

Analysis

Marius Müller, Luis Lohse

Akademisches Jahr 2025/2026

Inhaltsverzeichnis

1	Grundbegriffe	6
1.1	Aussagenlogik	6
1.2	Mengenlehre	10
1.3	Abbildungen	16
2	Vorkurs: Folgen	27
2.1	Definition einer Folge	34
2.2	Wachstum von Folgen	39
2.3	Konvergenz und Grenzwerte	65
2.4	Abzählungen	91
3	Die reellen Zahlen	95
3.1	Körper	95
3.2	Anordnungen	103
3.3	Natürliche, Ganze und Rationale Zahlen	120
3.4	Ordnungsvollständigkeit und Reelle Zahlen	145
3.5	Folgen in $\mathbb{R}$	156
4	Die komplexen Zahlen	179
4.1	Definition und Schreibweisen	179
4.2	Konjugation und Absolutbetrag	183
4.3	Folgen in $\mathbb{C}$	192
5	Reihen	197
5.1	Konvergenzkriterien	204
5.2	Potenzreihen	226
5.3	Das Cauchy-Produkt	234
5.4	Die Exponentialfunktion	238
5.5	Trigonometrische Funktionen	246

6	Funktionsgrenzwerte, Stetigkeit	266
6.1	Funktionsgrenzwerte	268
6.2	Stetigkeit	291
6.3	Anwendung: Logarithmus und allgemeine Potenz	311
6.4	Gleichmäßige Stetigkeit	325
7	Differentialrechnung	328
7.1	Ableitungsregeln	341
7.2	Lokale Extrema	352
7.3	Mittelwertsätze und Monotonieverhalten	362
7.4	Höhere Ableitungen	391
7.5	Konvexität	413
8	Integralrechnung	423
8.1	Definition des Riemann-Integrals	425
8.2	Eigenschaften des Integrals	442
8.3	Der Fundamentalsatz der Analysis	458
8.4	Substitution und partielle Integration	470
8.5	Uneigentliche Integrale	486
8.6	Integration komplexwertiger Funktionen	491
9	Funktionenfolgen	494
9.1	Punktweise und gleichmäßige Konvergenz	494
9.2	Vererbung der Stetigkeit	504
9.3	Vertauschung von Grenzprozessen	506
10	Mehrdimensionale Analysis	514
10.1	Die Euklidische Norm	515
10.2	Mehrdimensionale Folgen	518
10.3	Topologie des $\mathbb{R}^d$	524
10.4	Funktionsgrenzwerte in mehreren Variablen	548
10.5	Mehrdimensionales Ableiten	572
10.6	Höhere Ableitungen	611
10.7	Differenzierbare Umkehrfunktionen	645
10.8	Extrema unter Nebenbedingungen	648
11	Abstrakte Analysis	658
11.1	Topologische Räume	658
11.2	Metrische Räume	666
11.3	Normierte Räume	697

Vorwort

Herzlich Willkommen im Mathestudium! Ich freue mich sehr, dass Sie sich dazu entschieden haben, an der Uni Augsburg Mathematik zu studieren und hoffe, Ihnen in diesem Semester die Grundlagen der Analysis beizubringen. Um bei dieser Vorlesung gut folgen zu können möchte ich Ihnen einige Tipps vorab mitgeben:

1. Bearbeiten Sie die (wöchentlich erscheinenden) *Übungsaufgaben*! „Bearbeiten“ bedeutet: Befassen Sie sich mit der Aufgabe, schreiben Sie sich alle relevanten Definitionen und Sätze aus dem Skript heraus und *rätseln* Sie (gerne auch mal etwas länger und gerne auch gemeinsam mit Kommiliton*innen), wie Sie diese Konzepte verbinden können, um eine Lösung zu erhalten. *Geben Sie nicht direkt auf*, wenn Sie nicht auf die Lösung kommen. Je mehr Zeit und „Hirnschmalz“ Sie auf dem Weg zur richtigen Lösung investieren, desto größer wird am Ende der Lerneffekt (und auch Ihre Freude) sein. Wenn Sie eine Aufgabe gelöst haben, nehmen Sie sich auch danach die Zeit, die Lösung logisch stringent und *sauber* aufzuschreiben. Auch das ist ein wichtiges Lernziel im ersten Semester und wir werden auch in der Vorlesung besprechen, wie es funktioniert. Ein hilfreiches Zusatzangebot bei der Bearbeitung der Aufgaben ist der *offene Matheraum*. Und noch ein Tipp: Schreiben Sie nicht einfach ab und googlen Sie die Lösungen nicht. Immer wenn Sie googlen, nehmen Sie sich die Gelegenheit, mit Kraft Ihrer *eigenen Kreativität* auf eine Lösung zu kommen.

Geben Sie die Aufgaben mit Ihrer Lerngruppe ab und sehen Sie sich das *Feedback*, das Sie von Ihren Tutor*innen erhalten, gut an! Sprechen Sie uns bei Unklarheiten an und kommunizieren Sie in den Übungsgruppen auch, welche Aufgaben Ihnen Probleme bereitet haben!

2. Bilden Sie *Lerngruppen*! Am besten solche, die sich regelmäßig (z.B. in allen *Hohlstunden*) treffen und dort konstruktiv über die Übungsaufgaben diskutieren. Eine solche Lerngruppe muss nicht die ganze Zeit interaktiv sein: Zunächst kann man allein in Stillarbeit Definitionen und Sätze nachschlagen, kurz nachdenken, und danach erst über mögliche Lösungswege diskutieren. Auch Personen, die aus logistischen Gründen nicht oft an der Uni sind, sollten Zeit für einen solchen Austausch einplanen. Übrigens bin ich der Auffassung, dass *Teamwork* sich nirgends so gut erlernen lässt wie im Studium der Mathematik! Im Gegensatz zu anderen Teamwork-Situationen, wo es oft nur um Aufgabenverteilung/Ideensammlung geht, steht man hier vor Problemen, die man leichter mit *gemeinsamem Rätseln* lösen kann.
3. Gehen Sie in die *Vorlesung* und nehmen Sie auch aktiv daran Teil! Nehmen Sie ihr Skript mit und schreiben Sie wichtige Anmerkungen und zusätzliche Punkte entweder mit Bleistift ins Skript oder auf Extra-Zettel. Betrachten Sie dieses Skript keineswegs als vollständig: Nur Sie können es vollständig machen, indem Sie ergänzen, was Ihnen *persönlich* hilft, die Inhalte zu verstehen. Stellen Sie bitte auch aktiv *Fragen*, wenn Sie etwas nicht verstanden haben. Gerne können Sie sich dazu während der Vorlesung melden, alternativ kommen Sie nach der Vorlesung zu mir vor.

Gerade am Anfang des Studiums ist es manchmal schwer, eine Frage konkret zu formulieren. Dies löst bei vielen Studierenden Hemmungen aus, Fragen zu stellen, sollte es aber nicht! Kommen Sie in diesem Fall doch einfach nach der Stunde vor, zeigen Sie auf die Stelle im Skript, die Fragezeichen bei Ihnen hervorruft und sagen Sie „Ich habe das hier nicht verstanden“. Ich bin mir sicher, dass wir schnell gemeinsam herausfinden, wo Ihr Problem liegt! Scheuen Sie auch nicht, mit Fragen aus vergangenen Vorlesungen zu mir zu kommen. Für Fragen aller Art biete ich übrigens auch eine *Sprechstunde* an. Termine finden Sie unten.

4. Sie werden merken, dass ziemlich viel Zeit draufgeht, wenn Sie die oben genannten Tipps gewissenhaft befolgen. Das ist zum Teil beabsichtigt, aber dennoch sollten Sie sich *Ruhepausen und Ausgleich* gönnen und sich nicht unter Druck setzen lassen. Innere Ruhe ist sowohl wichtig, um die mathematischen Inhalte gut zu verstehen, als auch für die mentale Gesundheit. Wenn Sie den Eindruck haben, dass Sie nicht mehr zur Ruhe kommen können, dann sollten Sie darüber nachdenken, wie Sie ihr Leben besser mit dem Studium synchronisieren können. Wie das genau geht, können Sie nur sich selbst beantworten: Manchen von Ihnen würden mehr Hobbies oder Kontakt mit Freund*innen zum Abschalten gut tun. Andere von Ihnen neigen dazu, sich mit Hobbies oder Nebenjobs zu überfordern. Achten Sie auf sich, schaffen Sie Räume für Entlastung und überfordern Sie sich nicht! Nutzen Sie auch die Angebote der Studienberatung, wenn Sie das Gefühl haben, überfordert oder belastet zu sein.

Literaturhinweise

1. Forster, Otto. *Analysis 1. Differential- und Integralrechnung in einer Veränderlichen*. Springer-Lehrbuch. Springer-Verlag, 2011. ISBN: 978-3-8348-1251-3.
2. Beck, Lisa. *Analysis 1. Vorlesungsskript, Uni Augsburg* [LINK](#)
3. Beck, Lisa. *Analysis 2. Vorlesungsskript, Uni Augsburg* [LINK](#)
4. Mugnolo, Delio. *Analysis 1. Vorlesungsskript, Uni Ulm* [LINK](#)
5. Königsberger, Konrad. *Analysis 1*. Springer-Lehrbuch. Springer-Verlag, 2004. ISBN: 978-3-540-40371-5.
6. Hildebrandt, Stefan. *Analysis 1*. Springer-Lehrbuch. Springer-Verlag, 2006. ISBN: 978-3-540-25368-6

Das Buch von Forster ist ein äußerst kompakt und kurz gehaltenes Lehrbuch. Es eignet sich als Nachschlagewerk und bildet die inhaltliche Grundlage für die Vorlesung. Das Skript von Beck ist sehr ausführlich und enthält viele Illustrationen, an denen ich mich orientiert habe. An dieser Stelle Dankeschön für die Erlaubnis! Hier habe ich auch inhaltlich einige Add-Ons zum Buch von Forster her, die im Augsburger Analysis-Curriculum üblich sind. Das Skript von Mugnolo enthält die Vorlesung, mit der ich selbst Analysis gelernt habe. Es hat daher viel damit zu tun, wie ich Analysis verstehe und weiterzugeben versuche. Beck's Skript scheint

sich zu einem gewissen Grad auch an dem (wie ich finde auch sehr gelungenen) Lehrbuch von Königsberger zu orientieren. Es zeichnet sich durch hohe Sorgfalt und gute geometrische Erklärungen aus.

Herr Prof. Oldenburg (Professor für Mathematikdidaktik) rät auch zur Benutzung von *Serlo* [LINK ZU SERLO](#). Ich finde die Inhalte gut aufbereitet und finde es gut, dass *hinter die Kulissen* der Analysis geschaut wird (z.B. *warum* wir Sachen so definieren, wie sie definiert sind). Ich habe das Werk bei der Erstellung des Skriptes zwar nicht explizit verwendet, sehe es aber als gutes Querlesewerk.

Bevor wir anfangen möchte ich auch noch meinem Tutor Luis Lohse für das sorgfältige Erstellen der schönen Abbildungen in diesem Skript danken!

1 Grundbegriffe

13.10.2025

1.1 Aussagenlogik

In diesem (kurz gehaltenen) Kapitel wollen wir die Basis für unsere mathematische Arbeitsweise erklären: Die *Logik*. Am Anfang des Studiums benötigt man noch nicht so viel Logik. Und doch ist Logik ein umfangreiches Gebiet der Mathematik. Die Aussagenlogik beschäftigt sich mit sogenannten *Aussagen*.

Definition 1.1.1 — AUSSAGE

Eine (*logische*) *Aussage* ist eine Behauptung, die einen *Wahrheitsgehalt* hat, d.h., die entweder *wahr* (w) oder *falsch* (f) sein kann. Ist eine Aussage wahr, so sagen wir auch häufig, dass sie *gilt*.

Man beachte: Logische Aussagen müssen stets eindeutig auf ihren Wahrheitsgehalt überprüfbar sein. Meinungsäußerungen sind daher z.B. nicht zugelassen. Beispiele für Aussagen sind

A : Die Erde ist eine Scheibe.

B : In Augsburg gibt es mehr Brücken als in Venedig.

C : $2 + 2 = 5$.

Klarerweise ist Aussage A falsch, B ist (wie Sie eventuell schon wissen) wahr. Aussage C ist auch falsch, aber streng genommen können wir das noch gar nicht beurteilen: Wir haben nämlich noch nicht definiert, was eine „2“ oder eine „5“ ist ...

Es ist möglich, aus mehreren gegebenen Aussagen, neue Aussagen zu bilden. Dazu benutzt man *Junktoren*.

Definition 1.1.2 — JUNKTOREN

Es seien A und B zwei Aussagen.

- (i) Die Aussage $\neg A$ („nicht A “) ist die Aussage, die wahr ist, wenn A falsch ist und falsch ist wenn A wahr ist.
- (ii) Die Aussage $A \wedge B$ („ A und B “) ist wahr, wenn sowohl A als auch B wahr sind und falsch, wenn eine der beiden Aussagen falsch ist.
- (iii) Die Aussage $A \vee B$ („ A oder B “) ist wahr, wenn (mindestens) eine der beiden Aussagen A oder B wahr sind. Sie ist falsch, wenn beide Aussagen falsch sind.
- (iv) Die Aussage $A \implies B$ („ A impliziert B “) ist wahr, wenn A wahr ist und B wahr ist oder A falsch ist (unabhängig davon welchen Wahrheitsgehalt B hat). Sie ist falsch, wenn A wahr ist und B falsch.^a

(v) Die Aussage $A \iff B$ („A ist äquivalent zu B“) ist stets wahr wenn $(A \implies B) \wedge (B \implies A)$ wahr ist. Sie ist falsch wenn $(A \implies B) \wedge (B \implies A)$ falsch ist.

^aMit einfacheren Worten: Gilt „ $A \implies B$ “ so bedeutet das, dass aus der Wahrheit von A die Wahrheit von B *gefolgt* werden kann.

Am einfachsten kann man sich den Wahrheitsgehalt der oben definierten Aussagen an einer *Wahrheitstabelle* klarmachen.

Beispiel 1.1.3 — WAHRHEITSTABELLEN

Es seien A und B zwei Aussagen. Dann können wir die folgende *Wahrheitstabelle* aufstellen.

A	B	$A \wedge B$	$A \vee B$	$A \implies B$	$B \implies A$	$A \iff B$
w	w	w	w	w	w	w
w	f	f	w	f	w	f
f	w	f	w	w	f	f
f	f	f	f	w	w	w

Das bedeutet: Haben A und B den Wahrheitsgehalt aus den ersten beiden Spalten, so haben die Aussagen $A \wedge B$, $A \vee B$, $A \implies B$, $B \implies A$ und $A \iff B$ stets den Wahrheitsgehalt aus der entsprechenden Spalte.

Eine Lehre daraus: Äquivalente Aussagen haben stets denselben Wahrheitsgehalt! Das bedeutet äquivalente Aussagen sind entweder beide wahr oder beide falsch.

So kann man auch nachweisen, dass zwei Aussagen C und D äquivalent sind: Da die Äquivalenz von Aussagen über den Wahrheitsgehalt charakterisiert ist, genügt es zu nachzuweisen, dass C und D denselben Wahrheitsgehalt haben. Etwas kontraintuitiv: Die beiden in unserem Verständnis nicht korrelierten Aussagen

A : Die Erde ist eine Scheibe.

C : $2 + 2 = 5$.

sind äquivalent (d.h. $A \iff C$)! (Denn sie sind ja beide falsch...)

Proposition 1.1.4 — DEMORGAN’SCHES REGELN.

Es seien A und B zwei Aussagen.

(i) Die Aussagen $\neg(A \wedge B)$ und $(\neg A) \vee (\neg B)$ sind äquivalent.

(ii) Die Aussagen $\neg(A \vee B)$ und $(\neg A) \wedge (\neg B)$ sind äquivalent.

Beweis. Zu (i): Um die Äquivalenz zu zeigen, müssen wir uns davon überzeugen, dass die Aussagen stets denselben Wahrheitsgehalt haben. Der Wahrheitsgehalt der Aussagen $\neg(A \wedge B)$ und $(\neg A) \vee (\neg B)$ kann stets aus dem Wahrheitsgehalt der Aussagen A und B *ermittelt* werden. Hierfür nutzen wir wiederum eine Wahrheitstabelle:

A	B	$\neg A$	$\neg B$	$A \wedge B$	$\neg(A \wedge B)$	$(\neg A) \vee (\neg B)$
w	w	f	f	w	f	f
w	f	f	w	f	w	w
f	w	w	f	f	w	w
f	f	w	w	f	w	w

Wir beobachten: Egal welchen Wahrheitsgehalt A und B auch immer haben, der Wahrheitsgehalt der Aussagen $\neg(A \wedge B)$ und $(\neg A) \vee (\neg B)$ ist derselbe! Damit haben die Aussagen stets denselben Wahrheitsgehalt.

Zu (ii): Der Beweis funktioniert genauso. Wenn Sie möchten können Sie es selbst üben und sich die Wahrheitstabelle für die obige Aussage aufschreiben.

Mithilfe weiterer Überlegungen können wir *Rechengesetze* für Junktoren herleiten. Diese wollen wir nun zusammenfassen

Proposition 1.1.5 — RECHENREGELN FÜR UND-ODER-JUNKTOREN.

Es seien A, B, C Aussagen.

- (i) $A \wedge B$ ist äquivalent zu $B \wedge A$.
- (ii) $A \vee B$ ist äquivalent zu $B \vee A$.
- (iii) $A \wedge (B \wedge C)$ ist äquivalent zu $(A \wedge B) \wedge C$.
- (iv) $A \vee (B \vee C)$ ist äquivalent zu $(A \vee B) \vee C$.
- (v) $A \wedge (B \vee C)$ ist äquivalent zu $(A \wedge B) \vee (A \wedge C)$.
- (vi) $A \vee (B \wedge C)$ ist äquivalent zu $(A \vee B) \wedge (A \vee C)$.

Beweis. Ist den Lesenden als Übungsaufgabe überlassen (über Wahrheitstabellen).

Die Aussagen (iii) bzw. (iv) bedeuten, dass die Reihenfolge, in der man die Aussagen A, B, C mit „und“ bzw. „oder“ verknüpft keine Rolle spielt. Wir lassen daher die Klammern weg und schreiben einfach $A \wedge B \wedge C$ bzw. $A \vee B \vee C$.

Eine weitere bemerkenswerte Eigenschaft ist, dass sich die Implikation äquivalent mit Und-Oder-Junktoren ausdrücken lässt.

Proposition 1.1.6 — WIDERSPRUCHSREGEL

Die Aussage $A \implies B$ ist äquivalent zu $\neg(A \wedge \neg B)$.

Beweis. Über Wahrheitstabellen.

Ferner gibt es noch zwei wichtige Regeln für die logische Implikation, die wir häufig benutzen werden.

Proposition 1.1.7 — RECHENREGELN FÜR DIE IMPLIKATION.

Seien A , B und C Aussagen.

- (i) (*Kettenschluss*) $(A \implies B) \wedge (B \implies C)$ impliziert $A \implies C$.
Genauso gilt: $(A \iff B) \wedge (B \iff C)$ impliziert $A \iff C$.
- (ii) (*Kontraposition*) $A \implies B$ und $\neg B \implies \neg A$ sind äquivalent.

Beweis. Aussage (i) würde man wiederum mit Wahrheitstabellen zeigen. Aussage (ii) kann auch mithilfe der bisher bekannten Regeln gezeigt werden.

$$\begin{array}{rcl} & & A \implies B \\ \Longleftrightarrow & & \neg(A \wedge \neg B) \\ \text{Prop 1.1.6} & & \\ \Longleftrightarrow & & \neg(\neg B \wedge A) \\ \text{Prop 1.1.5(i)} & & \\ \Longleftrightarrow & & \neg(\neg B \wedge \neg(\neg A)) \\ A \iff \neg(\neg A) & & \\ \Longleftrightarrow & & \neg B \implies \neg A. \\ \text{Prop 1.1.6} & & \end{array}$$

Um nun aus dieser *Kette von Äquivalenzen* die Äquivalenz von $A \implies B$ und $\neg B \implies \neg A$ zu folgern, müssen wir die Aussage (i) mehrmals anwenden.

Beispiel 1.1.8

Die Aussagen $A \implies B$ und $\neg A \implies \neg B$ sind im Allgemeinen NICHT äquivalent. Dazu ein Beispiel aus dem Alltag:

A : In ganz Bayern regnet es
 B : Der Augsburger Dom wird nass

Klarerweise ist $A \implies B$ wahr, denn falls es in ganz Bayern regnet so wird der Augsburger Dom sicherlich nass. Nun bilden wir die Verneinungen

$\neg A$: Es regnet nicht in ganz Bayern
 $\neg B$: Der Augsburger Dom wird nicht nass

Man beachte, dass $\neg A \implies \neg B$ falsch ist! Man stelle sich dafür mal vor, dass es in ganz Schwaben und Oberbayern regnet, aber in Franken, Niederbayern und der Oberpfalz strahlender Sonnenschein ist. Dann ist $\neg A$ wahr, denn es regnet ja nicht in ganz Bayern. Die Aussage $\neg B$ ist allerdings falsch, denn der Regen in ganz Schwaben wird den Augsburger Dom sicherlich auch nass machen. Man sieht also, dass $\neg A \implies \neg B$ im Allgemeinen nicht denselben Wahrheitsgehalt hat wie $A \implies B$.

Ein Wort noch dazu: Aus Proposition 1.1.7 (ii) folgt, dass die Implikation $\neg B \implies \neg A$

wahr sein muss. Das ist auch intuitiv: Wenn der Augsburger Dom nicht nass wird, dann kann es nicht in ganz Bayern regnen.

Im folgenden Werden wir übrigens Aussagen wie „ $A \implies B$ “ häufig *in Worten* ausschreiben. Der Wortlaut dafür ist dann „Wenn A , dann B “ oder „Sei A wahr. Dann gilt B “. Diese klassischen *Wenn-dann-Aussagen* finden sich häufig in der Mathematik.

Etwas subtiler noch: Wir werden auch $A \iff B$ häufig in Worten ausschreiben. Der Wortlaut ist „ A genau dann wenn B “ oder „ A gilt dann und nur dann wenn B gilt“.

Wir schließen diesen Abschnitt mit einer weiteren Notation ab: Anstatt $\neg(A \implies B)$ werden wir auch häufig $A \nRightarrow B$ schreiben.

1.2 Mengenlehre

Das Konzept der Menge ist in der Mathematik essentiell. Viele mathematische Objekte (so auch z.B. Funktionen) lassen sich durch Mengen charakterisieren. Man muss dazu sagen, dass wir hier nicht die richtige Definition einer Menge besprechen, sondern nur die sogenannte *naive Mengendefinition*. Diese ist im strengen Sinne keine Definition, weil in ihr noch klärungsbedürftige Begriffe enthalten sind.

Definition 1.2.1 — NAIVE MENGENDEFINITION

Eine *Menge* M ist eine Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens. Diese Objekte nennen wir *Elemente* von M .

Für eine Menge M und ein Objekt x schreiben wir $x \in M$, wenn x ein Element der Menge M ist. Außerdem schreiben wir $x \notin M$ falls x kein Element der Menge M ist, d.h. $\neg(x \in M)$ gilt.

Das Wort *wohlunterschieden* bedeutet dass niemals zwei *gleiche* Objekte in einer Menge doppelt auftauchen dürfen. Man beachte, dass man dafür überhaupt erstmal sagen können muss, was es bedeutet, dass zwei Objekte gleich sind. Mathematischer: man müsste das Gleichheitszeichen „ $=$ “ definieren. Dies vertiefen wir hier an der Stelle nicht und verlassen uns auf unsere Intuition.

Bemerkung 1.2.2 — BESCHREIBUNG VON MENGEN

- Eine Möglichkeit, eine konkrete Menge zu beschreiben ist durch das **Aufzählen aller Elemente**. Die Aufzählung der Elemente wird dann stets in *Mengenklammern* $\{\dots\}$ geschrieben. Beispielsweise ist

$$M := \{\text{Sonne, Mond, Sterne}\}$$

eine Menge. Das Zeichen „ $:=$ “ ist hierbei als „definiert durch“ zu lesen. Ein etwas mathematischeres Beispiel ist die Menge aller *natürlichen Zahlen*

$$\mathbb{N} := \{1, 2, 3, 4, 5, 6, \dots\}. \tag{6}$$

Mit „...“ meinen wir stets, dass die Aufzählung nach der aus den ersten Elementen ersichtlichen Regel fortgesetzt werden soll. Ein Beispiel mit einer etwas komplizierteren Fortsetzungsregel ist die Menge der *Quadratzahlen* $\{1, 4, 9, 16, 25, \dots\}$.

- Eine weitere Möglichkeit, eine Menge zu beschreiben, ist durch **Angabe von Eigenschaften** der Elemente. So kann zum Beispiel bei der Menge aller Quadratzahlen genauer beschrieben werden durch

$$\{n \in \mathbb{N} : \text{Es existiert ein } m \in \mathbb{N} \text{ mit } n = m^2\}. \quad (7)$$

Der Doppelpunkt ist hier als „mit (der Eigenschaft)“ zu lesen.

- Eine letzte Alternative, die wir vorstellen möchten, ist das Angeben eines **Konstruktionsgesetzes für die Elemente**. Die Menge aller Quadratzahlen lässt sich zum Beispiel über das Konstruktionsgesetz

$$\{n^2 | n \in \mathbb{N}\}$$

angeben. Der Querstrich ist hier als „wobei“ zu lesen. Dies ist in diesem Fall wesentlich kürzer als in (7). Jedoch hat man nicht für alle Mengen ein Konstruktionsgesetz zur Verfügung.

Über die in (6) eingeführten natürlichen Zahlen müsste man eigentlich auch noch etwas mehr sagen. Wir wollen das an dieser Stelle aber nicht vertiefen und vertrauen hier zunächst unserer Intuition.

Eine Menge, die uns besonders oft begegnen wird ist die *leere Menge*.

Definition 1.2.3 — LEERE MENGE

Die Menge, die gar keine Elemente enthält wird stets mit $\emptyset$ bezeichnet und *leere Menge* genannt.

Als nächstes möchten wir den Begriff der *Teilmenge* einführen.

Definition 1.2.4 — TEILMENGEN UND MENGENGLEICHHEIT

Es seien M, N Mengen. Wir sagen $M \subseteq N$ („ M ist eine Teilmenge von N “) falls jedes Element von M auch ein Element von N ist. Mit anderen Worten: Es gilt $x \in M \implies x \in N$.

Man darf in diesem Fall übrigens auch $N \supseteq M$ schreiben^a und spricht auch „ N ist eine Obermenge von M “.

Wir nennen die Mengen M und N *gleich* falls $M \subseteq N$ und $N \subseteq M$. Wir schreiben dann $M = N$. Gilt $M \subseteq N$ und $M \neq N$ so schreibt man auch $M \subsetneq N$.

^adenn mathematische Symbole können sowieso stets in jede beliebige Richtung geschrieben werden!

Beachte: Falls $M = N$ so gilt für die Elemente $x \in M \implies x \in N$ und $x \in N \implies x \in M$. Mit anderen Worten: $x \in M \iff x \in N$.

Bemerkung 1.2.5

Für jede beliebige Menge M gilt $\emptyset \subseteq M$. In der Tat: Die Aussage „ $x \in \emptyset \implies x \in M$ “ ist immer wahr, denn die Aussage „ $x \in \emptyset$ “ ist stets falsch. (Hier haben wir Definition 1.1.2 (iv) benutzt, die besagt, dass eine Aussage der Form $A \implies B$ stets wahr ist wenn A falsch ist)

Bemerkung 1.2.6 — UMSORTIERUNG VON ELEMENTEN

Aus Definition 1.2.4 lernt man auch, dass beim Auflisten der Elemente einer Menge die *Reihenfolge der Auflistung* keine Rolle spielt. Ein Beispiel: Die Mengen

$$M := \{\text{Sonne, Mond, Sterne}\}$$

und

$$\widetilde{M} := \{\text{Sterne, Sonne, Mond}\}$$

sind gleich, d.h. $M = \widetilde{M}$.

Mengenoperationen

Wiederum gibt es *Operationen*, die aus gegebenen Mengen neue Mengen generieren.

Schnitte, Vereinigungen, Differenzen**Definition 1.2.7 — MENGENOPERATIONEN**

Es seien M und N zwei Mengen.

- (i) Wir definieren die *Schnittmenge* von M und N („ M geschnitten N “) durch

$$M \cap N := \{x : x \in M \wedge x \in N\}.$$

- (ii) Wir definieren die *Vereinigungsmenge* von M und N („ M vereinigt N “) durch

$$M \cup N := \{x : x \in M \vee x \in N\}.$$

- (iii) Wir definieren die *Differenzmenge* von M und N („ M ohne N “) durch

$$M \setminus N := \{x : x \in M \wedge x \notin N\}.$$

Die obigen Mengenoperationen können durch sogenannte *Venn-Diagramme* veranschaulicht werden, vgl. Abbildung 1.

16.10.2025

Auch für Mengenoperationen gibt es *Rechenregeln*.

Proposition 1.2.8 — RECHENREGELN FÜR MENGEN

Es seien L, M, N Mengen. Dann gilt

- (i) $M \cap N = N \cap M$.
- (ii) $M \cup N = N \cup M$.
- (iii) $L \cap (M \cap N) = (L \cap M) \cap N$.
- (iv) $L \cup (M \cup N) = (L \cup M) \cup N$.
- (v) $L \cap (M \cup N) = (L \cap M) \cup (L \cap N)$.
- (vi) $L \cup (M \cap N) = (L \cup M) \cap (L \cup N)$.

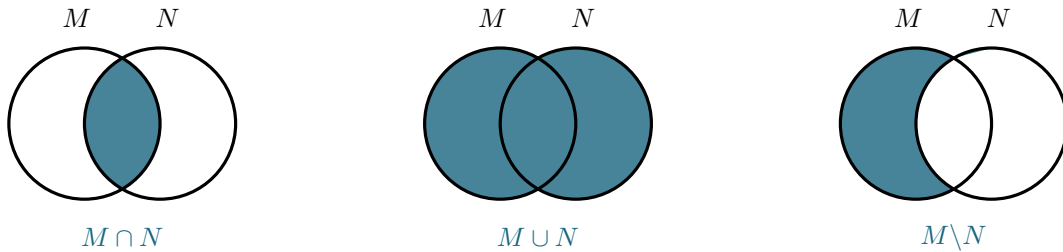


Abbildung 1: Venn-Diagramme

Beweis. Wir zeigen nur (i) und (v). Der Rest ist den Lesenden als Übungsaufgabe überlassen. Zu (i). Wir zeigen, dass $x \in M \cap N \iff x \in N \cap M$. Dazu:

$$\begin{aligned}
 & x \in M \cap N \\
 \iff & \text{Def.1.2.7(i)} \quad (x \in M) \wedge (x \in N) \\
 \iff & \text{Prop.1.1.5(i)} \quad (x \in N) \wedge (x \in M) \\
 \iff & \text{Def.1.2.7(i)} \quad x \in N \cap M.
 \end{aligned}$$

Zu (v). Wir zeigen wiederum, dass $x \in L \cap (M \cup N) \iff x \in (L \cap M) \cup (L \cap N)$. Dazu:

$$\begin{aligned}
 & x \in L \cap (M \cup N) \\
 \iff & \text{Def.1.2.7(i)} \quad (x \in L) \wedge (x \in M \cup N) \\
 \iff & \text{Def.1.2.7(ii)} \quad (x \in L) \wedge (x \in M \vee x \in N) \\
 \iff & \text{Prop.1.1.5(v)} \quad (x \in L \wedge x \in M) \vee (x \in L \wedge x \in N) \\
 \iff & \text{Def.1.2.7(i)} \quad (x \in L \cap M) \vee (x \in L \cap N)
 \end{aligned}$$

$$\stackrel{\text{Def. 1.2.7(ii)}}{\iff} x \in (L \cap M) \cup (L \cap N).$$

Im letzten Schritt muss man die benutzte Definition 1.2.7(ii) *rückwärts lesen*. Dies ist nicht unüblich: Unsere Schreibrichtung ist zwar von links nach rechts, jedoch sind Äquivalenzen stets in beide Richtungen gültig (sprich: $A \iff B$ ist äquivalent zu $B \iff A$ für alle Aussagen A und B).

Da für drei Mengen L, M, N die *Reihenfolge* der Schnittbildung oder Vereinigungsbildung keine Rolle spielt (vgl. Proposition 1.2.8 (iii),(iv)), werden wir wiederum anstatt $L \cap (M \cap N)$ oder $(L \cap M) \cap N$ einfach $L \cap M \cap N$ schreiben. Selbiges gilt für $L \cup M \cup N$.

Kartesisches Produkt Eine weitere Möglichkeit, eine neue Menge aus zwei gegebenen Mengen zu generieren ist das Bilden des *kartesischen Produktes*. Hierfür benötigen wir zunächst das Konzept eines *geordneten Paares*.

Definition 1.2.9 — GEORDNETES PAAR

Es seien x, y zwei Objekte. Dann nennt man die Auflistung (x, y) ein *geordnetes Paar*. Zwei geordnete Paare (x, y) und (x', y') heißen *gleich*, wenn $x = x'$ und $y = y'$.

Im Gegensatz zu einer Menge spielt also bei einem geordneten Paar die *Reihenfolge* eine Rolle.

Definition 1.2.10 — KARTESISCHES PRODUKT

Es seien M und N zwei Mengen. Dann definieren wir das *kartesische Produkt* $M \times N$ („ M kreuz N “) durch

$$M \times N := \{(x, y) : x \in M \text{ und } y \in N\}.$$

Auch für das kartesische Produkt gibt es Rechenregeln.

Proposition 1.2.11 — RECHENREGELN FÜR DAS KARTESISCHE PRODUKT

Es seien N, M und M_1, N_1, M_2, N_2 Mengen. Dann gilt

- (i) $M \times N = \emptyset \iff (M = \emptyset \vee N = \emptyset)$.
- (ii) $(M_1 \times N_1) \cap (M_2 \times N_2) = (M_1 \cap M_2) \times (N_1 \cap N_2)$.
- (iii) $(M_1 \times N_1) \cup (M_2 \times N_2) \subseteq (M_1 \cup M_2) \times (N_1 \cup N_2)$. Gleichheit gilt im Allgemeinen nicht.

Beweis. Ist den Lesenden als Übungsaufgabe überlassen.

Zwar beweisen wir die obige Proposition nicht, aber anhand von Abbildung 2 sollte es möglich sein, die Aussagen besser zu verstehen. Überzeugen Sie sich!

Zur Mathematik gehört es nicht nur, sich zu überlegen was gilt, sondern auch, was *nicht* gilt. Hierzu mal ein Beispiel zum kartesischen Produkt:

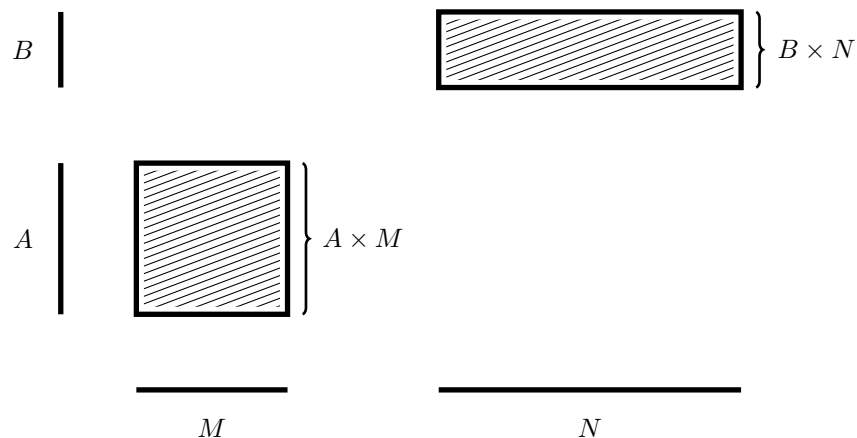


Abbildung 2: Visualisierung der Kartesischen Produkte $A \times M$ und $B \times N$.

Beispiel 1.2.12

Es seien M und N zwei Mengen. Die Behauptung „ $M \times N = N \times M$ “ ist *im Allgemeinen falsch*. Das bedeutet, sie ist nicht für alle möglichen Mengen M, N wahr. Dass eine Behauptung im Allgemeinen falsch ist, zeigt man durch ein *Gegenbeispiel*.

Als Gegenbeispiel kann man hier zum Beispiel $M = \{1\}$ und $N = \{2\}$ anführen. Dann ist $M \times N = \{(1, 2)\}$ und $N \times M = \{(2, 1)\}$. Da $(1, 2) \neq (2, 1)$ gilt in der Tat $M \times N \neq N \times M$.

Man beachte: Es kann durchaus Möglichkeiten geben, M und N so zu wählen, dass die Behauptung wahr wird. Zum Beispiel $M = N = \emptyset$. Das ändert aber nichts daran, dass die Behauptung *im Allgemeinen falsch* ist. Ein Beispiel ist kein Beweis!

Mengen von Mengen Objekte, die wir in Mengen zusammenfassen, können so ziemlich alles sein. Auch Mengen können selbst als Objekte auftreten! Man betrachte das Beispiel

$$M = \{1, \{1\}\}.$$

Auf den ersten Blick sieht das nicht zulässig aus, denn eine Menge darf niemals zwei gleiche Objekte enthalten, vgl. Definition 1.2.1. Jedoch sind die Objekte „1“ (die Ziffer Eins) und „ $\{1\}$ “ (die Menge, die nur aus der 1 besteht) verschiedene Objekte!

Eine wichtige Menge von Mengen ist die *Potenzmenge*.

Definition 1.2.13 — POTENZMENGE

Es sei M eine Menge. Dann heißt

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

die *Potenzmenge* von M oder *Menge aller Teilmengen* von M .

Als Beispiel kann man sich überlegen, dass

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

Die leere Menge sollte man hierbei nicht vergessen, vgl. auch Bemerkung 1.2.5.

Bonusmaterial. DIE RUSSELL'SCHE ANTINOMIE. Die folgende Beobachtung löste 1903 die sogenannte *Grundlagenkrise der Mathematik aus*, die eine Revolution der Logik und Mengenlehre notwendig machte. Als Vorüberlegung sehen wir uns das sogenannte *Barbierparadoxon* an.

Beobachtung 1.2.14 — BARBIERPARADOXON

Stellen Sie sich vor, Sie haben das Wort *Barbier* noch nie gehört. Sie fragen nun eine Person auf der Straße: „Was ist denn ein Barbier?“. Dann antwortet die Person: „Ein Barbier ist eine Person, der alle diejenigen (und nur diejenigen) rasiert, die sich nicht selbst rasieren“. Viele würde das vielleicht zufriedenstellen: Entweder man rasiert sich selber, oder man geht zum Barbier. . .

Sie mit Ihrem mathematischen Geschick entgegnen aber: „Eine solche Person kann es nicht geben! Stellen Sie sich mal vor, es gäbe einen solchen *Barbier*, wie Sie sagen. Wie würde der sich denn jetzt rasieren? Wenn er sich selbst rasiert, dann dürfte er sich gerade nach Ihrer Aussage *nicht* selbst rasieren. Denn er rasiert ja nur diejenigen Leute, die sich nicht selber rasieren. Es könnte ja noch sein, dass er sich nicht selbst rasiert. Aber dann müsste ihn nach Ihrer Aussage ja der Barbier rasieren. Er muss sich also doch selbst rasieren. Ihre Beschreibung eines Barbiers scheint mir absurd!“

Ein ähnliches Problem wie das Barbierparadoxon tritt bei der Definition der Folgenden Menge auf

$$\mathcal{N} := \{M : M \text{ ist eine Menge und } M \notin M\}. \quad (18)$$

Das bedeutet $\mathcal{N}$ ist die Menge aller Mengen, die sich nicht selbst als Element enthalten. Eine solche Menge kann es aber nicht geben! Warum? Stellen wir mal Frage ob $\mathcal{N} \in \mathcal{N}$, d.h. ob $\mathcal{N}$ sich selbst enthält. Wäre $\mathcal{N} \in \mathcal{N}$ so würde nach der Mengenbeschreibung in (18) (angewendet auf $M = \mathcal{N}$) gelten, dass $\mathcal{N} \notin \mathcal{N}$. Ein Widerspruch. Wäre hingegen $\mathcal{N} \notin \mathcal{N}$ so würde die Mengenbeschreibung in (18) gerade nicht auf $M = \mathcal{N}$ zutreffen, weswegen $\mathcal{N} \in \mathcal{N}$ gelten müsste. Wiederum ein Widerspruch. Die Definition in (18) ist also absurd. Dabei sollte es doch laut Bemerkung 1.2.2 eigentlich kein Problem sein, eine Menge wie in Gleichung (18) anhand der Eigenschaften ihrer Elemente zu charakterisieren... In der Realität ist aber gerade das problematisch! Wir werden in diesem Kurs allerdings die Probleme und deren Lösungen nicht besprechen können und verwenden daher ab sofort nur den naiven Mengenbegriff.

1.3 Abbildungen

Mithilfe unseres im vorigen Abschnitt eingeführten Mengenbegriffs können wir nun den Begriff einer *Abbildung* oder *Funktion* einführen. Die Definition klingt zunächst eventuell etwas kontraintuitiv. Wir werden aber schnell klären können, warum es mit dem uns bekannten Funktionsbegriff übereinstimmt.

Definition 1.3.1 — FUNKTION

Es seien D und Z zwei Mengen. Eine *Funktion* f von D nach Z , notiert als $f : D \rightarrow Z$ ist eine Teilmenge von $D \times Z$ mit folgender Eigenschaft:

Für alle $x \in D$ existiert genau ein $y \in Z$ mit $(x, y) \in f$.

Für $x \in D$ bezeichnet man diesen (eindeutig bestimmten) Wert y auch mit $f(x)$ und nennt ihn den *Funktionswert von f an der Stelle x* .

Anstatt $(x, y) \in f$ schreibt man dann auch $f : x \mapsto y$ und man sagt, f ordnet x den Wert y zu. Die Menge D nennt man im Allgemeinen *Definitionsbereich* und die Menge Z nennt man *Zielbereich*.

Bemerkung 1.3.2 — QUANTOREN

Die obige Definition ist leichter zu notieren, wenn man *Quantoren* verwendet. Dies sind nützliche Abkürzungen, die wir an dieser Stelle mal einführen wollen

- (i) $\forall$ bedeutet „Für alle“.
- (ii) $\exists$ bedeutet „Es existiert (mindestens) ein“. Das Wort „mindestens“ muss man in der Mathematik übrigens eigentlich nie mitschreiben: Man geht immer als *Default-Case* davon aus, dass es eventuell noch weitere Objekte mit der genannten Eigenschaft gibt.
- (iii) $\exists!$ bedeutet „Es existiert genau ein“. Das zusätzliche Ausrufezeichen besagt dann, dass es wirklich nur ein Objekt mit der genannten Eigenschaft geben soll. Man liest oft auch „Es existiert ein *eindeutiges*“.
- (vi) „ $\therefore$ “ bedeutet „so, dass“ oder „mit der Eigenschaft, dass“. Kurz liest man es oft einfach nur als „mit“ (auch wenn das nicht immer korrekte deutsche Grammatik erzeugt)

Die obige Definition einer Funktion ist dann leicht in „Quantorsalat“ aufzuschreiben: Eine Teilmenge f von $D \times Z$ ist eine Funktion genau dann wenn

$$\forall x \in D \quad \exists! y \in Z : (x, y) \in f.$$

Verwendet man nun noch die Kurznotation von oben, so heißt das

$$\forall x \in D \quad \exists! y \in Z : f : x \mapsto y.$$

Als Beispiel überlege man sich Folgendes. Seien $D = Z = \{1, 2\}$. Dann ist

$$f := \{(1, 2), (2, 2)\}$$

eine Funktion. Alternativ könnte man schreiben $f : 1 \mapsto 2, f : 2 \mapsto 2$. Oder auch $f(1) = 2$ und $f(2) = 2$.

Hingegen ist

$$g := \{(1, 2), (1, 1)\}$$

keine Funktion. Denn zu $x = 1$ gibt es mehrere Werte $y \in Z$ mit $(1, y) \in g$. Desweiteren gibt es gar keinen Wert $y \in Z$ mit $(2, y) \in g$. Das darf auch nicht sein, denn jedem Element des Definitionsbereiches sollte genau ein (und somit auch mindestens ein) Element aus dem Zielbereich zugeordnet werden.

Beachte: Definitionsbereich und Zielbereich sind stets Teil der Definition einer Funktion. Das wird in nachfolgendem Beispiel besonders deutlich. Die Menge $h := \{(1, 1), (2, 2)\}$ ist eine Funktion über $D = \{1, 2\}$, aber keine über $D = \{1, 2, 3\}$, denn es muss ja jedem Element des Definitionsbereichs „ein Wert zugewiesen werden“. Man muss also beim Definieren einer Funktion stets Definitionsbereich und Zielbereich gesondert angeben. Somit schreibt man also am besten immer $h : D \rightarrow Z$ anstatt nur h .

Bemerkung 1.3.3 — FUNKTIONSVORSCHRIFTEN

Wir wollen nun eine konkret gegebene Funktion $f : D \rightarrow Z$ beschreiben. Besteht D nur aus endlich vielen Elementen, etwa $D = \{x_1, \dots, x_n\}$ für ein $n \in \mathbb{N}^a$ so lässt sich f durch die Auflistung $x_1 \mapsto f(x_1), \dots, x_n \mapsto f(x_n)$ komplett beschreiben.

Typischerweise hat D aber unendlich viele Elemente. In diesem Fall gibt man oft eine *Vorschrift* an, mit der für alle $x \in D$ der Funktionswert $f(x)$ gefunden werden kann. Ein Beispiel: $f : \mathbb{N} \rightarrow \mathbb{N}$, $f : x \mapsto x^2 + 2x$. Dann gilt also $f(1) = 3, f(2) = 8, f(3) = 15, \dots$

Man beachte, dass es hier nicht genügt, nur die ersten drei Werte anzugeben. So zum Beispiel erfüllt $g : \mathbb{N} \rightarrow \mathbb{N}$, $g : x \mapsto 2x^2 - 2x + 4 - |x - 2|$ auch $g(1) = 3, g(2) = 8, g(3) = 15$. Allerdings ist $g \neq f$, denn $f(4) = 32$ und $g(4) = 34$.

^aWas die natürlichen Zahlen sind, wissen wir zu dem Zeitpunkt ja eigentlich noch gar nicht. Für dieses eine Beispiel begnügen wir uns aber mit unserer Intuition.

Bemerkung 1.3.4 — GLEICHHEIT VON FUNKTIONEN

Wer von der Gleichheit zweier Funktionen spricht, sollte sich noch einmal vor Augen führen was Funktionen überhaupt sind: Sie sind nichts anderes als Mengen!(vgl. Definition 1.3.1).

Da wir die Gleichheit von Mengen definiert haben, müssen wir also hier nichts neu definieren oder erklären. Allgemein sind zwei Funktionen $f, g : D \rightarrow Z$ genau dann gleich, wenn $f(x) = g(x)$ für alle $x \in D$ gilt. (Überzeugen Sie sich davon, dass dann f und g auch als Mengen aufgefasst gleich sind)

Ein wichtiges Beispiel für eine Funktion ist die *identische Abbildung*.

Definition 1.3.5 — IDENTISCHE ABBILDUNG

Es sei D eine Menge. Die *identische Abbildung von D* , notiert mit id_D , ist die Abbildung $\text{id}_D : D \rightarrow D, x \mapsto x$.

Als nächstes definieren wir die Begriffe *Bild* und *Urbild*.

Definition 1.3.6 — BILD UND URBILD

Es seien D und Z zwei Mengen und $f : D \rightarrow Z$ eine Funktion.

- (i) Für $U \subseteq D$ nennen wir $f(U) := \{f(x) : x \in U\}$ das *Bild* von U unter f . Mit anderen Worten $f(U) = \{y \in Z : \exists x \in U : f(x) = y\}$.
- (ii) Für $V \subseteq Z$ ist das *Urbild* von V unter f die Menge $f^{-1}(V) := \{x \in D : f(x) \in V\}$.

Beachte, dass für alle $U \subseteq D$ gilt $f(U) \subseteq Z$ und für alle $V \subseteq Z$ gilt $f^{-1}(V) \subseteq D$

Proposition 1.3.7

Es seien D und Z zwei Mengen und $f : D \rightarrow Z$ gegeben.

- (i) Ist U eine Teilmenge von D , so gilt $U \subseteq f^{-1}(f(U))$.
- (ii) Ist V eine Teilmenge von Z , so gilt $f(f^{-1}(V)) \subseteq V$.

Beweis. Zu (i). Zu beweisen ist $x \in U \implies x \in f^{-1}(f(U))$. Dazu:

$$\begin{array}{ccc}
 & x \in U & \\
 \xRightarrow{\text{Def.1.3.6(i)}} & f(x) \in f(U) & \\
 \xLeftrightarrow{\text{Def.1.3.6(ii) mit } V=f(U)} & x \in f^{-1}(f(U)). &
 \end{array}$$

Zu (ii). Zu beweisen ist $y \in f(f^{-1}(V)) \implies y \in V$. Dazu:

$$\begin{array}{ccc}
 & y \in f(f^{-1}(V)) & \\
 \xLeftrightarrow{\text{Def.1.3.6(i)}} & \exists x \in f^{-1}(V) : f(x) = y. & \\
 \xLeftrightarrow{\text{Def.1.3.6(ii)}} & \exists x \in D \text{ mit } f(x) \in V : f(x) = y. & \\
 \xRightarrow{y=f(x) \in V} & y \in V. &
 \end{array}$$

Bemerkung 1.3.8

Bei beiden Teilaussagen der obigen Proposition kann das $\subseteq$ -Zeichen nicht durch ein Mengengleichheitszeichen ersetzt werden! Dazu betrachte zum Beispiel $f : \{1, 2\} \rightarrow \{1, 2\}, 1 \mapsto 1, 2 \mapsto 1$.

- (i) Für $U := \{1\}$ ist $f^{-1}(f(U)) = f^{-1}(\{1\}) = \{1, 2\}$. Somit gilt $U \subsetneq f^{-1}(f(U))$.
- (ii) Für $V := \{2\}$ ist $f(f^{-1}(V)) \subsetneq V$, denn $f(f^{-1}(V)) = f(\emptyset) = \emptyset \subsetneq V$.

Wir werden nun die Begriffe *Surjektivität*, *Injektivität* und *Bijektivität* definieren. Diese werden hier, aber auch in der linearen Algebra noch eine Rolle spielen. 20.10.2025

Definition 1.3.9 — SURJEKTIVITÄT/INJEKTIVITÄT/BIJEKTIVITÄT.

Es seien D und Z zwei Mengen und $f : D \rightarrow Z$ eine Funktion.

- (i) f heißt *surjektiv* (von D nach Z), falls es zu jedem $y \in Z$ mindestens ein $x \in D$ gibt mit $f(x) = y$. Kürzer: $\forall y \in Z \exists x \in D : f(x) = y$.
- (ii) f heißt *injektiv* (von D nach Z), falls es zu jedem $y \in Z$ höchstens ein $x \in D$ gibt mit $f(x) = y$. Genauer: $\forall y \in Z : \neg(\exists x, x' \in D : x \neq x' \wedge f(x) = f(x') = y)$. ^(a)
- (iii) f heißt *bijektiv* (von D nach Z) oder *1zu1*, falls es zu jedem $y \in Z$ genau ein $x \in D$ gibt mit $f(x) = y$. Kürzer: $\forall y \in Z \exists! x \in D : f(x) = y$.

Sind D und Z aus dem Kontext ersichtlich, so schreibt man einfach nur *surjektiv*/*injektiv*/*bijektiv*.

^aDer Ausdruck „ $x, x' \in D$ “ meint nichts anderes als „ $x \in D$ sowie $x' \in D$ “. Bitte nicht verwechseln mit dem geordneten Paar (x, x') , vgl. Definition 1.2.9. Bei einem geordneten Paar sind aber auch Klammern drumherum, daher ist eine Verwechslung eigentlich ausgeschlossen.

Die nachfolgende Abbildung veranschaulicht das Konzept der Surjektivität und Injektivität.

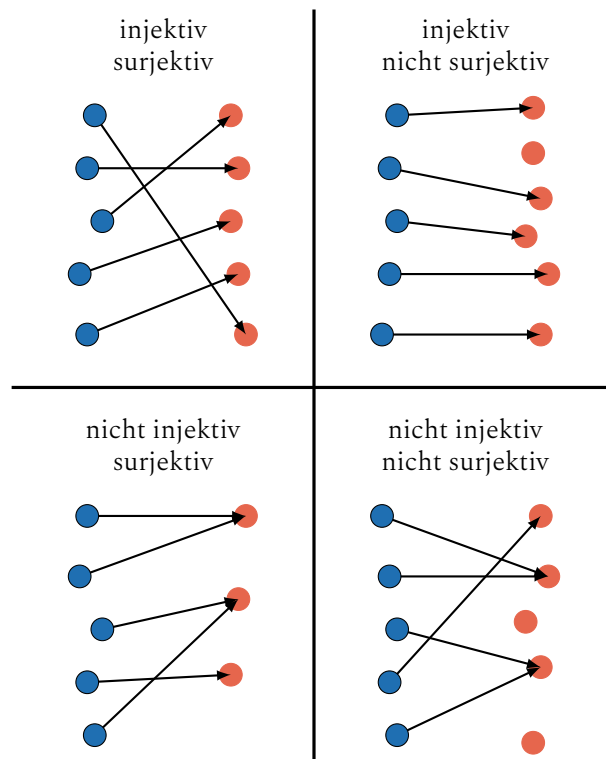


Abbildung 3: Surjektivität und Injektivität anschaulich

Wir besprechen nun eine alternative, logisch besser überprüfbare Charakterisierung der oben genannten Begriffe

Proposition 1.3.10 — ALTERNATIVE CHARAKTERISIERUNGEN FÜR SURJ., INJ., BIJ.

Es seien D, Z Mengen und $f : D \rightarrow Z$ eine gegebene Funktion.

(i) f ist surjektiv genau dann wenn $f(D) = Z$.

(ii) f ist injektiv genau dann wenn

$$x, x' \in D : f(x) = f(x') \implies x = x'. \quad (26)$$

(iii) f ist bijektiv genau dann wenn f surjektiv und injektiv ist.

Im folgenden Beweis verwenden wir die Abkürzung „Z“. Sie steht für „Zu Zeigen“.

Beweis. Zu (i). Wegen des Schlüsselwortes „genau dann wenn“ müssen wir die Äquivalenz der beiden Aussagen zeigen. Es muss dazu „ $\implies$ “ und „ $\impliedby$ “ gezeigt werden. Zu „ $\implies$ “. Sei f surjektiv. Z : $f(D) = Z$. Dass $f(D) \subseteq Z$ ist, folgt aus Definition 1.3.6. Nun zeigen wir noch $Z \subseteq f(D)$.

$$\begin{aligned} & y \in Z \\ \xRightarrow{\text{Def 1.3.9(i)}} & \exists x \in D : f(x) = y \\ \xRightarrow{\text{Def 1.3.6(i)}} & y \in f(D). \end{aligned}$$

Es gilt also $f(D) = Z$. Nun zu „ $\impliedby$ “. Es gelte $f(D) = Z$. Z : f ist surjektiv. Dazu:

$$\begin{aligned} & y \in Z \\ \xRightarrow{\text{Vorr. } f(D) = Z} & y \in f(D). \\ \xRightarrow{\text{Def 1.3.6(i)}} & y \in \{f(x) \mid x \in D\} \\ \xRightarrow{\text{Bem 1.2.2}} & \exists x \in D : y = f(x) \end{aligned}$$

Wir haben gezeigt, dass es für alle $y \in Z$ ein $x \in D$ gibt mit $y = f(x)$. Dies ist gerade die Surjektivität.

Zu Aussage (ii). Die Aussage

$$x, x' \in D : f(x) = f(x') \implies x = x'$$

ist nach Proposition 1.1.6 äquivalent zu

$$\neg(x, x' \in D : f(x) = f(x') \wedge x \neq x').$$

Die Äquivalenz zur Injektivität in Definition 1.3.9 überlegt man sich leicht. Aussage (iii) ergibt sich sofort aus Definition 1.3.9.

Als nächstes befassen wir uns mit der *Verkettung* von Funktionen.

Definition 1.3.11 — VERKETTUNG

Es seien D, M, N, Z Mengen und $f : D \rightarrow M$ sowie $g : N \rightarrow Z$ Funktionen. Ist $f(D) \subseteq N$, so definieren wir die *Verkettung* $g \circ f : D \rightarrow Z$ durch $(g \circ f)(x) := g(f(x))$ (für alle $x \in D$).

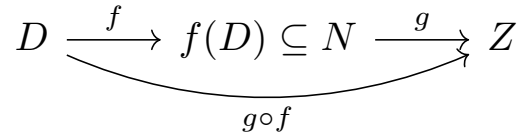


Abbildung 4: Die Verkettung $g \circ f$ in einem Diagramm.

Mit anderen Worten: Man führt die Funktionsvorschriften von g und f *hintereinander aus*. Die Reihenfolge ist bei dieser Hintereinanderausführung enorm wichtig. Außerdem muss gewährleistet sein, dass für alle $x \in D$ der Wert $f(x)$ auch im Definitionsbereich von g liegt. Daher kommt die Bedingung $f(D) \subseteq N$.

Beispiel 1.3.12 — VERKETTUNG MIT id

Sei $f : D \rightarrow Z$ gegeben. Man überzeugt sich leicht davon, dass $\text{id}_Z \circ f = f$ und $f \circ \text{id}_D = f$.

Bemerkung 1.3.13 — ASSOZIATIVITÄT DER VERKETTUNG

Sind $f_1 : M_1 \rightarrow N_1$, $f_2 : M_2 \rightarrow N_2$ und $f_3 : M_3 \rightarrow N_3$ Funktionen mit $f_3(M_3) \subseteq M_2$ und $f_2(M_2) \subseteq M_1$ so kann man $(f_1 \circ f_2) \circ f_3 : M_3 \rightarrow N_1$ bilden und $f_1 \circ (f_2 \circ f_3) : M_3 \rightarrow N_1$. Diese beiden Funktionen sind gleich! In der Tat: Für $x \in M_3$ gilt

$$((f_1 \circ f_2) \circ f_3)(x) \stackrel{\text{Def 1.3.11}}{=} (f_1 \circ f_2)(f_3(x)) \stackrel{\text{Def 1.3.11}}{=} f_1(f_2(f_3(x)))$$

und

$$(f_1 \circ (f_2 \circ f_3))(x) \stackrel{\text{Def 1.3.11}}{=} f_1((f_2 \circ f_3)(x)) \stackrel{\text{Def 1.3.11}}{=} f_1(f_2(f_3(x))).$$

Dies beweist, dass $((f_1 \circ f_2) \circ f_3)(x) = (f_1 \circ (f_2 \circ f_3))(x)$ für alle $x \in M_3$. Mit anderen Worten: $(f_1 \circ f_2) \circ f_3 = f_1 \circ (f_2 \circ f_3)$. Aufgrunddessen lassen wir die Klammern oft weg und schreiben $f_1 \circ f_2 \circ f_3$.

Wir werden uns jetzt damit beschäftigen, ob man Zuordnungen durch Funktionen *rückgängig machen*, d.h. *umkehren* kann.

Definition 1.3.14 — UMKEHRFUNKTION.

Es seien D und Z zwei Mengen und $f : D \rightarrow Z$ gegeben. Eine Funktion $g : Z \rightarrow D$ heißt *Umkehrfunktion* von f , falls $g \circ f = \text{id}_D$ und $f \circ g = \text{id}_Z$.

Bemerkung 1.3.15 — EINDEUTIGKEIT DER UMKEHRFUNKTION

Zu einer gegebenen Funktion $f : D \rightarrow Z$ existiert höchstens eine Umkehrfunktion. In der Tat: sind zwei Funktionen $g_1 : Z \rightarrow D$ und $g_2 : Z \rightarrow D$ Umkehrfunktionen von f , so gilt

$$g_1 \stackrel{\text{Bsp 1.3.12}}{=} g_1 \circ \text{id}_Z \stackrel{\text{Def 1.3.14}}{=} g_1 \circ (f \circ g_2) \stackrel{\text{Bem 1.3.13}}{=} (g_1 \circ f) \circ g_2 \stackrel{\text{Def 1.3.14}}{=} \text{id}_D \circ g_2 \stackrel{\text{Bsp 1.3.12}}{=} g_2.$$

Somit sind also alle Umkehrfunktionen gleich. Man folgert, dass es nur höchstens(!)^a eine Umkehrfunktion geben kann.

^aMan beachte: Es ist immer noch möglich, dass eine Funktion $f : D \rightarrow Z$ einfach gar keine Umkehrfunktion besitzt. Diese Bemerkung zeigt also eigentlich, dass es entweder keine oder eine einzigen Umkehrfunktion gibt.

Notation 1.3.16

Es sei eine Funktion $f : D \rightarrow Z$ gegeben. Besitzt f eine Umkehrfunktion, so notieren wir diese häufig mit $f^{-1} : Z \rightarrow D$.

Die nachfolgende Proposition charakterisiert die Funktionswerte einer Umkehrfunktion als *Lösung einer Gleichung*.

Proposition 1.3.17

Es habe $f : D \rightarrow Z$ eine Umkehrfunktion $f^{-1} : Z \rightarrow D$ und es seien $x \in D$ und $y \in Z$. Dann gilt

$$f(x) = y \quad \Longleftrightarrow \quad x = f^{-1}(y).$$

Beweis. Zu „ $\implies$ “. Man hat

$$\begin{aligned} & f(x) = y \\ \implies & f^{-1}(f(x)) = f^{-1}(y) \\ \stackrel{\text{Def 1.3.11}}{\implies} & (f^{-1} \circ f)(x) = f^{-1}(y) \\ \stackrel{\text{Def 1.3.14}}{\implies} & \text{id}_D(x) = f^{-1}(y) \\ \stackrel{\text{Def 1.3.5}}{\implies} & x = f^{-1}(y). \end{aligned}$$

Zu „ $\impliedby$ “. Man hat

$$x = f^{-1}(y)$$

$$\begin{aligned}
&\Longrightarrow f(x) = f(f^{-1}(y)) \\
&\xrightarrow{\text{Defl. 3.11}} f(x) = (f \circ f^{-1})(y) \\
&\xrightarrow{\text{Defl. 3.14}} f(x) = \text{id}_Z(y) \\
&\xrightarrow{\text{Defl. 3.5}} f(x) = y.
\end{aligned}$$

Wir fragen uns nun, unter welchen Umständen denn überhaupt eine Umkehrfunktion existiert.

Satz 1.3.18 — UMKEHRFUNKTIONEN UND BIJEKTIVITÄT

Es seien D und Z zwei Mengen und $f : D \rightarrow Z$ eine Funktion. Dann sind folgende Aussagen äquivalent:

1. f ist bijektiv.
2. f besitzt eine Umkehrfunktion.

Beweis. Wir müssen die Äquivalenz zweier Aussagen beweisen, d.h. es muss „1. $\implies$ 2.“ und „2. $\implies$ 1.“ gezeigt werden, vgl. Definition 1.1.2 (v).

Zu „1. $\implies$ 2.“. Es sei f bijektiv. Unsere Aufgabe ist es nun, eine Umkehrfunktion zu definieren (und zu zeigen, dass es sich tatsächlich um eine Umkehrfunktion handelt).

$$\begin{aligned}
&f \text{ ist bijektiv} \\
&\xrightarrow{\text{Defl. 3.9(iii)}} \forall y \in Z \quad \exists! x_y \in D : f(x_y) = y.
\end{aligned} \tag{34}$$

Beachte: Durch die Benennung „ x_y “ mit *Index* y haben wir explizit herausgestellt, dass x_y im Allgemeinen vom Wert $y \in Z$ abhängt. Definiere nun eine Funktion g durch die Vorschrift

$$g : Z \rightarrow D \quad g : y \mapsto x_y. \tag{35}$$

Behauptung. g ist Umkehrfunktion von f , d.h. $f \circ g = \text{id}_Z$ und $g \circ f = \text{id}_D$ (vgl. Definition 1.3.14). Zuerst zu „ $f \circ g = \text{id}_Z$ “. Sei $y \in Z$ beliebig. Dann gilt

$$(f \circ g)(y) \xrightarrow{\text{Def 1.3.11}} f(g(y)) \xrightarrow{(35)} f(x_y) \xrightarrow{(34)} y \xrightarrow{\text{Def 1.3.5}} \text{id}_Z(y).$$

Nun zu „ $g \circ f = \text{id}_D$ “. Sei $a \in D$ beliebig. Dann gilt

$$(g \circ f)(a) \xrightarrow{\text{Def 1.3.11}} g(f(a)) \xrightarrow{(35)} x_{f(a)}. \tag{36}$$

Wir werden nun sehen, dass $x_{f(a)} = a$. Mit (34) angewendet auf $y = f(a)$ gilt zunächst

$$f(x_{f(a)}) = f(a).$$

Beachte: f ist nach Proposition 1.3.10 (iii) auch injektiv. Nach Proposition 1.3.10 (ii) folgt

$$x_{f(a)} = a.$$

Somit liefert (36)

$$(g \circ f)(a) = x_{f(a)} = a = \text{id}_D(a) \quad \forall a \in D.$$

Man hat also $g \circ f = \text{id}_D$.

Nun zu “2. $\implies$ 1.” Es habe $f : D \rightarrow Z$ eine Umkehrfunktion, die wir nun mit $f^{-1} : Z \rightarrow D$ bezeichnen. Zu zeigen ist, dass f dann bijektiv ist. Nach Proposition 1.3.10 genügt es zu zeigen, dass f surjektiv und injektiv ist. Zuerst zur Surjektivität: Es sei $y \in Z$ beliebig. Wir müssen zeigen, dass es ein $x \in D$ gibt mit $f(x) = y$ (vgl. Definition 1.3.9). Hierzu probieren wir es mal mit $x = f^{-1}(y)$. In der Tat gilt für diese Wahl von x

$$f(x) = f(f^{-1}(y)) = (f \circ f^{-1})(y) = \text{id}_Z(y) = y.$$

Als nächstes zur Injektivität: Hierfür verwenden wir Proposition 1.3.10 (ii). Seien $x, x' \in D$ mit $f(x) = f(x')$. Wir müssen zeigen dass dann $x = x'$. Dazu

$$\begin{aligned} f(x) &= f(x') \\ \implies f^{-1}(f(x)) &= f^{-1}(f(x')) \\ \implies (f^{-1} \circ f)(x) &= (f^{-1} \circ f)(x') \\ \implies \text{id}_D(x) &= \text{id}_D(x') \\ \implies x &= x'. \end{aligned}$$

Die Behauptung folgt.

Eine spannende Konsequenz hieraus ist, dass die *Verkettung* zweier bijektiver Abbildungen wieder bijektiv ist

Proposition 1.3.19

Seien M_1, N_1, M_2, N_2 Mengen. Es seien $f_1 : M_1 \rightarrow N_1$ und $f_2 : M_2 \rightarrow N_2$ zwei bijektive Abbildungen mit $f_2(M_2) = N_1$. Dann ist $f_1 \circ f_2 : M_2 \rightarrow N_1$ bijektiv und es gilt $(f_1 \circ f_2)^{-1} = f_2^{-1} \circ f_1^{-1}$.

Beweis. Seien f_1, f_2 wie in der Aussage. Wegen der Surjektivität von f_2 (vgl. Definition 1.3.9) gilt

$$f_2(M_2) = N_2.$$

Mit der Voraussetzung $f_2(M_2) = N_1$ folgt

$$M_1 = N_2. \tag{37}$$

Nach Satz 1.3.18 gibt es eine Umkehrfunktionen $f_1^{-1} : N_1 \rightarrow M_1$ und $f_2^{-1} : N_2 \rightarrow M_2$.

Beachte: Es gilt

$$f_1^{-1}(N_1) \subseteq M_1 \underset{(37)}{=} N_2$$

Definiere nun

$$h := f_2^{-1} \circ f_1^{-1} : N_1 \rightarrow M_2 \quad (38)$$

Da $f_1^{-1}(N_1) \subseteq N_2$ ist h wohldefiniert (d.h. man *darf* die Verkettung so definieren, vgl. Definition 1.3.11). Dann gilt

$$\begin{aligned} h \circ (f_1 \circ f_2) &\overset{\text{Bem1.3.13}}{=} (h \circ f_1) \circ f_2 \\ &\overset{(38)}{=} ((f_2^{-1} \circ f_1^{-1}) \circ f_1) \circ f_2 \\ &\overset{\text{Bem1.3.13}}{=} (f_2^{-1} \circ (f_1^{-1} \circ f_1)) \circ f_2 \\ &\overset{\text{Def1.3.14}}{=} (f_2^{-1} \circ \text{id}_{M_1}) \circ f_2 \\ &\overset{(37)}{=} (f_2^{-1} \circ \text{id}_{N_2}) \circ f_2 \\ &\overset{\text{Bsp1.3.12}}{=} f_2^{-1} \circ f_2 \overset{\text{Def1.3.14}}{=} \text{id}_{M_2}. \end{aligned}$$

Man kann völlig analog vorgehen um zu zeigen, dass $(f_1 \circ f_2) \circ h = \text{id}_{N_1}$. Wir haben bewiesen, dass h eine Umkehrfunktion von $f_1 \circ f_2$ ist. Nach Satz 1.3.18 ist $f_1 \circ f_2$ dann also bijektiv und außerdem gilt $(f_1 \circ f_2)^{-1} = h = f_2^{-1} \circ f_1^{-1}$.

Da die Reihenfolge bei der Verkettung sehr wichtig ist, ist es auch wichtig, zu beachten, dass bei der Bildung der Umkehrfunktion in Proposition 1.3.19 die Reihenfolge verändert wird.

2 Vorkurs: Folgen

In diesem Kapitel werden wir für einen kurzen Moment die formale Strenge der Mathematik verlassen. Wir werden hier nämlich ohne jegliche Erklärung akzeptieren, was die *reellen Zahlen* sind. Diese sind eine *Zahlenmenge*, genannt $\mathbb{R}$, innerhalb welcher wir

- in der uns gewohnten Weise addieren/subtrahieren können (d.h. $a + b, a - b$ berechnen).
- in der uns gewohnten Weise multiplizieren/dividieren können (d.h. $a \cdot b, \frac{a}{b}$ berechnen).
- in der uns gewohnten Weise vergleichen können (d.h. die Ungleichung $a \leq b$ verifizieren können).

Was die „uns gewohnte Weise“ ist werden wir noch im nächsten Kapitel hinterfragen. Wir werden hier einfach wie aus der Schule bekannt damit arbeiten. Oft visualisiert man die reellen Zahlen anhand einer *Zahlengerade*.

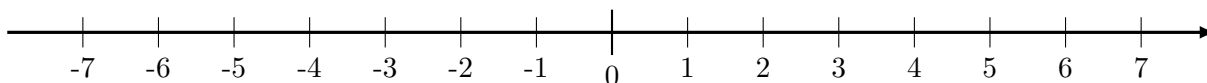


Abbildung 5: Zahlengerade

Mehr werden wir dazu an dieser Stelle nicht sagen. Wir werden übrigens auch die *natürlichen Zahlen* $\mathbb{N}$, die *ganzen Zahlen* $\mathbb{Z}$ und die *rationalen Zahlen* $\mathbb{Q}$ in diesem Kapitel ohne weitere Erklärung verwenden.

Zu zwei Konzepten müssen wir uns allerdings noch äußern, bevor es losgeht: Dem *Absolutbetrag* und dem Begriff der *(Un-)endlichkeit*.

Der Absolutbetrag

Definition 2.0.1 — ABSOLUTBETRAG

Es sei $x \in \mathbb{R}$. Dann definieren wir

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

und nennen $|x|$ den *(Absolut-)Betrag* von x .

23.10.2025

Beachte: In der Theorie könnte man beim zweiten Fall auch „ $x \leq 0$ “ anstatt „ $x < 0$ “ fordern. Dann hätte man den Fall $x = 0$ doppelt aufgeführt. Das macht aber nichts, da $-0 = 0$. Deswegen liefern beide Fälle im Fall $x = 0$ dasselbe Ergebnis.

Bemerkung 2.0.2 — BETRAG ALS MAXIMUM.

Für zwei Elemente $a, b \in \mathbb{R}$ definieren wir

$$\max\{a, b\} := \begin{cases} a & \text{falls } a \geq b \\ b & \text{sonst} \end{cases}.$$

Man kann sich leicht überlegen, dass dann

$$|x| = \max\{x, -x\} \quad \forall x \in \mathbb{R}. \quad (39)$$

Bevor wir loslegen wollen, halten wir noch einige Eigenschaften des Absolutbetrages fest.

Satz 2.0.3 — EIGENSCHAFTEN DES ABSOLUTBETRAGES

- (i) *Definitheit*: Für alle $x \in \mathbb{R}$ gilt $|x| \geq 0$ und $(|x| = 0 \iff x = 0)$.
- (ii) *Multiplikativität*: Für alle $x, y \in \mathbb{R}$ gilt $|x \cdot y| = |x| \cdot |y|$.
- (iii) *Dreiecksungleichung*: Für alle $x, y \in \mathbb{R}$ gilt $|x + y| \leq |x| + |y|$.

Beweis. Zu (i). Wir zeigen zuerst, dass $|x| \geq 0$ für alle $x \in \mathbb{R}$. Dies machen wir durch eine *Fallunterscheidung*.

$$\text{FALL 1. } x \geq 0. \quad x \geq 0 \xRightarrow{\text{Def 2.0.1}} |x| = x \geq 0.$$

$$\text{FALL 2. } x < 0. \quad x < 0 \implies -x > 0 \xRightarrow{\text{Def 2.0.1}} |x| = -x \geq 0.$$

Den Rest der Aussage $(|x| = 0 \iff x = 0)$ überlassen wir den Lesenden als Übungsaufgabe.

Zu (ii). Wiederum eine Fallunterscheidung

$$\text{FALL 1. } (x \geq 0) \wedge (y \geq 0). \quad (x \geq 0) \wedge (y \geq 0) \implies x \cdot y \geq 0.$$

$$\text{Mit Def 2.0.1: } |x \cdot y| = x \cdot y, \quad |x| \cdot |y| = x \cdot y, \text{ somit } |x \cdot y| = |x| \cdot |y|.$$

$$\text{FALL 2. } (x \geq 0) \wedge (y < 0). \quad (x \geq 0) \wedge (y < 0) \implies x \cdot y \leq 0.$$

$$\text{Mit Def 2.0.1: } |x \cdot y| = -x \cdot y, \quad |x| \cdot |y| = x \cdot (-y) = -x \cdot y, \text{ somit } |x \cdot y| = |x| \cdot |y|.$$

$$\text{FALL 3. } (x < 0) \wedge (y \geq 0). \quad (x < 0) \wedge (y \geq 0) \implies x \cdot y \leq 0.$$

$$\text{Mit Def 2.0.1: } |x \cdot y| = -x \cdot y, \quad |x| \cdot |y| = (-x) \cdot y = -x \cdot y, \text{ somit } |x \cdot y| = |x| \cdot |y|.$$

$$\text{FALL 4. } (x < 0) \wedge (y < 0). \quad (x < 0) \wedge (y < 0) \implies x \cdot y \geq 0.$$

$$\text{Mit Def 2.0.1: } |x \cdot y| = x \cdot y, \quad |x| \cdot |y| = (-x) \cdot (-y) = x \cdot y, \text{ somit } |x \cdot y| = |x| \cdot |y|.$$

Zu (iii). Man kann zunächst leicht aus Bemerkung 2.0.2 folgern, dass für alle $x \in \mathbb{R}$ gilt

$$x \leq |x| \text{ und } -x \leq |x|. \quad (40)$$

Nun zum Beweis der Ungleichung.

FALL 1. $x + y \geq 0$.

$$x + y \geq 0 \implies |x + y| \stackrel{\text{Def 2.0.1}}{=} x + y \underset{(40)}{\leq} |x| + y \underset{(40)}{\leq} |x| + |y|.$$

FALL 2. $x + y < 0$.

$$x + y \leq 0 \implies |x + y| \stackrel{\text{Def 2.0.1}}{=} -(x + y) = -x - y \underset{(40)}{\leq} |x| + (-y) \underset{(40)}{\leq} |x| + |y|.$$

Bemerkung 2.0.4 — FALLUNTERSCHIEDUNG

Ein Beweis via Fallunterscheidung fußt logisch auf der folgenden Beobachtung:

$$(A \implies C) \wedge (B \implies C) \text{ ist logisch äquivalent zu } A \vee B \implies C$$

In Worten: Möchte man die Aussage $D \implies C$ zeigen und lässt sich zusätzlich D in die Fälle A oder B aufteilen (d.h. $D \iff A \vee B$), so genügt es, $A \implies C$ und $B \implies C$ zu zeigen. Im konkreten Fall von Satz 2.0.3 schlüsselt man das logisch so auf: Die Voraussetzung „ $D : x \in \mathbb{R}$ “ lässt sich in die Fälle „ $A : x \geq 0$ “ und „ $B : x \leq 0$ “ aufteilen. Außerdem ist „ $C : |x| \geq 0$ “. Sie werden sehen, dass wir in der Fallunterscheidung nichts anderes gezeigt haben als $A \implies C$ („Fall 1“) und $B \implies C$ („Fall 2“).

Wir diskutieren nun zum Schluss noch weitere Eigenschaften des Absolutbetrages, die beim Rechnen hilfreich sein können.

Proposition 2.0.5 — WEITERE EIGENSCHAFTEN DES BETRAGES.

(i) Es gilt $|(-x)| = |x|$ für alle $x \in \mathbb{R}$.

(ii) Für alle $x, y \in \mathbb{R}$ mit $y \neq 0$ gilt $|\frac{x}{y}| = \frac{|x|}{|y|}$.

(iii) Für alle $x, y \in \mathbb{R}$ gilt

$$||x| - |y|| \leq |x - y|. \quad (41)$$

(iv) Für alle $x, y \in \mathbb{R}$ gilt

$$||x| - |y|| \leq |x + y|.$$

Punkt (iii) in der vorigen Proposition wird auch manchmal die *inverse Dreiecksungleichung* genannt.

Beweis. Zu (i). Sei $x \in \mathbb{R}$.

$$|(-x)| \stackrel{(39)}{=} \max\{-x, -(-x)\} = \max\{-x, x\} = \max\{x, -x\} \stackrel{(39)}{=} |x|.$$

Zu (ii). Seien $x, y \in \mathbb{R}$ mit $y \neq 0$. Dann gilt

$$|y| \cdot \left| \frac{x}{y} \right| \stackrel{\text{Prop2.0.3(ii)}}{=} \left| y \cdot \frac{x}{y} \right| = |x|.$$

Teilt man beide Seiten der vorigen Gleichung durch $|y|$ so folgt

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

Zu (iii). Wir zeigen $|x| - |y| \leq |x - y|$ und $-(|x| - |y|) \leq |x - y|$.

Zu Ersterem:

$$|x| - |y| = |x - y + y| - |y| = |(x - y) + y| - |y| \stackrel{\text{Prop2.0.3(iii)}}{\leq} |x - y| + |y| - |y| = |x - y|. \quad (43)$$

Zu Letzterem:

$$\begin{aligned} -(|x| - |y|) &= |y| - |x| = |(y - x) + x| - |x| \\ &\stackrel{\text{Prop2.0.3(iii)}}{\leq} |y - x| + |x| - |x| = |y - x| \stackrel{(i)}{=} |-(y - x)| = |x - y|. \end{aligned} \quad (44)$$

Nun gilt

$$||x| - |y|| = \max\{|x| - |y|, -(|x| - |y|)\} \stackrel{(43)}{\leq} |x - y| \stackrel{(44)}{\leq} |x - y|.$$

Punkt (iv) folgt aus Punkt (iii) mit $\tilde{x} := x$ und $\tilde{y} := -y$.

(Un-)Endlichkeit

Ein weiterer Begriff, den wir benötigen, ist der der *(Un-)Endlichkeit*. Im ersten Kapitel haben wir wenig mit unendlichen Mengen gearbeitet. Jedoch: Die Mengen, die wir seit Anfang dieses Kapitels betrachten ($\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$) sind (wie sich schnell herausstellen wird) alle unendlich.

Definition 2.0.6 — (UN-)ENDLICHKEIT

- Es sei M eine Menge. Wir sagen M ist *endlich*, falls entweder $M = \emptyset$ oder falls es ein $n \in \mathbb{N}$ gibt derart, dass man eine bijektive Funktion $f : \{1, \dots, n\} \rightarrow M$ finden kann.
- Eine Menge, die nicht endlich ist, nennen wir *unendlich*.

Beispiel 2.0.7 — BEISPIELE FÜR (UN-)ENDLICHE MENGEN

- (i) Die Menge
- $M := \{\text{Sonne, Mond, Sterne}\}$
- ist endlich.

Dazu: Man prüft leicht nach, dass $f : \{1, 2, 3\} \rightarrow M$, $f : 1 \mapsto \text{Sonne}$, $2 \mapsto \text{Mond}$, $3 \mapsto \text{Sterne}$, eine bijektive Abbildung ist. Definition 2.0.6 trifft also für $n = 3$ zu.

- (ii) Klarerweise ist für alle
- $n \in \mathbb{N}$
- die Menge
- $\{1, \dots, n\}$
- eine endliche Menge [denn
- $f = \text{id}_{\{1, \dots, n\}}$
- erfüllt die Bedingung in Definition 2.0.6]

- (iii) Die Menge
- $\mathbb{N}$
- ist unendlich.

Dazu muss gezeigt werden dass die Aussage „ $\mathbb{N}$ ist endlich“ falsch ist. Hierzu führen wir einen sog. *Widerspruchsbeweis*.

Angenommen, $\mathbb{N}$ wäre endlich. Dann gäbe es ein $n \in \mathbb{N}$ so, dass eine bijektive Abbildung $f : \{1, \dots, n\} \rightarrow \mathbb{N}$ existiert. Betrachte nun die natürliche Zahl

$$n^* := f(1) + \dots + f(n) + 1 \in \mathbb{N}. \quad (45)$$

Da f von $\{1, \dots, n\}$ nach $\mathbb{N}$ surjektiv ist und $n^* \in \mathbb{N}$ muss $n^* \in f(\{1, \dots, n\})$ gelten. Das heißt es gibt ein $m \in \{1, \dots, n\}$ so, dass $f(m) = n^*$. Man erhält

$$\begin{aligned} f(m) = n^* &\stackrel{(45)}{=} f(1) + \dots + f(n) + 1 \\ &= f(1) + \dots + f(m-1) + f(m) + f(m+1) + \dots + f(n) + 1. \end{aligned}$$

Der Wert $f(m)$ kommt als Summand auf beiden Seiten vor. Ziehen wir $f(m)$ auf beiden Seiten ab, so erhalten wir

$$0 = f(1) + \dots + f(m-1) + f(m+1) + \dots + f(n) + 1.$$

Alle Summanden auf der rechten Seite der obigen Gleichung sind natürliche Zahlen. Die Summe von natürlichen Zahlen kann aber niemals 0 ergeben!

Unsere oben getroffene *Annahme* erzeugt also einen mathematischen Widerspruch. Sie kann daher keine wahre Aussage sein.^a Somit ist die Annahme („ $\mathbb{N}$ ist endlich“) falsch. Das ist gerade das, was wir zeigen wollten.

^aDenn wäre sie wahr, so wäre alles, was man aus ihr folgern kann auch wahr! (vgl. Kapitel 1.1)

Proposition 2.0.8 — ENDLICHKEIT UNTER BIJEKTIVEN ABBILDUNGEN.

Es seien M_1, M_2 zwei Mengen derart, dass es eine bijektive Abbildung $g : M_1 \rightarrow M_2$ gibt. Dann gilt

$$M_1 \text{ endlich} \iff M_2 \text{ endlich}.$$

Beweis. Wir führen den Beweis im Spezialfall, dass $M_1, M_2 \neq \emptyset$. Der Fall, dass eine der beiden Mengen leer ist, ist den Lesenden als Übungsaufgabe überlassen. Zu „ $\implies$ “. Es sei

$g : M_1 \rightarrow M_2$ wie in der Voraussetzung bijektiv (von M_1 nach M_2). Nun gilt

$$M_1 \text{ endlich} \xrightarrow[\text{Def 2.0.6}]{=} \exists n \in \mathbb{N} : \exists f : \{1, \dots, n\} \rightarrow M_1 \text{ bijektiv.}$$

Definiere nun

$$h : \{1, \dots, n\} \rightarrow M_2, \quad h := g \circ f.$$

Diese Verkettung ist zulässig (vgl. Definition 1.3.11), da (wegen der Surjektivität von f)

$$f(\{1, \dots, n\}) = M_1 \quad (\text{was mit dem Definitionsbereich von } g \text{ übereinstimmt}). \quad (46)$$

Wegen (46) und Proposition 1.3.19 ist auch $h = g \circ f$ bijektiv. Somit existiert eine bijektive Abbildung von $\{1, \dots, n\}$ nach M_2 . Das zeigt gerade die Endlichkeit von M_2 .

Für „ $\Leftarrow$ “ können wir einfach den obigen Beweis abschreiben und die Rollen von M_1 und M_2 vertauschen. Details sind den Lesenden als Übungsaufgabe überlassen.

Beispiel 2.0.9 — WEITERE UNENDLICHE MENGEN

Die Menge aller geraden Zahlen

$$G := \{2m : m \in \mathbb{N}\}$$

ist unendlich. Dazu: Man überzeugt sich leicht, dass $g : \mathbb{N} \rightarrow G, g : m \mapsto 2m$ eine bijektive Abbildung ist. Wäre G endlich so wäre nach Proposition 2.0.8 auch $\mathbb{N}$ endlich. In Beispiel 2.0.7 (ii) haben wir aber das Gegenteil bewiesen. Somit ist G unendlich.

Analog kann man sich davon überzeugen, dass auch die Menge aller ungeraden Zahlen

$$U := \{2m - 1 : m \in \mathbb{N}\}$$

unendlich ist.

Endlichkeit ist außerdem *stabil unter Teilmengenbildung*, d.h. Teilmengen endlicher Mengen sind wiederum endlich.

Proposition 2.0.10 — TEILMENGEN ENDLICHER MENGEN SIND ENDLICH

Es sei M eine endliche Menge und $N \subseteq M$. Dann ist N endlich.

Beweis. Klar. (Wobei man hier fairerweise dazu sagen sollte dass ein formaler Beweis gar nicht so leicht aufzuschreiben ist)

Hat man eine endliche Menge A gegeben, so würde man normalerweise eine bijektive Funktion $f : \{1, \dots, n\} \rightarrow A$ wählen und schreiben

$$A = f(\{1, \dots, n\}) = \{f(1), \dots, f(n)\}.$$

Alternativ werden wir für eine endliche Menge auch häufig die Schreibweise

$$A = \{a_1, \dots, a_n\}$$

benutzen. Dahinter steckt dann die Wahl einer bijektiven Funktion $f : \{1, \dots, n\} \rightarrow A$ mit $f(i) := a_i$.

Bonusmaterial: HILBERT'S HOTEL.

Stellen Sie sich mal ein Hotel vor. In den üblichen Hotels, die wir auf der Erde finden, haben nur *endlich* viele Gäste Platz. Nicht so in *Hilbert's Hotel*! Dort gibt es für jede natürliche Zahl $n \in \mathbb{N}$ ein Zimmer. Nennen wir es Zimmer n .

Wenn bei einem üblichen Hotel alle Zimmer belegt sind, kann kein neuer Gast mehr aufgenommen werden. Das ist unglücklich: Dem Hotel entgehen finanzielle Profite und die Gäste müssen sich eine andere Bleibe suchen.

In Hilbert's Hotel ist das alles gar kein Problem: Sagen wir mal, alle Zimmer sind belegt. Mit anderen Worten:

Zimmer n ist belegt für alle $n \in \mathbb{N}$.

Ein neuer Gast. Wenn jetzt doch ein neuer Gast an der Rezeption ankommt und einchecken will, geschieht Folgendes: Man bittet jeden aktuellen Gast, in das Zimmer mit der nächsthöheren Nummer umzuziehen. Also:

Der Gast aus Zimmer n zieht in Zimmer $n + 1$ um.

Nachdem dieser Massenumzug vollzogen ist wird plötzlich Zimmer 1 frei! Der neue Gast kann in Zimmer 1 aufgenommen werden.

Ein Reisebus mit k Gästen ($k \in \mathbb{N}$). Stellen Sie sich jetzt mal vor es fährt ein Reisebus mit k Gästen vor und alle davon wollen einchecken. Können die Gäste aus dem Reisebus aufgenommen werden? Ja! Wiederholen Sie die obige Prozedur einfach k -mal, bis alle Gäste untergebracht sind. Mit anderen Worten:

Der Gast aus Zimmer n zieht in Zimmer $n + k$ um.

Die Zimmer $1, 2, \dots, k$ werden frei und unser Reisebus kommt unter!

Hilbert's Reisebus. Da sich Hilbert's Hotel zu rentieren scheint, greift nun auch ein Busunternehmen die Idee auf. In *Hilbert's Reisebussen* gibt es für jede natürliche Zahl $m \in \mathbb{N}$ einen Sitzplatz. Nennen wir ihn Sitzplatz m . Stellen Sie sich nun mal vor, Hilbert's Reisebus fährt vor Hilbert's Hotel vor und alle Busreisenden wollen ein Zimmer. Für die geschulte Rezeption in Hilbert's Hotel kein Problem: Sie bittet zunächst jeden aktuellen Gast, seine Zimmernummer zu verdoppeln. Sprich

Der Gast aus Zimmer n zieht in Zimmer $2n$ um.

Alle ungeraden Zimmernummern werden dadurch frei. Die clevere Rezeption gibt nun dem Busreisenden mit Sitzplatznummer m die Zimmernummer $2m - 1$. Da $2m - 1$ ungerade ist, ist dieses Zimmer stets frei!

Hilbert's Reisebüro. Jetzt treiben wir es auf die Spitze: Für eine Gruppenreise chartert *Hilbert's Reisebüro* für jede natürliche Zahl $j \in \mathbb{N}$ einen von Hilbert's Reisebussen. Nennen wir diesen Bus Nummer j . Alle diese Reisebusse funktionieren wie oben: Für jedes $m \in \mathbb{N}$ gibt es einen Sitzplatz! Stellt Sie sich nun vor, wie alle diese unendlich vielen Reisebusse mit jeweils unendlich vielen Gästen vor Hilbert's Hotel vorfahren! Die Rezeption ist von diesem Massentourismus recht unbeeindruckt und bekommt mit wenig Anstrengung alle unter. *Wie?* Eine Antwortmöglichkeit finden Sie in Proposition 2.4.6. Gedulden Sie sich bis dahin!

2.1 Definition einer Folge

Wir definieren nun den Begriff einer *Folge*, welchen wir in diesem Kapitel studieren möchten.

Definition 2.1.1 — FOLGE

Es sei M eine Menge. Eine *Folge in M* ist eine Abbildung $a : \mathbb{N} \rightarrow M$.

Anstatt $a(n)$ schreibt man auch häufig a_n . Eine solche Folge wird auch mit $(a_n)_{n \in \mathbb{N}}$ notiert. Der Wert $a_n \in M$ heißt dann *n -tes Folgenglied* von $(a_n)_{n \in \mathbb{N}}$.

Ist $M = \mathbb{R}$ so spricht man auch von einer *reellen Folge*.

Ist $X \subseteq M$ eine Teilmenge und gilt $a_n \in X$ für alle $n \in \mathbb{N}$ so schreibt man (salopp) $(a_n)_{n \in \mathbb{N}} \subseteq X$.

Wir stellen uns reelle Folgen als eine Ansammlung *unendlich vieler Punkte* auf der Zahlengerade vor. Diese Punkte sind *durchnummeriert* in der Reihenfolge $a_1, a_2, a_3, a_4, \dots$. Diese *unendliche Auflistung* werden wir oft anstelle von einer (unendlich großen) Wertetabelle angeben.

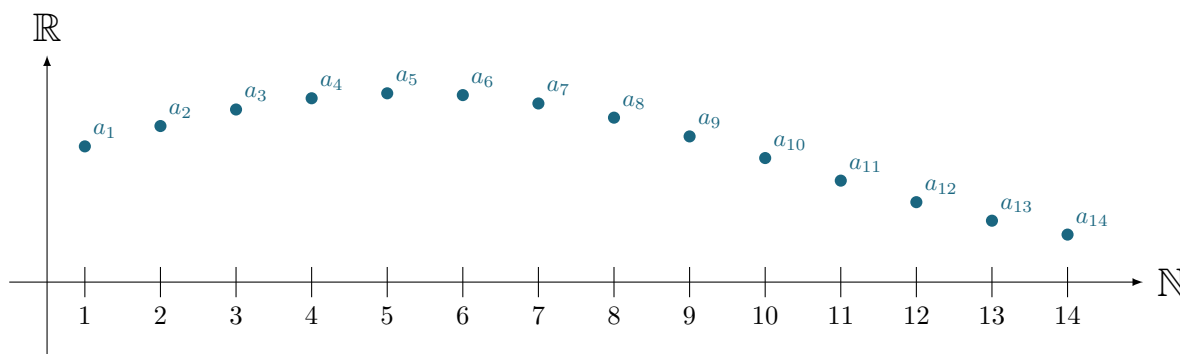


Abbildung 6: Darstellung einer reellen Folge

Beispiel 2.1.2 — REELLE FOLGEN

- (i) $(n^2)_{n \in \mathbb{N}}$, d.h. $1, 4, 9, 16, 25, \dots$
- (ii) $(2^n)_{n \in \mathbb{N}}$, d.h. $2, 4, 8, 16, 32, \dots$
- (iii) $(\frac{1}{n})_{n \in \mathbb{N}}$ d.h. $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$
- (iv) $((-1)^n)_{n \in \mathbb{N}}$, d.h. $-1, 1, -1, 1, -1, 1, \dots$
- (v) Für $c \in \mathbb{R}$ kann die *konstante Folge* $(c)_{n \in \mathbb{N}}$, d.h. $c, c, c, c, c, \dots$ betrachtet werden.
- (vii) Für $n \in \mathbb{N}$ definieren wir die *Fakultät* $n! := 1 \cdot 2 \cdot \dots \cdot n$. Dann ist auch $(n!)_{n \in \mathbb{N}}$ eine Folge. Die Werte sind $1, 2, 6, 24, 120, 720, \dots$.
Im Folgenden beschäftigen wir uns übrigens häufiger mit der Fakultät. Wir werden dabei auch die Konvention $0! := 1$ nutzen.
- (vi) $(n^n)_{n \in \mathbb{N}}$, d.h. $1, 4, 27, 256, 3125, \dots$

Wir werden uns im Folgenden verschiedenste Eigenschaften von Folgen anschauen. Die erste Eigenschaft, die wir betrachten ist die *Beschränktheit*.

Definition 2.1.3 — BESCHRÄNKTHEIT VON FOLGEN

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine Folge. Wir nennen $(a_n)_{n \in \mathbb{N}}$

1. *nach oben beschränkt*, falls

$$\exists K \in \mathbb{R} \forall n \in \mathbb{N} : a_n \leq K.$$

(In Worten: Es gibt eine Konstante $K \in \mathbb{R}$ so, dass $a_n \leq K$ für alle $n \in \mathbb{N}$).

2. *nach unten beschränkt*, falls

$$\exists k \in \mathbb{R} \forall n \in \mathbb{N} : a_n \geq k.$$

(In Worten: Es gibt eine Konstante $k \in \mathbb{R}$ so, dass $a_n \geq k$ für alle $n \in \mathbb{N}$).

3. *beschränkt*, falls $(a_n)_{n \in \mathbb{N}}$ nach oben und nach unten beschränkt ist.

Ist eine Folge nicht (nach oben/nach unten) beschränkt, so heißt sie (*nach oben/nach unten*) *unbeschränkt*.

Beispiel 2.1.4 — ZWEI BESCHRÄNKTE FOLGEN

Die Folge $((-1)^n)_{n \in \mathbb{N}}$ ist beschränkt. Sie ist nach oben beschränkt, denn

$$\forall n \in \mathbb{N} : (-1)^n \leq 1$$

und nach unten beschränkt weil

$$\forall n \in \mathbb{N} : \quad (-1)^n \geq -1.$$

Auch ist die Folge $(\frac{1}{n})_{n \in \mathbb{N}}$ beschränkt. Sie ist nach oben beschränkt weil

$$\forall n \in \mathbb{N} : \quad \frac{1}{n} \underset{n \geq 1}{\leq} \frac{1}{1} = 1,$$

und nach unten beschränkt weil

$$\forall n \in \mathbb{N} : \quad \frac{1}{n} > 0.$$

Beachte: Hier sieht man, dass man Folgen dringend von reellen Funktionen unterscheiden muss – die Funktion $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto \frac{1}{x}$ ist – wie bereits in der Schule besprochen – eine *unbeschränkte* Funktion^a. Man muss sich also wirklich dran gewöhnen, dass wir bei der Folge $(\frac{1}{n})_{n \in \mathbb{N}}$ wirklich nur natürliche Zahlen für n einsetzen (und nicht etwa beliebige reelle Zahlen).

^aWir haben hier noch gar nicht definiert was das überhaupt bedeutet

Beispiel 2.1.5 — EINE ANDERE BESCHRÄNKTE FOLGE

Etwas schwerer zu sehen: Die Folge $(\sqrt{n^2 + n} - n)_{n \in \mathbb{N}}$ ist beschränkt. Dazu formen wir den Ausdruck zunächst um

$$\begin{aligned} \sqrt{n^2 + n} - n &= (\sqrt{n^2 + n} - n)(\sqrt{n^2 + n} + n) \frac{1}{\sqrt{n^2 + n} + n} \\ &\stackrel{\text{3. Binom. Formel}}{=} [n^2 + n - n^2] \frac{1}{\sqrt{n^2 + n} + n} = \frac{n}{\sqrt{n^2 + n} + n}. \end{aligned}$$

Nun gilt für alle $n \in \mathbb{N}$

$$\frac{n}{\sqrt{n^2 + n} + n} \geq 0.$$

Daher ist die Folge nach unten durch 0 beschränkt. Außerdem gilt für alle $n \in \mathbb{N}$

$$\frac{n}{\sqrt{n^2 + n} + n} \leq \frac{n}{0 + n} \leq \frac{n}{n} = 1.$$

Deswegen ist die Folge nach oben durch 1 beschränkt.

Spannend an dem obigen Beispiel ist auch, dass es durch eine kleine Abänderung unbeschränkt wird: Die Folge $(\sqrt{n^2 + n^2} - n)_{n \in \mathbb{N}}$ ist nach oben unbeschränkt, denn

$$\sqrt{n^2 + n^2} - n = \sqrt{2n^2} - n = \sqrt{2}\sqrt{n^2} - n = \sqrt{2}n - n = (\sqrt{2} - 1)n \quad \forall n \in \mathbb{N}.$$

Überzeugen Sie sich als Übungsaufgabe davon, dass diese Folge nach oben unbeschränkt ist. Die nachfolgende Beobachtung ist nicht verwunderlich, aber trotzdem nennenswert. 27.10.2025

Proposition 2.1.6 — BESCHRÄNKTE FOLGEN

Sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine Folge. Dann gilt

$$(a_n)_{n \in \mathbb{N}} \text{ beschränkt} \iff \exists K \in \mathbb{R} \forall n \in \mathbb{N} : |a_n| \leq K.^a$$

^aBeachte: Ein solches K ist stets auch größergleich 0, da $|a_n| \geq 0$ für alle $n \in \mathbb{N}$.

Beweis. Um „ $\iff$ “ zu zeigen müssen stets beide Richtungen der Implikation nachgewiesen werden.

Zu „ $\implies$ “. Es sei $(a_n)_{n \in \mathbb{N}}$ beschränkt. Dann ist $(a_n)_{n \in \mathbb{N}}$ also nach oben und nach unten beschränkt. Aus Definition 2.1.3 folgt

$$\exists K_1 \in \mathbb{R} \forall n \in \mathbb{N} : a_n \leq K_1$$

und

$$\exists K_2 \in \mathbb{R} \forall n \in \mathbb{N} : a_n \geq K_2.$$

Beachte: $a_n \geq K_2 \iff -a_n \leq -K_2$. Wir haben also

$$\forall n \in \mathbb{N} : a_n \leq K_1 \quad \wedge \quad -a_n \leq -K_2.$$

Nun gilt

$$\forall n \in \mathbb{N} : |a_n| \stackrel{\text{Bem 2.0.2}}{=} \max\{a_n, -a_n\} \leq \max\{K_1, -K_2\}.$$

Definiert man $K := \max\{K_1, -K_2\}$ so ergibt sich

$$\forall n \in \mathbb{N} : |a_n| \leq K.$$

Zu „ $\impliedby$ “. Es gebe ein $K_0 \in \mathbb{R}$ so, dass

$$\forall n \in \mathbb{N} : |a_n| \leq K_0.$$

Da $|a_n| = \max\{a_n, -a_n\}$ folgern wir

$$\forall n \in \mathbb{N} : a_n \leq K_0 \quad \text{und} \quad \forall n \in \mathbb{N} : -a_n \leq K_0.$$

Man folgert

$$\forall n \in \mathbb{N} : a_n \leq K_0 \tag{47}$$

und

$$\forall n \in \mathbb{N} : \quad a_n \geq -K_0. \quad (48)$$

Mit Definition 2.1.3 folgt aus (47), dass $(a_n)_{n \in \mathbb{N}}$ nach oben beschränkt ist und aus (48), dass $(a_n)_{n \in \mathbb{N}}$ nach unten beschränkt ist.

Eine weitere wichtige Eigenschaft, die uns interessieren wird, ist die *Monotonie von Folgen*

Definition 2.1.7 — MONOTONIE VON FOLGEN

Eine Folge $(a_n)_{n \in \mathbb{N}}$ heißt

1. *monoton wachsend*, falls

$$\forall n \in \mathbb{N} : \quad a_{n+1} \geq a_n.$$

2. *monoton fallend*, falls

$$\forall n \in \mathbb{N} : \quad a_{n+1} \leq a_n.$$

3. *streng monoton wachsend*, falls

$$\forall n \in \mathbb{N} : \quad a_{n+1} > a_n.$$

4. *streng monoton fallend*, falls

$$\forall n \in \mathbb{N} : \quad a_{n+1} < a_n.$$

Um monotonies Wachstum einer Folge $(a_n)_{n \in \mathbb{N}}$ nachzuweisen muss also die *Abschätzung* $a_{n+1} \geq a_n$ für alle $n \in \mathbb{N}$ verifiziert werden. Manchmal kann es vorteilhaft sein, diese Abschätzung etwas umzuformulieren, etwa als

$$\forall n \in \mathbb{N} : \quad a_{n+1} - a_n \geq 0.$$

Falls $a_m > 0$ für alle $m \in \mathbb{N}$ ist eine weitere mögliche Umformulierung

$$\forall n \in \mathbb{N} : \quad \frac{a_{n+1}}{a_n} \geq 1.$$

Beispiel 2.1.8 — EINE MONOTONE FOLGE

Die Folge $(n^2)_{n \in \mathbb{N}}$ ist streng monoton wachsend. In der Tat: Definiert man $a_n := n^2$ für $n \in \mathbb{N}$ so gilt

$$a_{n+1} = (n+1)^2 \stackrel{\text{2. Binom. Formel}}{=} n^2 + 2n + 1 > n^2 + 0 + 0 = n^2 = a_n \quad \forall n \in \mathbb{N}.$$

QED

Bemerkung 2.1.9 — QED

QED steht für „quod erat demonstrandum“ – lateinisch für „was zu zeigen war“. Es signalisiert das Ende eines Beweises. Normalerweise hätten wir den Beweis jetzt vielleicht mit der Phrase „Die Behauptung folgt dann aus Definition 2.1.7“ beendet. Das fassen wir jetzt kürzer. Immer wenn Sie in Zukunft QED lesen, bitte überlegen Sie sich genau, wie aus den Ausführungen die Behauptung folgt. Moderner sieht man häufig ein kleines Kästchen $\square$, den sogenannten „Halmos“.

Beispiel 2.1.10 — EINE ANDERE MONTONE FOLGE

Die Folge $(\frac{n}{n+2})_{n \in \mathbb{N}}$ ist streng monoton wachsend. Dazu formen wir den Ausdruck erstmal wie folgt um: Wir definieren für alle $n \in \mathbb{N}$ die Größe $a_n := \frac{n}{n+2}$ und berechnen

$$\begin{aligned} a_{n+1} - a_n &= \frac{n+1}{(n+1)+2} - \frac{n}{n+2} = \frac{n+1}{n+3} - \frac{n}{n+2} = \frac{(n+1)(n+2) - n(n+3)}{(n+3)(n+2)} \\ &= \frac{(n^2 + 3n + 2) - (n^2 + 3n)}{(n+3)(n+2)} = \frac{2}{(n+3)(n+2)} > 0. \end{aligned}$$

Es folgt

$$a_{n+1} = a_{n+1} - a_n + a_n > 0 + a_n = a_n \quad \forall n \in \mathbb{N}.$$

QED

2.2 Wachstum von Folgen

In diesem Abschnitt interessieren uns die *Wachstumseigenschaften* von Folgen.

Welche Folge *wächst schneller*: Die Folge $(n^8)_{n \in \mathbb{N}}$ oder die Folge $(2^n)_{n \in \mathbb{N}}$? Ein möglicher Weg, das zu beantworten könnte es sein, sich die ersten Folgeglieder aufzuschreiben.

n	n^8	2^n
1	1	2
2	256	4
3	6561	8
4	65536	16
5	390625	32

Man hat nach Durchsicht dieser Tabelle den Eindruck, dass $(n^8)_{n \in \mathbb{N}}$ schneller wächst als $(2^n)_{n \in \mathbb{N}}$. Jedoch wendet sich irgendwann das Blatt:

n	n^8	2^n
1	1	2
2	256	4
3	6561	8
4	65536	16
5	390625	32
$\vdots$	$\vdots$	$\vdots$
100	10000000000000000	1267650600228229401496703205376

Wie können wir diesen Vorgang verstehen und begründen? Wie viele Folgenglieder müssen wir uns aufschreiben, bis wir die Wachstumsverhältnisse verstanden haben?

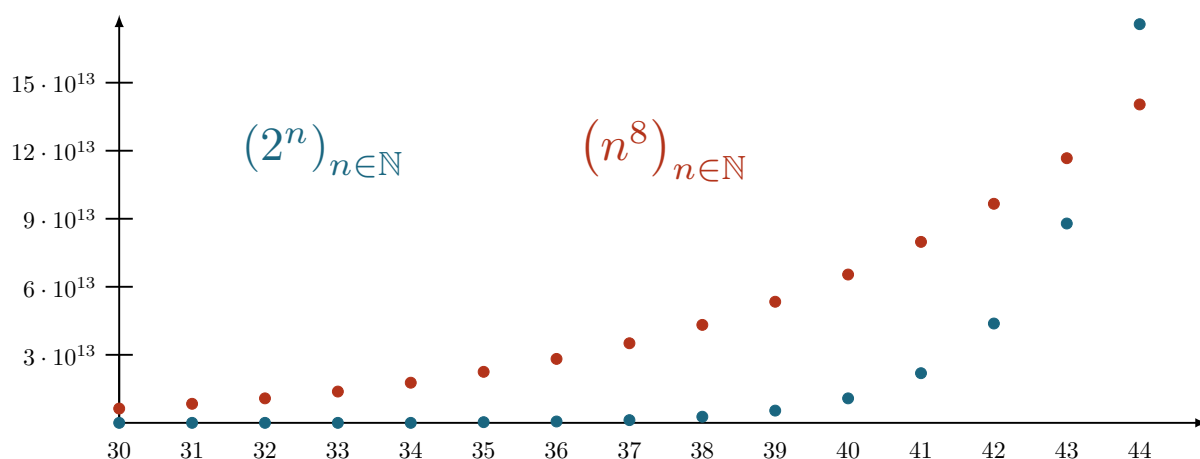


Abbildung 7: Die Folge $(2^n)_{n \in \mathbb{N}}$ überholt die Folge $(n^8)_{n \in \mathbb{N}}$ ab $n = 44$.

Definition 2.2.1 — MAJORISIERUNG/MINORISIERUNG

Es seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ reelle Folgen.

1. Wir sagen, $(b_n)_{n \in \mathbb{N}}$ *majorisiert* $(a_n)_{n \in \mathbb{N}}$, falls

$$\forall n \in \mathbb{N} : \quad a_n \leq b_n.$$

Eine weitere Sprechweise ist „ $(b_n)_{n \in \mathbb{N}}$ ist eine *Majorantenfolge* von $(a_n)_{n \in \mathbb{N}}$.“

2. Wir sagen, $(b_n)_{n \in \mathbb{N}}$ *majorisiert* $(a_n)_{n \in \mathbb{N}}$ *schließlich*, falls

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad a_n \leq b_n.$$

Eine weitere Sprechweise ist „ $(b_n)_{n \in \mathbb{N}}$ ist *schließlich eine Majorantenfolge* von $(a_n)_{n \in \mathbb{N}}$.“

Wir notieren $(a_n)_{n \in \mathbb{N}} \preceq (b_n)_{n \in \mathbb{N}}$.

3. Wir sagen, $(b_n)_{n \in \mathbb{N}}$ *minorisiert* $(a_n)_{n \in \mathbb{N}}$, falls

$$\forall n \in \mathbb{N} : \quad b_n \leq a_n.$$

Eine weitere Sprechweise ist „ $(b_n)_{n \in \mathbb{N}}$ ist eine *Minorantenfolge* von $(a_n)_{n \in \mathbb{N}}$.“

4. Wir sagen, $(b_n)_{n \in \mathbb{N}}$ *minorisiert* $(a_n)_{n \in \mathbb{N}}$ *schließlich*, falls

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad b_n \leq a_n$$

Eine weitere Sprechweise ist „ $(b_n)_{n \in \mathbb{N}}$ ist *schließlich eine Minorantenfolge* von $(a_n)_{n \in \mathbb{N}}$.“

Wir notieren $(a_n)_{n \in \mathbb{N}} \geq (b_n)_{n \in \mathbb{N}}$.

Beachte: Falls eine reelle Folge $(b_n)_{n \in \mathbb{N}}$ eine Majorantenfolge von $(a_n)_{n \in \mathbb{N}}$ ist, so majorisiert $(b_n)_{n \in \mathbb{N}}$ die Folge $(a_n)_{n \in \mathbb{N}}$ auch schließlich — man wähle einfach $n_0 = 1$ in der obigen Definition. Die Begrifflichkeiten, die wir im vorigen Teil dieses Kapitels eingeführt haben lassen sich auch mithilfe des Konzeptes der Majorisierung verstehen.

Bemerkung 2.2.2 — BESCHRÄNKTHEIT/MONOTONIE ÜBER MAJORISIERUNG

1. Eine Folge ist genau dann nach oben beschränkt, wenn sie eine konstante Folge als Majorantenfolge besitzt.
2. Eine Folge $(a_n)_{n \in \mathbb{N}}$ ist genau dann monoton wachsend, wenn $(a_{n+1})_{n \in \mathbb{N}}$ eine Majorantenfolge von $(a_n)_{n \in \mathbb{N}}$ ist.

Analoge Aussagen gelten auch für nach unten beschränkte und monoton fallende Folgen. Ferner könnte man mit dieser Beobachtung auch neue Begriffe definieren, zum Beispiel was es bedeutet, wenn eine Folge *schließlich monoton wachsend* ist.

Wir werden sehen: Die Folge $(2^n)_{n \in \mathbb{N}}$ ist zwar keine Majorantenfolge von $(n^8)_{n \in \mathbb{N}}$, aber dennoch ist sie *schließlich eine Majorantenfolge*.

Beispiel 2.2.3 — BEISPIELE FÜR MAJORISIERUNG

1. $(1)_{n \in \mathbb{N}}$ ist eine Majorantenfolge von $((-1)^n)_{n \in \mathbb{N}}$. In der Tat: Es gilt

$$\forall n \in \mathbb{N} : \quad (-1)^n \leq 1.$$

2. $(0)_{n \in \mathbb{N}}$ ist keine Majorantenfolge von $((-1)^n)_{n \in \mathbb{N}}$. In der Tat: Es gilt

$$(-1)^2 = 1 > 0.$$

Somit ist die Aussage $(-1)^n \leq 0$ für $n = 2$ nicht erfüllt. Wir folgern

$$\neg(\forall n \in \mathbb{N} : \quad (-1)^n \leq 0).$$

Wir würden gerne auch überprüfen, dass $(0)_{n \in \mathbb{N}}$ nichteinmal schließlich eine Majorantenfolge von $((-1)^n)_{n \in \mathbb{N}}$ ist. Bevor wir dazu kommen, müssen wir uns ein Paar Gedanken zur *logischen Verneinung* von Aussagen mit Quantoren machen.

Bemerkung 2.2.4 — VERHALTEN VON QUANTOREN BEI VERNEINUNG

Es sei M eine Menge. Für alle $y \in M$ sei $A(y)$ eine (von y abhängige) Aussage. Dann kann man mithilfe von Quantoren neue Aussagen formulieren, zum Beispiel die Folgende:

$$C: \quad (\forall y \in M : \quad A(y)) \quad (\text{In Worten : „Für alle } y \in M \text{ gilt } A(y)\text{“})$$

Frage: Wie würde man nun C verneinen, d.h. was ist $\neg C$?

Antwort: Mit Quantorschreibweise kann $\neg C$ wie folgt aufgeschrieben werden

$$\neg C: \quad (\exists y_0 \in M : \quad \neg A(y_0)) \quad (\text{In Worten: „Es existiert } y_0 \in M \text{ so, dass } A(y_0) \text{ nicht gilt“})$$

Liest man sich die Aussagen in Worten durch, wird hoffentlich einsichtig, dass die untere Aussage die logische Verneinung der oberen ist.

(!) Man sieht: Beim Verneinen muss ein anderer Quantor benutzt werden: Während in der Aussage C nur der Quantor $\forall$ vorkommt, kommt in $\neg C$ nur $\exists$ vor.

Aussagen mit der Gestalt von C begegnen uns häufig in der Praxis.

Ein (schon bekanntes) **Beispiel** (vgl. Definition 2.2.1): $(b_n)_{n \in \mathbb{N}}$ ist Majorantenfolge von $(a_n)_{n \in \mathbb{N}}$ genau dann wenn

$$\forall n \in \mathbb{N} : \quad a_n \leq b_n.$$

Hierbei ist also $M = \mathbb{N}$ und für $y \in \mathbb{N}$ ist $A(y) : a_y \leq b_y$. Mit der oben kennengelernten Verneinungsregel ergibt sich: $(b_n)_{n \in \mathbb{N}}$ ist keine Majorantenfolge von $(a_n)_{n \in \mathbb{N}}$ genau dann wenn

$$\exists n_0 \in \mathbb{N} : \quad \neg(a_{n_0} \leq b_{n_0}).$$

Einfacher ausgedrückt

$$\exists n_0 \in \mathbb{N} : \quad b_{n_0} < a_{n_0}.$$

Nichts anderes als das haben wir in Beispiel 2.2.3 (ii) überprüft (mit $n_0 = 2$).

Mit denselben Methoden wie oben besprochen kann auch die Aussage

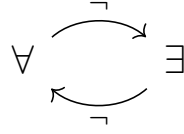
$$D: \quad (\exists y_0 \in M : \quad A(y_0)).$$

verneint werden durch

$$\neg D: \quad (\forall y \in M : \quad \neg A(y)) .$$

Dies können Sie sich selber klarmachen: Lesen Sie die Aussagen mal laut in Worten vor und überlegen Sie sich die Richtigkeit.

Als Faustregel beim Verneinen kann man also festhalten:



Mit dieser Faustregel können wir dann auch Aussagen verneinen, die mehrere Quantoren enthalten. Dazu gehen wir *Schritt für Schritt* vor:

Beispiel 2.2.5 — VERNEINUNG EINER AUSSAGE MIT MEHREREN QUANTOREN

Gemäß Definition 2.2.1 gilt $(a_n)_{n \in \mathbb{N}} \leqslant (b_n)_{n \in \mathbb{N}}$ genau dann wenn

$$\exists n_0 \in \mathbb{N} \forall m \geqslant n_0 : a_m \leqslant b_m.$$

Nun zum Verneinen: $(a_n)_{n \in \mathbb{N}} \not\leqslant (b_n)_{n \in \mathbb{N}}$ genau dann wenn

$$\neg(\exists n_0 \in \mathbb{N} \forall m \geqslant n_0 : a_m \leqslant b_m).$$

Äquivalent:

$$\forall n \in \mathbb{N} : \neg(\forall m \geqslant n : a_m \leqslant b_m). \quad (49)$$

Äquivalent:

$$\forall n \in \mathbb{N} \exists m_0 \geqslant n : \neg(a_{m_0} \leqslant b_{m_0}). \quad (50)$$

Äquivalent:

$$\forall n \in \mathbb{N} \exists m_0 \geqslant n : a_{m_0} > b_{m_0}. \quad (51)$$

In Worten gelesen: Zu jedem $n \in \mathbb{N}$ gibt es ein $m_0 \geqslant n$ mit der Eigenschaft, dass $a_{m_0} > b_{m_0}$.

Ein wenig verwirren könnte Sie hier noch der folgende Sachverhalt: In (49) stehen zwei Ungleichungen „ $a_m \leqslant b_m$ “ und „ $m \geqslant n$ “. Davon wird in (51) nur eine Ungleichung umgekehrt, die andere bleibt so. Um das einzusehen ist es eventuell sinnvoll, (49) wie folgt aufzuschreiben

$$\forall n \in \mathbb{N} \neg(a_m \leqslant b_m \quad \forall m \in \{m' \in \mathbb{N} : m' \geqslant n\}).$$

Mit dieser Darstellung kann man dann exakt verfahren wie in Vorüberlegung 2.2.4 mit $M = \{m' \in \mathbb{N} : m' \geqslant n\}$. Anstatt (50) würde man dann schreiben

$$\forall n \in \mathbb{N} \exists m_0 \in \{m' \in \mathbb{N} : m' \geqslant n\} : \neg(a_{m_0} \leqslant b_{m_0}).$$

Äquivalent:

$$\forall n \in \mathbb{N} \exists m_0 \in \{m' \in \mathbb{N} : m' \geqslant n\} : a_{m_0} > b_{m_0}.$$

Jetzt sieht man vielleicht noch besser, dass sich die Ungleichung in der Mengenbeschreibung nicht umkehrt, die andere jedoch schon.

Beispiel zu 2.2.3a.

Wir zeigen nun $((-1)^n)_{n \in \mathbb{N}} \not\leq (0)_{n \in \mathbb{N}}$. Hierfür weisen wir (51) nach.

$$\mathbb{Z} : \quad \forall n \in \mathbb{N} \quad \exists m_0 \geq n : \quad (-1)^{m_0} > 0.$$

Wir müssen also für jedes beliebige $n \in \mathbb{N}$ ein $m_0 \geq n$ finden mit $(-1)^{m_0} > 0$. Sei dazu $n \in \mathbb{N}$. Definiere $m_0 := 2n$. Dann gilt

$$(-1)^{m_0} = (-1)^{2n} = [(-1)^2]^n = [(-1)(-1)]^n = 1^n = 1 > 0.$$

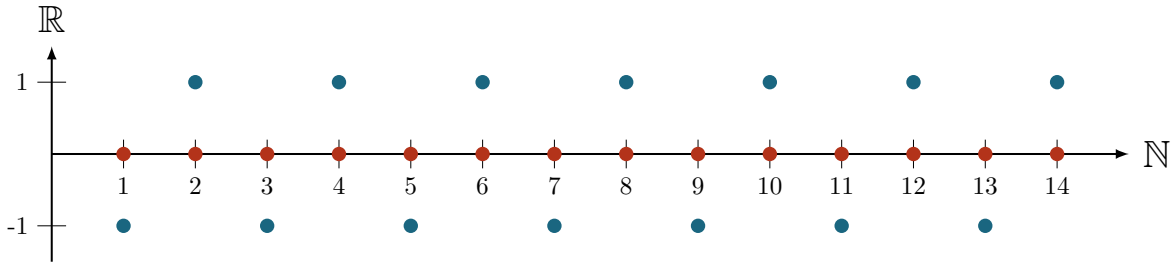


Abbildung 8: Die Folge $((-1)^n)_{n \in \mathbb{N}}$ (blau) und die Folge $(0)_{n \in \mathbb{N}}$ (orange)

Wir interessieren uns im Folgenden für verschiedenste Folgen (z.B. diejenigen aus Beispiel 2.1.2) und fragen uns ob sie einander (schließlich) majorisieren.
Zunächst untersuchen wir hier verschiedene *polynomiell wachsende Folgen*.

Proposition 2.2.6 — POLYNOMIELLES WACHSTUM

Es seien $\beta, \alpha, \varepsilon \in \mathbb{R}$ so, dass $\beta > \alpha \geq 0$ und $\varepsilon > 0$. Dann gilt ^a

$$(n^\alpha)_{n \in \mathbb{N}} \leq (\varepsilon n^\beta)_{n \in \mathbb{N}}$$

^aFür große Werte von ε ist diese Aussage absolut nicht verwunderlich. Für kleine Werte von ε ist hier schon etwas zu zeigen, vgl. auch Abbildung 9. Den griechischen Buchstaben ε (manchmal auch ϵ) verwendet man meist, wenn man an eine eher kleine Zahl denkt.

Beweis. Es seien $\beta > \alpha \geq 0$ und $\varepsilon > 0$. Dann gilt für $n \in \mathbb{N}$

$$n^\alpha \leq \varepsilon n^\beta \iff 1 \leq \varepsilon n^{\beta-\alpha} \iff \frac{1}{\varepsilon} \leq n^{\beta-\alpha} \iff \left(\frac{1}{\varepsilon}\right)^{\frac{1}{\beta-\alpha}} \leq n. \quad (52)$$

Fixiere nun ein $n_0 \in \mathbb{N}$ so^a, dass

$$n_0 \geq \left(\frac{1}{\varepsilon}\right)^{\frac{1}{\beta-\alpha}}. \quad (53)$$

Nun

$$\mathbb{Z}: \quad \forall n \geq n_0 : \quad n^\alpha \leq \varepsilon n^\beta. \quad (54)$$

Dazu:

$$n \geq n_0 \xRightarrow{(53)} n \geq \left(\frac{1}{\varepsilon}\right)^{\frac{1}{\beta-\alpha}} \xRightarrow{(52)} n^\alpha \leq \varepsilon n^\beta.$$

Mit anderen Worten

$$\forall n \geq n_0 : \quad n^\alpha \leq \varepsilon n^\beta.$$

Wir haben also gesehen

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad n^\alpha \leq \varepsilon n^\beta.$$

QED

^aDass man eine solche Zahl $n_0 \in \mathbb{N}$ überhaupt wählen kann, werden wir im nächsten Kapitel beweisen, wenn wir wirklich wissen, was $\mathbb{N}$ und $\mathbb{R}$ sind. Mit der Anschauung der Zahlengerade, sollte es aber relativ intuitiv sein.

Wir werden später noch weitere *Wachstumsvergleiche* wie in der vorigen Proposition anstellen.

Bemerkung 2.2.7

Im vorigen Beweis mussten wir eine Zahl $n_0 \in \mathbb{N}$ so wählen, dass (54) gilt. Manchmal interessiert man sich für die *kleinstmögliche* Wahl eines solchen $n_0 \in \mathbb{N}$. Dazu:

- Sei für $x \in \mathbb{R}$ die *obere Gauß-Klammer* $[x]$ die eindeutige^a ganze Zahl $z \in \mathbb{Z}$ so, dass $x \leq z < x + 1$.

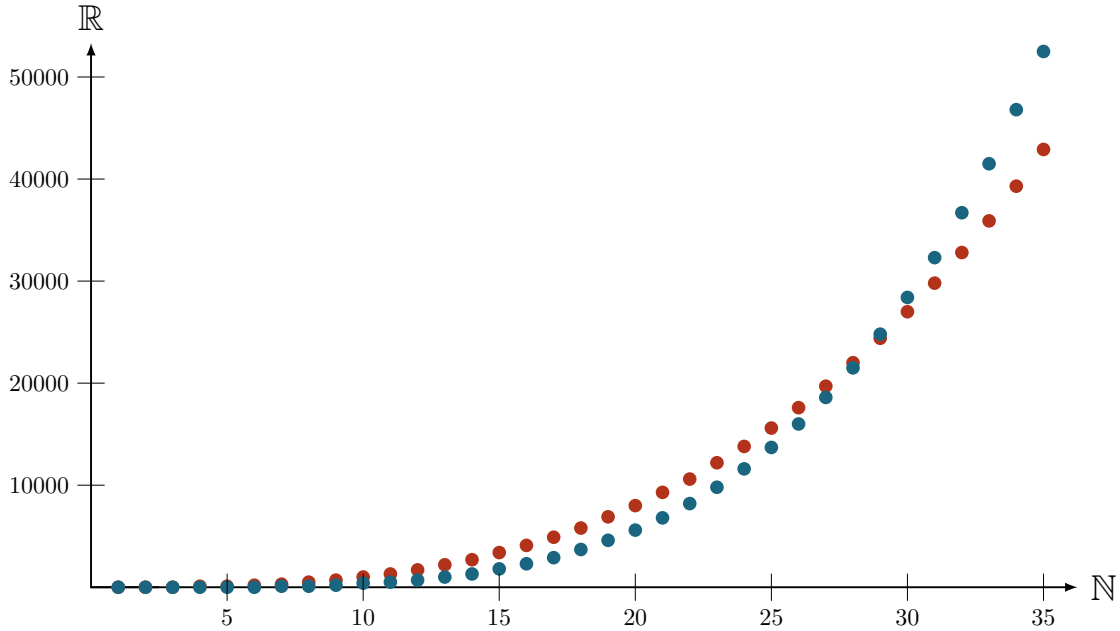


Abbildung 9: $(\frac{35}{1000}n^4)_{n \in \mathbb{N}}$ (blau) majorisiert $(n^3)_{n \in \mathbb{N}}$ (orange) schließlich.

Dann kann man sich leicht überlegen, dass die kleinstmögliche Wahl von n_0 so, dass (54) gilt, durch

$$n_0^* := \left\lceil \left(\frac{1}{\varepsilon}\right)^{\frac{1}{\beta-\alpha}} \right\rceil$$

gegeben ist.

^aÜberlegen Sie sich mal am Zahlenstrahl, dass es eine solche ganze Zahl gibt und sie eindeutig ist. Ein paar Beispiele: $\lceil 3.5 \rceil = 4$ oder $\lceil \frac{5}{2} \rceil = 3$.

Viel leichter kann man das Konzept der schließlichen Majorisierung übrigens auch mithilfe des Konzeptes der Endlichkeit fassen.

Proposition 2.2.8 — SCHLIESSLICHE MAJORISIERUNG UND ENDLICHKEIT

Es seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ reelle Folgen. Dann gilt

- (i) $(a_n)_{n \in \mathbb{N}} \leqslant (b_n)_{n \in \mathbb{N}} \iff \{m \in \mathbb{N} : a_m > b_m\} \text{ ist endlich.}$
- (ii) $(a_n)_{n \in \mathbb{N}} \geqslant (b_n)_{n \in \mathbb{N}} \iff \{m \in \mathbb{N} : a_m < b_m\} \text{ ist endlich.}$

Beweis. Wir zeigen nur (i), denn (ii) geht analog.

Zu „ $\implies$ “

$$(a_n)_{n \in \mathbb{N}} \leqslant (b_n)_{n \in \mathbb{N}}$$

$$\begin{aligned} \implies & \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : a_n \leq b_n \\ \implies & \{m \in \mathbb{N} : a_m > b_m\} \subseteq \{1, \dots, n_0\}. \end{aligned}$$

Da nun $\{1, \dots, n_0\}$ endlich ist, ist nach Proposition 2.0.10 auch $\{m \in \mathbb{N} : a_m > b_m\}$ endlich. Zu „ $\Leftarrow$ “. Es sei

$$M := \{m \in \mathbb{N} : a_m > b_m\} \quad (55)$$

eine endliche Menge. Dann gibt es ein $k \in \mathbb{N}$ so, dass man

$$\text{eine bijektive Funktion } f : \{1, \dots, k\} \rightarrow M \quad (56)$$

finden kann. Definiere

$$n_0 := f(1) + \dots + f(k) + 1. \quad (57)$$

Wir zeigen: $(\forall n \geq n_0 : a_n \leq b_n)$. Dazu: Angenommen das Gegenteil gilt, d.h. $\neg(\forall n \geq n_0 : a_n \leq b_n)$. Äquivalent

$$\exists n^* \geq n_0 : a_{n^*} > b_{n^*} \quad (58)$$

Mit der Mengenbeschreibung in (55) stellt man fest, dass $n^* \in M$. Wegen der Bijektivität von f (vgl. (56)) gibt es $\ell \in \{1, \dots, k\}$ mit $f(\ell) = n^*$. Man hat

$$f(\ell) = n^* \underset{(58)}{\geq} n_0 \underset{(57)}{=} f(1) + \dots + f(k) + 1.$$

Beachte: Da $\ell \in \{1, \dots, k\}$ kann man folgern

$$f(\ell) \geq f(1) + \dots + f(\ell - 1) + f(\ell) + f(\ell + 1) + \dots + f(k) + 1.$$

Man erhält

$$0 \geq f(1) + \dots + f(\ell - 1) + f(\ell + 1) + \dots + f(k) + 1.$$

Ein Widerspruch, denn aufgrund der Tatsache, dass $M \subseteq \mathbb{N}$ gilt $f(i) \geq 1$ für alle $i \in \{1, \dots, k\}$ (weswegen die rechte Seite eine Zahl strikt größer als 0 ergeben muss). **QED.**

Beispiel 2.2.9

Betrachte noch einmal die Folge $((-1)^n)_{n \in \mathbb{N}}$ und $(0)_{n \in \mathbb{N}}$. Dann gilt

$$((-1)^n)_{n \in \mathbb{N}} \not\leq (0)_{n \in \mathbb{N}} \quad \text{und} \quad ((-1)^n)_{n \in \mathbb{N}} \not\geq (0)_{n \in \mathbb{N}}.$$

Beides folgt direkt aus Proposition 2.2.8. Ersteres gilt weil

$$\{m \in \mathbb{N} : (-1)^m > 0\} = \{2m : m \in \mathbb{N}\} \quad \text{eine unendliche Menge ist.}$$

und letzteres gilt weil

$$\{m \in \mathbb{N} : (-1)^m < 0\} = \{2m - 1 : m \in \mathbb{N}\} \quad \text{eine unendliche Menge ist.}$$

Mit der Relation „ $\leq$ “ lassen sich also nicht alle Folgen miteinander *vergleichen*! Das ist ein großer Unterschied zu der Vergleichsrelation „ $\leq$ “ auf $\mathbb{R}$. Dort gilt stets eine der beiden Aussagen $a \leq b$ oder $b \leq a$.

Rechenregeln für das Folgenwachstum Wer zwei Folgen $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ gegeben hat, der kann daraus neue Folgen machen, etwa durch Bildung von $(a_n + b_n)_{n \in \mathbb{N}}$. Wir wollen auch das Wachstum dieser Folgen verstehen.

Proposition 2.2.10 — RECHENREGELN FÜR DAS FOLGENWACHSTUM

Es seien $(a_n)_{n \in \mathbb{N}}$, $(b_n)_{n \in \mathbb{N}}$, $(c_n)_{n \in \mathbb{N}}$ und $(d_n)_{n \in \mathbb{N}}$ reelle Folgen.

- (i) Ist $(a_n)_{n \in \mathbb{N}} \leq (b_n)_{n \in \mathbb{N}}$ und $(c_n)_{n \in \mathbb{N}} \leq (d_n)_{n \in \mathbb{N}}$ so ist $(a_n + c_n)_{n \in \mathbb{N}} \leq (b_n + d_n)_{n \in \mathbb{N}}$.
- (ii) Ist $(a_n)_{n \in \mathbb{N}} \leq (b_n)_{n \in \mathbb{N}}$ und $c_n \geq 0 \ \forall n \in \mathbb{N}$, so gilt $(c_n a_n)_{n \in \mathbb{N}} \leq (c_n b_n)_{n \in \mathbb{N}}$.
- (iii) Ist $(a_n)_{n \in \mathbb{N}} \leq (b_n)_{n \in \mathbb{N}}$ und $c_n \leq 0 \ \forall n \in \mathbb{N}$, so gilt $(c_n b_n)_{n \in \mathbb{N}} \leq (c_n a_n)_{n \in \mathbb{N}}$.
- (iv) Ist $(a_n)_{n \in \mathbb{N}} \leq (b_n c_n)_{n \in \mathbb{N}}$ und $b_n > 0 \ \forall n \in \mathbb{N}$, so ist $(\frac{a_n}{b_n})_{n \in \mathbb{N}} \leq (c_n)_{n \in \mathbb{N}}$.
- (v) $[(a_n)_{n \in \mathbb{N}} \leq (b_n)_{n \in \mathbb{N}}] \wedge [(b_n)_{n \in \mathbb{N}} \leq (c_n)_{n \in \mathbb{N}}] \implies (a_n)_{n \in \mathbb{N}} \leq (c_n)_{n \in \mathbb{N}}$.

Beweis. Wir zeigen nur (i) und (iv) und überlassen den Lesenden (ii), (iii) und (v) als Übungsaufgabe. Zu (i). Da $(a_n)_{n \in \mathbb{N}} \leq (b_n)_{n \in \mathbb{N}}$ ist, gilt (vgl. Definition 2.2.1)

$$\exists n_1 \in \mathbb{N} \quad \forall n \geq n_1 : \quad a_n \leq b_n \quad (59)$$

Da auch $(c_n)_{n \in \mathbb{N}} \leq (d_n)_{n \in \mathbb{N}}$ ist, gilt (wiederum vgl. Definition 2.2.1)

$$\exists n_2 \in \mathbb{N} \quad \forall n \geq n_2 : \quad c_n \leq d_n. \quad (60)$$

Schreiben wir nun kurz auf was zu zeigen ist:

$$\mathbb{Z} : \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad a_n + c_n \leq b_n + d_n \quad (61)$$

Wir wählen dazu $n_0 := \max\{n_1, n_2\}$. Wir behaupten

$$\forall n \geq n_0 : \quad a_n + c_n \leq b_n + d_n.$$

Wenn dies gezeigt ist, dann ist die Existenz eines Index n_0 wie in Definition 2.2.1 nachgewiesen (und somit wäre auch (i) gezeigt)! Sei also $n \geq n_0$ beliebig. Dann gilt

$$n \geq n_1 \quad \wedge \quad n \geq n_2.$$

Es folgt

$$a_n + c_n \stackrel{(59)}{\underset{(n \geq n_1)}{\leq}} b_n + c_n \stackrel{(60)}{\underset{(n \geq n_2)}{\leq}} b_n + d_n.$$

Gleichung (61) ist mit unserer getroffenen Wahl von n_0 nachgewiesen. Mit anderen Worten (vgl. Definition 2.2.1) $(a_n + c_n)_{n \in \mathbb{N}} \leq (b_n + d_n)_{n \in \mathbb{N}}$.

Zu (iv).

$$\begin{aligned}
 & (a_n)_{n \in \mathbb{N}} \leq (b_n c_n)_{n \in \mathbb{N}} \\
 \xRightarrow[\text{Def 2.2.1}]{} & \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad a_n \leq b_n c_n \\
 \xRightarrow[(b_n > 0 \ \forall n)]{} & \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \frac{a_n}{b_n} \leq c_n \\
 \xRightarrow[\text{Def 2.2.1}]{} & \left(\frac{a_n}{b_n}\right)_{n \in \mathbb{N}} \leq (c_n)_{n \in \mathbb{N}}.
 \end{aligned}$$

QED

30.10.2025

Beispiel 2.2.11

Behauptung. Es gilt für alle $\varepsilon > 0$, dass

$$\left(\frac{1}{n}\right)_{n \in \mathbb{N}} \leq (\varepsilon)_{n \in \mathbb{N}}.$$

Begründung. Aus Proposition 2.2.6 folgt

$$(1)_{n \in \mathbb{N}} = (n^0)_{n \in \mathbb{N}} \leq (\varepsilon n^1)_{n \in \mathbb{N}} = (\varepsilon n)_{n \in \mathbb{N}}.$$

Wendet man nun Proposition 2.2.10 (iv) mit $a_n := 1, b_n := n, c_n := \varepsilon$ an, so folgt die Behauptung.

Ein Fazit aus dem vorigen Beispiel: Egal wie klein doch $\varepsilon > 0$ sein mag, ab irgendeinem Index unterschreitet die Folge $\left(\frac{1}{n}\right)_{n \in \mathbb{N}}$ den Wert ε :

Aus dem Eingangsbeispiel von Sektion 2.2 haben wir noch ein wichtiges To-Do offen: Wir würden gerne zeigen, dass $(n^8)_{n \in \mathbb{N}} \leq (2^n)_{n \in \mathbb{N}}$. Um zwei so verschiedene Folgen miteinander vergleichen zu können, müssen wir zunächst ein paar wichtige *Ungleichungen* studieren.

Wichtige (Un)-Gleichungen

Unsere weiteren Überlegungen werden auf einer wichtigen Ungleichung basieren, die *exponentielles Wachstum* mit *linearem Wachstum* vergleicht — die *Bernoulli-Ungleichung*. Sie besagt

$$\forall x \geq 0 \quad \forall n \in \mathbb{N} : \quad (1 + x)^n \geq 1 + nx.$$

Eigentlich gilt sie sogar für alle $x \geq -1$. Das werden wir aber erst im nächsten Kapitel zeigen.¹ Den Ausdruck $(1 + x)^n$ werden wir hier *kombinatorisch* durch ausmultiplizieren studieren.

¹Wer möchte, kann sich aber mit GEOGEBRA davon überzeugen: [LINK](#)

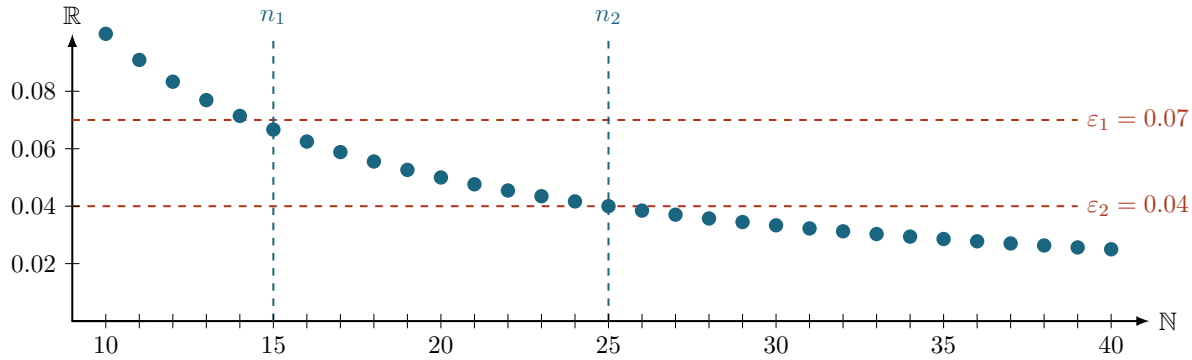


Abbildung 10: Nach $n_1 = 15$ sind alle Folgenglieder kleiner als $\varepsilon_1 = 0.07$. Nach $n_2 = 25$ kleiner als $\varepsilon_2 = 0.04$.

Dazu mal ein Experiment:

$$\begin{aligned}(1+x)^1 &= 1 + x \\ &= 1 + x.\end{aligned}\tag{62}$$

$$\begin{aligned}(1+x)^2 &= (1+x)(1+x) = 1 \cdot (1+x) + x \cdot (1+x) \\ &\quad \text{Alles mit allem} \\ &\quad \text{ausmultipliziert} \quad 1 \cdot 1 + 1 \cdot x + x \cdot 1 + x \cdot x \\ &= 1 + 2x + x^2.\end{aligned}\tag{63}$$

$$\begin{aligned}(1+x)^3 &= (1+x)(1+x)(1+x) \stackrel{(63)}{=} [1 \cdot 1 + 1 \cdot x + x \cdot 1 + x \cdot x](1+x) \\ &\quad \text{Alles mit allem} \\ &\quad \text{ausmultipliziert} \quad 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot x + 1 \cdot x \cdot 1 + 1 \cdot x \cdot x + x \cdot 1 \cdot 1 + x \cdot 1 \cdot x + x \cdot x \cdot 1 + x \cdot x \cdot x \\ &= 1 + 3x + 3x^2 + x^3.\end{aligned}$$

Wir sehen: beim Ausmultiplizieren entstehen Summanden aus *allen erdenklichen Kombinationen*. Man erhält z.B. bei $(1+x)^3$ sämtliche Produkte mit drei Faktoren, wobei man für jeden Faktor immer entweder die 1 oder x auswählen muss. Beim ersten Faktor hat man dann also die Wahl zwischen **blau** und **orange**, beim zweiten Faktor die Wahl zwischen **grün** und **rot** und beim dritten Faktor die Wahl zwischen **cyan** und **magenta**. Man muss alle diese Wahlmöglichkeiten berücksichtigen und aufsummieren.

Wir haben die Einser immer in einer *blau-artigen* Farbe (blau/grün/cyan) eingefärbt und die x -er immer in einer *orange-artigen* Farbe (orange/rot/magenta) eingefärbt. Lassen Sie uns mal alle Einser im gleichen blau und alle x -er im gleichen orange einfärben.

$$\begin{aligned}(1+x)^1 &= 1 + x \\ &= 1 + x.\end{aligned}$$

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

$$\begin{array}{l} \text{Alles mit allem} \\ \text{ausmultipliziert} \end{array} \quad 1 \cdot 1 + 1 \cdot x + x \cdot 1 + x \cdot x \\ = 1 + 2x + x^2.$$

$$(1+x)^3 = (1+x)(1+x)(1+x)$$

$$\begin{array}{l} \text{Alles mit allem} \\ \text{ausmultipliziert} \end{array} \quad 1 \cdot 1 \cdot 1 + 1 \cdot 1 \cdot x + 1 \cdot x \cdot 1 + 1 \cdot x \cdot x + x \cdot 1 \cdot 1 + x \cdot 1 \cdot x + x \cdot x \cdot 1 + x \cdot x \cdot x \\ = 1 + 3x + 3x^2 + x^3.$$

Jetzt hat man für die Summanden von $(1+x)^3$ bei jedem der drei Faktoren die Wahl zwischen **blau** und **orange**. Man spielt dann alle Kombinationen durch und summiert auf.

Für das Endergebnis $(1+x)^3 = 1 + 3x + 3x^2 + x^3$ ist es nur wichtig, zu zählen, wie oft null orangene Faktoren vorkommen (einmal), wie oft ein orangener Term vorkommt (dreimal), wie oft zwei orangene Faktoren vorkommen (dreimal), und wie oft drei orangene Faktoren vorkommen (einmal).

Das Problem sieht als *Urnenproblem* wie folgt aus:

Vorüberlegung — EIN ZUGEHÖRIGES URNENPROBLEM

In einer Urne liegen zwei Kugeln, eine blaue und eine orangene. Man darf dreimal ziehen. Es gibt also die folgenden Optionen.

Erster Ziehvorgang	Zweiter Ziehvorgang	Dritter Ziehvorgang
Blau oder Orange	Blau oder Orange	Blau oder Orange

Fragen: Wie viele Möglichkeiten gibt es

- (i) Insgesamt 0 orangene Kugeln zu ziehen?
- (ii) Insgesamt 1 orangene Kugel zu ziehen?
- (iii) Insgesamt 2 orangene Kugeln zu ziehen?
- (iv) Insgesamt 3 orangene Kugeln zu ziehen ?

Antworten. Zu (i). Wer null orangene Kugeln ziehen will, muss in jedem Zug eine blaue Ziehen. Es gibt also nur die Möglichkeit

Erster Ziehvorgang	Zweiter Ziehvorgang	Dritter Ziehvorgang
Blau	Blau	Blau

Zu (ii). Wer insgesamt eine orangene Kugel ziehen möchte, der kann entweder im ersten Ziehvorgang eine orangene Kugel ziehen (und sonst nur blaue) oder im zweiten Ziehvorgang eine orangene Kugel ziehen (und sonst nur blaue) oder im dritten Ziehvorgang eine orangene (und sonst nur blaue) zu ziehen. Man hat also drei Möglichkeiten.

Erster Ziehvorgang	Zweiter Ziehvorgang	Dritter Ziehvorgang
Orange	Blau	Blau
Blau	Orange	Blau
Blau	Blau	Orange

Zu (iii). Wer insgesamt zwei orangene Kugeln ziehen möchte, der darf insgesamt nur eine blaue ziehen. Ganz spiegelbildlich gedacht gibt es dafür auch nur drei Möglichkeiten

Erster Ziehvorgang	Zweiter Ziehvorgang	Dritter Ziehvorgang
Blau	Orange	Orange
Orange	Blau	Orange
Orange	Orange	Blau

Zu (iv). Überlegen Sie sich selber, dass es nur eine Möglichkeit gibt.

Allgemein: Für $(1+x)^n$ erhalten wir folgende Situation:

Vorüberlegung — EIN ZUGEHÖRIGES URNENPROBLEM: VERALLGEMEINERUNG
In einer Urne liegen zwei Kugeln, eine blaue und eine orangene. Wir ziehen insgesamt n -mal. Es gibt also die folgenden Optionen.

Erster Ziehvorgang	Zweiter Ziehvorgang	...	n -ter Ziehvorgang
Blau oder Orange	Blau oder Orange	...	Blau oder Orange

Frage: Es sei $k \in \{1, \dots, n\}$. Wie viele Möglichkeiten gibt es, insgesamt k orangene Kugeln zu ziehen?

Man müsste nun also sämtliche Möglichkeiten durchspielen. Eine Möglichkeit besteht hierbei daraus, für jeden der n Ziehvorgänge entweder **Blau** oder **Orange** in die jeweilige Spalte zu schreiben.

Hierzu ein Trick: Anstatt sich für jeden Ziehvorgang einzeln „Blau oder Orange“ zu notieren kann man sich auch einfach nur die **Nummern der Ziehvorgänge** aufschreiben, in denen man die orangene Kugel gezogen hat. Denn es ist ja offensichtlich, dass man dann in den nicht notierten Ziehvorgängen blau gezogen hat. Man erhält dadurch für jede Option eine Liste von Zahlen zwischen 1 und n . Die Liste wird stets *aufsteigend sortiert* sein, denn Sie notieren sich den frühesten orangenen Ziehvorgang ja zuerst.

Ein Beispiel: Nehmen Sie $n = 4$. Die Möglichkeit

Erster Ziehvorgang	Zweiter Ziehvorgang	Dritter Ziehvorgang	Vierter Ziehvorgang
Orange	Blau	Orange	Orange

wird jetzt einfach mit der Liste „1, 3, 4“ notiert.

Will man insgesamt k orangene Kugeln ziehen, so möchte man sich also k Nummern von 1 bis n in der aufsteigenden Liste notiert haben. Die Liste soll also die *Länge* k haben.

Es ergibt sich ein neues Urnenproblem:

Vorüberlegung 2.2a — EIN WEITERES ZUGEHÖRIGES URNENPROBLEM

In einer Urne liegen n Kugeln, alle mit verschiedenen Nummern $1, \dots, n$. Man darf k Kugeln (gleichzeitig) ziehen und notiert sich die Nummern in aufsteigender Reihenfolge in einer Liste.

Frage: Wie viele Möglichkeiten für solche aufsteigende Listen der Länge k hat man?

Anstatt am Ende aufsteigend zu ordnen könnte man auch nacheinander ziehen und festlegen dass „die Reihenfolge der gezogenen Zahlen keine Rolle spielt“ (denn wenn Sie die Zahlen am Ende der Größe nach ordnen, torpedieren Sie ja jegliche Reihenfolge der gezogenen Zahlen). In der Schule haben Sie dieses Urnenproblem eventuell (bundeslandabhängig) studiert und dafür den *Binomialkoeffizient* $\binom{n}{k}$ kennengelernt, welcher auch die Antwort auf die in Vorüberlegung 2.2a gestellte Frage ist. Falls nicht, ist das kein Problem: Wir werden es hier von Grund auf studieren.

Wir kennen übrigens bereits eine Möglichkeit, Elemente in einer Art und Weise aufzulisten in der die Reihenfolge keine Rolle spielt: Mithilfe von *Mengen*!

Vorüberlegung 2.2b — DAS TEILMENGENPROBLEM.

Es sei die Menge $\{1, \dots, n\}$ gegeben. *Frage:* Wie viele verschiedene Teilmengen von $\{1, \dots, n\}$ gibt es, die genau k Elemente enthalten?

Die Antwort auf diese Frage werden wir im nächsten Paragraphen finden.

Ein wenig Kombinatorik Wir definieren hier zunächst das Konzept der *Binomialkoeffizienten*.

Definition 2.2.12 — BINOMIALKOEFFIZIENT

Es sei $\alpha \in \mathbb{R}$ und $k \in \mathbb{N} \cup \{0\}$. Dann definieren wir den *Binomialkoeffizienten* „ α über k “ durch

$$\binom{\alpha}{0} := 1$$

und falls $k \in \mathbb{N}$

$$\binom{\alpha}{k} := \frac{\alpha \cdot (\alpha - 1) \cdot (\alpha - 2) \cdot \dots \cdot (\alpha - k + 1)}{k!}. \quad (a) \quad (64)$$

^aAchtung: falls $k \leq 2$ so taucht der Faktor $\alpha - 2$ im Produkt des Zählers gar nicht auf, da $\alpha - k + 1 > \alpha - 2$. Den Faktor $\alpha - 2$ haben wir allgemein nur zur Veranschaulichung des Musters mitgeschrieben. Es bedeutet nicht, dass er immer auftauchen muss.

Bevor wir lernen, wofür Binomialkoeffizienten gut sind, studieren wir zunächst ein paar Eigenschaften.

Proposition 2.2.13 — EIGENSCHAFTEN VON BINOMIALKOEFFIZIENTEN

Es seien $\alpha \in \mathbb{R}$ und $k \in \mathbb{N}$. Dann gilt

(i)

$$\binom{\alpha}{1} = \alpha.$$

(ii)

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

Beweis. Zu (i). Da $\alpha = \alpha - 1 + 1$ hat das Produkt im Zähler von (64) nur den Faktor α . Da $1! = 1$ haben wir

$$\binom{\alpha}{1} = \frac{\alpha}{1} = \alpha.$$

Nun zu (ii). Zunächst betrachten wir den Fall ' $k \geq 2$ '. Es gilt

$$\begin{aligned} & \binom{\alpha-1}{k-1} + \binom{\alpha-1}{k} \\ & \stackrel{(64)}{=} \frac{(\alpha-1) \cdots (\alpha-1-(k-1)+1)}{(k-1)!} + \frac{(\alpha-1) \cdots (\alpha-1-k+1)}{k!} \\ & = \frac{(\alpha-1) \cdots (\alpha-k+1)}{(k-1)!} + \frac{(\alpha-1) \cdots (\alpha-k)}{k!} \\ & = \frac{(\alpha-1) \cdots (\alpha-k+1) \cdot k}{(k-1)! \cdot k} + \frac{(\alpha-1) \cdots (\alpha-k)}{k!} \\ & \stackrel{\text{Bsp 2.2.21}}{=} \frac{(\alpha-1) \cdots (\alpha-k+1) \cdot k}{k!} + \frac{(\alpha-1) \cdots (\alpha-k)}{k!} \\ & = \frac{(\alpha-1) \cdots (\alpha-k+1) \cdot k}{k!} + \frac{(\alpha-1) \cdots (\alpha-k+1) \cdot (\alpha-k)}{k!} \\ & \stackrel{\text{Gemeinsame Faktoren ausklammern}}{=} \frac{(\alpha-1) \cdots (\alpha-k+1)}{k!} \cdot (k + (\alpha-k)) \\ & = \frac{(\alpha-1) \cdots (\alpha-k+1)}{k!} \cdot \alpha = \frac{\alpha \cdot (\alpha-1) \cdots (\alpha-k+1)}{k!} \stackrel{(64)}{=} \binom{\alpha}{k}. \end{aligned}$$

Nun müssen wir noch zusätzlich den Fall ‘ $k = 1$ ’ separat nachrechnen. Dazu:

$$\binom{\alpha-1}{0} + \binom{\alpha-1}{1} \stackrel{(i)}{=} 1 + (\alpha-1) = \alpha \stackrel{(i)}{=} \binom{\alpha}{1}.$$

QED

Weitere Eigenschaften ergeben sich im Spezialfall, dass $\alpha \in \mathbb{N}$.

Proposition 2.2.14 — BINOMIALKOEFFIZIENTEN MIT NATÜRLICHEM α

Es seien $n, k \in \mathbb{N} \cup \{0\}$.

(i) Falls $0 \leq k \leq n$ so gilt

$$\binom{n}{k} = \frac{n!}{(n-k)! \cdot k!}.$$

(Hierbei wird die Konvention $0! := 1$ benutzt)

(ii) Falls $0 \leq k \leq n$ so gilt

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

(iii) Falls $k > n$ so gilt

$$\binom{n}{k} = 0.$$

Beweis. Zu (i). Für $k \geq 1$ ergibt sich

$$\begin{aligned} \binom{n}{k} &\stackrel{(64)}{=} \frac{n(n-1) \cdot \dots \cdot (n-k+1)}{k!} \\ &= \frac{n(n-1) \cdot \dots \cdot (n-k+1) \cdot (n-k) \cdot (n-k-1) \cdot \dots \cdot 1}{k! \cdot (n-k) \cdot (n-k-1) \cdot \dots \cdot 1} \\ &= \frac{n(n-1) \cdot \dots \cdot 1}{k! \cdot (n-k)!} = \frac{n!}{(n-k)!k!}. \end{aligned}$$

Für $k = 0$ ergibt sich

$$\binom{n}{k} = \binom{n}{0} = 1 = \frac{n!}{n!0!} = \frac{n!}{(n-k)!k!}.$$

Zu (ii). Sei $0 \leq k \leq n$. Beachte, dass $0 \leq n-k \leq n$. Man hat

$$\binom{n}{n-k} \stackrel{(i)}{=} \frac{n!}{(n-(n-k))!(n-k)!} = \frac{n!}{k!(n-k)!} = \frac{n!}{(n-k)!k!} \stackrel{(i)}{=} \binom{n}{k}.$$

Zu (iii). Sei $k > n$, d.h. (weil k und n ganze Zahlen sind) $k \geq n + 1$. Dann gilt auch $n - k + 1 \leq 0$. Daher hat man

$$\binom{n}{k} \stackrel{(64)}{=} \frac{n(n-1) \cdot \dots \cdot (n-k+1)}{k!} = \frac{n(n-1) \cdot \dots \cdot 0 \cdot \dots \cdot (n-k+1)}{k!} = 0.$$

Bemerkung 2.2.15

Insbesondere folgt aus Proposition 2.2.14(ii) für alle $n \in \mathbb{N}_0$, dass

$$\binom{n}{n} \stackrel{\text{Prop 2.2.14}}{=} \binom{n}{0} \stackrel{\text{Def 2.2.12}}{=} 1.$$

Bemerkung 2.2.16 — PASCAL'SCHES DREIECK

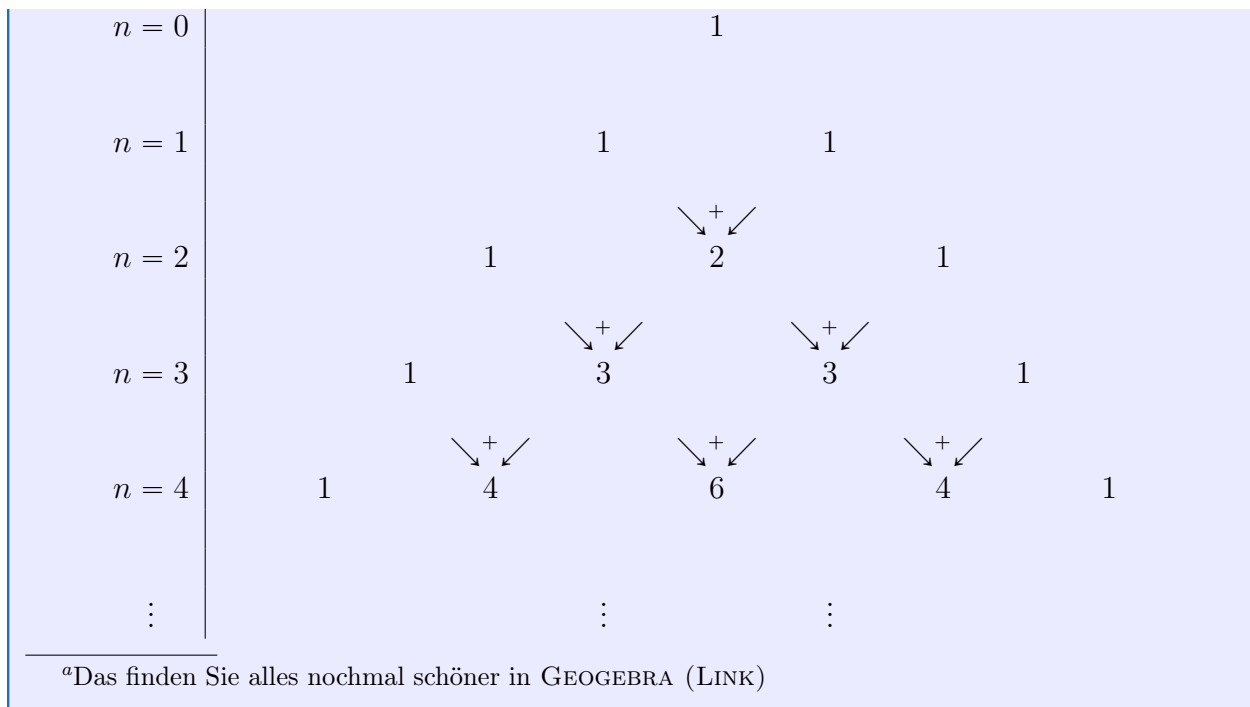
Hier^a lernen wir eine Methode, um für $n, k \in \mathbb{N} \cup \{0\}$ die Binomialkoeffizienten sukzessive anhand eines *Schemas* auszurechnen. Dazu ordne zunächst Binomialkoeffizienten im folgenden *Dreiecksschema* an:

$$\begin{array}{c|cccccc} n = 0 & & & & & & \binom{0}{0} \\ n = 1 & & & & \binom{1}{0} & & \binom{1}{1} \\ n = 2 & & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} \\ n = 3 & & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} \\ n = 4 & \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} \\ \vdots & & & & & \vdots & & \vdots & & \end{array}$$

Nun beachte

- Bemerkung 2.2.15 besagt, dass alle Koeffizienten am *Rand des Dreiecks* durch 1 gegeben sind.
- Proposition 2.2.13(ii) besagt, dass jeder Koeffizient, der nicht am Rand des Dreiecks liegt, die Summe aus den beiden schräg darüberliegenden Koeffizienten sein muss.

Mit nur diesen beiden Regeln kann man alle Binomialkoeffizienten sukzessive berechnen.



Nun endlich sind wir in der Lage, die Frage in Vorüberlegung 2.2b zu beantworten.

Proposition 2.2.17 — LÖSUNG DES TEILMENGENPROBLEMS

Es seien $k, n \in \mathbb{N} \cup \{0\}$ mit $0 \leq k \leq n$. Dann gibt es genau $\binom{n}{k}$ viele k -elementige Teilmengen von $\{1, \dots, n\}$.

Beweisskizze^a. Man muss zunächst k Elemente aus einer Menge mit n Elementen wählen. Man hat dafür

- n Möglichkeiten, das erste Element zu wählen,
- $n - 1$ Möglichkeiten, das zweite Element zu wählen (denn das als erstes gewählte Element darf kein weiteres Mal gewählt werden),
- $n - 2$ Möglichkeiten, das dritte Element zu wählen (analog),
- ...
- $n - k + 1$ Möglichkeiten, das k -te Element zu wählen.

Dies würde insgesamt $n \cdot (n - 1) \cdot \dots \cdot (n - k + 1)$ Möglichkeiten ergeben (*). Wenn wir so wie oben verfahren und alle Möglichkeiten durchspielen werden wir allerdings mehrmals dieselbe Menge erhalten! Denn: Werden dieselben Elemente in einer unterschiedlichen Reihenfolge ausgewählt, so ergibt sich am Ende dieselbe Menge. Die Anzahl der verschiedenen Möglichkeiten in (*) muss also noch reduziert werden!

Um herauszufinden, wie oft bei der obigen Prozedur dieselben Mengen auftauchen, stellen wir folgende Frage: Wie viele Möglichkeiten hat man, k gegebene Elemente in eine

andere Reihenfolge zu bringen? Betrachten Sie dazu k feste Elemente und zählen Sie die Möglichkeiten, diese in eine andere Reihenfolge zu bringen. Man hat

- k Möglichkeiten, das erste Element zu wählen,
- $k - 1$ Möglichkeiten, das zweite Element zu wählen (denn das als erstes gewählte Element darf kein weiteres Mal gewählt werden),
- $k - 2$ Möglichkeiten, das dritte Element zu wählen,
- ...
- 2 Möglichkeiten, das $(k - 1)$ -ste Element zu wählen,
- 1 Möglichkeit, das k -te (und letzte) Element zu wählen.

Es ergeben sich insgesamt $k \cdot (k - 1) \cdot \dots \cdot 2 \cdot 1 = k!$ Möglichkeiten. Diese Überlegung zeigt, dass wir in der obigen Prozedur $k!$ -mal dieselbe Menge wählen. Um die Anzahl der wirklich verschiedenen Teilmengen rauszufinden muss man die Formel in (*) also durch $k!$ teilen.

$$\frac{n(n-1) \cdot \dots \cdot (n-k+1)}{k!} \stackrel{(64)}{=} \binom{n}{k}$$

Möglichkeiten.

^aEine Beweisskizze ist kein logisch bis ins letzte Detail schlüssiger Beweis, sondern vielmehr eine überzeugende Erklärung, warum die Aussage richtig ist. Für mich als Dozierender ist es manchmal aus pädagogischen Gründen zulässig, einen Beweis nicht komplett zu zeigen. Sie als Übende sollten sich aber nicht mit Skizzen zufrieden geben. Die Behauptungen, die wir so „bewiesen“ haben, können Sie aber in den Übungen verwenden.

Der Binomische Lehrsatz Die folgende Beobachtung, auch *Binomischer Lehrsatz* genannt, ist eine wichtige Zutat zur Bernoulli-Ungleichung (jedenfalls in der Version wie wir sie zunächst kennenlernen werden). Nachdem wir die Kombinatorik dahinter verstanden haben ist der Satz vielleicht gar nicht mehr so überraschend.

Wir werden hier wieder nur eine Beweisskizze mit kombinatorischen Überlegungen zeigen. Später werden wir aber nochmal einen anderen *wasserdichten* Beweis führen. Alle, die das Prinzip der *vollständigen Induktion* schon kennen, können sich schonmal an einem wasserdichten Beweis probieren.

Satz 2.2.18 — BINOMISCHER LEHRSAZ

Es seien $x, y \in \mathbb{R}$ und $n \in \mathbb{N}$. Dann gilt

$$(x + y)^n = \binom{n}{0} x^0 y^{n-0} + \binom{n}{1} x^1 y^{n-1} + \binom{n}{2} x^2 y^{n-2} + \dots + \binom{n}{n} x^n y^{n-n}.$$

Beweisskizze. Um $(x + y)^n$ zu berechnen muss man alle Terme miteinander ausmultiplizieren, (vgl. (62)f mit $y = 1$). Man erhält Summen aus Produkten mit n Faktoren, von denen alle jeweils x oder y sind. Hierbei muss man sämtliche Kombinationen betrachten. Fixiere nun ein $k \in \{0, \dots, n\}$. Man kann dann alle Produkte zusammenfassen bei denen der Faktor x k -mal genommen wird (und der Faktor y dadurch klarerweise $(n - k)$ -mal auftauchen muss).

Aus Vorüberlegung 2.2a/b lernen wir: Für jede k -elementige Teilmenge von $\{1, \dots, n\}$ (bestehend aus gerade den Ziehvorgängen, in denen man x genommen hat) erhält man einen Summanden der Form $x^k y^{n-k}$.

Da es nach Proposition 2.2.17 genau $\binom{n}{k}$ solcher Teilmengen gibt, erhalten wir einen Summanden der Form $\binom{n}{k} x^k y^{n-k}$.

Da $k \in \{0, \dots, n\}$ beliebig ist, müssen wir über alle k von 0 bis n summieren und erhalten

$$(x + y)^n = \underbrace{\binom{n}{0} x^0 y^{n-0}}_{k=0} + \underbrace{\binom{n}{1} x^1 y^{n-1}}_{k=1} + \underbrace{\binom{n}{2} x^2 y^{n-2}}_{k=2} + \dots + \underbrace{\binom{n}{n} x^n y^{n-n}}_{k=n}.$$

Korollar 2.2.19 — MÄCHTIGKEIT DER POTENZMENGE.

Es sei $n \in \mathbb{N}$. Die Menge

$$\mathcal{P}(\{1, \dots, n\}) := \{M : M \subseteq \{1, \dots, n\}\}$$

hat 2^n Elemente.

Beweisskizze. Wir müssen zählen, wie viele unterschiedliche Teilmengen es von $\{1, \dots, n\}$ gibt. Wir wissen aus Proposition 2.2.17 — Es gibt

- $\binom{n}{0}$ Teilmengen von $\{1, \dots, n\}$ mit 0 Elementen,
- $\binom{n}{1}$ Teilmengen von $\{1, \dots, n\}$ mit 1 Element,
- $\binom{n}{2}$ Teilmengen von $\{1, \dots, n\}$ mit 2 Elementen,
- ...
- $\binom{n}{n}$ Teilmengen von $\{1, \dots, n\}$ mit n Elementen.

Insgesamt gibt es also

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n}$$

Teilmengen von $\{1, \dots, n\}$. Nun gilt

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \dots + \binom{n}{n}$$

$$= \binom{n}{0} 1^0 1^{n-0} + \binom{n}{1} 1^1 1^{n-1} + \binom{n}{2} 1^2 1^{n-2} + \cdots + \binom{n}{n} 1^n 1^{n-n}$$

$$\xlongequal{\text{Satz 2.2.18}} (1+1)^n = 2^n.$$

Die Behauptung folgt.

Die Bernoulli-Ungleichung Nun sind wir in der Lage, die (für uns zunächst ausreichende erste Version der) *Bernoulli-Ungleichung* zu beweisen.

Korollar 2.2.20 — BERNOULLI-UNGLEICHUNG

Es sei $n \in \mathbb{N}$ und $x \geq 0$. Dann gilt

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

Wie schon angesprochen gilt die Ungleichung auch für $x \geq -1$. Das beweisen wir aber erst später.

Beweis. Es sei $x \geq 0$. Dann gilt

$$\begin{aligned} (1+x)^n &= (x+1)^n \xlongequal{\text{Satz 2.2.18}} \binom{n}{0} x^0 1^{n-0} + \binom{n}{1} x^1 1^{n-1} + \underbrace{\binom{n}{2} x^2 1^{n-2} + \cdots + \binom{n}{n} x^n 1^{n-n}}_{\geq 0 \text{ (da } x \geq 0)} \\ &\geq \binom{n}{0} x^0 1^{n-0} + \binom{n}{1} x^1 1^{n-1} + 0 = \binom{n}{0} x^0 1^{n-0} + \binom{n}{1} x^1 1^{n-1} \\ &= \binom{n}{0} + \binom{n}{1} x \xlongequal{\text{Prop 2.2.13}} 1+nx. \end{aligned}$$

Beispiel 2.2.21

Es gilt

$$2^n \geq n+1 \quad \forall n \in \mathbb{N}.$$

In der Tat: Für $n \in \mathbb{N}$ hat man

$$2^n = (1+1)^n \xlongequal{\text{Prop 2.2.20}} \geq 1+n \cdot 1 = 1+n = n+1.$$

Insbesondere majorisiert die Folge $(2^n)_{n \in \mathbb{N}}$ die Folge $(n+1)_{n \in \mathbb{N}}$.

Wachstumsvergleiche

In diesem Abschnitt wollen wir einige *Wachstumsvergleiche* für Folgen aus Beispiel 2.2.21 anstellen. Dazu zunächst ein Vergleich zwischen *polynomiellen Wachstum* und *exponentiellem Wachstum*.

Proposition 2.2.22 — POLYNOMIELLES UND EXPONENTIELLES WACHSTUM.

Es sei $a > 1$ und $\theta \geq 0$ sowie $\varepsilon > 0$. Dann gilt

$$(n^\theta)_{n \in \mathbb{N}} \leq (\varepsilon a^n)_{n \in \mathbb{N}}.$$

Beweis. Wähle ein beliebiges $\theta' > \theta$. Dann gilt

$$\varepsilon a^n = \varepsilon a^{\frac{1}{\theta'} n \theta'} \stackrel{\text{Potenzgesetze}}{=} \varepsilon \left[\left(a^{\frac{1}{\theta'}} \right)^n \right]^{\theta'} = \varepsilon \left[\left(1 + \left(a^{\frac{1}{\theta'}} - 1 \right) \right)^n \right]^{\theta'}.$$

Verwendung der Bernoulli-Ungleichung (Proposition 2.2.20) mit $x = (a^{\frac{1}{\theta'}} - 1)$ ergibt für alle $n \in \mathbb{N}$ ^a

$$\varepsilon a^n \stackrel{\text{Prop 2.2.20}}{\geq} \varepsilon \left[1 + n(a^{\frac{1}{\theta'}} - 1) \right]^{\theta'} \geq \varepsilon \left[0 + n(a^{\frac{1}{\theta'}} - 1) \right]^{\theta'} = \varepsilon n^{\theta'} (a^{\frac{1}{\theta'}} - 1)^{\theta'}. \quad (65)$$

Definiere nun $\varepsilon' := \varepsilon (a^{\frac{1}{\theta'}} - 1)^{\theta'}$. Da $a > 1$ ist, gilt $\varepsilon' > 0$. Dann besagt (65), dass

$$\varepsilon a^n \geq \varepsilon' n^{\theta'} \quad \forall n \in \mathbb{N}. \quad (66)$$

Nun gilt nach Proposition 2.2.6 (angewendet mit $\beta = \theta'$ und $\alpha = \theta$), dass $(n^\theta)_{n \in \mathbb{N}} \leq (\varepsilon' n^{\theta'})_{n \in \mathbb{N}}$. [Beachte: Hier benutzen wir, dass $\theta' > \theta$!] Das bedeutet: Es gibt ein $n_0 \in \mathbb{N}$ so, dass

$$\forall n \geq n_0 : \quad n^\theta \leq \varepsilon' n^{\theta'}.$$

Mit (66) folgt dann

$$\forall n \geq n_0 : \quad n^\theta \leq \varepsilon a^n.$$

Die Behauptung ist gezeigt.

^aDie Voraussetzung „ $x \geq 0$ “ ist erfüllt, weil $a > 1$ ist (und somit gilt auch $a^{\frac{1}{\theta'}} > 1$).

Beispiel 2.2.23

Nun ist auch endlich unser Eingangsbeispiel aus Sektion 2.2 fertig studiert: Mit $\theta = 8$, $a = 2$ und $\varepsilon = 1$ folgt

$$(n^8)_{n \in \mathbb{N}} \leq (2^n)_{n \in \mathbb{N}}.$$

Als Korollar aus dem vorigen Wachstumsvergleich auch können wir nun auch das exponentielle Wachstum mit verschiedenen Exponenten miteinander vergleichen.

Korollar 2.2.24 — EXPONENTIELLES WACHSTUM MIT VERSCHIEDENEN EXPONENTEN.

Es seien $0 < a < b$ sowie $\varepsilon > 0$. Dann gilt

$$(a^n)_{n \in \mathbb{N}} \leq (\varepsilon b^n)_{n \in \mathbb{N}}$$

Beweis. Es sei $0 < a < b$ und $\varepsilon > 0$. Dann gilt insbesondere $\frac{b}{a} > 1$. Benutzen wir Proposition 2.2.22 mit $\theta = 0$ so ergibt sich

$$(1)_{n \in \mathbb{N}} = (n^0)_{n \in \mathbb{N}} \leq \left(\varepsilon \left(\frac{b}{a}\right)^n\right)_{n \in \mathbb{N}} = \left(\frac{1}{a^n} \varepsilon b^n\right)_{n \in \mathbb{N}}.$$

Mit Proposition 2.2.10 (iv) [mit $a_n = 1, b_n = \frac{1}{a^n}$ und $c_n = \varepsilon b^n$] erhält man

$$(a^n)_{n \in \mathbb{N}} \leq (\varepsilon b^n)_{n \in \mathbb{N}}.$$

Eine weitere spannende Folge ist $(\sqrt[n]{n})_{n \in \mathbb{N}}$, d.h. $1, \sqrt{2}, \sqrt[3]{3}, \sqrt[4]{4}, \dots$. Beachte: $\sqrt[n]{n} \geq \sqrt[n]{1} = 1$ für alle $n \in \mathbb{N}$. Wir werden im nächsten Korollar aber sehen, dass viele Folgenglieder nicht allzu weit über der Zahl 1 liegen. 03.11.2025

Proposition 2.2.25 — FOLGE DER n -TEN WURZELN

Es sei $\varepsilon > 0$. Dann gilt

$$(i) \quad (\sqrt[n]{n})_{n \in \mathbb{N}} \leq (1 + \varepsilon)_{n \in \mathbb{N}}.$$

$$(ii) \quad \text{Für alle } q > 1 \text{ gilt } (\sqrt[q]{q})_{n \in \mathbb{N}} \leq (1 + \varepsilon)_{n \in \mathbb{N}}.$$

Beweis. Zu (i). Sei $\varepsilon > 0$. Da $(n)_{n \in \mathbb{N}} \leq ((1 + \varepsilon)^n)_{n \in \mathbb{N}}$ (vgl. Proposition 2.2.22 angewandt mit $a = 1 + \varepsilon > 1$) gilt

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad n \leq (1 + \varepsilon)^n. \quad (67)$$

Sei nun $n \geq n_0$, wobei n_0 wie in der obigen Gleichung gewählt ist. Ziehen wir die n -te Wurzel auf beiden Seiten der Ungleichung in (67), so folgt

$$\sqrt[n]{n} \leq (1 + \varepsilon) \quad \forall n \geq n_0.$$

Man folgert

$$\exists n_0 \in \mathbb{N} : \quad \sqrt[n]{n} \leq (1 + \varepsilon) \quad \forall n \geq n_0.$$

Dies zeigt die Behauptung.

Zu (ii). Ist den Lesenden als Übungsaufgabe überlassen. Tipp: Für alle $n \in \mathbb{N}$ mit $n \geq q$ gilt $\sqrt[q]{q} \leq \sqrt[n]{n}$.

Als nächstes stellen wir Vergleiche mit der *Fakultätsfolge* an.

Proposition 2.2.26 — MAJORISIERUNG UND FAKULTÄT

Es sei $\varepsilon > 0$ beliebig.

(i) Für alle $a > 0$ gilt $(a^n)_{n \in \mathbb{N}} \leq (\varepsilon n!)_{n \in \mathbb{N}}$.

(ii) Es gilt $(n!)_{n \in \mathbb{N}} \leq (\varepsilon n^n)_{n \in \mathbb{N}}$.

Beweis. Zu (i). Es sei $a > 0$. Wähle eine natürliche Zahl $m \in \mathbb{N}$ mit $m > a$. Dann gilt auch

$$\forall n \in \mathbb{N} : \quad a^n \leq m^n \quad (68)$$

Für alle $n \geq 2m + 1$ gilt

$$\begin{aligned} \varepsilon n! &= \varepsilon n \cdot \underbrace{(n-1) \cdot \dots \cdot (2m+1)}_{n-2m \text{ Faktoren}} \cdot (2m) \cdot (2m-1) \cdot \dots \cdot 1 \\ &\geq \varepsilon \underbrace{(2m) \cdot (2m) \cdot \dots \cdot (2m)}_{n-2m \text{ Faktoren}} \cdot (2m) \cdot (2m-1) \cdot \dots \cdot 1 \\ &= \varepsilon (2m)^{n-2m} (2m)! = \varepsilon 2^{n-2m} m^{n-2m} (2m)! = \varepsilon 2^{-2m} m^{-2m} (2m)! 2^n m^n. \end{aligned} \quad (69)$$

Definiere nun $\varepsilon' := \varepsilon 2^{-2m} m^{-2m} (2m)! > 0$. Dann gilt mit (69)

$$\forall n \geq 2m + 1 : \quad \varepsilon n! \geq \varepsilon' 2^n m^n \underset{(68)}{\geq} \varepsilon' 2^n a^n.$$

Mit anderen Worten: Für $\varepsilon' > 0$ wie oben definiert gilt

$$(\varepsilon n!)_{n \in \mathbb{N}} \geq (\varepsilon' 2^n a^n)_{n \in \mathbb{N}} \quad (70)$$

Beachte: Nach Proposition 2.2.22 (mit $a = 2, \theta = 0$) gilt $(\varepsilon' 2^n)_{n \in \mathbb{N}} \geq (1)_{n \in \mathbb{N}}$, und deswegen ist (nach Proposition 2.2.10(ii) mit $c_n = a^n$) auch

$$(\varepsilon' 2^n a^n)_{n \in \mathbb{N}} \geq (a^n)_{n \in \mathbb{N}}. \quad (71)$$

Zusammengefasst haben wir in (70) und (71) gesehen, dass

$$(a^n)_{n \in \mathbb{N}} \leq (\varepsilon' 2^n a^n)_{n \in \mathbb{N}} \quad \text{und} \quad (\varepsilon' 2^n a^n)_{n \in \mathbb{N}} \leq (\varepsilon n!)_{n \in \mathbb{N}}.$$

Mit Proposition 2.2.10(v) folgt

$$(a^n)_{n \in \mathbb{N}} \leq (\varepsilon n!)_{n \in \mathbb{N}}.$$

Die Behauptung (i) ist gezeigt.

Zu (ii). Für $n \geq 2$ gilt

$$n! = \underbrace{1 \cdot 2 \cdot 3 \cdots n}_{n \text{ Faktoren}} = \underbrace{2 \cdot 3 \cdots n}_{n-1 \text{ Faktoren}} \leq \underbrace{n \cdot n \cdots n}_{n-1 \text{ Faktoren}} = n^{n-1} = \frac{1}{n} n^n. \quad (72)$$

Da nun $(\frac{1}{n})_{n \in \mathbb{N}} \leq (\varepsilon)_{n \in \mathbb{N}}$ (vgl. Beispiel 2.2.11) gilt

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \frac{1}{n} \leq \varepsilon.$$

Für $n \geq \max\{n_0, 2\}$ gilt dann wegen (72)

$$n! \leq \frac{1}{n} n^n \leq \varepsilon n^n.$$

Wir folgern (mit der Wahl $\tilde{n}_0 := \max\{n_0, 2\}$)

$$\exists \tilde{n}_0 \in \mathbb{N} \quad \forall n \geq \tilde{n}_0 : \quad n! \leq \varepsilon n^n.$$

QED

2.3 Konvergenz und Grenzwerte

Ein Konzept, welches uns im gesamten weiteren Studium der Analysis begleiten wird, ist das Konzept der *Konvergenz*.

Definition 2.3.1 — KONVERGENZ VON FOLGEN

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine reelle Folge. Wir sagen,

- (i) $(a_n)_{n \in \mathbb{N}}$ *konvergiert gegen ein* $a \in \mathbb{R}$ falls

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |a_n - a| \leq \varepsilon \quad (73)$$

In Worten:

Zu jedem $\varepsilon > 0$ existiert ein
(i.A. von ε abhängiger) Index n_0 so,
dass für alle größeren Indizes $n \geq n_0$ gilt,
dass $|a_n - a| \leq \varepsilon$.

- (ii) $(a_n)_{n \in \mathbb{N}}$ ist *konvergent*, falls es ein $a \in \mathbb{R}$ gibt so, dass $(a_n)_{n \in \mathbb{N}}$ gegen a konvergiert.

Falls $(a_n)_{n \in \mathbb{N}}$ gegen a konvergiert, nennt man a einen *Grenzwert* von $(a_n)_{n \in \mathbb{N}}$ und schreibt

$$\lim_{n \rightarrow \infty} a_n = a.$$

Auch gebräuchlich sind

$$a_n \xrightarrow{n \rightarrow \infty} a \quad \text{oder} \quad [a_n \longrightarrow a \quad (n \rightarrow \infty)].$$

Bemerkung 2.3.2 — GEOMETRISCHE DEUTUNG VON KONVERGENZ

Eine reelle Folge $(a_n)_{n \in \mathbb{N}}$ ist *konvergent*^a gegen a genau dann wenn

$$\forall \varepsilon > 0 : \quad (|a_n - a|)_{n \in \mathbb{N}} \leq (\varepsilon)_{n \in \mathbb{N}} \quad (74)$$

(vgl. Definition 2.2.1). Äquivalent dazu ist nach Proposition 2.2.8, dass

$$\forall \varepsilon > 0 : \quad \{n \in \mathbb{N} : |a_n - a| > \varepsilon\} \quad \text{ist endlich.}$$

Weiterhin gilt: $|a_n - a| > \varepsilon \iff a_n \notin [a - \varepsilon, a + \varepsilon]$. Daher ist auch äquivalent zu (74)

$$\forall \varepsilon > 0 : \quad \{n \in \mathbb{N} : a_n \notin [a - \varepsilon, a + \varepsilon]\} \quad \text{ist endlich.}$$

Mit anderen Worten: Für jedes fixierte $\varepsilon > 0$ müssen *bis auf endlich viele Ausnahmen* alle

Folglieder in $[a - \varepsilon, a + \varepsilon]$ liegen.

^aBei dem geometrischen Verständnis von Konvergenz kann auch GEOGEBRA behilflich sein [LINK](#). Da gibt's auch eine interaktive Version von Abbildung 11...

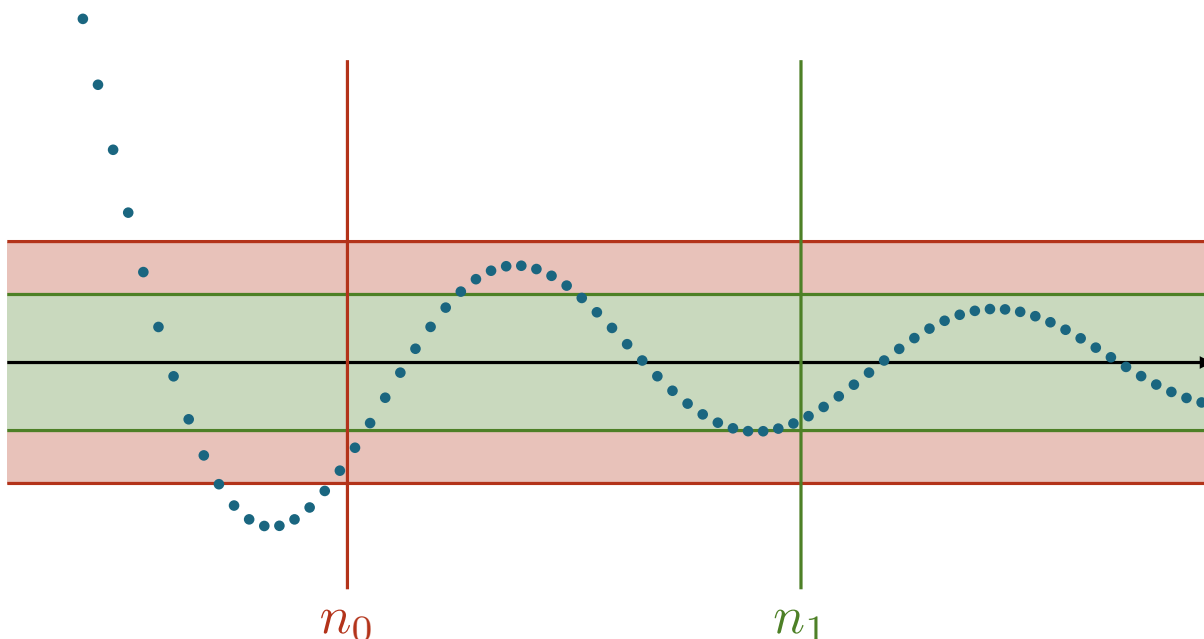


Abbildung 11: Eine gegen 0 konvergente Folge. Ab n_0 liegen alle Folgenglieder im großen roten ε -Schlauch. Ab n_1 liegen alle im kleineren grünen ε -Schlauch.

Beispiel 2.3.3

Behauptung. Die Folge $\left(\frac{1}{n}\right)_{n \in \mathbb{N}}$ konvergiert gegen 0.

Begründung.

$$\mathbb{Z}: \quad \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \left| \frac{1}{n} - 0 \right| \leq \varepsilon$$

Geben wir uns also ein beliebiges $\varepsilon > 0$ vor.

$$\mathbb{Z}: \quad \text{Es gibt } n_0 \in \mathbb{N} \text{ so, dass für alle } n \geq n_0 \text{ gilt } \left| \frac{1}{n} - 0 \right| < \varepsilon$$

Unsere Aufgabe ist es nun also, ein solches n_0 zu wählen. In diesem ersten Beispiel fällt die nachfolgende Wahl von n_0 *einfach vom Himmel*. Am Ende werden wir diskutieren, warum wir gerade diese Wahl getroffen haben. *Wähle.*

$$n_0 := \left\lceil \frac{1}{\varepsilon} \right\rceil. \quad (75)$$

Beachte: Nach Bemerkung 2.2.7 gilt $\lceil \frac{1}{\varepsilon} \rceil \geq \frac{1}{\varepsilon}$. Für $n \geq n_0$ gilt dann

$$|\frac{1}{n} - 0| = \frac{1}{n} \leq \frac{1}{n_0} = \frac{1}{\lceil \frac{1}{\varepsilon} \rceil} \leq \frac{1}{\frac{1}{\varepsilon}} = \varepsilon.$$

Wir haben also gezeigt

$$\forall n \geq n_0 : \quad |\frac{1}{n} - 0| \leq \varepsilon$$

Da $\varepsilon > 0$ beliebig war, folgern wir

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |\frac{1}{n} - 0| \leq \varepsilon.$$

Das heißt nach Definition 2.3.1 nichts anderes als

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0.$$

Bemerkung 2.3.4

Eine große Frage bleibt: Wie sind wir auf die Wahl in (75) gekommen? Dazu:

Sei $\varepsilon > 0$ vorgegeben.

Vorhaben: Finde einen Index $n_0 \in \mathbb{N}$ so, dass für alle $n \geq n_0$ gilt $|\frac{1}{n} - 0| \leq \varepsilon$.

Dazu folgende *Wolke*:

$$\begin{aligned} |\frac{1}{n} - 0| \leq \varepsilon & \iff \frac{1}{n} \leq \varepsilon \\ & \iff \frac{1}{\varepsilon} \leq n \\ & \iff n \geq \frac{1}{\varepsilon}. \end{aligned}$$

Wir lernen aus der Wolke: n_0 muss eine natürliche Zahl sein, die auch größer als $\frac{1}{\varepsilon}$ ist (damit $n \geq n_0$ auch direkt $|\frac{1}{n} - 0| < \varepsilon$ impliziert). Somit ist $n_0 = \lceil \frac{1}{\varepsilon} \rceil$ eine geeignete Wahl. Man hätte auch durchaus auch noch größere natürliche Zahlen als n_0 wählen dürfen — größer geht immer!

Beispiel 2.3.5

Behauptung. Die Folge $(\frac{n}{n+1})_{n \in \mathbb{N}}$ konvergiert gegen 1.

Begründung. Geben wir uns ein beliebiges $\varepsilon > 0$ vor.

$\mathbb{Z}$: Es gibt $n_0 \in \mathbb{N}$ so, dass $|\frac{n}{n+1} - 1| < \varepsilon$ für alle $n \geq n_0$.

Unsere Aufgabe ist es nun also, ein solches n_0 zu wählen. Dazu zunächst wieder eine kurze Ideensammlung in Form einer Wolke.

$$\begin{array}{lcl}
 |\frac{n}{n+1} - 1| \stackrel{!}{\leq} \varepsilon & \iff & |\frac{n}{n+1} - \frac{n+1}{n+1}| \stackrel{!}{\leq} \varepsilon \\
 & \iff & |-\frac{1}{n+1}| \stackrel{!}{\leq} \varepsilon \\
 & \iff & \frac{1}{n+1} \stackrel{!}{\leq} \varepsilon \\
 & \iff & \frac{1}{\varepsilon} \stackrel{!}{\leq} n+1 \\
 & \iff & n \stackrel{!}{\geq} \frac{1}{\varepsilon} - 1.
 \end{array}$$

Wähle.

$$n_0 := \max\{1, \lceil \frac{1}{\varepsilon} - 1 \rceil\} \quad (a)$$

Für $n \geq n_0$ gilt dann:

FALL 1. $\varepsilon < 1$ (d.h. $n_0 = \lceil \frac{1}{\varepsilon} - 1 \rceil$).

$$|\frac{n}{n+1} - 1| = |\frac{n}{n+1} - \frac{n+1}{n+1}| = |-\frac{1}{n+1}| = \frac{1}{n+1} \leq \frac{1}{n_0+1} = \frac{1}{\lceil \frac{1}{\varepsilon} - 1 \rceil + 1} \leq \frac{1}{\frac{1}{\varepsilon}} = \varepsilon.$$

FALL 2. $\varepsilon \geq 1$ (d.h. $n_0 = 1$).

$$|\frac{n}{n+1} - 1| = |\frac{n}{n+1} - \frac{n+1}{n+1}| = \frac{1}{n+1} \leq \frac{1}{n_0+1} = \frac{1}{1+1} \leq \frac{1}{2} < 1 \leq \varepsilon.$$

Wir haben also in allen Fällen gezeigt

$$\forall n \geq n_0 : \quad |\frac{n}{n+1} - 1| \leq \varepsilon.$$

Da $\varepsilon > 0$ beliebig war folgern wir

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |\frac{n}{n+1} - 1| \leq \varepsilon.$$

Das heißt nach Definition 2.3.1 nichts anderes als

$$\lim_{n \rightarrow \infty} \frac{n}{n+1} = 1.$$

^aIn der Wahl steht $\lceil \frac{1}{\varepsilon} - 1 \rceil$, weil die Zahl größergleich als $\frac{1}{\varepsilon} - 1$ sein muss, siehe Wolke. Das Maximum wird gewählt, weil es sich stets um eine natürliche Zahl handeln muss. Beachte: Für $\varepsilon \geq 1$ wäre $\lceil \frac{1}{\varepsilon} - 1 \rceil \notin \mathbb{N}$, daher wird das Maximum wirklich benötigt.

Beispiel 2.3.6 — KONSTANTE FOLGEN.

Es sei $a \in \mathbb{R}$ beliebig. Definiere für alle $n \in \mathbb{N}$ $a_n := a$. Die konstante Folge $(a_n)_{n \in \mathbb{N}}$ (d.h. $a, a, a, a, \dots$) konvergiert dann klarerweise gegen a . Dies wollen wir hier beweisen

Beweis. Sei $\varepsilon > 0$ beliebig vorgegeben. Wähle $n_0 := 1$. Für $n \geq n_0$ gilt

$$|a_n - a| = |a - a| = 0 \leq \varepsilon.$$

Wir haben gesehen

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |a_n - a| \leq \varepsilon.$$

Mit Definition 2.3.1 folgt sofort $a_n \xrightarrow{n \rightarrow \infty} a$.

Gleichung (74) verbindet Konvergenzaussagen mit Majorisierungsaussagen. Dies liefert einige nützliche Grenzwert-Resultate für für uns interessante Folgen.

Proposition 2.3.7 — SPEZIELLE FOLGENGRENZWERTE.

- (i) Für alle $\theta > 0$ gilt $\lim_{n \rightarrow \infty} \frac{1}{n^\theta} = 0$.
- (ii) Für alle $\theta \geq 0$ und $a > 1$ gilt $\lim_{n \rightarrow \infty} \frac{n^\theta}{a^n} = 0$.
- (iii) Es gilt $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$.
- (iv) Für alle $a \in \mathbb{R}$ gilt $\lim_{n \rightarrow \infty} \frac{a^n}{n!} = 0$.
- (v) Es gilt $\lim_{n \rightarrow \infty} \frac{n!}{n^n} = 0$.

Beweis. Wir zeigen (ii) und (iii). Die anderen Behauptungen folgen demselben Beweisprinzip (unter Verwendung der Propositionen 2.2.22, 2.2.25, 2.2.26)

Zu (ii). Sei $\varepsilon > 0$ beliebig vorgegeben. Nach Proposition 2.2.22 ist

$$(n^\theta)_{n \in \mathbb{N}} \leq (\varepsilon a^n)_{n \in \mathbb{N}}.$$

Das heißt

$$\exists n_0 \in \mathbb{N} \forall n \geq n_0 : \quad n^\theta \leq \varepsilon a^n.$$

Man folgert

$$\exists n_0 \in \mathbb{N} \forall n \geq n_0 : \quad \frac{n^\theta}{a^n} \leq \varepsilon.$$

Äquivalent

$$\exists n_0 \in \mathbb{N} \forall n \geq n_0 : \quad \left| \frac{n^\theta}{a^n} - 0 \right| \leq \varepsilon.$$

Man hat also (weil $\varepsilon > 0$ beliebig war)

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \left| \frac{n^\theta}{a^n} - 0 \right| \leq \varepsilon.$$

Das heißt nach Definition 2.3.1 nichts anderes als

$$\lim_{n \rightarrow \infty} \frac{n^\theta}{a^n} = 0.$$

Zu (iii). Sei $\varepsilon > 0$ beliebig vorgegeben. Nach Proposition 2.2.25 gilt

$$(\sqrt[n]{n})_{n \in \mathbb{N}} \leq (1 + \varepsilon)_{n \in \mathbb{N}}.$$

Das bedeutet

$$\exists n_0 \in \mathbb{N} \forall n \geq n_0 : \quad \sqrt[n]{n} \leq 1 + \varepsilon.$$

Ferner gilt auch für alle $n \in \mathbb{N}$, dass $\sqrt[n]{n} \geq \sqrt[n]{1} = 1$. Das bedeutet

$$\exists n_0 \in \mathbb{N} \forall n \geq n_0 : \quad 1 \leq \sqrt[n]{n} \leq 1 + \varepsilon$$

Äquivalent

$$\exists n_0 \in \mathbb{N} \forall n \geq n_0 : \quad 0 \leq \sqrt[n]{n} - 1 \leq \varepsilon.$$

Sicherlich hat jede Zahl zwischen 0 und ε einen Betrag, der kleiner gleich ε ist. Es folgt

$$\exists n_0 \in \mathbb{N} : \quad \left| \sqrt[n]{n} - 1 \right| \leq \varepsilon \quad \forall n \geq n_0.$$

Weil $\varepsilon > 0$ beliebig war, gilt

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad \left| \sqrt[n]{n} - 1 \right| \leq \varepsilon \quad \forall n \geq n_0.$$

Das heißt nach Definition 2.3.1 nichts anderes als

$$\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1.$$

Beispiel 2.3.8

Es sei $x \in \mathbb{R}$ mit $0 \leq x < 1$. Dann gilt

$$\lim_{n \rightarrow \infty} x^n = 0.$$

Beweis.

FALL 1. $x = 0$. Da $0^n = 0$ für alle $n \in \mathbb{N}$, d.h. die Folge ist die konstante Folge $(0)_{n \in \mathbb{N}}$ und konvergiert bereits nach Beispiel 2.3.6 gegen 0.

FALL 2. $0 < x < 1$. Definiere nun $a := \frac{1}{x}$ und beobachte, dass $a > \frac{1}{1} = 1$.

$$x^n = \frac{1}{\frac{1}{x^n}} = \frac{1}{\left(\frac{1}{x}\right)^n} = \frac{n^0}{a^n} \xrightarrow[n \rightarrow \infty]{\text{Proposition 2.3.7(ii)}} 0.$$

Man beachte, dass die Voraussetzungen von Proposition 2.3.7 nur erfüllt sind weil $a > 1$.

06.11.2025

Es gibt auch Folgen $(a_n)_{n \in \mathbb{N}}$, die nicht konvergieren! Für diese macht die Schreibweise $\lim_{n \rightarrow \infty} a_n$ keinen Sinn, denn der Grenzwert ist gar nicht definiert (vgl. Definition 2.3.1)!

Definition 2.3.9 — DIVERGENZ

Eine Folge, die nicht konvergent ist, heißt *divergent*.

Divergenz ist (ohne Weiteres) nicht so leicht nachzuweisen wie Konvergenz! Während man beim Nachweis der Konvergenz stets einen Kandidaten für einen Grenzwert vor Augen hat, muss man beim Nachweis der Divergenz beweisen, dass kein Wert $a \in \mathbb{R}$ als Grenzwert in Frage kommt. Eine Herausforderung der wir uns nun stellen wollen!

Beispiel 2.3.10

Behauptung. Die Folge $((-1)^n)_{n \in \mathbb{N}}$ (d.h. $-1, 1, -1, 1, \dots$) ist divergent.

Beweis. Angenommen $((-1)^n)_{n \in \mathbb{N}}$ ist konvergent. Dann gibt es ein $a \in \mathbb{R}$ so, dass

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |(-1)^n - a| \leq \varepsilon.$$

Betrachtet man nun den Spezialfall $\varepsilon = \frac{1}{2}$, so gilt

$$\exists n^* \in \mathbb{N} \quad \forall n \geq n^* : \quad |(-1)^n - a| \leq \frac{1}{2}. \quad (76)$$

Dann gilt

$$\begin{aligned} \frac{1}{2} &\stackrel{(76)}{\geq} |(-1)^{n^*+1} - a| = |(-1)^{n^*+1} - (-1)^{n^*} + (-1)^{n^*} - a| \\ &\stackrel{(41)}{\geq} ||(-1)^{n^*+1} - (-1)^{n^*}| - |(-1)^{n^*} - a|| \\ &\geq |(-1)^{n^*+1} - (-1)^{n^*}| - \underbrace{|(-1)^{n^*} - a|}_{\leq \frac{1}{2}, \text{vgl. (76)}} \end{aligned}$$

$$\begin{aligned} &\geq |(-1)^{n^*+1} - (-1)^{n^*}| - \frac{1}{2} \\ &= |(-1)^{n^*}(-1-1)| - \frac{1}{2} = |(-1)^{n^*}| |(-2)| - \frac{1}{2} = 2 - \frac{1}{2} = \frac{3}{2}. \end{aligned}$$

Wir erhalten den Widerspruch $\frac{1}{2} \geq \frac{3}{2}$. Daher muss die Annahme falsch sein. Die Annahme war nun aber, dass $((-1)^n)_{n \in \mathbb{N}}$ konvergent ist. Wir haben also die Divergenz gezeigt.

Bemerkung 2.3.11 — VERNEINUNG DER GRENZWERTDEFINITION.

Zum besseren Verständnis der Definition 2.3.1 bilden wir mal die logische Verneinung von (73): Eine Folge $(a_n)_{n \in \mathbb{N}}$ *konvergiert nicht gegen ein* $a \in \mathbb{R}$ falls

$$\exists \varepsilon_0 > 0 : \quad \forall m \in \mathbb{N} \quad \exists N \geq m : \quad |a_N - a| > \varepsilon_0. \quad (77)$$

Wir verweisen hier auf Beispiel 2.2.5 und die Diskussion davor.

Man sollte hier auf Folgendes hinweisen: Nur weil eine Folge *nicht gegen ein gegebenes* $a \in \mathbb{R}$ konvergiert, bedeutet das nicht, dass die Folge nicht konvergent ist! Denn schließlich könnte die Folge ja doch noch gegen ein anderes $a' \in \mathbb{R}$ mit $a' \neq a$ konvergieren ...

Um die Divergenz einer Folge zu beweisen, muss man (77) also gleich für sämtliche $a \in \mathbb{R}$ nachweisen. Mit anderen Worten müssten wir zeigen

$$\forall a \in \mathbb{R} \quad \exists \varepsilon_0 > 0 : \quad \forall m \in \mathbb{N} \quad \exists N \geq m : \quad |a_N - a| > \varepsilon_0.$$

Erste Eigenschaften von Grenzwerten Hier diskutieren wir einige Eigenschaften von Grenzwerten, die man direkt aus der Definition folgern kann.

Eine wichtige Eigenschaft ist, dass man bei einer konvergenten Folge stets endlich viele Folgenglieder abändern kann, ohne dabei das Konvergenzverhalten zu verändern.

Proposition 2.3.12 — KONVERGENZ UNTER ABÄNDERUNG ENDLICH VIELER FOLGEGLIEDER.

Es seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ Folgen so, dass

$$\exists n^* \in \mathbb{N} \quad \forall n \geq n^* : \quad a_n = b_n \quad (78)$$

Dann gilt

$$a_n \xrightarrow{n \rightarrow \infty} a \quad \implies \quad b_n \xrightarrow{n \rightarrow \infty} a.$$

Beweis. Es sei $a_n \xrightarrow{n \rightarrow \infty} a$, d.h. (nach Definition 2.3.1)

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |a_n - a| \leq \varepsilon. \quad (79)$$

Sei nun $\varepsilon > 0$ beliebig vorgegeben und $n_0 \in \mathbb{N}$ gewählt wie in (79).

$$\mathbb{Z} : \quad \exists n_1 \in \mathbb{N} \quad \forall n \geq n_1 : \quad |b_n - a| \leq \varepsilon.$$

Dazu: Wähle $n_1 := \max\{n^*, n_0\}$. Dann gilt für $n \geq n_1$ (d.h. $n \geq n^*$ und $n \geq n_0$)

$$|b_n - a| \stackrel{(78)}{=} |a_n - a| \stackrel{(79)}{\leq}_{n \geq n_0} \varepsilon.$$

Wir haben gesehen:

$$\forall \varepsilon > 0 \quad \exists n_1 \in \mathbb{N} \quad \forall n \geq n_1 : \quad |b_n - a| \leq \varepsilon.$$

QED

Das bedeutet insbesondere, dass uns die ersten endlich vielen Folgenglieder für die Konvergenz irrelevant sind. Wir werden daher auch manchmal von der Konvergenz der Folge $(a_n)_{n \geq N}$ (für $N \in \mathbb{N}$) sprechen, falls es angenehmer ist, die Folgenglieder erst ab einem gewissen Index $N \in \mathbb{N}$ zu definieren. Die ersten $N - 1$ Folgenglieder darf man sich dann frei ausdenken — Für die Konvergenzfrage spielen sie keine Rolle!

Eine weitere wichtige Eigenschaft ist die *Eindeutigkeit* des Grenzwertes. Das bedeutet: Eine Folge besitzt maximal einen Grenzwert. (Wie oben schon gezeigt ist es aber durchaus auch möglich dass sie gar keinen besitzt!). Eigentlich hätten wir das schon viel früher diskutieren müssen, denn wenn eine Folge verschiedene Grenzwerte haben könnte, wären Formeln für Grenzwerte wie in 2.3.7 ziemlich nutzlos – sie würden ja schließlich dann nur einen möglichen Grenzwert aufzeigen... Die Notation $\lim_{n \rightarrow \infty} a_n = a$ würde ohne die Eindeutigkeit auch keinen Sinn machen. Man dürfte sie eigentlich erst nach gezeigter Eindeutigkeit einführen.

Proposition 2.3.13 — EINDEUTIGKEIT DES GRENZWERTES.

Es sei $(a_n)_{n \in \mathbb{N}}$ eine Folge und $a, b \in \mathbb{R}$. Dann gilt

$$(a_n \xrightarrow{n \rightarrow \infty} a) \quad \wedge \quad (a_n \xrightarrow{n \rightarrow \infty} b) \quad \implies \quad a = b.$$

Beweis. Die Voraussetzung impliziert (vgl. Definition 2.3.1)

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |a_n - a| \leq \varepsilon. \quad (80)$$

und

$$\forall \varepsilon > 0 \quad \exists n_1 \in \mathbb{N} \quad \forall n \geq n_1 : \quad |a_n - b| \leq \varepsilon. \quad (81)$$

Angenommen, dass $a \neq b$. Dann gilt $\frac{|a-b|}{3} > 0$. Betrachten wir nun in den beiden obigen Gleichungen den Fall $\varepsilon = \frac{|a-b|}{3}$. Wir wissen aus (80)

$$\exists n_0^* \in \mathbb{N} \quad \forall n \geq n_0^* : \quad |a_n - a| \leq \frac{|a-b|}{3}. \quad (82)$$

und aus (81)

$$\exists n_1^* \in \mathbb{N} \quad \forall n \geq n_1^* : \quad |a_n - b| \leq \frac{|a-b|}{3}. \quad (83)$$

Setze nun $n^* := \max\{n_0^*, n_1^*\}$. Dann gilt mit der Dreiecksungleichung (Satz 2.0.3 (iii))

$$|a - b| = |a - a_{n^*} + a_{n^*} - b| = |(a - a_{n^*}) + (a_{n^*} - b)| \underset{\text{Satz 2.0.3(iii)}}{\leq} |a - a_{n^*}| + |a_{n^*} - b|. \quad (84)$$

Da $n^* \geq n_0^*$ gilt $|a - a_{n^*}| = |a_{n^*} - a| \leq \frac{|a-b|}{3}$, vgl. (82). Da außerdem $n^* \geq n_1^*$ gilt $|a_{n^*} - b| \leq \frac{|a-b|}{3}$, vgl. (83). Wir erhalten mit (84)

$$|a - b| \leq |a - a_{n^*}| + |a_{n^*} - b| \leq \frac{|a-b|}{3} + \frac{|a-b|}{3} = \frac{2}{3}|a - b|.$$

Betrachten wir die obige Formel als Ungleichung und ziehen auf beiden Seiten $\frac{2}{3}|a - b|$ ab so folgt

$$\frac{|a-b|}{3} \leq 0.$$

Dies impliziert $a = b$. Man erhält einen Widerspruch zur Annahme $a \neq b$. Die Behauptung ist gezeigt.

Ein *notwendiges Kriterium* für die Konvergenz einer Folge ist deren *Beschränktheit*.

Proposition 2.3.14 — KONVERGENTE FOLGEN SIND BESCHRÄNKT

Es sei $(a_n)_{n \in \mathbb{N}}$ eine konvergente Folge. Dann ist $(a_n)_{n \in \mathbb{N}}$ beschränkt.

Beweis. Es sei $(a_n)_{n \in \mathbb{N}}$ konvergent, etwa gelte $a = \lim_{n \rightarrow \infty} a_n$. Dann gilt nach Definition 2.3.1

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |a_n - a| \leq \varepsilon. \quad (85)$$

Wir benutzen nun Proposition 2.1.6, um die Beschränktheit zu zeigen. Das heißt wir zeigen folgendes

$$\mathbb{Z} : \quad \exists K > 0 \quad \forall n \in \mathbb{N} : \quad |a_n| \leq K.$$

Benutzen wir (85) mit $\varepsilon = 1$ so finden wir

$$\exists n_0^* \in \mathbb{N} \quad \forall n \geq n_0^* : \quad |a_n - a| \leq 1. \quad (86)$$

Für alle $n \geq n_0^*$ gilt dann

$$|a_n| = |a_n - a + a| = |(a_n - a) + a| \leq |a_n - a| + |a| \underset{(86)}{\leq} 1 + |a|. \quad (87)$$

Definiere nun

$$K := |a_1| + \cdots + |a_{n_0^*}| + (1 + |a|). \quad (88)$$

Wir behaupten nun, dass für alle $n \in \mathbb{N}$ gilt $|a_n| \leq K$.

Dazu:

FALL 1. $n < n_0^*$. Dann gilt

$$|a_n| \leq |a_1| + \cdots + |a_{n_0^*}| \underset{(88)}{\leq} K.$$

FALL 2. $n \geq n_0^*$. Dann gilt

$$|a_n| \underset{(87)}{\leq} 1 + |a| \underset{(88)}{\leq} K.$$

Die Behauptung folgt.

Bemerkung 2.3.15

Wir haben in der vorigen Proposition gezeigt

$$(a_n)_{n \in \mathbb{N}} \text{ konvergent} \implies (a_n)_{n \in \mathbb{N}} \text{ beschränkt.}$$

Die (nach Proposition 1.1.7(ii) dazu äquivalente) Kontraposition besagt

$$(a_n)_{n \in \mathbb{N}} \text{ unbeschränkt} \implies (a_n)_{n \in \mathbb{N}} \text{ divergent.}$$

Jede unbeschränkte Folge ist also divergent. Damit kann man zeigen, dass die Folgen $(n)_{n \in \mathbb{N}}$, $((-1)^n n)_{n \in \mathbb{N}}$, etc. etc. divergent sind.

Beachten Sie: Jede konvergente Folge ist beschränkt, aber nicht jede beschränkte Folge ist auch konvergent (vgl. Beispiel 2.3.10)!

Grenzwertregeln

Im nachfolgenden Abschnitt lernen wir gewisse Regeln kennen, die uns erlauben, aus gegebenen konvergenten Folgen weitere konvergente Folgen zu basteln, z.B. durch *arithmetische Operationen*. Solche Regeln nennen wir *Grenzwertregeln* (abgekürzt mit GWR).

Rechenregeln Zunächst beschäftigen wir uns mit Regeln für Summen, Produkten und Quotienten von konvergenten Folgen.

Grenzwertregel 1 — SUMMEN KONVERGENTER FOLGEN

Es seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ zwei konvergente Folgen. Dann ist $(a_n + b_n)_{n \in \mathbb{N}}$ eine konvergente Folge und es gilt

$$\lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n.$$

Beweis. Es sei

$$a := \lim_{n \rightarrow \infty} a_n \quad b := \lim_{n \rightarrow \infty} b_n, \tag{89}$$

d.h. $a_n \xrightarrow{n \rightarrow \infty} a$ und $b_n \xrightarrow{n \rightarrow \infty} b$. Wir formulieren die Aussage der Proposition mit der soeben

eingeführten Notation um:

$$\mathbb{Z} : \quad (a_n + b_n) \xrightarrow{n \rightarrow \infty} (a + b).$$

Es sei nun $\varepsilon > 0$ beliebig vorgegeben. Dann ist auch $\varepsilon' := \frac{\varepsilon}{2} > 0$. Mit (73)^a (angewendet mit ε' statt ε) finden wir

$$\exists n_1 \in \mathbb{N} \quad \forall n \geq n_1 : \quad |a_n - a| \leq \frac{\varepsilon}{2} \quad (90)$$

und

$$\exists n_2 \in \mathbb{N} \quad \forall n \geq n_2 : \quad |b_n - b| \leq \frac{\varepsilon}{2}. \quad (91)$$

Setze nun

$$n_0 := \max\{n_1, n_2\}.$$

Für alle $n \geq n_0$ (d.h. auch $n \geq n_1$ und $n \geq n_2$) gilt dann

$$|(a_n + b_n) - (a + b)| = |(a_n - a) + (b_n - b)| \underset{\text{Satz 2.0.3(iii)}}{\leq} |a_n - a| + |b_n - b| \overset{(90)}{\underset{(91)}{\leq}} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

(Beachte: Bei der Benutzung von (90) und (91) haben wir auch verwendet, dass $n \geq n_1$ und $n \geq n_2$!). Wir haben also gesehen

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |(a_n + b_n) - (a + b)| \leq \varepsilon.$$

Da $\varepsilon > 0$ beliebig war erhalten wir

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad |(a_n + b_n) - (a + b)| \leq \varepsilon \quad \forall n \geq n_0.$$

Das bedeutet gerade nach Definition 2.3.1, dass $(a_n + b_n) \xrightarrow{n \rightarrow \infty} (a + b)$. Es folgt

$$\lim_{n \rightarrow \infty} (a_n + b_n) = a + b \underset{(89)}{=} \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n.$$

^aMan profitiert an dieser Stelle massivst davon, dass man in (73) jede beliebige positive Zahl anstelle von ε einsetzen kann — ebenso auch das oben gewählte $\varepsilon' = \frac{\varepsilon}{2} > 0$! Das ist ein wichtiger Teil der Mechanik der Definition der Folgenkonvergenz!

Bemerkung 2.3.16

Eine unabdingbare Voraussetzung bei der vorigen Proposition ist die *Konvergenz der Folgen* $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$!

Über die Summe zweier *divergenter* Folgen kann man nämlich *keine Aussage* treffen! Eine solche Summe kann *sowohl konvergent als auch divergent* sein.

Beispiel 2.3.17

$((-1)^n)_{n \in \mathbb{N}}$ und $((-1)^{n+1})_{n \in \mathbb{N}}$ sind beide divergent (vgl. Beispiel 2.3.10).

Jedoch ist

$$((-1)^n + (-1)^{n+1})_{n \in \mathbb{N}} = ((-1)^n - (-1)^n)_{n \in \mathbb{N}} = (0)_{n \in \mathbb{N}}$$

konvergent (gegen 0), vgl. Beispiel 2.3.6.

Grenzwertregel 2 — PRODUKTE KONVERGENTER FOLGEN

Es seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ zwei konvergente Folgen. Dann ist $(a_n b_n)_{n \in \mathbb{N}}$ eine konvergente Folge und es gilt

$$\lim_{n \rightarrow \infty} (a_n b_n) = \left(\lim_{n \rightarrow \infty} a_n \right) \cdot \left(\lim_{n \rightarrow \infty} b_n \right).$$

Beweis. Es sei wiederum

$$a := \lim_{n \rightarrow \infty} a_n \quad b := \lim_{n \rightarrow \infty} b_n, \quad (92)$$

d.h. $a_n \xrightarrow{n \rightarrow \infty} a$ und $b_n \xrightarrow{n \rightarrow \infty} b$.

$$\mathbb{Z}\mathbb{L}: \quad a_n b_n \xrightarrow{n \rightarrow \infty} ab.$$

Vorab: $(a_n)_{n \in \mathbb{N}}$ ist als konvergente Folge beschränkt (vgl. Proposition 2.3.14). Somit finden wir (nach Proposition 2.1.6) ein $K \geq 0$ so, dass

$$\forall n \in \mathbb{N}: \quad |a_n| \leq K. \quad (93)$$

Es sei nun $\varepsilon > 0$ beliebig vorgegeben. Nun gilt mit (73) (angewendet mit $\varepsilon' := \frac{\varepsilon}{2(|b|+1)} > 0$)

$$a_n \xrightarrow{n \rightarrow \infty} a \quad \implies \quad \exists n_1 \in \mathbb{N} \quad \forall n \geq n_1: \quad |a_n - a| \leq \frac{\varepsilon}{2(|b|+1)}. \quad (94)$$

und wiederum mit (73) (angewendet mit $\varepsilon' := \frac{\varepsilon}{2(K+1)} > 0$)

$$b_n \xrightarrow{n \rightarrow \infty} b \quad \implies \quad \exists n_2 \in \mathbb{N} \quad \forall n \geq n_2: \quad |b_n - b| \leq \frac{\varepsilon}{2(K+1)}. \quad (95)$$

Definiere abermals

$$n_0 := \max\{n_1, n_2\}. \quad (96)$$

Für $n \geq n_0$ gilt dann

$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - a_n b + a_n b - ab| = |a_n(b_n - b) + b(a_n - a)| \\ &\stackrel{\text{Satz 2.0.3(iii)}}{\leq} |a_n(b_n - b)| + |b(a_n - a)| \stackrel{\text{Satz 2.0.3(ii)}}{=} |a_n| \cdot |b_n - b| + |b| \cdot |a_n - a| \\ &\stackrel{(93)}{\leq} K \cdot |b_n - b| + |b| \cdot |a_n - a| \stackrel{(94)}{\stackrel{(95)}}{\leq} K \frac{\varepsilon}{2(K+1)} + |b| \frac{\varepsilon}{2(|b|+1)} \\ &= \frac{\varepsilon}{2} \frac{K}{K+1} + \frac{\varepsilon}{2} \frac{|b|}{|b|+1} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

(Beachte: (94) und (95) konnten nur benutzt werden weil $n \geq n_0$, d.h. auch $n \geq n_1$ und $n \geq n_2$, vgl. (96)).

Da $\varepsilon > 0$ beliebig war, haben wir nachgewiesen

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |a_n b_n - ab| \leq \varepsilon.$$

Es folgt die behauptete Konvergenz, d.h.

$$\lim_{n \rightarrow \infty} (a_n b_n) = ab \stackrel{(92)}{=} \left(\lim_{n \rightarrow \infty} a_n \right) \cdot \left(\lim_{n \rightarrow \infty} b_n \right).$$

Korollar 2.3.18 — KONSTANTE VIELFACHE VON FOLGEN.

Es sei $(a_n)_{n \in \mathbb{N}}$ eine konvergente Folge und $\lambda \in \mathbb{R}$. Dann ist auch $(\lambda a_n)_{n \in \mathbb{N}}$ eine konvergente Folge und es gilt

$$\lim_{n \rightarrow \infty} (\lambda a_n) = \lambda \lim_{n \rightarrow \infty} a_n.$$

Beweis. Wende die vorige Proposition mit $(b_n)_{n \in \mathbb{N}} = (\lambda)_{n \in \mathbb{N}}$ an.

Korollar 2.3.19 — LINEARKOMBINATIONEN VON FOLGEN.

Es seien $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}$ konvergente Folgen und $\lambda, \mu \in \mathbb{R}$. Dann ist auch $(\lambda a_n + \mu b_n)_{n \in \mathbb{N}}$ eine konvergente Folge und es gilt

$$\lim_{n \rightarrow \infty} (\lambda a_n + \mu b_n) = \lambda \lim_{n \rightarrow \infty} a_n + \mu \lim_{n \rightarrow \infty} b_n.$$

Beweis. Verwende das vorige Korollar und Grenzwertregel 1.

Die dritte Grenzwertregel beschäftigt sich mit Quotienten konvergenter Folgen. Hierbei ist besondere Vorsicht geboten: Die untenstehende Regel ist nicht anwendbar, wenn die Folge der Nenner den Grenzwert 0 besitzt.

Grenzwertregel 3 — QUOTIENTEN KONVERGENTER FOLGEN

Es seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ konvergente Folgen mit $\lim_{n \rightarrow \infty} b_n \neq 0$. Dann gilt

(i) Es existiert ein $N \in \mathbb{N}$ mit $b_n \neq 0$ für alle $n \geq N$.

(ii) Die Folge $(\frac{a_n}{b_n})_{n \geq N}$ ist konvergent und es gilt

$$\lim_{n \rightarrow \infty} \left(\frac{a_n}{b_n} \right) = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n}. \quad (a)$$

^aDa man die Folge auf der linken Seite erst ab dem Index N betrachten darf (weil man ja sonst eventuell durch 0 teilen würde) schreiben hier manche Bücher auch „ $\lim_{n \rightarrow \infty, n \geq N} (\frac{a_n}{b_n})$ “. Wir werden das hier nicht aktiv

verwenden: Da wegen Proposition 2.3.12 die ersten $N - 1$ Folgenglieder sowieso irrelevant sind, müssen wir die Notation an dieser Stelle nicht unnötig verkomplizieren.

Beweis. Zu (i). Es sei

$$b := \lim_{n \rightarrow \infty} b_n \quad [\neq 0 \text{ laut Voraussetzung}].$$

Dann gilt $|b| > 0$ und somit auch $\varepsilon' := \frac{|b|}{2} > 0$. Da $b_n \xrightarrow{n \rightarrow \infty} b$ haben wir

$$\exists N \in \mathbb{N} \quad \forall n \geq N : \quad |b_n - b| \leq \frac{|b|}{2} \quad (97)$$

Für $n \geq N$ gilt dann

$$|b_n| = |b + b_n - b| = |b - (b - b_n)| \underset{(41)}{\geq} ||b| - |b - b_n|| \geq |b| - |b - b_n| \underset{(97)}{\geq} |b| - \frac{|b|}{2} = \frac{|b|}{2}. \quad (98)$$

Man folgert aus der vorigen Gleichung, dass $|b_n| \geq \frac{|b|}{2} > 0$ für alle $n \geq N$. Das impliziert $b_n \neq 0$ für alle $n \geq N$.

Zu (ii). Wir untersuchen zunächst einen Spezialfall und folgern aus diesem dann den allgemeinen Fall.

SPEZIALFALL. $a_n = 1 \quad \forall n \in \mathbb{N}$.

Es sei $b := \lim_{n \rightarrow \infty} b_n \quad [\neq 0 \text{ laut Voraussetzung}]$.

$$\mathbb{Z} : \quad \frac{1}{b_n} \xrightarrow{n \rightarrow \infty} \frac{1}{b}.$$

Sei dazu $\varepsilon > 0$ beliebig vorgegeben. Wie schon in (98) leiten wir her

$$\exists N \in \mathbb{N} \quad \forall n \geq N : \quad |b_n| \geq \frac{|b|}{2}. \quad (99)$$

Da $b_n \xrightarrow{n \rightarrow \infty} b$ und $\varepsilon' := \frac{\varepsilon|b|^2}{2} > 0$ gilt

$$\exists n_1 \in \mathbb{N} \quad \forall n \geq n_1 : \quad |b_n - b| \leq \frac{\varepsilon|b|^2}{2}. \quad (100)$$

Definiere nun

$$n_0 := \max\{N, n_1\}.$$

Für $n \geq n_0$ gilt dann

$$\left| \frac{1}{b_n} - \frac{1}{b} \right| = \left| \frac{b}{b_n b} - \frac{b_n}{b_n b} \right| = \left| \frac{b - b_n}{b_n b} \right| \underset{\substack{\text{Satz 2.0.3(ii)} \\ \text{Prop 2.0.5(ii)}}}{\leq} \frac{|b - b_n|}{|b_n| |b|} \underset{(99)}{\leq} \frac{|b - b_n|}{\frac{|b|}{2} |b|} = \frac{2}{|b|^2} |b_n - b| \underset{(100)}{\leq} \frac{2}{|b|^2} \frac{\varepsilon|b|^2}{2} = \varepsilon.$$

Wir haben gesehen

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \left| \frac{1}{b_n} - \frac{1}{b} \right| \leq \varepsilon.$$

Da $\varepsilon > 0$ beliebig war folgt $\frac{1}{b_n} \xrightarrow{n \rightarrow \infty} \frac{1}{b}$. Somit gilt

$$\lim_{n \rightarrow \infty} \left(\frac{1}{b_n} \right) = \frac{1}{b} = \frac{1}{\lim_{n \rightarrow \infty} b_n}$$

ALLGEMEINER FALL. Nun sei $(a_n)_{n \in \mathbb{N}}$ eine beliebige konvergente Folge.

Sei

$$a := \lim_{n \rightarrow \infty} a_n \quad b := \lim_{n \rightarrow \infty} b_n.$$

Dann gilt $\frac{a_n}{b_n} = a_n \cdot \frac{1}{b_n}$ für alle $n \geq N$. Da wir aus dem oben gezeigten Spezialfall schon wissen, dass $\frac{1}{b_n} \xrightarrow{n \rightarrow \infty} \frac{1}{b}$, kann $\left(\frac{a_n}{b_n} \right)_{n \geq N}$ nun als Produkt konvergenter Folgen angesehen werden. Wir folgern mit Grenzwertregel 2

$$\frac{a_n}{b_n} = \underbrace{a_n}_{\rightarrow a} \cdot \underbrace{\frac{1}{b_n}}_{\rightarrow \frac{1}{b}} \xrightarrow[\text{GWR2}]{n \rightarrow \infty} a \cdot \frac{1}{b} = \frac{a}{b}.$$

Es folgt

$$\lim_{n \rightarrow \infty} \left(\frac{a_n}{b_n} \right) = \frac{a}{b} = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n}.$$

Beispiel 2.3.20

Wir untersuchen die Folge

$$\left(\frac{3n^2 + 13n}{n^2 - 2} \right)_{n \in \mathbb{N}}$$

auf Konvergenz.

