbb{1}_{n-1} & 0 \\ \hline c^t \circ D^{-1} & a_{nn} - c^t \circ D^{-1} \circ b \end{array}\right)$$

und

$$C = \left(\begin{array}{c|c} D & b \\ \hline 0 & 1 \end{array}\right).$$

Die Abbildungen $\varphi = f_A$, $\beta = f_B$ und $\gamma = f_C$ erfüllen dann die Voraussetzungen von Lemma 30.18. Insbesondere gilt

$$\varphi = f_A = f_B \circ f_C = \beta \circ \gamma,$$

da man leicht $A = B \circ C$ nachrechnet.

Lemma 30.20 (Zerlegungssatz)
Für jeden Punkt in $x \in U$ gibt es (nach eventuellem Umsortieren der Koordinaten) eine offene Umgebung V sowie Diffeomorphismen

$$\gamma : V \longrightarrow \mathbb{R}^n : x \mapsto \big(\varphi_1(x), \dots, \varphi_{n-1}(x), x_n\big)^t$$

und

$$\beta : \gamma(V) \longrightarrow \mathbb{R}^n : x \mapsto \big(x_1, \dots, x_{n-1}, \beta_n(x)\big)^t$$

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 Umnumerieren 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 26.19 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 $\alpha_1, \ldots, \alpha_n$, d.h.

$$\gamma^{-1} : \gamma(V) \longrightarrow \mathbb{R}^n : y \mapsto \big(\alpha_1(y), \ldots, \alpha_n(y)\big)^t.$$

Dann definieren wir

$$\beta_n : \gamma(V) \longrightarrow \mathbb{R} : y \mapsto \varphi_n\big(\alpha_1(y), \ldots, \alpha_{n-1}(y), y_n\big)^t$$

und erhalten damit

$$\begin{aligned}\beta\big(\gamma(z)\big) &= \big(\varphi_1(z), \ldots, \varphi_{n-1}(z), \varphi_n(\alpha_1(\gamma(z)), \ldots, \alpha_{n-1}(\gamma(z)), z_n)\big)^t \\ &= \big(\varphi_1(z), \ldots, \varphi_{n-1}(z), \varphi_n(z_1, \ldots, z_{n-1}, z_n)\big)^t = \varphi(z).\end{aligned}$$

Zudem ist $\beta = \varphi \circ \gamma^{-1}$ als Komposition zweier Diffeomorphismen offenbar ein Diffeomorphismus. □

Lemma 30.21
(152) gilt für ein beliebiges B und φ.

Beweis: Wegen Lemma 30.16 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 30.6). Für den Induktionsschritt können wir also annehmen, daß (152) in $\mathbb{R}^{n-1}$ bereits gilt.

Aus dem Zerlegungssatz 30.20 wissen wir, daß jeder Punkt $x \in B$ eine Umgebung V_x besitzt, auf der φ sich (bis auf die Reihenfolge der Variablen) wie in Lemma 30.18 zerlegen läßt, und durch Verkleinern können wir ohne Einschränkung annehmen, daß V_x ein offener Quader ist. Die V_x überdecken B und da B kompakt ist, reichen dazu schon endlich viele V_{x_i}, $i = 1, \ldots, k$, aus. Wir können nun nach Lemma 28.13 eine Zerlegung Z des Quaders B wählen, die so fein ist, daß jeder Teilquader von Z ganz in einem der V_{x_i} enthalten ist. Dann können wir Lemma 30.18 auf jeden dieser Teilquader anwenden, wobei die Rolle von U durch den zugehörigen offenen Quader V_{x_i} übernommen wird, und wir erhalten, daß (152) auf jedem Teilquader gilt. Dann folgt aber (152) auf B

mittels der speziellen Additivität des Integrals 29.29:

$$\begin{aligned}\int_{\varphi(B)} f(x)\,dx &= \sum_{Q\in \mathrm{TQ}(Z)} \int_{\varphi(Q)} f(x)\,dx \\ &= \sum_{Q\in \mathrm{TQ}(Z)} \int_{Q} f(\varphi(x))\cdot|\det(D\varphi(x))|\,dx \\ &= \int_B f(\varphi(x))\cdot|\det(D\varphi(x))|\,dx.\end{aligned}$$

□

Damit ist der Transformationssatz für Integrale 30.7 gezeigt. Es bleibt noch, die allgemeine Fassung des Satzes mit den abgeschwächten Voraussetzungen in Bemerkung 30.8 zu zeigen. Wegen Lemma 30.15 reicht es dazu, die Transformationsformel (152) für diesen Fall herzuleiten.

Lemma 30.22
(152) gilt auch unter den abgeschwächten Voraussetzungen von Bemerkung 30.8.

Beweis: Wir geben uns eine positive natürliche Zahl m vor. Die Jordan-Nullmenge $\overline{N}$ (siehe Aufgabe 29.40) läßt sich nach Bemerkung 29.18 durch das Innere endlich vieler Quader $Q_1,\ldots,Q_k$ in U überdecken mit

$$\sum_{i=1}^{k} V(Q_i) < \frac{1}{m}.$$

Wir setzen nun

$$X := \bigcup_{i=1}^{k} \mathring{Q}_i \subseteq U$$

und

$$Y := B \setminus X.$$

Dann ist Y abgeschlossen und beschränkt und nach dem Satz von Heine-Borel 22.44 deshalb kompakt. Zudem ist φ auf der offenen Umgebung $U\setminus\overline{N}$ von Y ein Diffeomorphismus, so daß die Transformationsformel nach Satz 30.7 auf Y gilt.

Zudem ist $V(X) < \frac{1}{m}$ und wie im Beweis des 4. Schritts von Lemma 30.16 zeigt man, daß

$$V(\varphi(X)) \le L^n \cdot V(X) < \frac{L^n}{m}$$

gilt für eine geeignete Konstante L, die unabhängig von m ist.

Dann kann man wie im Beweis des 4. Schrittes von Lemma 30.16 fortfahren und erhält schließlich die Transformationsformel (152) auch in diesem Fall. □

Aufgaben

Aufgabe 30.23

Sei $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ ein Diffeomorphismus, $a \in \mathbb{R}^3$ und $\overline{U_r(a)}$ bezeichne den Würfel um a mit Kantenlänge $2r$ (siehe Bemerkung 28.3). Zeige

$$\lim_{r \to 0} \frac{V\big(f\big(\overline{U_r(a)}\big)\big)}{8 \cdot r^3} = |\det Df(a)|.$$

Aufgabe 30.24

Sei $A \subseteq \mathbb{R}^3$ die obere Hälfte der Kugel mit dem Radius R um den Nullpunkt, d.h. $A = \{(x, y, z)^t \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq R^2,\ z \geq 0\}$. Berechne

$$\int_A z \, d(x, y, z).$$

Kapitel IV

Gewöhnliche Differentialgleichungen

§ 31 Gewöhnliche Differentialgleichungen

In diesem Abschnitt betrachten wir $\mathbb{R}^n$ stets als normierten Raum bezüglich der euklidischen Norm $||\cdot||_2$, sofern nichts anderes gesagt wird.

A) System gewöhnlicher Differentialgleichungen

Notation 31.1 (k-te Ableitungen)
Wir werden im folgenden m-fach stetig differenzierbare Abbildungen

$$x : I \longrightarrow \mathbb{R}^n$$

von einem offenen Intervall I in $\mathbb{R}$ in den $\mathbb{R}^n$ betrachten für ein $m \geq 1$. Der Bezeichnung x für die Funktion und t für die Variable liegt die Vorstellung zugrunde, daß die Variable die Zeit beschreibt und x die Änderung des Ortes in Abhängigkeit von der Zeit. x ist dann eine stetig differenzierbare Kurve $\mathbb{R}^n$ im Sinne von Definition B1.1.

Die Ableitung von x

$$Dx : I \longrightarrow \mathbb{R}^n : t \mapsto \left(\frac{\partial x_1}{\partial t}(t), \ldots, \frac{\partial x_n}{\partial t}(t)\right)^t$$

ist wieder eine Abbildung von I in den $\mathbb{R}^n$. Wir bezeichen sie im folgenden kurz mit

$$\dot{x} := Dx$$

und entsprechend bezeichnen wir mit

$$x^{(k)} : I \longrightarrow \mathbb{R}^n$$

die k-te Ableitung von x als Abbildung von I nach $\mathbb{R}^n$ für $k \leq m$.

Definition 31.2 (Systeme gewöhnlicher DGL m-ter Ordnung)
Sei $U \subseteq \mathbb{R}^n$ offen, $I \subseteq \mathbb{R}$ ein offenes Intervall und

$$f : I \times U \times (\mathbb{R}^n)^{m-1} \longrightarrow \mathbb{R}^n$$

eine stetige Abbildung. Eine m-fach stetig differenzierbare Abbildung

$$x : I \longrightarrow U$$

heißt *Lösung des Systems gewöhnlicher Differentialgleichungen m-ter Ordnung*

$$x^{(m)} = f(t, x, \dot{x}, \dots, x^{(m-1)}), \tag{171}$$

wenn für alle $t \in I$

$$x^{(m)}(t) = f(t, x(t), \dot{x}(t), \dots, x^{(m-1)}(t))$$

gilt. Hält man einen Punkt $t_0 \in I$ fest und gibt sich ein Tupel $(\eta_0, \dots, \eta_{m-1}) \in (\mathbb{R}^n)^m$ von m Vektoren vor und gilt

$$x^{(k)}(t_0) = \eta_k$$

für $k = 0, \dots, m-1$, so heißt x eine Lösung des Systems zum *Anfangswert* $(\eta_0, \dots, \eta_{m-1})$. Wir sprechen dann auch von einem *Anfangswertproblem* (AWP). Hängt die Funktion f nicht von der Variablen t ab, so nennt man das System ein *autonomes* System, sonst heißt es *nicht-autonom*. Manchmal sprechen wir einfach von einer Differentialgleichung statt von einem System von Differentialgleichungen, insbesondere im Fall $n = 1$.

Bemerkung 31.3 (Reduktion auf ein System erster Ordnung)
Jedes System gewöhnlicher Differentialgleichungen m-ter Ordnung wie in (171) kann auf ein System gewöhnlicher Differentialgleichungen erster Ordnung zurück geführt werden. Mit der Notation aus Definition 31.2 betrachten wir die offene Menge

$$V = U \times (\mathbb{R}^n)^{m-1} \subseteq \mathbb{R}^{nm}$$

in $\mathbb{R}^{nm}$ und die stetige Abbildung

$$F : I \times V \longrightarrow \mathbb{R}^{nm} : (t, y_0, \dots, y_{m-1})^t \mapsto \big(y_1, \dots, y_{m-1}, f(t, y_0, \dots, y_{m-1})\big)^t,$$

wobei die y_i Vektoren in $\mathbb{R}^n$ sind. Nun ist

$$z : I \longrightarrow V$$

genau dann eine Lösung des Systems gewöhnlicher Differentialgleichungen erster Ordnung

$$\dot{z} = F(t, z),$$

wenn z_0 eine Lösung des Systems gewöhnlicher Differentialgleichungen m-ter Ordnung in (171) ist und zudem $z_i = \dot{z}_{i-1}$ für $i = 1, \dots, m-1$ gilt. Die Anfangswerte übertragen sich ebenso.

Beispiel 31.4 (Klassische Mechanik)
Betrachten wir ein Kraftfeld $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ und die Bewegung eines Teilchens der Masse m in dem Kraftfeld. Dann löst die Bahnkurve

$$x : \mathbb{R} \longrightarrow \mathbb{R}^3$$

des Teilchens aufgrund des physikalischen Gesetzes "Masse x Beschleunigung = Kraft" die autonome Differentialgleichung zweiter Ordnung

$$m \cdot \ddot{x} = f(x).$$

Mit dem Ansatz in Bemerkung 31.3 überführen wir sie in ein System von Differentialgleichungen erster Ordnung

$$\dot{z} = F(z) \tag{172}$$

mit

$$F : \mathbb{R}^6 \longrightarrow \mathbb{R}^6 : (y_0, y_1)^t \mapsto \left(y_1, \tfrac{1}{m} \cdot f(y_0)\right)^t.$$

Dabei stehen die Bahnkurven x mit den Lösungen von (172) vermittels $z = (x, \dot{x})^t$ in Beziehung. Gelingt es uns (172) zu lösen, so erhalten wir durch Projektion auf die ersten drei Koordinaten also eine mögliche Bahnkurve des Teilchens.

Bemerkung 31.5 (Reduktion auf autonome Systeme)
Jedes nicht-autonome System gewöhnlicher Differentialgleichungen läßt sich zudem auf ein autonomes System zurückführen. Wegen Bemerkung 31.3 reicht es dabei, Systeme erster Ordnung zu betrachten. Sei also

$$f : I \times U \longrightarrow \mathbb{R}^n$$

stetig differenzierbar wie in Definition 31.2 für $m = 1$ gegeben. Dann betrachten wir offene Menge

$$V = I \times U \subseteq \mathbb{R}^{n+1}$$

in $\mathbb{R}^{n+1}$ und die stetige Abbildung

$$F : V \longrightarrow \mathbb{R}^{n+1} : (t, x)^t \mapsto (1, f(t, x))^t.$$

Damit ist

$$z : I \longrightarrow \mathbb{R}^{n+1} : t \mapsto (z_0(t), \ldots, z_n(t))^t$$

genau dann eine Lösung des Anfangswertproblems

$$\dot{z} = F(z)$$

mit $z(t_0) = (t_0, \eta_0)$, wenn

$$x : I \longrightarrow \mathbb{R}^n : t \mapsto (z_1(t), \ldots, z_n(t))^t$$

eine Lösung des Anfangswertproblems

$$\dot{x} = f(t, x)$$

mit $x(t_0) = \eta_0$ ist, da aus $\dot{z}_0 = 1$ und $z_0(t_0) = t_0$ unmittelbar $z_0(t) = t$ folgt.

Bemerkung 31.6 (Konsequenz aus den Reduktionen)
Als Konsequenz aus den Bemerkungen 31.3 und 31.5 können wir bei theoretischen Untersuchen etwa zur Lösbarkeit von Systemen von Differentialgleichungen stets voraussetzen, daß die Systeme autonom und von erster Ordnung sind. Ein solches System wird durch eine Abbildung

$$f : U \longrightarrow \mathbb{R}^n$$

für eine offene Menge $U \subseteq \mathbb{R}^n$ gegeben, und eine solche Abbildung nennen wir auch ein *Vektorfeld*.

B) Der Satz von Picard-Linedlöf

Wir wollen uns nun der Frage der Lösbarkeit und der Eindeutigkeit der Lösungen zuwenden. Der Satz von Picard-Lindelöf wird die Frage unter geeigneten Voraussetzungen lokal für uns lösen.

Beispiel 31.7 (Einfache Differentialgleichungen)

a. Betrachten wir den einfachen Fall

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto ax$$

für eine Konstante $a \in \mathbb{R}$ und die zugehörige Differentialgleichung der Form

$$\dot{x} = a \cdot x,$$

dann finden wir die Funktionen

$$x(t) = e^{at} \cdot c$$

mit $c \in \mathbb{R}$ beliebig als Ein-Parameter-Familie von Lösungen der Differentialgleichung. Wählt man nun einen Anfangswert $x(0) = \eta_0$, so erhält man

$$\eta_0 = x(0) = e^{a \cdot 0} \cdot c = c$$

und somit die Lösung

$$x : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto e^{at} \cdot \eta_0.$$

Wir werden sehen, daß dies die einzige Lösung des Anfangswertproblems ist.

b. Betrachten wir nun die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sqrt{|x|}$$

und das zugehörige Anfangswertproblem

$$\dot{x} = f(x)$$

mit $x(0) = 0$. Dann sind die Funktionen

$$x : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto 0$$

sowie

$$y : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto \begin{cases} \frac{t^2}{4}, & \text{für } t \geq 0, \\ -\frac{t^2}{4}, & \text{für } t < 0 \end{cases}$$

zwei verschiedene Lösungen des Anfangswertproblems.

c. Betrachten wir die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$$

und das zugehörige Anfangswertproblem

$$\dot{x} = f(x) = x^2$$

mit $x(-1) = 1$. Dann ist die Funktion

$$x : (-\infty, 0) \longrightarrow \mathbb{R} : t \mapsto -\frac{1}{t}$$

eine Lösung des Anfangswertproblems auf $I = (-\infty, 0)$. Auf ganz $\mathbb{R}$ besitzt das Anfangswertproblem hingegen keine Lösung, wie wir in Beispiel 31.16 sehen werden.

Bemerkung 31.8 (Zielvorgaben für die Lösbarkeit von AWP)
Wir suchen Bedingungen, die die Abbildung f erfüllen sollte, damit hinsichtlich der Lösbarkeit von Anfangswertproblemen zu f folgendes sichergestellt ist:

(1) Es gibt stets ein Intervall I, in dem das (AWP) eine Lösung besitzt.

(2) In diesem Intervall I ist die Lösung eindeutig.

(3) Die Lösungen hängen stetig von den Eingabedaten (f und Anfangswert) ab.

Lipschitzbedingungen an f stellen sich dabei als geeignet heraus.

Für den Beweis des lokalen Existenz- und Eindeutigkeitssatzes von Picard-Lindelöf benötigen wir komponentenweises Integrieren stetiger Kurven im $\mathbb{R}^n$ (zum Begriff der Kurve siehe auch Definition B1.1).

Definition 31.9 (Komponentenweises Integral einer Kurve)
Für eine stetige Abbildung $h : I \longrightarrow \mathbb{R}^n$ für ein Intervall I und $a, b \in I$ erinnern wir uns an die Definition des Integrals (siehe Definition 25.29)

$$\int_a^b h(s)\, ds = \left(\int_a^b h_1(s)\, ds, \ldots, \int_a^b h_1(s)\, ds \right)^t \in \mathbb{R}^n.$$

Für dieses gilt dann

$$\left|\left|\int_a^b h(s)\,ds\right|\right|_2 \leq \left|\int_a^b ||h(s)||_2\,ds\right|, \tag{173}$$

wie die folgende einfache Rechnung mit $y = \int_a^b h(s)\,ds \in \mathbb{R}^n$ für $a \leq b$ unter Anwendung der Ungleichung von Cauchy-Schwarz A14.8 zeigt

$$\begin{aligned}||y||_2^2 =& \langle y, y\rangle = \sum_{i=1}^n y_i \cdot \int_a^b h_i(s)\,ds = \int_a^b \sum_{i=1}^n y_i \cdot h_i(s)\,ds = \int_a^b \langle y, h(s)\rangle\,ds \\ \leq& \int_a^b |\langle y, h(s)\rangle|\,ds \leq \int_a^b ||y||_2 \cdot ||h(s)||_2\rangle|\,ds = ||y||_2 \cdot \int_a^b ||h(s)||_2\,ds.\end{aligned}$$

Satz 31.10 (Existenz- und Eindeutigkeitssatz von Picard-Lindelöf)
Sei $U \subseteq \mathbb{R}^n$ offen, $f : U \longrightarrow \mathbb{R}^n$ lokal Lipschitz-stetig und $\eta_0 \in U$.
Dann gibt es ein $\delta > 0$, so daß es genau eine stetig differenzierbare Lösung

$$x : (-\delta, \delta) \longrightarrow U$$

auf $(-\delta, \delta)$ des Anfangswertproblems $\dot{x} = f(x)$ mit $x(0) = \eta_0$ gibt.

Bemerkung 31.11 (Grundidee des Beweises des Satzes von Picard-Lindelöf)
Ist $x : (-\delta, \delta) \longrightarrow U$ eine stetig differenzierbare Lösung des Anfangswertproblems $\dot{x} = f(x)$ mit $x(0) = \eta_0$, so folgt aus dem Hauptsatz der Differential- und Integralrechnung angewendet auf die Komponentenfunktionen

$$x(t) - x(0) = \int_0^t \dot{x}(s)\,ds = \int_0^t f(x(s))\,ds$$

und damit

$$x(t) = \eta_0 + \int_0^t f(x(s))\,ds.$$

Wir haben die Differentialgleichung damit in eine Integralgleichung umgewandelt, die wir besser handhaben können. Betrachten wir nämlich den Operator

$$\mathcal{I} : \mathcal{C}((-\delta, \delta), U) \longrightarrow \mathcal{C}((-\delta, \delta), \mathbb{R}^n) : y \mapsto \left(t \mapsto \eta_0 + \int_0^t f(y(s))\,ds\right),$$

der einer stetigen Funktion y die durch das Integral definierte stetige Funktion zuordnet, dann suchen wir letzten Endes einfach einen *Fixpunkt* dieses Operators:

$$\mathcal{I}(x) = x.$$

Der Fixpunktsatz von Banach 26.4 liefert uns die gewünschte Exisitenz und Eindeutigkeit, wenn wir den Funktionenraum $\mathcal{C}((-\delta, \delta), U)$ so einschränken, daß der Raum vollständig und der Operator $\mathcal{I}$ eine strikte Kontraktion auf dem Raum wird. Der Beweis des Satzes von Picard-Lindelöf besteht deshalb im wesentlichen daraus, zu zeigen, daß und wie dies machbar ist.

Beweis des Satzes von Picard-Lindelöf 31.10: Da U offen ist gibt es ein $r > 0$ mit

$$M = \overline{U_r(\eta_0)} \subset U$$

und nach Proposition 23.38 ist f als lokal Lipschitz-stetige Abbildung auf dem Kompaktum M dann Lipschitz-stetig mit einer Lipschitz-Konstanten $q > 0$. Damit ist f dann aber auch stetig und somit beschränkt auf M. Es gibt also eine Konstante $C > 0$ mit

$$||f(z)||_2 \leq C$$

für alle $z \in M$. Wir setzen nun

$$\delta := \min\left\{\tfrac{1}{q}, \tfrac{r}{C}\right\}$$

und wählen ein $0 < \varepsilon < \delta$ beliebig. Wie in Bemerkung 23.42 sieht man, daß der Vektorraum

$$V = \mathcal{C}([-\varepsilon, \varepsilon], \mathbb{R}^n)$$

bezüglich der Maximumsnorm

$$||x||_\infty := \max_{|t| \leq \varepsilon} ||x(t)||_2$$

ein Banachraum ist. Die Teilmenge

$$A = \{x \in V \mid x(0) = \eta_0, \mathrm{Im}(x) \subset M\}$$

von V ist abgeschlossen, da Konvergenz bezüglich der Maximumsnorm gleichmäßige Konvergenz ist und der gleichmäßige Grenzwert stetiger Abbildungen wieder stetig ist (siehe Satz 23.41), womit die beiden Bedingungen in A von der Grenzfunktion ebenfalls erfüllt werden müssen. Als abgeschlossene Teilmenge eines vollständigen Raums ist A dann aber nach Aufgabe 22.57 ebenfalls vollständig bezüglich der Maximumsnorm.

Wir wollen nun zeigen, daß der Integraloperator

$$\mathcal{I} : A \longrightarrow A : x \mapsto \left(t \mapsto \eta_0 + \int_0^t f(x(s))\, ds\right),$$

wirklich eine Selbstabbildung ist. Dabei beachten wir für $x \in A$, daß $\mathcal{I}(x)$ wegen des Hauptsatzes der Differential- und Integralrechnung stetig differenzierbar und damit insbesondere stetig ist und daß

$$\mathcal{I}(x)(0) = \eta_0 + \int_0^0 f(x(s))\, ds = \eta_0$$

gilt. Es bleibt also zu zeigen, daß $\mathrm{Im}(\mathcal{I}(x))$ in $M = \overline{U_r(\eta_0)}$ enthalten ist, daß also der Abstand von $\mathcal{I}(x)(t)$ von η_0 stets durch r beschränkt ist:

$$\begin{aligned} ||\mathcal{I}(x)(t) - \eta_0||_2 &= \left|\left|\int_0^t f(x(s))\, ds\right|\right|_2 \overset{(173)}{\leq} \left|\int_0^t ||f(x(s))||_2\, ds\right| \\ &\leq \left|\int_0^t C\, ds\right| = |t| \cdot C \leq \varepsilon \cdot C < \delta \cdot C \leq r. \end{aligned}$$

Als nächstes wollen wir zeigen, daß $\mathcal{I}$ ein strikte Kontraktion ist. Dazu geben wir uns zwei Abbildungen $x, y \in A$ vor und erhalten

$$\begin{aligned}||\mathcal{I}(x) - \mathcal{I}(y)||_\infty &= \max_{|t|\leq\varepsilon} \left|\left| \int_0^t f(x(s)) - f(y(s))\, ds \right|\right|_2 \\ &\overset{(173)}{\leq} \max_{|t|\leq\varepsilon} \left| \int_0^t ||f(x(s)) - f(y(s))||_2\, ds \right| \leq \max_{|t|\leq\varepsilon} \left| \int_0^t q \cdot ||x(s) - y(s)||_2\, ds \right| \\ &\leq \max_{|t|\leq\varepsilon} \left| \int_0^t q \cdot ||x - y||_\infty\, ds \right| = \max_{|t|\leq\varepsilon} |t| \cdot q \cdot ||x - y||_\infty \leq \varepsilon \cdot q \cdot ||x - y||_\infty\end{aligned}$$

und aufgrund der Wahl von ε gilt

$$\varepsilon \cdot q < \delta \cdot q \leq 1.$$

Mithin ist $\mathcal{I}$ eine strikte Kontraktion auf dem vollständigen Raum A, so daß der Fixpunktsatz von Banach 26.4 uns einen eindeutigen Fixpunkt $x_\varepsilon \in A$ von $\mathcal{I}$ liefert. Aus

$$x_\varepsilon(t) = \mathcal{I}(x_\varepsilon)(t) = \eta_0 + \int_0^t f(x_\varepsilon(s))\, ds$$

folgt mittels des Hauptsatzes der Differential- und Integralrechnung dann, daß x_ε nicht nur stetig, sondern stetig differenzierbar ist und daß

$$\dot{x_\varepsilon}(t) = f(x_\varepsilon(t))$$

sowie $x_\varepsilon(0) = \eta_0$ gilt. x_ε ist also eine Lösung des Anfangswertproblems auf $[-\varepsilon, \varepsilon]$. Man beachte nun, daß für $\varepsilon < \nu < \delta$ die Einschränkung von x_ν auf $[-\varepsilon, \varepsilon]$ ebenfalls eine Lösung des Anfangswertproblems auf $[-\varepsilon, \varepsilon]$ und damit ein Fixpunkt von $\mathcal{I}$ ist. Wegen der Eindeutigkeit des Fixpunktes muß x_ν auf $[-\varepsilon, \varepsilon]$ also mit x_ε übereinstimmen. Wir können die Lösung des Anfangswertproblems deshalb auf eindeutige Weise auf das Intervall

$$(-\delta, \delta) = \bigcup_{0\leq\varepsilon<\delta} [-\varepsilon, \varepsilon]$$

fortsetzen durch

$$x : (-\delta, \delta) \longrightarrow \mathbb{R}^n : t \mapsto \eta_0 + \int_0^t f(x_\varepsilon(s))\, ds.$$

□

Bemerkung 31.12 (Satz von Picard-Lindelöf)

a. Der Satz von Picard-Lindelöf liefert die Existenz einer Lösung des Anfangswertproblems im allgemeinen nur lokal in η_0, d.h. nur für eine kurze Zeit $t \in (-\delta, \delta)$. Wenn wir die Konstanten r, C und q aus dem Beweis des Satzes von Picard-Lindelöf kennen, dann können wir eine untere Schranke für die Länge δ des Existenzintervalls angeben:

$$\delta = \min\left\{\tfrac{1}{q}, \tfrac{r}{C}\right\}.$$

Das reicht aus, um die Punkte (1) und (2) in Bemerkung 31.8 zu erfüllen.

b. Der Fixpunktsatz von Banach liefert uns ein Rekursionsverfahren zur Approximation der eindeutigen Lösung x, indem wir $x_0 = \eta_0$ setzen und dann

$$x_k(t) = \mathcal{I}(x_{k-1})(t) = \eta_0 + \int_0^t f(x_{k-1}(s))\, ds$$

für $k \geq 1$. Die Folge $(x_k)_{k\in\mathbb{N}}$ konvergiert dann bezüglich der Maximumsnorm, also gleichmäßig, gegen x, und wir könnten mit Hilfe der a priori Abschätzung und der a posteriori Abschätzung im Fixpunktsatz von Banach sogar Abschätzungen für die Qualität der Konvergenz angeben.

Beispiel 31.13

Die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto \sqrt{|x|}$$

aus Beispiel 31.7 Teil b. ist in 0 nicht lokal Lipschitz-stetig, was man u.a. daran sehen kann, daß das dort beschriebene Anfangswertproblem mehr als eine Lösung besitzt. Es folgt aber auch daraus, daß die Steigung des Graphen im Punkt $(0, 0)$ unbeschränkt ist.

C) Existenz und Eindeutigkeit im Großen

Wir wenden uns nun der Frage zu, inwieweit die eindeutigen Lösungen sich fortsetzen lassen und ob dabei die Eindeutigkeit erhalen bleibt.

Definition 31.14 (Maximale Lösung)

Sei $f : U \longrightarrow \mathbb{R}^n$ stetig und seien $y : I \longrightarrow U$ und $z : J \longrightarrow U$ zwei Lösungen des Anfangswertproblems $\dot{x} = f(x)$ und $x(0) = \eta_0$ mit $I \subseteq J$ und $z_{|I} = y$, dann heißt z eine *Fortsetzung* von y.

Besitzt y keine echte Fortsetzung, so heißt y eine *maximale Lösung* des (AWP) und I heißt das zugehörige *maximale Existenzintervall*.

Korollar 31.15 (Existenz- und Eindeutigkeitssatz im Großen)

Sei $U \subseteq \mathbb{R}^n$ offen mit $\eta_0 \in U$ und $f : U \longrightarrow \mathbb{R}^n$ sei lokal Lipschitz-stetig.

Dann läßt sich jede Lösung des Anfangswertproblems $\dot{x} = f(x)$ mit $x(0) = \eta_0$ zu einer maximalen Lösung fortsetzen und es existiert genau eine maximale Lösung

$$x : (t_-(\eta_0), t_+(\eta_0)) \longrightarrow U.$$

Beweis: Seien $y : I \longrightarrow U$ und $z : J \longrightarrow U$ zwei Lösungen des Anfangswertproblems. Dann betrachten wir das Intervall

$$K := \{s \in I \cap J \mid y(t) = z(t)\ \forall\ 0 \leq t \leq s\} \cup \{s \in I \cap J \mid y(t) = z(t)\ \forall\ 0 \geq t \geq s\}$$

mit den Randpunkten $a = \inf(K)$ und $b = \sup(K)$.

Angenommen $a \in K$, dann ist a kein Randpunkt der offenen Intervalle I und J, aber da das (AWP) $\dot{x} = f(x)$ mit $x(a) = y(a) \stackrel{!}{=} z(a)$ lokal in a mit dem Satz von Picard-Lindelöf 31.10 eindeutig lösbar ist, würden y und z dann noch über a hinaus übereinstimmen im Widerspruch zur Definition von a als Infimum von K.

Nehmen wir $a \in I \cap J$ an, so folgt aus der Stetigkeit von y und z in a unmittelbar

$$y(a) = \lim_{t \to a+} y(t) = \lim_{t \to a+} z(t) = z(a),$$

und damit $a \in K$, was nicht sein kann. Also ist a ein Randpunkt von I oder von J.

Analog sehen wir, daß b ein Randpunkt von I oder von J ist, und wir erhalten insgesamt, daß y und z auf $I \cap J$ übereinstimmen, so daß die Funktion

$$x : I \cup J \longrightarrow U : t \mapsto \begin{cases} y(t), & \text{falls } t \in I, \\ z(t), & \text{falls } t \in J, \end{cases}$$

eine Fortsetzung von y und von z ist. Damit können wir nun die Menge

$$\mathcal{Y} = \{y : I_y \longrightarrow U \mid y \text{ ist Lösung des (AWP) } \dot{x} = f(x) \text{ mit } x(0) = \eta_0\}$$

aller Lösungen des Anfangswertproblems betrachten und die Abbildung

$$x : \bigcup_{y \in \mathcal{Y}} I_y \longrightarrow U : t \mapsto y(t),$$

für $t \in I_y$, ist dann wohldefiniert und ist offenbar die eindeutige maximale Lösung des Anfangswertproblems. □

Beispiel 31.16

a. In Beispiel 31.7 Teil a. haben wir das Anfangswertproblem

$$\dot{x} = a \cdot x$$

mit $x(0) = \eta_0$ auf $\mathbb{R}$ betrachtet und die Funktion

$$x : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto e^{at} \cdot \eta_0$$

als Lösung gefunden. Da die zugehörige Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto ax$$

Lipschitz-stetig ist, folgt aus Korollar 31.15, daß dies die einzige Lösung des Anfangswertproblems ist. Das maximale Existenzintervall ist $\mathbb{R}$.

b. In Teil c. von Beispiel 31.7 haben wir die Funktion

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x^2$$

betrachtet, die stetig differenzierbar und damit lokal Lipschitz-stetig ist. Die dort gefundene Lösung des Anfangswertproblems ist also eindeutig und das maximale

Existenzintervall ist $(-\infty, 0)$, da sich die Funktion nicht stetig nach 0 fortsetzen läßt. Mithin gibt es auch keine Lösung des (AWP) auf ganz $\mathbb{R}$.

Bemerkung 31.17 (Stetige + differenzierbare Abhängigkeit der Lösungen)
Ist $U \subseteq \mathbb{R}^n$ offen und ist $f : U \longrightarrow \mathbb{R}^n$ lokal Lipschitz-stetig, so bezeichnen wir mit

$$I(\eta_0) = (t_-(\eta_0), t_+(\eta_0))$$

das maximale Existenzintervall des (AWP) $\dot{x} = f(x)$ mit $x(0) = \eta_0$ und wenn $x : I(\eta_0) \longrightarrow U$ die zugehörige maximale Lösung ist, so setzen wir

$$\varphi^t(\eta_0) := x(t).$$

Man kann dann zeigen, daß die Menge

$$\Omega = \{(t, \eta) \in \mathbb{R} \times U \mid t \in I(\eta)\}$$

offen ist und daß die Endpunkte der maximalen Existenzintervalle stetig von den Anfangswerten abhängen. Ferner können wir eine Abbildung

$$\varphi : \Omega \longrightarrow U : (t, \eta) \mapsto \varphi^t(\eta) \tag{174}$$

definieren, und man kann zeigen, daß diese stetig ist. Ist f stetig differenzierbar, so ist φ sogar stetig differenzierbar — siehe etwa [Ama83, §§8-10]). Dies zeigt, daß die maximalen Lösungen je nach Qualität von f stetig bzw. stetig differenzierbar vom Anfangswert abhängen. Entsprechende Abhängigkeiten von f kann man ebenfalls formulieren und beweisen, wenn man f mittels eines Parameters stetig differenzierbar variiert. Damit ist dann auch Punkt (3) in Bemerkung 31.8 erfüllt.

D) Flüsse und dynamische Systeme

In Bemerkung 31.17 haben wir gesehen, daß die Lösungen eines Anfangswertproblems bei Variation des Anfangswertes zu einer Familie von Kurven führt. Diesen Umstand wollen wir formalisieren, indem wir den Begriff des Flusses einführen.

Definition 31.18 (Dynamisches System / Fluß)
Sei $U \subseteq \mathbb{R}^n$ offen und für $x \in U$ sei $I(x)$ ein offenes Intervall in $\mathbb{R}$ mit $0 \in I(x)$. Ferner sei

$$\Omega = \bigcup_{x \in U} I(x) \times \{x\} \subseteq \mathbb{R} \times U$$

offen in $\mathbb{R}^{n+1}$. Dann heißt eine stetig differenzierbare Abbildung

$$\varphi : \Omega \longrightarrow U : (t, x) \mapsto \varphi^t(x)$$

ein *dynamisches System* oder ein *Fluß* auf U, wenn

(1) $\varphi^0 = \mathrm{id}_U$.

(2) Für $t \in I(x)$ gilt $s \in I(\varphi^t(x))$ genau dann, wenn $s+t \in I(x)$.

(3) $(\varphi^s \circ \varphi^t)(x) = \varphi^{s+t}(x)$ für alle $t, s+t \in I(x)$.

Wir nennen U den *Phasenraum* des Flusses und

$$x^\varphi := \{\varphi^t(x) \mid t \in I(x)\}$$

die *Bahn*, den *Orbit* oder die *Trajektorie* von x unter dem Fluß. Ist $I(x) = \mathbb{R}$ für alle $x \in U$ und damit $\Omega = \mathbb{R} \times U$, so nennen wir den Fluß φ *global*.

Beispiel 31.19 (Ein globaler Fluß)
Auf der Menge $\Omega = \mathbb{R} \times \mathbb{R}$ ist durch

$$\varphi : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} : (t, x) \mapsto e^t \cdot x$$

ein globaler Fluß gegeben. Der Phasenraum ist in diesem Fall ganz $\mathbb{R}$ und die Bahn x^φ von x unter φ ist

$$x^\varphi = \begin{cases} (-\infty, 0), & \text{wenn } x < 0, \\ \{0\}, & \text{wenn } x = 0, \\ (0, \infty), & \text{wenn } x > 0. \end{cases}$$

Daß der Phasenraum die disjunkte Vereinigung der Bahnen ist, ist kein Zufall, wie das folgende Lemma zeigt.

Lemma 31.20 (Zerlegung des Phasenraums in Bahnen)
Ist $\varphi : \Omega \longrightarrow U$ ein Fluß, dann sind je zwei Bahnen von φ entweder disjunkt oder identisch, und die Bahnen liefern eine disjunkte Zerlegung des Phasenraums.

Beweis: Wir nennen zwei Vektoren $x, y \in U$ im Phasenraum äquivalent, wenn $y \in x^\varphi$ in der Bahn von x liegt. Es reicht, zu zeigen, daß dadurch eine Äquivalenzrelation auf U definiert wird, denn die Bahnen sind dann offensichtlich die Äquivalenzklassen und liefern somit die gesuchte disjunkte Zerlegung.

Ist $x \in U$, so ist wegen (1) $x = \varphi^0(x) \in x^\varphi$ in der Bahn von x unter φ und ist somit äquivalent zu x. Dies zeigt die Reflexivität der Relation.

Sind $x, y \in U$ und ist $y = \varphi^t(x) \in x^\varphi$ in der Bahn von x, so ist wegen (2)

$$-t \in I(\varphi^t(x)) = I(y)$$

und mit (1) und (3) ist dann

$$y^\varphi \ni \varphi^{-t}(y) = \varphi^{-t}(\varphi^t(x)) = \varphi^{-t+t}(x) = \varphi^0(x) = x,$$

d.h. x ist auch in der Bahn von y. Die Relation ist also auch symmetrisch.

Sind $x, y, z \in U$ mit $y = \varphi^t(x)$ in der Bahn von x und $z = \varphi^s(y)$ in der Bahn von y, so gilt

$$s \in I(y) = I(\varphi^t(x)),$$

woraus mit (2) unmittelbar $s + t \in I(x)$ und mit (3) dann

$$z = \varphi^s(y) = \varphi^t(\varphi^s(x)) = \varphi^{s+t}(x)$$

folgt, d.h z ist auch in der Bahn von x. Die Relation ist also transitiv. □

Lemma 31.21 (Flüsse und Diffeomorphismen)
Ist $\varphi : \Omega \longrightarrow U$ ein stetig differenzierbarer Fluß und ist

$$U_t := \{x \in U \mid t \in I(x)\},$$

so ist U_t offen in $\mathbb{R}^n$ und

$$\varphi^t : U_t \longrightarrow U_{-t}$$

ist ein Diffeomorphismus.
Insbesondere, wenn der Fluß global ist, dann ist $U_t = U$ für alle t und

$$(\mathbb{R}, +) \longrightarrow (\mathrm{Diff}(U), \circ) : t \mapsto \varphi^t$$

ist ein Gruppenhomomorphismus.

Beweis: Ist $x \in U_t$, so ist eine ganze Umgebung von $(t, x) \in \Omega$ und diese geschnitten mit $\{t\} \times \mathbb{R}^n$ liefert eine offene Umgebung von x in U_t. Also ist U_t offen in $\mathbb{R}^n$.

Wir beachten zunächst, daß für $t \in I(x)$ und $t + (-t) = 0 \in I(x)$ auch $-t \in I(\varphi^t(x))$ gilt und umgekehrt. Damit erhalten wir

$$\varphi^t(U_t) = \{\varphi^t(x) \mid t \in I(x)\} = \{y \mid \ -t \in I(y)\} = U_{-t},$$

und zudem ist φ^{-t} wegen

$$\varphi^t \circ \varphi^{-t} = \varphi^0 = \mathrm{id}$$

und

$$\varphi^{-t} \circ \varphi^t = \varphi^0 = \mathrm{id}$$

die Inverse von φ^t, so daß φ^t ein Diffeomorphismus ist. □

Beispiel 31.22
Der globale Fluß aus Beispiel 31.19 liefert uns den Gruppenhomomorphismus

$$(\mathbb{R}, +) \longrightarrow (\mathrm{Diff}(\mathbb{R}), \circ) : t \mapsto \varphi^t$$

mit

$$\varphi^t : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto x \cdot e^t,$$

und die φ^t sind dabei einfache lineare Abbildungen.

E) Differentialgleichungen und Flüsse

Proposition 31.23 (Differentialgleichungen liefern Flüsse)
Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \longrightarrow \mathbb{R}^n$ stetig differenzierbar. Ferner sei $I(\eta)$ das maximale Existenzintervall der Lösung des (AWP) $\dot{x} = f(x)$ mit $x(0) = \eta$ und

$$\Omega = \bigcup_{\eta \in U} I(\eta) \times \{\eta\}.$$

Dann ist φ aus (174) in Bemerkung 31.17 ein stetig differenzierbarer Fluß auf U.

Beweis: Aus Bemerkung 31.17 wissen wir, daß Ω offen in $\mathbb{R}^{n+1}$ ist und daß φ stetig differenzierbar ist, wobei

$$I(\eta) \longrightarrow U : t \mapsto \varphi^t(\eta)$$

die eindeutig bestimmte maximale Lösung des (AWP) $\dot{x} = f(x)$ mit $x(0) = \eta$ ist. Damit gilt insbesondere

$$\varphi^0(\eta) = \eta$$

für alle $\eta \in U$, d.h. $\varphi^0 = \mathrm{id}_U$ und (1) in der Definition eines Flusses ist erfüllt. Ist ferner $t \in I(\eta)$ und $\tau_t : \mathbb{R} \longrightarrow \mathbb{R} : s \mapsto s - t$, so lösen

$$\tau_t(I(\eta)) \longrightarrow U : s \mapsto \varphi^{s+t}(\eta)$$

und

$$I(\varphi^t(\eta)) \longrightarrow U : s \mapsto \varphi^s(\varphi^t(\eta))$$

beide das (AWP) $\dot{x} = f(x)$ mit $x(0) = \varphi^t(\eta)$ und sind nicht verlängerbar. Wegen der Eindeutigkeit der maximalen Lösung des (AWP) folgen daraus (2) und (3) in der Definition eines Flusses. □

Beispiel 31.24 (Beispiel eines Flusses in $\mathbb{R}$)
Wenden wir Proposition 31.23 auf die stetig differenzierbare Abbildung

$$f : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto ax$$

an, so erhalten wir den globalen Fluß

$$\varphi : \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} : (t, \eta) \mapsto e^{at} \cdot \eta$$

mit $\mathbb{R}$ als Phasenraum. Ist $a \neq 0$, so hat der Fluß drei Bahnen

$$-1^\varphi = (-\infty, 0), \quad 0^\varphi = \{0\} \quad \text{und} \quad 1^\varphi = (0, \infty).$$

Die Bahnen sind die Spuren von Kurven und haben insofern eine Richtung, in der sie durchlaufen werden. Diese Richtung hängt vom Vorzeichen von a ab (siehe Abbildung 1).

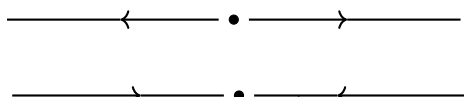


Abbildung 1. Bahnen des Flusses $e^{at} \cdot \eta$ für $a < 0$ und für $a > 0$

Proposition 31.25 (Jeder Fluß ist der Fluß zu einer DGL.)
Sei $\varphi : \Omega \longrightarrow U$ ein stetig differenzierbarer Fluß, dann ist

$$f : U \longrightarrow \mathbb{R}^n : \eta \mapsto \frac{\partial \varphi}{\partial t}(0, \eta)$$

stetig und für jedes $\eta \in U$ ist

$$x : I(\eta) \longrightarrow U : t \mapsto \varphi^t(\eta)$$

eine Lösung des Anfangswertproblems $\dot{x} = f(x)$ mit $x(0) = \eta$.

Beweis: Da φ stetig differenzierbar ist, ist die partielle Ableitung $\frac{\partial \varphi}{\partial t}$ stetig und somit ist auch f stetig. In der folgenden Rechnung wenden wir die Bedingung (3) aus der Definition eines Flusses in der Form

$$\varphi(t, \varphi(s, \eta)) = \varphi^t(\varphi^s(\eta)) = \varphi^{t+s}(\eta) = \varphi(t + s, \eta)$$

an und erhalten

$$f(x(s)) = \frac{\partial}{\partial t}\varphi(t, \varphi(s, \eta))_{|t=0} = \frac{\partial}{\partial t}\varphi(t + s, \eta)_{|t=0} \overset{KR}{=} \frac{\partial}{\partial t}\varphi(t, \eta)_{|t=s} = \dot{x}(s)$$

sowie

$$x(0) = \varphi^0(\eta) = \eta.$$

□

Bemerkung 31.26 (Zusammenhang zwischen Flüssen und DGL)
Proposition 31.23 sagt, daß jede Differentialgleichung einen Fluß definiert, und Proposition 31.25 sagt, daß seinerseits jeder Fluß eine Differentialgleichung liefert, deren Fluß den gegebenen fortsetzt. Differentialgleichungen und Flüsse sind also zwei unterschiedliche Beschreibungen desselben dynamischen Phänomens, und da die Differentialgleichung $\dot{x} = f(x)$ durch das Vektorfeld f festgelegt ist, ist das Vektorfeld eine dritte Möglichkeit das Phänomen zu beschreiben. Die Beschreibung als Differentialgleichung kommt in natürlicher Weise aus den Anwendungen, die Beschreibung als Fluß eignet sich gut für globale Betrachtungen und als Vektorfeld für lokale.

F) Lösung separabler Differentialgleichungen

Wir wollen den Abschnitt mit einem analytischen Verfahren zum Berechnen von Lösungen von Differentialgleichungen von recht einfachem Typ über $\mathbb{R}$ beenden. Dabei werden

Definition 31.27 (Separable Differentialgleichungen)
Es seien $g : \mathbb{R} \longrightarrow \mathbb{R}$ und $h : I \longrightarrow \mathbb{R}$ zwei stetig differenzierbare reellwertige Funktionen mit $I \subseteq \mathbb{R}$ ein offenes Intervall. Eine nicht-autonome Differentialgleichung erster Ordnung der Form

$$\dot{x} = g(x) \cdot h(t)$$

heißt eine *separable Differentialgleichung*.

Bemerkung 31.28 (Lösung durch Separation der Variablen)
Für die Lösung einer separablen Differentialgleichung über $\mathbb{R}$ kann man folgenden Ansatz machen, der als Separation der Variablen bekannt ist: aus der Gleichung

$$\dot{x}(t) = g(x(t)) \cdot h(t)$$

leiten wir

$$h(t) = \frac{\dot{x}(t)}{g(x(t))}$$

her und durch Integration nach t und Substitution von $x = x(t)$ erhalten wir

$$\int^t h(t)\, dt = \int^t \frac{\dot{x}(t)}{g(x(t))}\, dt = \int^x \frac{1}{g(x)}\, dx =: \phi(x). \tag{175}$$

Wenn wir die Funktion ϕ invertieren können und die Stammfunktionen von h kennen, dann können wir die Gleichung nach x auflösen und erhalten eine Ein-Parameter-Familie von Lösungen für die Differentialgleichung.

Der Ansatz kann nur zum Ziel führen, wenn $g(x)$ nie Null ist. In dem Fall ist $g(x)$ dann wegen des Zwischenwertsatzes stets positiv oder stets negativ und damit ist die Ableitung der Funktion

$$\phi : x \mapsto \int^x \frac{1}{g(x)}\, dx$$

stets positiv oder stets negativ. Sie ist dann aber streng monoton und stetig differenzierbar, woraus mit dem Umkehrsatz 14.21 und dem Satz über die Ableitung der Umkehrfunktion 17.14 folgt, daß sie invertierbar ist und eine stetig invertierbare Umkehrabbildung besitzt. Aus der Existenz folgt aber natürlich nicht, daß wir im konkreten Fall die Umkehrabbildung auch in Form bekannter Funktionen ausdrücken können.

Beispiel 31.29 (Lösung mittels Separation der Variablen)

a. Mittels Separation der Variablen können wir die Differentialgleichung

$$\dot{x} = ax$$

lösen. Wir haben $g(x) = x$ und $h(t) = a$ und müssen wegen (175) mithin die Gleichung

$$\phi(x) = \ln|x| = \int^t a\, dt = at + C$$

nach x lösen, woraus

$$|x(t)| = e^{at+C} = e^{at} \cdot e^{C}$$

für eine Konstante C folgt. Setzen wir $c = \pm e^C$, so haben wir die Lösungen aus Beispiel 31.7 a. wiedergefunden.

b. Auch für die Differentialgleichung

$$\dot{x} = x^2$$

aus Beispiel 31.7 c. können wir auf dem Weg lösen mit $g(x) = x^2$ und $h(t) = 1$. Wir erhalten aus (175)

$$\phi(x) = \int^x \frac{1}{x^2}\, dx = \int^t 1\, dt = t + C$$

und damit

$$-\frac{1}{x} = t + C,$$

also

$$x(t) = -\frac{1}{t+C}.$$

Aufgaben

Aufgabe 31.30
Sei $I \subseteq \mathbb{R}$ ein Intervall, $h : I \longrightarrow \mathbb{R}^n$ stetig und $a, b \in I$. Zeige,

$$\left|\left| \int_a^b h(s)\, ds \right|\right|_\infty \leq \int_a^b ||h(s)||_\infty\, ds.$$

Aufgabe 31.31
Sei $y : \mathbb{R} \to \mathbb{R}$ eine stetig differenzierbare nicht konstante Funktion und $g : \mathbb{R} \to \mathbb{R}$ eine stetige Funktion, so daß die Differentialgleichung

$$\dot{y}(t) = g(y(t))$$

für alle $t \in \mathbb{R}$ erfüllt ist. Zeige, daß $y : \mathbb{R} \to \mathbb{R}$ nicht periodisch ist.

Aufgabe 31.32
Sei $f : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ lokal Lipschitz-stetig und sei $x : (t_-, t_+) \longrightarrow \mathbb{R}^n$ eine maximale Lösung von $\dot{x} = f(x)$. Zeige, wenn x beschränkt ist, dann ist $(t_-, t_+) = \mathbb{R}$.

Aufgabe 31.33

Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ lokal Lipschitz-stetig und es gelte

$$\langle x, f(x)\rangle \leq 0$$

alle $x \in \mathbb{R}^n$ mit $\|x\| \leq R$.
Sei $u : (t_-, t_+) \to \mathbb{R}^n$ eine maximale Lösung des Anfangswertproblems

$$\dot{u}(t) = f(u(t)) \text{ mit } u(0) = u_0$$

wobei $\|u_0\| \leq R$. Zeige die folgenden Aussagen:

a. $u(t) \in \overline{U_R(0)}$ für alle $t \in [0, t_+)$.

b. Wenn $t_+ < \infty$ gilt, dann ist u stetig nach t_+ fortsetzbar.

c. Wenn $t_+ < \infty$ gilt, dann ist die Fortsetzung von u in t_+ stetig differenzierbar mit $\dot{u}(t_+) = f(u(t_+))$.

d. $t_+ = \infty$.

Aufgabe 31.34

Gegeben sei ein Fluß mit der Abbildungsvorschrift

$$\varphi(t, x_1, x_2) = \left(x_1 \cdot e^t, \tfrac{x_2}{1-x_2 t}\right)^t.$$

Bestimme eine Differentialgleichung, deren zugehöriger Fluß φ ist.

Aufgabe 31.35

Gegeben sei ein Fluß mit der Abbildungsvorschrift

$$\varphi(t, x) = \frac{x}{\sqrt{1 - 2x^2 t}}.$$

Bestimme eine Differentialgleichung, deren zugehöriger Fluß φ ist.

Aufgabe 31.36

Zeige, daß durch die Abbildungsvorschrift

$$\phi(t, x, y) = \left(\frac{x}{e^t}, \frac{tx + y}{e^t}\right)$$

ein globaler Fluß definiert wird.

Aufgabe 31.37

Sei $\varphi : \Omega \longrightarrow U$ ein Fluß und sei $\eta_0 \in U$. Zeige, wenn es ein $T > 0$ mit $\varphi^T(\eta_0) = \eta_0$ gibt, so gilt $\varphi^T(\eta) = \eta$ für alle $\eta \in \eta_0^\varphi$.

Aufgabe 31.38

Zeige, daß durch die Abbildungsvorschrift

$$\varphi(t,x) = \frac{x \cdot e^t}{x \cdot e^t - x + 1}$$

ein Fluß auf $I = [0,1]$ definiert wird.

Aufgabe 31.39

Berechne die allgemeine Lösung der Differentialgleichung

$$\dot{x} = \frac{\sqrt{1-x^2}}{t}.$$

Aufgabe 31.40

Berechne die allgemeinen Lösungen der folgenden Differentialgleichungen über $\mathbb{R}$:

a. $\dot{x} = t^2 \cdot x$,

b. $\dot{x} + t \cdot x^2 = 0$.

Aufgabe 31.41

Berechne eine Lösung des Anfangswertproblems

$$\dot{x} = -\frac{1}{2t} \cdot \frac{x^2 - 6x + 5}{x - 3}$$

mit $x(1) = 2$ und $x(t) > 1$ für alle t.

Aufgabe 31.42

a. Sei $f : \mathbb{R} \longrightarrow \mathbb{R}$ stetig differenzierbar. Überführe die Differentialgleichung

$$\dot{x} = f\left(\tfrac{x}{t}\right)$$

durch Variablentransformation in eine separable Differentialgleichung.

b. Berechne die alle Lösungen der Differentialgleichung

$$\dot{x} = \frac{2x - t}{t}.$$

Sind die Anfangswertprobleme stets eindeutig lösbar?

§ 32 Lineare gewöhnliche Differentialgleichungen

In diesem Abschnitt betrachten wir $\mathbb{R}^n$ stets als normierten Raum bezüglich der euklidischen Norm $||\cdot||_2$. Ferner betrachten wir auf $\mathrm{Mat}_n(\mathbb{R})$ stets eine Norm $||\cdot||$, die submultiplikativ ist, für die also $||A \circ B|| \leq ||A|| \cdot ||B||$ gilt, wie etwa die euklidische Norm oder eine beliebige Operatornorm.

A) Lineare gewöhnliche Differentialgleichungen erster Ordnung

Definition 32.1 (Lineare gewöhnliche DGL erster Ordnung)
Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und seien $A : I \longrightarrow \mathrm{Mat}_n(\mathbb{R})$ und $b : I \longrightarrow \mathbb{R}^n$ stetig.

a. Ein System gewöhnlicher Differentialgleichungen erster Ordnung der Form

$$\dot{x} = A(t) \circ x + b(t)$$

heißt ein *lineares System gewöhnlicher Differentialgleichungen erster Ordnung mit veränderlichen Koeffizienten.*

b. Ist die Abbildung A konstant, so heißt

$$\dot{x} = Ax + b(t)$$

ein *lineares System gewöhnlicher Differentialgleichungen erster Ordnung mit konstanten Koeffizienten.*

c. Ist b die Nullabbildung, so heißen die linearen Systeme gewöhnlicher Differentialgleichungen in a. und b. *homogen*, sonst heißen sie *inhomogen.*

Beispiel 32.2 (Lösung einer linearen DGL mit konstanten Koeffizienten)
Die Differentialgleichung

$$\dot{x} = a \cdot x$$

aus Beispiel 31.7 ist eine homogene lineare gewöhnliche Differentialgleichung erster Ordnung mit konstanten Koeffizienten. Die allgemeine Lösung ist

$$x : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto e^{at} \cdot c$$

für einen Parameter $c \in \mathbb{R}$. Der Lösungsraum

$$L_h = \{e^{at} \cdot c \mid c \in \mathbb{R}\}$$

aller Lösungen der Differentialgleichung ist also ein Vektorraum der Dimension 1.

Satz 32.3 (Existenz und Eindeutigkeit der Lösung im homogenen Fall)
Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $A : I \longrightarrow \mathrm{Mat}_n(\mathbb{R})$ stetig, $\eta_0 \in \mathbb{R}^n$ und $t_0 \in I$. Dann besitzt das Anfangswertproblem $\dot{x} = A(t) \circ x$ mit $x(t_0) = \eta_0$ genau eine maximale Lösung $x : I \longrightarrow \mathbb{R}^n$.

Beweis: Wir betrachten zunächst ein beliebiges Teilintervall $[a, b] \subseteq I$ mit $t_0 \in (a, b)$. Der Vektorraum $V = \mathcal{C}([a, b], \mathbb{R}^n)$ ist mit der Maximumsnorm

$$||x||_\infty = \max_{t\in[a,b]} ||x(t)||_2$$

ein Banachraum.

Wie in Bemerkung 31.11 sehen wir mit Hilfe des Hauptsatzes der Differential- und Integralrechnung, daß $x \in \mathcal{C}^1([a, b], \mathbb{R}^n)$ genau dann eine Lösung des (AWP) $\dot{x} = A(t) \circ x$ mit $x(t_0) = \eta_0$ ist, wenn $x \in V$ eine Lösung der Integralgleichung

$$x(t) = \eta_0 + \int_{t_0}^{t} A(s) \circ x(s)\, ds$$

ist. Dies wiederum ist gleichwertig dazu, daß $x \in V$ eine Lösung der Gleichung

$$(\mathrm{id}_V - \mathcal{I})(x) = \eta_0$$

ist mit

$$\mathcal{I} : V \longrightarrow V : x \mapsto \left(t \mapsto \int_{t_0}^{t} A(s) \circ x(s)\, ds \right).$$

Wenn wir zeigen können, daß $\mathrm{id}_V - \mathcal{I}$ invertierbar ist, dann besitzt das Anfangswertproblem auf $[a, b]$ die eindeutig bestimmte Lösung

$$x = (\mathrm{id}_V - \mathcal{I})^{-1}(\eta_0),$$

wobei wir η_0 als konstante Abbildung $t \mapsto \eta_0$ auffassen. Die Invertierbarkeit von $\mathrm{id}_V - \mathcal{I}$ wollen wir nun mit Hilfe der Neumannschen Reihe 23.56 zeigen, d.h. wir müssen zeigen, daß die Reihe $\sum_{k=0}^{\infty} ||\mathcal{I}^k||$ konvergiert, wobei $||\mathcal{I}||$ die Operatornorm von $\mathcal{I}$ bezügich der Maximumsnorm $|| \cdot ||_\infty$ auf V bezeichnet. Man beachte dabei, daß damit insbesondere $||\mathcal{I}|| < \infty$ und somit $\mathcal{I} \in L(V, V)$ gezeigt ist.

Dazu zeigen wir zunächst per Induktion nach k, daß

$$||\mathcal{I}^k(x)(t)||_2 \leq \frac{C^k \cdot |t - t_0|^k}{k!} \cdot ||x||_\infty \tag{176}$$

für ein fest gegebenes $x \in V$ und $t \in [a, b]$ gilt, wobei

$$C = \max_{t\in[a,b]} ||A(t)||$$

und $||A(t)||$ die Operatornorm von $A(t)$ bezüglich der euklidischen Norm auf $\mathbb{R}^n$ bezeichnet. Man beachte, daß das Maximum existiert, da die Abbildung $t \mapsto ||A(t)||$

stetig und $[a,b]$ kompakt ist. Für $k=0$ hat die Ungleichung die Form

$$||\mathcal{I}^0(x)(t)||_2 = ||x(t)||_2 \leq ||x||_\infty = \frac{C^0 \cdot (t-t_0)^0}{0!} \cdot ||x||_\infty$$

und ist erfüllt. Wir nehmen nun an, daß sie für k bereits erfüllt ist, und erhalten

$$\begin{aligned} ||\mathcal{I}^{k+1}(x)(t)||_2 =& ||\mathcal{I}\big(\mathcal{I}^k(x)\big)(t)||_2 = \left|\left|\int_{t_0}^t A(s) \circ \mathcal{I}^k(x)(s)\, ds\right|\right|_2 \\ &\overset{(173)}{\leq} \left|\int_{t_0}^t ||A(s) \circ \mathcal{I}^k(x)(s)||_2\, ds\right| \overset{23.51}{\leq} \left|\int_{t_0}^t ||A(s)|| \circ ||\mathcal{I}^k(x)(s)||_2\, ds\right| \\ &\overset{Ind.}{\leq} \left|\int_{t_0}^t C \cdot \frac{C^k \cdot |s-t_0|^k}{k!} \cdot ||x||_\infty\, ds\right| = \frac{C^{k+1} \cdot |t-t_0|^{k+1}}{(k+1)!} \cdot ||x||_\infty. \end{aligned}$$

Die Ungleichung (176) ist also mit Induktion gezeigt. Daraus leiten unmittelbar

$$||\mathcal{I}^k(x)||_\infty = \max_{t\in[a,b]} ||\mathcal{I}^k(x)(t)||_2 \leq \frac{C^k \cdot (b-a)^k}{k!} \cdot ||x||_\infty$$

und damit

$$||\mathcal{I}^k|| = \sup_{0\neq x\in V} \frac{||\mathcal{I}^k(x)||_\infty}{||x||_\infty} \leq \frac{C^k \cdot (b-a)^k}{k!}$$

ab, was wiederum

$$\sum_{k=0}^\infty ||\mathcal{I}^k|| \leq \sum_{k=0}^\infty \frac{C^k \cdot (b-a)^k}{k!} = e^{C\cdot(b-a)} < \infty$$

die für die Konvergenz der Neumannschen Reihe von $\mathcal{I}$ in Proposition 23.56 geforderte Konvergenz der Reihe $\sum_{k=0}^\infty ||\mathcal{I}^k||$ liefert.

Wir haben die Existenz und Eindeutigkeit der Lösung x auf $[a,b]$ somit gezeigt. Vergrößern wir das Intervall nun innerhalb von I, so erhalten wir auf dem größeren Intervall ebenfalls eine eindeutige Lösung, die auf $[a,b]$ mit x übereinstimmen muß. Auf dem Weg können wir die Lösung schließlich auf ganz I auf eindeutige Weise fortsetzen, wie wir das auch im Beweis des Satzes von Picard-Lindelöf getan haben. □

Bemerkung 32.4 (Lösung im eindimensionalen Fall)
Im eindimensionalen Fall können wir das Anfangswertproblem

$$\dot{x} = a(t) \cdot x$$

mit $x(t_0) = x_0$ auch explizit lösen mittels

$$x(t) = e^{\int_{t_0}^t a(s)\, ds} \cdot x_0.$$

B) Fundamentalsysteme

Korollar 32.5 (Fundamentalsysteme und Lösungsräume)
Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und seien $A : I \longrightarrow \mathrm{Mat}_n(\mathbb{R})$ und $b : I \longrightarrow \mathbb{R}^n$ stetig.

a. Der Lösungsraum

$$L_h := \{x \in \mathcal{C}^1(I, \mathbb{R}^n) \mid \dot{x} = A(t) \circ x\}$$

der homogenen linearen gewöhnlichen Differentialgleichung $\dot{x} = A(t) \circ x$ ist ein Unterraum von $\mathcal{C}^1(I, \mathbb{R}^n)$.

b. Für Lösungen $x^1, \ldots, x^k \in L_h$ sind die folgenden Aussagen gleichwertig:
   (i) $(x^1, \ldots, x^k)$ ist linear unabhängig in L_h.
   (ii) Für jedes $t \in I$ ist $(x^1(t), \ldots, x^k(t))$ linear unabhängig in $\mathbb{R}^n$.
   (iii) Es gibt ein $t_0 \in I$, so daß $(x^1(t_0), \ldots, x^k(t_0))$ linear unabhängig in $\mathbb{R}^n$ ist.

c. $\dim_{\mathbb{R}}(L_h) = n$.

d. Jede Basis $(x^1, \ldots, x^n)$ von L_h liefert stetig differenzierbare Abbildungen

$$X : I \longrightarrow \mathrm{Gl}_n(\mathbb{R}) : t \mapsto X(t) = (x^1(t) \ldots x^n(t))$$

und

$$I \longrightarrow \mathrm{Gl}_n(\mathbb{R}) : t \mapsto X(t)^{-1}.$$

Wir nennen X auch ein *Fundamentalsystem* von $\dot{x} = A(t) \circ x$.

e. Ist $x : I \longrightarrow \mathbb{R}^n$ eine Lösung von $\dot{x} = A(t) \circ x + b(t)$, dann ist

$$L_i := x + L_h := \{x + y \mid y \in L_h\}$$

die Menge aller Lösungen der Differentialgleichung auf I.

Beweis: a. Sind $x, y \in L_h$ und $\lambda, \mu \in \mathbb{R}$, so gilt

$$(\lambda \cdot x + \mu \cdot y)' = \lambda \cdot \dot{x} + \mu \cdot \dot{y} = \lambda \cdot A(t) \circ x + \mu \cdot A(t) \circ y = A(t) \circ (\lambda \cdot x + \mu \cdot y),$$

also ist L_h ein Unterraum von $\mathcal{C}^1(I, \mathbb{R}^n)$.

b. Wir setzen zunächst voraus, daß $(x^1, \ldots, x^k)$ linear unabhängig in L_h ist und wählen ein $t_0 \in I$. Dann betrachten wir Skalare $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$, für die

$$\lambda_1 \cdot x^1(t_0) + \ldots + \lambda_k \cdot x^k(t_0) = 0$$

gilt. Dann sind $\lambda_1 \cdot x^1 + \ldots + \lambda_k \cdot x^k$ und $x = 0$ zwei Lösungen des Anfangswertproblems $\dot{x} = A(t) \circ x$ mit $x(t_0) = 0$. Wegen der Eindeutigkeit der Lösung folgt dann

$$\lambda_1 \cdot x^1 + \ldots + \lambda_k \cdot x^k = 0,$$

und aus der linearen Unabhängigkeit von $(x^1, \ldots, x^k)$ folgt

$$\lambda_1 = \ldots = \lambda_k = 0.$$

Mithin ist $(x^1(t_0), \ldots, x^k(t_0))$ für jedes $t_0 \in I$ linear unabhängig, (i) impliziert also (ii). Offenbar impliziert (ii) auch (iii).
Sei also nun $t_0 \in I$ so, daß $(x^1(t_0), \ldots, x^k(t_0))$ linear unabhängig in $\mathbb{R}^n$ ist. Wir betrachten nun Skalare $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ mit

$$\lambda_1 \cdot x^1 + \ldots + \lambda_k \cdot x^k = 0.$$

Dann gilt insbesondere

$$\lambda_1 \cdot x^1(t_0) + \ldots + \lambda_k \cdot x^k(t_0) = 0$$

und aus der linearen Unabhängigkeit der Vektoren in $\mathbb{R}^n$ folgt

$$\lambda_1 = \ldots = \lambda_k = 0,$$

so daß $(x^1, \ldots, x^k)$ auch linear unabhängig in L_h ist. Damit impliziert also (iii) auch (i).

c. Sei $x^i \in L_h$ die eindeutig bestimmte maximale Lösung des (AWP) $\dot{x} = A(t) \circ x$ mit $x^i(t_0) = e_i \in \mathbb{R}^n$ für ein festes $t_0 \in I$. Dann ist $(x^1, \ldots, x^n)$ wegen b. linear unabhängig in L_h, woraus

$$\dim_{\mathbb{R}}(L_h) \geq n$$

folgt. Ist $x^0 \in L_h$ eine weitere Lösung von $\dot{x} = A(t) \circ x$, so ist $(x^0(t_0), \ldots, x^n(t_0))$ aus Dimensionsgründen linear abhängig in $\mathbb{R}^n$, so daß $(x^0, \ldots, x^n)$ wegen b. in L_h ebenfalls linear abhängig. Es folgt also

$$\dim_{\mathbb{R}}(L_h) = n.$$

d. Ist $(x^1, \ldots, x^n)$ eine Basis von L_h und ist

$$X(t) = (x^1(t) \ldots x^n(t))$$

die Matrix, deren Spalten die $x^i(t)$ sind, dann ist $X(t)$ nach b. invertierbar und wir erhalten eine stetig differenzierbare Abbildung

$$X : I \longrightarrow \mathrm{Gl}_n(\mathbb{R}) : t \mapsto X(t)$$

in die offene Teilmenge $\mathrm{Gl}_n(\mathbb{R})$ von $\mathrm{Mat}_n(\mathbb{R})$. Daß die Abbildung

$$I \longrightarrow \mathrm{Gl}_n(\mathbb{R}) : t \mapsto X(t)^{-1}$$

ebenfalls stetig differenzierbar ist, liegt an der Adjunktenformel A9.29 für die Inverse

$$X(t)^{-1} = \frac{X(t)^{\#}}{\det(X(t))},$$

die zeigt, daß die Inverse von $X(t)$ eine rationale Funktion in den stetig differenzierbaren Einträgen von $X(t)$ ist.

e. Ist $x : I \longrightarrow \mathbb{R}^n$ eine Lösung der inhomogenen Differentialgleichung und ist $y \in L_h$, so gilt

$$(x+y)' = \dot{x} + \dot{y} = A(t) \circ x + b(t) + A(t) \circ y = A(t) \circ (x+y) + b(t),$$

so daß $x + y$ also auch eine Lösung der inhomogenen Differentialgleichung ist Ist umgekehrt $z : I \longrightarrow \mathbb{R}^n$ eine zweite Lösung der inhomogenen Differentialgleichung, so gilt

$$(z-x)' = \dot{z} - \dot{x} = A(t) \circ z + b(t) - A(t) \circ x - b(t) = A(t) \circ (z-x)$$

und somit $z - x \in L_h$, woraus

$$z = x + (z-x) \in x + L_h$$

folgt.

□

C) Variation der Konstanten

Satz 32.6 (Existenz und Eindeutigkeit im inhomogenen Fall)
Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, seien $A : I \longrightarrow \mathrm{Mat}_n(\mathbb{R})$ und $b : I \longrightarrow \mathbb{R}^n$ stetig, sei X ein Fundamentalsystem von $\dot{x} = A(t) \circ x$ und seien $t_0 \in I$ und $\eta_0 \in \mathbb{R}^n$. Dann besitzt das (AWP) $\dot{x} = A(t) \circ x + b(t)$ mit $x(t_0) = \eta_0$ genau eine maximale Lösung $x : I \longrightarrow \mathbb{R}^n$, und es gilt

$$x(t) = X(t) \circ \left(X(t_0)^{-1} \circ \eta_0 + \int_{t_0}^{t} X(s)^{-1} \circ b(s)\, ds \right). \tag{177}$$

Die Formel wird auch *Variation der Konstanten* genannt.

Beweis: Aus Korollar 32.5 wissen wir, daß X und die Abbildung $t \mapsto X(t)^{-1}$ auf I stetig differenzierbar sind. Die Abbildung

$$x : I \longrightarrow \mathbb{R}^n : t \mapsto X(t) \circ \left(X(t_0)^{-1} \circ \eta_0 + \int_{t_0}^{t} X(s)^{-1} \circ b(s)\, ds \right)$$

ist also stetig differenzierbar. Setzen wir

$$c : I \longrightarrow \mathbb{R}^n : t \mapsto X(t_0)^{-1} \circ \eta_0 + \int_{t_0}^{t} X(s)^{-1} \circ b(s)\, ds,$$

so gilt

$$x(t) = X(t) \circ c(t)$$

und

$$\dot{c}(t) = X(t)^{-1} \circ b(t).$$

Damit ergibt sich x die Ableitung

$$\begin{aligned}\dot{x}(t) =& \dot{X}(t) \circ c(t) + X(t) \circ \dot{c}(t) \\ =& A(t) \circ X(t) \circ c(t) + X(t) \circ X(t)^{-1} \circ b(t) = A(t) \circ x(t) + b(t),\end{aligned}$$

so daß x eine Lösung der Differentialgleichung ist. Die Eindeutigkeit ergibt sich dann aus der Eindeutigkeit der Lösung der homogenen Differentialgleichung. Denn ist $z : J \longrightarrow \mathbb{R}^n$ eine zweite maximale Lösung mit $J \subseteq I$, so ist $z - x$ auf J eine Lösung der homogenen Differentialgleichung und die ist eindeutig, so daß $x = z$ auf J gilt, und aus der Maximalität folgt dann auch $I = J$. □

Bemerkung 32.7 (Variation der Konstanten)
Man kann die Formel (177) aus dem Ansatz

$$x(t) = X(t) \circ c(t)$$

mit $c : I \longrightarrow \mathbb{R}^n$ stetig differenzierbar herleiten. Denn wir möchten c so bestimmen, daß

$$\dot{X}(t) \circ c(t) + X(t) \circ \dot{c}(t) = A(t) \circ X(t) \circ c(t) + b(t)$$

gilt. Wegen $\dot{X}(t) = A(t) \circ X(t)$ ist dies gleichwertig dazu, daß

$$X(t) \circ \dot{c}(t) = b(t)$$

oder alternativ

$$\dot{c}(t) = X(t)^{-1} \circ b(t)$$

gilt. Dann läßt sich c aber mittels Integration als

$$c(t) = \int_{t_0}^{t} X^{-1}(s) \circ b(s)\, ds + \tilde{c}$$

für einen konstanten Vektor $\tilde{c}$ ermitteln. Den konstanten Vektor erhält man aus der Anfangswertbedingung

$$\eta_0 = x(t_0) = X(t_0) \circ \tilde{c}$$

oder alternativ

$$\tilde{c} = X(t_0)^{-1} \circ \eta_0.$$

Der Ansatz mittels der Abbildung c ist ein Ansatz mit variierender Konstanten und erläutert den Begriff *Variation der Konstanten* für die Formel (177).

Beispiel 32.8 (Variation der Konstanten)
Wir wollen nun mit Hilfe von Bemerkung 32.4 und Variation der Konstanten das Anfangswertproblem

$$\dot{x} = 2t \cdot x - 2t^3$$

mit $x(0) = 1$ lösen. Als Fundamentalsystem für die homogene Differentialgleichung erhalten wir

$$x_h(t) = e^{\int_0^t 2s\, ds} = e^{t^2}$$

und als spezielle inhomogene Lösung dann

$$x_i(t) = e^{t^2} \cdot \left(e^{-0^2} \cdot 1 + \int_0^t e^{-s^2} \cdot (-2s^3)\, ds \right) = e^{t^2} + e^{t^2} \cdot \int_0^t e^{-s^2} \cdot (-2s^3)\, ds.$$

Das Integral können wir nun mittels partieller Integration lösen

$$\int_0^t e^{-s^2} \cdot (-2s^3)\, ds = e^{-t^2} \cdot t^2 - \int_0^t e^{-s^2} \cdot 2s\, ds = e^{-t^2} \cdot t^2 + e^{-t^2} - 1,$$

und erhalten damit

$$x_i(t) = t^2 + 1.$$

D) Lineare gewöhnliche Differentialgleichungen höherer Ordnung

Definition 32.9 (Lineare gewöhnliche DGL m-ter Ordnung)
Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und seien $a_0, \ldots, a_{m-1}, b : I \longrightarrow \mathbb{R}$ stetig.

a. Eine gewöhnliche Differentialgleichungen m-ter Ordnung der Form

$$x^{(m)} = a_{m-1}(t) \cdot x^{(m-1)} + \ldots + a_0(t) \cdot x + b(t)$$

heißt eine *lineare gewöhnliche Differentialgleichung m-ter Ordnung mit veränderlichen Koeffizienten.*

b. Sind die Abbildungen $a_0, \ldots, a_{m-1}$ konstant, so heißt

$$x^{(m)} = a_{m-1} \cdot x^{(m-1)} + \ldots + a_0 \cdot x + b(t)$$

eine *lineare gewöhnliche DGL m-ter Ordnung mit konstanten Koeffizienten.*

c. Ist b die Nullabbildung, so heißen die linearen gewöhnlichen Differentialgleichungen in a. und b. *homogen*, sonst heißen sie *inhomogen.*

Bemerkung 32.10 (Reduktion höherer Ordnung auf erste Ordnung)
Mit dem Ansatz aus Bemerkung 31.3 erhalten wir, daß $z : I \longrightarrow \mathbb{R}^m$ genau dann eine Lösung des Anfangswertproblems

$$\dot{z} = A(t) \circ z + \tilde{b}(t)$$

mit $\tilde{b}(t) = (0, \ldots, 0, b(t))^t$ und

$$A(t) = \begin{pmatrix} 0 & 1 & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & \ldots & 0 & 0 & 1 \\ a_0(t) & \ldots & \ldots & \ldots & \ldots & a_{m-2}(t) & a_{m-1}(t) \end{pmatrix}$$

sowie $z(t_0) = (\eta_0, \ldots, \eta_{m-1}) \in \mathbb{R}^m$ ist, wenn z_0 eine Lösung des Anfangswertproblems

$$x^{(m)} = a_{m-1}(t) \cdot x^{(m-1)} + \ldots + a_0(t) \cdot x + b(t)$$

mit $x(t_0) = \eta_0$ ist und zudem $z_i = \dot{z}_{i-1}$ mit $z_0^{(i)}(t_0) = \eta_i$ gilt. Aussagen zur Lösbarkeit für Anfangswertprobleme linearer gewöhnlicher Differentialgleichungen erster Ordnung übertragen sich also auf lineare Differentialgleichungen m-ter Ordnung.

Korollar 32.11 (Existenz und Eindeutigkeit der Lösung)
Sei $I \subseteq \mathbb{R}$ ein offenes Intervall und seien $a_0, \ldots, a_{m-1}, b : I \longrightarrow \mathbb{R}$ stetig.

a. Die Menge L_h der Lösungen der homogenen linearen Differentialgleichung

$$x^{(m)} = a_{m-1}(t) \cdot x^{(m-1)} + \ldots + a_0(t) \cdot x$$

ist ein Unterraum von $\mathcal{C}^m(I, \mathbb{R})$ der Dimension m.

b. Ist $x \in \mathcal{C}^m(I, \mathbb{R}^n)$ ein Lösung der inhomogenen linearen Differentialgleichung

$$x^{(m)} = a_{m-1}(t) \cdot x^{(m-1)} + \ldots + a_0(t) \cdot x + b(t),$$

so ist die Menge

$$L_i := x + L_h := \{x + y \mid y \in L_h\}$$

die Menge aller Lösungen der Differentialgleichung.

c. Für jedes $t_0 \in I$ und jedes $\eta = (\eta_0, \ldots, \eta_{m-1}) \in \mathbb{R}^m$ besitzt das (AWP)

(178) $$x^{(m)} = a_{m-1}(t) \cdot x^{(m-1)} + \ldots + a_0(t) \cdot x + b(t),$$

mit $x^{(i)}(t_0) = \eta_i$, $i = 0, \ldots, m-1$, genau eine maximale Lösung $x \in \mathcal{C}^m(I, \mathbb{R})$.

d. Ist $(x_1, \ldots, x_m)$ eine Basis von L_h, so erhalten wir durch *Variation der Konstanten* eine Lösung

$$x : I \longrightarrow \mathbb{R} : t \mapsto \sum_{i=1}^{m} x_i(t) \cdot (-1)^{m+i} \cdot \int_{t_0}^{t} \frac{W_i(s) \cdot b(s)}{W(s)}\, ds$$

der Differentialgleichung (178), wobei

$$W(t) = \det(X(t))$$

mit

$$X = \begin{pmatrix} x_1 & \ldots & x_m \\ \dot{x}_1 & \ldots & \dot{x}_m \\ \vdots & & \vdots \\ x_1^{(m-1)} & \ldots & x_m^{(m-1)} \end{pmatrix}$$

die sogenannte *Wronski-Determinante* des *Fundamentalsystems* $(x_1, \ldots, x_m)$ ist und $W_i(t)$ die Determinante der Matrix ist, die aus $X(t)$ durch Streichung der letzten Zeile und der i-ten Spalte entsteht.

Beweis: Die Teile a.-c. folgen unmittelbar aus Satz 32.6 und Korollar 32.5. Für den Beweis von Teil d. verwenden wir die Reduktion auf eine lineare Differentialgleichung erster Ordnung aus Bemerkung 32.10 und betrachten für diese die Variation der Konstanten aus Bemerkung 32.7 mit dem Ansatz

$$z(t) = X(t) \circ c(t)$$

und der Reduktion auf die Gleichung

$$X(t) \circ \dot{c}(t) = \tilde{b}(t).$$

Aus der Cramerschen Regel A9.36 erhalten wir

$$\dot{c}_i(t) = \frac{\det(B(t))}{\det(X(t))} = (-1)^{m+i} \cdot \frac{W_i(t) \cdot b(t)}{W(t)},$$

wobei

$$B(t) = \begin{pmatrix} x_1 & \dots & x_{i-1} & 0 & x_{i+1} & \dots & x_m \\ \dot{x}_1 & \dots & x_{i-1} & 0 & x_{i+1} & \dots & \dot{x}_m \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ x_1^{(m-1)} & \dots & x_{i-1}^{(m-1)} & b(t) & x_{i+1}^{(m-1} & \dots & x_m^{(m-1)} \end{pmatrix}.$$

Integration liefert dann

$$c_i(t) = (-1)^{m+i} \cdot \int_{t_0}^{t} \frac{W_i(s) \cdot b(s)}{W(s)}\, ds$$

als eine Lösung für die i-te Komponente von c und wir erhalten dann wie gewünscht mit $z(t) = X(t) \circ c(t)$ die Lösung

$$x(t) = z_0(t) = \sum_{i=1}^{m} x_i(t) \cdot c_i(t).$$

□

E) Lineare DGL mit konstanten Koeffizienten - 1. Ordnung

Beispiel 32.12 (Gekoppelter elektrischer Schwingkreis)

Betrachten wir einen gekoppelten elektrischen Schwingkreis wie in Abbildung 2 mit Spannung $E : \mathbb{R} \longrightarrow \mathbb{R}$, Widerständen R_1 und R_2 sowie Spuleninduktivitäten L_1 und L_2, so erhalten wir für die Stromstärken die lineare Differentialgleichung mit konstanten Koeffizienten

$$\begin{pmatrix} \dot{I}_2 \\ \dot{I}_3 \end{pmatrix} = - \begin{pmatrix} \frac{R_1+R_2}{L_1} & \frac{R_1}{L_1} \\ \frac{R_1}{L_2} & \frac{R_1}{L_2} \end{pmatrix} \circ \begin{pmatrix} I_2 \\ I_3 \end{pmatrix} + \begin{pmatrix} \frac{E(t)}{L_1} \\ \frac{E(t)}{L_2} \end{pmatrix}.$$

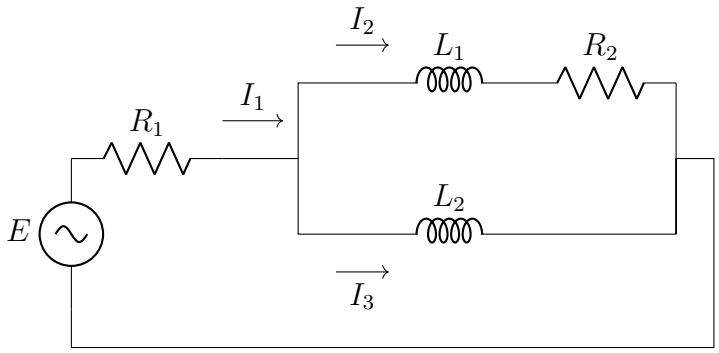


ABBILDUNG 2. Gekoppelter elektrischer Schwingkreis

Satz 32.13 (Lineare DGL mit konstanten Koeffizienten)
Sei $I \subseteq \mathbb{R}$ ein offenes Intervall, $A \in \mathrm{Mat}_n(\mathbb{R})$, $b : I \longrightarrow \mathbb{R}^n$ stetig, $t_0 \in I$, $\eta_0 \in \mathbb{R}^n$.

a. Dann ist
$$x : \mathbb{R} \longrightarrow \mathbb{R}^n : t \mapsto e^{A\cdot(t-t_0)} \cdot \eta_0$$
die eindeutige maximale Lösung des (AWP) $\dot{x} = Ax$ mit $x(t_0) = \eta_0$.
Zudem ist $X : \mathbb{R} \longrightarrow \mathrm{Gl}_n(\mathbb{R}) : t \mapsto e^{A\cdot t}$ ein Fundamentalsystem von $\dot{x} = Ax$ und der zugehörige globale Fluß ist
$$\varphi : \mathbb{R}^{n+1} \longrightarrow \mathbb{R}^n : (t, \eta)^t \mapsto e^{A\cdot t} \circ \eta.$$

b. Mittels *Variation der Konstanten* erhalten wir die eindeutige Lösung
$$x : I \longrightarrow \mathbb{R}^n : t \mapsto e^{A\cdot(t-t_0)} \circ \eta_0 + \int_{t_0}^{t} e^{A\cdot(t-s)} \circ b(s)\, ds$$
des Anfangswertproblems $\dot{x} = Ax + b(t)$ mit $x(0) = \eta_0$.

Beweis: a. Aus Korollar 25.32 wissen wir, daß die Abbildung
$$x : \mathbb{R} \longrightarrow \mathbb{R}^n : t \mapsto e^{A\cdot(t-t_0)} \cdot \eta_0$$
unendlich oft differenzierbar ist und daß für ihre Ableitung
$$\dot{x}(t) = A \circ e^{A\cdot(t-t_0)} \cdot \eta_0 = A \circ x(t)$$
gilt. Damit ist x eine Lösung, die auf ganz $\mathbb{R}$ existiert, und nach Satz 32.3 ist sie zudem eindeutig. Das angegebene Fundamentalsystem erhalten wir, indem wir die Lösungen der Anfangswertprobleme zu $x(0) = e_i$ für $i = 1, \ldots, n$ berechnen, die wegen Korollar 32.5 eine Basis von L_h bilden.

b. Unter Beachtung von Korollar 25.32 und dem Hauptsatz der Differential- und Integralrechnung erhalten wir als Ableitung der Abbildung
$$x : I \longrightarrow \mathbb{R}^n : t \mapsto e^{A\cdot t} \circ c(t)$$

mit

$$c : I \longrightarrow \mathbb{R}^n : t \mapsto e^{-A\cdot t_0} \circ \eta_0 + \int_{t_0}^{t} e^{-A\cdot s} \circ b(s)\, ds$$

den Ausdruck

$$\dot{x}(t) = A \circ e^{A\cdot t} \circ c(t) + e^{A\cdot t} \circ \dot{c}(t) = A \circ x(t) + e^{A\cdot t} \circ \dot{c}(t).$$

Wegen

$$\dot{c}(t) = e^{-A\cdot t} \circ b(t)$$

und wegen $x(t_0) = \eta_0$ ist damit gezeigt, daß x das (AWP) löst. Die Eindeutigkeit der Lösung folgt dann aus Satz 32.6.

□

Bemerkung 32.14 (Berechnung von $e^{A\cdot t}$)
Ist $A \in \mathrm{Mat}_n(\mathbb{R})$, so liefert uns $e^{A\cdot t}$ ein Fundamentalsystem der Differentialgleichung $\dot{x} = Ax$, und wir sind deshalb daran interessiert, $e^{A\cdot t}$ berechnen zu können. Im folgenden schreiben wir $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ für eine Diagonalmatrix mit den Diagonaleinträgen $\lambda_1, \ldots, \lambda_n$.

a. Ist $A = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, dann ist

$$e^{A\cdot t} = \mathrm{diag}(e^{\lambda_1 t}, \ldots, e^{\lambda_n t})$$

ebenfalls eine Diagonalmatrix.

b. Ist A symmetrisch, so gibt es wegen des Spektralsatzes für symmetrische Matrizen A15.32 eine invertierbare Matrix $T \in \mathrm{Gl}_n(\mathbb{R})$, so daß

$$T^{-1} \circ A \circ T = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$$

eine Diagonalmatrix ist, auf deren Diagonale die Eigenwerte von A stehen. Aus Korollar 25.32 erhalten wir

$$T^{-1} \circ e^{A\cdot t} \circ T = e^{T^{-1}\circ A\circ T\cdot t} = \mathrm{diag}(e^{\lambda_1 t}, \ldots, e^{\lambda_n t})$$

und damit

$$e^{A\cdot t} \circ T = T \circ \mathrm{diag}(e^{\lambda_1 t}, \ldots, e^{\lambda_n t}) = (e^{\lambda_1 t} \cdot u_1 \ldots e^{\lambda_n t} \cdot u_n), \tag{179}$$

wenn $u_1, \ldots, u_n$ die Spalten von T sind. Da die Multiplikation mit T von rechts den Spaltenraum von e^{At} nicht ändert, liefert uns die Matrix in (179) ebenfalls ein Fundamentalsystem von $\dot{x} = Ax$.

c. Für die nilpotente Matrix

$$N = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ \vdots & & & & \ddots & 1 \\ 0 & \dots & \dots & \dots & \dots & 0 \end{pmatrix}$$

zeigt eine leichte Rechnung

$$e^{N\cdot t} = \begin{pmatrix} 1 & t & \frac{t^2}{2} & \dots & \dots & \frac{t^{n-1}}{(n-1)!} \\ 0 & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \frac{t^2}{2} \\ \vdots & & & \ddots & \ddots & t \\ 0 & \dots & \dots & \dots & 0 & 1 \end{pmatrix}$$

d. Ist A ein Jordanblock der Form

$$A = \lambda \cdot \mathbb{1}_n + N,$$

dann kommutiert $\mathbb{1}_n$ mit N und wir erhalten aus Korollar 25.32

$$e^{A\cdot t} = e^{\lambda\cdot\mathbb{1}_n\cdot t} \circ e^{N\cdot t} = e^{\lambda\cdot t} \cdot e^{N\cdot t}.$$

e. Ist $A = \bigoplus_{i=1}^{k} J(\lambda_i)$ eine Matrix in Jordanscher Normalform mit Jordanblöcken $J(\lambda_i)$, dann ist

$$e^{A\cdot t} = \bigoplus_{i=1}^{k} e^{J(\lambda_i)\cdot t}$$

eine Blockdiagonalmatrix, deren Blöcke wie in d. aussehen. Das Fundamentalsystem hat dann Einträge, die von der Form

$$p(t) \cdot e^{\lambda t}$$

sind für einen Eigenwert λ von A und ein Polynom, dessen Grad echt kleiner als der Nilpotenzindex des Eigenwertes ist.

f. Der Satz über die Jordansche Normalform A13.2 sagt, daß wir über $\mathbb{C}$ jede Matrix A mittels einer invertierbaren Matrix $T \in \mathrm{Gl}_n(\mathbb{C})$ in Jordansche Normalform $T^{-1} \circ A \circ T$ transformieren können. Ist die Transformationsmatrix reell, so erhalten wir aus d. mit dem Ansatz aus b. ein Fundamentalsystem, dessen Einträge wieder von der Form

$$p(t) \cdot e^{\lambda t}$$

sind für einen Eigenwert λ von A und ein Polynom, dessen Grad kleiner als der Nilpotenzindex des Eigenwertes ist. Ist die Transformationsmatrix T hingegen komplex und ist

$$\lambda = \alpha + i \cdot \beta$$

eine Zerlegung des Eigenwertes λ in Real- und Imaginärteil, dann ist

$$e^{\lambda t} = e^{\alpha t} \cdot \cos(\beta t) + i \cdot e^{\alpha t} \cdot \sin(\beta t).$$

Wir erhalten dann ein Fundamentalsystem dessen Einträge von der Form

$$p(t) \cdot e^{\alpha t} \cdot \cos(\beta t) \quad \text{bzw.} \quad p(t) \cdot e^{\alpha t} \cdot \sin(\beta t)$$

sind. Die Anzahl der Lösungen im Fundamentalsytem zu einem reellen Eigenwert λ ist die algebraische Vielfachheit von λ als Eigenwert, die Anzahl der Lösungen zu einem Paar komplex konjugierter nicht-reeller Eigenwerte ist das doppelte der algebraischen Vielfachheit eines der Eigenwerte.

Bemerkung 32.15 (2-dimensionale lineare Flüsse)
Sei $A = \operatorname{diag}(\lambda, \mu) \in \operatorname{Mat}_2(\mathbb{R})$ eine 2×2-Diagonalmatrix, dann definiert sie den linearen Fluß

$$\varphi : \mathbb{R}^3 \longrightarrow \mathbb{R}^2 : (t, x, y) \mapsto (e^{\lambda t} \cdot x, e^{\mu t} \cdot y).$$

Der Ursprung $(x, y)^t = (0, 0)^t$ ist eine nulldimensionale Bahn, die übrigen Bahnen sind eindimensional. In den Abbildungen 3, 4 und 5 zeigen einige Beispiele für Phasenporträts in Abhängigkeit von λ und μ.

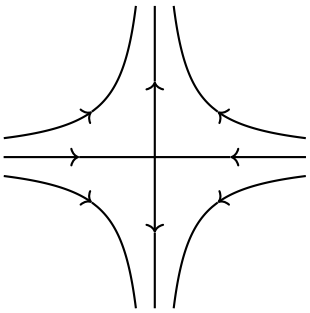


Abbildung 3. Linearer Fluß zu $A = \operatorname{diag}(\lambda, \mu)$ mit $\lambda < 0 < \mu$

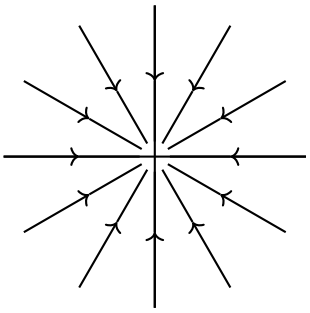


Abbildung 4. Linearer Fluß zu $A = \operatorname{diag}(\lambda, \mu)$ mit $\lambda = \mu < 0$

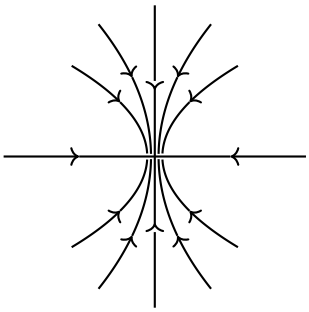


ABBILDUNG 5. Linearer Fluß zu $A = \operatorname{diag}(\lambda, \mu)$ mit $\lambda < \mu < 0$

Hat die Matrix A zwei nicht-reelle Eigenwerte, etwa

$$A = \begin{pmatrix} 0 & -\beta \\ \beta & 0 \end{pmatrix}$$

mit

$$\varphi(t, x, y) = \begin{pmatrix} \cos(\beta t) & -\sin(\beta t) \\ sin(\beta t) & \cos(\beta t) \end{pmatrix} \circ \begin{pmatrix} x \\ y \end{pmatrix}$$

oder

$$A = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$$

mit

$$\varphi(t, x, y) = e^{\alpha t} \cdot \begin{pmatrix} \cos(\beta t) & -\sin(\beta t) \\ \sin(\beta t) & \cos(\beta t) \end{pmatrix} \circ \begin{pmatrix} x \\ y \end{pmatrix}$$

dann können auch periodische Phasenporträts wie in Abbildung 6 oder Spiralen wie in Abbildung 7 als Phasenportäts auftreten. Für weitere Details sei auf [**Ama83**, §13] verwiesen.

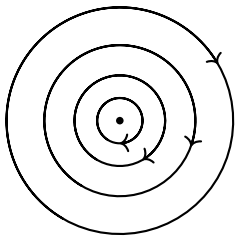


ABBILDUNG 6. Linearer Fluß zu $\sigma(A) = \{\pm i\beta\}$ mit $\beta < 0$

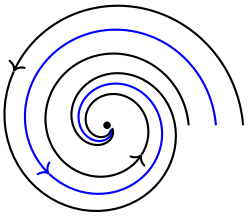


ABBILDUNG 7. Linearer Fluß zu $\sigma(A) = \{\alpha \pm i\beta\}$ mit $\alpha < 0 < \beta$

F) Lineare DGL mit konstanten Koeffizienten - höhere Ordnung

Satz 32.16 (Lineare DGL höherer Ordnung mit konstanten Koeffiz.)
Seien $a_0, \ldots, a_{m-1} \in \mathbb{R}$.

a. Ist $\lambda \in \mathbb{R}$ eine k-fache Nullstelle des *charakteristischen Polynoms*

$$\chi = t^m - a_{m-1} \cdot t^{m-1} - \ldots - a_0,$$

so sind die Funktionen

$$x_i : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto t^i \cdot e^{\lambda t}$$

für $i = 0, \ldots, k-1$ linear unabhängige Lösungen der homogenen linearen gewöhnlichen Differentialgleichung

$$x^{(m)} = a_{m-1} \cdot x^{(m-1)} + \ldots + a_0 \cdot x.$$

b. Sind $\lambda = \alpha + i \cdot \beta \in \mathbb{C} \setminus \mathbb{R}$ und $\overline{\lambda} = \alpha - i \cdot \beta$ zwei komplex konjugierte k-fache Nullstellen des *charakteristischen Polynoms*

$$\chi = t^m - a_{m-1} \cdot t^{m-1} - \ldots - a_0,$$

so sind die Funktionen

$$x_i^c : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto t^i \cdot e^{\alpha t} \cdot \cos(\beta t)$$

und

$$x_i^s : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto t^i \cdot e^{\alpha t} \cdot \sin(\beta t)$$

für $i = 0, \ldots, k-1$ linear unabhängige Lösungen der homogenen linearen gewöhnlichen Differentialgleichung

$$x^{(m)} = a_{m-1} \cdot x^{(m-1)} + \ldots + a_0 \cdot x.$$

c. Führt man das Verfahren in a. und b. für alle Nullstellen des charakteristischen Polynoms der Differentialgleichung durch, so erhält man eine Basis von L_h und damit ein Fundamentalsystem.

Beweis: Überführen wir die Differentialgleichung höherer Ordnung in ein System von Differentialgleichungen erster Ordnung (Bemerkung 32.10), erhalten wir die Matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & 0 & 1 & 0 & \ldots & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & \ldots & 0 & 0 & 1 \\ a_0 & \ldots & \ldots & \ldots & \ldots & a_{m-2} & a_{m-1} \end{pmatrix} \in \mathrm{Mat}_m(\mathbb{R}),$$

deren charakteristisches Polynom das charakterstische Polynom der Differentialgleichung ist (siehe Aufgabe A11.23):

$$\chi_A = \det(t \cdot \mathbb{1}_m - A) = t^m - a_{m-1} \cdot t^{m-1} - \ldots - a_0 = \chi.$$

Ist nun $\lambda \in \mathbb{R}$ eine k-fache Nullstelle von χ, so gibt es nach Bemerkung 32.14 ein Fundamentalsystem von $\dot{x} = Ax$, bei dem zum Eigenwert λ genau k linear unabhängige Lösungen dabei sind, für die die ersten Komponenten von der Form $p_j(t) \cdot e^{\lambda t}$ für Polynome p_j sind, deren Grad kleiner als der Nilpotenzindex des Eigenwertes und damit auch durch $k-1$ beschränkt ist. Dann bilden die Polynome $p_1, \ldots, p_k$ aber eine Basis des Vektorraums der Polynome vom Grad höchstens k und wir können deshalb

$$p_j(t) = t^{j-1}$$

für $j = 1, \ldots, k$ wählen. Damit ist a. gezeigt. Der Teil b. folgt analog mit Bemerkung 32.14, und der Teil c. folgt dann ebenfalls aus Bemerkung 32.14. □

Beispiel 32.17 (Das Hooksche Gesetz)
Betrachten wir eine Feder mit Federkonstante $k > 0$, an der ein Gewicht der Masse m hängt. Ferner beschreibe

$$x : \mathbb{R} \longrightarrow \mathbb{R}$$

die vertikale Auslenkung des Gewichtes aus der Ruhelage im Verlauf der Zeit. Auf das Gewicht wirkt die Kraft $F(x) = -kx$, so daß wir die Bewegungsgleichung (Masse x Beschleunigung = Rückstellkraft)

$$\ddot{x} = -\frac{k}{m} \cdot x \tag{180}$$

erhalten. Daraus ergibt sich das charakteristische Polynom

$$\chi = t^2 + \frac{k}{m},$$

das die zwei rein imaginären Nullstellen $\pm i \cdot \omega$ mit $\omega = \sqrt{\frac{k}{m}}$ besitzt. Dann bilden also

$$x^c : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto \cos(\omega t)$$

und

$$x^s : \mathbb{R} \longrightarrow \mathbb{R} : t \mapsto \sin(\omega t)$$

ein Fundamentalsystem für (180).

Betrachten wir das zu (180) äquivalente zweidimensionale System

$$\begin{pmatrix} \dot{z_1} \\ \dot{z_2} \end{pmatrix} = \dot{z} = \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} \circ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} z_2 \\ -\omega^2 \cdot z_1 \end{pmatrix}$$

mit

$$(z_1, z_2) = (x, \dot{x}),$$

dann hat der zugehörige Fluß das qualitative Verhalten wie in Abbildung 6.

Aufgaben

Aufgabe 32.18

Sei $A : I \longrightarrow \mathrm{Mat}_n(\mathbb{R})$ stetig, I ein offenes Intervall und sei $B \in Gl_n(\mathbb{R})$. Zeige, $X : I \longrightarrow \mathrm{Gl}_n(\mathbb{R})$ ist genau dann ein Fundamentalsytem von $\dot{x} = A(t) \circ x$, wenn

$$I \longrightarrow \mathrm{Gl}_n(\mathbb{R}) : t \mapsto B \circ X(t)$$

ein Fundamentalsystem von $\dot{x} = BA(t)B^{-1} \circ x$ ist.

Aufgabe 32.19

Gegeben seien die Abbildung $b : \mathbb{R} \longrightarrow \mathbb{R}^3 : t \mapsto (0, 0, e^t)^t$ und die Matrix

$$A = \begin{pmatrix} 2 & -2 & 0 \\ 4 & -2 & 0 \\ 3 & -1 & -1 \end{pmatrix} \in \mathrm{Gl}_3(\mathbb{R}).$$

a. Zeige, die folgende Abbildung ist ein Fundamentalsystem von $\dot{x} = Ax$

$$X : \mathbb{R} \longrightarrow \mathrm{Gl}_3(\mathbb{R}) : t \mapsto \begin{pmatrix} 0 & \cos(2t) & \sin(2t) \\ 0 & \cos(2t) + \sin(2t) & -\cos(2t) + \sin(2t) \\ e^t & \sin(2t) & -\cos(2t) \end{pmatrix}.$$

b. Löse das Anfangswertproblem $\dot{x} = Ax + b(t)$ mit $x(\pi/2) = (-1, 1, 1)^t$.

Aufgabe 32.20

Finde ein Beispiel für eine nicht lokal Lipschitz-stetige Abbildung f, so dass das Anfangswertproblem $\dot{x} = f(t, x)$ mit $x(0) = \eta$ für jeden Anfangswert η eine eindeutige globale Lösung besitzt.

Aufgabe 32.21

Bestimme die allgemeine Lösung der Differentialgleichung

$$\dot{x} + cos(t) \cdot x = \frac{sin(2t)}{2}$$

auf dem Intervall $I = (0, \pi)$.

Teil C

Anhang

Kapitel A

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 liegen die Körper der reellen und komplexen Zahlen zugrunde. Wesentlichster Baustein neben der Addition und Multiplikation sind dort der Absolutbetrag mit Werten in $\mathbb{R}$ und die Ordnungsrelation auf $\mathbb{R}$, die $\mathbb{R}$ zu einem vollständigen angeordneten Körper machen. 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 $\mathbb{F}_2$ oder allgemeiner $\mathbb{F}_p$ in Beispiel 7.6.

A1 Grundlegendes zu Gruppen und Körpern

Es ist ein wichtiges Ziel der Mathematik, *Strukturen* auf Mengen zu studieren. Was eine *Struktur auf einer Menge* ganz konkret ist, hängt letztlich sehr vom Zweig der Mathematik und der betrachteten Fragestellung ab. In dieser Vorlesung wird die Struktur stets aus einer oder mehreren *zweistelligen Operationen* auf der Menge bestehen, die dann bestimmten *Gesetzmäßigkeiten*, sogenannten *Axiomen*, genügen sollen. Dabei ist eine *zweistellige Operation* auf einer Menge G eine Abbildung, die einem Paar (g, h) von Elementen aus G wieder ein Element in G zuordnet, also eine Abbildung $G \times G \to G$.

A) Gruppen

Die grundlegendste und wichtigste algebraische Struktur auf einer Menge ist die *Gruppenstruktur*. Im vorliegenden Abschnitt wollen wir einige wesentliche elementare Eigenschaften zusammenstellen, die auch in der *linearen* Algebra benötigt werden.

Definition A1.1 (Gruppen)

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

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

so daß die folgenden *Gruppenaxiome* gelten:

G1: $(g * h) * k = g * (h * k) \quad \forall\, g, h, k \in G,$ ("Assoziativgesetz")

G2: $\exists\, e \in G \;:\; \forall\, g \in G \;:\; e * g = g,$ ("Existenz eines Neutralen")

G3: $\forall\, g \in G\, \exists\, g^{-1} \in G \;:\; g^{-1} * g = e.$ ("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:

G4: $g * h = h * g \quad \forall\, g, h \in G$ ("Kommutativgesetz")

Beispiel A1.2

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. Die einfachste Gruppe ist die *einelementige Gruppe* $G = \{e\}$, deren Gruppenoperation durch $e * e = e$ definiert ist.

d. Ist M eine Menge, so ist die Menge

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

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

Bemerkung A1.3
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 $\left(g^{-1}\right)^{-1} = g$ und $(g * h)^{-1} = h^{-1} * g^{-1}$.

d. Für $g \in G$ und zwei ganze Zahlen $n, m \in \mathbb{Z}$ gelten die Potenzgesetze:

$$g^n * g^m = g^{n+m} \qquad \text{und} \qquad (g^m)^n = g^{m \cdot n}.$$

e. 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 $g - h$.

f. In Ermangelung eines besseren Namens nennen wir auch "$*$" oft einfach die *Gruppenmultiplikation*.

Die Aussagen wollen wir in dieser Vorlesung nicht beweisen, aber wir fügen für den interessierten Leser einen Beweis ein.[1]

Beweis von Bemerkung A1.3: 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.

[1] Die Begriffe *Lemma*, *Proposition*, *Satz* und *Korollar* sind in der Mathematik übliche *Ordnungsstrukturen* (vergleichbar etwa einer Karteikarte), in denen bewiesene Aussagen festgehalten werden. Dabei werden die Aussagen, die als Propositionen formuliert werden, meist als wichtiger erachtet, als Aussagen in einem Lemma, und entsprechend sind Aussagen in einem Satz meist wesentlicher als Aussagen in einer Proposition. Das Korollar fällt etwas aus dem Rahmen, da es übersetzt *Folgerung* bedeutet und somit andeutet, daß es aus einer der unmittelbar zuvor getroffenen Aussagen abgeleitet werden kann. – Es kann vorkommen, daß der Beweis einer Aussage den Rahmen dieser Vorlesung sprengen würde oder wir aus anderen Gründen auf den Beweis verzichten wollen oder müssen. In einem solchen Fall werden wir die Aussage nur als *Bemerkung* formulieren und deutlich machen, weshalb wir auf einen Beweis verzichten.

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

$$(g^{-1})^{-1} * g^{-1} = e. \tag{181}$$

Also folgt:

(182)

$$g * g^{-1} \overset{G2}{=} e * \left(g * g^{-1}\right) \overset{(181)}{=} \left((g^{-1})^{-1} * g^{-1}\right) * \left(g * g^{-1}\right) \overset{G1}{=} (g^{-1})^{-1} * \left(g^{-1} * (g * g^{-1})\right)$$
$$\overset{G1}{=} (g^{-1})^{-1} * \left((g^{-1} * g) * g^{-1}\right) \overset{G3}{=} (g^{-1})^{-1} * \left(e * g^{-1}\right) \overset{G2}{=} (g^{-1})^{-1} * g^{-1} \overset{(181)}{=} e.$$

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

$$g * e \overset{G3}{=} g * \left(g^{-1} * g\right) \overset{G1}{=} \left(g * g^{-1}\right) * g \overset{(182)}{=} e * g \overset{G2}{=} g. \tag{183}$$

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.

$$\tilde{e} * h = h \tag{184}$$

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. (183), mit $\tilde{e}$ in der Rolle von g anwenden, und anschließend (184) mit e in der Rolle von h:

$$\tilde{e} \overset{(183)}{=} \tilde{e} * e \overset{(184)}{=} e.$$

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

$$\tilde{g}^{-1} * g = e \tag{185}$$

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

$$\tilde{g}^{-1} \overset{(183)}{=} \tilde{g}^{-1} * e \overset{(182)}{=} \tilde{g}^{-1} * \left(g * g^{-1}\right) \overset{G1}{=} \left(\tilde{g}^{-1} * g\right) * g^{-1} \overset{(185)}{=} e * g^{-1} \overset{G2}{=} g^{-1}.$$

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:

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

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

$$\left(g^{-1}\right)^{-1} = g.$$

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:

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

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

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

Die Potenzgesetze zeigt man dann mit Hilfe vollständiger Induktion. Wer den Beweis ausgeführt sehen möchte, sei auf [Mar08, Lemma 1.10] verwiesen. □

Lemma A1.4 (Kürzungsregeln)
Sei $(G, *)$ eine Gruppe, $g, a, b \in G$. Dann gelten die *Kürzungsregeln*:

a. Wenn $g * a = g * b$ gilt, dann gilt $a = b$.

b. Wenn $a * g = b * g$ gilt, dann gilt $a = b$.

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

$$\begin{aligned}a \overset{G2}{=} e * a \overset{G3}{=} \left(g^{-1} * g\right) * a &\overset{G1}{=} g^{-1} * (g * a)\\ &\overset{Vor.}{=} g^{-1} * (g * b) \overset{G1}{=} \left(g^{-1} * g\right) * b \overset{G3}{=} e * b \overset{G2}{=} b.\end{aligned}$$

Entsprechend folgt die zweite Kürzungsregel durch Multiplikation mit g^{-1} von rechts und unter Berücksichtigung der zusätzlichen Eigenschaft des Inversen in Bemerkung A1.3. Die Details überlassen wir dem Leser. □

B) Untergruppen

Ein wichtiges Prinzip in der Mathematik ist es, zu jeder betrachteten Struktur auch ihre *Unter-* oder *Teilstrukturen* zu betrachten. Für eine Menge sind das einfach ihre Teilmengen, für eine Gruppe werden es ihre *Untergruppen* sein – das sind Teilmengen, die die zusätzliche Struktur *respektieren*. Eine *Gruppe* besteht aus einer Menge G und zusätzlich einer zweistelligen Operation $\cdot : G \times G \to G$, die gewissen Axiomen genügt. Ist $U \subseteq G$ eine Teilmenge von G, so kann man die Abbildung "$\cdot$" auf $U \times U$ einschränken und erhält eine Abbildung

$$U \times U \longrightarrow G : (u, v) \mapsto u \cdot v,$$

wobei der Ausdruck $u \cdot v$ ein Element aus G ist, in aller Regel aber nicht in U liegt. Letzteres bedeutet, daß die Einschränkung von "$\cdot$" auf $U \times U$ in aller Regel keine

zweistellige Operation auf U definiert! Das ist aber sicher eine Minimalforderung an U um zu sagen, daß U die Gruppenoperation "$\cdot$" respektiert. Nehmen wir nun an, daß wider Erwarten die Einschränkung von "$\cdot$" tatsächlich eine zweistellige Operation auf U definiert, dann stellt sich die Frage, ob das Paar bestehend aus U und der Einschränkung von "$\cdot$" den Gruppenaxiomen G1-G3 genügt, sprich selbst eine Gruppe ist – und erst in letzterem Fall kann man wirklich guten Gewissens sagen, die Teilmenge U respektiere die zusätzliche Struktur. Diese Überlegungen führen zum Begriff der *Untergruppe*, auch wenn dies aus unserer Definition nicht unmittelbar ersichtlich ist. Sie lassen sich im Übrigen auf alle weiteren von uns betrachteten algebraischen Strukturen übertragen.

Definition A1.5

Sei $(G, \cdot)$ eine Gruppe. Eine nicht-leere Teilmenge $U \subseteq G$ heißt *Untergruppe* von G, wenn

$$u \cdot v \in U \quad \text{und} \quad u^{-1} \in U$$

für alle $u, v \in U$. Ist U eine Untergruppe von G, so schreiben wir dafür $U \leq G$.

Bemerkung A1.6

Genau dann ist eine Teilmenge U von G mit der Einschränkung der Gruppenoperation selbst eine Gruppe, wenn U eine Untergruppe von G ist.

Beweis von Bemerkung A1.6: Sei zunächst U mit der Einschränkung der Gruppenoperation selbst eine Gruppe. Dies bedeutet, daß das Bild von $U \times U$ unter der Abbildung "$\cdot$" in U liegt, d. h. für $u, v \in U$ gilt $uv \in U$. Außerdem gelten in U die Gruppenaxiome. Sei also $e_U \in U$ das Neutrale in U und $e_G \in G$ das Neutrale in G. Ferner bezeichne zu $u \in U \subseteq G$ u_G^{-1} das Inverse zu u in G und u_U^{-1} das Inverse zu u in U, d. h. $u_G^{-1}u = uu_G^{-1} = e_G$ und $u_U^{-1}u = uu_U^{-1} = e_U$. In der folgenden Gleichung benötigen wir das Inverse von e_U in der Gruppe G, was in unserer Notation zu dem etwas unübersichtlichen $(e_U)_G^{-1}$ wird. Mit dieser Schreibweise gilt nun:

(186)
$$e_U \overset{G2\,in\,G}{=} e_G e_U \overset{G3\,in\,G}{=} \big((e_U)_G^{-1} e_U\big) e_U \overset{G1\,in\,G}{=} (e_U)_G^{-1}(e_U e_U) \overset{G2\,in\,U}{=} (e_U)_G^{-1} e_U \overset{G3\,in\,G}{=} e_G.$$

Zudem gilt aber

$$u_U^{-1}u \overset{G3\,in\,U}{=} e_U \overset{(186)}{=} e_G,$$

also ist $u_U^{-1} = u_G^{-1}$ wegen der Eindeutigkeit des Inversen in G, und damit $u_G^{-1} \in U$.

Sei nun umgekehrt vorausgesetzt, daß U eine Untergruppe von G ist. Da $uv \in U$ für alle $u, v \in U$, ist das Bild von $U \times U$ unter der Abbildung "$\cdot$" in der Tat in U enthalten. Es bleibt also, die Axiome G1-G3 nachzuprüfen. Dabei überträgt sich G1 von der größeren Menge G auf die Teilmenge U. Da $U \neq \emptyset$, existiert ein $u \in U$. Nach Voraussetzung gilt dann aber $u_G^{-1} \in U$ und damit $e_G = u_G^{-1}u \in U$. Da aber $e_G u = u$ für alle $u \in U$, ist auch G2 erfüllt und es gilt $e_U = e_G$. Ferner haben wir bereits bemerkt, daß für $u \in U$

auch $u_G^{-1} \in U$, und es gilt

$$u_G^{-1} \cdot u = e_G = e_U.$$

Somit ist auch G3 erfüllt und die Inversen von u in U bzw. in G stimmen überein. □

Man interessiert sich auch deshalb für die Untergruppen einer Gruppe, weil die Kenntnis dieser wichtige Informationen über die Struktur der Gruppe selbst liefert.

Beispiel A1.7

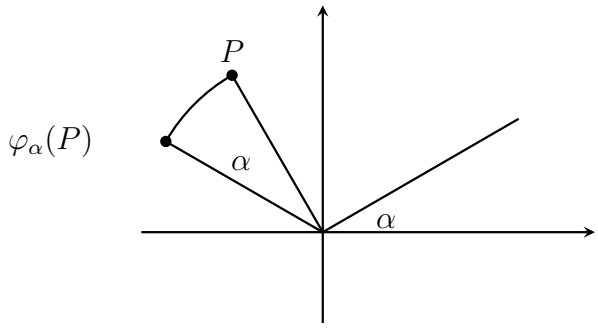
a. Ist $(G, \cdot)$ eine Gruppe mit neutralem Element e_G, so sind die beiden Teilmengen $\{e_G\}$ und G von G stets Untergruppen. Man nennt sie deshalb auch die *trivialen Untergruppen*. Sie geben keine zusätzliche Information über die Struktur der Gruppe selbst.

b. $(\{-1, 1\}, \cdot)$ ist eine Untergruppe von $(\mathbb{Q} \setminus \{0\}, \cdot)$.

c. Für $\alpha \in \mathbb{R}$ bezeichnet

$$\varphi_\alpha : \mathbb{R}^2 \to \mathbb{R}^2 : (x, y) \to \big(\cos(\alpha) \cdot x - \sin(\alpha) \cdot y, \sin(\alpha) \cdot x + \cos(\alpha) \cdot y\big)$$

die Drehung der Ebene $\mathbb{R}^2$ um den Nullpunkt um den Winkel α im Bogenmaß.

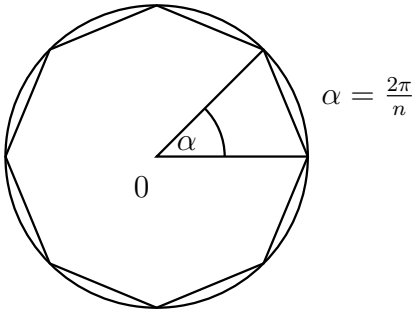


Offensichtlich gilt $\varphi_\alpha \circ \varphi_\beta = \varphi_{\alpha+\beta}$ für $\alpha, \beta \in \mathbb{R}$, und für $\alpha \in \mathbb{R}$ ist somit $\varphi_{-\alpha} = (\varphi_\alpha)^{-1}$, da $\varphi_0 = \mathrm{id}_{\mathbb{R}^2}$. Insbesondere ist φ_α also bijektiv für jedes $\alpha \in \mathbb{R}$. Also ist

$$\mathrm{SO}(2) := \{\varphi_\alpha : \mathbb{R}^2 \to \mathbb{R}^2 \mid \alpha \in \mathbb{R}\}$$

eine Untergruppe von $\mathrm{Sym}\,(\mathbb{R}^2)$.

d. Sei $E_n \subset \mathbb{R}^2$ das *reguläre n-Eck*.



Die Menge

$$U := \{\varphi_\alpha \in \mathrm{SO}(2) \mid \varphi_\alpha(E_n) = E_n\}$$

ist dann eine Untergruppe von $\big(\mathrm{SO}(2), \circ\big)$.
Denn für $\varphi_\alpha, \varphi_\beta \in U$ gilt

$$(\varphi_\alpha \circ \varphi_\beta)(E_n) = \varphi_\alpha\big(\varphi_\beta(E_n)\big) = \varphi_\alpha(E_n) = E_n$$

und

$$\varphi_\alpha^{-1}(E_n) = \varphi_\alpha^{-1}\big(\varphi_\alpha(E_n)\big) = \big(\varphi_\alpha^{-1} \circ \varphi_\alpha\big)(E_n) = \mathrm{id}_{\mathbb{R}^2}(E_n) = E_n.$$

Also gilt $\varphi_\alpha \circ \varphi_\beta \in U$ und $\varphi_\alpha^{-1} \in U$, und da $\mathrm{id}_{\mathbb{R}^2} = \varphi_0 \in U$, ist $U \neq \emptyset$ und folglich ist U eine Untergruppe von $\mathrm{SO}(2)$.
Man überzeugt sich leicht, daß U aus allen Drehungen φ_α mit $\alpha = k \cdot \frac{2\pi}{n}$, $k = 0, \ldots, n-1$, besteht. Insbesondere gilt also, $|U| = n$.

e. Die Untergruppen von $(\mathbb{Z}, +)$ sind genau die Mengen $n\mathbb{Z} := \{nz \mid z \in \mathbb{Z}\}$ aller Vielfachen von n für $n \in \mathbb{N}$, siehe Aufgabe A1.33.

f. Die Inklusionen $\mathbb{Z} \subset \mathbb{Q}$, $\mathbb{Z} \subset \mathbb{R}$ und $\mathbb{Q} \subset \mathbb{R}$ machen die Teilmenge bezüglich der Addition als Gruppenstruktur jeweils zu einer Untergruppe.

C) Gruppenhomomorphismen

Immer wenn man eine Struktur auf einer Menge definiert hat, spielen die *strukturerhaltenden Abbildungen* eine besondere Rolle. Diese werden (Struktur-)*Morphismen* oder (Struktur-)*Homomorphismen* genannt.

Definition A1.8
Es seien $(G, \cdot)$ und $(H, *)$ zwei Gruppen. Eine Abbildung $\alpha : G \to H$ heißt *Gruppenhomomorphismus* (oder kürzer *Homomorphismus* oder nur *Morphismus*), falls für alle $g, h \in G$ gilt:

$$\alpha(g \cdot h) = \alpha(g) * \alpha(h).$$

Wieder wollen wir uns zunächst Beispiele anschauen.

Beispiel A1.9

a. Ist $(G, \cdot)$ eine Gruppe und $U \leq G$ eine Untergruppe, dann ist die kanonische Inklusion $i_U : U \to G$ ein Gruppenhomomorphismus, da für $g, h \in U$ gilt $i_U(g \cdot h) = g \cdot h = i_U(g) \cdot i_U(h)$.

b. Sei $a \in \mathbb{R}$ und $m_a : (\mathbb{R}, +) \to (\mathbb{R}, +) : g \mapsto ag$ die Multiplikation mit a, dann ist m_a ein Gruppenhomomorphismus, da für $g, h \in \mathbb{R}$ gilt

$$m_a(g + h) = a(g + h) = ag + ah = m_a(g) + m_a(h).$$

c. Ist $(G,\cdot)$ eine Gruppe und $g \in G$, so definiert

$$i_g : G \to G : h \mapsto ghg^{-1} =: h^g$$

einen Gruppenhomomorphismus, die sogenannte *Konjugation* mit g, denn für $h,k \in G$ gilt

$$\begin{aligned} i_g(hk) =& g(hk)g^{-1} = g\big(hek\big)g^{-1} = g\big(h\big(g^{-1}g\big)k\big)g^{-1} \\ =& \big(ghg^{-1}\big)\big(gkg^{-1}\big) = i_g(h)\cdot i_g(k), \end{aligned}$$

Das folgende Lemma sagt, daß die Komposition zweier Homomorphismen stets wieder ein Homomorphismus ist.

Lemma A1.10
Sind $\alpha_1 : (G_1,\cdot) \to (G_2,*)$ und $\alpha_2 : (G_2,*) \to (G_3,\times)$ Gruppenhomomorphismen, so ist auch $\alpha_2 \circ \alpha_1 : (G_1,\cdot) \to (G_3,\times)$ ein Gruppenhomomorphismus.

Beweis: Seien $g,h \in G_1$, dann gilt:

$$\begin{aligned} (\alpha_2\circ\alpha_1)(g\cdot h) &= \alpha_2\big(\alpha_1(g\cdot h)\big) = \alpha_2\big(\alpha_1(g) * \alpha_1(h)\big) = \alpha_2\big(\alpha_1(g)\big) \times \alpha_2\big(\alpha_1(h)\big) \\ &= (\alpha_2\circ\alpha_1)(g) \times (\alpha_2\circ\alpha_1)(h). \end{aligned}$$

□

Der Umstand, daß die Gruppenhomomorphismen die Gruppenstruktur *erhalten*, hat einige einfache, aber ungemein wichtige Auswirkungen.

Proposition A1.11
Es sei $\alpha : (G,\cdot) \to (H,*)$ ein Gruppenhomomorphismus. Dann gelten:

a. $\alpha(e_G) = e_H$.

b. $\alpha\big(g^{-1}\big) = \big(\alpha(g)\big)^{-1}$ für $g \in G$.

c. Ist α bijektiv, so ist $\alpha^{-1} : H \to G$ ein Gruppenhomomorphismus.

d. Ist $U \leq G$, dann ist $\alpha(U) \leq H$. $\alpha(U)$ heißt das *Bild* von U unter α.

e. Ist $V \leq H$, dann ist $\alpha^{-1}(V) \leq G$. $\alpha^{-1}(V)$ heißt das *Urbild* von V unter α.

f. $\mathrm{Im}(\alpha) := \alpha(G)$, das *Bild* von α, ist eine Untergruppe von H.

g. $\mathrm{Ker}(\alpha) := \alpha^{-1}(e_H)$, der *Kern* von α, ist eine Untergruppe von G.

Beweis: a. Es gilt

$$e_H * \alpha(e_G) = \alpha(e_G) = \alpha(e_G\cdot e_G) = \alpha(e_G) * \alpha(e_G).$$

Mit Hilfe der Kürzungsregel 7.4 folgt dann $e_H = \alpha(e_G)$.

b. Für $g \in G$ gilt:

$$\alpha\big(g^{-1}\big) * \alpha(g) = \alpha\big(g^{-1} \cdot g\big) = \alpha(e_G) = e_H.$$

Wegen der Eindeutigkeit der Inversen in H folgt die Behauptung.

c. Ist $\alpha : G \to H$ bijektiv, so existiert die Umkehrabbildung $\alpha^{-1} : H \to G$. Seien $u, v \in H$. Setze $g := \alpha^{-1}(u)$ und $h := \alpha^{-1}(v)$, also $u = \alpha(g)$ und $v = \alpha(h)$. Dann gilt:

$$\alpha^{-1}(u * v) = \alpha^{-1}\big(\alpha(g) * \alpha(h)\big) = \alpha^{-1}\big(\alpha(g \cdot h)\big) = g \cdot h = \alpha^{-1}(u) \cdot \alpha^{-1}(v).$$

Also ist α^{-1} ein Gruppenhomomorphismus.

d. Sind $u, v \in \alpha(U)$, dann existieren $g, h \in U$ mit $\alpha(g) = u$ und $\alpha(h) = v$. Da $g \cdot h \in U$, gilt:

$$u * v = \alpha(g) * \alpha(h) = \alpha(g \cdot h) \in \alpha(U).$$

Außerdem gilt $g^{-1} \in U$ und somit:

$$u^{-1} = \big(\alpha(g)\big)^{-1} = \alpha\big(g^{-1}\big) \in \alpha(U).$$

Da zudem $\alpha(e_G) \in \alpha(U)$, also $\alpha(U) \neq \emptyset$, folgt, daß $\alpha(U)$ eine Untergruppe von H ist.

e. Seien $g, h \in \alpha^{-1}(V)$, so gilt $\alpha(g \cdot h) = \alpha(g) * \alpha(h) \in V$, da V eine Untergruppe ist. Also gilt $g \cdot h \in \alpha^{-1}(V)$. Außerdem gilt $\alpha\big(g^{-1}\big) = \big(\alpha(g)\big)^{-1} \in V$, wieder da V eine Untergruppe ist. Somit liegt auch g^{-1} in $\alpha^{-1}(V)$. Da das Urbild von V unter α ferner nicht leer ist, weil wegen $\alpha(e_G) = e_H \in V$ gilt, daß $e_G \in \alpha^{-1}(V)$, folgt wieder, daß $\alpha^{-1}(V)$ eine Untergruppe von G ist.

f. Dies folgt aus d., da G eine Untergruppe von G ist.

g. Dies folgt aus e., da $\{e_H\}$ eine Untergruppe von H ist.

□

Nach Definition muß man für die Injektivität einer Abbildung nachprüfen, daß jedes Element im Bild nur ein Urbild hat. Bei Gruppenhomomorphismen gibt es ein einfacheres Kriterium.

Lemma A1.12
Ein Gruppenhomomorphismus $\alpha : (G, \cdot) \to (H, *)$ ist genau dann injektiv, wenn $\mathrm{Ker}(\alpha) = \{e_G\}$.

Beweis: Ist α injektiv, so ist $\alpha^{-1}(e_H)$ höchstens einelementig, und wegen $\alpha(e_G) = e_H$ gilt dann $\mathrm{Ker}(\alpha) = \alpha^{-1}(e_H) = \{e_G\}$.

Gilt umgekehrt $\mathrm{Ker}(\alpha) = \{e_G\}$, und sind $g, h \in G$ mit $\alpha(g) = \alpha(h)$, so folgt wegen:

$$e_H = \alpha(g) * \big(\alpha(h)\big)^{-1} = \alpha(g) * \alpha\big(h^{-1}\big) = \alpha\big(g \cdot h^{-1}\big),$$

daß $g \cdot h^{-1} = e_G$, also $g = h$. Somit ist α injektiv. □

D) Körper

Als nächstes wollen wir eine etwas komplexere algebraische Struktur mit zwei zweistelligen Operationen $+$ und $\cdot$ einführen, die grundlegend für die Theorie der Vektorräume ist und die die Eigenschaften der rationalen Zahlen verallgemeinert.

Definition A1.13 (Körper)
Ein *Körper* ist ein Tripel $(K, +, \cdot)$ bestehend aus einer Menge K zusammen mit zwei zweistelligen Operationen

$$+ : K \times K \to K : (x, y) \mapsto x + y, \qquad (\text{“}\textit{Addition}\text{”})$$

und

$$\cdot : K \times K \to K : (x, y) \mapsto x \cdot y, \qquad (\text{“}\textit{Multiplikation}\text{”})$$

so daß folgende Axiome erfüllt sind:

a. $(K, +)$ ist eine abelsche Gruppe mit neutralem Element 0.

b. $(K \setminus \{0\}, \cdot)$ ist eine abelsche Gruppe mit neutralem Element 1.

c. 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 A1.14 (Die endlichen Körper $\mathbb{F}_p$)

a. 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 $\mathbb{R}$.

b. Die ganzen Zahlen $(\mathbb{Z}, +, \cdot)$ sind kein Körper, da z.B. der Zahl 2 ein multiplikatives Inverses fehlt.

c. Auf der Menge $\mathbb{F}_2 := \{0, 1\}$ definieren wir zwei Operationen durch folgende Additions- und Multiplikationstafeln:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

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 Verknüpfungstafeln besteht, wenn man einen Körper mit genau zwei Elementen haben möchte. — Beachte auch, daß $\mathbb{F}_2$ kein Teilkörper von $\mathbb{R}$ ist, da das Ergebnis von $1+1$ in den beiden Körpern nicht übereinstimmt.

d. Allgemeiner zeigt man, daß man für eine Primzahl p die Menge

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

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 "$\cdot$" auf der rechten Seite jeweils die Operation in den ganzen Zahlen bezeichnet, während das "$+$" und das "$\cdot$" 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

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

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 weiterführenden Vorlesungen, 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

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

Man schreibt übrigens meist $\mathbb{Z}/p\mathbb{Z}$ oder $\mathbb{Z}_p$ anstatt $\mathbb{F}_p$, und die Zahl a wird meist mit $\overline{a}$ bezeichnet. Man kann den Körper mit der Menge der Äquivalenzklassen der Kongruenz modulo p identifizieren kann (siehe Aufgabe 6.16).

Notation A1.15

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 A1.16 (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. $\left(x^{-1}\right)^{-1} = x$, 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 aus Bemerkung A1.3 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.:

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

g. Für $x, y, z \in K$ impliziert e.:

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

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 wie die Kürzungsregeln für Gruppen (siehe Lemma 7.4).

k. Unter Beachtung der Assoziativität und Kommutativität der Multiplikation sowie der Notation A1.15 gilt

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

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

□

E) Der Körper der komplexen Zahlen

In der Vorlesung zur Analysis wird der Aufbau des Zahlsystems

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$$

thematisiert, wie man etwa die natürlichen Zahlen axiomatisch definieren und aus diesen die ganzen und dann die rationalen Zahlen mit Hilfe von Äquivalenzrelationen gewinnen kann, um den Problemen zu begegnen, beliebige Zahlen subtrahieren und durch beliebige Zahlen ungleich null dividieren zu können. Daß auch die rationalen Zahlen für die Lösung von Gleichungen höheren Grades nicht ausreichen , etwa

$$x^2 = 2,$$

haben wir schon in der Schule gelernt und als Motivation für die Einführung der reellen Zahlen genommen. Ebenso motiviert der Umstand, daß die Gleichung

$$x^2 = -1$$

in den reellen Zahlen immer noch keine Lösung besitzt, die erneute Erweiterung des Zahlsystems zum Körper der komplexen Zahlen. Der Fundamentalsatz der Algebra A10.17 zeigt, daß wir damit dann aber auch fertig sind und jede polynomiale Gleichung über $\mathbb{C}$ vollständig lösbar ist.

Bemerkung A1.17 (Konstruktion der komplexen Zahlen)

Es ist unser erklärtes Ziel, auf der reellen Zahlenebene $\mathbb{R}^2$ mit der Vektoraddtion

$$(x, y) + (u, v) := (x + u, y + v)$$

eine *Multiplikation* zu definieren, so daß die üblichen Rechenregeln (Assoziativgesetze, Kommutativgesetze, Distributivgesetze) gelten und daß zudem der Vektor

$$i := (0, 1)$$

eine Lösung der Gleichung

$$z^2 = -1$$

ist. Um letzteres richtig zu interpretieren, denken wir uns die reelle Zahlengerade $\mathbb{R}$ als Teilmenge von $\mathbb{R}^2$, indem wir sie mit der x-Achse identifizieren, d.h.

$$\mathbb{R} \mathrel{\hat{=}} \{(a,0) \mid a \in \mathbb{R}\} = x\text{-Achse}.$$

Die Multiplikation soll also der Bedingung

$$i^2 = (0,1) \cdot (0,1) = (-1,0) \mathrel{\hat{=}} -1$$

genügen. Außerdem würden wir uns sicher wünschen, daß die Multiplikation eines Vektors mit der reellen Zahl

$$a \mathrel{\hat{=}} (a,0)$$

wie die Streckung des Vektors um den Faktor a funktioniert, d.h.

$$(a,0) \cdot (x,y) \mathrel{\hat{=}} a \cdot (x,y) = (ax, ay).$$

Wenn eine Multiplikation diese Wunschliste erfüllt, so gilt offenbar:

$$\begin{aligned}
(x,y) \cdot (u,v) &= \big((x,0) + (0,y)\big) \cdot \big((u,0) + (0,v)\big) \\
&= \big((x,0) + (y,0) \cdot (0,1)\big) \cdot \big((u,0) + (v,0) \cdot (0,1)\big) \\
&= (x,0) \cdot (u,0) + (y,0) \cdot (0,1) \cdot (u,0) + (x,0) \cdot (v,0) \cdot (0,1) \\
&\qquad + (y,0) \cdot (0,1) \cdot (v,0) \cdot (0,1) \\
&= (xu,0) + (yu,0) \cdot (0,1) + (xv,0) \cdot (0,1) + (yv,0) \cdot (0,1) \cdot (0,1) \\
&= (xu,0) + (yu,0) \cdot (0,1) + (xv,0) \cdot (0,1) + (yv,0) \cdot (-1,0) \\
&= (xu,0) + (0,yu) + (0,xv) + (-yv,0) = (xu - yv, xv + yu).
\end{aligned}$$

Wir haben für die Definition der Multiplikation also nur *eine einzige* Möglichkeit, und die funktioniert zum Glück auch.

Satz A1.18 (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

$$(x,y) \cdot (u,v) := (xu - yv, xv + yu), \quad \text{für } (x,y),(u,v) \in \mathbb{C},$$

definierten Addition und Multiplikation ist ein Körper, den wir den Körper der *komplexen Zahlen* nennen. .

Beweis: Dies folgt aus Aufgabe A1.42. □

Bemerkung A1.19

a. Daß $\mathbb{C}$ mit den beiden Operationen ein *Körper* ist, bedeutet, daß die oben erwähnten üblichen Rechenregeln bezüglich der Addition, Subtraktion, Multiplikation und Division gelten, so wie wir sie von den reellen Zahlen her kennen. Man beachte dabei, daß die reelle Zahl $0 \mathrel{\hat{=}} (0,0)$ bei der Addtion nichts tut und die reelle Zahl $1 \mathrel{\hat{=}} (1,0)$ bei der Multiplikation ohne Wirkung ist:

$$(x,y) + (0,0) = (x+0, y+0) = (x,y)$$

und

$$(x,y) \cdot (1,0) = (x \cdot 1 - y \cdot 0, x \cdot 0 + y \cdot 1) = (x,y).$$

Das multiplikative Inverse der Zahl $(0,0) \neq (x,y) \in \mathbb{C}$ ist

$$(x,y)^{-1} = \left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2} \right).$$

b. Die Abbildung

$$\iota : \mathbb{R} \longrightarrow \mathbb{C} : x \mapsto (x,0)$$

ist mit der Addition und der Multiplikation verträglich und identifiziert den Körper der reellen Zahlen $\mathbb{R}$ mit dem Teilkörper $\mathbb{R} \times \{0\}$ von $\mathbb{C}$. Wir fassen $\mathbb{R}$ in diesem Sinne als Teilmenge von $\mathbb{C}$ auf.

c. 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 $z = (x,y) \in \mathbb{C}$

$$z = (x,y) = (x,0) + (0,y) = (x,0) + (0,1) \cdot (y,0) \mathrel{\hat{=}} x + iy.$$

Diese Schreibweise wollen wir künftig für komplexe Zahlen verwenden. Damit gilt dann:

$$i^2 = (0,1) \cdot (0,1) = -1.$$

Ferner ergibt sich die etwas willkürlich anmutende Definition der Multiplikation ganz "natürlich" aus

$$(x+iy)(u+iv) = \big(xu + i^2 yv\big) + i(xv+yu) = (xu - yv) + i(xv+yu).$$

Wir wollen nun noch einige nützliche Operationen auf dem Körper der komplexen Zahlen einführen, die wir in unserer Vorlesung nicht oft benötigen werden, die aber für das Rechnen mit komplexen Zahlen durchaus wichtig sind und die wir deshalb nicht unerwähnt lassen wollen.

Definition A1.20 (Der Betrag und die komplexe Konjugation)

a. Wir definieren die *Betragsfunktion* auf $\mathbb{C}$ durch

$$|\cdot| : \mathbb{C} \longrightarrow \mathbb{R}_{\geq 0} : x + iy \mapsto \sqrt{x^2 + y^2}$$

und nennen $|x|$ auch den *Absolutbetrag* von x.
Beachte zudem, für $x \in \mathbb{R}$ gilt

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

b. Wir definieren die *komplexe Konjugation* als

$$\overline{\cdot} : \mathbb{C} \longrightarrow \mathbb{C} : z = x + iy \mapsto \overline{z} := x - iy.$$

Für $z \in \mathbb{C}$ heißt $\overline{z}$ die zu z *konjugiert komplexe Zahl*.

c. Wir definieren die Abbildungen *Realteil*

$$\mathrm{Re} : \mathbb{C} \longrightarrow \mathbb{R} : x + iy \mapsto x$$

und *Imaginärteil*

$$\mathrm{Im} : \mathbb{C} \longrightarrow \mathbb{R} : x + iy \mapsto y$$

und nennen $\mathrm{Re}(x + iy) = x$ den *Realteil* von z und $\mathrm{Im}(x + iy) = y$ den *Imaginärteil* von z.

Beispiel A1.21

Wir betrachten die komplexe Zahl

$$z = i - 1 = -1 + i.$$

Dann gilt

$$\mathrm{Re}(z) = -1,$$

sowie

$$\mathrm{Im}(z) = 1$$

und

$$\overline{z} = -1 - i = -(1 + i).$$

Für den Betrag von z rechnen wir

$$|z| = \sqrt{\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2} = \sqrt{1 + 1} = \sqrt{2}$$

und damit erhalten wir die Gleichung

$$z \cdot \overline{z} = (-1 + i) \cdot (-1 - i) = 2 = |z|^2.$$

Lemma A1.22 (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. $z \cdot \overline{z} = |z|^2$.

e. Wenn $z \neq 0$, dann ist $z^{-1} = \frac{1}{z} = \frac{\overline{z}}{|z|^2}$.

f. Die komplexe Konjugation ist additiv und multiplikativ, d.h.
$$\overline{z} + \overline{w} = \overline{z + w} \qquad \text{und} \qquad \overline{z} \cdot \overline{w} = \overline{z \cdot w}.$$

g. $\overline{\overline{z}} = z$.

h. $\overline{z} = z \iff z \in \mathbb{R}$.

Beweis: Den Beweis der Aussagen in den Teilen a. und c.-h. überlassen wir dem Leser als Übungsaufgabe.

a. Wir wollen nun die Dreiecksungleichung unter Verwendung der übrigen Aussagen zeigen. Es gilt

$$\begin{aligned}
|z + w|^2 &\overset{d.}{=} (z + w) \cdot \overline{(z + w)} \\
&\overset{f.}{=} z \cdot \overline{z} + (z \cdot \overline{w} + \overline{z} \cdot w) + w \cdot \overline{w} \\
&\overset{d.,g.}{=} |z|^2 + \left(z \cdot \overline{w} + \overline{z \cdot \overline{w}}\right) + |w|^2 \\
&= |z|^2 + 2 \cdot \mathrm{Re}(z \cdot \overline{w}) + |w|^2 \\
&\leq |z|^2 + 2 \cdot |z \cdot \overline{w}| + |w|^2 \\
&\overset{a.}{=} |z|^2 + 2 \cdot |z| \cdot |\overline{w}| + |w|^2 \\
&= |z|^2 + 2 \cdot |z| \cdot |w| + |w|^2 \\
&= \left(|z| + |w|\right)^2.
\end{aligned}$$

Da dies eine Ungleichung von nicht-negativen Zahlen in dem angeordneten Körper $\mathbb{R}$ ist, folgt
$$|z + w| \leq |z| + |w|.$$

□

Beispiel A1.23

a. Gegeben seien $z = 3 + 2i$ und $w = 5 - i$. Dann gelten

$$z \cdot w = (3 \cdot 5 - 2 \cdot (-1)) + (3 \cdot (-1) + 2 \cdot 5) \cdot i = 17 + 7i$$

sowie

$$|w| = \sqrt{5^2 + (-1)^2} = \sqrt{26}$$

und

$$\begin{aligned} \frac{z}{w} =& z \cdot \frac{\overline{w}}{|w|^2} = (3 + 2i) \cdot \left(\frac{5}{26} + \frac{1}{26} \cdot i \right) \\ =& \left(3 \cdot \frac{5}{26} - 2 \cdot \frac{1}{26} \right) + \left(3 \cdot \frac{1}{26} + 2 \cdot \frac{5}{26} \right) \cdot i \\ =& \frac{13}{26} + \frac{13}{26} \cdot i = \frac{1}{2} + \frac{1}{2} \cdot i. \end{aligned}$$

b. Für die komplexen Zahlen $z = 3 + 4i$ und $w = 5 - 12i$ gilt

$$z + w = (3 + 5) + (4 - 12) \cdot i = 8 - 8i$$

und somit

$$\begin{aligned} |z + w| =& \sqrt{8^2 + 8^2} = \sqrt{2} \cdot 8 < 16 < 18 = 5 + 13 \\ =& \sqrt{25} + \sqrt{169} = \sqrt{3^2 + 4^2} + \sqrt{5^2 + 12^2} = |z| + |w|. \end{aligned}$$

Außerdem gilt

$$\frac{z + \overline{z}}{2} = \frac{(3 + 4i) + (3 - 4i)}{2} = \frac{6}{2} = 3 = \mathrm{Re}(z).$$

Bemerkung A1.24 (Geometrische Deutung und Polarkoordinaten)

Wir wollen hier einige der bisher eingeführten Operationen auf den komplexen Zahlen und der angeführten Eigenschaften derselben geometrisch interpretieren.

- Die Addition ist einfach die komponentenweise Addition, also die Addition der Vektoren (siehe Abbildung 1).

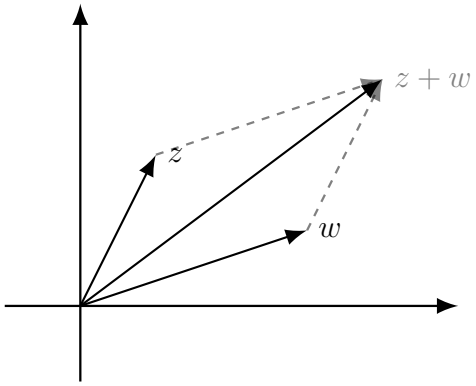


Abbildung 1: Addition in $\mathbb{C}$ als Vektoraddition

- Die komplexe Konjugation ist die Spiegelung an der x-Achse.

- Der Realteil ist die orthogonale Projektion auf die x-Achse und der Imaginärteil die orthogonale Projektion auf die y-Achse.
- Der Betrag $|z| = \sqrt{x^2 + y^2}$ einer komplexen Zahl $z = x + iy$ ist die euklidische Länge des Vektors z, d.h. der Abstand von z zum Ursprung. Dies ergibt sich unmittelbar aus dem Satz von Pythagoras (siehe Abbildung 2).

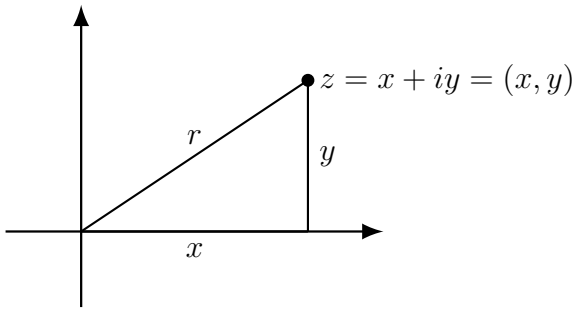


Abbildung 2: Pythagoras: $x^2 + y^2 = r^2$

- 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.
- Damit ist die Menge

$$K := \{z \in \mathbb{C} \mid |z| = 1\}$$

die Menge der Punkte in der Ebene, deren Abstand zum Ursprung genau 1 ist, d.h. K ist der Einheitskreis um den Ursprung. Man beachte, daß bei einem Punkt

$$z = x + iy,$$

der auf dem Einheitskreis liegt, die kartesichen Koordinaten x und y schon vollständig durch den Winkel $\alpha \in [0, 2\pi)$ bestimmt sind, den der Vektor z mit der x-Achse einschließt. Es gilt nämlich (siehe Abbildung 3)

$$x = \cos(\alpha)$$

und

$$y = \sin(\alpha)$$

und somit

$$z = \cos(\alpha) + i \cdot \sin(\alpha).$$

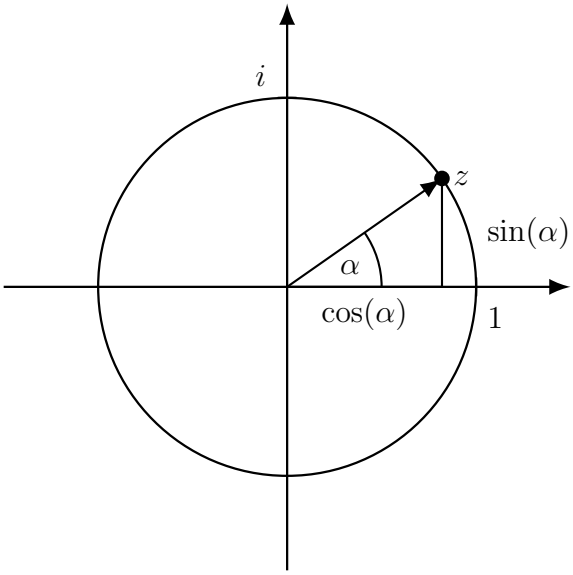


Abbildung 3: Koordinaten eines Punktes $z = \cos(\alpha) + i \cdot \sin(\alpha)$ auf dem Einheitskreis

- Es bleibt, die Multiplikation zweier komplexer Zahlen $0 \neq z, w \in \mathbb{C}$ geometrisch zu deuten. Dazu schreiben wir die Zahl z als

$$z = |z| \cdot \frac{z}{|z|} = r \cdot z'$$

mit $r = |z|$ und $z' = \frac{z}{|z|}$. Man beachte, daß die Zahl z' den Betrag 1 hat, so daß es genau einen Winkel $\alpha \in [0, 2\pi)$ gibt mit

$$z' = (\cos(\alpha), \sin(\alpha)) = \cos(\alpha) + i \cdot \sin(\alpha).$$

Die komplexe Zahl $z \neq 0$ ist also eindeutig durch ihren Betrag und den Winkel α bestimmt. Wir nennen

$$\arg(z) := \alpha = \arctan\left(\frac{\mathrm{Im}(z)}{\mathrm{Re}(z)}\right)$$

das *Argument* von z und das Paar

$$(r, \alpha) = \big(|z|, \arg(z)\big)$$

die *Polarkoordinaten* von z.

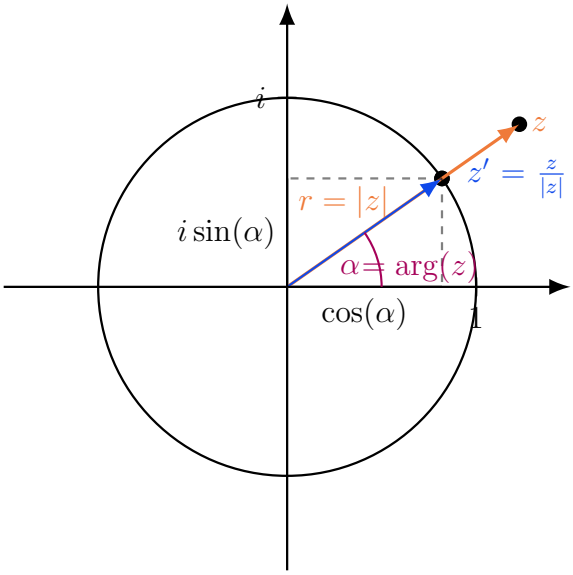


Abbildung 4: Polarkoordinaten von $z = r \cdot (\cos(\alpha) + i \cdot \sin(\alpha))$

Wir erinnern hier an die beiden Additionstheoreme für den Sinus und den Cosinus:

$$\cos(\alpha + \beta) = \cos(\alpha) \cdot \cos(\beta) - \sin(\alpha) \cdot \sin(\beta) \tag{187}$$

und

$$\sin(\alpha + \beta) = \cos(\alpha) \cdot \sin(\beta) + \sin(\alpha) \cdot \cos(\beta). \tag{188}$$

Betrachten wir zunächst die Multiplikation von zwei komplexen Zahlen $z = |z| \cdot \big(\cos(\alpha) + i \cdot \sin(\alpha)\big)$ und $w = |w| \cdot \big(\cos(\beta) + i \cdot \sin(\beta)\big)$:

$$\begin{aligned} z \cdot w &= |z| \cdot |w| \cdot \big(\cos(\alpha) + i \cdot \sin(\alpha)\big) \cdot \big(\cos(\beta) + i \cdot \sin(\beta)\big) \\ &= |z| \cdot |w| \cdot \big(\cos(\alpha) \cdot \cos(\beta) - \sin(\alpha) \cdot \sin(\beta)\big) + i \cdot \big(\cos(\alpha) \cdot \sin(\beta) + \sin(\alpha) \cdot \cos(\beta)\big) \\ &\overset{(187),(188)}{=} |z| \cdot |w| \cdot \big(\cos(\alpha + \beta) + i \cdot \sin(\alpha + \beta)\big). \end{aligned}$$

Die beiden Zahlen werden also multipliziert, indem man die Argumente addiert und die Beträge multipliziert (siehe Abbildung 5).

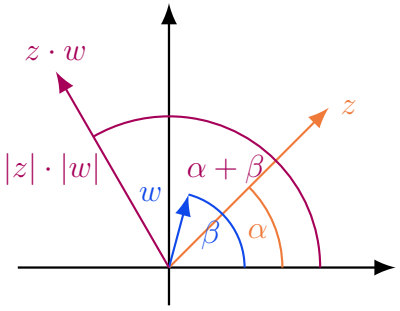


Abbildung 5: Multiplikation zweier komplexer Zahlen

In Polarkoordinaten könnte man dies schreiben als

$$(r, \alpha) \cdot (s, \beta) = (r \cdot s, \alpha + \beta).$$

Beispiel A1.25

Zur Ermittlung von $\alpha = \arg(z)$ für $z = i - 1$ betrachten wir die Zahl

$$\frac{z}{|z|} = -\frac{\sqrt{2}}{2} + \frac{\sqrt{2}}{2} i$$

vom Betrag 1, für die gilt

$$\frac{z}{|z|} = \cos(\alpha) + i \sin(\alpha),$$

d.h. $\cos(\alpha) = -\frac{\sqrt{2}}{2}$ und $\sin(\alpha) = \frac{\sqrt{2}}{2}$, also $\alpha = \frac{3}{4}\pi$.

Aufgaben

Aufgabe A1.26

Untersuche, ob die folgende zweistellige Operation die Menge $G := \mathbb{Q} \times \mathbb{Q}$ zu einer Gruppe macht:

$$G \times G \longrightarrow G : \big((a,b),(a',b')\big) \mapsto (a,b) \cdot (a',b') := (aa', bb').$$

Aufgabe A1.27

Es seien $(G, \cdot)$ und $(H, *)$ zwei Gruppen. Wir definieren auf der Menge $G \times H = \{(x,y) \mid x \in G, y \in H\}$ eine zweistellige Operation durch

$$(x,y) \circ (x',y') := (x \cdot x', y * y')$$

für $(x,y),(x',y') \in G \times H$. Zeige, daß dann $(G \times H, \circ)$ eine Gruppe ist.

Aufgabe A1.28

Untersuche, welche der folgenden zweistelligen Operationen Gruppen definieren:

a. $G = \big(\mathbb{Q}\setminus\{0\}\big) \times \big(\mathbb{Q}\setminus\{0\}\big)$ mit $(a,b)\cdot(a',b') = (ab', ba')$ für $a, a', b, b' \in \mathbb{Q}\setminus\{0\}$,

b. $G = \mathbb{R}\times\mathbb{R}\setminus\big\{(0,0)\big\}$ mit $(a,b)\cdot(a',b') = (aa'-bb', ab'+ba')$ für $a, a', b, b' \in \mathbb{R}$.

Aufgabe A1.29

Finde alle möglichen zweistelligen Operationen auf der Menge $G = \{e, a, b, c\}$, bezüglich derer G eine Gruppe mit neutralem Element e wird. Dabei sollten nur Möglichkeiten aufgelistet werden, die nicht durch Vertauschung der Buchstaben a, b und c ineinander überführt werden können.

Aufgabe A1.30 (Boolsche Gruppe)
Es sei M eine Menge.

a. Sind $X, Y, Z \subseteq M$, dann gelten

$$X \setminus \big((Y \setminus Z) \cup (Z \setminus Y)\big) = \big(X \setminus (Y \cup Z)\big) \cup (X \cap Y \cap Z)$$

und

$$\big((X \setminus Y) \cup (Y \setminus X)\big) \setminus Z = \big(X \setminus (Y \cup Z)\big) \cup \big(Y \setminus (X \cup Z)\big).$$

b. Wir definieren auf der Potenzmenge $G = \mathcal{P}(M) = \{A \mid A \subseteq M\}$ von M eine zweistellige Operation durch

$$A + B := (A \setminus B) \cup (B \setminus A) = (A \cup B) \setminus (A \cap B)$$

für $A, B \in G$. Zeige, $(G, +)$ ist eine abelsche Gruppe.

Aufgabe A1.31
Es sei $(G, \cdot)$ eine Gruppe, $g \in G$ und $\emptyset \neq U \subseteq G$ eine endliche Teilmenge von G.

a. Ist $\{g^n \mid n > 0\}$ endlich, so gibt es ein $n > 0$ mit $g^n = e_G$.

b. Genau dann ist U eine Untergruppe von G, wenn für alle $u, v \in U$ auch $u \cdot v \in U$.

Aufgabe A1.32
Sei M eine Menge, $m \in M$, $k \in \mathbb{N}$ und $\sigma \in \mathrm{Sym}(M)$ mit $\sigma^k(m) = m$. Zeige, dann ist auch $\sigma^{q \cdot k}(m) = m$ für alle $q \in \mathbb{Z}$.

Aufgabe A1.33
Zeige, eine Teilmenge $U \subseteq \mathbb{Z}$ ist genau dann eine Untergruppe von $(\mathbb{Z}, +)$, wenn es eine ganze Zahl $n \geq 0$ gibt mit $U = n\mathbb{Z} = \langle n \rangle$.

Aufgabe A1.34
Für zwei reelle Zahlen $a, b \in \mathbb{R}$ definieren wir die Abbildung

$$f_{a,b} : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto a \cdot x + b.$$

Welche der folgenden Mengen sind Untergruppen von $\big(\mathrm{Sym}(\mathbb{R}), \circ\big)$?

a. $U = \{f_{a,b} \mid a, b \in \mathbb{R}, a \neq 0\}$,

b. $V = \{f_{a,1} \mid a \in \mathbb{R}, a \neq 0\}$.

Aufgabe A1.35

Wir betrachten die Gruppe $U = \{f_{a,b} : \mathbb{R} \longrightarrow \mathbb{R} : x \mapsto ax+b \mid a, b \in \mathbb{R}, a \neq 0\}$ aus Aufgabe A1.34, wobei die Gruppenoperation die Verknüpfung von Abbildungen ist, sowie die Gruppe $(\mathbb{R} \setminus \{0\}, \cdot)$. Zeige, daß die Abbildung

$$\alpha : U \longrightarrow \mathbb{R} \setminus \{0\} : f_{a,b} \mapsto a$$

ein Gruppenhomomorphismus ist.

Aufgabe A1.36

Es sei $(G, \cdot)$ eine Gruppe und $h, k \in G$ gegeben. Prüfe, welche Bedingungen für h und k gelten müssen, damit α bzw. β ein Gruppenhomomorphismen ist:

a. $\alpha : G \to G : g \mapsto h \cdot g$,

b. $\beta : G \to G : g \mapsto h^{-1} \cdot g \cdot k$,

Aufgabe A1.37

Es sei $(G, \cdot)$ eine Gruppe. Zeige, $\mathrm{inv} : G \longrightarrow G : g \mapsto g^{-1}$ ist genau dann ein Gruppenhomomorphismus, wenn G abelsch ist.

Aufgabe A1.38

Es sei $\alpha : (G, \cdot) \longrightarrow (H, *)$ ein Gruppenhomomorphismus, $g \in G$ und $g' \in \mathrm{Ker}(\alpha)$. Zeige, dann gilt $g^{-1} \cdot g' \cdot g \in \mathrm{Ker}(\alpha)$.

Aufgabe A1.39

Es sei $(G, \cdot)$ eine Gruppe und $g \in G$.

a. Die Abbildung $L_g : G \longrightarrow G : h \mapsto g \cdot h$ ist bijektiv.

b. Die Abbildung $\alpha : G \longrightarrow \mathrm{Sym}(G) : g \mapsto L_g$ ist ein injektiver Gruppenhomomorphismus.

Aufgabe A1.40

Bestimme alle Gruppenhomomorphismen von $(\mathbb{Z}, +)$ nach $(\mathbb{R}, +)$.

Aufgabe A1.41

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

Aufgabe A1.42

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

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

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

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

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 "·" wie in a. und b. definiert sind.

Aufgabe A1.43 (Die projektive Gerade als Gruppe)

Wir haben in Aufgabe 6.15 die Projektive Gerade $\mathbb{P}^1_{\mathbb{R}}$ 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 $\mathbb{P}^1_{\mathbb{R}}$ mit dieser Operation eine Gruppe ist.

Aufgabe A1.44

Bestimme für die folgenden komplexen Zahlen $\operatorname{Re} z$, $\operatorname{Im} z$, $\arg z$, $|z|$, $\overline{z}$ und z^{-1}:

(a) $z = i - 1$.

(b) $z = \frac{4i}{1+i}$.

(c) $z = \frac{(2+2i)^7}{(1-i)^3}$.

Aufgabe A1.45 (Endliche Integritätsbereiche sind Körper.)

Es sei K eine endliche, nicht-leere Menge mit zwei zweistelligen Operationen "+" und "·", so daß für $(K,+,\cdot)$ bis auf die Existenz von multiplikativen Inversen alle Körperaxiome erfüllt sind. Zeige, wenn aus $x \cdot y = 0$ für $x, y \in K$ stets $x = 0$ oder $y = 0$ folgt, dann ist $(K,+,\cdot)$ eine Körper.

A2 Rechnen mit Matrizen und der Gauß-Algorithmus

Wir führen in diesem Abschnitt den K^n ein, der das wichtigste Beispiel für einen K-Vektorraum darstellt, und allgemeiner den Begriff der $m \times n$-Matrix, mit deren Hilfe sich lineare Abbildungen effizient beschreiben und untersuchen lassen. Diese Strukturen bilden die Grundlage für alles, was im weiteren Verlauf des Kapitels betrachtet wird.

A) Grundlegende Begriffe und Operationen

Definition A2.1 (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 $\mathrm{Mat}(m \times n, K)$ bezeichnet, und falls $m = n$, dann auch kurz mit $\mathrm{Mat}_n(K) = \mathrm{Mat}(n, K)$ und man spricht von *quadratischen Matrizen*.

c. Ist $A = (a_{ij})$ eine $m \times n$-Matrix, dann bezeichnen wir

$$a_i := (a_{i1}, \dots, a_{in})$$

als den i-ten *Zeilenvektor* von A und

$$a^j := \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$$

als den j-ten *Spaltenvektor* von A.

d. Ist $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$, 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

$$K^n := \mathrm{Mat}(n \times 1, K) = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \,\middle|\, x_1, \ldots, x_n \in K \right\}.$$

Die Elemente von K^n heißen *Vektoren* oder *Punkte* im K^n. x_i heißt die i-te *Komponente* des Vektors x.

Definition A2.2 (Operationen mit Matrizen)

a. Es seien $A = (a_{ij}), B = (b_{ij}) \in \mathrm{Mat}(m \times n, K)$ und $\lambda \in K$. Dann definiert man

$$A + B := (a_{ij} + b_{ij}) = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \ldots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \ldots & a_{2n} + b_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \ldots & a_{mn} + b_{mn} \end{pmatrix},$$

sowie

$$\lambda \cdot A := (\lambda a_{ij}) = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \ldots & \lambda a_{1n} \\ \lambda a_{21} & \lambda a_{22} & \ldots & \lambda a_{2n} \\ \vdots & \vdots & & \vdots \\ \lambda a_{m1} & \lambda a_{m2} & \ldots & \lambda a_{mn} \end{pmatrix}.$$

b. Sind $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$ und $B = (b_{jk}) \in \mathrm{Mat}(n \times p, K)$ zwei Matrizen, wobei A genauso viele Spalten wie B Zeilen hat. Dann definieren wir das *Matrixprodukt* durch

$$A \circ B := C, \quad \text{mit } C = (c_{ik}) \in \mathrm{Mat}(m \times p, K) \text{ und } c_{ik} := \sum_{j=1}^{n} a_{ij} b_{jk}.$$

Beispiel A2.3

Folgende Matrizen $A, B \in \mathrm{Mat}(2 \times 3, \mathbb{K})$ und $C \in \mathrm{Mat}(3 \times 2, \mathbb{K})$ 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:

$$A + B = \begin{pmatrix} 4 & 2 & 3 \\ 3 & 5 & 7 \end{pmatrix}, \; 3 \cdot A = \begin{pmatrix} 3 & 0 & 6 \\ 9 & 3 & 6 \end{pmatrix} \text{ und } A \circ C = \begin{pmatrix} 7 & 2 \\ 11 & 3 \end{pmatrix}.$$

Bemerkung A2.4

a. Die in Definition A2.2 a. definierte Addition zweier Matrizen definiert auf $\mathrm{Mat}(m \times n, K)$ offensichtlich eine zweistellige Operation, bezüglich derer $\mathrm{Mat}(m \times n, K)$ eine abelsche Gruppe $(\mathrm{Mat}(m \times n, K), +)$ wird, wie man leicht nachprüft.

b. Wir werden meist kurz λA bzw. λx schreiben, statt $\lambda \cdot A$ bzw. $\lambda \cdot x$, wenn $\lambda \in K$, $A \in \mathrm{Mat}(m \times n, K)$ und $x \in K^n$.

c. Wir schreiben statt $A \circ B$ häufig kurz AB, insbesondere auch Ax statt $A \circ x$.

d. Spaltenvektoren nehmen im Skript sehr viel Raum ein. Um platzsparend arbeiten zu können, werden wir deshalb statt den Spaltenvektor $x \in K^n$ als

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

anzugeben, meist den *transponierten* Zeilenvektor

$$x = (x_1 \ \ldots \ x_n)^t$$

betrachten, und um Mißverständnissen vorzubeugen, fügen wir zudem meist Kommata als Trennsymbole ein

$$x = (x_1, \ldots, x_n)^t.$$

e. 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.

B) Die Abbildung zu einer Matrix

Jede Matrix definiert wie folgt eine Abbildung.

Definition A2.5 (Die Abbildung f_A)
Ist $A \in \mathrm{Mat}(m \times n, K)$, so definieren wir

$$f_A : K^n \to K^m : x \mapsto Ax = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n a_{1j}x_j \\ \vdots \\ \sum_{j=1}^n a_{mj}x_j \end{pmatrix}.$$

f_A heißt die *zu A assoziierte* oder *zu A gehörige Abbildung*.

Beispiel A2.6

Die Matrix

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

definiert die Abbildung

$$f_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x_1, x_2)^t \mapsto (x_1 + 2x_2, 3x_2)^t.$$

Wir wollen uns nun mit der Frage beschäftigen, ob zwei verschiedene Matrizen dieselbe Abbildung definieren können.

Bemerkung A2.7 (Einheitsvektoren)
Um den Zusammenhang zwischen A und f_A besser zu verstehen, betrachten wir für $i = 1, \ldots, n$ den i-ten *Einheitsvektor* $e_i = (\delta_{1i}, \cdots, \delta_{ni})^t \in K^n$, 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.

C) Rechenregeln für Matrizen

Lemma A2.8 (Einfache Rechenregeln für Matrizen)
Für $x, y \in K^n$, $A, B \in \mathrm{Mat}(m \times n, K)$, $C \in \mathrm{Mat}(n \times p, K)$ und $\lambda \in K$ gelten:

a. $A(x + y) = Ax + Ay$ und $A(\lambda x) = \lambda Ax$,

b. $\lambda \cdot (A \circ C) = (\lambda \cdot A) \circ C = A \circ (\lambda \cdot C)$,

c. $f_{A+B} = f_A + f_B$, und

d. $f_{\lambda A} = \lambda f_A$.

Beweis: Der Beweis ist eine Übungsaufgabe. □

Wir wollen jetzt sehen, wie sich die Multiplikation von Matrizen mit den zugehörigen Abbildungen verträgt.

Satz A2.9 (Matrixmultiplikation und Komposition)
Für $A \in \mathrm{Mat}(m \times n, K)$ und $B \in \mathrm{Mat}(n \times p, K)$ 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 $x = (x_1, \ldots, x_p)^t \in K^p$ gegeben.

$$(f_{A \circ B})(x) = (A \circ B)x = \begin{pmatrix} \sum_{j=1}^n a_{1j} b_{j1} & \ldots & \sum_{j=1}^n a_{1j} b_{jp} \\ \vdots & & \vdots \\ \sum_{j=1}^n a_{mj} b_{j1} & \ldots & \sum_{j=1}^n a_{mj} b_{jp} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_p \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{k=1}^p \sum_{j=1}^n a_{1j} b_{jk} x_k \\ \vdots \\ \sum_{k=1}^p \sum_{j=1}^n a_{mj} b_{jk} x_k \end{pmatrix}.$$

Ferner gilt:

$$(f_A \circ f_B)(x) = f_A(Bx) = A(Bx) = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} \sum_{k=1}^p b_{1k} x_k \\ \vdots \\ \sum_{k=1}^p b_{nk} x_k \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^n a_{1j} \sum_{k=1}^p b_{jk} x_k \\ \vdots \\ \sum_{j=1}^n a_{mj} \sum_{k=1}^p b_{jk} x_k \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^n \sum_{k=1}^p a_{1j} b_{jk} x_k \\ \vdots \\ \sum_{j=1}^n \sum_{k=1}^p a_{mj} b_{jk} x_k \end{pmatrix}.$$

Beide Ausdrücke stimmen (bis auf die Reihenfolge der Summation) überein, was zu zeigen war. □

Korollar A2.10 (Die Matrixmultiplikation ist assoziativ.)
Für $A \in \mathrm{Mat}(m \times n, K)$, $B \in \mathrm{Mat}(n \times p, K)$ und $C \in \mathrm{Mat}(p \times q, K)$ gilt

$$A \circ (B \circ C) = (A \circ B) \circ C.$$

Beweis: Dies folgt aus Satz A2.9, da die Komposition von Abbildungen assoziativ ist und da eine Matrix A durch die Abbildung f_A eindeutig bestimmt ist. □

Man kann die Aussage des Korollars natürlich auch direkt nachweisen, was auf die gleiche Rechnung wie in A2.9 führt - statt des einen Vektors x hat man die q Spaltenvektoren von C zu multiplizieren, was etwas mehr Schreibarbeit bedeutet.

Lemma A2.11 (Distributivgesetze)
Sind $A, B \in \mathrm{Mat}(m \times n, K)$ und $C, D \in \mathrm{Mat}(n \times p, K)$, 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 A2.10 aus Lemma A2.8 und Satz A2.9 abgeleitet werden und sei dem Leser als Übung anempfohlen. □

D) Invertierbare Matrizen und die allgemeine lineare Gruppe

Definition A2.12 (Invertierbare Matrizen)
Eine Matrix $A \in \mathrm{Mat}_n(K)$ heißt *invertierbar*, falls es eine Matrix $A^{-1} \in \mathrm{Mat}_n(K)$ gibt, so daß

$$A \circ A^{-1} = A^{-1} \circ A = \mathbb{1}_n,$$

wobei die Matrix $\mathbb{1}_n = (\delta_{ij})_{i,j=1,\ldots,n} \in \mathrm{Mat}_n(K)$ 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 A2.13 (Die allgemeine lineare Gruppe $\mathrm{Gl}_n(K)$)
Die Menge der invertierbaren $n \times n$-Matrizen

$$\mathrm{Gl}_n(K) = \{A \in \mathrm{Mat}_n(K) \mid A \text{ ist invertierbar}\}$$

ist eine Gruppe mit neutralem Element $\mathbb{1}_n$, die für $n > 1$ nicht kommutativ ist. Insbesondere ist die Inverse zu A eindeutig bestimmt und es gelten für $A, B \in \mathrm{Gl}_n(K)$

$$(AB)^{-1} = B^{-1}A^{-1} \quad \text{und} \quad (A^{-1})^{-1} = A.$$

Beweis: Die Assoziativität der Operation folgt aus Korollar A2.10. Daß $\mathbb{1}_n$ das neutrale Element ist, folgt unmittelbar aus der Definition der Multiplikation, und aus der Definition der Invertierbarkeit folgt dann, daß A^{-1} das zu A Inverse ist und in $\mathrm{Gl}_n(K)$ liegt. Also ist $(\mathrm{Gl}_n(K), \circ)$ eine Gruppe. Daß diese nicht kommutativ ist, überlassen wir dem Leser als Übungsaufgabe. □

E) Zeilen-Stufen-Form einer Matrix

Will man in der linearen Algebra konkrete Beispiele rechnen, so wird man in der Regel die Zeilen-Stufen-Form einer Matrix benötigen. Wir wollen nun einführen, was die Zeilen-Stufen-Form einer Matrix ist und uns dann mit dem Gauß-Algorithmus zu ihrer Berechnung beschäftigen.

Definition A2.14 (Zeilen-Stufen-Form)
Es sei $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$.

a. A ist in *Zeilen-Stufen-Form*, kurz ZSF, falls es ein r, mit $0 \leq r \leq m$ und Indizes $j_1, \ldots, j_r$ mit $1 \leq j_1 < j_2 < \ldots < j_r \leq m$ 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 $j = 1, \ldots, n$, und
   (iii) $a_{ij_i} \neq 0$ für $i = 1, \ldots, r$.
   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 $r = \mathrm{rang}(A)$!

b. Eine Zeilen-Stufen-Form von A heißt *reduziert*, falls zusätzlich gilt:
   (iv) $a_{ij_i} = 1$ für $i = 1, \ldots, r$, und
   (v) $a_{kj_i} = 0$ für $k < i$ und $i = 1, \ldots, r$.

Bemerkung A2.15 (Zeilen-Stufen-Form)
Eine Matrix A in Zeilen-Stufen-Form ist also von der folgenden Gestalt:

$$A = \begin{pmatrix}
0 & \ldots & 0 & a_{1j_1} & * & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & * \\
0 & \ldots & \ldots & 0 & \ldots & 0 & a_{2j_2} & * & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & * \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & \ldots & 0 & a_{3j_3} & * & \ldots & \ldots & \ldots & * \\
\vdots & & & & & & & & & & \ddots & & & & \vdots \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 & a_{rj_r} & * & \ldots & * \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0 \\
\vdots & & & & & & & & & & & & & & \vdots \\
0 & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & 0
\end{pmatrix}$$

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 A2.16

Betrachte die Matrizen $A, B, C \in \mathrm{Mat}(4 \times 5, K)$ mit

$$A = \left(\begin{array}{c|cccc} 0 & 1 & 0 & 3 & 0 \\ \hline 0 & 0 & 1 & 2 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 \end{array}\right), B = \left(\begin{array}{c|cccc} 1 & 2 & 3 & 4 & 5 \\ \hline 0 & 2 & 3 & 4 & 5 \\ \hline 0 & 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 0 & 4 & 1 \end{array}\right) \text{ und } C = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 3 & 4 & 0 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Die Matrix A ist in reduzierter ZSF mit $\mathrm{rang}(A) = r = 3$, $j_1 = 2$, $j_2 = 3$ und $j_3 = 5$.
Die Matrix B ist in ZSF mit $\mathrm{rang}(B) = r = 4$ 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.

F) Elementare Zeilen- und Spaltenoperationen

Für die folgende Definition erinnern wir uns (siehe Beispiel A4.11) an die Matrizen

$$E_i^j = (e_{lk})_{l,k=1,\ldots,n} = (\delta_{il} \cdot \delta_{jk})_{l,k=1,\ldots,n} \in \mathrm{Mat}_n(K),$$

die an der Stelle (i, j) den Eintrag 1 und sonst nur Nullen als Einträge haben.

Definition A2.17 (Elementarmatrizen)

Es seien $0 \neq \lambda \in K$, $n > 0$ und $1 \leq i, j \leq n$ mit $i \neq j$. Wir definieren die folgenden quadratischen Matrizen in $\mathrm{Mat}(n, K)$, 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 A2.18 (Elementare Zeilen- und Spaltenoperationen)

Es seien $0 \neq \lambda \in K$, $1 \leq i, j \leq n$ mit $i \neq j$ und $A \in \mathrm{Mat}(n \times m, K)$.

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 A2.19 (Elementarmatrizen sind invertierbar.)
Es seien $0 \neq \lambda \in K$, $1 \leq i, j \leq n$ mit $i \neq j$ und $A \in \mathrm{Mat}(n \times m, K)$. Dann gelten:

a. $S_i\left(\lambda^{-1}\right) \circ S_i(\lambda) = \mathbb{1}_n$,

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 $0 \neq \lambda \in K$ gilt, vermittels der Distributivität der Matrixmultiplikation:

$$Q_i^j(-\lambda) \circ Q_i^j(\lambda) = \left(\mathbb{1}_n - \lambda \cdot E_i^j\right) \circ \left(\mathbb{1}_n + \lambda \cdot E_i^j\right) = \mathbb{1}_n - \lambda^2 \cdot E_i^j \circ E_i^j = \mathbb{1}_n,$$

da $E_i^j \circ E_i^j = 0$ wegen $i \neq j$. Beachte dazu, daß für $E_i^j \circ E_i^j = (c_{lk})$ gilt:

$$c_{lk} = \sum_{p=1}^{n} \delta_{il}\delta_{jp}\delta_{ip}\delta_{jk},$$

und daß für $i \neq j$ und p beliebig gilt $\delta_{jp}\delta_{ip} = 0$. □

G) Der Gauß-Algorithmus

Wir wollen nun zeigen, daß man Matrizen über einem Körper K mittels elementarer Zeilenoperationen auf Zeilen-Stufen-Form bringen kann. Der Beweis ist konstruktiv; der sich daraus ergebende Algorithmus heißt Gauß-Algorithmus.

Satz A2.20 (Existenz der reduzierten Zeilen-Stufen-Form)
Jede Matrix $A \in \mathrm{Mat}(m \times n, K)$ läßt sich mittels endlich vieler elementarer Zeilenoperationen in reduzierte Zeilen-Stufen-Form $\mathrm{rZSF}(A)$ überführen, d.h. es gibt Elementarmatrizen $T_1, \ldots, T_k$, so daß

$$\mathrm{rZSF}(A) = T_1 \circ \ldots \circ T_k \circ A.$$

Beweis: Sei $A \in \mathrm{Mat}(m \times n, K)$. Ist $A = 0$, so hat A bereits ZSF mit $r = 0$ und wir sind fertig.

Ist $A \neq 0$, so führe folgende Schritte durch:

1. 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:

$$\begin{pmatrix} 0 & \dots & 0 & 0 & * & \dots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \vdots & & \vdots & 0 & * & \dots & * \\ \vdots & & \vdots & a_{i_1 j_1} & * & \dots & * \\ \vdots & & \vdots & * & * & \dots & * \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \dots & 0 & * & * & \dots & * \end{pmatrix}$$

2. 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 heiße $\widetilde{A}_1 = (\tilde{a}_{ij})$. Dann ist $\tilde{a}_{1 j_1}$ unser erstes Pivot.

3. Schritt: Erzeuge in der Spalte $\tilde{a}^{j_1}$ von $\widetilde{A}_1$ unterhalb von $\tilde{a}_{1 j_1}$ Nullen durch elementare Operationen vom Typ II, d. h. addiere für $k = 2, \dots, m$ zur k-ten Zeile das $-\frac{\tilde{a}_{k j_1}}{\tilde{a}_{1 j_1}}$-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:

$$A^{(1)} := \left(\begin{array}{cccc|ccc} 0 & \dots & 0 & a^{(1)}_{1 j_1} & * & \dots & * \\ \hline 0 & \dots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & A_2 & \\ 0 & \dots & 0 & 0 & & & \end{array}\right),$$

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:

$$A^{(2)} = \left(\begin{array}{ccccccc|ccc} 0 & \dots & 0 & a^{(2)}_{1 j_1} & * & \dots & \dots & \dots & \dots & * \\ 0 & \dots & 0 & 0 & \dots & 0 & a^{(2)}_{2 j_2} & * & \dots & * \\ \hline 0 & \dots & 0 & 0 & \dots & 0 & 0 & & & \\ \vdots & & \vdots & \vdots & & \vdots & \vdots & & A_3 & \\ 0 & \dots & 0 & 0 & \dots & 0 & 0 & & & \end{array}\right)$$

mit einem Pivot $a^{(2)}_{2 j_2}$ 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, $A^{(4)}$, etc.. 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 $A^{(r)} = \left(a_{ij}^{(r)}\right)$ in reduzierte ZSF zu bringen, multiplizieren wir zunächst die Zeilen $a_i^{(r)}$, für $i = 1, \ldots, r$, mit $\frac{1}{a_{ij_i}^{(r)}}$, was einer elementaren Zeilenoperation vom Typ I entspricht. Die so entstehende Matrix heiße $A' = (a'_{ij})$. Sodann addiert man für $i = 1, \ldots, r$ und $k = 1, \ldots, i-1$ zur k-ten Zeile das $-a'_{kj_i}$-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 $A'' = \left(a''_{ij}\right)$ reduzierte ZSF hat, da in Spalte j_i die Elemente a'_{kj_i} in $a''_{kj_i} = 0$, für $k < i$, übergegangen sind. □

Bemerkung A2.21 (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. Für den Beweis der Aussage benögtigt man etwas mehr Theorie, weshalb dieser an einer späteren Stelle in der Ausarbeitung zu finden ist.

Beispiel A2.22 (Der Gauß-Algorithmus zur Berechnung von rZSF(A))

Wir überführen nun die folgende Matrix in reduzierte ZSF.

$$\begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ \underline{-1} & 1 & -3 & 0 & 2 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} \xrightarrow{I \leftrightarrow II} \begin{pmatrix} \underline{-1} & 1 & -3 & 0 & 2 \\ 0 & 0 & -1 & 2 & 3 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} \xrightarrow{III \mapsto III+I}$$

$$\left(\begin{array}{cc|ccc} -1 & 1 & -3 & 0 & 2 \\ \hline 0 & 0 & \underline{-1} & 2 & 3 \\ 0 & 0 & -2 & 4 & 5 \end{array}\right) \xrightarrow{III \mapsto III-2\cdot II} \begin{pmatrix} \underline{-1} & 1 & -3 & 0 & 2 \\ 0 & 0 & \underline{-1} & 2 & 3 \\ 0 & 0 & 0 & 0 & \underline{-1} \end{pmatrix} \xrightarrow[III \mapsto -III]{I \mapsto -I, II \mapsto -II}$$

$$\begin{pmatrix} 1 & -1 & 3 & 0 & -2 \\ 0 & 0 & 1 & -2 & -3 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[II \mapsto II+3\cdot III]{I \mapsto I+2\cdot III} \begin{pmatrix} 1 & -1 & 3 & 0 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xrightarrow{I \mapsto I-3\cdot II}$$

$$\begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

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.

$$\begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ \underline{1} & -1 & 1 & 4 & 3 \end{pmatrix} \xmapsto{I \leftrightarrow III} \begin{pmatrix} 1 & -1 & 1 & 4 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ 0 & 0 & -1 & 2 & 3 \end{pmatrix} \xmapsto{II \mapsto II+I}$$

$$\left(\begin{array}{cc|ccc} 1 & -1 & 1 & 4 & 3 \\ \hline 0 & 0 & \underline{-2} & 4 & 5 \\ 0 & 0 & -1 & 2 & 3 \end{array}\right) \xmapsto{III \mapsto III - \frac{1}{2} \cdot II} \begin{pmatrix} \underline{1} & -1 & 1 & 4 & 3 \\ 0 & 0 & \underline{-2} & 4 & 5 \\ 0 & 0 & 0 & 0 & \underline{\frac{1}{2}} \end{pmatrix} \xmapsto[III \mapsto 2 \cdot III]{II \mapsto -\frac{1}{2} \cdot II}$$

$$\begin{pmatrix} 1 & -1 & 1 & 4 & 3 \\ 0 & 0 & 1 & -2 & -\frac{5}{2} \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xmapsto[II \mapsto II + \frac{5}{2} \cdot III]{I \mapsto I - 3 \cdot III} \begin{pmatrix} 1 & -1 & 1 & 4 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \xmapsto{I \mapsto I - II}$$

$$\begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Bemerkung A2.23

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 werden im weiteren Verlauf der Ausarbeitung viele Algorithmen kennenlernen, die darauf fußen, daß wir die reduzierte Zeilen-Stufen-Form einer Matrix berechnen können und mit deren Hilfe man interessante Objekte der linearen Algebra explizit ausrechnen kann. Deshalb wollen wir abschließend den Gauß-Algorithmus zur Berechnung der reduzierten Zeilen-Stufen-Form einer Matrix noch einmal explizit und in leicht abgewandelter Form als rekursiven Algorithmus formulieren.

Algorithmus A2.24 (Gauß-Algorithmus)

INPUT: $A \in \mathrm{Mat}(m \times n, K)$.

OUTPUT: $\mathrm{rZSF}(A)$, die reduzierte Zeilen-Stufen-Form von A.

0. Schritt: Falls $A = 0$, gehe zu Schritt 8.

1. Schritt: Falls $m = 1$, gehe zu Schritt 7.

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 $k = 2, \ldots, m$ addiere zur k-ten Zeile das $-\frac{a_{k1}}{a_{11}}$-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 $k = 1, \ldots, i-1$ zur k-ten Zeile das $-a_{kj}$-fache der i-ten Zeile.

8. Schritt: Gib die (veränderte) Matrix A zurück.

Aufgaben

Aufgabe A2.25

Zeige, daß die Matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_2(K)$$

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 A2.26

Es sei $A \in \mathrm{Mat}(m \times n, K)$ und $B \in \mathrm{Mat}(n \times p, K)$. Zeige, $(AB)^t = B^t A^t$.

[2] Dies 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.

Aufgabe A2.27 (Nilpotente Matrizen)

Es sei $N = (n_{ij}) \in \mathrm{Mat}_n(K)$ eine Matrix, für die die Einträge auf der oberen Nebendiagonale alle 1 sind und für die alle anderen Einträge 0 sind, d.h. $n_{ij} = \delta_{j-i,1}$.
Zeige für $k = 1, \ldots, n$, daß die Einträge der Matrix $N^k = \left(n_{ij}^{(k)}\right)$ auf der k-ten oberen Nebendiagonale alle 1 und alle anderen Einträge 0 sind, d.h. $n_{ij}^{(k)} = \delta_{j-i,k}$. Insbesondere ist $N^n = 0$ und $N^k \neq 0$ für $k < n$.

Aufgabe A2.28

Es seien $0 \neq \lambda \in K$, $1 \leq i, j \leq n$ mit $i \neq j$. Dann gelten:

$$Q_i^j(\lambda) = S_j\left(\lambda^{-1}\right) \circ Q_i^j(1) \circ S_j(\lambda),$$

und

$$P_i^j = Q_j^i(1) \circ Q_i^j(-1) \circ Q_j^i(1) \circ S_j(-1).$$

Aufgabe A2.29

Überführe die folgende Matrix in reduzierte Zeilen-Stufen-Form:

$$A = \begin{pmatrix} 0 & 0 & 2 & 6 & 4 & 0 \\ -1 & -2 & 1 & 2 & 1 & 0 \\ 2 & 4 & 0 & 2 & 2 & 1 \\ 1 & 2 & 0 & 1 & 1 & 0 \\ -2 & -4 & 2 & 4 & 2 & 0 \end{pmatrix} \in \mathrm{Mat}(5 \times 6, \mathbb{Q}).$$

Aufgabe A2.30

Gegeben seien folgende Matrizen und Vektoren über den reellen Zahlen:

$$A = \begin{pmatrix} 1 & 2 & 8 & -2 \\ 0 & 1 & -1 & 0 \\ 2 & 1 & 0 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 3 & 0 \\ 2 & 1 & 4 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix},$$

$$C = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \\ -1 & -1 & -1 \end{pmatrix}, \quad x = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix}.$$

Berechne die Terme:

$$A + B, \quad B \circ x, \quad A \circ C \quad \text{und} \quad B \circ B^t.$$

Aufgabe A2.31

Zeige, für $A \in \mathrm{Mat}(m \times n, K)$ und $B \in \mathrm{Mat}(n \times p, K)$ gilt $(A \circ B)^t = B^t \circ A^t$.

Aufgabe A2.32

Bestimme eine Matrix $A \in \mathrm{Mat}_2(\mathbb{Q})$, so dass

$$\begin{pmatrix} 1 & 4 \\ 2 & 7 \end{pmatrix} \circ A = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Aufgabe A2.33

Bestimme eine reduzierte Zeilen-Stufen-Form der folgenden Matrizen über $\mathbb{R}$ (das geht auch ohne zu rechnen):

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 2 \\ 2 & 1 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 3 & 0 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{pmatrix}, \quad D = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

A3 Vektorräume und lineare Abbildungen

A) Vektorräume

Definition A3.1 (Vektorräume)
Ein *K-Vektorraum* (oder Vektorraum über K) besteht aus einer nicht-leeren Menge V sowie einer zweistelligen Operation

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

die *Vektoraddition* genannt wird, und einer zweistelligen Operation

$$\cdot : K \times V \to V : (\lambda, x) \mapsto \lambda \cdot x = \lambda x,$$

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 A3.2

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 $\mathrm{Mat}(m \times n, K)$ der $m \times n$-Matrizen über K mittels der in Definition A2.2 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 $0_{K^n} = (0, \ldots, 0)^t$.
Speziell sind $\mathbb{R}^n$, $\mathbb{C}^n$ und $\mathbb{F}_2^n$ Vektorräume über $\mathbb{R}$, $\mathbb{C}$ bzw. $\mathbb{F}_2$.

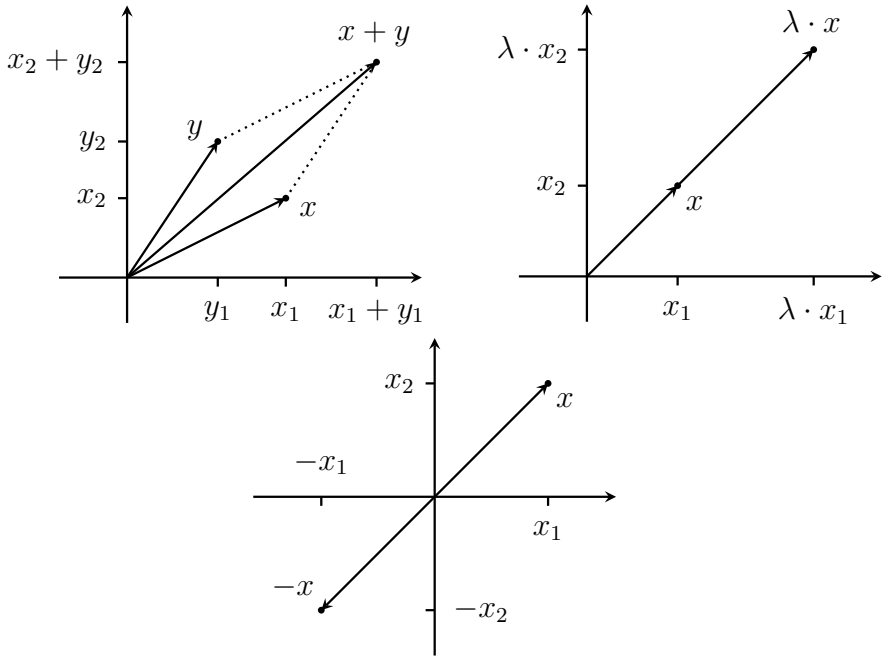


Abbildung 6: Addition und Skalarmultiplikation in $\mathbb{R}^2$

f. Ist M eine Menge und V ein K-Vektorraum, so wird die Menge

$$V^M = \{f : M \longrightarrow V \mid f \text{ ist eine Abbildung}\}$$

durch die Operationen

$$+ : V^M \times V^M \to V^M : (f, g) \mapsto \big(f + g : M \to V : x \mapsto f(x) + g(x)\big)$$

und

$$\cdot : K \times V^M \to V^M : (\lambda, f) \mapsto \big(\lambda \cdot f : M \to V : x \mapsto \lambda \cdot f(x)\big)$$

zu einem K-Vektorraum, wie man leicht nachrechnet.
Ist z.B. $M = \mathbb{N}$ und $K = V = \mathbb{R}$, so ist

$$\mathbb{R}^{\mathbb{N}} = \{a : \mathbb{N} \longrightarrow \mathbb{R} \mid a \text{ ist eine Abbildung}\} = \big\{(a_n)_{n\in\mathbb{N}} \;\big|\; a_n \in \mathbb{R}\big\}$$

der Vektorraum der Folgen in $\mathbb{R}$. Unsere Definitionen sorgen dafür, daß Folgen komponentenweise addiert und mit Skalaren multipliziert werden.

g. Ist t eine Veränderliche und sind $a_0, \ldots, a_n \in K$, so nennen wir einen Ausdruck der Form

$$\sum_{k=0}^{n} a_k \cdot t^k = a_n \cdot t^n + a_{n-1} \cdot t^{n-1} + \ldots + a_1 \cdot t + a_0$$

ein *Polynom* in der Veränderlichen t mit Koeffizienten in K. 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 $\deg(0) := -\infty$. Die Menge

$$K[t] := \left\{ \sum_{k=0}^{n} a_k \cdot t^k \;\middle|\; n \in \mathbb{N}, a_k \in K \right\}$$

der Polynome mit Koeffizienten in K wird durch die Addition

$$\sum_{k=0}^{m} a_k \cdot t^k + \sum_{k=0}^{n} b_k \cdot t^k := \sum_{k=0}^{\max\{m,n\}} (a_k + b_k) \cdot t^k,$$

wobei $a_k = 0$ für $k > m$ und $b_k = 0$ für $k > n$ gelten soll, und durch die Skalarmultiplikation

$$\lambda \cdot \sum_{k=0}^{n} a_k \cdot t^k := \sum_{k=0}^{n} (\lambda \cdot a_k) \cdot t^k$$

zu einem K-Vektorraum, wie man leicht nachprüft.

h. Da man für $M = \{1, \ldots, n\}$ eine Abbildung $f : M \to K$ in eindeutiger Weise durch das Tupel der Bildelemente $\big(f(1), \ldots, f(n)\big)$ beschreiben kann, sieht man leicht, daß die Zuordnung

$$K^M \to K^n : f \mapsto \big(f(1), \ldots, f(n)\big)^t$$

in diesem Falle eine Bijektion ist. Man prüft überdies leicht nach, daß diese Abbildung ein Vektorraumhomomorphismus im Sinne von Definition A3.19 ist. K^M und K^n sind dann also isomorph.

i. Ist $(V_i \mid i \in I)$ eine Familie von K-Vektorräumen, dann wird das *direkte Produkt* der V_i

$$\prod_{i \in I} V_i = \left\{ (x_i)_{i \in I} \;\middle|\; x_i \in V_i \text{ für } i \in I \right\}$$

mit der komponentenweisen Addition und Skalarmultiplikation ein K-Vektorraum, d. h. durch

$$(x_i)_{i \in I} + (y_i)_{i \in I} := (x_i + y_i)_{i \in I}$$

und

$$\lambda \cdot (x_i)_{i \in I} := (\lambda x_i)_{i \in I}$$

für $\lambda \in K$ und $(x_i)_{i \in I}, (y_i)_{i \in I} \in V$.

Lemma A3.3 (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:

$$\lambda \cdot x = 0_V \quad \Longrightarrow \quad \lambda = 0 \quad \text{oder} \quad x = 0.$$

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

$$0_V = \lambda^{-1} \cdot 0_V = \lambda^{-1} \cdot (\lambda \cdot x) = \left(\lambda^{-1} \cdot \lambda\right) \cdot x = 1 \cdot x = x.$$

c. Für $x \in K$ gilt:

$$x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = \big(1 + (-1)\big) \cdot x = 0_K \cdot x = 0_V.$$

Also ist $(-1) \cdot x$ das (eindeutig bestimmte) additive Inverse zu x.

□

B) Unterräume

Definition A3.4 (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

$$\lambda \cdot x \in U \quad \text{und} \quad x + y \in U. \tag{189}$$

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 A3.5 (Unterräume sind Vektorräume.)
Jeder Unterraum eines K-Vektorraums ist selbst ein K-Vektorraum.

Beweis: Aus der Abgeschlossenheit bezüglich der Skalarmultiplikation folgt insbesondere

$$-x = (-1) \cdot x \in U$$

für alle $x \in U$, d.h. die Abgeschlossenheit bezüglich der Inversenbildung. Zusammen mit der Abgeschlossenheit bezüglich der Addtion ist U also eine Untergruppe von $(V, +)$ und damit selbst eine abelsche Gruppe nach Bemerkung A1.6.

Die verbleibenden Axiome aus Teil b. in Definition A3.1 vererben sich analog zum Assoziativgesetz unmittelbar von V auf U. Damit ist U also ein K-Vektorraum. □

Beispiel A3.6

a. 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*.

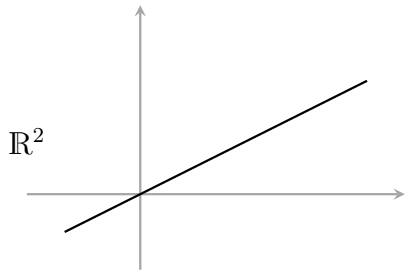
b. Eine Gerade G durch den Ursprung in der Ebene $\mathbb{R}^2$ mit Steigung m genügt der Gleichung $y = m \cdot x$ und ist deshalb

$$G = \left\{ (x, mx)^t \;\middle|\; x \in \mathbb{R} \right\}.$$

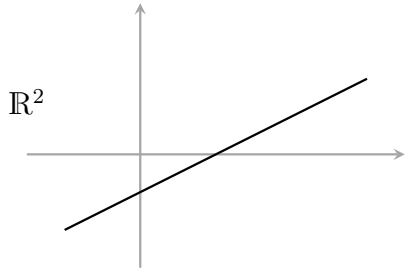
Für $v = (x, mx)^t, w = (x', mx')^t \in G$ und $\lambda \in \mathbb{R}$ gilt dann

$$v + w = \left(x + x', m \cdot (x + x')\right)^t,\ \lambda \cdot v = (\lambda x, m\lambda x)^t \in G.$$

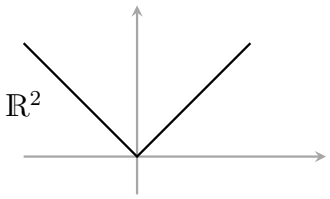
Mithin ist G ein Unterraum von $\mathbb{R}^2$.



c. 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

$$U := \{(a_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid (a_n)_{n\in\mathbb{N}} \text{ ist konvergent}\}$$

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

$$P_n := \left\{ \sum_{k=0}^{n} a_k \cdot t^k \;\middle|\; a_k \in K \right\}$$

einen Unterraum des Vektorraums der Polynome $K[t]$.[3]

h. Ist $(V_i \mid i \in I)$ eine Familie von K-Vektorräumen, dann ist

$$\bigoplus_{i\in I} V_i = \left\{ (x_i)_{i\in I} \in \prod_{i\in I} V_i \;\middle|\; x_i \neq 0 \text{ für höchstens endlich viele } i \in I \right\}$$

ein Unterraum des direkten Produktes $\prod_{i\in I} V_i$, der die *äußere direkte Summe* der V_i genannt wird.

Lemma A3.7 (Durchschnitt von Unterräumen)
Der Durchschnitt beliebig vieler Unterräume ist wieder ein Unterraum.

Beweis: Es seien U_i für $i \in I$ Unterräume eines K-Vektorraums V. Da $0_V \in U_i$ für alle $i \in I$, ist $U := \bigcap_{i\in I} U_i$ nicht die leere Menge. Es bleibt also zu zeigen, daß für $x, y \in U$ und $\lambda \in K$ gilt:

$$x + y \in U \quad \text{und} \quad \lambda x \in U.$$

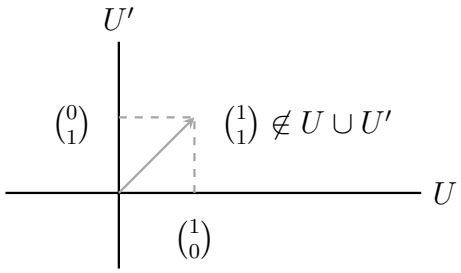
Für ein beliebiges $i \in I$ gilt, da U_i ein Unterraum von V ist und da $x, y \in U \subseteq U_i$, daß $x + y \in U_i$ und $\lambda x \in U_i$. Also liegen die Vektoren im Durchschnitt U. □

[3] Man beachte, obwohl $K[t]$ auch ein Ring ist, ist P_n für $n \geq 1$ kein Unterring von $K[t]$, weil $t^n \cdot t^n = t^{2n} \notin P_n$.

Bemerkung A3.8

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 Unterräume von $\mathbb{R}^2$. Es gilt aber $(1,1)^t = e_1 + e_2 \notin U \cup U'$ und kann $U \cup U'$ kein Unterraum von $\mathbb{R}^2$ sein.



C) Die lineare Hülle und Summen von Unterräumen

Definition A3.9 (Linearkombination und lineare Hülle)

Es sei V ein K-Vektorraum.

a. Wir nennen $x \in V$ eine *Linearkombination* von $x_1, \ldots, x_r \in V$, falls es $\lambda_1, \ldots, \lambda_r \in K$ gibt mit

$$x = \lambda_1 x_1 + \ldots + \lambda_r x_r.$$

Ist eines der λ_i ungleich Null, so nennen wir die Linearkombination *nicht-trivial*.

b. Ist $M \subseteq V$, so nennen wir den Durchschnitt

$$\mathrm{Lin}\,(M) := \langle M \rangle := \bigcap_{M \subseteq U \leq V} U$$

aller Unterräume von V, die M enthalten, die *lineare Hülle* oder das *Erzeugnis* von M.

Bemerkung A3.10

a. Man beachte, daß die lineare Hülle von M wegen Lemma A3.7 ein Unterraum von V ist. Aufgrund der Definition ist es *der kleinste* Unterraum, der M enthält.

b. Ist $M = \emptyset$, so ist $\mathrm{Lin}\,(M) = \mathrm{Lin}\,(\emptyset) = \{0\}$.

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 (189) unmittelbar, daß ein Unterraum U abgeschlossen bezüglich endlicher Linearkombinationen von Vektoren aus U ist.

Proposition A3.11 (Lineare Hülle = Menge der Linearkombinationen)
Ist V ein K-Vektorraum und $\emptyset \neq M \subseteq V$, so ist die lineare Hülle von M

$$\mathrm{Lin}\,(M) = \{\lambda_1 \cdot x_1 + \ldots + \lambda_r \cdot x_r \mid r \in \mathbb{N}, x_1, \ldots, x_r \in M, \lambda_1, \ldots, \lambda_r \in K\} \leq V$$

die Menge aller Linearkombinationen von Elementen in M.

Beweis: Wir setzen

$$U := \{\lambda_1 \cdot x_1 + \ldots + \lambda_r \cdot x_r \mid r \in \mathbb{N}, x_1, \ldots, x_r \in M, \lambda_1, \ldots, \lambda_r \in K\}.$$

Als Unterraum, der M enthält, enthält $\mathrm{Lin}\,(M)$ auch jede endliche Linearkombination von Elementen in M, also auch U.

Wir wollen nun zeigen, daß U ein Unterraum von V ist, der M enthält, da er aufgrund der Definition des Erzeugnisses dann auch das Erzeugnis von M enthält. Dazu beachten wir zunächst, daß für $x \in M$ auch $x = 1 \cdot x \in U$ gilt. Also ist $M \subseteq U$ und somit ist U auch nicht leer. Seien nun $x = \sum_{i=1}^{r} \lambda_i \cdot x_i$ und $y = \sum_{j=1}^{s} \mu_j \cdot y_j$ mit $x_i, y_j \in M$ und $\lambda_i, \mu_i \in K$ sowie $\lambda \in K$ gegeben. Dann ist

$$x + y = \lambda_1 \cdot x_1 + \ldots + \lambda_r \cdot x_r + \mu_1 \cdot y_1 + \ldots + \mu_s \cdot y_s \in U,$$

weil es eine endliche Linearkombination von Elementen in M ist, und ebenso ist

$$\lambda \cdot x = \sum_{i=1}^{r} (\lambda \cdot \lambda_i) \cdot x_i \in U.$$

Also ist U ein Unterraum von V. □

Beispiel A3.12

a. Ist $M = \{x_1, \ldots, x_n\} \subseteq V$ endlich, so ist das Erzeugnis von M

$$\mathrm{Lin}\,(x_1, \ldots, x_n) := \mathrm{Lin}\,(M) = \{\lambda_1 \cdot x_1 + \ldots + \lambda_n \cdot x_n \mid \lambda_1, \ldots, \lambda_n \in K\}.$$

Insbesondere gilt $\mathrm{Lin}\,(x) := \mathrm{Lin}\,(\{x\}) = \{\lambda \cdot x \mid \lambda \in K\}$.

b. Die Lineare Hülle der Vektoren $x_1 = (1, 0)^t$ und $x_2 = (0, 1)^t$ in $\mathbb{R}^2$ ist

$$\mathrm{Lin}\,(x_1, x_2) = \{\lambda_1 \cdot x_1 + \lambda_2 \cdot x_2 = (\lambda_1, \lambda_2)^t \mid \lambda_1, \lambda_2 \in \mathbb{R}\} = \mathbb{R}^2.$$

c. Die Lineare Hülle von $x = (1, m)^t \in \mathbb{R}^2$ ist die Gerade

$$\mathrm{Lin}\,(x) = \{(\lambda, \lambda m)^t \mid \lambda \in \mathbb{R}\}.$$

d. Es gilt offenbar $\mathrm{Lin}\,(t^0, t^1, \ldots, t^n) = P_n$.

Proposition A3.13 (Summe zweier Unterräume)
Es seien U und U' Unterräume des K-Vektorraums V. Dann gilt

$$U + U' := \{u + u' \mid u \in U, u' \in U'\} = \mathrm{Lin}\,(U \cup U') \leq V,$$

und wir nennen diesen Unterraum von V die *Summe* von U und U'.

Beweis: Wegen Proposition A3.11 ist $U+U'$ in $\mathrm{Lin}\,(U \cup U')$ enthalten, da die Elemente von $U + U'$ Linearkombinationen von Elementen in $U \cup U'$ sind.

Umgekehrt ist jede Linearkombination von Elementen in $U \cup U'$ von der Form $\sum_{i=1}^{r} \lambda_i \cdot u_i + \sum_{j=1}^{s} \mu_j \cdot u'_j$ mit $u_i \in U$, $u'_j \in U'$ und $\lambda_i, \mu_j \in K$. Da U und U' aber Unterräume sind, ist

$$u := \sum_{i=1}^{r} \lambda_i \cdot u_i \in U$$

und

$$u' := \sum_{j=1}^{s} \mu_j \cdot u'_j \in U'.$$

Deshalb ist die Linearkombination

$$\sum_{i=1}^{r} \lambda_i \cdot u_i + \sum_{j=1}^{s} \mu_j \cdot u'_j = u + u' \in U + U'$$

in $U + U'$, und mit Proposition A3.11 enthält $U + U'$ auch $\mathrm{Lin}\,(U \cup U')$. □

Bemerkung A3.14 (Summen von Unterräumen)

a. Die Summe zweier Unterräume ersetzt ihre Vereinigung in der Theorie der Vektorräume. Sie ist der kleinste Unterraum, der beide enthält. Im Beispiel aus Bemerkung A3.8 ist die Summe der beiden Unterräume ganz $\mathbb{R}^2$.

b. Analog zu Proposition A3.13 zeigt man allgemeiner: Sind $U_1, \ldots, U_n$ Unterräume des K-Vektorraums V, so gilt

$$U_1 + \ldots + U_n := \{u_1 + \ldots + u_n \mid u_i \in U_i\} = \mathrm{Lin}\,(U_1 \cup \ldots \cup U_n)\,.$$

Beispiel A3.15
Jeder Vektor x in $U + U'$ läßt sich schreiben als $x = u + u'$ mit $u \in U$ und $u' \in U'$, diese Darstellung muß aber nicht eindeutig sein.

Sind z.B. $U = \mathrm{Lin}\,((1,0,0)^t, (0,1,1)^t)$ und $U' = \mathrm{Lin}\,((1,1,0)^t, (1,0,1)^t)$ als Unterräume von $\mathbb{R}^3$ gegeben, so können wir den Vektor $x = (1,0,-1)^t$ auf folgende beiden Weisen als Summe zweier Vektoren in U und U' schreiben:

$$x = (0,-1,-1)^t + (1,1,0)^t = (2,0,0)^t + (-1,0,-1)^t.$$

Definition A3.16 (Innere direkte Summe)
Es seien $U_1, \ldots, U_n$ Unterräume des K-Vektorraums V. Wir nennen die Summe $U = U_1 + \ldots + U_n$ eine *(innere) direkte Summe*, wenn sich jeder Vektor $x \in U_1 + \ldots + U_n$ auf eindeutige Weise als Summe $x = u_1 + \ldots + u_n$ mit $u_i \in U_i$ schreiben läßt. Wir schreiben dann $U = U_1 \oplus \ldots \oplus U_n$. (Siehe auch Aufgabe A3.51 und Aufgabe A3.52.)

Proposition A3.17 (Direkte Summe zweier Unterräume)
Es seien U, U' und W Unterräume des K-Vektorraums V.
Genau dann gilt $W = U \oplus U'$, wenn $W = U + U'$ und $U \cap U' = \{0\}$.

Beweis: Ist die Summe $W = U \oplus U'$, so gilt insbesondere $W = U + U'$. Für $x \in U \cap U'$, gilt zudem

$$x = x + 0 = 0 + x \in U + U',$$

und wegen der Eindeutigkeit der Darstellung in $U + U'$ muß $x = 0$ 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. □

Beispiel A3.18
Betrachte die Unterräume $U = \mathrm{Lin}\,((1,1,1)^t)$ und $U' = \mathrm{Lin}\,((1,0,1)^t)$ 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.

D) Lineare Abbildungen

Zu jeder Struktur gehören die strukturerhaltenden Abbildungen.

Definition A3.19 (Lineare Abbildungen)
Es seien V und W zwei K-Vektorräume.

a. Eine Abbildung $f : V \to W$ 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 $\mathrm{Hom}_K(V, W)$ und die Menge aller Endomorphismen von V mit $\mathrm{End}_K(V)$.

Bemerkung A3.20
Die beiden Bedingungen in Definition A3.19 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 A3.21
Wir wollen uns nun einige Beispiele für lineare Abbildungen anschauen.

a. Die Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x_1, x_2)^t \mapsto x_1 - x_2$ ist $\mathbb{R}$-linear.
Denn für $x = (x_1, x_2)$ und $y = (y_1, y_2)$ sowie $\lambda \in \mathbb{R}$ gilt

$$\begin{aligned} f(x+y) &= f\big((x_1 + y_1, x_2 + y_2)\big) = x_1 + y_1 - x_2 - y_2 \\ &= x_1 - x_2 + y_1 - y_2 = f(x) + f(y) \end{aligned}$$

und

$$f(\lambda x) = f\big((\lambda x_1, \lambda x_2)\big) = \lambda x_1 - \lambda x_2 = \lambda \cdot (x_1 - x_2) = \lambda \cdot f(x).$$

b. Ist I ein Intervall, so ist die Abbildung

$$D : \mathcal{C}^1(I, \mathbb{R}) \longrightarrow \mathcal{C}(I, \mathbb{R}) : f \mapsto f'$$

$\mathbb{R}$-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

$$f : \mathbb{R}[t] \longrightarrow \mathbb{R}^{\mathbb{R}} : \sum_{k=0}^{n} a_k \cdot t^k \mapsto (a_k)_{k \in \mathbb{N}},$$

wobei $a_k = 0$ für $k > n$ gelten soll, ist eine injektive $\mathbb{R}$-lineare Abbildung, also ein Monomorphismus. Ihr Bild, der Unterraum

$$\mathrm{Im}(f) = \{(a_k)_{k\in\mathbb{N}} \mid \text{ nur endlich viele } a_k \text{ sind nicht } 0\}$$

der *abbrechenden Folgen* in $\mathbb{R}$, ist mithin isomorph zu $K[t]$.

d. Die formale Ableitung

$$d : K[t] \longrightarrow K[t] : \sum_{k=0}^{n} a_k \cdot t^k \mapsto \sum_{k=1}^{n} k \cdot a_k \cdot t^{k-1}$$

ist eine K-lineare Abbildung, wie man leicht nachrechnet.