Dazu: Berechne zuerst für $n \in \mathbb{N}$

$$\frac{3n^2 + 13n}{n^2 - 2} = \frac{(3n^2 + 13n) \cdot \frac{1}{n^2}}{(n^2 - 2) \cdot \frac{1}{n^2}} = \frac{3 + \frac{13}{n}}{1 - \frac{2}{n^2}}. \quad (101)$$

Nun betrachten wir Zähler und Nenner einzeln: Da $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$ haben wir nach Korollar 2.3.18

$$\lim_{n \rightarrow \infty} \frac{13}{n} = 13 \cdot \lim_{n \rightarrow \infty} \frac{1}{n} = 13 \cdot 0 = 0.$$

Mit Grenzwertregel 1 (und der Tatsache, dass $\lim_{n \rightarrow \infty} 3 = 3$) folgt

$$\lim_{n \rightarrow \infty} \left(3 + \frac{13}{n}\right) = \lim_{n \rightarrow \infty} 3 + \lim_{n \rightarrow \infty} \frac{13}{n} = 3 + 0 = 3.$$

Analog zeigt man für den Nenner (unter Zuhilfenahme der Tatsache, dass $\lim_{n \rightarrow \infty} \frac{1}{n^2} = 0$, vgl. Proposition 2.3.7(i))

$$\lim_{n \rightarrow \infty} \left(1 - \frac{2}{n^2}\right) = 1.$$

Mit Grenzwertregel 3 erhält man schließlich

$$\lim_{n \rightarrow \infty} \frac{3 + \frac{13}{n}}{1 - \frac{2}{n^2}} = \frac{\lim_{n \rightarrow \infty} \left(3 + \frac{13}{n}\right)}{\lim_{n \rightarrow \infty} \left(1 - \frac{2}{n^2}\right)} = \frac{3}{1} = 3.$$

Es folgt mit (101), dass

$$\lim_{n \rightarrow \infty} \frac{3n^2 + 13n}{n^2 - 2} = \lim_{n \rightarrow \infty} \frac{3 + \frac{13}{n}}{1 - \frac{2}{n^2}} = 3.$$

QED

Ab sofort reicht es übrigens, wie folgt zu argumentieren:

$$\frac{3n^2 + 13n}{n^2 - 2} = \frac{3 + \overbrace{\frac{13}{n}}^{\rightarrow 0}}{1 - \underbrace{\frac{2}{n^2}}_{\rightarrow 0}} \xrightarrow[n \rightarrow \infty]{\text{GWR}} 3.$$

Das „GWR“ steht hierbei als Sammelbegriff für sämtliche Grenzwertregeln, die wir in diesem Abschnitt kennengelernt haben.

10.11.2025

Vergleichsregeln Eine weitere wichtige Rechenregel für Grenzwerte besagt, dass gewisse *Ungleichungen* unter Grenzwertbildung erhalten bleiben. Hierzu zunächst ein vorbereitendes Lemma.

Lemma 2.3.21 — POSITIVITÄT UNTER GRENZWERTBILDUNG

Es sei $(c_n)_{n \in \mathbb{N}}$ eine konvergente Folge mit $c_n \geq 0$ für alle $n \in \mathbb{N}$. Dann gilt auch

$$\lim_{n \rightarrow \infty} c_n \geq 0.$$

Beweis. Es gelte

$$c_n \geq 0 \quad \forall n \in \mathbb{N}. \tag{102}$$

Definiere $c := \lim_{n \rightarrow \infty} c_n$.

$$\mathbb{Z}: \quad c \geq 0.$$

Angenommen $c < 0$. Definiere $\varepsilon' := -\frac{c}{2}$. Dann gilt klarerweise $\varepsilon' > 0$. Da $c_n \xrightarrow{n \rightarrow \infty} c$ gilt (wegen Definition 2.3.1)

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0: \quad |c_n - c| \leq \varepsilon' = -\frac{c}{2}. \quad (103)$$

Dann gilt für dieses n_0

$$c_{n_0} = c_{n_0} - c + c \leq |c_{n_0} - c| + c \underset{(103)}{\leq} -\frac{c}{2} + c = \frac{c}{2} < 0.$$

Wir folgern $c_{n_0} < 0$, was (102) widerspricht. Somit muss die getroffene Annahme “ $c < 0$ ” falsch sein und es folgt $c \geq 0$. Mit anderen Worten:

$$\lim_{n \rightarrow \infty} c_n \geq 0.$$

Grenzwertregel 4 — UNGLEICHUNGEN UND GRENZWERTE

Es seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ zwei konvergente Folgen mit $a_n \leq b_n$ für alle $n \in \mathbb{N}$. Dann gilt

$$\lim_{n \rightarrow \infty} a_n \leq \lim_{n \rightarrow \infty} b_n.$$

Beweis. Definiere für $n \in \mathbb{N}$ die reelle Zahl $c_n := b_n - a_n$. Da $a_n \leq b_n$ für alle $n \in \mathbb{N}$ gilt

$$c_n \geq 0 \text{ für alle } n \in \mathbb{N}.$$

Betrachte nun $(c_n)_{n \in \mathbb{N}}$ als Folge (insbesondere gilt $(c_n)_{n \in \mathbb{N}} = (b_n - a_n)_{n \in \mathbb{N}}$). Diese ist nach Korollar 2.3.19 (mit $\lambda = 1$ und $\mu = -1$) auch konvergent und aus Lemma 2.3.21 folgt, dass

$$\lim_{n \rightarrow \infty} c_n \geq 0. \quad (104)$$

Nach Korollar 2.3.19 (mit $\lambda = 1$ und $\mu = -1$) gilt auch

$$\begin{aligned} \lim_{n \rightarrow \infty} c_n &= \lim_{n \rightarrow \infty} (b_n - a_n) = \lim_{n \rightarrow \infty} (1 \cdot b_n + (-1) \cdot a_n) \\ &\underset{\text{Kor2.3.19}}{=} 1 \cdot \lim_{n \rightarrow \infty} b_n + (-1) \cdot \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n - \lim_{n \rightarrow \infty} a_n. \end{aligned}$$

Zusammen mit (104) folgern wir

$$\lim_{n \rightarrow \infty} b_n - \lim_{n \rightarrow \infty} a_n \geq 0.$$

Addieren wir $\lim_{n \rightarrow \infty} a_n$ auf beiden Seiten so folgt

$$\lim_{n \rightarrow \infty} b_n \geq \lim_{n \rightarrow \infty} a_n.$$

Bemerkung 2.3.22 — STRIKTE UNGLEICHUNGEN BLEIBEN IM GRENZWERT NICHT ERHALTEN

Es seien $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ zwei konvergente Folgen. Wir haben in der vorigen Proposition gesehen, dass

$$a_n \leq b_n \quad \forall n \in \mathbb{N} \quad \implies \quad \lim_{n \rightarrow \infty} a_n \leq \lim_{n \rightarrow \infty} b_n.$$

Jedoch ist Vorsicht geboten, denn

$$a_n < b_n \quad \forall n \in \mathbb{N} \quad \not\implies \quad \lim_{n \rightarrow \infty} a_n < \lim_{n \rightarrow \infty} b_n!$$

Beispiel 2.3.23

Definiere $a_n := 0$ und $b_n := \frac{1}{n}$ für alle $n \in \mathbb{N}$. Dann gilt

$$a_n < b_n \quad \text{für alle } n \in \mathbb{N},$$

aber

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = 0.$$

Somit wäre die Aussage „ $\lim_{n \rightarrow \infty} a_n < \lim_{n \rightarrow \infty} b_n$ “ nicht erfüllt.

Bemerkung 2.3.24

Da die ersten endlich vielen Folgenglieder für die Konvergenz keine Rolle spielen (vgl. Proposition 2.3.12), kann in Grenzwertregel 4 die Voraussetzung „ $a_n \leq b_n$ für alle $n \in \mathbb{N}$ “ durch „ $(a_n)_{n \in \mathbb{N}} \leq (b_n)_{n \in \mathbb{N}}$ “ ersetzt werden.

Korollar 2.3.25

Es sei $(c_n)_{n \in \mathbb{N}}$ eine konvergente Folge und $x \in \mathbb{R}$.

- (i) Gilt $c_n \leq x$ für alle $n \in \mathbb{N}$, so gilt auch $\lim_{n \rightarrow \infty} c_n \leq x$.
- (ii) Gilt $x \leq c_n$ für alle $n \in \mathbb{N}$, so gilt auch $x \leq \lim_{n \rightarrow \infty} c_n$.

Beweis. Zu (i). Verwende Proposition 4 mit $(a_n)_{n \in \mathbb{N}} = (c_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}} = (x)_{n \in \mathbb{N}}$.
Zu (ii). Verwende Proposition 4 mit $(a_n)_{n \in \mathbb{N}} = (x)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}} = (c_n)_{n \in \mathbb{N}}$.

Sandwichlemma Alle Resultate, die wir vorhin kennengelernt haben, sind nur gültig, wenn man bereits weiß, dass alle vorkommenden Folgen konvergieren. Das *Sandwichlemma* hingegen ist ein Tool, mit dem man Konvergenz einer Folge zeigen kann. Hierfür ist vorausgesetzt, dass diese Folge *zwischen zwei konvergenten Folgen mit demselben Grenzwert eingesperrt ist* (vgl. Abbildung 12)

Lemma 2.3.26 — SANDWICHLEMMA

Es seien $(a_n)_{n \in \mathbb{N}}$, $(x_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ reelle Folgen mit der Eigenschaft, dass

$$a_n \leq x_n \leq b_n \quad \forall n \in \mathbb{N}. \quad (105)$$

Sind $(a_n)_{n \in \mathbb{N}}$ und $(b_n)_{n \in \mathbb{N}}$ beide konvergent und gilt zusätzlich

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n \quad (*)$$

so konvergiert auch $(x_n)_{n \in \mathbb{N}}$ gegen den Wert aus (*).

Beweis. Definiere

$$a := \lim_{n \rightarrow \infty} a_n \quad \left[\text{was nach Voraussetzung auch } = \lim_{n \rightarrow \infty} b_n \text{ ist} \right]$$

Wir zeigen nun, dass $x_n \xrightarrow{n \rightarrow \infty} a$. Sei dazu $\varepsilon > 0$ beliebig vorgegeben.

$$\mathbb{Z}: \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |x_n - a| \leq \varepsilon.$$

Da $a_n \xrightarrow{n \rightarrow \infty} a$ und $b_n \xrightarrow{n \rightarrow \infty} a$ gilt (vgl. Definition 2.3.1)

$$\exists n_1 \in \mathbb{N} \quad \forall n \geq n_1 : \quad |a_n - a| \leq \varepsilon \quad (106)$$

und

$$\exists n_2 \in \mathbb{N} \quad \forall n \geq n_2 : \quad |b_n - a| \leq \varepsilon. \quad (107)$$

Definiere nun

$$n_0 := \max\{n_1, n_2\}.$$

Für $n \geq n_0$ gilt dann

$$\begin{aligned} |x_n - a| &\stackrel{(39)}{=} \max\{x_n - a, -(x_n - a)\} = \max\{\underbrace{x_n}_{\leq b_n} - a, a - x_n\} \\ &\stackrel{(105)}{\leq} \max\{b_n - a, \underbrace{a - x_n}_{\leq -a_n}\} \stackrel{(105)}{\leq} \max\{b_n - a, a - a_n\} \\ &\stackrel{(106)}{\leq} \max\{|b_n - a|, |a - a_n|\} \stackrel{(107)}{\leq} \max\{\varepsilon, \varepsilon\} = \varepsilon. \end{aligned}$$

Wir haben also gezeigt

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |x_n - a| \leq \varepsilon.$$

Da $\varepsilon > 0$ beliebig war folgt $x_n \xrightarrow{n \rightarrow \infty} a$, also ist die Behauptung gezeigt.

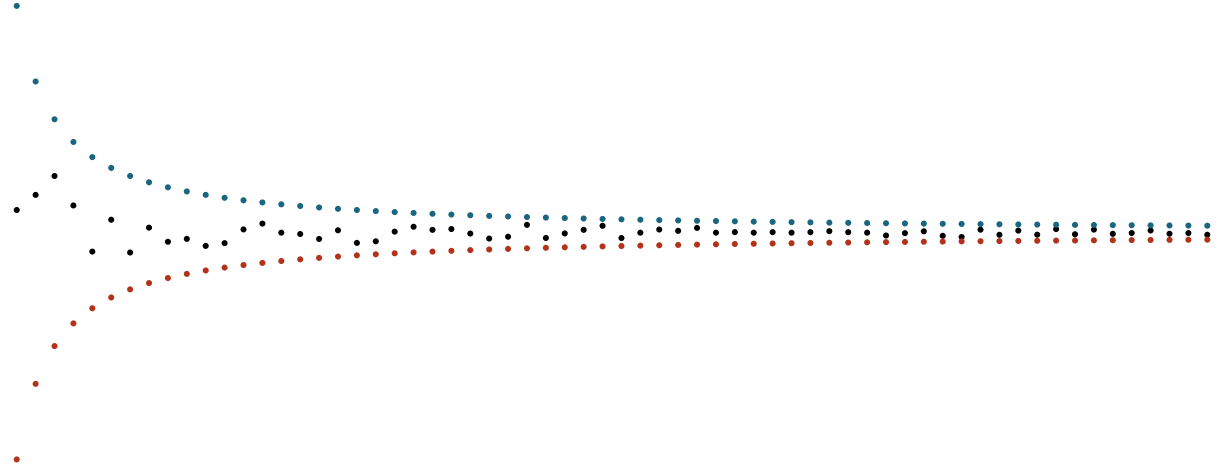


Abbildung 12: Grafische Illustration des Sandwichlemmas.
Die Folge $(a_n)_{n \in \mathbb{N}}$ in rot, $(b_n)_{n \in \mathbb{N}}$ in blau und $(x_n)_{n \in \mathbb{N}}$ in schwarz.

Beispiel 2.3.27

Behauptung.

$$\frac{\sqrt{n-1}}{n} \xrightarrow{n \rightarrow \infty} 0.$$

Beweis. Betrachte für $n \in \mathbb{N}$ die folgende Ungleichungskette

$$0 \leq \frac{\sqrt{n-1}}{n} \leq \frac{\sqrt{n-0}}{n} = \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}}. \quad (108)$$

Definiert man für $n \in \mathbb{N}$ nun $a_n := 0$ und $b_n := \frac{1}{\sqrt{n}} = \frac{1}{n^{\frac{1}{2}}}$, so gilt (vgl. Proposition 2.3.7(i))

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = 0 \quad (109)$$

Dank (108) erfüllt die Folge $(x_n)_{n \in \mathbb{N}}$ gegeben durch $x_n := \frac{\sqrt{n-1}}{n}$ die Voraussetzungen des Sandwichlemmas (Proposition 2.3.26). Man lernt aus (109), dass

$$\lim_{n \rightarrow \infty} x_n = 0.$$

Beispiel 2.3.28*Behauptung.*

$$\frac{\sqrt{n+1}}{n} \xrightarrow{n \rightarrow \infty} 0$$

Dazu beobachte

$$0 \leq \frac{\sqrt{n+1}}{n} \leq \frac{\sqrt{n+n}}{n} = \frac{\sqrt{2n}}{n} = \frac{\sqrt{2}\sqrt{n}}{n} = \sqrt{2} \frac{1}{\sqrt{n}}. \quad (110)$$

Definiere für alle $n \in \mathbb{N}$ nun $a_n := 0$ und $b_n := \sqrt{2} \frac{1}{\sqrt{n}}$. Dann gilt

$$\lim_{n \rightarrow \infty} a_n = 0$$

und

$$\lim_{n \rightarrow \infty} b_n \stackrel{\text{Kor 2.3.18}}{=} \sqrt{2} \lim_{n \rightarrow \infty} \frac{1}{\sqrt{n}} \stackrel{\text{Prop 2.3.7}}{=} \sqrt{2} \cdot 0 = 0.$$

Wiederum erfüllt nach (110) die Folge $(\frac{\sqrt{n+1}}{n})_{n \in \mathbb{N}}$ die Voraussetzungen des Sandwichlemmas (Proposition 2.3.26) und wir lernen

$$\lim_{n \rightarrow \infty} \frac{\sqrt{n+1}}{n} = 0.$$

Nullfolgen Im nachfolgenden Abschnitt interessieren uns Folgen, die *gegen 0* konvergieren. Nach Definition 2.3.1 konvergiert eine Folge $(a_n)_{n \in \mathbb{N}}$ genau dann gegen 0, wenn

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : |a_n| \leq \varepsilon$$

Definition 2.3.29 — NULLFOLGEEine Folge $(a_n)_{n \in \mathbb{N}}$ heißt *Nullfolge*, wenn $a_n \xrightarrow{n \rightarrow \infty} 0$.

Die 0 nimmt (wie wir später auch noch sehen werden) auf unserer Zahlengerade eine besondere Rolle ein. Deswegen gilt die nachfolgende Aussage, auch *nur* für den Grenzwert 0.

Proposition 2.3.30Es sei $(a_n)_{n \in \mathbb{N}}$ eine Folge. Dann gilt

$$(a_n \xrightarrow{n \rightarrow \infty} 0) \iff (|a_n| \xrightarrow{n \rightarrow \infty} 0).$$

In Worten: $(a_n)_{n \in \mathbb{N}}$ ist eine Nullfolge genau dann wenn $(|a_n|)_{n \in \mathbb{N}}$ eine Nullfolge ist.

Beweis. Beachte: „ $a_n \xrightarrow{n \rightarrow \infty} 0$ “ ist äquivalent zu (vgl. Definition 2.3.1)

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |a_n| \leq \varepsilon. \quad (111)$$

und „ $|a_n| \xrightarrow{n \rightarrow \infty} 0$ “ ist äquivalent zu

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad ||a_n|| \leq \varepsilon. \quad (112)$$

Da $||a_n|| = |a_n|$ sind auch (111) und (112) äquivalent. Die Behauptung folgt.

Bemerkung 2.3.31

Für andere Grenzwerte als 0 ist die Aussage falsch. Beispielweise ist

$$(a_n \xrightarrow{n \rightarrow \infty} 1) \not\iff (|a_n| \xrightarrow{n \rightarrow \infty} 1). \quad (113)$$

Dazu betrachte $(a_n)_{n \in \mathbb{N}} = ((-1)^n)_{n \in \mathbb{N}}$. Es gilt zwar

$$\lim_{n \rightarrow \infty} |a_n| = \lim_{n \rightarrow \infty} |(-1)^n| = \lim_{n \rightarrow \infty} 1 = 1.$$

Jedoch gilt auch: $\lim_{n \rightarrow \infty} a_n$ existiert gar nicht (vgl. Beispiel 2.3.10)!

In (113) kann also „ $\iff$ “ nicht gelten. Interessanterweise gilt „ $\implies$ “ aber schon, wie man der nachfolgenden Proposition entnehmen kann.

Proposition 2.3.32

Es sei $(a_n)_{n \in \mathbb{N}}$ eine Folge und $a \in \mathbb{R}$. Dann gilt

$$(a_n \xrightarrow{n \rightarrow \infty} a) \implies (|a_n| \xrightarrow{n \rightarrow \infty} |a|).$$

Beweis. $\lim_{n \rightarrow \infty} a_n = a$ bedeutet nach Definition 2.3.1

$$\forall \varepsilon > 0 \quad \exists n_1 \in \mathbb{N} \quad \forall n \geq n_1 : \quad |a_n - a| \leq \varepsilon. \quad (114)$$

$\mathbb{Z}\mathbb{L}$: $|a_n| \xrightarrow{n \rightarrow \infty} |a|$ ($n \rightarrow \infty$), d.h.

$$\mathbb{Z}\mathbb{L} : \quad \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad ||a_n| - |a|| \leq \varepsilon$$

Sei dazu $\varepsilon > 0$ beliebig vorgegeben. Wir wissen aus (114): Es gibt ein $n_1 \in \mathbb{N}$, so dass

$$\forall n \geq n_1 : \quad |a_n - a| \leq \varepsilon. \quad (115)$$

Definiere nun $n_0 := n_1$. Für $n \geq n_0$ (d.h. $n \geq n_1$) gilt dann

$$||a_n| - |a|| \underset{\text{Prop 2.0.5(iii)}}{\leq} |a_n - a| \underset{(115)}{\leq} \varepsilon.$$

Wir haben gezeigt

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad ||a_n| - |a|| \leq \varepsilon.$$

Eine weitere nützliche Erkenntnis über Nullfolgen ist, dass die Multiplikation mit einer *beschränkten Folge* die Nullfollgeneigenschaft erhält.

Proposition 2.3.33 — PRODUKTE VON BESCHRÄNKTEN FOLGEN UND NULLFOLGEN

Es sei $(c_n)_{n \in \mathbb{N}}$ eine beschränkte^a Folge und $(a_n)_{n \in \mathbb{N}}$ eine Nullfolge. Dann ist auch $(c_n a_n)_{n \in \mathbb{N}}$ eine Nullfolge.

^aBeachte: $(c_n)_{n \in \mathbb{N}}$ wird nur als beschränkt und nicht als konvergent vorausgesetzt. $(c_n)_{n \in \mathbb{N}}$ darf also im allgemeinen durchaus divergent sein!

Beweis. ALTERNATIVE I. MIT DEM SANDWICHLEMMA. Es sei $(a_n)_{n \in \mathbb{N}}$ eine Nullfolge. Dann ist nach Proposition 2.3.30 auch $(|a_n|)_{n \in \mathbb{N}}$ eine Nullfolge. Da $(c_n)_{n \in \mathbb{N}}$ beschränkt ist gibt es (vgl. Proposition 2.1.6) ein $K > 0$ mit

$$\forall n \in \mathbb{N} : \quad |c_n| \leq K. \quad (116)$$

Nun gilt für alle $n \in \mathbb{N}$

$$0 \leq |c_n a_n| = |c_n| \cdot |a_n| \leq K |a_n|.$$

Die Folgen $(0)_{n \in \mathbb{N}}$ und $(K|a_n|)_{n \in \mathbb{N}}$ schachteln also die Folge $(|c_n a_n|)_{n \in \mathbb{N}}$ ein. Da

$$\lim_{n \rightarrow \infty} 0 = \lim_{n \rightarrow \infty} K |a_n| = 0$$

gilt nach dem Sandwichlemma (Proposition 2.3.26) auch

$$\lim_{n \rightarrow \infty} |c_n a_n| = 0.$$

Mit Proposition 2.3.30 folgt auch

$$\lim_{n \rightarrow \infty} c_n a_n = 0.$$

*ALTERNATIVE II. DIREKT MIT DER DEFINITION. Da $(a_n)_{n \in \mathbb{N}}$ eine Nullfolge ist gilt (vgl. Definition 2.3.1)

$$\forall \varepsilon > 0 \quad \exists n_1 \in \mathbb{N} \quad \forall n \geq n_1 : \quad |a_n| \leq \varepsilon. \quad (117)$$

Da $(c_n)_{n \in \mathbb{N}}$ beschränkt ist, gibt es (vgl. Proposition 2.1.6) ein $K > 0$ mit

$$\forall n \in \mathbb{N} : \quad |c_n| \leq K. \quad (118)$$

Es sei nun $\varepsilon > 0$ beliebig vorgegeben.

$$\text{Z:} \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |c_n a_n| \leq \varepsilon.$$

Definiere zuerst $\varepsilon' := \frac{\varepsilon}{K} > 0$. Nach (117) gibt es ein $n_1 \in \mathbb{N}$ so, dass

$$\forall n \geq n_1 : \quad |a_n| \leq \varepsilon' = \frac{\varepsilon}{K}. \quad (119)$$

Definiere nun $n_0 := n_1$. Für $n \geq n_0$ (d.h. $n \geq n_1$) gilt dann

$$|c_n a_n| = |c_n| \cdot |a_n| \underset{(118)}{\leq} K |a_n| \underset{(119)}{\leq} K \frac{\varepsilon}{K} = \varepsilon.$$

Wir haben gezeigt

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |c_n a_n| \leq \varepsilon.$$

Da $\varepsilon > 0$ beliebig war folgt die Behauptung.

Beispiel 2.3.34

Es gilt

$$\lim_{n \rightarrow \infty} \frac{(-1)^n}{n} = 0.$$

In der Tat: Wählt man für alle $n \in \mathbb{N}$

- $c_n := (-1)^n$ und
- $a_n := \frac{1}{n}$

so ist $(c_n)_{n \in \mathbb{N}}$ beschränkt (vgl. Beispiel 2.1.4) und $(a_n)_{n \in \mathbb{N}}$ eine Nullfolge (vgl. Proposition 2.3.7 (i)). Man erhält, dass $(c_n a_n)_{n \in \mathbb{N}} = \left(\frac{(-1)^n}{n}\right)_{n \in \mathbb{N}}$ eine Nullfolge ist, was zu zeigen war.

Bestimmte Divergenz Im Folgenden wollen wir uns über das Konzept der *bestimmten Divergenz* gegen $+\infty$ oder $-\infty$ unterhalten. Manche Lehrbücher sprechen dabei übrigens auch von einer „(uneigentlichen)“ *Konvergenz gegen ∞ oder $-\infty$* . Es handelt sich allerdings nicht um Konvergenz im Sinne von Definition 2.3.1 und die Grenzwertregeln gelten so wie wir sie kennen auch nicht. Daher werden wir nur den Begriff der *bestimmten Divergenz* verwenden.

Definition 2.3.35 — BESTIMMTE DIVERGENZ

Eine reelle Folge $(a_n)_{n \in \mathbb{N}}$

- (i) *divergiert bestimmt gegen ∞* falls

$$\forall L > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad a_n \geq L$$

- (ii) *divergiert bestimmt gegen $-\infty$* falls

$$\forall K > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad a_n \leq -K$$

Man schreibt in Fall (i) auch

$$”\lim_{n \rightarrow \infty}” a_n = \infty \quad \text{oder} \quad a_n ”\xrightarrow{n \rightarrow \infty}” \infty.$$

In Fall (ii) verwendet man dieselbe Notation mit $-\infty$. Die Anführungszeichen darf man weglassen, wenn aus dem Kontext klar ist, was gemeint ist.

Mit anderen Worten: $(a_n)_{n \in \mathbb{N}}$ ist bestimmt divergent gegen ∞ falls für alle (noch so großen) $L > 0$ gilt

$$(L)_{n \in \mathbb{N}} \leq (a_n)_{n \in \mathbb{N}}.$$

Mit noch anderen Worten (vgl. Proposition 2.2.8):

$$\{n \in \mathbb{N} : a_n < L\} \text{ ist endlich für alle } L > 0.$$

Beispiel 2.3.36 — EINE BESTIMMT DIVERGENTE FOLGE

Die Folge $(n)_{n \in \mathbb{N}}$ ist bestimmt divergent gegen ∞ . In der Tat: für beliebiges aber festes $L > 0$ gilt

$$\{n \in \mathbb{N} : n < L\} \subseteq \{1, 2, \dots, [L]\},$$

was eine endliche Menge ist. Nach Proposition 2.0.10 ist somit auch $\{n \in \mathbb{N} : n < L\}$ endlich.

Bemerkung 2.3.37 — KEINE GRENZWERTREGELN BEI BESTIMMTER DIVERGENZ

Warum verwenden wir die umständliche Notation $”\lim_{n \rightarrow \infty}” a_n = \infty$ mit den Anführungszeichen? Ganz einfach: Weil wir den Begriff der bestimmten Divergenz von der Theorie konvergenter Folgen stark abgrenzen müssen! Für die bestimmte Divergenz gelten zum Beispiel **keine Grenzwertregeln**. Wer die Anführungszeichen nicht mitschreibt und die Grenzwertregeln unhinterfragt für die bestimmte Divergenz verwendet landet oft bei folgendem Trugschluss

$$1 = \lim_{n \rightarrow \infty} 1 = \lim_{n \rightarrow \infty} n \cdot \frac{1}{n} = \left(\lim_{n \rightarrow \infty} n\right) \cdot \left(\lim_{n \rightarrow \infty} \frac{1}{n}\right) = \infty \cdot 0 = 0.$$

Die Aussage $1 = 0$ ist natürlich falsch. Der Grund für diesen Fehler ist, dass die Grenzwertregel 2 sich nicht auf bestimmte Divergenz verallgemeinern lässt. Genau aus diesem Grund verwenden wir die Anführungszeichen, damit man nicht in Versuchung gerät, solche Regeln zu verwenden.

$\infty \cdot 0$ ist übrigens auch nicht gleich 0. Mit dem Symbol “ ∞ ” sollte man nicht anfangen zu rechnen. Wir werden im Verlauf der gesamten Vorlesung kein einzige Rechenregel für das Symbol ∞ studieren, weil keine Konvention hier zufriedenstellend ist! Sollte ihr in den Übungsaufgaben also in Versuchung geraten, Ausdrücke wie $0 \cdot \infty$ oder $\infty - \infty$ zu berechnen, seid ihr auf dem Holzweg – diese Ausdrücke sind nicht definiert.

2.4 Abzählungen

In diesem Abschnitt setzen wir uns mit der *Abzählbarkeit* von Mengen auseinander. Mengen heißen intuitiv gesprochen abzählbar, wenn ich alle ihre Elemente *mit einer Folge durchnummerieren kann*, ohne dabei eines zu vergessen. Wir werden sehen, dass das nicht für alle Mengen machbar ist.

Definition 2.4.1 — ABZÄHLBARKEIT

Eine Menge M heißt (höchstens) abzählbar, falls entweder $M = \emptyset$ ist oder es eine Folge $(a_n)_{n \in \mathbb{N}} \subseteq M$ gibt so, dass

$$\{a_n : n \in \mathbb{N}\} = M. \quad (120)$$

Eine solche Folge nennt man dann eine *Abzählung* von M . Eine Menge, die nicht abzählbar ist, nennt man *überabzählbar*.

Es genügt übrigens, in (120) „ $\supseteq$ “ nachzuweisen. Die Inklusion „ $\subseteq$ “ folgt ja schon allein dadurch, dass $a_n \in M$ für alle $n \in \mathbb{N}$.

Bemerkung 2.4.2

Per Definition ist eine Folge nichts anderes als eine Abbildung $a : \mathbb{N} \rightarrow M$ (vgl. Definition 2.1.1). Statt $a(n)$ haben wir nur immer a_n geschrieben. Ist $(a_n)_{n \in \mathbb{N}}$ nun eine Folge wie in (120), so gilt

$$M = \{a_n : n \in \mathbb{N}\} = \{a(n) : n \in \mathbb{N}\} = a(\mathbb{N}).$$

Mit anderen Worten: $a : \mathbb{N} \rightarrow M$ ist surjektiv (vgl. Proposition 1.3.10). Wir halten fest

M ist genau dann abzählbar wenn es eine surjektive Abbildung $a : \mathbb{N} \rightarrow M$ gibt.

Ein erstes Beispiel für abzählbare Mengen sind *endliche* Mengen.

13.11.2025

Proposition 2.4.3 — ENDLICHE MENGEN SIND ABZÄHLBAR

Es sei M eine endliche Menge. Dann ist M abzählbar.

Beweis. Es sei M eine endliche Menge. Falls $M = \emptyset$ so ist die Aussage klar, daher dürfen wir für den Rest des Beweises annehmen, dass $M \neq \emptyset$. Nach Definition 2.0.6 gibt es ein $N \in \mathbb{N}$ so, dass man eine bijektive Abbildung $f : \{1, \dots, N\} \rightarrow M$ finden kann. Mit anderen Worten gilt

$$M = \{f(1), \dots, f(N)\}.$$

Definiere nun die Folge $(a_n)_{n \in \mathbb{N}}$ durch

$$a_n := \begin{cases} f(n) & \text{falls } n \leq N \\ f(1) & \text{falls } n > N \end{cases}.$$

Dann gilt $(a_n)_{n \in \mathbb{N}} \subseteq M$ und außerdem

$$\{a_n : n \in \mathbb{N}\} \supseteq \{a_1, \dots, a_N\} = \{f(1), \dots, f(N)\} = M.$$

Somit ist M abzählbar.

Es gibt aber auch unendliche abzählbare Mengen, das Paradebeispiel dafür ist $\mathbb{N}$.

Beispiel 2.4.4

Die natürlichen Zahlen $\mathbb{N}$ sind abzählbar. In der Tat: Definiert man $a_n := n$ für alle $n \in \mathbb{N}$ so ist

$$\{a_n : n \in \mathbb{N}\} = \{n : n \in \mathbb{N}\} = \mathbb{N}.$$

Bei Mengen, die intuitiv *größer* sind als die natürlichen Zahlen sollte es mit der Durchnummerierung schwierig werden. Manchmal klappt es allerdings erstaunlicherweise doch

Beispiel 2.4.5

Die ganzen Zahlen $\mathbb{Z}$ sind abzählbar. Dazu betrachte die Folge

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

Genauer: Ist $(a_n)_{n \in \mathbb{N}}$ gegeben durch

$$a_n := (-1)^{n-1} \left(\left\lceil \frac{n}{2} \right\rceil - 1 \right) \quad \forall n \in \mathbb{N},$$

so gilt

$$\{a_n : n \in \mathbb{N}\} = \mathbb{Z}.$$

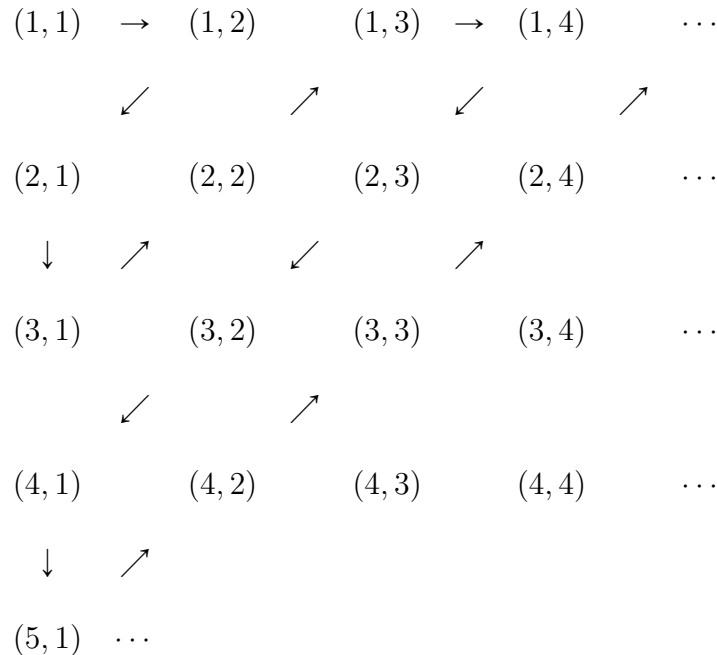
Proposition 2.4.6 — ABZÄHLBARKEIT VON $\mathbb{N} \times \mathbb{N}$

Die Menge

$$\mathbb{N} \times \mathbb{N} = \{(m, n) : m \in \mathbb{N} \wedge n \in \mathbb{N}\}$$

ist abzählbar.

Beweisskizze. Ordne die Elemente von $\mathbb{N} \times \mathbb{N}$ in folgendem *Rechtecksschema* an



Die durch die Pfeile angedeutete Abzählungsvorschrift liefert eine Folge

$(1, 1), (1, 2), (2, 1), (3, 1), (2, 2), (1, 3), (1, 4), (2, 3), (3, 2), (4, 1), (5, 1), (4, 2), (3, 3), (2, 4), \dots$

Diese nummeriert alle Elemente von $\mathbb{N} \times \mathbb{N}$ durch.

Bemerkung 2.4.7

Mit der Überlegung der vorigen Proposition folgt, dass auch die rationalen Zahlen $\mathbb{Q}$ abzählbar sind. In der Tat: Eine positive rationale Zahl lässt sich stets schreiben als $\frac{m}{n}$ mit $m, n \in \mathbb{N}$. Identifizieren wir den Bruch $\frac{m}{n}$ mit dem Paar $(m, n) \in \mathbb{N} \times \mathbb{N}$ und wählen eine Abzählung wie in Proposition 2.4.6, so haben wir eine Folge, die alle positiven rationalen Zahlen durchnummeriert.

Kann man nun auch eine Folge finden, die die alle rationalen Zahlen (also auch die negativen und die Zahl 0) durchnummeriert? Ich behaupte ja und lade Sie in diesem Zusammenhang dazu ein, nochmal genau über Beispiel 2.4.5 nachzudenken!

Nun haben wir bisher nur Beispiele für abzählbare Mengen gesehen. Jetzt sehen wir mal ein Beispiel für eine überabzählbare Menge: Die *Potenzmenge der natürlichen Zahlen* $\mathcal{P}(\mathbb{N})$. Hierfür machen wir zunächst eine vorbereitende viel allgemeinere Beobachtung

Lemma 2.4.8

Es sei M eine beliebige nichtleere Menge. Dann gibt es keine surjektive Abbildung $f : M \rightarrow \mathcal{P}(M)$.

Beweis. Angenommen es gibt doch eine solche surjektive Abbildung $f : M \rightarrow \mathcal{P}(M)$. Beachte: Für alle $x \in M$ ist dann $f(x) \in \mathcal{P}(M)$ eine Teilmenge von M ! Betrachte nun die Menge

$$N := \{x \in M : x \notin f(x)\}. \quad (a) \tag{121}$$

Nun ist klarerweise $N \subseteq M$. Mit anderen Worten $N \in \mathcal{P}(M)$. Da f surjektiv ist, muss es ein $x^* \in M$ geben mit

$$N = f(x^*). \tag{122}$$

Nun erhalten wir folgenden Widerspruch:

- Wäre $x^* \in N$, so wäre nach (121) $x^* \notin f(x^*)$. Aus (122) folgt dann $x^* \notin N$. Ein Widerspruch.
- Wäre $x^* \notin N$, so wäre nach (121) $x^* \in f(x^*)$. Aus (122) folgt dann $x^* \in N$. Ein Widerspruch.

In allen Fällen ergibt sich ein Widerspruch. Somit muss die Annahme (dass es eine surjektive Abbildung $f : M \rightarrow \mathcal{P}(M)$ gibt) falsch sein.

^aNoch ein weiteres Mal gebe ich zu beachten, dass die Aussage $x \notin f(x)$ Sinn ergibt, weil $f(x)$ eine Teilmenge von M ist.

Korollar 2.4.9

Die Menge

$$\mathcal{P}(\mathbb{N}) = \{A : A \subseteq \mathbb{N}\}$$

ist überabzählbar.

Beweis. Wäre $\mathcal{P}(\mathbb{N})$ abzählbar, so gäbe es nach Bemerkung 2.4.2 eine surjektive Abbildung $a : \mathbb{N} \rightarrow \mathcal{P}(\mathbb{N})$. Die Existenz einer solchen Abbildung ist aber wegen Lemma 2.4.8 ausgeschlossen! Daher ist $\mathcal{P}(\mathbb{N})$ überabzählbar.

Später (wahrscheinlich erst in Analysis III) werden wir sehen, dass die reellen Zahlen $\mathbb{R}$ auch überabzählbar sind. Um dies richtig verstehen zu können, müssen wir aber die reellen Zahlen erst genau einführen. Das soll das Ziel des nachfolgenden Kapitels sein.

3 Die reellen Zahlen

Im Folgenden möchten wir die *reellen Zahlen*, von uns notiert mit $\mathbb{R}$, studieren. Wir möchten verstehen, mit welchen *Rechengesetzen* wir in den reellen Zahlen operieren können und wie man diese aus einer möglichst kleinen Zahl von *Axiomen* herleiten kann.

Sie werden sehen, dass wir die Rechengesetze, die wir in diesem Kapitel kennenlernen, bereits in Kapitel 2 ohne Kommentar richtig benutzt haben. Und doch ist es notwendig, sie kritisch zu hinterfragen!

Vergessen Sie ab hier für einen Moment einmal alles, was Sie in Kapitel 2 gelernt haben und auch alles, was Sie aus der Schule über das Rechnen mit Zahlen wissen. Sie dürfen in den Übungsaufgaben zu diesem Kapitel Erkenntnisse aus Kapitel 2 und aus der Schule nicht verwenden, sondern nur das, was bis zum Zeitpunkt der Übungsaufgabe in den Kapiteln 1 und 3 gelernt wurde.

Wir wollen die Zahlenmenge $\mathbb{R}$ im Folgenden rigoros definieren. Ich werde dazu hier eine Definition präsentieren, die jedoch noch sehr viele unbekannte Begriffe enthalten wird. Wir werden diese (in blau gekennzeichneten) Begriffe im Laufe dieser Sektion definieren und studieren.

Definition 3.0.1 — REELLE ZAHLEN

Eine Menge $\mathbb{R}$ heißt *Realisierung der reellen Zahlen* falls

- (i) $\mathbb{R}$ ein **angeordneter Körper** ist (mit gewissen **Verknüpfungen** $+$, $\cdot$ und einer **Ordnungsrelation** $\leq$).
- (ii) $\mathbb{R}$ **ordnungsvollständig** ist.

Wenn wir die oben genannten Begriffe eingeführt haben, dann würde sich wohl zunächst die Frage stellen, ob eine solche Menge $\mathbb{R}$ überhaupt existiert. Dieser können wir in dieser Vorlesung aus Zeitgründen nicht nachgehen.

Tatsächlich wird es mehrere verschiedene Realisierungen der reellen Zahlen geben, d.h. mehrere Mengen, die die Eigenschaften aus Definition 3.0.1 haben. Wir werden aber nach diesem Kapitel eine einzige *fixe Realisierung der reellen Zahlen* betrachten und mit dieser rechnen.

3.1 Körper

Ein Körper ist eine Menge von Objekten mit zwei *Verknüpfungen*, für die bestimmte *Rechengesetze* gelten. Die Rechenregeln, die wir in der Definition fordern, bilden die *minimale Anzahl von Spielregeln*, aus denen wir alle uns geläufigen Rechenregeln herleiten können.

Bevor wir uns über Rechenregeln unterhalten, müssen wir uns aber erstmal verstehen was *rechnen* bedeutet. Für uns meint „rechnen“ stets das Anwenden *innerer Verknüpfungen*.

Definition 3.1.1 — INNERE VERKNÜPFUNG

Es sei M eine Menge. Eine *innere Verknüpfung* $\odot$ auf M ist eine Abbildung $\odot : M \times M \rightarrow M$. Anstatt $\odot(x, y)$ schreibt man auch oft $x \odot y$. Somit hat man auch die Zuordnungsvorschrift

$$\odot : M \times M \rightarrow M, (x, y) \mapsto x \odot y.$$

Beispiel 3.1.2 — INNERE VERKNÜPFUNGEN

Es ist

$$+ : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}, \quad + : (x, y) \mapsto x + y$$

eine innere Verknüpfung. Zum Beispiel gilt $+(1, 2) = 1 + 2 = 3$. Die Notation $1 + 2$ ist Ihnen hierbei wahrscheinlich wesentlich geläufiger als $+(1, 2)$. Daher kommt auch die Notation aus Definition 3.1.1.

Es kann aber auch ganz absurde innere Verknüpfungen geben:

Beispiel 3.1.3 — ETWAS ABSURD

Man kann für $M = \{\text{Sonne, Mond, Sterne}\}$ eine Verknüpfung $\odot$ mithilfe einer *Verknüpfungstabelle* definieren

$\odot$	Sonne	Mond	Sterne
Sonne	Sonne	Mond	Sterne
Mond	Mond	Sterne	Sonne
Sterne	Sterne	Sonne	Mond

Diese Tabelle funktioniert so: verknüpft man ein Element aus der ersten Spalte mit einem Element aus der ersten Zeile, so ist das Ergebnis der jeweilige Eintrag der Tabelle.

Nun sind wir bereit, unsere *Minimalanforderungen* an unser Rechnen zu formulieren.

Definition 3.1.4 — KÖRPER

Ein *Körper* $(K, \oplus, \odot)$ ist eine Menge K mit zwei inneren Verknüpfungen

$$\oplus : K \times K \rightarrow K \quad \text{und} \quad \odot : K \times K \rightarrow K$$

so, dass folgende Gesetze gelten.

(A1) *Assoziativgesetz*. Für alle $x, y, z \in K$ gilt $x \oplus (y \oplus z) = (x \oplus y) \oplus z$.

(A2) *Kommutativgesetz*. Für alle $x, y \in K$ gilt $x \oplus y = y \oplus x$.

(A3) *Nullgesetz*. Es existiert ein Element $0_K \in K$ mit $x \oplus 0_K = x$ für alle $x \in K$.

(A4) *Gesetz des Negativen*. Für alle $x \in K$ gibt es ein Element $-x \in K$ derart, dass $x \oplus -x = 0_K$.

(M1) *Assoziativgesetz*. Für alle $x, y, z \in K$ gilt $x \odot (y \odot z) = (x \odot y) \odot z$.

(M2) *Kommutativgesetz*. Für alle $x, y \in K$ gilt $x \odot y = y \odot x$.

(M3) *Einsgesetz*. Es existiert ein Element $1_K \in K$ mit $x \odot 1_K = x$ für alle $x \in K$.

(M4) *Gesetz des Kehrwerts*. Für alle $x \in K \setminus \{0_K\}$ gibt es ein Element $x^{-1} \in K \setminus \{0_K\}$ so, dass $x \odot x^{-1} = 1_K$.

(D) *Distributivgesetz*. Für alle $x, y, z \in K$ gilt $x \odot (y \oplus z) = (x \odot y) \oplus (x \odot z)$.

Ein Element 0_K wie in (A3) nennen wir *Nullelement* und ein Element 1_K wie in (M3) nennen wir *Einselement*. Für $x \in K$ nennen wir ein Element $-x$ wie in (A4) stets ein *Negatives* oder *additiv Inverses* von x und für $x \in K \setminus \{0_K\}$ nennen wir ein Element wie in (M4) einen *Kehrwert* oder *multiplikativ Inverses* von x .

Die Regeln (A1) bis (A4) nennen wir auch *Rechenregeln für die Addition* und die Regeln (M1) bis (M4) nennen wir *Rechenregeln für die Multiplikation*. Im Distributivgesetz (D) werden beide innere Verknüpfungen benutzt, daher fällt es unter keine Kategorie. Die Klammern auf der rechten Seite im Distributivgesetz (D) können weggelassen werden, da wir uns ab sofort auf die Konvention *Punktrechnung vor Strichrechnung* einigen, d.h. $\odot$ hat Vorfahrt vor $\oplus$.

Folgendes Beispiel können wir eigentlich erst machen, wenn wir die reellen Zahlen bereits verstanden haben. Daher ist es an dieser Stelle noch nicht rigoros. Trotzdem ist es wohl hilfreich, die genannten Beispiele vor Augen zu haben:

Zukunfts-Beispiel 3.1.5 — DIE REELLEN ZAHLEN

Es bildet $(\mathbb{R}, +, \cdot)$ mit den uns geläufigen Rechenoperationen „+“ und „ $\cdot$ “ einen Körper. Ferner gilt $0_{\mathbb{R}} = 0$ und $1_{\mathbb{R}} = 1$.

Zukunfts-Beispiel 3.1.6 — DIE RATIONALEN ZAHLEN

Ein anderes Beispiel dieser Gattung ist $(\mathbb{Q}, +, \cdot)$, wobei $\mathbb{Q}$ die *rationalen Zahlen* bezeichnen (welche wir auch noch nicht eingeführt haben). Bei diesem Beispiel sind „+“ und „ $\cdot$ “ die uns geläufigen Verknüpfungen.

Nicht-Beispiel 3.1.7 — DIE GANZEN ZAHLEN

$(\mathbb{Z}, +, \cdot)$ ist *kein* Körper, denn (M4) ist nicht erfüllt: So besitzt zum Beispiel $2 \in \mathbb{Z} \setminus \{0\}$ kein multiplikativ Inverses innerhalb $\mathbb{Z} \setminus \{0\}$ (da $\frac{1}{2} \notin \mathbb{Z} \setminus \{0\}$)^a.

^aMan müsste hier eigentlich genauer argumentieren, denn schließlich könnte es ja sein, dass trotzdem eine andere Zahl aus $\mathbb{Z} \setminus \{0\}$ das Gewünschte leistet. Die präzisere Argumentation geht so: Es gibt kein $z \in \mathbb{Z} \setminus \{0\}$ so, dass $2 \cdot z = 1$. Das ist leicht verstanden: $2 \cdot z$ muss ja für jedes ganzzahlige z eine gerade

Zahl sein und 1 ist eine ungerade Zahl.

Wir wollen an dieser Stelle einmal ein etwas anderes Beispiel eines Körpers besprechen, bei dem die Rechenoperationen uns nicht so geläufig sind.

Beispiel 3.1.8

Betrachte die Menge $K := \{\text{Sonne, Mond, Sterne}\}$ mit Verknüpfungen $\oplus, \odot : K \times K \rightarrow K$ gegeben durch folgende Verknüpfungstabellen:

$\oplus$	Sonne	Mond	Sterne
Sonne	Sonne	Mond	Sterne
Mond	Mond	Sterne	Sonne
Sterne	Sterne	Sonne	Mond

$\odot$	Sonne	Mond	Sterne
Sonne	Sonne	Sonne	Sonne
Mond	Sonne	Mond	Sterne
Sterne	Sonne	Sterne	Mond

Man kann sich davon überzeugen, dass $(K, \oplus, \odot)$ einen Körper bildet. Es gilt $0_K = \text{Sonne}$ und $1_K = \text{Mond}$.

Sämtliche Rechenregeln, die Sie aus Ihrer Intuition gewohnt sind zu benutzen, können Sie aus den oben genannten herleiten.

Proposition 3.1.9 — RECHENREGELN IN KÖRPERN

Es sei $(K, \oplus, \odot)$ ein Körper.

- (i) Es gibt genau ein Element $x_0 \in K$, welches dem Nullgesetz genügt (d.h. $x \oplus x_0 = x$ für alle $x \in K$). Dieses ist gegeben durch $x_0 = 0_K$.
- (ii) Es gibt genau ein Element $x_1 \in K$, welches dem Einsgesetz genügt (d.h. $x \odot x_1 = x$ für alle $x \in K$). Dieses ist gegeben durch $x_1 = 1_K$.
- (iii) Zu jedem $x \in K$ gibt es genau ein Element $x' \in K$, welches der Gleichung $x \oplus x' = 0_K$ genügt. Dieses ist gegeben durch $x' = -x$. Insbesondere gibt es genau ein einziges additiv Inverses von x .
- (iv) Zu jedem $x \in K \setminus \{0_K\}$ gibt es genau ein Element $\bar{x} \in K$, welches der Gleichung $x \odot \bar{x} = 1_K$ genügt. Dieses ist gegeben durch $\bar{x} = x^{-1}$. Insbesondere gibt es genau ein einziges multiplikativ Inverses von x .
- (v) Es gilt $-0_K = 0_K$.
- (vi) Für alle $x \in K$ ist $-(-x) = x$.
- (vii) Für alle $x \in K \setminus \{0_K\}$ ist $(x^{-1})^{-1} = x$.
- (viii) Für alle $x, y \in K$ gilt $-(x \oplus y) = (-x) \oplus (-y)$.
- (ix) Für alle $x, y, z \in K$ gilt $(x \oplus y) \odot z = x \odot z \oplus y \odot z$.

- (x) Für alle $x \in K$ gilt $x \odot 0_K = 0_K$.
- (xi) Für $x, y \in K$ gilt $x \odot y = 0_K$ genau dann wenn $x = 0_K \vee y = 0_K$.
- (xii) Für alle $x, y \in K \setminus \{0_K\}$ ist $(x \odot y)^{-1} = x^{-1} \odot y^{-1}$.
- (xiii) Für alle $x \in K$ gilt $-x = (-1_K) \odot x$.
- (xiv) Für alle $x, y \in K$ gilt $(-x) \odot (-y) = x \odot y$.
- (xv) Falls $K \neq \{0_K\}$ so gilt $1_K \neq 0_K$.

Beweis. Wir verwenden hierfür nur die Spielregeln (A1)-(A4),(M1)-(M4) und (D) aus Definition 3.1.4.

Zu (i). Dass $x_0 = 0_K$ dem Nullgesetz genügt ist klar mit (A3).

Noch $\mathbb{Z}$: Es gibt keine weiteren Elemente die dem Nullgesetz genügen.

Seien dazu x_0 und x'_0 Elemente aus K , die dem Nullgesetz genügen, d.h.

$$x \oplus x_0 = x \quad \forall x \in K \quad (123)$$

und

$$x \oplus x'_0 = x \quad \forall x \in K. \quad (124)$$

Dann gilt

$$x_0 \stackrel{(124)}{=} x_0 \oplus x'_0 \stackrel{(A2)}{=} x'_0 \oplus x_0 \stackrel{(123)}{=} x'_0.$$

Wir folgern $x_0 = x'_0$.

Somit kann nur ein einziges Element dem Nullgesetz genügen, nämlich $x_0 = 0_K$.

Zu (ii). Analog zu (i).

Zu (iii). Dass das Element $x' = -x$ die Gleichung $x \oplus x' = 0_K$ erfüllt, ist klar mit (A4).

Noch $\mathbb{Z}$: Es gibt keine weiteren Elemente, die diese Gleichung erfüllen.

Seien dazu $x' \in K$ und $x'' \in K$ mit

$$x \oplus x' = 0_K \quad (125)$$

und

$$x \oplus x'' = 0_K. \quad (126)$$

Dann gilt

$$x' \underset{(A3)}{=} x' \oplus 0_K \underset{(126)}{=} x' \oplus (x \oplus x'') \underset{(A1)}{=} (x' \oplus x) \oplus x'' \underset{(A2)}{=} (x \oplus x') \oplus x'' \underset{(125)}{=} 0_K \oplus x'' \underset{(A2)}{=} x'' \oplus 0_K \underset{(A3)}{=} x''.$$

Es folgt $x' = x''$.

Somit kann nur ein einziges Element x' die Gleichung $x \oplus x' = 0_K$ lösen, nämlich $x' = -x$.

Zu (iv). Analog zu (iii).

Zu (v). Es gilt

$$-0_K \underset{(A3)}{=} -0_K \oplus 0_K \underset{(A2)}{=} 0_K \oplus (-0_K) \underset{(A4)}{=} 0_K.$$

Zu (vi). Es gilt

$$\begin{aligned} -(-x) &\underset{(A3)}{=} -(-x) \oplus 0_K \underset{(A4)}{=} -(-x) \oplus (x \oplus (-x)) \underset{(A2)}{=} -(-x) \oplus (-x \oplus x) \\ &\underset{(A1)}{=} (-(-x) \oplus (-x)) \oplus x \underset{(A2)}{=} (-x \oplus -(-x)) \oplus x \underset{(A4)}{=} 0_K \oplus x \underset{(A2, A3)}{=} x. \end{aligned}$$

Zu (vii). Analog zu (vi).

Zu (viii). Um zu zeigen, dass $-(x \oplus y) = (-x) \oplus (-y)$ genügt es, zu zeigen, dass $(-x) \oplus (-y)$ ein additiv Inverses von $x \oplus y$ ist. Denn laut Teil (iii) dieser Proposition gibt es nur ein einziges additiv Inverses von $x \oplus y$, nämlich $-(x \oplus y)$! Damit hätte man dann $-(x \oplus y) = (-x) \oplus (-y)$.

Also ist

$$\mathbb{Z}\text{Z: } (x \oplus y) \oplus ((-x) \oplus (-y)) = 0_K.$$

Dazu:

$$\begin{aligned} (x \oplus y) \oplus ((-x) \oplus (-y)) &\underset{(A1)}{=} x \oplus (y \oplus ((-x) \oplus (-y))) \underset{(A2)}{=} x \oplus (y \oplus ((-y) \oplus (-x))) \\ &\underset{(A1)}{=} x \oplus ((y \oplus (-y)) \oplus (-x)) \underset{(A4)}{=} x \oplus (0_K \oplus (-x)) \\ &\underset{(A2)}{=} x \oplus ((-x) \oplus 0_K) \underset{(A3)}{=} x \oplus (-x) \underset{(A3)}{=} 0_K. \end{aligned}$$

Die Behauptung folgt.

Zu (ix).

$$(x \oplus y) \odot z \stackrel{(M2)}{=} z \odot (x \oplus y) \stackrel{(D)}{=} z \odot x \oplus z \odot y \stackrel{(M2)}{=} x \odot z \oplus y \odot z.$$

Zu (x). Wir beobachten zuerst, dass

$$x \odot 0_K \stackrel{(A3)}{=} x \odot (0_K \oplus 0_K) \stackrel{(D)}{=} x \odot 0_K \oplus x \odot 0_K. \quad (127)$$

Wir folgern

$$\begin{aligned} 0_K &\stackrel{(A4)}{=} x \odot 0_K \oplus (-(x \odot 0_K)) \stackrel{(127)}{=} (x \odot 0_K \oplus x \odot 0_K) + (-(x \odot 0_K)) \\ &\stackrel{(A1)}{=} x \odot 0_K \oplus (x \odot 0_K + (-(x \odot 0_K))) \stackrel{(A4)}{=} x \odot 0_K \oplus 0_K \stackrel{(A3)}{=} x \odot 0_K. \end{aligned}$$

Zu (xi). Wir müssen hier eine Äquivalenz zeigen, nämlich $x \odot y = 0_K \iff (x = 0_K \vee y = 0_K)$. Hierzu müssen zwei Implikationen gezeigt werden, nämlich

$$„x \odot y = 0_K \implies (x = 0_K \vee y = 0_K)“ \text{ und } „x \cdot y = 0_K \iff (x = 0_K \vee y = 0_K)“.$$

Zu „ $\iff$ “. Es gelte $x = 0_K \vee y = 0_K$. Falls $x = 0_K$ so gilt $x \odot y \stackrel{\text{Vorr.}}{=} 0_K \odot y \stackrel{(M2)}{=} y \odot 0_K \stackrel{(x)}{=} 0_K$. Falls $y = 0_K$ so gilt $x \odot y \stackrel{\text{Vorr.}}{=} x \odot 0_K \stackrel{(x)}{=} 0_K$. Die Behauptung ist gezeigt.

Zu „ $\implies$ “. Wir zeigen die nach Proposition 1.1.6 äquivalente Aussage $\neg(x \odot y = 0_K \wedge \neg(x = 0_K \vee y = 0_K))$. Nach den De Morgan'schen Regeln ist diese Aussage äquivalent zu

$$\neg(x \odot y = 0_K \wedge x \neq 0_K \wedge y \neq 0_K).$$

Wir müssen also zeigen, dass diese Aussage wahr ist. Alternativ können wir zeigen, dass die Aussage

$$x \odot y = 0_K \wedge x \neq 0_K \wedge y \neq 0_K \quad (128)$$

falsch ist. Angenommen die Aussage in (128) ist wahr. Dann gilt auch

$$x \stackrel{(M3)}{=} x \odot 1_K \stackrel{(M4)}{=} x \odot (y \odot y^{-1}) \stackrel{(M1)}{=} (x \odot y) \odot y^{-1} \stackrel{(128)}{=} 0_K \odot y^{-1} \stackrel{(M2)}{=} y^{-1} \odot 0_K \stackrel{(xi)}{=} 0_K.$$

Dies widerspricht $x \neq 0_K$, vgl. (128). Wir folgern (xi).

Zu (xii). Analog zu (viii).

Zu (xiii). Es genügt wiederum zu zeigen, dass $(-1_K) \odot x$ ein additiv Inverses von x ist. Dazu:

$$\begin{aligned} x \oplus [(-1_K) \odot x] &\stackrel{(M2)}{=} x \oplus [x \odot (-1_K)] \stackrel{(M3)}{=} [x \odot 1_K] \oplus [x \odot (-1_K)] \\ &\stackrel{(D)}{=} x \odot (1_K \oplus (-1_K)) \stackrel{(A3)}{=} x \odot 0_K \stackrel{(x)}{=} 0_K. \end{aligned}$$

Dies beweist, dass $(-1_K) \odot x$ ein additiv Inverses von x ist. Mit der Eindeutigkeitsaussage (iii) schließen wir dann $(-1_K) \odot x = -x$.

Zu (xiv). Zunächst eine Beobachtung:

$$1_K \stackrel{(vi)}{=} -(-1_K) \stackrel{(xiii)}{=} (-1_K) \odot (-1_K). \quad (129)$$

Nun gilt

$$\begin{aligned} (-x) \odot (-y) &\stackrel{(xiii)}{=} ((-1_K) \odot x) \odot ((-1_K) \odot y) \stackrel{(M1)}{=} ((-1_K) \odot x) \odot (-1_K) \odot y \\ &\stackrel{(M2)}{=} ((x \odot (-1_K)) \odot (-1_K)) \odot y \stackrel{(M1)}{=} (x \odot ((-1_K) \odot (-1_K))) \odot y \\ &\stackrel{(129)}{=} (x \odot 1_K) \odot y \stackrel{(M2)}{=} x \odot y. \end{aligned}$$

Zu (xv). Wir zeigen die *Kontraposition*, d.h. $1_K = 0_K \Rightarrow K = \{0_K\}$. Aus Proposition 1.1.7 folgt dann die Behauptung^a. Sei also $1_K = 0_K$. Dann gilt für alle $x \in K$

$$x \stackrel{(M3)}{=} x \odot 1_K \stackrel{\text{Vorr.}}{=} x \odot 0_K \stackrel{(x)}{=} 0_K.$$

Somit muss jedes beliebige Element aus K mit 0_K übereinstimmen. Es folgt $K = \{0_K\}$.

^aWenn wir anstatt einer Aussage der Form $A \Rightarrow B$ die (äquivalente) Aussage $\neg B \Rightarrow \neg A$ zeigen, dann nennt man das einen *Beweis durch Kontraposition*.

17.11.2025 Wir haben in der vorigen Proposition gesehen, dass bei einem Körper das Nullelement und das Einselement eindeutig festgelegt sind. Dies rechtfertigt die Ausdrucksweise *das Nullelement* (anstatt *ein Nullelement*) und *das Einselement* (anstatt *ein Einselement*). Den Index „ K “ bei $0_K, 1_K$ werden wir häufig weglassen und ab sofort kurz 0 und 1 schreiben.

Notation 3.1.10 — NOTATION IN KÖRPERN

Wir treffen einige (gewohnte) Notations-konventionen für das Rechnen in Körpern.

Die Verknüpfungen $\oplus$ bzw. $\odot$ notieren wir ab sofort mit $+$ bzw. $\cdot$. Wir weisen aber an dieser

Stelle nochmal darauf hin, dass es sich nicht zwingend um die Addition oder Multiplikation auf reellen Zahlen handelt, sondern um abstrakte innere Verknüpfungen, die man für jeden Körper gesondert definieren muss.

Oft schreiben wir in einem Körper anstatt $x \cdot y$ nur xy . Ferner setzen wir für $y \in K \setminus \{0_K\}$

$$\frac{x}{y} := x \cdot y^{-1} = xy^{-1}.$$

Für $x, y, z \in K$ schreiben wir statt $x + (y + z)$ oder gleichwertig $(x + y) + z$ einfach

$$x + y + z := (x + y) + z.$$

Analoge Konvention treffen wir für

$$xyz := (xy)z.$$

Für beliebig viele Summanden $x_1, \dots, x_n \in K$ setzen wir auch

$$x_1 + x_2 + x_3 + \dots + x_n := (\dots((x_1 + x_2) + x_3) + \dots + x_{n-1}) + x_n$$

Analoge Konvention treffen wir für „“. Man beachte, dass hierbei die Reihenfolge der Summanden eigentlich keine Rolle spielt.

3.2 Anordnungen

Mengen heißen *angeordnet*, wenn man Elemente miteinander *vergleichen*, d.h. auf eine bestimmte Weise *in eine Relation setzen* kann. Was das bedeutet definieren wir nun.

Definition 3.2.1 — RELATION

Es sei M eine Menge. Eine *Relation* R auf M ist eine Teilmenge von $M \times M$. Anstatt $(x, y) \in R$ schreiben wir häufig $x R y$ und sagen *x steht in Relation R zu y* .

Nicht alle Relationen erlauben uns, Mengen anzuordnen. Wir werden nun Axiome vorstellen, die es uns erlauben, von einer *Ordnungsrelation* zu sprechen.

Definition 3.2.2 — ORDNUNGSRELATION

Es sei M eine Menge. Eine Relation $\leq$ heißt *Ordnungsrelation* auf M falls die folgenden Gesetze gelten.

(O1) *Reflexivität.* $x \leq x$ für alle $x \in M$.

(O2) *Antisymmetrie*. Es gilt $(x \leq y) \wedge (y \leq x) \Rightarrow x = y$.

(O3) *Transitivität*. Es gilt $(x \leq y) \wedge (y \leq z) \Rightarrow x \leq z$.

(O4) *Totalität*. Für alle $x, y \in M$ gilt eine der beiden Aussagen: $x \leq y$ oder $y \leq x$.

Ist $\leq$ eine Ordnungsrelation auf M , so heißt das Paar $(M, \leq)$ *angeordnete Menge*.

Erfüllt $\leq$ nur (O1)-(O3) (und nicht zwingend (O4)), so heißt $\leq$ eine *Halbordnungsrelation* und $(M, \leq)$ heißt eine *halbangeordnete Menge*.

Ist $(M, \leq)$ eine (halb-)angeordnete Menge, so definieren wir außerdem

$$x < y \quad :\Longleftrightarrow \quad (x \leq y) \wedge (x \neq y).$$

Wie bei allen mathematischen Symbolen darf man das Symbol „ $\leq$ “ in alle Richtungen lesen. Insbesondere bedeutet $x \geq y$ dasselbe wie $y \leq x$ (ohne, dass wir es gesondert definieren müssen).

Zukunfts-Beispiel 3.2.3

Wieder mal ein *Zukunfts-Beispiel* für eine Ordnungsrelation wäre die *Kleinergleich-Relation* $\leq$ auf $\mathbb{R}$ und $\mathbb{Q}$ (jedenfalls sofern wir diese Zahlenmengen schon richtig eingeführt hätten).

Beispiel 3.2.4 — BEISPIELE FÜR ORDNUNGSRELATIONEN

Es gibt aber noch weitere weniger geläufige Beispiele für (Halb-)Ordnungsrelationen. Ein Beispiel ist auf $M = \mathbb{N}$ die sogenannte *Teilbarkeitsrelation*

$$x \leq y \quad :\Longleftrightarrow \quad \exists m \in \mathbb{N} : y = mx.$$

Diese Relation ist

- *reflexiv*, da für alle $x \in \mathbb{N}$ gilt $x = 1 \cdot x$ und somit $x \leq x$,
- *antisymmetrisch*, denn sind $x, y \in \mathbb{N}$ mit $x \leq y$ und $y \leq x$, so gibt es $m_1 \in \mathbb{N}$ und $m_2 \in \mathbb{N}$ mit $y = m_1 x$ und $x = m_2 y$. Man folgert

$$x = m_2 y = m_2(m_1 x) = (m_2 m_1) x. \quad \Rightarrow \quad m_2 m_1 = 1.$$

Dies impliziert (wie man sich leicht überlegt) $m_2 = m_1 = 1$ und somit $x = m_2 y = y$ und $y = m_1 x = x$.

- *transitiv*, denn sind $x, y, z \in \mathbb{N}$ mit $x \leq y$ und $y \leq z$, so gibt es $m_1 \in \mathbb{N}$ mit $y = m_1 x$ und $m_2 \in \mathbb{N}$ mit $z = m_2 y$. Daher gilt $z = m_2 y = m_2 m_1 x$. Somit gilt $z = m_3 x$ für ein $m_3 \in \mathbb{N}$ (gegeben durch $m_3 = m_1 m_2$) Man folgert $x \leq z$.

- *nicht total*, denn es gilt z.B. weder $2 \leq 3$ noch $3 \leq 2$.

Wir haben also eine *Halbordnung* vorliegen, die nichts mit unserer klassischen *Kleiner-Relation* zu tun hat.

Wir kennen auch noch eine weitere Halbordnung. Es sei $\mathcal{M}$ die Menge aller Mengen^a. Für $M_1, M_2 \in \mathcal{M}$ definieren wir^b

$$M_1 \leq M_2 \quad :\Longleftrightarrow \quad M_1 \subseteq M_2. \quad (130)$$

Die Details überlassen wir den Lesenden als Übungsaufgabe. Man kann sich auch leicht davon überzeugen, dass es sich nicht um eine Ordnungsrelation sondern wirklich nur um eine Halbordnungsrelation handelt!

^aWenn wir ein bisschen mehr über Mengenlehre wüssten, käme jetzt wohl der berechtigte Einwand, dass es eine solche Menge gar nicht gibt. Um das Beispiel rigoros zu machen kann man aber die Menge aller Teilmengen einer gewissen *Grundmenge* A betrachten.

^bDie Notation $M_1 \leq M_2$ für Mengen M_1, M_2 benutzen wir NUR(!) in diesem Beispiel. Betrachten Sie dies nicht als global gebräuchliche Notation und fangen Sie bitte schonmal gar nicht an, diese Notation in Ihren Übungsabgaben zu benutzen. Es geht hier lediglich um die Illustration des Begriffes der Halbordnung.

Ein wichtiges Konzept in einer halbangeordneten Menge ist das einer Schranke.

Definition 3.2.5 — SCHRANKEN, MAXIMA, MINIMA

Es sei $(M, \leq)$ eine halbangeordnete Menge und $N \subseteq M$.

- (i) Ein Element $b \in M$ heißt *obere Schranke* von N , falls

$$\forall x \in N : \quad x \leq b.$$

Besitzt eine Menge N eine obere Schranke, so nennen wir N *nach oben beschränkt*.

- (ii) Ein Element $b \in M$ heißt *untere Schranke* von N , falls

$$\forall x \in N : \quad x \geq b.$$

Besitzt eine Menge N eine untere Schranke, so nennen wir N *nach unten beschränkt*.

- (iii) Eine Menge $N \subseteq M$ heißt *beschränkt*, wenn sie sowohl nach oben als auch nach unten beschränkt ist.

- (iv) Ein Element $b \in M$ heißt *Maximum* von N , falls

$$b \in N \quad \text{und} \quad b \text{ ist eine obere Schranke von } N.$$

(v) Ein Element $b \in M$ heißt *Minimum* von N ,

$$b \in N \quad \text{und} \quad b \text{ ist eine untere Schranke von } N.$$

Diese Begrifflichkeiten erlauben uns *Abschätzungen* für Elemente aus N . Optimale Kandidaten für Abschätzungen von Elementen von N sind stets Maxima und Minima. Allerdings gibt es einen Haken: Nicht jede Menge besitzt ein Maximum!

Zukunfts-Beispiel 3.2.6 — EINE MENGE OHNE MAXIMUM

Betrachte die angeordnete Menge $(\mathbb{R}, \leq)$. Die Menge

$$N := \{x \in \mathbb{R} : x < 2\}$$

besitzt kein Maximum.

Zukunfts-Beweis. Angenommen N besitzt ein Maximum, nennen wir es m . Da $m \in N$ gilt $m < 2$. Definiere nun $\tilde{m} := \frac{m+2}{2}$. Beachte: $\tilde{m} \in N$, da $\tilde{m} = \frac{m+2}{2} < \frac{2+2}{2} = \frac{4}{2} = 2$. Da m eine obere Schranke von N ist muss folglich gelten $\tilde{m} \leq m$. Jedoch gilt auch

$$\tilde{m} = \frac{m+2}{2} > \frac{m+m}{2} = \frac{2m}{2} = m.$$

Ein Widerspruch zu $\tilde{m} \leq m$.

Sofern es ein Maximum gibt, so ist dieses aber *eindeutig bestimmt*.

Proposition 3.2.7 — EINDEUTIGKEIT VON MAXIMUM UND MINIMUM

Es sei $(M, \leq)$ eine (halb-)angeordnete Menge und $N \subseteq M$.

(i) N besitzt höchstens ein Maximum.

(ii) N besitzt höchstens ein Minimum.

Beweis. Wir zeigen nur (i), denn der Beweis von (ii) geht analog. Es seien b und b' Maxima von N . Wir werden zeigen, dass $b = b'$. Wir wissen aus Definition 3.2.5

$$b \in N \text{ und } b \text{ ist obere Schranke von } N$$

sowie

$$b' \in N \text{ und } b' \text{ ist obere Schranke von } N.$$

Weil $b \in N$ und b' eine obere Schranke von N ist, gilt

$$b \leq b'. \tag{131}$$

Ferner: Weil $b' \in N$ und b eine obere Schranke von N ist, gilt

$$b' \leq b. \quad (132)$$

Aus (131) $\wedge$ (132) und (O2) in Definition 3.2.2 folgt, dass

$$b = b'.$$

Die Behauptung ist gezeigt.

Aufgrund dieser Eindeutigkeit sind auch die folgenden Notationen gerechtfertigt

Notation 3.2.8

Es sei $(M, \leq)$ eine (halb-)angeordnete Menge und $N \subseteq M$. Besitzt N ein Maximum und ist dieses gegeben durch $b \in N$ so schreiben wir

$$b = \max N.$$

Analog: Besitzt N ein Minimum und ist dieses gegeben durch $b \in N$ so schreiben wir

$$b = \min N.$$

Beispiel 3.2.9 — MAXIMUM ZWEELEMENTIGER MENGEN

Es sei $(M, \leq)$ eine (halb-) angeordnete Menge und $a, b \in M$ mit $a < b$.

Behauptung. Es gilt $\max(\{a, b\}) = b$. ^(a)

Beweis. b ist eine obere Schranke von $\{a, b\}$ denn $a \leq b$ und $b \leq b$. Außerdem gilt $b \in \{a, b\}$. Mit Definition 3.2.5(iii) folgt, dass $b = \max\{a, b\}$.

^aDas zeigt auch die Konsistenz der hier eingeführten Maximumsbegriffes mit dem aus Bemerkung 2.0.2

Auch wenn eine Menge mal kein Maximum besitzt, würden wir gerne *möglichst optimal abschätzen*. In Beispiel 3.2.6 ist die Zahl 2 doch irgendwie schon eine optimale obere Schranke von N ...

Um dieses Phänomen besser zu fassen gibt es das Konzept vom *Supremum und Infimum*.

Definition 3.2.10 — SUPREMUM UND INFIMUM

Es sei $(M, \leq)$ eine (halb-)angeordnete Menge und $N \subseteq M$, $N \neq \emptyset$.

- (i) Ein Element $s \in M$ heißt *Supremum von N* falls
 - (i.a) s eine obere Schranke von N ist und
 - (i.b) Für jede weitere obere Schranke s' von N gilt, dass $s \leq s'$.

(ii) Ein Element $a \in M$ heißt *Infimum von N* falls

(ii.a) a eine untere Schranke von N ist und

(ii.b) Für jede weitere untere Schranke a' von N gilt, dass $a' \leq a$.

Mit anderen Worten: Ein Supremum ist eine *kleinstmögliche obere Schranke* und ein Infimum ist eine *größtmögliche untere Schranke*.

Wir werden nun herausfinden, dass jedes Maximum auch ein Supremum ist. Allerdings gibt es (wie wir auch sehen werden) deutlich mehr Mengen, die ein Supremum besitzen als diejenigen, die ein Maximum besitzen.

Proposition 3.2.11

Es sei $(M, \leq)$ eine (halb-)angeordnete Menge und $N \subseteq M$, $N \neq \emptyset$.

(i) Ist $m = \max N$, so ist m auch ein Supremum von N .

(ii) Ist $m = \min N$, so ist m auch ein Infimum von N .

Beweis. Wieder mal beweisen wir nur (i). Es sei $m = \max N$. Um zu zeigen, dass m ein Supremum ist, müssen die Bedingungen (i.a) und (i.b) aus Definition 3.2.10 nachgerechnet werden.

ZU BEDINGUNG (i.a). Diese ist erfüllt, denn jedes Maximum von N ist nach Definition 3.2.5 eine obere Schranke von N .

ZU BEDINGUNG (i.b). Es sei s' eine beliebige obere Schranke von N . Da $m \in N$ muss nach Definition 3.2.5 (i) [angewendet auf $b = s'$] gelten, dass $m \leq s'$. Es folgt also auch (i.b).

Wie bereits angekündigt, greifen wir nun Beispiel 3.2.6 und diskutieren, dass die dortige Menge zwar kein Maximum aber ein Supremum besitzt.

Zukunfts-Beispiel 3.2.12 — EINE MENGE OHNE MAXIMUM (ABER MIT SUPREMUM)

Betrachte wie in Beispiel 3.2.6 die angeordnete Menge $(\mathbb{R}, \leq)$ und

$$N := \{x \in \mathbb{R} : x < 2\}.$$

Wir haben schon angekündigt, dass N kein Maximum besitzt. Jedoch besitzt N ein Supremum, nämlich 2! Auch das zeigen wir später.

Auffällig ist, dass das Supremum 2 in diesem Falle kein Element von N ist!

Leider besitzen trotzdem nicht alle Mengen ein Supremum. Klar: Nach oben unbeschränkte Mengen besitzen kein Supremum, denn dort gibt es ja nichtmal irgendeine obere Schranke. Aber auch nach oben beschränkte Mengen müssen kein Supremum haben:

Zukunfts-Beispiel 3.2.13 — EINE MENGE OHNE SUPREMUM

Betrachte die angeordnete Menge $(\mathbb{Q}, \leq)$. Die Menge

$$N := \{x \in \mathbb{Q} : x^2 \leq 2\}$$

besitzt kein Supremum. Dies werden wir später beweisen.

Falls eine Menge jedoch ein Supremum besitzt, so ist es eindeutig bestimmt.

Proposition 3.2.14 — EINDEUTIGKEIT DES SUPREMUMS

Es sei $(M, \leq)$ eine (halb-)angeordnete Menge und $N \subseteq M, N \neq \emptyset$.

- (i) N besitzt höchstens ein Supremum.
- (ii) N besitzt höchstens ein Infimum.

Beweis. Wieder mal zeigen wir nur (i). Aussage (ii) ist analog.

Zu (i). Der Beweis benutzt einen cleveren Trick. Definiere zunächst die Menge

$$S := \{b \in M : b \text{ ist obere Schranke von } N\}. \quad (133)$$

Wir zeigen, dass jedes Supremum von N ein Minimum von S ist. Da nun S nach Proposition 3.2.7 höchstens ein Minimum besitzen darf, kann es somit auch nur höchstens ein Supremum von N geben. Also

$$\text{Zz: } s \text{ ist Supremum von } N \implies s = \min S.$$

Sei nun s ein Supremum von N . Wir müssen beweisen, dass $s \in S$ und s ist untere Schranke von S , vgl. Definition 3.2.5 (v).

Zu „ $s \in S$ “:

$$\begin{array}{ll} & s \text{ ist Supremum von } N \\ \xRightarrow{\text{Def3.2.10(i.a)}} & s \text{ ist obere Schranke von } N \\ \xRightarrow{(133)} & s \in S. \end{array}$$

Zu „ s ist untere Schranke von S “:

$$\begin{array}{ll} & s' \in S \\ \xRightarrow{(133)} & s' \text{ ist obere Schranke von } N \\ \xRightarrow{\text{Def3.2.10(i.b)}} & s \leq s'. \end{array}$$

Somit gilt $s \leq s'$ für alle $s' \in S$. Mit anderen Worten s ist eine untere Schanke von S , vgl. Definition 3.2.5 (ii).

Notation 3.2.15

Es sei $(M, \leq)$ eine (halb-)angeordnete Menge und $N \subseteq M$, $N \neq \emptyset$.

- (i) Besitzt N ein Supremum $s \in M$, so bezeichnen wir es mit $s = \sup N$.
- (ii) Besitzt N ein Infimum $a \in M$, so bezeichnen wir es mit $a = \inf N$.

Angeordnete Körper

Ziel des nachfolgenden Abschnittes ist es, die *Körperstruktur* mit der *Anordnungsstruktur* zu verflechten. Hierfür lernen wir das Konzept eines *angeordneten Körpers* kennen.

Definition 3.2.16 — ANGEORDNETER KÖRPER

Ein *angeordneter Körper* $(K, +, \cdot, \leq)$ ist eine Menge K mit zwei inneren Verknüpfungen $+, \cdot : K \times K \rightarrow K$ und einer Relation $\leq$ auf K so, dass

- (i) $(K, +, \cdot)$ ist ein Körper mit Nullelement 0 und Einselement 1 [d.h. es gelten (A1)-(A4), (M1)-(M4) und (D) aus Definition 3.1.4].
 - (ii) $(K, \leq)$ ist eine angeordnete Menge [d.h. es gelten (O1)-(O4) aus Definition 3.2.2]
- (V1) *Verträglichkeit von $+$ und $\leq$* . Es gilt für alle $x, y, z \in K$

$$x \leq y \implies x + z \leq y + z.$$

- (V2) *Verträglichkeit von $\cdot$ und $\leq$* . Es gilt für alle $x, y \in K$

$$(x \geq 0 \wedge y \geq 0) \implies (x \cdot y \geq 0).$$

Erfüllt ein $x \in K$ die Ungleichung $x \geq 0$ (bzw. $x > 0$) so nennen wir x *schwach positiv* (bzw. *strikt positiv*). Erfüllt hingegen ein $x \in K$ die Ungleichung $x \leq 0$ (bzw. $x < 0$) so nennen wir x *negativ* (bzw. *strikt negativ*).

Bemerkung 3.2.17 — TRICHOTOMIE

Sei $(K, +, \cdot, \leq)$ ein angeordneter Körper. Wegen der Totalität der Ordnungsrelation $\leq$ [(O4) in Definition 3.2.2] gilt für alle $x \in K$, dass $x \geq 0$ oder $x \leq 0$. Sprich: Jedes $x \in K$ ist positiv oder negativ.

Beides gleichzeitig kann aber nur für $x = 0$ gelten, denn mit (O2) aus Definition 3.2.2

folgt man

$$(x \geq 0) \wedge (x \leq 0) \implies x = 0.$$

Man sieht, dass für jedes $x \in K$ genau eine der folgenden drei Eigenschaften zutrifft

$$x > 0, \quad x = 0, \quad x < 0.$$

Wir nennen dies eine *Trichotomie*.

Zukunfts-Beispiel 3.2.18

Wir werden sehen, dass $(\mathbb{R}, +, \cdot, \leq)$ und $(\mathbb{Q}, +, \cdot, \leq)$ angeordnete Körper sind (sobald wir $\mathbb{R}$ und $\mathbb{Q}$ richtig eingeführt haben).

Nicht-Beispiel 3.2.19

Betrachte $K = \{\text{Sonne, Mond, Sterne}\}$ mit $+$ $= \oplus$, $\cdot = \odot$ definiert wie in Beispiel 3.1.3. Wir werden sehen: Es ist nicht möglich, eine Ordnungsrelation $\leq$ zu definieren, so, dass $(K, +, \cdot, \leq)$ ein angeordneter Körper ist.

Proposition 3.2.20 — RECHNEN IN ANGEORDNETEN KÖRPERN

Es sei $(K, +, \cdot, \leq)$ ein angeordneter Körper. Dann gilt

- (i) Für $x, y, z \in K$ hat man $x < y \implies x + z < y + z$.
- (ii) Für $x, y \in K$ hat man $(x > 0) \wedge (y > 0) \implies xy > 0$.
- (iii) Für $x, y \in K$ gilt $x \leq y \iff x - y \leq 0$. Selbige Behauptung gilt mit $<$.
- (iv) Für $x, y \in K$ gilt $x \geq y \iff x - y \geq 0$. Selbige Behauptung gilt mit $>$.
- (v) Für $y \in K$ gilt $y \geq 0 \iff -y \leq 0$ sowie $y \leq 0 \iff -y \geq 0$. Selbige Behauptungen gelten mit $>$ bzw. $<$.
- (vi) Für $x, y \in K$ gilt $x \leq y \iff -y \leq -x$.
- (vii) Für $x, y, a, b \in K$ gilt $(x \leq y) \wedge (a \leq b) \implies x + a \leq y + b$.
- (viii) Für $x, y, a \in K$ gilt $(x \leq y) \wedge (0 \leq a) \implies ax \leq ay$.
- (ix) Für $x, y, a \in K$ gilt $(x \leq y) \wedge (a \leq 0) \implies ax \geq ay$.
- (x) Für alle $x \in K$ ist $x \cdot x \geq 0$.
- (xi) Ist $K \neq \{0\}$ so ist $1 > 0$.
- (xii) Für $x \in K \setminus \{0\}$ hat man $x > 0 \iff x^{-1} > 0$.

(xiii) Für $x, y \in K$ mit $x > 0$ und $y > 0$ ^(a) gilt $(x \leq y \implies y^{-1} \leq x^{-1})$.

^aDies würden wir ab sofort mit „Für $x, y > 0$ “ abkürzen.

20.11.2025

Beweis. Mal wieder spielen wir nur nach den *Spielregeln* (A1)-(A4), (M1)-(M4), (D) aus Definition 3.1.4 und (O1)-(O4) aus Definition 3.2.2 sowie (V1) und (V2) aus Definition 3.2.16.

Zu (i). Wir zeigen „ $x < y \implies x + z \leq y + z$ “ und „ $x < y \implies x + z \neq y + z$ “. Zur ersten Aussage:

$$x < y \implies x \leq y \xRightarrow{(V1)} x + z \leq y + z.$$

Die zweite Aussage beweisen wir durch Kontraposition.

$$x + z = y + z \implies (x + z) - z = (y + z) - z \xRightarrow{(A1), (A3)} x = y \xRightarrow{(A1), (A3)} \neg(x < y).$$

Die Behauptung ist gezeigt.

Zu (ii). Sei $(x > 0) \wedge (y > 0)$. Dass $xy \geq 0$ folgt direkt aus (V2). Dass $xy \neq 0$ beweisen wir durch Widerspruch. Wäre $xy = 0$, so wäre laut Proposition 3.1.9(xi) $x = 0$ oder $y = 0$, was $x > 0 \wedge y > 0$ widerspricht.

Zu (iii). Zunächst zu „ $\implies$ “.

$$x \leq y \xRightarrow{(V1)} x + (-y) \leq y + (-y) \xRightarrow{(A3)} x + (-y) \leq 0 \xRightarrow{\text{Not3.1.10}} x - y \leq 0.$$

Nun zu „ $\impliedby$ “.

$$x - y \leq 0 \xRightarrow{(V1)} (x - y) + y \leq 0 + y \xRightarrow{(A1, A3)} x + (-y) + y \leq y \xRightarrow{(A4)} x \leq y.$$

Dass dieselbe Behauptung auch mit der strikten Ungleichung gilt, überlassen wir den Lesenden als Übungsaufgabe.

Zu (iv). Geht analog zu (iii).

Zu (v). Die erste Aussage folgt aus (iii) mit $x = 0$. Die zweite Aussage folgt aus (iv) mit $x = 0$.

Zu (vi). Wir berechnen zuerst

$$x - y \stackrel{\text{Not3.1.10}}{=} x + (-y) \stackrel{\text{Prop3.1.9(vi)}}{=} -(-x) + (-y) \stackrel{(A2)}{=} (-y) - (-x). \quad (134)$$

$$\begin{aligned} x \leq y & \stackrel{(iii)}{\iff} x - y \leq 0 \\ & \stackrel{(134)}{\iff} (-y) - (-x) \leq 0 \\ & \stackrel{(iii)}{\iff} -y \leq -x. \end{aligned}$$

Zu (vii). Es gilt

$$\begin{aligned} & (x \leq y) \wedge (a \leq b) \\ \implies_{(V1)} & (x + a \leq y + a) \wedge (a + y \leq b + y) \\ \implies_{(A2)} & (x + a \leq y + a) \wedge (y + a \leq y + b) \\ \implies_{(O3)} & x + a \leq y + b. \end{aligned}$$

Zu (viii). Als Vorüberlegung beachte

$$ax \leq ay \iff ay \geq ax \stackrel{(iv)}{\iff} ay - ax \geq 0 \stackrel{(D)}{\iff} a(y - x) \geq 0. \quad (135)$$

Nun gilt

$$\begin{aligned} & (x \leq y) \wedge (0 \leq a) \\ \iff & (y \geq x) \wedge (a \geq 0) \\ \iff_{(iv)} & (y - x \geq 0) \wedge (a \geq 0) \\ \iff & (a \geq 0) \wedge (y - x \geq 0) \\ \implies_{(V2)} & a(y - x) \geq 0 \\ \implies_{(135)} & ax \leq ay. \end{aligned}$$

Zu (ix).

$$(x \leq y) \wedge (a \leq 0)$$

$$\begin{aligned}
& \xLeftrightarrow[(v)] (x \leq y) \wedge (-a \geq 0) \\
& \xRightarrow[(viii)] (-a)x \leq (-a)y \\
& \xLeftrightarrow[\text{Prop3.1.9(xiii)}] (-1)ax \leq (-1)ay \\
& \xLeftrightarrow[\text{Prop3.1.9(xiii)}] -(ax) \leq -(ay) \\
& \xLeftrightarrow[(vi)] ay \leq ax
\end{aligned}$$

Zu (x). Sei $x \in K$. Wegen (O4) gilt entweder $x \geq 0$ oder $x \leq 0$. Wir unterscheiden diese beiden Fälle.

$$\text{FALL 1: } x \geq 0. \quad x \geq 0 \xRightarrow[(V2)]{ } x \cdot x \geq 0.$$

$$\begin{aligned}
\text{FALL 2: } x \leq 0. \quad x \leq 0 & \xRightarrow[(v)]{ } (-x) \geq 0 \\
& \xRightarrow[\text{Fall 1}]{ } (-x) \cdot (-x) \geq 0 \\
& \xRightarrow[\text{Prop3.1.9(xiv)}]{ } x \cdot x \geq 0.
\end{aligned}$$

Zu (xi). Sei $K \neq \{0\}$. Dann ist nach Proposition 3.1.9(xv) $1 \neq 0$. Außerdem gilt

$$1 \xRightarrow[(M3)]{ } 1 \cdot 1 \geq 0. \quad (x)$$

Daher haben wir $1 > 0$.

Zu (xii). Zuerst zu „ $\implies$ “. Dass $x^{-1} \neq 0$ folgt direkt aus (M3) in der Definition eines Körpers. Wir müssen also nur $x^{-1} \geq 0$ zeigen. Beachte: Nach (x) gilt

$$x^{-1} \cdot x^{-1} \geq 0. \quad (136)$$

Damit folgern wir

$$\begin{aligned}
& x > 0 \\
& \implies x \geq 0 \\
& \xRightarrow[(136)]{ } (x \geq 0) \quad \wedge \quad (x^{-1} \cdot x^{-1} \geq 0) \\
& \xRightarrow[(V2)]{ } (x^{-1} \cdot x^{-1}) \cdot x \geq 0
\end{aligned}$$

$$\stackrel{\text{(M4)}}{\implies} x^{-1} \geq 0.$$

Nun zu „ $\Leftarrow$ “. Hierfür benutzen wir, dass wir „ $\implies$ “ bereits gezeigt haben.

$$x^{-1} > 0 \stackrel{\text{„}\Leftarrow\text{“}}{\implies} (x^{-1})^{-1} > 0 \stackrel{\text{Prop 3.1.9(vii)}}{\implies} x > 0.$$

Zu (xiii). Zunächst eine Vorüberlegung:

$$(x > 0) \wedge (y > 0) \stackrel{\text{(ii)}}{\implies} xy > 0 \stackrel{\text{(xii)}}{\implies} (xy)^{-1} > 0 \stackrel{\text{Prop 3.1.9(ix)}}{\implies} x^{-1}y^{-1} > 0.$$

Daher gilt

$$x \leq y \stackrel{\text{(viii)}}{\implies} (x^{-1}y^{-1})x \leq (x^{-1}y^{-1})y \stackrel{\text{(M1)-(M4)}}{\implies} y^{-1} \leq x^{-1}.$$

Im Folgenden wollen wir noch einige spezielle Elemente von angeordneten Körpern genauer studieren.

Beispiel 3.2.21 — DIE ELEMENTE 2 UND 3.

Es sei $(K, +, \cdot, \leq)$ ein angeordneter Körper mit $K \neq \{0\}$. Das Einselement notieren wir wie immer mit $1 := 1_K$. Wir definieren zusätzlich die Elemente $2 := 1 + 1$ und $3 := 2 + 1$.

Behauptung: $3 > 2 > 1 > 0$. Insbesondere sind alle diese vier Elemente verschieden.

Beweis. Dass $1 > 0$ ist gerade Proposition 3.2.20 (xi). Dass $2 > 1$ folgt so:

$$2 = 1 + 1 \stackrel{\text{Prop 3.2.20(i)}}{>} 1 + 0 = 1.$$

Analog folgt $3 > 2$, denn

$$3 = 2 + 1 \stackrel{\text{Prop 3.2.20(i)}}{>} 2 + 0 = 2.$$

Beispiel 3.2.22

Mithilfe des vorigen Beispiels sieht man auch, dass der Körper $(\{\text{Sonne, Mond, Sterne}\}, \oplus, \otimes)$ aus Beispiel 3.1.8 durch keine Anordnungsrelation zu einem angeordneten Körper gemacht werden kann! Denn wie oben gezeigt hat jeder angeordnete Körper mindestens vier verschiedene Elemente!

Sie können sogar als Übungsaufgabe zeigen, dass jeder angeordnete Körper $K \neq \{0\}$ unendlich viele Elemente enthalten muss.

Beispiel 3.2.23 — DAS ELEMENT $\frac{1}{2}$.

Sei wie in den vorigen Beispielen $(K, +, \cdot, \leq)$ angeordneter Körper und

$$2 := 1 + 1. \quad (137)$$

Wir definieren (konsistent mit unserer vorherigen Notation) $\frac{1}{2} := 2^{-1}$. Beachte: Nach Proposition 3.2.20 (xii) gilt $\frac{1}{2} > 0$. Wir lernen hier eine weitere nützliche Eigenschaft des Elementes $\frac{1}{2}$ kennen.

Behauptung. Für alle $x, y \in K$ mit $x < y$ gilt $x < \frac{1}{2}(x + y) < y$.

Beweis. Wir zeigen nur $x < \frac{1}{2}(x + y)$. Die andere Ungleichung überlasse ich den Lesenden als Übungsaufgabe. Dazu:

$$x < \frac{1}{2}(x + y) \quad \xLeftrightarrow[\text{Prop 3.2.20 (iv)}] \quad \frac{1}{2}(x + y) - x > 0.$$

Nun gilt

$$\begin{aligned} \frac{1}{2}(x + y) - x & \stackrel{(M3)}{=} \frac{1}{2}(x + y) - 1x \\ & \stackrel{(M4)}{=} \frac{1}{2}(x + y) - \frac{1}{2}(2x) \\ & \stackrel{(D)}{=} \frac{1}{2}((x + y) - 2x) \\ & \stackrel{(137)}{=} \frac{1}{2}((x + y) - (1 + 1)x) \\ & \stackrel{(D)}{=} \frac{1}{2}(x + y - (x + x)) \\ & \stackrel{\text{Prop 3.1.9(viii)}}{=} \frac{1}{2}(x + y - x - x) \\ & \stackrel{(A3)}{=} \frac{1}{2}(y - x). \end{aligned}$$

Nun gilt (da $x < y$), dass $y - x > 0$, vgl. Proposition 3.2.20 (iv). Da auch $\frac{1}{2} > 0$ (s.o.) folgt $\frac{1}{2}(y - x) > 0$, vgl. Proposition 3.2.20 (ii). Es ergibt sich

$$\frac{1}{2}(x + y) - x = \frac{1}{2}(y - x) > 0.$$

Die Behauptung ist gezeigt.

Solche Rechnungen wie in dem obigen Beispiel zeigen, dass wir in angeordneten Körpern so rechnen können, wie wir das gewohnt sind. Dies wollen wir ab sofort machen. Wir werden auch die benutzten Rechenregeln nicht mehr (oder nur noch in Ausnahmefällen) unter das Gleichheitszeichen schreiben.

Intervalle Wichtige Teilmengen angeordneter Körper sind *Intervalle*.

Definition 3.2.24 — INTERVALLE

Es sei $(K, +, \cdot, \leq)$ ein angeordneter Körper und $a, b \in K$, $a \leq b$. Dann definieren wir die folgenden Mengen

$$[a, b]_K := \{x \in K : (x \geq a) \wedge (x \leq b)\}, \quad (138)$$

$$(a, b)_K := \{x \in K : (x > a) \wedge (x < b)\}, \quad (139)$$

$$[a, b)_K := \{x \in K : (x \geq a) \wedge (x < b)\},$$

$$(a, b]_K := \{x \in K : (x > a) \wedge (x \leq b)\}.$$

Solche Mengen nennen wir *Intervalle*. Ist der angeordnete Körper K aus dem Kontext ersichtlich, so lassen wir den Index K auch weg.

Proposition 3.2.25

Es sei $(K, +, \cdot, \leq)$ ein angeordneter Körper und $a, b \in K$, $a < b$. Dann gilt

- (i) $\max([a, b]_K) = \sup([a, b]_K) = b$.
- (ii)
 - $\sup([a, b)_K) = b$.
 - Das Maximum von $[a, b)_K$ existiert nicht.

Analoge Aussagen gelten für andere Intervalltypen.

Beweis. Zu (i). Es genügt wegen Proposition 3.2.2 (i) zu zeigen, dass $\max([a, b]_K) = b$. Dazu: $b \in [a, b]_K$ ist richtig, denn $b \geq a$ und $b \leq b$. Ferner ist b eine obere Schranke von $[a, b]_K$, denn

$$x \in [a, b]_K \quad \xRightarrow{(138)} \quad (x \geq a) \wedge (x \leq b) \quad \xRightarrow{(a)} \quad x \leq b.$$

Zu (ii). TEIL1.

Zunächst zeigen wir, dass $\sup([a, b)_K) = b$. Wir prüfen dazu die Bedingungen (i.a) und (i.b) aus Definition 3.2.10.

Zu (i.a). b ist eine obere Schranke von $[a, b)_K$, denn

$$x \in [a, b)_K \quad \xRightarrow{(139)} \quad (x \geq a) \wedge (x < b) \quad \Rightarrow \quad x < b \quad \Rightarrow \quad x \leq b.$$

Zu (i.b). Es sei s' eine weitere obere Schranke von $[a, b)_K$. Beachte: Es gilt $s' \geq a$, denn $a \in [a, b)_K$.

$$\text{ZZ: } s' \geq b.$$

Mal angenommen, dass $s' < b$. Definiere

$$s'' := \frac{1}{2}(s' + b).$$

Dann gilt nach Beispiel 3.2.23, dass

$$s' < s'' < b. \quad (140)$$

Beachte: Da $s' \geq a$ folgt auch $s'' \geq a$. Aus diesen Überlegungen folgt $(s'' \geq a) \wedge (s'' < b)$. Mit anderen Worten: $s'' \in [a, b)_K$, vgl. Definition 3.2.24. Da s' aber eine obere Schranke von $[a, b)_K$ ist, würde folgen

$$s'' \leq s'.$$

Ein Widerspruch (140).

TEIL2.

Jetzt zeigen wir noch Nichtexistenz des Maximums: Angenommen $[a, b)_K$ besitzt ein Maximum. Dann würden nach Proposition 3.2.2 Supremum und Maximum übereinstimmen, d.h. $\max([a, b)_K) = \sup([a, b)_K) = b$. Da ein Maximum einer Menge nach Definition 3.2.5 stets in der Menge selbst enthalten ist, folgt $b \in [a, b)_K$. Jedoch gilt $b \notin [a, b)_K$ (da $b \not< b$).

Ein Widerspruch

QED

^aIn diesem Schritt benutzen wir, dass die Aussage $A \wedge B \Rightarrow B$ für beliebige Aussagen A, B stets wahr ist.

Sie können sich analog überlegen, ob Maxima/Minima/Suprema/Infima von sämtlichen Intervallen aus Definition 3.2.24 existieren und welche Werte sie annehmen.

Suprema und Infima in angeordneten Körpern Sobald man eine Anordnung vorliegen hat, kann man auch von Suprema und Infima sprechen, wie wir im vorigen Abschnitt schon getan haben. Im Folgenden wollen wir eine nützliche Charakterisierung von Suprema und Infima in angeordneten Körpern besprechen.

Proposition 3.2.26

Es sei $(K, +, \cdot, \leq)$ ein angeordneter Körper und $N \subseteq K$, $N \neq \emptyset$.

(i) Sei $s \in K$ eine obere Schranke von N . Dann gilt

$$s = \sup N \iff \forall \varepsilon > 0 \exists x_0 \in N : s - \varepsilon < x_0.$$

(ii) Sei $a \in K$ eine untere Schranke von N . Dann gilt

$$a = \inf N \iff \forall \varepsilon > 0 \exists x_0 \in N : x_0 < a + \varepsilon.$$

Beweis. Mal wieder zeigen wir nur (i).

Zu „ $\implies$ “. Es sei $s = \sup N$.

$\mathbb{Z}$: Zu jedem $\varepsilon > 0$ gibt es ein Element x_0 aus N mit $s - \varepsilon < x_0$.

Dazu fixiere ein beliebiges $\varepsilon > 0$. Dann gilt

$$s - \varepsilon < s - 0 = s.$$

Mit Definition 3.2.10 (i.b) folgt

$s - \varepsilon$ ist keine obere Schranke von N .

Nach Definition 3.2.5 bedeutet das

$$\neg(x \leq s - \varepsilon \quad \forall x \in N)$$

Wie man eine solche Aussage verneint haben wir schon in Vorüberlegung 2.2.4 gesehen ^{a!}
Wir folgern

$$\exists x_0 \in N : s - \varepsilon < x_0.$$

Da wir diese Aussage für *jedes beliebige* $\varepsilon > 0$ verifiziert haben, gilt

$$\forall \varepsilon > 0 \quad \exists x_0 \in N : s - \varepsilon < x_0.$$

Nun zu „ $\impliedby$ “. Es sei $s \in K$ eine obere Schranke von N und ferner gelte

$$\forall \varepsilon > 0 \quad \exists x_0 \in N : s - \varepsilon < x_0. \tag{141}$$

$$\mathbb{Z} : s = \sup N.$$

Wir prüfen dazu wiederum (i.a) und (i.b) aus Definition 3.2.10. Dass s eine obere Schranke von N ist folgt bereits aus der Voraussetzung, also ist (i.a) abgehakt. Noch zu (i.b). Es sei nun s' eine weitere obere Schranke von N .

$$\mathbb{Z} : s \leq s'.$$

Angenommen $s' < s$. Dann ist $\varepsilon_0 := s - s' > 0$. Mit (141) (angewendet auf $\varepsilon = \varepsilon_0$) folgt

$$\exists x_0 \in N : s - \varepsilon_0 < x_0.$$

Man folgert (wegen der Wahl von ε_0)

$$\exists x_0 \in N : s - (s - s') < x_0.$$

Die Rechenregeln in Körpern implizieren

$$\exists x_0 \in N : s' < x_0.$$

Dann ist s' aber keine obere Schranke von N . Ein Widerspruch.

QED

^aDiese Vorüberlegung kommt aus Kapitel 2, welches wir eigentlich nicht verwenden dürfen. Sie benutzt aber ausschließlich Erkenntnisse aus Kapitel 1. Daher ist die Verwendung zulässig!

Bemerkung 3.2.27

Seien $(K, +, \cdot, \leq)$, N und s wie in Proposition 3.2.26(i). Da s nach Voraussetzung eine obere Schranke ist, könnte man Teil (i) dieser Proposition auch schreiben als

$$s = \sup N \iff \forall \varepsilon > 0 \exists x_0 \in N : s - \varepsilon < x_0 \leq s.$$

Anders ausgedrückt:

$$s = \sup N \iff \forall \varepsilon > 0 \exists x_0 \in N : x_0 \in (s - \varepsilon, s]_K.$$

3.3 Natürliche, Ganze und Rationale Zahlen

In diesem Abschnitt werden wir eine Menge einführen, die uns intuitiv eigentlich geläufig sein müsste, nämlich die Menge $\mathbb{N}$ der *natürlichen Zahlen*. Wir haben sie in Gleichung (6) schonmal intuitiv benutzt, jedoch gibt es dort ein Problem: Um $\mathbb{N} = \{1, 2, 3, 4, 5, 6, \dots\}$ schreiben zu können, müssen wir eigentlich bereits wissen, was die natürlichen Zahlen sind — Denn die „ $\dots$ “-Notation geht stets davon aus, dass wir bereits eine *Fortsetzungsregel* kennen. Über eine solche Regel haben wir aber noch nicht reflektiert.

Wir werden hier die natürlichen Zahlen zunächst abstrakt über die sogenannten *Peano-Axiome* einführen.

Definition 3.3.1 — NATÜRLICHE ZAHLEN

Ein Tripel $(\mathbb{N}, \alpha, s)$ heißt *Realisierung der natürlichen Zahlen*, falls $\mathbb{N}$ eine Menge ist, α ein Element von $\mathbb{N}$ ist (genannt *Alpha-Element*), und $s : \mathbb{N} \rightarrow \mathbb{N}$ (genannt *Nachfolgerabbildung*) eine Abbildung ist mit den folgenden Eigenschaften

(P1) $s(n) \neq \alpha$ für alle $n \in \mathbb{N}$.

(P2) Es gilt für alle $n, m \in \mathbb{N}$, dass $s(n) = s(m) \implies n = m$

(P3) Erfüllt eine Teilmenge $X \subseteq \mathbb{N}$

(a) $\alpha \in X$ und

(b) $n \in X \implies s(n) \in X$

so ist $X = \mathbb{N}$.

Für $n \in \mathbb{N}$ nennen wir $s(n)$ dann stets *Nachfolger* von n .

Bemerkung 3.3.1. Die natürlichen Zahlen sind auf diese Weise sehr abstrakt definiert. Dinge, die Sie in den natürlichen Zahlen gerne tun würden (rechnen/zählen/etc.) werden durch die Axiome (P1), (P2), (P3) zunächst nicht abgebildet.

Das *Zählen* ist vielleicht schon machbar, wenn Sie

$$\alpha, \quad s(\alpha), \quad s(s(\alpha)), \quad s(s(s(\alpha))), \quad \dots$$

betrachten. Man kann zeigen, dass alle oben aufgelisteten Elemente verschieden sind^a. Jedoch wird aus (P1),(P2),(P3) nicht ersichtlich, ob sie mit dieser Zählbarkeit *alle* natürlichen Zahlen auflisten können. Ein Dilemma, welches uns in diesem Kapitel begleiten wird.

^aÜberlegen Sie sich das mal als Übungsaufgabe! Eine Starthilfe: Nach (P1) ist $s(\alpha) \neq \alpha$. Aus (der Kontraposition von) (P2) folgt dann $s(s(\alpha)) \neq s(\alpha)$ und aus (P1) folgt auch $s(s(\alpha)) \neq \alpha$.

Das Gesetz (P2) besagt nichts anderes als dass die Nachfolgerabbildung $s : \mathbb{N} \rightarrow \mathbb{N}$ injektiv ist. Beachte: s ist aber nicht surjektiv, da nach (P1) gilt $\alpha \notin s(\mathbb{N})$.

Bonusmaterial. DAS BILD DER NACHFOLGERABBILDUNG.

Wie wir gesehen haben ist die Nachfolgerabbildung in den natürlichen Zahlen nicht surjektiv. Schön wäre es doch, wenn wir das gesamte *Bild der Nachfolgerabbildung* aus (P1),(P2),(P3) bestimmen könnten. Und in der Tat können wir zeigen: Jedes Element außer das α -Element wird von der Nachfolgerabbildung getroffen. Dies rechtfertigt in meinen Augen das *rückwärts-zählen* was Sie sicher im Alltag (und wir auch schon in Kapitel 2) oft getan haben.

Proposition 3.3.2 — BILD DER NACHFOLGERABBILDUNG

Es sei $(\mathbb{N}, \alpha, s)$ eine Realisierung der natürlichen Zahlen. Dann gilt $s(\mathbb{N}) = \mathbb{N} \setminus \{\alpha\}$.

Beweis. Wir müssen zeigen, dass $s(\mathbb{N}) \subseteq \mathbb{N} \setminus \{\alpha\}$ und $s(\mathbb{N}) \supseteq \mathbb{N} \setminus \{\alpha\}$. Zuerst zu „ $\subseteq$ “.

$$(P1) : \alpha \neq s(n) \quad \forall n \in \mathbb{N} \quad \xRightarrow{\text{Def 1.3.6(i)}} \alpha \notin s(\mathbb{N}) \quad \xRightarrow{\text{Def 1.2.7 (iii)}} s(\mathbb{N}) \subseteq \mathbb{N} \setminus \{\alpha\}.$$

Nun zu „ $\supseteq$ “. Es genügt zu zeigen, dass $s(\mathbb{N}) \cup \{\alpha\} = \mathbb{N}$ (denn dann gilt $\mathbb{N} \setminus \{\alpha\} = (s(\mathbb{N}) \cup \{\alpha\}) \setminus \{\alpha\} \subseteq s(\mathbb{N})$ ^a). Dazu sei $X := s(\mathbb{N}) \cup \{\alpha\}$. Beachte, dass $X \subseteq \mathbb{N}$. Wir werden mithilfe von (P3) zeigen, dass $X = \mathbb{N}$. Dazu müssen wir die Eigenschaften (a) und (b) nachprüfen.

Zu (a).

$$\alpha \in X \quad \checkmark.$$

Zu (b).

$$\begin{aligned} & n \in X \\ \xRightarrow{X \subseteq \mathbb{N}} & n \in \mathbb{N} \\ \xRightarrow{\text{Def 1.3.6(i)}} & s(n) \in s(\mathbb{N}) \\ \xRightarrow{s(\mathbb{N}) \subseteq X} & s(n) \in X. \end{aligned}$$

Mit dem Peano-Axiom (P3) folgt $X = \mathbb{N}$ und somit die Behauptung.

QED

^awobei man sich die hier behauptete Teilmengenrelation noch überlegen sollte. Wir überlassen den Lesenden die Details als Übungsaufgabe

Vollständige Induktion Oftmals möchte man für alle $n \in \mathbb{N}$ eine (von n abhängige) Aussage $A(n)$ beweisen. Das wird für uns Menschen ein schwieriges Unterfangen: Wir müssen unendlich viele Aussagen beweisen, aber unsere Lebenszeit ist endlich.

Eine noch elementarere Herausforderung: Man möchte (für eine Menge M) eine Abbildung $f : \mathbb{N} \rightarrow M$ definieren. Würden wir das mit einer Wertetabelle machen wollen, wären wir *unendlich lang beschäftigt*. Das wollen wir nicht!

Um diese Herausforderungen zu meistern, lernen wir das *Prinzip der vollständigen Induktion* kennen. Ein kleines Beispiel zum Verständnis:

Stellen Sie sich den *Domino Day* mit unendlich vielen Domino-Steinen vor. Würden Sie jeden der Steine einzeln umwerfen wollen, so wären Sie unendlich lange beschäftigt. Beim Domino Day wird aber nicht jeder Stein einzeln umgeworfen: Man wirft nur den ersten Stein um und sorgt zusätzlich dafür, dass jeder Stein seinen *Nachfolger* umwirft. Wenn das gelingt, fallen auch alle Steine um!

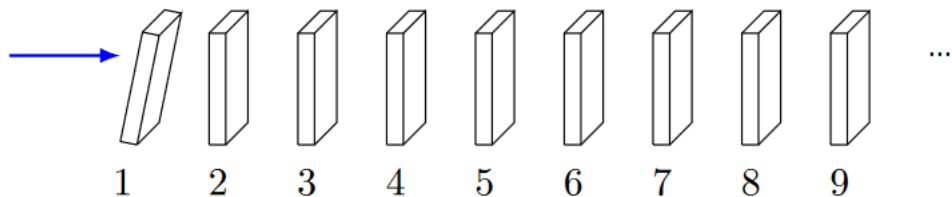


Abbildung 13: Domino Day

Proposition 3.3.3 — PRINZIP DER VOLLSTÄNDIGEN INDUKTION

Es seien für alle $n \in \mathbb{N}$ (von n abhängige) Aussagen $A(n)$ gegeben. Dann

$$A(\alpha) \quad \wedge \quad \left[\forall n \in \mathbb{N} : A(n) \implies A(s(n)) \right] \quad \implies \quad (\forall m \in \mathbb{N} : A(m)).$$

Mit anderen Worten: Wenn $A(\alpha)$ gilt und $A(n)$ stets $A(s(n))$ impliziert, dann gilt $A(m)$ für alle $m \in \mathbb{N}$.

Beweis. Es gelte $A(\alpha) \wedge (\forall n \in \mathbb{N} : A(n) \implies A(s(n)))$. Definiere

$$X := \{m \in \mathbb{N} : A(m) \text{ ist wahr}\} \quad (142)$$

Wir werden mithilfe von (P3) in Definition 3.3.1 beweisen, dass $X = \mathbb{N}$. Dazu sind die Bedingungen (a) und (b) aus (P3) nachzuprüfen.

Zu (a). Da $A(\alpha)$ nach Voraussetzung wahr ist, ist $\alpha \in X$.

Zu (b).

$$\begin{array}{ccc} & n \in X & \\ \xRightarrow{(142)} & A(n) \text{ ist wahr} & \\ \xRightarrow{A(n) \implies A(s(n))} & A(s(n)) \text{ ist wahr} & \\ \xRightarrow{(142)} & s(n) \in X. & \end{array}$$

Mit (P3) folgt $X = \mathbb{N}$. Das bedeutet nach (142), dass $A(m)$ wahr ist für alle $m \in \mathbb{N}$.

Bemerkung 3.3.4 — INDUKTIONSBEWEISE

Aus der vorigen Proposition lässt sich ein Beweisprinzip ableiten, der *Beweis durch vollständige Induktion*. Will man $A(m)$ für alle $m \in \mathbb{N}$ beweisen, so kann man wie folgt vorgehen

- (i) *Induktionsanfang*. Man zeigt $A(\alpha)$.
- (ii) *Induktionsschritt*. Man zeigt $A(n) \implies A(s(n))$ für beliebiges $n \in \mathbb{N}$.

Die Aussage $A(n)$ heißt innerhalb des Induktionsschrittes *Induktionshypothese*. Man darf sie dort als wahr annehmen und muss $A(s(n))$ aus ihr folgern.

Bemerkung 3.3.5 — REKURSIVE DEFINITION

Eine analoge Vorgehensweise erlaubt es, Funktionen $f : \mathbb{N} \rightarrow M$ (für eine beliebige Menge M) ohne Angabe einer Wertetabelle zu definieren.^a Nämlich wie folgt:

- (i) Man definiert $f(\alpha)$,
- (ii) Man erklärt, wie $f(s(n))$ aus $f(n)$ berechnet werden kann.

Eine solche Funktion $f : \mathbb{N} \rightarrow M$ nennt man dann *rekursiv definiert*.

“Solche Funktionen hatten wir in Kapitel 2 Folgen genannt, vgl. Definition 2.1.1!

Das Prinzip der vollständigen Induktion werden wir im folgenden noch üben. Zunächst müssen wir aber dafür lernen, mit natürlichen Zahlen zu *rechnen*. Um Rechenoperationen auf $\mathbb{N}$ zur Verfügung zu haben, müssen wir $\mathbb{N}$ als Teilmenge eines (angeordneten) Körpers verstehen können. Wie das geht klären wir im nächsten Abschnitt.

Natürliche Zahlen in angeordneten Körpern Wie vorhin angekündigt soll es hier darum gehen, die natürlichen Zahlen in einem angeordneten Körper wiederzufinden.

Proposition 3.3.6 — NATÜRLICHE ZAHLEN IN ANGEORDNETEN KÖRPERN

Es sei $(K, +, \cdot, \leq)$, $K \neq \{0\}$ ein angeordneter Körper. Dann existiert eine eindeutige Teilmenge $\mathbb{N} \subseteq K$ so, dass für

- $\alpha := 1$
- $s : \mathbb{N} \rightarrow K$ definiert durch $s(n) := n + 1$.

das Tripel $(\mathbb{N}, \alpha, s)$ eine Realisierung der natürlichen Zahlen ist.

Beweis. Wir müssen hier zwei Dinge zeigen: Erstens die *Existenz* einer solchen Menge $\mathbb{N}$ und zweitens die *Eindeutigkeit*, d.h., dass es keine weitere Menge gibt, die die o.g. Bedingungen erfüllt.

Schritt 1. Existenz. Es sei $\mathcal{G}$ die Menge aller Mengen $N \subseteq K$ mit den Eigenschaften

$$(a') \quad 1 \in N$$

$$(b') \quad m \in N \quad \Rightarrow \quad m + 1 \in N.$$

(Kurz würde man schreiben: $\mathcal{G} := \{N \subseteq K : N \text{ erfüllt (a')} \text{ und (b')}\}$). Definiere dann

$$\mathbb{N} := \{x \in K \mid \forall N \in \mathcal{G} : x \in N\}. \quad (143)$$

Definiere auch $\alpha := 1$ und die Abbildung $s : \mathbb{N} \rightarrow K$, $s(n) := n + 1$.

Behauptung. $(\mathbb{N}, \alpha, s)$ ist eine Realisierung der natürlichen Zahlen.

Nach Definition 3.3.1 muss dann Folgendes gezeigt werden:

$$\mathbb{Z}: \quad \alpha \in \mathbb{N}, s : \mathbb{N} \rightarrow \mathbb{N} \text{ (d.h. } s(\mathbb{N}) \subseteq \mathbb{N}) \text{ und } s \text{ erfüllt (P1),(P2),(P3).}$$

Dies arbeiten wir nacheinander ab.

Zu „ $\alpha \in \mathbb{N}$ “. Es gilt $1 \in \mathbb{N}$, da $1 \in N$ für alle $N \in \mathcal{G}$ (vgl. Bedingung (a')).

Zu „ $s : \mathbb{N} \rightarrow \mathbb{N}$ “ (d.h. insbesondere $s(\mathbb{N}) \subseteq \mathbb{N}$). Die Abbildung s erfüllt, dass $s : \mathbb{N} \rightarrow \mathbb{N}$, denn

$$\begin{aligned}
 & n \in \mathbb{N} \\
 \xRightarrow{(143)} & \forall N \in \mathcal{G} : \quad n \in N \\
 \xRightarrow{\text{Bed(b')}} & \forall N \in \mathcal{G} : \quad n + 1 \in N \\
 \xRightarrow{} & \forall N \in \mathcal{G} : \quad s(n) \in N \\
 \xRightarrow{(143)} & s(n) \in \mathbb{N}.
 \end{aligned}$$

Nun zu (P1).

$$\mathbb{Z} : \quad s(n) \neq 1 \text{ für alle } n \in \mathbb{N}$$

Angenommen es gibt ein $n \in \mathbb{N}$ mit $s(n) = 1$. Es gilt dann

$$s(n) = 1 \quad \Longrightarrow \quad n + 1 = 1 \quad \Longrightarrow \quad n = 0.$$

Wir folgern $0 \in \mathbb{N}$. Mit (143) schließen wir

$$0 \in \mathbb{N}. \quad \text{Das bedeutet} \quad \forall N \in \mathcal{G} : \quad 0 \in N. \quad (144)$$

Betrachte nun die Menge

$$N^* := \{x \in K : x \geq 1\}.$$

Dann gilt $0 \notin N^*$ da $1 > 0$ (vgl. Proposition 3.2.20 (xi)). Wir behaupten nun, dass $N^* \in \mathcal{G}$. Dazu: Bedingung (a') ist erfüllt, denn $1 \geq 1$ (und somit $1 \in N^*$). Bedingung (b') ist auch erfüllt, denn

$$\begin{aligned}
 m \in N^* & \Longrightarrow m \geq 1 \Longrightarrow m + 1 \geq \underbrace{1 + 1}_{\geq 1 \text{ (Bsp 3.2.21)}} \xRightarrow{(O3)} m + 1 \geq 1 \\
 & \Longrightarrow m + 1 \in N^*.
 \end{aligned}$$

Wir haben gezeigt, dass $N^* \in \mathcal{G}$ und $0 \notin N^*$. Dies widerspricht (144). Ein Widerspruch zur Annahme. Wir folgern daraus (P1).

Zu (P2). Seien $n, m \in \mathbb{N}$. Dann gilt

$$s(n) = s(m)$$

$$\begin{aligned}
&\implies n + 1 = m + 1 \\
&\implies (n + 1) - 1 = (m + 1) - 1 \\
&\implies n = m.
\end{aligned}$$

Zu (P3). Es sei $X \subseteq \mathbb{N}$ mit

- $\alpha \in X$ und
- $n \in X \implies s(n) \in X$.

$$\mathbb{Z}: \quad X = \mathbb{N}.$$

Dazu: Es genügt zu zeigen, dass $X \supseteq \mathbb{N}$ (denn „ $\subseteq$ “ ist nach Definition wahr). Wir zeigen zuerst $X \in \mathcal{G}$. In der Tat:

(a') $1 \in X$ da $\alpha = 1$ und $\alpha \in X$.

(b') Sei $m \in X$. $\mathbb{Z}: m + 1 \in X$. Dazu: Da $s(m) \in X$ und $s(m) = m + 1$ folgt $m + 1 \in X$.

Nun zu $X \supseteq \mathbb{N}$.

$$x \in \mathbb{N} \implies x \in N \quad \forall N \in \mathcal{G} \implies_{X \in \mathcal{G}} x \in X.$$

Es folgt wie oben erklärt $X = \mathbb{N}$ und somit (P3). Die Behauptung ist gezeigt.

Schritt 2. Eindeutigkeit. Seien $\mathbb{N}_1$ und $\mathbb{N}_2$ Teilmengen von K so, dass für $\alpha = 1$ und $s_1 : \mathbb{N}_1 \rightarrow K, s_1(n) := n + 1$ bzw. $s_2 : \mathbb{N}_2 \rightarrow K, s_2(n) := n + 1$ definiert wie in der Aussage die Tripel $(\mathbb{N}_1, \alpha, s_1)$ und $(\mathbb{N}_2, \alpha, s_2)$ Realisierungen der natürlichen Zahlen sind.

$$\mathbb{Z}: \quad \mathbb{N}_1 = \mathbb{N}_2.$$

Dazu: Aufgrund der Tatsache, dass beide der obigen Tripel Realisierungen der natürlichen Zahlen sind, bildet s_1 Elemente von $\mathbb{N}_1$ in $\mathbb{N}_1$ ab und s_2 Elemente von $\mathbb{N}_2$ in $\mathbb{N}_2$ ab. Definiere nun $X := \mathbb{N}_1 \cap \mathbb{N}_2$. (**)

Beachte

- $\alpha = 1 \in X$, da $\alpha \in \mathbb{N}_1$ und $\alpha \in \mathbb{N}_2$.
- Sei $n \in X$. Wir zeigen $s_1(n) \in X$. Dass $s_1(n) \in \mathbb{N}_1$ liegt daran, dass s_1 von $\mathbb{N}_1$ nach $\mathbb{N}_1$ abbildet. Dass $s_1(n) \in \mathbb{N}_2$ folgt so:

$$n \in X \implies (n \in \mathbb{N}_1) \wedge (n \in \mathbb{N}_2) \implies s_1(n) = n + 1 = s_2(n) \in \mathbb{N}_2.$$

Wir haben gezeigt, dass $s_1(n) \in \mathbb{N}_1 \cap \mathbb{N}_2 = X$. Benutzen wir dass für $\mathbb{N}_1$ das Peano-Axiom (P3) gilt, so folgern wir aus den oberen Bullet Points, dass $X = \mathbb{N}_1$. Schreiben wir den Beweis ab (**) nochmal mit vertauschten Rollen von $\mathbb{N}_1$ und $\mathbb{N}_2$ ab, so können wir analog zeigen, dass $X = \mathbb{N}_2$. Es ergibt sich $\mathbb{N}_1 = X = \mathbb{N}_2$. Da $X = \mathbb{N}_1 \cap \mathbb{N}_2$ ist dies nur möglich wenn $\mathbb{N}_1 = \mathbb{N}_2$. Also folgt die Behauptung.

Sobald wir uns in einem angeordneten Körper $(K, +, \cdot, \leq)$, $K \neq \{0\}$ befinden, bezeichnen wir mit $\mathbb{N}$ einfach die (eindeutig bestimmte) Realisierung der natürlichen Zahlen aus der obigen Proposition.

In der Mathematik ist es übrigens stets kontrovers, ob das Alpha-Element der natürlichen Zahlen bei 1 oder 0 liegen soll. In der Logik macht es keinen Unterschied: Genau wie in der obigen Proposition könnte man die Existenz einer eindeutig bestimmten Realisierung der natürlichen Zahlen mit Alpha-Element 0 zeigen. Es ist alles nur eine Frage der Konvention. Wir können jetzt wie gewohnt die Zahlen $2 := 1 + 1, 3 := 2 + 1, 4 := 3 + 1, 5 := 4 + 1$ usw. definieren. Da jene natürlichen Zahlen nun Elemente aus einem angeordneten Körper $(K, +, \cdot, \leq)$ sind, ist es uns auch möglich, mit ihnen zu *rechnen*.

Bemerkung 3.3.7 — INDUKTIONSBEWEISE II

Im Folgenden werden wir auf $\mathbb{N}$ (wie in der vorigen Proposition) *Beweise durch vollständige Induktion* durchführen. Für unsere konkret gegebene Realisierung der natürlichen Zahlen geht das wie folgt (vgl. Bemerkung 3.3.4): Will man $A(m)$ für alle $m \in \mathbb{N}$ beweisen, so kann man wie folgt vorgehen

- (i) *Induktionsanfang*. Man zeigt $A(1)$.
- (ii) *Induktionsschritt*. Man zeigt $A(n) \implies A(n + 1)$ für beliebiges $n \in \mathbb{N}$.

Den Induktionsanfang versteht man meistens mit der Überschrift „IndA: $n = 1$ “ und den Induktionsschritt mit dem „IndS: $n \longrightarrow n + 1$ “. Jeder Induktionsschritt startet bei uns mit der Formulierung der *Induktionshypothese*, genauer mit den Worten: „Es gelte für ein $n \in \mathbb{N}$ die Induktionshypothese $A(n)$.“ Meist schreiben wir auch noch „ $\mathbb{Z}$: Dann gilt auch $A(n + 1)$ “ dazu. Das Wort „Induktionshypothese“ darf auch mit IndH abgekürzt werden.

Dasselbe Prinzip können wir bei der Definition von Abbildungen verwenden, vergleiche Bemerkung 3.3.5. Ein Beispiel dafür ist die Definition der *Potenz*.

Definition 3.3.8 — POTENZEN

Es sei $(K, +, \cdot, \leq)$ ein angeordneter Körper und $x \in K$. Wir möchten für $n \in \mathbb{N}$ die Größe x^n definieren. Dazu setzen wir

- (i) $x^1 := x$
- (ii) x^{n+1} berechnet sich aus x^n durch die Formel $x^{n+1} := x \cdot x^n$.

Nach Bemerkung 3.3.5 ist damit x^n für alle $n \in \mathbb{N}$ vollends definiert. Wir setzen außerdem auch $x^0 := 1$ für alle $x \in K$. Ferner definieren wir für $x \in K \setminus \{0\}$ und $n \in \mathbb{N}$ auch

$$x^{-n} := (x^{-1})^n.$$

Eine erste Induktionsübung ist die folgende

Proposition 3.3.9 — RECHENREGELN FÜR POTENZEN

Es sei $(K, +, \cdot, \leq), K \neq \{0\}$ ein angeordneter Körper und $x, y \in K$ sowie $n, m \in \mathbb{N}$. Dann gilt

- (i) $(xy)^n = x^n y^n$.
- (ii) $x^{m+n} = x^m x^n$.
- (iii) Falls $y \in K \setminus \{0\}$ so gilt $(\frac{x}{y})^n = \frac{x^n}{y^n}$.
- (iv) Falls $x \in K \setminus \{0\}$ so gilt $x^{m-n} = \frac{x^m}{x^n}$ sowie $x^{-m-n} = x^{-m} x^{-n}$.
- (v) $x^{mn} = (x^m)^n$.

Beweis. Zu (i). Wir zeigen die Behauptung durch vollständige Induktion. Die Aussage $A(n)$ ist hierbei gegeben durch $(xy)^n = x^n y^n$.

IndA: $n = 1$.

Es gilt

$$(xy)^1 \stackrel{\text{Def3.3.8}}{=} xy$$

und

$$x^1 y^1 \stackrel{\text{Def3.3.8 x2}}{=} xy.$$

IndS: $n \longrightarrow n + 1$.

Es gelte für ein $n \in \mathbb{N}$ die Induktionshypothese

$$(xy)^n = x^n y^n. \quad (\text{IndH})$$

$$\mathbb{Z}: \quad \text{Dann gilt auch} \quad (xy)^{n+1} = x^{n+1} y^{n+1}.$$

Dazu:

$$(xy)^{n+1} \stackrel{\text{Def3.3.8}}{=} (xy)(xy)^n \stackrel{\text{IndH}}{=} (xy)(x^n y^n) = (xx^n)(yy^n) \stackrel{\text{Def3.3.8 x2}}{=} x^{n+1} y^{n+1}.$$

Die Behauptung ist (durch vollständige Induktion nach n) gezeigt.

QED

Zu (ii). Fixiere $m \in \mathbb{N}$ und $x \in K$. Sei dann für $n \in \mathbb{N}$ die Aussage $A(n)$ gegeben durch $x^{m+n} = x^m x^n$. Wir beweisen $A(n) \forall n \in \mathbb{N}$ durch vollständige Induktion. Dazu:

IndA: $n = 1$.

Es gilt

$$x^{m+1} \stackrel{\text{Def3.3.8(ii)}}{=} xx^m \stackrel{\text{Def3.3.8(i)}}{=} x^1 x^m.$$

Somit ist $A(1)$ gezeigt.

IndS: $n \longrightarrow n + 1$.

Es gelte für ein $n \in \mathbb{N}$ die Induktionshypothese

$$x^{m+n} = x^m x^n. \quad (\text{IndH})$$

$$\mathbb{Z}\mathbb{Z}: \quad \text{Dann gilt auch } x^{m+(n+1)} = x^m x^{n+1}.$$

Dazu:

$$x^{m+(n+1)} = x^{(m+n)+1} \stackrel{\text{Def3.3.8(ii)}}{=} xx^{m+n} \stackrel{\text{IndH}}{=} x(x^m x^n) = x^m (xx^n) \stackrel{\text{Def3.3.8(ii)}}{=} x^m x^{n+1}.$$

Die Behauptung ist mit vollständiger Induktion gezeigt.

QED

Zu (iii). Könnte man auch wieder mit vollständiger Induktion zeigen. Eine einfachere Methode ist aber die Folgende. Wir berechnen erst:

$$y^n \left(\frac{x}{y}\right)^n \stackrel{(i)}{=} \left(y \cdot \frac{x}{y}\right)^n = x^n.$$

Teilen wir durch y^n (was ungleich 0 ist, wie man sich auch leicht mit einer Induktion basierend auf Proposition 3.1.9 (xi) überlegen kann) so folgt

$$\left(\frac{x}{y}\right)^n = \frac{x^n}{y^n}.$$

Zu (iv), (v). Überlassen wir den Lesenden als Übungsaufgabe.

24.11.2025 Eine zweite (auch nicht allzu schwere) Induktionsaufgabe ist

Proposition 3.3.10 — POSITIVITÄT DER NATÜRLICHEN ZAHLEN

Es sei $(K, +, \cdot, \leq), K \neq \{0\}$ ein angeordneter Körper. Ferner sei $\mathbb{N}$ wie in Proposition 3.3.6^a. Dann gilt für alle $n \in \mathbb{N}$, dass $n \geq 1$.

^awie angekündigt werden wir das ab sofort nicht mehr dazuschreiben

Beweis. Wir beweisen die Behauptung durch vollständige Induktion. Die Aussage $A(n)$ ist hierbei gegeben durch $A(n) : n \geq 1$.

IndA: $n = 1$.

Es gilt klarerweise $1 \geq 1$.

IndS: $n \longrightarrow n + 1$.

Es gelte für ein $n \in \mathbb{N}$ die Induktionshypothese $n \geq 1$. (IndH).

$\mathbb{Z}$: Dann gilt auch $n + 1 \geq 1$.

Dazu:

$$n + 1 \underset{1 > 0}{\geq} n + 0 \underset{(A3)}{=} n \underset{\text{IndH}}{\geq} 1.$$

Die Behauptung folgt.

Aus der obigen Proposition (und $1 > 0$) folgt auch $n > 0$ für alle $n \in \mathbb{N}$.

Eine etwas kompliziertere Induktion ist die *echte* Bernoulli-Ungleichung, d.h. der allgemeine Fall von Proposition 2.2.20.

Proposition 3.3.11 — BERNOULLI'SCHE UNGLEICHUNG

Es sei $(K, +, \cdot, \leq), K \neq \{0\}$ ein angeordneter Körper. Sei ferner $x \in K$ mit $x \geq -1$. Dann gilt

$$\forall n \in \mathbb{N} : \quad (1 + x)^n \geq 1 + n \cdot x.$$

Beweis. Wir beweisen die Behauptung durch vollständige Induktion.

IndA: $n = 1$. ($\mathbb{Z} : (1 + x)^1 \geq 1 + 1 \cdot x$).

$$(1 + x)^1 = 1 + x = 1 + 1 \cdot x.$$

IndS: $n \longrightarrow n + 1$.

Es gelte für ein $n \in \mathbb{N}$ die Induktionshypothese

$$(1 + x)^n \geq 1 + n \cdot x \quad (\text{IndH})$$

$\mathbb{Z}$: Dann gilt auch $(1 + x)^{n+1} \geq 1 + (n + 1) \cdot x$.

Dazu:

$$(1 + x)^{n+1} = (1 + x) \cdot (1 + x)^n$$

$$\begin{aligned}
& \underset{\text{IndH}^a}{\geq} (1+x) \cdot (1+n \cdot x) \\
& = (1+x) + (1+x) \cdot n \cdot x \\
& \underset{(D)}{=} 1+x+n \cdot x+n \cdot x \cdot x \\
& \underset{(D)}{=} 1+(n+1)x+n \cdot \underbrace{x \cdot x}_{\geq 0(\text{Prop 3.2.20(x)})} \\
& \underset{n>0}{\geq} 1+(n+1)x+n \cdot 0 = 1+(n+1)x.
\end{aligned}$$

Die Behauptung ist mit vollständiger Induktion gezeigt.

QED

^aHier benutzen wir außerdem, dass $1+x \geq 0$ und Proposition 3.2.20 (viii).

^aHier benutzen wir außerdem, dass $1+x \geq 0$ und Proposition 3.2.20 (viii).

Wir üben noch einen weiteren Induktionsbeweis.

Proposition 3.3.12

Für alle $n \in \mathbb{N}$ ist $n(n+1)$ gerade, d.h. $\frac{n(n+1)}{2} \in \mathbb{N}$.

***Beweis.** Wir beweisen die Behauptung durch vollständige Induktion.

IndA: $n = 1$. ($\mathbb{Z} : 1 \cdot (1+1)$ ist gerade.)

Dazu:

$$1 \cdot (1+1) = 1 \cdot 2 = 2.$$

Nun gilt

$$\frac{2}{2} = 2 \cdot 2^{-1} = 1 \in \mathbb{N}.$$

IndS: $n \longrightarrow n+1$.

Es gelte für ein $n \in \mathbb{N}$ die Induktionshypothese (IndH: $n(n+1)$ ist gerade).

$\mathbb{Z} :$ Dann gilt auch $(n+1)((n+1)+1)$ ist gerade.

Dazu:

$$(n+1)((n+1)+1) = (n+1)(n+2) \underset{(D)}{=} (n+1)n + (n+1) \cdot 2 = n(n+1) + 2(n+1). \quad (145)$$

Nach der Induktionshypothese gilt $\frac{n(n+1)}{2} \in \mathbb{N}$, nennen wir etwa $k := \frac{n(n+1)}{2}$. Mit (145) berechnen wir

$$\frac{(n+1)((n+1)+1)}{2} = \frac{n(n+1) + 2(n+1)}{2}$$

$$\begin{aligned}
&= \frac{n(n+1)}{2} + \frac{2(n+1)}{2} \\
&= \frac{n(n+1)}{2} + (n+1) = k + n + 1.
\end{aligned} \tag{146}$$

Da $k, n \in \mathbb{N}$ folgt $k + n + 1 \in \mathbb{N}$.^a Dies und (146) zeigt die Behauptung.

^aAn dieser Stelle müsste man eigentlich noch zeigen, dass die Summe zweier natürlicher Zahlen wieder eine natürliche Zahl ist. Die Lesenden dürfen dies gerne mit einer Induktion als Übungsaufgabe probieren.

Ganze Zahlen Wer natürliche Zahlen definiert hat, kann auch *ganze Zahlen* definieren.

Definition 3.3.13 — GANZE ZAHLEN

Sei $(K, +, \leq)$ ein angeordneter Körper. Wir definieren die *ganzen Zahlen (in K)* durch

$$\mathbb{Z} := \mathbb{N} \cup \{0\} \cup \{-n : n \in \mathbb{N}\}.$$

Beachten Sie: Da $n > 0$ für alle $n \in \mathbb{N}$ und $-n < 0$ für alle $n \in \mathbb{N}$ gilt

$$n \in \mathbb{N} \iff n \in \mathbb{Z} \wedge n > 0. \tag{147}$$

Differenzen zweier natürlicher Zahlen sind stets Elemente der ganzen Zahlen.

Proposition 3.3.14

Es seien $m, n \in \mathbb{N}$. Dann gilt $m - n \in \mathbb{Z}$. Ferner gilt für $z_1, z_2 \in \mathbb{Z}$, dass auch $z_1 z_2 \in \mathbb{Z}$.

***Beweis.** Induktion, die wir den Lesenden als Übungsaufgabe überlassen. (Der erste Teil ist übrigens gar nicht so leicht und verwendet im Induktionsanfang Proposition 3.3.2).

Ganze Zahlen spielen in so einigen Beweisen eine Rolle, denn es kommt öfters mal vor, dass man natürliche Zahlen voneinander subtrahieren muss. So zum Beispiel in dieser Folgerung aus Proposition 3.3.12.

Proposition 3.3.15

Es sei $m \in \mathbb{N}$. Dann gilt^a

$$m^2 \text{ gerade} \iff m \text{ gerade}.$$

^aZur Erinnerung aus Proposition 3.3.12: Eine Zahl $j \in \mathbb{N}$ heißt *gerade* falls $\frac{j}{2} \in \mathbb{N}$.

***Beweis.** *Vorüberlegung.* Da nach Proposition 3.3.12 $m(m+1)$ gerade ist, gilt

$$\mathbb{N} \ni \frac{m(m+1)}{2} \stackrel{(D)}{=} \frac{m^2 + m}{2} = \frac{m^2}{2} + \frac{m}{2}.$$

Sei etwa

$$k := \frac{m^2}{2} + \frac{m}{2} \quad [\in \mathbb{N}]. \quad (148)$$

Nun zu „ $\implies$ “. Falls m^2 gerade ist, so ist $\frac{m^2}{2} \in \mathbb{N}$. Dann ist nach (148)

$$\frac{m}{2} = k - \frac{m^2}{2} \stackrel{\text{Prop 3.3.14}}{\in} \mathbb{Z}.$$

Da aber $\frac{m}{2} > 0$ (vgl. Proposition 3.3.10) schließen wir nach (147), dass $\frac{m}{2} \in \mathbb{N}$. Es folgt, dass m gerade ist.

Nun zu „ $\impliedby$ “. Falls m gerade ist, so ist $\frac{m}{2} \in \mathbb{N}$. Dann ist nach (148)

$$\frac{m^2}{2} = k - \frac{m}{2} \stackrel{\text{Prop 3.3.14}}{\in} \mathbb{Z}.$$

Da aber $\frac{m^2}{2} > 0$ schließen wir nach (147), dass $\frac{m}{2} \in \mathbb{N}$. Es folgt, dass m gerade ist.

Summen -und Produktnotation Im folgenden Abschnitt sei stets $(K, +, \cdot, \leq), K \neq \{0\}$ ein angeordneter Körper.

In diesem Abschnitt führen wir das *Summenzeichen* ein. Intuitiv soll für $a_1, \dots, a_n \in K$ das Summenzeichen den folgenden Ausdruck auswerten:

$$\sum_{k=1}^n a_k = a_1 + \dots + a_n. \quad (149)$$

Um diesen Vorgang aber richtig zu definieren, verwenden wir die Idee der vollständigen Induktion, ähnlich wie bei Definition 3.3.8.

Definition 3.3.16 — SUMMENZEICHEN

Es seien für $k \in \mathbb{Z}$ Elemente $a_k \in K$ gegeben. Für $n \in \mathbb{N}$ definieren wir den Ausdruck

$$\sum_{k=1}^n a_k \in K$$

rekursiv durch die folgenden Eigenschaften:

(i)

$$\sum_{k=1}^1 a_k := a_1.$$

(ii)

$$\sum_{k=1}^{n+1} a_k := \sum_{k=1}^n a_k + a_{n+1}.$$

Ferner definieren wir für $m, n \in \mathbb{Z}$ mit $n \geq m$ den Ausdruck

$$\sum_{k=m}^n a_k := \sum_{k=1}^{n-m+1} a_{k+m-1}. \quad (150)$$

Eine platzsparende Notation, die wir manchmal auch im Skript verwenden werden, ist $\sum_{k=m}^n a_k$.

Proposition 3.3.17 — RECHNEN MIT SUMMEN

Es sei für alle $k \in \mathbb{Z}$ Elemente $a_k, b_k \in K$ gegeben und $c \in K$. Ferner seien $m, n \in \mathbb{Z}$ so, dass $m \leq n$. Dann gilt

(i) *Additivität.*

$$\sum_{k=m}^n (a_k + b_k) = \sum_{k=m}^n a_k + \sum_{k=m}^n b_k.$$

(ii) *erweitertes Distributivgesetz.*

$$\sum_{k=m}^n (c \cdot a_k) = c \sum_{k=m}^n a_k.$$

(iii) *Indexverschiebung.*

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

(iv) *Rückwärtssummation.*

$$\sum_{k=m}^n a_k = \sum_{k=m}^n a_{m+n-k}.$$

(v) *Teleskopsumme.*

$$\sum_{k=m}^n (a_k - a_{k+1}) = a_m - a_{n+1}.$$

(vi) *Zählsumme.*

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

(vii) *Aufteilen einer Summe.* Für alle $j \in \mathbb{Z}$ mit $m \leq j < n$ gilt

$$\sum_{k=m}^n a_k = \sum_{k=m}^j a_k + \sum_{k=j+1}^n a_k.$$

Beweisskizze. Dank (153) kann man die Behauptungen stets auf den Fall $m = 1$ zurückführen. Sobald das gemacht ist, sind die Aussagen dankbar durch vollständige Induktion beweisbar. Wir führen die Details nicht vor und überlassen sie den Lesenden als Übungsaufgabe.

Bemerkung 3.3.18 — ABSPALTUNG EINZELNER SUMMANDEN

Nutzt man Proposition 3.3.17 (vii) mit $j = m$ so sieht man

$$\sum_{k=m}^n a_k = \sum_{k=m}^m a_k + \sum_{k=m+1}^n a_k \stackrel{\text{Def 3.3.16(i)}}{=} a_m + \sum_{k=m+1}^n a_k.$$

Somit ist es insbesondere möglich, einzelne Summanden am Anfang der Summe *abzuspalten*. Am Ende der Summe ist das Abspalten eines Summanden nach Definition 3.3.16 (ii) sowieso immer möglich.

Diese Regeln kann man verwenden um zwei sehr bekannte Summenformeln zu zeigen: Die *Gauß'sche Summenformel* und die *Geometrische Summenformel*.

Proposition 3.3.19 — GAUSS'SCHE SUMMENFORMEL

Es gilt für alle $n \in \mathbb{N}$

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

Beweis. Wir verwenden einige der Identitäten der obigen Proposition mit $a_k = k$.

$$\begin{aligned} 2 \left(\sum_{k=1}^n k \right) &\stackrel{\text{Prop 3.3.17(iv)}}{=} \sum_{k=1}^n k + \sum_{k=1}^n k \\ &\stackrel{\text{Prop 3.3.17(i)}}{=} \sum_{k=1}^n [k + (n+1-k)] \\ &\stackrel{\text{Prop 3.3.17(ii)}}{=} \sum_{k=1}^n (n+1) \\ &\stackrel{\text{Prop 3.3.17(ii)}}{=} (n+1) \sum_{k=1}^n 1 \\ &\stackrel{\text{Prop 3.3.17(ii)}}{=} (n+1)n. \end{aligned}$$

Multipliziert man die linke Seite und die rechte Seite der Gleichung mit $\frac{1}{2}$, so folgt die Behauptung.

Die obige Summenformel könnte man übrigens auch durch vollständige Induktion zeigen.

***Beispiel 3.3.2.** GAUSS'SCHE SUMMENFORMEL MIT INDUKTION. *Behauptung.* Es gilt für alle $n \in \mathbb{N}$

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

Beweis. Vollständige Induktion.

IndA: $n = 1$. ($\mathbb{Z}$: $\sum_{k=1}^1 k = \frac{1(1+1)}{2}$)

Dazu berechne

$$\sum_{k=1}^1 k = 1$$

und

$$\frac{1(1+1)}{2} = \frac{1 \cdot 2}{2} = \frac{1}{2} \cdot 2 = 1. \quad \checkmark$$

IndS: $n \longrightarrow n+1$.

Es gelte für ein $n \in \mathbb{N}$ die Induktionshypothese

$$\sum_{k=1}^n k = \frac{n(n+1)}{2}. \quad (\text{IndH})$$

$$\mathbb{Z} : \quad \text{Dann gilt auch} \quad \sum_{k=1}^{n+1} k = \frac{(n+1)((n+1)+1)}{2}.$$

Dazu:

$$\begin{aligned} \sum_{k=1}^{n+1} k &\stackrel{\text{Def 3.3.16}}{=} \sum_{k=1}^n k + (n+1) \stackrel{\text{IndH}}{=} \frac{n(n+1)}{2} + (n+1) \\ &\stackrel{(\text{D})}{=} (n+1) \left(\frac{n}{2} + 1 \right) = (n+1) \left(\frac{n+2}{2} \right) = \frac{(n+1)(n+2)}{2} = \frac{(n+1)((n+1)+1)}{2}. \end{aligned}$$

Die Behauptung ist gezeigt.

Beispiel 3.3.3. SUMMIERUNG ÜBER UNGERADE ZAHLEN. Gemäß (149) ist

$$\sum_{k=1}^n (2k-1) = 1 + 3 + 5 + \dots + (2n-1)$$

die Summe über die *ersten n ungeraden Zahlen*. Verwenden wir unsere Rechenregeln und die Gauß'sche Summenformel, so erhalten wir eine erstaunliche Formel, nämlich

$$\sum_{k=1}^n (2k-1) = n^2.$$

Beweis.

$$\begin{aligned} \sum_{k=1}^n (2k-1) &= \sum_{k=1}^n 2 \cdot k + (-1) \cdot 1 \\ &\stackrel{\text{Prop 3.3.17(i)}}{=} \sum_{k=1}^n 2 \cdot k + \sum_{k=1}^n (-1) \cdot 1 \\ &\stackrel{\text{Prop 3.3.17(ii)}}{=} 2 \sum_{k=1}^n k + (-1) \sum_{k=1}^n 1 \\ &\stackrel{\text{Prop 3.3.19}}{=} 2 \frac{n(n+1)}{2} - \sum_{k=1}^n 1 \\ &\stackrel{\text{Prop 3.3.17(vi)}}{=} n(n+1) - n = n^2 + n - n = n^2. \end{aligned}$$

Nun zur *geometrischen Summe*.

Proposition 3.3.20 — GEOMETRISCHE SUMMENFORMEL

Es sei $q \in K$ und $q \neq 1$. Dann gilt für alle $n \in \mathbb{N} \cup \{0\}$

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

Beweis. Wir berechnen

$$\begin{aligned} (1-q) \sum_{k=0}^n q^k &\stackrel{\text{Prop 3.3.17(ii)}}{=} \sum_{k=0}^n (1-q)q^k \\ &= \sum_{k=0}^n (q^k - q \cdot q^k) \end{aligned}$$

$$\begin{aligned} &= \sum_{k=0}^n (q^k - q^{k+1}) \\ &= q^0 - q^{n+1} = 1 - q^{n+1}. \end{aligned}$$

Prop3.3.17(v)

Die Behauptung folgt indem man beide Seiten mit $\frac{1}{1-q}$ multipliziert.

Eine wichtige Konsequenz ist die *Monotonie der Potenzfunktion*

Proposition 3.3.21 — MONOTONIE DER POTENZFUNKTION

(i) Seien $x, y \in K$ mit $x \neq y$. Dann gilt für alle $n \in \mathbb{N}$

$$\frac{y^n - x^n}{y - x} = \sum_{k=0}^{n-1} x^k y^{n-1-k}.$$

(ii) Für alle $x, y \in K$ mit $0 \leq x < y$ und alle $n \in \mathbb{N}$ gilt $x^n < y^n$.

Beweis. Zu (i)^a. Seien $x, y \in K$ mit $x \neq y$.

FALL 1. $y = 0$.

Dann gilt

$$\frac{y^n - x^n}{y - x} = \frac{-x^n}{-x} = \frac{(-1)x^n}{(-1)x} = \frac{x^n}{x} = x^{n-1}. \quad (151)$$

Beachte: Nach Definition 3.3.8 ist $y^0 = 1$ aber — wie man leicht mit Induktion zeigt — $y^m = 0$ für alle $m \in \mathbb{N}$. Daher

$$\sum_{k=0}^{n-1} x^k y^{n-1-k} \stackrel{\text{Bem3.3.18}}{=} \sum_{k=0}^{n-2} x^k y^{n-1-k} + x^{n-1} y^0 = \sum_{k=0}^{n-2} 0 + x^{n-1} y^0 = 0 + x^{n-1} = x^{n-1}. \quad (152)$$

Da (151) und (152) für alle $x \in K$ übereinstimmen ist die Behauptung in diesem Fall gezeigt.

FALL 2. $y \neq 0$.

Definiere $q := \frac{x}{y}$. Dann ist $q \neq 1$. Es gilt nun

$$\begin{aligned} \frac{y^n - x^n}{y - x} &= \frac{y^n(1 - \frac{x^n}{y^n})}{y(1 - \frac{x}{y})} \\ &= \frac{y^n}{y} \frac{1 - \frac{x^n}{y^n}}{1 - \frac{x}{y}} \end{aligned}$$

$$\begin{aligned}
& \xlongequal[\text{Prop3.3.9(iii)}]{y^n \frac{1 - (\frac{x}{y})^n}{1 - \frac{x}{y}}} \\
& \xlongequal[\text{mit } q = \frac{x}{y} \neq 1!]{\text{Prop3.3.20}} \frac{y^n}{y} \sum_{k=0}^{n-1} \left(\frac{x}{y}\right)^k \\
& \xlongequal[\text{Prop3.3.9(iv)}]{y^{n-1} \sum_{k=0}^{n-1} \left(\frac{x}{y}\right)^k} \\
& \xlongequal[\text{Prop3.3.9(iii),(iv)}]{y^{n-1} \sum_{k=0}^{n-1} x^k y^{-k}} \\
& \xlongequal[\text{Prop3.3.17(ii)}]{\sum_{k=0}^{n-1} x^k y^{n-1-k}} \\
& \xlongequal[\text{Prop3.3.9(iv)}]{\sum_{k=0}^{n-1} x^k y^{n-1-k}}.
\end{aligned}$$

Die Behauptung ist gezeigt.

Zu (ii). Es seien $x, y \in K$ mit $0 \leq x < y$.

FALL 1. $x \neq 0$.

Aus (i) lernen wir

$$y^n - x^n = \underbrace{(y - x)}_{>0} \sum_{k=0}^{n-1} \underbrace{x^k y^{n-1-k}}_{>0} > 0,$$

wobei wir hier einige Rechenregeln (z.B. Proposition 3.2.20 (ii)) benutzt haben. Wir folgern, dass $y^n - x^n > 0$. Addieren wir x^n auf beiden Seiten und verwenden Proposition 3.2.20 (i), so erhalten wir

$$y^n > x^n.$$

Die Behauptung folgt.

FALL 2. $x = 0$.

Dann ist $x^n = 0$. Mit einer Induktion, die auf Proposition 3.2.20(ii) basiert, zeigt man leicht, dass $y^m > 0$ für alle $m \in \mathbb{N}$. Es folgt, dass $x^n = 0 < y^n$. QED

^aEine Möglichkeit, diese Aussage zu beweisen ist, beide Seiten mit $x - y$ zu multiplizieren und dann vorzugehen wie in 3.3.20. Eine andere Möglichkeit präsentieren wir hier!

Im vorigen Beweis haben wir das Rechnen mit Potenzen (d.h. die Verwendung der Potenzgesetze aus Proposition 3.3.9) geübt. Im Folgenden werden wir die Potenzgesetze, die wir benutzen nicht, mehr explizit unter das Gleichheitszeichen schreiben.

Ähnlich wie das Summenzeichen können wir auch das *Produktzeichen* definieren.

Definition 3.3.22 — PRODUKTZEICHEN

Es seien für $k \in \mathbb{Z}$ Elemente $a_k \in K$ gegeben. Für $n \in \mathbb{N}$ definieren wir den Ausdruck

$$\prod_{k=1}^n a_k \in K$$

rekursiv durch die folgenden Eigenschaften:

(i)

$$\prod_{k=1}^1 a_k := a_1.$$

(ii)

$$\prod_{k=1}^{n+1} a_k := \left(\prod_{k=1}^n a_k \right) \cdot a_{n+1}.$$

Ferner definieren wir für $m, n \in \mathbb{Z}$ mit $n \geq m$ den Ausdruck

$$\prod_{k=m}^n a_k := \prod_{k=1}^{n-m+1} a_{k+m-1}. \quad (153)$$

Eine platzsparende Notation, die wir manchmal auch im Skript verwenden werden, ist $\prod_{k=m}^n a_k$.

Proposition 3.3.23 — RECHNEN MIT PRODUKTEN

Es sei für alle $k \in \mathbb{Z}$ Elemente $a_k, b_k \in K$ gegeben und $c \in K$. Ferner seien $m, n \in \mathbb{Z}$ so, dass $m \leq n$. Dann gilt

(i) *Multiplikativität.*

$$\prod_{k=m}^n (a_k \cdot b_k) = \left(\prod_{k=m}^n a_k \right) \cdot \left(\prod_{k=m}^n b_k \right).$$

(ii) *Vorziehen von Konstanten mit Potenz.*

$$\prod_{k=m}^n (c \cdot a_k) = c^{n-m+1} \prod_{k=m}^n a_k.$$

(iii) *Indexverschiebung.*

$$\prod_{k=m}^n a_k = \prod_{k=m+1}^{n+1} a_{k-1}.$$

(iv) *Rückwärtsmultiplikation.*

$$\prod_{k=m}^n a_k = \prod_{k=m}^n a_{m+n-k}.$$

(v) *Teleskopprodukt.*

$$\prod_{k=m}^n \frac{a_k}{a_{k+1}} = \frac{a_m}{a_{n+1}}.$$

(vi) *Einsprodukt.*

$$\prod_{k=m}^n 1 = 1.$$

(vii) *Aufteilen eines Produktes.* Für alle $j \in \mathbb{Z}$ mit $m \leq j < n$ gilt

$$\prod_{k=m}^n a_k = \left(\prod_{k=m}^j a_k \right) \cdot \left(\prod_{k=j+1}^n a_k \right).$$

Beweis. Führen wir nicht vor (vgl. Proposition 3.3.17.)

Fakultät und Binomialkoeffizient Im Folgenden definieren wir zwei spezielle (uns bereits aus Kapitel 2 bekannte) Produkte: die *Fakultät* und den *Binomialkoeffizient*. Wir haben diese im Kapitel 2 schonmal im Kontext des *binomischen Lehrsatzes* betrachtet. Diesen wollen wir hier nun rigoros mithilfe des Prinzips der vollständigen Induktion beweisen. Wie immer sei $(K, +, \cdot, \leq)$ ein angeordneter Körper und $\mathbb{N}$ die oben fixierte Realisierung der natürlichen Zahlen in K .

Definition 3.3.24 — FAKULTÄT

Für $n \in \mathbb{N}$ setzt man

$$n! := \prod_{k=1}^n k.$$

Außerdem ist $0! := 1$.

Definition 3.3.25 — BINOMIALKOEFFIZIENT

Für $\alpha \in K$ und $k \in \mathbb{N} \cup \{0\}$ und setzt man

$$\binom{\alpha}{k} := \begin{cases} 1 & \text{falls } k = 0 \\ \frac{1}{k!} \prod_{\ell=1}^k (\alpha - \ell + 1) & \text{falls } k \in \mathbb{N}. \end{cases}$$

Genau wie in Kapitel 2 können wir die nachfolgenden Eigenschaften von Binomialkoeffizienten nachweisen.²

²Wir müssen hier eben nur evtl etwas genauer darauf achten, nur die bekannten Rechenregeln in angeordneten Körpern zu benutzen. Das gelingt aber ohne große Modifikationen der Techniken!

Proposition 3.3.26 — EIGENSCHAFTEN VON BINOMIALKOEFFIZIENTEN

Es seien $\alpha \in K$ und $k \in \mathbb{N}$. Dann gilt

(i)

$$\binom{\alpha}{1} = \alpha.$$

(ii)

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

Ist außerdem $n \in \mathbb{N} \cup \{0\}$, so gilt ferner

(iii) Falls $0 \leq k \leq n$ so gilt

$$\binom{n}{k} = \frac{n!}{(n-k)!k!}.$$

(Hierbei wird die Konvention $0! := 1$ benutzt)

(iv) Falls $0 \leq k \leq n$ so gilt

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

(v) Falls $k > n$ so gilt

$$\binom{n}{k} = 0.$$

Beweis. Haargenau wie in Kapitel 2.

Nutzen wir dies und die vollständige Induktion, so finden wir auch einen *rigorosen Beweis* des Binomischen Lehrsatzes (Satz 2.2.18) ohne kombinatorische Überlegungen.

Satz 3.3.27 — BINOMISCHER LEHRSATZ, RIGOROS

Es seien $x, y \in K$ und $n \in \mathbb{N}$. Dann gilt

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

***Beweis.** Beweis durch vollständige Induktion nach n .

IndA: $n = 1$. ($\mathbb{Z} : (x + y)^1 = \sum_{k=0}^1 \binom{1}{k} x^k y^{1-k}$.)

Dazu berechne

$$(x + y)^1 = x + y$$

und

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

IndS: $n \longrightarrow n + 1$.

Es gelte für ein $n \in \mathbb{N}$ die Induktionshypothese

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

$$\text{ZZ:} \quad \text{Dann gilt auch} \quad (x + y)^{n+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k}.$$

Wir berechnen

$$\begin{aligned} (x + y)^{n+1} &= (x + y)(x + y)^n \stackrel{\text{IndH}}{=} (x + y) \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \\ &= x \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} + y \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} = \sum_{k=0}^n \binom{n}{k} (x \cdot x^k) y^{n-k} + \sum_{k=0}^n \binom{n}{k} x^k (y \cdot y^{n-k}) \\ &= \sum_{k=0}^n \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^n \binom{n}{k} x^k y^{n+1-k} \\ &\stackrel{\text{Prop 3.3.17(iii)}}{=} \sum_{k=1}^{n+1} \binom{n}{k-1} x^k y^{n-(k-1)} + \sum_{k=0}^n \binom{n}{k} x^k y^{n+1-k} \\ &= \sum_{k=1}^{n+1} \binom{n}{k-1} x^k y^{n+1-k} + \sum_{k=0}^n \binom{n}{k} x^k y^{n+1-k} \\ &\stackrel{\text{Bem 3.3.18}}{=} \binom{n}{n} x^{n+1} y^{n+1-(n+1)} + \sum_{k=1}^n \binom{n}{k-1} x^k y^{n+1-k} + \sum_{k=1}^n \binom{n}{k} x^k y^{n+1-k} + \binom{n}{0} x^0 y^{n+1-0} \\ &= x^{n+1} + \sum_{k=1}^n \left(\binom{n}{k-1} + \binom{n}{k} \right) x^k y^{n+1-k} + y^{n+1} \\ &\stackrel{\text{Prop 3.3.26(ii) mit } \alpha=n+1}{=} x^{n+1} + \sum_{k=1}^n \binom{n+1}{k} x^k y^{n+1-k} + y^{n+1} \\ &= y^{n+1} + \sum_{k=1}^n \binom{n+1}{k} x^k y^{n+1-k} + x^{n+1} \end{aligned}$$

$$\begin{aligned}
&= \binom{n+1}{0} x^0 y^{n+1-0} + \sum_{k=1}^n \binom{n+1}{k} x^k y^{n+1-k} + \binom{n+1}{n+1} x^{n+1} y^{n+1-(n+1)} \\
&\stackrel{\text{Bem 3.3.18}}{=} \sum_{k=0}^n \binom{n+1}{k} x^k y^{n+1-k} + \binom{n+1}{n+1} x^{n+1} y^{n+1-(n+1)} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} x^k y^{n+1-k}.
\end{aligned}$$

Die Behauptung folgt.

Rationale Zahlen

Nun wollen wir noch die rationalen Zahlen $\mathbb{Q}$, welche es in jedem angeordneten Körper $(K, +, \cdot, \leq)$ mit $K \neq \{0\}$ gibt, definieren. Fixiere dazu wie immer die in Proposition 3.3.6 konstruierte Realisierung $\mathbb{N}$ der natürlichen Zahlen und die ganzen Zahlen wie in Definition 3.3.13.

Die rationalen Zahlen, notiert mit $\mathbb{Q}$ sind im wesentlichen alle Zahlen, die sich als Bruch einer ganzen Zahl und einer natürlichen Zahl darstellen lassen.

Definition 3.3.28 — RATIONALE ZAHLEN

Wir definieren die *rationalen Zahlen (in K)* durch

$$\mathbb{Q} := \left\{ \frac{z}{n} \mid z \in \mathbb{Z}, n \in \mathbb{N} \right\}.$$

An dieser Stelle möchte ich betonen, dass man in *jedem* angeordneten Körper $(K, +, \cdot, \leq)$ eine solche Teilmenge $\mathbb{Q} \subseteq K$ der rationalen Zahlen konstruieren kann. Wie diese konkret aussehen, hängt maßgeblich von K ab. Man spricht wie immer von unterschiedlichen *Realisierungen der rationalen Zahlen*. In einigen Beispielen im Laufe dieses Kapitels haben wir stets auf „die“ rationalen Zahlen verwiesen. Alle diese Beispiele können in jeder beliebigen Realisierung der rationalen Zahlen betrachtet werden.

Sobald wir die rationalen Zahlen gebildet haben, ist leicht festzustellen, dass diese selbst ein angeordneter Körper sind (mit den Operationen $+$, $\cdot$, die sich von K erben).

Proposition 3.3.29 — RATIONALE ZAHLEN

Es ist $(\mathbb{Q}, +, \cdot, \leq)$ ein angeordneter Körper. (Man spricht auch von einem *angeordneten Unterkörper* von K).

Beweisskizze. Das Nachprüfen ist den Lesenden als Übungsaufgabe überlassen. Hierbei ist es hauptsächlich wichtig, nachzuprüfen, dass die Körper-Operationen (Addition, Multiplikation, Bildung des Inversen) innerhalb von $\mathbb{Q}$ stattfinden. Bei den Anordnungsaxiomen ergeben sich übrigens keine Schwierigkeiten: Schränkt man eine Ordnungsrelation auf eine

Teilmenge ein, so erhält man stets wieder eine Ordnungsrelation.

Erinnern wir uns an unser Ziel zurück: Wir wollen die reellen Zahlen einführen. Auf dem Weg dahin haben wir das Konzept eines angeordneten Körpers studiert. Dieses Konzept reicht aber nicht aus, um die reellen Zahlen zu charakterisieren. Denn schließlich sind – wie wir ja oben gesehen haben – auch die rationalen Zahlen schon ein angeordneter Körper. Es fehlt noch eine entscheidende Zutat, welche wir im nächsten Abschnitt besprechen.

3.4 Ordnungsvollständigkeit und Reelle Zahlen

Am Ende dieses Abschnittes werden wir endlich die reellen Zahlen definieren können. Das einzige was uns noch fehlt ist das Konzept der *Ordnungsvollständigkeit*.

Definition 3.4.1 — ORDNUNGSVOLLSTÄNDIGKEIT

Es sei $(M, \leq)$ eine (halb-)angeordnete Menge. Man nennt M *ordnungsvollständig*, falls jede nichtleere, nach oben beschränkte Teilmenge $A \subseteq M$ ein Supremum besitzt.

Genau HIER liegt am Ende der Unterschied zwischen den rationalen und den reellen Zahlen! Wir werden im folgenden (bereits seit einiger Zeit aufgeschobenen) Beispiel sehen, dass $\mathbb{Q}$ nicht ordnungsvollständig ist.

Beispiel 3.4.2 — EINE MENGE OHNE SUPREMUM

Betrachte die angeordnete Menge $(\mathbb{Q}, \leq)$. Wir haben bereits in Beispiel 3.2.13 behauptet, dass die Menge

$$A := \{x \in \mathbb{Q} : x^2 \leq 2\}$$

kein Supremum in $\mathbb{Q}$ besitzt (und wollen dies hier beweisen). Insbesondere gilt: A ist nach oben beschränkt und besitzt kein Supremum. Die Menge $\mathbb{Q}$ ist also insbesondere *nicht ordnungsvollständig*.

Beweis. Dass A nach oben beschränkt ist, überlassen wir den Lesenden als Übungsaufgabe. Nun zur Nicht-Existenz des Supremums. Angenommen es existiert $\sup(A) \in \mathbb{Q}$, etwa nennen wir

$$s := \sup(A) [\in \mathbb{Q}].$$

Beachte: Da $1 \in A$ und s eine obere Schranke von A ist gilt $s \geq 1$ (vgl. Definition 3.2.5). Wir unterscheiden nun 3 Fälle: $s^2 < 2$, $s^2 > 2$ und $s^2 = 2$ und führen alle drei Fälle zum Widerspruch.

$$\text{FALL 1. } s^2 < 2.$$

Definiere

$$d := \frac{2s + 2}{s + 2} \in \mathbb{Q}.$$

(Beachte: Für die Aussage „ $d \in \mathbb{Q}$ “ wurde die Körpereigenschaft von $\mathbb{Q}$ benutzt). Dann gilt

$$\begin{aligned} d - s &= \frac{2s+2}{s+2} - s = \frac{2s+2}{s+2} - \frac{s(s+2)}{s+2} \\ &= \frac{2s+2-s(s+2)}{s+2} = \frac{2s+2-s^2-2s}{s+2} = \frac{2-s^2}{s+2} > 0. \end{aligned}$$

Also

$$s < d. \quad (154)$$

Zusätzlich gilt

$$\begin{aligned} d^2 - 2 &= \left(\frac{2s+2}{s+2} \right)^2 - 2 = \frac{(2s+2)^2}{(s+2)^2} - 2 = \frac{(2s+2)^2 - 2(s+2)^2}{(s+2)^2} \\ &= \frac{4(s+1)^2 - 2(s^2 + 4s + 4)}{(s+2)^2} = \frac{4(s^2 + 2s + 1) - 2(s^2 + 4s + 4)}{(s+2)^2} \\ &= \frac{4s^2 + 8s + 4 - 2s^2 - 8s - 8}{(s+2)^2} = \frac{2s^2 - 4}{(s+2)^2} = 2 \frac{s^2 - 2}{(s+2)^2} = \underbrace{\frac{2}{(s+2)^2}}_{>0} \underbrace{(s^2 - 2)}_{<0} < 0. \end{aligned}$$

Das bedeutet

$$d^2 < 2. \quad (155)$$

Aus (155) folgt $d \in A$. Da s eine obere Schranke von A ist folgt $d \leq s$, vgl. Definition 3.2.5. Dies widerspricht (154). Ein Widerspruch.

$$\text{FALL 2. } s^2 > 2.$$

Definiere wieder

$$d := \frac{2s+2}{s+2} \in \mathbb{Q}. \quad (156)$$

und berechne analog zu oben

$$d - s = \frac{2-s^2}{s+2} < 0.$$

Somit

$$d < s.$$

Weil $s = \sup(A)$ und $d < s$ gilt, ist d keine obere Schranke von A . Somit gibt es ein $x \in A$ mit $d < x$. Da $d > 0$ (wie man leicht an (156) sieht) gilt mit Korollar 3.3.21, dass

$$d^2 < x^2 \leq 2. \quad (157)$$

Berechne haargenau wie in Fall 1, dass

$$d^2 - 2 = \underbrace{\frac{2}{(s+2)^2}}_{>0} \underbrace{(s^2 - 2)}_{>0} > 0,$$

Das bedeutet

$$d^2 > 2,$$

was (157) widerspricht. Auch in diesem Fall erhalten wir einen Widerspruch.

FALL 3. $s^2 = 2$.

Dieser Fall ist ausgeschlossen, denn es gibt keine rationale Zahl mit $s^2 = 2$, wie die nachfolgende Proposition beweist. Man erhält auch einen Widerspruch.

In allen Fällen haben wir einen Widerspruch erhalten. Die Annahme muss also falsch sein.

Proposition 3.4.3 — KEINE RATIONALE WURZEL AUS 2

Es gibt kein $x \in \mathbb{Q}$ mit $x^2 = 2$. ^(a)

^aAnstatt sich den Beweis unten durchzulesen, können Sie ihn sich auch auf YouTube *vorsingen* lassen: MATHE-SONG VON DORFUCHS

***Beweis.** Angenommen es existiert ein $x \in \mathbb{Q}$ mit $x^2 = 2$. Da $x^2 = (-x)^2$ dürfen wir oBdA annehmen, dass $x > 0$ (sonst ersetze eben x durch $-x$). Sei nun gemäß Definition 3.3.28 $x = \frac{z}{n}$ für ein $z \in \mathbb{Z}$ und $n \in \mathbb{N}$. Da $x > 0$ gilt auch $z \in \mathbb{N}$.

Man darf oBdA annehmen, dass z und n nicht beides gerade Zahlen sind. In der Tat: Falls beide gerade sind ersetze z durch $z' := \frac{z}{2} \in \mathbb{N}$ und n durch $n' := \frac{n}{2} \in \mathbb{N}$ und beobachte, dass

$$\frac{z'}{n'} = \frac{z}{n} = x.$$

Fahre so lange mit dieser Ersetzung fort, bis eine der beiden Zahlen nicht mehr durch zwei teilbar ist^a. Die Annahme liefert

$$2 = x^2 = \left(\frac{z}{n}\right)^2 = \frac{z^2}{n^2}.$$

Multiplizieren wir mit n^2 so erhalten wir

$$z^2 = 2n^2. \tag{158}$$

Wir folgern: z^2 ist gerade. Nach Korollar 3.3.15 ist dann auch z gerade. Somit ist $k := \frac{z}{2} \in \mathbb{N}$. Formt man um, so ergibt sich

$$\exists k \in \mathbb{N} : \quad z = 2k.$$

Dies und (158) ergeben

$$2n^2 = z^2 = (2k)^2 = 2 \cdot 2k^2.$$

Teilt man durch 2, so hat man

$$n^2 = 2k^2.$$

Man folgert, dass n^2 gerade ist. Nun folgt wiederum aus Korollar 3.3.15, dass n gerade ist. Wir haben gezeigt, dass z, n beides gerade Zahlen sind. Ein Widerspruch zu unserer Annahme, dass z und n nicht beides gerade Zahlen sind. Die Behauptung ist gezeigt.

^aWir müssten hier eigentlich noch begründen, dass wir nach endlich vielen Ersetzungen diesen Zustand erreichen können. Das liegt aber daran, dass eine gerade natürliche Zahl stets kleiner wird wenn ich sie durch 2 teile. Da natürliche Zahlen aber stets größergleich 1 sind, kann man die Ersetzung nur endlich oft durchführen.

Nun wiederholen wir die Definition der reellen Zahlen, wie wir sie am Anfang des Kapitels getroffen haben. Dank unserer Vorarbeit ergibt sie nämlich jetzt Sinn:

Definition 3.4.4 — REELLE ZAHLEN (VGL. DEFINITION 3.0.1)

$(\mathbb{R}, +, \cdot, \leq)$ heißt *Realisierung der reellen Zahlen* falls

- (i) $(\mathbb{R}, +, \cdot, \leq)$ ein angeordneter Körper ist,
- (ii) $(\mathbb{R}, \leq)$ ordnungsvollständig ist.

Dass es eine Menge wie in Definition 3.0.1 überhaupt gibt, beweisen wir nicht. Wir fixieren aber ab sofort eine solche Menge $\mathbb{R}$ (genau wie deren Operationen $+, \cdot, \leq$) und rechnen innerhalb dieser mit den Gesetzen, die wir bereits studiert haben. Wenn wir im Folgenden von $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$ sprechen, meinen wir stets die natürlichen, ganzen bzw. rationalen Zahlen *in* $\mathbb{R}$ (vgl. Abschnitt 3.3).

Bemerkung 3.4.5

Wir möchten an dieser Stelle noch festhalten, dass aus den Punkten (i),(ii) in Definition 3.0.1 folgt:

Jede nichtleere nach unten beschränkte Teilmenge $A \subseteq \mathbb{R}$ besitzt ein Infimum.

Diese Aussage ist gewissermaßen *spiegelbildlich* zum Begriff der Ordnungsvollständigkeit und hätte stattdessen gefordert werden können.

Beweisskizze. Definiere für eine beliebige Teilmenge $A \subseteq \mathbb{R}$ die Menge

$$A' := \{-x : x \in A\}$$

und mache die folgenden zwei Beobachtungen

- (a) A ist nach unten beschränkt genau dann wenn A' nach oben beschränkt ist
- (b) Falls $\sup A'$ existiert, so existiert auch $\inf A$ und $\inf A = -\sup A'$.

Ein Beweis der Punkte (a) und (b) ist den Lesenden als Übungsaufgabe überlassen. Die Behauptung folgt dann direkt aus der Ordnungsvollständigkeit.

Eigenschaften von $\mathbb{R}$

Im Folgenden studieren wir ein paar Eigenschaften von $\mathbb{R}$, denn Definition 3.0.1 ist sehr abstrakt und muss noch mit Leben gefüllt werden. Nachfolgend werden wir übrigens ohne Kommentar die Erkenntnisse zum reellen Absolutbetrag (vgl. Sektion 2) verwenden.

Wurzelziehen Eine sehr wichtige Operation, die in $\mathbb{R}$ zur Verfügung steht, in $\mathbb{Q}$ aber i.A. nicht, ist das *Wurzelziehen*, d.h. die Lösbarkeit der Gleichung

$$x^n = \alpha \quad \text{für gegebenes } \alpha \geq 0 \text{ und } n \in \mathbb{N}. \quad (159)$$

Proposition 3.4.6 — EXISTENZ UND EINDEUTIGKEIT VON WURZELN

Es sei $\alpha \in \mathbb{R}$ mit $\alpha \geq 0$ und $n \in \mathbb{N}$. Dann gibt es ein eindeutiges $x \in \mathbb{R}$ mit $x \geq 0$ und

$$x^n = \alpha.$$

Beweis. Zwei Dinge sind zu zeigen: Existenz und Eindeutigkeit der n -ten Wurzel (d.h. des Elementes x aus der Beschreibung).

TEIL A. Existenz.

Definiere die Menge

$$A := \{y \in \mathbb{R} : y \geq 0 \text{ und } y^n \leq \alpha\}.$$

Dann ist A nichtleer, da $0 \in A$, denn $0^n = 0 \leq \alpha$.

ZWISCHENBEHAUPTUNG. A ist nach oben beschränkt.

Dazu: Fixiere eine beliebige Zahl $z \geq 1 + \alpha$. Nach Korollar 3.3.21 gilt

$$z^n \geq (1 + \alpha)^n.$$

Es folgt mit demselben Korollar

$$z^n \geq (1 + \alpha)^n = (1 + \alpha) \cdot (1 + \alpha)^{n-1} \underset{\text{Kor3.3.21}}{\geq} (1 + \alpha) \cdot 1^{n-1} = (1 + \alpha) > \alpha.$$

Das bedeutet $z \notin A$. Wir haben gezeigt

$$z \geq 1 + \alpha \implies z \notin A.$$

Äquivalent (nach Kontraposition, Proposition 1.1.7(ii))

$$z \in A \implies z < 1 + \alpha.$$

Mit anderen Worten: Es gilt $z < 1 + \alpha$ für alle $z \in A$. Somit ist $1 + \alpha$ eine obere Schranke von A .

ZWISCHENBEHAUPTUNG ENDE.

Da $\mathbb{R}$ nach Definition 3.0.1 ordnungsvollständig ist, besitzt A ein Supremum (vgl. Definition 3.4.1). Sei etwa

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

Wir werden nun noch zeigen, dass $x^n = \alpha$. Dazu:

Schritt 1. Wir zeigen $x^n \leq \alpha$. Angenommen $x^n > \alpha$. Insbesondere gilt $x \neq 0$, da $x^n > 0$. Da $x = \sup(A)$ gilt wegen Proposition 3.2.26, dass

$$\forall \varepsilon > 0 \quad \exists x_0 \in A : \quad x - \varepsilon < x_0.$$

Betrachtet man nun $\varepsilon' := \frac{x^n - \alpha}{(1+|x|)^n} > 0$, so^a gilt nach Proposition 3.2.26

$$\exists x_0 \in A : \quad x - \varepsilon' < x_0.$$

Insbesondere hat man $(x - \varepsilon')^n < x_0^n \leq \alpha$ (weil $x_0 \in A$). Nun gilt nach dem Binomischen Lehrsatz (Satz 3.3.27)

$$\begin{aligned} (x - \varepsilon')^n &= (-\varepsilon' + x)^n = \sum_{k=0}^n \binom{n}{k} (-\varepsilon')^k x^{n-k} \\ &= x^n + \sum_{k=1}^n \binom{n}{k} (-\varepsilon')^k x^{n-k} = x^n - \varepsilon' \sum_{k=1}^n \binom{n}{k} (-\varepsilon')^{k-1} x^{n-k}. \end{aligned}$$

Man folgert

$$x^n - \varepsilon' \sum_{k=1}^n \binom{n}{k} (-\varepsilon')^{k-1} x^{n-k} = (x - \varepsilon')^n \leq \alpha.$$

Addiert man den Summenausdruck auf beiden Seiten, so folgt

$$\begin{aligned} x^n &\leq \alpha + \varepsilon' \sum_{k=1}^n \binom{n}{k} (-\varepsilon')^{k-1} x^{n-k} \\ &\leq \alpha + \varepsilon' \sum_{k=1}^n \binom{n}{k} |(-\varepsilon')^{k-1} x^{n-k}| \leq \alpha + \varepsilon' \sum_{k=1}^n \binom{n}{k} (\varepsilon')^{k-1} |x|^{n-k}. \end{aligned}$$

Nutzt man nun, dass $\varepsilon' < 1$ (siehe Fußnote), so erhält man

$$x^n \leq \alpha + \varepsilon' \sum_{k=1}^n \binom{n}{k} |x|^{n-k} \underset{x \neq 0}{<} \alpha + \varepsilon' \sum_{k=0}^n \binom{n}{k} |x|^{n-k} \underset{\text{Bin. Lehrsatz}}{=} \alpha + \varepsilon'(1 + |x|^n)$$

$$\underset{\text{Wahl von } \varepsilon'}{=} \alpha + \frac{x^n - \alpha}{(1 + |x|)^n} (1 + |x|)^n = x^n.$$

Ein Widerspruch ($x^n < x^n$). Die Annahme „ $x^n > \alpha$ “ muss also falsch sein und wir erhalten $x^n \leq \alpha$.

Schritt 2. Wir zeigen $x^n \not< \alpha$. Nun angenommen, dass $x^n < \alpha$. Definieren Sie nun $\varepsilon' := \min\{\frac{\alpha - x^n}{(1 + |x|)^n}, 1\} > 0$. Mit dem Binomischen Lehrsatz (Satz 3.3.27) ergibt sich dann

$$\begin{aligned} (x + \varepsilon')^n &= (\varepsilon' + x)^n = \sum_{k=0}^n \binom{n}{k} (\varepsilon')^k x^{n-k} = x^n + \sum_{k=1}^n \binom{n}{k} (\varepsilon')^k x^{n-k} \\ &= x^n + \varepsilon' \sum_{k=1}^n \binom{n}{k} (\varepsilon')^{k-1} x^{n-k} \leq x^n + \varepsilon' \sum_{k=1}^n \binom{n}{k} (\varepsilon')^{k-1} |x|^{n-k} \\ &\underset{\varepsilon' \leq 1}{\leq} x^n + \varepsilon' \sum_{k=1}^n \binom{n}{k} |x|^{n-k} \leq x^n + \varepsilon' \sum_{k=0}^n \binom{n}{k} |x|^{n-k} \\ &\underset{\text{Satz 3.3.27}}{=} x^n + \varepsilon'(1 + |x|)^n \leq x^n + \frac{\alpha - x^n}{(1 + |x|)^n} (1 + |x|)^n = \alpha. \end{aligned}$$

Das bedeutet $x + \varepsilon' \in A$, was der Tatsache, dass x obere Schranke von A ist, widerspricht. Es bleibt also nur noch „ $x^n = \alpha$ “ übrig.

TEIL B. Eindeutigkeit.

Folgt direkt aus Korollar 3.3.21. In der Tat: Angenommen es gibt zwei verschiedene $x_1, x_2 \in \mathbb{R}$, $x_1 \geq 0, x_2 \geq 0$ mit $x_1 \neq x_2$ und $x_1^n = x_2^n = \alpha$. OBdA gilt $x_1 < x_2$ (sonst sortiere die beiden Zahlen nach der Größe um). Nach Korollar 3.3.21 (ii) gilt nun aber auch

$$x_1^n < x_2^n,$$

was den Widerspruch $\alpha < \alpha$ ergibt.

^aBeachte: $\varepsilon' \leq \frac{x^n}{(1 + |x|)^n} \leq 1$.

Definition 3.4.7 — WURZELN

Es sei $\alpha \in \mathbb{R}$ mit $\alpha \geq 0$ und $n \in \mathbb{N}$. Die eindeutige Zahl $x \in \mathbb{R}$ mit $x \geq 0$ und $x^n = \alpha$ bezeichnen wir als n -te Wurzel von α und schreiben $x = \sqrt[n]{\alpha}$. Im Fall $n = 2$ sagen wir auch (Quadrat-)Wurzel von α und schreiben nur $x = \sqrt{\alpha}$.

Für Wurzeln gelten erwartbare Rechengesetze, wie zum Beispiel die folgende *Multiplikativität*.

Proposition 3.4.8 — RECHENREGELN FÜR DIE WURZEL

Es seien $\alpha, \beta \in \mathbb{R}$ mit $\alpha \geq 0$ und $\beta \geq 0$. Ferner sei $n \in \mathbb{N}$.

(A) Ist $\alpha \leq \beta$ so gilt $\sqrt[n]{\alpha} \leq \sqrt[n]{\beta}$

(B) Es gilt $\sqrt[n]{\alpha\beta} = \sqrt[n]{\alpha} \sqrt[n]{\beta}$

Beweis. Zu (A). Sei $\alpha \leq \beta$. Nehmen wir zunächst das Gegenteil an, d.h. $\sqrt[n]{\alpha} > \sqrt[n]{\beta}$. Nach Proposition 3.3.21 (anwendbar weil es sich bei beiden Wurzeln um nichtnegative Zahlen handelt) gilt dann auch $(\sqrt[n]{\alpha})^n > (\sqrt[n]{\beta})^n$. Daraus folgt $\alpha > \beta$, ein Widerspruch.

Zu (B). Sei $x := \sqrt[n]{\alpha\beta}$. Dann ist x eindeutig charakterisiert durch die Eigenschaften

(i) $x \geq 0$

(ii) $x^n = \alpha\beta$.

Wir werden zeigen, dass $\sqrt[n]{\alpha} \sqrt[n]{\beta}$ auch dieselben Eigenschaften (i) und (ii) besitzt. Wegen der Eindeutigkeit folgt dann $\sqrt[n]{\alpha} \sqrt[n]{\beta} = x = \sqrt[n]{\alpha\beta}$. Dazu

- Zu (i). $\sqrt[n]{\alpha} \sqrt[n]{\beta}$ erfüllt (i), da sowohl $\sqrt[n]{\alpha} \geq 0$ als auch $\sqrt[n]{\beta} \geq 0$ (vergleiche auch Axiom (V2) in Definition 3.2.16)
- Zu (ii). Wir rechnen mithilfe der Potenzgesetze aus Proposition 3.3.8

$$(\sqrt[n]{\alpha} \sqrt[n]{\beta})^n = (\sqrt[n]{\alpha})^n (\sqrt[n]{\beta})^n = \alpha\beta,$$

wobei wir im letzten Rechenschritt Definition 10.6.14 verwendet haben.

Die Behauptung folgt wie oben angekündigt aus der Eindeutigkeit der Wurzel.

Bemerkung 3.4.9

Es sei $x \in \mathbb{R}$ (eventuell auch negativ). Dann gilt stets $x^2 \geq 0$ (vgl. Proposition 3.2.20(x)). Somit kann man auch die Quadratwurzel $\sqrt{x^2}$ bilden. Diese stimmt aber nicht zwingend mit x überein, denn nach Definition 10.6.14 muss $\sqrt{x^2}$ zwingend eine positive Zahl sein. Stattdessen gilt

$$\sqrt{x^2} = |x|$$

Davon überzeugt man sich auch leicht mithilfe der Eindeutigkeit der Quadratwurzel.

In $\mathbb{Q}$ ist Gleichung (159) nicht für alle α , n lösbar — Schließlich haben wir in Lemma 3.4.3 gesehen, dass die Gleichung $x^2 = 2$ keine rationale Lösung besitzt! Ein fundamentaler Unterschied zwischen $\mathbb{Q}$ und $\mathbb{R}$! Und übrigens auch ein Beweis, dass $\mathbb{Q} \neq \mathbb{R}$ – denn $\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$!

Definition 3.4.10 — IRRATIONALE ZAHLEN

Eine Zahl $z \in \mathbb{R} \setminus \mathbb{Q}$ nennen wir *irrational*.

***Die Mitternachtsformel.** In der Schule haben Sie die *Mitternachtsformel* gelernt, um für $a \in \mathbb{R} \setminus \{0\}$ sowie $b, c \in \mathbb{R}$ alle Lösungen $x \in \mathbb{R}$ der quadratischen Gleichung

$$ax^2 + bx + c = 0$$

zu bestimmen. Diese Formel wollen wir hier noch einmal festhalten.

Proposition 3.4.11 — MITTERNACHTSFORMEL

Es seien $a \in \mathbb{R} \setminus \{0\}$ sowie $b, c \in \mathbb{R}$.

(i) Ist $b^2 - 4ac < 0$ so gibt es kein $x \in \mathbb{R}$ mit $ax^2 + bx + c = 0$.

(ii) Ist $b^2 - 4ac \geq 0$ so gilt

$$ax^2 + bx + c = 0 \quad \Longleftrightarrow \quad x \in \left\{ \frac{-b - \sqrt{b^2 - 4ac}}{2a}, \frac{-b + \sqrt{b^2 - 4ac}}{2a} \right\}$$

Beweis. Zuerst eine Vorüberlegung. Es sei $x \in \mathbb{R}$. Dann gilt

$$\begin{aligned} ax^2 + bx + c &= 0 \\ \Longleftrightarrow x^2 + \frac{b}{a}x + \frac{c}{a} &= 0 \\ \Longleftrightarrow x^2 + \frac{b}{a}x &= -\frac{c}{a} \\ \Longleftrightarrow x^2 + \frac{b}{a}x + \left(\frac{b}{2a}\right)^2 &= \left(\frac{b}{2a}\right)^2 - \frac{c}{a} \\ \Longleftrightarrow x^2 + 2\frac{b}{2a}x + \left(\frac{b}{2a}\right)^2 &= \left(\frac{b}{2a}\right)^2 - \frac{c}{a} \\ \stackrel{1. \text{ Binom. }}{\Longleftrightarrow} \left(x + \frac{b}{2a}\right)^2 &= \frac{b^2}{4a^2} - \frac{c}{a} = \frac{b^2 - 4ac}{4a^2}. \end{aligned}$$

Definiert man $y := x + \frac{b}{2a}$ und $\theta := \frac{b^2 - 4ac}{4a^2}$ so gilt also $y^2 = \theta$.

Nun zu (i). Es sei $b^2 - 4ac < 0$. Angenommen es gibt ein $x \in \mathbb{R}$ mit $ax^2 + bx + c = 0$. Definiere dann wie oben $y := x + \frac{b}{2a}$. Weil $4a^2 > 0$ (vgl. Proposition 3.2.20(x)) ist auch $\theta = \frac{b^2 - 4ac}{4a^2} < 0$, vgl. Proposition 3.2.20(ii). Nun gilt nach Proposition 3.2.20(x), dass $y^2 = y \cdot y \geq 0$. Ein Widerspruch zu $y^2 = \theta$.

Nun zu (ii). Sei $b^2 - 4ac \geq 0$. Ähnlich wie in der obigen Argumentation sieht man dann $\theta \geq 0$. Wie in Definition 10.6.14 können wir dann auch $\sqrt{\theta}$ bilden und es gilt $\theta = (\sqrt{\theta})^2$. Es sei dann wiederum $x \in \mathbb{R}$ mit $ax^2 + bx + c = 0$. Definiere wiederum $y := x + \frac{b}{2a}$. Aus der Vorüberlegung folgt

$$\begin{aligned}
 ax^2 + bx + c &= 0 \\
 \iff y^2 &= \theta \\
 \iff y^2 - \theta &= 0 \\
 \iff y^2 - (\sqrt{\theta})^2 &= 0 \\
 \iff (y - \sqrt{\theta})(y + \sqrt{\theta}) &= 0 \\
 \iff y &\in \{-\theta, \theta\} \\
 \text{Prop 3.1.9 (ix)} \quad \iff x + \frac{b}{2a} &\in \{-\theta, \theta\} \\
 \iff x \in \{-\frac{b}{2a} - \theta, \frac{b}{2a} + \theta\}.
 \end{aligned}$$

Aus Blatt 9 Aufgabe 1d wissen wir, dass

$$\theta = \sqrt{\frac{b^2 - 4ac}{4a^2}} = \frac{\sqrt{b^2 - 4ac}}{\sqrt{4a^2}} = \frac{\sqrt{b^2 - 4ac}}{\sqrt{4}\sqrt{a^2}} = \frac{b^2 - 4ac}{2|a|}$$

Dies zusammen mit der Tatsache, dass $|a| \in \{+a, -a\}$ liefert

$$\begin{aligned}
 \{-\frac{b}{2a} - \theta, \frac{b}{2a} + \theta\} &= \left\{ \frac{-b - \sqrt{b^2 - 4ac}}{2|a|}, \frac{-b + \sqrt{b^2 - 4ac}}{2|a|} \right\} \\
 &= \left\{ \frac{-b - \sqrt{b^2 - 4ac}}{2a}, \frac{-b + \sqrt{b^2 - 4ac}}{2a} \right\},
 \end{aligned}$$

(wobei wir im Fall $|a| = +a$ die Reihenfolge behalten können und im Fall $|a| = -a$ die Reihenfolge nochmal tauschen müssen).

Die Archimedes-Eigenschaft Die Archimedes-Eigenschaft ist etwas für unsere Intuition alltägliches, was aber trotzdem eines Beweises bedarf. 27.11.2025

Proposition 3.4.12 — ARCHIMEDES-EIGENSCHAFT

$\mathbb{R}$ besitzt die sogenannte *Archimedes-Eigenschaft*, d.h.

$$\forall x \in \mathbb{R} \quad \exists n \in \mathbb{N} : \quad x \leq n.$$

Beweis. Es sei $x \in \mathbb{R}$ beliebig aber fest. Angenommen, das Gegenteil ist der Fall, d.h.

$$\neg(\exists n \in \mathbb{N} : \quad x \leq n).$$

Äquivalent zur obigen Aussage ist

$$\forall n \in \mathbb{N} : \quad x > n.$$

Es folgt

x ist eine obere Schranke von $\mathbb{N}$.

Somit ist $\mathbb{N}$ nach oben beschränkt^a. Weil $\mathbb{R}$ nun ordnungsvollständig ist, existiert das Supremum $s := \sup(\mathbb{N})$ (vgl. Definition 3.4.1). Verwenden wir Proposition 3.2.26 mit $\varepsilon = \frac{1}{2}$, so gilt

$$\exists x_0 \in \mathbb{N} : s - \frac{1}{2} < x_0. \quad (160)$$

Definiere nun

$$x^* := x_0 + 1.$$

Dann gilt $x^* \in \mathbb{N}$ (wegen der Definition der Nachfolgerabbildung in Proposition 3.3.6). Ferner gilt wegen (160)

$$x^* = x_0 + 1 > (s - \frac{1}{2}) + 1 = s + \frac{1}{2} > s.$$

Wir erhalten einen Widerspruch: Denn s muss eine obere Schranke von $\mathbb{N}$ sein und daher muss $x^* \leq s$ gelten (wegen Definition 3.2.5 und weil $x^* \in \mathbb{N}$). Die Behauptung folgt.

^aDiese Aussage liest sich schon ziemlich abwegig. Wir müssen hier allerdings zu einem echten logischen Widerspruch kommen und müssen daher weiterargumentieren...

Für jede reelle Zahl können wir also eine natürliche Zahl finden, die größer ist. Dies haben wir in Kapitel 2 exzessiv benutzt! Zum Beispiel in Proposition 2.2.6 und Proposition 2.2.26. Die Definition der *oberen Gauß-Klammer* in Bemerkung 2.2.7 wird auch erst sinnvoll, wenn man die Archimedes-Eigenschaft gezeigt hat. An dieser Stelle bräuchte man sogar noch ein paar mehr Details. Wir verzichten aber zugunsten der Kompaktheit des Skriptes darauf und benutzen ab jetzt die obere Gauß-Klammer gemäß der folgenden Definition

Definition 3.4.13 — OBERE GAUSS-KLAMMER

Für $x \in \mathbb{R}$ ist $\lceil x \rceil$ die eindeutige ganze Zahl $z \in \mathbb{Z}$ mit $x \leq z < x + 1$.

Eine wichtige Folgerung aus der Archimedes-Eigenschaft ist, dass es zwischen je zwei beliebigen reellen Zahlen stets eine rationale Zahl gibt. Man sagt auch, die rationalen Zahlen *liegen dicht* in $\mathbb{R}$.

Proposition 3.4.14 — DIE RATIONALEN ZAHLEN LIEGEN DICHT

Es seien $x, y \in \mathbb{R}$ mit $x < y$. Dann gibt es ein $q \in \mathbb{Q}$ mit $x < q < y$. Insbesondere gilt $q \in (x, y)_{\mathbb{R}}$.

Beweis. Wegen Proposition 3.4.12 (angewendet auf $x' = \frac{2}{y-x}$) gilt

$$\exists m \in \mathbb{N} : \quad \frac{2}{y-x} \leq m.$$

Definiere $m' := m + 1$. Dann gilt

$$\frac{2}{y-x} < m'. \quad \xRightarrow{y-x>0} \quad 2 < m'(y-x) \quad \Longrightarrow \quad 2 < m'y - m'x \quad \Longrightarrow \quad m'x + 2 < m'y. \quad (161)$$

Definiere nun $z := \lceil m'x + 1 \rceil \in \mathbb{Z}$. Nach Definition der oberen Gaußklammer (vgl. Definition 3.4.13) gilt

$$m'x + 1 \leq z < m'x + 2.$$

Es folgt

$$m'x < m'x + 1 \leq z < m'x + 2 \underset{(161)}{<} m'y.$$

Das heißt zusammengefasst: $m'x < z < m'y$ (*). Teilt man in (*) durch m' so ergibt sich

$$x < \frac{z}{m'} < y.$$

Setzen wir $q := \frac{z}{m'} \in \mathbb{Q}$ so folgt die Behauptung.

3.5 Folgen in $\mathbb{R}$

Ab hier erinnern wir uns nun an Kapitel 2, in dem wir über Folgen reeller Zahlen $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ gesprochen haben. Sie dürfen in den Übungen nun wieder die Erkenntnisse aus Kapitel 2 verwenden. Ich lade Sie an dieser Stelle dazu ein, durch Kapitel 2 durchzugehen und alles was wir dort gemacht haben mithilfe der gelernten Eigenschaften der reellen Zahlen auf *Herz und Nieren* zu überprüfen. Sie werden damit Erfolg haben, bis auf zwei (kleine) Ausnahmen:

- Bei der Definition der Endlichkeit müssten Sie die Menge $\{1, \dots, n\}$ stets durch die Menge $\{m \in \mathbb{N} : m \leq n\}$ ersetzen. Streng genommen wissen wir nicht, dass diese beiden Mengen übereinstimmen (da wir $\mathbb{N}$ ja nur axiomatisch eingeführt haben). Wir werden dieses Problem im Folgenden ignorieren.
- In Proposition 2.2.6 müssen wir den Potenz-Ausdruck n^θ für ein $\theta \in \mathbb{R}$ definieren. Wir haben Potenzen aber bisher nur für Exponenten $\theta \in \mathbb{Z}$ definiert... Diese Diskussion wollen wir aber noch vertagen.

Ansonsten haben wir aus Kapitel 2 keine *Altlasten* mehr, also nichts aus Kapitel 2, was ohne Beweis geblieben ist.

Mit den herausgefundenen Eigenschaften von $\mathbb{R}$ können wir mehr über Konvergenz von Folgen sagen. Wiederholen wir zunächst die Definition der Folgenkonvergenz.

Definition 3.5.1 — KONVERGENZ VON FOLGEN (VGL. DEFINITION 2.3.1)

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine reelle Folge. Wir sagen, $(a_n)_{n \in \mathbb{N}}$ *konvergiert* gegen ein $a \in \mathbb{R}$ falls

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |a_n - a| \leq \varepsilon. \quad (162)$$

Wir nennen $(a_n)_{n \in \mathbb{N}}$ *konvergent*, falls es ein $a \in \mathbb{R}$ gibt, sodass $(a_n)_{n \in \mathbb{N}}$ gegen a konvergiert.

Bisher haben wir immer (162) und unsere in Kapitel 2 hergeleiteten Grenzwertregeln benutzt, um Folgen auf Konvergenz zu untersuchen. Ein wesentlicher Nachteil von (162) ist, dass der Grenzwert a bereits bei der Überprüfung bekannt sein muss. Man muss also bereits im Vorhinein eine Vermutung haben, was der Grenzwert einer Folge ist, damit man deren Konvergenz nachweisen kann.

In diesem Kapitel werden wir Methoden kennenlernen, mit denen wir von diesem Prinzip abweichen können: Wir formulieren Konvergenzkriterien, für die wir im Vorhinein keine Vermutung über den Grenzwert anstellen müssen.

Monotone Folgen

Ein erstes Konvergenzkriterium betrifft monotone und beschränkte Folgen.

Proposition 3.5.2 — MONOTONE BESCHRÄNKTE FOLGEN KONVERGIEREN

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine Folge.

- (i) Ist $(a_n)_{n \in \mathbb{N}}$ monoton wachsend und nach oben beschränkt, so ist $(a_n)_{n \in \mathbb{N}}$ konvergent. Ferner gilt

$$\lim_{n \rightarrow \infty} a_n = \sup\{a_n : n \in \mathbb{N}\}.$$

- (ii) Ist $(a_n)_{n \in \mathbb{N}}$ monoton fallend und nach unten beschränkt, so konvergiert $(a_n)_{n \in \mathbb{N}}$. Ferner gilt

$$\lim_{n \rightarrow \infty} a_n = \inf\{a_n : n \in \mathbb{N}\}.$$

Beweis. Wir zeigen nur (i). Punkt (ii) geht dann analog. Es sei $(a_n)_{n \in \mathbb{N}}$ eine monoton wachsende und nach oben beschränkte Folge. Definiere

$$a := \sup\{a_n : n \in \mathbb{N}\}. \quad (163)$$

(Dass dieses Supremum existiert liegt daran, dass $\{a_n : n \in \mathbb{N}\}$ nach oben beschränkt ist und $\mathbb{R}$ ordnungsvollständig ist, vgl. Definition 3.4.1). Um die Konvergenz zu zeigen verwenden wir (162). Dazu sei $\varepsilon > 0$ beliebig vorgegeben.

$$\mathbb{Z}: \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |a_n - a| \leq \varepsilon. \quad (164)$$

Nach Proposition 3.2.26 und der Wahl von a als Supremum (vgl. (163)) gilt

$$\exists x_0 \in \{a_n : n \in \mathbb{N}\} : \quad a - \varepsilon < x_0.$$

Beachte: $x_0 = a_{n_1}$ für ein $n_1 \in \mathbb{N}$. Somit stellen wir fest

$$\exists n_1 \in \mathbb{N} : \quad a - \varepsilon < a_{n_1}.$$

Sei nun $n \geq n_1$. Dann gilt $a_{n_1} \leq a_n$ und somit

$$\forall n \geq n_1 : \quad a - \varepsilon \leq a_{n_1} \leq a_n \quad . \quad (165)$$

Da a nach (163) eine obere Schranke von $\{a_n : n \in \mathbb{N}\}$ ist, gilt auch $a_n \leq a$ für alle $n \in \mathbb{N}$. Dies und (165) implizieren

$$\forall n \geq n_1 : \quad a - \varepsilon \leq a_n \leq a.$$

Zieht man a auf beiden Seiten ab, so ergibt sich

$$\forall n \geq n_1 : \quad -\varepsilon \leq a_n - a \leq 0.$$

Man folgert

$$\forall n \geq n_1 : \quad |a_n - a| \leq \varepsilon.$$

Wählt man nun $n_0 := n_1$, so ist (164) gezeigt.

Bemerkung 3.5.3

Auf die Monotonie-Voraussetzung in Proposition 3.5.2 kann nicht verzichtet werden! Beschränktheit alleine reicht für die Konvergenz nicht aus. Ein gutes Beispiel dafür ist die Folge $((-1)^n)_{n \in \mathbb{N}}$. Diese ist zwar beschränkt (vgl. Beispiel 2.1.4) aber nicht konvergent (vgl. Beispiel 2.3.10).

Bemerkung 3.5.4

Auf die Beschränktheits-Voraussetzung in Proposition 3.5.2 kann nicht verzichtet werden! Aus Proposition 2.3.14 wissen wir nämlich, dass konvergente Folgen stets automatisch beschränkt sein müssen!

Man kann allerdings zeigen:

- (i) Ist eine Folge $(a_n)_{n \in \mathbb{N}}$ monoton wachsend und nicht nach oben beschränkt, so divergiert sie bestimmt gegen ∞ (im Sinne von Definition 2.3.35).
- (ii) Ist eine Folge $(a_n)_{n \in \mathbb{N}}$ monoton fallend und nicht nach unten beschränkt, so divergiert sie bestimmt gegen $-\infty$ (im Sinne von Definition 2.3.35).

Der Beweis ist ähnlich zu Proposition 3.5.2 und ist den Lesenden als Übungsaufgabe überlassen.

Verbindet man diese Aussage mit Proposition 3.5.2, so ergibt sich folgendes

FAZIT:

Monotone Folgen sind entweder konvergent oder sie divergieren bestimmt.

Die Euler'sche Zahl Als nächstes lernen wir die Euler'sche Zahl e kennen. Sie kommt häufig vor, wenn man *Wachstumsprozesse* beschreiben will, siehe auch die folgende Überlegung.

Beispiel 3.5.5 — MOTIVATION DER e -FUNKTION

Stellen Sie sich vor, Sie haben ein Konto mit einem Guthaben von 1€. Um die Zinsen zu optimieren, schauen Sie sich nach verschiedenen Banken um. Sie holen folgende Angebote der Banken B1, B2, B3, ... ein

B1: Es gibt 1-mal im Jahr 100% Zinsen.

B2: Es gibt 2-mal im Jahr 50% Zinsen.

B3: Es gibt 3-mal im Jahr $33\frac{1}{3}\%$ Zinsen.

B4: Es gibt 4-mal im Jahr 25% Zinsen.

... ..

B n : Es gibt n -mal im Jahr $\frac{100}{n}\%$ Zinsen.

... ..

Zu welcher Bank würden Sie gehen? Ist es besser einmal im Jahr den hohen Zinssatz zu erhalten oder öfters den kleineren Zinssatz zu erhalten? Wie reich können Sie damit maximal werden?

Erhalten Sie auf einen Kontostand von x eine Zinszahlung von $\frac{100}{n}\%$, so ist ihr Kontostand danach $x + \frac{x}{n} = (1 + \frac{1}{n}) \cdot x$. Legen wir nun mal unseren 1€ bei der Bank B n an und sehen uns an, wie viel nach einem Jahr dazugekommen ist. Beachten Sie: Bei B n gibt es n Zinszahlungen im Jahr

Zinszahlung no.	Vorher	1	2	...	n
Kontostand	1€	$(1 + \frac{1}{n}) \cdot 1€$	$(1 + \frac{1}{n})[(1 + \frac{1}{n}) \cdot 1]€$	...	$(1 + \frac{1}{n})^n \cdot 1 €$

Mit anderen Worten haben wir am Ende des Jahres bei Bank B n ein Vermögen von $(1 + \frac{1}{n})^n €$.

Die obige Überlegung motiviert die Betrachtung der *Exponentialfolge* $((1 + \frac{1}{n})^n)_{n \in \mathbb{N}}$. Wir werden zeigen, dass diese Folge konvergent ist. Hierzu verwenden wir Proposition 3.5.2. In den

nächsten beiden Lemmata prüfen wir deren Voraussetzungen (Monotonie und Beschränktheit)

Lemma 3.5.6 — MONOTONIE DER EXPONENTIALFOLGE

Für $n \in \mathbb{N}$ definiere

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

Dann ist $(a_n)_{n \in \mathbb{N}}$ monoton wachsend.

Beweis. Für $n \in \mathbb{N}$ berechne

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= \frac{\left(1 + \frac{1}{n+1}\right)^{n+1}}{\left(1 + \frac{1}{n}\right)^n} = \frac{\left(1 + \frac{1}{n+1}\right)^{n+1}}{\left(1 + \frac{1}{n}\right)^{n+1}} \left(1 + \frac{1}{n}\right) = \left[\frac{\left(1 + \frac{1}{n+1}\right)}{\left(1 + \frac{1}{n}\right)} \right]^{n+1} \left(1 + \frac{1}{n}\right) \\ &= \left[\frac{\left(\frac{n+2}{n+1}\right)}{\left(\frac{n+1}{n}\right)} \right]^{n+1} \left(1 + \frac{1}{n}\right) = \left[\frac{(n+2)n}{(n+1)^2} \right]^{n+1} \left(1 + \frac{1}{n}\right) = \left[\frac{n^2 + 2n}{(n+1)^2} \right]^{n+1} \left(1 + \frac{1}{n}\right) \\ &= \left[\frac{n^2 + 2n + 1 - 1}{(n+1)^2} \right]^{n+1} \left(1 + \frac{1}{n}\right) \stackrel{\text{2.Binom.}}{=} \left[\frac{(n+1)^2 - 1}{(n+1)^2} \right]^{n+1} \left(1 + \frac{1}{n}\right) \\ &= \left[1 - \frac{1}{(n+1)^2} \right]^{n+1} \left(1 + \frac{1}{n}\right) \stackrel{\text{Prop3.3.11}}{\geq} \left[1 - (n+1) \frac{1}{(n+1)^2} \right] \left(1 + \frac{1}{n}\right) \\ &= \left[1 - \frac{1}{n+1} \right] \left(1 + \frac{1}{n}\right) = \left[\frac{n}{n+1} \right] \left(\frac{n+1}{n} \right) = 1. \end{aligned}$$

Multipliziert man mit a_n (und benutzt, dass $a_n > 0$), so erhält man $a_{n+1} \geq a_n$.

Lemma 3.5.7 — BESCHRÄNKTHEIT DER EXPONENTIALFOLGE

Für $n \in \mathbb{N}$ definiere

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

Dann ist $(a_n)_{n \in \mathbb{N}}$ nach oben beschränkt. Konkreter gilt $a_n \leq 3$ für alle $n \in \mathbb{N}$.

Beweis. Nach dem Binomischen Lehrsatz (Satz 3.3.27) mit $x = \frac{1}{n}$ und $y = 1$ gilt

$$\begin{aligned} a_n &= \left(1 + \frac{1}{n}\right)^n = \left(\frac{1}{n} + 1\right)^n \stackrel{\text{Satz3.3.27}}{=} \sum_{k=0}^n \binom{n}{k} \left(\frac{1}{n}\right)^k = \underbrace{\binom{n}{0} \left(\frac{1}{n}\right)^0}_{=1} + \sum_{k=1}^n \binom{n}{k} \frac{1}{n^k} \\ &\stackrel{\text{Def3.3.25}}{=} 1 + \sum_{k=1}^n \frac{1}{k!} \left(\prod_{\ell=1}^k (n - \ell + 1) \right) \frac{1}{n^k} \leq 1 + \sum_{k=1}^n \frac{1}{k!} \left(\prod_{\ell=1}^k n \right) \frac{1}{n^k} \\ &\stackrel{\text{Prop3.3.23(ii)}}{=} 1 + \sum_{k=1}^n \frac{1}{k!} n^k \frac{1}{n^k} = 1 + \sum_{k=1}^n \frac{1}{k!}. \end{aligned} \tag{166}$$

Nun verwenden wir die folgende Ungleichung: Für $k \in \mathbb{N}$ mit $k \geq 2$ gilt

$$k! = \prod_{\ell=1}^k \ell = 1 \cdot \prod_{\ell=2}^k \ell = \prod_{\ell=2}^k \ell \geq \prod_{\ell=2}^k 2 \stackrel{\text{Prop 3.3.23(ii)}}{=} 2^{k-1}.$$

Leicht kann man sich auch davon überzeugen, dass dieselbe Ungleichung auch für $k = 1$ gilt. Dies und (166) ergeben

$$\begin{aligned} a_n &\leq 1 + \sum_{k=1}^n \frac{1}{k!} = 1 + \sum_{k=1}^n \frac{1}{k!} \leq 1 + \sum_{k=1}^n \frac{1}{2^{k-1}} = 1 + \sum_{k=0}^{n-1} \frac{1}{2^k} \\ &= 1 + \sum_{k=0}^{n-1} \left(\frac{1}{2}\right)^k \stackrel{\text{Prop 3.3.20}}{=} 1 + \frac{1 - \left(\frac{1}{2}\right)^n}{1 - \frac{1}{2}} = 1 + \frac{1 - \left(\frac{1}{2}\right)^n}{\frac{1}{2}} = 1 + 2 \left(1 - \left(\frac{1}{2}\right)^n\right) \\ &\leq 1 + 2 \cdot 1 = 3. \end{aligned}$$

Nun endlich können wir die Konvergenz von $\left((1 + \frac{1}{n})^n\right)_{n \in \mathbb{N}}$ folgern. Den Grenzwert dieser Folge nennen wir e , die *Euler'sche Zahl*, welche uns später nochmal begegnen wird, wenn wir über die Exponentialfunktion sprechen.

Proposition 3.5.8 — KONVERGENZ DER EXPONENTIALFOLGE

Für $n \in \mathbb{N}$ definiere $a_n := (1 + \frac{1}{n})^n$. Dann existiert $e := \lim_{n \rightarrow \infty} a_n$ und es gilt $2 \leq e \leq 3$.

Beweis. Wie in der Aussage definiere

$$a_n := \left(1 + \frac{1}{n}\right)^n \quad \forall n \in \mathbb{N}. \quad (167)$$

Nach Lemma 3.5.6 und Lemma 3.5.7 ist $(a_n)_{n \in \mathbb{N}}$ monoton wachsend und nach oben beschränkt. Nach Proposition 3.5.2 folgt, dass $(a_n)_{n \in \mathbb{N}}$ konvergent ist und

$$\lim_{n \rightarrow \infty} a_n = \sup\{a_n : n \in \mathbb{N}\}.$$

Mit der Definition $e := \lim_{n \rightarrow \infty} a_n$ haben wir also

$$e = \sup\{a_n : n \in \mathbb{N}\}. \quad (168)$$

Nun ist

$$a_1 \stackrel{(167)}{=} \left(1 + \frac{1}{1}\right)^1 = 2^1 = 2.$$

Da e nach (168) eine obere Schranke von $\{a_n : n \in \mathbb{N}\}$ ist gilt

$$2 = a_1 \stackrel{\text{Def 3.2.5}}{\leq} e. \quad (169)$$

Ferner ist nach Lemma 3.5.7 die Zahl 3 eine obere Schranke von $\{a_n : n \in \mathbb{N}\}$. Wendet man Definition 3.2.10 (i.b) an mit $s = e = \sup_{(168)} \{a_n : n \in \mathbb{N}\}$ und $s' = 3$ an, so folgt aus (168)

$$e \leq 3.$$

Dies und (169) ergeben $2 \leq e \leq 3$.

Definition 3.5.9 — EULER'SCHE ZAHL

Die Zahl e aus der vorigen Proposition nennen wir *Euler'sche Zahl*.

Teilfolgen und Häufungswerte

Anstatt einer ganzen Folge $(a_n)_{n \in \mathbb{N}}$ können wir uns auch *Teilfolgen* anschauen. Diese entstehen durch Auswahl von unendlich vielen (aber nicht notwendigerweise allen) Folgengliedern a_n .

Definition 3.5.10 — TEILFOLGEN

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq M$ eine Folge in einer beliebigen Menge M . Ferner sei $(\ell_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}$ eine Folge natürlicher Zahlen so, dass $\ell_{n+1} \geq \ell_n + 1$ für alle $n \in \mathbb{N}$. Dann nennt man die Folge $(a_{\ell_n})_{n \in \mathbb{N}} \subseteq M$ eine *Teilfolge* von $(a_n)_{n \in \mathbb{N}}$.

Anstatt der gesamten Folge $a_1, a_2, \dots$ betrachten wir nun also $a_{\ell_1}, a_{\ell_2}, \dots$. Die Bedingung $\ell_{n+1} \geq \ell_n + 1$ bedeutet nichts anderes, als dass die Indexfolge $(\ell_n)_{n \in \mathbb{N}}$ *strikt aufsteigend* ist.

Beispiel 3.5.11

Es sei $(a_n)_{n \in \mathbb{N}}$ eine beliebige Folge.

- (i) Wählt man in Definition 3.5.10 $\ell_n := n$ für alle $n \in \mathbb{N}$ so ergibt sich, dass $(a_n)_{n \in \mathbb{N}}$ selbst eine Teilfolge von $(a_n)_{n \in \mathbb{N}}$ ist. Die Wahl $\ell_n := n$ ist zulässig, weil

$$\ell_{n+1} = n + 1 = \ell_n + 1.$$

- (ii) Wählt man in Definition 3.5.10 $\ell_n := 2n$ für alle $n \in \mathbb{N}$ so ergibt sich, dass $(a_{2n})_{n \in \mathbb{N}}$ (d.h. alle *geraden Indizes*) eine Teilfolge von $(a_n)_{n \in \mathbb{N}}$ ist. Die Wahl $\ell_n := 2n$ ist zulässig, weil

$$\ell_{n+1} = 2(n+1) = 2n + 2 \geq 2n + 1 = \ell_n + 1.$$

- (iii) Wählt man in Definition 3.5.10 $\ell_n := 2n + 1$ für alle $n \in \mathbb{N}$ so ergibt sich, dass $(a_{2n+1})_{n \in \mathbb{N}}$ (d.h. alle *ungeraden Indizes*) eine Teilfolge von $(a_n)_{n \in \mathbb{N}}$ ist. Überzeugen Sie sich als Übungsaufgabe davon, dass die Wahl $\ell_n = 2n + 1$ zulässig ist.

- (iv) Allgemein ist für $k \in \mathbb{N}$ und $r \in \mathbb{N} \cup \{0\}$ die Wahl $\ell_n = kn + r$ zulässig. Wir werden derartige Teilfolgen oft betrachten.

Bemerkung 3.5.12

Es sei $(\ell_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}$ eine Folge wie in 3.5.10, d.h. $\ell_{n+1} \geq \ell_n + 1$ für alle $n \in \mathbb{N}$. Mit Induktion zeigt man leicht, dass

$$\ell_n \geq n \quad \forall n \in \mathbb{N}.$$

Diese Ungleichung werden wir oft benutzen.

Eine erste wichtige Beobachtung ist, dass Teilfolgen konvergenter Folgen stets konvergent sind (und denselben Grenzwert haben).

Proposition 3.5.13 — TEILFOLGEN KONVERGENTER FOLGEN

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine konvergente Folge. Dann ist jede Teilfolge $(a_{\ell_n})_{n \in \mathbb{N}}$ konvergent und es gilt

$$\lim_{n \rightarrow \infty} a_{\ell_n} = \lim_{n \rightarrow \infty} a_n.$$

Beweis. Es sei $a := \lim_{n \rightarrow \infty} a_n$. Sei nun $\varepsilon > 0$ beliebig vorgegeben.

$$\mathbb{Z}: \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |a_{\ell_n} - a| \leq \varepsilon.$$

Wir wissen aus (162)

$$\exists n_1 \in \mathbb{N} \quad \forall n \geq n_1 : \quad |a_n - a| \leq \varepsilon. \quad (170)$$

Nun wähle $n_0 := n_1$ (wobei n_1 wie in (170)). Für $n \geq n_0$ gilt dann

$$\ell_n \underset{\text{Bem 3.5.12}}{\geq} n \geq n_0 = n_1.$$

Weil $\ell_n \geq n_1$ und folgt dann aus (170)

$$|a_{\ell_n} - a| \leq \varepsilon.$$

Wir haben also gesehen

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |a_{\ell_n} - a| \leq \varepsilon.$$

Die Behauptung folgt.

Beispiel 3.5.1. Betrachte die Folge $((-1)^n)_{n \in \mathbb{N}}$.

(a) Betrachte die Teilfolge mit geraden Indizes

$$((-1)^{2n})_{n \in \mathbb{N}} = [(-1)^2]^n_{n \in \mathbb{N}} = (1^n)_{n \in \mathbb{N}} = (1)_{n \in \mathbb{N}}.$$

(b) Betrachte die Teilfolge mit ungeraden Indizes

$$\left((-1)^{2n+1}\right)_{n \in \mathbb{N}} = \left((-1)^{2n} \cdot (-1)\right)_{n \in \mathbb{N}} = \left(1 \cdot (-1)\right)_{n \in \mathbb{N}} = (-1)_{n \in \mathbb{N}}.$$

Mit den Beobachtungen (a) und (b) können wir auch zeigen, dass $\left((-1)^n\right)_{n \in \mathbb{N}}$ divergiert! In der Tat: Wäre $\left((-1)^n\right)_{n \in \mathbb{N}}$ konvergent, so müsste nach Proposition 3.5.13 jede Teilfolge gegen denselben Grenzwert konvergieren.

In Beobachtung (a) haben wir jedoch eine Teilfolge entdeckt, die gegen 1 konvergiert. Beobachtung (b) zeigt hingegen eine Teilfolge auf, die gegen -1 konvergiert. Das widerspricht der Annahme. Somit divergiert $\left((-1)^n\right)_{n \in \mathbb{N}}$.

Dieses Argument ist wesentlich einfacher als in Beispiel 2.3.10. Finden Sie nicht?

Wenn eine Folge $(a_n)_{n \in \mathbb{N}}$ eine Eigenschaft (Konvergenz/Monotonie/Beschränktheit/etc.) nicht hat, bleibt stets noch zu hoffen, dass man eine *Teilfolge* finden kann, die die gesuchte Eigenschaft erfüllt.

Ein Beispiel dafür: Eine monotone Teilfolge finden wir immer!

Proposition 3.5.14 — MONOTONE TEILFOLGEN

Jede Folge $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ besitzt eine monotone Teilfolge.^a

^aMonoton bedeutet monoton wachsend oder monoton fallend. Man kann im Vorhinein nicht sagen, welches von beiden zutreffen wird.

Beweis. Wir nennen einen Index $r \in \mathbb{N}$ in diesem Beweis eine *Spitze*, falls

$$a_m \leq a_r \quad \forall m \geq r + 1. \quad (171)$$

Definiere

$$\mathcal{S} := \{r \in \mathbb{N} : r \text{ ist eine Spitze}\}.$$

Wir unterscheiden nun zwei Fälle.

FALL 1. $\mathcal{S}$ ist eine unendliche Menge.

Sei $\ell_1 \in \mathbb{N}$ eine beliebige Spitze. Wähle ferner für jedes $n \in \mathbb{N}$ ein Element ℓ_{n+1} welches eine Spitze ist und gleichzeitig $\ell_{n+1} \geq \ell_n + 1$ erfüllt. Ein solches Element ℓ_{n+1} existiert stets, denn wären alle Spitzen kleiner als $\ell_n + 1$, so wäre $\mathcal{S} \subseteq \{1, 2, \dots, \ell_n, \ell_n + 1\}$ und somit wäre $\mathcal{S}$ endlich. Ein Widerspruch. Induktiv ist nun also eine Folge $(\ell_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}$ von Spitzen definiert (vgl. Bemerkung 3.3.5). Wegen unserer strikt aufsteigenden Wahl der $(\ell_n)_{n \in \mathbb{N}}$ und Definition 3.5.10 ist $(a_{\ell_n})_{n \in \mathbb{N}}$ eine zulässige Teilfolge. Sei nun $n \in \mathbb{N}$ beliebig. Da ℓ_n eine Spitze ist und $\ell_{n+1} \geq \ell_n + 1$ folgt aus (171), dass

$$a_{\ell_{n+1}} \leq a_{\ell_n}.$$

Mit anderen Worten: $(a_{\ell_n})_{n \in \mathbb{N}}$ ist monoton fallend.

FALL 2. $\mathcal{S}$ ist endlich, etwa $\mathcal{S} = \{s_1, \dots, s_n\}$.

Wähle $k_1 := 1 + \sum_{i=1}^n s_i$. Dann ist $k_1 > s$ für alle $s \in \mathcal{S}$. Also ist k_1 keine Spitze. Verneint man die Aussage in (171) (mit $r = k_1$) so erhält man

$$\exists k_2 \geq k_1 + 1 : \quad a_{k_2} > a_{k_1}.$$

Auch k_2 ist keine Spitze, da $k_2 > k_1 > s$ für alle $s \in \mathcal{S}$. Betrachtet man wieder die Verneinung der Aussage in (171) so erhält man

$$\exists k_3 \geq k_2 + 1 : \quad a_{k_3} > a_{k_2}.$$

Fahren wir sukzessive fort, so erhalten wir eine Folge $(k_n)_{n \in \mathbb{N}}$ natürlicher Zahlen, die *strikt aufsteigend* ist, d.h. $k_{n+1} \geq k_n + 1$. Außerdem gilt

$$a_{k_{n+1}} > a_{k_n} \quad \forall n \in \mathbb{N}.$$

Mit anderen Worten: Die Teilfolge $(a_{k_n})_{n \in \mathbb{N}}$ ist (streng) monoton wachsend.

Wir haben in jedem der beiden Fälle eine monotone Teilfolge entdeckt. Die Behauptung folgt.

01.12.2025

Bemerkung 3.5.15

Im Anbetracht der Technik aus Fall 1 des obigen Beweises können wir folgende, allgemeingültige Erkenntnis festhalten:

Für jede unendliche Menge $A \subseteq \mathbb{N}$ findet man eine Folge $(\ell_n)_{n \in \mathbb{N}} \subseteq A$ mit $\ell_{n+1} \geq \ell_n + 1$.