Lemma A3.22 (Einfache Eigenschaften linearer Abbildungen)
Seien U, V und W K-Vektorräume und $f : U \longrightarrow V$ und $g : V \longrightarrow W$ seien K-linear. Ferner seien $x, x_1, \ldots, x_n \in U$ und $\lambda_1, \ldots, \lambda_n \in K$. Dann gelten:

a. $f(0_U) = 0_V$ und $f(-x) = -f(x)$.

b. $f(\lambda_1 x_1 + \ldots + \lambda_n x_n) = \lambda_1 f(x_1) + \ldots + \lambda_n f(x_n)$.

c. Ist f bijektiv, so ist $f^{-1} : V \longrightarrow U$ K-linear.

d. $g \circ f : U \longrightarrow W$ ist K-linear.

e. $\mathrm{Hom}_K(U, V)$ 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 $y, y' \in V$ und $\lambda, \lambda' \in K$ sowie $x = f^{-1}(y)$ und $x' = f^{-1}(y')$. Wegen der Linearität von f gilt

$$f(\lambda \cdot x + \lambda' \cdot x') = \lambda \cdot f(x) + \lambda' \cdot f(x').$$

Wenden wir auf beiden Seiten f^{-1} an, so erhalten wir

$$\begin{aligned} \lambda \cdot f^{-1}(y) + \lambda' \cdot f^{-1}(y') &= \lambda \cdot x + \lambda' \cdot x' = f^{-1}\big(f(\lambda \cdot x + \lambda' \cdot x')\big) \\ &= f^{-1}\big(\lambda \cdot f(x) + \lambda' \cdot f(x')\big) = f^{-1}\big(\lambda \cdot y + \lambda' \cdot y'\big). \end{aligned}$$

Mithin ist f^{-1} eine lineare Abbildung.

d. Seien $\lambda, \mu \in K$ und $x, y \in U$, so gelten

$$\begin{aligned}(g \circ f)(\lambda x + \mu y) &= g\big(f(\lambda x + \mu y)\big) = g(\lambda f(x) + \mu f(y)\big) \\ &= \lambda g\big(f(x)\big) + \mu g(\big(f(y)\big) = \lambda(g \circ f)(x) + \mu(g \circ f)(y).\end{aligned}$$

e. Dies folgt aus Aufgabe A3.45.

□

Proposition A3.23 (f_A ist linear.)
Für $A \in \mathrm{Mat}(m \times n, K)$ ist $f_A : K^n \longrightarrow K^m$ eine K-lineare Abbildung.

Beweis: Aus Lemma A2.8 folgt für $x, y \in K^n$ und $\lambda \in K$

$$f_A(x + y) = A(x + y) = Ax + Ay = f_A(x) + f_A(y)$$

und

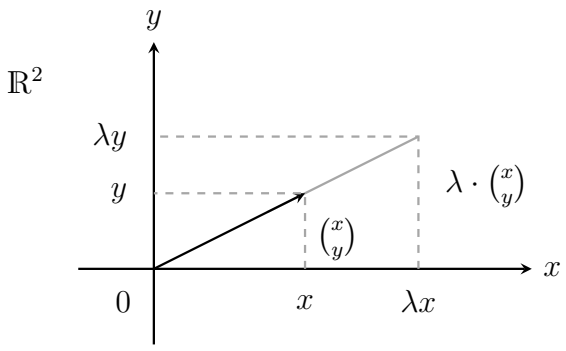
$$f_A(\lambda x) = A(\lambda x) = \lambda(Ax) = \lambda f_A(x).$$

□

Beispiel A3.24

a. Im Fall $n = 1$ und $A = (a)$ ist die K-lineare Abbildung $f_A : K \to K : x \mapsto a \cdot x$ gerade die Multiplikation mit a.

b. Die lineare Abbildung $f_A : \mathbb{R}^2 \mapsto \mathbb{R}^2 : (x, y)^t \mapsto (\lambda x, \lambda y)^t$ zu $A = \lambda \mathbb{1}_2$ ist eine *Streckung* um den Faktor λ.



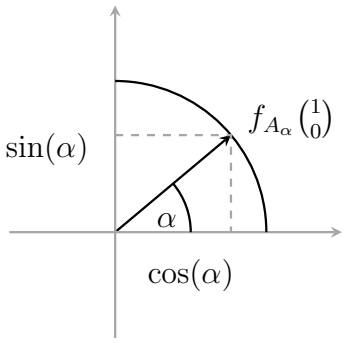
c. Für $\alpha \in \mathbb{R}$ setzen wir

$$A_\alpha := \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Dann ist die lineare Abbildung $f_{A_\alpha} : \mathbb{R}^2 \to \mathbb{R}^2$ eine Drehung um den Winkel α. Beachte dazu, daß

$$A_\alpha e_1 = \big(\cos(\alpha), \sin(\alpha)\big)^t \quad \text{und} \quad A_\alpha e_2 = \big(-\sin(\alpha), \cos(\alpha)\big)^t,$$

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: $f_{A_\alpha}\big((x,y)^t\big) = xf_{A_\alpha}(e_1) + yf_{A_\alpha}(e_2)$.

d. Ist $n \geq m$, so ist die Abbildung

$$\mathrm{pr} : K^n \to K^m : (x_1,\ldots,x_n)^t \mapsto (x_1,\ldots,x_m)^t$$

eine K-lineare Abbildung, genannt die kanonische *Projektion*.
Ist $m \geq n$, dann ist die kanonische *Inklusion*

$$i_{K^n} : K^n \to K^m : (x_1,\ldots,x_n)^t \mapsto (x_1,\ldots,x_n,0,\ldots,0)^t$$

ebenfalls K-linear. Beides prüft man leicht nach.

Proposition A3.25 (Kern und Bild sind Unterräume)
Es seien V und W K-Vektorräume und $f : V \longrightarrow W$ sei K-linear.

a. Ist U ein Unterraum von V, so ist $f(U)$ ein Unterraum von W.

b. Ist U ein Unterraum von W, so ist $f^{-1}(U)$ ein Unterraum von V.

c. Das *Bild* $\mathrm{Im}(f) = f(V)$ von f ist ein Unterraum von W.

d. Der *Kern* von f, $\mathrm{Ker}(f) = \{x \in V \mid f(x) = 0\}$, ist ein Unterraum von V.

Beweis:

a. 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 $\lambda \in K$ 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

$$\lambda u = \lambda f(x) = f(\lambda x) \in f(U).$$

Also ist $f(U)$ ein Unterraum von W.

b. Es sei U ein Unterraum von W. Dann ist $0_W \in U$ und wegen $f(0_V) = 0_W$ ist dann $0_V \in f^{-1}(U)$, so daß $f^{-1}(U)$ nicht leer ist. Sind $\lambda \in K$ und $x,y \in f^{-1}(U)$, so gilt

$f(x), f(y) \in U$ und somit

$$f(x+y) = f(x) + f(y) \in U$$

und

$$f(\lambda x) = \lambda f(x) \in U.$$

Also auch $x + y \in f^{-1}(U)$ und $\lambda x \in f^{-1}(U)$, so daß $f^{-1}(U)$ ein Unterraum von V ist.

c. Dies folgt aus a. mit $U = V$.

d. Dies folgt aus b. mit $U = \{0_W\}$.

□

Beispiel A3.26

Die Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x_1, x_2) \mapsto x_1 - x_2$ aus Beispiel A3.21 hat den Kern

$$\mathrm{Ker}(f) = \{x \in \mathbb{R}^2 \mid f(x) = x_1 - x_2 = 0\} = \left\{(\lambda, \lambda)^t \;\middle|\; \lambda \in K\right\}$$

und das Bild ist $\mathrm{Im}(f) = \mathbb{R}$.

Proposition A3.27 (Injektivität linearer Abbildungen)

Eine lineare Abbildung $f : V \longrightarrow W$ ist genau dann injektiv, wenn

$$\mathrm{Ker}(f) = \{0_V\}.$$

Beweis: Die Aussage folgt unmittelbar aus Lemma A1.12, da f auch ein Gruppenhomomorphismus der additiven Gruppen ist. □

E) Faktorräume

Definition A3.28 (Faktorraum)

Es sei V ein K-Vektorraum und U ein Unterraum von V.

a. Für $x \in V$ heißt

$$\overline{x} := x + U := \{x + u \mid u \in U\}$$

die *Restklasse* oder *Nebenklasse* von x modulo U und x heißt ein *Vertreter* der Restklasse. Man nennt $x + U$ auch einen *affinen Raum* parallel zum Unterraum U mit Aufpunkt x.

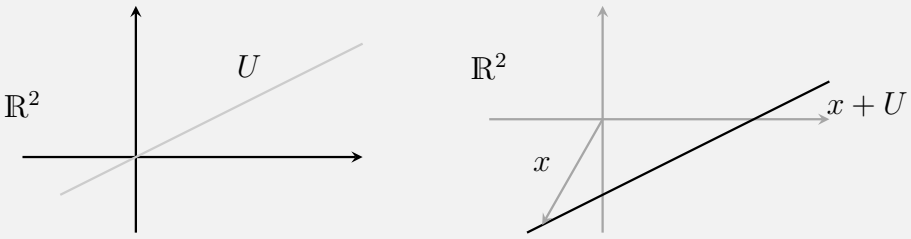


Abbildung 7: Ein affiner Raum $x + U$ zum Unterraum U.

Man beachte, daß aus der Notation $\overline{x}$ 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 U

$$V/U := \{x + U \mid x \in V\} = \{\overline{x} \mid x \in V\}$$

auch den *Faktorraum* von V modulo U.

Bemerkung A3.29 (Restklassen als Äquivalenzklassen)
Ist U ein Unterraum des K-Vektorraums V, so wird durch

$$x \sim y \iff x - y \in U$$

für $x, y \in V$ eine *Äquivalenzrelation* auf V definiert, wie man leicht nachprüft. Die Äquivalenzklasse von x ist dann gerade $x + U$, und V/U ist die Menge aller Äquivalenzklassen der Äquivalenzrelation. Dies ist konzeptionell der richtige Weg, die Restklassen von V modulo U sowie den Faktorraum einzuführen.

Beweis: Wir wollen zunächst nachweisen, daß $\sim$ eine Äquivalenzrelation ist. Dazu wählen wir $x, y, z \in V$. Wegen $x - x = 0 \in U$ gilt $x \sim x$ und $\sim$ ist reflexiv. Gilt $x \sim y$, d.h. $x - y \in U$, so gilt auch $y - x = -(x - y) \in U$ und somit $y \sim x$, d.h. $\sim$ ist auch symmetrisch. Gilt schließlich $x \sim y$ und $y \sim z$, d.h. $x - y, y - z \in U$, so gilt ebenfalls $x - z = (x - y) + (y - z) \in U$ und somit $x \sim z$, d.h. $\sim$ ist auch transitiv. Wir haben also gezeigt, daß $\sim$ eine Äquivalenzrelation ist, und hierbei haben wir die Unterraumeigenschaften von U ausgenutzt.

Es bleibt noch für $x \in V$ zu zeigen, daß $x + U$ die zu x gehörige Äquivalenzklasse ist. Sei dazu zunächst y in der Äquivalenzklasse von x gegeben, d.h. $y \sim x$ und damit $y - x \in U$. Dann ist aber auch $y = x + (y - x) \in x + U$. Sei nun umgekehrt $y \in x + U$ gegeben. Dann gibt es ein $u \in U$ mit $y = x + u$ und somit ist $y - x = u \in U$ und $y \sim x$, was zur Folge hat, daß y zur Äquivalenzklasse von x gehört. □

Lemma A3.30 (Rechnen mit Restklassen)
Es sei V ein K-Vektorraum, U ein Unterraum, $x, x', y, y' \in V$ und $\lambda \in K$. In V/U gelten dann die folgenden Aussagen:

a. Entweder $\overline{x} = \overline{y}$ oder $\overline{x} \cap \overline{y} = \emptyset$.

b. Es gilt:
$$\overline{x} = \overline{y} \quad \Longleftrightarrow \quad x - y \in U.$$
Insbesondere, $\overline{x} = \overline{0} = U$ genau dann, wenn $x \in U$.

c. Gilt $\overline{x} = \overline{x'}$ und $\overline{y} = \overline{y'}$, so gelten auch
$$\overline{x + y} = \overline{x' + y'} \quad \text{und} \quad \overline{\lambda x} = \overline{\lambda x'}.$$

Beweis:

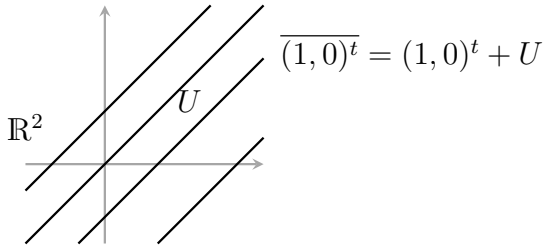
a. Dies folgt unmittelbar aus Proposition 6.10, da $\overline{x}$ und $\overline{y}$ nach Bemerkung A3.29 Äquivalenzklassen sind.

b. Dies folgt aus Bemerkung A3.29.

c. Wir wollen hier zur Verdeutlichung $x + U$ und $y + U$ statt $\overline{x}$ und $\overline{y}$ schreiben. Aus $x + U = x' + U$ sowie $y + U = y' + U$ folgt nach b.
$$x - x', y - y' \in U.$$
Damit gilt dann auch
$$(x + y) - (x' + y') = (x - x') + (y - y') \in U$$
und
$$(\lambda x - \lambda x') = \lambda \cdot (x - x') \in U.$$
Wegen b. gilt dann wieder $(x + y) + U = (x' + y') + U$ und $\lambda x + U = \lambda x' + U$.

□

Beispiel A3.31
Sei $V = \mathbb{R}^2$ und $U = \mathrm{Lin}\,((1,1)^t)$. Dann gilt
$$\overline{(1,0)^t} = (1,0)^t + U = (2,1)^t + U = \overline{(2,1)^t}$$
und
$$\overline{(1,0)^t} = (1,0)^t + U \neq (0,1)^t + U = \overline{(0,1)^t}.$$
Die Restklassen in V/U sind in diesem Fall genau die zu U parallelen Geraden.



Satz A3.32 (Der Faktorraum ist ein Vektorraum.)
Es sei V ein K-Vektorraum und U ein Unterraum. Dann definiert

$$\overline{x} + \overline{y} := \overline{x+y} \tag{190}$$

und

$$\lambda \cdot \overline{x} := \overline{\lambda x} \tag{191}$$

für $\overline{x}, \overline{y} \in V/U$ 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

$$\pi : V \longrightarrow V/U : x \mapsto \overline{x}$$

eine surjektive K-lineare Abbildung mit $\mathrm{Ker}(\pi) = U$, die wir die *Restklassenabbildung* nennen.

Beweis: Bevor wir uns den Vektorraumaxiomen widmen können, müssen wir zeigen, daß durch (190) und (191) 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 $\overline{x} = \overline{x'}$ und $\overline{y} = \overline{y'}$ auch $\overline{x+y} = \overline{x'+y'}$ und $\overline{\lambda x} = \overline{\lambda x'}$ gelten. Dies folgt aber aus Lemma A3.30.

Wir wollen nun zeigen, daß V/U den Vektorraumaxiomen genügt. Dazu seien $\overline{x}, \overline{y}, \overline{z} \in V/U$ und $\lambda, \mu \in K$ gegeben. Dann gilt

$$\big(\overline{x} + \overline{y}\big) + \overline{z} = \overline{x+y} + \overline{z} = \overline{(x+y)+z} = \overline{x+(y+z)} = \overline{x} + \overline{y+z} = \overline{x} + \big(\overline{y} + \overline{z}\big).$$

Die Addition ist also assoziativ. Zudem gilt

$$\overline{0} + \overline{x} = \overline{0+x} = \overline{x}$$

und

$$\overline{-x} + \overline{x} = \overline{-x+x} = \overline{0},$$

so daß $\overline{0}$ der Nullvektor ist und $\overline{x}$ ein Inverses besitzt. Da zudem

$$\overline{x} + \overline{y} = \overline{x+y} = \overline{y+x} = \overline{y} + \overline{x}$$

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.

$$(\lambda+\mu)\cdot\overline{x} = \overline{(\lambda+\mu)\cdot x} = \overline{\lambda x+\mu x} = \overline{\lambda x}+\overline{\mu x} = \lambda\cdot\overline{x}+\mu\cdot\overline{x}$$

und

$$\lambda\cdot\overline{x+y} = \overline{\lambda\cdot(x+y)} = \overline{\lambda x+\lambda y} = \overline{\lambda x}+\overline{\lambda y} = \lambda\cdot\overline{x}+\lambda\cdot\overline{y}$$

und

$$(\lambda\cdot\mu)\cdot\overline{x} = \overline{(\lambda\cdot\mu)\cdot x} = \overline{\lambda\cdot(\mu\cdot x)} = \lambda\overline{\mu\cdot x} = \lambda\cdot\big(\mu\cdot\overline{x}\big)$$

und

$$1\cdot\overline{x} = \overline{1\cdot x} = \overline{x}.$$

Also ist V/U ein K-Vektorraum.

Es bleibt, die Aussagen zur Restklassenabbildung π zu zeigen. Die Linearität von π folgt aus der Definition der Operationen auf V/U. Sind $x, y \in V$ und $\lambda \in K$, dann gilt

$$\pi(x+y) = \overline{x+y} = \overline{x}+\overline{y} = \pi(x)+\pi(y)$$

und

$$\pi(\lambda x) = \overline{\lambda x} = \lambda\cdot\overline{x} = \lambda\cdot\pi(x).$$

Außerdem ist π surjektiv, da jedes $\overline{x} \in V/U$ sich schreiben läßt als $\overline{x} = \pi(x)$. Ds gilt

$$x \in \operatorname{Ker}(\pi) \iff \overline{x} = \pi(x) = \overline{0} \iff x \in U.$$

Damit sind alle Aussagen des Satzes bewiesen. □

Bemerkung A3.33 (Die vier Rechenregeln für den Faktorraum)
Um mit dem Faktorraum rechnen zu können, braucht man nur die Rechenregeln:

a. $\overline{0} = 0+U = U$ ist der Nullvektor.

b. $\overline{x}+\overline{y} = \overline{x+y}$.

c. $\lambda\cdot\overline{x} = \overline{\lambda x}$.

d. $\overline{x} = \overline{y} \iff x-y \in U$.

Satz A3.34 (Homomorphiesatz)
Ist $f : V \longrightarrow W$ eine K-lineare Abbildung, so ist

$$\overline{f} : V/\operatorname{Ker}(f) \longrightarrow \operatorname{Im}(f) : \overline{x} \mapsto f(x)$$

ein Isomorphismus. Insbesondere gilt also $V/\operatorname{Ker}(f) \cong \operatorname{Im}(f)$.

Beweis: Da wir für die Definition von $\overline{f}(\overline{x})$ 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 $\overline{f}$ zeigen.

Seien dazu $\overline{x} = x + \operatorname{Ker}(f) = y + \operatorname{Ker}(f) = \overline{y}$ gegeben. Dann gilt

$$x - y \in \operatorname{Ker}(f),$$

und mithin $0 = f(x-y) = f(x) - f(y)$, oder alternativ $f(x) = f(y)$. Die Abbildung $\overline{f}$ ist also wohldefiniert.

Die Linearität von $\overline{f}$ folgt dann aus der Linearität von f. Seien dazu $\overline{x}, \overline{y} \in V/\operatorname{Ker}(f)$ und $\lambda \in K$, dann gilt

$$\overline{f}\big(\overline{x}+\overline{y}\big) = \overline{f}\big(\overline{x+y}\big) = f(x+y) = f(x) + f(y) = \overline{f}\big(\overline{x}\big) + \overline{f}\big(\overline{y}\big)$$

und

$$\overline{f}\big(\lambda \cdot \overline{x}\big) = \overline{f}\big(\overline{\lambda \cdot x}\big) = f(\lambda \cdot x) = \lambda \cdot f(x) = \lambda \cdot \overline{f}\big(\overline{x}\big).$$

Es bleibt noch zu zeigen, daß $\overline{f}$ surjektiv und injektiv ist.

Ist $y \in \operatorname{Im}(f)$, so gibt es ein $x \in V$ mit $f(x) = y$, und somit gilt

$$y = f(x) = \overline{f}\big(\overline{x}\big).$$

Also ist $\overline{f}$ surjektiv.

Für die Injektivität nutzen wir Proposition A3.27. Es gilt

$$\overline{x} \in \operatorname{Ker}\big(\overline{f}\big) \iff 0 = \overline{f}\big(\overline{x}\big) = f(x) \iff x \in \operatorname{Ker}(f) \iff \overline{x} = \overline{0}.$$

Also enthält der Kern von $\overline{f}$ nur den Nullvektor, und somit ist $\overline{f}$ injektiv. □

F) Direkte Komplemente

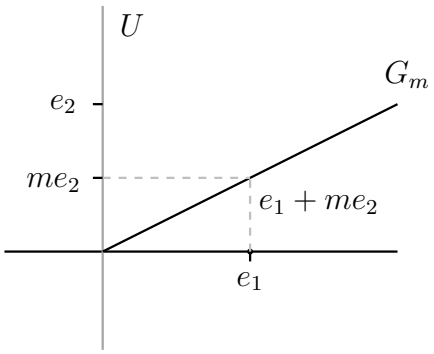
Definition A3.35 (Direkte Komplemente)
Es sei V ein K-Vektorraum und U und U' seien Unterräume von V. Dann heißt U' ein (direktes) *Komplement* von U, falls $V = U \oplus U'$.

Beispiel A3.36 (Komplemente sind nicht eindeutig.)
Ist $V = \mathbb{R}^2$ und $U = \operatorname{Lin}(e_2)$ die y-Achse, dann ist die Ursprungsgerade mit Steigung m

$$G_m := \operatorname{Lin}(e_1 + me_2)$$

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 A3.37 (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

$$\pi_| : U' \to V/U : x \mapsto \overline{x}$$

auf U' ein Isomorphismus. Insbesondere sind je zwei Komplemente von U isomorph.

Beweis: Es ist klar, daß $\pi_|$ als Einschränkung einer K-linearen Abbildung wieder K-linear ist.

Wir zeigen zunächst, daß $\pi_|$ surjektiv ist. Sei dazu $\overline{x} \in V/U$ gegeben. Wegen $V = U \oplus U'$ läßt sich x als $x = y + z$ mit $y \in U$ und $z \in U'$ schreiben. Damit gilt:

$$\overline{x} = \overline{z} = \pi_|(z) \in \operatorname{Im}\left(\pi_|\right).$$

Also ist $\pi_|$ surjektiv.

Es bleibt zu zeigen, daß $\pi_|$ injektiv ist, d. h. $\operatorname{Ker}\left(\pi_|\right) = \{0\}$. Sei dazu $z \in \operatorname{Ker}\left(\pi_|\right)$, dann gilt

$$\overline{0} = \pi_|(z) = \overline{z}.$$

D. h. $z \in U$. Damit gilt aber $z \in U \cap U' = \{0\}$, also $z = 0$.

Seien schließlich U' und U'' zwei Komplemente von U, dann ist die Komposition

$$U' \xrightarrow{\pi_{|U'}} V/U \xrightarrow{\pi^{-1}_{|U''}} U''$$

ein Isomorphismus von U' nach U'' und die Komplemente sind isomorph zueinander. □

Bemerkung A3.38
Da V/U isomorph zu jedem Komplement von U ist, kann man in vielen Situationen ein Komplement schlicht durch den Faktorraum ersetzen. Während es sehr viele Komplemente von U geben kann, gibt es nur *einen* Faktorraum. Dieser ist durch U eindeutig bestimmt. Das ist unter Umständen ein großer Vorteil!

G) Ringe und Moduln

Bemerkung A3.39 (Kommutative Ringe mit Eins)

Eine Menge 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.

$$\mathbb{R}[t] = \left\{ \sum_{k=0}^{n} a_k \cdot t^k \;\middle|\; a_k \in \mathbb{R}, n \in \mathbb{N} \right\},$$

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 A3.40 (Matrizen über Ringen)

Setzen wir in Abschnitt A2 nur voraus, daß K ein kommutativer Ring mit Eins ist, so bleiben alle Definitionen und Aussagen korrekt.

Bemerkung A3.41 (Moduln und lineare Abbildungen)

Setzen wir in Definition A3.1 nur voraus, daß K ein kommutativer Ring mit Eins ist, so nennen wir V einen *K-Modul*. Der Name ist dabei aber auch das einzige, was sich ändert. Entsprechend wird man in Definition A3.4 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:

$$\lambda \cdot v = 0 \quad \not\Longrightarrow \quad \lambda = 0 \text{ oder } v = 0.$$

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 A3.42

Wir betrachten die Menge $V = \mathbb{R}_{>0}$ aller positiven reellen Zahlen und definieren für $x, y \in V$ und $\lambda \in \mathbb{R}$

$$x \oplus y := x \cdot y$$

und

$$\lambda \odot y := y^\lambda = \exp(\lambda \cdot \ln(y)).$$

Überprüfe, ob $(V, \oplus, \odot)$ ein $\mathbb{R}$-Vektorraum ist.

Aufgabe A3.43

Welche der folgenden Teilmengen von K^3 sind Unterräume des K^3? Begründe Deine Anworten.

a. $\{(x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid x_1 \cdot x_2 = 2x_3\}$ für $K = \mathbb{R}$.

b. $\{(x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid ax_1 + x_2 + x_3 = a + 1\}$ für ein festes $a \in \mathbb{R}$ für $K = \mathbb{R}$.

c. $\{(x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid x_1 \leq 0\}$ für $K = \mathbb{R}$.

d. $\{(1, 0, 0)^t, (0, 1, 0)^t, (1, 1, 0)^t, (0, 0, 0)^t\}$ für $K = \mathbb{R}$ oder $K = \mathbb{F}_2$.

Aufgabe A3.44

Gegeben seien die folgenden Teilmengen des $\mathbb{Q}$–Vektorraums $V = \mathbb{Q}^4$:

a. $U_1 = \{(x, x + 1, x + 2, x + 4)^t \mid x \in \mathbb{Q}\}$,

b. $U_2 = \{(x, 2x, 3x, 4x)^t \mid x \in \mathbb{Q}\}$,

c. $U_3 = \{(x_1, x_2, x_3, x_4)^t \mid x_1, x_2, x_3, x_4 \in \mathbb{Q},\ x_3 = x_1 + x_2,\ x_4 = x_2 + x_3\}$,

d. $U_4 = \{(x_1, x_2, x_3, x_4)^t \mid x_1, x_2, x_3, x_4 \in \mathbb{Q},\ x_2 \geq 0\}$,

e. $U_5 = \{(x_1, x_2, x_3, x_4)^t \mid x_1, x_2, x_3, x_4 \in \mathbb{Z}\}$.

Welche dieser Mengen sind Unterräume von V? Begründe Deine Aussage.

Aufgabe A3.45

Seien U, V und W K-Vektorräume, $\lambda, \lambda' \in K$ und $f, f' \in \mathrm{Hom}_K(U, V)$ und $g, g' \in \mathrm{Hom}_K(V, W)$. Dann gelten:

a. $f + f', \lambda \cdot f \in \mathrm{Hom}_K(U, V)$, d.h. $\mathrm{Hom}_K(U, V)$ ist ein Unterraum von V^U.

b. $g \circ (\lambda f + \lambda' f') = \lambda(g \circ f) + \lambda'(g \circ f')$ 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 A3.46

Seien V, W K–Vektorräume und $f, g \in \mathrm{Hom}_K(V, W)$. Zeige

$$\mathrm{Ker}(f+g) \supseteq \mathrm{Ker}(f) \cap \mathrm{Ker}(g)$$

und

$$\mathrm{Im}(f+g) \subseteq \mathrm{Im}(f) + \mathrm{Im}(g).$$

Finde außerdem Beispiele, so dass die Inklusionen strikt sind.

Aufgabe A3.47 (Direkte Summe)

Seien $U_1, \ldots, U_n$ Unterräume von V. Zeige, die Summe $U_1 + \ldots + U_n$ ist genau dann eine direkte Summe, wenn der Nullvektor sich auf eindeutige Weise als Summe von Vektoren in $U_1, \ldots, U_n$ schreiben läßt.

Aufgabe A3.48 (f-invariante Unterräume)

Ist V ein K-Vektorraum, $f : V \longrightarrow V$ 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

$$f_U : U \longrightarrow U : x \mapsto f(x)$$

und

$$f_{V/U} : V/U \longrightarrow V/U : \overline{x} \mapsto \overline{f(x)}$$

K-lineare Abbildungen definiert werden.

Aufgabe A3.49 (Erster Isomorphiesatz)

Sei V ein K–Vektorraum und $U, U' \leq V$. Zeige

$$U/(U \cap U') \cong (U + U')/U'.$$

Aufgabe A3.50 (Projektionen)

Es sei V ein K-Vektorraum. $f \in \mathrm{End}_K(V)$ heißt *Projektion*, falls $f^2 = f$ gilt.
Zeige, die folgenden Aussagen sind äquivalent:

a. f ist eine Projektion,

b. $\mathrm{id}_V - f$ ist eine Projektion,

c. $\mathrm{Im}\left(\mathrm{id}_V - f\right) = \mathrm{Ker}(f)$,

d. $\mathrm{Ker}\left(\mathrm{id}_V - f\right) = \mathrm{Im}(f)$.

Zeige auch, sind obige Bedingungen erfüllt, so gilt zudem $V = \mathrm{Ker}(f) \oplus \mathrm{Im}(f)$.

Aufgabe A3.51 (Innere direkte Summe)

Sei $(U_i)_{i\in I}$ eine Familie von Unterräumen des K-Vektorraums V.

a. Dann gilt

$$\sum_{i\in I} U_i := \operatorname{Lin}\left(\bigcup_{i\in I} U_i\right) \stackrel{!}{=} \left\{ \sum_{\substack{i\in I\\ \text{endlich}}} x_i \mid x_i \in U_i \right\}.$$

Man nennt diesen Unterraum die *Summe* der Unterräume U_i.

b. In $\sum_{i\in I} U_i$ ist genau dann jeder Vektor auf eindeutige Weise als Summe von Vektoren in den U_i darstellbar, wenn dies für den Nullvektor gilt. In dem Fall nennt man die Summe eine *(innere) direkte Summe* und schreibt

$$\bigoplus_{i\in I} U_i = \sum_{i\in I} U_i.$$

Aufgabe A3.52 (Äußere und innere direkte Summe)

Es sei $(V_i \mid i \in I)$ eine Familie von K-Vektorräumen. Zeige die folgenden Aussagen:

a. Zeige, die natürliche Abbildung

$$\delta_j : V_j \longrightarrow \bigoplus_{i\in I} V_i : x \mapsto (x_i \mid x_j = x, x_i = 0 \,\forall\, i \neq j)$$

ist ein Vektorraummonomorphism.

b. Die äußere direkte Summe $\bigoplus_{i\in I} V_i$ ist die *innere* direkte Summe der $\delta_i(V_i)$.

Aufgabe A3.53 (Äußere und innere direkte Summe)

Sei V ein K-Vektorraum und $(V_i \mid i \in I)$ eine Familie von Unterräumen von V. Ferner bezeichne $U = \bigoplus_{i\in I} V_i$ die äußere direkte Summe der Vektorräume V_i, $i \in I$, und $W = \sum_{i\in I} V_i$ die Summe der V_i in V.
Zeige, genau dann ist W eine innere direkte Summe, wenn die Abbildung

$$\delta : U \longrightarrow W : (x_i)_{i\in I} \mapsto \sum_{\substack{i\in I\\ \text{endlich}}} x_i$$

ein Isomorphismus ist.

Aufgabe A3.54

Sei M eine Menge und $\mathcal{P}(M) = \{A \mid A \subseteq M\}$ die Potenzmenge von M. Wir definieren auf $\mathcal{P}(M)$ eine Addition mittels der symmetrischen Differenz

$$A \triangle B := (A \cup B) \setminus (A \cap B)$$

für $A, B \in \mathcal{P}(M)$ und eine Multiplikation mit Skalaren aus dem Körper $\mathbb{Z}_2 = \{\overline{0}, \overline{1}\}$ durch

$$\lambda \cdot A := \begin{cases} A, & \text{wenn } \lambda = \overline{1}, \\ \emptyset, & \text{wenn } \lambda = \overline{0}. \end{cases}$$

Zeige, daß $(\mathcal{P}(M), \triangle, \cdot)$ ein $\mathbb{Z}_2$-Vektorraum ist.

Aufgabe A3.55

Zeige, die Mengen

$$U = \{(x_1, x_2, x_3)^t \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$$

und

$$U' = \text{Lin}\left((1, 0, 0)^t\right)$$

sind Unterräume von $\mathbb{R}^3$ und es gilt $\mathbb{R}^3 = U \oplus U'$.

Aufgabe A3.56

Es sei $V = \mathbb{R}^{\mathbb{R}}$ der $\mathbb{R}$-Vektorraum aller Abbildungen von $\mathbb{R}$ nach $\mathbb{R}$ und

$$U = \{f : \mathbb{R} \longrightarrow \mathbb{R} \mid f(-x) = -f(x) \; \forall \, x \in \mathbb{R}\}$$

sowie

$$U' = \{f : \mathbb{R} \longrightarrow \mathbb{R} \mid f(-x) = f(x) \; \forall \, x \in \mathbb{R}\}.$$

Zeige, daß U und U' Unterräume von V sind mit $V = U \oplus U'$.

Aufgabe A3.57

Berechne das Bild und den Kern der linearen Abbildung $f_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$, die durch die folgende Matrix gegeben ist:

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \in \text{Mat}_2(\mathbb{R}).$$

Aufgabe A3.58

Zeige, daß die Abbildung

$$f : \mathbb{Z}_2^3 \longrightarrow \mathbb{Z}_2^3 : \left(\overline{x}, \overline{y}, \overline{z}\right)^t \mapsto \left(\overline{x} - \overline{y}, \overline{z}^2, \overline{x} - \overline{y}\right)^t$$

linear ist und berechne Bild und Kern von f.

Aufgabe A3.59

Berechne das Bild und den Kern der linearen Abbildung

$$D : P_2 \longrightarrow P_2 : p \mapsto p' - p'',$$

wobei P_2 der $\mathbb{R}$-Vektorraum der Polynome vom Grad kleiner oder gleich zwei ist und p' bzw. p'' die Ableitung bzw. zweite Ableitung des Polynoms p bezeichnet.

Aufgabe A3.60

Interpretiere die Wirkung der linearen Abbildung $f_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ geometrisch, wenn A eine der folgenden Matrizen ist:

$$\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}.$$

Wende die Abbildung dazu zunächst auf die Vektoren $e_1 = (1,0)^t$, $e_2 = (0,1)^t$ und $x = (2,3)^t$ an.

Aufgabe A3.61

Welche der folgenden Abbildungen ist eine lineare Abbildung?

a. $f : K^3 \longrightarrow K^2 : (x,y,z)^t \mapsto (x + y \cdot z, x - y + z)^t$ mit $K = \mathbb{R}$ oder $K = \mathbb{F}_2$.

b. $f : K^3 \longrightarrow K^3 : (x,y,z)^t \mapsto (2x - z, x - 2y + z, 3y - z)^t$ mit $K = \mathbb{Q}$ oder $K = \mathbb{F}_5$.

c. $f : K^2 \longrightarrow K^2 : (x,y)^t \mapsto (x + y, 2x - y + 3)^t$ mit $K = \mathbb{C}$, $K = \mathbb{F}_5$ oder $K = \mathbb{F}_3$.

d. $I : \mathcal{C}([0,1], \mathbb{R}) \longrightarrow \mathbb{R}$ mit

$$I(f) := \int_0^1 f(t)\, dt.$$

A4 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 A4.1 (Familien)
Es seien I und X zwei Mengen.

a. Wir nennen ein Tupel der Form $F = (x_i)_{i\in I}$ 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. 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 $i \in I$.

c. 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 $F' \subseteq F$. Ebenso schreiben wir $x \in F$, um auszudrücken, daß $x = x_i$ für ein $i \in I$ gilt.

d. Wir schreiben kurz $\mathrm{Lin}\,(F)$ für die lineare Hülle $\mathrm{Lin}\,(\{x_i \mid i \in I\})$ und nennen $\mathrm{Lin}\,(F)$ die *lineare Hülle* von F.

Beispiel A4.2
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 $(x_1, x_2, x_3) = \big((1, 0)^t, (1, 1)^t, (1, 0)^t\big)$ definiert. $(x_1, x_3) = \big((1, 0)^t, (1, 0)^t\big)$ ist eine Teilfamilie.

Bemerkung A4.3 (Familien von Vektoren)

a. In einer Familie können Elemente auch *mehrfach* auftreten, in einer Menge geht das nicht. Z.B. $F = \big((1, 0)^t, (1, 1)^t, (1, 0)^t\big)$.

b. Ist die Menge I geordnet, so ordnen wir die Mitglieder der Familie F in der gleichen Weise. Z.B. $\big((1, 0)^t, (1, 1)^t, (1, 0)^t\big) \neq \big((1, 0)^t, (1, 0)^t, (1, 1)^t\big)$.

c. In unseren Anwendungen wird die Menge I meist $\{1, \ldots, n\}$ 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 $F : I \longrightarrow X : i \mapsto x_i$ angeben. Die Tupelschreibweise ist aber suggestiver als die Schreibweise als Abbildung.

Definition A4.4 (Lineare Unabhängigkeit)

Es sei V ein K-Vektorraum und F eine Familie von Vektoren in V.

a. Eine endliche Familie $(x_1, \ldots, x_n)$ von Vektoren in V heißt *linear unabhängig*, wenn aus

$$\lambda_1 \cdot x_1 + \ldots + \lambda_n \cdot x_n = 0$$

stets

$$\lambda_1 = \ldots = \lambda_n = 0$$

folgt, d.h. wenn nur die triviale Linearkombination der x_i Null ergibt.
Wir sagen dann oft einfach, die Vektoren $x_1, \ldots, x_n$ seien linear unabhängig.

b. Eine endliche Familie $(x_1, \ldots, x_n)$ von Vektoren in V heißt *linear abhängig*, wenn es Skalare $\lambda_1, \ldots, \lambda_n \in K$ gibt, so daß

$$\lambda_1 \cdot x_1 + \ldots + \lambda_n \cdot x_n = 0,$$

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 $x_1, \ldots, x_n$ 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 A4.5 (Lineare Unabhängigkeit)

a. Die Einheitsvektoren $e_1, \ldots, e_n \in K^n$ sind linear unabhängig. Denn aus

$$(0, \ldots, 0)^t = \lambda_1 \cdot e_1 + \ldots + \lambda_n \cdot e_n = (\lambda_1, \ldots, \lambda_n)^t$$

folgt unmittelbar $\lambda_1 = \ldots = \lambda_n = 0$.

b. Die Familie $\big((1,0)^t, (0,1)^t, (1,1)^t\big)$ 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 $(e_{k_1}, \ldots, e_{k_l})$ für $0 \leq k_1 < \ldots < k_l$. Dann folgt aus

$$\lambda_{k_1} \cdot e_{k_1} + \ldots + \lambda_{k_l} \cdot e_{k_l} = (0)_{n\in\mathbb{N}}$$

unmittelbar $\lambda_{k_1} = \ldots = \lambda_{k_l} = 0$, 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 A4.6 (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. □

Beispiel A4.7
In Beispiel A4.5 b. gilt

$$(1,0)^t = -(0,1)^t + (1,1)^t,$$

woraus ebenfalls die lineare Abhängigkeit der Familie folgt.

Notation A4.8 (Linearkombination)
Sei $F = (x_i)_{i\in I}$ eine Familie von Vektoren in V und I sei nicht notwendigerweise endlich. Wir werden des öfteren

$$x = \sum_{\substack{i\in I\\ \text{endlich}}} \lambda_i \cdot x_i \tag{192}$$

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 (192) 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 A4.9 (Ergänzung linear unabhängiger Familien)
Ist $B = (x_i)_{i\in I}$ eine linear unabhängige Familie in V mit $\mathrm{Lin}\,(B) \subsetneqq V$, so ist die Familie $(x, x_i \mid i \in I)$ für jedes $x \in V \setminus \mathrm{Lin}\,(B)$ 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

$$x = \sum_{\substack{i\in I\\ \text{endlich}}} -\frac{\lambda_i}{\lambda} \cdot x_i \in \mathrm{Lin}\,(B)$$

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. □

B) Erzeugendensysteme und Basen

Definition A4.10 (Erzeugendensystem und Basis)
Es sei V ein K-Vektorraum und F eine Familie von Vektoren in V.

a. F heißt ein *Erzeugendensystem* von V, wenn $V = \mathrm{Lin}(F)$, d.h. wenn jeder Vektor in V eine Linearkombination von Vektoren in F ist.

b. F heißt eine *Basis* von V, wenn F ein linear unabhängiges Erzeugendensystem von V ist.

c. V heißt *endlich erzeugt*, wenn V ein endliches Erzeugendensystem besitzt.

Beispiel A4.11 (Erzeugendensystem und Basis)

a. Die Familie $B = (e_1, \ldots, e_n)$ 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 A4.5 ist B linear unabhängig und zudem ist ein beliebiger Vektor $x = (x_1, \ldots, x_n)^t \in K^n$ eine Linearkombination

$$x = x_1 \cdot e_1 + \ldots + x_n \cdot e_n$$

der Vektoren in B.

b. Analog sieht man, daß für $n, m \geq 1$ die Familie

$$\left(E_i^j \mid i = 1, \ldots, m; j = 1, \ldots, n\right),$$

wobei $E_i^j = (e_{lk})_{l=1,\ldots,m;k=1,\ldots,n}$ 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 $\mathrm{Mat}(m \times n, K)$ ist.

c. Die Familie $\left((1,0)^t, (0,1)^t, (1,1)^t\right)$ in Beispiel A4.5 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 A4.5 c. ist *kein* Erzeugendensystem von $\mathbb{R}^{\mathbb{N}}$. Es scheint zwar, als gelte für eine beliebige Folge $(a_n)_{n\in\mathbb{N}}$

$$(a_n)_{n\in\mathbb{N}} = (a_0, a_1, a_2, \ldots) = \sum_{n\in\mathbb{N}} a_n \cdot e_n,$$

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 $B = (t^0, t^1, t^2, \ldots)$ ist eine Basis von $K[t]$.

Proposition A4.12 (Charakterisierung von Basen)
Für eine Familie B von Vektoren in V sind die folgenden Aussagen gleichwertig:

a. B ist eine Basis von V.

b. B ist ein minimales Erzeugendensystem von V.

c. B ist eine maximale linear unabhängige Familie in V.

Bemerkung A4.13
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 unabhänige Familie enthält! Es gibt nicht *die* maximale linear unabhängige Familie.

Beweis von Proposition A4.12: Es sei $B = (x_i)_{i \in I}$.

a. ⇒ b.: Ist B eine Basis, so erzeugt B den Vektorraum V per definitionem. Ist $(x_j \mid j \in J)$ eine echte Teilfamilie von B und ist $i \in I \setminus J$, so gibt es wegen der linearen Unabhängigkeit von B keine Darstellung

$$x_i - \sum_{\substack{j \in J \\ \text{endlich}}} \lambda_j x_j = 0$$

also ist $x_i \notin \operatorname{Lin}(x_j \mid j \in J)$.

b. ⇒ c.: Wir zeigen zunächst, daß B linear unabhängig ist. Angenommen, dies sei nicht der Fall, dann läßt sich nach Lemma A4.6 ein x_i als Linearkombination

$$x_i = \sum_{\substack{i \neq j \in I \\ \text{endlich}}} \lambda_j x_j$$

der übrigen Vektoren in B darstellen. Damit gilt dann aber

$$\operatorname{Lin}(x_j \mid j \in I \setminus \{i\}) = \operatorname{Lin}(x_j \mid j \in I) = V,$$

im Widerspruch zur Minimalität des Erzeugendensystems B.

Sei nun $(x_j \mid j \in J)$ mit $I \subsetneqq J$ 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 A4.6.

c. ⇒ a.: Da B linear unabhängig ist, bleibt zu zeigen, daß $\operatorname{Lin}(B) = V$. Gäbe es ein $x \in V \setminus \operatorname{Lin}(B)$, so wäre wegen Lemma A4.9 auch $(x, x_i \mid i \in I)$ linear unabhängig, im Widerspruch zur Maximalität von B.

□

Beispiel A4.14

Wir betrachten die Familie $B = \big((1,1)^t, (1,-1)^t\big)$ in $\mathbb{R}^2$. 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 $\mathbb{R}^2$.

Bemerkung A4.15 (Existenz von Basen)

Besitzt ein K-Vektorraum ein endliches Erzeugendensystem, so folgt aus Proposition A4.12 unmittelbar, daß er auch ein Basis besitzt, da wir aus dem Erzeugendensystem sukkzessive Elemente entfernen können, bis wir ein minimales Erzeugendensystem erhalten. Man kann in der Tat aber für einen beliebigen K-Vektorraum V die folgenden Aussagen zeigen (siehe Unterabschnitt F):

- Jedes Erzeugendensystem von V enthält eine Basis.
- Jede linear unabhängige Familie in V läßt sich zu einer Basis ergänzen.
- Insbesondere besitzt V eine Basis.
- Jeder Unterraum U von V besitzt ein direktes Komplement.

Für die letzte Aussage wählt man eine Basis B von U ergänzt sie durch eine Familie B' zu einer Basis von V; dann erzeugt B' ein direktes Komplement.

Man beachte, daß Basen durchaus kompliziert sein können. Betrachten wir etwa $\mathbb{R}$ als $\mathbb{Q}$-Vektorraum, so wird eine Basis von $\mathbb{R}$ notwendigerweise überabzählbar viele Elemente enthalten, da jede lineare Hülle von abzählbar vielen Elementen aus $\mathbb{R}$ mit Koeffizienten aus $\mathbb{Q}$ nur abzählbar viele Elemente enthalten kann.

C) Algorithmus zur Berechnung einer Basis

Der folgende Algorithmus beruht auf der Tatsache, daß elementare Zeilenoperationen den Zeilenraum nicht verändern - vgl. Aufgabe A6.52.

Algorithmus A4.16 (Basisberechnung)

INPUT: Ein endliches 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 Nicht-Null-Zeilen als Spaltenvektoren zurück.

Beweis: Weil die elementaren Zeilenoperationen den Zeilenraum nicht ändern, sind die Zeilen der Zeilen-Stufen-Form ein Erzeugendensystem von U.

Die Zeilen-Stufen-Form habe r Zeilen $z_1, \ldots, z_r$. Wir zeigen mit abbrechender Induktion nach k, daß die Familie $(z_{r-k}^t, \ldots, z_r^t)$ linear unabhängig ist, und erhalten damit insbesondere, dass $(z_1^t, \ldots, z_r^t)$ linear unabhängig, also eine Basis von U ist. Für $k = 0$ ist die Aussage korrekt, da z_r^t nicht der Nullvektor ist. Sei nun $(z_{r-k+1}^t, \ldots, z_r^t)$ bereits linear unabhängig. Dann haben alle Vektoren im Erzeugnis $\mathrm{Lin}(z_{r-k+1}^t, \ldots, z_r^t)$ eine Null in der Zeile, die dem Pivot in z_{r-k} entspricht. Also ist z_{r-k} nicht in der linearen Hülle enthalten. Wegen Lemma A4.9 ist die Familie $(z_{r-k}^t, \ldots, z_r^t)$ dann auch linear unabhängig. □

Beispiel A4.17

Betrachte $U = \mathrm{Lin}\left((1,0,-1,2,3)^t, (1,-1,1,4,3)^t, (0,2,-4,-4,0)^t\right) \leq \mathbb{R}^5$:

$$\begin{pmatrix} 1 & 0 & -1 & 2 & 3 \\ 1 & -1 & 1 & 4 & 3 \\ 0 & 2 & -4 & -4 & 0 \end{pmatrix} \longmapsto \ldots \longmapsto \begin{pmatrix} 1 & 0 & -1 & 2 & 3 \\ 0 & -1 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Also ist $B = \left((1,0,-1,2,3)^t, (0,-1,2,2,0)^t\right)$ eine Basis von U.

D) Eindeutige Darstellbarkeit bezüglich einer Basis

Proposition A4.18 (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. □

Beispiel A4.19

Kommen wir auf das Beispiel $B = \big((1,1)^t, (1,-1)^t\big)$ im $\mathbb{R}^2$ zurück. Wir haben bereits gesehen, dass 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. Da B eine Basis ist, ist diese Darstellung auch eindeutig, wie man auch leicht direkt nachprüfen könnte.

Satz A4.20 (Existenz- und Eindeutigkeitssatz 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 $f : V \longrightarrow W$, 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 A4.18 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

$$f(x) := \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot y_i. \tag{193}$$

Wir wollen nun zeigen, daß $f : V \longrightarrow W$ 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

$$\begin{aligned} f\big(\lambda x + \lambda' x'\big) &= f\left(\sum_{\substack{i \in I \\ \text{endlich}}} (\lambda \cdot \lambda_i + \lambda' \cdot \lambda'_i) \cdot x_i \right) \\ &= \sum_{\substack{i \in I \\ \text{endlich}}} (\lambda \cdot \lambda_i + \lambda' \cdot \lambda'_i) \cdot y_i \\ &= \lambda \cdot \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \cdot y_i + \lambda' \cdot \sum_{\substack{i \in I \\ \text{endlich}}} \lambda'_i \cdot y_i \\ &= \lambda \cdot f(x) + \lambda' \cdot f(x'). \end{aligned}$$

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 $g : V \longrightarrow W$ 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 A4.21 (Existenz- und Eindeutigkeitssatz für lineare Abb.)
Satz A4.20 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 (193) sagt man auch, daß sich f aus der Vorschrift $f(x_i) = y_i$, $i \in I$, durch *lineare Fortsetzung* ergibt.

Beispiel A4.22
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 $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ geben mit

$$f\big((1,1)^t\big) = f(x_1) = y_1 = (1,1)^t \quad \text{und} \quad f\big((1,-1)^t\big) = f(x_2) = y_2 = (3,1)^t.$$

Diese besitzt die Abbildungsvorschrift

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x,y)^t \mapsto (2x - y, x)^t.$$

Wir werden später sehen, wie man die Abbildungsvorschrift systematisch bestimmen kann.

Korollar A4.23 (Alle linearen Abb. $K^n \to K^m$ sind von der Form f_A.)
Jede lineare Abbildung $f : K^n \longrightarrow K^m$ ist von der Form $f = f_A$ für eine eindeutig bestimmte Matrix $A \in \mathrm{Mat}(m \times n, K)$.

Beweis: Ist $f : K^n \longrightarrow K^m$ eine lineare Abbildung, so setzen wir $a^i := f(e_i) \in K^m$ für $i = 1, \ldots, n$ und bilden eine Matrix A mit den a^i als Spaltenvektoren. Dann ist f_A eine lineare Abbildung, mit

$$f_A(e_i) = Ae_i = a^i = f(e_i),$$

so daß aus der Eindeutigkeitsaussage in A4.20 unmittelbar $f_A = f$ folgt. Die Eindeutigkeit der Matrix A folgt aus der Tatsache, daß A die Abbildung f_A eindeutig festlegt (siehe Bemerkung A2.7). □

E) Algorithmus zur Berechnung des Bildes von f_A

Ist $f = f_A$ für eine $m \times n$-Matrix A, dann wird das Bild von f von den Spalten von A erzeugt. Wir können eine Basis des Bildes also wie folgt bestimmen.

Algorithmus A4.24 (Bild von f_A)
INPUT: $A \in \mathrm{Mat}(m \times n, K)$.
OUTPUT: Eine Basis von $\mathrm{Im}(f_A)$.

1. **Schritt:** Transponiere A und überführe die Transponierte in ZSF.
2. **Schritt:** Transponiere das Ergebnis wieder und gib die ersten $\mathrm{rang}(A)$ Spaltenvektoren zurück.

F) Existenz von Basen in beliebigen Vektorräumen

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 A4.25 (Ketten)
Es sei $(M, \leq)$ eine teilgeordnete Menge. Eine total geordnete Teilmenge K von M heißt eine *Kette* in M.

Axiom A4.26 (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 A4.27 (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 eine Kette in M. Wir setzen $X := \bigcup_{G \in K} G$. Dann gilt sicher $G \subseteq X$ für alle $G \in K$. Sobald wir gezeigt haben, daß $X \in M$, ist also X eine obere Schranke von K. Wegen $F \subseteq G \subseteq E$ für alle $G \in K$, gilt auch

$$F \subseteq X \subseteq E.$$

Es bleibt also noch zu zeigen, daß X linear unabhängig ist. Seien dazu $(x_1, \ldots, x_r)$ eine endliche Teilfamilie von X. Dann gibt es zu jedem $i \in \{1, \ldots, r\}$ ein $G_i \in K$ mit $x_i \in G_i$. Da K eine Kette ist, gibt es aber ein $G \in K$ mit $G_i \subseteq G$ für alle $i = 1, \ldots, r$, so daß $(x_1, \ldots, x_r)$ eine Teilfamilie der linear unabhängigen Familie G ist. Damit ist auch $(x_1, \ldots, x_r)$ 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 $E \not\subseteq \operatorname{Lin}(B)$ und es gibt ein Element $x \in E \setminus \operatorname{Lin}(B)$. Nach Lemma A4.9 gilt dann aber, daß $(x, y \mid y \in B)$ linear unabhängig ist und somit, wegen $F \subset (x, y \mid y \in B) \subseteq E$,

$$(x, y \mid y \in B) \in M,$$

im Widerspruch zur Maximalität von B. Also war die Annahme falsch und B ist eine Basis. □

Korollar A4.28 (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 A4.27 $F = F'$ und $E = V$, und für b. wähle zu gegebenem E' dort $F = \emptyset$ und $E = E'$. □

Beispiel A4.29

Das Erzeugendensystem $E = \big((1,0)^t, (1,1)^t, (0,1)^t\big)$ von $\mathbb{R}^2$ enthält die kanonische Basis,

und die linear unabhängige Familie $F = \big((1,1)^t\big)$ kann zur Basis $B = \big((1,1)^t, (1,-1)^t\big)$ von $\mathbb{R}^2$ ergänzt werden.

Korollar A4.30 (Existenz von Basen)
Jeder Vektorraum besitzt eine Basis.

Korollar A4.31 (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 Familie B' zu einer Basis $B \cup B'$ von V gemäß dem Basisergänzungssatz A4.28. Dann ist $U' := \mathrm{Lin}\,(B')$ ein Komplement von U, denn

$$U + U' = \mathrm{Lin}\,(U \cup U') \supseteq \mathrm{Lin}\,(B \cup B') = V$$

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 $U \cap U' = \{0\}$. □

Beispiel A4.32 ($\mathbb{R}$ 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.

a. $\mathbb{R}$ *ist überabzählbar* nach Proposition 5.8.

b. *Die abzählbare Vereinigung höchstens abzählbarer Mengen ist wieder höchstens abzählbar.* Seien dazu $M_i = \{a_{i0}, a_{i1}, a_{i2}, \ldots\}$, $i \in \mathbb{N}$, (ohne Einschränkung) abzählbare Mengen, dann schreiben wir sie wie folgt auf:

$$\begin{array}{lccccccc}
M_0 & : & a_{00} & a_{01} \to a_{02} & & a_{03} \to a_{04} & & \ldots \\
 & & \downarrow\nearrow & \swarrow\quad\nearrow & & \swarrow & & \\
M_1 & : & a_{10} & a_{11} & a_{12} & a_{13} & a_{14} & \ldots \\
 & & \swarrow & \nearrow & \swarrow & & & \\
M_2 & : & a_{20} & a_{21} & a_{22} & a_{23} & a_{24} & \ldots \\
 & & \downarrow\nearrow & \swarrow & & & & \\
M_3 & : & a_{30} & a_{31} & a_{32} & a_{33} & a_{34} & \ldots \\
 & & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{array}$$

Abzählen der Elemente wie angedeutet, wobei man Elemente, die mehrfach vorkommen, nur beim ersten Mal berücksichtigt, liefert eine Bijektion von $\mathbb{N} \to \bigcup_{i\in\mathbb{N}} M_i$, mithin ist die Vereinigung abzählbar.

c. Es gilt also $\mathbb{Z} = \mathbb{N} \cup (-\mathbb{N})$ und

$$\mathbb{Q} = \bigcup_{0\neq q\in\mathbb{N}} \tfrac{1}{q}\mathbb{Z} = \bigcup_{0\neq q\in\mathbb{N}} \left\{ \tfrac{p}{q} \mid p \in \mathbb{Z} \right\}$$

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 $N \to \{m\} \times N : n \mapsto (m, n)$ 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

$$V_n := \operatorname{Lin}(x_1, \ldots, x_n).$$

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.

G) Ringe und Moduln

Bemerkung A4.33 (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 A4.1, A4.4, A4.8 und A4.10 können für Moduln über Ringen in der gleichen Weise gegeben werden. Beispiele A4.5 und A4.11 bleiben dann ebenso richtig wie die Aussagen in Lemma A4.6 a. und c. und in den wichtigen Sätzen A4.18, A4.20 und A4.23.

Die Aussagen in Lemma A4.6 und A4.9, Proposition A4.12 und A4.27 sowie in den Korollaren A4.28 und A4.30 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 A4.34
Welche der folgenden Familien sind linear unabhängig / Erzeugendensysteme / Basen von $\mathbb{R}^2$?

a. $\big((1,0)^t,(0,1)^t,(1,1)^t\big)$.

b. $\big((1,1)^t,(2,2)^t\big)$.

c. $\big((1,3)^t\big)$.

d. $\big((1,1)^t,(1,-2)^t\big)$.

e. $\big((1,1)^t,(0,0)^t\big)$.

f. $\big((1,1)^t,(0,0)^t,(1,-2)^t\big)$.

g. $\big((1,2)^t,(2,1)^t\big)$.

Aufgabe A4.35
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 A4.36
Ist $f : V \longrightarrow W$ eine K-lineare Abbildung, F eine Familie von Vektoren in V, so ist

$$f\big(\operatorname{Lin}(F)\big) = \operatorname{Lin}\big(f(x) \mid x \in F\big).$$

Aufgabe A4.37

Bestimme eine Basis und die Dimension des Unterraums

$$U = \mathrm{Lin}(t, t^2 + t^3, t^3 + t, t^2 - t)$$

von $\mathbb{R}[t]$. Ist $6t^3 + 3t + 9$ ein Element von U?

Aufgabe A4.38

Es seien V und W zwei K-Vektorräume, $f : V \longrightarrow W$ 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 A4.39

Es seien $U_1, \ldots, U_k$ Unterräume eines K-Vektorraums V mit Basen $B_1, \ldots, B_k$. Zeige, genau dann ist $V = U_1 \oplus \ldots \oplus U_k$ die direkte Summe der U_i, wenn $B = B_1 \cup \ldots \cup B_k$ eine Basis von V ist.

Aufgabe A4.40

Sei V ein $\mathbb{C}$–Vektorraum, dann ist V offensichtlich auch ein $\mathbb{R}$–Vektorraum, und seien $x_1, \ldots, x_n \in V$. Zeige, daß $(x_1, \ldots, x_n)$ genau dann linear unabhängig über $\mathbb{C}$ ist, wenn $(x_1, ix_1, \ldots, x_n, ix_n)$ linear unabhängig über $\mathbb{R}$ ist.

Aufgabe A4.41

Es sei $(V, +, \cdot)$ ein K-Vektorraum, und $x_1, \ldots, x_n \in V$ seien linear abhängige Vektoren mit der Eigenschaft, daß je $n - 1$ der Vektoren linear unabhängig sind. Zeige:

a. Es gibt $\lambda_1, \ldots, \lambda_n \in K \setminus \{0\}$ mit der Eigenschaft

$$\sum_{i=1}^{n} \lambda_i x_i = 0.$$

b. Gilt für $\mu_1, \ldots, \mu_n \in K$ ebenfalls $\sum_{i=1}^{n} \mu_i x_i = 0$, so gibt es ein $\nu \in K$ mit $\mu_i = \lambda_i \cdot \nu$ für alle $i = 1, \ldots, n$.

Aufgabe A4.42

Es sei $A \in \mathrm{Mat}_n(K)$.

a. Genau dann ist f_A bijektiv, wenn $A \in \mathrm{Gl}_n(K)$.

b. Ist $A \in \mathrm{Gl}_n(K)$, so gilt $(f_A)^{-1} = f_{A^{-1}}$.

Aufgabe A4.43

Welche der folgenden Familien von Vektoren sind linear abhängig, welche sind linear unabhängig, welche sind ein Erzeugendensystem, welche sind eine Basis des jeweils angegebenen K-Vektorraums V.

a. $\big((1,2,1)^t, (1,1,1)^t\big)$ mit $V = \mathbb{R}^3$ und $K = \mathbb{R}$.

b. $\big((1,2,3)^t, (0,2,5)^t, (1,2,4)^t, (1,1,0)^t\big)$ mit $V = \mathbb{R}^3$ und $K = \mathbb{R}$.

c. $\big((1,2,3)^t, (0,2,5)^t, (0,0,0)^t, (1,1,0)^t\big)$ mit $V = \mathbb{R}^3$ und $K = \mathbb{R}$.

d. $\big((1,0,1)^t, (1,0,-1)^t, (1,1,0)^t\big)$ mit $V = K^3$ und $K = \mathbb{R}$ sowie $K = \mathbb{F}_2$.

e. $\big((1,1,1,1,1)^t, (1,0,1,0,1)^t, (0,1,0,1,1)^t, (1,1,1,1,1)^t\big)$ mit $V = \mathbb{Q}^5$ und $K = \mathbb{Q}$.

f. $\big((1,0,1)^t, (1,1,0)^t\big)$ mit $K = \mathbb{R}$ und $V = \mathrm{Ker}(f)$ für

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R} : (x,y,z)^t \mapsto x - y - z.$$

Aufgabe A4.44

Bestimme eine Basis des Vektorraums

$$\mathrm{Lin}\big((1,1,2,0,1)^t, (2,-1,3,1,0)^t, (0,-3,-3,1,-2)^t, (1,2,-1,3,0)^t, (1,-1,-4,4,-2)^t\big) \leq \mathbb{R}^5.$$

Aufgabe A4.45 (Im Lemma von Zorn reichen abzählbare Ketten nicht!)

Finde ein Beispiel für eine teilgeordnete Menge $(M, \leq)$, so daß jede abzählbare Kette

$$K_1 \leq K_2 \leq K_3 \leq \ldots$$

von Elementen $K_i \in M$ eine obere Schranke in M besitzt, daß aber M selbst kein maximales Element hat.

Beispiel von Thomas Seiferling und Mario Hefter: Betrachte die Menge M aller abzählbaren Teilmengen von $\mathbb{R}$ mit der Inklusion als Teilordnung. Für eine abzählbare Kette $K_1 \subseteq K_2 \subseteq \ldots$ ist $\bigcup_{i=1}^{\infty} K_i$ eine abzählbare Teilmenge von $\mathbb{R}$, die somit eine obere Schranke der Kette in M ist. Hätte M aber ein maximales Element X, so wäre dies insbesondere eine abzählbare Teilmenge von $\mathbb{R}$ und somit gäbe es ein Element $x \in \mathbb{R} \setminus X$. Dann wäre aber $X \cup \{x\}$ auch abzählbar und $X \subsetneqq X \cup \{x\}$, im Widerspruch zur Maximalität von X. □

A5 Endlich-dimensionale Vektorräume

Wir betrachten jetzt *endlich erzeugte Vektorräume* V, d. h. Vektorräume, die ein endliches Erzeugendensystem besitzen. Nach Bemerkung A4.15 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) Austauschsatz von Steinitz

Lemma A5.1 (Austauschlemma)
Sei $B = (x_1, \ldots, x_n)$ eine Basis von V, $y = \sum_{i=1}^{n} \lambda_i x_i$ und $\lambda_j \neq 0$ für ein j.
Dann ist $(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n)$ 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

$$V = \operatorname{Lin}(x_1, \ldots, x_n) = \operatorname{Lin}(y, x_1, \ldots, x_n) = \operatorname{Lin}(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n).$$

Bleibt also zu zeigen, daß $(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n)$ linear unabhängig ist. Seien dazu $\mu_i \in K$, $i = 1, \ldots, n$, gegeben mit

$$0 = \mu_j y + \sum_{i \neq j} \mu_i x_i = \sum_{i=1}^{n} \mu_j \lambda_i x_i + \sum_{i \neq j} \mu_i x_i$$

$$= \mu_j \lambda_j x_j + \sum_{i \neq j} (\mu_j \lambda_i + \mu_i) x_i.$$

Dann folgt aus der linearen Unabhängigkeit von $x_1, \ldots, x_n$

$$\mu_j \lambda_j = 0 \quad \text{und} \quad \mu_i = -\mu_j \lambda_i, \text{ für } i \neq j.$$

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 $(x_1, \ldots, x_{j-1}, y, x_{j+1}, \ldots, x_n)$ gezeigt. □

Beispiel A5.2

Ist zum Beispiel $E = (e_1, \ldots, e_n)$ die kanonische Basis des K^n und $x = (\lambda_1, \ldots, \lambda_n)^t \in K^n$ 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 A5.3 (Austauschsatz von Steinitz)
Sei $(x_1, \ldots, x_n)$ eine Basis von V und $(y_1, \ldots, y_r)$ sei linear unabhängig in V. Dann lassen sich die $x_1, \ldots, x_n$ so umnumerieren, daß $(y_1, \ldots, y_r, x_{r+1}, \ldots, x_n)$ eine Basis von V ist. Insbesondere gilt $r \leq n$.

Beweis von Satz A5.3: 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 $(y_1, \ldots, y_{r-1}, x_r, \ldots, x_n)$ eine Basis von V. Dann besitzt y_r eine Darstellung der Form

$$y_r = \lambda_1 y_1 + \ldots + \lambda_{r-1} y_{r-1} + \lambda_r x_r + \ldots + \lambda_n x_n,$$

mit $\lambda_i \in K$. Angenommen, $\lambda_r = \ldots = \lambda_n = 0$, dann wäre $(y_1, \ldots, y_r)$ linear abhängig, im Widerspruch zur Voraussetzung. Also gibt es ein $j \in \{r, \ldots, n\}$ mit $\lambda_j \neq 0$. Durch Umnumerieren können wir annehmen, daß $j = r$ gilt. Dann können wir aber nach dem Austauschlemma A5.1 y_r gegen x_r austauschen, und die Behauptung ist bewiesen. □

Bemerkung A5.4

a. Der Austauschsatz von Steinitz besagt also, daß man - nach eventuellem Umnumerieren - die linear unabhängigen Vektoren $x_1, \ldots, x_r$ durch $y_1, \ldots, y_r$ ersetzen kann.

b. 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 $i_1, \ldots, i_r$ aus. Im j-ten Schritt wissen wir, daß wir eine Darstellung

$$y_j = \sum_{l=1}^{j-1} \lambda_{i_l} y_l + \sum_{l \notin \{i_1, \ldots, i_{j-1}\}} \lambda_l x_l$$

haben mit $\lambda_l \neq 0$ für ein $l \notin \{i_1, \ldots, i_{j-1}\}$, 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 Steinitzschen Austauschsatzes.

Aus dem Satz von Steinitz folgen unmittelbar die folgenden beiden Korollare.

Korollar A5.5 (Basisergänzungssatz)
Ist $(y_1, \ldots, y_r)$ eine linear unabhängige Familie im endlich erzeugten Vektorraum V, so kann diese zu einer Basis B von V ergänzt werden.

Beweis: Nach Voraussetzung besitzt V ein endliches Erzeugendensystem und dieses kann zu einer endlichen Basis $(x_1, \ldots, x_n)$ von V ausgedünnt werden. Dann wenden wir auf $(y_1, \ldots, y_r)$ und $(x_1, \ldots, x_n)$ den Satz von Steinitz an. □

Korollar A5.6 (Existenz eines Komplementes)
Ist V ein endlich erzeugter Vektorraum und U ein Unterraum von V, dann besitzt U ein direktes Komplement.

Beweis: Wir wählen eine Basis B von U und ergänzen sie durch eine Familie B' zu einer Basis von V, dann erzeugt B' ein direktes Komplement von U in V. □

B) Die Dimension eines endlich-erzeugten Vektorraums

Als Folgerung des Steinitzschen Austauschsatzes erhalten wir den folgenden Satz.

Satz A5.7 (Alle Basen sind gleichmächtig.)

a. Ist V endlich erzeugt, so ist jede Basis von V endlich und alle Basen haben gleich viele Elemente.

b. Ist V nicht endlich erzeugt, so hat jede Basis unendlich viele Elemente.

Beweis: a. Nach Voraussetzung besitzt V ein endliches Erzeugendensystem E und nach Bemerkung A4.15 folgt dann auch, daß V eine endliche Basis $B = (x_1, \ldots, x_n)$ 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 $(y_1, \ldots, y_{n+1})$ in B', im Widerspruch zum Austauschsatz von Steinitz, der verlangt $n + 1 \leq n$.

b. Dies ist offensichtlich, da jede Basis V erzeugt.

□

Satz A5.7 rechtfertigt die folgende Definition.

Definition A5.8 (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 A5.7 und Definition A5.8 folgt unmittelbar das folgende Korollar.

Korollar A5.9
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 Korollar A4.28 ist F in einer Basis von V enthalten und E enthält eine Basis von V. Die Ungleichungen folgen dann aus Satz A5.7, 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 A4.12 eine Basis. □

Beispiel A5.10

a. Es gilt:
$$\dim_K(V) = 0 \Leftrightarrow V = \mathrm{Lin}(\emptyset) \Leftrightarrow V = \{0\}.$$

b. $\dim_K\left(K^n\right) = n$, da die kanonische Basis $E = (e_1, \ldots, e_n)$ 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 $B = (t^0, t^1, \ldots, t^n)$ eine Basis ist.

Satz A5.11 (Karten eines Vektorraums)
Es sei $B = (x_1, \ldots, x_n)$ eine Basis von V und $E = (e_1, \ldots, e_n)$ die kanonische Basis des K^n. Dann bestimmt B einen Isomorphismus

$$\phi_B : V \to K^n : x_i \mapsto e_i, \text{ für } i = 1, \ldots, n,$$

durch *lineare Fortsetzung*. Man nennt ϕ_B die *Karte* von V zur Basis B.

Beweis: Nach Satz A4.20 bestimmen die Zuordnungen

$$x_i \mapsto e_i,\ i = 1, \ldots, n, \quad \text{und} \quad e_i \mapsto x_i,\ i = 1, \ldots, n,$$

zwei lineare Abbildungen $\phi_B : V \to K^n$ und $\phi^B : K^n \to V$. Es bleibt zu zeigen, daß

$$\phi_B \circ \phi^B = \mathrm{id}_{K^n} \quad \text{und} \quad \phi^B \circ \phi_B = \mathrm{id}_V\,.$$

Dazu reicht es wegen Satz A4.20, daß die beiden Seiten jeweils auf einer Basis übereinstimmen, und das tun sie offenbar. □

Daraus ergibt sich unmittelbar folgendes Korollar.

Korollar A5.12 (Dimension als einzige Invariante eines Vektorraums)
Für zwei endlich-dimensionale K-Vektorräume V und W sind gleichwertig:

a. $V \cong W$.

b. $\dim_K(V) = \dim_K(W)$.

Insbesondere ist jeder n-dimensionale K-Vektorraum isomorph zum K^n.

Beweis: Es seien $n = \dim_K(V)$ und $m = \dim_K(W)$.

Ist $f : V \to W$ ein Isomorphismus, so überführt er laut Aufgabe A4.38 eine Basis von V, die n Elemente enthält, in eine Basis von W, die m Elemente enthält. Mithin gilt $n = m$.

Ist umgekehrt $n = m$, dann sind nach Satz A5.11 beide Vektorräume isomorph zu K^n und mithin zueinander. □

Beispiel A5.13
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 $\mathrm{Lin}\,(\sin, \cos)$ die Dimension zwei, da $(\sin, \cos)$ eine Basis ist, und deshalb gilt

$$\mathrm{Lin}\,(\sin, \cos) \cong \mathbb{R}^2.$$

C) 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 A5.9 den folgenden Algorithmus.

Algorithmus A5.14 (Test auf lineare Unabhängigkeit)
INPUT: Eine Familie F von 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 Zeilen in eine Matrix A.
2. **Schritt:** Berechne eine Zeilen-Stufen-Form von A.
3. **Schritt:** Falls diese keine Nullzeile enthält, so gib Eins zurück, sonst Null.

Beweis: Da elementare Zeilenoperationen den Zeilenraum nicht ändern, erzeugen die Zeilen der Zeilen-Stufen-Form den Unterraum $\mathrm{Lin}(F)$. Sie sind zudem linear unabhängig (siehe Beweis von Algorithmus A4.16) und somit eine Basis von $\mathrm{Lin}(F)$. Da das Erezgendensystem F genauso viele Vektoren enthält, muss es ebenfalls eine Basis sein. □

D) Dimensionsformel für Summen von Unterräumen

Lemma A5.15
Ist $\dim_K(V) < \infty$ und $U \subseteq V$ ein Unterraum. Dann gelten

$$\dim_K(U) \leq \dim_K(V),$$

und

$$U = V \iff \dim_K(U) = \dim_K(V).$$

Beweis: Ist U ein Unterraum, so kann eine Basis B von U zu einer Basis B' von V ergänzt werden, so daß $\dim_K(U) = |B| \leq |B'| = \dim_K(V)$ gelten muß.

Ist $U = V$, so ist offenbar auch $\dim_K(U) = \dim_K(V)$. Gilt umgekehrt $\dim_K(U) = \dim_K(V)$ und ist $B = (x_1, \ldots, x_n)$ eine Basis von U, so können wir sie nach dem Basisergänzungssatz A4.28 zu einer Basis B' von V ergänzen. Wegen $B \subseteq B'$ und $|B| = n = |B'|$ folgt dann aber notwendigerweise $B = B'$, und somit $U = \mathrm{Lin}\,(B) = \mathrm{Lin}\,(B') = V$. □

Satz A5.16 (Dimensionsformel für Summen von Unterräumen)
Ist $\dim_K(V) < \infty$ und sind U und U' Unterräume von V, dann gilt:

$$\dim_K\left(U + U'\right) = \dim_K(U) + \dim_K\left(U'\right) - \dim_K\left(U \cap U'\right).$$

Beweis: Wir beweisen mehr, nämlich wie wir geeignete Basen von U, U' und $U \cap U'$ wählen können. Sei $B_{U\cap U'} := (x_1, \ldots, x_r)$ eine Basis von $U \cap U'$. Wir ergänzen $B_{U\cap U'}$ zu einer Basis $B_U := (x_1, \ldots, x_r, y_1, \ldots, y_s)$ von U, und zu einer Basis $B_{U'} := (x_1, \ldots, x_r, z_1, \ldots, z_t)$ von U'. Das geht nach dem Basisergänzungssatz A4.28.

Behauptung: $B_{U+U'} := (x_1, \ldots, x_r, y_1, \ldots, y_s, z_1, \ldots, z_t)$ 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 $\lambda_i, \lambda'_i, \mu_j, \mu'_k \in K$, $i = 1, \ldots, r$, $j = 1, \ldots, s$, $k = 1, \ldots, t$. Daraus folgt:

$$x + x' = \sum_{i=1}^{r} \left(\lambda_i + \lambda'_i\right) x_i + \sum_{j=1}^{s} \mu_j y_j + \sum_{k=1}^{t} \mu'_k z_k \in \mathrm{Lin}\left(B_{U+U'}\right).$$

Dann müssen wir noch zeigen, daß $B_{U+U'}$ linear unabhängig ist. Sei dazu

$$\sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j y_j + \sum_{k=1}^{t} \nu_k z_k = 0 \tag{194}$$

eine Linearkombination der Null. Dann ist

$$\underbrace{\sum_{i=1}^{r} \lambda_i x_i + \sum_{j=1}^{s} \mu_j y_j}_{\in U} = \underbrace{\sum_{k=1}^{t} -\nu_k z_k}_{\in U'} \in U \cap U'.$$

Da $B_{U\cap U'}$ eine Basis von $U \cap U'$ 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_U linear unabhängig ist, ergibt ein Koeffizientenvergleich auf beiden Seiten insbesondere $\mu_j = 0$ für alle $j = 1, \ldots, s$. Damit erhalten wir aus (194) dann

$$\sum_{i=1}^{r} \lambda_i x_i + \sum_{k=1}^{t} \nu_k z_k = 0,$$

und da $B_{U'}$ linear unabhängig ist, müssen notwendigerweise alle λ_i und ν_k Null sein. Damit haben wir dann auch gezeigt, daß $B_{U+U'}$ linear unabhängig ist.

Aus der Behauptung folgt,

$$\dim_K\left(U+U'\right) = r+s+t = (r+s)+(r+t)-r = \dim_K(U)+\dim_K\left(U'\right)-\dim_K\left(U\cap U'\right).$$

□

Beispiel A5.17

Für die Unterräume $U = \mathrm{Lin}\left((1,0,0)^t, (1,1,1)^t\right)$ und $U' = \mathrm{Lin}\left((1,1,1)^t, (0,0,1)^t\right)$ von $\mathbb{R}^3$ sieht man leicht, daß

$$U \cap U' = \mathrm{Lin}\left((1,1,1)^t\right)$$

ein Vektorraum von Dimension eins ist, während U und U' jeweils Dimension zwei haben, da die angegebenen Erzeugendensysteme auch linear unabhängig sind. Mithin erhalten wir

$$\dim_{\mathbb{R}}(U + U') = 2 + 2 - 1 = 3 = \dim_{\mathbb{R}}(\mathbb{R}^3),$$

so daß

$$U + U' = \mathbb{R}^3$$

gelten muß. Da zudem

$$U + U' = \mathrm{Lin}\left((1,0,0)^t, (1,1,1)^t, (0,0,1)^t\right)$$

gilt, sehen wir, daß dieses Erzeugendensystem eine Basis des $\mathbb{R}^3$ ist.

E) 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 A5.18 (Summe zweier Unterräume)

INPUT: Erzeugendensysteme F und G von Unterräumen U und U' des K^n.

OUTPUT: Eine Basis von $U + U'$ und $\dim_K(U + U')$.

1. **Schritt:** Bilde aus F und G ein Erzeugendensystem und berechne mittels A4.16 eine Basis von $U + U' = \mathrm{Lin}(F \cup G)$.
2. **Schritt:** Gib diese Basis zurück und die Anzahl der Basisvektoren zurück.

Wir greifen das Beispiel von oben noch mal auf und berechnen eine Basis der Summe mit unserem Algorithmus.

Beispiel A5.19 (Basis von $U + U'$)

Wir betrachten wieder die beiden Unterräume $U = \mathrm{Lin}\left((1,0,0)^t, (1,1,1)^t\right)$ und $U' = \mathrm{Lin}\left((1,1,1)^t, (0,0,1)^t\right)$ von $\mathbb{R}^3$. Schreiben wir die Erzeuger als Zeilen in eine Matrix A und überführen diese in Zeilen-Stufen-Form, so erhalten

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \mapsto \ldots \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} = \mathrm{rZSF}(A).$$

Also gilt $U + U' = \mathbb{R}^3$.

F) Weitere Dimensionsformeln

Korollar A5.20 (Dimensionsformel für Komplemente)
Ist $\dim_K(V) < \infty$, dann sind für Unterräume U und U' von V die folgenden Aussagen äquivalent:

a. $V = U \oplus U'$.

b. $V = U + U'$ und $U \cap U' = \{0\}$.

c. $V = U + U'$ und $\dim_K(V) = \dim_K(U) + \dim_K\left(U'\right)$.

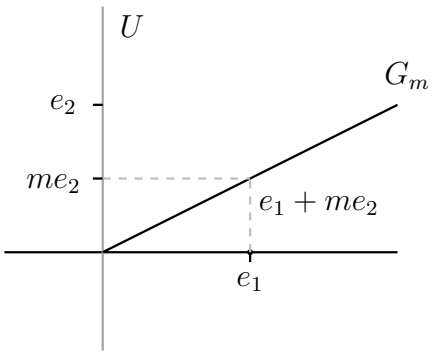
d. $U \cap U' = \{0\}$ und $\dim_K(V) = \dim_K(U) + \dim_K\left(U'\right)$.

Beweis: Dies ist eine direkte Folgerung aus Lemma A3.17, Lemma A5.15 und Satz A5.16. □

Beispiel A5.21
Wir erinnern uns an Beispiel A3.36. Dort haben wir in $V = \mathbb{R}^2$ den Unterraum $U = \mathrm{Lin}\,(e_2)$, die y-Achse, betrachtet und gezeigt, daß jede Ursprungsgerade mit Steigung m

$$G_m := \mathrm{Lin}\,(e_1 + me_2)$$

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 A5.22 (Dimensionsformel für Faktorräume)
Es sei V ein endlich-dimensionaler K-Vektorraum und U ein Unterraum von V. Dann gilt

$$\dim_K(V/U) = \dim_K(V) - \dim_K(U).$$

Beweis: Nach Bemerkung A4.15 besitzt U ein Komplement U', und nach Proposition A3.37 gilt $U' \cong V/U$. Aus Korollar A5.12 und Korollar A5.20 folgt dann

$$\dim_K(V) = \dim_K(U) + \dim_K(U') = \dim_K(U) + \dim_K(V/U).$$

□

Bemerkung A5.23 (Basis von V/U)

Es sei V ein K-Vektorraum und U ein Unterraum von V mit Basis $(x_1, \ldots, x_r)$. Dann sind die folgenden Aussagen für $y_1, \ldots, y_s \in V$ gleichwertig:

a. $(x_1, \ldots, x_r, y_1, \ldots, y_s)$ ist eine Basis von V.

b. $(y_1, \ldots, y_s)$ ist Basis eines Komplementes von U.

c. $\left(\overline{y_1}, \ldots, \overline{y_s}\right)$ ist eine Basis von V/U.

Beweis: Die Äquivalenz von b. und c. folgt aus Proposition A3.37 und Aufgabe A4.38, und die von a. und b. folgt aus dem Beweis von Korollar A5.6 und Korollar A5.20. □

Satz A5.24 (Dimensionsformel für lineare Abbildungen)

Es sei $f : V \longrightarrow W$ eine K-lineare Abbildung und $\dim_K(V) < \infty$. Dann gilt

$$\dim_K(V) = \dim_K\left(\operatorname{Ker}(f)\right) + \dim_K\left(\operatorname{Im}(f)\right).$$

Beweis: Aus dem Homomorphiesatz A3.34 erhalten wir den Isomorphismus

$$V/\operatorname{Ker}(f) \cong \operatorname{Im}(f),$$

so daß die Formel dann aus Korollar A5.22 folgt. □

Beispiel A5.25

Die lineare Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x_1, x_2)^t \mapsto x_1 - x_2$ hat den Kern $\operatorname{Lin}\left((1,1)^t\right)$ von Dimension eins und ist surjektiv. Wir erhalten also die Formel

$$\dim_{\mathbb{R}}(\mathbb{R}^2) = 2 = 1 + 1 = \dim_{\mathbb{R}}(\operatorname{Ker}(f)) + \dim_{\mathbb{R}}\left(\operatorname{Im}(f)\right).$$

G) Bijektivität linearer Abbildungen

Korollar A5.26 (Injektiv = surjektiv = bijektiv)

Es seien V und W zwei endlich-dimensionale K-Vektorräume gleicher Dimension und $f : V \longrightarrow W$ sei K-linear. Dann sind gleichwertig:

a. f ist bijektiv,

b. f ist injektiv,

c. f ist surjektiv.

Beweis: Ist f injektiv, so ist $\mathrm{Ker}(f) = \{0\}$, und wir erhalten aus der Dimensionsformel für lineare Abbildungen A5.24

$$\dim_K(W) = \dim_K(V) = \dim_K\big(\mathrm{Ker}(f)\big) + \dim_K\big(\mathrm{Im}(f)\big) = \dim_K\big(\mathrm{Im}(f)\big).$$

Wegen Lemma A5.15 gilt dann $W = \mathrm{Im}(f)$ und f ist surjektiv.

Ist f surjektiv, so ist $W = \mathrm{Im}(f)$ und wir erhalten aus der Dimensionsformel für lineare Abbildungen A5.24

$$\dim_K\big(\mathrm{Ker}(f)\big) = \dim_K(V) - \dim_K\big(\mathrm{Im}(f)\big) = \dim_K(V) - \dim_K(W) = 0.$$

Dann ist aber $\mathrm{Ker}(f) = \{0\}$ und somit ist f injektiv. □

Korollar A5.27 (Invertierbare Matrizen)
Sind $A, B \in \mathrm{Mat}_n(K)$ mit $AB = \mathbb{1}_n$, so gilt auch $BA = \mathbb{1}_n$ und $A \in \mathrm{Gl}_n(K)$.

Beweis: Aus $AB = \mathbb{1}_n$ folgt

$$f_A \circ f_B = f_{AB} = f_{\mathbb{1}_n} = \mathrm{id}_{K^n},$$

so daß f_B injektiv mit Linksinverser f_A ist. Nach Korollar A5.26 ist f_B dann aber schon bijektiv, und die Linksinverse ist die Inverse von f_B. Damit folgt dann auch

$$f_{\mathbb{1}_n} = \mathrm{id}_{K^n} = f_B \circ f_A = f_{BA},$$

und damit $BA = \mathbb{1}_n$. □

H) Ringe und Moduln

Bemerkung A5.28 (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 A5.27 bleibt wahr, allerdings braucht man einen neuen Beweis.

Aufgaben

Aufgabe A5.29
Bestimme die Dimension der von den Familien in Aufgabe A4.43 aufgespannten Vektorräume.

Aufgabe A5.30

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 A5.31

Finde einen K-Vektorraum V sowie zwei K-lineare Abbildungen $f, g : V \longrightarrow V$, so daß folgendes gilt:

a. f ist injektiv, aber nicht surjektiv.

b. g ist surjektiv, aber nicht injektiv.

Aufgabe A5.32

Es sei V ein K-Vektorraum mit $1 \leq \dim_K(V) = n < \infty$ und $g \in \mathrm{End}_K(V)$. Zeige, es gibt eine Zahl $0 \leq k \leq n$, so daß für alle $i \geq 1$ gilt:

$$\mathrm{Ker}(g^0) \subsetneqq \mathrm{Ker}(g^1) \subsetneqq \ldots \subsetneqq \mathrm{Ker}(g^k) = \mathrm{Ker}(g^{k+i}).$$

Aufgabe A5.33

Es sei $B := \big((3,5,2)^t, (1,1,-1)^t, (2,4,1)^t\big)$.

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 A5.34

Sei V ein K-Vektorraum und $F = (v_1, \ldots, v_5)$ eine linear unabhängige Familie in V. Welchen der Vektoren $v_1, \ldots, v_5$ 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 A5.35

Sei K ein Körper.

a. Begründe, weshalb die Mengen $U := \{(a_1, \ldots, a_n)^t \in K^n \mid a_1 = \ldots = a_n\}$ und $U' := \big\{(a_1, \ldots, a_n)^t \in K^n \;\big|\; a_1 + \ldots + a_n = 0\big\}$ Unterräume des K^n sind.

b. Bestimme $\dim_K(U)$, $\dim_K(U')$, $\dim_K(U \cap U')$ und $\dim_K(U + U')$.

Aufgabe A5.36

Es sei $B := \big((2,0,3)^t, (1,2,1)^t, (4,2,1)^t\big)$.

a. Zeige, B ist eine Basis von $\mathbb{F}_5^3$.

b. Ersetze *mit Hilfe des Austauschsatzes von Steinitz* zwei Vektoren in B durch die Vektoren $(3,3,2)^t$ und $(1,1,3)^t$.

Aufgabe A5.37

Beschreibe die folgenden Unterräume des $\mathbb{R}^3$ geometrisch und bestimme ihre Dimension:

a. $\operatorname{Lin}\big((0,0,0)^t\big)$.

b. $\operatorname{Lin}\big((1,0,0)^t\big)$.

c. $\operatorname{Lin}\big((0,0,0)^t, (-2,0,0)^t\big)$.

a. $\operatorname{Lin}\big((1,1,0)^t, (0,0,1)^t\big)$.

b. $\operatorname{Lin}\big((1,1,0)^t, (0,0,1)^t, (1,1,1)^t\big)$.

c. $\operatorname{Lin}\big((1,0,0)^t, (0,0,1)^t, (0,1,0)^t\big)$.

A6 Lineare Abbildungen und Matrizen

In diesem Abschnitt sind V und W zwei endlich-dimensionale K-Vektorräume mit Basen $B = (b_1, \ldots, b_n)$ und $D = (d_1, \ldots, d_m)$, sofern nichts anderes gesagt wird.
Ferner bezeichnen wir mit $E = (e_1, \ldots, e_n)$ die kanonische Basis von K^n und mit $F = (f_1, \ldots, f_m)$ die kanonische Basis von K^m.

A) Matrixdarstellung linearer Abbildungen

Definition A6.1 (Matrixdarstellung einer linearen Abbildung)
Gegeben seien eine Basis $B = (b_1, \ldots, b_n)$ von V und eine Basis $D = (d_1, \ldots, d_m)$ von W sowie eine lineare Abbildung $f : V \longrightarrow W$.

a. Ist $x \in V$, so läßt sich x nach Proposition A4.18 auf eindeutige Weise darstellen als Linearkombination der Basis B

$$x = \lambda_1 \cdot b_1 + \ldots \lambda_n \cdot b_n.$$

Wir nennen den Vektor $M_B(x) = (\lambda_1, \ldots, \lambda_n)^t$ den *Koordinatenvektor* oder die *Koordinaten* von x bezüglich B.

b. Für jeden Basisvektor b_j in B läßt sich das Bild $f(b_j) \in W$ unter f nach Proposition A4.18 auf eindeutige Weise als Linearkombination der Basis D darstellen

$$f(b_j) = a_{1j} \cdot d_1 + \ldots + a_{mj} \cdot d_m.$$

Schreiben wir die Koeffizienten als Spalten in eine Matrix, so erhalten wir die Matrix

$$M_D^B(f) := \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \in \mathrm{Mat}(m \times n, K),$$

die sogenannte *Matrixdarstellung* von f bezüglich der Basen B und D.

Beispiel A6.2

a. Der Koordinatenvektor eines Vektors $x = (x_1, \ldots, x_n)^t \in K^n$ bezüglich der kanonischen Basis $E = (e_1, \ldots, e_n)$ ist der Vektor $M_E(x) = x$ selbst.

b. Um den Koordinatenvektor von $x = (4, 0)^t \in \mathbb{R}^2$ bezüglich der Basis $B = \big((1, 1)^t, (1, -1)^t\big)$ zu bestimmen, muß man ihn als Linearkombination bezüglich

der Basis darstellen. Man sieht leicht, daß

$$x = 2 \cdot (1,1)^t + 2 \cdot (1,-1)^t$$

gilt. Damit erhalten wir

$$M_B(x) = (2,2)^t.$$

c. Sei $V = \mathbb{R}^2$ mit Basis $B = (b_1, b_2) = \big((1,2)^t, (1,1)^t\big)$ und $W = \mathbb{R}^3$ mit Basis $D = (d_1, d_2, d_3) = \big((1,1,0)^t, (0,1,1)^t, (0,0,1)^t\big)$, und sei $f : V \to W$ 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:

$$M_D^B(f) = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix}.$$

d. Ist $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$ eine Matrix und $f_A : K^n \longrightarrow K^m$ die zugehörige Abbildung, dann ist

$$f_A(e_j) = A \cdot e_j = j\text{-te Spalte von } A = a_{1j} \cdot f_1 + \ldots + a_{mj} \cdot f_m.$$

Die Matrixdarstellung f_A bezüglich der kanonischen Basen ist also

$$M_F^E(f_A) = A.$$

e. Ist $V = \mathrm{Lin}(\cos, \sin)$ der Unterraum des $\mathbb{R}^{\mathbb{R}}$ aus Beispiel A5.13 mit Basis $B = (\cos, \sin)$ und ist

$$D : V \longrightarrow V : f \mapsto f'$$

der Ableitungsoperator, dann gilt

$$D(\cos) = \cos' = -\sin = 0 \cdot \cos + (-1) \cdot \sin$$

und

$$D(\sin) = \sin' = \cos = 1 \cdot \cos + 0 \cdot \sin,$$

so daß wir die Matrixdarstellung

$$M_B^B(D) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

erhalten.

Bemerkung A6.3

a. Mit der Notation aus Satz A5.11 gilt

$$M_B(x) = \phi_B(x),$$

d. h. der Koordinatenvektor von x unter B ist das Bild unter der Karte ϕ_B. Die Zuordnung

$$\phi_B : V \longrightarrow K^n : x \mapsto M_B(x)$$

ist ein Isomorphismus von Vektorräumen, der die Vektoren in einem festgelegten Koordinatensystem darstellt.

b. Nach Definition gilt zudem:

> Die j-te Spalte von $M_D^B(f)$ ist der Koordinatenvektor von $f(b_j)$ bez. D.

Proposition A6.4 (Rechnen in Koordinaten)
Ist $f : V \longrightarrow W$ eine lineare Abbildung und $x \in V$, so gilt

$$M_D\big(f(x)\big) = M_D^B(f) \circ M_B(x),$$

d. h. der Koordinatenvektor $M_D\big(f(x)\big)$ 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.

Beweis: Für einen Vektor $x = \sum_{j=1}^n \lambda_j b_j$ und eine lineare Abbildung mit Matrixdarstellung $M_D^B(f) = (a_{ij})$ gilt

$$f(x) = \sum_{j=1}^{n} \lambda_j \cdot f(b_j) = \sum_{j=1}^{n} \lambda_j \cdot \sum_{i=1}^{m} a_{ij} \cdot d_i = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij} \cdot \lambda_j \right) \cdot d_i.$$

Daraus folgt dann

$$M_D\big(f(x)\big) = (a_{ij}) \circ (\lambda_1, \ldots, \lambda_n)^t = M_D^B(f) \circ M_B(x).$$

□

Beispiel A6.5
Betrachten wir die lineare Abbildung $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^3$ aus Beispiel A6.2 und den Vektor $x = (0, 1)^t$, so gilt

$$x = 1 \cdot (1, 2)^t - 1 \cdot (1, 1)^t = 1 \cdot b_1 - 1 \cdot b_2$$

und damit

$$M_B(x) = (1, -1)^t.$$

Daraus leiten wir

$$M_D(f(x)) = M_D^B(f) \cdot M_B(x) = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \circ \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 2 \end{pmatrix}$$

ab, woraus

$$f(x) = 0 \cdot d_1 - 1 \cdot d_2 + 2 \cdot d_3 = 0 \cdot (1,1,0)^t - 1 \cdot (0,1,1)^t + 2 \cdot (0,0,1)^t = (0,-1,1)^t$$

folgt. Wir können also die Bilder beliebiger Vektoren ausrechnen, obwohl wir die Abbildungsvorschrift nicht als geschlossene Formel in den Standardkoordinaten kennen. Das ist allerdings mühsam, und wir werden weiter unten sehen, wie man diese Abbildungsvorschrift aus der Matrixdarstellung gewinnen kann.

Satz A6.6 (Matrixdarstellung linearer Abbildungen)
Die Abbildung

$$M_D^B : \mathrm{Hom}_K(V,W) \longrightarrow \mathrm{Mat}(m \times n, K) : f \mapsto M_D^B(f)$$

ist ein Isomorphismus von K-Vektorräumen.

Beweis: Wir zeigen zunächst, daß die Abbildung linear ist. Sind $f, g \in \mathrm{Hom}_K(V,W)$ mit $M_D^B(f) = (a_{ij})$ und $M_D^B(g) = (b_{ij})$ und sind $\lambda, \mu \in K$, so gilt

$$(\lambda f + \mu g)(b_j) = \lambda \cdot f(b_j) + \mu \cdot g(b_j) = \lambda \cdot \sum_{i=1}^{m} a_{ij} \cdot d_i + \mu \cdot \sum_{i=1}^{m} b_{ij} \cdot d_i = \sum_{i=1}^{m} (\lambda \cdot a_{ij} + \mu \cdot b_{ij}) \cdot d_i,$$

woraus wir die Matrixdarstellung

$$M_D^B(\lambda f + \mu g) = (\lambda \cdot a_{ij} + \mu \cdot b_{ij}) = \lambda \cdot (a_{ij}) + \mu \cdot (b_{ij}) = \lambda \cdot M_D^B(f) + \mu \cdot M_D^B(g)$$

erhalten. Die Abbildung M_D^B ist also K-linear.

Es bleibt, zu zeigen, daß M_D^B bijektiv ist. Sei dazu $A = (a_{ij}) \in \mathrm{Mat}(m \times n, K)$ eine beliebige $m \times n$-Matrix und setzen wir

$$y_j = \sum_{i=1}^{m} a_{ij} \cdot d_i \in W,$$

so gibt es nach dem Existenz- und Eindeutigkeitssatz für lineare Abbildungen A4.20 genau eine lineare Abbildung $f \in \mathrm{Hom}_K(V,W)$ mit

$$f(b_j) = y_j = \sum_{i=1}^{m} a_{ij} \cdot d_i$$

für $j = 1, \ldots, n$, d. h. es gibt genau eine lineare Abbildung f mit

$$M_D^B(f) = (a_{ij}).$$

Die Abbildung M_D^B ist also bijektiv. □

Bemerkung A6.7 (Matrixdarstellung bezüglich der kanonischen Basen)
Für $f \in \mathrm{Hom}_K\left(K^n, K^m\right)$ definieren wir eine Matrix $A_f \in \mathrm{Mat}(m \times n, K)$, deren j-te Spalte das Bild $f(e_j)$ des j-ten Einheitsvektors ist. Dann ist

$$M_F^E : \mathrm{Hom}_K\left(K^n, K^m\right) \longrightarrow \mathrm{Mat}(m \times n, K) : f \mapsto A_f$$

die Umkehrabbildung von

$$\mathrm{Mat}(m \times n, K) \longrightarrow \mathrm{Hom}_K\left(K^n, K^m\right) : A \mapsto f_A.$$

Diesen Spezialfall von Satz A6.6 hatten wir schon in Korollar A4.23 bewiesen, wobei die Linearität der Abbildung dabei unmittelbar aus Lemma A2.8 folgt.

Lemma A6.8 (Verträglichkeit von Matrixdarstellung und Komposition)
Sind $f \in \mathrm{Hom}_K(U, V)$ und $g \in \mathrm{Hom}_K(V, W)$ und sind B, C bzw. D Basen von U, V bzw. W, dann gilt

$$M_D^B(g \circ f) = M_D^C(g) \circ M_C^B(f).$$

Beweis: Ist $B = (b_1, \ldots, b_n)$, so gilt für den j-ten Einheitsvektor $e_j \in K^n$

$$M_B(b_j) = e_j$$

und mit Proposition A6.4 folgt dann

$$\begin{aligned} M_D^B(g \circ f) \circ e_j =& M_D^B(g \circ f) \circ M_B(b_j) = M_D\big((g \circ f)(b_j)\big) \\ =& M_D\big(g(f(b_j))\big) = M_D^C(g) \circ M_C\big(f(b_j)\big) \\ =& M_D^C(g) \circ \big(M_C^B(f) \circ M_B(b_j)\big) = \big(M_D^C(g) \circ M_C^B(f)\big) \circ e_j. \end{aligned}$$

Da die Multiplikation einer Matrix mit e_j die j-te Spalte dieser Matrix liefert, stimmen in $M_D^B(g \circ f)$ und in $M_D^C(g) \circ M_C^B(f)$ also die j-te Spalte überein und das für alle $j = 1, \ldots, n$. Die Matrizen sind also identisch. □

Bemerkung A6.9 (K-Algebren)
Ein K-Vektorraum $(B, +, \cdot)$, auf dem zusätzlich eine Multiplikation

$$\circ : B \times B \to B : (x, y) \mapsto x \circ y$$

definiert ist, so daß $(B, +, \circ)$ ein (nicht unbedingt kommutativer) Ring mit Eins 1_B ist, heißt eine *K-Algebra*, falls die Skalarmultiplikation mit der Ringmultiplikation verträglich ist, d. h. für $\lambda \in K$ und $x, y \in B$ gelten:

$$\lambda \cdot (x \circ y) = (\lambda \cdot x) \circ y = x \circ (\lambda \cdot y).$$

Ein *K-Algebrenisomorphismus* ist eine bijektive Abbildung $\varphi : A \to B$ zwischen zwei K-Algebren A und B, die mit allen drei Operationen verträglich ist und die 1_A auf 1_B abbildet.

$\left(\mathrm{End}_K(V), +, \cdot, \circ\right)$ und $\left(\mathrm{Mat}(n, K), +, \cdot, \circ\right)$ sind Beispiele für K-Algebren, die für unsere Vorlesung von Bedeutung sind, und das folgende Korollar besagt, daß diese isomorph zueinander sind.

Korollar A6.10 (M_B^B ist ein K-Algebrenisomorphismus)
Für zwei Endomorphismen $f, g \in \mathrm{End}_K(V)$ gilt

$$M_B^B(f \circ g) = M_B^B(f) \circ M_B^B(g).$$

Insbesondere, $M_B^B : \mathrm{End}_K(V) \longrightarrow \mathrm{Mat}_n(K)$ ist ein K-Algebrenisomorphismus.

Beweis: Die Aussage folgt unmittelbar aus Lemma A6.8. □

Bemerkung A6.11
In Abschnitt A7 werden wir lernen, wie man eine Matrixdarstellung oder einen Basiswechsel ausrechnen kann.

B) Basiswechsel

Unser Ziel ist es nun, in obigem Beispiel A6.5 aus $M_D^B(f)$ die Matrix $A_f = M_F^E(f)$ zu bestimmen. Dazu führen wir folgende allgemeine Begriffsbildung ein.

Definition A6.12 (Basiswechsel)
Sind $B = (b_1, \ldots, b_n)$ und $B' = (b'_1, \ldots, b'_n)$ zwei Basen von V, so besitzt jedes b_j eine eindeutige Darstellung als Linearkombination der Basis B'

$$b_j = a_{1j} \cdot b'_1 + \ldots + a_{nj} \cdot b'_n.$$

Schreiben wir die Koeffizienten als Spalten in eine Matrix, so erhalten wir die Matrix

$$T_{B'}^B := \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix} \in \mathrm{Mat}_n(K),$$

den *Basiswechsel* oder die *Koordinatentransformationsmatrix* bezüglich (B, B').

Es gilt also:

> Die j-te Spalte von $T_{B'}^B$ ist der Koordinatenvektor $M_{B'}(b_j)$ von b_j bez. B'.

Bemerkung A6.13 (Basiswechsel als Matrixdarstellung)
Offensichtlich ist der Basiswechsel ein Spezialfall einer Matrixdarstellung

$$T_{B'}^B = M_{B'}^B(\mathrm{id}_V),$$

nämlich die Matrixdarstellung der Identität bezüglich der Basen B und B'.

Lemma A6.14 (Basiswechselmatrizen sind invertierbar)
Sind B und B' zwei Basen von V, so ist $T_{B'}^{B}$ invertierbar mit

$$\left(T_{B'}^{B}\right)^{-1} = T_{B}^{B'}.$$

Beweis: Mit Lemma A6.8

$$T_{B}^{B'} \circ T_{B'}^{B} = M_{B}^{B'}(\mathrm{id}_V) \circ M_{B'}^{B}(\mathrm{id}_V) = M_{B}^{B}(\mathrm{id}_V \circ \mathrm{id}_V) = M_{B}^{B}(\mathrm{id}_V) = \mathbb{1}_n.$$

□

Satz A6.15 (Basiswechsel bei Matrixdarstellungen)
Seien B und B' Basen von V, D und D' Basen von W und $f \in \mathrm{Hom}_K(V, W)$. Dann gilt:

$$M_{D'}^{B'}(f) = T_{D'}^{D} \circ M_{D}^{B}(f) \circ T_{B}^{B'}.$$

Beweis: Die Aussage folgt unmittelbar aus Lemma A6.8

$$\begin{aligned} M_{D'}^{B'}(f) =& M_{D'}^{B'}(\mathrm{id}_W \circ f \circ \mathrm{id}_V) \\ =& M_{D'}^{D}(\mathrm{id}_W) \circ M_{D}^{B}(f) \circ M_{B}^{B'}(\mathrm{id}_V) \\ =& T_{D'}^{D} \circ M_{D}^{B}(f) \circ T_{B}^{B'}. \end{aligned}$$

□

Korollar A6.16 (Basiswechsel bei Endomorphismen)
Sind B und B' Basen von V, ist $T = T_{B}^{B'}$ und ist $f \in \mathrm{End}_K(V)$, so gilt

$$M_{B'}^{B'}(f) = T^{-1} \circ M_{B}^{B}(f) \circ T.$$

Beweis: Dies folgt aus Satz A6.15, weil nach Lemma A6.14 $\left(T_{B}^{B'}\right)^{-1} = T_{B'}^{B}$. □

Beispiel A6.17
Wir wollen nun für die Abbildung in Beispiel A6.5 die Matrixdarstellung $M_F^E(f)$ bezüglich der kanonischen Basen berechnen. Nach Satz A6.15 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 A2.25),

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix},$$

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

$$T_B^E = \left(T_E^B\right)^{-1} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix}.$$

Also gilt:

$$M_F^E(f) = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \circ \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix} \circ \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 1 & -1 \\ 0 & 1 \end{pmatrix}.$$

C) f-invariante Unterräume

Definition A6.18 (f-invarianter Unterraum)
Ist $f : V \longrightarrow V$ eine K-lineare Abbildung und $U \leq V$ ein Unterraum mit $f(U) \subseteq U$, so nennen wir U einen *f-invarianten Unterraum*.