Sprich: man findet eine *strikt aufsteigende Folge* in einer jeden unendlichen Menge A . Das wird für uns später noch nützlich sein.

Häufungswerte und der Satz von Bolzano-Weierstraß Eine wichtige und für uns sehr zentrale Konsequenz aus der vorigen Proposition 3.5.14 ist der sogenannte *Satz von Bolzano-Weierstraß*.

Satz 3.5.16 — SATZ VON BOLZANO-WEIERSTRASS

Es sei $(a_n)_{n \in \mathbb{N}}$ eine beschränkte Folge. Dann besitzt $(a_n)_{n \in \mathbb{N}}$ (mindestens) eine konvergente Teilfolge.

Beweis. Es sei $(a_n)_{n \in \mathbb{N}}$ eine beschränkte Folge. Nach Proposition 3.5.14 gibt es eine monotone Teilfolge $(a_{\ell_n})_{n \in \mathbb{N}}$. Man überzeugt sich leicht (mithilfe von Definition 2.1.3), dass auch $(a_{\ell_n})_{n \in \mathbb{N}}$ beschränkt ist.

FALL 1. $(a_{\ell_n})_{n \in \mathbb{N}}$ ist monoton wachsend.

Dann ist $(a_{\ell_n})_{n \in \mathbb{N}}$ monoton wachsend und nach oben beschränkt (vgl. Definition 2.1.3). Mit Proposition 3.5.2 folgt, dass $(a_{\ell_n})_{n \in \mathbb{N}}$ konvergiert.

FALL 2. $(a_{\ell_n})_{n \in \mathbb{N}}$ ist monoton fallend.

Dann ist $(a_{\ell_n})_{n \in \mathbb{N}}$ monoton fallend und nach unten beschränkt (vgl. Definition 2.1.3). Mit Proposition 3.5.2 folgt, dass $(a_{\ell_n})_{n \in \mathbb{N}}$ konvergiert.

In allen Fällen haben wir eine konvergente Teilfolge gefunden. Die Behauptung folgt.

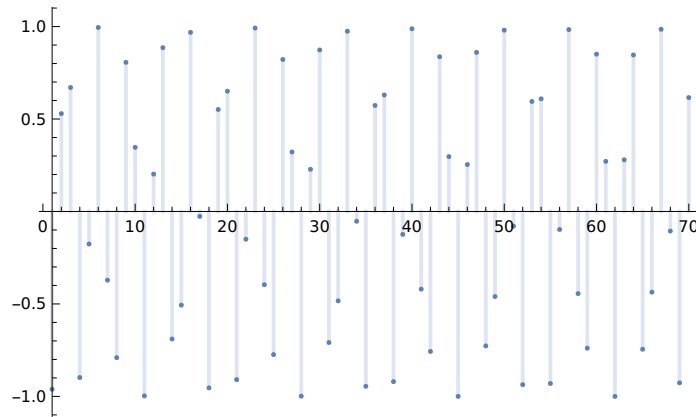


Abbildung 14: Die (beschränkte) Folge $(\sin(17n))_{n \in \mathbb{N}}$. Mit dem bloßen Auge fällt es uns nicht leicht, eine konvergente Teilfolge zu extrahieren. Und doch wissen wir, dass es eine gibt...

Man muss dazu sagen: Eine beschränkte Folge $(a_n)_{n \in \mathbb{N}}$ hat eigentlich i.A. *mehr als nur eine* konvergente Teilfolge. In der Tat: Hat man mal eine konvergente Teilfolge $(a_{\ell_n})_{n \in \mathbb{N}}$ gewählt, so konvergiert ja jede Teilfolge dieser Folge auch! Somit sind z.B. auch die Folgen $(a_{\ell_{2n}})_{n \in \mathbb{N}}$, $(a_{\ell_{3n}})_{n \in \mathbb{N}}$, $(a_{\ell_{4n}})_{n \in \mathbb{N}}$, ... konvergente Teilfolgen!

Grenzwerte konvergenter Teilfolgen nennen wir *Häufungswerte*.

Definition 3.5.17 — HÄUFUNGSWERT

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine Folge. Eine reelle Zahl $a \in \mathbb{R}$ heißt *Häufungswert* von $(a_n)_{n \in \mathbb{N}}$ falls es eine Teilfolge $(a_{\ell_n})_{n \in \mathbb{N}}$ gibt, die gegen a konvergiert.

Beispiel 3.5.18

Die Folge $((-1)^n)_{n \in \mathbb{N}}$ besitzt die Häufungswerte 1 und -1 . Dass 1 und -1 Häufungswerte sind entnehmen wir aus Beispiel 3.5.1 (a),(b). Wir müssten aber eigentlich auch noch zeigen, dass es keine weiteren Häufungswert gibt. Dies überlassen wir den Lesenden als Übungsaufgabe.

Der Satz von Bolzano-Weierstraß (Satz 3.5.16) kann mit Definition 6.1.1 auch wie folgt umformuliert werden: Jede beschränkte Folge besitzt mindestens einen Häufungswert. Für unbeschränkte Folgen ist dies im Allgemeinen nicht der Fall.

Beispiel 3.5.19

Behauptung. Die Folge $(n)_{n \in \mathbb{N}}$ besitzt keine Häufungswerte.

In der Tat: Sei $(\ell_n)_{n \in \mathbb{N}}$ eine Teilfolge von $(n)_{n \in \mathbb{N}}$. Nach Bemerkung 3.5.12 gilt $\ell_n \geq n$ und somit ist $(\ell_n)_{n \in \mathbb{N}}$ unbeschränkt. Dann kann auch keine Konvergenz vorliegen, vgl. Proposition 2.3.14.

Beispiel 3.5.20

Nichtsdestotrotz *kann* eine unbeschränkte Folge durchaus einen Häufungswert haben. Betrachte z.B. die Folge $(c_m)_{m \in \mathbb{N}}$ gegeben durch

$$c_m := \begin{cases} \frac{1}{m} & \text{falls } m \text{ gerade} \\ m & \text{falls } m \text{ ungerade} \end{cases}.$$

Diese Folge ist zwar unbeschränkt, besitzt aber 0 als Häufungswert, da

$$\lim_{n \rightarrow \infty} c_{2n} \stackrel{\text{GWR}}{=} \lim_{n \rightarrow \infty} \frac{1}{2n} \stackrel{\text{GWR}}{=} 0.$$

Beispiel 3.5.21

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine konvergente Folge.

Behauptung. Dann ist $a := \lim_{n \rightarrow \infty} a_n$ der einzige Häufungswert.

Beweis. Es sei $(a_{\ell_n})_{n \in \mathbb{N}}$ eine beliebige Teilfolge von $(a_n)_{n \in \mathbb{N}}$. Nach Proposition 3.5.13 dann $(a_{\ell_n})_{n \in \mathbb{N}}$ auch gegen a konvergieren. Somit kommt kein weiterer Häufungswert in Frage.

³Eine Art *Umkehrung* der obigen Beobachtung gilt auch für beschränkte Folgen: Hat eine beschränkte Folge nur einen Häufungswert, so ist sie konvergent.

Proposition 3.5.22

Besitzt eine beschränkte Folge $(a_n)_{n \in \mathbb{N}}$ nur einen Häufungswert $a \in \mathbb{R}$, so ist sie konvergent (gegen a).

Beweis. Sei $(a_n)_{n \in \mathbb{N}}$ eine beschränkte Folge mit nur einem Häufungswert a . Angenommen $(a_n)_{n \in \mathbb{N}}$ ist nicht konvergent. Dann gilt also (vgl. Bemerkung 2.3.2)

$$\neg(\forall \varepsilon > 0 : (|a_n - a|)_{n \in \mathbb{N}} \leq (\varepsilon)_{n \in \mathbb{N}})$$

Unsere Verneinungsregeln (vgl. Überlegung 2.2.4) ergeben

$$\exists \varepsilon_0 > 0 : \neg((|a_n - a|)_{n \in \mathbb{N}} \leq (\varepsilon_0)_{n \in \mathbb{N}}).$$

Mit Proposition 2.2.8 ergibt sich

$$\exists \varepsilon_0 > 0 : \neg(\{n \in \mathbb{N} : |a_n - a| > \varepsilon_0\} \text{ ist endlich}).$$

Also:

$$\exists \varepsilon_0 > 0 : \{n \in \mathbb{N} : |a_n - a| > \varepsilon_0\} \text{ ist unendlich.}$$

Definiere ab jetzt $A := \{n \in \mathbb{N} : |a_n - a| > \varepsilon_0\}$. Mit Bemerkung 3.5.15 finden wir eine strikt aufsteigende Folge $(\ell_n)_{n \in \mathbb{N}} \subseteq A$. Betrachte nun die Teilfolge $(a_{\ell_n})_{n \in \mathbb{N}}$. Da $\ell_n \in A$ für alle $n \in \mathbb{N}$ ergibt sich

$$|a_{\ell_n} - a| > \varepsilon_0 \quad \forall n \in \mathbb{N}. \quad (172)$$

Beachte auch, dass $(a_{\ell_n})_{n \in \mathbb{N}}$ eine beschränkte Folge ist (weil ja nach Voraussetzung die gesamte Folge beschränkt ist). Nach dem Satz von Bolzano-Weierstraß (vgl. Satz 3.5.16) hat dann $(a_{\ell_n})_{n \in \mathbb{N}}$ eine konvergente Teilfolge $(a_{\ell_{k_n}})_{n \in \mathbb{N}}$. Da $(a_{\ell_{k_n}})_{n \in \mathbb{N}}$ aber nun auch eine Teilfolge von $(a_n)_{n \in \mathbb{N}}$ ist, muss deren Grenzwert auch ein Häufungswert von $(a_n)_{n \in \mathbb{N}}$ sein. Weil jedoch nach der Voraussetzung a der einzige Häufungswert ist, gilt

$$\lim_{n \rightarrow \infty} a_{\ell_{k_n}} = a.$$

Mit den Grenzwertregeln folgt

$$\lim_{n \rightarrow \infty} (a_{\ell_{k_n}} - a) = 0.$$

Daraufhin erhalten wir mit Proposition 2.3.30

$$\lim_{n \rightarrow \infty} |a_{\ell_{k_n}} - a| = 0. \quad (173)$$

³Wer noch mehr Beispiele von Folgen und ihren Häufungswerten sehen will, kann das auf GEOGEBRA tun: [LINK](#)

Auf der anderen Seite ist wegen (172) $|a_{\ell_{k_n}} - a| \geq \varepsilon_0$ für alle $n \in \mathbb{N}$. Mit Grenzwertregel 4 folgt

$$\lim_{n \rightarrow \infty} |a_{\ell_{k_n}} - a| \geq \varepsilon_0.$$

Mit (173) ergibt sich der Widerspruch

$$0 \geq \varepsilon_0 > 0.$$

Auf die Beschränktheit kann in der obigen Proposition nicht verzichtet werden, siehe Beispiel 3.5.20.

Limes Superior und Limes Inferior Bei einer Folge $(a_n)_{n \in \mathbb{N}}$ werden uns zwei Häufungswerte besonders interessieren: Der *größte* und der *kleinste*. Diese werden wir mit *Limes Superior* und *Limes Inferior* bezeichnen.

Wir werden diese beiden Begriffe zunächst anders definieren als die obige Erklärung es suggerieren würde — aus der untigen Definition (Definition 3.5.24) wird z.B. nicht direkt ersichtlich, dass es sich um Häufungswerte handelt. Dennoch wird die in der Definition gegebene *Berechnungsformel* sehr nützlich sein.

Als Vorbereitung der Berechnungsformel benötigen wir die folgende Proposition

Proposition 3.5.23 — EINE HILFSFOLGE

Es sei $(a_n)_{n \in \mathbb{N}}$ eine reelle Folge.

- (i) Sei zusätzlich $(a_n)_{n \in \mathbb{N}}$ nach oben beschränkt. Definiere $(\hat{a}_n)_{n \in \mathbb{N}}$ durch

$$\hat{a}_n := \sup\{a_k : k \geq n\} \quad \forall n \in \mathbb{N}. \quad (174)$$

Dann ist $(\hat{a}_n)_{n \in \mathbb{N}}$ monoton fallend.

- (ii) Sei nun $(a_n)_{n \in \mathbb{N}}$ nach unten beschränkt. Definiere $(\check{a}_n)_{n \in \mathbb{N}}$ durch

$$\check{a}_n := \inf\{a_k : k \geq n\} \quad \forall n \in \mathbb{N}.$$

Dann ist $(\check{a}_n)_{n \in \mathbb{N}}$ monoton wachsend.

Beweis. Wir beweisen nur (i). Für (ii) kann man analog vorgehen.

Zu (i). Dass $(\hat{a}_n)_{n \in \mathbb{N}}$ eine *wohldefinierte*^a reelle Folge ist liegt daran, dass $\{a_k : k \geq n\}$ nach Voraussetzung nach oben beschränkt ist. Deswegen existiert das Supremum in (174).

Nun zur Monotonie. Sei $n \in \mathbb{N}$. Für alle $\ell \geq n + 1$ gilt (weil auch $\ell \geq n$)

$$a_\ell \underset{\text{Def 3.2.10(i.a)}}{\leq} \sup\{a_k : k \geq n\} = \hat{a}_n.$$

Somit ist $\hat{a}_n$ eine obere Schranke von $\{a_\ell : \ell \geq n + 1\}$. Nach Definition 3.2.10(i.b) (mit

$s = \sup\{a_\ell : \ell \geq n + 1\}$ und $s' = \hat{a}_n$) folgt

$$\sup\{a_\ell : \ell \geq n + 1\} \leq \hat{a}_n.$$

Verwenden wir nun, dass nach (174) die rechte Seite der Ungleichung mit $\hat{a}_{n+1}$ übereinstimmt, so erhalten wir

$$\hat{a}_{n+1} \leq \hat{a}_n.$$

^aDas Wort *wohldefiniert* verwendet man, um zu sagen, dass die Definition des betrachteten Objektes keinen mathematischen Widerspruch erzeugt. In unserem konkreten Fall müssen wir überprüfen, ob das Supremum in (174) auch wirklich existiert. Ansonsten hätten wir die Folge $(b_n)_{n \in \mathbb{N}}$ gar nicht definieren dürfen und können dementsprechend auch nicht damit arbeiten.

Mit Bemerkung 3.5.4 wissen wir nun: Die Folge $(\hat{a}_n)_{n \in \mathbb{N}}$ aus der obigen Proposition ist entweder konvergent oder divergiert bestimmt gegen $-\infty$. Dies motiviert die nachfolgende folgende Definition.

Definition 3.5.24 — LIMES SUPERIOR UND LIMES INFERIOR

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine reelle Folge. Dann definieren wir den *Limes Superior* von $(a_n)_{n \in \mathbb{N}}$ durch

$$\limsup_{n \rightarrow \infty} a_n := \begin{cases} \infty & \text{falls } (a_n)_{n \in \mathbb{N}} \text{ nach oben unbeschränkt} \\ \lim_{n \rightarrow \infty} (\sup\{a_k : k \geq n\}) & \text{falls } (\sup\{a_k : k \geq n\})_{n \in \mathbb{N}} \text{ konvergiert} \\ -\infty & \text{falls } (\sup\{a_k : k \geq n\})_{n \in \mathbb{N}} \text{ divergiert} \end{cases}$$

und den *Limes Inferior* von $(a_n)_{n \in \mathbb{N}}$ durch

$$\liminf_{n \rightarrow \infty} a_n := \begin{cases} -\infty & \text{falls } (a_n)_{n \in \mathbb{N}} \text{ nach unten unbeschränkt} \\ \lim_{n \rightarrow \infty} (\inf\{a_k : k \geq n\}) & \text{falls } (\inf\{a_k : k \geq n\})_{n \in \mathbb{N}} \text{ konvergiert} \\ \infty & \text{falls } (\inf\{a_k : k \geq n\})_{n \in \mathbb{N}} \text{ divergiert} \end{cases}$$

Aus Proposition 3.5.23 folgt, dass stets einer der drei Fälle in der obigen Definition eintritt.

Beispiel 3.5.25

Betrachte die Folge $(a_n)_{n \in \mathbb{N}}$ gegeben durch $a_n := (-1)^n(1 + \frac{1}{n})$. Für $n \in \mathbb{N}$ gilt

$$\sup\{a_k : k \geq n\} = \begin{cases} 1 + \frac{1}{n} & \text{falls } n \text{ gerade} \\ 1 + \frac{1}{n+1} & \text{falls } n \text{ ungerade} \end{cases}.$$

Der Beweis hierzu sei Ihnen als Übungsaufgabe überlassen. Wir verweisen auch auf Abbildung 15. Man hat also

$$1 \leq \sup\{a_k : k \geq n\} \leq 1 + \frac{1}{n} \quad \forall n \in \mathbb{N}.$$

Mit dem Sandwichlemma (Proposition 2.3.26) folgt

$$\lim_{n \rightarrow \infty} \sup\{a_k : k \geq n\} = 1$$

und somit (vgl. Definition 3.5.24)

$$\limsup_{n \rightarrow \infty} a_n = 1.$$

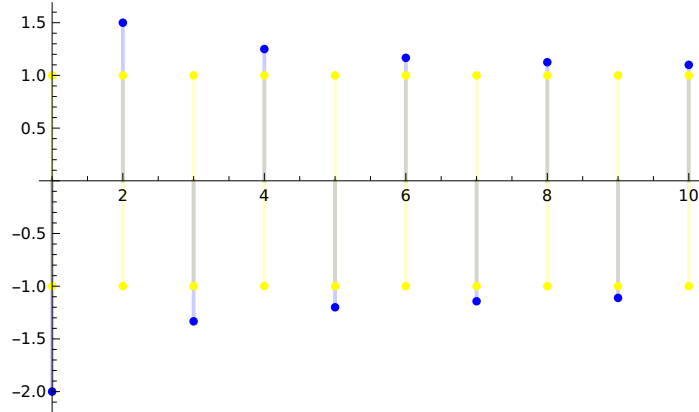


Abbildung 15: Die Folge $((-1)^n(1 + \frac{1}{n}))_{n \in \mathbb{N}}$.

Was aus Definition 3.5.24 nicht klar herauskommt, ist dass $\limsup$ und $\liminf$ (im Falle von beschränkten Folgen) stets Häufungswerte sind. Diese Information macht es aber oft leichter, $\limsup$ und $\liminf$ zu berechnen: Man muss sie nur unter den Häufungswerten suchen. Bevor wir das beweisen können noch eine kurze Vorabbemerkung:

Bemerkung 3.5.26

Bei beschränkten Folgen tritt stets der mittlere Fall in Definition 3.5.24 ein. In der Tat: Da $(a_n)_{n \in \mathbb{N}}$ nach oben beschränkt ist, ist $(\sup\{a_k : k \geq n\})_{n \in \mathbb{N}}$ eine monoton fallende Folge, vgl. Proposition 3.5.23. Da $(a_n)_{n \in \mathbb{N}}$ auch nach unten beschränkt ist, gilt

$$\exists L > 0 : \quad a_n \geq -L \quad \forall n \in \mathbb{N}.$$

Man folgert

$$\sup\{a_k : k \geq n\} \geq -L \quad \forall n \in \mathbb{N}.$$

Somit ist $(\sup\{a_k : k \geq n\})_{n \in \mathbb{N}}$ eine monoton fallende, nach unten beschränkte Folge und konvergiert nach Proposition 3.5.2. Es gilt also (vgl. Definition 3.5.24)

$$\limsup_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \sup\{a_k : k \geq n\}.$$

Proposition 3.5.27

Es sei $(a_n)_{n \in \mathbb{N}}$ eine beschränkte Folge. Dann sind $\limsup_{n \rightarrow \infty} a_n$ und $\liminf_{n \rightarrow \infty} a_n$ Häufungswerte von $(a_n)_{n \in \mathbb{N}}$.

Beweis. Wir zeigen nur, dass $\limsup_{n \rightarrow \infty} a_n$ ein Häufungswert ist. Definiere dazu

$$\hat{a}_n := \sup\{a_k : k \geq n\} \quad \forall n \in \mathbb{N}$$

Dann gilt (vgl. Anmerkung 3.5.26)

$$\limsup_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \hat{a}_n. \quad (175)$$

Fixiere nun ein $n \in \mathbb{N}$. Wendet man Proposition 3.2.26 mit $\varepsilon = \frac{1}{n}$ an, so sieht man, dass ein Element aus $\{a_k : k \geq n\}$ größergleich als $\hat{a}_n - \frac{1}{n}$ sein muss. Nennen wir dieses (von n abhängige) Element a_{k_n} . Mit anderen Worten

$$\exists k_n \in \mathbb{N}, k_n \geq n : \quad \hat{a}_n - \frac{1}{n} \leq a_{k_n}.$$

Da $\hat{a}_n$ auch eine obere Schranke von $\{a_k : k \geq n\}$ ist, gilt auch

$$\hat{a}_n - \frac{1}{n} \leq a_{k_n} \leq \hat{a}_n \quad \forall n \in \mathbb{N}.$$

Da wegen der Grenzwertregeln

$$\lim_{n \rightarrow \infty} \left(\hat{a}_n - \frac{1}{n} \right) = \lim_{n \rightarrow \infty} \hat{a}_n \stackrel{(175)}{=} \limsup_{n \rightarrow \infty} a_n$$

folgt man aus dem Sandwichlemma (Proposition 2.3.26)

$$\lim_{n \rightarrow \infty} a_{k_n} = \lim_{n \rightarrow \infty} \hat{a}_n = \limsup_{n \rightarrow \infty} a_n.$$

Dies ist schon *fast* die Behauptung. Einziges Problem: $(k_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}$ ist nicht zwingend eine strikt aufsteigende Folge! Somit kann man bei $(a_{k_n})_{n \in \mathbb{N}}$ nicht zwingend von einer Teilfolge von $(a_n)_{n \in \mathbb{N}}$ im Sinne von Definition 6.1.1 sprechen. Hierzu machen wir eine Modifikation. Definiere

$$\ell_1 := k_1$$

Beachte: Weil $k_n \geq n$ gilt auch $k_{\ell_1+1} \geq \ell_1 + 1$. Definieren wir

$$\ell_2 := k_{\ell_1+1},$$

so erhalten wir also $\ell_2 \geq \ell_1 + 1$. Wiederum weil $k_n \geq n$ gilt nun $k_{\ell_2+1} \geq \ell_2 + 1$. Definieren wir

$$\ell_3 := k_{\ell_2+1},$$

so gilt $\ell_3 \geq \ell_2 + 1$. Wir können weitermachen und definieren induktiv

$$\ell_{n+1} := k_{\ell_n+1} \quad \forall n \in \mathbb{N}.$$

Dank dieser Konstruktion ist $\ell_{n+1} \geq \ell_n + 1$, d.h. $(a_{\ell_n})_{n \in \mathbb{N}}$ eine Teilfolge von $(a_n)_{n \in \mathbb{N}}$ im Sinne von Definition 6.1.1. Zusätzlich ist aber $(a_{\ell_n})_{n \in \mathbb{N}} = (a_{k_{\ell_{n-1}+1}})_{n \in \mathbb{N}}$ auch eine Teilfolge von $(a_{k_n})_{n \in \mathbb{N}}$. Daher gilt

$$\lim_{n \rightarrow \infty} a_{\ell_n} = \lim_{n \rightarrow \infty} a_{k_n} = \limsup_{n \rightarrow \infty} a_n.$$

Wir haben eine Teilfolge $(a_{\ell_n})_{n \in \mathbb{N}}$ von $(a_n)_{n \in \mathbb{N}}$ gefunden, die gegen $\limsup_{n \rightarrow \infty} a_n$ konvergiert. Das zeigt dass dieser Wert ein Häufungswert ist (vgl. Definition 6.1.1).

Eine wichtige Konsequenz dieses Satzes ist die Folgende: Stellen wir uns vor eine Folge $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ ist konvergent (insbesondere also auch beschränkt, vgl. Proposition 2.3.14. Dann hat nach Beispiel 3.5.21 $(a_n)_{n \in \mathbb{N}}$ nur einen Häufungswert, nämlich $\lim_{n \rightarrow \infty} a_n$. Da aber nach voriger Proposition auch $\limsup_{n \rightarrow \infty} a_n$ und $\liminf_{n \rightarrow \infty} a_n$ Häufungswerte von $(a_n)_{n \in \mathbb{N}}$ sind sehen wir

$$\text{Für eine konvergente Folge } (a_n)_{n \in \mathbb{N}} \text{ gilt } \liminf_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} a_n. \quad (176)$$

Bemerkung 3.5.28

Die obige Proposition liefert einen alternativen Beweis zum Satz von Bolzano-Weierstraß 3.5.16! Denn jede beschränkte Folge besitzt einen Limes Superior (vgl. Bemerkung 3.5.26) und dieser ist nach der obigen Proposition auch ein Häufungswert. Nach Definition 6.1.1 muss es auch eine gegen diesen Häufungswert konvergente Teilfolge geben.

Man hat im Falle eine beschränkten Folge sogar noch mehr Informationen: Der Limes Superior ist stets der *größte Häufungswert* und der Limes Inferior ist stets der *kleinste Häufungswert*.

Proposition 3.5.29 — LIMSUP/LIMINF ALS EXTREMALE HÄUFUNGSWERTE

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine beschränkte Folge. Es bezeichne H die Menge aller Häufungswerte von $(a_n)_{n \in \mathbb{N}}$. Dann gilt^a

$$(i) \quad \limsup_{n \rightarrow \infty} a_n = \max H,$$

$$(ii) \quad \liminf_{n \rightarrow \infty} a_n = \min H.$$

^aHier kann wieder GEOGEBRA beim Verständnis helfen: [LINK](#)

Beweis. Wir zeigen nur (i), denn (ii) geht analog. Nach Definition 3.2.5 muss dafür gezeigt werden, dass

- $\limsup_{n \rightarrow \infty} a_n \in H$ und
- $\limsup_{n \rightarrow \infty} a_n$ ist eine obere Schranke von H

Ersteres folgt sofort aus Proposition 3.5.27. Nun zu Letzterem: Es sei $h \in H$ ein weiterer Häufungswert.

$$\underline{ZL} : \quad h \leq \limsup_{n \rightarrow \infty} a_n.$$

Nach Definition 6.1.1 gibt es eine Teilfolge $(a_{\ell_n})_{n \in \mathbb{N}}$ mit

$$\lim_{n \rightarrow \infty} a_{\ell_n} = h.$$

Da nach Bemerkung 3.5.12 stets $\ell_n \geq n$ gelten muss, gilt auch

$$a_{\ell_n} \leq \sup\{a_k : k \geq n\} \quad \forall n \in \mathbb{N}.$$

Daraus folgt

$$h = \lim_{n \rightarrow \infty} a_{\ell_n} \leq \lim_{n \rightarrow \infty} \sup\{a_k : k \geq n\} \stackrel{\text{Bem 3.5.26}}{=} \limsup_{n \rightarrow \infty} a_n.$$

Wir haben also $h \leq \limsup_{n \rightarrow \infty} a_n$ gezeigt.

Auch die Folgenkonvergenz lässt sich (bei beschränkten Folgen) anhand der Konzepte von Limes Superior und Limes Inferior überprüfen.

Korollar 3.5.30 — KONVERGENZ VIA GLEICHHEIT VON LIMSUP/LIMINF

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine beschränkte Folge. Dann gilt

$$\liminf_{n \rightarrow \infty} a_n \leq \limsup_{n \rightarrow \infty} a_n. \quad (177)$$

Gleichheit in (177) gilt genau dann wenn $(a_n)_{n \in \mathbb{N}}$ konvergent ist. In diesem Fall stimmt $\lim_{n \rightarrow \infty} a_n$ mit den beiden obigen Größen überein.

Beweis. Es sei H wie in Proposition 3.5.29 die Menge aller Häufungswerte von $(a_n)_{n \in \mathbb{N}}$. Die Ungleichung (177) folgt so:

$$\liminf_{n \rightarrow \infty} a_n \stackrel{\text{Prop 3.5.29}}{=} \min H \leq \max H \stackrel{\text{Prop 3.5.29}}{=} \limsup_{n \rightarrow \infty} a_n.$$

Dass aus Konvergenz von $(a_n)_{n \in \mathbb{N}}$ Gleichheit in (177) folgt, haben wir bereits in (176) diskutiert

Noch $\mathbb{Z}$: Gilt Gleichheit in (177) so ist $(a_n)_{n \in \mathbb{N}}$ konvergent.

Gilt Gleichheit in (177) so folgt aus Proposition 3.5.29, dass $\min H = \max H$. Man überlegt sich leicht als Übungsaufgabe, dass H dann einelementig sein muss. Sprich: Die Folge $(a_n)_{n \in \mathbb{N}}$ besitzt nur einen Häufungswert. Dank der Annahme, dass $(a_n)_{n \in \mathbb{N}}$ beschränkt war, impliziert Proposition 3.5.22 nun, dass $(a_n)_{n \in \mathbb{N}}$ gegen diesen Häufungswert konvergiert. Die Behauptung ist gezeigt.

Cauchy-Folgen

Als nächstes lernen wir das *Cauchy-Kriterium* für die Folgenkonvergenz kennen. Es ist äquivalent zur Folgenkonvergenz und somit eine weitere Möglichkeit, festzustellen, ob eine Folge konvergiert. Der große Vorteil: Um das Cauchy-Kriterium zu überprüfen muss man im Vorhinein nicht wissen, was der Grenzwert der Folge sein wird!

Definition 3.5.31 — CAUCHY-FOLGEN

Eine Folge $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ ist eine *Cauchy-Folge*, falls

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \geq n_0 : \quad |a_n - a_m| \leq \varepsilon$$

Grob gesprochen: Eine Folge ist eine Cauchy-Folge, wenn die Folgeglieder untereinander nur beliebig wenig abweichen, wenn nur die Indizes genügend groß sind. Es genügt hierbei aber nicht, dass nur $|a_{n+1} - a_n|$ beliebig klein wird! Stattdessen muss $|a_m - a_n|$ für alle hinreichend großen Indizes m, n beliebig klein sein.⁴

Wir werden nun sehen, dass die Konvergenz einer Folge die Cauchy-Eigenschaft impliziert. Danach werden wir uns auch der logischen Umkehrung dieser Aussage widmen.

Proposition 3.5.32 — KONVERGENTE FOLGEN SIND CAUCHY-FOLGEN

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine konvergente Folge. Dann ist $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ auch eine Cauchy-Folge.

Beweis. Sei $(a_n)_{n \in \mathbb{N}}$ eine konvergente Folge. Etwa sei $a := \lim_{n \rightarrow \infty} a_n$. Es sei $\varepsilon > 0$ beliebig vorgegeben.

$$\mathbb{Z} : \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \geq n_0 : \quad |a_n - a_m| \leq \varepsilon \quad (\text{vgl. Definition 3.5.31}).$$

Wir wissen aus der Definition der Folgenkonvergenz (162) mit $\varepsilon' = \frac{\varepsilon}{2}$, dass

$$\exists n_1 \in \mathbb{N} : \quad |a_k - a| \leq \frac{\varepsilon}{2} \quad \forall k \geq n_1.$$

⁴Wer das Cauchy-Kriterium mal in action sehen möchte, geht am besten auf GEOGEBRA: LINK.

Wähle nun $n_0 := n_1$. Für $n, m \geq n_0$ (d.h. $n, m \geq n_1$) gilt dann

$$|a_n - a_m| = |a_n - a + a - a_m| \leq |a_n - a| + |a - a_m| = \underbrace{|a_n - a|}_{\leq \varepsilon/2 \text{ da } n \geq n_1} + \underbrace{|a_m - a|}_{\leq \varepsilon/2 \text{ da } m \geq n_1} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Wir haben gesehen:

$$\exists n_0 \in \mathbb{N} : \quad |a_n - a_m| \leq \varepsilon \quad \forall n, m \geq n_0.$$

Die Behauptung folgt.

Für den Rest des Abschnitts ist unser Ziel, zu zeigen, dass alle Cauchy-Folgen auch konvergent sind. Ein wichtiger Schritt in diese Richtung ist die *Beschränktheit von Cauchy-Folgen*.

Lemma 3.5.33 — BESCHRÄNKTHEIT VON CAUCHY-FOLGEN

Jede Cauchy-Folge ist beschränkt.

Beweis. Es sei $(a_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge. Das heißt (vgl. Definition 3.5.31)

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \geq n_0 : \quad |a_n - a_m| \leq \varepsilon$$

Betrachte nun $\varepsilon = 1$. Dann gibt es also $n_1 \in \mathbb{N}$ mit

$$\forall n, m \geq n_1 : \quad |a_n - a_m| \leq 1. \quad (178)$$

Nun definiere

$$K := \sum_{k=1}^{n_1} |a_k| + 1. \quad (179)$$

Wir werden zeigen, dass $|a_n| \leq K$ für alle $n \in \mathbb{N}$. Dazu:

FALL 1: $n \leq n_1$. Dann gilt

$$|a_n| \leq \sum_{k=1}^{n_1} |a_k| \stackrel{(179)}{<} K.$$

FALL 2: $n > n_1$. Dann gilt

$$|a_n| = |a_n - a_{n_1} + a_{n_1}| \leq \underbrace{|a_n - a_{n_1}|}_{\leq 1, \text{ wg. (178) mit } m = n_1} + |a_{n_1}| \leq 1 + |a_{n_1}| \leq \sum_{k=1}^{n_1} |a_k| + 1 = K.$$

In allen Fällen gilt also $|a_n| \leq K$. Proposition 2.1.6 liefert die Behauptung.

Mit dieser Vorbereitung können wir nun zeigen, dass Cauchy-Folgen in $\mathbb{R}$ sogar stets konvergent sind.

Proposition 3.5.34 — CAUCHY-FOLGEN IN $\mathbb{R}$ KONVERGIEREN

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine Cauchy-Folge. Dann ist $(a_n)_{n \in \mathbb{N}}$ konvergent.

Beweis. Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine Cauchy-Folge. Nach Lemma 3.5.33 ist $(a_n)_{n \in \mathbb{N}}$ beschränkt. Nach dem Satz von Bolzano-Weierstraß (Satz 3.5.16) gibt es eine konvergente Teilfolge $(a_{\ell_n})_{n \in \mathbb{N}}$. Sei etwa

$$a := \lim_{n \rightarrow \infty} a_{\ell_n}. \quad (180)$$

Dass eine Teilfolge konvergiert heißt natürlich noch nicht, dass auch die gesamte Folge konvergiert. In unserer Situation dürfen wir aber noch die Cauchy-Eigenschaft benutzen, und können daher mehr zeigen.

Behauptung. Die gesamte Folge $(a_n)_{n \in \mathbb{N}}$ konvergiert gegen a .

Sei dafür $\varepsilon > 0$ beliebig vorgegeben.

$$\mathbb{Z}: \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |a_n - a| \leq \varepsilon.$$

Dazu: Wir wissen nach (180)

$$\exists n_1 \in \mathbb{N} \quad \forall n \geq n_1 : \quad |a_{\ell_n} - a| \leq \frac{\varepsilon}{2}. \quad (181)$$

Außerdem gilt wegen der Cauchy-Eigenschaft (Definition 3.5.31)

$$\exists n_2 \in \mathbb{N} \quad \forall n, m \geq n_2 : \quad |a_n - a_m| \leq \frac{\varepsilon}{2}. \quad (182)$$

Wähle nun

$$n_0 := \max\{n_1, n_2\}.$$

Sei jetzt $n \geq n_0$ gegeben. Beachte: Laut Bemerkung 3.5.12 gilt auch $\ell_n \geq n \geq n_0$. Die Dreiecksungleichung liefert

$$|a_n - a| = |a_n - a_{\ell_n} + a_{\ell_n} - a| \leq |a_n - a_{\ell_n}| + |a_{\ell_n} - a|. \quad (183)$$

Da n (und somit auch ℓ_n) größergleich n_2 ist folgt aus (182) (mit $m = \ell_n$), dass $|a_n - a_{\ell_n}| \leq \frac{\varepsilon}{2}$. Ferner folgt aus (181) (und $n \geq n_1$), dass $|a_{\ell_n} - a| \leq \frac{\varepsilon}{2}$. Zusammen mit (183) ergibt sich für alle $n \geq n_0$

$$|a_n - a| \leq \underbrace{|a_n - a_{\ell_n}|}_{\leq \varepsilon/2} + \underbrace{|a_{\ell_n} - a|}_{\leq \varepsilon/2} \leq \varepsilon.$$

Die Behauptung folgt.

Bemerkung 3.5.35

Proposition 3.5.34 und Proposition 3.5.32 ergeben zusammen, dass

$$(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R} \quad \text{ist eine Cauchy-Folge} \quad \Longleftrightarrow \quad (a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R} \quad \text{ist konvergent.}$$

Man nennt diese Eigenschaft auch die *(Norm-)Vollständigkeit* von $\mathbb{R}$.

04.12.2025

4 Die komplexen Zahlen

Wir haben in Abschnitt 3.4 gelernt, dass man im Reellen aus positiven Zahlen eine *Quadratwurzel* ziehen kann. Warum geht das eigentlich nur für positiven Zahlen?

Die Antwort darauf kennen wir: In jedem angeordneten Körper $(K, +, \cdot, \leq)$ gilt (vgl. Proposition 3.2.20 (x))

$$z^2 \geq 0 \quad \forall z \in K.$$

Insbesondere kann die Gleichung

$$z^2 = -1$$

in keinem angeordneten Körper eine Lösung besitzen. Wenn wir jedoch bereit sind, die *Anordnungseigenschaften* aufzugeben, so werden in der Lage sein den Körper $\mathbb{R}$ so zu *vergrößern*, dass auch die Gleichung $z^2 = -1$ eine Lösung besitzt. Diese Vergrößerung werden wir dann als die *komplexen Zahlen* $\mathbb{C}$ bezeichnen.

Von einer *Quadratwurzel von -1* zu sprechen ist allerdings irreführend. Denn die Rechenregeln, die man für eine Quadratwurzel erwarten würde, gelten nicht. Dies zeigt eine verhängnisvolle Beobachtung von Euler.

Warnung 4.0.1 — $\sqrt{-1}$ IST KEINE SINNVOLLE SCHREIBWEISE

Für Lösungen der Gleichung $z^2 = -1$ können wir nicht alle uns bekannten Rechengesetze verwenden, sonst wäre

$$-1 = (\sqrt{-1})^2 = \sqrt{-1}\sqrt{-1} \xrightarrow{\text{erwartbare Rechenregeln für } \sqrt{\cdot}} \sqrt{(-1)(-1)} = \sqrt{1} = 1.$$

4.1 Definition und Schreibweisen

Wir werden die komplexen Zahlen aus einer speziellen Körperstruktur auf $\mathbb{R} \times \mathbb{R}$ herauszaubern. Diese werden wir im Folgenden besprechen.

Proposition 4.1.1 — KÖRPER DER KOMPLEXEN ZAHLEN

Betrachte die Menge $\mathbb{R} \times \mathbb{R}$ mit den wie folgt definierten inneren Verknüpfungen

$$+ : (\mathbb{R} \times \mathbb{R}) \times (\mathbb{R} \times \mathbb{R}) \rightarrow (\mathbb{R} \times \mathbb{R}), \quad (x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2) \quad (184)$$

und

$$\cdot : (\mathbb{R} \times \mathbb{R}) \times (\mathbb{R} \times \mathbb{R}) \rightarrow (\mathbb{R} \times \mathbb{R}), \quad (x_1, y_1) \cdot (x_2, y_2) := (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1). \quad (185)$$

Dann ist $(\mathbb{R} \times \mathbb{R}, +, \cdot)$ ein Körper (vgl. Definition 3.1.4). Ferner gilt

- (i) Das Nullelement ist gegeben durch $(0, 0)$.
- (ii) Das Einselement ist gegeben durch $(1, 0)$.
- (iii) Für ein $(x, y) \in \mathbb{R} \times \mathbb{R}$ ist das additiv inverse Element $-(x, y) = (-x, -y)$.

(iv) Für ein $(x, y) \neq (0, 0)$ ist das multiplikativ inverse Element $(x, y)^{-1} = (\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2})$.

Beweisskizze. Wir werden nicht alle Axiome beweisen (vgl. Definition 3.1.4). Das Körper-Axiom (M4) möchte ich aber doch beweisen, denn die Formel für das multiplikativ Inverse ist lang und überraschend kompliziert. Dazu:

$$\begin{aligned} (x, y) \cdot (\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}) &\stackrel{(185)}{=} (x \frac{x}{x^2+y^2} - y \frac{-y}{x^2+y^2}, x \frac{-y}{x^2+y^2} + y \frac{x}{x^2+y^2}) \\ &= (\frac{x^2}{x^2+y^2} + \frac{y^2}{x^2+y^2}, \frac{-xy}{x^2+y^2} + \frac{xy}{x^2+y^2}) = (\frac{x^2+y^2}{x^2+y^2}, 0) = (1, 0). \end{aligned}$$

Definition 4.1.2 — KOMPLEXE ZAHLEN

Der in der obigen Proposition definierte Körper heißt *Körper der komplexen Zahlen* und wird mit $(\mathbb{C}, +, \cdot)$ bezeichnet.

Bemerkung 4.1.3 — $\mathbb{C}$ ALS ERWEITERUNG VON $\mathbb{R}$

Der oben definierte Körper $(\mathbb{C}, +, \cdot)$ ist im folgenden Sinne eine *Erweiterung* der reellen Zahlen: Für $x \in \mathbb{R}$ kann man das Element $(x, 0) \in \mathbb{R} \times \mathbb{R}$ betrachten. Mit (184) und (185) überzeugt man sich leicht, dass für $x_1, x_2 \in \mathbb{R}$ gilt

$$(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0)$$

und

$$(x_1, 0) \cdot (x_2, 0) = (x_1 x_2, 0).$$

Mit anderen Worten: Für das Rechnen in $\{(x, 0) : x \in \mathbb{R}\}$ gelten dieselben Additions- und Multiplikationsregeln wie für das Rechnen in $\mathbb{R}$. Auf diese Weise lassen sich die reellen Zahlen als *Teilkörper* in den komplexen Zahlen wiederfinden. Man würde auch sagen: $\mathbb{C}$ ist eine *Körpererweiterung* von $\mathbb{R}$. Mehr dazu wird sicherlich in einer Algebra-Vorlesung besprochen.

Eine spezielle Rolle im oben definierten Körper spielt die sogenannte *imaginäre Zahl* $i \in \mathbb{C}$, welche wir im Folgenden definieren

Definition 4.1.4 — IMAGINÄRE EINHEIT

Wir definieren die *imaginäre Einheit* i durch

$$i := (0, 1) \in \mathbb{C}.$$

Warum diese Zahl eine Sonderrolle einnimmt, ergibt sich, wenn wir $i^2 := i \cdot i$ berechnen.

Proposition 4.1.5

Es gilt

$$i^2 = -1_{\mathbb{C}}.$$

Beweis.

$$i \cdot i = (0, 1) \cdot (0, 1) \underset{(185)}{=} (0 \cdot 0 - 1 \cdot 1, 0 \cdot 1 + 1 \cdot 0) = (-1, 0) \underset{\text{Prop 4.1.1}}{=} -(1, 0) \underset{\text{Prop 4.1.1}}{=} -1_{\mathbb{C}}.$$

Notation 4.1.6

Im Folgenden werden wir für $x \in \mathbb{R}$ anstatt $(x, 0)$ nur noch x schreiben. Auf diese Weise wird $\mathbb{R}$ mit $\{(x, 0) : x \in \mathbb{R}\}$ *identifiziert*. Dies ist zulässig, da laut Bemerkung 4.1.3 in beiden Mengen dieselben Rechengesetze gelten!

Insbesondere wird nun auch $1 \in \mathbb{R}$ mit $(1, 0) = 1_{\mathbb{C}}$ identifiziert. Diese Identifikation erlaubt uns, die Aussage von Proposition 4.1.5 in Form von

$$i^2 = -1$$

umzuschreiben.

Proposition 4.1.7

Für alle $x, y \in \mathbb{R}$ gilt

$$(x, y) = (x, 0) + i \cdot (y, 0). \quad (186)$$

Mit unserer oben beschriebenen Identifikation bedeutet das also

$$(x, y) = x + i \cdot y. \quad (187)$$

Beweis. Wir rechnen (186) *rückwärts gelesen* nach.

$$\begin{aligned} (x, 0) + i \cdot (y, 0) &= (x, 0) + (0, 1) \cdot (y, 0) \\ &\underset{(185)}{=} (x, 0) + (0 \cdot y - 1 \cdot 0, 0 \cdot 0 + 1 \cdot y) = (x, 0) + (0, y) \underset{(184)}{=} (x, y). \end{aligned}$$

Wir haben in (187) gesehen, dass jedes Element von $\mathbb{C}$ eine Darstellung in Form von $x + i \cdot y$ für $x, y \in \mathbb{R}$ besitzt. Wir nennen diese Darstellung im Folgenden *Normalform* einer komplexen Zahl. Den Malpunkt lassen wir im Allgemeinen weg und schreiben

$$\mathbb{C} = \{x + iy : x, y \in \mathbb{R}\} \quad (188)$$

Diese Darstellung von $\mathbb{C}$ ist besonders einfach zum Rechnen: Man kann im Wesentlichen mit denselben Körperaxiomen rechnen wie in $\mathbb{R}$ nur mit der Zusatzinfo, dass $i^2 = -1$.

Bemerkung 4.1.8

Wer die Darstellung in (188) verinnerlicht hat muss sich auch das komplizierte Multiplikationsgesetz aus (185) nicht merken. Stattdessen kann man einfach drauflos-multiplizieren wie in $\mathbb{R}$ gewohnt:

$$\begin{aligned}(x_1 + iy_1) \cdot (x_2 + iy_2) &= x_1(x_2 + iy_2) + iy_1(x_2 + iy_2) = x_1x_2 + ix_1y_2 + iy_1x_2 + i^2y_1y_2 \\ &= x_1x_2 + ix_1y_2 + iy_1x_2 + (-1)y_1y_2 \\ &= x_1x_2 - y_1y_2 + i(x_1y_2 + y_1x_2).\end{aligned}\tag{189}$$

Übersetzen wir nun zwischen den Darstellungen gemäß (187) so ergibt sich

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2)$$

Genau wie in (185)!

Ab jetzt wird die Normalform (188) also unsere präferierte Darstellung sein.

Definition 4.1.9 — REALTEIL UND IMAGINÄRTEIL

Es sei $z \in \mathbb{C}$ eine beliebige komplexe Zahl, etwa $z = x + iy$ mit $x, y \in \mathbb{R}$. Dann heißt $x \in \mathbb{R}$ *Realteil von z* , im Zeichen $x = \operatorname{Re}(z)$. Ferner heißt $y \in \mathbb{R}$ *Imaginärteil von z* , im Zeichen $y = \operatorname{Im}(z)$.

Eine Zahl $z \in \mathbb{C}$ heißt *reell* wenn $\operatorname{Im}(z) = 0$ und *rein imaginär* wenn $\operatorname{Re}(z) = 0$.

Achtung: Der Imaginärteil ist (genauso wie der Realteil) stets eine *reelle* Zahl! Man beachte: Nach (187) sind zwei komplexe Zahlen z und z' genau dann gleich wenn $\operatorname{Re}(z) = \operatorname{Re}(z')$ und $\operatorname{Im}(z) = \operatorname{Im}(z')$.

Proposition 4.1.10 — EIGENSCHAFTEN DES REAL- UND IMAGINÄRTEILS

Es seien $z, w \in \mathbb{C}$ sowie $\lambda \in \mathbb{R}$. Dann gilt

- (i) $\operatorname{Re}(z + w) = \operatorname{Re}(z) + \operatorname{Re}(w)$ und $\operatorname{Im}(z + w) = \operatorname{Im}(z) + \operatorname{Im}(w)$.
- (ii) $\operatorname{Re}(\lambda z) = \lambda \operatorname{Re}(z)$ und $\operatorname{Im}(\lambda z) = \lambda \operatorname{Im}(z)$.
- (iii) $\operatorname{Re}(iz) = -\operatorname{Im}(z)$ und $\operatorname{Im}(iz) = \operatorname{Re}(z)$.

Beweis. Wir beweisen nur $\operatorname{Re}(z + w) = \operatorname{Re}(z) + \operatorname{Re}(w)$. Der Rest ist den Lesenden als Übungsaufgabe überlassen. Sei etwa $z = x + iy$ für $x, y \in \mathbb{R}$ und $w = x' + iy'$ für $x', y' \in \mathbb{R}$. Dann gilt

$$\operatorname{Re}(z) = x \quad \text{und} \quad \operatorname{Re}(w) = x'.\tag{190}$$

Nun gilt

$$\operatorname{Re}(z + w) = \operatorname{Re}((x + iy) + (x' + iy')) = \operatorname{Re}((x + x') + i(y + y')) = x + x' \stackrel{(190)}{=} \operatorname{Re}(z) + \operatorname{Re}(w).$$

Im folgenden Beispiel wollen wir üben, komplexe Zahlen auf die Gestalt in (188) (d.h. die *Normalform*) zu bringen, das heißt wir wollen insbesondere Realteil und Imaginärteil bestimmen.

Beispiel 4.1.3. Wir möchten für die Zahl $z = \frac{2+3i}{3+2i}$ eine Darstellung in *Normalform*, d.h. wie in (188). Das heißt wir wollen $\operatorname{Re}(z)$ und $\operatorname{Im}(z)$ berechnen. Hierzu

$$\begin{aligned} \frac{2+3i}{3+2i} &= (2+3i)(3+2i)^{-1} \stackrel{\text{Prop 4.1.1(iv)}}{=} (2+3i) \left(\frac{3}{3^2+2^2} + i \frac{-2}{3^2+2^2} \right) \\ &= (2+3i) \left(\frac{3}{13} + i \frac{-2}{13} \right) \stackrel{(189)}{=} \frac{6}{13} + \frac{6}{13}i + i \left(\frac{9}{13} - \frac{4}{13} \right) = \frac{12}{13} + \frac{5}{13}i. \end{aligned}$$

Es gilt also $\operatorname{Re}(z) = \frac{12}{13}$ und $\operatorname{Im}(z) = \frac{5}{13}$.

Für das Rechnen im Körper der komplexen Zahlen treffen wir ab sofort dieselben Konventionen wie im Reellen auch. Beispielsweise gilt für $z, w \in \mathbb{C}, w \neq 0$, $\frac{z}{w} := z \cdot w^{-1}$. Auch können wir für $z \in \mathbb{C}$ und $n \in \mathbb{N}$ die *Potenzen* z^n wie gewohnt induktiv definieren (vgl. Definition 3.3.8). Desweiteren verwenden wir das *Summenzeichen* $\sum_{k=1}^n$ und das *Produktzeichen* $\prod_{k=1}^n$ wie im Reellen. Man beachte: In der Definition des Summen- und Produktzeichens haben wir eigentlich vorausgesetzt, dass die Summanden in einem angeordneten(!) Körper leben müssen (was bei $\mathbb{C}$ nicht der Fall wäre). Die Anordnung wird aber in der Realität nicht benötigt: Es wird lediglich benötigt, dass der Körper, in dem wir arbeiten, eine Realisierung der natürlichen Zahlen besitzt. Da letzteres auch auf $\mathbb{C}$ zutrifft, ist es legitim, mit Summen- und Produktzeichen zu arbeiten. Die geometrische Summenformel (vgl. Proposition 3.3.20) überträgt sich dann übrigens auch auf den Fall $q \in \mathbb{C}, q \neq 1$. Ferner kann auch der binomische Lehrsatz verwendet werden.

4.2 Konjugation und Absolutbetrag

Im Folgenden sprechen wir über *Rechen-Operationen* auf komplexen Zahlen.

Komplexe Konjugation

Eine erste wichtige Operation auf den komplexen Zahlen ist die komplexe *Konjugation*.

Definition 4.2.1 — KOMPLEXE KONJUGATION

Es sei $z \in \mathbb{C}$, etwa $z = x + iy$ (mit $x, y \in \mathbb{R}$). Dann heißt

$$\bar{z} := x - iy$$

das *Konjugiertkomplexe* von z . (Sprich „ z konjugiert“).

Mit anderen Worten: Es gilt $\operatorname{Re}(\bar{z}) = \operatorname{Re}(z)$ und $\operatorname{Im}(\bar{z}) = -\operatorname{Im}(z)$.

Eine besonders nützliche Angelegenheit an der komplexen Konjugation ist, dass sich $\operatorname{Re}(z)$ und $\operatorname{Im}(z)$ stets aus z und $\bar{z}$ errechnen lassen. Die Operation des Konjugierens erlaubt uns daher, Ausdrücke *innerhalb der Welt der komplexen Zahlen* zu beschreiben (anstatt immer über Realteil und Imaginärteil sprechen zu müssen).

Proposition 4.2.2

Es sei $z \in \mathbb{C}$. Dann gilt

$$(i) \operatorname{Re}(z) = \frac{z + \bar{z}}{2},$$

$$(ii) \operatorname{Im}(z) = \frac{z - \bar{z}}{2i}.$$

Beweis. Zu (i). Es sei $z \in \mathbb{C}$ etwa

$$z = x + iy \quad \text{für ein } x, y \in \mathbb{R}.$$

Nun rechnen wir (i) rückwärts gelesen nach. Dazu:

$$\frac{z + \bar{z}}{2} = \frac{(x + iy) + (x - iy)}{2} = \frac{2x}{2} = x = \operatorname{Re}(z).$$

Für (ii) gehen wir analog vor. Dazu:

$$\frac{z - \bar{z}}{2i} = \frac{(x + iy) - (x - iy)}{2} = \frac{2iy}{2i} = y = \operatorname{Im}(z).$$

Für die komplexe Konjugation gibt auch es einige nützliche Rechenregeln

Proposition 4.2.3 — RECHENREGELN FÜR DIE KONJUGATION

Es seien $z, w \in \mathbb{C}$. Dann gilt

$$(i) \bar{\bar{z}} = z.$$

$$(ii) \overline{z + w} = \bar{z} + \bar{w}.$$

$$(iii) \overline{z \cdot w} = \bar{z} \cdot \bar{w}$$

$$(iv) \text{ Falls } w \neq 0 \text{ so gilt } \overline{\left(\frac{z}{w}\right)} = \frac{\bar{z}}{\bar{w}}.$$

Beweis. Zu (i). Es sei $z \in \mathbb{C}$ beliebig, etwa $z = x + iy$ für $x, y \in \mathbb{R}$. Dann gilt

$$\bar{\bar{z}} \stackrel{\text{Def4.2.1}}{=} \overline{x - iy} = \overline{x + i(-y)} \stackrel{\text{Def4.2.1}}{=} x - i(-y) = x + iy.$$

Zu (ii). Es seien $z, w \in \mathbb{C}$ beliebig, etwa $z = x + iy$ und $w = x' + iy'$ für $x, x', y, y' \in \mathbb{R}$. Dann gilt

$$\begin{aligned}\overline{z + w} &= \overline{(x + iy) + (x' + iy')} \\ &= \overline{(x + x') + i(y + y')} \stackrel{\text{Def4.2.1}}{=} (x + x') - i(y + y') \\ &= (x - iy) + (x' - iy') \stackrel{\text{Def4.2.1}}{=} \bar{z} + \bar{w}.\end{aligned}$$

Zu (iii). Es seien $z, w \in \mathbb{C}$ beliebig, etwa $z = x + iy$ und $w = x' + iy'$ für $x, x', y, y' \in \mathbb{R}$. Dann gilt

$$\begin{aligned}\overline{z \cdot w} &\stackrel{\text{Def4.2.1}}{=} \overline{(x + iy)(x' + iy')} \\ &\stackrel{(189)}{=} \overline{(xx' - yy') + i(xy' + yx')} \\ &\stackrel{\text{Def4.2.1}}{=} (xx' - yy') - i(xy' + yx') \\ &\stackrel{\text{Def4.2.1}}{=} (xx' - (-y)(-y')) + i(x(-y') + (-y)x') \\ &\stackrel{(189)}{=} (x + i(-y))(x' + i(-y')) \\ &\stackrel{\text{Def4.2.1}}{=} (x - iy)(x' - iy') = \bar{z} \cdot \bar{w}.\end{aligned}$$

Die letzte Anwendung von (189) würde ich wiederum rückwärts lesen.

Zu (iv). Es sei $w \neq 0$. Dann gilt klarerweise auch $\bar{w} \neq 0$ (vgl. Definition 4.2.1). Man erhält

$$\bar{w} \cdot \overline{\left(\frac{z}{w}\right)} \stackrel{\text{(iii)}}{=} \overline{w \cdot \left(\frac{z}{w}\right)} = \overline{\frac{zw}{w}} = \bar{z}.$$

Teilt man beide Seiten dieser Gleichung durch $\bar{w}$ so folgt

$$\overline{\left(\frac{z}{w}\right)} = \frac{\bar{z}}{\bar{w}}.$$

Komplexer Absolutbetrag

Während wir die Reellen Zahlen $\mathbb{R}$ mithilfe einer *Zahlengerade* visualisieren, lassen sich die komplexen Zahlen $\mathbb{C}$ durch eine *komplexe Zahlenebene* $\mathbb{R} \times \mathbb{R}$ (auch genannt $\mathbb{R}^2$) visualisieren, vgl. Proposition 4.1.1. Eine wichtige Größe in der Ebene ist die *euklidische Norm*, welche üblicherweise den Abstand zwischen einem Element $(x, y) \in \mathbb{R} \times \mathbb{R}$ und dem *Koordinatenursprung* $(0, 0)$ misst.

In der Welt der komplexen Zahlen nennen wir diese Größe den *komplexen Absolutbetrag*.

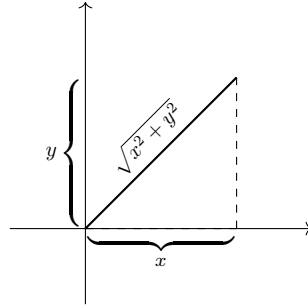


Abbildung 16: Die Euklidische Norm

Definition 4.2.4 — KOMPLEXER ABSOLUTBETRAG

Es sei $z \in \mathbb{C}$, etwa $z = x + iy$ für $x, y \in \mathbb{R}$. Dann nennen wir die reelle(!) Zahl

$$|z| := \sqrt{x^2 + y^2} \in \mathbb{R}$$

den *komplexen Absolutbetrag* von z . Man beachte, dass die Wurzel $\sqrt{\cdot}$ hier nur von einer reellen Zahl gezogen wird (und deswegen solche Probleme wie in Warnung 4.0.1 besprochen bei obiger Definition nicht auftreten).

Mit anderen Worten gilt

$$|z| = \sqrt{\operatorname{Re}(z)^2 + \operatorname{Im}(z)^2} \quad \forall z \in \mathbb{C}.$$

Wir haben nicht nur den komplexen, sondern auch den *reellen* Absolutbetrag mit $|\cdot|$ notiert. Damit unsere Notation widerspruchsfrei bleibt, sollten wir uns von der *Konsistenz* der Notationen überzeugen. Dazu: Für $x \in \mathbb{R}$ ist

$$|x + i \cdot 0| = \sqrt{x^2 + 0^2} = \sqrt{x^2} = |x|.$$

Hierbei ist auf der linken Seite der komplexe Absolutbetrag und auf der rechten Seite der reelle Absolutbetrag gemeint. Wir haben also gezeigt, dass die beiden Absolutbeträge auf $\mathbb{R}$ miteinander übereinstimmen, was unsere Notation rechtfertigt.

Bemerkung 4.2.5

Man folgert auch leicht, dass für alle $z \in \mathbb{C}$ gilt, dass

$$|\operatorname{Re}(z)| \leq |z|$$

und

$$|\operatorname{Im}(z)| \leq |z|.$$

Die nachfolgende Proposition verknüpft die komplexe Konjugation und den Absolutbetrag.

Proposition 4.2.6 — KONJUGATION UND ABSOLUTBETRAG

Es sei $z \in \mathbb{C}$. Dann gilt

- (i) $|z| = |\bar{z}|$,
- (ii) $z\bar{z} = |z|^2$.

Beweis. Behauptung (i) ist den Lesenden als Übungsaufgabe überlassen. Zu (ii). Es sei etwa $z = x + iy$ für $x, y \in \mathbb{R}$. Dann gilt

$$z\bar{z} = (x + iy)(x - iy) \underset{(189)}{=} x^2 + y^2 + i(xy - xy) = x^2 + y^2 = \left(\sqrt{x^2 + y^2}\right)^2 = |z|^2.$$

Bemerkung 4.2.7

Sei $z \in \mathbb{C} \setminus \{0\}$. Die obige Proposition und die Tatsache, dass $|z|^2 \in \mathbb{R}$, liefert eine gute Möglichkeit, um den Kehrwert $\frac{1}{z}$ in Normalform auszudrücken. Dazu beobachte

$$\frac{1}{z} = \frac{\bar{z}}{z\bar{z}} \underset{\text{Prop4.2.6(ii)}}{=} \frac{\bar{z}}{|z|^2}.$$

Nun sei $z = x + iy$ für $x, y \in \mathbb{R}$. Dann hat man mit der vorigen Gleichung

$$\frac{1}{z} = \frac{\bar{z}}{|z|^2} \underset{\text{Def4.2.4}}{\overset{\text{Def4.2.1}}{=}} \frac{x - iy}{x^2 + y^2} = \frac{x}{x^2 + y^2} + i \frac{-y}{x^2 + y^2}.$$

Dies deckt sich mit der Regel zur Kehrwertbildung in Proposition 4.1.1 (iv). Um den Kehrwert einer komplexen Zahl zu bilden, muss man daher die Formel in Proposition 4.1.1 (iv) nicht auswendig lernen! Stattdessen kann man auch einfach wie in der obigen Rechnung *mit dem Konjugiertkomplexen erweitern*.

Dazu ein Zahlenbeispiel:

$$\frac{1}{3 + 2i} = \frac{3 - 2i}{(3 + 2i)(3 - 2i)} \underset{\text{Prop4.2.6}}{=} \frac{3 - 2i}{3^2 + 2^2} = \frac{3}{13} - \frac{2}{13}i.$$

Eine weitere wichtige Formel ist die sogenannte *Parallelogrammidentität*, vgl. Abbildung 17.

Proposition 4.2.8

Es seien $z, w \in \mathbb{C}$. Dann gilt

- (i) $|z + w|^2 = |z|^2 + 2\operatorname{Re}(z\bar{w}) + |w|^2$
- (ii) (*Parallelogrammidentität.*) $|z + w|^2 + |z - w|^2 = 2(|z|^2 + |w|^2)$

Beweis. Seien $z, w \in \mathbb{C}$. Zu (i).

$$\begin{aligned}
 |z + w|^2 &= (z + w)\overline{(z + w)} \\
 &= (z + w)(\bar{z} + \bar{w}) \\
 &\stackrel{(D)}{=} z\bar{z} + z\bar{w} + w\bar{z} + w\bar{w} \\
 &= |z|^2 + |w|^2 + z\bar{w} + \bar{z}w \\
 &\stackrel{\text{Prop4.2.3(i)}}{=} |z|^2 + |w|^2 + z\bar{w} + \overline{z\bar{w}} \\
 &\stackrel{\text{Prop4.2.3(iii)}}{=} |z|^2 + |w|^2 + z\bar{w} + \overline{z\bar{w}} \\
 &= |z|^2 + |w|^2 + 2\frac{z\bar{w} + \overline{z\bar{w}}}{2} \\
 &\stackrel{\text{Prop4.2.2}}{=} |z|^2 + |w|^2 + 2\text{Re}(z\bar{w}).
 \end{aligned}$$

Zu (ii).

$$\begin{aligned}
 |z + w|^2 + |z - w|^2 &= |z + w|^2 + |z + (-w)|^2 \\
 &\stackrel{(i)}{=} |z|^2 + 2\text{Re}(z\bar{w}) + |w|^2 + [|z|^2 + 2\text{Re}(z\overline{(-w)})] + \underbrace{|(-w)|^2}_{=|w|^2} \\
 &= 2(|z|^2 + |w|^2) + 2[\text{Re}(z\bar{w}) + \text{Re}(z\overline{(-w)})] \\
 &\stackrel{\text{Prop4.1.2(i)}}{=} 2(|z|^2 + |w|^2) + 2\text{Re}(z\bar{w} + z\overline{(-w)}) \\
 &= 2(|z|^2 + |w|^2) + 2\text{Re}(z(\bar{w} + \overline{(-w)})) \\
 &\stackrel{\text{Prop4.2.3(ii)}}{=} 2(|z|^2 + |w|^2) + 2\text{Re}(z(\bar{w} + \overline{(-w)})) \\
 &= 2(|z|^2 + |w|^2) + \underbrace{2\text{Re}(z \cdot \bar{0})}_{=2\text{Re}(0)=0} = 2(|z|^2 + |w|^2).
 \end{aligned}$$

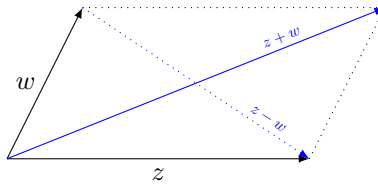


Abbildung 17: Die Größen, die in der Parallelogrammidentität auftauchen sind z, w sowie die Parallelogrammdiagonalen $z + w, z - w$ (vorläufige Abbildung)

Der komplexe Absolutbetrag genießt abgesehen davon noch ähnliche Eigenschaften wie der reelle Absolutbetrag. Der Beweis ist hierbei aber völlig anders.

Satz 4.2.9 — EIGENSCHAFTEN DES KOMPLEXEN ABSOLUTBETRAGES

- (i) *Definitheit.* Für alle $z \in \mathbb{C}$ gilt $|z| \geq 0$ und $(|z| = 0 \Leftrightarrow z = 0)$.
- (ii) *Multiplikativität.* Für alle $z, w \in \mathbb{C}$ gilt $|z \cdot w| = |z| \cdot |w|$.
- (iii) *Dreiecksungleichung.* Für alle $z, w \in \mathbb{C}$ gilt $|z + w| \leq |z| + |w|$.

Beweis. Behauptung (i) ist den Lesenden als Übungsaufgabe überlassen. Zu (ii). Es seien $z, w \in \mathbb{C}$. Dann gilt

$$|z \cdot w|^2 \stackrel{\text{Prop 4.2.6}}{=} z \cdot w \cdot \overline{z \cdot w} \stackrel{\text{Prop 4.2.3}}{=} z \cdot w \cdot \bar{z} \cdot \bar{w} = z \cdot \bar{z} \cdot w \cdot \bar{w} \stackrel{\text{Prop 4.2.6}}{=} |z|^2 \cdot |w|^2 = (|z| \cdot |w|)^2.$$

Wir können nun auf beiden Seiten die Wurzel ziehen und erhalten (da sowohl $|z \cdot w| \geq 0$ als auch $|z| \cdot |w| \geq 0$), dass

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

Zu (iii). Es gilt

$$\begin{aligned} |z + w|^2 &\stackrel{\text{Prop 4.2.8(i)}}{=} |z|^2 + |w|^2 + 2\operatorname{Re}(z\bar{w}) \\ &\leq |z|^2 + |w|^2 + 2|\operatorname{Re}(z\bar{w})| \\ &\leq |z|^2 + |w|^2 + 2|z\bar{w}| \\ &\stackrel{(ii)}{=} |z|^2 + |w|^2 + 2|z| \cdot |w| \\ &\stackrel{1. \text{ Binom. Formel}}{=} (|z| + |w|)^2. \end{aligned}$$

Durch Wurzelziehen (und aus der Monotonie der Wurzel, vgl. Blatt 9 Aufgabe 1d) folgt

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

Beispiel 4.2.10

Es gilt $|i| = 1$. Man hat zwei Möglichkeiten, diesen Sachverhalt herzuleiten. Eine Alternative ist direkt mit der Definition des Betrages, nämlich gilt

$$|i| = |0 + 1i| = \sqrt{0^2 + 1^2} = 1.$$

Alternativ kann man diese Formel auch aus der Multiplikativität folgern:

$$|i|^2 = |i| \cdot |i| \stackrel{\text{Prop 4.2.9(ii)}}{=} |i \cdot i| = |i^2| = |-1| = 1.$$

Nun können wir (in $\mathbb{R}$) die Wurzel ziehen. Da $|i| \geq 0$ erhalten wir

$$|i| = \sqrt{1} = 1.$$

Die Folgenden weiteren Eigenschaften des Betrags folgen direkt aus dem obigen Satz, und zwar genau wie im Reellen.

Proposition 4.2.11 — WEITERE EIGENSCHAFTS DES KOMPLEXEN ABSOLUTBETRAGES

(i) Es gilt $|(-z)| = |z|$ für alle $z \in \mathbb{C}$.

(ii) Für alle $z, w \in \mathbb{C}$ mit $w \neq 0$ gilt $|\frac{z}{w}| = \frac{|z|}{|w|}$.

(iii) Für alle $z, w \in \mathbb{C}$ gilt

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

(iv) Für alle $z, w \in \mathbb{C}$ gilt

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

(Für (iii) und (iv) sollte man dazu sagen, dass mit den äußeren Betragsstrichen auf der linken Seite ein reeller Absolutbetrag und mit den inneren Betragsstrichen ein komplexer Absolutbetrag gemeint ist)

Beweis. Folgt 1zu1 den Argumenten im Reellen, vgl. Proposition 2.0.5.

Beispiel 4.2.12

Wir möchten nun für die Zahl $z = \frac{2+3i}{3+2i}$ aus Beispiel 4.1.3 den Absolutbetrag berechnen. Hierzu gibt es zwei Möglichkeiten. Die erste ist häufig viel mühseliger.

MÖGLICHKEIT 1. Über die Normalform. Wir haben (mit etwas Mühe) in Beispiel 4.1.3 gesehen, dass

$$\frac{2+3i}{3+2i} = \frac{12}{13} + \frac{5}{13}i$$

Dann gilt

$$\left| \frac{2+3i}{3+2i} \right| = \left| \frac{12}{13} + \frac{5}{13}i \right| \stackrel{\text{Def 4.2.4}}{=} \sqrt{\left(\frac{12}{13}\right)^2 + \left(\frac{5}{13}\right)^2} = \sqrt{\frac{144}{169} + \frac{25}{169}} = \sqrt{\frac{169}{169}} = 1.$$

MÖGLICHKEIT 2. Mit den gelernten Rechenregeln.

$$\left| \frac{2+3i}{3+2i} \right| \stackrel{\text{Prop 4.2.11(ii)}}{=} \frac{|2+3i|}{|3+2i|} = \frac{\sqrt{2^2+3^2}}{\sqrt{3^2+2^2}} = 1.$$

Zum Abschluss noch eine kleine Anmerkung zu Satz 4.2.9: Die Dreiecksungleichung lässt sich übrigens auch mit beliebig vielen Summanden *iteriert anwenden*.

Proposition 4.2.13 — ITERIERTE DREIECKSUNGLEICHUNG

Seien für alle $k \in \mathbb{Z}$ Elemente $a_k \in \mathbb{C}$ gegeben, sowie $m, n \in \mathbb{Z}$ mit $m \leq n$. Dann gilt

$$\left| \sum_{k=m}^n a_k \right| \leq \sum_{k=m}^n |a_k|.$$

Beweis. Ohne Beschränkung der Allgemeinheit können wir annehmen, dass $m = 1$. Wir zeigen dann die Behauptung mit vollständiger Induktion.

IndA: $n = 1$.

Es gilt

$$\left| \sum_{k=1}^1 a_k \right| = |a_1| = \sum_{k=1}^1 |a_k|.$$

Da Gleichheit herrscht, gilt insbesondere „ $\leq$ “.

IndS: $n \longrightarrow n+1$.

Es gelte für ein $n \in \mathbb{N}$ die Induktionshypothese

$$\left| \sum_{k=1}^n a_k \right| \leq \sum_{k=1}^n |a_k|. \quad (\text{IndH})$$

$$\mathbb{Z} : \quad \text{Dann gilt auch} \quad \left| \sum_{k=1}^{n+1} a_k \right| \leq \sum_{k=1}^{n+1} |a_k|.$$

Dazu:

$$\begin{aligned} \left| \sum_{k=1}^{n+1} a_k \right| &= \left| \sum_{k=1}^n a_k + a_{n+1} \right| \stackrel{\text{Satz 4.2.9(iii)}}{\leq} \left| \sum_{k=1}^n a_k \right| + |a_{n+1}| \\ &\stackrel{\text{IndH}}{\leq} \sum_{k=1}^n |a_k| + |a_{n+1}| = \sum_{k=1}^{n+1} |a_k|. \end{aligned}$$

4.3 Folgen in $\mathbb{C}$

Nun möchten wir uns *Folgen* komplexer Zahlen zuwenden. Genauer möchten wir sagen, was *Konvergenz* von Folgen komplexer Zahlen bedeutet. Hierfür *kopieren* wir im Wesentlichen die Konvergenzdefinition für Folgen in $\mathbb{R}$. Im Reellen haben wir auch das Konzept einer *Cauchy-Folge* kennengelernt. Auch diese Definition können wir kopieren.

Im Reellen haben wir gezeigt, dass die Cauchy-Eigenschaft äquivalent zur Konvergenz ist. Im Komplexen wird diese Aussage genauso gelten.

Definition 4.3.1 — CAUCHY-EIGENSCHAFT UND KONVERGENZ

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{C}$ eine Folge.

- (i) Wir sagen $(a_n)_{n \in \mathbb{N}}$ *konvergiert* gegen ein $a \in \mathbb{C}$ falls

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |a_n - a| \leq \varepsilon.$$

Ein solcher Wert $a \in \mathbb{C}$ heißt *Grenzwert* von $(a_n)_{n \in \mathbb{N}}$ und wir schreiben auch

$$\lim_{n \rightarrow \infty} a_n = a \quad \text{oder} \quad a_n \xrightarrow{n \rightarrow \infty} a.$$

Wie immer nennen wir $(a_n)_{n \in \mathbb{N}}$ *konvergent* falls es ein $a \in \mathbb{C}$ gibt, sodass $(a_n)_{n \in \mathbb{N}}$ gegen a konvergiert.

- (ii) Wir sagen $(a_n)_{n \in \mathbb{N}}$ ist eine *Cauchy-Folge* falls

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \geq n_0 : \quad |a_n - a_m| \leq \varepsilon.$$

Man kann einige *Grenzwertregeln* für reelle Folgen auf Folgen in $\mathbb{C}$ verallgemeinern. Wir wiederholen hier diejenigen, die auch in $\mathbb{C}$ Gültigkeit haben. Überzeugen Sie sich davon, dass die Beweise sich aus dem Reellen übertragen lassen.

Proposition 4.3.2 — GRENZWERTREGELN IN $\mathbb{C}$

Es seien $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{C}$ und $(b_n)_{n \in \mathbb{N}} \subseteq \mathbb{C}$ zwei konvergente Folgen und $\lambda, \mu \in \mathbb{C}$. Dann gilt

- (o) Der Grenzwert von $(a_n)_{n \in \mathbb{N}}$ ist eindeutig und ändert sich nicht, wenn man endlich viele Folgenglieder abändert. Ferner ist $(a_n)_{n \in \mathbb{N}}$ in $\mathbb{C}$ *beschränkt*^a, d.h. $\exists K \geq 0$ mit $|a_n| \leq K$ für alle $n \in \mathbb{N}$.
- (i) $(a_n + b_n)_{n \in \mathbb{N}}$ ist konvergent und es gilt $\lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n$.
- (ii) $(a_n b_n)_{n \in \mathbb{N}}$ ist konvergent und es gilt $\lim_{n \rightarrow \infty} a_n b_n = \left(\lim_{n \rightarrow \infty} a_n \right) \cdot \left(\lim_{n \rightarrow \infty} b_n \right)$.
- (iii) $(\lambda a_n + \mu b_n)_{n \in \mathbb{N}}$ ist konvergent und es gilt $\lim_{n \rightarrow \infty} (\lambda a_n + \mu b_n) = \lambda \lim_{n \rightarrow \infty} a_n + \mu \lim_{n \rightarrow \infty} b_n$.

(iv) Ist $\lim_{n \rightarrow \infty} b_n \neq 0$, so gibt es ein $N \in \mathbb{N}$ so, dass $b_n \neq 0 \ \forall n \geq N$. Ferner ist $(\frac{a_n}{b_n})_{n \geq N}$

konvergent und es gilt $\lim_{n \rightarrow \infty} \left(\frac{a_n}{b_n} \right) = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n}$.

(v) Es gilt $\lim_{n \rightarrow \infty} a_n = 0 \iff \lim_{n \rightarrow \infty} |a_n| = 0$. Für andere Grenzwerte als 0 gilt eine solche Aussage nicht.

(vi) Falls $\lim_{n \rightarrow \infty} a_n = 0$ und $(c_n)_{n \in \mathbb{N}} \subset \mathbb{C}$ beschränkt ist (vgl. Punkt (i)), dann gilt $\lim_{n \rightarrow \infty} (a_n c_n) = 0$. Für andere Grenzwerte als 0 gilt eine solche Aussage nicht.

^aBeachte: Weil $\mathbb{C}$ keine Anordnungsstruktur hat, lässt sich unsere ursprüngliche Definition der Beschränktheit in Definition 2.1.3 NICHT auf komplexwertige Folgen übertragen. Die äquivalente Charakterisierung in Proposition 2.1.6 jedoch schon!

Beweis. Ziemlich analog zu den Beweisen im Reellen, vgl. Kapitel 2.3.

Eine Grenzwertregel die sich NICHT auf komplexwertige Folgen verallgemeinern lässt ist die ‘GWR no. 4’ (vgl. Proposition 4). Sie beschäftigt sich mit dem Verhalten von Ungleichungen unter Grenzwertbildung. Der Grund, dass diese in $\mathbb{C}$ nicht gilt, ist, dass man in $\mathbb{C}$ überhaupt nicht von einer Ungleichung sprechen kann — denn auf $\mathbb{C}$ haben wir ja keine Anordnung zur Verfügung! Daher ist auch genausowenig an ein *Sandwichlemma* zu denken!

Eine Grenzwertregel, die aber dazukommt ist eine Rechenregel für das Konjugiertkomplexe

Proposition 4.3.3

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{C}$ eine konvergente Folge. Dann ist auch $(\overline{a_n})_{n \in \mathbb{N}}$ konvergent und

$$\lim_{n \rightarrow \infty} \overline{a_n} = \overline{\lim_{n \rightarrow \infty} a_n}$$

Beweis. Überlassen wir den Lesenden als Übungsaufgabe.

Die folgende Proposition führt die Begriffe aus Definition 4.3.1 auf die Konvergenz/Cauchy-Eigenschaft in $\mathbb{R}$ zurück — durch Betrachtung von Realteil und Imaginärteil!

Proposition 4.3.4

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{C}$ eine Folge.

(i) $(a_n)_{n \in \mathbb{N}}$ ist konvergent $\iff (\operatorname{Re}(a_n))_{n \in \mathbb{N}} \subseteq \mathbb{R}$ und $(\operatorname{Im}(a_n))_{n \in \mathbb{N}} \subseteq \mathbb{R}$ sind konvergent.

Ist eine der beiden obigen Aussagen wahr, so gilt

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} \operatorname{Re}(a_n) + i \lim_{n \rightarrow \infty} \operatorname{Im}(a_n). \quad (191)$$

(ii) $(a_n)_{n \in \mathbb{N}}$ ist eine Cauchy-Folge $\iff (\operatorname{Re}(a_n))_{n \in \mathbb{N}} \subseteq \mathbb{R}$ und $(\operatorname{Im}(a_n))_{n \in \mathbb{N}} \subseteq \mathbb{R}$ sind Cauchy-Folgen.

Beweis. Zu (i). Zuerst zu „ $\implies$ “. Es sei $(a_n)_{n \in \mathbb{N}}$ konvergent, etwa sei $a_n \xrightarrow{n \rightarrow \infty} a$. Das heißt

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad |a_n - a| \leq \varepsilon \quad \forall n \geq n_0.$$

ZWISCHENBEHAUPTUNG. $\operatorname{Re}(a_n) \xrightarrow{n \rightarrow \infty} \operatorname{Re}(a)$ und $\operatorname{Im}(a_n) \xrightarrow{n \rightarrow \infty} \operatorname{Im}(a)$.

Zu Ersterem. Fixiere $\varepsilon > 0$. Wähle $n_0 \in \mathbb{N}$ so, dass

$$|a_n - a| \leq \varepsilon \quad \forall n \geq n_0.$$

Dann gilt für $n \geq n_0$

$$|\operatorname{Re}(a_n) - \operatorname{Re}(a)| \underset{\text{Prop 4.1.2}}{=} |\operatorname{Re}(a_n - a)| \underset{\text{Bem 4.2.5}}{\leq} |a_n - a| \leq \varepsilon.$$

Mit anderen Worten

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad |\operatorname{Re}(a_n) - \operatorname{Re}(a)| \leq \varepsilon.$$

Mit anderen Worten haben wir nachgewiesen, dass $\operatorname{Re}(a_n) \xrightarrow{n \rightarrow \infty} \operatorname{Re}(a)$. Dass $\operatorname{Im}(a_n) \xrightarrow{n \rightarrow \infty} \operatorname{Im}(a)$ zeigt man ganz analog.

ZWISCHENBEHAUPTUNG ENDE.

Dass $(\operatorname{Re}(a_n))_{n \in \mathbb{N}}$ und $(\operatorname{Im}(a_n))_{n \in \mathbb{N}}$ konvergieren, ist also schon nachgewiesen. Wir folgern aus der Zwischenbehauptung auch, dass

$$\lim_{n \rightarrow \infty} a_n = a = \operatorname{Re}(a) + i\operatorname{Im}(a) = \lim_{n \rightarrow \infty} \operatorname{Re}(a_n) + i \lim_{n \rightarrow \infty} \operatorname{Im}(a_n).$$

Somit können wir auch (191) schon als bewiesen betrachten.

Nun aber noch zu „ $\impliedby$ “. Es seien $(\operatorname{Re}(a_n))_{n \in \mathbb{N}}$ und $(\operatorname{Im}(a_n))_{n \in \mathbb{N}}$ konvergent. Dann ist

$$a_n = \operatorname{Re}(a_n) + i\operatorname{Im}(a_n) = 1 \cdot \operatorname{Re}(a_n) + i\operatorname{Im}(a_n)$$

Nach Proposition 4.3.2(iii) angewendet mit $\lambda = 1, \mu = i, (\tilde{a}_n)_{n \in \mathbb{N}} := (\operatorname{Re}(a_n))_{n \in \mathbb{N}}, (\tilde{b}_n)_{n \in \mathbb{N}} := (\operatorname{Im}(a_n))_{n \in \mathbb{N}}$ ist dann $(a_n)_{n \in \mathbb{N}}$ auch konvergent.

Zu (ii). Zu „ $\implies$ “. Es sei $(a_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge. Das heißt

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad |a_n - a_m| \leq \varepsilon \quad \forall n, m \geq n_0.$$

Sei nun $\varepsilon > 0$. Wähle $n_0 \in \mathbb{N}$ so, dass

$$|a_n - a_m| \leq \varepsilon \quad \forall n, m \geq n_0.$$

Dann gilt für $n, m \geq n_0$ auch

$$|\operatorname{Re}(a_n) - \operatorname{Re}(a_m)| \stackrel{\text{Prop 4.1.2}}{=} |\operatorname{Re}(a_n - a_m)| \stackrel{\text{Bem 4.2.5}}{\leq} |a_n - a_m| \leq \varepsilon.$$

Wir haben gezeigt

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad |\operatorname{Re}(a_n) - \operatorname{Re}(a_m)| \leq \varepsilon.$$

Sprich: $(\operatorname{Re}(a_n))_{n \in \mathbb{N}}$ ist eine Cauchy-Folge. Analog zeigt man, dass $(\operatorname{Im}(a_n))_{n \in \mathbb{N}}$ eine Cauchy-Folge ist.

Nun zu „ $\Leftarrow$ “. Es seien $(\operatorname{Re}(a_n))_{n \in \mathbb{N}}$ und $(\operatorname{Im}(a_n))_{n \in \mathbb{N}}$ beide Cauchy-Folgen. Es sei $\varepsilon > 0$ beliebig vorgegeben. Dann gilt wegen Definition 4.3.1 (angewendet mit $\varepsilon' = \frac{\varepsilon}{2} > 0$)

$$\exists n_1 \in \mathbb{N} : \quad |\operatorname{Re}(a_n) - \operatorname{Re}(a_m)| \leq \frac{\varepsilon}{2} \quad \forall n, m \geq n_1 \quad (192)$$

$$\exists n_2 \in \mathbb{N} : \quad |\operatorname{Im}(a_n) - \operatorname{Im}(a_m)| \leq \frac{\varepsilon}{2} \quad \forall n, m \geq n_2 \quad (193)$$

Definiere nun $n_0 := \max\{n_1, n_2\}$. Für $n, m \geq n_0$ (d.h. $n, m \geq n_1$ und $n, m \geq n_2$) gilt

$$\begin{aligned} |a_n - a_m| &= |\operatorname{Re}(a_n) + i\operatorname{Im}(a_n) - (\operatorname{Re}(a_m) + i\operatorname{Im}(a_m))| \\ &= |\operatorname{Re}(a_n) - \operatorname{Re}(a_m) + i(\operatorname{Im}(a_n) - \operatorname{Im}(a_m))| \\ &\stackrel{\text{Satz 4.2.9(iii)}}{\leq} |\operatorname{Re}(a_n) - \operatorname{Re}(a_m)| + |i(\operatorname{Im}(a_n) - \operatorname{Im}(a_m))| \\ &\stackrel{\text{Satz 4.2.9(ii)}}{=} |\operatorname{Re}(a_n) - \operatorname{Re}(a_m)| + |i||\operatorname{Im}(a_n) - \operatorname{Im}(a_m)| \\ &\stackrel{\text{Bsp 4.2.10}}{=} |\operatorname{Re}(a_n) - \operatorname{Re}(a_m)| + |\operatorname{Im}(a_n) - \operatorname{Im}(a_m)| \\ &\stackrel{(192), (193)}{\leq} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Es ist gezeigt, dass

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad |a_n - a_m| \leq \varepsilon \quad \forall n, m \geq n_0.$$

Sprich: $(a_n)_{n \in \mathbb{N}}$ ist eine Cauchy-Folge.

Die obige Beobachtung erlaubt uns auch die Diskussion der *(Norm-)Vollständigkeit* von $\mathbb{C}$, vgl. Bemerkung 3.5.35.

Proposition 4.3.5 — (NORM-)VOLLSTÄNDIGKEIT VON $\mathbb{C}$

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{C}$ eine Folge. Dann gilt

$$(a_n)_{n \in \mathbb{N}} \text{ ist eine Cauchy-Folge} \iff (a_n)_{n \in \mathbb{N}} \text{ ist konvergent.}$$

Beweis.

$(a_n)_{n \in \mathbb{N}}$ ist eine Cauchy-Folge

$$\stackrel{\text{Prop 4.3.4}}{\iff} (\operatorname{Re}(a_n))_{n \in \mathbb{N}} \text{ ist Cauchy-Folge} \quad \text{und} \quad (\operatorname{Im}(a_n))_{n \in \mathbb{N}} \text{ ist Cauchy-Folge.}$$

$$\stackrel{\text{Bem 3.5.35}}{\iff} (\operatorname{Re}(a_n))_{n \in \mathbb{N}} \text{ ist konvergent} \quad \text{und} \quad (\operatorname{Im}(a_n))_{n \in \mathbb{N}} \text{ ist konvergent.}$$

$$\stackrel{\text{Prop 4.3.4}}{\iff} (a_n)_{n \in \mathbb{N}} \text{ ist konvergent.}$$

08.12.2025

5 Reihen

Folgen sind (intuitiv) *nicht abbrechende* Auflistungen von Zahlen. *Unendliche Reihen*, welche wir hier studieren werden, sind intuitiv *nicht abbrechende* Summen von Zahlen.

Mit den herkömmlichen Körperoperationen können wir aber zunächst nur *endlich* viele Zahlen aufsummieren. Für alles weitere müssen wir uns an dem Konzept des *Grenzwertes* bedienen. Bevor wir dazu kommen definieren wir aber erstmal das Konzept der *Partialsumme*.

Definition 5.0.1 — PARTIALSUMME

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{C}$ eine Folge und $m \in \mathbb{N}$. Wir definieren die *m-te Partialsumme* $s_m \in \mathbb{C}$ durch

$$s_m := \sum_{k=1}^m a_k. \quad (194)$$

Die Folge $(s_m)_{m \in \mathbb{N}}$ heißt *Partialsummenfolge* zugehörig zu $(a_n)_{n \in \mathbb{N}}$.

Mit anderen Worten: Die *m-te* Partialsumme entsteht durch aufsummieren der ersten *m* Folgenglieder. Also intuitiv

$$s_m = a_1 + a_2 + \dots + a_m \quad \forall m \in \mathbb{N}$$

Definition 5.0.2 — UNENDLICHE REIHE

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{C}$ eine Folge. Sei $(s_m)_{m \in \mathbb{N}}$ die zugehörige Partialsummenfolge. Wir sagen,

- (i) $\sum_{k \geq 1} a_k$ ist eine *konvergente (unendliche) Reihe*, falls die Partialsummenfolge $(s_m)_{m \in \mathbb{N}} \subseteq \mathbb{C}$ konvergiert.
- (ii) $\sum_{k \geq 1} a_k$ ist eine *divergente (unendliche) Reihe*, falls die Partialsummenfolge $(s_m)_{m \in \mathbb{N}} \subseteq \mathbb{C}$ divergiert.

Im Falle, dass $\sum_{k \geq 1} a_k$ eine konvergente unendliche Reihe ist, nennen wir

$$\sum_{k=1}^{\infty} a_k := \lim_{m \rightarrow \infty} s_m \quad (195)$$

den *Wert der Reihe*.

Achtung: Bei der Konvergenz der unendlichen Reihe $\sum_{k \geq 1} a_k$ geht es also *nicht* um die Konvergenz der Folge $(a_n)_{n \in \mathbb{N}}$ sondern stattdessen um die Konvergenz der *Partialsummenfolge* $(s_m)_{m \in \mathbb{N}}$, die erst gebildet werden muss.

Nutzt man (194) in (195), so erhält man eine intuitive Gleichung für den Wert der Reihe

$$\sum_{k=1}^{\infty} a_k = \lim_{m \rightarrow \infty} \left(\sum_{k=1}^m a_k \right).$$

Man beachte aber: Damit die obige Gleichung Sinn ergibt, muss es sich bei $\sum_{k \geq 1} a_k$ um eine konvergente Reihe handeln. Ansonsten darf man gar nicht über den Grenzwert auf der rechten Seite sprechen.

Beispiel 5.0.3

Behauptung. Die Reihe $\sum_{k \geq 1} 1$ ist eine divergente Reihe.

Beweis. Dazu bilden wir zunächst die Partialsummenfolge. Für $m \in \mathbb{N}$ ist

$$s_m := \sum_{k=1}^m 1 = m.$$

Die Folge $(s_m)_{m \in \mathbb{N}} = (m)_{m \in \mathbb{N}}$ ist divergent. Das könnte man zum Beispiel damit begründen, dass diese Folge unbeschränkt ist (vgl. Beispiel 2.3.15).

Jetzt ein erstes Beispiel für eine konvergente Reihe.

Beispiel 5.0.4

Behauptung. Die Reihe $\sum_{k \geq 1} \frac{1}{k(k+1)}$ ist eine konvergente Reihe.

Beweis. Dazu bilden wir zunächst die Partialsummenfolge. Für $m \in \mathbb{N}$ sei

$$s_m := \sum_{k=1}^m \frac{1}{k(k+1)}.$$

Nun ein Trick: Für alle $k \in \mathbb{N}$ ist $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$. Bringen Sie dazu die rechte Seite der Gleichung auf einen Hauptnenner, dann werden Sie das nachvollziehen können!^a

Wir erhalten

$$s_m = \sum_{k=1}^m \frac{1}{k(k+1)} = \sum_{k=1}^m \left(\frac{1}{k} - \frac{1}{k+1} \right) \stackrel{\text{Teleskopsumme}}{=} \frac{1}{1} - \frac{1}{m+1} = 1 - \frac{1}{m+1}.$$

Diese Partialsummenfolge ist konvergent! Somit konvergiert $\sum_{k \geq 1} a_k$ und es gilt

$$\sum_{k=1}^{\infty} a_k \stackrel{(195)}{=} \lim_{m \rightarrow \infty} s_m = \lim_{m \rightarrow \infty} \left(1 - \frac{1}{m+1} \right) = 1.$$

^aNachvollziehen ist das Eine, aber drauf zu kommen ist etwas anderes. Und doch gibt es eine Möglichkeit, auf eine solche Zerlegung zu kommen. Sie nennt sich *Partialbruchzerlegung*. Machen Sie an der Stelle doch einen *leap of faith* und wählen Sie den Ansatz $\frac{1}{k(k+1)} \stackrel{!}{=} \frac{A}{k} + \frac{B}{k+1} \quad \forall k \in \mathbb{N}$ (für zu wählende $A, B \in \mathbb{R}$). Bringen Sie auf einen Hauptnenner zurück und finden Sie Möglichkeiten für A und B durch

einen Koeffizientenvergleich. Im Detail:

$$\begin{aligned}
 \frac{1}{k(k+1)} &\stackrel{!}{=} \frac{A}{k} + \frac{B}{k+1} \\
 \Leftrightarrow \frac{1}{k(k+1)} &\stackrel{!}{=} \frac{A(k+1)}{k(k+1)} + \frac{Bk}{k(k+1)} \\
 \Leftrightarrow \frac{1}{k(k+1)} &\stackrel{!}{=} \frac{A(k+1) + Bk}{k(k+1)} \\
 \Leftrightarrow 1 &\stackrel{!}{=} A(k+1) + Bk = (A+B)k + A
 \end{aligned}$$

Um diese Gleichung für alle $k \in \mathbb{N}$ wahr zu machen, bietet es sich an, A und B so zu wählen, dass $A = 1$ und $A + B = 0$. Mit anderen Worten $A = 1$ und $B = -1$. Man erhält mit dieser Wahl $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$.

Wie bei endlichen Summen kann man übrigens auch andere Startindizes betrachten, d.h. $\sum_{k \geq \ell} a_k$ für beliebige $\ell \in \mathbb{Z}$. Sie können die Begriffe „Partialsomme“, „konvergente/divergente unendliche Reihe“, „Wert der Reihe“ dann analog zu Definition 5.0.1 und 5.0.2 definieren.

Beispiel 5.0.5

GEOMETRISCHE REIHEN.^a Es sei $z \in \mathbb{C}$ mit $|z| < 1$.

Behauptung. Dann ist die sog. *geometrische Reihe* $\sum_{k \geq 0} z^k$ eine konvergente Reihe. Ihr Reihenwert ist gegeben durch

$$\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}.$$

Beweis. Dazu bilden wir zunächst die Partialsommenfolge. Für $m \in \mathbb{N}$ sei

$$s_m := \sum_{k=0}^m z^k.$$

Nach der geometrischen Summenformel (vgl. Proposition 3.3.20) gilt

$$s_m = \sum_{k=0}^m z^k = \frac{1 - z^{m+1}}{1 - z}$$

Beachte: Nach Beispiel 2.3.8 gilt (weil $0 \leq |z| < 1$)

$$\lim_{m \rightarrow \infty} |z^{m+1}| \stackrel{\text{Satz 4.2.9(ii)}}{=} \lim_{m \rightarrow \infty} |z|^{m+1} = \lim_{m \rightarrow \infty} |z|^m |z| \stackrel{\text{Bsp 2.3.8}}{=} 0 \cdot |z| = 0.$$

Mithilfe der Grenzwertregeln erhalten wir auch $\lim_{m \rightarrow \infty} z^{m+1} = 0$ und somit

$$\lim_{m \rightarrow \infty} s_m = \lim_{m \rightarrow \infty} \frac{1 - z^{m+1}}{1 - z} = \frac{1}{1 - z}.$$

Daher ist $\sum_{k \geq 0} z^k$ konvergent und

$$\sum_{k=0}^{\infty} z^k = \lim_{m \rightarrow \infty} s_m = \frac{1}{1-z}.$$

^aAuch hier hilft GEOGEBRA bei der Anschauung.

Sehr spannend wird ist die Konvergenzfrage auch bei $\sum_{k \geq 1} \frac{1}{k}$, der sogenannten *harmonischen Reihe*.

Proposition 5.0.6 — DIE HARMONISCHE REIHE

Die Reihe $\sum_{k \geq 1} \frac{1}{k}$ ist eine divergente Reihe.

Beweis. Dazu betrachten wir wieder die Partialsummenfolge, d.h für $m \in \mathbb{N}$ definieren wir

$$s_m := \sum_{k=1}^m \frac{1}{k}.$$

Wir behaupten, dass $(s_m)_{m \in \mathbb{N}}$ divergiert. Dies werden wir zeigen, indem wir zeigen, dass $(s_m)_{m \in \mathbb{N}}$ unbeschränkt ist. Dazu betrachte für $\ell \in \mathbb{N} \cup \{0\}$

$$s_{2^\ell} = \sum_{k=1}^{2^\ell} \frac{1}{k} = \sum_{k=1}^{2^0} \frac{1}{k} + \sum_{k=2^0+1}^{2^1} \frac{1}{k} + \sum_{k=2^1+1}^{2^2} \frac{1}{k} + \dots + \sum_{k=2^{\ell-1}+1}^{2^\ell} \frac{1}{k}$$

Wir haben die Summe s_{2^ℓ} also in $\ell + 1$ Teilsummen aufgeteilt. Die erste hat nur einen Summanden, weil $2^0 = 1$. Ferner ist $\sum_{k=1}^1 \frac{1}{k} = \frac{1}{1} = 1$.

Sei nun $j \geq 2$. Die j -te Teilsumme, genauer gegeben durch

$$\sum_{k=2^{j-2}+1}^{2^{j-1}} \frac{1}{k},$$

hat genau $2^{j-1} - 2^{j-2} = 2 \cdot 2^{j-2} - 2^{j-2} = 2^{j-2}$ Summanden. Für die j -te Teilsumme gilt außerdem die Abschätzung

$$\sum_{k=2^{j-2}+1}^{2^{j-1}} \frac{1}{k} \geq \sum_{k=2^{j-2}+1}^{2^{j-1}} \frac{1}{2^{j-1}} = 2^{j-2} \frac{1}{2^{j-1}} = 2^{j-2-(j-1)} = 2^{-1} = \frac{1}{2}.$$

Man erhält

$$\begin{aligned}
 s_{2^\ell} &= \sum_{k=1}^{2^\ell} \frac{1}{k} = \underbrace{\sum_{k=1}^{2^0} \frac{1}{k}}_{=1} + \underbrace{\sum_{k=2^0+1}^{2^1} \frac{1}{k}}_{\geq \frac{1}{2}} + \underbrace{\sum_{k=2^1+1}^{2^2} \frac{1}{k}}_{\geq \frac{1}{2}} + \dots + \underbrace{\sum_{k=2^{\ell-1}+1}^{2^\ell} \frac{1}{k}}_{\geq \frac{1}{2}} \\
 &\geq 1 + \underbrace{\frac{1}{2} + \dots + \frac{1}{2}}_{\ell\text{-mal}} = 1 + \ell \cdot \frac{1}{2} = 1 + \frac{\ell}{2}.
 \end{aligned}$$

Da $(1 + \frac{\ell}{2})_{\ell \in \mathbb{N}}$ nach oben unbeschränkt ist, ist also auch $(s_{2^\ell})_{\ell \in \mathbb{N}}$ nach oben unbeschränkt. Man überzeugt sich leicht davon, dass dann auch $(s_m)_{m \in \mathbb{N}}$ nach oben unbeschränkt sein muss. Diese Folge muss also divergent sein. Somit ist auch $\sum_{k \geq 1} \frac{1}{k}$ eine divergente Reihe (vgl. Definition 5.0.2).

Die harmonische Reihe ist gewissermaßen ein *Grenzfall*, wie die folgende Proposition zeigt:

Proposition 5.0.7

Es sei $\theta > 1$. Dann ist $\sum_{k \geq 1} \frac{1}{k^\theta}$ eine konvergente Reihe.

Beweis. Wiederum definieren wir die Partialsummen, d.h. für alle $m \in \mathbb{N}$ ist

$$s_m := \sum_{k=1}^m \frac{1}{k^\theta}.$$

Beachte: $(s_m)_{m \in \mathbb{N}}$ ist monoton wachsend, da jeder Summand positiv ist. Genauer gilt für $m \in \mathbb{N}$

$$s_{m+1} - s_m = \sum_{k=1}^{m+1} \frac{1}{k^\theta} - \sum_{k=1}^m \frac{1}{k^\theta} = \frac{1}{(m+1)^\theta} > 0, \quad \text{das heißt} \quad s_{m+1} > s_m \quad \forall m \in \mathbb{N}.$$

Wenn wir jetzt noch zeigen können, dass $(s_m)_{m \in \mathbb{N}}$ zusätzlich auch nach oben beschränkt ist, folgt die Konvergenz aus Proposition 3.5.2. Also

$$\mathbb{Z}: \quad (s_m)_{m \in \mathbb{N}} \text{ ist nach oben beschränkt.}$$

Dazu: Da $m \leq 2^m$ (vgl. Beispiel 2.2.21) und $(s_m)_{m \in \mathbb{N}}$ monoton wachsend ist, gilt auch

$$s_m \leq s_{2^m} = \sum_{k=1}^{2^m} \frac{1}{k^\theta} = \sum_{k=1}^{2^0} \frac{1}{k^\theta} + \sum_{k=2^0+1}^{2^1} \frac{1}{k^\theta} + \dots + \sum_{k=2^{m-1}+1}^{2^m} \frac{1}{k^\theta}. \quad (196)$$

Wir haben hier wieder $m + 1$ Teilsummen. Die erste hat wiederum den Wert $\sum_{k=1}^1 \frac{1}{k^\theta} = 1$.

Für $j \geq 2$ besteht die j -te Teilsumme aus

$$\sum_{k=2^{j-2}+1}^{2^{j-1}} \frac{1}{k^\theta}.$$

Beachte: Da in der j -ten Teilsumme nur Summationsindizes $k \geq 2^{j-2}$ vorkommen gilt

$$\begin{aligned} \sum_{k=2^{j-2}+1}^{2^{j-1}} \frac{1}{k^\theta} &\leq \sum_{k=2^{j-2}+1}^{2^{j-1}} \frac{1}{(2^{j-2})^\theta} = \frac{1}{(2^{j-2})^\theta} (2^{j-1} - 2^{j-2}) \\ &= \frac{1}{(2^{j-2})^\theta} (2 \cdot 2^{j-2} - 2^{j-2}) = \frac{1}{(2^{j-2})^\theta} 2^{j-2} = \frac{1}{(2^\theta)^{j-2}} 2^{j-2} \\ &= \left(\frac{2}{2^\theta} \right)^{j-2} = \alpha^{j-2}, \end{aligned}$$

wobei $\alpha := \frac{2}{2^\theta} < 1$. Schätzt man für alle $j \geq 2$ die j -te Teilsumme in (196) durch α^{j-2} ab, so erhält man

$$s_m \leq 1 + \sum_{j=2}^{m+1} \alpha^{j-2} \stackrel{\text{Indexverschiebung}}{=} 1 + \sum_{j=0}^{m-1} \alpha^j \stackrel{\text{Prop 3.3.20}}{=} 1 + \frac{1 - \alpha^m}{1 - \alpha} \leq 1 + \frac{1}{1 - \alpha} \quad \forall m \in \mathbb{N}.$$

Somit ist $(s_m)_{m \in \mathbb{N}}$ durch $1 + \frac{1}{1-\alpha}$ nach oben beschränkt. Proposition 3.5.2 liefert nun die Konvergenz von $(s_m)_{m \in \mathbb{N}}$. Somit ist $\sum_{k \geq 1} \frac{1}{k^\theta}$ konvergent.

Eine wichtige Rechenregel für konvergente Reihen ist die Konvergenz von Linearkombinationen.

Proposition 5.0.8 — LINEARKOMBINATIONEN KONVERGENTER REIHEN

Es seien $\sum_{k \geq 1} a_k, \sum_{k \geq 1} b_k$ zwei konvergente Reihen und $\lambda, \mu \in \mathbb{C}$. Dann ist $\sum_{k \geq 1} (\lambda a_k + \mu b_k)$ eine konvergente Reihe und es gilt

$$\sum_{k=1}^{\infty} (\lambda a_k + \mu b_k) = \lambda \sum_{k=1}^{\infty} a_k + \mu \sum_{k=1}^{\infty} b_k.$$

Beweis. Es definieren für $m \in \mathbb{N}$ die Partialsummen

$$\sigma_m := \sum_{k=1}^m a_k$$

und

$$\tau_m := \sum_{k=1}^m b_k.$$

Nach Definition 5.0.2 sind $(\sigma_m)_{m \in \mathbb{N}}$ und $(\tau_m)_{m \in \mathbb{N}}$ konvergent und

$$\lim_{m \rightarrow \infty} \sigma_m = \sum_{k=1}^{\infty} a_k, \quad \text{und} \quad \lim_{m \rightarrow \infty} \tau_m = \sum_{k=1}^{\infty} b_k. \quad (197)$$

Wir definieren nun

$$s_m := \sum_{k=1}^m (\lambda a_k + \mu b_k)$$

und stellen fest, dass nach Proposition 3.3.17

$$s_m = \sum_{k=1}^m (\lambda a_k + \mu b_k) = \lambda \sum_{k=1}^m a_k + \mu \sum_{k=1}^m b_k = \lambda \sigma_m + \mu \tau_m.$$

Wegen der Grenzwertregeln und (197) erhalten wir

$$\lim_{m \rightarrow \infty} s_m = \lim_{m \rightarrow \infty} (\lambda \sigma_m + \mu \tau_m) = \lambda \lim_{m \rightarrow \infty} \sigma_m + \mu \lim_{m \rightarrow \infty} \tau_m \stackrel{(197)}{=} \lambda \sum_{k=1}^{\infty} a_k + \mu \sum_{k=1}^{\infty} b_k.$$

Nutzen wir nun Definition 5.0.2 (iii) so folgt auch

$$\sum_{k=1}^{\infty} (\lambda a_k + \mu b_k) = \lambda \sum_{k=1}^{\infty} a_k + \mu \sum_{k=1}^{\infty} b_k.$$

Beispiel 5.0.9

Ist $\sum_{k \geq 1} a_k$ eine divergente Reihe, so ist auch für alle $c \neq 0$ die Reihe $\sum_{k \geq 1} c \cdot a_k$ eine divergente Reihe.

Beweis. Sei $c \neq 0$. Wäre $\sum_{k \geq 1} c \cdot a_k$ eine konvergente Reihe, so wäre auch (nach Proposition 5.0.8 mit $\lambda = \frac{1}{c}$ und $\mu = 0$ sowie $b_k = 0$ für alle k) die Reihe

$$\sum_{k \geq 1} \lambda \cdot c \cdot a_k = \sum_{k \geq 1} \frac{1}{c} \cdot c \cdot a_k = \sum_{k \geq 1} a_k$$

eine konvergente Reihe. Ein Widerspruch.

Insbesondere ist also auch zum Beispiel $\sum_{k \geq 1} \frac{1}{2k} = \sum_{k \geq 1} \frac{1}{2} \cdot \frac{1}{k}$ eine divergente Reihe.

Genau wie bei Folgen endlich viele Folgenglieder für die Konvergenzfrage keine Rolle spielen, spielen endlich viele Summanden für die Reihenkonvergenz keine Rolle.

Proposition 5.0.10

Es seien $\ell_1, \ell_2 \in \mathbb{Z}$ beliebig. Dann gilt

$$\sum_{k \geq \ell_1} a_k \text{ ist konvergente Reihe} \iff \sum_{k \geq \ell_2} a_k \text{ ist konvergente Reihe}$$

Konvergiert eine der beiden Reihen und ist $\ell_1 < \ell_2$ so gilt ferner

$$\sum_{k=\ell_1}^{\infty} a_k = \sum_{k=\ell_1}^{\ell_2-1} a_k + \sum_{k=\ell_2}^{\infty} a_k$$

Beweis. OBdA sei $\ell_1 < \ell_2$. Definiere für $m \in \mathbb{N}$ die Partialsummen

$$s_m := \sum_{k=\ell_1}^m a_k \quad \text{und} \quad \sigma_m := \sum_{k=\ell_2}^m a_k.$$

Dann gilt für alle $m \geq \ell_2$

$$s_m = \sum_{k=\ell_1}^m a_k = \sum_{k=\ell_1}^{\ell_2-1} a_k + \sum_{k=\ell_2}^m a_k = \underbrace{\sum_{k=\ell_1}^{\ell_2-1} a_k}_{=:c} + \sigma_m = c + \sigma_m.$$

Somit unterscheiden sich $(s_m)_{m \geq \ell_2}$ und $(\sigma_m)_{m \geq \ell_2}$ nur um eine additive Konstante. Man überzeugt sich leicht (Grenzwertregeln), dass eine solche Konstante das Konvergenzverhalten nicht beeinflusst. Außerdem gilt nach den Grenzwertregeln

$$\sum_{k=\ell_1}^{\infty} a_k = \lim_{m \rightarrow \infty} s_m = \lim_{m \rightarrow \infty} (c + \sigma_m) = c + \lim_{m \rightarrow \infty} \sigma_m = c + \sum_{k=\ell_2}^{\infty} a_k = \sum_{k=\ell_1}^{\ell_2-1} a_k + \sum_{k=\ell_2}^{\infty} a_k.$$

5.1 Konvergenzkriterien

Mithilfe der Kriterien, die wir in diesem Abschnitt lernen, können wir Reihen auf Konvergenz/Divergenz überprüfen. Und das ohne, dass wir den Wert der Reihe angeben können!

Etwas ähnliches haben wir schonmal gesehen: Das *Cauchy-Kriterium* erlaubt uns, eine Folge auf Konvergenz zu überprüfen, ohne, dass wir den Grenzwert angeben können. Die Kriterien, die wir im Folgenden studieren, werden genau auf diesem Cauchy-Kriterium basiert sein. Im folgenden Abschnitt formulieren wir dies erstmal noch aus.

Das Cauchy-Kriterium für Reihenkonvergenz

Proposition 5.1.1 — CAUCHY-KRITERIUM FÜR REIHENKONVERGENZ

Die Reihe $\sum_{k \geq 1} a_k$ ist eine konvergente Reihe genau dann wenn

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m : n > m \geq n_0 : \left| \sum_{k=m+1}^n a_k \right| \leq \varepsilon.$$

Beweis. Definiere für $m \in \mathbb{N}$

$$s_m := \sum_{k=1}^m a_k \tag{198}$$

Dann gilt

$$\begin{aligned} \sum_{k \geq 1} a_k \text{ ist eine konvergente Reihe} \\ \stackrel{\text{Def 5.0.2}}{\iff} (s_m)_{m \in \mathbb{N}} \text{ konvergiert} \\ \stackrel{\text{Bem 3.5.35}}{\iff} (s_m)_{m \in \mathbb{N}} \text{ ist eine Cauchy-Folge} \\ \stackrel{\text{Def 3.5.31}}{\iff} \forall \varepsilon > 0 : \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \geq n_0 : \quad |s_n - s_m| \leq \varepsilon. \end{aligned}$$

Beachte: Wenn in der unteren Formel $n = m$ betrachtet wird, gilt $|s_n - s_n| = 0 \leq \varepsilon$. Dieser Fall muss also nicht gesondert betrachtet werden, weil die Ungleichung sowieso stimmt. Man erhält

$$\begin{aligned} \sum_{k \geq 1} a_k \text{ ist eine konvergente Reihe} \\ \iff \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \geq n_0 : n \neq m : \quad |s_n - s_m| \leq \varepsilon. \end{aligned}$$

Beachte auch: Da $|s_n - s_m| = |s_m - s_n|$ kann man stets erreichen, dass der größere Index von beiden in der Differenz vorne steht. Daher kann $n > m$ auch ohne Einschränkung angenommen werden. Man erhält

$$\begin{aligned} \sum_{k \geq 1} a_k \text{ ist eine konvergente Reihe} \\ \iff \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m : n > m \geq n_0 : \quad |s_n - s_m| \leq \varepsilon. \end{aligned}$$

Weiter:

$$\begin{aligned} \stackrel{(198)}{\iff} \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m : n > m \geq n_0 : \quad \left| \sum_{k=1}^n a_k - \sum_{k=1}^m a_k \right| \leq \varepsilon \\ \stackrel{\text{Prop 3.3.17}}{\iff} \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m : n > m \geq n_0 : \quad \left| \sum_{k=m+1}^n a_k \right| \leq \varepsilon \end{aligned}$$

Eine genauere Erklärung des letzten Schritts: Nach Proposition 3.3.17 gilt $\sum_{k=1}^n a_k = \sum_{k=1}^m a_k + \sum_{k=m+1}^n a_k$.

Das Abkling-Kriterium

Als erstes Kriterium zeigen wir das *Abkling-Kriterium*, welches *nur Divergenz* nachweisen kann und nicht Konvergenz.

Stellt man sich die Folgenglieder als Töne auf einem Xylophon vor und die Reihe als *Akkord (Überlagerung) aus allen diesen unendlich vielen Tönen*, so kann der Akkord nur schön klingen wenn die Töne *im unendlichen sehr leise werden*. Dies wird durch Teil (ii) des nachfolgenden Satzes erfasst.

Anders ausgedrückt: Wenn das, was wir dort aufsummieren wollen, nicht *am Ende sehr klein ist*, dann kann die Reihe nicht konvergieren.

Satz 5.1.2 — ABKLING-KRITERIUM

$$(i) \quad \sum_{k \geq 1} a_k \text{ ist eine konvergente Reihe} \implies a_n \xrightarrow{n \rightarrow \infty} 0. \quad (a)$$

$$(ii) \quad \neg(a_n \xrightarrow{n \rightarrow \infty} 0) \implies \sum_{k \geq 1} a_k \text{ ist eine divergente Reihe.}$$

^aKommen Sie an dieser Stelle bitte nicht in Versuchung, aus diesem Folgerungspfeil einen Äquivalenzpfeil zu machen! Vergessen Sie niemals Bemerkung 5.1.3, s.u.!

Teil (i) ist nicht als Konvergenzkriterium nicht nutzbar: Denn die Implikation geht in die falsche Richtung! Die Rückrichtung gilt auch nicht, wie das nachfolgende Beispiel zeigt

Bemerkung 5.1.3

$$a_n \xrightarrow{n \rightarrow \infty} 0 \quad \not\Rightarrow \quad \sum_{k \geq 1} a_k \text{ konvergiert!!!}$$

In der Tat: Wir haben schon gesehen, dass $\frac{1}{n} \xrightarrow{n \rightarrow \infty} 0$, aber $\sum_{k \geq 1} \frac{1}{k}$ divergiert (vgl. Proposition 5.0.6).

Fazit.

Teil (ii) des Abkling-Kriteriums (Satz 5.1.2) ist nützlich, um Divergenz einer Reihe zu zeigen.

Teil (i) des Kriteriums kann aber NICHT verwendet werden, um Konvergenz zu zeigen!

Diese Bemerkung war mir wichtig zu machen *bevor* wir das Abkling-Kriterium zeigen, da dieser Trugschluss oft Probleme bereitet.

Beweis von Satz 5.1.2. (ii) folgt aus (i) direkt über die Kontrapositionsregel (Proposition 1.1.7 (ii)). Daher genügt es, (i) zu zeigen. Zu (i). Es sei $\sum_{k \geq 1} a_k$ eine konvergente Reihe. Mit Proposition 5.1.1 folgern wir

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m : n > m \geq n_0 : \quad \left| \sum_{k=m+1}^n a_k \right| \leq \varepsilon.$$

Fixieren Sie ein $\varepsilon > 0$ beliebig und wählen Sie $n_0 \in \mathbb{N}$ wie in der obigen Gleichung. Dann gilt

$$\forall n, m : n > m \geq n_0 : \quad \left| \sum_{k=m+1}^n a_k \right| \leq \varepsilon.$$

Betrachten Sie nun den Spezialfall $m = n - 1$. Dann wäre $\left| \sum_{k=m+1}^n a_k \right| = \left| \sum_{k=n}^n a_k \right| = |a_n|$. Damit $m \geq n_0$ ist muss $n \geq n_0 + 1$ sein. Wir folgern aus dieser Überlegung

$$\forall n \geq n_0 + 1 : \quad |a_n| \leq \varepsilon.$$

Definiere nun $\tilde{n}_0 := n_0 + 1$. Dann haben wir

$$\forall n \geq \tilde{n}_0 : \quad |a_n - 0| \leq \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, haben wir gezeigt

$$\forall \varepsilon > 0 \quad \exists \tilde{n}_0 \in \mathbb{N} \quad \forall n \geq \tilde{n}_0 : \quad |a_n - 0| \leq \varepsilon.$$

Das heißt nach Definition 2.3.1 nichts anderes als $a_n \xrightarrow{n \rightarrow \infty} 0$.

Beispiel 5.1.4

Die Reihe $\sum_{k \geq 1} (-1)^k$ ist nach dem Abkling-Kriterium (Satz 5.1.2 (ii)) eine divergente Reihe, denn $(-1)^n \not\rightarrow 0$ (für $n \rightarrow \infty$).

Beispiel 5.1.5 — GEOMETRISCHE REIHEN - II

Betrachte wie in Beispiel 5.0.5 die *geometrische Reihe* $\sum_{k \geq 0} z^k$, nur diesmal für $z \in \mathbb{C}$ mit $|z| \geq 1$ (anstatt $|z| < 1$). Dann ist diese Reihe divergent, denn für $n \in \mathbb{N}$ gilt

$$|z^n| = |z|^n \geq 1^n = 1 \quad \text{und somit} \quad |z^n| \not\rightarrow 0 \quad (n \rightarrow \infty).$$

11.12.2025

Das Leibniz-Kriterium

Das nächste Kriterium, was wir studieren wollen ist das *Leibniz-Kriterium für alternierende Reihen*. Was eine *alternierende Reihe* ist, wird aus der Aussage des nachfolgenden Satzes

ersichtlich.

Beachte: Es ist in dieser Vorlesung eines der wenigen Konvergenzkriterien, bei dem nur *reellwertige Summanden* zugelassen sind!

Satz 5.1.6 — LEIBNIZ-KRITERIUM

Gegeben sei eine reelle Folge $(b_k)_{k \in \mathbb{N}} \subseteq \mathbb{R}$ mit folgenden Eigenschaften

(L1) $(b_k)_{k \in \mathbb{N}}$ ist monoton fallend,

(L2) $\lim_{n \rightarrow \infty} b_n = 0$.

Dann bildet die sogenannte *alternierende Reihe* $\sum_{k \geq 1} (-1)^k b_k$ eine konvergente Reihe.

Beweis. Als Vorabbemerkung besprechen wir zunächst eine wichtige Folgerung aus (L1) und (L2). Für alle $k \in \mathbb{N}$ gilt

$$b_k \geq \inf\{b_j : j \in \mathbb{N}\} \stackrel{\text{Prop 3.5.2}}{\underset{\text{(L1)}}{=}} \lim_{j \rightarrow \infty} b_j \stackrel{\text{(L2)}}{=} 0.$$

Das bedeutet also

$$b_k \geq 0 \quad \forall k \in \mathbb{N}. \quad (199)$$

Definiere nun wieder einmal für $m \in \mathbb{N}$ die Partialsumme

$$s_m := \sum_{k=1}^m (-1)^k b_k. \quad (200)$$

$\mathbb{Z} : (s_m)_{m \in \mathbb{N}}$ ist konvergent.

Dazu:

ZWISCHENBEHAUPTUNG 1. $(s_{2\ell})_{\ell \in \mathbb{N}}$ ist monoton fallend und $(s_{2\ell+1})_{\ell \in \mathbb{N}}$ ist monoton wachsend.

Zu Ersterem:

$$s_{2(\ell+1)} = s_{2\ell+2} = \sum_{k=1}^{2\ell+2} (-1)^k b_k = \underbrace{\sum_{k=1}^{2\ell} (-1)^k b_k}_{= s_{2\ell} \text{ vgl. (200)}} - b_{2\ell+1} + b_{2\ell+2} = s_{2\ell} + \underbrace{(b_{2\ell+2} - b_{2\ell+1})}_{\leq 0, \text{ vgl. (L1)}} \leq s_{2\ell}.$$

Im letzten Schritt haben wir benutzt: wegen der Monotonie von $(b_k)_{k \in \mathbb{N}}$ gilt $b_{2\ell+2} \leq b_{2\ell+1}$.

Zu Letzerem:

$$s_{2(\ell+1)+1} = s_{2\ell+3} = \sum_{k=1}^{2\ell+3} (-1)^k b_k = \underbrace{\sum_{k=1}^{2\ell+1} (-1)^k b_k}_{=s_{2\ell+1} \text{ vgl. (200)}} + b_{2\ell+2} - b_{2\ell+3} = s_{2\ell+1} + \underbrace{(b_{2\ell+2} - b_{2\ell+3})}_{\geq 0, \text{ vgl. (L1)}} \geq s_{2\ell+1}.$$

ZWISCHENBEHAUPTUNG 1 ENDE.

ZWISCHENBEHAUPTUNG 2. $s_{2\ell} \geq s_1$ für alle $\ell \in \mathbb{N}$ und $s_{2\ell+1} \leq s_2$ für alle $\ell \in \mathbb{N}$.

Zu Ersterem: Iteriert man Zwischenbehauptung 1 so findet man

$$s_1 \underset{\text{ZB.1}}{\leq} s_3 \underset{\text{ZB.1}}{\leq} \dots \underset{\text{ZB.1}}{\leq} s_{2\ell-1} \underset{(200)}{=} \sum_{k=1}^{2\ell-1} (-1)^k b_k = \sum_{k=1}^{2\ell} (-1)^k b_k - [(-1)^{2\ell} b_{2\ell}] \underset{(200)}{=} s_{2\ell} - b_{2\ell} \underset{(199)}{\leq} s_{2\ell}.$$

Zu Letzerem:

$$\begin{aligned} s_2 \underset{\text{ZB.1}}{\geq} s_4 \underset{\text{ZB.1}}{\geq} \dots \underset{\text{ZB.1}}{\geq} s_{2\ell} &\underset{(200)}{=} \sum_{k=1}^{2\ell} (-1)^k b_k \\ &= \sum_{k=1}^{2\ell+1} (-1)^k b_k - [(-1)^{2\ell+1} b_{2\ell+1}] \underset{(200)}{=} s_{2\ell+1} + b_{2\ell+1} \underset{(199)}{\geq} s_{2\ell+1}. \end{aligned}$$

ZWISCHENBEHAUPTUNG 2 ENDE.

Aus Zwischenbehauptung 1 und 2 folgt, dass $(s_{2\ell})_{\ell \in \mathbb{N}}$ eine monoton fallende und nach unten beschränkte Folge ist. Nach Proposition 3.5.2 existiert

$$\lim_{\ell \rightarrow \infty} s_{2\ell} =: S$$

Desweiteren ist $(s_{2\ell+1})_{\ell \in \mathbb{N}}$ eine monoton wachsende und nach oben beschränkte Folge. Wiederum nach Proposition 3.5.2 existiert

$$\lim_{\ell \rightarrow \infty} s_{2\ell+1} =: S'$$

Nach den Grenzwertregeln gilt

$$\begin{aligned} S' - S &= \lim_{\ell \rightarrow \infty} s_{2\ell+1} - \lim_{\ell \rightarrow \infty} s_{2\ell} = \lim_{\ell \rightarrow \infty} (s_{2\ell+1} - s_{2\ell}) \underset{(200)}{=} \lim_{\ell \rightarrow \infty} \left(\sum_{k=1}^{2\ell+1} (-1)^k b_k - \sum_{k=1}^{2\ell} (-1)^k b_k \right) \\ &= \lim_{\ell \rightarrow \infty} (-b_{2\ell+1}) = \lim_{\ell \rightarrow \infty} (-1)^{2\ell+1} b_{2\ell+1} \underset{\text{(L2) und Prop 3.5.13}}{=} 0. \end{aligned}$$

Also $S = S'$. Mit anderen Worten

$$\lim_{\ell \rightarrow \infty} s_{2\ell} = \lim_{\ell \rightarrow \infty} s_{2\ell+1} = S. \quad (201)$$

Behauptung. $(s_m)_{m \in \mathbb{N}}$ konvergiert gegen S .

Wer die entsprechende Übungsaufgabe (Blatt 08 Aufgabe 2b) bearbeitet hat, kann das direkt aus dieser folgern und wäre mit dem Beweis fertig. Für alle anderen präsentieren wir hier nochmal das Argument. Es sei $\varepsilon > 0$. Wegen (201) gilt

$$\exists n_1 \in \mathbb{N} : \quad |s_{2\ell} - S| \leq \varepsilon \quad \forall \ell \geq n_1 \quad (202)$$

und

$$\exists n_2 \in \mathbb{N} : \quad |s_{2\ell+1} - S| \leq \varepsilon \quad \forall \ell \geq n_2. \quad (203)$$

Wähle nun $n_0 := \max\{2n_1, 2n_2 + 1\}$. Nun gilt

$$m \geq n_0 \quad \Rightarrow \quad (m \geq 2n_1) \wedge (m \geq 2n_2 + 1).$$

FALL 1. m gerade, d.h. $m = 2j$ für ein $j \in \mathbb{N}$. Aus $2j = m \geq 2n_1$ folgt dann $j \geq n_1$. Deswegen gilt

$$|s_m - S| = |s_{2j} - S| \underset{(202)}{\leq} \varepsilon.$$

FALL 2. m ungerade, d.h. $m = 2j + 1$ für ein $j \in \mathbb{N}$. Aus $2j + 1 = m \geq 2n_2 + 1$ folgt dann $j \geq n_2$. Deswegen gilt

$$|s_m - S| = |s_{2j+1} - S| \underset{(203)}{\leq} \varepsilon.$$

Nimmt man beide Fälle zusammen, haben wir also gezeigt

$$|s_m - S| \leq \varepsilon \quad \forall m \geq n_0.$$

Da $\varepsilon > 0$ beliebig war gilt

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad |s_m - S| \leq \varepsilon \quad \forall m \geq n_0.$$

Es folgt also, dass $(s_m)_{m \in \mathbb{N}}$ (gegen S) konvergiert. Definition 5.0.2 impliziert dann, dass die alternierende Reihe $\sum_{k \geq 1} (-1)^k b_k$ eine konvergente Reihe ist.

Beispiel 5.1.7

Die Reihe

$$\sum_{k \geq 1} (-1)^k \frac{1}{k}$$

bildet nach dem Leibniz-Kriterium eine konvergente Reihe.

Beweis. Definieren wir für $k \in \mathbb{N}$ die reelle Zahl $b_k := \frac{1}{k}$, so gilt

(L1) $(b_k)_{k \in \mathbb{N}}$ ist monoton fallend, da $b_{k+1} = \frac{1}{k+1} \leq \frac{1}{k} = b_k$.

(L2) $\lim_{k \rightarrow \infty} b_k = \lim_{k \rightarrow \infty} \frac{1}{k} = 0$ (vgl. Beispiel 2.3.3)

Auf die Voraussetzungen (L1) und (L2) im Konvergenzkriterium von Leibniz kann im Allgemeinen nicht verzichtet werden.

Beispiel 5.1.8

Auf (L1) kann nicht verzichtet werden. In der Tat: betrachte den Fall $b_k = (-1)^k \frac{1}{k}$. Dann ist (L1) nicht erfüllt, (L2) hingegen schon (vgl. Beispiel 2.3.34). Jedoch ist

$$\sum_{k \geq 1} (-1)^k b_k = \sum_{k \geq 1} (-1)^k (-1)^k \frac{1}{k} = \sum_{k \geq 1} [(-1)(-1)]^k \frac{1}{k} = \sum_{k \geq 1} \frac{1}{k}$$

eine divergente Reihe.

Beispiel 5.1.9

Auf (L2) kann auch nicht verzichtet werden. Denn falls (L2) nicht erfüllt ist, d.h. $b_n \not\rightarrow 0$ (für $n \rightarrow \infty$), so gilt auch,

$$|(-1)^n b_n| = |b_n| \not\rightarrow 0 \quad (n \rightarrow \infty)$$

und somit auch $(-1)^n b_n \not\rightarrow 0$ (für $n \rightarrow \infty$) [vgl. Proposition 2.3.30]. Nach dem Abklingkriterium (Satz 5.1.2 (ii)) ist dann $\sum_{k \geq 1} (-1)^k b_k$ eine divergente Reihe.

Absolute Konvergenz von Reihen

Als nächstes lernen wir das Konzept der *absoluten Konvergenz* kennen.

Definition 5.1.10 — ABSOLUT KONVERGENTE REIHEN

Wir sagen $\sum_{k \geq 1} a_k$ ist eine *absolut konvergente Reihe*, wenn $\sum_{k \geq 1} |a_k|$ eine konvergente Reihe ist^a.

^avgl. Definition 5.0.2

Jetzt könnte man zunächst einen Einwand hervorbringen: $\sum_{k \geq 1} |a_k|$ ist ja im Allgemeinen eine völlig andere Reihe als $\sum_{k \geq 1} a_k$ — warum sollte man diese beiden Reihen überhaupt mit einer Begrifflichkeit miteinander verknüpfen?
Die Antwort geben wir in der nachfolgenden Proposition

Proposition 5.1.11 — ABSOLUTE KONVERGENZ IMPLIZIERT KONVERGENZ

Es sei $\sum_{k \geq 1} a_k$ eine absolut konvergente Reihe. Dann ist $\sum_{k \geq 1} a_k$ eine konvergente Reihe (im Sinne von Definition 5.0.2).

Beweis. Es sei $\sum_{k \geq 1} a_k$ eine absolut konvergente Reihe. Das bedeutet $\sum_{k \geq 1} |a_k|$ ist eine konvergente Reihe. Nach Proposition 5.1.1 gilt

$$\forall \varepsilon > 0 \quad \exists n_1 \in \mathbb{N} \quad \forall n, m \in \mathbb{N}, n > m \geq n_1 : \quad \left| \sum_{k=m+1}^n |a_k| \right| \leq \varepsilon.$$

Da $\sum_{k=m+1}^n |a_k| \in \mathbb{R}$ und $\sum_{k=m+1}^n |a_k| \geq 0$ sind die äußeren Betragsstriche nicht notwendig und man erhält

$$\forall \varepsilon > 0 \quad \exists n_1 \in \mathbb{N} \quad \forall n, m \in \mathbb{N}, n > m \geq n_1 : \quad \sum_{k=m+1}^n |a_k| \leq \varepsilon. \quad (204)$$

Nun ist zu zeigen, dass $\sum_{k \geq 1} a_k$ eine konvergente Reihe ist. Dazu verwenden wir Proposition 5.1.1. Das heißt

$$\mathbb{Z} : \quad \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \in \mathbb{N}, n > m \geq n_0 : \quad \left| \sum_{k=m+1}^n a_k \right| \leq \varepsilon.$$

Sei dazu $\varepsilon > 0$ beliebig vorgegeben. Wir wissen aus (204)

$$\exists n_1 \in \mathbb{N} \quad \forall n, m \in \mathbb{N}, n > m \geq n_1 : \quad \sum_{k=m+1}^n |a_k| \leq \varepsilon.$$

Wähle $n_0 := n_1$. Für zwei Indizes $n, m \in \mathbb{N}$ mit $n > m \geq n_0$ gilt dann (wegen der iterierten Dreiecksungleichung, vgl. 4.2.13)

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

Wir haben also gesehen

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \in \mathbb{N} : n > m \geq n_0 : \left| \sum_{k=m+1}^n a_k \right| \leq \varepsilon.$$

Nach dem Cauchy-Kriterium für Reihenkonvergenz (Proposition 5.1.1) folgt somit, dass $\sum_{k \geq 1} a_k$ eine konvergente Reihe ist.

Bemerkung 5.1.12

Wir haben in der vorigen Proposition gesehen, dass jede absolut konvergente Reihe auch eine konvergente Reihe ist. Die Umkehrung gilt nicht:

Nicht jede konvergente Reihe ist eine absolut konvergente Reihe.

Betrachten Sie dazu

$$\sum_{k \geq 1} \frac{(-1)^k}{k}$$

Diese ist nach dem Leibnizkriterium eine konvergente Reihe (vgl. Beispiel 5.1.7). Jedoch ist es keine absolut konvergente Reihe, denn

$$\sum_{k \geq 1} \left| \frac{(-1)^k}{k} \right| = \sum_{k \geq 1} \frac{1}{k}$$

ist eine divergente Reihe, vgl. Proposition 5.0.6.

Umordnungen von Reihen Als nächstes sprechen wir über das Konvergenzverhalten von *Umordnungen*. Eine *Umordnung* einer Reihe entsteht durch vertauschen der Summanden gemäß einer bijektiven Abbildung $\tau : \mathbb{N} \rightarrow \mathbb{N}$.

Definition 5.1.13 — UMORDNUNG

Es sei eine Reihe $\sum_{k \geq 1} a_k$ gegeben. Für eine bijektive Abbildung $\tau : \mathbb{N} \rightarrow \mathbb{N}$ nennen wir die Reihe $\sum_{k \geq 1} a_{\tau(k)}$ die *aus τ entstehende Umordnung von $\sum_{k \geq 1} a_k$* .

Bei endlichen Summen spielt die Vertauschung von Summanden keine Rolle. Das wissen wir aus den Körperaxiomen! Bei konvergenten unendlichen Reihen ist das allerdings nicht so klar — und im allgemeinen auch falsch!

Proposition 5.1.14 — UMORDNUNGEN MIT ANDEREM WERT

- (i) Es bildet $\sum_{k \geq 1} \frac{(-1)^{k+1}}{k}$ eine konvergente Reihe und für den Reihenwert gilt $S :=$

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \geq \frac{1}{2}.$$

(ii) Es gibt $\tau : \mathbb{N} \rightarrow \mathbb{N}$ bijektiv mit $\sum_{k=1}^{\infty} \frac{(-1)^{\tau(k)+1}}{\tau(k)} \neq S$. Dies gilt zum Beispiel für $\tau : \mathbb{N} \rightarrow \mathbb{N}$ gegeben durch

$$\tau(k) := \begin{cases} 4\ell - 3 & \text{falls } k = 3\ell - 2 \text{ für ein } \ell \in \mathbb{N} \\ 4\ell - 1 & \text{falls } k = 3\ell - 1 \text{ für ein } \ell \in \mathbb{N} \\ 2\ell & \text{falls } k = 3\ell \text{ für ein } \ell \in \mathbb{N} \end{cases}$$

Beweis(skizze). Zu (i). Die Konvergenz folgt aus Beispiel 5.1.7 und Proposition 5.0.8 mit $\lambda = -1, \mu = 0$, da $\sum_{k \geq 1} \frac{(-1)^{k+1}}{k} = \sum_{k \geq 1} (-1) \frac{(-1)^k}{k}$. Definiere für $j \in \mathbb{N}$ die j -te Partialsumme

$$s_j := \sum_{k=1}^j \frac{(-1)^{k+1}}{k}.$$

Zur unteren Abschätzung des Wertes der Reihe betrachte für $m \in \mathbb{N}$

$$\begin{aligned} s_{2m} &= \sum_{k=1}^{2m} \frac{(-1)^{k+1}}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots - \dots + \frac{1}{2m-1} - \frac{1}{2m} = \\ &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \dots + \left(\frac{1}{2m-1} - \frac{1}{2m}\right) \\ &= \sum_{k=1}^m \left(\frac{1}{2k-1} - \frac{1}{2k}\right) = \sum_{k=1}^m \left(\frac{2k - (2k-1)}{(2k-1)2k}\right) \\ &= \sum_{k=1}^m \frac{1}{2k(2k-1)} = \frac{1}{2} + \sum_{k=2}^m \frac{1}{2k(2k-1)} \geq \frac{1}{2}. \end{aligned}$$

Nun gilt

$$S = \lim_{j \rightarrow \infty} s_j \stackrel{\text{Prop 3.5.13}}{=} \lim_{m \rightarrow \infty} s_{2m} \stackrel{\text{GWR}}{\geq} \frac{1}{2}.$$

Zu (ii). Diesen Sachverhalt skizzieren wir nur. Ein formaler Beweis kann aus dem hier aufgezeigten Phänomen entwickelt werden. Die Bijektivität von τ zeigen wir nicht. Beobachte nun

$$\begin{aligned} S &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \frac{1}{9} - \frac{1}{10} + \frac{1}{11} - \frac{1}{12} + \dots \\ \frac{1}{2}S &= \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} + \dots \end{aligned}$$

$$\begin{aligned}
S + \frac{1}{2}S \stackrel[\text{rot rauskürzen}]{\text{blautöne summieren}} 1 + \frac{1}{3} \underbrace{\frac{1}{4} - \frac{1}{4}}_{=-\frac{1}{2}} + \frac{1}{5} + \frac{1}{7} \underbrace{\frac{1}{8} - \frac{1}{8}}_{=-\frac{1}{4}} + \frac{1}{9} + \frac{1}{11} \underbrace{\frac{1}{12} - \frac{1}{12}}_{=-\frac{1}{6}} + \dots \\
= 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \dots = \sum_{k=1}^{\infty} \frac{(-1)^{\tau(k)+1}}{\tau(k)}.
\end{aligned}$$

Man erhält

$$\sum_{k=1}^{\infty} \frac{(-1)^{\tau(k)+1}}{\tau(k)} = \frac{3}{2}S.$$

Nun gilt $\frac{3}{2}S \neq S$, denn es ist $S \geq \frac{1}{2} > 0$ und deswegen gilt $\frac{3}{2}S > S$.

Wir sehen: Beim Umordnen von Reihen kann sich der Wert ändern. Übrigens kann sich sogar das Konvergenzverhalten verändern! Es ist tatsächlich möglich, die obige Reihe so umzuordnen, dass sie divergent wird!

Bei absolut konvergenten Reihen muss man sich beim Umordnen allerdings keine Sorgen machen, wie wir aus dem folgenden Satz lernen.

Satz 5.1.15 — UMORDNUNGSSATZ

Es sei $\sum_{k \geq 1} a_k$ eine absolut konvergente Reihe und $\tau : \mathbb{N} \rightarrow \mathbb{N}$ eine bijektive Abbildung. Dann ist auch die aus τ entstehende Umordnung $\sum_{k \geq 1} a_{\tau(k)}$ eine konvergente Reihe und für den Wert der Umordnung gilt

$$\sum_{k=1}^{\infty} a_{\tau(k)} = \sum_{k=1}^{\infty} a_k.$$

Beweis. Für $m \in \mathbb{N}$ sei

$$s_m := \sum_{k=1}^m a_k \quad \text{und} \quad \sigma_m := \sum_{k=1}^m a_{\tau(k)}.$$

Definiere ferner

$$S := \lim_{m \rightarrow \infty} s_m = \sum_{k=1}^{\infty} a_k \in \mathbb{C}.$$

Zu zeigen ist dass die umgeordnete Reihe denselben Wert hat, d.h.

$$\mathbb{Z} : \quad \sigma_m \xrightarrow{m \rightarrow \infty} S.$$

Dazu sei $\varepsilon > 0$ beliebig vorgegeben. Verwendet man das Cauchy-Kriterium für die konver-

gente Reihe $\sum_{k=1}^{\infty} |a_k|$ (vgl. Proposition 5.1.1) mit $\varepsilon' = \frac{\varepsilon}{2}$, so folgert man

$$\exists n_0 \in \mathbb{N} \quad \forall n, m \in \mathbb{N} : n > m \geq n_0 : \sum_{k=m+1}^n |a_k| \leq \frac{\varepsilon}{2}. \quad (205)$$

Beachte: Weil $(s_j - s_{n_0})_{j \in \mathbb{N}}$ für $j \rightarrow \infty$ gegen $S - s_{n_0}$ konvergiert, konvergiert auch $(|s_j - s_{n_0}|)_{j \in \mathbb{N}}$ gegen $|S - s_{n_0}|$ (vgl. Proposition 2.3.32). Man hat

$$|S - s_{n_0}| = \lim_{j \rightarrow \infty} |s_j - s_{n_0}| = \lim_{j \rightarrow \infty} \left| \sum_{k=1}^j a_k - \sum_{k=1}^{n_0} a_k \right| = \lim_{j \rightarrow \infty} \left| \sum_{k=n_0+1}^j a_k \right|. \quad (206)$$

Nun gilt für alle $j \geq n_0 + 1$

$$\left| \sum_{k=n_0+1}^j a_k \right| \underset{\text{Prop 4.2.13}}{\leq} \sum_{k=n_0+1}^j |a_k| \leq \frac{\varepsilon}{2}$$

Aus den Grenzwertvergleichsregeln (Proposition 4) und Gleichung (206) folgt

$$|S - s_{n_0}| \underset{(206)}{=} \lim_{j \rightarrow \infty} \left| \sum_{k=n_0+1}^j a_k \right| \leq \lim_{j \rightarrow \infty} \frac{\varepsilon}{2} = \frac{\varepsilon}{2}.$$

Da τ bijektiv ist, gibt es für alle $i \in \{1, \dots, n_0\}$ genau ein $k_i \in \mathbb{N}$ mit $i = \tau(k_i)$. Definiert man nun $N := k_1 + \dots + k_{n_0} + 1$, so gilt

$$\{\tau(1), \dots, \tau(N)\} \supseteq \{\tau(k_1), \dots, \tau(k_{n_0})\} = \{1, \dots, n_0\}. \quad (207)$$

Dann gilt für $m \geq N + 1$

$$|\sigma_m - S| = |\sigma_m - s_{n_0} + s_{n_0} - S| \leq |\sigma_m - s_{n_0}| + \underbrace{|s_{n_0} - S|}_{\leq \frac{\varepsilon}{2}} \leq |\sigma_m - s_{n_0}| + \frac{\varepsilon}{2}. \quad (208)$$

Nun hat man

$$|\sigma_m - s_{n_0}| = \left| \sum_{k=1}^m a_{\tau(k)} - \sum_{k=1}^{n_0} a_k \right|.$$

Aufgrund von (207) kommen in der ersten Summe alle Summanden $a_1, \dots, a_{n_0}$ vor! Wegen der Bijektivität von τ kommen Sie dort sogar genau einmal vor. Verrechnen wir diese mit den Summanden der zweiten Summe so kürzt sich die komplette zweite Summe weg. Auf die übrig bleibenden Summanden der ersten Summe können wir die Dreiecksungleichung anwenden und erhalten

$$|\sigma_m - s_{n_0}| \leq \sum_{k=n_0+1}^{m^*} |a_k|,$$

wobei $m^* = \max\{\tau(\ell) : \ell \in \{1, \dots, m\}\} > n_0$ (weil alle Werte von 1 bis n_0 schon in $\{\tau(1), \dots, \tau(N)\}$ vorkommen und τ wegen der Bijektivität keinen Wert doppelt annimmt). Verwenden wir nun (205) mit $n = m^*$ und $m = n_0$ so ergibt sich

$$|\sigma_m - s_{n_0}| \leq \sum_{k=n_0+1}^{m^*} |a_k| \leq \frac{\varepsilon}{2}.$$

Dies benutzen wir auf der rechten Seite (208) und finden heraus, dass für alle $m \geq N+1 =: \tilde{N}$ gilt

$$|\sigma_m - S| \underset{(208)}{\leq} |\sigma_m - s_{n_0}| + \frac{\varepsilon}{2} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Somit haben wir gesehen

$$\forall \varepsilon > 0 \quad \exists \tilde{N} \in \mathbb{N} : \quad |\sigma_m - S| \leq \varepsilon \quad \forall m \geq \tilde{N}.$$

Dies zeigt, dass $\sigma_m \xrightarrow{m \rightarrow \infty} S$. Die Behauptung folgt.

Majorantenkriterium

Eine wichtige Grundlage für die weiteren Konvergenzkriterien, die wir studieren werden, bildet das *Majorantenkriterium*. Es untersucht die *absolute Konvergenz von Reihen*. Es besagt, dass man absolute Konvergenz durch das Finden einer *konvergenten Majorantenreihe* nachweisen kann.

Satz 5.1.16 — MAJORANTENKRITERIUM

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{C}$ so, dass $(|a_n|)_{n \in \mathbb{N}} \leq (c_n)_{n \in \mathbb{N}}$ für eine Folge reeller Zahlen $(c_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$. Dann

Ist $\sum_{k \geq 1} c_k$ eine konvergente Reihe, so ist $\sum_{k \geq 1} a_k$ eine absolut konvergente Reihe.

Will man dieses Kriterium anwenden, um Konvergenz einer Reihe $\sum_{k \geq 1} a_k$ zu zeigen, so muss man also stets eine Folge $(c_n)_{n \in \mathbb{N}}$ finden mit

- $(|a_n|)_{n \in \mathbb{N}} \leq (c_n)_{n \in \mathbb{N}}$
- $\sum_{k \geq 1} c_k$ ist konvergent.

Wir nennen eine solche Reihe $\sum_{k \geq 1} c_k$ dann *konvergente Majorantenreihe* zu $\sum_{k \geq 1} a_k$.

Beweis. Es sei also $\sum_{k \geq 1} c_k$ eine konvergente Reihe. Wir müssen zeigen, dass $\sum_{k \geq 1} a_k$ konver-

giert. Das bedeutet nach dem Cauchy-Kriterium für Reihenkonvergenz:

$$\mathbb{Z}: \quad \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \in \mathbb{N}, n > m \geq n_0 : \quad \sum_{k=m+1}^n |a_k| \leq \varepsilon.$$

Dazu geben wir uns $\varepsilon > 0$ beliebig vor. Da $(|a_n|)_{n \in \mathbb{N}} \leq (c_n)_{n \in \mathbb{N}}$ gilt

$$\exists n_1 \in \mathbb{N} \quad \forall k \geq n_1 : \quad |a_k| \leq c_k. \quad (209)$$

Da $\sum_{k \geq 1} c_k$ eine konvergente Reihe ist, gilt außerdem

$$\exists n_2 \in \mathbb{N} \quad \forall n, m \in \mathbb{N}, n > m \geq n_2 : \quad \left| \sum_{k=m+1}^n c_k \right| \leq \varepsilon.$$

Wähle nun $n_0 := \max\{n_1, n_2\}$. Für $n, m \in \mathbb{N}$ mit $n > m \geq n_0$ gilt dann

$$\sum_{k=m+1}^n |a_k| \underset{(209)}{\leq} \sum_{k=m+1}^n c_k. \quad (210)$$

Beachte nun: Da für $k \geq n_1$ gilt $c_k \geq |a_k| \geq 0$ hat man für $n > m \geq n_1$

$$\sum_{k=m+1}^n c_k = \left| \sum_{k=m+1}^n c_k \right|.$$

Somit impliziert (210), dass für $n > m \geq n_0$

$$\sum_{k=m+1}^n |a_k| \underset{(209)}{\leq} \left| \sum_{k=m+1}^n c_k \right| \leq \varepsilon.$$

Wir haben gesehen:

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \in \mathbb{N} : n > m \geq n_0 : \quad \sum_{k=m+1}^n |a_k| \leq \varepsilon.$$

Beispiel 5.1.17

Wir untersuchen

$$\sum_{k \geq 1} \frac{1}{\sqrt{2k^3 + 1}}$$

auf Konvergenz. Dazu: Beachte, dass für alle $k \in \mathbb{N}$

$$\left| \frac{1}{\sqrt{2k^3 + 1}} \right| = \frac{1}{\sqrt{2k^3 + 1}} \leq \frac{1}{\sqrt{2k^3 + 0}} = \frac{1}{\sqrt{2}} \frac{1}{k^{\frac{3}{2}}} \leq \frac{1}{k^{\frac{3}{2}}}$$

Somit gilt also auch $(\frac{1}{\sqrt{2k^3+1}})_{k \in \mathbb{N}} \leq (\frac{1}{k^{\frac{3}{2}}})_{k \in \mathbb{N}}$. Nach Proposition 5.0.7 ist $\sum_{k \geq 1} \frac{1}{k^{\frac{3}{2}}}$ eine konvergente Reihe. Nach dem Majorantenkriterium (Satz 5.1.16) gilt somit auch

$$\sum_{k \geq 1} \frac{1}{\sqrt{2k^3+1}} \text{ ist eine konvergente Reihe.}$$

Beispiel 5.1.18

Wir untersuchen

$$\sum_{k \geq 1} \frac{1}{\sqrt{2k^3-1}}$$

auf Konvergenz. Dazu: Beachte, dass für alle $k \in \mathbb{N}$

$$\left| \frac{1}{\sqrt{2k^3-1}} \right| = \frac{1}{\sqrt{2k^3-1}} \leq \frac{1}{\sqrt{2k^3-k^3}} = \frac{1}{\sqrt{k^3}} = \frac{1}{k^{\frac{3}{2}}}$$

Somit gilt also auch $(\frac{1}{\sqrt{2k^3-1}})_{k \in \mathbb{N}} \leq (\frac{1}{k^{\frac{3}{2}}})_{k \in \mathbb{N}}$. Nach Proposition 5.0.7 ist $\sum_{k \geq 1} \frac{1}{k^{\frac{3}{2}}}$ eine konvergente Reihe. Nach dem Majorantenkriterium (Satz 5.1.16) gilt somit auch

$$\sum_{k \geq 1} \frac{1}{\sqrt{2k^3-1}} \text{ ist eine konvergente Reihe.}$$

Es gibt auch ein *Minorantenkriterium*, welches die *Divergenz* von Reihen mit positiven reellen Summanden zeigen kann. Hierfür muss man eine *divergente Minorante* finden.

Proposition 5.1.19 — MINORANTENKRITERIUM

Es seien $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}$ reelle Folgen. Ferner gelte $b_n \geq 0$ für alle $n \in \mathbb{N}$ und $(a_n)_{n \in \mathbb{N}} \geq (b_n)_{n \in \mathbb{N}}$. Dann gilt

Ist $\sum_{k \geq 1} b_k$ eine divergente Reihe, so ist auch $\sum_{k \geq 1} a_k$ eine divergente Reihe.

Beweis. Sei also $\sum_{k \geq 1} b_k$ eine divergente Reihe. Angenommen $\sum_{k \geq 1} a_k$ wäre eine konvergente Reihe. Da

$$(|b_k|)_{k \in \mathbb{N}} \stackrel{b_k \geq 0 \forall k}{=} (b_k)_{k \in \mathbb{N}} \leq (a_k)_{k \in \mathbb{N}}$$

wäre nach dem Majorantenkriterium dann auch $\sum_{k \geq 1} b_k$ eine absolut konvergente Reihe (vgl. Satz 5.1.16) und somit auch eine konvergente Reihe. Dies widerspricht der Voraussetzung, dass $\sum_{k \geq 1} b_k$ eine divergente Reihe ist.

Beispiel 5.1.1. Die Reihe

$$\sum_{k \geq 1} \frac{1}{\sqrt{k^2 + 1}}$$

ist eine divergente Reihe. Dazu: Man beobachte zuerst, dass

$$\frac{1}{\sqrt{k^2 + 1}} \geq \frac{1}{\sqrt{k^2 + k^2}} = \frac{1}{\sqrt{2k^2}} = \frac{1}{\sqrt{2}} \frac{1}{k} \quad \forall k \in \mathbb{N}.$$

Somit gilt also $(\frac{1}{\sqrt{k^2+1}})_{k \in \mathbb{N}} \geq (\frac{1}{\sqrt{2}} \frac{1}{k})_{k \in \mathbb{N}}$. Da

$$\sum_{k \geq 1} \frac{1}{\sqrt{2}} \frac{1}{k}$$

eine divergente Reihe ist (vgl. Beispiel 5.0.9 und Proposition 5.0.6), ist nach dem Minorantenkriterium auch $\sum_{k \geq 1} \frac{1}{\sqrt{k^2+1}}$ eine divergente Reihe.

Wurzelkriterium

Das nächste studienwürdige Kriterium ist das *Wurzelkriterium*, welches wiederum benutzt werden kann um die *absolute Konvergenz* (oder auch die Divergenz) von Reihen zu zeigen.

Satz 5.1.20 — WURZELKRITERIUM

(i) Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{C}$ so, dass

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} < 1.$$

Dann ist $\sum_{k \geq 1} a_k$ eine absolut konvergente Reihe.

(ii) Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{C}$ so, dass

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1.$$

Dann ist $\sum_{k \geq 1} a_k$ eine divergente Reihe.

Beweis. Zu (i). Definiere

$$q := \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} < 1$$

Beachte: Da jedes Folgenglied der oben vorkommenden Folge nichtnegativ ist, ist auch $q \geq 0$. Ferner gilt mit Definition 3.5.24

$$\sup\{\sqrt[k]{|a_k|} : k \geq n\} \xrightarrow{n \rightarrow \infty} q.$$

Definiere nun $\varepsilon^* := \frac{1-q}{2} > 0$. Dann gilt

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |\sup\{\sqrt[k]{|a_k|} : k \geq n\} - q| \leq \varepsilon^*.$$

Insbesondere gilt für $n \geq n_0$

$$\begin{aligned} \sup\{\sqrt[k]{|a_k|} : k \geq n\} &= \sup\{\sqrt[k]{|a_k|} : k \geq n\} - q + q \\ &\leq |\sup\{\sqrt[k]{|a_k|} : k \geq n\} - q| + q \leq \varepsilon^* + q \\ &= \frac{1-q}{2} + q = \frac{1+q}{2}. \end{aligned} \tag{211}$$

Somit gilt für $n \geq n_0$ auch

$$\sqrt[n]{|a_n|} \leq \sup\{\sqrt[k]{|a_k|} : k \geq n\} \underset{(211)}{\leq} \frac{1+q}{2}.$$

Nimmt man bei der linken und rechten Seite der Ungleichung die n -te Potenz so erhält man

$$\forall n \geq n_0 : \quad |a_n| \leq \left(\frac{1+q}{2}\right)^n. \tag{212}$$

Wir haben gezeigt, dass $(|a_n|)_{n \in \mathbb{N}} \leq \left(\left(\frac{1+q}{2}\right)^n\right)_{n \in \mathbb{N}}$. Da $\frac{1+q}{2} < \frac{1+1}{2} = 1$, ist $\sum_{k \geq 1} \left(\frac{1+q}{2}\right)^k$ eine konvergente Reihe (vgl. Beispiel 5.0.5). Nach dem Majorantenkriterium (Satz 5.1.16) ist dann auch $\sum_{k \geq 1} a_k$ eine absolut konvergente Reihe.

Zu (ii). Definiere

$$q := \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1.$$

FALL 1. $(\sqrt[n]{|a_n|})_{n \in \mathbb{N}}$ ist nach oben beschränkt.

Da dann q nach Proposition 3.5.27 ein Häufungswert von $(\sqrt[n]{|a_n|})_{n \in \mathbb{N}}$ ist, gilt für eine Teilfolge $\sqrt[\ell]{|a_{\ell_n}|} \xrightarrow{n \rightarrow \infty} q$.

ZWISCHENBEHAUPTUNG. $\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \sqrt[\ell]{|a_{\ell_n}|} \geq 1$

Dazu wähle $\varepsilon^* := q - 1$. Dann gilt nach der Definition der Folgenkonvergenz

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |\sqrt[\ell]{|a_{\ell_n}|} - q| \leq \varepsilon^*.$$

Für $n \geq n_0$ hat man daher

$$\sqrt[\ell]{|a_{\ell_n}|} = \sqrt[\ell]{|a_{\ell_n}|} - q + q \geq -|\sqrt[\ell]{|a_{\ell_n}|} - q| + q \geq q - \varepsilon^* = q - (q - 1) = 1.$$

ZWISCHENBEHAUPTUNG ENDE.

Insbesondere gilt für alle $n \geq n_0$ auch

$$|a_{\ell_n}| \geq 1^{\ell_n} = 1. \quad (213)$$

Mit anderen Worten: Es gibt eine Teilfolge $(a_{\ell_n})_{n \in \mathbb{N}}$ mit $|a_{\ell_n}| \geq 1$ für alle $n \geq n_0$. Als Konsequenz erhalten wir $\neg(a_{\ell_n} \xrightarrow{n \rightarrow \infty} 0)^b$. Somit gilt aber auch $\neg(a_n \xrightarrow{n \rightarrow \infty} 0)$ (denn wäre $(a_n)_{n \in \mathbb{N}}$ gegen 0 konvergent so müssten auch alle Teilfolgen gegen 0 konvergieren). Mit dem Abkling-Kriterium (Satz 5.1.2) folgt dann die Divergenz der Reihe $\sum_{k \geq 1} a_k$.

FALL 2. $(\sqrt[n]{|a_n|})_{n \in \mathbb{N}}$ ist nach oben unbeschränkt.

In diesem Fall ist klarerweise $A := \{m \in \mathbb{N} : \sqrt[m]{|a_m|} \geq 1\}$ eine unendliche Menge.^c Nach Bemerkung 3.5.15 findet man eine Teilfolge $(\ell_n)_{n \in \mathbb{N}} \subseteq A$ mit $\ell_{n+1} \geq \ell_n + 1$ für alle $n \in \mathbb{N}$. Man erhält dadurch eine Teilfolge $(\sqrt[\ell_n]{|a_{\ell_n}|})_{n \in \mathbb{N}}$ mit

$$\sqrt[\ell_n]{|a_{\ell_n}|} \geq 1 \quad \forall n \in \mathbb{N}.$$

Wiederum folgert man

$$|a_{\ell_n}| \geq 1^{\ell_n} = 1 \quad \forall n \in \mathbb{N}.$$

Ab hier kann man wie in Fall 1 (Diskussion nach (213)) vorgehen und erhält wiederum die Divergenz der Reihe $\sum_{k \geq 1} a_k$.

^aWenn man es ganz genau nimmt, müsste man das vermutlich noch mit Proposition 3.5.27 und den Grenzwertregeln begründen. Sind alle Folgenglieder nichtnegativ, so ist auch jeder Häufungswert nichtnegativ – Überlegen Sie sich das mal!

^bÜberlegen Sie sich die Details dazu als Übungsaufgabe!

^cAuch eine gute Übung!

Bemerkung 5.1.21

Gegeben sei eine Reihe $\sum_{k \geq 1} a_k$. Um das (absolute) Konvergenzverhalten gemäß Satz 5.1.20 zu verstehen berechnen wir

$$q := \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}.$$

Wir wissen aus dem vorigen Satz

$$\begin{cases} q < 1 & \implies \sum_{k \geq 1} a_k \text{ ist eine absolut konvergente Reihe,} \\ q > 1 & \implies \sum_{k \geq 1} a_k \text{ ist eine divergente Reihe.} \end{cases}$$

Frage: Kann auch eine Konvergenz-Aussage getroffen werden, wenn $q = 1$?

Antwort: NEIN!

Dazu betrachten wir die Reihen

$$\sum_{k \geq 1} \frac{1}{k} \quad \text{und} \quad \sum_{k \geq 1} \frac{1}{k^2}$$

Erstere Reihe ist divergent, letztere ist absolut konvergent. Ferner gilt

$$\lim_{n \rightarrow \infty} \sqrt[n]{\frac{1}{n}} = \lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{n}} \stackrel{\text{GWR3}}{=} \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{n}} \stackrel{\text{Prop 2.3.7(iii)}}{=} \frac{1}{1} = 1 \quad \stackrel{\text{Kor 3.5.30}}{\implies} \quad \limsup_{n \rightarrow \infty} \sqrt[n]{\frac{1}{n}} = 1.$$

und

$$\begin{aligned} \lim_{n \rightarrow \infty} \sqrt[n]{\frac{1}{n^2}} &= \lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{n^2}} \stackrel{\text{GWR3}}{=} \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{n^2}} = \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{n \cdot n}} \\ &= \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{n} \sqrt[n]{n}} \stackrel{\text{GWR2}}{=} \frac{1}{1 \cdot 1} = 1. \quad \stackrel{\text{Kor 3.5.30}}{\implies} \quad \limsup_{n \rightarrow \infty} \sqrt[n]{\frac{1}{n^2}} = 1. \end{aligned}$$

Beide obigen Reihen wären also im Fall „ $q = 1$ “ anzusiedeln — aber ihr Konvergenzverhalten ist gänzlich unterschiedlich. Somit kann im Fall $q = 1$ keine Konvergenz-Aussage getroffen werden.

15.12.2025

Quotientenkriterium

Ähnlich wie das Wurzelkriterium kann das *Quotientenkriterium* bewiesen werden. Dieses ist wiederum hinreichend für die absolute Konvergenz von Reihen

Satz 5.1.22 — QUOTIENTENKRITERIUM

Es sei $(a_n)_{n \in \mathbb{N}}$ derart, dass es ein $N \in \mathbb{N}$ gibt mit $a_n \neq 0$ für alle $n \geq N$.

(i) Ist

$$\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1,$$

so ist $\sum_{k \geq 1} a_k$ eine absolut konvergente Reihe.

(ii) Gilt

$$\liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| > 1, \quad (a)$$

so ist $\sum_{k \geq 1} a_k$ eine divergente Reihe.

^aAchtung: Hier steht wirklich $\liminf$ und nicht $\limsup$! Mehr dazu in Beispiel 5.1.24!

Beweis. Zu (i). Definiere $q := \limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|$. Dann gilt (vgl. Definition 3.5.24), dass $q \in [0, 1)$ und

$$\sup \left\{ \left| \frac{a_{k+1}}{a_k} \right| : k \geq n \right\} \xrightarrow{n \rightarrow \infty} q.$$

Ganz ähnlich wie in (212) zeigt man die Existenz eines $n_0 \in \mathbb{N}, n_0 \geq N$ mit

$$\forall n \geq n_0 : \quad \left| \frac{a_{n+1}}{a_n} \right| \leq \frac{1+q}{2}.$$

Definiere nun $\theta := \frac{1+q}{2}$. Dann gilt $0 < \theta < 1$ (weil $\frac{1+q}{2} < \frac{1+1}{2} = 1$). Ferner gilt

$$\forall n \geq n_0 : \quad \left| \frac{a_{n+1}}{a_n} \right| \leq \theta.$$

Das bedeutet

$$|a_{n+1}| \leq \theta |a_n| \quad \forall n \geq n_0 \quad (214)$$

Für alle $n \geq n_0$ gilt dann

$$|a_n| \underset{(214)}{\leq} \theta |a_{n-1}| \underset{(214)}{\leq} \theta^2 |a_{n-2}| \underset{(214)}{\leq} \dots \underset{(214)}{\leq} \theta^{n-n_0} |a_{n-(n-n_0)}| = \theta^{n-n_0} |a_{n_0}|.$$

Mit anderen Worten gilt $|a_n| \leq \theta^{n-n_0} |a_{n_0}|$ für alle $n \geq n_0$ ^a. In unserer Sprache heißt das

$$(|a_n|)_{n \in \mathbb{N}} \leq (\theta^{n-n_0} |a_{n_0}|)_{n \in \mathbb{N}}. \quad (215)$$

Nun ist

$$\sum_{k \geq 1} \theta^{k-n_0} |a_{n_0}| = \sum_{k \geq 1} \underbrace{\frac{|a_{n_0}|}{\theta^{n_0}}}_{=: c} \theta^k = \sum_{k \geq 1} c \cdot \theta^k$$

ein konstantes Vielfaches einer geometrischen Reihe mit $z = \theta < 1$ und damit eine konvergente Reihe, vgl. Beispiel 5.0.5. Diese Beobachtung, zusammen mit (215) impliziert die absolute Konvergenz von $\sum_{k \geq 1} a_k$ nach dem Majorantenkriterium.

Zu (ii). Dieses Argument skizzieren wir nur. Ist $\liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| > 1$ so muss^b es ein $n_0 \in \mathbb{N}, n_0 \geq N$ geben mit

$$\forall n \geq n_0 : \quad \left| \frac{a_{n+1}}{a_n} \right| \geq 1.$$

Insbesondere gilt für alle $n \geq n_0$, dass $|a_{n+1}| \geq |a_n|$. Eine Folgerung daraus ist dass

$$\forall n \geq n_0 : \quad |a_n| \geq |a_{n_0}|. \quad (216)$$

Da $n_0 \geq N$ gilt $|a_{n_0}| > 0$. Nutzt man dies und (216) so sieht man: $(a_n)_{n \in \mathbb{N}}$ kann nicht gegen 0 konvergieren. Nach dem Abkling-Kriterium (Satz 5.1.2) folgt die Divergenz von $\sum_{k \geq 1} a_k$.

^aRigoros müsste man diese Ungleichung noch mit vollständiger Induktion zeigen. Dies ist den Lesenden als Übungsaufgabe überlassen

^bIn der Tat: Falls es ein solches n_0 nicht gibt, so müsste $A := \left\{ n \in \mathbb{N} : \left| \frac{a_{n+1}}{a_n} \right| < 1 \right\}$ eine unendliche Menge sein. Wählt man sich dann eine Teilfolge $(\ell_n)_{n \in \mathbb{N}} \subseteq A, \ell_{n+1} \geq \ell_n + 1$, so gilt $\ell_n \geq n$ und $\left| \frac{a_{\ell_{n+1}}}{a_{\ell_n}} \right| \leq 1$. Daraus folgern wir

$$\inf \left\{ \left| \frac{a_{k+1}}{a_k} \right| : k \geq n \right\} \leq \left| \frac{a_{\ell_{n+1}}}{a_{\ell_n}} \right| < 1$$

und somit

$$\liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \lim_{n \rightarrow \infty} \inf \left\{ \left| \frac{a_{k+1}}{a_k} \right| : k \geq n \right\} \leq 1.$$

Ein Widerspruch.

Beispiel 5.1.23

Behauptung. Die Reihe $\sum_{k \geq 1} \frac{k^2}{2^k}$ ist eine absolut konvergente Reihe.

Beweis. Definiere für $n \in \mathbb{N}$ den Wert $a_n := \frac{n^2}{2^n}$ und berechne

$$\begin{aligned} \left| \frac{a_{n+1}}{a_n} \right| &= \frac{\frac{(n+1)^2}{2^{n+1}}}{\frac{n^2}{2^n}} = \frac{(n+1)^2 2^n}{n^2 2^{n+1}} = \frac{(n+1)^2}{n^2} 2^{n-(n+1)} = \frac{(n+1)^2}{n^2} \frac{1}{2} = \left(\frac{n+1}{n} \right)^2 \frac{1}{2} \\ &= \frac{n+1}{n} \cdot \frac{n+1}{n} \cdot \frac{1}{2} = \underbrace{\left(1 + \frac{1}{n} \right)}_{\rightarrow 1} \underbrace{\left(1 + \frac{1}{n} \right)}_{\rightarrow 1} \frac{1}{2} \xrightarrow{n \rightarrow \infty} \frac{1}{2}. \end{aligned}$$

Wir folgern (da der Limes Superior im Falle einer konvergenten Folge mit dem Grenzwert übereinstimmt, vgl. Korollar 3.5.30)

$$\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \frac{1}{2} < 1.$$

Nach dem Quotientenkriterium folgt die Konvergenz der zu untersuchenden Reihe.

Beispiel 5.1.24

In diesem Beispiel geht es darum, zu zeigen, dass in Teil (ii) des Quotientenkriterium

5.1.22 $\liminf$ nicht durch $\limsup$ ersetzt werden kann. Betrachte dazu

$$a_n := \begin{cases} \left(\frac{1}{3}\right)^n & \text{falls } n \text{ gerade} \\ \left(\frac{1}{2}\right)^n & \text{falls } n \text{ ungerade} \end{cases}$$

Die Reihe $\sum_{k \geq 1} a_k$ konvergiert absolut nach dem Majorantenkriterium, denn $|a_n| \leq \left(\frac{1}{2}\right)^n$ für alle $n \in \mathbb{N}$ und $\sum_{k \geq 1} \left(\frac{1}{2}\right)^k$ ist eine konvergente Reihe, vgl. Beispiel 5.0.5. Allerdings ist

$$\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \infty > 1! \quad (217)$$

In der Tat: Wir können zeigen, dass $\left(\left| \frac{a_{n+1}}{a_n} \right| \right)_{n \in \mathbb{N}}$ nach oben unbeschränkt ist, denn

$$\left| \frac{a_{2n+1}}{a_{2n}} \right| = \left| \frac{\left(\frac{1}{2}\right)^{2n+1}}{\left(\frac{1}{3}\right)^{2n}} \right| = \frac{3^{2n}}{2^{2n+1}} = \frac{1}{2} \left(\frac{3}{2} \right)^{2n}.$$

Da der Ausdruck auf der rechten Seite offensichtlich unbeschränkt ist folgt (217) direkt aus Definition 3.5.24.

5.2 Potenzreihen

Die Reihen, die wir im folgenden betrachten, sind besonders nützlich, da man mit ihnen die meisten für uns gängigen Funktionen beschreiben kann. Wie genau das funktioniert werden wir im Folgenden sehen.

Definition 5.2.1 — POTENZREIHEN

Es sei $(c_n)_{n \geq 0} \subseteq \mathbb{C}$ und $z_0 \in \mathbb{C}$. Eine *Potenzreihe* mit *Referenzpunkt* $z_0 \in \mathbb{C}$ ist eine von einem Parameter $z \in \mathbb{C}$ abhängige Reihe der Form

$$\sum_{k \geq 0} c_k (z - z_0)^k.$$

Wir nennen ferner

$$D_K := \left\{ z \in \mathbb{C} : \sum_{k \geq 0} c_k (z - z_0)^k \text{ ist eine konvergente Reihe} \right\}$$

den *Konvergenzbereich* der Potenzreihe und

$$D_A := \left\{ z \in \mathbb{C} : \sum_{k \geq 0} c_k (z - z_0)^k \text{ ist eine absolut konvergente Reihe} \right\}$$

den *absoluten Konvergenzbereich* der Potenzreihe. Außerdem definieren wir die *Wertefunk-*

tion $P : D_K \rightarrow \mathbb{C}$ durch

$$P(z) := \sum_{k=0}^{\infty} c_k (z - z_0)^k \quad (z \in D_K).$$

Beispiel 5.2.2 — DIE GEOMETRISCHE REIHE

Ein Beispiel für eine Potenzreihe kennen wir schon — die *geometrische Reihe*, vgl. Beispiel 5.0.5. Es ist nämlich

$$\sum_{k \geq 0} z^k = \sum_{k \geq 0} 1 \cdot (z - 0)^k$$

eine Potenzreihe mit $(c_n)_{n \geq 0} = (1)_{n \geq 0}$ und Referenzpunkt $z_0 = 0$. Wir berechnen als nächstes D_K und D_A .

Zuerst zu D_K . In Beispiel 5.1.5 und 5.0.5 haben wir gesehen, dass

$$\sum_{k \geq 0} z^k \quad \begin{cases} \text{ist eine konvergente Reihe} & \text{falls } |z| < 1, \\ \text{ist eine divergente Reihe} & \text{falls } |z| \geq 1. \end{cases} \quad (218)$$

Somit gilt

$$D_K = \{z \in \mathbb{C} : |z| < 1\}.$$

Nun zu D_A . Wir verwenden Definition 5.1.10 und untersuchen daher das Konvergenzverhalten von

$$\sum_{k \geq 0} |z^k| = \sum_{k \geq 0} |z|^k.$$

Wir sehen, dass es sich wiederum um eine geometrische Reihe (mit $z' = |z|$) handelt. Mit derselben Erkenntnis wie in (218) erhalten wir

$$\sum_{k \geq 0} |z|^k \quad \begin{cases} \text{ist eine konvergente Reihe} & \text{falls } |z| < 1, \\ \text{ist eine divergente Reihe} & \text{falls } |z| \geq 1. \end{cases}$$

Mit anderen Worten

$$D_A = \{z \in \mathbb{C} : |z| < 1\}.$$

Man hat also hier $D_A = D_K = \{z \in \mathbb{C} : |z| < 1\}$. Auf dieser Menge kennen wir dank Beispiel 5.0.5 auch die Wertefunktion, nämlich gilt

$$P(z) = \sum_{k=0}^{\infty} z^k = \frac{1}{1-z} \quad \forall z \in \mathbb{C} : |z| < 1.$$

Wir werden später sehen, dass im Allgemeinen D_A nicht zwingend mit D_K übereinstimmt. Jedoch gilt stets $D_A \subseteq D_K$, weil nach Proposition 5.1.11 jede absolut konvergente Reihe auch konvergent ist.

Der Konvergenzbereich $\{z \in \mathbb{C} : |z| < 1\}$ aus dem vorigen Beispiel ist eine *Kreisscheibe* in $\mathbb{C}$. Dies ist kein Zufall, wie wir in der folgenden Proposition sehen werden.

18.12.2025

Proposition 5.2.3 — BEOBACHTUNG VON CAUCHY-HADAMARD

Es sei eine Potenzreihe

$$\sum_{k \geq 0} c_k (z - z_0)^k$$

gegeben. Definiere nun

$$R := \begin{cases} 0 & \limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|} = \infty \\ \frac{1}{\limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|}} & \limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|} \in (0, \infty) \\ \infty & \limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|} = 0. \end{cases} \quad (219)$$

Dann gilt

$$\{z \in \mathbb{C} : |z - z_0| < R\} \subseteq D_A \subseteq D_K \subseteq \{z \in \mathbb{C} : |z - z_0| \leq R\}. \quad (a) \quad (b)$$

Anders ausgedrückt

$$\sum_{k \geq 0} c_k (z - z_0)^k \quad \begin{cases} \text{ist eine absolut konvergente Reihe} & \text{falls } |z - z_0| < R, \\ \text{ist eine divergente Reihe} & \text{falls } |z - z_0| > R. \end{cases}$$

^aIm Fall $R = \infty$ sind alle Mengen gleich $\mathbb{C}$

^bDie Menge $\{z \in \mathbb{C} : |z - z_0| < R\}$ nennen wir manchmal auch *Konvergenzscheibe*: Es handelt sich um die größtmögliche Kreisscheibe im Konvergenzbereich.

Beweis. Es sei R wie in (219).

FALL 1. $R \in (0, \infty)$

Wir wissen aus dem Wurzelkriterium (Satz 5.1.20), dass

$$\sum_{k \geq 0} c_k (z - z_0)^k \quad \begin{cases} \text{ist eine abs. konvergente Reihe} & \text{falls } \limsup_{n \rightarrow \infty} \sqrt[n]{|c_n (z - z_0)^n|} < 1, \\ \text{ist eine divergente Reihe} & \text{falls } \limsup_{n \rightarrow \infty} \sqrt[n]{|c_n (z - z_0)^n|} > 1. \end{cases} \quad (220)$$

Nun gilt

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|c_n (z - z_0)^n|} = \limsup_{n \rightarrow \infty} \sqrt[n]{|c_n| |z - z_0|^n} = \limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|} |z - z_0|.$$

Da $R \in (0, \infty)$ sind wir im mittleren Fall von (219) und daher gilt

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|} |z - z_0| \stackrel{(219)}{=} \frac{|z - z_0|}{R} \quad \begin{cases} < 1 & \text{falls } |z - z_0| < R \\ > 1 & \text{falls } |z - z_0| > R \end{cases}.$$

Mit (220) folgern wir

$$\sum_{k \geq 0} c_k (z - z_0)^k \quad \begin{cases} \text{ist eine absolut konvergente Reihe} & \text{falls } |z - z_0| < R, \\ \text{ist eine divergente Reihe} & \text{falls } |z - z_0| > R. \end{cases}$$

Die erste Zeile zeigt nun $\{z \in \mathbb{C} : |z - z_0| < R\} \subseteq D_A$ und die zweite Zeile zeigt $D_K \subseteq \{z \in \mathbb{C} : |z - z_0| \leq R\}$. Dass $D_A \subseteq D_K$ folgt dann wie besprochen noch aus Proposition 5.1.11.

FALL 2 BZW. FALL 3. $R = \infty$ bzw. $R = 0$.

Überlassen wir den Lesenden als Übungsaufgabe.

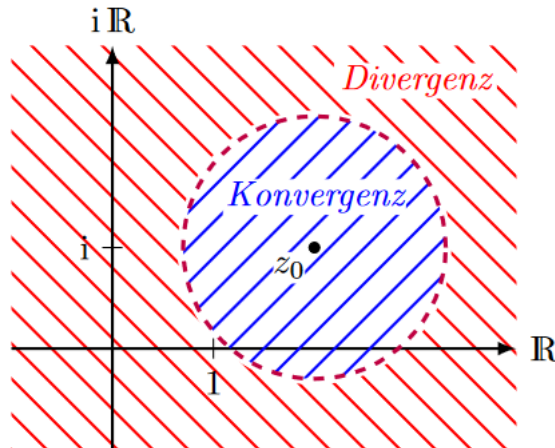


Abbildung 18: Konvergenzverhalten von Potenzreihen: In der blauen Kreisscheibe hat man (sogar absolute) Konvergenz. Im roten Bereich hat man Divergenz. Nur auf dem violetten Rand ist unklar was passiert. Der Radius der blauen Kreisscheibe ist gerade R aus (219)

Definition 5.2.4 — KONVERGENZRADIUS

Die Zahl $R \in [0, \infty) \cup \{\infty\}$ ^a aus (219) nennen wir *Konvergenzradius* der Potenzreihe $\sum_{k \geq 0} c_k(z - z_0)^k$.

^aVon jetzt an schreiben wir auch als Kurzschreibweise $[0, \infty] := [0, \infty) \cup \{\infty\}$

Durch Bestimmung des Konvergenzradius können wir den Konvergenzbereich bereits sehr gut *eingrenzen*: Man weiß, dass die Konvergenzscheibe $\{z \in \mathbb{C} : |z - z_0| < R\}$ im Konvergenzbereich drin liegt und $\{z \in \mathbb{C} : |z - z_0| > R\}$ draußen liegt. Lediglich der Rand der Konvergenzscheibe muss dann noch gesondert untersucht werden.

Mit diesem Verfahren können wir D_K und D_A für sehr viele Potenzreihen bestimmen.

Als Anwendung können wir auch endlich eine Potenzreihe sehen, für die $D_K \neq D_A$ gilt.

Beispiel 5.2.5

Es sei die Potenzreihe

$$\sum_{k \geq 0} \frac{z^k}{k+1}$$

gegeben (d.h. $z_0 = 0$ und $c_k = \frac{1}{k+1}$ für alle $k \in \mathbb{N}$).

Behauptung. $D_A = \{z \in \mathbb{C} : |z| < 1\}$ und $D_K \subsetneq D_A$.

Beweis. Wir unterteilen den Beweis in zwei Schritte.

SCHRITT 1. Konvergenzradius R bestimmen.

Wir folgern zunächst leicht aus dem Sandwichlemma, dass

$$\lim_{n \rightarrow \infty} \sqrt[n]{n+1} = 1. \quad (a)$$

Deswegen ist auch

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|} = \limsup_{n \rightarrow \infty} \sqrt[n]{\frac{1}{n+1}} = \limsup_{n \rightarrow \infty} \frac{1}{\sqrt[n]{n+1}} \stackrel{\text{Prop 3.5.30}}{=} \lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{n+1}} \stackrel{\text{GWR}}{=} \frac{1}{1} = 1.$$

Der Konvergenzradius ist also nach (219)

$$R = \frac{1}{\limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|}} = \frac{1}{1} = 1.$$

Man folgert aus der Proposition 5.2.3 dass

$$\{z \in \mathbb{C} : |z| < 1\} \subseteq D_A \subseteq D_K \subseteq \{z \in \mathbb{C} : |z| \leq 1\}$$

Jetzt, wo wir die Konvergenzbereiche schon so arg eingegrenzt haben, bleibt nur noch zu untersuchen, was für ein Konvergenzverhalten im *Randbereich der Konvergenzscheibe* $\{z \in \mathbb{C} : |z| = 1\}$ vorliegt.

SCHRITT 2. Untersuchung des *Randbereiches* $\{z \in \mathbb{C} : |z| = R\}$.

SCHRITT 2.1 Bestimmung von D_A

Zuerst untersuchen wir auf absolute Konvergenz (mithilfe von Definition 5.1.10). Ist $z \in \mathbb{C}$ mit $|z| = R = 1$ so gilt

$$\sum_{k \geq 0} \left| \frac{z^k}{k+1} \right| = \sum_{k \geq 0} \frac{|z^k|}{k+1} = \sum_{k \geq 0} \frac{|z|^k}{k+1} = \sum_{k \geq 0} \frac{1^k}{k+1} = \sum_{k \geq 0} \frac{1}{k+1} \quad (221)$$

Für die Partialsummen dieser Reihe gilt

$$s_m := \sum_{k=0}^m \frac{1}{k+1} \stackrel{\text{Indexverschiebung}}{=} \sum_{k=1}^{m+1} \frac{1}{k} \quad \forall m \in \mathbb{N}.$$

Der Ausdruck auf der rechten Seite ist die $(m+1)$ -ste Partialsumme der harmonischen Reihe. Da diese Partialsummen unbeschränkt sind folgt auch, dass $(s_m)_{m \in \mathbb{N}}$ unbeschränkt ist. Somit ist $\sum_{k \geq 0} \frac{1}{k+1}$ eine divergente Reihe. Es folgt

$$D_A = \{z \in \mathbb{C} : |z| < 1\}$$

SCHRITT 2.2 Bestimmung von D_K

Dies ist deutlich komplizierter als Schritt 2.1! Für D_A mussten wir im Fall $|z| = 1$ stets nur eine Reihe betrachten, wie man an der Rechnung (221) sieht. Für D_K müsste man aber *unendlich viele Reihen* betrachten, für jeden Punkt auf $\{z \in \mathbb{C} : |z| < 1\}$ kann etwas anderes herauskommen...

Oft begnügt man sich daher mit $D_K \cap \mathbb{R}$ — hier muss man nämlich nur zwei Reihen betrachten. In unserem Fall sind das die Reihen für $z = 1$ und $z = -1$. Für $z = 1$ hat man

$$\sum_{k \geq 0} \frac{1^k}{k+1} = \sum_{k \geq 0} \frac{1}{k+1}.$$

Dass diese Reihe divergiert, haben wir bereits in Schritt 2.1 gesehen. Für $z = -1$ ist

$$\sum_{k \geq 0} \frac{(-1)^k}{k+1}$$

eine konvergente Reihe, wie man mit dem Leibniz-Kriterium oder auch mit einer Indexverschiebung und Proposition 5.1.14 (i) leicht überprüfen kann. Wir überlassen die Details den Lesenden als Übungsaufgabe.

Festhalten kann man nun

$$D_K \supsetneq D_A \quad \text{denn} \quad -1 \in D_K \setminus D_A.$$

Mit etwas mehr kann man sogar den gesamten Konvergenzbereich D_K bestimmen und sehen dass $D_K = \{z \in \mathbb{C} : |z| \leq 1\} \setminus \{1\}$.

^aIn der Tat: Für $n \in \mathbb{N}$ gilt $\sqrt[n]{n} \leq \sqrt[n]{n+1} \leq \sqrt[n]{2n} = \sqrt[n]{2} \sqrt[n]{n}$ und beide Seiten der Ungleichung streben gegen 1 für $n \rightarrow \infty$ (vgl. Proposition 2.3.7).

Eine spannende Beobachtung, die uns bei der Untersuchung von Potenzreihen oft von Nutzen sein wird ist die Folgende:

Proposition 5.2.6

Es sei eine Potenzreihe $\sum_{k \geq 0} c_k(z - z_0)^k$ mit Konvergenzradius $R \in [0, \infty]$ gegeben. Dann haben für alle $j \in \mathbb{N} \cup \{0\}$ auch die Potenzreihen

$$\sum_{k \geq 0} k^j c_k(z - z_0)^k \quad \text{und} \quad \sum_{k \geq 0} k^j |c_k| (z - z_0)^k$$

denselben Konvergenzradius R .

Beweis. Wir zeigen die Behauptung nur für $\sum_{k \geq 0} k^j c_k(z - z_0)^k$ im Spezialfall $R \in (0, \infty)$. Der Rest bleibt den Lesenden als Übungsaufgabe überlassen. Es sei $\tilde{R}$ der Konvergenzradius von $\sum_{k \geq 0} k^j c_k(z - z_0)^k$. Definiere außerdem für $k \in \mathbb{N}$ die Koeffizienten $\tilde{c}_k := k^j c_k$, d.h. unsere Potenzreihe ist gegeben durch $\sum_{k \geq 0} \tilde{c}_k(z - z_0)^k$.

$$\mathbb{Z}: \quad \tilde{R} = R.$$

Zunächst erstmal folgt aus (219), dass

$$\frac{1}{R} = \limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|} \quad (222)$$

$$\text{Zwischenbehauptung:} \quad \limsup_{n \rightarrow \infty} \sqrt[n]{|\tilde{c}_n|} = \frac{1}{R} \in (0, \infty)$$

Dazu: Für $n \in \mathbb{N}$ gilt

$$\sqrt[n]{|\tilde{c}_n|} = \sqrt[n]{n^j |c_n|} = \sqrt[n]{n^j} \sqrt[n]{|c_n|} = (\sqrt[n]{n})^j \sqrt[n]{|c_n|}. \quad (223)$$

Beachte: Nach den Grenzwertregeln gilt

$$\lim_{n \rightarrow \infty} (\sqrt[n]{n})^j = 1^j = 1.$$

Selbiges gilt dann auch für jede Teilfolge, vgl Proposition 3.5.13. Ist nun h ein Häufungswert von $(\sqrt[n]{|c_n|})_{n \in \mathbb{N}}$, d.h. ein Grenzwert einer Teilfolge $(\sqrt[\ell]{|c_{\ell_n}|})_{n \in \mathbb{N}}$, so gilt auch

$$\lim_{n \rightarrow \infty} \sqrt[\ell]{|\tilde{c}_{\ell_n}|} = \lim_{n \rightarrow \infty} \underbrace{(\sqrt[\ell]{\ell_n})^j}_{\rightarrow 1} \underbrace{\sqrt[\ell]{|c_{\ell_n}|}}_{\rightarrow h} = 1 \cdot h = h.$$

Also ist h auch ein Häufungswert von $(\sqrt[n]{|\tilde{c}_n|})_{n \in \mathbb{N}}$. Analog zeigt man, dass jeder Häufungswert von $(\sqrt[n]{|\tilde{c}_n|})_{n \in \mathbb{N}}$ auch einer von $(\sqrt[n]{|c_n|})_{n \in \mathbb{N}}$ sein muss. Man sieht also:

$$(\sqrt[n]{|\tilde{c}_n|})_{n \in \mathbb{N}} \text{ und } (\sqrt[n]{|c_n|})_{n \in \mathbb{N}} \text{ haben dieselben Häufungswerte.}$$

Da beide Folgen beschränkt sind^a ist der $\limsup$ nichts anderes als der größte Häufungswert, vgl. Proposition 3.5.29. Man folgert aus der Gleichheit aller Häufungswerte, dass

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|\tilde{c}_n|} = \limsup_{n \rightarrow \infty} \sqrt[n]{|c_n|} \stackrel{(222)}{=} \frac{1}{R}.$$

Die Zwischenbehauptung ist gezeigt.

Aus der Zwischenbehauptung folgt tatsächlich die Aussage der Proposition, denn mit (222) erhält man

$$\tilde{R} = \frac{1}{\limsup_{n \rightarrow \infty} \sqrt[n]{|\tilde{c}_n|}} \stackrel{ZwBeh}{=} \frac{1}{\frac{1}{R}} = R.$$

^awie man sich anhand von (222) und (223) leicht überlegen kann (Übungsaufgabe!)

Zum Schluss lernen wir noch eine andere Formel kennen, um den Konvergenzradius auszurechnen, die sogenannte *Konvergenzradiusformel von Euler*. Ganz ähnlich wie Proposition 5.2.3 (nur mit dem Quotientenkriterium anstatt Wurzelkriterium) zeigt man

Proposition 5.2.7 — FORMEL VON EULER

Es sei

$$\sum_{k \geq 0} c_k (z - z_0)^k$$

mit Konvergenzradius $R \in [0, \infty]$ gegeben (siehe (219)). Ferner sei die Folge $(|\frac{c_{k+1}}{c_k}|)_{k \in \mathbb{N}}$ konvergent oder bestimmt divergent. Definiere

$$\hat{R} = \begin{cases} 0 & \text{falls } \lim_{k \rightarrow \infty} |\frac{c_{k+1}}{c_k}| = \infty \\ \frac{1}{\lim_{k \rightarrow \infty} |\frac{c_{k+1}}{c_k}|} & \text{falls } \lim_{k \rightarrow \infty} |\frac{c_{k+1}}{c_k}| \in (0, \infty) \\ \infty & \text{falls } \lim_{k \rightarrow \infty} |\frac{c_{k+1}}{c_k}| = 0 \end{cases} \quad (224)$$

Dann gilt $\hat{R} = R$.

Beweisidee. Man zeigt (Übungsaufgabe) mithilfe des Quotientenkriteriums

$$\{z \in \mathbb{C} : |z - z_0| < \hat{R}\} \subseteq D_A \subseteq D_K \subseteq \{z \in \mathbb{C} : |z - z_0| \leq \hat{R}\}.$$

Dies und Proposition 5.2.3 impliziert $\hat{R} = R$. Beachte: Im Gegensatz zur Formel von Cauchy-Hadamard (Proposition 5.2.3) muss hier der Limes $\lim_{k \rightarrow \infty} |\frac{c_{k+1}}{c_k}|$ existieren, ein Limes Superior genügt nicht. Das liegt daran, dass im Quotientenkriterium (Satz 5.1.22) auch nicht in beiden Fällen der Limes Superior verwendet wird, sondern nur im Konvergenzfall (i).

5.3 Das Cauchy-Produkt

Als nächstes soll es um eine *Produktformel* für konvergente Reihen gehen.

Satz 5.3.1 — CAUCHY-PRODUKT VON REIHEN

Es seien $\sum_{k \geq 0} a_k$ und $\sum_{k \geq 0} b_k$ absolut konvergente Reihen. Man definiere für $k \in \mathbb{N} \cup \{0\}$

$$c_k := \sum_{\ell=0}^k a_\ell b_{k-\ell}. \quad (225)$$

Dann ist $\sum_{k \geq 0} c_k$ absolut konvergent und

$$\sum_{k=0}^{\infty} c_k = \left(\sum_{k=0}^{\infty} a_k \right) \cdot \left(\sum_{k=0}^{\infty} b_k \right). \quad (226)$$

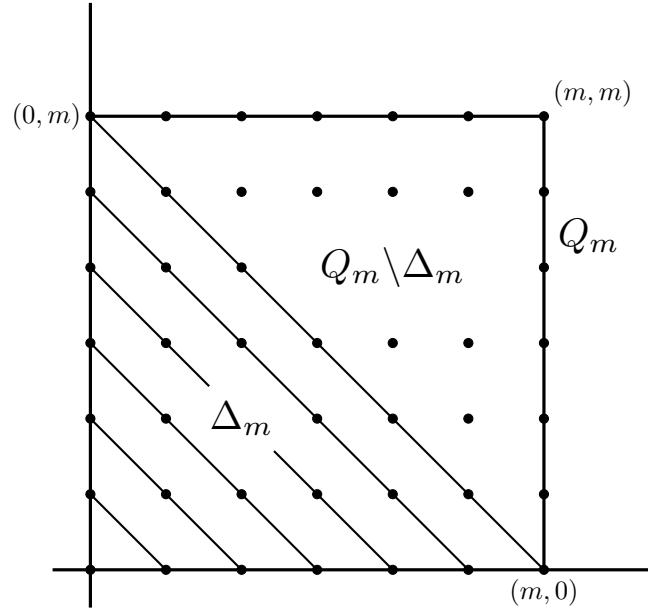


Abbildung 19: Abbildung zum Beweis von Satz 5.3.1.

***Beweisskizze.** Zu (226). Wir definieren für $m \in \mathbb{N}$ die Partialsummen

$$A_m := \sum_{k=0}^m a_k \quad \text{sowie} \quad A := \sum_{k=0}^{\infty} a_k [= \lim_{m \rightarrow \infty} A_m]$$

und

$$B_m := \sum_{k=0}^m b_k \quad \text{sowie} \quad B := \sum_{k=0}^{\infty} b_k [= \lim_{m \rightarrow \infty} B_m]$$

sowie

$$\tilde{A}_m := \sum_{k=0}^m |a_k| \quad \text{sowie} \quad \tilde{A} := \sum_{k=0}^{\infty} |a_k| [= \lim_{m \rightarrow \infty} \tilde{A}_m],$$

$$\tilde{B}_m := \sum_{k=0}^m |b_k| \quad \text{sowie} \quad \tilde{B} := \sum_{k=0}^{\infty} |b_k| [= \lim_{m \rightarrow \infty} \tilde{B}_m],$$

und

$$C_m := \sum_{k=0}^m c_k$$

SCHRITT 1. RECHTECKE UND DREIECKE DEFINIEREN.

Wir definieren $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ und

$$Q_m := \{(j, k) \in \mathbb{N}_0 \times \mathbb{N}_0 : 0 \leq j \leq m, 0 \leq k \leq m\}$$

sowie

$$D_m := \{(j, k) \in \mathbb{N}_0 \times \mathbb{N}_0 : j + k = m\}$$

und

$$\Delta_m := D_0 \cup D_1 \cup \dots \cup D_m = \{(j, k) \in \mathbb{N}_0 \times \mathbb{N}_0 : j + k \leq m\}.$$

Man beachte: Für gerade m gilt stets

$$Q_{\frac{m}{2}} \subseteq \Delta_m, \quad (227)$$

denn

$$(j, k) \in Q_{\frac{m}{2}} \implies j \leq \frac{m}{2}, k \leq \frac{m}{2} \implies j + k \leq \frac{m}{2} + \frac{m}{2} = m.$$

SCHRITT 2. UMSCHREIBEN VON C_m .

Man beachte: Für $n \in \mathbb{N}$ ist

$$c_n \stackrel{(225)}{=} \sum_{\ell=0}^n a_\ell b_{n-\ell} = a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \dots + a_n b_0 = \sum_{(j,k) \in D_n} a_j b_k,$$

wobei hier $\sum_{(j,k) \in D_n}$ meint, dass über alle infrage kommenden Indizes summiert werden muss.

Somit ist auch

$$C_m = \sum_{n=0}^m c_n = \sum_{n=0}^m \sum_{(j,k) \in D_n} a_j b_k = \sum_{(j,k) \in \Delta_m} a_j b_k.$$

SCHRITT 3. UMSCHREIBEN VON $A_m B_m$.

Ferner gilt für alle $m \in \mathbb{N}$

$$\begin{aligned} A_m B_m &= \left(\sum_{j=0}^m a_j \right) \left(\sum_{k=0}^m b_k \right) = (a_0 + \dots + a_m)(b_0 + \dots + b_m) \\ &= a_0(b_0 + \dots + b_m) + a_1(b_0 + \dots + b_m) + \dots + a_m(b_0 + \dots + b_m) \\ &= \begin{aligned} &a_0 b_0 + \dots + a_0 b_m \\ &+ a_1 b_0 + \dots + a_1 b_m \\ &\vdots \\ &+ a_m b_0 + \dots + a_m b_m \end{aligned} \end{aligned}$$

$$= \sum_{(j,k) \in Q_m} a_j b_k.$$

Analog kann man übrigens zeigen

$$\tilde{A}_m \tilde{B}_m = \sum_{(j,k) \in Q_m} |a_j| |b_k| \quad (228)$$

Somit gilt

$$\begin{aligned} |A_m B_m - C_m| &= \left| \sum_{(j,k) \in Q_m} a_j b_k - \sum_{(j,k) \in \Delta_m} a_j b_k \right| = \left| \sum_{(j,k) \in Q_m \setminus \Delta_m} a_j b_k \right| \\ &\leq \sum_{(j,k) \in Q_m \setminus \Delta_m} |a_j| |b_k| \end{aligned} \quad (229)$$

Wegen (227) gilt für m gerade

$$\begin{aligned} \sum_{(j,k) \in Q_m \setminus \Delta_m} |a_j| |b_k| &\leq \sum_{(j,k) \in Q_m \setminus Q_{\frac{m}{2}}} |a_j| |b_k| = \sum_{(j,k) \in Q_m} |a_j| |b_k| - \sum_{(j,k) \in Q_{\frac{m}{2}}} |a_j| |b_k| \\ &\stackrel{(228)}{=} \tilde{A}_m \tilde{B}_m - \tilde{A}_{\frac{m}{2}} \tilde{B}_{\frac{m}{2}} \end{aligned}$$

Dies und (229) implizieren für gerade $m \in \mathbb{N}$

$$0 \leq |A_m B_m - C_m| \leq \tilde{A}_m \tilde{B}_m - \tilde{A}_{\frac{m}{2}} \tilde{B}_{\frac{m}{2}} \quad (230)$$

Da

$$\lim_{m \rightarrow \infty} 0 = 0 \quad \text{und} \quad \lim_{m \rightarrow \infty} (\tilde{A}_m \tilde{B}_m - \tilde{A}_{\frac{m}{2}} \tilde{B}_{\frac{m}{2}}) = \tilde{A} \tilde{B} - \tilde{A} \tilde{B} = 0 \quad (a)$$

folgt aus dem Sandwichlemma und (230), dass

$$\lim_{m \rightarrow \infty} |A_m B_m - C_m| = 0.$$

Wir folgern, dass

$$C_m = \underbrace{C_m - A_m B_m}_{\rightarrow 0} + \underbrace{A_m B_m}_{\rightarrow AB \text{ [GWR2]}} \xrightarrow[\text{GWR}]{n \rightarrow \infty} 0 + AB = \left(\sum_{k=0}^{\infty} a_k \right) \cdot \left(\sum_{k=0}^{\infty} b_k \right)$$

Mit anderen Worten

$$\sum_{k=0}^{\infty} c_k = \lim_{m \rightarrow \infty} C_m = \left(\sum_{k=0}^{\infty} a_k \right) \cdot \left(\sum_{k=0}^{\infty} b_k \right).$$

Gleichung (226) ist gezeigt. Dass $\sum_{k \geq 0} c_k$ absolut konvergent ist, überlassen wir den Lesenden als Übungsaufgabe.

^aEigentlich müsste man hier auf die Teilfolge $\tilde{j}_m = 2m$ umsteigen, damit man nur die geraden Indizes durchzählt. Um die Notation schlank zu halten werden wir das aber nicht vertiefen.

5.4 Die Exponentialfunktion

Als nächstes werden wir die *Exponentialfunktion* als Potenzreihe einführen und danach ihre Eigenschaften studieren. In der Realität ist die Exponentialfunktion nichts anderes als die aus der Schule bekannte *e*-Funktion. Mit einem wichtigen Unterschied: Wir können auch komplexwertige Argumente zulassen!

Satz 5.4.1 — EXPONENTIALREIHE

Die Potenzreihe

$$\sum_{k \geq 0} \frac{z^k}{k!}$$

hat Konvergenzradius $R = \infty$, insbesondere gilt also $D_K = \mathbb{C}$. Ihre Wertefunktion ist also auf ganz $\mathbb{C}$ definiert.

Beweis. Wir verwenden Proposition 5.2.7 (mit $(c_n)_{n \in \mathbb{N}} = (\frac{1}{n!})_{n \in \mathbb{N}}$ und $z_0 = 0$) um $R(= \hat{R})$ zu berechnen. Dazu betrachte

$$\left| \frac{c_{k+1}}{c_k} \right| = \frac{\frac{1}{(k+1)!}}{\frac{1}{k!}} = \frac{k!}{(k+1)!} = \frac{1}{k+1} \xrightarrow{k \rightarrow \infty} 0$$

Mit Fall 3 in (224) erhalten wir $R = \hat{R} = \infty$. Dass die Wertefunktion auf ganz $\mathbb{C}$ definiert ist folgt dann aus Proposition 5.2.3.

Definition 5.4.2 — EXPONENTIALFUNKTION

Die Funktion $\exp : \mathbb{C} \rightarrow \mathbb{C}$ gegeben durch

$$\exp(z) := \sum_{k=0}^{\infty} \frac{z^k}{k!} \quad (z \in \mathbb{C}).$$

Beispiel 5.4.3

Es gilt

$$\exp(0) = \sum_{k=0}^{\infty} \frac{0^k}{k!} = \frac{0^0}{0!} + \sum_{k=1}^{\infty} \frac{0^k}{k!} = 1 + \sum_{k=1}^{\infty} 0 = 1.$$

Wir werden später auch noch sehen, dass $\exp(1) = e$. Dies wird den Zusammenhang zwischen der hier definierten Exponentialfunktion und der evtl aus der Schule bekannten

e-Funktion ergründen.

Bemerkung 5.4.4

Auch wenn die Exponentialfunktion jetzt auf ganz $\mathbb{C}$ definiert ist: Man kann sie *auf eine reelle Funktion einschränken*. Dazu beachte: Für alle $x \in \mathbb{R}$ und alle $m \in \mathbb{N}$ ist

$$\sum_{k=0}^m \frac{x^k}{k!} \in \mathbb{R}, \quad \text{da alle Summanden reellwertig sind.}$$

Man überzeugt sich leicht davon (Übungsaufgabe), dass dann auch

$$\lim_{m \rightarrow \infty} \sum_{k=0}^m \frac{x^k}{k!} \in \mathbb{R}$$

Somit gilt

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \lim_{m \rightarrow \infty} \sum_{k=0}^m \frac{x^k}{k!} \in \mathbb{R}.$$

Es ist also $\exp(\mathbb{R}) \subseteq \mathbb{R}$. Somit können wir auch von der reellen Funktion

$$\exp : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \exp(x)$$

sprechen. Man spricht auch manchmal von $\exp_{\mathbb{R}}$ oder $\exp|_{\mathbb{R}}$ wenn man diese Funktion notationsmäßig von der komplexen Exponentialfunktion unterscheiden will.

Im Folgenden soll es aber vor Allem um die komplexe Exponentialfunktion gehen. Und doch werde ich Sie am Ende überzeugen, dass $\exp_{\mathbb{R}}$ genau die von Ihnen in der Schule studierte *e-Funktion* ist.

Funktionalgleichung und erste Eigenschaften Die wichtigste charakterisierende Eigenschaft der Exponentialfunktion ist die sog. *Funktionalgleichung*. Mit ihr werden wir auch die *Exponentialgesetze* und später die Rechenregeln für Potenzfunktionen herleiten.

Satz 5.4.5 — FUNKTIONALGLEICHUNG DER EXPONENTIALFUNKTION

Für alle $z_1, z_2 \in \mathbb{C}$ gilt

$$\exp(z_1 + z_2) = \exp(z_1) \exp(z_2).$$

Beweis. Es gilt

$$\exp(z_1) \exp(z_2) \stackrel{\text{Def 5.4.2}}{=} \left(\sum_{k=0}^{\infty} \frac{z_1^k}{k!} \right) \left(\sum_{k=0}^{\infty} \frac{z_2^k}{k!} \right) \stackrel{\text{Satz 5.3.1}}{=} \sum_{k=0}^{\infty} c_k \quad (231)$$

wobei

$$\begin{aligned} c_k &= \sum_{\ell=0}^k \frac{z_1^\ell}{\ell!} \frac{z_2^{k-\ell}}{(k-\ell)!} = \sum_{\ell=0}^k \frac{1}{\ell!(k-\ell)!} z_1^\ell z_2^{k-\ell} = \sum_{\ell=0}^k \frac{1}{k!} \frac{k!}{\ell!(k-\ell)!} z_1^\ell z_2^{k-\ell} \\ &= \frac{1}{k!} \sum_{\ell=0}^k \frac{k!}{\ell!(k-\ell)!} z_1^\ell z_2^{k-\ell} \stackrel{\text{Prop 2.2.14}}{=} \frac{1}{k!} \sum_{\ell=0}^k \binom{k}{\ell} z_1^\ell z_2^{k-\ell} \stackrel{\text{Satz 2.2.18}}{=} \frac{1}{k!} (z_1 + z_2)^k. \end{aligned} \quad (232)$$

Man hat also

$$\exp(z_1) \exp(z_2) \stackrel{(231)}{=} \sum_{k=0}^{\infty} c_k \stackrel{(232)}{=} \sum_{k=0}^{\infty} \frac{(z_1 + z_2)^k}{k!} \stackrel{\text{Def 5.4.2}}{=} \exp(z_1 + z_2).$$

Korollar 5.4.6

- (i) Für alle $z \in \mathbb{C}$ gilt $\exp(z) \neq 0$.
- (ii) Für alle $z \in \mathbb{C}$ gilt $\frac{1}{\exp(z)} = \exp(-z)$.

Beweis. Zu (i). Sei $z \in \mathbb{C}$ beliebig. Dann gilt

$$\exp(z) \exp(-z) \stackrel{\text{Satz 5.4.5}}{=} \exp(z + (-z)) = \exp(0) \stackrel{\text{Bsp 5.4.3}}{=} 1. \quad (233)$$

Dass $\exp(z) \neq 0$ folgt nun so: Wäre $\exp(z) = 0$, so wäre auch $\exp(z) \exp(-z) = 0 \cdot \exp(-z) = 0$. Ein Widerspruch zu (233).

Jetzt noch zu (ii). Teilt man in (233) durch $\exp(z)$, so folgt die Behauptung sofort.

Eine noch weitreichendere Erkenntnis als Korollar 5.4.6 ergibt sich für die *reelle Exponentialfunktion*. Hier sind die Werte nicht nur ungleich Null, sondern sogar *strikt positiv*.

Korollar 5.4.7 — POSITIVITÄT VON $\exp_{\mathbb{R}}$

Es sei $x \in \mathbb{R}$. Dann gilt $\exp(x) \in (0, \infty)_{\mathbb{R}}$.

Beweis. Dass $\exp(x) \in \mathbb{R}$ folgt aus Beispiel 5.4.4. Dass $\exp(x) \geq 0$ folgt so:

$$\exp(x) = \exp\left(\frac{x}{2} + \frac{x}{2}\right) = \exp\left(\frac{x}{2}\right) \exp\left(\frac{x}{2}\right) = [\exp\left(\frac{x}{2}\right)]^2.$$

Da $\exp(\frac{x}{2})$ eine reelle(!) Zahl ist, ist $[\exp(\frac{x}{2})]^2 \geq 0$ (vgl. Proposition 3.2.20 (x)). Wir erhalten $\exp(x) \geq 0$. Dass $\exp(x) > 0$ folgt dann schlussendlich aus Korollar 5.4.6 (i).

Als nächstes studieren wir das Verhalten der Exponentialfunktion unter komplexer Konjugation und Absolutbetrag.

22.12.2025

Proposition 5.4.8 — EXPONENTIALFUNKTION UND KONJUGATION

- (i) Für alle $z \in \mathbb{C}$ gilt $\overline{\exp(z)} = \exp(\bar{z})$.
- (ii) Für alle $z \in \mathbb{C}$ gilt $|\exp(z)| = \exp(\operatorname{Re}(z))$.
- (iii) Für alle $y \in \mathbb{R}$ gilt $|\exp(iy)| = 1$.

Beweis. Zu (i). Für $z \in \mathbb{C}$ gilt

$$\begin{aligned} \overline{\exp(z)} &= \overline{\sum_{k=0}^{\infty} \frac{z^k}{k!}} = \overline{\lim_{m \rightarrow \infty} \sum_{k=0}^m \frac{z^k}{k!}} \stackrel{\text{Prop 4.3.3}}{=} \lim_{m \rightarrow \infty} \overline{\sum_{k=0}^m \frac{z^k}{k!}} \stackrel{\text{Prop 4.2.3(ii)}}{=} \lim_{m \rightarrow \infty} \sum_{k=0}^m \overline{\left(\frac{z^k}{k!}\right)} \\ &\stackrel{\text{Prop 4.2.3(iii)}}{=} \lim_{m \rightarrow \infty} \sum_{k=0}^m \frac{\bar{z}^k}{k!} \stackrel{k! \in \mathbb{R}}{=} \lim_{m \rightarrow \infty} \sum_{k=0}^m \frac{\bar{z}^k}{k!} = \sum_{k=0}^{\infty} \frac{\bar{z}^k}{k!} = \exp(\bar{z}). \end{aligned}$$

Zu (ii). Für $z \in \mathbb{C}$ gilt

$$\begin{aligned} |\exp(z)|^2 &\stackrel{\text{Prop 4.2.6(ii)}}{=} \exp(z) \overline{\exp(z)} \stackrel{(i)}{=} \exp(z) \exp(\bar{z}) \stackrel{\text{Satz 5.4.5}}{=} \exp(z + \bar{z}) \\ &\stackrel{\text{Prop 4.2.2(i)}}{=} \exp(2\operatorname{Re}(z)) = \exp(\operatorname{Re}(z) + \operatorname{Re}(z)) \\ &\stackrel{\text{Satz 5.4.5}}{=} \exp(\operatorname{Re}(z)) \exp(\operatorname{Re}(z)) = \exp(\operatorname{Re}(z))^2 \end{aligned}$$

Ziehen wir auf beiden Seiten die Wurzel so erhalten wir

$$|\exp(z)| = |\exp(\operatorname{Re}(z))| \stackrel{\text{Korollar 5.4.7}}{\underset{\operatorname{Re}(z) \in \mathbb{R}}{=}} \exp(\operatorname{Re}(z)).$$

Zu (iii). Es gilt für $y \in \mathbb{R}$

$$|\exp(iy)| \stackrel{(ii)}{=} \exp(\operatorname{Re}(iy)) = \exp(\operatorname{Re}(0 + iy)) = \exp(0) \stackrel{\text{Bsp 5.4.3}}{=} 1.$$

Eine weitere nützliche Aussage ist das *Exponentialgesetz*, welches direkt aus der Funktionalgleichung folgt.

Proposition 5.4.9 — EXPONENTIALGESETZ

Es sei $z \in \mathbb{C}$. Dann gilt

$$\exp(z)^n = \exp(nz) \quad \forall n \in \mathbb{Z}$$

Beweis.

FALL 1. $n = 0$. Dann gilt

$$\exp(z)^0 = 1 = \exp(0) = \exp(0 \cdot z).$$

FALL 2. $n \in \mathbb{N}$. In diesem Fall zeigen wir die Behauptung mit vollständiger Induktion nach n .

IndA: $n = 1$.

$$\exp(z)^1 = \exp(z) = \exp(1 \cdot z) \quad \checkmark.$$

IndS: $n \longrightarrow n + 1$.

Es gelte für ein $n \in \mathbb{N}$ die Induktionshypothese

$$\exp(z)^n = \exp(nz) \quad (\text{IndH}).$$

$$\mathbb{Z}\mathbb{Z}: \quad \text{Dann gilt auch} \quad \exp(z)^{n+1} = \exp((n+1)z).$$

Dazu:

$$\exp(z)^{n+1} = \exp(z)^n \exp(z) \underset{\text{IndH}}{=} \exp(nz) \exp(z) \underset{\text{Satz 5.4.5}}{=} \exp(nz + z) = \exp((n+1)z).$$

Die Behauptung ist für alle $n \in \mathbb{N}$ gezeigt.

FALL 3: $n = -j$ FÜR EIN $j \in \mathbb{N}$. Dann gilt

$$\exp(z)^n = \exp(z)^{-j} = \frac{1}{\exp(z)^j} \underset{\text{Fall 2}}{=} \frac{1}{\exp(jz)} \underset{\text{Kor 5.4.6(ii)}}{=} \exp(-jz) = \exp((-j)z) = \exp(nz).$$

Zusammenhang mit der e -Funktion In diesem Paragraphen studieren wir (wie angekündigt) den Zusammenhang zwischen $\exp_{\mathbb{R}}$ und der e -Funktion. Dazu zunächst die folgende Beobachtung

Proposition 5.4.10 — EULER'SCHE ZAHL UND EXPONENTIALFUNKTION

Es sei e wie in Definition 3.5.9. Dann gilt

$$\exp(1) = e.$$

Beweis. Beachte: Nach Definition 5.4.2 ist

$$\exp(1) = \sum_{k=0}^{\infty} \frac{1^k}{k!} = \sum_{k=0}^{\infty} \frac{1}{k!}. \quad (234)$$

Nach dem Binomischen Lehrsatz (vgl. Satz 3.3.27) gilt für alle $n \in \mathbb{N}$

$$\left(1 + \frac{1}{n}\right)^n = \left(\frac{1}{n} + 1\right)^n = \sum_{k=0}^n \binom{n}{k} \left(\frac{1}{n}\right)^k 1^{n-k} = \sum_{k=0}^n \binom{n}{k} \frac{1}{n^k} = \sum_{k=0}^n \frac{1}{k!} \frac{1}{n^k} \prod_{\ell=1}^k (n - \ell + 1).$$

Schreibt man $n^k = \prod_{\ell=1}^k n$, so erhält man

$$\left(1 + \frac{1}{n}\right)^n = \sum_{k=0}^n \frac{1}{k!} \frac{\prod_{\ell=1}^k (n - \ell + 1)}{\prod_{\ell=1}^k n} = \sum_{k=0}^n \frac{1}{k!} \prod_{\ell=1}^k \frac{n - \ell + 1}{n}.$$

Daher gilt

$$\left(1 + \frac{1}{n}\right)^n - \sum_{k=0}^n \frac{1}{k!} = \sum_{k=0}^n \frac{1}{k!} \prod_{\ell=1}^k \frac{n - \ell + 1}{n} - \sum_{k=0}^n \frac{1}{k!} = \sum_{k=0}^n \frac{1}{k!} \left(\prod_{\ell=1}^k \frac{n - \ell + 1}{n} - 1 \right).$$

Es sei nun $\varepsilon > 0$. Für jedes feste $k \in \mathbb{N}$ gilt

$$\lim_{n \rightarrow \infty} \prod_{\ell=1}^k \frac{n - \ell + 1}{n} \stackrel{\text{GWR}}{=} \prod_{\ell=1}^k 1 = 1^k = 1.$$

Daher gilt für jedes $k \in \mathbb{N}$

$$\exists \tilde{n}_k \in \mathbb{N} \quad \forall n \geq \tilde{n}_k : \quad \left| \prod_{\ell=1}^k \frac{n - \ell + 1}{n} - 1 \right| \leq \frac{\varepsilon}{8} \quad (a) \quad (235)$$

Wir wissen ferner aus Satz 5.4.1, dass $\sum_{k \geq 1} \frac{1}{k!}$ eine konvergente Reihe ist. Das Cauchy-Kriterium (Proposition 5.1.1) gibt uns dann

$$\exists n_0 \in \mathbb{N} \quad \forall n, m \in \mathbb{N} : n > m \geq n_0 : \quad \left| \sum_{k=m+1}^n \frac{1}{k!} \right| \leq \frac{\varepsilon}{4}. \quad (236)$$

Wähle nun $N := \max\{\tilde{n}_1, \dots, \tilde{n}_{n_0}, n_0\}$. Für $n \geq N$ gilt dann

$$\left| \left(1 + \frac{1}{n}\right)^n - \sum_{k=0}^n \frac{1}{k!} \right|$$

$$\begin{aligned}
&= \left| \sum_{k=0}^n \frac{1}{k!} \left(\prod_{\ell=1}^k \frac{n-\ell+1}{n} - 1 \right) \right| \leq \sum_{k=0}^n \frac{1}{k!} \left| \prod_{\ell=1}^k \frac{n-\ell+1}{n} - 1 \right| \\
&\leq \sum_{k=0}^{n_0} \frac{1}{k!} \left| \prod_{\ell=1}^k \frac{n-\ell+1}{n} - 1 \right| + \sum_{k=n_0+1}^n \frac{1}{k!} \left| \prod_{\ell=1}^k \frac{n-\ell+1}{n} - 1 \right| \quad (237)
\end{aligned}$$

Beachte: Da $n \geq \tilde{n}_1, \dots, n \geq \tilde{n}_{n_0}$ gilt für alle $k \in \{1, \dots, n_0\}$ (also in der ersten Summe), dass

$$\left| \prod_{\ell=1}^k \frac{n-\ell+1}{n} - 1 \right| \leq \frac{\varepsilon}{8}, \quad (238)$$

vgl. (235). Für $k \geq n_0$ und $k \leq n$ (also in der zweiten Summe) können wir eine solche Abschätzung zwar nicht machen, aber es gilt immerhin noch

$$\begin{aligned}
\left| \prod_{\ell=1}^k \frac{n-\ell+1}{n} - 1 \right| &\stackrel{\text{Satz 2.0.3(iii)}}{\leq} \left| \prod_{\ell=1}^k \frac{n-\ell+1}{n} \right| + 1 \\
&= \prod_{k \leq n} \frac{n-\ell+1}{n} + 1 \leq \prod_{\ell=1}^k \frac{n}{n} + 1 = \prod_{\ell=1}^k 1 + 1 = 2. \quad (239)
\end{aligned}$$

Verwendet man (238) und (239) in (237), so erhält man dass für alle $n \geq N$

$$\begin{aligned}
&\left| \left(1 + \frac{1}{n} \right)^n - \sum_{k=0}^n \frac{1}{k!} \right| \\
&\leq \sum_{k=0}^{n_0} \frac{1}{k!} \underbrace{\left| \prod_{\ell=1}^k \frac{n-\ell+1}{n} - 1 \right|}_{\leq \frac{\varepsilon}{8}} + \sum_{k=n_0+1}^n \frac{1}{k!} \underbrace{\left| \prod_{\ell=1}^k \frac{n-\ell+1}{n} - 1 \right|}_{\leq 2} \\
&\leq \frac{\varepsilon}{8} \sum_{k=0}^{n_0} \frac{1}{k!} + 2 \sum_{k=n_0+1}^n \frac{1}{k!} \quad (240)
\end{aligned}$$

Nun sagt (236) (mit $m = n_0$), dass

$$\sum_{k=n_0+1}^n \frac{1}{k!} \leq \frac{\varepsilon}{4}. \quad (241)$$

Desweiteren ist (weil $k! \geq 2^{k-1}$ für alle $k \in \mathbb{N}^b$)

$$\sum_{k=0}^{n_0} \frac{1}{k!} = 1 + \sum_{k=1}^{n_0} \frac{1}{k!} \leq 1 + \sum_{k=1}^{n_0} \frac{1}{2^{k-1}} = 1 + \sum_{k=0}^{n_0-1} \left(\frac{1}{2} \right)^k = 1 + \frac{1 - \left(\frac{1}{2} \right)^{n_0}}{1 - \frac{1}{2}} \leq 1 + \frac{1}{1 - \frac{1}{2}} = 3 < 4. \quad (242)$$

Mit (241) und (242) in (240) sieht man

$$\forall n \geq N : \quad \left| \left(1 + \frac{1}{n}\right)^n - \sum_{k=0}^n \frac{1}{k!} \right| \leq \frac{\varepsilon}{8} \sum_{k=0}^{n_0} \frac{1}{k!} + 2 \sum_{k=n_0+1}^n \frac{1}{k!} \leq \frac{\varepsilon}{8} \cdot 4 + 2 \frac{\varepsilon}{4} = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Wir haben gesehen

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geq N : \quad \left| \left(1 + \frac{1}{n}\right)^n - \sum_{k=0}^n \frac{1}{k!} \right| \leq \varepsilon.$$

Mit anderen Worten

$$0 = \lim_{n \rightarrow \infty} \left[\left(1 + \frac{1}{n}\right)^n - \sum_{k=0}^n \frac{1}{k!} \right]$$

Mit den Grenzwertregeln (vgl Grenzwertregel 2) folgt

$$0 = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n - \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{1}{k!} \stackrel{\text{Prop 3.5.8}}{=} e - \sum_{k=0}^{\infty} \frac{1}{k!} \stackrel{(234)}{=} e - \exp(1).$$

Die Behauptung folgt.

^aAm liebsten würde man jetzt vielleicht das Maximum $\max\{\tilde{n}_k : k \in \mathbb{N}\}$ betrachten. Dieses muss aber nicht zwingend existieren — denn die Menge kann unendlich sein! Daher kommen wir so nicht weiter...

^bwie sie leicht mit Induktion zeigen könnten

Bemerkung 5.4.11

Ähnlich wie im Beweis der vorigen Proposition kann man auch zeigen, dass

$$\forall z \in \mathbb{C} : \quad \exp(z) = \lim_{n \rightarrow \infty} \left(1 + \frac{z}{n}\right)^n.$$

Sie können gerne als Übung versuchen, den obigen Beweis dafür anzupassen.

Die Konsistenz mit der e -Funktion folgt nun direkt aus der Funktionalgleichung

Satz 5.4.12 — KONSISTENZ MIT DER e -FUNKTION

Es seien $n \in \mathbb{Z}, m \in \mathbb{N}$. Dann gilt

$$\exp\left(\frac{n}{m}\right) = (\sqrt[m]{e})^n \quad (a)$$

^aDer Ausdruck $(\sqrt[m]{e})^n$ wird daher auch oft mit $e^{\frac{n}{m}}$ notiert.

Beweis. Für $m \in \mathbb{N}$ gilt

$$\exp\left(\frac{1}{m}\right)^m \stackrel{\text{Prop 5.4.9}}{=} \exp\left(m \cdot \frac{1}{m}\right) = \exp(1) \stackrel{\text{Prop 5.4.10}}{=} e.$$

Da nach Korollar 5.4.7 $\exp(\frac{1}{m}) > 0$, können wir in der obigen Gleichung die m -te Wurzel ziehen und erhalten

$$\exp(\frac{1}{m}) = \sqrt[m]{e}. \quad (243)$$

Nun gilt für $n \in \mathbb{Z}$ und $m \in \mathbb{N}$

$$\exp(\frac{n}{m}) \stackrel{\text{Prop 5.4.9}}{=} \exp(\frac{1}{m})^n \stackrel{(243)}{=} (\sqrt[m]{e})^n.$$

Bemerkung 5.4.13

Man kann nun dank des vorigen Satzes (siehe Fußnote) auch schreiben, dass

$$\exp(q) = e^q \quad \forall q \in \mathbb{Q}.$$

Bevor man diese Notation einführt, sollte man aber eine Spitzfindigkeit diskutieren: Ist $q = \frac{n}{m} = \frac{n'}{m'}$ auf zwei verschiedene Arten als Bruch dargestellt, so muss man überprüfen, dass

$$(\sqrt[m]{e})^n = (\sqrt[m']{e})^{n'}. \quad (244)$$

Andernfalls wäre der Ausdruck e^q nicht *wohldefiniert*, da man je nach gewählter Darstellung von q ein anderes Resultat erhalten könnte.

Die Gültigkeit von (244) sieht man mit dem vorigen Satz aber sehr leicht:

$$(\sqrt[m']{e})^{n'} = \exp(\frac{n'}{m'}) = \exp(q) = \exp(\frac{n}{m}) = (\sqrt[m]{e})^n$$

Die obige Bemerkung motiviert die folgende *Kurznotation* für $\exp$.

Notation 5.4.1. Für $z \in \mathbb{C}$ schreiben wir anstatt $\exp(z)$ oft auch einfach e^z .

5.5 Trigonometrische Funktionen

Als nächstes definieren wir die trigonometrischen Funktionen Kosinus und Sinus. Aus dem hier gesammelten Eigenschaften wird noch nicht vollends klar, dass es sich bei Kosinus und Sinus um die Funktionen handelt, die Sie in der Schule anhand des Einheitskreises definiert haben. Dies werden wir erst sehen, wenn wir am Ende des Abschnittes die *Kreiszahl* π eingeführt haben.

Definition 5.5.1 — KOSINUS UND SINUS

Es sei $z \in \mathbb{C}$. Wir definieren

$$(i) \text{ Kosinus. } \cos(z) := \frac{e^{iz} + e^{-iz}}{2}.$$

$$(ii) \text{ Sinus. } \sin(z) := \frac{e^{iz} - e^{-iz}}{2i}.$$

Dadurch haben wir zwei Funktionen $\cos : \mathbb{C} \rightarrow \mathbb{C}$ und $\sin : \mathbb{C} \rightarrow \mathbb{C}$ erklärt.