Bemerkung A6.19 (f-invariante Unterräume)
Aus Aufgabe A3.48 wissen wir, daß jeder f-invariante Unterraum U zwei K-lineare Abbildungen

$$f_U : U \longrightarrow U : x \mapsto f(x)$$

und

$$f_{V/U} : V/U \longrightarrow V/U : \overline{x} \mapsto \overline{f(x)}$$

induziert. Mit Hilfe dieser Abbildungen erhalten wir eine vorteilhafte Blockgestalt bei der Matrixdarstellung.

Proposition A6.20 (Matrixdarstellung in Blockform)
Es sei $f : V \longrightarrow V$ eine K-lineare Abbildung und $U \leq V$ ein f-invarianter Unterraum. Ferner sei $B' = (x_1, \ldots, x_k)$ eine Basis von U und $B = (x_1, \ldots, x_n)$ eine Ergänzung von B' zu einer Basis von V.
Dann ist $B'' = \left(\overline{x_{k+1}}, \ldots, \overline{x_n}\right)$ 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 $0 \in \mathrm{Mat}((n-k) \times k, K)$ die Nullmatrix ist und $* \in \mathrm{Mat}(k \times (n-k), K)$ eine geeignete Matrix ist.

Beweis: Daß B'' eine Basis von V/U ist, wissen wir bereits aus Bemerkung A5.23. Sei nun $M_B^B(f) = (a_{ij}) \in \mathrm{Mat}_n(K)$, $M_{B'}^{B'}(f_U) = (b_{ij}) \in \mathrm{Mat}_k(K)$ und $M_{B''}^{B''}(f_{V/U}) = (c_{ij}) \in \mathrm{Mat}_{n-k}(K)$.

Für $j = 1, \ldots, k$ 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

$$(a_{ij})_{i,j=1,\ldots,k} = (b_{ij})_{i,j=1,\ldots,k} = M_{B'}^{B'}(f_U)$$

und

$$(a_{ij})_{i=k+1,\ldots,n,\ j=1,\ldots,k} = 0 \in \mathrm{Mat}((n-k) \times k, K).$$

Für $j = k+1, \ldots, n$ erhalten wir analog

$$\sum_{i=k+1}^{n} c_{ij}\overline{x_i} = f_{V/U}\left(\overline{x_j}\right) = \overline{f(x_j)} = \overline{\sum_{i=1}^{n} a_{ij} \cdot x_i} = \sum_{i=1}^{n} a_{ij} \cdot \overline{x_i} = \sum_{i=k+1}^{n} a_{ij} \cdot \overline{x_i},$$

da $\overline{x_i} = \overline{0}$ für $i = 1, \ldots, k$. Aus der Eindeutigkeit der Darstellung als Linearkombination von Basisvektoren erhalten wir also wieder

$$(a_{ij})_{i,j=k+1,\ldots,n} = (c_{ij})_{i,j=k+1,\ldots,n} = M_{B''}^{B''}(f_{V/U}).$$

Insgesamt haben wir damit die Behauptung

$$M_B^B(f) = (a_{ij})_{i,j=1,\ldots,n} = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right)$$

gezeigt. □

Proposition A6.21 (Matrixdarstellung in Blockdiagonalgestalt)
Sei $f : V \longrightarrow V$ eine K-lineare Abbildung und $V = U_1 \oplus \ldots \oplus U_k$ die direkte Summe nicht-trivialer f-invarianter Unterräume $U_1, \ldots, U_k$ mit Basen $B_1, \ldots, B_k$. Dann ist $B = B_1 \cup \ldots \cup B_k$ eine Basis von V und es gilt

$$M_B^B(f) = \left(\begin{array}{c|ccc} M_{B_1}^{B_1}(f_{U_1}) & 0 & \cdots & 0 \\ \hline 0 & \boxed{M_{B_2}^{B_2}(f_{U_2})} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{M_{B_k}^{B_k}(f_{U_k})} \end{array}\right).$$

Beweis: Daß B eine Basis von V ist, wissen wir aus Aufgabe A4.39.

Es sei $B = (x_1, \ldots, x_n)$ und $M_B^B(f) = (a_{ij}) \in \mathrm{Mat}_n(K)$. Halten wir ein $1 \leq l \leq k$ fest, so ist $B_l = (x_r, \ldots, x_s)$ für geeignete $1 \leq r < s \leq n$ und $M_{B_l}^{B_l}(f_{U_l}) = (b_{ij})_{i,j=r,\ldots,s} \in \mathrm{Mat}_{s-r+1}(K)$. Wie im Beweis von Proposition A6.20 sehen wir für $j = r, \ldots, s$ dann

$$\sum_{i=1}^{r-1} 0 \cdot x_i + \sum_{i=r}^{s} b_{ij} \cdot x_i + \sum_{i=s+1}^{n} 0 \cdot x_i = \sum_{i=1}^{k} b_{ij} \cdot x_i = f_{U_l}(x_j) = f(x_j) = \sum_{i=1}^{n} a_{ij} \cdot x_i,$$

woraus wieder

$$(a_{ij})_{i,j=r,\ldots,s} = (b_{ij})_{i,j=r,\ldots,s} = M_{B_l}^{B_l}(f_{U_l})$$

sowie $a_{ij} = 0$ für alle $i < r$ und $i > s$ folgt. Damit ist die Behauptung gezeigt. □

D) Äquivalenz von Matrizen und der Rang von Matrizen

Die Koordinatentransformationen in Vektorräumen mit Basen führen auf folgende Äquivalenzbegriffe für Matrizen.

Definition A6.22 (Äquivalenz von Matrizen)
Eine Matrix $A' \in \mathrm{Mat}(m \times n, K)$ heißt *äquivalent* zu $A \in \mathrm{Mat}(m \times n, K)$, falls es invertierbare Matrizen $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$ gibt mit

$$A' = S \circ A \circ T.$$

Bemerkung A6.23 (Äquivalenz von Matrizen als Äquivalenzrelation)
Die Äquivalenz von Matrizen eine Äquivalenzrelation auf $\mathrm{Mat}(m \times n, K)$ ist.

Denn für $A, B, C \in \mathrm{Mat}(m \times n, K)$ gelten:

- $A = \mathbb{1}_m \circ A \circ \mathbb{1}_n$ und mithin ist A äquivalent zu A;

- wenn B äquivalent zu A ist, gibt es $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$ mit $B = S \circ A \circ T$ und mithin $S^{-1} \circ B \circ T^{-1} = A$, so daß auch A äquivalent zu B;
- wenn B äquivalent zu A und C äquivelent zu B ist, so gibt es Matrizen $S, U \in \mathrm{Gl}_m(K)$ und $T, V \in \mathrm{Gl}_n(K)$ mit $B = S \circ A \circ T$ und $C = U \circ B \circ V$ und mithin gilt auch $C = (U \circ S) \circ A \circ (T \circ V)$ mit $U \circ S \in \mathrm{Gl}_m(K)$ und $T \circ V \in \mathrm{Gl}_n(K)$, so daß C auch äquivalent zu A ist.

Definition A6.24 (Rang)

a. Ist $f \in \mathrm{Hom}_K(V, W)$, so definieren wir den *Rang* von f als

$$\mathrm{rang}(f) := \dim_K \big(\mathrm{Im}(f)\big).$$

b. Ferner definieren wir für eine Matrix $A \in \mathrm{Mat}(m \times n, K)$ den *Rang* von A durch:

$$\mathrm{rang}(A) := \mathrm{rang}(f_A).$$

Bemerkung A6.25 (Rangabschätzung)

a. Für $f \in \mathrm{Hom}_K(V, W)$ gilt wegen der Dimensionsformel für lineare Abbildungen

$$\mathrm{rang}(f) = \dim_K(V) - \dim_K \big(\mathrm{Ker}(f)\big) \leq \dim_K(V)$$

und da $\mathrm{Im}(f)$ ein Unterraum von W ist, gilt auch $\mathrm{rang}(f) \leq \dim_K(W)$.

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 $A \in \mathrm{Mat}(m \times n, K)$

$$\mathrm{rang}(A) \leq \min\{m, n\}.$$

Beispiel A6.26

Da der Rang einer Matrix die Anzahl linear unabhängiger Spalten ist, gelten

$$\mathrm{rang}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 0, \quad \mathrm{rang}\begin{pmatrix} 1 & 0 & 2 \\ 1 & 0 & 2 \end{pmatrix} = 1, \quad \mathrm{rang}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 2 \end{pmatrix} = 2.$$

Satz A6.27 (Invertierbare Matrizen haben vollen Rang.)

Eine Matrix $A \in \mathrm{Mat}_n(K)$ ist genau dann invertierbar, wenn $\mathrm{rang}(A) = n$.
Damit sind die Spalten einer invertierbaren $n \times n$-Matrix eine Basis des K^n.

Beweis: Nach Aufgabe A4.42 ist A genau dann invertierbar, wenn $f_A : K^n \longrightarrow K^n$ bijektiv ist. Wegen Korollar A5.26 ist dies genau dann der Fall, wenn f_A surjektiv ist, d.h. wenn $\mathrm{Im}(f_A) = K^n$. Wegen $\mathrm{Im}(f_A) \subseteq K^n$ und Lemma A5.15 ist dies wiederum

genau dann der Fall, wenn

$$n = \dim_K(\mathrm{Im}(f_A)) = \mathrm{rang}(A).$$

Also ist A genau dann invertierbar, wenn $\mathrm{rang}(A) = n$ gilt. In diesem Fall sind die Spalten von A linear unabhängig in K^n und bilden mithin eine Basis des K^n. □

Bemerkung A6.28

Sind die Matrizen $A, A' \in \mathrm{Mat}(m \times n, K)$ 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 $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$ 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 A6.27. 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

$$M_D^B(f_A) = T_D^F \circ M_F^E(f_A) \circ T_E^B = S \circ A \circ T = A'.$$

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 A6.29 (Rang einer Matrixdarstellung)

Ist $f \in \mathrm{Hom}_K(V, W)$, so gilt

$$\mathrm{rang}(f) = \mathrm{rang}\left(M_D^B(f)\right).$$

Insbesondere, Multiplikation einer Matrix A mit einer invertierbaren Matrix ändert den Rang von A nicht und äquivalente Matrizen haben denselben Rang.

Beweis: Wir betrachten die Karten $\phi_B : V \longrightarrow K^n$ und $\phi_D : W \longrightarrow K^m$. Da ϕ_D ein Isomorphismus ist, erhält ϕ_D die Dimension eines Vektorraumes und es folgt

$$\begin{aligned}
\mathrm{rang}\left(M_D^B(f)\right) &= \mathrm{rang}\left(f_{M_D^B(f)}\right) = \mathrm{rang}\left(\phi_D \circ f \circ \phi_B^{-1}\right) \\
&= \dim_K\left(\phi_D\left(f\left(\phi_B^{-1}(K^n)\right)\right)\right) = \dim_K\left(\phi_D\left(f(V)\right)\right) \\
&= \dim_K\left(f(V)\right) = \mathrm{rang}(f)
\end{aligned}$$

Sind A und A' äquivalent, so sind sie nach Bemerkung A6.28 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. □

Beispiel A6.30

Die Abbildung f in Beispiel A6.2 hat den Rang zwei, wie man an ihrer Matrixdarstellung sieht:

$$M_D^B(f) = \begin{pmatrix} 3 & 3 \\ -4 & -3 \\ 6 & 4 \end{pmatrix}.$$

Denn die beiden Spalten sind keine Vielfachen voneinander und sind deshalb linear unabhängig, so daß der Rang als Maximalanzahl linear unabhängiger Spalten der Matrix zwei sein muß.

E) Algorithmus zur Berechnung des Rangs einer Matrix

Lemma A6.31

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 A6.29 ändert dies den Rang der Matrix nicht. Die Aussage folgt also mit Proposition A2.19. □

Algorithmus A6.32 (zur Bestimmung des Rangs)

INPUT: $A \in \mathrm{Mat}(m \times n, K)$.

OUTPUT: $\mathrm{rang}(A)$

1. **Schritt:** Überführe A in ZSF.
2. **Schritt:** Zähle die Anzahl r der Nicht-Nullzeilen in der ZSF.
3. **Schritt:** Gib r zurück.

Beispiel A6.33

In Beispiel A2.22 haben wir eine ZSF berechnet:

$$A := \begin{pmatrix} 0 & 0 & -1 & 2 & 3 \\ -1 & 1 & -3 & 0 & 2 \\ 1 & -1 & 1 & 4 & 3 \end{pmatrix} \longmapsto \quad \ldots \quad \longmapsto \begin{pmatrix} 1 & -1 & 0 & 6 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Die Matrix A hat also Rang 3.

F) Algorithmus zur Berechnung der Inversen einer Matrix

Satz A6.34 (Kriterium für die Invertierbarkeit einer Matrix)
Es sei $A \in \mathrm{Mat}(n, K)$. Dann sind gleichwertig:

a. A ist invertierbar.

b. $\mathrm{rZSF}(A) = \mathbb{1}_n$.

c. Es gibt Elementarmatrizen $T_1, \ldots, T_k \in \mathrm{Mat}(n, K)$ mit:
$$T_k \circ \ldots \circ T_1 \circ A = \mathbb{1}_n.$$

d. Es gibt Elementarmatrizen $T_1', \ldots, T_k' \in \mathrm{Mat}(n, K)$ mit:
$$A = T_1' \circ \ldots \circ T_k'.$$

Insbesondere wird die Gruppe $\mathrm{Gl}_n(K)$ also von den Elementarmatrizen erzeugt.

Beweis: Nach Korollar A6.27 gilt, daß A genau dann invertierbar ist, wenn $\mathrm{rang}(A) = n$. Also folgt die Äquivalenz von a.-d. aus Satz A2.20 unter Berücksichtigung von Proposition A2.19. □

Aus Satz A6.34 leitet sich folgendes Verfahren zur Bestimmung der Inversen einer invertierbaren Matrix ab. Hierzu beachte man, daß für Elementarmatrizen $T_1, \ldots, T_k$, für die gilt, daß $T_k \circ \ldots \circ T_1 \circ A = \mathbb{1}_n$, auch gilt, daß

$$T_k \circ \ldots \circ T_1 \circ (A, \mathbb{1}_n) = (\mathbb{1}_n, T_k \circ \ldots \circ T_1) = \left(\mathbb{1}_n, A^{-1}\right).$$

Algorithmus A6.35 (zur Bestimmung der Inversen)
INPUT: $A \in \mathrm{Mat}(n, K)$.
OUTPUT: Inverse von A, falls sie existiert, eine Fehlermeldung sonst.

1. **Schritt:** Erweitere die Matrix A um $\mathbb{1}_n$ zu $C = (A, \mathbb{1}_n) \in \mathrm{Mat}(n \times 2n, K)$.
2. **Schritt:** Überführe C in reduzierte ZSF $C' = \left(A', B\right)$.
3. **Schritt:** Falls rang $\left(A'\right) = n$, dann gib B zurück, sonst eine Fehlermeldung.

Beispiel A6.36
Wir betrachten die 3×3-Matrix

$$\begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \in \mathrm{Mat}(3 \times 3, K)$$

und versuchen die Inverse mittels des Algorithmus A6.35 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

$$A^{-1} = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \\ -1 & 1 & 1 \end{pmatrix}.$$

G) Normalform einer Matrix bezüglich Äquivalenz

Satz A6.37 (Normalform einer Matrixdarstellung bezüglich Äquivalenz) Es sei $f \in \mathrm{Hom}_K(V, W)$ mit $\mathrm{rang}(f) = r$. 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.[4] Wir bezeichnen die rechte Seite der obigen Gleichung auch als die *Normalform* von f bezüglich Äquivalenz.

[4] Man bezeichnet die vier Matrizen $\mathbb{1}_r \in \mathrm{Mat}_r(K)$, $0 \in \mathrm{Mat}\,(r \times (n-r), K)$, $0 \in \mathrm{Mat}\,((m-r) \times r, K)$ und $0 \in \mathrm{Mat}((m-r) \times (n-r), K)$ auch als *Blöcke* von $M_D^B(f)$ und die Matrix $M_D^B(f)$ als eine *Blockmatrix*.

Beweis: Wähle vermöge Lemma A5.6 ein Komplement U von $\mathrm{Ker}(f)$. Nach Satz A3.34 und Lemma A3.37 ist die folgende Abbildung ein Isomorphismus

$$f_{|U} : U \to \mathrm{Im}(f) : x \mapsto f(x).$$

Wähle eine Basis $(d_1, \ldots, d_r)$ von $\mathrm{Im}(f)$. Dann ist $(b_1, \ldots, b_r)$ mit $b_i := (f_{|U})^{-1}(d_i)$ eine Basis von U, nach Aufgabe A4.38. Wähle nun eine Basis $(b_{r+1}, \ldots, b_n)$ von $\mathrm{Ker}(f)$, dann ist $B = (b_1, \ldots, b_n)$ eine Basis von $V = U \oplus \mathrm{Ker}(f)$. Ergänze ferner $(d_1, \ldots, d_r)$ zu einer Basis $D = (d_1, \ldots, d_m)$ von W vermöge Korollar A4.28. Dann:

$$f(b_i) = \begin{cases} d_i, & i = 1, \ldots, r, \\ 0, & i = r+1, \ldots, n. \end{cases}$$

Also hat $M_D^B(f)$ die gewünschte Gestalt. □

So wie für eine lineare Abbildung eine Matrixdarstellung in Normalform existiert, existiert auch für jede Matrix eine.

Korollar A6.38 (Normalform einer Matrix bezüglich Äquivalenz)
Zu $A \in \mathrm{Mat}(m \times n, K)$ mit $r = \mathrm{rang}(A)$ existieren Matrizen $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$ mit

$$S \circ A \circ T = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right). \tag{195}$$

Die rechte Seite heißt die *Normalform* von A bezüglich Äquivalenz (siehe Bemerkung A6.40).

Beweis: Anwendung des Satzes A6.37 auf $f_A : K^n \to K^m$ liefert B und D von K^n bzw. K^m mit

$$\left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right) = M_D^B(f_A) = T_D^F \circ M_F^E(f_A) \circ T_E^B = T_D^F \circ A \circ T_E^B.$$

Die Behauptung folgt also, da $S := T_D^F$ und $T := T_E^B$ invertierbar sind. □

Beispiel A6.39
Die folgende Matrix $A \in \mathrm{Mat}(3 \times 4, \mathbb{R})$ 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 A6.40 (Normalform als guter Repräsentant der Äquivalenzklasse)
Aus Korollar A6.38 folgt, daß zwei Matrizen genau äquivalent sind, wenn sie den gleichen Rang haben. $\mathrm{Mat}(m \times n, K)$ 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 $\overline{A}$, $A \in \mathrm{Mat}(m \times n, K)$, 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*.

Wir wollen die Normalform einer Matrix bezüglich Äquivalenz nun ausnutzen, um Eigenschaften wie den Rang oder die Invertierbarkeit der Matrix und ihrer Transponierten zueinander in Beziehung zu setzen.

Korollar A6.41 (Zeilen- und Spaltenrang)

a. $A \in \mathrm{Mat}_n(K)$ ist genau dann invertierbar, wenn A^t invertierbar ist. In dem Fall gilt
$$\left(A^t\right)^{-1} = \left(A^{-1}\right)^t.$$

b. Für eine Matrix $A \in \mathrm{Mat}(m \times n, K)$ gilt
$$\mathrm{rang}(A) = \mathrm{rang}\left(A^t\right).$$
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 A5.27 invertierbar ist mit Inverser $\left(A^t\right)^{-1} = \left(A^{-1}\right)^t$. Ist umgekehrt A^t invertierbar, so ist nach dem eben gezeigten auch $A = (A^t)^t$ invertierbar.

b. Nach Korollar A6.38 finden wir invertierbare Matrizen $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$, 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 $r = \mathrm{rang}(A)$ gegeben ist. Dann ist aber

$$T^t \circ A^t \circ S^t = \left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right)^t \in \mathrm{Mat}(n \times m, K)$$

ebenfalls eine Matrix in Normalform. Es gilt also

$$\mathrm{rang}\left(T^t \circ A^t \circ S^t\right) = r$$

und wegen Teil a. ist die Matrix $T^t \circ A^t \circ S^t$ äquivalent zu A^t, so daß sie nach Proposition A6.29 den gleichen Rang hat wie A^t.

□

Beispiel A6.42

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 Rang 3; die drei Spalten sind linear unabhängig:

$$A^t = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \\ 3 & 2 & 3 \\ 5 & 2 & 1 \end{pmatrix}.$$

H) Algorithmus zur Berechnung der Normalform einer Matrix

Korollar A6.43 (Normalform einer Matrix)

Sei $A \in \mathrm{Mat}(m \times n, K)$ mit $r = \mathrm{rang}(A)$, so läßt sich A durch endlich viele elementare Zeilen- und Spaltenoperationen auf die folgende Form bringen:

(196)
$$\left(\begin{array}{c|c} \mathbb{1}_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

Beweis: Die Aussage folgt aus Korollar A6.38 und Korollar A6.34, da elementare Operationen nach Bemerkung A2.18 durch Multiplikation mit Elementarmatrizen realisierbar sind. □

Algorithmus A6.44 (Normalform-Algorithmus)

INPUT: $A \in \mathrm{Mat}(m \times n, K)$.

OUTPUT: Normalform $\mathrm{NF}(A)$ von A bezüglich Äquivalenz sowie die zugehörigen Transformationsmatrizen $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$

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.

Wir wollen nun noch ein Beispiel für die Berechung einer Normalform (196) geben.

Beispiel A6.45

Durch elementare Zeilen und Spaltenoperationen überführt man A_λ, $\lambda \in K$, in Normalform:

$\mathbb{1}_m$	A_λ	$\mathbb{1}_n$	
$\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}$	$\begin{matrix}1&0&\lambda\\0&1&0\\\lambda&0&1\end{matrix}$	$\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}$	$ZIII \mapsto ZIII - \lambda \cdot ZI$
$\begin{matrix}1&0&0\\0&1&0\\-\lambda&0&1\end{matrix}$	$\begin{matrix}1&0&\lambda\\0&1&0\\0&0&1-\lambda^2\end{matrix}$	$\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}$	$SIII \mapsto SIII - \lambda \cdot SI$
$\begin{matrix}1&0&0\\0&1&0\\-\lambda&0&1\end{matrix}$	$\begin{matrix}1&0&0\\0&1&0\\0&0&1-\lambda^2\end{matrix}$	$\begin{matrix}1&0&-\lambda\\0&1&0\\0&0&1\end{matrix}$	falls $\lambda = \pm 1$ fertig, sonst $SIII \mapsto \frac{1}{1-\lambda^2} \cdot SIII$
$\begin{matrix}1&0&0\\0&1&0\\-\lambda&0&1\end{matrix}$	$\begin{matrix}1&0&0\\0&1&0\\0&0&1\end{matrix}$	$\begin{matrix}1&0&-\frac{\lambda}{1-\lambda^2}\\0&1&0\\0&0&\frac{1}{1-\lambda^2}\end{matrix}$	
S	$\mathrm{NF}(A_\lambda)$	T	

Für die Normalform $\mathrm{NF}(A) = SAT$ erhalten wir also

$$\mathrm{NF}(A_\lambda) = \begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}, \quad S = \begin{pmatrix}1&0&0\\0&1&0\\-\lambda&0&1\end{pmatrix}, \quad T = \begin{pmatrix}1&0&-\frac{\lambda}{1-\lambda^2}\\0&1&0\\0&0&\frac{1}{1-\lambda^2}\end{pmatrix},$$

falls $\lambda \neq \pm 1$, und sonst

$$\mathrm{NF}(A_\lambda) = \begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}, \quad S = \begin{pmatrix}1&0&0\\0&1&0\\-\lambda&0&1\end{pmatrix}, \quad T = \begin{pmatrix}1&0&-\lambda\\0&1&0\\0&0&1\end{pmatrix}.$$

Insbesondere gilt, $\mathrm{rang}(A) = 3$ für $\lambda^2 \neq 1$ und $\mathrm{rang}(A) = 2$ sonst.

I) Algorithmus zum Test auf Injektivität und Surjektivität

Bemerkung A6.46

Sei $A \in \mathrm{Mat}(m \times n, K)$. Aus der Dimensionsformel für lineare Abbildungen A5.24

$$\dim_K \big(\mathrm{Ker}(f_A)\big) = n - \mathrm{rang}(A)$$

folgt unmittelbar:

- f_A ist injektiv $\iff$ $\mathrm{rang}(A) = n$.
- f_A ist surjektiv $\iff$ $\mathrm{rang}(A) = m$.
- f_A ist bijektiv $\iff$ $\mathrm{rang}(A) = n = m$.

Algorithmus A6.47 (Test auf Injektivität / Surjektivität / Bijektivität)
INPUT: $A \in \mathrm{Mat}(m \times n, K)$.
OUTPUT: Meldung, ob f_A injektiv, surjektiv oder bijektiv ist.

1. **Schritt:** Bestimme den Rang r von A.
2. **Schritt:** Ist $r = m = n$, gib "f_A ist bijektiv" zurück. Ist $r = m < n$, gib "f_A ist surjektiv" zurück. Ist $r = n < m$, gib "f_A ist injektiv" zurück.

J) Ringe und Moduln

Bemerkung A6.48 (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 A6.6, A6.7, A6.15 und A6.16 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 A6.49
Für Matrizen $A \in \mathrm{Mat}(n \times p, K)$ und $B \in \mathrm{Mat}(m \times n, K)$ gilt:

$$\mathrm{rang}(B \circ A) \leq \min\big\{\mathrm{rang}(A), \mathrm{rang}(B)\big\}.$$

Aufgabe A6.50 (f-invarianter Unterraum)
Sei $f : V \longrightarrow V$ eine K-lineare Abbildung und $U \leq V$ ein Unterraum mit $f(U) \subseteq U$; U heißt dann ein *f-invarianter Unterraum*. Zeige, daß die Abbildungen

$$f_U : U \longrightarrow U : x \mapsto f(x)$$

und

$$f_{V/U} : V/U \longrightarrow V/U : \overline{x} \mapsto \overline{f(x)}$$

wohldefiniert und K-linear sind.

Aufgabe A6.51 (Zyklische Unterräume)
Sei $f \in \mathrm{End}_K(V)$, $0 \neq x \in V$ und $m \in \mathbb{N}$ minimal mit $f^{m-1}(x) \neq 0$ und $f^m(x) = 0$.

a. Zeige, $B = \big(f^{m-1}(x), f^{m-2}(x), \ldots, f(x), x\big)$ ist eine Basis von $U = \mathrm{Lin}\,(B)$.

b. Zeige, U ist f-invariant.

c. Bestimme $M_B^B(f_U)$.

Wir nennen U einen *zyklischen Unterraum* von V.

Aufgabe A6.52
Für eine Matrix $A \in \mathrm{Mat}(m \times n, K)$ bezeichne $\mathrm{ZR}(A)$ die lineare Hülle der Zeilen von A und $\mathrm{SR}(A)$ die lineare Hülle der Spalten von A.
Zeige für $A \in \mathrm{Mat}(m \times n, K)$, $S \in \mathrm{Gl}_m(K)$ und $T \in \mathrm{Gl}_n(K)$

$$\mathrm{ZR}(A) = \mathrm{ZR}(SA) \quad \text{und} \quad \mathrm{SR}(A) = \mathrm{SR}(AT).$$

Aufgabe A6.53
Betrachte den Vektorraum P_n der Polynome vom Grad höchstens n (siehe Beispiel A3.6) mit Basis $B = (t^0, t^1, \ldots, t^n)$ 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 A3.21 bereits wissen, daß sie K-linear ist.

a. Berechne die Matrixdarstellung $M_B^B(d)$ und den Rang von d.

b. 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 $M_D^D(d)$.

Aufgabe A6.54
Berechne den Rang der folgenden Matrix in Abhängigkeit von a und b:

$$\begin{pmatrix} 0 & b & b & b \\ a & 0 & b & b \\ a & a & 0 & b \end{pmatrix} \in \mathrm{Mat}(3 \times 4, \mathbb{R}).$$

Aufgabe A6.55

Bestimme die Inverse der Matrix

$$\begin{pmatrix} 1 & 3 & -1 & 4 \\ 2 & 5 & -1 & 3 \\ 0 & 4 & -3 & 1 \\ -3 & 1 & -5 & -2 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{R}).$$

Aufgabe A6.56

Transformiere die folgende Matrix A in Normalform bezüglich Äquivalenz und gib auch die Transformationsmatrizen S und T an:

$$A = \begin{pmatrix} 1 & -2 & 3 & 0 \\ 2 & -7 & 10 & -1 \\ -2 & 4 & -7 & 2 \\ 3 & -5 & 7 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{R}).$$

Aufgabe A6.57

Überprüfe die folgende Abbildung auf Injektivität und Surjektivität:

$$g : \mathbb{R}^4 \to \mathbb{R}^4 \ : \ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} 2x_1 + x_2 + x_4 \\ 3x_1 + x_2 + x_3 + x_4 \\ 2x_1 + x_3 + x_4 \\ 2x_1 + x_2 + x_3 \end{pmatrix}.$$

Aufgabe A6.58

Berechne den Rang der Matrix

$$A = \begin{pmatrix} 0 & 0 & 2 & 6 & 4 & 0 \\ -1 & -2 & 1 & 2 & 1 & 0 \\ 2 & 4 & 0 & 2 & 2 & 1 \\ 1 & 2 & 0 & 1 & 1 & 0 \\ -2 & -4 & 2 & 4 & 2 & 0 \end{pmatrix} \in \mathrm{Mat}(5 \times 6, \mathbb{Q}).$$

Aufgabe A6.59

Berechne die Inverse der folgenden Matrix

$$A = \begin{pmatrix} 3 & -2 & 0 & 1 \\ 4 & -3 & 0 & 2 \\ 0 & 1 & 1 & 0 \\ 2 & -2 & 0 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{R}).$$

Aufgabe A6.60

Transformiere die folgende Matrix B in Normalform bezüglich Äquivalenz und gib auch die Transformationsmatrizen S und T an:

$$B = \begin{pmatrix} 1 & 2 & 1 & 2 \\ 0 & 1 & 1 & 0 \\ 2 & 1 & 2 & 1 \end{pmatrix} \in \mathrm{Mat}(3 \times 4, \mathbb{F}_3).$$

Bestimme insbesondere den Rang der Matrix. Beachte: Grundkörper ist $\mathbb{F}_3$!

Aufgabe A6.61

Sei $V = \mathrm{Mat}_2(\mathbb{R})$ der $\mathbb{R}$-Vektorraum der 2×2-Matrizen über $\mathbb{R}$, $E = (E_1^1, E_1^2, E_2^1, E_2^2)$ die kanonische Basis von V und

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}).$$

Zeige, dass die Abbildung

$$f : V \longrightarrow V \quad \text{mit} \quad f(X) = A \circ X$$

$\mathbb{R}$-linear ist und berechne die Matrixdarstellung $M_E^E(f)$.

A7 Lineare Gleichungssysteme

Definition A7.1 (Lineare Gleichungssysteme)

a. Ein *lineares Gleichungssystem* über K

$$\text{(LGS)} \qquad \begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \dots & + & a_{2n}x_n & = & b_2 \\ & & & & \vdots & & & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \dots & + & a_{mn}x_n & = & b_m \end{array}$$

besteht aus m Gleichungen in n *Unbestimmten* oder *Variablen* $x_1, \dots, x_n$ mit $a_{ij}, b_i \in K$ für $1 \le i \le m$ und $1 \le j \le n$.
Da sich (LGS) mit $A = (a_{ij})$, $b = (b_1, \dots, b_m)^t$ und $x = (x_1, \dots, x_n)^t$ vermittels Matrixmultiplikation auch kurz schreiben läßt als

$$Ax = b,$$

sprechen wir meist von dem *linearen Gleichungssystem* $Ax = b$.

b. Die Matrix

$$A = (a_{ij}) = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in \mathrm{Mat}(m \times n, K)$$

heißt *Koeffizientenmatrix* und der Vektor $b = (b_1, \dots, b_m)^t \in K^m$ die *Inhomogenität* des Gleichungssystems (LGS). Ferner heißt die Matrix

$$(A \mid b) := \left(\begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array}\right) \in \mathrm{Mat}\big(m \times (n+1), K\big)$$

die *erweiterte Koeffizientenmatrix* von (LGS).

c. Das lineare Gleichungssystem (LGS) heißt *homogen*, falls $b = 0_{K^m}$ der Nullvektor in K^m ist. Ansonsten heißt das System *inhomogen*.

d. Ist ein lineares Gleichungssystem $Ax = b$ gegeben, so heißt das Gleichungssystem $Ax = 0$ (mit $0 = 0_{K^m}$) das *zugehörige homogene Gleichungssystem*.

e. Ein Vektor $c = (c_1, \dots, c_n)^t \in K^n$ heißt *Lösung* von (LGS), wenn die Gleichung $Ac = b$ erfüllt ist. Die Menge aller Lösungen von (LGS) wird mit

$$\mathrm{Lös}(A, b) := \big\{c \in K^n \mid Ac = b\big\}.$$

bezeichnet.

Bei einem LGS sind also Körperelemente a_{ij} und b_i fest vorgegeben und für die Unbestimmten x_j werden Körperelemente c_j gesucht, die das Gleichungssystem lösen.

Wir werden im Folgenden mehrere Verfahren zur Lösung eines LGS kennenlernen und uns mit der Struktur der Lösungsmenge $\text{Lös}(A,b)$ beschäftigen. Eine wichtige Rolle spielt dabei die lineare Abbildung $f_A : K^n \to K^m$.

Bemerkung A7.2 (Struktur des Lösungraums)
Es sei $A \in \text{Mat}(m \times n, K)$ und $b \in K^m$.

a. Aus den Definitionen folgt unmittelbar

$$\text{Lös}(A,0) = \{c \in K^n \mid Ac = 0\} = \text{Ker}(f_A),$$

so daß $\text{Lös}(A,0)$ ein Unterraum des K^n ist mit Dimension

$$\dim_K \big(\text{Lös}(A,0)\big) = \dim_K \big(\text{Ker}(f_A)\big) = n - \text{rang}(A).$$

Insbesondere ist $Ax = 0$ genau dann eindeutig lösbar, wenn $\text{rang}(A) = n$.

b. Ebenfalls anhand der Definitionen sieht man, daß das lineare Gleichungssystem $Ax = b$ genau dann eine Lösung besitzt, wenn $b \in \text{Im}(f_A) = \{Ac \mid c \in K^n\}$.

Beispiel A7.3
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

$$(A \mid b) = \left(\begin{array}{ccc|c} 1 & 2 & 1 & 1 \\ 2 & 3 & 0 & 1 \\ 0 & 1 & -1 & 0 \end{array}\right) \in \text{Mat}(3 \times 4, \mathbb{R}).$$

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 $\big(0, \frac{1}{3}, \frac{1}{3}\big)^t$ die einzige Lösung ist.

Wir geben zunächst ein Kriterium für die Lösbarkeit eines Gleichungssystems.

Satz A7.4 (Kriterium für die Lösbarkeit eines LGS)
Ein Gleichungssystem $Ax = b$ ist genau dann lösbar, wenn $\text{rang}(A) = \text{rang}(A \mid b)$.

Beweis: Wir beachten, daß $\mathrm{Im}(f_A)$ von den Spaltenvektoren $a^1, \ldots, a^n$ von A erzeugt wird, und erhalten deshalb:

$$\begin{aligned} Ax = b \text{ lösbar} &\iff b \in \mathrm{Im}(f_A) = \mathrm{Lin}\left(a^1, \ldots, a^n\right) \\ &\iff b \text{ ist Linearkombination von } a^1, \ldots, a^n \\ &\iff \mathrm{Im}(f_A) = \mathrm{Lin}\left(a^1, \ldots, a^n\right) = \mathrm{Lin}\left(a^1, \ldots, a^n, b\right) = \mathrm{Im}(f_{(A|b)}) \\ &\iff \mathrm{rang}(A) = \dim_K\left(\mathrm{Im}(f_A)\right) = \dim_K\left(\mathrm{Im}(f_{(A|b)})\right) = \mathrm{rang}(A \mid b), \end{aligned}$$

wobei wir für die letzte Äquivalenz berücksichtigen, daß $\mathrm{Im}(f_A) \subseteq \mathrm{Im}(f_{(A|b)})$ gilt. □

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 A7.5 (Struktur von Lös(A, b) als affiner Raum)
Seien $A \in \mathrm{Mat}(m \times n, K)$, $b \in K^m$ und sei $c \in K^n$ eine Lösung des linearen Gleichungssystems $Ax = b$. Dann gilt:

$$\mathrm{Lös}(A, b) = c + \mathrm{Lös}(A, 0).$$

Beweis: Sei zunächst $y \in \mathrm{Lös}(A, 0)$. Dann gilt:

$$A\big(c + y\big) = Ac + Ay = b + 0 = b,$$

also ist $c + y \in \mathrm{Lös}(A, b)$.

Ist umgekehrt $x \in \mathrm{Lös}(A, b)$. Dann gilt für $y := x - c$

$$Ay = A\big(x - c\big) = Ax - Ac = b - b = 0,$$

also ist $y \in \mathrm{Lös}(A, 0)$. Aber damit ist $x = c + y \in c + \mathrm{Lös}(A, 0)$. □

Wir wollen nun einen Algorithmus kennenlernen, der es uns erlaubt, die Lösung eines linearen Gleichungssystems $Ax = b$ 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 $(A \mid b)$ in reduzierte Zeilen-Stufen-Form.

Lemma A7.6 (Element. Zeilenoperationen ändern Lösungsraum nicht.)
Sind $A, A' \in \mathrm{Mat}(m \times n, K)$ und $b, b' \in K^m$ und entsteht die Matrix $(A' \mid b')$ aus $(A \mid b)$ durch elementare Zeilenoperationen, so gilt

$$\mathrm{Lös}(A, b) = \mathrm{Lös}\left(A', b'\right).$$

Beweis: Daß $(A' \mid b')$ aus $(A \mid b)$ durch elementare Zeilenoperationen hervorgeht, bedeutet, daß es eine invertierbare Matrix $S \in \mathrm{Gl}_m(K)$ gibt mit $A' = SA$ und $b' = Sb$.

Ist nun $c \in \mathrm{Lös}(A, b)$, dann gilt $Ac = b$ und damit

$$b' = Sb = SAc = A'c.$$

Also ist $c \in \mathrm{Lös}\left(A', b'\right)$.

Ist andererseits $c \in \mathrm{Lös}\left(A', b'\right)$, dann gilt $A'c = b'$ und damit

$$b = S^{-1}b' = S^{-1}A'c = Ac.$$

Also ist $c \in \mathrm{Lös}(A, b)$. □

Bemerkung A7.7

Aus Lemma A7.6 und Satz A2.20 folgt, daß wir die erweiterte Koeffizientenmatrix eines Gleichungssystems $Ax = b$ 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 A7.8 (Lösbarkeitskriterium für LGS mittels Gauß-Algorithmus)

Sei $A \in \mathrm{Mat}(m \times n, K)$ eine Matrix in Zeilen-Stufen-Form und $b \in K^m$. Die erweiterte Koeffizientenmatrix habe die Gestalt

(197)

$$(A \mid b) = \left(\begin{array}{ccccccccccc|c}
0 & \dots & 0 & a_{1j_1} & * & \dots & \dots & \dots & \dots & \dots & * & b_1 \\
0 & \dots & \dots & 0 & \dots & 0 & a_{2j_2} & * & \dots & \dots & * & b_2 \\
\vdots & & & & & & & \ddots & & & \vdots & \vdots \\
0 & \dots & \dots & \dots & \dots\dots & \dots & \dots & 0 & a_{rj_r} & * & * & b_r \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & b_{r+1} \\
\vdots & & & & & & & & & & \vdots & \vdots \\
0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & b_m
\end{array}\right)$$

mit Pivots $a_{ij_i} \neq 0$ für $i = 1, \dots, r$. Dann gilt:

a. $Ax = b$ ist genau dann lösbar, wenn $b_{r+1} = \dots = b_m = 0$.

b. Sind $b_{r+1} = \dots = b_m = 0$ und $r = n$, so hat $Ax = b$ *genau eine* Lösung.

c. Sind $b_{r+1} = \dots = b_m = 0$ und ist $r < n$, so hat $Ax = b$ mehr als eine Lösung. Genauer $\mathrm{Lös}(A, b) = c + \mathrm{Lös}(A, 0)$, wobei c eine spezielle Lösung ist und $\mathrm{Lös}(A, 0)$ die Dimension $n - r$ hat.

Beweis: Die Aussagen folgen aus Satz A7.4, Satz A7.5 und Bemerkung A7.2. □

A) Der Gauß-Algorithmus zur Lösung eines (LGS)

Algorithmus A7.9 (Algorithmus zur Lösung eines LGS)

INPUT: Die erweiterte Koeffizientenmatrix $(A \mid b)$ eines LGS $Ax = b$.

OUTPUT: Eine spezielle Lösung c von $Ax = b$ und eine Basis B von $\text{Lös}(A, 0)$, sofern das Gleichungssystem lösbar ist.

1. **Schritt:** Berechne eine reduzierte Zeilen-Stufen-Form $(A' \mid b')$ von $(A \mid b)$ mit $r = \text{rang}(A')$.
2. **Schritt:** Ist $b'_{r+1} \neq 0$, dann ist das LGS nicht lösbar.
3. **Schritt:** Überführe $(A' \mid b')$ in eine $n \times (n+1)$-Matrix $(A'' \mid b'')$ 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 $c := b''$ und die Spalten von A'', die eine -1 auf der Diagonale haben, sind eine Basis von $\text{Lös}(A, 0)$.

Beispiel A7.10

Wir betrachten das Gleichungssystem:

$$\begin{aligned} x_1 + x_2 + x_3 - 2x_4 &= 1 \\ x_1 + x_2 - x_3 &= -1 \\ 3x_1 + 3x_2 + x_3 - 4x_4 &= 1 \end{aligned} \tag{198}$$

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 $\text{rang}(A) = \text{rang}(A \mid b) = 2$, 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 (198) und die Spalten 2 und 4 als Basis

$$B = \big((1, -1, 0, 0)^t, (-1, 0, -1, -1)^t\big)$$

des Lösungsraums $\text{Lös}(A, 0)$ des homogenen Gleichungsystems $Ax = 0$. Insgesamt gilt damit

$$\text{Lös}(A, b) = c + \text{Lös}(A, 0) = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} + s \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + t \cdot \begin{pmatrix} -1 \\ 0 \\ -1 \\ -1 \end{pmatrix} \;\middle|\; s, t \in K \right\}.$$

Wir wollen nun einige Algorithmen angeben, denen der Algorithmus zur Lösung eines linearen Gleichungssystems zugrunde liegt.

B) Algorithmus zur Berechnung des Kerns von f_A

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 A7.11 (Kern von f_A)
INPUT: $A \in \text{Mat}(m \times n, K)$.
OUTPUT: Eine Basis von $\text{Ker}(f_A)$.

1. **Schritt:** Bestimme eine Lösung $\big(c, B\big)$ von $Ax = 0$ gemäß A7.9.
2. **Schritt:** Gib B als Basis zurück.

Beispiel A7.12
Wir wollen den Kern der Linearen Abbildung $f_A : K^4 \longrightarrow K^3$ berechnen, die durch die Koeffizientenmatrix

$$A = \begin{pmatrix} 1 & 1 & 1 & -2 \\ 1 & 1 & -1 & 0 \\ 3 & 3 & 1 & -4 \end{pmatrix}$$

in Beispiel A7.10 gegeben ist. Dann gehen wir wie in Beispiel A7.10 vor, wobei wir die Inhomogenität durch den Nullvektor ersetzen oder einfach gänzlich ignorieren können. Die Rechnungen ändern sich nicht und wir erhalten wie dort

$$B = \big((1,-1,0,0)^t, (-1,0,-1,-1)^t\big)$$

als Basis von $\mathrm{Ker}(f_A) = \mathrm{Lös}(A,0)$.

C) Algorithmus zur Berechnung der Transformationsmatrix $T^B_{B'}$

Sind $B = (b_1,\ldots,b_n)$ und $B' = (b'_1,\ldots,b'_n)$ zwei Basen des 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 Gleichungsysteme Lösen, bei denen die Koeffizientenmatrix stets $b'_1,\ldots,b'_n$ als Spaltenvektoren hat und bei denen die Inhomogenitäten durch die Vektoren $b_1,\ldots,b_n$ gegeben werden. Da die Koeffizientenmatrix sich nicht ändert, können wir die n Gleichungssysteme simultan lösen, indem wir der erweiterten Koeffizientenmatrix gleich alle Vektoren $b_1,\ldots,b_n$ als zusätzliche Spalten anhängen.

Algorithmus A7.13 (Transformationsmatrix $T^B_{B'}$)
INPUT: Zwei Basen $B = (b_1,\ldots,b_n)$ und $B' = (b'_1,\ldots,b'_n)$ im K^n.
OUTPUT: Die Transformationsmatrix $T^B_{B'}$.

1. **Schritt:** Schreibe die Vektoren $b'_1,\ldots,b'_n,b_1,\ldots,b_n$ in dieser Reihenfolge als Spalten in eine Matrix A.
2. **Schritt:** Bringe A auf reduzierte ZSF.
3. **Schritt:** Die letzten n Spalten von $\mathrm{rZSF}(A)$ sind $T^B_{B'}$.

Beispiel A7.14
Seien die zwei Basen $B = \big((1,1)^t,(1,-1)^t\big)$ und $B' = \big((1,2)^t,(-1,0)^t\big)$ 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 $f : K^n \longrightarrow K^m$ bezüglich zweier Basen $B = (b_1, \ldots, b_n)$ von K^n und $D = (d_1, \ldots, d_m)$ von K^m berechnet. Die Grundidee ist ähnlich wie beim Algorithmus zur Berechnung der Transformationsmatrix.

Algorithmus A7.15 (Matrixdarstellung $M_D^B(f)$)

INPUT: Eine lineare Abbildung $f : K^n \longrightarrow K^m$, eine Basis $B = (b_1, \ldots, b_n)$ von K^n und eine Basis $D = (d_1, \ldots, d_m)$ im K^m.

OUTPUT: Die Matrixdarstellung $M_D^B(f)$.

1. **Schritt:** Schreibe die Vektoren $d_1, \ldots, d_m, f(b_1), \ldots, f(b_n)$ 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 A7.16

Für die Basen $B = \big((1,0,1)^t, (1,1,0)^t, (0,0,1)^t\big)$ des K^3 und $D = \big((1,1)^t, (1,-1)^t\big)$ des K^2 sowie die lineare Abbildung

$$f : K^3 \longrightarrow K^2 : (x,y,z)^t \mapsto (x+y+z, x-z)^t$$

wollen wir die Matrixdarstellung $M_D^B(f)$ berechnen.

D		$f(B)$		
1	1	2	2	1
1	−1	0	1	−1
1	1	2	2	1
0	−2	−2	−1	−2
1	1	2	2	1
0	1	1	$\frac{1}{2}$	1
1	0	1	$\frac{3}{2}$	0
0	1	1	$\frac{1}{2}$	1
$\mathbb{1}_2$		$M_D^B(f)$		

Bemerkung A7.17

Natürlich könnte man auch zunächst die Matrixdarstellung $M_F^E(f)$ 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^F 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 $F = (y_1, \ldots, y_r)$, die wir in die Basis $B = (x_1, \ldots, x_n)$ hineintauschen wollen, sukzessive als Linearkombination der Basisvektoren in (dem veränderten) B darstellen, d.h. wir müssen immer wieder lineare Gleichungssysteme lösen.

Algorithmus A7.18 (Austauschverfahren von Steinitz)

INPUT: Eine Basis $B = (x_1, \ldots, x_n)$ und eine linear unabhängige Familie $F = (y_1, \ldots, y_r)$ von Vektoren in $V = \mathrm{Lin}\,(B) \subseteq K^n$.

OUTPUT: Eine Basis B' von V, die F enthält.

1. Schritt: Für $i = 1, \ldots, r$ 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 A7.19

Betrachte die linear unabhängige Familie $F = (y_1, y_2) = \big((1,2,1)^t, (1,2,2)^t\big)$ und die Basis $B = (x_1, x_2, x_3) = \big((1,1,0)^t, (1,0,1)^t, (0,0,1)^t\big)$. 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 A7.20 (Berechnung eines Komplementes / einer Basis für K^n/U)
Will man ein Komplement eines Unterraums 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 A5.23). Der obige Algorithmus erlaubt also auch die Berechnung eine Komplementes und einer Basis eines Faktorraums.

F) Algorithmus zur Berechnung von 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 A7.21 (Gleichungen eines Unterraumes)
INPUT: Eine Familie $F = (x_1, \ldots, x_m)$ von Vektoren im K^n.
OUTPUT: Eine Matrix $A \in \mathrm{Mat}(k \times n, K)$ mit $\mathrm{Lös}(A, 0) = \mathrm{Lin}(F)$.

1. **Schritt:** Schreibe die Vektoren aus F als Zeilen in eine Matrix $B \in \mathrm{Mat}(m \times n, K)$ und bestimme eine Basis $(y_1, \ldots, y_k)$ von $\mathrm{Ker}(f_B) = \mathrm{Lös}(B, 0)$.
2. **Schritt:** Schreibe die $y_1, \ldots, y_k$ als Zeilenvektoren in eine Matrix A.
3. **Schritt:** Gib A zurück.

Beispiel A7.22
Finde ein lineares Gleichungssystem $Ax = 0$ mit Lösungsmenge

$$\mathrm{Lös}(A, 0) = \mathrm{Lin}\left((1, 2, 1)^t, (0, 1, 0)^t\right) \leq \mathbb{R}^3.$$

Dazu Bilden wir die 2×3-Matrix

$$B = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

und berechnen ihren Kern:

$$\mathrm{rZSF}(B, 0) = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \end{pmatrix}.$$

Somit ist der Vektor in der dritten Spalte eine Basis von $\mathrm{Lös}(B, 0)$ und wir erhalten

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

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 A7.23 (Durchschnitt zweier Unterräume)
INPUT: Zwei Familien F und G von Vektoren in K^n.
OUTPUT: Eine Basis des Schnitts von $\mathrm{Lin}\,(F)$ und $\mathrm{Lin}\,(G)$.

1. **Schritt:** Bestimme Matrizen A und A' gemäß A7.21, so daß $\mathrm{Lin}\,(F) = \mathrm{Lös}(A, 0)$ und $\mathrm{Lin}\,(G) = \mathrm{Lös}\,(A', 0)$.
2. **Schritt:** Bilde aus den Zeilen von A und A' eine gemeinsame Matrix A''.
3. **Schritt:** Bestimme eine Basis B von $\mathrm{Ker}(f_{A''}) = \mathrm{Lös}\,(A'', 0)$ gemäß A7.11 und gib B zurück.

Beispiel A7.24
Wir wollen den Durchschnitt der Unterräume

$$U = \mathrm{Lin}\,\big((1, 2, 1)^t, (0, 1, 0)^t\big)$$

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 $\mathrm{Lös}(A, 0)$ bereits in Beispiel A7.22 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

$$(A'',0)=\left(\begin{array}{cccc}1&0&-1&0\\1&1&1&0\end{array}\right)\mapsto\left(\begin{array}{cccc}1&0&-1&0\\0&1&2&0\end{array}\right)\mapsto\left(\begin{array}{cccc}1&0&-1&0\\0&1&2&0\\0&0&-1&0\end{array}\right).$$

Damit ist die dritte Spalte

$$B=\left((-1,2,-1)^t\right)$$

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 A7.25 (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:

$$v=4s,\ s=t+5,\ (v+5)+(s+5)=6(t+5).$$

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:

$$\left(\begin{array}{ccc|c}1&-4&0&0\\0&1&-1&5\\1&1&-6&20\end{array}\right)\mapsto\left(\begin{array}{ccc|c}1&-4&0&0\\0&1&-1&5\\0&5&-6&20\end{array}\right)\mapsto\left(\begin{array}{ccc|c}1&-4&0&0\\0&1&-1&5\\0&0&1&5\end{array}\right)$$

$$\mapsto\left(\begin{array}{ccc|c}1&-4&0&0\\0&1&0&10\\0&0&1&5\end{array}\right)\mapsto\left(\begin{array}{ccc|c}1&0&0&40\\0&1&0&10\\0&0&1&5\end{array}\right)$$

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 A7.26 (Schnitt zweier Ebenen)

Wir definieren eine *Ebene* im $\mathbb{R}^3$ als Lösungsmenge einer linearen Gleichung

$$E : a_1x_1 + a_2x_2 + a_3x_3 = b$$

mit $a_1, a_2, a_3, b \in \mathbb{R}$ und $a_i \neq 0$ für mindestens ein i.

Dies stimmt mit der Anschauung überein (sind alle a_i und b gleich 0, so erhalten wir als Lösungsmenge den ganzen $\mathbb{R}^3$, sind alle $a_i = 0$ und $b \neq 0$, 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 A7.27 (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

$$E_1 : x_1 + x_2 + 2x_2 = 3$$

und E_2 eine beliebige Ebene

$$E_2 : a_1x_1 + a_2x_2 + a_3x_3 = b.$$

Wir wollen feststellen für welche a_1, a_2, a_3, b entweder $E_1 \cap E_2$ eine Gerade, leer oder E_1 ist:

$$\left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ a_1 & a_2 & a_3 & b \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 1 & 2 & 3 \\ 0 & a_2 - a_1 & a_3 - 2a_1 & b - 3a_1 \end{array}\right).$$

Die letzte Gleichung lautet

$$(a_2 - a_1)x_2 + (a_3 - 2a_1)x_3 = b - 3a_1.$$

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):

$$E_1 \cap E_2 = \emptyset \quad \Leftrightarrow \quad a_2 - a_1 = a_3 - 2a_1 = 0, (b - 3a_1) \neq 0, \tag{199}$$

$$E_1 = E_2 \quad \Leftrightarrow \quad a_2 - a_1 = a_3 - 2a_1 = b - 3a_1 = 0. \tag{200}$$

In allen anderen Fällen ist $E_1 \cap E_2$ eine Gerade.

Im Fall $E_1 = E_2$ haben wir wieder ein Gleichungssystem (200) mit drei Gleichungen in den vier Unbestimmten a_1, a_2, a_3, b zu lösen:

$$\left(\begin{array}{cccc|c} -1 & 1 & 0 & 0 & 0 \\ -2 & 0 & 1 & 0 & 0 \\ -3 & 0 & 0 & 1 & 0 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} -1 & 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & -3 & 0 & 1 & 0 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} -1 & 1 & 0 & 0 & 0 \\ 0 & -2 & 1 & 0 & 0 \\ 0 & 0 & -\frac{3}{2} & 1 & 0 \end{array}\right)$$

$$\mapsto \left(\begin{array}{cccc|c} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & -\frac{2}{3} & 0 \end{array}\right) \mapsto \left(\begin{array}{cccc|c} 1 & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 1 & 0 & -\frac{1}{3} & 0 \\ 0 & 0 & 1 & -\frac{2}{3} & 0 \end{array}\right)$$

Als Lösung ergibt sich $a_1 = -\frac{b}{3}$, $a_2 = \frac{b}{3}$ und $a_3 = \frac{2b}{3}$, oder kurz

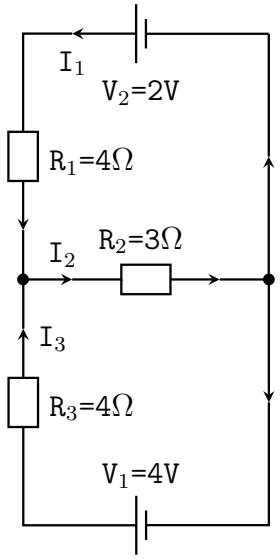
$$(a_1, a_2, a_3, b) = t \cdot (-1, 1, 2, 3)$$

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.

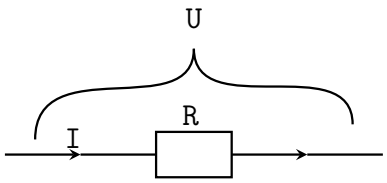
Beispiel A7.28 (Elektrische Netzwerke)

In einem einfachen elektrischen Netzwerk, wie z. B.

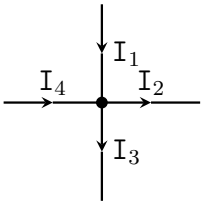


bezeichnet man mit U die Spannung, mit I den Strom und mit R den Widerstand, gemessen in Volt (V), Ampere (A) und Ohm (Ω) respektive. Dabei gelten folgende Gesetze:

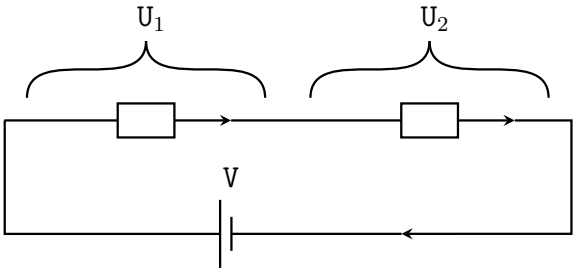
- *Ohmsches Gesetz*: Der Spannungsabfall über einen Widerstand ist das Produkt von Widerstand und Strom, $U=R\cdot I$.



- *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:

$$\begin{aligned} I_1 + I_3 &= I_2, && \text{(Knotengleichung)} \\ 4I_1 + 3I_2 &= 2, && \text{(1. Maschengleichung)} \\ 4I_3 + 3I_2 &= 4. && \text{(2. Maschengleichung)} \end{aligned}$$

Wir erhalten das folgende Gleichungssystem:

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 4 & 3 & 0 & 2 \\ 0 & 3 & 4 & 4 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 7 & -4 & 2 \\ 0 & 3 & 4 & 4 \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 7 & -4 & 2 \\ 0 & 0 & 40 & 22 \end{array}\right)$$
$$\mapsto \left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 0 & 1 & -\frac{4}{7} & \frac{2}{7} \\ 0 & 0 & 1 & \frac{11}{20} \end{array}\right) \mapsto \left(\begin{array}{ccc|c} 1 & 0 & 0 & \frac{1}{20} \\ 0 & 1 & 0 & \frac{3}{5} \\ 0 & 0 & 1 & \frac{11}{20} \end{array}\right)$$

woraus sich die folgende Lösung ergibt:

$$I_3 = \tfrac{11}{20}, I_2 = \tfrac{3}{5} \text{ und } I_1 = \tfrac{1}{20}.$$

Beispiel A7.29 (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

$$f'(x) = 3ax^2 + 2bx + c$$

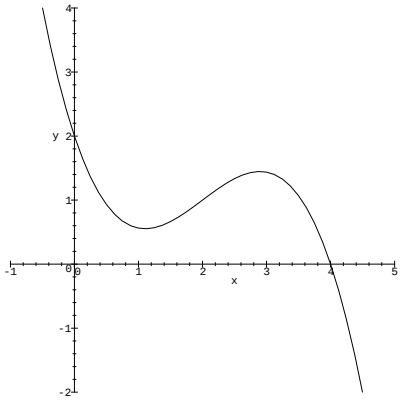
und

$$f(0) = 2, f(4) = 0, f'(0) = -3 \text{ und } f'(4) = -3$$

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 A7.30 (Leontieff-Modell)

Die folgende Planungsaufgabe zeigt, daß durchaus Gleichungen mit vielen Veränderlichen in der Praxis auftauchen.

Ein Konzern besitzt n Fabriken $F_1, \ldots, F_n$, 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, $k = 1, \ldots, n$; wir haben also einen Produktionsvektor $x = (x_1, \ldots, x_n)$.

Zur Herstellung von $x = (x_1, \ldots, x_n)$ werden

$$\sum_{k=1}^{n} a_{jk} x_k = a_{j1} x_1 + \cdots + a_{jn} x_n$$

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 $y = (y_1, \ldots, y_n)$ ist vorgegeben. Gesucht ist ein Produktionsvektor $x = (x_1, \ldots, x_n)$ mit

$$\begin{array}{rcl} x_1 - (a_{11} x_1 + \cdots + a_{1n} x_n) & = & y_1 \\ \vdots & & \vdots \\ x_n - (a_{n1} x_1 + \cdots + a_{nn} x_n) & = & y_n. \end{array}$$

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 $a_{j_k} \geq 0$).

(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

$$\left(\begin{array}{cc|c} 1 & -\frac{1}{2} & y_1 \\ -\frac{1}{3} & 1 & y_2 \end{array}\right) \mapsto \left(\begin{array}{cc|c} 1 & -\frac{1}{2} & y_1 \\ 0 & \frac{5}{6} & \frac{1}{3}y_1 + y_2 \end{array}\right) \mapsto \left(\begin{array}{cc|c} 1 & -\frac{1}{2} & y_1 \\ 0 & 1 & \frac{2}{5}y_1 + \frac{6}{5}y_2 \end{array}\right)$$
$$\mapsto \left(\begin{array}{cc|c} 1 & 0 & \frac{6}{5}y_1 + \frac{3}{5}y_2 \\ 0 & 1 & \frac{2}{5}y_1 + \frac{6}{5}y_2 \end{array}\right)$$

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 A7.31 (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:

(1) $$a_{11} + a_{13} = b_1,$$

(2) $$a_{23} = b_2.$$

$(1,1,1)^t$ 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:

$$\begin{array}{c} \\ (1) \\ (3) \\ (4) \\ (2) \end{array} \begin{array}{c} \begin{array}{ccccccc} a_{11} & a_{12} & a_{13} & a_{22} & a_{23} & b_1 & b_2 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & -1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix} \end{array} \mapsto \begin{array}{c} \begin{array}{ccccccc} a_{11} & a_{12} & a_{13} & a_{22} & a_{23} & b_1 & b_2 \end{array} \\ \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -1 \end{pmatrix} \end{array}.$$

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 A7.32 (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 A7.5 bleibt wahr. Alle Aussagen, die den Rang einer Matrix verwenden, sind jedoch nicht mehr richtig. Außerdem kann sich die Lösungmenge 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 A7.33

Berechne die Lösungsmenge des folgenden Gleichungssystems, sofern es lösbar ist:

$$\begin{array}{rcrcrcl} -x & + & 6y & + & 2z & = & 4 \\ 2x & - & 2y & - & z & = & 2 \\ 3x & - & 4y & - & 2z & = & 1 \end{array}$$

Aufgabe A7.34

Für welche $a, b \in \mathbb{R}$ besitzt das lineare Gleichungssystem

$$\begin{array}{rcrcrcl} ax & + & & & z & = & ab \\ -2x & + & by & + & az & = & -b \\ & & by & + & (a+1)z & = & b \end{array}$$

außer $(b, 1, 0)$ noch weitere Lösungen. Bestimme sie.

Aufgabe A7.35

Bestimme die Lösungsmenge des folgenden linearen Gleichungssystems über $\mathbb{R}$ in Abhängigkeit vom Parameter $t \in \mathbb{R}$:

$$\begin{array}{rcrcrcl} x & + & y & + & z & = & 1 \\ & & ty & + & z & = & 1 \\ tx & + & ty & + & z & = & 1+t \end{array}$$

Aufgabe A7.36

Bestimme eine Basis des Kerns und des Bildes von f_A mit

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ -1 & 0 & 1 & 1 & 1 \\ -1 & -1 & 0 & 1 & 1 \\ -1 & -1 & -1 & 0 & 1 \\ -1 & -1 & -1 & -1 & 0 \end{pmatrix} \in \mathrm{Mat}_5(\mathbb{R}).$$

Aufgabe A7.37

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 A7.38

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 A7.39

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 A7.40

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 A7.41

Wir betrachten

$$B = \big((1, 1, 1, 1)^t, (-1, 0, 0, 1)^t, (0, -1, 0, 1)^t, (0, 0, -1, 1)^t\big)$$

und

$$D = \big((1, 1, 0)^t, (0, 1, 1)^t, (0, 0, 1)^t\big).$$

a. Zeige, dass B eine Basis des $\mathbb{R}^4$ und D eine Basis des $\mathbb{R}^3$ ist.

b. Bestimme $M_D^B(f)$ für $f : \mathbb{R}^4 \longrightarrow \mathbb{R}^3 : (x_1, x_2, x_3, x_4)^t \mapsto (x_1 - x_2, x_3, x_2 + x_4)^t$.

c. Bestimme umgekehrt die Funktionsvorschrift für $g \in \mathrm{Hom}_{\mathbb{R}}\left(\mathbb{R}^4, \mathbb{R}^3\right)$ mit

$$M_D^B(g) = \begin{pmatrix} 1 & 0 & -1 & 0 \\ 2 & 1 & 3 & 1 \\ 0 & -1 & 2 & 0 \end{pmatrix}.$$

Aufgabe A7.42

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 $f \in \mathrm{End}_{\mathbb{R}}(V)$ 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 A7.43

Seien $B = \big((1,1,1)^t, (1,1,0)^t, (1,0,-1)^t\big)$ und $B' = \big((2,1)^t, (1,1)^t\big)$. E bzw. E' seien die kanonischen Basen des $\mathbb{R}^3$ bzw. des $\mathbb{R}^2$. Ferner sei $f \in \mathrm{Hom}_K\left(\mathbb{R}^3, \mathbb{R}^2\right)$ gegeben durch $f\big((x,y,z)^t\big) = (x - y + z, 2x + y)^t$.

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 $T_{B'}^{E'} \cdot M_{E'}^{E}(f) \cdot T_E^B = M_{B'}^{B}(f)$.

Aufgabe A7.44

Bestimme die Lösungsmenge des folgenden linearen Gleichungssystems über $\mathbb{R}$:

$$\begin{aligned} x_4 - x_5 &= 0, \\ x_1 + x_3 - x_4 + 2x_5 &= 1, \\ -x_1 + x_2 - x_3 &= 1, \\ x_1 + x_3 + x_5 &= 1. \end{aligned}$$

Aufgabe A7.45

Es seien $U = \mathrm{Lin}\big((1,1,0)^t, (1,0,2)^t\big)$ und $U' = \mathrm{Lin}\big((2,1,2)^t, (1,1,1)^t\big)$ zwei Unterräume des $\mathbb{R}^3$ und ferner sei

$$A = \begin{pmatrix} 2 & -2 & -1 \\ 3 & -2 & -2 \\ 5 & -4 & -3 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R}).$$

a. Bestimme Gleichungen für die Unterräume U und U'.

b. Berechne eine Basis von $U \cap U'$.

c. Zeige, daß $\mathrm{Ker}(f_A) = U \cap U'$ gilt.

Aufgabe A7.46

Bestimme die Lösungsmenge des folgenden linearen Gleichungssystems über $\mathbb{R}$:

$$\begin{aligned} x_1 + 2x_2 + x_4 + x_5 &= 0, \\ 2x_1 + 4x_2 + x_3 + x_4 + 2x_5 &= 1, \\ x_1 + 2x_2 + x_3 + x_5 &= 1, \\ 5x_1 + 10x_2 + 2x_3 + 3x_4 + 5x_5 &= 2. \end{aligned}$$

A8 Die symmetrische Gruppe

Im vorliegenden Abschnitt stellen wir die für uns relevanten Ergebnisse zur symmetrischen Gruppe $\mathbb{S}_n$ zusammen. Wir verzichten auf Beweise und erläutern sie stattdessen an Beispielen. Die Beweise der Aussagen finden sich in einer etwas ausführlicheren Version des Abschnitts im Lehrskript zu den Algebraischen Strukturen.

Definition A8.1 (Permutationen)
Eine bijektive Abbildung $\sigma : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}$ nennen wir eine *Permutation* der Menge $\{1, \ldots, n\}$, und wir bezeichnen mit

$$\mathbb{S}_n = \mathrm{Sym}(\{1, \ldots, n\}) = \big\{\sigma : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\} \;\big|\; \sigma \text{ ist bijektiv}\big\}$$

die Menge aller Permutation der Menge $\{1, \ldots, n\}$.
Eine Permutation $\sigma \in \mathbb{S}_n$ kann durch eine *Wertetabelle* der folgenden Form beschrieben werden:

$$\begin{pmatrix} 1 & 2 & \ldots & n \\ \sigma(1) & \sigma(2) & \ldots & \sigma(n) \end{pmatrix}$$

bzw.

$$\begin{pmatrix} a_1 & a_2 & \ldots & a_n \\ \sigma(a_1) & \sigma(a_2) & \ldots & \sigma(a_n) \end{pmatrix},$$

falls $a_1, \ldots, a_n$ irgendeine Anordnung der Zahlen $1, \ldots, n$ ist.

Bemerkung A8.2 (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 A8.3 ($\mathbb{S}_n$ nicht abelsch für $n \geq 3$)
Die Gruppe $\mathbb{S}_n$ ist für $n \geq 3$ nicht abelsch. In $\mathbb{S}_3$ gilt für die Permutationen

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \pi = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix} \in \mathbb{S}_3$$

nämlich

$$(\sigma \circ \pi)(1) = \sigma(\pi(1)) = \sigma(2) = 1 \neq 3 = \pi(2) = \pi(\sigma(1)) = (\pi \circ \sigma)(1)$$

und deshalb

$$\pi \circ \sigma \neq \sigma \circ \pi.$$

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 A8.4 (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 A8.5 (Zyklen und Transpositionen)

a. Sei $\{1, \dots, n\} = \{a_1, \dots, a_k\} \dot\cup \{b_1, \dots, b_{n-k}\}$, $k \geq 2$, 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:

(201) $$\sigma = (a_1 \dots a_k).$$

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 $\mathrm{id}_{\{1,\dots,n\}}$, wollen wir der Einfachheit halber mit id bezeichnen.

Bemerkung A8.6 (Zykelschreibweise)

a. Die Interpretation der Schreibweise in Gleichung (201) 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 Zyklenschreibweise 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 Zyklenschreibweise 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 $\tau^2 = \mathrm{id}$.

d. Für kleine Werte n ist $\mathbb{S}_n$ sehr übersichtlich, für große Werte n wird $\mathbb{S}_n$ jedoch riesig. $\mathbb{S}_1 = \{\mathrm{id}\}$ und $\mathbb{S}_2 = \{\mathrm{id}, (1\ 2)\}$. $\mathbb{S}_3 = \{\mathrm{id}, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$ hat schon sechs Elemente, $\mathbb{S}_4$ gar 24 und $\mathbb{S}_{64}$ ungefähr 10^{89}. Letztere Zahl entspricht in etwa der angenommenen Anzahl der Atome im Universum.

Satz A8.7 (Zyklenzerlegung und Signum)

a. Jede Permutation läßt sich als Produkt von disjunkten Zyklen schreiben.

b. Jede Permutation läßt sich als Produkt von Transpositionen schreiben.

c. Jede Permutation läßt sich als Produkt von Transpositionen benachbarter Zahlen schreiben.

d. Es gibt genau einen Gruppenhomomorphismus, das *Signum* genannt,

$$\mathrm{sgn} : (\mathbb{S}_n, \circ) \longrightarrow \big(\{1, -1\}, \cdot\big)$$

mit $\mathrm{sgn}(\tau) = -1$ für jede Transposition $\tau \in \mathbb{S}_n$. Insbesondere gilt, ist $\sigma \in \mathbb{S}_n$ ein Produkt von k Transpositionen, dann gilt mithin

$$\mathrm{sgn}(\sigma) = (-1)^k.$$

e. Für $\sigma \in \mathbb{S}_n$ gilt $\mathrm{sgn}(\sigma) = \mathrm{sgn}\big(\sigma^{-1}\big)$.

f. Ist $\mathbb{A}_n = \{\sigma \in \mathbb{S}_n \mid \mathrm{sgn}(\sigma) = 1\}$ und ist $\tau = (i\ j)$ eine Transposition, so gilt

$$\mathbb{S}_n = \mathbb{A}_n \dot\cup \tau\mathbb{A}_n$$

wobei $\tau\mathbb{A}_n = \{\tau \circ \sigma \mid \sigma \in \mathbb{A}_n\}$.

Beweis: Für den formalen Beweis der Aussagen verweisen wir auf das Lehrskript zu den Algebraischen Strukturen. □

Bemerkung A8.8 (sgn als Gruppenhomomorphismus)

a. Daß die Abbildung sgn ein *Gruppenhomomorphismus* ist, heißt

$$\operatorname{sgn}(\sigma \cdot \pi) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\pi)$$

für alle $\sigma, \pi \in \mathbb{S}_n$. Ist also $\sigma = \tau_1 \circ \cdots \circ \tau_k \in \mathbb{S}_n$ ein Produkt von k Transpositionen, dann gilt

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau_1) \cdot \ldots \cdot \operatorname{sgn}(\tau_k) = (-1)^k.$$

b. Einen Beweis der Aussagen findet der interessierte Leser im Lehrskript zu den Algebraichen Strukturen. 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 A8.9 (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

$$\sigma = (1\ 2\ 5) \circ (3\ 4) = (3\ 4) \circ (1\ 2\ 5). \tag{202}$$

Eine berechtigte Frage ist, wie wir die Zyklenzerlegung in (202) 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 $\{1, \ldots, 5\}$, 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. □

Beispiel A8.10 (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 A8.9 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:

$$(3\ 7) = (3\ 4)\circ(4\ 5)\circ(5\ 6)\circ(6\ 7)\circ(5\ 6)\circ(4\ 5)\circ(3\ 4).$$

Beispiel A8.11 (Symmetrische Gruppe $\mathbb{S}_3$)

Wir können alle $6 = 3!$ Elemente der $\mathbb{S}_3$ auflisten:

$$\mathbb{S}_3 = \left\{ \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} \right\}$$
$$= \{\mathrm{id} = (1), (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}\,.$$

Wir können dann auch für alle Elemente das Signum ausrechnen:

σ	id	(1 2)	(1 3)	(2 3)	(1 2 3)	(1 3 2)
$\mathrm{sgn}(\sigma)$	1	-1	-1	-1	1	1

Damit gilt dann

$$\mathbb{A}_3 = \{\mathrm{id}, (1\ 2\ 3), (1\ 3\ 2)\}$$

und

$$(1\ 2)\mathbb{A}_3 = \{(1\ 2) \circ \mathrm{id}, (1\ 2) \circ (1\ 2\ 3), (1\ 2) \circ (1\ 3\ 2)\} = \{(1\ 2), (2\ 3), (1\ 3)\},$$

woraus sich dann unmittelbar die folgende Gleichung ergibt:

$$\mathbb{S}_3 = \mathbb{A}_3 \dot\cup (1\ 2)\mathbb{A}_3.$$

Aufgaben

Aufgabe A8.12

Betrachte die Permutationen

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 2 & 6 & 3 & 7 & 5 & 1 & 4 \end{pmatrix}, \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 2 & 7 & 1 & 4 & 3 & 6 \end{pmatrix} \in \mathbb{S}_7.$$

a. Berechne $\sigma \circ \pi$, $\pi \circ \sigma$, σ^{-1}, π^{-1}.

b. Bestimme für jede der Permutationen in a. die Zyklenzerlegung.

c. Schreibe $\sigma \circ \pi$ als ein Produkt von Transpositionen.

d. Schreibe π^{-1} als ein Produkt von Transpositionen aufeinander folgender Zahlen.

e. Berechne für jede der Permutationen in a. das Signum.

Aufgabe A8.13

Bestimme die Elemente der Untergruppe $\mathbb{D}_8 = \langle (1\ 2\ 3\ 4), (1\ 4) \circ (2\ 3) \rangle$ von $\mathbb{S}_4$.

Aufgabe A8.14

Bestimme die Elemente der Untergruppe $\mathbb{D}_{10} = \langle (1\ 2\ 3\ 4\ 5), (2\ 4) \rangle$ von $\mathbb{S}_5$.

Aufgabe A8.15

Betrachte die Permutationen

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 5 & 4 & 3 & 1 & 7 & 8 & 6 \end{pmatrix}, \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 4 & 8 & 3 & 1 & 7 & 5 & 2 & 6 \end{pmatrix} \in \mathbb{S}_8.$$

a. Berechne $\sigma \circ \pi$, $\pi \circ \sigma$, σ^{-1}, π^{-1}.

b. Bestimme für jede der Permutationen in a. die Zyklenzerlegung.

c. Schreibe $\sigma \circ \pi$ als ein Produkt von Transpositionen.

d. Schreibe π^{-1} als ein Produkt von Transpositionen aufeinander folgender Zahlen.

e. Berechne für jede der Permutationen in a. das Signum.

A9 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. Eine weitere wichtige Bedeutung in der Analysis hat die Determinante für die Volumenberechnung (siehe auch Bemerkung A14.31).

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 Leibniz-Formel für die Determinante

Definition A9.1 (Determinante)
Wir definieren für $A \in \mathrm{Mat}_n(K)$ die *Determinante* von A durch die *Leibniz-Formel*

$$\det(A) := |A| := \sum_{\sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) \cdot a_{1\sigma(1)} \cdot \ldots \cdot a_{n\sigma(n)}. \tag{203}$$

Beispiel A9.2 (Determinanten für $n = 1, 2, 3$)

a. Ist $n = 1$, dann ist $A = (a) \in \mathrm{Mat}(1, K)$ und $\det(A) = a$.

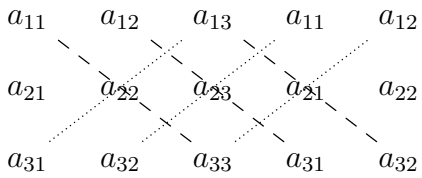
b. Ist $n = 2$, dann ist $\mathbb{S}_2 = \{\mathrm{id}, (1\ 2)\}$ und damit folgt:

$$\det(A) = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21},$$

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 $\mathbb{S}_n$ bereits sechs Elemente. Man berechnet in diesem Fall die Determinante mit der *Regel von Sarrus*:



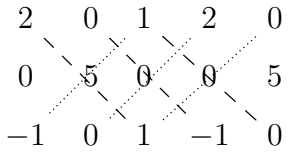
Die Produkte der Elemente längs der gestrichelten Linien tauchen bei der Berechnung der Determinante als positive Summanden auf, die Produkte der Elemente längs der gepunkteten Linien als negative Summanden. D. h., wir erhalten:

$$\begin{aligned} \det(A) &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &\quad -a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}. \end{aligned}$$

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



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.

Haben Matrizen eine spezielle Form, so ist es unter Umständen leichter, ihre Determinante zu berechnen. Sehr einfach ist dies für obere und untere Dreiecksmatrizen.

Proposition A9.3 (Determinanten von Dreiecksmatrizen)
Ist $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ 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 $\mathrm{id} \neq \sigma \in \mathbb{S}_n$, 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 (203) also nur der für id übrig. □

Beispiel A9.4

$$\det \begin{pmatrix} 2 & 3 & 12 & -3 \\ 0 & -1 & 5 & -111 \\ 0 & 0 & 3 & 5 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 2 \cdot (-1) \cdot 3 \cdot 1 = -6.$$

Lemma A9.5 (Alternative Leibniz-Formel)
Für die Determinante von $A \in \mathrm{Mat}_n(K)$ gilt

$$\det(A) = \sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) \cdot a_{\sigma(1)1} \cdot \ldots \cdot a_{\sigma(n)n}. \tag{204}$$

Beweis: Man beachte, daß für $\sigma \in \mathbb{S}_n$ auch σ^{-1} eine Permutation der Zahlen $1, \ldots, n$ ist, d. h. $\{1, \ldots, n\} = \{\sigma^{-1}(1), \ldots, \sigma^{-1}(n)\}$. Zudem wissen wir aus Satz A8.7 e., daß $\mathrm{sgn}(\sigma) = \mathrm{sgn}\left(\sigma^{-1}\right)$, 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:

$$\begin{aligned}
\det(A) &= \textstyle\sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) \cdot a_{\sigma^{-1}(1)\sigma(\sigma^{-1}(1))} \cdot \ldots \cdot a_{\sigma^{-1}(n)\sigma(\sigma^{-1}(n))} \\
&= \textstyle\sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}(\sigma) \cdot a_{\sigma^{-1}(1)1} \cdot \ldots \cdot a_{\sigma^{-1}(n)n} \\
&\overset{A8.7e.}{=} \textstyle\sum_{\sigma \in \mathbb{S}_n} \mathrm{sgn}\left(\sigma^{-1}\right) \cdot a_{\sigma^{-1}(1)1} \cdot \ldots \cdot a_{\sigma^{-1}(n)n} \\
&= \textstyle\sum_{\sigma^{-1} \in \mathbb{S}_n} \mathrm{sgn}\left(\sigma^{-1}\right) \cdot a_{\sigma^{-1}(1)1} \cdot \ldots \cdot a_{\sigma^{-1}(n)n} \\
&= \textstyle\sum_{\pi \in \mathbb{S}_n} \mathrm{sgn}(\pi) \cdot a_{\pi(1)1} \cdot \ldots \cdot a_{\pi(n)n}.
\end{aligned}$$

□

Proposition A9.6 (Die Determinante der Transponierten)
Für $A \in \mathrm{Mat}_n(K)$ gilt:

$$\det(A) = \det\left(A^t\right).$$

Beweis: Sei $A = (a_{ij})$ und $A^t = (a'_{ij})$, dann gilt $a'_{ij} = a_{ji}$. Mithin erhalten wir mit Hilfe von Lemma A9.5

$$\begin{aligned} \det(A) &= \sum_{\sigma\in\mathbb{S}_n} \mathrm{sgn}(\sigma)\cdot a_{1\sigma(1)}\cdot\ldots\cdot a_{n\sigma(n)} \\ &= \sum_{\sigma\in\mathbb{S}_n} \mathrm{sgn}(\sigma)\cdot a'_{\sigma(1)1}\cdot\ldots\cdot a'_{\sigma(n)n} \quad = \quad \det\left(A^t\right). \end{aligned}$$

□

Beispiel A9.7
Beispiel A9.2 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 A9.8 (Multilineare Abbildungen)
Es seien V und W zwei K-Vektorräume.

a. Eine Abbildung

$$f : V^n = V\times \overset{n}{\ldots} \times V \to W$$

heißt *multilinear*, falls f in jedem Argument linear ist, d. h. es gelten

$$f(x_1,\ldots,x_i+y_i,\ldots,x_n) = f(x_1,\ldots,x_i,\ldots,x_n) + f(x_1,\ldots,y_i,\ldots,x_n)$$

und

$$f(x_1,\ldots,\lambda x_i,\ldots,x_n) = \lambda\cdot f(x_1,\ldots,x_i,\ldots,x_n)$$

für jedes $i \in \{1,\ldots,n\}$ und für alle $x_1,\ldots,x_n,y_i \in V$ und $\lambda \in K$.

b. Eine multilineare Abbildung $f : V^n \to W$ heißt *alternierend*, falls für $(x_1,\ldots,x_n) \in V^n$ mit $x_i = x_j$ für ein $i \neq j$, gilt:

$$f(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_n) = 0.$$

Lemma A9.9
Ist $f : V^n \to W$ eine alternierende multilineare Abbildung, dann gilt für $\sigma \in \mathbb{S}_n$

$$f(x_{\sigma(1)},\ldots,x_{\sigma(n)}) = \mathrm{sgn}(\sigma)\cdot f(x_1,\ldots,x_n).$$

Insbesondere gilt $f(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_n) = -f(x_1,\ldots,x_j,\ldots,x_i,\ldots,x_n)$.

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

$$\begin{aligned} 0 &= f(x_1, \ldots, x_i + x_j, \ldots, x_i + x_j, \ldots, x_n) \\ &= f(x_1, \ldots, x_i, \ldots, x_i, \ldots, x_n) + f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) \\ &\quad + f(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n) + f(x_1, \ldots, x_j, \ldots, x_j, \ldots, x_n) \\ &= f(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n) + f(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n). \end{aligned}$$

Ist $\sigma \in \mathbb{S}_n$ beliebig, so können wir $\sigma = \tau_1 \circ \ldots \circ \tau_k$ 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 $\pi = \tau_2 \circ \ldots \tau_k$, so folgt mit der Vorüberlegung

$$\begin{aligned} f(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) &= f(x_{\tau_1(\pi(1))}, \ldots, x_{\tau_1(\pi(n))}) = -f(x_{\pi(1)}, \ldots, x_{\pi(n)}) \\ &\overset{Ind.}{=} -\operatorname{sgn}(\pi) \cdot f(x_1, \ldots, x_n) \overset{Satz\ A8.7}{=} \operatorname{sgn}(\tau_1 \circ \pi) \cdot f(x_1, \ldots, x_n) \\ &= \operatorname{sgn}(\sigma) \cdot f(x_1, \ldots, x_n). \end{aligned}$$

□

Bemerkung A9.10

Wir können $\mathrm{Mat}_n(K)$ auf recht natürliche Weise mit $K^n \times \overset{n}{\ldots} \times K^n$ identifizieren, indem wir eine Matrix $A = (a_{ij})$ mit dem n-Tupel ihrer Spaltenvektoren $(a^1, \ldots, a^n)$ gleichsetzen. Das wollen wir im folgenden tun.

Satz A9.11 (Die Determinante als Volumenform)

a. Die Determinante

$$\det : \mathrm{Mat}_n(K) \to K : A \mapsto \det(A)$$

ist eine alternierende multilineare Abbildung mit $\det(\mathbb{1}_n) = 1$.

b. Ist $f : \mathrm{Mat}_n(K) \longrightarrow K$ eine alternierende multilineare Abbildung und $A \in \mathrm{Mat}_n(K)$, so gilt

$$f(A) = f(\mathbb{1}_n) \cdot \det(A).$$

Beweis:

a. Wir werden im Beweis die Formel (204) aus Lemma A9.5 zur Berechnung der Determinante verwenden, da sie auf die Bedürfnisse der Determinante als multilineare Abbildung bezüglich der Spalten zugeschnitten ist.
Es seien $a^j = (a_{1j}, \ldots, a_{nj})^t$, $j = 1, \ldots, n$, und $b^i = (b_{1i}, \ldots, b_{ni})^t$. Wir setzen $A := (a^1 \ldots a^i \ldots a^n)$, $B := (a^1 \ldots b^i \ldots a^n)$ und $C := (a^1 \ldots \lambda a^i + \mu b^i \ldots a^n)$. Dann

gilt

$$\begin{aligned}\det(C) &= \sum_{\sigma\in\mathbb{S}_n} \operatorname{sgn}(\sigma)\cdot a_{\sigma(1)1}\cdot\ldots\cdot(\lambda a_{\sigma(i)i}+\mu b_{\sigma(i)i})\cdot\ldots\cdot a_{\sigma(n)n}\\ &= \lambda\cdot\sum_{\sigma\in\mathbb{S}_n}\operatorname{sgn}(\sigma)\cdot a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(i)i}\cdot\ldots\cdot a_{\sigma(n)n}\\ &\quad +\mu\cdot\sum_{\sigma\in\mathbb{S}_n}\operatorname{sgn}(\sigma)\cdot a_{\sigma(1)1}\cdot\ldots\cdot b_{\sigma(i)i}\cdot\ldots\cdot a_{\sigma(n)n}\\ &= \lambda\cdot\det(A)+\mu\cdot\det(B),\end{aligned}$$

so daß det multilinear ist.
Sei nun $a^i = a^j$, für ein $i \neq j$. Ist $\tau = (i\ j)$, die Transposition, die i und j vertauscht, dann besitzt $\mathbb{S}_n$ nach Satz A8.7 die Zerlegung $\mathbb{S}_n = \mathbb{A}_n \cup \mathbb{A}_n\tau$. Ferner gilt für $\sigma \in \mathbb{A}_n$

$$\operatorname{sgn}(\sigma) = 1 \quad \text{und} \quad \operatorname{sgn}(\sigma\tau) = -1.$$

Wir erhalten also

$$\begin{aligned}\det(A) &= \sum_{\sigma\in\mathbb{A}_n} a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(i)i}\cdot\ldots\cdot a_{\sigma(j)j}\cdot\ldots\cdot a_{\sigma(n)n}\\ &\quad -\sum_{\sigma\in\mathbb{A}_n} a_{\sigma\tau(1)1}\cdot\ldots\cdot a_{\sigma\tau(i)i}\cdot\ldots\cdot a_{\sigma\tau(j)j}\cdot\ldots\cdot a_{\sigma\tau(n)n}\\ &= \sum_{\sigma\in\mathbb{A}_n} a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(i)i}\cdot\ldots\cdot a_{\sigma(j)j}\cdot\ldots\cdot a_{\sigma(n)n}\\ &\quad -\sum_{\sigma\in\mathbb{A}_n} a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(j)i}\cdot\ldots\cdot a_{\sigma(i)j}\cdot\ldots\cdot a_{\sigma(n)n} = 0,\end{aligned}$$

und somit ist det alternierend.
Außerdem folgt $\det(\mathbb{1}_n) = 1\cdot\ldots\cdot 1 = 1$ aus Proposition A9.3.

b. Mit den Notationen von a. gilt $a^i = \sum_{j=1}^n a_{ji}e^j$, wenn e^j der j-te Einheitsvektor ist. Aus der Multilinearität von f folgt:

$$f(A) = \sum_{j_1=1}^n a_{j_1 1} f\left(e^{j_1}\ a^2\ldots a^n\right) = \sum_{j_1=1}^n a_{j_1 1}\sum_{j_2=1}^n a_{j_2 2} f\left(e^{j_1}\ e^{j_2}\ a^3\ldots a^n\right)$$

$$= \ldots = \sum_{j_1,\ldots,j_n=1}^n a_{j_1 1}\cdot\ldots\cdot a_{j_n n} f\left(e^{j_1}\ \ldots\ e^{j_n}\right).$$

Genau dann, wenn die $j_1,\ldots,j_n$ paarweise verschieden sind, existiert eine Permutation $\sigma \in \mathbb{S}_n$ mit $\left(e^{j_1}\ldots e^{j_n}\right) = \left(e^{\sigma(1)}\ldots e^{\sigma(n)}\right)$, und wegen Lemma A9.9 gilt dann

$$f(e^{\sigma(1)}\ldots e^{\sigma(n)}) = \operatorname{sgn}(\sigma)\cdot f(e^1\ldots e^n) = \operatorname{sgn}(\sigma)\cdot f(\mathbb{1}_n).$$

Andernfalls stimmen zwei der j_i überein und $f\left(e^{j_1}\ \ldots\ e^{j_n}\right) = 0$, da f alternierend ist. Insgesamt haben wir damit gezeigt:

$$f(A) = \sum_{\sigma\in\mathbb{S}_n} a_{\sigma(1)1}\cdot\ldots\cdot a_{\sigma(n)n}\cdot\operatorname{sgn}(\sigma)\cdot f(\mathbb{1}_n) = \det(A)\cdot f(\mathbb{1}_n).$$

□

Bemerkung A9.12 (Das Volumen des Parallelotops)
Eine alternierende multilineare Abbildung $f : \mathrm{Mat}_n(K) \longrightarrow K$ wird auch eine *Volumenform* genannt. Aus Satz A9.11 b. folgt, daß die Determinante die einzige Volumenform f mit $f(\mathbb{1}_n) = 1$ ist, d.h. det ist durch die Eigenschaften in Satz A9.11 a. eindeutig bestimmt.

Die Determinante hat eine wichtige geometrische Interpretation, die den Begriff *Volumenform* rechtfertigt. Seien $x_1, \ldots, x_n \in \mathbb{R}^n$ und sei

$$P\big(x_1, \ldots, x_n\big) := \big\{\lambda_1 x_1 + \ldots + \lambda_n x_n \in \mathbb{R}^n \mid 0 \leq \lambda_i \leq 1, i = 1, \ldots, n\big\}$$

das von den Vektoren $x_1, \ldots, x_n$ aufgespannte *Parallelotop* (siehe Abbildung 8).

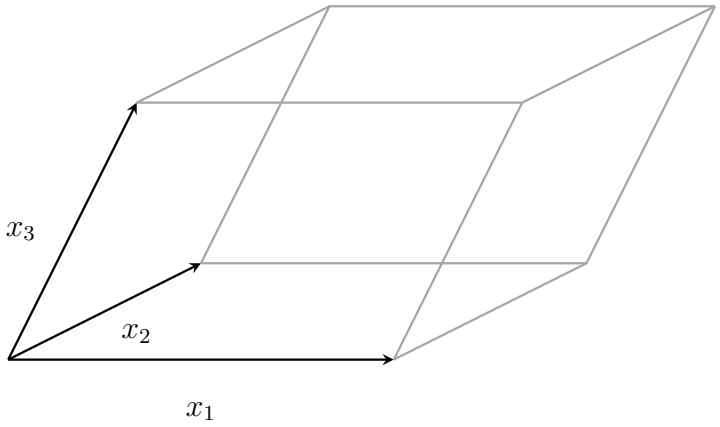


Abbildung 8: Das Parallelotop $P(x_1, x_2, x_3)$ im $\mathbb{R}^3$

Dann definiert man das n-dimensionale Volumen von $P(x_1, \ldots, x_n)$ mit Hilfe der Determinante als

$$\text{Volumen}\,\big(P(x_1, \ldots, x_n)\big) = |\det\big(x_1 \ldots x_n\big)|.$$

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 A14.31 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. □

C) Der Gauß-Algorithmus zur Berechnung der Determinante

Wir haben oben gesehen, daß die Leibniz-Formel ungeeignet ist zur Berechnung der Determinante einer $n \times n$-Matrix für großes n, weil die Symmetrische Gruppe dann viel zu viele Elmenete enthält. Wir wollen deshalb nun zeigen, wie die Determinante effizient mit Hilfe des Gauß-Algorithmus' berechnet werden kann.

Korollar A9.13 (Spaltenoperationen und die Determinante)
Es sei $A \in \mathrm{Mat}_n(K)$ und $\lambda \in K$.

a. Bei Vertauschung zweier Spalten von A ändert sich das Vorzeichen von $\det(A)$.

b. Bei Multiplikation einer Spalte von A mit λ multipliziert sich $\det(A)$ mit λ.

c. Bei Addition des λ-fachen einer Spalte zu einer anderen Spalte ändert sich $\det(A)$ nicht.

d. Enthält A eine Nullspalte, so ist $\det(A) = 0$.

e. Sind zwei Spalten von A gleich, so ist $\det(A) = 0$.

Beweis:

a. Das ist ein Spezialfall von Lemma A9.9, da det nach Satz A9.11 alternierend ist.

b. Dies folgt aus der Multilinearität von det, siehe Satz A9.11.

c. Für $A = \begin{pmatrix} a^1 \dots a^n \end{pmatrix}$ und $A' = \begin{pmatrix} a^1 \dots a^j + \lambda a^i \dots a^n \end{pmatrix}$ folgt aus der Multilinearität und da det alternierend ist:

$$\det(A') = \det(A) + \lambda \cdot \det(a^1 \dots a^i \dots a^i \dots a^n) = \det(A) + \lambda \cdot 0 = \det(A).$$

d. Ist eine Spalte von A Null, so folgt $\det(A) = 0$ aus b. mit $\lambda = 0$.

e. Das folgt, da det alternierend ist.

□

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 A9.14 (Zeilenoperationen und die Determinante)
Wir können $\det : \mathrm{Mat}_n(K) \to K$ auch als multilineare Abbildung auf den Zeilen einer Matrix A auffassen. Entsprechend gilt Korollar A9.13 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. Dabei müssen wir aber das Tauschen von Zeilen kompensieren und wir müssen auf das Mulitplizieren von Zeilen mit Skalaren verzichten.

Algorithmus A9.15 (Algorithmus zur Berechnung der Determinante)
INPUT: $A \in \mathrm{Mat}_n(K)$.
OUTPUT: $\det(A)$.

1. **Schritt:** Überführe A mittels Gauß-Algorithmus in nicht-reduzierte ZSF, d. h. führe im Gauß-Algorithmus A2.24 Schritt sieben nicht aus. Jedesmal, wenn dabei zwei Zeilen vertauscht werden, ersetze eine davon durch ihr Negatives. - Wird bei der Gaußreduktion ein Pivotelement zu Null, gib Null zurück und brich ab.
2. **Schritt:** Gib das Produkt der Diagonalelemente der ZSF zurück.

Beispiel A9.16
Wir wollen die Determinante der Matrix

$$A = \begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 0 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$$

berechnen. Dazu überführen wir sie mittels des Gauß-Algorithmus in ZSF und merken uns die Zeilenvertauschungen.

$$\begin{pmatrix} 4 & 5 & 6 \\ 1 & 2 & 3 \\ 7 & 8 & 0 \end{pmatrix} \xmapsto[d=-1]{I \leftrightarrow II} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{pmatrix} \xmapsto[III \to III-7I]{II \to II-4I} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -21 \end{pmatrix}$$

$$\xmapsto{III \to III-2II} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & -9 \end{pmatrix}$$

Damit gilt dann

$$\det(A) = d \cdot 1 \cdot (-3) \cdot (-9) = -27.$$

Beispiel A9.17
Sei $A \in \mathrm{Mat}_{n+1}(K)$ 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 $i = 1, \ldots, n$ von der i-ten Zeile die $(i+1)$-te Zeile ab. Wir erhalten:

$$A' = \begin{pmatrix} -1 & 1 & 1 & 1 & \ldots & 1 & 1 \\ -1 & -1 & 1 & 1 & \ldots & 1 & 1 \\ -1 & -1 & -1 & 1 & \ldots & 1 & 1 \\ -1 & -1 & -1 & -1 & \ldots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -1 & -1 & -1 & -1 & \ldots & -1 & 1 \\ n & n-1 & n-2 & n-3 & \ldots & 1 & 0 \end{pmatrix}.$$

Addiere nun für $i = 2, \ldots, n+1$ die erste Spalte zur i-ten Spalte. Dann erhalten wir:

$$A'' = \begin{pmatrix} -1 & 0 & 0 & 0 & \ldots & 0 & 0 \\ * & -2 & 0 & 0 & \ldots & 0 & 0 \\ * & * & -2 & 0 & \ldots & 0 & 0 \\ * & * & * & -2 & \ldots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ & * & * & * & \ldots & -2 & 0 \\ * & * & * & * & \ldots & * & n \end{pmatrix}.$$

Es folgt:

$$\det(A) = \det(A'') = (-1) \cdot (-2)^{n-1} \cdot n = -n \cdot (-2)^{n-1}.$$

Bemerkung A9.18

In Beispiel A9.17 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

Satz A9.19 (Determinantenmultiplikationssatz)

Für Matrizen $A, B \in \mathrm{Mat}_n(K)$ gilt

$$\det(A \circ B) = \det(A) \cdot \det(B).$$

Beweis: Wähle $A \in \mathrm{Mat}_n(K)$ fest und betrachte die Abbildung

$$f : \mathrm{Mat}_n(K) \to K : B \mapsto \det(A \circ B).$$

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 A9.11:

$$\det(A \circ B) = f(B) = f(\mathbb{1}_n) \cdot \det(B) = \det(A) \cdot \det(B).$$

□

Beispiel A9.20

In Beispiel A9.2 b. gilt

$$A = \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix} \circ \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}),$$

so folgt aus dem Determinantenmultiplikationssatz A9.19 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 A2.25.

Korollar A9.21 (Determinante und Invertierbarkeit)
Genau dann ist $A \in \mathrm{Mat}_n(K)$ invertierbar, wenn $\det(A) \neq 0$. In diesem Fall gilt

$$\det\left(A^{-1}\right) = \frac{1}{\det(A)}.$$

Beweis: Ist A invertierbar, so gilt

$$1 = \det(\mathbb{1}_n) = \det\left(A \circ A^{-1}\right) = \det(A) \cdot \det\left(A^{-1}\right).$$

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 A9.13 c. ändert sich dabei der Wert der Determinante nicht, und nach Korollar A9.13 d. muß er somit 0 sein. □

Beispiel A9.22

Die Matrix A in Beispiel A9.20 ist invertierbar und ihre Inverse hat Determinante $\frac{1}{3}$.

Dies wissen wir, ohne die Inverse auszurechnen. Diese können wir mit Hilfe von Aufgabe A2.25 berechnen. Für eine invertierbare 2×2-Matrix gilt

$$B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_n(K)$$

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 A9.23 (det ist ein Gruppenepimorphismus.)
In der Sprache der Vorlesung Algebraische Strukturen folgt aus Satz A9.19 und Korollar A9.21, daß

$$\det : \big(\mathrm{Gl}_n(K), \circ\big) \to \big(K^*, \cdot\big)$$

ein Gruppenepimorphismus ist. Dazu beachte man, daß det surjektiv ist wegen

$$\det \left(\begin{array}{c|ccc} \lambda & 0 & \dots & 0 \\ \hline 0 & & & \\ \vdots & & \mathbb{1}_{n-1} & \\ 0 & & & \end{array}\right) = \lambda.$$

E) Der Kästchensatz

Satz A9.24 (Kästchensatz)
Es sei $A \in \mathrm{Mat}_n(K)$ eine *Blockmatrix* der Form

$$A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right)$$

mit $B \in \mathrm{Mat}_k(K)$, $C \in \mathrm{Mat}(k \times l, K)$, $D \in \mathrm{Mat}_l(K)$, $0 \in \mathrm{Mat}(l \times k, K)$ und $n = k + l$. Dann gilt:

$$\det(A) = \det(B) \cdot \det(D).$$

Beweis: Man beachte, daß

$$A = \left(\begin{array}{c|c} B & C \\ \hline 0 & D \end{array}\right) = \left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D \end{array}\right) \circ \left(\begin{array}{c|c} B & C \\ \hline 0 & \mathbb{1}_l \end{array}\right).$$

Wegen des Determinantenmultiplikationssatzes A9.19 reicht es mithin zu zeigen:

$$\det\left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D \end{array}\right) = \det(D) \tag{205}$$

und

$$\det\left(\begin{array}{c|c} B & C \\ \hline 0 & \mathbb{1}_l \end{array}\right) = \det(B). \tag{206}$$

Die Abbildung

$$f : \mathrm{Mat}_l(K) \to K : D' \mapsto \det\left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D' \end{array}\right)$$

ist offensichtlich multilinear und alternierend, und wegen Satz A9.11 b. gilt mithin

$$\det\left(\begin{array}{c|c} \mathbb{1}_k & 0 \\ \hline 0 & D \end{array}\right) = f(D) = f(\mathbb{1}_l) \cdot \det(D) = \det(\mathbb{1}_n) \cdot \det(D) = \det(D),$$

d. h. (205) ist erfüllt.

Analog ist die Abbildung

$$g : \mathrm{Mat}_k(K) \to K : B' \mapsto \det\left(\begin{array}{c|c} B' & C \\ \hline 0 & \mathbb{1}_l \end{array}\right)$$

alternierend und multilinear in den Spalten von B', also eine Volumenform. Wieder folgt aus Satz A9.11 mit Hilfe von Proposition A9.3, daß

$$\det\left(\begin{array}{c|c} B & C \\ \hline 0 & \mathbb{1}_l \end{array}\right) = g(B) = g(\mathbb{1}_k) \cdot \det(B) = \det\left(\begin{array}{c|c} \mathbb{1}_k & C \\ \hline 0 & \mathbb{1}_l \end{array}\right) \cdot \det(B) = \det(B),$$

womit auch (206) gezeigt ist. □

Beispiel A9.25 (Vandermonde-Determinante)

Wir wollen mit Hilfe des Gauß-Algorithmus zeigen, daß

$$det\begin{pmatrix} 1 & a_0 & a_0^2 & a_0^3 & \dots & a_0^n \\ 1 & a_1 & a_1^2 & a_1^3 & \dots & a_1^n \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & a_n^2 & a_n^3 & \dots & a_n^n \end{pmatrix} = \prod_{0 \le i < j \le n} (a_j - a_i)$$

für $a_0, \dots, a_n \in K$ und $n \ge 1$ gilt. Die Determinate dieser Matrix ist als *Vandermonde-Determinante* bekannt.

Beweis: Wir beweisen die Aussage mit Hilfe von Induktion nach n. Für $n = 1$ ist

$$\det(A) = \begin{vmatrix} 1 & a_0 \\ 1 & a_1 \end{vmatrix} = a_1 - a_0$$

und die Aussage stimmt. Sei also $n > 1$ und die Aussage sei für Matrizen dieser Gestalt der Größe n (beachte, daß A die Größe $n + 1$ hat) bereits gezeigt. Addieren wir für

$j = n+1, \ldots, 2$ zur j-ten Spalte das $-a_0$-fache der $j-1$-ten Spalte, so ändert sich die Determinante nicht und wir erhalten

$$\det(A) = \begin{vmatrix} 1 & a_0 & a_0^2 & a_0^3 & \ldots & a_0^n \\ 1 & a_1 & a_1^2 & a_1^3 & \ldots & a_1^n \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & a_n^3 & \ldots & a_n^n \end{vmatrix}$$

$$= \left|\begin{array}{c|cccc} 1 & 0 & 0 & 0 & \ldots & 0 \\ \hline 1 & a_1 - a_0 & a_1^2 - a_1 a_0 & a_1^3 - a_1^2 a_0 & \ldots & a_1^n - a_1^{n-1} a_0 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n - a_0 & a_n^2 - a_n a_0 & a_n^3 - a_n^2 a_0 & \ldots & a_n^n - a_n^{n-1} a_0 \end{array}\right|$$

Aufgrund des Kästchensatzes und wegen $\det(1) = 1$ gilt dann

$$\det(A) = \begin{vmatrix} a_1 - a_0 & a_1^2 - a_1 a_0 & a_1^3 - a_1^2 a_0 & \ldots & a_1^n - a_1^{n-1} a_0 \\ \vdots & \vdots & \vdots & & \vdots \\ a_n - a_0 & a_n^2 - a_n a_0 & a_n^3 - a_n^2 a_0 & \ldots & a_n^n - a_n^{n-1} a_0 \end{vmatrix}$$

$$= \begin{vmatrix} a_1 - a_0 & (a_1 - a_0) \cdot a_1 & (a_1 - a_0) \cdot a_1^2 & \ldots & (a_1 - a_0) \cdot a_1^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ a_n - a_0 & (a_n - a_0) \cdot a_n & (a_n - a_0) \cdot a_n^2 & \ldots & (a_n - a_0) \cdot a_n^{n-1} \end{vmatrix}$$

Klammern wir nun in der i-ten Zeile $a_i - a_0$ aus, so erhalten wir

$$\det(A) = \prod_{i=1}^{n} (a_i - a_0) \cdot \begin{vmatrix} 1 & a_1 & a_1^2 & a_1^3 & \ldots & a_1^{n-1} \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 1 & a_n & a_n^2 & a_n^3 & \ldots & a_n^{n-1} \end{vmatrix}.$$

Auf die letzte Determinante können wir Induktion anwenden und erhalten

$$\det(A) = \prod_{i=1}^{n} (a_i - a_0) \cdot \prod_{1 \le i < j \le n} (a_j - a_i) = \prod_{0 \le i < j \le n} (a_j - a_i).$$

□

F) Der Satz über die Adjunkte

Unser nächstes Ziel ist die Herleitung einer alternativen Berechnung der Determinante, die im Gegensatz zum Gaußalgorithmus ohne Division auskommt und deshalb über jedem kommutativen Ring mit Eins funktioniert (siehe Bemerkung A9.37). Zu ihrer Herleitung führen wir zunächst verschiedene Hilfmatrizen ein.