Für $x \in \mathbb{R}$ liegt e^{ix} stets auf dem *Einheitskreis* $\{z \in \mathbb{C} : |z| = 1\}$, vgl. Proposition 5.4.8 (iii). Eine wichtige Grundlage für die (Ihnen eventuell schon bekannte) Anschauung von Kosinus und Sinus *am Einheitskreis* ist die folgende Beobachtung:

Proposition 5.5.2 — KOSINUS, SINUS UND DER EINHEITSKREIS

Es sei $x \in \mathbb{R}$. Dann gilt

- (i) $\cos(x) = \operatorname{Re}(e^{ix})$. Insbesondere gilt auch $\cos(x) \in \mathbb{R}$.
- (ii) $\sin(x) = \operatorname{Im}(e^{ix})$. Insbesondere gilt auch $\sin(x) \in \mathbb{R}$.
- (iii) $e^{ix} = \cos(x) + i \sin(x)$.

Beweis. Sei $x \in \mathbb{R}$. Zu (i). Es gilt

$$\begin{aligned} \cos(x) &\stackrel{\text{Def5.5.1}}{=} \frac{e^{ix} + e^{-ix}}{2} = \frac{e^{ix} + e^{0-ix}}{2} = \frac{e^{ix} + e^{\overline{0+ix}}}{2} \\ &= \frac{e^{ix} + e^{\overline{ix}}}{2} \stackrel{\text{Prop5.4.8(i)}}{=} \frac{e^{ix} + \overline{e^{ix}}}{2} \stackrel{\text{Prop4.2.2}}{=} \operatorname{Re}(e^{ix}). \end{aligned}$$

Zu (ii).

$$\begin{aligned} \sin(x) &\stackrel{\text{Def5.5.1}}{=} \frac{e^{ix} - e^{-ix}}{2i} = \frac{e^{ix} - e^{0-ix}}{2i} = \frac{e^{ix} - e^{\overline{0+ix}}}{2i} \\ &= \frac{e^{ix} - e^{\overline{ix}}}{2i} \stackrel{\text{Prop5.4.8(i)}}{=} \frac{e^{ix} - \overline{e^{ix}}}{2i} \stackrel{\text{Prop4.2.2}}{=} \operatorname{Im}(e^{ix}). \end{aligned}$$

Behauptung (iii) folgt sofort aus (i) und (ii). In der Tat

$$e^{ix} = \operatorname{Re}(e^{ix}) + i \operatorname{Im}(e^{ix}) = \cos(x) + i \sin(x).$$

Die Anschauung von Sinus und Kosinus am Einheitskreis können wir erst konkreter studieren, wenn wir die *Kreiszahl* π eingeführt haben.

Im Folgenden werden wir noch einige *Eigenschaften* von Kosinus und Sinus studieren.

Proposition 5.5.3 — EIGENSCHAFTEN VON KOSINUS UND SINUS

- (i) $\cos(0) = 1$, $\sin(0) = 0$.
- (ii) Für alle $z \in \mathbb{C}$ gilt $\cos(-z) = \cos(z)$ und $\sin(-z) = -\sin(z)$.

(iii) Für alle $z \in \mathbb{C}$ gilt $(\cos(z))^2 + (\sin(z))^2 = 1$.^a

^aIn Zukunft schreiben wir statt $(\cos(z))^2$ auch $\cos^2(z)$ (und analog für den Sinus), womit Sie diese Regel auch als $\cos^2(z) + \sin^2(z) = 1$ lesen können

Beweis. (i) ist klar mit Definition 5.5.1. Zu (ii). Es gilt für $z \in \mathbb{C}$

$$\cos(-z) \stackrel{\text{Def 5.5.1}}{=} \frac{e^{i(-z)} + e^{-i(-z)}}{2} = \frac{e^{-iz} + e^{iz}}{2} = \frac{e^{iz} + e^{-iz}}{2} \stackrel{\text{Def 5.5.1}}{=} \cos(z).$$

Die Gleichung für den Sinus ist ganz analog.

Zu (iii).

$$\begin{aligned} (\cos(z))^2 + (\sin(z))^2 &= \left(\frac{e^{iz} + e^{-iz}}{2} \right)^2 + \left(\frac{e^{iz} - e^{-iz}}{2i} \right)^2 \\ &= \frac{(e^{iz} + e^{-iz})^2}{4} + \frac{(e^{iz} - e^{-iz})^2}{-4} \\ &= \frac{(e^{iz} + e^{-iz})^2 - (e^{iz} - e^{-iz})^2}{4} \\ &= \frac{(e^{iz})^2 + 2e^{iz}e^{-iz} + (e^{-iz})^2 - [(e^{iz})^2 - 2e^{iz}e^{-iz} + (e^{-iz})^2]}{4} \\ &= \frac{4e^{iz}e^{-iz}}{4} = e^{iz}e^{-iz} \stackrel{\text{Kor 5.4.6(ii)}}{=} e^{iz} \frac{1}{e^{iz}} = 1 \end{aligned}$$

Weitere wichtige Werkzeuge für das Rechnen mit Kosinus und Sinus sind die *Additionstheoreme*.

Proposition 5.5.4 — ADDITIONSTHEOREME

Es seien $z, w \in \mathbb{C}$. Dann gilt

(i) $\cos(z + w) = \cos(z)\cos(w) - \sin(z)\sin(w).$

(ii) $\sin(z + w) = \sin(z)\cos(w) + \cos(z)\sin(w).$

Ferner gelten (als Spezialfälle der oberen Regeln) die folgenden *Verdopplungsformeln*.

(iii) $\cos(2z) = \cos^2(z) - \sin^2(z).$

(iv) $\sin(2z) = 2\sin(z)\cos(z).$

Beweis. Wir zeigen nur (i) und (iii). Den Rest überlassen wir den Lesenden als Übungs-

aufgabe. Zu (i).

$$\begin{aligned}
 \cos(z) \cos(w) - \sin(z) \sin(w) &= \frac{e^{iz} + e^{-iz}}{2} \frac{e^{iw} + e^{-iw}}{2} - \frac{e^{iz} - e^{-iz}}{2i} \frac{e^{iw} - e^{-iw}}{2i} \\
 &= \frac{(e^{iz} + e^{-iz})(e^{iw} + e^{-iw})}{2 \cdot 2} - \frac{(e^{iz} - e^{-iz})(e^{iw} - e^{-iw})}{(2i) \cdot (2i)} \\
 &= \frac{(e^{iz} + e^{-iz})(e^{iw} + e^{-iw})}{4} - \frac{(e^{iz} - e^{-iz})(e^{iw} - e^{-iw})}{-4} \\
 &= \frac{(e^{iz} + e^{-iz})(e^{iw} + e^{-iw})}{4} + \frac{(e^{iz} - e^{-iz})(e^{iw} - e^{-iw})}{4} \\
 &= \frac{e^{iz}e^{iw} + e^{iz}e^{-iw} + e^{-iz}e^{iw} + e^{-iz}e^{-iw}}{4} \\
 &\quad + \frac{e^{iz}e^{iw} - e^{iz}e^{-iw} - e^{-iz}e^{iw} + e^{-iz}e^{-iw}}{4} \\
 &= \frac{2e^{iz}e^{iw} + 2e^{-iz}e^{-iw}}{4} = \frac{e^{iz}e^{iw} + e^{-iz}e^{-iw}}{2} \\
 &\stackrel{\text{Satz 5.4.5}}{=} \frac{e^{iz+iw} + e^{-i(z+w)}}{2} = \frac{e^{i(z+w)} + e^{-i(z+w)}}{2} = \cos(z+w).
 \end{aligned}$$

Zu (iii) Betrachten Sie (i) mit $z = w$. Genauer

$$\cos(2z) = \cos(z+z) \stackrel{(i)}{=} \cos(z) \cos(z) - \sin(z) \sin(z) = \cos^2(z) - \sin^2(z).$$

Bemerkung 5.5.5

Die Formeln in Proposition 5.5.4 sind leicht zu zeigen, wenn man sie schon vor sich liegen hat. Wie aber ist man auf eine solche Formel gekommen? Dazu nimm $x, y \in \mathbb{R}$. Dann gilt wegen der Funktionalgleichung (Satz 5.4.5)

$$e^{i(x+y)} = e^{ix+iy} = e^{ix}e^{iy}.$$

Nutzen wir nun Proposition 5.5.2(iii) für alle Ausdrücke, so sehen wir

$$\cos(x+y) + i \sin(x+y) = (\cos(x) + i \sin(x))(\cos(y) + i \sin(y))$$

Multiplizieren wir nun das Produkt auf der rechten Seite aus (vgl. (189)) so erhalten wir

$$\cos(x+y) + i \sin(x+y) = (\cos(x) \cos(y) - \sin(x) \sin(y)) + i(\sin(x) \cos(y) + \cos(x) \sin(y))$$

Vergleichen wir nun Realteile und Imaginärteile, so erhalten wir unsere Additionstheoreme

$$\cos(x+y) = \cos(x) \cos(y) - \sin(x) \sin(y)$$

und

$$\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y).$$

Bemerkung 5.5.6

Laut Proposition 5.5.3 (ii) ist für alle $z \in \mathbb{C}$

$$1 = \cos^2(z) + \sin^2(z).$$

Ferner gilt nach Proposition 5.5.4 (iii)

$$\cos(2z) = \cos^2(z) - \sin^2(z).$$

Addieren wir die beiden obigen Gleichungen auf, so erhalten wir

$$1 + \cos(2z) = [\cos^2(z) + \sin^2(z)] + [\cos^2(z) - \sin^2(z)] = 2\cos^2(z).$$

Dies beschert uns die *Formel des Quadrates über den doppelten Winkel*

$$\cos^2(z) = \frac{1 + \cos(2z)}{2}.$$

Als nächstes leiten wir *Potenzreihendarstellungen* für Kosinus und Sinus her. Diese werden uns in Zukunft dabei helfen, Aussagen über $\cos/\sin$ zu treffen.

Proposition 5.5.7 — KOSINUSREIHE UND SINUSREIHE

Es gilt für alle $z \in \mathbb{C}$

(i)

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

(ii)

$$\sin(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} z^{2k+1}.$$

Beide Reihen konvergieren für alle $z \in \mathbb{C}$ absolut.

Beweis. Sei $z \in \mathbb{C}$. Zu (i).

$$\begin{aligned} \cos(z) &= \frac{e^{iz} + e^{-iz}}{2} = \frac{1}{2} (e^{iz} + e^{-iz}) \\ &\stackrel{\text{Def 5.4.2}}{=} \frac{1}{2} \left(\sum_{\ell=0}^{\infty} \frac{(iz)^\ell}{\ell!} + \sum_{\ell=0}^{\infty} \frac{(-iz)^\ell}{\ell!} \right) \end{aligned}$$

$$\begin{aligned}
& \stackrel{\text{Prop 5.0.8}}{=} \frac{1}{2} \sum_{\ell=0}^{\infty} \left(\frac{(iz)^{\ell}}{\ell!} + \frac{(-iz)^{\ell}}{\ell!} \right) \\
&= \frac{1}{2} \sum_{\ell=0}^{\infty} \left(\frac{i^{\ell} z^{\ell}}{\ell!} + \frac{(-1)^{\ell} i^{\ell} z^{\ell}}{\ell!} \right) \\
&= \frac{1}{2} \sum_{\ell=0}^{\infty} \left(\underbrace{[1 + (-1)^{\ell}]}_{\substack{=2 \text{ falls } \ell \text{ gerade,} \\ =0 \text{ falls } \ell \text{ ungerade}}} \frac{i^{\ell} z^{\ell}}{\ell!} \right) \stackrel{\ell=2k}{=} \frac{1}{2} \sum_{k=0}^{\infty} 2 \frac{i^{2k} z^{2k}}{(2k)!} \\
&= \sum_{k=0}^{\infty} \frac{(i^2)^k}{(2k)!} z^{2k} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} z^{2k}
\end{aligned}$$

Zu (ii).

$$\begin{aligned}
\sin(z) &= \frac{e^{iz} - e^{-iz}}{2i} = \frac{1}{2i} (e^{iz} - e^{-iz}) \\
& \stackrel{\text{Def 5.4.2}}{=} \frac{1}{2i} \left(\sum_{\ell=0}^{\infty} \frac{(iz)^{\ell}}{\ell!} - \sum_{\ell=0}^{\infty} \frac{(-iz)^{\ell}}{\ell!} \right) \\
& \stackrel{\text{Prop 5.0.8}}{=} \frac{1}{2i} \sum_{\ell=0}^{\infty} \left(\frac{(iz)^{\ell}}{\ell!} - \frac{(-iz)^{\ell}}{\ell!} \right) \\
&= \frac{1}{2i} \sum_{\ell=0}^{\infty} \left(\frac{i^{\ell} z^{\ell}}{\ell!} + \frac{(-1)^{\ell} i^{\ell} z^{\ell}}{\ell!} \right) \\
&= \frac{1}{2i} \sum_{\ell=0}^{\infty} \left(\underbrace{[1 - (-1)^{\ell}]}_{\substack{=0 \text{ falls } \ell \text{ gerade,} \\ =2 \text{ falls } \ell \text{ ungerade}}} \frac{i^{\ell} z^{\ell}}{\ell!} \right) \stackrel{\ell=2k+1}{=} \frac{1}{2i} \sum_{k=0}^{\infty} 2 \frac{i^{2k+1} z^{2k+1}}{(2k+1)!} \\
&= \sum_{k=0}^{\infty} \frac{i^{2k}}{(2k+1)!} z^{2k+1} = \sum_{k=0}^{\infty} \frac{(i^2)^k}{(2k+1)!} z^{2k+1} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} z^{2k+1}.
\end{aligned}$$

Die absolute Konvergenz der Reihen überlassen wir den Lesenden als Übungsaufgabe.

Bemerkung 5.5.8

Man sieht den Reihen in Proposition 5.5.7 auf den ersten Blick nicht an, dass es sich um Potenzreihen handelt. Eigentlich müsste man es wie folgt aufschreiben

$$\cos(z) = \sum_{\ell=0}^{\infty} a_{\ell} z^{\ell} \quad \text{wobei } a_{\ell} = \begin{cases} \frac{(-1)^{\frac{\ell}{2}}}{\ell!} & \ell \text{ gerade} \\ 0 & \ell \text{ ungerade} \end{cases}$$

Dabei handelt es sich dann um eine Potenzreihe gemäß Definition 5.2.1. Wenn Sie $\ell = 2k$,

$k \in \mathbb{N} \cup \{0\}$, oben einsetzen, können Sie sich überzeugen, dass am Ende dieselbe Formel wie in Proposition 5.5.7 herauskommt. Analog gilt

$$\sin(z) = \sum_{\ell=0}^{\infty} a_{\ell} z^{\ell} \quad \text{wobei } a_{\ell} = \begin{cases} \frac{(-1)^{\frac{\ell-1}{2}}}{\ell!} & \ell \text{ ungerade} \\ 0 & \ell \text{ gerade} \end{cases}.$$

Wiederum können Sie $\ell = 2k + 1$, $k \in \mathbb{N} \cup \{0\}$, setzen und sich davon überzeugen, dass sie bei derselben Formel wie in Proposition 5.5.7 landen. Damit handelt es sich wirklich um Potenzreihen.

Die Kreiszahl π . Mit dem Wissen, das wir über trigonometrische Funktionen haben, können wir (mit einer kleinen Lücke) die *Kreiszahl* π studieren. Hierzu möchten wir in diesem Abschnitt folgenden Satz beweisen

Satz 5.5.9

Die Funktion $\cos : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \cos(x)$ hat im Intervall $[0, 2]_{\mathbb{R}}$ genau eine Nullstelle.

Sobald die Nullstelle im vorigen Satz gefunden ist, können wir π definieren, nämlich so:

Definition 5.5.10 — DIE KREISZAHL π

Es sei $\theta \in [0, 2]_{\mathbb{R}}$ die Nullstelle aus dem vorherigen Satz. Dann definieren wir $\pi := 2\theta$.

Bevor wir zum Beweis des obigen Satzes kommen benötigen wir einige Vorbereitungen. Nun zunächst ein Lemma zur *Fehlerabschätzung* bei den Partialsummen der Kosinus- und Sinusreihe.

Lemma 5.5.11

Es sei $x \in \mathbb{R}$ und $m \in \mathbb{N}_0$ mit $|x| \leq 2m + 3$. Dann gilt

- (i) $\left| \cos(x) - \sum_{k=0}^m \frac{(-1)^k}{(2k)!} x^{2k} \right| \leq \frac{|x|^{2m+2}}{(2m+2)!}$
- (ii) $\left| \sin(x) - \sum_{k=0}^m \frac{(-1)^k}{(2k+1)!} x^{2k+1} \right| \leq \frac{|x|^{2m+3}}{(2m+3)!}$

Beweisskizze. Wir zeigen nur (i), denn (ii) ist analog. Es gilt

$$\begin{aligned} \cos(x) - \sum_{k=0}^m \frac{(-1)^k}{(2k)!} x^{2k} &= \sum_{k=m+1}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} \\ &= \frac{(-1)^{m+1}}{(2m+2)!} x^{2m+2} + \frac{(-1)^{m+2}}{(2m+4)!} x^{2m+4} + \frac{(-1)^{m+3}}{(2m+6)!} x^{2m+6} + \dots \end{aligned}$$

$$= \frac{(-1)^{m+1}}{(2m+2)!} x^{2m+2} \left(1 - \frac{x^2}{(2m+3)(2m+4)} + \frac{x^4}{(2m+3)(2m+4)(2m+5)(2m+6)} - \dots \right) \quad (245)$$

Definiere nun für $\ell \in \mathbb{N}$ die Zahl $b_\ell := \frac{x^{2\ell}}{(2m+3)(2m+4) \cdot \dots \cdot (2m+2\ell+2)}$. Dann gilt nach (245)

$$\cos(x) - \sum_{k=0}^m \frac{(-1)^k}{(2k)!} x^{2k} = \frac{(-1)^{m+1}}{(2m+2)!} x^{2m+2} (1 - b_1 + b_2 - b_3 + \dots).$$

Wir werden nun zeigen, dass $(1 - b_1 + b_2 - b_3 + \dots) \in (0, 1]$. Hierzu benötigen wir zunächst eine Zwischenüberlegung zum Monotonieverhalten.

ZWISCHENBEHAUPTUNG. $(b_\ell)_{\ell \in \mathbb{N}}$ ist monoton fallend.

Dazu berechnen wir für $\ell \in \mathbb{N}$

$$\begin{aligned} b_{\ell+1} &= \frac{x^{2(\ell+1)}}{(2m+3)(2m+4) \cdot \dots \cdot (2m+2(\ell+1)+2)} \\ &= \frac{x^{2\ell+2}}{(2m+3)(2m+4) \cdot \dots \cdot (2m+2\ell+4)} \\ &= b_\ell \frac{x^2}{(2m+2\ell+3)(2m+2\ell+4)} \underset{|x| \leq 2m+3}{\leq} b_\ell \frac{(2m+3)^2}{(2m+2\ell+3)(2m+2\ell+4)} \\ &\leq b_\ell \frac{(2m+3)^2}{(2m+3)(2m+3)} = b_\ell. \end{aligned}$$

Aufgrund dieser Monotonie gilt zum Einen

$$\begin{aligned} 1 - b_1 + \underbrace{b_2 - b_3}_{\geq 0} + \underbrace{b_4 - b_5}_{\geq 0} + \dots &\geq 1 - b_1 \\ &\geq 1 - \frac{x^2}{(2m+3)(2m+4)} \underset{|x| \leq 2m+3}{\geq} 1 - \frac{(2m+3)^2}{(2m+3)(2m+4)} > 0. \end{aligned}$$

und zum Anderen

$$\underbrace{1 - b_1 + b_2}_{\leq 0} \underbrace{- b_3 + b_4}_{\leq 0} + \dots \leq 1.$$

Also gilt $1 - b_1 + b_2 - b_3 + \dots \in (0, 1]$. Man erhält

$$\left| \cos(x) - \sum_{k=0}^m \frac{(-1)^k}{(2k)!} x^{2k} \right| = \left| \frac{(-1)^{m+1}}{(2m+2)!} x^{2m+2} \right| \underbrace{\left| 1 - b_1 + b_2 - b_3 + \dots \right|}_{\in [0,1]} \leq \frac{1}{(2m+2)!} |x|^{2m+2}.$$

Ein erstes Korollar aus dieser Fehlerabschätzung ist, dass $\cos(2) < 0$. Es geht sogar noch genauer als das.

Korollar 5.5.12

Es gilt $\cos(2) \leq -\frac{1}{3}$.

Beweis. Wenden wir Lemma 5.5.11 (i) mit $m = 1$ an, so finden wir, dass für alle $x \in [-5, 5]$ gilt

$$\left| \cos(x) - \sum_{k=0}^1 \frac{(-1)^k}{(2k)!} x^{2k} \right| \leq \frac{|x|^4}{4!}$$

Mit anderen Worten

$$\left| \cos(x) - \left(1 - \frac{x^2}{2} \right) \right| \leq \frac{|x|^4}{4!} \quad \forall x \in [-5, 5] \quad (246)$$

Man erhält

$$\begin{aligned} \cos(2) &= \cos(2) - \left(1 - \frac{2^2}{2} \right) + \left(1 - \frac{2^2}{2} \right) \\ &\leq \left| \cos(2) - \left(1 - \frac{2^2}{2} \right) \right| + \left(1 - \frac{4}{2} \right) \\ &\stackrel{(246)}{\leq} \frac{2^4}{4!} + \left(1 - \frac{4}{2} \right) = \frac{16}{24} - 1 = \frac{-8}{24} = -\frac{1}{3}. \end{aligned}$$

Ein zweites Korollar aus der Fehlerabschätzung in Lemma 5.5.11 ist die Positivität des Sinus auf $(0, 2]$.

Korollar 5.5.13

Es gilt $\sin(x) > 0$ für alle $x \in (0, 2]$.

Beweis. Wenden wir Lemma 5.5.11 mit $m = 0$ an, so finden wir, dass für alle $x \in [-3, 3]$ gilt

$$\left| \sin(x) - \sum_{k=0}^0 \frac{(-1)^k}{(2k+1)!} x^{2k+1} \right| \leq \frac{|x|^3}{3!}.$$

Mit anderen Worten

$$|\sin(x) - x| \leq \frac{|x|^3}{3!} \quad \forall x \in [-3, 3].$$

Daher gilt für $x \in (0, 2]$

$$\begin{aligned} \sin(x) &= \underbrace{\sin(x) - x}_{\geq -|\sin(x) - x|} + x \geq x - |\sin(x) - x| \geq x - \frac{|x|^3}{3!} = x - \frac{x^3}{3!} \\ &= x \left(1 - \frac{x^2}{6}\right) \geq x \underbrace{\left(1 - \frac{2^2}{6}\right)}_{=\frac{1}{3}} = \frac{1}{3}x > 0. \end{aligned}$$

Lemma 5.5.14

Es seien $x, y \in [0, 2]$ mit $x < y$. Dann gilt $\cos(x) > \cos(y)$.

Beweis.

$$\begin{aligned} \cos(y) - \cos(x) &= \cos\left(\underbrace{\frac{y+x}{2} + \frac{y-x}{2}}_{=y}\right) - \cos\left(\underbrace{\frac{y+x}{2} - \frac{y-x}{2}}_{=x}\right) \\ &\stackrel{\text{Prop 5.5.4}}{=} \cos\left(\frac{y+x}{2}\right) \cos\left(\frac{y-x}{2}\right) - \sin\left(\frac{y+x}{2}\right) \sin\left(\frac{y-x}{2}\right) - \left[\cos\left(\frac{y+x}{2}\right) \cos\left(-\frac{y-x}{2}\right) - \sin\left(\frac{y+x}{2}\right) \sin\left(-\frac{y-x}{2}\right)\right] \\ &= \cos\left(\frac{y+x}{2}\right) \cos\left(\frac{y-x}{2}\right) - \sin\left(\frac{y+x}{2}\right) \sin\left(\frac{y-x}{2}\right) - \left[\cos\left(\frac{y+x}{2}\right) \underbrace{\cos\left(-\frac{y-x}{2}\right)}_{=\cos\left(\frac{y-x}{2}\right)} - \sin\left(\frac{y+x}{2}\right) \underbrace{\sin\left(-\frac{y-x}{2}\right)}_{=-\sin\left(\frac{y-x}{2}\right)}\right] \\ &\stackrel{\text{Prop 5.5.3}}{=} \cos\left(\frac{y+x}{2}\right) \cos\left(\frac{y-x}{2}\right) - \sin\left(\frac{y+x}{2}\right) \sin\left(\frac{y-x}{2}\right) - \left[\cos\left(\frac{y+x}{2}\right) \cos\left(\frac{y-x}{2}\right) + \sin\left(\frac{y+x}{2}\right) \sin\left(\frac{y-x}{2}\right)\right] \\ &= -2 \sin\left(\frac{y+x}{2}\right) \sin\left(\frac{y-x}{2}\right) \end{aligned}$$

Da $0 < x < y < 2$ gilt $\frac{x+y}{2} \in (0, 2]$ und $\frac{y-x}{2} \in (0, 2]^a$. Nach Korollar 5.5.13 ist dann $\sin\left(\frac{y+x}{2}\right), \sin\left(\frac{y-x}{2}\right) > 0$. Wir folgern aus der obigen Gleichung

$$\cos(y) - \cos(x) < 0.$$

^aDie Details dazu sind den Lesenden als Übungsaufgabe überlassen.

Nun können wir erklären wie Satz 5.5.9 bewiesen wird.

Beweis (zu Satz 5.5.9).

EXISTENZ DER NULLSTELLE. Wir wissen aus Proposition 5.5.3, dass $\cos(0) = 1$. Ferner wissen wir aus Korollar 5.5.1 dass $\cos(2) \leq -\frac{1}{3}$. Jetzt werden wir den sogenannten *Zwischenwertsatz* einsetzen, welchen wir später erst beweisen können. Dieser geht so:

Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine *stetige*^a Funktion. Dann wird jeder Wert zwischen $f(a)$ und $f(b)$ im Intervall $[a, b]$ mindestens einmal angenommen.

Hier wenden wir dies an für $f = \cos$ auf $[a, b] = [0, 2]$. Dann wird jeder Wert zwischen $\cos(0) = 1$ und $\cos(2) \leq -\frac{1}{3}$ angenommen. Insbesondere muss der Wert 0 angenommen werden. Existenz der Nullstelle ist gezeigt.

EINDEUTIGKEIT DER NULLSTELLE. Es seien $\theta_1, \theta_2 \in [0, 2]$ mit $\cos(\theta_1) = \cos(\theta_2) = 0$. Wäre $\theta_1 \neq \theta_2$ (etwa z.B. $\theta_1 < \theta_2$) so wäre

$$0 = \cos(\theta_1) \underset{\text{Lemma 5.5.14}}{>} \cos(\theta_2) = 0.$$

Ein Widerspruch. Analog erhalten wir auch einen Widerspruch wenn $\theta_2 < \theta_1$ (mit vertauschten Rollen von θ_1, θ_2). Somit muss die Annahme $\theta_1 \neq \theta_2$ falsch sein.

^awas „stetig“ bedeutet, müssen wir noch erklären. Und warum $f = \cos$ auf $[0, 2]$ stetig ist müssen wir auch noch erklären.

Die Lücke, die im obigen Beweis ist, möchten wir noch schließen. Das motiviert das nächste Kapitel, in dem wir uns über *stetige Funktionen* unterhalten wollen.

Bevor wir dazu kommen fassen wir aber noch ein paar Konsequenzen aus Satz 5.5.9 zusammen.

Proposition 5.5.15 — VERSCHIEDENE WERTE VON KOSINUS/SINUS

Es sei π wie in Definition 5.5.10. Dann gilt

- (i) $\cos \frac{\pi}{2} = 0$ und $\sin \frac{\pi}{2} = 1$
- (ii) $e^{i\frac{\pi}{2}} = i$
- (iii) $e^{i\pi} = -1$
- (iv) $\cos \pi = -1$ und $\sin \pi = 0$.
- (v) $e^{\frac{3}{2}\pi i} = -i$
- (vi) $\cos \frac{3}{2}\pi = 0$ und $\sin \frac{3}{2}\pi = -1$
- (vii) $e^{2\pi i} = 1$
- (viii) $\cos 2\pi = 1$ und $\sin 2\pi = 0$

Beweis. Zu (i). Dass $\cos \frac{\pi}{2} = 0$ folgt direkt aus Definition 5.5.10. Nun beachte, dass aus Proposition 5.5.3 (iii) folgt

$$1 = \underbrace{\cos^2 \frac{\pi}{2}}_{=0} + \sin^2 \frac{\pi}{2} = \sin^2 \frac{\pi}{2}.$$

Somit gilt

$$0 = 1 - \left(\sin \frac{\pi}{2}\right)^2 \stackrel{\text{3.Bin.Formel}}{=} (1 - \sin \frac{\pi}{2})(1 + \sin \frac{\pi}{2}).$$

Das bedeutet, dass $\sin \frac{\pi}{2} = 1$ oder $\sin \frac{\pi}{2} = -1$. Da aber nach Definition 5.5.10 $\frac{\pi}{2} \in (0, 2]$, gilt nach 5.5.13, dass $\sin \frac{\pi}{2} > 0$. Wir schließen, dass $\sin \frac{\pi}{2} = 1$

Zu (ii). Es gilt nach Proposition 5.5.2 (iii)

$$e^{i\frac{\pi}{2}} = \cos \frac{\pi}{2} + i \sin \frac{\pi}{2} = 0 + i \cdot 1 = i$$

Zu (iii). Es gilt

$$e^{i\pi} = e^{i(\frac{\pi}{2} + \frac{\pi}{2})} = e^{i\frac{\pi}{2} + i\frac{\pi}{2}} \stackrel{\text{Satz 5.4.5}}{=} e^{i\frac{\pi}{2}} e^{i\frac{\pi}{2}} = (e^{i\frac{\pi}{2}})^2 = i^2 = -1.$$

Zu (iv). Nach Proposition 5.5.2 gilt

$$\cos \pi = \operatorname{Re}(e^{i\pi}) \stackrel{\text{(iii)}}{=} \operatorname{Re}(-1) = -1 \quad \text{und} \quad \sin \pi = \operatorname{Im}(e^{i\pi}) \stackrel{\text{(iii)}}{=} \operatorname{Im}(-1) = 0.$$

Die restlichen Behauptungen (v)-(viii) überlassen wir den Lesenden als Übungsaufgabe.

Zusammen mit den oben kennengelernten Werten und den Additionstheoremen ergeben sich *Symmetrien und Periodizitäten* für Kosinus und Sinus.

Korollar 5.5.16 — SYMMETRIEN UND PERIODIZITÄT

Es gilt für $z \in \mathbb{C}$

- (i) $\cos(\frac{\pi}{2} + z) = -\sin(z)$, $\sin(\frac{\pi}{2} + z) = \cos(z)$.
- (ii) $\cos(\pi + z) = -\cos(z)$, $\sin(\pi + z) = -\sin(z)$.
- (iii) $\cos(2\pi + z) = \cos(z)$, $\sin(2\pi + z) = \sin(z)$.

Beweis. Die Behauptungen folgen stets aus den Additionstheoremen (Proposition 5.5.4) und aus Proposition 5.5.15. Zu (i). Es gilt für $z \in \mathbb{C}$

$$\cos(\frac{\pi}{2} + z) \stackrel{\text{Prop 5.5.4(i)}}{=} \cos(\frac{\pi}{2}) \cos(z) - \sin(\frac{\pi}{2}) \sin(z) \stackrel{\text{Prop 5.5.15(i)}}{=} 0 \cdot \cos(z) - 1 \cdot \sin(z) = -\sin(z).$$

Die restlichen Behauptungen sind den Lesenden als Übungsaufgabe überlassen.

Mit den gegebenen Informationen könnte man nun auch weitere Werte von $\cos, \sin$ berechnen. Ein Beispiel wollen wir hier aufführen

Proposition 5.5.17

Es gilt

$$\cos\left(\frac{\pi}{4}\right) = \sin\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}.$$

Beweis. Es gilt

$$\cos\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{2} - \frac{\pi}{4}\right) \stackrel{\text{Kor5.5.16(i)}}{=} -\sin\left(-\frac{\pi}{4}\right) \stackrel{\text{Prop5.5.3}}{=} \sin\left(\frac{\pi}{4}\right) \quad (247)$$

Die erste Gleichheit in der Behauptung ist also schonmal gezeigt. Außerdem gilt auch

$$2 \sin^2\left(\frac{\pi}{4}\right) = \sin^2\left(\frac{\pi}{4}\right) + \sin^2\left(\frac{\pi}{4}\right) \stackrel{(247)}{=} \sin^2\left(\frac{\pi}{4}\right) + \cos^2\left(\frac{\pi}{4}\right) \stackrel{\text{Prop5.5.3}}{=} 1.$$

Man folgert also $\sin^2\left(\frac{\pi}{4}\right) = \frac{1}{2}$. Da nun $\frac{\pi}{4} \leq \frac{\pi}{2} \leq 2$ (vgl. Definition 5.5.10), gilt nach Korollar 5.5.13 auch $\sin\left(\frac{\pi}{4}\right) > 0$. Daher können wir (mit positivem Vorzeichen) Wurzelziehen und erhalten

$$\sin\left(\frac{\pi}{4}\right) = \frac{1}{\sqrt{2}}.$$

Zusammen mit (247) folgt die Behauptung.

08.01.2025

Bonuskapitel. KOSINUS UND SINUS AM EINHEITSKREIS, POLARKOORDINATEN.Für sämtliche $x \in \mathbb{R}$ gilt

$$e^{ix} = \cos(x) + i \sin(x).$$

Die komplexe Zahl e^{ix} wird also auf der komplexen Zahlenebene als Vektor $(\cos(x), \sin(x))$ eingezeichnet. Ferner liegt die Zahl e^{ix} stets auf dem Einheitskreis. Desweiteren kennen wir nun einige Werte für $x = 0, \frac{\pi}{2}, \pi, \frac{3}{2}\pi, 2\pi$ (und wegen der Periodizität auch noch viele mehr). Diese Werte legen folgende Beobachtung nahe: Die Abbildung $x \mapsto e^{ix}$ durchläuft den Einheitskreis innerhalb des Intervalls $[0, 2\pi]$ *einmal gegen den Uhrzeigersinn*. Es bleiben aber noch Fragen offen

- Wird der Einheitskreis komplett als Ganzes durchlaufen?
- Wird die Durchlaufrichtung stets beibehalten oder macht die Abbildung eine Kehrtwende?
- Wird der Einheitskreis ohne Sprünge durchlaufen?
- Wird der Einheitskreis mit *gleichbleibender Geschwindigkeit* durchlaufen?

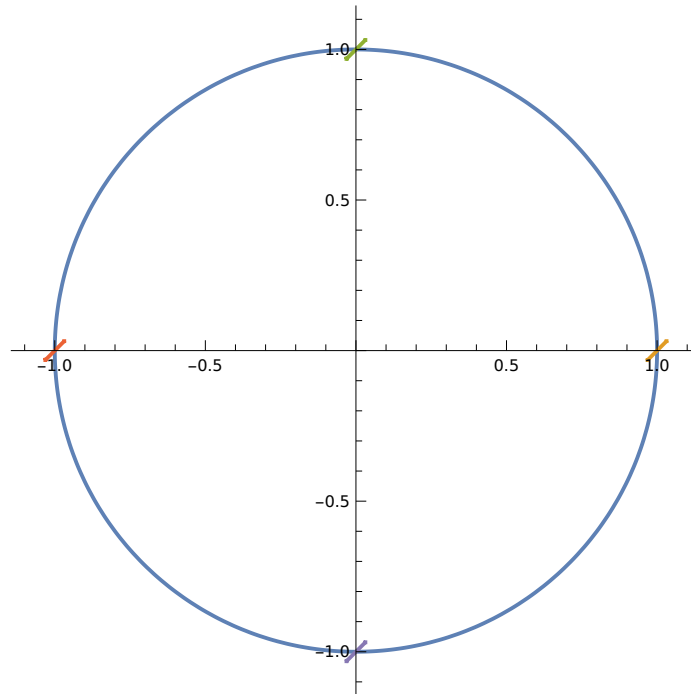


Abbildung 20: Der Einheitskreis mit markierten Werten $e^{i0} = 1$ (gelb), $e^{i\frac{\pi}{2}} = i$ (grün), $e^{i\pi} = -1$ (orange), $e^{i\frac{3}{2}\pi} = -i$ (lila)

Auf all diese Fragen können wir noch in diesem Semester Antworten geben. Frage 1 und 2 können wir mit der folgenden Proposition klären. Die dritte Frage können wir leider nur mit dem Begriff der *Stetigkeit* beantworten und für die vierte Frage brauchen wir *Differentialrechnung* ... Wir müssen das an dieser Stelle offen lassen, werden aber an geeigneter Stelle nochmal darauf zurückkommen.

Proposition 5.5.18

Es sei $\mathbb{S}^1 := \{z \in \mathbb{C} : |z| = 1\}$. Dann ist

$$e : [0, 2\pi) \rightarrow \mathbb{S}^1, \quad e(x) := e^{ix}$$

eine bijektive Abbildung. ^a

^aWie diese aussieht kann man auch interaktiv auf GEOGEBRA verstehen.

Dass der Einheitskreis komplett durchlaufen wird durch die *Surjektivität* garantiert. Dass es keine Kehrtwenden geben kann folgt aus der *Injektivität*. Somit ist obiges exakt das was wir brauchen um die ersten beiden Fragen zu beantworten.

Beweis. Definiere zunächst die *vier Quadranten*

$$Q_1 := \{z \in \mathbb{C} : \operatorname{Re}(z) \geq 0, \operatorname{Im}(z) \geq 0\}$$

$$Q_2 := \{z \in \mathbb{C} : \operatorname{Re}(z) \leq 0, \operatorname{Im}(z) \geq 0\}$$

$$Q_3 := \{z \in \mathbb{R} : \operatorname{Re}(z) \leq 0, \operatorname{Im}(z) \leq 0\}$$

$$Q_4 := \{z \in \mathbb{R} : \operatorname{Re}(z) \geq 0, \operatorname{Im}(z) \leq 0\}.$$

SCHRITT 1. Wir zeigen, dass $e : [0, \frac{\pi}{2}] \rightarrow \mathbb{S}^1 \cap Q_1$ bijektiv ist.

Dazu:

1. Injektivität.

Angenommen es gibt zwei verschiedene $x_1, x_2 \in [0, \frac{\pi}{2}]$ (OBdA $x_1 < x_2$) mit $e(x_1) = e(x_2)$, d.h. $e^{ix_1} = e^{ix_2}$. Nimmt man den Realteil, so folgt $\cos(x_1) = \cos(x_2)$. Da nach Definition 5.5.10 aber $\frac{\pi}{2} \leq 2$ gilt, ist $\cos$ auf $[0, \frac{\pi}{2}]$ streng monoton fallend (vgl. Lemma 5.5.14). Dies würde $\cos(x_1) > \cos(x_2)$ implizieren und liefert einen Widerspruch zu $\cos(x_1) = \cos(x_2)$.

2. Surjektivität.

Es sei $z \in Q_1 \cap \mathbb{S}^1$, etwa $z = a + ib$ für $a, b \in \mathbb{R}$. Insbesondere gilt $a \geq 0, b \geq 0$ (da $z \in Q_1$) und $a^2 + b^2 = |z|^2 = 1$ (da $z \in \mathbb{S}^1$). Man erhält daher auch $a \in [0, 1]$ und $b \in [0, 1]$. Wir behaupten nun, dass es ein $x \in [0, \frac{\pi}{2}]$ gibt mit $e^{ix} = z$. Dazu: Wir wissen, dass $\cos(0) = 1$ und $\cos(\frac{\pi}{2}) = 0$. Der (noch zu beweisende) *Zwischenwertsatz* (siehe auch den Beweis von Satz 5.5.9) liefert dann, dass die Kosinusfunktion im Intervall $[0, \frac{\pi}{2}]$ jeden Wert zwischen 0 und 1 mindestens einmal annimmt. Da nun $a \in [0, 1]$ erhalten wir

$$\exists t \in [0, \frac{\pi}{2}] : \quad \cos(t) = a.$$

Dann gilt auch nach Proposition 5.5.3 (iii)

$$b^2 = 1 - a^2 = 1 - (\cos(t))^2 = \sin(t)^2.$$

Das bedeutet $\sin(t) = \pm b$. Da jedoch nach Korollar 5.5.13 $\sin(t) \geq 0$, haben wir $\sin(t) = b$. Somit gilt

$$e^{it} = \cos(t) + i \sin(t) = a + ib = z.$$

Die Behauptung folgt.

SCHRITT 2. Wir zeigen, dass

$$e : [\frac{\pi}{2}, \pi] \rightarrow \mathbb{S}^1 \cap Q_2,$$

$$e : [\pi, \frac{3}{2}\pi] \rightarrow \mathbb{S}^1 \cap Q_3 \text{ und}$$

$$e : [\frac{3}{2}\pi, 2\pi] \rightarrow \mathbb{S}^1 \cap Q_4$$

auch bijektiv sind.

Dies folgt direkt aus der Bijektivität von $e : [0, \frac{\pi}{2}] \rightarrow \mathbb{S}^1 \cap Q_1$ den Symmetrien von Kosinus und Sinus, die wir in Korollar 5.5.16 besprochen haben. Details sind den Lesenden als Übungsaufgabe überlassen.

SCHRITT 3. Wir folgern, dass $e : [0, 2\pi) \rightarrow \mathbb{S}^1$ bijektiv ist.

Zur Surjektivität: Da $e(0) = e(2\pi)$ gilt $e([0, 2\pi)) = e([0, 2\pi])$. Ferner gilt

$$\begin{aligned} e([0, 2\pi]) &= e([0, \frac{\pi}{2}] \cup [\frac{\pi}{2}, \pi] \cup [\pi, \frac{3}{2}\pi] \cup [\frac{3}{2}\pi, 2\pi]) \\ &= e([0, \frac{\pi}{2}]) \cup e([\frac{\pi}{2}, \pi]) \cup e([\pi, \frac{3}{2}\pi]) \cup e([\frac{3}{2}\pi, 2\pi]) \\ &= (\mathbb{S}^1 \cap Q_1) \cup (\mathbb{S}^1 \cap Q_2) \cup (\mathbb{S}^1 \cap Q_3) \cup (\mathbb{S}^1 \cap Q_4) \\ &= \mathbb{S}^1 \cap (Q_1 \cup Q_2 \cup Q_3 \cup Q_4) = \mathbb{S}^1. \end{aligned}$$

Für die Injektivität muss man sich überlegt man sich aus Schritt 2, dass

$$\begin{aligned} e : [0, \frac{\pi}{2}) &\rightarrow (\mathbb{S}^1 \cap Q_1) \setminus \{i\}, \\ e : [\frac{\pi}{2}, \pi) &\rightarrow (\mathbb{S}^1 \cap Q_2) \setminus \{-1\} \\ e : [\pi, \frac{3}{2}\pi) &\rightarrow (\mathbb{S}^1 \cap Q_3) \setminus \{-i\} \\ e : [\frac{3}{2}\pi, 2\pi) &\rightarrow (\mathbb{S}^1 \cap Q_4) \setminus \{1\} \end{aligned}$$

alle bijektiv sind. Insbesondere wird jedes Element in den jeweiligen Wertebereichen nur einmal getroffen. Da $(\mathbb{S}^1 \cap Q_1) \setminus \{i\}$, $(\mathbb{S}^1 \cap Q_2) \setminus \{-1\}$, $(\mathbb{S}^1 \cap Q_3) \setminus \{-i\}$, $(\mathbb{S}^1 \cap Q_4) \setminus \{1\}$ paarweise leeren Schnitt haben kann jedes Element aus $\mathbb{S}^1$ insgesamt nur einmal getroffen werden.

Als Konsequenz erhalten wir die sogenannte *Polarkoordinatendarstellung* für Zahlen aus $\mathbb{C} \setminus \{0\}$ und $\mathbb{R}^2 \setminus \{(0, 0)\}$, welche besagt, dass wir jede solche Zahl mithilfe von einem *Radius* $r > 0$ und einem *Winkel* $\theta \in [0, 2\pi)$ ausdrücken können.

Satz 5.5.19 — POLARKOORDINATEN

- (i) Es sei $z \in \mathbb{C} \setminus \{0\}$. Dann gibt es ein eindeutiges Paar (r, θ) mit $r \in (0, \infty)$ und $\theta \in [0, 2\pi)$ sodass

$$z = re^{i\theta}.$$

- (ii) Es sei $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$. Dann gibt es ein eindeutiges Paar (r, θ) mit $r \in (0, \infty)$ und $\theta \in [0, 2\pi)$ sodass

$$x = r \cos \theta \quad \text{und} \quad y = r \sin \theta.$$

Beweis. Zu (i).

EXISTENZ.

Es sei $z \in \mathbb{C} \setminus \{0\}$. Definiere $r := |z|$. Dann ist $r \in (0, \infty)$. Setze nun $z' := \frac{z}{r}$. Dann gilt

$$|z'| = \left| \frac{z}{r} \right| = \frac{|z|}{r} = \frac{r}{r} = 1.$$

Somit gilt also $z' \in \mathbb{S}^1$. Nach Proposition 5.5.18 gibt es also ein $\theta \in [0, 2\pi)$ mit $e(\theta) = z'$

(wobei e gewählt ist wie in Proposition 5.5.18, d.h. $e : [0, 2\pi) \rightarrow \mathbb{S}^1, x \mapsto e^{ix}$). Man erhält

$$e^{i\theta} = e(\theta) = z' = \frac{z}{r}$$

Formt man um, so ergibt sich $z = re^{i\theta}$.

EINDEUTIGKEIT.

Es seien $(r_1, \theta_1), (r_2, \theta_2)$ so, dass $r_1, r_2 \in (0, \infty), \theta_1, \theta_2 \in [0, 2\pi)$ und

$$z = r_1 e^{i\theta_1} = r_2 e^{i\theta_2}.$$

Also gilt $r_1 e^{i\theta_1} = r_2 e^{i\theta_2}$. Da $|e^{ix}| = 1$ für alle $x \in \mathbb{R}$ gilt

$$r_1 = r_1 \cdot 1 = r_1 \cdot |e^{i\theta_1}| = |r_1 e^{i\theta_1}| = |r_2 e^{i\theta_2}| = r_2 |e^{i\theta_2}| = r_2 \quad (248)$$

Definiere nun $r_0 := r_1 [= r_2]$. Folglich erhalten wir

$$r_0 e^{i\theta_1} = r_0 e^{i\theta_2}$$

und somit (da $r_0 > 0$) auch $e^{i\theta_1} = e^{i\theta_2}$. Mit anderen Worten gilt $e(\theta_1) = e(\theta_2)$. Da e auf $[0, 2\pi)$ injektiv ist folgt auch $\theta_1 = \theta_2$. Dies zusammen mit (248) impliziert, dass $(r_1, \theta_1) = (r_2, \theta_2)$.

Zu (ii).

EXISTENZ.

Zu gegebenem $(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\}$ assoziiere $x + iy \in \mathbb{C} \setminus \{0\}$. Stelle nun mithilfe von (i) $x + iy$ als $re^{i\theta}$ für ein Paar (r, θ) mit $r \in (0, \infty)$ und $\theta \in [0, 2\pi)$ dar. Dann gilt

$$x + iy = re^{i\theta} = r(\cos \theta + i \sin \theta) = r \cos \theta + ir \sin \theta$$

Vergleicht man nun Realteile und Imaginärteile so folgt

$$x = r \cos \theta \quad \text{und} \quad y = r \sin \theta.$$

EINDEUTIGKEIT.

Funktioniert auch durch die Identifikation von $\mathbb{R}^2$ mit $\mathbb{C}$. Details überlassen wir den Lesenden als Übungsaufgabe.

Wir haben also nun neben der Normalform $z = a + ib$ (wobei $a, b \in \mathbb{R}$) noch eine weitere Möglichkeit, komplexe Zahlen $z \in \mathbb{C} \setminus \{0\}$ eindeutig darzustellen: Nämlich durch die Polarkoordinatenform $z = re^{i\theta}$ (wobei $r \in (0, \infty)$ und $\theta \in [0, 2\pi)$). Genau wie wir a Realteil und b Imaginärteil nennen, können wir auch r und θ Namen geben. Bevor wir das tun: Beachte, dass wir für r schon einen Namen haben, denn r ist immer der *komplexe Absolutbetrag*! In

der Tat, ist $z = re^{i\theta}$ so gilt

$$|z| = |re^{i\theta}| = r |e^{i\theta}| = r. \quad (249)$$

Den Winkel θ nennen wir meistens *Polarwinkel* oder *Argument*.

Definition 5.5.20

Es sei $z \in \mathbb{C} \setminus \{0\}$. Ferner seien $r_z \in (0, \infty)$, $\theta_z \in [0, 2\pi)$ wie in Satz 5.5.19, d.h. $z = r_z e^{i\theta_z}$. Dann nennen wir r_z den *Absolutbetrag* von z und $\theta_z =: \arg(z)$ den *Polarwinkel* oder das *Argument* von z .

Nun mal ein konkretes Beispiel

Beispiel 5.5.21

Wir stellen $z = 1 + i$ in Polarkoordinaten dar. Dazu beobachten wir zunächst, dass nach (249)

$$r_z = |z| = |1 + i| = \sqrt{1^2 + 1^2} = \sqrt{2}.$$

Nun noch zum Polarwinkel. Wir wollen $\theta_z \in [0, 2\pi)$ mit $z = r_z e^{i\theta_z}$. Insbesondere gilt

$$e^{i\theta_z} = \frac{z}{r_z} = \frac{1+i}{\sqrt{2}} = \frac{1}{\sqrt{2}} + i \frac{1}{\sqrt{2}}$$

Da $e^{i\theta_z} = \cos(\theta_z) + i \sin(\theta_z)$ muss gelten, dass

$$\cos(\theta_z) = \frac{1}{\sqrt{2}}, \quad \sin(\theta_z) = \frac{1}{\sqrt{2}}.$$

Nach Proposition 5.5.17 ist dies gerade für $\theta_z = \frac{\pi}{4}$ der Fall. Wegen der Eindeutigkeit von θ_z ist dies auch die einzige Möglichkeit θ_z zu wählen und wir folgern

$$1 + i = \sqrt{2} e^{i\frac{\pi}{4}}.$$

Polarkoordinaten werden uns später im Studium noch häufiger begegnen. Aber auch schon jetzt eröffnen sie einige neue Perspektiven. Ein Beispiel dazu: Hat man zwei komplexe Zahlen $z_1 = a_1 + ib_1$, $z_2 = a_2 + ib_2$ in Normalform gegeben (d.h. $a_1, a_2, b_1, b_2 \in \mathbb{R}$ so kann man die Multiplikation in Normalform wie folgt ausrechnen

$$z_1 \cdot z_2 = (a_1 + ib_1)(a_2 + ib_2) = (a_1 a_2 - b_1 b_2) + i(a_1 b_2 + a_2 b_1). \quad (250)$$

Dies lässt sich zwar angenehm ausrechnen aber eine geometrische Vorstellung aus dieser Formel zu entwickeln ist sehr kompliziert. Viel einfacher zu interpretieren wird es, wenn wir die Zahlen in Polarkoordinaten vorliegen haben. Etwa sei $z_1 = r_1 e^{i\theta_1}$ und $z_2 = r_2 e^{i\theta_2}$ (wobei $r_1, r_2 \in (0, \infty)$ und $\theta_1, \theta_2 \in [0, 2\pi)$). Dann gilt

$$z_1 z_2 = r_1 e^{i\theta_1} r_2 e^{i\theta_2} = r_1 r_2 e^{i\theta_1} e^{i\theta_2} = (r_1 r_2) e^{i(\theta_1 + \theta_2)}. \quad (251)$$

Diese Formel lässt sich nun leicht in eine Anschauung überführen. Wer zwei komplexe Zahlen miteinander multipliziert muss...

- ... die komplexen Absolutbeträge miteinander multiplizieren,
- ... die Polarwinkel aufaddieren,

und erhält damit die Polarkoordinatendarstellung des Produktes. Diese Anschauung ist geometrisch leicht umzusetzen.

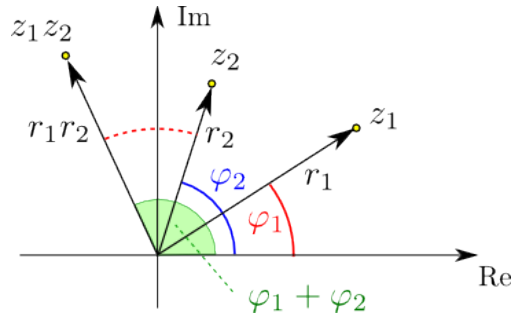


Abbildung 21: Veranschaulichung der komplexen Multiplikation von $z_1 = r_1 e^{i\varphi_1}$, $z_2 = r_2 e^{i\varphi_2}$

Einen kleinen Fallstrick gibt es hierbei: Betrachte wieder zwei komplexe Zahlen $z_1 = r_1 e^{i\theta_1}$ und $z_2 = r_2 e^{i\theta_2}$. Ist die Summe der Polarwinkel $\theta_1 + \theta_2$ größer als 2π , so ist der resultierende Winkel in (251) kein Polarwinkel mehr. Das lässt sich aber leicht reparieren: Wegen der 2π -Periodizität von $\theta \mapsto e^{i\theta}$ liefert der Winkel $\theta_1 + \theta_2 - 2\pi$ (der wieder ein zulässiger Polarwinkel ist) dasselbe Ergebnis.

Man kann sich nun fragen, was für einen Vorteil eine solche geometrische Anschauung der Multiplikation überhaupt hat. Die nächsten zwei Beispiele sollten (hoffentlich) von der Nützlichkeit überzeugen.

Beispiel 5.5.22

Wir berechnen $(1 + i)^{2026}$. Mit der üblichen Multiplikation in Normalform (250) könnte dies ein langwieriges Unterfangen werden. Nicht so, wenn man die Berechnung in Polarkoordinaten durchführt. Mit Beispiel 5.5.21 wissen wir, dass $1 + i = \sqrt{2} e^{i\frac{\pi}{4}}$ und somit

$$(1 + i)^{2026} = (\sqrt{2} e^{i\frac{\pi}{4}})^{2026} = 2^{1013} e^{2026i\frac{\pi}{4}}. \quad (252)$$

Schreiben wir nun (mittels Division durch 4 mit Rest) $2026 = 506 \cdot 4 + 2$ so sehen wir, dass $2026 \cdot \frac{\pi}{4} = (506 \cdot 4 + 2) \frac{\pi}{4} = 506\pi + \frac{\pi}{2}$. Daher gilt

$$e^{2026i\frac{\pi}{4}} = e^{506i\pi} e^{i\frac{\pi}{2}} = (e^{2\pi i})^{253} e^{i\frac{\pi}{2}} = 1^{253} i = i.$$

Das bedeutet in (252)

$$(1 + i)^{2026} = 2^{1013} i.$$

Beispiel 5.5.23

Wir stehen vor folgender Herausforderung

Aufgabe. Finde eine komplexe Zahl $z \in \mathbb{C}$ mit $z^{17} = 1 + i$.

Natürlich könnte man nun den Ansatz $z = a + ib$, mühsam z^{17} in Termen von a und b ausdrücken und bekäme dann ein nichtlineares Gleichungssystem in mehreren Variablen. Das wollen wir uns ersparen: Mit Polarkoordinaten ist es viel einfacher als das. Schreibe zunächst (mit Beispiel 5.5.21) $1 + i = \sqrt{2}e^{i\frac{\pi}{4}} = 2^{\frac{1}{2}}e^{i\frac{\pi}{4}}$. Ein $z \in \mathbb{C}$ mit $z^{17} = 2^{\frac{1}{2}}e^{i\frac{\pi}{4}}$ ist leicht gefunden: Definiere

$$z := 2^{\frac{1}{34}}e^{i\frac{\pi}{4 \cdot 17}}$$

Dann gilt

$$z^{17} = (2^{\frac{1}{34}}e^{i\frac{\pi}{4 \cdot 17}})^{17} = (2^{\frac{1}{34}})^{17}(e^{i\frac{\pi}{4 \cdot 17}})^{17} = 2^{\frac{1}{2}}e^{i\frac{\pi}{4}} = 1 + i.$$

6 Funktionsgrenzwerte, Stetigkeit

08.01.2026

Der nächste Abschnitt widmet sich dem Studium von *Funktionen* der Form $f : \mathbb{R} \rightarrow \mathbb{R}$ oder $f : \mathbb{C} \rightarrow \mathbb{C}$ oder $f : \mathbb{R} \rightarrow \mathbb{C}$ oder $f : \mathbb{C} \rightarrow \mathbb{R}$. Diese Objekte sind wesentlich komplizierter als Folgen, weil der Definitionsbereich viel *reichhaltiger* ist. Im Folgenden seien stets $\mathbb{K}, \mathbb{K}_1, \mathbb{K}_2 \in \{\mathbb{R}, \mathbb{C}\}$.

Definition 6.0.1 — $(\mathbb{K}_1, \mathbb{K}_2)$ -FUNKTION

Es sei $D \subseteq \mathbb{K}_1$. Dann nennen wir eine Abbildung $f : D \rightarrow \mathbb{K}_2$ eine $(\mathbb{K}_1, \mathbb{K}_2)$ -Funktion mit *Definitionsbereich* D . Ist $\mathbb{K}_1 = \mathbb{K}_2 = \mathbb{K}$ so spricht man auch von einer $\mathbb{K}$ -Funktion.

Im Fall von $\mathbb{K}_1 = \mathbb{K}_2 = \mathbb{R}$ spricht man manchmal auch von einer *reellen Funktion*.

Beispiel 6.0.2 — GÄNGIGE $\mathbb{K}$ -FUNKTIONEN

Im Folgenden geben wir eine Zusammenfassung bereits bekannter $\mathbb{K}$ -Funktionen.

- (i) *Konstante Funktionen*^a. Für $c \in \mathbb{K}_2$ betrachte $f : \mathbb{K}_1 \rightarrow \mathbb{K}_2, x \mapsto c$.
- (ii) *Identische Abbildungen*. Für $D \subseteq \mathbb{K}$ betrachte $id_D : D \rightarrow \mathbb{K}, x \mapsto x$.
- (iii) *Absolutbetrag*. Auch der Absolutbetrag $abs : \mathbb{K} \rightarrow \mathbb{R}, x \mapsto |x|$ definiert eine $(\mathbb{K}, \mathbb{R})$ -Funktion
- (iv) *Polynome*. Für $n \in \mathbb{N}$ und $a_0, \dots, a_n \in \mathbb{K}$ definiert $p : \mathbb{K} \rightarrow \mathbb{K}, x \mapsto \sum_{k=0}^n a_k x^k$ eine $\mathbb{K}$ -Funktion.
- (v) *Rationale Funktionen*. Es seien $p : \mathbb{K} \rightarrow \mathbb{K}$ und $q : \mathbb{K} \rightarrow \mathbb{K}$ zwei Polynome. Ferner sei $D := \{x \in \mathbb{K} : q(x) \neq 0\}$. Dann nennen wir $r : D \rightarrow \mathbb{K}, x \mapsto \frac{p(x)}{q(x)}$ eine rationale Funktion.
- (vi) *Exponentialfunktion*. Auch $\exp : \mathbb{K} \rightarrow \mathbb{K}, x \mapsto \exp(x)$ ist eine $\mathbb{K}$ -Funktion.
- (vii) *Trigonometrische Funktionen*. Auch $\cos : \mathbb{K} \rightarrow \mathbb{K}$ und $\sin : \mathbb{K} \rightarrow \mathbb{K}$ definieren $\mathbb{K}$ -Funktionen. Desweiteren kann man noch den *Tangens* definieren: Sei $D := \{x \in \mathbb{K} : \cos(x) \neq 0\}$. Dann ist $\tan : D \rightarrow \mathbb{K}$ definiert durch $\tan(x) := \frac{\sin(x)}{\cos(x)}$.

Im Fall $\mathbb{K} = \mathbb{R}$ kann man noch weitere $\mathbb{K}$ -Funktionen definieren.

- (viii) *Wurzelfunktionen*. Für $n \in \mathbb{N}$ definieren wir $\text{sqrt}_n : [0, \infty) \rightarrow \mathbb{R}, x \mapsto \sqrt[n]{x}$.
- (ix) *Gaußklammer*. Wir definieren $\text{ceil} : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto [x]$.

Auch im Fall $\mathbb{K} = \mathbb{C}$ hat man Möglichkeiten.

(x) *Konjugation*. Definiere $\text{conj} : \mathbb{C} \rightarrow \mathbb{C}, z \mapsto \bar{z}$.

^aIst f eine konstante mit dem Wert $c \in \mathbb{K}$ so schreiben wir auch $f \equiv c$.

Aus diesen Funktionen lassen sich noch weitere Funktionen durch sogenannte *Kompositionen* bauen. Streng genommen meint das Wort *Komposition* nichts anderes als eine Verkettung. Salopp wird es aber manchmal auch für andere Operationen (Addition, Multiplikation,...) verwendet, die man eigentlich besser *rationale Operationen* nennen sollte. Wer sich aber mit mehrdimensionaler Analysis auskennt weiß, dass auch diese Operationen als Verkettungen aufgefasst werden können. Mehr dazu im nächsten Semester!

Definition 6.0.3 — KOMPOSITIONEN VON FUNKTIONEN

Es sei $D \subseteq \mathbb{K}_1$ und seien $f, g : D \rightarrow \mathbb{K}_2$ zwei $(\mathbb{K}_1, \mathbb{K}_2)$ -Funktionen und $\lambda \in \mathbb{K}_2$.

- (i) Wir definieren $f + g : D \rightarrow \mathbb{K}_2, x \mapsto f(x) + g(x)$.
- (ii) Wir definieren $fg : D \rightarrow \mathbb{K}_2, x \mapsto f(x) \cdot g(x)$.
- (iii) Wir definieren $\lambda f : D \rightarrow \mathbb{K}_2, x \mapsto \lambda \cdot f(x)$.
- (iv) Für $\hat{D} := \{x \in D : g(x) \neq 0\}$ definieren wir $\frac{f}{g} : \hat{D} \rightarrow \mathbb{K}_2, x \mapsto \frac{f(x)}{g(x)}$.
- (v) Ist ferner $h : \tilde{D} \rightarrow \mathbb{K}$ eine $(\mathbb{K}_2, \mathbb{K})$ -Funktion und $\tilde{D} \supseteq f(D)$, so definieren wir $h \circ f : D \rightarrow \mathbb{K}, x \mapsto h(f(x))$.

Im Fall $\mathbb{K}_2 = \mathbb{C}$ kann man auch *Realteil* -und *Imaginärteiloperationen* definieren.

Definition 6.0.4 — REAL/IMAGINÄRTEIL VON FUNKTIONEN

Es sei $D \subseteq \mathbb{K}_1$ und sei $f : D \rightarrow \mathbb{C}$ eine $(\mathbb{K}_1, \mathbb{C})$ -Funktion.

- (i) Wir definieren $\text{Re}(f) : D \rightarrow \mathbb{R}, x \mapsto \text{Re}(f(x))$
- (ii) Wir definieren $\text{Im}(f) : D \rightarrow \mathbb{R}, x \mapsto \text{Im}(f(x))$

Um Funktionen zu visualisieren betrachten wir oft den *Funktionsgraphen*.

Definition 6.0.5 — FUNKTIONSGRAPH

Es sei $D \subseteq \mathbb{K}_1$ und $f : D \rightarrow \mathbb{K}_2$ eine $(\mathbb{K}_1, \mathbb{K}_2)$ -Funktion. Dann heißt die Menge

$$\Gamma_f := \{(x, y) \in D \times \mathbb{K}_2 : y = f(x)\} \subseteq \mathbb{K}_1 \times \mathbb{K}_2$$

der *Graph* von f .

Im Fall $\mathbb{K}_1 = \mathbb{K}_2 = \mathbb{R}$ ist der Graph stets eine Teilmenge von $\mathbb{R} \times \mathbb{R} = \mathbb{R}^2$ und lässt sich somit in die Ebene zeichnen. Beispielhaft tun wir dies mal für die Funktion abs aus dem vorigen Beispiel.

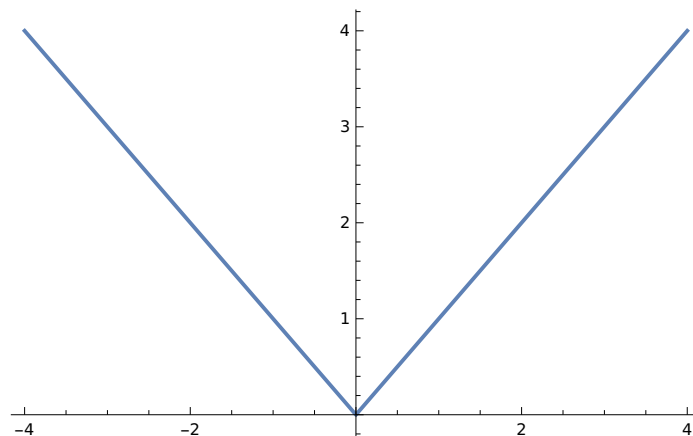


Abbildung 22: Der Graph Γ_{abs}

Nicht alle Funktionsgraphen lassen sich gut zeichnen. Ein Beispiel für einen nicht dankbar zu zeichnenden Graphen ist der der *Dirichlet-Funktion*

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}. \quad (253)$$

Da es zwischen zwei reellen Zahlen stets eine rationale Zahl gibt (vgl. Proposition 3.4.14), muss die Funktion beliebig nah an jeder reellen Zahl zwischen 0 und 1 hin- und herspringen. Unsere Zeichenkünste können dieses Phänomen unmöglich abbilden.

Unser Ziel in diesem Abschnitt ist es, herauszufinden, welche Funktionsgraphen sich *dankbar* mit unseren Zeichenkünsten aufzeichnen lassen.

6.1 Funktionsgrenzwerte

Damit man einen Funktionsgraph gut zeichnen kann, darf die Funktion offenbar nicht zu viele Sprünge verursachen.

Um das Sprungverhalten einer Funktion zu erfassen, studieren wir das *Grenzverhalten* von $(\mathbb{K}_1, \mathbb{K}_2)$ -Funktionen um einen gegebenen Punkt $x_0 \in \mathbb{K}_1$.

Damit ein solches Grenzverhalten studiert werden kann, darf x_0 allerdings nicht zu weit weg vom Definitionsbereich liegen. Genauer: x_0 muss ein *Häufungspunkt des Definitionsbereiches* sein.

Definition 6.1.1 — HÄUFUNGSPUNKT

Es sei $D \subseteq \mathbb{K}$ eine Menge. Ein Punkt $x_0 \in \mathbb{K}$ heißt *Häufungspunkt von D* falls es eine Folge $(x_n)_{n \in \mathbb{N}} \subseteq D$ gibt mit $\lim_{n \rightarrow \infty} x_n = x_0$. (a)

^aBeachte: Im Allgemeinen gibt es mehr als nur eine einzige Folge $(x_n)_{n \in \mathbb{N}}$ die obiges erfüllt. Schließlich würde jede Teilfolge einer solchen Folge auch obiges erfüllen.

Das Begriff *Häufungspunkt einer Menge* sollte nicht mit dem Begriff *Häufungswert einer Folge* verwechselt werden. Das sind zwei verschiedene Begriffe, die auch nichts miteinander zu tun haben.

Bemerkung 6.1.2

- Jedes Element von D ist klarerweise ein Häufungspunkt von D . Denn für $x_0 \in D$ konvergiert die konstante Folge $(x_0)_{n \in \mathbb{N}}$ klarerweise gegen x_0 .
- Ein Häufungspunkt von D muss aber nicht zwingend ein Element von D sein. Zum Beispiel ist 0 ein Häufungspunkt von $(0, \infty)_{\mathbb{R}}$ ^a, jedoch gilt $0 \notin (0, \infty)_{\mathbb{R}}$.

^aUm das mit Definition 6.1.1 zu sehen würde ich die Folge $(x_n)_{n \in \mathbb{N}} = (\frac{1}{n})_{n \in \mathbb{N}} \subseteq (0, \infty)$ betrachten!

Nun sind wir bereit, Funktionsgrenzwerte an Häufungspunkten des Definitionsbereiches zu definieren. Die Definition nutzt das uns bereits bekannte Konzept der *Folngengrenzwerte*.

Definition 6.1.3 — FUNKTIONSGRENZWERT

Es sei $f : D \rightarrow \mathbb{K}_2$ eine $(\mathbb{K}_1, \mathbb{K}_2)$ -Funktion und $x_0 \in \mathbb{K}_1$ ein Häufungspunkt von D . Wir sagen, f besitzt bei x_0 den Grenzwert $y_0 \in \mathbb{K}_2$ falls für jede Folge $(x_n)_{n \in \mathbb{N}} \subseteq D$ mit $\lim_{n \rightarrow \infty} x_n = x_0$ gilt, dass $\lim_{n \rightarrow \infty} f(x_n) = y_0$. Trifft obiges zu so schreibt man

$$\lim_{x \rightarrow x_0} f(x) = y_0.$$

Falls der Definitionsbereich D nicht aus dem Kontext klar sein sollte, schreibt man auch

$$\lim_{\substack{x \rightarrow x_0 \\ x \in D}} f(x) = y_0 \quad \text{oder} \quad \lim_{x \rightarrow x_0, x \in D} f(x) = y_0.$$

Außerdem führen wir auch die folgenden Schreibweisen ein:

$$f(x) \xrightarrow{x \rightarrow x_0} y_0 \quad \text{und} \quad [f(x) \longrightarrow y_0 \quad (x \rightarrow x_0)]$$

ein.

Die Aussage „ $\lim_{x \rightarrow x_0} f(x) = y_0$ “ kann man dann in Quantorschreibweise auch so formulieren:

$$(x_n)_{n \in \mathbb{N}} \subseteq D, \quad x_n \xrightarrow{n \rightarrow \infty} x_0 \quad \implies \quad f(x_n) \xrightarrow{n \rightarrow \infty} y_0.$$

Es sollte darauf hingewiesen werden, dass im Allgemeinen $y_0 \neq f(x_0)$ gilt. Das ist auch klar: Der Wert x_0 muss ja nicht einmal zwingend ein Element des Definitionsbereiches von f sein! In diesem Fall ist der Ausdruck $f(x_0)$ überhaupt nicht definiert. Ein Funktionsgrenzwert kann aber in diesem Fall trotzdem existieren.

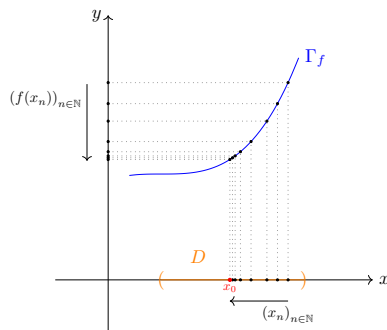


Abbildung 23: Funktionsgrenzwert anschaulich. Ein interaktives Schaubild dazu hat auch GEOGEBRA.

Bemerkung 6.1.4 — EINDEUTIGKEIT VON FUNKTIONSGRENZWERTEN

Funktionsgrenzwerte sind stets eindeutig!

Das erbt sich sofort aus der Eindeutigkeit von Folngrenzwerten, vgl. Proposition 2.3.13.

Genauer: Es sei $f : D \rightarrow \mathbb{K}_2$ eine $(\mathbb{K}_1, \mathbb{K}_2)$ -Funktion. Angenommen man hätte bei $x_0 \in \mathbb{K}_1$, einem Häufungswert von D , dass

$$\lim_{x \rightarrow x_0} f(x) = y_1 \quad \text{und} \quad \lim_{x \rightarrow x_0} f(x) = y_2$$

für $y_1, y_2 \in \mathbb{K}_2$. Wähle nun eine Folge $(x_n)_{n \in \mathbb{N}} \subseteq D$ mit $\lim_{n \rightarrow \infty} x_n = x_0$. (Eine solche Folge gibt es weil x_0 Häufungspunkt von D ist, vgl. Definition 6.1.1). Nach Definition 6.1.3 gilt dann

$$\lim_{n \rightarrow \infty} f(x_n) = y_1 \quad \text{und} \quad \lim_{n \rightarrow \infty} f(x_n) = y_2.$$

Aus Proposition 2.3.13 folgt dann $y_1 = y_2$. Die Eindeutigkeit von Funktionsgrenzwerten ist damit gezeigt.

Beispiel 6.1.5

Behauptung: $\lim_{x \rightarrow 2} (x^2 - 4) = 0.$

Genauer müsste man wohl schreiben: Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2 - 4$ erfüllt $\lim_{x \rightarrow 2} f(x) = 0$.

Beweis. Sei $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine beliebige Folge mit $\lim_{n \rightarrow \infty} x_n = 2$.

$$\mathbb{Z} : \quad \lim_{n \rightarrow \infty} (x_n^2 - 4) = 0.$$

Dazu:

$$x_n^2 - 4 = \underbrace{x_n}_{\rightarrow 2} \cdot \underbrace{x_n}_{\rightarrow 2} - 4 \xrightarrow[\text{GWR}]{n \rightarrow \infty} 2 \cdot 2 - 4 = 4 - 4 = 0.$$

Da $(x_n)_{n \in \mathbb{N}}$ beliebig war folgt die Behauptung.

Auf dieselbe Art und Weise kann man bei sämtlichen Polynomen alle Funktionsgrenzwerte berechnen. Probieren Sie es aus!

Macht eine Funktion an einer Stelle x_0 einen *Sprung*, so existiert der Funktionsgrenzwert dort nicht.

Beispiel 6.1.6

Gegeben sei $f : \mathbb{R} \rightarrow \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} 1 & \text{falls } x \leq 0 \\ 2 & \text{falls } x > 0. \end{cases}$$

Dann besitzt f bei 0 keinen Grenzwert.

Beweis. Angenommen, f besitzt einen Grenzwert $y_0 \in \mathbb{R}$. Dann gilt

$$(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}, \quad x_n \xrightarrow{n \rightarrow \infty} 0 \quad \implies \quad f(x_n) \xrightarrow{n \rightarrow \infty} y_0. \quad (254)$$

Betrachte nun die Folge $(\frac{1}{n})_{n \in \mathbb{N}} \subseteq \mathbb{R}$. Da $\frac{1}{n} \xrightarrow{n \rightarrow \infty} 0$ haben wir nach (256)

$$y_0 = \lim_{n \rightarrow \infty} f\left(\frac{1}{n}\right) = \lim_{n \rightarrow \infty} 2 = 2.$$

Als nächstes betrachte die Folge $(\frac{-1}{n})_{n \in \mathbb{N}} \subseteq \mathbb{R}$. Da $\frac{-1}{n} \xrightarrow{n \rightarrow \infty} 0$ haben wir nach (256)

$$y_0 = \lim_{n \rightarrow \infty} f\left(\frac{-1}{n}\right) = \lim_{n \rightarrow \infty} 1 = 1.$$

Ein Widerspruch, denn wir haben vorher bereits $y_0 = 2$ gezeigt.

Beispiel 6.1.7 — DIE DIRICHLET-FUNKTION

Es sei wie in (253) die *Dirichlet-Funktion*

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) := \begin{cases} 1 & \text{falls } x \in \mathbb{Q} \\ 0 & \text{falls } x \notin \mathbb{Q} \end{cases}.$$

gegeben.

Behauptung: f besitzt bei 0 keinen Grenzwert.

Beweis. Angenommen das Gegenteil, d.h.

$$\exists \lim_{x \rightarrow 0} f(x) =: y_0. \quad (255)$$

Dann gilt wie im vorigen Beispiel

$$(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}, \quad x_n \xrightarrow{n \rightarrow \infty} 0 \quad \implies \quad f(x_n) \xrightarrow{n \rightarrow \infty} y_0. \quad (256)$$

Betrachte nun zuerst die Folge $(x_n)_{n \in \mathbb{N}} = (\frac{1}{n})_{n \in \mathbb{N}}$, die klarerweise gegen 0 konvergiert. Da $\frac{1}{n} \in \mathbb{Q}$ für alle $n \in \mathbb{N}$ gilt $f(\frac{1}{n}) = 1$ für alle $n \in \mathbb{N}$. Aus (255) folgt

$$y_0 \stackrel{\text{Def 6.1.3}}{=} \lim_{n \rightarrow \infty} f(\frac{1}{n}) = \lim_{n \rightarrow \infty} 1 = 1. \quad (257)$$

Betrachte nun als nächstes die Folge $(x_n)_{n \in \mathbb{N}} = (\frac{\sqrt{2}}{n})_{n \in \mathbb{N}}$, die klarerweise auch gegen 0 konvergiert. Beachte: Für alle $n \in \mathbb{N}$ ist $\frac{\sqrt{2}}{n} \notin \mathbb{Q}$, denn wäre $\frac{\sqrt{2}}{n} \in \mathbb{Q}$ so wäre (weil $\mathbb{Q}$ ein Körper ist) auch das Produkt $n \cdot \frac{\sqrt{2}}{n} \in \mathbb{Q}$. Dies ergibt jedoch $\sqrt{2} \in \mathbb{Q}$, ein Widerspruch zu Lemma 3.4.3. Aus dieser Diskussion nehmen wir mit, dass $f(\frac{\sqrt{2}}{n}) = 0$ für alle $n \in \mathbb{N}$. Es folgt

$$y_0 \stackrel{\text{Def 6.1.3}}{=} \lim_{n \rightarrow \infty} f(\frac{\sqrt{2}}{n}) = \lim_{n \rightarrow \infty} 0 = 0.$$

Die vorige Gleichung widerspricht (257).

Mit ähnlichen Methoden kann man zeigen, dass f bei keinem einzigen Wert $x_0 \in \mathbb{R}$ einen Grenzwert besitzt. Details überlassen wir den Lesenden als Übungsaufgabe.

Die nachfolgende Proposition formuliert *Rechenregeln für Funktionsgrenzwerte*. Diese sind direkte Konsequenzen aus den Grenzwertregeln für Folgen.

Proposition 6.1.8 — RECHENREGELN FÜR FUNKTIONSGRENZWERTE

Es seien $f, g : D \rightarrow \mathbb{K}_2$ zwei $(\mathbb{K}_1, \mathbb{K}_2)$ -Funktionen und $x_0 \in \mathbb{K}_1$ ein Häufungspunkt von D . Ferner seien die Grenzwerte $\lim_{x \rightarrow x_0} f(x)$ und $\lim_{x \rightarrow x_0} g(x)$ existent. Dann existieren auch

- (i) $\lim_{x \rightarrow x_0} (f + g)(x)$ und es gilt $\lim_{x \rightarrow x_0} (f + g)(x) = \lim_{x \rightarrow x_0} f(x) + \lim_{x \rightarrow x_0} g(x)$.
- (ii) $\lim_{x \rightarrow x_0} (fg)(x)$ und es gilt $\lim_{x \rightarrow x_0} (fg)(x) = \lim_{x \rightarrow x_0} f(x) \cdot \lim_{x \rightarrow x_0} g(x)$.
- (iii) (für alle $\lambda \in \mathbb{K}_2$) $\lim_{x \rightarrow x_0} (\lambda f)(x)$ und $\lim_{x \rightarrow x_0} (\lambda f)(x) = \lambda \cdot \lim_{x \rightarrow x_0} f(x)$.
- (iv) (sofern $\lim_{x \rightarrow x_0} g(x) \neq 0$) $\lim_{x \rightarrow x_0} \left(\frac{f}{g} \right)(x)$ und $\lim_{x \rightarrow x_0} \left(\frac{f}{g} \right)(x) = \frac{\lim_{x \rightarrow x_0} f(x)}{\lim_{x \rightarrow x_0} g(x)}$

Beweis. Wir zeigen nur (i) und (iv). Der Rest ist analog. Dazu seien

$$y_1 := \lim_{x \rightarrow x_0} f(x) \quad \text{und} \quad y_2 := \lim_{x \rightarrow x_0} g(x). \quad (258)$$

Zu (i).

$$\mathbb{Z}: \quad \lim_{x \rightarrow x_0} (f + g)(x) = y_1 + y_2.$$

Hierfür arbeiten wir mit Definition 6.1.3. Es sei dazu $(x_n)_{n \in \mathbb{N}} \subseteq D$ eine Folge mit $\lim_{n \rightarrow \infty} x_n = x_0$. Wir müssen zeigen, dass $\lim_{n \rightarrow \infty} (f + g)(x_n) = y_1 + y_2$. Dazu:

$$(f + g)(x_n) = \underbrace{f(x_n)}_{\rightarrow y_1, \text{ Def 6.1.3}} + \underbrace{g(x_n)}_{\rightarrow y_2, \text{ Def 6.1.3}} \xrightarrow[\text{GWR}]{n \rightarrow \infty} y_1 + y_2.$$

Da die Folge $(x_n)_{n \in \mathbb{N}}$ beliebig war folgt aus Definition 6.1.3, dass

$$\lim_{x \rightarrow x_0} (f + g)(x) = y_1 + y_2 \stackrel{(258)}{=} \lim_{x \rightarrow x_0} f(x) + \lim_{x \rightarrow x_0} g(x).$$

Zu (iv). Hier könnte ein Problem entstehen, denn $\frac{f}{g}$ ist a priori nur auf

$$\hat{D} := \{x \in D : g(x) \neq 0\} \quad (259)$$

definiert. Damit der Ausdruck $\lim_{x \rightarrow x_0} \left(\frac{f}{g}\right)(x)$ überhaupt Sinn ergibt muss man zunächst argumentieren, dass x_0 ein Häufungspunkt von $\hat{D}$ ist.

ZWISCHENBEHAUPTUNG. x_0 ist ein Häufungspunkt von $\hat{D}$.

Dazu wähle zuerst eine Folge $(x_n)_{n \in \mathbb{N}} \subseteq D$ mit $\lim_{n \rightarrow \infty} x_n = x_0$ (die es ja gibt, weil x_0 ja ein Häufungspunkt von D ist). Definieren nun

$$M := \{n \in \mathbb{N} : x_n \notin \hat{D}\} \stackrel{(259)}{=} \{n \in \mathbb{N} : g(x_n) = 0\}. \quad (260)$$

Wir behaupten, dass M endlich ist. In der Tat: Wäre M unendlich, so gäbe es in M eine strikt aufsteigende Folge $(\ell_n)_{n \in \mathbb{N}} \subseteq M$, vgl. Bemerkung 3.5.15. Nun gilt für $n \in \mathbb{N}$

$$\ell_n \in M \quad \stackrel{(260)}{\implies} \quad g(x_{\ell_n}) = 0.$$

Es folgt $g(x_{\ell_n}) = 0$ für alle $n \in \mathbb{N}$ und somit auch

$$\lim_{n \rightarrow \infty} g(x_{\ell_n}) = 0. \quad (261)$$

Da aber $(x_{\ell_n})_{n \in \mathbb{N}}$ als Teilfolge von $(x_n)_{n \in \mathbb{N}}$ aber auch gegen x_0 konvergiert, gilt nach Definition 6.1.3

$$\lim_{n \rightarrow \infty} g(x_{\ell_n}) = \lim_{x \rightarrow x_0} g(x) \underset{\text{Vorr.}}{\neq} 0,$$

was (261) widerspricht. Somit ist M endlich und daher gibt es ein $N \in \mathbb{N}$ mit $M \subseteq \{1, \dots, N\}$. Mit anderen Worten gilt für alle $n \geq N + 1$, dass $n \notin M$, d.h.

$$x_n \in \hat{D} \quad \forall n \geq N + 1.$$

Somit ist $(x_n)_{n \geq N+1} \subseteq \hat{D}$ eine Folge in $\hat{D}$, die gegen x_0 konvergiert. Mit anderen Worten: x_0 ist Häufungspunkt von $\hat{D}$.

ZWISCHENBEHAUPTUNG ENDE.

Es sei nun $(x_n)_{n \in \mathbb{N}}$ eine beliebige Folge mit $\lim_{n \rightarrow \infty} x_n = x_0$. Dann gilt wegen (258), dass $\lim_{n \rightarrow \infty} f(x_n) = y_1$ und $\lim_{n \rightarrow \infty} g(x_n) = y_2 [\neq 0]$. Aus den Grenzwertregeln folgt

$$\left(\frac{f}{g}\right)(x_n) = \frac{f(x_n)}{g(x_n)} \xrightarrow{n \rightarrow \infty} \frac{y_1}{y_2}$$

Da die Folge $(x_n)_{n \in \mathbb{N}}$ beliebig war gilt

$$\lim_{x \rightarrow x_0} \left(\frac{f}{g}\right)(x) = \frac{y_1}{y_2} = \frac{\lim_{x \rightarrow x_0} f(x)}{\lim_{x \rightarrow x_0} g(x)}.$$

Beispiel 6.1.9

Wir zeigen, dass

$$\lim_{x \rightarrow 2} \frac{x^2 - 3x + 2}{x^2 - 4} = \frac{1}{4}. \quad (a)$$

Beweis. Die Regel aus Proposition 6.1.8 (iv) ist nicht anwendbar, denn nach Beispiel 6.1.5 gilt

$$\lim_{x \rightarrow 2} (x^2 - 4) = 0.$$

Schreibt man jedoch $x^2 - 4 = (x - 2)(x + 2)$ und $x^2 - 3x + 2 = (x - 1)(x - 2)$, so gilt für

alle $x \in \mathbb{R} \setminus \{2, -2\}$

$$\frac{x^2 - 3x + 2}{x^2 - 4} = \frac{(x-1)(x-2)}{(x-2)(x+2)} = \frac{x-1}{x+2}$$

Man zeigt leicht (wie in Beispiel 6.1.5, sogar etwas leichter), dass

$$\lim_{x \rightarrow 2} (x-1) = 1 \quad \text{und} \quad \lim_{x \rightarrow 2} (x+2) = 4.$$

Dann folgt aus Proposition 6.1.8 (iv), dass

$$\lim_{x \rightarrow 2} \frac{x-1}{x+2} = \frac{1}{4}.$$

^aGenauer müsste man hier eigentlich

$$\lim_{\substack{x \rightarrow 2, \\ x \in \mathbb{R} \setminus \{-2, 2\}}} \frac{x^2 - 3x - 4}{x^2 - 4}$$

schreiben, um den Definitionsbereich zu spezifizieren. Außerdem müsste man eigentlich überprüfen, dass es sich bei der Zahl 2 um einen Häufungspunkt des Definitionsbereiches handelt. Überlegen Sie sich gerne als Übungsaufgabe, dass 2 ein Häufungspunkt von $\mathbb{R} \setminus \{-2, 2\}$ ist!

Beispiel 6.1.10

Wir zeigen, dass für alle $n \in \mathbb{N}$ gilt

$$\lim_{z \rightarrow 1} \frac{z^n - 1}{z - 1} = n. \quad (a)$$

Beachte: Proposition 6.1.8 (iv) ist wieder nicht anwendbar, da $\lim_{z \rightarrow 1} (z-1) = 0$. Allerdings gilt wegen der geometrischen Summenformel (Proposition 3.3.20) für alle $z \in \mathbb{C} \setminus \{1\}$

$$\frac{z^n - 1}{z - 1} = \frac{1 - z^n}{1 - z} = \sum_{k=0}^{n-1} z^k. \quad (262)$$

Daraus folgt

$$\lim_{z \rightarrow 1} \frac{z^n - 1}{z - 1} = \lim_{z \rightarrow 1} \sum_{k=0}^{n-1} z^k \stackrel{\text{Prop 6.1.8(i)}}{=} \sum_{k=0}^{n-1} \lim_{z \rightarrow 1} z^k \stackrel{\text{Bsp 6.1.5}}{=} \sum_{k=0}^{n-1} 1 = n.$$

^aGenauer müsste man eigentlich wieder

$$\lim_{\substack{z \rightarrow 1 \\ z \in \mathbb{C} \setminus \{1\}}} \frac{z^n - 1}{z - 1}$$

schreiben, um den Definitionsbereich zu spezifizieren. Klarheit über den Definitionsbereich ist insbesondere wichtig, wenn wir den Ausdruck wie hier vor Betrachtung des Grenzwertes noch umformen wollen wie zum Beispiel in (262).

Im Fall $\mathbb{K}_2 = \mathbb{C}$ kommen übrigens noch weitere Grenzwertregeln dazu.

Proposition 6.1.11

Es sei $f : D \rightarrow \mathbb{C}$ eine $(\mathbb{K}_1, \mathbb{C})$ -Funktion und $x_0 \in \mathbb{K}_1$ ein Häufungspunkt von D . Dann gilt

$$\exists \lim_{x \rightarrow x_0} f(x) \iff \left(\exists \lim_{x \rightarrow x_0} \operatorname{Re}(f(x)) \wedge \exists \lim_{x \rightarrow x_0} \operatorname{Im}(f(x)) \right).$$

Ist eine der obigen Aussagen wahr so gilt auch

$$\lim_{x \rightarrow x_0} \operatorname{Re}(f(x)) = \operatorname{Re} \left(\lim_{x \rightarrow x_0} f(x) \right) \quad \text{und} \quad \lim_{x \rightarrow x_0} \operatorname{Im}(f(x)) = \operatorname{Im} \left(\lim_{x \rightarrow x_0} f(x) \right).$$

Beweis. Nutze Definition 6.1.3 und Proposition 4.3.4. Details sind den Lesenden als Übungsaufgabe überlassen.

Der Spezialfall $\mathbb{K} = \mathbb{R}$

Im Spezialfall $\mathbb{K}_1 = \mathbb{R}$ und/oder $\mathbb{K}_2 = \mathbb{R}$ haben wir zusätzlich noch eine *Ordnungsrelation* $\leq$ auf Definitionsbereich und/oder Zielbereich zur Verfügung. Daraus ergeben sich einige Konsequenzen für Funktionsgrenzwerte.

Vergleichsregeln und Sandwichlemma (im Fall $\mathbb{K}_2 = \mathbb{R}$) Hier sehen wir uns zusätzliche Resultate an, wenn der *Zielbereich* reell ist, d.h. $\mathbb{K}_2 = \mathbb{R}$.

Aus der Vergleichsregel für Folgengrenzwerte erbt sich eine Vergleichsregel für Funktionsgrenzwerte.

Proposition 6.1.12 — VERGLEICHSREGEL FÜR FUNKTIONSGRENZWERTE

Es sei $D \subseteq \mathbb{K}_1$ und seien $f, g : D \rightarrow \mathbb{R}$ Funktionen mit

$$f(x) \leq g(x) \text{ für alle } x \in D.$$

Wenn $\lim_{x \rightarrow x_0} f(x)$ und $\lim_{x \rightarrow x_0} g(x)$ existieren, dann gilt

$$\lim_{x \rightarrow x_0} f(x) \leq \lim_{x \rightarrow x_0} g(x).$$

Beweis. Wir benennen zunächst

$$y_1 := \lim_{x \rightarrow x_0} f(x) \quad \text{und} \quad y_2 := \lim_{x \rightarrow x_0} g(x).$$

$$\mathbb{Z} : \quad y_1 \leq y_2.$$

Es sei $(x_n)_{n \in \mathbb{N}} \subseteq D$ eine Folge mit $\lim_{n \rightarrow \infty} x_n = x_0$. Aus der Voraussetzung folgt, dass

$$f(x_n) \leq g(x_n) \quad \forall n \in \mathbb{N}. \quad (263)$$

Nun beachte, dass nach Definition 6.1.3 gilt

$$\lim_{n \rightarrow \infty} f(x_n) = y_1 \quad \text{und} \quad \lim_{n \rightarrow \infty} g(x_n) = y_2.$$

Nach den Grenzwert(vergleichs)regeln folgt aus (263), dass

$$\lim_{n \rightarrow \infty} f(x_n) \leq \lim_{n \rightarrow \infty} g(x_n),$$

also $y_1 \leq y_2$, was zu zeigen war.

Desweiteren gibt es ein *Sandwichlemma* für Funktionsgrenzwerte.

Proposition 6.1.13 — SANDWICHLEMA FÜR FUNKTIONSGRENZWERTE

Es sei $D \subseteq \mathbb{K}_1$ und seien $f, g, h : D \rightarrow \mathbb{R}$ Funktionen mit

$$f(x) \leq g(x) \leq h(x) \quad \forall x \in D.$$

Falls

$$\lim_{x \rightarrow x_0} f(x) \quad \text{und} \quad \lim_{x \rightarrow x_0} h(x) \quad \text{existieren und miteinander übereinstimmen,}$$

so existiert auch $\lim_{x \rightarrow x_0} g(x)$ und stimmt mit den beiden obigen Werten überein.

Beweis. Überlassen wir den Lesenden als Übungsaufgabe.

Einseitige Grenzwerte (im Fall $\mathbb{K}_1 = \mathbb{R}$) Hier interessiert uns, was für weitere Konsequenzen die Reellwertigkeit des *Definitionsgebietes* hat, d.h. $\mathbb{K}_1 = \mathbb{R}$.

Wegen der Anordnung des Definitionsbereiches ergibt sich dadurch die Möglichkeit, *links-* und *rechtsseitige* Grenzwerte zu betrachten.

Definition 6.1.14 — EINSEITIGE GRENZWERTE

Es sei $D \subseteq \mathbb{R}$ und $f : D \rightarrow \mathbb{K}_2$ eine Funktion. Sei dazu $x_0 \in \mathbb{R}$.

(i) Definiere den *Links-Cut* von f bei x_0 durch

$$f_\ell : D \cap (-\infty, x_0] \rightarrow \mathbb{K}_2, \quad f_\ell(x) := f(x).$$

Wir sagen f besitzt bei x_0 den *linksseitigen Grenzwert* y_ℓ , falls $\lim_{x \rightarrow x_0} f_\ell(x) = y_\ell$. ^(a)

Wir bezeichnen

$$\lim_{x \rightarrow x_0 -} f(x) := y_\ell.$$

(ii) Definiere den *punktierten Links-Cut* von f bei x_0 durch

$$\dot{f}_\ell : D \cap (-\infty, x_0) \rightarrow \mathbb{K}_2, \quad \dot{f}_\ell(x) := f(x).$$

Wir sagen f besitzt bei x_0 den *punktierten linksseitigen Grenzwert* $\dot{y}_\ell$, falls $\lim_{x \rightarrow x_0} \dot{f}_\ell(x) = \dot{y}_\ell$. Wir bezeichnen

$$\lim_{x \rightarrow x_0 -}^{\circ} f(x) := \dot{y}_\ell.$$

(iii) Definiere den *Rechts-Cut* von f bei x_0 durch

$$f_r : D \cap [x_0, \infty) \rightarrow \mathbb{K}_2, \quad f_r(x) := f(x).$$

Wir sagen f besitzt bei x_0 den *rechtsseitigen Grenzwert* y_r , falls $\lim_{x \rightarrow x_0} f_r(x) = y_r$. Wir bezeichnen

$$\lim_{x \rightarrow x_0 +} f(x) := y_r.$$

(iv) Definiere den *punktierten Rechts-Cut* von f bei x_0 durch

$$\dot{f}_r : D \cap (x_0, \infty) \rightarrow \mathbb{K}_2, \quad \dot{f}_r(x) := f(x).$$

Wir sagen f besitzt bei x_0 den *punktierten rechtsseitigen Grenzwert* $\dot{y}_r$, falls $\lim_{x \rightarrow x_0} \dot{f}_r(x) = \dot{y}_r$. Wir bezeichnen

$$\lim_{x \rightarrow x_0 +}^{\circ} f(x) := \dot{y}_r.$$

^aBeachte: Hierfür muss vorausgesetzt werden, dass x_0 ein Häufungspunkt des Definitionsbereiches von f_ℓ , also von $D \cap (-\infty, x_0]$ ist. Analoges in Fall (ii),(iii),(iv).

Mit anderen Worten gilt

$$\lim_{x \rightarrow x_0 -} f(x) = \lim_{\substack{x \rightarrow x_0 \\ x \in D \cap (-\infty, x_0]}} f(x) \quad \text{und} \quad \lim_{x \rightarrow x_0 +} f(x) = \lim_{\substack{x \rightarrow x_0 \\ x \in D \cap [x_0, \infty)}} f(x). \quad (264)$$

Für die punktierten Grenzwerte muss man $(-\infty, x_0]$ durch $(-\infty, x_0)$ und $[x_0, \infty)$ durch (x_0, ∞) ersetzen.

Definition 6.1.14 klingt zunächst etwas kontraintuitiv, wenn man bedenkt, dass die Funktionswerte von f_ℓ , $\dot{f}_\ell$, f_r und $\dot{f}_r$, stets mit denen von f übereinstimmen und sich die Funktionen

lediglich im Definitionsbereich unterscheiden. Am nachfolgenden Beispiel sieht man aber, dass das Bilden von links-cuts und rechts-cuts die Grenzwertsituation wirklich verändert.

Beispiel 6.1.15

Betrachte noch einmal $f : \mathbb{R} \rightarrow \mathbb{R}$ aus Beispiel 6.1.6 definiert durch

$$f(x) = \begin{cases} 1 & \text{falls } x \leq 0 \\ 2 & \text{falls } x > 0. \end{cases}$$

Nehmen wir nun $x_0 = 0$. Dann ist der links-cut $f_\ell : (-\infty, 0] \rightarrow \mathbb{R}$ gegeben durch

$$f_\ell(x) = 1 \quad \forall x \in (-\infty, 0].$$

Man überzeugt sich leicht davon, dass dann

$$\lim_{x \rightarrow 0-} f(x) = \lim_{x \rightarrow 0} f_\ell(x) = 1.$$

Ferner ist der rechts-cut $f_r : [0, \infty) \rightarrow \mathbb{R}$ gegeben durch

$$f_r(x) = \begin{cases} 2 & \text{falls } x > 0 \\ 1 & \text{falls } x = 0 \end{cases} \quad \forall x \in [0, \infty).$$

Man sieht, dass f_r bei 0 keinen Grenzwert besitzt. Details dazu sind den Lesenden als Übungsaufgabe überlassen. Wir folgern

$$\lim_{x \rightarrow 0+} f(x) \quad \text{existiert nicht.}$$

Nun bilden wir noch den *punktierten* rechts-cut $\dot{f}_r : (0, \infty) \rightarrow \mathbb{R}$. Dieser erfüllt

$$\dot{f}_r(x) = 2 \quad \forall x \in (0, \infty).$$

Es gilt dann

$$\lim_{x \rightarrow 0+}^{\circ} f(x) = \lim_{x \rightarrow 0} \dot{f}_r(x) = 2.$$

Wir haben gesehen, dass es im Allgemeinen einen Unterschied macht, ob man punktierte oder nicht-punktierte Grenzwerte betrachtet. Manchmal macht es aber auch keinen Unterschied: Ist $x_0 \notin D$, so ist $D \cap (-\infty, x_0] = D \cap (-\infty, x_0)$. Somit stimmen Links-Cut und punktierter Links-Cut miteinander überein! Selbiges gilt für den Rechts-Cut und den punktierten Rechts-Cut.

Existieren links -und rechtsseitiger Grenzwert und stimmen überein, so existiert auch der eigentliche Grenzwert, wie die nachfolgende Proposition besagt.

12.01.2026

Proposition 6.1.16 — EINSEITIGE GRENZWERTE UND BEIDSEITIGER GRENZWERT

Es sei $D \subseteq \mathbb{R}$ und $f : D \rightarrow \mathbb{K}_2$ eine Funktion. Ferner sei $x_0 \in \mathbb{R}$ so, dass die einseitigen Grenzwerte $\lim_{x \rightarrow x_0-} f(x)$ und $\lim_{x \rightarrow x_0+} f(x)$ existieren^a. Dann gilt

$$\exists \lim_{x \rightarrow x_0} f(x) \iff \lim_{x \rightarrow x_0-} f(x) = \lim_{x \rightarrow x_0+} f(x).$$

Ferner stimmt in diesem Fall auch $\lim_{x \rightarrow x_0} f(x)$ mit den beiden einseitigen Grenzwerten überein.

Zusatz. Die Aussage bleibt richtig wenn man einen der beiden einseitigen Grenzwerte durch den punktierten Grenzwert $\lim_{x \rightarrow x_0}$ ersetzt (jedoch nicht beide!).

^aInsbesondere muss x_0 ein Häufungspunkt von $D \cap (-\infty, x_0]$ und $D \cap [x_0, \infty)$ sein, vgl. Definition 6.1.14 (Fußnote)

***Beweis.** Erst zum Fall, dass keiner der beiden Grenzwerte punktiert ist (also ohne Zusatz).

Zuerst zu „ $\Rightarrow$ “. Sei etwa

$$y_0 := \lim_{x \rightarrow x_0} f(x).$$

Betrachte nun den Links-Cut $f_\ell : D \cap (-\infty, x_0] \rightarrow \mathbb{K}_2$, vgl Definition 6.1.14. Wir behaupten, dass $\lim_{x \rightarrow x_0} f_\ell(x) = y_0$. Es sei dazu $(x_n)_{n \in \mathbb{N}} \subseteq D \cap (-\infty, x_0]$ eine Folge mit $\lim_{n \rightarrow \infty} x_n = x_0$. Dann gilt insbesondere auch $(x_n)_{n \in \mathbb{N}} \subseteq D$. Man erhält

$$f_\ell(x_n) \xrightarrow[\text{Def 6.1.14}]{} f(x_n) \xrightarrow[\text{Def 6.1.3}]{n \rightarrow \infty} y_0.$$

Nun zu „ $\Leftarrow$ “. Definiere

$$y^* := \lim_{x \rightarrow x_0-} f(x) = \lim_{x \rightarrow x_0+} f(x).$$

Zu zeigen ist, dass $\lim_{x \rightarrow x_0} f(x) = y^*$. Angenommen es gilt $\neg(\lim_{x \rightarrow x_0} f(x) = y^*)$. Dann gibt es eine Folge $(x_n)_{n \in \mathbb{N}} \subseteq D$ mit $\lim_{n \rightarrow \infty} x_n = x_0$ so, dass $f(x_n) \not\rightarrow y^*$ (mit $n \rightarrow \infty$). In Quantorenschreibweise bedeutet das

$$\neg(\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall n \geq n_0 : |f(x_n) - y^*| \leq \varepsilon)$$

Nutzen wir unser Wissen über Verneinungen mit Quantoren (vgl. Überlegung 2.2.4) so erhalten wir

$$\exists \varepsilon_0 > 0 : \quad \forall m \in \mathbb{N} \quad \exists n_m \geq m : \quad |f(x_{n_m}) - y^*| > \varepsilon_0. \quad (265)$$

Betrachte nun die Folge $(x_{n_m})_{m \in \mathbb{N}}$. Man überzeugt sich leicht davon, dass $\lim_{m \rightarrow \infty} x_{n_m} = x_0$. Definiere

$$S_\ell := \{m \in \mathbb{N} : x_{n_m} \leq x_0\} \quad \text{und} \quad S_r := \{m \in \mathbb{N} : x_{n_m} \geq x_0\}. \quad (266)$$

Beachte, dass $S_\ell \cup S_r = \mathbb{N}$. Insbesondere ist mindestens eine der beiden Mengen unendlich. (Denn wären beide endlich, so wäre auch die Vereinigung endlich. Diese ist hier jedoch $\mathbb{N}$ und somit sicherlich unendlich...)

FALL 1. S_ℓ ist unendlich.

Nach Bemerkung 3.5.15 gibt es dann eine strikt aufsteigende Folge $(k_m)_{m \in \mathbb{N}} \subseteq S_\ell$. Für alle $n \in \mathbb{N}$ gilt dann $x_{n_{k_m}} \leq x_0$ und somit $(x_{n_{k_m}})_{m \in \mathbb{N}} \subseteq D \cap (-\infty, x_0]$. Da $(x_{n_{k_m}})_{m \in \mathbb{N}}$ Teilfolge von $(x_{n_m})_{m \in \mathbb{N}}$ ist, gilt $\lim_{m \rightarrow \infty} x_{n_{k_m}} = x_0$. Beobachte nun

$$y^* = \lim_{x \rightarrow x_0-} f(x) = \lim_{x \rightarrow x_0} f_\ell(x) \stackrel{\text{Def 6.1.3}}{=} \lim_{n \rightarrow \infty} f_\ell(x_{n_{k_m}}) = \lim_{n \rightarrow \infty} f(x_{n_{k_m}}).$$

Dies impliziert $\lim_{m \rightarrow \infty} |f(x_{n_{k_m}}) - y^*| = 0$. Allerdings folgt aus (265), dass $|f(x_{n_{k_m}}) - y^*| \geq \varepsilon_0$ für alle $m \in \mathbb{N}$. Es ergibt sich ein Widerspruch, denn dann ist

$$0 = \lim_{m \rightarrow \infty} |f(x_{n_{k_m}}) - y^*| \geq \lim_{m \rightarrow \infty} \varepsilon_0 = \varepsilon_0.$$

Somit ist die Annahme $\neg(\lim_{x \rightarrow x_0} f(x) = y^*)$ falsch und es folgt die Behauptung.

FALL 2. S_ℓ ist endlich. Dann ist insbesondere S_r unendlich.

Verfahre komplett spiegelbildlich zu Fall 1, und ersetze f_ℓ durch f_r . Folgere denselben Widerspruch und erhalte die Behauptung.

Nun noch zum Zusatz. Falls einer der beiden Grenzwerte punktiert ist, ersetze in den Definitionen von S_ℓ und S_r in (266) eine der beiden Ungleichungen durch eine strikte Ungleichung. Nach wie vor gilt dann $S_\ell \cup S_r = \mathbb{N}$. Ansonsten lässt sich der Beweis Zeile für Zeile übertragen.

Beispiel 6.1.17

Es sei $f : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$f(x) := \begin{cases} x^2 & \text{falls } x < 1 \\ x^3 & \text{falls } x \geq 1 \end{cases}$$

$$\text{Behauptung:} \quad \lim_{x \rightarrow 1} f(x) = 1.$$

Beweis. Der punktierte(!) links-cut $\dot{f}_\ell : (-\infty, 1) \rightarrow \mathbb{R}$ gegeben durch $\dot{f}_\ell(x) = x^2$ und der (nicht-punktierte) rechts-cut $f_r : [1, \infty) \rightarrow \mathbb{R}$ ist gegeben durch $f_r(x) = x^3$. Man erhält

$$\lim_{x \rightarrow 1-}^\circ f(x) = \lim_{x \rightarrow 1} \dot{f}_\ell(x) = \lim_{x \rightarrow 1} x^2 \stackrel{\text{Bsp 6.1.5}}{=} 1^2 = 1$$

und

$$\lim_{x \rightarrow 1+} f(x) = \lim_{x \rightarrow 1} f_r(x) = \lim_{x \rightarrow 1} x^3 \stackrel{\text{Bsp 6.1.5}}{=} 1^3 = 1.$$

Wir lernen

$$\lim_{x \rightarrow 1-}^\circ f(x) = \lim_{x \rightarrow 1+} f(x) = 1.$$

Proposition 6.1.16 impliziert dann

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

Man sieht, dass es in den Anwendungen manchmal vorteilhaft sein kann, einen der Grenzwerte in Proposition 6.1.16 durch einen punktierten Grenzwert zu ersetzen. Man sollte allerdings drauf hinweisen, dass es nicht zulässig ist, beide Grenzwerte zu ersetzen!

Beispiel 6.1.18

Die Aussage von Proposition 6.1.16 wird falsch, wenn man beide einseitige Grenzwerte durch punktierte Grenzwerte ersetzt. In der Tat: Betrachte $f : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$f(x) = \begin{cases} 1 & x \neq 0 \\ 2 & x = 0 \end{cases}$$

Behauptung. Die einseitigen punktierten Grenzwerte $\lim_{x \rightarrow 0\pm}^\circ f(x)$ existieren und stimmen überein. Jedoch ist $\lim_{x \rightarrow 0} f(x)$ nicht existent.

Beweis. Da $f(x) = 1$ für alle $x \in (-\infty, 0)$ gilt $\dot{f}_\ell(x) = 0$ für alle $x \in (-\infty, 0)$ und somit

$$\lim_{x \rightarrow 0-}^\circ f(x) = \lim_{x \rightarrow 0} \dot{f}_\ell(x) = \lim_{x \rightarrow 0} 0 = 1.$$

Ganz analog zeigt man (da $f(x) = 1$ für alle $x \in (0, \infty)$)

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

Jedoch ist $\lim_{x \rightarrow 0} f(x)$ nicht existent. Um das zu sehen kann man die Folgen $(x_n^{(1)})_{n \in \mathbb{N}} = (\frac{1}{n})_{n \in \mathbb{N}}$ und $(x_n^{(2)})_{n \in \mathbb{N}} = (0)_{n \in \mathbb{N}}$ betrachten, die beide gegen 0 konvergieren. Dann gilt $f(x_n^{(1)}) = 1$ für alle $n \in \mathbb{N}$ und $f(x_n^{(2)}) = 2$ für alle $n \in \mathbb{N}$. Man folgert

$$\lim_{n \rightarrow \infty} f(x_n^{(1)}) \neq \lim_{n \rightarrow \infty} f(x_n^{(2)}) \quad (\text{da } 1 \neq 2). \quad (267)$$

Würde $\lim_{x \rightarrow 0} f(x)$ existieren, so müsste aber für jede Folge $(x_n)_{n \in \mathbb{N}}$ die gegen 0 konvergiert auch $\lim_{n \rightarrow \infty} f(x_n)$ existieren und mit $\lim_{x \rightarrow 0} f(x)$ übereinstimmen. Das kann nicht gelten wenn wie in (267) zwei verschiedene Grenzwerte möglich sind.

FAZIT: In Proposition 6.1.16 darf man wirklich nur maximal einen Grenzwert durch einen punktierten Grenzwert ersetzen.

Grenzwerte im Unendlichen Für reelle Folgen $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ haben wir auch die *bestimmte Divergenz* „ $\lim_{n \rightarrow \infty} x_n = \pm \infty$ “ definiert, vgl. Definition 2.3.35. Aufbauend darauf können wir auch Funktionsgrenzwerte im Unendlichen studieren.

Definition 6.1.19 — FUNKTIONSGRENZWERTE IM UNENDLICHEN

- (i) Es sei $D \subseteq \mathbb{R}$ nach oben unbeschränkt und $f : D \rightarrow \mathbb{K}_2$ eine Funktion. Wir sagen f besitzt bei ∞ den *Grenzwert* $y^* \in \mathbb{K}_2$, falls für jede Folge $(x_n)_{n \in \mathbb{N}} \subseteq D$ mit „ $\lim_{n \rightarrow \infty} x_n = \infty$ “ gilt, dass $\lim_{n \rightarrow \infty} f(x_n) = y^*$. Wir schreiben in diesem Fall

$$\lim_{x \rightarrow \infty} f(x) = y^*.$$

- (ii) Es sei $D \subseteq \mathbb{R}$ nach unten unbeschränkt und $f : D \rightarrow \mathbb{K}_2$ eine Funktion. Wir sagen f besitzt bei $-\infty$ den *Grenzwert* $y_* \in \mathbb{K}_2$, falls für jede Folge $(x_n)_{n \in \mathbb{N}} \subseteq D$ mit „ $\lim_{n \rightarrow \infty} x_n = -\infty$ “ gilt, dass $\lim_{n \rightarrow \infty} f(x_n) = y_*$. Wir schreiben in diesem Fall

$$\lim_{x \rightarrow -\infty} f(x) = y_*.$$

Beispiel 6.1.20

Es gilt

$$\lim_{x \rightarrow \infty} \frac{1}{x} = 0.$$

Beweis. Es sei $(x_n)_{n \in \mathbb{N}}$ eine bestimmt gegen ∞ divergente Folge, d.h. „ $\lim_{n \rightarrow \infty} x_n = \infty$ “. Das bedeutet nach Definition 2.3.35

$$\forall L > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad x_n \geq L. \quad (268)$$

$$\mathbb{Z} : \quad \frac{1}{x_n} \xrightarrow{n \rightarrow \infty} 0.$$

Dazu sei $\varepsilon > 0$. Betrachte $L := \frac{1}{\varepsilon} (*)$. Dann gibt es nach (268) ein $n_0 \in \mathbb{N}$ mit

$$x_n \geq L \quad \forall n \geq n_0. \quad (269)$$

Für $n \geq n_0$ gilt dann auch

$$\left| \frac{1}{x_n} - 0 \right| = \frac{1}{x_n} \stackrel{(269)}{\leq} \frac{1}{L} \stackrel{(*)}{=} \frac{1}{\frac{1}{\varepsilon}} = \varepsilon.$$

Wir haben gesehen

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \left| \frac{1}{x_n} - 0 \right| \leq \varepsilon.$$

Mit anderen Worten $\frac{1}{x_n} \xrightarrow{n \rightarrow \infty} 0$, was zu zeigen war.

Beispiel 6.1.21

Das obige Beispiel kann man auch wie folgt erweitern: Es sei $f : \mathbb{R} \rightarrow \mathbb{C}$ eine Funktion so, dass $\lim_{y \rightarrow 0} f(y)$ existiert. Dann gilt

$$\lim_{x \rightarrow \infty} f\left(\frac{1}{x}\right) = \lim_{y \rightarrow 0} f(y).$$

Beweis. Ist $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ eine Folge mit „ $\lim_{n \rightarrow \infty} x_n = \infty$ “, so gilt nach Beispiel 6.1.20, dass $\lim_{n \rightarrow \infty} \frac{1}{x_n} = 0$. Mit $y_n := \frac{1}{x_n}$ für alle $n \in \mathbb{N}$ und $\lim_{n \rightarrow \infty} y_n = 0$ erhalten wir

$$\lim_{n \rightarrow \infty} f\left(\frac{1}{x_n}\right) = \lim_{n \rightarrow \infty} f(y_n) \stackrel{\text{Def 6.1.3}}{=} \lim_{y \rightarrow 0} f(y).$$

Da $(x_n)_{n \in \mathbb{N}}$ eine beliebige, bestimmt gegen unendlich divergente Folge war folgt

$$\lim_{x \rightarrow \infty} f\left(\frac{1}{x}\right) = \lim_{y \rightarrow 0} f(y).$$

Bestimmte Divergenz von Funktionen Auch kann man von *bestimmter Divergenz von Funktionen* sprechen.

Definition 6.1.22 — BESTIMMTE DIVERGENZ VON FUNKTIONEN

- (i) Es sei $D \subseteq \mathbb{K}$ und $x_0 \in \mathbb{K}$ ein Häufungspunkt von D . Man sagt eine Funktion $f : D \rightarrow \mathbb{R}$ divergiert bei x_0 bestimmt gegen $\pm\infty$, falls für jede Folge $(x_n)_{n \in \mathbb{N}} \subseteq D$

mit $\lim_{n \rightarrow \infty} x_n = x_0$ gilt, dass „ $\lim_{n \rightarrow \infty} f(x_n) = \pm\infty$ “. Man schreibt dann

$$\text{„}\lim_{x \rightarrow x_0} f(x) = \pm\infty.$$

- (ii) Es sei $D \subseteq \mathbb{K}$ nach oben (bzw. unten) unbeschränkt. Man sagt eine Funktion $f : D \rightarrow \mathbb{R}$ divergiert bei ∞ (bzw. $-\infty$) bestimmt gegen $\pm\infty$ falls für jede Folge $(x_n)_{n \in \mathbb{N}} \subseteq D$ mit „ $\lim_{n \rightarrow \infty} x_n = \infty$ “ (bzw. „ $\lim_{n \rightarrow \infty} x_n = -\infty$ “) gilt, dass „ $\lim_{n \rightarrow \infty} f(x_n) = \pm\infty$ “. Man schreibt dann

$$\text{„}\lim_{x \rightarrow \infty} f(x) = \pm\infty.$$

$$(\text{bzw. } \text{„}\lim_{x \rightarrow -\infty} f(x) = \pm\infty).$$

Wie schon bei Folgen lässt man die Anführungszeichen auch manchmal weg.

Das Epsilon-Delta-Kriterium

Im Folgenden lernen wir eine geometrischere Sichtweise auf Funktionsgrenzwerte von $(\mathbb{K}_1, \mathbb{K}_2)$ -Funktionen $f : D \rightarrow \mathbb{K}$ kennen.

Für uns ist irgendwo intuitiv, was es bedeuten soll, dass eine Funktion f bei $x_0 \in \mathbb{K}_1$ den Grenzwert y_0 besitzt. Man denkt nämlich an Folgendes:

$f(x)$ liegt nah an y_0 für x hinreichend nah an x_0 .

Diese Intuition können wir mathematisch mithilfe des sogenannten *Epsilon-Delta-Kriteriums* in Worte fassen.

Satz 6.1.23 — EPSILON-DELTA-KRITERIUM

Es sei $D \subseteq \mathbb{K}_1$ und $f : D \rightarrow \mathbb{K}_2$ eine $(\mathbb{K}_1, \mathbb{K}_2)$ -Funktion. Ferner sei $x_0 \in \mathbb{K}_1$ ein Häufungspunkt von D . Dann gilt für $y_0 \in \mathbb{K}_2$

$$\lim_{x \rightarrow x_0} f(x) = y_0$$

$$\iff$$

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x \in D \text{ mit } |x - x_0| < \delta \text{ gilt } |f(x) - y_0| < \varepsilon. \quad (270)$$

Beweis. Wir starten mit „ $\Leftarrow$ “. Es gelte also das Epsilon-Delta-Kriterium in (547).

$$\mathbb{Z}: \quad \lim_{x \rightarrow x_0} f(x) = y_0.$$

Dazu sei $(x_n)_{n \in \mathbb{N}} \subseteq D$ eine beliebige Folge mit $\lim_{n \rightarrow \infty} x_n = x_0$. Zu zeigen ist, dass $\lim_{n \rightarrow \infty} f(x_n) = y_0$. Dazu sei $\varepsilon > 0$.

$$\mathbb{Z}: \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |f(x_n) - y_0| \leq \varepsilon.$$

Dazu wähle zunächst $\delta > 0$ wie in (547), d.h. so, dass

$$\forall x \in D \text{ mit } |x - x_0| < \delta \text{ gilt } |f(x) - y_0| < \varepsilon. \quad (271)$$

Da $(x_n)_{n \in \mathbb{N}}$ gegen x_0 konvergiert gibt es nun ein $n_1 \in \mathbb{N}$ mit

$$\forall n \geq n_1 : \quad |x_n - x_0| \leq \frac{\delta}{2}.$$

Für $n \geq n_1$ gilt dann $x_n \in D$ und $|x_n - x_0| \leq \frac{\delta}{2} < \delta$. Aus (271) folgt dann

$$\forall n \geq n_1 : \quad |f(x_n) - y_0| < \varepsilon.$$

Wählt man $n_0 := n_1$ so folgt die Behauptung.

Nun noch zu „ $\Rightarrow$ “. Gelte $\lim_{x \rightarrow x_0} f(x) = y_0$ (*). Angenommen das Epsilon-Delta-Kriterium aus (547) gilt nicht, d.h.

$$\neg(\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x \in D \text{ mit } |x - x_0| < \delta \text{ gilt } |f(x) - y_0| < \varepsilon). \quad (272)$$

Unsere Verneinungsregeln für Quantoren (vgl. Überlegung 2.2.4) implizieren dann

$$\exists \varepsilon_0 > 0 : \quad \forall \delta > 0 \quad \neg(\forall x \in D \text{ mit } |x - x_0| < \delta \text{ gilt } |f(x) - y_0| < \varepsilon_0).$$

Wir verneinen weiter:

$$\exists \varepsilon_0 > 0 : \quad \forall \delta > 0 \quad \exists x_\delta \in D \text{ mit } |x_\delta - x_0| < \delta \text{ und } |f(x_\delta) - y_0| \geq \varepsilon_0. \quad (273)$$

Beachte: Mit dem Ausdruck x_δ stellen wir explizit heraus, dass das gewählte Element von δ abhängen darf. Wir betrachten nun für $n \in \mathbb{N}$ den Spezialfall $\delta = \frac{1}{n}$ und definieren $\tilde{x}_n := x_{\frac{1}{n}} \in D$. Beachte: Nach der vorigen Gleichung gilt dann

$$|\tilde{x}_n - x_0| = |x_{\frac{1}{n}} - x_0| < \frac{1}{n} \quad \forall n \in \mathbb{N}.$$

Wir fassen $(\tilde{x}_n)_{n \in \mathbb{N}}$ in eine Folge zusammen und stellen folgende Behauptungen auf

- $(\tilde{x}_n)_{n \in \mathbb{N}} \subseteq D$
- $\lim_{n \rightarrow \infty} \tilde{x}_n = x_0$

Ersteres ist klar, da nach Definition $\tilde{x}_n = x_{\frac{1}{n}} \in D$ für alle $n \in \mathbb{N}$. Zu letzterem: Da

$$0 \leq |\tilde{x}_n - x_0| \leq \frac{1}{n} \quad \forall n \in \mathbb{N}$$

folgt nach dem Sandwichlemma, dass $\lim_{n \rightarrow \infty} |\tilde{x}_n - x_0| = 0$. Nach Proposition 2.3.30 gilt dann auch $\lim_{n \rightarrow \infty} (\tilde{x}_n - x_0) = 0$. Addieren wir x_0 auf beiden Seiten und nutzen die Grenzwertregeln,

so folgt auch $\lim_{n \rightarrow \infty} \tilde{x}_n = x_0$. Wegen den obigen Bullet Points ist $(\tilde{x}_n)_{n \in \mathbb{N}}$ nun eine zulässige Folge in Definition 6.1.3. Wir folgern aus (*), dass

$$\lim_{n \rightarrow \infty} f(\tilde{x}_n) = y_0.$$

Insbesondere gilt auch $\lim_{n \rightarrow \infty} f(\tilde{x}_n) - y_0 = 0$. Wiederum nach Proposition 2.3.30 gilt auch

$$\lim_{n \rightarrow \infty} |f(\tilde{x}_n) - y_0| = 0. \quad (274)$$

Jedoch gilt wegen (273), dass

$$|f(\tilde{x}_n) - y_0| = |f(x_{\frac{1}{n}}) - y_0| \geq \varepsilon_0 \quad \forall n \in \mathbb{N}. \quad (275)$$

Ein Widerspruch, denn

$$0 \underset{(274)}{=} \lim_{n \rightarrow \infty} |f(\tilde{x}_n) - y_0| \underset{\substack{(275) \\ \text{GWR}}}{\geq} \varepsilon_0.$$

Somit muss unsere Annahme (272) falsch sein und daraus folgt sofort die Behauptung.

Bemerkung 6.1.24

Grenzwerte im Unendlichen können mit dem Epsilon-Delta Kriterium nicht beschrieben werden, es muss $x_0 \in \mathbb{K}_1$ vorausgesetzt werden.

Wir werden hier mal ein Beispiel diskutieren, wie Funktionsgrenzwerte mithilfe des Epsilon-Delta-Kriteriums nachgewiesen werden können.

Beispiel 6.1.25

Für alle $x_0 \in (0, \infty)$ gilt $\lim_{\substack{x \rightarrow x_0 \\ x \in (0, \infty)}} \frac{1}{x} = \frac{1}{x_0}$.

BEWEISMÖGLICHKEIT 1. Mit Proposition 6.1.8

Dazu sei $x_0 \in (0, \infty)$ beliebig vorgegeben. Insbesondere $x_0 = \lim_{x \rightarrow x_0} x \neq 0$. Mit Proposition 6.1.8 (iv) folgert man

$$\lim_{x \rightarrow x_0} \frac{1}{x} = \frac{\lim_{x \rightarrow x_0} 1}{\lim_{x \rightarrow x_0} x} = \frac{1}{x_0}$$

BEWEISMÖGLICHKEIT 2. Mit dem Epsilon-Delta-Kriterium.

Sei $x_0 \in (0, \infty)$ beliebig vorgegeben. Ferner sei $\varepsilon > 0$ beliebig vorgegeben. ^a

Drücke nun $\left|\frac{1}{x} - \frac{1}{x_0}\right|$ erstmal in Termen von $|x - x_0|$ aus.

$$\left|\frac{1}{x} - \frac{1}{x_0}\right| = \left|\frac{x_0 - x}{xx_0}\right| = \frac{|x - x_0|}{xx_0} = \frac{|x - x_0|}{[x_0 + (x - x_0)]x_0}.$$

Falls nun $|x - x_0| < x_0$ so hat man auch

$$\left|\frac{1}{x} - \frac{1}{x_0}\right| \leq \frac{|x - x_0|}{(x_0 - |x - x_0|)x_0}.$$

Ist nun $|x - x_0| < \delta$ (mit $\delta \in (0, x_0)$) so gilt

$$\left|\frac{1}{x} - \frac{1}{x_0}\right| \leq \frac{|x - x_0|}{(x_0 - |x - x_0|)x_0} \leq \frac{\delta}{(x_0 - |x - x_0|)x_0} \leq \frac{\delta}{(x_0 - \delta)x_0}.$$

Das soll kleiner als ε werden.

$$\begin{aligned} \frac{\delta}{(x_0 - \delta)x_0} &\stackrel{!}{<} \varepsilon \\ \Leftrightarrow \delta &\stackrel{!}{<} \varepsilon x_0 (x_0 - \delta) \\ \Leftrightarrow \delta &\stackrel{!}{<} \varepsilon x_0^2 - \delta \varepsilon x_0 \\ \Leftrightarrow \delta(1 + \varepsilon x_0) &\stackrel{!}{<} \varepsilon x_0^2 \\ \Leftrightarrow \delta &\stackrel{!}{<} \frac{\varepsilon x_0^2}{1 + \varepsilon x_0}. \end{aligned}$$



Wähle nun $\delta := \frac{1}{2} \frac{\varepsilon x_0^2}{1 + \varepsilon x_0} > 0$. Beachte: Dann gilt insbesondere $\delta < \frac{\varepsilon x_0^2}{0 + \varepsilon x_0} = x_0$. Für $x \in (0, \infty)$ mit $|x - x_0| < \delta$ gilt dann

$$\begin{aligned} \left|\frac{1}{x} - \frac{1}{x_0}\right| &= \frac{|x_0 - x|}{xx_0} = \frac{|x - x_0|}{xx_0} = \frac{|x - x_0|}{[x_0 + (x - x_0)]x_0} \leq \frac{|x - x_0|}{[x_0 - |x - x_0|]x_0} \leq \\ &= \frac{\delta}{[x_0 - |x - x_0|]x_0} \leq \frac{\delta}{[x_0 - \delta]x_0} \\ &= \frac{\frac{1}{2} \frac{\varepsilon x_0^2}{1 + \varepsilon x_0}}{[x_0 - \frac{1}{2} \frac{\varepsilon x_0^2}{1 + \varepsilon x_0}]x_0} = \frac{\varepsilon x_0^2}{[2x_0(1 + \varepsilon x_0) - \varepsilon x_0^2]x_0} = \frac{\varepsilon x_0}{2x_0(1 + \varepsilon x_0) - \varepsilon x_0^2} \\ &= \frac{\varepsilon x_0}{2x_0 + 2\varepsilon x_0^2 - \varepsilon x_0^2} = \frac{\varepsilon x_0}{2x_0 + \varepsilon x_0^2} < \frac{\varepsilon x_0}{2x_0} = \frac{\varepsilon}{2} < \varepsilon. \end{aligned}$$

FAZIT.

Man stellt fest, dass es oft nicht ganz einfach ist, das Epsilon-Delta-Kriterium explizit zu überprüfen. Für Beweise wird es aber trotzdem extrem nützlich sein.

^aAn dieser Stelle danke ich meinem Tutor Luis für die TikZ-Überarbeitung der allseits beliebten *Gedankenwolke*

15.01.2026

Nun wollen wir das Epsilon-Delta-Kriterium noch geometrisch interpretieren. Hierfür definieren wir zunächst, was eine *Umgebung* ist.

Definition 6.1.26 — r -UMGEBUNG

Es sei $x_0 \in \mathbb{K}$ und $r > 0$. Dann nennen wir

$$B_r(x_0) := \{x \in \mathbb{K} : |x - x_0| < r\}$$

eine r -Umgebung (oder auch r -Ball) mit *Mittelpunkt* x_0 .

Das Epsilon-Delta-Kriterium für Funktionsgrenzwerte wird auch *Umgebungskriterium* genannt. Den Grund dafür verstehen wir wenn wir die nachfolgende leicht abgewandelte Version des Kriteriums betrachten:

Proposition 6.1.27 — UMGEBUNGSKRITERIUM

Es sei $D \subseteq \mathbb{K}_1$ und $f : D \rightarrow \mathbb{K}_2$ eine $(\mathbb{K}_1, \mathbb{K}_2)$ -Funktion. Ferner sei $x_0 \in \mathbb{K}_1$ ein Häufungspunkt von D . Dann gilt für $y_0 \in \mathbb{K}_2$

$$\lim_{x \rightarrow x_0} f(x) = y_0$$

$$\Longleftrightarrow$$

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad f(D \cap B_\delta(x_0)) \subseteq B_\varepsilon(y_0).$$

Beweis. Nach dem Satz 6.1.23 gilt

$$\lim_{x \rightarrow x_0} f(x) = y_0$$

$$\Longleftrightarrow$$

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x \in D \text{ mit } |x - x_0| < \delta \text{ gilt } |f(x) - y_0| < \varepsilon. \quad (276)$$

Beachte: Nach Definition 6.1.26 gilt

$$|x - x_0| < \delta \quad \Longleftrightarrow \quad x \in B_\delta(x_0)$$

und

$$|f(x) - y_0| < \varepsilon \iff f(x) \in B_\varepsilon(x_0).$$

Die Aussage in (276) lässt sich also äquivalent umschreiben in

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x \in D \text{ mit } x \in B_\delta(x_0) \text{ gilt } f(x) \in B_\varepsilon(y_0).$$

Mit anderen Worten:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \{f(x) \mid x \in D \text{ und } x \in B_\delta(x_0)\} \subseteq B_\varepsilon(y_0).$$

Die Behauptung folgt daraus sofort, vgl. Definition 1.3.6.

Eine Folgerung, die man sicherlich auch anders hätte herleiten können, ist, dass für die Existenz eines Funktionsgrenzwertes *nur das lokale Verhalten* der Funktion um x_0 eine Rolle spielt. Man spricht von einer *lokalen Eigenschaft*. Genauer:

Korollar 6.1.28

Es sei $D \subseteq \mathbb{K}_1$ und $f, g : D \rightarrow \mathbb{K}_2$, sowie $x_0 \in \mathbb{K}_1$ ein Häufungspunkt von D . Ferner existiere $\lim_{x \rightarrow x_0} f(x)$ und f und g stimmen in einer Umgebung von x_0 überein, d.h. es gibt ein $r > 0$ so, dass

$$\forall x \in D \cap B_r(x_0) : \quad g(x) = f(x). \quad (277)$$

Dann existiert auch $\lim_{x \rightarrow x_0} g(x)$ und

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

Fazit: Für den Grenzwert $\lim_{x \rightarrow x_0} f(x)$ spielt nur das Verhalten von f in einer kleinen Umgebung von x_0 eine Rolle.

Beweis. Definiere

$$y_0 := \lim_{x \rightarrow x_0} f(x)$$

Dann gilt nach Proposition 6.1.27

$$\forall \tilde{\varepsilon} > 0 \quad \exists \tilde{\delta} > 0 : \quad f(D \cap B_{\tilde{\delta}}(x_0)) \subseteq B_{\tilde{\varepsilon}}(y_0) \quad (278)$$

$$\mathbb{Z} : \quad \lim_{x \rightarrow x_0} g(x) = y_0.$$

Dazu verwenden wir auch Proposition 6.1.27 (diesmal die Rückrichtung). Sei $\varepsilon > 0$ beliebig vorgegeben.

$$\mathbb{Z} : \quad \exists \delta > 0 : \quad g(D \cap B_\delta(x_0)) \subseteq B_\varepsilon(y_0).$$

Beobachte zunächst mit (278) (mit $\tilde{\varepsilon} = \varepsilon$), dass es $\tilde{\delta} > 0$ gibt mit

$$f(D \cap B_{\tilde{\delta}}(x_0)) \subseteq B_\varepsilon(y_0). \quad (279)$$

Wähle dann $\delta := \min\{\tilde{\delta}, r\}$, wobei r wie in (277) gewählt ist. Da $D \cap B_\delta(x_0) \subseteq D \cap B_r(x_0)$ stimmen f und g auf $D \cap B_\delta(x_0)$ überein. Das heißt auch

$$g(B_\delta(x_0) \cap D) = f(B_\delta(x_0) \cap D). \quad (280)$$

Da nun $\delta \leq \tilde{\delta}$ gilt

$$f(B_\delta(x_0) \cap D) \subseteq f(B_{\tilde{\delta}}(x_0) \cap D) \underset{(279)}{\subseteq} B_\varepsilon(y_0)$$

Man erhält mit (280)

$$g(B_\delta(x_0) \cap D) \subseteq B_\varepsilon(y_0).$$

QED

6.2 Stetigkeit

Jetzt, wo wir über Funktionsgrenzwerte gesprochen haben, können wir auch das Konzept der *Stetigkeit* einführen.

Definition 6.2.1 — STETIGKEIT

Es sei $D \subseteq \mathbb{K}_1$ und $f : D \rightarrow \mathbb{K}_2$ eine $(\mathbb{K}_1, \mathbb{K}_2)$ -Funktion. Ferner sei $x_0 \in D$. Wir nennen f *stetig bei* $x_0 \in D$ falls

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

Wir nennen f außerdem stetig auf D falls f an jedem Punkt $x_0 \in D$ stetig ist.

Bemerkung 6.2.2

Entpacken wir die Definition der Stetigkeit mithilfe von Definition 6.1.3 so erhalten wir:

$$f : D \rightarrow \mathbb{K}_2 \text{ ist stetig bei } x_0 \in D$$

$$\iff$$

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

$$\iff$$

$$\text{Für jede Folge } (x_n)_{n \in \mathbb{N}} \subseteq D \text{ mit } \lim_{n \rightarrow \infty} x_n = x_0 \text{ gilt, dass } \lim_{n \rightarrow \infty} f(x_n) = f(x_0).$$

Salopp gesagt: Die Funktion f überführt Folgen, die gegen x_0 konvergieren, in Folgen, die gegen $f(x_0)$ konvergieren.

Noch plakativer könnte man das auch so formulieren: Ist $(x_n)_{n \in \mathbb{N}}$ eine konvergente Folge und f bei $\lim_{n \rightarrow \infty} x_n$ stetig, so gilt

$$\lim_{n \rightarrow \infty} f(x_n) = f\left(\lim_{n \rightarrow \infty} x_n\right). \quad (281)$$

Ein *Vertauschungsergebnis*!

Im Folgenden denken wir bei D stets an eine Teilmenge von $\mathbb{K}_1$, sofern nichts anderes gesagt ist.

Bemerkung 6.2.3

Es sei $f : D \rightarrow \mathbb{K}_2$ wie in Definition 6.2.1 und $x_0 \in D$. Eigentlich hätte man bei der Definition der Stetigkeit in x_0 auch nur fordern können, dass

$$\lim_{x \rightarrow x_0} f(x) \quad \text{existiert.}$$

In der Tat: Nimm an, dass $\lim_{x \rightarrow x_0} f(x)$. Betrachten wir die konstante Folge $(x_n)_{n \in \mathbb{N}} = (x_0)_{n \in \mathbb{N}} \subseteq D$ in Definition 6.1.3, so ergibt sich automatisch, dass $\lim_{x \rightarrow x_0} f(x) = f(x_0)$.^a

^aAchtung: Dies ist nur möglich weil $x_0 \in D$. Würde x_0 nicht in D liegen, so wäre die Betrachtung der obigen Folge $(x_n)_{n \in \mathbb{N}}$ nicht zulässig. Daher haben wir diesen Fakt im vorigen Kapitel (wo allgemeine Häufungspunkte von D zugelassen sind) noch nicht diskutiert.

Bemerkung 6.2.4

Auch könnten wir das Epsilon-Delta-Kriterium (Satz 6.1.23) bemühen um eine äquivalente Formulierung des Stetigkeitsbegriffs zu erhalten.^a

$$f : D \rightarrow \mathbb{K}_2 \text{ ist stetig bei } x_0 \in D$$

$$\iff$$

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x \in D \text{ mit } |x - x_0| < \delta \text{ gilt } |f(x) - f(x_0)| < \varepsilon.$$

^aDiese Formulierung ist sogar sehr anschaulich, vgl. GEOGEBRA

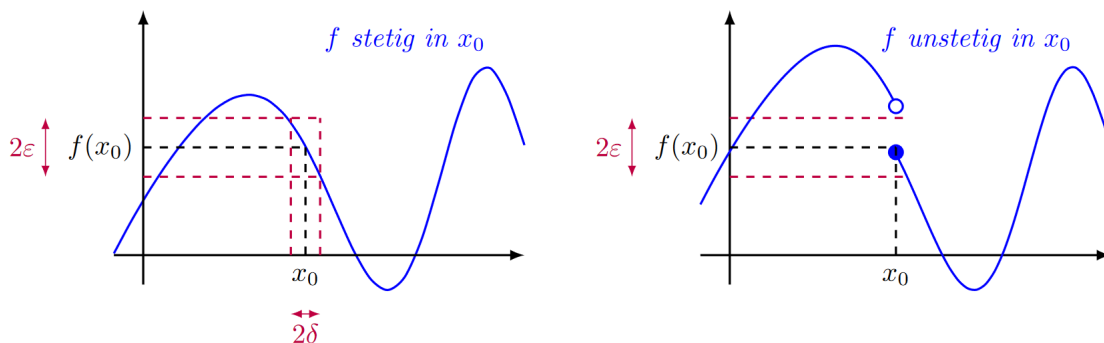


Abbildung 24: Stetige und unstetige Funktionen anschaulich.

Als nächstes sehen wir erste Beispiele stetiger Funktionen

Beispiel 6.2.5

Konstante Funktionen, identische Abbildungen und der Absolutbetrag sind auf ihrem gesamten Definitionsbereich stetig.

Beweis. (Nur für den Absolutbetrag auf $\mathbb{R}$). Wir erinnern uns, dass $\text{abs} : \mathbb{R} \rightarrow \mathbb{R}$ gegeben ist durch $\text{abs}(x) = |x|$. Sei nun $x_0 \in \mathbb{R}$ beliebig.

$$\text{Zu : } \quad \lim_{x \rightarrow x_0} \text{abs}(x) = \text{abs}(x_0) \quad \text{d.h.} \quad \lim_{x \rightarrow x_0} |x| = |x_0|.$$

Dazu: Sei $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ mit $\lim_{n \rightarrow \infty} x_n = x_0$

$$\text{Proposition 2.3.32} \quad \implies \quad \lim_{n \rightarrow \infty} |x_n| = |x_0|$$

Da $(x_n)_{n \in \mathbb{N}}$ beliebig war, folgt nach Definition 6.1.3, dass $\lim_{x \rightarrow x_0} |x| = |x_0|$. QED

Es gibt eine Vielzahl von Möglichkeiten, aus gegebenen stetigen Funktionen, neue stetige Funktionen zu basteln. Jegliche *Kompositionsoperationen* (vgl. Definition 6.0.3) erhalten die Stetigkeit!

Proposition 6.2.6 — KOMPOSITIONEN STETIGER FUNKTIONEN SIND STETIG

Es sei $D \subseteq \mathbb{K}_1$ und $f, g : D \rightarrow \mathbb{K}_2$. Ferner sei $x_0 \in D$ so, dass f und g bei x_0 stetig sind.

(I) (*Rationale Operationen*) Dann sind auch

$f + g, fg$ und (sofern $g(x_0) \neq 0$) auch $\frac{f}{g}$ bei x_0 stetig.

(II) (*Verkettungen*) Ist ferner $h : \tilde{D} \rightarrow \mathbb{K}$ so, dass $\tilde{D} \supseteq f(D)$ und h bei $f(x_0)$ stetig ist, so ist auch

$(h \circ f) : D \rightarrow \mathbb{K}$ bei x_0 stetig.

Beweis.

ZU (I). Rationale Operationen.

Zeigen wir zunächst, dass $f + g$ bei x_0 stetig ist. Dazu berechnen wir

$$\lim_{x \rightarrow x_0} (f + g)(x) = \lim_{x \rightarrow x_0} \underbrace{f(x)}_{\xrightarrow{x \rightarrow x_0} f(x_0)} + \underbrace{g(x)}_{\xrightarrow{x \rightarrow x_0} g(x_0)} \stackrel{\text{Prop 6.1.8(i)}}{=} f(x_0) + g(x_0) \stackrel{\text{Def 6.0.3}}{=} (f + g)(x_0).$$

Die Behauptung folgt dann direkt aus Definition 6.2.1. Die Stetigkeitsaussagen für Produkte

und Quotienten funktionieren ganz analog und sind den Lesenden als Übungsaufgabe überlassen.

ZU (II). Verkettungen.

Nun zur Stetigkeit der Verkettung $(h \circ f)$ bei x_0 .

$$\mathbb{Z}: \quad \lim_{x \rightarrow x_0} (h \circ f)(x) = (h \circ f)(x_0).$$

Es sei hierzu $(x_n)_{n \in \mathbb{N}} \subseteq D$ mit $\lim_{n \rightarrow \infty} x_n = x_0$. Da f bei x_0 stetig ist gilt

$$\lim_{n \rightarrow \infty} f(x_n) = f(x_0).$$

Da $f(D) \subseteq \tilde{D}$ gilt, ist $(f(x_n))_{n \in \mathbb{N}}$ eine Folge in $\tilde{D}$ mit $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$. Aus der Stetigkeit von $h : \tilde{D} \rightarrow \mathbb{K}$ bei $f(x_0)$ folgt dann

$$\lim_{n \rightarrow \infty} h(f(x_n)) = h(f(x_0))$$

Wir folgern

$$\lim_{n \rightarrow \infty} (h \circ f)(x_n) = \lim_{n \rightarrow \infty} h(f(x_n)) = h(f(x_0)) = (h \circ f)(x_0).$$

Viele Funktionen sind durch *Potenzreihen* darstellbar, so z.B. alle Polynome, die Exponentialfunktion, Kosinus und Sinus etc... Die nächste Proposition zeigt, dass sämtliche Potenzreihen innerhalb ihrer *Konvergenzscheibe* stetig sind.

Proposition 6.2.7 — STETIGKEIT VON POTENZREIHEN

Es sei eine Potenzreihe $\sum_{k \geq 0} c_k(z - w)^k$ gegeben mit Referenzpunkt $w \in \mathbb{C}$ und Konvergenzradius $R > 0$. Dann ist die Wertefunktion $P : B_R(w) \rightarrow \mathbb{C}, z \mapsto \sum_{k=0}^{\infty} c_k(z - w)^k$ auf ganz $B_R(w)$ stetig. ^a

^aIn Definition 5.2.1 war P eigentlich auf dem ganzen *Konvergenzbereich* $D_K \supseteq B_R(w)$ definiert. Wir werden allerdings die Stetigkeit von P nur auf der eventuell kleineren Menge $B_R(w) := \{z \in \mathbb{C} : |z - w| < R\}$ nachweisen. Ist hingegen $R = \infty$, so kann $B_R(w)$ durch $\mathbb{C}$ ersetzt werden.

Beweis. Es sei $z_0 \in B_R(w)$, also gilt $|z_0 - w| < R$. Definiere $r := |z_0 - w|$. Zu zeigen ist

$$\lim_{z \rightarrow z_0} P(z) = P(z_0), \quad \text{d.h.} \quad \lim_{z \rightarrow z_0} \sum_{k=0}^{\infty} c_k(z - w)^k = \sum_{k=0}^{\infty} c_k(z_0 - w)^k.$$

Hierzu sei $(z_n)_{n \in \mathbb{N}} \subseteq B_R(w)$ eine Folge mit $\lim_{n \rightarrow \infty} z_n = z_0$.

$$\mathbb{Z}_L: \quad \lim_{n \rightarrow \infty} P(z_n) = P(z_0).$$

Nun beachte: Für alle $n \in \mathbb{N}$ mit $z_n \neq z_0$ gilt

$$\begin{aligned} P(z_n) - P(z_0) &= \sum_{k=0}^{\infty} c_k (z_n - w)^k - \sum_{k=0}^{\infty} c_k (z_0 - w)^k = \sum_{k=0}^{\infty} [c_k (z_n - w)^k - c_k (z_0 - w)^k] \\ &= \sum_{k=0}^{\infty} c_k [(z_n - w)^k - (z_0 - w)^k] \\ &\stackrel{\substack{\text{0-ter Summand} \\ \text{verschwindet}}}{=} \sum_{k=1}^{\infty} c_k [(z_n - w)^k - (z_0 - w)^k] \\ &= \underbrace{((z_n - w) - (z_0 - w))}_{=(z_n - z_0)} \sum_{k=1}^{\infty} c_k \left[\frac{(z_n - w)^k - (z_0 - w)^k}{(z_n - w) - (z_0 - w)} \right] \\ &\stackrel{\text{Kor 3.3.21(i)}}{=} (z_n - z_0) \underbrace{\sum_{k=1}^{\infty} c_k \left(\sum_{\ell=0}^{k-1} (z_n - w)^\ell (z_0 - w)^{k-1-\ell} \right)}_{=: a_n} = (z_n - z_0) a_n. \end{aligned} \tag{282}$$

Selbige Gleichung gilt übrigens auch für den Fall, dass $z_n = z_0$, denn in diesem Fall sind beide Seiten gleich 0. Beachte: $(z_n - z_0)_{n \in \mathbb{N}}$ ist eine Nullfolge. Um Konvergenz von $(P(z_n) - P(z_0))_{n \in \mathbb{N}}$ nachzuweisen, genügt es nach Proposition 2.3.33 zu zeigen, dass $(a_n)_{n \in \mathbb{N}}$ beschränkt ist.

ZWISCHENBEHAUPTUNG: $(a_n)_{n \in \mathbb{N}}$ ist beschränkt.

Dazu: Da $z_n \xrightarrow{n \rightarrow \infty} z_0$ gilt auch $|z_n - w| \xrightarrow{n \rightarrow \infty} |z_0 - w| = r$.^(a) Wähle nun $r' \in (r, R)$ beliebig. Wegen der Konvergenz von $|z_n - w|$ gegen r folgt (Übungsaufgabe^{b)})

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |z_n - w| \leq r'. \tag{283}$$

Für $n \geq n_0$ gilt dann

$$\begin{aligned} |a_n| &\stackrel{\Delta\text{-Ungl.}}{\leq} \sum_{k=1}^{\infty} |c_k| \left| \sum_{\ell=0}^{k-1} (z_n - w)^\ell (z_0 - w)^{k-1-\ell} \right| \\ &\stackrel{\Delta\text{-Ungl.}}{\leq} \sum_{k=1}^{\infty} |c_k| \sum_{\ell=0}^{k-1} |z_n - w|^\ell |z_0 - w|^{k-1-\ell} \\ &\leq \sum_{k=1}^{\infty} |c_k| \sum_{\ell=0}^{k-1} (r')^\ell r^{k-1-\ell} \end{aligned}$$

$$\begin{aligned}
&\leq \sum_{k=1}^{\infty} |c_k| \sum_{\ell=0}^{k-1} (r')^{\ell} (r')^{k-1-\ell} \\
&\leq \sum_{k=1}^{\infty} |c_k| \sum_{\ell=0}^{k-1} (r')^{k-1} = \sum_{k=1}^{\infty} |c_k| k (r')^{k-1} \\
&= \frac{1}{r'} \sum_{k=1}^{\infty} |c_k| k (r')^k.
\end{aligned}$$

Der obige Ausdruck hängt nicht mehr von n ab und könnte daher als obere Schranke fungieren. Jedoch fehlt noch etwas: Wir haben noch gar nicht nachgewiesen, dass es sich bei dem Ausdruck überhaupt um eine endliche reelle Zahl handelt. Dafür müssen wir noch nachweisen, dass

$$\sum_{k \geq 1} |c_k| k (r')^k \quad \text{eine konvergente Reihe ist.} \quad (284)$$

Dazu: Die Potenzreihe $\sum_{k \geq 0} |c_k| k z^k (*)$ hat nach Proposition 5.2.6 denselben Konvergenzradius wie $\sum_{k \geq 0} c_k z^k$, also R . Nun gilt $|r'| = r' < R$ und somit liegt $z = r'$ im Konvergenzbereich der betrachteten Reihe (*). Also muss diese Reihe für $z = r'$ konvergieren und (284) folgt. Wir erhalten die Beschränktheit von $(a_n)_{n \geq n_0}$. Sprich: Wir finden $\tilde{K} > 0$ mit

$$\forall n \geq n_0 : \quad |a_n| \leq \tilde{K}$$

Für alle $n \in \mathbb{N}$ gilt dann $|a_n| \leq \max\{|a_1|, \dots, |a_{n_0}|, \tilde{K}\} =: K$.^c Somit ist $(a_n)_{n \in \mathbb{N}}$ beschränkt.

ZWISCHENBEHAUPTUNG ENDE.

Mit (282) gilt dann

$$\lim_{n \rightarrow \infty} (P(z_n) - P(z_0)) \stackrel{(282)}{=} \lim_{n \rightarrow \infty} a_n (z_n - z_0) \stackrel{\text{Prop 2.3.33}}{=} 0.$$

Die Grenzwertregeln implizieren $\lim_{n \rightarrow \infty} P(z_n) = P(z_0)$, was zu zeigen war.

^aWer diesen Schritt genau begründen will kann jetzt zum Beispiel die Stetigkeit des Absolutbetrages (vgl. Bemerkung 6.2.5) verwenden: Es gilt nach Bemerkung 6.2.2

$$\lim_{n \rightarrow \infty} |z_n - w| = \left| \lim_{n \rightarrow \infty} (z_n - w) \right| \stackrel{\text{GWR}}{=} \left| \lim_{n \rightarrow \infty} z_n - w \right| = |z_0 - w| = r$$

^bDiese Übungsaufgabe kann man zum Beispiel so lösen: Würde (283) nicht gelten so wäre $A := \{n \in \mathbb{N} : |z_n - w| > r'\}$ unendlich. Man würde dann (vgl. Bemerkung 3.5.15) eine strikt aufsteigende Folge $(\ell_n)_{n \in \mathbb{N}} \subseteq A$ finden d.h. $|z_{\ell_n} - w| > r'$ für alle $n \in \mathbb{N}$. Da $(|z_{\ell_n} - w|)_{n \in \mathbb{N}}$ aber gegen r konvergieren muss liefern die Grenzwertvergleichsregeln den Widerspruch $r \geq r'$.

^cDas überlegt man sich am besten mal genau mit einer Fallunterscheidung (Fall 1 $n < n_0$ und Fall 2 $n \geq n_0$).

Als Korollar möchten wir nochmal zusammenfassen, dass alle Funktionen, die als Potenzreihe darstellbar sind (wir haben sehr viele Beispiele davon) stetig sind.

Korollar 6.2.8

Alle Polynome und auch $\exp, \cos, \sin$ sind auf $\mathbb{C}$ stetig. Kompositionen aus diesen Funktionen sind auf ihrem jeweiligen Definitionsbereich stetig.

Beweis. Folgt sofort aus den vorigen beiden Propositionen und dem Fakt, dass Polynome, $\exp, \sin, \cos$ als Potenzreihe darstellbar sind (vgl. Definition 5.4.2 und Proposition 5.5.7).

Beispiel 6.2.9

Diese Beobachtung und (281) hilft uns bei so einigen Grenzwertberechnungen. Ein Beispiel:

$$\text{Behauptung.} \quad \lim_{n \rightarrow \infty} \cos\left(\frac{1}{2^n}\right) = 1.$$

Beweis. Nach dem vorigen Korollar ist $\cos : \mathbb{C} \rightarrow \mathbb{C}$ stetig. Nach (281) gilt dann

$$\lim_{n \rightarrow \infty} \cos\left(\frac{1}{2^n}\right) \underset{(281)}{=} \cos\left(\lim_{n \rightarrow \infty} \frac{1}{2^n}\right) = \cos\left(\lim_{n \rightarrow \infty} \left(\frac{1}{2}\right)^n\right) = \cos(0) = 1.$$

Wichtig bei diesen Grenzwertvertauschungen ist, dass der Grenzwert der reingetauscht auch wirklich ein Stetigkeitspunkt der Funktion ist (vgl. Voraussetzung von (281)). Was demnach keinen Sinn macht ist sowas wie die folgende

$$\text{Falsche Rechnung:} \quad \lim_{n \rightarrow \infty} \sin(\pi n^2) = \sin\left(\lim_{n \rightarrow \infty} \pi n^2\right) = \sin(\infty),$$

denn ∞ ist kein Punkt an dem der Sinus stetig ist. Das von uns erfundene Symbol $\sin(\infty)$ ist offenbar auch gar keine wohldefinierte reelle Zahl. Stattdessen würde man richtigerweise wie folgt argumentieren: Da für alle $n \in \mathbb{N}$ die Zahl πn^2 ein Vielfaches von π ist gilt $\forall n \in \mathbb{N} : \sin(\pi n^2) = 0$ (wie man sich leicht mit Proposition 5.5.15 überlegt) Es folgt sofort, dass $\lim_{n \rightarrow \infty} \sin(\pi n^2) = 0$. Auch darf man natürlich keinen Grenzwert reintauschen, der nicht existiert. Beispiel

$$\text{Falsche Rechnung:} \quad \lim_{n \rightarrow \infty} \sin(\pi(-1)^n) = \sin\left(\lim_{n \rightarrow \infty} \pi(-1)^n\right) \quad \text{existiert nicht,}$$

da $\lim_{n \rightarrow \infty} \pi(-1)^n$ nicht existiert.

Hier wurde falsch gemacht, dass der reingetauschte Grenzwert gar nicht existiert (dies wird aber in (281) vorausgesetzt). Richtig würde man auch wieder so argumentieren: für jedes $n \in \mathbb{N}$ ist $\pi(-1)^n$ ein ganzzahliges Vielfaches von π und somit gilt wieder $\forall n \in \mathbb{N} : \sin(\pi(-1)^n) = 0$. Daraus folgt wieder sofort, dass $\lim_{n \rightarrow \infty} \sin(\pi(-1)^n) = 0$.

19.01.2026

Der Satz von Weierstraß

Im Schulunterricht haben Sie sich oft dafür interessiert, ob und wo Funktionen *Extremwerte* (d.h. Maxima/Minima) annehmen. Der *Satz von Weierstraß* liefert ein erstes Resultat zu genau dieser Frage. Allerdings beschränkt er sich nur mit dem „ob“ und nicht mit dem „wo“. Wir definieren nun zunächst, was es bedeutet wenn eine Funktion ihre Extremwerte annimmt.

Definition 6.2.10

Es sei $D \subseteq \mathbb{K}_1$ und $f : D \rightarrow \mathbb{R}$ eine $(\mathbb{K}_1, \mathbb{R})$ -Funktion. Wir sagen

- (i) f nimmt auf D ein Maximum an falls $\max f(D)$ existiert. Ein Punkt $x_0 \in D$ mit

$$f(x_0) = \max f(D)$$

heißt dann *Maximumsstelle* von f . Man notiert $\max f(D)$ gelegentlich auch mit $\max_D f$ oder $\max_{x \in D} f(x)$

- (ii) f nimmt auf D ein Minimum an falls $\min f(D)$ existiert. Ein Punkt $x_0 \in D$ mit

$$f(x_0) = \min f(D)$$

heißt dann *Minimumsstelle* von f . Man notiert $\min f(D)$ gelegentlich auch mit $\min_D f$ oder $\min_{x \in D} f(x)$

Der Satz von Weierstraß besagt konkret, dass stetige Funktionen auf abgeschlossenen beschränkten Intervallen $D = [a, b]$ ein Maximum und Minimum annehmen. Ein solches abgeschlossenes beschränktes Intervall nennt man übrigens auch *kompaktes Intervall*.

Um den Satz von Weierstraß zu beweisen benötigen wir zunächst ein vorbereitendes Lemma über kompakte Intervalle, welches auch später noch von Nutzen sein wird.

Lemma 6.2.11

Es sei für $a, b \in \mathbb{R}, a \leq b$ das abgeschlossene Intervall $[a, b] \subseteq \mathbb{R}$ gegeben. Sei außerdem $(x_n)_{n \in \mathbb{N}} \subseteq [a, b]$ eine Folge. Dann gibt es eine konvergente Teilfolge $(x_{\ell_n})_{n \in \mathbb{N}}$ von $(x_n)_{n \in \mathbb{N}}$ und

$$\lim_{n \rightarrow \infty} x_{\ell_n} \in [a, b].$$

Beweis. Beobachte zunächst, dass

$$(x_n)_{n \in \mathbb{N}} \subseteq [a, b] \implies (x_n)_{n \in \mathbb{N}} \text{ beschränkt.}$$

In der Tat: eine obere Schranke ist durch b gegeben und eine untere Schranke durch a ! Dann liefert der Satz von Bolzano-Weierstraß (Satz 3.5.16) eine konvergente Teilfolge

$(x_{\ell_n})_{n \in \mathbb{N}}$. Setze

$$x_0 := \lim_{n \rightarrow \infty} x_{\ell_n}$$

$$\text{NOCH } \mathbb{Z}: \quad x_0 \in [a, b].$$

Beachte, dass auch $(x_{\ell_n})_{n \in \mathbb{N}} \subseteq [a, b]$. Also

$$x_{\ell_n} \geq a \quad \forall n \in \mathbb{N} \quad \implies \quad x_0 \geq a$$

und

$$x_{\ell_n} \leq b \quad \forall n \in \mathbb{N} \quad \implies \quad x_0 \leq b.$$

Mit anderen Worten: $x_0 \in [a, b]$!

QED

Die obige Beobachtung nennt man auch *Folgenkompaktheit* des Intervalles $[a, b]$. Mehr dazu in Analysis II. Nun sind wir in der Lage, den Satz von Weierstraß zu beweisen.

Satz 6.2.12 — SATZ VON WEIERSTRASS

Es seien $a, b \in \mathbb{R}$ mit $a \leq b$ und $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$. Dann nimmt f auf $[a, b]$ ein Maximum und ein Minimum an.

Beweis. Wir zeigen nur, dass f ein Maximum annimmt. Die Aussage für das Minimum geht dann ganz analog.

SCHRITT 1. Wir zeigen: $f([a, b])$ ist nach oben beschränkt.

Angenommen $f([a, b])$ ist nicht nach oben beschränkt. Dann muss jede natürliche Zahl innerhalb der Menge $f([a, b])$ überboten werden, d.h.:

$$\text{Es gibt für alle } n \in \mathbb{N} \text{ ein } y_n \in f([a, b]) \text{ mit } y_n \geq n. \quad (285)$$

Sei nun $n \in \mathbb{N}$. Da $y_n \in f([a, b])$ gibt es stets ein $x_n \in [a, b]$ mit $y_n = f(x_n)$. Betrachte nun $(x_n)_{n \in \mathbb{N}} \subseteq [a, b]$ als Folge. Nach Lemma 6.2.11 gibt es eine konvergente Teilfolge $(x_{\ell_n})_{n \in \mathbb{N}}$ mit

$$x_0 := \lim_{n \rightarrow \infty} x_{\ell_n} \in [a, b].$$

Beachte: Nach Voraussetzung ist f bei x_0 stetig. Nutzen wir die Folge $(x_{\ell_n})_{n \in \mathbb{N}}$ wie in Bemerkung 6.2.2, so erhalten wir

$$\lim_{n \rightarrow \infty} f(x_{\ell_n}) = f(x_0)$$

Insbesondere ist also $(f(x_{\ell_n}))_{n \in \mathbb{N}}$ eine konvergente (und somit beschränkte) Folge. Jedoch gilt auch

$$f(x_{\ell_n}) = y_{\ell_n} \underset{(285)}{\geq} \ell_n \geq n \quad \forall n \in \mathbb{N},$$

was der Beschränktheit von $(f(x_{\ell_n}))_{n \in \mathbb{N}}$ widerspricht. Unsere Annahme muss also falsch sein und daher ist $f([a, b])$ nach oben beschränkt.

SCHRITT 2. f nimmt auf $[a, b]$ ein Maximum an, d.h. $\exists \max f([a, b])$.

Zunächst folgern wir aus Schritt 1 und der Ordnungsvollständigkeit von $\mathbb{R}$ (vgl. Sektion 3.4), dass $\sup f([a, b])$ existiert. Etwa sei $y^* := \sup f([a, b])$. Nach Bemerkung 3.2.27 (mit $\varepsilon = \frac{1}{n}$) gibt es nun für alle $n \in \mathbb{N}$ ein $y_n \in f([a, b])$ mit

$$y^* - \frac{1}{n} < y_n \leq y^*.$$

Das Sandwichlemma impliziert, dass dann

$$\lim_{n \rightarrow \infty} y_n = y^*.$$

Da $y_n \in f([a, b])$ gilt wiederum, dass $y_n = f(x_n)$ für ein $x_n \in [a, b]$. Betrachte nun wiederum $(x_n)_{n \in \mathbb{N}} \subseteq [a, b]$ als Folge. Wiederum mit Lemma 6.2.11 finden wir eine konvergente Teilfolge $(x_{\ell_n})_{n \in \mathbb{N}}$ mit

$$x_0 := \lim_{n \rightarrow \infty} x_{\ell_n} \in [a, b].$$

Die Stetigkeit von f bei x_0 liefert dann wieder einmal

$$\lim_{n \rightarrow \infty} f(x_{\ell_n}) = f(x_0). \quad (286)$$

Nun erinnern wir uns daran, dass $(x_n)_{n \in \mathbb{N}}$ so gewählt ist, dass $f(x_n) = y_n$ für alle $n \in \mathbb{N}$. Daher gilt

$$f(x_0) \stackrel{(286)}{=} \lim_{n \rightarrow \infty} f(x_{\ell_n}) = \lim_{n \rightarrow \infty} y_{\ell_n} \stackrel{\text{Prop 3.5.13}}{=} \lim_{n \rightarrow \infty} y_n = y^* = \sup f([a, b]).$$

Somit muss $f(x_0)$ eine obere Schranke von $f([a, b])$ sein (vgl. Definition 3.2.10). Da aber auch $f(x_0) \in f([a, b])$ folgt (vgl. Definition 3.2.5)

$$f(x_0) = \max f([a, b]).$$

Bemerkung 6.2.13

Eine wichtige Konsequenz ist, dass eine stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ auch automatisch *beschränkt* ist, d.h.

$$\exists K > 0 : \quad |f(x)| \leq K \quad \forall x \in [a, b]. \quad (287)$$

In der Tat: Da nach Satz 6.2.12 $\max f([a, b])$ und $\min f([a, b])$ existieren, ist $f([a, b])$ nach oben und nach unten beschränkt. Genau wie in Proposition 2.1.6 kann man dann (287) nachweisen.

Bemerkung 6.2.14

Der obige Satz verliert seine Gültigkeit, wenn man anstatt $[a, b]$ das offene Intervall (a, b) oder die halboffenen Intervalle $[a, b)/(a, b]$ betrachtet. In der Tat: Betrachte

$$f : [0, 1) \rightarrow \mathbb{R}, \quad f(x) = x.$$

Dann ist f auf $[0, 1)$ stetig, nimmt aber dort kein Maximum an, denn

$$\max f([0, 1)) \stackrel{f=\text{id}}{=} \max [0, 1)$$

existiert nicht, vgl. Proposition 3.2.25.

Wir möchten am Ende dieses Abschnittes noch eine Anmerkung zu Definition 6.2.10 machen: Das Maximum $\max f(D)$ ist (sofern existent) eindeutig bestimmt, vgl. Proposition 3.2.7. Jedoch kann es durchaus mehrere *Maximumsstellen* geben, d.h. mehrere Punkte $x_0 \in D$ mit

$$f(x_0) = \max f(D).$$

Ein Extrembeispiel: Bei einer konstanten Funktion ist jeder Wert $x_0 \in D$ eine Maximumsstelle. Überzeugen Sie sich davon als Übungsaufgabe!

Der Zwischenwertsatz

Der *Zwischenwertsatz* ist ein wichtiges Werkzeug zum Lösen von Gleichungen, welches wir bereits bei der Konstruktion von π in Kapitel 5.5 ohne Beweis benutzt haben. Hier wollen wir diese Lücke schließen.

Satz 6.2.15 — ZWISCHENWERTSATZ

Es seien $a, b \in \mathbb{R}, a \leq b$ und $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion. Dann nimmt f jeden Wert zwischen $f(a)$ und $f(b)$ auf dem Intervall $[a, b]$ mindestens einmal an. Konkreter

$$\forall y_0 \in [\min\{f(a), f(b)\}, \max\{f(a), f(b)\}] \quad \exists x_0 \in [a, b] : \quad f(x_0) = y_0.$$

Beweis.

$$\text{FALL 1. } f(a) \leq f(b), \quad \text{d.h. } [\min\{f(a), f(b)\}, \max\{f(a), f(b)\}] = [f(a), f(b)].$$

Sei $y_0 \in [f(a), f(b)]$ beliebig. Definiere die Menge

$$A := \{x \in [a, b] : f(x) \leq y_0\}. \tag{288}$$

Dann ist A nichtleer, denn $a \in A$ (weil $f(a) \leq y_0$). Außerdem ist A durch b nach oben beschränkt. Somit existiert

$$x_0 := \sup(A) \quad [= \sup\{x \in [a, b] : f(x) \leq y_0\}].$$

Beachte, dass $x_0 \geq a$ (da $a \in A$) und $x_0 \leq b$ (da b klarerweise eine obere Schranke von A ist). Also gilt $x_0 \in [a, b]$. Wir werden nun zeigen, dass gerade an dieser Stelle x_0 der Wert y_0 angenommen wird.

$$\text{Behauptung:} \quad f(x_0) = y_0.$$

SCHRITT 1. $f(x_0) \leq y_0$.

Da $x_0 = \sup(A)$ gibt es für alle $n \in \mathbb{N}$ ein $x_n \in A$ mit

$$x_0 - \frac{1}{n} < x_n \leq x_0$$

(vgl. Bemerkung 3.2.27 $\varepsilon = \frac{1}{n}$). Insbesondere folgt aus dem Sandwichlemma

$$\lim_{n \rightarrow \infty} x_n = x_0. \quad (289)$$

Ferner gilt wegen (289) und der Stetigkeit von f bei x_0 , dass

$$f(x_0) = \lim_{n \rightarrow \infty} f(x_n). \quad (290)$$

Beachte: Da $x_n \in A$ gilt nach (288) $f(x_n) \leq y_0 \quad \forall n \in \mathbb{N}$. Die Grenzwertvergleichsregeln liefern $\lim_{n \rightarrow \infty} f(x_n) \leq y_0$. Mit (290) folgt

$$f(x_0) \leq y_0.$$

SCHRITT 2. $f(x_0) = y_0$.

Angenommen $f(x_0) \neq y_0$. Dann folgt aus Schritt 1, dass $f(x_0) < y_0$. Insbesondere gilt $x_0 \neq b$, denn wäre $x_0 = b$ so wäre $f(b) < y_0$ (was $y_0 \in [f(a), f(b)]$ widersprechen würde). Ab sofort nehmen wir also an, dass $x_0 < b$. Definiere nun

$$\varepsilon := \frac{y_0 - f(x_0)}{2} [> 0].$$

Nach dem Epsilon-Delta-Kriterium 6.2.4 gibt es dann ein $\delta > 0$ so, dass

$$\forall x \in [a, b] : \quad \text{mit } |x - x_0| < \delta \quad \text{gilt} \quad |f(x) - f(x_0)| < \varepsilon = \frac{y_0 - f(x_0)}{2}. \quad (291)$$

Wähle nun $x' := \min\{x_0 + \frac{\delta}{2}, b\}$. Dann gilt $x' \in [a, b]$ und $|x' - x_0| < \delta$. Ferner gilt $x' > x_0$, da $b > x_0$ und $\delta > 0$. Wenden wir (291) mit $x = x'$ an, so ergibt sich

$$f(x') = f(x') - f(x_0) + f(x_0) \leq |f(x') - f(x_0)| + f(x_0)$$

$$< \frac{y_0 - f(x_0)}{2} + f(x_0) = \frac{y_0 + f(x_0)}{2} < y_0.$$

Man folgert dass $f(x') < y_0$. Dies impliziert dann aber nach der Definition von A in (288), dass $x' \in A$. Jedoch gilt auch nach der Konstruktion von x' dass $x' > x_0 = \sup(A)$. Ein Widerspruch, denn ein Element aus A kann nicht echt größer als eine obere Schranke von A sein. Die Behauptung folgt.

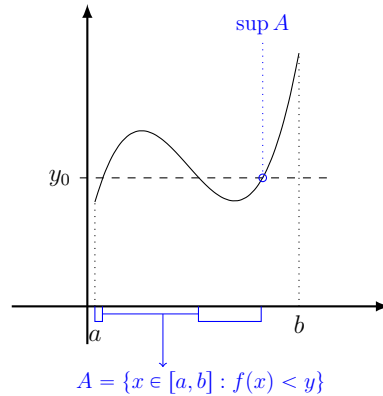


Abbildung 25: Bild zum Zwischenwertsatz

Der Zwischenwertsatz kann oft verwendet werden, um *Surjektivität* von Funktionen zu beweisen. So zum Beispiel in dieser Anwendung:

Proposition 6.2.16

Die Funktion $\exp : [0, \infty) \rightarrow [1, \infty)$ ist surjektiv.

Beweis. Es sei $y \in [1, \infty)$. Zu zeigen ist, dass es ein $x \in [0, \infty)$ gibt mit $\exp(x) = y$. Vorab eine Beobachtung: Für alle $x \in [0, \infty)$ gilt

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \sum_{k=0}^1 \frac{x^k}{k!} + \underbrace{\sum_{k=2}^{\infty} \frac{x^k}{k!}}_{>0} \geq \sum_{k=0}^1 \frac{x^k}{k!} = 1 + x.$$

Insbesondere gilt

$$\exp(y) \geq 1 + y > y. \quad (292)$$

Definiere nun $a := 0$ und $b := y$. Dann gilt

$$[\min\{\exp(a), \exp(b)\}, \max\{\exp(a), \exp(b)\}] = [1, \exp(y)]$$

Aus (292) folgt dann, dass $y \in [\min\{\exp(a), \exp(b)\}, \max\{\exp(a), \exp(b)\}]$. Aus dem Zwischenwertsatz (Satz 6.2.15, der anwendbar ist weil $\exp$ stetig ist, vgl Korollar 6.2.8) folgt

dann

$$\exists x \in [a, b] : \quad \exp(x) = y.$$

Da $[a, b] = [0, y] \subseteq [0, \infty)$ folgt auch

$$\exists x \in [0, \infty) : \quad \exp(x) = y.$$

QED

Die Funktion aus der Aussage der vorigen Proposition war nicht auch einem *kompakten Intervall der Form* $[a, b]$ definiert. Wir mussten sie im Beweis der vorigen Proposition erst auf ein kompaktes Intervall *einschränken*. Ansonsten wäre keine direkte Anwendung des Zwischenwertsatzes erlaubt, vgl. Satz 6.2.15.

Für andere Intervalltypen möchte man allerdings auch gern die Surjektivität diskutieren. Dabei ist die nachfolgende Proposition sehr hilfreich, die auch bei einer klassischen Kurvendiskussion häufig benutzt wird (Diskussion des *Globalverhaltens*)

Proposition 6.2.17 — GLOBALVERHALTEN

Seien $x_0 \in \mathbb{R} \cup \{-\infty\}$ und $x_1 \in \mathbb{R} \cup \{\infty\}$, $x_0 < x_1$ und sei $f : (x_0, x_1) \rightarrow \mathbb{R}$ eine stetige Funktion.

(i) Sind $y_0, y_1 \in \mathbb{R} \cup \{\pm\infty\}$, $y_0 < y_1$ so, dass

$$\lim_{x \rightarrow x_0} f(x) \leq y_0 \quad \text{und} \quad \lim_{x \rightarrow x_1} f(x) \geq y_1. \quad (293)$$

Dann gilt $f((x_0, x_1)) \supseteq (y_0, y_1)$.

(ii) Sind $y_0, y_1 \in \mathbb{R} \cup \{\pm\infty\}$, $y_0 > y_1$ so, dass

$$\lim_{x \rightarrow x_0} f(x) \geq y_0 \quad \text{und} \quad \lim_{x \rightarrow x_1} f(x) \leq y_1. \quad (294)$$

Dann gilt $f((x_0, x_1)) \supseteq (y_1, y_0)$.

Beweis. Wir beweisen nur (i), denn Teil (ii) folgt direkt aus Anwendung von (i) auf $-f$. Sei $y \in (y_0, y_1)$.

$$\underline{\mathbb{Z}}: \quad y \in f((x_0, x_1)).$$

Wähle nun eine Folge $(\xi_n)_{n \in \mathbb{N}} \subseteq (x_0, x_1)$ mit $\lim_{n \rightarrow \infty} \xi_n = x_0$. Dann gilt wegen Definition 6.1.3 und (294)

$$\lim_{n \rightarrow \infty} f(\xi_n) \leq y_0 < y = \lim_{n \rightarrow \infty} y.$$

Nach Aufgabe 1(a) auf Blatt 6^a gilt dann auch $(f(\xi_n))_{n \in \mathbb{N}} \leq (y)_{n \in \mathbb{N}}$, d.h.

$$\exists n_0 \in \mathbb{N} : \quad f(\xi_n) \leq y \quad \forall n \geq n_0. \quad (295)$$

Definiere dann

$$a := \xi_{n_0} \in (x_0, x_1).$$

Dann gilt nach (295)

$$f(a) \leq y \tag{296}$$

Wähle als nächstes eine Folge $(\zeta_n)_{n \in \mathbb{N}}$ mit $\lim_{n \rightarrow \infty} \zeta_n = x_1$. Dann gilt

$$\lim_{n \rightarrow \infty} f(\zeta_n) \geq y_1 > y = \lim_{n \rightarrow \infty} y.$$

Nach Blatt 6 Aufgabe 1(a) gilt dann auch $(y)_{n \in \mathbb{N}} \leq (f(\zeta_n))_{n \in \mathbb{N}}$, d.h.

$$\exists n_1 \in \mathbb{N} : \quad f(\zeta_n) \leq y \quad \forall n \geq n_1.$$

Da $\zeta_n \xrightarrow{n \rightarrow \infty} x_1$ und $x_1 > a$ gibt es auch ein $n_2 \in \mathbb{N}$ mit $\zeta_n \geq a$ für alle $n \geq n_2$ (auch wieder dank Blatt 6 Aufgabe 1(a), Details sind Ihnen überlassen). Definiere nun

$$b := \zeta_{\max\{n_1, n_2\}} \in (x_0, x_1).$$

Dann gilt $b \geq a$ und $f(b) \geq y$. Dies zusammen mit (296) zeigt

$$y \in [f(a), f(b)].$$

Nach dem Zwischenwertsatz gilt dann (weil $a, b \in \mathbb{R}, a \leq b$, vgl Satz 6.2.15(!))

$$\exists x \in [a, b] : \quad f(x) = y.$$

Da $[a, b] \subseteq (x_0, x_1)$ haben wir auch

$$\exists x \in (x_0, x_1) : \quad f(x) = y.$$

Das bedeutet $y \in f((x_0, x_1))$. Die Behauptung folgt.

^aEine kleine Spitzfindigkeit erzeugen wir an dieser Stelle, wenn $\lim_{n \rightarrow \infty} f(\xi_n)$ nur im Sinne der bestimmten Divergenz, d.h. als „ $\lim_{n \rightarrow \infty} f(\xi_n)$ “ existiert. In dem Fall müsste man die Aussage von Aufgabe 1(a) von Blatt 6 auch herleiten, wenn die dort betrachteten Folgen nicht konvergieren sondern bestimmt divergieren. Dies ist möglich und den Lesenden als Übungsaufgabe überlassen.

Eine wichtige Beobachtung, die aus dem Zwischenwertsatz folgt, ist, dass für eine stetige Funktion $f : I \rightarrow \mathbb{R}$, definiert auf einem Intervall $I \subseteq \mathbb{R}$, auch das Bild $f(I)$ wieder ein Intervall ist

Korollar 6.2.18 — INTERVALLTREUE

Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ stetig. Dann ist $f(I)$ ein Intervall.

Beweisskizze.

FALL 1. I ist abgeschlossenes beschränktes Intervall, d.h. $I = [a, b]$ für $a, b \in \mathbb{R}, a \leq b$.

Wir wissen nach dem Satz von Weierstraß (Satz 6.2.12), dass f ein Maximum und ein Minimum annimmt. Seien etwa $x_1, x_2 \in [a, b]$ mit

$$f(x_1) = \min_{x \in [a, b]} f(x) \quad \text{und} \quad f(x_2) = \max_{x \in [a, b]} f(x). \quad (297)$$

BEHAUPTUNG. $f([a, b]) = [f(x_1), f(x_2)]$

Zuerst zu „ $\subseteq$ “. Für $x \in [a, b]$ gilt nach (297), dass $f(x_1) \leq f(x) \leq f(x_2)$, d.h. $f(x) \in [f(x_1), f(x_2)]$.

Nun zu „ $\supseteq$ “. Es sei $y \in [f(x_1), f(x_2)]$. OBdA gelte auch $x_1 \leq x_2$, sonst kann man $-f$ betrachten. Die Einschränkung $f : [x_1, x_2] \rightarrow \mathbb{R}$ ist dann nach wie vor eine stetige Funktion. Nach dem Zwischenwertsatz (Satz 6.2.15) gibt es ein $x \in [x_1, x_2]$ mit $f(x) = y$. Somit gilt $y \in f([x_1, x_2]) \subseteq f([a, b])$.

*FALL 2. I ist ein allgemeines Intervall.

Diese Fragestellung kann stets mit *Ausschöpfung* behandelt werden. Sei zum Beispiel $I = [a, b]$ mit $a < b$. In dem Fall verfahren wir so: Für $n \in \mathbb{N}$ genügend groß (so, dass $b - \frac{1}{n} > a$, etwa ab einem Index $n_0 \in \mathbb{N}$) ist $J_n := f([a, b - \frac{1}{n}])$ ein Intervall. Außerdem gilt $J_n \subseteq J_{n+1}$. Insbesondere ist nach Fall 1 $J_n = [m_n, M_n]$ für alle $n \geq n_0$. Weil die Intervalle immer größer werden ist $(m_n)_{n \geq n_0}$ monoton fallend und $(M_n)_{n \geq n_0}$ monoton wachsend. Definiere nun

$$J := \{y \in \mathbb{R} : m_n \leq y \leq M_n \ \forall n \geq n_0\}.$$

Mit etwas Arbeit überzeugt man sich, dass J ein Intervall ist. Details dazu sind den Lesenden als Übungsaufgabe überlassen. Ähnlich wie in Blatt 2 Aufgabe 2(a) sieht man auch, dass

$$f([a, b)) = J,$$

was die Behauptung zeigt. Für andere Intervalltypen kann man ähnlich vorgehen.

Bemerkung 6.2.19

Wir möchten noch eine wichtige Erkenntnis aus dem obigen Beweis herausstellen: Ist $I = [a, b]$ (für $a, b \in \mathbb{R}, a \leq b$) und $f : [a, b] \rightarrow \mathbb{R}$ stetig, so ist

$$f([a, b]) = \left[\min_{x \in [a, b]} f(x), \max_{x \in [a, b]} f(x) \right].$$

Stetigkeit der Umkehrfunktion

Stetigkeit ist ein *Qualitätsmerkmal* einer Funktion. Solche Merkmale nennt man auch oft die *Regularität* einer Funktion. In diesem Abschnitt zeigen wir, dass diese Regularität beim Bilden der Umkehrabbildung nicht verloren geht. Genauer zeigen wir

Satz 6.2.20 — STETIGKEIT DER UMKEHRFUNKTION

Es sei $Z \subseteq \mathbb{R}$ und $f : [a, b] \rightarrow Z$ bijektiv und stetig^a. Dann ist $f^{-1} : Z \rightarrow [a, b]$ auch stetig.

^aBeachte: Um von Stetigkeit sprechen zu können müssten wir eigentlich $f : [a, b] \rightarrow \mathbb{R}$ haben. Da $Z \subseteq \mathbb{R}$ kann man das aber so auffassen.

Insbesondere ist $f^{-1} : Z \rightarrow [a, b]$ sogar *stetig und bijektiv*, genau wie f ! In der Tat: f^{-1} ist auch bijektiv, weil f^{-1} auch eine Umkehrfunktion besitzt (nämlich f). Da alle Funktionen, die eine Umkehrfunktion besitzen, bijektiv sind (vgl. Satz 1.3.18), folgt die Bijektivität von f^{-1} .

Den obigen Satz werden wir am Ende des Abschnittes beweisen. Zunächst eine Diskussion der Voraussetzungen. Die Bijektivität wird für die Existenz von f^{-1} benötigt, dies ist also eine sehr natürliche Voraussetzung. Etwas restriktiv mag die Annahme erscheinen, dass f auf einem Intervall definiert sein muss...

Bemerkung 6.2.21

Auf die Voraussetzung im obigen Satz, dass f auf einem kompakten Intervall $[a, b]$ definiert ist, kann nicht ohne Weiteres verzichtet werden. Betrachte zum Beispiel

$$f : [1, 2] \cup (3, 4] \rightarrow [1, 3]$$

definiert durch

$$f(x) := \begin{cases} x & \text{falls } x \in [1, 2] \\ x - 1 & \text{falls } x \in (3, 4] \end{cases}$$

Überzeugen Sie sich: Dann ist f auf $D := [1, 2] \cup (3, 4]$ stetig und von D nach $[1, 3]$ bijektiv. Ferner ist $f^{-1} : [1, 3] \rightarrow D$ gegeben durch

$$f^{-1}(x) = \begin{cases} x & \text{falls } x \in [1, 2] \\ x + 1 & \text{falls } x \in (2, 3] \end{cases}.$$

Diese Funktion ist bei $x_0 = 2$ jedoch unstetig (wie man sich leicht überzeugen kann). Die Aussage von Satz 6.2.20 gilt also nicht, weil D kein Intervall ist.

Bevor wir diesen Satz beweisen benötigen wir ein wenig Vorbereitung. Wir müssen uns mit dem *Monotonieverhalten* von Funktionen auseinandersetzen.

Definition 6.2.22 — MONOTONIE VON FUNKTIONEN

Es sei $D \subseteq \mathbb{R}$ und $f : D \rightarrow \mathbb{R}$ eine Funktion. Wir sagen f ist auf D

- *monoton wachsend* wenn für alle $x_1, x_2 \in D$ mit $x_1 \leq x_2$ gilt $f(x_1) \leq f(x_2)$,
- *monoton fallend* wenn für alle $x_1, x_2 \in D$ mit $x_1 \leq x_2$ gilt $f(x_1) \geq f(x_2)$,
- *streng monoton wachsend* wenn für alle $x_1, x_2 \in D$ mit $x_1 < x_2$ gilt $f(x_1) < f(x_2)$,
- *streng monoton fallend* wenn für alle $x_1, x_2 \in D$ mit $x_1 < x_2$ gilt $f(x_1) > f(x_2)$.

Bemerkung 6.2.23 — MONOTONIE UND INJEKTIVITÄT

Das Monotonieverhalten einer Funktion kann verwendet werden um die Injektivität zu diskutieren: Eine streng monoton wachsende/fallende Funktion ist sicherlich auch injektiv. In der Tat: Ist $f : D \rightarrow \mathbb{R}$ etwa streng monoton wachsend und $x_1, x_2 \in D$ so gilt

$$x_1 \neq x_2 \Rightarrow (x_1 < x_2) \vee (x_2 < x_1) \Rightarrow (f(x_1) < f(x_2)) \vee (f(x_2) < f(x_1)) \Rightarrow f(x_1) \neq f(x_2).$$

Die Kontraposition ergibt

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

also die Injektivität von f (vgl. Proposition 1.3.10).

Bei stetigen Funktionen gilt nun auch eine Rückrichtung: Jede stetige injektive Funktion muss monoton sein!

Lemma 6.2.24 — INJEKTIVITÄT UND MONOTONIE

Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und injektiv. Dann ist f entweder streng monoton wachsend oder streng monoton fallend. Insbesondere gilt für alle Teilintervalle $[c, d] \subseteq [a, b]$ stets $f([c, d]) = [f(c), f(d)]$ oder $f([c, d]) = [f(d), f(c)]$.

Beweis. Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und injektiv. Da f injektiv ist, gilt $f(a) \neq f(b)$. Wir unterscheiden nun zwei Fälle.

FALL 1. $f(a) < f(b)$.

Wir behaupten, dass f dann streng monoton wachsend ist. Dazu seien $x_1, x_2 \in [a, b]$ mit $x_1 < x_2$. Zu zeigen ist, dass $f(x_1) < f(x_2)$. Angenommen $f(x_1) \geq f(x_2)$. Da f injektiv ist, ist Gleichheit nicht erlaubt. Sprich: $f(x_1) > f(x_2)$.

- Ist $f(x_2) < f(b)$, so wähle ein $y_0 \in (f(x_2), \min\{f(x_1), f(b)\})$. Ein solcher Wert y_0 wird nun $[x_1, x_2]$ angenommen, denn

$$y_0 \in [f(x_2), f(x_1)] \xrightarrow{\text{Satz 6.2.15}} \exists x_* \in [x_1, x_2] : f(x_*) = y_0$$

Jedoch gilt $x_* \neq x_2$, da $y_0 \neq f(x_2)$, weswegen wir $x_* \in [x_1, x_2)$ erhalten. Außerdem

wird y_0 nochmal in $[x_2, b]$ angenommen, denn

$$y_0 \in [f(x_2), f(b)] \xrightarrow[\text{Satz 6.2.15}]{} \exists x^* \in [x_2, b] : f(x^*) = y_0.$$

Da $x_* \in [x_1, x_2)$ und $x^* \in [x_2, b]$ gilt $x_* \neq x^*$. Man erhält einen Widerspruch zur Injektivität von f , da $f(x_*) = f(x^*) = y_0$.

- Ist $f(x_2) \geq f(b)$ so gilt insbesondere $f(x_1) > f(x_2) \geq f(b)$. Setze nun $y_0 = f(b)$. Da $f(b) > f(a)$ gilt $y_0 \in [f(a), f(x_1)]$. Wiederum hat man

$$y_0 \in [f(a), f(x_1)] \xrightarrow[\text{Satz 6.2.15}]{} \exists x^* \in [a, x_1] : f(x^*) = y_0.$$

Somit wird $y_0 = f(b)$ einmal im Intervall $[a, x_1]$ und einmal beim Wert b angenommen. Ein Widerspruch zur Injektivität von f .

FALL 2. $f(a) > f(b)$.

Betrachten Sie $g := -f$. Dann ist auch g stetig und injektiv und $g(a) < g(b)$. Man folgert aus Fall 1, dass g streng monoton wachsend ist. Daraufhin überlegt man sich leicht, dass f monoton fallend ist.

Noch $\mathbb{Z}$: Für alle $c, d \in [a, b]$ mit $c < d$ gilt $f([c, d]) = [f(c), f(d)]$

Dies folgt direkt aus der Monotonie und aus Bemerkung 6.2.19:

$$f([c, d]) = \left[\min_{x \in [c, d]} f(x), \max_{x \in [c, d]} f(x) \right] \xrightarrow[\text{Monotonie}]{} [f(c), f(d)].$$

Auf die Stetigkeit kann im vorigen Lemma übrigens nicht verzichtet werden! Auch dafür verweisen wir auf die Übungen!

Jetzt sind wir endlich in der Lage, Satz 6.2.20 zu beweisen.

Beweis (von Satz 6.2.20). Es sei $f : [a, b] \rightarrow Z$ stetig und bijektiv. Nach Lemma 6.2.1 ist f entweder streng monoton wachsend oder streng monoton fallend. Wir nehmen oBdA an, dass f streng monoton wachsend ist, sonst betrachte einfach $-f$. Insbesondere gilt $Z = f([a, b]) = [f(a), f(b)]$. Sei nun $y_0 \in [f(a), f(b)]$

$\mathbb{Z}$: f^{-1} ist bei y_0 stetig.

Hierzu verwenden wir das Epsilon-Delta-Kriterium, d.h.

$\mathbb{Z}$: $\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall y \in Z \text{ mit } |y - y_0| < \delta \quad \text{gilt } |f^{-1}(y) - f^{-1}(y_0)| < \varepsilon.$

FALL 1. $y_0 \in (f(a), f(b))$.

Zunächst: Wegen der Bijektivität von f gibt es ein (eindeutiges) $x_0 \in [a, b]$ mit $f(x_0) = y_0$. Da $y_0 \notin \{f(a), f(b)\}$ muss außerdem gelten $x_0 \notin \{a, b\}$. Das bedeutet also $x_0 \in (a, b)$. Nun sei $\varepsilon > 0$ beliebig vorgegeben. Wähle dann $\varepsilon' \leq \varepsilon$ so, dass $x_0 - \varepsilon' > a$ und $x_0 + \varepsilon' < b$ (zum Beispiel durch $\varepsilon' = \frac{1}{2} \min\{\varepsilon, x_0 - a, b - x_0\}$). Beachte: Da $x_0 - \varepsilon' < x_0 < x_0 + \varepsilon'$ gilt wegen der strengen Monotonie

$$f(x_0 - \varepsilon') < \underbrace{f(x_0)}_{=y_0} < f(x_0 + \varepsilon'). \quad (298)$$

Definiere nun

$$\delta := \min\{y_0 - f(x_0 - \varepsilon'), f(x_0 + \varepsilon') - y_0\}. \quad (299)$$

Dass $\delta > 0$ folgt sofort aus (298). Es sei nun $y \in Z$ mit $|y - y_0| < \delta$, d.h. insbesondere auch

$$y \in (y_0 - \delta, y_0 + \delta). \quad (300)$$

Ferner gibt es wegen der Bijektivität auch ein $x \in [a, b]$ mit $y = f(x)$. Beachte auch: $x = f^{-1}(y)$.

ZWISCHENBEHAUPTUNG. $x > x_0 - \varepsilon'$ und $x < x_0 + \varepsilon'$,
(d.h. auch $x \in (x_0 - \varepsilon', x_0 + \varepsilon')$ bzw. $|x - x_0| < \varepsilon'$)

In der Tat: Wäre $x \leq x_0 - \varepsilon'$ so wäre wegen der Monotonie

$$f(x) \leq f(x_0 - \varepsilon') = y_0 - (y_0 - f(x_0 - \varepsilon')) \underset{(299)}{\leq} y_0 - \delta \underset{(300)}{<} y$$

Ein Widerspruch zu $f(x) = y$. Es folgt $x > x_0 - \varepsilon'$. Genauso kann man auch zeigen, dass $x < x_0 + \varepsilon'$. [In der Tat: Wäre $x \geq x_0 + \varepsilon'$ so wäre wegen der Monotonie

$$f(x) \geq f(x_0 + \varepsilon') = y_0 + (f(x_0 + \varepsilon') - y_0) \geq y_0 + \delta > y$$

Wiederum ein Widerspruch zu $f(x) = y$.]

Die Zwischenbehauptung folgt.

Damit erhalten wir auch für alle $y \in Z$ mit $|y - y_0| < \delta$

$$|f^{-1}(y) - f^{-1}(y_0)| = |x - x_0| < \varepsilon' \leq \varepsilon.$$

Die Stetigkeit bei y_0 ist gezeigt. Nun fehlt nur noch Stetigkeit an den Randpunkten:

$$\text{FALL 2. } y_0 \in \{f(a), f(b)\}$$

Dieser Fall geht mit ähnlicher Technik, nur mit einseitigen Intervallen. Die Details sind den Lesenden als Übungsaufgabe überlassen.

Als Korollar erhalten wir die Stetigkeit der *Wurzelfunktionen*.

Korollar 6.2.25 — STETIGKEIT DER WURZELFUNKTION

Es sei $k \in \mathbb{N}$. Dann ist $f : [0, \infty) \rightarrow [0, \infty), x \mapsto \sqrt[k]{x}$ auf ganz $[0, \infty)$ stetig.

Beweis. Es sei $x_0 \in [0, \infty)$. Zu zeigen ist, dass f bei x_0 stetig ist. Dazu wähle zunächst einen Wert $b > \sqrt[k]{x_0}$. Dann ist nach Korollar 3.3.21(ii) auch

$$b^k > x_0. \quad (301)$$

Nach Korollar 6.2.8 ist das Polynom

$$h : [0, b] \rightarrow [0, b^k], x \mapsto x^k$$

auf ganz $[0, b]$ stetig und bijektiv (da streng monoton wachsend, vgl. Korollar 3.3.21). Somit ist die Umkehrfunktion h^{-1} auf $Z = [0, b^k]$ stetig, also insbesondere bei x_0 , vgl. (301). Man überzeugt sich leicht, dass $h^{-1} = f$ und erhält die behauptete Stetigkeit von f bei x_0 .

6.3 Anwendung: Logarithmus und allgemeine Potenz

Der natürliche Logarithmus

Im Folgenden werden wir zeigen, dass die *reelle Exponentialfunktion* $\exp : \mathbb{R} \rightarrow (0, \infty)$ eine Umkehrfunktion besitzt und dass diese auch stetig ist. Wir werden diese *natürlicher Logarithmus* nennen. Dies wird ein nützliches Tool sein, um die *allgemeine Potenz* x^α für $\alpha \in \mathbb{R}, x \in (0, \infty)$ zu definieren.

Proposition 6.3.1

Die Funktion $\exp : \mathbb{R} \rightarrow (0, \infty)$ ist bijektiv und streng monoton wachsend. Ferner ist die Umkehrfunktion $\exp^{-1} : (0, \infty) \rightarrow \mathbb{R}$ stetig.

Beweis.

ZUR SURJEKTIVITÄT.

Wir zeigen $\exp(\mathbb{R}) = (0, \infty)$. Die Teilmengen-Inklusion „ $\subseteq$ “ folgt bereits aus Korollar 5.4.7. Nun also zu „ $\supseteq$ “. Es sei $y \in (0, \infty)$ beliebig.

FALL 1. $y \in [1, \infty)$.

Dann gibt es nach Proposition 6.2.16 ein $x \in [0, \infty)$ mit $y = \exp(x)$. Somit gilt sicher $y \in \exp(\mathbb{R})$.

FALL 2. $y \in (0, 1)$.

Dann ist $\frac{1}{y} \in (1, \infty)$ und somit gibt es nach Proposition 6.2.16 ein $x' \in [0, \infty)$ mit $\exp(x') = \frac{1}{y}$. Nun gilt

$$\exp(-x') = \frac{1}{\exp(x')} = \frac{1}{\frac{1}{y}} = y.$$

Somit ist $y = \exp(-x') \in \exp(\mathbb{R})$.

ZUR INJEKTIVITÄT (UND STRENGEN MONOTONIE).

Da die strenge Monotonie die Injektivität impliziert (vgl. Bemerkung 6.2.23) genügt es, die strenge Monotonie nachzuweisen. Es seien $x_1, x_2 \in \mathbb{R}, x_1 < x_2$. Dann gilt nach der Funktionalgleichung

$$\exp(x_2) = \exp(x_1 + x_2 - x_1) = \exp(x_1) \exp(x_2 - x_1). \quad (302)$$

Nun ist außerdem

$$\begin{aligned} \exp(x_2 - x_1) &= \sum_{k=0}^{\infty} \frac{(x_2 - x_1)^k}{k!} = \frac{(x_2 - x_1)^0}{0!} + \frac{(x_2 - x_1)^1}{1!} + \underbrace{\sum_{k=2}^{\infty} \frac{(x_2 - x_1)^k}{k!}}_{>0} \\ &\geq \frac{(x_2 - x_1)^0}{0!} + \frac{(x_2 - x_1)^1}{1!} = 1 + (x_2 - x_1) > 1. \end{aligned}$$

Da $\exp(x_1) > 0$ gilt auch

$$\exp(x_2) \underset{(302)}{=} \exp(x_1) \exp(x_2 - x_1) > \exp(x_1) \cdot 1 = \exp(x_1).$$

ZUR STETIGKEIT DER UMKEHRFUNKTION.

Folgt direkt aus Satz 6.2.20.

22.01.2026

Definition 6.3.2 — NATÜRLICHER LOGARITHMUS

Die Umkehrfunktion aus Proposition 6.3.1 nennen wir (*natürlicher*) *Logarithmus* und bezeichnen sie mit $\log : (0, \infty) \rightarrow \mathbb{R}$ oder $\ln : (0, \infty) \rightarrow \mathbb{R}$ ^a

^aWir werden hier eigentlich immer $\log$ als Bezeichnung verwenden. Dies weicht eventuell von der Notation ab, die Sie aus der Schule kennen — dort ist $\ln$ oft üblicher.

Im folgenden wollen wir die Eigenschaften des Logarithmus studieren. Ein wichtiger Wert des Logarithmus ist der Folgende

Proposition 6.3.3
 $\log(1) = 0, \log(e) = 1.$

Beweis. Da $\log$ die Umkehrfunktion von $\exp$ ist, sind die Behauptungen äquivalent zu $\exp(0) = 1$ und $\exp(1) = e$. Dies sind aber gerade die Aussagen von Beispiel 5.4.3 und Proposition 5.4.10.

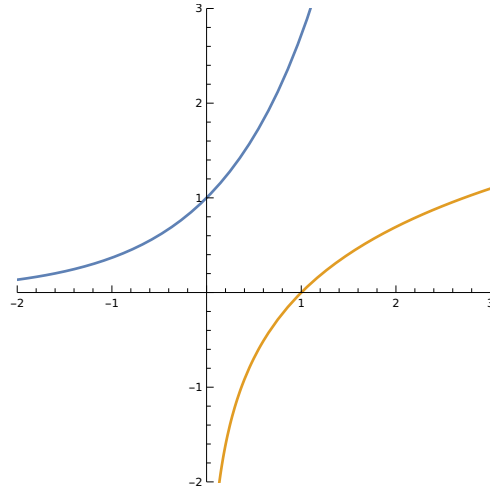


Abbildung 26: Graph der Exponentialfunktion (blau) und der Logarithmusfunktion (gelb)

Aus der Funktionalgleichung für $\exp$ erbt sich auch eine Funktionalgleichung für den $\log$:

Proposition 6.3.4 — FUNKTIONALGLEICHUNG FÜR DEN LOGARITHMUS

Es seien $x, y \in (0, \infty)$. Dann gilt

$$\log(xy) = \log(x) + \log(y).$$

Beweis. Es gilt wegen der Funktionalgleichung von $\exp$ (vgl. Satz 5.4.5)

$$\exp(\log(x) + \log(y)) = \underbrace{\exp(\log(x))}_{=x} \underbrace{\exp(\log(y))}_{=y} = xy = \exp(\log(xy))$$

Nutzen wir nun die Injektivität von $\exp : \mathbb{R} \rightarrow (0, \infty)$ so folgt

$$\log(x) + \log(y) = \log(xy),$$

was zu zeigen war.

Als Korollar ergibt sich ein weiteres *Logarithmusgesetz*.

Korollar 6.3.5

Es seien $x, y \in (0, \infty)$. Dann gilt

$$\log\left(\frac{x}{y}\right) = \log(x) - \log(y). \quad (303)$$

Insbesondere gilt für $x \in (0, \infty)$, dass $\log\left(\frac{1}{x}\right) = -\log(x)$.

Beweis. Zuerst zu (303). Seien x, y wie in der Aussage. Dann gilt

$$\log\left(\frac{x}{y}\right) + \log(y) \stackrel{\text{Prop 6.3.4}}{=} \log\left(\frac{x}{y} \cdot y\right) = \log(x).$$

Subtrahieren wir $\log(y)$ auf beiden Seiten so folgt die Behauptung.

Nun noch zum letzten Satz der Aussage. Nach (303) gilt

$$\log\left(\frac{1}{x}\right) = \log(1) - \log(x) \stackrel{\text{Prop 6.3.3}}{=} 0 - \log(x) = -\log(x).$$

Auch das Monotonieverhalten des Logarithmus kann studiert werden

Proposition 6.3.6 — MONOTONIE DES LOGARITHMUS

$\log : (0, \infty) \rightarrow \mathbb{R}$ ist streng monoton wachsend.

Beweis. Es seien $x_1, x_2 \in (0, \infty)$ mit $x_1 < x_2$.

$$\text{Z:} \quad \log(x_1) < \log(x_2)$$

Dazu: Angenommen $\log(x_1) \geq \log(x_2)$. Dann gilt wegen der Monotonie von $\exp$ auch

$$\exp(\log(x_1)) \geq \exp(\log(x_2))$$

Da $\exp$ die Umkehrfunktion von $\log$ ist folgt

$$x_1 \geq x_2,$$

was ein Widerspruch zu $x_1 < x_2$ ist.

QED

Die allgemeine Potenz

Mithilfe des Logarithmus können wir auch die *allgemeine Potenz* einführen.

Definition 6.3.7 — ALLGEMEINE POTENZ

Für $x \in (0, \infty)$ und $\alpha \in \mathbb{C}$ definieren wir

$$x^\alpha := \exp(\alpha \log(x)). \quad (304)$$

An dieser Stelle sollten wir eines klären: Für $z \in \mathbb{C}$ haben wir e^z jetzt auf zwei verschiedene Weisen definiert. Zum einen wäre nach (304)

$$e^z = \exp(z \log(e))$$

und zum anderen ist nach Notation 5.4.1

$$e^z := \exp(z).$$

Da nach $\log(e) = 1$ (vgl. Proposition 6.3.3) sind die beiden vorigen Notationen miteinander konsistent — unsere Definition erzeugt also keinen Widerspruch! Ein weiterer Widerspruch der entstehen könnte ist, dass für $n \in \mathbb{N}$ und $x \in (0, \infty)$ die Größe x^n doppelt definiert ist (einmal in Definition 6.3.7 als $x^n = \exp(n \log(x))$ und einmal in Definition 3.3.8 über eine Rekursion). Dass diese Definitionen miteinander konsistent sind folgt direkt aus Proposition 5.4.9. Die Details sind den Lesenden als Übungsaufgabe überlassen.

Bemerkung 6.3.8

Für $\alpha \in \mathbb{R}$ und $x \in (0, \infty)$ folgt aus Korollar 5.4.7, dass

$$x^\alpha = \exp(\alpha \log(x)) \in (0, \infty).$$

Für $\alpha \in \mathbb{C} \setminus \mathbb{R}$ ist diese Aussage im Allgemeinen falsch.

Bemerkung 6.3.9

Aus Definition 6.3.7 ergibt sich direkt auch ein weiteres *Logarithmusgesetz*. Nimm $x \in (0, \infty)$, $\alpha \in \mathbb{R}$ und wende $\log$ auf beiden Seiten von (304) an.^a Man erhält

$$\log(x^\alpha) = \alpha \log(x).$$

^aBeachte: Hier ist $\alpha \in \mathbb{R}$ wichtig! Ist nämlich $\alpha \in \mathbb{C} \setminus \mathbb{R}$, so können wir uns nicht sicher sein, dass der in (304) betrachtete Ausdruck in $(0, \infty)$ liegt. Ist dieser Ausdruck jedoch nicht in $(0, \infty)$, so dürfen wir den Logarithmus nicht anwenden.

Als nächstes sprechen wir über die *Potenzgesetze* für die verallgemeinerte Potenz, vgl. auch Proposition 3.3.9.

Proposition 6.3.10 — POTENZGESETZE ALLGEMEIN

Es seien $x, y \in (0, \infty)$ und $\alpha, \beta \in \mathbb{C}$. Dann gilt

(i) $(xy)^\alpha = x^\alpha y^\alpha,$

(ii) $x^{\alpha+\beta} = x^\alpha x^\beta,$

- (iii) Falls $y \neq 0$ so gilt $(\frac{x}{y})^\alpha = \frac{x^\alpha}{y^\alpha}$,
- (iv) $x^{-\beta} = \frac{1}{x^\beta}$,
- (v) Falls $\alpha \in \mathbb{R}$ so gilt $x^{\alpha\beta} = (x^\alpha)^\beta$,
- (vi) $x^0 = 1$.
- (vii) Für $n \in \mathbb{Z}, m \in \mathbb{N}$ gilt $x^{\frac{n}{m}} = (\sqrt[m]{x})^n$

Beweis. Seien $x, y \in (0, \infty)$.

Zu (i).

$$\begin{aligned} (xy)^\alpha &\stackrel{(304)}{=} \exp(\alpha \log(xy)) \stackrel{\text{Prop 6.3.4}}{=} \exp(\alpha [\log(x) + \log(y)]) \\ &= \exp(\alpha \log(x) + \alpha \log(y)) = \exp(\alpha \log(x)) \exp(\alpha \log(y)) \stackrel{(304)}{=} x^\alpha y^\alpha. \end{aligned}$$

Zu (ii).

$$\begin{aligned} x^{\alpha+\beta} &\stackrel{(304)}{=} \exp((\alpha + \beta) \log(x)) = \exp(\alpha \log(x) + \beta \log(x)) \\ &= \exp(\alpha \log(x)) \exp(\beta \log(x)) \stackrel{(304)}{=} x^\alpha x^\beta. \end{aligned}$$

Zu (iii).

$$\left(\frac{x}{y}\right)^\alpha y^\alpha \stackrel{(i)}{=} \left(\frac{x}{y}y\right)^\alpha = x^\alpha.$$

Zu (iv). Es gilt

$$x^{-\beta} = \exp(-\beta \log(x)) \stackrel{\text{Kor 5.4.6}}{=} \frac{1}{\exp(\beta \log(x))} = \frac{1}{x^\beta}.$$

Zu (v). Es gilt für $\alpha \in \mathbb{R}$

$$(x^\alpha)^\beta = \exp(\beta \log(x^\alpha)) \stackrel{\text{Bem 6.3.9}}{=} \exp(\beta \alpha \log(x)) = x^{\beta\alpha}.$$

Zu (vi).

$$x^0 = \exp(0 \cdot \log(x)) = \exp(0) = 1.$$

Zu (vii). Wegen der Multiplikativität der Wurzel (vgl. Proposition 3.4.8) gilt $(\sqrt[n]{x})^n = \sqrt[n]{x^n}$. Nun gilt nach der Regel (v)

$$(x^{\frac{n}{m}})^m \underset{(v)}{=} x^{\frac{n}{m} \cdot m} = x^n$$

Somit ist $y := x^{\frac{n}{m}}$ eine Lösung der Gleichung $y^m = x^n$. Ferner ist $y > 0$, vgl. Bemerkung 6.3.8. Wir erhalten aus der Eindeutigkeit Lösung von $y^m = \alpha$, vgl. Proposition 3.4.1, dass $y = \sqrt[m]{x^n}$. Die Behauptung folgt.

Bemerkung 6.3.11

Das bekannte Potenzgesetz Nummer (v) der vorigen Proposition gilt *nicht* für $\alpha \in \mathbb{C} \setminus \mathbb{R}$! Ein Beispiel (welches die Kreizahl π benutzt, vgl. Übungsaufgabe 2(a) auf Blatt 13): Wäre Proposition 6.3.10(v) auch für $\alpha, \beta \in \mathbb{C}$ gültig so wäre für alle $x \in \mathbb{R}$

$$e^{2\pi i x} = (e^{2\pi i})^x = 1^x = 1.$$

Setzen wir jedoch z.B. $x = \frac{1}{2}$ so ergibt sich

$$e^{2\pi i x} = e^{i\pi} = -1.$$

Eine Sache, die ich noch dazu erwähnen möchte: Falls $\beta \in \mathbb{N}$, so gilt Potenzgesetz (v) doch für alle $\alpha \in \mathbb{C}$. Das heißt für alle $\alpha \in \mathbb{C}$ und $n \in \mathbb{N}$ gilt

$$(x^\alpha)^n = x^{n\alpha}.$$

Dies zu zeigen ist eine leichte Induktion über $n \in \mathbb{N}$.

Als nächstes folgt eine Stetigkeitsuntersuchung für die allgemeine Potenz.

Proposition 6.3.12

- (i) Es sei $x \in (0, \infty)$ beliebig vorgegeben. Dann ist die Funktion $f : \mathbb{C} \rightarrow \mathbb{C}$, $f(\alpha) := x^\alpha$ auf ganz $\mathbb{C}$ stetig.
- (ii) Es sei $\alpha \in \mathbb{C}$ beliebig vorgegeben. Dann ist die Funktion $g : (0, \infty) \rightarrow \mathbb{C}$, $g(x) := x^\alpha$ auf ganz $(0, \infty)$ stetig.

Beweis. Zu (i). Beachte, dass für alle $\alpha \in \mathbb{C}$

$$f(\alpha) = x^\alpha = \exp(\alpha \log(x)) \quad (305)$$

Nun ist $\alpha \mapsto \alpha \log(x)$ als Polynom in α auf $\mathbb{C}$ stetig und auch $\exp$ ist auf $\mathbb{C}$ stetig. Da Kompositionen stetiger Funktionen stetig sind (vgl. Proposition 6.2.6) folgt die Stetigkeit von f aus (305).

Zu (ii). Beachte, dass für alle $x \in (0, \infty)$

$$g(x) = x^\alpha = \exp(\alpha \log(x))$$

Da $\log$ auf $(0, \infty)$ stetig ist und $\alpha \in \mathbb{C}$ eine feste Konstante, ist auch $x \mapsto \alpha \log(x)$ stetig. Mit der Stetigkeit von $\exp$ und Proposition 6.2.6 folgt wiederum die Stetigkeit von g .

Diskussion der Funktionsgraphen

Nun können wir Monotonie und Grenzwertverhalten für die allgemeine Potenz untersuchen. Ziel dabei ist, mehr über den Funktionsgraphen bzw. das Funktionswachstum der Potenzfunktionen zu lernen.

Proposition 6.3.13

- (i) Sei $\alpha \in \mathbb{R} : \alpha > 0$. Dann ist $g : (0, \infty) \rightarrow \mathbb{R} \ x \mapsto x^\alpha$ streng monoton wachsend.
- (ii) Sei $\alpha \in \mathbb{R} : \alpha < 0$. Dann ist $f : (0, \infty) \rightarrow \mathbb{R} \ x \mapsto x^\alpha$ streng monoton fallend.
- (iii) Für $\alpha \in \mathbb{R} : \alpha > 0$ gilt

$$\lim_{x \rightarrow 0} x^\alpha = 0.$$

Beweis. Zu (i). Sei $\alpha > 0$. Es seien $x_1, x_2 \in (0, \infty)$ mit $x_1 < x_2$. Nach Proposition 6.3.6 ist auch $\log(x_1) < \log(x_2)$. Weil $\alpha > 0$ gilt dann auch $\alpha \log(x_1) < \alpha \log(x_2)$. Wegen der Monotonie der Exponentialfunktion ist dann

$$\exp(\alpha \log(x_1)) < \exp(\alpha \log(x_2)).$$

Mit anderen Worten $x_1^\alpha < x_2^\alpha$, vgl. Definition 6.3.7.

Zu (ii). Dies folgt direkt aus (i) und Proposition 6.3.10 (iv).

Zu (iii). Es sei $\alpha > 0$. Dann gibt es eine natürliche Zahl $n \in \mathbb{N}$ mit

$$\frac{1}{n} < \alpha \quad (306)$$

(wie man sich mithilfe der Aussage $\lim_{j \rightarrow \infty} \frac{1}{j} = 0$ leicht überlegt). Sei nun $x \in (0, 1)$. Dann gilt nach Proposition 6.3.6, dass

$$\log(x) < \log(1) = 0 \quad (\text{vgl. auch Proposition 6.3.3}).$$

Multiplizieren wir (306) mit $\log(x)$, so dreht sich also das Ungleichheitszeichen um. Man erhält

$$\alpha \log(x) < \frac{1}{n} \log(x).$$

Wegen der strengen Monotonie und $\exp(\mathbb{R}) = (0, \infty)$ folgt

$$0 < \exp(\alpha \log(x)) < \exp\left(\frac{1}{n} \log(x)\right)$$

also

$$0 < x^\alpha < x^{\frac{1}{n}} \stackrel{\text{Prop 6.3.10}}{=} \sqrt[n]{x}. \quad (307)$$

Da $x \mapsto \sqrt[n]{x}$ bei 0 stetig ist (vgl. Korollar 6.2.25) gilt

$$\lim_{x \rightarrow 0} \sqrt[n]{x} = \sqrt[n]{0} = 0$$

Dies und (307) machen das Sandwichlemma (Proposition 6.1.13) anwendbar und wir erhalten

$$\lim_{x \rightarrow 0} x^\alpha = 0.$$

Bemerkung 6.3.14

Aufgrund von Punkt (iii) in der vorigen Propostion treffen wir für $\alpha > 0$ auch die Konvention, dass

$$0^\alpha := 0.$$

Zum Abschluss diskutieren wir das *asymptotische Verhalten* von Logarithmus und allgemeiner Potenz. Hierzu zunächst einige Ungleichungen, die uns dabei von Nutzen sein werden.

Proposition 6.3.15 — WACHSTUMSUNGLEICHUNGEN FÜR EXP/LOG

- (i) Für alle $x \in [0, \infty)$ gilt $\exp(x) \geq 1 + x$
- (ii) Für alle $x \in (-\infty, 0]$ gilt $\exp(x) \leq \frac{1}{1-x}$.
- (iii) Für alle $y \in [1, \infty)$ gilt $\log(y) \leq y - 1$.

Beweis. Die ersten beiden Ungleichungen haben wir auf Übungsblatt 11 (Aufgabe 3) gezeigt. For convenience wiederholen wir hier die Beweise.

Zu (i). Wegen $x \in [0, \infty)$ ist auch $\frac{1}{k!}x^k \geq 0 \ \forall k \in \mathbb{N}$ und somit

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \sum_{k=0}^1 \frac{x^k}{k!} + \underbrace{\sum_{k=2}^{\infty} \frac{x^k}{k!}}_{\geq 0} = \sum_{k=0}^1 \frac{x^k}{k!} = 1 + x.$$

Zu (ii). Sei $x \in (-\infty, 0]$, dann ist $-x \in [0, \infty)$ und somit gilt nach (i)

$$\exp(-x) \geq 1 + (-x) = 1 - x$$

Weil $1 - x \geq 0$ folgt

$$\frac{1}{1-x} \geq \frac{1}{\exp(-x)}$$

Mit Korollar 5.4.6 (ii) folgt $\frac{1}{\exp(-x)} = \exp(-(-x)) = \exp(x)$ und somit

$$\frac{1}{1-x} \geq \exp(x).$$

Nun zu (iii). Angenommen es gibt ein $y_0 \in [1, \infty)$ mit $\log(y_0) > y_0 - 1$. Nach Proposition 6.2.16 gibt es ein $x_0 \in [0, \infty)$ mit $y_0 = \exp(x_0)$ [insbesondere gilt auch $x_0 = \log(y_0)$]. Dann gilt aber

$$y_0 = \exp(x_0) \underset{(i)}{\geq} x_0 + 1 = \log(y_0) + 1 > (y_0 - 1) + 1 = y_0.$$

Ein Widerspruch.

Man könnte die Ungleichung (i) übrigens auch für alle $x \in \mathbb{R}$ herleiten, (ii) für alle $x \in (-\infty, 1)$ und (iii) für alle $y \in (0, \infty)$. Wir benötigen dies an der Stelle nicht.

Proposition 6.3.16 — ASYMPTOTISCHES VERHALTEN VON POTENZ UND LOG

(i) Es gilt für alle $\beta > 1$

$$\lim_{x \rightarrow -\infty} \beta^x = 0 \quad \text{und} \quad \lim_{x \rightarrow \infty} \beta^x = \infty.$$

(ii) Es gilt für alle $\gamma \in (0, 1)$

$$\lim_{x \rightarrow -\infty} \gamma^x = \infty \quad \text{und} \quad \lim_{x \rightarrow \infty} \gamma^x = 0.$$

(iii) Es gilt

$$\lim_{x \rightarrow 0} \log(x) = -\infty \quad \text{und} \quad \lim_{x \rightarrow \infty} \log(x) = \infty.$$

(iv) Es gilt für alle $\alpha \in [0, \infty)$ und $\beta \in (1, \infty)$

$$\lim_{x \rightarrow \infty} \frac{x^\alpha}{\beta^x} = 0.$$

(v) Es gilt für alle $\alpha \in (0, \infty)$ und $\beta \in (0, \infty)$

$$\lim_{x \rightarrow \infty} \frac{\log(x)^\beta}{x^\alpha} = 0. \quad (a)$$

(vi) Für alle $\alpha \in (0, \infty)$ und $n \in \mathbb{N}$ gilt auch

$$\lim_{x \rightarrow 0+} x^\alpha \log(x)^n = 0.$$

^aAnders ausgedrückt: Der Logarithmus wächst langsamer als jede positive Potenz. Jedoch divergiert er trotzdem noch gegen ∞ , vgl. Punkt (iii)

05.02.2024

Beweis. Zu (i). *Erster Grenzwert.* Es sei $(x_n)_{n \in \mathbb{N}}$ eine bestimmt gegen $-\infty$ divergente Folge, d.h. „ $\lim_{n \rightarrow \infty} x_n = -\infty$ “.

$$\mathbb{Z}\mathbb{Z}: \quad \lim_{n \rightarrow \infty} \beta^{x_n} = 0.$$

Dazu beobachte zuerst: Es gibt ein $N \in \mathbb{N}$ mit $x_n < 0$ für alle $n \geq N$.^a Für $n \geq N$ gilt dann auch, dass $x_n \log \beta < 0$ (da $\log \beta > \log 1 = 0$). Man folgert aus Bemerkung 6.3.8

$$0 < \beta^{x_n} = \exp(x_n \log \beta) \underset{\text{Lemma 6.3.15(ii)}}{\leq} \frac{1}{1 - x_n \log \beta}$$

Man überzeugt sich leicht, dass $(1 - x_n \log \beta)_{n \in \mathbb{N}}$ bestimmt gegen ∞ divergiert. Wegen Beispiel 6.1.20 folgt dann $\lim_{n \rightarrow \infty} \frac{1}{1 - x_n \log \beta} = 0$. Das Sandwichlemma für Folgen impliziert

$$\lim_{n \rightarrow \infty} \beta^{x_n} = 0$$

Zweiter Grenzwert. Der andere Grenzwert kann ähnlich hergeleitet werden (diesmal mit Lemma 6.3.15(i) anstatt Lemma 6.3.15(ii)). Wir überlassen den Lesenden Details als Übungsaufgabe.

Zu (ii). Kann leicht aus (i) hergeleitet werden: Ist $\gamma \in (0, 1)$, so ist $\beta := \frac{1}{\gamma} > 1$. Den Rest überlassen wir den Lesenden als Übungsaufgabe.

Zu (iii). *Zweiter Grenzwert.* Zuerst zu $\lim_{x \rightarrow \infty} \log(x) = \infty$. Es sei $(x_n)_{n \in \mathbb{N}} \subseteq (0, \infty)$ mit „ $\lim_{n \rightarrow \infty} x_n = \infty$ “. Zu zeigen ist, dass „ $\lim_{n \rightarrow \infty} \log(x_n) = \infty$ “, d.h. (vgl. Definition 2.3.35)

$$\mathbb{Z}_L : \quad \forall L > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \log(x_n) \geq L \quad . \quad (308)$$

Dazu: Sei $L > 0$. Setze $M := e^L > 0$. Da „ $\lim_{n \rightarrow \infty} x_n = \infty$ “ gilt

$$\exists n_1 \in \mathbb{N} \quad \forall n \geq n_1 : \quad x_n \geq M = e^L.$$

Da $\log$ streng monoton wachsend ist gilt für $n \geq n_1$ auch

$$\log(x_n) \geq \log(e^L) = \log(\exp(L)) = L.$$

Wählt man $n_0 := n_1$ so folgt (308) direkt.

Erster Grenzwert. Die nachfolgende Gleichungskette (von der man eigentlich jeden Schritt noch als Übungsaufgabe begründen müsste) liefert

$$\lim_{y \rightarrow 0} \log(y) = \lim_{x \rightarrow \infty} \log\left(\frac{1}{x}\right) \stackrel{\text{Kor6.3.5}}{=} \lim_{x \rightarrow \infty} (-\log(x)) = -\infty.$$

Zu (iv). Sei $\alpha' := \alpha + 1$. Dann gilt nach Proposition 6.3.10 (v)

$$\beta^x = \left(\beta^{\frac{x}{\alpha'}}\right)^{\alpha'} = \left(\exp\left(\frac{x}{\alpha'} \log \beta\right)\right)^{\alpha'} \stackrel{\text{Prop6.3.13(i)}}{\geq} \left(1 + \frac{x}{\alpha'} \log \beta\right)^{\alpha'} = \left(1 + \frac{x}{\alpha'} \log \beta\right) \left(1 + \frac{x}{\alpha'} \log \beta\right)^{\alpha}$$

Dann gilt für $x \in (0, \infty)$ auch

$$\begin{aligned} 0 \leq \frac{x^\alpha}{\beta^x} &\leq \frac{x^\alpha}{\left(1 + \frac{x}{\alpha'} \log \beta\right) \left(1 + \frac{x}{\alpha'} \log \beta\right)^\alpha} = \frac{1}{1 + \frac{x}{\alpha'} \log \beta} \left(\frac{x}{1 + \frac{x}{\alpha'} \log \beta}\right)^\alpha \\ &\stackrel{\text{Prop6.3.13(i)}}{\leq} \frac{1}{1 + \frac{x}{\alpha'} \log \beta} \left(\frac{1}{\frac{1}{x} + \frac{1}{\alpha'} \log \beta}\right)^\alpha \leq \frac{1}{1 + \frac{x}{\alpha'} \log \beta} \underbrace{\left(\frac{1}{\frac{1}{\alpha'} \log \beta}\right)^\alpha}_{=: c} = \frac{c}{1 + \frac{x}{\alpha'} \log \beta} \end{aligned}$$

Man überzeugt sich leicht, dass $\lim_{x \rightarrow \infty} \frac{1}{1 + \frac{x}{\alpha'} \log \beta} = 0$. Wie beim Sandwichlemma folgt dann aus der obigen Ungleichungskette

$$\lim_{x \rightarrow \infty} \frac{x^\alpha}{\beta^x} = 0.$$

Zu (v). Wähle $\eta \in (0, \frac{\alpha}{\beta})$. Dann gilt für $x \in (1, \infty)$

$$\log(x) = \log(x^1) = \log\left((x^\eta)^{\frac{1}{\eta}}\right)$$

$$\stackrel{\text{Bem6.3.9}}{=} \frac{1}{\eta} \log(x^\eta) \leq \frac{1}{\eta} \log(x^\eta) \stackrel{\text{Lemma6.3.15}}{\leq} \frac{1}{\eta} (x^\eta - 1) \leq \frac{1}{\eta} x^\eta.$$

Deswegen gilt für $x \in (1, \infty)$ (d.h. auch $\log(x) > \log(1) = 0$), dass

$$0 \leq \frac{\log(x)^\beta}{x^\alpha} \leq \frac{(\frac{1}{\eta} x^\eta)^\beta}{x^\alpha} = \left(\frac{1}{\eta}\right)^\beta \frac{x^{\eta\beta}}{x^\alpha} = \left(\frac{1}{\eta}\right)^\beta x^{\eta\beta - \alpha}. \quad (309)$$

Beachte: $\eta\beta - \alpha < \frac{\alpha}{\beta}\beta - \alpha = 0$. Man folgert

$$\lim_{x \rightarrow \infty} x^{\eta\beta - \alpha} = \lim_{x \rightarrow \infty} \left(\frac{1}{\eta}\right)^{\alpha - \eta\beta} \stackrel{\text{Bsp6.1.21}}{=} \lim_{y \rightarrow 0} y^{\alpha - \eta\beta} \stackrel{\text{Prop6.3.13(ii)}}{=} 0.$$

Wieder kann man dank (309) und der vorigen Gleichung wie im Sandwichlemma vorgehen. Man folgert

$$\lim_{x \rightarrow \infty} \frac{\log(x)^\beta}{x^\alpha} = 0.$$

Zu (vi). Sei $x \in (0, \infty)$. Definiere $y := \frac{1}{x} \in (0, \infty)$. Das heißt auch $x = \frac{1}{y}$. berechnen wir

$$x^\alpha \log(x)^n = \left(\frac{1}{y}\right)^\alpha \log\left(\frac{1}{y}\right)^n \stackrel{\text{Kor6.3.5}}{=} \frac{1}{y^\alpha} (-\log(y))^n = (-1)^n \frac{\log(y)^n}{y^\alpha}.$$

Man überzeugt sich aufgrunddessen leicht davon, dass

$$\lim_{x \rightarrow 0} x^\alpha \log(x)^n = \lim_{y \rightarrow \infty} (-1)^n \frac{\log(y)^n}{y^\alpha} \stackrel{(v)}{=} (-1)^n \cdot 0 = 0.$$

^aDies können Sie als Übungsaufgabe beweisen! Man würde es sich wohl wie folgt klar machen: Da $(x_n)_{n \in \mathbb{N}}$ bestimmt gegen $-\infty$ divergiert gibt es nach Definition 2.3.35 zu jedem $K > 0$ einen Index $n_0 \in \mathbb{N}$ so, dass $\forall n \geq n_0 : x_n \leq -K$. Betrachtet man nun $K = 1$ so gibt es also einen Index $N \in \mathbb{N}$ so, dass $\forall n \geq N : x_n \leq -1$. Für $n \geq N$ gilt dann sicherlich auch $x_n < 0$.