Definition A9.26

Es sei $A = (a_{ij}) = (a^1 \dots a^n) \in \mathrm{Mat}_n(K)$, $n \geq 2$, 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 A9.27

Für $A \in \mathrm{Mat}_n(K)$, $n \geq 2$, $1 \leq i, j \leq n$, gilt:

$$\det\Big(A_i(e_j)\Big) = \det\big(S_{ji}(A)\big) = (-1)^{i+j} \det(A_{ji}).$$

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 A9.13:

$$\det\Big(A_i(e_j)\Big) = \det\big(S_{ji}(A)\big).$$

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 A9.24 unter Beachtung der Korollare A9.13 und A9.14

$$\det\big(S_{ji}(A)\big) = (-1)^{i+j}\det(A_{ji}).$$

□

In der folgenden Definition beachte man die Vertauschung der Indizes!

Definition A9.28

Für $A \in \mathrm{Mat}_n(K)$, $n \geq 2$, $1 \leq i, j \leq n$ heißt

$$a_{ij}^{\#} := (-1)^{i+j}\det(A_{ji})$$

ein *Kofaktor* von A. Die Matrix der Kofaktoren

$$A^{\#} := \big(a_{ij}^{\#}\big) \in \mathrm{Mat}_n(K)$$

heißt die *Adjunkte* oder *Komplementärmatrix* von A.

Satz A9.29 (Satz über die Adjunkte)

Für $A \in \mathrm{Mat}_n(K)$, $n \geq 2$, gilt:

$$A^{\#} \circ A = A \circ A^{\#} = \det(A) \cdot \mathbb{1}_n.$$

Beweis: Sei $A^{\#} \circ A = (c_{ik})$. Dann gilt mit Lemma A9.27:

$$c_{ik} = \sum_{j=1}^{n} a_{ij}^{\#} \cdot a_{jk} = \sum_{j=1}^{n} a_{jk} \cdot \det\big(a^1 \dots a^{i-1}\ e_j\ a^{i+1} \dots a^n\big)$$

$$= \det\left(a^1 \dots a^{i-1} \sum_{j=1}^{n} a_{jk} e_j\ a^{i+1} \dots a^n\right)$$

$$= \det\big(a^1 \dots a^{i-1}\ a^k\ a^{i+1} \dots a^n\big) = \delta_{ik} \cdot \det(A),$$

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 $A \circ A^{\#} = \det(A) \cdot \mathbb{1}_n$ geht analog. □

Korollar A9.30
Es sei $A \in \mathrm{Mat}_n(K)$ 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 A9.31
Für eine Matrix $A \in \mathrm{Mat}(n, K)$ sind gleichwertig:

a. A ist invertierbar.
b. $\mathrm{rang}(A) = n$.
c. $\det(A) \neq 0$.
d. f_A ist bijektiv.
e. f_A ist injektiv.
f. f_A ist surjektiv.
g. $\mathrm{rZSF}(A) = \mathbb{1}_n$.
h. A ist das Produkt endlich vieler Elementarmatrizen.
i. Es gibt eine Matrix $B \in \mathrm{Mat}(n, K)$ mit $A \circ B = \mathbb{1}_n$.

Beweis: Die unterschiedlichen Äquivalenzen sind in den Sätzen A4.42, A5.26, A5.27, A6.27, A6.34, A7.2 und A9.21 gezeigt worden. □

Beispiel A9.32
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 A2.25 bewiesen.

Sei nun konkret $K = \mathbb{Q}$ und

$$A = \begin{pmatrix} 1 & 1 \\ 3 & 4 \end{pmatrix}.$$

Dann ist $\det(A) = 1$ und somit gilt

$$A^{-1} = \begin{pmatrix} 4 & -1 \\ -3 & 1 \end{pmatrix}.$$

G) Laplacescher Entwicklungssatz und Cramersche Regel

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 (203) zur Definition der Determinante.

Satz A9.33 (Laplacescher Entwicklungssatz)
Es sei $A \in \mathrm{Mat}_n(K)$.

a. Wir nennen die folgende Formel, die *Entwicklung nach der i-ten Zeile*:

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij}). \tag{207}$$

b. Entsprechend nennen wir die folgende Formel, die *Entwicklung nach der j-ten Spalte*:

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij}). \tag{208}$$

Beweis: Nach Satz A9.29 gilt für $A \circ A^{\#} = (c_{ik})$

$$\det(A) = c_{ii} = \sum_{j=1}^{n} a_{ij} \cdot a_{ji}^{\#} = \sum_{j=1}^{n} (-1)^{i+j} \cdot a_{ij} \cdot \det(A_{ij}).$$

Damit folgt (207), und (208) zeigt man analog durch die Betrachtung von $A^{\#} \circ A$. □

Bemerkung A9.34

Entwickelt man $A = (a_{ij})$ 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

$$A = \begin{pmatrix} 0 & 2 & 0 \\ 1 & 3 & 4 \\ 2 & 5 & 3 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$$

mit Hilfe der Entwicklung nach der ersten Zeile. Dann gilt

$$\begin{aligned} \det(A) =& 0 \cdot \det(A_{11}) - 2 \cdot \det(A_{12}) + 0 \cdot \det(A_{13}) \\ =& -2 \cdot \det \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} = (-2) \cdot (-5) = 10. \end{aligned}$$

Wir leiten nun aus dem Laplaceschen Entwicklungssatz einen Algorithmus ab.

Algorithmus A9.35 (Laplace-Entwicklung)

INPUT: $A \in \mathrm{Mat}_n(K)$.

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 $i = 1, \ldots, n$:
   - 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 23.21).

Satz A9.36 (Cramersche Regel)

Es sei $A \in \mathrm{Mat}_n(K)$ invertierbar und $b \in K^n$.

Für die eindeutig bestimmte Lösung $x = (x_1, \ldots, x_n)^t \in K^n$ von $Ax = b$ gilt dann

$$\begin{aligned} x_i &= \tfrac{1}{\det(A)} \cdot \det\big(A_i(b)\big) \\ &= \tfrac{1}{\det(A)} \cdot \det \begin{pmatrix} a_{11} & \ldots & a_{1\,i-1} & b_1 & a_{1\,i+1} & \ldots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \ldots & a_{n\,i-1} & b_n & a_{n\,i+1} & \ldots & a_{nn} \end{pmatrix}. \end{aligned}$$

Beweis: Wegen Korollar A9.30 ist

$$x = A^{-1}b = \frac{1}{\det(A)} \cdot A^{\#}b$$

die eindeutig bestimmte Lösung des linearen Gleichungssystems. Also folgt mit Lemma A9.27 und der Multilinearität der Determinante

$$\begin{aligned} x_i &= \tfrac{1}{\det(A)} \cdot \sum_{j=1}^{n} a_{ij}^{\#} \cdot b_j = \tfrac{1}{\det(A)} \cdot \sum_{j=1}^{n} \det\big(A_i(e_j)\big) \cdot b_j \\ &= \tfrac{1}{\det(A)} \cdot \sum_{j=1}^{n} \det\big(a^1 \dots a^{i-1}\ e_j\ a^{i+1} \dots a^n\big) \cdot b_j \\ &= \tfrac{1}{\det(A)} \cdot \det\big(a^1 \dots a^{i-1}\ b\ a^{i+1} \dots a^n\big) \\ &= \tfrac{1}{\det(A)} \cdot \det\big(A_i(b)\big). \end{aligned}$$

□

H) Ringe und Moduln

Bemerkung A9.37 (Determinanten über kommutativen Ringen mit Eins)
Ist K nur ein kommutativer Ring mit Eins, so können wir die Determinante einer Matrix in $\mathrm{Mat}_n(K)$ 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 A9.15 über beliebigen Ringen in der angegebenen Form *nicht* mehr anwenden, da dabei Divisionen nötig sind. Außerdem gilt Korollar A9.21 *nicht* mehr in der angegebenen Form, und ebenso gilt Korollar A9.31 in *nicht* in vollem Umfang.

Alle anderen Aussagen gelten jedoch ohne jede Änderung. Dies trifft insbesondere auf den Satz zur Adjunkten A9.29 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 A9.30 beachte man, daß aus dem Determinantenmultiplikationssatz A9.19 und dem Satz zur Adjunkten A9.29 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 A9.30 ebenso zu wie auf die Cramersche Regel A9.36.

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 A9.32 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 $\mathrm{Mat}_2(\mathbb{Q})$ also invertierbar, als Matrix in $\mathrm{Mat}_2(\mathbb{Z})$ aber nicht.

Aufgaben

Aufgabe A9.38

Berechne die Determinanten der folgenden Matrizen:

a. $\begin{pmatrix} 1 & -4 \\ -3 & 8 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$.

b. $\begin{pmatrix} -1 & 0 & -1 \\ 0 & 5 & 0 \\ 3 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$.

c. $\begin{pmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$.

Aufgabe A9.39

Sei K ein Körper und $\lambda \in K$. Bestimme die Determinante der Matrix

$$\begin{pmatrix} 1 & \lambda & \lambda^2 & \dots & \lambda^{n-1} \\ \lambda^{n-1} & 1 & \lambda & \dots & \lambda^{n-2} \\ \lambda^{n-2} & \lambda^{n-1} & 1 & \dots & \lambda^{n-3} \\ \vdots & \vdots & \vdots & & \vdots \\ \lambda & \lambda^2 & \lambda^3 & \dots & 1 \end{pmatrix} \in \mathrm{Mat}(n \times n, K).$$

Aufgabe A9.40

Für $n \in \mathbb{N} \setminus \{0\}$ definieren wir

$$A_n = \begin{pmatrix} 1 & 1 & 0 & \dots & \dots & \dots & 0 \\ 1 & 1 & 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 1 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 1 & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 & 1 & 1 \\ 0 & \dots & \dots & \dots & 0 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_n(\mathbb{R})$$

als die Matrix, deren Einträge auf der Diagonalen sowie auf der oberen und unteren Nebendiagonalen alle eins sind, während alle anderen Einträge null sind. Ferner setzen wir $d_n = \det(A_n)$.

a. Zeige, für $n \geq 3$ gilt die Rekursionsformel $d_n = d_{n-1} - d_{n-2}$.

b. Zeige, für $k \in \mathbb{N}$ gilt

$$d_n = \begin{cases} 1, & \text{falls } n \equiv 1 (\mathrm{mod}\ 6) \text{ oder } n \equiv 0 (\mathrm{mod}\ 6), \\ 0, & \text{falls } n \equiv 2 (\mathrm{mod}\ 6) \text{ oder } n \equiv 5 (\mathrm{mod}\ 6), \\ -1, & \text{falls } n \equiv 3 (\mathrm{mod}\ 6) \text{ oder } n \equiv 4 (\mathrm{mod}\ 6). \end{cases}$$

Aufgabe A9.41

Berechne die Determinante folgender Matrix mit Hilfe des Gauß-Algorithmus':

$$A = \begin{pmatrix} 1 & 3 & 0 & 2 & 1 \\ -2 & 0 & 3 & 1 & 3 \\ 0 & 2 & 5 & 0 & -1 \\ 0 & 0 & 1 & 1 & 1 \\ 1 & 0 & 3 & 4 & 1 \end{pmatrix} \in \mathrm{Mat}_5(\mathbb{R}).$$

Aufgabe A9.42

Berechne die folgende Determinante mit Hilfe des Determinantenentwicklungssatzes:

$$B = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & 3 & 1 & 3 \\ 2 & 5 & 0 & -1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q}).$$

Aufgabe A9.43

Sei V ein n-dimensionaler $\mathbb{C}$-Vektorraum und $f : V \longrightarrow V$ $\mathbb{C}$-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 $(v_1, \ldots, v_n)$ von V betrachte man die zugehörige $\mathbb{R}$-Basis $(v_1, \ldots, v_n, iv_1, \ldots, iv_n)$ sowie jeweils die zugehörige Matrixdarstellung von f. Wem der allgemeine Fall zu schwer ist, der beschränke sich auf die Abbildung $f : \mathbb{C} \longrightarrow \mathbb{C} : z \longmapsto (a + ib) \cdot z$ mit $a, b \in \mathbb{R}$ fest vorgegeben. Was ist eine Basis von $\mathbb{C}$ als $\mathbb{R}$-Vektorraum?

Aufgabe A9.44

Bestimme für welche $s \in \mathbb{R}$ die Abbildung

$$f_s : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (x, y, z)^t \mapsto (x + z, x + 2y + z, sx + y - z)^t$$

invertierbar ist und berechne für diese die Inverse mit Hilfe der Adjunkten.

Aufgabe A9.45

Löse das folgende lineare Gleichungssystem mit Hilfe der Cramerschen Regel:

$$x + 2z = 3, \quad 3x + y = 5 \quad \text{und} \quad -x + y = 1.$$

Aufgabe A9.46

Berechne die Determinante der folgenden Matrix $A \in \mathrm{Mat}_4(K)$ mit Hilfe des Gauß-Algorithmus für $K = \mathbb{Q}$, $K = \mathbb{F}_5$ und $K = \mathbb{F}_3$:

$$A = \begin{pmatrix} 1 & 0 & 2 & 1 \\ 0 & 1 & 3 & 1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

Ist A invertierbar? Falls ja, berechne die Inverse von A.

Aufgabe A9.47

Es sei $A \in \mathrm{Mat}_n(\mathbb{R})$ 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 A9.48

Zeige, ist $A \in \mathrm{Mat}_n(K)$ eine invertierbare obere Dreiecksmatrix, so ist auch A^{-1} eine obere Dreiecksmatrix.

Aufgabe A9.49

Betrachte das folgende lineare Gleichungssystem über $\mathbb{R}$:

$$\begin{aligned} 4x_1 - x_2 - x_3 &= 1, \\ x_2 + 2x_3 &= 0, \\ 3x_1 - x_2 - x_3 &= 1. \end{aligned}$$

a. Berechne die Determinante der Koeffizientenmatrix A des Gleichungssystems.

b. Zeige, daß das Gleichungssystem eindeutig lösbar ist.

c. Bestimme die Lösung des Gleichungssystems mit Hilfe des Gauß-Algorithmus'.

d. Bestimme die Lösung mit Hilfe der Cramerschen Regel.

e. Berechne A^{-1} mit Hilfe des Gauß-Algorithmus' und leite daraus die Lösung des Gleichungssystems her.

f. Berechne A^{-1} mit Hilfe der Adjunkten und leite daraus die Lösung her.

Aufgabe A9.50

Es sei K ein Körper und $\lambda \in K$. Für $n \in \mathbb{N} \setminus \{0\}$ definieren wir

$$A_{\lambda,n} = \begin{pmatrix} \lambda & -1 & 0 & \dots & \dots & \dots & 0 \\ -1 & \lambda & -1 & 0 & \dots & \dots & 0 \\ 0 & -1 & \lambda & -1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & -1 & \lambda & -1 & 0 \\ 0 & \dots & \dots & 0 & -1 & \lambda & -1 \\ 0 & \dots & \dots & \dots & 0 & -1 & \lambda \end{pmatrix} \in \mathrm{Mat}_n(\mathbb{R})$$

und $d_{\lambda,n} = \det(A_{\lambda,n})$. Bearbeite **eine** der beiden Teilaufgaben.

a. Beweise die folgende Rekursionsformel für $d_{\lambda,n}$ für $n \geq 3$:

$$d_{\lambda,n} = \lambda \cdot d_{\lambda,n-1} - d_{\lambda,n-2}.$$

b. Zeige mittels Induktion nach n, daß $d_{1,3\cdot n+2} = 0$ für alle $n \in \mathbb{N}$ gilt.

Aufgabe A9.51

Für $n \geq 1$ definieren wir die Zahl

$$d_n := \sum_{\sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) \cdot 2^{|\{k \mid \sigma(k)=k\}|}.$$

a. Finde eine Matrix $A_n \in \mathrm{Mat}_n(\mathbb{Z})$ ohne Null-Eintrag, so daß $\det(A_n) = d_n$ gilt.

b. Berechne d_n mit Hilfe elementarer Zeilenoperationen.

Aufgabe A9.52

Die Menge $R = \mathcal{C}^\infty(\mathbb{R}, \mathbb{R})$ der auf $\mathbb{R}$ unendlich oft differenzierbaren Funktionen ist ein kommutativer Ring mit 1 mit der üblichen punktweisen Addition und Multiplikation von Funktionen. Für eine solche Funktion g bezeichne g' die Ableitungsfunktion, und für eine Matrix $G = (g_{ij})_{i,j=1,\ldots,n} \in \mathrm{Mat}_n(R)$ definieren wir

$$d_k(G) := \det \begin{pmatrix} g_{11} & \cdots & g_{1k-1} & g'_{1k} & g_{1k+1} & \cdots & g_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ g_{n1} & \cdots & g_{nk-1} & g'_{nk} & g_{nk+1} & \cdots & g_{nn} \end{pmatrix}.$$

a. Zeige, $\det(G)' = d_1(G) + \ldots + d_n(G)$.

b. Zeige, $\det(G)' = \mathrm{Spur}(G^\# \circ G')$.

Die Spur einer quadratischen Matrix ist dabei die Summe der Diagonalelemente.

Aufgabe A9.53

Für welche Werte $x \in \mathbb{R}$ ist die folgende Familie von Vektoren im $\mathbb{R}^3$ linear abhängig

$$\left(\left(\sin(\pi), \cos\left(x + \frac{\pi}{4}\right), \tan\left(x - \frac{\pi}{4}\right) \right)^t, \left(\sin\left(x - \frac{\pi}{4}\right), -\cos\left(\frac{\pi}{2}\right), \ln\left(\frac{4}{\pi}\right) \right)^t, \right.$$
$$\left. \left(\cot\left(x + \frac{\pi}{4}\right), \ln\left(\frac{\pi}{4}\right), \tan(\pi) \right)^t \right).$$

A10 Der Polynomring $K[t]$

In diesem Abschnitt sei K stets ein Körper.

A) Grundlegende Eigenschaften des Polynomrings

Bei der Behandlung von Endomorphismen spielt das charakteristische Polynom des Endomorphismus eine wichtige Rolle. Wir führen deshalb hier den Begriff des Polynoms noch einmal ein und listen einige wichtige Eigenschaften des Polynomrings. Auf Beweise werden wir weitgehend verzichten, da diese Bestandteil der Linearen Algebra 2 sind.

Definition A10.1 (Der Polynomring)
Wir nennen einen Ausdruck der Form

$$\sum_{k=0}^{n} a_k \cdot t^k = a_n \cdot t^n + a_{n-1} \cdot t^{n-1} + \ldots + a_1 \cdot t^1 + a_0 \cdot t^0$$

mit $a_0, \ldots, a_n \in K$ und $n \in \mathbb{N}$ ein *Polynom* in der Unbestimmten t mit Koeffizienten in K.
Die Menge

$$K[t] = \left\{ \sum_{k=0}^{n} a_k \cdot t^k \;\middle|\; a_k \in K, n \in \mathbb{N} \right\}$$

aller Polynome in der Unbestimmten t mit Koeffizienten in K heißt der *Polynomring* in der Unbestimmten t über dem Körper K.

Bemerkung A10.2 (Rechenoperationen im Polynomring)
Für zwei Polynome $f = \sum_{k=0}^{n} a_k \cdot t^k \in K[t]$ und $g = \sum_{k=0}^{m} b_k \cdot t^k \in K[t]$ gelten

$$f + g = \sum_{k=0}^{\max\{m,n\}} (a_k + b_k) \cdot t^k \tag{209}$$

mit $a_k = 0$ für $k > n$ und $b_k = 0$ für $k > m$ sowie

$$f \cdot g = \sum_{k=0}^{m+n} \left(\sum_{i=0}^{k} a_i \cdot b_{k-i} \right) \cdot t^k = \sum_{k=0}^{m+n} \left(\sum_{i+j=k} a_i \cdot b_j \right) \cdot t^k \tag{210}$$

wobei $a_i = 0$ für $i > n$ und $b_j = 0$ für $j > m$.

Ein Skalar $\lambda \in K$ ist ein konstantes Polynom und für das Produkt mit λ gilt dann

$$\lambda \cdot f = \sum_{k=0}^{n} (\lambda \cdot a_k) \cdot t^k.$$

Dies wird im Zusammenhang mit Vektorräumen interessant sein.

Wir erinnern zudem daran: für zwei Polynome $f = \sum_{k=0}^{n} a_k \cdot t^k$ und $g = \sum_{k=0}^{m} b_k \cdot t^k$ gilt

$$f = g \iff a_k = b_k \quad \forall\, k = 0, \ldots, \max\{m, n\}, \tag{211}$$

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*.

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 A6.9).

Proposition A10.3 (Der Polynomring als K-Algebra.)
Der Polynomring $K[t]$ ist eine kommutative K-Algebra mit Basis $B = \left(t^k \mid k \in \mathbb{N}\right)$.

Beweis: In weiterführenden Vorlesungen 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 A10.24) 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 $\left(t^k \mid k \in \mathbb{N}\right)$ ein Erzeugendensystem des K-Vektorraumes $K[t]$, und wegen (211) 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 $k = 0, \ldots, n$. □

Bemerkung A10.4
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

$$a_n \cdot t^n + \ldots + a_1 \cdot t^1 + a_0 \cdot t^0 = a_n \cdot t^n + \ldots + a_1 \cdot t + a_0.$$

Definition A10.5 (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

$$\mathrm{lc}(f) := a_n$$

der *Leitkoeffizient* von f. Zudem setzen wir $\deg(0) := -\infty$ und $\mathrm{lc}(0) := 0$. Ist $\mathrm{lc}(f) = 1$ oder $f = 0$, so nennen wir f *normiert*.

Bemerkung A10.6 (Konstante Polynome)

Beachte, ein Polynom f ist genau dann konstant, wenn $\deg(f) \leq 0$.

Lemma A10.7 (Gradformeln)

Seien $f, g \in K[t]$. Dann gelten:

a. $\deg(f+g) \leq \max\{\deg(f), \deg(g)\}$.

b. $\deg(f \cdot g) = \deg(f) + \deg(g)$.

Beweis: Ist $f = 0$ oder $g = 0$, so sind die Aussagen offenbar korrekt und wir können deshalb $f \neq 0 \neq g$ annehmen. Dann folgt a. unmittelbar aus (209) und b. aus (210), wobei wir für b. beachten, daß der Koeffizient von t^{m+n} gerade $a_m \cdot b_n = \mathrm{lc}(f) \cdot \mathrm{lc}(g)$ ist. □

Beispiel A10.8 (Grandformeln)

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

B) Nullstellen und Division mit Rest

Definition A10.9 (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 A10.10

Sei $f = t^2 - 4 = t^2 - 4 \cdot t^0 \in \mathbb{R}[t]$, $L = \mathrm{Mat}_2(\mathbb{R})$ und $b = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$, dann gilt

$$f(b) = b^2 - 4 \cdot b^0 = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}^2 - 4 \cdot \mathbb{1}_2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Also ist b eine Nullstelle von f in $\mathrm{Mat}_2(\mathbb{R})$.

Satz A10.11 (Division mit Rest)
Es seien $f, g \in K[t] \setminus \{0\}$ und $I \subseteq K[t]$ ein Ideal.

a. Es gibt eindeutige Polynome $q, r \in K[t]$ mit $f = q \cdot g + r$ und $\deg(r) < \deg(g)$.

b. 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.

c. Ist $\deg(f) = n$, so hat f höchstens n Nullstellen in K.

d. Es gibt genau ein normiertes Polynom $\mu \in K[t]$ mit $I \stackrel{!}{=} \{\mu \cdot p \mid p \in K[t]\}$.

Beweis: Für den Beweis verweisen wir auf [Mar08, Satz 7.27, Proposition 7.39, Satz 7.42, Korollar 7.51]. 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. In Teil d. wählt man μ als normiertes Nicht-Null-Polynom kleinsten Grades, wenn $I \neq 0$, und zeigt mit Hilfe von Teil a., daß jedes Polynom in I ein Vielfaches von μ ist. □

Beispiel A10.12
Sei $f = t^3 - 1 \in \mathbb{R}[t]$, dann gilt offenbar $f(1) = 1^3 - 1 = 0$. Polynomdivision liefert:

$$\begin{array}{rl} (t^3 & -1) : (t-1) = t^2 + t + 1. \\ \underline{t^3 - t^2} & \\ t^2 & \\ \underline{t^2 - t} & \\ t - & 1 \\ \underline{t - & 1} \\ - & \end{array}$$

Also gilt $f = (t^2 + t + 1) \cdot (t - 1)$.

C) Das Minimalpolynom

Proposition A10.13 (Das Minimalpolynom)
Es sei L eine K-Algebra und $b \in L$.

a. Der *Einsetzhomomorphismus*

$$\phi_b : K[t] \longrightarrow L : f \mapsto f(b)$$

ist ein K-Algebrenhomomorphismus, d.h. $(f+g)(b) = f(b)+g(b)$, $(f \cdot g)(b) = f(b) \cdot g(b)$, $(\lambda \cdot f)(b) = \lambda \cdot f(b)$ und $1(b) = 1$.

b. Es gibt ein eindeutig bestimmtes normiertes Polynom $\mu_b \in K[t]$, so daß

$$\mu_b K[t] := \{\mu_b \cdot g \mid g \in K[t]\} \stackrel{!}{=} \{h \in K[t] \mid h(b) = 0\} =: \mathrm{Ker}(\phi_b).$$

μ_b heißt das *Minimalpolynom* von 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

$$\mathrm{Ker}(\phi_b) = \{h \in K[t] \mid h(b) = 0\}$$

von ϕ_b ein Ideal im Ring $K[t]$, siehe [Mar08, Satz 6.43]. Aus der Algebra 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

$$\mathrm{Ker}(\phi_b) = \mu_b K[t] = \{\mu_b \cdot g \mid g \in K[t]\}.$$

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. □

Bemerkung A10.14 (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 A10.11 somit das Nullpolynom sein.

D) Der Fundamentalsatz der Algebra

Definition A10.15 (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* $\mathrm{mult}(f,\lambda) = m$.

b. Es sei $K \subseteq L$ ein Teilkörper des Körpers L und $f \in K[t]$ mit $n = \deg(f) > 0$. Gibt es $b_1, \ldots, b_n \in L$ und ein $0 \neq c \in L$ mit $f = c \cdot (t-b_1) \cdot \ldots \cdot (t-b_n)$ so sagen wir, daß f über L in *Linearfaktoren* zerfällt.

c. Wir nennen einen Körper K *algebraisch abgeschlossen*, wenn jedes nichtkonstante Polynom in $K[t]$ über K in Linearfaktoren zerfällt.

Beispiel A10.16

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 $\mathrm{mult}(f,1) = 2$. 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 A10.17 (Fundamentalsatz der Algebra)

a. Der Köper $\mathbb{C}$ 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* $\overline{K}$ 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. □

E) Irreduzible Polynome

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 A10.18 (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 A10.19
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 A10.20 (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: Für den Beweis verweisen wir auf [Mar08, Satz 7.65]. □

Beispiel A10.21
Das Polynom $f = t^4 - 2t^3 + 2t^2 - 2t + 1$ aus Beispiel A10.16 hat in $\mathbb{R}[t]$ die Primfaktorzerlegung

$$f = (t-1)^2 \cdot \left(t^2 + 1\right)$$

und in $\mathbb{C}[t]$ die Primfaktorzerlegung

$$f = (t-1)^2 \cdot (t-i) \cdot (t+i).$$

Satz A10.22 (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 $q_1, \ldots, q_r \in K[t]$ normierte Polynome und gibt es keinen Primfaktor, den alle gemeinsam haben, dann gibt es Polynome $p_1, \ldots, p_r \in K[t]$ mit

$$1 = p_1 \cdot q_1 + \ldots + p_r \cdot q_r.$$

Beweis: Für den Beweis verweisen wir auf [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 A10.23 (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 Algebra, 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

$$K(t) = \left\{ \frac{f}{g} \;\middle|\; f, g \in K[t], g \neq 0 \right\}$$

der *rationalen Funktionen*. Das Kürzen von Brüchen funktioniert wie in den rationalen Zahlen, und gleiches gilt für die Addition und die Multiplikation.

Bemerkung A10.24 (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

$$V = \mathrm{Lin}\,(e_k \mid k \in \mathbb{N}) = \mathrm{Lin}\,(e_0, e_1, e_2, \ldots)$$

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 A10.25 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 Ojekte, 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 A10.25 (Existenz- und Eindeutigkeitssatz für bilineare Abbildungen)
Sei V ein K-Vektorraum mit Basis $B = (x_i \mid i \in I)$ und sei $F = \big(y_{ij} \mid (i,j) \in I \times I\big)$ eine Familie von Vektoren im K-Vektorraum W. Dann gibt es genau eine multilineare Abbildung $f : V \times V \to W$ mit $f(x_i, x_j) = y_{ij}$ für alle $(i,j) \in I \times I$. Sind $x = \sum\limits_{\substack{i \in I \\ \text{endlich}}} a_i x_i, y = \sum\limits_{\substack{i \in I \\ \text{endlich}}} b_i x_i \in V$, so gilt

$$f(x,y) = \sum_{\substack{i \in I \\ \text{endlich}}} \sum_{\substack{j \in I \\ \text{endlich}}} a_i b_j y_{ij}. \tag{212}$$

Aufgabe A10.26 (Polynominterpolation)
Es seien $b_0, \ldots, b_n \in K$ paarweise verschieden und $c_0, \ldots, c_n \in K$ 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 $i = 0, \ldots, n$.

Aufgabe A10.27
Zerlege das Polynom $f = t^4 + t^3 + 2t - 4 \in \mathbb{C}[t]$ in Linearfaktoren.

Aufgabe A10.28
Bestimme das Minimalpolynom $\mu_b \in \mathbb{Q}[t]$ von $b = \sqrt{2} \in \mathbb{R}$.

Aufgabe A10.29

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 $g = (t - \lambda) \cdot (t - \overline{\lambda})$.

Aufgabe A10.30

Es sei $f \in \mathrm{End}_K(V)$ und $x \in V$.

a. Zeige, daß die Menge

$$I_{f,x} := \{p \in K[t] \mid p(f)(x) = 0\}$$

ein Ideal in $K[t]$ ist.

b. Zeige, daß die Menge

$$U_{f,x} := \{p(f)(x) \mid p \in K[t]\}$$

ein Unterraum von V ist.

c. Zeige, ist $m \in \mathbb{N}$ minimal mit $f^m(x) = 0$, so gilt

$$I_{f,x} = \{t^m \cdot p \mid p \in K[t]\}$$

und

$$U_{f,x} = \mathrm{Lin}\left(f^{m-1}(x), f^{m-2}(x), \ldots, f(x), x\right)$$

ist ein zyklischer Unterraum wie in Aufgabe A6.51.

Aufgabe A10.31

Zeige, ist K ein Körper, dann gilt $K[t]^* = K \setminus \{0\}$.

Aufgabe A10.32

Zeige, $f = \overline{2} \cdot t - \overline{1} \in \mathbb{Z}_4[[t]]$ ist eine Einheit und bestimme die Inverse von f.

Aufgabe A10.33

Bestimme alle Nullstellen des Polynoms

$$f = t^4 - t^3 + t^2 - t \in K[t]$$

für die Fälle $K = \mathbb{C}$, $K = \mathbb{R}$, $K = \mathbb{Z}_2$, $K = \mathbb{Z}_7$ und $K = \mathbb{Z}_{13}$. Was sind jeweils die irreduziblen Faktoren des Polynoms?

Aufgabe A10.34
Ist K ein Körper und $f \in K[t]$ ein Polynom mit $deg(f) \in \{2, 3\}$.
Zeige, f ist genau dann irreduzibel, wenn f keine Nullstelle hat.

Aufgabe A10.35
Bestimme alle irreduziblen Polynome f in $\mathbb{Z}_2[t]$ vom Grad $\deg(f) \leq 4$.

Aufgabe A10.36

a. Zeige, hat ein normiertes Polynom $f = t^n + a_{n-1}t^{n-1} + \ldots + a_0$ mit $a_0, \ldots, a_{n-1} \in \mathbb{Z}$ eine Nullstelle in $\lambda \in \mathbb{Q}$, dann gilt schon $\lambda \in \mathbb{Z}$.

b. Ist $f = t^3 + 10t^2 - 881t + 2623 \in \mathbb{Q}[t]$ irreduzibel?

Aufgabe A10.37
Ist die Menge $I = \{f \in \mathbb{Q}[t] \mid f(0) = 1\}$ ein Ideal in $\mathbb{Q}[t]$?

Aufgabe A10.38
Teile das Polynom $f = 2t^5 - t^4 + 2t^3 - t + 1 \in K[t]$ mit Rest durch $g = t^2 + 2t + 3$ für $K = \mathbb{Q}$ und $K = \mathbb{F}_5$.

Aufgabe A10.39
Aus Aufgabe A10.34 wissen wir, daß das Polynom $t^2 + t + \overline{1} \in \mathbb{Z}_2[t]$ irreduzibel ist, weil es in $\mathbb{Z}_2$ keine Nullstelle hat.

a. Zeige, daß die Menge

$$I = \{f \cdot (t^2 + t + \overline{1}) \mid f \in \mathbb{Z}_2[t]\}$$

ein Ideal in $\mathbb{Z}_2[t]$ und damit insbesondere eine Untergruppe von $(\mathbb{Z}_2[t], +)$ ist.

b. Zeige, daß die Faktorgruppe $K := \mathbb{Z}_2[t]/I$ durch die Multiplikation

$$\overline{g} \cdot \overline{h} = \overline{g \cdot h}$$

ein kommutativer Ring mit Eins wird. Wie im Beweis von Satz 25.11 ist i.w. nur die Wohldefiniertheit der Multiplikation zu zeigen.

c. Zeige, daß K in der Tat ein Körper ist und genau vier Elemente enthält.

d. Sind die Gruppen $(K, +)$ und $(\mathbb{Z}_4, +)$ isomorph?

Aufgabe A10.40

Für $n \geq 2$ definieren wir den Polynomring $K[x_1, \ldots, x_n]$ rekursiv als

$$K[x_1, \ldots, x_n] = \big(K[x_1, \ldots, x_{n-1}]\big)[x_n]$$

Wir nennen $K[x_1, \ldots, x_n]$ den *Polynomring in n Veränderlichen*, und die Elemente von $K[x_1, \ldots, x_n]$ nennen wir *Polynome in n Veränderlichen*. $K[x_1, \ldots, x_n]$ ist ein K-Vektorraum, mehr noch, sogar eine kommutative K-Algebra mit 1_K als Eins.

Wir setzen für $\nu = (\nu_1, \ldots, \nu_n) \in \mathbb{N}^n$ und nun

$$|\nu| = \nu_1 + \ldots + \nu_n \quad \text{und} \quad x^\nu := x_1^{\nu_1} \cdots x_n^{\nu_n},$$

wobei x ein neues Symbol ist. Zeige:

a. $\big(x^\nu \;\big|\; \nu \in \mathbb{N}^n\big)$ ist eine Basis von $K[x_1, \ldots, x_n]$ als K-Vektorraum. Insbesondere, jedes Polynom $f \in K[x_1, \ldots, x_n]$ hat eine eindeutige Darstellung als endliche Linearkombination der Form

$$f = \sum_{\nu \in \mathbb{N}^n}{}' a_\nu x^\nu = \sum_{|\nu|=0}^{d} a_\nu x_1^{\nu_1} \cdots x_n^{\nu_n}.$$

Die Basiselemente $x^\nu = x_1^{\nu_1} \cdots x_n^{\nu_n}$ nennen wir *Monome*.

b. Man leite die offensichtlichen Formeln für das Produkt und die Summe zweier Polynome sowie für das skalare Vielfache eines Polynoms her.

c. Ist f nicht das Nullpolynom, so heißt

$$\deg(f) := \max\{|\nu| \mid a_\nu \neq 0\}$$

Grad des Polynoms f und wir setzen $\deg(0) = -\infty$. Man leite Gradformeln für $K[x_1, \ldots, x_n]$ her.

d. Ist $\lambda = (\lambda_1, \ldots, \lambda_n)^t \in K^n$, so ist der Einsetzhomomorphismus

$$\phi_\lambda : K[x_1, \ldots, x_n] \to K : f = \sum_{|\nu|=0}^{d} a_\nu x_1^{\nu_1} \cdots x_n^{\nu_n} \mapsto f(\lambda_1, \ldots, \lambda_n)$$

mit $f(\lambda_1, \ldots, \lambda_n) := \sum_{|\nu|=0}^{d} a_\nu \lambda_1^{\nu_1} \cdots \lambda_n^{\nu_n}$ ein K-Algebrenhomomorphismus.

e. Ein Polynom $f \in \mathbb{R}[x_1, \ldots, x_n]$ heißt *homogen vom Grad d*, wenn für alle $\lambda \in \mathbb{R}$ gilt $f(\lambda \cdot x) = \lambda^d \cdot f$. Zeige, $f \in \mathbb{R}[x_1, \ldots, x_n]$ ist genau dann homogen vom Grad d, wenn $f = \sum_{|\nu|=d} a_\nu x^\nu$, d. h. wenn in der Darstellung von f nur Monome vom Grad d mit Koeffizienten ungleich Null vorkommen können.

Aufgabe A10.41

Es sei A eine K-Algebra und $a \in A$. Wir nennen A *frei in* a, falls für jede K-Algebra B und jedes $b \in B$ gilt, daß es *genau einen* K-Algebrenhomomorphismus $\phi_b : A \to B$ gibt mit $\phi_b(a) = b$. Zeige:

a. $K[t]$ ist frei in t.

b. Ist A frei in a, so gibt es genau einen *Isomorphismus* $\phi_b : K[t] \to A$ mit $\phi_b(t) = a$.

Damit ist dann $K[t]$ bis auf eindeutige Isomorphie die einzige freie K-Algebra in einer Veränderlichen.

A11 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 A11.1 (Endomorphismen)
Wir erinnern uns, daß K-lineare Abbildungen

$$f : V \longrightarrow V$$

auch *Endomorphismen* des K-Vektorraums V genannt werden (siehe Definition A3.19) und daß

$$\mathrm{End}_K(V) = \{f : V \longrightarrow V \mid f \text{ ist } K\text{-linear}\}$$

die *K-Algebra* der Endomorphismen von V ist (siehe Bemerkung A6.9).

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 A6.16)

$$M_D^D(f) = T^{-1} \circ M_B^B(f) \circ T.$$

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 A11.2 (Konjugiert oder ähnlich)
Zwei quadratische Matrizen $A, B \in \mathrm{Mat}_n(K)$ heißen *konjugiert* oder *ähnlich*, wenn es eine invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ gibt, so daß $B = T^{-1} \circ A \circ T$ ist.

Bemerkung A11.3 (Konjugation ist eine Äquivalenzrelation)
Konjugation von Matrizen ist ein Beispiel für eine Äquivalenzrelation auf der Menge $\mathrm{Mat}_n(K)$ 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.

Definition A11.4 (Das Charakteristische Polynom einer Matrix)
Es sei $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ eine quadratische Matrix.

a. Wir nennen
$$\chi_A := \det(t \cdot \mathbb{1}_n - A) \in K[t]$$
das *charakteristische Polynom* von A, wobei
$$t \cdot \mathbb{1}_n - A = \begin{pmatrix} t - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & t - a_{22} & \dots & -a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ -a_{n1} & -a_{n2} & \dots & t - a_{nn} \end{pmatrix} \in \mathrm{Mat}_n\left(K[t]\right)$$
eine quadratische Matrix mit Polynomen als Einträgen ist.

b. Wir definieren die *Spur* von A als Summe der Diagonalelemente, d.h.
$$\mathrm{Spur}(A) := \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + a_{33} + \ldots + a_{nn}.$$

Beispiel A11.5
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 $-\mathrm{Spur}(A) = -8$ ist.

Proposition A11.6 (Charakteristisches Polynom)
Es sei $A \in \mathrm{Mat}_n(K)$ eine quadratische Matrix. Dann ist das Polynom
$$\chi_A = t^n + \alpha_{n-1} \cdot t^{n-1} + \alpha_{n-2} \cdot t^{n-2} + \ldots + \alpha_1 \cdot t + \alpha_0 \in K[t]$$
ein normiert vom Grad n mit $\alpha_{n-1} = -\mathrm{Spur}(A)$ und $\alpha_0 = (-1)^n \cdot \det(A)$.

Beweis: Ist $A = (a_{ij})$ und $t \cdot \mathbb{1}_n - A = \big(p_{ij}\big)$, dann folgt aus der Leibnitzschen Formel für die Determinante

$$\chi_A = \det(t \cdot \mathbb{1}_n - A) = (t - a_{11}) \cdots (t - a_{nn}) + \sum_{\mathrm{id} \neq \sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) \cdot p_{1\sigma(1)} \cdots p_{n\sigma(n)}.$$

Da für $\sigma \neq \mathrm{id}$ mindestens zwei Faktoren in $p_{1\sigma(1)} \cdots p_{n\sigma(n)}$ konstante Polynome sind, ergibt $\sum_{\mathrm{id} \neq \sigma \in \mathbb{S}_n} \operatorname{sgn}(\sigma) \cdot p_{1\sigma(1)} \cdots p_{n\sigma(n)}$ 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

$$\alpha_{n-1} = -a_{11} - a_{22} - \ldots - a_{nn} = -\operatorname{Spur}(A).$$

Ferner ist

$$\alpha_0 = \chi_A(0) = \det(-A) = (-1)^n \cdot \det(A)$$

der konstante Term im charakteristischen Polynom. □

Bemerkung A11.7

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. Diese Tatsache haben wir im obigen Beweis bei der Berechnung des konstanten Terms des charakteristischen Polynoms verwendet.

Proposition A11.8 (Charakteristisches Polynom konjugierter Matrizen)

Sind $A, B \in \mathrm{Mat}_n(K)$ konjugiert, so gelten

$$\chi_A = \chi_B, \quad \det(A) = \det(B) \quad \text{und} \quad \operatorname{Spur}(A) = \operatorname{Spur}(B).$$

Beweis: Sei $T \in \mathrm{Gl}_n(K)$ mit $B = T^{-1} \circ A \circ T$, dann gilt auch

$$T^{-1} \circ (t \cdot \mathbb{1}_n - A) \circ T = t \cdot T^{-1} \circ \mathbb{1}_n \circ T - T^{-1} \circ A \circ T = t \cdot \mathbb{1}_n - B.$$

Um den Determinantenmultiplikationssatz A9.19 anwenden zu können, betrachten wir die Matrix $t \cdot \mathbb{1}_n - A$ als Matrix mit Einträgen im Körper $K(t)$ der rationalen Funktionen über K. Damit erhalten wir dann

$$\begin{aligned}
\chi_B &= \det(t \cdot \mathbb{1}_n - B) = \det\big(T^{-1} \circ (t \cdot \mathbb{1}_n - A) \circ T\big) \\
&= \det(T^{-1}) \cdot \det(t \cdot \mathbb{1}_n - A) \cdot \det\big(T\big) \\
&= \frac{1}{\det(T)} \cdot \det(t \cdot \mathbb{1}_n - A) \cdot \det(T) = \chi_A.
\end{aligned}$$

Die Aussagen zur Determinante und Spur folgen dann aus Proposition A11.8. □

Damit können wir das charakteristische Polynom, die Determinante und die Spur eines Endomorphismus definieren.

Definition A11.9 (Charakteristisches Polynom eines Endomorphismus)
Sei V ein endlich-dimensionaler K-Vektorraum der Dimension mindestens Eins mit Basis B und $f \in \mathrm{End}_K(V)$. Wir definieren

- das *charakteristische Polynom* von f durch
$$\chi_f := \chi_{M_B^B(f)},$$
- die *Determinante* von f durch
$$\det(f) := \det\left(M_B^B(f)\right)$$
- und die *Spur* von f durch
$$\mathrm{Spur}(f) := \mathrm{Spur}\left(M_B^B(f)\right).$$

Da die Matrixdarstellungen eines Endomorphismus f zu verschiedenen Basen nach Bemerkung A11.3 konjugiert sind, sind diese Definitionen unter Berücksichtigung von Proposition A11.8 unabhängig von der Wahl der Basis B.

Beispiel A11.10
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

$$M_E^E(f) = \begin{pmatrix} 5 & 2 \\ 6 & 3 \end{pmatrix}.$$

Alternativ könnte man die Basis $B = \left((1, 1)^t, (0, 1)^t\right)$ betrachten und erhält dann

$$M_B^B(f) = \begin{pmatrix} 7 & 2 \\ 2 & 1 \end{pmatrix}.$$

Das charakteristische Polynom von f erhalten wir dann als

$$\chi_f = \chi_{M_E^E(f)} = \det \begin{pmatrix} t-5 & -2 \\ -6 & t-3 \end{pmatrix} = t^2 - 8t + 3$$

oder alternativ als

$$\chi_f = \chi_{M_B^B(f)} = \det \begin{pmatrix} t-7 & -2 \\ -2 & t-1 \end{pmatrix} = (t-7) \cdot (t-1) - 4 = t^2 - 8t + 3.$$

Das charakteristische Polynom verträgt sich gut mit f-invarianten Unterräumen.

Proposition A11.11 (Charkteristisches Polynom und f-Invarianz)
Es sei $1 \leq \dim_K(V) < \infty$, $f \in \mathrm{End}_K(V)$.

a. Ist $U \subseteq V$ ein f-invarianter Unterraum, dann gilt

$$\chi_f = \chi_{f_U} \cdot \chi_{f_{V/U}}.$$

b. Ist $V = U_1 \oplus \ldots \oplus U_k$, wobei die U_i f-invariant seien, dann gilt

$$\chi_f = \chi_{f_{U_1}} \cdot \ldots \cdot \chi_{f_{U_k}}.$$

Beweis:

a. Wir wählen eine Basis $B' = (x_1, \ldots, x_k)$ von U und ergänzen diese zu einer Basis $B = (x_1, \ldots, x_n)$ von V. Dann ist $B'' = (\overline{x_{k+1}}, \ldots, \overline{x_n})$ eine Basis von V/U und es gilt

$$M_B^B(f) = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right)$$

nach Proposition A6.20. Für das charakteristische Polynom erhalten wir mit Hilfe des Kästchensatzes A9.24 dann

$$\begin{aligned} \chi_f = \chi_{M_B^B(f)} &= \det \left(\begin{array}{c|c} t \cdot \mathbb{1}_k - M_{B'}^{B'}(f_U) & ** \\ \hline 0 & t \cdot \mathbb{1}_{n-k} - M_{B''}^{B''}(f_{V/U}) \end{array}\right) \\ &= \det\left(t \cdot \mathbb{1}_k - M_{B'}^{B'}(f_U)\right) \cdot \det\left(t \cdot \mathbb{1}_{n-k} - M_{B''}^{B''}(f_{V/U})\right) = \chi_{f_U} \cdot \chi_{f_{V/U}}. \end{aligned}$$

b. Die Aussage folgt analog zu Teil a. aus Proposition A6.21.

□

B) Eigenwerte, Eigenvektoren und Eigenräume

Der Begriff des Eigenwertes ist von zentraler Bedeutung für die Untersuchung von Endomorphismen und quadratischen Matrizen in Bemerkung A12.1 angestrebte Klassifikation.

Definition A11.12 (Eigenwerte und Eigenvektoren)
Sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. $\lambda \in K$ 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, und die Menge

$$\mathrm{Eig}(f, \lambda) := \{y \in V \mid f(y) = \lambda y\}$$

heißt der *Eigenraum* von f zum Eigenwert λ. Die Menge

$$\sigma(f) := \{\lambda \in K \mid \lambda \text{ ist Eigenwert von } f\}$$

der Eigenwerte von f heißt das *Spektrum* von f.

b. $\lambda \in K$ heißt *Eigenwert* von A, falls es ein $0 \neq x \in K^n$ mit $Ax = \lambda x$ gibt. Der Vektor x heißt dann *Eigenvektor* zum Eigenwert λ von A, und die Menge

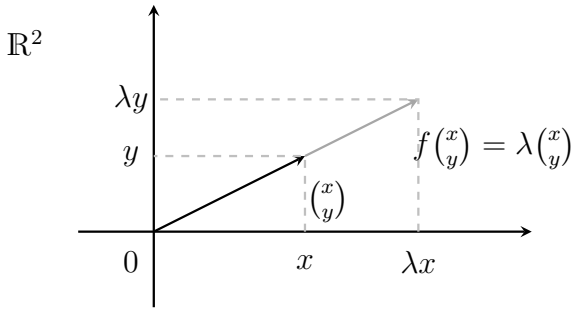
$$\mathrm{Eig}(A, \lambda) := \{y \in V \mid Ay = \lambda y\}$$

heißt *Eigenraum* von A zum Eigenwert λ. Die Menge

$$\sigma(A) := \{\lambda \in K \mid \lambda \text{ ist Eigenwert von } A\}$$

der Eigenwerte von A heißt das *Spektrum* von A.

Bemerkung A11.13 (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 $0 < \lambda < 1$.

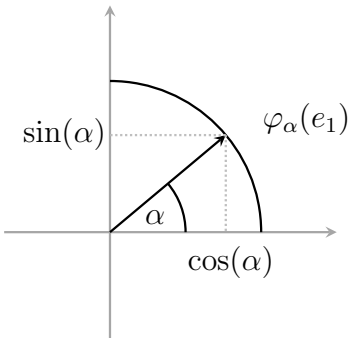


Beispiel A11.14

a. Ist $\dim_K(V) = 1$, so ist jeder Vektor ungleich Null ein Eigenvektor von f, da f schlicht die Multiplikation mit einer Konstanten ist.

b. Ist $\dim_K(V) \geq 2$, so braucht f hingegen keine Eigenwerte und Eigenvektoren zu besitzen. Dabei hängt die Frage der Existenz wesentlich vom Grundkörper K ab. Betrachte etwa die Drehung $\varphi_\alpha = f_{A_\alpha} : \mathbb{R}^2 \to \mathbb{R}^2$ um den Winkel $\alpha \in \mathbb{R}$ aus Beispiel A3.24. 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 A11.15 (Eigenräume)
Es seien $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

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 $f - \lambda\,\mathrm{id}_V \in \mathrm{End}_K(V)$ liegt, gilt also

$$\mathrm{Eig}(f, \lambda) = \mathrm{Ker}(f - \lambda\,\mathrm{id}_V) = \mathrm{Ker}(\lambda\,\mathrm{id}_V - f).$$

Analog erhält man:

$$\mathrm{Eig}(A, \lambda) = \mathrm{Ker}(f_A - \lambda\,\mathrm{id}_V) = \mathrm{Lös}(A - \lambda\mathbb{1}_n, 0) = \mathrm{Lös}(\lambda\mathbb{1}_n - A, 0).$$

b. Aus der Definition folgt unmittelbar, daß λ genau dann ein Eigenwert von A ist, wenn λ ein Eigenwert von f_A ist.

c. Ebenso folgt unmittelbar, daß der Eigenraum $\mathrm{Eig}(f, \lambda)$ von f zum Eigenwert λ f-invariant ist.

d. Kennt man einen Eigenwert $\lambda \in K$ von A, so kann man das lineare Gleichungssystem

$$(A - \lambda\mathbb{1}_n)x = 0 \quad \text{oder} \quad (\lambda\mathbb{1}_n - A)x = 0$$

lösen und damit eine Basis des Eigenraumes $\mathrm{Eig}(A, \lambda) = \mathrm{Lös}(A - \lambda\mathbb{1}_n, 0)$ 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?

C) Eigenwerte als Nullstellen des charakteristischen Polynoms

Satz A11.16 (Eigenwerte und das charakteristische Polynom)
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

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.

Insbesondere, f und A haben höchstens n paarweise verschiedene Eigenwerte.

Beweis: Für $\lambda \in K$ gilt unter Berücksichtigung von Korollar A5.26:

$$\begin{aligned}
\lambda \text{ ist Eigenwert von } f &\iff \mathrm{Ker}(\lambda\,\mathrm{id}_V - f) = \mathrm{Eig}(f,\lambda) \neq \{0\} \\
&\iff \lambda\,\mathrm{id}_V - f \text{ ist nicht injektiv} \\
&\overset{A5.26}{\iff} \lambda\,\mathrm{id}_V - f \text{ ist nicht bijektiv} \\
&\iff \chi_f(\lambda) = \det(\lambda\,\mathrm{id}_V - f) = 0.
\end{aligned}$$

Der Beweis für die Matrizen geht analog. □

Bevor wir das charakteristische Polynom weiter untersuchen, wollen wir zunächst einige Beispiele betrachten.

Beispiel A11.17

a. Betrachten wir zunächst die folgende Matrix:

$$A = \begin{pmatrix} 0 & 1 & 1 \\ -1 & 2 & 1 \\ -1 & 1 & 2 \end{pmatrix} \in \mathrm{Mat}(3,\mathbb{Q}).$$

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 $t\mathbb{1}_n - A \in \mathrm{Mat}_n\big(\mathbb{Q}(t)\big)$ auch als Matrix über dem Körper $\mathbb{Q}(t)$ auffassen (siehe Bemerkung A10.23). Da $\mathbb{Q}(t)$ ein Körper ist, dürfen wir die Determinante mittels des Gaußschen Algorithmus' A9.15

bestimmen. Insbesondere dürfen wir dabei durch Polynome dividieren!

$$
\begin{aligned}
&\begin{pmatrix} t & -1 & -1 \\ 1 & t-2 & -1 \\ 1 & -1 & t-2 \end{pmatrix} \xrightarrow[III \mapsto III - \frac{1}{t} I]{II \mapsto II - \frac{1}{t} I} \begin{pmatrix} t & -1 & -1 \\ 0 & t-2+\frac{1}{t} & \frac{1}{t}-1 \\ 0 & \frac{1}{t}-1 & t-2+\frac{1}{t} \end{pmatrix} = \\
&\begin{pmatrix} t & -1 & -1 \\ 0 & \frac{(t-1)^2}{t} & -\frac{t-1}{t} \\ 0 & -\frac{t-1}{t} & \frac{(t-1)^2}{t} \end{pmatrix} \xrightarrow{III \mapsto III + \frac{1}{t-1} II} \begin{pmatrix} t & -1 & -1 \\ 0 & \frac{(t-1)^2}{t} & -\frac{t-1}{t} \\ 0 & 0 & t-2 \end{pmatrix}.
\end{aligned} \tag{213}
$$

Entsprechend erhalten wir für das charakteristische Polynom

$$\chi_A = t \cdot \tfrac{(t-1)^2}{t} \cdot (t-2) = (t-1)^2 \cdot (t-2).$$

Das charakteristische Polynom hat also die Nullstellen $\lambda = 1$ und $\lambda = 2$, wobei $\lambda = 1$ eine zweifache Nullstelle ist.

Wir können jetzt für $\lambda = 1$ und für $\lambda = 2$ jeweils den Eigenraum $\mathrm{Eig}(A, \lambda) = \mathrm{Lös}(\lambda \mathbb{1}_n - A, 0)$ mit Hilfe des Gauß-Algorithmus bestimmen.[5]

Der Algorithmus zur Bestimmung von $\mathrm{Eig}(A, 1) = \mathrm{Lös}(\mathbb{1}_n - A, 0)$ sieht vor, daß wir die Matrix zunächst auf reduzierte ZSF bringen und dann in den Nullzeilen die Diagonalelemente durch -1 ersetzen:

$$\begin{pmatrix} -1 & 1 & 1 \\ -1 & 1 & 1 \\ -1 & 1 & 1 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 & -1 & -1 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Die letzten beiden Spalten, d.h. die, bei denen eine -1 auf der Diagonalen steht, bilden dann eine Basis des Eigenraumes zum Eigenwert 1:

$$\mathrm{Eig}(A, 1) = \mathrm{Lin}\left((-1, -1, 0)^t, (-1, 0, -1)^t\right).$$

$\mathrm{Eig}(A, 1)$ ist also zweidimensional.

Analog ergibt sich $\mathrm{Eig}(A, 2)$ 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 $\mathrm{Eig}(A, 2) = \mathrm{Lin}\left((-1, -1, -1)^t\right)$.

[5] Man beachte, daß es zur Berechnung der reduzierten Zeilen-Stufen-Form von $\lambda \mathbb{1}_n - A$ für $\lambda = 1$ *nicht* erlaubt ist, in (213) in der letzten Matrix t etwa durch $\lambda = 1$ zu ersetzen, um die ZSF zu erhalten, da wir bei den vorgenommenen Umformungen zur Ermittelung 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!

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

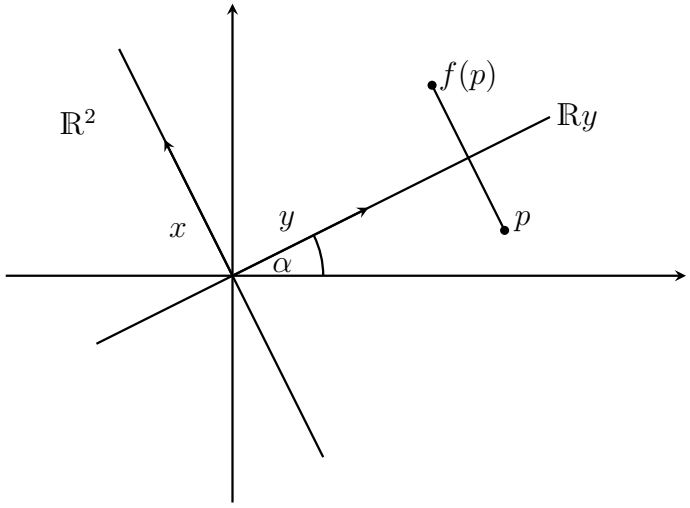
$$A_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}.$$

Aber

$$\chi_{A_\alpha} = \big(t - \cos(\alpha)\big)^2 + \sin^2(\alpha) = t^2 - 2\cos(\alpha)t + 1.$$

Die Nullstellen von χ_{A_α} in $\mathbb{C}$ sind $\cos(\alpha) \pm i \cdot |\sin(t)|\sqrt{\cos^2(\alpha) - 1}$. 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 $f \in \mathrm{End}_{\mathbb{R}}\left(\mathbb{R}^2\right)$ die *Spiegelung* an einer Geraden $\mathrm{Lin}\,(y) = \mathbb{R} \cdot y \subset \mathbb{R}^2$ mit $0 \neq y = (y_1, y_2)^t \in \mathbb{R}^2$.



Wir setzen $x = (-y_2, y_1)^t \in \mathbb{R}^2$. Dann steht x senkrecht auf y und $B = (y, x)$ ist eine Basis von $\mathbb{R}^2$. Die Spiegelung f bildet mithin y auf sich selbst und x auf $-x$ ab, da x senkrecht auf $\mathrm{Lin}\,(y)$ 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 die Eigenwerte 1 und -1.
Beschreiben wir f in den Standardkoordinaten $E = (e_1, e_2)$ von $\mathbb{R}^2$, so ist f die Spiegelung an $\mathrm{Lin}\,(e_1) = \mathbb{R} \cdot e_1$ gefolgt von der Drehung um den Winkel 2α, wenn

α der Winkel ist, den $\mathrm{Lin}\,(y)$ mit $\mathrm{Lin}\,(e_1)$ 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

$$\big(t - \cos(2\alpha)\big) \cdot \big(t + \cos(2\alpha)\big) - \sin^2(2\alpha) = t^2 - 1 = (t-1) \cdot (t+1).$$

Korollar A11.18 (Eigenwerte einer Dreiecksmatrix)
Ist $A = (a_{ij}) \in \mathrm{Mat}_n(K)$ eine obere oder untere Dreiecksmatrix, dann ist

$$\chi_A = (t - a_{11}) \cdot \ldots \cdot (t - a_{nn})$$

und die Einträge auf der Diagonalen sind genau die Eigenwerte von A.

Beweis: Für eine obere Dreiecksmatrix $A = (a_{ij})$ gilt

$$\chi_A = \det \begin{pmatrix} t - a_{11} & * & \ldots & \ldots & * \\ 0 & t - a_{22} & * & \ldots & * \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & \ldots & \ldots & 0 & t - a_{nn} \end{pmatrix} = \prod_{i=1}^{n} (t - a_{ii}),$$

und dieses Polynom hat genau die Nullstellen $a_{11}, \ldots, a_{nn}$. Der Beweis für untere Dreiecksmatrizen geht analog. □

D) Geometrische und algebraische Vielfachheit von Eigenwerten

Für die Definition der Vielfachheit $\mathrm{mult}(p, \lambda)$ einer Zahl λ als Nullstelle eines Polynoms p verweisen wir auf Definition A10.15.

Definition A11.19 (Vielfachheit von Eigenwerten)
Es sei $f \in \mathrm{End}_K(V)$, $A \in \mathrm{Mat}_n(K)$ und $\lambda \in K$.

a. $\mathrm{mult}(\chi_f, \lambda)$ heißt *algebraische Vielfachheit* von λ als Eigenwert von f.
$\dim_K \mathrm{Eig}(f, \lambda)$ heißt *geometrische Vielfachheit* von λ als Eigenwert von f.

b. $\mathrm{mult}(\chi_A, \lambda)$ heißt *algebraische Vielfachheit* von λ als Eigenwert von A.
$\dim_K \mathrm{Eig}(A, \lambda)$ heißt *geometrische Vielfachheit* von λ als Eigenwert von A.

Die algebraischen Vielfachheiten nennt man auch *arithmetische Vielfachheiten*.

Beispiel A11.20

Die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C})$$

aus Beispiel A12.5 hat nur den Eigenwert 0, da $\chi_A = t^2$. Die algebraische Vielfachheit von 0 als Eigenwert von A ist

$$\mathrm{mult}(\chi_A, 0) = \mathrm{mult}(t^2, 0) = 2,$$

während die geometrische Vielfachheit

$$\dim_{\mathbb{C}} \mathrm{Eig}(A, 0) = \dim_{\mathbb{C}} \mathrm{Lös}(A, 0) = \dim_{\mathbb{C}} \mathrm{Lin}\left((1, 0)^t\right) = 1$$

ist.

Lemma A11.21 (Geometrische und algebraische Vielfachheit)
Es sei $f \in \mathrm{End}_K(V)$, $A \in \mathrm{Mat}_n(K)$ und $\lambda \in K$. Dann gilt stets

$$\dim_K \mathrm{Eig}(f, \lambda) \leq \mathrm{mult}(\chi_f, \lambda)$$

und

$$\dim_K \mathrm{Eig}(A, \lambda) \leq \mathrm{mult}(\chi_A, \lambda),$$

d.h. die geometrische Vielfachheit eines Eigenwertes ist stets nach oben durch die algebraische Vielfachheit beschränkt.

Beweis: Man beachte, daß $U := \mathrm{Eig}(f, \lambda)$ ein f-invarianter Unterraum ist und daß $f_U = \lambda \cdot \mathrm{id}_U$ gilt. Mithin ist

$$\chi_{f_U} = \chi_{\lambda \cdot \mathrm{id}_U} = \det(t \cdot \mathrm{id}_U - \lambda \cdot \mathrm{id}_U) = \det\left((t - \lambda) \cdot \mathrm{id}_U\right) = (t - \lambda)^s$$

wobei $s = \dim_K(U) = \dim_K \mathrm{Eig}(f, \lambda)$. Außerdem gilt nach Proposition A11.11

$$\chi_f = \chi_{f_U} \cdot \chi_{f_{V/U}} = (t - \lambda)^s \cdot \chi_{f_{V/U}}.$$

Daraus folgt unmittelbar

$$\mathrm{mult}(f, \lambda) \geq s = \dim_K \mathrm{Eig}(f, \lambda).$$

Die analoge Aussage für A folgt hieraus mit $f = f_A$. □

Lemma A11.22 (Eigenwerte bei konjugierten Matrizen)
Für $A, B \in \mathrm{Mat}_n(K)$ und $T \in \mathrm{Gl}_n(K)$ mit $B = T^{-1} \circ A \circ T$ sowie $\lambda \in K$ gelten:

a. $\sigma(A) = \sigma(B)$.

b. $\mathrm{mult}(\chi_A, \lambda) = \mathrm{mult}(\chi_B, \lambda)$.

c. $\dim_K \mathrm{Eig}(A, \lambda) = \dim_K \mathrm{Eig}(B, \lambda)$.

d. $x \in \mathrm{Eig}(A, \lambda) \iff T^{-1}x \in \mathrm{Eig}(B, \lambda)$.

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 A11.8 haben A und B die gleichen charakteristischen Polynome. Mithin stimmen wegen Satz A11.16 die Eigenwerte von A und B sowie deren algebraische Vielfachheiten überein. Damit sind a. und b. gezeigt. Ferner gilt

$$\begin{aligned} x \in \mathrm{Eig}(A,\lambda) &\iff \lambda x = Ax = ATT^{-1}x \\ &\iff \lambda T^{-1}x = T^{-1}ATT^{-1}x = BT^{-1}x \iff T^{-1}x \in \mathrm{Eig}(B,\lambda). \end{aligned}$$

Damit ist d. gezeigt und außerdem folgt, daß der Isomorphismus $f_{T^{-1}}$ den Eigenraum $\mathrm{Eig}(A,\lambda)$ isomorph auf den Eigenraum $\mathrm{Eig}(B,\lambda)$ abbildet. Die beiden müssen also die gleiche Dimension haben, womit auch c. gezeigt ist. □

Aufgaben

Aufgabe A11.23

Berechne das charakteristische Polynom der Matrix

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & 0 & 1 \\ a_0 & \dots & \dots & \dots & \dots & a_{n-2} & a_{n-1} \end{pmatrix} \in \mathrm{Mat}_n(\mathbb{R})$$

für $a_0, \dots, a_{n-1} \in K$.

Aufgabe A11.24

Bestimme die Eigenwerte und die Eigenräume der folgenden Matrix A und entscheide, ob sie diagonalisierbar ist:

$$A = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 3 & 0 & 0 \\ -1 & 1 & 2 & 1 \\ -1 & 1 & 0 & 3 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q}).$$

Aufgabe A11.25 (Nilpotente Endomorphismen und Matrizen)

Es sei $A \in \mathrm{Mat}_n(K)$ und $f \in \mathrm{End}_K(V)$ mit $1 \leq \dim_K(V) < \infty$.

a. Zeige, gibt es ein $r \in \mathbb{N}$ mit $f^r = 0$, so gilt $\mathrm{Spur}(f) = 0$.

b. Zeige, gibt es ein $r \in \mathbb{N}$ mit $A^r = 0$, so gilt $\mathrm{Spur}(A) = 0$.

c. 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 $n = \dim_K(V)$. Dazu zeige man, daß $M_B^B(f)$ für eine geeignete Wahl von B Blockgestalt mit einem Nullblock in der oberen linken Ecke hat. Aufgabe A6.20b. mit $U = \mathrm{Ker}(f)$ ist dabei hilfreich.

Aufgabe A11.26

Es sei $1 \leq \dim_K(V) < \infty$, $f \in \mathrm{End}_K(V)$.

a. Ist $U \subseteq V$ ein f-invarianter Unterraum, dann gilt

$$\det(f) = \det(f_U) \cdot \det(f_{V/U})$$

und

$$\chi_f = \chi_{f_U} \cdot \chi_{f_{V/U}}.$$

b. Ist $V = U_1 \oplus \ldots \oplus U_k$, wobei die U_i f-invariant seien, dann gilt

$$\det(f) = \det(f_{U_1}) \cdot \ldots \cdot \det(f_{U_k})$$

und

$$\chi_f = \chi_{f_{U_1}} \cdot \ldots \cdot \chi_{f_{U_k}}.$$

Aufgabe A11.27

Für ein Polynom $p \in K[t]$ und zwei konjugierte Matrizen $A, B \in \mathrm{Mat}_n(K)$ gilt

$$p(A) = 0 \iff p(B) = 0.$$

Aufgabe A11.28 (Zyklische Unterräume)

Zeige, $\chi_{f_U} = \mu_{f_U} = t^m$ für den Endomorphismus f_U aus Aufgabe A6.51.

Aufgabe A11.29

Sind $f, g \in \mathrm{End}_K(V)$, so ist λ genau dann ein Eigenwert von $f \circ g$, wenn λ ein Eigenwert von $g \circ f$ ist.

Aufgabe A11.30

Bestimme die Eigenwerte und die Eigenräaume der Matrix

$$A = \begin{pmatrix} -1 & 0 & 0 & -1 \\ -4 & 1 & 0 & -2 \\ 0 & 0 & -2 & 0 \\ 4 & 1 & 0 & 4 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{R})$$

und überprüfe A auf Diagonalisierbarkeit.

Aufgabe A11.31

a. Berechne für die folgenden Matrizen das charakteristische Polynom, bestimme die Eigenwerte und entscheide, ob die Matrizen diagonalisierbar sind:

$$A = \begin{pmatrix} 3 & 2 \\ 0 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 3 & 3 \\ 0 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 3 & 2 \\ 2 & 3 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}).$$

b. Berechne für die folgenden Matrizen das charakteristische Polynom, bestimme die Eigenwerte und entscheide, ob die Matrizen diagonalisierbar sind:

$$D = \begin{pmatrix} 3 & 2 & 5 \\ 2 & 3 & 7 \\ 0 & 0 & 2 \end{pmatrix}, \quad E = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R}).$$

c. Zeige, $\lambda \in K$ ist genau dann ein Eigenwert von $A \in \mathrm{Mat}_n(K)$, wenn λ ein Eigenwert von A^t ist.

Aufgabe A11.32

Gegeben sei die folgende Matrix

$$A = \begin{pmatrix} 3 & 0 & 4 & 0 \\ 4 & 2 & 9 & -2 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & 4 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{R}).$$

a. Bestimme das charakteristische Polynom und die Eigenwerte von A.

b. Bestimme Basen der Eigenräume von A.

c. Zeige, daß A diagonalisierbar ist und bestimme eine invertierbare Matrix $T \in \mathrm{Gl}_4(\mathbb{R})$, so dass $T^{-1} \circ A \circ T$ in Diagonalgestalt ist.

Aufgabe A11.33

Es sei $V = \mathrm{Mat}_2(\mathbb{R})$ der $\mathbb{R}$-Vektorraum der 2×2-Matrizen über $\mathbb{R}$. Ferner sei

$$A = \begin{pmatrix} 2 & 1 \\ -1 & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$$

gegeben und wir betrachten den Endomorphismus

$$T_A : V \longrightarrow V \ : \ X \mapsto A \circ X.$$

a. Berechne $\det(T_A)$ und vergleiche den Wert mit $\det(A)$.

b. Berechne $\mathrm{Spur}(T_A)$ und vergleiche den Wert mit $\mathrm{Spur}(A)$.

c. Berechne χ_{T_A} und vergleiche das Polynom mit χ_A.

d. Bestimme die Eigenwerte von T_A.

e. Bestimme für jeden Eigenwert mindestens einen Eigenvektor.

f. Bestimme die Eigenräume von T_A. Ist T_A diagonalisierbar?

Hinweis, man berechne zunächst die Matrixdarstellung $M_B^B(T_A)$ für die Basis $B = (E_1^1, E_2^1, E_1^2, E_2^2)$ von V. Beachte dabei die Reihenfolge der Basisvektoren.

Aufgabe A11.34 (Die Eigenräume bilden eine direkte Summe.)

Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Sind $x_1, \ldots, x_r \in V$ Eigenvektoren von f zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r \in K$, dann ist die Familie $(x_1, \ldots, x_r)$ linear unabhängig. Insbesondere gilt

$$\mathrm{Eig}(f, \lambda_1) + \ldots + \mathrm{Eig}(f, \lambda_r) = \mathrm{Eig}(f, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(f, \lambda_r).$$

b. Sind $x_1, \ldots, x_r \in K^n$ Eigenvektoren von A zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_r \in K$, dann ist die Familie $(x_1, \ldots, x_r)$ linear unabhängig. Insbesondere gilt

$$\mathrm{Eig}(A, \lambda_1) + \ldots + \mathrm{Eig}(A, \lambda_r) = \mathrm{Eig}(A, \lambda_1) \oplus \ldots \oplus \mathrm{Eig}(A, \lambda_r).$$

Aufgabe A11.35

Zeige mit Hilfe der Eigenwerte der Matrix

$$A = \begin{pmatrix} \cos(1) & -\sin(1) \\ \sin(1) & \cos(1) \end{pmatrix},$$

daß die Folgen $(\cos(n))_{n \in \mathbb{N}}$ und $(\sin(n))_{n \in \mathbb{N}}$ divergent sind.

Aufgabe A11.36

Sind $f, g \in \mathrm{End}_K(V)$, so gilt $\sigma(f \circ g) = \sigma(g \circ f)$.

A12 Diagonalisierbarkeit und Trigonalisierbarkeit

In diesem Abschnitt sei V ein K-Vektorraum mit $1 \leq \dim_K(V) = n < \infty$.

Bemerkung A12.1 (Normalformen bezüglich Konjugation als Ziel)
Die nächsten beiden Abschnitte Abschnitte sind folgender Aufgabe gewidmet:

Finde für $f \in \mathrm{End}_K(V)$ 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 für $A \in \mathrm{Mat}_n(K)$ ein invertierbares $T \in \mathrm{Gl}_n(K)$ 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*. Eine Diagonalmatrix als Normalform wie in Satz A12.22 wäre optimal einfach, ist aber sehr speziell, wie selbiger Satz zeigt. Wir wollen nun untersuchen, was unter welchen Voraussetzung erreichbar ist, und suchen dabei nach Normalformen mit möglichst vielen Nullen.

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 A6.37 und Korollar A6.38 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.

Definition A12.2 (Diagonalisierbar und trigonalisierbar)
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. f heißt *trigonalisierbar*, falls es eine Basis B von V gibt, so daß $M_B^B(f)$ eine obere Dreiecksmatrix ist.

b. A heißt *trigonalisierbar*, falls es eine Matrix $T \in \mathrm{Gl}_n(K)$ gibt, so daß $T^{-1} \circ A \circ T$ eine obere Dreiecksmatrix ist.

A) Trigonalisierbarkeit

Satz A12.3 (Trigonalisierbarkeit)
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

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.

Zerfalle 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 A11.14 sogar diagonalisierbar.

Sei also $n > 1$ und sei $0 \neq x_1 \in V$ ein Eigenvektor von f zum Eigenwert λ_1. Wir setzen $U := \mathrm{Lin}(x_1) \leq V$. 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 A11.26

$$\chi_{f_{V/U}} = (t - \lambda_2) \cdots (t - \lambda_n),$$

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 $B'' = \left(\overline{x_2}, \dots, \overline{x_n}\right)$ von V/U, so daß $M_{B''}^{B''}\left(f_{V/U}\right)$ eine obere Dreiecksmatrix ist. Dann ist aber $B = (x_1, \dots, x_n)$ eine

Basis von V und mit $B' = (x_1)$ gilt wegen Aufgabe A6.20

$$M_B^B(f) = \left(\begin{array}{c|c} M_{B'}^{B'}(f_U) & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right) = \left(\begin{array}{c|c} \lambda_1 & * \\ \hline 0 & M_{B''}^{B''}(f_{V/U}) \end{array}\right).$$

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. □

Bemerkung A12.4

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 A12.5

a. Die Drehmatrix

$$A_\alpha = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

hat das charakteristische Polynom $\chi_{A_\alpha} = t^2 - 2\cos(\alpha)t + 1 = (t-\lambda)\cdot\left(t-\overline{\lambda}\right)$ 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

$$\begin{pmatrix} \lambda & 0 \\ 0 & \overline{\lambda} \end{pmatrix}.$$

Ist α kein ganzzahliges Vielfaches von π, so besitzt A_α zwei verschiedene Eigenwerte, so daß zugehörige Eigenvektoren nach Aufgabe A11.34 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 $A_\alpha = \mathbb{1}_2$ oder $A_\alpha = -\mathbb{1}_2$ und hat bereits Diagonalgestalt.

b. Die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C})$$

ist hingegen auch über $\mathbb{C}$ nicht diagonalisierbar. Denn, gäbe es eine Matrix $T \in \mathrm{Gl}_2(\mathbb{C})$ mit

$$T^{-1} \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{C}),$$

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:

$$0 = \operatorname{rang} \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} = \operatorname{rang}\left(T^{-1} \circ A \circ T\right) = \operatorname{rang}(A) = 1,$$

da $T \in \mathrm{Gl}_2(\mathbb{C})$. Dies ist jedoch ein Widerspruch.

B) Diagonalblockmatrizen

Definition A12.6 (Diagonalblockmatrizen)

Wir werden im Folgenden sehr häufig mit Blockmatrizen der folgenden Form arbeiten:

$$A = \left(\begin{array}{c|ccc} A_1 & 0 & \cdots & 0 \\ \hline 0 & \boxed{A_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{A_r} \end{array}\right) \in \mathrm{Mat}_n(K),$$

wobei $A_i \in \mathrm{Mat}_{n_i}(K)$, $i = 1, \ldots, r$ mit $n = n_1 + \ldots + n_r$. Es empfiehlt sich deshalb, eine Kurzschreibweise für solche *Diagonalblockmatrizen* einzuführen. Wir schreiben kurz:

$$A = A_1 \oplus \ldots \oplus A_r = \bigoplus_{i=1}^{r} A_i.$$

Bemerkung A12.7 (Diagonalblockmatrizen)

a. 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!

b. Mit Hilfe dieser Notation gilt beispielsweise, daß eine Matrix A genau dann diagonalisierbar ist, wenn es Körperelemente $\lambda_1, \ldots, \lambda_r \in K$ und positive natürliche Zahlen $n_1, \ldots, n_r \in \mathbb{N}$ gibt sowie eine invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ 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

$$A = \left(\begin{array}{c|ccc} A_1 & * & \cdots & * \\ \hline 0 & \boxed{A_2} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{A_r} \end{array}\right) \in \mathrm{Mat}_n(K),$$

gilt, daß

$$p(A) = \left(\begin{array}{c|ccc} p(A_1) & * & \cdots & * \\ \hline 0 & \boxed{p(A_2)} & & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & \boxed{p(A_r)} \end{array}\right) \in \mathrm{Mat}_n(K),$$

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 $\dim_K \big(\mathrm{Mat}_n(K)\big) = n^2$ gilt, sind die $n^2 + 1$ Matrizen

$$\mathbb{1}_n = A^0, A^1, A^2, \dots, A^{n^2}$$

in $\mathrm{Mat}_n(K)$ linear abhängig. D. h. es existieren $\lambda_0, \dots, \lambda_{n^2} \in K$, nicht alle null, mit

$$\lambda_0 A^0 + \lambda_1 A^1 + \dots + \lambda_{n^2} A^{n^2} = 0 \in \mathrm{Mat}_n(K).$$

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 A12.8 (Cayley-Hamilton)
Für $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$ gilt $\chi_f(f) = 0$ und $\chi_A(A) = 0$.

Beweis: Da für eine Basis D von V die Abbildung $M_D^D : \mathrm{End}_K(V) \longrightarrow \mathrm{Mat}_n(K)$ ein K-Algebrenhomomorphismus ist, gilt

$$M_D^D\big(\chi_f(f)\big) = \chi_f\big(M_D^D(f)\big).$$

Dann gilt aber $\chi_f(f) = 0$ genau dann, wenn

$$0 = M_D^D\big(\chi_f(f)\big) = \chi_f\big(M_D^D(f)\big) = \chi_{M_D^D(f)}\big(M_D^D(f)\big).$$

Es reicht deshalb, die Aussage für Matrizen zu beweisen.

Betrachte dazu die Matrix

$$B_t := t \cdot \mathbb{1}_n - A \in \mathrm{Mat}_n(K[t])$$

sowie die Adjunkte $B_t^\# = (p_{ij}) \in \mathrm{Mat}_n(K[t])$ von B_t, die auch *Busadjunkte* von A genannt wird. Nach dem Satz über die Adjunkte A9.29 in $\mathrm{Mat}_n(K[t])$ gilt die Adjunktengleichung

$$B_t \circ B_t^\# = (t\mathbb{1}_n - A) \circ (t\mathbb{1}_n - A)^\# = \det(t\mathbb{1}_n - A) \cdot \mathbb{1}_n = \chi_A \cdot \mathbb{1}_n. \tag{214}$$

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 $B_0, \ldots, B_{n-1} \in \mathrm{Mat}_n(K)$ mit

$$B_t^\# = B_{n-1}t^{n-1} + \ldots + B_1 t + B_0.$$

Ist $\chi_A = t^n + \alpha_{n-1}t^{n-1} + \ldots + \alpha_0$, so folgt aus der Adjunktengleichung (214)

$$(\mathbb{1}_n t - A) \circ \big(B_{n-1}t^{n-1} + \ldots + B_1 t + B_0\big) \stackrel{!}{=} \mathbb{1}_n t^n + \alpha_{n-1}\mathbb{1}_n t^{n-1} + \ldots + \alpha_0 \mathbb{1}_n \tag{215}$$

durch Koeffizientenvergleich für die t^i, $i = 0, \ldots, n$:

$$\begin{aligned} B_{n-1} &= \mathbb{1}_n \\ -AB_{n-1} + B_{n-2} &= \alpha_{n-1}\mathbb{1}_n \\ -AB_{n-2} + B_{n-3} &= \alpha_{n-2}\mathbb{1}_n \\ &\vdots \\ -AB_1 + B_0 &= \alpha_1 \mathbb{1}_n \\ -AB_0 &= \alpha_0 \mathbb{1}_n \end{aligned} \tag{216}$$

Multipliziert man die i-te Zeile in (216) mit A^{n-i+1} und summiert die beiden Seiten auf, so erhält man die Behauptung:

$$\begin{array}{rcl} A^n B_{n-1} & = & A^n \\ -A^n B_{n-1} + A^{n-1} B_{n-2} & = & \alpha_{n-1} A^{n-1} \\ -A^{n-1} B_{n-2} + A^{n-2} B_{n-3} & = & \alpha_{n-2} A^{n-2} \\ \vdots & & \\ -A^2 B_1 + A B_0 & = & \alpha_1 A \\ -A B_0 & = & \alpha_0 \mathbb{1}_n \\ \hline 0 & = & \chi_A(A). \end{array}$$

□

Bemerkung A12.9

a. Man beachte, daß der folgende *offensichtliche* Beweis für $\chi_A(A) = 0$, nämlich

$$\chi_A(A) = \det(A * \mathbb{1}_n - A) = \det(0) = 0$$

falsch ist, da "$*$" beim Einsetzen von A in $\det(t\mathbb{1}_n - A) \in K[t]$ 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 $\chi_A = t^n + \alpha_{n-1} t^{n-1} + \ldots + \alpha_0$, so läßt sich daraus mittels (215) und der Rekursionsformel (216) die Busadjunkte

$$(t\mathbb{1}_n - A)^{\#} = B_{n-1} t^{n-1} + \ldots + B_1 t + B_0$$

von A bestimmen. Für die B_{n-k}, $k = 1, \ldots, n$, gilt dabei explizit:

$$B_{n-k} = A^{k-1} + \alpha_{n-1} A^{k-2} + \ldots + \alpha_{n-k+1} A^0,$$

und wegen $\alpha_0 = (-1)^n \cdot \det(A)$ erhalten wir die Adjunkte von A als

$$A^{\#} = (-1)^{n+1} \cdot B_0 = (-1)^{n+1} \cdot \left(A^{n-1} + \alpha_{n-1} A^{n-2} + \ldots + \alpha_1 A^0\right).$$

Diese Formel zur Berechnung der Adjunkten von A ist weit effizienter, als die Determinanten sämtlicher Streichungsmatrizen zu berechnen.

D) Das Minimalpolynom

Satz A12.10 (Das Minimalpolynom)
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

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 A10.13, da $\mathrm{End}_K(V)$ sowie $\mathrm{Mat}_n(K)$ 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.

Zum besseren Verständnis skizzieren wir hier die wesentliche Beweisidee für μ_f. Wir betrachten die Menge
$$I := \{p \in K[t] \mid p(f) = 0\}$$
und beachten, daß für zwei Polynome p und q in I sowie ein beliebiges Polynom r stets auch $p + q \in I$ und $r \cdot p \in I$ gilt, wegen
$$(p + q)(f) = p(f) + q(f) = 0 + 0 = 0$$
und
$$(r \cdot p)(f) = r(f) \cdot p(f) = r(f) \cdot 0 = 0.$$
Aufgrund des Satzes von Cayley-Hamilton enthält I ein Nicht-Null-Polynom und wir können deshalb in I ein Nicht-Null-Polynom $0 \neq g \in I$ von minimalem Grad wählen. Normieren wir das Polynom g, indem wir durch seinen Leitkoeffizienten teilen, so erhalten wir ein Polynom $0 \neq \mu_f \in I$ vom selben minimalen Grad. Jedes Vielfache $\mu_f \cdot q$ mit $q \in K[t]$ liegt ebenfalls in I, was die Inklusion
$$I = \{p \in K[t] \mid p(f) = 0\} \supseteq \{\mu_f \cdot q \mid q \in K[t]\}$$
zeigt. Sei nun umgekehrt $p \in I$ gegeben, so können wir p mit Polynomdivision durch μ_f mit Rest teilen
$$p = q \cdot \mu_f + r,$$
wobei der Grad des Restes r echt kleiner ist als der von μ_f. Wegen
$$r = p - q \cdot \mu_f \in I$$
und wegen der Wahl von μ_f mit minimalem Grad, muß dann $r = 0$ gelten und $p = q \cdot \mu_f$ ist ein polynomielles Vielfaches von q, was die umgekehrte Inklusion zeigt. □

Bemerkung A12.11 (Minimalpolynome konjugierter Matrizen)

a. Sei $f \in \mathrm{End}_K(V)$, B eine Basis von V und $p \in K[t]$, dann gilt

$$M_B^B\big(p(f)\big) = p\big(M_B^B(f)\big),$$

da M_B^B ein K-Algebrenhomomorphismus ist.
Insbesondere gilt daher $p(f) = 0$ genau dann, wenn $p\big(M_B^B(f)\big) = 0$, 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 A11.27 und Satz A12.10 gilt für konjugierte Matrizen $A, B \in \mathrm{Mat}_n(K)$

$$\mu_A \cdot K[t] = \{p \in K[t] \mid p(A) = 0\} = \{p \in K[t] \mid p(B) = 0\} = \mu_B \cdot K[t].$$

Korollar A12.12 (Primfaktorzerlegung von χ und μ)
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Hat das charakteristische Polynom von f die Primfaktorzerlegung

$$\chi_f = p_1^{n_1} \cdot \ldots \cdot p_r^{n_r},$$

so hat das Minimalpolynom von f eine Primfaktorzerlegung der Form

$$\mu_f = p_1^{m_1} \cdot \ldots \cdot p_r^{m_r}$$

mit $1 \le m_i \le n_i$ für $i = 1, \ldots, r$.

b. Hat das charakteristische Polynom von A die Primfaktorzerlegung

$$\chi_A = p_1^{n_1} \cdot \ldots \cdot p_r^{n_r},$$

so hat das Minimalpolynom von A eine Primfaktorzerlegung der Form

$$\mu_A = p_1^{m_1} \cdot \ldots \cdot p_r^{m_r}$$

mit $1 \le m_i \le n_i$ für $i = 1, \ldots, r$.

Beweis: Der Satz von Cayley-Hamilton besagt, daß $\chi_A \in \mathrm{Ker}(\phi_A) = \mu_A K[t]$ gilt. Also gibt es ein $h \in K[t]$ mit $p_1^{n_1} \cdot \ldots \cdot p_r^{n_r} = \chi_A = \mu_A \cdot h$.

Daraus folgt $\mu_A = p_1^{m_1} \cdots p_r^{m_r}$ für $0 \le m_i \le n_i$, $i = 1, \ldots, r$, geeignet. Wir müssen zeigen, daß jedes p_i in μ_A auch vorkommt, d. h., daß $m_i \ge 1$ für alle $i = 1, \ldots, r$.

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 A10.22 Polynome $p, q \in K[t]$ mit

$$1 = p \cdot \mu_A + q \cdot p_i.$$

Wir führen die Annahme zum Widerspruch, indem wir zum algebraischen Abschluß $\overline{K}$ von K übergehen. Da p_i vom Grad $\deg(p_i) \geq 1$ ist, besitzt es eine Nullstelle $\lambda \in \overline{K}$. 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 $A \in \mathrm{Mat}_n\left(\overline{K}\right)$. Sei nun $0 \neq x \in \overline{K}^n$ ein Eigenvektor von A zum Eigenwert λ. Dann gilt für das Polynom $\mu_A = \sum_{i=0}^{m} a_i t^i \in K[t] \subseteq \overline{K}[t]$

$$\mu_A(A)x = \sum_{i=0}^{m} a_i\left(A^i x\right) = \sum_{i=0}^{m} a_i\left(\lambda^i x\right) = \mu_A(\lambda) \cdot x \neq 0,$$

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 $M_B^B(f)$. □

Korollar A12.13 (Die Eigenwerte sind die Nullstellen des Minimalpolynoms.)
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Die Eigenwerte von f sind genau die Nullstellen des Minimalpolynoms μ_f, und χ_f zerfällt genau dann über K in Linearfaktoren, wenn μ_f zerfällt.

b. Die Eigenwerte von A sind genau die Nullstellen des Minimalpolynoms μ_A, und χ_A zerfällt genau dann über K in Linearfaktoren, wenn μ_A zerfällt.

Beweis: Genau dann ist $\lambda \in K$ ein Eigenwert von f, wenn $t - \lambda$ ein Primfaktor von χ_f ist, und dies ist wegen Korollar A12.12 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 A12.12 zerfällt χ_f zudem genau dann über K in Linearfaktoren, wenn μ_f in Linearfaktoren zerfällt. Die entsprechenden Aussagen für eine Matrix A folgen aus a. mit $f = f_A$.

□

Beispiel A12.14

a. Ist $A = \lambda \mathbb{1}_n \in \mathrm{Mat}_n(K)$ 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 $\lambda \in K$ und

$$J := J_n(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{pmatrix} \in \mathrm{Mat}_n(K),$$

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 A12.12 ist mithin $\mu_J = (t-\lambda)^m$ für ein $1 \le m \le n$. Dabei ist m die kleinste natürliche Zahl mit $(J - \lambda \mathbb{1}_n)^m = 0$. Nun ist aber

$$J - \lambda \mathbb{1}_n = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & 0 \end{pmatrix} =: N$$

und man sieht mittels einer einfachen Induktion, daß $N^k \neq 0$ für $k < n$, aber $N^n = 0$ (vgl. Aufgabe A2.27). Also gilt

$$\mu_J = (t-\lambda)^n.$$

c. Ist $A = A_1 \oplus \ldots \oplus A_r \in \mathrm{Mat}_n(K)$ eine Diagonalblockmatrix mit $A_i \in \mathrm{Mat}_{n_i}(K)$, so folgt aus der Definition des charakteristischen Polynoms unmittelbar (vgl. Aufgabe A11.26)

$$\chi_A = \prod_{i=1}^{r} \chi_{A_i}.$$

Die entsprechende Formel für das Minimalpolynom gilt nicht. Sei etwa $A_1 = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right) \in \mathrm{Mat}_2(K)$ und $A_2 = (1) \in \mathrm{Mat}_1(K)$, 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 $\mu_{A_1}, \ldots, \mu_{A_r}$ im Sinne der Vorlesung Algebraische Strukturen ist. Darauf wollen wir hier nicht näher eingehen.

Bemerkung A12.15 (Berechnung des Minimalpolynoms)

Zur praktischen Berechnung des Minimalpolynoms von $A \in \mathrm{Mat}_n(K)$ kann man wie folgt vorgehen. Aufgrund des Satzes von Cayley-Hamilton wissen wir, daß die Matrizen $A^0, \ldots, A^n$ 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 $x_0, \ldots, x_m$ linear abhängig sind, und wir suchen ferner geeignete $\beta_0, \ldots, \beta_{m-1}$ mit

$$x_m + \beta_{m-1}x_{m-1} + \ldots + \beta_0 x_0 = 0.$$

Dies ist dann gleichbedeutend damit, daß

$$t^m + \beta_{m-1}t^{m-1} + \ldots + \beta_0 \in K[t]$$

das gesuchte Minimalpolynom von A ist.

Bezeichne $X = (x_0 \ldots x_n) \in \mathrm{Mat}\left(n^2 \times (n+1), K\right)$ die Matrix, deren Spalten $x_0, \ldots, x_n$ sind, dann suchen wir eine Lösung des linearen Gleichungssystems

$$X \cdot \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_n \end{pmatrix} = 0 \in K^{n^2} \tag{217}$$

mit $\beta_{m+1} = \ldots = \beta_n = 0$ und $\beta_m = 1$ und so, daß m minimal mit dieser Eigenschaft ist. Da $(x_0, \ldots, x_{m-1})$ nach Definition von m linear unabhängig, $(x_0, \ldots, x_m)$ aber linear abhängig ist, bedeutet dies, daß in einer ZSF von X die Zahlen $1, \ldots, m$ Pivotindizes sind, während $m+1$ kein Pivotindex mehr ist.

Berechnet man eine Basis des Lösungsraums von (217) mittels des Algorithmus A7.9, 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 $A \in \mathrm{Mat}_n(K)$.

Algorithmus A12.16 (Algorithmus zur Berechnung des Minimalpolynoms)

INPUT: $A \in \mathrm{Mat}_n(K)$

OUTPUT: μ_A

1. **Schritt:** Falls A nicht quadratisch ist, gib 0 zurück.
2. **Schritt:** Bilde die Potenzen $A^0, \ldots, A^n$ und schreibe die Matrizen in Form von Spaltenvektoren der Länge n^2 in eine Matrix $X \in \mathrm{Mat}\left(n^2 \times (n+1), K\right)$.
3. **Schritt:** Berechne eine Basis von $\mathrm{Lös}(X, 0)$.
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 A5.32)

$$\operatorname{Ker}\left(f - \lambda \operatorname{id}_V\right) \subseteq \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^2\right) \subseteq \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^3\right) \subseteq \ldots \subseteq V$$

und

$$\operatorname{Lös}\left(A - \lambda \mathbb{1}_n, 0\right) \subseteq \operatorname{Lös}\left((A - \lambda \mathbb{1}_n)^2, 0\right) \subseteq \operatorname{Lös}\left((A - \lambda \mathbb{1}_n)^3, 0\right) \subseteq \ldots \subseteq V$$

Die Vereinigung all dieser Unterräume ist offenbar wieder ein Unterraum und führt zu folgender Definition.

Definition A12.17 (Hauptraum)
Es sei $f \in \operatorname{End}_K(V)$, $A \in \operatorname{Mat}_n(K)$ und $\lambda \in K$. Dann heißen

$$\operatorname{Hau}(f, \lambda) = \bigcup_{k \in \mathbb{N}} \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^k\right) \quad \text{und} \quad \operatorname{Hau}(A, \lambda) = \bigcup_{k \in \mathbb{N}} \operatorname{Lös}\left((A - \lambda \mathbb{1}_n)^k, 0\right)$$

der *Hauptraum* oder *verallgemeinerte Eigenraum* von f bzw. A zu λ.

Lemma A12.18 (Nilpotenzindex und Hauptraum)
Es sei $\lambda \in K$ gegeben.

a. Es gibt eine natürliche Zahl $0 \leq m \leq n$ mit

$$\operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^0\right) \subsetneqq \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^1\right) \subsetneqq \ldots \subsetneqq \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^m\right)$$

und für $k > m$

$$\operatorname{Hau}(f, \lambda) = \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^m\right) = \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^k\right).$$

Die Zahl m heißt *Nilpotenzindex* von $f - \lambda \operatorname{id}_V$ und erfüllt $m \leq \operatorname{mult}(\mu_f, \lambda)$.

b. Für jedes $k \in \mathbb{N}$ ist $\operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^k\right)$ ein f-invarianter Unterraum von V. Insbesondere sind also Eigenräume und Haupträume von f auch f-invariant.

Die entsprechenden Aussagen für eine Matrix $A \in \operatorname{Mat}_n(K)$ 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 A5.32 wissen wir, daß es eine natürliche Zahl $0 \leq m \leq n$ gibt mit

$$\operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^0\right) \subsetneqq \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^1\right) \subsetneqq \ldots \subsetneqq \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^m\right)$$

und

$$\operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^m\right) = \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^k\right)$$

für $k > m$. Aus der Definition des Hauptraumes folgt dann unmittelbar

$$\operatorname{Hau}(f, \lambda) = \bigcup_{k \in \mathbb{N}} \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^k\right) = \operatorname{Ker}\left((f - \lambda \operatorname{id}_V)^m\right).$$

Es bleibt zu zeigen, daß

$$m \leq \operatorname{mult}(\mu_f, \lambda).$$

Angenommen, $l := \operatorname{mult}(\mu_f, \lambda) < m$. Dann gibt es ein $x \in \operatorname{Ker}\big((f - \lambda \operatorname{id}_V)^m\big)$ mit $y := (f - \lambda \operatorname{id}_V)^l(x) \neq 0$. 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 A10.22 gibt es Polynome $p, q \in K[t]$ mit

$$p \cdot (t - \lambda)^{m-l} + q \cdot h = 1.$$

Es folgt

$$\big(p(f) \circ (f - \lambda \operatorname{id}_V)^{m-l} + q(f) \circ h(f)\big)(y) = \operatorname{id}_V(y) = y. \tag{218}$$

Andererseits gilt aber $(f - \lambda \operatorname{id}_V)^{m-l}(y) = (f - \lambda \operatorname{id}_V)^m(x) = 0$ sowie

$$h(f)(y) = \big(h(f) \circ (f - \lambda \operatorname{id}_V)^l\big)(x) = \mu_f(f)(x) = 0.$$

Aus (218) 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 $x \in \operatorname{Ker}\big((f - \lambda \operatorname{id}_V)^k\big)$

$$(f - \lambda \operatorname{id}_V)^k\big(f(x)\big) = f\big((f - \lambda \operatorname{id}_V)^k(x)\big) = f(0) = 0,$$

woraus die Behauptung folgt.

□

Satz A12.19 (Hauptraumzerlegung)
Es sei $f \in \operatorname{End}_K(V)$ so, daß χ_f über K in Linearfaktoren zerfällt, d. h. es gibt paarweise verschiedene $\lambda_1, \ldots, \lambda_r \in K$ 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:

a. $V = \operatorname{Hau}(f, \lambda_1) \oplus \ldots \oplus \operatorname{Hau}(f, \lambda_r)$,

b. $n_i = \operatorname{mult}(\chi_f, \lambda_i) = \dim_K \big(\operatorname{Hau}(f, \lambda_i)\big)$ und

c. $m_i = \operatorname{mult}(\mu_f, \lambda_i)$ 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 $V_i := \mathrm{Hau}(f, \lambda_i)$, 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: $V = V_1 + \ldots + V_r$ und $B = B_1 \cup \ldots \cup B_r$ erzeugt V.

Offenbar gibt es keinen Primfaktor, den alle $q_1, \ldots, q_r$ gemeinsam haben, so daß aus der Bézout-Identität A10.22 die Existenz von Polynomen $p_1, \ldots, p_r \in K[t]$ folgt mit

$$p_1 q_1 + \ldots + p_r q_r = 1.$$

Setzt man $Q_i := q_i p_i$, dann folgt

$$Q_1(f) + \ldots + Q_r(f) = \mathrm{id}_V .$$

Sei $x \in V$ beliebig. Wegen $(f - \lambda_i \,\mathrm{id}_V)^{m_i} \circ Q_i(f) = p_i(f) \circ \mu_f(f) = 0$ gilt

$$Q_i(f)(x) \in \mathrm{Ker}(f - \lambda_i \,\mathrm{id}_V)^{m_i} = V_i,$$

und somit

$$x = Q_1(f)(x) + \ldots + Q_r(f)(x) \in V_1 + \ldots + V_r.$$

Damit ist $V = V_1 + \ldots + V_r$ gezeigt, und insbesondere ist $B = B_1 \cup \ldots \cup B_r$ ein Erzeugendensystem von V, da die B_i die V_i erzeugen.

Zeige: $\dim_K(V_i) = |B_i| \leq n_i$.

Wir wissen aus Lemma A12.18, daß V_i ein f-invarianter Unterraum ist, und aus der Definition von $V_i = \mathrm{Ker}\left((f - \lambda \,\mathrm{id}_V)^{m_i}\right)$ folgt unmittelbar

$$(f_{V_i} - \lambda_i \,\mathrm{id}_{V_i})^{m_i} = 0.$$

Also muß nach Satz A12.10 $(t-\lambda_i)^{m_i}$ 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 A12.12 gilt dann aber auch

$$\chi_{f_{V_i}} = (t-\lambda_i)^k$$

und dabei muß dann $k = \dim_K(V_i) = |B_i|$ gelten. Aus Aufgabe A11.26 wissen wir, daß χ_f ein Vielfaches von $\chi_{f_{V_i}}$ ist, und mithin muß $k \leq n_i$ gelten.

Zeige: $V = V_1 \oplus \ldots \oplus V_r$ und $n_i = \dim_K(V_i)$ (d.h. Teil a. und b.).

Da B ein Erzeugendensystem von V ist und da $|B_i| \leq n_i$ gilt, erhalten wir

$$n \leq |B| \leq |B_1| + \ldots + |B_r| \leq n_1 + \ldots + n_r = \deg(\chi_f) = n.$$

Dies zeigt, daß alle beteiligten Ungleichheitszeichen in der Tat Gleichheitszeichen waren, d.h. $\dim_K(V_i) = |B_i| = n_i$ für $i = 1, \ldots, n$. Aber als Erzeugendensystem mit genau $n = \dim_K(V)$ Elementen ist B dann eine Basis von V und aus Aufgabe A4.39 folgt dann, daß V die direkte Summe der V_i ist.

Zeige: m_i ist der Nilpotenzindex von $f - \lambda_i \mathrm{id}_V$ (d.h. Teil c.).

Für den Nilpotenzindex m von $f - \lambda_i \mathrm{id}_V$ gilt nach Lemma A12.18 $m_i \geq m$. Angenommen, $m_i > m$. Sei nun $x = x_1 + \ldots + x_r \in V$ beliebig mit $x_j \in V_j$ für $j = 1, \ldots, r$. Dann gilt $(f - \lambda_i \mathrm{id}_V)^m(x_i) = 0$ und für $j \neq i$ gilt $q_i(f)(x_j) = 0$, also folgt für $p := (t - \lambda_i)^m q_i \in K[t]$

$$p(f)(x) = 0.$$

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 A12.10.

Die entsprechende Aussage für eine Matrix A läßt sich unmittelbar auf die Aussage für f_A zurückführen. □

Bemerkung A12.20 (Hauptraumzerlegung)

Man kann im Beweis von Satz A12.19 auch ohne Rückgriff auf Aufgabe A4.39 zeigen, daß die Summe $V_1 + \ldots + V_r$ eine direkte Summe ist:

Mit der Notation aus dem Beweis gilt für $x \in V_j$ und $i \neq j$

$$Q_i(f)(x) = p_i(f) \circ \prod_{k \neq i,j} (f - \lambda_k \mathrm{id}_V)^{m_k} \circ (f - \lambda_j \mathrm{id}_V)^{m_j}(x) = 0$$

und für $x \in V_i$ gilt deshalb

$$Q_i(f)(x) = \sum_{j=1}^{r} Q_j(f)(x) = \mathrm{id}_V(x) = x.$$

Sind nun also $x = x_1 + \ldots + x_r = y_1 + \ldots + y_r$ zwei Darstellung des Vektors x mit $x_i, y_i \in V_i$, $i = 1, \ldots, r$, dann gilt

$$0 = Q_i(f)(x - x) = Q_i(f)\big((x_1 - y_1) + \ldots + (x_r - y_r)\big) = Q_i(f)(x_i - y_i) = x_i - y_i,$$

d.h. $x_i = y_i$ für alle $i = 1, \ldots, r$. 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 $\mathrm{Ker}\big(Q_i(f)\big) = \bigoplus_{j \neq i} V_j$ ist - vgl. Aufgabe A3.50.

Aus Satz A12.19 Teil b. und c. folgt, da die Haupträume von f f-invariant sind, unmittelbar das folgende Korollar.