Bemerkung 6.3.17

Die Aussage von Proposition 6.3.16(iv) gilt eigentlich auch für $\alpha < 0$. In der Tat: Für $\alpha < 0$ und $x \in (0, \infty)$ ist $\beta^x \geq \beta^0 = 1$. Daher gilt

$$0 \leq \frac{x^\alpha}{\beta^x} \leq x^\alpha = \left(\frac{1}{x}\right)^{-\alpha} = \left(\frac{1}{x}\right)^{\tilde{\alpha}} \quad \text{mit } \tilde{\alpha} := -\alpha > 0. \quad (310)$$

Nun gilt

$$\lim_{x \rightarrow \infty} \left(\frac{1}{x}\right)^{\tilde{\alpha}} \stackrel{\text{Bsp6.1.21}}{=} \lim_{y \rightarrow 0} y^{\tilde{\alpha}} \stackrel{\text{Prop6.3.13(iii)}}{=} 0.$$

Daraus und aus (310) folgert man wie im Sandwichlemma

$$\lim_{x \rightarrow \infty} \frac{x^\alpha}{\beta^x} = 0.$$

Beispiel 6.3.1. Es gilt für alle $\alpha \in \mathbb{R}$

$$\lim_{\substack{x \rightarrow 0 \\ x \in (0, \infty)}} x^\alpha e^{-\frac{1}{x^2}} = 0.$$

Beweis.

FALL 1. $\alpha > 0$.

Dann gilt (weil $e^{-\frac{1}{x^2}} \leq e^0 = 1$ für alle $x \in (0, \infty)$)

$$0 \leq x^\alpha e^{-\frac{1}{x^2}} \leq x^\alpha,$$

Da nach Proposition 6.3.13 gilt $\lim_{x \rightarrow 0} x^\alpha = 0$ folgt nach dem Sandwichlemma für Funktionsgrenzwerte, dass

$$\lim_{x \rightarrow 0} x^\alpha e^{-\frac{1}{x^2}} = 0.$$

FALL 2. $\alpha \leq 0$.

Wir gehen vor wie im Beweis von Proposition 6.3.16(vi) Definiere $y := \frac{1}{x}$ und beobachte, dass

$$x^\alpha e^{-\frac{1}{x^2}} = \left(\frac{1}{y}\right)^\alpha e^{-y^2} = \frac{e^{-y^2}}{y^\alpha} = y^{-\alpha} e^{-y^2}. \quad (311)$$

Anstatt $\lim_{x \rightarrow 0} x^\alpha e^{-\frac{1}{x^2}}$ berechnen wir nun also $\lim_{y \rightarrow \infty} y^{-\alpha} e^{-y^2}$. Für $y \in (1, \infty)$ gilt nun $e^{-y^2} \leq e^{-y}$. Somit gilt auch

$$0 \leq y^{-\alpha} e^{-y^2} \leq y^{-\alpha} e^{-y} = \frac{y^{-\alpha}}{e^y} \quad (312)$$

Da $\lim_{y \rightarrow \infty} \frac{y^{-\alpha}}{e^y} = 0$ (nach Proposition 6.3.16(iv), denn $-\alpha \in [0, \infty)$) folgert man wie im Sandwichlemma aus (312), dass

$$\lim_{y \rightarrow \infty} y^{-\alpha} e^{-y^2} = 0.$$

Dank (311) gilt dann auch

$$\lim_{x \rightarrow 0} x^\alpha e^{-\frac{1}{x^2}} = 0.$$

6.4 Gleichmäßige Stetigkeit

Gegeben sei für $D \subseteq \mathbb{K}_1$ eine Funktion $f : D \rightarrow \mathbb{K}_2$. Wir erinnern uns an das Epsilon-Delta-Kriterium für Stetigkeit (Bemerkung 6.2.4). Eine Funktion ist bei einem $x_0 \in D$ stetig genau dann wenn

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x \in D \text{ mit } |x - x_0| < \delta \text{ gilt } |f(x) - f(x_0)| < \varepsilon. \quad (313)$$

Stellen wir uns nun mal vor, wir wollen die Stetigkeit bei einem Punkt $x_0 \in D$ mit dem Epsilon-Delta-Kriterium nachweisen. Zu gegebenem $\varepsilon > 0$ müssen wir also ein (von ε abhängiges) $\delta > 0$ finden so, dass obiges gilt. Dieses δ kann (und wird im Allgemeinen auch) von dem fixierten Punkt x_0 abhängen (vgl. dazu auch Beweis 2 in 6.1.25).

Nur zu gern würde man ein $\delta > 0$ wählen, welches *unabhängig von x_0 ist* und somit für alle Punkte $x_0 \in D$ gleichermaßen verwendet werden kann. Falls das möglich ist, nennen wir eine Funktion *gleichmäßig stetig*.

Definition 6.4.1 — GLEICHMÄSSIGE STETIGKEIT

Es sei $D \subseteq \mathbb{K}_1$. Eine Funktion $f : D \rightarrow \mathbb{K}_2$ heißt auf D *gleichmäßig stetig*^a wenn

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x, x' \in D \text{ mit } |x - x'| < \delta \text{ gilt } |f(x) - f(x')| < \varepsilon.$$

^aEine schöne Illustration dieses Konzeptes bietet mal wieder GEOGEBRA.

Die Variable x' hätten wir in der obigen Definition genausogut x_0 nennen können. Dann steht dort dasselbe wie in (313) nur eben mit der Zusatz-Bedingung, dass das oben gewählte $\delta > 0$ für alle $x_0 \in D$ verwendbar sein muss. Dies zeigt insbesondere

Jede gleichmäßig stetige Funktion ist auch stetig.

Die Rückrichtung gilt im Allgemeinen nicht, wie wir noch sehen werden. Jedoch: Im Fall, dass $D = [a, b] \subseteq \mathbb{R}$ ein kompaktes Intervall ist, besteht gar kein Unterschied zwischen Stetigkeit und gleichmäßiger Stetigkeit!

Satz 6.4.2 — SATZ VON HEINE

Es seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a, b] \rightarrow \mathbb{K}$ eine stetige Funktion. Dann ist f auf $[a, b]$ auch gleichmäßig stetig.

Beweis. Angenommen, f ist nicht gleichmäßig stetig. Dann gilt also

$$\neg (\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x, x' \in D \text{ mit } |x - x'| < \delta \text{ gilt } |f(x) - f(x')| < \varepsilon).$$

Verneinung mit Quantoren ergibt (vgl. Überlegung 2.2.4 und auch die Diskussion um

(272))

$$\exists \varepsilon_0 > 0 : \quad \forall \delta > 0 \quad \exists x_\delta, x'_\delta \in D \text{ mit } |x_\delta - x'_\delta| < \delta \text{ und } |f(x_\delta) - f(x'_\delta)| \geq \varepsilon_0.$$

Betrachte nun für $n \in \mathbb{N}$ den Fall $\delta = \frac{1}{n}$ und wähle $\tilde{x}_n := x_{\frac{1}{n}}$ und $\tilde{x}'_n := x'_{\frac{1}{n}}$. Insbesondere gilt für alle $n \in \mathbb{N}$

$$|\tilde{x}_n - \tilde{x}'_n| < \frac{1}{n} \quad \text{und} \quad |f(\tilde{x}_n) - f(\tilde{x}'_n)| \geq \varepsilon_0. \quad (314)$$

Betrachte nun $(\tilde{x}_n)_{n \in \mathbb{N}} \subseteq [a, b]$ und $(\tilde{x}'_n)_{n \in \mathbb{N}} \subseteq [a, b]$ als Folgen in $[a, b]$. Nach Lemma 6.2.11 gibt es eine konvergente Teilfolge $(\tilde{x}_{\ell_n})_{n \in \mathbb{N}}$ von $(\tilde{x}_n)_{n \in \mathbb{N}}$ mit

$$x_1 := \lim_{n \rightarrow \infty} \tilde{x}_{\ell_n} \in [a, b]. \quad (315)$$

Wendet man nun wiederum Lemma 6.2.11 auf die Folge $(\tilde{x}'_{\ell_n})_{n \in \mathbb{N}} \subseteq [a, b]$ an, so gibt es eine weitere Teilfolge $(\tilde{x}'_{\ell_{k_n}})_{n \in \mathbb{N}}$ mit

$$x_2 := \lim_{n \rightarrow \infty} \tilde{x}'_{\ell_{k_n}} \in [a, b]. \quad (316)$$

ZWISCHENBEHAUPTUNG $x_1 = x_2$.

Dazu beachte, dass nach Proposition 3.5.13 auch $x_1 = \lim_{n \rightarrow \infty} \tilde{x}_{\ell_{k_n}}$. Also gilt

$$|x_1 - x_2| = \left| \lim_{n \rightarrow \infty} \tilde{x}_{\ell_{k_n}} - \lim_{n \rightarrow \infty} \tilde{x}'_{\ell_{k_n}} \right| \stackrel{\text{GWR}}{=} \left| \lim_{n \rightarrow \infty} (\tilde{x}_{\ell_{k_n}} - \tilde{x}'_{\ell_{k_n}}) \right| \stackrel{\text{Bem 6.2.2}}{=} \lim_{n \rightarrow \infty} |\tilde{x}_{\ell_{k_n}} - \tilde{x}'_{\ell_{k_n}}| \quad (317)$$

Nun gilt, da $0 \leq |\tilde{x}_{\ell_{k_n}} - \tilde{x}'_{\ell_{k_n}}| \stackrel{(314)}{\leq} \frac{1}{\ell_{k_n}} \stackrel{\text{Bem 3.5.12}}{\leq} \frac{1}{k_n} \stackrel{\text{Bem 3.5.12}}{\leq} \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0$ auch

$$\lim_{n \rightarrow \infty} |\tilde{x}_{\ell_{k_n}} - \tilde{x}'_{\ell_{k_n}}| = 0,$$

vgl. Sandwichlemma. Aus (317) folgt dann

$$|x_1 - x_2| = 0 \quad \text{d.h.} \quad x_1 = x_2.$$

ZWISCHENBEHAUPTUNG ENDE.

Nennen wir nun $x_0 := x_1 [= x_2]$. Das heißt nach (315) und (316)

$$\lim_{n \rightarrow \infty} \tilde{x}_{\ell_{k_n}} = \lim_{n \rightarrow \infty} \tilde{x}'_{\ell_{k_n}} = x_0 \in [a, b].$$

Da f bei x_0 stetig ist, folgt auch

$$\lim_{n \rightarrow \infty} f(\tilde{x}_{\ell_{k_n}}) = \lim_{n \rightarrow \infty} f(\tilde{x}'_{\ell_{k_n}}) = f(x_0).$$

Man folgert aus den Grenzwertregeln, dass

$$\lim_{n \rightarrow \infty} [f(\tilde{x}_{\ell_{k_n}}) - f(\tilde{x}'_{\ell_{k_n}})] = \lim_{n \rightarrow \infty} f(\tilde{x}_{\ell_{k_n}}) - \lim_{n \rightarrow \infty} f(\tilde{x}'_{\ell_{k_n}}) = f(x_0) - f(x_0) = 0.$$

Somit gilt auch (wegen der Stetigkeit der Betragsfunktion)

$$\lim_{n \rightarrow \infty} |f(\tilde{x}_{\ell_{k_n}}) - f(\tilde{x}'_{\ell_{k_n}})| = 0.$$

Gleichzeitig folgt aber aus (314), dass

$$|f(\tilde{x}_{\ell_{k_n}}) - f(\tilde{x}'_{\ell_{k_n}})| \geq \varepsilon_0 \quad \forall n \in \mathbb{N}.$$

Ein Widerspruch, denn dann hat man

$$0 = \lim_{n \rightarrow \infty} \underbrace{|f(\tilde{x}_{\ell_{k_n}}) - f(\tilde{x}'_{\ell_{k_n}})|}_{\geq \varepsilon_0} \geq \varepsilon_0 > 0.$$

Die Behauptung ist gezeigt.

Innerhalb eines kompakten Intervalles der Form $[a, b]$ kann also das δ in (313) stets unabhängig von x_0 gewählt werden. Von ε hängt es jedoch im Allgemeinen trotzdem noch ab!

Auf die Kompaktheit des Intervalles kann nicht verzichtet werden, wie wir an folgendem Beispiel sehen wollen:

Beispiel 6.4.3

Die Funktion $f : (0, 1] \rightarrow \mathbb{R}, x \mapsto \frac{1}{x}$ ist auf $(0, 1]$ stetig, aber nicht gleichmäßig stetig.

Beweis. Die Stetigkeit von f auf $(0, 1]$ folgt direkt aus Proposition 6.2.6 oder auch aus Beispiel 6.1.25. Angenommen nun f wäre gleichmäßig stetig. Dann gilt (mit $\varepsilon = 1$)

$$\exists \delta > 0 : \quad \forall x, x' \in D \text{ mit } |x - x'| < \delta \text{ gilt } |f(x) - f(x')| < 1. \quad (318)$$

Wähle nun ein $n \in \mathbb{N}$ mit $\frac{1}{n} < \delta$.^a Dann gilt für $x := \frac{1}{n}$ und $x' = \frac{1}{2n}$, dass $|x - x'| = \frac{1}{n} - \frac{1}{2n} = \frac{1}{2n} < \delta$. Jedoch ist

$$|f(x) - f(x')| = |f(\frac{1}{n}) - f(\frac{1}{2n})| = |n - 2n| = |-n| = n \geq 1.$$

Ein Widerspruch zu (318).

^aEin solches $n \in \mathbb{N}$ existiert da $\frac{1}{j} \xrightarrow{j \rightarrow \infty} 0$.

7 Differentialrechnung

29.01.2026 In der Differentialrechnung geht es zunächst darum, eine Funktion $f : D \rightarrow \mathbb{R}$ in einer kleinen Umgebung von einem Punkt $x_0 \in D$ durch eine (*affin-*)*lineare Funktion* zu *approximieren*. Eine gute lineare Approximation des Graphen Γ_f ist gerade die *Tangente* von Γ_f bei $(x_0, f(x_0))$. Deren Steigung ist durch die *Ableitung* gegeben, welche wir im Folgenden definieren wollen. Wir werden das Konzept der Ableitung erst definieren und danach die oben umrissene geometrische Anschauung diskutieren.

In diesem Abschnitt seien wieder $\mathbb{K}_1, \mathbb{K}_2 \in \{\mathbb{R}, \mathbb{C}\}$.

Definition 7.0.1 — DIFFERENZIERBARKEIT

Es sei $D \subseteq \mathbb{K}_1$ und $f : D \rightarrow \mathbb{K}_2$ eine Funktion. Ferner sei $x_0 \in D$. Wir nennen f bei x_0 *differenzierbar*, falls x_0 ein Häufungspunkt von $D \setminus \{x_0\}$ ist und der Grenzwert

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

in $\mathbb{C}$ existiert. ^a In diesem Fall nennen wir den obigen Grenzwert *Ableitung von f in x₀* und schreiben dafür $f'(x_0)$. Alternative Notationen sind $\frac{df}{dx}(x_0)$, $\frac{d}{dx}f(x)|_{x=x_0}$ oder $Df(x_0)$.

Wir nennen f auf D differenzierbar, falls f bei jedem Punkt $x_0 \in D$ differenzierbar ist.

^aGenauer müssten wir hier wohl schreiben

$$\lim_{\substack{x \rightarrow x_0 \\ x \in D \setminus \{x_0\}}} \frac{f(x) - f(x_0)}{x - x_0},$$

bzw. noch genauer $\lim_{x \rightarrow x_0} g(x)$, wobei $g : D \setminus \{x_0\} \rightarrow \mathbb{C}, x \mapsto \frac{f(x) - f(x_0)}{x - x_0}$. Achtung: $\mathbb{C}$ als Bildbereich von g ist kein Schreibfehler — ist einer der beiden Körper $\mathbb{K}_1$ oder $\mathbb{K}_2$ gleich $\mathbb{C}$, so ist die Abbildungsvorschrift nur als Quotient komplexer Zahlen zu verstehen. Im Spezialfall $\mathbb{K}_1 = \mathbb{K}_2 = \mathbb{R}$ kann der Grenzwert als Grenzwert in $\mathbb{R}$ aufgefasst werden.

Im Folgenden betrachten wir häufig den Spezialfall $\mathbb{K}_1 = \mathbb{K}_2 = \mathbb{R}$. Im Fall $\mathbb{K}_1 = \mathbb{K}_2 = \mathbb{C}$ bringt obige Definition in der Tat sehr wundersame Phänomene hervor, welche man in der Vorlesung FUNKTIONENTHEORIE studiert.

Als nächstes sehen wir einige Beispiele differenzierbarer Funktionen.

Beispiel 7.0.2

Es sei $c \in \mathbb{K}_2$. Dann ist die *konstante Funktion* $f : \mathbb{K}_1 \rightarrow \mathbb{K}_2, x \mapsto c$ auf $\mathbb{K}_1$ differenzierbar.

Beweis. Sei $x_0 \in \mathbb{K}_1$. Für $x \in \mathbb{K}_1 \setminus \{x_0\}$ gilt

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{c - c}{x - x_0} = \frac{0}{x - x_0} = 0.$$

Wir erhalten

$$\lim_{\substack{x \rightarrow x_0 \\ x \in \mathbb{K}_1 \setminus \{x_0\}}} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{\substack{x \rightarrow x_0 \\ x \in \mathbb{K}_1 \setminus \{x_0\}}} 0 = 0. \quad (319)$$

Also ist f bei x_0 differenzierbar mit $f'(x_0) = 0$.^a

^aMan beachte: Wir müssten hier, wenn wir streng nach Definition 7.0.1 vorgehen, auch noch überprüfen, ob x_0 ein Häufungspunkt von $\mathbb{K}_1 \setminus \{x_0\}$ ist — denn ansonsten wäre das Bilden des Grenzwertes in (319) nicht zulässig. Dass x_0 ein Häufungspunkt von $\mathbb{K}_1 \setminus \{x_0\}$ folgt aber direkt, wenn man in Definition 6.1.1 mit der Folge $(x_n)_{n \in \mathbb{N}} := (x_0 + \frac{1}{n})_{n \in \mathbb{N}} \subseteq \mathbb{K}_1 \setminus \{x_0\}$ betrachtet. Wir werden diese Überprüfung in Zukunft meist nicht explizit aufschreiben.

Beispiel 7.0.3

Die *identische Abbildung* $f := \text{id}_{\mathbb{K}} : \mathbb{K} \rightarrow \mathbb{K}, x \mapsto x$ ist auf ganz $\mathbb{K}$ differenzierbar.

Beweis. Sei $x_0 \in \mathbb{K}$. Für $x \in \mathbb{K} \setminus \{x_0\}$ gilt

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x - x_0}{x - x_0} = 1$$

Wir erhalten

$$\lim_{\substack{x \rightarrow x_0 \\ x \in \mathbb{K} \setminus \{x_0\}}} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{\substack{x \rightarrow x_0 \\ x \in \mathbb{K} \setminus \{x_0\}}} 1 = 1.$$

Also ist f bei x_0 differenzierbar mit $f'(x_0) = 1$.

Beispiel 7.0.4

Die *Quadratfunktion* $f : \mathbb{K} \rightarrow \mathbb{K}, x \mapsto x^2$ ist auf ganz $\mathbb{K}$ differenzierbar.

Beweis. Sei $x_0 \in \mathbb{K}$. Für $x \in \mathbb{K} \setminus \{x_0\}$ gilt

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{x^2 - x_0^2}{x - x_0} = \frac{(x - x_0)(x + x_0)}{x - x_0} = x + x_0.$$

Wir erhalten

$$\lim_{\substack{x \rightarrow x_0 \\ x \in \mathbb{K} \setminus \{x_0\}}} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{\substack{x \rightarrow x_0 \\ x \in \mathbb{K} \setminus \{x_0\}}} (x + x_0) = \lim_{x \rightarrow x_0} x + \lim_{x \rightarrow x_0} x_0 = 2x_0.$$

Also ist f bei x_0 differenzierbar mit $f'(x_0) = 2x_0$.

Nun einmal ein Beispiel einer nicht-differenzierbaren Funktion.

Beispiel 7.0.5

Der reelle Absolutbetrag $\text{abs} : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto |x|$ ist bei $x_0 = 0$ nicht differenzierbar.

Beweis. Für $x \in \mathbb{R} \setminus \{0\}$ gilt

$$\frac{|x| - |0|}{x - 0} = \frac{|x|}{x}$$

Beachte: für $x \in (0, \infty)$ gilt $\frac{|x|}{x} = \frac{x}{x} = 1$. Hingegen gilt für $x \in (-\infty, 0)$, dass $\frac{|x|}{x} = \frac{-x}{x} = -1$. Mit anderen Worten

$$\frac{|x|}{x} = \begin{cases} 1 & \text{falls } x > 0 \\ -1 & \text{falls } x < 0. \end{cases}$$

Wir erhalten

$$\lim_{x \rightarrow 0+} \frac{|x|}{x} = \lim_{x \rightarrow 0+} 1 = 1 \quad ^a$$

und

$$\lim_{x \rightarrow 0-} \frac{|x|}{x} = \lim_{x \rightarrow 0-} -1 = -1$$

Die einseitigen Grenzwerte stimmen also nicht überein. Nach Proposition 6.1.16 kann dann

$$\lim_{x \rightarrow 0} \frac{|x| - |0|}{x - 0} = \lim_{x \rightarrow 0} \frac{|x|}{x}$$

nicht existieren. Es folgt, dass abs in 0 nicht differenzierbar ist.

^aBeachte: Der Definitionsbereich vom Rechts-Cut ist hier gerade $\mathbb{R} \setminus \{0\} \cap [0, \infty) = (0, \infty)$. Analog ist der Definitionsbereich vom Links-Cut gerade $(-\infty, 0)$. Es handelt sich daher wirklich um *nicht-punktierte* Grenzwerte, wie hier behauptet. Somit ist Proposition 6.1.16 wirklich anwendbar.

Zunächst zur geometrischen Interpretation der Ableitung als *Tangentensteigung* (für $\mathbb{K}_1 = \mathbb{K}_2 = \mathbb{R}$). Für $x \in D$ mit $x \neq x_0$ kann man eine *Sekante*, d.h. eine gerade Verbindungsstrecke zwischen dem Punkt $(x, f(x))$ und $(x_0, f(x_0))$ betrachten. Diese ist ein Funktionsgraph, nämlich von der affin-linearen Funktion

$$\ell_x : \mathbb{R} \rightarrow \mathbb{R}, \quad t \mapsto f(x_0) + \frac{f(x) - f(x_0)}{x - x_0}(t - x_0). \quad (320)$$

In der Tat: Man überzeugt sich leicht, dass $\ell_x(x) = f(x)$ und $\ell_x(x_0) = f(x_0)$. Da ℓ_x affin-linear ist muss es sich bei Γ_{ℓ_x} also gerade um die Verbindungsstrecke zwischen $(x, f(x))$ und $(x_0, f(x_0))$ handeln.

Lässt man $x \rightarrow x_0$ streben, so werden die Sekanten zu Tangenten, siehe Abbildung 27 ⁵. Betrachte nun also für $t \in \mathbb{R}$

$$\lim_{x \rightarrow x_0} \ell_x(t) = \lim_{x \rightarrow x_0} \left(f(x_0) + \frac{f(x) - f(x_0)}{x - x_0}(t - x_0) \right) = f(x_0) + f'(x_0)(t - x_0).$$

⁵Diese Abbildung gibt es auch nochmal interaktiv auf GEOGEBRA

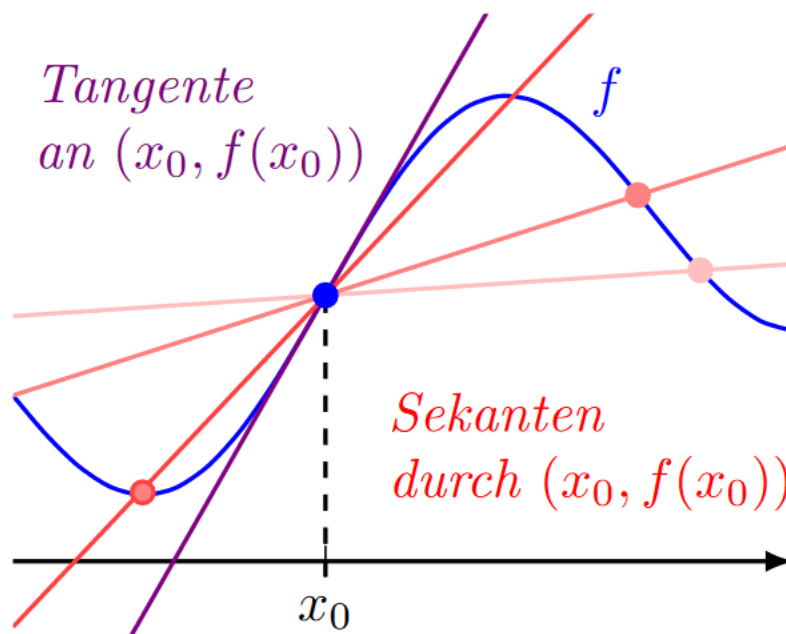


Abbildung 27: In Blau: Der Graph Γ_f . In Rottönen: Sekanten Γ_{ℓ_x} .

Somit ist die Tangente an Γ_f bei $(x_0, f(x_0))$ gegeben durch

$$t \mapsto f(x_0) + f'(x_0)(t - x_0).$$

Dies motiviert die folgende äquivalente Formulierung der Differenzierbarkeit als *Linearisierbarkeit*.

Proposition 7.0.6 — ABLEITUNG ALS LINEARISIERUNG

Es sei $D \subseteq \mathbb{K}_1$ und $f : D \rightarrow \mathbb{K}_2$. Dann sind für $x_0 \in D$ (einen Häufungspunkt von $D \setminus \{x_0\}$) folgende Aussagen äquivalent

- (a) f ist bei x_0 differenzierbar.
- (b) Es gibt ein $c \in \mathbb{C}$ und eine in x_0 stetige Abbildung $\phi : D \rightarrow \mathbb{C}$ mit^a $\phi(x_0) = 0$ so, dass

$$\forall x \in D : \quad f(x) = f(x_0) + c(x - x_0) + \phi(x)(x - x_0).$$

Trifft eine der beiden Aussagen zu, so ist c eindeutig bestimmt durch $c = f'(x_0)$.

^aFalls $\mathbb{K}_1 = \mathbb{K}_2 = \mathbb{R}$ so sind hier ϕ und c reellwertig, d.h. $c \in \mathbb{R}$ und $\phi : D \rightarrow \mathbb{R}$.

Wir nennen $x \mapsto \phi(x)(x - x_0)$ den *Linearisierungsfehler* von f um x_0 . Da ϕ nah an x_0 sehr klein ist, ist dieser Fehler *asymptotisch kleiner* als jede lineare Funktion.

Beweis. Zu „(b) $\Rightarrow$ (a)“. Es seien c und ϕ wie in der Aussage (b). Dann gilt also

$$\forall x \in D : \quad f(x) = f(x_0) + c(x - x_0) + \phi(x)(x - x_0).$$

Mit dieser Formel berechnen wir für $x \in D \setminus \{x_0\}$

$$\begin{aligned} \frac{f(x) - f(x_0)}{x - x_0} &= \frac{[f(x_0) + c(x - x_0) + \phi(x)(x - x_0)] - f(x_0)}{x - x_0} \\ &= \frac{c(x - x_0) + \phi(x)(x - x_0)}{x - x_0} = c + \phi(x). \end{aligned}$$

Es folgt

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} [c + \phi(x)] = c + \lim_{x \rightarrow x_0} \phi(x) \stackrel[\text{Definition 6.2.1}]{\phi \text{ stetig bei } x_0} c + \phi(x_0) \stackrel{\phi(x_0) = 0}{=} c.$$

Gemäß Definition 7.0.1 ist dann f bei x_0 differenzierbar und $f'(x_0) = c$.

Zu „(a) $\Rightarrow$ (b)“. Es sei f bei x_0 differenzierbar. Definiere $c := f'(x_0)$ und $\phi : D \rightarrow \mathbb{C}$ durch

$$\phi(x) := \begin{cases} \frac{f(x) - f(x_0)}{x - x_0} - c & x \neq x_0 \\ 0 & x = x_0 \end{cases} \quad (321)$$

Dann gilt für alle $x \in D \setminus \{x_0\}$

$$\begin{aligned} f(x) &= f(x_0) + f(x) - f(x_0) = f(x_0) + \frac{f(x) - f(x_0)}{x - x_0}(x - x_0) \\ &= f(x_0) + [\phi(x) + c](x - x_0) = f(x_0) + c(x - x_0) + \phi(x)(x - x_0) \end{aligned}$$

Dieselbe Gleichung gilt auch noch, wenn man $x = x_0$ einsetzt, wie man leicht durch Einsetzen von x_0 überprüft. Somit gilt sie also für alle $x \in D$.

$$\text{Noch } \underline{Z}_1 : \quad \phi \text{ ist in } x_0 \text{ stetig und } \phi(x_0) = 0.$$

Dass $\phi(x_0) = 0$ gilt, folgt direkt aus (321). Nun also noch zur Stetigkeit von ϕ in x_0 .

$$\underline{Z}_2 : \quad \lim_{x \rightarrow x_0} \phi(x) = \phi(x_0) [= 0].$$

Am liebsten würden wir in (321) nur den ersten Fall betrachten. Dann hätte man leicht

$$\lim_{x \rightarrow x_0} \phi(x) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} - c \stackrel{\text{Def 7.0.1}}{=} f'(x_0) - c = f'(x_0) - f'(x_0) = 0.$$

Man kann aber für Grenzwerte nicht einfach nur einen Fall betrachten.^a Daher müssen wir an der Stelle mehr machen. Es sei nun $(x_n)_{n \in \mathbb{N}} \subseteq D$ eine beliebige Folge mit $\lim_{n \rightarrow \infty} x_n = x_0$.

$$\mathbb{Z}: \quad \lim_{n \rightarrow \infty} \phi(x_n) = 0.$$

Da x_0 auch ein Häufungspunkt von $D \setminus \{x_0\}$ ist gibt es eine (eventuell andere) Folge $(y_n)_{n \in \mathbb{N}} \subseteq D \setminus \{x_0\}$ mit $\lim_{n \rightarrow \infty} y_n = x_0$. Definiere nun für $n \in \mathbb{N}$

$$\tilde{x}_n := \begin{cases} x_n & \text{falls } x_n \neq x_0 \\ y_n & \text{falls } x_n = x_0 \end{cases} \quad (322)$$

Insbesondere gilt $\tilde{x}_n \neq x_0$ für alle $n \in \mathbb{N}$.

$$\text{Behauptung 1:} \quad \tilde{x}_n \xrightarrow{n \rightarrow \infty} x_0$$

Beachte zuerst: Da $(x_n)_{n \in \mathbb{N}}$ und $(y_n)_{n \in \mathbb{N}}$ beide gegen x_0 konvergieren gilt $\lim_{n \rightarrow \infty} |x_n - x_0| = \lim_{n \rightarrow \infty} |y_n - x_0| = 0$. Sei nun $n \in \mathbb{N}$. Da $\tilde{x}_n$ entweder mit x_n oder y_n übereinstimmt gilt

$$|\tilde{x}_n - x_0| \leq \max\{|x_n - x_0|, |y_n - x_0|\} \leq |x_n - x_0| + |y_n - x_0|.$$

Aus diesem Grund gilt

$$0 \leq |\tilde{x}_n - x_0| \leq |x_n - x_0| + |y_n - x_0| \xrightarrow{n \rightarrow \infty} 0 + 0 = 0.$$

Also (nach Sandwichlemma) $\lim_{n \rightarrow \infty} |\tilde{x}_n - x_0| = 0$. Das zeigt auch $\lim_{n \rightarrow \infty} x_n - x_0 = 0$ (vgl. Proposition 2.3.30), also $\lim_{n \rightarrow \infty} x_n = x_0$.

Behauptung 1 Ende

Behauptung 2: Es gilt für alle $n \in \mathbb{N}$, dass $|\phi(x_n)| \leq |\phi(\tilde{x}_n)|$!

Dazu: Im ersten Fall von (322) gilt $x_n = \tilde{x}_n$. Damit gilt in diesem Fall sogar $|\phi(x_n)| = |\phi(\tilde{x}_n)|$. Im zweiten Fall von (322) gilt $x_n = x_0$. Somit gilt $|\phi(x_n)| = 0$. Dann gilt klarerweise $|\phi(\tilde{x}_n)| \geq 0 = |\phi(x_n)|$.

Behauptung 2 Ende.

Wir erhalten für $n \in \mathbb{N}$

$$0 \leq |\phi(x_n)| \leq |\phi(\tilde{x}_n)| \stackrel{(321)}{\leq}_{\tilde{x}_n \neq x_0} \left| \frac{f(\tilde{x}_n) - f(x_0)}{\tilde{x}_n - x_0} - c \right|. \quad (323)$$

Nun gilt $\lim_{n \rightarrow \infty} 0 = 0$ und

$$\begin{aligned} \lim_{n \rightarrow \infty} \left| \frac{f(\tilde{x}_n) - f(x_0)}{\tilde{x}_n - x_0} - c \right| &= \left| \lim_{n \rightarrow \infty} \frac{f(\tilde{x}_n) - f(x_0)}{\tilde{x}_n - x_0} - c \right| \\ &= \left| \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} - c \right| = |f'(x_0) - c| = 0. \end{aligned}$$

Nach dem Sandwichlemma folgt aus (323), dass $\lim_{n \rightarrow \infty} |\phi(x_n)| = 0$, also auch

$$\lim_{n \rightarrow \infty} \phi(x_n) = 0.$$

Da $(x_n)_{n \in \mathbb{N}}$ beliebig war, folgt $\lim_{x \rightarrow x_0} \phi(x) = 0 = \phi(x_0)$, was zu zeigen war.

^aEin gutes Beispiel ist hier Beispiel 6.1.6: Betrachtet man jeden der Fälle einzeln so haben die dort stehenden Funktionen Grenzwerte. Insgesamt existiert der Grenzwert aber nicht.

Nun ein sehr wichtiges Korollar: Differenzierbare Funktionen sind auch stetig!

Korollar 7.0.7 — DIFFERENZIERBARKEIT IMPLIZIERT STETIGKEIT

Es sei $D \subseteq \mathbb{K}_1$ und $f : D \rightarrow \mathbb{K}_2$ eine Funktion.

Ist f in $x_0 \in D$ differenzierbar, so ist f in x_0 auch stetig.

Beweis. Es sei f in x_0 differenzierbar. Dann gibt es nach voriger Proposition $c \in \mathbb{C}$ und $\phi : D \rightarrow \mathbb{C}$ stetig in x_0 (mit $\phi(x_0) = 0$) so, dass

$$\forall x \in D : \quad f(x) = f(x_0) + c(x - x_0) + \phi(x)(x - x_0).$$

Nun ist

$$\ell : D \rightarrow \mathbb{C}, \quad x \mapsto f(x_0) + c(x - x_0)$$

als Polynom sicherlich in x_0 stetig. Außerdem ist

$$\psi : D \rightarrow \mathbb{C}, \quad x \mapsto \phi(x)(x - x_0)$$

als Produkt stetiger Funktionen in x_0 stetig, vgl. Proposition 6.2.6. Da $f = \ell + \psi$ folgt auch die Stetigkeit von f in x_0 , vgl. noch einmal Proposition 6.2.6.

Insbesondere sagt (die Kontraposition von) Korollar 7.0.7, dass *unstetige Funktionen* nicht differenzierbar sind. Zum Beispiel ist die Funktion

$$f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \begin{cases} -1 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

an der Stelle $x_0 = 0$ nicht differenzierbar, da sie dort nicht stetig ist.

Bemerkung 7.0.8

Die Rückrichtung in Korollar 7.0.7 gilt nicht. Mit anderen Worten:

Nicht jede stetige Funktion ist automatisch differenzierbar!

In der Tat: Der reelle Absolutbetrag $\text{abs} : \mathbb{R} \rightarrow \mathbb{R}$ ist zwar bei $x_0 = 0$ stetig (vgl. Beispiel 6.2.5), jedoch ist er bei $x_0 = 0$ nicht differenzierbar, vgl. Beispiel 7.0.5.

Beispiel 7.0.9

Es gibt auch stetige Funktionen, die nirgends differenzierbar sind. In $\mathbb{R}$ lässt sich hierfür die *Weierstraß-Funktion*

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) := \sum_{k=0}^{\infty} \frac{2^k \cos(2^k x)}{3^k}.$$

anführen. Diese werden wir zu gegebenem Zeitpunkt diskutieren.

Auf $\mathbb{C}$ gibt es einfachere Beispiele von stetigen Funktionen, die nirgends differenzierbar sind.

Behauptung. $f : \mathbb{C} \rightarrow \mathbb{C}, z \mapsto \text{Re}(z)$ ist auf $\mathbb{C}$ stetig, aber an keinem Punkt $z_0 \in \mathbb{C}$ diff'bar.

Beweis. Es sei $z_0 \in \mathbb{C}$. Dann gilt für $z \in \mathbb{C} \setminus \{z_0\}$

$$\frac{\text{Re}(z) - \text{Re}(z_0)}{z - z_0} = \frac{\text{Re}(z - z_0)}{z - z_0}.$$

Betrachte nun die Folge $(z_n^{(1)})_{n \in \mathbb{N}} = (z_0 + \frac{1}{n})_{n \in \mathbb{N}}$. Dann gilt

$$\frac{\text{Re}(z_n^{(1)}) - \text{Re}(z_0)}{z_n^{(1)} - z_0} = \frac{\text{Re}(z_n^{(1)} - z_0)}{z_n^{(1)} - z_0} = \frac{\text{Re}(\frac{1}{n})}{\frac{1}{n}} = \frac{\frac{1}{n}}{\frac{1}{n}} = 1 \xrightarrow{n \rightarrow \infty} 1.$$

Betrachte als nächstes die Folge $(z_n^{(2)})_{n \in \mathbb{N}} = (z_0 + \frac{i}{n})_{n \in \mathbb{N}}$. Dann gilt

$$\frac{\text{Re}(z_n^{(2)}) - \text{Re}(z_0)}{z_n^{(2)} - z_0} = \frac{\text{Re}(z_n^{(2)} - z_0)}{z_n^{(2)} - z_0} = \frac{\text{Re}(\frac{i}{n})}{\frac{i}{n}} = \frac{0}{\frac{i}{n}} = 0 \xrightarrow{n \rightarrow \infty} 0.$$

Man sieht, dass $\lim_{n \rightarrow \infty} z_n^{(1)} = \lim_{n \rightarrow \infty} z_n^{(2)} = z_0$, aber

$$\lim_{n \rightarrow \infty} \frac{\text{Re}(z_n^{(1)}) - \text{Re}(z_0)}{z_n^{(1)} - z_0} \neq \lim_{n \rightarrow \infty} \frac{\text{Re}(z_n^{(2)}) - \text{Re}(z_0)}{z_n^{(2)} - z_0}.$$

Somit kann $\lim_{z \rightarrow z_0} \frac{\operatorname{Re}(z) - \operatorname{Re}(z_0)}{z - z_0}$ nicht existieren (vgl. Definition 6.1.3).

Es drängt sich nun die Frage auf, was für Funktionen differenzierbar sind. Im Folgenden zeigen wir, dass *Potenzreihen* stets differenzierbare Funktionen erzeugen.

Proposition 7.0.10 — POTENZREIHEN SIND DIFFERENZIERBAR

Es sei $\sum_{k \geq 0} c_k(z - w)^k$ eine gegebene Potenzreihe mit Referenzpunkt $w \in \mathbb{C}$ und Konvergenzradius $R > 0$. Dann ist die Wertefunktion $P : B_R(w) \rightarrow \mathbb{C}, z \mapsto \sum_{k=0}^{\infty} c_k(z - w)^k$ auf ganz $B_R(w)$ differenzierbar und es gilt

$$P'(z_0) = \sum_{k=1}^{\infty} k c_k (z_0 - w)^{k-1} \quad \forall z_0 \in B_R(w).$$

Im Falle $R = \infty$ kann wie immer $B_R(w)$ durch $\mathbb{C}$ ersetzt werden.

Mit anderen Worten gilt für alle $z \in B_R(w)$

$$\left. \frac{d}{dz} \sum_{k=0}^{\infty} c_k (z - w)^k \right|_{z=z_0} = \sum_{k=1}^{\infty} k c_k (z_0 - w)^{k-1}.$$

Das heißt auch: Potenzreihen können *summandenweise* abgeleitet werden. Genauer: Wer eine Potenzreihe differenzieren will, der muss nur für jedes $k \in \mathbb{N}$ den k -ten Summanden mithilfe der aus der Schule gewohnten Formel $\frac{d}{dz}(z^k)|_{z=z_0} = k z^{k-1}$ ableiten. Der nullte Summand fällt einfach weg. Dass das einfach so geht ist schon nichttrivial: Denn sowohl in der Definition einer Reihe als auch in der Definition der Ableitung versteckt sich ein *Grenzwert*. Man hat es also mit zwei Grenzwerten zu tun, deren Interaktion man verstehen muss.

Beweis. Wir zeigen die Behauptung nur im Spezialfall $w = 0$, d.h.

$$P(z) = \sum_{k=0}^{\infty} c_k z^k \quad (z \in B_R(0)).$$

Der allgemeine Fall geht genauso, ist nur etwas mehr Schreibaarbeit. Sei $z_0 \in B_R(0)$. Dann ist für $z \in B_R(0) \setminus \{z_0\}$

$$\begin{aligned} \frac{P(z) - P(z_0)}{z - z_0} &= \frac{1}{z - z_0} \left(\sum_{k=0}^{\infty} c_k z^k - \sum_{k=0}^{\infty} c_k z_0^k \right) = \frac{1}{z - z_0} \sum_{k=0}^{\infty} (c_k z^k - c_k z_0^k) \\ &\stackrel{c_0 z^0 - c_0 z_0^0 = 0}{=} \frac{1}{z - z_0} \sum_{k=1}^{\infty} (c_k z^k - c_k z_0^k) = \frac{1}{z - z_0} \sum_{k=1}^{\infty} c_k (z^k - z_0^k) \\ &= \sum_{k=1}^{\infty} c_k \frac{z^k - z_0^k}{z - z_0} \stackrel{\text{Kor 3.3.21}}{=} \sum_{k=1}^{\infty} c_k \left(\sum_{\ell=0}^{k-1} z^{\ell} z_0^{k-1-\ell} \right) \end{aligned}$$

$$\begin{aligned}
&= \sum_{k=1}^{\infty} c_k \left(\sum_{\ell=0}^{k-1} (z^\ell - z_0^\ell) z_0^{k-1-\ell} \right) \\
&= \sum_{k=1}^{\infty} c_k \left(\sum_{\ell=0}^{k-1} (z^\ell - z_0^\ell) z_0^{k-1-\ell} + \sum_{\ell=0}^{k-1} z_0^\ell z_0^{(k-1-\ell)} \right) \\
&= \sum_{k=1}^{\infty} c_k \left(\sum_{\ell=0}^{k-1} (z^\ell - z_0^\ell) z_0^{k-1-\ell} + \sum_{\ell=0}^{k-1} z_0^{k-1} \right) \\
&= \sum_{k=1}^{\infty} c_k \left(\sum_{\ell=0}^{k-1} (z^\ell - z_0^\ell) z_0^{k-1-\ell} + k z_0^{k-1} \right) \\
&= \underbrace{\sum_{k=1}^{\infty} c_k \left(\sum_{\ell=0}^{k-1} (z^\ell - z_0^\ell) z_0^{k-1-\ell} \right)}_{=: A(z)} + \sum_{k=1}^{\infty} c_k k z_0^{k-1} = A(z) + \sum_{k=1}^{\infty} c_k k z_0^{k-1} \quad (324)
\end{aligned}$$

ZWISCHENBEHAUPTUNG. $\lim_{z \rightarrow z_0} A(z) = 0.$

Beachte: Der Summand für $k = 1$ von $A(z)$ ist gegeben durch

$$\sum_{\ell=0}^0 (z^\ell - z_0^\ell) z_0^{0-\ell} = (z^0 - z_0^0) z_0^0 = (1 - 1) \cdot 1 = 0.$$

Somit kann man die Summe auch erst bei $k = 2$ starten lassen. Außerdem darf in der inneren Summe auch stets erst ab $\ell = 1$ summiert werden, da der Summand für $\ell = 0$ stets verschwindet. Man erhält

$$\begin{aligned}
A(z) &= \sum_{k=2}^{\infty} c_k \left(\sum_{\ell=1}^{k-1} (z^\ell - z_0^\ell) z_0^{k-1-\ell} \right) = (z - z_0) \sum_{k=2}^{\infty} c_k \left(\sum_{\ell=1}^{k-1} \left(\frac{z^\ell - z_0^\ell}{z - z_0} \right) z_0^{k-1-\ell} \right) \\
&\stackrel{\text{Kor 3.3.21}}{=} (z - z_0) \sum_{k=2}^{\infty} c_k \left(\sum_{\ell=1}^{k-1} \left(\sum_{j=0}^{\ell-1} z^j z_0^{\ell-1-j} \right) z_0^{k-1-\ell} \right) \\
&= (z - z_0) \sum_{k=2}^{\infty} c_k \left(\sum_{\ell=1}^{k-1} \sum_{j=0}^{\ell-1} z^j z_0^{\ell-1-j+k-1-\ell} \right) = (z - z_0) \sum_{k=2}^{\infty} c_k \left(\sum_{\ell=1}^{k-1} \sum_{j=0}^{\ell-1} z^j z_0^{k-2-j} \right)
\end{aligned}$$

An dieser Stelle möchten wir kurz darauf hinweisen, dass in obigen Term nur positive Potenzen von z_0 vorkommen, da $k-2-j$ bei jeder möglichen Konstellation von Summanden nichtnegativ ist (In der Tat: $j \leq \ell-1 \leq (k-1)-1 = k-2$). Dies werden wir bei der Abschätzung in (*) unten auch noch benutzen. Nun argumentieren wir endlich, dass $\lim_{z \rightarrow z_0} A(z) = 0$. Da $z_0 \in B_R(0)$ gilt $r := |z_0| < R$. Wähle nun eine Zahl $r' \in (r, R)$. Es sei $(z_n)_{n \in \mathbb{N}} \subseteq B_R(0) \setminus \{z_0\}$ eine Folge mit $\lim_{n \rightarrow \infty} z_n = z_0$. Insbesondere gilt $\lim_{n \rightarrow \infty} |z_n| = |z_0| = r < r'$.

Dann gibt es (nach Aufgabe 1 auf Blatt 6) auch ein $n_0 \in \mathbb{N}$ mit $\forall n \geq n_0 : |z_n| \leq r'$. Für $n \geq n_0$ gilt dann

$$\begin{aligned}
|A(z_n)| &= |z_n - z_0| \left| \sum_{k=2}^{\infty} c_k \left(\sum_{\ell=1}^{k-1} \sum_{j=0}^{\ell-1} z_n^j z_0^{k-2-j} \right) \right| \\
&\stackrel{\text{Satz 4.2.9}}{\leq} |z_n - z_0| \sum_{k=2}^{\infty} |c_k| \left| \sum_{\ell=1}^{k-1} \sum_{j=0}^{\ell-1} z_n^j z_0^{k-2-j} \right| \\
&\stackrel{\Delta\text{-Ungl.}}{\leq} |z_n - z_0| \sum_{k=2}^{\infty} |c_k| \left(\sum_{\ell=1}^{k-1} \sum_{j=0}^{\ell-1} \underbrace{|z_n|^j}_{\leq (r')^j} \underbrace{|z_0|^{k-2-j}}_{\leq (r')^{k-2-j} (*)} \right) \quad a \\
&\leq |z_n - z_0| \sum_{k=2}^{\infty} |c_k| \left(\sum_{\ell=1}^{k-1} \sum_{j=0}^{\ell-1} (r')^j (r')^{k-2-j} \right) \\
&= |z_n - z_0| \sum_{k=2}^{\infty} |c_k| \left(\sum_{\ell=1}^{k-1} \sum_{j=0}^{\ell-1} (r')^{k-2} \right) = |z_n - z_0| \sum_{k=2}^{\infty} |c_k| \left(\sum_{\ell=1}^{k-1} (r')^{k-2} \underbrace{\left(\sum_{j=0}^{\ell-1} 1 \right)}_{=\ell} \right) \\
&= |z_n - z_0| \sum_{k=2}^{\infty} |c_k| \left(\sum_{\ell=1}^{k-1} (r')^{k-2} \ell \right) = |z_n - z_0| \sum_{k=2}^{\infty} |c_k| (r')^{k-2} \underbrace{\left(\sum_{\ell=1}^{k-1} \ell \right)}_{=\frac{(k-1)k}{2}} \\
&\stackrel{\text{Prop 3.3.19}}{=} |z_n - z_0| \sum_{k=2}^{\infty} \underbrace{\frac{k(k-1)}{2}}_{\leq k^2/2} |c_k| (r')^{k-2} \leq |z_n - z_0| \frac{1}{2(r')^2} \sum_{k=2}^{\infty} k^2 |c_k| (r')^k.
\end{aligned}$$

Die Reihe, nun am Ende dieser Abschätzung steht ist konvergent, weil $\sum_{k \geq 2} k^2 |c_k| z^k$ nach Proposition 5.2.6 auch den Konvergenzradius R hat und $|r' - 0| = r' < R$ (weswegen die vorhin genannte Reihe für $z = r'$ konvergiert). Dadurch ist $c := \frac{1}{2(r')^2} \sum_{k=2}^{\infty} k^2 |c_k| (r')^k$ eine reelle nichtnegative Zahl. Man erhält

$$0 \leq |A(z_n)| \leq c |z_n - z_0| \quad \forall n \in \mathbb{N}.$$

Mit $\lim_{n \rightarrow \infty} |z_n - z_0| = 0$ und dem Sandwichlemma folgt

$$\lim_{n \rightarrow \infty} |A(z_n)| = 0. \quad \text{Das heißt auch} \quad \lim_{n \rightarrow \infty} A(z_n) = 0.$$

Da $(z_n)_{n \in \mathbb{N}}$ beliebig war folgt

$$\lim_{z \rightarrow z_0} A(z) = 0.$$

ZWISCHENBEHAUPTUNG ENDE.

Aus (324) folgt nun

$$\lim_{z \rightarrow z_0} \frac{P(z) - P(z_0)}{z - z_0} = \lim_{z \rightarrow z_0} \left(A(z) + \sum_{k=1}^{\infty} k c_k z_0^{k-1} \right) = 0 + \sum_{k=1}^{\infty} k c_k z_0^{k-1} = \sum_{k=1}^{\infty} k c_k z_0^{k-1}.$$

Also gilt

$$P'(z_0) = \sum_{k=1}^{\infty} k c_k z_0^{k-1},$$

was zu zeigen war.

^aBei (*) mussten wir benutzen, dass $r^{k-2-j} \leq (r')^{k-2-j}$. Hierfür muss man überprüfen, dass $k-2-j \geq 0$ (denn wäre der Exponent negativ, so würde sich die Abschätzung umkehren. Beachte dazu: In der Summation gilt stets $j \leq \ell - 1$ und $\ell \leq k - 1$. Beides zusammen ergibt $j \leq k - 2$).

^aBei (*) mussten wir benutzen, dass $r^{k-2-j} \leq (r')^{k-2-j}$. Hierfür muss man überprüfen, dass $k-2-j \geq 0$ (denn wäre der Exponent negativ, so würde sich die Abschätzung umkehren. Beachte dazu: In der Summation gilt stets $j \leq \ell - 1$ und $\ell \leq k - 1$. Beides zusammen ergibt $j \leq k - 2$).

Ein wichtiges Korollar ist die Differenzierbarkeit der *üblichen Verdächtigen*: Polynome, exp, cos, sin,...

Korollar 7.0.11

- (i) Polynome sind auf ganz $\mathbb{C}$ differenzierbar. Genauer gilt für $n \in \mathbb{N}$, $a_0, \dots, a_n \in \mathbb{C}$ und $z_0 \in \mathbb{C}$

$$\left. \frac{d}{dz} \sum_{k=0}^n a_k z^k \right|_{z=z_0} = \sum_{k=1}^n k a_k z_0^{k-1}. \quad (325)$$

Ein wichtiger Spezialfall: Für alle $n \in \mathbb{N}$ gilt $\left. \frac{d}{dz} (z^n) \right|_{z=z_0} = n z_0^{n-1}$. ^(a)

- (ii) Die Exponentialfunktion exp ist auf $\mathbb{C}$ differenzierbar und für $z_0 \in \mathbb{C}$ gilt

$$\exp'(z_0) = \exp(z_0). \quad ^b$$

- (iii) Kosinus und Sinus sind auf $\mathbb{C}$ differenzierbar und es gilt für $z_0 \in \mathbb{C}$, dass

$$\cos'(z_0) = -\sin(z_0) \quad \text{und} \quad \sin'(z_0) = \cos(z_0).$$

^aDiesen Spezialfall folgert man leicht aus (325) mit $a_n = 1$, $a_0 = \dots = a_{n-1} = 0$.

^bEine andere Schreibweise mit Notation 5.4.1 wäre $\left. \frac{d}{dz} e^z \right|_{z=z_0} = e^{z_0}$.

Beweis. Zu (i). Definiere

$$c_k := \begin{cases} a_k & k \leq n \\ 0 & k > n \end{cases}$$

Dann hat (wie man sich leicht überlegt) $\sum_{k \geq 0} c_k z^k$ den Konvergenzradius $R = \infty$. Man erhält für $z_0 \in \mathbb{C}$ aus Proposition 7.0.10

$$\left. \frac{d}{dz} \sum_{k=0}^n a_k z^k \right|_{z=z_0} = \left. \frac{d}{dz} \sum_{k=0}^{\infty} c_k z^k \right|_{z=z_0} = \sum_{k=1}^{\infty} k c_k z_0^{k-1} = \sum_{k=1}^n k a_k z_0^{k-1}.$$

Zu (ii). Für $z_0 \in \mathbb{C}$ gilt

$$\begin{aligned} \exp'(z_0) &\stackrel{\text{Def 5.4.2}}{=} \left. \frac{d}{dz} \sum_{k=0}^{\infty} \frac{z^k}{k!} \right|_{z=z_0} = \sum_{k=1}^{\infty} k \frac{z_0^{k-1}}{k!} = \sum_{k=1}^{\infty} \frac{z_0^{k-1}}{(k-1)!} \\ &= \lim_{m \rightarrow \infty} \sum_{k=1}^m \frac{z_0^{k-1}}{(k-1)!} \stackrel{\text{Indexverschiebung}}{=} \lim_{m \rightarrow \infty} \sum_{k=0}^{m-1} \frac{z_0^k}{k!} = \sum_{k=0}^{\infty} \frac{z_0^k}{k!} = \exp(z_0). \end{aligned}$$

Das obige Argument mit der Indexverschiebung werden wir im Folgenden nicht mehr genau ausführen.

Zu (iii). Für $z_0 \in \mathbb{C}$ berechnen wir

$$\begin{aligned} \cos'(z_0) &\stackrel{\text{Prop 5.5.7}}{=} \left. \frac{d}{dz} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} z^{2k} \right|_{z=z_0} \stackrel{a}{=} \sum_{k=1}^{\infty} (2k) \frac{(-1)^k}{(2k)!} z_0^{2k-1} = \sum_{k=1}^{\infty} \frac{(-1)^k}{(2k-1)!} z_0^{2k-1} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k-1)!} z_0^{2k-1} = \sum_{k=0}^{\infty} \frac{(-1)^{k+1}}{(2(k+1)-1)!} z_0^{2(k+1)-1} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^{k+1}}{(2k+1)!} z_0^{2k+1} = - \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} z_0^{2k+1} \stackrel{\text{Prop 5.5.7}}{=} -\sin(z_0) \end{aligned}$$

Den Sinus können wir analog differenzieren. Dies ist den Lesenden als Übungsaufgabe überlassen.

^aBeachte: Um zu sehen, warum hier $2k$ als Vorfaktor in der differenzierten Summe steht und nicht nur k , sollte man die Reihe wirklich als Potenzreihe darstellen, d.h. als $\sum_{\ell=0}^{\infty} a_{\ell} z^{\ell}$ mit $a_{\ell} = \frac{(-1)^k}{(2k)!}$ falls $\ell = 2k$ gerade ist und $a_{\ell} = 0$ falls ℓ ungerade ist. Man sieht, dass nur bei geraden Summanden ein Beitrag zur Ableitung geleistet wird.

^aBeachte: Um zu sehen, warum hier $2k$ als Vorfaktor in der differenzierten Summe steht und nicht nur k , sollte man die Reihe wirklich als Potenzreihe darstellen, d.h. als $\sum_{\ell=0}^{\infty} a_{\ell} z^{\ell}$ mit $a_{\ell} = \frac{(-1)^k}{(2k)!}$ falls $\ell = 2k$ gerade ist und $a_{\ell} = 0$ falls ℓ ungerade ist. Man sieht, dass nur bei geraden Summanden ein Beitrag zur Ableitung geleistet wird.

02.02.2026

Ist eine Funktion auf ihrem gesamten Definitionsbereich differenzierbar so können wir die Werte der Ableitungen an den jeweiligen Stellen in einer Funktion zusammenfassen. Diese nennen wir dann *Ableitungsfunktion*.

Definition 7.0.12 — ABLEITUNGSFUNKTION

Es sei $D \subseteq \mathbb{K}_1$ und $f : D \rightarrow \mathbb{K}_2$ auf ganz D differenzierbar. Dann heißt $f' : D \rightarrow \mathbb{C}^a, x \mapsto f'(x)$ die *Ableitungsfunktion* von f .

^aWieder mal ist $\mathbb{C}$ kein Schreibfehler, vgl Anmerkung zu Definition 7.0.1.

Mit obiger Definition können wir jetzt auch ganz salopp $\exp' = \exp$, $\cos' = -\sin$ und $\sin' = \cos$ schreiben. Außerdem müssen wir nicht mehr an einem fixierten Punkt z_0 auswerten. Die Schreibweisen

$$\begin{aligned}\exp'(z) &= \exp(z) & \forall z \in \mathbb{C}, \\ \cos'(z) &= -\sin(z) & \forall z \in \mathbb{C}, \\ \sin'(z) &= \cos(z) & \forall z \in \mathbb{C}\end{aligned}$$

werden dadurch also auch zulässig.

7.1 Ableitungsregeln

In diesem Abschnitt lernen wir *Rechenregeln* für differenzierbare Funktionen, d.h. wir zeigen die Differenzierbarkeit von Kompositionen differenzierbarer Funktionen und geben eine Formel für die Ableitung an. Die Differenzierbarkeit der meisten Funktionen kann dann aus diesen sogenannten *Ableitungsregeln* und der Differenzierbarkeit der *üblichen Verdächtigen* (Korollar 7.0.11) gefolgert werden.

Summenregel, Produktregel, Quotientenregel

Proposition 7.1.1 — SUMMEN/PRODUKT/QUOTIENTENREGEL

Es sei $D \subseteq \mathbb{K}_1$ und $x_0 \in D$. Ferner seien $f, g : D \rightarrow \mathbb{K}_2$ bei x_0 differenzierbar. Dann gilt

- (i) *Summenregel*. Die Funktion $f + g : D \rightarrow \mathbb{K}_2, x \mapsto f(x) + g(x)$ ist bei x_0 differenzierbar mit

$$(f + g)'(x_0) = f'(x_0) + g'(x_0).$$

- (ii) *Produktregel*. Die Funktion $f \cdot g : D \rightarrow \mathbb{K}_2, x \mapsto f(x)g(x)$ ist bei x_0 differenzierbar mit

$$(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

- (iii) *Quotientenregel*. Ist $g(x_0) \neq 0$ so ist $\frac{f}{g} : \hat{D} \rightarrow \mathbb{K}_2, x \mapsto \frac{f(x)}{g(x)}$ (wobei $\hat{D} := \{x \in D : g(x) \neq 0\}$) bei x_0 differenzierbar und es gilt

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}.$$

Beweis. Da f und g bei x_0 differenzierbar sind, existieren nach Definition 7.0.1

$$f'(x_0) := \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \quad \text{und} \quad g'(x_0) := \lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0}. \quad (326)$$

Zu (i). Wir bilden nun

$$\begin{aligned} \lim_{x \rightarrow x_0} \frac{(f+g)(x) - (f+g)(x_0)}{x - x_0} &= \lim_{x \rightarrow x_0} \frac{f(x) + g(x) - (f(x_0) + g(x_0))}{x - x_0} \\ &= \lim_{x \rightarrow x_0} \frac{f(x) + g(x) - f(x_0) - g(x_0)}{x - x_0} \\ &= \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) + g(x) - g(x_0)}{x - x_0} \\ &= \lim_{x \rightarrow x_0} \underbrace{\frac{f(x) - f(x_0)}{x - x_0}}_{\substack{\xrightarrow{x \rightarrow x_0} f'(x_0) \\ (326)}} + \underbrace{\frac{g(x) - g(x_0)}{x - x_0}}_{\substack{\xrightarrow{x \rightarrow x_0} g'(x_0) \\ (326)}} \\ &\stackrel{\text{Prop 6.1.8}}{=} f'(x_0) + g'(x_0). \end{aligned} \quad (327)$$

Man erhält

$$(f+g)'(x_0) \stackrel{\text{Def 7.0.1}}{=} \lim_{x \rightarrow x_0} \frac{(f+g)(x) - (f+g)(x_0)}{x - x_0} \stackrel{(327)}{=} f'(x_0) + g'(x_0).$$

Zu (ii). Für $x \in D \setminus \{x_0\}$ gilt

$$\begin{aligned} \frac{(f \cdot g)(x) - (f \cdot g)(x_0)}{x - x_0} &= \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \frac{f(x)g(x) - f(x_0)g(x) + f(x_0)g(x) - f(x_0)g(x_0)}{x - x_0} \\ &= \frac{(f(x) - f(x_0))g(x) + f(x_0)(g(x) - g(x_0))}{x - x_0} \\ &= \frac{(f(x) - f(x_0))g(x)}{x - x_0} + \frac{f(x_0)(g(x) - g(x_0))}{x - x_0} \\ &= \frac{f(x) - f(x_0)}{x - x_0} g(x) + f(x_0) \frac{g(x) - g(x_0)}{x - x_0}. \end{aligned} \quad (328)$$

Da g bei x_0 differenzierbar ist, ist g bei x_0 auch stetig, vgl. Korollar 7.0.7. Das bedeutet

$$\lim_{x \rightarrow x_0} g(x) = g(x_0), \quad (329)$$

vgl. Definition 6.2.1. Man erhält daher

$$\begin{aligned} \lim_{x \rightarrow x_0} \frac{(f \cdot g)(x) - (f \cdot g)(x_0)}{x - x_0} &= \lim_{x \rightarrow x_0} \underbrace{\frac{f(x) - f(x_0)}{x - x_0}}_{\substack{\xrightarrow{x \rightarrow x_0} f'(x_0) \\ \text{vgl. (329)}}} \underbrace{g(x)}_{\substack{\xrightarrow{x \rightarrow x_0} g(x_0) \\ \text{vgl. (329)}}} + f(x_0) \underbrace{\frac{g(x) - g(x_0)}{x - x_0}}_{\xrightarrow{x \rightarrow x_0} g'(x_0)} \\ &\stackrel{\text{Prop 6.1.8}}{=} f'(x_0)g(x_0) + f(x_0)g'(x_0). \end{aligned}$$

Man erhält wieder mit Definition 7.0.1

$$(f \cdot g)'(x_0) = \lim_{x \rightarrow x_0} \frac{(f \cdot g)(x) - (f \cdot g)(x_0)}{x - x_0} = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

Zu (iii). Zunächst studieren wir den Spezialfall $f \equiv 1$, d.h. wir betrachten die Funktion $\frac{1}{g}$. Man überzeugt sich leicht davon, dass x_0 auch ein Häufungspunkt von $\hat{D} \setminus \{x_0\}$ ist ^a. Nun berechnen wir für $x \in \hat{D} \setminus \{x_0\}$

$$\begin{aligned} \frac{\frac{1}{g(x)} - \frac{1}{g(x_0)}}{x - x_0} &= \frac{\frac{g(x_0)}{g(x)g(x_0)} - \frac{g(x)}{g(x)g(x_0)}}{x - x_0} = \frac{\frac{1}{g(x)g(x_0)}(g(x_0) - g(x))}{x - x_0} = \frac{1}{g(x)g(x_0)} \frac{g(x_0) - g(x)}{x - x_0} \\ &= -\frac{1}{g(x)g(x_0)} \frac{g(x) - g(x_0)}{x - x_0}. \end{aligned}$$

Wie schon in (329) diskutiert, folgt aus Korollar 7.0.7, dass $\lim_{x \rightarrow x_0} g(x) = g(x_0)$. Nach Proposition 6.1.8 gilt dann auch $\lim_{x \rightarrow x_0} \frac{1}{g(x)} = \frac{1}{g(x_0)}$. Dies und (326) und die Grenzwertregeln implizieren

$$\begin{aligned} \lim_{x \rightarrow x_0} \frac{\frac{1}{g(x)} - \frac{1}{g(x_0)}}{x - x_0} &= \lim_{x \rightarrow x_0} -\frac{1}{g(x)g(x_0)} \frac{g(x) - g(x_0)}{x - x_0} \\ &= -\frac{1}{g(x_0)} \left(\lim_{x \rightarrow x_0} \frac{1}{g(x)} \right) \left(\lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0} \right) = -\frac{1}{g(x_0)^2} g'(x_0). \end{aligned}$$

Damit haben wir gezeigt, dass $\frac{1}{g}$ bei x_0 differenzierbar ist und

$$\left(\frac{1}{g} \right)'(x_0) = \frac{-g'(x_0)}{g(x_0)^2} \quad (330)$$

gilt. Nun studieren wir den allgemeinen Fall. Seien f und g wie in der Voraussetzung. Wir haben oben gesehen, dass $\frac{1}{g}$ bei x_0 differenzierbar ist. Nach der Produktregel (ii) ist dann

auch $\frac{f}{g} = f \cdot \frac{1}{g}$ bei x_0 differenzierbar und es gilt

$$\begin{aligned} \left(\frac{f}{g}\right)'(x_0) &= \left(f \cdot \frac{1}{g}\right)'(x_0) \\ &\stackrel{(ii)}{=} f'(x_0) \frac{1}{g(x_0)} + f(x_0) \left(\frac{1}{g}\right)'(x_0) \\ &\stackrel{(330)}{=} f'(x_0) \frac{1}{g(x_0)} + f(x_0) \frac{-g'(x_0)}{g(x_0)^2} \\ &= \frac{f'(x_0)g(x_0)}{g(x_0)^2} + \frac{-f(x_0)g'(x_0)}{g(x_0)^2} \\ &= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2}. \end{aligned}$$

^aDies ist den Lesenden als Übungsaufgabe überlassen. Die Technik ist ähnlich wie im Beweis der Zwischenbehauptung von Proposition 6.1.8

Beispiel 7.1.2

Es sei $n \in \mathbb{N}$ und $q : \mathbb{K} \setminus \{0\} \rightarrow \mathbb{K}, x \mapsto x^{-n}$ gegeben.

Behauptung. q ist bei jedem $x_0 \in \mathbb{K} \setminus \{0\}$ differenzierbar und es gilt $q'(x_0) = -nx_0^{-n-1}$.

Beweis. Wir schreiben $q(x)$ für $x \in \mathbb{K} \setminus \{0\}$ wie folgt um:

$$q(x) = x^{-n} = \frac{1}{x^n} = \frac{1}{g(x)},$$

wobei $g : \mathbb{K} \rightarrow \mathbb{K}, x \mapsto x^n$. Für $x_0 \in \mathbb{K} \setminus \{0\}$ ist g bei x_0 differenzierbar und $g(x_0) \neq 0$. Außerdem hat man $g'(x_0) = nx_0^{n-1}$. Aus Proposition 7.1.1 (iii) mit $f \equiv 1$ folgt, dass q bei x_0 differenzierbar ist und

$$\begin{aligned} q'(x_0) &= \frac{0 \cdot g(x_0) - 1 \cdot g'(x_0)}{g(x_0)^2} = \frac{-g'(x_0)}{g(x_0)^2} \\ &= \frac{-(nx_0^{n-1})}{(x_0^n)^2} = -n \frac{x_0^{n-1}}{x_0^{2n}} = -nx_0^{n-1-2n} = -nx_0^{-n-1}. \end{aligned}$$

Beispiel 7.1.3

Es sei $\hat{D} := \{x \in \mathbb{R} : \cos(x) \neq 0\}$ und $\tan : \hat{D} \rightarrow \mathbb{R}, x \mapsto \tan(x)$ gegeben. Zur Erinnerung: Für $x \in \hat{D}$ definieren wir $\tan(x) := \frac{\sin(x)}{\cos(x)}$.

Behauptung. $\tan$ ist differenzierbar und es gilt $\tan'(x) = 1 + \tan^2(x)$ für alle $x \in D$.

Beweis. Für $x \in D$ gilt nach Proposition 7.1.1 (iv)

$$\begin{aligned} \tan'(x) &= \left(\frac{\sin}{\cos} \right)'(x) = \frac{\sin'(x) \cos(x) - \sin(x) \cos'(x)}{\cos(x)^2} \\ &\stackrel{\text{Kor7.0.11}}{=} \frac{\cos(x) \cos(x) - \sin(x)(-\sin(x))}{\cos^2(x)} = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} \quad (331) \\ &= 1 + \frac{\sin^2(x)}{\cos^2(x)} = 1 + \left(\frac{\sin(x)}{\cos(x)} \right)^2 = 1 + \tan(x)^2 = 1 + \tan^2(x). \end{aligned}$$

Eine andere hilfreiche Darstellung der obigen Ableitung des Tangens ist

$$\tan'(x) \stackrel{(331)}{=} \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)}.$$

Ein wichtiges Korollar aus der Produktregel ist auch, dass *konstante Vielfache* differenzierbarer Funktion wieder differenzierbar sind.

Korollar 7.1.4

Es sei $D \subseteq \mathbb{K}_1$ und $x_0 \in D$. Ferner sei $f : D \rightarrow \mathbb{K}_2$ bei x_0 differenzierbar und $\lambda \in \mathbb{K}_2$. Dann ist auch $(\lambda f) : D \rightarrow \mathbb{K}_2$ bei x_0 differenzierbar und es gilt

$$(\lambda f)'(x_0) = \lambda f'(x_0).$$

Beweis. Betrachte die konstante Funktion $g : D \rightarrow \mathbb{K}_2, x \mapsto \lambda$. Dann ist g differenzierbar bei x_0 , vgl. Beispiel 7.0.2. Außerdem gilt $g'(x_0) = 0$. Nach der Produktregel (Proposition 7.1.1) ist dann auch $\lambda f = gf$ bei x_0 differenzierbar und man hat

$$(\lambda f)'(x_0) = (gf)'(x_0) = g'(x_0)f(x_0) + f'(x_0)g(x_0) = 0 \cdot f(x_0) + f'(x_0)\lambda = \lambda f'(x_0).$$

Die Behauptung ist gezeigt.

Kettenregel

Die nächste Regel die wir kennenlernen ist die *Kettenregel*, welche das Differenzieren von *Verkettungen* erlaubt. Im Folgenden seien wie immer $\mathbb{K}_1, \mathbb{K}_2, \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

Proposition 7.1.5 — KETTENREGEL

Es seien $D \subseteq \mathbb{K}_1$ und $E \subseteq \mathbb{K}_2$ und $f : D \rightarrow \mathbb{K}_2$ sowie $g : E \rightarrow \mathbb{K}$ gegeben. Ferner sei $f(D) \subseteq E$ und $x_0 \in D$ so, dass f bei x_0 differenzierbar ist und g bei $f(x_0)$ differenzierbar

ist. Dann ist auch $(g \circ f)$ bei x_0 differenzierbar und es gilt

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Beweis. Dazu sei x_0 wie in der Voraussetzung. Da g bei $f(x_0)$ differenzierbar ist, gibt es nach Proposition 7.0.6 eine bei $f(x_0)$ stetige Funktion $\phi : E \rightarrow \mathbb{C}$, mit $\phi(f(x_0)) = 0$ so, dass

$$\forall y \in E : \quad g(y) = g(f(x_0)) + c(y - f(x_0)) + \phi(y)(y - f(x_0)), \quad \text{wobei } c = g'(f(x_0)).$$

Zusammengefasst gilt also

$$\forall y \in E : \quad g(y) = g(f(x_0)) + g'(f(x_0))(y - f(x_0)) + \phi(y)(y - f(x_0)).$$

Sei nun $x \in D$ beliebig. Insbesondere gilt dann $f(x) \in f(D) \subseteq E$. Dann kann man $y = f(x)$ in obige Gleichung einsetzen und erhält

$$\forall x \in D : \quad g(f(x)) = g(f(x_0)) + g'(f(x_0))(f(x) - f(x_0)) + \phi(f(x))(f(x) - f(x_0)). \quad (332)$$

Wir folgern für alle $x \in D \setminus \{x_0\}$

$$\begin{aligned} \frac{(g \circ f)(x) - (g \circ f)(x_0)}{x - x_0} &= \frac{g(f(x)) - g(f(x_0))}{x - x_0} \\ &\stackrel{(332)}{=} \frac{[g(f(x_0)) + g'(f(x_0))(f(x) - f(x_0)) + \phi(f(x))(f(x) - f(x_0))] - g(f(x_0))}{x - x_0} \\ &= \frac{g'(f(x_0))(f(x) - f(x_0)) + \phi(f(x))(f(x) - f(x_0))}{x - x_0} \\ &= g'(f(x_0)) \frac{f(x) - f(x_0)}{x - x_0} + \phi(f(x)) \frac{f(x) - f(x_0)}{x - x_0}. \end{aligned} \quad (333)$$

Nun beachte: Da f bei x_0 stetig ist und ϕ bei $f(x_0)$ stetig ist, ist auch $\phi \circ f$ bei x_0 stetig. Das bedeutet, dass

$$\lim_{x \rightarrow x_0} \phi(f(x)) = \lim_{x \rightarrow x_0} (\phi \circ f)(x) = (\phi \circ f)(x_0) = \phi(f(x_0)) = 0. \quad (334)$$

Wir folgern aus (333), dass

$$\lim_{x \rightarrow x_0} \frac{(g \circ f)(x) - (g \circ f)(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} g'(f(x_0)) \underbrace{\frac{f(x) - f(x_0)}{x - x_0}}_{\xrightarrow{x \rightarrow x_0} f'(x_0)} + \underbrace{\phi(f(x))}_{\xrightarrow{x \rightarrow x_0} 0 \text{ vgl. (334)}} \underbrace{\frac{f(x) - f(x_0)}{x - x_0}}_{\xrightarrow{x \rightarrow x_0} f'(x_0)}$$

$$= g'(f(x_0))f'(x_0) + 0 \cdot f'(x_0) = g'(f(x_0))f'(x_0).$$

Es folgt mit Definition 7.0.1, dass $g \circ f$ bei x_0 differenzierbar ist und

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Bei der Verkettung $g \circ f$ bezeichnet man auch oft g als *äußere Funktion* und f als *innere Funktion*. Bei der Ableitungsformel

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$$

nennt man den Term $g'(f(x_0))$ auch *äußere Ableitung* und den Term $f'(x_0)$ *innere Ableitung*. Das Multiplizieren mit der inneren Ableitung nennt man manchmal auch *Nachdifferenzieren*.

Beispiel 7.1.6

Ist $D \subseteq \mathbb{K}_1$ und $f : D \rightarrow \mathbb{K}_2$ auf D differenzierbar, so ist $\exp \circ f$ auch auf D differenzierbar und für alle $x \in D$ gilt

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

Eine Folgerung aus obigem Beispiel ist

Beispiel 7.1.7

Es sei $\beta \in (0, \infty)$.

Behauptung. Dann ist $h : \mathbb{C} \rightarrow \mathbb{C}, z \mapsto \beta^z$ auf ganz $\mathbb{C}$ differenzierbar mit $h'(z) = \beta^z \log(\beta)$.

Beweis. Es gilt für alle $z \in \mathbb{C}$

$$h(z) = \beta^z = \exp(z \log(\beta)) = \exp(f(z)), \quad \text{wobei } f : \mathbb{C} \rightarrow \mathbb{C}, z \mapsto z \log(\beta).$$

Wendet man nun Beispiel 7.1.6 an, so folgt, dass $h = \exp \circ f$ differenzierbar ist und für alle $z \in \mathbb{C}$ gilt

$$h'(z) = \exp(f(z))f'(z) = \exp(z \log(\beta)) \log(\beta) = \beta^z \log(\beta).$$

Als Merkregel können wir festhalten

$$\forall z \in \mathbb{C} : \quad \frac{d}{dz} \beta^z = \beta^z \log(\beta).$$

Ableitung der Umkehrfunktion

Wir wissen bereits aus Satz 6.2.20: Umkehrfunktionen von stetigen, bijektiven Funktionen sind automatisch stetig. Man sagt, die Stetigkeit *überträgt* sich auf die Umkehrfunktion.

Hier wollen wir uns nun fragen, ob sich auch die Differenzierbarkeit auf Umkehrfunktionen überträgt.

Satz 7.1.8 — ABLEITUNG DER UMKEHRFUNKTION

Es sei $Z \subseteq \mathbb{R}$ und $f : [a, b] \rightarrow Z$ eine stetige und bijektive Funktion mit Umkehrfunktion $f^{-1} : Z \rightarrow [a, b]$. Ist f an einer Stelle $x_0 \in [a, b]$ differenzierbar und $f'(x_0) \neq 0$ so ist f^{-1} bei $y_0 := f(x_0)$ auch differenzierbar und es gilt

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

Anders ausgedrückt gilt

$$(f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))}. \quad (335)$$

Beweis. Wir wollen die Existenz von

$$\lim_{y \rightarrow y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0}$$

beweisen. Hierzu sei $(y_n)_{n \in \mathbb{N}} \subseteq Z \setminus \{y_0\}$ eine beliebige Folge mit $\lim_{n \rightarrow \infty} y_n = y_0$. Definiere für $n \in \mathbb{N}$ auch

$$x_n := f^{-1}(y_n). \quad (336)$$

Beachte: Es gilt $x_n \neq x_0$ für alle $n \in \mathbb{N}$, denn wäre $x_n = x_0$ für ein $n \in \mathbb{N}$, so wäre auch $f(x_n) = f(x_0)$. Das hieße dann aber wegen (336)

$$y_n = y_0. \quad \nexists \text{ (widerspricht } y_n \in Z \setminus \{y_0\}).$$

Ferner gilt wegen der Stetigkeit von f^{-1} (vgl. Satz 6.2.20)

$$\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} f^{-1}(y_n) \stackrel[\text{Def 6.2.1}]{\text{Satz 6.2.20}} f^{-1}(y_0) = x_0. \quad (337)$$

Nun hat man

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} &= \lim_{n \rightarrow \infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{f(f^{-1}(y_n)) - f(f^{-1}(y_0))} \\ &= \lim_{n \rightarrow \infty} \frac{x_n - x_0}{f(x_n) - f(x_0)} \\ &\stackrel{x_n \neq x_0 \ \forall n}{=} \lim_{n \rightarrow \infty} \frac{1}{\frac{f(x_n) - f(x_0)}{x_n - x_0}}. \end{aligned} \quad (338)$$

Dann gilt wegen (337), dass

$$\lim_{n \rightarrow \infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0).$$

Mit (338) und den Grenzwertregeln (anwendbar da $f'(x_0) \neq 0$) folgt

$$\lim_{n \rightarrow \infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} = \frac{1}{f'(x_0)}.$$

Da $(y_n)_{n \in \mathbb{N}}$ beliebig war folgt

$$\lim_{y \rightarrow y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{1}{f'(x_0)}.$$

Mit Definition 7.0.1 ergibt sich

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}.$$

Ein häufiger Irrglaube ist, dass die Umkehrfunktion f^{-1} dasselbe ist wie $\frac{1}{f}$. Dies ist **falsch**. Ein Gegenbeispiel ist $f = \text{id}_{\mathbb{R}}$ (denn dann ist $f^{-1} = \text{id}_{\mathbb{R}} \neq \frac{1}{f}$). Wäre der Irrglaube richtig, so müsste man den obigen Satz gar nicht beweisen, sondern könnte einfach auf Proposition 7.1.1 (iv) verweisen... So einfach ist es leider nicht!

Bemerkung 7.1.9

Wer sich die Formel (335) nicht merken kann, kann sich die Formel einfach aus der Kettenregel herleiten. Sei dazu $f : [a, b] \rightarrow Z$ wie in Satz 7.1.8. Wenn wir schon von der Differenzierbarkeit von f^{-1} wissen, dann gilt für alle $y \in Z$

$$1 = \text{id}'(y) = (f \circ f^{-1})'(y) \stackrel{\text{Prop 7.1.5}}{=} f'(f^{-1}(y)) \cdot (f^{-1})'(y).$$

Teilt man durch den ersten Faktor, so erhält man unsere Formel

$$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}.$$

Eine Voraussetzung, die man bei der Anwendung von Satz 7.1.8 stets überprüfen muss, ist $f'(x_0) \neq 0$ (für x_0 wie in der Aussage des Satzes). Diese ist unabdingbar, wie wir im nachfolgenden Beispiel sehen werden.

Bemerkung 7.1.10

Verzichtet man in Satz 7.1.8 auf die Voraussetzung $f'(x_0) \neq 0$, so überträgt sich die Differenzierbarkeit *im Allgemeinen nicht* auf die Umkehrfunktion. Dazu ein Beispiel: Betrachte die (bijektive stetige) Funktion

$$f : [-1, 1] \rightarrow [-1, 1], \quad x \mapsto x^3 \quad \text{und} \quad x_0 = 0.$$

Dann ist $y_0 := f(x_0) = 0$. Außerdem ist f differenzierbar und $f'(x_0) = 3x_0^2 = 0$.

Behauptung: f^{-1} ist bei y_0 nicht differenzierbar.

Beweis. Es gilt nach der Definition der dritten Wurzel (Definition 10.6.14)

$$f^{-1}(x) = \sqrt[3]{x} \quad \forall x \in [0, 1]. \quad {}^a$$

Damit sieht man, dass f^{-1} bei 0 nicht differenzierbar ist, denn

$$\lim_{x \rightarrow 0+} \frac{f^{-1}(x) - f^{-1}(0)}{x - 0} = \lim_{x \rightarrow 0+} \frac{\sqrt[3]{x} - 0}{x - 0} \stackrel{\text{Prop 6.3.10(vii)}}{=} \lim_{x \rightarrow 0+} \frac{x^{\frac{1}{3}}}{x} \stackrel{\text{Prop 6.3.10(iv)}}{=} \lim_{x \rightarrow 0} \frac{1}{x^{\frac{2}{3}}} \stackrel{\text{Prop 6.3.16}}{=} \infty.$$

Damit wir von einer Ableitung sprechen können, muss der Grenzwert aber in $\mathbb{C}$ liegen, vgl. Definition 7.0.1. (Ist eine Funktion bestimmt divergent gegen $\pm\infty$ so besitzt sie nämlich in unserem Sinne kein Grenzwert, vgl. 6.1.22)

^aDies würde eigentlich sogar für $x \in [-1, 1]$ gelten, aber wir haben bisher die dritte Wurzel nur für positive Zahlen definiert, vgl. Definition 10.6.14.

Beispiel 7.1.11 — ABLEITUNG DES LOGARITHMUS

Der natürliche Logarithmus $\log : (0, \infty) \rightarrow \mathbb{R}$ ist differenzierbar und es gilt

$$\log'(x) = \frac{1}{x} \quad \text{für alle } x \in (0, \infty).$$

Beweis. Sei $y_0 \in (0, \infty)$. Dann gibt es nach Proposition 6.3.1 ein $x_0 \in \mathbb{R}$ mit $\exp(x_0) = y_0$. Wende Satz 7.1.8 mit $f = \exp$ und $[a, b] = [x_0 - 1, x_0 + 1]$ an^a und erhalte

$$\log'(y_0) = (\exp^{-1})'(y_0) = \frac{1}{\exp'(\exp^{-1}(y_0))} \stackrel{\text{Kor 7.0.11}}{=} \frac{1}{\exp(\exp^{-1}(y_0))} = \frac{1}{y_0}.$$

^aAls Voraussetzung muss hier eigentlich geprüft werden, dass $\exp'(x_0) \neq 0$. Das ist klar, da $\exp'(x_0) = \exp(x_0) \in (0, \infty)$, vgl. Kor 7.0.11

Mit dem obigen Beispiel und der Kettenregel kann man auch allgemeine Potenzfunktionen ableiten.

Proposition 7.1.12 — ABLEITUNG DER ALLGEMEINEN POTENZ

Sei $\alpha \in \mathbb{C}$. Dann ist die Funktion $p : (0, \infty) \rightarrow \mathbb{C}, x \mapsto x^\alpha$ auf ganz $(0, \infty)$ differenzierbar mit $p'(x) = \alpha x^{\alpha-1}$ für alle $x \in (0, \infty)$.

Beweis. Für alle $x \in (0, \infty)$ gilt

$$p(x) = x^\alpha \stackrel{\text{Def 6.3.7}}{=} \exp(\alpha \log(x)).$$

Mit der Kettenregel folgt dass p für alle $x \in (0, \infty)$ differenzierbar ist und

$$\begin{aligned} p'(x) &= \exp'(\alpha \log(x))(\alpha \log)'(x) \stackrel{\text{Kor7.1.4}}{=} \exp(\alpha \log(x)) \alpha \log'(x) \\ &= \exp(\alpha \log(x)) \alpha \frac{1}{x} = \alpha x^\alpha \frac{1}{x} \stackrel{\text{Prop6.3.10}}{=} \alpha x^{\alpha-1}. \end{aligned}$$

In Proposition 5.4.10 haben wir mühsam bewiesen, dass

$$\exp(1) = e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n.$$

Mithilfe der Theorie der Differenzierbarkeit kann man *viel eleganter* eine allgemeinere Aussage zeigen (vgl. auch Bemerkung 5.4.11).

Proposition 7.1.13

Für alle $x \in \mathbb{R}$ gilt

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = \exp(x).$$

Beweis. Für $x = 0$ ist die Aussage klar. Es sei also $x \in \mathbb{R} \setminus \{0\}$. Definiere für $n \in \mathbb{N}$

$$y_n := 1 + \frac{x}{n} \in \mathbb{R} \setminus \{1\}$$

Dann gilt für $n > N := [|x|] + 1$, dass $y_n \in (0, \infty)$. Somit ist $(y_n)_{n \geq N} \subseteq (0, \infty) \setminus \{1\}$ eine Folge mit $\lim_{n \rightarrow \infty} y_n = 1$. Man folgert

$$\lim_{n \rightarrow \infty} \frac{\log(y_n) - \log(1)}{y_n - 1} = \lim_{y \rightarrow 1} \frac{\log(y) - \log(1)}{y - 1} = \log'(1) \stackrel{\text{Bsp7.1.11}}{=} \frac{1}{1} = 1.$$

Somit gilt

$$\begin{aligned} 1 &= \lim_{n \rightarrow \infty} \frac{\log(y_n) - \log(1)}{y_n - 1} = \lim_{n \rightarrow \infty} \frac{\log(1 + \frac{x}{n}) - 0}{(1 + \frac{x}{n}) - 1} = \lim_{n \rightarrow \infty} \frac{\log(1 + \frac{x}{n}) - 0}{\frac{x}{n}} \\ &= \lim_{n \rightarrow \infty} \frac{1}{x} n \log\left(1 + \frac{x}{n}\right) = \frac{1}{x} \lim_{n \rightarrow \infty} n \log\left(1 + \frac{x}{n}\right) \stackrel{\text{Bem6.3.9}}{=} \frac{1}{x} \lim_{n \rightarrow \infty} \log\left(1 + \frac{x}{n}\right)^n. \end{aligned}$$

Multiplizieren wir mit x , so folgt

$$x = \lim_{n \rightarrow \infty} \log\left(1 + \frac{x}{n}\right)^n.$$

Wegen der Stetigkeit von $\exp$ ergibt sich

$$\exp(x) = \exp\left(\lim_{n \rightarrow \infty} \log\left(1 + \frac{x}{n}\right)^n\right) \stackrel{(281)}{=} \lim_{n \rightarrow \infty} \exp\left(\log\left(1 + \frac{x}{n}\right)^n\right) = \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n.$$

Die Behauptung folgt.

7.2 Lokale Extrema

Im Folgenden wollen wir besprechen, wie Ableitungen bei der *Kurvendiskussion* helfen können. Kurvendiskussion meint hierbei eine Diskussion des *Verlaufs* des Funktionsgraphen Γ_f einer gegebenen Funktion $f : D \rightarrow \mathbb{R}$ (d.h. hier stets ist $\mathbb{K}_2 = \mathbb{R}$). Was alles zu einer solchen Kurvendiskussion dazugehört werden wir später noch besprechen. Ein wichtiger Schritt auf dem Weg dorthin ist jedenfalls die Bestimmung *lokaler Extrema*, welche wir gleich definieren wollen.

Definition 7.2.1 — LOKALE EXTREMA

Sei $D \subseteq \mathbb{K}_1$ sowie $f : D \rightarrow \mathbb{R}$ und $x_0 \in D$. Wir sagen

- (i) f nimmt an der Stelle x_0 ein *lokales Maximum* an, falls es ein $\delta > 0$ gibt mit

$$\forall x \in D \cap B_\delta(x_0) : \quad f(x) \leq f(x_0)$$

- (ii) f nimmt an der Stelle x_0 ein *lokales Minimum* an, falls es ein $\delta > 0$ gibt mit

$$\forall x \in D \cap B_\delta(x_0) : \quad f(x) \geq f(x_0)$$

- (iii) f nimmt an der Stelle x_0 ein *striktes lokales Maximum* an, falls es ein $\delta > 0$ gibt mit

$$\forall x \in D \cap B_\delta(x_0) \setminus \{x_0\} : \quad f(x) < f(x_0).$$

- (iv) f nimmt an der Stelle x_0 ein *striktes lokales Minimum* an, falls es ein $\delta > 0$ gibt mit

$$\forall x \in D \cap B_\delta(x_0) \setminus \{x_0\} : \quad f(x) > f(x_0).$$

Wir nennen in den obigen Fällen x_0 eine (strikte) *lokale Maximumsstelle* bzw. *lokale Minimumsstelle*. Allgemein spricht man auch von einer (strikten) *lokalen Extremstelle*, falls x_0 eine (strikte) *lokale Maximumsstelle* oder *lokale Minimumsstelle* ist.

Bemerkung 7.2.2

In Definition 6.2.10 haben wir definiert, was es bedeutet, ein (sog. *globales*) Maximum auf D anzunehmen. Wir sollten nun an dieser Stelle dazu sagen, dass jede globale Maximumsstelle auch eine lokale Maximumsstelle ist. In der Tat: Ist x_0 eine globale Maximumsstelle von

$f : D \rightarrow \mathbb{R}$ im Sinne von Definition 6.2.10, so gilt

$$\forall x \in D : \quad f(x) \leq f(x_0).$$

Insbesondere gilt auch

$$\forall x \in D \cap B_1(x_0) : \quad f(x) \leq f(x_0).$$

Somit ist Definition 7.2.1 mit $\delta = 1$ erfüllt und x_0 ist auch eine lokale Maximumsstelle.

Allerdings muss nicht jede lokale Maximumsstelle auch gleich eine globale Maximumsstelle sein, wie etwa Abbildung 28 eindrücklich zeigt.

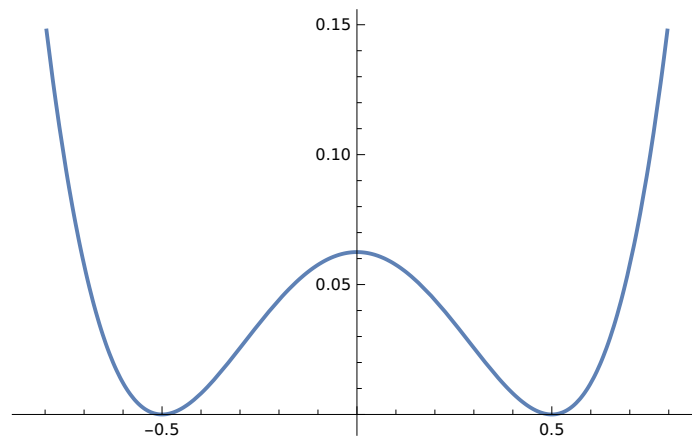


Abbildung 28: Der Graph der Funktion $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto (1 - x^2)^2$. Bei $x_0 = 0$ liegt ein lokales Maximum vor, welches aber kein globales Maximum ist. Man kann den Wert δ aus Definition 7.2.1 als 0.5 wählen.

Bemerkung 7.2.3

Es sei $D \subseteq \mathbb{K}_1$ und $f : D \rightarrow \mathbb{R}$ sowie $x_0 \in D$. Man überzeugt sich leicht

f nimmt in x_0 ein lokales Maximum an

$$\Longleftrightarrow$$

$-f$ nimmt in x_0 ein lokales Minimum an.

Spiegelbildlich gilt auch

f nimmt in x_0 ein lokales Minimum an

$$\Longleftrightarrow$$

$-f$ nimmt in x_0 ein lokales Maximum an.

Selbiges gilt auch für strikte lokale Maxima/Minima.

Notwendiges Kriterium für innere lokale Extremstellen

Falls $D \subseteq \mathbb{R}$ so kann man mit dem sog. „notwendigen Kriterium für lokale Extremstellen“ *Kandidaten* für lokale Extremstellen ausfindig machen. Was mit *Kandidaten* gemeint ist klären wir später auf.

Zunächst ist aber zu sagen, dass unser Kriterium nur sog. *innere Punkte* des Definitionsbereiches untersuchen kann:

Definition 7.2.4 — INNERER PUNKT

Es sei $D \subseteq \mathbb{K}$ eine nichtleere Menge. Wir nennen $x_0 \in D$ einen *inneren Punkt* von D wenn es ein $r > 0$ gibt mit $B_r(x_0) \subseteq D$.^a Die Menge aller inneren Punkte von D wird mit $\overset{\circ}{D}$ oder $\text{int}(D)$ bezeichnet.

^aZur Erinnerung: $B_r(x_0) = \{x \in \mathbb{K} : |x - x_0| < r\}$, vgl. Definition 6.1.26

Beispiel 7.2.5

Sei $\mathbb{K} = \mathbb{R}$ und $a, b \in \mathbb{R}$. Wir bestimmen alle inneren Punkte von $D := [a, b]$.

Behauptung. x_0 innerer Punkt von $[a, b] \iff x_0 \in (a, b)$.

Mit anderen Worten gilt $\overset{\circ}{[a, b]} = (a, b)$.

***Beweis.** Zu „ $\Leftarrow$ “. Es sei $x_0 \in (a, b)$, d.h. $x_0 > a$ und $x_0 < b$. Wähle $r := \frac{1}{2} \min\{x_0 - a, b - x_0\} > 0$. Wir behaupten, dass dann $B_r(x_0) \subseteq [a, b]$ gilt. In der Tat: Es sei $x \in B_r(x_0)$. Man hat

$$x \in B_r(x_0) \implies |x - x_0| < r \implies x_0 - r < x < x_0 + r \implies x \in (x_0 - r, x_0 + r). \quad (339)$$

Nun gilt

$$x_0 - r \geq x_0 - \frac{1}{2}(x_0 - a) > x_0 - (x_0 - a) = a.$$

Außerdem gilt

$$x_0 + r \leq x_0 + \frac{1}{2}(b - x_0) < x_0 + b - x_0 = b.$$

Wir haben also gesehen, dass $(x_0 - r, x_0 + r) \subseteq (a, b)$. Man folgert aus (339), dass

$$x \in (a, b) \quad \text{und somit auch} \quad x \in [a, b].$$

Da also $B_r(x_0) \subseteq [a, b]$, ist x_0 innerer Punkt von $[a, b]$, vgl. Definition 7.2.4.

Für „ $\implies$ “ zeigen wir die Kontraposition. Das bedeutet für $x_0 \in [a, b]$ zeigen wir

$$x_0 \notin (a, b) \implies x_0 \text{ kein innerer Punkt von } [a, b].$$

Sei dazu $x_0 \in [a, b]$ so, dass $x_0 \notin (a, b)$. Mit anderen Worten gilt entweder $x_0 = a$ oder $x_0 = b$. Ist $x_0 = a$ so gilt für alle $r > 0$

$$B_r(x_0) = B_r(a) = (a - r, a + r) \not\subseteq [a, b]$$

denn zum Beispiel ist $a - \frac{r}{2} \in (a - r, a + r) \setminus [a, b]$. Also kann x_0 kein innerer Punkt sein. Ist $x_0 = b$, so gilt für alle $r > 0$

$$B_r(x_0) = B_r(b) = (b - r, b + r) \not\subseteq [a, b],$$

denn zum Beispiel ist $b + \frac{r}{2} \in (b - r, b + r) \setminus [a, b]$.

Nun zum angekündigten Kriterium:

Satz 7.2.6 — NOTWENDIGES KRITERIUM FÜR (INNERE) LOKALE EXTREMSTELLEN

Es sei $D \subseteq \mathbb{K}$ nichtleer und $f : D \rightarrow \mathbb{R}$ gegeben. Sei außerdem $x_0 \in D$ so, dass

- (a) f bei x_0 differenzierbar ist und
- (b) x_0 ein innerer Punkt von D ist.

Dann gilt

$$x_0 \text{ ist eine lokale Extremstelle von } f \implies f'(x_0) = 0.$$

13.04.2026

Beweis. Sei x_0 wie in der Aussage eine lokale Extremstelle, die (a) und (b) erfüllt. OBdA dürfen wir annehmen, dass x_0 eine lokale Maximumsstelle von f ist. (Falls x_0 nämlich eine lokale Minimumsstelle von f ist können wir $-f$ betrachten, vgl. Bemerkung 7.2.3). Etwa sei $\delta > 0$ so, dass

$$\forall x \in D \cap B_\delta(x_0) : f(x) \leq f(x_0). \quad (340)$$

Da f bei x_0 differenzierbar ist gilt (vgl. Definition 7.0.1)

$$f'(x_0) = \lim_{\substack{x \rightarrow x_0 \\ x \in D \setminus \{x_0\}}} \frac{f(x) - f(x_0)}{x - x_0}.$$

Definiere nun für $n \in \mathbb{N}$ die Zahlen $x_n := x_0 + \frac{1}{n}$ und $y_n := x_0 - \frac{1}{n}$.

ZWISCHENBEHAUPTUNG. Es gibt ein $N \in \mathbb{N}$ so, dass $(x_n)_{n \geq N}, (y_n)_{n \geq N} \subseteq D \cap B_\delta(x_0) \setminus \{x_0\}$.

Dass x_n und y_n stets verschieden von x_0 sind folgt direkt aus der Definition. Noch zu zeigen ist, dass es ein $N \in \mathbb{N}$ gibt mit $x_n, y_n \in D$ für alle $n \geq N$. Da x_0 nach Voraussetzung (b) ein innerer Punkt von D ist gibt es $r > 0$ so, dass $B_r(x_0) \subseteq D$. Definiere nun $N := \lceil \frac{1}{\min\{r, \delta\}} \rceil + 1$. Für $n \geq N$ gilt dann

$$|x_n - x_0| = |\frac{1}{n}| = \frac{1}{n} \leq \frac{1}{N} < \frac{1}{\frac{1}{\min\{r, \delta\}}} = \min\{r, \delta\}.$$

Das bedeutet zum Einen, dass $x_n \in B_r(x_0) \subseteq D$ und zum Anderen, dass $x_n \in B_\delta(x_0)$. Wir folgern also $x_n \in D \cap B_\delta(x_0)$. Analog zeigt man für $n \geq N$ auch $y_n \in D \cap B_\delta(x_0)$.

ZWISCHENBEHAUPTUNG ENDE.

Es gilt nun $(x_n)_{n \geq N}, (y_n)_{n \geq N} \subseteq D \setminus \{x_0\}$ und $\lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} y_n = x_0$.

SCHRITT 1. Wir zeigen $f'(x_0) \leq 0$.

Dazu berechnen wir zunächst

$$f'(x_0) = \lim_{\substack{x \rightarrow x_0 \\ x \in D \setminus \{x_0\}}} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{n \rightarrow \infty} \frac{f(x_n) - f(x_0)}{x_n - x_0}. \quad (341)$$

Wir analysieren nun das Vorzeichen vom Zähler und Nenner des Ausdrucks auf der rechten Seite der obigen Gleichung. Zuerst zum Nenner. Es gilt $x_n - x_0 = \frac{1}{n} > 0$. Nun zum Zähler. Für $n \geq N$ gilt nach der Zwischenbehauptung auch $x_n \in D \cap B_\delta(x_0)$ und daher impliziert (340) auch $f(x_n) \leq f(x_0)$, d.h. $f(x_n) - f(x_0) \leq 0$. Also gilt

$$\frac{\overbrace{f(x_n) - f(x_0)}^{\leq 0, \text{vgl. 340}}}{\underbrace{x_n - x_0}_{> 0}} \leq 0 \quad \forall n \geq N.$$

Die Grenzwertvergleichsregeln (Proposition 4) implizieren dann

$$\lim_{n \rightarrow \infty} \frac{f(x_n) - f(x_0)}{x_n - x_0} \leq 0.$$

Das bedeutet nach (341), dass

$$f'(x_0) \leq 0. \quad (342)$$

SCHRITT 2. Wir zeigen $f'(x_0) \geq 0$.

Hierfür verwenden wir nun die Folge $(y_n)_{n \in \mathbb{N}}$ anstatt $(x_n)_{n \in \mathbb{N}}$. Wie in (341) zeigt man, dass

$$f'(x_0) = \lim_{\substack{x \rightarrow x_0 \\ x \in D \setminus \{x_0\}}} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{n \rightarrow \infty} \frac{f(y_n) - f(x_0)}{y_n - x_0}. \quad (343)$$

Nun gilt $y_n - x_0 = -\frac{1}{n} < 0$ für alle $n \in \mathbb{N}$. Hingegen gilt für $n \geq N$ wiederum $f(y_n) \leq f(x_0)$, da $y_n \in D \cap B_\delta(x_0)$ (vgl. (340) und Zwischenbehauptung). Also hat man für $n \geq N$, dass $f(y_n) - f(x_0) \leq 0$ und folgert

$$\frac{\overbrace{f(y_n) - f(x_0)}^{\leq 0, \text{ vgl. 340}}}{\underbrace{y_n - x_0}_{< 0}} \geq 0 \quad \forall n \geq N.$$

Die Grenzwertvergleichsregeln (Proposition 4) implizieren dann wieder

$$\lim_{n \rightarrow \infty} \frac{f(y_n) - f(x_0)}{y_n - x_0} \geq 0.$$

Aus (343) folgt dann auch

$$f'(x_0) \geq 0. \quad (344)$$

Gleichung (342) und (344) gemeinsam lassen nur den Schluss zu, dass $f'(x_0) = 0$.

Auf die Voraussetzung (a) kann beim obigen Satz klarerweise nicht verzichtet werden. Denn schließlich könnte man ja sonst die Ableitung gar nicht bilden. Dass auf Voraussetzung (b) nicht verzichtet werden kann ist etwas schwieriger zu begreifen. Und doch ist auch diese Voraussetzung unverzichtbar.

Bemerkung 7.2.7

Auf Voraussetzung (b) kann in Satz 7.2.6 nicht verzichtet werden. Betrachte dazu $f : [-1, 1] \rightarrow \mathbb{R}, x \mapsto x$. Dann nimmt f bei $x_0 = 1$ ein lokales (sogar globales) Maximum an. Prüfen wir nun die Voraussetzungen des Satzes 7.2.6:

(a) f ist bei x_0 differenzierbar ✓.

(b) x_0 ist innerer Punkt von $D = [-1, 1]$ ✗.

Voraussetzung (b) ist also nicht erfüllt. Die Aussage von Satz 7.2.6 ist aber nicht mehr zutreffend: Denn nach Beispiel 7.0.3 gilt

$$f'(x_0) = 1 \neq 0.$$

In Satz 7.2.6 kann man also auf Voraussetzung (b) nicht verzichten.

(Lokale) Maximumsstellen, die keine inneren Punkte von D sind, nennen wir (lokale) *Rand-maximumsstellen*. Analog nennen wir (lokale) Minimumsstellen, die keine inneren Punkte

von D sind, (lokale) *Randminimumsstellen*. Allgemein spricht man von *Randextrema*.

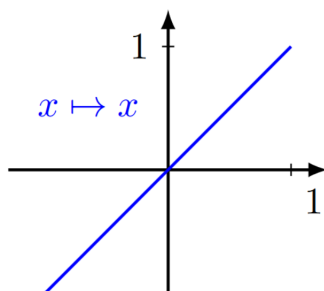


Abbildung 29: Die Funktion aus Beispiel 7.2.7

Um nun nach lokalen Extrema einer Funktion $f : D \rightarrow \mathbb{R}$ zu suchen, die keine Randextrema sind, kann man also Lösungen der Gleichung

$$f'(x) = 0 \quad (345)$$

betrachten. Allerdings ist nicht jede Lösung $x_0 \in D$ dieser Gleichung auch gleich eine lokale Extremstelle! Auch dann nicht, wenn die Voraussetzungen (a) und (b) von Satz 7.2.6 erfüllt sind!

Das behauptet unser Satz 7.2.6 auch gar nicht. Denn schließlich behandelt der Satz nur die „ $\implies$ “ Implikationsrichtung

$$x_0 \text{ ist lokale Extremstelle von } f \implies f'(x_0) = 0.$$

Die Implikation „ $\impliedby$ “ ist nicht richtig, wie das folgende Beispiel zeigt.

Beispiel 7.2.8

Im Allgemeinen gibt es mehr Lösungen von (345) als innere lokale Extrema. Betrachte die Funktion $f : [-1, 1] \rightarrow \mathbb{R}, x \mapsto x^3$. Dann gilt

$$\forall x \in [-1, 1] : \quad f'(x) = 3x^2$$

und somit auch

$$f'(0) = 0.$$

Die Stelle $x_0 = 0$ erfüllt außerdem Voraussetzung (a) und (b) von Satz 7.2.6. Jedoch ist $x_0 = 0$ keine lokale Extremstelle von f .

In der Tat: Wäre $x_0 = 0$ eine lokale Maximumsstelle von f , so gäbe es nach Definition 7.2.1 ein $\delta > 0$ mit

$$f(x) \leq f(0) = 0 \quad \forall x \in D \cap B_\delta(0). \quad (346)$$

Nun gilt $D \cap B_\delta(0) = [-1, 1] \cap (-\delta, \delta)$. Definiere nun $\delta' := \min\{\delta, 1\}$. Man sieht leicht ein,

dass $(-\delta', \delta') \subseteq D \cap B_\delta(0)$. Insbesondere wäre $\frac{\delta'}{2} \in D \cap B_\delta(0)$. Nach (347) wäre dann

$$0 = f(0) \geq f\left(\frac{\delta'}{2}\right) = \left(\frac{\delta'}{2}\right)^3 = \frac{(\delta')^3}{8} > 0.$$

Ein Widerspruch. Also kann 0 keine lokale Maximumsstelle gewesen sein.

Wäre hingegen $x_0 = 0$ eine lokale Minimumsstelle von f , so gäbe es nach Definition 7.2.1 ein $\delta > 0$ mit

$$f(x) \geq f(0) = 0 \quad \forall x \in D \cap B_\delta(0). \quad (347)$$

Wiederum gilt $D \cap B_\delta(0) = [-1, 1] \cap (-\delta, \delta)$. Definiere abermals $\delta' := \min\{\delta, 1\}$ und beobachte dass $(-\delta', \delta') \subseteq D \cap B_\delta(0)$. Insbesondere wäre $-\frac{\delta'}{2} \in D \cap B_\delta(0)$. Nach (347) wäre dann

$$0 = f(0) \leq f\left(-\frac{\delta'}{2}\right) = \left(-\frac{\delta'}{2}\right)^3 = -\frac{(\delta')^3}{8} < 0.$$

Ein Widerspruch. Also kann 0 keine lokale Maximumsstelle gewesen sein.

Man folgert: $x_0 = 0$ ist keine lokale Extremstelle.

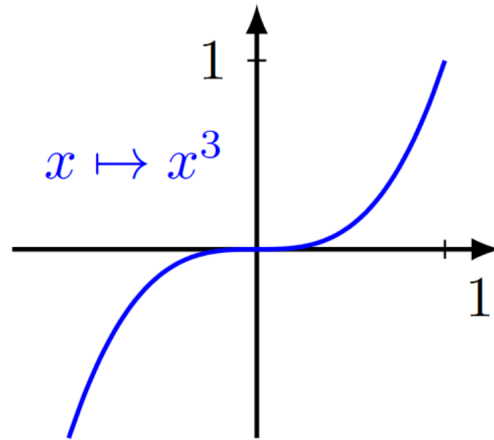


Abbildung 30: Die Funktion aus Beispiel 7.2.8

FAZIT. Das Verschwinden der ersten Ableitung ist für das Vorliegen einer inneren lokalen Extremstelle *notwendig*, aber nicht *hinreichend*!

Nullstellen der ersten Ableitung von f sind also nur *Kandidaten* für lokale Extrema. Das bedeutet: Alle lokalen Extrema sind unter diesen Nullstellen von f' zu suchen. Jedoch sind nicht alle Nullstellen von f' gleich lokale Extrema. Nullstellen von f' , die keine lokalen Extrema sind, nennt man im Allgemeinen *Sattelpunkte*.

Beispiel 7.2.9

Gegeben sei für $b > 1$

$$f : [0, b] \rightarrow \mathbb{R}, x \mapsto x^2(1-x)^2.$$

Aufgabe. Zeige, dass $\max_{x \in [0, b]} f(x)$ existiert und berechne den Wert (in Abhängigkeit von b).

Die Existenz von $\max_{x \in [0, b]} f(x)$ folgt direkt aus dem Satz von Weierstraß (vgl. Satz 6.2.12), weil f auf $[0, b]$ stetig ist und $[0, b]$ ein kompaktes Intervall ist.

SCHRITT 1. Kandidaten für innere lokale Extremstellen

Wir berechnen mit Satz 7.2.6 Kandidaten für innere lokale Extremstellen.

$$\begin{aligned} f'(x) &= \frac{d}{dx}(x^2(1-x)^2) \stackrel{\text{Prop 7.1.1}}{=} 2x(1-x)^2 + x^2 \frac{d}{dx}(1-x)^2 \\ &\stackrel{\text{Prop 7.1.5}}{=} 2x(1-x)^2 + x^2 \cdot 2(1-x)(-1) \\ &= 2x(1-x)[(1-x) - x] = 2x(1-x)(1-2x). \end{aligned}$$

Man folgert

$$f'(x) = 0 \iff x(1-x)(1-2x) = 0 \iff x = 0 \vee 1-x = 0 \vee 1-2x = 0.$$

Somit sind $x_0 = 0$, $x_1 = \frac{1}{2}$ und $x_2 = 1$ die Nullstellen von f' . Allerdings sind nur x_1 und x_2 innere Punkte von $[0, b]$. Daher sind x_1 und x_2 Kandidaten für lokale Extrema. Beachte

$$f(x_1) = f\left(\frac{1}{2}\right) = \frac{1}{2^2}\left(1 - \frac{1}{2}\right)^2 = \frac{1}{4} \cdot \frac{1}{4} = \frac{1}{16}, \quad f(x_2) = f(1) = 0. \quad (348)$$

SCHRITT 2. Kandidaten für Randextrema.

Wir vergleichen Funktionswerte der *Randpunkte* 0 und b mit den Funktionswerten der Kandidaten in (348). Dazu:

$$f(0) = 0^2(1-0)^2 = 0 \quad \text{und} \quad f(b) = b^2(1-b)^2.$$

SCHRITT 3. Vergleich der Kandidaten.

Wir haben gesehen: f nimmt auf $[0, b]$ sein globales Maximum an. Außerdem muss jedes globale Maximum ein lokales Maximum sein. Lokale Maxima können sich, wie wir in Schritt 1 und 2 gesehen haben, aber nur bei $x_1, x_2, 0, b$ befinden. Jetzt kommen also nur noch 4 Werte als Maximum infrage: Nämlich $f(x_1), f(x_2), f(0)$ und $f(b)$. Beachte:

$$f(0) = f(x_2) = 0, \quad f(x_1) > 0 \quad \text{und} \quad f(b) > 0.$$

Somit kann $\max_{x \in [0, b]} f(x)$ nicht bei 0 und x_2 angenommen werden, denn offensichtlich werden in $[0, b]$ ja auch noch größere Werte als 0 angenommen. Somit kommen nur noch $f(x_1)$ und $f(b)$ infrage. Man sieht

$$\max_{x \in [0, b]} f(x) \underset{(*)}{=} \max\{f(x_1), f(b)\} = \max\left\{\frac{1}{16}, b^2(1-b)^2\right\}. \quad (349)$$

Beachte: Ohne den Satz von Weierstraß (Satz 6.2.12) wäre die Argumentation in $(*)$ nicht möglich gewesen. (sonst hätte es nämlich auch passieren können, dass $\max_{x \in [0, b]} f(x)$ nicht existiert und daher keiner der Kandidaten den maximalen Wert liefert).

Eigentlich wären wir jetzt fertig. Es drängt sich aber noch die Frage auf, für welche Werte von $b > 1$ welcher der beiden Werte in (349) der Maximale ist. Dazu berechnen wir zunächst ein $b_* > 1$ so, dass $b_*^2(1-b_*)^2 = \frac{1}{16}$. Dazu

$$b^2(1-b)^2 = \frac{1}{16} \quad \Longleftrightarrow \quad b(1-b) = \pm \frac{1}{4}$$

Da $b > 1$ gilt $b(1-b) < 0$ und somit kann nur $\pm = -$ zutreffen. Man erhält

$$b(1-b) = -\frac{1}{4}$$

und formt äquivalent um zu $b^2 - b - \frac{1}{4} = 0$. Diese Gleichung gilt für

$$b = \frac{1}{2} \pm \sqrt{\frac{1}{4} + \frac{1}{4}} = \frac{1}{2} \pm \sqrt{\frac{1}{2}}.$$

Da jedoch $b_* > 1$ ist, ist hier nur $\pm = +$ zulässig und man folgert

$$b_* = \frac{1}{2} + \frac{1}{\sqrt{2}} = \frac{1 + \sqrt{2}}{2}$$

(Beachte: Da $\sqrt{2} > 1$ gilt auch wirklich $b_* = \frac{1+\sqrt{2}}{2} > \frac{1+1}{2} = 1$!) Man überzeugt sich leicht davon, dass

$$b^2(1-b)^2 \begin{cases} < \frac{1}{16} & b \in [1, b_*) \\ = \frac{1}{16} & b = b_* \\ > \frac{1}{16} & b \in (b_*, \infty) \end{cases}. \quad (a)$$

Zusammengefasst gilt mit (349)

$$\max_{x \in [0, b]} f(x) = \begin{cases} \frac{1}{16} & b \leq \frac{1+\sqrt{2}}{2} \\ b^2(1-b^2) & b > \frac{1+\sqrt{2}}{2} \end{cases}$$

^aHierfür braucht man keine Kurvendiskussion – es reicht völlig der Zwischenwertsatz!

15.04.2026

7.3 Mittelwertsätze und Monotonieverhalten

Mit dem sogenannten *ersten Mittelwertsatz der Differentialrechnung*⁶ werden wir uns in die Lage versetzen, einen Funktionsgraphen auf Monotonie zu diskutieren. Geometrisch setzt er die *Sekantensteigung* und *Tangentensteigung* miteinander in Relation.

Satz 7.3.1 — ERSTER MITTELWERTSATZ DER DIFFERENTIALRECHNUNG

Es sei $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ so, dass

- (a) f ist stetig auf $[a, b]$
- (b) f ist differenzierbar auf (a, b)

Dann gibt es ein $\xi \in (a, b)$ so, dass

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}. \quad (350)$$

Bemerkung 7.3.2

Ein paar Anmerkungen zur Aussage des Mittelwertsatzes.

- (i) Voraussetzung (a) kann abgeschwächt werden. Es genügt zu fordern, dass f bei den Randpunkten a und b stetig ist. In der Tat: Auf (a, b) ist f ja schließlich differenzierbar und somit auch automatisch stetig! (Vgl. Proposition 7.0.7)
- (ii) Im weiteren Verlauf wollen wir den Mittelwertsatz verwenden, um *Abschätzungen* zu zeigen. Hierzu ist es ganz nützlich die Formel

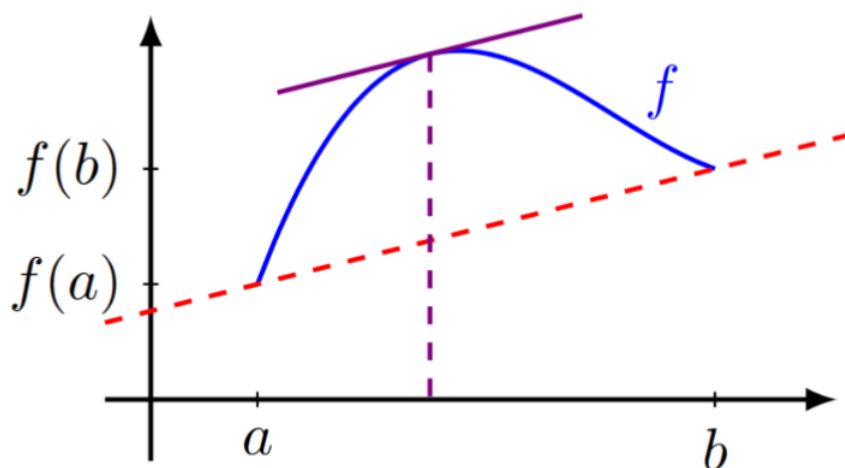
$$f'(\xi) = \frac{f(b) - f(a)}{b - a}$$

aus Satz 7.3.1 umgestellt zu betrachten, nämlich als

$$f(b) - f(a) = f'(\xi)(b - a) \quad \text{für ein } \xi \in (a, b). \quad (351)$$

⁶Auch für diesen Satz bietet GEOGEBRA eine gute Verständnishilfe.

Bevor wir den Beweis sehen besprechen wir die geometrische Interpretation. Der Wert $\frac{f(b)-f(a)}{b-a}$ ist die Steigung der Sekante von f zwischen a und b . Also wird in (350) behauptet, dass diese Sekantensteigung mit der Tangentensteigung an einem gewissen Punkt $\xi \in (a, b)$, den wir nicht kennen, übereinstimmt.



Abbildungung 31: Ein Bild zum Mittelwertsatz. Die Zwischenstelle ξ ist in lila markiert. Die Grafik hat man auch nochmal interaktiv auf GEOGEBRA.

Der Punkt x_0 kann auf dem Graphen Γ_f stets so gefunden werden: Man legt ein Geodreieck auf die Sekante zwischen a und b (in Abbildung 31 in rot) und verschiebt diese Sekante so lange nach oben/unten, bis man den Graphen *nur noch berührt*. Diese verschobene Sekante ist dann eine Tangente an Γ_f , welche dieselbe Steigung hat. Der Punkt ξ ist also der x -Wert dieses Berührungspunktes.

Für den Beweis des Mittelwertsatzes benötigen wir vorbereitend ein wichtiges Lemma, den sogenannten *Satz von Rolle*⁷, welcher eigentlich bereits ein Spezialfall des Mittelwertsatzes ist.

Lemma 7.3.3 — SATZ VON ROLLE

Es sei $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ so, dass

- (a) f ist stetig auf $[a, b]$,
- (b) f ist differenzierbar auf (a, b) ,
- (c) $f(a) = f(b)$.

Dann gibt es ein $\xi \in (a, b)$ mit $f'(\xi) = 0$.

⁷Die Aussage ist auch nochmal sehr anschaulich in GEOGEBRA dargestellt.

Beweis. Da f auf dem kompakten Intervall $[a, b]$ stetig ist, nimmt f dem Satz von Weierstraß (Satz 6.2.12) ein Maximum und Minimum an. Mit anderen Worten gibt es $x_1, x_2 \in [a, b]$ mit

$$f(x_1) = \max_{x \in [a, b]} f(x) \quad \text{und} \quad f(x_2) = \min_{x \in [a, b]} f(x). \quad (352)$$

Nun unterscheiden wir verschiedene Fälle

FALL 1. Einer der beiden Punkte x_1, x_2 ist ein innerer Punkt von $[a, b]$.

Ist etwa x_1 innerer Punkt von $[a, b]$ so gilt $x_1 \in (a, b)$, vgl. Beispiel 7.2.5. Nun ist also x_1 eine (globale und somit auch lokale) Extremstelle von f mit

- (a) f ist differenzierbar bei x_1 (da $x_1 \in (a, b)$) und
- (b) x_1 ist innerer Punkt von $[a, b]$.

Aus Satz 7.2.6 folgt dann $f'(x_1) = 0$. Man kann also $\xi = x_1$ wählen und erhält die Behauptung. Ist stattdessen x_2 ein innerer Punkt von $[a, b]$ so folgt mit derselben Begründung $f'(x_2) = 0$.

FALL 2. Keiner der beiden Punkte ist ein innerer Punkt von $[a, b]$.

Somit gilt also $x_1, x_2 \notin \overset{\circ}{[a, b]} = (a, b)$, vgl. Beispiel 7.2.5. Es muss also gelten, dass x_1 und x_2 am Rand des Intervalles liegen, d.h. $x_1, x_2 \in \{a, b\}$. Es gilt also $x_1 = x_2 = a$ oder $x_1 = x_2 = b$ oder $(x_1 = a \text{ und } x_2 = b)$ oder $(x_2 = a \text{ und } x_1 = b)$. Da $f(a) = f(b)$ folgert man in jedem Fall $f(x_1) = f(x_2)$. Aus (352) folgt dann wiederum

$$\max_{x \in [a, b]} f(x) = \min_{x \in [a, b]} f(x).$$

Aus Bemerkung 6.2.19 lernt man dann, dass f auf $[a, b]$ konstant sein muss. Nach Beispiel 7.0.2 gilt dann $f'(x) = 0$ für alle $x \in (a, b)$. Somit gilt zum Beispiel für $\xi = \frac{a+b}{2}$, dass $f'(\xi) = 0$.

Nun sind wir in der Lage den Mittelwertsatz zu beweisen. Wir werden dabei die Aussage durch betrachten einer *geschickten Hilfsfunktion* auf den Satz von Rolle zurückführen.

Beweis von Satz 7.3.1. Wir führen die Aussage auf den Satz von Rolle zurück. Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf (a, b) . Definiere die Hilfsfunktion $h : [a, b] \rightarrow \mathbb{R}$ durch

$$h(x) := f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Wir wollen nun den Satz von Rolle (Lemma 7.3.3) auf h anwenden. Dazu prüfen wir die Voraussetzungen.

- (a) h ist als Komposition stetiger Funktionen stetig auf $[a, b]$,
- (b) h ist als Komposition differenzierbarer Funktionen auf (a, b) auch wieder differenzierbar auf (a, b) , vgl Proposition 7.1.1.
- (c) $h(a) = h(b)$, denn

$$h(a) = f(a) - \frac{f(b) - f(a)}{b - a}(a - a) = f(a),$$

$$h(b) = f(b) - \frac{f(b) - f(a)}{b - a}(b - a) = f(b) - (f(b) - f(a)) = f(a).$$

Aus dem Satz von Rolle folgt, dass es in $\xi \in (a, b)$ gibt mit $h'(\xi) = 0$. Nun gilt nach Proposition 7.1.1

$$h'(x) = f'(x) - \frac{f(b) - f(a)}{b - a} \cdot 1 \quad \forall x \in (a, b).$$

Man erhält

$$0 = h'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}$$

und somit auch

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Als nächstes sehen wir zwei Beispiele für Abschätzungen, die durch den Mittelwertsatz möglich gemacht werden.

Beispiel 7.3.4

Behauptung. Für alle $x \in (0, \infty)$ gilt $\sin(x) \leq x$.

Beweis. Sei $x \in (0, \infty)$. Da $\sin(0) = 0$ (vgl. Proposition 5.5.3) gilt

$$\sin(x) = \sin(x) - \sin(0).$$

Da nun $\sin$ auf $[0, x]$ stetig und auf $(0, x)$ differenzierbar ist, kann man den Mittelwertsatz mit $[a, b] = [0, x]$ und $f = \sin$ verwenden. Man folgert aus (351), dass es ein $\xi \in (0, x)$ gibt mit

$$\sin(x) - \sin(0) = \sin'(\xi)(x - 0).$$

Also gilt

$$\sin(x) = \sin(x) - \sin(0) = \sin'(\xi)(x - 0) \stackrel{\text{Kor7.0.11}}{=} \cos(\xi)x.$$

Da $|\cos(\xi)| = |\operatorname{Re}(e^{i\xi})| \leq |e^{i\xi}| = 1$ (vgl. Proposition 5.5.2) gilt

$$\sin(x) = \cos(\xi)x \leq |\cos(\xi)x| = |\cos(\xi)|x \leq x.$$

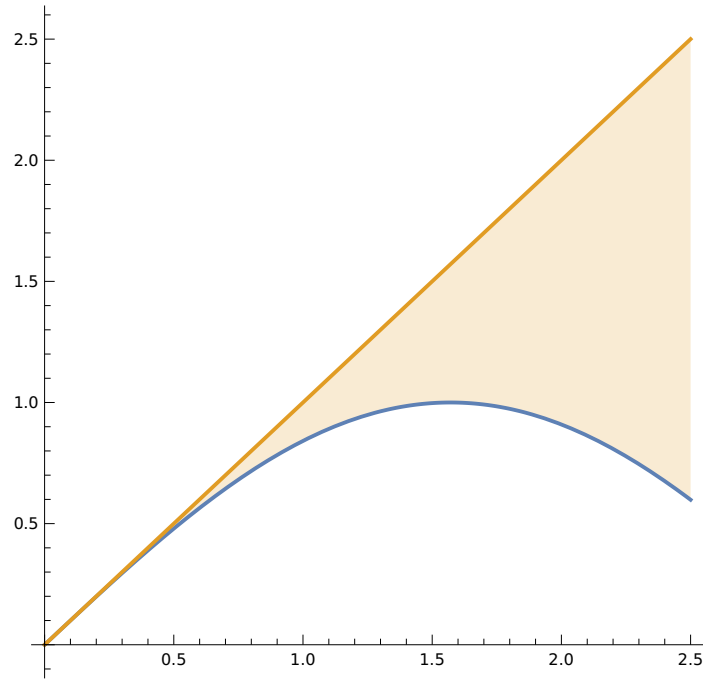


Abbildung 32: Plot der Graphen von $x \mapsto x$ (gelb) und $x \mapsto \sin(x)$ (blau). Man sieht: Für kleine Werte von x ist die Abschätzung in Beispiel 7.3.4 ziemlich genau.

Eine weitere schöne Anwendung ist, dass man das *Wachstum* der Partialsummen der harmonischen Reihe verstehen kann.

Beispiel 7.3.5

Behauptung. Für alle $n \in \mathbb{N}$ mit $n \geq 2$ gilt $\log(n) \leq \sum_{k=1}^n \frac{1}{k} \leq \log(n) + 1$.

Beweis. Für $n \in \mathbb{N}$ mit $n \geq 2$ gilt stets

$$\log(n) = \log(n) - 0 = \log(n) - \log(1) \stackrel{\text{Teleskopsumme}}{=} \sum_{k=2}^n \log(k) - \log(k-1). \quad (353)$$

Nun gibt es nach dem Mittelwertsatz für alle $k \in \{2, \dots, n\}$ ein $\xi_k \in (k-1, k)$ so, dass

$$\log(k) - \log(k-1) \stackrel{(351) \text{ mit } b=k, a=k-1}{=} \log'(\xi_k)[k - (k-1)] \stackrel{\text{Bsp 7.1.11}}{=} \frac{1}{\xi_k} \cdot [k - (k-1)].$$

Da $k - (k-1) = 1$ gilt also

$$\log(k) - \log(k-1) = \frac{1}{\xi_k} \quad \forall k \in \{2, \dots, n\}.$$

Also gilt nach (353)

$$\log(n) = \sum_{k=2}^n \frac{1}{\xi_k}. \quad (354)$$

Mithilfe von (354) findet man

$$\log(n) = \sum_{k=2}^n \frac{1}{\xi_k} \leq \sum_{k=2}^n \frac{1}{k-1} \stackrel{\text{Indexverschiebung}}{=} \sum_{k=1}^{n-1} \frac{1}{k} \leq \sum_{k=1}^n \frac{1}{k}. \quad (355)$$

Außerdem ergibt sich als weitere Konsequenz aus (354)

$$\log(n) = \sum_{k=2}^n \frac{1}{\xi_k} \geq \sum_{k=2}^n \frac{1}{k} = \left(\sum_{k=1}^n \frac{1}{k} \right) - \frac{1}{1} = \sum_{k=1}^n \frac{1}{k} - 1.$$

Formt man um so ergibt sich

$$\sum_{k=1}^n \frac{1}{k} \leq \log(n) + 1.$$

Zusammen mit (355) ergibt sich

$$\log(n) \leq \sum_{k=1}^n \frac{1}{k} \leq \log(n) + 1.$$

Man kann den ersten Mittelwertsatz noch weiter verallgemeinern. Diese Verallgemeinerung nennt sich dann *zweiter Mittelwertsatz der Differentialrechnung*

Satz 7.3.6 — ZWEITER MITTELWERTSATZ DER DIFFERENTIALRECHNUNG

Es sei $a, b \in \mathbb{R}$ mit $a < b$ und $f, g : [a, b] \rightarrow \mathbb{R}$ so, dass

- (a) f, g sind stetig auf $[a, b]$
- (b) f, g sind differenzierbar auf (a, b)
- (c) $g'(x) \neq 0$ für alle $x \in (a, b)$

Dann gibt es ein $\xi \in (a, b)$ so, dass

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Beweis. Wir verfahren genau wie beim Beweis von Satz 7.3.1, nur mit einer etwas modifizierten Hilfsfunktion. Es seien $f, g : [a, b] \rightarrow \mathbb{R}$ wie in der Aussage.

Zwischenbehauptung. $g(a) \neq g(b)$.

Wäre $g(a) = g(b)$ so gäbe es nach dem Satz von Rolle (Lemma 7.3.3) ein $x_0 \in (a, b)$ mit

$g'(x_0) = 0$. Dies widerspricht aber Voraussetzung (c).

Zwischenbehauptung Ende.

Definiere die Hilfsfunktion $h : [a, b] \rightarrow \mathbb{R}$ durch

$$h(x) := f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(x) - g(a)).$$

Wir wollen nun den Satz von Rolle (Lemma 7.3.3) auf h anwenden. Dazu prüfen wir die Voraussetzungen.

- (a) h ist als Komposition stetiger Funktionen stetig auf $[a, b]$,
- (b) h ist als Komposition differenzierbarer Funktionen auf (a, b) auch wieder differenzierbar auf (a, b) , vgl. Proposition 7.1.1.
- (c) $h(a) = h(b)$, denn

$$\begin{aligned} h(a) &= f(a) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(a) - g(a)) = f(a), \\ h(b) &= f(b) - \frac{f(b) - f(a)}{g(b) - g(a)}(g(b) - g(a)) = f(b) - (f(b) - f(a)) = f(a). \end{aligned}$$

Aus dem Satz von Rolle folgt, dass es ein $\xi \in (a, b)$ gibt mit $h'(\xi) = 0$. Nun gilt nach Proposition 7.1.1

$$h'(x) = f'(x) - \frac{f(b) - f(a)}{g(b) - g(a)} \cdot g'(x) \quad \forall x \in (a, b).$$

Man erhält

$$0 = h'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(\xi).$$

Teilt man durch $g'(\xi)$ und formt um so erhält man

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Beachte: Man darf durch $g'(\xi)$ teilen weil nach Voraussetzung (c) gilt $g'(\xi) \neq 0$.

Bemerkung 7.3.7

Der zweite Mittelwertsatz (Satz 7.3.6) verallgemeinert in der Tat den ersten Mittelwertsatz. In der Tat: Im Spezialfall $g = \text{id}_{[a,b]}$ ergibt sich der klassische erste Mittelwertsatz (Satz 7.3.1).

Kurvendiskussion mit dem ersten Mittelwertsatz

In diesem Abschnitt sprechen wir darüber, wie der Mittelwertsatz verwendet werden kann, um das Verhalten von Funktionen zu diskutieren. Im Folgenden seien stets $a, b \in \mathbb{R}$ und $a < b$. Eine erste wichtige Erkenntnis ist, dass die *Konstanz* einer Funktion anhand der Ableitung abgelesen werden kann.

Proposition 7.3.8 — KONSTANZSATZ

Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion, die auf dem offenen Intervall (a, b) differenzierbar ist. Dann gilt

$$f'(x) = 0 \quad \text{für alle } x \in (a, b) \quad \Longleftrightarrow \quad f \text{ ist auf } [a, b] \text{ konstant.}^a$$

^aHierbei nennen wir eine Funktion $f : D \rightarrow \mathbb{R}$ *konstant*, falls $f(D)$ nur aus einem Element besteht. Mit anderen Worten: Es gibt ein $c \in \mathbb{R}$ mit $f(x) = c$ für alle $x \in D$.

Beweis. „ $\Leftarrow$ “ folgt direkt aus Beispiel 7.0.2: Eine konstante Funktion hat verschwindende Ableitung. Nun zu „ $\Rightarrow$ “. Es gelte

$$f'(x) = 0 \quad \text{für alle } x \in (a, b) \tag{356}$$

Es sei außerdem $y \in [a, b]$ beliebig. Um zu zeigen, dass f konstant ist, muss gezeigt werden, dass der Wert $f(y)$ nicht von y abhängt. Hierzu:

$$\text{ZZ: } f(y) = f(a).$$

Das sieht man so: Für $y = a$ ist die Behauptung klar. Für $y \in (a, b]$ beobachten wir zunächst, dass f auf $[a, y]$ stetig und auf (a, y) differenzierbar ist. Aus dem Mittelwertsatz (bzw. aus (351)) folgt

$$\exists \xi \in (a, y) : \quad f(y) - f(a) = f'(\xi)(y - a). \tag{357}$$

Wegen (356) und weil $\xi \in (a, y) \subseteq (a, b)$ gilt, dass $f'(\xi) = 0$. Mit (357) folgern wir

$$f(y) - f(a) = 0,$$

d.h. $f(y) = f(a)$, was zu zeigen war.

Bemerkung 7.3.9

Der obige Konstanzsatz kann auch verallgemeinert werden, sodass als Definitionsbereich allgemeinere Intervalle der Form $[a, b)$, $(a, b]$, (a, b) und auch unbeschränkte Intervalle zugelassen werden können. Genauer gilt:

Verallgemeinerung. (von Proposition 7.3.8) Ist $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ eine

stetige Funktion, die auf $\overset{\circ}{I}$ differenzierbar ist, so gilt

$$f'(x) = 0 \text{ f\"ur alle } x \in \overset{\circ}{I} \quad \Longleftrightarrow \quad f \text{ ist auf } I \text{ konstant}$$

Beweisskizze. Es sei $I \subset \mathbb{R}$ ein beliebiges Intervall. „ $\Leftarrow$ “ folgt nach wie vor direkt aus Beispiel 7.0.2. Nun zu „ $\Rightarrow$ “. Nimm $x_1, x_2 \in I$ beliebig mit $x_1 \neq x_2$, OBdA $x_1 < x_2$. Wir zeigen $f(x_1) = f(x_2)$. Dazu: Da $[x_1, x_2] \subseteq I$ ist f auf $[x_1, x_2]$ stetig. Außerdem überzeugt man sich leicht, dass $(x_1, x_2) \subseteq \overset{\circ}{I}$, weswegen f auf (x_1, x_2) auch differenzierbar ist. Nach Proposition 7.3.8 ist f dann auf $[x_1, x_2]$ konstant. Das bedeutet auch $f(x_1) = f(x_2)$. Da $x_1, x_2 \in I$ beliebig waren folgt dass f konstant ist.

Achtung: Der Definitionsbereich von f muss ein Intervall sein, sonst gilt der Konstanzsatz (bzw seine Verallgemeinerung) nicht. Betrachte etwa $f : (1, 2) \cup (3, 4) \rightarrow \mathbb{R}$ gegeben durch

$$f(x) = \begin{cases} 1 & x \in (1, 2) \\ 18 & x \in (3, 4) \end{cases}$$

Dann ist $f'(x) = 0$ für alle $x \in (1, 2) \cup (3, 4)$, aber f ist nicht konstant auf $(1, 2) \cup (3, 4)$.

Wir sehen im nachfolgenden Beispiel, wie der Konstanzsatz benutzt werden kann um *andere Darstellungen* für Funktionen zu finden.

Beispiel 7.3.10 — REIHENDARSTELLUNG DES LOGARITHMUS

Wir beweisen mithilfe des Konstanzsatzes (Proposition 7.3.8), dass

$$\forall x \in (-1, 1) : \quad \log(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k.$$

Beweis. Definiere $f : (-1, 1) \rightarrow \mathbb{R}$ durch

$$f(x) := \log(1+x) - \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k. \quad (358)$$

Wir werden zuerst zeigen, dass f konstant ist. Dazu bilden wir für $x \in (-1, 1)$

$$f'(x) = \frac{d}{dx} \log(1+x) - \sum_{k=1}^{\infty} \frac{(-1)^k}{k} x^{k-1} \stackrel{\text{Prop 7.1.1}}{=} \frac{d}{dx} \log(1+x) - \frac{d}{dx} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k \quad (359)$$

Mit Beispiel 7.1.11 und der Kettenregel ergibt sich

$$\frac{d}{dx} \log(1+x) = \frac{1}{1+x} \cdot 1 = \frac{1}{1+x}$$

Mit Proposition 7.0.10 ergibt sich für $x \in (-1, 1)^a$

$$\begin{aligned} \frac{d}{dx} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k &\stackrel{\text{Prop 7.0.10}}{=} \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} k x^{k-1} = \sum_{k=1}^{\infty} (-1)^{k-1} x^{k-1} \\ &= \sum_{k=1}^{\infty} (-x)^{k-1} = \sum_{k=0}^{\infty} (-x)^k \stackrel{\text{geom. Reihe Bsp 5.0.5}}{=} \frac{1}{1 - (-x)} = \frac{1}{1+x}. \end{aligned}$$

Setzt man die beiden vorigen Gleichungen in (359) ein so erhält man

$$f'(x) = 0 \quad \forall x \in (-1, 1).$$

Aus dem Konstanzsatz (bzw. aus der Verallgemeinerung in Bemerkung 7.3.9) folgt, dass f auf $(-1, 1)$ konstant ist. Das bedeutet

$$\exists c \in \mathbb{R} \quad \forall x \in (-1, 1) : \quad f(x) = c \quad . \quad (360)$$

$$\text{Noch } \mathbb{Z}\text{:} \quad c = 0.$$

Hierzu setzen wir $x = 0$ ein. Man erhält

$$c = f(0) \stackrel{(358)}{=} \log(1+0) - \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} 0^k = 0 - \sum_{k=1}^{\infty} 0 = 0.$$

Somit gilt $c = 0$ und daher nach (360)

$$0 = f(x) = \log(1+x) - \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k \quad \forall x \in (-1, 1).$$

Addiert man $\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k$ auf beiden Seiten so folgt die Behauptung.

^aDamit die Gleichung tatsächlich für $x \in (-1, 1)$ gilt müsste man an der Stelle überprüfen, dass der Konvergenzradius der Potenzreihe 1 ist. Dies überlassen wir den Lesenden als Übungsaufgabe.

Neben der Frage ob eine Funktion konstant ist kann auch das *Monotonieverhalten* mithilfe des Mittelwertsatzes studiert werden. Wir erinnern an der Stelle nochmal an Definition 6.2.22, wo der Begriff *Monotonie* definiert ist.

Mit dem Mittelwertsatz ergeben sich folgende Kriterien für Monotonie und strenge Monotonie.

Proposition 7.3.11 — MONOTONIESATZ

Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion, die auf (a, b) differenzierbar ist. Dann gilt

- (i) $f'(x) \geq 0 \quad \forall x \in (a, b) \iff f$ ist monoton wachsend auf $[a, b]$.
- (ii) $f'(x) \leq 0 \quad \forall x \in (a, b) \iff f$ ist monoton fallend auf $[a, b]$.
- (iii) $f'(x) > 0 \quad \forall x \in (a, b) \implies f$ ist streng monoton wachsend auf $[a, b]$.
- (iv) $f'(x) < 0 \quad \forall x \in (a, b) \implies f$ ist streng monoton fallend auf $[a, b]$.

Beweis. Zu (i).

Zuerst zu „ $\Leftarrow$ “. Es sei f monoton wachsend und $x_0 \in (a, b)$ beliebig.

$$\mathbb{Z}\mathbb{Z}: \quad f'(x_0) \geq 0.$$

Hierzu schreiben wir zuerst

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0+} \frac{f(x) - f(x_0)}{x - x_0} \stackrel{(264)}{=} \lim_{\substack{x \rightarrow x_0 \\ x \in (x_0, b]}} \frac{f(x) - f(x_0)}{x - x_0}. \quad (361)$$

Nun gilt wegen der Monotonie für alle $x \in (x_0, b]$, dass $f(x) \geq f(x_0)$ (vgl. Definition 6.2.22). Also hat man auch $f(x) - f(x_0) \geq 0$ für alle $x \in (x_0, b]$. Da für $x \in (x_0, b]$ auch $x - x_0 > 0$ gilt erhält man

$$\overbrace{\frac{f(x) - f(x_0)}{x - x_0}}^{\geq 0} \geq 0 \quad \forall x \in (x_0, b]$$

Mit den Grenzwertvergleichsregeln folgt

$$\lim_{\substack{x \rightarrow x_0 \\ x \in (x_0, b]}} \frac{f(x) - f(x_0)}{x - x_0} \geq 0$$

und somit nach (361) auch $f'(x_0) \geq 0$.

Nun noch zu „ $\implies$ “. Es sei $f'(x) \geq 0$ für alle $x \in (a, b)$.

$$\mathbb{Z}\mathbb{Z}: \quad f \text{ ist auf } [a, b] \text{ monoton wachsend.}$$

Wir überprüfen dazu Definition 6.2.22(i). Seien $x_1, x_2 \in [a, b]$ mit $x_1 \leq x_2$. Falls $x_1 = x_2$ ist, so ist auch klarerweise $f(x_1) = f(x_2)$, also insbesondere $f(x_1) \leq f(x_2)$. Wir dürfen

also ObdA annehmen, dass $x_1 < x_2$. Dann ist klarerweise f auf $[x_1, x_2] \subseteq [a, b]$ stetig und auf (x_1, x_2) differenzierbar. Es folgt aus dem Mittelwertsatz (bzw. (351))

$$\exists \xi \in (x_1, x_2) : \quad f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1).$$

Nun gilt $f'(\xi) \geq 0$, denn $\xi \in (x_1, x_2) \subseteq (a, b)$. Daher hat man

$$f(x_2) - f(x_1) = \underbrace{f'(\xi)}_{\geq 0} \underbrace{(x_2 - x_1)}_{\geq 0} \geq 0.$$

Addiert man $f(x_1)$ auf beiden Seiten so folgt $f(x_2) \geq f(x_1)$.

Zu (ii). Es gilt

$$\begin{aligned} & f \text{ ist monoton fallend auf } [a, b] \\ \iff & -f \text{ ist monoton wachsend auf } [a, b] \\ \iff & (-f)'(x) \geq 0 \quad \forall x \in (a, b) \\ \text{(i)} & \\ \iff & -f'(x) \geq 0 \quad \forall x \in (a, b) \\ \text{Kor 7.1.4} & \\ \iff & f'(x) \leq 0 \quad \forall x \in (a, b) \end{aligned}$$

(iii) funktioniert ähnlich wie die Richtung „ $\implies$ “ von (i). Wir überlassen den Lesenden Details als Übungsaufgabe.

20.04.2026

Bemerkung 7.3.12

Bei den Punkten (iii) und (iv) vom Monotoniesatz (Proposition 7.3.11) gilt die Rückrichtung nicht. In der Tat ist zum Beispiel

$$f : [-1, 1] \rightarrow \mathbb{R}, x \mapsto x^3$$

streng monoton wachsend (Übungsaufgabe). Jedoch gilt

$$\neg \quad (f'(x) > 0 \quad \forall x \in (-1, 1)),$$

denn wir haben bereits in Beispiel 7.2.8 gesehen, dass $f'(0) = 0$.

Bemerkung 7.3.13

Der Monotoniesatz (Proposition 7.3.11) kann mit wenig Aufwand auch auf Intervalle der Form $[a, b)$, $(a, b]$, (a, b) und auch auf unbeschränkte Intervalle verallgemeinert werden. Voraussetzung hierbei ist stets, dass f' an allen inneren Punkten dieser Intervalle das

entsprechende Vorzeichen haben muss. Genauer:

Verallgemeinerung. (von Proposition 7.3.11). Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ stetig und differenzierbar auf $\overset{\circ}{I}$. Dann gilt

- (i) $f'(x) \geq 0 \quad \forall x \in \overset{\circ}{I} \iff f$ ist monoton wachsend auf I ,
- (ii) $f'(x) \leq 0 \quad \forall x \in \overset{\circ}{I} \iff f$ ist monoton fallend auf I ,
- (iii) $f'(x) > 0 \quad \forall x \in \overset{\circ}{I} \implies f$ ist streng monoton wachsend auf I ,
- (iv) $f'(x) < 0 \quad \forall x \in \overset{\circ}{I} \implies f$ ist streng monoton fallend auf I .

Beispiel 7.3.14

Wir betrachten $f : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$f(x) = xe^x.$$

Dann gilt nach der Produktregel

$$f'(x) = 1 \cdot e^x + x \cdot e^x = (x+1)e^x \quad \forall x \in \mathbb{R}.$$

Man folgert

$$f'(x) \begin{cases} < 0 & x \in (-\infty, -1) \\ = 0 & x = -1 \\ > 0 & x \in (-1, \infty) \end{cases}$$

Zum Monotonieverlauf lässt sich dann folgendes sagen

$$f(x) \begin{cases} \text{ist streng monoton fallend auf } (-\infty, -1] \\ \text{ist streng monoton wachsend auf } [-1, \infty) \end{cases}$$

Mit diesem Monotonieverlauf und dem *Grenzverhalten* für $x \rightarrow \pm\infty$ kann man sich bereits ein relativ genaues Bild vom Funktionsgraphen Γ_f machen (vgl. Abbildung 33) .

Wenn man das Monotonieverhalten einer Funktion kennt, dann kann man auch besser verstehen, wo lokale Extrema angenommen werden. Diese werden nämlich unter anderem dort angenommen, wo sich das Monotonieverhalten *umkehrt*. Da wir das Monotonieverhalten auch über die Ableitung beschreiben können ergibt sich das folgende, sog. *Vorzeichenwechselkriterium* für Maximierungsaufgaben.

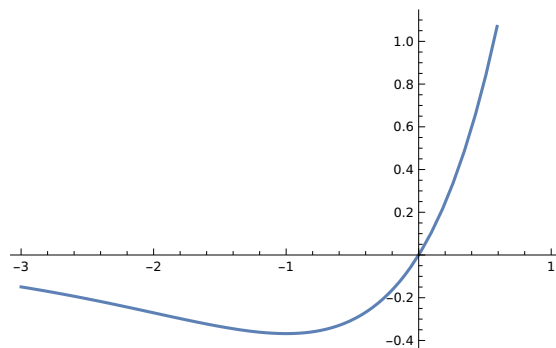


Abbildung 33: Die Funktion f aus Beispiel 7.3.14.

Korollar 7.3.15 — VORZEICHENWECHSELKRITERIUM

Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ stetig. Ferner sei $x_0 \in I$ und f differenzierbar auf $I \setminus \{x_0\}$.

(i)

$$f'(x) \begin{cases} \geq 0 & \forall x \in I : x < x_0 \\ \leq 0 & \forall x \in I : x > x_0 \end{cases} \implies x_0 \text{ ist Maximumsstelle von } f \text{ auf } I$$

(ii)

$$f'(x) \begin{cases} \leq 0 & \forall x \in I : x < x_0 \\ \geq 0 & \forall x \in I : x > x_0 \end{cases} \implies x_0 \text{ ist Minimumsstelle von } f \text{ auf } I$$

(iii)

$$f'(x) \begin{cases} > 0 & \forall x \in I : x < x_0 \\ < 0 & \forall x \in I : x > x_0 \end{cases} \implies x_0 \text{ ist strikte Max.-Stelle von } f \text{ auf } I$$

(iv)

$$f'(x) \begin{cases} < 0 & \forall x \in I : x < x_0 \\ > 0 & \forall x \in I : x > x_0 \end{cases} \implies x_0 \text{ ist strikte Min.-Stelle von } f \text{ auf } I$$

Beweis. Wir zeigen nur (i) und überlassen den Rest den Lesenden als Übungsaufgabe. Definiere die beiden Intervalle

$$I_1 := I \cap (-\infty, x_0] \text{ und } I_2 := I \cap [x_0, \infty).$$

Beide Intervalle I_1 und I_2 sind nichtleer. Besteht eines dieser Intervalle nur aus einem Punkt,

so kann es im Folgenden ignoriert werden (denn Monotonie auf einer Ein-Punkt-Menge ist trivial). Man überzeugt sich leicht, dass $x_0 \notin \overset{\circ}{I}_1$ und $x_0 \notin \overset{\circ}{I}_2$. Somit ist f in $\overset{\circ}{I}_1$ und $\overset{\circ}{I}_2$ differenzierbar und es gilt nach Voraussetzung

$$f'(x) \geq 0 \quad \forall x \in \overset{\circ}{I}_1 \quad \text{und} \quad f'(x) \leq 0 \quad \forall x \in \overset{\circ}{I}_2.$$

Mit Bemerkung 7.3.13 folgert man

$$f(x) \begin{cases} \text{monoton wachsend auf } I_1 \\ \text{monoton fallend auf } I_2 \end{cases}$$

Behauptung. $f(x) \leq f(x_0)$ für alle $x \in I$.

Es sei dazu $x \in I$ beliebig. Dann gilt $x \in I_1$ oder $x \in I_2$.

FALL 1. $x \in I_1$.

Dann gilt (nach der Definition von I_1) auch $x \leq x_0$. Da f auf I_1 monoton wachsend ist und $x, x_0 \in I_1$ sind gilt

$$x \leq x_0 \quad \implies \quad f(x) \leq f(x_0),$$

was zu zeigen war.

FALL 2. $x \in I_2$.

Dann gilt (nach der Definition von I_2) auch $x \geq x_0$. Da f auf I_2 monoton fallend ist und $x, x_0 \in I_2$ sind gilt

$$x \geq x_0 \quad \implies \quad f(x) \leq f(x_0),$$

was zu zeigen war.

Bemerkung 7.3.16

Es sei nun $D \subseteq \mathbb{R}$ beliebig und $f : D \rightarrow \mathbb{R}$ gegeben. Das vorige Kriterium kann auch verwendet werden um *lokale* Extrema zu bestimmen. In der Tat: Sei $x_0 \in D$. Falls

- (a) ein $\delta > 0$ existiert so, dass $D \cap B_\delta(x_0)$ ein Intervall ist und
- (b) dieses Intervall so ist dass einer der in Korollar 7.3.15 behandelten Fälle zutrifft,

so können wir zeigen, dass x_0 eine Extremstelle von f in $D \cap B_\delta(x_0)$ ist. Das bedeutet aber nichts anderes als, dass f eine lokale Extremstelle von f auf D ist, vgl. Definition 7.2.1!

Beachte: Wegen Bedingung (a) kann man obige Aussage im Allgemeinen nicht bei jedem Punkt $x_0 \in D$ anwenden. Allerdings ist trotzdem ein wichtiger Fall mit eingeschlossen — nämlich der Fall der *inneren lokalen Extrema*!

In der Tat: Ist x_0 ein innerer Punkt von D so gibt es nach Definition 7.2.4 ein $r > 0$ so, dass $B_r(x_0) \subseteq D$. Wählt man dann $\delta \leq r$ so ist $B_\delta(x_0) \subseteq B_r(x_0) \subseteq D$. Somit ist

$$D \cap B_\delta(x_0) = B_\delta(x_0) = (x_0 - \delta, x_0 + \delta)$$

tatsächlich ein Intervall. Bedingung (a) ist also erfüllt.

Anwendung: Umkehrfunktionen der trigonometrischen Funktionen

Als nächstes möchten wir die obige Monotoniediskussion nutzen, um die *Arcus-Funktionen*, die Umkehrfunktionen der trigonometrischen Funktionen $\cos, \sin, \tan$, einzuführen. Um eine Umkehrfunktion finden zu können, müssen wir die o.g. Funktionen zunächst auf geeignete Definitions- und Bildbereiche einschränken, sodass sie dort bijektiv sind.

Proposition 7.3.17

- (i) $\cos : [0, \pi] \rightarrow [-1, 1]$ ist bijektiv.
- (ii) $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [-1, 1]$ ist bijektiv.
- (iii) $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}$ ist bijektiv.

Bevor wir diese Proposition beweisen, schieben wir ein hilfreiches Lemma ein, welches den Beweis erleichtern wird.

Lemma 7.3.18

Es gilt $\sin(x) > 0$ für alle $x \in (0, \pi)$.

Beweis. Für $x \in (0, \frac{\pi}{2}]$ folgt $\sin(x) > 0$ bereits aus Korollar 5.5.13, denn wir wissen nach Definition 5.5.10 dass $\frac{\pi}{2} \leq 2$ und Korollar 5.5.13 besagt gerade, dass $\sin(x) > 0$ für alle $x \in (0, 2]$. Um den Sachverhalt auch für $x \in [\frac{\pi}{2}, \pi)$ zu klären betrachten wir zunächst die folgende Rechnung:

$$\sin(\pi - x) \stackrel{\text{Prop 5.5.4}}{=} \sin(\pi) \cos(-x) + \cos(\pi) \sin(-x) \stackrel{\text{Kor 5.5.15}}{=} -\sin(-x) \stackrel{\text{Prop 5.5.3}}{=} \sin(x).$$

Ist nun $x \in (\frac{\pi}{2}, \pi)$ so ist $\pi - x \in (0, \frac{\pi}{2})$ und man folgert aus obiger Rechnung

$$\sin(x) = \sin(\pi - x) > 0 \quad [\text{da } \pi - x \in (0, \frac{\pi}{2})].$$

Nun sind wir bereit, Proposition 7.3.17 zu beweisen.

Beweis (von Proposition 7.3.17). Zu (i). Wir wissen aus Lemma 7.3.18, dass

$$\sin(x) > 0 \quad \forall x \in (0, \pi).$$

Wir folgern

$$\cos'(x) = -\sin(x) < 0 \quad \forall x \in (0, \pi).$$

Mit Proposition 7.3.11 folgt

$\cos$ ist streng monoton fallend auf $[0, \pi]$.

Somit ist $\cos$ auf $[0, \pi]$ auch injektiv. Desweiteren gilt mit Lemma 6.2.1

$$\cos([0, \pi]) = [\cos(\pi), \cos(0)] \stackrel{\text{Kor 5.5.15}}{=} [-1, 1].$$

Mit anderen Worten ist $\cos : [0, \pi] \rightarrow [-1, 1]$ surjektiv und injektiv, also bijektiv.

Zu (ii). Aus Korollar 5.5.16 wissen wir, dass $\sin(z) = -\cos(\frac{\pi}{2} + z)$ für alle $z \in \mathbb{R}$. Nutzt man dies und (i) so folgt leicht die Bijektivität von $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [-1, 1]$.

Zu (iii). Beachte: $(-\frac{\pi}{2}, \frac{\pi}{2})$ liegt im Definitionsbereich des Tangens, d.h. $\cos(x) \neq 0$ für alle $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$. In der Tat: Angenommen es gibt ein $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ mit $\cos(x) = 0$. Beachte: Es gilt $|x| \in [0, \frac{\pi}{2})$. Nach Proposition 5.5.3 gilt $\cos(x) = \cos(-x)$ also auch $\cos(x) = \cos(|x|)$. Wir folgern

$$\cos(|x|) = \cos(x) = 0 = \cos(\frac{\pi}{2}).$$

Da $\cos$ auf $[0, \frac{\pi}{2}]$ injektiv ist, folgt $|x| = \frac{\pi}{2}$, was $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ widerspricht. Nun zur Bijektivität Aus Beispiel 7.1.3 wissen wir bereits, dass $\tan'(x) = 1 + \tan^2(x) > 0$ für alle x aus dem Definitionsbereich des Tangens. Also gilt auch

$$\tan'(x) > 0 \quad \text{für alle } x \in (-\frac{\pi}{2}, \frac{\pi}{2}).$$

Somit ist $\tan$ auch streng monoton wachsend auf $(-\frac{\pi}{2}, \frac{\pi}{2})$. Wir müssen uns noch davon überzeugen, dass

$$\tan((-\frac{\pi}{2}, \frac{\pi}{2})) = \mathbb{R}.$$

Dazu bilde

$$\lim_{x \rightarrow \frac{\pi}{2}-} \tan(x) = \lim_{x \rightarrow \frac{\pi}{2}-} \frac{\sin(x)}{\cos(x)}.$$

Daraus dass der Zähler gegen 1 und der Nenner gegen 0 steht können wir leicht folgern, dass $\lim_{x \rightarrow \frac{\pi}{2}-} \tan(x) = \infty$. Die Details überlassen wir den Lesenden als Übungsaufgabe. Analog gilt $\lim_{x \rightarrow (-\frac{\pi}{2})+} \tan(x) = -\infty$. Dies und Proposition 6.2.17 impliziert die Behauptung.

Da $\sin, \cos, \tan$ auf den in der vorigen Proposition genannten Intervallen stetig und bijektiv sind haben sie auch stetige Umkehrfunktionen, vgl. Satz 6.2.20. Diesen werden wir nun Namen geben.

Definition 7.3.19 — DIE ARCUS-FUNKTIONEN

- (i) Der *Arcus-Kosinus* $\arccos : [-1, 1] \rightarrow [0, \pi]$ bezeichne die Umkehrfunktion von $\cos : [0, \pi] \rightarrow [-1, 1]$
- (ii) Der *Arcus-Sinus* $\arcsin : [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$ bezeichne die Umkehrfunktion von $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [-1, 1]$
- (iii) Der *Arcus-Tangens* $\arctan : \mathbb{R} \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$ bezeichne die Umkehrfunktion von $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}$.

Im Inneren der Definitionsintervalle sind die oben definierten Funktionen auch differenzierbar, wie wir in der nachfolgenden Proposition sehen werden.

Proposition 7.3.20 — DIFFERENTIATION DER ARCUS-FUNKTIONEN

- (i) $\arccos$ ist auf $(-1, 1)$ differenzierbar und es gilt

$$\arccos'(x) = -\frac{1}{\sqrt{1-x^2}} \quad \forall x \in (-1, 1).$$

- (ii) $\arcsin$ ist auf $(-1, 1)$ differenzierbar und es gilt

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}} \quad \forall x \in (-1, 1).$$

- (iii) $\arctan$ ist auf $\mathbb{R}$ differenzierbar und es gilt

$$\arctan'(x) = \frac{1}{1+x^2} \quad \forall x \in \mathbb{R}.$$

Beweis. Zu (i). Wir werden Satz 7.1.8 benutzen. Sei dazu $y \in (-1, 1)$. Weil $\cos([0, \pi]) = [-1, 1]$ gibt es ein $x \in [0, \pi]$ mit $\cos(x) = y$. Die Werte 0 und π kommen aber für x nicht infrage, da $\cos(0) = 1 \neq y$ und $\cos(\pi) = -1 \neq y$. Mit anderen Worten

$$y = \cos(x) \quad \text{für ein } x \in (0, \pi) \tag{362}$$

Voraussetzung für die Differenzierbarkeit des Umkehrfunktion bei y ist nach Satz 7.1.8, dass $\cos'(x) \neq 0$. Nun gilt für alle $x \in (0, \pi)$, dass

$$\cos'(x) = -\sin(x) \neq 0. \quad (\text{vgl. Lemma 7.3.18})$$

Daher sind die Voraussetzungen von Satz 7.1.8 überprüft und es folgt

$$\arccos'(y) = \frac{1}{\cos'(x)} \stackrel{x=\arccos(y)}{\substack{(362)}} \frac{1}{\cos'(\arccos(y))} = -\frac{1}{\sin(\arccos(y))}. \tag{363}$$

Diesen Ausdruck vereinfachen wir nun weiter: Es gilt für $x = \arccos(y) \in (0, \pi)$ (vgl. (362))

$$\sin(x) > 0 \quad \text{und} \quad \sin^2(x) + \cos^2(x) = 1.$$

Wir folgern

$$\sin(x) > 0 \quad \text{und} \quad \sin^2(x) = 1 - \cos^2(x),$$

woraus sich durch Wurzelziehen ergibt

$$\sin(x) = \sqrt{1 - \cos^2(x)}.$$

Dies in (363) eingesetzt liefert

$$\begin{aligned} \arccos'(y) &= -\frac{1}{\sin(\arccos(y))} = -\frac{1}{\sqrt{1 - \cos^2(\arccos(y))}} \\ &= -\frac{1}{\sqrt{1 - (\cos(\arccos(y)))^2}} = -\frac{1}{\sqrt{1 - y^2}}, \end{aligned}$$

was zu zeigen war.

(ii) ist den Lesenden als Übungsaufgabe überlassen (und funktioniert analog zu (i)).

Nun zu (iii). Da nach Beispiel 7.1.3 gilt $\tan'(x) = 1 + \tan^2(x) > 0$ für alle $x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ sind die Voraussetzungen von Satz 7.1.8 stets erfüllt. Man folgert für alle $y \in \mathbb{R}$, dass

$$\arctan'(y) = \frac{1}{\tan'(\arctan(y))} \stackrel{\text{Bsp 7.1.3}}{=} \frac{1}{1 + \tan^2(\arctan(y))} = \frac{1}{1 + \tan(\arctan(y))^2} = \frac{1}{1 + y^2}.$$

Die Regeln von de L'Hospital

In diesem Abschnitt, der eine weitere Anwendung der Mittelwertsätze präsentiert, beschäftigen wir uns mit einer nützlichen Regel, die die Berechnung von Funktionsgrenzwerten mithilfe von Ableitungen ermöglicht. Sie beschäftigt sich mit Grenzwerten von *Quotienten von Funktionen*, d.h. mit Grenzwerten der Form

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

Existieren die Grenzwerte von Zähler und Nenner und ist der Grenzwert des Nenners ungleich 0 so kann der Grenzwert mithilfe von Proposition 6.1.8 berechnet werden.

Ist der Grenzwert des Nenners aber gleich 0, so haben wir bisher keine Möglichkeit, den Grenzwert auszurechnen... Hier kann in vielen Fällen aber die sogenannte *L'Hospital'sche* $\frac{0}{0}$ -Regel weiterhelfen!

Satz 7.3.21 — L'HOSPITAL'SCHE $\frac{0}{0}$ -REGEL

Es sei $I \subseteq \mathbb{R}$ ein offenes Intervall. Es seien $f, g : I \rightarrow \mathbb{R}$ auf I differenzierbar und $x_0 \in \mathbb{R}$

ein Häufungspunkt von I . Ferner gelte die sogenannte „ $\frac{0}{0}$ -Bedingung“

$$\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = 0. \quad (364)$$

Falls $g'(x) \neq 0$ für alle $x \in I$ und falls $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$ existiert, so existiert auch $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}$ und es gilt

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}.$$

Zusatz. Ist I nach oben (bzw. nach unten) unbeschränkt, so gilt selbiges auch für $x_0 = \infty$ (bzw. $x_0 = -\infty$).

Beweis. Wir zeigen zuerst die Behauptung ohne Zusatz. Das bedeutet: Wir dürfen oBdA annehmen dass I beschränkt ist. (Denn wenn I unbeschränkt ist, kann man f und g stattdessen auf dem (offenen und beschränkten) Intervall $I \cap (x_0 - 1, x_0 + 1)$ betrachten. Dass diese Verkleinerung des Intervalles nichts am Grenzwert ändert liegt an Korollar 6.1.28.

Es sei also dann etwa $I = (a, b)$ für $a, b \in \mathbb{R}$ mit $a < b$.

FALL 1. $x_0 \in (a, b)$.

$$\mathbb{Z} : \quad \lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}.$$

Ist $x_0 \in (a, b)$ so sind f und g bei x_0 diffbar und daher auch stetig. Man erhält aus der Voraussetzung (374)

$$f(x_0) = \lim_{x \rightarrow x_0} f(x) = 0 \quad (365)$$

und

$$g(x_0) = \lim_{x \rightarrow x_0} g(x) = 0. \quad (366)$$

Ferner gilt $g'(x_0) \neq 0$. Beachte: Da $x_0 \in (a, b)$ und $g'(x_0) \neq 0$ darf man die Folge $(x_n)_{n \in \mathbb{N}} = (x_0)_{n \in \mathbb{N}} \subseteq (a, b)$ betrachten und erhält

$$\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} = \lim_{n \rightarrow \infty} \frac{f'(x_n)}{g'(x_n)} = \lim_{n \rightarrow \infty} \frac{f'(x_0)}{g'(x_0)} = \frac{f'(x_0)}{g'(x_0)}. \quad (367)$$

Nun berechnet man mit (365) und (366), dass für alle $x \in (a, b) \setminus \{x_0\}$

$$\frac{f(x)}{g(x)} \stackrel{(365), (366)}{=} \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{\frac{f(x) - f(x_0)}{x - x_0}}{\frac{g(x) - g(x_0)}{x - x_0}} \quad (368)$$

Da $\lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0} = g'(x_0) \neq 0$ kann die Grenzwertregel aus Proposition 6.1.8 angewendet werden und man erhält

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} \stackrel{(368)}{=} \lim_{x \rightarrow x_0} \frac{\overbrace{\frac{f(x) - f(x_0)}{x - x_0}}^{\xrightarrow{x \rightarrow x_0} f'(x_0)}}{\underbrace{\frac{g(x) - g(x_0)}{x - x_0}}_{\xrightarrow{x \rightarrow x_0} g'(x_0)}} = \frac{f'(x_0)}{g'(x_0)} \stackrel{(367)}{=} \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)},$$

was zu zeigen war.

FALL 2. $x_0 \notin (a, b)$.

Da x_0 aber ein Häufungspunkt von (a, b) ist, muss gelten, dass entweder $x_0 = a$ oder $x_0 = b$. Wir zeigen die Behauptung nur im Fall, dass $x_0 = b$, der andere Fall geht analog. Definiere hierzu die Funktionen

$$\tilde{f} : (a, b] \rightarrow \mathbb{R}, \quad \tilde{f}(x) := \begin{cases} f(x) & x \in (a, b) \\ 0 & x = b \end{cases} \quad (369)$$

sowie

$$\tilde{g} : (a, b] \rightarrow \mathbb{R}, \quad \tilde{g}(x) := \begin{cases} g(x) & x \in (a, b) \\ 0 & x = b \end{cases}. \quad (370)$$

Zwischenbehauptung. $\tilde{f}$ und $\tilde{g}$ sind auf $(a, b]$ stetig.

Dazu: Stetigkeit auf (a, b) ist klar, weil f und g dort stetig sind. Es muss also nur noch die Stetigkeit von $\tilde{f}$ und $\tilde{g}$ bei b gezeigt werden, d.h.

$$\lim_{x \rightarrow b} \tilde{f}(x) = \tilde{f}(b) [= 0].$$

Dazu beachte, dass

$$\lim_{x \rightarrow b-} \overset{\circ}{\tilde{f}}(x) = \lim_{x \rightarrow b} f(x) = \lim_{x \rightarrow x_0} f(x) \stackrel{(374)}{=} 0.$$

Da ferner $(a, b] \cap [b, \infty] = \{b\}$ eine Einpunktmenge ist und $\tilde{f}(b) = 0$, gilt für den rechts-cut $\tilde{f}_r : \{b\} \rightarrow \mathbb{R}$, dass $\tilde{f}_r(x) = \tilde{f}(x) = 0$ für alle $x \in \{b\}$. Man erhält

$$\lim_{x \rightarrow b+} \tilde{f}(x) = \lim_{x \rightarrow b} \tilde{f}_r(x) = 0$$

Die beiden vorigen Gleichungen implizieren mit 6.1.16, dass

$$\lim_{x \rightarrow b} \tilde{f}(x) = 0 = f(b).$$

Dieselbe Beweistechnik zeigt die Stetigkeit von $\tilde{g}$. Man nennt $\tilde{f}$ und $\tilde{g}$ *stetige Fortsetzungen* von f und g .

Zwischenbehauptung Ende.

Es sei nun $(x_n)_{n \in \mathbb{N}} \subseteq (a, b)$ eine beliebige Folge mit $\lim_{n \rightarrow \infty} x_n = b$.

$$\mathbb{Z}\mathbb{Z}: \quad \lim_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}.$$

Dank der Zwischenbehauptung wissen wir: $\tilde{f}$ und $\tilde{g}$ auf $[x_n, b]$ stetig und auch auf (x_n, b) differenzierbar (da sie dort ja mit den differenzierbaren Funktionen f bzw. g übereinstimmen). Ferner gilt nach Voraussetzung $g'(x) \neq 0$ für alle $x \in (x_n, b)$. Aus dem zweiten Mittelwertsatz (Satz 7.3.6) folgt nun dass es ein $\xi_n \in (x_n, b)$ gibt so, dass

$$\frac{\tilde{f}(x_n) - \tilde{f}(b)}{\tilde{g}(x_n) - \tilde{g}(b)} = \frac{\tilde{f}'(\xi_n)}{\tilde{g}'(\xi_n)} \stackrel{(369), \xi_n \in (a, b)}{=} \frac{f'(\xi_n)}{g'(\xi_n)}. \quad (371)$$

Wegen (369) und (370) gilt $\tilde{f}(x_n) = f(x_n)$, $\tilde{g}(x_n) = g(x_n)$ sowie $\tilde{f}(b) = \tilde{g}(b) = 0$. Nutzt man dies in (371) so erhält man

$$\frac{f(x_n)}{g(x_n)} = \frac{f'(\xi_n)}{g'(\xi_n)} \quad \forall n \in \mathbb{N}. \quad (372)$$

Nun gilt $\xi_n \xrightarrow{n \rightarrow \infty} b$. Dies folgt aus $x_n \leq \xi_n \leq b$ und dem Sandwichlemma (vgl. Proposition 2.3.26 — anwendbar, da $\lim_{n \rightarrow \infty} x_n = b = \lim_{n \rightarrow \infty} b$). Da nach der Voraussetzung $\lim_{x \rightarrow b} \frac{f'(x)}{g'(x)}$ existiert folgern wir

$$\lim_{n \rightarrow \infty} \frac{f'(\xi_n)}{g'(\xi_n)} = \lim_{x \rightarrow b} \frac{f'(x)}{g'(x)}$$

Dies und (372) impliziert

$$\lim_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} = \lim_{x \rightarrow b} \frac{f'(x)}{g'(x)},$$

was zu zeigen war.

Nun noch eine Beweisskizze zum Zusatz. Wir skizzieren nur die Beweisidee im Fall $x_0 = \infty$. Man überzeugt sich leicht, dass (sofern einer der beiden Grenzwerte existiert)

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

Außerdem folgt aus (374) (mit $x_0 = \infty$) auch, dass

$$\lim_{x \rightarrow 0} f\left(\frac{1}{x^2}\right) = \lim_{x \rightarrow 0} g\left(\frac{1}{x^2}\right) = 0.$$

(Details dazu sind den Lesenden als Übungsaufgabe überlassen). Nun lässt sich die bereits beweisene Aussage für $x_0 = 0 \in \mathbb{R}$ anwenden ^a. Man erhält

$$\lim_{x \rightarrow 0} \frac{f\left(\frac{1}{x^2}\right)}{g\left(\frac{1}{x^2}\right)} = \lim_{x \rightarrow 0} \frac{\frac{d}{dx} f\left(\frac{1}{x^2}\right)}{\frac{d}{dx} g\left(\frac{1}{x^2}\right)} \stackrel{\text{Prop 7.1.5}}{=} \lim_{x \rightarrow 0} \frac{f'\left(\frac{1}{x^2}\right)\left[-\frac{2}{x^3}\right]}{g'\left(\frac{1}{x^2}\right)\left[-\frac{2}{x^3}\right]} = \lim_{x \rightarrow 0} \frac{f'\left(\frac{1}{x^2}\right)}{g'\left(\frac{1}{x^2}\right)} = \lim_{x \rightarrow \infty} \frac{f'(x)}{g'(x)}.$$

^aAn dieser Stelle müsste man eigentlich noch mehr Voraussetzungen überprüfen. Wir überlassen auch dies den Lesenden als Übungsaufgabe.

Bemerkung 7.3.22

Die Voraussetzung (V) “ $g'(x) \neq 0$ für alle $x \in I$ ” kann abgeschwächt werden zu (V’) “ $g'(x) \neq 0$ für alle $x \in I \setminus \{x_0\}$ ”. In der Tat: Falls $f, g : I \rightarrow \mathbb{R}$ alle Voraussetzungen von Satz 7.3.21 mit (V’) anstatt von (V) erfüllen, so kann man I in $I_1 := I \cap (-\infty, x_0)$ und $I_2 := I \cap (x_0, \infty)$ aufteilen. Auf beiden Teilintervallen sind alle Voraussetzungen von Satz 7.3.21 dann komplett erfüllt. (Beachte: der Punkt x_0 ist stets ein Randpunkt von I_1 und I_2 und damit insbesondere ein Häufungspunkt). Die Regel von l’Hospital erlaubt uns somit also

$$\lim_{\substack{x \rightarrow x_0 \\ x \in I_1}} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0^-} \frac{f(x)}{g(x)} \quad \text{ sowie } \quad \lim_{\substack{x \rightarrow x_0 \\ x \in I_2}} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0^+} \frac{f(x)}{g(x)}$$

zu berechnen. Danach kann der Gesamtgrenzwert dank Proposition 6.1.16 auch bestimmt werden.

Bemerkung 7.3.23

In Fall 2 des obigen Beweises haben wir eine sogenannte *stetige Fortsetzung* von f bzw g gewählt. Dies ist oft nützlich und kann als allgemeine Erkenntnis festgehalten werden:

Fakt. Gilt für eine Funktion $f : (a, b) \rightarrow \mathbb{R}$ und einen Wert $y_0 \in \mathbb{R}$, dass

$$\lim_{x \rightarrow b} f(x) = y_0,$$

so ist $\tilde{f} : (a, b] \rightarrow \mathbb{R}$ definiert durch

$$\tilde{f}(x) := \begin{cases} f(x) & x \in (a, b) \\ y_0 & x = b \end{cases}$$

eine auf $(a, b]$ stetige Funktion.

Der Beweis ist analog zur Zwischenbehauptung in Satz 7.3.21 und ist den Lesenden als Übungsaufgabe überlassen.

Nun sehen wir einige Beispiele, wo uns die L'Hospital'sche $\frac{0}{0}$ -Regel helfen kann.

Beispiel 7.3.24

Betrachte nochmal den Grenzwert aus Beispiel 6.1.9, d.h.

$$\lim_{x \rightarrow 2} \frac{x^2 - 3x + 2}{x^2 - 4}.$$

In Beispiel 6.1.9 mussten wir mühsam die Nullstellen von Zähler und Nenner ausfaktorisieren. Bei Polynomen höheren Grades wäre das sehr kompliziert. Mit der l'Hospital'schen $\frac{0}{0}$ -Regel lässt sich das ganze aber sehr angenehm lösen.

Berechnung des Grenzwertes. Wir prüfen die Voraussetzungen von Satz 7.3.21. Zuerst zur „ $\frac{0}{0}$ -Bedingung“ (vgl. (374)):

$$\lim_{x \rightarrow 2} (x^2 - 3x + 2) = 2^2 - 3 \cdot 2 + 2 = 0,$$

$$\lim_{x \rightarrow 2} (x^2 - 4) = 2^2 - 4 = 0.$$

Nun bilden wir den Quotienten der Ableitungen, d.h.

$$\frac{\frac{d}{dx}(x^2 - 3x + 2)}{\frac{d}{dx}(x^2 - 4)} = \frac{2x - 3}{2x}$$

und stellen fest, dass

$$\lim_{x \rightarrow 2} \frac{2x - 3}{2x} = \frac{2 \cdot 2 - 3}{2 \cdot 2} = \frac{4 - 3}{4} = \frac{1}{4}.$$

Wendet man nun Satz 7.3.21 mit $f : x \mapsto x^2 - 3x + 2$ und $g : x \mapsto x^2 - 4$ an, so folgt

$$\lim_{x \rightarrow 2} \frac{x^2 - 3x + 2}{x^2 - 4} \stackrel{\text{„}\frac{0}{0}\text{-IH“}}{=} \lim_{x \rightarrow 2} \frac{\frac{d}{dx}(x^2 - 3x + 2)}{\frac{d}{dx}(x^2 - 4)} = \lim_{x \rightarrow 2} \frac{2x - 3}{2x} \stackrel{\text{GWR}}{=} \frac{1}{4}. \quad (373)$$

In Zukunft ist die Rechnung in (373) völlig ausreichend als Begründung. Sie dürfen das in den Übungsaufgaben so aufschreiben. Allerdings sollten Sie an der Stelle wo „ $\frac{0}{0}$ -IH“ steht auch wirklich die $\frac{0}{0}$ -Bedingung überprüfen! Wir werden nämlich im nächsten Beispiel sehen, dass diese essentiell ist.

Wichtig ist auch: Wenn in einer Rechnung wie in (373) der Grenzwert des Quotienten der Ableitungen nicht existieren sollte, muss die Rechnung abgebrochen werden. In dem

Fall ist nämlich die Regel von L'Hospital NICHT anwendbar! Schließlich ist die Existenz von $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$ in Satz 7.3.21 explizit Voraussetzung und muss eigentlich separat überprüft werden.

Wir haben in der obigen Rechnung ein wenig unterschlagen, die Bedingung „ $g'(x) \neq 0$ für alle $x \in I$ “ aus Satz 7.3.21 zu überprüfen. Das liegt auch daran, dass wir dafür zunächst einmal I wählen müssten. Beachte

$$g'(x) = \frac{d}{dx}(x^2 - 4) = 2x \neq 0 \quad \forall x \in (0, \infty).$$

Wählt man also $I = (0, \infty)$, so ist diese Voraussetzung erfüllt. Man muss bei dieser Wahl allerdings noch sicherstellen, dass x_0 ein Häufungspunkt von I ist. Diese ist hier erfüllt, da $2 \in (0, \infty)$. Aufgrund dieser zusätzlichen Bedingung hätte man das Intervall $I = (-\infty, 0)$ nicht betrachten dürfen (obwohl gleichermaßen $g'(x) \neq 0$ für alle $x \in (-\infty, 0)$ gilt).

Beispiel 7.3.25

Auf die „ $\frac{0}{0}$ -Bedingung“ darf in l'Hospitals $\frac{0}{0}$ -Regel (Satz 7.3.21) nicht verzichtet werden. In der Tat: Es gilt

$$\lim_{x \rightarrow 1} \frac{x^2}{x} = \lim_{x \rightarrow 1} x = 1,$$

aber

$$\lim_{x \rightarrow 1} \frac{\frac{d}{dx}(x^2)}{\frac{d}{dx}(x)} = \lim_{x \rightarrow 1} \frac{2x}{1} = 2 \neq 1.$$

Die Erklärung dafür, dass die beiden Grenzwerte nicht übereinstimmen, ist in der Tat, dass die $\frac{0}{0}$ -Bedingung nicht erfüllt ist, denn

$$\lim_{x \rightarrow 1} x^2 = 1 \neq 0 \quad \text{und} \quad \lim_{x \rightarrow 1} x = 1 \neq 0.$$

Beispiel 7.3.26

Ein weiteres klassisches Beispiel für die Anwendung l'Hospitalsche $\frac{0}{0}$ -Regel (Satz 7.3.21) ist die Berechnung

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x}.$$

Zunächst prüfen wir die $\frac{0}{0}$ -Bedingung. Dazu

$$\lim_{x \rightarrow 0} \sin(x) = \sin(0) = 0 \quad \text{und} \quad \lim_{x \rightarrow 0} x = 0$$

hat man mit l'Hospitals $\frac{0}{0}$ -Regel

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} \stackrel{„\frac{0}{0}\text{-l'H}“}{=} \lim_{x \rightarrow 0} \frac{\frac{d}{dx} \sin(x)}{\frac{d}{dx} (x)} = \lim_{x \rightarrow 0} \frac{\cos(x)}{1} = \lim_{x \rightarrow 0} \cos(x) = \cos(0) = 1.$$

22.04.2026

Beispiel 7.3.27

Die $\frac{0}{0}$ -Regel von l'Hospital kann auch *iteriert* angewendet werden. Das bedeutet, dass erst ein mehrfaches Anwenden der Regel einen berechenbaren Grenzwert liefert. In diesem Fall sollte man aber nach jeder Anwendung noch überprüfen ob die „ $\frac{0}{0}$ -Bedingung“ noch erfüllt ist. Ein Beispiel:

$$\lim_{x \rightarrow 0} \frac{\sin(x) - x}{x \sin(x)}.$$

Da (wie man leicht als Übungsaufgabe sieht) $\lim_{x \rightarrow 0} \sin(x) - x = \lim_{x \rightarrow 0} x \sin(x) = 0$ haben wir

$$\lim_{x \rightarrow 0} \frac{\sin(x) - x}{x \sin(x)} \stackrel{„\frac{0}{0}\text{-l'H}“}{=} \lim_{x \rightarrow 0} \frac{\frac{d}{dx}(\sin(x) - x)}{\frac{d}{dx}(x \sin(x))} = \lim_{x \rightarrow 0} \frac{\cos(x) - 1}{\sin(x) + x \cos(x)}.$$

Da (wie man leicht als Übungsaufgabe sieht) $\lim_{x \rightarrow 0} (\cos(x) - 1) = \lim_{x \rightarrow 0} (\sin(x) + x \cos(x)) = 0$ hat man wiederum

$$\lim_{x \rightarrow 0} \frac{\cos(x) - 1}{\sin(x) + x \cos(x)} \stackrel{„\frac{0}{0}\text{-l'H}“}{=} \lim_{x \rightarrow 0} \frac{-\sin(x)}{\cos(x) + \cos(x) - x \sin(x)} = \frac{\sin(0)}{2 \cos(0) - 0 \cdot \sin(0)} = \frac{0}{2} = 0.$$

Wir folgern nach zweimaliger Anwendung von der $\frac{0}{0}$ -Regel von l'Hospital, dass

$$\lim_{x \rightarrow 0} \frac{\sin(x) - x}{x \sin(x)} = 0.$$

Beispiel 7.3.28

Manche Grenzwerte, die mit der Regel von l'Hospital berechnet werden können, stehen nicht direkt als Quotient da. Man muss sie erst als Quotient umformen. Ein Beispiel:

$$\lim_{x \rightarrow \infty} x(e^{\frac{1}{x}} - 1).$$

Als Quotient umgeschrieben hat man

$$\lim_{x \rightarrow \infty} \frac{e^{\frac{1}{x}} - 1}{\frac{1}{x}}$$

Mit $\lim_{x \rightarrow \infty} (e^{\frac{1}{x}} - 1) = \lim_{y \rightarrow 0} (e^y - 1) = 0$ und $\lim_{x \rightarrow \infty} \frac{1}{x} = 0$ erhält man

$$\lim_{x \rightarrow \infty} \frac{e^{\frac{1}{x}} - 1}{\frac{1}{x}} \stackrel{0/0\text{-L'H}}{=} \lim_{x \rightarrow \infty} \frac{\frac{d}{dx}(e^{\frac{1}{x}} - 1)}{\frac{d}{dx}(\frac{1}{x})} = \lim_{x \rightarrow \infty} \frac{e^{\frac{1}{x}}(-\frac{1}{x^2})}{-\frac{1}{x^2}} = \lim_{x \rightarrow \infty} e^{\frac{1}{x}} \stackrel{\text{Bsp 6.1.21}}{=} \lim_{y \rightarrow 0} e^y = 1.$$

Es gibt noch eine weitere Regel von l'Hospital, die auch sehr hilfreich bei der Berechnung von Grenzwerten ist. Sie beschäftigt sich wiederum mit Grenzwerten der Form

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

Diesmal ist aber $\lim_{x \rightarrow x_0} g(x) = \infty$ vorausgesetzt.

Satz 7.3.29 — L'HOSPITAL'SCHE $\frac{\infty}{\infty}$ -REGEL.

Es sei $I \subseteq \mathbb{R}$ ein offenes Intervall. Es seien $f, g : I \rightarrow \mathbb{R}$ auf $I = (a, b)$ differenzierbar und $x_0 \in \{a, b\}$ ein Randpunkt von I . Ferner gelte die sogenannte „ $\frac{\infty}{\infty}$ -Bedingung“

$$\lim_{x \rightarrow x_0} g(x) = \infty. \quad (374)$$

Falls $g'(x) \neq 0$ für alle $x \in I$ und falls $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$ existiert, so existiert auch $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)}$ und es gilt

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}.$$

Zusatz. Ist I nach oben (bzw. nach unten) unbeschränkt, so gilt selbiges auch für $x_0 = \infty$ (bzw. $x_0 = -\infty$).

***Beweis.** Wir beweisen die Behauptung nur ohne Zusatz (d.h. für $x_0 \in \mathbb{R}$). Für den Zusatz verweisen wir auf den Beweis von Satz 7.3.21. Es seien f und g wie in dem Satz und $I = (a, b)$, $x_0 \in \{a, b\}$. Wir beweisen die Behauptung nur für $x_0 = b$, der Fall $x_0 = a$ geht analog. Definiere nun

$$L := \lim_{x \rightarrow b} \frac{f'(x)}{g'(x)}. \quad (375)$$

Das bedeutet nach Satz 6.1.23

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x \in I : \quad |x - b| < \delta \quad \text{gilt} \quad \left| \frac{f'(x)}{g'(x)} - L \right| < \varepsilon. \quad (376)$$

$$\text{ZZ : } \lim_{x \rightarrow b} \frac{f(x)}{g(x)} = L.$$

Sei dazu $\varepsilon > 0$ beliebig vorgegeben und δ wie in (376) gewählt. Beachte, dass $\{x \in I :$

$|x - b| < \delta\} = (b - \delta, b)$. Deswegen lernen wir aus (376), dass

$$\forall x \in (b - \delta, b) : \quad L - \varepsilon \leq \frac{f'(x)}{g'(x)} \leq L + \varepsilon. \quad (377)$$

Wähle dazu nun eine beliebige Folge $(x_n)_{n \in \mathbb{N}} \subseteq (a, b)$ mit $\lim_{n \rightarrow \infty} x_n = b$. Man überzeugt sich leicht (als Übungsaufgabe) dass es ein $N \in \mathbb{N}$ gibt mit $(x_n)_{n \geq N} \subseteq (b - \delta, b)$. Nun gibt es nach dem zweiten Mittelwertsatz (Satz 7.3.6)^a für alle $n \geq N$ einen Wert $\xi_n \in (b - \delta, b)$ so, dass

$$\frac{f(x_n) - f(b - \delta)}{g(x_n) - g(b - \delta)} = \frac{f'(\xi_n)}{g'(\xi_n)}.$$

Formt man um so ergibt sich

$$f(x_n) - f(b - \delta) = \frac{f'(\xi_n)}{g'(\xi_n)}(g(x_n) - g(b - \delta))$$

d.h.

$$f(x_n) = f(b - \delta) + \frac{f'(\xi_n)}{g'(\xi_n)}(g(x_n) - g(b - \delta)).$$

Teilen wir nun durch $g(x_n)$ auf beiden Seiten so ergibt sich

$$\frac{f(x_n)}{g(x_n)} = \frac{f(b - \delta)}{g(x_n)} + \frac{f'(\xi_n)}{g'(\xi_n)} \left(1 - \frac{g(b - \delta)}{g(x_n)}\right).$$

Nun gilt wegen der „ $\frac{?}{\infty}$ -Bedingung“

$$\lim_{n \rightarrow \infty} \frac{f(b - \delta)}{g(x_n)} = 0 \quad \text{und} \quad \lim_{n \rightarrow \infty} \left(1 - \frac{g(b - \delta)}{g(x_n)}\right) = 1. \quad (378)$$

Somit gibt es auch ein $N_1 \geq N$ mit $\left(1 - \frac{g(b - \delta)}{g(x_n)}\right) > 0$ für alle $n \geq N_1$. Für $n \geq N_1$ hat man dann

$$\frac{f(x_n)}{g(x_n)} = \frac{f(b - \delta)}{g(x_n)} + \frac{f'(\xi_n)}{g'(\xi_n)} \left(1 - \frac{g(b - \delta)}{g(x_n)}\right) \stackrel{b - \delta < \xi_n < b}{\underset{(377)}{\leq}} \frac{f(b - \delta)}{g(x_n)} + (L + \varepsilon) \left(1 - \frac{g(b - \delta)}{g(x_n)}\right)$$

und

$$\frac{f(x_n)}{g(x_n)} = \frac{f(b - \delta)}{g(x_n)} + \frac{f'(\xi_n)}{g'(\xi_n)} \left(1 - \frac{g(b - \delta)}{g(x_n)}\right) \stackrel{b - \delta < \xi_n < b}{\underset{(377)}{\geq}} \frac{f(b - \delta)}{g(x_n)} + (L - \varepsilon) \left(1 - \frac{g(b - \delta)}{g(x_n)}\right).$$

Aus der ersten Ungleichung schließt man

$$\limsup_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} \leq \limsup_{n \rightarrow \infty} \frac{f(b - \delta)}{g(x_n)} + (L + \varepsilon) \left(1 - \frac{g(b - \delta)}{g(x_n)}\right)$$

$$\begin{aligned}
& \stackrel{(378)}{=} \lim_{n \rightarrow \infty} \frac{f(b - \delta)}{g(x_n)} + (L + \varepsilon) \left(1 - \frac{g(b - \delta)}{g(x_n)} \right) \\
& \stackrel{(378)}{=} 0 + (L + \varepsilon) \cdot 1 = L + \varepsilon
\end{aligned}$$

Aus der zweiten Ungleichung hingegen folgert man analog

$$\begin{aligned}
\liminf_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} & \geq \liminf_{n \rightarrow \infty} \frac{f(b - \delta)}{g(x_n)} + (L - \varepsilon) \left(1 - \frac{g(b - \delta)}{g(x_n)} \right) \\
& \stackrel{(378)}{=} \lim_{n \rightarrow \infty} \frac{f(b - \delta)}{g(x_n)} + (L - \varepsilon) \left(1 - \frac{g(b - \delta)}{g(x_n)} \right) \\
& \stackrel{(378)}{=} 0 + (L - \varepsilon) \cdot 1 = L - \varepsilon.
\end{aligned}$$

Die vorigen zwei Gleichungen liefern

$$L - \varepsilon \leq \liminf_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} \leq \limsup_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} \leq L + \varepsilon.$$

Nun kann $\varepsilon > 0$ aber beliebig klein gewählt werden. Da L die einzige Zahl ist, die sich für alle $\varepsilon > 0$ zwischen $L - \varepsilon$ und $L + \varepsilon$ aufhält, folgert man

$$\liminf_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} = \limsup_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} = L$$

Nach Proposition 3.5.30 muss bei Gleichheit von $\limsup$ und $\liminf$ auch der Grenzwert existieren und man folgert

$$\lim_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} = L$$

Da die Folge $(x_n)_{n \in \mathbb{N}} \subseteq (a, b)$ mit $\lim_{n \rightarrow \infty} x_n = b$ beliebig war gilt

$$\lim_{x \rightarrow b} \frac{f(x)}{g(x)} = L \stackrel{(375)}{=} \lim_{x \rightarrow b} \frac{f'(x)}{g'(x)}.$$

^aÜberprüfen Sie als Übungsaufgabe, warum dieser anwendbar ist!

Beispiel 7.3.30

Wir berechnen

$$\lim_{x \rightarrow \infty} (e^x + x)^{\frac{1}{x}}.$$

Dazu formen wir zunächst um: Es gilt für $x > 0$, dass

$$(e^x + x)^{\frac{1}{x}} = \exp\left(\frac{1}{x} \log(e^x + x)\right) \tag{379}$$

Wir werden daher zuerst $\lim_{x \rightarrow \infty} \frac{1}{x} \log(e^x + x)$ untersuchen. Hierzu verwenden wir die l'Hospital'sche

$\frac{?}{\infty}$ -Regel (Satz 7.3.29). Dazu: Da $\lim_{x \rightarrow \infty} x = \infty$ gilt

$$\lim_{x \rightarrow \infty} \frac{\log(e^x + x)}{x} \stackrel{„\frac{?}{\infty}\text{-l'H}“}{=} \lim_{x \rightarrow \infty} \frac{\frac{d}{dx} \log(e^x + x)}{\frac{d}{dx} x} = \lim_{x \rightarrow \infty} \frac{\frac{e^x + 1}{e^x + x}}{1} = \lim_{x \rightarrow \infty} \frac{e^x + 1}{e^x + x} = \lim_{x \rightarrow \infty} \frac{1 + \frac{1}{e^x}}{1 + \frac{x}{e^x}}.$$

Da nach Proposition 6.3.16 $\lim_{x \rightarrow \infty} \frac{1}{e^x} = \lim_{x \rightarrow \infty} \frac{x}{e^x} = 0$, gilt (mit den Grenzwertregeln)

$$\lim_{x \rightarrow \infty} \frac{\log(e^x + x)}{x} = 1.$$

Aus (379) folgt

$$\lim_{x \rightarrow \infty} (e^x + x)^{\frac{1}{x}} = \lim_{x \rightarrow \infty} \exp\left(\frac{1}{x} \log(e^x + x)\right) \stackrel{\text{exp stetig bei } x_0 = 1}{=} \exp(1) = e.$$

Beispiel 7.3.31

Wir wissen bereits aus Proposition 6.3.16, dass

$$\lim_{x \rightarrow \infty} \frac{x^m}{e^x} = 0 \quad \forall m \in \mathbb{N}.$$

Dies könnte man aber auch mit der Regel von l'Hospital zeigen. Dazu beobachtet man für $m \in \mathbb{N}$

$$\lim_{x \rightarrow \infty} \frac{x^m}{e^x} \stackrel{„\frac{?}{\infty}\text{-l'H}“}{=} \lim_{x \rightarrow \infty} \frac{\frac{d}{dx}(x^m)}{\frac{d}{dx}(e^x)} = \lim_{x \rightarrow \infty} \frac{mx^{m-1}}{e^x} = m \lim_{x \rightarrow \infty} \frac{x^{m-1}}{e^x}, \quad (380)$$

sofern der Grenzwert auf der rechten Seite existiert (vgl. die Voraussetzungen in 7.3.29). Für die Berechnung des Grenzwertes $\lim_{x \rightarrow \infty} \frac{x^{m-1}}{e^x}$ können wir nun dieselbe Rechnung wie in (380) (mit $m - 1$ anstatt m) reproduzieren und erhalten (sofern der letztgenannte Grenzwert existiert)

$$\lim_{x \rightarrow \infty} \frac{x^m}{e^x} \stackrel{„\frac{?}{\infty}\text{-l'H}“}{=} m \lim_{x \rightarrow \infty} \frac{x^{m-1}}{e^x} \stackrel{„\frac{?}{\infty}\text{-l'H}“}{=} m(m-1) \lim_{x \rightarrow \infty} \frac{x^{m-2}}{e^x}.$$

Iteriert man diese Rechnung m -mal so erhält man

$$\lim_{x \rightarrow \infty} \frac{x^m}{e^x} = \dots = [m(m-1) \cdot \dots \cdot 1] \lim_{x \rightarrow \infty} \frac{x^0}{e^x} = m! \lim_{x \rightarrow \infty} \frac{1}{e^x} = 0.$$

Formell richtig würde man das wohl nun noch mit Induktion beweisen müssen.

7.4 Höhere Ableitungen

Viele Funktionen kann man auch mehrmals ableiten. Die Ableitungsfunktion der Ableitungsfunktion nennen wir im folgenden *zweite Ableitung*. Ist diese Ableitung wiederum differenzierbar, so können wir wiederum eine Ableitung bilden.

Definition 7.4.1 — HÖHERE ABLEITUNGEN

Es sei $D \subseteq \mathbb{K}_1$ und $f : D \rightarrow \mathbb{K}_2$ eine auf D differenzierbare Funktion. Ist die Ableitungsfunktion $f' : D \rightarrow \mathbb{C}$ bei $x_0 \in D$ differenzierbar, so heißt f bei $x_0 \in D$ *zweimal differenzierbar* und man schreibt

$$f''(x_0) := (f')'(x_0)$$

Wir sagen f ist zweimal auf D differenzierbar, falls f an jedem Punkt $x_0 \in D$ zweimal differenzierbar ist. In diesem Fall nennen wir $f'' : D \rightarrow \mathbb{C}$ die *zweite Ableitungsfunktion*. Eine alternative Schreibweise ist $f^{(2)} : D \rightarrow \mathbb{C}$.

Allgemein können wir induktiv Folgendes definieren: Sei $k \in \mathbb{N}$. Wir sagen eine Funktion $f : D \rightarrow \mathbb{K}_2$ ist bei einem Punkt $x_0 \in D$ $(k+1)$ -mal differenzierbar falls f auf D k -mal differenzierbar ist und die k -te Ableitungsfunktion $f^{(k)} : D \rightarrow \mathbb{C}$ bei x_0 noch ein weiteres Mal differenzierbar ist.

Auch hier nennt man f auf ganz D *k -mal differenzierbar*, falls f bei jedem Punkt $x_0 \in D$ k -mal differenzierbar ist.

Weitere Schreibweisen für $f^{(k)}(x_0)$ sind

$$\frac{d^k f}{dx^k}(x_0), \quad \text{und} \quad \left. \frac{d^k}{dx^k} f(x) \right|_{x=x_0}.$$

Man kann auch definieren, dass man mit der *0-ten Ableitung* $f^{(0)}$ die Funktion selbst meint. Dies werden wir im Folgenden verwenden.

Bemerkung 7.4.2

An dieser Stelle könnte man der Vollständigkeit halber noch folgende Behauptung mit der vollständigen Induktion zeigen: Ist $f : D \rightarrow \mathbb{K}_2$ eine k -mal auf D differenzierbare Funktion ($k \in \mathbb{N}, k \geq 2$), so sind die ersten $(k-1)$ Ableitungen $f^{(1)}, \dots, f^{(k-1)}$ auf D differenzierbar und es gilt

$$(f^{(k-1)})'(x) = f^{(k)}(x) \quad \text{für alle } x \in D.$$

Oft ist es auch hilfreich wenn die (höheren) Ableitungen auch stetig sind. Wir sprechen dann davon, dass eine Funktion (k -mal) *stetig differenzierbar* ist.

Definition 7.4.3 — STETIG DIFFERENZIERBAR

Eine Funktion $f : D \rightarrow \mathbb{K}_2$ heißt auf D *stetig differenzierbar*, falls f auf ganz D differenzierbar ist und die Ableitungsfunktion $f' : D \rightarrow \mathbb{C}$ auf D stetig ist. Eine Funktion $f : D \rightarrow \mathbb{K}_2$ heißt auf D *k -mal stetig differenzierbar*, falls f auf ganz D k -mal differenzierbar ist und

die k -te Ableitungsfunktion $f^{(k)} : D \rightarrow \mathbb{C}$ auf D stetig ist.

Bemerkung 7.4.4

Ist eine Funktion $f : D \rightarrow \mathbb{K}_2$ k -mal stetig differenzierbar, so ist nicht nur die k -te Ableitungsfunktion stetig, sondern auch die ersten $k-1$ Ableitungsfunktionen $f^{(1)}, \dots, f^{(k-1)}$. In der Tat: Nach Bemerkung 7.4.2 sind $f^{(1)}, \dots, f^{(k-1)}$ differenzierbar. Da differenzierbare Funktionen auf D auch stetig sind (Korollar 7.0.7) folgt die Stetigkeit von $f^{(1)}, \dots, f^{(k-1)}$.

Der Begriff *stetig differenzierbar* sollte nicht mit *stetig und differenzierbar* verwechselt werden. Die Aussage „ $f : D \rightarrow \mathbb{K}_2$ ist auf D stetig und differenzierbar“ würde ja nichts anderes bedeuten, als dass f differenzierbar ist (denn differenzierbare Funktionen sind automatisch stetig, vgl. Korollar 7.0.7). Die Aussage „ $f : D \rightarrow \mathbb{K}_2$ ist auf D stetig differenzierbar“ bedeutet allerdings mehr: Zum einen muss $f : D \rightarrow \mathbb{K}_2$ differenzierbar sein aber zum anderen muss auch $f' : D \rightarrow \mathbb{C}$ stetig sein.

Beispiel 7.4.5

Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$f(x) := \begin{cases} x^2 \sin \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

ist auf ganz $\mathbb{R}$ differenzierbar. Allerdings ist f auf $\mathbb{R}$ nicht stetig differenzierbar.

Beweis. Die Differenzierbarkeit auf $\mathbb{R} \setminus \{0\}$ folgt mit den elementaren Ableitungsregeln aus Sektion 7.1. Man sieht, dass für alle $x \in \mathbb{R} \setminus \{0\}$ gilt

$$f'(x) = \frac{d}{dx}(x^2 \sin(\frac{1}{x})) = 2x \sin(\frac{1}{x}) + x^2 \cos(\frac{1}{x})(-\frac{1}{x^2}) = 2x \sin(\frac{1}{x}) - \cos(\frac{1}{x}). \quad (381)$$

Zur Differenzierbarkeit von f bei $x_0 = 0$ beobachte, dass für $x \in \mathbb{R} \setminus \{0\}$ gilt

$$\frac{f(x) - f(0)}{x - 0} = \frac{x^2 \sin(\frac{1}{x}) - 0}{x - 0} = \frac{x^2 \sin(\frac{1}{x})}{x} = x \sin(\frac{1}{x}).$$

Deswegen folgert man, dass

$$f'(0) = \lim_{x \rightarrow 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \rightarrow 0} x \sin(\frac{1}{x}) = 0.$$

Dies und (381) impliziert, dass $f' : \mathbb{R} \rightarrow \mathbb{R}$ gegeben ist durch

$$f'(x) = \begin{cases} 2x \sin(\frac{1}{x}) - \cos(\frac{1}{x}) & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Wir behaupten nun, dass f' bei $x_0 = 0$ nicht stetig ist. Dazu betrachte die Folge $(x_n)_{n \in \mathbb{N}} = (\frac{1}{2\pi n})_{n \in \mathbb{N}}$. Dann gilt $\lim_{n \rightarrow \infty} x_n = 0$ und

$$\lim_{n \rightarrow \infty} f'(x_n) = \lim_{n \rightarrow \infty} [2x_n \sin(\frac{1}{x_n}) - \cos(\frac{1}{x_n})] = \lim_{n \rightarrow \infty} [\frac{2}{2\pi n} \sin(2\pi n) - \cos(2\pi n)] = \lim_{n \rightarrow \infty} -1 = -1.$$

Beachte: $f'(0) = 0 \neq -1$! Somit haben wir eine Folge $(x_n)_{n \in \mathbb{N}}$ gefunden mit $\lim_{n \rightarrow \infty} x_n = 0$ und

$$\neg (\lim_{n \rightarrow \infty} f'(x_n) = f'(0)).$$

Es folgt aus Definition 6.1.3

$$\neg (\lim_{x \rightarrow 0} f'(x) = f'(0)).$$

Sprich: f' ist bei $x_0 = 0$ unstetig.

Der Satz von Taylor

Wir haben in Proposition 7.0.6 gesehen: Mithilfe der ersten Ableitung kann man eine Funktion f an einem Punkt x_0 durch eine *affin-lineare* Funktion, also durch ein *Polynom ersten Grades* approximieren. Mithilfe der ersten k Ableitungen lässt sich f (unter bestimmten Umständen) analog durch ein *Polynom k -ten Grades* approximieren. Dies besagt der *Approximationssatz von Taylor*. Er folgt aus der sogenannten *Taylor-Formel*, welche wir im Folgenden diskutieren.

Satz 7.4.6 — TAYLOR-FORMEL

Es seien $a, b \in \mathbb{R} \cup \{\pm\infty\}$, $a < b$ und $f : (a, b) \rightarrow \mathbb{R}$ auf (a, b) k -mal differenzierbar ($k \in \mathbb{N}$). Ferner sei $x_0 \in (a, b)$. Dann gibt es für alle $x \in (a, b)$, $x \neq x_0$ ein $\xi \in (\min(x, x_0), \max(x, x_0))$ so, dass

$$f(x) = \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell + \frac{f^{(k)}(\xi)}{k!} (x - x_0)^k. \quad (382)$$

Beweis.

FALL 1. $k = 1$.

Für $k = 1$ liest sich (382) als

$$f(x) = f^{(0)}(x_0) + \frac{1}{1!} f^{(1)}(\xi)(x - x_0)^1 = f(x_0) + f'(\xi)(x - x_0)$$

Ziehen wir $f(x_0)$ auf beiden Seiten ab, so erhalten wir

$$f(x) - f(x_0) = f'(\xi)(x - x_0) \quad \text{für ein } \xi \in (\min(x, x_0), \max(x, x_0)). \quad (383)$$

Diese Formel (*) folgt direkt aus dem ersten Mittelwertsatz (Satz 7.3.1), vgl. auch (351) mit $a = \min\{x, x_0\}$ und $b = \max\{x, x_0\}$. In der Tat: Da f auf $[\min(x, x_0), \max(x, x_0)] \subseteq (a, b)$

stetig und auf $(\min(x, x_0), \max(x, x_0))$ differenzierbar ist gilt

$$f(\max\{x, x_0\}) - f(\min\{x, x_0\}) = f'(\xi)(\max\{x, x_0\} - \min\{x, x_0\})$$

für ein $\xi \in (\min(x, x_0), \max(x, x_0))$. Ist nun $x \geq x_0$ so liefert obige Gleichung

$$f(x) - f(x_0) = f'(\xi)(x - x_0) \quad \text{für ein } \xi \in (\min(x, x_0), \max(x, x_0)).$$

Ist hingegen $x \leq x_0$ so hat man zunächst

$$f(x_0) - f(x) = f'(\xi)(x_0 - x) \quad \text{für ein } \xi \in (\min(x, x_0), \max(x, x_0)),$$

woraus sich (383) durch Durchmultiplizieren mit -1 ergibt.

FALL 2. $k > 1$.

Fixiere $x \in (a, b)$ mit $x \neq x_0$. Definiere die reelle Konstante

$$\theta := \frac{k!}{(x - x_0)^k} \left[f(x) - \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell \right] \in \mathbb{R}. \quad (384)$$

Formt man die Formel in (382) an, so erhält man eine äquivalente Formulierung der Behauptung:

$$\mathbb{Z}\mathbb{Z}: \quad \theta = f^{(k)}(\xi) \quad \text{für ein } \xi \in (\min\{x, x_0\}, \max\{x, x_0\})$$

Definiere dazu die Hilfsfunktion

$$h : (a, b) \rightarrow \mathbb{R}, \quad h(t) := \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(t)}{\ell!} (x - t)^\ell + \frac{\theta}{k!} (x - t)^k.$$

Dann ist h nach Bemerkung 7.4.2 auf (a, b) differenzierbar.

Zwischenbehauptung. Es gilt $h(x) = h(x_0)$

Dazu:

$$h(x) = \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x)}{\ell!} \underbrace{(x - x)^\ell}_{=0 \text{ falls } \ell \neq 0} + \frac{\theta}{k!} \underbrace{(x - x)^k}_{=0} = \frac{f^{(0)}(x)}{0!} (x - x)^0 = \frac{f(x)}{1} \cdot 0^0 = f(x)$$

und

$$h(x_0) = \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell + \frac{\theta}{k!} (x - x_0)^k$$

$$\begin{aligned}
& \stackrel{(384)}{=} \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x-x_0)^\ell + \frac{(x-x_0)^k}{k!} \frac{k!}{(x-x_0)^k} \left[f(x) - \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x-x_0)^\ell \right] \\
& = \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x-x_0)^\ell + f(x) - \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x-x_0)^\ell = f(x).
\end{aligned}$$

Dies und die vorige Gleichung ergibt

$$h(x) = h(x_0) = f(x).$$

Zwischenbehauptung Ende.

Nun ist h auch $[\min\{x, x_0\}, \max\{x, x_0\}] \subseteq (a, b)$ stetig und im Inneren des Intervalls differenzierbar. Dies uns die Zwischenbehauptung zeigen, dass die Voraussetzungen für den Satz von Rolle (Lemma 7.3.3) erfüllt sind. Aus dem Satz von Rolle folgt dann

$$h'(\xi) = 0 \quad \text{für ein } \xi \in (\min\{x, x_0\}, \max\{x, x_0\}). \quad (385)$$

Nun berechnen wir für $t \in (a, b)$

$$\begin{aligned}
h'(t) &= \frac{d}{dt} \left(\sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(t)}{\ell!} (x-t)^\ell + \frac{\theta}{k!} (x-t)^k \right) \\
&\stackrel{\text{Prop 7.1.1}}{=} \frac{d}{dt} \left(\frac{f^{(0)}(t)}{0!} (x-t)^0 + \sum_{\ell=1}^{k-1} \frac{f^{(\ell)}(t)}{\ell!} (x-t)^\ell \right) + \frac{\theta}{k!} k(x-t)^{k-1}(-1) \\
&= \frac{d}{dt} \left(f^{(0)}(t) + \sum_{\ell=1}^{k-1} \frac{f^{(\ell)}(t)}{\ell!} (x-t)^\ell \right) - \frac{\theta}{(k-1)!} (x-t)^{k-1} \\
&\stackrel{\text{Prop 7.1.1}}{=} f^{(1)}(t) + \sum_{\ell=0}^{k-1} \frac{d}{dt} \left(\frac{f^{(\ell)}(t)}{\ell!} (x-t)^\ell \right) - \frac{\theta}{(k-1)!} (x-t)^{k-1} \\
&\stackrel{\text{Prop 7.1.1}}{=} f^{(1)}(t) + \sum_{\ell=1}^{k-1} \frac{f^{(\ell+1)}(t)}{\ell!} (x-t)^\ell + \frac{f^{(\ell)}(t)}{\ell!} \frac{d}{dt} (x-t)^\ell - \frac{\theta}{(k-1)!} (x-t)^{k-1} \\
&\stackrel{\text{Prop 7.1.5}}{=} f^{(1)}(t) + \sum_{\ell=1}^{k-1} \frac{f^{(\ell+1)}(t)}{\ell!} (x-t)^\ell + \frac{f^{(\ell)}(t)}{\ell!} \ell (x-t)^{\ell-1}(-1) - \frac{\theta}{(k-1)!} (x-t)^{k-1} \\
&= f^{(1)}(t) + \sum_{\ell=1}^{k-1} \frac{f^{(\ell+1)}(t)}{\ell!} (x-t)^\ell - \frac{f^{(\ell)}(t)}{(\ell-1)!} (x-t)^{\ell-1} - \frac{\theta}{(k-1)!} (x-t)^{k-1} \\
&\stackrel{\text{Teleskopsumme}}{=} f^{(1)}(t) + \frac{f^{(k)}(t)}{(k-1)!} (x-t)^{k-1} - \frac{f^{(1)}(t)}{0!} (x-t)^0 - \frac{\theta}{(k-1)!} (x-t)^{k-1} \\
&= f^{(1)}(t) + \frac{f^{(k)}(t)}{(k-1)!} (x-t)^{k-1} - f^{(1)}(t) \frac{\theta}{(k-1)!} (x-t)^{k-1}
\end{aligned}$$

$$= \frac{f^{(k)}(t)}{k!}(x-t)^k - \frac{\theta}{(k-1)!}(x-t)^{k-1} = \frac{f^{(k)}(t) - \theta}{(k-1)!}(x-t)^{k-1}$$

Es sei nun ξ wie in (385). Dann gilt $h'(\xi) = 0$ also

$$\frac{f^{(k)}(\xi) - \theta}{(k-1)!}(x-\xi)^{k-1} = 0.$$

Da $\xi \in (\min\{x, x_0\}, \max\{x, x_0\})$ gilt $\xi \neq x$. Wir folgern daher

$$f^{(k)}(\xi) - \theta = 0.$$

Das bedeutet

$$f^{(k)}(\xi) = \theta \stackrel{(384)}{=} \frac{k!}{(x-x_0)^k} \left[f(x) - \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x-x_0)^\ell \right]$$

Stellt man dies nach $f(x)$ um, so ergibt sich

$$f(x) = \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x-x_0)^\ell + \frac{f^{(k)}(\xi)}{k!} (x-x_0)^k,$$

was zu zeigen war.

Schaut man sich den Ausdruck auf der rechten Seite von (382) an, so hat man auf den ersten Blick ein Polynom k -ten Grades in x . Hierbei gibt es aber ein Problem: Die Zwischenstelle ξ wird noch von x abhängig sein! Daher ist der obige Ausdruck in der Realität kein Polynom in x . Nichtsdestotrotz können wir auf Basis der obigen Formel ein sogenanntes *Taylor-Polynom* finden, was f gut approximiert.

Proposition 7.4.7 — TAYLOR-APPROXIMATION

Sei $k \in \mathbb{N}$. Es sei $f : (a, b) \rightarrow \mathbb{R}$ k -mal stetig differenzierbar und sei $x_0 \in (a, b)$. Dann gibt es eine stetige Funktion $\phi : (a, b) \rightarrow \mathbb{R}$ mit $\phi(x_0) = 0$ so, dass

$$f(x) = \sum_{\ell=0}^k \frac{f^{(\ell)}(x_0)}{\ell!} (x-x_0)^\ell + \phi(x)(x-x_0)^k \quad \forall x \in (a, b) \quad (386)$$

Man beachte: Im Fall $k = 1$ ergibt sich die Formel

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \phi(x)(x-x_0),$$

die wir auch aus der linearen Approximation differenzierbarer Funktionen (vgl. Proposition 7.0.6) kennen. Der Term $\phi(x)(x-x_0)^k$ wird *Approximationsfehler k -ter Ordnung bei x_0* genannt. Für $x \rightarrow x_0$ geht dieser Fehler schneller gegen 0 als $(x-x_0)^k$.

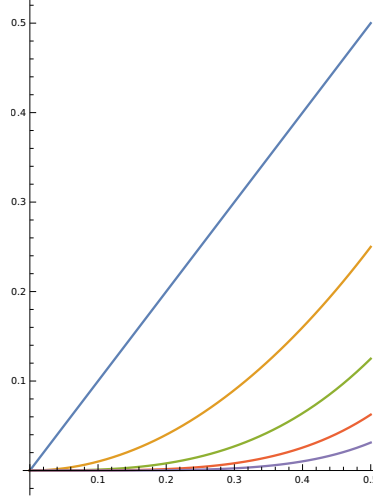


Abbildung 34: Die Monome $(x - 0)^k$ für $k = 1$ (blau), $k = 2$ (gelb), $k = 3$ (grün), $k = 4$ (rot), $k = 5$ (lila). Für $x \approx x_0$ werden diese Terme immer schneller klein.

Beweis von Proposition 7.4.7. Definiere $\phi : (a, b) \rightarrow \mathbb{R}$ durch

$$\phi(x) := \begin{cases} \frac{1}{(x-x_0)^k} \left[f(x) - \sum_{\ell=0}^k \frac{f^{(\ell)}(x_0)}{\ell!} (x-x_0)^\ell \right] & \text{falls } x \neq x_0 \\ 0 & \text{falls } x = x_0 \end{cases} \quad (387)$$

Beachte: ϕ ist bereits so gewählt, dass $\phi(x_0) = 0$. Außerdem gilt für $x \in (a, b) : x \neq x_0$ klarerweise, dass

$$\phi(x)(x-x_0)^k = f(x) - \sum_{\ell=0}^k \frac{f^{(\ell)}(x_0)}{\ell!} (x-x_0)^\ell.$$

Man erhält daraufhin (386) für $x \neq x_0$ wenn man nach $f(x)$ umstellt. Gleichung (386) gilt auch für $x = x_0$. Davon kann man sich leicht durch Einsetzen überzeugen: Auf der linken Seite steht dann $f(x_0)$ und auf der rechten Seite

$$\sum_{\ell=0}^k \frac{f^{(\ell)}(x_0)}{\ell!} (x_0 - x_0)^\ell + \phi(x_0)(x_0 - x_0)^k = f^{(0)}(x_0) = f(x_0).$$

Nun muss man nur noch zeigen, dass ϕ auf (a, b) stetig ist. Auf $(a, b) \setminus \{x_0\}$ ist ϕ nach dem ersten Fall von (387) als Komposition stetiger Funktionen stetig. Noch zu zeigen ist die Stetigkeit von ϕ bei x_0 , d.h.

$$\mathbb{Z} : \quad \lim_{x \rightarrow x_0} \phi(x) = \phi(x_0) = 0. \quad (388)$$

Es sei dazu $x \in (a, b), x \neq x_0$. Aus Satz 7.4.6 wissen wir, dass es ein (im Allgemeinen von

x abhängiges) $\xi_x \in (\min\{x, x_0\}, \max\{x, x_0\})$ gibt mit

$$f(x) = \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell + \frac{f^{(k)}(\xi_x)}{k!} (x - x_0)^k.$$

Damit folgt

$$\begin{aligned} f(x) - \sum_{\ell=0}^k \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell &= \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell + \frac{f^{(k)}(\xi_x)}{k!} (x - x_0)^k - \sum_{\ell=0}^k \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell \\ &= \frac{f^{(k)}(\xi_x)}{k!} (x - x_0)^k - \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = \frac{f^{(k)}(\xi_x) - f^{(k)}(x_0)}{k!} (x - x_0)^k. \end{aligned}$$

Somit gilt für alle $x \in (a, b)$ mit $x \neq x_0$, dass

$$\phi(x) = \frac{1}{(x - x_0)^k} \left[f(x) - \sum_{\ell=0}^k \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell \right] = \frac{f^{(k)}(\xi_x) - f^{(k)}(x_0)}{k!} \quad (389)$$

Diese Gleichung stimmt auch für $x = x_0$ wenn man $\xi_{x_0} := x_0$ definiert. Man beachte: Wegen dieser Wahl gilt künftig im Allgemeinen nicht mehr $\xi_x \in (\min\{x, x_0\}, \max\{x, x_0\})$ sondern $\xi_x \in [\min\{x, x_0\}, \max\{x, x_0\}]$. Es sei nun $(x_n)_{n \in \mathbb{N}} \subseteq (a, b)$ eine beliebige Folge mit $\lim_{n \rightarrow \infty} x_n = x_0$. Da $\min\{x_n, x_0\} \leq \xi_{x_n} \leq \max\{x_n, x_0\}$ gilt auch nach dem Sandwichlemma (vgl. Proposition 2.3.26)

$$\lim_{n \rightarrow \infty} \xi_{x_n} = x_0. \quad (390)$$

Wir folgern aus (389), dass

$$\lim_{n \rightarrow \infty} \phi(x_n) = \lim_{n \rightarrow \infty} \frac{f^{(k)}(\xi_{x_n}) - f^{(k)}(x_0)}{k!} \stackrel{(390) \ \& \ \text{Stetigkeit von } f^{(k)}}{=} \frac{f^{(k)}(x_0) - f^{(k)}(x_0)}{k!} = 0.$$

Da die Folge $(x_n)_{n \in \mathbb{N}}$ beliebig gewählt war, folgt die noch zu zeigende Gleichung (388).

Das approximierende Polynom, welches wir in Proposition 7.4.7 identifiziert haben, nennen wir *k-tes Taylor-Polynom*. Genauer:

Definition 7.4.8 — TAYLOR-POLYNOM

Es sei $f : (a, b) \rightarrow \mathbb{R}$ bei $x_0 \in (a, b)$ k -mal differenzierbar ($k \in \mathbb{N}$). Dann ist das *k-te Taylor-Polynom von f bei x_0* (im Zeichen $T_k[f, x_0]$) gegeben durch

$$T_k[f, x_0](x) := \sum_{\ell=0}^k \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell.$$

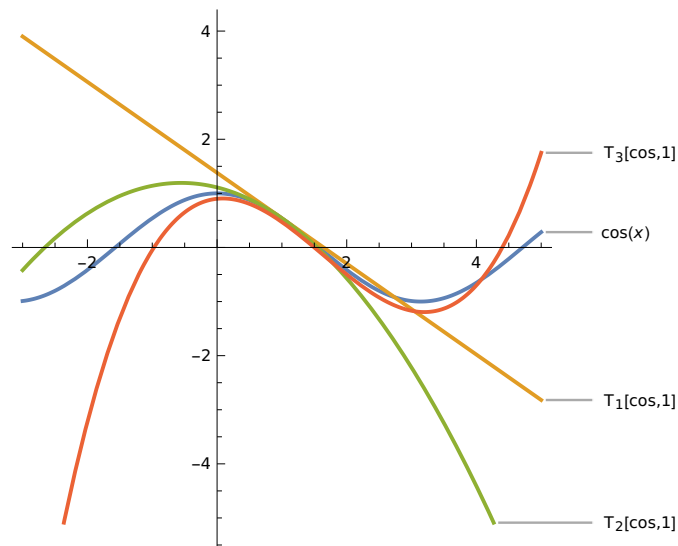


Abbildung 35: Der Kosinus und seine Taylorpolynome bei $x_0 = 1$, mehr davon in GEOGEBRA.

Bemerkung 7.4.9

Proposition 7.4.7 kann mit der obigen Definition umformuliert werden: Ist $f : (a, b) \rightarrow \mathbb{R}$ k -mal stetig differenzierbar so gilt

$$f(x) = T_k[f, x_0](x) + \phi(x)(x - x_0)^k,$$

wobei $\phi : (a, b) \rightarrow \mathbb{R}$ stetig ist mit $\phi(x_0) = 0$. Mit anderen Worten:

Bis auf einen Approximationsfehler k -ter Ordnung bei x_0 stimmt f mit seinem k -ten Taylorpolynom überein.

Somit erhalten wir in einer *kleinen Umgebung von* x_0 gute Approximationen verschiedener Funktionen. Ob diese Polynome auch in einer größeren Umgebung von x_0 eine gute Approximation sind werden wir später besprechen.

Es gilt also für eine genügend oft differenzierbare Funktion $f : (a, b) \rightarrow \mathbb{R}$

$$\begin{aligned} T_1[f, x_0](x) &= f(x_0) + f'(x_0)(x - x_0), \\ T_2[f, x_0](x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 \\ T_3[f, x_0](x) &= f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + \frac{1}{6}f'''(x_0)(x - x_0)^3, \end{aligned} \quad (391)$$

und so weiter.

27.04.2026

Beispiel 7.4.10

Wir berechnen für $k \in \mathbb{N}$ das Taylorpolynom

$$T_k[\exp, 0]$$

Dazu berechne mithilfe der Regel $\exp' = \exp$ (vgl. Korollar 7.0.11 (ii))

$$\exp(0) = 1, \quad \exp'(0) = \exp(0) = 1, \quad \exp''(0) = (\exp')'(0) = \exp'(0) = \exp(0) = 1, \dots$$

Man überzeugt sich also leicht, dass

$$\exp^{(\ell)}(0) = \exp(0) = 1 \quad \forall \ell \in \mathbb{N} \quad (392)$$

Man erhält für $k \in \mathbb{N}$

$$T_k[\exp, 0](x) = \sum_{\ell=0}^k \frac{\exp^{(\ell)}(0)}{\ell!} (x-0)^\ell \stackrel{(392)}{=} \sum_{\ell=0}^k \frac{x^\ell}{\ell!}. \quad (393)$$

Gleichung (393) motiviert die Formel, die wir in Definition 5.4.2 genommen haben, um die Exponentialfunktion zu definieren. Mehr Details dazu (z.B. zum Grenzübergang $k \rightarrow \infty$) besprechen wir später!

Kriterien höherer Ordnung für lokale Extrema

Die Taylor-Approximation kann benutzt werden um ein hinreichendes Kriterium für lokale Extrema von Funktionen $f : (a, b) \rightarrow \mathbb{R}$ zu beweisen. In der Schule haben wir dieses Kriterium ein *Kriterium zweiter Ordnung* genannt und oft angewendet.

Proposition 7.4.11 — KRITERIUM ZWEITER ORDNUNG FÜR LOKALE EXTREMA

Es sei $f : (a, b) \rightarrow \mathbb{R}$ eine 2-mal stetig differenzierbare Funktion und $x_0 \in (a, b)$.

- (i) Ist $f'(x_0) = 0$ und $f''(x_0) > 0$, so ist x_0 ein striktes lokales Minimum von f
- (ii) Ist $f'(x_0) = 0$ und $f''(x_0) < 0$, so ist x_0 ein striktes lokales Maximum von f

Beweis. Es muss nur (i) gezeigt werden, denn (ii) folgt aus (i) durch Betrachtung von $-f$. Mit Bemerkung 7.4.9 erhalten wir: Es gibt ein $\phi : (a, b) \rightarrow \mathbb{R}$ stetig mit $\phi(x_0) = 0$ so, dass für alle $x \in (a, b)$ gilt

$$f(x) = T_2[f, x_0] + \phi(x)(x-x_0)^2 \stackrel{(391)}{=} f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2}f''(x_0)(x-x_0)^2 + \phi(x)(x-x_0)^2.$$

Da $f'(x_0) = 0$ erhalten wir für $x \in (a, b)$

$$f(x) = f(x_0) + [\tfrac{1}{2}f''(x_0) + \phi(x)](x-x_0)^2. \quad (394)$$

Weil ϕ stetig ist und $\phi(x_0) = 0$ folgt, dass $\lim_{x \rightarrow x_0} \phi(x) = 0$. Mit dem Epsilon-Delta-Kriterium (Satz 6.1.23), angewandt mit $\varepsilon = \frac{1}{4}f''(x_0) > 0$ erhält man

$$\exists \delta > 0 : \quad \forall x \in (a, b) \text{ mit } |x - x_0| < \delta \quad \text{gilt} \quad |\phi(x) - 0| < \frac{1}{4}f''(x_0).$$

Für $x \in (a, b) \cap B_\delta(x_0)$ (d.h. für $x \in (a, b)$ mit $|x - x_0| < \delta$) gilt dann also $|\phi(x)| < \frac{1}{4}f''(x_0)$ und somit

$$\frac{1}{2}f''(x_0) + \phi(x) \geq \frac{1}{2}f''(x_0) - |\phi(x)| \geq \frac{1}{2}f''(x_0) - \frac{1}{4}f''(x_0) = \frac{1}{4}f''(x_0).$$

Mit (394) ergibt sich für alle $x \in (a, b) \cap B_\delta(x_0)$, dass

$$f(x) \geq f(x_0) + \frac{1}{4}f''(x_0)(x - x_0)^2 \quad (395)$$

Da $f''(x_0) > 0$ gilt $\frac{1}{4}f''(x_0)(x - x_0)^2 > 0$ für alle $x \neq x_0$. Wir folgern aus (395), dass

$$f(x) > f(x_0) \quad \forall x \in (a, b) \cap B_\delta(x_0) \setminus \{x_0\}.$$

Das heißt nach Definition 7.2.1 nichts anderes als dass f ein striktes lokales Minimum bei x_0 annimmt.

Das Kriterium zweiter Ordnung aus (Proposition 7.4.11) haben Sie in der Schule oft angewendet um Funktionen auf strikte lokale Extrema zu untersuchen. Allerdings kann nicht jedes strikte lokale Extremum auch mit diesem Kriterium gefunden werden! Ein Beispiel dazu:

Bemerkung 7.4.12

Die Funktion

$$f : (-1, 1) \rightarrow \mathbb{R}, x \mapsto x^4$$

hat ein striktes lokales (und sogar globales) Minimum bei $x_0 = 0$ (denn klarerweise gilt für alle $x \in (-1, 1) \setminus \{0\}$, dass $f(x) = x^4 > 0 = f(0)$). Dieses Minimum kann aber nicht mit dem obigen Kriterium gefunden werden, weil die Ungleichung aus Proposition 7.4.11 (i) ($f''(0) > 0$) hier nicht erfüllt ist. In der Tat gilt

$$f''(0) = \frac{d^2}{dx^2}x^4 \Big|_{x=0} = \frac{d}{dx}4x^3 \Big|_{x=0} = 12x^2 \Big|_{x=0} = 0.$$

Somit ist Proposition 7.4.11 hier nicht anwendbar.

Es ist sicherlich etwas unzufriedenstellend, dass das obige in der Schule erlernte Kriterium nicht *universell* ist, d.h. nicht auf jede Funktion angewendet werden kann. Jedoch kann das Kriterium in Proposition 7.4.11 noch verallgemeinert werden, sodass zumindest der Fall im vorigen Beispiel abgedeckt ist.

Proposition 7.4.13 — KRITERIUM GERADER ORDNUNG FÜR LOKALE EXTREMA

Es sei $k \in \mathbb{N}$, $f : (a, b) \rightarrow \mathbb{R}$ eine $2k$ -mal stetig differenzierbare Funktion und $x_0 \in (a, b)$.

- (i) Ist $f'(x_0) = \dots = f^{(2k-1)}(x_0) = 0$ und $f^{(2k)}(x_0) > 0$, so ist x_0 ein striktes lokales Minimum von f .

(ii) Ist $f'(x_0) = \dots = f^{(2k-1)}(x_0) = 0$ und $f^{(2k)}(x_0) < 0$, so ist x_0 ein striktes lokales Maximum von f .

Beweis. Es muss wiederum nur (i) gezeigt werden, denn (ii) folgt aus (i) durch Betrachtung von $-f$. Es seien f, x_0 wie in der Aussage. Mit Bemerkung 7.4.9 erhalten wir: Es gibt ein $\phi : (a, b) \rightarrow \mathbb{R}$ stetig mit $\phi(x_0) = 0$ so, dass für alle $x \in (a, b)$ gilt

$$f(x) = T_{2k}[f, x_0] + \phi(x)(x - x_0)^{2k} = \sum_{\ell=0}^{2k} \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell + \phi(x)(x - x_0)^{2k}.$$

Da $f'(x_0) = \dots = f^{(2k-1)}(x_0) = 0$ verschwinden die Summanden für $\ell = 1, \dots, 2k - 1$. Wir erhalten für $x \in (a, b)$

$$f(x) = f(x_0) + \frac{1}{(2k)!} f^{(2k)}(x_0)(x - x_0)^{2k} + \phi(x)(x - x_0)^{2k} = f(x_0) + \left[\frac{1}{(2k)!} f^{(2k)}(x_0) + \phi(x) \right] (x - x_0)^{2k}.$$

Weil ϕ stetig ist und $\phi(x_0) = 0$ folgt, dass $\lim_{x \rightarrow x_0} \phi(x) = 0$. Mit dem Epsilon-Delta-Kriterium (Satz 6.1.23), angewandt mit $\varepsilon = \frac{1}{2(2k)!} f^{(2k)}(x_0) > 0$ erhält man

$$\exists \delta > 0 : \quad \forall x \in (a, b) \text{ mit } |x - x_0| < \delta \quad \text{gilt} \quad |\phi(x) - 0| < \frac{1}{2(2k)!} f^{(2k)}(x_0).$$

Für $x \in (a, b) \cap B_\delta(x_0)$ (d.h. für $x \in (a, b)$ mit $|x - x_0| < \delta$) gilt dann also $|\phi(x)| < \frac{1}{2(2k)!} f^{(2k)}(x_0)$ und somit

$$\frac{1}{(2k)!} f^{(2k)}(x_0) + \phi(x) \geq \frac{1}{(2k)!} f^{(2k)}(x_0) - |\phi(x)| \geq \frac{1}{(2k)!} f^{(2k)}(x_0) - \frac{1}{2(2k)!} f^{(2k)}(x_0) = \frac{1}{2(2k)!} f^{(2k)}(x_0).$$

Man erhält daraus für alle $x \in (a, b) \cap B_\delta(x_0)$, dass

$$f(x) \geq f(x_0) + \frac{1}{2(2k)!} f^{(2k)}(x_0)(x - x_0)^{2k} = f(x_0) + \frac{1}{2(2k)!} f^{(2k)}(x_0)[(x - x_0)^k]^2. \quad (396)$$

Da $f^{(2k)}(x_0) > 0$ gilt $\frac{1}{2(2k)!} f^{(2k)}(x_0)[(x - x_0)^k]^2 > 0$ für alle $x \neq x_0$. Wir folgern aus (396), dass

$$f(x) > f(x_0) \quad \forall x \in (a, b) \cap B_\delta(x_0) \setminus \{x_0\}.$$

Das heißt nach Definition 7.2.1 nichts anderes als dass f ein striktes lokales Minimum bei x_0 annimmt.

Beispiel 7.4.14

Wir gehen nochmal zurück zu der Funktion aus Bemerkung 7.4.12, d.h. zu

$$f : (-1, 1) \rightarrow \mathbb{R}, x \mapsto x^4,$$

die klarerweise bei $x_0 = 0$ ein lokales Minimum hat. Dies kann nun mit Proposition 7.4.13

auch tatsächlich überprüft werden: Es gilt nämlich (Übungsaufgabe)

$$f'(0) = f''(0) = f'''(0) = 0 \quad \text{und} \quad f^{(4)}(0) = 24 > 0.$$

Somit ist die Bedingung aus Proposition 7.4.13 (i) für $k = 2$ überprüft und man sieht, dass f bei $x_0 = 0$ ein striktes lokales Minimum besitzt.

Unendlich oft differenzierbare Funktionen

Die Klasse der Funktionen mit der wir uns in diesem Abschnitt beschäftigen wollen ist die der *unendlich oft differenzierbaren* oder auch *glatten* Funktionen.

Definition 7.4.15 — GLATTE FUNKTIONEN

Es sei $D \subseteq \mathbb{K}_1$. Eine Funktion $f : D \rightarrow \mathbb{K}_2$ heißt *unendlich oft differenzierbar*, *beliebig oft differenzierbar* oder auch *glatt*, wenn f für jedes $k \in \mathbb{N}$ auf D k -mal differenzierbar ist.

Viele Funktionen, die wir bereits kennen, z.B. Polynome, $\exp$, $\sin$, $\cos$ sind glatt: Denn wir sehen an der Formel für die Ableitung (Korollar 7.0.11), dass wir die Ableitungsfunktionen *immer weiter differenzieren* können.

Wir wissen: Jede k -mal (stetig) differenzierbare Funktion lässt sich durch ein Polynom k -ten Grades approximieren. Der Fehlerterm, der dabei entsteht, kann aus der Taylorformel abgelesen werden, vgl. Satz 7.4.6. Wird dieser Fehlerterm für große Werte von k klein, so kann man das Grenzverhalten der Taylorpolynome $(T_k[f, x_0])_{k \in \mathbb{N}}$ für $k \rightarrow \infty$ studieren. Im Grenzfall ergibt sich dabei eine *Potenzreihe*. Man spricht von einer sogenannten *Taylorreihe*. Solche Taylorreihen bilden im Folgenden das Objekt unseres Studiums.

Proposition 7.4.16 — KONVERGENZ VON TAYLORREIHEN

Es sei $f : (a, b) \rightarrow \mathbb{R}$ eine glatte Funktion und $x_0 \in (a, b)$. Ferner sei $\delta > 0$ so, dass $B_\delta(x_0) \subseteq (a, b)$ und

$$\lim_{k \rightarrow \infty} \sup_{y \in B_\delta(x_0)} |f^{(k)}(y)| \frac{\delta^k}{k!} = 0. \quad (397)$$

Dann hat die Potenzreihe $\sum_{\ell \geq 0} \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell$ einen Konvergenzradius $R \geq \delta$ und es gilt

$$f(x) = \sum_{\ell=0}^{\infty} \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell \quad \forall x \in (x_0 - \delta, x_0 + \delta).$$

Beweis. Sei $x \in (x_0 - \delta, x_0 + \delta)$ und $k \in \mathbb{N}$. Die $(k - 1)$ -ste Partialsumme der zu untersuchenden Potenzreihe ist gegeben durch

$$\sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell = T_{k-1}[f, x_0](x).$$

Die Taylorformel (Satz 7.4.6) liefert, dass es für alle $k \in \mathbb{N}$ ein $\xi_k \in (\min\{x_0, x\}, \max\{x_0, x\})$ gibt mit

$$f(x) - T_{k-1}[f, x_0](x) = f(x) - \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell \stackrel{\text{Satz 7.4.6}}{=} \frac{f^{(k)}(\xi_k)}{k!} (x - x_0)^k.$$

Da nun $x \in (x_0 - \delta, x_0 + \delta)$ und ξ_k zwischen x und x_0 liegt gilt klarerweise auch $\xi_k \in (x_0 - \delta, x_0 + \delta)$ und daher gilt

$$|f(x) - T_{k-1}[f, x_0](x)| = \frac{|f^{(k)}(\xi_k)|}{k!} |x - x_0|^k \leq \sup_{y \in (x_0 - \delta, x_0 + \delta)} |f^{(k)}(y)| \frac{|x - x_0|^k}{k!}$$

Da $|x - x_0| < \delta$ gilt dann auch

$$0 \leq |f(x) - T_{k-1}[f, x_0](x)| \leq \sup_{y \in (x_0 - \delta, x_0 + \delta)} |f^{(k)}(y)| \frac{\delta^k}{k!} \xrightarrow{k \rightarrow \infty} 0. \quad (397)$$

Aus dem Sandwichlemma folgt, dass

$$\lim_{k \rightarrow \infty} |f(x) - T_{k-1}[f, x_0](x)| = 0.$$

Nach Proposition 2.3.30 heißt das auch $\lim_{k \rightarrow \infty} (f(x) - T_{k-1}[f, x_0](x)) = 0$. Die Grenzwertregeln liefern dann

$$f(x) = \lim_{k \rightarrow \infty} T_{k-1}[f, x_0](x) = \lim_{k \rightarrow \infty} \sum_{\ell=0}^{k-1} \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell \quad \forall x \in (x_0 - \delta, x_0 + \delta)$$

Somit ist $\sum_{\ell \geq 0} \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell$ auch für alle $x \in (x_0 - \delta, x_0 + \delta)$ konvergent. Sei nun $R \in [0, \infty]$ der Konvergenzradius dieser Potenzreihe und D_K der Konvergenzbereich (siehe Definition 5.2.1). Wir haben dann vorhin gesehen, dass

$$(x_0 - \delta, x_0 + \delta) \subseteq D_K \stackrel{\text{Prop 5.2.3}}{\subseteq} \{z \in \mathbb{C} : |z - x_0| \leq R\} \quad (398)$$

Dies impliziert $R \geq \delta$. (In der Tat: Wäre $R < \delta$ so wäre $x_0 + \frac{R+\delta}{2}$ ein Element von $(x_0 - \delta, x_0 + \delta)$ aber kein Element von $\{z \in \mathbb{C} : |z - x_0| \leq R\}$, was (398) widersprechen würde.)

Die Potenzreihe aus Proposition 7.4.16 nennen wir *Taylorreihe* von f .

Definition 7.4.17 — TAYLORREIHE

Es sei $f : (a, b) \rightarrow \mathbb{R}$ eine glatte Funktion. Dann heißt die Potenzreihe

$$T[f, x_0] := \sum_{\ell \geq 0} \frac{f^{(\ell)}(x_0)}{\ell!} (x - x_0)^\ell$$

die *Taylorreihe* von f mit Referenzpunkt x_0 .

Eine Frage drängt sich nun förmlich auf:

Frage. Unter welchen Bedingungen an eine Funktion $f : (a, b) \rightarrow \mathbb{R}$ stimmt sie auf (a, b) mit der Wertefunktion ihrer Taylorreihe überein?

Proposition 7.4.16 gibt darauf sogar schon eine partielle Antwort. Allerdings beinhaltet sie die unter Umständen schwer zu überprüfende Voraussetzung (397). Wegen dieser Voraussetzung gilt die Übereinstimmung von Funktion und Wertefunktion der Taylorreihe auch eventuell nicht auf dem gesamten Definitions-Intervall, sondern nur auf einem Teilbereich $B_\delta(x_0) \subseteq (a, b)$.

Ein wichtiger Fakt, den wir aber an dieser Stelle schonmal festhalten können ist, dass f und die Wertefunktion von $T[f, x_0]$ auf jeden Fall beim Referenzpunkt x_0 übereinstimmen – unabhängig von der genauen Gestalt von f !

Beispiel 7.4.18

Es sei $f : (a, b) \rightarrow \mathbb{R}$ eine glatte Funktion und $x_0 \in (a, b)$. Am Referenzpunkt x_0 stimmt f auf jeden Fall mit der Wertefunktion der Taylorreihe $T[f, x_0]$ überein, denn

$$\sum_{\ell=0}^{\infty} \frac{f^{(\ell)}(x_0)}{\ell!} (x_0 - x_0)^\ell = \frac{f^{(0)}(x_0)}{0!} = f(x_0).$$

Leider werden wir am Ende des nächsten Abschnittes (Beispiel 7.4.25) sehen, dass es passieren kann, dass f und die Wertefunktion von $T[f, x_0]$ *ausschließlich am Referenzpunkt* übereinstimmen (und sonst nicht). Man lernt daraus: Die Bedingung (397) muss unbedingt überprüft werden!

Um nun Funktionen zu studieren, bei denen $T[f, x_0]$ auf einem größeren Bereich konvergiert, sprechen wir über *reell-analytische* Funktionen.

Reell-Analytische Funktionen

Wir haben in der vorigen Sektion gelernt, dass sich eine gegebene glatte Funktion $f : (a, b) \rightarrow \mathbb{R}$ durch eine *Taylorreihe* unter gewissen Bedingungen ausdrücken lässt. Erfreulich daran ist, dass man so eine *Potenzreihendarstellung* von f finden kann. Die Klasse von Funktionen, die wir als Potenzreihe darstellen können, nennen wir *reell-analytisch*.

Definition 7.4.19 — REELL-ANALYTISCHE FUNKTION

Eine Funktion $f : (a, b) \rightarrow \mathbb{R}$ heißt *reell-analytisch* auf (a, b) , wenn es für ein $x_0 \in (a, b)$

eine Potenzreihe $\sum_{k \geq 0} c_k(x - x_0)^k$ gibt, die auf ganz (a, b) konvergiert und

$$f(x) = \sum_{k=0}^{\infty} c_k(x - x_0)^k \quad \forall x \in (a, b)$$

erfüllt.

Beispiel 7.4.20

Dass Polynome, $\exp$, $\sin$, $\cos$ auf $\mathbb{R}$ reell-analytisch sind ist bereits bekannt, vgl. z.B. Proposition 5.5.7 für $\sin$ und $\cos$. Sämtliche reell-analytische Funktionen sind glatt, da Potenzreihen unendlich oft differenzierbar sind.^a Wir werden aber später sehen, dass nicht alle glatten Funktionen auch automatisch reell-analytisch sind.

^aSchließlich sind ja alle Ableitungen von Potenzreihen wieder Potenzreihen und somit abermals differenzierbar, vgl. Proposition 7.0.10.

Beispiel 7.4.21

Behauptung. Auch $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \cos((x + 1)^2)$ ist reell-analytisch.

Beweis. Für $x \in \mathbb{R}$ hat man

$$\cos((x + 1)^2) \stackrel{\text{Prop 5.5.7}}{=} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} [(x + 1)^2]^{2k} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} (x + 1)^{4k} = \sum_{\ell=0}^{\infty} c_{\ell} (x - (-1))^{\ell},$$

wobei

$$c_{\ell} = \begin{cases} \frac{(-1)^k}{(2k)!} & \text{falls } \ell = 4k \text{ für ein } k \in \mathbb{N}_0 \\ 0 & \text{sonst} \end{cases}$$

Somit ist Definition 7.4.19 mit $x_0 = -1$ überprüft.

Man kann sich nun fragen, ob es mehrere Möglichkeiten gibt, dieselbe Funktion als Potenzreihe darzustellen. Mit verschiedenen Entwicklungspunkten x_0 ist das sicher möglich. Aber – wie wir sehen werden – gibt es mit festgelegtem Entwicklungspunkt nur eine einzige Option. Genauer werden wir in der nächsten Proposition zeigen: Die Taylorreihe $T[f, x_0]$ aus Definition 7.4.17 ist unsere *einzige Möglichkeit* für eine Potenzreihendarstellung mit Entwicklungspunkt x_0 . Sollte die Taylorreihe auf (a, b) nicht gegen die Funktion f konvergieren, gibt es also gar keine Potenzreihendarstellung von f .

Proposition 7.4.22 — WENN POTENZREIHE, DANN TAYLORREIHE

Es sei $f : (a, b) \rightarrow \mathbb{R}$ reell-analytisch und $x_0 \in (a, b)$ sowie $\sum_{k \geq 0} c_k(x - x_0)^k$ so gewählt, dass

$$f(x) = \sum_{k=0}^{\infty} c_k(x - x_0)^k \quad \text{für alle } x \in (a, b). \quad (399)$$

Dann gilt

$$c_k = \frac{f^{(k)}(x_0)}{k!} \quad \forall k \in \mathbb{N}_0. \quad (400)$$

Insbesondere stimmt die obige Potenzreihe mit der Taylorreihe $T[f, x_0]$ überein.

Beweisskizze. Es sei $f : (a, b) \rightarrow \mathbb{R}$ wie in der Aussage und $\ell \in \mathbb{N}_0$ beliebig vorgegeben. Wir zeigen, dass $c_\ell = \frac{f^{(\ell)}(x_0)}{\ell!}$. Falls $\ell = 0$ gilt klarerweise

$$c_0 = \sum_{k=0}^{\infty} c_k(x_0 - x_0)^k \stackrel{(399)}{=} f(x_0) = \frac{f^{(0)}(x_0)}{0!}.$$

Wir nehmen also im Folgenden $\ell \geq 1$ an. Wendet man die Differentiationsregel für Potenzreihen aus Proposition 7.0.10 ℓ -mal an, so hat man für $x \in (a, b)$

$$f^{(\ell)}(x) = \frac{d^\ell}{dx^\ell} \left(\sum_{k=0}^{\infty} c_k(x - x_0)^k \right) = \sum_{k=\ell}^{\infty} c_k[k(k-1) \cdots (\ell - \ell + 1)](x - x_0)^{k-\ell}.$$

Setzt man $x = x_0$ ein so ist in der Summe auf der rechten Seite jeder Term 0 mit Ausnahme des Terms für $k = \ell$. Man erhält

$$f^{(\ell)}(x_0) = c_\ell[\ell(\ell-1) \cdots (\ell - \ell + 1)](x - x_0)^{\ell-\ell} = c_\ell \ell!.$$

Teilen wir durch $\ell!$ so erhalten wir

$$c_\ell = \frac{f^{(\ell)}(x_0)}{\ell!} \quad \forall \ell \in \mathbb{N}.$$

Damit ist (400) gezeigt. Insbesondere ist auch

$$\sum_{k \geq 0} c_k(x - x_0)^k = \sum_{k \geq 0} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k = T[f, x_0].$$

Beispiel 7.4.23

Insbesondere können wir ohne Rechnung die Taylorreihen $T[\exp, 0]$, $T[\sin, 0]$ und $T[\cos, 0]$ angeben: Denn wenn eine Funktion bereits eine bekannte Potenzreihendarstellung hat, müssen Taylorreihe und diese Potenzreihe laut Proposition 7.4.22 miteinander übereinstimmen.

Wir folgern dann zum Beispiel aus Proposition 5.5.7

$$T[\sin, 0] = \sum_{k \geq 0} \frac{(-1)^k}{(2k+1)!} x^{2k+1}$$

$$T[\cos, 0] = \sum_{k \geq 0} \frac{(-1)^k}{(2k)!} x^{2k}$$

Wir lernen aus Proposition 7.4.22: Wer eine neue Potenzreihendarstellung für eine gegebene Funktion finden möchte, muss es mit der Taylorreihe versuchen. Deren Konvergenzverhalten kann mit Proposition 7.4.16 untersucht werden.

Beispiel 7.4.24

Es gilt für alle $\alpha \in \mathbb{R} \setminus \{0\}$ und $x \in (-\frac{1}{2}, \frac{1}{2})$, dass

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k.$$

Insbesondere gilt die folgende *Reihenentwicklung der Wurzel*

$$\sqrt{1+x} = \sum_{k=0}^{\infty} \binom{\frac{1}{2}}{k} x^k \quad \forall x \in (-\frac{1}{2}, \frac{1}{2}).$$

Beweis.

FALL 1. $\alpha \in \mathbb{N}$.

Falls $\alpha \in \mathbb{N}$ ist das Berechnen der Taylorreihe nicht notwendig: Die obige Formel folgt direkt aus dem Binomischen Lehrsatz (vgl. Satz 3.3.27). In der Tat erhält man dann (sogar für $x \in \mathbb{R}$)

$$\sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \stackrel{\text{Prop 3.3.26(v)}}{=} \sum_{k=0}^{\alpha} \binom{\alpha}{k} x^k = \sum_{k=0}^{\alpha} \binom{\alpha}{k} x^k 1^{\alpha-k} \stackrel{\text{Satz 3.3.27}}{=} (1+x)^\alpha.$$

FALL 2. $\alpha \notin \mathbb{N}$.

Es sei $f : (-1, 1) \rightarrow \mathbb{R}$, $x \mapsto (1+x)^\alpha$. Wir zeigen mit Proposition 7.4.16, dass f mit der Wertefunktion der Taylorreihe übereinstimmt. und $k \in \mathbb{N}$ berechnen wir

$$\begin{aligned} f'(x) &= \alpha(1+x)^{\alpha-1} \\ f''(x) &= \alpha(\alpha-1)(1+x)^{\alpha-2} \\ &\dots \\ f^{(k)}(x) &= \alpha(\alpha-1) \cdots (\alpha-k+1)(1+x)^{\alpha-k}. \end{aligned} \tag{401}$$

Es sei nun $\delta \in (0, \frac{1}{2})$ beliebig. Wir überprüfen dann (397). Dazu: Für alle $k \geq [\alpha]$ gilt $\alpha - k \leq 0$ und daher

$$\begin{aligned} \sup_{y \in (-\delta, \delta)} |f^{(k)}(y)| & \sup_{y \in (-\delta, \delta)} |\alpha(\alpha-1) \cdot \dots \cdot (\alpha-k+1)| |1+y|^{\alpha-k} \\ & \leq |\alpha(\alpha-1) \cdot \dots \cdot (\alpha-k+1)| (1-\delta)^{\alpha-k} \end{aligned}$$

Man erhält

$$\sup_{y \in B_\delta(x_0)} |f^{(k)}(y)| \frac{\delta^k}{k!} = \frac{1}{k!} |\alpha(\alpha-1) \cdot \dots \cdot (\alpha-k+1)| (1-\delta)^{\alpha-k} \delta^k = (1-\delta)^\alpha \left| \binom{\alpha}{k} \right| \frac{\delta^k}{(1-\delta)^k}. \quad (402)$$

Um (397) zu überprüfen müssen wir nun noch zeigen, dass obiger Ausdruck für $k \rightarrow \infty$ gegen 0 strebt. Dazu zeigen wir

$$\text{ZWISCHENBEHAUPTUNG.} \quad \lim_{k \rightarrow \infty} \left| \binom{\alpha}{k} \right| \frac{\delta^k}{(1-\delta)^k} = 0 \quad (403)$$

Eine Möglichkeit, das zu beweisen ist dadurch, dass $\sum_{k \geq 1} \left| \binom{\alpha}{k} \right| \frac{\delta^k}{(1-\delta)^k}$ konvergiert, vgl Satz 5.1.2 (i). Dies kann leicht mit dem Quotientenkriterium verifiziert werden (vgl. Satz 5.1.22). Dazu:

$$\begin{aligned} \frac{\left| \binom{\alpha}{k+1} \right| \frac{\delta^{k+1}}{(1-\delta)^{k+1}}}{\left| \binom{\alpha}{k} \right| \frac{\delta^k}{(1-\delta)^k}} &= \frac{\delta}{1-\delta} \frac{\left| \binom{\alpha}{k+1} \right|}{\left| \binom{\alpha}{k} \right|} = \frac{\delta}{1-\delta} \frac{\frac{|\alpha(\alpha-1) \cdot \dots \cdot (\alpha-(k+1)+1)|}{(k+1)!}}{\frac{|\alpha(\alpha-1) \cdot \dots \cdot (\alpha-k+1)|}{k!}} \\ &= \frac{\delta}{1-\delta} \left| \frac{\alpha(\alpha-1) \cdot \dots \cdot (\alpha-k)}{\alpha(\alpha-1) \cdot \dots \cdot (\alpha-k+1)} \right| \frac{k!}{(k+1)!} = \frac{\delta}{1-\delta} \left| \frac{\alpha-k}{k+1} \right| = \frac{\delta}{1-\delta} \left| \frac{\frac{\alpha}{k} - 1}{1 + \frac{1}{k}} \right|. \end{aligned}$$

Es folgt

$$\lim_{k \rightarrow \infty} \left| \frac{\left| \binom{\alpha}{k+1} \right| \frac{\delta^{k+1}}{(1-\delta)^{k+1}}}{\left| \binom{\alpha}{k} \right| \frac{\delta^k}{(1-\delta)^k}} \right| = \frac{\delta}{1-\delta} | -1 | = \frac{\delta}{1-\delta} = \frac{1 - (1-\delta)}{1-\delta} = \frac{1}{1-\delta} - 1.$$

Da $\delta < \frac{1}{2}$ gilt $\frac{1}{1-\delta} < 2$ und somit $\frac{1}{1-\delta} - 1 < 1$. Die Reihe $\sum_{k \geq 1} \left| \binom{\alpha}{k} \right| \frac{\delta^k}{(1-\delta)^k}$ konvergiert also nach dem Quotientenkriterium und (403) folgt dann aus Satz 5.1.2 (i).

ZWISCHENBEHAUPTUNG ENDE.

Aus der Zwischenbehauptung und (402) lernen wir: Die Voraussetzungen von Proposition 7.4.16 sind erfüllt. Das heißt

$$f(x) = \sum_{\ell=0}^{\infty} \frac{f^{(\ell)}(0)}{\ell!} (x-0)^\ell \quad \forall x \in (-\delta, \delta).$$

Nun berechnen wir aus (401)

$$\frac{f^{(\ell)}(0)}{\ell!} \stackrel{(401)}{=} \frac{\alpha(\alpha-1) \cdots (\alpha-\ell+1)}{\ell!} = \binom{\alpha}{\ell} \quad \forall \ell \in \mathbb{N}.$$

Also gilt mit den beiden vorigen Gleichungen

$$(1+x)^\alpha = f(x) = \sum_{\ell=0}^{\infty} \binom{\alpha}{\ell} x^\ell \quad \forall x \in (-\delta, \delta).$$

Da $\delta \in (0, \frac{1}{2})$ beliebig war folgt auch

$$(1+x)^\alpha = \sum_{\ell=0}^{\infty} \binom{\alpha}{\ell} x^\ell \quad \forall x \in (-\tfrac{1}{2}, \tfrac{1}{2}).$$

In der Vorlesung *Funktionentheorie* werden Sie lernen dass diese Gleichheit sogar auf einer viel größeren Menge als $(-\frac{1}{2}, \frac{1}{2})$ gilt.

Eine essentielle Voraussetzung von Proposition 7.4.16 ist Gleichung (397). Ohne diese Voraussetzung muss die Taylorreihe von f nichteinmal konvergieren und — selbst wenn sie konvergiert — nicht mit f übereinstimmen.

Wir sehen im Folgenden was schief gehen kann, wenn Gleichung (397) nicht erfüllt ist.

Beispiel 7.4.25

Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$f(x) := \begin{cases} e^{-\frac{1}{x^2}} & x \neq 0 \\ 0 & x = 0 \end{cases} \quad (404)$$

ist glatt auf $\mathbb{R}$ und erfüllt

$$T[f, 0] = 0.$$

Insbesondere stimmt f einzig und allein auf dem Referenzpunkt $x_0 = 0$ mit der Wertefunktion der Taylorreihe überein.^a Also ist f glatt aber nicht reell-analytisch.

Beweis.

SCHRITT 1. Glattheit von f .

Dass f auf $\mathbb{R} \setminus \{0\}$ glatt ist, sieht man leicht. Die Glattheit bei 0 ist jedoch etwas komplizierter. Man zeigt leicht mit Induktion: Für alle $n \in \mathbb{N}_0$ gibt es ein Polynom $p_n : \mathbb{R} \rightarrow \mathbb{R}$ so, dass für alle $x \in \mathbb{R} \setminus \{0\}$ gilt

$$f^{(n)}(x) = p_n\left(\frac{1}{x}\right)e^{-\frac{1}{x^2}}. \quad (405)$$

Darauf basierend zeigen wir mit Induktion

Zwischenbehauptung. $\forall n \in \mathbb{N}_0 : f$ ist bei 0 n -mal diff'bar und $f^{(n)}(0) = 0$.

Induktionsanfang $n = 0$.

$$f^{(0)}(0) = f(0) \underset{(404)}{=} 0.$$

Induktionsschritt $n \longrightarrow n + 1$.

Es gelte für ein $n \in \mathbb{N}$ die Induktionshypothese $f^{(n)}(0) = 0$. Damit f bei $x_0 = 0$ auch $(n + 1)$ -mal differenzierbar ist muss der Grenzwert

$$f^{(n+1)}(0) = \lim_{\substack{x \rightarrow 0 \\ x \in \mathbb{R} \setminus \{0\}}} \frac{f^{(n)}(x) - f^{(n)}(0)}{x - 0} \underset{\text{Induktionshypothese}}{=} \lim_{\substack{x \rightarrow 0 \\ x \in \mathbb{R} \setminus \{0\}}} \frac{f^{(n)}(x)}{x}$$

existieren. Nutzen wir nun (405) so ergibt sich

$$\lim_{\substack{x \rightarrow 0 \\ x \in \mathbb{R} \setminus \{0\}}} \frac{f^{(n)}(x)}{x} = \lim_{\substack{x \rightarrow 0 \\ x \in \mathbb{R} \setminus \{0\}}} \frac{\frac{1}{x} p_n\left(\frac{1}{x}\right) e^{-\frac{1}{x^2}}}{x} = \lim_{\substack{x \rightarrow 0 \\ x \in \mathbb{R} \setminus \{0\}}} \frac{f^{(n)}(x)}{x} = \lim_{\substack{x \rightarrow 0 \\ x \in \mathbb{R} \setminus \{0\}}} \tilde{p}_n\left(\frac{1}{x}\right) e^{-\frac{1}{x^2}},$$

wobei $\tilde{p}_n : \mathbb{R} \rightarrow \mathbb{R}, \tilde{p}_n(y) := y p_n(y)$ auch ein Polynom ist. Aus Beispiel 6.3.1 wissen wir, dass für alle Potenzen $\alpha \in \mathbb{R}$

$$\lim_{x \rightarrow 0} x^\alpha e^{-\frac{1}{x^2}} = 0.$$

Da $\tilde{p}_n\left(\frac{1}{x}\right)$ eine Summe ist, die in jedem Summanden eine Potenz von x und einen konstanten Faktor beinhaltet, gilt

$$\lim_{\substack{x \rightarrow 0 \\ x \in \mathbb{R} \setminus \{0\}}} \tilde{p}_n\left(\frac{1}{x}\right) e^{-\frac{1}{x^2}} = 0.$$

Alles in Allem folgern wir, dass f bei $x_0 = 0$ $(n + 1)$ -mal differenzierbar ist und

$$f^{(n+1)}(0) = 0.$$

Zwischenbehauptung Ende.

SCHRITT 2. Taylorreihe von f .

Mit der Zwischenbehauptung folgt

$$T[f, 0] = \sum_{\ell \geq 0} \frac{f^{(\ell)}(0)}{\ell!} x^\ell = 0.$$

Wir müssen noch klären warum f unter diesem Umständen nicht reell-analytisch ist: Wäre f reell-analytisch, so würde f nach Proposition 7.4.22 mit der Wertefunktion seiner Taylorreihe übereinstimmen. Wir haben hier aber nachgewiesen, dass die Wertefunktion der Taylorreihe (die konstante Nullfunktion) nicht mit f übereinstimmt.

^adenn für $x \neq 0$ gilt $f(x) = \exp(-\frac{1}{x^2}) > 0$.

Wir haben also bewiesen, dass es auf $\mathbb{R}$ eine glatte Funktion gibt, die nicht reell-analytisch ist. Spannenderweise gibt es so etwas auf $\mathbb{C}$ nicht – mehr dazu lernen Sie in der Vorlesung *Funktionentheorie*.

29.04.2026

7.5 Konvexität

Im Folgenden sei stets $I \subseteq \mathbb{R}$ ein Intervall. In diesem Abschnitt wollen wir verstehen, was es bedeutet, wenn sich eine Funktion $f : I \rightarrow \mathbb{R}$ *konvex* bzw. *konkav* verhält. Anschaulich heißt das, dass der Funktionsgraph eine *Linkskurve* bzw. *Rechtskurve* macht. Unsere Definition fußt auf folgender:

Beobachtung. Macht ein Graph eine Linkskurve, so liegen alle Sekanten über dem Graphen, vgl. Abbildung 36.

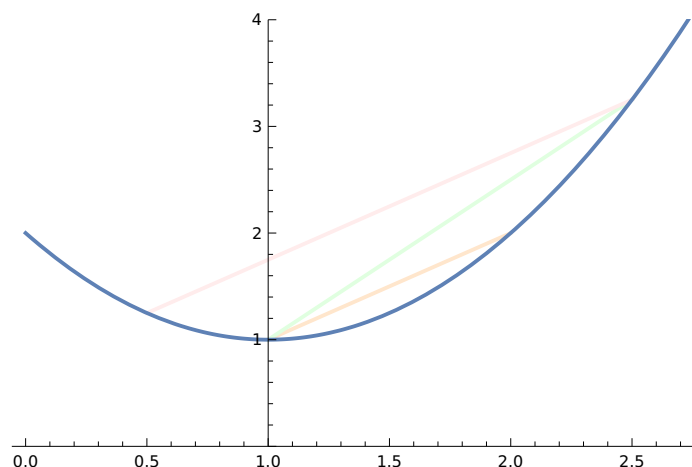


Abbildung 36: Ein Graph einer konvexen Funktion (blau) und einige Sekanten

Aus (320) erinnern wir uns an die Sekantengleichung von f zwischen zwei Punkten $x \in I$ und $x_0 \in I$, $x \neq x_0$.

$$\ell_{x,x_0}(t) = f(x_0) + \frac{f(x) - f(x_0)}{x - x_0}(t - x_0) \quad t \in [\min\{x, x_0\}, \max\{x, x_0\}].$$

Konvexität sollte also bedeuten, dass

$$\forall t \in [\min\{x, x_0\}, \max\{x, x_0\}] : \quad f(t) \leq \ell_{x,x_0}(t)$$

Diese Gleichung formen wir nun etwas um. Für $t \in [\min\{x, x_0\}, \max\{x, x_0\}]$ gilt

$$f(t) \leq f(x_0) + (f(x) - f(x_0)) \frac{t-x_0}{x-x_0} = \frac{t-x_0}{x-x_0} f(x) + \left(1 - \frac{t-x_0}{x-x_0}\right) f(x_0). \quad (406)$$

Man kann obige Bedingung auch einfacher schreiben, indem man sich des Konzeptes der *Linearkombinationen* bedient, welches Sie bereits aus der Linearen Algebra kennen: Liegt t zwischen x und x_0 so gibt es ein $\lambda \in [0, 1]$ so, dass

$$t = \lambda x + (1 - \lambda)x_0.$$

Dann gilt auch

$$\frac{t-x_0}{x-x_0} = \frac{\lambda x + (1-\lambda)x_0 - x_0}{x-x_0} = \frac{\lambda(x-x_0)}{x-x_0} = \lambda.$$

Setzen wir die obigen beiden Gleichungen für t in (406) ein so folgt

$$f(\lambda x + (1 - \lambda)x_0) \leq \lambda f(x) + (1 - \lambda)f(x_0).$$

Ist $\lambda \in [0, 1]$ beliebig, so erreicht die Linearkombination $\lambda x + (1 - \lambda)x_0$ alle Werte zwischen x und x_0 . Dies motiviert die nachfolgende Definition der *Konvexität*.

Definition 7.5.1 — KONVEXITÄT / KONKAVITÄT

Es sei $I \subseteq \mathbb{R}$ ein Intervall. Eine Funktion $f : I \rightarrow \mathbb{R}$ heißt

(i) *konvex*^a auf I falls für alle $x_1, x_2 \in I$ und $\lambda \in [0, 1]$ gilt, dass

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

(ii) *konkav* auf I falls für alle $x_1, x_2 \in I$ und $\lambda \in [0, 1]$ gilt, dass

$$f(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

^aWie diese Definition anschaulich funktioniert kann man auch nochmal auf GEOGEBRA nachschauen.

Klarerweise ist f konkav genau dann wenn $-f$ konvex ist und umgekehrt.

Bemerkung 7.5.2

Die *affin-lineare Abbildung* $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \alpha x + \beta$ (für gewisse $\alpha, \beta \in \mathbb{R}$) ist sowohl konvex als auch konkav. In der Tat gilt für $x_1, x_2 \in \mathbb{R}$ und $\lambda \in [0, 1]$

$$\begin{aligned} f(\lambda x_1 + (1 - \lambda)x_2) &= \alpha(\lambda x_1 + (1 - \lambda)x_2) + \beta \\ &= \lambda \alpha x_1 + (1 - \lambda)\alpha x_2 + \beta \\ &= \lambda \alpha x_1 + (1 - \lambda)\alpha x_2 + (\lambda + 1 - \lambda)\beta \end{aligned}$$

$$\begin{aligned}
&= \lambda \alpha x_1 + (1 - \lambda) \alpha x_2 + \lambda \beta + (1 - \lambda) \beta \\
&= \lambda(\alpha x_1 + \beta) + (1 - \lambda)(\alpha x_2 + \beta) = \lambda f(x_1) + (1 - \lambda) f(x_2).
\end{aligned}$$

Wegen dieser Gleichheit sind sowohl Punkt (i) als auch Punkt (ii) aus Definition 7.5.1 erfüllt. Ergo ist f sowohl konvex als auch konkav.

In der Schule wurden die Begriffe konvex / konkav stets mit der zweiten Ableitung in Verbindung gebracht. Nach Definition 7.5.1 können aber auch Funktionen konvex sein, die gar keine zweite Ableitung haben, weil sie z.B. nicht differenzierbar sind. Den Zusammenhang mit der zweiten Ableitung wollen wir aber später noch klären.

Beispiel 7.5.3

Behauptung. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto |x|$ ist konvex.

Beweis. Wir gehen streng nach Definition 7.5.1 vor. Es seien dazu $x_1, x_2 \in \mathbb{R}$ und $\lambda \in [0, 1]$. Dann gilt

$$\begin{aligned}
f(\lambda x_1 + (1 - \lambda)x_2) &= |\lambda x_1 + (1 - \lambda)x_2| \\
&\stackrel{\Delta\text{-Ungl.}}{\leq} |\lambda x_1| + |(1 - \lambda)x_2| \\
&\stackrel{\text{Satz 2.0.3}}{=} |\lambda| |x_1| + |1 - \lambda| |x_2| \\
&\stackrel{\text{Satz 2.0.3}}{=} \lambda |x_1| + (1 - \lambda) |x_2| = \lambda f(x_1) + (1 - \lambda) f(x_2).
\end{aligned}$$

Differenzierbare konvexe Funktionen

Als nächstes kombinieren wir die Begriffe „Differenzierbarkeit“ und „Konvexität“. Bei einer differenzierbaren Funktion konnten wir stets eine *Tangente* an dem Graphen betrachten. Wir werden sehen, dass eine konvexe differenzierbare Funktion stets *über ihrer Tangenten liegt*, siehe auch Abbildung (37).

Die nachfolgenden Resultate formulieren wir nur für ein offenes Intervall (a, b) , wobei $a, b \in \mathbb{R} \cup \{\pm\infty\}$ mit $a < b$. Verallgemeinerungen auf andere Intervalltypen können aber auch ohne große Probleme gefunden werden. Desweiteren werden wir nur konvexe Funktionen $f : (a, b) \rightarrow \mathbb{R}$ besprechen. Man findet aber leicht analoge Formulierungen der Resultate für konkave Funktionen, indem man einfach $-f$ betrachtet.

Proposition 7.5.4

Es sei $f : (a, b) \rightarrow \mathbb{R}$ differenzierbar. Dann sind folgende Aussagen äquivalent

- (a) f ist auf (a, b) konvex

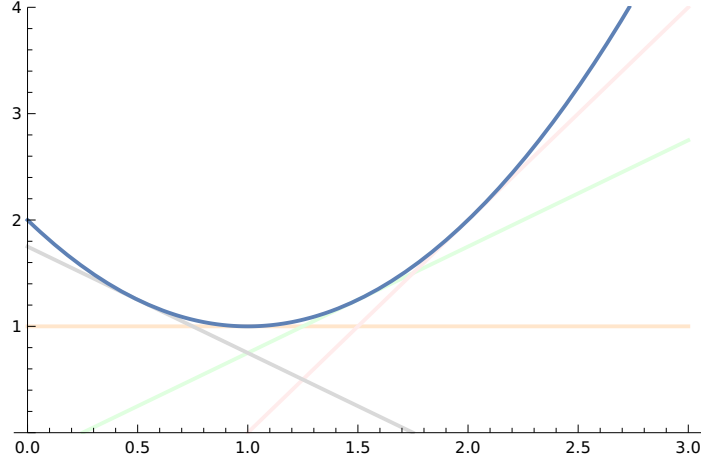


Abbildung 37: Dieselbe (konvexe) Funktion wie in Abbildung 36 zusammen mit einigen Tangenten.

(b) Für alle $x, x_0 \in (a, b)$ gilt

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0).$$

Beweis. Zuerst zu „(a) $\implies$ (b)“. Es gelte (a). Ferner seien $x, x_0 \in (a, b)$.

$$\mathbb{Z}: \quad f(x) \geq f(x_0) + f'(x_0)(x - x_0).$$

Nach der Kettenregel gilt für alle $t \in [0, 1]$

$$\frac{d}{dt}f(tx + (1 - t)x_0) = f'(tx + (1 - t)x_0)(x - x_0)$$

Liest man dies rückwärts und setzt $t = 0$ ein so folgt

$$f'(x_0)(x - x_0) = \frac{d}{dt}f(tx + (1 - t)x_0)\Big|_{t=0}. \quad (407)$$

Mit der Definition der Ableitung gilt

$$\begin{aligned} \frac{d}{dt}f(tx + (1 - t)x_0)\Big|_{t=0} & \stackrel{\text{Def 7.0.1}}{=} \lim_{t \rightarrow 0} \frac{f(tx + (1 - t)x_0) - f(0 \cdot x + (1 - 0)x_0)}{t - 0} \\ & = \lim_{t \rightarrow 0} \frac{1}{t} (f(tx + (1 - t)x_0) - f(x_0)) \end{aligned} \quad (408)$$

Wähle nun eine Folge $(t_n)_{n \in \mathbb{N}}$ mit $t_n \in (0, 1]$ für alle $n \in \mathbb{N}$ und $\lim_{n \rightarrow \infty} t_n = 0$. Dann gilt

$$f'(x_0)(x - x_0) \stackrel{(407)}{=} \frac{d}{dt}f(tx + (1 - t)x_0)\Big|_{t=0} \stackrel{(408)}{=} \lim_{n \rightarrow \infty} \frac{1}{t_n} (f(t_n x + (1 - t_n)x_0) - f(x_0))$$

$$\stackrel{(a)}{\underset{\text{Def 7.5.1}}{\leq}} \lim_{n \rightarrow \infty} \frac{1}{t_n} (t_n f(x) + (1 - t_n) f(x_0) - f(x_0)) = \lim_{n \rightarrow \infty} \frac{t_n (f(x) - f(x_0))}{t_n} = f(x) - f(x_0).$$

Addiert man $f(x_0)$ auf beiden Seiten so findet man

$$f(x_0) + f'(x_0)(x - x_0) \leq f(x),$$

was zu zeigen war.

Nun noch zu „(b) $\implies$ (a)“. Es gelte (b). Es seien $x_1, x_2 \in (a, b)$ und $\lambda \in [0, 1]$.

$$\mathbb{Z}: \quad f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

Dazu: Verwenden wir (b) mit $x = x_1$ und $x_0 = \lambda x_1 + (1 - \lambda)x_2$ so erhalten wir

$$\begin{aligned} f(x_1) &\geq f(\lambda x_1 + (1 - \lambda)x_2) + f'(\lambda x_1 + (1 - \lambda)x_2)[x_1 - (\lambda x_1 + (1 - \lambda)x_2)] \\ &= f(\lambda x_1 + (1 - \lambda)x_2) + f'(\lambda x_1 + (1 - \lambda)x_2)(1 - \lambda)(x_1 - x_2). \end{aligned}$$

Verwenden wir außerdem (b) mit $x = x_2$ und $x_0 = \lambda x_1 + (1 - \lambda)x_2$ so erhalten wir

$$\begin{aligned} f(x_2) &\geq f(\lambda x_1 + (1 - \lambda)x_2) + f'(\lambda x_1 + (1 - \lambda)x_2)[x_2 - (\lambda x_1 + (1 - \lambda)x_2)] \\ &= f(\lambda x_1 + (1 - \lambda)x_2) + f'(\lambda x_1 + (1 - \lambda)x_2)\lambda(x_2 - x_1). \end{aligned}$$

Man erhält mit den vorigen beiden Abschätzungen

$$\begin{aligned} \lambda f(x_1) + (1 - \lambda)f(x_2) &\geq \lambda[f(\lambda x_1 + (1 - \lambda)x_2) + f'(\lambda x_1 + (1 - \lambda)x_2)(1 - \lambda)(x_1 - x_2)] \\ &\quad + (1 - \lambda)[f(\lambda x_1 + (1 - \lambda)x_2) + f'(\lambda x_1 + (1 - \lambda)x_2)\lambda(x_2 - x_1)] \\ &= f(\lambda x_1 + (1 - \lambda)x_2)[\lambda + (1 - \lambda)] \\ &\quad + f'(\lambda x_1 + (1 - \lambda)x_2)[\lambda(1 - \lambda)(x_1 - x_2) + (1 - \lambda)\lambda(x_2 - x_1)] \\ &= f(\lambda x_1 + (1 - \lambda)x_2). \end{aligned}$$

Die obige Erkenntnis erlaubt uns, den Zusammenhang zwischen Konvexität und dem Verhalten der ersten Ableitung zu studieren.

Korollar 7.5.5

Es sei $f : (a, b) \rightarrow \mathbb{R}$ differenzierbar. Dann sind folgende Aussagen äquivalent

- (a) f ist auf (a, b) konvex
- (b) Für alle $x, x_0 \in (a, b)$ gilt $(f'(x) - f'(x_0))(x - x_0) \geq 0$

(c) f' ist auf (a, b) monoton wachsend.

Beweis. Wir werden die Aussage beweisen, indem wir „ $(a) \implies (b)$ “, „ $(b) \implies (c)$ “ und „ $(c) \implies (a)$ “ zeigen. Aus diesen drei Folgerungen folgt die Äquivalenz aller drei Aussagen über die Transitivität der Folgerung (vgl. Proposition 1.1.7). In der Tat: Die Aussage „ $(b) \implies (a)$ “ folgt zum Beispiel direkt aus „ $[(b) \implies (c)] \wedge [(c) \implies (a)]$ “. Man nennt diese Beweistechnik einen *Ringschluss*.

Zu „ $(a) \implies (b)$ “. Sei f auf (a, b) konvex und seien $x, x_0 \in (a, b)$. Aus Proposition 7.5.4 folgern wir

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0). \quad (409)$$

Wenden wir Proposition 7.5.4 (b) noch einmal mit vertauschten Rollen von x und x_0 an so ergibt sich

$$f(x_0) \geq f(x) + f'(x)(x_0 - x).$$

Addieren wir die vorigen Gleichungen (409) und so erhalten wir

$$\begin{aligned} f(x) + f(x_0) &\geq [f(x_0) + f'(x_0)(x - x_0)] + [f(x) + f'(x)(x_0 - x)] \\ &= f(x) + f(x_0) + f'(x_0)(x - x_0) + f'(x)(x_0 - x) \end{aligned}$$

Subtrahieren wir $f(x) + f(x_0)$ von beiden Seiten, so ergibt sich

$$0 \geq f'(x_0)(x - x_0) + f'(x)(x_0 - x) = f'(x_0)(x - x_0) - f'(x)(x - x_0) = (f'(x_0) - f'(x))(x - x_0).$$

Multiplizieren wir auf beiden Seiten mit (-1) (wodurch sich das Ungleichheitszeichen ändert, vgl. Proposition 3.2.20), so erhalten wir

$$0 \leq -(f'(x_0) - f'(x))(x - x_0) = (f'(x) - f'(x_0))(x - x_0)$$

Zu „ $(b) \implies (c)$ “. Es gelte (b). Seien nun $x_1, x_2 \in (a, b)$ mit $x_1 \leq x_2$.

$$\text{Z:} \quad f'(x_1) \leq f'(x_2)$$

OBdA können wir annehmen, dass $x_1 < x_2$ denn falls $x_1 = x_2$ ist die obige Ungleichung klarerweise erfüllt. Es gilt nun

$$f'(x_2) - f'(x_1) = \underbrace{\frac{1}{x_2 - x_1}}_{>0} \underbrace{(f'(x_2) - f'(x_1))(x_2 - x_1)}_{\geq 0, \text{ wegen (b)}} \geq 0.$$

Addieren wir $f'(x_1)$ auf beiden Seiten so folgt $f'(x_2) \geq f'(x_1)$.

Zu „(c) $\implies$ (a)“. Es sei f' monoton wachsend. Wir überprüfen die Konvexität indem wir Proposition 7.5.4 (b) nachweisen. Es seien dazu $x, x_0 \in (a, b)$.

$$\mathbb{Z}\mathbb{Z}: \quad f(x) \geq f(x_0) + f'(x_0)(x - x_0)$$

OBdA gilt $x \neq x_0$, denn im Fall $x = x_0$ ist die obige Gleichung klarerweise erfüllt. Da f dann auf $[\min\{x, x_0\}, \max\{x, x_0\}] \subseteq (a, b)$ stetig und auf $(\min\{x, x_0\}, \max\{x, x_0\})$ differenzierbar ist, findet man nach dem ersten Mittelwertsatz (Satz 7.3.1) oder auch alternativ nach der Taylorformel mit $k = 1$ (Satz 7.4.6) ein $\xi \in (\min\{x, x_0\}, \max\{x, x_0\})$ so, dass

$$f(x) - f(x_0) = f'(\xi)(x - x_0)$$

Nun formen wir um

$$f(x) - f(x_0) = f'(\xi)(x - x_0) = f'(x_0)(x - x_0) + (f'(\xi) - f'(x_0))(x - x_0) \quad (410)$$

Fall 1. $x > x_0$.

In dem Fall ist $\xi \in (x_0, x)$. Wegen der Monotonie von f' gilt dann $f'(\xi) \geq f'(x_0)$. somit haben wir

$$\underbrace{(f'(\xi) - f'(x_0))}_{\geq 0} \underbrace{(x - x_0)}_{\geq 0} \geq 0.$$

Nutzen wir diese Abschätzung in (410) so ergibt sich

$$f(x) - f(x_0) \geq f'(x_0)(x - x_0),$$

was zu zeigen war.

Fall 2. $x < x_0$.

In diesem Fall ist $\xi \in (x, x_0)$. Wegen der Monotonie von f' gilt dann $f'(\xi) \leq f'(x_0)$. somit haben wir

$$\underbrace{(f'(\xi) - f'(x_0))}_{\leq 0} \underbrace{(x - x_0)}_{\leq 0} \geq 0.$$

Nutzen wir diese Abschätzung wiederum in (410) so ergibt sich

$$f(x) - f(x_0) \geq f'(x_0)(x - x_0),$$

was zu zeigen war.

Monotonie der ersten Ableitung bedeutet Positivität der zweiten Ableitung, vgl. Bemerkung 7.3.13. Deshalb können wir endlich den Zusammenhang zwischen Konvexität und der zweiten Ableitung herausstellen.

Korollar 7.5.6 — KONVEXITÄT/KONKAVITÄT UND ZWEITE ABLEITUNG

Es sei $f : (a, b) \rightarrow \mathbb{R}$ zweimal differenzierbar. Dann gilt

(i)

$$f \text{ ist auf } (a, b) \text{ konvex} \iff f''(x) \geq 0 \quad \forall x \in (a, b).$$

(ii)

$$f \text{ ist auf } (a, b) \text{ konkav} \iff f''(x) \leq 0 \quad \forall x \in (a, b).$$

Beweis. Zu (i). Es gilt

f ist auf (a, b) konvex.

$$\iff \text{Korollar 7.5.5}$$

f' ist auf (a, b) monoton wachsend.

$$\iff \text{Bemerkung 7.3.13}$$

$(f')'(x) \geq 0$ für alle $x \in (a, b)$.

$$\iff \text{Definition 7.4.1}$$

$f''(x) \geq 0$ für alle $x \in (a, b)$.

Zu (ii).

f ist auf (a, b) konkav.

$$\iff$$

$-f$ ist auf (a, b) konvex.

$$\iff \text{(i)}$$

$(-f)''(x) \geq 0$ für alle $x \in (a, b)$.

$$\iff (-f)'' = -f''$$

$$-f''(x) \geq 0 \text{ für alle } x \in (a, b).$$

$$\Longleftrightarrow$$

$$f''(x) \leq 0 \text{ für alle } x \in (a, b).$$

Beispiel 7.5.1. Es sei $n \in \mathbb{N}$ beliebig.

Behauptung. $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^{2n}$ konvex.

Beweis. Wir nutzen Korollar 7.5.6. Dazu: für alle $x \in \mathbb{R}$ gilt

$$f''(x) = (2n)(2n-1)x^{2n-2} = 2n(2n-1)(x^{n-1})^2 \geq 0.$$

Mit Korollar 7.5.6 folgt die Konvexität sofort.

Zusammenfassung – Kurvendiskussion

In den letzten Abschnitten haben wir viel darüber gelernt, wie wir anhand bestimmter Eigenschaften einer gegebenen Funktion $f : D \rightarrow \mathbb{R}$ Rückschlüsse auf das Aussehen des Graphen Γ_f ziehen können. Dieses Unterfangen nenne wir *Kurvendiskussion*. Sie beinhaltet für uns die Beantwortung der nachfolgenden Fragen

- **Definitions -und Wertebereich.**

- Toolbox (für den Wertebereich): Zwischenwertsatz (Satz 6.2.15, Satz 6.2.17)

- **Monotonieverhalten.** Auf welchen Teilbereichen von D ist f monoton wachsend/fallend?

- Toolbox: Monotoniesatz (Proposition 7.3.11).

- **globale Extremwerte.** Was für *globale Extremwerte* nimmt f an, d.h. was sind

$$\sup_{x \in D} f(x) := \sup\{f(x) : x \in D\} \quad \text{bzw.} \quad \inf_{x \in D} f(x) := \inf\{f(x) : x \in D\}$$

und werden diese Werte auch wirklich angenommen?

- Toolbox: Satz von Weierstraß (Satz 6.2.12) und die Methoden für lokale Extremwerte (s.u.).

- **lokale Extremwerte.** Was für lokale Extrema werden angenommen?

- Toolbox: Notwendiges Kriterium für innere lokale Extremstellen (Satz 7.2.6), Vorzeichenwechselkriterium (Proposition 7.3.15), Kriterien höherer Ordnung (Proposition 7.4.11 und Proposition 7.4.13)
- **Konvexe/Konkave Bereiche.** Gibt es Bereiche, in denen die Funktion konvex oder konkav ist?
 - Toolbox: Konvexität und zweite Ableitung (Proposition 7.5.6).

8 Integralrechnung

In der gymnasialen Oberstufe haben Sie neben der Differentialrechnung auch die *Integralrechnung* besprochen. Dieser widmen wir uns nun. Ziel der Integralrechnung ist es für einer *hinreichend guten* (z.B. *stetigen*) Funktion $f : [a, b] \rightarrow \mathbb{R}$ eine Zahl namens

$$\int_a^b f(x) \, dx \in \mathbb{R}$$

zuzuweisen, sodass Folgendes gilt

- Falls $f(x) \geq 0$ für alle $x \in [a, b]$ so ist $\int_a^b f(x) \, dx$ der *Flächeninhalt* der Fläche unter dem Graphen Γ_f . Wir schreiben

$$\int_a^b f(x) \, dx = \text{Area}(\{(x, y) \in \mathbb{R}^2 : x \in [a, b] \wedge 0 \leq y \leq f(x)\}).$$

- Falls f auch negative Werte annimmt so soll $\int_a^b f(x) \, dx$ der Flächeninhalt *mit Vorzeichen* sein, d.h. der Flächeninhalt des Bereiches oberhalb der x -Achse geht mit positivem Vorzeichen ein und der Flächeninhalt des Bereiches unterhalb der x -Achse geht mit negativem Vorzeichen ein, siehe Abbildung 38.

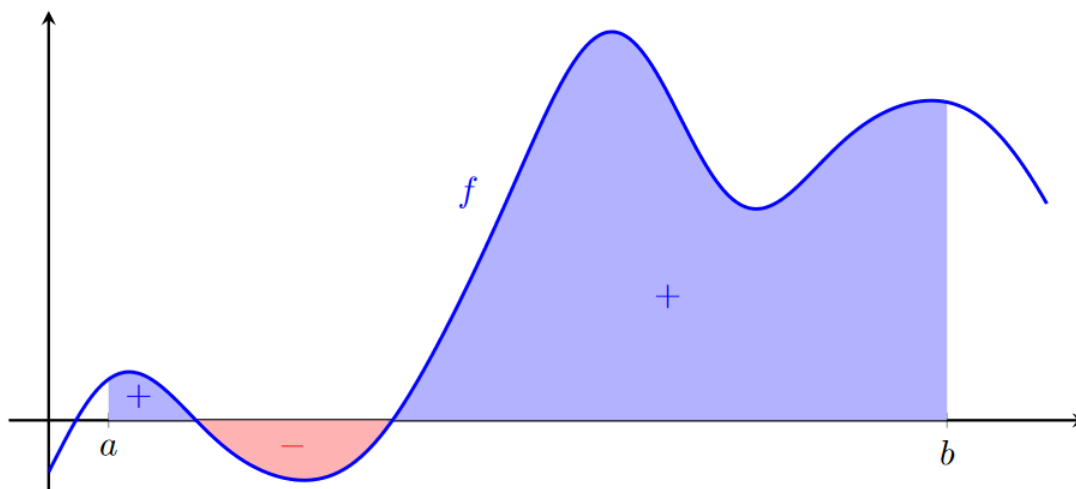


Abbildung 38: Für das Integral $\int_a^b f(x) \, dx$ berechnet man den Flächeninhalt der blauen und des roten Bereiches. Man subtrahiert dann die rote Fläche von der blauen Fläche.

Die Flächen die wir berechnen müssen sind für allgemeine Funktionen sehr *gekrümmt*. Das macht. Daher müssen wir die zu berechnenden Flächen approximieren mit *Referenzflächen* deren Flächeninhalt wir kennen. In unserem Fall werden diese Referenzflächen stets *Rechtecke*

der Form $[a, b] \times [c, d]$ sein. Der Flächeninhalt eines solchen Rechtecks ist gegeben durch das *Produkt der Seitenlängen*. Konkreter gilt für $\alpha_1, \beta_1, \alpha_2, \beta_2 \in \mathbb{R}, (\beta_1 > \alpha_1, \beta_2 > \alpha_2)$,

$$\text{Area}([\alpha_1, \beta_1] \times [\alpha_2, \beta_2]) = (\beta_1 - \alpha_1) \cdot (\beta_2 - \alpha_2).$$

Der Funktionsgraph aus Abbildung 38 kann wie in Abbildung 39 mit Rechtecken approximiert werden. Summiert man über die Flächeninhalte der dabei auftretenden Rechtecke mit den entsprechenden Vorzeichen, so erhält man eine approximation an das Integral.

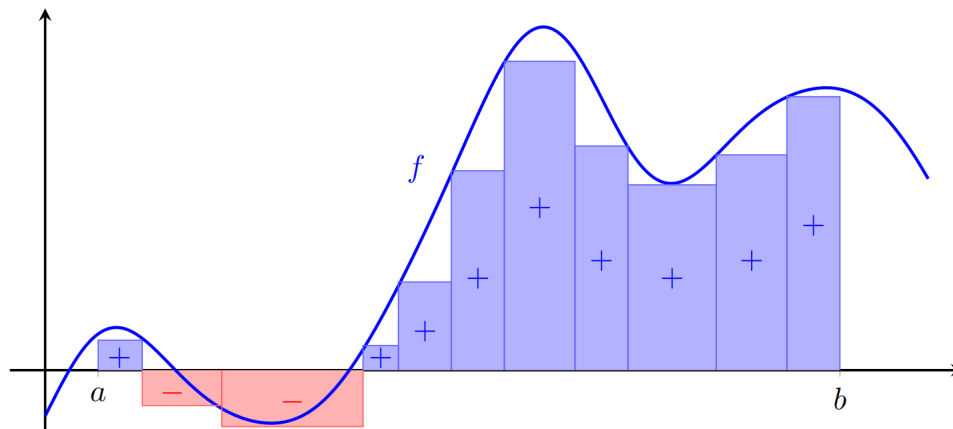


Abbildung 39: Rechtecke können benutzt werden, um die zu berechnende Fläche anzunähern

Beachte: Sämtliche in Abbildung 39 genutzten Rechtecke liegen *ganz unterhalb* des Funktionsgraphen. Das bedeutet: Die Summe von der wir oben gesprochen haben *unterschätzt* das Integral. Man spricht auch von einer *Untersumme*. Nennen wir nun die auftretenden Rechtecke $R_i, i = 1, \dots, m$. Sehen wir uns Abbildung 39 genauer an so stellen wir auch folgendes fest:

- Das Intervall $[a, b] \times \{0\}$ wird durch die Rechtecke komplett ausgepflastert, weil der x -Endpunkt eines jeden Rechtecks stets mit dem x -Anfangspunkt des nachfolgenden Rechtecks übereinstimmt. (Eine Ausnahme bildet hier natürlich das „letzte Rechteck“ was kein nachfolgendes Rechteck besitzt).
- Den x -Endpunkt des i -ten Rechtecks bezeichnen wir stets mit $x_i, i = 1, \dots, m$. Insbesondere gilt $x_m = b$. Wir setzen außerdem $x_0 := a$. Dadurch erhalten wir eine sogenannte *Zerlegung* $Z = (x_0, x_1, \dots, x_m)$ des Intervalles $[a, b]$, d.h. eine angeordnete Liste von Punkten mit

$$x_0 = a, \quad x_m = b, \quad x_0 < x_1 < x_2 < \dots < x_m.$$

- Außerdem berühren alle Rechtecke den Funktionsgraphen Γ_f an dem kleinsten Wert, den f im Intervall $[x_{i-1}, x_i]$ annimmt. Somit ist stets einer der y -Endpunkte des Rechtecks auf der durch

$$\inf_{[x_{i-1}, x_i]} f := \inf f([x_{i-1}, x_i])$$

gegeben (und der andere durch 0).

- Die Rechtecke aus Abbildung 2 lassen sich dann wie folgt aufschreiben: Für $i = 1, \dots, m$ ist das i -te Rechteck gegeben durch

$$R_i := \begin{cases} [x_{i-1}, x_i] \times [0, \inf_{[x_{i-1}, x_i]} f] & \text{falls } \inf_{[x_{i-1}, x_i]} f \geq 0 \text{ (blauer Bereich)} \\ [x_{i-1}, x_i] \times [\inf_{[x_{i-1}, x_i]} f, 0] & \text{falls } \inf_{[x_{i-1}, x_i]} f < 0 \text{ (roter Bereich)} \end{cases}$$

Berechnet man die Flächeninhalte der Rechtecke so erhält man

$$\text{Area}(R_i) := \begin{cases} \left(\inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) & \text{falls } \inf_{[x_{i-1}, x_i]} f \geq 0 \text{ (blauer Bereich)} \\ - \left(\inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) & \text{falls } \inf_{[x_{i-1}, x_i]} f < 0 \text{ (roter Bereich)} \end{cases} \quad (411)$$

Um die oben besprochene Approximation zu bilden müssen wir die blauen Rechtecke summieren und die roten Rechtecke abziehen. Das Minus-Vorzeichen vor den roten Rechtecken kürzt sich mit dem Minuszeichen im zweiten Fall von (411) weg. Wir erhalten dadurch eine *Formel für die obige Untersumme*

$$\sum_{i=1}^m \left(\inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1})$$

Analog kann man auch eine *Obersumme* bilden, die das Integral überschätzt. Wir werden diese Summen nutzen, um das Integral zu approximieren. Leider wird es so sein, dass dieses Approximations-Vorgehen nicht für alle Funktionen erfolgreich ist. Diejenigen Funktionen, bei denen es aber klappt, nennen wir *Riemann-integrierbar*. 04.05.2026

8.1 Definition des Riemann-Integrals

Hier definieren nun rigoros, was wir unter dem *Riemann-Integral* einer Funktion $f : [a, b] \rightarrow \mathbb{R}$ verstehen und für welche Funktionen wir es überhaupt bilden können. Zunächst zum Begriff einer *Zerlegung*.

Definition 8.1.1 — ZERLEGUNG

Es seien $a, b \in \mathbb{R}$, $a < b$. Eine *geordnete Liste von Punkten* $Z = (x_0, \dots, x_m)$, $m \in \mathbb{N}$, heißt *Zerlegung des Intervalles* $[a, b]$, falls

$$x_0 = a, \quad x_m = b, \quad x_0 < x_1 < x_2 < \dots < x_m. \quad a$$

Die Punkte $x_0, x_1, \dots, x_m$ nennen wir *Unterteilungsstellen*.

^aInsbesondere gilt auch $x_i \in [a, b]$ für alle $i = 0, \dots, m$.

Beispiel 8.1.2 — ÄQUIDISTANTE ZERLEGUNG

Es sei $[a, b]$ ein gegebenes Intervall (d.h. $a, b \in \mathbb{R}, a < b$) und $m \in \mathbb{N}$. Definiere für

$k = 0, \dots, m$

$$x_k := a + \frac{k}{m}(b - a).$$

Dann ist $Z = (x_0, \dots, x_m)$ eine Zerlegung von $[a, b]$. Wir nennen sie die *äquidistante Zerlegung* von $[a, b]$ mit $m + 1$ Unterteilungsstellen^a. Wir bezeichnen sie im Folgenden mit $Z_m^{(E)}$. Die Zerlegung $Z_T := Z_1^{(E)}$ besteht nur aus den Punkten a und b und heißt auch *triviale Zerlegung* von $[a, b]$.

^aDer Begriff *äquidistant* kommt daher, dass die Distanz zwischen zwei aufeinanderfolgenden Unterteilungsstellen immer dieselbe ist. In der Tat berechnet man für alle $k \in \{1, \dots, m\}$, dass

$$x_k - x_{k-1} = \left(a + \frac{k}{m}(b - a)\right) - \left(a + \frac{k-1}{m}(b - a)\right) = \left(\frac{k}{m} - \frac{k-1}{m}\right)(b - a) = \frac{b - a}{m},$$

was nicht von k abhängig ist.

Oft wird es nützlich sein, eine Zerlegung zu *verfeinern*, d.h. mehr Zerlegungspunkte in eine Zerlegung hineinzunehmen.

Im folgenden sei stets $a, b \in \mathbb{R}$ mit $a < b$, sodass $[a, b] \subseteq \mathbb{R}$ stets ein abgeschlossenes beschränktes Intervall ist.

Definition 8.1.3 — VERFEINERUNG

Es sei $Z = (x_0, \dots, x_m)$ eine Zerlegung von $[a, b]$. Eine Zerlegung $Z' = (y_0, \dots, y_{m'})$ von $[a, b]$ heißt *Verfeinerung* von Z falls

$$\{x_0, \dots, x_m\} \subseteq \{y_0, \dots, y_{m'}\}.$$

Sprich: Jede Unterteilungsstelle von Z ist auch eine Unterteilungsstelle von Z' .

Eine Möglichkeit, eine Zerlegung zu verfeinern, ist es sich mit einer weiteren Zerlegung *zusammenzufügen*.

Definition 8.1.4 — ZUSAMMENFÜGUNG

Es seien $Z_1 = (x_0, \dots, x_{m_1})$ und $Z_2 = (y_0, \dots, y_{m_2})$ Zerlegungen von $[a, b]$. Die *zusammengefügte Zerlegung* $Z_1 \vee Z_2$ ist eine Zerlegung die wie folgt entsteht

- (i) Bilde die Vereinigung $A := \{x_0, \dots, x_{m_1}\} \cup \{y_0, \dots, y_{m_2}\}$
- (ii) Ordne die Elemente aus A nach der Größe, etwa sei $A = \{z_0, \dots, z_r\}$ mit $r \in \mathbb{N}$ und $z_0 < \dots < z_r$.
- (iii) Dann gilt $Z_1 \vee Z_2 := (z_0, \dots, z_r)$.

Bemerkung 8.1.5

Sind Z_1, Z_2 zwei Zerlegungen von $[a, b]$, so ist $Z_1 \vee Z_2$ eine Verfeinerung sowohl von Z_1 als

auch von Z_2 . Davon überzeugt man sich leicht, wir überlassen Details den Lesenden als Übungsaufgabe.

Haben wir eine Zerlegung $Z = (x_0, \dots, x_m)$ des Intervalles $[a, b]$ gegeben, so können wir bezüglich dieser Zerlegung die *Obersumme* und *Untersumme* definieren.

Definition 8.1.6 — OBER -UND UNTERSUMMEN

Seien $a, b \in \mathbb{R}, a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion. Ferner sei $Z = (x_0, \dots, x_m)$ eine Zerlegung von $[a, b]$. Dann definieren wir

(i) die *Obersumme*^a von f bezüglich Z durch

$$\overline{S}_Z(f) := \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1})$$

(ii) die *Untersumme* von f bezüglich Z durch

$$\underline{S}_Z(f) := \sum_{i=1}^m \left(\inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1})$$

^aIch verweise für ein anschauliches Bild auf GEOGEBRA

Bemerkung 8.1.7

Wir haben f in der obigen Definition als beschränkt vorausgesetzt, damit $f([x_{i-1}, x_i])$ stets eine beschränkte Menge ist. Wegen der Ordnungsvollständigkeit von $\mathbb{R}$ (vgl. Definition 3.4.1) ist dann auch

$$\sup_{[x_{i-1}, x_i]} f := \sup f([x_{i-1}, x_i])$$

stets wohldefiniert (also stets eine reelle Zahl). Ohne die Voraussetzung der Beschränktheit kann davon im Allgemeinen nicht ausgegangen werden!

Bemerkung 8.1.8

Man sieht leicht, dass für alle Zerlegungen Z von $[a, b]$ gilt

- $\underline{S}_Z(-f) = -\overline{S}_Z(f)$ und
- $\overline{S}_Z(-f) = -\underline{S}_Z(f)$.

Wir beschäftigen uns nun mit einigen Eigenschaften von Ober -und Untersummen.

Proposition 8.1.9

Es sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt.

- (i) Für jede Zerlegung Z von $[a, b]$ gilt $\underline{S}_Z(f) \leq \overline{S}_Z(f)$.

(ii) Ist Z' eine Verfeinerung von Z , so gilt $\underline{S}_Z(f) \leq \underline{S}_{Z'}(f)$.

(iii) Ist Z' eine Verfeinerung von Z , so gilt $\overline{S}_Z(f) \geq \overline{S}_{Z'}(f)$.

Beweis. Zu (i). Es sei $Z = (x_0, \dots, x_m)$. Dann gilt für alle $i \in \{1, \dots, m\}$

$$\inf_{[x_{i-1}, x_i]} f \leq \sup_{[x_{i-1}, x_i]} f.$$

In der Tat: Dies kann man aus den Ungleichungen $\inf_{[x_{i-1}, x_i]} f \leq f(x_i)$ und $\sup_{[x_{i-1}, x_i]} f \geq f(x_i)$.

Diese sind gültig, weil das Infimum stets eine untere Schranke und das Supremum stets eine obere Schranke ist, vgl Definition 3.2.10.

Zu (ii). Es sei $Z = (x_0, \dots, x_m)$ und $Z' = (y_0, \dots, y_r)$ eine Verfeinerung, d.h. $\{x_0, \dots, x_m\} \subseteq \{y_0, \dots, y_r\}$, vgl Definition 8.1.3 . Das heißt

$$\text{Für alle } i \in \{0, \dots, m\} \text{ gibt es eine Zahl } \ell_i \in \{0, \dots, r\} \text{ so, dass } x_i = y_{\ell_i}. \quad (412)$$

Da beide Zerlegungslisten nach der Größe sortiert sind gilt auch $\ell_0 < \ell_1 < \ell_2 < \dots < \ell_m$. Da außerdem $x_0 = a = y_0$ gilt $\ell_0 = 0$ und da $x_m = b = y_r$ gilt $\ell_m = r$. Man erhält

$$\begin{aligned} x_0 &= y_0 < y_1 < \dots < y_{\ell_1} = x_1, \\ x_1 &= y_{\ell_1} < y_{\ell_1+1} < \dots < y_{\ell_2} = x_2, \\ x_2 &= y_{\ell_2} < y_{\ell_2+1} < \dots < y_{\ell_3} = x_3, \\ &\dots \\ x_{m-1} &= y_{\ell_{m-1}} < y_{\ell_{m-1}+1} < \dots < y_{\ell_m} = x_m. \end{aligned} \quad (413)$$

Nun gilt nach Definition 8.1.6 und da $0 < \ell_1 < \dots < \ell_m = r$

$$\underline{S}_{Z'}(f) = \sum_{k=1}^r \left(\inf_{[y_{k-1}, y_k]} f \right) (y_k - y_{k-1}) = \sum_{i=1}^m \sum_{k=\ell_{i-1}+1}^{\ell_i} \left(\inf_{[y_{k-1}, y_k]} f \right) (y_k - y_{k-1}).$$

Beachte: Für $k \in \{\ell_{i-1} + 1, \ell_i\}$ gilt stets, dass $[y_{k-1}, y_k] \subseteq [x_{i-1}, x_i]$, vgl (413). Daher gilt auch $f([y_{k-1}, y_k]) \subseteq f([x_{i-1}, x_i])$ und somit auch

$$\inf_{[y_{k-1}, y_k]} f \geq \inf_{[x_{i-1}, x_i]} f.$$

Nutzen wir diese Abschätzung in (413) so erhalten wir

$$\underline{S}_{Z'}(f) = \sum_{i=1}^m \sum_{k=\ell_{i-1}+1}^{\ell_i} \left(\inf_{[y_{k-1}, y_k]} f \right) (y_k - y_{k-1}) \geq \sum_{i=1}^m \sum_{k=\ell_{i-1}+1}^{\ell_i} \left(\inf_{[x_{i-1}, x_i]} f \right) (y_k - y_{k-1})$$

$$\begin{aligned}
& \xlongequal{\text{Prop 3.3.17(ii) in innerer Summe}} \sum_{i=1}^m \left(\inf_{[x_{i-1}, x_i]} f \right) \sum_{k=\ell_{i-1}+1}^{\ell_i} (y_k - y_{k-1}) \\
& \xlongequal{\text{Teleskopsumme}} \sum_{i=1}^m \left(\inf_{[x_{i-1}, x_i]} f \right) (y_{\ell_i} - y_{\ell_{i-1}}) \\
& \xlongequal{\text{Definition der } \ell_i, (412)} \sum_{i=1}^m \left(\inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) = \underline{S}_Z(f).
\end{aligned}$$

Es folgt $\underline{S}_Z(f) \leq \underline{S}_{Z'}(f)$, was zu zeigen war.

Zu (iii).

$$\overline{S}_Z(f) \xlongequal{\text{Bem 8.1.8}} -\underline{S}_Z(-f) \underset{(ii)}{\geq} -\underline{S}_{Z'}(-f) \xlongequal{\text{Bem 8.1.8}} \overline{S}_{Z'}(f).$$

Nutzen wir die Punkte (ii) und (iii) in Proposition 8.1.9 geschickt, so können wir eine Verbesserung der Aussage (i) erreichen, mit der sogar verschiedene Zerlegungen miteinander verglichen werden können.

Korollar 8.1.10

Es sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt und Z_1, Z_2 zwei Zerlegungen von $[a, b]$. Dann gilt

$$\underline{S}_{Z_1}(f) \leq \overline{S}_{Z_2}(f)$$

Beweis. Der Trick ist, mit der Zusammenfügung $Z_1 \vee Z_2$ (vgl Definition 8.1.4) zu arbeiten. Es gilt

$$\underline{S}_{Z_1}(f) \underset{\text{Prop 8.1.9(ii)}}{\leq} \underline{S}_{Z_1 \vee Z_2}(f) \underset{\text{Prop 8.1.9(i)}}{\leq} \overline{S}_{Z_1 \vee Z_2}(f) \underset{\text{Prop 8.1.9(iii)}}{\leq} \overline{S}_{Z_2}(f).$$

Bemerkung 8.1.11

Es sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt und Z_T wie in Bemerkung 8.1.2. As dem vorigen Korollar folgt auch, dass

$$\overline{S}_Z(f) \geq \underline{S}_{Z_T}(f) = \left(\inf_{[a,b]} f \right) (b - a) \quad \text{für alle Zerlegungen } Z \text{ von } [a, b]. \quad (414)$$

und

$$\underline{S}_Z(f) \leq \overline{S}_{Z_T}(f) = \left(\sup_{[a,b]} f \right) (b - a) \quad \text{für alle Zerlegungen } Z \text{ von } [a, b].$$

Die Menge aller Obersummen $\{\overline{S}_Z(f) : Z \text{ Zerlegung von } [a, b]\} \subseteq \mathbb{R}$ ist also nach unten beschränkt. Gleichsam ist die Menge aller Untersummen $\{\underline{S}_Z(f) : Z \text{ Zerlegung von } [a, b]\} \subseteq$

$\mathbb{R}$ nach oben beschränkt.

Diese Beschränktheit erlaubt die Definition des sogenannten *Oberintegrals* und *Unterintegrals*. Das Oberintegral ist der *kleinstmögliche Wert, der durch Obersummen approximiert werden kann* und das Unterintegral ist der *größtmögliche Wert, der durch Untersummen approximiert werden kann*.

Definition 8.1.12 — OBERINTEGRAL UND UNTERINTEGRAL

Es seien $a, b \in \mathbb{R}$, $a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ beschränkt. Dann definieren wir

(i) das *Oberintegral* von f auf $[a, b]$ durch

$$\int_a^b f(x) \, dx := \inf \{ \overline{S}_Z(f) : Z \text{ Zerlegung von } [a, b] \} \quad (415)$$

(ii) das *Unterintegral* von f auf $[a, b]$ durch

$$\int_a^b f(x) \, dx := \sup \{ \underline{S}_Z(f) : Z \text{ Zerlegung von } [a, b] \}$$

Das Oberintegral (bzw. Unterintegral) verstehen wir als bestmögliche Approximation von oben (bzw. von unten) an das Integral. Wir werden später fordern, dass diese Approximationen bei einer integrierbaren Funktion übereinstimmen sollen.

Bemerkung 8.1.13

Aufgrund von (414) ist die Menge über die in (415) das Infimum gebildet wird stets nach unten beschränkt. Daher existiert $\int_a^b f(x) \, dx$ stets (als reelle Zahl, vgl. Bemerkung 3.4.5!). Genauso kann man argumentieren, dass das Unterintegral stets eine wohldefinierte reelle Zahl ist.

Eine direkte Folgerung aus Korollar 8.1.10 ist eine Ungleichung zwischen Ober- und Unterintegral.

Proposition 8.1.14 — OBERINTEGRAL IST GRÖßER ALS UNTERINTEGRAL

Es sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt. Dann gilt

$$\int_a^b f(x) \, dx \leq \int_a^b f(x) \, dx.$$

Beweis. Es sei $n \in \mathbb{N}$. Nach Definition 8.1.12 (i) und Proposition 3.2.26 (mit $\varepsilon = \frac{1}{n}$) gibt es eine Zerlegung Z_n von $[a, b]$ mit

$$\overline{S}_{Z_n}(f) \leq \int_a^b f(x) \, dx + \frac{1}{n}.$$

Nach Definition 8.1.12 (ii) und Proposition 3.2.26 (mit $\varepsilon = \frac{1}{n}$) gibt es eine Zerlegung Z'_n von $[a, b]$ mit

$$\underline{S}_{Z'_n}(f) \geq \int_a^b f(x) \, dx - \frac{1}{n}.$$

Nach Korollar 8.1.10 gilt nun auch $\underline{S}_{Z'_n}(f) \leq \overline{S}_{Z_n}(f)$ für alle $n \in \mathbb{N}$. Wir folgern

$$\int_a^b f(x) \, dx - \frac{1}{n} \leq \underline{S}_{Z'_n}(f) \leq \overline{S}_{Z_n}(f) \leq \int_a^b f(x) \, dx + \frac{1}{n} \quad \forall n \in \mathbb{N}$$

Addiert man $\frac{1}{n}$ auf beiden Seiten, so findet man

$$\int_a^b f(x) \, dx \leq \int_a^b f(x) \, dx + \frac{2}{n} \quad \forall n \in \mathbb{N}.$$

Nun können wir beide Seiten als Folgen betrachten, die von einem Index n abhängen. Tun wir dies, so können wir auch zum Grenzwert $n \rightarrow \infty$ übergehen. Die Grenzwertvergleichsregeln liefern

$$\lim_{n \rightarrow \infty} \int_a^b f(x) \, dx \leq \lim_{n \rightarrow \infty} \left(\int_a^b f(x) \, dx + \frac{2}{n} \right).$$

Ergo gilt

$$\int_a^b f(x) \, dx \leq \int_a^b f(x) \, dx.$$

Wenn wir von einem *Riemann-Integral* sprechen, dann wollen wir, dass sowohl Obersummen als auch Untersummen den gleichen Wert approximieren können. Sprich: Oberintegral und Unterintegral müssen übereinstimmen. Den gemeinsamen Wert nennen wir *Riemann-Integral*.

Definition 8.1.15 — RIEMANN-INTEGRIERBARKEIT

Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion. Falls gilt, dass

$$\int_a^b f(x) \, dx = \int_a^b f(x) \, dx, \tag{416}$$

so nennen wir f auf $[a, b]$ *Riemann-integrierbar*. Den gemeinsamen Wert aus (416) nennen wir dann das *Riemann-Integral* von f und bezeichnen ihn mit

$$\int_a^b f(x) \, dx.$$

Die Funktion f nennen wir hierbei stets *Integrand*.

Nun haben wir das Riemann-Integral vollständig definiert. Ein Problem ist, dass die Definition an eine Bedingung geknüpft ist, nämlich an die *Riemann-Integrierbarkeit*, vgl. (416). Es wäre

mit unserem jetzigen Kenntnisstand durchaus noch möglich, dass keine einzige Funktion überhaupt Riemann-integrierbar ist. Daher müssen wir den Begriff näher studieren.

Bemerkung 8.1.16

Die Bedingung (416) kann abgeschwächt werden, nämlich zu

$$\int_a^b f(x) \, dx \geq \overline{\int_a^b f(x) \, dx}.$$

Die andere Ungleichung „ $\leq$ “ ist nämlich dank Proposition 8.1.14 automatisch erfüllt.

Riemann-Integrierbarkeit

Nicht alle Funktionen sind Riemann-Integrierbar. In diesem Abschnitt werden wir aber verstehen, welche Funktionen es sind und welche nicht.

Beispiel 8.1.17

Behauptung. Konstante Funktionen sind Riemann-integrierbar.

Beweis. Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine konstante Funktion, etwas $f(x) = c$ für alle $x \in [a, b]$. Dann gilt für jede beliebige Zerlegung $Z = (x_0, \dots, x_m)$ von $[a, b]$, dass

$$\overline{S}_Z(f) \stackrel{\text{Def 8.1.6}}{=} \sum_{i=1}^m \underbrace{\left(\sup_{[x_{i-1}, x_i]} f \right)}_{=c} (x_i - x_{i-1}) = c \sum_{i=1}^m (x_i - x_{i-1}) \stackrel{\text{Telekopsumme}}{=} c(x_m - x_0) = c(b - a).$$

Nun nutzen wir (415), um das Oberintegral zu bestimmen. Da jede Obersumme den gleichen Wert $c(b - a)$ hat, ist die Menge über die in (415) das Infimum gebildet wird einelementig und man erhält

$$\overline{\int_a^b} f(x) \, dx = c(b - a). \quad (417)$$

Analog zeigt man (Übungsaufgabe), dass für jede beliebige Zerlegung Z von $[a, b]$ gilt, dass

$$\underline{S}_Z(f) = c(b - a).$$

Wir schließen, dass auch

$$\underline{\int_a^b} f(x) \, dx = c(b - a). \quad (418)$$

Aus (417) und (418) folgt, dass Oberintegral und Unterintegral übereinstimmen und den

Wert $c(b - a)$ annehmen. Aus Definition 8.1.15 folgt, dass f Riemann integrierbar ist und

$$\int_a^b f(x) \, dx = c(b - a).$$

Wenn man einsetzt, dass $f(x) = c$ für alle $x \in [a, b]$ kann man diese Gleichung nochmal umschreiben, nämlich als

$$\int_a^b c \, dx = c(b - a).$$

Wir hoffen natürlich, dass mehr als nur konstante Funktionen Riemann-integrierbar sind. Um die Klasse der Riemann-integrierbaren Funktionen weiter zu studieren leiten wir nun ein hilfreiches *Integrierbarkeits-Kriterium* her.

Proposition 8.1.18 — INTEGRIERBARKEITS-KRITERIUM

Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion. Dann sind folgende Aussagen äquivalent:

- (a) f ist Riemann-integrierbar
- (b) Zu jedem $\varepsilon > 0$ gibt es eine Zerlegung Z von $[a, b]$ so, dass gilt $\overline{S}_Z(f) - \underline{S}_Z(f) \leq \varepsilon$.

Die Größe $\overline{S}_Z(f) - \underline{S}_Z(f)$ wird auch manchmal *Schwankungssumme* von f bezüglich der Zerlegung Z genannt.

Beweis. Zu „(b) $\implies$ (a)“. Es sei $f : [a, b] \rightarrow \mathbb{R}$ wie in der Aussage und es gelte (b).

$$\underline{Z}: \quad f \text{ ist Riemann-integrierbar, d.h. (nach Bemerkung 8.1.16) } \underline{\int_a^b} f(x) \, dx \geq \overline{\int_a^b} f(x) \, dx.$$

Es sei dazu $n \in \mathbb{N}$. Verwendet man (b) für $\varepsilon = \frac{1}{n}$ so findet man eine Zerlegung Z_n von $[a, b]$ so, dass

$$\overline{S}_{Z_n}(f) - \underline{S}_{Z_n}(f) \leq \frac{1}{n}. \quad (419)$$

Nun berechnen wir für beliebiges $n \in \mathbb{N}$

$$\begin{aligned} \overline{\int_a^b} f(x) \, dx &\stackrel{\text{Def 8.1.12}}{=} \inf \{ \overline{S}_Z(f) : Z \text{ Zerlegung von } [a, b] \} \leq \overline{S}_{Z_n}(f) \\ &\stackrel{(419)}{\leq} \underline{S}_{Z_n}(f) + \frac{1}{n} \leq \sup \{ \underline{S}_Z(f) : Z \text{ Zerlegung von } [a, b] \} + \frac{1}{n} = \underline{\int_a^b} f(x) \, dx + \frac{1}{n}. \end{aligned}$$

Wir haben also gesehen

$$\overline{\int_a^b} f(x) \, dx \leq \underline{\int_a^b} f(x) \, dx + \frac{1}{n} \quad \forall n \in \mathbb{N}.$$

Wiederum fassen wir beide Seiten dieser Ungleichung als Folgen auf. Da obige Ungleichung für alle Folgenglieder gilt, gilt sie auch für die Grenzwerte. Wir gehen zum Grenzwert über und erhalten

$$\overline{\int_a^b f(x) \, dx} \leq \underline{\int_a^b f(x) \, dx},$$

was zu zeigen war.

Nun zu „(a) $\implies$ (b)“. Es sei f Riemann-integrierbar. Wir wollen nun (b) zeigen. Hierzu sei $\varepsilon > 0$ beliebig aber fest.

Z_Z: Es gibt eine Zerlegung Z von $[a, b]$ mit $\overline{S}_Z(f) - \underline{S}_Z(f) \leq \varepsilon$.

Wegen Definition 8.1.12 und Proposition 3.2.26 gibt es zunächst eine Zerlegung Z_1 von $[a, b]$ so, dass

$$\overline{S}_{Z_1}(f) \leq \overline{\int_a^b f(x) \, dx} + \frac{\varepsilon}{2} \stackrel[\text{Def 8.1.15}]{f \text{ nach Vorr. Riemann-integrierbar}} \int_a^b f(x) \, dx + \frac{\varepsilon}{2}.$$

Wiederum mit Definition 8.1.12 und Proposition 3.2.26 gibt es eine Zerlegung Z_2 von $[a, b]$ so dass

$$\underline{S}_{Z_2}(f) \geq \underline{\int_a^b f(x) \, dx} - \frac{\varepsilon}{2} \stackrel[\text{Def 8.1.15}]{f \text{ nach Vorr. Riemann-integrierbar}} \int_a^b f(x) \, dx - \frac{\varepsilon}{2}.$$

Mit den beiden vorigen Gleichungen gilt

$$\overline{S}_{Z_1}(f) - \underline{S}_{Z_2}(f) \leq \left(\int_a^b f(x) \, dx + \frac{\varepsilon}{2} \right) - \left(\int_a^b f(x) \, dx - \frac{\varepsilon}{2} \right) = \varepsilon. \quad (420)$$

Dies ist noch nicht die Behauptung, denn Z_1 und Z_2 stimmen im Allgemeinen nicht überein. In Aussage (b) kommt hingegen nur eine Zerlegung Z vor. Definiere nun noch $Z := Z_1 \vee Z_2$. Dann ist Z auch eine Zerlegung von $[a, b]$ und auch eine Verfeinerung von sowohl Z_1 als auch Z_2 , vgl. Bemerkung 8.1.5. Nach Proposition 8.1.9 gilt dann

$$\overline{S}_Z(f) \leq \overline{S}_{Z_1}(f) \quad \text{und} \quad \underline{S}_Z(f) \geq \underline{S}_{Z_2}(f). \quad (421)$$

Wir folgern, dass

$$\overline{S}_Z(f) - \underline{S}_Z(f) \stackrel{(421)}{\leq} \overline{S}_{Z_1}(f) - \underline{S}_Z(f) \stackrel{(421)}{\leq} \overline{S}_{Z_1}(f) - \underline{S}_{Z_2}(f) \stackrel{(420)}{\leq} \varepsilon.$$

Um die Riemann-Integrierbarkeit zu beweisen genügt es also, Zerlegungen mit beliebig kleinen Schwankungssummen zu finden (vgl. Proposition 8.1.18). Hat man solche Zerlegungen einmal gefunden, so kann man sie auch dazu benutzen um den Wert des Riemann-Integrals zu approximieren. Der Fehler ist dabei nicht größer als die Schwankungssumme.

Lemma 8.1.19

Ist $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar, so gilt für jede Zerlegung Z von $[a, b]$

(i)

$$\left| \int_a^b f(x) \, dx - \underline{S}_Z(f) \right| \leq \overline{S}_Z(f) - \underline{S}_Z(f).$$

(ii)

$$\left| \overline{S}_Z(f) - \int_a^b f(x) \, dx \right| \leq \overline{S}_Z(f) - \underline{S}_Z(f).$$

Beweis. Wir zeigen nun (i), denn (ii) ist analog. Es sei $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar und Z eine Zerlegung von $[a, b]$. Zunächst gilt

$$\begin{aligned} \int_a^b f(x) \, dx - \underline{S}_Z(f) &\stackrel{\text{Def8.1.15}}{=} \int_a^b f(x) \, dx - \underline{S}_Z(f) \\ &\stackrel{\text{Def8.1.12}}{=} \sup\{\underline{S}_{Z_0}(f) : Z_0 \text{ Zerlegung von } [a, b]\} - \underline{S}_Z(f) \geq 0. \end{aligned}$$

Außerdem gilt

$$\begin{aligned} \int_a^b f(x) \, dx - \underline{S}_Z(f) &\stackrel{\text{Def8.1.15}}{=} \int_a^b f(x) \, dx - \underline{S}_Z(f) \\ &\stackrel{\text{Def8.1.12}}{=} \inf\{\overline{S}_{Z_0}(f) : Z_0 \text{ Zerlegung von } [a, b]\} \leq \overline{S}_Z(f) - \underline{S}_Z(f). \end{aligned}$$

Wir folgern aus den beiden Rechnungen

$$0 \leq \int_a^b f(x) \, dx - \underline{S}_Z(f) \leq \overline{S}_Z(f) - \underline{S}_Z(f).$$

Dies impliziert (i).

Bemerkung 8.1.20

Hat man für ein vorgegebenes $\varepsilon > 0$ eine Zerlegung Z von $[a, b]$ mit $\overline{S}_Z(f) - \underline{S}_Z(f) \leq \varepsilon$ (vgl. Proposition 8.1.18) so gilt nach dem vorigen Lemma

$$\left| \int_a^b f(x) \, dx - \underline{S}_Z(f) \right| \leq \varepsilon \quad \text{und} \quad \left| \overline{S}_Z(f) - \int_a^b f(x) \, dx \right| \leq \varepsilon.$$

Mit anderen Worten: Das Integral wird bis auf einen Fehler von ε approximiert. Hat man also bei der Überprüfung des Integrierbarkeitskriteriums (Proposition 8.1.18) eine

Zerlegung gefunden die Proposition 8.1.18 (b) für ein kleines ε erfüllt, so kann mit dieser Zerlegung auch bereits das Integral gut approximiert werden.

Nun nutzen wir das oben hergeleitete Kriterium um die Riemann-Integrierbarkeit diverser Klassen von Funktionen zu beweisen.

Proposition 8.1.21 — MONOTONE FUNKTIONEN SIND RIEMANN-INTEGRIERBAR

Es sei $f : [a, b] \rightarrow \mathbb{R}$ monoton^a und beschränkt. Dann ist f Riemann-integrierbar. Ferner gilt (mit der Notation wie aus Beispiel 8.1.2)

$$\int_a^b f(x) \, dx = \lim_{m \rightarrow \infty} \bar{S}_{Z_m^{(E)}}(f) = \lim_{m \rightarrow \infty} \underline{S}_{Z_m^{(E)}}(f). \quad (422)$$

^a„Monoton“ bedeutet hier wie immer „monoton wachsend oder monoton fallend“.

06.05.2026

Beweis.

SCHRITT 1. Zur Riemann-Integrierbarkeit.

Wir zeigen die Behauptung nur für den Fall, dass $f : [a, b] \rightarrow \mathbb{R}$ monoton wachsend (und beschränkt) ist. Bei monoton fallenden Funktionen kann man ganz analog vorgehen. Wir werden Proposition 8.1.18 (b) nachweisen. Betrachte dazu für ein $m \in \mathbb{N}$ die Zerlegung $Z_m^{(E)} = (x_0, \dots, x_m)$, wobei

$$x_k = a + \frac{k}{m}(b - a) \quad \forall k \in \{0, \dots, m\}. \quad (423)$$

Wir berechnen im Folgenden für ein beliebiges $m \in \mathbb{N}$ die Schwankungssumme $\bar{S}_{Z_m^{(E)}}(f) - \underline{S}_{Z_m^{(E)}}(f)$. Dazu erinnern wir uns zunächst, dass

$$\bar{S}_{Z_m^{(E)}}(f) = \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}). \quad (424)$$

Nun gilt

$$\sup_{[x_{i-1}, x_i]} f = \sup f([x_{i-1}, x_i]) \xlongequal[f \text{ monoton wachsend}]{} \sup([f(x_{i-1}), f(x_i)]) \xlongequal[\text{Prop 3.2.25}]{} f(x_i). \quad (425)$$

Außerdem gilt mit (423) für alle $i \in \{1, \dots, m\}$

$$x_i - x_{i-1} = \frac{b - a}{m}, \quad (426)$$

vgl. Fußnote von Beispiel 8.1.2. Nutzen wir die vorigen beiden Gleichungen um (424) umzuformen so gilt

$$\bar{S}_{Z_m^{(E)}}(f) = \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) \stackrel{(425), (426)}{=} \sum_{i=1}^m f(x_i) \frac{b-a}{m} = \frac{b-a}{m} \sum_{i=1}^m f(x_i).$$

Völlig analog leitet man her, dass

$$\underline{S}_{Z_m^{(E)}}(f) = \sum_{i=1}^m \left(\inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) \stackrel{(426)}{=} \sum_{i=1}^m f(x_{i-1}) \frac{b-a}{m} = \frac{b-a}{m} \sum_{i=1}^m f(x_{i-1}).$$

Mit den obigen Formeln können wir die Schwankungssumme gut ausrechnen. Man erhält

$$\begin{aligned} \bar{S}_{Z_m^{(E)}}(f) - \underline{S}_{Z_m^{(E)}}(f) &= \frac{b-a}{m} \sum_{i=1}^m f(x_i) - \frac{b-a}{m} \sum_{i=1}^m f(x_{i-1}) \\ &= \frac{b-a}{m} \left(\sum_{i=1}^m f(x_i) - \sum_{i=1}^m f(x_{i-1}) \right) \\ &= \frac{b-a}{m} \sum_{i=1}^m (f(x_i) - f(x_{i-1})) \\ &\stackrel{\text{Teleskopsumme}}{=} \frac{b-a}{m} (f(x_m) - f(x_0)) \stackrel{\text{Def 8.1.1}}{=} \frac{b-a}{m} (f(b) - f(a)) \end{aligned}$$

Diese Gleichung gilt für alle $m \in \mathbb{N}$. Wir wissen also

$$0 \underset{\text{Prop 8.1.9}}{\leq} \bar{S}_{Z_m^{(E)}}(f) - \underline{S}_{Z_m^{(E)}}(f) \leq \frac{b-a}{m} (f(b) - f(a)) \quad \forall m \in \mathbb{N}.$$

Beachte: Es gilt $\lim_{m \rightarrow \infty} \frac{b-a}{m} (f(b) - f(a)) = 0$. Nach dem Sandwichlemma folgt dann auch

$$\lim_{m \rightarrow \infty} \left(\bar{S}_{Z_m^{(E)}}(f) - \underline{S}_{Z_m^{(E)}}(f) \right) = 0. \quad (427)$$

Erinnern wir uns an die Definition der Folgenkonvergenz so folgt

$$\forall \varepsilon > 0 \quad \exists m_0 \in \mathbb{N} \quad \forall m \geq m_0 : \quad |\bar{S}_{Z_m^{(E)}}(f) - \underline{S}_{Z_m^{(E)}}(f)| \leq \varepsilon.$$

Es sei nun $\varepsilon > 0$ beliebig vorgegeben. Dann gilt für m_0 gewählt wie oben

$$\bar{S}_{Z_{m_0}^{(E)}}(f) - \underline{S}_{Z_{m_0}^{(E)}}(f) \leq \varepsilon.$$

Das Kriterium aus Proposition 8.1.18 (b) ist also mit $Z = Z_{m_0}^{(E)}$ erfüllt. Folglich ist f Riemann-integrierbar.

SCHRITT 2. Die Approximationseigenschaft (422).

Es gilt nach Lemma 8.1.19 (i) für alle $m \in \mathbb{N}$

$$\left| \int_a^b f(x) \, dx - \underline{S}_{Z_m^{(E)}}(f) \right| \leq \overline{S}_{Z_m^{(E)}}(f) - \underline{S}_{Z_m^{(E)}}(f).$$

Da die rechte Seite nach (427) gegen 0 konvergiert, gilt auch (wieder nach dem Sandwichlemma)

$$\lim_{m \rightarrow \infty} \left| \int_a^b f(x) \, dx - \underline{S}_{Z_m^{(E)}}(f) \right| = 0.$$

Daraus folgert man (vgl. Proposition 2.3.30), dass

$$\lim_{m \rightarrow \infty} \underline{S}_{Z_m^{(E)}}(f) = \int_a^b f(x) \, dx.$$

Ein analoges Vorgehen mit Lemma 8.1.19(ii) zeigt auch

$$\lim_{m \rightarrow \infty} \overline{S}_{Z_m^{(E)}}(f) = \int_a^b f(x) \, dx.$$

Die Behauptung folgt.

Damit gibt es nun schonmal sehr viele integrierbare Funktionen. Nun kommen noch mehr dazu: Alle *stetigen Funktionen* sind auch Riemann-integrierbar.

Proposition 8.1.22 — STETIGE FUNKTIONEN SIND RIEMANN-INTEGRIERBAR

Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Dann ist f auch Riemann-integrierbar.

Beweis. Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig auf $[a, b]$. Normalerweise müsste man für die Riemann-integrierbarkeit erst sicherstellen, dass $f : [a, b] \rightarrow \mathbb{R}$ beschränkt ist. Dies folgt für eine stetige Funktion auf $[a, b]$ aber sofort aus dem Satz von Weierstraß, vgl. Satz 6.2.12. Wiederum nutzen wir das Integrierbarkeits-Kriterium aus Proposition 8.1.18(b). Dazu sei $\varepsilon > 0$ beliebig vorgegeben.

$$\textbf{Z}: \quad \text{Es gibt eine Zerlegung } Z \text{ von } [a, b] \text{ mit } \overline{S}_Z(f) - \underline{S}_Z(f) \leq \varepsilon.$$

Da f auf dem abgeschlossenen und beschränkten Intervall $[a, b]$ stetig ist, ist f dort auch gleichmäßig stetig, vgl. Satz 6.4.2. Wir folgern aus der Definition der gleichmäßigen Stetigkeit (Definition 6.4.1, mit $\varepsilon' := \frac{\varepsilon}{b-a} > 0$).

$$\exists \delta > 0 : \quad \forall x, y \in [a, b] \text{ mit } |x - y| < \delta \quad \text{gilt} \quad |f(x) - f(y)| < \frac{\varepsilon}{b-a}. \quad (428)$$

Es sei nun $Z = (x_0, \dots, x_m)$ eine beliebige Zerlegung mit der Eigenschaft, dass

$$x_i - x_{i-1} < \delta \quad \forall i \in \{1, \dots, m\}. \quad a \quad (429)$$

Wir bilden nun die Schankungssumme $\bar{S}_Z(f) - \underline{S}_Z(f)$. Dazu

$$\bar{S}_Z(f) - \underline{S}_Z(f) = \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) - \sum_{i=1}^m \left(\inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) \quad (430)$$

$$= \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}). \quad (431)$$

Nun ist f auf $[x_{i-1}, x_i]$ stetig und nimmt daher nach dem Satz von Weierstraß Maximum und Minimum an (vgl. Satz 6.2.12). Somit gilt für alle $i \in \{1, \dots, m\}$

$$\sup_{[x_{i-1}, x_i]} f \stackrel{\text{Prop 3.2.2}}{=} \max_{[x_{i-1}, x_i]} f = f(\eta_i) \quad \text{für eine Maximumsstelle } \eta_i \in [x_{i-1}, x_i] \quad (432)$$

und

$$\inf_{[x_{i-1}, x_i]} f \stackrel{\text{Prop 3.2.2}}{=} \min_{[x_{i-1}, x_i]} f = f(\theta_i) \quad \text{für eine Minimumsstelle } \theta_i \in [x_{i-1}, x_i]. \quad (433)$$

Wir erhalten

$$\begin{aligned} \bar{S}_Z(f) - \underline{S}_Z(f) &\stackrel{(431)}{=} \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) \\ &\stackrel{(432), (433)}{=} \sum_{i=1}^m (f(\eta_i) - f(\theta_i)) (x_i - x_{i-1}). \end{aligned} \quad (434)$$

Beachte: Da für alle i gilt $\eta_i, \theta_i \in [x_{i-1}, x_i]$ gilt

$$|\eta_i - \theta_i| = \max\{\eta_i, \theta_i\} - \min\{\eta_i, \theta_i\} \leq x_i - x_{i-1} \stackrel{(429)}{<} \delta.$$

Somit gilt nach (428) auch $|f(\eta_i) - f(\theta_i)| < \frac{\varepsilon}{b-a}$. Setzen wir dies in (434) ein so erhalten wir

$$\begin{aligned} \bar{S}_Z(f) - \underline{S}_Z(f) &\leq \sum_{i=1}^m \frac{\varepsilon}{b-a} (x_i - x_{i-1}) \\ &= \frac{\varepsilon}{b-a} \sum_{i=1}^m (x_i - x_{i-1}) \stackrel{\text{Teleskopsumme}}{=} \frac{\varepsilon}{b-a} (x_m - x_0) \stackrel{\text{Def 8.1.1}}{=} \frac{\varepsilon}{b-a} (b - a) = \varepsilon. \end{aligned}$$

^aBeachte: Eine solche Zerlegung gibt es stets, denn es gilt zum Beispiel für $Z = Z_m^{(E)}$ mit $m := \lceil \frac{b-a}{\delta} \rceil + 1$, dass

$$x_i - x_{i-1} \stackrel{\text{Bsp 8.1.2}}{=} \left(a + \frac{i}{m}(b-a) \right) - \left(a + \frac{i-1}{m}(b-a) \right) = \frac{b-a}{m} < \frac{b-a}{\frac{b-a}{\delta}} = \delta.$$

Bemerkung 8.1.23

Wir wissen nun, dass eine stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar ist. Mal wieder sollte man sich an dieser Stelle fragen, für was für Zerlegungen Z das Integral

$$\int_a^b f(x) \, dx$$

gut von $\overline{S}_Z(f)$ und $\underline{S}_Z(f)$ approximiert wird. Hierfür nutzt man stets die Fehlerabschätzung in Lemma 8.1.19 und die Tatsache, dass man im Beweis der Riemann-Integrierbarkeit Zerlegungen Z mit kleiner Schwankungssumme $\overline{S}_Z(f) - \underline{S}_Z(f)$ gefunden hat. Geht man den Beweis von Proposition 8.1.22 noch einmal durch so kann man die Zerlegungen identifizieren, die eine gute Approximation an das Integral liefern. Dazu:

Definition. Für eine gegebene Zerlegung $Z = (x_0, \dots, x_m)$ von $[a, b]$ definieren wir das *Feinheitsmaß* $\Delta(Z) := \max\{x_i - x_{i-1} : i = 1, \dots, m\}$.

Beobachtung. Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und $(Z_n)_{n \in \mathbb{N}}$ eine Folge von Zerlegungen derart, dass $\lim_{n \rightarrow \infty} \Delta(Z_n) = 0$. Dann gilt

$$\lim_{n \rightarrow \infty} \overline{S}_{Z_n}(f) = \lim_{n \rightarrow \infty} \underline{S}_{Z_n}(f) = \int_a^b f(x) \, dx.$$

Der Beweis ist den Lesenden als Übungsaufgabe überlassen. Er basiert auf der Technik von Proposition 8.1.22.

Wir haben nun gesehen, dass ziemlich viele (beschränkte) Funktionen Riemann-integrierbar sind: monotone, stetige, etc etc... Zum Schluss dieses Abschnittes wollen wir uns auch einmal davon überzeugen, dass es *nicht-Riemann-integrierbare* Funktionen gibt.

Klarerweise sind unbeschränkte Funktionen nicht Riemann-integrierbar, da Definition 8.1.15 Beschränktheit benötigt. Es gibt aber auch *beschränkte nicht-Riemann-integrierbare* Funktionen.

Beispiel 8.1.24

Behauptung. Die sogenannte *Dirichlet-Funktion*

$$f : [0, 1] \rightarrow \mathbb{R}, \quad f(x) := \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases} \quad (435)$$

ist beschränkt, aber nicht Riemann-integrierbar.

Beweis. Dass f beschränkt ist folgt ziemlich direkt aus der Definition von f (schließlich nimmt f ja nur zwei Werte an). Es sei nun $Z = (x_0, \dots, x_m)$ eine Zerlegung von $[0, 1]$. Das heißt insbesondere $x_0 < x_1 < \dots < x_m$ (vgl. Definition 8.1.1). Nach Proposition 3.4.14 gibt es nun

- eine rationale Zahl $q_1 \in \mathbb{Q}$ mit $x_0 < q_1 < x_1$,
- eine rationale Zahl $q_2 \in \mathbb{Q}$ mit $x_1 < q_2 < x_2$,
- ...
- und eine rationale Zahl $q_m \in \mathbb{Q}$ mit $x_{m-1} < q_m < x_m$.

Jedes Unterteilungsintervall enthält also einen rationalen Punkt. Daher gilt für alle $i \in \{1, \dots, m\}$

$$\sup_{[x_{i-1}, x_i]} f \geq f(q_i) \stackrel{(435)}{=} 1.$$

Somit hat man

$$\overline{S}_Z(f) = \sum_{i=1}^m \left(\sup_{[x_i, x_{i-1}]} f \right) (x_i - x_{i-1}) \geq \sum_{i=1}^m (x_i - x_{i-1}) = x_m - x_0 = 1 - 0 = 1. \quad (436)$$

Nun enthält auch jedes Intervall $[x_{i-1}, x_i]$ auch eine irrationale Zahl, nennen wir sie ξ_i ^a. Daher gilt für alle $i \in \{1, \dots, m\}$

$$\inf_{[x_{i-1}, x_i]} f \leq f(\xi_i) = 0.$$

Für die Untersummen hat man dann

$$\underline{S}_Z(f) = \sum_{i=1}^m \left(\inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) \leq \sum_{i=1}^m 0 = 0. \quad (437)$$

Nun überlegen wir uns was obige Erkenntnis für Oberintegral und Unterintegral bedeuten. Für das Oberintegral gilt wegen (436)

$$\int_a^b f(x) dx = \inf \{ \overline{S}_Z(f) : Z \text{ Zerlegung von } [a, b] \} \geq 1.$$

Für das Unterintegral folgern wir aus (437)

$$\int_a^b f(x) dx = \sup \{ \underline{S}_Z(f) : Z \text{ Zerlegung von } [a, b] \} \leq 0.$$

^aDiese kann man so konstruieren: Wähle ein hinreichend großes $n_i \in \mathbb{N}$ so, dass $\frac{\sqrt{2}}{n_i} < q_i - x_{i-1}$ (*). Definiere dann $\xi_i := q_i - \frac{\sqrt{2}}{n_i}$.

- ξ_i ist irrational. Denn angenommen ξ_i wäre rational, so wäre auch $n_i(q_i - \xi_i)$ rational (weil $\mathbb{Q}$ ein Körper ist). Man überzeugt sich aber leicht anhand der Definition von ξ_i , dass $n_i(q_i - \xi_i) = \sqrt{2}$. Ein Widerspruch zu Proposition 3.4.3
- $\xi_i < q_i < x_i$ und $\xi_i = q_i - \frac{\sqrt{2}}{n_i} \underset{(*)}{>} q_i - (q_i - x_{i-1}) = x_{i-1}$. Das bedeutet $x_{i-1} < \xi_i < x_i$.

8.2 Eigenschaften des Integrals

Nun widmen wir uns nützlichen Eigenschaften des Integrals, die wir bei expliziten Berechnungen häufig verwenden werden.

Linearität des Integrals

Eine erste wichtige Eigenschaft ist das Verhalten des Integrals unter *linearen Operationen* (d.h. Addition und Bildung von Vielfachen). Wir werden sehen, dass Summen und reelle Vielfache Riemann-integrierbarer Funktionen wieder Riemann-integrierbar sind. In der Sprache der Linearen Algebra würde man nun sagen, dass die Menge der Riemann-integrierbaren Funktionen einen *Vektorraum* über $\mathbb{R}$ bildet. Ferner wird sich herausstellen, dass das Integral eine *lineare Abbildung* ist.

Hierzu müssen wir zunächst eine Vorüberlegung über die Supremumsbildung anstellen

Lemma 8.2.1

Es seien $A, B \subseteq \mathbb{R}$ beschränkte Mengen und $\lambda \in \mathbb{R}$. Definiere $A + B := \{a + b : a \in A, b \in B\}$ und $\lambda A := \{\lambda a : a \in A\}$. Dann gilt

(i)

$$\sup(A + B) \leq \sup(A) + \sup(B) \quad \text{und} \quad \inf(A + B) \geq \inf(A) + \inf(B)$$

(ii)

$$\sup(\lambda A) = \begin{cases} \lambda \sup(A) & \text{falls } \lambda \geq 0 \\ \lambda \inf(A) & \text{falls } \lambda < 0 \end{cases} \quad \text{und} \quad \inf(\lambda A) = \begin{cases} \lambda \inf(A) & \text{falls } \lambda \geq 0 \\ \lambda \sup(A) & \text{falls } \lambda < 0 \end{cases}$$

Beweis. Wir zeigen nur (i) für das Supremum. Für das Infimum funktioniert alles analog und (ii) ist den Lesenden als Übungsaufgabe überlassen. Um zu zeigen, dass $\sup(A + B) \leq \sup(A) + \sup(B)$, genügt es zu zeigen, dass $\sup(A) + \sup(B)$ eine obere Schranke von $A + B$ ist. Ist das gezeigt, kann man nämlich mit Definition 3.2.10 (i.b) (angewendet mit $s' := \sup(A) + \sup(B)$) folgern, dass $\sup(A + B) \leq \sup(A) + \sup(B)$. Da $\sup(A)$ eine obere Schranke von A ist und $\sup(B)$ eine obere Schranke von B ist gilt für alle $a \in A$ und $b \in B$

$$a + b = \underbrace{a}_{\leq \sup(A)} + \underbrace{b}_{\leq \sup(B)} \leq \sup(A) + \sup(B).$$

Mit anderen Worten

$$\forall c \in A + B \quad \text{gilt} \quad c \leq \sup(A) + \sup(B)$$

Somit ist $\sup(A) + \sup(B)$ eine obere Schranke von $A + B$ und die Behauptung ist gezeigt.

Basierend auf obigem Lemma können wir studieren, wie sich Obersummen und Untersummen unter linearen Operationen verhalten.

Proposition 8.2.2

Es seien $f, g : [a, b] \rightarrow \mathbb{R}$ beschränkt, $\lambda \in \mathbb{R}$ und Z eine Zerlegung von $[a, b]$. Dann gilt

$$(i) \quad \overline{S}_Z(f + g) \leq \overline{S}_Z(f) + \overline{S}_Z(g) \quad \text{und} \quad \underline{S}_Z(f + g) \geq \underline{S}_Z(f) + \underline{S}_Z(g).$$

$$(ii) \quad \overline{S}_Z(\lambda f) = \begin{cases} \lambda \overline{S}_Z(f) & \lambda \geq 0 \\ \lambda \underline{S}_Z(f) & \lambda \leq 0 \end{cases} \quad \text{und} \quad \underline{S}_Z(\lambda f) = \begin{cases} \lambda \underline{S}_Z(f) & \lambda \geq 0 \\ \lambda \overline{S}_Z(f) & \lambda \leq 0 \end{cases}.$$

Beweis. Wir zeigen nur (i) für die Obersumme. Der Rest bleibt wieder mal als Übungsaufgabe an die Lesenden gestellt. Sei etwa $Z = (x_0, \dots, x_m)$. Dann gilt

$$\overline{S}_Z(f + g) = \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} (f + g) \right) (x_i - x_{i-1}) \quad (438)$$

Nun gilt mit der Notation von Lemma 8.2.1

$$\begin{aligned} \sup_{[x_{i-1}, x_i]} (f + g) &= \sup\{(f + g)(x) : x \in [x_{i-1}, x_i]\} = \sup\{f(x) + g(x) : x \in [x_{i-1}, x_i]\} \\ &\leq \sup(f([x_{i-1}, x_i]) + g([x_{i-1}, x_i])) \\ (*) & \\ &\stackrel{\text{Lem 8.2.1}}{\leq} \sup f([x_{i-1}, x_i]) + \sup g([x_{i-1}, x_i]) \\ &= \sup_{[x_{i-1}, x_i]} f + \sup_{[x_{i-1}, x_i]} g. \end{aligned}$$

In (*) haben wir benutzt, dass $\{(f + g)(x) : x \in [x_{i-1}, x_i]\} \subseteq f([x_{i-1}, x_i]) + g([x_{i-1}, x_i])$.^(a) Mit (438) erhalten wir

$$\begin{aligned} \overline{S}_Z(f + g) &\leq \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} f + \sup_{[x_{i-1}, x_i]} g \right) (x_i - x_{i-1}) \\ &= \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) + \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} g \right) (x_i - x_{i-1}) \\ &= \overline{S}_Z(f) + \overline{S}_Z(g). \end{aligned}$$

Die behauptete Ungleichung folgt.

^aBeachte, dass hier wirklich nur die Teilmengeninklusion gilt, denn

$$f([x_{i-1}, x_i]) + g([x_{i-1}, x_i]) = \{a + b : a \in f([x_{i-1}, x_i]), b \in g([x_{i-1}, x_i])\} = \{f(x) + g(y) \mid x, y \in [x_{i-1}, x_i]\}.$$

Dies ist offensichtlich eine Obermenge von $\{f(x) + g(x) \mid x \in [x_{i-1}, x_i]\}$, wie man sieht wenn man den Spezialfall $y = x$ betrachtet.

Die Ungleichungen die wir nun für Ober/Untersummen hergeleitet haben übertragen sich leicht auf Ober/Unter-integrale

Proposition 8.2.3

Es seien $f, g : [a, b] \rightarrow \mathbb{R}$ beschränkt und $\lambda \in \mathbb{R}$. Dann gilt

(i)

$$\begin{aligned} \overline{\int_a^b} (f + g)(x) \, dx &\leq \overline{\int_a^b} f(x) \, dx + \overline{\int_a^b} g(x) \, dx \\ \text{und} \quad \underline{\int_a^b} (f + g)(x) \, dx &\geq \underline{\int_a^b} f(x) \, dx + \underline{\int_a^b} g(x) \, dx. \end{aligned}$$

(ii)

$$\begin{aligned} \overline{\int_a^b} (\lambda f)(x) \, dx &= \begin{cases} \lambda \overline{\int_a^b} f(x) \, dx & \lambda \geq 0 \\ \lambda \underline{\int_a^b} f(x) \, dx & \lambda \leq 0 \end{cases} \\ \text{und} \quad \underline{\int_a^b} (\lambda f)(x) \, dx &= \begin{cases} \lambda \underline{\int_a^b} f(x) \, dx & \lambda \geq 0 \\ \lambda \overline{\int_a^b} f(x) \, dx & \lambda \leq 0 \end{cases} \end{aligned}$$

Beweis. Wieder einmal beweisen wir nur die Ungleichung (i) für das Oberintegral. Dazu: Nach Definition 8.1.12 gilt

$$\overline{\int_a^b} f(x) \, dx = \inf \{ \overline{S}_Z(f) : Z \text{ Zerlegung von } [a, b] \},$$

$$\underline{\int_a^b} g(x) \, dx = \inf \{ \overline{S}_Z(g) : Z \text{ Zerlegung von } [a, b] \}.$$

Es sei nun $\varepsilon > 0$ beliebig. Nach Proposition 3.2.26 findet man eine Zerlegung Z_1 von $[a, b]$

so, dass

$$\overline{S}_{Z_1}(f) \leq \int_a^b f(x) \, dx + \frac{\varepsilon}{2} \quad (439)$$

und

$$\overline{S}_{Z_2}(g) \leq \int_a^b g(x) \, dx + \frac{\varepsilon}{2}. \quad (440)$$

Definiere nun $Z := Z_1 \vee Z_2$. Wiederum nach Definition 8.1.12 gilt

$$\int_a^b (f+g)(x) \, dx \leq \overline{S}_Z(f+g) \stackrel{\text{Prop 8.2.2}}{\leq} \overline{S}_Z(f) + \overline{S}_Z(g). \quad (441)$$

Da Z eine Verfeinerung von Z_1 ist, folgt aus Proposition 8.1.9 (ii), dass $\overline{S}_Z(f) \leq \overline{S}_{Z_1}(f)$. Analog gilt $\overline{S}_Z(g) \leq \overline{S}_{Z_2}(g)$, da Z auch eine Verfeinerung von Z_2 ist. Wir folgern aus (441)

$$\begin{aligned} \int_a^b (f+g)(x) \, dx &\leq \overline{S}_{Z_1}(f) + \overline{S}_{Z_2}(g) \stackrel{(439),(440)}{\leq} \int_a^b f(x) \, dx + \frac{\varepsilon}{2} + \int_a^b g(x) \, dx + \frac{\varepsilon}{2} \\ &= \int_a^b f(x) \, dx + \int_a^b g(x) \, dx + \varepsilon. \end{aligned}$$

Nun ist $\varepsilon > 0$ beliebig. Man könnte nun wieder $\varepsilon = \frac{1}{n}$ für sämtliche $n \in \mathbb{N}$ betrachten, d.h.

$$\int_a^b (f+g)(x) \, dx \leq \int_a^b f(x) \, dx + \int_a^b g(x) \, dx + \frac{1}{n} \quad \forall n \in \mathbb{N}.$$

Geht man dann zum Grenzwert $n \rightarrow \infty$ über so folgert man

$$\int_a^b (f+g)(x) \, dx = \int_a^b f(x) \, dx + \int_a^b g(x) \, dx,$$

was zu zeigen war.

Bei Teil (i) der Proposition spricht man auch von *Subadditivität* des Oberintegrals bzw. *Superadditivität* des Unterintegrals. Mit dieser Vorbereitung können wir die *Linearität des Integrals* nun beweisen.

Satz 8.2.4 — LINEARITÄT DES RIEMANN-INTEGRALS

Es seien $f, g : [a, b] \rightarrow \mathbb{R}$ (beschränkte und) Riemann-integrierbare Funktionen und $\lambda \in \mathbb{R}$. Dann gilt

(i) $f + g$ ist auf $[a, b]$ Riemann-integrierbar und

$$\int_a^b (f+g)(x) \, dx = \int_a^b f(x) \, dx + \int_a^b g(x) \, dx.$$

(ii) λf ist auf $[a, b]$ Riemann-integrierbar und

$$\int_a^b (\lambda f)(x) \, dx = \lambda \int_a^b f(x) \, dx.$$

Beweis. Wir zeigen nur (i). Da f und g Riemann-integrierbar sind gilt nach Definition 8.1.15

$$\int_a^b f(x) \, dx = \overline{\int_a^b f(x) \, dx} = \int_a^b f(x) \, dx \quad \text{und} \quad \int_a^b g(x) \, dx = \overline{\int_a^b g(x) \, dx} = \int_a^b g(x) \, dx. \quad (442)$$

SCHRITT 1. Riemann-Integrierbarkeit von $f + g$.

Nach Bemerkung 8.1.16 genügt es zu zeigen, dass

$$\overline{\int_a^b (f + g)(x) \, dx} \leq \int_a^b (f + g)(x) \, dx.$$

Dazu:

$$\overline{\int_a^b (f + g)(x) \, dx} \underset{\text{Prop 8.2.3}}{\leq} \overline{\int_a^b f(x) \, dx} + \overline{\int_a^b g(x) \, dx} \quad (443)$$

$$\begin{aligned} & \underset{(442)}{=} \int_a^b f(x) \, dx + \int_a^b g(x) \, dx \\ & \underset{(442)}{=} \int_a^b f(x) \, dx + \int_a^b g(x) \, dx \underset{\text{Prop 8.2.3}}{\leq} \int_a^b (f + g)(x) \, dx, \end{aligned} \quad (444)$$

was zu zeigen war. Somit ist $f + g$ Riemann-integrierbar. Insbesondere gilt nach Definition 8.1.15

$$\overline{\int_a^b (f + g)(x) \, dx} = \int_a^b (f + g)(x) \, dx = \int_a^b (f + g)(x) \, dx \quad (445)$$

SCHRITT 2. Additivität des Integrals.

Noch $\mathbb{Z}$.
$$\int_a^b (f + g)(x) \, dx = \int_a^b f(x) \, dx + \int_a^b g(x) \, dx.$$

Dazu beobachten wir zuerst

$$\int_a^b (f + g)(x) \, dx \underset{(445)}{=} \overline{\int_a^b (f + g)(x) \, dx} \underset{(443)}{\leq} \overline{\int_a^b f(x) \, dx} + \overline{\int_a^b g(x) \, dx}$$

$$= \int_a^b f(x) \, dx + \int_a^b g(x) \, dx$$

Zusätzlich gilt noch

$$\begin{aligned} \int_a^b f(x) \, dx + \int_a^b g(x) \, dx &= \int_a^b f(x) \, dx + \int_a^b g(x) \, dx \stackrel{(444)}{\leq} \int_a^b (f+g)(x) \, dx \\ &\stackrel{(445)}{=} \int_a^b (f+g)(x) \, dx \end{aligned}$$

Die beiden vorigen Gleichungen ergeben — wie behauptet —

$$\int_a^b (f+g)(x) \, dx = \int_a^b f(x) \, dx + \int_a^b g(x) \, dx.$$

Monotonie des Integrals

11.05.2026

Als nächstes diskutieren wir die *Monotonie des Integrals*. Das bedeutet: Je größer die Werte des Integranden, desto größer der Wert des Integrals.

Proposition 8.2.5 — MONOTONIE DES INTEGRALS

Es seien $f, g : [a, b] \rightarrow \mathbb{R}$ (beschränkt und) Riemann-integrierbar mit

$$f(x) \leq g(x) \quad \forall x \in [a, b].$$

Dann gilt auch

$$\int_a^b f(x) \, dx \leq \int_a^b g(x) \, dx.$$

Beweis.

SCHRITT 1. Spezialfall $f(x) = 0 \, \forall x \in [a, b]$, d.h. auch $g(x) \geq 0$ für alle $x \in [a, b]$

In dem Fall gilt

$$\int_a^b f(x) \, dx = \int_a^b 0 \, dx \stackrel{\text{Bsp8.1.17}}{=} 0 \cdot (b-a) = 0.$$

Daher ist zu zeigen

$$\mathbb{Z}\mathbb{Z} : \quad 0 \leq \int_a^b g(x) \, dx.$$

Dies überlegt man sich leicht. Für jede beliebige Zerlegung $Z = (x_0, \dots, x_m)$ von $[a, b]$ ist

$$\bar{S}_Z(g) = \sum_{i=1}^m \underbrace{\left(\sup_{[x_{i-1}, x_i]} g \right)}_{\geq g(x_i) \geq 0} (x_i - x_{i-1}) \geq \sum_{i=1}^m 0 \cdot (x_i - x_{i-1}) = 0.$$

Aus diesem Grund ist 0 eine untere Schranke von $\{\bar{S}_Z(g) : Z \text{ Zerlegung von } [a, b]\}$. Es folgt, dass

$$\int_a^b g(x) \, dx = \overline{\int_a^b g(x) \, dx} = \inf \{\bar{S}_Z(g) : Z \text{ Zerlegung von } [a, b]\} \geq 0.$$

Somit ist die Behauptung gezeigt.

SCHRITT 2. Allgemeiner Fall.

Es seien f und g wie in der Aussage, d.h. sie sind Riemann-integrierbar und $f(x) \leq g(x)$ für alle $x \in [a, b]$. Nach Satz 8.2.4 ist auch $g - f$ Riemann-integrierbar. Ferner gilt $(g - f)(x) = g(x) - f(x) \geq 0$ für alle $x \in [a, b]$. Schritt 1 (mit $\tilde{g} := g - f$) impliziert dann

$$0 \leq \int_a^b (g - f)(x) \, dx$$

Addieren wir $\int_a^b f(x) \, dx$ auf beiden Seiten so folgt

$$\int_a^b f(x) \, dx \leq \int_a^b f(x) \, dx + \int_a^b (g - f)(x) \, dx \stackrel{\text{Satz 8.2.4}}{=} \int_a^b (f + (g - f))(x) \, dx = \int_a^b g(x) \, dx.$$

Eine spannende Konsequenz aus der oben diskutierten Monotonie ist auch eine *Dreiecksungleichung* für Integrale. Wir erinnern, dass die Dreiecksungleichung für Summen besagt, dass für beliebige reelle Zahlen $a_1, \dots, a_n \in \mathbb{R}$ gilt (vgl. Proposition 4.2.13)

$$\left| \sum_{k=1}^n a_k \right| \leq \sum_{k=1}^n |a_k|$$

Da man sich Integrale als Grenzfälle von Summen vorstellen kann, ist nachfolgendes Resultat nicht überraschend:

Korollar 8.2.6 — DREIECKSUNGLEICHUNG FÜR INTEGRALE

Es sei $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar. Dann ist auch $\text{abs} \circ f : [a, b] \rightarrow \mathbb{R}, x \mapsto |f(x)|$ ^(a) Riemann-integrierbar und es gilt

$$\left| \int_a^b f(x) \, dx \right| \leq \int_a^b |f(x)| \, dx.$$

^aAls andere Bezeichnung für $\text{abs} \circ f$ werden wir im Folgenden auch $|f|$ verwenden.

Beweis.

SCHRITT 1. $\text{abs} \circ f$ ist Riemann-integrierbar.

Wir behaupten, dass für alle Zerlegungen Z von $[a, b]$ gilt

$$\overline{S}_Z(|f|) - \underline{S}_Z(|f|) \leq \overline{S}_Z(f) - \underline{S}_Z(f). \quad (446)$$

Dazu: Sei $Z = (x_0, \dots, x_m)$ eine Zerlegung von $[a, b]$. Für $i \in \{1, \dots, m\}$ und $x, y \in [x_{i-1}, x_i]$ gilt nach der inversen Dreiecksungleichung

$$|f(x)| - |f(y)| \leq ||f(x)| - |f(y)|| \leq |f(x) - f(y)| \leq \sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f.$$

Man kann daraus folgern, dass

$$\sup_{[x_{i-1}, x_i]} |f| - \inf_{[x_{i-1}, x_i]} |f| \leq \sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f.$$

Nun gilt

$$\begin{aligned} \overline{S}_Z(|f|) - \underline{S}_Z(|f|) &= \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} |f| - \inf_{[x_{i-1}, x_i]} |f| \right) (x_i - x_{i-1}) \\ &= \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) = \overline{S}_Z(f) - \underline{S}_Z(f). \end{aligned}$$

Jetzt, wo (446) gezeigt ist kann die Riemannintegrierbarkeit mit Proposition 8.1.18(b) nachgewiesen werden. Dazu sei $\varepsilon > 0$ beliebig aber fest. Nutzt man Proposition 8.1.18 für f so findet man eine Zerlegung Z von $[a, b]$ mit $\overline{S}_Z(f) - \underline{S}_Z(f) \leq \varepsilon$. Für diese Zerlegung gilt nach der obigen Gleichung

$$\overline{S}_Z(|f|) - \underline{S}_Z(|f|) \leq \overline{S}_Z(f) - \underline{S}_Z(f) \leq \varepsilon.$$

Das impliziert nach Proposition 8.1.18 die Riemann-Integrierbarkeit von $|f|$.

SCHRITT 2. Zur Ungleichung.

Für alle $x \in [a, b]$ gilt $f(x) \leq |f(x)|$. Nach Proposition 8.2.5 gilt dann also

$$\int_a^b f(x) \, dx \leq \int_a^b |f(x)| \, dx. \quad (447)$$

Desweiteren gilt für alle $x \in [a, b]$, dass $-f(x) \leq |f(x)|$. Proposition 8.2.5 liefert dann

$$\int_a^b (-f(x)) \, dx \leq \int_a^b |f(x)| \, dx.$$

Mit der Linearität aus Satz 8.2.4 folgt

$$-\int_a^b f(x) \, dx \leq \int_a^b |f(x)| \, dx.$$

Nun gilt

$$\left| \int_a^b f(x) \, dx \right| = \max \left\{ \int_a^b f(x) \, dx, -\int_a^b f(x) \, dx \right\} \stackrel{(447)}{\leq} \int_a^b |f(x)| \, dx.$$

Die Behauptung folgt.

BONUSKAPITEL. *Strenge* Monotonie des Integrals?

Eine Frage, die man sich noch stellen könnte ist, ob das Integral *streng monoton ist*. Das bedeutet: Angenommen

$f(x) \leq g(x) \quad \forall x \in [a, b]$ und an mindestens einem Punkt $x_0 \in [a, b]$ gilt sogar $f(x_0) < g(x_0)$.

Gilt dann auch

$$\int_a^b f(x) \, dx < \int_a^b g(x) \, dx \quad ?$$

Die Antwort ist nein, wie das folgende Beispiel zeigt.

Beispiel 8.2.7

Es sei $g : [-1, 1] \rightarrow \mathbb{R}$ so, dass

$$g(x) = \begin{cases} 1 & x = 0 \\ 0 & \text{sonst} \end{cases}.$$

Behauptung. g ist Riemann-integrierbar und $\int_{-1}^1 g(x) \, dx = 0$.

Insbesondere haben g und $f : [-1, 1] \rightarrow \mathbb{R}, x \mapsto 0$ dasselbe Integral, obwohl $f(x) \leq g(x)$ für alle $x \in [-1, 1]$ und $f(0) < g(0)$.

Beweis. Betrachte für $n \in \mathbb{N}$ die Zerlegung $Z_n := (-1, -\frac{1}{n}, \frac{1}{n}, 1)$. Dann gilt

$$\begin{aligned}\overline{S}_{Z_n}(g) &= \left(\sup_{[-1, -\frac{1}{n}]} g \right) \left(-\frac{1}{n} - (-1) \right) + \left(\sup_{[-\frac{1}{n}, \frac{1}{n}]} g \right) \left(\frac{1}{n} - (-\frac{1}{n}) \right) + \left(\sup_{[\frac{1}{n}, 1]} g \right) \left(1 - \frac{1}{n} \right) \\ &= 0 \cdot \left(-\frac{1}{n} - (-1) \right) + 1 \cdot \left(\frac{1}{n} - (-\frac{1}{n}) \right) + 0 \cdot \left(1 - \frac{1}{n} \right) = \frac{2}{n}\end{aligned}$$

und

$$\begin{aligned}\underline{S}_{Z_n}(g) &= \left(\inf_{[-1, -\frac{1}{n}]} g \right) \left(-\frac{1}{n} - (-1) \right) + \left(\inf_{[-\frac{1}{n}, \frac{1}{n}]} g \right) \left(\frac{1}{n} - (-\frac{1}{n}) \right) + \left(\inf_{[\frac{1}{n}, 1]} g \right) \left(1 - \frac{1}{n} \right) \\ &= 0 \cdot \left(-\frac{1}{n} - (-1) \right) + 0 \cdot \left(\frac{1}{n} - (-\frac{1}{n}) \right) + 0 \cdot \left(1 - \frac{1}{n} \right) = 0.\end{aligned}\tag{448}$$

Man erhält

$$\overline{S}_{Z_n}(g) - \underline{S}_{Z_n}(g) = \frac{2}{n} - 0 = \frac{2}{n}.\tag{449}$$

Nun lässt sich Proposition 8.1.18 (b) anwenden, um die Riemann-Integrierbarkeit zu zeigen. Es sei dazu $\varepsilon > 0$ beliebig vorgegeben. Dann gibt es ein $n_0 \in \mathbb{N}$ mit $\frac{2}{n_0} < \varepsilon$. Nach (449) hat man dann, dass

$$\overline{S}_{Z_{n_0}}(g) - \underline{S}_{Z_{n_0}}(g) = \frac{2}{n_0} < \varepsilon.$$

Somit ist nach Proposition 8.1.18 $g : [-1, 1] \rightarrow \mathbb{R}$ Riemann integrierbar. Außerdem gilt für alle $n \in \mathbb{N}$

$$\left| \int_{-1}^1 g(x) \, dx \right| \stackrel{(448)}{=} \left| \int_{-1}^1 g(x) \, dx - \underline{S}_{Z_n}(g) \right| \stackrel{\text{Prop 8.1.19}}{\leq} \overline{S}_{Z_n}(g) - \underline{S}_{Z_n}(g) = \frac{2}{n}.$$

Lässt man auf beiden Seiten $n \rightarrow \infty$ streben so folgert man

$$\left| \int_{-1}^1 g(x) \, dx \right| = 0 \quad \text{d.h. auch} \quad \int_{-1}^1 g(x) \, dx = 0.$$

Man sieht also: Das Integral ist im allgemeinen nicht streng monoton. Ein größerer Integrand erzwingt nicht zwingen ein größeres Integral. Unter einer Einschränkung ist das allerdings doch richtig: Wenn beide Integranden stetig sind!

Proposition 8.2.8

Es seien $f, g : [a, b] \rightarrow \mathbb{R}$ stetig. Ferner gelte $f(x) \leq g(x)$ für alle $x \in [a, b]$ und $f \neq g$.^a Dann gilt

$$\int_a^b f(x) \, dx < \int_a^b g(x) \, dx.$$

^aDas bedeutet in unserem Fall, dass an mindestens einem Punkt $x_0 \in [a, b]$ die strikte Ungleichung $f(x_0) < g(x_0)$ gelten muss!

Beweis. Es sei $x_0 \in [a, b]$ ein beliebiger Wert mit $f(x_0) < g(x_0)$. Da f und g stetig sind ist auch $g - f$ stetig. Ferner gilt $(g - f)(x) \geq 0$ für alle $x \in [a, b]$ und $(g - f)(x_0) > 0$. Definiert man $\varepsilon := \frac{1}{2}(g - f)(x_0) > 0$ und wendet Bemerkung 6.2.4 an so lernt man

$$\exists \delta > 0 : \quad \forall x \in [a, b] : |x - x_0| < \delta \text{ gilt } |(g - f)(x) - (g - f)(x_0)| < \frac{1}{2}(g - f)(x_0).$$

Man erhält für alle $x \in (x_0 - \delta, x_0 + \delta) \cap [a, b]$, dass

$$\begin{aligned} (g - f)(x) &= (g - f)(x_0) - [(g - f)(x) - (g - f)(x_0)] \\ &\geq (g - f)(x_0) - |(g - f)(x) - (g - f)(x_0)| \\ &> (g - f)(x_0) - \frac{1}{2}(g - f)(x_0) \\ &= \frac{1}{2}(g - f)(x_0). \end{aligned}$$

Definiere nun $a_0 := \min\{a, x_0 - \delta\}$ und $b_0 := \max\{b, x_0 + \delta\}$ und ferner

$$h(x) := \begin{cases} \frac{1}{2}(g - f)(x_0) & x \in (a_0, b_0) \\ 0 & \text{sonst} \end{cases}.$$

Die Monotonie des Integrals (vgl. Proposition 8.2.5) impliziert dann

$$\int_a^b (g - f)(x) \, dx \geq \int_a^b h(x) \, dx \quad (450)$$

Man überlegt sich leicht als Übungsaufgabe (eventuell unter Verwendung des nächsten Abschnittes), dass

$$\int_a^b h(x) \, dx = \frac{1}{2}(g - f)(x_0)(b_0 - a_0) \quad (451)$$

Nun gilt $b_0 - a_0 \geq \min\{\delta, b - a\} > 0$.^a Da desweiteren $(g - f)(x_0) > 0$ folgt aus (451), dass $\int_a^b h(x) \, dx > 0$. Wegen (450) gilt

$$\int_a^b (g - f)(x) \, dx \geq \int_a^b h(x) \, dx > 0. \quad (452)$$

Nutzt man die Linearität des Integrals so erhält man

$$\int_a^b g(x) \, dx - \int_a^b f(x) \, dx = \int_a^b (g - f)(x) \, dx \stackrel{(452)}{>} 0.$$

^aDies überlegt man sich leicht: Falls $a_0 = x_0 - \delta, b_0 = x_0 + \delta$ so hat man $b_0 - a_0 = 2\delta \geq \min\{\delta, b - a\}$. Falls $a_0 = a$ und $b_0 = x_0 + \delta$ so hat man $b_0 - a_0 = x_0 + \delta - a = (x_0 - a) + \delta \geq 0 + \delta = \delta \geq \min\{b - a, x_0\}$. Die anderen Fälle sind den Lesenden als Übungsaufgabe überlassen.

Teilungseigenschaft des Integrals

Es ist desweiteren möglich, bei Riemann-integrierbaren Funktionen $f : [a, b] \rightarrow \mathbb{R}$ das Integral $\int_a^b f(x) \, dx$ in mehrere integrale *aufzuteilen*, d.h. für alle $c \in (a, b)$ gilt

$$\int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx. \quad (453)$$

Um diese Aufteilung zu begründen müssen wir uns aber erstmal überlegen, dass f auf $[a, c]$ und $[c, b]$ überhaupt Riemann-integrierbar ist.

Definition 8.2.9 — EINSCHRÄNKUNG

Es sei $f : D \rightarrow \mathbb{R}$ eine Funktion und $\tilde{D} \subseteq D$. Dann definieren wir die *Einschränkung* von f auf $\tilde{D}$ durch

$$f|_{\tilde{D}} : \tilde{D} \rightarrow \mathbb{R}, x \mapsto f(x).$$

Die Einschränkung auf Teilintervalle zerstört die Riemann-Integrierbarkeit nicht, wie wir nun sehen werden.

Proposition 8.2.10 — EINSCHRÄNKUNGEN RIEMANN-INTEGRIERBARER FUNKTIONEN

Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion und $c \in (a, b)$. Dann ist auch $f|_{[a, c]} : [a, c] \rightarrow \mathbb{R}$ und $f|_{[c, b]} : [c, b] \rightarrow \mathbb{R}$ Riemann-integrierbar.

Beweis. Es sei $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar und $c \in (a, b)$. Zuerst zur Riemann-integrierbarkeit von $f|_{[a, c]}$. Es sei $\varepsilon > 0$ beliebig vorgegeben. Nach Proposition 8.1.18 gibt es eine Zerlegung $Z = (x_0, \dots, x_m)$ von $[a, b]$ mit

$$\overline{S}_Z(f) - \underline{S}_Z(f) \leq \varepsilon.$$

Beachte:

$$\begin{aligned} \overline{S}_Z(f) - \underline{S}_Z(f) &= \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) - \sum_{i=1}^m \left(\inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) \\ &= \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}). \end{aligned} \quad (454)$$

Wir müssen zeigen (vgl. Proposition 8.1.18), dass es eine Zerlegung Z_1 von $[a, c]$ gibt mit

$$\overline{S}_{Z_1}(f|_{[a, c]}) - \underline{S}_{Z_1}(f|_{[a, c]}) \leq \varepsilon.$$

Dazu unterscheiden wir zwei Fälle

FALL 1. $c \in \{x_0, \dots, x_m\}$. Etwa gelte $c = x_r$ für ein $r \in \{0, \dots, m\}$.

Es sei dann $Z_1 := (x_0, \dots, x_r)$. Dann ist Z_1 eine Zerlegung von $[a, c]$ im Sinne von Definition 8.1.1, denn $x_0 = a, x_r = c$ und $x_0 < \dots < x_r$. Nun gilt

$$\begin{aligned}\overline{S}_{Z_1}(f|_{[a,c]}) - \underline{S}_{Z_1}(f|_{[a,c]}) &= \sum_{i=1}^r \left(\sup_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) - \sum_{i=1}^r \left(\inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) \\ &= \sum_{i=1}^r \left(\sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1})\end{aligned}\quad (455)$$

Beachte: Jeder Summand der Summe in (454) ist positiv (*), da für alle $i \in \{1, \dots, m\}$ gilt $\sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f \geq 0$ und $x_i - x_{i-1} > 0$. Mit dieser Information können wir berechnen

$$\begin{aligned}\overline{S}_{Z_1}(f|_{[a,c]}) - \underline{S}_{Z_1}(f|_{[a,c]}) &\stackrel{(455)}{=} \sum_{i=1}^r \left(\sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) \\ &\stackrel{(\leq)}{\stackrel{(*)}{\leq}} \sum_{i=1}^m \left(\sup_{[x_{i-1}, x_i]} f - \inf_{[x_{i-1}, x_i]} f \right) (x_i - x_{i-1}) = \overline{S}_Z(f) - \underline{S}_Z(f) \leq \varepsilon.\end{aligned}$$

FALL 2. $c \notin \{x_0, \dots, x_m\}$.

Es sei Z' die Zerlegung von $[a, b]$, die aus allen Unterteilungspunkten von Z und zusätzlich noch c besteht.^a Klarerweise ist Z' eine Verfeinerung von Z . Nach Proposition 8.1.9 gilt $\overline{S}_{Z'}(f) \leq \overline{S}_Z(f)$ und $\underline{S}_{Z'}(f) \geq \underline{S}_Z(f)$. Das bedeutet auch

$$\overline{S}_{Z'}(f) - \underline{S}_{Z'}(f) \leq \overline{S}_Z(f) - \underline{S}_Z(f) \leq \varepsilon.$$

Zusätzlich dazu ist auch c eine Unterteilungsstelle in Z' . Somit erfüllt Z' alle Annahmen die in FALL 1 getroffen worden. Wie in Fall 1 kann man dann auch eine Zerlegung Z_1 von $[a, c]$ konstruieren so, dass

$$\overline{S}_{Z_1}(f|_{[a,c]}) - \underline{S}_{Z_1}(f|_{[a,c]}) \leq \varepsilon.$$

Da wir nun in beiden Fällen eine solche Zerlegung gefunden haben folgt die Behauptung.

Nun müsste man eigentlich noch zeigen, dass $f|_{[c,b]}$ Riemann-integrierbar ist. Das geht analog und ist den Lesenden als Übungsaufgabe überlassen.

^aGenauer sollte man wohl schreiben: $Z' = Z \vee (a, c, b)$. Noch genauer: Man findet zuerst einen Index $\ell \in \{1, \dots, m\}$ so, dass $x_{\ell-1} < c < x_\ell$ und setzt dann $Z' := (x_0, \dots, x_{\ell-1}, c, x_\ell, \dots, x_m)$.

Bemerkung 8.2.11

Ganz ähnlich kann man auch die Riemann-Integrierbarkeit der Einschränkung auf beliebige Teilintervalle zeigen:

Es sei $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar und $c, d \in [a, b]$ mit $c < d$.
Dann ist auch $f|_{[c, d]}$ auf $[c, d]$ Riemann-integrierbar.

Diesen Fakt werden wir im Folgenden oft ohne weiteren Kommentar verwenden.

Nun ist zwar die Riemann-integrierbarkeit der Einschränkungen $f|_{[a, c]}$ und $f|_{[c, b]}$ geklärt, einen Beweis der Formel (453) sind wir aber noch schuldig.

Proposition 8.2.12 — TEILUNGSFORMEL

Es sei $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar und $c \in (a, b)$. Dann gilt

$$\int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx \quad (a)$$

^aWer es ganz korrekt möchte, würde hier wohl $\int_a^b f(x) \, dx = \int_a^c f|_{[a, c]}(x) \, dx + \int_c^b f|_{[c, b]}(x) \, dx$ schreiben, damit auch klar ist, dass in allen Integranden eine Riemann-integrierbare Funktion steht (vgl. Proposition 8.2.10). Das werden wir aber in Zukunft nicht machen – die Werte $f|_{[a, c]}(x)$ und $f(x)$ sind ja schließlich für alle $x \in [a, c]$ dieselben.

Beweis.

$$\text{SCHRITT 1. } \int_a^b f(x) \, dx \geq \int_a^c f(x) \, dx + \int_c^b f(x) \, dx.$$

Es sei $Z = (x_0, \dots, x_m)$ eine beliebige Zerlegung von $[a, b]$. Außerdem sei Z' die Zerlegung, die daraus entsteht, dass man c zu Z hinzufügt (falls c nicht schon drin ist). Es sei nun $Z' = (\tilde{x}_0, \dots, \tilde{x}_{\tilde{m}})$ mit $c = \tilde{x}_r$ (für ein $r \in \{0, \dots, \tilde{m}\}$). Definiere $Z_1 := (\tilde{x}_0, \dots, \tilde{x}_r)$ und $Z_2 := (\tilde{x}_r, \dots, \tilde{x}_{\tilde{m}})$. Insbesondere ist Z_1 eine Zerlegung von $[a, c]$ und Z_2 eine Zerlegung von $[c, b]$. Dann gilt

$$\begin{aligned} \overline{S}_Z(f) &\underset{\text{Prop 8.1.9(iii)}}{\geq} \overline{S}_{Z'}(f) \underset{\text{Übungsaufgabe}}{=} \overline{S}_{Z_1}(f) + \overline{S}_{Z_2}(f) \underset{\text{Def 8.1.12}}{\geq} \int_a^c f(x) \, dx + \int_c^b f(x) \, dx. \\ &\underset{\text{Prop 8.2.10}}{=} \int_a^c f(x) \, dx + \int_c^b f(x) \, dx. \end{aligned} \quad (456)$$

Somit ist $\int_a^c f(x) \, dx + \int_c^b f(x) \, dx$ eine untere Schranke von $\{\overline{S}_Z(f) : Z \text{ Zerlegung von } [a, b]\}$.

Daher gilt

$$\overline{\int_a^b f(x) \, dx} \stackrel{\text{Def 8.1.12}}{=} \inf \{ \overline{S}_Z(f) : Z \text{ Zerlegung von } [a, b] \} \stackrel{(456)}{\geq} \int_a^c f(x) \, dx + \int_c^b f(x) \, dx.$$

Nutzt man die Riemann-Integrierbarkeit von f auf $[a, b]$ so folgt

$$\int_a^b f(x) \, dx \geq \int_a^c f(x) \, dx + \int_c^b f(x) \, dx.$$

SCHRITT 2. $\int_a^b f(x) \, dx \leq \int_a^c f(x) \, dx + \int_c^b f(x) \, dx.$

Für diese Ungleichung arbeiten wir mit den Untersummen. Sei wieder Z eine beliebige Zerlegung von $[a, b]$ und Z' die Zerlegung die durch Hinzunahme von c zu Z entsteht. Wieder einmal sei $Z' = (\tilde{x}_0, \dots, \tilde{x}_{\tilde{m}})$ mit $c = \tilde{x}_r$ und wir definieren $Z_1 := (\tilde{x}_0, \dots, \tilde{x}_r)$, $Z_2 := (\tilde{x}_r, \dots, \tilde{x}_{\tilde{m}})$. Dann gilt

$$\begin{aligned} \underline{S}_Z(f) &\stackrel{\text{Prop 8.1.9(ii)}}{\leq} \underline{S}_{Z'}(f) \stackrel{\text{Übungsaufgabe}}{=} \underline{S}_{Z_1}(f) + \underline{S}_{Z_2}(f) \leq \int_{\underline{a}}^c f(x) \, dx + \int_{\underline{c}}^b f(x) \, dx \\ &\stackrel{\text{Prop 8.2.10}}{=} \int_a^c f(x) \, dx + \int_c^b f(x) \, dx \end{aligned}$$

Nun nimmt man das Supremum über alle Zerlegungen auf der rechten Seite, so findet man wiederum

$$\int_{\underline{a}}^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx.$$

Nutzt man wiederum die Riemann-integrierbarkeit auf $[a, b]$ so hat man

$$\int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx.$$

Dies zusammen mit Schritt 1 liefert

$$\int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx.$$

Wir können nun an dieser Stelle unseren Integralbegriff noch erweitern und das sogenannte *orientierte Riemann'sche Integral* einführen. Das bedeutet wir können dem Ausdruck $\int_a^b f(x) \, dx$ auch einen Sinn geben falls $a \nless b$.

Definition 8.2.13 — ORIENTIERTES RIEMANN-INTEGRAL

Es sei $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar und $c, d \in [a, b]$, $c < d$. Dann definieren wir

$$\int_d^c f(x) \, dx := - \int_c^d f(x) \, dx \quad (a)$$

und außerdem

$$\int_c^c f(x) \, dx := 0. \quad (457)$$

^aDass $\int_c^d f(x) \, dx$ wohldefiniert ist liegt an Bemerkung 8.2.11.

Auch dieses orientierte Integral hat die vorher diskutierte Teilungseigenschaft.

Proposition 8.2.14 — ALLGEMEINE TEILUNGSFORMEL

Es sei $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar und $c, d, \theta \in [a, b]$ beliebig. Dann gilt

$$\int_c^d f(x) \, dx = \int_c^\theta f(x) \, dx + \int_\theta^d f(x) \, dx.$$

Beweis. Sind zwei der Parameter gleich, ist die Behauptung klar, da dann auf jeden Fall nach (457) eines der Integrale gleich 0 ist. Daher nehmen wir im Folgenden an, dass c, θ, d verschieden sind.

FALL 1. $c < \theta < d$.

Falls $c < \theta < d$, so gehen wir wie folgt vor: f ist nach Bemerkung 8.2.11 auf $[c, d]$ Riemann-integrierbar und es gilt dann (weil $\theta \in [c, d]$) nach der Teilungsformel (Proposition 8.2.12)

$$\int_c^d f(x) \, dx = \int_c^\theta f(x) \, dx + \int_\theta^d f(x) \, dx$$

FALL 2. $c < d < \theta$.

Da $[c, \theta] \subseteq [a, b]$ ist f nach Bemerkung 8.2.11 auf $[c, \theta]$ Riemann-integrierbar. Nach der Teilungsformel (Proposition 8.2.12) gilt

$$\int_c^\theta f(x) \, dx = \int_c^d f(x) \, dx + \int_d^\theta f(x) \, dx.$$

Nutzen wir nun, dass nach Definition 8.2.13 $\int_d^\theta f(x) \, dx = - \int_\theta^d f(x) \, dx$ so folgern wir

$$\int_c^\theta f(x) \, dx = \int_c^d f(x) \, dx - \int_\theta^d f(x) \, dx.$$

Addieren wir $\int_{\theta}^d f(x) \, dx$ auf beiden Seiten, so folgt

$$\int_c^{\theta} f(x) \, dx + \int_{\theta}^d f(x) \, dx.$$

FALL 3. $d < c < \theta$.

Da $[d, \theta] \subseteq [a, b]$ ist f auf $[d, \theta]$ Riemann-integrierbar (vgl. Bemerkung 8.2.11) und es gilt nach der Teilungsformel (Proposition 8.2.12)

$$\int_d^{\theta} f(x) \, dx = \int_d^c f(x) \, dx + \int_c^{\theta} f(x) \, dx.$$

Nutzt man wieder Definition 8.2.13 so erhält man

$$-\int_{\theta}^d f(x) \, dx = -\int_c^d f(x) \, dx + \int_c^{\theta} f(x) \, dx.$$

Addiert man beide Terme mit negativem Vorzeichen auf die andere Seite so erhält man

$$\int_c^d f(x) \, dx = \int_c^{\theta} f(x) \, dx + \int_{\theta}^d f(x) \, dx.$$

Die verbleibenden Fälle

FALL 4. $d < \theta < c$,

FALL 5. $\theta < c < d$,

FALL 6. $\theta < d < c$,

sind den Lesenden als Übungsaufgabe überlassen.

Man überzeugt sich auch leicht davon, dass für das orientierte Integral auch die Linearität gilt. Eine Eigenschaft, die das orientierte Integral allerdings *nicht* besitzt, ist die Monotonie!

In der Tat: Es gilt zwar $1 \leq 2$, aber $\int_4^3 1 \, dx > \int_4^3 2 \, dx$, denn

$$\int_4^3 1 \, dx = -\int_3^4 1 \, dx = -1(4-3) = -1 \quad \text{und} \quad \int_4^3 2 \, dx = -\int_3^4 2 \, dx = -2(4-3) = -2.$$

8.3 Der Fundamentalsatz der Analysis

Nun kommen wir zu einem sehr wichtigen Satz, dem sogenannten *Fundamentalsatz der Analysis*. Der Grund warum man den Satz einen *Fundamentalsatz* nennt ist, dass er alle wichtigen Konzepte der Analysis von Funktionen – Stetigkeit, Differentialrechnung und Integralrechnung – miteinander verbindet.

Satz 8.3.1 — FUNDAMENTALSATZ DER ANALYSIS

Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Definiere

$$F : [a, b] \rightarrow \mathbb{R}, \quad F(x) := \int_a^x f(t) \, dt. \quad (458)$$

Dann ist F auf $[a, b]$ differenzierbar und es gilt

$$F'(x) = f(x) \quad \forall x \in [a, b].$$

Mit anderen Worten: Das Differenzieren macht das Integrieren in (458) *rückgängig* (jedenfalls sofern es sich beim Integranden um eine stetige Funktion handelt). Dieser Satz hat große Auswirkungen darauf, wie wir Integrale in der Praxis ausrechnen.

Bemerkung 8.3.2

Damit die in (458) gebildete Funktion F wohldefiniert ist, muss man überprüfen, dass für alle $x \in [a, b]$ die Funktion f auf $[a, x]$ Riemann-integrierbar ist. Das kann man zum Beispiel mit der Stetigkeit von f auf $[a, x]$ begründen, vgl. Proposition 8.1.22.

Beachte: Für $x = a$ ist qua Definition 8.2.13

$$F(a) = \int_a^a f(x) \, dx = 0.$$

Für den Beweis des Fundamentalsatzes benötigen wir den sogenannten *Mittelwertsatz der Integralrechnung* dem wir den nächsten Abschnitt widmen.

Mittelwertsatz der Integralrechnung**Proposition 8.3.3** — ERSTER MITTELWERTSATZ DER INTEGRALRECHNUNG

Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und $c, d \in [a, b]$. Dann gibt es ein $\xi \in [\min\{c, d\}, \max\{c, d\}]$ mit

$$\int_c^d f(x) \, dx = f(\xi)(d - c). \quad (a)$$

^aEine geometrische Interpretation liefert GEOGEBRA

Beweis. Ist $c = d$ so ist nichts zu zeigen (dann sind nämlich beide Seiten gleich 0, vgl. Definition 8.2.13).

FALL 1. $c < d$.

Definiere

$$y := \frac{1}{d-c} \int_c^d f(x) \, dx. \quad (459)$$

Da f auf $[c, d]$ stetig ist, nimmt f auf $[c, d]$ außerdem nach dem Satz von Weierstraß (Satz 6.2.12) ein Maximum und Minimum an. Seien $x_1, x_2 \in [c, d]$ so, dass

$$f(x_1) = \min_{x \in [c, d]} f(x) \quad \text{und} \quad f(x_2) = \max_{x \in [c, d]} f(x).$$

Mit anderen Worten gilt

$$f(x_1) \leq f(x) \leq f(x_2) \quad \forall x \in [c, d].$$

Nach Proposition 8.2.5 gilt dann auch

$$\int_c^d f(x_1) \, dx \leq \int_c^d f(x) \, dx \leq \int_c^d f(x_2) \, dx.$$

Da $f(x_1), f(x_2)$ für die Integration über x Konstanten sind, kann man die Formel aus Beispiel 8.1.17 anwenden und erhält

$$f(x_1)(d-c) \leq \int_c^d f(x) \, dx \leq f(x_2)(d-c).$$

Teilt man durch $d-c$ und nutzt $d-c > 0$ so erhält man

$$f(x_1) \leq \frac{1}{d-c} \int_c^d f(x) \, dx \leq f(x_2)$$

Mit der Definition von y in (459) folgern wir dass $y \in [f(x_1), f(x_2)]$. Nach dem Zwischenwertsatz (Satz 6.2.15)^a gibt es dann ein $\xi \in [\min\{x_1, x_2\}, \max\{x_1, x_2\}]$ mit $f(\xi) = y$. Da $x_1, x_2 \in [c, d]$ ist auch $[\min\{x_1, x_2\}, \max\{x_1, x_2\}] \subseteq [c, d]$ und somit gilt klarerweise auch $\xi \in [c, d]$. Ferner hat man

$$f(\xi) = y \stackrel{(459)}{=} \frac{1}{d-c} \int_c^d f(x) \, dx.$$

Multipliziert man mit $d-c$ so folgt die Behauptung.

FALL 2. $c > d$.

Dann gilt nach Definition 8.2.13

$$\int_c^d f(x) \, dx = - \int_d^c f(x) \, dx. \quad (460)$$

Nach Fall 1 gibt es nun ein $\xi \in [d, c] = [\min\{c, d\}, \max\{c, d\}]$ mit

$$\int_d^c f(x) \, dx = f(\xi)(c - d).$$

Zusammen mit (460) ergibt sich

$$\int_c^d f(x) \, dx = - \int_d^c f(x) \, dx = -f(\xi)(c - d) = f(\xi)(d - c),$$

was zu zeigen war.

^aHier wurde die Stetigkeit von f auch benutzt: Der Zwischenwertsatz (Satz 6.2.15) ist nur anwendbar weil f stetig ist

Bemerkung 8.3.4

Auf die Stetigkeit von f kann im obigen Mittelwertsatz (Proposition 8.3.3) nicht verzichtet werden. In der Tat: Betrachte

$$f : [-2, 2] \rightarrow \mathbb{R}, \quad f(x) := \begin{cases} 1 & x \in [-2, 0] \\ -1 & x \in (0, 2] \end{cases} \quad (461)$$

Man kann sich leicht davon überzeugen, dass f

- (i) Riemann-integrierbar ist (z.B. als monotone Funktion, vgl Proposition 8.1.21)
- (ii) Bei $x_0 = 0$ nicht stetig ist (vgl. Beispiel 6.1.15) .

Nun gilt^a

$$\int_{-2}^2 f(x) \, dx = 1 \cdot (0 - (-2)) + (-1) \cdot (2 - 0) = 2 - 2 = 0. \quad (462)$$

Nun gibt es aber kein $\xi \in [-2, 2]$ mit

$$f(\xi)(2 - (-2)) = \int_{-2}^2 f(x) \, dx = 0,$$

denn das würde ja bedeuten, dass $4f(\xi) = 0$, also $f(\xi) = 0$. Dies widerspricht aber (461).

^aDie Formel (462) müsste man wohl noch geeignet begründen. Dazu: Wegen der Teilungseigenschaft können wir schreiben

$$\int_{-2}^2 f(x) \, dx = \int_{-2}^0 f(x) \, dx + \int_0^2 f(x) \, dx.$$

Für das erste Integral ist der Fall klar: Nach Beispiel 8.1.17 gilt

$$\int_{-2}^0 f(x) \, dx = 1 \cdot (0 - (-2)) = 2.$$

Beim zweiten Integral geht es darum, die Funktion

$$f|_{[0,2]} : [0,2] \rightarrow \mathbb{R}, x \mapsto \begin{cases} -1 & x = 0 \\ 1 & x \in (0,2] \end{cases}$$

zu integrieren. Hierzu gehen wir ähnlich vor wie in Beispiel 8.2.7. Betrachte für $n \in \mathbb{N}$ die Zerlegung $Z_n := (0, \frac{1}{n}, 2)$ von $[0,2]$. Dann gilt

$$\overline{S}_{Z_n}(f) = \left(\sup_{[0, \frac{1}{n}]} f \right) \cdot \frac{1}{n} + \left(\sup_{[\frac{1}{n}, 2]} f \right) \left(2 - \frac{1}{n} \right) = 1 \cdot \frac{1}{n} + (-1) \cdot \left(2 - \frac{1}{n} \right) = -2 + \frac{2}{n},$$

$$\underline{S}_{Z_n}(f) = \left(\inf_{[0, \frac{1}{n}]} f \right) \cdot \frac{1}{n} + \left(\inf_{[\frac{1}{n}, 2]} f \right) \left(2 - \frac{1}{n} \right) = (-1) \cdot \frac{1}{n} + (-1) \cdot \left(2 - \frac{1}{n} \right) = -2.$$

Nun gilt nach Lemma 8.1.19

$$\left| \int_0^2 f(x) \, dx - (-2) \right| = \left| \int_0^2 f(x) \, dx - \underline{S}_{Z_n}(f) \right| \leq \overline{S}_{Z_n}(f) - \underline{S}_{Z_n}(f) = \frac{2}{n}$$

Lässt man nun auf beiden Seiten $n \rightarrow \infty$ streben so folgert man

$$\left| \int_0^2 f(x) \, dx - (-2) \right| = 0,$$

spricht: $\int_0^2 f(x) \, dx = -2$.

Es gibt auch noch eine Verallgemeinerung dieses Satzes, der sogenannte *zweite Mittelwertsatz der Integralrechnung*. Wir beweisen ihn hier nicht weil wir ihn vorerst nicht brauchen. Der Beweis nutzt auch dieselben Techniken wie Proposition 8.3.3 und ist daher nicht interessant.

Proposition 8.3.5 — ZWEITER MITTELWERTSATZ DER INTEGRALRECHNUNG

Es seien $f, g : [a, b] \rightarrow \mathbb{R}$ stetig mit

$$g(x) \geq 0 \quad \forall x \in [a, b].$$

Ferner seien $c, d \in [a, b]$. Dann gibt es ein $\xi \in [\min\{c, d\}, \max\{c, d\}]$ mit

$$\int_c^d f(x)g(x) \, dx = f(\xi) \int_c^d g(x) \, dx.$$

Beweis. Übungsaufgabe.

Beweis zum Fundamentalsatzes der Analysis

Nun sind wir endlich in der Lage, den Fundamentalsatz der Analysis zu beweisen. Wir wiederholen an dieser Stelle nochmal die Formulierung:

Wiederholung Satz 8.3.1. FUNDAMENTALSATZ DER ANALYSIS.

Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig. Definiere

$$F : [a, b] \rightarrow \mathbb{R}, \quad F(x) := \int_a^x f(t) \, dt. \quad (463)$$

Dann ist F auf $[a, b]$ differenzierbar und es gilt

$$F'(x) = f(x) \quad \forall x \in [a, b].$$

Beweis. Es sei $x_0 \in [a, b]$ beliebig und F wie in (463) definiert. Klarerweise ist dann x_0 ein Häufungspunkt von $[a, b] \setminus \{x_0\}$. Wir müssen nun zeigen, dass

$$F'(x_0) = \lim_{\substack{x \rightarrow x_0 \\ x \in [a, b] \setminus \{x_0\}}} \frac{F(x) - F(x_0)}{x - x_0} \text{ existiert und mit } f(x_0) \text{ übereinstimmt.}$$

Dazu sei $(x_n)_{n \in \mathbb{N}} \subseteq [a, b] \setminus \{x_0\}$ eine beliebige Folge mit $\lim_{n \rightarrow \infty} x_n = x_0$. Dann gilt für alle $n \in \mathbb{N}$

$$\begin{aligned} \frac{F(x_n) - F(x_0)}{x_n - x_0} &= \frac{1}{x_n - x_0} \left(\int_a^{x_n} f(t) \, dt - \int_a^{x_0} f(t) \, dt \right) \\ &\stackrel{\text{Prop 8.2.14}}{=} \frac{1}{x_n - x_0} \left(\int_a^{x_0} f(t) \, dt + \int_{x_0}^{x_n} f(t) \, dt - \int_a^{x_0} f(t) \, dt \right) \\ &= \frac{1}{x_n - x_0} \int_{x_0}^{x_n} f(t) \, dt. \end{aligned} \quad (464)$$

Nach dem Mittelwertsatz der Integralrechnung (Proposition 8.3.3) gibt es nun für jedes $n \in \mathbb{N}$ ein $\xi_n \in [\min\{x_0, x_n\}, \max\{x_0, x_n\}]$ so, dass

$$\frac{1}{x_n - x_0} \int_{x_0}^{x_n} f(t) \, dt = \frac{1}{x_n - x_0} f(\xi_n)(x_n - x_0) = f(\xi_n). \quad (465)$$

Wie man sich leicht überzeugt gilt $\min\{x_0, x_n\} \xrightarrow{n \rightarrow \infty} x_0$ und $\max\{x_0, x_n\} \xrightarrow{n \rightarrow \infty} x_0$. Da für alle $n \in \mathbb{N}$ gilt $\min\{x_0, x_n\} \leq \xi_n \leq \max\{x_0, x_n\}$ erhalten wir mit dem Sandwichlemma, dass $\lim_{n \rightarrow \infty} \xi_n = x_0$. Wegen der Stetigkeit von f , vgl. Bemerkung 6.2.2 gilt dann auch

$$\lim_{n \rightarrow \infty} f(\xi_n) = f(x_0). \quad (466)$$

Alles in allem haben wir dann

$$\lim_{n \rightarrow \infty} \frac{F(x_n) - F(x_0)}{x_n - x_0} \stackrel{(464)}{=} \lim_{n \rightarrow \infty} \frac{1}{x_n - x_0} \int_{x_0}^{x_n} f(t) \, dt \stackrel{(465)}{=} \lim_{n \rightarrow \infty} f(\xi_n) \stackrel{(466)}{=} f(x_0).$$

Da $(x_n)_{n \in \mathbb{N}} \subseteq [a, b] \setminus \{x_0\}$ beliebig war gilt

$$F'(x_0) = \lim_{\substack{x \rightarrow x_0 \\ x \in [a, b] \setminus \{x_0\}}} \frac{F(x) - F(x_0)}{x - x_0} = f(x_0).$$

Die Behauptung folgt.

13.05.2026

Stammfunktionen und Integrale

Der Fundamentalsatz der Analysis kann benutzt werden um Integrale zu berechnen. Wie das geht klären wir in diesem Kapitel über *Stammfunktionen*.

Definition 8.3.6 — STAMMFUNKTION

Sei $D \subseteq \mathbb{R}$ und $f : D \rightarrow \mathbb{R}$ eine Funktion. Eine Funktion $F : D \rightarrow \mathbb{R}$ heißt *Stammfunktion* von f auf D , falls F auf D differenzierbar ist und

$$F'(x) = f(x) \quad \forall x \in D.$$

Beispiel 8.3.7

Es sei $D = (0, \infty)$ und $f : D \rightarrow \mathbb{R}, x \mapsto \frac{1}{x}$. Dann ist $F : D \rightarrow \mathbb{R}, x \mapsto \log(x)$ eine Stammfunktion von F , denn es gilt nach Beispiel 7.1.11

$$F'(x) = \frac{d}{dx} \log(x) \stackrel{\text{Bsp 7.1.11}}{=} \frac{1}{x} = f(x) \quad \forall x \in (0, \infty).$$

Wir werden später noch eine ausführliche Liste von uns bisher bekannten Stammfunktionen geben. Der Fundamentalsatz der Analysis kann übrigens mithilfe des Begriffes einer Stammfunktion umformuliert werden, wie wir in der nachfolgenden Bemerkung festhalten werden

Bemerkung 8.3.8

Der Fundamentalsatz der Analysis (Satz 8.3.1) besagt, dass jede stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ eine Stammfunktion besitzt. Ferner kann man eine solche Stammfunktion $F : [a, b] \rightarrow \mathbb{R}$ auch explizit angeben, nämlich durch die Vorschrift

$$F(x) := \int_a^x f(x) \, dx. \tag{467}$$

Bemerkung 8.3.9

Allerdings ist die Funktion F aus (467) nicht die einzige Stammfunktion einer gegebenen

stetigen Funktion $f : [a, b] \rightarrow \mathbb{R}$. In der Tat gibt es *unendlich viele Stammfunktionen*, wie wir in der folgenden Proposition sehen werden!

Proposition 8.3.10

Es sei $D \subseteq \mathbb{R}$ und $f : D \rightarrow \mathbb{R}$ gegeben und $F : D \rightarrow \mathbb{R}$ eine Stammfunktion von f . Ferner sei $k \in \mathbb{R}$ eine beliebige Konstante. Definiere $G : D \rightarrow \mathbb{R}$ durch

$$G(x) := F(x) + k$$

Dann ist auch G eine Stammfunktion von f auf D .

Beweis. Wir verifizieren Definition 8.3.6 für G , d.h. wir zeigen

$$G'(x) = f(x) \quad \forall x \in D.$$

Dazu berechnen wir

$$G'(x) = \frac{d}{dx} G(x) = \frac{d}{dx} (F(x) + k) \stackrel{\text{Prop 7.1.1}}{=} \frac{d}{dx} F(x) + \frac{d}{dx} k \stackrel{\text{Bsp 7.0.2}}{=} F'(x) + 0 = F'(x).$$

Da F bereits eine Stammfunktion von f ist, gilt $F'(x) = f(x)$ für alle $x \in D$. Verwenden wir dieses Erkenntnis in der obigen Gleichung so folgt

$$G'(x) = f(x) \quad \forall x \in D.$$

Man sieht: Durch das Addieren einer Konstante zu einer Stammfunktion kann man stets eine neue Stammfunktion konstruieren.

Ist die Definitionsmenge D ein Intervall, so ist das Addieren einer Konstante auch die *einzige Möglichkeit* eine weitere Stammfunktion zu erzeugen. Das ist eine Folgerung aus dem *Konstanzsatz* (Proposition 7.3.8).

Proposition 8.3.11

Es sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ eine Funktion. Ferner seien $F, G : I \rightarrow \mathbb{R}$ zwei Stammfunktionen von f . Dann ist $G - F$ konstant. Insbesondere gibt es ein $k \in \mathbb{R}$ mit

$$G(x) = F(x) + k \quad \forall x \in I. \tag{468}$$

Beweis. Wir benutzen den Konstanzsatz für $G - F$. Dazu bilden wir für $x \in I$

$$(G - F)'(x) \stackrel{\text{Prop 7.1.1}}{=} G'(x) - F'(x) = f(x) - f(x) = 0.$$

Die Voraussetzungen für den Konstanzsatz sind also erfüllt und wir schließen, dass $G - F$

konstant ist. Deswegen gibt es ein $k \in \mathbb{R}$ so, dass

$$G(x) - F(x) = k \quad \forall x \in I.$$

Addieren wir $F(x)$ auf beiden Seiten so folgt auch (468).

Man sagt, Stammfunktionen auf Intervallen sind *bis auf eine additive Konstante eindeutig*. Nun endlich können wir klären, wie Stammfunktionen benutzt werden können um Integrale zu berechnen. Das Werkzeug dafür nennen wir *Hauptsatz der Differential- und Integralrechnung*, kurz: HDI.

Satz 8.3.12 — HAUPTSATZ DER DIFFERENTIAL- UND INTEGRALRECHNUNG

Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion und F eine beliebige Stammfunktion von f auf $[a, b]$. Dann gilt

$$\int_a^b f(t) \, dt = F(b) - F(a).$$

Beweis. Nach dem Fundamentalsatz der Analysis (Satz 8.3.1) ist

$$G : [a, b] \rightarrow \mathbb{R}, \quad G(x) := \int_a^x f(t) \, dt \quad (469)$$

eine weitere Stammfunktion von f . Nach Proposition 8.3.11 können sich nun die beiden Stammfunktionen G und F nur um eine additive Konstante unterscheiden. Das bedeutet es gibt ein $k \in \mathbb{R}$ mit

$$G(x) = F(x) + k \quad \forall x \in [a, b].$$

Das heißt auch, dass $F(x) = G(x) - k$ für alle $x \in [a, b]$. Man erhält

$$\begin{aligned} F(b) - F(a) &= (G(b) - k) - (G(a) - k) \\ &= G(b) - k - G(a) + k = G(b) - G(a) \stackrel{(469)}{=} \int_a^b f(t) \, dt - \int_a^a f(t) \, dt. \end{aligned}$$

Da $\int_a^a f(t) \, dt = 0$ folgt die Behauptung.

Mit anderen Worten: Wer eine Stammfunktion einer gegebenen Funktion kennt, kann damit Integrale berechnen.

Bemerkung 8.3.1. Der HDI lässt sich auch für Integration über Teilintervalle oder mit umgekehrt angeordneten Integrationsgrenzen anwenden:

Verallgemeinerung. Ist $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion mit Stammfunktion F :

$[a, b] \rightarrow \mathbb{R}$ sowie $c, d \in [a, b]$ beliebig, so gilt

$$\int_c^d f(t) \, dt = F(d) - F(c).$$

Beispiel 8.3.13

Wir berechnen das Integral

$$\int_2^3 \frac{1}{x} \, dx$$

Da nach Beispiel 8.3.7 der natürliche Logarithmus eine Stammfunktion von $x \mapsto \frac{1}{x}$ ist gilt nach dem HDI (Satz 8.3.12)

$$\int_2^3 \frac{1}{x} \, dx = \log(3) - \log(2) \stackrel{\text{Kor6.3.5}}{=} \log\left(\frac{3}{2}\right).$$

Auf diese Weise können wir Integrale über sämtliche Funktionen bestimmen, deren Stammfunktionen wir kennen. Wir müssen also mehr Stammfunktionen kennenlernen... Bevor wir dies tun führen wir aber noch eine Notation

Definition 8.3.14 — AUSWERTUNGSKLAMMERN

Für eine Funktion $F : [a, b] \rightarrow \mathbb{R}$ definieren wir

$$\left[F(x) \right]_{x=a}^{x=b} := F(b) - F(a).$$

Insbesondere kann man den HDI (Satz 8.3.12) kürzer formulieren: Man hat für eine stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ mit Stammfunktion $F : [a, b] \rightarrow \mathbb{R}$, dass

$$\int_a^b f(x) \, dx = \left[F(x) \right]_{x=a}^{x=b}.$$

Hier fassen wir nun einmal zusammen für welche Funktionen wir im Laufe dieses Skriptes Stammfunktionen kennengelernt haben.

Proposition 8.3.15 — STAMMFUNKTIONEN BEKANNTER FUNKTIONEN

Es sei $c \in \mathbb{R}$.

- (i) Ist $x_0 \in \mathbb{R}$ und $\sum_{k \geq 0} a_k (x - x_0)^k$ eine Potenzreihe mit Konvergenzradius $R > 0$, so ist $\sum_{k \geq 0} \frac{1}{k+1} a_k (x - x_0)^{k+1} + c$ eine Potenzreihe mit demselben Konvergenzradius. Ist

$$f : (x_0 - R, x_0 + R) \rightarrow \mathbb{R}, x \mapsto \sum_{k=0}^{\infty} a_k x^k$$

$$\text{und } F_c : (x_0 - R, x_0 + R) \rightarrow \mathbb{R}, x \mapsto \sum_{k=0}^{\infty} \frac{1}{k+1} a_k (x - x_0)^{k+1} + c$$

so ist F_c eine Stammfunktion von f .

- (ii) Ist $\alpha \in \mathbb{R} \setminus \{-1\}$ und $f : (0, \infty) \rightarrow \mathbb{R}, x \mapsto x^\alpha$, so ist $F_c : (0, \infty) \rightarrow \mathbb{R}, x \mapsto \frac{1}{\alpha+1} x^{\alpha+1} + c$ eine Stammfunktion von f .
- (iii) Falls $\alpha \in \mathbb{Z} \setminus \{-1\}$ und ist $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto x^\alpha$, so ist $F_c : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto \frac{1}{\alpha+1} x^{\alpha+1} + c$ eine Stammfunktion von f .
- (iv) Für $\alpha = -1$ gilt eine andere Formel: Ist $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto x^{-1}$, so ist $F_c : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto \log(|x|) + c$ eine Stammfunktion von f .
- (v) Ist $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \beta^x$ für ein $\beta \in (0, \infty) \setminus \{1\}$, so ist $F_c : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \frac{1}{\log(\beta)} \beta^x$ eine Stammfunktion.
- (vi) Ist $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \exp(x)$, so ist $F_c : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \exp(x) + c$ eine Stammfunktion von f .
- (vii) Ist $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \sin(x)$, so ist $F_c : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto -\cos(x) + c$ eine Stammfunktion von f .
- (viii) Ist $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \cos(x)$, so ist $F_c : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \sin(x) + c$ eine Stammfunktion von f .
- (ix) Ist $f : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}, x \mapsto \frac{1}{\cos^2(x)}$, so ist $F_c : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \tan(x) + c$ eine Stammfunktion von f .
- (x) Ist $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \frac{1}{1+x^2}$ so ist $F_c : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \arctan(x) + c$ eine Stammfunktion von f .
- (xi) Ist $f : (-1, 1) \rightarrow \mathbb{R}, x \mapsto \frac{1}{\sqrt{1-x^2}}$ so ist $F_c : (-1, 1) \rightarrow \mathbb{R}, x \mapsto \arcsin(x)$ eine Stammfunktion von f .

Beweis. Alle Beweise funktionieren stur nach Definition 8.3.6.

Zu (i). Dass die beiden Potenzreihen denselben Konvergenzradius haben folgt wie in Proposition 5.2.6. Außerdem gilt nach Proposition 7.0.10 für alle $x \in B_R(x_0) \cap \mathbb{R} = (x_0 - R, x_0 + R)$, dass

$$\frac{d}{dx} \left(\sum_{k=0}^{\infty} \frac{1}{k+1} a_k (x - x_0)^{k+1} \right) = \sum_{k=0}^{\infty} a_k (x - x_0)^k = f(x).$$

Die Behauptung folgt mit Proposition 8.3.10.

Zu (ii). Für $x \in (0, \infty)$ gilt

$$\frac{d}{dx} \left(\frac{1}{\alpha + 1} x^{\alpha+1} \right) \stackrel{\text{Prop 7.1.12}}{=} \frac{1}{\alpha + 1} (\alpha + 1) x^{\alpha+1-1} = x^\alpha = f(x).$$

Die Behauptung folgt mit Proposition 8.3.10.

Zu (iii). Ist den Lesenden als Übungsaufgabe überlassen.

Zu (iv). Für $x \in (0, \infty)$ haben wir schon in Beispiel 8.3.7 nachgerechnet, dass

$$\frac{d}{dx} \log(|x|) = \frac{d}{dx} \log(x) = \frac{1}{x} = x^{-1} = f(x).$$

Für $x \in (-\infty, 0)$ gilt

$$\frac{d}{dx} \log(|x|) = \frac{d}{dx} \log(-x) \stackrel{\text{Prop 7.1.5}}{=} \frac{1}{-x} (-1) = \frac{1}{x} = x^{-1}.$$

Die Behauptung folgt mit Proposition 8.3.10.

Zu (v). Folgt direkt mit Beispiel 7.1.7.

Zu (vi). Für $x \in \mathbb{R}$ gilt

$$\frac{d}{dx} \exp(x) \stackrel{\text{Prop 7.0.12}}{=} \exp(x) = f(x).$$

Die Behauptung folgt mit Proposition 8.3.10.

Zu (vii). Für $x \in \mathbb{R}$ gilt

$$\frac{d}{dx} (-\cos(x)) \stackrel{\text{Prop 7.0.12}}{=} -(-\sin(x)) = \sin(x) = f(x).$$

Die Behauptung folgt mit Proposition 8.3.10.

Zu (viii). Ist den Lesenden als Übungsaufgabe überlassen.

Zu (ix). Ist den Lesenden als Übungsaufgabe überlassen.

Zu (x). Für $x \in \mathbb{R}$ gilt

$$\frac{d}{dx} \arctan(x) \stackrel{\text{Prop 7.3.20}}{=} \frac{1}{1+x^2} = f(x).$$

Die Behauptung folgt mit Proposition 8.3.10.

Zu (xi). Für $x \in (-1, 1)$ gilt

$$\frac{d}{dx} \arcsin(x) \stackrel{\text{Prop 7.3.20}}{=} \frac{1}{\sqrt{1-x^2}} = f(x).$$

Die Behauptung folgt mit Proposition 8.3.10.

Oft gelingt es die zu integrierenden Ausdrücke mithilfe von Umformungen und den Rechenregeln aus Kapitel 8.2 auf die obige Gestalt zurückzubringen.

Beispiel 8.3.2. Wir berechnen

$$\int_{-1}^1 \frac{x^2}{1+x^2} dx.$$

Dazu:

$$\begin{aligned} \int_{-1}^1 \frac{x^2}{1+x^2} dx &= \int_{-1}^1 \frac{1+x^2-1}{1+x^2} dx = \int_{-1}^1 \left(1 - \frac{1}{1+x^2}\right) dx = \int_{-1}^1 1 dx - \int_{-1}^1 \frac{1}{1+x^2} dx \\ &= \left[x\right]_{x=-1}^{x=1} - \left[\arctan(x)\right]_{x=-1}^{x=1} = (1 - (-1)) - (\arctan(1) - \arctan(-1)) \end{aligned}$$

Da nach Proposition 5.5.17^a $\arctan(1) = \frac{\pi}{4}$ und $\arctan(-1) = -\frac{\pi}{4}$ gilt erhalten wir

$$\int_{-1}^1 \frac{x^2}{1+x^2} dx = 2 - 2\frac{\pi}{4} = 2 - \frac{\pi}{2}.$$

^aDiese Proposition muss man hier so benutzen: Sie besagt, dass $\cos(\frac{\pi}{4}) = \sin(\frac{\pi}{4})$, woraus wir schließen, dass $\tan(\frac{\pi}{4}) = 1$. Durch Anwenden der Umkehrfunktion ergibt sich dann $\arctan(1) = \frac{\pi}{4}$.

8.4 Substitution und partielle Integration

Im Folgenden möchten wir unsere Integrationstechniken noch weiter verfeinern. Hierfür lernen wir die *Substitutionsregel* und die *partielle Integrationsregel*. Diese sind direkte Folgerungen aus dem HDI (Satz 8.3.12). Wenn man sie aber richtig beherrscht kann man damit sehr viele Integrale berechnen.

Substitutionen

Satz 8.4.1 — SUBSTITUTIONSREGEL

Es sei $f : D \rightarrow \mathbb{R}$ eine stetige Funktion und $\phi : [a, b] \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion mit $\phi([a, b]) \subseteq D$. Dann gilt

$$\int_a^b f(\phi(x)) \cdot \phi'(x) \, dx = \int_{\phi(a)}^{\phi(b)} f(t) \, dt. \quad (470)$$

Beweis. Beachte: Da f stetig ist und ϕ stetig differenzierbar ist, ist der Integrand auf der linken Seite von (470), gegeben durch $(f \circ \phi) \cdot \phi'$, als Komposition stetiger Funktionen auf $[a, b]$ stetig. Nach Proposition 6.2.18 ist $I := \phi([a, b])$ ein abgeschlossenes beschränktes Intervall mit $\phi(a), \phi(b) \in I$. Es sei $F : I \rightarrow \mathbb{R}$ eine Stammfunktion von f .^a Dann gilt nach der Kettenregel für alle $x \in [a, b]$

$$\frac{d}{dx}(F \circ \phi)(x) = F'(\phi(x))\phi'(x) = f(\phi(x))\phi'(x).$$

Das bedeutet: Der Integrand auf der linken Seite von (470) besitzt gerade $F \circ \phi$ als Stammfunktion. Es folgt aus dem HDI (Satz 8.3.12), dass

$$\begin{aligned} \int_a^b f(\phi(x))\phi'(x) \, dx &= \left[(F \circ \phi)(x) \right]_{x=a}^{x=b} = (F \circ \phi)(b) - (F \circ \phi)(a) = F(\phi(b)) - F(\phi(a)) \\ &\stackrel{\text{Def 8.3.14}}{=} \left[F(t) \right]_{t=\phi(a)}^{t=\phi(b)} \stackrel{\text{Satz 8.3.12}}{=} \int_{\phi(a)}^{\phi(b)} f(t) \, dt, \end{aligned}$$

wobei man für die letzte Gleichheit die Aussage des HDI (Satz 8.3.12, genauer eigentlich Bemerkung 8.3.1 mit $c = \phi(a), d = \phi(b)$) rückwärts lesen muss. Die Behauptung ist gezeigt.

^aEine solche existiert dank Bemerkung 8.3.9.

Die Substitutionsformel vereinfacht viele Integrale — Betrachtet man eine geschickt gewählte Funktion ϕ , so ist das Integral auf der rechten Seite von (470) einfacher als das ursprüngliche Integral auf der linken Seite. Hier zunächst ein Beispiel.

Beispiel 8.4.2

Wir berechnen mithilfe einer Substitution

$$\int_0^{\frac{\pi}{2}} \frac{\cos(x)}{2 + \sin(x)} \, dx. \quad (471)$$

Definiere dazu $\phi : [0, \frac{\pi}{2}] \rightarrow \mathbb{R}$ durch $\phi(x) := 2 + \sin(x)$. Dann ist ϕ stetig differenzierbar mit $\phi'(x) = 0 + \cos(x) = \cos(x)$ für alle $x \in [0, \frac{\pi}{2}]$. Das Integral in (471) ist dann gerade

$$\int_0^{\frac{\pi}{2}} \frac{1}{\phi(x)} \phi'(x) \, dx.$$

Nutzen wir die Substitutionsregel (Satz 8.4.1) mit $f : (0, \infty) \rightarrow \mathbb{R}$, $f(t) = \frac{1}{t}$ so erhalten wir

$$\int_0^{\frac{\pi}{2}} \frac{\cos(x)}{2 + \sin(x)} dx = \int_0^{\frac{\pi}{2}} \frac{1}{\phi(x)} \phi'(x) dx = \int_0^{\frac{\pi}{2}} f(\phi(x)) \phi'(x) dx \stackrel{\text{Satz 8.4.1}}{=} \int_{\phi(0)}^{\phi(\frac{\pi}{2})} f(t) dt.$$

Da $\phi(0) = 2 + \sin(0) = 2$ und $\phi(1) = 2 + \sin(\frac{\pi}{2}) = 2 + 1 = 3$ erhalten wir

$$\int_0^{\frac{\pi}{2}} \frac{\cos(x)}{2 + \sin(x)} dx = \int_2^3 f(t) dt = \int_2^3 \frac{1}{t} dt \stackrel{\text{Bsp 8.3.13}}{=} \left[\log(t) \right]_{t=2}^{t=3} = \log(3) - \log(2) = \log\left(\frac{3}{2}\right).$$

Um sich die Substitutionsformel besser merken zu können, notiert man eine Substitution oft wie folgt

$$\int_a^b f(\phi(x)) \cdot \phi'(x) dx \stackrel{\substack{\text{Subst.: } t = \phi(x) \\ \frac{dt}{dx} = \phi'(x) \\ dt = \phi'(x) dx}}{=} \int_{\phi(a)}^{\phi(b)} f(t) dt.$$

Man stellt sich daher vor, dass man $\phi(x)$ mit t ersetzt und dt dabei den Platz von $\phi'(x) dx$ einnimmt (so wie es in der dritten Zeile unter dem obigen Gleichheitszeichen steht). Die zweite Zeile unter dem Gleichheitszeichen ist eine *Hilfszeile*, in der man den Ausdruck für t (der von x abhängt) nach x ableiten muss. Farblich hervorgehoben:

$$\int_a^b f(\phi(x)) \cdot \phi'(x) dx \stackrel{\substack{\text{Subst.: } t = \phi(x) \\ \frac{dt}{dx} = \phi'(x) \\ dt = \phi'(x) dx}}{=} \int_{\phi(a)}^{\phi(b)} f(t) dt.$$

Bei dieser Prozedur sollte man allerdings nicht vergessen, die Integralgrenzen anzupassen!

Beispiel 8.4.3

Wir berechnen

$$\int_0^2 x e^{x^2} dx.$$

Dazu machen wir erst eine kleine Umformung und verwenden danach die Substitutionsregel.

$$\int_0^2 x e^{x^2} dx = \frac{1}{2} \int_0^2 e^{x^2} (2x) dx \stackrel{\substack{\text{Subst.: } t = x^2 \\ \frac{dt}{dx} = 2x \\ dt = 2x dx}}{=} \frac{1}{2} \int_{0^2}^{2^2} e^t dt = \frac{1}{2} \left[e^t \right]_{t=0}^{t=4} = \frac{e^4 - e^0}{2} = \frac{e^4 - 1}{2}.$$

Manchmal braucht man ein geübtes Auge um zu sehen, was man wie genau substituieren soll. Außerdem kann es durchaus hilfreich sein, mehr als einmal zu substituieren.

Beispiel 8.4.4

Wir berechnen

$$\int_{-1}^1 \frac{x^5}{\sqrt{2-x^3}} dx.$$

Dazu:

$$\int_{-1}^1 \frac{x^5}{\sqrt{2-x^3}} dx = \int_{-1}^1 \frac{x^3 \cdot x^2}{\sqrt{2-x^3}} dx = \frac{1}{3} \int_{-1}^1 \frac{x^3 \cdot (3x^2)}{\sqrt{2-x^3}} dx \stackrel{\substack{\text{Subst.: } t = x^3 \\ \frac{dt}{dx} = 3x^2 \\ dt = 3x^2 dx}}{=} \frac{1}{3} \int_{(-1)^3}^{1^3} \frac{t}{\sqrt{2-t}} dt. \quad (472)$$

Da $(-1)^3 = -1$ und $1^3 = 1$ ist nun noch $\int_{-1}^1 \frac{t}{\sqrt{2-t}} dt$ zu berechnen. Dazu

$$\begin{aligned} \int_{-1}^1 \frac{t}{\sqrt{2-t}} dt &= \int_{-1}^1 \frac{t-2+2}{\sqrt{2-t}} dt = \int_{-1}^1 \frac{t-2}{\sqrt{2-t}} dt + \int_{-1}^1 \frac{2}{\sqrt{2-t}} dt \\ &= - \int_{-1}^1 \frac{2-t}{\sqrt{2-t}} dt + 2 \int_{-1}^1 \frac{1}{\sqrt{2-t}} dt = - \int_{-1}^1 \sqrt{2-t} dt + 2 \int_{-1}^1 \frac{1}{\sqrt{2-t}} dt \end{aligned}$$

$$\begin{aligned} &\stackrel{\substack{\text{in beiden Integralen} \\ \text{Subst.: } u = 2-t \\ \frac{du}{dt} = -1 \\ du = -dt}}{=} \int_3^1 \sqrt{u} du - 2 \int_3^1 \frac{1}{\sqrt{u}} du = \int_3^1 u^{\frac{1}{2}} du - 2 \int_3^1 u^{-\frac{1}{2}} du \\ &= \left[\frac{1}{1+\frac{1}{2}} u^{1+\frac{1}{2}} \right]_{u=3}^{u=1} - 2 \left[\frac{1}{1-\frac{1}{2}} u^{1-\frac{1}{2}} \right]_{u=3}^{u=1} = \frac{2}{3} (1^{\frac{3}{2}} - 3^{\frac{3}{2}}) - 4 (1^{\frac{1}{2}} - 3^{\frac{1}{2}}). \end{aligned}$$

Zusammen mit (472) ergibt sich dann

$$\int_{-1}^1 \frac{x^5}{\sqrt{2-x^3}} dx = \frac{1}{3} \left(\frac{2}{3} (1 - 3^{\frac{3}{2}}) - 4 (1 - 3^{\frac{1}{2}}) \right)$$

Vor der Substitution muss auch manchmal umgeformt werden

Beispiel 8.4.5

Wir berechnen

$$\int_{-\pi}^{\pi} \cos^2(x) dx.$$

Dazu: Die Formel des Quadrates über den doppelten Winkel (vgl. Bemerkung 5.5.6) besagt, dass $\cos^2(x) = \frac{1+\cos(2x)}{2}$ für alle $x \in \mathbb{R}$. Dann gilt

$$\begin{aligned} \int_{-\pi}^{\pi} \cos^2(x) dx &= \int_{-\pi}^{\pi} \frac{1+\cos(2x)}{2} dx = \frac{1}{2} \int_{-\pi}^{\pi} (1+\cos(2x)) dx \\ &= \frac{1}{2} \left(\int_{-\pi}^{\pi} 1 dx + \int_{-\pi}^{\pi} \cos(2x) dx \right) = \frac{1}{2} \left(2\pi + \int_{-\pi}^{\pi} \cos(2x) dx \right) \\ &\stackrel{\substack{\text{Subst.: } t = 2x \\ \frac{dt}{dx} = 2 \\ \frac{1}{2} dt = dx}}{=} \frac{1}{2} \left(2\pi + \int_{-2\pi}^{2\pi} \cos(t) \frac{1}{2} dt \right) \\ &= \frac{1}{2} \left(2\pi + \frac{1}{2} \left[\sin(t) \right]_{t=-2\pi}^{t=2\pi} \right) = \frac{1}{2} (2\pi) = \pi. \end{aligned}$$

Partielle Integration

Eine weitere hilfreiche Integrationstechnik ist die *Partielle Integration*. Hierbei kann man ein Integral über ein Produkt von Funktionen umformen.

Satz 8.4.6 — PARTIELLE INTEGRATION

Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion und $g : [a, b] \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion. Ferner sei $F : [a, b] \rightarrow \mathbb{R}$ eine Stammfunktion von f . Dann gilt

$$\int_a^b f(x) \cdot g(x) \, dx = \left[F(x)g(x) \right]_{x=a}^{x=b} - \int_a^b F(x) \cdot g'(x) \, dx. \quad (473)$$

Beweis. Nach den Voraussetzungen ist $fg + Fg' : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion. [Beachte hierbei: F ist auf $[a, b]$ stetig weil es auf $[a, b]$ sogar differenzierbar ist, vgl. Proposition 7.0.7].

ZWISCHENBEHAUPTUNG. Es ist Fg eine Stammfunktion von $fg + Fg'$.

Dazu: Es gilt für alle $x \in [a, b]$ nach der Produktregel

$$\frac{d}{dx} F(x)g(x) = F'(x)g(x) + F(x)g'(x) \stackrel{\text{Def 8.3.6}}{=} f(x)g(x) + F(x)g'(x).$$

Somit ist Fg eine Stammfunktion von $fg + Fg'$ auf $[a, b]$.

ZWISCHENBEHAUPTUNG ENDE.

Der HDI (Satz 8.3.12) impliziert dann

$$\int_a^b (fg + Fg')(x) \, dx = \left[(Fg)(x) \right]_{x=a}^{x=b}.$$

Anders ausgedrückt

$$\int_a^b (f(x)g(x) + F(x)g'(x)) \, dx = \left[F(x)g(x) \right]_{x=a}^{x=b}.$$

Nutzt man nun noch die Linearität aus Proposition 8.2.4 auf der linken Seite so ergibt sich

$$\int_a^b f(x)g(x) \, dx + \int_a^b F(x)g'(x) \, dx = \left[F(x)g(x) \right]_{x=a}^{x=b}.$$

Zieht man dann noch $\int_a^b F(x)g'(x) \, dx$ auf beiden Seiten ab so folgt die Behauptung.

Die partielle Integration ist nur hilfreich, wenn das Integral auf der rechten Seite von (473) leichter zu lösen ist als das Integral auf der linken Seite von (473). Hierzu ist es wichtig, die

Rollen der Funktion f (die Funktion die *aufgeleitet* wird) und g (die Funktion die abgeleitet wird) richtig zuzuweisen.

Beispiel 8.4.7

Wir berechnen

$$\int_0^{\pi} x \sin(x) \, dx.$$

Dazu definiere $f : [0, \pi] \rightarrow \mathbb{R}, x \mapsto \sin(x)$ und $g : [0, \pi] \rightarrow \mathbb{R}, x \mapsto x$. Eine Stammfunktion von f ist nach Proposition 8.3.15 gegeben durch $F : [0, \pi] \rightarrow \mathbb{R}, x \mapsto -\cos(x)$. Die partielle Integration 8.4.6 liefert

$$\begin{aligned} \int_0^{\pi} x \sin(x) \, dx &= \int_0^{\pi} \sin(x) \cdot x \, dx = \left[(-\cos(x)) \cdot x \right]_{x=0}^{x=\pi} - \int_0^{\pi} (-\cos(x)) \cdot 1 \, dx \\ &= \left[-x \cos x \right]_{x=0}^{x=\pi} + \int_0^{\pi} \cos(x) \, dx \\ &= -\pi \cos(\pi) + 0 \cos(0) + \int_0^{\pi} \cos(x) \, dx = \pi + \int_0^{\pi} \cos(x) \, dx \\ &= \pi + \left[\sin(x) \right]_{x=0}^{x=\pi} = \pi + \sin(\pi) - \sin(0) = \pi. \end{aligned}$$

In der Theorie hätte man auch die Rollen von f und g tauschen können, d.h. $f : [0, \pi] \rightarrow \mathbb{R}, x \mapsto x$ und $g : [0, \pi] \rightarrow \mathbb{R}, x \mapsto \sin(x)$ wählen können. Dann wäre die partielle Integration wie folgt abgelaufen

$$\int_0^{\pi} x \sin(x) \, dx = \left[\frac{x^2}{2} \sin(x) \right]_{x=0}^{x=\pi} - \int_0^{\pi} \frac{x^2}{2} \cos(x) \, dx.$$

Diese Umformung wäre nicht zielführend gewesen, da $\int_0^{\pi} \frac{x^2}{2} \cos(x) \, dx$ nicht leichter zu lösen ist als das ursprüngliche Integral.

Man sieht: Man muss sich richtig entscheiden, welche Funktion man bei der partiellen Integration *aufleitet* und welche man *ableitet*. Eine Möglichkeit ist, die getroffene Wahl mit Pfeilen zu kennzeichnen. Das heie in (473)

$$\int_a^b \overset{\uparrow}{f}(x) \overset{\downarrow}{g}(x) \, dx \stackrel{\text{PI}}{=} \left[F(x)g(x) \right]_{x=a}^{x=b} - \int_a^b F(x)g'(x) \, dx.$$

Beispiel 8.4.8

Wir berechnen

$$\int_1^2 x \log(x) \, dx$$

Dazu verfahren wir wie folgt:

$$\begin{aligned} \int_1^2 \overset{\uparrow}{x} \cdot \overset{\downarrow}{\log(x)} \, dx &\stackrel{\text{PI}}{=} \left[\frac{x^2}{2} \log(x) \right]_{x=1}^{x=2} - \int_1^2 \frac{x^2}{2} \cdot \frac{1}{x} \, dx = \left[\frac{x^2}{2} \log(x) \right]_{x=1}^{x=2} - \int_1^2 \frac{x}{2} \, dx \\ &= \left[\frac{x^2}{2} \log(x) \right]_{x=1}^{x=2} - \left[\frac{x^2}{4} \right]_{x=1}^{x=2} = 2 \log(2) - \left(1 - \frac{1}{4} \right) = 2 \log(2) - \frac{3}{4}. \end{aligned}$$

Auch hier stellt sich die Frage, warum wir entschieden haben $x \mapsto x$ aufzuleiten und $x \mapsto \log(x)$ abzuleiten. Naheliegender ist dies vor allem weil $\log$ schwer aufzuleiten ist. Beachte: Im vorigen Beispiel (Beispiel 8.4.7) haben wir $x \mapsto x$ genau die entgegengesetzte Rolle zugewiesen (nämlich die Funktion die abgeleitet wird). Man sieht dass die Zuweisung situationsabhängig ist. Es gibt kein allgemeines Rezept außer die eigene Kreativität.

Manchmal ist es auch nicht so offensichtlich, warum ein Integrand als Produkt aufgefasst werden kann. 18.05.2026

Beispiel 8.4.9

Wir berechnen

$$\int_1^2 \log(x) \, dx$$

Dazu verfahren wir wie folgt:

$$\begin{aligned} \int_1^2 \log(x) \, dx &= \int_1^2 \overset{\uparrow}{1} \cdot \overset{\downarrow}{\log(x)} \, dx \stackrel{\text{PI}}{=} \left[x \log(x) \right]_{x=1}^{x=2} - \int_1^2 x \cdot \frac{1}{x} \, dx = \left[x \log(x) \right]_{x=1}^{x=2} - \int_1^2 1 \, dx \\ &= \left[x \log(x) \right]_{x=1}^{x=2} - \left[x \right]_{x=1}^{x=2} = 2 \log(2) - (2 - 1) = 2 \log(2) - 1. \end{aligned}$$

Findung von Stammfunktionen

Wer die Rechentricks, die wir im Laufe dieses Abschnittes kennengelernt haben richtig anwendet, kann damit auch weitere Stammfunktionen finden, die nicht in Proposition 8.3.15 aufgeführt sind. Hierfür muss integrieren und die obere Integralgrenze vom Argument x abhängig wählen, vgl. Bemerkung 8.3.9. Die untere Integralgrenze kann dabei beliebig sein, wie die folgende Proposition zeigt:

Proposition 8.4.10

Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und $c \in [a, b]$ beliebig. Dann ist

$$G : [a, b] \rightarrow \mathbb{R}, x \mapsto \int_c^x f(t) \, dt$$

eine Stammfunktion von f auf $[a, b]$. Das Intervall $[a, b]$ kann hier durchgängig durch ein

allgemeines Intervall $I \subseteq \mathbb{R}$ ersetzt werden.

Beweis. Wir wissen aus der Bemerkung 8.3.9, dass $F : [a, b] \rightarrow \mathbb{R}, x \mapsto \int_a^x f(t) \, dt$ eine Stammfunktion von f ist. Nun gilt für $x \in [a, b]$

$$G(x) = \int_c^x f(t) \, dt \stackrel{\text{Prop 8.2.14}}{=} \int_a^x f(t) \, dt + \int_c^a f(t) \, dt = F(x) + \int_c^a f(t) \, dt$$

Definiert man nun $k := \int_c^a f(t) \, dt \in \mathbb{R}$ so gilt

$$G(x) = F(x) + k \quad \forall x \in [a, b].$$

Nach Proposition 8.3.10 ist G dann eine weitere Stammfunktion.

Dass $[a, b]$ durch ein beliebiges (nicht-zwingend abgeschlossenes und beschränktes Intervall I ersetzt werden kann liegt daran, dass für die Integration eigentlich nur das Verhalten von f auf dem abgeschlossenen Teilintervall $[\min\{c, x\}, \max\{c, x\}] \subseteq I$ berücksichtigt werden muss. Wir überlassen den Lesenden die Details als Übungsaufgabe.

Beispiel 8.4.11

Wir finden eine Stammfunktion von

$$f : (0, \infty) \rightarrow \mathbb{R}, x \mapsto \log(x).$$

Dazu: Nach Proposition 8.4.10 (mit $c = 1$) ist eine solche gegeben durch

$$F : (0, \infty) \rightarrow \mathbb{R}, x \mapsto \int_1^x \log(t) \, dt.$$

Man berechnet wie im vorigen Beispiel

$$\begin{aligned} \int_1^x \log(t) \, dt &= \int_1^x \overset{\uparrow}{1} \cdot \overset{\downarrow}{\log(t)} \, dt \stackrel{\text{PI}}{=} \left[t \log(t) \right]_{t=1}^{t=x} - \int_1^x t \cdot \frac{1}{t} \, dt = \left[t \log(t) \right]_{t=1}^{t=x} - \int_1^x 1 \, dt \\ &= \left[t \log(t) \right]_{t=1}^{t=x} - \left[t \right]_{t=1}^{t=x} = x \log(x) - x + 1. \end{aligned}$$

Das „+1“ am Ende könnten wir auch weglassen, da additive Konstanten stets zu Stammfunktionen addiert/subtrahiert werden können, vgl. Proposition 8.3.10.

Fazit. $G : (0, \infty) \rightarrow \mathbb{R}, G(x) := x \log(x) - x$ definiert eine Stammfunktion von f .

Weitere Integrationstricks

Hier lernen wir einige weitere trickreiche Integrationen kennen.

(1) Substitution und partielle Integration zusammen.

Manchmal ist ein Trick nicht genug und es müssen zwei Integrationstricks geschickt zusammenspielen.

Beispiel 8.4.12

Wir berechnen

$$\int_1^2 e^{\sqrt{x}} dx.$$

Dazu erst eine nützliche Vorüberlegung

$$\frac{d}{dx} \sqrt{x} = \frac{d}{dx} x^{\frac{1}{2}} = \frac{1}{2} x^{-\frac{1}{2}} = \frac{1}{2\sqrt{x}}. \quad (474)$$

Wir formen erst um:

$$\int_1^2 e^{\sqrt{x}} dx = \int_1^2 \frac{2\sqrt{x}}{2\sqrt{x}} e^{\sqrt{x}} dx.$$

Dies und (474) legt nahe, dass eine Substitution weiterhelfen kann.

$$\int_1^2 e^{\sqrt{x}} dx = \int_1^2 2\sqrt{x} e^{\sqrt{x}} \cdot \frac{1}{2\sqrt{x}} dx \stackrel{\substack{\text{Subst.: } t = \sqrt{x} \\ \frac{dt}{dx} = \frac{1}{2\sqrt{x}} \\ dt = \frac{1}{2\sqrt{x}} dx}}{\int_{\sqrt{1}}^{\sqrt{2}}} 2te^t dt = 2 \int_1^{\sqrt{2}} te^t dt. \quad (475)$$

Das Integral was dadurch aufgetaucht ist lässt sich dann mit einer partellen Integration lösen. In der Tat gilt

$$\int_1^{\sqrt{2}} \overset{\uparrow}{t} \cdot \overset{\downarrow}{e^t} dt \stackrel{\text{PI}}{=} \left[e^t \cdot t \right]_{t=1}^{t=\sqrt{2}} - \int_1^{\sqrt{2}} e^t \cdot 1 dt = \left[e^t \cdot t \right]_{t=1}^{t=\sqrt{2}} - \int_1^{\sqrt{2}} e^t dt = \left[e^t \cdot t \right]_{t=1}^{t=\sqrt{2}} - \left[e^t \right]_{t=1}^{t=\sqrt{2}}.$$

Wertet man den Ausdruck auf der rechten Seite aus so ergibt sich

$$\int_1^{\sqrt{2}} te^t dt = (\sqrt{2} - 1)e^{\sqrt{2}}.$$

Mit (475) ergibt sich

$$\int_1^2 e^{\sqrt{x}} dx = 2(\sqrt{2} - 1)e^{\sqrt{2}}.$$

In obigen Beispiel war zuerst eine Substitution notwendig und danach eine partielle Integration. Es kann aber auch andersherum sein...

Beispiel 8.4.13

Wir berechnen

$$\int_0^1 \arctan(x) \, dx.$$

Dazu zunächst

$$\int_0^1 \arctan(x) \, dx = \int_0^1 \overset{\uparrow}{1} \cdot \overset{\downarrow}{\arctan}(x) \, dx = \left[x \cdot \arctan(x) \right]_{x=0}^{x=1} - \int_0^1 x \frac{1}{1+x^2} \, dx \quad {}^a \quad (476)$$

Nun muss man noch $\int_0^1 \frac{x}{1+x^2} \, dx$ berechnen. Dazu:

$$\int_0^1 \frac{x}{1+x^2} \, dx = \frac{1}{2} \int_0^1 \frac{2x}{1+x^2} \, dx \underset{\substack{\text{Subst.: } t = 1+x^2 \\ \frac{dt}{dx} = 2x \\ dt = 2x \, dx}}{=} \frac{1}{2} \int_1^2 \frac{1}{t} \, dt = \frac{1}{2} \left[\log(t) \right]_{t=1}^{t=2} = \frac{1}{2} \log(2).$$

Man erhält mit (476), dass

$$\int_0^1 \arctan(x) \, dx = \left[x \cdot \arctan(x) \right]_{x=0}^{x=1} - \frac{1}{2} \log(2) = \arctan(1) - \frac{1}{2} \log(2) \underset{\text{Prop 5.5.17}}{=} \frac{\pi}{4} - \frac{1}{2} \log(2).$$

^aHierbei haben wir benutzt, dass laut Proposition 7.3.20 gilt $\frac{d}{dx} \arctan(x) = \frac{1}{1+x^2}$

(2) Partialbruchzerlegung.

Oft ist es ratsam, Ausdrücke geschickt umzuformen bevor man sie integriert. Ein Beispiel dafür ist die Integration *rationaler Funktionen*. Hier kann stets eine *Partialbruchzerlegung vor dem Integrieren* hilfreich sein.

Beispiel 8.4.14

Wir berechnen

$$\int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{1-x^2} \, dx$$

Dazu: Für $x \in (-1, 1)$ gilt

$$\frac{1}{1-x^2} = \frac{1}{(1-x)(1+x)}$$

NEBENRECHNUNG. Partialbruchzerlegung.

Wir gehen wie in der Fußnote von Beispiel 5.0.4 vor um eine Partialbruchzerlegung zu

erhalten. Das heißt wir machen den Ansatz

$$\frac{1}{(1-x)(1+x)} \stackrel{!}{=} \frac{A}{1-x} + \frac{B}{1+x} \quad \text{für gewisse } A, B \in \mathbb{R}. \quad (477)$$

Sind A, B wie in obiger Gleichung so gilt

$$\frac{1}{(1-x)(1+x)} = \frac{A(1-x)}{(1-x)(1+x)} + \frac{B(1+x)}{(1-x)(1+x)} = \frac{(A+B) + (B-A)x}{(1-x)(1+x)}$$

Damit dies für alle $x \in (-1, 1)$ gilt ist es ratsam A und B so zu wählen, dass

$$\begin{cases} A + B = 1 & \text{und} \\ B - A = 0 \end{cases}$$

Es ergibt sich ein lineares Gleichungssystem, dessen eindeutige Lösung gegeben ist durch $A = B = \frac{1}{2}$.^a Man folgert mit (477)

$$\frac{1}{1-x^2} = \frac{1}{(1-x)(1+x)} = \frac{1}{2} \frac{1}{1-x} + \frac{1}{2} \frac{1}{1+x}$$

Somit hat man

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

Nun könnte man noch mit Substitutionen arbeiten (z.B. $t = 1 - x$ im ersten Integral und $t = 1 + x$ im zweiten Integral) oder einfach direkt die Stammfunktionen *raten*. Wir machen an der stelle mal letzteres und folgern

$$\begin{aligned} \int_{-\frac{1}{2}}^{\frac{1}{2}} \frac{1}{1-x^2} dx &= \frac{1}{2} \left[-\log(1-x) \right]_{x=-\frac{1}{2}}^{x=\frac{1}{2}} + \frac{1}{2} \left[\log(1+x) \right]_{x=-\frac{1}{2}}^{\frac{1}{2}} \\ &= \frac{1}{2} \left(-\log \frac{1}{2} + \log \frac{3}{2} \right) + \frac{1}{2} \left(\log \frac{3}{2} - \log \frac{1}{2} \right) \\ &= \log \frac{3}{2} - \log \frac{1}{2} = \log 3. \end{aligned}$$

^aFür verschiedene Wege um auf diese Lösung zu kommen verweisen wir auf die lineare Algebra.

Eine berechtigte Frage ist, wie der Ansatz für eine Partialbruchzerlegung (vgl. (477)) allgemein zu wählen ist, sodass er wirklich zu einer Partialbruchzerlegung führt.

Bemerkung 8.4.15

Wir werden unten sehen: Für $x_0, \dots, x_m \in \mathbb{R}$ und $r_0, \dots, r_m \in \mathbb{N}$ kann jeder Ausdruck der

Form

$$\frac{p(x)}{(x-x_0)^{r_0}(x-x_1)^{r_1} \cdots (x-x_m)^{r_m}} \quad (\text{wobei } p \text{ ein beliebiges Polynom ist})$$

in Partialbrüche zerlegt werden, d.h. umgeschrieben werden als Summe der Form

$$\sum_{k=0}^N \frac{c_k}{(x-z_k)^{\mu_k}}, \quad (478)$$

wobei $N \in \mathbb{N}$ und für alle $k \in \{0, \dots, N\}$ gilt $\mu_k \in \mathbb{Z}$, $z_k \in \{x_0, \dots, x_m\}$ und $c_k \in \mathbb{R}$.

Für alle Summanden auf der rechten Seite von (479) können wir mithilfe von Proposition 8.3.15 (iii) und (iv) Stammfunktionen angeben. Daher ist diese Zerlegung vorteilhaft für's Integrieren!

Zu der Partialbruchzerlegung aus Bemerkung 8.4.15 präsentieren wir ein systematisches Vorgehen. Mit diesem Vorgehen kann die Gestalt aus (478) immer erreicht werden.

1. Für $r_0, r_1 \in \mathbb{N}$ und $x_0, x_1 \in \mathbb{R}$ führt der Ansatz

$$\frac{1}{(x-x_0)^{r_0}(x-x_1)^{r_1}} \stackrel{!}{=} \sum_{k=1}^{r_0} \frac{A_k}{(x-x_0)^k} + \sum_{k=1}^{r_1} \frac{B_k}{(x-x_1)^k} \quad (479)$$

für gewisse $A_1, \dots, A_{r_0}, B_1, \dots, B_{r_1}$ stets zum Ziel und liefert die Gestalt aus (478). Das kann auch leicht mit Induktion nach $n := r_0 + r_1$ bewiesen werden.

2. Wer mit einem Ausdruck der Form $\frac{1}{(x-x_0)^{r_0}(x-x_1)^{r_1}(x-x_2)^{r_2}}$ konfrontiert ist, kann diesen zum Beispiel umschreiben als $\frac{1}{(x-x_2)^{r_2}} \left(\frac{1}{(x-x_0)^{r_0}(x-x_1)^{r_1}} \right)$ und zuerst das, was in der Klammer steht, gemäß (479) zerlegen. Multipliziert man jeden Summanden in dem Ergebnis dann mit $\frac{1}{(x-x_2)^{r_2}}$ so haben wir (vgl. (479))

$$\frac{1}{(x-x_2)^{r_2}} \left(\frac{1}{(x-x_0)^{r_0}(x-x_1)^{r_1}} \right) = \sum_{k=1}^{r_0} \frac{A_k}{(x-x_0)^{r_2}(x-x_0)^k} + \sum_{k=1}^{r_1} \frac{B_k}{(x-x_0)^{r_2}(x-x_1)^k}$$

für gewisse Zahlen $A_1, \dots, A_{r_0}, B_1, \dots, B_{r_1} \in \mathbb{R}$, die wir nach Schritt 1 ausrechnen können. Nun haben alle Summanden auf der rechten Seite wiederum die Gestalt aus Schritt 1. Daher lassen sie sich wie in Schritt 1 weiter zerlegen.

3. So kann man iterativ für alle $m \in \mathbb{N}$ Terme der Form

$$\frac{1}{(x-x_0)^{r_0} \cdots (x-x_m)^{r_m}} \quad x_0, \dots, x_m \in \mathbb{R}, r_0, \dots, r_m \in \mathbb{N}$$

in Partialbrüche zerlegen.