Korollar A12.21

Sei f wie in Satz A12.19, dann gilt

$$\chi_{f_{\mathrm{Hau}(f,\lambda_i)}} = (t - \lambda_i)^{n_i}$$

und

$$\mu_{f_{\mathrm{Hau}(f,\lambda_i)}} = (t - \lambda_i)^{m_i}.$$

F) Diagonalisierbarkeit

Satz A12.22 (Diagonalisierbarkeit von Endomorphismen)
Für einen Endomorphismus $f \in \mathrm{End}_K(V)$ sind die folgenden Aussagen äquivalent:

a. f ist diagonalisierbar.

b. V hat eine Basis aus Eigenvektoren von f.

c. Sind $\lambda_1, \ldots, \lambda_r$ die paarweise verschiedenen Eigenwerte von f in K, dann gilt
$$V = \bigoplus_{i=1}^{r} \mathrm{Eig}(f, \lambda_i).$$

d. Das charakteristische Polynom von f zerfällt über K in Linearfaktoren und für jeden Eigenwert λ stimmen algebraische und geometrische Vielfachheit überein.

e. Das Minimalpolynom von f zerfällt über K in paarweise verschiedene Linearfaktoren.

Beweis:
a. $\Rightarrow$ e.: 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 $p = (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_r) \in K[t]$, so ist
$$p\big(M_B^B(f)\big) = p(\lambda_1 \mathbb{1}_{n_1}) \oplus \ldots \oplus p(\lambda_r \mathbb{1}_{n_r})$$
wegen Bemerkung A12.7 eine Diagonalblockmatrix, 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
$$\mu_f = (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_r),$$
so zerfällt wegen Korollar A12.13 auch χ_f in Linearfaktoren
$$\chi_f = (t - \lambda_1)^{n_1} \cdot \ldots \cdot (t - \lambda_r)^{n_r}.$$
Aus dem Satz zur Hauptraumzerlegung A12.19 folgt zudem $\mathrm{Hau}(f, \lambda_i) = \mathrm{Eig}(f, \lambda_i)$, da der Nilpotenzindex von $f - \lambda_i \,\mathrm{id}_V$ eins ist, und
$$\mathrm{mult}(\chi_f, \lambda_i) = n_i = \dim_K \mathrm{Hau}(f, \lambda_i) = \dim_K \mathrm{Eig}(f, \lambda_i),$$

d.h. die geometrische und die algebraische Vielfachheit jedes Eigenwertes stimmen überein.

d. ⇒ c.: Das charakteristische Polynom habe die Primfaktorzerlegung

$$\chi_f = (t-\lambda_1)^{n_1} \cdot \ldots \cdot (t-\lambda_r)^{n_r}.$$

Aus dem Satz über die Hauptraumzerlegung und der Voraussetzung folgt dann

$$\dim_K \operatorname{Hau}(f,\lambda_i) = n_i = \operatorname{mult}(\chi_f,\lambda_i) = \dim_K \operatorname{Eig}(f,\lambda_i),$$

und da stets $\operatorname{Eig}(f,\lambda_i) \subseteq \operatorname{Hau}(f,\lambda_i)$ gilt, folgt

$$\operatorname{Eig}(f,\lambda_i) = \operatorname{Hau}(f,\lambda_i).$$

Aus dem Satz über die Hauptraumzerlegung folgt dann aber wiederum

$$V = \operatorname{Hau}(f,\lambda_1) \oplus \ldots \oplus \operatorname{Hau}(f,\lambda_r) = \operatorname{Eig}(f,\lambda_1) \oplus \ldots \oplus \operatorname{Eig}(f,\lambda_r).$$

c. ⇒ b.: Es sei B_i eine Basis von $\operatorname{Eig}(f,\lambda_i)$, dann ist nach Aufgabe A4.39 $B = B_1 \cup \ldots \cup B_r$ eine Basis von V, die aus Eigenvektoren besteht.

b. ⇒ a.: Ist $B = (x_1,\ldots,x_n)$ 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 $\lambda_1,\ldots,\lambda_n$. □

Korollar A12.23 (Diagonalisierbarkeit von Matrizen)
Für eine quadratische Matrix $A \in \operatorname{Mat}_n(K)$ sind die folgenden Aussagen äquivalent:

a. A ist diagonalisierbar.

b. K^n hat eine Basis aus Eigenvektoren von A.

c. Sind $\lambda_1,\ldots,\lambda_r$ die paarweise verschiedenen Eigenwerte von A, dann gilt
$$K^n = \bigoplus_{i=1}^{r} \operatorname{Eig}(A,\lambda_i).$$

d. Das charakteristische Polynom von A zerfällt über K in Linearfaktoren und für jeden Eigenwert λ stimmen algebraische und geometrische Vielfachheit überein.

e. Das Minimalpolynom von A zerfällt über K in paarweise verschiedene Linearfaktoren.

Insbesondere, genau dann ist $T \in \operatorname{Gl}_n(K)$ so, daß $T^{-1} \circ A \circ T$ eine Diagonalmatrix ist, wenn die Spalten von T eine Basis des K^n aus Eigenvektoren von A sind.

Beweis: Wende Satz A12.22 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 A12.24 (Diagonalisierbarkeit)
Es sei $f \in \mathrm{End}_K(V)$ und $A \in \mathrm{Mat}_n(K)$.

a. Hat f genau n paarweise verschiedene Eigenwerte, so ist f diagonalisierbar.

b. Hat A genau n paarweise verschiedene Eigenwerte, so ist A diagonalisierbar.

Beweis: Hat f genau n paarweise verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_n \in K$, so muß

$$\chi_f = \mu_f = (t - \lambda_1) \cdot \ldots \cdot (t - \lambda_n)$$

gelten. D.h. μ_f zerfällt in paarweise verschiedene Linearfaktoren und f ist diagonalisierbar. Der Beweis für A geht analog. □

Aus Korollar A12.23 können wir ein Verfahren ableiten, das es uns erlaubt, eine Matrix zu diagonalisieren und die Transformationsmatrix T zu berechnen.

Algorithmus A12.25 (Algorithmus zur Diagonalisierung)
INPUT: $A \in \mathrm{Mat}_n(K)$.
OUTPUT: 0, falls A über K nicht diagonalisierbar ist,
$1, D, T$, falls A diagonaliserbar 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 $\lambda_1, \ldots, \lambda_r$ sowie ihre algebraischen Vielfachheiten $n_1, \ldots, n_r$.

3. Schritt: Bestimme für jeden Eigenwert λ_i eine Basis des Eigenraums $\mathrm{Eig}(A, \lambda_i)$ als $\mathrm{Lös}(A - \lambda_i \mathbb{1}_n, 0)$ - vgl. Algorithmus A7.9 - sowie seine Dimension - vgl. Algorithmus A4.16 -, 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 $\lambda_1, \ldots, \lambda_r$ entsprechend ihren algebraischen Vielfachheiten in der Diagonalen einer Nullmatrix einträgt.

Beispiel A12.26

Gegeben sei die Matrix

$$A = \begin{pmatrix} 2 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 1 & -1 & -1 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q}).$$

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 $\mathrm{mult}(\chi_A, 1) = \mathrm{mult}(\chi_A, 2) = 2$.

Als nächstes berechnen wir den Eigenraum $\mathrm{Lös}(2\mathbb{1}_4 - A, 0)$ zum Eigenwert $\lambda = 2$:

$$2\cdot\mathbb{1}_4 - A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 1 & 1 \end{pmatrix} \overset{\substack{I\leftrightarrow IV \\ I\mapsto -I \\ IV\mapsto IV-II \\ I\mapsto I+II}}{\longmapsto} \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{-1\text{'en einfügen}}{\longmapsto} \begin{pmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Mithin ist

$$\mathrm{Eig}(A, 2) = \mathrm{Lin}\left((-1, 0, -1, 0)^t, (-1, 0, 0, -1)^t\right)$$

und

$$\dim_{\mathbb{Q}} \mathrm{Eig}(A, 2) = 2 = \mathrm{mult}(\chi_A, 2).$$

Dann berechnen wir den Eigenraum $\mathrm{Lös}(\mathbb{1}_4 - A, 0)$ zum Eigenwert $\lambda = 1$:

$$1\cdot\mathbb{1}_4 - A = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ -1 & 1 & 1 & 0 \end{pmatrix} \overset{\substack{IV\mapsto IV-I+III \\ I\mapsto -I \\ III\mapsto -III}}{\longmapsto} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{-1\text{'en einfügen}}{\longmapsto} \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Mithin ist

$$\mathrm{Eig}(A, 1) = \mathrm{Lin}\left((-1, -1, 0, 0)^t, (0, 0, 0, -1)^t\right)$$

und

$$\dim_{\mathbb{Q}} \mathrm{Eig}(A, 1) = 2 = \mathrm{mult}(\chi_A, 1).$$

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 A12.27

Es sei $E = (E_{11}, E_{12}, E_{21}, E_{22})$ die kanonische Basis von $V = \mathrm{Mat}_2(K)$ und $T = E_{11} + E_{12} + E_{22} \in \mathrm{Gl}_2(K)$. Zeige, daß der Endomorphismus $f : V \to V : A \mapsto T \circ A \circ T^{-1}$ trigonalisierbar, aber nicht diagonalisierbar ist, und bestimme eine Basis B von V, so daß $M_B^B(f)$ eine obere Dreiecksmatrix ist.

Aufgabe A12.28

Zeige, ist $A \in \mathrm{Gl}_n(K)$, so gibt es ein Polynom $p \in K[t]$ mit $A^{-1} = p(A)$.

Aufgabe A12.29

Zeige, ist $1 \leq \dim_K(V) = n < \infty$, so sind für $\mathcal{A} \subseteq \mathrm{End}_K(V)$ die folgenden beiden Aussagen gleichwertig:

a. $\mathcal{A}$ 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.

b. 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.

A13 Die Jordansche Normalform

Eine Matrix $A \in \mathrm{Mat}_n(K)$, deren charakteristisches Polynom zerfällt, was über einem algebraisch abgeschlossenen Körper wie $\mathbb{C}$ stets der Fall ist, ist zu einer besonders einfachen Matrix konjugiert. Aus dieser lassen sich Invarianten von A einfach ablesen und diese Invarianten bestimmen die Matrix A bis auf Konjugation eindeutig.

Satz A13.1 (Jordansche Normalform eines Endomorphismus)
Es sei $f \in \mathrm{End}_K(V)$ ein Endomorphismus, dessen charakteristisches Polynom über K zerfällt, $\chi_f = (t-\lambda_1)^{n_1} \cdot \ldots \cdot (t-\lambda_r)^{n_r}$, und es sei $\mu_f = (t-\lambda_1)^{m_1} \cdot \ldots \cdot (t-\lambda_r)^{m_r}$. Dann gibt es für jedes $i = 1, \ldots, r$ und $1 \le j \le m_i$, je eine natürliche Zahl t_{ij} und es gibt eine Basis B so, daß

(1) $\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K \mathrm{Hau}(f, \lambda_i)$,

(2) $\sum_{j=1}^{m_i} t_{ij} = \dim_K \mathrm{Eig}(f, \lambda_i)$,

(3) $t_{im_i} \ge 1$ 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 A13.2 (Jordansche Normalform einer quadratischen Matrix)
Es sei $A \in \mathrm{Mat}_n(K)$ eine quadratische Matrix, deren charakteristisches Polynom über K zerfällt, $\chi_A = (t-\lambda_1)^{n_1} \cdot \ldots \cdot (t-\lambda_r)^{n_r}$, und es sei $\mu_A = (t-\lambda_1)^{m_1} \cdot \ldots \cdot (t-\lambda_r)^{m_r}$.
Dann gibt es für jedes $i = 1, \ldots, r$ und $1 \le j \le m_i$, je eine natürliche Zahl t_{ij} und es gibt ein invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ so, daß

(1) $\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K \mathrm{Hau}(A, \lambda_i)$,

(2) $\sum_{j=1}^{m_i} t_{ij} = \dim_K \mathrm{Eig}(A, \lambda_i)$,

(3) $t_{im_i} \ge 1$ 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 A13.1 mit $f = f_A$. □

A) Anmerkungen und Beispiele zur Jordanschen Normalform

Es scheint angebracht, den Satz zunächst etwas zu erläutern, um ihn verständlicher zu machen.

Bemerkung A13.3 (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 A12.14).
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 $\sum_{j=1}^{m_i} t_{ij} = \dim_K \operatorname{Eig}(A, \lambda_i)$ 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 A13.26, daß die Elementarteiler übereinstimmen, da für eine invertierbare Matrix $T \in \mathrm{Gl}_n(K)$ und ein $k \in \mathbb{N}$ gilt

$$\begin{aligned}\operatorname{rang}\left(\left(T^{-1} \circ A \circ T - \lambda \mathbb{1}_n\right)^k\right) &= \operatorname{rang}\left(T^{-1} \circ (A - \lambda \mathbb{1}_n)^k \circ T\right) \\ &= \operatorname{rang}\left((A - \lambda \mathbb{1}_n)^k\right).\end{aligned}$$

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 A13.4 (Jordansche Normalform)
Wir wollen nun in einigen einfachen Fällen die Jordansche Normalform bestimmen.

a. Die Matrix

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 5 & 6 & 7 \\ 0 & 0 & 8 & 9 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q})$$

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 A12.23 bereits, daß die

Spalten von T Eigenvektoren zu den vier Eigenwerten sein müssen. Wir könnten T also leicht berechnen.

b. Die Matrix

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q})$$

hat offenbar den Rang eins. Deshalb gilt für die geometrische Vielfachheit von 0 als Eigenwert von A

$$\dim_K \mathrm{Eig}(A, 0) = \dim_K \mathrm{Lös}(A, 0) = 4 - \mathrm{rang}(A) = 3. \tag{219}$$

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 A11.6 wissen wir aber, daß der zweithöchste Koeffizient des charakteristischen Polynoms das Negative der Spur der Matrix ist, d.h. $\lambda = \mathrm{Spur}(A) = 4$. Wir haben also

$$\chi_A = t^3 \cdot (t - 4).$$

Aus (219) 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.

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{F}_2).$$

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

$$\chi_A = t^3 \cdot \big(t - \mathrm{Spur}(A)\big).$$

Allerdings ist die Spur diesmal

$$\mathrm{Spur}(A) = 1 + 1 + 1 + 1 = 0 \in \mathbb{F}_2,$$

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

$$J_A = \left(\begin{array}{cc|cc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right).$$

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 $A \in \mathrm{Mat}_3(\mathbb{Q})$ 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 A12.10 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	$\left(\begin{array}{c\|cc} 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 0 \end{array}\right)$	$\left(\begin{array}{c\|cc} 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)$	$\left(\begin{array}{c\|cc} 0 & 0 & 0 \\ \hline 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right)$, $\left(\begin{array}{c\|cc} 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & 1 \end{array}\right)$	$\left(\begin{array}{cc\|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \end{array}\right)$	$\left(\begin{array}{cc\|c} 0 & 1 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 1 \end{array}\right)$

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 A12.23 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.

B) Nilpotente Endomorphismen und Young-Diagramme

Wir werden Satz A13.1 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.

Definition A13.5 (Nilpotent)
Wir nennen einen Endomorphismus $f \in \mathrm{End}_K(V)$ bzw. eine Matrix $A \in \mathrm{Mat}_n(K)$ *nilpotent*, wenn es ein $r \in \mathbb{N}$ gibt, so daß $f^r = 0$ bzw. $A^r = 0$. Offenbar gilt dann $\mu_f = t^m$ bzw. $\mu_A = t^m$ für ein $1 \leq m \leq r$.

Beispiel A13.6 (Ein nilpotentes Jordankästchen)
Sei $f \in \mathrm{End}_K(V)$ und sei $B = (x_1, \ldots, x_n)$ eine Basis von V, so daß

$$M_B^B(f) = J_n(0) := \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \ldots & \ldots & 0 & 0 \end{pmatrix}$$

ein Jordankästchen der Größe n zum Eigenwert 0 ist, dann folgt aus der Matrixdarstellung zunächst

$$f(x_{i+1}) = x_i$$

und damit

$$x_i = f^{n-i}(x_n)$$

für $i = 1, \ldots, n-1$. Das heißt, V ist ein zyklischer Unterraum mit seiner kanonischen Basis

$$B = \big(f^{n-1}(x_n), f^{n-2}(x_n), \ldots, f(x_n), x_n\big)$$

wie wir sie in Aufgabe A6.51 betrachtet haben. Man beachte auch, daß die Matrix $M_B^B(f)$ und damit der Endomorphismus f nilpotent mit Nilpotenzindex n ist (siehe Aufgabe A2.27).

Wir wollen im folgenden zeigen, daß die Jordansche Normalform eines nilpotenten Endomorphismus eine Blockdiagonalmatrix ist, deren Diagonalblöcke Jordankästchen der obigen Form sind. Das Beispiel sagt uns also, welche Gestalt der Anteil der Basis haben muß, der zu einem solchen Kästchen gehört!

Definition A13.7 (Partitionen)
Eine *Partition* der positiven natürlichen Zahl n ist ein Tupel $P = (k_1, \ldots, k_m)$ natürlicher Zahlen, so daß $k_1 \geq k_2 \geq \ldots \geq k_m \geq 1$ und $k_1 + k_2 + \ldots + k_m = n$.

Lemma A13.8 (Die duale Partition)
Ist $P = (k_1, \ldots, k_m)$ eine Partition von n, ist auch $P^* = (l_1, \ldots, l_s)$ mit $s = k_1$ und

$$l_i = |\{j \mid 1 \leq j \leq m, k_j \geq i\}|$$

eine Partition von n, die sogenannte *duale Partition* zu P.

Beweis: Man kann die Partition P durch ihr *Young-Diagramm* veranschaulichen. Dieses besteht aus n Kästen, die in r Reihen übereinander angeordnet sind, wobei in der i-ten Reihe genau k_i Kästchen sind (siehe Abbildung 9).

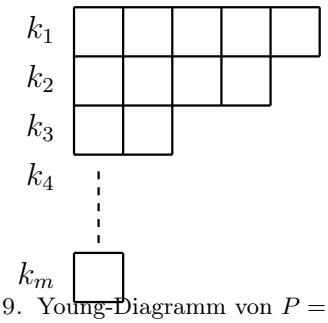


Abbildung 9. Young-Diagramm von $P = (k_1, \ldots, k_m)$

Die Anzahl an Kästchen in der i-ten Spalte ist dann genau l_i. Damit ist die Summe der l_i gerade n und $l_1 \geq l_2 \geq \ldots \geq l_s$. □

Beispiel A13.9
$P = (5, 4, 2, 2, 2)$ ist eine Partition von $n = 15$ mit dem folgenden Young-Diagramm (siehe Abbildung 10).

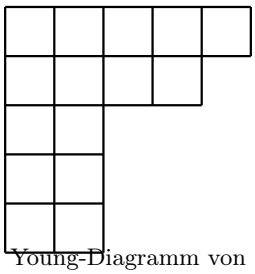


Abbildung 10. Young-Diagramm von $P = (5, 4, 2, 2, 2)$

Das Young-Diagramm der dualen Partition $P^* = (5, 5, 2, 2, 1)$ entsteht durch Spiegelung an der Winkelhalbierenden (siehe Abbildung 11).

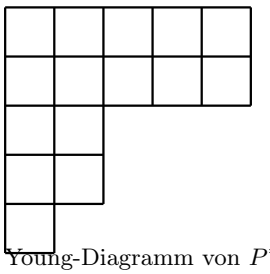


Abbildung 11. Young-Diagramm von $P^* = (5,5,2,2,1)$

Bemerkung A13.10 (Anzahl der Partitionen von n)
Die Funktion

$$\pi : \mathbb{Z}_{>0} \longrightarrow \mathbb{Z}_{>0},$$

die einer natürlichen Zahl n die Anzahl der Partitionen von n zuordnet, ist eine interessante zahlentheoretische Funktion. Wir wollen einige Werte von π zur Veranschaulichung ihrer Komplexität angeben:

n	1	2	3	4	5	6	7	8	9	10	100
$\pi(n)$	1	2	3	5	7	11	15	22	30	42	190569292

Für große n gilt

$$\pi(n) \approx \tfrac{1}{4\sqrt{3}n} e^{\pi\sqrt{\frac{2n}{3}}}.$$

Das folgende Lemma zusammen mit Bemerkung A13.3 besagt, daß $\pi(n)$ zugleich die Anzahl der Konjugationsklassen nilpotenter Matrizen der Größe $n \times n$ angibt.

C) Jordansche Normalform für nilpotente Endomorphismen

Lemma A13.11 (Jordansche NF nilpotenter Endomorphismen)
Sei $f \in \mathrm{End}_K(V)$ ein nilpotenter Endomorphismus mit $m_f = t^m$.

a. Setzen wir $U_i = \mathrm{Ker}(f^i)$, $i = 0, \ldots, m$, dann induziert f für $i = 2, \ldots, m$ eine injektive lineare Abbildung

$$f_i : U_i/U_{i-1} \longrightarrow U_{i-1}/U_{i-2} : \overline{x} \mapsto \overline{f(x)}.$$

Zudem ist $P = (k_1, \ldots, k_m)$ eine Partition von n mit

$$k_i = \dim_K(U_i/U_{i-1}) = \mathrm{rang}(f^{i-1}) - \mathrm{rang}(f^i),$$

die wir die *Jordan-Partition* des nilpotenten Endomorphismus nennen wollen.

b. Ist $P^* = (l_1, \ldots, l_s)$ die zu P duale Partition, dann gibt es eine Basis B von V, so daß

$$M_B^B(f) = J_{l_1}(0) \oplus J_{l_2}(0) \oplus \ldots \oplus J_{l_s}(0).$$

Die analogen Aussagen für Matrizen $A \in \mathrm{Mat}(n \times n, K)$ gelten ebenfalls.

Beweis: Wir beweisen zunächst Teil a. und beachten dazu, daß wir aus Lemma A12.18

$$U_0 \subsetneqq U_1 \subsetneqq \ldots \subsetneqq U_m$$

wissen.

Wir müssen zunächst zeigen, daß f_i wohldefiniert ist. Für $\overline{x} = \overline{y} \in U_i/U_{i-1}$ ist

$$x - y \in U_{i-1} = \mathrm{Ker}(f^{i-1}),$$

so daß

$$f(x) - f(y) = f(x-y) \in \mathrm{Ker}(f^{i-2}) = U_{i-2}$$

folgt, d.h. $\overline{f(x)} = \overline{f(y)} \in U_{i-1}/U_{i-2}$. Da mit $x \in U_i = \mathrm{Ker}(f^i)$ zudem $f(x) \in \mathrm{Ker}(f^{i-1}) = U_{i-1}$ gilt, ist f_i wohldefiniert.

Mit f ist dann aber auch f_i eine lineare Abbildung und für die Injektivität reicht es, zu zeigen

$$\mathrm{Ker}(f_i) = \{\overline{0}\}.$$

Nun ist aber $\overline{x} \in \mathrm{Ker}(f_i)$ gleichwertig zu $f(x) \in U_{i-2} = \mathrm{Ker}(f^{i-2})$, was wiederum nur für $x \in \mathrm{Ker}(f^{i-1}) = U_{i-1}$, d.h. für $\overline{x} = \overline{0} \in U_i/U_{i-1}$, zutrifft.

Wir müssen noch zeigen, daß $P = (k_1, \ldots, k_m)$ eine Partition von n ist. Aus

$$0 \neq U_m/U_{m-1} \hookrightarrow U_{m-1}/U_{m-2} \hookrightarrow \ldots \hookrightarrow U_1/U_0 = U_1$$

folgt auch

$$1 \leq \dim_K(U_m/U_{m-1}) \leq \dim_K(U_{m-1}/U_{m-2}) \leq \ldots \leq \dim_K(U_1/U_0),$$

d.h.

$$1 \leq k_m \leq k_{m-1} \leq \ldots \leq k_1.$$

Man beachte dabei, daß $U_m/U_{m-1} \neq 0$ gilt, weil m der Nilpotenzindex von f ist!

Außerdem folgt aus der Dimensionsformel für Vektorräume

$$n = \dim_K(V) = \dim_K(V/U_{m-1}) + \dim_K(U_{m-1}) = \dim_K(U_m/U_{m-1}) + \dim_K(U_{m-1}).$$

Mit Induktion nach m folgt dann

$$\begin{aligned} n =& \dim_K(U_m/U_{m-1}) + \dim_K(U_{m-1}) \\ =& \dim_K(U_m/U_{m-1}) + \dim_K(U_{m-1}/U_{m-2}) + \ldots + \dim_K(U_1/U_0) + \dim_K(U_0) \\ =& k_m + k_{m-1} + \ldots + k_1 + 0. \end{aligned}$$

Damit ist P eine Partition von n und Teil a. ist bewiesen.

Wenden wir uns nun Teil b. zu und konstruieren die Basis B.

Dazu wählen wir zunächst Vektoren

(220) $$x_1^m, \ldots, x_{k_m}^m,$$

deren Restklassen eine Basis von U_m/U_{m-1} bilden. Dann wenden wir f auf diese an und erhalten Vektoren

$$x_1^{m-1} := f(x_1^m), \ldots, x_{k_m}^{m-1} := f(x_{k_m}^m) \in U_{m-1},$$

deren Restklassen in U_{m-1}/U_{m-2} linear unabhängig sind, weil die Abbildung

$$f_m : U_m/U_{m-1} \hookrightarrow U_{m-1}/U_{m-2}$$

eine injektive lineare Abbildung ist. Nun ergänzen wir die Restklassen von diesen durch die Restklassen von Vektoren

$$x_{k_m+1}^{m-1}, \ldots, x_{k_{m-1}}^{m-1},$$

zu einer Basis von U_{m-1}/U_{m-2}. Mit den so gewonnenen Vektoren

$$x_1^{m-1}, \ldots, x_{k_{m-1}}^{m-1}$$

verfahren wir analog und konstruieren so rekursiv Vektoren

$$x_1^i, \ldots, x_{k_i}^i, \tag{221}$$

deren Restklassen jeweils eine Basis für U_i/U_{i-1} sind, für $i = 1, \ldots, m$. Diese $n = k_1 + \ldots + k_m$ Vektoren ordnen wir der Übersichtlichkeit halber in dem Young-Diagramm von P an (siehe Abbildung 12). Wir können die Vektoren im Young-Diagramm auch

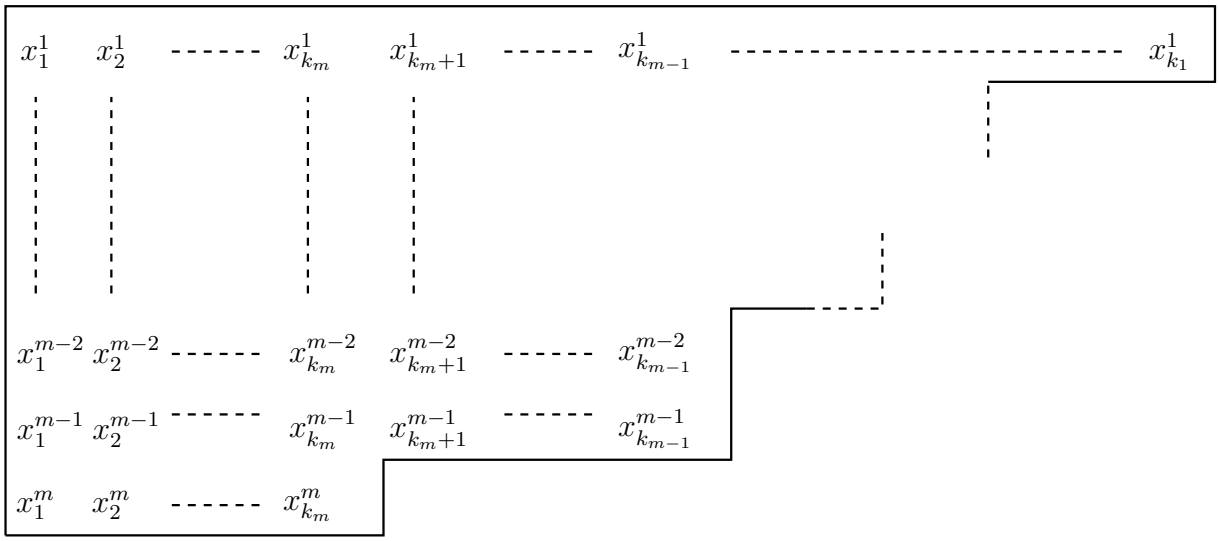


Abbildung 12. Anordnung der Basis B im Young-Diagramm zu P

als Bilder unter der Abbildung f schreiben und erhalten Abbildung 13. Schließlich benennen wir die Vektoren um, wie in Abbildung 14 angegeben, d.h. wir lesen das Diagramm aus, indem wir, in der linken oberen Ecke beginnend, die Spalten sukzessive von oben nach unten durchlaufen. Wir müssen nun nur noch zeigen, daß

$$B = (x_1, \ldots, x_n)$$

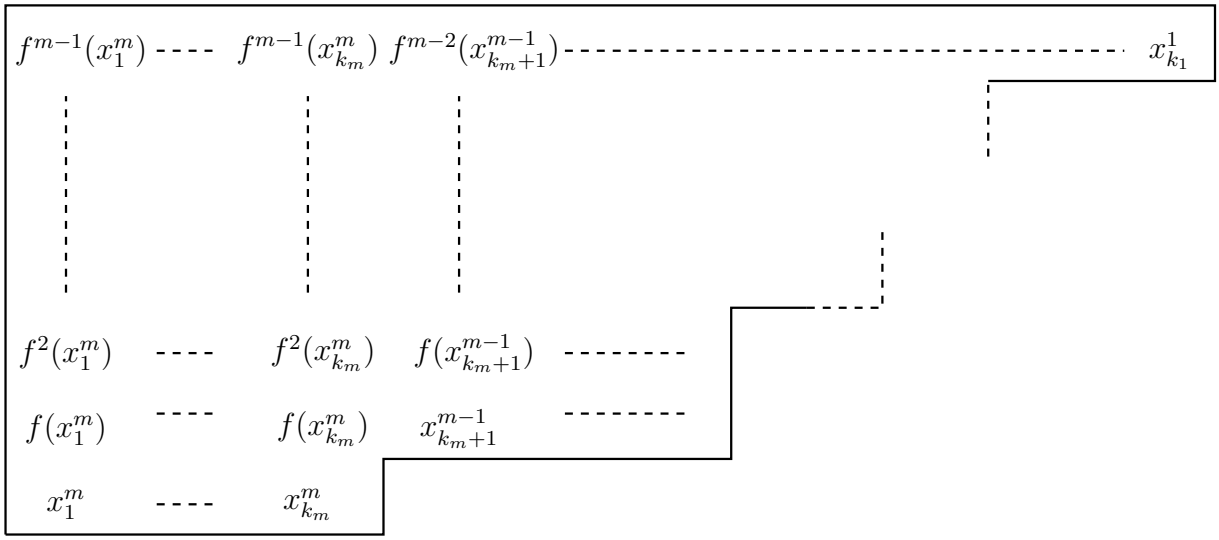


Abbildung 13. Anordnung der Basis B im Young-Diagramm zu P

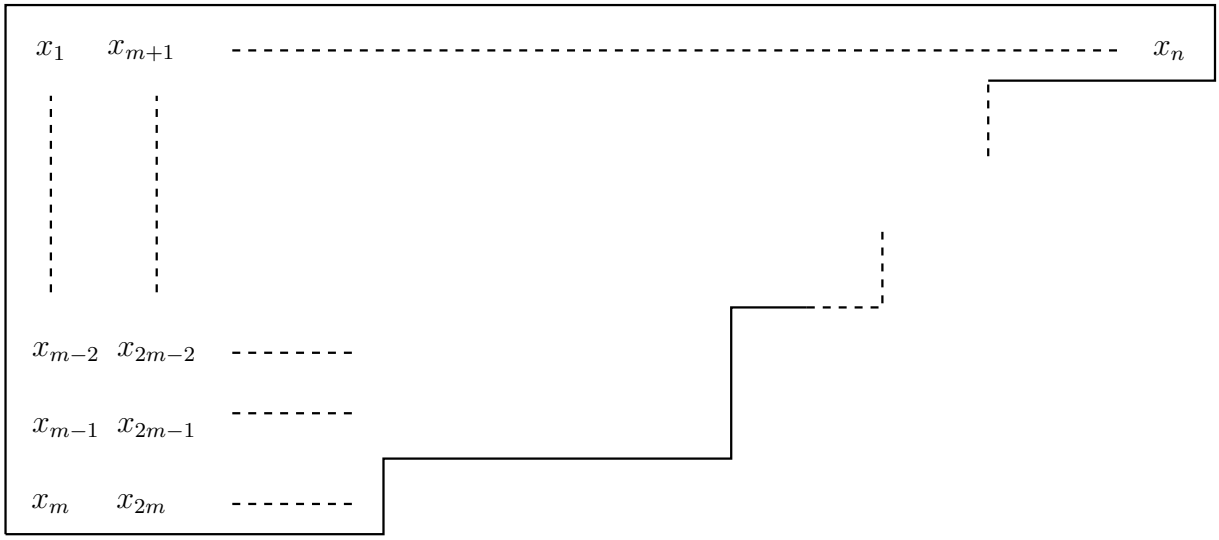


Abbildung 14. Anordnung der Basis B im Young-Diagramm zu P

linear unabhängig ist, dann ist B eine Basis des n-dimensionalen Vektorraums V und die Matrix-Darstellung hat offenbar die gewünschte Gestalt, wie wir aus Abbildung 13 sehen. Dazu beachten wir, daß die Spalten des Diagramms jeweils die kanonische Basis eines zyklischen Unterraums sind und somit einen Jordanblock liefern. Das zeigt insbesondere, daß die zu P duale Partition die Größen der Jordanblöcke liefert.

Um zu zeigen, daß B eine Basis ist, betrachten wir eine Linearkombination

$$\sum_{i=1}^{m} \sum_{j=1}^{k_j} \lambda_{i,j} \cdot x_j^i = 0 \tag{222}$$

der Vektoren in B, die den Nullvektor ergibt. Beachten wir, daß $x_j^i \in U_{m-1}$ für $i < m$ gilt, und betrachten wir die Gleichung in U_m/U_{m-1}, so reduziert sie sich auf

$$\sum_{j=1}^{k_m} \lambda_{m,j} \cdot \overline{x_j^m} = \overline{0} \in U_m/U_{m-1}.$$

Da die Vektoren $\overline{x_1^m}, \ldots, \overline{x_{k_m}^m}$ linear unabhängig sind (siehe (220)), gilt also

$$\lambda_{m,1} = \ldots = \lambda_{m,k_m} = 0$$

und (222) reduziert sich zu

$$\sum_{i=1}^{m-1} \sum_{j=1}^{k_j} \lambda_{i,j} \cdot x_j^i = 0.$$

Diese Gleichung können wir mit demselben Argument in U_{m-1}/U_{m-2} betrachten und erhalten

$$\sum_{j=1}^{k_{m-1}} \lambda_{m-1,i} \cdot \overline{x_j^{m-1}} = \overline{0} \in U_{m-1}/U_{m-2}.$$

Die Restklassen der beteiligten Vektoren sind nach Konstruktion (siehe (221)) linear unabhängig in U_{m-1}/U_{m-2} und somit gilt

$$\lambda_{m-1,1} = \ldots = \lambda_{m-1,k_{m-1}} = 0.$$

Fahren wir so fort erhalten wir insgesamt, daß alle $\lambda_{i,j}$ Null sein müssen und B ist linear unabhängig. □

Wir können damit nun auch Satz A13.1 für nilpotente Endomorphismen beweisen.

Lemma A13.12 (Jordansche NF nilpotenter Endomorphismen)
Es sei $f \in \mathrm{End}_K(V)$ ein nilpotenter Endomorphismus mit $\mu_f = t^m$. 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) $\sum_{j=1}^{m} j \cdot t_j = n = \dim_K \mathrm{Hau}(f, 0) = \dim_K(V)$,

(2) $\sum_{j=1}^{m} t_j = \dim_K \mathrm{Eig}(f, 0)$,

(3) $t_m \geq 1$ und

$$J_f := M_B^B(f) = \bigoplus_{j=1}^{m} \bigoplus_{k=1}^{t_j} J_j(0).$$

Beweis: Sei die Partition $P = (k_1, \ldots, k_m)$ wie in Lemma A13.11 gegeben. Setzen wir

$$t_j := k_j - k_{j+1}$$

für $j = 1, \ldots, m$ mit der Konvention $k_{m+1} = 0$, dann ist t_j gerade die Anzahl der Jordanblöcke der Größe $j \times j$ in der Matrixdarstellung

$$M_B^B(f) = J_{l_1}(0) \oplus \ldots \oplus J_{l_s}(0) \tag{223}$$

in Lemma A13.11 (siehe Abbildung 12). Mithin gilt

$$\sum_{j=1}^{m} t_j = k_1 - k_{m+1} = k_1 = \dim_K(U_1) = \dim_K \operatorname{Ker}(f) = \dim_K \operatorname{Eig}(f, 0)$$

und

$$\sum_{j=1}^{m} j \cdot t_j = n = \dim_K(V) = \dim_K \operatorname{Hau}(f, 0),$$

weil die Summe der Größen der Kästchen mit ihren Vielfachheiten die Größe der Gesamtmatrix ist. Außerdem ist

$$t_m = k_m - k_{m+1} = k_m - 0 = k_m \geq 1$$

und die Matrixdarstellung in (223) kann dann auch geschrieben werden als

$$M_B^B(f) = \bigoplus_{j=1}^{m} \bigoplus_{k=1}^{t_j} J_j(0).$$

□

Bemerkung A13.13 (Jordanbasis einer nilpotenten Matrix)
Ist A eine nilpotente Matrix mit $\mu_A = t^m$ und bestimmt man wie im Beweis von Lemma A13.11 (siehe auch Abbildung 13) linear unabhängige Familien

$$B_{j,l} = \left(A^{j-1}x_l^j, A^{j-2}x_l^j, \ldots, Ax_l^j, x_l^j\right) \subset K^n$$

in $\text{Lös}(A^j, 0)$ für $j = 1, \ldots, m$ und $l = k_{j+1} + 1, \ldots, k_j$, dann ist die Matrix $T \in \mathrm{Gl}_n(K)$, deren Spalten gerade all diese Vektoren sind, eine Transformationsmatrix, die A in Jordansche Normalform überführt.

Beispiel A13.14 (Jordansche Normalform einer nilpotenten Matrix)
Wir wollen nun für die folgende nilpotente Matrix

$$A = \begin{pmatrix} -2 & 0 & 0 & 1 & 0 \\ -2 & -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ -4 & 0 & -1 & 2 & 0 \\ -2 & -1 & 0 & 1 & 1 \end{pmatrix} \in \mathrm{Mat}_5(\mathbb{Q})$$

die Jordansche Normalform sowie die Transformationsmatrix $T \in \mathrm{Gl}_5(\mathbb{Q})$ 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

$$A^3 = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

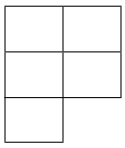
Der Nilpotenzindex von A ist also $m = 3$, $\mu_A = t^3$ und

$$\mathrm{Hau}(A, 0) = \mathbb{Q}^5.$$

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:

$$J_A = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad \text{oder} \quad J_A = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Bestimmen wir die Ränge der Potenzen $A^0, \ldots, A^3$ von A (siehe weiter unten), so können wir die Jordan-Partition



von A dann als

$$\begin{aligned} P =& \big(\operatorname{rang}(A^0) - \operatorname{rang}(A^1), \operatorname{rang}(A^1) - \operatorname{rang}(A^2), \operatorname{rang}(A^2) - \operatorname{rang}(A^3)\big) \\ =& (5-3, 3-1, 1-0) = (2, 2, 1) \end{aligned}$$

berechnen, und wir erhalten als duale Partition

$$P^* = (3, 2),$$

woraus sich unmittelbar die Jordansche Normalform

$$J_A = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right)$$

ergibt sowie die Elementarteiler

$$t_1 = 0, \quad t_2 = 1, \quad t_3 = 1.$$

Um die Jordanbasis B von f_A oder alternativ die Transformationsmatrix T von A zu bestimmen, reicht es gemäß Lemma A13.11 im wesentlichen, geeignete Basen der Vektorräume U_j/U_{j-1} für $j = 3, 2, 1$ zu bestimmen, wobei $U_j = \text{Lös}(A^j, 0) = \text{Ker}(f_A^j)$ ist.

Wir beginnen mit $j = 3$ und $U_3/U_2 = \mathbb{Q}^5/U_2$, wobei wir $U_3 = \mathbb{Q}^5$ beachten.

Um eine Basis von U_3 zu finden, muß man einerseits eine Basis von U_2 berechnen und diese dann zu einer Basis von U_3 ergänzen, indem man sie mit Steinitz in eine Basis von U_3 hineintauscht. Bestimmen wir also zunächst eine Basis B_2' von $U_2 = \text{Lös}(A^2, 0)$. Aufgrund der einfachen Form von A^2 mit Rang 1 geschieht dies durch einfaches Draufschauen — vier der Einheitsvektoren tun es offenbar:

$$B_2' = \big((1,0,0,0,0)^t, (0,1,0,0,0)^t, (0,0,0,1,0)^t, (0,0,0,0,1)^t\big).$$

Für $U_3 = \text{Lös}(A^3, 0) = \mathbb{Q}^5$ ist es noch einfacher, eine Basis zu bestimmen, die kanonische Basis tut's:

$$B_3' = \big((1,0,0,0,0)^t, (0,1,0,0,0)^t, (0,0,1,0,0)^t, (0,0,0,1,0)^t, (0,0,0,0,1)^t\big).$$

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

$$x_1^3 = e_3 = (0,0,1,0,0)^t.$$

Damit erhalten wir $k_3 = \dim_{\mathbb{Q}}(U_3/U_2) = 1$ und die erste Teilbasis der Jordanbasis:

$$B_{3,1} = \big(A^2x_1^3, Ax_1^3, x_1^3\big) = \big((-1,0,0,-2,0)^t, (0,-1,0,-1,0)^t, (0,0,1,0,0)^t\big).$$

Diese können wir in das Young-Diagramm der Jordan-Partition eintragen:

$A^2x_1^3$	
Ax_1^3	
x_1^3	

Als nächstes betrachten wir $j = 2$ und U_2/U_1.

Um eine geeignete Basis von U_2/U_1 zu bestimmen, müssen wir eine Basis von U_1 berechnen und diese zusammen mit der zweiten Ebene des bereits befüllten Young-Diagramms, d.h. mit der $(Ax_1^3) = (Ae_3)$, zu einer Basis von U_2 ergänzen. Dazu berechnen wir zunächst eine Basis B_1' von $U_1 = \text{Lös}(A, 0)$:

$$A = \begin{pmatrix} -2 & 0 & 0 & 1 & 0 \\ -2 & -1 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ -4 & 0 & -1 & 2 & 0 \\ -2 & -1 & 0 & 1 & 1 \end{pmatrix} \overset{\text{Gauß}}{\longmapsto} \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \overset{\substack{-1\text{'en} \\ \text{einfügen}}}{\longmapsto} \begin{pmatrix} 1 & 0 & 0 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{pmatrix}$$

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 $U_1 = \text{Eig}(A, 0) = \text{Lös}(A, 0)$ ist dann

$$B_1' = \big((1, 0, 0, 2, 0)^t, (0, 1, 0, 0, 1)^t\big).$$

Wir müssen also die Familie

$$B_2'' = B_1' \cup (Ae_3) = \big((1, 0, 0, 2, 0)^t, (0, 1, 0, 0, 1)^t, (0, -1, 0, -1, 0)^t\big),$$

zu einer Basis von U_2 ergänzen. Dazu können wir sie mit Hilfe des Austauschsatzes von Steinitz in die Basis B_2' von U_2 hineintauschen, oder alternativ kann man auch einfach genau hinschauen. Man sieht nämlich leicht, daß der Vektor

$$x_2^2 = (0, 0, 0, 0, 1)^t$$

von den drei Vektoren in B_2'' linear unabhängig ist, und somit ergänzt er B_2'' zu einer Basis von 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 x_2^2 bestimmt haben, erhalten wir die zweite Teilbasis

$$B_{2,2} = \big(Ax_2^2, x_2^2\big) = \big((0, 1, 0, 0, 1)^t, (0, 0, 0, 0, 1)^t\big)$$

der Jordanbasis und das fertig ausgefüllte Young-Diagramm der Jordan-Partition:

$$\begin{array}{|c|c|} \hline A^2x_1^3 & Ax_2^2 \\ \hline Ax_1^3 & x_2^2 \\ \hline x_1^3 \\ \cline{1-1} \end{array} \quad = \quad \begin{array}{|c|c|} \hline x_1 & x_4 \\ \hline x_2 & x_5 \\ \hline x_3 \\ \cline{1-1} \end{array}$$

Im Prinzip bliebe noch der Fall $j = 1$ zu untersuchen, aber da die zu P duale Partition nur zwei Spalten hat und damit $t_1 = 0$ gilt, sind wir fertig.

Wir haben also die Jordanbasis $B = B_{3,1} \cup B_{2,2}$ bestimmt und damit auch die Transformationsmatrix T, deren Spalten die Vektoren in B sind. Wir haben

$$T = \begin{pmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ -2 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{pmatrix} \in \mathrm{Gl}_5(\mathbb{Q})$$

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

$$J_A = T^{-1} \circ A \circ T = \left(\begin{array}{ccc|cc} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

D) Jordanschen Normalform für allgemeine Endomorphimen

Kommen wir nun zum Beweis von Satz A13.1.

Beweis von Satz A13.1: Nach Satz A12.19 zerfällt V in die direkte Summe der Haupträume $V_i := \mathrm{Hau}(f, \lambda_i)$, $i = 1, \ldots, r$, und diese sind nach Lemma A12.18 invariant unter f und $f - \lambda_i \mathrm{id}_V$. Betrachten wir nun die Abbildungen

$$(f - \lambda_i \mathrm{id}_V)_{V_i} : V_i \to V_i$$

für $i = 1, \ldots, r$, so sind diese nilpotent mit $\chi_{(f-\lambda_i \mathrm{id}_V)_{V_i}} = t^{n_i}$ und $\mu_{(f-\lambda_i \mathrm{id}_V)_{V_i}} = t^{m_i}$ (vgl. Korollar A12.21). Nach Lemma A13.12 gibt es dann aber für jedes $i = 1, \ldots, r$ Basen B_i von V_i und natürliche Zahlen t_{ij}, $j = 1, \ldots, m_i$, so daß gilt

(1) $\displaystyle\sum_{j=1}^{m_i} j \cdot t_{ij} = n_i = \dim_K \mathrm{Hau}(f, \lambda_i)$,

(2) $\displaystyle\sum_{j=1}^{m_i} t_{ij} = \dim_K \mathrm{Eig}\left((f - \lambda_i \mathrm{id}_V)_{V_i}, 0\right) = \dim_K \mathrm{Eig}(f, \lambda_i)$,

(3) $t_{im_i} \geq 1$ und

$$\begin{aligned} M_{B_i}^{B_i}(f_{V_i}) =& \lambda_i \mathbb{1}_{n_i} + M_{B_i}^{B_i}((f - \lambda_i \operatorname{id}_V)_{V_i}) \\ =& \lambda_i \mathbb{1}_{n_i} + \left(\bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(0) \right) = \bigoplus_{j=1}^{m_i} \bigoplus_{k=1}^{t_{ij}} J_j(\lambda_i). \end{aligned}$$

Damit folgt die Behauptung, da für $B = B_1 \cup \ldots \cup B_r$ 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 A13.15 (Jordansche Normalform - I)

INPUT: $A \in \operatorname{Mat}_n(\mathbb{Q})$ mit μ_A zerfällt in Linearfaktoren.

OUTPUT: J_A und eine Transformationsmatrix $T \in \operatorname{Gl}_n(K)$ mit $T^{-1} \circ A \circ T = J_A$.

1. Schritt: Bestimme das Minimalpolynom μ_A von A und faktorisiere 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 $i = 1, \ldots, r$ bilde man die Matrix $A_i = A - \lambda_i \mathbb{1}_n$ und führe folgende Schritte aus:

Schritt a.: Berechne die Partition $P = (k_1, \ldots, k_{m_i})$ von $n - \operatorname{rang}(A_i^{m_i})$ mit $k_j = \operatorname{rang}(A_i^{j-1}) - \operatorname{rang}(A_i^j)$ gemäß Lemma A13.11 sowie das zugehörige Young-Diagramm.

Schritt b.: Bestimme eine Basis B_{m_i} von Lös $\left(A_i^{m_i}, 0\right)$ sowie eine Basis B_{m_i-1} von Lös $\left(A_i^{m_i-1}, 0\right)$.

Schritt c.: Tausche B_{m_i-1} mittels des Satzes von Steinitz in B_{m_i} hinein und bestimme die in B_{m_i} verbliebenen Vektoren $x_1^{m_i}, \ldots, x_{k_{m_i}}^{m_i}$.

Schritt d.: Dann fülle man die ersten k_{m_i} Spalten des Young-Diagramms von P durch die Vektoren $A_i^{m_i-1} x_l^{m_i}, \ldots, A_i^0 x_l^{m_i}$ auf, $l = 1, \ldots, k_{m_i}$, wie in Abbildung 15.

Schritt e.: Für $j = m_i - 1, \ldots, 1$ führe man folgendes aus:

- bestimme eine Basis B_{j-1} von Lös $\left(A_i^{j-1}, 0\right)$;
- tausche B_{j-1} sowie die auf der j-ten Ebene des Young-Diagramms bereits eingetragenen Vektoren mittels des Satzes von Steinitz in B_j hinein;

- bestimme die in B_j verbliebenen Vektoren $x^j_{k_{j+1}+1}, \dots, x^j_{k_j}$;
- für $l = k_{j+1}+1, \dots, k_j$ fülle die Spalten des Young-Diagramms von P mit den Vektoren $A_i^{j-1}x_l^j, \dots, A_i^0 x_l^j$.

Schritt f.: Füge die Vektoren aus dem Young-Diagramm als Spalten in die Matrix T ein, beginnend in der linken oberen Ecke und die Spalten des Young-Diagramms von oben nach unten nacheinander durchlaufend.

4. Schritt: Gib $T^{-1} \circ A \circ T$ und T zurück.

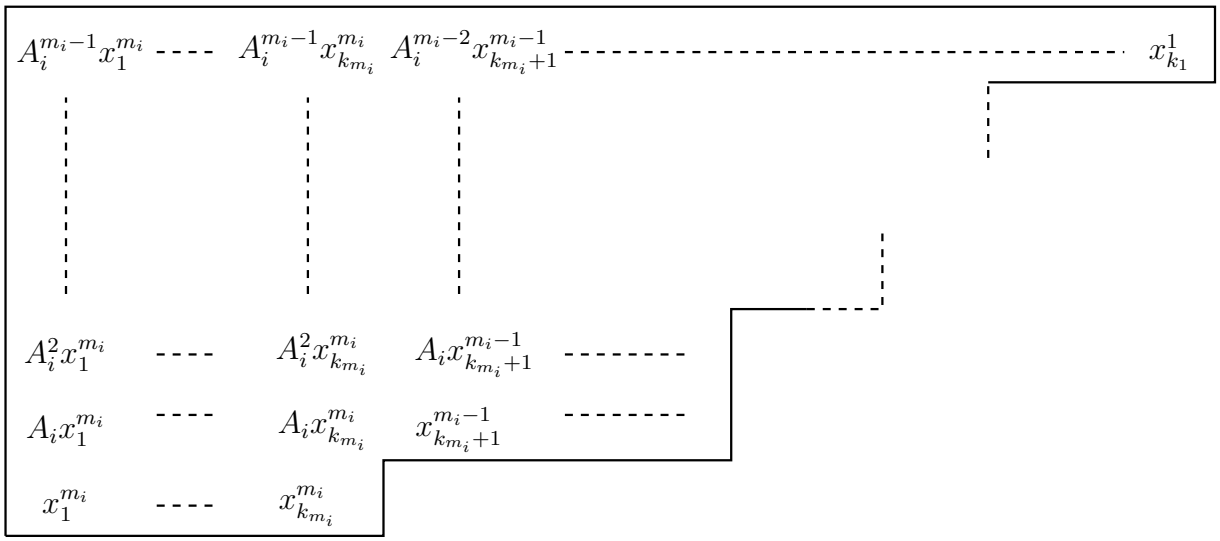


Abbildung 15. Anordnung der Basis B_i im Young-Diagramm zu P

Beispiel A13.16 (Jordansche Normalform)

Wir wollen nun die Jordansche Normalform und die Transformationsmatrix von

$$A = \begin{pmatrix} 2 & 1 & 1 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathrm{Mat}_4(\mathbb{Q})$$

bestimmen.

Das charakteristische Polynom berechnet man mit Hilfe des Kästchensatzes als

$$\chi_A = \begin{vmatrix} t-2 & -1 & -1 & -2 \\ 0 & t-1 & 0 & 0 \\ 0 & -1 & t-1 & -1 \\ 0 & 0 & 0 & t-2 \end{vmatrix} = (t-2) \cdot \begin{vmatrix} t-1 & 0 \\ -1 & t-1 \end{vmatrix} \cdot (t-2) = (t-2)^2 \cdot (t-1)^2.$$

Wir betrachten nun zunächst den Eigenwert $\lambda = 1$:

Wie im nilpotenten Fall können wir zunächst die Partition von

$$\dim_{\mathbb{R}} \operatorname{Hau}(A, 1) = \operatorname{mult}(\chi_A, 1) = 2$$

ausrechnen, die Jordansche Normalform der Einschränkung von f_A auf $\operatorname{Hau}(A, 1)$ festlegt und damit auch die Jordankästchen zum Eigenwert 1 in J_A. Es kommen nur die beiden Partitionen $P_1 = (2)$ und $P_1 = (1, 1)$ mit den in Abbildung 16 gegebenen Young-Diagrammen in Frage. Wir erhalten die Partition formal als

Abbildung 16. Mögliche Young-Diagramme zum Eigenwert 1

$$P_1 = \Big(\operatorname{rang}\big((A-\mathbb{1}_4)^0\big) - \operatorname{rang}\big((A-\mathbb{1}_4)^1\big), \ldots, \operatorname{rang}\big((A-\mathbb{1}_4)^{m-1}\big) - \operatorname{rang}\big((A-\mathbb{1}_4)^m\big),$$

wobei m der Nilpotenzindex des Eigenwertes 1 ist. Da die Matrix

$$A - \mathbb{1}_4 = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

offenbar den Rang 3 hat und somit

$$\operatorname{rang}\big((A - \mathbb{1}_4)^0\big) - \operatorname{rang}\big((A - \mathbb{1}_4)^1\big) = 4 - 13 = 1$$

gilt, muß die Partition $P_1 = (1, 1)$ sein mit dualer Partition $P_1^* = (2)$. Daraus folgt insbesondere dann, daß J_A zum Eigenwert 1 genau ein Kästchen der Größe 2 hat. Die zugehörige Jordanteilbasis wird von der Form

$$B' = \big((A - \mathbb{1}_4) \circ x_1^2, x_1^2\big)$$

sein für einen Basisvektor

$$x_1^2 \in U_2 \setminus U_1$$

mit

$$U_i = \operatorname{Ker}\big((A - \mathbb{1}_4)^2\big).$$

Eingefüllt ins Young-Diagramm von P_1 ergibt sich das Bild:

$(A - \mathbb{1}_4) \circ x_1^2$
x_1^2

Wir wollen nun also Basen für $U_1 = \operatorname{Eig}(A, 1)$ und $U_2 = \operatorname{Hau}(A, 1)$ berechnen und wollen die Basis von U_1 in die von U_2 hineintauschen. Der in der Basis von U_2 verbleibende Vektor ist dann der gesuchte Vektor x_1^2.

Berechnen wir also eine Basis von $U_1 = \mathrm{Eig}(A,1) = \mathrm{Lös}(A - \mathbb{1}_4, 0)$:

$$A - \mathbb{1}_4 = \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \overset{\text{Gauß}}{\longmapsto} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{\text{-1'en einfügen}}{\longmapsto} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Die gesuchte Basis von U_1 ist dann

$$B_1' = (1, 0, -1, 0)^t.$$

Als nächstes berechnen wir $U_2 = \mathrm{Hau}(A,1) = \mathrm{Lös}\left((A - \mathbb{1}_4)^2, 0\right)$:

$$(A - \mathbb{1}_4)^2 = \begin{pmatrix} 1 & 2 & 1 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \overset{\text{Gauß}}{\longmapsto} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{\text{-1'en einfügen}}{\longmapsto} \begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Als Basis für $\mathrm{Hau}(A,1)$ erhalten wir also

$$B_2' = \left((2, -1, 0, 0)^t, (1, 0, -1, 0)^t\right).$$

Der erste der beiden Vektoren ist nicht in $\mathrm{Eig}(A,1)$, so daß wir ihn als x_1^2 wählen können. Damit erhalten wir

$$x_1 = (A - \mathbb{1}_4)x_1^2 = (1, 0, -1, 0)^t, \quad x_2 = x_1^2 = (2, -1, 0, 0)^t$$

als Jordanteilbasis für den Hauptraum zum Eigenwert 1 und damit als die ersten beiden Spalten von T.

Für den zweiten Eigenwert $\lambda = 2$ gehen wir analog vor:

Die Matrix

$$A - 2 \cdot \mathbb{1}_4 = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

hat ebenfalls offenbar Rang 3, so daß mit demselben Argument wie oben die gesuchte Partition $P_2 = (1,1)$ sein muß. Wir können also genau wie oben vorgehen, wobei wir Notationen wie

$$U_i = \mathrm{Ker}\left((A - 2 \cdot \mathbb{1}_4)^2\right)$$

der Einfachheit wieder verwenden, nun angepaßt auf den neuen Eigenwert.

Zunächst berechnen wir wieder eine Basis B_1' von $U_1 = \mathrm{Eig}(A, 2)$:

$$A - 2 \cdot \mathbb{1}_4 = \begin{pmatrix} 0 & 1 & 1 & 2 \\ 0 & -1 & 0 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{\text{Gauß}}{\longmapsto} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{\substack{-1\text{'en} \\ \text{einfügen}}}{\longmapsto} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Wir erhalten also

$$B_1' = (-1, 0, 0, 0)^t.$$

Die Berechnung der Basis B_2' von $U_2 = \mathrm{Hau}(A, 2) = \mathrm{Lös}\left((A - \mathbb{1}_4)^2, 0\right)$ ergibt

$$(A - 2 \cdot \mathbb{1}_4)^2 = \begin{pmatrix} 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & -2 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{\text{Gauß}}{\longmapsto} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \overset{\substack{-1\text{'en} \\ \text{einfügen}}}{\longmapsto} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

und damit

$$B_2' = \left((-1, 0, 0, 0)^t, (0, 0, -1, -1)^t\right).$$

Wir können also $x_1^2 = (0, 0, 1, 1)^t$ wählen und erhalten dann die Jordanteilbasis mit den Vektoren

$$x_3 = (A - 2 \cdot \mathbb{1}_4) \circ x_1^2 = (3, 0, 0, 0)^t, \quad x_4 = x_1^2 = (0, 0, 1, 1)^t$$

als die Spalten 3 und 4 der Matrix T.

Insgesamt haben wir also

$$T = \begin{pmatrix} 1 & 2 & 3 & 0 \\ 0 & -1 & 0 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \in \mathrm{Gl}_4(\mathbb{Q})$$

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 A13.26 sehr viel einfacher zu bewerkstelligen ist. Dies führt auf folgenden Algorithmus zur

Bestimmung der Jordanschen Normalform einer Matrix A, deren charakteristisches Polynom zerfällt.

Algorithmus A13.17 (Jordansche Normalform - II)

INPUT: $A \in \mathrm{Mat}_n(\mathbb{Q})$ 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 $\mathrm{mult}(\mu_A, \lambda_i) = m_i$ bestimme man für $j = 0, \ldots, m_i + 1$ die Zahlen $\mathrm{rang}\left((A - \lambda_i \mathbb{1}_n)^j\right)$ und berechne daraus den Vektor der Elementarteiler $(t_{i1}, \ldots, t_{im_i})$. 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 A13.18 (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 $i = 1, \ldots, r$ 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 $A \in \mathrm{Mat}_n(K)$ 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 A13.19 (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.

E) Nicht-konstruktiver Beweis der Jordanschen Normalform

Wir haben oben einen *konstruktiven* Beweis des Satzes der Jordanschen Normalform für nilpotente Endomorphismen A13.12 gegeben. Aus dem Beweis ließ sich unmittelbar ableiten, wie man die Jordansche Normalform eines Endomorphismus sowie die zugehörige Jordanbasis berechnen kann. Bezahlt haben wir dies mit einer Vielzahl an Indizes, die den Beweis nicht unbedingt übersichtlich machen. Wir wollen das Kapitel damit abschließen, einen alternativen *nicht-konstruktiven* Beweis des Satzes A13.12 zu geben. Dazu erinnern wir uns zunächst der Notationen

$$I_{f,x} = \{p \in K[t] \mid p(f)(x) = 0\}$$

und

$$U_{f,x} = \{p(f)(x) \mid p \in K[t]\}$$

aus Aufgabe A10.30 und beweisen dann folgende Hilfsaussage.

Lemma A13.20 (Zykl. Unterräume haben invariante Komplemente.)
Ist $f \in \mathrm{End}_K(V)$ mit $\mu_f = t^m$, so gibt es einen Vektor $0 \neq x \in V$, so daß der f-invariante Vektorraum $U_{f,x}$ Dimension m hat und ein f-invariantes direktes Komplement besitzt.

Beweis: Da m nach Satz A12.19 der Nilpotenzindex von f ist, gibt es nach Lemma A12.18 ein

$$x \in \ker(f^m) \setminus \ker(f^{m-1})$$

und nach Aufgabe A10.30 gilt dann

$$I_{f,x} = \{t^m \cdot p \mid p \in K[t]\} \tag{224}$$

und

$$U_{f,x} = \mathrm{Lin}\left(f^{m-1}(x), f^{m-2}(x), \ldots, f(x), x\right)$$

ist ein f-invarianter zyklischer Unterraum von V, der nach Aufgabe A6.51 die Dimension m hat.

Wir betrachten nun die nicht-leere Menge

$$M = \{U \leq V \mid f(U) \subseteq U, U \cap U_{f,x} = 0\}$$

und wählen in dieser ein $U \in M$ von maximaler Dimension. Wir wollen zeigen, daß der f-invariante Unterraum U ein Komplement von $U_{f,x}$ ist. Dazu setzen wir

$$W = U \oplus U_{f,x}$$

und zeigen zunächst die folgende Behauptung:

Behauptung
Wenn $y \in V$ ist mit $f(y) \in W$, dann gilt schon $y \in W$.

Nach Voraussetzung gibt es einen Vektor $u \in U$ und ein Polynom $p \in K[t]$ mit

$$f(y) = u + p(f)(x). \tag{225}$$

Da das Minimalpolynom von f die Form $\mu_f = t^m$ hat, gilt

$$\begin{aligned} 0 = \mu_f(f)(y) = f^m(y) = f^{m-1}\big(f(y)\big) &= f^{m-1}\big(u + p(f)(x)\big) \\ &= f^{m-1}(u) + \big(f^{m-1} \circ p(f)\big)(x) = f^{m-1}(u) + q(f)(x) \in U + U_{f,x}, \end{aligned}$$

wobei $q = t^{m-1} \cdot p$ ist. Damit ist aber

$$q(f)(x) = -f^{m-1}(u) \in U \cap U_{f,x} = \{0\},$$

und wir erhalten

$$q(f)(x) = 0.$$

Letzteres impliziert

$$t^{m-1} \cdot p = q \in I_{f,x} = \{t^m \cdot g \mid g \in K[t]\}.$$

Mithin muß p durch t teilbar sein und es gibt ein Polynom $g \in K[t]$ mit

$$p = t \cdot g. \tag{226}$$

Wir setzen nun

$$z = y - g(f)(x). \tag{227}$$

Dieser Vektor z erfüllt wegen (225), (226) und (227) die Gleichung:

$$f(z) \overset{(227)}{=} f\big(y - g(f)(x)\big) = f(y) - \big(f \circ g(f)\big)(x) \overset{(226)}{=} f(y) - p(f)(x) \overset{(225)}{=} u. \tag{228}$$

Gelingt es uns, zu zeigen, daß z in W liegt, so ist auch

$$y = z + g(f)(x) \in W + U_{f,x} = W$$

und die Behauptung ist gezeigt. Wir nehmen also

$$z \notin W \tag{229}$$

an, und wollen dies zum Widerspruch führen.

Der Unterraum

$$U' = U + U_{f,z}$$

ist als Summe zweier f-invarianter Unterräume f-invariant und seine Dimension ist größer als die von U, da U' zusätzlich zu U noch den Vektor z enthält, denn aus $z \in U$ würde auch $z \in W$ folgen. Die Maximalitätsbedinung der Wahl von U impliziert also, daß

$$U' \cap U_{f,x} \neq 0$$

gilt, und wir finden mithin zwei Polynome $h, k \in K[t]$ sowie einen Vektor $v \in U$, so daß

$$0 \neq v + h(f)(z) = k(f)(x) \in U' \cap U_{f,x}. \tag{230}$$

Daraus folgt unmittelbar

$$h(f)(z) = k(f)(x) - v \in U_{f,x} + U = W. \tag{231}$$

Wir betrachten nun zunächst den Fall, daß t kein Teiler von h ist, so daß die Bézout-Identität A10.22 uns zwei Polynome $r, s \in K[t]$ schenkt mit

$$1 = r \cdot t + s \cdot h.$$

Daraus ergibt sich dann die Gleichung

$$\begin{aligned} z \;&=\; \mathrm{id}(z) = (r \cdot t + s \cdot h)(f)(z) = r(f)\big(f(z)\big) + s(f)\big(h(f)(z)\big) \\ &\overset{(228)}{=} r(f)(u) + s(f)\big(h(f)(z)\big) \overset{(231)}{\in} r(f)(U) + s(f)(W) = U + W = W, \end{aligned}$$

da U und W f-invariant sind. Dies steht aber im Widerspruch zur Annahme (229).

Als nächstes betrachten wir den Fall, daß t ein Teiler von h ist und erhalten ein Polynom $r \in K[t]$ mit $h = r \cdot t$. Damit gilt dann aber die Gleichung

$$U_{f,x} \ni k(f)(x) = v + h(f)(z) = v + r(f)\big(f(z)\big) \overset{(228)}{=} v + r(f)(u) \in U,$$

und wegen (230) hätten wir somit einen Vektor ungleich 0 in $U \cap U_{f,x} = \{0\}$ gefunden, was ein offensichtlicher Widerspruch ist.

Damit haben wir die obige Behauptung gezeigt, und wir wollen nun daraus

$$W = V$$

herleiten, was gleichbedeutend dazu ist, daß U ein f-invariantes direktes Komplement von $U_{f,x}$ in V ist. Sei dazu $y \in V$ beliebig gegeben. Wegen $f^m(y) = 0 \in W$ gibt es eine kleinste natürliche Zahl k mit $f^k(y) \in W$. Wäre diese Zahl k nicht 0, so leiten wir aus

$$f\big(f^{k-1}(y)\big) = f^k(y) \in W$$

und obiger Behauptung her, daß schon $f^{k-1}(y) \in W$ gilt, im Widerspruch zur Minimalität von k. Also ist $k = 0$ und damit $y = f^0(y) \in W$, woraus wie gewünscht $V = W$ folgt. □

Alternativer Beweis von Lemma A13.12: Wir zeigen zunächst mit Induktion nach $n = \dim_K(V)$, daß es eine Basis B gibt, bezüglich derer die Matrixdarstellung für f die Form

$$M_B^B(f) = \bigoplus_{j=1}^{n} \bigoplus_{k=1}^{t_j} J_j(0), \tag{232}$$

wobei t_j gerade die Anzahl der Jordanblöcke der Größe j ist. Für $n = 1$ ist dabei nichts zu zeigen, da dann f der Nullhomomorphismus ist. Sei also $n > 1$.

Nach Voraussetzung ist $f \in \mathrm{End}_K(V)$ ein nilpotenter Endomorphismus mit $\mu_f = t^m$. Wir wählen $0 \neq x \in V$ wie in Lemma A13.20 und betrachten den m-dimensionalen f-invarianten Unterraum $U_{f,x}$ mit Basis

$$B' = \big(f^{m-1}(x), f^{m-2}(x), \ldots, f(x), x\big)$$

bezüglich derer die Matrixdarstellung von $f_{U_{f,x}}$ ein Jordanblock der Größe m ist,

$$M_{B'}^{B'}\big(f_{U_{f,x}}\big) = J_m(0).$$

Falls $\dim_K(U_{f,x}) = m = n = \dim_K(V)$ gilt, so ist

$$U_{f,x} = V = \mathrm{Hau}(f, 0)$$

und $B = B'$ ist die gesuchte Jordanbasis von V mit $J_f = M_B^B(f) = J_n(0)$.

Ist $\dim_K(U_{f,x}) = m < n = \dim_K(V)$, so finden wir mit Lemma A13.20 ein f-invariantes Komplement U zu $U_{f,x}$ und wegen

$$t^n = \chi_f = \chi_{f_U} \cdot \chi_{f_{U_{f,x}}}$$

ist auch $f_{U_{f,x}}$ dann wieder nilpotent. Da die Dimension von U echt kleiner als die von V ist, finden wir nun mit Induktion eine Jordanbasis B'' von U mit

$$M_{B''}^{B''}(f_U) = \bigoplus_{j=1}^{n''} \bigoplus_{k=1}^{t''_j} J_j(0),$$

wobei $n'' = \dim_K(U) < \dim_K(V) = n$. Dann ist

$$B = B' \cup B''$$

eine Basis von V und die Matrixdarstellung $M_B^B(f) = M_{B'}^{B'}(f_{U_{f,x}}) \oplus M_{B''}^{B''}(f_U)$ ist eine Matrix in jordanscher Normalform (232).

Da $M_B^B(f)^m = M_B^B(f^m) = 0$ die Nullmatrix ist, kann $M_B^B(f)$ keinen Jordanblock $J_j(0)$ einer Größe $j > m$ enthalten, so daß die erste Summe in (232) nur bis m geht. Zudem gilt $t_m > 0$, da der Jordanblock zu $f_{U_{f,x}}$ die Größe m hat, und die Größe der Matrix ist die Summe der Größen der Jordanblöcke,

$$n = \sum_{j=1}^{m} j \cdot t_j.$$

Schließlich erniedrigt sich der Rang von f mit jedem Jordanblock um 1, da ein Jordanblock der Größe j gerade Rang $j-1$ hat. Damit gilt aber

$$\operatorname{rang}(f) = n - \#\text{Jordanblöcke in } M_B^B(f) = n - \sum_{j=0}^{m} t_j,$$

und da der Eigenraum von f zum Eigenwert 0 gerade der Kern von f ist, folgt

$$\dim_K \operatorname{Eig}(f,0) = \dim_K \operatorname{Ker}(f) = n - \operatorname{rang}(f) = \sum_{j=0}^{m} t_j.$$

Damit ist Lemma A13.12 gezeigt. □

Aufgaben

Aufgabe A13.21

Bestimme die Jordansche Normalform und die zugehörige Transformationsmatrix für die Matrix A

$$A = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & 2 & 0 \\ -2 & 0 & 1 & 0 & 0 \\ 0 & -4 & 0 & 0 & -2 \end{pmatrix} \in \operatorname{Mat}_5(\mathbb{Q}).$$

Aufgabe A13.22

Bestimme die Jordansche Normalform und die zugehörige Transformationsmatrix $T^{-1} \in \operatorname{Gl}_4(\mathbb{Q})$ für

$$A = \begin{pmatrix} 3 & 6 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 9 & 3 & 0 \\ 0 & 12 & 0 & 3 \end{pmatrix} \in \operatorname{Mat}_4(\mathbb{Q}).$$

Aufgabe A13.23

Es sei $A \in \operatorname{Mat}(5,K)$ mit $\chi_A = t(t-1)^4$, $\mu_A = t(t-1)^2$ und $\operatorname{rang}(A - \mathbb{1}_5) = 2$. Bestimme die Jordansche Normalform von A.

Aufgabe A13.24

Zeige, ist $A \in \operatorname{Mat}_n(K)$ so, daß χ_A über K in Linearfaktoren zerfällt, so sind A und A^t konjugiert.

Aufgabe A13.25

Beweise oder widerlege die folgende Aussage für zwei Matrizen $A, B \in \mathrm{Mat}_n(K)$:

$$A \text{ ist konjugiert zu } B \iff \chi_A = \chi_B,\ \mu_A = \mu_B \text{ und } \operatorname{rang}(A) = \operatorname{rang}(B).$$

Aufgabe A13.26 (Berechnung der Elementarteiler)

Mit den Bezeichnungen aus Satz A13.1 zeige man, für $i = 1, \ldots, r$ und $1 \leq j \leq m_i$ gilt:

$$t_{ij} = \operatorname{rang}\left((f - \lambda_i \operatorname{id}_V)^{j-1}\right) - 2 \cdot \operatorname{rang}\left((f - \lambda_i \operatorname{id}_V)^{j}\right) + \operatorname{rang}\left((f - \lambda_i \operatorname{id}_V)^{j+1}\right)$$

bzw.

$$t_{ij} = \operatorname{rang}\left((A - \lambda_i \mathbb{1}_n)^{j-1}\right) - 2 \cdot \operatorname{rang}\left((A - \lambda_i \mathbb{1}_n)^{j}\right) + \operatorname{rang}\left((A - \lambda_i \mathbb{1}_n)^{j+1}\right).$$

Hinweise: 1. Zeige, $J_j(0)^l = \left(\delta_{\mu+l,\nu}\right)_{\mu,\nu=1,\ldots,j}$ und $\operatorname{rang}\left(J_j(0)^l\right) = \max\{0, j-l\}$ 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 $g_i := (f - \lambda_i \operatorname{id}_V)_{\mathrm{Hau}(f,\lambda_i)}$ zurück.

A14 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

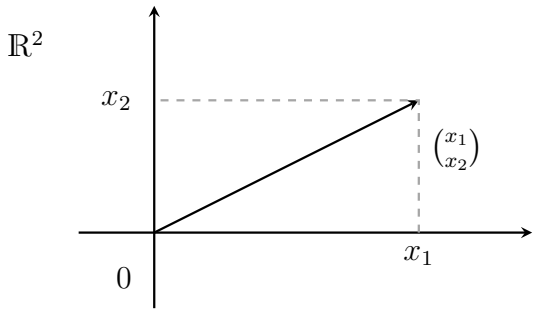
$$||\cdot|| : \mathbb{R}^2 \to \mathbb{R}_{\geq 0},$$

die einem Vektor $x = (x_1, x_2)^t \in \mathbb{R}^2$ seine Länge $||x||$ zuordnet, sowie

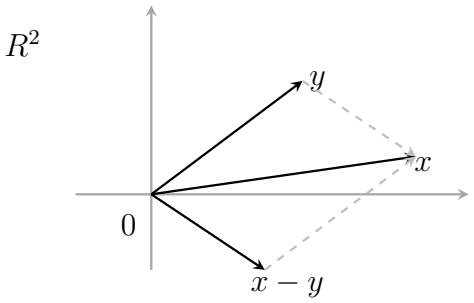
$$d : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}_{\geq 0},$$

die zwei Punkten $x \in \mathbb{R}^2$ und $y \in \mathbb{R}^2$ ihren Abstand $d(x, y)$ zuweist.

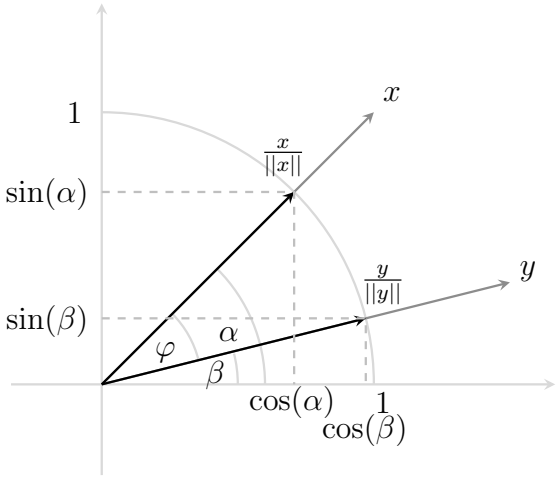
Der Satz von Pythagoras liefert dann $||x|| = \sqrt{x_1^2 + x_2^2}$.



Wir nennen $||x||$ auch die *euklidische Norm* des Vektors x. Da der Abstand der Punkte $x = (x_1, x_2)^t$ und $y = (y_1, y_2)^t$ gerade die Länge des Vektors $x - y$ ist, folgt somit $d(x, y) = ||x - y|| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}$.



Mit Hilfe der Norm können wir - nach einigen geometrischen Überlegungen - auch den Winkel $\measuredangle(x, y)$, den zwei Vektoren x und y miteinander einschließen, bestimmen.



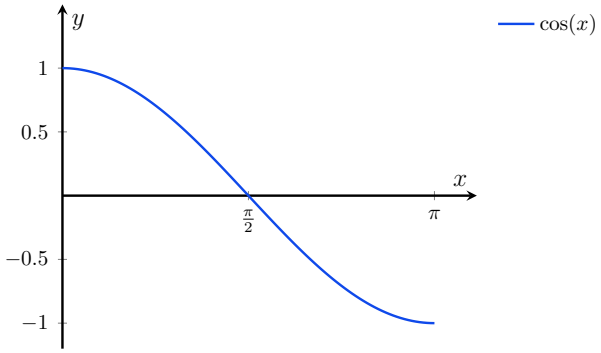
Dazu gehen wir zunächst zu den normierten Vektoren $\frac{x}{||x||}$ und $\frac{y}{||y||}$ über, die beide die Länge eins haben, wobei wir $x \neq 0$ und $y \neq 0$ voraussetzen. Mit den Bezeichnungen in der Skizze gilt dann

$$\measuredangle(x,y) = \measuredangle\left(\frac{x}{||x||}, \frac{y}{||y||}\right) = \alpha - \beta = \varphi.$$

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

$$\cos : [0,\pi] \to \mathbb{R} : \varphi \mapsto \cos(\varphi)$$



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

$$\cos(\varphi) = \cos(\alpha-\beta) = \cos(\alpha)\cos(\beta) + \sin(\alpha)\sin(\beta) = \frac{x_1y_1 + x_2y_2}{||x|| \cdot ||y||}.$$

Dies führt zur Definition einer weiteren Abbildung

$$\langle \cdot, \cdot \rangle : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} : (x,y) = \big((x_1,x_2)^t, (y_1,y_2)^t\big) \mapsto \langle x,y\rangle := x_1y_1 + x_2y_2,$$

welche wir *Skalarprodukt* nennen. Mit deren Hilfe erhalten wir

$$\cos(\varphi) = \frac{\langle x, y\rangle}{||x|| \cdot ||y||}$$

oder alternativ

$$\measuredangle(x, y) = \varphi = \arccos\left(\frac{\langle x, y\rangle}{||x|| \cdot ||y||}\right).$$

Wir sind also mittels recht einfacher Abbildungen in der Lage, Längen und Winkel auszudrücken. Dieses Beispiel motiviert die folgenden Begriffsbildungen.

A) Skalarprodukte

Definition A14.1 (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 $x, y, z \in V$ gilt:

(1) $\langle x, \lambda y + \mu z\rangle = \lambda \cdot \langle x, y\rangle + \mu \cdot \langle x, z\rangle$,

(2) $\langle x, y\rangle = \overline{\langle y, x\rangle}$,

(3) $\langle x, x\rangle \in \mathbb{R}$ und $\langle x, x\rangle > 0$ für $x \neq 0$.

Ist $\mathbb{K} = \mathbb{R}$, so nennen wir das Quadrupel $(V, +, \cdot, \langle \cdot, \cdot \rangle)$ einen *euklidischen Raum*; ist $\mathbb{K} = \mathbb{C}$, so nennen wir es einen *unitären Raum*.

Beispiel A14.2 (Skalarprodukte)

a. Wir nennen

$$\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} : (x, y) \mapsto x^t \circ y = \sum_{i=1}^{n} x_i \cdot y_i$$

das *kanonische Skalarprodukt* oder *Standardskalarprodukt* auf $\mathbb{R}^n$. $\left(\mathbb{R}^n, +, \cdot, \langle \cdot, \cdot \rangle\right)$ ist ein euklidischer Raum.

b. Analog definieren wir das *kanonische Skalarprodukt* auf $\mathbb{C}^n$ durch

$$\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \longrightarrow \mathbb{C} : (x, y) \mapsto \overline{x}^t \circ y = x^* \circ y = \sum_{i=1}^{n} \overline{x_i} \cdot y_i,$$

und $\left(\mathbb{C}^n, +, \cdot, \langle \cdot, \cdot \rangle\right)$ ist ein unitärer Raum.

c. 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 A14.32) und $\big(\mathcal{C}([0,1],\mathbb{R}),+,\cdot,\langle\cdot,\cdot\rangle\big)$ ist ein euklidischer Raum.

Definition A14.3 (Symmetrische und hermitesche Matrizen)

a. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ heißt *symmetrisch*, falls $A = A^t$.

b. Für eine Matrix $A = (a_{ij})_{i,j} \in \mathrm{Mat}_n(\mathbb{C})$ heißt die Matrix
$$A^* = \overline{A}^t = \big(\overline{a_{ji}}\big)_{i,j} \in \mathrm{Mat}_n(\mathbb{C})$$
die *adjungierte Matrix* zu A.

c. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ heißt *hermitesch* oder *selbstadjungiert*, falls $A = A^*$.

Beachte, eine reelle Matrix ist genau dann symmetrisch, wenn sie hermitesch ist.

Bemerkung A14.4

Ist $A \in \mathrm{Mat}_n(\mathbb{K})$, so erfüllt die Abbildung
$$\langle\cdot,\cdot\rangle_A : \mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K} : (x,y) \mapsto x^* \circ A \circ y = \overline{x}^t \circ A \circ y$$
die Bedingnung (1) eines Skalarproduktes, weil die Matrixmultiplikation linear ist. Ist die Matrix A zudem hermitesch, so erfüllt $\langle\cdot,\cdot\rangle_A$ auch die Bedingung (2) eines Skalarproduktes, weil
$$\langle x,y\rangle_A = \overline{x}^t \circ A \circ y = (\overline{x}^t \circ A \circ y)^t = \overline{\overline{y}^t \circ \overline{A}^t \circ x} = \overline{\overline{y}^t \circ A \circ x} = \overline{\langle y,x\rangle_A}.$$
Betrachten wir konkret $\mathbb{K} = \mathbb{R}$ und die symmetrische Matrix
$$A = \begin{pmatrix} 1 & 1 \\ 1 & 4 \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R}),$$
dann gilt zudem für $x = (x_1,x_2)^t \in \mathbb{R}^2$ mit $x \neq (0,0)$
$$\langle x,x\rangle_A = x_1^2 + 2x_1x_2 + 4x_2^2 = (x_1+x_2)^2 + 3x_2^2 > 0$$
und damit Bedinung (3), so daß $\langle\cdot,\cdot\rangle_A$ in dem Fall ein Skalarprodukt auf $\mathbb{R}^3$ ist.

Bemerkung A14.5

Wir behandeln im Folgenden oft euklidische und eines unitäre Räume parallel. Dabei machen wir uns zunutze, daß für eine reelle Zahl λ gilt $\lambda = \overline{\lambda}$. Mithin gilt in einem *reellen* Vektorraum V
$$\langle \lambda x + \mu y, z\rangle = \lambda\langle x,z\rangle + \mu\langle y,z\rangle \quad\Longleftrightarrow\quad \langle \lambda x + \mu y, z\rangle = \overline{\lambda}\langle x,z\rangle + \overline{\mu}\langle y,z\rangle,$$
und für $A \in \mathrm{Mat}(n,\mathbb{R})$ gilt genau dann $A = A^t$, wenn $A = \overline{A}^t$ erfüllt ist.

B) Die euklidische Norm zu einem Skalarprodukt

Definition A14.6 (Normierter Raum)
Es sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung

$$||\cdot|| : V \to \mathbb{R}_{\geq 0},$$

heißt eine *Norm* auf V, falls für alle $x, y \in V$ und $\lambda \in \mathbb{K}$

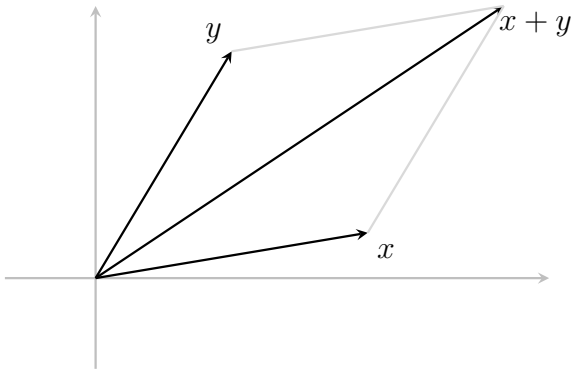
(1) $||x|| = 0 \quad \Leftrightarrow \quad x = 0,$ ("Positive Definitheit")

(2) $||\lambda x|| = |\lambda| \cdot ||x||$, und ("Homogenität")

(3) $||x + y|| \leq ||x|| + ||y||.$ ("Dreiecksungleichung")

Das Quadrupel $\big(V, +, \cdot, ||\cdot||\big)$ heißt dann ein *normierter Raum*.

Bemerkung A14.7 (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 und Abstände von Punkten messen soll, nicht unbillig. Und in der Tat reichen diese Forderungen auch bereits aus, um einen vernünftigen Längenbegriff zu erhalten.

Satz A14.8 (Cauchy-Schwarzsche Ungleichung)
Ist V ein euklidischer oder unitärer Raum, dann gilt für alle $x, y \in V$

$$\big|\langle x, y\rangle\big| \leq \sqrt{\langle x, x\rangle} \cdot \sqrt{\langle y, y\rangle}, \tag{233}$$

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 $\lambda \in \mathbb{K}$

$$0 \leq \langle x - \lambda y, x - \lambda y\rangle = \langle x, x\rangle - \overline{\lambda}\overline{\langle x, y\rangle} - \lambda\langle x, y\rangle + \overline{\lambda}\lambda\langle y, y\rangle. \tag{234}$$

Wählen wir nun speziell $\lambda = \frac{\overline{\langle x,y\rangle}}{\langle y,y\rangle} \in \mathbb{K}$, dann folgt

$$\begin{aligned} 0 &\leq \langle x, x\rangle - \tfrac{\langle x,y\rangle}{\langle y,y\rangle} \cdot \overline{\langle x, y\rangle} - \tfrac{\overline{\langle x,y\rangle}}{\langle y,y\rangle} \cdot \langle x, y\rangle + \tfrac{\overline{\langle x,y\rangle}}{\langle y,y\rangle}\tfrac{\langle x,y\rangle}{\langle y,y\rangle} \cdot \langle y, y\rangle \\ &= \langle x, x\rangle - \tfrac{|\langle x,y\rangle|^2}{\langle y,y\rangle} - \tfrac{|\langle x,y\rangle|^2}{\langle y,y\rangle} + \tfrac{|\langle x,y\rangle|^2}{\langle y,y\rangle}, \end{aligned}$$

also

$$\big|\langle x, y\rangle\big|^2 \leq \langle x, x\rangle \cdot \langle y, y\rangle. \tag{235}$$

Durch Ziehen der positiven Wurzel folgt die gesuchte Ungleichung (233).

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 (234) das Gleichheitszeichen gilt. Dieses λ ist eindeutig bestimmt, und erfüllt

$$\lambda = \lambda \cdot \frac{\langle y, y\rangle}{\langle y, y\rangle} = \frac{\langle y, \lambda y\rangle}{\langle y, y\rangle} = \frac{\langle y, x\rangle}{\langle y, y\rangle} = \frac{\overline{\langle x, y\rangle}}{\langle y, y\rangle}.$$

Damit ist die Gleichheit in (234) gleichwertig zur Gleichheit in (235). □

Satz A14.9 (Die euklidische Norm zu einem Skalarprodukt)
Es sei V ein euklidischer oder unitärer Raum. Dann wird durch

$$||\cdot|| : V \to \mathbb{R}_{\geq 0} : x \mapsto \sqrt{\langle x, x\rangle}$$

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:

$$\big|\langle x, y\rangle\big| \leq ||x|| \cdot ||y||.$$

Beweis: Seien $x \in V$ und $\lambda \in \mathbb{K}$. Aus der positiven Definitheit von $\langle \cdot, \cdot\rangle$ folgt, daß $\langle x, x\rangle \geq 0$ und somit $||x||$ definiert und stets nicht-negativ ist. Ferner folgt, daß $||x|| = 0$ genau dann gilt, wenn x der Nullvektor ist. Aus der Bilinearität bzw. Sesquilinearität von $\langle \cdot, \cdot\rangle$ leiten wir her, daß

$$\langle \lambda x, \lambda x\rangle = \overline{\lambda}\lambda\langle x, x\rangle = |\lambda|^2\langle x, x\rangle,$$

und somit $||\lambda x|| = |\lambda| \cdot ||x||$.

Allein, die Dreiecksungleichung ist etwas schwieriger zu zeigen. Wir verwenden hierfür die Cauchy-Schwarzsche Ungleichung aus Satz A14.8. Beachten wir noch, daß für eine

komplexe Zahl $c \in \mathbb{C}$ stets

$$\mathrm{Re}(c) = \tfrac{1}{2} \cdot \big(c + \overline{c}\big)$$

gilt, so erhalten wir für $x, y \in V$

$$\begin{aligned} ||x+y||^2 &= \langle x+y, x+y\rangle = \langle x,x\rangle + \langle x,y\rangle + \overline{\langle x,y\rangle} + \langle y,y\rangle \\ &= ||x||^2 + 2 \cdot \mathrm{Re}\big(\langle x,y\rangle\big) + ||y||^2 \\ &\leq ||x||^2 + 2 \cdot \big|\langle x,y\rangle\big| + ||y||^2 \\ &\leq ||x||^2 + 2 \cdot ||x|| \cdot ||y|| + ||y||^2 \\ &= \big(||x|| + ||y||\big)^2. \end{aligned}$$

Hieraus folgt dann die Dreiecksungleichung. □

Bemerkung A14.10 (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 (233) folgt für $0 \neq x, y \in V$

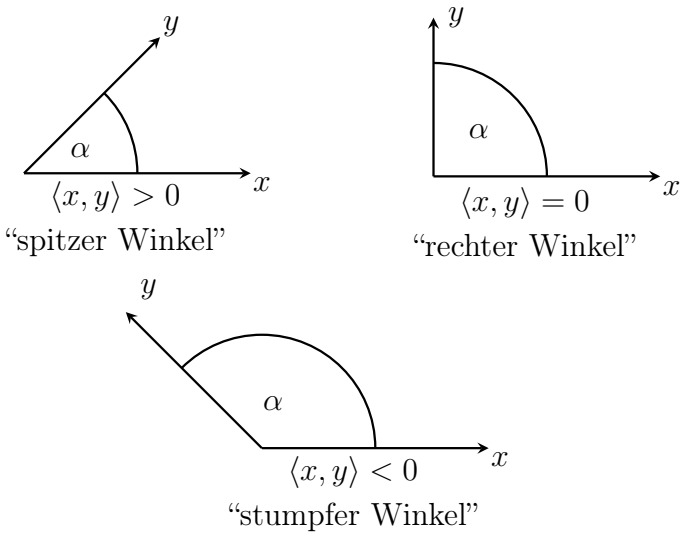
$$-1 \leq \frac{\langle x,y\rangle}{||x|| \cdot ||y||} \leq 1. \tag{236}$$

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

$$\measuredangle(x,y) = \arccos\left(\frac{\langle x,y\rangle}{||x|| \cdot ||y||}\right) \in [0,\pi]$$

als den *Winkel* zwischen x und y.

Ist $\langle x,y\rangle > 0$, also $\measuredangle(x,y) \in \big[0, \frac{\pi}{2}\big[$, so spricht man von einem *spitzen Winkel*.
Ist $\langle x,y\rangle < 0$, also $\measuredangle(x,y) \in \big]\frac{\pi}{2}, \pi\big]$, so spricht man von einem *stumpfen Winkel*.
Ist $\langle x,y\rangle = 0$, also $\measuredangle(x,y) = \frac{\pi}{2}$, so spricht man von einem *rechten Winkel*.



C) Orthonormalbasen und Parsevalsche Gleichung

Definition A14.11 (Orthogonal)
Sei V ein euklidischer oder unitärer Raum, $x, y \in V$, $M, N \subseteq V$ und $U \leq V$.

a. x heißt *orthogonal* zu y, falls $\langle x, y\rangle = 0$. Wir schreiben dann $x \perp y$.

b. M heißt *orthogonal* zu N, falls $m \perp n$ für alle $m \in M$ und $n \in N$. Wir schreiben dann $M \perp N$.

c. Wir nennen $U^\perp := \{z \in V \mid z \perp U\}$ das *orthogonale Komplement* von U.

Lemma A14.12 (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 $z \in U$

$$\langle \lambda x + \mu y, z\rangle = \overline{\lambda}\langle x, z\rangle + \overline{\mu}\langle y, z\rangle = 0,$$

Also $\lambda x + \mu y \in U^\perp$. Damit ist $U^\perp$ ein Unterraum von V. □

Definition A14.13 (Orthonormalbasis)
Sei V ein euklidischer oder unitärer Raum und $B = (x_i \mid i \in I)$ eine Familie in V.

a. B heißt *orthogonal*, falls $x_i \perp x_j$ für alle $i, j \in I$ mit $i \neq j$ gilt.

b. B heißt *orthonormal*, falls B orthogonal ist und $||x_i|| = 1$ für alle $i \in I$ gilt.

c. Ist B eine Basis und orthonormal, so heißt B eine *Orthonormalbasis*, kurz *ONB*.

Beispiel A14.14 (ONB)

Betrachten wir $\mathbb{K}^n$ mit dem kanonischen Skalarprodukt, dann ist die kanonische Basis $E = (e_1, \ldots, e_n)$ offenbar eine Orthonormalbasis von $\mathbb{K}^n$, da $\langle e_i, e_j \rangle = \delta_{ij}$ für $i, j \in \{1, \ldots, n\}$.

Lemma A14.15 (Orthogonal impliziert linear unabhängig.)

Sei V ein euklidischer oder unitärer Raum und $B = (x_i \mid i \in I)$ eine Familie paarweise orthogonaler Vektoren 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 A14.16 (Parsevalsche Gleichung)

Sei V ein euklidischer oder unitärer Raum, $B = (x_i \mid i \in I)$ eine ONB und $x \in V$, dann gilt

$$x = \sum_{\substack{i \in I \\ \text{endlich}}} \langle x_i, x \rangle \cdot x_i. \tag{237}$$

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

$$\langle x_j, x \rangle = \Big\langle x_j, \sum\nolimits_{\substack{i \in I \\ \text{endlich}}} \lambda_i x_i \Big\rangle = \sum_{\substack{i \in I \\ \text{endlich}}} \lambda_i \langle x_j, x_i \rangle = \lambda_j \langle x_j, x_j \rangle = \lambda_j.$$

□

Bemerkung A14.17

Ist B eine ONB von V, so erlaubt es die Gleichung (237), 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 A14.18 (Gram-Schmidt)
Sei $B = (x_1, \ldots, x_n)$ eine Basis des endlich-dimensionalen euklidischen oder unitären Raums V, dann gibt es eine ONB $B' = (z_1, \ldots, z_n)$ von V, so daß

$$\mathrm{Lin}(x_1, \ldots, x_r) = \mathrm{Lin}(z_1, \ldots, z_r)$$

für jedes $r = 1, \ldots, n$ gilt.

Beweis: Wir konstruieren die z_r rekursiv und beginnen mit

$$z_1 := \frac{x_1}{||x_1||},$$

Wir können nun voraussetzen, dass wir die orthnormale Familie $(z_1, \ldots, z_r)$ bereits konstruiert haben, und konstruieren uns den Vektor z_{r+1} passend.

Dazu setzen wir zunächst

$$y_{r+1} := x_{r+1} - \sum_{i=1}^{r} \langle z_i, x_{r+1} \rangle \cdot z_i. \tag{238}$$

Da x_{r+1} linear unabhängig von B ist, ist $y_{r+1} \neq 0$, und wir können deshalb

$$z_{r+1} := \tfrac{1}{||y_{r+1}||} \cdot y_{r+1} \tag{239}$$

setzen. Dann ist $||z_{r+1}|| = 1$ und außerdem gilt für $i = 1, \ldots, r$

$$\begin{aligned}
\langle z_i, z_{r+1} \rangle &= \frac{1}{||y_{r+1}||} \cdot \langle z_i, y_{r+1} \rangle \\
&= \frac{1}{||y_{r+1}||} \cdot \left(\langle z_i, x_{r+1} \rangle - \sum_{j=1}^{r} \langle z_j, x_{r+1} \rangle \cdot \langle z_i, z_j \rangle \right) \\
&= \frac{1}{||y_{r+1}||} \cdot \big(\langle z_i, x_{r+1} \rangle - \langle z_i, x_{r+1} \rangle \big) = 0.
\end{aligned}$$

Dann ist aber $(z_1, \ldots, z_{r+1})$ orthonormal. Setzen wir das Vefahren bis $r = n$ fort, so erhalten wir die gesuchte ONB B'. □

Der Beweis von Satz A14.18 ist konstruktiv und wird auch das *Gram-Schmidtsche Orthonormalisierungsverfahren* genannt. Es erlaubt, aus einem gegebenen Erzeugendensystem eine ONB zu konsturieren.

Algorithmus A14.19 (Gram-Schmidt-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 $B = (x_1, \ldots, x_r)$ von $\mathrm{Lin}(M)$, z. B. mittels Algorithmus A4.16.

2. Schritt: Für $i = 1, \ldots, r$ 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 $(z_1, \ldots, z_r)$ zurück.

Bemerkung A14.20

a. Will man in der Praxis ein Skalarprodukt übergeben, so wird man im reellen Fall eine symmetrische Matrix A übergeben und im komplexen Fall eine hermitesche (siehe Bemerkung A14.4). Das Skalarprodukt wird dann durch

$$\langle x, y \rangle = x^t \circ A \circ \overline{y}$$

gebildet.

b. Um zu normieren, ist in Algorithmus A14.19 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

$$y_i = x_i - \sum_{j=1}^{i-1} \frac{\langle y_j, x_i \rangle}{\langle y_j, y_j \rangle} \cdot y_j,$$

dann liefert Algorithmus A14.19 eine Basis $(y_1, \ldots, y_r)$ paarweise orthogonaler Vektoren von $\mathrm{Lin}(M)$. Das hat den Vorteil, daß man exakt rechnen kann - etwa in SINGULAR, wenn die Eingabedaten rationale Zahlen waren.

Beispiel A14.21 (Gram-Schmidt-Orthonormalisierungsverfahren)
Es sei $B = (x_1, x_2, x_3) = \{(1,0,1)^t, (1,1,1)^t, (0,0,4)^t\} \subseteq \mathbb{R}^3$, 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 $y_1 := (1,0,1)^t$, dann ist $\langle y_1, y_1 \rangle = 2$ und somit ersetzen wir x_1 in B durch

$$z_1 = \frac{1}{||y_1||} \cdot y_1 = \tfrac{1}{\sqrt{2}} \cdot y_1 = \left(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)^t.$$

Im nächsten Schritt setzen wir

$$y_2 = x_2 - \langle z_1, x_2 \rangle \cdot z_1 = (1,1,1)^t - \tfrac{2}{\sqrt{2}} \cdot \left(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)^t = (0,1,0)^t.$$

Dann ist $\langle y_2, y_2 \rangle = 1$ und somit ersetzen wir x_2 in B durch $z_2 = y_2$.

Schließlich bilden wir

$$\begin{aligned} y_3 &= x_3 - \langle z_1, x_3\rangle \cdot z_1 - \langle z_2, x_3\rangle \cdot z_2 \\ &= (0,0,4)^t - \tfrac{4}{\sqrt{2}} \cdot \left(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)^t - 0 \cdot (0,1,0)^t \\ &= (-2,0,2)^t, \end{aligned}$$

und erhalten $\langle y_3, y_3\rangle = 8$. Somit müssen wir x_3 durch den Vektor

$$z_3 = \tfrac{1}{\|y_3\|} \cdot y_3 = \tfrac{1}{\sqrt{8}} \cdot (-2,0,2)^t = \left(-\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)^t$$

ersetzen. Damit ergibt sich aus dem Gram-Schmidtschen Orthonormalisierungsverfahren die ONB

$$\left(\left(\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)^t, (0,1,0)^t, \left(-\tfrac{1}{\sqrt{2}}, 0, \tfrac{1}{\sqrt{2}}\right)^t\right).$$

E) Orthogonale und unitäre Matrizen

Schreibt man die Vektoren einer Orthonormalbasis als Spalten in eine Matrix, so erhält man eine invertierbare Matrix, deren Inverse sich durch einfaches Transponieren und komplexes Konjugieren berechnen läßt. Dies führt zum Begriff der orthognalen bzw. unitären Matrizen.

Definition A14.22 (Orthogonale / unitäre Matrizen)

a. Für eine Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ heißt $A^* := \overline{A}^t$ die zu A *adjungierte Matrix*.

b. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ heißt *orthogonal*, wenn $A^t \circ A = \mathbb{1}_n$ gilt. Wir nennen $\mathrm{O}(n) := \{A \in \mathrm{Mat}_n(\mathbb{R}) \mid A \text{ ist orthogonal}\}$ *orthogonale Gruppe* vom Grad n.

c. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ heißt *unitär*, wenn $A^* \circ A = \mathbb{1}_n$ gilt, und wir nennen $\mathrm{U}(n) := \{A \in \mathrm{Mat}_n(\mathbb{C}) \mid A \text{ ist unitär}\}$ die *unitäre Gruppe* vom Grad n.

Proposition A14.23 (Determinante orthogonaler / unitärer Matrizen)
Ist $A \in \mathrm{Mat}_n(\mathbb{K})$ 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 A14.24 (Orthogonale / unitäre Matrizen)
Für eine quadratische Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ sind äquivalent:

a. A ist orthogonal bzw. unitär.

b. A ist invertierbar mit $A^* = A^{-1}$.

c. Die Spalten von A sind eine ONB von $\mathbb{K}^n$ mit kanonischem Skalarprodukt.

d. 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 a^i der i-te Spaltenvektor von A, so ist $\overline{a^i}^t$ der i-te Zeilenvektor von A^* und deshalb ist

$$\langle a^i, a^j \rangle = \overline{a^i}^t \circ a^j$$

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 a_i der i-te Zeilenvektor von A, so ist $\overline{a_i}^t$ der i-te Spaltenvektor von A^*. Also ist

$$\overline{\langle a_i^t, a_j^t \rangle} = a_i \circ \overline{a_j}^t$$

der Eintrag von $A \circ A^*$ an der Stelle (i,j). Dies zeigt schließlich, daß b. und d. äquivalent sind. □

Beispiel A14.25 (Orthogonale Matrix)
Die Matrix

$$\begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$$

ist orthogonal, da ihre Spalten nach Beispiel A14.21 eine ONB von $\mathbb{R}^3$ bezüglich des kanonischen Skalarproduktes sind.

Korollar A14.26 (Die orthogonale und die unitäre Gruppe)
$(\mathrm{O}(n), \circ)$ und $(\mathrm{U}(n), \circ)$ sind Gruppen.

Beweis: Es reicht, zu zeigen, daß sie Untergruppen von $\mathrm{Gl}_n(\mathbb{C})$ sind. Offenbar sind $\mathrm{O}(n)$ und $\mathrm{U}(n)$ nicht-leere Teilmengen von $\mathrm{Gl}_n(\mathbb{C})$. Sind nun A und B in $\mathrm{O}(n)$ bzw. in $\mathrm{U}(n)$, so gilt

$$(A \circ B)^* = B^* \circ A^* = B^{-1} \circ A^{-1} = (A \circ B)^{-1},$$

und

$$\left(A^{-1}\right)^* = A^{**} = A = \left(A^{-1}\right)^{-1}.$$

Mithin liegen auch $A \circ B$ und A^{-1} in $\mathrm{O}(n)$ bzw. in $\mathrm{U}(n)$. Damit ist gezeigt, daß $\mathrm{O}(n)$ und $\mathrm{U}(n)$ Untergruppen von $\mathrm{Gl}_n(\mathbb{C})$ sind. □

Bemerkung A14.27 (Die orthogonale Gruppe $\mathrm{O}(2)$)
Die Determinante

$$\det : \mathrm{O}(2) \longrightarrow \{1, -1\}$$

ist wegen des Determinantenmultiplikationssatzes ein Gruppenepimorphismus. Der Kern von det ist der Normalteiler

$$\mathrm{SO}(2) := \{A \in \mathrm{O}(2) \mid \det(A) = 1\}$$

von $\mathrm{O}(2)$ und wird die *spezielle orthogonale Gruppe* vom Grad 2 genannt. Wir werden unten zeigen, daß

$$\mathrm{SO}(2) = \{T(\alpha) \mid \alpha \in [0, 2\pi)\}$$

und

$$\mathrm{O}(2) \setminus \mathrm{SO}(2) = \{S(\alpha) \mid \alpha \in [0, 2\pi)\},$$

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 $\mathrm{Lin}\left(\left(\cos\left(\frac{\alpha}{2}\right), \sin\left(\frac{\alpha}{2}\right)\right)^t\right)$. Insbesondere ist im Fall $n = 2$ also jede orthogonale Matrix eine Drehung oder eine Spieglung.