4. Nun zum allgemeinen Fall: Man hat eine rationale Funktion der Form

$$\frac{p(x)}{(x-x_0)^{r_0} \cdot \dots \cdot (x-x_m)^{r_m}} \quad x_0, \dots, x_m \in \mathbb{R}, r_0, \dots, r_m \in \mathbb{N}$$

(wobei p ein beliebiges Polynom ist) als Integrand gegeben. Da man im Nenner gemäß Schritt 3 eine Partialbruchzerlegung machen kann, kann man schreiben

$$\frac{1}{(x-x_0)^{r_0} \cdot \dots \cdot (x-x_m)^{r_m}} = \sum_{k=0}^N \frac{c_k}{(x-z_k)^{\mu_k}},$$

wobei $N \in \mathbb{N}$ und für alle $k \in \{1, \dots, N\}$ gilt $c_k \in \mathbb{R}$, $z_k \in \{x_0, \dots, x_m\}$ und $\mu_k \in \mathbb{Z}$. Multipliziert man mit dem Polynom p so hat man

$$\frac{p(x)}{(x-x_0)^{r_0} \cdot \dots \cdot (x-x_m)^{r_m}} = \sum_{k=0}^N c_k \frac{p(x)}{(x-z_k)^{\mu_k}}.$$

Alle Summanden auf der rechten Seite sind von der Form $\frac{p(x)}{(x-z)^\mu}$ (für ein $z \in \mathbb{R}$ und $\mu \in \mathbb{Z}$). Daher genügt es, wenn man solche Terme in die Form aus (478) bringen kann. Da das Polynom p eine Summe von Monomen der Form x^ℓ ($\ell \in \mathbb{N}_0$) ist, genügt es auch, wenn man $\frac{x^\ell}{(x-z)^\mu}$ in die Form aus (478) bringen kann. Dazu berechne für $\mu \in \mathbb{N}$ und $\ell \in \mathbb{N}_0$

$$\begin{aligned} \frac{x^\ell}{(x-z)^\mu} &= \frac{(x-z+z)^\ell}{(x-z)^\mu} \stackrel{\text{Satz 3.3.27}}{=} \frac{1}{(x-z)^\mu} \sum_{k=0}^{\ell} \binom{\ell}{k} (x-z)^k z^{\ell-k} \\ &= \sum_{k=0}^{\ell} \binom{\ell}{k} z^{\ell-k} (x-z)^{k-\mu} = \sum_{k=0}^{\ell} \frac{\binom{\ell}{k} z^{\ell-k}}{(x-z)^{\mu-k}}. \end{aligned}$$

Dies ist die Form aus (478) mit $N = \ell$, $c_k = \binom{\ell}{k} z^{\ell-k}$ und $\mu_k = \mu - k$.

Beachte: Für unsere Partialbruchzerlegung aus Bemerkung 8.4.15 brauchen wir, dass das Polynom im Nenner in *Linearfaktoren* zerfällt. Ist dies nicht der Fall, so können die obigen Schritte nicht angewendet werden. In der linearen Algebra werden Sie aber besprechen, dass sich *über den komplexen Zahlen* jedes nichtkonstante Polynom als Produkt von Linearfaktoren schreiben lässt. Über $\mathbb{C}$ würde die obige Partialbruchzerlegung also immer gelingen! Dies kann man als eine Motivation sehen, unsere Integrationstheorie auf komplexwertige Funktionen zu verallgemeinern. Wir kommen später dazu.

(3) Der Phönix.

Als *Phönix* bezeichnet man eine Technik, bei der man zweimal partiell integriert und das Integral danach als *Gleichung* löst. Sie bietet sich immer an wenn man Integrale hat bei denen $\cos, \sin, \exp$ doppelt und mit verschiedenen Argumenten vorkommen (ist aber nicht immer die einfachste Methode). Am besten lernt man sie, indem man sich ein Beispiel dazu anschaut.

Beispiel 8.4.16

Wir berechnen

$$I := \int_0^\pi \sin(2x) \cos(x) \, dx. \quad (480)$$

Dazu rechnen wir zunächst

$$I = \int_0^\pi \overset{\downarrow}{\sin(2x)} \overset{\uparrow}{\cos(x)} \, dx = \left[\sin(2x) \sin(x) \right]_{x=0}^{x=\pi} - \int_0^\pi 2 \cos(2x) \sin(x) \, dx. \quad (481)$$

Das zweite Integral kann weiter durch partielle Integration umgeformt werden

$$\begin{aligned} \int_0^\pi 2 \overset{\downarrow}{\cos(2x)} \overset{\uparrow}{\sin(x)} \, dx &= \left[(2 \cos(2x))(-\cos(x)) \right]_{x=0}^{x=\pi} - \int_0^\pi (-4 \sin(2x))(-\cos(x)) \, dx \\ &= \left[-2 \cos(2x) \cos(x) \right]_{x=0}^{x=\pi} - 4 \int_0^\pi \sin(2x) \cos(x) \, dx. \end{aligned}$$

Setzen wir diese Erkenntnis in (481) ein so erhalten wir

$$\begin{aligned} I &= \left[\sin(2x) \sin(x) \right]_{x=0}^{x=\pi} - \left(\left[-2 \cos(2x) \cos(x) \right]_{x=0}^{x=\pi} - 4 \int_0^\pi \sin(2x) \cos(x) \, dx \right) \\ &= \left[\sin(2x) \sin(x) \right]_{x=0}^{x=\pi} - \left[-2 \cos(2x) \cos(x) \right]_{x=0}^{x=\pi} + 4 \int_0^\pi \sin(2x) \cos(x) \, dx. \end{aligned}$$

Schauen wir die Formel für (480) noch einmal an, so sehen wir, dass das einzige noch verbleibende Integral in der vorigen Gleichung auch mit I übereinstimmt. Das heißt

$$I = \left[\sin(2x) \sin(x) \right]_{x=0}^{x=\pi} - \left[-2 \cos(2x) \cos(x) \right]_{x=0}^{x=\pi} + 4I.$$

Dies können wir als Gleichung für I auffassen. Subtrahieren wir $4I$ auf beiden Seiten, so ergibt sich

$$-3I = \left[\sin(2x) \sin(x) \right]_{x=0}^{x=\pi} - \left[-2 \cos(2x) \cos(x) \right]_{x=0}^{x=\pi}.$$

Teilen wir dann noch durch -3 so erhalten wir

$$I = -\frac{1}{3} \left(\left[\sin(2x) \sin(x) \right]_{x=0}^{x=\pi} - \left[-2 \cos(2x) \cos(x) \right]_{x=0}^{x=\pi} \right) = \frac{1}{3} (0 - (2 - (-2))) = \frac{4}{3}.$$

Also gilt mit (480)

$$\int_0^\pi \sin(2x) \cos(x) \, dx = \frac{4}{3}.$$

Der Phönix ist nicht immer die einfachste Lösung. Oft kann man Integrale mit $\cos, \sin, \exp$ auch auf eine andere Art und Weise lösen.

Beispiel 8.4.17

Im Folgenden berechnen wir noch einmal

$$\int_0^\pi \sin(2x) \cos(x) \, dx,$$

aber diesmal ohne Phönix. Dazu beobachte zuerst, dass nach Proposition 5.5.4 (iv) für alle $x \in \mathbb{R}$ gilt

$$\sin(2x) = 2 \sin(x) \cos(x).$$

Man erhält somit

$$\int_0^\pi \sin(2x) \cos(x) \, dx = \int_0^\pi 2 \sin(x) \cos^2(x) \, dx = (-2) \int_0^\pi (-\sin(x)) \cos^2(x) \, dx \quad (482)$$

Dieses Integral lässt sich durch eine Substitution lösen. In der Tat gilt

$$\int_0^\pi (-\sin(x)) \cos^2(x) \, dx \stackrel{\substack{\text{Subst.: } t = \cos(x) \\ \frac{dt}{dx} = -\sin(x) \\ dt = (-\sin(x)) \, dx}}{=} \int_1^{-1} t^2 \, dt = \left[\frac{1}{3} t^3 \right]_{t=1}^{t=-1} = -\frac{2}{3}.$$

Zusammen mit (482) ergibt sich

$$\int_0^\pi \sin(2x) \cos(x) \, dx = \frac{4}{3},$$

genau wie in Beispiel 8.4.16.

***(4) Die Weierstrass-Substitution**

Oftmals hat man Integrale der Form $\int_a^b f(\cos(x), \sin(x)) \, dx$ für $a, b \in (-\pi, \pi)$. Hierbei bezeichnet $f(\cos(x), \sin(x))$ einen Ausdruck in dem nur $\cos(x)$ und $\sin(x)$ vorkommen und der als ganzes in x stetig ist. All diese Integrale können mithilfe einer genialen Substitution vereinfacht werden. Hierzu benötigen wir zunächst einige Vorbereitung

Lemma 8.4.18 — EINIGE UMFORMUNGEN FÜR $\cos$ UND $\sin$

- (i) Für alle $z \in (-\frac{\pi}{2}, \frac{\pi}{2})$ gilt $\cos(z) = \frac{1}{\sqrt{1+\tan^2(z)}}$.
- (ii) Für alle $z \in (-\frac{\pi}{2}, \frac{\pi}{2})$ gilt $\sin(z) = \frac{\tan(z)}{\sqrt{1+\tan^2(z)}}$.
- (iii) Für alle $x \in (-\pi, \pi)$ gilt $\cos(x) = \frac{1-\tan^2(\frac{x}{2})}{1+\tan^2(\frac{x}{2})}$

(iv) Für alle $x \in (-\pi, \pi)$ gilt $\sin(x) = \frac{2 \tan(\frac{x}{2})}{1 + \tan^2(\frac{x}{2})}$

Beweis. Zu (i). Es sei $z \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Dann gilt

$$\frac{1}{1 + \tan^2(z)} = \frac{1}{1 + \frac{\sin^2(z)}{\cos^2(z)}} = \frac{\cos^2(z)}{\cos^2(z)(1 + \frac{\sin^2(z)}{\cos^2(z)})} = \frac{\cos^2(z)}{\cos^2(z) + \sin^2(z)} = \frac{\cos^2(z)}{1} = \cos^2(z). \quad (483)$$

Weil $z \in (-\frac{\pi}{2}, \frac{\pi}{2})$ gilt $\cos(z) > 0$ und daher

$$\cos(z) = \sqrt{\cos^2(z)} \stackrel{(483)}{=} \sqrt{\frac{1}{1 + \tan^2(z)}} = \frac{1}{\sqrt{1 + \tan^2(z)}}.$$

Zu (ii). Es gilt für $z \in (-\frac{\pi}{2}, \frac{\pi}{2})$

$$\sin(z) = \frac{\sin(z)}{\cos(z)} \cos(z) = \tan(z) \cos(z) \stackrel{(i)}{=} \tan(z) \frac{1}{\sqrt{1 + \tan^2(z)}}.$$

Zu (iii). Es sei $x \in (-\pi, \pi)$. Es gilt nach Proposition 5.5.4 (iv)

$$\cos(x) = \cos^2(\frac{x}{2}) - \sin^2(\frac{x}{2}) = \cos^2(\frac{x}{2}) \left(1 - \frac{\sin^2(\frac{x}{2})}{\cos^2(\frac{x}{2})}\right) = \cos^2(\frac{x}{2}) (1 - \tan^2(\frac{x}{2}))$$

Da $\frac{x}{2} \in (-\frac{\pi}{2}, \frac{\pi}{2})$ können wir den ersten Faktor mit (i) umschreiben, als $\cos^2(\frac{x}{2}) = \left(\frac{1}{\sqrt{1 + \tan^2(\frac{x}{2})}}\right)^2 = \frac{1}{1 + \tan^2(\frac{x}{2})}$. Somit gilt

$$\cos(x) = \frac{1}{1 + \tan^2(\frac{x}{2})} (1 - \tan^2(\frac{x}{2})),$$

was zu zeigen war.

Zu (iv). Sei $x \in (-\pi, \pi)$ [insbesondere also auch $\frac{x}{2} \in (-\frac{\pi}{2}, \frac{\pi}{2})$]. Dann gilt

$$\sin(x) = 2 \sin(\frac{x}{2}) \cos(\frac{x}{2}) \stackrel{(i)}{\stackrel{(ii)}}{=} 2 \frac{\tan(\frac{x}{2})}{\sqrt{1 + \tan^2(\frac{x}{2})}} \frac{1}{\sqrt{1 + \tan^2(\frac{x}{2})}} = \frac{2 \tan(\frac{x}{2})}{1 + \tan^2(\frac{x}{2})}$$

Nun sind wir so weit, dass wir die *Weierstrass-Substitution* besprechen können. Es handelt sich dabei um die Substitution

$$t = \tan(\frac{x}{2}). \quad (484)$$

Diese erlaubt uns, Integrale die Ausdrücke der Form $\cos(x)$, $\sin(x)$ enthalten, mit dem vorigen Lemma in einen anderen Ausdruck umzuwandeln mit den Formeln

$$\cos(x) = \frac{1-t^2}{1+t^2}, \quad \sin(x) = \frac{2t}{1+t^2} \quad (\text{vgl. Lemma 8.4.18(iii),(iv)}). \quad (485)$$

Eine weitere hilfreiche Eigenschaft an (484) ist, dass sich die Ableitung von $x \mapsto \tan(\frac{x}{2})$ wieder mithilfe von $\tan(\frac{x}{2})$ ausdrücken lässt. Genauer:

$$\frac{dt}{dx} = \frac{1}{2} (1 + \tan^2(\frac{x}{2})) = \frac{1}{2}(1+t^2), \quad \text{das heißt} \quad dx = \frac{2}{1+t^2} dt.$$

Beispiel 8.4.19

Wir berechnen für $a, b \in (0, \pi)$

$$\int_a^b \frac{1}{1 - \cos(x)} dx.$$

Dazu nutzen wir die Weierstraß-Substitution $t = \tan(\frac{x}{2})$, und drücken mit (485) $\cos(x)$ durch t aus. Dazu:

$$\begin{aligned} \int_a^b \frac{1}{1 - \cos(x)} dx & \stackrel{\text{Subst.: } t = \tan(\frac{x}{2})}{\substack{\frac{dt}{dx} = \frac{1}{2}(1+t^2) \\ dx = \frac{2}{1+t^2} dt}} \int_{\tan \frac{a}{2}}^{\tan \frac{b}{2}} \frac{1}{1 - \frac{1-t^2}{1+t^2}} \frac{2}{1+t^2} dt = \int_{\tan \frac{a}{2}}^{\tan \frac{b}{2}} \frac{2}{(1+t^2) - (1-t^2)} dt \\ & = \int_{\tan \frac{a}{2}}^{\tan \frac{b}{2}} \frac{2}{2t^2} dt = \int_{\tan \frac{a}{2}}^{\tan \frac{b}{2}} \frac{1}{t^2} dt = \left[-\frac{1}{t} \right]_{t=\tan \frac{a}{2}}^{t=\tan \frac{b}{2}} = \frac{1}{\tan \frac{a}{2}} - \frac{1}{\tan \frac{b}{2}}. \end{aligned}$$

8.5 Uneigentliche Integrale

Manchmal kommt es vor, dass Integranden am Rand des Integrationsbereiches nicht definiert sind. In diesem Fall bildet man das sogenannte *uneigentliche Integral*.

Definition 8.5.1 — UNEIGENTLICHE INTEGRALE

Es seien $a, b \in \mathbb{R} \cup \{\pm\infty\}$ mit $a < b$. Ferner sei $f : (a, b) \rightarrow \mathbb{R}$ eine Funktion. Wir nennen f auf (a, b) *uneigentlich Riemann-integrierbar*, falls

- für alle $\alpha, \beta \in (a, b)$ mit $\alpha < \beta$ die Funktion $f|_{[\alpha, \beta]}$ Riemann-integrierbar ist und
- es ein $c \in (a, b)$ gibt so, dass die Grenzwerte

$$\lim_{\substack{\alpha \rightarrow a \\ \alpha \in (a, b)}} \int_{\alpha}^c f(x) dx \quad \text{sowie} \quad \lim_{\substack{\beta \rightarrow b \\ \beta \in (a, b)}} \int_c^{\beta} f(x) dx$$

beide existieren.

Ist f uneigentlich Riemann integrierbar so definieren wir das *uneigentliche Riemann-integral*

$$\int_a^b f(x) \, dx := \lim_{\substack{\alpha \rightarrow a \\ \alpha \in (a,b)}} \int_{\alpha}^c f(x) \, dx + \lim_{\substack{\beta \rightarrow b \\ \beta \in (a,b)}} \int_c^{\beta} f(x) \, dx \quad (486)$$

Oft schreibt man auch nur salopp $\int_a^b f(x) \, dx$ anstatt $\int_a^b f(x) \, dx$.

Bemerkung 8.5.2

Es sei $f : (a, b) \rightarrow \mathbb{R}$ uneigentlich Riemann-integrierbar. Damit das uneigentliche Integral in (486) *wohldefiniert* ist, muss man zeigen, dass die Definition von $\int_a^b f(x) \, dx$ nicht von der Wahl von c abhängt. Sprich: Jedes zulässige c liefert in (486) dasselbe Ergebnis.

Dazu: Es seien c und c' Elemente von (a, b) so, dass

$$\lim_{\substack{\alpha \rightarrow a \\ \alpha \in (a,b)}} \int_{\alpha}^c f(x) \, dx \quad \text{sowie} \quad \lim_{\substack{\beta \rightarrow b \\ \beta \in (a,b)}} \int_c^{\beta} f(x) \, dx$$

und

$$\lim_{\substack{\alpha \rightarrow a \\ \alpha \in (a,b)}} \int_{\alpha}^{c'} f(x) \, dx \quad \text{sowie} \quad \lim_{\substack{\beta \rightarrow b \\ \beta \in (a,b)}} \int_{c'}^{\beta} f(x) \, dx$$

alle existieren.

$$\mathbb{Z} : \quad \lim_{\substack{\alpha \rightarrow a \\ \alpha \in (a,b)}} \int_{\alpha}^c f(x) \, dx + \lim_{\substack{\beta \rightarrow b \\ \beta \in (a,b)}} \int_c^{\beta} f(x) \, dx = \lim_{\substack{\alpha \rightarrow a \\ \alpha \in (a,b)}} \int_{\alpha}^{c'} f(x) \, dx + \lim_{\substack{\beta \rightarrow b \\ \beta \in (a,b)}} \int_{c'}^{\beta} f(x) \, dx.$$

Dazu:

$$\begin{aligned} & \lim_{\substack{\alpha \rightarrow a \\ \alpha \in (a,b)}} \int_{\alpha}^c f(x) \, dx + \lim_{\substack{\beta \rightarrow b \\ \beta \in (a,b)}} \int_c^{\beta} f(x) \, dx \\ &= \lim_{\substack{\alpha \rightarrow a \\ \alpha \in (a,b)}} \left(\int_{\alpha}^{c'} f(x) \, dx + \int_{c'}^c f(x) \, dx \right) + \lim_{\substack{\beta \rightarrow b \\ \beta \in (a,b)}} \left(\int_{c'}^{\beta} f(x) \, dx + \int_c^{c'} f(x) \, dx \right) \\ &\stackrel{\text{GWR}}{=} \lim_{\substack{\alpha \rightarrow a \\ \alpha \in (a,b)}} \int_{\alpha}^{c'} f(x) \, dx + \int_{c'}^c f(x) \, dx + \lim_{\substack{\beta \rightarrow b \\ \beta \in (a,b)}} \int_{c'}^{\beta} f(x) \, dx + \int_c^{c'} f(x) \, dx \\ &= \lim_{\substack{\alpha \rightarrow a \\ \alpha \in (a,b)}} \int_{\alpha}^{c'} f(x) \, dx + \lim_{\substack{\beta \rightarrow b \\ \beta \in (a,b)}} \int_{c'}^{\beta} f(x) \, dx + \int_{c'}^c f(x) \, dx + \int_c^{c'} f(x) \, dx \end{aligned}$$

$$\begin{aligned}
& \stackrel{\text{Def 8.2.13}}{=} \lim_{\substack{\alpha \rightarrow a \\ \alpha \in (a,b)}} \int_{\alpha}^{c'} f(x) \, dx + \lim_{\substack{\beta \rightarrow b \\ \beta \in (a,b)}} \int_{c'}^{\beta} f(x) \, dx + \int_{c'}^c f(x) \, dx - \int_{c'}^c f(x) \, dx \\
& = \lim_{\substack{\alpha \rightarrow a \\ \alpha \in (a,b)}} \int_{\alpha}^{c'} f(x) \, dx + \lim_{\substack{\beta \rightarrow b \\ \beta \in (a,b)}} \int_{c'}^{\beta} f(x) \, dx.
\end{aligned}$$

Bemerkenswert an Definition 8.5.1 ist, dass auch Integranden zugelassen sind, die auf unbeschränkten Intervallen definiert sind.

Beispiel 8.5.3

Behauptung. $f : (1, \infty) \rightarrow \mathbb{R}, x \mapsto \frac{1}{x^2}$ ist auf $(1, \infty)$ uneigentlich Riemann-integrierbar und $\int_1^{\infty} \frac{1}{x^2} \, dx = 1$.

Beweis. Für alle $\alpha, \beta \in (1, \infty)$ mit $\alpha < \beta$ ist f auf $[\alpha, \beta]$ Riemann-integrierbar, da f dort stetig ist (vgl. Proposition 6.3.12). Wir wählen nun $c = 2$. Dann gilt für $\alpha \in (1, \infty)$

$$\int_{\alpha}^c \frac{1}{x^2} \, dx = \int_{\alpha}^2 x^{-2} \, dx \stackrel{\text{Prop 8.3.15(ii)}}{=} \left[-\frac{1}{x} x^{-1} \right]_{x=\alpha}^{x=2} = -\frac{1}{2} + \frac{1}{\alpha}.$$

Somit hat man

$$\lim_{\alpha \rightarrow 1} \int_{\alpha}^c \frac{1}{x^2} \, dx = \lim_{\alpha \rightarrow 1} \left(-\frac{1}{2} + \frac{1}{\alpha} \right) = -\frac{1}{2} + 1 = \frac{1}{2}. \quad (487)$$

Ferner gilt für $\beta \in (1, \infty)$

$$\int_c^{\beta} \frac{1}{x^2} \, dx = \int_2^{\beta} x^{-2} \, dx \stackrel{\text{Prop 8.3.15(ii)}}{=} \left[-\frac{1}{x} x^{-1} \right]_{x=2}^{x=\beta} = -\frac{1}{\beta} + \frac{1}{2}.$$

Somit hat man

$$\lim_{\beta \rightarrow \infty} \int_c^{\beta} \frac{1}{x^2} \, dx = \lim_{\beta \rightarrow \infty} \left(-\frac{1}{\beta} + \frac{1}{2} \right) = \frac{1}{2}. \quad (488)$$

Mit Definition 8.5.1 ergibt sich, dass f auf $(1, \infty)$ uneigentlich Riemann-integrierbar ist und

$$\int_a^b f(x) \, dx = \lim_{\alpha \rightarrow 1} \int_{\alpha}^c \frac{1}{x^2} \, dx + \lim_{\beta \rightarrow \infty} \int_c^{\beta} \frac{1}{x^2} \, dx \stackrel{(488), (487)}{=} \frac{1}{2} + \frac{1}{2} = 1.$$

Eine Frage die sich aufdrängen sollte ist, ob der Begriff des uneigentlichen Integrals *konsistent* ist mit dem vorher eingeführten Riemann-Integral. Genauer fragen wir uns, ob (eigentlich) Riemann-integrierbare Funktionen $f : [a, b] \rightarrow \mathbb{R}$ auch uneigentlich Riemann-integrierbar sind. Nur wenn das erfüllt ist kann das uneigentliche Riemann-integral als Erweiterung unseres bisher definierten Integralkonzeptes gesehen werden.

Proposition 8.5.4 — KONSISTENZ DER INTEGRALBEGRIFFE

Es seien $a, b \in \mathbb{R}$ und $a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar^a. Dann ist $f|_{(a,b)}$ auch uneigentlich Riemann-integrierbar und

$$\int_a^b f(x) \, dx = \int_a^b f(x) \, dx \quad (489)$$

^aInsbesondere ist f also auch beschränkt, vgl. Definition 8.1.15

Beweis. Wir überprüfen die beiden Punkte in Definition 8.5.1.

SCHRITT 1. Riemann-Integrierbarkeit auf Teilintervallen $[\alpha, \beta]$ (mit $a < \alpha < \beta < b$).

Die Riemann-Integrierbarkeit von f auf jedem Teilintervall $[\alpha, \beta] \subseteq [a, b]$ folgt sofort aus Proposition 8.2.10 (bzw. Bemerkung 8.2.11).

SCHRITT 2. Grenzwerte der Riemann-Integrale.

Es sei $c \in (a, b)$ beliebig. Wir zeigen nun, dass die Grenzwerte aus Definition 8.5.1 existieren (und bestimmen sie dabei auch). Dazu sei $\alpha \in (a, b)$. Dann gilt

$$\left| \int_{\alpha}^c f(x) \, dx - \int_a^c f(x) \, dx \right| = \left| \int_{\alpha}^c f(x) \, dx + \int_c^a f(x) \, dx \right| \stackrel{\text{Prop 8.2.14}}{=} \left| \int_{\alpha}^a f(x) \, dx \right|$$

Wegen Definition 8.2.13 kann man die Grenzen auch vertauschen und erhält

$$\left| \int_{\alpha}^c f(x) \, dx - \int_a^c f(x) \, dx \right| = \left| \int_{\alpha}^a f(x) \, dx \right| = \left| \int_a^{\alpha} f(x) \, dx \right| \quad (490)$$

Da f beschränkt ist, gibt es ein $K \geq 0$ so, dass $|f(x)| \leq K$ für alle $x \in [a, b]$. Dies und die Dreiecksungleichung für Integrale (vgl. Korollar 8.2.6) impliziert

$$\left| \int_a^{\alpha} f(x) \, dx \right| \leq \int_a^{\alpha} |f(x)| \, dx \leq \int_a^{\alpha} K \, dx = K(\alpha - a).$$

Nutzt man dies in (490) so sieht man

$$0 \leq \left| \int_{\alpha}^c f(x) \, dx - \int_a^c f(x) \, dx \right| = \left| \int_a^{\alpha} f(x) \, dx \right| \leq K(\alpha - a) \xrightarrow{\alpha \rightarrow a} 0.$$

Aus dem Sandwichlemma folgt

$$\lim_{\alpha \rightarrow a} \left| \int_{\alpha}^c f(x) \, dx - \int_a^c f(x) \, dx \right| = 0$$

Man folgert daraus direkt (vgl. Proposition 2.3.30)

$$\lim_{\alpha \rightarrow a} \int_{\alpha}^c f(x) \, dx = \int_a^c f(x) \, dx. \quad (491)$$

Ganz analog kann man zeigen, dass

$$\lim_{\beta \rightarrow b} \int_c^{\beta} f(x) \, dx = \int_c^b f(x) \, dx. \quad (492)$$

Die uneigentliche Riemann-Integrierbarkeit folgt dann aus den zwei vorigen Gleichungen.

SCHRITT 3. Wert des uneigentlichen Riemann-Integrals

Es muss nun nur noch (489) nachgewiesen werden. Dazu

$$\begin{aligned} \int_a^b f(x) \, dx &= \lim_{\alpha \rightarrow a} \int_{\alpha}^c f(x) \, dx + \lim_{\beta \rightarrow b} \int_c^{\beta} f(x) \, dx \\ &\stackrel{(491), (492)}{=} \int_a^c f(x) \, dx + \int_c^b f(x) \, dx \stackrel{\text{Prop 8.2.14}}{=} \int_a^b f(x) \, dx \end{aligned}$$

Für uneigentliche Integrale gelten ähnliche Rechenregeln wie für eigentliche: *Linearität*, *Monotonie* und *Teilungseigenschaft*. Dies fassen wir noch in der nachfolgenden Proposition zusammen

Proposition 8.5.5 — EIGENSCHAFTEN DES UNEIGENTLICHEN INTEGRALS

Es seien $f, g : (a, b) \rightarrow \mathbb{R}$ uneigentlich Riemann-integrierbar und $\lambda, \mu \in \mathbb{R}$.

(i) *Linearität*. $\lambda f + \mu g$ ist uneigentlich Riemann integrierbar und

$$\int_a^b (\lambda f + \mu g)(x) \, dx = \lambda \int_a^b f(x) \, dx + \mu \int_a^b g(x) \, dx.$$

(ii) *Monotonie*. Ist $f(x) \leq g(x)$ für alle $x \in (a, b)$ so gilt

$$\int_a^b f(x) \, dx \leq \int_a^b g(x) \, dx$$

(iii) *Teilungsformel*. Für $c \in (a, b)$ sind auch $f|_{(a,c)}$, $f|_{(c,b)}$ uneigentlich Riemann-integrierbar und es gilt

$$\int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx$$

Beweis. Übungsaufgabe. Folgt ziemlich direkt aus den Erkenntnissen aus Sektion 8.2 über das *eigentliche* Riemann-integral.

8.6 Integration komplexwertiger Funktionen

Bisher können wir nur reellwertige Funktionen integrieren. Das wollen wir nun ändern. Der Definitionsbereich ist allerdings nach wie vor ein abgeschlossenes beschränktes Intervall $[a, b] \subseteq \mathbb{R}$ (oder ein offenes Intervall (a, b) falls man über uneigentliche Integrale spricht). Wir werden uns der Einfachheit halber auf *eigentliche* Riemann-Integrale beschränken.

Definition 8.6.1 — INTEGRATION KOMPLEXWERTIGER FUNKTIONEN

Es sei $f : [a, b] \rightarrow \mathbb{C}$ eine beschränkte Funktion.^a Wir nennen f *Riemann-integrierbar* falls $\operatorname{Re}(f) : [a, b] \rightarrow \mathbb{R}$ und $\operatorname{Im}(f) : [a, b] \rightarrow \mathbb{R}$ beide Riemann integrierbar (im Sinne von Definition 8.1.15) sind. In diesem Fall definieren wir

$$\int_a^b f(x) \, dx := \int_a^b \operatorname{Re}(f(x)) \, dx + i \int_a^b \operatorname{Im}(f(x)) \, dx.$$

^aWir weisen an dieser Stelle noch einmal gesondert darauf hin, dass der *Definitionsbereich* $[a, b]$ trotzdem ein *reelles* Intervall sein muss. Lediglich der Wertebereich ist hier komplexwertig zugelassen!

Die Definition der Integrierbarkeit ist zwar intuitiv sinnvoll, jedoch ist sie für die *Berechnung* von Integralen nicht optimal. Unser hauptsächliches Hilfsmittel zur Berechnung von Integralen ist der *Hauptsatz der Differential und Integralrechnung* (HDI, vgl. Satz 8.3.12). Diesen werden wir nun auf komplexwertige Funktionen verallgemeinern.

Satz 8.6.2 — HDI – KOMPLEXWERTIGE VERALLGEMEINERUNG

Es sei $f : [a, b] \rightarrow \mathbb{C}$ eine stetige Funktion und $F : [a, b] \rightarrow \mathbb{C}$ eine beliebige Stammfunktion von f auf $[a, b]$ (d.h. F ist diffbar und $F'(x) = f(x)$ für alle $x \in [a, b]$). Dann gilt

$$\int_a^b f(t) \, dt = F(b) - F(a).$$

Beweis. Es sei $F : [a, b] \rightarrow \mathbb{C}$ wie in der Aussage eine Stammfunktion von f , d.h.

$$F'(x) = f(x) \quad \forall x \in [a, b].$$

ZWISCHENBEHAUPTUNG. $\operatorname{Re}(F)$ bzw. $\operatorname{Im}(F)$ sind Stammfunktionen von $\operatorname{Re}(f)$ bzw. $\operatorname{Im}(f)$ auf $[a, b]$.

Dies zeigen wir hier mal exemplarisch nur für $\operatorname{Re}(F)$. Dazu sei $x_0 \in [a, b]$.

$$\frac{\operatorname{Re}(F(x)) - \operatorname{Re}(F(x_0))}{x - x_0} = \frac{\operatorname{Re}(F(x) - F(x_0))}{x - x_0} = \operatorname{Re}\left(\frac{F(x) - F(x_0)}{x - x_0}\right)$$

Nun gilt $\lim_{x \rightarrow x_0} \frac{F(x) - F(x_0)}{x - x_0} = F'(x_0) = f(x_0)$. Dies impliziert nach Grenzwertregeln

$$\lim_{x \rightarrow x_0} \frac{\operatorname{Re}(F(x)) - \operatorname{Re}(F(x_0))}{x - x_0} = \lim_{x \rightarrow x_0} \operatorname{Re}\left(\frac{F(x) - F(x_0)}{x - x_0}\right) = \operatorname{Re}(f(x_0))$$

Dies zeigt, dass $\operatorname{Re}(F)$ bei x_0 differenzierbar ist und

$$\operatorname{Re}(F)'(x_0) = \operatorname{Re}(f)(x_0).$$

Die Behauptung folgt.

Wir haben also Stammfunktionen von $\operatorname{Re}(f)$ und $\operatorname{Im}(f)$. Beachte: Desweiteren sind $\operatorname{Re}(f)$ und $\operatorname{Im}(f)$ als Kompositionen stetiger Funktionen auch stetig. Nun können wir mit dem reellen Hauptsatz (Satz 8.3.12) den Beweis vollenden:

$$\begin{aligned} \int_a^b f(t) \, dt &= \int_a^b \operatorname{Re}(f(t)) \, dt + i \int_a^b \operatorname{Im}(f(t)) \, dt \\ &\stackrel{(\text{??})}{\underset{\text{Satz 8.3.12}}{=}} (\operatorname{Re}(F(b)) - \operatorname{Re}(F(a)) + i(\operatorname{Im}(F(b)) - \operatorname{Im}(F(a)))) \end{aligned}$$

Sortieren wir um so erhalten wir

$$\begin{aligned} \int_a^b f(t) \, dt &= \operatorname{Re}(F(b)) + i\operatorname{Im}(F(b)) - \operatorname{Re}(F(a)) - i\operatorname{Im}(F(a)) \\ &= \operatorname{Re}(F(b)) + i\operatorname{Im}(F(b)) - [\operatorname{Re}(F(a)) + i\operatorname{Im}(F(a))] = F(b) - F(a). \end{aligned}$$

Somit können wir also auch Integrale über komplexwertige Funktionen wie gewohnt durch Bilden einer *Stammfunktion* bestimmen

Beispiel 8.6.3

Wir berechnen

$$\int_0^{2\pi} e^{it} \, dt.$$

Dazu: Mit der Kettenregel überlegt man sich leicht, dass für alle $x \in \mathbb{R}$

$$\frac{d}{dx} \left(\frac{1}{i} e^{ix} \right) = \frac{1}{i} i e^{ix} = e^{ix}.$$

Somit ist also $F : [0, 2\pi] \rightarrow \mathbb{C}, t \mapsto \frac{1}{i} e^{it}$ eine Stammfunktion von $f : [0, 2\pi] \rightarrow \mathbb{C}, t \mapsto e^{it}$.

Wir erhalten

$$\int_0^{2\pi} e^{it} \, dt = \int_0^{2\pi} f(t) \, dt = F(2\pi) - F(0) = \frac{1}{i}e^{2\pi i} - \frac{1}{i}e^0 \stackrel{\text{Kor5.5.15}}{=} \frac{1}{i} - \frac{1}{i} = 0.$$

20.05.2026

9 Funktionenfolgen

Es seien im Folgenden wieder $\mathbb{K}_1, \mathbb{K}_2, \mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$.

In der Vergangenheit haben wir Zahlenfolgen $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{K}$ betrachtet und für diese einen *Konvergenzbegriff* definiert. Nun wollen wir Konvergenzbegriffe für *Folgen von Funktionen* $(f_n)_{n \in \mathbb{N}}$ definieren und studieren.

Definition 9.0.1 — FUNKTIONENFOLGE

Es sei $D \subseteq \mathbb{K}_1$ und für alle $n \in \mathbb{N}$ sei $f_n : D \rightarrow \mathbb{K}_2$ eine Funktion. Die Folge $(f_n)_{n \in \mathbb{N}}$ heißt dann *Funktionenfolge* auf D . Man schreibt auch

$$(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}.$$

Beachte: Eigentlich ist eine Funktionenfolge eine Art *Ansammlung* von Zahlenfolgen: Ist $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ eine Funktionenfolge und $x \in D$ beliebig, so bildet $(f_n(x))_{n \in \mathbb{N}} \subseteq \mathbb{K}_2$ eine Zahlenfolge.

Bemerkung 9.0.2

Eigentlich haben wir bereits einige Funktionenfolgen studiert und auch in gewisser Weise auf Konvergenz untersucht. Nämlich können wir die Frage nach der Konvergenz von *Potenzreihen* mithilfe von Funktionenfolgen betrachten. In der Tat: Es sei $\sum_{k \geq 0} c_k(x - x_0)^k$ eine gegebene Potenzreihe mit Konvergenzradius $R > 0$. Definiere nun für $D := B_R(x_0)$ und $n \in \mathbb{N}$ die Funktion

$$f_n : D \rightarrow \mathbb{C}, \quad f_n(x) := \sum_{k=0}^n c_k(x - x_0)^k.$$

Diese *Partialsummen* der Potenzreihe lassen sich somit als Funktionenfolge $(f_n : D \rightarrow \mathbb{C})_{n \in \mathbb{N}}$ auffassen.

9.1 Punktweise und gleichmäßige Konvergenz

Für eine Funktionenfolge gibt es zwei zu unterscheidende Konvergenzbegriffe, die *punktweise* und *gleichmäßige* Konvergenz.

Definition 9.1.1 — PUNKTWEISE UND GLEICHMÄSSIGE KONVERGENZ

Es sei $D \subseteq \mathbb{K}_1$ und $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ eine Funktionenfolge auf D . Ferner sei $f : D \rightarrow \mathbb{K}_2$ eine Funktion.

- (i) Wir sagen $(f_n)_{n \in \mathbb{N}}$ konvergiert *punktweise* auf D gegen f wenn

$$\forall x \in D : \quad f_n(x) \xrightarrow{n \rightarrow \infty} f(x).$$

(Sprich: Für jeden Punkt $x \in D$ konvergiert die Zahlenfolge $(f_n(x))_{n \in \mathbb{N}} \subseteq \mathbb{K}_2$ gegen

$f(x)$.^a). Anders ausgedrückt:

Für jeden Punkt $x \in D$ gilt $|f_n(x) - f(x)| \xrightarrow{n \rightarrow \infty} 0$.

(ii) Wir sagen $(f_n)_{n \in \mathbb{N}}$ konvergiert *gleichmäßig* auf D gegen f wenn

$$\sup_{y \in D} |f_n(y) - f(y)| \xrightarrow{n \rightarrow \infty} 0. \quad (b)$$

(Sprich: Der maximale y -Abstand, den der Graph von f_n vom Graphen von f hat, muss gegen 0 konvergieren)

Man schreibt auch $f_n \xrightarrow{n \rightarrow \infty} f$ punktweise, bzw. $f_n \xrightarrow{n \rightarrow \infty} f$ gleichmäßig. Alternativ verwendet man auch $f_n \xrightarrow[n \rightarrow \infty]{\text{pktw}} f$ bzw. $f_n \xrightarrow[n \rightarrow \infty]{\text{glm}} f$.

Die Funktion f nennt man in beiden Fällen *Grenzfunktion* von $(f_n)_{n \in \mathbb{N}}$ und schreibt auch

(i) $f = \text{p-}\lim_{n \rightarrow \infty} f_n$ wenn $(f_n)_{n \in \mathbb{N}}$ punktweise gegen f konvergiert.

(ii) $f = \text{u-}\lim_{n \rightarrow \infty} f_n$ wenn $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen f konvergiert.

Ist die Art der Konvergenz aus dem Zusammenhang ersichtlich, so schreibt man auch einfach $f = \lim_{n \rightarrow \infty} f_n$.

Ähnlich in der Vergangenheit auch nennen wir eine Funktionenfolge $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ *punktweise konvergent* (bzw. gleichmäßig konvergent), falls es eine Funktion $f : D \rightarrow \mathbb{K}_2$ gibt, sodass $(f_n)_{n \in \mathbb{N}}$ gegen f punktweise konvergiert (bzw. gleichmäßig konvergiert).

^aDiese Konvergenz kann allerdings für jedes $x \in D$ *unterschiedlich schnell* sein.

^bDamit der Ausdruck $\sup_{y \in D} |f_n(y) - f(y)|$ als reelle Zahl Sinn ergibt, muss $\{|f_n(y) - f(y)| : y \in D\}$ eine beschränkte Teilmenge von $\mathbb{R}$ sein (vgl. Definition 3.4.1). Wir nehmen hier implizit an, dass dies zumindest für genügend große n der Fall ist, damit vom Grenzwert dieser Größe gesprochen werden kann.

Gleichmäßige Konvergenz ist stärker als punktweise Konvergenz. Dies werden wir nun zuerst zeigen. Danach werden wir aber auch sehen, dass es tatsächlich Funktionenfolgen gibt, die punktweise konvergieren aber nicht gleichmäßig.

Proposition 9.1.2 — GLEICHMÄSSIGE KONVERGENZ IMPLIZIERT PUNKTWEISE

Es konvergiere $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ gleichmäßig gegen ein $f : D \rightarrow \mathbb{K}_2$. Dann konvergiert $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ auch punktweise gegen f .

Beweis. Es sei $x \in D$ beliebig.

$$\mathbb{Z} : \quad f_n(x) \xrightarrow{n \rightarrow \infty} f(x).$$

Dazu beobachten wir zuerst

$$|f_n(x) - f(x)| \leq \sup_{y \in D} |f_n(y) - f(y)|. \quad (a)$$

Nutzt man nun, dass f_n gleichmäßig gegen f konvergiert so erhält man

$$0 \leq |f_n(x) - f(x)| \leq \sup_{y \in D} |f_n(y) - f(y)| \xrightarrow[\text{Def 9.1.1(ii)}]{n \rightarrow \infty} 0$$

Nach dem Sandwichlemma folgt man $|f_n(x) - f(x)| \xrightarrow{n \rightarrow \infty} 0$ und daher auch $f_n(x) \xrightarrow{n \rightarrow \infty} f(x)$ (vgl. Proposition 2.3.30).

^adenn $\sup_{y \in D} |f_n(y) - f(y)|$ ist nach Definition 3.2.10 eine obere Schranke von $\{|f_n(y) - f(y)| : y \in D\}$.

Als Folgerung erhält man die *Eindeutigkeit von Grenzfunktionen*. Das bedeutet, dass man nicht einfach gegen zwei verschiedene Grenzfunktionen konvergieren kann. Dies rechtfertigt auch erst die Schreibweisen $\lim_{n \rightarrow \infty} f_n$ und $\lim_{n \rightarrow \infty} f_n$.

Proposition 9.1.3 — EINDEUTIGKEIT DER GRENZFUNKTION

Es sei $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ eine Funktionenfolge und $f, g : D \rightarrow \mathbb{K}_2$ Funktionen.

- (i) Konvergiert $(f_n)_{n \in \mathbb{N}}$ punktweise sowohl gegen f als auch gegen g , so gilt $f = g$. ^a
- (ii) Konvergiert $(f_n)_{n \in \mathbb{N}}$ gleichmäßig sowohl gegen f als auch gegen g , so gilt $f = g$.

^aDas bedeutet für uns, dass $f(x) = g(x)$ für alle $x \in D$.

Beweis. Zu (i). Es sei $x \in D$. Dann gilt nach Definition 9.1.1, dass

$$f_n(x) \xrightarrow{n \rightarrow \infty} f(x)$$

und

$$f_n(x) \xrightarrow{n \rightarrow \infty} g(x).$$

Da nach Proposition 4.3.2(o) der Grenzwert von Zahlenfolgen in $\mathbb{K}_2$ eindeutig ist, muss $f(x) = g(x)$ gelten. Die Behauptung ist gezeigt.

Zu (ii). Konvergiert $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen f und g , so konvergiert nach Proposition 9.1.2 $(f_n)_{n \in \mathbb{N}}$ auch punktweise gegen f und g . Wir lernen dann aus Punkt (i), dass $f = g$. Die Behauptung folgt.

Die Umkehrung in Proposition 9.1.2 gilt übrigens nicht: Aus der punktweisen Konvergenz kann nicht die gleichmäßige Konvergenz gefolgert werden!

Beispiel 9.1.4

Es sei $f_n : (0, 1) \rightarrow \mathbb{R}, x \mapsto x^n$ gegeben. Wir behaupten:^a

$(f_n)_{n \in \mathbb{N}}$ konvergiert pktw. gegen $f : (0, 1) \rightarrow \mathbb{R}, x \mapsto 0$, ist aber nicht glm. konvergent.

In Kurzschreibweise: $f_n \xrightarrow[\text{pktw}]{n \rightarrow \infty} 0$, jedoch nicht gleichmäßig.

Beweis. Dass $f_n \xrightarrow[\text{pktw}]{n \rightarrow \infty} 0$ folgt so: Es sei $x \in (0, 1)$ beliebig. Dann kann man Proposition 6.3.16(ii) mit $\gamma = x$ anwenden und folgert

$$\lim_{n \rightarrow \infty} x^n = 0.$$

Nutzen wir dann noch die Definition von f_n und f aus so haben wir gezeigt

$$\lim_{n \rightarrow \infty} f_n(x) = f(x).$$

Nun zeigen wir noch dass $(f_n)_{n \in \mathbb{N}}$ nicht gleichmäßig konvergent ist. Wäre $(f_n)_{n \in \mathbb{N}}$ gleichmäßig konvergent gegen eine Funktion $g : (0, 1) \rightarrow \mathbb{R}$, so wäre

$$f = \text{u-} \lim_{n \rightarrow \infty} f_n \stackrel{\text{Prop 9.1.2}}{=} \text{p-} \lim_{n \rightarrow \infty} f_n = 0.$$

Schreiben wir nun mit Definition 9.1.1 aus, was die gleichmäßige Konvergenz gegen $g = 0$ bedeutet, so erhalten wir

$$0 = \lim_{n \rightarrow \infty} \sup_{y \in (0,1)} |f_n(y) - g(y)| = \lim_{n \rightarrow \infty} \sup_{y \in (0,1)} |y^n - 0| = \lim_{n \rightarrow \infty} \sup_{y \in (0,1)} |y|^n. \quad (493)$$

Nun beachte aber, dass für alle $n \in \mathbb{N}$ gilt $(1 - \frac{1}{n}) \in (0, 1)$. Somit hat man für alle $n \in \mathbb{N}$

$$\sup_{x \in (0,1)} |y|^n \geq (1 - \frac{1}{n})^n.$$

Lässt man auf beiden Seiten $n \rightarrow \infty$ streben so gilt

$$0 \stackrel{(493)}{=} \lim_{n \rightarrow \infty} \sup_{y \in (0,1)} |y|^n \geq \lim_{n \rightarrow \infty} (1 - \frac{1}{n})^n \xrightarrow[\text{Prop 7.1.13}]{n \rightarrow \infty} e^{-1}.$$

Wir erhalten den Widerspruch $0 \geq e^{-1} > 0$.

^aZu diesem Beispiel gibt es auch eine interaktive Grafik auf GEOGEBRA.

Als Beispiele für gleichmäßig konvergente Funktionenfolgen $(f_n)_{n \in \mathbb{N}}$ könnte man vieles anführen. Zum Beispiel nicht von n abhängige Funktionenfolgen... Wir rechnen mal ein interessanteres Beispiel.

Beispiel 9.1.5

Es sei $f_n : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch $x \mapsto \sqrt{x^2 + \frac{1}{n}}$.

Behauptung. $(f_n)_{n \in \mathbb{N}}$ konvergiert gleichmäßig gegen $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto |x|$.

Beweis. Sei $n \in \mathbb{N}$. Es gilt für alle $y \in \mathbb{R}$

$$\sqrt{y^2 + \frac{1}{n}} - |y| = (\sqrt{y^2 + \frac{1}{n}} - |y|) \frac{\sqrt{y^2 + \frac{1}{n}} + |y|}{\sqrt{y^2 + \frac{1}{n}} + |y|} = \frac{y^2 + \frac{1}{n} - |y|^2}{\sqrt{y^2 + \frac{1}{n}} + |y|} \stackrel{y^2 = |y|^2}{=} \frac{\frac{1}{n}}{\sqrt{y^2 + \frac{1}{n}} + |y|}.$$

Nutzt man nun, dass $y^2 \geq 0$ und $|y| \geq 0$ so erhält man für alle $y \in \mathbb{R}$

$$0 \leq \sqrt{y^2 + \frac{1}{n}} - |y| = \frac{\frac{1}{n}}{\sqrt{y^2 + \frac{1}{n}} + |y|} \leq \frac{\frac{1}{n}}{\sqrt{\frac{1}{n}}} = \sqrt{\frac{1}{n}}.$$

Man erhält

$$\sup_{y \in \mathbb{R}} |f_n(y) - f(y)| = \sup_{y \in \mathbb{R}} \left| \sqrt{y^2 + \frac{1}{n}} - |y| \right| \leq \sqrt{\frac{1}{n}} \xrightarrow{n \rightarrow \infty} 0.$$

Punktweise und gleichmäßige Konvergenz lassen sich auch *quantorisiert* aufschreiben.

Proposition 9.1.6

Es sei $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ eine Funktionenfolge und $f : D \rightarrow \mathbb{K}_2$. Dann gilt

(i) $(f_n)_{n \in \mathbb{N}}$ konvergiert punktweise gegen f genau dann wenn

$$\forall x \in D \quad \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |f_n(x) - f(x)| \leq \varepsilon.$$

(Hierbei hängt die Wahl von n_0 im Allgemeinen von ε und x ab)

(ii) $(f_n)_{n \in \mathbb{N}}$ konvergiert gleichmäßig gegen f genau dann wenn

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 \quad \forall x \in D : \quad |f_n(x) - f(x)| \leq \varepsilon.$$

(Hierbei hängt also n_0 im Allgemeinen nur von ε ab) ^a

^aUm die Abhängigkeiten genau zu verstehen können folgende Geogebra-Links hilfreich sein: [LINK1: PUNKTWEISE KONVERGENZ](#) und [LINK2: GLEICHMÄSSIGE KONVERGENZ](#)

Beweis. Zu (i).

$$f_n \xrightarrow{n \rightarrow \infty} f \quad \text{punktweise}$$

$$\Longleftrightarrow$$

$$\forall x \in D : \quad f_n(x) \xrightarrow{n \rightarrow \infty} f(x).$$

$$\stackrel{\text{Def 2.3.1}}{\Longleftrightarrow}$$

$$\forall x \in D : \quad \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad |f_n(x) - f(x)| \leq \varepsilon.$$

Zu (ii).

$$f_n \xrightarrow{n \rightarrow \infty} f \quad \text{gleichm\u00e4\u00dfig}$$

$$\Longleftrightarrow$$

$$\sup_{y \in D} |f_n(y) - f(y)| \xrightarrow{n \rightarrow \infty} 0$$

$$\stackrel{\text{Def 2.3.1}}{\Longleftrightarrow}$$

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \sup_{y \in D} |f_n(y) - f(y)| \leq \varepsilon. \quad (494)$$

Beachte nun, dass f\u00fcr eine beliebige Menge $A \subseteq \mathbb{R}$ gilt

$$\sup(A) \leq \varepsilon \quad \stackrel{\text{Def 3.2.10(i.b)}}{\Longleftrightarrow} \quad \varepsilon \text{ ist obere Schranke von } A \quad \Longleftrightarrow \quad \forall v \in A : \quad v \leq \varepsilon.$$

Wendet man das auf $A := \{|f_n(y) - f(y)| : y \in D\}$ an, so erh\u00e4lt man mit (494)

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad |f_n(x) - f(x)| \leq \varepsilon \quad \forall x \in D \quad \forall n \geq n_0.$$

Ein Cauchy-Kriterium f\u00fcr gleichm\u00e4\u00dfige Konvergenz

F\u00fcr Zahlenfolgen konnten wir den Konvergenz \u00e4quivalent \u00fcber ein *Cauchy-Kriterium* erfassen. Der Vorteil am Cauchy-Kriterium ist, dass man bei der \u00dcberpr\u00fcfung keine Kenntnis \u00fcber den Grenzwert ben\u00f6tigt. Nun wollen wir Cauchy-Kriterien f\u00fcr die (punktweise bzw. gleichm\u00e4\u00dfige) Konvergenz von Funktionenfolgen herleiten.

Bemerkung 9.1.7

Für die punktweise Konvergenz einer Funktionenfolge $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ ist es leicht, ein passendes Cauchy-Kriterium zu formulieren, weil es ja lediglich um die Konvergenz der Zahlenfolgen $(f_n(x))_{n \in \mathbb{N}} \subseteq \mathbb{K}_2$ geht. Mit anderen Worten gilt

$$\begin{aligned}
& (f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}} \text{ ist punktweise konvergent} \\
& \iff \forall x \in D : (f_n(x))_{n \in \mathbb{N}} \subseteq \mathbb{K}_2 \text{ ist konvergent} \\
& \xLeftrightarrow{\text{Prop 4.3.5}} \forall x \in D : (f_n(x))_{n \in \mathbb{N}} \subseteq \mathbb{K}_2 \text{ ist Cauchy-Folge.} \\
& \xLeftrightarrow{\text{Def 4.3.1}} \forall x \in D : \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \geq n_0 : |f_n(x) - f_m(x)| \leq \varepsilon.
\end{aligned}$$

Nun formulieren und beweisen wir auch ein Cauchy-Kriterium für die gleichmäßige Konvergenz. Wie man auf diese Formulierung kommt, werden wir später noch genauer hinterfragen.

Satz 9.1.8 — CAUCHY-KRITERIUM FÜR GLEICHMÄSSIGE KONVERGENZ

Es sei $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ Funktionenfolge so, dass

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \geq n_0 : \sup_{y \in D} |f_n(y) - f_m(y)| \leq \varepsilon. \quad (495)$$

Dann ist $(f_n)_{n \in \mathbb{N}}$ gleichmäßig konvergent, d.h. es gibt $f : D \rightarrow \mathbb{K}_2$ so, dass $f_n \xrightarrow[\text{glm}]{n \rightarrow \infty} f$.

Erfüllt eine Funktionenfolge $(f_n : D \rightarrow \mathbb{K}_2)$ die Bedingung in Gleichung (495), so sagen wir, $(f_n)_{n \in \mathbb{N}}$ erfüllt das *Cauchy-Kriterium für gleichmäßige Konvergenz*.

Beweis. Es sei $(f_n)_{n \in \mathbb{N}}$ wie in der Aussage.

SCHRITT 1. Punktweise Konvergenz

Wir zeigen zunächst etwas schwächeres, nämlich, dass $(f_n)_{n \in \mathbb{N}}$ *punktweise konvergiert*. Es sei dazu zunächst $x \in D$ beliebig. Dann gilt für alle $n, m \in \mathbb{N}$

$$|f_n(x) - f_m(x)| \leq \sup_{y \in D} |f_n(y) - f_m(y)|.$$

Wir folgern daher aus Bedingung (495), dass

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n, m \geq n_0 : |f_n(x) - f_m(x)| \leq \varepsilon.$$

Mit anderen Worten ist für alle $x \in D$ die Zahlenfolge $(f_n(x))_{n \in \mathbb{N}} \subseteq \mathbb{K}_2$ eine Cauchy-Folge und somit konvergent (vgl. Proposition 4.3.5). Wir definieren nun $f : D \rightarrow \mathbb{K}_2$ durch

$$f(x) := \lim_{n \rightarrow \infty} f_n(x) \quad \forall x \in D. \quad (496)$$

Klarerweise gilt dann $f_n \xrightarrow[\text{pktw}]{n \rightarrow \infty} f$ (da ja für alle $x \in D$ der Wert $f(x)$ gerade so gewählt ist, dass $f_n(x) \xrightarrow{n \rightarrow \infty} f(x)$).

SCHRITT 2. Gleichmäßige Konvergenz.

Nun haben wir schonmal einen Kandidaten für die Grenzfunktion gefunden, nämlich f aus Schritt 1. Noch zu zeigen ist, dass $f_n \xrightarrow[\text{glm}]{n \rightarrow \infty} f$. Dazu prüfen wir die Bedingung in Proposition 9.1.6 (ii). Sei $\varepsilon > 0$ beliebig vorgegeben.

$$\mathbb{Z}: \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \forall x \in D \quad |f_n(x) - f(x)| \leq \varepsilon. \quad (497)$$

Wir wissen aus Bedingung (495), dass es ein $n_1 \in \mathbb{N}$ gibt mit

$$\forall n, m \geq n_1 : \quad \sup_{y \in D} |f_n(y) - f_m(y)| \leq \varepsilon.$$

Wähle nun $n_0 := n_1$.

$$\mathbb{Z}: \forall n \geq n_0 \quad \forall x \in D : \quad |f_n(x) - f(x)| \leq \varepsilon.$$

Dafür sei $x \in D$ und $n \geq n_0$ beliebig (d.h. auch $n \geq n_1$). Wegen (496) gilt

$$|f_n(x) - f(x)| \stackrel{(496)}{=} |f_n(x) - \lim_{\substack{m \rightarrow \infty \\ m \geq n_1}} f_m(x)| = \lim_{\substack{m \rightarrow \infty \\ m \geq n_1}} |f_n(x) - f_m(x)| \quad (498)$$

Da $n \geq n_1$, gilt für sämtliche $m \geq n_1$, dass

$$|f_n(x) - f_m(x)| \leq \sup_{y \in D} |f_n(y) - f_m(y)| \stackrel{\substack{n, m \geq n_1 \\ (495)}}{\leq} \varepsilon.$$

Somit gilt für den Grenzwert in (498), dass

$$\lim_{\substack{m \rightarrow \infty \\ m \geq n_1}} |f_n(x) - f_m(x)| \leq \varepsilon.$$

Wiederum aus (498) folgern wir dann dass

$$|f_n(x) - f(x)| \leq \varepsilon.$$

Da $x \in D$ und $n \geq n_0$ beliebig gewählt waren ist (497) bewiesen.

Ein erstes wichtiges Korollar ist der *Weierstraß'sche M-Test*, der anwendbar ist, wenn die Folgeglieder der Funktionenfolge die Partialsummen einer Reihe sind. Man spricht dann von einer *Funktionenreihe*. Wir müssen dies hier aber nicht weiter vertiefen.

Korollar 9.1.9 — WEIERSTRASS'SCHER M-TEST

Es sei $(g_k : D \rightarrow \mathbb{K}_2)_{k \in \mathbb{N}}$ eine Funktionenfolge. Definiere für $n \in \mathbb{N}$

$$f_n : D \rightarrow \mathbb{K}_2, \quad f_n(x) := \sum_{k=1}^n g_k(x). \quad (499)$$

Ferner gebe es $(M_k)_{k \in \mathbb{N}} \subseteq \mathbb{R}$ so, dass $\sup_{y \in D} |g_k(y)| \leq M_k$ und

$$\sum_{k \geq 0} M_k \quad \text{ist konvergent.}$$

Dann ist $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ gleichmäßig konvergent und die Grenzfunktion $f : D \rightarrow \mathbb{K}_2$ erfüllt

$$\forall x \in D : \quad f(x) := \sum_{k=1}^{\infty} g_k(x). \quad (500)$$

27.05.2026

Beweis. Es sei $(f_n)_{n \in \mathbb{N}}$ gegeben wie in (499). Um die gleichmäßige Konvergenz nachzuweisen verwenden wir das Cauchy-Kriterium (Satz 9.1.8). Hierzu seien $m, n \in \mathbb{N}$ und OBdA $n > m$. Dann gilt für $y \in D$

$$|f_n(y) - f_m(y)| = \left| \sum_{k=1}^n g_k(y) - \sum_{k=1}^m g_k(y) \right| = \left| \sum_{k=m+1}^n g_k(y) \right| \leq \sum_{k=m+1}^n |g_k(y)| \leq \sum_{k=m+1}^n M_k.$$

Aus diesem Grund haben wir

$$\sup_{y \in D} |f_n(y) - f_m(y)| \leq \sum_{k=m+1}^n M_k. \quad (501)$$

Sei nun $\varepsilon > 0$. Nach dem Cauchy-Kriterium für Reihenkonvergenz (vgl Satz 5.1.1) gibt es $n_0 \in \mathbb{N}$ so, dass

$$\forall n > m \geq n_0 : \quad \left| \sum_{k=m+1}^n M_k \right| \leq \varepsilon$$

Für $n, m \geq n_0$ (OBdA $n > m \geq n_0$) gilt dann nach (501)

$$\sup_{y \in D} |f_n(y) - f_m(y)| \leq \sum_{k=m+1}^n M_k \leq \left| \sum_{k=m+1}^n M_k \right| \leq \varepsilon.$$

Die gleichmäßige Konvergenz folgt. Nun noch zu (500). Es sei $f := \text{u-}\lim_{n \rightarrow \infty} f_n$. Dann gilt $f_n \xrightarrow[\text{glm}]{n \rightarrow \infty} f$ aber somit auch (nach Proposition 9.1.2) $f_n \xrightarrow[\text{pktw}]{n \rightarrow \infty} f$. Aus diesem Grund gilt

für $x \in D$

$$f(x) = \lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \sum_{k=1}^n g_k(x) = \sum_{k=1}^{\infty} g_k(x).$$

Zunächst ein Beispiel zum Weierstraß-M-Test

Beispiel 9.1.10

Gegeben sei für $n \in \mathbb{N}$ die Funktionenfolge $(f_n : \mathbb{R} \rightarrow \mathbb{R})_{n \in \mathbb{N}}$ durch

$$f_n(x) := \sum_{k=1}^n \frac{\sin(3^k x)}{k^2}$$

Wir behaupten, dass $(f_n)_{n \in \mathbb{N}}$ gleichmäßig konvergent ist.

Beweis. Wir definieren $g_k : \mathbb{R} \rightarrow \mathbb{R}$, $g_k(x) := \frac{\sin(3^k x)}{k^2}$. Dann gilt

$$\sup_{y \in \mathbb{R}} |g_k(y)| = \sup_{y \in \mathbb{R}} \frac{|\sin(3^k y)|}{k^2} \leq \frac{1}{k^2} \quad \forall k \in \mathbb{N}.$$

Da $\sum_{k \geq 1} \frac{1}{k^2}$ konvergiert, können wir den Weierstraß-M-Test (Korollar 9.1.9) mit $M_k = \frac{1}{k^2}$ verwenden und folgern die gleichmäßige Konvergenz von $(f_n)_{n \in \mathbb{N}}$.

Der Weierstraß-M-Test hilft insbesondere dabei, Partialsummen von Potenzreihen auf gleichmäßige zu untersuchen.

Korollar 9.1.11 — GLEICHMÄSSIGE KONVERGENZ BEI POTENZREIHEN

Es sei $\sum_{k \geq 0} c_k(x - x_0)^k$ eine Potenzreihe mit Konvergenzradius $R > 0$. Sei $r \in (0, R)$ und definiere wie in Bemerkung 9.0.2 die Funktionenfolge $f_n : B_r(x_0) \rightarrow \mathbb{C}$ durch

$$f_n(x) := \sum_{k=0}^n c_k(x - x_0)^k.$$

Dann ist $(f_n)_{n \in \mathbb{N}}$ auf $B_r(x_0)$ gleichmäßig konvergent.

Beweis. Es sei $R > 0$, $r \in (0, R)$ wie in der Aussage. Wir verwenden den Weierstraß-M-Test (Korollar 9.1.9). Definiere hierzu $g_k : B_r(x_0) \rightarrow \mathbb{C}$, $x \mapsto c_k(x - x_0)^k$ für $k \in \mathbb{N}_0$. Dann gilt für $y \in B_r(x_0)$ und $k \in \mathbb{N}_0$

$$|g_k(y)| = |c_k| |y - x_0|^k = |c_k| r^k,$$

d.h. also auch $\sup_{y \in B_r(x_0)} |g_k(y)| \leq |c_k| r^k$. Da $r < R$ und da nach Proposition 5.2.6 die

Potenzreihe $\sum_{k \geq 0} |c_k| z^k$ auch denselben Konvergenzradius R hat, muss die Reihe $\sum_{k \geq 0} |c_k| r^k$ konvergent sein. Definiert man nun $M_k := |c_k| r^k$ (für alle $k \in \mathbb{N}$) so sind die Voraussetzungen von Korollar 9.1.9 erfüllt. Wir folgern, dass die Funktionenfolge $(f_n : B_r(x_0) \rightarrow \mathbb{C})_{n \in \mathbb{N}}$ gegeben durch

$$f_n(x) = \sum_{k=0}^n g_k(x) = \sum_{k=0}^n c_k (x - x_0)^k$$

gleichmäßig konvergent ist.

9.2 Vererbung der Stetigkeit

Eine wichtige Frage, die man untersuchen kann ist, was für Eigenschaften von einer konvergenten Funktionenfolge $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ auf die Grenzfunktion *vererbt* werden. Hier soll es um die Vererbung der *Stetigkeit* gehen. Das meint konkret Folgendes:

FRAGE. Es sei $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ eine Funktionenfolge, so, dass alle f_n stetige Funktionen sind. Ferner gelte $f_n \xrightarrow{n \rightarrow \infty} f$ (punktweise/gleichmäßig). Ist dann f automatisch stetig?

Bei punktweiser Konvergenz ist die Antwort auf diese Frage „Nein“.

Beispiel 9.2.1

Es sei für $n \in \mathbb{N}$ die Funktion $f_n : [0, 1] \rightarrow \mathbb{R}$ gegeben durch $f_n(x) := x^n$. Dann konvergiert $(f_n)_{n \in \mathbb{N}}$ auf $[0, 1]$ punktweise gegen $f : [0, 1] \rightarrow \mathbb{R}$ mit

$$f(x) := \begin{cases} 0 & x \in [0, 1) \\ 1 & x = 1 \end{cases}$$

Alle Funktionen in der Folge $(f_n)_{n \in \mathbb{N}}$ sind auf $[0, 1]$ stetig, aber f ist es nicht! $x_0 = 1$ ist eine Unstetigkeitsstelle!

Fazit. Bei punktweiser Konvergenz wird die Stetigkeit im Allgemeinen nicht auf die Grenzfunktion vererbt!

Wir werden gleich sehen, dass bei gleichmäßiger Konvergenz die Stetigkeit hingegen doch vererbt wird.

Beachte: Wir wissen bereits, dass die Funktionenfolge $(f_n)_{n \in \mathbb{N}}$ aus dem vorigen Beispiel nicht gleichmäßig konvergiert, denn

$$\sup_{y \in [0, 1]} |f_n(y) - f(y)| \geq \sup_{y \in (0, 1)} |f_n(y) - f(y)| = \sup_{y \in (0, 1)} |y^n - 0| \underset{\text{Bsp 9.1.4}}{\geq} e^{-1} \xrightarrow{n \rightarrow \infty} 0.$$

Satz 9.2.2 — GLEICHMÄSSIGE KONVERGENZ VERERBT STETIGKEIT

Es sei $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ eine gleichmäßig gegen ein $f : D \rightarrow \mathbb{K}_2$ konvergente Funktionenfolge. Ist für alle $n \in \mathbb{N}$ die Funktion f_n stetig auf D so ist auch f stetig auf D .

Beweis. Wir weisen die Stetigkeit mit dem Epsilon-Delta-Kriterium nach, vgl. Bemerkung 6.2.4. Es sei $x_0 \in D$. Zu zeigen ist, dass f bei x_0 stetig ist, d.h.

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x \in D \text{ mit } |x - x_0| < \delta \quad \text{gilt} \quad |f(x) - f(x_0)| < \varepsilon.$$

Sei $\varepsilon > 0$ beliebig vorgegeben.

$$\text{Z:} \quad \exists \delta > 0 : \quad \forall x \in D \text{ mit } |x - x_0| < \delta \quad \text{gilt} \quad |f(x) - f(x_0)| < \varepsilon. \quad (502)$$

Da $f_n \xrightarrow{n \rightarrow \infty} f$ gleichmäßig gilt $\sup_{y \in D} |f_n(y) - f(y)| \xrightarrow{n \rightarrow \infty} 0$ (vgl. Definition 9.1.1). Schreiben wir dies mit Definition 2.3.1 aus so erhalten wir

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \sup_{y \in D} |f_n(y) - f(y)| \leq \frac{\varepsilon}{4}. \quad (503)$$

Fixieren wir $n_0 \in \mathbb{N}$ nun wie oben. Da f_{n_0} (wie alle Funktionen aus der Folge $(f_n)_{n \in \mathbb{N}}$) stetig ist, gibt es nach Bemerkung 6.2.4 ein $\delta > 0$ so, dass

$$\forall x \in D \text{ mit } |x - x_0| < \delta \quad \text{gilt} \quad |f_{n_0}(x) - f_{n_0}(x_0)| < \frac{\varepsilon}{4}. \quad (504)$$

Es sei nun $x \in D$ mit $|x - x_0| < \delta$. Wir werden zeigen, dass $|f(x) - f(x_0)| < \varepsilon$. In der Tat gilt

$$\begin{aligned} |f(x) - f(x_0)| &= |f(x) - f_{n_0}(x) + f_{n_0}(x) - f_{n_0}(x_0) + f_{n_0}(x_0) - f(x_0)| \\ &= |(f(x) - f_{n_0}(x)) + (f_{n_0}(x) - f_{n_0}(x_0)) + (f_{n_0}(x_0) - f(x_0))| \\ &\stackrel{\Delta\text{-Ungl}}{\leq} |f(x) - f_{n_0}(x)| + |f_{n_0}(x) - f_{n_0}(x_0)| + |f_{n_0}(x_0) - f(x_0)|. \end{aligned}$$

Der erste und der dritte Summand kann durch $\sup_{y \in D} |f(y) - f_{n_0}(y)|$ abgeschätzt werden. Wir folgern

$$|f(x) - f(x_0)| \leq 2 \sup_{y \in D} |f(y) - f_{n_0}(y)| + |f_{n_0}(x) - f_{n_0}(x_0)|. \quad (505)$$

Wegen (503) ist $\sup_{y \in D} |f(y) - f_{n_0}(y)| \leq \frac{\varepsilon}{4}$. Da $|x - x_0| < \delta$ gilt auch $|f_{n_0}(x) - f_{n_0}(x_0)| < \frac{\varepsilon}{4}$, vgl. (504). Wegen (505) gilt

$$|f(x) - f(x_0)| \leq 2 \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \frac{3\varepsilon}{4} < \varepsilon.$$

Da $x \in D$ beliebig mit $|x - x_0| < \delta$ war folgt (502).

Gleichmäßige Konvergenz vererbt also Stetigkeit.

Bemerkung 9.2.3

Mithilfe von Satz 9.2.2 könnte man die Stetigkeit von Potenzreihen ganz leicht beweisen: Es sei $\sum_{k \geq 0} c_k(x - x_0)^k$ eine Potenzreihe mit Konvergenzradius $R > 0$. Sei $r \in (0, R)$ und $(f_n)_{n \in \mathbb{N}}$ definiert wie in Korollar 9.1.11, d.h.

$$f_n : B_r(x_0) \rightarrow \mathbb{R}, x \mapsto \sum_{k=0}^n c_k(x - x_0)^k.$$

Die Funktionen f_n sind für alle $n \in \mathbb{N}$ stetig, da es sich um Polynome handelt. Nach Korollar 9.1.11 konvergiert $(f_n)_{n \in \mathbb{N}}$ nun auf $B_r(x_0)$ gegen die Wertefunktion der Potenzreihe. Daher muss diese Wertefunktion auf $B_r(x_0)$ stetig sein. Da $r < R$ beliebig gewählt werden kann, folgert man auch Stetigkeit auf $B_R(x_0)$.

Nun stellt sich natürlicherweise auch die Frage ob noch andere Eigenschaften unter gleichmäßiger Konvergenz vererbt werden. Unsere viel studierte *Differenzierbarkeit* gehört allerdings nicht zu den vererbten Eigenschaften.

Bemerkung 9.2.4

Es sei $(f_n : D \rightarrow \mathbb{K}_2)_{n \in \mathbb{N}}$ eine gleichmäßig konvergente Funktionenfolge, die nur aus differenzierbaren Funktionen besteht. Dann ist die Grenzfunktion $f = \lim_{n \rightarrow \infty} f_n$ nicht automatisch differenzierbar! Als Gegenbeispiel betrachte

$$f_n : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \sqrt{x^2 + \frac{1}{n}}.$$

Dann ist für jedes $n \in \mathbb{N}$ die Funktion f_n auf ganz $\mathbb{R}$ differenzierbar (vgl. Proposition 7.1.5). Jedoch konvergiert $(f_n)_{n \in \mathbb{N}}$ nach Beispiel 9.1.5 gegen die nicht-differenzierbare Funktion $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto |x|$.

9.3 Vertauschung von Grenzprozessen

Als nächstes soll es um *Vertauschungsergebnisse* gehen. Vertauschungsergebnisse handeln zum Beispiel von der Vertauschung von *Integralen* und *(Funktions-)Grenzwerten*. Das Spannende daran ist, dass man hierbei zwei verschiedene Grenzprozesse hat, deren Interaktion man verstehen muss. Was wir genau damit meinen schreiben wir im Folgenden auf.

Vertauschung von Integral und Funktionsgrenzwert

Proposition 9.3.1 — VERTAUSCHUNG VON INTEGRAL UND FUNKTIONSGRENZWERT

Es seien $a, b \in \mathbb{R}$, $a < b$ und $(f_n : [a, b] \rightarrow \mathbb{R})_{n \in \mathbb{N}}$ eine Folge Riemann-integrierbarer Funktionen, die gleichmäßig gegen eine Riemann-integrierbare Funktion $f : [a, b] \rightarrow \mathbb{R}$

konvergiert. Dann gilt

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) \, dx = \int_a^b f(x) \, dx.$$

Würde man dies als Vertauschungsergebnis formulieren, so würde man schreiben

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) \, dx = \int_a^b (\lim_{n \rightarrow \infty} f_n)(x) \, dx.$$

Beweis (von Proposition 9.3.1). Es sei $(f_n)_{n \in \mathbb{N}}$ und f wie in der Aussage. Dann gilt mithilfe der Dreiecksungleichung für Integrale (vgl. Korollar 8.2.6)

$$\left| \int_a^b f_n(x) \, dx - \int_a^b f(x) \, dx \right| \stackrel{\text{Satz 8.2.4}}{=} \left| \int_a^b (f_n(x) - f(x)) \, dx \right| \leq \int_a^b |f_n(x) - f(x)| \, dx. \quad (506)$$

Nun kann man für jedes $x \in [a, b]$ abschätzen, dass

$$|f_n(x) - f(x)| \leq \sup_{y \in D} |f_n(y) - f(y)|.$$

Die Monotonie des Integrals (vgl. Proposition 8.2.5) liefert dann

$$\begin{aligned} \int_a^b |f_n(x) - f(x)| \, dx &\leq \int_a^b \sup_{y \in [a, b]} |f_n(y) - f(y)| \, dx \\ &= \sup_{y \in [a, b]} |f_n(y) - f(y)| \int_a^b 1 \, dx = (b - a) \sup_{y \in [a, b]} |f_n(y) - f(y)|. \end{aligned} \quad (507)$$

Mit (506) schließen wir

$$0 \leq \left| \int_a^b f_n(x) \, dx - \int_a^b f(x) \, dx \right| \stackrel{(506)}{\leq} \int_a^b |f_n(x) - f(x)| \, dx \stackrel{(507)}{\leq} (b - a) \sup_{y \in [a, b]} |f_n(y) - f(y)|.$$

Wegen der gleichmäßigen Konvergenz von $(f_n)_{n \in \mathbb{N}}$ gegen f konvergiert die rechte Seite der obigen Gleichung gegen 0 für $n \rightarrow \infty$ (vgl. Definition 9.1.1). Es folgt

$$\lim_{n \rightarrow \infty} \left| \int_a^b f_n(x) \, dx - \int_a^b f(x) \, dx \right| = 0.$$

Daraus folgt direkt mit Proposition 2.3.30, dass $\lim_{n \rightarrow \infty} \left(\int_a^b f_n(x) \, dx - \int_a^b f(x) \, dx \right) = 0$ und somit (nach den Grenzwertregeln)

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) \, dx = \int_a^b f(x) \, dx,$$

was zu zeigen war.

Voraussetzung für das obige Vertauschungsergebnis in Proposition 9.3.1 ist die gleichmäßige Konvergenz der Integranden $(f_n)_{n \in \mathbb{N}}$. Konvergieren die Integranden $(f_n)_{n \in \mathbb{N}}$ nur punktweise, so kann diese Vertauschung nicht garantiert werden!

Beispiel 9.3.2

Definiere für $k \in \mathbb{N}, k \geq 2$ die Funktion $f_k : [0, 1] \rightarrow \mathbb{R}$

$$f_k(x) := \begin{cases} k & \text{falls } x \in (\frac{1}{k}, \frac{2}{k}) \\ 0 & \text{sonst} \end{cases} \quad \text{siehe Abbildung 40.} \quad (508)$$

Man kann sich davon überzeugen, dass f_k für jedes $k \in \mathbb{N}$ Riemann-integrierbar ist.

Behauptung. $(f_k)_{k \in \mathbb{N}}$ konvergiert punktweise (aber nicht gleichmäßig) gegen $f : [0, 1] \rightarrow \mathbb{R}, x \mapsto 0$. Jedoch gilt

$$\int_0^1 f_k(x) \, dx \xrightarrow{k \rightarrow \infty} \int_0^1 f(x) \, dx.$$

Beweis.

TEIL 1. $f_k \xrightarrow{k \rightarrow \infty} f$ punktweise, d.h. für alle $x \in [0, 1]$ gilt $\lim_{k \rightarrow \infty} f_k(x) = 0$.

Es sei $x \in [0, 1]$ beliebig. Falls $x = 0$ so gilt $f_k(x) = 0$ für alle $k \in \mathbb{N}$, vgl. (508). Also gilt auch

$$\lim_{k \rightarrow \infty} f_k(x) = \lim_{k \rightarrow \infty} 0 = 0.$$

Falls $x \neq 0$ (d.h. $x > 0$) so gibt es ein $k_0 \in \mathbb{N}$ mit $\frac{2}{k_0} < x$.^a Für alle $k \geq k_0$ gilt dann nach (508), dass $f_k(x) = 0$. Da die ersten k_0 Folgenglieder für das Grenzverhalten keine Rolle spielen gilt

$$\lim_{k \rightarrow \infty} f_k(x) = \lim_{\substack{k \rightarrow \infty \\ k \geq k_0}} f_k(x) = \lim_{\substack{k \rightarrow \infty \\ k \geq k_0}} 0 = 0.$$

$$\text{TEIL 2.} \quad \int_0^1 f_k(x) \, dx \xrightarrow{k \rightarrow \infty} \int_0^1 f(x) \, dx.$$

Für $k \in \mathbb{N}, k \geq 2$ gilt nach der Teilungsformel für Riemann-Integrale (Proposition 8.2.12)

$$\int_0^1 f_k(x) \, dx = \int_0^{\frac{1}{k}} f_k(x) \, dx + \int_{\frac{1}{k}}^{\frac{2}{k}} f_k(x) \, dx + \int_{\frac{2}{k}}^1 f_k(x) \, dx. \quad (509)$$

Das erste und das dritte Integral verschwindet, da $f_k(x) = 0$ für alle $x \in [0, \frac{1}{k}] \cup [\frac{2}{k}, 1]$. Beim zweiten Integral ist es so, dass $f_k(x) = k$ für alle $x \in (\frac{1}{k}, \frac{2}{k})$, lediglich an den Randpunkten nimmt das Integral einen anderen Wert an. Nach Proposition 8.5.4 stimmt das zweite Integral allerdings mit dem uneigentlichen Integral von $f_k|_{(\frac{1}{k}, \frac{2}{k})}$ überein. Für dieses spielt das Verhalten an den Randpunkten keine Rolle. Daher erhält man

$$\int_{\frac{1}{k}}^{\frac{2}{k}} f_k(x) \, dx = \int_{\frac{1}{k}}^{\frac{2}{k}} k \, dx = k \left(\frac{2}{k} - \frac{1}{k} \right) = k \cdot \frac{1}{k} = 1.$$

Mit (509) erhalten wir für alle $k \in \mathbb{N}, k \geq 2$

$$\int_0^1 f_k(x) \, dx = 1.$$

Somit gilt auch

$$\lim_{k \rightarrow \infty} \int_0^1 f_k(x) \, dx = 1 \neq 0 = \int_0^1 f(x) \, dx.$$

Die Behauptung folgt.

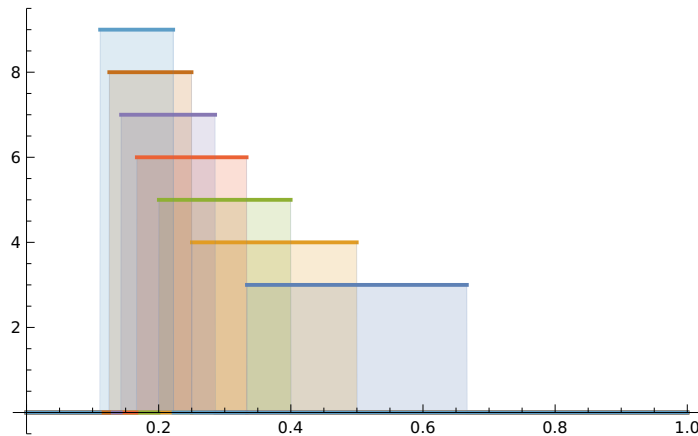
TEIL 3. $(f_k)_{k \in \mathbb{N}}$ konvergiert nicht gleichmäßig gegen f .

Würde $f_k \xrightarrow{k \rightarrow \infty} f$ gleichmäßig konvergieren, so würde nach Proposition 9.3.1 auch gelten, dass

$$\lim_{k \rightarrow \infty} \int_0^1 f_k(x) \, dx = \int_0^1 f(x) \, dx.$$

Wie haben aber bereits in Teil 2 gesehen, dass dies nicht der Fall ist.

^aMan wähle dafür etwa $k_0 := \lceil \frac{2}{x} \rceil + 1$.



Abbildungung 40: Die Funktionen f_k aus Beispiel 9.3.2 für $k = 3, \dots, 9$

Beispiel 9.3.3

Wir berechnen

$$\lim_{n \rightarrow \infty} \int_0^1 \arctan(\cos(\frac{x^4}{n})) \, dx$$

Eine Stammfunktion der Funktion im Integranden zu bestimmen ist sehr schwierig! Daher können wir das Integral nicht explizit berechnen. Und trotzdem lässt sich der Grenzwert mithilfe von Proposition 9.3.1 berechnen. Dazu definieren wir für $n \in \mathbb{N}$ die Funktion $f_n : [0, 1] \rightarrow \mathbb{R}, x \mapsto \arctan(\cos(\frac{x^4}{n}))$ und beweisen folgende Konvergenz:

$$\text{ZWISCHENBEHAUPTUNG. } f_n \xrightarrow{n \rightarrow \infty} \frac{\pi}{4} \text{ gleichmäßig auf } [0, 1]. \quad (a)$$

Wir müssen also zeigen, dass

$$\sup_{y \in [0, 1]} |f_n(y) - \frac{\pi}{4}| \xrightarrow{n \rightarrow \infty} 0$$

Wir nutzen dafür, dass für alle $z_1, z_2 \in \mathbb{R}$ gilt

$$|\arctan(z_1) - \arctan(z_2)| \leq |z_1 - z_2| \quad (510)$$

und

$$|\cos(z_1) - \cos(z_2)| \leq |z_1 - z_2|. \quad (511)$$

Beide Ungleichungen folgen direkt aus dem ersten Mittelwertsatz (Satz 7.3.1, insb. (351)) und der Tatsache, dass $|\arctan'(x)| = \frac{1}{1+x^2} \leq 1$ für alle $x \in \mathbb{R}$ sowie $|\cos'(x)| = |-\sin(x)| \leq 1$ für alle $x \in \mathbb{R}$. Es sei nun $x \in [0, 1]$. Dann gilt

$$|f_n(y) - \frac{\pi}{4}| = |\arctan(\cos(\frac{y^4}{n})) - \arctan(1)| \underset{(510)}{\leq} |\cos(\frac{y^4}{n}) - 1| = |\cos(\frac{y^4}{n}) - \cos(0)| \underset{(511)}{\leq} |\frac{y^4}{n} - 0|.$$

Nun gilt weiterhin für alle $y \in [0, 1]$

$$|\frac{y^4}{n} - 0| = \frac{y^4}{n} \leq \frac{1^4}{n} = \frac{1}{n}.$$

Die beiden vorigen Gleichungen zusammen ergeben

$$\sup_{y \in [0, 1]} |f_n(y) - \frac{\pi}{4}| \leq \frac{1}{n} \quad \forall n \in \mathbb{N}.$$

Mit dem Sandwichlemma folgt

$$\lim_{n \rightarrow \infty} \sup_{y \in [0, 1]} |f_n(y) - \frac{\pi}{4}| = 0.$$

Die Zwischenbehauptung folgt direkt aus Definition 9.1.1.

ZWISCHENBEHAUPTUNG ENDE.

Nun gilt nach Proposition 9.3.1, dass

$$\lim_{n \rightarrow \infty} \int_0^1 \arctan(\cos(\frac{x^4}{n})) \, dx = \lim_{n \rightarrow \infty} \int_0^1 f_n(x) \, dx \stackrel{\text{Prop 9.3.1}}{=} \int_0^1 \frac{\pi}{4} \, dx = \frac{\pi}{4}(1 - 0) = \frac{\pi}{4}.$$

^aWenn man es genau nimmt müsste man eigentlich schreiben $f_n \xrightarrow{n \rightarrow \infty} f$, wobei $f : [0, 1] \rightarrow \mathbb{R}, x \mapsto \frac{\pi}{4}$. Es stellt sich noch die Frage, wie man an dieser Stelle auf die Grenzfunktion $\frac{\pi}{4}$ kommt. Eine Möglichkeit dafür ist, zunächst erstmal die punktweise Konvergenz zu untersuchen. Für alle $x \in [0, 1]$ gilt $\frac{x^4}{n} \xrightarrow{n \rightarrow \infty} 0$. Wegen der Stetigkeit von $\arctan \circ \cos$ bei 0 gilt dann

$$\lim_{n \rightarrow \infty} f_n(x) = \lim_{n \rightarrow \infty} \arctan(\cos(\frac{x^4}{n})) = \arctan(\cos(0)) = \arctan(1) = \frac{\pi}{4}.$$

Der Vertauschungssatz aus Proposition 9.3.1 kann nicht auf uneigentliche Integrale angewendet werden! Die *eigentliche* Riemann-Integrierbarkeit aller Folgenglieder f_n ist eine wichtige Voraussetzung.

Beispiel 9.3.4

Es sei $(a, b) = (0, \infty)$ und

$$f_n(x) := \begin{cases} \frac{1}{n} & \text{falls } x \in (0, n) \\ 0 & \text{sonst.} \end{cases}$$

Dann ist f_n auf $(0, \infty)$ uneigentlich Riemann-integrierbar (siehe z.B. Übungsaufgabe 4 auf Blatt 6). Ferner konvergiert f_n auf $(0, \infty)$ gleichmäßig gegen 0, denn

$$\sup_{x \in (0, \infty)} |f_n(x) - 0| = \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0.$$

Jedoch gilt

$$\int_0^\infty f_n(x) \, dx = \int_0^n f_n(x) \, dx = \int_0^n \frac{1}{n} \, dx = \frac{1}{n} \cdot n = 1.$$

Also gilt

$$\lim_{n \rightarrow \infty} \int_0^\infty f_n(x) \, dx = 1 \neq 0 = \int_0^\infty 0 \, dx = \int_0^\infty (\text{u-}\lim_{n \rightarrow \infty} f_n)(x) \, dx.$$

Man sieht: Obwohl $(f_n)_{n \in \mathbb{N}}$ gleichmäßig konvergiert, können Integral und Grenzübergang nicht vertauscht werden. Beachte noch einmal: Dies ist kein Widerspruch zu Proposition 9.3.1, weil diese nicht für uneigentliche Integrale anwendbar ist!.

01.06.2026

Vertauschung von Ableitung und Funktionsgrenzwert

Wie wir im vorigen gesehen haben, lassen sich Integrale und Funktionsgrenzwerte bei gleichmäßiger Konvergenz vertauschen. Jetzt untersuchen wir dieselbe Frage für Ableitungen. Hierbei

gibt es aber Probleme: Gleichmäßige Konvergenz vererbt die Differenzierbarkeit nämlich nicht auf die Grenzfunktion, vgl. Bemerkung 9.2.4. Wer die Grenzfunktion nicht ableiten kann, kann auch Ableitung und Grenzwertbildung nicht vertauschen! Dies legt nahe, dass andere Voraussetzungen als nur die gleichmäßige Konvergenz benötigt werden

Proposition 9.3.5

Es sei $I \subseteq \mathbb{R}$ ein Intervall und $(f_n : I \rightarrow \mathbb{R})_{n \in \mathbb{N}}$ eine Folge stetig differenzierbarer Funktionen, so dass

- $(f_n)_{n \in \mathbb{N}}$ ist auf I punktweise konvergent,
- $(f'_n)_{n \in \mathbb{N}}$ ist auf I gleichmäßig konvergent.

Dann ist $f := \text{p-}\lim_{n \rightarrow \infty} f_n$ auch stetig differenzierbar und

$$f'(x) = \lim_{n \rightarrow \infty} f'_n(x) \quad \forall x \in I. \quad (512)$$

Würde man (512) wieder als Vertauschungsergebnis aufschreiben wollen, so würde man sagen

$$(\text{p-}\lim_{n \rightarrow \infty} f_n)'(x) = \lim_{n \rightarrow \infty} f'_n(x) \quad \forall x \in I.$$

Wir weisen hier abermals darauf hin, dass diese Gleichung aber nur unter den Voraussetzungen von Proposition 9.3.5 (d.h. vor allem bei gleichmäßiger Konvergenz der Ableitungen) gilt.

Beweis (von Proposition 9.3.5). Es sei $(f_n)_{n \in \mathbb{N}}$ wie in der Aussage. Definiere $f := \text{p-}\lim_{n \rightarrow \infty} f_n$ und $g := \text{u-}\lim_{n \rightarrow \infty} f'_n$. Wähle einen beliebigen Punkt $x_0 \in I$. Beachte: $f'_n : I \rightarrow \mathbb{R}$ ist eine stetige Funktion und besitzt f_n als Stammfunktion. Nach Bemerkung 8.3.1 gilt für alle $x \in I$

$$f_n(x) - f_n(x_0) = \int_{x_0}^x f'_n(t) \, dt.$$

Umgeformt heißt das

$$f_n(x) = f_n(x_0) + \int_{x_0}^x f'_n(t) \, dt \quad \forall x \in I. \quad (513)$$

Wir wollen nun das Grenzverhalten dieses Ausdrucks für $n \rightarrow \infty$ verstehen. Sei dazu $x \in I$ beliebig aber fest. Da $f = \text{p-}\lim_{n \rightarrow \infty} f_n$ gilt

$$\lim_{n \rightarrow \infty} f_n(x) = f(x) \quad (514)$$

und auch

$$\lim_{n \rightarrow \infty} f_n(x_0) = f(x_0). \quad (515)$$

Wir erinnern uns nun, dass $(f'_n)_{n \in \mathbb{N}}$ gleichmäßig gegen g konvergiert. Da die Folgenglieder f'_n für alle $n \in \mathbb{N}$ stetig sind^a, ist g nach Satz 9.2.2 auch stetig und somit

auch Riemann-Integrierbar (vgl. Proposition 8.1.22). Wendet man Proposition 9.3.1 mit $[a, b] = [\min\{x, x_0\}, \max\{x, x_0\}]$ an so folgt man

$$\lim_{n \rightarrow \infty} \int_{x_0}^x f'_n(t) \, dt = \int_{x_0}^x g(t) \, dt. \quad (516)$$

Die Gleichungen (514), (515) und (516) erlauben es, auf beiden Seiten von (513) zum Grenzwert $n \rightarrow \infty$ überzugehen. Man erhält für alle $x \in I$

$$f(x) = f(x_0) + \int_{x_0}^x g(t) \, dt. \quad (517)$$

Nach dem Fundamentalsatz der Analysis (Satz 8.3.1) ist nun die Abbildung $x \mapsto \int_{x_0}^x g(t) \, dt$ auf I differenzierbar^b mit

$$\frac{d}{dx} \int_{x_0}^x g(t) \, dt = g(x) \quad \forall x \in I.$$

Gleichung (517) und die Summenregel des Differenzierens implizieren dann, dass auch f auf I differenzierbar ist und

$$f'(x) = g(x) \quad \forall x \in I. \quad (518)$$

Da $(f'_n)_{n \in \mathbb{N}}$ gleichmäßig auf I gegen g konvergiert, konvergiert es auch punktweise gegen g (vgl. Proposition 9.1.2). Daher gilt nach Definition 9.1.1, dass $g(x) = \lim_{n \rightarrow \infty} f'_n(x)$ für alle $x \in I$. Nutzt man dies in (518) so erhält man

$$f'(x) = \lim_{n \rightarrow \infty} f'_n(x) \quad \forall x \in I.$$

Somit ist auch (512) gezeigt und die Behauptung folgt.

^aHier wird benutzt, dass f_n stetig differenzierbar ist.

^bFür den Fundamentalsatz der Analysis wurde noch einmal benutzt, dass g stetig ist!

10 Mehrdimensionale Analysis

Im Folgenden wollen wir die Analysis, die wir vorher für Folgen / Funktionen auf $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ betrieben haben nun *mehrdimensional*, d.h. auf $\mathbb{K}^d$ für ein $d \in \mathbb{N}$ betreiben. Der Einfachheit halber ignorieren wir in diesem Abschnitt die komplexen Zahlen und beschränken uns auf

$$\mathbb{R}^d := \mathbb{R} \times \dots \times \mathbb{R} := \{(x^{(1)}, \dots, x^{(d)}) : x^{(i)} \in \mathbb{R} \ \forall i = 1, \dots, d\}. \quad (519)$$

Elemente aus $\mathbb{R}^d$ nennen wir *Vektoren*. Wir werden Sie im Folgenden nicht mehr als Tupel reeller Zahlen wie in (519) sondern in *Spaltenvektornotation* schreiben, d.h.

$$x = \begin{pmatrix} x^{(1)} \\ \vdots \\ x^{(d)} \end{pmatrix}. \quad (520)$$

Den Eintrag $x^{(i)}$ nennen wir *i-te Komponente* von x . Eine besondere Rolle spielen der *Nullvektor*

$$0 := 0_{\mathbb{R}^d} := \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

und die *kanonischen Einheitsvektoren*

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad e_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots, \quad e_d := \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}.$$

Da wir häufig mit *Spaltenvektoren* arbeiten werden schreiben wir für den Spaltenvektor x aus (520) auch häufig als Kurznotation $x = (x^{(1)}, \dots, x^{(d)})^T$.

Desweiteren können wir Vektoren aus $\mathbb{R}^d$ *addieren* und *mit einem Skalar* $\lambda \in \mathbb{R}$ *multiplizieren*, nämlich wie folgt: Ist

$$x = \begin{pmatrix} x^{(1)} \\ \vdots \\ x^{(d)} \end{pmatrix} \quad \text{und} \quad y = \begin{pmatrix} y^{(1)} \\ \vdots \\ y^{(d)} \end{pmatrix}$$

so definieren wir

$$x + y := \begin{pmatrix} x^{(1)} + y^{(1)} \\ \vdots \\ x^{(d)} + y^{(d)} \end{pmatrix}$$

und

$$\lambda x := \lambda \cdot x := \begin{pmatrix} \lambda x^{(1)} \\ \vdots \\ \lambda x^{(d)} \end{pmatrix}$$

Zwei Vektoren miteinander multiplizieren oder gar durcheinander teilen kann man allerdings nicht!

Für Vektoren darf man dann natürlich auch Dinge wie das *Summenzeichen* komponentenweise definieren. Somit kann man einen Vektor $x = (x^{(1)}, \dots, x^{(d)})^T$ auch schreiben als

$$x = \sum_{i=1}^d x^{(i)} e_i.$$

10.1 Die Euklidische Norm

Während die obigen Begrifflichkeiten größtenteils aus der linearen Algebra bekannt vorkommen, gibt es doch ein Objekt, welches es sich für die mehrdimensionale Analysis genauer zu studieren lohnt, nämlich die *Euklidische Norm* und deren Eigenschaften. Anschaulich ist die Euklidische Norm der Abstand eines Vektors x zum Koordinatenursprung $(0, \dots, 0)^T$.

Definition 10.1.1 — EUKLIDISCHE NORM

Für einen Vektor $x = (x^{(1)}, \dots, x^{(d)})^T \in \mathbb{R}^d$ wie in (520) definieren wir die *Euklidische Norm* durch

$$\|x\| := \|x\|_{\mathbb{R}^d} := \sqrt{\sum_{i=1}^d |x^{(i)}|^2} = \sqrt{|x^{(1)}|^2 + \dots + |x^{(d)}|^2}. \quad (a) \quad (521)$$

^adie reellen Absolutbeträge auf der rechten Seite dieser Gleichung werden hier nicht zwingend benötigt, schließlich ist $|x^{(i)}|^2 = (x^{(i)})^2$ für alle $i = 1, \dots, d$

Die Euklidische Norm verallgemeinert den reellen Absolutbetrag $|x| = \sqrt{|x|^2}$ (für $x \in \mathbb{R}$). Sie hat auch ähnliche Eigenschaften.

Satz 10.1.2 — EIGENSCHAFTEN DER EUKLIDISCHEN NORM

Es seien $x, y \in \mathbb{R}^d$ und $\lambda \in \mathbb{R}$. Dann gilt

- (i) *Definitheit.* $\|x\| \geq 0$ und $(\|x\| = 0 \iff x = 0_{\mathbb{R}^d})$.
- (ii) *Homogenität.* $\|\lambda x\| = |\lambda| \|x\|$.^a
- (iii) *Dreiecksungleichung.* $\|x + y\| \leq \|x\| + \|y\|$.

^aHierbei ist mit $|\lambda|$ wie immer der reelle Absolutbetrag von λ gemeint

Die Punkte (i) und (ii) könnten wir leicht mit der Definition der Euklidischen Norm beweisen. Für (iii) benötigen wir ein vorbereitendes Lemma.

Lemma 10.1.3 — CAUCHY-SCHWARZ-UNGLEICHUNG

Es sei $n \in \mathbb{N}$ sowie $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{R}$. Dann gilt

$$\left| \sum_{k=1}^n a_k b_k \right| \leq \sqrt{\sum_{k=1}^n a_k^2} \sqrt{\sum_{k=1}^n b_k^2}.$$

Beweis. ObdA gilt $\sum_{k=1}^n a_k^2 > 0$ und $\sum_{k=1}^n b_k^2 > 0$. (Ist nämlich eine der beiden Summen gleich 0 so gilt entweder $a_1 = \dots = a_n = 0$ oder $b_1 = \dots = b_n = 0$. In diesem Fall ist dann auch jeder der Summanden der Summe auf der linken Seite gleich 0. Die Ungleichung ist damit erfüllt, da $0 \leq 0$.) Definiere nun

$$\theta := \frac{\sum_{k=1}^n a_k b_k}{\sum_{k=1}^n a_k^2}. \quad (522)$$

Nun beobachten wir

$$\begin{aligned} 0 &\leq \sum_{k=1}^n (b_k - \theta a_k)^2 \\ &= \sum_{k=1}^n (b_k^2 - 2\theta b_k a_k + \theta^2 a_k^2) \\ &= \sum_{k=1}^n b_k^2 - 2\theta \sum_{k=1}^n b_k a_k + \theta^2 \sum_{k=1}^n a_k^2 \\ &\stackrel{(522)}{=} \sum_{k=1}^n b_k^2 - 2 \frac{\sum_{k=1}^n a_k b_k}{\sum_{k=1}^n a_k^2} \underbrace{\sum_{k=1}^n b_k a_k}_{=\sum_{k=1}^n a_k b_k} + \frac{\left(\sum_{k=1}^n a_k b_k\right)^2}{\left(\sum_{k=1}^n a_k^2\right)^2} \sum_{k=1}^n a_k^2. \end{aligned}$$

Kürzen wir im letzten Summanden einmal $\sum_{k=1}^n a_k^2$ weg, so erhalten wir

$$0 \leq \sum_{k=1}^n b_k^2 - 2 \frac{\left(\sum_{k=1}^n a_k b_k\right)^2}{\sum_{k=1}^n a_k^2} + \frac{\left(\sum_{k=1}^n a_k b_k\right)^2}{\sum_{k=1}^n a_k^2}$$

Das bedeutet

$$0 \leq \sum_{k=1}^n b_k^2 - \frac{\left(\sum_{k=1}^n a_k b_k\right)^2}{\sum_{k=1}^n a_k^2}.$$

Umgeformt erhält man

$$\frac{\left(\sum_{k=1}^n a_k b_k\right)^2}{\sum_{k=1}^n a_k^2} \leq \sum_{k=1}^n b_k^2.$$

Multipliziert man mit $\sum_{k=1}^n a_k^2$ so folgt

$$\left(\sum_{k=1}^n a_k b_k\right)^2 \leq \left(\sum_{k=1}^n a_k^2\right) \left(\sum_{k=1}^n b_k^2\right).$$

Durch Wurzelziehen folgert man die Behauptung.

Nun sind wir in der Lage, die Eigenschaften der Norm in Satz 10.1.2 zu beweisen.

Beweis (von Satz 10.1.2). Zu (i). Es sei $x = (x^{(1)}, \dots, x^{(d)})^T \in \mathbb{R}^d$. Nach (521) gilt

$$\|x\|^2 = (x^{(1)})^2 + \dots + (x^{(d)})^2 \quad (523)$$

Insbesondere gilt

$$\|x\|^2 \geq (x^{(i)})^2 \quad \forall i \in \{1, \dots, d\}. \quad (524)$$

Aus (523) und Proposition 3.2.20(x) folgern wir $\|x\| \geq 0$ für alle $x \in \mathbb{R}^d$. Ist ferner $\|x\| = 0$ so gilt nach (524), dass $x^{(i)} = 0$ für alle $i \in \{1, \dots, d\}$. Mit anderen Worten gilt $x = 0$. Damit kann man Teil (i) der Behauptung folgern.

Zu (ii). Es gilt für $\lambda \in \mathbb{R}$ und $x = (x^{(1)}, \dots, x^{(d)})^T \in \mathbb{R}^d$, dass $\lambda x = (\lambda x^{(1)}, \dots, \lambda x^{(d)})^T$. Daher gilt

$$\|\lambda x\|^2 \stackrel{(521)}{=} (\lambda x^{(1)})^2 + \dots + (\lambda x^{(d)})^2 = \lambda^2 [(x^{(1)})^2 + \dots + (x^{(d)})^2] = \lambda^2 \|x\|^2.$$

Die Behauptung folgt durch Wurzelziehen.

Zu (iii). Es seien nun $x = (x^{(1)}, \dots, x^{(d)})^T$ und $y = (y^{(1)}, \dots, y^{(d)})^T$ gegeben. Dann gilt

$$\begin{aligned} \|x + y\|^2 &= \sum_{i=1}^d (x^{(i)} + y^{(i)})^2 = \sum_{i=1}^d (x^{(i)})^2 + 2x^{(i)}y^{(i)} + (y^{(i)})^2 \\ &= \sum_{i=1}^d (x^{(i)})^2 + 2 \sum_{i=1}^d x^{(i)}y^{(i)} + \sum_{i=1}^d (y^{(i)})^2. \end{aligned} \quad (525)$$

Der erste Summand ist gerade $\|x\|^2$, der dritte Summand ist $\|y\|^2$. Den zweiten Summanden

können wir mit der Cauchy-Schwarz-Ungleichung nach oben abschätzen. Dazu:

$$2 \sum_{i=1}^d x^{(i)} y^{(i)} \leq 2 \left| \sum_{i=1}^d x^{(i)} y^{(i)} \right| \stackrel{\text{Lem 10.1.3}}{\leq} 2 \sqrt{\sum_{i=1}^d (x^{(i)})^2} \sqrt{\sum_{i=1}^d (y^{(i)})^2} = 2 \|x\| \|y\|.$$

Dies und (525) liefern

$$\|x + y\|^2 \leq \|x\|^2 + 2 \|x\| \|y\| + \|y\|^2 = (\|x\| + \|y\|)^2.$$

Die Behauptung folgt durch Wurzelziehen.

Mithilfe der Euklidischen Norm können wir auch wieder einmal *r-Umgebungen* im Mehrdimensionalen definieren.

Definition 10.1.4 — *r*-UMGEBUNG

Es sei $r > 0$ gegeben und $x_0 \in \mathbb{R}^d$. Die *r-Umgebung* um x_0 (bzw. *r-Ball* um x_0) ist gegeben durch die Menge

$$B_r(x_0) := \{x \in \mathbb{R}^d : \|x - x_0\| < r\}.$$

Wenn die Dimension d des Balles nicht aus dem Kontext klar ist, schreiben wir auch oft $B_r^d(x_0)$.

10.2 Mehrdimensionale Folgen

Wie auch in der eindimensionalen Analysis beginnen wir unser Studium mit Folgen, diesmal mit Folgen $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$. Beachte: Eine Folge in $\mathbb{R}^d$ lässt sich eigentlich als d reelle Folgen auffassen, denn für $n \in \mathbb{N}$ kann man den Vektor

$$a_n = \begin{pmatrix} a_n^{(1)} \\ \vdots \\ a_n^{(d)} \end{pmatrix}$$

in seine einzelnen Komponenten ausschreiben und jede Komponente die *Komponentenfolgen* $(a_n^{(i)})_{n \in \mathbb{N}} \subseteq \mathbb{R}$ als Folge in $\mathbb{R}$ betrachten.

Definition 10.2.1 — KOMPONENTENFOLGEN

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ eine Folge, etwa $a_n = (a_n^{(1)}, \dots, a_n^{(d)})^T$. Für $i \in \{1, \dots, d\}$ und $n \in \mathbb{N}$ sei $a_n^{(i)}$ die *i*-te Komponente von a_n . Dann heißt $(a_n^{(i)})_{n \in \mathbb{N}}$ die *i*-te *Komponentenfolge*.

Nun können wir intuitiv einen *Konvergenzbegriff* für Folgen in $\mathbb{R}^d$ definieren.

Definition 10.2.2 — KONVERGENZ MEHRDIMENSIONALER FOLGEN

Wir sagen eine Folge $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ *konvergiert* gegen ein $a = (a^{(1)}, \dots, a^{(d)})^T \in \mathbb{R}^d$ wenn für alle $i \in \{1, \dots, d\}$ die *i*-te Komponentenfolge $(a_n^{(i)})_{n \in \mathbb{N}} \subseteq \mathbb{R}$ gegen $a^{(i)}$ konvergiert.

In diesem Fall schreiben wir $a_n \xrightarrow{n \rightarrow \infty} a$ oder $\lim_{n \rightarrow \infty} a_n = a$. Der Vektor a heißt dann *Grenzwert* von $(a_n)_{n \in \mathbb{N}}$.

Direkt aus dieser Definition leiten sich die folgenden *Rechenregeln für mehrdimensionale Grenzwerte* ab.

Proposition 10.2.3 — GRENZWERTREGELN IM MEHRDIMENSIONALEN

Es seien $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ und $(b_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ konvergente Folgen. Dann gilt

- (o) Der Grenzwert $\lim_{n \rightarrow \infty} a_n$ ist eindeutig bestimmt und ändert sich nicht, wenn man nur endlich viele Folgenglieder abändert.
- (i) $(a_n + b_n)_{n \in \mathbb{N}}$ ist konvergent und $\lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n$.
- (ii) Ist $(\lambda_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ in $\mathbb{R}$ konvergent, so ist $(\lambda_n a_n)_{n \in \mathbb{N}}$ konvergent und $\lim_{n \rightarrow \infty} \lambda_n a_n = \left(\lim_{n \rightarrow \infty} \lambda_n \right) \left(\lim_{n \rightarrow \infty} a_n \right)$.
- (iii) Ist $(c_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ beschränkt und gilt $\lim_{n \rightarrow \infty} a_n = 0$, so gilt auch $\lim_{n \rightarrow \infty} c_n a_n = 0$

Beweis. Die Aussagen (o),(i),(ii),(iii) folgen direkt aus der Definition der Folgenkonvergenz und den Rechenregeln für reelle Zahlenfolgen. Beispielhaft zeigen wir hier (i). Es seien etwa $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}} \subset \mathbb{R}^d$ wie in der Aussage. Definiere

$$a := \lim_{n \rightarrow \infty} a_n, \quad b := \lim_{n \rightarrow \infty} b_n.$$

Dann konvergiert für alle $i \in \{1, \dots, d\}$ die Folge $(a_n^{(i)})_{n \in \mathbb{N}}$ gegen $a^{(i)}$ und $(b_n^{(i)})_{n \in \mathbb{N}}$ gegen $b^{(i)}$. Sei nun $i \in \{1, \dots, d\}$ beliebig. Für die i -te Komponentenfolge von $(a_n + b_n)_{n \in \mathbb{N}}$ gilt dann

$$(a_n + b_n)^{(i)} = a_n^{(i)} + b_n^{(i)} \xrightarrow{n \rightarrow \infty} a^{(i)} + b^{(i)} = (a + b)^{(i)}.$$

Da $i \in \{1, \dots, d\}$ beliebig war erhalten wir nach Definiton 10.2.2, dass $(a_n + b_n)_{n \in \mathbb{N}}$ konvergiert und

$$\lim_{n \rightarrow \infty} (a_n + b_n) = a + b = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n.$$

Punkt (iii) kann auch verallgemeinert werden, wenn der vektorwertige Teil beschränkt ist und die skalarwertige Folge gegen 0 konvergiert.

Proposition 10.2.4

Es sei $(c_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ in $\mathbb{R}$ gegen 0 konvergent und $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ eine *beschränkte Folge*^a

Dann gilt $\lim_{n \rightarrow \infty} c_n a_n = 0$.

^ad.h. es gibt ein $L > 0$ mit $\|a_n\| \leq L$ für alle $n \in \mathbb{N}$.

Beweis. Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ beschränkt. Etwa sei $L > 0$ so, dass $\|a_n\| \leq L$ für alle $n \in \mathbb{N}$. Sei $i \in \{1, \dots, d\}$ beliebig aber fest. Dann gilt auch, dass

$$|a_n^{(i)}| \leq \sqrt{(a_n^{(i)})^2} \leq \sqrt{(a_n^{(1)})^2 + \dots + (a_n^{(d)})^2} = \|a_n\| \leq L \quad \forall n \in \mathbb{N}.$$

Somit ist die i -te Komponentenfolge $(a_n^{(i)})_{n \in \mathbb{N}} \subset \mathbb{R}$ von $(a_n)_{n \in \mathbb{N}}$ eine in $\mathbb{R}$ beschränkte Folge. Die i -te Komponentenfolge von $(c_n a_n)_{n \in \mathbb{N}}$, gegeben durch $(c_n a_n^{(i)})_{n \in \mathbb{N}}$, konvergiert dann wegen Proposition 4.3.2 (vi) gegen Null. Wir haben also gezeigt, dass alle Komponentenfolgen gegen 0 konvergieren. Es folgt $\lim_{n \rightarrow \infty} c_n a_n = 0$.

Konvergenz und Euklidische Norm

Es ist möglich, den obigen Konvergenzbegriff geometrischer zu erfassen, wenn man die Euklidische Norm verwendet. Hierzu formulieren wir zunächst ein Lemma über *Nullfolgen* (d.h. gegen 0 konvergente Folgen) in $\mathbb{R}^d$. Dieses besagt, dass die Nullfolgeneigenschaft komplett mithilfe der Euklidischen Norm erfasst werden kann

Lemma 10.2.5 — NULLFOLGEN IN $\mathbb{R}^d$

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ eine Folge. Dann gilt

$$a_n \xrightarrow{n \rightarrow \infty} 0 \quad (\text{in } \mathbb{R}^d) \quad \Longleftrightarrow \quad \|a_n\| \xrightarrow{n \rightarrow \infty} 0 \quad (\text{in } \mathbb{R}).$$

Beweis. Wir fangen mit „ $\implies$ “ an.

$$\begin{aligned} & \xRightarrow{\text{Definition 10.2.2}} a_n \xrightarrow{n \rightarrow \infty} 0 \\ & \xRightarrow{\text{Quadratfunktion stetig}} a_n^{(i)} \xrightarrow{n \rightarrow \infty} 0 \quad \forall i \in \{1, \dots, d\} \\ & \xRightarrow{\text{GWR}} (a_n^{(i)})^2 \xrightarrow{n \rightarrow \infty} 0^2 = 0 \quad \forall i \in \{1, \dots, d\} \\ & \xRightarrow{\text{Definition 10.1.1}} (a_n^{(1)})^2 + \dots + (a_n^{(d)})^2 \xrightarrow{n \rightarrow \infty} 0 + \dots + 0 = 0 \\ & \xRightarrow{\text{Stetigkeit der Wurzel}} \|a_n\|^2 \xrightarrow{n \rightarrow \infty} 0 \\ & \xRightarrow{\text{Stetigkeit der Wurzel}} \|a_n\| \xrightarrow{n \rightarrow \infty} \sqrt{0} = 0. \end{aligned}$$

Nun zu „ $\impliedby$ “. Es gelte $\|a_n\| \xrightarrow{n \rightarrow \infty} 0$. Für $i \in \{1, \dots, d\}$ gilt dann (wegen der Monotonie der

Wurzel, vgl Proposition 3.4.8(A))

$$0 \leq |a_n^{(i)}| = \sqrt{(a_n^{(i)})^2} \leq \sqrt{(a_n^{(1)})^2 + \dots + (a_n^{(d)})^2} = \|a_n\| \xrightarrow{n \rightarrow \infty} 0.$$

Wegen dem Sandwichlemma folgt

$$|a_n^{(i)}| \xrightarrow{n \rightarrow \infty} 0.$$

Nach Proposition 2.3.30 gilt dann auch

$$a_n^{(i)} \xrightarrow{n \rightarrow \infty} 0.$$

Da $i \in \{1, \dots, d\}$ beliebig war folgt aus Definition 10.2.2, dass

$$a_n \xrightarrow{n \rightarrow \infty} 0.$$

Die Behauptung ist gezeigt.

Der Vorteil an der Folge $(\|a_n\|)_{n \in \mathbb{N}}$ ist, dass es sich um eine *reelle Zahlenfolge* in $\mathbb{R}^1$ handelt. Somit kann man diese Folge mit *eindimensionalen Methoden* untersuchen.

Nun kommen wir zu der Formulierung der Konvergenz mithilfe der Euklidischen Norm

Satz 10.2.6 — KONVERGENZ UND NORM

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ eine Folge und $a \in \mathbb{R}^d$. Dann sind folgende Aussagen äquivalent

- (a) $a_n \xrightarrow{n \rightarrow \infty} a.$
- (b) $\|a_n - a\| \xrightarrow{n \rightarrow \infty} 0$
- (c) $\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \|a_n - a\| \leq \varepsilon.$

Beweis.

$$\begin{array}{ll}
 & a_n \xrightarrow{n \rightarrow \infty} a \\
 \Longleftrightarrow & a_n - a \xrightarrow{n \rightarrow \infty} 0 \\
 \text{Prop10.2.3} & \\
 \Longleftrightarrow & \|a_n - a\| \xrightarrow{n \rightarrow \infty} 0 \quad (\text{in } \mathbb{R}) \\
 \text{Lem10.2.5} & \\
 \Longleftrightarrow & \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \|a_n - a\| - 0 \leq \varepsilon \\
 \text{Def2.3.1} & \\
 \Longleftrightarrow & \forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \quad \|a_n - a\| \leq \varepsilon \\
 \text{Satz10.1.2(i)} &
 \end{array}$$

Ein wichtiges Korollar ist zum Beispiel die *Beschränktheit* konvergenter Folgen, die ganz ähnlich wie im Eindimensionalen gezeigt werden kann.

Korollar 10.2.7 — KONVERGENTE FOLGEN SIND BESCHRÄNKT

Es sei $(a_n)_{n \in \mathbb{N}}$ eine konvergente Folge. Dann ist $(a_n)_{n \in \mathbb{N}}$ *beschränkt*, d.h. es gibt ein $L > 0$ so, dass

$$\forall n \in \mathbb{N} : \|a_n\| \leq L.$$

Beweis. Es sei $a := \lim_{n \rightarrow \infty} a_n$. Wenden wir Satz 10.2.6 mit $\varepsilon = 1$ an so schließen wir

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \|a_n - a\| \leq 1. \quad (526)$$

Für $n \geq n_0$ gilt dann

$$\|a_n\| = \|a_n - a + a\| \leq \|a_n - a\| + \|a\| \underset{(526)}{\leq} 1 + \|a\|.$$

Setzt man nun $L := \max\{\|a_1\|, \dots, \|a_{n_0}\|, 1 + \|a\|\}$ so folgert man leicht, dass $\|a_n\| \leq L$ für alle $n \in \mathbb{N}$

Der Satz von Bolzano-Weierstraß

Ein wichtiges Resultat für reelle Zahlenfolgen, welches auch eine mehrdimensionale Verallgemeinerung für Folgen auf $\mathbb{R}^d$ besitzt, ist der *Satz von Bolzano-Weierstraß* (vgl. Satz 3.5.16). Er besagt, dass jede beschränkte Folge in $\mathbb{R}$ eine konvergente Teilfolge besitzt. Wir erinnern an dieser Stelle an die Definition einer Teilfolge.

Definition 10.2.8 — TEILFOLGE

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ eine Folge und $(\ell_n)_{n \in \mathbb{N}} \subseteq \mathbb{N}$ eine *strikt aufsteigende Folge* (d.h. $\ell_{n+1} \geq \ell_n + 1$ für alle $n \in \mathbb{N}$). Dann heißt $(a_{\ell_n})_{n \in \mathbb{N}}$ *Teilfolge* von $(a_n)_{n \in \mathbb{N}}$.

Man überzeugt sich auch wie im $\mathbb{R}^1$ leicht davon, dass Teilfolgen konvergenter Folgen wieder konvergieren.

Proposition 10.2.9 — TEILFOLGEN KONVERGENTER FOLGEN KONVERGIEREN

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ eine konvergente Folge. Dann konvergiert jede Teilfolge von $(a_n)_{n \in \mathbb{N}}$ gegen denselben Grenzwert.

Beweis. Ist den Lesenden als Übungsaufgabe überlassen.

Nun sind wir bereit, den Satz von Bolzano-Weierstraß auf $\mathbb{R}^d$ zu verallgemeinern. Der Beweis kann nicht aus dem Reellen übernommen werden, denn dort wird die *Anordnung der reellen Zahlen* verwendet (welche wir auf $\mathbb{R}^d$ nicht zur Verfügung haben)

Satz 10.2.10 — SATZ VON BOLZANO-WEIERSTRASS IN $\mathbb{R}^d$

Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ eine beschränkte Folge^a. Dann besitzt $(a_n)_{n \in \mathbb{N}}$ eine konvergente Teilfolge.

^ad.h. es gibt ein $L > 0$ so, dass $\|a_n\| \leq L$ für alle $n \in \mathbb{N}$.

03.06.2026

Beweis. Es sei $(a_n)_{n \in \mathbb{N}}$ wie in der Aussage und $L > 0$ so, dass $\|a_n\| \leq L$ für alle $n \in \mathbb{N}$. Für alle $i \in \{1, \dots, d\}$ und $n \in \mathbb{N}$ gilt dann auch

$$|a_n^{(i)}| \leq \|a_n\| \leq L. \quad (527)$$

Somit ist auch jede Komponentenfolge $(a_n^{(i)})_{n \in \mathbb{N}}$ beschränkt. Deswegen hat nach dem herkömmlichen, eindimensionalen Satz von Bolzano-Weierstraß die Folge $(a_n^{(1)})_{n \in \mathbb{N}}$ eine konvergente Teilfolge (vgl. Satz 3.5.16). Fixiere eine solche Teilfolge und nenne den Grenzwert $a^{(1)} \in \mathbb{R}$. Das bedeutet

$$\exists (\ell_{1,n})_{n \in \mathbb{N}} \subseteq \mathbb{N} \text{ strikt aufsteigend : } a_{\ell_{1,n}}^{(1)} \xrightarrow{n \rightarrow \infty} a^{(1)}. \quad (528)$$

Nun ist $(a_{\ell_{1,n}}^{(2)})_{n \in \mathbb{N}}$ nach (527) auch eine beschränkte Folge. Wiederum finden wir nach dem eindimensionalen Satz von Weierstraß eine konvergente Teilfolge von $(a_{\ell_{1,n}}^{(2)})_{n \in \mathbb{N}}$. Wir fixieren wiederum eine solche Folge und nennen den Grenzwert $a^{(2)} \in \mathbb{R}$. Mit anderen Worten:

$$\exists (\ell_{2,n})_{n \in \mathbb{N}} \subseteq \mathbb{N} \text{ Teilfolge von } (\ell_{1,n})_{n \in \mathbb{N}} : a_{\ell_{2,n}}^{(2)} \xrightarrow{n \rightarrow \infty} a^{(2)}. \quad (529)$$

Nun ist außerdem $(a_{\ell_{2,n}}^{(3)})_{n \in \mathbb{N}}$ eine beschränkte Folge und wir folgern wiederum aus dem eindimensionalen Satz von Weierstraß, dass (für ein gewisses $a^{(3)} \in \mathbb{R}$) gilt

$$\exists (\ell_{3,n})_{n \in \mathbb{N}} \subseteq \mathbb{N} \text{ Teilfolge von } (\ell_{2,n})_{n \in \mathbb{N}} : a_{\ell_{3,n}}^{(3)} \xrightarrow{n \rightarrow \infty} a^{(3)}.$$

Für alle weiteren $i \in \{4, \dots, d\}$ können wir weitermachen und folgern (für ein gewisse $a^{(i)} \in \mathbb{R}$)

$$\exists (\ell_{i,n})_{n \in \mathbb{N}} \subseteq \mathbb{N} \text{ Teilfolge von } (\ell_{i-1,n})_{n \in \mathbb{N}} : a_{\ell_{i,n}}^{(i)} \xrightarrow{n \rightarrow \infty} a^{(i)}. \quad (530)$$

Nun betrachte die Folge $(k_n)_{n \in \mathbb{N}} := (\ell_{d,n})_{n \in \mathbb{N}}$. Diese ist Teilfolge von allen vorher betrachteten Folgen $(\ell_{1,n})_{n \in \mathbb{N}}, (\ell_{2,n})_{n \in \mathbb{N}}, \dots, (\ell_{d-1,n})_{n \in \mathbb{N}}$.^a Nach (528) und Proposition 10.2.9 gilt dann

$$a_{k_n}^{(1)} = a_{\ell_{d,n}}^{(1)} \xrightarrow{n \rightarrow \infty} a^{(1)} \quad \text{weil } (a_{\ell_{d,n}}^{(1)})_{n \in \mathbb{N}} \text{ Teilfolge von } (a_{\ell_{1,n}}^{(1)})_{n \in \mathbb{N}} \text{ ist.}$$

Ferner gilt nach (529) und Proposition 10.2.9

$$a_{k_n}^{(2)} = a_{\ell_{d,n}}^{(2)} \xrightarrow{n \rightarrow \infty} a^{(2)} \quad \text{weil } (a_{\ell_{d,n}}^{(2)})_{n \in \mathbb{N}} \text{ Teilfolge von } (a_{\ell_{2,n}}^{(2)})_{n \in \mathbb{N}} \text{ ist.}$$

Analog gilt für $i \in \{3, \dots, d\}$ wegen Proposition 10.2.9 (da $(\ell_{d,n})_{n \in \mathbb{N}}$ Teilfolge von $(\ell_{i,n})_{n \in \mathbb{N}}$ ist)

$$a_{k_n}^{(i)} = a_{\ell_{d,n}}^{(i)} \xrightarrow{n \rightarrow \infty} a^{(i)}.$$

Definiert man nun $a := (a^{(1)}, \dots, a^{(d)})^T$ so folgt aus den vorigen drei Gleichungen und Definition 10.2.2, dass

$$a_{k_n} \xrightarrow{n \rightarrow \infty} a.$$

^aIn der Tat: Für alle $i \in \{1, \dots, d\}$ ist ja $(\ell_{i,n})_{n \in \mathbb{N}}$ eine Teilfolge von $(\ell_{i-1,n})_{n \in \mathbb{N}}$, vgl. (530). Man spricht von *geschachtelten Teilfolgen*.

10.3 Topologie des $\mathbb{R}^d$

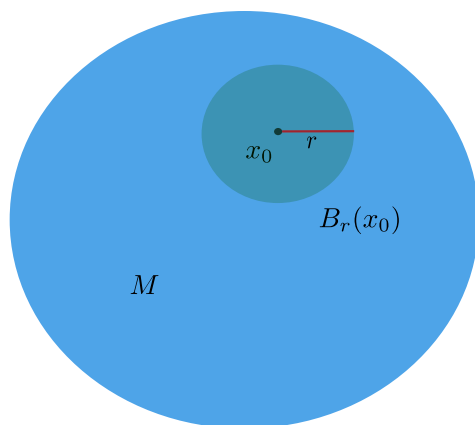
Das Wort Topologie kommt von griechischen Wort *topos*(=Ort) ab. Im Folgenden wollen wir also über spezielle Orte bei Punktmengen in $\mathbb{R}^d$ sprechen und dabei verschiedene Konzepte verallgemeinern, die wir in der eindimensionalen Analysis nur *on the fly* eingeführt haben. Im Folgenden ist stets $M \subseteq \mathbb{R}^d$ eine Menge.

Definition 10.3.1

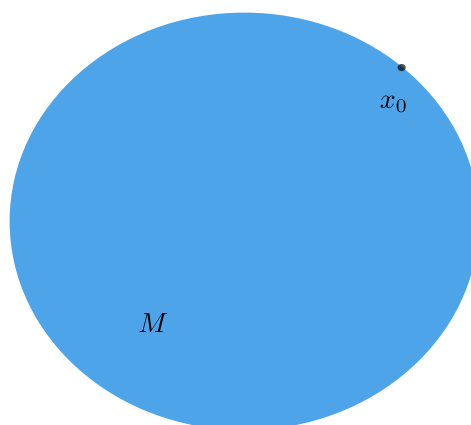
Es sei $M \subseteq \mathbb{R}^d$ eine Menge. Ein Punkt $x_0 \in \mathbb{R}^d$ heißt

- (i) *Innerer Punkt* von M falls ein $r > 0$ existiert mit $B_r(x_0) \subseteq M$. Die Menge aller inneren Punkte von M bezeichnet man mit $\overset{\circ}{M}$ oder $\text{int}(M)$ und nennt sie das *Innere* von M .
- (ii) *Berührungspunkt*^a von M falls es eine Folge $(x_n)_{n \in \mathbb{N}} \subseteq M$ gibt mit $\lim_{n \rightarrow \infty} x_n = x_0$. Die Menge aller Berührungspunkte nennt man $\overline{M}$ und spricht vom *Abschluss* von M .
- (iii) *Randpunkt* von M , falls x_0 ein Berührungspunkt von M aber kein innerer Punkt von M ist. Die Menge aller Randpunkte nennen wir Rand von M und bezeichnen sie mit ∂M .

^aIn der Analysis I habe ich hierfür das Wort *Häufungspunkt* verwendet, bin damit aber jetzt nicht mehr so glücklich. Daher werde ich ab jetzt das Wort Berührungspunkt verwenden.



$$x_0 \in \overset{\circ}{M}$$



$$x_0 \in \partial M$$

Bemerkung 10.3.2

Ist $x_0 \in \mathbb{R}^d$ ein innerer Punkt einer Menge M so gilt stets $x_0 \in M$. In der Tat: Ist $r > 0$ so, dass $B_r(x_0) \subseteq M$ so gilt $x_0 \in B_r(x_0) \subseteq M$. Wir haben also gesehen, dass

$$\overset{\circ}{M} \subseteq M.$$

Ein Berührungspunkt von M muss aber nicht zwingend in M liegen. Man nehme dazu als Beispiel $M = \mathbb{R}^d \setminus \{0\}$. Dann ist $0 \notin M$, aber 0 ist durchaus ein Berührungspunkt von M , denn es gilt $(\frac{1}{n}e_1)_{n \in \mathbb{N}} \subseteq M$ und

$$\lim_{n \rightarrow \infty} \frac{1}{n}e_1 \stackrel{\text{Prop 10.2.3}}{=} 0 \cdot e_1 = 0.$$

Jedoch gilt die umgekehrte Inklusion: Jeder Punkt $x_0 \in M$ ist auch ein Berührungspunkt von M ! In der Tat: Für $x_0 \in M$ kann man stets die *konstante Folge* $(x_n)_{n \in \mathbb{N}} = (x_0)_{n \in \mathbb{N}} \subset M$ betrachten und mit dieser Definition 10.3.1 (ii) verifizieren. Mit anderen Worten gilt

$$M \subseteq \overline{M}.$$

Fazit. Für alle Mengen $M \subseteq \mathbb{R}^d$ gilt

$$\overset{\circ}{M} \subseteq M \subseteq \overline{M}.$$

Man beachte auch, dass nach Definition 10.3.1(iii) stets gilt, dass

$$\partial M = \overline{M} \setminus \overset{\circ}{M}.$$

Zum Warmwerden erstmal ein Beispiel im Reellen (d.h. $d = 1$).

Beispiel 10.3.3

Es sei $A \subseteq \mathbb{R}$ beschränkt.

Behauptung. $\sup(A)$ und $\inf(A)$ sind stets Berührungspunkte von A .

Beweis. Wir weisen das nur für $\sup(A)$ nach. Sei $n \in \mathbb{N}$. Nutzt man Bemerkung 3.2.27 mit $\varepsilon = \frac{1}{n}$ so findet man zu jedem $n \in \mathbb{N}$ ein $a_n \in A$ mit

$$\sup(A) - \frac{1}{n} \leq a_n \leq \sup(A). \quad (531)$$

Man kann nun $(a_n)_{n \in \mathbb{N}} \subseteq A$ als Folge in A auffassen. Mit (531) und dem Sandwichlemma kann man feststellen, dass diese gegen $\sup(A)$ konvergiert. Nach Definition 10.3.1 ist $\sup(A)$ dann also ein Berührungspunkt.

Nun ein weiteres Beispiel zum Abschluss einer Menge

Beispiel 10.3.4

Es sei $x_1 \in \mathbb{R}^d$ und $s > 0$.

Behauptung. $\overline{B_s(x_1)} = \{x \in \mathbb{R}^d : \|x - x_1\| \leq s\}$.

Beweis. Wir müssen eine Mengengleichheit zeigen. Dies machen wir, indem wir die Teilmengeninklusionen ' $\subseteq$ ' und ' $\supseteq$ ' zeigen.

SCHRITT 1. $\overline{B_s(x_0)} \subseteq \{x \in \mathbb{R}^d : \|x - x_1\| \leq s\}$

Es sei dazu $x \in \overline{B_s(x_1)}$. Zu zeigen ist, dass $\|x - x_1\| \leq s$. Wähle dazu eine Folge $(y_n)_{n \in \mathbb{N}} \subseteq B_s(x_1)$ mit $\lim_{n \rightarrow \infty} y_n = x$.^a Dann gilt für alle $n \in \mathbb{N}$

$$\|x - x_1\| = \|x - y_n + y_n - x_1\| \leq \|x - y_n\| + \underbrace{\|y_n - x_1\|}_{< s \text{ da } y_n \in B_s(x_1)} \leq \|x - y_n\| + s. \quad (532)$$

Nutzt man nun Satz 10.2.6 und $y_n \xrightarrow{n \rightarrow \infty} x_1$, so folgt $\|x - y_n\| \xrightarrow{n \rightarrow \infty} 0$. Lässt man in der Ungleichung (532) n aus beiden Seiten gegen unendlich streben so folgert man

$$\|x - x_1\| \leq s.$$

SCHRITT 2. $\overline{B_s(x_0)} \supseteq \{x \in \mathbb{R}^d : \|x - x_1\| \leq s\}$

Dazu sei $x \in \mathbb{R}^d$ mit $\|x - x_1\| \leq s$. Zu zeigen ist dass $x \in \overline{B_s(x_1)}$, d.h. (nach Definition 10.3.1), dass eine Folge $(y_n)_{n \in \mathbb{N}} \subseteq B_s(x_1)$ existiert mit $\lim_{n \rightarrow \infty} y_n = x$. Definiere dazu für $n \in \mathbb{N}$

$$y_n := x - \frac{1}{n}(x - x_1) \quad (533)$$

Man überzeugt sich leicht mit Proposition 10.2.3 (ii) davon, dass $\lim_{n \rightarrow \infty} y_n = x$. Es muss noch gezeigt werden, dass $(y_n)_{n \in \mathbb{N}} \subseteq B_s(x_1)$. Dazu bilden wir

$$\begin{aligned} \|y_n - x_1\| &\stackrel{(533)}{=} \left\| x - \frac{1}{n}(x - x_1) - x_1 \right\| = \|x - x_1 - \frac{1}{n}(x - x_1)\| \\ &= \left\| \left(1 - \frac{1}{n}\right)(x - x_1) \right\| \stackrel{\text{Satz 10.1.2(ii)}}{=} \left(1 - \frac{1}{n}\right) \|x - x_1\|. \end{aligned}$$

Nutzen wir, dass $\|x - x_1\| \leq s$ so folgern wir

$$\|y_n - x_1\| \leq \left(1 - \frac{1}{n}\right)s = s - \frac{s}{n} < s.$$

Das zeigt, dass $y_n \in B_s(x_1)$ für alle $n \in \mathbb{N}$, d.h. $(y_n)_{n \in \mathbb{N}} \subseteq B_s(x_1)$. Somit haben wir eine Folge $(y_n)_{n \in \mathbb{N}}$ wie oben gewünscht gefunden.

^aEine solche Folge $(y_n)_{n \in \mathbb{N}}$ existiert gerade nach Definition 10.3.1

Offene und Abgeschlossene Mengen

Die im vorigen Abschnitt definierten Begriffe sind im Fall $d = 1$ bereits aus vorherigen Abschnitten des Skriptes bekannt, wurden dort aber nur wenig diskutiert. Beim Studium offener und abgeschlossener Mengen können wir vertiefte Einblicke in die obigen Begrifflichkeiten ermöglichen.

Definition 10.3.5

Es sei $M \subseteq \mathbb{R}^d$ eine Menge.

- (i) M heißt *offen*, wenn jeder Punkt $x_0 \in M$ ein innerer Punkt von M ist, d.h. $M = \overset{\circ}{M}$ ^{ab}
- (ii) M heißt *abgeschlossen*, wenn jeder Berührungspunkt von M auch in M liegt, d.h. $M = \overline{M}$ _c.

^aEigentlich steht in der Definition nur, dass $M \subseteq \overset{\circ}{M}$. Die andere Teilmengeninklusion gilt aber automatisch nach Bemerkung 10.3.2.

^bIn GEOGEBRA können Sie eine offene Menge selber erkunden!

^cEigentlich steht in der Definition nur, dass $\overline{M} \subseteq M$. Die andere Teilmengeninklusion gilt aber automatisch nach Bemerkung 10.3.2.

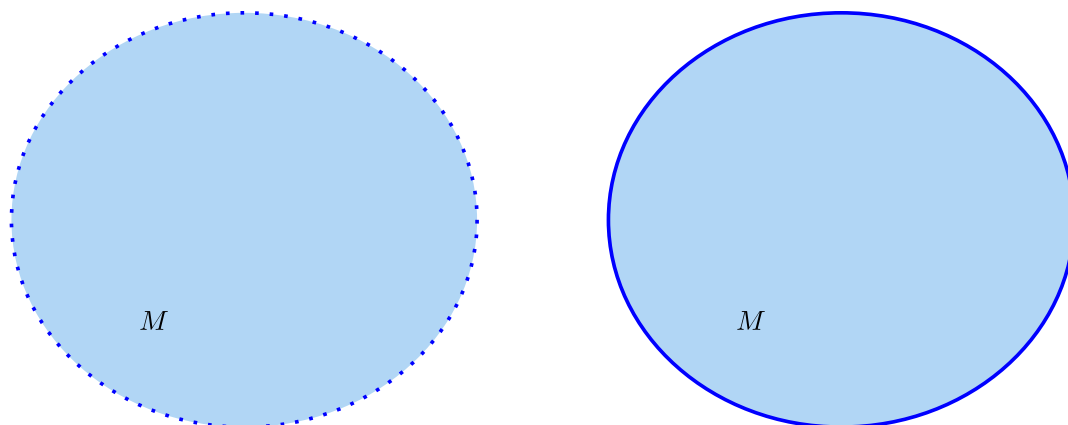


Abbildung 41: Bei einer offenen Menge (links) ist jeder Punkt ein innerer Punkt. Insbesondere sind die Randpunkte (gepunktet) nicht enthalten. Bei einer abgeschlossenen Menge (rechts) ist jeder Berührungspunkt enthalten. Insbesondere sind die Randpunkte (durchgezogen) auch drin.

Bemerkung 10.3.6

Fügt man die obige Definition und Definition 10.3.1 zusammen, so kann man gut *quantifiziert* aufschreiben, was *offen* und *abgeschlossen* bedeutet. Es gilt

(i)

$$\begin{aligned}
 &M \subseteq \mathbb{R}^d \text{ ist offen} \\
 \stackrel{\text{Def10.3.5}}{\iff} &\forall x_0 \in M : \quad x_0 \text{ ist innerer Punkt von } M \\
 \stackrel{\text{Def10.3.1}}{\iff} &\forall x_0 \in M : \quad \exists r > 0 : \quad B_r(x_0) \subseteq M.
 \end{aligned}$$

Man verbalisiert letztere Bedingung auch häufig so: *Jeder Punkt aus M besitzt eine Umgebung, die in M liegt.*^a

(ii)

$$\begin{aligned}
 &M \subseteq \mathbb{R}^d \text{ ist abgeschlossen} \\
 \stackrel{\text{Def10.3.5}}{\iff} &(x_0 \in \mathbb{R}^d \text{ ist Berührungspunkt von } M \implies x_0 \in M) \\
 \stackrel{\text{Def10.3.1}}{\iff} &\left(\exists (x_n)_{n \in \mathbb{N}} \subseteq M : \lim_{n \rightarrow \infty} x_n = x_0 \implies x_0 \in M \right).
 \end{aligned}$$

Man verbalisiert letztere Bedingung auch häufig so: *Jede konvergente Folge in M besitzt ihren Grenzwert in M .*

^aAnschaulich kann man die Bedingung auch mit folgender interaktiver GEOGEBRA-Grafik verstehen.

Beispiel 10.3.7

Es sei $x_1 \in \mathbb{R}^d$ und $s > 0$.

Behauptung. $B_s(x_1)$ ist offen.

Beweis. Es sei $x_0 \in B_s(x_1)$, d.h. $\|x_0 - x_1\| < s$ (vgl. Definition 10.1.4). Zu zeigen ist, dass x_0 ein innerer Punkt von $B_s(x_1)$ ist, d.h. es gibt ein $r > 0$ mit $B_r(x_0) \subseteq B_s(x_1)$. Definiere hierzu $r := s - \|x_0 - x_1\| > 0$. Wir behaupten, dass $B_r(x_0) \subseteq B_s(x_1)$ (siehe auch Abbildung 42). Dazu: Sei $x \in B_r(x_0)$ beliebig (d.h. $\|x - x_0\| < r$).

$$\underline{Z}: \quad x \in B_s(x_1) \quad (\text{d.h. } \|x - x_1\| < s).$$

Dazu:

$$\begin{aligned} \|x - x_1\| &= \|(x - x_0) + (x_0 - x_1)\| \underset{\text{Satz 10.1.2}}{\leq} \|x - x_0\| + \|x_0 - x_1\| < r + \|x_0 - x_1\| \\ &= (s - \|x_0 - x_1\|) + \|x_0 - x_1\| = s. \end{aligned}$$

Somit ist gezeigt, dass $\|x - x_1\| < s$.

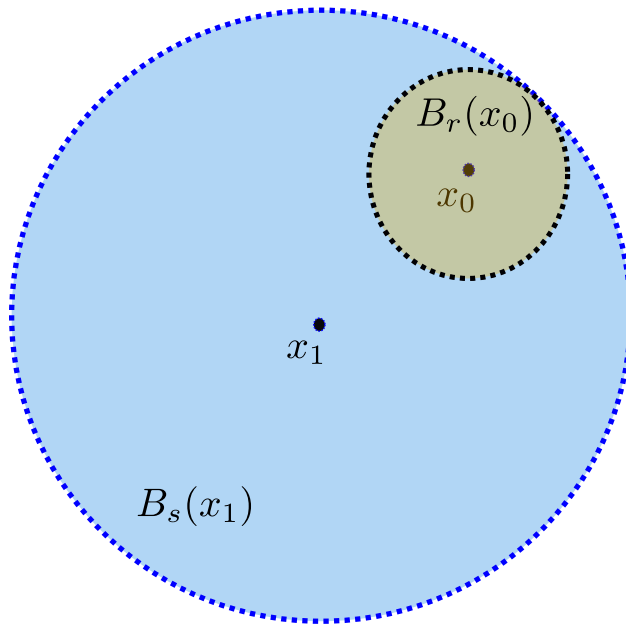


Abbildung 42: Abbildung zu Beispiel 10.3.7

Nicht jede Teilmenge in $\mathbb{R}^d$ ist offen oder abgeschlossen, d.h. *weder-noch* ist möglich.

Beispiel 10.3.8

Sei $d = 1$.

Behauptung. $(0, 1] \subseteq \mathbb{R}^1$ ist weder abgeschlossen noch offen.

Beweis.

SCHRITT 1. $(0, 1]$ ist nicht offen.

Um zu zeigen, dass $(0, 1]$ nicht offen ist, zeigen wir, dass 1 kein innerer Punkt von $(0, 1]$ ist. Dazu: Für jedes $r > 0$ gilt

$$B_r(1) = (1 - r, 1 + r) \not\subseteq (0, 1], \quad \text{da z.B. } 1 + \frac{r}{2} \in (1 - r, 1 + r) \setminus (0, 1].$$

SCHRITT 2. $(0, 1]$ ist nicht abgeschlossen.

Um zu zeigen, dass $(0, 1]$ nicht abgeschlossen ist, finden wir einen Berührungspunkt von $(0, 1]$, der nicht in $(0, 1]$ liegt. Dazu: 0 ist Berührungspunkt von $(0, 1]$, denn $(\frac{1}{n})_{n \in \mathbb{N}} \subseteq (0, 1]$ und

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0.$$

Eine Menge kann aber durchaus gleichzeitig offen und abgeschlossen sein.

Beispiel 10.3.9

Es sei $d \in \mathbb{N}$.

Behauptung. $\emptyset, \mathbb{R}^d$ sind sowohl offen als auch abgeschlossen.

Beweis. Zuerst für $\emptyset$. Zur Offenheit muss gezeigt werden, dass jeder Punkt aus $\emptyset$ ein innerer Punkt von $\emptyset$ ist. Das ist klar, da $\emptyset$ keine Punkte enthält und somit hier kein einziger Punkt überprüft werden muss. Für die Abschlossenheit muss gezeigt werden, dass jeder Berührungspunkt von $\emptyset$ in $\emptyset$ liegt. Jedoch hat $\emptyset$ gar keine Berührungspunkte, da man innerhalb $\emptyset$ gar keine Folge $(x_n)_{n \in \mathbb{N}} \subseteq \emptyset$ finden kann. Eine solche Folge bräuchte man aber um einen Berührungspunkt zu approximieren, vgl. Definition 10.3.1 (ii).

Der Beweis für $\mathbb{R}^d$ ist den Lesenden als Übungsaufgabe überlassen oder kann auch direkt mit der nachfolgenden Proposition gefolgert werden.

Ist eine Menge M offen, so ist ihre *Komplementmenge* $M^C := \mathbb{R}^d \setminus M$ abgeschlossen und andersherum. Dieses Resultat bedeutet (wie bereits diskutiert) NICHT, dass eine Menge entweder offen oder abgeschlossen ist. Leider wird es aber häufig so verstanden. Es handelt sich aber um ein völlig anderes Resultat: Das Bilden einer Komplementmenge ist eben etwas anderes als die Verneinung einer Aussage und daran sollte man sich an dieser Stelle immer erinnern!

Proposition 10.3.10

Es sei $M \subseteq \mathbb{R}^d$ und $M^C := \mathbb{R}^d \setminus M$.

- (i) Ist M offen, so ist M^C abgeschlossen.

(ii) Ist M abgeschlossen, so ist M^C offen.

Beweis. Zu (i). Es sei M offen. Wir müssen nun zeigen, dass M^C abgeschlossen ist. Es sei dazu $x_0 \in \mathbb{R}^d$ ein beliebiger Berührungspunkt von M^C , d.h. es gibt $(x_n)_{n \in \mathbb{N}} \subseteq M^C$ mit $\lim_{n \rightarrow \infty} x_n = x_0$.

$$\mathbb{Z}: \quad x_0 \in M^C. \quad (a)$$

Angenommen das Gegenteil ist der Fall, d.h. $x_0 \in M$. Da M nach Voraussetzung offen ist, ist x_0 auch ein innerer Punkt von M , vgl. Definition 10.3.5 (i). Das bedeutet (vgl. Definition 10.3.1) es gibt ein $r > 0$ mit $B_r(x_0) \subseteq M$. Da $x_n \xrightarrow{n \rightarrow \infty} x_0$, können wir nun Satz 10.2.6(c) mit $\varepsilon = \frac{r}{2} > 0$ nutzen und sehen

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0: \quad \|x_n - x_0\| \leq \frac{r}{2}.$$

Insbesondere gilt $\|x_{n_0} - x_0\| \leq \frac{r}{2} < r$. Das bedeutet nach Definition 10.1.4 aber nichts anderes als $x_{n_0} \in B_r(x_0)$. Dann gilt aber $x_{n_0} \in B_r(x_0) \subseteq M$. Das widerspricht der Tatsache, dass $(x_n)_{n \in \mathbb{N}} \subseteq M^C$. Wir erhalten einen Widerspruch und folgern $x \in M^C$. Die Behauptung ist gezeigt.

Zu (ii). Es sei M abgeschlossen. Zu zeigen ist, dass M^C offen ist. Dazu sei $x_0 \in M^C$ beliebig. Wir müssen nach Definition 10.3.5 zeigen, dass x_0 ein innerer Punkt von M^C ist, d.h.

$$\mathbb{Z}: \quad \exists r > 0: \quad B_r(x_0) \subseteq M^C.$$

Dazu machen wir wieder einen Widerspruchsbeweis. Angenommen, ein solches $r > 0$ gibt es nicht. Dann enthält für jedes $r > 0$ die Umgebung $B_r(x_0)$ einen Punkt aus M . Genauer gilt

$$M \cap B_r(x_0) \neq \emptyset \quad \forall r > 0.$$

Deswegen kann man für alle $n \in \mathbb{N}$ einen Punkt $x_n \in M \cap B_{\frac{1}{n}}(x_0)$ wählen. Dann definiert $(x_n)_{n \in \mathbb{N}} \subseteq M$ eine Folge in M und

$$\|x_n - x_0\| < \frac{1}{n} \quad \forall n \in \mathbb{N} \quad (\text{da } x_n \in B_{\frac{1}{n}}(x_0)).$$

Wir folgern, dass $\|x_n - x_0\| \xrightarrow{n \rightarrow \infty} 0$ und somit gilt (nach Satz 10.2.6) auch $x_n \xrightarrow{n \rightarrow \infty} x_0$. Da $(x_n)_{n \in \mathbb{N}} \subseteq M$ ist x_0 also ein Berührungspunkt von M . Wegen der Abgeschlossenheit von M gilt dann $x_0 \in M$. Ein Widerspruch, denn x_0 war ein Punkt aus M^C .

^aDann ist nämlich jeder beliebige Berührungspunkt von M^C in M^C enthalten und somit ist M^C abgeschlossen, vgl. Definition 10.3.5 (ii).

Topologie und Mengenoperationen

Nun sprechen wir darüber, wie Offenheit und Abgeschlossenheit sich unter den bekannten *Mengenoperationen* (d.h. Durchschnitt, Vereinigung, etc.) verhält. Um dies in voller Allge-

meinheit diskutieren zu können, führen wir die oben genannten Mengenoperationen nun zunächst für beliebig große *Mengenfamilien* ein.

Definition 10.3.11 — OPERATIONEN AUF MENGENFAMILIEN

Es sei I eine Menge, im Folgenden *Indexmenge* genannt. Ferner sei für jedes $i \in I$ eine Menge $A_i \subseteq \mathbb{R}^d$ gegeben.^a Dann definieren wir

$$\bigcup_{i \in I} A_i := \{x \in \mathbb{R}^d \mid \exists i \in I : x \in A_i\} \quad \text{bzw.} \quad \bigcap_{i \in I} A_i := \{x \in \mathbb{R}^d \mid \forall i \in I : x \in A_i\}$$

und sprechen von der *Vereinigung* bzw. vom *Schnitt* von $(A_i)_{i \in I}$.

^aWir sprechen dann von einer sogenannten *Mengenfamilie* $(A_i)_{i \in I}$.

Ist I endlich, etwa $I = \{i_1, \dots, i_n\}$, so können wir wie vorher

$$\bigcup_{i \in I} A_i = A_{i_1} \cup \dots \cup A_{i_n} \quad \text{bzw.} \quad \bigcap_{i \in I} A_i = A_{i_1} \cap \dots \cap A_{i_n}$$

schreiben.

Bemerkung 10.3.12

Die aus Kapitel 1.2 bekannten Mengen-Rechenregeln verallgemeinern sich auf Vereinigungen und Schnitte beliebiger Familien. Zwei Regeln, auf die ich an dieser Stelle gesondert hinweisen möchte, sind die *De Morgan'schen Regeln*

Fakt. DE MORGAN'SCHE REGELN. Sei I eine (Index-)Menge und sei für alle $i \in I$ eine Menge $A_i \subseteq \mathbb{R}^d$ gegeben. Dann gilt

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

(Hierbei ist für eine Menge $M \subseteq \mathbb{R}^d$ die Komplementmenge M^C wie immer definiert durch $M^C := \mathbb{R}^d \setminus M$.)

Beweis. Wir zeigen nur die erste De Morgan'sche Regel. Der Rest ist den Lesenden als Übungsaufgabe überlassen.

$$\begin{aligned} & x \in \left(\bigcup_{i \in I} A_i \right)^C \\ \iff & x \in \mathbb{R}^d \quad \wedge \quad x \notin \bigcup_{i \in I} A_i \\ \iff & x \in \mathbb{R}^d \quad \wedge \quad \neg(\exists i \in I : x \in A_i) \\ & \text{Def 10.3.11} \end{aligned}$$

$$\begin{aligned}
& \stackrel{\text{Bem2.2.4}}{\iff} x \in \mathbb{R}^d \quad \wedge \quad \forall i \in I : \neg(x \in A_i) \\
& \iff x \in \mathbb{R}^d \quad \wedge \quad \forall i \in I : x \notin A_i \\
& \iff \forall i \in I : \quad x \in \mathbb{R}^d \quad \wedge \quad x \notin A_i \\
& \iff \forall i \in I : \quad x \in A_i^C. \\
& \stackrel{\text{Def10.3.11}}{\iff} x \in \bigcap_{i \in I} A_i^C.
\end{aligned}$$

Proposition 10.3.13 — TOPOLOGIE UND MENGENOPERATIONEN

Es sei I eine Menge und für alle $i \in I$ seien $A_i \subseteq \mathbb{R}^d$ Mengen.

- (i) Ist A_i offen für alle $i \in I$, so ist auch $\bigcup_{i \in I} A_i$ offen.
- (ii) Ist I endlich und A_i offen für alle $i \in I$, so ist auch $\bigcap_{i \in I} A_i$ offen.
- (iii) Ist A_i abgeschlossen für alle $i \in I$, so ist auch $\bigcap_{i \in I} A_i$ abgeschlossen.
- (iv) Ist I endlich und A_i abgeschlossen für alle $i \in I$, so ist auch $\bigcup_{i \in I} A_i$ abgeschlossen.

Beweis. Zu (i). Es sei $x \in \bigcup_{i \in I} A_i$ beliebig, etwa gelte $x \in A_{i_0}$ für ein $i_0 \in I$. Zu zeigen ist, dass x ein innerer Punkt von $\bigcup_{i \in I} A_i$ ist, d.h.

$$\mathbb{Z}: \quad \exists r > 0 : \quad B_r(x) \subseteq \bigcup_{i \in I} A_i.$$

Dazu: Da A_{i_0} offen ist, ist x bereits ein innerer Punkt von A_{i_0} , d.h. es gibt $r > 0$ mit $B_r(x) \subseteq A_{i_0}$. Dann gilt aber klarerweise auch

$$B_r(x) \subseteq A_{i_0} \subseteq \bigcup_{i \in I} A_i.$$

Dies zeigt die Behauptung.

Zu (ii). Es sei $I = \{i_1, \dots, i_n\}$, $n \in \mathbb{N}$, d.h. $\bigcap_{i \in I} A_i = A_{i_1} \cap \dots \cap A_{i_n}$. Nun sei $x \in A_{i_1} \cap \dots \cap A_{i_n}$ (d.h. $x \in A_{i_j}$ für alle $j = 1, \dots, n$). Zu zeigen ist wieder, dass x ein innerer Punkt von $A_{i_1} \cap \dots \cap A_{i_n}$ ist, d.h.

$$\mathbb{Z}: \quad \exists r > 0 : \quad B_r(x) \subseteq A_{i_1} \cap \dots \cap A_{i_n}$$

Dazu: Da A_{i_1} offen ist und $x \in A_{i_1}$, ist x auch ein innerer Punkt von A_{i_1} . Das bedeutet

$$\exists r_1 > 0 : \quad B_{r_1}(x) \subseteq A_{i_1}$$

Da A_{i_2} offen ist und $x \in A_{i_2}$, ist x auch ein innerer Punkt von A_{i_2} . Das bedeutet

$$\exists r_2 > 0 : \quad B_{r_2}(x) \subseteq A_{i_2}$$

Analog sieht man: Für alle $j \in \{1, \dots, n\}$ gilt

$$\exists r_j > 0 : \quad B_{r_j}(x) \subseteq A_{i_j} \tag{534}$$

Definiere nun $r := \min\{r_1, \dots, r_n\}$. Dann ist $r > 0$, da das Minimum innerhalb dieser endlich vieler strikt positiver Zahlen angenommen wird.^a Wir behaupten nun, dass $B_r(x) \subseteq A_{i_1} \cap \dots \cap A_{i_n}$. Dazu sei $y \in B_r(x)$ beliebig. Für $j \in \{1, \dots, n\}$ gilt dann

$$y \in B_r(x) \underset{r \leq r_j}{\subseteq} B_{r_j}(x) \underset{(534)}{\subseteq} A_{i_j}.$$

Wir haben also gezeigt, dass $y \in A_{i_j}$ für alle $j \in \{1, \dots, n\}$. Das bedeutet nach Definition 10.3.11 nichts anderes als

$$y \in A_{i_1} \cap \dots \cap A_{i_n}.$$

Da $y \in B_r(x)$ beliebig war folgt

$$B_r(x) \subseteq A_{i_1} \cap \dots \cap A_{i_n}.$$

Die Behauptung ist gezeigt.

Zu (iii). Es sei A_i abgeschlossen für alle $i \in I$. Dann ist nach Proposition 10.3.10 $A_i^C := \mathbb{R}^d \setminus A_i$ offen für alle $i \in I$. Nach dem bereits bewiesenen Teil (i) ist dann auch $\bigcup_{i \in I} A_i^C$ offen.

Wiederum mit Proposition 10.3.10 ist

$$\left(\bigcup_{i \in I} A_i^C \right)^C \quad \text{abgeschlossen.} \tag{535}$$

Nach den Regeln von De Morgan (vgl. Bemerkung 10.3.12) gilt

$$\left(\bigcup_{i \in I} A_i^C \right)^C = \bigcap_{i \in I} (A_i^C)^C = \bigcap_{i \in I} A_i.$$

Das zusammen mit (535) zeigt

$$\bigcap_{i \in I} A_i \quad \text{ist abgeschlossen.}$$

Zu (iv). Funktioniert genauso wie (iii) über die De Morgan'schen Regeln. Details sind den Lesenden als Übungsaufgabe überlassen.

^aHier würde der Beweis schiefgehen, wenn die Indexmenge I nicht endlich wäre!

Bemerkung 10.3.14

In Punkt (ii) und (iv) der vorigen Proposition (Proposition 10.3.13) kann auf die vorausgesetzte Endlichkeit der Indexmenge I nicht verzichtet werden. Wir geben ein Gegenbeispiel zur Aussage von Punkt (ii) wenn I nicht endlich ist. Dazu sei $I = \mathbb{N}$ und für $i \in I$ definiere

$$A_i := B_{\frac{1}{i}}(0) \subseteq \mathbb{R}^d.$$

Wie wir in Beispiel 10.3.7 gesehen haben sind dann alle Mengen A_i offen.

Behauptung: $\bigcap_{i \in I} A_i = \{0\}$ (und somit ist $\bigcap_{i \in I} A_i$ nicht offen)^a.

Beweis. Ist $x \in \bigcap_{i \in I} A_i$, so gilt für alle $i \in \mathbb{N}$, dass $x \in A_i$, d.h.

$$\|x\| = \|x - 0\| < \frac{1}{i} \quad \forall i \in \mathbb{N}.$$

Das ist dann und nur dann möglich wenn $\|x\| = 0$, d.h. $x = 0$ (vgl. Satz 10.1.2(i)). Wir haben gesehen, dass $\bigcap_{i \in I} A_i = \{0\}$.

^aGenau genommen müsste man hier natürlich rigoros zeigen, dass $\{0\}$ nicht offen ist. Hier ein kurzes Argument dazu: Wäre $\{0\}$ offen, so wäre $\mathbb{R}^d \setminus \{0\}$ abgeschlossen. Allerdings ist $(\frac{1}{n}e_1)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d \setminus \{0\}$ eine Folge, deren Grenzwert nicht in $\mathbb{R}^d \setminus \{0\}$ liegt. Aus diesem Grund kann $\{0\}$ nicht offen sein.

Kompaktheit

Häufig wird uns das Prinzip der *Kompaktheit* begegnen. Eigentlich haben wir in den vorigen Betrachtungen der eindimensionalen Analysis schon sehr häufig davon profitiert. Dort haben wir nämlich sehr häufig Resultate über *kompakte Intervalle* $[a, b] \subseteq \mathbb{R}$ gezeigt. Diese Resultate werden wir nun im Laufe der Zeit verallgemeinern.

Definition 10.3.15 — KOMPAKTHEIT IN $\mathbb{R}^d$

Eine Teilmenge $K \subseteq \mathbb{R}^d$ heißt *kompakt in $\mathbb{R}^d$* , falls sie

- abgeschlossen ist und
- beschränkt ist, d.h. es gibt ein $L > 0$ mit $\forall x \in K : \|x\| \leq L$.

Später, wenn wir die abstrakte Analysis behandeln, werden wir Kompaktheit anders definieren. Für hier ist diese Definition aber genug. Kompaktheit erweist sich oft als hilfreich, wenn man *Folgen* in kompakten Mengen betrachtet. Den Grund dafür halten wir in der nächsten Proposition fest. Diese Proposition ist eine Verallgemeinerung von Lemma 6.2.11, welches zB beim Beweis vom Satz von Weierstraß sehr nützlich war. Daher werden wir die Rolle dieser Aussage am Ende dieses Abschnittes auch noch genauer hinterfragen.

Proposition 10.3.16 — FOLGEN IN KOMPAKTEN MENGEN

Es sei $K \subseteq \mathbb{R}^d$ eine kompakte Menge. Ist $(x_n)_{n \in \mathbb{N}} \subseteq K$ eine Folge in K so besitzt $(x_n)_{n \in \mathbb{N}}$ eine konvergente Teilfolge, deren Grenzwert wieder in K liegt. Genauer: Es gibt eine Teilfolge $(x_{\ell_n})_{n \in \mathbb{N}}$ und ein $x \in K$ so, dass $x_{\ell_n} \xrightarrow{n \rightarrow \infty} x$.

Beweis. Es sei $(x_n)_{n \in \mathbb{N}} \subseteq K$ eine Folge. Da K beschränkt ist, ist auch $(x_n)_{n \in \mathbb{N}}$ beschränkt. Nach dem Satz von Bolzano-Weierstraß (Satz 10.2.10) gibt es eine konvergente Teilfolge $(x_{\ell_n})_{n \in \mathbb{N}}$. Sei etwa $x := \lim_{n \rightarrow \infty} x_{\ell_n}$. Da $(x_{\ell_n})_{n \in \mathbb{N}} \subseteq K$ ist somit x ein Berührungspunkt von K , vgl. Definition 10.3.1. Da K abgeschlossen ist, enthält K alle seine Berührungspunkte und somit gilt $x \in K$. Die Behauptung ist gezeigt.

Natürlich klären wir auch wieder auf wie sich Kompaktheit unter Mengenoperationen verhält.

Proposition 10.3.17 — KOMPAKTHEIT UNTER MENGENOPERATIONEN

Sei $(K_i)_{i \in I} \subseteq \mathbb{R}^d$ eine Familie kompakter Mengen (d.h. K_i ist kompakt für alle $i \in I$).

- (i) Der Schnitt $\bigcap_{i \in I} K_i$ ist wieder kompakt.
- (ii) Ist I endlich, so ist auch die Vereinigungsmenge $\bigcup_{i \in I} K_i$ wieder kompakt.

Beweis. Wir zeigen nur (ii) und überlassen (i) den Lesenden als Übungsaufgabe. Es sei $I = \{i_1, \dots, i_m\}$ eine endliche Menge. Wir müssen nach Definition 10.3.15 zeigen, dass $\bigcup_{i \in I} K_i = K_{i_1} \cup \dots \cup K_{i_m}$ abgeschlossen und beschränkt ist. Die Abgeschlossenheit folgt direkt aus der Abgeschlossenheit von $K_{i_1}, \dots, K_{i_m}$ und Proposition 10.3.13. Zur Beschränktheit. Sei $j \in \{1, \dots, m\}$. Da K_{i_j} beschränkt ist (vgl. Definition 10.3.15) gibt es ein $L_j > 0$ so, dass

$$\forall x \in K_{i_j} : \|x\| \leq L_j. \quad (536)$$

Definiere nun $L := \max\{L_1, \dots, L_m\} > 0$. Wir behaupten, dass $\|x\| \leq L$ für alle $x \in K_{i_1} \cup \dots \cup K_{i_m}$. Sei dazu $x \in K_{i_1} \cup \dots \cup K_{i_m}$. Dann gibt es $j \in \{1, \dots, m\}$ mit $x \in K_{i_j}$. Somit gilt nach (536)

$$\|x\| \leq L_j \leq L.$$

Die Behauptung ist gezeigt.

In diesem Abschnitt (und auch im Folgenden) werden wir uns fragen, wie man die Kompaktheit einer Menge anders charakterisieren kann. Es stellt sich schnell heraus, dass die Eigenschaft in Proposition 10.3.16 bereits *charakterisierend* für die Kompaktheit ist. Mehr dazu im nächsten Abschnitt.

Folgenkompaktheit Wir führen nun eine weitere Eigenschaft einer Menge ein: die *Folgenkompaktheit*.

Definition 10.3.18 — FOLGENKOMPAKTHEIT

Eine Menge $M \subseteq \mathbb{R}^d$ heißt *folgenkompakt*, falls es zu jeder Folge $(x_n)_{n \in \mathbb{N}} \subseteq M$ in M eine Teilfolge $(x_{\ell_n})_{n \in \mathbb{N}}$ gibt, die konvergiert und deren Grenzwert wiederum in M liegt.

Beispiele folgenkompakter Mengen kennen wir genug: Alle kompakten Mengen (vgl. 10.3.16)!

Bemerkung 10.3.19

Nach Proposition 10.3.16 ist jede kompakte Menge $M \subseteq \mathbb{R}^d$ auch folgenkompakt.

Die Rückrichtung gilt auch: Folgenkompakte Mengen sind auch kompakt. Somit ist die Folgenkompaktheit in $\mathbb{R}^d$ einfach ein äquivalentes Konzept der Kompaktheit.

Proposition 10.3.20

Es sei $M \subseteq \mathbb{R}^d$ folgenkompakt. Dann ist M auch kompakt.

Beweis. Es sei $M \subseteq \mathbb{R}^d$ folgenkompakt. Wir müssen nach Definition 10.3.15 nachweisen, dass M abgeschlossen und beschränkt ist.

SCHRITT 1. M ist beschränkt.

Dazu müssen wir ein $L > 0$ finden, sodass $\|x\| \leq L$ für alle $x \in M$. Angenommen ein solches L gibt es nicht. Dann muss es zu jedem $n \in \mathbb{N}$ ein $x_n \in M$ geben mit

$$\|x_n\| > n. \quad (537)$$

(denn gäbe es ein solches x_n nicht, so hätte man $\|x\| \leq n$ für alle $x \in M$ und könnte somit $L = n$ wählen). Nun können wir $(x_n)_{n \in \mathbb{N}} \subseteq M$ als Folge in M auffassen. Da M folgenkompakt ist, muss es eine Teilfolge $(x_{\ell_n})_{n \in \mathbb{N}}$ geben, die gegen ein $x \in M$ konvergiert. Beachte: Nach Korollar 10.2.7, sind konvergente Folgen aber beschränkt. Man folgert, dass es ein $R > 0$ gibt so, dass

$$\forall n \in \mathbb{N} : \quad \|x_{\ell_n}\| \leq R.$$

Dies und (537) ergibt $\ell_n < \|x_{\ell_n}\| \leq R$ für alle $n \in \mathbb{N}$. Man erhält einen Widerspruch, da $(\ell_n)_{n \in \mathbb{N}}$ eine strikt aufsteigende Folge natürlicher Zahlen ist.

SCHRITT 2. M ist abgeschlossen.

Es sei $y \in \mathbb{R}^d$ ein Berührungspunkt von M . Etwa sei $(x_n)_{n \in \mathbb{N}} \subseteq M$ mit $\lim_{n \rightarrow \infty} x_n = y$.

$$\mathbb{Z}: \quad y \in M.$$

Dazu: Wegen der Folgenkompaktheit von M gibt es eine Teilfolge $(x_{\ell_n})_{n \in \mathbb{N}}$ und ein $x \in M$ mit $\lim_{n \rightarrow \infty} x_{\ell_n} = x$. Nun gilt (weil Teilfolgen konvergenter Folgen stets denselben Grenzwert haben, vgl. Proposition 10.2.9)

$$x = \lim_{n \rightarrow \infty} x_{\ell_n} \stackrel{\text{Prop 10.2.9}}{=} \lim_{n \rightarrow \infty} x_n = y.$$

Da $x \in M$ folgt auch $y \in M$, was zu zeigen war.

Bemerkung 10.3.21

Wir halten die Erkenntnis aus Bemerkung 10.3.19 und Proposition 10.3.20 nochmal gesondert fest. Für eine Menge $K \subseteq \mathbb{R}^d$ gilt

$$K \text{ kompakt} \quad \Longleftrightarrow \quad K \text{ folgenkompakt.}$$

Das Schachtelungsprinzip. Eine weitere interessante und häufiger genutzte Eigenschaft kompakter Mengen ist das *Schachtelungsprinzip*. Es wird auch im nachfolgenden Abschnitt zum Einsatz kommen.

Proposition 10.3.22 — SCHACHTELUNGSPRINZIP

Es sei $(K_i)_{i \in \mathbb{N}}$ eine Familie nichtleerer kompakter Mengen, die *geschachtelt sind*, d.h. es gilt

$$K_{i+1} \subseteq K_i \quad \forall i \in \mathbb{N} \quad (a). \quad (538)$$

Dann gilt $\bigcap_{i \in \mathbb{N}} K_i \neq \emptyset$, d.h. es gibt einen Punkt $x \in \mathbb{R}^d$, der in allen Mengen K_i enthalten ist.

^aSalopp kann man hier auch schreiben $K_1 \supseteq K_2 \supseteq K_3 \supseteq \dots$

Beweis. Für sämtliche $i \in \mathbb{N}$ wähle einen Punkt $x_i \in K_i$. Dieser Punkt lässt sich für alle $i \in \mathbb{N}$ wählen, da $K_i \neq \emptyset$ für alle i . Beachte: Für alle $n \in \mathbb{N}$ gilt

$$x_n \in K_n \stackrel{(538)}{\subseteq} K_1.$$

Somit dürfen wir die Folge $(x_n)_{n \in \mathbb{N}} \subseteq K_1$ als Folge in K_1 auffassen. Da K_1 kompakt ist gibt es nach Proposition 10.3.16 eine Teilfolge $(x_{\ell_n})_{n \in \mathbb{N}}$ und ein $x \in K_1$ mit $\lim_{n \rightarrow \infty} x_n = x$.

BEHAUPTUNG. $x \in K_N$ für alle $N \in \mathbb{N}$.

Dazu: Für beliebiges aber festes $N \in \mathbb{N}$ gilt nach Bemerkung 3.5.12, dass $\ell_N \geq N$. Somit gilt auch für alle $n \geq N$, dass

$$x_{\ell_n} \in K_{\ell_n} \subseteq K_{\ell_N} \subseteq K_N.$$

Daher lässt sich $(x_{\ell_n})_{n \geq N}$ als Folge in K_N sehen. Da $(x_{\ell_n})_{n \geq N} \subseteq K_N$ gegen x konvergiert, ist x ein Berührungspunkt von K_N . Da K_N abgeschlossen ist folgt, dass $x \in K_N$, vgl. Definition 10.3.5. Wir haben gezeigt, dass $x \in K_N$ für alle $N \in \mathbb{N}$. Mit anderen Worten

$$x \in \bigcap_{N \in \mathbb{N}} K_N = \bigcap_{i \in \mathbb{N}} K_i.$$

Die Behauptung folgt.

Überdeckungskompaktheit und der Satz von Heine-Borel. Für die abstrakte Analysis und auch für die Analysis III sehr nützlich ist ein weiteres Kompaktheits-Konzept, die sogenannte *Überdeckungskompaktheit*. Hierzu definieren wir zunächst, was eine Überdeckung ist.

Definition 10.3.23 — ÜBERDECKUNG

Es sei $M \subseteq \mathbb{R}^d$ eine Menge. Eine Familie $(U_i)_{i \in I}$ (I beliebige Indexmenge) heißt *Überdeckung* von M falls

$$M \subseteq \bigcup_{i \in I} U_i.$$

Eine Überdeckung von M $(U_i)_{i \in I}$ heißt *offene Überdeckung (von M)* falls U_i für alle $i \in I$ offen ist.

Beispiel 10.3.24

Es sei $M := (0, 1) \subseteq \mathbb{R}^1$ gegeben und für $n \in \mathbb{N}$ sei $U_n := (\frac{1}{3n}, 1 - \frac{1}{3n})$.

Behauptung. Dann ist $(U_n)_{n \in \mathbb{N}}$ eine Überdeckung von M .

Beweis. Sei $x \in (0, 1)$ beliebig aber fest.

$$\mathbb{Z}: \quad x \in \bigcup_{n \in \mathbb{N}} U_n$$

Beachte: Da $0 < x < 1$, gilt $\min\{x, 1 - x\} > 0$. Wählt man dann ein $n_0 \in \mathbb{N}$ so, dass

$\frac{1}{3n_0} < \min\{x, 1-x\}^a$, so gilt

$$x \geq \min\{x, 1-x\} > \frac{1}{3n_0}$$

und

$$x = 1 - (1-x) \leq 1 - \min\{x, 1-x\} < 1 - \frac{1}{3n_0}.$$

Mit anderen Worten zeigen die beiden vorigen Gleichungen, dass $x \in (\frac{1}{3n_0}, 1 - \frac{1}{3n_0})$. Also gilt

$$x \in U_{n_0} \subseteq \bigcup_{n \in \mathbb{N}} U_n.$$

Da $x \in (0, 1)$ beliebig war folgt

$$(0, 1) \subseteq \bigcup_{n \in \mathbb{N}} U_n.$$

Somit ist $(U_n)_{n \in \mathbb{N}}$ nach Definition 10.3.23 eine Überdeckung von M . Es handelt sich sogar um eine offene Überdeckung, weil U_n für alle $n \in \mathbb{N}$ offen ist.^b

^aDass es ein solches n_0 gibt können Sie leicht als Übungsaufgabe zeigen.

^bDas sieht man zum Beispiel, wenn man U_n als $B_{\frac{1}{2} - \frac{1}{3n}}(\frac{1}{2})$ umschreibt und Beispiel 10.3.7 benutzt.

Nun zum Konzept der *Überdeckungskompaktheit*.

Definition 10.3.25 — ÜBERDECKUNGSKOMPAKTHEIT

Es sei $M \subseteq \mathbb{R}^d$ eine Menge. Wir nennen M *überdeckungskompakt*, falls es zu jeder offenen Überdeckung $(U_i)_{i \in I}$ eine endliche Teilmenge $I_0 \subseteq I$ gibt so, dass $(U_i)_{i \in I_0}$ auch eine Überdeckung von M ist. Kurz gesagt:

Jede offene Überdeckung besitzt eine endliche Teilüberdeckung.

Bitte beachten Sie: *Jede* offene Überdeckung muss eine endliche Teilüberdeckung besitzen! Es reicht hier NICHT, dass M nur eine einzige endliche Überdeckung besitzt. Letzteres wäre nämlich für jede beliebige Menge $M \subseteq \mathbb{R}^d$ der Fall, denn jede Menge $M \subseteq \mathbb{R}^d$ lässt sich mit nur einer Menge (nämlich $\mathbb{R}^d$ selbst) überdecken.

Beispiel 10.3.26

Wir betrachten nochmal die Menge $M = (0, 1) \subseteq \mathbb{R}^1$ aus Beispiel 10.3.24.

Behauptung. M ist nicht überdeckungskompakt.

Beweis. Betrachte dazu die offene Überdeckung $(U_n)_{n \in \mathbb{N}}$ aus Beispiel 10.3.24 (d.h. $U_n = (\frac{1}{3n}, 1 - \frac{1}{3n})$ für alle $n \in \mathbb{N}$). Wäre M überdeckungskompakt, so gäbe es eine endliche

Teilmenge $I_0 \subseteq \mathbb{N}$ so, dass $(U_n)_{n \in I_0}$ auch eine Überdeckung von M ist, d.h.

$$M \subseteq \bigcup_{n \in I_0} U_n.$$

Sei nun etwa $I_0 = \{n_1, \dots, n_k\} \subseteq \mathbb{N}$ und definiere $m := \max\{n_1, \dots, n_k\} \in \mathbb{N}$. Dann gilt für alle $j \in \{1, \dots, k\}$, dass $U_{n_j} = (\frac{1}{3n_j}, 1 - \frac{1}{3n_j}) \subseteq (\frac{1}{3m}, 1 - \frac{1}{3m}) = U_m$. Somit gilt auch

$$M \subseteq \bigcup_{n \in I_0} U_n = U_{n_1} \cup \dots \cup U_{n_k} \subseteq U_m \cup \dots \cup U_m = U_m.$$

Wir erhalten, dass

$$(0, 1) \subseteq (\frac{1}{3m}, 1 - \frac{1}{3m}) \quad \text{für ein } m \in \mathbb{N}.$$

Ein Widerspruch.

Im Folgenden besprechen wir den *Satz von Heine-Borel*, welcher die Äquivalenz der Überdeckungskompaktheit zur Kompaktheit zeigt.

Satz 10.3.27 — SATZ VON HEINE-BOREL

Es sei $K \subseteq \mathbb{R}^d$. Dann gilt

$$K \text{ ist kompakt in } \mathbb{R}^d \quad \Longleftrightarrow \quad K \text{ ist überdeckungskompakt.}$$

Wir werden im nachfolgenden Abschnitt beide Richtungen einzeln beweisen.

Davor noch ein Wort zum Zweck des Konzeptes der Überdeckungskompaktheit. Auch wenn die Überdeckungskompaktheit komplizierter derzeit wohl erscheint als das Konzept der Kompaktheit in $\mathbb{R}^d$, hat es einen entscheidenden Vorteil: Es ist wesentlich universeller und lässt sich in viel abstrakteren Kontexten einsetzen. Mehr dazu im Kapitel über abstrakte Analysis.

Beweis von „ $\implies$ “ in Satz 10.3.27. Ein nützliches Tool hierfür ist die sogenannte *Quader-
teilung*.

Definition 10.3.28 — REGULÄRER QUADER

Es sei $R > 0$. Eine Menge $Q \subset \mathbb{R}^d$ heißt *regulärer Quader in $\mathbb{R}^d$* mit *Seitenlänge $2R$* falls es ein $x_0 \in \mathbb{R}^d$ gibt so, dass $Q = x_0 + [-R, R]^d$.^a (Genauer: Für jedes $x \in Q$ liegt jede Komponente von $x - x_0$ im Intervall $[-R, R]$). Der Punkt x_0 heißt *Mittelpunkt* des Quaders.

^aHierbei ist wie üblich

$$[-R, R]^d := \underbrace{[-R, R] \times \dots \times [-R, R]}_{d\text{-mal}} := \{(x^{(1)}, \dots, x^{(d)})^T \in \mathbb{R}^d : x^{(i)} \in [-R, R] \forall i = 1, \dots, d\}$$

Im folgenden werden wir die folgende *Aufteilung von Quadern* in Teilquader häufig benutzen.

Lemma 10.3.29 — QUADERTEILUNG

Es seien $R > 0$, $d \in \mathbb{N}$ und Q ein regulärer Quader mit Seitenlänge $2R$. Dann gibt es 2^d reguläre Quader $Q_1, \dots, Q_{2^d}$ mit Seitenlänge R so, dass

$$Q = Q_1 \cup Q_2 \cup \dots \cup Q_{2^d}. \quad (539)$$

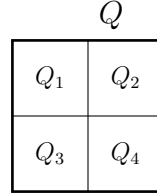


Abbildung 43: Quaderteilung für $d = 2$.

Beweis. Es sei $x_0 \in \mathbb{R}^d$ so, dass $Q = x_0 + [-R, R]^d$. Man beachte: Die Menge $\{-\frac{R}{2}, \frac{R}{2}\}^d = \{-\frac{R}{2}, \frac{R}{2}\} \times \{-\frac{R}{2}, \frac{R}{2}\} \times \dots \times \{-\frac{R}{2}, \frac{R}{2}\}$ besteht aus 2^d Elementen^a. Diese Elemente nennen wir $x_1, \dots, x_{2^d}$. Für alle $j = 1, \dots, 2^d$ definiere nun

$$Q_j := (x_0 + x_j) + [-\frac{R}{2}, \frac{R}{2}]^d.$$

Nach Definition 10.3.28 ist Q_j ein regulärer Quader mit Seitenlänge $2 \cdot \frac{R}{2} = R$. Nun zeigen wir (539).

Zu „ $\subseteq$ “ in (539). Sei $x \in Q$. Dann gilt also $(x - x_0)^{(i)} \in [-R, R]$ für alle $i \in \{1, \dots, d\}$. Definiere nun

$$\zeta^{(i)} := \begin{cases} -\frac{R}{2} & \text{falls } (x - x_0)^{(i)} \in [-R, 0) \\ \frac{R}{2} & \text{falls } (x - x_0)^{(i)} \in [0, R] \end{cases}. \quad (540)$$

Der Vektor $(\zeta^{(1)}, \dots, \zeta^{(d)})^T$ liegt somit in $\{-\frac{R}{2}, \frac{R}{2}\}^d$. Daher gibt es ein $j \in \{1, \dots, 2^d\}$ so, dass $(\zeta^{(1)}, \dots, \zeta^{(d)})^T = x_j$.

$$\text{Behauptung. } x \in Q_j = (x_0 + x_j) + [-\frac{R}{2}, \frac{R}{2}]^d.$$

Dazu prüfen wir Definition 10.3.29. Für $i \in \{1, \dots, d\}$ gilt

$$\begin{aligned} (x - (x_0 + x_j))^{(i)} &= (x - x_0)^{(i)} - x_j^{(i)} = (x - x_0)^{(i)} - \zeta^{(i)} \\ &\stackrel{(540)}{=} \begin{cases} (x - x_0)^{(i)} + \frac{R}{2} & \text{falls } (x - x_0)^{(i)} \in [-R, 0) \\ (x - x_0)^{(i)} - \frac{R}{2} & \text{falls } (x - x_0)^{(i)} \in [0, R] \end{cases} \end{aligned}$$

Man überzeugt sich leicht, dass in beiden Fällen ein Wert in $[-\frac{R}{2}, \frac{R}{2}]$ angenommen wird. Also haben wir gezeigt, dass $(x - (x_0 + x_j))^{(i)} \in [-\frac{R}{2}, \frac{R}{2}]$ für alle $i = 1, \dots, d$. Die Behauptung folgt.

Nun noch zu „ $\supseteq$ “ in (539). Man überzeugt sich leicht davon, dass $Q_j \subseteq Q$ für alle $j = 1, \dots, 2^d$. Die Inklusion ist damit leicht gezeigt!

^awie man wohl eigentlich mit Induktion nach d zeigen müsste

Nun sind wir bereit die Richtung „ $\implies$ “ von Satz 10.3.27 zu beweisen.

Beweis von Satz 10.3.27 „ $\implies$ “. Es sei $K \subseteq \mathbb{R}^d$ kompakt im Sinne von Definition 10.3.15.

ZZ: K ist überdeckungskompakt im Sinne von Definition 10.3.25.

Es sei dazu $(U_i)_{i \in I}$ eine beliebige offene Überdeckung von K . Zu zeigen ist, dass es eine endliche Teilüberdeckung gibt, d.h. es gibt ein $I_0 \subseteq I$ endlich, sodass $(U_i)_{i \in I_0}$ auch noch eine Überdeckung von K ist. Für einen Widerspruch nehmen wir an, dass es ein solches I_0 nicht gibt. Da K beschränkt ist, gibt es ein $R > 0$ so, dass $K \subseteq [-R, R]^d$.^a Insbesondere gilt $K = K \cap [-R, R]^d$. Es seien nun $Q_1, \dots, Q_{2^d}$ reguläre Quader mit Seitenlänge R so dass $[-R, R]^d = Q_1 \cup \dots \cup Q_{2^d}$. Definiere für $j = 1, \dots, 2^d$

$$K_j^{(1)} := K \cap Q_j.$$

Klarerweise ist $(U_i)_{i \in I}$ auch eine Überdeckung von $K_j^{(1)}$. Beachte, dass für alle $j = 1, \dots, 2^d$ die Mengen $K_j^{(1)}$ als Schnitte kompakter Mengen kompakt sind, vgl. Proposition 10.3.16. Außerdem gilt $K = K_1^{(1)} \cup \dots \cup K_{2^d}^{(1)}$.

ZWISCHENBEHAUPTUNG. Mindestens eine der Mengen $K_1^{(1)}, \dots, K_{2^d}^{(1)}$ kann nicht durch endlich viele Mengen aus der Familie $(U_i)_{i \in I}$ überdeckt werden.

In der Tat: Könnten alle diese Mengen durch endlich viele Mengen überdeckt werden, d.h. gäbe es für alle $j = 1, \dots, 2^d$ ein $I_j \subset I$ endlich mit $(U_i)_{i \in I_j}$ ist Überdeckung von $K_j^{(1)}$, so wäre

$$K = K_1^{(1)} \cup \dots \cup K_{2^d}^{(1)} \subseteq \bigcup_{i \in I_1} U_i \cup \dots \cup \bigcup_{i \in I_{2^d}} U_i = \bigcup_{i \in I_0} U_i,$$

wobei wir hier $I_0 = I_1 \cup \dots \cup I_{2^d}$ gesetzt haben. Dieses I_0 ist dann aber als Vereinigung endlich vieler endlicher Mengen wiederum endlich. Somit wäre $(U_i)_{i \in I_0}$ eine endliche Teilüberdeckung. Ein Widerspruch zur Annahme, dass es eine solche endliche Teilüberdeckung nicht gibt.

ZWISCHENBEHAUPTUNG ENDE.

Es sei nun $j_0 \in \{1, \dots, 2^d\}$ so, dass $K_{j_0}^{(1)}$ nicht durch endlich viele Mengen der Familie $(U_i)_{i \in I}$ überdeckt werden kann. Definiere $Q^{(2)} := Q_{j_0}$ und $K^{(2)} := K_{j_0}^{(1)} \subseteq K$. Nun wähle mit Lemma 10.3.29 reguläre Quader $Q_1^{(2)}, \dots, Q_{2^d}^{(2)}$ mit Seitenlänge $\frac{R}{2}$ so, dass $Q^{(2)} = Q_1^{(2)} \cup \dots \cup Q_{2^d}^{(2)}$. Definiere für $j = 1, \dots, 2^d$

$$K_j^{(2)} := K^{(2)} \cap Q_j^{(2)}.$$

Klarerweise sind diese Mengen alle kompakt (vgl. Proposition 10.3.16) und $(U_i)_{i \in I}$ ist immer noch eine Überdeckung von $K_j^{(2)}$ für alle $j = 1, \dots, 2^d$. Außerdem gilt $K^{(2)} = K_1^{(2)} \cup \dots \cup K_{2^d}^{(2)}$.

ZWISCHENBEHAUPTUNG. Mindestens eine der Mengen $K_1^{(2)}, \dots, K_{2^d}^{(2)}$ kann nicht durch endlich viele Mengen in der Familie $(U_i)_{i \in I}$ überdeckt werden.

Der Beweis hierzu ist analog zu der vorigen Zwischenbehauptung: Könnte man alle diese Mengen endlich überdecken, so könnte man auch die Vereinigung endlich überdecken. Die Vereinigung ist aber $K^{(2)}$ und für $K^{(2)}$ finden wir keine endliche Teilüberdeckung.

ZWISCHENBEHAUPTUNG ENDE.

Wähle nun wieder ein $j_0 \in \{1, \dots, 2^d\}$ so, dass $K_{j_0}^{(2)}$ nicht endlich überdeckt werden kann und definiere $Q^{(3)} := Q_{j_0}^{(2)}$ und $K^{(3)} := K_{j_0}^{(2)}$. Wähle wiederum mit Lemma 10.3.29 reguläre Quader $Q_1^{(3)}, \dots, Q_{2^d}^{(3)}$ mit Seitenlänge $\frac{R}{4}$ so, dass $Q^{(3)} = Q_1^{(3)} \cup \dots \cup Q_{2^d}^{(3)}$. Definiere für $j = 1, \dots, 2^d$

$$K_j^{(3)} := K^{(3)} \cap Q_j^{(3)}$$

Wiederum sind diese Mengen alle kompakt (vgl. Proposition 10.3.16) und $(U_i)_{i \in I}$ ist immer noch eine Überdeckung von $K_j^{(3)}$ für alle $j = 1, \dots, 2^d$. Außerdem gilt $K^{(3)} = K_1^{(3)} \cup \dots \cup K_{2^d}^{(3)}$. Wie gehabt zeigen wir dann

Mindestens eine der Mengen $K_1^{(3)}, \dots, K_{2^d}^{(3)}$ kann nicht durch endlich viele Mengen in der Familie $(U_i)_{i \in I}$ überdeckt werden.

Wiederum würden wir dann ein $j_0 \in \{1, \dots, 2^d\}$ suchen sodass $K_{j_0}^{(3)}$ nicht endlich überdeckt werden kann und definieren $Q^{(4)} := Q_{j_0}^{(3)}$ und $K^{(4)} := K_{j_0}^{(3)}$.

Wir iterieren den Prozess nun und machen immer weiter. Wir finden dadurch eine Folge von Mengen $K^{(1)} := K, K^{(2)}, K^{(3)}, K^{(4)}, \dots$ so, dass für alle $n \in \mathbb{N}$ gilt

- $K^{(n)}$ ist kompakt
- $K^{(n+1)} \subseteq K^{(n)}$ für alle $n \in \mathbb{N}$.
- $K^{(n)}$ ist in einem regulären Quader mit Seitenlänge $\frac{R}{2^{n-2}}$ enthalten.

- $K^{(n)}$ kann nicht durch endlich viele Mengen in der Familie $(U_i)_{i \in I}$ überdeckt werden.

Wegen den ersten zwei Punkten ist $(K^{(n)})_{n \in \mathbb{N}}$ eine geschachtelte Familie kompakter Mengen. Nach Proposition 10.3.22 gibt es dann ein Element x mit

$$x \in \bigcap_{n \in \mathbb{N}} K^{(n)}. \quad (541)$$

Da dann insbesondere $x \in K$ ist und $(U_i)_{i \in I}$ eine Überdeckung von K ist, muss es ein $i_0 \in I$ geben mit $x \in U_{i_0}$. Da U_{i_0} offen ist, ist x auch ein innerer Punkt von U_{i_0} , d.h.

$$\exists r > 0 : \quad B_r(x) \subseteq U_{i_0}. \quad (542)$$

Wähle nun $n_0 \in \mathbb{N}$ so groß, dass

$$\frac{R\sqrt{d}}{2^{n_0-3}} < r. \quad (543)$$

BEHAUPTUNG. $K^{(n_0)} \subseteq U_{i_0}$.

Diese Behauptung würde zu einem Widerspruch führen, denn $K^{(n_0)}$ kann ja nicht durch endlich viele Mengen der Familie $(U_i)_{i \in I}$ überdeckt werden — also schonmal gar nicht nur durch eine einzige! In der Behauptung steht aber gerade, dass $K^{(n)}$ durch die Menge U_{i_0} überdeckt wird. Ein Widerspruch! Nun also zum Beweis. Es sei $y \in K^{(n_0)}$. Es bezeichne Q^* einen regulären Quader mit Seitenlänge $\frac{R}{2^{n_0-2}}$ so, dass $K^{(n_0)} \subseteq Q^*$. Ferner sei x^* der Mittelpunkt von Q^* . Beachte: Nach (541) gilt $x \in K^{(n_0)}$ weswegen auch $x \in Q^*$. Somit gilt für $i \in \{1, \dots, d\}$

$$\begin{aligned} (x - y)^{(i)} &= (x - x^* - (y - x^*))^{(i)} \\ &= \underbrace{(x - x^*)^{(i)}}_{\in [-\frac{R}{2^{n_0-2}}, \frac{R}{2^{n_0-2}}]} - \underbrace{(y - x^*)^{(i)}}_{\in [-\frac{R}{2^{n_0-2}}, \frac{R}{2^{n_0-2}}]} \in \left[-2\frac{R}{2^{n_0-2}}, 2\frac{R}{2^{n_0-2}}\right] = \left[-\frac{R}{2^{n_0-3}}, \frac{R}{2^{n_0-3}}\right]. \end{aligned}$$

Wir folgern

$$\|x - y\| = \sqrt{\sum_{i=1}^d |(x - y)^{(i)}|^2} \leq \sqrt{\sum_{i=1}^d \left(\frac{R}{2^{n_0-3}}\right)^2} = \sqrt{\left(\frac{R}{2^{n_0-3}}\right)^2 \cdot d} = \frac{R\sqrt{d}}{2^{n_0-3}} \stackrel{(543)}{<} r.$$

Mit anderen Worten gilt $y \in B_r(x)$. Nach (542) gilt dann auch

$$y \in B_r(x) \subseteq U_{i_0}.$$

Wir folgern, da $y \in K^{(n_0)}$ beliebig war, dass $K^{(n_0)} \subseteq U_{i_0}$ und somit folgt die Behauptung. Wie bereits hinter der Behauptung erklärt, haben wir dann den Beweis abgeschlossen.

^aIn der Tat: Wir wissen dass es ein $L > 0$ gibt so, dass $\|x\| \leq L$ für alle $x \in K$. Sei nun $x \in K$ und $i \in \{1, \dots, d\}$. Dann gilt $|x^{(i)}| \leq \|x\| \leq L$. Das bedeutet nichts anderes als $x^{(i)} \in [-L, L]$. Man folgert $x \in [-L, L]^d$. Mit $R := L$ folgt die Behauptung.

Beweis von „ $\Leftarrow$ “ in Satz 10.3.27. Dies geht wesentlich schneller als die andere Richtung. Hierzu nehmen wir uns eine Überdeckungskompakte Menge $M \subseteq \mathbb{R}^d$. Wir können für diese dann einzeln ihre Abgeschlossenheit und Beschränktheit zeigen. Somit folgt die Kompaktheit im Sinne von Definition 10.3.15. Zuerst zur Beschränktheit.

Proposition 10.3.30

Es sei $M \subseteq \mathbb{R}^d$ überdeckungskompakt. Dann ist M beschränkt.

Beweis. Man überlege sich zunächst, dass

$$\mathbb{R}^d = \bigcup_{n \in \mathbb{N}} B_n(0).$$

In der Tat: Ist $x \in \mathbb{R}^d$ beliebig vorgegeben, so gilt für jede natürliche Zahl $n_0 > \|x\|$, dass

$$\|x - 0\| = \|x\| < n_0 \quad \text{d.h.} \quad x \in B_{n_0}(0) \subseteq \bigcup_{n \in \mathbb{N}} B_n(0).$$

Sei nun $M \subseteq \mathbb{R}^d$ wie in der Aussage überdeckungskompakt. Nun gilt offenbar

$$M \subseteq \mathbb{R}^d = \bigcup_{n \in \mathbb{N}} B_n(0).$$

Da $B_n(0)$ für alle $n \in \mathbb{N}$ offen ist (vgl. Beispiel 10.3.7), ist also $(B_n(0))_{n \in \mathbb{N}}$ eine offene Überdeckung von M . Wegen der Überdeckungskompaktheit (vgl. Definition 10.3.25) gibt es eine endliche Menge $I_0 \subseteq \mathbb{N}$, etwa $I_0 = \{n_1, \dots, n_k\}$ ($k \in \mathbb{N}$), so, dass auch $(B_n(0))_{n \in I_0}$ eine Überdeckung ist, d.h.

$$M \subseteq \bigcup_{n \in I_0} B_n(0) = B_{n_1}(0) \cup \dots \cup B_{n_k}(0). \quad (544)$$

Definiere nun $L := \max\{n_1, \dots, n_k\}$. Dann gilt $B_{n_j}(0) \subseteq B_L(0)$ für alle $j \in \{1, \dots, k\}$. Nach (544) gilt dann auch

$$M \subseteq B_L(0).$$

Insbesondere gilt für alle $x \in M$, dass

$$\|x\| = \|x - 0\| < L.$$

Die behauptete Beschränktheit folgt.

Nun noch zur Abgeschlossenheit.

Proposition 10.3.31

Es sei $M \subseteq \mathbb{R}^d$ überdeckungskompakt. Dann ist M abgeschlossen.

Beweis. Es sei $M \subseteq \mathbb{R}^d$ überdeckungskompakt. Wir zeigen, dass $M^C := \mathbb{R}^d \setminus M$ offen ist. Die Abgeschlossenheit von M folgt dann mit Proposition 10.3.10. Sei dazu $y \in M^C$ beliebig. Zu zeigen ist, dass y ein innerer Punkt von M^C ist, d.h.

$$\underline{Z}: \quad \exists r > 0 : \quad B_r(y) \subseteq M^C.$$

Für $x \in M$ definiere $r_x := \frac{1}{2}\|x - y\|$. Dann gilt $r_x > 0$, denn $x \neq y$ impliziert $\|x - y\| > 0$.

$$\text{ZWISCHENBEHAUPTUNG. } B_{r_x}(x) \cap B_{r_x}(y) = \emptyset.$$

In der Tat: Gäbe es einen Punkt $w \in B_{r_x}(x) \cap B_{r_x}(y)$ so wäre

$$\|x - y\| = \|x - w + w - y\| \leq \underbrace{\|x - w\|}_{< r_x, \text{ da } w \in B_{r_x}(x)} + \underbrace{\|w - y\|}_{< r_x, \text{ da } w \in B_{r_x}(y)} < 2r_x = 2 \cdot \frac{1}{2}\|x - y\| = \|x - y\|.$$

Wir erhalten den Widerspruch $\|x - y\| < \|x - y\|$.

$$\text{ZWISCHENBEHAUPTUNG ENDE.}$$

Nun gilt

$$M \subseteq \bigcup_{x \in M} \{x\} \subseteq \bigcup_{x \in M} B_{r_x}(x).$$

Da $r_x > 0$ ist $B_{r_x}(x)$ offen (vgl. Beispiel 10.3.7) und somit ist $(B_{r_x}(x))_{x \in M}$ eine offene Überdeckung von M . Wegen der Überdeckungskompaktheit gibt es dann eine endliche Teilmenge $\{x_1, \dots, x_k\}$ von M so, dass

$$M \subseteq B_{r_{x_1}}(x_1) \cup \dots \cup B_{r_{x_k}}(x_k). \quad (545)$$

Wähle nun $r := \min\{r_{x_1}, \dots, r_{x_k}\} > 0$. Wir behaupten, dass $B_r(y) \subseteq M^C$. In der Tat: Gäbe es ein $z \in B_r(y)$ mit $z \notin M^C$ (d.h. $z \in M$) so gäbe es wegen (545) auch ein $j \in \{1, \dots, k\}$ mit $z \in B_{r_{x_j}}(x_j)$. Dann wäre aber

$$z \in B_{r_{x_j}}(x_j) \cap B_r(y) \stackrel{r \leq r_{x_j}}{\subseteq} B_{r_{x_j}}(x_j) \cap B_{r_{x_j}}(y) \stackrel{\text{Zwischenbehauptung}}{=} \emptyset.$$

Wir erhalten einen Widerspruch zur Annahme und folgern $B_r(y) \subseteq M^C$. Das war zu zeigen.

Beweis von Satz 10.3.27, „ $\Leftarrow$ “. Es sei $M \subseteq \mathbb{R}^d$ überdeckungskompakt. Mit Proposition 10.3.30 folgt, dass M beschränkt ist und mit Proposition 10.3.31 folgt, dass M abgeschlossen ist. Nach Definition 10.3.15 ist M dann kompakt. Die Behauptung folgt.

10.4 Funktionsgrenzwerte in mehreren Variablen

Wir haben nun über Folgen in $\mathbb{R}^d$ gesprochen. Genau wie in der eindimensionalen Analysis erklären wir dann wieder was Funktionsgrenzwerte sind und Stetigkeit bedeutet. Unsere Funktionen gehen in diesem Kapitel stets von einer Menge $D \subseteq \mathbb{R}^d$ nach $\mathbb{R}^m$, $d, m \in \mathbb{N}$. Bei einer solchen Funktion $f : D \rightarrow \mathbb{R}^m$ und $(x^{(1)}, \dots, x^{(d)})^T \in D$ definieren wir als Kurznotation

$$f(x^{(1)}, \dots, x^{(d)}) := f((x^{(1)}, \dots, x^{(d)})^T) = f\left(\begin{pmatrix} x^{(1)} \\ \vdots \\ x^{(d)} \end{pmatrix}\right)$$

Für $d = 2$ bzw. $d = 3$ schreibt man dann auch häufig $f(x, y)$ bzw. $f(x, y, z)$. Für jedes Argument $x = (x^{(1)}, \dots, x^{(d)}) \in D$ ist $f(x^{(1)}, \dots, x^{(d)})$ ein Vektor in $\mathbb{R}^m$. Dessen Koordinaten notieren wir standardmäßig so:

$$f(x^{(1)}, \dots, x^{(d)}) = \begin{pmatrix} f^{(1)}(x^{(1)}, \dots, x^{(d)}) \\ \vdots \\ f^{(m)}(x^{(1)}, \dots, x^{(d)}) \end{pmatrix}$$

Hierbei kann man dann die sogenannten *Komponentenfunktionen* $f^{(1)}, \dots, f^{(m)} : D \rightarrow \mathbb{R}$ als Funktionen von D nach $\mathbb{R}$ betrachten.⁸

Funktionsgrenzwerte

Definition 10.4.1 — FUNKTIONSGRENZWERTE

Es sei $D \subseteq \mathbb{R}^d$ und $f : D \rightarrow \mathbb{R}^m$ eine Funktion. Ferner sei $x_0 \in \mathbb{R}^d$ ein Berührungspunkt von D . Wir sagen f besitzt bei x_0 den Grenzwert $y_0 \in \mathbb{R}^m$ falls für jede Folge $(x_n)_{n \in \mathbb{N}} \subseteq D$ mit $\lim_{n \rightarrow \infty} x_n = x_0$ gilt, dass $\lim_{n \rightarrow \infty} f(x_n) = y_0$. Trifft obiges zu so schreibt man

$$\lim_{x \rightarrow x_0} f(x) = y_0 \quad \text{oder auch} \quad \lim_{\substack{x \rightarrow x_0 \\ x \in D}} f(x) = y_0$$

Außerdem führen wir auch die folgenden Schreibweisen ein:

$$f(x) \xrightarrow{x \rightarrow x_0} y_0 \quad \text{und} \quad [f(x) \longrightarrow y_0 \quad (x \rightarrow x_0)]$$

⁸Hier mal ein GEOGEBRA-Beispiel für eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$

Will man alle Koordinaten des vektorwertigen Argumentes $x = (x^{(1)}, \dots, x^{(d)})^T$ einzeln auflisten, so schreibt man auch oft

$$\lim_{(x^{(1)}, \dots, x^{(d)}) \rightarrow (x_0^{(1)}, \dots, x_0^{(d)})} f(x^{(1)}, \dots, x^{(d)}).$$

In $\mathbb{R}^d$ haben wir unendlich viele Richtungen zur Verfügung, von denen eine Folge $(x_n)_{n \in \mathbb{N}}$ einen Wert approximieren kann. Alle diese Richtungen müssen überprüft werden.

Beispiel 10.4.2

Definiere^a $f : \mathbb{R}^2 \setminus \{(0, 0)^T\} \rightarrow \mathbb{R}$ durch

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

Für einen Plot von f siehe Abbildung 44.

Behauptung: $\lim_{(x, y) \rightarrow (0, 0)} f(x, y)$ existiert nicht.

Beweis. Es sei für $n \in \mathbb{N}$ die $a_n := (\frac{1}{n}, 0)^T$ und $b_n := (\frac{1}{n}, \frac{1}{n})^T$. Es gilt $a_n \xrightarrow{n \rightarrow \infty} (0, 0)^T$ und $b_n \xrightarrow{n \rightarrow \infty} (0, 0)^T$. Nun gilt für alle $n \in \mathbb{N}$

$$f(a_n) = f\left(\frac{1}{n}, 0\right) = \frac{\frac{1}{n} \cdot 0}{\left(\frac{1}{n}\right)^2 + 0^2} = \frac{0}{\frac{1}{n^2}} = 0$$

und somit $\lim_{n \rightarrow \infty} f(a_n) = 0$. Ferner gilt für alle $n \in \mathbb{N}$

$$f(b_n) = f\left(\frac{1}{n}, \frac{1}{n}\right) = \frac{\frac{1}{n} \cdot \frac{1}{n}}{\left(\frac{1}{n}\right)^2 + \left(\frac{1}{n}\right)^2} = \frac{\frac{1}{n^2}}{\frac{1}{n^2} + \frac{1}{n^2}} = \frac{\frac{1}{n^2}}{\frac{2}{n^2}} = \frac{1}{2}$$

und somit $\lim_{n \rightarrow \infty} f(b_n) = \frac{1}{2}$. Der Funktionsgrenzwert kann daher nicht existieren, denn würde $\lim_{(x, y) \rightarrow (0, 0)} f(x, y) =: \xi$ existieren so wäre nach Definition 10.4.1 (und weil $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = (0, 0)^T$)

$$\lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} f(b_n) = \xi.$$

Eine weitere wichtige Erkenntnis, die man aus diesem Beispiel ziehen kann ist, dass es für die Existenz des Grenzwertes $\lim_{(x, y) \rightarrow (0, 0)} f(x, y)$ NICHT ausreicht, nur Grenzwerte entlang der Koordinatenrichtungen zu betrachten. In der Tat: Für f wie oben gilt

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

denn für alle $t \in \mathbb{R} \setminus \{0\}$ hat man

$$f(t, 0) = \frac{t \cdot 0}{t^2 + 0^2} = 0 \quad \text{und} \quad f(0, t) = \frac{0 \cdot t}{0^2 + t^2} = 0.$$

FAZIT. Unser Grenzwertbegriff kann nicht auf Grenzwerte entlang einzelner Koordinatenrichtungen zurückgeführt werden. Das ist insofern spannend, weil man bei der Definition der Folgenkonvergenz gerade von dieser Rückführung Gebrauch macht, vgl. Definition 10.2.2. Mal wieder zeigt sich, dass Funktionsgrenzwerte ein stärkeres Konzept sind als Folgengrenzwerte.

^aEin weiteres Beispiel dieser Art finden Sie auf GEOGEBRA.

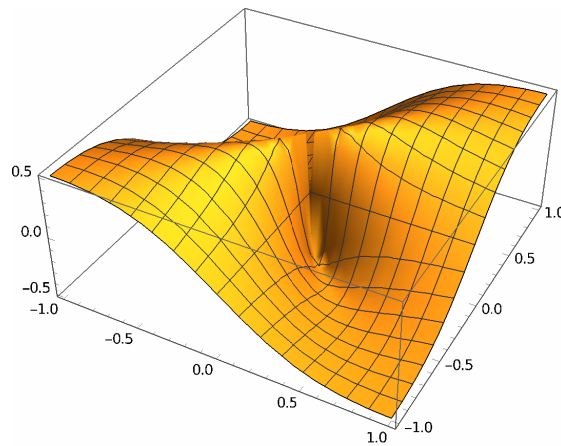


Abbildung 44: Die Funktion aus Beispiel 10.4.2.

Wir haben gesehen, dass es im Definitionsbereich für die Grenzwertbildung nicht ausreicht, nur einzelne Koordinatenrichtungen zu betrachten. Im *Bildbereich* hingegen gelingt trotzdem eine solche Rückführung auf Koordinaten:

Proposition 10.4.3

Es sei $D \subseteq \mathbb{R}^d$, $x_0 \in \mathbb{R}^d$ ein Berührungspunkt von D und $f : D \rightarrow \mathbb{R}^m$ eine Funktion. Ferner seien $f^{(1)}, \dots, f^{(m)} : D \rightarrow \mathbb{R}$ die Komponentenfunktionen. Sei nun $y_0 = (y_0^{(1)}, \dots, y_0^{(m)})^T$. Dann gilt

$$\lim_{x \rightarrow x_0} f(x) = y_0 \quad \Longleftrightarrow \quad \lim_{x \rightarrow x_0} f^{(i)}(x) = y_0^{(i)} \quad \forall i = 1, \dots, m$$

Beweis.

$$\lim_{x \rightarrow x_0} f(x) = y_0$$

$$\iff \text{Für jede Folge } (x_n)_{n \in \mathbb{N}} \subseteq D \text{ mit } \lim_{n \rightarrow \infty} x_n = x_0 \text{ gilt } \lim_{n \rightarrow \infty} f(x_n) = y_0. \quad (546)$$

Nun beachte, dass $(f(x_n))_{n \in \mathbb{N}} \subseteq \mathbb{R}^m$ eine Folge mit Komponentenfolgen $(f^{(1)}(x_n))_{n \in \mathbb{N}}, \dots, (f^{(m)}(x_n))_{n \in \mathbb{N}}$ ist. Mit Definition 10.2.2 können wir dann (546) äquivalent umschreiben und finden heraus

$$\lim_{x \rightarrow x_0} f(x) = y_0$$

$$\iff \text{Für jede Folge } (x_n)_{n \in \mathbb{N}} \subseteq D \text{ mit } \lim_{n \rightarrow \infty} x_n = x_0 \text{ gilt}$$

$$\lim_{n \rightarrow \infty} f^{(i)}(x_n) = y_0^{(i)} \text{ für alle } i \in \{1, \dots, m\}$$

$$\iff \text{Für alle } i \in \{1, \dots, m\} \text{ und für jede Folge } (x_n)_{n \in \mathbb{N}} \subseteq D$$

$$\text{mit } \lim_{n \rightarrow \infty} x_n = x_0 \text{ gilt } \lim_{n \rightarrow \infty} f^{(i)}(x_n) = y_0^{(i)}.$$

$$\iff \text{Für alle } i \in \{1, \dots, m\} \text{ gilt } \lim_{x \rightarrow x_0} f^{(i)}(x) = y_0^{(i)}.$$

Eine weitere nicht überraschende Tatsache ist der Zusammenhang zwischen Funktionsgrenzwerten und der Euklidischen Norm.

Proposition 10.4.4 — FUNKTIONSGRENZWERT UND NORM

Es sei $D \subseteq \mathbb{R}^d$, $x_0 \in \mathbb{R}^d$ ein Berührungspunkt von D und $f : D \rightarrow \mathbb{R}^m$ eine Funktion. Dann gilt für $y_0 \in \mathbb{R}^m$

$$\lim_{x \rightarrow x_0} f(x) = y_0 \iff \lim_{x \rightarrow x_0} \|f(x) - y_0\| = 0$$

Beweis. Dies ist eine direkte Folgerung aus Definition 10.4.1 und Satz 10.2.6 „(a) $\iff$ (b)“. Die Details sind den Lesenden als Übungsaufgabe überlassen.

Wir formulieren an dieser Stelle keine Grenzwertregeln für Funktionsgrenzwerte, obwohl man es könnte. Wir vertrauen darauf, dass Sie aus den Rechenregeln für die Folgenkonvergenz und mit den Techniken, die Sie gelernt haben, mittlerweile selbst Grenzwertregeln formulieren können. Was wir aber schon formulieren können ist ein ε - δ -Kriterium für mehrdimensionale Funktionsgrenzwerte.

Satz 10.4.5 — EPSILON-DELTA-KRITERIUM IN $\mathbb{R}^d$

Es sei $D \subseteq \mathbb{R}^d$ und $f : D \rightarrow \mathbb{R}^m$ eine Funktion. Ferner sei $x_0 \in \mathbb{R}^d$ ein Berührungspunkt von D . Dann gilt für $y_0 \in \mathbb{R}^m$

$$\lim_{x \rightarrow x_0} f(x) = y_0$$

$$\iff$$

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x \in D \text{ mit } \|x - x_0\|_{\mathbb{R}^d} < \delta \text{ gilt } \|f(x) - y_0\|_{\mathbb{R}^m} < \varepsilon. \quad (547)$$

Beweis. Der Beweis läuft genau so ab wie der Beweis von Satz 6.1.23 und muss daher an dieser Stelle nicht gezeigt werden.

Die obige Aussage lässt sich auch noch genau wie im Eindimensionalen mithilfe von *Umgebungen* erfassen.

Proposition 10.4.6 — UMGEBUNGSKRITERIUM IN $\mathbb{R}^d$

Es sei $D \subseteq \mathbb{R}^d$ und $f : D \rightarrow \mathbb{R}^m$ eine Funktion. Ferner sei $x_0 \in \mathbb{R}^d$ ein Berührungspunkt von D . Dann gilt für $y_0 \in \mathbb{R}^m$

$$\lim_{x \rightarrow x_0} f(x) = y_0$$

$$\iff$$

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad f(D \cap B_\delta^d(x_0)) \subseteq B_\varepsilon^m(y_0). \quad (a)$$

^aFür die Notationen $B_\delta^d(x_0)$ und $B_\varepsilon^m(y_0)$ erinnern wir an Definition 10.1.4

Beweis. Auch dieser Beweis ist eine Kopie aus der eindimensionalen Analysis (Proposition 6.1.27) und kann an dieser Stelle weggelassen werden.

10.06.2026

Stetigkeit

Wenn Funktionsgrenzwerte definiert sind, kann man auch natürlicherweise über Stetigkeit sprechen.

Definition 10.4.7 — STETIGKEIT IM MEHRDIMENSIONALEN

Es sei $D \subseteq \mathbb{R}^d$, $f : D \rightarrow \mathbb{R}^m$ eine Funktion und $x_0 \in D$. Wir sagen f ist stetig im Punkt x_0 falls

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

Wir nennen $f : D \rightarrow \mathbb{R}^m$ stetig auf D falls f in jedem Punkt aus D stetig ist.

Man nochmal mithilfe von Definition 10.4.1 ausführlicher schreiben was Stetigkeit bei einem Punkt x_0 bedeutet: Für jede Folge $(x_n)_{n \in \mathbb{N}} \subseteq D$ mit $\lim_{n \rightarrow \infty} x_n = x_0$ gilt $\lim_{n \rightarrow \infty} f(x_n) = f(x_0)$.

Bemerkung 10.4.8

Mal wieder kann man das Epsilon-Delta-Kriterium (Satz 10.4.5) auch als Kriterium für die mehrdimensionale Stetigkeit sehen: Für $D \subseteq \mathbb{R}^d$ ist eine Funktion $f : D \rightarrow \mathbb{R}^m$ genau dann stetig bei $x_0 \in D$ wenn

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x \in D \text{ mit } \|x - x_0\|_{\mathbb{R}^d} < \delta \text{ gilt } \|f(x) - f(x_0)\|_{\mathbb{R}^m} < \varepsilon.$$

Auch äquivalent ist das Umgebungskriterium (vgl. Proposition 10.4.6), d.h.

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad f(D \cap B_\delta^d(x_0)) \subseteq B_\varepsilon^m(f(x_0)).$$

Beispiel 10.4.9

Betrachte $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch

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

Wie wir bereits in Beispiel 10.4.2 gesehen haben ist

$$\lim_{n \rightarrow \infty} f\left(\frac{1}{n}, \frac{1}{n}\right) = \lim_{n \rightarrow \infty} \frac{\frac{1}{n} \cdot \frac{1}{n}}{\left(\frac{1}{n}\right)^2 + \left(\frac{1}{n}\right)^2} = \frac{1}{2} \neq 0 = f(0, 0).$$

Die Funktion f ist also bei $x_0 = (0, 0)^T$ nicht stetig. Und das, obwohl die Funktionen

$$t \mapsto f(t, 0) \quad \text{und} \quad t \mapsto f(0, t)$$

beide in $t = 0$ stetig wären.^a Wie man nach Beispiel 10.4.2 erwarten würde reicht es also auch hier nicht aus, die Funktion nur entlang von Koordinatenrichtungen zu betrachten.

^aSchließlich sind beide Funktionen konstant gleich 0 und daher sicherlich stetig. Dies zu überprüfen überlassen wir den Lesenden als Übungsaufgabe und verweisen auf Beispiel 10.4.2.

Im Bildbereich ist die Rückführung auf Koordinaten allerdings wieder erlaubt.

Bemerkung 10.4.10

Es sei $D \subseteq \mathbb{R}^d$, $x_0 \in D$ und $f : D \rightarrow \mathbb{R}^m$ eine Funktion. Ferner seien $f^{(1)}, \dots, f^{(m)} : D \rightarrow \mathbb{R}$ die Komponentenfunktionen. Aus Proposition 10.4.3 lernt man, dass f bei x_0 stetig ist genau dann wenn $f^{(1)}, \dots, f^{(m)}$ bei x_0 stetig sind.

Ein Paradebeispiel für eine stetige Funktion sind die *Koordinatenfunktionen*.

Beispiel 10.4.11

Es sei $i \in \{1, \dots, d\}$ beliebig und

$$f : \mathbb{R}^d \rightarrow \mathbb{R} \quad f(x) = x^{(i)}$$

Behauptung. f ist auf ganz $\mathbb{R}^d$ stetig.

Beweis. Es sei $x_0 \in \mathbb{R}^d$. Wir zeigen, dass f in x_0 stetig ist, d.h.

$$\lim_{x \rightarrow x_0} f(x) = f(x_0). \quad (548)$$

Dazu sei $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ eine Folge mit $\lim_{n \rightarrow \infty} x_n = x_0$. Nach Definition 10.2.2 gilt dann auch

$$\lim_{n \rightarrow \infty} x_n^{(i)} = x_0^{(i)}. \quad (549)$$

Das bedeutet

$$\lim_{n \rightarrow \infty} f(x_n) = \lim_{n \rightarrow \infty} x_n^{(i)} \stackrel{(549)}{=} x_0^{(i)} = f(x_0).$$

Da $(x_n)_{n \in \mathbb{N}}$ eine beliebige gegen x_0 konvergente Folge war, folgt (548).

An dieser Stelle können wir auch wieder Rechenregeln für die Stetigkeit wie im Eindimensionalen formulieren. Zum Beispiel die Stetigkeit der Verkettung:

Proposition 10.4.12 — KOMPOSITIONEN STETIGER FUNKTIONEN SIND STETIG

Es sei $D \subseteq \mathbb{R}^d$ und $f : D \rightarrow \mathbb{R}^m$ eine Funktion, die stetig bei einem $x_0 \in D$ ist. Ist ferner $\tilde{D} \subseteq \mathbb{R}^m$ mit $f(D) \subseteq \tilde{D}$ und $h : \tilde{D} \rightarrow \mathbb{R}^k$ stetig bei $f(x_0)$, so ist auch $(h \circ f) : D \rightarrow \mathbb{R}^k$ bei x_0 stetig.

Beweis. Dies kann genauso bewiesen werden wie im Eindimensionalen (vgl. Proposition 6.2.6(II)) und ist den Lesenden als Übungsaufgabe überlassen.

Stetigkeit von Summen, Produkten und Quotienten folgt nun direkt als Anwendung der obigen Kompositionsregel.

Korollar 10.4.13

Es sei $D \subseteq \mathbb{R}^d$ und $f, g : D \rightarrow \mathbb{R}^a$ beide stetig in einem Punkt $x_0 \in D$. Dann sind $f + g$ und $f \cdot g$ und (sofern $g(x_0) \neq 0$) auch $\frac{f}{g}$ stetig in x_0 .

^aDas bedeutet $m = 1!$ (Ansonsten sind die Operationen $f \cdot g$ und $\frac{f}{g}$ nämlich gar nicht definiert!)

Beweis. Auch dieser Beweis kann wie im Eindimensionalen geführt werden (vgl. Proposition 6.2.6(I)). Es gibt jedoch eine erfrischende andere Möglichkeit zu argumentieren, die wir im Folgenden präsentieren wollen.

SCHRITT 1. Zeige die mehrdimensionale Stetigkeit der Abbildungen

$$\text{add} : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \rightarrow x + y,$$

$$\text{mult} : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \rightarrow x \cdot y,$$

und

$$\text{div} : \mathbb{R}^2 \setminus \{(t, 0)^T : t \in \mathbb{R}\} \rightarrow \mathbb{R}, (x, y) \rightarrow \frac{x}{y}.$$

Wir zeigen das nur exemplarisch für add . Es sei $(a_0, b_0)^T \in \mathbb{R}^2$ beliebig und $((a_n, b_n)^T)_{n \in \mathbb{N}}$ eine beliebige Folge mit $\lim_{n \rightarrow \infty} (a_n, b_n)^T = (a_0, b_0)^T$. Das bedeutet nach Definition 10.2.2 auch

$$\lim_{n \rightarrow \infty} a_n = a_0 \quad \text{und} \quad \lim_{n \rightarrow \infty} b_n = b_0.$$

Zu zeigen ist, dass $\text{add}(a_n, b_n) \xrightarrow{n \rightarrow \infty} \text{add}(a_0, b_0)$. Dazu:

$$\text{add}(a_n, b_n) = \underbrace{a_n}_{\rightarrow a_0} + \underbrace{b_n}_{\rightarrow b_0} \xrightarrow[\text{GWR}]{n \rightarrow \infty} a_0 + b_0 = \text{add}(a_0, b_0).$$

Analog kann man auch die Stetigkeit von mult und div zeigen.

SCHRITT 2. Fasse $f + g, f \cdot g, \frac{f}{g}$ als Verkettung auf.

Dazu: Definiere zunächst $h : D \rightarrow \mathbb{R}^2$ durch $h(x) := (f(x), g(x))$. Beachte: h ist stetig weil jede Komponentenfunktion (also f und g) nach Voraussetzung stetig ist, vgl. Bemerkung 10.4.10. Wir behaupten: Dann ist $f + g = \text{add} \circ h$, $f \cdot g = \text{mult} \circ h$ und $\frac{f}{g} = \text{div} \circ h$. In der Tat: Beispielsweise gilt für alle $x \in D$

$$(\text{add} \circ h)(x) = \text{add}(h(x)) = \text{add}(f(x), g(x)) = f(x) + g(x) = (f + g)(x).$$

Für die anderen Behauptungen kann man analog vorgehen.

SCHRITT 3. Beweis der Stetigkeit.

Dank Proposition 10.4.12 und Schritt 1 sind nun $\text{add} \circ h$, $\text{mult} \circ h$ und (falls $g(x_0) \neq 0$, d.h. $(f(x_0), g(x_0)) \notin \{(t, 0) : t \in \mathbb{R}\}$) auch $\text{div} \circ h$ stetig. Die Stetigkeit von $f + g, f \cdot g$ und $\frac{f}{g}$ (sofern $g(x_0) \neq 0$) folgt.

Ist der Definitionsbereich D offen, so können wir den obigen Stetigkeitsbegriff auch sehr schön und einfach *topologisch* charakterisieren. Diese Charakterisierung ist neu! *Topologisch* bedeutet hier dass es nur Begriffe aus der Topologie verwendet (d.h. offen, abgeschlossen und Mengenoperationen)

Satz 10.4.14 — TOPOLOGISCHE CHARAKTERISIERUNG DER STETIGKEIT

Es sei $D \subseteq \mathbb{R}^d$ eine offene Menge und $f : D \rightarrow \mathbb{R}^m$ eine Funktion. Dann gilt

f ist auf D stetig $\iff$ Für jede offene Menge $U \subseteq \mathbb{R}^m$ ist $f^{-1}(U) \subseteq \mathbb{R}^d$ wieder offen.

Beweis. Zuerst zu „ $\implies$ “. Es sei $f : D \rightarrow \mathbb{R}^m$ stetig.

Z: Für jede offene Menge $U \subseteq \mathbb{R}^m$ ist $f^{-1}(U) \subseteq \mathbb{R}^d$ wieder offen.

Es sei dazu $U \subseteq \mathbb{R}^m$ eine offene Menge. Wir wollen zeigen, dass $f^{-1}(U)$ offen ist. Dazu nehmen wir ein beliebiges $x_0 \in f^{-1}(U)$, d.h. $x_0 \in D$ mit $f(x_0) \in U$. Wir müssen nach Definition 10.3.5 zeigen, dass x_0 ein innerer Punkt von $f^{-1}(U)$ ist, d.h.

$$\text{Z: } \exists \theta > 0 : B_\theta(x_0) \subseteq f^{-1}(U). \quad (550)$$

Wie wir bereits angemerkt haben gilt $f(x_0) \in U$. Wegen der Offenheit von U ist $f(x_0)$ sogar ein innerer Punkt von U , d.h.

$$\exists \varepsilon > 0 : B_\varepsilon(f(x_0)) \subseteq U. \quad (551)$$

Mit Bemerkung 10.4.8 gibt es aber ein $\delta > 0$ so, dass $f(D \cap B_\delta^d(x_0)) \subseteq B_\varepsilon(f(x_0))$. Wegen (551) gilt dann aber

$$f(D \cap B_\delta(x_0)) \subseteq U. \quad (552)$$

Betrachten wir auf beiden Seiten das Urbild und nutzen Proposition 1.3.7(i) so gilt

$$D \cap B_\delta(x_0) \underset{\text{Prop 1.3.7(i)}}{\subseteq} f^{-1}(f(D \cap B_\delta(x_0))) \underset{(552)}{\subseteq} f^{-1}(U). \quad (553)$$

Nun ist $x_0 \in D$. Weil D offen ist, ist x_0 auch ein innerer Punkt von D und wir folgern dass es ein $r > 0$ gibt mit $B_r(x_0) \subseteq D$. Definiere nun $\theta := \min\{r, \delta\} > 0$. Dann gilt (wie man sich leicht als Übungsaufgabe überlegt)

$$B_\theta(x_0) \subseteq B_r(x_0) \cap B_\delta(x_0) \subseteq D \cap B_\delta(x_0) \underset{(553)}{\subseteq} f^{-1}(U).$$

Somit ist (550) nachgewiesen. Da x_0 beliebig war folgt die Offenheit von $f^{-1}(U)$.