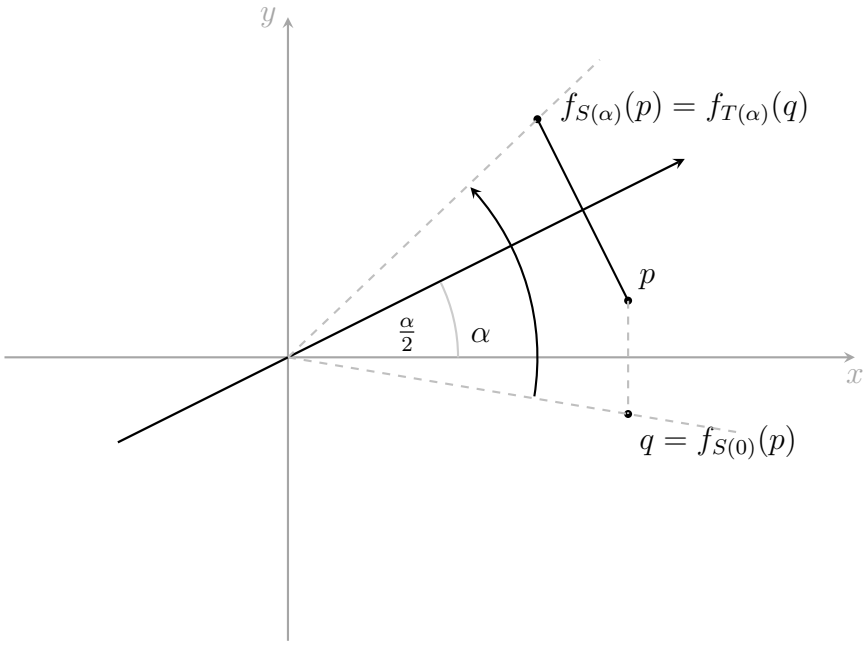


Abbildung 17: Die Spielung $S(\alpha) = T(\alpha) \circ S(0)$

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.

Beweis: Die Matrix

$$A = \begin{pmatrix} a & c \\ b & d \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$$

ist genau dann orthogonal, wenn die beiden Spaltenvektoren $x = (a, b)^t$ und $y = (c, d)^t$ eine Orthonormalbasis von $\mathbb{R}^2$ bezüglich des Standardskalarproduktes sind. Insbesondere muß y also senkrecht auf x stehen. In der Ebene ist ein Vektor, der senkrecht steht auf x aber bis auf einen Skalarfaktor eindeutig bestimmt und $(-b, a)^t$ ist ein solcher Vektor. Es muß also

$$y = \lambda \cdot (-b, a)^t$$

gelten. Aus

$$1 = ||y|| = |\lambda| \cdot \sqrt{(-b)^2 + a^2} = |\lambda| \cdot ||x|| = |\lambda|$$

folgt dann $\lambda = 1$ oder $\lambda = -1$. Also ist die Matrix A von der Form

$$A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \quad \text{oder} \quad A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix},$$

wobei die Zahlen $a, b \in \mathbb{R}$ nur die Bedingung

$$a^2 + b^2 = ||x||^2 = 1$$

erfüllen müssen. Aus dem Satz von Pythagoras wissen wir aber, daß es dann genau einen Winkel $\alpha \in [0, 2\pi)$ gibt mit

$$a = \cos(\alpha) \quad \text{und} \quad b = \sin(\alpha),$$

und somit

$$A = T(\alpha) \quad \text{oder} \quad A = S(\alpha).$$

Beachten wir nun noch, daß

$$\det(T(\alpha)) = \cos(\alpha)^2 + \sin(\alpha)^2 = 1$$

und

$$\det(S(\alpha)) = -\cos(\alpha)^2 - \sin(\alpha)^2 = -1$$

ist, so ist

$$\mathrm{SO}(2) = \{T(\alpha) \mid \alpha \in [0, 2\pi)\}$$

und

$$\mathrm{O}(2) \setminus \mathrm{SO}(2) = \{S(\alpha) \mid \alpha \in [0, 2\pi)\},$$

gezeigt. □

F) Orthogonale Summe und orthogonale Projektion

Definition A14.28 (Orthogonale Summe)
Sei V ein euklidischer oder unitärer Raum. Wir nennen V die *orthogonale Summe* der Unterräume $U_1, \ldots, U_r$, falls $V = U_1 \oplus \ldots \oplus U_r$ und $U_i \perp U_j$ für $i \neq j$. In diesem Fall schreiben wir $V = U_1 \perp \ldots \perp U_r$.

Proposition A14.29 (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 A14.18 eine ONB $(x_1, \ldots, x_r)$ von U wählen und diese zu einer ONB $(x_1, \ldots, x_n)$ von V ergänzen. Dann gilt aber

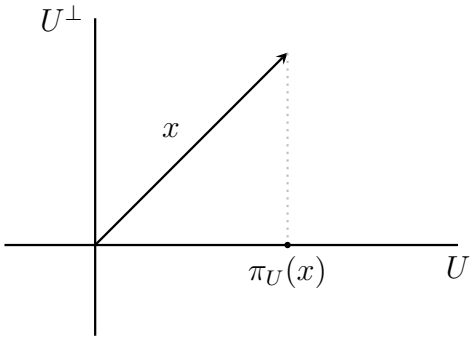
$$V = \mathrm{Lin}\,(x_1, \ldots, x_r) + \mathrm{Lin}\,(x_{r+1}, \ldots, x_n) \subseteq U + U^\perp,$$

da nach Wahl $x_{r+1}, \ldots, x_n \in U^\perp$. Hierbei beachte man, daß ein Vektor, der orthogonal zu einer Basis von U ist, automatisch orthogonal zu jedem Vektor in U ist. □

Bemerkung A14.30
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

$$\pi_U : V \to V$$

definieren durch $\pi\big(u + u'\big) = u$ für $u \in U$ und $u' \in U^\perp$.



In Aufgabe A14.33 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 $\mathrm{Im}(\pi_U) = U$ das Bild von π_U und $\mathrm{Ker}(\pi_U) = U^\perp$ der Kern.

Man beachte auch, daß beim Verfahren von Gram-Schmidt der Vektor y_{r+1} aus dem Vektor x_{k+1} durch orthogonale Projektion auf den Unterraum $\mathrm{Lin}(z_1, \ldots, z_r)$ konstruiert wird (siehe Beweis von Satz A14.18).

Bemerkung A14.31 (Determinante als Volumenform)
In Bemerkung A9.12 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 18).

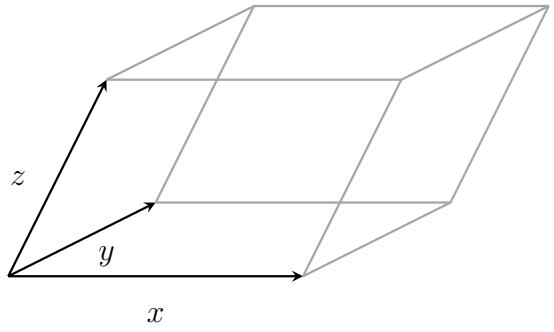


Abbildung 18: Das Parallelotop $P(x, y, z)$ im $\mathbb{R}^3$

Wir wollen zeigen, daß das Volumen

$$\text{Volumen}\,\big(P(x, y, z)\big) = |\det(x \;\; y \;\; z)|$$

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 $U = \mathrm{Lin}\,(x)$ 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 19).

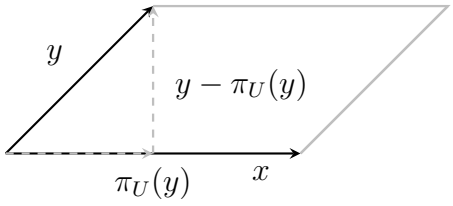


Abbildung 19: Die Höhe im Parallelogramm zu x und y.

Für die Fläche A des Parallelogramms gilt deshalb

$$A = ||x|| \cdot ||y - \pi_U(y)||.$$

Auf ähnliche Weise kann man die Höhe h des Parallelotops $P(x, y, z)$ bestimmen. Hierzu betrachten wir den Unterraum $W = \mathrm{Lin}\,(x, y)$ und die orthogonale Projektion π_W auf W. Dann ist h die Länge des Vektors $z - \pi_W(z)$ (siehe Abbildung 20).

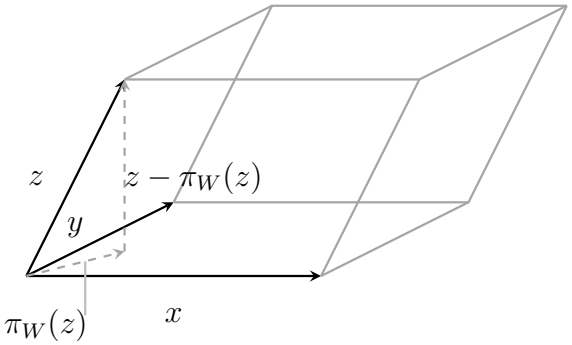


Abbildung 20: Die Höhe in $P(x, y, z)$

Für das Volumen von $P(x, y, z)$ erhalten wir deshalb

$$\text{Volumen}\,\big(P(x, y, z)\big) = A \cdot h = ||x|| \cdot ||y - \pi_U(y)|| \cdot ||z - \pi_W(z)||.$$

Wegen $\pi_U(y) \in U = \mathrm{Lin}\,(x)$ und $\pi_W(z) \in W = \mathrm{Lin}\,(x, y)$, gibt es $\lambda, \mu, \nu \in \mathbb{R}$ 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 19 und 20). 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

$$\begin{aligned} |\det(x \;\; y \;\; z)| &= ||x|| \cdot ||y - \pi_U(y)|| \cdot ||z - \pi_W(z)|| \cdot \left|\det\left(\tfrac{x}{||x||} \;\; \tfrac{y-\pi_U(y)}{||y-\pi_U(y)||} \;\; \tfrac{z-\pi_W(z)}{||z-\pi_W(z)||}\right)\right| \\ &= ||x|| \cdot ||y - \pi_U(y)|| \cdot ||z - \pi_W(z)||. \end{aligned}$$

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 A14.32

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 A14.33 (Orthogonale Projektion)

Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$.

a. Zeige, $\pi_U \in \mathrm{End}_{\mathbb{K}}(V)$ ist eine Projektion mit $\mathrm{Ker}(\pi_U) = U^\perp$ und $\mathrm{Im}(\pi_U) = U$.

b. Zeige, ist $\pi \in \mathrm{End}_{\mathbb{K}}(V)$ eine Projektion mit $\mathrm{Ker}(\pi) = U^\perp$ und $\mathrm{Im}(\pi) = U$, dann ist $\pi = \pi_U$.

c. Ist $(x_1, \ldots, x_r)$ 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 A14.34

Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $U \leq V$. Dann gilt $(U^\perp)^\perp = U$.

Aufgabe A14.35

Zeige, durch $\left\langle (x_1, x_2, x_3)^t, (y_1, y_2, y_3)^t \right\rangle := x_1y_1 + x_1y_2 + x_2y_1 + 2x_2y_2 + x_2y_3 + x_3y_2 + 2x_3y_3$ 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 A14.36

Bestimme eine Orthonormalbasis des $\mathbb{R}$-Vektorraums $U = \{(v, w, x, y, z)^t \in \mathbb{R}^5 \mid v + w + x + y + z = 0\}$ bezüglich des Standardskalarproduktes auf $\mathbb{R}^5$ und berechne die orthogonale Projektion des Vektors $x = (1, 2, 3, 4, 5)^t$ auf den Unterraum U.

Aufgabe A14.37 (Legendre-Polynome)
Betrachte den Vektorraum $U = \{f : [0,1] \to \mathbb{R} : x \mapsto ax^2 + bx + c \mid a, b, c \in \mathbb{R}\}$ und bestimme eine ONB bezüglich des Skalarproduktes aus Aufgabe A14.32.

Aufgabe A14.38 (Tschebyscheff-Polynome)

a. Zeige, daß auf $V = \mathcal{C}([-1,1], \mathbb{R})$ durch
$$V \times V \longrightarrow \mathbb{R} : (f,g) \mapsto \langle f, g\rangle := \int_{-1}^{1} f(x) \cdot g(x) \cdot \frac{1}{\sqrt{1-x^2}}\, dx$$
ein Skalarprodukt definiert wird.

b. Berechne für den Unterraum $U\{f : [-1,1] \to \mathbb{R} : x \mapsto ax^2 + bx + c \mid a, b, c \in \mathbb{R}\}$ von V eine Orthonormalbasis bezüglich des Skalarproduktes aus Teil a..

Hinweis, substituiere in $x = \cos(t)$, um die Konvergenz des uneigentlichen Integrals zu zeigen.

Aufgabe A14.39
Für $V = \mathrm{Mat}_n(\mathbb{R})$ definieren wir $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R} : (A, B) \mapsto \mathrm{Spur}\left(A^t \circ B\right)$.

a. Zeige, $\langle \cdot, \cdot \rangle$ ist ein Skalarprodukt auf V.

b. Zeige, für $U = \{A \in V \mid A^t = A\}$ gilt $U^\perp = \{A \in V \mid A^t = -A\}$.

Aufgabe A14.40
Sei V ein euklidischer oder unitärer Raum und $||\cdot|| : V \to \mathbb{R} : x \mapsto \sqrt{\langle x, x\rangle}$ die durch das Skalarprodukt definierte Norm. Zeige, für $x, y \in V$ gelten:

a. Die Parallelogrammgleichung: $||x+y||^2 + ||x-y||^2 = 2 \cdot \left(||x||^2 + ||y||^2\right)$.

b. Der Satz des Pythagoras': $x \perp y \implies ||x||^2 + ||y||^2 = ||x+y||^2$.

Aufgabe A14.41
Zeige, für $a, b \in \mathbb{R}$ gibt es genau dann zwei normierte Vektoren $x = (u, v, a)^t \in \mathbb{R}^3$ und $y = (r, s, b)^t \in \mathbb{R}^3$ die bezüglich des Standardskalarproduktes orthogonal zueinander sind, wenn $a^2 + b^2 \leq 1$.

Aufgabe A14.42
Zeige, dass durch
$$\langle (x_1, x_2)^t, (y_1, y_2)^t \rangle = 2x_1y_1 + x_2y_1 + x_1y_2 + x_2y_2$$
für $(x_1, x_2)^t, (y_1, y_2)^t \in \mathbb{R}^2$ ein Skalarprodukt auf $\mathbb{R}^2$ definiert wird.

Aufgabe A14.43

Bestimme eine Orthonormalbasis des $\mathbb{R}$-Vektorraums

$$U = \{(x, y, z)^t \in \mathbb{R}^3 \mid x + y + 2z = 0\}$$

bezüglich des Standardskalarproduktes auf $\mathbb{R}^3$.

Aufgabe A14.44

Berechne eine orthogonale Matrix $A \in \mathrm{O}(4)$, die den Vektor $\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)^t$ als erste Spalte enthält.

Aufgabe A14.45 (Der p-adische Betrag)

Sei p eine Primzahl. Für $0 \neq a \in \mathbb{Z}$ bezeichne $\nu_p(a)$ die höchste Potenz von p, die a teilt, und für $0 \neq q = \frac{b}{c} \in \mathbb{Q}$ setzen wir $\nu_p(q) = \nu_p(b) - \nu_p(c)$. Zeige, die Abbildung

$$|\cdot|_p : \mathbb{Q} \longrightarrow \mathbb{R} : q \mapsto \begin{cases} p^{-\nu_p(q)}, & \text{wenn } q \neq 0, \\ 0, & \text{wenn } q = 0. \end{cases}$$

ist positiv definit, multiplikativ und genügt der Dreiecksungleichung.

A15 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 A15.1 (Die adjungierte Abbildung)
Zu jedem Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ gibt es genau ein $f^* \in \mathrm{End}_{\mathbb{K}}(V)$, so daß

$$\langle f(x), y \rangle = \langle x, f^*(y) \rangle \tag{240}$$

für alle $x, y \in V$ gilt. Ist $B = (x_1, \ldots, x_n)$ eine ONB von V, so gilt für $y \in V$

$$f^*(y) = \sum_{i=1}^{n} \langle f(x_i), y \rangle \cdot x_i. \tag{241}$$

Die Abbildung f^* heißt die *adjungierte Abbildung* zu f.

Beweis: Wir wollen zunächst zeigen, daß es einen Endomorphismus $f^* \in \mathrm{End}_{\mathbb{K}}(V)$ mit der Eigenschaft (240) gibt. Dazu wählen wir eine ONB $B = (x_1, \ldots, x_n)$ 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 (241). 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 A14.16 gilt dann

$$\begin{aligned}
\langle f(x), y \rangle &\overset{A14.16}{=} \left\langle f\left(\sum_{i=1}^{n} \langle x_i, x \rangle \cdot x_i \right), y \right\rangle = \left\langle \sum_{i=1}^{n} \langle x_i, x \rangle \cdot f(x_i), y \right\rangle \\
&= \sum_{i=1}^{n} \overline{\langle x_i, x \rangle} \cdot \langle f(x_i), y \rangle = \sum_{i=1}^{n} \langle x, x_i \rangle \cdot \langle f(x_i), y \rangle \\
&= \sum_{i=1}^{n} \langle x, \langle f(x_i), y \rangle \cdot x_i \rangle = \left\langle x, \sum_{i=1}^{n} \langle f(x_i), y \rangle \cdot x_i \right\rangle = \langle x, f^*(y) \rangle.
\end{aligned}$$

Damit ist gezeigt, daß der durch (241) definierte Endomorphismus die Gleichung (240) erfüllt.

Es bleibt noch die Eindeutigkeit zu zeigen. Sei dazu $h \in \mathrm{End}_{\mathbb{K}}(V)$ mit

$$\langle f(x), y\rangle = \langle x, h(y)\rangle \tag{242}$$

für alle $x, y \in V$. Wegen der Parsevalschen Gleichung A14.16 gilt für $y \in V$ dann

$$h(y) \overset{A14.16}{=} \sum_{i=1}^{n} \langle x_i, h(y)\rangle \cdot x_i \overset{(242)}{=} \sum_{i=1}^{n} \langle f(x_i), y\rangle \cdot x_i \overset{(241)}{=} f^*(y).$$

Mithin stimmen f^* und h überein, so daß f^* eindeutig bestimmt ist. □

Korollar A15.2 ($f^{**} = f$)
Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$, 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

$$\langle f^*(x), y\rangle = \overline{\langle y, f^*(x)\rangle} \overset{(240)}{=} \overline{\langle f(y), x\rangle} = \langle x, f(y)\rangle.$$

Damit erfüllt f die Bedingung, durch die die Abbildung f^{**} eindeutig festgelegt ist. Also muß $f = f^{**}$ gelten. □

Korollar A15.3 (Matrixdarstellung der adjungierten Abbildung)
Sei B eine ONB von V und $f \in \mathrm{End}_{\mathbb{K}}(V)$, 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 A14.16 gilt

$$\sum_{i=1}^{n} a_{ij} \cdot x_i = f(x_j) \overset{A14.16}{=} \sum_{i=1}^{n} \langle x_i, f(x_j)\rangle \cdot x_i \tag{243}$$

und

$$\sum_{j=1}^{n} b_{ji} \cdot x_j = f^*(x_i) \overset{A14.16}{=} \sum_{j=1}^{n} \langle x_j, f^*(x_i)\rangle \cdot x_j. \tag{244}$$

Da die Darstellung als Linearkombination einer Basis eindeutig ist, erhalten wir

$$\overline{a_{ij}} \overset{(243)}{=} \overline{\langle x_i, f(x_j)\rangle} = \langle f(x_j), x_i\rangle \overset{(240)}{=} \langle x_j, f^*(x_i)\rangle \overset{(244)}{=} b_{ji}.$$

Daraus folgt

$$M_B^B(f^*) = (b_{ji})_{j,i} = \left(\overline{a_{ij}}\right)_{j,i} = \left(\overline{a_{ij}}\right)_{i,j}^t = M_B^B(f)^*.$$

□

Beispiel A15.4 (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 A15.3 die Adjungierte von f.

B) Der Spektralsatz für normale Endomorphismen

Definition A15.5
Es sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ und $A \in \mathrm{Mat}_n(\mathbb{K})$.

a. f heißt *normal*, falls $f^* \circ f = f \circ f^*$.

b. A heißt *normal*, falls $A^* \circ A = A \circ A^*$.

Bemerkung A15.6

a. Jede symmetrische Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ und jede hermitesche Matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ ist normal, denn wegen $A = A^*$ gilt auch

$$A \circ A^* = A \circ A = A^* \circ A.$$

b. Jede orthogonale Matrix $A \in \mathrm{O}(n)$ und jede unitäre Matrix $A \in \mathrm{U}(n)$ 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 A15.7 (Matrixdarstellung normaler Endomorphismen)
Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ und B eine ONB von V.
Genau dann ist f normal, wenn $M_B^B(f)$ normal ist.

Beweis: Ist f normal, so gilt

$$\begin{aligned} M_B^B(f)^* \circ M_B^B(f) &\overset{A15.3}{=} M_B^B(f^*) \circ M_B^B(f) = M_B^B(f^* \circ f) \\ &= M_B^B(f \circ f^*) = M_B^B(f) \circ M_B^B(f^*) \overset{A15.3}{=} M_B^B(f) \circ M_B^B(f)^* \end{aligned}$$

und somit ist $M_B^B(f)$ normal. Ist umgekehrt $M_B^B(f)$ normal, so gilt

$$\begin{aligned} M_B^B(f^* \circ f) &= M_B^B(f^*) \circ M_B^B(f) \overset{A15.3}{=} M_B^B(f)^* \circ M_B^B(f) \\ &= M_B^B(f) \circ M_B^B(f)^* \overset{A15.3}{=} M_B^B(f) \circ M_B^B(f^*) = M_B^B(f \circ f^*). \end{aligned}$$

Dann stimmen aber die Abbildungen $f^* \circ f$ und $f \circ f^*$ überein, und somit ist f normal. □

Beispiel A15.8 (Normale Abbildung)

In Beispiel A15.4 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 A15.9 (Normale Abbildungen)

Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ normal. Genau dann gilt $x \in \mathrm{Eig}(f, \lambda)$, wenn $x \in \mathrm{Eig}\left(f^*, \overline{\lambda}\right)$.

Beweis: Für ein beliebiges $x \in V$ gilt

$$\begin{aligned} \langle f^*(x) - \overline{\lambda}x, f^*(x) - \overline{\lambda}x \rangle &= \langle f^*(x), f^*(x) \rangle - \langle f^*(x), \overline{\lambda}x \rangle - \langle \overline{\lambda}x, f^*(x) \rangle + \langle \overline{\lambda}x, \overline{\lambda}x \rangle \\ &= \langle f \circ f^*(x), x \rangle - \overline{\lambda}\langle f^*(x), x \rangle - \lambda \langle x, f^*(x) \rangle + \lambda\overline{\lambda}\langle x, x \rangle \\ &= \langle f^* \circ f(x), x \rangle - \overline{\lambda}\langle x, f(x) \rangle - \lambda \langle f(x), x \rangle + \lambda\overline{\lambda}\langle x, x \rangle \\ &= \langle f(x), f(x) \rangle - \langle \lambda x, f(x) \rangle - \langle f(x), \lambda x \rangle + \langle \lambda x, \lambda x \rangle \\ &= \langle f(x) - \lambda x, f(x) - \lambda x \rangle. \end{aligned}$$

Da das Skalarprodukt positiv definit ist, gilt somit

$$\begin{aligned} x \in \mathrm{Eig}(f, \lambda) &\iff f(x) - \lambda x = 0 \iff \langle f(x) - \lambda x, f(x) - \lambda x \rangle = 0 \\ &\iff \langle f^*(x) - \overline{\lambda}x, f^*(x) - \overline{\lambda}x \rangle = 0 \iff f^*(x) - \overline{\lambda}x = 0 \\ &\iff x \in \mathrm{Eig}\left(f^*, \overline{\lambda}\right). \end{aligned}$$

□

Satz A15.10 (Spektralsatz für normale Endomorphismen)
Für $f \in \mathrm{End}_{\mathbb{K}}(V)$ sind die folgenden beiden Aussagen gleichwertig:

a. f ist normal und χ_f zerfällt über $\mathbb{K}$ in Linearfaktoren.

b. V besitzt eine ONB aus Eigenvektoren von f.

Insbesondere ist f dann bezüglich einer ONB diagonalisierbar.

Beweis:
b. $\Longrightarrow$ a.: Besitzt V eine ONB B aus Eigenvektoren, so zerfällt χ_f nach Satz A12.22 über $\mathbb{K}$ in Linearfaktoren. Zudem ist dann $M_B^B(f)$ eine Diagonalmatrix, und wegen Korollar A15.3 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 A15.7 ist dann aber auch f normal.

a. $\Longrightarrow$ 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 $U := \mathrm{Lin}\,(x)$ ein f-invarianter Unterraum der Dimension 1 und es gilt

$$V = U \perp U^\perp = U \oplus U^\perp$$

nach Proposition A14.29.

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

$$\langle f(y), x\rangle = \langle y, f^*(x)\rangle \overset{A15.9}{=} \langle y, \overline{\lambda}x\rangle = \overline{\lambda} \cdot \langle y, x\rangle = 0,$$

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 $(x_2, \ldots, x_n)$ aus Eigenvektoren von $f_{U^\perp}$. Dann ist aber $B = (x_1, x_2, \ldots, x_n)$ mit $x_1 = \frac{x}{||x||}$ eine ONB aus Eigenvektoren von f. □

Korollar A15.11 (Spektralsatz für normale Matrizen)
Für eine Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ sind die folgenden beiden Aussagen gleichwertig:

a. A ist normal und χ_A zerfällt über $\mathbb{K}$ in Linearfaktoren.

b. Es gibt ein T in $\mathrm{O}(n)$ bzw. $\mathrm{U}(n)$, so daß $T^{-1} \circ A \circ T$ eine Diagonalmatrix ist.

Beweis: Wenden wir den Spektralsatz A15.10 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 A14.24 daher orthogonal bzw. unitär. □

Der Beweis ist konstruktiv, sofern man die Eigenwerte von A exakt kennt. Man leitet daraus folgenden prinzipiellen Algorithmus zur Bestimmung von T her.

Algorithmus A15.12
INPUT: $A \in \mathrm{Mat}_n(\mathbb{K})$ normal mit χ_A zerfällt über $\mathbb{K}$.
OUTPUT: T in $\mathrm{O}(n)$ bzw. $\mathrm{U}(n)$, 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 A15.13 (Diagonalisierung einer normalen Abbildung)
Die Abbildung f in Beispiel A15.4 ist nach Beispiel A15.8 normal.

$$\chi_f = \det\left(t \cdot \mathbb{1}_2 - M_E^E(f)\right) = \begin{vmatrix} t-2 & -4 \\ 4 & t-2 \end{vmatrix} = t^2 - 4t + 20 = \big(t-(2+4i)\big) \cdot \big(t-(2-4i)\big)$$

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 $\mathrm{Eig}(f, 2+4i)$ und $\mathrm{Eig}(f, 2-4i)$ ausgehend von der Matrixdarstellung $M_E^E(f)$ in Beispiel A15.4.

Für $\mathrm{Eig}(f, 2+4i) = \mathrm{Lös}\left(M_E^E(f) - (2+4i)\cdot \mathbb{1}_2, 0\right)$ liefert unser Algorithmus

$$\begin{pmatrix} -4i & 4 \\ -4 & -4i \end{pmatrix} \overset{\substack{II\mapsto II+i\cdot I \\ I\mapsto \frac{1}{-4i}\cdot I}}{\longmapsto} \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} \overset{\substack{-1\text{'en} \\ \text{ergänzen}}}{\longmapsto} \begin{pmatrix} 1 & i \\ 0 & -1 \end{pmatrix},$$

so daß $x_1 = (i, -1)^t$ eine Basis von $\mathrm{Eig}(f, 2+4i)$ ist. Analog erhalten wir für $\mathrm{Eig}(f, 2-4i) = \mathrm{Lös}\left(M_E^E(f) - (2-4i)\cdot \mathbb{1}_2, 0\right)$

$$\begin{pmatrix} 4i & 4 \\ -4 & 4i \end{pmatrix} \overset{\substack{II\mapsto II-i\cdot I \\ I\mapsto \frac{1}{4i}\cdot I}}{\longmapsto} \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix} \overset{\substack{-1\text{'en} \\ \text{ergänzen}}}{\longmapsto} \begin{pmatrix} 1 & -i \\ 0 & -1 \end{pmatrix},$$

so daß $x_2 = (-i, -1)^t$ eine Basis von $\mathrm{Eig}(f, 2-4i)$ 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

$$B = \left(\tfrac{1}{||x_1||}\cdot x_1, \tfrac{1}{||x_2||}\cdot x_2\right) = \left(\left(\tfrac{i}{\sqrt{2}}, \tfrac{-1}{\sqrt{2}}\right)^t, \left(\tfrac{-i}{\sqrt{2}}, \tfrac{-1}{\sqrt{2}}\right)^t\right).$$

Es ist übrigens kein Zufall, daß die beiden Eigenvektoren im letzten Beispiel orthogonal zueinander standen.

Lemma A15.14
Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ normal, $x \in \mathrm{Eig}(f, \lambda)$, $y \in \mathrm{Eig}(f, \mu)$ und $\lambda \neq \mu$, dann gilt $x \perp y$.

Beweis: Nach Voraussetzung gilt

$$\lambda\cdot\langle x, y\rangle = \langle\overline{\lambda}x, y\rangle \overset{A15.9}{=} \langle f^*(x), y\rangle \overset{A15.2}{=} \langle x, f(y)\rangle = \langle x, \mu y\rangle = \mu\cdot\langle x, y\rangle.$$

Für die Differenz der beiden Seiten erhalten wir dann $(\lambda-\mu)\cdot\langle x, y\rangle = 0$, wobei nach Voraussetzung $\lambda - \mu \neq 0$ gilt. Also ist $\langle x, y\rangle = 0$ und somit $x \perp y$. □

Satz A15.15 (Spektralzerlegung für normale Endomorphismen)
Sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ normal mit zerfallendem charakteristischem Polynom und seien $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ die paarweise verschiedenen Eigenwerte von f. Ferner bezeichne

$$\pi_i : V \longrightarrow V$$

die orthogonale Projektion von V auf $\mathrm{Eig}(f, \lambda_i)$.
Dann ist

$$V = \mathrm{Eig}(f, \lambda_1) \perp \ldots \perp \mathrm{Eig}(f, \lambda_r)$$

die *orthogonale Summe* der Eigenräume von f und es gilt

$$f = \lambda_1 \cdot \pi_1 + \ldots + \lambda_r \cdot \pi_r.$$

Man nennt dies die *Spektralzerlegung* von f.

Beweis: Nach dem Spektralsatz für normale Abbildungen ist f diagonalisierbar. Deshalb folgt aus Satz A12.22, daß V die direkte Summe der Eigenräume von f ist. Wegen Lemma A15.14 ist diese Summe dann eine orthogonale Summe.

Ist nun $x = x_1 + \ldots + x_r \in V$ mit $x_i \in \mathrm{Eig}(f, \lambda_i)$ gegeben, so gilt

$$\pi_j(x_i) = \delta_{ij} \cdot x_i,$$

da $x_j \in \mathrm{Eig}(f, \lambda_j)$ und $x_i \perp \mathrm{Eig}(f, \lambda_j)$ für $i \neq j$, und somit

$$\pi_j(x) = \pi_j(x_1) + \ldots + \pi_j(x_r) = x_j.$$

Damit erhalten wir dann

$$\begin{aligned}(\lambda_1 \cdot \pi_1 + \ldots + \lambda_r \cdot \pi_r)(x) =& \lambda_1 \cdot \pi_1(x) + \ldots + \lambda_r \cdot \pi_r(x) \\ =& \lambda_1 \cdot x_1 + \ldots + \lambda_r \cdot x_r = f(x_1) + \ldots + f(x_r) = f(x).\end{aligned}$$

□

Bemerkung A15.16 (Spektralzerlegung und Hauptraumzerlegung)
Die Projektionen π_i stimmen übrigens mit den Projektionen $Q_i(f)$ aus Bemerkung A12.20 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 A15.17 (Orthogonale / unitäre Abbildungen)

a. Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{R}}(V)$ heißt *orthogonal*, falls $f^* \circ f = \mathrm{id}_V$ gilt. $\mathrm{O}(V) := \{f \in \mathrm{End}_{\mathbb{R}}(V) \mid f \text{ ist orthogonal}\}$ heißt die *orthogonale Gruppe* von V.

b. Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{C}}(V)$ heißt *unitär*, falls $f^* \circ f = \mathrm{id}_V$ gilt. Wir nennen $\mathrm{U}(V) := \{f \in \mathrm{End}_{\mathbb{C}}(V) \mid f \text{ ist unitär}\}$ die *unitäre Gruppe* von V.

Proposition A15.18 (Matrixdarstellung orthogonaler / unitärer Abbildungen)

Es sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ 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:

$$\begin{aligned} f \text{ ist orthogonal bzw. unitär} &\iff f^* \circ f = \mathrm{id}_V \\ &\iff M_B^B(f^*) \circ M_B^B(f) = M_B^B(f^* \circ f) = \mathbb{1}_n \\ &\overset{\text{A15.3}}{\iff} M_B^B(f)^* \circ M_B^B(f) = \mathbb{1}_n \\ &\iff M_B^B(f) \text{ ist orthogonal bzw. unitär.} \end{aligned}$$

□

Korollar A15.19 (Orthogonal / unitär $\Longrightarrow$ normal)

Orthogonale und unitäre Abbildungen sind normal.

Beweis: Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ orthogonal oder unitär und B eine ONB von V, so ist $M_B^B(f)$ nach Proposition A15.18 orthogonal oder unitär. Nach Bemerkung A15.6 ist dann $M_B^B(f)$ auch normal, und mit Lemma A15.7 ist f deshalb normal. □

Proposition A15.20 (Charakterisierung orthogonaler / unitärer Endomorphismen)

Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ 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

$$\langle x, y\rangle = \langle x, \mathrm{id}_V(y)\rangle = \langle x, f^* \circ f(y)\rangle = \langle f(x), f(y)\rangle.$$

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

$$\langle (f^* \circ f)(x) - \mathrm{id}_V(x), y\rangle = \langle (f^* \circ f)(x), y\rangle - \langle x, y\rangle = 0$$

für alle $x, y \in V$. Für $x \in V$ beliebig setzen wir $y = (f^* \circ f)(x) - \mathrm{id}_V(x)$ und erhalten

$$\langle (f^* \circ f)(x) - \mathrm{id}_V(x), (f^* \circ f)(x) - \mathrm{id}_V(x)\rangle = 0,$$

was $(f^* \circ f)(x) - \mathrm{id}_V(x) = 0$ und $(f^* \circ f)(x) = \mathrm{id}_V(x)$ zur Folge hat. Mithin ist $f^* \circ f = \mathrm{id}_V$ und f ist orthogonal bzw. unitär. □

Proposition A15.21 (Eigenschaften orthogonaler / unitärer Abbildungen)

Es seien $f, g \in \mathrm{End}_{\mathbb{K}}(V)$ orthogonal bzw. unitär und $x, y \in V$

a. $||f(x)|| = ||x||$.

b. $x \perp y$ genau dann, wenn $f(x) \perp f(y)$.

c. Jeder Eigenwert von f hat Betrag 1.

d. f ist bijektiv.

e. f^{-1} und $f \circ g$ sind orthogonal bzw. unitär, d.h. $\mathrm{O}(V)$ und $\mathrm{U}(V)$ sind Gruppen.

Insbesondere, orthogonale und unitäre Abbildungen erhalten Längen und Abstände.

Beweis:

a. $||f(x)||^2 = \langle f(x), f(x)\rangle = \langle x, x\rangle = ||x||^2$.

b. $x \perp y \iff \langle f(x), f(y)\rangle = \langle x, y\rangle = 0 \iff f(x) \perp f(y)$.

c. Ist $0 \neq x \in V$ ein Eigenvektor zum Eigenwert $\lambda \in \mathbb{K}$, so gilt nach a.

$$||x|| = ||f(x)|| = ||\lambda x|| = |\lambda| \cdot ||x||,$$

also $|\lambda| = 1$, da $x \neq 0$.

d. Ist $x \in \mathrm{Ker}(f)$, so gilt nach a. $0 = ||f(x)|| = ||x||$, und somit $x = 0$. Also ist f injektiv, und da V endlich-dimensional ist, ist f somit auch bijektiv.

e. f^{-1} und $f \circ g$ sind orthogonal bzw. unitär, da für $x, y \in V$ gelten

$$\left\langle f^{-1}(x), f^{-1}(y)\right\rangle = \left\langle f\left(f^{-1}(x)\right), f\left(f^{-1}(y)\right)\right\rangle = \langle x, y\rangle$$

und

$$\left\langle (f \circ g)(x), (f \circ g)(y)\right\rangle = \left\langle f\left(g(x)\right), f\left(g(y)\right)\right\rangle = \left\langle g(x), g(y)\right\rangle = \langle x, y\rangle.$$

□

Bemerkung A15.22 (Orthogonale Abbildungen sind winkeltreu.)
Orthogonale Abbildungen erhalten Winkel, d.h. für $f \in \mathrm{O}(V)$ und $0 \neq x, y \in V$ gilt

$$\measuredangle\big(f(x), f(y)\big) = \measuredangle(x, y),$$

da f das Skalarprodukt und die euklidische Norm erhält. Es gilt nämlich

$$\measuredangle\big(f(x), f(y)\big) = \arccos\left(\tfrac{\langle f(x), f(y)\rangle}{||f(x)||\cdot||f(y)||}\right) = \arccos\left(\tfrac{\langle x, y\rangle}{||x||\cdot||y||}\right) = \measuredangle(x, y).$$

Satz A15.23 (Spektralsatz für unitäre Abbildungen)
Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{C}}(V)$ 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 A15.19 normal und χ_f zerfällt nach dem Fundamentalsatz der Algebra über $\mathbb{C}$ in Linearfaktoren. Aufgrund des Spektralsatzes für normale Abbildungen A15.10 besitzt V dann eine ONB aus Eigenvektoren von f. Zudem haben die Eigenwerte nach Proposition A15.21 alle Betrag 1.

Besitzt umgekehrt V eine ONB $B = (x_1, \ldots, x_n)$ aus Eigenvektoren von f zu Eigenwerten $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ vom Betrag 1, so gilt

$$M_B^B(f)^* \circ M_B^B(f) = \begin{pmatrix} \overline{\lambda_1} & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \overline{\lambda_n} \end{pmatrix} \circ \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

$$= \begin{pmatrix} \overline{\lambda_1}\lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \overline{\lambda_n}\lambda_n \end{pmatrix} = \begin{pmatrix} |\lambda_1|^2 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & |\lambda_n|^2 \end{pmatrix} = \mathbb{1}_n.$$

Mithin ist $M_B^B(f)$ unitär, und damit ist auch f unitär nach Proposition A15.18. □

Korollar A15.24 (Spektralsatz für unitäre Matrizen)
Ist $A \in \mathrm{U}(n)$, dann gibt es ein $T \in \mathrm{U}(n)$ mit

$$T^{-1} \circ A \circ T = T^* \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

wobei $\lambda_i \in \mathbb{C}$ mit $|\lambda_i| = 1$, $i = 1, \ldots, n$, 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 $T \in \mathrm{U}(n)$ und T transformiert A in eine Diagonalmatrix. □

Beispiel A15.25 (Unitäre Matrix)

Betrachten wir $\mathbb{C}^3$ mit dem kanonischen Skalarprodukt sowie die Matrix

$$A = \frac{1}{9} \cdot \begin{pmatrix} 1 & 8 & -4 \\ -4 & 4 & 7 \\ 8 & 1 & 4 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{C}).$$

Man rechnet sofort nach, daß $A^* \circ A = \mathbb{1}_3$, 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 $T \in \mathrm{U}(3)$ 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 $\mathrm{Eig}(A, 1)$ und finden $\left(\frac{1}{3}, \frac{2}{3}, \frac{2}{3}\right)^t$ als ONB. Durch Einsetzen in das Gleichungssystem überzeugt man sich, daß $(4, -1+3i, -1-3i)^t$ eine Lösung von $(A - i\mathbb{1}_3)x = 0$ ist, und durch Normierung erhalten wir dann $\left(\frac{2}{3}, \frac{-1+3i}{6}, \frac{-1-3i}{6}\right)^t$ als ONB von $\mathrm{Eig}(A, i)$. Da A eine reelle Matrix ist, muß somit $-i$ gerade den konjugiert komplexen Vektor als Eigenvektor haben, d. h. $\left(\frac{2}{3}, \frac{-1-3i}{6}, \frac{-1+3i}{6}\right)^t$ ist eine ONB von $\mathrm{Eig}(A, -i)$.

Wir erhalten also

$$T = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & \frac{-1+3i}{6} & \frac{-1-3i}{6} \\ \frac{2}{3} & \frac{-1-3i}{6} & \frac{-1+3i}{6} \end{pmatrix} \in \mathrm{U}(3)$$

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 A15.26 (Spektralsatz für orthogonale Abbildungen)

Mit dem gleichen Beweis wie in Satz A15.23 kann man zeigen, daß eine orthogonale Abbildung $f \in \mathrm{O}(V)$, 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 A11.14 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 $f \in \mathrm{End}_{\mathbb{R}}(V)$ orthogonal, dann gibt es eindeutig bestimmte Zahlen $r, s, t \in \mathbb{N}$ sowie Winkel $\alpha_1, \ldots, \alpha_t \in (0, 2\pi) \setminus \{\pi\}$ und eine ONB B von V, so daß

$$M_B^B(f) = \mathbb{1}_r \oplus -\mathbb{1}_s \oplus T(\alpha_1) \oplus \ldots \oplus T(\alpha_t).$$

D) Spektralsatz für selbstadjungierte Endomorphismen

Definition A15.27 (Selbstadjungierter Endomorphismus)
$f \in \mathrm{End}_{\mathbb{K}}(V)$ heißt *selbstadjungiert* oder *hermitesch*, wenn $f = f^*$ gilt.

Proposition A15.28 (Matrixdarstellung selbstadjungierter Endomorph.)
Es sei $f \in \mathrm{End}_{\mathbb{K}}(V)$ und B eine ONB von V.
Genau dann ist f selbstadjungiert, wenn $M_B^B(f)$ symmetrisch bzw. hermitesch ist.

Beweis: Aus Korollar A15.3 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 A15.29 (Selbstadjungiert impliziert normal.)
Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ 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 A15.30 (Selbstadjungierte EM haben nur reelle Eigenwerte.)
Ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ selbstadjungiert, dann ist $\chi_f \in \mathbb{R}[t]$ und χ_f zerfällt über $\mathbb{R}$.
Insbesondere 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 $\mathrm{Mat}_n(\mathbb{C})$ auffassen können und daß $A = A^*$ gilt. Über $\mathbb{C}$ zerfällt das charakteristische Polynom von A, d. h. es gibt $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ mit

$$\chi_A = (t - \lambda_1) \cdots (t - \lambda_n).$$

Es reicht, zu zeigen, daß $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ 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

$$\begin{aligned} \lambda_i \cdot \left(\overline{x_i}^t \circ x_i\right) &= \overline{x_i}^t \circ (\lambda_i x_i) = \overline{x_i}^t \circ (Ax_i) = \overline{x_i}^t \circ A \circ x_i \\ &= \overline{x_i}^t \circ \overline{A}^t \circ x_i = \overline{Ax_i}^t \circ x_i = \overline{\lambda_i x_i}^t \circ x_i = \overline{\lambda_i} \cdot \left(\overline{x_i}^t \circ x_i\right). \end{aligned}$$

Aus $\overline{x_i}^t \circ x_i \neq 0$ folgt dann $\lambda_i = \overline{\lambda_i}$, d. h. $\lambda_i \in \mathbb{R}$. □

Satz A15.31 (Spektralsatz für selbstadjungierte Endomorphismen)
Ein Endomorphismus $f \in \mathrm{End}_{\mathbb{K}}(V)$ 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 A15.29 normal und nach Lemma A15.30 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 $B = (x_1, \ldots, x_n)$ aus Eigenvektoren von V mit reellen Eigenwerten $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$, so gilt

$$M_B^B(f) = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix} = \begin{pmatrix} \overline{\lambda_1} & 0 & \ldots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & \overline{\lambda_n} \end{pmatrix} = M_B^B(f)^*.$$

Mithin ist $M_B^B(f)$ symmetrisch bzw. hermitesch, und somit ist f selbstadjungiert. □

Korollar A15.32 (Spektralsatz für symmetrische/hermitesche Matrizen)
Zu jeder symmetrischen bzw. hermiteschen Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ gibt es eine Matrix $T \in \mathrm{O}(n)$ bzw. $T \in \mathrm{U}(n)$ mit

$$T^{-1} \circ A \circ T = T^* \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \ldots & 0 & \lambda_n \end{pmatrix}$$

und $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$. 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. □

Dies ist eine wichtige Ergänzung des Satzes über die Jordansche Normalform.

Beispiel A15.33

Wir betrachten $\mathbb{C}^n$ mit dem kanonischen Skalarprodukt sowie die Matrix

$$A = \begin{pmatrix} 0 & -1 & i \\ -1 & 0 & -i \\ -i & i & 0 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{C}).$$

Da $A = A^*$ gilt, ist A hermitesch, und wir wollen mit Hilfe des Algorithmus A15.12 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

$$\dim_{\mathbb{C}} \mathrm{Eig}(A, 2) = \mathrm{mult}(\chi_A, 2) = 1$$

und

$$\dim_{\mathbb{C}} \mathrm{Eig}(A, -1) = \mathrm{mult}(\chi_A, -1) = 2$$

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:

$$B' = \big((-i, i, -1)^t\big)$$

ist eine Basis von $\mathrm{Eig}(A, 2)$ und

$$B'' = \big((1, 1, 0)^t, (0, i, 1)^t\big)$$

ist eine Basis von $\mathrm{Eig}(A, -1)$.

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 $\mathrm{Eig}(A, 2)$ deshalb

$$z_1 = \left(\tfrac{-i}{\sqrt{3}}, \tfrac{i}{\sqrt{3}}, \tfrac{-1}{\sqrt{3}}\right)^t.$$

Für $B'' = (x_2, x_3)$ ist es etwas mehr Aufwand. Wir setzen zunächst

$$z_2 = \frac{1}{||x_2||} \cdot x_2 = \left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0\right)^t.$$

Als nächstes setzen wir

$$y_3 = x_3 - \langle z_2, x_3\rangle \cdot z_2 = (0,i,1)^t - \tfrac{i}{\sqrt{2}} \cdot \left(\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0\right)^t = \left(\tfrac{-i}{2}, \tfrac{i}{2}, 1\right)^t$$

und normieren diesen Vektor anschließend zu

$$z_3 = \frac{1}{||y_3||} \cdot y_3 = \left(\tfrac{-i}{\sqrt{6}}, \tfrac{i}{\sqrt{6}}, \tfrac{2}{\sqrt{6}}\right)^t.$$

Die Vektoren z_2 und z_3 bilden eine ONB von $\mathrm{Eig}(A,-1)$, 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

$$T = \begin{pmatrix} \frac{-i}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-i}{\sqrt{6}} \\ \frac{i}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{i}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & 0 & \frac{2}{\sqrt{6}} \end{pmatrix} \in \mathrm{U}(3).$$

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 A15.34 (Definite Matrizen)

a. Eine symmetrische Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ heißt *positiv definit* bzw. *negativ definit* bzw. *indefinit*, wenn die Bilinearform b_A positiv bzw. negativ definit bzw. indefinit ist.

b. Eine hermitesche Matrix $A \in \mathrm{Mat}_n(\mathbb{C})$ heißt *positiv* bzw. *negativ definit* bzw. *indefinit*, wenn die Sesquilinearform b_A^s positiv bzw. negativ definit bzw. indefinit ist.

c. Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ 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 $\det\big(A(k)\big)$ den k-ten *Hauptminor* von A.

Bemerkung A15.35 (Negativ definite Matrizen)

Man beachte, daß eine symmetrische oder hermitesche Matrix A genau dann positiv definit ist, wenn für alle $0 \neq x \in \mathbb{K}^n$

$$\overline{x}^t \circ A \circ x > 0$$

gilt, und daß A genau dann negativ definit ist, wenn

$$\overline{x}^t \circ A \circ x < 0$$

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 A15.36 (Positiv definite Matrizen)
Sei $A \in \mathrm{Mat}_n(\mathbb{K})$ symmetrisch bzw. hermitesch und $T \in \mathrm{Gl}_n(\mathbb{K})$ invertierbar. Genau dann ist A positiv definit, wenn $T^* \circ A \circ T$ positiv definit ist.

Beweis: Wir beachten, daß

$$\mathbb{K}^n \setminus \{0\} = \{T \circ x \mid 0 \neq x \in \mathbb{K}^n\} \tag{245}$$

gilt, da T invertierbar ist, und daß die Matrix $T^* \circ A \circ T$ symmetrisch bzw. hermitesch ist, wegen

$$(T^* \circ A \circ T)^* = T^* \circ A^* \circ T^{**} = T^* \circ A \circ T.$$

Wir erhalten deshalb

$$\begin{aligned} A \text{ ist positiv definit} &\iff \overline{x}^t \circ A \circ x > 0 \quad \forall\, 0 \neq x \in \mathbb{K}^n \\ &\overset{(245)}{\iff} \overline{T \circ x}^t \circ A \circ (T \circ x) > 0 \quad \forall\, 0 \neq x \in \mathbb{K}^n \\ &\iff \overline{x}^t \circ T^* \circ A \circ T \circ x > 0 \quad \forall\, 0 \neq x \in \mathbb{K}^n \\ &\iff T^* \circ A \circ T \text{ ist positiv definit} \end{aligned}$$

□

Satz A15.37 (Hurwitz-Kriterium für positive Definitheit)
Für eine symmetrische bzw. hermitesche Matrix $A \in \mathrm{Mat}_n(\mathbb{K})$ sind die folgenden Aussagen gleichwertig:

a. A ist positiv definit.

b. Alle Eigenwerte von A sind positiv.

c. Alle Hauptminoren von A sind positiv.

Beweis:
a. $\iff$ b.: Wegen des Spektralsatzes für symmetrische und hermitesche Matrizen

A15.32 gibt es eine orthogonale bzw. unitäre Matrix $T \in \mathrm{Mat}_n(\mathbb{K})$, so daß

$$T^* \circ A \circ T = T^{-1} \circ A \circ T = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_n \end{pmatrix} \tag{246}$$

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 (246)

$$\lambda_i = \overline{x_i}^t \circ A \circ x_i > 0.$$

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

$$T \circ M_B(x) = \sum_{i=1}^{n} \mu_i \cdot x_i = x$$

und

$$\begin{aligned} \overline{x}^t \circ A \circ x &= \overline{T \circ M_B(x)}^t \circ A \circ (T \circ M_B(x)) \\ &= \overline{M_B(x)}^t \circ (T^* \circ A \circ T) \circ M_B(x) \\ &\overset{(246)}{=} \sum_{i=1}^{n} \overline{\mu_i} \cdot \mu_i \cdot \lambda_i = \sum_{i=1}^{n} |\mu_i|^2 \cdot \lambda_i > 0, \end{aligned}$$

da nicht alle μ_i null sind. Damit ist die Äquivalenz von a. und b. gezeigt.

a. $\Longrightarrow$ 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 A15.32 gibt es eine Matrix $T \in \mathrm{Gl}_n(\mathbb{K})$, so daß (246) erfüllt ist, und deshalb gilt

$$\det(A) = \det(T^{-1} \circ A \circ T) = \lambda_1 \cdot \ldots \cdot \lambda_n > 0.$$

Jede positiv definite symmetrische oder hermitesche Matrix hat also eine positive Determinante.

Für Vektoren $0 \neq y \in \mathbb{K}^k$ ist $x = (y_1, \dots, y_k, 0, \dots, 0) \in \mathbb{K}^n \setminus \{0\}$ und es gilt

$$\overline{y}^t \circ A(k) \circ y = \overline{x}^t \circ A \circ x > 0,$$

so daß auch $A(k)$ positiv definit ist. Dann ist aber ihre Determinante, der k-te Hauptminor von A, positiv.

c. $\Longrightarrow$ 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 A15.32 existiert für die symmetrische bzw. hermitesche Matrix $A(n-1)$ eine Matrix S in $O(n-1)$ bzw. $U(n-1)$, 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 $\lambda_1, \ldots, \lambda_{n-1}$ von $A(n-1)$ positiv.

Wir setzen nun $T = S \oplus \mathbb{1}_1 \in \mathrm{Gl}_n(\mathbb{K})$. Dann gilt

$$T^* \circ A \circ T = \left(\begin{array}{ccc|c} & & & a_1 \\ & D & & \vdots \\ & & & a_{n-1} \\ \hline \overline{a_1} & \ldots & \overline{a_{n-1}} & a_n \end{array}\right) =: B$$

für geeignete $a_1, \ldots, a_n \in \mathbb{K}$. Setzen wir ferner $c_j = -\frac{a_j}{\lambda_j}$, $j = 1, \ldots, n-1$, und

$$C = \left(\begin{array}{ccc|c} & & & c_1 \\ & \mathbb{1}_{n-1} & & \vdots \\ & & & c_{n-1} \\ \hline 0 & \ldots & 0 & 1 \end{array}\right) \in \mathrm{Gl}_n(\mathbb{K}),$$

dann folgt mit $a = (a_1, \ldots, a_{n-1})^t$ und $c = (c_1, \ldots, c_{n-1})^t$

$$\begin{aligned} C^* \circ B \circ C &= \left(\begin{array}{c|c} \mathbb{1}_{n-1} & 0 \\ \hline c^* & 1 \end{array}\right) \circ \left(\begin{array}{c|c} D & a \\ \hline a^* & a_n \end{array}\right) \circ \left(\begin{array}{c|c} \mathbb{1}_{n-1} & c \\ \hline 0 & 1 \end{array}\right) \\ &= \left(\begin{array}{c|c} \mathbb{1}_{n-1} \circ D + 0 \circ a^* & \mathbb{1}_{n-1} \circ a + 0 \cdot a_n \\ \hline c^* \circ D + 1 \cdot a^* & c^* \circ a + 1 \cdot a_n \end{array}\right) \circ \left(\begin{array}{c|c} \mathbb{1}_{n-1} & c \\ \hline 0 & 1 \end{array}\right) \\ &= \left(\begin{array}{c|c} D & a \\ \hline 0 & c^* \circ a + a_n \end{array}\right) \circ \left(\begin{array}{c|c} \mathbb{1}_{n-1} & c \\ \hline 0 & 1 \end{array}\right) \qquad (247) \\ &= \left(\begin{array}{c|c} D \circ \mathbb{1}_{n-1} + a \circ 0 & D \circ c + a \cdot 1 \\ \hline 0 \circ \mathbb{1}_{n-1} + (c^* \circ a + a_n) \cdot 0 & 0 \circ c + (c^* \circ a + a_n) \cdot 1 \end{array}\right) \\ &= \left(\begin{array}{c|c} D & 0 \\ \hline 0 & c^* \circ a + a_n \end{array}\right) \end{aligned}$$

und damit

$$E := (T \circ C)^* \circ A \circ (T \circ C) = C^* \circ T^* \circ A \circ T \circ C = C^* \circ B \circ C \overset{(247)}{=} \bigoplus_{i=1}^{n} \lambda_i \mathbb{1}_1,$$

für $\lambda_n = c^* \circ a + a_n \in \mathbb{R}$ geeignet ist. Man beachte, daß damit $\lambda_1, \ldots, \lambda_n$ die Eigenwerte von E sind, und daß

$$\lambda_1 \cdot \ldots \cdot \lambda_n = \det\big((T \circ C)^* \circ A \circ (T \circ C)\big) = \det(A) \cdot |\det(C \circ T)|^2 > 0,$$

da $\det(A) > 0$ der n-te Hauptminor von A ist. Da aber $\lambda_1, \ldots, \lambda_{n-1}$ nach Voraussetzung positiv waren, ist dann auch λ_n positiv. E hat somit nur positive Eigenwerte und ist deshalb positiv definit. Aber mit Lemma A15.36 ist dann auch A positiv definit. □

Bemerkung A15.38 (Negativ definite und indefinite Matrizen)
Wie im Beweis von "a. $\Longleftrightarrow$ b." im Beweis von Satz A15.37 sieht man:

$$A \text{ ist } \textit{negativ definit} \iff A \text{ hat nur negative Eigenwerte}$$

und

$$A \text{ ist } \textit{indefinit} \iff A \text{ hat einen positiven und einen negativen Eigenwert.}$$

F) Hauptachsentransformationssatz für quadratische Formen

Definition A15.39 (Bilinearformen)
Es sei V ein $\mathbb{R}$-Vektorraum. Eine Abbildung

$$b : V \times V \to \mathbb{R},$$

die linear in beiden Argumenten ist, nennen wir *bilinear* oder eine *Bilinearform*, d. h. für $x, y, z \in V$ und $\lambda, \mu \in \mathbb{R}$ gilt (vgl. Definition A9.8):

$$b(\lambda x + \mu y, z) = \lambda b(x, z) + \mu b(y, z)$$

und

$$b(z, \lambda x + \mu y) = \lambda b(z, x) + \mu b(z, y).$$

Gilt zudem

$$b(x, y) = b(y, x)$$

für alle $x, y \in V$, so heißt b eine *symmetrische* Bilinearform, und wir nennen die Abbildung

$$q_b : V \longrightarrow \mathbb{R} : x \mapsto b(x, x)$$

die zu b gehörende *quadratische Form*.

Bemerkung A15.40 (Polarisierung einer Bilinearform)
Ist $b : V \times V \longrightarrow \mathbb{R}$ eine symmetrische Bilinearform, dann gilt

$$b(x, y) = \tfrac{1}{2}\big(q_b(x + y) - q_b(x) - q_b(y)\big)$$

für $x, y \in V$, wie man durch Einsetzen der Definition leicht nachrechnet. Die Bilinearform ist durch die zugehörige quadratische Form also eindeutig bestimmt. Man nennt die Gleichung die *Polarisierung* der Bilinearform.

Beispiel A15.41
Jedes Skalarprodukt auf einem $\mathbb{R}$-Vektorraum ist eine symmetrische Bilinearform.

Zudem definiert jede symmetrische Matrix $A \in \mathrm{Mat}_n(\mathbb{R})$ durch

$$b_A : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R} : (x, y) \mapsto x^t \circ A \circ y$$

eine symmetrische Bilinearform. Die zu b_A gehörende quadratische Form bezeichnen wir mit

$$q_A : \mathbb{R}^n \longrightarrow \mathbb{R} : x \mapsto x^t \circ A \circ x.$$

Satz A15.42 (Hauptachsentransformationssatz für quadrat. Formen)
Ist $A \in \mathrm{Mat}_n(\mathbb{R})$ eine symmetrische Matrix, so gibt es eine ONB $B = (x_1, \ldots, x_n)$ von $\mathbb{R}^n$ zum kanonischen Skalarprodukt mit

$$A \circ x_i = \lambda_i \cdot x_i$$

für $i = 1, \ldots, n$ und

$$q_A(x) = \sum_{i=1}^{n} \lambda_i \cdot \langle x_i, x \rangle^2.$$

Die x_i werden die *Hauptachsen* der quadratischen Form genannt.

Beweis: Aus dem Spektralsatz für selbstadjungierte Endomorphismen A15.31 wissen wir, daß es eine ONB $B = (x_1, \ldots, x_n)$ von $\mathbb{R}^n$ aus Eigenvektoren von f_A gibt. Diese wählen wir, so daß

$$A \circ x_i = f_A(x_i) = \lambda_i \cdot x_i$$

für geeignete $\lambda_i \in \mathbb{R}$ gilt. Sei $x \in V$, so folgt aus der Parsevalschen Gleichung A14.16

$$x = \sum_{i=1}^{n} \langle x_i, x \rangle \cdot x_i.$$

Für die quadratische Form q_A ausgewertet in x ergibt sich dann

$$\begin{aligned} q_A(x) =& x^t \circ A \circ x = x^t \circ A \circ \sum_{i=1}^{n} \langle x_i, x \rangle \cdot x_i = \sum_{i=1}^{n} \langle x_i, x \rangle \cdot x^t \circ A \circ x_i \\ =& \sum_{i=1}^{n} \langle x_i, x \rangle \cdot x^t \circ \lambda_i \cdot x_i = \sum_{i=1}^{n} \langle x_i, x \rangle \cdot \lambda_i \cdot \langle x, x_i \rangle = \sum_{i=1}^{n} \lambda_i \cdot \langle x, x_i \rangle^2. \end{aligned}$$

□

Bemerkung A15.43 (Geometrische Interpretation der Hauptachsentransf.)
Ist $A \in \mathrm{Mat}_n(\mathbb{R})$ eine positiv definite symmetrische Matrix mit quadratischer Form q_A, so interessieren wir uns für die Einheitssphäre zu q_A

$$S_{q_A} = \{y \in \mathbb{R}^n \mid q_A(y) = y^t \circ A \circ y = 1\}.$$

Man beachte, daß die Bilinearform b_A in diesem Fall ein Skalarprodukt ist und daß S_{q_A} die Menge der Vektoren in $\mathbb{R}^n$ ist, die bezüglich der zu diesem Skalarprodukt gehörenden Norm die Länge 1 haben.

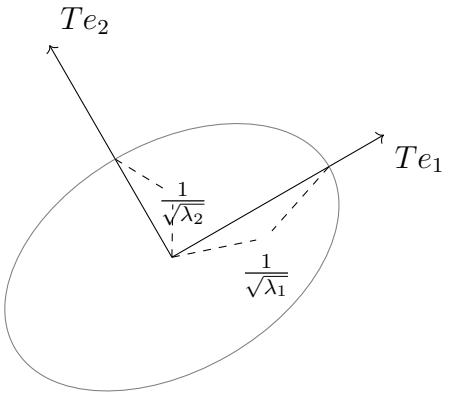


Abbildung 21. Ellipse S_{q_A} mit Hauptachsen Te_1 und Te_2

Der Spektralsatz für selbstadjungiert Endomorphismen liefert die Existenz einer orthogonalen Matrix $T \in \mathrm{O}(n)$, so daß

$$T^t \circ A \circ T = \lambda_1 \cdot \mathbb{1}_1 \oplus \ldots \oplus \lambda_n \cdot \mathbb{1}_1$$

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

$$q = \lambda_1 \cdot y_1^2 + \ldots + \lambda_n \cdot y_n^2$$

hat. Die Einheitssphäre zu q ist dann ein n-dimensionales Ellipsoid

$$S_q = \{y \in \mathbb{R}^n \mid q(y) = \lambda_1 \cdot y_1^2 + \ldots + \lambda_n \cdot y_n^2 = 1\}.$$

Man kann sich S_{q_A} deshalb als eine Einheitskugel vorstellen, die in Richtung $x_i = Te_i$ um den Faktor $\frac{1}{\sqrt{\lambda_i}}$ gestreckt wurde. Die Koordinatenvektoren x_i dind dann die *Hauptachsen* des Ellipsoids (siehe Abbildung 21). 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*.

Aufgaben

Aufgabe A15.44 (Lineare Funktionale)
Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum.
Dann gibt es für jedes $g \in \mathrm{Hom}_{\mathbb{K}}(V, \mathbb{K})$ genau ein $y \in V$, so daß für alle $x \in V$ gilt

$$g(x) = \langle y, x \rangle.$$

Aufgabe A15.45
Zeige, wenn für $f \in \mathrm{End}_{\mathbb{K}}(V)$ stets $||f(x)|| = ||x||$ gilt, so ist f orthogonal bzw. unitär.

Aufgabe A15.46 (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 $f : V \longrightarrow W$ genau eine lineare Abbildung $f^* : W \longrightarrow V$, so daß

$$\langle f(x), y \rangle_W = \langle x, f^*(y) \rangle_V \tag{248}$$

für alle $x \in V$ und $y \in W$. Die Abbildung f^* heißt die *adjungierte Abbildung* zu f.

Aufgabe A15.47 (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 $f : V \longrightarrow W$

$$M_B^D(f^*) = \big(M_D^B(f)\big)^*,$$

d.h. die Matrixdarstellung der adjungierten Abbildung ist die Adjungierte der Matrixdarstellung.

Aufgabe A15.48
Sei V ein endlich-dimensionaler euklidischer oder unitärer Raum.
Zeige, ist $f \in \mathrm{End}_{\mathbb{K}}(V)$ normal, so gelten

$$\mathrm{Ker}(f) = \mathrm{Ker}(f^*)$$

und

$$V = \mathrm{Ker}(f) \perp \mathrm{Im}(f).$$

Aufgabe A15.49
Es sei V ein endlich-dimensionaler euklidischer oder unitärer Raum und $f \in \mathrm{End}_{\mathbb{K}}(V)$ bijektiv. Zeige, die folgenden Aussagen sind äquivalent:

a. Für $x, y \in V$ mit $x \perp y$ gilt $f(x) \perp f(y)$.

b. Für $x, y \in V$ mit $||x|| = ||y||$ gilt $||f(x)|| = ||f(y)||$.

c. Es gibt ein $\lambda \in \mathbb{R}_{>0}$ und ein $g \in \mathrm{O}(V)$ bzw. $g \in \mathrm{U}(V)$ mit $f = \lambda g$.

Aufgabe A15.50

Es sei $V \neq 0$ ein endlich-dimensionaler unitärer Raum und $f \in \mathrm{End}_{\mathbb{C}}(V)$. Zeige, die folgenden Aussagen sind gleichwertig:

a. $f^* = -f$.

b. Für alle $x \in V$ gilt: $\langle f(x), x\rangle \in i\mathbb{R}$.

c. Es gibt eine Orthonormalbasis von V aus Eigenvektoren von f und der Realteil aller Eigenwerte ist Null.

Aufgabe A15.51

Überprüfe, ob die folgende symmetrische Matrix $A \in \mathrm{Mat}_3(\mathbb{R})$ positiv definit ist:

$$A = \begin{pmatrix} 9 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

Aufgabe A15.52

Bestimme eine orthogonale Matrix $T \in \mathrm{O}(3)$, 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?

Aufgabe A15.53

Es sei $\sigma \in \mathbb{S}_n$ ein n-Zykel, $E = (e_1, \ldots, e_n)$ die kanonische Basis des $\mathbb{C}^n$ und $f \in \mathrm{End}_{\mathbb{C}}(\mathbb{C}^n)$ der eindeutig bestimmte Endomorphismus mit $f(e_i) = e_{\sigma(i)}$ für $i = 1, \ldots, n$. Wir betrachten $\mathbb{C}^n$ als unitären Raum mit dem kanonischen Skalarprodukt. Zeige die folgenden Aussagen:

a. f ist unitär.

b. $\chi_f = t^n - 1$ und die Eigenwerte von f sind die n-ten Einheitswurzeln.

c. Bestimme eine ONB B, so daß $M_B^B(f)$ eine Diagonalmatrix ist.

d. Wie verallgemeinert sich die Aussage, wenn σ kein n-Zykel ist?

Hinweis, in Teil c. betrachte man den Vektor $b_k = \sum_{j=1}^{n} \xi^{(1-k)\cdot(j-1)} \cdot e_{\sigma^{j-1}(1)}$ mit $\xi = e^{\frac{2\pi i}{n}}$.

Aufgabe A15.54

Sei V ein endlich-dimensionaler unitärer Raum und $f \in \mathrm{End}_{\mathbb{C}}(V)$ normal. Zeige, es gibt ein Polynom $p \in \mathbb{C}[t]$ mit $f^* = p(f)$.

A16 Klassifikation der Kegelschnitte in der euklidischen Ebene

Man kann die Überlegungen in A15.43 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 A16.1

Es sei $\big(V, \langle\cdot,\cdot\rangle\big)$ ein euklidischer Raum.

a. Eine Abbildung $f : V \to V$ heißt eine *affine Abbildung* auf V, falls es ein $y \in V$ gibt und ein $g \in \mathrm{End}_{\mathbb{R}}(V)$ mit $f(x) = y + g(x)$ für alle $x \in V$.

b. Für $y \in V$ nennen wir die affine Abbildung

$$\tau_y : V \to V : x \mapsto x + y$$

die *Translation* um den Vektor y.

c. Eine Abbildung $f : V \to V$ heißt eine *Ähnlichkeit*, wenn es einen Vektor $y \in V$ gibt und eine orthogonale Abbildung $g \in \mathrm{O}(V)$ mit $f = \tau_y \circ g$, d. h.

$$f(x) = \tau_y\big(g(x)\big) = y + g(x) \quad \forall\, x \in V.$$

d. Ist $V = \mathbb{R}^n$ und sei $f = \tau_y \circ g$ mit $g \in \mathrm{End}_{\mathbb{R}}(V)$ eine bijektive affine Abbildung auf V, dann nennen wir die induzierte Abbildung

$$\mathbb{R}[t_1, \ldots, t_n] \to \mathbb{R}[t_1, \ldots, t_n] : p \mapsto p\big(f(t_1, \ldots, t_n)\big)$$

auf der Menge der Polynome in den Veränderlichen $t_1, \ldots, t_n$ einen *affinen Koordinatenwechsel* von $\mathbb{R}[t_1, \ldots, t_n]$.

Bemerkung A16.2

a. Jede affine Abbildung $f : \mathbb{R}^n \to \mathbb{R}^n$ läßt sich offenbar in eindeutiger Weise schreiben, als $f = \tau_y \circ g$ mit $y = f(0) \in V$ und $g \in \mathrm{End}_{\mathbb{R}}(V)$.

b. Ist $f = \tau_y \circ g : \mathbb{R}^n \to \mathbb{R}^n$ eine affine Abbildung mit $y \in \mathbb{R}^n$ und $g \in \mathrm{End}_{\mathbb{R}}(\mathbb{R}^n)$ bijektiv, dann gibt es eine eindeutig bestimmte Matrix $T \in \mathrm{Gl}_n(\mathbb{R})$ mit $g = f_T$. Damit gilt für $p \in \mathbb{R}[t_1, \ldots, t_n]$ und $t = (t_1, \ldots, t_n)^t$

$$p\big(f(t_1, \ldots, t_n)\big) = p(Tt + y).$$

Ist beispielsweise $p = t_1^2 + 3t_2 - 1 \in \mathbb{R}[t_1, t_2]$, $T = T\big(\frac{\pi}{2}\big)$ die Drehung um 90° und $y = (2, -2)$, dann ist für $f = \tau_y \circ f_T$

$$p\big(f(t_1, t_2)\big) = p(-t_2 + 2, t_1 - 2) = (-t_2 + 2)^2 + 3(t_1 - 2) - 1.$$

Definition A16.3
Es sei $p \in \mathbb{R}[t_1, \ldots, t_n]$ dann nennen wir die Menge

$$N(p) = \{\lambda = (\lambda_1, \ldots, \lambda_n)^t \in \mathbb{R}^n \mid p(\lambda) = 0\}$$

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 A16.4
Wir definieren auf $\mathbb{R}[t_1, \ldots, t_n]$ eine Relation durch

$$p \equiv q \quad :\Leftrightarrow \quad \exists\, c \in \mathbb{R} \setminus \{0\} : p = c \cdot q$$

für $p, q \in \mathbb{R}[t_1, \ldots, t_n]$. Wir nennen p und q mit $p \equiv q$ auch *äquivalent*.

Bemerkung A16.5
Man sieht sofort, daß $\equiv$ eine Äquivalenzrelation auf $\mathbb{R}[t_1, \ldots, t_n]$ definiert.

Ferner gilt offensichtlich, daß für zwei äquivalente Polynome $p, q \in \mathbb{R}[t_1, \ldots, t_n]$ 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 A16.6
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

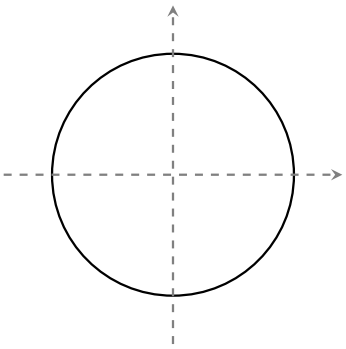
$$p = \alpha_{11}t_1^2 + 2\alpha_{12}t_1t_2 + \alpha_{22}t_2^2 + \alpha_1t_1 + \alpha_2t_2 + \alpha = \langle t, St\rangle + \langle a, t\rangle + \alpha,$$

wobei $t = (t_1, t_2)^t$, $0 \neq S = (\alpha_{ij})_{i,j\in\{1,2\}} \in \mathrm{Mat}_2(\mathbb{R})$ und $a = (\alpha_1, \alpha_2)^t$.

Beispiel A16.7
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,

$$N\big(t_1^2 + t_2^2 - 1\big) = \big\{\lambda = (\lambda_1, \lambda_2)^t \in \mathbb{R}^2 \mid \lambda_1^2 + \lambda_2^2 = 1\big\},$$

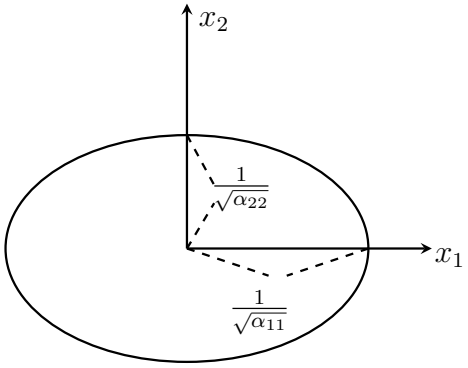
ist offenbar der Einheitskreis.



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

$$N\Big(\big(\sqrt{\alpha_{11}}t_1\big)^2 + \big(\sqrt{\alpha_{22}}t_2\big)^2 - 1\Big) = \Big\{(\lambda_1,\lambda_2)^t \in \mathbb{R}^2 \;\Big|\; \big(\sqrt{\alpha_{11}}\lambda_1\big)^2 + \big(\sqrt{\alpha_{22}}\lambda_2\big)^2 = 1\Big\}$$

eine Ellipse.



Satz A16.8

Es sei

$$p = \langle t, St\rangle + \langle a, t\rangle + \alpha \in \mathbb{R}[t_1, t_2] \tag{249}$$

ein Polynom zweiten Grades mit symmetrischer Matrix $0 \neq S = (\alpha_{ij}) \in \mathrm{Mat}_2(\mathbb{R})$. Dann gibt es eine affine Koordinatentransformation mittels einer Ähnlichkeit $f = \tau_y \circ f_T$ von $\mathbb{R}^2$ mit $T \in \mathrm{SO}(2)$, so daß $q := p\big(f(t_1,t_2)\big)$ 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: $\det(S) = 0$, $a \neq (0,0)^t$. Dann ist $q \equiv t_1^2 - \lambda t_2$ und $N(q)$ ist eine *Parabel*.

IV: $\det(S) = 0$, $a = (0,0)^t$.
- 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 *Doppelgraden*, d. h. einer Geraden, die man doppelt zählt.

Bemerkung A16.9

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

$$N\big(t_1^2 + t_2^2 - t_3^2\big) \subset \mathbb{R}^3$$

mit einer geeigneten Ebene im $\mathbb{R}^3$ realisierbar sind (siehe Abbildung 22).

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 A16.8:

1. Fall: $a = (0,0)^t$: Wir betrachten zunächst den Fall $a = (0,0)^t$.

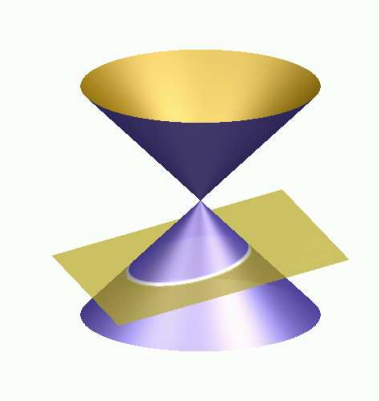
Nach dem Satz über die Hauptachsentransformation A15.32 existiert ein $T \in \mathrm{SO}(2)$, so daß

$$T^t \circ S \circ T = T^{-1} \circ S \circ T = \begin{pmatrix} \mu_1 & 0 \\ 0 & \mu_2 \end{pmatrix}.$$

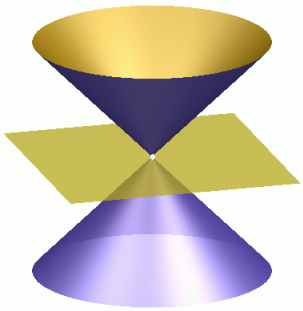
Man beachte noch, daß nicht beide Eigenwerte μ_1 und μ_2 null sein können, da $S \neq 0$. Also können wir o. E. annehmen, daß $\mu_1 \neq 0$ und daß $\mu_1 \geq \mu_2$ gilt, falls $\mu_2 \neq 0$.

Die lineare Abbildung $f_T : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto Tx$ ist eine Drehung und es gilt

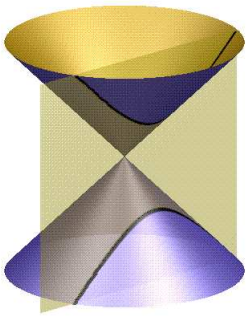
$$\begin{aligned} p(Tt) &= \big\langle Tt, (S \circ T)t \big\rangle + \alpha \\ &= \big\langle t, \big(T^t \circ S \circ T\big)t \big\rangle + \alpha \\ &= \mu_1 t_1^2 + \mu_2 t_2^2 + \alpha. \end{aligned}$$



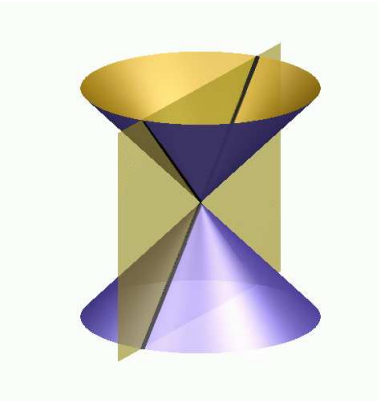
I.1: Ellipse



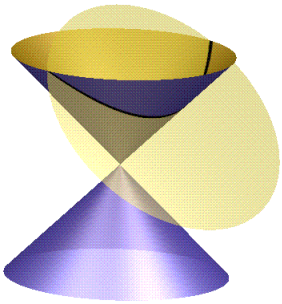
I.3: Punkt



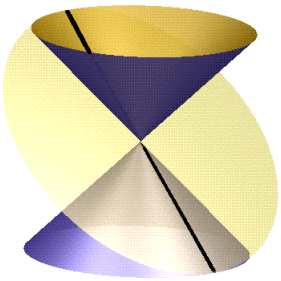
II.1: Hyperbel



II.2: Zwei Geraden mit Schnitt



III: Parabel



IV.3: Doppelgerade

Abbildung 22. Kegelschnitte

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: $a \neq (0,0)^t$: Sind wir im Fall $a = (0,0)^t$ noch ohne Translation ausgekommen, so werden wir jetzt doch Translationen betrachten müssen.

Für $c \in \mathbb{R}^2$ bewirkt die Translation $\tau_c : \mathbb{R}^2 \to \mathbb{R}^2 : x \mapsto x + c$ folgende Koordinatentransformation für p

$$\begin{aligned} p(t+c) &= \langle t+c, St+Sc\rangle + 2\langle a, t+c\rangle + \alpha \\ &= \langle t, St\rangle + 2\langle a+Sc, t\rangle + \langle c, Sc\rangle + 2\langle a, c\rangle + \alpha \\ &= \langle t, St\rangle + 2\langle b, t\rangle + \beta, \end{aligned} \tag{250}$$

wenn wir $b = a + Sc$ und $\beta = \langle c, Sc\rangle + 2\langle a, c\rangle + \alpha$ setzen.

Fall 2.1: $\exists\, c \in \mathbb{R}^2 : b = a + Sc = (0,0)^t$: Dann haben wir p durch $p\big(\tau_c(t)\big)$ auf den ersten Fall "$a = (0,0)^t$" zurückgeführt. Es gibt also ein $T \in \mathrm{SO}(2)$, so daß $q = p\big((\tau_c \circ f_T)(t)\big)$ äquivalent zu einem der Fälle I, II oder IV ist.

Fall 2.2: $\forall\, c \in \mathbb{R}^2 : b = a + Sc \neq (0,0)^t$: Aus Lemma A16.10 folgt, daß es ein $c \in \mathbb{R}^2$ gibt mit $Sb = S^2c + Sa = 0$. Setzen wir nun noch $\delta := -\frac{\beta}{2\langle b,b\rangle}$, dann gilt für die Translation $\tau_{c+\delta b}$[6]

$$\begin{aligned} p(t+c+\delta b) =& \langle t, St\rangle + 2\big\langle a + S(c+\delta b), t\big\rangle + \big\langle c+\delta b, S(c+\delta b)\big\rangle + 2\langle a, c+\delta b\rangle + \alpha \\ =& \langle t, St\rangle + 2\langle b + \delta Sb, t\rangle + \delta^2\langle b, Sb\rangle + 2\delta\langle b, b\rangle + \beta \\ =& \langle t, St\rangle + 2\langle b, t\rangle + 2\delta\langle b, b\rangle + \beta \\ =& \langle t, St\rangle + 2\langle b, t\rangle. \end{aligned}$$

Beachtet man, daß, wegen $Sb = 0$, Null auf alle Fälle ein Eigenwert von S ist und daß $S \neq 0$, so folgt aus dem Satz über Hauptachsentransformation A15.32 die Existenz eines $T \in \mathrm{SO}(2)$, 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 $\det(S) = 0$.

Ferner gilt für $T^t b =: (\mu, \lambda)^t$ unter Berücksichtigung, daß $T^t = T^{-1}$,

$$(\mu_1\mu, 0) = \big(T^t \circ S \circ T\big) \circ (T^t b) = T^t \circ (Sb) = 0,$$

und mithin ist $T^t b = (0, \lambda)^t$, wobei $\lambda \neq 0$, da T^t invertierbar und $b \neq (0,0)^t$. Aber dann überführt $t \mapsto Tt$ das Polynom $\langle t, St\rangle + 2\langle b, t\rangle$ in das Polynom

$$\big\langle Tt, (S \circ T)t\big\rangle + 2\big\langle b, Tt\big\rangle = \big\langle t^t, Dt\big\rangle + 2\big\langle T^t b, t\big\rangle = \mu_1 t_1^2 + 2\lambda t_2.$$

[6] Man setze zunächst in der Gleichung (250) für c den Wert $c + \delta b$ ein. Dann ziehe man die Skalarprodukte auseinander und gruppiere sie neu, so daß man $b = a + Sc$, $Sb = 0$ sowie die Definition von β verwenden kann. Man beachte auch, daß S symmetrisch, also selbstadjungiert, ist.

D. h. dann aber, daß

$$q := p\big((\tau_{c+\delta b} \circ f_T)(t)\big) = \mu_1 t_1^2 + 2\lambda t_2,$$

und damit sind wir genau im Fall III. □

Lemma A16.10

Ist $S \in \mathrm{Mat}_n(\mathbb{R})$ symmetrisch, so gilt für die lineare Abbildung $f_S : \mathbb{R}^n \to \mathbb{R}^n$.

a. $\mathrm{Ker}\left(f_S^2\right) = \mathrm{Ker}(f_S)$ und $\mathrm{Im}\left(f_S^2\right) = \mathrm{Im}(f_S)$.

b. Zu jedem $a \in \mathbb{R}^n$ existiert ein $c \in \mathbb{R}^n$, so daß $S^2c + Sa = 0$.

Beweis: a. Für $x \in \mathrm{Ker}\left(f_S^2\right)$ ergibt sich aus

$$0 = \langle x, S^2x\rangle = \langle Sx, Sx\rangle,$$

also $f_S(x) = Sx = 0$ und $x \in \mathrm{Ker}(f_S)$. Die umgekehrte Inklusion ist klar.
Wir wissen bereits, daß $\mathrm{Im}(f_S) \supseteq \mathrm{Im}\left(f_S^2\right)$ gilt. Da nun ferner

$$\begin{aligned}\dim_{\mathbb{R}}\left(\mathrm{Im}(f_S)\right) &= n - \dim_{\mathbb{R}}\left(\mathrm{Ker}(f_S)\right)\\ &= n - \dim_{\mathbb{R}}\left(\mathrm{Ker}\left(f_S^2\right)\right) = \dim_{\mathbb{R}}\left(\mathrm{Im}\left(f_S^2\right)\right)\end{aligned}$$

gilt, folgt also die Gleichheit.

b. Es gilt für $a \in \mathbb{R}^n$, daß $S(-a) = f_S(-a) \in \mathrm{Im}(f_S) = \mathrm{Im}\left(f_S^2\right)$, also gibt es nach a. ein $c \in \mathbb{R}^n$ mit $S^2c + Sa = f_S^2(c) - f_S(-a) = 0$.

□

A17 Singulärwertzerlegung

Die essentiellen Informationen einer diagonalisierbaren quadratischen Matrix sind in den Eigenwerten und Eigenvektoren der Matrix enthalten. Mit Hilfe der Eigenwertzerlegung können wir auf diese Information zugreifen. Allerdings ist nicht jede quadratische Matrix diagonalisierbar.

In diesem Abschnitt wollen wir nun eine Verallgemeinerung der Eigenwertzerlegung betrachten, die für jede (auch nicht-quadratische) Matrix über $\mathbb{R}$ existiert und die von großem praktischen Interesse ist, die *Singulärwertzerlegung*. Da eine Singulärwertzerlegung einer Matrix auch für nicht-quadratische Matrizen existiert, werden wir von links und rechts von einander verschiedene Transformationsmatrizen zulassen müssen. Damit dabei wesentliche Informationen über den Rang hinaus erhalten bleiben, werden wir für diese aber fordern, daß sie orthogonal sind.

Die Singulärwertzerlegung hat viele Anwendungen in den Bereichen der Linearen Algebra, des Maschinellen Lernens, der Statistik, der Bildverarbeitung, der Geographie (z.B. Wettervorhersage), der Chemie, ...

Satz A17.1 (Singulärwertzerlegung)
Für eine Matrix $A \in \mathrm{Mat}(m \times n, \mathbb{R})$ vom Rang r gibt es orthogonale Matrizen $U \in \mathrm{O}(m)$ und $V \in \mathrm{O}(n)$ sowie positive Zahlen $\sigma_1, \ldots, \sigma_r \in \mathbb{R}$ mit

$$\sigma_1 \geq \ldots \geq \sigma_r > 0,$$

so daß

$$A = U \circ \Sigma \circ V^t \tag{251}$$

gilt, wobei Σ eine Blockdiagnoalmatrix der Form

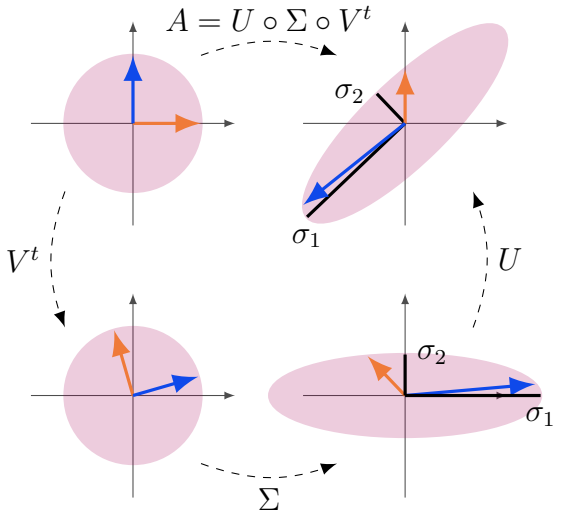
$$\Sigma = \left(\begin{array}{c|c} D & 0 \\ \hline 0 & 0 \end{array}\right) \in \mathrm{Mat}(m \times n, \mathbb{R})$$

ist mit einer Diagonalmatrix

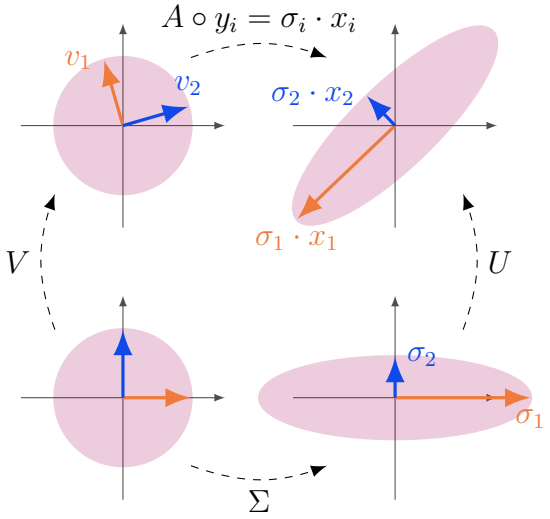
$$D = \begin{pmatrix} \sigma_1 & 0 & \ldots & \ldots & 0 \\ 0 & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \ldots & \ldots & 0 & \sigma_r \end{pmatrix} \in \mathrm{Mat}_r(\mathbb{R}).$$

Wir nennen eine Zerlegung wie in (251) eine *Singulärwertzerlegung* von A, die Spalten von U heißen *linke Singulärvektoren*, die Spalten von V *rechte Singulärvektoren* und die Diagonaleinträge $\sigma_1 \geq \ldots \geq \sigma_r > 0$ heißen *Singulärwerte*.

Bemerkung A17.2 (Geometrische Interpretation der Singulärwertzerlegung)
Geometrisch stellt eine Singulärwertzerlegung die lineare Abbildung A als Komposition einer Drehspiegelung, einer Skalierung der Koordinatenachsen und einer weiteren Drehspiegelung dar. Diese Interpretation ist in Abbildung 23(a) visualisiert.



(a) Eine Singulärwertzerlegung stellt die lineare Abbildung A als eine Komposition einer Drehspiegelung V^t, einer Skalierungsmatrix Σ, und einer weiteren Drehspiegelung U dar.



(b) Bezüglich der Basen aus Singulärvektoren ist die lineare Abbildung A besonders einfach. Es gilt $Ay_i = \sigma_i x_i$.

ABBILDUNG 23. Geometrie einer Singulärwertzerlegung.

Alternativ zur Zerlegung in ein Produkt von Matrizen lässt sich eine Singulärwertzerlegung von A auch wie folgt schreiben:

$$A = U \circ \Sigma \circ V^t = \sum_{i=1}^{r} \sigma_i \cdot x_i \circ y_i^t ,$$

wobei $r = \operatorname{rang}(A)$ ist, $x_1, \ldots, x_m$ die Spalten von U und $y_1, \ldots, y_n$ die Spalten von V sind. Eine weitere Darstellungsmöglichkeit ist

$$A \circ V = U \circ \Sigma$$

oder

$$A \circ y_i = \sigma_i \cdot x_i \quad \text{für} \quad i = 1, \ldots, r.$$

Letztere zeigt, daß die lineare Abbildung A bezüglich der Basen $(y_1, \ldots, y_n)$ von $\mathbb{R}^n$ und $(x_1, \ldots, x_m)$ von $\mathbb{R}^m$ eine besonders einfache Darstellung hat, nämlich eine Skalierung der Basisvektoren. Diese Charakterisierung ist in Abbildung 23(b) visualisiert. Dabei können die Ellipsen (bzw. Ellipsoiden) auch niedrigdimensionale Einbettungen sein, z.B. eine Scheibe im $\mathbb{R}^3$.

Bemerkung A17.3 (Singulärwerte und Eigenwerte)

a. Für eine symmetrische positiv semi-definite Matrix A liefert die Eigenwertzerlegung eine Singulärwertzerlegung.

b. Für eine symmetrische Matrix A sind die Singulärwerte gerade die Beträge der Eigenwerte von A.

c. Sei $A = U \circ \Sigma \circ V^t$ eine Singulärwertzerlegung. Dann gilt:

$$A \circ A^t = (U \circ \Sigma \circ V^t) \circ (V \circ \Sigma^t \circ U^t) = U \circ \Sigma \circ \Sigma^t \circ U^t$$

und analog

$$A^t \circ A = V \circ \Sigma^t \circ \Sigma \circ V^t .$$

Daher entsprechen die Quadrate der Singulärwerte $\sigma_1, \ldots, \sigma_r$ Eigenwerten von $A \circ A^t$ und $A^t \circ A$ zu den Eigenvektoren $x_1, \ldots, x_r$ bzw. $y_1, \ldots, y_r$.

Für den Beweis der Existenz der Singulärwertzerlegung nutzen wir einige Eigenschaften der Matrix $A^t \circ A$, die wir zunächst zeigen wollen.

Lemma A17.4 ($A^t \circ A$ ist symmetrisch.)
Sei $A \in \mathrm{Mat}(m \times n, \mathbb{R})$ eine Matrix. Dann ist die Matrix $A^t \circ A \in \mathrm{Mat}_n(\mathbb{R})$ symmetrisch, alle Eigenwerte sind nicht-negativ und es gelten

$$\mathrm{Ker}(f_A) = \mathrm{Ker}(f_{A^t \circ A})$$

sowie

$$\mathrm{rang}(A) = \mathrm{rang}(A^t \circ A).$$

Beweis: Aus der Gleichung

$$(A^t \circ A)^t = A^t \circ (A^t)^t = A^t \circ A$$

folgt, daß $A^t \circ A$ symmetrisch ist. Zudem gilt

$$x^t \circ (A^t \circ A) \circ x = \langle A \circ x, A \circ x \rangle \geq 0$$

für alle $x \in \mathbb{R}^n$.[7] Ist $\lambda \in \mathbb{R}$ ein Eigenwert von $A^t \circ A$ mit Eigenvektor $x \in \mathbb{R}^n$, dann folgt zudem

$$\lambda \cdot \circ \langle x, x \rangle = \lambda \cdot x^t \circ x = x^t \circ (\lambda \cdot x) = x^t \circ A^t \circ A \circ x \geq 0.$$

Als nächstes wollen wir die folgende Gleichung zeigen:

$$\mathrm{Ker}(f_A) = \mathrm{Ker}(f_{A^t \circ A}).$$

Ist $x \in \mathrm{Ker}(f_A)$, so gilt auch

$$f_{A^t \circ A}(x) = A^t \circ A \circ x = A^t \circ 0 = 0.$$

Für die umgekehrte Inklusion wählen wir nun einen Vektor $x \in \mathrm{Ker}(f_{A^t \circ A})$ und erhalten

$$||f_A(x)|| = \langle A \circ x, A \circ x \rangle = x^t \circ A^t \circ A \circ x = x^t \circ 0 = 0,$$

woraus $A \circ x = 0$ und somit $x \in \mathrm{Ker}(f_A)$ folgt.

[7] Eine solche Matrix nennt man *positiv semidefinit*.

Damit ist dann schließlich aber auch

$$\operatorname{rang}(A) = n - \dim_{\mathbb{R}} \operatorname{Ker}(f_A) = n - \dim_{\mathbb{R}}(\operatorname{Ker}(f_{A^t \circ A})) = \operatorname{rang}(A^t \circ A)$$

gezeigt. □

Beweis von Satz A17.1: Nach dem Spektralsatz für symmetrische Matrizen A15.32 gibt es für die symmetrische Matrix $A^t \circ A$ eine orthogonalen Matrix $V \in \mathrm{O}(n)$, so daß

$$V^t \circ A^t \circ A \circ V = V^{-1} \circ A^t \circ A \circ V = \Lambda$$

eine Diagonalmatrix $\Lambda \in \mathrm{Mat}_n(\mathbb{R})$ ist, auf deren Diagonale die Eigenwerten $\lambda_1 \geq \ldots \geq \lambda_r > \lambda_{r+1} = \ldots = \lambda_n = 0$ von $A^t \circ A$ stehen, die wir o.E. der Größe nach sortiert annehmen dürfen. Wegen Lemma A17.4 ist dabei

$$r = \operatorname{rang}(A^t \circ A) = \operatorname{rang}(A).$$

Die Spalten $y_1, \ldots, y_n$ von V sind zudem die zugehörigen Eigenvektoren und sie bilden eine ONB von $\mathbb{R}^n$. Bringen wir V^t und V auf die andere Seite, so erhalten wir die Eigenwertzerlegung

$$A^t \circ A = V \circ \Lambda \circ V^t.$$

Wir konstruieren die gewünschte orthogonale Matrix U mit Spalten $(x_1, \ldots, x_m)$ explizit. Für $i = 1, \ldots, r$ setzen wir $x_i = \frac{1}{\sigma_i} \cdot A \circ y_i$ mit $\sigma_i = \sqrt{\lambda_i} > 0$. Diese Vektoren sind orthonormal wegen

$$\langle x_i, x_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle A y_i, A y_j \rangle = \frac{1}{\sigma_i \sigma_j} \langle A^t A y_i, y_j \rangle = \frac{\lambda_i}{\sigma_i \sigma_j} \langle y_i, y_j \rangle = \begin{cases} 1\,, & \text{wenn } i = j\,, \\ 0\,, & \text{sonst}\,. \end{cases}$$

Mittels des Gram–Schmidt Orthogonalisierungsverfahrens ergänzen wir $x_1, \ldots, x_r$ zu einer Orthonormalbasis des $\mathbb{R}^m$.

Es bleibt zu zeigen, daß die Konstruktion das Richtige tut. Dazu beachten wir, daß

$$y_{r+1}, \ldots, y_n \in \operatorname{Ker}(f_{A^t \circ A}) = \operatorname{Ker}(f_A)$$

und daß deshalb

$$A \circ y_i = f_A(y_i) = 0 \tag{252}$$

gilt für $i = r+1, \ldots, n$. Da $(y_1, \ldots, y_n)$ eine ONB von $\mathbb{R}^n$ bezüglich des Standardskalarproduktes ist, gilt zudem

$$\sum_{i=1}^{n} y_i \circ y_i^t = \mathbb{1}_n.$$

Damit erhalten wir dann

$$U \circ \Sigma \circ V^t = \sum_{i=1}^{r} \sigma_i \cdot x_i \circ y_i^t = \sum_{i=1}^{r} (A \circ y_i) \circ y_i^t$$
$$\overset{(252)}{=} \sum_{i=1}^{n} (A \circ y_i) \circ y_i^t = A \circ \sum_{i=1}^{n} y_i \circ y_i^t = A.$$

□

Korollar A17.5

Sei $A = U \circ \Sigma \circ V^t$ wie in Satz A17.1 mit orthogonalen Matrizen $U = (x_1, \ldots, x_m)$, $V = (y_1, \ldots, y_n)$ und $\sigma_1 \geq \ldots \geq \sigma_r > 0$ mit $r = \operatorname{rang}(A)$. Dann gelten:

a. $\operatorname{Im}(f_A) = \operatorname{Lin}(x_1, \ldots, x_r)$,

b. $\operatorname{Ker}(f_A) = \operatorname{Lin}(y_{r+1}, \ldots, y_n)$.

Beweis: Folgt direkt aus dem Beweis von Satz A17.1. □

Bemerkung A17.6

a. Die Singulärwertzerlegung ist im allgemeinen nicht eindeutig. Kommt ein Eigenwert von $A^t A$ mit höherer Vielfachheit vor, so sind die zugehörigen orthonormierten Eigenvektoren nicht eindeutig. Auch die Ergänzung von $(x_1, \ldots, x_r)$ zu einer Basis ist bezüglich ihrer Reihenfolge nicht eindeutig bestimmt. Die Singulärwerte jedoch sind als Eigenwerte von $A^t \circ A$ eindeutig.

b. Die Singulärwertzerlegung lässt sich auch für komplexe Matrizen bestimmen. In diesem Fall sind U, V unitäre Matrizen und transponieren von V muss durch adjungieren ersetzt werden. Die Singulärwerte sind aber trotzdem reell.

Beispiel A17.7

a. Die Matrix $A = (-1, 2, 2)^t$ bestehend aus einer Spalte hat $\operatorname{rang}(A) = 1$. Aus $A^t A = (9)$ sehen wir, daß A genau einen positiven Singulärwert, $\sigma_1 = 3$, hat. Die zugehörige ONB des $\mathbb{R}^1$ aus Eigenvektoren ist $(y_1) = (1)$ und somit ist auch

$$V = (1).$$

Daraus ergibt sich für die erste Spalte von U dann

$$x_1 = \frac{1}{\sigma_1} \cdot A \circ y_1 = \frac{1}{3} \cdot (-1, 2, 2)^t.$$

Diese ergänzen wir zu einer ONB von $\mathbb{R}^3$ durch

$$x_2 = \frac{1}{3} \cdot (2, -1, 2)^t$$

und

$$x_3 = \frac{1}{3} \cdot (2, 2, -1)^t.$$

Als Singulärwertzerlegung von A erhalten wir dann

$$\begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix} = A = U \circ \Sigma \circ V^t = \frac{1}{3} \begin{pmatrix} -1 & 2 & 2 \\ 2 & -1 & 2 \\ 2 & 2 & -1 \end{pmatrix} \circ \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} \circ \begin{pmatrix} 1 \end{pmatrix}.$$

b. Wir möchten nun eine Singulärwertzerlegung für die Matrix

$$A = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \in \mathrm{Mat}(2 \times 3, \mathbb{R})$$

bestimmen. Dazu berechnen wir zunächst die Matrix

$$A^t \circ A = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{pmatrix} \in \mathrm{Mat}_3(\mathbb{R})$$

sowie deren charateristisches Polynom

$$\begin{aligned} \chi_{A^t A} =& \det(t \cdot \mathbb{1}_3 - A^t A) = \det \begin{pmatrix} t-1 & 1 & 0 \\ 1 & t-2 & 1 \\ 0 & 1 & t-1 \end{pmatrix} = t^3 - 4t^2 + 3t \\ =& t \cdot (t-1) \cdot (t-3). \end{aligned}$$

Die Eigenräume von $A^t \circ A$ zu den drei Eigenwerten $\lambda_1 = 3$, $\lambda_2 = 1$ und $\lambda_3 = 0$ müssen also jeweils ein-dimensional sein und es gilt

$$\mathrm{rang}(A) = \mathrm{rang}(A^t \circ A) = 2.$$

Wir können jeweils eine Basis mit Hilfe des Gauß-Algorithmus' bestimmen und erhalten:

$$\begin{aligned} \mathrm{Eig}(A, 3) =& \mathrm{Lin}\left((-1, 2, -1)^t\right), \\ \mathrm{Eig}(A, 1) =& \mathrm{Lin}\left((1, 0, -1)^t\right), \\ \mathrm{Eig}(A, 0) =& \mathrm{Lin}\left((-1, -1, -1)^t\right). \end{aligned}$$

Für die Singulärwerte erhalten wir

$$\sigma_1 = \sqrt{3}$$

und

$$\sigma_2 = \sqrt{1} = 1.$$

Die Matrix V erhalten wir, indem wir das Orthognonalisierungsverfahren von Gram-Schmidt auf die Basen der Eigenräume anwenden und diese dann als Spalten in die

Matrix V schreiben. Im vorliegenden Fall müssen wir die Vektoren nur normieren,

$$\begin{aligned} y_1 &= \frac{1}{\sqrt{6}} \cdot (-1, 2, -1)^t, \\ y_2 &= \frac{1}{\sqrt{2}} \cdot (1, 0, -1)^t, \\ y_3 &= \frac{1}{\sqrt{3}} \cdot (-1, -1, -1)^t \end{aligned}$$

und damit

$$V = \begin{pmatrix} -\frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{3}} \end{pmatrix} \in \mathrm{O}(3).$$

Die beiden Spalten von U ergeben sich aus den Singulärwerten also

$$\begin{aligned} x_1 &= \frac{1}{\sigma_1} \cdot A \circ y_1 = \frac{1}{\sqrt{2}} \cdot (1, -1)^t, \\ x_2 &= \frac{1}{\sigma_2} \cdot A \circ y_2 = \frac{1}{\sqrt{2}} \cdot (-1, -1)^t \end{aligned}$$

und somit

$$U = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \in \mathrm{O}(2).$$

Als Singulärwertzerlegung erhalten wir damit

$$\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{pmatrix} \circ \begin{pmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \circ \begin{pmatrix} -\frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & -\frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} & -\frac{1}{\sqrt{3}} \end{pmatrix}.$$

Beispiel A17.8 (Anwendung in der 3D-Rekonstruktion)
Wir betrachten eine Problemstellung der 3D-Rekonstruktion. Seien $P_1, \ldots, P_s$ Punkte auf einem Objekt in $\mathbb{R}^3$. Diese Punkte werden von t verschiedenen Kameras (bzw. von einer Kamera zu verschiedenen Zeitpunkten und aus verschiedenen Positionen) aufgenommen. Die "Projektion" eines Punktes P_j in die i-te Kamera bezeichnen wir mit $p_{ij} \in \mathbb{R}^2$, $i = 1, \ldots, t$ und $j = 1, \ldots, s$. Das Ziel der Anwendung ist die Rekonstruktion der 3D-Koordinaten basierend auf den in die Kameras projizierten 2D-Koordinaten.

Unter der Annahme eines affinen Kameramodells gilt für alle Punkte:

$$p_{ij} = M_i \circ P_j + q_i\,, \quad j = 1, \ldots, s\,,$$

wobei $M_i \in \mathrm{Mat}(2 \times 3, \mathbb{R})$ und $q_i \in \mathbb{R}^2$ ist; dabei beschreibt M_i die orthogonale Projektion auf eine Ebene, die, wenn wir sie um einen geeigneten Punkt q_i verschieben, die Aufnahmeebene der Kamera ist.

Da wir den Ursprung des 3D-Koordinatensystems frei wählen können, nehmen wir an, daß die Punkte P_j um den Ursprung herum zentriert sind, d.h. es gilt

$$\frac{1}{s} \cdot \sum_{j=1}^{s} P_j = 0.$$

Daraus ergibt sich mit obiger Gleichung für jede Kamera i direkt, daß die projizierten Punkte p_{ij} um

$$q_i = \frac{1}{s} \cdot \sum_{j=1}^{s} p_{ij}$$

zentriert sind, denn

$$\frac{1}{s} \cdot \sum_{j=1}^{s} p_{ij} = \frac{1}{s} \cdot \sum_{j=1}^{s} (M_i \circ P_j + q_i) = M_i \circ \underbrace{\left(\frac{1}{s} \cdot \sum_{j=1}^{s} P_j\right)}_{=0} + q_i = q_i.$$

In der Praxis haben wir nur die 2D-Koordinaten gegeben. Wir können diese 2D-Koordinaten problemlos zentrieren. Die resultierenden zentrierten 2D-Koordinaten sind

$$\hat{p}_{ij} := p_{ij} - \frac{1}{s} \cdot \sum_{j=1}^{s} p_{ij}\,.$$

Wir erhalten das folgende Modell für den Zusammenhang von 2D- und 3D-Koordinaten:

$$\underbrace{\begin{pmatrix} \hat{p}_{11} & \hat{p}_{12} & \cdots & \hat{p}_{1s} \\ \hat{p}_{21} & \hat{p}_{22} & \cdots & \hat{p}_{2s} \\ \vdots & \vdots & & \vdots \\ \hat{p}_{t1} & \hat{p}_{t2} & \cdots & \hat{p}_{ts} \end{pmatrix}}_{=:W \in \mathrm{Mat}(2t \times s, \mathbb{R})} = \underbrace{\begin{pmatrix} M_1 \\ M_2 \\ \vdots \\ M_t \end{pmatrix}}_{=:M \in \mathrm{Mat}(2t \times 3, \mathbb{R})} \underbrace{\begin{pmatrix} \hat{P}_1 & \hat{P}_2 & \cdots & \hat{P}_s \end{pmatrix}}_{=:S \in \mathrm{Mat}(3 \times s, \mathbb{R})},$$

wobei M die Matrix der Kamerabewegungen ist und S als Strukturmatrix bezeichnet wird. Offensichtlich ist der Rang von M höchstens 3. Daher liegen die so genannten Trajektorien[8] $(\hat{p}_{1j}, \ldots, \hat{p}_{tj})^t$, die Spalten von W, auf einem 3-dimensionalen Unterraum von $\mathbb{R}^{2t}$.

Ist W gegeben, so können wir M und S mithilfe der Singulärwertzerlegung $W = U \circ \Sigma \circ V^t$ bestimmen. Wir setzen

$$M = (\sigma_1 u^1, \sigma_2 u^2, \sigma_3 u^3) \in \mathrm{Mat}(2t \times 3, \mathbb{R})$$

und

$$S = (v^1, v^2, v^3)^t \in \mathrm{Mat}(3 \times s, \mathbb{R}),$$

und die Matrix S liefert also die gewünschte Rekonstruktion der 3D-Punkte. Diese ist bis auf affine Ambiguität eindeutig, d.h. wir können S durch $A \circ S$ für eine beliebige

[8] Eine Trajektorie besteht aus allen 2D-Koordinaten eines 3D Punktes in den verschiedenen Kameras.

invertierbare Matrix $A \in \mathrm{Gl}_3(\mathbb{R})$ ersetzen, wenn wir zugleich M durch $M \circ A^{-1}$ ersetzen. Insbesondere, wenn wir die M_j kennen, können wir auf dem Weg S eindeutig rekonstruieren.

Aufgaben

Aufgabe A17.9

Berechne eine Singulärwertzerlegung der Matrix

$$A = \begin{pmatrix} 2 & 0 & -10 \\ -11 & 0 & 5 \\ 0 & 3 & 0 \\ 0 & -4 & 0 \end{pmatrix} \in \mathrm{Mat}(4 \times 3, \mathbb{R}).$$

A18 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 Befehlt `exit`. Hilfe zur Syntax von SINGULAR findet man im Manual auf der SINGULAR-Webseite oder durch den Befehl `help`.

Beispiel A18.1 (Reduzierte Zeilen-Stufen-Form und Rang einer Matrix)
Wir wollen eine reduzierte Zeilen-Stufenform und damit den Rang der folgenden Matrix berechnen:

$$A = \begin{pmatrix} 1 & 2 & 1 & 3 & 1 \\ 2 & 4 & 7 & 3 & 0 \\ 4 & 8 & 9 & 9 & 2 \\ 3 & 6 & 0 & 2 & 1 \end{pmatrix} \in \mathrm{Mat}(4 \times 5, \mathbb{Q})$$

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 A18.2 (Kern einer Matrix)

Mit dem Befehl `syz` können wir eine Basis des Kerns der Matrix in Beispiel A18.1 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 A18.3 (Lösung eines linearen Gleichungssystems)

Wir setzen nun $b = (2,2,6,4)^t$ und wollen das lineare Gleichungssystem $Ax = b$ 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

$$c = (0, 0, 2, -4, 12)^t.$$

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 A18.2 wissen.

Beispiel A18.4 (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 A18.5 (Minimalpolynom einer Matrix)

Als nächstes wollen wir das Minimalpolynom der Matrix A in Beispiel A18.4 berechnen. Dazu verwenden wir den Algorithmus A12.16 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 A18.6 (Diagonalisierung einer Matrix)

Wir haben in Beispiel A18.5 gesehen, daß die Matrix A aus Beispiel A18.4 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 A18.7 (Jordansche Normalform)

In diesem Beispiel wollen wir die Jordansche Normalform und die Transformationsmatrix

T für die Matrix

$$A = \begin{pmatrix} 21 & 5 & 1 & 21 & 5 \\ -23 & 4 & 8 & -31 & 1 \\ -2 & -1 & -2 & -1 & -1 \\ -17 & -4 & -1 & -17 & -4 \\ 22 & -2 & -8 & 30 & 1 \end{pmatrix} \in \mathrm{Mat}_5(\mathbb{Q})$$

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

$$J_A = \left(\begin{array}{cc|c|cc} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right).$$

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 A13.15 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 $\mathrm{Ker}\left((A-3\cdot\mathbb{1}_5)^2\right)$ nicht im $\mathrm{Ker}(A-3\cdot\mathbb{1}_5)$ 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

$$(A - 3 \cdot \mathbb{1}_5)x_1 = (5, 1, -1, -4, -2)^t \quad \text{und} \quad x_1 = (0, 1, 0, 0, 0)^t,$$

und diese sind die ersten beiden Spalten der Matrix T.

Nun müssen wir noch den Vektor $(A - 3 \cdot \mathbb{1}_5)x_1$ zu einer Basis von $\mathrm{Ker}(A - 3 \cdot \mathbb{1}_5)$ ergänzen. Ein Blick auf die Basis von $\mathrm{Ker}(A-3\cdot\mathbb{1}_5)$ 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

$$x_3 = (0,-2,1,0,2)^t \in \operatorname{Ker}\left(A + \mathbb{1}_5^2\right) \setminus \operatorname{Ker}(A + \mathbb{1}_5)$$

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

$$J_A = T^{-1} \circ A \circ T = \left(\begin{array}{cc|c|cc} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ \hline 0 & 0 & 3 & 0 & 0 \\ \hline 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{array}\right).$$

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 A18.8 (Näherungsweise Bestimmung von Eigenwerten)

Eine zufällig ausgewählte Matrix in $\mathrm{Mat}_n(\mathbb{Q})$ wird keine rationalen Eigenwerte haben. Betrachen wir sie aber als Matrix in $\mathrm{Mat}_n(\mathbb{C})$, 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 häufit 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 A18.9

Bestimme mit Hilfe von SINGULAR eine Basis B von $\mathbb{Q}^5$, bezüglich derer die Matrixdarstellung der Abbildung $f : \mathbb{Q}^5 \to \mathbb{Q}^5$ Jordansche Normalform hat, wo:

$$f(x_1, x_2, x_3, x_4, x_5) =$$
$$(x_1 - x_2, x_1 + 2x_2 - x_3, -x_1 + 3x_3, -x_1 - 2x_2 - 2x_3 + 2x_4 - x_5, x_1 - x_3 + 2x_5)^t.$$

Kapitel B

Die Integralsätze von Green, Gauß und Stokes

B1 Kurvenintegrale im $\mathbb{R}^n$ und der Integralsatz von Green

A) Stückweise $\mathcal{C}^1$-Kurven

Definition B1.1 (Stückweise $\mathcal{C}^1$-Kurven)
Es sei $[a,b]$ ein Intervall in $\mathbb{R}$ und $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ eine Abbildung mit den Komponentenfunktionen $\gamma_i : [a,b] \longrightarrow \mathbb{R}$, $i = 1, \ldots, n$.

a. Ist γ stetig, so nennen wir γ *Parameterdarstellung einer Kurve* oder schlicht eine *Kurve* im $\mathbb{R}^n$, und wir nennen das Bild $\Gamma = \mathrm{Im}(\varphi)$ von γ die *Spur* der Kurve.

b. Ist γ stetig differenzierbar, so nennen wir die Kurve *stetig differenzierbar* oder eine $\mathcal{C}^1$-*Kurve* und wir bezeichnen

$$\gamma'(t) = \big(\gamma_1'(t), \ldots, \gamma_n'(t)\big)^t \in \mathbb{R}^n$$

als den *Tangentialvektor* oder *Geschwindigkeitsvektor* an die Kurve γ im Punkt $\gamma(t)$ und mit

$$l(\gamma) = \int_a^b ||\gamma'(t)||_2 \, dt$$

die *Länge der Kurve* γ.

c. Sind $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ und $\delta : [b,c] \longrightarrow \mathbb{R}^n$ zwei Kurven mit $\gamma(b) = \delta(c)$, so ist die *Summe* der beiden

$$\gamma + \delta : [a,c] \longrightarrow \mathbb{R}^n : t \mapsto \begin{cases} \gamma(t), & t \leq b, \\ \delta(t), & t > b \end{cases}$$

eine Kurve im $\mathbb{R}^n$, und wenn Γ die Spur von γ und Δ die Spur von δ ist, so bezeichnen wir die Spur von $\gamma + \delta$ kurz mit $\Gamma + \Delta$. Man beachte, daß hier mit $\gamma + \delta$ *nicht* die punktweise Summe der beiden Abbildungen gemeint ist!

d. Eine Kurve γ heißt *stückweise stetig differenzierbar* oder eine *stückweise $\mathcal{C}^1$-Kurve*, wenn es eine Zerlegung $a = t_0 < t_1 < \ldots < t_n = b$ von $[a, b]$ in endlich viele Teilintervalle $[t_{i-1}, t_i]$ gibt, so daß $\gamma_{|[t_{i-1},t_i]}$ stetig differenzierbar ist, d.h. wenn γ Summe von endlich vielen $\mathcal{C}^1$-Kurven ist. Wir bezeichnen mit

$$l(\gamma) = \sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} ||\gamma'(t)||_2 \, dt$$

dann die *Länge der Kurve* γ.

Beispiel B1.2 (Kurven und ihre Länge)

a. Sind $p, q \in \mathbb{R}^n$ zwei Punkte im $\mathbb{R}^n$, so ist die Parametrisierung

$$\gamma : [0, 1] \longrightarrow \mathbb{R}^n : t \mapsto t \cdot q + (1 - t) \cdot p$$

der Strecke von p nach q eine $\mathcal{C}^1$-Kurve im $\mathbb{R}^n$ mit konstantem Tangentialvektor

$$\gamma'(t) = q - p$$

und ihre Länge

$$l(\gamma) = \int_0^1 ||q - p||_2 \, dt = ||q - p||_2$$

ist der Abstand der beiden Punkte.

b. Der Kreis mit Radius $r > 0$ um den Mittelpunkt $p \in \mathbb{R}^2$ in der Ebene ist die Spur der $\mathcal{C}^1$-Kurve

$$\gamma : [0, 2\pi] \longrightarrow \mathbb{R}^2 : t \mapsto \big(r \cdot \cos(t) + p_1, r \cdot \sin(t) + p_2\big)^t$$

mit Tangentialvektor

$$\gamma'(t) = \big(- r \cot \sin(t), r \cdot \cos(t)\big)^t$$

in $\gamma(t)$ und Länge

$$l(\gamma) = \int_0^{2\pi} \sqrt{(-r)^2 \cdot \sin^2(t) + r^2 \cos^2(t)} \, dt = \int_0^{2\pi} r \, dt = 2\pi r.$$

c. Das Quadrat mit den Eckpunkten $(0, 0)^t, (1, 0)^t, (1, 1)^t, (0, 1)^t \in \mathbb{R}^2$ ist die Spur einer stückweisen $\mathcal{C}^1$-Kurve und hat wegen Teil a. die Länge 4.

d. Die Helix

$$\gamma : [0, 6\pi] \longrightarrow \mathbb{R}^3 : t \mapsto \big(\cos(t), \sin(t), t\big)^t$$

ist eine $\mathcal{C}^1$-Kurve im $\mathbb{R}^3$ der Länge

$$l(\gamma) = \int_0^{6\pi} \sqrt{\sin^2(t) + \cos^2(t) + 1}\, dt = \int_0^{6\pi} \sqrt{2}\, dt = 6\pi\sqrt{2}.$$

Bemerkung B1.3 (Länge der Kurve)
Um die Definition der Länge einer Kurve zu verstehen, approximieren wir die Kurve $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ durch einen Polygonzug, indem wir für das Intervall $[a,b]$ eine äquidistante Zerlegung $a = t_0 < t_1 < \ldots < t_n = b$ wählen und den zugehörigen Polygonzug Γ mit den Ecken $\gamma(t_0), \ldots, \gamma(t_n)$ betrachten. Die Länge des Polygonzugs ist wegen Beispiel B1.2 die Summe

$$l(\Gamma) = \sum_{i=0}^{n-1} ||\gamma(t_{i+1}) - \gamma(t_i)||_2 = \sum_{i=0}^{n-1} \left|\left| \frac{\gamma(t_i + \Delta t) - \gamma(t_i)}{\Delta t} \right|\right|_2 \cdot \Delta t,$$

wenn Δt den Abstand von zwei aufeinanderfolgenden Stützstellen bezeichnet. In der Infinitesimalrechnung wird beim Grenzübergang aus dem Differenzenquotienten die Ableitung, aus der Summe wird das Integral und aus Δt wird dt und wir erhalten

$$l(\Gamma) = \int_a^b ||\gamma'(t)||_2\, dt.$$

Definition B1.4 (Äquivalenz und Orientierung von Kurven)
Zwei stückweise $\mathcal{C}^1$-Kurven $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ und $\delta : [c,d] \longrightarrow \mathbb{R}^n$ heißen *äquivalent*, wenn es eine bijektive, stetig differenzierbare Abbildung

$$\varphi : [a,b] \longrightarrow [c,d]$$

mit $\varphi'(t) \neq 0$ für $t \in [a,b]$ und

$$\gamma = \delta \circ \varphi$$

gibt. Ist φ' stets positiv, so sagen wir, daß γ und δ *dieselbe Orientierung* haben; ist φ' stets negativ, so sagen wir, daß γ und δ *entgegengesetzte Orientierung* haben. Wegen des Zwischenwertsatzes muß einer der beiden Fälle eintreten. Man beachte zudem, daß äquivalente Kurven dieselbe Spur haben.

Lemma B1.5 (Unabhängigkeit der Länge von der Parametrisierung)
Zwei äquivalente Kurven haben sie dieselbe Länge.

Beweis: Verwenden wir die Bezeichnungen aus Definition B1.4, so folgt dies unmittelbar aus der Kettenregel und der Substitutionsregel

$$\begin{aligned} l(\gamma) &= \int_a^b ||\gamma'(t)||_2 \, dt = \int_a^b ||(\delta \circ \varphi)'(t)||_2 \, dt \\ &= \int_a^b ||\delta'(\varphi(t)) \cdot \varphi'(t)||_2 \, dt = \int_a^b ||\delta'(\varphi(t))||_2 \cdot |\varphi'(t)| \, dt \\ &= \int_c^d ||\delta'(s)||_2 \, ds, \end{aligned}$$

wobei man beachte, daß das Vorzeichen von $\varphi'(t)$ ggf. durch die Reihenfolge der Integrationsgrenzen kompensiert wird. □

Bemerkung B1.6 (Bogenlänge)

Ist $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ eine $\mathcal{C}^1$-Kurve, deren Ableitung nie der Nullvektor wird, so ist die *Bogenlänge*

$$s : [a,b] \longrightarrow \mathbb{R} : t \mapsto \int_a^t ||\gamma'(\tau)||_2 \, d\tau$$

eine stetig differenzierbare Funktion mit $s'(t) = ||\gamma'(t)||_2 > 0$ für alle $t \in [a,b]$ und somit streng monoton wachsend. Sie besitzt wegen des Umkehrsatzes für streng monotone stetige Funktionen 14.21 und des Satzes über die Ableitung der Umkehrfunktion 17.14 eine stetig differenzierbare Umkehrfunktion

$$s^{-1} : [0, l(\gamma)] \longrightarrow \mathbb{R}$$

und die Kurve

$$\widetilde{\gamma} = \gamma \circ s^{-1}$$

ist eine zu γ äquivalente Kurve mit Geschwindigkeitsvektor der konstanten Länge

$$||\widetilde{\gamma}'(t)||_2 = \left|\left|\gamma'(s^{-1}(t)) \cdot (s^{-1})'(t)\right|\right|_2 = \left|\left|\gamma'(s^{-1}(t)) \cdot \frac{1}{s'(s^{-1}(t))}\right|\right|_2 = 1.$$

Bemerkung B1.7 (Das Negative einer Kurve)

Ist $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ eine stückweise $\mathcal{C}^1$-Kurve mit Spur Γ, so nennen wir die Kurve

$$-\gamma : [a,b] \longrightarrow \mathbb{R}^n : t \mapsto \gamma(a+b-t) \tag{253}$$

das *Negative* von γ und bezeichnen ihre Spur mit $-\Gamma$. Sie ist äquivalent zu γ mit umgekehrter Orientierung. Man beachte, daß hier mit $-\gamma$ *nicht* das punktweise Negative der Abbildung γ gemeint ist!

B) Kurvenintegrale

Definition B1.8 (Kurvenintegrale längs einer Kurve)

a. Eine Abbildung $f : A \longrightarrow \mathbb{R}^n$ mit $A \subseteq \mathbb{R}^n$ nennen wir ein *Vektorfeld*.

b. Eine Abbildung $f : A \longrightarrow \mathbb{R}$ mit $A \subseteq \mathbb{R}^n$ nennen wir ein *Skalarfeld*.

c. Ist $\gamma : [a, b] \longrightarrow \mathbb{R}^n$ eine $\mathcal{C}^1$-Kurve mit Spur Γ und $f : \Gamma \longrightarrow \mathbb{R}^n$ ein stetiges Vektorfeld, so nennen wir

$$\int_\gamma f \cdot \mathbf{dx} := \int_\gamma \sum_{i=1}^{n} f_i \, dx_i := \int_a^b \langle f(\gamma(t)), \gamma'(t) \rangle \, dt$$

das *(vektorielle) Kurvenintegral* von f längs der Kurve γ.

d. Ist $\gamma : [a, b] \longrightarrow \mathbb{R}^n$ eine $\mathcal{C}^1$-Kurve mit Spur Γ und $f : \Gamma \longrightarrow \mathbb{R}$ ein stetiges Skalarfeld, so nennen wir

$$\int_\gamma f \, ds := \int_a^b f(\gamma(t)) \cdot ||\gamma'(t)||_2 \, dt$$

das *(skalare) Kurvenintegral* von f längs der Kurve γ *bezüglich der Bogenlänge*.

e. Ist die Kurve γ in b. nur eine stückweise $\mathcal{C}^1$-Kurve, so definieren wir das *Kurvenintegral* von f längs der Kurve γ als die Summe der Kurvenintegrale von f längs der endlich vielen $\mathcal{C}^1$-Stücke der Kurve γ.

Bemerkung B1.9 (Vektorfeld)

Hinter der Bezeichnung Vektorfeld für eine Abbildung von $\mathbb{R}^n$ nach $\mathbb{R}^n$ steht die Interpretation, daß die Abbildung f an jeden Vektor x im Definitionsbereich einen Vektor, nämlich das Bild $f(x)$ von x unter f anheftet (siehe Abbildung 1). Das Vektorfeld kann etwa Richtung und Geschwindigkeit einer fließenden Flüssigkeit beschreiben.

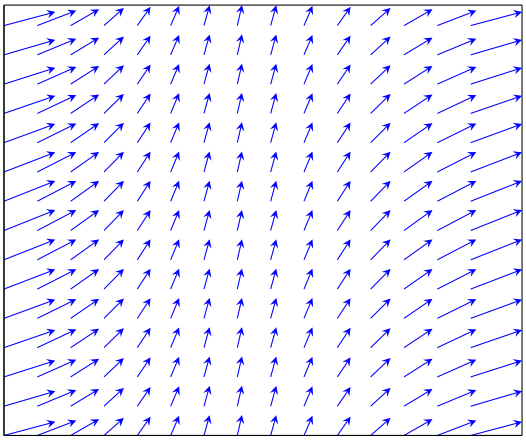


Abbildung 1. Das Vektorfeld $f(x, y) = \left(\frac{1+10x^2}{10}, \frac{1}{2+x^2y^2}\right)^t$

Bemerkung B1.10 (Arbeit = Kraft x Weg)

Bewegen wir einen Massepunkt durch ein Kraftfeld $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ entlang einer Kurve γ, so wird die verrichtete Arbeit durch das vektorielle Kurvenintegral

$$\int_\gamma f \cdot \mathbf{dx}$$

von f längs γ beschrieben. Die Idee dabei ist, daß im Punkt $\gamma(t)$ das Skalarprodukt aus dem Vektor $f(\gamma(t))$ der Kraft mit dem Geschwindigkeitsvektor $\gamma'(t)$ des Massepunktes die geleistete Arbeit zum Zeitpunkt t beschreibt.

Bemerkung B1.11 (Masse eines Drahtes)

Wird die Lage eines Drahtes im Raum durch die Kurve γ beschrieben und beschreibt $f(t)$ für das Skalarfeld $f : \mathbb{R}^3 \longrightarrow \mathbb{R}$ die Massendichte des Drahtes im Punkt $\gamma(t)$, so ist die totale Masse des Drahtes durch das skalare Kurvenintegral

$$\int_\gamma f \, ds$$

von f längs γ bezüglich der Bogenlänge gegeben. Dies ergibt sich aus der offensichtlichen Verallgemeinerung des Ansatzes in Bemerkung B1.3.

Beispiel B1.12 (Kurvenintegrale)

Gegeben seien das Vektorfeld

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x, y)^t \mapsto (x, y - x)^t$$

und die Kurve

$$\gamma : [0, 1] \longrightarrow \mathbb{R}^2 : t \mapsto (t, t^2)^t,$$

d.h. das Stück der Normalparabel vom Ursprung zum Punkt $(1, 1)^t$. Das vektorielle Kurvenintegral von f längs γ hat dann den Wert

$$\begin{aligned}\int_\gamma f \cdot \mathbf{dx} = \int_\gamma x \, dx + (y - x) \, dy &= \int_0^1 \langle (t, t^2 - t)^t, (1, 2t)^t \rangle \, dt \\ &= \int_0^1 t + 2t^3 - 2t^2 \, dt = \frac{t^2}{2} + \frac{t^4}{2} - \frac{2t^3}{3}\Big|_0^1 = \frac{1}{3}.\end{aligned}$$

Sei zudem das Skalarfeld

$$g : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y)^t \mapsto 2x \cdot \sqrt{1 + 4y}$$

gegeben, so hat das skalare Kurvenintegral von g längs γ den Wert

$$\int_\gamma g \, ds = \int_0^1 2t \cdot \sqrt{1 + 4t^2} \cdot \sqrt{1 + 4t^2} \, dt = \int_0^1 2t + 8t^3 \, dt = t^2 + 2t^4\Big|_0^1 = 3.$$

Proposition B1.13 (Rechenregeln für Kurvenintegrale)
Seien $\gamma : [a,b] \longrightarrow \mathbb{R}^n$ und $\delta : [c,d] \longrightarrow \mathbb{R}^n$ äquivalente stückweise $\mathcal{C}^1$-Kurve mit Spur Γ und sei $\epsilon = 1$, wenn sie dieselbe Orientierung haben, und sei $\epsilon = -1$ sonst.

a. Für ein stetiges Vektorfeld $f : \Gamma \longrightarrow \mathbb{R}^n$ gelten die Formeln

$$\int_\gamma f \cdot \mathbf{dx} = \epsilon \int_\delta f \cdot \mathbf{dx}$$

und

$$\left| \int_\gamma f \cdot \mathbf{dx} \right| \leq \sup_{x \in \Gamma} ||f(x)||_2 \cdot l(\gamma).$$

Für eine Zerlegung $\gamma = \kappa + \omega$ der Kurve gilt zudem

$$\int_\gamma f \cdot \mathbf{dx} = \int_\kappa f \cdot \mathbf{dx} + \int_\omega f \cdot \mathbf{dx}.$$

b. Für ein stetiges Skalarfeld $f : \Gamma \longrightarrow \mathbb{R}$ gelten die Formeln

$$\int_\gamma f \, ds = \int_\delta f \, ds$$

und

$$\left| \int_\gamma f \, ds \right| \leq \sup_{x \in \Gamma} |f(x)| \cdot l(\gamma).$$

Für eine Zerlegung $\gamma = \kappa + \omega$ der Kurve gilt zudem

$$\int_\gamma f \, ds = \int_\kappa f \, ds + \int_\omega f \, ds.$$

Beweis: a. Die letzte Gleichung folgt unmittelbar aus der Definition, und für die verbleibenden Formeln können wir deshalb annehmen, daß die Kurve stetig differenzierbar ist.

Ist φ wie in Definition B1.4, so folgt mit der Kettenregel und der Substitutionsregel

$$\begin{aligned}
\int_\gamma f \cdot \mathbf{dx} &= \int_a^b \langle f(\gamma(t)), \gamma'(t) \rangle \, dt = \int_a^b \langle f((\delta \circ \varphi)(t)), (\delta \circ \varphi)'(t) \rangle \, dt \\
&= \int_a^b \langle f(\delta(\varphi(t))), \delta'(\varphi(t)) \cdot \varphi'(t) \rangle \, dt \\
&= \int_a^b \langle f(\delta(\varphi(t))), \delta'(\varphi(t)) \rangle \cdot \varphi'(t) \, dt \\
&= \int_{\varphi(a)}^{\varphi(b)} \langle f(\delta(s)), \delta'(s) \rangle \, ds = \epsilon \int_c^d \langle f(\delta(s)), \delta'(s) \rangle \, ds = \epsilon \int_\delta f \cdot \mathbf{dx}.
\end{aligned}$$

Für die Abschätzung verwenden wir die Ungleichung von Cauchy-Schwarz A14.8

$$\begin{aligned}\left|\int_\gamma f\cdot\mathbf{dx}\right| &= \left|\int_a^b \langle f(\gamma(t)),\gamma'(t)\rangle\,dt\right| \\ &\le \int_a^b |\langle f(\gamma(t)),\gamma'(t)\rangle|\,dt \le \int_a^b ||f(\gamma(t))||_2\cdot||\gamma'(t)||_2\,dt \\ &\le \int_a^b \sup_{x\in\Gamma}||f(x)||_2\cdot||\gamma'(t)||_2\,dt = \sup_{x\in\Gamma}||f(x)||_2\cdot l(\gamma).\end{aligned}$$

b. Die letzte Gleichung folgt wieder unmittelbar aus der Definition, und für die verbleibenden Formeln können wir deshalb wieder annehmen, daß die Kurve stetig differenzierbar ist.
Für skalare Kurvenintegrale erhalten wir die Wegunabhängigkeit wie im Beweis von Lemma B1.5, wobei wir φ wie in Definition B1.4 wählen:

$$\begin{aligned}\int_\gamma f\,ds &= \int_a^b f(\gamma(t))\cdot||\gamma'(t)||_2\,dt = \int_a^b f((\delta\circ\varphi)(t))\cdot||(\delta\circ\varphi)'(t)||_2\,dt \\ &= \int_a^b f(\delta(\varphi(t)))\cdot||\delta'(\varphi(t))\cdot\varphi'(t)||_2\,dt \\ &= \int_a^b f(\delta(\varphi(t)))\cdot||\delta'(\varphi(t))||_2\cdot|\varphi'(t)|\,dt = \int_c^d f(\delta(s))\cdot||\delta'(s)||_2\,ds.\end{aligned}$$

Die Abschätzung folgt aus der Dreiecksungleichung für Integrale:

$$\begin{aligned}\left|\int_\gamma f\,ds\right| &= \left|\int_a^b f(\gamma(t))\cdot||\gamma'(t)||_2\,dt\right| \le \int_a^b |f(\gamma(t))|\cdot||\gamma'(t)||_2\,dt \\ &\le \int_a^b \sup_{x\in\Gamma}|f(x)|\cdot||\gamma'(t)||_2\,dt = \sup_{x\in\Gamma}|f(x)|\cdot l(\gamma).\end{aligned}$$

□

Bemerkung B1.14 (Zusammenhang skalares und vektorielles Kurvenintegral)
Ist $\gamma:[a,b]\longrightarrow\mathbb{R}^n$ eine $\mathcal{C}^1$-Kurve mit $\gamma'(t)\neq 0$ für alle $t\in[a,b]$ und ist

$$T(\gamma(t)) := \frac{\gamma'(t)}{||\gamma'(t)||_2}$$

der normierte Tangentialvektor von γ, dann gilt

$$\int_\gamma f\cdot\mathbf{dx} = \int_a^b \langle f(\gamma(t)),\gamma'(t)\rangle\,dt = \int_a^b \langle f(\gamma(t)),T(\gamma(t))\rangle\cdot||\gamma'(t)||_2\,dt = \int_\gamma \langle f,T\rangle\,ds.$$

Beispiel B1.15
Wir können den Zusammenhang in Bemerkung B1.14 auch verwenden, um vektorielle Kurvenintegrale leichter zu berechnen. Integrieren wir etwa das Vektorfeld

$$\mathrm{id}:\mathbb{R}^2\longrightarrow\mathbb{R}^2:(x,y)^t\mapsto(x,y)^t$$

längs des Einheitskreises mit Standardparametrisierung γ, dann wissen wir, daß das Vektorfeld in jedem Punkt $(x,y)^t$ des Einheitskreises senkrecht auf dem normierten Tangentialvektor $(-y,x)^t$ steht. Mithin gilt

$$\int_\gamma \mathrm{id}\cdot\mathbf{dx} = \int_\gamma 0\; ds = 0.$$

C) Konservative Vektorfelder

Definition B1.16

Sei G eine Teilmengen von $\mathbb{R}^n$.

a. G heißt *wegzusammenhängend*, wenn es zu je zwei Punkten $p,q \in G$ eine Kurve $\gamma : [a,b] \longrightarrow G$ mit $\gamma(a) = p$ und $\gamma(b) = q$ gibt.

b. G heißt ein *Gebiet*, wenn G offen in $\mathbb{R}^n$ und wegzusammenhängend ist.

c. G heißt *sternförmig*, wenn es einen Punkt $p \in G$ gibt, so daß für jeden Punkt $q \in G$ die Verbindungsstrecke $\overline{pq}$ in G liegt.

Beispiel B1.17

a. Offene konvexe Mengen sind sternförmig.

b. Die geschlitzte Ebene $\mathbb{R}^2 \setminus \{(x,0)^t \mid x \leq 0\}$ ist sternförmig mit $p = (1,0)^t$, ist aber nicht konvex.

c. Sternförmige Mengen sind Gebiete.

d. Die punktierte Ebene $\mathbb{R}^2 \setminus \{(0,0)^t\}$ ist ein Gebiet, ist aber nicht sternförmig.

Lemma B1.18 (Wege in Gebieten)

Ist $G \subseteq \mathbb{R}^n$ ein Gebiet, so gibt es für je zwei Punkte $p,q \in G$ eine stückweise $\mathcal{C}^1 - Kurve$ in G, die p und q verbindet.

Beweis: Seien $p,q \in G$ gegeben. Da G wegzusammenhängend ist, gibt es eine stetige Abbildung

$$\gamma : [a,b] \longrightarrow G$$

mit $\gamma(a) = p$ und $\gamma(b) = q$. Die Spur $\Gamma = \mathrm{Im}(\gamma)$ von γ ist als stetiges Bild eines kompakten Intervalls eine kompakte Teilmenge von $\mathbb{R}^n$. Da G offen ist mit $\Gamma \subset G$, gibt es für jeden Punkt $x \in \Gamma$ eine Umgebung $U_{\varepsilon_x}(x) \subseteq G$, die ganz in G enthalten ist. Die offene Überdeckung

$$\Gamma \subset \bigcup_{x\in\Gamma} U_{\varepsilon_x}(x)$$

der kompakten Menge Γ besitzt dann eine endliche Teilüberdeckung

$$\Gamma \subseteq \bigcup_{i=1}^{m} U_{\varepsilon_{x_i}}(x_i).$$

In diesen können wir den Weg γ durch einen Polygonzug ersetzen. □

Definition B1.19 (Gradientenfeld und konservatives Feld)
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \longrightarrow \mathbb{R}^n$ ein stetiges Vektorfeld.

a. f heißt ein *Gradientenfeld*, wenn es eine stetig differenzierbare Funktion $F : G \longrightarrow \mathbb{R}$ gibt mit $f = \operatorname{grad}(F)$. In diesem Fall heißt F eine *Stammfunktion* von f und $-F$ heißt ein *Potential* von f.

b. f heißt *konservativ* oder *wegunabhängig*, wenn für je zwei stückweise $\mathcal{C}^1$-Kurven γ und δ in G mit Anfangspunkt p und Endpunkt q schon

$$\int_\gamma f \cdot \mathbf{dx} = \int_\delta f \cdot \mathbf{dx}$$

gilt. Wir setzen dann

$$\int_p^q f \cdot \mathbf{dx} := \int_\gamma f \cdot \mathbf{dx}.$$

Lemma B1.20 (Wegunabhängigkeit)
Ein stetiges Vektorfeld $f : G \longrightarrow \mathbb{R}^n$ auf einem Gebiet G ist genau dann konservativ, wenn

$$\int_\gamma f \cdot \mathbf{dx} = 0$$

für jede geschlossene stückweise $\mathcal{C}^1$-Kurve γ in G.

Beweis: Ist $\gamma : [a, b] \longrightarrow G$ eine geschlossene stückweise $\mathcal{C}^1$-Kurve und ist f konservativ, so stimmen das Kurvenintegral von f längs γ und das längs des konstanten Weges

$$\delta : [a, b] \longrightarrow G : t \mapsto \gamma(a)$$

überein, aber wegen $\delta'(t) = 0$ für alle t gilt offenbar

$$\int_\delta f \cdot \mathbf{dx} = 0.$$

Ist umgekehrt das Kurvenintegral von f längs jeder geschlossenen stückweise $\mathcal{C}^1$-Kurve Null und sind γ und δ zwei stückweise $\mathcal{C}^1$-Kurven in G mit jeweils demselben Anfangspunkt und Endpunkt, so ist $\gamma - \delta$ eine geschlossene stückweise $\mathcal{C}^1$-Kurve und mit Proposition B1.13 folgt dann

$$0 = \int_{\gamma-\delta} f \cdot \mathbf{dx} = \int_\gamma f \cdot \mathbf{dx} - \int_\delta f \cdot \mathbf{dx}.$$

□

Bemerkung B1.21 (Konservatives Feld)
Wenn ein Feld konservativ ist, dann bedeutet das, daß die Kraft entlang eines geschlossenen Weges keine Arbeit leistet, oder alternativ, daß die Energie eines ruhenden Körpers nur von seiner Position abhängt und nicht davon, wie der Körper dorthin gekommen ist.

Beispiel B1.22 (Ein nicht konservatives Feld)
Wir betrachten das stetig differenzierbare Vektorfeld

$$f : \mathbb{R}^2 \setminus \{(0,0)^t\} \longrightarrow \mathbb{R}^2 : (x,y) \mapsto \left(\frac{-y}{x^2+y^2}, \frac{x}{x^2+y^2} \right)^t$$

und die $\mathcal{C}^1$-Kurve

$$\gamma : [0, 2\pi] \longrightarrow \mathbb{R}^2 : t \mapsto \big(\cos(t), \sin(t)\big)^t,$$

deren Spur der Einheitskreis ist. Das Kurvenintegral von f längs γ hat den Wert

$$\int_\gamma f \cdot \mathbf{dx} = \int_0^{2\pi} \langle (-\sin(t), \cos(t))^t, (-\sin(t), \cos(t))^t \rangle \, dt = \int_0^{2\pi} 1 \, dt = 2\pi.$$

Das Vektorfeld ist also nicht konservativ auf $\mathbb{R}^2 \setminus \{(0,0)^t\}$.

Satz B1.23 (Gradientenfelder sind die konservativen Felder.)
Ein stetiges Vektorfeld $f : G \longrightarrow \mathbb{R}^n$ auf einem Gebiet $G \subseteq \mathbb{R}^n$ ist genau dann konservativ, wenn f ein Gradientenfeld ist.
Zudem ist dann für jeden Punkt $p \in G$ die Abbildung

$$G \longrightarrow \mathbb{R} : x \mapsto \int_p^x f \cdot \mathbf{dx}$$

eine Stammfunktion von f, und für jede Stammfunktion F von f und jede stückweise $\mathcal{C}^1$-Kurve $\gamma : [a,b] \longrightarrow G$ gilt

$$\int_\gamma f \cdot \mathbf{dx} = F(\gamma(b)) - F(\gamma(a)).$$

Beweis: Sei zunächst F eine Stammfunktion von f und $\gamma : [a,b] \longrightarrow G$ eine $\mathcal{C}^1$-Kurve. Dann folgt mit Hilfe der Kettenregel und des Hauptsatzes der Differential- und Integralrechnung

$$\begin{aligned}
\int_\gamma f \cdot \mathbf{dx} &= \int_a^b \langle f(\gamma(t)), \gamma'(t) \rangle \, dt = \int_a^b \langle \operatorname{grad}(F)(\gamma(t)), \gamma'(t) \rangle \, dt \\
&= \int_a^b DF(\gamma(t)) \circ \gamma'(t) \, dt = \int_a^b D(F \circ \gamma)(t) \, dt \\
&= \int_a^b (F \circ \gamma)'(t) \, dt = F(\gamma(b)) - F(\gamma(a)).
\end{aligned}$$

Insbesondere hängt der Wert des Kurvenintegrals nur vom Anfangs- und vom Endpunkt der Kurve ab, so daß das Vektorfeld konservativ ist. Ist die Kurve nur stückweise stetig

differenzierbar, so wenden wir das Ergebnis auf die $\mathcal{C}^1$-Teilstücke an und erhalten die Wegunabhängigkeit.

Sei nun umgekehrt das Vektorfeld konservativ und sei $p \in G$ ein fest vorgegebener Punkt. Nach Lemma B1.18 können wir für jeden Punkt $x \in G$ eine stückweise $\mathcal{C}^1$-Kurve γ_x von p nach x wählen und definieren uns so die Abbildung

$$F : G \longrightarrow \mathbb{R} : x \mapsto \int_a^x f \cdot \mathbf{dx} := \int_{\gamma_x} f \cdot \mathbf{dx}.$$

Da das Vektorfeld konservativ ist, ist die Definition unabhängig von der konkreten Wahl der stückweise $\mathcal{C}^1$-Kurve γ_x. Wir müssen noch zeigen, daß F stetig differenzierbar ist mit

$$\operatorname{grad}(F) = f.$$

Sei dazu $x_0 \in G$ gegeben. Da G offen ist, gibt es ein $\varepsilon > 0$ mit

$$U_\varepsilon(x_0) \subseteq G.$$

Für jeden Punkt $x \in U_\varepsilon(x_0)$ bezeichne

$$\delta_x : [0,1] \longrightarrow U_\varepsilon(x_0) : t \mapsto t \cdot x + (1-t) \cdot x_0 = x_0 + t \cdot (x - x_0)$$

den Streckenzug von x_0 nach x. Dann ist $\gamma_{x_0} + \delta_x$ eine stückweise $\mathcal{C}^1$-Kurve von p nach x und da f konservativ ist, folgt

$$F(x) = \int_p^x f \cdot \mathbf{dx} = \int_{\gamma_{x_0}+\delta_x} f \cdot \mathbf{dx} = \int_{\gamma_{x_0}} f \cdot \mathbf{dx} + \int_{\delta_x} f \cdot \mathbf{dx} = F(x_0) + \int_{\delta_x} f \cdot \mathbf{dx}.$$

Wir erhalten damit

$$\begin{aligned}
\rho(x) :=& F(x) - F(x_0) - f(x_0)^t \circ (x - x_0) \\
=& \int_{\delta_x} f \cdot \mathbf{dx} - \langle f(x_0), x - x_0 \rangle \\
=& \int_0^1 \langle f(\delta_x(t)), \delta_x'(t) \rangle \, dt - \int_0^1 \langle f(x_0), x - x_0 \rangle \, dt \\
=& \int_0^1 \langle f(x_0 + t \cdot (x - x_0)), x - x_0 \rangle - \langle f(x_0), x - x_0 \rangle \, dt \\
=& \int_0^1 \langle f(x_0 + t \cdot (x - x_0)) - f(x_0), x - x_0 \rangle \, dt.
\end{aligned}$$

Mit Hilfe der Cauchy-Schwarzschen Ungleichung A14.8 folgt damit

$$\begin{aligned}\frac{|\rho(x)|}{||x-x_0||_2} &\leq \frac{\int_0^1 |\langle f(x_0+t\cdot(x-x_0))-f(x_0), x-x_0\rangle|\,dt}{||x-x_0||_2}\\ &\leq \frac{\int_0^1 ||f(x_0+t\cdot(x-x_0))-f(x_0)||_2\cdot||x-x_0||_2\,dt}{||x-x_0||_2}\\ &= \int_0^1 ||f(x_0+t\cdot(x-x_0))-f(x_0)||_2\,dt\\ &\leq \int_0^1 \sup\left\{||f(y)-f(x_0)||_2 \;\middle|\; ||y-x_0||_2 \leq ||x-x_0||_2\right\} dt\\ &= \sup\left\{||f(y)-f(x_0)||_2 \;\middle|\; ||y-x_0||_2 \leq ||x-x_0||_2\right\}.\end{aligned}$$

Da f in x_0 stetig ist, konvergiert die rechte Seite der Ungleichung für $x \to x_0$ gegen 0, somit gilt dies auch für die linke Seite. Damit ist dann aber gezeigt, daß F in x_0 differenzierbar ist mit $DF(x_0) = f(x_0)^t$. □

D) Integrabilitätskriterium

Wir wollen nun ein einfaches Kriterium dafür geben, daß ein Vektorfeld konservativ, d.h. ein Gradientenfeld, ist.

Satz B1.24 (Integrabilitätskriterium)
Sei $G \subseteq \mathbb{R}^n$ ein Gebiet und $f : G \longrightarrow \mathbb{R}^n$ sei ein stetig differenzierbares Vektorfeld.

a. Ist f ein Gradientenfeld, so gilt für alle $i, j \in \{1, \ldots, n\}$

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}. \tag{254}$$

b. Ist G sternförmig und gilt (254) für alle i, j, so ist f ein Gradientenfeld.

Beweis: a. Ist $F : G \longrightarrow \mathbb{R}$ eine Stammfunktion von f, so ist sie zweifach stetig differenzierbar und aus dem Satz von Schwarz 25.5 folgt

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial^2 F}{\partial x_j \partial x_i} = \frac{\partial^2 F}{\partial x_i \partial x_j} = \frac{\partial f_j}{\partial x_i}$$

für alle $i, j \in \{1, \ldots, n\}$.

b. Sei $p \in G$ ein Punkt, so daß für jedes $x \in G$ der Streckenzug γ_x von p nach x in G liegt. Wir können ohne Einschränkung annehmen, daß $p = 0$ der Ursprung ist. Dann gilt

$$\gamma_x : [0,1] \longrightarrow G : t \mapsto t \cdot x$$

und wir definieren

$$F : G \longrightarrow \mathbb{R} : x \mapsto \int_{\gamma_x} f \cdot \mathbf{dx} = \int_0^1 \langle f(tx), x\rangle \, dt = \sum_{i=1}^n \int_0^1 f_i(tx) \cdot x_i \, dt.$$

Wir wollen zeigen, daß F stetig differenzierbar mit $D_j F = f_j$ ist.
Die stetige Differenzierbarkeit nach x_j erhalten wir aus Aufgabe B1.38 angewandt auf die einzelnen Summanden, und die Aufgabe liefert uns zudem die Ableitung

$$\frac{\partial}{\partial x_j} \int_0^1 f_i(tx) \cdot x_i \, dt = \int_0^1 \frac{\partial}{\partial x_j}(f_i(tx) \cdot x_i) \, dt, \tag{255}$$

wobei

$$\frac{\partial}{\partial x_j}(f_i(tx) \cdot x_i) = \begin{cases} t \cdot \frac{\partial f_j}{\partial x_j}(tx) \cdot x_j) + f_j(tx), & i = j, \\ t \cdot \frac{\partial f_i}{\partial x_j}(tx) \cdot x_i), & i \neq j. \end{cases} \tag{256}$$

Für die partielle Ableitung von F ergibt sich unter Beachtung der Kettenregel (KR) und der Produktregel (PR) damit

$$\begin{aligned} D_j F(x) &= \sum_{i=1}^n \frac{\partial}{\partial x_j} \int_0^1 f_i(tx) \cdot x_i \, dt \overset{(255),(256)}{=} \int_0^1 f_j(tx) + \sum_{i=1}^n t \cdot \frac{\partial f_i}{\partial x_j}(tx) \cdot x_i) \, dt \\ &\overset{(254)}{=} \int_0^1 f_j(tx) + t \cdot \sum_{i=1}^n \frac{\partial f_j}{\partial x_i}(tx) \cdot x_i) \, dt \overset{(KR)}{=} \int_0^1 f_j(tx) + t \cdot \frac{\partial f_j}{\partial t}(tx) \, dt \\ &\overset{(PR)}{=} \int_0^1 \frac{\partial}{\partial t}(t \cdot f_j)(tx) \, dt = f_j(x). \end{aligned}$$

□

Beispiel B1.25 (Ein nicht-konservatives Feld mit Integrabilitätskriterium)
Daß die Bedingung G sternförmig in Satz B1.24 wesentlich ist, sehen wir an Beispiel B1.22. Das Vektorfeld

$$f : \mathbb{R}^2 \setminus \{(0,0)^t\} \longrightarrow \mathbb{R}^2 : (x_1, x_2) \mapsto \left(\frac{-x_2}{x_1^2 + x_2^2}, \frac{x_1}{x_1^2 + x_2^2} \right)^t$$

dort ist stetig differenzierbar und es gilt

$$\frac{\partial f_1}{\partial x_2} = \frac{x_2^2 - x_1^2}{(x_1^2 + x_2^2)^2} = \frac{\partial f_2}{\partial x_1}.$$

Aber, das Vektorfeld ist nicht konservativ, wie wir in Beispiel B1.22 gesehen haben, und ist mithin wegen Satz B1.23 auch kein Gradientenfeld.

Beispiel B1.26 (Berechnen einer Stammfunktion)
Wir betrachten das stetig differenzierbare Vektorfeld

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x_1, x_2)^t \mapsto (2x_1 + 3x_2^2, 6x_1 x_2)^t.$$

Man rechnet leicht nach, daß

$$\frac{\partial f_1}{\partial x_2} = 6x_2 = \frac{\partial f_2}{\partial x_1}$$

gilt. Da $\mathbb{R}^2$ sternförmig ist, ist f also konservativ und besitzt eine Stammfunktion. Um diese zu berechnen, integrieren wir die Komponentenfunktion f_i nach der Variablen x_i:

$$\int^{x_1} f_1(x)\, dx_1 = x_1^2 + 3x_1x_2^2 + c(x_2)$$

und

$$\int^{x_2} f_2(x)\, dx_2 = 3x_1x_2^2 + d(x_1).$$

Dabei sind $c(x_2)$ und $d(x_1)$ Funktionen in der Veränderlichen x_2 bzw. x_1, die beim Differenzieren nach der jeweils anderen Variablen als Konstanten wegfallen. Da eine Funktion F gesucht wird, die beide Gleichungen erfüllt, setzen wir diese gleich und erhalten

$$x_1^2 + 3x_1x_2^2 + c(x_2) = 3x_1x_2^2 + d(x_1).$$

Wir können also $c(x_2) = 0$ und $d(x_1) = x_1^2$ wählen und haben dann eine Stammfunktion

$$F : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x_1, x_2)^t \mapsto x_1^2 + 3x_1x_2^2$$

gefunden.

Definition B1.27 (Rotation eines Vektorfeldes)

a. Wir definieren das *Vektorprodukt* von $x = (x_1, x_2, x_3)^t$ und $y = (y_1, y_2, y_3)^t$ als

$$x \times y := (x_2y_3 - x_3y_2, x_3y_1 - x_1y_3, x_1y_2 - x_2y_1)^t.$$

b. Für ein Vektorfeld $f : U \longrightarrow \mathbb{R}^3$ mit $U \subseteq \mathbb{R}^3$ definieren wir die *Rotation* als

$$\operatorname{rot}(f) := \nabla \times f = \begin{pmatrix} \frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3} \\ \frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1} \\ \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2} \end{pmatrix}.$$

Sie ist damit ein Vektorfeld.

c. Für ein Vektorfeld $f : U \longrightarrow \mathbb{R}^2$ mit $U \subseteq \mathbb{R}^2$ definieren wir die *Rotation* als

$$\operatorname{rot}(f) := \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}.$$

Sie ist damit ein Skalarfeld.

Korollar B1.28

Sei $G \subseteq \mathbb{R}^n$ ein Gebiet mit $n \in \{2, 3\}$ und $f : G \longrightarrow \mathbb{R}^n$ sei stetig differenzierbar.

a. Ist f ein Gradientenfeld, so gilt $\operatorname{rot}(f) = 0$.

b. Ist G sternförmig und gilt $\operatorname{rot}(f) = 0$, so ist f ein Gradientenfeld.

Beweis: Die Bedingung $\operatorname{rot}(f) = 0$ ist Kurzschreibweise der Bedingung (254) in Satz B1.24. □

E) Der Integralsatz von Green

Definition B1.29 ($\mathcal{C}^1$-Normalbereich)
Wir nennen eine Teilmenge $B \subset \mathbb{R}^2$ einen $\mathcal{C}^1$-*Normalbereich*, wenn es stetig differenzierbare Funktionen

$$\varphi, \psi : [a, b] \longrightarrow \mathbb{R}$$

und

$$\Phi, \Psi : [c, d] \longrightarrow \mathbb{R}$$

gibt, so daß

$$\begin{aligned} B =& \{(x, y)^t \mid a \leq x \leq b, \varphi(x) \leq y \leq \psi(x)\} \\ =& \{(x, y)^t \mid c \leq y \leq d, \Phi(y) \leq x \leq \Psi(y)\}. \end{aligned}$$

Bemerkung B1.30 ($\mathcal{C}^1$-Normalbereiche)
Ein $\mathcal{C}^1$-Normalbereich B wie in Definition B1.29 ist sowohl ein Normalbereich

$$B = NB(a, b, \varphi, \psi)$$

bezüglich (x, y), als auch ein Normalbereich

$$B = NB(c, d, \Phi, \Psi)$$

bezüglich (y, x) im Sinne von Definition 29.35. Der Rand ∂B von B ist die Spur einer stückweisen $\mathcal{C}^1$-Kurve $\gamma = \delta + \eta - \mu - \nu$ mit

$$\begin{aligned} \delta &: [a, b] \longrightarrow \mathbb{R}^2 : t \mapsto (t, \varphi(t))^t, \\ \eta &: [\varphi(b), \psi(b)] \longrightarrow \mathbb{R}^2 : t \mapsto (b, t)^t, \\ \mu &: [a, b] \longrightarrow \mathbb{R}^2 : t \mapsto (t, \psi(t))^t, \\ \nu &: [\varphi(a), \psi(a)] \longrightarrow \mathbb{R}^2 : t \mapsto (a, t)^t. \end{aligned}$$

Wir nennen diese Parametrisierung der Randkurve *positiv orientiert*, da der Normalenvektor $(-\gamma_2'(t), \gamma_1'(t))^t$ stets in den Normalbereich hinein zeigt. Analog kann man eine positiv orientierte Parametrisierung des Randes mit Hilfe von Φ und Ψ definieren, und bei geeigneter Wahl von Anfangs- und Endpunkt sind die beiden Parametrisierungen äquivalent mit derselben Orientierung.

Satz B1.31 (Integralsatz von Green)
Sei $U \subseteq \mathbb{R}^2$ offen, $f : U \longrightarrow \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld und $B \subset U$ ein $\mathcal{C}^1$-Normalbereich mit positiv orientierter Randkurve γ wie in Bemerkung B1.30. Dann gilt

$$\int_\gamma f \cdot \mathbf{dx} = \int_B D_x f_2(x,y) - D_y f_1(x,y)\, d(x,y).$$

Bemerkung B1.32 (Integralsatz von Green)
In der Literatur werden die Komponentenfunktionen des Vektorfelds f oft P und Q genannt, und daß der Rand ∂B in positiver Orientierung durchlaufen werden soll, wird durch ein spezielles Integralsymbol ausgedrückt, so daß die Formel im Integralsatz dann die folgende Form bekommt:

$$\oint_{\partial B} P\, dx + Q\, dy = \int_B D_x Q(x,y) - D_y P(x,y)\, d(x,y).$$

Man kann die Voraussetzungen an den Integrationsbereich B deutlich abschwächen. Es reicht, daß der Rand von B die Spur einer injektiven, stückweise stetig differenzierbaren Kurve γ mit positiver Orientierung (d.h. der Normalenvektor $(-\gamma_2'(t), \gamma_1'(t))$ zeigt ins Innere von B) ist. Solche Kurven nennt man auch *Jordankurven*; sie sind als endliche Vereinigung stetig differenzierbarer Bilder kompakter Nullmengen wieder Nullmengen im $\mathbb{R}^2$, so daß B dann Jordan-meßbar ist. Das ist z.B. dann sicher der Fall, wenn B sich in endliche viele $\mathcal{C}^1$-Normalbereiche zerlegen läßt.

Diese Verallgemeinerung des Integralsatzes von Green ist wichtig, weil schon einfache Normalbereiche wie der Einheitskreis keine $\mathcal{C}^1$-Normalbereiche in unserem Sinne sind. In den Randpunkten sind die Tangentialvektoren der Kurven senkrecht und die Funktionen φ und ψ sind dort nicht mehr differenzierbar.

Beweis des Satzes von Green B1.31: Für den Beweis verwenden wir die Notation $f = (P, Q)^t$ und nehmen ohne Einschränkung an, daß die Funktionen in der Definition des Normalbereichs auf ihren Intervallen stetig differenzierbar sind.

Wir zeigen die Aussage zunächst für den Fall $Q = 0$, indem wir den Hauptsatz der Differential- und Integralrechnung 20.5 und den Satz von Fubini 29.37 für Normalbereiche anwenden:

$$\begin{aligned}
\int_\gamma (P,0)^t \cdot \mathbf{dx} &= \int_\delta (P,0)^t \cdot \mathbf{dx} + \int_\eta (P,0)^t \cdot \mathbf{dx} - \int_\mu (P,0)^t \cdot \mathbf{dx} - \int_\nu (P,0)^t \cdot \mathbf{dx} \\
&= \int_a^b P(\delta(t)) \cdot \delta_1'(t)\ dt + \int_{\varphi(b)}^{\psi(b)} P(\eta(t)) \cdot \eta_1'(t)\ dt \\
&\quad - \int_a^b P(\mu(t)) \cdot \mu_1'(t)\ dt - \int_{\varphi(a)}^{\psi(a)} P(\nu(t)) \cdot \nu_1'(t)\ dt \\
&= \int_a^b P(t,\varphi(t)) \cdot 1\ dt + \int_{\varphi(b)}^{\psi(b)} P(b,t) \cdot 0\ dt \\
&\quad - \int_a^b P(t,\psi(t)) \cdot 1\ dt + \int_{\varphi(a)}^{\psi(a)} P(a,t) \cdot 0\ dt \\
&= -\int_a^b P(t,\psi(t)) - P(t,\varphi(t))\ dt \\
&\overset{HDIR}{=} -\int_a^b \int_{\varphi(a)}^{\psi(a)} D_y P(t,y)\ dy\ dt \\
&\overset{Fubini}{=} -\int_B D_y P(x,y)\ d(x,y).
\end{aligned}$$

Analog zeigt man im Fall $P = 0$

$$\int_\gamma (0,Q)^t \cdot \mathbf{dx} = \int_B D_x Q(x,y)\ d(x,y).$$

Den allgemeinen Fall erhalten wir dann aus der Additivität des Integrals:

$$\int_\gamma f \cdot \mathbf{dx} = \int_\gamma (P,0)^t \cdot \mathbf{dx} + \int_\gamma (0,Q)^t \cdot \mathbf{dx} = \int_B D_x Q(x,y) - D_y P(x,y)\ d(x,y).$$

□

Beispiel B1.33

Mit Hilfe des Satzes von Green und der Bemerkung B1.32 kann man das Kurvenintegral des Vektorfeldes

$$f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x,y)^t \mapsto (y - cos(e^x), 2x + \sqrt{\cos(y)^2 + 1})^t$$

längs des Einheitskreises leicht berechnen. Es gilt nämlich

$$\oint_{\partial U_1(0)} f \cdot \mathbf{dx} = \int_{\overline{U_1(0)}} 2 - 1\ d(x,y) = \pi.$$

Definition B1.34 (Divergenz)
Ist $f : U \longrightarrow \mathbb{R}^n$ mit $U \subseteq \mathbb{R}^n$ offen ein stetig differenzierbares Vektorfeld, so heißt das Skalarfeld

$$\operatorname{div}(f) : U \longrightarrow \mathbb{R} : x \mapsto \sum_{i=1}^{n} \frac{\partial f_i}{\partial x_i}(x)$$

die *Divergenz* von f. Ist $\operatorname{div}(f) = 0$, so nennt man f *quellenfrei*; Punkte $x \in U$ mit $\operatorname{div}(f)(x) > 0$ nennt man *Quellen*, Punkte $x \in U$ mit $\operatorname{div}(f)(x) < 0$ nennt man *Senken*.

Korollar B1.35 (Integralsatz von Green in Divergenzform)
Sei $U \subseteq \mathbb{R}^2$ offen, $f : U \longrightarrow \mathbb{R}^2$ ein stetig differenzierbares Vektorfeld und $B \subset U$ ein $\mathcal{C}^1$-Normalbereich mit positiv orientierter Randkurve γ wie in Bemerkung B1.30. Dann gilt

$$\oint_{\partial B} \langle f, n\rangle \, ds := \int_{\gamma} \langle f, n\rangle \, ds = \int_B \operatorname{div}(f)(x,y) \, d(x,y),$$

wenn $n(\gamma(t)) = \frac{(\gamma_2'(t), -\gamma_1'(t))^t}{||\gamma'(t)||_2}$ der äußere Normaleneinheitsvektor auf ∂B in $\gamma(t)$ ist.

Beweis: Wir nehmen ohne Einschränkung an, daß $\gamma : [\alpha, \beta] \longrightarrow \mathbb{R}^2$ stetig differenzierbar ist. Dann gilt

$$\begin{aligned}
\int_{\gamma} \langle f, n\rangle \, ds &= \int_{\alpha}^{\beta} \left\langle f(\gamma(t)), \frac{(\gamma_2'(t), -\gamma_1'(t))^t}{||\gamma'(t)||_2} \right\rangle \cdot ||\gamma'(t)||_2 \, dt \\
&= \int_{\alpha}^{\beta} f_1(\gamma(t)) \cdot \gamma_2'(t) - f_2(\gamma(t)) \cdot \gamma_1'(t) \, dt \\
&= \int_{\gamma} (-f_2, f_1)^t \cdot \mathbf{dx} \overset{B1.31}{=} \int_B \operatorname{div}(f)(x,y) \, d(x,y).
\end{aligned}$$

□

Bemerkung B1.36 (Integralsatz von Green in Divergenzform)
Auch in Korollar B1.35 kann man die Voraussetzungen an den Integrationsbereich B wieder abschwächen, siehe Bemerkung B1.32. Man beachte zudem, daß der Normaleneinheitsvektor n nur außerhalb der endlich vielen Stellen definiert ist, an denen γ nicht differenzierbar ist.

Bemerkung B1.37 (Quellenfreie Vektorfelder)
Beschreibe $v : G \longrightarrow \mathbb{R}^2$ für ein Gebiet G die Geschwindigkeit eines ebenen Flüssigkeitsfeldes und sei die Dichte ρ der Flüssigkeit konstant. Dann beschreibt die Flußdichte

$$f := \rho \cdot v : G \longrightarrow \mathbb{R}^2,$$

wie viel Masse sich pro Länge und Zeiteinheit an der Stelle x in Richtung $v(x)$ verschiebt. Das Kurvenintegral

$$\oint_{\partial B} \langle f, n\rangle \, ds = \int_B \operatorname{div}(f)(x,y) \, d(x,y)$$

gibt dann die Differenz aus ausfließender und einfließender Masse in $B \subset G$ pro Zeiteinheit an. Wenn das Flüssigkeitsfeld in B keine Quellen und Senken hat, dann sollte das Kurvenintegral 0 sein. Wegen der Gleichheit der Integrale folgt dies sicher aus $\operatorname{div}(f) = 0$, d.h. wenn f quellenfrei ist.

Aufgaben

Aufgabe B1.38 (Parameterintegrale)
Sei $U \subseteq \mathbb{R}^2$ offen mit $[a,b] \times [c,d] \subset U$ und $f : U \longrightarrow \mathbb{R}$ stetig.

a. Zeige, die Funktion

$$h : [a,b] \times [c,d] \longrightarrow \mathbb{R} : (x,y)^t \mapsto \int_a^x f(t,y) \, dt$$

ist stetig.

b. Zeige, existiert zudem $D_y f : G \longrightarrow \mathbb{R}$ und ist stetig, so ist die Funktion

$$g : [c,d] \longrightarrow \mathbb{R} : y \mapsto \int_a^b f(t,y) \, dt$$

stetig differenzierbar mit

$$g'(y) = D_y g(y) = \int_a^b (D_y f)(t,y) \, dt.$$

Aufgabe B1.39
Zeige, ist B ein $\mathcal{C}^1$-Normalbereich mit positiv orientierter Parametrisierung γ des Randes und ist $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x,y)^t \mapsto \frac{(-y,x)^t}{2}$, dann ist $\int_\gamma f \cdot \mathbf{dx}$ der Flächeninhalt von B.

Aufgabe B1.40 (Greensche Formeln)
Seien $F, G : U \longrightarrow \mathbb{R}$ zweifach stetig differenzierbare Skalarfelder mit $U \subseteq \mathbb{R}^2$ offen und sei $B \subseteq U$ ein $\mathcal{C}^1$-Normalbereich.

a. Zeige, $\int_B \Delta(F)(x,y) \, d(x,y) = \oint_{\partial B} \frac{\partial F}{\partial n} \, ds$, wobei $\frac{\partial F}{\partial n}$ die Richtungsableitung von F in Richtung des äußeren Normaleneinheitsvektors ist.

b. Zeige, $\int_B (F \cdot \Delta(G) + \langle \nabla F, \nabla G\rangle)(x,y) \, d(x,y) = \oint_{\partial B} F \cdot \frac{\partial G}{\partial n} \, ds$.

c. Zeige, $\int_B (F \cdot \Delta(G) - G \cdot \Delta(F))(x,y) \, d(x,y) = \oint_{\partial B} F \cdot \frac{\partial G}{\partial n} - G \cdot \frac{\partial F}{\partial n} \, ds$.

Aufgabe B1.41 (Das Kreuzprodukt von Vektoren)

Seien $x, y, z \in \mathbb{R}^3$, $\lambda \in \mathbb{R}$ und $T \in \mathrm{O}(3)$ eine orthogonale 3×3-Matrix. Zeige die folgenden Rechenregeln für das Kreuzprodukt:

a. $x \times y = -y \times x$.

b. $(x + y) \times z = (x \times z) + (y \times z)$.

c. $(\lambda \cdot x) \times y = x \times (\lambda y) = \lambda \cdot (x \times y)$.

d. $\langle x \times y, x \rangle = \langle x \times y, y \rangle = 0$, d.h. $x \times y \perp x$ und $x \times y \perp y$.

e. $||x \times y||_2^2 = ||x||_2^2 \cdot ||y||_2^2 - \langle x, y \rangle^2$.

f. x und y sind genau dann linear abhängig, wenn $x \times y = 0$ gilt.

g. $||Tx \times Ty||_2 = ||x \times y||_2$.

h. Wenn $x_3 = y_3 = 0$, dann gilt

$$||x \times y||_2 = \left| \det \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \end{pmatrix} \right|.$$

i. $||x \times y||_2$ ist der Flächeninhalt des von x und y aufgespannten Parallelogramms.

Aufgabe B1.42 (Rotation)

Seien $f, g : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ und $h : \mathbb{R}^3 \longrightarrow \mathbb{R}$.

a. Zeige, ist h zweifach stetig differenzierbar, dann gilt $\mathrm{rot}(\mathrm{grad}(h)) = 0$.

b. Zeige, ist f zweifach stetig differenzierbar, dann gelten

$$\mathrm{div}(\mathrm{rot}(f)) = 0$$

und

$$\mathrm{rot}(\mathrm{rot}(f)) = \mathrm{grad}(\mathrm{div}(f)) - \Delta f.$$

c. Zeige, sind f, g und h stetig differenzierbar, dann gelten

$$\mathrm{rot}(f + g) = \mathrm{rot}(f) + \mathrm{rot}(g)$$

und

$$\mathrm{rot}(h \cdot f) = h \cdot \mathrm{rot}(f) + (\mathrm{grad}(f) \times f)$$

sowie

$$\mathrm{div}(f \times g) = \langle g, \mathrm{rot}(f) \rangle - \langle f, \mathrm{rot}(g) \rangle.$$

Aufgabe B1.43 a. Sei γ die Helix $\gamma : [0, 4\pi] \longrightarrow \mathbb{R}^3 : t \mapsto (\cos t, \sin t, t)^t$, und sei F das Vektorfeld $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (x, y, z)^t \mapsto (yz, xy, 0)^t$. Berechne $\int_\gamma F \cdot \boldsymbol{dx}$.

b. Seien $a, b > 0$ und sei γ der Polygonzug durch die Punkte $(0, 0)^t$, $(a, 0)^t$, $(0, b)^t$ und $(0, 0)^t$. Sei außerdem f das Skalarfeld $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y)^t \mapsto xe^{-y}$. Berechne $\int_\gamma f \boldsymbol{ds}$.

Aufgabe B1.44 a. Sei $\gamma : [1, 2] \longrightarrow \mathbb{R}^2 : t \mapsto (1, \frac{e^t+e^{-t}}{2})^t$ und $f : \mathbb{R}^2 \longrightarrow \mathbb{R} : (x, y) \mapsto y$. Berechne $\int_\gamma f ds$.

b. Sei $\gamma : [1, 2] \longrightarrow \mathbb{R}^3 : t \mapsto (1, t, \ln(t))^t$ und $f : \mathbb{R}^3 \longrightarrow \mathbb{R} : (x, y, z) \mapsto y^{-3}$. Berechne $\int_\gamma f ds$.

Aufgabe B1.45
Zeige, ist B ein $\mathcal{C}^1$-Normalbereich mit positiv orientierter Parametrisierung γ des Randes und ist $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2 : (x, y)^t \mapsto \frac{(-y,x)^t}{2}$, dann ist $\int_\gamma f \cdot \boldsymbol{dx}$ der Flächeninhalt von B.

B2 Die Integralsätze von Gauß und Stokes

Im letzten Abschnitt haben wir Integrale von Funktionen über zweidimensionale Bereiche auf Integrale über deren Randkurven zurückgeführt. Analog dazu wollen wir in diesem Abschnitt Integrale über Körper im $\mathbb{R}^3$ auf Integrale über deren Ober- oder Randflächen zurückführen. Dazu benötigen wir zunächst eine gute Beschreibung zulässiger Oberflächen und müssen dann analog zum Kurvenintegral den Begriff des Oberflächenintegrals einführen.

A) Parametrisierte Flächen

Definition B2.1 (Glatte Flächen)
Sei $D \subset \mathbb{R}^2$ kompakt und Jordan-meßbar und der Rand ∂D von D sei die Spur einer stückweise stetig differenzierbaren Kurve; ferner sei $U \subseteq \mathbb{R}^2$ offen mit $D \subset U$; und schließlich sei $\phi : U \longrightarrow \mathbb{R}^3$ stetig differenzierbar, so daß $\phi_{|\mathring{D}}$ injektiv ist und $\operatorname{rang}(D\phi(u,v)) = 2$ für fast alle $u, v \in D$.
Dann nennen wir ϕ *Parameterdarstellung* der *glatten Fläche* $\mathcal{F} = \phi(D)$ mit *Parameterbereich* D. Die Vektoren $\phi_u(u,v)$ und $\phi_v(u,v)$ heißen *Tangentialvektoren* an die Fläche $\mathcal{F}$ im Punkt $\phi(u,v)$, und der Vektor $(\phi_u \times \phi_v)(u,v)$ heißt *Normalenvektor* der Fläche $\mathcal{F}$ in $\phi(u,v)$.
Ist eine Fläche im $\mathbb{R}^3$ Vereinigung mehrerer glatter Flächen im $\mathbb{R}^3$, die sich nur in ihrem Rand schneiden, so nennen wir die Fläche *stückweise stetig differenzierbar* oder *stückweise glatt*.

Bemerkung B2.2 (Glatte Flächen)
Betrachten wir einen Punkt $\phi(u,v) \in \mathcal{F}$, halten den zweiten Parameter fest und variieren den ersten leicht, so erhalten wir eine Kurve

$$\gamma : [-\varepsilon, \varepsilon] \longrightarrow \mathbb{R}^3 : t \mapsto \phi(u+t, v)$$

in der Fläche $\mathcal{F}$ und

$$\gamma'(0) = \phi_u(u,v)$$

ist der Tangentialvektor an diese Kurve, so daß $\phi_u(u,v)$ in der Tat tangential an die Fläche $\mathcal{F}$ anliegt. Analog sieht man dies für $\phi_v(u,v)$.

Die Rangbedingung an $D\phi(u,v)$ liefert, daß die beiden Tangentialvektoren $\phi_u(u,v)$ und $\phi_v(u,v)$ linear unabhängig sind. Damit ist der Tangentialraum an $\mathcal{F}$ in $\phi(u,v)$ zweidimensional und wird von den beiden Vektoren aufgespannt und $\mathcal{F}$ ist lokal in $\phi(u,v)$ in der Tat eine Fläche. Zudem ist wegen Aufgabe B1.41 das Kreuzprodukt $(\phi_u \times \phi_v)(u,v)$ damit nicht der Nullvektor und steht senkrecht auf dem Tangentialraum

an $\mathcal{F}$ in $\phi(u,v)$, ist also der Normalenvektor zum Tangentialraum und damit zur Fläche in $\phi(u,v)$.

Man beachte übrigens, daß die Rangbedingung an $D\phi(u,v)$ wegen Aufgabe B1.41 äquivalent dazu ist, daß der Normalenvektor in $\phi(u,v)$ nicht der Nullvektor ist. Dies ist also in fast allen Punkten von $\mathcal{F}$ sicher gestellt, d.h. außerhalb einer Jordan-Nullmenge, die bei der Integration keine Rolle spielt.

Beispiel B2.3

a. Ist $\varphi : \mathbb{R}^2 \longrightarrow \mathbb{R}$ stetig differenzierbar und ist D wie in Definition B2.1, dann ist

$$\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}^3 : (u,v)^t \mapsto (u,v,\varphi(u,v))^t$$

Parameterdarstellung einer glatten Fläche, nämlich des Graphen von ϕ über D, da der Normalenvektor in $\phi(u,v)$

$$(\phi_u \times \phi_v)(u,v) = \begin{pmatrix} 1 \\ 0 \\ \varphi_u(u,v) \end{pmatrix} \times \begin{pmatrix} 0 \\ 1 \\ \varphi_v(u,v) \end{pmatrix} = \begin{pmatrix} -\varphi_u(u,v) \\ -\varphi_v(u,v) \\ 1 \end{pmatrix}$$

nicht der Nullvektor ist. Die Abbildung ϕ ist dabei auf ganz $\mathbb{R}^2$ injektiv.

b. Der Kreiszylinder mit Radius r und Höhe h ist eine glatte Fläche mit der Parameterdarstellung

$$\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}^3 : (\theta,v)^t \mapsto (r\cos(\theta), r\sin(\theta), v)^t$$

und dem Normalenvektor

$$(\phi_\theta \times \phi_v)(\theta,v) = \begin{pmatrix} -r\sin(\theta) \\ r\cos(\theta) \\ 0 \end{pmatrix} \times \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \\ 0 \end{pmatrix} \neq 0$$

und dem Parameterbereich

$$D = [0, 2\pi] \times [0, h].$$

Man beachte, daß ϕ im Inneren von D injektiv ist.

c. Die Oberfläche der Kugel mit Radius r ist eine glatte Fläche mit der Parameterdarstellung

$$\phi : \mathbb{R}^2 \longrightarrow \mathbb{R}^3 : (\theta,\vartheta)^t \mapsto \big(r\cdot\cos(\theta)\cdot\cos(\vartheta), r\cdot\sin(\theta)\cdot\cos(\vartheta), r\cdot\sin(\vartheta)\big)^t$$

mit Parameterbereich

$$D = [0, 2\pi] \times \left[-\tfrac{\pi}{2}, \tfrac{\pi}{2}\right]$$

und Normalenvektor

$$(\phi_\theta \times \phi_\vartheta)(\theta,\vartheta) = r\cdot\cos(\vartheta)\cdot\phi(\theta,\vartheta).$$

Man beachte, daß der Normalenvektor im Inneren von D ungleich null ist und daß φ dort auch injektiv ist.

B) Oberflächenintegrale

Definition B2.4 (Oberflächenintegrale und Oberflächeninhalt)
Sei ϕ Parameterdarstellung einer glatten Fläche $\mathcal{F}$ mit Parameterbereich D.

a. Ist $f : \mathcal{F} \longrightarrow \mathbb{R}$ ein stetiges Skalarfeld, dann heißt

$$\int_\phi f \, dS := \int_D f(\phi(u,v)) \cdot ||(\phi_u \times \phi_v)(u,v)||_2 \, d(u,v)$$

das *(skalare) Oberflächenintegral* von f über die glatte Fläche $\mathcal{F}$.
Im Spezialfall $f = 1$ nennen wir das Integral

$$I(\mathcal{F}) := \int_\phi 1 \, dS = \int_D ||(\phi_u \times \phi_v)(u,v)||_2 \, d(u,v)$$

auch den *(Oberflächen-)Inhalt* von $\mathcal{F}$.

b. Ist $f : \mathcal{F} \longrightarrow \mathbb{R}^3$ ein stetiges Vektorfeld, dann heißt

$$\int_\phi f \cdot \mathbf{dS} := \int_D \langle f(\phi(u,v)), (\phi_u \times \phi_v)(u,v) \rangle \, d(u,v)$$

das *(vektorielle) Oberflächenintegral* von f über die glatte Fläche $\mathcal{F}$.

c. Ist eine Fläche im $\mathbb{R}^3$ stückweise stetig differenzierbar, d.h. Vereinigung mehrerer glatter Flächen im $\mathbb{R}^3$, die sich nur in ihrem Rand schneiden, so definieren wir das Oberflächenintegral über die Gesamtfläche als die Summe der einzelnen Oberflächenintegrale.

Bemerkung B2.5 (Oberflächenintegrale)

a. Es ist in der Literatur üblich statt $\int_\phi f \, dS$ und $\int_\phi f \cdot \mathbf{dS}$ einfach $\int_\mathcal{F} f \, dS$ und $\int_\mathcal{F} f \cdot \mathbf{dS}$ zu schreiben. Wir sollten aber im Blick behalten, daß die Definition zunächst nicht nur von der Fläche $\mathcal{F}$, sondern von deren konkreter Parametrisierung abhängt, auch wenn wir zeigen werden, daß hinreichend gute Parameterwechsel den Wert des Integrals unangetastet lassen (siehe Proposition B2.7).

b. Ist $D = [0,1] \times [0,1]$ und ist

$$\phi : D \longrightarrow \mathbb{R}^3 : (u,v) \mapsto u \cdot a + v \cdot b$$

für zwei linear unabhängige Vektoren $a, b \in \mathbb{R}^3$, so ist $\mathcal{F} = \phi(D)$ ein Vektorparallelogramm und aus Aufgabe B1.41 wissen wir, daß

$$I(\mathcal{F}) = \int_D ||(\phi_u \times \phi_v)(u,v)||_2 \, ds = \int_D ||a \times b||_2 \, ds = ||a \times b||_2$$

in der Tat der Flächeninhalt der Fläche ist. Das Kreuzprodukt hat hier die Änderung des Flächeninhalts $V(D) = 1$ von D unter φ gemessen. In Verallgemeinerung von Bemerkung B1.3 ist es daher naheliegend, $||\phi_u \times \phi_v||_2$ zu verwenden, um die infinitesimale Flächenänderung zu messen und über diese zu integrieren, um die Gesamtänderung und damit den Flächeninhalt zu erhalten.

c. Wie bei Kurvenintegralen gibt es auch bei Oberflächenintegralen einen einfachen Zusammenhang zwischen dem skalaren und dem vektoriellen Oberflächenintegral. Führt man den Normaleneinheitsvektor

$$n(\phi(u,v)) = \frac{(\phi_u \times \phi_v)(u,v)}{||(\phi_u \times \phi_v)(u,v)||_2}$$

ein, dann gilt

$$\begin{aligned} \int_\phi f \cdot \mathbf{dS} &= \int_D \langle f(\phi(u,v)), (\phi_u \times \phi_v)(u,v)\rangle \, d(u,v) \\ &= \int_D \langle f(\phi(u,v)), n(\phi(u,v))\rangle \cdot ||\phi_u \times \phi_u)(u,v)||_2 \, d(u,v) \\ &= \int_\phi \langle f, n\rangle \, dS. \end{aligned} \tag{257}$$

d. Man beachte, daß die Definition in Teil c. nur dann sinnvoll ist, wenn man dabei eine konsistente Wahl der Normalenvektoren trifft (siehe etwa Bemerkung B2.10), da das Vorzeichen der einzelnen Summanden von dieser Wahl abhängt.

Beispiel B2.6

Mit Hilfe von Beispiel B2.3 können wir den Oberflächeninhalt der Kugeloberfläche berechnen als

$$\int_D ||(\phi_\theta \times \phi_\vartheta)(\theta,\vartheta)||_2 \, d(\theta,\vartheta) = \int_0^{2\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} r^2 \cdot \cos(\vartheta) \, d\vartheta d\theta = 4\pi r^2.$$

Proposition B2.7 (Parameterwechsel)

Seien $\phi : U \longrightarrow \mathbb{R}^3$ und $\psi : V \longrightarrow \mathbb{R}^3$ zwei Parameterdarstellungen der glatten Fläche $\mathcal{F}$ mit Parameterbereichen D und C in $\mathbb{R}^2$, und sei

$$\varphi : V \longrightarrow U$$

eine Abbildung die die Voraussetzungen des Transformationssatzes für Integrale in der Fassung von Bemerkung 30.8 oder von Korollar 30.10 mit fast überall positiver Funktionaldeterminante $\det(D\varphi(s,t))$ erfüllt und für die $\psi = \phi \circ \varphi$ gilt.

a. Für $(s,t) \in V$ gilt dann

$$(\psi_s \times \psi_t)(s,t) = (\phi_u \times \phi_v)(\varphi(s,t)) \cdot \det(D\varphi(s,t)).$$

b. Wenn $f : \mathcal{F} \longrightarrow \mathbb{R}$ ein stetiges Skalarfeld ist, dann ist

$$\int_\phi f \, dS = \int_\psi f \, dS.$$

c. Wenn $f : \mathcal{F} \longrightarrow \mathbb{R}^3$ stetiges Vektorfeld ist, dann ist

$$\int_\phi f \cdot \mathbf{dS} = \int_\psi f \cdot \mathbf{dS}.$$

Beweis: a. Aus der Kettenregel erhalten wir $D\psi = D\phi \circ D\varphi$ und somit

$$\psi_s = D_s\psi = D\phi \circ D_s\varphi = \phi_u \cdot \varphi_{1s} + \phi_v \cdot \varphi_{2s}$$

und

$$\psi_t = D_t\psi = D\phi \circ D_t\varphi = \phi_u \cdot \varphi_{1t} + \phi_v \cdot \varphi_{2t}.$$

Für das Kreuzprodukt erhalten wir damit

$$\begin{aligned}\psi_s \times \psi_t =& (\phi_u \cdot \varphi_{1s} + \phi_v \cdot \varphi_{2s}) \times (\phi_u \cdot \varphi_{1t} + \phi_v \cdot \varphi_{2t}) \\ =& (\phi_u \times \phi_v) \cdot (\varphi_{1s}\varphi_{2t} - \varphi_{2s}\varphi_{1t}) = (\phi_u \times \phi_v) \cdot \det(D\varphi),\end{aligned}$$

wenn wir beachten, daß das Kreuzprodukt linear abhängiger Vektoren stets Null ist, daß das Kreuzprodukt in seinen Komponenten linear ist und daß $\phi_v \times \phi_u = -\phi_u \times \phi_v$ gilt (siehe Aufgabe B1.41).

b. Die Gleichheit der Oberflächenintegrale folgt dann aus Teil a. und dem Transformationssatz für Integrale:

$$\begin{aligned}\int_\phi f \, dS &= \int_D f(\phi(u,v)) \cdot ||(\phi_u \times \phi_v)(u,v)||_2 \, d(u,v) \\ &\overset{TSI}{=} \int_C f(\phi(\varphi(s,t))) \cdot ||(\phi_u \times \phi_v)(\varphi(s,t))||_2 \cdot |\det(D\varphi(s,t))| \, d(s,t) \\ &\overset{a.}{=} \int_C f(\psi(s,t)) \cdot ||(\psi_s \times \psi_t)(s,t)||_2 \, d(s,t) = \int_\psi f \, dS.\end{aligned}$$

c. Dies folgt aus b. unter Beachtung von (257) in Bemerkung B2.5

□

Bemerkung B2.8 (Parameterwechsel)

Man beachte zunächst, daß die Abbildung ϕ auf dem Inneren von D injektiv ist und daß auf dem Inneren von C deshalb notwendigerweise

$$\varphi = \phi^{-1} \circ \psi$$

gilt, und daß diese Abbildung aufgrund der Rangbedingung an die Parameterdarstellungen fast überall eine nicht-verschwindende Funktionaldeterminante hat. Daß wir die Transformationsabbildung φ über das Innere von C hinaus auf einer offenen Umgebung benötigen, ist nur unserer Fassung des Transformationssatzes für Riemann-Integrale

geschuldet. Man kann die Voraussetzungen hier mit etwas Arbeit deutlich abschwächen. Die Verwendung des Lebesgue-Integrals anstelle des Riemann-Integrals macht die Arbeit dabei erheblich leichter. Wir wollen Proposition B2.7 deshalb als Indikator dafür nehmen, daß bei einem hinreichend guten, orientierungserhaltenden Parameterwechsel der Wert des Oberflächenintegrals sich nicht ändert.

C) Der Integralsatz von Gauß

Wir wollen nun den Begriff des $\mathcal{C}^1$-Normalbereichs auf den $\mathbb{R}^3$ verallgemeinern.

Definition B2.9 ($\mathcal{C}^1$-Normalbereich)
Eine Teilmenge $B \subset \mathbb{R}^3$, die bezüglich jeder Anordnung der Koordinaten ein Normalbereich im Sinne Definition 29.35 mit stetig differenzierbaren Randfunktionen ist, heißt ein *$\mathcal{C}^1$-Normalbereich* im $\mathbb{R}^3$.

Bemerkung B2.10 ($\mathcal{C}^1$-Normalbereich)
Wählen wir in der Normalbereichdarstellung die Variable z als letzte Variable, so erhalten wir eine Darstellung des *$\mathcal{C}^1$-Normalbereichs* B der Form

$$B = \{(x,y,z)^t \mid (x,y)^t \in D, \varphi(x,y) \leq z \leq \psi(x,y)\}$$

für einen $\mathcal{C}^1$-Normalbereich D in $\mathbb{R}^2$ und zwei stetig differenzierbare Funktionen

$$\varphi, \psi : D \longrightarrow \mathbb{R}$$

mit $\varphi(x,y) \leq \psi(x,y)$ für alle $(x,y)^t \in D$. Die Oberfläche $\mathcal{F}$ von B zerlegt sich dann in natürlicher Weise in die beiden Graphen der Funktionen φ

$$\phi^- : D \longrightarrow \mathrm{Graph}(\varphi) : (u,v)^t \mapsto (u,v,\varphi(u,v))^t$$

und ψ

$$\phi^+ : D \longrightarrow \mathrm{Graph}(\psi) : (u,v)^t \mapsto (u,v,\psi(u,v))^t$$

über D sowie den Zylindermantel Z über der Randkurve von D

$$\phi^Z : [a,b]\times[0,1] \longrightarrow Z : (u,v) \mapsto \big(x(u), y(u), v\cdot\varphi(x(u),y(u)) + (1-v)\cdot\psi(x(u),v(u))\big)^t,$$

wobei

$$[a,b] \longrightarrow \mathbb{R}^2 : u \mapsto (x(u), v(u))^t$$

eine Parameterdarstellung der Randkurve von D ist. Aus den angegebenen Parametrisierungen können wir den äußeren Normaleneinheitsvektor auf die Oberfläche B in den einzelnen Teilstücken definieren. Auf dem Graphen von φ erhalten wir

$$-\phi_u^- \times \phi_v^- = -(1,0,\varphi_u)^t \times (0,1,\varphi_v)^t = (\varphi_u, \varphi_v, -1)^t$$

und damit

$$n = \frac{\phi_u^- \times \phi_v^-}{||\phi_u^- \times \phi_v^-||_2} = \frac{(\varphi_u, \varphi_v, -1)^t}{\sqrt{1 + \varphi_u^2 + \varphi_v^2}}.$$

Auf dem Graphen von ψ erhalten wir

$$\phi_u^+ \times \phi_v^+ = (1, 0, \psi_u)^t \times (0, 1, \psi_v)^t = (-\psi_u, -\psi_v, 1)^t$$

und damit

$$n = \frac{\phi_u^- \times \phi_v^-}{||\phi_u^- \times \phi_v^-||_2} = \frac{(-\psi_u, -\psi_v, 1)^t}{\sqrt{1 + \psi_u^2 + \psi_v^2}}.$$

Auf dem Zylindermantel Z erhalten wir

$$\phi_u^Z \times \phi_v^Z = (0, 0, \varphi - \psi)^t \times (x', y', v \cdot x' + (1 - v) \cdot y')^t = (y' \cdot (\psi - \varphi), x' \cdot (\varphi - \psi), 0)^t$$

und damit

$$n = \frac{\phi_u^- \times \phi_v^-}{||\phi_u^- \times \phi_v^-||_2} = \frac{(y', -x', 0)^t}{\sqrt{x'^2 + y'^2}}.$$

In analoger Weise können wir vorgehen, wenn wir die Variablen x und y als letzte Variablen in der Normalbereichdarstellung wählen.

Satz B2.11 (Integralsatz von Gauß in Divergenzform)
Sei $B \subset \mathbb{R}^3$ ein $\mathcal{C}^1$-Normalbereich mit Oberfläche $\mathcal{F}$ und äußerem Normaleneinheitsvektor n (siehe Bemerkung B2.10) und sei $f : U \longrightarrow \mathbb{R}^3$ ein stetig differenzierbares Vektorfeld mit $U \subseteq \mathbb{R}^3$ offen und $B \subset U$. Dann gilt

$$\int_B \operatorname{div}(f)(x, y, z)\, d(x, y, z) = \int_{\mathcal{F}} \langle f, n \rangle\, dS = \int_{\mathcal{F}} f \cdot \mathbf{dS}.$$

Beweis: Wir gehen wie im Beweis des Satzes von Green B1.31 vor und betrachten B zunächst als $\mathcal{C}^1$-Normalbereich mit z als letzter Koordinate wie in Bemerkung B2.10. Mit der Notation aus Bemerkung B2.10, dem Hauptsatz der Differenzial- und Integralrechnung 20.5 sowie dem Satz von Fubini 29.37 für Normalbereiche erhalten wir

damit:

$$\begin{aligned}
\int_{\mathcal{F}} (0,0,f_3)^t \cdot \mathbf{dS} &= -\int_{\phi^-} (0,0,f_3)^t \cdot \mathbf{dS} + \int_{\phi^+} (0,0,f_3)^t \cdot \mathbf{dS} + \int_{\phi^Z} (0,0,f_3)^t \cdot \mathbf{dS} \\
&= -\int_D f_3(u,v,\varphi(u,v)) \cdot 1 \; d(u,v) + \int_D f_3(u,v,\psi(u,v)) \cdot 1 \; d(u,v) \\
&\quad + \int_{[a,b]\times[0,1]} f_3(\phi^Z(u,v)) \cdot 0 \; d(u,v) \\
&= \int_D f_3(u,v,\psi(u,v)) - f_3(u,v,\varphi(u,v)) \; d(u,v) \\
&\overset{HDIR}{=} \int_D \int_{\varphi(u,v)}^{\psi(u,v)} \frac{\partial f_3}{\partial z}(u,v,z) \; dz \; d(u,v) \\
&\overset{Fubini}{=} \int_B \frac{\partial f_3}{\partial z}(u,v,z) \; d(u,v,z).
\end{aligned}$$

Analog zeigt man

$$\int_{\mathcal{F}} (0,f_2,0)^t \cdot \mathbf{dS} = \int_B \frac{\partial f_2}{\partial y}(x,y,z) \; d(x,y,z)$$

und

$$\int_{\mathcal{F}} (f_1,0,0)^t \cdot \mathbf{dS} = \int_B \frac{\partial f_1}{\partial x}(x,y,z) \; d(x,y,z).$$

Kombinieren wir die drei Gleichungen, so folgt die Aussage des Satzes. — Eigentlich muß man hier noch zeigen, daß $\int_{\mathcal{F}}$ beim Wechsel der Parameterdarstellung erhalten bleibt, und Proposition B2.7 ist in der vorliegenden Form hier nicht anwendbar, aber dieses technisch aufwendige Detail wollen wir an der Stelle ignorieren. □

Bemerkung B2.12 (Integralsatz von Gauß)

Wie beim Integralsatz von Green möchten wir auch für den Integralsatz von Gauß anmerken, daß die Voraussetzung, daß es sich bei B um einen $\mathcal{C}^1$-Normalbereich handeln muß, viel zu restriktiv ist. Es reicht, daß es sich um eine kompakte Menge handelt, deren Oberfläche, d.h. deren Rand ∂B, eine stückweise stetig differenzierbare Fläche ist. Insbesondere gilt er etwa, wenn B ein Kugel ist, deren Oberfläche kein $\mathcal{C}^1$-Normalbereich im Sinne unserer Definition ist.

Wir können dann auf jedem Flächenstück den äußeren Normaleneinheitsvektor so wählen, daß er aus B hinauszeigt, wie wir das in Bemerkung B2.10 für den $\mathcal{C}^1$-Normalbereich getan haben.

Man beachte auch, daß ∂B dann als endliche Vereinigung stetig differenzierbarer Bilder von Jordan-Nullmengen im $\mathbb{R}^3$ eine Jordan-Nullmenge ist und daß B damit Jordan-meßbar ist. Besitzt B eine Zerlegung in endlich viele $\mathcal{C}^1$-Normalbereiche, so ist die Voraussetzung sicher erfüllt.

Beispiel B2.13

Wir betrachten das Vektorfeld

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (x, y, z)^t \mapsto (x^2 + e^{y^2+z^2}, y^2 + x^2z^2, z^2 - e^y)^t$$

und den Einheitswürfel mit Kantenlänge 1 und Eckpunkten $(0, 0, 0)^t$ und $(1, 1, 1)^t$. Um das Oberflächenintegral $\int_{\partial B} f \cdot \mathbf{dS}$ zu berechnen müßte man die Oberflächenintegrale über sechs Seiten berechnen und diese dazu parametrisieren. Mit Hilfe des Satzes von Gauß können wir stattdessen die Divergenz

$$\operatorname{div}(f) : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (x, y, z)^t \mapsto 2x + 2y + 2z$$

über den Einheitswürfel integrieren und erhalten

$$\int_{\partial B} f \cdot \mathbf{dS} = \int_0^1 \int_0^1 \int_0^1 2x + 2y + 2z \; dz \; dy \; dx = 3.$$

D) Der Integralsatz von Stokes

Der Integralsatz von Stokes setzt das Oberflächenintegral der Rotation eines Vektorfeldes über eine Fläche mit dem Kurvenintegral des Vektorfeldes über die Randkurve der Fläche in Beziehung. Dazu muß die Fläche aber orientierbar sein.

Definition B2.14 (Glatte orientierbare Flächen)
Eine glatte Fläche $\mathcal{F}$ im $\mathbb{R}^3$ heißt *orientierbar*, wenn sie ein stetiges Normaleneinheitsvektorfeld $n : \mathcal{F} \longrightarrow \mathbb{R}^3$ besitzt. Wir nennen n dann eine *Orientierung* von $\mathcal{F}$.

Bemerkung B2.15 (Orientierbare Flächen)

a. Ist die glatte Fläche $\mathcal{F}$ durch die Parameterdarstellung $\phi : U \longrightarrow \mathbb{R}^3$ auf dem Parameterbereich D gegeben und habe $D\phi$ stets vollen Rang. Dann ist $\mathcal{F}$ *orientierbar*, wenn für je zwei Punkte $p, q \in D$ mit $\phi(p) = \phi(q)$ auch

$$\frac{\phi_u(p) \times \phi_v(p)}{||\phi_u(p) \times \phi_v(p)||_2} = \frac{\phi_u(q) \times \phi_v(q)}{||\phi_u(q) \times \phi_v(q)||_2} \tag{258}$$

gilt. Der Einheitsnormalenvektor an $\mathcal{F}$, der durch ϕ definiert wird, hängt dann nämlich nicht von dem Parameter in D ab, durch den wir ihn definieren.

b. Wir interpretieren die Definition der Orientierbarkeit wie folgt: Eine Fläche soll orientierbar heißen, wenn man ihre beiden Seiten voneinander unterscheiden kann. Es soll nicht möglich sein, mit einem Boot auf auf der Fläche entlang zu segeln und dabei an einem Punkt zweimal vorbei zu kommen, einmal auf der einen Seite, einmal auf der anderen. Denken wir uns den Normaleneinheitsvektor der Fläche als den Mast des Bootes, dann sollen wir diesen nicht entlang einer geschlossenen stetig differenzierbaren Kurve auf der Fläche transportieren können und am Endpunkt

einen Normaleneinheitsvektor haben, der dem am Anfangspunkt entgegengesetzt ist. Ist die Fläche durch eine Parameterdarstellung gegeben, so stellt die Bedingung in (258) sicher, daß durch die Parametrisierung eine Orientierung definiert wird.

c. Ist die Fläche nur stückweise stetig differenzierbar und haben wir für die einzelnen Stücke eine Orientierung, müssen diese an den Bruchstellen konsistent gewählt sein, damit wir eine Orientierung für die ganze Fläche erhalten.

Beispiel B2.16 (Das Möbiusband)
Ein einfaches Beispiel für eine nicht-orientierbare Fläche im $\mathbb{R}^3$ erhalten wir, indem wir einen Papierstreifen ab einem Ende einmal um 180 Grad um seine Achse drehen und dann zusammenkleben. Dies wird durch folgende Abbildung beschrieben

$$\phi : [-1,1] \times [0,2\pi] \longrightarrow \mathbb{R}^3 : (u,v)^t \mapsto \begin{pmatrix} \left(1+\frac{u}{2}\cdot\cos\left(\frac{v}{2}\right)\right)\cdot\cos(v) \\ \left(1+\frac{u}{2}\cdot\cos\left(\frac{v}{2}\right)\right)\cdot\sin(v) \\ \frac{u}{2}\cdot\sin\left(\frac{v}{2}\right) \end{pmatrix}.$$

Für konstante Werte von v erhalten wir Geradenstücke, die bei Variation der Variablen v von 0 bis 2π wegen des Terms $\cos(\frac{v}{2})$ um 180 Grad gedreht werden.

Satz B2.17 (Integralsatz von Stokes)
Sei $\mathcal{F}$ eine glatte orientierbare Fläche im $\mathbb{R}^3$ mit zweifach stetig differenzierbarer Parameterdarstellung $\phi : U \longrightarrow \mathbb{R}^3$ und Parameterbereich $D \subset U$ und $U \subseteq \mathbb{R}^2$ offen, und sei dabei D ein $\mathcal{C}^1$-Normalbereich im $\mathbb{R}^2$ mit positiv orientierter Randkurve $\gamma : [a,b] \longrightarrow \mathbb{R}^2$. Für ein stetig differenzierbares Vektorfeld f auf einer offenen Umgebung von $\mathcal{F}$ gilt dann:

$$\int_{\mathcal{F}} \mathrm{rot}(f)\cdot\mathbf{dS} := \int_{\phi} \mathrm{rot}(f)\cdot\mathbf{dS} = \int_{\phi\circ\gamma} f\cdot\mathbf{dx}.$$

Beweis: Der Beweis ergibt sich aus einer Anwendung des Integralsatzes von Green B1.31 und einer Nebenrechnung (259), wobei wir die Hilfsabbildung

$$g : U \longrightarrow \mathbb{R}^2 : (u,v)^t) \mapsto g(u,v) = \left(f(\phi(u,v))^t \circ D\phi(u,v)\right)^t$$

verwenden und die Integrale in Kurform schreiben:

$$\begin{aligned}\int_{\phi\circ\gamma} f\cdot\mathbf{dx} &= \int_a^b \langle f(\phi\circ\gamma),(\phi\circ\gamma)'\rangle\,dt\\ &= \int_a^b f(\phi\circ\gamma)^t\circ D\phi(\gamma)\circ\gamma'\,dt = \int_a^b g(\gamma)^t\circ\gamma'\,dt\\ &= \int_a^b \langle g(\gamma),\gamma'\rangle\,dt = \int_\gamma g\cdot\mathbf{dx}\\ &\overset{B1.31}{=} \int_D \frac{\partial g_2}{\partial u}-\frac{\partial g_1}{\partial v}\,d(u,v)\\ &\overset{(259)}{=} \int_D \langle \mathrm{rot}(f),\phi_u\times\phi_v\rangle\,d(u,v) = \int_{\mathcal{F}} \mathrm{rot}(f)\cdot\mathbf{dS}.\end{aligned}$$

Es bleibt die Nebenrechnung durchzuführen. Dazu berechnen wir zunächst die benötigten partiellen Ableitungen von g:

$$\begin{aligned}\frac{\partial g_2}{\partial u} &= \frac{\partial}{\partial u}\sum_{i=1}^3 (f_i\circ\phi)\cdot\phi_{iv}\\ &= \sum_{i=1}^3\left((f_i\circ\phi)\cdot\phi_{ivu}+\sum_{j=1}^3(f_{ix_j}\circ\phi)\cdot\phi_{ju}\cdot\phi_{iv}\right)\end{aligned}$$

sowie

$$\begin{aligned}\frac{\partial g_1}{\partial v} &= \frac{\partial}{\partial v}\sum_{i=1}^3 (f_i\circ\phi)\cdot\phi_{iu}\\ &= \sum_{i=1}^3\left((f_i\circ\phi)\cdot\phi_{iuv}+\sum_{j=1}^3(f_{ix_j}\circ\phi)\cdot\phi_{jv}\cdot\phi_{iu}\right).\end{aligned}$$

Unter Beachtung des Satzes von Schwarz 25.5 $\phi_{iuv}=\phi_{ivu}$ erhalten wir für die Differenz der beiden Ausdrücke dann

$$\frac{\partial g_2}{\partial u}-\frac{\partial g_1}{\partial v} = \sum_{i=1}^3\sum_{j=1}^3 (f_{ix_j}\circ\phi)\cdot\big(\phi_{ju}\cdot\phi_{iv}-\phi_{jv}\cdot\phi_{iu}\big) = \langle \mathrm{rot}(f),\phi_u\times\phi_v\rangle \tag{259}$$

wobei wir für das letzte Gleichheitszeichen

$$\mathrm{rot}(f) = (f_{3x_2}-f_{2x_3}, f_{1x_3}-f_{3x_1}, f_{2x_1}-f_{1x_2})^t$$

und

$$\phi_u\times\phi_v = (\phi_{2u}\phi_{3v}-\phi_{2v}\phi_{3u}, \phi_{3u}\phi_{1v}-\phi_{3v}\phi_{1u}, \phi_{1u}\phi_{2v}-\phi_{1v}\phi_{2u})^t$$

beachten. □

Bemerkung B2.18 (Integralsatz von Stokes)

Wie beim Integralsatz von Green B1.31 und beim Integralsatz von Gauß B2.11 kann man die Voraussetzungen an den Integralsatz von Stokes B2.17 deutlich abschwächen. Es reicht, daß die Fläche stückweise stetig differenzierbar und orientierbar ist, und

der Parameterbereich braucht nur eine Jordan-Kurve als Randkurve zu haben (siehe Bemerkung B1.32).

Beispiel B2.19
Das Vektorfeld $f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (x, y, z)^t \mapsto (y, -x, 0)^t$ hat das Vektorfeld

$$\mathrm{rot}(f) : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (x, y, z)^t \mapsto (0, 0, 2)^t$$

als Rotation, so daß wir etwa den Fluß von letzterem Vektorfeld durch die Oberfläche der oberen Halbkugel B mit Radius 1 als

$$\int_B \mathrm{rot}(f) \cdot \mathbf{dS} = \int_{\partial K} f \cdot \mathbf{dx}$$

von f über den Rand des Einheitskreises berechnen können. Für dieses Integral haben wir aber im Prinzip in Beispiel B1.22 berechnet und es hat den Wert 2π.

Aufgaben

Aufgabe B2.20 (Greensche Formeln)
Seien $F, G : U \longrightarrow \mathbb{R}$ zweifach stetig differenzierbare Skalarfelder mit $U \subseteq \mathbb{R}^3$ offen und sei $B \subseteq U$ ein $\mathcal{C}^1$-Normalbereich.

a. Zeige, $\int_B \Delta(F)(x, y, z)\, d(x, y, z) = \int_{\partial B} \frac{\partial F}{\partial n}\, dS$, wobei $\frac{\partial F}{\partial n}$ die Richtungsableitung von F in Richtung des äußeren Normaleneinheitsvektors ist.

b. Zeige, $\int_B (F \cdot \Delta(G) + \langle \nabla F, \nabla G \rangle)(x, y, z)\, d(x, y, z) = \int_{\partial B} F \cdot \frac{\partial G}{\partial n}\, dS$.

c. Zeige, $\int_B (F \cdot \Delta(G) - G \cdot \Delta(F))(x, y, z)\, d(x, y, z) = \int_{\partial B} F \cdot \frac{\partial G}{\partial n} - G \cdot \frac{\partial F}{\partial n}\, dS$.

Aufgabe B2.21
Sei $\mathcal{F}$ die Oberfläche des Einheitswürfels mit den Eckpunkte $(0, 0, 0)^t$ und $(1, 1, 1)^t$. Berechne das Oberflächenintegral $\int_{\mathcal{F}} f \cdot \mathbf{dS}$ für das Vektorfeld

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (x, y, z)^t \mapsto (4xz, -y^2, yz)^t.$$

Aufgabe B2.22
Sei $\mathcal{F}$ die Oberfläche der nördlichen Hemisphäre der Einheitskugel mit Mittelpunkt im Ursprung und sei $q = (0, 0, z)$ ein fester Punkt mit $z > 1$. Berechne das Oberflächenintegral $\int_{\mathcal{F}} f\, dS$ für das Skalarfeld

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R} : p \mapsto ||p - q||_2.$$

Aufgabe B2.23

Sei $\mathcal{F}$ wie in Aufgabe 50 und sei

$$f : \mathbb{R}^3 \longrightarrow \mathbb{R}^3 : (x, y, z)^t \mapsto (1, xz, xy)^t.$$

Berechne das Oberflächenintegral $\int_{\mathcal{F}} \operatorname{rot}(f) \cdot \mathbf{dS}$.

Aufgabe B2.24

Sei $f : [a, b] \longrightarrow \mathbb{R}_{\geq 0}$ eine stetig differenzierbare Funktion und sei $\mathcal{F}$ die glatte Fläche im $\mathbb{R}^3$, die entsteht, indem man den Graphen von f um die x-Achse rotieren läßt. Zeige, daß der Oberflächeninhalt von $\mathcal{F}$ durch die folgende Formel berechnet werden kann:

$$I(\mathcal{F}) = 2\pi \cdot \int_a^b f(x) \cdot \sqrt{1 + (f'(x))^2}\, dx.$$

Literaturverzeichnis

[Ama83] Herbert Amann, *Gewöhnliche Differentialgleichungen*, Walter de Gruyter, 1983.

[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, TU Kaiserslautern, 2010.

[Ebb92] Heinz-Dieter Ebbinghaus (ed.), *Zahlen*, 3 ed., Springer, 1992.

[Gat08] Andreas Gathmann, *Grundlagen der Mathematik*, Vorlesungsskript WS2007/08, TU Kaiserslautern, 2008.

[Har15] Peter Hartmann, *Mathematik für Informatiker*, Vieweg, 2015.

[Hei19] Matthias Hein, *Mathematik für Informatiker I*, Vorlesungsskript 2018/19, Universität Tübingen, 2019.

[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, TU Kaiserslautern, 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 Debreccen **1** (1949/50), 254–57.