Nun noch zu „ $\impliedby$ “. Es sei nun $f : D \rightarrow \mathbb{R}^m$ so, dass jede offene Menge $U \subseteq \mathbb{R}^m$ ein offenes Urbild $f^{-1}(U) \subseteq \mathbb{R}^d$ besitzt. Zu zeigen ist, dass f auf D stetig ist. Hierzu sei $x_0 \in D$ beliebig. Um die Stetigkeit von f bei x_0 zu zeigen weisen wir das Umgebungskriterium nach, d.h. (vgl. Bemerkung 10.4.8)

$$\text{Z: } \forall \varepsilon > 0 \quad \exists \delta > 0 : f(D \cap B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0)).$$

Fixiere hierzu ein $\varepsilon > 0$. Betrachte die Menge $U := B_\varepsilon(f(x_0))$. Diese ist nach Beispiel 10.3.7 offen. Nach unserer Voraussetzung ist dann auch $f^{-1}(U)$ offen. Beachte: $x_0 \in f^{-1}(U)$, da $f(x_0) \in B_\varepsilon(f(x_0)) = U$. Somit ist x_0 auch ein innerer Punkt von $f^{-1}(U)$. Es folgt

$$\exists \delta > 0 : B_\delta(x_0) \subseteq f^{-1}(U).$$

Klarerweise gilt dann auch $D \cap B_\delta(x_0) \subseteq f^{-1}(U)$. Bildet man auf beiden Seiten das Bild unter f so findet man

$$f(D \cap B_\delta(x_0)) \subseteq f(f^{-1}(U)) \underset{\text{Prop 1.3.7(ii)}}{\subseteq} U.$$

Erinnern wir uns an die Definition von U so lernen wir

$$f(D \cap B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0)).$$

Wir haben also gesehen

$$\exists \delta > 0 : \quad f(D \cap B_\delta(x_0)) \subseteq B_\varepsilon(f(x_0)).$$

Die Behauptung folgt.

Es folgt ein spannendes Beispiel dazu mit einigen Verweisen auf die lineare Algebra.

Beispiel 10.4.15

^a Hat man eine Matrix $A = (a_{ij})_{i,j=1}^d \in \mathbb{R}^{d \times d}$ gegeben, so kann man sie als ein Vektor in $\mathbb{R}^{d^2}$ aufschreiben. Die *Determinante* einer Matrix kann dann als Abbildung

$$\det : \mathbb{R}^{d^2} \rightarrow \mathbb{R}, \quad A \mapsto \det(A)$$

aufgefasst werden. Man überzeugt sich leicht mit der sogenannten *Leibniz-Formel*, dass die Abbildung $\det$ stetig ist. Beachte: Die Menge $\mathbb{R} \setminus \{0\}$ ist offen (da $(\mathbb{R} \setminus \{0\})^C = \{0\}$ abgeschlossen ist). Somit ist nach Satz 10.4.14 das Urbild $\det^{-1}(\mathbb{R} \setminus \{0\})$ offen in $\mathbb{R}^{d^2}$. Ein Fakt aus der linearen Algebra besagt, dass für $A \in \mathbb{R}^{d \times d}$ gilt

$$\det(A) \neq 0 \quad \Longleftrightarrow \quad A \text{ ist invertierbar}$$

Dies zeigt, dass

$$\det^{-1}(\mathbb{R} \setminus \{0\}) = \{A \in \mathbb{R}^{d^2} : \det(A) \neq 0\} = \{A \in \mathbb{R}^{d^2} : A \text{ invertierbar}\}$$

Wir haben gesehen, dass die Menge aller invertierbaren Matrizen offen in $\mathbb{R}^{d^2}$ sind.

^aIn diesem Beispiel werden Kenntnisse aus der linearen Algebra vorausgesetzt. Eigentlich ist unser Skript in sich geschlossen und nutzt keine Resultate von anderen Vorlesungen. Dieses Bonus-Beispiel ist eine kleine Ausnahme.

Sätze über stetige Funktionen

Hier fragen wir uns, was für Sätze über stetige Funktionen wir aufs Mehrdimensionale verallgemeinern können. Im Wesentlichen haben wir in der Analysis I vier wichtige Sätze über stetige Funktionen besprochen

- Satz von Weierstraß (Satz 6.2.12)

- Zwischenwertsatz (Satz 6.2.15)
- Stetigkeit der Umkehrfunktion (Satz 6.2.20)
- Stetigkeit und gleichmäßige Stetigkeit (Satz 6.4.2)

Wir kommentieren im Folgenden alle vier Sätze und untersuchen sie auf ihre Verallgemeinerbarkeit in unser mehrdimensionales Setting.

Zum Zwischenwertsatz. Der Zwischenwertsatz macht für $\mathbb{R}^m$ -wertige Funktionen nicht direkt Sinn, denn $\mathbb{R}^m$ ist nicht mit einer Ordnungsrelation angeordnet und somit kann man nicht direkt davon sprechen, was es in $\mathbb{R}^m$ bedeutet *zwischen* zwei Vektoren zu liegen. Man könnte mehr auf diesen Sachverhalt eingehen, wenn man genauer über das Konzept einer *zusammenhängenden Menge* wüsste. Wir benötigen dies an dieser Stelle aber noch nicht und vertagen daher die Diskussion.

Zur Stetigkeit der Umkehrfunktion. Der Satz über die Stetigkeit der Umkehrfunktion Satz 6.2.20 besagte in einer Dimension folgendes: Ist $f : D \rightarrow Z$ stetig und bijektiv, so ist auch $f^{-1} : Z \rightarrow \mathbb{R}$ stetig. Eine Voraussetzung an D war hierbei, dass es sich um ein kompaktes Intervall $[a, b] \subseteq \mathbb{R}$ handeln muss. Diese hätte allerdings abgeschwächt werden können — man die Aussage durchaus auf offene/halboffene Intervalle verallgemeinern können! Eine solche Verallgemeinerung haben wir uns auch erarbeitet – zum Beispiel beim Beweis der Stetigkeit der Wurzel (Korollar 6.2.25)!

Leider lässt sich die Aussage von Satz 6.2.20 nicht ohne Weiteres auf höhere Dimensionen verallgemeinern!

Beispiel 10.4.16

Es sei $D := [0, 2\pi)$ und $Z := \mathbb{S}^1 := \{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 = 1\}$. Betrachte

$$f : [0, 2\pi) \rightarrow \mathbb{S}^1, \quad f(t) = (\cos(t), \sin(t))^T.$$

Klarerweise ist $f : D \rightarrow \mathbb{R}^2$ stetig (denn jede Komponentenfunktion ist stetig, vgl. Bemerkung 10.4.10). Dass $f : D \rightarrow \mathbb{S}^1$ auch bijektiv ist, haben wir in Prop 5.5.18 gezeigt.^a

Behauptung. $f^{-1} : \mathbb{S}^1 \rightarrow \mathbb{R}$ ist bei $(1, 0)^T \in \mathbb{S}^1$ unstetig.

Beweis. Betrachte

$$(a_n)_{n \in \mathbb{N}} := ((\cos(2\pi - \frac{1}{n}), \sin(2\pi - \frac{1}{n}))^T)_{n \in \mathbb{N}}. \quad (554)$$

Dann gilt wegen der Stetigkeit von $\cos$ und $\sin$

$$a_n \xrightarrow{n \rightarrow \infty} (\cos(2\pi), \sin(2\pi))^T = (1, 0)^T.$$

Wir behaupten, dass $f^{-1}(a_n) \xrightarrow{n \rightarrow \infty} f^{-1}((1, 0)^T)$. Dazu: Für alle $n \in \mathbb{N}$ gilt $f^{-1}(a_n) = 2\pi - \frac{1}{n}$, denn nach (554) gilt $a_n = f(2\pi - \frac{1}{n})$. Außerdem gilt $f^{-1}((1, 0)^T) = 0$, denn $f(0) = (\cos(0), \sin(0))^T = (1, 0)^T$. Daher gilt

$$\lim_{n \rightarrow \infty} f^{-1}(a_n) = 2\pi \neq 0 = f^{-1}((1, 0)^T).$$

Die behauptete Unstetigkeit folgt.

^aEin Bonuskapitel. Wer dieses nicht gelesen hat, kann an dieser Stelle auch einfach glauben, dass das Bild von f der gesamte Einheitskreis ist.

Ist der Definitionsbereich von f aber kompakt, so klappt es mit der Stetigkeit der Umkehrfunktion!

Proposition 10.4.17

Es sei $K \subseteq \mathbb{R}^d$ kompakt und $W \subseteq \mathbb{R}^m$. Ferner sei $f : K \rightarrow W$ stetig und bijektiv. Dann ist die Umkehrfunktion $f^{-1} : W \rightarrow K$ auch stetig.

Beweis. Es sei $y_0 \in W$.

$$\mathbb{Z}\mathbb{Z} : \quad \lim_{y \rightarrow y_0} f^{-1}(y) = f^{-1}(y_0).$$

Hierzu sei $(y_n)_{n \in \mathbb{N}} \subseteq W$ eine beliebige Folge mit $\lim_{n \rightarrow \infty} y_n = y_0$.

$$\mathbb{Z}\mathbb{Z} : \quad f^{-1}(y_n) \xrightarrow{n \rightarrow \infty} f^{-1}(y_0).$$

Angenommen das ist nicht der Fall. Das bedeutet

$$\neg \quad (\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad \|f^{-1}(y_n) - f^{-1}(y_0)\| \leq \varepsilon \quad \forall n \geq n_0)$$

Verneint ergibt sich

$$\exists \varepsilon_0 > 0 : \quad \forall N \in \mathbb{N} \quad \exists k_N \geq N : \quad \|f^{-1}(y_{k_N}) - f^{-1}(y_0)\| > \varepsilon_0. \quad (555)$$

Definiere nun für $N \in \mathbb{N}$ den Wert $x_N := f^{-1}(y_{k_N}) \in K$. Somit ist $(x_N)_{N \in \mathbb{N}} \subseteq K$ eine Folge. Nach Proposition 10.3.16 hat $(x_N)_{N \in \mathbb{N}}$ eine konvergente Teilfolge $(x_{\ell_N})_{n \in \mathbb{N}}$ deren Grenzwert wieder in K liegt. Etwa sei $x_0 := \lim_{N \rightarrow \infty} x_{\ell_N}$. Da $x_0 \in K$ ist f bei x_0 stetig. Wir folgern

$$f(x_0) = \lim_{N \rightarrow \infty} f(x_{\ell_N}) = \lim_{N \rightarrow \infty} f(f^{-1}(y_{k_{\ell_N}})) = \lim_{N \rightarrow \infty} y_{k_{\ell_N}} \xrightarrow{\text{Prop 10.2.9}} \lim_{N \rightarrow \infty} y_N = y_0.$$

Also gilt $f(x_0) = y_0$ und somit (wegen der Bijektivität von f) auch $x_0 = f^{-1}(y_0)$. Es zeigt sich

$$f^{-1}(y_{k_{\ell_N}}) = x_{\ell_N} \xrightarrow{N \rightarrow \infty} f^{-1}(y_0).$$

Insbesondere gilt $\|f^{-1}(y_{\ell_{k_N}}) - f^{-1}(y_0)\| \xrightarrow{N \rightarrow \infty} 0$. Nach (555) gilt aber

$$\|f^{-1}(y_{\ell_{k_N}}) - f^{-1}(y_0)\| \geq \varepsilon_0 \quad \forall N \in \mathbb{N}.$$

Ein Widerspruch (denn nach den Grenzwertvergleichsregeln müsste der Grenzwert dann größer oder gleich ε_0 sein!)

Eine stetige bijektive Abbildung $f : D_1 \rightarrow D_2$, deren Umkehrabbildung wiederum stetig ist, nennt man übrigens einen *Homöomorphismus*.

Zur gleichmäßigen Stetigkeit. Hierzu definieren wir zunächst durch Analogie das Konzept der gleichmäßigen Stetigkeit im Mehrdimensionalen.

Definition 10.4.18 — GLEICHMÄSSIGE STETIGKEIT

Es sei $D \subseteq \mathbb{R}^d$. Eine Funktion $f : D \rightarrow \mathbb{R}^m$ heißt auf D *gleichmäßig stetig* falls

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x, y \in D \text{ mit } \|x - y\|_{\mathbb{R}^d} < \delta \text{ gilt } \|f(x) - f(y)\|_{\mathbb{R}^m} < \varepsilon.$$

Satz 6.4.2, der besagt, dass stetige Funktionen auf kompakten Intervallen gleichmäßig stetig sind, lässt sich tatsächlich ins Mehrdimensionale verallgemeinern — Das kompakte Intervall muss hierbei durch eine kompakte Teilmenge des $\mathbb{R}^d$.

Satz 10.4.19 — STETIGKEIT UND GLEICHMÄSSIGE STETIGKEIT

Es sei $K \subseteq \mathbb{R}^d$ eine kompakte Menge und $f : K \rightarrow \mathbb{R}^m$ eine stetige Funktion. Dann ist f auf K auch gleichmäßig stetig.

Beweis. Es handelt sich mal wieder um eine minimale Modifikation des Beweises von Satz 6.4.2. Details sind den Lesenden als Übungsaufgabe überlassen.

Zum Satz von Weierstraß. Wiederum fragen wir uns beim Satz von Weierstraß, unter welchen Umständen stetige Funktionen $f : D \rightarrow \mathbb{R}$ Maxima oder Minima annehmen. (Beachte: Dass die Funktion f nur in $\mathbb{R}^1$ abbildet ist kein Schreibfehler, denn sonst könnte man gar nicht von Maxima oder Minima sprechen!) Aufgrund unserer Fortschritte in der Topologie können wir jetzt aber noch etwas viel Allgemeineres zeigen:

Satz 10.4.20 — ERHALTUNG VON KOMPAKTHEIT UNTER STETIGEN ABBILDUNGEN

Es sei $K \subseteq \mathbb{R}^d$ eine kompakte Menge und $f : K \rightarrow \mathbb{R}^m$ eine stetige Funktion. Dann ist $f(K) \subseteq \mathbb{R}^m$ kompakt.

Beweis. Es sei $K \subseteq \mathbb{R}^d$ kompakt. Zu zeigen ist, dass $f(K)$ kompakt ist. Hierzu zeigen wir, dass $f(K)$ folgenkompakt ist (und folgern die Aussage dann mit Bemerkung 10.3.21). Es sei dann $(y_n)_{n \in \mathbb{N}} \subseteq f(K)$ eine Folge.

$\mathbb{Z}\mathbb{Z}$: Es gibt eine Teilfolge $(y_{\ell_n})_{n \in \mathbb{N}}$ und ein $y \in f(K)$ mit $y_{\ell_n} \xrightarrow{n \rightarrow \infty} y$. (vgl. Def. 10.3.18)

Dazu: Da $y_n \in f(K)$ gibt es ein $x_n \in K$ mit $y_n = f(x_n)$. Die Folge $(x_n)_{n \in \mathbb{N}} \subseteq K$ definiert nun eine Folge in K . Da K bereits folgenkompakt ist gilt

Es gibt eine Teilfolge $(x_{\ell_n})_{n \in \mathbb{N}}$ von $(x_n)_{n \in \mathbb{N}}$ und ein $x \in K$ mit $x_{\ell_n} \xrightarrow{n \rightarrow \infty} x$.

Da f stetig ist gilt, dann auch

$$f(x_{\ell_n}) \xrightarrow{n \rightarrow \infty} f(x).$$

Da für alle $n \in \mathbb{N}$ gilt $f(x_{\ell_n}) = y_{\ell_n}$ haben wir somit eine konvergente Teilfolge $(y_{\ell_n})_{n \in \mathbb{N}}$ mit Grenzwert $y := f(x) \in f(K)$ gefunden. Die Behauptung folgt mit Bemerkung 10.3.21.

Als Korollar ergibt sich nun der *mehrdimensionale Satz von Weierstraß*.

Proposition 10.4.21 — SATZ VON WEIERSTRASS — MEHRDIMENSIONALE VERSION

Es sei $K \subseteq \mathbb{R}^d$ kompakt und $f : K \rightarrow \mathbb{R}$ eine stetige Funktion^a. Dann nimmt f auf K Maximum und Minimum an, d.h. es gibt $x^*, x_* \in K$ mit

$$f(x^*) = \max f(K) \quad \text{und} \quad f(x_*) = \min f(K)$$

^aDas bedeutet insbesondere $m = 1$ — der Zielbereich muss reellwertig sein, sonst kann von Maxima und Minima nicht gesprochen werden.

Beweis. Da $K \subseteq \mathbb{R}^d$ kompakt ist ist auch $f(K) \subseteq \mathbb{R}$ kompakt, vgl. Satz 10.4.20. Das bedeutet $f(K) \subseteq \mathbb{R}$ ist abgeschlossen und beschränkt (vgl. Definition 10.3.15). Wegen der Beschränktheit von $f(K)$ gibt es $\sup f(K)$ und $\inf f(K)$. Nach Beispiel 10.3.3 sind $\sup f(K)$ und $\inf f(K)$ auch Berührungspunkte von $f(K)$. Da die Menge $f(K)$ abgeschlossen ist, enthält sie alle ihre Berührungspunkte. Das zeigt, dass

$$\sup f(K), \inf f(K) \in f(K).$$

Obere (bzw. untere) Schranken, die bereits in der Menge liegen, sind aber gerade Maxima (bzw. Minima), vgl. Definition 3.2.5 (iii). Wir erhalten die Existenz von $\max f(K)$ und $\min f(K)$.

Exkurs. Normen und deren Äquivalenz auf $\mathbb{R}^d$. Wir haben uns zu Beginn dieses Kapitels dazu entschieden, Abstände im $\mathbb{R}^d$ mit der *Euklidischen Norm*

$$\|x\|_{\mathbb{R}^d} := \sqrt{\sum_{i=1}^d (x^{(i)})^2}$$

zu messen. In diesem Kapitel sind wir in der Lage, die *Tragweite* dieser Entscheidung zu verstehen: Was wäre mit unserer Analysis passiert, wenn wir Abstände anders messen würden? Hierzu definieren wir zunächst abstrakt, was wir unter einem Abstandsbegriff verstehen wollen und untersuchen dann die Zusammenhänge zwischen den verschiedenen Begriffen.

Definition 10.4.22 — NORM

Eine Abbildung $N : \mathbb{R}^d \rightarrow \mathbb{R}$ heißt *Norm auf $\mathbb{R}^d$* falls folgende Bedingungen erfüllt sind

- (i) *Definitheit.* Für $x \in \mathbb{R}^d$ gilt $N(x) \geq 0$ und $(N(x) = 0 \iff x = 0)$
- (ii) *Homogenität.* Für $x \in \mathbb{R}^d$ und $\lambda \in \mathbb{R}$ gilt $N(\lambda x) = |\lambda|N(x)$.
- (iii) *Dreiecksungleichung.* Für $x, y \in \mathbb{R}^d$ gilt $N(x + y) \leq N(x) + N(y)$.

Es gibt mehrere Beispiele für Normen auf $\mathbb{R}^d$.

Beispiel 10.4.23

Definiere für $x = (x^{(1)}, \dots, x^{(d)})^T \in \mathbb{R}^d$

- (i) *Eins-Norm.* $N_1(x) := \sum_{i=1}^d |x^{(i)}|$
- (ii) *Zwei-Norm (bzw. Euklidische Norm).* $N_2(x) := \left(\sum_{i=1}^d |x^{(i)}|^2 \right)^{\frac{1}{2}}$
- (iii) *Unendlich-Norm.* $N_\infty(x) := \max_{i=1, \dots, d} |x^{(i)}|$.

Behauptung. N_1, N_2, N_∞ definieren allesamt Normen auf $\mathbb{R}^d$

Beweis. Für N_2 ist dies exakt in Satz 10.1.2 bewiesen worden. N_∞ überlassen wir den Lesenden als Übungsaufgabe. Also bleibt nur N_1 zu zeigen.

SCHRITT 1. Definitheit von N_1 .

Sei $x \in \mathbb{R}^d$. Dass $N_1(x) \geq 0$ gilt ist klar (da $|\cdot|$ stets nur positive Werte annimmt). Ist $N_1(x) = 0$ so muss jeder Summand in der Summe gleich Null sein (denn jeder Summand ist nichtnegativ). Wir folgern

$$|x^{(1)}| = \dots = |x^{(d)}| = 0 \quad \text{und somit} \quad x^{(1)} = \dots = x^{(d)} = 0.$$

Das bedeutet gerade $x = 0$. Die Definitheit ist gezeigt.

SCHRITT 2. Homogenität von N_1 .

Es sei $x \in \mathbb{R}^d$ und $\lambda \in \mathbb{R}$. Dann gilt

$$N_1(\lambda x) = \sum_{i=1}^d |\lambda x^{(i)}| = \sum_{i=1}^d |\lambda| |x^{(i)}| = |\lambda| \sum_{i=1}^d |x^{(i)}| = |\lambda| N_1(x).$$

SCHRITT 3. Dreiecksungleichung für N_1 .

Es seien $x = (x^{(1)}, \dots, x^{(d)})^T$ und $y = (y^{(1)}, \dots, y^{(d)})^T \in \mathbb{R}^d$. Dann gilt

$$N_1(x + y) = \sum_{i=1}^d |(x + y)^{(i)}| = \sum_{i=1}^d |x^{(i)} + y^{(i)}| \stackrel{(*)}{\leq} \sum_{i=1}^d |x^{(i)}| + |y^{(i)}| = \sum_{i=1}^d |x^{(i)}| + \sum_{i=1}^d |y^{(i)}|$$

wobei wir in $(*)$ die Dreiecksungleichung für den reellen Absolutbetrag verwendet haben. Nutzen wir auf der rechten Seite die Definition von N_1 so folgern wir

$$N_1(x + y) \leq N_1(x) + N_1(y).$$

15.06.2026

In vielen Büchern schreibt man für $x \in \mathbb{R}^d$ auch $\|x\|_1 := N_1(x)$, $\|x\|_2 := N_2(x)$ bzw. $\|x\|_\infty := N_\infty(x)$. Wir haben aber für $N_2(x)$ bereits in Definition 10.1.1 die Notation $\|x\|$ ohne Index eingeführt, daher werden wir diese Notation nicht verwenden.

Bemerkung 10.4.24

Die Dreiecksungleichung aus Definition 10.4.22(iii) kann mit einer einfachen Induktion auf eine beliebige Anzahl Summanden verallgemeinert werden. In der Tat: Sind $x_1, \dots, x_n$ Vektoren in $\mathbb{R}^d$, $n \in \mathbb{N}$, so gilt die *iterierte Dreiecksungleichung*

$$N\left(\sum_{k=1}^n x_k\right) \leq \sum_{k=1}^n N(x_k).$$

Normen haben viele Gemeinsamkeiten, zum Beispiel gilt die *sog. Inverse Dreiecksungleichung*.

Proposition 10.4.25 — INVERSE DREIECKSUNGLEICHUNG

Es sei $N : \mathbb{R}^d \rightarrow \mathbb{R}$ eine Norm. Dann gilt für alle $x, y \in \mathbb{R}^d$

$$|N(x) - N(y)| \leq N(x - y).$$

Beweis. Es seien $x, y \in \mathbb{R}^d$. Dann gilt

$$N(x) = N(x - y + y) \stackrel{\text{Def 10.4.22(iii)}}{\leq} N(x - y) + N(y).$$

Subtrahiert man $N(y)$ auf beiden Seiten so erhält man

$$N(x) - N(y) \leq N(x - y) \tag{556}$$

Dieselbe Rechnung mit vertauschten Rollen von x und y liefert

$$N(y) - N(x) \leq N(y - x). \quad (557)$$

Beachte: Wegen der Homogenität (Definition 10.4.22 (ii)) gilt

$$N(y - x) = N((-1)(x - y)) = |(-1)|N(x - y) = N(x - y).$$

Somit kann (557) umgeschrieben werden als

$$N(y) - N(x) \leq N(x - y). \quad (558)$$

Wir folgern

$$\begin{aligned} |N(x) - N(y)| &= \max\{N(x) - N(y), -(N(x) - N(y))\} \\ &= \max\{N(x) - N(y), N(y) - N(x)\} \underset{(556), (558)}{\leq} N(x - y). \end{aligned}$$

Eine weitere Gemeinsamkeit ist, dass die Normen alle stetig sind.

Proposition 10.4.26

Es sei $N : \mathbb{R}^d \rightarrow \mathbb{R}$ eine Norm. Dann ist N auf $\mathbb{R}^d$ stetig.

Beweis. Es sei $x_0 \in \mathbb{R}^d$.

$$\mathbb{Z} : \quad \lim_{x \rightarrow x_0} N(x) = N(x_0) \quad (559)$$

Hierzu zunächst eine hilfreiche Abschätzung: Für $x \in \mathbb{R}^d$ gilt

$$\begin{aligned} |N(x) - N(x_0)| &\underset{\text{Prop 10.4.25}}{\leq} N(x - x_0) = N\left(\sum_{i=1}^d (x - x_0)^{(i)} e_i\right) \\ &\underset{\text{Bem 10.4.24}}{\leq} \sum_{i=1}^d N((x - x_0)^{(i)} e_i) \underset{\text{Def 10.4.22(ii)}}{=} \sum_{i=1}^d |(x - x_0)^{(i)}| N(e_i) \\ &= \sum_{i=1}^d |x^{(i)} - x_0^{(i)}| N(e_i). \end{aligned} \quad (560)$$

Nun zum eigentlichen Nachweis von (559). Sei $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ eine Folge mit $\lim_{n \rightarrow \infty} x_n = x_0$.

Dann gilt für alle $i \in \{1, \dots, d\}$ auch $x_n^{(i)} \xrightarrow{n \rightarrow \infty} x_0^{(i)}$. Wegen der Stetigkeit der Betragsfunktion gilt dann auch

$$|x_n^{(i)} - x_0^{(i)}| \xrightarrow{n \rightarrow \infty} 0 \quad \forall i = 1, \dots, d. \quad (561)$$

Wegen (560) gilt

$$0 \leq |N(x_n) - N(x_0)| \stackrel{(560)}{\leq} \sum_{i=1}^d \underbrace{|x_n^{(i)} - x_0^{(i)}|}_{\rightarrow 0, \text{ vgl. (561)}} N(e_i) \xrightarrow{n \rightarrow \infty} \sum_{i=1}^d 0 \cdot N(e_i) = 0.$$

Nach dem Sandwichlemma gilt dann also

$$\lim_{n \rightarrow \infty} |N(x_n) - N(x_0)| = 0.$$

Nach Satz 10.2.6 (angewendet für $d = 1$) folgt

$$\lim_{n \rightarrow \infty} N(x_n) = N(x_0).$$

Da $(x_n)_{n \in \mathbb{N}}$ beliebig war (mit $\lim_{n \rightarrow \infty} x_n = x_0$) gilt

$$\lim_{x \rightarrow x_0} N(x) = N(x_0).$$

Wir werden nun aufklären, wie sich verschiedene Normen auf $\mathbb{R}^d$ voneinander unterscheiden können. Es stellt sich heraus, dass alle Normen Längen auf *vergleichbare* Weise messen. Das bedeutet gewissermaßen, dass das *Verhältnis der Normen* nach unten und nach oben beschränkt ist. Was damit genau gemeint ist formulieren wir im Folgenden Lemma.

Lemma 10.4.27

Es sei $N : \mathbb{R}^d \rightarrow \mathbb{R}$ eine Norm. Dann gibt es Konstanten $k, K > 0$ so, dass

$$\forall x \in \mathbb{R}^d : \quad k\|x\| \leq N(x) \leq K\|x\|. \quad (562)$$

Beweis. Für den Beweis definieren wir (zunächst ohne erkennbaren Grund)

$$\mathbb{S} := \{x \in \mathbb{R}^d : \|x\| = 1\} \quad (563)$$

Wir behaupten zuerst, dass $\mathbb{S}$ kompakt ist. In der Tat: $\mathbb{S}$ ist offensichtlich beschränkt da für jeden Vektor $x \in \mathbb{S}$ gilt $\|x\| \leq 1$. Desweiteren ist $\mathbb{S}$ abgeschlossen: Dazu: Sei $x_0 \in \mathbb{R}^d$ ein Berührungspunkt von $\mathbb{S}$, etwa sei $(x_n)_{n \in \mathbb{N}} \subseteq \mathbb{S}$ mit $\lim_{n \rightarrow \infty} x_n = x_0$. Zu zeigen ist, dass $x_0 \in \mathbb{S}$.

Wegen der Stetigkeit von N_2 (vgl. Proposition 10.4.26) gilt

$$\|x_0\| = \left\| \lim_{n \rightarrow \infty} x_n \right\| \stackrel{\text{Prop 10.4.26}}{=} \lim_{n \rightarrow \infty} \|x_n\| \stackrel{x_n \in \mathbb{S}, (563)}{=} \lim_{n \rightarrow \infty} 1 = 1.$$

Nun zum Beweis. Betrachte die Funktion $N|_{\mathbb{S}} : \mathbb{S} \rightarrow \mathbb{R}$. Diese ist stetig (weil N ja sogar auf der größeren Menge $\mathbb{R}^d$ stetig ist, vgl. Proposition 10.4.26). Nach Proposition 10.4.21 nimmt $N|_{\mathbb{S}}$ ein Maximum und ein Minimum an. Das bedeutet es gibt $x^* \in \mathbb{S}$ und $x_* \in \mathbb{S}$

so, dass

$$N(x_*) = \min_{x \in \mathbb{S}} N(x) \quad \text{und} \quad N(x^*) = \max_{x \in \mathbb{S}} N(x).$$

Mit anderen Worten gilt

$$N(x_*) \leq N(x) \leq N(x^*) \quad \forall x \in \mathbb{S}. \quad (564)$$

Beachte: Wegen der Definitheit von N (vgl. Definition 10.4.22 (i)) gilt $N(x_*), N(x^*) > 0$ (da $x_*, x^* \neq 0$ weil $0 \notin \mathbb{S}$). Definiere nun

$$k := N(x_*) \text{ und } K := N(x^*).$$

Dann gilt insbesondere $k, K > 0$. Mit diesen Konstanten zeigen wir (562). Sei dazu $x \in \mathbb{R}^d$ beliebig. Ist $x = 0$, so ist nichts zu zeigen.^a Für $x \neq 0$ gilt $\frac{x}{\|x\|} \in \mathbb{S}$, da $\|\frac{x}{\|x\|}\| \stackrel{\text{Def 10.4.22(ii)}}{=} \frac{1}{\|x\|} \|x\| = 1$. Nach (564) gilt dann

$$k \leq N\left(\frac{x}{\|x\|}\right) \leq K.$$

Multiplizieren wir mit $\|x\|$ so ergibt sich

$$k\|x\| \leq N\left(\frac{x}{\|x\|}\right) \|x\| \leq K\|x\|.$$

Wegen der Homogenität der Norm (vgl. Definition 10.4.22 (ii)) gilt $N\left(\frac{x}{\|x\|}\right) \|x\| = N\left(\|x\| \frac{x}{\|x\|}\right) = N(x)$. Zusammen mit der vorigen Gleichung erhalten wir

$$k\|x\| \leq N(x) \leq K\|x\|.$$

^aDenn $\|0\| = N(0) = 0$ (vgl. Definition 10.4.22 (i)) und somit ist (562) für $x = 0$ klarerweise erfüllt.

Wir haben also gesehen, dass sich die Messung von Abständen bezüglich einer Norm N nicht zu sehr von der Abstandsmessung mit der Euklidischen Norm unterscheidet. Auch die Resultate, in denen wir in den vorigen Abschnitten die Euklidische Norm benutzt haben, können nun auf beliebige Normen übertragen werden. Ein Beispiel dafür bildet der Satz über *Konvergenz und Norm* (Satz 10.2.6)

Beispiel 10.4.28 — ALLE NORMEN AUF $\mathbb{R}^d$ KÖNNEN KONVERGENZ ÜBERPRÜFEN

Es sei $N : \mathbb{R}^d \rightarrow \mathbb{R}$ eine beliebige Norm. Für eine Folge $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ sind äquivalent

$$a_n \xrightarrow{n \rightarrow \infty} a \quad \Longleftrightarrow \quad N(a_n - a) \xrightarrow{n \rightarrow \infty} 0.$$

Beweis. Die Richtung „ $\implies$ “ folgt direkt aus der Stetigkeit der Norm, vgl. Proposition 10.4.26. Für „ $\impliedby$ “ sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ so, dass $N(a_n - a) \xrightarrow{n \rightarrow \infty} 0$. Ferner seien $k, K > 0$ wie

in Lemma 10.4.27. Dann gilt

$$0 \leq k\|a_n - a\| \underset{\text{Lem10.4.27}}{\leq} N(a_n - a) \xrightarrow{n \rightarrow \infty} 0 \quad \text{d.h.} \quad k\|a_n - a\| \xrightarrow{n \rightarrow \infty} 0.$$

Nach den Grenzwertregeln gilt dann auch

$$\frac{1}{k} (k\|a_n - a\|) \xrightarrow{n \rightarrow \infty} 0.$$

Also gilt $\|a_n - a\| \xrightarrow{n \rightarrow \infty} 0$ und somit nach Satz 10.2.6 auch $a_n \xrightarrow{n \rightarrow \infty} a$.

Konvergenz, und damit auch unsere gesamte Theorie der *Topologie* könnte daher auch mit einer beliebigen Norm N auf $\mathbb{R}^d$ entwickelt werden. Somit hätte es für unsere Theorie keinen Unterschied gemacht, ob wir die Euklidische Norm oder eine beliebige Norm verwendet hätten. All das liegt daran, dass jede beliebige Norm mit der Euklidischen Norm *vergleichbar* ist, im Sinne von (562).

Dieses Konzept der Vergleichbarkeit wird *Äquivalenz von Normen auf $\mathbb{R}^d$* genannt. Was wir damit meinen wollen wir im nachfolgenden Satz noch einmal zusammenfassen

Satz 10.4.29 — NORMEN AUF $\mathbb{R}^d$ SIND ÄQUIVALENT

Es seien $N, \tilde{N} : \mathbb{R}^d \rightarrow \mathbb{R}$ zwei Normen auf $\mathbb{R}^d$. Dann sind N und $\tilde{N}$ *äquivalent*, das bedeutet es gibt $c, C > 0$ so, dass

$$\forall x \in \mathbb{R}^d : \quad cN(x) \leq \tilde{N}(x) \leq CN(x).$$

Beweis. Nach Lemma 10.4.27 gibt es Konstanten $k_1, K_1 > 0$ sowie $k_2, K_2 > 0$ so, dass

$$k_1\|x\| \leq N(x) \leq K_1\|x\|$$

und

$$k_2\|x\| \leq \tilde{N}(x) \leq K_2\|x\|. \quad (565)$$

Nun gilt

$$\tilde{N}(x) \underset{(565)}{\leq} K_2\|x\| = \frac{K_2}{k_1} k_1\|x\| \leq \frac{K_2}{k_1} N(x)$$

und

$$\tilde{N}(x) \underset{(565)}{\geq} k_2\|x\| = \frac{k_2}{K_1} K_1\|x\| \geq \frac{k_2}{K_1} N(x).$$

Definiert man $c := \frac{k_2}{K_1} > 0$ und $C := \frac{K_2}{k_1} > 0$ so erhält man mit den beiden vorigen Gleichungen

$$cN(x) \leq \tilde{N}(x) \leq CN(x),$$

was zu zeigen war.

BONUSKAPITEL. p-Normen auf $\mathbb{R}^d$. In diesem zusätzlichen Kapitel wollen wir weitere Normen auf $\mathbb{R}^d$ kennen lernen — die sogenannten *p-Normen*. Diese sind für $p \in [1, \infty)$ definiert durch $N_p : \mathbb{R}^d \rightarrow \mathbb{R}$

$$N_p(x) := \left(\sum_{i=1}^d |x^{(i)}|^p \right)^{\frac{1}{p}}.$$

Wir werden in diesem Kapitel zeigen, dass es sich bei diesen Abbildungen tatsächlich um Normen handelt. Für $p = 1$ ergibt sich z.B. die Einsnorm aus Beispiel 11.3.2. Für $p = 2$ ergibt sich die Euklidische Norm $N_2(x) = \|x\|$. Somit ist unsere Notation auch konsistent mit Beispiel 11.3.2. Man könnte sogar zeigen, dass für alle $x \in \mathbb{R}^d$ gilt

$$\lim_{p \rightarrow \infty} N_p(x) = \max_{i=1, \dots, d} |x^{(i)}| = N_\infty(x) \quad (\text{vgl. Beispiel 11.3.2}).$$

Das rechtfertigt dann also auch die Notation N_∞ aus Beispiel 11.3.2. Wir wollen in diesem Abschnitt beweisen:

Proposition 10.4.30

Für alle $p \in [1, \infty)$ definiert die Abbildung $N_p : \mathbb{R}^d \rightarrow \mathbb{R}$, $N_p(x) := \left(\sum_{i=1}^d |x^{(i)}|^p \right)^{\frac{1}{p}}$ eine Norm auf $\mathbb{R}^d$ (im Sinne von Definition 10.4.22).

Für $p = 1$ haben wir diese Behauptung bereits in Beispiel 11.3.2 gezeigt. Daher beschränken sich unsere nachfolgenden Ausführungen auf den Fall $p \in (1, \infty)$. Definitheit und Homogenität werden leicht zu zeigen sein. Nur bei der Dreiecksungleichung sind weitere Vorbereitungen erforderlich.

Hierzu benötigen wir zunächst einige Vorbereitungen, zum Beispiel die *Young'sche Ungleichung*.

Lemma 10.4.31 — YOUNG'SCHE UNGLEICHUNG

Es seien $p, q \in (1, \infty)$ so, dass $\frac{1}{p} + \frac{1}{q} = 1$ ^a. Dann gilt für nichtnegative reelle Zahlen $x, y \in [0, \infty)$ stets

$$xy \leq \frac{1}{p}x^p + \frac{1}{q}y^q.$$

^aKonkret bedeutet das $q = \frac{p}{p-1}$. Denn dann ist $\frac{1}{p} + \frac{1}{q} = \frac{1}{p} + \frac{p-1}{p} = \frac{p}{p} = 1$.

Beweis. Falls eine der beiden Zahlen x, y gleich 0 ist, so ist nichts zu zeigen (denn die linke Seite ist dann bereits gleich 0 – kleiner als das geht's nicht!) Also dürfen wir annehmen $x, y \in (0, \infty)$.

ZWISCHENBEHAUPTUNG. $\log$ ist auf $(0, \infty)$ konkav.

Dazu: Für $x \in (0, \infty)$ gilt

$$\log''(x) = \log''(x) = \frac{d^2}{dx^2} \log(x) = \frac{d}{dx} \left(\frac{1}{x} \right) = \left(-\frac{1}{x^2} \right) < 0.$$

Nach Proposition 7.5.6 ist $\log$ dann konkav

ZWISCHENBEHAUPTUNG ENDE.

Nutze nun die Definition der Konkavität mit $\lambda = \frac{1}{p}$ aus und folgere

$$\begin{aligned} \log \left(\frac{1}{p} x^p + \frac{1}{q} y^q \right) &= \log \left(\frac{1}{p} x^p + \left(1 - \frac{1}{p} \right) y^q \right) \stackrel{\text{Def 7.5.1}}{\geq} \frac{1}{p} \log(x^p) + \left(1 - \frac{1}{p} \right) \log(y^q) \\ &\stackrel{\text{Bem 6.3.9}}{=} \frac{1}{p} \cdot p \log(x) + \left(1 - \frac{1}{p} \right) \cdot q \log(y) = \frac{1}{p} \cdot p \log(x) + \frac{1}{q} \cdot q \log(y) \\ &= \log(x) + \log(y) \stackrel{\text{Prop 6.3.4}}{=} \log(xy). \end{aligned}$$

Nun können wir auf beiden Seiten dieser Ungleichung die (auf $\mathbb{R}$ streng monoton wachsende) Exponentialfunktion anwenden und erhalten

$$\frac{1}{p} x^p + \frac{1}{q} y^q \geq xy,$$

was zu zeigen war.

Eine Folgerung aus der Young'schen Ungleichung ist die *Hölder'sche Ungleichung*.

Lemma 10.4.32 — HÖLDER'SCHE UNGLEICHUNG

Es sei $n \in \mathbb{N}$ und $x_1, \dots, x_n, y_1, \dots, y_n \in \mathbb{R}$. Ferner seien $p, q \in (1, \infty)$ mit $\frac{1}{p} + \frac{1}{q} = 1$. Dann gilt

$$\left| \sum_{k=1}^n x_k y_k \right| \leq \left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^n |y_k|^q \right)^{\frac{1}{q}}.$$

Wir bemerken: Im Falle $p = 2$ (also auch $q = 2$, da $\frac{1}{2} + \frac{1}{2} = 1$) ergibt sich gerade die *Cauchy-Schwarz'sche Ungleichung* aus Lemma 10.1.3, welche wir bereits für die Dreiecksungleichung in der Euklidischen Norm verwendet haben.

Beweis von Lemma 10.4.32. Zunächst verwenden wir die iterierte Dreiecksungleichung und schätzen ab

$$\left| \sum_{k=1}^n x_k y_k \right| \leq \sum_{k=1}^n |x_k y_k| = \sum_{k=1}^n |x_k| |y_k|.$$

Es genügt also zu zeigen, dass

$$\sum_{k=1}^n |x_k| |y_k| \leq \left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^n |y_k|^q \right)^{\frac{1}{q}}. \quad (566)$$

Ist einer der Faktoren $\left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}}$ oder $\left(\sum_{k=1}^n |y_k|^q\right)^{\frac{1}{q}}$ gleich 0 so ist nichts zu zeigen ^a.

Wir dürfen also OBdA annehmen, dass $\left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}}, \left(\sum_{k=1}^n |y_k|^q\right)^{\frac{1}{q}} > 0$. Definiere nun

$A := \left(\sum_{k=1}^n |x_k|^p\right)^{\frac{1}{p}}$ und $B := \left(\sum_{k=1}^n |y_k|^q\right)^{\frac{1}{q}}$. Wir werden nun den Term

$$\sum_{k=1}^n \frac{|x_k|}{A} \frac{|y_k|}{B}$$

mithilfe der Young'schen Ungleichung (Lemma 10.4.31) abschätzen. Dazu wenden wir sie in jedem Summanden an und erhalten

$$\sum_{k=1}^n \frac{|x_k|}{A} \frac{|y_k|}{B} \leq \sum_{k=1}^n \left(\frac{1}{p} \left(\frac{|x_k|}{A} \right)^p + \frac{1}{q} \left(\frac{|y_k|}{B} \right)^q \right) = \frac{1}{p} \frac{1}{A^p} \sum_{k=1}^n |x_k|^p + \frac{1}{q} \frac{1}{B^q} \sum_{k=1}^n |y_k|^q. \quad (567)$$

Nutzen wir nun, dass

$$A^p = \left(\left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} \right)^p = \sum_{k=1}^n |x_k|^p \quad \text{und analog} \quad B^q = \sum_{k=1}^n |y_k|^q,$$

so können wir beiden Summanden von (567) die Summen im Zähler durch die Nenner rauskürzen. Wir erhalten

$$\sum_{k=1}^n \frac{|x_k|}{A} \frac{|y_k|}{B} \leq \frac{1}{p} + \frac{1}{q} = 1.$$

Jetzt können wir noch A und B als konstante Faktoren vor die Summe schreiben und erhalten

$$\frac{1}{AB} \sum_{k=1}^n |x_k| |y_k| \leq 1. \quad \text{Es folgt:} \quad \sum_{k=1}^n |x_k| |y_k| \leq AB.$$

Nutzen wir die Definition von A und B so haben wir gezeigt, dass

$$\sum_{k=1}^n |x_k| |y_k| \leq \left(\sum_{k=1}^n |x_k|^p \right)^{\frac{1}{p}} \left(\sum_{k=1}^n |y_k|^q \right)^{\frac{1}{q}}.$$

Also folgt (566) und somit auch die Behauptung.

^aDenn dann gilt entweder $x_1 = \dots = x_n = 0$ oder $y_1 = \dots = y_n = 0$. Wenn das so ist, ist aber auch die linke Seite der Gleichung $\sum_{k=1}^n x_k y_k = 0$, da jeder Summand verschwindet

Nun sind wir endlich bereit, Proposition 10.4.30 zu beweisen.

Beweis von Proposition 10.4.30. Für $p = 1$ folgt die Behauptung bereits aus Beispiel 11.3.2. Daher beschränken wir uns wie angekündigt auf den Fall $p \in (1, \infty)$. Wir überprüfen die in Definition 10.4.22 geforderten Eigenschaften.

SCHRITT 1. Definitheit.

Diese zu überprüfen ist unspektakulär und den Lesenden an dieser Stelle als Übungsaufgabe überlassen.

SCHRITT 2. Homogenität.

Es sei $x = (x^{(1)}, \dots, x^{(d)})^T \in \mathbb{R}^d$ und $\lambda \in \mathbb{R}$, d.h. auch $\lambda x = (\lambda x^{(1)}, \dots, \lambda x^{(d)})^T$. Dann gilt

$$\begin{aligned} N_p(\lambda x) &= \left(\sum_{i=1}^d |\lambda x^{(i)}|^p \right)^{\frac{1}{p}} = \left(\sum_{i=1}^d |\lambda|^p |x^{(i)}|^p \right)^{\frac{1}{p}} = \left(|\lambda|^p \sum_{i=1}^d |x^{(i)}|^p \right)^{\frac{1}{p}} \\ &= (|\lambda|^p)^{\frac{1}{p}} \left(\sum_{i=1}^d |x^{(i)}|^p \right)^{\frac{1}{p}} = |\lambda| \left(\sum_{i=1}^d |x^{(i)}|^p \right)^{\frac{1}{p}} = |\lambda| N_p(x). \end{aligned}$$

Die Behauptung ist gezeigt.

SCHRITT 3. Dreiecksungleichung.

Es seien $x = (x^{(1)}, \dots, x^{(d)})^T, y = (y^{(1)}, \dots, y^{(d)})^T \in \mathbb{R}^d$.

$$\mathbb{Z}: \quad N_p(x + y) \leq N_p(x) + N_p(y).$$

ObdA ist $N_p(x + y) > 0$, denn andernfalls ist nichts zu zeigen weil die linke Seite dann bereits den kleinstmöglichen Wert 0 annimmt. Definiere nun $q := \frac{p}{p-1} \in (1, \infty)$. Dann gilt (wie bereits bemerkt) $\frac{1}{p} + \frac{1}{q} = \frac{1}{p} + \frac{p-1}{p} = \frac{p}{p} = 1$. Wir berechnen zunächst

$$\begin{aligned} N_p(x + y)^p &= \sum_{i=1}^d |x^{(i)} + y^{(i)}|^p = \sum_{i=1}^d |x^{(i)} + y^{(i)}| |x^{(i)} + y^{(i)}|^{p-1} \\ &\stackrel{\Delta\text{-Ungl}}{\leq} \sum_{i=1}^d (|x^{(i)}| + |y^{(i)}|) |x^{(i)} + y^{(i)}|^{p-1} \\ &= \sum_{i=1}^d |x^{(i)}| |x^{(i)} + y^{(i)}|^{p-1} + \sum_{i=1}^d |y^{(i)}| |x^{(i)} + y^{(i)}|^{p-1} \end{aligned} \quad (568)$$

Nun wenden wir im ersten Summanden die Hölder'sche Ungleichung (Lemma 10.4.32) an und schätzen ab

$$\sum_{i=1}^d |x^{(i)}| |x^{(i)} + y^{(i)}|^{p-1} \leq \left(\sum_{i=1}^d |x^{(i)}|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^d (|x^{(i)} + y^{(i)}|^{p-1})^q \right)^{\frac{1}{q}} \quad (569)$$

Nutzen wir, dass $(p-1) \cdot q = (p-1) \frac{p}{p-1} = p$ so erhalten wir, dass

$$\begin{aligned} \left(\sum_{i=1}^d (|x^{(i)} + y^{(i)}|^{p-1})^q \right)^{\frac{1}{q}} &= \left(\sum_{i=1}^d |x^{(i)} + y^{(i)}|^{(p-1)q} \right)^{\frac{1}{q}} = \left(\sum_{i=1}^d |x^{(i)} + y^{(i)}|^p \right)^{\frac{1}{q}} \\ &= \left(\sum_{i=1}^d |x^{(i)} + y^{(i)}|^p \right)^{1 - \frac{1}{p}} = \left(\sum_{i=1}^d |x^{(i)} + y^{(i)}|^p \right)^{\frac{p-1}{p}} \\ &= \left(\left(\sum_{i=1}^d |x^{(i)} + y^{(i)}|^p \right)^{\frac{1}{p}} \right)^{p-1} = N_p(x+y)^{p-1}. \end{aligned}$$

Die Ungleichung (569) zeigt also

$$\sum_{i=1}^d |x^{(i)}| |x^{(i)} + y^{(i)}|^{p-1} \leq \left(\sum_{i=1}^d |x^{(i)}|^p \right)^{\frac{1}{p}} N_p(x+y)^{p-1} = N_p(x) N_p(x+y)^{p-1}$$

Wir können für den zweiten Summanden in (568) analog zeigen, dass

$$\sum_{i=1}^d |y^{(i)}| |x^{(i)} + y^{(i)}|^{p-1} \leq N_p(y) N_p(x+y)^{p-1}$$

Mit der Ungleichungskette in (568) folgern wir

$$N_p(x+y)^p \leq N_p(x) N_p(x+y)^{p-1} + N_p(y) N_p(x+y)^{p-1} = N_p(x+y)^{p-1} (N_p(x) + N_p(y)).$$

Da wir angenommen haben, dass $N_p(x+y) > 0$ können wir durch $N_p(x+y)^{p-1}$ teilen und erhalten

$$N_p(x+y) \leq N_p(x) + N_p(y),$$

was zu zeigen war.

10.5 Mehrdimensionales Ableiten

In diesem Abschnitt kommen wir zur *mehrdimensionalen Differentialrechnung*, welche uns im Wesentlichen bis zum Ende dieser Vorlesung beschäftigen wird. Wir haben in Kapitel (7) die Ableitung als *Tangentensteigung* bei einem Punkt x_0 kennengelernt. Bei mehrdimensionalen Funktionsgraphen $f : \mathbb{R}^d \rightarrow \mathbb{R}$ kann man nun unter Umständen *mehrere Tangenten* anlegen — denn man kann bei einem Punkt $x_0 \in \mathbb{R}^d$ *in unendlich viele Richtungen schauen* und sich fragen, was für eine Tangentensteigung in dieser Richtung zu beobachten ist.

Mathematisch bedeutet das Folgendes: Für einen Richtungsvektor $v \in \mathbb{R}^d \setminus \{0\}$ kann man die sogenannte *v-Sektorfunktion von f bei x_0*

$$t \mapsto f(x_0 + tv) \quad (t \in \mathbb{R})$$

betrachten. Diese Funktion hängt dann nur von einem reellen Parameter t ab. Ist sie bei

$t = 0$ differenzierbar, so können wir dort die Ableitung bilden. Diese Ableitung entspricht gewissermaßen der *Tangentensteigung von f in Richtung v* . Tun wir das für jedes v , so erhalten wir die sogenannten *Richtungsableitungen*. Ein wichtiger Spezialfall der Richtungsableitungen sind die *partiellen Ableitungen*. Dies sind Richtungsableitungen in Richtung der *kanonischen Einheitsvektoren* $e_1, \dots, e_d$.

Es wird sich allerdings herausstellen, dass dieses Konzept der partiellen Ableitungen einige *Unzulänglichkeiten* besitzt, weswegen es für uns kein zufriedenstellendes Konzept der Ableitung ist. Ein Beispiel dafür: Im Eindimensionalen impliziert die Differenzierbarkeit die Stetigkeit (vgl Proposition 7.0.7). Ein solches Resultat hätte man gerne auch im Mehrdimensionalen. Allerdings verhält es sich so: Selbst wenn in einem Punkt x_0 alle partiellen Ableitungen existieren, ist die Funktion nicht zwingend stetig! Das hat damit zu tun, dass es für die Feststellung der Stetigkeit nicht ausreicht, nur einzelne Koordinatenrichtungen zu betrachten, vgl. Beispiel 10.4.9.

Diese Unannehmlichkeit wollen wir später beseitigen, indem wir ein weiteres, stärkeres Konzept der Ableitung einführen, die sogenannte *totale Ableitung*.

Richtungsableitungen

Definition 10.5.1 — SEKTORFUNKTIONEN

Es sei $D \subseteq \mathbb{R}^d$ und $f : D \rightarrow \mathbb{R}$ eine Funktion. Ferner sei $x_0 \in D$ und $v \in \mathbb{R}^d \setminus \{0\}$. Definiere $D_v := \{t \in \mathbb{R} : x_0 + tv \in D\} \subseteq \mathbb{R}$. Dann ist die *v -Sektorfunktion* $\gamma_v : D_v \rightarrow \mathbb{R}$ von f bei x_0 gegeben durch

$$\gamma_v(t) := f(x_0 + tv).$$

Ist $D = \mathbb{R}^d$, so ist stets $D_v = \mathbb{R}$. Ansonsten ist allerdings im Allgemeinen nicht klar, wie groß der Definitionsbereich D_v tatsächlich ist. Es könnte stark von v abhängen.

Bemerkung 10.5.2

Es sei $D \subseteq \mathbb{R}^d$ und $x_0 \in D$ ein innerer Punkt von D . Dann ist für $v \in \mathbb{R}^d \setminus \{0\}$ auch der Punkt 0 ein innerer Punkt von $D_v \subseteq \mathbb{R}$ (wobei D_v gegeben ist wie in Definition 10.5.1).

In der Tat: Wenn x_0 innerer Punkt von D ist, gibt es ein $r > 0$ mit $B_r(x_0) \subseteq D$.

$$\text{Behauptung.} \quad \left(-\frac{r}{\|v\|}, \frac{r}{\|v\|}\right) \subseteq D_v.$$

Dazu sei $t \in \left(-\frac{r}{\|v\|}, \frac{r}{\|v\|}\right)$, d.h. $|t| < \frac{r}{\|v\|}$. Wir müssen zeigen, dass $t \in D_v$, d.h. $x_0 + tv \in D$. Hierzu beobachten wir

$$\|(x_0 + tv) - x_0\| = \|tv\| = |t| \|v\| < \frac{r}{\|v\|} \|v\| = r.$$

Dies zeigt, dass $x_0 + tv \in B_r(x_0)$. Da $B_r(x_0) \subseteq D$ folgt auch $x_0 + tv \in D$. Wir haben somit gesehen, dass $\left(-\frac{r}{\|v\|}, \frac{r}{\|v\|}\right) = B_{\frac{r}{\|v\|}}^1(0) \subseteq D_v$. Das heißt, dass 0 ein innerer Punkt von D_v ist.

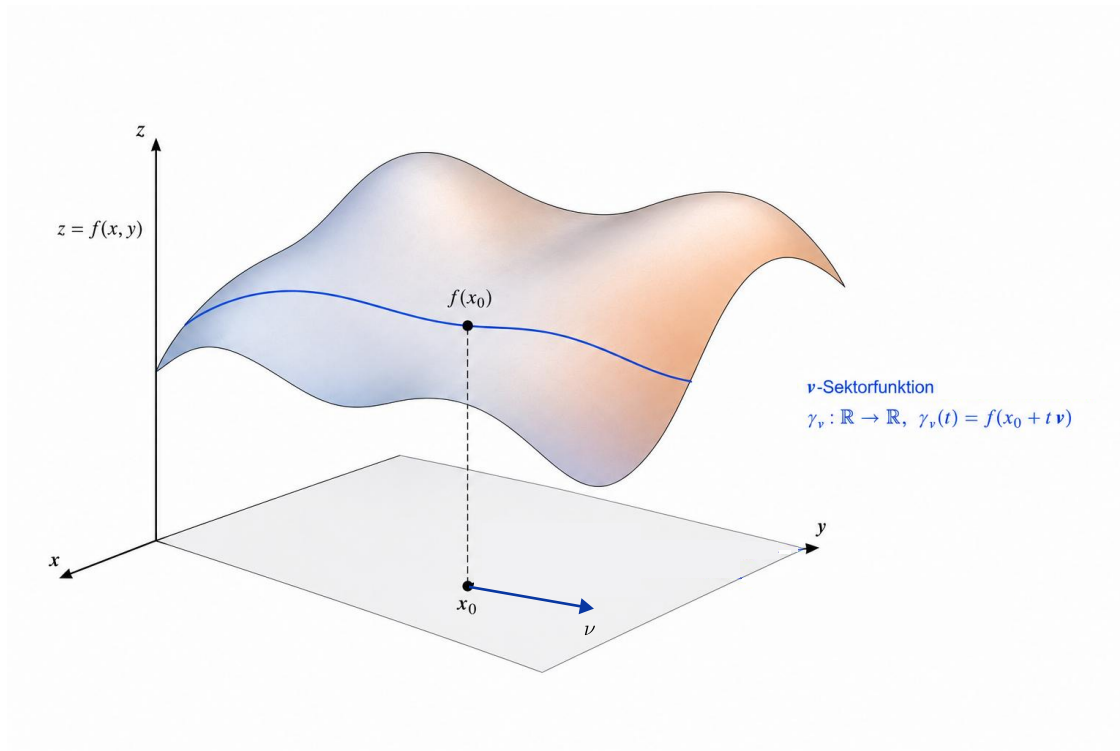


Abbildung 45: Ein Graph einer Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit Sektorfunktion.

Nun können wir die Richtungsableitung einführen.

Definition 10.5.3 — RICHTUNGSABLEITUNG

Es sei $D \subseteq \mathbb{R}^d$.

- (i) Sei $f : D \rightarrow \mathbb{R}$ eine Funktion, $x_0 \in D$ und $v \in \mathbb{R}^d \setminus \{0\}$. Wir sagen f ist bei x_0 in Richtung v *richtungsdifferenzierbar*, wenn die (eindimensionale) v -Sektorfunktion $\gamma_v : D_v \rightarrow \mathbb{R}, t \mapsto f(x_0 + tv)$ im Punkt $0 \in D_v$ differenzierbar ist. In dem Fall nennen wir

$$\partial_v f(x_0) := \gamma'_v(0) \quad (570)$$

die *Richtungsableitung* von f bei x_0 .

- (ii) Eine Funktion $f : D \rightarrow \mathbb{R}^m$ heißt bei $x_0 \in D$ in Richtung $v \in \mathbb{R}^d \setminus \{0\}$ *richtungsdifferenzierbar*, wenn alle Komponentenfunktionen $f^{(1)}, \dots, f^{(m)}$ in Richtung v richtungsdifferenzierbar ist. Die Richtungsableitung ist dann gegeben durch

$$\partial_v f(x_0) := \begin{pmatrix} \partial_v f^{(1)}(x_0) \\ \vdots \\ \partial_v f^{(m)}(x_0) \end{pmatrix} \in \mathbb{R}^m \quad (571)$$

(iii) f heißt bei $x_0 \in D$ *richtungsdifferenzierbar*, falls f in alle Richtungen $v \in \mathbb{R}^d \setminus \{0\}$ differenzierbar ist.

Da man die Funktionen immer komponentenweise betrachten kann, werden wir in den nachfolgenden Ausführungen hauptsächlich den Fall $m = 1$ untersuchen.

Bemerkung 10.5.4

Es sei $D \subseteq \mathbb{R}^d$, $x_0 \in D$ und $f : D \rightarrow \mathbb{R}$ eine gegebene Funktion mit v -Sektorfunktion $\gamma_v : D_v \rightarrow \mathbb{R}, t \mapsto f(x_0 + tv)$ bei x_0 . Damit γ_v bei 0 differenzierbar sein kann muss (gemäß Definition 7.0.1) der Punkt 0 ein Berührungspunkt^a von $D_v \setminus \{0\}$ sein. Dies ist zum Beispiel der Fall wenn x_0 ein innerer Punkt von D ist (wie man leicht aus Bemerkung 10.5.2 folgert).

FAZIT. Man muss sich keine Sorgen um die Berührungspunkte von $D_v \setminus \{0\}$ machen, wenn x_0 innerer Punkt von D ist. Am einfachsten ist das Differenzieren daher wenn einfach *jeder Punkt* ein innerer Punkt von D ist. Sprich: Wenn D offen ist (vgl. Definition 10.3.5)! Wir werden die Offenheit von D daher ab jetzt aus *Faulheitsgründen* oft annehmen.

^aIch erinnere nochmal daran dass in der Analysis II Häufungspunkte in Berührungspunkte umbenannt wurden!

Bemerkung 10.5.5

Es sei $D \subseteq \mathbb{R}^d$ offen, $x_0 \in D$ und $f : D \rightarrow \mathbb{R}$ eine Funktion. Für $v \in \mathbb{R}^d \setminus \{0\}$ bilden wir wieder die v -Sektorfunktion $\gamma_v : D_v \rightarrow \mathbb{R}, t \mapsto f(x_0 + tv)$. Schreibt man (570) als Grenzwert aus so erhält man

$$\partial_v f(x_0) = \gamma'_v(0) = \lim_{t \rightarrow 0} \frac{\gamma_v(t) - \gamma_v(0)}{t - 0} = \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t}.$$

Merkregel. Für die Richtungsableitung nach $v \in \mathbb{R}^d \setminus \{0\}$ erhält man die folgende explizite Formel

$$\partial_v f(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t}. \quad (572)$$

Betrachtet man jede Komponente in (571) einzeln so kann man zeigen, dass die Formel (572) auf für Funktionen $f : D \rightarrow \mathbb{R}^m$ Gültigkeit hat.

Beispiel 10.5.6

Es sei $j \in \mathbb{N}, j \geq 2$.

Behauptung. Die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch

$$f(x, y) = \begin{cases} \frac{x^j y}{x^2 + y^2} & (x, y) \neq (0, 0)^T \\ 0 & \text{sonst} \end{cases} \quad (573)$$

ist bei $x_0 = (0, 0)^T$ in jede Richtung $v = (v^{(1)}, v^{(2)})^T \in \mathbb{R}^2 \setminus \{0\}$ richtungsdifferenzierbar und es gilt

$$\partial_v f(0, 0) = \begin{cases} \frac{(v^{(1)})^2 v^{(2)}}{(v^{(1)})^2 + (v^{(2)})^2} & j = 2 \\ 0 & j \geq 3 \end{cases}$$

Beweis. Es sei $v = (v^{(1)}, v^{(2)})^T \in \mathbb{R}^2 \setminus \{0\}$. Wir bilden die v -Sektorfunktion $\gamma_v : \mathbb{R} \rightarrow \mathbb{R}$

$$\gamma_v(t) = f(x_0 + tv) = f(0 + tv^{(1)}, 0 + tv^{(2)}) = f(tv^{(1)}, tv^{(2)}).$$

Da $(v^{(1)}, v^{(2)})^T \neq 0$ gilt für $t \in \mathbb{R} \setminus \{0\}$, dass $(tv^{(1)}, tv^{(2)})^T \neq 0$. Somit gilt nach (573), dass

$$\gamma_v(t) = f(tv^{(1)}, tv^{(2)}) = \frac{(tv^{(1)})^j (tv^{(2)})}{(tv^{(1)})^2 + (tv^{(2)})^2} = \frac{t^{j+1}}{t^2} \frac{(v^{(1)})^j v^{(2)}}{(v^{(1)})^2 + (v^{(2)})^2} = t^{j-1} \frac{(v^{(1)})^j v^{(2)}}{(v^{(1)})^2 + (v^{(2)})^2}.$$

Für $t = 0$ gilt $\gamma_v(0) = f(x_0) = f(0, 0) = 0$. Da $j \geq 2$ ist $0^{j-1} = 0$ und somit gilt

$$\gamma_v(t) = t^{j-1} \frac{(v^{(1)})^2 v^{(2)}}{(v^{(1)})^2 + (v^{(2)})^2} \quad \text{sogar für alle } t \in \mathbb{R}.$$

Für $j = 2$ ist $j - 1 = 1$ und es folgt

$$\gamma'_v(0) = \frac{d}{dt} \left(t \frac{(v^{(1)})^2 v^{(2)}}{(v^{(1)})^2 + (v^{(2)})^2} \right) \Big|_{t=0} = \frac{(v^{(1)})^2 v^{(2)}}{(v^{(1)})^2 + (v^{(2)})^2}$$

Für $j \geq 3$ ist

$$\gamma'_v(0) = \frac{d}{dt} \left(t^{j-1} \frac{(v^{(1)})^j v^{(2)}}{(v^{(1)})^2 + (v^{(2)})^2} \right) \Big|_{t=0} = \left((j-1) t^{j-2} \frac{(v^{(1)})^j v^{(2)}}{(v^{(1)})^2 + (v^{(2)})^2} \right) \Big|_{t=0} = 0.$$

Beispiel 10.5.7

Es sei $f : \mathbb{R}^d \rightarrow \mathbb{R}$ gegeben durch $f(x) = x^{(i)}$ für ein festes $i \in \{1, \dots, d\}$. Ferner sei $x_0 \in \mathbb{R}^d$ beliebig.

Behauptung. f ist bei x_0 richtungsdifferenzierbar und für $v \in \mathbb{R}^d \setminus \{0\}$ gilt $\partial_v f(x_0) = v^{(i)}$.

Beweis. Dies zeigen wir mit (572). Es gilt für $t \in \mathbb{R} \setminus \{0\}$

$$\frac{f(x_0 + tv) - f(x_0)}{t} = \frac{(x_0 + tv)^{(i)} - x_0^{(i)}}{t} = \frac{x_0^{(i)} + tv^{(i)} - x_0^{(i)}}{t} = \frac{tv^{(i)}}{t} = v^{(i)}.$$

Wir folgern

$$\partial_v f(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} = \lim_{t \rightarrow 0} v^{(i)} = v^{(i)}.$$

Die Ableitungsformel in dem vorigen Beispiel kann man auch aufschreiben als

$$\partial_v(x^{(i)}) = v^{(i)} \quad \forall i = 1, \dots, d. \quad (574)$$

Beispiel 10.5.8

Es sei $f : \mathbb{R}^d \rightarrow \mathbb{R}$ gegeben durch $f(x) = \|x\|^2$ und $x_0 \in \mathbb{R}^d$.

Behauptung. Für alle $v \in \mathbb{R}^d \setminus \{0\}$, ist f bei x_0 in Richtung v richtungsdifferenzierbar.

Beweis. Dies zeigen wir mit (572). Nach der Definition der Euklidischen Norm gilt

$$\begin{aligned} \partial_v f(x_0) &\stackrel{(572)}{=} \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} = \lim_{t \rightarrow 0} \frac{\|x_0 + tv\|^2 - \|x_0\|^2}{t} \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \left(\sum_{i=1}^d (x_0^{(i)} + tv^{(i)})^2 - \sum_{i=1}^d (x_0^{(i)})^2 \right) \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \left(\sum_{i=1}^d \left[(x_0^{(i)})^2 + 2tx_0^{(i)}v^{(i)} + t^2(v^{(i)})^2 \right] - \sum_{i=1}^d (x_0^{(i)})^2 \right) \\ &= \lim_{t \rightarrow 0} \frac{1}{t} \left(2t \sum_{i=1}^d x_0^{(i)}v^{(i)} + t^2 \sum_{i=1}^d (v^{(i)})^2 \right) = \lim_{t \rightarrow 0} \left(2 \sum_{i=1}^d x_0^{(i)}v^{(i)} + t \sum_{i=1}^d (v^{(i)})^2 \right) \\ &= 2 \sum_{i=1}^d x_0^{(i)}v^{(i)}. \end{aligned}$$

17.06.2026

Für die Richtungsableitung gibt es einige Rechenregeln, mehr oder weniger direkt aus Regeln für das eindimensionale Differenzieren folgen.

Proposition 10.5.9 — RECHENREGELN FÜR RICHTUNGSABLEITUNGEN

Es sei $D \subseteq \mathbb{R}^d$ offen und $f, g : D \rightarrow \mathbb{R}$ Funktionen, die an einer Stelle $x_0 \in D$ in Richtung $v \in \mathbb{R}^d \setminus \{0\}$ richtungsdifferenzierbar sind. Ferner seien $\lambda, \mu \in \mathbb{R}$.

- (i) *Summenregel.* Die Funktion $(\lambda f + \mu g) : D \rightarrow \mathbb{R}$ ist an der Stelle x_0 in Richtung v richtungsdifferenzierbar und es gilt

$$\partial_v(\lambda f + \mu g)(x_0) = \lambda \partial_v f(x_0) + \mu \partial_v g(x_0)$$

- (ii) *Produktregel.* Die Funktion $(f \cdot g) : D \rightarrow \mathbb{R}$ ist an der Stelle x_0 in Richtung v richtungsdifferenzierbar und es gilt

$$\partial_v(f \cdot g)(x_0) = \partial_v f(x_0)g(x_0) + f(x_0)\partial_v g(x_0)$$

- (iii) *Quotientenregel.* Ist $g(x_0) \neq 0$ so ist $\frac{f}{g} : D \rightarrow \mathbb{R}$ an der Stelle x_0 in Richtung v richtungsdifferenzierbar und es gilt

$$\partial_v\left(\frac{f}{g}\right)(x_0) = \frac{g(x_0)\partial_v f(x_0) - f(x_0)\partial_v g(x_0)}{g(x_0)^2}$$

- (iv) *Kettenregel.^a* Ist $I \subseteq \mathbb{R}$ so dass $f(D) \subseteq I$ und $h : I \rightarrow \mathbb{R}$ bei $f(x_0)$ (eindimensional) differenzierbar, so ist auch $h \circ f : D \rightarrow \mathbb{R}$ an der Stelle x_0 in Richtung v richtungsdifferenzierbar und

$$\partial_v(h \circ f)(x_0) = h'(f(x_0))\partial_v f(x_0)$$

- (v) *Homogenität in v .* Ist $\lambda \in \mathbb{R} \setminus \{0\}$ so ist f an der Stelle x_0 auch in Richtung λv richtungsdifferenzierbar und es gilt

$$\partial_{\lambda v} f(x_0) = \lambda \partial_v f(x_0).$$

^aWichtig bei dieser Kettenregel ist, dass wir hier derzeit nicht zwei mehrdimensionale Funktionen miteinander verketten können sondern nur eine mehrdimensionale reellwertige Funktion mit einer klassisch differenzierbaren Funktion in 1D. Mit der Differenzierbarkeit der Verkettung zweier mehrdimensionaler Funktionen werden wir uns erst später in Satz 10.5.33 auseinandersetzen können und brauchen dafür sogar einen anderen Begriff der Differenzierbarkeit.

Beweis. Wir zeigen nur (ii) und (v). Der Rest ist den Lesenden als Übungsaufgabe überlassen.

Zu (ii). Es sei $\gamma_v : D_v \rightarrow \mathbb{R}, t \mapsto f(x_0 + tv)$ die v -Sektorfunktion von f und $\tilde{\gamma}_v : D_v \rightarrow \mathbb{R}, t \mapsto g(x_0 + tv)$ die v -Sektorfunktion von g . Die v -Sektorfunktion von $(f \cdot g)$ lässt sich dann für $t \in D_v$ berechnen als

$$t \mapsto (f \cdot g)(x_0 + tv) = f(x_0 + tv)g(x_0 + tv) = \gamma_v(t)\tilde{\gamma}_v(t) = (\gamma_v\tilde{\gamma}_v)(t).$$

Nach (572) ergibt sich

$$\begin{aligned} \partial_v(f \cdot g)(x_0) &= (\gamma_v\tilde{\gamma}_v)'(0) \stackrel{\text{Prop 7.1.1}}{=} \gamma_v'(0)\tilde{\gamma}_v(0) + \gamma_v(0)\tilde{\gamma}_v'(0) \\ &= \partial_v f(x_0)g(x_0 + 0 \cdot v) + f(x_0 + 0 \cdot v)\partial_v g(x_0) = \partial_v f(x_0)g(x_0) + f(x_0)\partial_v g(x_0). \end{aligned}$$

Zu (v). Es seien $\gamma_v : D_v \rightarrow \mathbb{R}, t \mapsto f(x_0 + tv)$ die v -Sektorfunktion von f und $\gamma_{\lambda v} : D_{\lambda v} \rightarrow \mathbb{R}, t \mapsto f(x_0 + t(\lambda v))$ die λv -Sektorfunktion von f . Dann gilt für t in einer hinreichend kleinen Umgebung von 0, dass

$$\gamma_{\lambda v}(t) = f(x_0 + t(\lambda v)) = f(x_0 + (\lambda t)v) = \gamma_v(\lambda t). \quad (575)$$

Mit der Kettenregel folgern wir

$$\begin{aligned} \partial_{\lambda v} f(x_0) &\stackrel{(570)}{=} \gamma'_{\lambda v}(0) = \left. \frac{d}{dt} \gamma_{\lambda v}(t) \right|_{t=0} = \left. \frac{d}{dt} \gamma_v(\lambda t) \right|_{t=0} \stackrel{(575)}{=} \\ &= \lambda \gamma'_v(\lambda \cdot 0) = \lambda \gamma'_v(0) \stackrel{(570)}{=} \lambda \partial_v f(x_0) \end{aligned}$$

Beispiel 10.5.10

Mithilfe der obigen Regeln und der Formel $\partial_v(x^{(i)}) = v^{(i)}$ (vgl. (574)) lassen sich sehr viele Richtungsableitungen ohne komplizierte Grenzwertbetrachtung berechnen. Als Beispiel betrachten wir nochmal die Funktion $f : \mathbb{R}^d \rightarrow \mathbb{R}, x \mapsto \|x\|^2$ aus Beispiel 10.5.8. Für $v \in \mathbb{R}^d \setminus \{0\}$ und $x \in \mathbb{R}^d$ gilt nun

$$\partial_v(\|x\|^2) = \partial_v \left(\sum_{i=1}^d (x^{(i)})^2 \right) \stackrel{\text{Prop 10.5.9(i)}}{=} \sum_{i=1}^d \partial_v[(x^{(i)})^2] \stackrel{\text{Prop 10.5.9(iv)}}{=} \sum_{i=1}^d 2x^{(i)} \partial_v x^{(i)} \stackrel{(574)}{=} 2 \sum_{i=1}^d x^{(i)} v^{(i)}$$

Partielle Ableitungen

Spezielle Richtungsableitungen, die uns häufiger interessieren werden, sind Richtungsableitungen in Richtung der *kanonischen Einheitsvektoren* $e_1, \dots, e_d$. Diese nennen wir *partielle Ableitungen*.

Definition 10.5.11 — PARTIELLE ABLEITUNGEN

Es sei $D \subseteq \mathbb{R}^d$ und $f : D \rightarrow \mathbb{R}^m$. Ferner sei $i \in \{1, \dots, d\}$. Wir nennen f an einer Stelle $x_0 \in D$ in die i -te Koordinatenrichtung partiell differenzierbar, wenn f bei x_0 in die Richtung e_i richtungsdifferenzierbar ist. Wir nennen in diesem Fall

$$\partial_i f(x_0) := \partial_{e_i} f(x_0)$$

die *partielle Ableitung von f an der Stelle x_0 in die i -te Koordinatenrichtung*.

Weitere Bezeichnungen sind

$$\frac{\partial f}{\partial x^{(i)}}(x_0) \quad \text{oder} \quad \left. \frac{\partial}{\partial x^{(i)}} f(x) \right|_{x=x_0}.$$

Ist f an der Stelle x_0 in jede Koordinatenrichtung partiell differenzierbar, so nennen wir f an der Stelle x_0 *partiell differenzierbar*. Ferner nennen wir $f : D \rightarrow \mathbb{R}^m$ wie üblich *auf D partiell differenzierbar*, wenn f an jedem Punkt $x_0 \in D$ partiell differenzierbar ist. In diesem Fall heißt

$$\partial_i f : D \rightarrow \mathbb{R}^m \quad x \rightarrow \partial_i f(x)$$

die i -te *partielle Ableitungsfunktion*. Desweiteren heißt f auf D *stetig partiell differenzierbar*, falls alle partiellen Ableitungsfunktionen $\partial_1 f, \dots, \partial_d f : D \rightarrow \mathbb{R}^m$ auf D stetig sind.

Bemerkung 10.5.12 — BERECHNUNG PARTIELLER ABLEITUNGEN

Es sei $f : D \rightarrow \mathbb{R}$ and einer Stelle $x_0 \in D$ partiell differenzierbar. Nach Bemerkung 10.5.5 gilt für $i \in \{1, \dots, d\}$

$$\begin{aligned} \partial_i f(x_0) &= \partial_{e_i} f(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + te_i) - f(x_0)}{t} \\ &= \lim_{t \rightarrow 0} \frac{f(x_0^{(1)}, \dots, x_0^{(i-1)}, x_0^{(i)} + t, x_0^{(i+1)}, \dots, x_0^{(d)}) - f(x_0^{(1)}, \dots, x_0^{(i-1)}, x_0^{(i)}, x_0^{(i+1)}, \dots, x_0^{(d)})}{t}. \end{aligned}$$

Substituieren wir gedanklich $x := x_0^{(i)} + t$ so erhalten wir

$$\partial_i f(x_0) = \lim_{x \rightarrow x_0^{(i)}} \frac{f(x_0^{(1)}, \dots, x_0^{(i-1)}, x, x_0^{(i+1)}, \dots, x_0^{(d)}) - f(x_0^{(1)}, \dots, x_0^{(i-1)}, x_0^{(i)}, x_0^{(i+1)}, \dots, x_0^{(d)})}{x - x_0^{(i)}}. \quad (576)$$

Definiere nun (dort wo dies erlaubt ist) $h(x) := f(x_0^{(1)}, \dots, x_0^{(i-1)}, x, x_0^{(i+1)}, \dots, x_0^{(d)})$ und beachte: h ist so definiert, dass alle Komponenten von f bis auf die i -te fest gelassen werden und nur die i -te Komponente als Variable aufgefasst wird. Nach (576) gilt

$$\partial_i f(x_0) = \lim_{x \rightarrow x_0^{(i)}} \frac{h(x) - h(x_0^{(i)})}{x - x_0^{(i)}} = h'(x_0^{(i)}). \quad (577)$$

Diese Formel lehrt uns, wie die die partielle Ableitung $\partial_i f(x_0)$ im Allgemeinen zu berechnen ist

- Wir halten alle Variablen bis auf die i -te fest (mit den Werten $(x_0^{(1)}, \dots, x_0^{(i-1)}, x_0^{(i+1)}, \dots, x_0^{(d)})$).
- Die resultierende Funktion (hier h) hängt dann nur noch von einer einzigen Variablen ab.
- Bilde dann die (eindimensionale) Ableitung dieser resultierenden Funktion bei $x_0^{(i)}$ um $\partial_i f(x_0)$ zu erhalten (vgl. (577)).

Wie angekündigt schreiben wir bei einer Auswertung einer Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ statt

$f(x^{(1)}, x^{(2)})$ auch oft $f(x, y)$. In dem Fall definieren wir auch

$$\partial_x f(x, y) := \partial_1 f(x, y)$$

und

$$\partial_y f(x, y) := \partial_2 f(x, y).$$

Die anderen Bezeichnungen aus Definition 10.5.11 sind dann mit derselben Konvention auch erlaubt, so wie $\frac{\partial}{\partial x} f(x, y)$, $\frac{\partial}{\partial y} f(x, y)$ etc.

Beispiel 10.5.13

Die vorige Bemerkung macht das partielle Ableiten leicht: Gegeben sei z.B.

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := xe^{3y}.$$

Dann gilt

$$\partial_x f(x, y) = e^{3y} \quad \partial_y f(x, y) = 3xe^{3y}.$$

Wir haben hierbei die erste partielle Ableitung so bestimmt, dass wir die y -Variable *fest gelassen* (d.h. als Konstante betrachtet) haben und nach der x -Variablen abgeleitet haben — genau wie in Bemerkung 10.5.12 beschrieben! Bei der zweiten partiellen Ableitung haben wir die x -Variable fest gelassen und nach der y -Variablen abgeleitet.

Wenn wir alle partiellen Ableitungen aller Komponentenfunktionen bilden können, können wir all diese partiellen Ableitungen in eine Matrix schreiben, die sogenannte *Jacobi-Matrix*. Sie wird im Laufe dieses Kapitels noch eine große Rolle spielen. Kurz zur Erinnerung: Eine reellwertige $m \times d$ -Matrix ist ein Rechtecksschema von Zahlen mit m Zeilen und d Spalten. Die Menge aller solcher reellwertiger $m \times d$ -Matrizen bezeichnen wir mit $\mathbb{R}^{m \times d}$.

Definition 10.5.14 — JACOBI-MATRIX

Es sei $D \subseteq \mathbb{R}^d$ und $f : D \rightarrow \mathbb{R}^m$ an einer Stelle $x_0 \in D$ partiell differenzierbar. Dann definieren wir die *Jacobi-Matrix* von f an der Stelle x_0 als

$$J_f(x_0) := \begin{pmatrix} \partial_1 f^{(1)}(x_0) & \partial_2 f^{(1)}(x_0) & \cdots & \partial_d f^{(1)}(x_0) \\ \partial_1 f^{(2)}(x_0) & & & \vdots \\ \vdots & \ddots & & \vdots \\ \partial_1 f^{(m)}(x_0) & \cdots & \cdots & \partial_d f^{(m)}(x_0) \end{pmatrix} \in \mathbb{R}^{m \times d}$$

Mit anderen Worten ist $J_f(x_0)$ eine $m \times d$ -Matrix mit Einträgen $\partial_j f^{(i)}(x_0)$ in der i -ten Zeile und j -ten Spalte.

Bemerkung 10.5.15 — GRADIENT

Es sei $D \subseteq \mathbb{R}^d$ und $f : D \rightarrow \mathbb{R}$ an einer Stelle $x_0 \in D$ partiell differenzierbar (d.h. also $m = 1$). Dann ist $J_f(x_0)$ eine $1 \times d$ -Matrix bestehend aus

$$J_f(x_0) = (\partial_1 f(x_0) \quad \cdots \quad \partial_d f(x_0))$$

Eine $1 \times d$ Matrix ist ja eigentlich nichts anderes als ein Vektor in $\mathbb{R}^d$. Gerne würden wir $J_f(x_0)$ daher in diesem Fall auch als Vektor auffassen. In der Vorlesung betrachten wir aber lieber Spaltenvektoren als Zeilenvektoren. Daher definieren wir den sogenannten *Gradienten* von f an der Stelle x_0 durch den Spaltenvektor

$$\nabla f(x_0) := (\partial_1 f(x_0), \dots, \partial_d f(x_0))^T = \begin{pmatrix} \partial_1 f(x_0) \\ \vdots \\ \partial_d f(x_0) \end{pmatrix}.$$

Ist $D \subseteq \mathbb{R}^d$ und $f = (f^{(1)}, \dots, f^{(m)})^T : D \rightarrow \mathbb{R}^m$ bei $x_0 \in D$ partiell differenzierbar so gilt für $j \in \{1, \dots, m\}$

$$\nabla f^{(j)}(x_0)^T = (\partial_1 f^{(j)}(x_0) \quad \cdots \quad \partial_d f^{(j)}(x_0)) \quad \hat{=} \quad j\text{-te Zeile der Jacobimatrix.}$$

Die Zeilen der Jacobimatrix $J_f(x_0)$ sind also gegeben durch $\nabla f^{(1)}(x_0)^T, \dots, \nabla f^{(m)}(x_0)^T$.

Bemerkung 10.5.16 — KURVEN

Ein Spezialfall, der häufiger auftauchen wird sind sogenannte *Kurven*, d.h. man hat $D \subseteq \mathbb{R}^1$ und eine („partiell“ differenzierbare) Abbildung $\gamma : D \rightarrow \mathbb{R}^m, t \mapsto \gamma(t)$. Da $D \subseteq \mathbb{R}^1$ gibt es nur eine einzige Variable und man hat für $t_0 \in D$

$$J_\gamma(t_0) = \begin{pmatrix} \partial_1 \gamma^{(1)}(t_0) \\ \vdots \\ \partial_1 \gamma^{(m)}(t_0) \end{pmatrix} \quad \text{eine } m \times 1\text{-Matrix!}$$

Bemerkung 10.5.12 sagt uns nun wie die partielle Ableitung der Komponentenfunktion $t \mapsto \gamma^{(j)}(t)$ bei t_0 nach t berechnet werden muss.^a Man erhält

$$J_\gamma(t_0) = \begin{pmatrix} (\gamma^{(1)})'(t_0) \\ \vdots \\ (\gamma^{(m)})'(t_0) \end{pmatrix}. \quad (578)$$

Im Fall von Kurven definieren wir dann auch

$$\gamma'(t_0) := J_\gamma(t_0) = \begin{pmatrix} (\gamma^{(1)})'(t_0) \\ \vdots \\ (\gamma^{(m)})'(t_0) \end{pmatrix}. \quad (579)$$

^aNämlich so: Man hält alle Variablen außer t fest (in unserem eindimensionalen Fall gibt es gar keine weiteren Variablen) und leitet dann nach t ab.

Beispiel 10.5.17

Gemäß (579) würde zum Beispiel für die Kurve

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^2, \quad t \mapsto \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

gelten, dass

$$\gamma'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} \quad \forall t \in \mathbb{R}$$

Ein weiteres Beispiel: Seien $a, b \in \mathbb{R}^m$ beliebig. Man betrachte die Kurve

$$\gamma : \mathbb{R} \rightarrow \mathbb{R}^m, \quad t \mapsto a + t(b - a).$$

Behauptung. Dann gilt für alle $t \in \mathbb{R}$ $\gamma'(t) = b - a$.

Dazu: Für $i \in \{1, \dots, m\}$ gilt

$$\gamma^{(i)}(t) = (a + t(b - a))^{(i)} = a^{(i)} + t(b^{(i)} - a^{(i)}) \quad \implies \quad (\gamma^{(i)})'(t) = b^{(i)} - a^{(i)}.$$

Somit gilt

$$\gamma'(t) = \begin{pmatrix} (\gamma^{(1)})'(t) \\ \vdots \\ (\gamma^{(m)})'(t) \end{pmatrix} = \begin{pmatrix} b^{(1)} - a^{(1)} \\ \vdots \\ b^{(m)} - a^{(m)} \end{pmatrix} = b - a.$$

Warnung: Partielle Differenzierbarkeit impliziert nicht die Stetigkeit! Wie bereits in der Einleitung angesprochen impliziert die partielle Differenzierbarkeit nicht die Stetigkeit. Diese Problematik wird uns, wie schon in der Einleitung angesprochen, dazu verleiten, einen weiteren Differenzierbarkeitsbegriff einzuführen.

Beispiel 10.5.18

Wir betrachten nochmal die Funktion aus Beispiel 10.4.9

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := \begin{cases} \frac{xy}{x^2+y^2} & (x, y)^T \neq (0, 0)^T \\ 0 & (x, y)^T = (0, 0)^T \end{cases}.$$

Behauptung. f ist auf ganz $\mathbb{R}^2$ partiell differenzierbar aber in $(0, 0)^T$ nicht stetig.

Beweis. Dass f in $(0,0)^T$ nicht stetig ist haben wir bereits in Beispiel 10.4.9 gesehen. Dass f auf $\mathbb{R}^2 \setminus \{(0,0)\}$ partiell differenzierbar ist, ist leicht. (Betrachtet man zum Beispiel y als Konstante so kann die Funktion $x \mapsto \frac{xy}{x^2+y^2}$ bequem mithilfe der Ableitungsregeln aus Kapitel 7.1 nach x abgeleitet werden, jedenfalls an allen Stellen x für die gilt $x^2 + y^2 \neq 0$).

Noch $\mathbb{Z}$: f ist auch bei $(0,0)^T$ partiell diffbar.

Dazu berechnen wir für $t \in \mathbb{R} \setminus \{0\}$

$$\frac{f(t, 0) - f(0, 0)}{t} = \frac{\frac{t \cdot 0}{t^2 + 0^2} - 0}{t} = \frac{0}{t} = 0. \quad (580)$$

Daher existiert

$$\partial_x f(0, 0) = \partial_{e_1} f(0, 0) \stackrel{(572)}{=} \lim_{t \rightarrow 0} \frac{f(0_{\mathbb{R}^2} + te_1) - f(0_{\mathbb{R}^2})}{t} = \lim_{t \rightarrow 0} \frac{f(t, 0) - f(0, 0)}{t} \stackrel{(580)}{=} 0.$$

Analog gilt für $t \in \mathbb{R} \setminus \{0\}$

$$\frac{f(0, t) - f(0, 0)}{t} = \frac{\frac{0 \cdot t}{0^2 + t^2} - 0}{t} = \frac{0}{t} = 0.$$

Damit überlegt man sich ganz analog, dass

$$\partial_y f(0, 0) = 0.$$

Es folgt, dass f auch in $x_0 = (0,0)^T$ partiell differenzierbar ist. Somit haben wir gezeigt dass es eine partiell differenzierbare Funktion gibt, die nicht stetig ist.

Es wäre schon wünschenswert, dass wir einen Differenzierbarkeitsbegriff finden, der auch die Stetigkeit impliziert. Leider leistet das der Begriff der partiellen Differenzierbarkeit nicht. Ein stärkeres Konzept, welches wir bereits kennen, ist das Konzept der Richtungs-differenzierbarkeit in jede Richtung. Aber auch dies impliziert die Stetigkeit nicht.

Beispiel 10.5.19

Wir betrachten die Funktion

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := \begin{cases} \frac{xy^2}{x^2+y^4} & (x, y)^T \neq (0, 0)^T \\ 0 & (x, y)^T = (0, 0)^T \end{cases}$$

Behauptung. f ist auf ganz $\mathbb{R}^2$ richtungs-differenzierbar aber in $(0,0)^T$ nicht stetig.

Beweis.

Zur Richtungs-differenzierbarkeit.

Zunächst untersuchen wir die Richtungs-differenzierbarkeit. Auf $\mathbb{R}^2 \setminus \{(0, 0)^T\}$ ist diese wieder eine direkte Konsequenz aus den Ableitungsregeln in Kapitel 7.1. Nun zur Richtungs-differenzierbarkeit in $(0, 0)^T$. Für $v = (v^{(1)}, v^{(2)}) \in \mathbb{R}^2 \setminus \{0\}$ ist die v -Sektorfunktion $\gamma_v : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$\begin{aligned}\gamma_v(t) &= f(0_{\mathbb{R}^2} + tv) = f(tv^{(1)}, tv^{(2)}) = \frac{t^3 v^{(1)} (v^{(2)})^2}{t^2 (v^{(1)})^2 + t^4 (v^{(2)})^4} \\ &= t \frac{v^{(1)} (v^{(2)})^2}{(v^{(1)})^2 + t^2 (v^{(2)})^4} \quad (t \in \mathbb{R} \setminus \{0\})\end{aligned}\tag{581}$$

Ferner gilt $\gamma_v(0) = f(0_{\mathbb{R}^2}) = 0$. Um die Ableitung zu bilden unterscheiden wir nun zwei Fälle.

FALL 1. $v^{(1)} = 0$ (insbesondere $v^{(2)} \neq 0$, da $v \in \mathbb{R}^2 \setminus \{0_{\mathbb{R}^2}\}$).

Dann gilt nach (581) $\gamma_v(t) = 0$ für alle $t \in \mathbb{R}$. Deswegen haben wir

$$\partial_v f(0_{\mathbb{R}^2}) = \gamma'_v(0) = 0.$$

FALL 2. $v^{(1)} \neq 0$.

Dann sieht man leicht, dass die Formel in (581) auch für $t = 0$ Sinn ergibt und dort mit $\gamma_v(0) = 0$ übereinstimmt. Aus diesem Grund berechnen wir leicht

$$\begin{aligned}\partial_v f(0_{\mathbb{R}^2}) &= \gamma'_v(0) = \lim_{t \rightarrow 0} \frac{\gamma_v(t) - \gamma_v(0)}{t} \\ &\stackrel{(581)}{=} \lim_{t \rightarrow 0} \frac{1}{t} \left(t \frac{v^{(1)} (v^{(2)})^2}{(v^{(1)})^2 + t^2 (v^{(2)})^4} \right) = \lim_{t \rightarrow 0} \frac{v^{(1)} (v^{(2)})^2}{(v^{(1)})^2 + t^2 (v^{(2)})^4} = \frac{(v^{(2)})^2}{v^{(1)}}.\end{aligned}$$

Wir haben also gesehen, dass f auch bei $x_0 = 0_{\mathbb{R}^2}$ in jede Richtung $v \in \mathbb{R}^2 \setminus \{0_{\mathbb{R}^2}\}$ richtungs-differenzierbar sein muss.

Zur Unstetigkeit bei $(0, 0)^T$.

Nun bleibt noch zu zeigen, dass f in $0_{\mathbb{R}^2}$ nicht stetig ist. Betrachte dazu die Folge $(a_n)_{n \in \mathbb{N}}$ gegeben durch $a_n := (\frac{1}{n^2}, \frac{1}{n})^T$ für alle $n \in \mathbb{N}$. Klarerweise haben wir $a_n \xrightarrow{n \rightarrow \infty} 0_{\mathbb{R}^2}$. Jedoch gilt

$$f(a_n) = \frac{\frac{1}{n^2} (\frac{1}{n})^2}{(\frac{1}{n^2})^2 + (\frac{1}{n})^4} = \frac{\frac{1}{n^4}}{\frac{1}{n^4} + \frac{1}{n^4}} = \frac{\frac{1}{n^4}}{\frac{2}{n^4}} = \frac{1}{2} \xrightarrow{n \rightarrow \infty} \frac{1}{2} \neq 0 = f(0_{\mathbb{R}^2})$$

Wir haben also gesehen $\lim_{n \rightarrow \infty} f(a_n) \neq f(\lim_{n \rightarrow \infty} a_n)$. Stetigkeit kann also bei $\lim_{n \rightarrow \infty} a_n = 0_{\mathbb{R}^2}$ nicht vorliegen.

Intermezzo. Matrixalgebra. Im Folgenden werden wir viel mit *Matrizen* rechnen. Unter einer $m \times d$ Matrix A verstehen wir wie schon besprochen ein *Rechtecksschema* aus m Zeilen und d Spalten. Ist für $i \in \{1, \dots, m\}$ und $j \in \{1, \dots, d\}$ die reelle Zahl a_{ij} der Eintrag in der

i -ten Zeile und j -ten Spalte, so schreiben wir

$$A = (a_{ij})_{j=1, \dots, d}^{i=1, \dots, m} \quad \text{oder auch} \quad A = \begin{pmatrix} a_{11} & \cdots & a_{1d} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{md} \end{pmatrix}.$$

Bemerkung 10.5.20 — MATRIXPRODUKTE

(A) Matrix-Vektor-Produkt.

In den nachfolgenden Abschnitten werden wir häufig das *Matrix-Vektor-Produkt* aus der Linearen Algebra verwenden. Dieses ist wie folgt definiert. Ist $A \in \mathbb{R}^{m \times d}$ eine $m \times d$ -Matrix mit Eintrag a_{ij} in der i -ten Zeile und der j -ten Spalte und ist $v = (v^{(1)}, \dots, v^{(d)})^T \in \mathbb{R}^d$ ein Vektor so definieren wir

$$Av := A \begin{pmatrix} v^{(1)} \\ \vdots \\ v^{(d)} \end{pmatrix} := \begin{pmatrix} \sum_{j=1}^d a_{1j}v^{(j)} \\ \sum_{j=1}^d a_{2j}v^{(j)} \\ \vdots \\ \sum_{j=1}^d a_{mj}v^{(j)} \end{pmatrix} \quad (582)$$

Mit anderen Worten gilt

$$Av = \sum_{i=1}^m \left(\sum_{j=1}^d a_{ij}v^{(j)} \right) e_i = \sum_{i=1}^m \sum_{j=1}^d a_{ij}v^{(j)} e_i. \quad (583)$$

Eine hilfreiche Beobachtung ergibt sich, wenn man den Spezialfall $v = e_k$ für ein $k \in \{1, \dots, d\}$ betrachtet. Dann gilt

$$Ae_k = \sum_{i=1}^m \sum_{j=1}^d a_{ij}e_k^{(j)} e_i = \sum_{i=1}^m a_{ik}e_i = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix}$$

Mit anderen Worten

$$Ae_k \text{ entspricht der } k\text{-ten Spalte von } A. \quad (584)$$

Matrix-Matrix-Produkt. Auch nicht ganz irrelevant für uns ist das *Matrix-Matrix-Produkt*: Ist $A = (a_{ij})_{j=1, \dots, d_2}^{i=1, \dots, d_1}$ eine $d_1 \times d_2$ -Matrix und $B = (b_{ij})_{j=1, \dots, d_3}^{i=1, \dots, d_2}$ eine $d_2 \times d_3$ -Matrix

so ist AB eine $d_1 \times d_3$ Matrix, deren Eintrag in der i -ten Zeile und j -ten Spalte gegeben ist durch

$$(AB)_{ij} := \sum_{k=1}^{d_2} a_{ik} b_{kj} \quad i = 1, \dots, d_1 \quad j = 1, \dots, d_3$$

Man kann zeigen, dass obiges Matrix-Matrix-Produkt einem *Assoziativgesetz* genügt, d.h. für drei Matrizen A, B, C , so, dass untiges Matrixprodukt wohldefiniert ist, gilt auch

$$A(BC) = (AB)C.$$

Transponieren. Eine weitere Operation auf Matrizen, die wir benutzen werden, ist das *Transponieren*! Ist $A = (a_{ij})_{j=1, \dots, d}^{i=1, \dots, m}$ eine $d \times m$ -Matrix so ist A^T diejenige $m \times d$ -Matrix, die in der i -ten Zeile und j -ten Spalte den Eintrag hat, der in der j -ten Zeile und i -ten Spalte von A steht. Mit anderen Worten hat A^T die Einträge

$$(A^T)_{ij} := a_{ji} \quad j = 1, \dots, d \quad i = 1, \dots, m.$$

22.06.2026

Eine wichtige Rolle spielen Matrizen bei der *Darstellung von linearen Abbildungen*.

Definition 10.5.21 — LINEARE ABBILDUNG

Eine Abbildung $L : \mathbb{R}^d \rightarrow \mathbb{R}^m$ heißt *linear*, wenn für alle $v, w \in \mathbb{R}^d$ und alle $\lambda, \mu \in \mathbb{R}$ gilt

$$L(\lambda v + \mu w) = \lambda L(v) + \mu L(w)$$

Der Zusammenhang mit Matrizen ist der Folgende: Ist $L : \mathbb{R}^d \rightarrow \mathbb{R}^m$ eine lineare Abbildung, so hat man für $v \in \mathbb{R}^d$

$$L(v) = L\left(\sum_{j=1}^d v^{(j)} e_j\right) = \sum_{j=1}^d v^{(j)} L(e_j) = \sum_{j=1}^d v^{(j)} \left(\sum_{i=1}^m L(e_j)^{(i)} e_i\right) = \sum_{i=1}^m \sum_{j=1}^d (L(e_j)^{(i)} v^{(j)}) e_i. \quad (585)$$

Definiert man nun die $m \times d$ -Matrix $A := (a_{ij})_{j=1, \dots, d}^{i=1, \dots, m} := (L(e_j)^{(i)})_{j=1, \dots, d}^{i=1, \dots, m}$, so gilt für alle $v \in \mathbb{R}^d$

$$Av \stackrel{(583)}{=} \sum_{i=1}^m \sum_{j=1}^d a_{ij} v^{(j)} e_i = \sum_{i=1}^m \sum_{j=1}^d L(e_j)^{(i)} v^{(j)} e_i \stackrel{(585)}{=} L(v).$$

FAZIT. Zu jeder linearen Abbildung $L : \mathbb{R}^d \rightarrow \mathbb{R}^m$ kann man eine $m \times d$ -Matrix A assoziieren mit

$$L(v) = Av \quad \text{für alle } v \in \mathbb{R}^d.$$

Für die Einträge dieser Matrix gilt

$$a_{ij} = L(e_j)^{(i)}. \quad (586)$$

Die obige Formel ist nicht nur sehr nützlich um lineare Abbildungen in einer einfachen Form zu schreiben sondern auch, um ihre Eigenschaften qualitativ zu analysieren.

Bemerkung 10.5.22

Jede lineare Abbildung $L : \mathbb{R}^d \rightarrow \mathbb{R}^m$ auf ganz $\mathbb{R}^d$ stetig! Um das einzusehen, definiert man sich zunächst wie oben eine $m \times d$ -Matrix A mit

$$L(v) = Av \quad \forall v \in \mathbb{R}^d.$$

Dass die Abbildung $v \mapsto Av$ stetig ist folgt leicht aus (582). Somit ist die Stetigkeit von L bewiesen.

Totale Ableitungen

Das Konzept der partiellen Differenzierbarkeit hat den Nachteil, dass partielle Differenzierbarkeit nicht zwingend die Stetigkeit impliziert (vgl. Beispiel 10.5.18). Wir definieren nun das stärkere Konzept der *totalen Differenzierbarkeit*, welches hingegen notwendigerweise die Stetigkeit impliziert.

Es sollte allerdings an dieser Stelle gesagt werden, dass es im Allgemeinen komplizierter ist, die totale Differenzierbarkeit nachzuweisen. Wir werden in diesem Abschnitt aber auch studieren, wie ein solcher Nachweis gelingen kann.

Ausgangspunkt für unsere Definition der totalen Ableitung ist, dass differenzierbare Funktionen sich durch *affin-lineare Abbildungen* approximieren lassen sollen.

Definition 10.5.23 — TOTALE DIFFERENZIERBARKEIT

Es sei $D \subseteq \mathbb{R}^d$ und $f : D \rightarrow \mathbb{R}^m$ eine Funktion. f heißt an einem Punkt $x_0 \in D$ total differenzierbar, falls x_0 ein Berührungspunkt von $D \setminus \{x_0\}$ ist und es eine lineare Abbildung $L_{x_0} : \mathbb{R}^d \rightarrow \mathbb{R}^m$ gibt mit

$$\lim_{\substack{x \rightarrow x_0 \\ x \in D \setminus \{x_0\}}} \frac{f(x) - f(x_0) - L_{x_0}(x - x_0)}{\|x - x_0\|} = 0. \quad a \quad (587)$$

Eine lineare Abbildung $L_{x_0} : \mathbb{R}^d \rightarrow \mathbb{R}^m$, die (587) genügt wird *totale Ableitung* von f an der Stelle x_0 genannt.

^aIn der eindimensionalen Analysis (vgl. Definition 7.0.1) hatten wir hier im Nenner $x - x_0$ stehen. Im Mehrdimensionalen ist $x - x_0 \in \mathbb{R}^d$ aber im Allgemeinen ein Vektor, durch den wir nicht teilen dürfen!

Moralisch bedeutet das: f lässt sich in einer Umgebung von x_0 durch die affin-lineare Funktion

$$x \mapsto f(x_0) + L_{x_0}(x - x_0)$$

approximieren. Die Approximation ist so gut, dass der Approximationsfehler „schneller gegen 0 geht als $\|x - x_0\|$ “.

Diese Heuristik kann man in eine Proposition (welche Proposition 7.0.6 aus Analysis I gewissermaßen verallgemeinert) gießen:

Proposition 10.5.24

Es sei $D \subseteq \mathbb{R}^d$ und $f : D \rightarrow \mathbb{R}^m$ sowie $x_0 \in D$ derart, dass x_0 ein Berührungspunkt von $D \setminus \{x_0\}$ ist. Dann sind äquivalent

- (a) f ist an der Stelle x_0 total differenzierbar.
- (b) Es gibt eine lineare Abbildung $L : \mathbb{R}^d \rightarrow \mathbb{R}^m$ und eine Funktion $\phi : D \rightarrow \mathbb{R}^m$, die bei x_0 stetig ist und $\phi(x_0) = 0$ erfüllt, mit

$$\forall x \in D : \quad f(x) = f(x_0) + L(x - x_0) + \|x - x_0\|\phi(x). \quad (588)$$

Trifft eine der beiden Aussagen zu, so ist die lineare Abbildung L aus Teil (b) eine totale Ableitung von f an der Stelle x_0 .

Beweisskizze. Wir schildern hier nur den Beweis von „(a) $\implies$ (b)“. Es sei $f : D \rightarrow \mathbb{R}^m$ bei x_0 total differenzierbar. Etwa gelte für eine lineare Abbildung $L_{x_0} : \mathbb{R}^d \rightarrow \mathbb{R}^m$, dass

$$\lim_{\substack{x \rightarrow x_0 \\ x \in D \setminus \{x_0\}}} \frac{f(x) - f(x_0) - L_{x_0}(x - x_0)}{\|x - x_0\|} = 0. \quad (589)$$

Definiere nun $L := L_{x_0}$ und $\phi : D \rightarrow \mathbb{R}^m$ durch

$$\phi(x) := \begin{cases} \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|} & x \in D \setminus \{x_0\} \\ 0 & x = x_0 \end{cases}$$

Mit dieser Definition zeigt man (588) leicht, denn für $x \in D \setminus \{x_0\}$ hat man

$$\begin{aligned} \phi(x) = \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|} &\Rightarrow \|x - x_0\|\phi(x) = f(x) - f(x_0) - L(x - x_0) \\ &\Rightarrow f(x_0) + L(x - x_0) + \|x - x_0\|\phi(x) = f(x). \\ &\Rightarrow f(x) = f(x_0) + L(x - x_0) + \|x - x_0\|\phi(x). \end{aligned}$$

Man überzeugt sich auch leicht davon, dass obige Gleichung auch für $x = x_0$ stimmt. Es ist per Definition auch klar, dass $\phi(x_0) = 0$.

Noch $\mathbb{Z}\mathbb{Z}$: ϕ ist bei x_0 stetig, d.h.

$$\lim_{x \rightarrow x_0} \phi(x) = \phi(x_0) (= 0).$$

Dies folgt aber direkt aus (589). Zumindest gilt

$$\begin{aligned} \lim_{\substack{x \rightarrow x_0 \\ x \in D \setminus \{x_0\}}} \phi(x) &= \lim_{\substack{x \rightarrow x_0 \\ x \in D \setminus \{x_0\}}} \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|} \\ &= \lim_{\substack{x \rightarrow x_0 \\ x \in D \setminus \{x_0\}}} \frac{f(x) - f(x_0) - L_{x_0}(x - x_0)}{\|x - x_0\|} = 0. \end{aligned}$$

Dass man die Zusatzbedingung „ $x \in D \setminus \{x_0\}$ “ weglassen kann zeigen wir hier nicht und verweisen auf die Technik in Proposition 7.0.6.

Zu „(b) $\Rightarrow$ (a)“. Es gelte (588). Somit gilt für $x \in D \setminus \{x_0\}$

$$\frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|} \stackrel{(588)}{=} \frac{\|x - x_0\| \phi(x)}{\|x - x_0\|} = \phi(x).$$

Wir erhalten

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - L(x - x_0)}{\|x - x_0\|} = \lim_{x \rightarrow x_0} \phi(x) \stackrel{\text{lt. Voraussetzung}}{\underset{\phi \text{ bei } x_0 \text{ stetig}}{=}} \phi(x_0) = 0.$$

Als Korollar ergibt sich dass total differenzierbare Funktionen immer stetig sind

Korollar 10.5.25

Es sei $D \subseteq \mathbb{R}^d$ und $f : D \rightarrow \mathbb{R}^m$ an einer Stelle $x_0 \in D$ total differenzierbar. Dann ist f bei x_0 auch stetig.

Beweis. Es sei f an der Stelle x_0 total differenzierbar. Nach Proposition 10.5.24 gibt es eine lineare Abbildung $L : \mathbb{R}^d \rightarrow \mathbb{R}^m$ und eine bei x_0 stetige Abbildung $\phi : D \rightarrow \mathbb{R}^m$ mit $\phi(x_0) = 0$ so, dass

$$\forall x \in D : \quad f(x) = f(x_0) + L(x - x_0) + \|x - x_0\| \phi(x).$$

Nun ist $x \mapsto L(x - x_0) = L(x) - L(x_0)$ nach Bemerkung 10.5.22 stetig. Auch ist klarerweise $x \mapsto \|x - x_0\|$ stetig bei x_0 . Da nun auch ϕ bei x_0 stetig ist, ist f als Komposition stetiger Funktionen stetig.

Bisher haben wir keine Möglichkeit, die lineare Abbildung L_{x_0} in Definition 10.5.23 auszurechnen. Wir wissen nichteinmal, ob L_{x_0} überhaupt eindeutig bestimmt ist! Dies werden wir nun aufklären, wenn wir den Zusammenhang zwischen totaler Differenzierbarkeit und Richtungs-differenzierbarkeit verstehen.

Satz 10.5.26 — TOTALE DIFFBARKEIT UND RICHTUNGSDIFFBARKEIT

Es sei $D \subseteq \mathbb{R}^d$ offen^a und $f : D \rightarrow \mathbb{R}^m$ an der Stelle $x_0 \in D$ total differenzierbar mit totaler Ableitung $L_{x_0} : \mathbb{R}^d \rightarrow \mathbb{R}^m$. Dann gilt

- (i) f ist an der Stelle x_0 auch in jede Richtung $v \in \mathbb{R}^d \setminus \{0\}$ richtungsdifferenzierbar und es gilt

$$L_{x_0}(v) = \partial_v f(x_0) \quad \forall v \in \mathbb{R}^d \setminus \{0\}.$$

Insbesondere ist f also auch partiell differenzierbar (vgl. Definition 10.5.11).

- (ii) Die totale Ableitung $L_{x_0} : \mathbb{R}^d \rightarrow \mathbb{R}^m$ ist eindeutig bestimmt durch

$$L_{x_0}(v) = J_f(x_0)v \quad \forall v \in \mathbb{R}^d.$$

^aDie Offenheit wird hier nur benötigt, damit wir uns um die Spitzfindigkeit aus Bemerkung 10.5.4 keine Gedanken machen müssen.

Beweis. Zu (i). Es seien f, x_0 wie in der Aussage und $L_{x_0} : \mathbb{R}^d \rightarrow \mathbb{R}^m$ eine totale Ableitung, d.h. (nach Definition 10.5.23)

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - L_{x_0}(x - x_0)}{\|x - x_0\|} = 0. \quad (590)$$

Sei nun $v \in \mathbb{R}^d \setminus \{0\}$. Wir nutzen die Formel in Bemerkung 10.5.5. Wir berechnen dazu zuerst für $t \in \mathbb{R} \setminus \{0\}$

$$\frac{f(x_0 + tv) - f(x_0)}{t} = \frac{f(x_0 + tv) - f(x_0) - L_{x_0}(tv)}{t} + \frac{L_{x_0}(tv)}{t}.$$

Da L_{x_0} linear ist gilt $\frac{L_{x_0}(tv)}{t} = \frac{tL_{x_0}(v)}{t} = L_{x_0}(v)$. Somit ergibt sich

$$\begin{aligned} \frac{f(x_0 + tv) - f(x_0)}{t} &= L_{x_0}(v) + \frac{f(x_0 + tv) - f(x_0) - L_{x_0}((x_0 + tv) - x_0)}{t} \\ &= L_{x_0}(v) + \frac{|t| \|v\|}{t} \frac{f(x_0 + tv) - f(x_0) - L_{x_0}((x_0 + tv) - x_0)}{|t| \|v\|} \\ &= L_{x_0}(v) + \frac{|t| \|v\|}{t} \frac{f(x_0 + tv) - f(x_0) - L_{x_0}((x_0 + tv) - x_0)}{\|tv\|} \\ &= L_{x_0}(v) + \frac{|t| \|v\|}{t} \frac{f(x_0 + tv) - f(x_0) - L_{x_0}((x_0 + tv) - x_0)}{\|(x_0 + tv) - x_0\|} \end{aligned} \quad (591)$$

Wir wollen nun zeigen, dass

$$\partial_v f(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} = L_{x_0}(v).$$

Dazu sei $(t_n)_{n \in \mathbb{N}} \subseteq \mathbb{R} \setminus \{0\}$ eine beliebige Folge so, dass $\lim_{n \rightarrow \infty} t_n = 0$. Definiere für $n \in \mathbb{N}$ $x_n := x_0 + t_n v$. Dann gilt klarerweise $\lim_{n \rightarrow \infty} x_n = x_0$. Nach (591) gilt

$$\frac{f(x_0 + t_n v) - f(x_0)}{t_n} = L_{x_0}(v) + \frac{|t_n| \|v\|}{t_n} \frac{f(x_n) - f(x_0) - L_{x_0}(x_n - x_0)}{\|x_n - x_0\|}. \quad (592)$$

Nun gilt (da $x_n \xrightarrow{n \rightarrow \infty} x_0$)

$$\lim_{n \rightarrow \infty} \frac{f(x_n) - f(x_0) - L_{x_0}(x_n - x_0)}{\|x_n - x_0\|} \stackrel{(590)}{=} \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - L_{x_0}(x - x_0)}{\|x - x_0\|} = 0.$$

da $\left(\frac{|t_n| \|v\|}{t_n}\right)_{n \in \mathbb{N}} \subseteq \mathbb{R}$ beschränkt ist, gilt (wegen Proposition 10.2.3)

$$\lim_{n \rightarrow \infty} \frac{|t_n| \|v\|}{t_n} \frac{f(x_n) - f(x_0) - L_{x_0}(x_n - x_0)}{\|x_n - x_0\|} = 0.$$

Daher gilt mit (592)

$$\lim_{n \rightarrow \infty} \frac{f(x_0 + t_n v) - f(x_0)}{t_n} = L_{x_0}(v) + 0 = L_{x_0}(v).$$

Da $(t_n)_{n \in \mathbb{N}} \subseteq \mathbb{R} \setminus \{0\}$ beliebig war folgern wir

$$\partial_v f(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t} = L_{x_0}(v).$$

Nun noch zu (ii). In (586) haben wir gesehen, dass jede lineare Abbildung $L : \mathbb{R}^d \rightarrow \mathbb{R}^m$ die Gestalt $v \mapsto Av$ besitzt, wobei die $m \times d$ -Matrix A aus den Einträgen $a_{ij} := L(e_j)^{(i)}$ besteht ($i = 1, \dots, m, j = 1, \dots, d$). Für $L = L_{x_0}$ ergibt sich

$$a_{ij} = L_{x_0}(e_j)^{(i)} \stackrel{(i)}{=} \partial_{e_j} f(x_0)^{(i)} \stackrel{(571)}{=} \partial_{e_j} f^{(i)}(x_0) = \partial_j f^{(i)}(x_0).$$

Dies ist aber nach Definition 10.5.14 auch der entsprechende Eintrag der Jacobi-Matrix $J_f(x_0)$ in der i -ten Zeile und j -ten Spalte. Es folgt, dass $A = J_f(x_0)$ ist und daher

$$L_{x_0}(v) = Av = J_f(x_0)v.$$

Jetzt wo wir wissen, dass die totale Ableitung eindeutig bestimmt ist, können wir ihr einen Namen geben.

Definition 10.5.27 — NOTATION FÜR DIE TOTALE ABLEITUNG

Es sei $D \subseteq \mathbb{R}^d$ offen und $f : D \rightarrow \mathbb{R}^m$ bei $x_0 \in D$ total differenzierbar. Dann notieren wir die totale Ableitung von f als

$$Df(x_0) : \mathbb{R}^d \rightarrow \mathbb{R}^m.$$

Insbesondere gilt für $v \in \mathbb{R}^d$ (nach Satz 10.5.26), dass

$$Df(x_0)(v) = J_f(x_0)v \quad \forall v \in \mathbb{R}^d.$$

Gewöhnungsbedürftig ist die *doppelte Klammer bei* $Df(x_0)(v)$ auf der linken Seite der vorigen Gleichung. Wenn man sich überlegt, mit was für Objekten man hier jongliert, ergibt aber alles Sinn: $Df(x_0)$ ist eine (von x_0 abhängige) lineare Abbildung von $\mathbb{R}^d$ nach $\mathbb{R}^m$. Wenn man diese an einem Vektor $v \in \mathbb{R}^d$ auswerten möchte, muss man eben $Df(x_0)(v)$ schreiben.

Beispiel 10.5.28

Betrachte wieder aus Beispiel 10.5.6 für $j \in \mathbb{N}$ die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch

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

Behauptung. f ist bei $(0, 0)^T$ genau dann total differenzierbar, wenn $j \geq 3$.

Fall 1, $j = 1$. Für $j = 1$ kann f in $x_0 = (0, 0)^T$ nicht total differenzierbar sein, da f dort gar nicht stetig ist, vgl. Beispiel 10.4.9.

Fall 2, $j \geq 3$. Für $j \geq 3$ ist f in $(0, 0)^T$ total differenzierbar. Um das zu sehen, raten wir zunächst einen Kandidaten für die totale Ableitung $Df(0, 0)$. Falls es diese gibt, so müsste nach Satz 10.5.26(i) der Wert $Df(0, 0)(v)$ (für $v \in \mathbb{R}^d \setminus \{0\}$) mit $\partial_v f(0, 0)$ übereinstimmen. In Beispiel 10.5.6 haben wir aber gesehen, dass alle diese Richtungsableitungen gleich 0 sind. Unsere Vermutung ist also $Df(0)(v) = 0$ für alle $v \in \mathbb{R}^d$. Wir weisen daher nach

$$\lim_{(x, y) \rightarrow (0, 0)} \frac{f(x, y) - f(0, 0) - 0}{\|(x, y)^T - (0, 0)^T\|} = 0.$$

Dazu: Für $(x, y)^T \in \mathbb{R}^2 \setminus \{(0, 0)^T\}$ gilt

$$\frac{f(x, y) - f(0, 0) - 0}{\|(x, y)^T\|} = \frac{\frac{x^j y}{x^2 + y^2}}{\sqrt{x^2 + y^2}} = \frac{x^j y}{(x^2 + y^2)^{\frac{3}{2}}}.$$

Nun gilt

$$\begin{aligned} 0 &\leq \left| \frac{f(x, y) - f(0, 0) - 0}{\|(x, y)^T\|} \right| = \frac{|x|^j |y|}{(x^2 + y^2)^{\frac{3}{2}}} \leq \frac{(\sqrt{x^2 + y^2})^j \sqrt{x^2 + y^2}}{(x^2 + y^2)^{\frac{3}{2}}} \\ &= \frac{(x^2 + y^2)^{\frac{j+1}{2}}}{(x^2 + y^2)^{\frac{3}{2}}} = (x^2 + y^2)^{\frac{j-2}{2}} = \|(x, y)^T\|^{j-2} = \|(x, y)^T - (0, 0)^T\|^{j-2}. \end{aligned} \quad (593)$$

Da $j \geq 3$ gilt $j - 2 > 0$. Dies und (593) implizieren

$$\lim_{(x,y) \rightarrow (0,0)} \left| \frac{f(x,y) - f(0,0) - 0}{\|(x,y)^T\|} \right| = 0. \quad (a)$$

Man überlegt sich dann leicht, dass auch

$$\lim_{(x,y) \rightarrow (0,0)} \frac{f(x,y) - f(0,0) - 0}{\|(x,y)^T\|} = 0.$$

Es folgt, dass f total differenzierbar ist mit Ableitung $Df(x_0)(v) = 0$ für alle $v \in \mathbb{R}^d$

Fall 3, $j = 2$. Für $j = 2$ ist f in $(0,0)^T$ nicht total differenzierbar. Hier muss man allerdings anders argumentieren als für $j = 1$, denn f ist tatsächlich im Nullpunkt stetig^b). Wir argumentieren stattdessen so: Wäre f in $(0,0)^T$ total differenzierbar, so wäre für $v \in \mathbb{R}^d \setminus \{0\}$

$$Df(0,0)(v) \stackrel{\text{Satz 10.5.26}}{=} \partial_v f(0,0) \stackrel{\text{Bsp 10.5.6}}{=} \frac{(v^{(1)})^2 v^{(2)}}{(v^{(1)})^2 + (v^{(2)})^2}$$

Es ergibt sich bereits ein Widerspruch, denn $Df(0,0) : \mathbb{R}^2 \rightarrow \mathbb{R}$ muss auf jeden Fall eine *lineare*(!) Abbildung sein! Was auf der rechten Seite steht ist aber sicher keine lineare Abbildung. Wer das genau begründen wollen würde, würde wohl sagen, dass wegen der obigen Formel $Df(0,0)((1,0)^T) = 0 = Df(0,0)((0,1)^T)$. Jedoch gilt

$$Df(0,0)((1,1)^T) = \frac{1^2 \cdot 1}{1^2 + 1^2} = \frac{1}{2} \neq 0 = Df(0,0)((1,0)^T) + Df(0,0)((0,1)^T).$$

Die Linearität von $Df(0,0)$ ist also verletzt (im Anbetracht der Tatsache, dass $(1,1)^T = (1,0)^T + (0,1)^T$).

^aDies im Detail aus (593) zu folgern, ist den Lesenden als Übungsaufgabe überlassen. Probieren Sie es mal: Nehmen Sie sich dafür eine beliebige Folge $((x_n, y_n)^T)_{n \in \mathbb{N}}$ mit $\lim_{n \rightarrow \infty} (x_n, y_n)^T = (0,0)^T$ und folgern Sie aus (593) und dem Sandwichlemma, dass $\lim_{n \rightarrow \infty} \left| \frac{f(x_n, y_n) - f(0,0) - 0}{\|(x_n, y_n)^T\|} \right| = 0$.

^bAuch wenn es für die Aufgabe nicht relevant ist zeigen wir die Stetigkeit kurz: Es gilt für alle $(x,y)^T \in \mathbb{R}^2$

$$0 \leq |f(x,y)| = \left| \frac{x^2 y}{x^2 + y^2} \right| = \frac{x^2 |y|}{x^2 + y^2} = |y| \underbrace{\frac{x^2}{x^2 + y^2}}_{\leq 1} \leq |y|$$

Da $\lim_{(x,y) \rightarrow (0,0)} 0 = 0$ und $\lim_{(x,y) \rightarrow (0,0)} |y| = 0$ folgt auch $\lim_{(x,y) \rightarrow (0,0)} |f(x,y)| = 0$. Man folgert, dass $\lim_{(x,y) \rightarrow (0,0)} f(x,y) = 0 = f(0,0)$ und somit die Stetigkeit von f in $(0,0)^T$.

Das Vorgehen aus Fall 3 des vorigen Beispiels möchten wir noch im folgenden Korollar zusammenfassen, welches direkt aus Satz 10.5.26 folgt.

Korollar 10.5.29 — EIN NOTWENDIGES KRITERIUM FÜR TOTALE DIFFBARKEIT

Es sei $D \subseteq \mathbb{R}^d$ offen und $f : D \rightarrow \mathbb{R}^m$ bei $x_0 \in D$ total differenzierbar. Dann ist f bei x_0 auch richtungsdifferenzierbar (und somit natürlich auch partiell differenzierbar). Desweiteren gilt

$$\partial_v f(x_0) = J_f(x_0)v \quad \forall v \in \mathbb{R}^d \setminus \{0\}$$

Insbesondere ist die Abbildung

$$T : \mathbb{R}^d \rightarrow \mathbb{R}^m, \quad v \mapsto \begin{cases} \partial_v f(x_0) & v \in \mathbb{R}^d \setminus \{0\} \\ 0 & v = 0 \end{cases} \quad (594)$$

eine lineare Abbildung.

Beweis. Es sei f wie in der Aussage bei x_0 total differenzierbar. Nach Satz 10.5.26 ((i) und (ii)) gilt dann für $v \in \mathbb{R}^d \setminus \{0\}$

$$\partial_v f(x_0) = Df(x_0)(v) \quad \text{und} \quad Df(x_0)(v) = J_f(x_0)v.$$

Es folgt

$$\partial_v f(x_0) = J_f(x_0)v \quad \forall v \in \mathbb{R}^d \setminus \{0\}. \quad (595)$$

Ist nun T gegeben wie in (594) so gilt für $v \in \mathbb{R}^d$

$$T(v) = \begin{cases} \partial_v f(x_0) & v \in \mathbb{R}^d \setminus \{0\} \\ 0 & v = 0 \end{cases} \stackrel{(595)}{=} \begin{cases} J_f(x_0)v & v \in \mathbb{R}^d \setminus \{0\} \\ 0 & v = 0 \end{cases} = \begin{cases} J_f(x_0)v & v \in \mathbb{R}^d \\ J_f(x_0) \cdot 0 & v = 0 \end{cases} = J_f(x_0)v.$$

Die Abbildung $v \mapsto J_f(x_0)v$ ist klarerweise linear und somit folgt die Behauptung.

Wir haben in Beispiel 10.5.28 (für $j = 1, 2$) gesehen: partiell differenzierbare Funktionen müssen nicht zwingend total differenzierbar sein. Wenn die partiellen Ableitungen aber zusätzlich stetig sind, kann die totale Differenzierbarkeit doch gefolgert werden!

24.06.2026

Satz 10.5.30 — STETIGE PARTIELLE ABLEITUNGEN $\implies$ TOTALE DIFFBARKEIT

Es sei $D \subseteq \mathbb{R}^d$ offen und $f : D \rightarrow \mathbb{R}^m$ eine Funktion. Ferner sei $x_0 \in D$. Ist f auf D partiell differenzierbar und sind alle partiellen Ableitungsfunktionen $\partial_1 f, \dots, \partial_d f : D \rightarrow \mathbb{R}^m$ in x_0 stetig, so ist f bei x_0 auch total differenzierbar.

Beweis. Da D offen ist, ist x_0 innerer Punkt von D , d.h. es gibt ein $r > 0$ so, dass $B_r(x_0) \subseteq D$. Es sei $x \in B_r(x_0)$. Dann kann man schreiben

$$x - x_0 = \sum_{j=1}^d (x - x_0)^{(j)} e_j,$$

also gilt

$$x = x_0 + \sum_{j=1}^d (x - x_0)^{(j)} e_j.$$

Definiere nun für $k = 1, \dots, d$

$$x_k := x_0 + \sum_{j=1}^k (x - x_0)^{(j)} e_j.$$

Beobachte zuerst: Für $k = 1, \dots, d$ gilt $x_k \in D$ da $x_k \in B_r(x_0)$. In der Tat: Es gilt

$$\|x_k - x_0\| = \left\| \sum_{j=1}^k (x - x_0)^{(j)} e_j \right\| = \sqrt{\sum_{j=1}^k [(x - x_0)^{(j)}]^2} \leq \sqrt{\sum_{j=1}^d [(x - x_0)^{(j)}]^2} = \|x - x_0\| < r. \quad (596)$$

Beachte nun, dass für $k = 2, \dots, d$ gilt

$$x_k - x_{k-1} = x_0 + \sum_{j=1}^k (x - x_0)^{(j)} e_j - \left(x_0 + \sum_{j=1}^{k-1} (x - x_0)^{(j)} e_j \right) = (x - x_0)^{(k)} e_k. \quad (597)$$

Man rechnet auch leicht nach, dass dieselbe Formel für $k = 1$ gilt^a, d.h. $x_1 - x_0 = (x - x_0)^{(1)} e_1$. Außerdem ist $x_d = x$ und deswegen gilt auch

$$f(x) - f(x_0) = f(x_d) - f(x_0) = \sum_{k=1}^d f(x_k) - f(x_{k-1}) = \sum_{k=1}^d \sum_{i=1}^m (f^{(i)}(x_k) - f^{(i)}(x_{k-1})) e_i.$$

Ferner ist nach (583)

$$J_f(x_0)(x - x_0) = \sum_{i=1}^m \sum_{k=1}^d \partial_k f^{(i)}(x_0) (x - x_0)^{(k)} e_i.$$

Somit gilt (nach den beiden vorigen Gleichungen)

$$f(x) - f(x_0) - J_f(x_0)(x - x_0) = \sum_{i=1}^m \sum_{k=1}^d (f^{(i)}(x_k) - f^{(i)}(x_{k-1}) - \partial_k f^{(i)}(x_0) (x - x_0)^{(k)}) e_i. \quad (598)$$

Definiere nun für $k \in \{1, \dots, d\}$ und $i = \{1, \dots, m\}$ die Hilfsfunktion $h_{i,k} : [\min\{0, (x - x_0)^{(k)}\}, \max\{0, (x - x_0)^{(k)}\}] \rightarrow \mathbb{R}$ durch

$$h_{i,k}(t) := f^{(i)}(x_{k-1} + t e_k) \quad (b)$$

Dann ist $h_{i,k}(0) = f^{(i)}(x_{k-1})$ und

$$h_{i,k}((x - x_0)^{(k)}) = f^{(i)}(x_{k-1} + (x - x_0)^{(k)} e_k) \stackrel{(597)}{=} f^{(i)}(x_{k-1} + (x_k - x_{k-1})) = f^{(i)}(x_k).$$

Nun gibt es nach dem ersten Mittelwertsatz der Differentialrechnung (Satz 7.3.1) ein $\xi_{i,k} \in [\min\{0, (x - x_0)^{(k)}\}, \max\{0, (x - x_0)^{(k)}\}]$ (*) so, dass

$$f^{(i)}(x_k) - f^{(i)}(x_{k-1}) = h_{i,k}((x - x_0)^{(k)}) - h_{i,k}(0) = h'_{i,k}(\xi_{i,k})[(x - x_0)^{(k)} - 0]$$

Da $h_{i,k}$ nur die k -te Variable von f verändert, hat man $h'_{i,k}(t) = \partial_k f^{(i)}(x_{k-1} + t e_k)$ (vgl. auch Bemerkung 10.5.12). Wir haben also

$$f^{(i)}(x_k) - f^{(i)}(x_{k-1}) = h'_{i,k}(\xi_{i,k})(x - x_0)^{(k)} = \partial_k f^{(i)}(x_{k-1} + \xi_{i,k} e_k)(x - x_0)^{(k)}.$$

Setzen wir dies in Gleichung (598) ein so erhalten wir

$$\begin{aligned} f(x) - f(x_0) - J_f(x_0)(x - x_0) &= \sum_{i=1}^m \sum_{k=1}^d (f^{(i)}(x_k) - f^{(i)}(x_{k-1}) - \partial_k f^{(i)}(x_0)(x - x_0)^{(k)}) e_i \\ &= \sum_{i=1}^m \sum_{k=1}^d (\partial_k f^{(i)}(x_{k-1} + \xi_{i,k} e_k)(x - x_0)^{(k)} - \partial_k f^{(i)}(x_0)(x - x_0)^{(k)}) e_i \\ &= \sum_{i=1}^m \sum_{k=1}^d (\partial_k f^{(i)}(x_{k-1} + \xi_{i,k} e_k) - \partial_k f^{(i)}(x_0)) (x - x_0)^{(k)} e_i. \end{aligned}$$

Bildet man die Euklidische Norm und nutzt die Dreiecksungleichung, so erhält man

$$\begin{aligned} &\|f(x) - f(x_0) - J_f(x_0)(x - x_0)\| \\ &= \left\| \sum_{i=1}^m \sum_{k=1}^d (\partial_k f^{(i)}(x_{k-1} + \xi_{i,k} e_k) - \partial_k f^{(i)}(x_0)) (x - x_0)^{(k)} e_i \right\| \\ &\leq \sum_{i=1}^m \sum_{k=1}^d \|(\partial_k f^{(i)}(x_{k-1} + \xi_{i,k} e_k) - \partial_k f^{(i)}(x_0)) (x - x_0)^{(k)} e_i\| \\ &= \sum_{i=1}^m \sum_{k=1}^d |\partial_k f^{(i)}(x_{k-1} + \xi_{i,k} e_k) - \partial_k f^{(i)}(x_0)| |(x - x_0)^{(k)}| \|e_i\| \\ &= \sum_{i=1}^m \sum_{k=1}^d |\partial_k f^{(i)}(x_{k-1} + \xi_{i,k} e_k) - \partial_k f^{(i)}(x_0)| |(x - x_0)^{(k)}| \end{aligned}$$

Nun wenden wir in der inneren Summe die Cauchy-Schwarz-Ungleichung (vgl. Lemma 10.1.3) an und erhalten

$$\|f(x) - f(x_0) - J_f(x_0)(x - x_0)\|$$

$$\leq \sum_{i=1}^m \sqrt{\sum_{k=1}^d |\partial_k f^{(i)}(x_{k-1} + \xi_{i,k} e_k) - \partial_k f^{(i)}(x_0)|^2} \sqrt{\sum_{k=1}^d |(x - x_0)^{(k)}|^2}.$$

Der letzte Faktor ist gerade die Euklidische Norm $\|x - x_0\|$. Teilen wir durch diese so finden wir

$$\left\| \frac{f(x) - f(x_0) - J_f(x_0)(x - x_0)}{\|x - x_0\|} \right\| \leq \sum_{i=1}^m \sqrt{\sum_{k=1}^d |\partial_k f^{(i)}(x_{k-1} + \xi_{i,k} e_k) - \partial_k f^{(i)}(x_0)|^2} \quad (599)$$

Nun interessiert uns der Grenzübergang $x \rightarrow x_0$. Dazu beobachte zuerst, dass für $k = 1, \dots, d$ gilt (mit der Konvention, dass im Spezialfall $k = 1 \sum_{j=1}^0 (x - x_0)^{(j)} e_j = 0$)

$$\begin{aligned} \|x_{k-1} + \xi_{i,k} e_k - x_0\| &= \left\| x_0 + \sum_{j=1}^{k-1} (x - x_0)^{(j)} e_j + \xi_{i,k} e_k - x_0 \right\| = \left\| \sum_{j=1}^{k-1} (x - x_0)^{(j)} e_j + \xi_{i,k} e_k \right\| \\ &= \sqrt{\sum_{j=1}^{k-1} [(x - x_0)^{(j)}]^2 + \xi_{i,k}^2}. \end{aligned} \quad (600)$$

Wegen (*) gilt, außerdem, dass $|\xi_{i,k}| \leq |(x - x_0)^{(k)}|$, d.h. auch $\xi_{i,k}^2 \leq [(x - x_0)^{(k)}]^2$. Dies erlaubt uns in (600) weiter nach oben abzuschätzen. Wir erhalten

$$\|x_{k-1} + \xi_{i,k} e_k - x_0\| \leq \sqrt{\sum_{j=1}^k [(x - x_0)^{(j)}]^2} \leq \sqrt{\sum_{j=1}^d [(x - x_0)^{(j)}]^2} = \|x - x_0\|. \quad (601)$$

Es sei nun $\varepsilon > 0$ beliebig vorgegeben. Da $\partial_k f^{(i)}$ bei x_0 stetig ist, wissen wir, dass es ein $\delta_{i,k} > 0$ gibt so, dass

$$\text{Für alle } \eta \in D \text{ mit } \|\eta - x_0\| < \delta_{i,k} \text{ gilt, dass } |\partial_k f^{(i)}(\eta) - \partial_k f^{(i)}(x_0)| < \frac{\varepsilon}{m\sqrt{d}}. \quad (602)$$

Definiere nun $\delta := \min(\{r\} \cup \{\delta_{i,k} : i = 1, \dots, m, k = 1, \dots, d\})$. Wir zeigen dann

$$\forall x \in D \setminus \{x_0\} : \|x - x_0\| < \delta \quad \text{gilt} \quad \left\| \frac{f(x) - f(x_0) - J_f(x_0)(x - x_0)}{\|x - x_0\|} \right\| < \varepsilon. \quad (603)$$

Es sei dazu $x \in D$ mit $\|x - x_0\| < \delta$. Definieren wir für $i \in \{1, \dots, m\}$ und $k \in \{1, \dots, d\}$ die Zahl $\eta_{i,k} := x_{k-1} + \xi_{i,k} e_k$, so lernen wir aus (601), dass

$$\|\eta_{i,k} - x_0\| \leq \|x - x_0\| < \delta \leq \delta_{i,k}.$$

Nach (602) gilt dann $|\partial_k^{(i)} f(\eta_{i,k}) - \partial_k^{(i)} f(x_0)| < \frac{\varepsilon}{m\sqrt{d}}$. Diese Abschätzung können wir dann in (599) verwenden. Es gilt

$$\begin{aligned} \left\| \frac{f(x) - f(x_0) - J_f(x_0)(x - x_0)}{\|x - x_0\|} \right\| &\leq \sum_{i=1}^m \sqrt{\sum_{k=1}^d |\partial_k f^{(i)}(x_{k-1} + \xi_{i,k} e_k) - \partial_k f^{(i)}(x_0)|^2} \\ &= \sum_{i=1}^m \sqrt{\sum_{k=1}^d |\partial_k f^{(i)}(\eta_{i,k}) - \partial_k f^{(i)}(x_0)|^2} \\ &\stackrel{(602)}{<} \sum_{i=1}^m \sqrt{\sum_{k=1}^d \left(\frac{\varepsilon}{m\sqrt{d}}\right)^2} = \sum_{i=1}^m \sqrt{d \left(\frac{\varepsilon}{m\sqrt{d}}\right)^2} = \sum_{i=1}^m \frac{\varepsilon}{m} = \varepsilon. \end{aligned}$$

Somit ist (603) gezeigt. Mit Proposition 10.4.4 folgt

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

Die Behauptung folgt und die totale Ableitung ist (wie zu erwarten war) gegeben durch

$$L_{x_0}(v) := J_f(x_0)v \quad \forall v \in \mathbb{R}^d.$$

^aDas einzige Problem, warum wir das nicht mit obiger Rechnung abarbeiten können ist, dass $\sum_{j=1}^{1-1} \dots$ nicht definiert ist. Wenn wir es für diesen Beweis als 0 definieren geht aber obiges Argument bereits für $k = 1$ durch

^bMan müsste an dieser Stelle eigentlich überprüfen, dass $h_{i,k}$ wohldefiniert ist. Dafür müsste man zeigen, dass $x_{k-1} + te_k \in D$ für alle t aus dem Definitionsbereich von $h_{i,k}$ (denn sonst dürfte $f^{(i)}$ dort gar nicht ausgewertet werden). Dies kann man sich ähnlich wie in (596) überlegen. Details sind den Lesenden als Übungsaufgabe überlassen.

Die Aussage des vorigen Satzes können wir kurz und einprägsam wie folgt zusammenfassen:

FAZIT. Stetig partiell differenzierbare Funktionen sind total differenzierbar.

Wenn man das so kurz und prägnant formuliert muss man sich aber einer Sache bewusst sein: Die Aussage " $f : D \rightarrow \mathbb{R}^m$ ist stetig partiell differenzierbar" bedeutet nicht nur, dass f stetig und partiell differenzierbar sein muss, sondern die partiellen Ableitungen $\partial_1 f, \dots, \partial_d f$ müssen auch stetig sein!

Beispiel 10.5.31

Wir betrachten noch ein weiteres Mal die Funktion aus Beispiel 10.5.28, d.h. (für $j \in \mathbb{N}$)

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

Wir haben bereits in Beispiel 10.5.28 gesehen, dass f bei $x_0 = (0, 0)^T$ nur für $j \geq 3$ total

diffbar ist. An anderen Punkten als $(0,0)^T$ haben wir die totale Diffbarkeit aber noch gar nicht untersucht!

Behauptung. Auf $\mathbb{R}^2 \setminus \{(0,0)^T\}$ ist f (unabhängig von der Wahl von j) total diffbar!

Beweis. Wir bilden bei $(x,y)^T \in \mathbb{R}^2 \setminus \{(0,0)^T\}$ die partiellen Ableitungen von f und überprüfen, ob sie stetig sind. Man hat

$$\begin{aligned}\partial_x \left(\frac{x^j y}{x^2 + y^2} \right) &\stackrel{\text{Bem 10.5.12}}{=} \frac{j x^{j-1} y}{x^2 + y^2} - \frac{x^j y}{(x^2 + y^2)^2} (2x) = \frac{(j x^{j-1} (x^2 + y^2) - 2x^{j+1}) y}{(x^2 + y^2)^2}, \\ \partial_y \left(\frac{x^j y}{x^2 + y^2} \right) &= \frac{x^j}{x^2 + y^2} - \frac{x^j y}{(x^2 + y^2)^2} (2y) = \frac{x^j (x^2 + y^2) - 2x^j y^2}{(x^2 + y^2)^2}.\end{aligned}$$

Somit sind $\partial_x f(x,y)$ und $\partial_y f(x,y)$ auf $\mathbb{R}^2 \setminus \{(0,0)^T\}$ stetig. Die totale Differenzierbarkeit auf $\mathbb{R}^2 \setminus \{(0,0)\}$ folgt aus Satz 10.5.30.

Zum Schluss dieses Abschnittes sprechen wir noch über die *Konsistenz* mit der in der Analysis I definierten Ableitung. Schließlich soll der mehrdimensionale Ableitungsbegriff sich ja nicht vom eindimensionalen unterscheiden (sondern den Begriff nur erweitern).

Proposition 10.5.32 — KONSISTENZ MIT DER EINDIMENSIONALEN ABLEITUNG

Es sei $D \subseteq \mathbb{R}$ und $f : D \rightarrow \mathbb{R}$ eine Funktion. Ferner sei $x_0 \in D$. Dann sind folgende Aussagen äquivalent

- (a) f ist bei x_0 differenzierbar (im Sinne von Definition 7.0.1)
- (b) f ist bei x_0 total differenzierbar

Trifft eine der beiden obigen Aussagen zu so gilt für die totale Ableitung von f

$$Df(x_0)(v) = f'(x_0) \cdot v \quad \forall v \in \mathbb{R}.$$

Beweis. Zuerst zu „(a) $\implies$ (b)“. Es sei f im Sinne von 7.0.1 differenzierbar mit Ableitung $f'(x_0)$. Das heißt es gilt

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0). \quad (604)$$

Wir überprüfen nun Definition 10.5.23 mit der linearen Abbildung $L_{x_0} : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch $L_{x_0}(v) := f'(x_0) \cdot v$. Dazu: Für $x \in D \setminus \{x_0\}$ berechnen wir

$$\begin{aligned}\frac{f(x) - f(x_0) - L_{x_0}(x - x_0)}{\|x - x_0\|_{\mathbb{R}^1}} &= \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{|x - x_0|} \\ &= \frac{x - x_0}{|x - x_0|} \left(\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right).\end{aligned} \quad (605)$$

Nun gilt nach (604)

$$\lim_{x \rightarrow x_0} \left(\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right) = 0.$$

Ferner ist $x \mapsto \frac{x-x_0}{|x-x_0|}$ auf $D \setminus \{x_0\}$ beschränkt. Wir folgern aus Blatt 13 Aufgabe 3f Analysis I, dass

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

Mit (605) gilt

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - L_{x_0}(x - x_0)}{\|x - x_0\|_{\mathbb{R}^1}} = 0.$$

Die totale Differenzierbarkeit von f nach Definition 10.5.23 ist gezeigt. Die Formel für die totale Ableitung ergibt sich aus der getroffenen Wahl

$$Df(x_0)(v) = L_{x_0}(v) = f'(x_0) \cdot v.$$

Nun zu „(b) $\implies$ (a)“. Ist f total differenzierbar so ist nach Satz 10.5.26 auch partiell differenzierbar nach jeder Variable. Wenn es nur eine Variable gibt, stimmt die partielle Ableitung aber mit der Ableitung im Sinne von Definition 7.0.1 überein, wie man aus Bemerkung 10.5.12 im Spezialfall $d = 1$ rauslesen kann. Die Behauptung folgt.

Zusammenfassung. Zusammenhang zwischen den Differenzierbarkeitsbegriffen

Zwischen den behandelten Differenzierbarkeitskonzepten besteht nun der folgende hierarchische Zusammenhang. Für eine offene Menge $D \subseteq \mathbb{R}^d$ und eine Funktion $f : D \rightarrow \mathbb{R}^m$ gilt

f ist stetig partiell differenzierbar (vgl. Definition 10.5.11)

$\Downarrow$ (Satz 10.5.30)

f ist total differenzierbar (vgl. Definition 10.5.23)

$\Downarrow$ (Korollar 10.5.29)

f ist richtungsdifferenzierbar und $v \mapsto \partial_v f(x)$ ist linear für alle $x \in D$

$\Downarrow$

f ist richtungsdifferenzierbar (vgl. Definition 10.5.3)

⇓ (Definition 10.5.11)

f ist partiell differenzierbar

Die umgekehrten Implikationen gelten allesamt nicht

- Partiiell differenzierbare Funktionen sind nicht zwingend richtungsdifferenzierbar (man überzeugt sich zum Beispiel leicht, dass die Funktion aus Beispiel 10.4.9 partiell diffbar in $(0, 0)^T$ aber nicht richtungsdiffbar ist)
- Richtungsdifferenzierbare Funktionen haben nicht zwingend eine lineare Richtungsableitung (vgl. Beispiel 10.5.28 im Fall $j = 2$.)
- Richtungsdifferenzierbare Funktionen mit linearer Richtungsableitung sind nicht zwingend total differenzierbar (vgl. Blatt 12 Aufgabe 1)
- Total differenzierbare Funktionen sind nicht zwingend stetig partiell differenzierbar (vgl. Blatt 11 Aufgabe 3)

Außerdem sollte man vielleicht noch die Bedingung „ f ist stetig“ in obiges Schaubild mit aufnehmen. Man erhält dann

f ist stetig partiell differenzierbar

⇓ (Satz 10.5.30)

f ist total differenzierbar $\xrightarrow[\text{Kor10.5.25}]{} f$ ist stetig

⇓ (Korollar 10.5.29)

f ist richtungsdifferenzierbar und $v \mapsto \partial_v f(x)$ ist linear für alle $x \in D$

⇓

f ist richtungsdifferenzierbar

⇓ (Definition 10.5.11)

f ist partiell differenzierbar

Auch hier sollte man die fehlenden Implikationen klären:

- Richtungs-differenzierbare Funktionen sind nicht zwingend stetig (vgl. Beispiel 10.5.19)
- Stetige Funktionen sind nicht zwingend partiell differenzierbar (das stimmt ja bereits für $d = 1$ nicht, wie man am Beispiel $\text{abs} : \mathbb{R}^1 \rightarrow \mathbb{R}, x \mapsto |x|$ sehen kann, vgl. auch Beispiel 7.0.5)

Kettenregel und Konstanzsatz

In diesem Abschnitt besprechen wir noch die *mehrdimensionale Kettenregel* für die totale Ableitung.

Satz 10.5.33 — MEHRDIMENSIONALE KETTENREGEL

Es sei $D \subseteq \mathbb{R}^d$ offen und $f : D \rightarrow \mathbb{R}^m$ eine Funktion, die an einem Punkt $x_0 \in D$ total differenzierbar ist. Ferner sei $\tilde{D} \subseteq \mathbb{R}^m$ offen mit $f(D) \subseteq \tilde{D}$ und $h : \tilde{D} \rightarrow \mathbb{R}^k$ eine Funktion, die bei $f(x_0)$ total differenzierbar ist. Dann ist $h \circ f : D \rightarrow \mathbb{R}^k$ bei x_0 total differenzierbar und für die totale Ableitung gilt

$$D(h \circ f)(x_0) = Dh(f(x_0)) \circ Df(x_0). \quad (a)$$

^aBeachte: Diese Verkettung ist wohldefiniert, weil $Df(x_0) : \mathbb{R}^d \rightarrow \mathbb{R}^m$ und $Dh(f(x_0)) : \mathbb{R}^m \rightarrow \mathbb{R}^k$.

Beweis. Nach Proposition 10.5.24 gibt es eine in x_0 stetige Funktion $\phi_1 : D \rightarrow \mathbb{R}^m$ mit $\phi_1(x_0) = 0$ und

$$\forall x \in D : \quad f(x) = f(x_0) + Df(x_0)(x - x_0) + \|x - x_0\| \phi_1(x). \quad (606)$$

Desweiteren gibt es eine in $f(x_0)$ stetige Funktion $\phi_2 : \tilde{D} \rightarrow \mathbb{R}^k$ mit $\phi_2(f(x_0)) = 0$ so, dass

$$\forall y \in \tilde{D} : \quad h(y) = h(f(x_0)) + Dh(f(x_0))(y - f(x_0)) + \|y - f(x_0)\| \phi_2(y).$$

Da für alle $x \in D$ gilt, dass $f(x) \in \tilde{D}$, erhalten wir aus der vorigen Gleichung

$$\forall x \in D : \quad h(f(x)) = h(f(x_0)) + Dh(f(x_0))(f(x) - f(x_0)) + \|f(x) - f(x_0)\| \phi_2(f(x)).$$

Nutzen wir (606) so ergibt sich für alle $x \in D$

$$h(f(x)) = h(f(x_0)) + Dh(f(x_0))[Df(x_0)(x - x_0) + \|x - x_0\| \phi_1(x)] + \|f(x) - f(x_0)\| \phi_2(f(x))$$

Wir nutzen die Linearität von $Dh(f(x_0))$ und erhalten für $x \in D$

$$\begin{aligned} (h \circ f)(x) &= (h \circ f)(x_0) + Dh(f(x_0))[Df(x_0)(x - x_0)] \\ &\quad + \|x - x_0\| Dh(f(x_0))(\phi_1(x)) + \|f(x) - f(x_0)\| \phi_2(f(x)). \end{aligned}$$

Formt man um so ergibt sich für $x \in D$

$$(h \circ f)(x) - (h \circ f)(x_0) - Dh(f(x_0))[Df(x_0)(x - x_0)]$$

$$= \|x - x_0\| Dh(f(x_0))(\phi_1(x)) + \|f(x) - f(x_0)\| \phi_2(f(x)).$$

Beachte: $Dh(f(x_0))[Df(x_0)(x-x_0)]$ kann auch als $[Dh(f(x_0)) \circ Df(x_0)](x-x_0)$ geschrieben werden. Für $x \in D \setminus \{x_0\}$ kann man durch $\|x - x_0\|$ teilen und erhält

$$\begin{aligned} & \frac{(h \circ f)(x) - (h \circ f)(x_0) - [Dh(f(x_0)) \circ Df(x_0)](x - x_0)}{\|x - x_0\|} \\ &= Dh(f(x_0))(\phi_1(x)) + \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} \phi_2(f(x)) \end{aligned} \quad (607)$$

Wir zeigen nun in zwei Schritten, dass der Term in (607) für $x \rightarrow x_0$ den Grenzwert 0 besitzt.

SCHRITT 1. Wir zeigen $\lim_{x \rightarrow x_0} Dh(f(x_0))(\phi_1(x)) = 0$

Dazu analysieren wir zunächst den ersten Summanden. Nach Bemerkung 10.5.22 ist $Dh(f(x_0))$ stetig auf $\mathbb{R}^m$. Nach Proposition 10.4.12 ^a ist dann auch $Dh(f(x_0)) \circ \phi_1$ stetig in x_0 . Daher gilt

$$\begin{aligned} \lim_{x \rightarrow x_0} Dh(f(x_0))(\phi_1(x)) &= \lim_{x \rightarrow x_0} [Dh(f(x_0)) \circ \phi_1](x) = [Dh(f(x_0)) \circ \phi_1](x_0) \\ &= Dh(f(x_0))(\phi_1(x_0)) = Dh(f(x_0))(0) \\ &\stackrel{\text{Satz 10.5.26}}{=} J_h(f(x_0)) \cdot 0 = 0. \end{aligned} \quad (608)$$

SCHRITT 2. Wir zeigen $\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} \phi_2(f(x)) = 0$.

Nun noch zum zweiten Summanden $x \mapsto \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} \phi_2(f(x))$. Beobachte zunächst, dass wegen der Stetigkeit von f in x_0 (vgl. Korollar 10.5.25) und von ϕ_2 in $f(x_0)$ auch $\phi_2 \circ f$ stetig in x_0 ist. Daher gilt

$$\lim_{x \rightarrow x_0} \phi_2(f(x)) = \lim_{x \rightarrow x_0} (\phi_2 \circ f)(x) = (\phi_2 \circ f)(x_0) = \phi_2(f(x_0)) = 0. \quad (609)$$

Ferner gilt

$$\begin{aligned} \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} \|\phi_2(f(x))\| &= \frac{\|Df(x_0)(x - x_0) + \|x - x_0\| \phi_1(x)\|}{\|x - x_0\|} \|\phi_2(f(x))\| \\ &\stackrel{\text{Satz 10.1.2(ii)}}{=} \left\| Df(x_0) \left(\frac{x - x_0}{\|x - x_0\|} \right) + \phi_1(x) \right\| \|\phi_2(f(x))\| \\ &\stackrel{\text{Satz 10.1.2(iii)}}{\leq} \left(\left\| Df(x_0) \left(\frac{x - x_0}{\|x - x_0\|} \right) \right\| + \|\phi_1(x)\| \right) \|\phi_2(f(x))\| \end{aligned} \quad (610)$$

Nun ist^b $v \mapsto \|Df(x_0)(v)\|$ eine stetige Abbildung von $\mathbb{R}^d$ nach $\mathbb{R}$. Ferner ist $\mathbb{S} := \{y \in \mathbb{R}^d : \|y\| = 1\}$ eine kompakte Menge^c. Nach dem Satz von Weierstraß gibt es dann

$$C := \max_{v \in \mathbb{S}} \|Df(x_0)(v)\| \in \mathbb{R}. \quad (611)$$

Beachte: Für alle $x \in D \setminus \{x_0\}$ ist $\frac{x-x_0}{\|x-x_0\|} \in \mathbb{S}$. Nach (611) gilt dann auch $\|Df(x_0)(\frac{x-x_0}{\|x-x_0\|})\| \leq \max_{v \in \mathbb{S}} \|Df(x_0)(v)\| = C$. Dann lässt sich (610) weiter abschätzen und man erhält

$$\begin{aligned} 0 &\leq \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} \|\phi_2(f(x))\| \leq (C + \|\phi_1(x)\|) \|\phi_2(f(x))\| \\ &= C \|\phi_2(f(x))\| + \|\phi_1(x)\| \|\phi_2(f(x))\|. \end{aligned}$$

Wegen (609) und der Stetigkeit von ϕ_1 und von der euklidischen Norm haben wir

$$\lim_{x \rightarrow x_0} (C \|\phi_2(f(x))\| + \|\phi_1(x)\| \|\phi_2(f(x))\|) = C \cdot 0 + \|\phi_1(x_0)\| \cdot 0 = 0.$$

Wir folgern

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} \|\phi_2(f(x))\| = 0.$$

Mithilfe von Lemma 10.2.5 kann man dann auch zeigen, dass

$$\lim_{x \rightarrow x_0} \frac{\|f(x) - f(x_0)\|}{\|x - x_0\|} \phi_2(f(x)) = 0.$$

Verwenden wir dies zusammen mit (608) in (607) so gilt

$$\lim_{x \rightarrow x_0} \frac{(h \circ f)(x) - (h \circ f)(x_0) - [Dh(f(x_0)) \circ Df(x_0)](x - x_0)}{\|x - x_0\|} = 0.$$

Nach Definition 10.5.23 ist $h \circ f$ dann total differenzierbar mit totaler Ableitung^d

$$D(h \circ f)(x_0) = Dh(f(x_0)) \circ Df(x_0).$$

^aund weil ϕ_1 stetig in x_0 ist

^bvgl. Bemerkung 10.5.22

^cDavon haben wir uns bereits überzeugt, nämlich im Beweis von Lemma 10.4.27

^dAn dieser Stelle müssten wir uns eigentlich davon überzeugen, dass die Abbildung $Dh(f(x_0)) \circ Df(x_0)$ auch die in Definition 10.5.23 geforderten Eigenschaften besitzt. Zum Beispiel müsste man überprüfen, dass es sich um eine *lineare* Abbildung handelt. Wir überlassen den Lesenden als Übungsaufgabe, zu beweisen, dass die Verkettung linearer Abbildungen stets wieder linear ist.

Als Korollar ergibt sich eine Formel für die Richtungsableitungen bzw Jacobimatrix der Verkettung total differenzierbarer Funktionen

Korollar 10.5.34

Es sei $D \subseteq \mathbb{R}^d$ offen und $f : D \rightarrow \mathbb{R}^m$ eine Funktion, die an einem Punkt $x_0 \in D$ total differenzierbar ist. Ferner sei $\tilde{D} \subseteq \mathbb{R}^m$ offen mit $f(D) \subseteq \tilde{D}$ und $h : \tilde{D} \rightarrow \mathbb{R}^k$ eine Funktion, die bei $f(x_0)$ total differenzierbar ist. Dann gilt

$$J_{h \circ f}(x_0) = J_h(f(x_0))J_f(x_0).$$

Beweis. Wir wissen bereits aus Satz 10.5.33, dass $(h \circ f)$ bei x_0 total differenzierbar ist. Somit ist nach Satz 10.5.26 $h \circ f$ auch partiell differenzierbar und für $v \in \mathbb{R}^d$ gilt

$$J_{h \circ f}(x_0)v = D(h \circ f)(x_0)(v) \stackrel{\text{Satz 10.5.33}}{=} [Dh(f(x_0)) \circ Df(x_0)](v) = Dh(f(x_0))(Df(x_0)v).$$

Wendet man nochmal Satz 10.5.26 an, so hat man $Df(x_0)(v) = J_f(x_0)v$. Man folgert

$$J_{h \circ f}(x_0)v = Dh(f(x_0))(J_f(x_0)v) \stackrel{\text{Satz 10.5.26}}{=} J_h(f(x_0))(J_f(x_0)v)$$

Wegen der Assoziativität der Matrixmultiplikation gilt

$$J_{h \circ f}(x_0)v = (J_h(f(x_0))J_f(x_0))v \quad (612)$$

Nach (584) gilt für alle $k \in \{1, \dots, d\}$

$$\begin{aligned} k\text{-te Spalte von } J_{h \circ f}(x_0) &\hat{=} J_{h \circ f}(x_0)e_k \\ &\stackrel{(612)}{=} (J_h(f(x_0))J_f(x_0))e_k \hat{=} k\text{-te Spalte von } J_h(f(x_0))J_f(x_0). \end{aligned}$$

Die Behauptung folgt.

Beispiel 10.5.35

Es sei $D \subseteq \mathbb{R}^d$ offen und $f : D \rightarrow \mathbb{R}$ auf D total differenzierbar. Ferner seien $a, b \in D$ so, dass $\text{Seg}_{a,b} := \{a + t(b - a) | t \in [0, 1]\} \subseteq D$. Definiere

$$h : [0, 1] \rightarrow \mathbb{R}, \quad h(t) := f(a + t(b - a)).$$

Behauptung. h ist auf $(0, 1)$ differenzierbar und $h'(t) = \nabla f(a + t(b - a))^T(b - a) \quad \forall t \in (0, 1)$.

Beweis. Definiere hierfür $\gamma : (0, 1) \rightarrow \mathbb{R}^d$ durch $\gamma(t) := a + t(b - a)$. Insbesondere gilt $\gamma((0, 1)) \subseteq D$. Beobachte nun, dass $h = f \circ \gamma$. Da f total differenzierbar ist und γ als stetig partiell differenzierbare Funktionen auch total differenzierbar ist, ist auch h auf $(0, 1)$ total differenzierbar (vgl. Satz 10.5.33 und Satz 10.5.30). Nach der Kettenregel gilt für $t \in (0, 1)$

$$Dh(t) = Df(\gamma(t)) \circ D\gamma(t).$$

Nach Satz 10.5.26 wird $D\gamma(t)$ durch die Jacobi-Matrix $J_\gamma(t)$ dargestellt. Diese stimmt nach Bemerkung 10.5.16 mit dem Vektor $\gamma'(t)$ überein. Nach Proposition 10.5.32 ist h auch im üblichen Sinne eindimensional differenzierbar und es gilt

$$h'(t) = J_h(t) = J_{f \circ \gamma}(t) \stackrel{\text{Kor 10.5.34}}{=} J_f(\gamma(t))J_\gamma(t) = J_f(\gamma(t))\gamma'(t) = \nabla f(\gamma(t))^T \gamma'(t) \quad \forall t \in (0, 1).$$

Die Behauptung folgt nun durch Einsetzen, denn für $t \in (0, 1)$ gilt $\gamma(t) = a + t(b - a)$ und $\gamma'(t) = b - a$ (vgl. Beispiel 10.5.17).

29.06.2026

Eine Folgerung aus der Kettenregel ist der *mehrdimensionale Konstanzsatz*. Im Eindimensionalen besagte der Konstanzsatz (Proposition 7.3.8), dass das Verschwinden der Ableitung auf einem Intervall impliziert, dass die Funktion dort konstant ist. Wir erinnern daran (vgl. Bemerkung 7.3.9), dass für den Konstanzsatz essentiell ist, dass der Definitionsbereich ein Intervall ist.

Im Mehrdimensionalen wollen wir ein ähnliches Resultat beweisen. Jedoch müssen wir erstmal den Begriff des *Intervalles* auf's Mehrdimensionale verallgemeinern.

Definition 10.5.36 — WEGZUSAMMENHANG

Es sei $D \subseteq \mathbb{R}^d$ eine offene Menge. D heißt *C^1 -wegzusammenhängend*^a, falls es zu je zwei Punkten $x_0, x_1 \in D$ eine stetig differenzierbare^b Kurve $\gamma : [0, 1] \rightarrow \mathbb{R}^d$ gibt mit

- $\gamma(t) \in D$ für alle $t \in [0, 1]$
- $\gamma(0) = x_0, \gamma(1) = x_1$.

Eine solche Kurve wird *stetig differenzierbare Verbindungsstrecke* zwischen x_0 und x_1 genannt.

^aDas Symbol „ C^1 “ ist eine Abkürzung für *stetig differenzierbar*

^bIn dem Sinne dass jede Komponente stetig differenzierbar ist, vgl. auch Bemerkung 10.5.16.

In nachfolgendem Beispiel klären wir zunächst inwiefern die Aussage “ $D \subseteq \mathbb{R}^d$ ist C^1 -wegzusammenhängend” eine gute mehrdimensionale Version der Aussage “ $D \subseteq \mathbb{R}$ ist ein Intervall” ist.

Beispiel 10.5.37

Wir betrachten den Fall $d = 1$. Dort ist für alle $a, b \in \mathbb{R}$ mit $a < b$ das offene Intervall (a, b) eine C^1 -wegzusammenhängende Menge. In der Tat: Für $x_0, x_1 \in (a, b)$ betrachten wir die stetig differenzierbare Kurve $\gamma : [0, 1] \rightarrow \mathbb{R}, t \mapsto (1-t)x_0 + tx_1$. Dann gilt offenbar $\gamma(0) = x_0$ und $\gamma(1) = x_1$. Man muss sich aber auch noch davon überzeugen, dass $\gamma(t) \in (a, b)$ für alle $t \in [0, 1]$. Dazu sei $t \in [0, 1]$ beliebig. Weil $t, 1 - t$ beides nichtnegative Zahlen sind, gilt

$$\gamma(t) = (1-t)x_0 + tx_1 \leq (1-t)\max\{x_0, x_1\} + t\max\{x_0, x_1\} = \max\{x_0, x_1\}$$

und

$$\gamma(t) = (1-t)x_0 + tx_1 \geq (1-t)\min\{x_0, x_1\} + t\min\{x_0, x_1\} = \min\{x_0, x_1\}.$$

Da $x_0, x_1 \in (a, b)$ folgt

$$a < \min\{x_0, x_1\} \leq \gamma(t) \leq \max\{x_0, x_1\} < b.$$

Mit anderen Worten gilt $\gamma(t) \in (a, b)$.

Offene Teilmengen von $\mathbb{R}$, die keinen Intervallcharakter haben, sind im Allgemeinen nicht C^1 -wegzusammenhängend. In der Tat ist beispielsweise $(0, 1) \cup (1, 2) \subset \mathbb{R}$ nicht C^1 -wegzusammenhängend. In der Tat: Wir zeigen dass es keine stetig differenzierbare Kurve $\gamma : [0, 1] \rightarrow \mathbb{R}$ geben kann mit

- $\gamma(t) \in (0, 1) \cup (1, 2)$ für alle $t \in [0, 1]$
- $\gamma(0) = \frac{1}{2}, \gamma(1) = \frac{3}{2}$.

Denn nach dem eindimensionalen Zwischenwertsatz (Satz 6.2.15) müsste eine stetige Kurve $\gamma : [0, 1] \rightarrow \mathbb{R}$ mit $\gamma(0) = \frac{1}{2}$ und $\gamma(1) = \frac{3}{2}$ alle Werte zwischen $\frac{1}{2}$ und $\frac{3}{2}$ annehmen. Insbesondere gibt es ein $t_0 \in [0, 1]$ mit $\gamma(t_0) = 1$. Das widerspricht der Tatsache, dass $\gamma(t) \in (0, 1) \cup (1, 2)$ für alle $t \in [0, 1]$,

Man kann sogar zeigen, dass die (offenen) C^1 -wegzusammenhängenden Mengen in $\mathbb{R}^1$ genau die offenen Intervalle sein müssen.

Nun können wir einen Konstanzsatz für C^1 -wegzusammenhängende Definitionsbereiche formulieren.

Satz 10.5.38 — MEHRDIMENSIONALER KONSTANZSATZ

Es sei $D \subseteq \mathbb{R}^d$ offen und C^1 -wegzusammenhängend und $f : D \rightarrow \mathbb{R}^m$ auf D total differenzierbar. Dann gilt

$$Df(x) = 0 \text{ für alle } x \in D^a \quad \Longleftrightarrow \quad f \text{ ist auf } D \text{ konstant.}$$

^aHiermit ist gemeint, dass $Df(x)(v) = 0$ für alle $v \in \mathbb{R}^d$. Mit anderen Worten: $Df(x)$ ist die Abbildung, die jeden Vektor $v \in \mathbb{R}^d$ auf 0 schickt.

Beweis. Zuerst zu „ $\Leftarrow$ “. Es sei f auf D konstant, etwa gelte für $c \in \mathbb{R}^m$, dass $f(x) = c \forall x \in D$. Sei nun $x_0 \in D$ beliebig. Dann gilt für $x \in D \setminus \{0\}$

$$\frac{f(x) - f(x_0) - 0}{\|x - x_0\|} = \frac{c - c}{\|x - x_0\|} = 0.$$

Somit ist die Definition der totalen Differenzierbarkeit (Definition 10.5.23) mit $L_{x_0}(v) := 0, (v \in \mathbb{R}^d)$ erfüllt. Man folgert

$$Df(x_0)(v) = 0 \quad \forall v \in \mathbb{R}^d.$$

Da $x_0 \in D$ beliebig war folgt die Behauptung sofort.

Noch zu „ $\implies$ “. Es sei $Df(x) = 0$ für alle $x \in D$. Insbesondere gilt auch $J_f(x) = 0 \ \forall x \in D$.

$$\mathbb{Z} \quad f \text{ ist auf } D \text{ konstant, d.h. für } x_0, x_1 \in D \text{ gilt } f(x_0) = f(x_1).$$

Dazu: Sei $\gamma : [0, 1] \rightarrow \mathbb{R}^d$ eine stetig differenzierbare Kurve mit $\gamma(t) \in D$ für alle $t \in [0, 1]$ und $\gamma(0) = x_0, \gamma(1) = x_1$. Eine solche Kurve existiert, weil D ja C^1 -wegzusammenhängend ist (vgl. Definition 10.5.36). Beachte: $f \circ \gamma : [0, 1] \rightarrow \mathbb{R}^m$ ist eine Kurve. Nun ist γ als stetig (partiell) differenzierbare Funktion auf $(0, 1)$ auch total differenzierbar. Nach Satz 10.5.33 ist $f \circ \gamma$ auch total differenzierbar und es gilt für $t \in (0, 1)$

$$D(f \circ \gamma)(t) = Df(\gamma(t)) \circ D\gamma(t). \quad (613)$$

Da $\gamma(t) \in D$ ist $Df(\gamma(t))$ nach Voraussetzung die Nullabbildung. Deswegen schickt auch die Verkettung in (613) jeden Vektor auf den Nullvektor. Man erhält

$$D(f \circ \gamma)(t) = 0 \quad \forall t \in (0, 1).$$

Somit ist auch $J_{f \circ \gamma}(t) = 0$ für alle $t \in (0, 1)$. Mit (578) folgt

$$0 = J_{f \circ \gamma}(t) \stackrel{(578)}{=} \begin{pmatrix} [(f \circ \gamma)^{(1)}]'(t) \\ \vdots \\ [(f \circ \gamma)^{(m)}]'(t) \end{pmatrix}$$

Wir folgern, dass

$$[(f \circ \gamma)^{(1)}]'(t) = \dots = [(f \circ \gamma)^{(m)}]'(t) = 0 \quad \forall t \in (0, 1)$$

Nach dem eindimensionalen Konstanzsatz (Proposition 7.3.8) sind $(f \circ \gamma)^{(1)}, \dots, (f \circ \gamma)^{(m)}$ auf $[0, 1]$ konstant.^a Also gilt

$$(f \circ \gamma)^{(j)}(1) = (f \circ \gamma)^{(j)}(0) \quad \forall j = 1, \dots, m.$$

Mit anderen Worten gilt

$$(f \circ \gamma)(1) = (f \circ \gamma)(0).$$

Deswegen gilt

$$f(x_0) = f(\gamma(0)) = (f \circ \gamma)(0) = (f \circ \gamma)(1) = f(\gamma(1)) = f(x_1).$$

Da $x_0, x_1 \in D$ beliebig waren folgt die Behauptung.

^aHierfür müsste man eigentlich noch überprüfen, dass $(f \circ \gamma)^{(j)}$ für alle $j \in \{1, \dots, m\}$ stetig auf $[0, 1]$ ist. Dazu würde man argumentieren, dass $f \circ \gamma$ als Verkettung stetiger Funktionen wieder stetig ist und dann Bemerkung 10.4.10 verwenden.

*Mehrdimensionale Mittelwertsätze

Der Konstanzsatz wurde im Eindimensionalen über den ersten Mittelwertsatz der Differentialrechnung (Satz 7.3.1) bewiesen. In diesem Abschnitt wollen wir studieren, ob es zu diesem Mittelwertsatz mehrdimensionale Verallgemeinerungen gibt. Im Fall $m = 1$ findet man eine (unter zusätzlichen Voraussetzungen an das Gebiet) eine Verallgemeinerung.

Proposition 10.5.39 — EIN MITTELWERTSATZ IM MEHRDIMENSIONALEN

Es sei $D \subseteq \mathbb{R}^d$ offen und $f : D \rightarrow \mathbb{R}^a$ auf D total differenzierbar. Ferner seien $a, b \in D$ so, dass $\text{Seg}_{a,b} := \{a + t(b - a) | t \in [0, 1]\} \subseteq D$. Dann gibt es ein $\xi \in \text{Seg}_{a,b}$ so, dass

$$f(b) - f(a) = Df(\xi)(b - a).$$

^ad.h. $m = 1!$

Beweis. Definiere $h : [0, 1] \rightarrow \mathbb{R}$ durch

$$h(t) := f(a + t(b - a)).$$

Die Funktion h ist wohldefiniert, da $a + t(b - a) \in \text{Seg}_{a,b} \subseteq D$ für alle $t \in [0, 1]$. Nach Beispiel 10.5.35 ist h auf $(0, 1)$ differenzierbar und es gilt Mit dieser Information ergibt sich

$$h'(t) = \nabla f(a + t(b - a))^T(b - a) = J_f(\gamma(t))(b - a) = Df(\gamma(t))(b - a). \quad (614)$$

Mit dem eindimensionalen Mittelwertsatz finden wir ein $\xi_0 \in (0, 1)$ so, dass

$$h(1) - h(0) = h'(\xi_0) \cdot (1 - 0)$$

Mit der Definition von h und (614) folgern wir

$$f(b) - f(a) = h(1) - h(0) = h'(\xi_0) = Df(\gamma(\xi_0))(b - a)$$

Definiert man nun $\xi := \gamma(\xi_0) = a + \xi_0(b - a) \in \text{Seg}_{a,b}$ so folgt

$$f(b) - f(a) = Df(\xi)(b - a).$$

Auf die Voraussetzung „ $m = 1$ “ kann nicht verzichtet werden. Schon für $m = 2$ kann man keinen Mittelwertsatz mehr herleiten.

Beispiel 10.5.40

Es sei $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$ gegeben durch

$$\gamma(t) := \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

Behauptung. Es gibt kein $\xi \in \text{Seg}_{0,2\pi}$ ^a so, dass

$$\gamma(2\pi) - \gamma(0) = D\gamma(\xi)(2\pi - 0).$$

Beweis. Dazu beobachte, dass

$$\gamma(2\pi) = \begin{pmatrix} \cos(2\pi) \\ \sin(2\pi) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{und} \quad \gamma(0) = \begin{pmatrix} \cos(0) \\ \sin(0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Somit gilt

$$\gamma(2\pi) - \gamma(0) = 0 \tag{615}$$

Hingegen gilt für $\xi \in [0, 2\pi]$

$$D\gamma(\xi)(2\pi - 0) = \gamma'(\xi)(2\pi) \stackrel{\text{Bsp10.5.17}}{=} 2\pi \begin{pmatrix} -\sin(\xi) \\ \cos(\xi) \end{pmatrix} \tag{616}$$

Dies kann jedoch nicht der Nullvektor sein, denn

$$\left\| \begin{pmatrix} -\sin(\xi) \\ \cos(\xi) \end{pmatrix} \right\| = \sqrt{(-\sin(\xi))^2 + (\cos(\xi))^2} = \sqrt{\sin^2(\xi) + \cos^2(\xi)} = \sqrt{1} = 1 \neq 0.$$

Es zeigt sich, dass (615) und (616) für kein $\xi \in [0, 2\pi]$ übereinstimmen können. Die Behauptung folgt.

Selbstverständlich dürfte man den eindimensionalen Mittelwertsatz in jeder einzelnen Komponente anwenden. Die Zwischenwerte ξ , die man bei dieser Anwendung bekommt wären allerdings für jede Komponente verschiedene! Daher enthält obiges Beispiel keinen Widerspruch zum eindimensionalen Mittelwertsatz (Satz 7.3.1).

^aBeachte, dass $\text{Seg}_{0,2\pi} = [0, 2\pi]$!

10.6 Höhere Ableitungen

Nun, nachdem wir die verschiedenen Konzepte der Differenzierbarkeit besprochen haben, können wir uns fragen, wie man *höhere Ableitungen* definieren kann. Wir beschränken uns in diesem Abschnitt auf das Konzept der partiellen Differenzierbarkeit. Das liegt nicht zuletzt an Satz 10.5.30, der besagt, dass stetige partielle Differenzierbarkeit auch automatisch totale Differenzierbarkeit impliziert.

Definition 10.6.1 — HÖHERE PARTIELLE ABLEITUNGEN

Es sei $D \subseteq \mathbb{R}^d$ offen und $f : D \rightarrow \mathbb{R}^m$ eine Funktion. Ferner sei $k \in \mathbb{N}$.

- (i) f heißt bei $x_0 \in D$ *zweimal partiell differenzierbar* falls f auf ganz D einmal partiell differenzierbar ist und alle partiellen Ableitungen $\partial_1 f, \dots, \partial_d f$ bei x_0 wiederum partiell differenzierbar sind. Wir bezeichnen

$$\partial_j \partial_i f(x_0) := \partial_j (\partial_i f)(x_0)$$

- (ii) Rekursiv bezeichnen wir f bei $x_0 \in D$ als $(k+1)$ -mal partiell differenzierbar, wenn alle partiellen Ableitungen $\partial_{i_1} \partial_{i_2} \dots \partial_{i_k} f$ für beliebige $i_1, \dots, i_k \in \{1, \dots, n\}$ auf ganz D existieren und in x_0 ein weiteres Mal partiell differenzierbar sind.
- (iii) Wir nennen f auf D *k -mal partiell differenzierbar*, falls f bei jedem $x_0 \in D$ k -mal partiell differenzierbar ist.
- (iv) Wir nennen f auf D *k -mal stetig partiell differenzierbar*, wenn f auf D k -mal partiell differenzierbar ist und alle partiellen Ableitungen stetig sind

Der Satz von Schwarz

Wenn man nach Definition 10.6.1 (i) vorgeht würde die Reihenfolge der partiellen Ableitungen eine Rolle spielen. Gemäß Definition 10.6.1 (i) sind nämlich $\partial_i \partial_j f(x_0)$ und $\partial_j \partial_i f(x_0)$ verschiedene Ausdrücke. Das heißt insbesondere, dass man sich auch merken muss, nach welcher der beiden Variablen man zuerst ableitet. Allerdings gibt es den *Satz von Schwarz*, der besagt, dass man sich bei zweimal stetig partiell differenzierbaren Funktionen über die Reihenfolge keine Gedanken machen muss.

Satz 10.6.2 — SATZ VON SCHWARZ

Es sei $D \subseteq \mathbb{R}^d$ offen, und $f : D \rightarrow \mathbb{R}^m$ eine auf D zweimal partiell differenzierbare Funktion. Ferner sei $x_0 \in D$. Sind für $i, j \in \{1, \dots, d\}$ die partiellen Ableitungen $\partial_i \partial_j f$ und $\partial_j \partial_i f$ in x_0 stetig, so gilt

$$\partial_j \partial_i f(x_0) = \partial_i \partial_j f(x_0).$$

Beweis. Wir führen den Beweis zuerst im Spezialfall $m = 1$ und $d = 2$. Da nur der Fall $i \neq j$ interessant ist, darf man dann auch $i = 1$ und $j = 2$ annehmen. OBdA kann man auch $x_0 = 0$ annehmen (denn falls $x_0 \neq 0$ ist, kann man stattdessen

$$\tilde{f} : \{v \in \mathbb{R}^d : v + x_0 \in D\} \rightarrow \mathbb{R}, x \mapsto f(x + x_0)$$

betrachten).

$$\mathbb{Z} : \quad \partial_2 \partial_1 f(0, 0) = \partial_1 \partial_2 f(0, 0).$$

Da D offen ist und $0 \in D$ gibt es ein $r > 0$ so, dass $B_r(0) \subseteq D$. Dann ist für $r' := \frac{r}{\sqrt{2}}$ auch

$$(-r', r') \times (-r', r') \subseteq D.^a$$

ZWISCHENBEHAUPTUNG. Für alle $x, y \in (-r', r') \setminus \{0\}$ gilt

$$\exists \xi^{(1)} \in (-|x|, |x|), \xi^{(2)} \in (-|y|, |y|) : f(x, y) - f(0, y) - f(x, 0) + f(0, 0) = xy \partial_1 \partial_2 f(\xi^{(1)}, \xi^{(2)}) \quad (617)$$

und

$$\exists \eta^{(1)} \in (-|x|, |x|), \eta^{(2)} \in (-|y|, |y|) : f(x, y) - f(0, y) - f(x, 0) + f(0, 0) = xy \partial_2 \partial_1 f(\eta^{(1)}, \eta^{(2)}). \quad (618)$$

Zuerst zu (617). Seien $x, y \in (-r', r') \setminus \{0\}$. Definiere nun die Hilfsfunktion $g : (-r', r') \rightarrow \mathbb{R}$ durch

$$g(t) := f(x, t) - f(0, t) \quad (t \in (-r', r')).$$

Dann gilt

$$\begin{aligned} f(x, y) - f(0, y) - f(x, 0) + f(0, 0) &= f(x, y) - f(0, y) - (f(x, 0) - f(0, 0)) \\ &= g(y) - g(0). \end{aligned}$$

Nach dem ersten Mittelwertsatz der Differentialrechnung (Satz 7.3.1) gibt es ein $\xi^{(2)} \in (\min\{y, 0\}, \max\{y, 0\}) \subseteq (-|y|, |y|)$ so, dass

$$g(y) - g(0) = g'(\xi^{(2)})(y - 0) = g'(\xi^{(2)})y.$$

Da nach Bemerkung 10.5.12 gilt $g'(t) = \partial_2 f(x, t) - \partial_2 f(0, t)^b$ erhalten wir

$$f(x, y) - f(0, y) - f(x, 0) + f(0, 0) = g(y) - g(0) = g'(\xi^{(2)})y = y(\partial_2 f(x, \xi^{(2)}) - \partial_2 f(0, \xi^{(2)})) \quad (619)$$

Nun definiere $h : (-r', r') \rightarrow \mathbb{R}$ durch

$$h(t) := \partial_2 f(t, \xi^{(2)}) \quad t \in (-r', r').$$

Aus (619) erhalten wir, dass

$$f(x, y) - f(0, y) - f(x, 0) + f(0, 0) = y(h(x) - h(0)) \quad (620)$$

Wiederum nach dem ersten Mittelwertsatz (Satz 7.3.1) gibt es ein $\xi^{(1)} \in (\min\{x, 0\}, \max\{x, 0\}) \subseteq (-|x|, |x|)$ so, dass

$$h(x) - h(0) = h'(\xi^{(1)})(x - 0) = h'(\xi^{(1)})x$$

Nach Bemerkung 10.5.12 gilt wieder $h'(t) = \partial_1(\partial_2 f)(t, \xi^{(2)}) = \partial_1 \partial_2 f(t, \xi^{(2)})$. Setzen wir dies in (620) ein, so erhalten wir

$$f(x, y) - f(x, 0) - f(0, y) + f(0, 0) = y(h'(\xi^{(1)})x) = xy\partial_1\partial_2 f(\xi^{(1)}, \xi^{(2)}).$$

Es folgt (617). Gleichung (618) kann analog mit vertauschten Rollen von x und y hergeleitet werden.

ZWISCHENBEHAUPTUNG ENDE.

Aus der Zwischenbehauptung lernen wir, dass es für alle $x, y \in (-r', r')$ Zahlen $\xi^{(1)}, \eta^{(1)} \in (-|x|, |x|)$ und $\xi^{(2)}, \eta^{(2)} \in (-|y|, |y|)$ gibt mit

$$xy\partial_2\partial_1 f(\eta^{(1)}, \eta^{(2)}) = f(x, y) - f(x, 0) - f(0, y) + f(0, 0) = xy\partial_1\partial_2 f(\xi^{(1)}, \xi^{(2)}).$$

Betrachte nun für $n \in \mathbb{N}$ beliebig mit $\frac{1}{n} < r'$ den Fall $x = y = \frac{1}{n}$. Wir lernen: Es gibt $\xi_n^{(1)}, \xi_n^{(2)}, \eta_n^{(1)}, \eta_n^{(2)} \in (-\frac{1}{n}, \frac{1}{n})$ (*) so, dass

$$\frac{1}{n^2}\partial_2\partial_1 f(\eta_n^{(1)}, \eta_n^{(2)}) = \frac{1}{n^2}\partial_1\partial_2 f(\xi_n^{(1)}, \xi_n^{(2)}).$$

Multiplizieren wir mit n^2 so ergibt sich

$$\partial_2\partial_1 f(\eta_n^{(1)}, \eta_n^{(2)}) = \partial_1\partial_2 f(\xi_n^{(1)}, \xi_n^{(2)}).$$

Als nächstes lassen wir $n \rightarrow \infty$ streben. Beachte: Wegen (*) gilt $(\xi_n^{(1)}, \xi_n^{(2)})^T \xrightarrow{n \rightarrow \infty} (0, 0)^T$ und $(\eta_n^{(1)}, \eta_n^{(2)})^T \xrightarrow{n \rightarrow \infty} (0, 0)^T$. Wegen der Stetigkeit der beiden zweiten partiellen Ableitungen in $(0, 0)^T$ folgt

$$\partial_2\partial_1 f(0, 0) = \lim_{n \rightarrow \infty} \partial_2\partial_1 f(\eta_n^{(1)}, \eta_n^{(2)}) = \lim_{n \rightarrow \infty} \partial_1\partial_2 f(\xi_n^{(1)}, \xi_n^{(2)}) = \partial_1\partial_2 f(0, 0).$$

Das war zu zeigen!

Nun müssen wir noch klären, was in den Fällen $m > 1$ und $d > 2$ passiert. In diesen Fällen dürfen wir wiederum OBdA $x_0 = 0$ annehmen (siehe oben). Ist $d > 2$ (aber noch $m = 1$) und $i, j \in \{1, \dots, d\}, i \neq j$ so kann man die zweidimensionale Funktion

$$\tilde{f}: \hat{D} \rightarrow \mathbb{R} \quad \tilde{f}(s, t) := f(0, \dots, 0, \underbrace{s}_{i\text{-te Stelle}}, 0, \dots, 0, \underbrace{t}_{j\text{-te Stelle}}, 0, \dots, 0) = f(se_i + te_j)$$

betrachten. Hierbei ist $\hat{D} := \{(s, t) \in \mathbb{R}^2 : se_i + te_j \in D\}$. Man überzeugt sich leicht: $\hat{D}$ ist offen und (vgl. auch Bemerkung 10.5.12)

$$\partial_i\partial_j f(0) = \partial_1\partial_2 \tilde{f}(0, 0) \quad \text{bzw.} \quad \partial_j\partial_i f(0) = \partial_2\partial_1 \tilde{f}(0, 0)$$

Ferner sieht man leicht dass $\partial_1\partial_2\tilde{f}, \partial_2\partial_1\tilde{f}$ bei 0 stetig sind und dann folgt aus dem Fall $d = 2, m = 1$

$$\partial_i\partial_j f(0) = \partial_1\partial_2\tilde{f}(0, 0) = \partial_2\partial_1\tilde{f}(0, 0) = \partial_j\partial_i f(0).$$

Die Behauptung folgt so auch für $d > 2$ und $m = 1$. Für $m > 1$ zeigen wir, dass bei allen Komponentenfunktionen $f^{(1)}, \dots, f^{(m)}$ die partiellen Ableitungen vertauscht werden können. Dazu: Wegen Bemerkung 10.4.10 und der Voraussetzung sind $\partial_i\partial_j f^{(1)}, \dots, \partial_i\partial_j f^{(m)}$ und $\partial_j\partial_i f^{(1)}, \dots, \partial_j\partial_i f^{(m)}$ bei 0 stetig. Daher können wir aus dem Fall $m = 1$ folgern, dass

$$\partial_i\partial_j f^{(\ell)}(0) = \partial_j\partial_i f^{(\ell)}(0) \quad \forall \ell = 1, \dots, m$$

Die Behauptung folgt!

^aDenn für $(x, y)^T \in (-r', r') \times (-r', r')$ gilt $\|(x, y)^T\| = \sqrt{x^2 + y^2} < \sqrt{(r')^2 + (r')^2} = \sqrt{2(r')^2} = \sqrt{2}r'$.

^bEigentlich müsste man die Bemerkung 10.5.12 schon vorher benutzen: Für die Anwendung des ersten Mittelwertsatzes müsste man nämlich begründen, warum g auf $(\min\{y, 0\}, \max\{y, 0\})$ differenzierbar ist. Das liegt gerade an der partiellen Differenzierbarkeit von f und daran und an Bemerkung 10.5.12.

FAZIT. Ist eine Funktion zweimal stetig (!) partiell differenzierbar, so kann die Reihenfolge der Ableitungen vertauscht werden.

Auf die Stetigkeit der zweiten Ableitungen kann bei dem obigen Vertauschungsergebnis nicht verzichtet werden

Beispiel 10.6.3

Es sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ gegeben durch

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

Wir werden sehen, dass f auf $\mathbb{R}^2$ zweimal differenzierbar ist, aber

$$\partial_x\partial_y f(0, 0) \neq \partial_y\partial_x f(0, 0).$$

Das kann einzig und allein daran liegen, dass die obigen partiellen Ableitungen bei $(0, 0)^T$ nicht stetig sind (vgl. Satz 10.6.2).

Dazu: Für $(x, y) \neq (0, 0)$ berechnet man leicht die partiellen Ableitungen

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

$$= y \frac{x^2 - y^2}{x^2 + y^2} + \frac{4x^2 y^3}{(x^2 + y^2)^2} = y \left(\frac{x^2 - y^2}{x^2 + y^2} + \frac{4x^2 y^2}{(x^2 + y^2)^2} \right),$$

$$\begin{aligned} \partial_y f(x, y) &= x \frac{x^2 - y^2}{x^2 + y^2} + xy \left(\frac{(-2y)}{x^2 + y^2} - \frac{2y(x^2 - y^2)}{(x^2 + y^2)^2} \right) \\ &= x \frac{x^2 - y^2}{x^2 + y^2} - 2x^2 y \left(\frac{1}{x^2 + y^2} + \frac{x^2 - y^2}{(x^2 + y^2)^2} \right) \\ &= x \frac{x^2 - y^2}{x^2 + y^2} - 2x^2 y \left(\frac{(x^2 + y^2) + (x^2 - y^2)}{(x^2 + y^2)^2} \right) \\ &= x \frac{x^2 - y^2}{x^2 + y^2} - \frac{4x^3 y^2}{(x^2 + y^2)^2} = x \left(\frac{x^2 - y^2}{x^2 + y^2} - \frac{4x^2 y^2}{(x^2 + y^2)^2} \right). \end{aligned}$$

Beachten Sie außerdem dass $f(t, 0) = f(0, t) = 0$ für alle $t \in \mathbb{R}$. Deswegen gilt

$$\partial_x f(0, 0) = \lim_{t \rightarrow 0} \frac{f(te_1) - f(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{f(t, 0) - f(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{0 - 0}{t} = \lim_{t \rightarrow 0} 0 = 0.$$

Analog zeigt man $\partial_y f(0, 0) = 0$. Wir erhalten

$$\partial_x f(x, y) = \begin{cases} y \left(\frac{x^2 - y^2}{x^2 + y^2} + \frac{4x^2 y^2}{(x^2 + y^2)^2} \right) & \text{falls } (x, y)^T \neq (0, 0)^T \\ 0 & \text{falls } (x, y)^T = (0, 0)^T \end{cases}$$

und

$$\partial_y f(x, y) = \begin{cases} x \left(\frac{x^2 - y^2}{x^2 + y^2} - \frac{4x^2 y^2}{(x^2 + y^2)^2} \right) & \text{falls } (x, y)^T \neq (0, 0)^T \\ 0 & \text{falls } (x, y)^T = (0, 0)^T. \end{cases}$$

Beachte: f ist zweimal partiell differenzierbar. Für $(x, y) \neq (0, 0)$ ist das klar. Für $(x, y) = (0, 0)$ berechnen wir

$$\partial_x \partial_x f(0, 0) = \lim_{t \rightarrow 0} \frac{\partial_x f(t, 0) - \partial_x f(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{0 - 0}{t} = 0$$

und analog

$$\partial_y \partial_y f(0, 0) = 0.$$

Für die *gemischten Ableitungen* gilt nun

$$\partial_x \partial_y f(0, 0) = \lim_{t \rightarrow 0} \frac{\partial_y f(t, 0) - \partial_y f(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{t \left(\frac{t^2 - 0^2}{t^2 + 0^2} + \frac{4 \cdot t^2 \cdot 0^2}{(t^2 + 0^2)^2} \right)}{t} = \lim_{t \rightarrow 0} \frac{t^2 - 0^2}{t^2 + 0^2} = \lim_{t \rightarrow 0} 1 = 1$$

und

$$\partial_y \partial_x f(0, 0) = \lim_{t \rightarrow 0} \frac{\partial_x f(0, t) - \partial_x f(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{t \left(\frac{0^2 - t^2}{0^2 + t^2} + \frac{4 \cdot 0^2 \cdot t^2}{(0^2 + t^2)^2} \right)}{t} = \lim_{t \rightarrow 0} \frac{0^2 - t^2}{0^2 + t^2} = -1.$$

Es folgt

$$\partial_x \partial_y f(0, 0) = 1 \neq -1 = \partial_y \partial_x f(0, 0).$$

Zum Abschluss bemerken wir noch, dass der Satz von Schwarz bedingt, dass die Reihenfolge höherer Ableitungen bei genügend oft stetig partiell differenzierbaren Funktionen beliebig vertauscht werden kann

Beispiel 10.6.4

Es sei $D \subseteq \mathbb{R}^d$ und $f : D \rightarrow \mathbb{R}$ nun dreimal stetig partiell differenzierbar. Ferner seien $i_1, i_2, i_3 \in \{1, \dots, d\}$.

Behauptung. Für alle $x_0 \in D$ gilt

$$\begin{aligned} \partial_{i_1} \partial_{i_2} \partial_{i_3} f(x_0) &\stackrel{(*)}{=} \partial_{i_2} \partial_{i_1} \partial_{i_3} f(x_0) = \partial_{i_1} \partial_{i_3} \partial_{i_2} f(x_0) \\ &= \partial_{i_3} \partial_{i_1} \partial_{i_2} f(x_0) \stackrel{(**)}{=} \partial_{i_3} \partial_{i_2} \partial_{i_1} f(x_0) = \partial_{i_2} \partial_{i_3} \partial_{i_1} f(x_0). \end{aligned}$$

Beweis. Alle Vertauschungen können mit dem Satz von Schwarz (bzw dessen iterativer Anwendung) begründet werden. Wir begründen beispielhaft die Gleichheiten in $(*)$ und $(**)$. Die Gleichheit $(*)$ folgt zum Beispiel so: Da f dreimal stetig partiell diffbar ist, ist $\partial_{i_3} f$ zweimal stetig partiell diffbar. Wir folgern aus dem Satz von Schwarz, dass

$$\partial_{i_1} \partial_{i_2} (\partial_{i_3} f)(x_0) = \partial_{i_2} \partial_{i_1} (\partial_{i_3} f)(x_0)$$

Zusammen mit Definition 10.6.1 folgt $(*)$. Für $(**)$ würde man zum Beispiel so argumentieren: Weil f zweimal stetig partiell diffbar ist gilt $\partial_{i_1} \partial_{i_2} f(x) = \partial_{i_2} \partial_{i_1} f(x)$ für alle $x \in D$. Deswegen gilt

$$\partial_{i_3} (\partial_{i_1} \partial_{i_2} f)(x_0) = \partial_{i_3} (\partial_{i_2} \partial_{i_1} f)(x_0).$$

Zusammen mit Definition 10.6.1 folgt wiederum $(**)$.

Für öfter differenzierbare Funktionen kann man analog weitere Vertauschungsergebnisse zeigen.

Wir können diese Vertauschung oft geschickt anwenden um Ableitungen effizient zu berechnen. Ein Beispiel dazu: Gegeben sei

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = xe^{\sqrt{y^2 + \arctan(17y) + \sin(\exp(y))}}$$

Wir wollen $\partial_x \partial_x \partial_y f$ berechnen. Dazu verwenden wir den Satz von Schwarz!

$$\begin{aligned} \partial_x \partial_x \partial_y f(x, y) &= \partial_y \partial_x \partial_x f(x, y) = \partial_y \partial_x \partial_x \left(xe^{\sqrt{y^2 + \arctan(17y) + \sin(\exp(y))}} \right) \\ &= \partial_y \partial_x \left(e^{\sqrt{y^2 + \arctan(17y) + \sin(\exp(y))}} \right) = \partial_y 0 = 0. \end{aligned}$$

Wer erst nach y differenziert hätte, hätte an dieser Stelle wesentlich mehr Arbeit gehabt!

Hesse-Matrix

Bei einer skalarwertigen zweimal partiell differenzierbaren Funktion $f : D \rightarrow \mathbb{R}$ fasst man die zweiten partiellen Ableitungen oft in der sogenannten *Hesse-Matrix* zusammen

Definition 10.6.5 — HESSE-MATRIX

Es sei $D \subseteq \mathbb{R}^d$ offen und $f : D \rightarrow \mathbb{R}$ bei $x_0 \in D$ zweimal partiell differenzierbar. Dann definieren wir die sogenannte *Hesse-Matrix* von f an der Stelle x_0 durch

$$H_f(x_0) := \begin{pmatrix} \partial_1 \partial_1 f(x_0) & \cdots & \partial_d \partial_1 f(x_0) \\ \vdots & \ddots & \vdots \\ \partial_1 \partial_d f(x_0) & \cdots & \partial_d \partial_d f(x_0) \end{pmatrix} \in \mathbb{R}^{d \times d}$$

Genauer: $H_f(x_0)$ ist eine $d \times d$ -Matrix mit Eintrag $\partial_j \partial_i f(x_0)$ in der i -ten Zeile und j -ten Spalte.

Die Hesse-Matrix ist ein mehrdimensionales Analogon zur zweiten Ableitung. Wie bereits in der eindimensionalen Analysis wird diese zweite Ableitung eine wichtige Rolle bei der Klassifikation von lokalen Extrema (lokale Maxima / Minima) hilfreich sein.

Bemerkung 10.6.6

Ist $f : D \rightarrow \mathbb{R}$ auf D sogar zweimal stetig partiell differenzierbar, so ist nach dem Satz von Schwarz (Satz 10.6.2) $H_f(x_0)$ eine *symmetrische* Matrix, d.h.

$$H_f(x_0)^T = H_f(x_0)$$

In der Tat: Der Eintrag in der i -ten Zeile und j -ten Spalte ($\partial_j \partial_i f(x_0)$) stimmt dann mit dem Eintrag in der j -ten Spalte und i -ten Zeile ($\partial_i \partial_j f(x_0)$) überein.

An dieser Stelle rechnen wir nochmal ein Beispiel, was die Rolle der Hesse-Matrix als zweite Ableitung illustrieren soll.

Beispiel 10.6.7

Es sei $D \subseteq \mathbb{R}^d$ eine offene Menge und $f : D \rightarrow \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion. Für $x, x_0 \in D$ so, dass $\text{Seg}_{x,x_0} := \{x_0 + t(x - x_0) : t \in [0, 1]\} \subseteq D$ betrachte $h : [0, 1] \rightarrow \mathbb{R}$

$$h(t) := f(x_0 + t(x - x_0)).$$

Behauptung. h ist auf $[0, 1]$ zweimal stetig differenzierbar. Ferner gilt

$$(i) \quad h'(t) = \nabla f(x_0 + t(x - x_0))^T(x - x_0) \text{ und}$$

$$(ii) \quad h''(t) = (x - x_0)^T H_f(x_0 + t(x - x_0))(x - x_0)$$

Insbesondere gilt $h'(0) = \nabla f(x_0)^T(x - x_0)$ und $h''(0) = (x - x_0)^T H_f(x_0)(x - x_0)$. Wir werden sogar zeigen, dass die Definition 10.6.7 auch für $t \in (a, b)$ für ein offenes Intervall $(a, b) \supseteq [0, 1]$ zulässig ist und die Ableitungsformeln in (i),(ii) sogar auf (a, b) gelten.

Beweis. Zu (i). Beachte zunächst: Die Situation ist leicht anders als in Beispiel 10.5.35. Dort behaupten wir nämlich nur Differenzierbarkeit auf $(0, 1)$ und nicht auf dem Abschluss $[0, 1]$. Um die Differenzierbarkeit sogar bis zum Abschluss zu zeigen gehen wir wie folgt vor: Definiere $\gamma : \mathbb{R} \rightarrow \mathbb{R}^d$ durch $\gamma(t) := x_0 + t(x - x_0)$. Da γ stetig ist und D offen ist, ist nach Satz 10.4.14 $\gamma^{-1}(D) \subseteq \mathbb{R}$ wiederum offen. Da $\gamma([0, 1]) = \text{Seg}_{x,x_0} \subseteq D$ gilt aber auch $[0, 1] \subseteq \gamma^{-1}(D)$. Wegen der Offenheit von $\gamma^{-1}(D)$ gibt es dann auch ein $a < 0$ und $b > 1$ so, dass $(a, b) \subseteq \gamma^{-1}(D)$.^a Somit kann $f \circ \gamma$ auf ganz (a, b) definiert werden und klarerweise stimmt h auf $[0, 1] \subset (a, b)$ mit $f \circ \gamma$ überein. Man kann nun wie in Beispiel 10.5.35 vorgehen und $f \circ \gamma$ auf (a, b) differenzieren mit derselben Ableitungsformel aus Beispiel 10.5.35, d.h.

$$\frac{d}{dt}(f \circ \gamma)(t) = \nabla f(x_0 + t(x - x_0))^T(x - x_0) \quad \forall t \in (a, b).$$

Schränkt man sich dann wieder auf $[0, 1]$ ein so folgt die Behauptung.

Zu (ii). Es sei das Intervall (a, b) konstruiert wie in (i). Schreiben wir das obige Matrixprodukt in der Formel aus (i) aus, so haben wir für $t \in (a, b)$

$$h'(t) = \nabla f(x_0 + t(x - x_0))^T(x - x_0) = \sum_{i=1}^d \partial_i f(x_0 + t(x - x_0))(x - x_0)^{(i)}. \quad (621)$$

Definieren wir für $t \in (a, b)$

$$g_i(t) := \partial_i f(x_0 + t(x - x_0)) \quad (622)$$

so haben wir mit (621)

$$h'(t) := \sum_{i=1}^d g_i(t)(x - x_0)^{(i)} \quad \forall t \in (a, b).$$

Daher gilt auch

$$h''(t) = \sum_{i=1}^d g'_i(t)(x - x_0)^{(i)} \quad \forall t \in (a, b). \quad (623)$$

Nun kann g'_i mithilfe von (622) wie in Punkt (i) (mit $\partial_i f$ statt f)^b berechnet werden. Man hat nach (i)

$$g'_i(t) = \nabla(\partial_i f)(x_0 + t(x - x_0))^T (x - x_0) = \sum_{j=1}^d \partial_j(\partial_i f)(x_0 + t(x - x_0))(x - x_0)^{(j)}.$$

Setzen wir dies in (623) ein so erhalten wir

$$\begin{aligned} h''(t) &= \sum_{i=1}^d \left(\sum_{j=1}^d \partial_j(\partial_i f)(x_0 + t(x - x_0))(x - x_0)^{(j)} \right) (x - x_0)^{(i)} \\ &= \sum_{i=1}^d \sum_{j=1}^d (x - x_0)^{(i)} \partial_j \partial_i f(x_0 + t(x - x_0))(x - x_0)^{(j)}. \end{aligned} \quad (624)$$

Wir müssen uns noch davon überzeugen, dass dieser Ausdruck mit $(x - x_0)^T H_f(x_0 + t(x - x_0))(x - x_0)$ übereinstimmt. Dazu berechnen wir für eine beliebige $d \times d$ -Matrix $A = (a_{ij})_{i,j=1,\dots,d}^{i=1,\dots,d}$ und $v \in \mathbb{R}^d$

$$v^T A v \stackrel{(583)}{=} v^T \left(\sum_{i=1}^d \sum_{j=1}^d a_{ij} v^{(j)} e_i \right) = \sum_{i=1}^d \sum_{j=1}^d a_{ij} v^{(j)} (v^T e_i) \quad (625)$$

Nun liefert $v^T e_i$ nach (584) gerade die i -te Spalte von v^T , welche nur aus $v^{(i)}$ besteht. Wir erhalten

$$v^T A v = \sum_{i=1}^d \sum_{j=1}^d a_{ij} v^{(j)} v^{(i)} = \sum_{i=1}^d \sum_{j=1}^d v^{(i)} a_{ij} v^{(j)}. \quad (626)$$

Mit $v = (x - x_0)$ und $A = H_f(x_0)$ [d.h. insbesondere $a_{ij} = \partial_j \partial_i f(x_0 + t(x - x_0))$], vgl. Definition 10.6.5] gilt

$$(x - x_0)^T H_f(x_0 + t(x - x_0))(x - x_0) = \sum_{i=1}^d \sum_{j=1}^d (x - x_0)^{(i)} \partial_j \partial_i f(x_0 + t(x - x_0))(x - x_0)^{(j)}.$$

Die rechte Seite stimmt gerade mit der Formel für $h''(t)$ aus (624) überein. Wir folgern also

$$h''(t) = (x - x_0)^T H_f(x_0 + t(x - x_0))(x - x_0).$$

^aIn der Tat: Ist $U \subseteq \mathbb{R}$ offen mit $[0, 1] \subseteq U$ so haben wir

$$0 \in U \xRightarrow[\text{Def10.3.5}]{U \text{ offen}} 0 \in \overset{\circ}{U} \xRightarrow{\text{Def10.3.1}} \exists r_1 > 0 : B_{r_1}(0) \subseteq U \implies \exists r_1 > 0 : (-r_1, r_1) \subseteq U.$$

Analog zeigt man

$$1 \in U \implies \exists r_2 > 0 : (1 - r_2, 1 + r_2) \subseteq U$$

Man erhält $(-r_1, r_1) \cup [0, 1] \cup (1 - r_2, 1 + r_2) \subseteq U$. Definiert man nun $a := -r_1 < 0$ und $b := 1 + r_2 > 1$ so folgt klarerweise $(a, b) \subseteq U$.

^bBei der Herleitung von (i) muss man an einem Punkt begründen, dass $\partial_i f$ total differenzierbar ist. Das folgt wiederum aus der angenommenen stetigen(!) partiellen Differenzierbarkeit von $\partial_i f$ und Satz 10.5.30. Man sieht, dass für die Aussage wirklich angenommen werden muss, dass f zweimal *stetig* partiell diffbar ist!

Wir haben gesehen, dass sich die zweite Ableitung mithilfe der Hesse-Matrix darstellen lässt. Ein Spezialfall: Setzt man $x := x_0 + v$ und betrachtet h wie im vorigen Beispiel, so ist

$$h(t) = f(x_0 + t(x - x_0)) = f(x_0 + tv) = \gamma_v(t)$$

gerade die *v-Sektorfunktion* von f , vgl. Definition 10.5.1. Dann gilt insbesondere nach dem vorigen Beispiel

$$\gamma_v''(0) = h''(0) = v^T H_f(x_0) v. \quad (627)$$

Den Ausdruck auf der rechten Seite nennen wir in Zukunft *quadratische Form*. Er wird später in der Vorlesung noch eine Rolle spielen

Definition 10.6.8 — QUADRATISCHE FORM

Es sei $A \in \mathbb{R}^{d \times d}$ eine $d \times d$ -Matrix. Dann definieren wir die *quadratische Form* $q_A : \mathbb{R}^d \rightarrow \mathbb{R}$ von A durch

$$q_A(v) := v^T A v.$$

Beachte: Jede quadratische Form ist eine Art „*Polynom zweiten Grades*“ in v . In der Tat: Nach (625) und (626) hat man

$$q_A(v) = v^T A v = \sum_{i=1}^d \sum_{j=1}^d a_{ij} v^{(i)} v^{(j)}. \quad (628)$$

Man sieht: Alle Summanden enthalten genau zwei Einträge von v , die miteinander multipliziert werden. Daher kommt der Begriff *Polynom zweiten Grades*. Insbesondere wollen wir an dieser Stelle bemerken, dass q_A auf $\mathbb{R}^d$ stetig ist (was direkt aus (628) folgt)

01.07.2026

Der Mehrdimensionale Satz von Taylor

In einer Dimension haben wir eine k -mal differenzierbare Funktion in einer Umgebung eines Punktes x_0 gut durch ein Polynom k -ten Grades approximieren können. Dieses Polynom nannten wir *Taylorpolynom* k -ten Grades (vgl. Definition 7.4.8).

Eine solche Taylorapproximation können wir auch für mehrdimensionale Funktionen erhalten, indem wir die Taylorentwicklung auf die jeweiligen Sektorfunktionen anwenden. Wir beschränken uns im Folgenden auf den Fall $k = 2$, da dieser für die Anwendungen am relevantesten ist.

Satz 10.6.9 — MEHRDIMENSIONALE TAYLORFORMEL

Es sei $D \subseteq \mathbb{R}^d$ offen und $x, x_0 \in D$ so, dass $\text{Seg}_{x, x_0} = \{x_0 + t(x - x_0) : t \in [0, 1]\} \subseteq D$. Ferner sei $f : D \rightarrow \mathbb{R}$ zweimal stetig differenzierbar. Dann gibt es ein $\xi \in (0, 1)$ so, dass

$$f(x) = f(x_0) + \nabla f(x_0)^T(x - x_0) + \frac{1}{2}(x - x_0)^T H_f(x_0 + \xi(x - x_0))(x - x_0).$$

Beweis. Wie in Beispiel 10.6.7 lässt sich ein offenes Intervall $(a, b) \supseteq [0, 1]$ finden, sodass die $x_0 + t(x - x_0) \in D$ für alle $t \in (a, b)$. Basierend darauf definieren wir wie in Beispiel 10.6.7

$$h : (a, b) \rightarrow \mathbb{R}, \quad h(t) := f(x_0 + t(x - x_0)).$$

Nach Beispiel 10.6.7 ist h auf (a, b) zweimal differenzierbar. Deswegen besagt die eindimensionale Taylorformel (Satz 7.4.6), dass es ein $\xi \in (\min\{0, 1\}, \max\{0, 1\}) = (0, 1)$ gibt mit

$$h(1) = h(0) + h'(0)(1 - 0) + \frac{1}{2}h''(\xi)(1 - 0)^2 = h(0) + h'(0) + \frac{1}{2}h''(\xi). \quad (629)$$

Nun hat man $h(1) = f(x)$, $h(0) = f(x_0)$ und nach Beispiel 10.6.7

- $h'(0) = \nabla f(x_0)^T(x - x_0)$,
- $h''(\xi) = (x - x_0)^T H_f(x_0 + \xi(x - x_0))(x - x_0)$.

Setzt man das alles in (629) ein, so erhält man

$$f(x) = f(x_0) + \nabla f(x_0)^T(x - x_0) + \frac{1}{2}(x - x_0)^T H_f(x_0 + \xi(x - x_0))(x - x_0).$$

Die Behauptung ist gezeigt.

Als Korollar fällt wie schon im Eindimensionalen eine *Taylor-Approximation* mit quadratischem Fehler raus.

Korollar 10.6.10 — TAYLOR-APPROXIMATION

Es sei $D \subseteq \mathbb{R}^d$ offen und $r > 0$ so, dass $B_r(x_0) \subseteq D$.^a Ferner sei $f : \mathbb{R}^d \rightarrow \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion. Dann gibt es eine im Punkt x_0 stetige Funktion

$\phi : B_r(x_0) \rightarrow \mathbb{R}$ mit $\phi(x_0) = 0$ so, dass

$$f(x) = f(x_0) + \nabla f(x_0)^T(x - x_0) + (x - x_0)^T H_f(x_0)(x - x_0) + \|x - x_0\|^2 \phi(x) \quad \forall x \in B_r(x_0).$$

^aFür die Existenz eines solchen $r > 0$ verweisen wir auf Bemerkung 10.3.6.

Beweis. Es sei $f : D \rightarrow \mathbb{R}$ wie in der Aussage und $r > 0$ so, dass $B_r(x_0) \subseteq D$. Beachte: Für alle $x \in B_r(x_0)$ gilt dann $\text{Seg}_{x, x_0} = \{x_0 + t(x - x_0) : t \in [0, 1]\} \subseteq B_r(x_0) \subseteq D$.^a Man erhält dann nach Satz 10.6.9 für jedes $x \in B_r(x_0)$ ein $\xi_x \in [0, 1]$ so, dass

$$f(x) = f(x_0) + \nabla f(x_0)^T(x - x_0) + \frac{1}{2}(x - x_0)^T H_f(x_0 + \xi_x(x - x_0))(x - x_0).$$

Dann gilt

$$\begin{aligned} f(x) &= f(x_0) + \nabla f(x_0)^T(x - x_0) + \frac{1}{2}(x - x_0)^T H_f(x_0)(x - x_0) \\ &\quad + \frac{1}{2}(x - x_0)^T [H_f(x_0 + \xi_x(x - x_0)) - H_f(x_0)](x - x_0). \end{aligned} \quad (630)$$

Definiere nun $\phi : B_r(x_0) \rightarrow \mathbb{R}$ durch

$$\phi(x) := \begin{cases} \frac{1}{2\|x - x_0\|^2}(x - x_0)^T [H_f(x_0 + \xi_x(x - x_0)) - H_f(x_0)](x - x_0) & x \neq x_0 \\ 0 & x = x_0 \end{cases} \quad (631)$$

Dann gilt für $x \in B_r(x_0)$ mit $x \neq x_0$, dass

$$\|x - x_0\|^2 \phi(x) = \frac{1}{2}(x - x_0)^T [H_f(x_0 + \xi_x(x - x_0)) - H_f(x_0)](x - x_0)$$

Dieselbe Gleichung gilt aber auch für $x = x_0$, da dann auf beiden Seiten 0 steht. Nutzt man dies in (630) so gilt

$$f(x) = f(x_0) + \nabla f(x_0)^T(x - x_0) + \frac{1}{2}(x - x_0)^T H_f(x_0)(x - x_0) + \|x - x_0\|^2 \phi(x) \quad \forall x \in B_r(x_0).$$

Noch $\mathbb{Z}$: ϕ ist stetig bei x_0 , d.h. $\lim_{x \rightarrow x_0} \phi(x) = \phi(x_0) (= 0)$.

Dazu: Für $x \in B_r(x_0)$ mit $x \neq x_0$ definiere $A_x := H_f(x_0 + \xi_x(x - x_0)) - H_f(x_0)$ und schreibe $A_x = (a_{ij}(x))_{j=1, \dots, d}^{i=1, \dots, m}$ wobei (vgl. Definition 10.6.5)

$$a_{ij}(x) = \partial_j \partial_i f(x_0 + \xi_x(x - x_0)) - \partial_j \partial_i f(x_0).$$

Insbesondere gilt dann nach (631), dass $\phi(x) = \frac{(x-x_0)^T A_x (x-x_0)}{2\|x-x_0\|^2}$ für alle $x \in B_r(x_0) \setminus \{x_0\}$. Man erhält

$$\begin{aligned}
|\phi(x)| &= \frac{|(x-x_0)^T A_x (x-x_0)|}{2\|x-x_0\|^2} \stackrel{(628)}{=} \frac{1}{2\|x-x_0\|^2} \left| \sum_{i,j=1}^d a_{ij}(x) (x-x_0)^{(i)} (x-x_0)^{(j)} \right| \\
&\leq \frac{1}{2\|x-x_0\|^2} \sum_{i,j=1}^d |a_{ij}(x)| |(x-x_0)^{(i)}| |(x-x_0)^{(j)}| \\
&\leq \frac{1}{2\|x-x_0\|^2} \sum_{i,j=1}^d |a_{ij}(x)| (\|x-x_0\|)(\|x-x_0\|) \\
&= \frac{1}{2\|x-x_0\|^2} \sum_{i,j=1}^d |a_{ij}(x)| \|x-x_0\|^2 \\
&= \frac{1}{2} \sum_{i,j=1}^d |a_{ij}(x)|. \tag{632}
\end{aligned}$$

ZWISCHENBEHAUPTUNG. Für alle $i, j \in \{1, \dots, d\}$ gilt $\lim_{x \rightarrow x_0} a_{ij}(x) = 0$

Dazu seien $i, j \in \{1, \dots, d\}$ beliebig und sei $(x_n)_{n \in \mathbb{N}} \subseteq B_r(x_0)$ eine Folge mit $\lim_{n \rightarrow \infty} x_n = x_0$. Dann gilt

$$\|(x_0 + \xi_{x_n}(x_n - x_0)) - x_0\| = \|\xi_{x_n}(x_n - x_0)\| = |\xi_{x_n}| \|x_n - x_0\| \leq 1 \|x_n - x_0\| \xrightarrow{n \rightarrow \infty} 0.$$

Es folgt also mit Satz 10.2.6, dass $\lim_{n \rightarrow \infty} (x_0 + \xi_{x_n}(x_n - x_0)) = x_0$. Da f zweimal stetig partiell differenzierbar ist, ist $\partial_j \partial_i f$ bei x_0 stetig und deswegen gilt auch

$$\lim_{n \rightarrow \infty} \partial_j \partial_i f(x_0 + \xi_{x_n}(x_n - x_0)) = \partial_j \partial_i f(x_0).$$

Daraus ergibt sich

$$\lim_{n \rightarrow \infty} a_{ij}(x_n) = \lim_{n \rightarrow \infty} (\partial_j \partial_i f(x_0 + \xi_{x_n}(x_n - x_0)) - \partial_j \partial_i f(x_0)) = 0.$$

Da $(x_n)_{n \in \mathbb{N}}$ beliebig (mit $\lim_{n \rightarrow \infty} x_n = x_0$) war gilt

$$\lim_{x \rightarrow x_0} a_{ij}(x) = 0.$$

ZWISCHENBEHAUPTUNG ENDE.

Wir wissen aus (632), dass

$$0 \leq |\phi(x)| \leq \frac{1}{2} \sum_{i,j=1}^d |a_{ij}(x)|.$$

Wegen der Zwischenbehauptung strebt die rechte Seite gegen 0 für $x \rightarrow x_0$. Aus dem Sandwichlemma folgt

$$\lim_{x \rightarrow x_0} |\phi(x)| = 0$$

Somit gilt auch

$$\lim_{x \rightarrow x_0} \phi(x) = 0 = \phi(x_0).$$

Die Stetigkeit von ϕ bei x_0 — und somit die Behauptung — ist gezeigt.

^aIn der Tat: Es gilt für alle $t \in [0, 1]$, dass

$$\|x_0 + t(x - x_0) - x_0\| = \|t(x - x_0)\| = |t|\|x - x_0\| \leq \|x - x_0\| < r.$$

Somit hat man $x_0 + t(x - x_0) \in B_r(x_0) \subseteq D$ und folgert, dass $\text{Seg}_{x,x_0} \subseteq D$.

Extremwertaufgaben

Sobald die Taylor-Approximation geglückt ist, können wir uns wie im Eindimensionalen *Extremwertaufgaben* widmen. Hierzu wiederholen wir die Definition aus der eindimensionalen Analysis für lokale Extrema

Definition 10.6.11 — LOKALE EXTREMSTELLEN IM MEHRDIMENSIONALEN

Es sei $D \subseteq \mathbb{R}^d$ und $f : D \rightarrow \mathbb{R}$ eine Funktion sowie $x_0 \in D$. Wir sagen,

- (i) x_0 ist eine *lokale Maximumsstelle* von f , falls es ein $r > 0$ gibt mit

$$f(x) \leq f(x_0) \quad \forall x \in D \cap B_r(x_0)$$

- (ii) x_0 ist eine *lokale Minimumsstelle* von f , falls es ein $r > 0$ gibt mit

$$f(x) \geq f(x_0) \quad \forall x \in D \cap B_r(x_0)$$

- (iii) x_0 ist eine *strikte lokale Maximumsstelle* von f , falls es ein $r > 0$ gibt mit

$$f(x) < f(x_0) \quad \forall x \in D \cap B_r(x_0) \setminus \{x_0\}$$

- (iv) x_0 ist eine *strikte lokale Minimumsstelle* von f , falls es ein $r > 0$ gibt mit

$$f(x) > f(x_0) \quad \forall x \in D \cap B_r(x_0) \setminus \{x_0\}$$

Ferner nennen wir x_0 eine (*strikte*) *lokale Extremstelle* von f falls f eine (strikte) lokale Maximumsstelle oder Minimumsstelle ist.

Bereits im Eindimensionalen haben wir Kriterien für lokale Extrema aus der Taylor-Approximation gefolgert. Dies wollen wir auch im Mehrdimensionalen tun. Zunächst wenden wir uns aber dem (nicht überraschenden) notwendigen Kriterium für lokale Extrema zu. Im Eindimensionalen sprachen wir sogar über ein notwendiges Kriterium für *innere* lokale Extrema (vgl. Satz 7.2.6). Da wir die Resultate im Mehrdimensionalen (aus den in Bemerkung 10.5.4 genannten Gründen) nur für offene Mengen entwickeln können wir das Wort „innere“ hier weglassen. Schließlich ist ja dann jeder Punkt ein innerer Punkt!

Satz 10.6.12 — NOTWENDIGES KRITERIUM FÜR LOKALE EXTREMA

Es sei $D \subseteq \mathbb{R}^d$ offen und $f : D \rightarrow \mathbb{R}$ eine Funktion, die an einem Punkt $x_0 \in D$ partiell differenzierbar ist. Ist x_0 eine lokale Extremstelle von f so gilt

$$\nabla f(x_0) = 0.$$

Beweis. Es sei $f : D \rightarrow \mathbb{R}$ wie in der Aussage. ObdA sei x_0 eine lokale Maximumsstelle (sonst betrachte $-f$). Sei etwa (vgl. Definition 10.6.11) $r > 0$ so, dass $f(x) \leq f(x_0)$ für alle $x \in D \cap B_r(x_0)$ (*). Nach Bemerkung 10.5.15 gilt

$$\nabla f(x_0) = \begin{pmatrix} \partial_1 f(x_0) \\ \vdots \\ \partial_d f(x_0) \end{pmatrix}.$$

Nach Definition 10.5.11 und Definition 10.5.3 hat man für $i \in \{1, \dots, d\}$

$$\partial_i f(x_0) = \partial_{e_i} f(x_0) = \gamma'_{e_i}(0), \quad (633)$$

wobei $\gamma_{e_i} : D_{e_i} \rightarrow \mathbb{R}$ gegeben ist durch $\gamma_{e_i}(t) = f(x_0 + te_i)$.

ZWISCHENBEHAUPTUNG. γ_{e_i} hat bei $t = 0$ eine lokale Maximumsstelle.

Es sei dazu $t \in (-r, r) \cap D_{e_i}$. Dann ist (weil $t \in D_{e_i}$) $x_0 + te_i \in D$, vgl. Definition 10.5.1. Desweiteren gilt $x_0 + te_i \in B_r(x_0)$, denn

$$\|(x_0 + te_i) - x_0\| = \|te_i\| = |t|\|e_i\| = |t| < r.$$

Da also $x_0 + te_i \in D \cap B_r(x_0)$ gilt nach (*)

$$f(x_0 + te_i) \leq f(x_0).$$

Das bedeutet

$$\gamma_{e_i}(t) \leq \gamma_{e_i}(0) \quad \forall t \in (-r, r) \cap D_{e_i}.$$

Da γ_{e_i} somit bei 0 ein inneres^a lokales Maximum annimmt gilt nach Satz 7.2.6 $\gamma'_{e_i}(0) = 0$. Aus (633) folgt dann $\partial_{e_i} f(x_0) = 0$. Da $i \in \{1, \dots, d\}$ beliebig war erhalten wir

$$\nabla f(x_0) = 0.$$

^aUm das Wort „Inneres“ zu rechtfertigen, benötigt man, dass 0 ein innerer Punkt von D_{e_i} ist. Dies folgt an dieser Stelle aus der Offenheit von D und Bemerkung 10.5.2!

Der obige Satz motiviert die nachfolgende Definition eines *kritischen Punktes*.

Definition 10.6.13 — KRITISCHER PUNKT

Es sei $D \subseteq \mathbb{R}^d$ offen und $f : D \rightarrow \mathbb{R}$ eine partiell differenzierbare Funktion. Wir nennen $x_0 \in D$ einen *kritischen Punkt* oder auch *stationären Punkt* von f falls

$$\nabla f(x_0) = 0.$$

Neben dem obigen notwendigen Kriterium gibt es wie im Eindimensionalen auch ein *hinreichendes Kriterium zweiter Ordnung*, welches auch eine Aussage über die *Art* des kritischen Punktes (Maximum/Minimum) treffen kann, vgl. 7.4.11. Dieses Kriterium wollen wir nun im mehrdimensionalen Verallgemeinern. Dafür müssen wir aber zuerst studieren, was *Definitheit von Matrizen* bedeutet.

Definition 10.6.14 — DEFINITHEIT

Es sei $A \in \mathbb{R}^{d \times d}$ eine $d \times d$ -Matrix. A heißt

- (i) *positiv semidefinit* wenn für alle $v \in \mathbb{R}^d$ gilt, dass $v^T A v \geq 0$.
- (ii) *negativ semidefinit* wenn für alle $v \in \mathbb{R}^d$ gilt, dass $v^T A v \leq 0$.
- (iii) *positiv definit* wenn für alle $v \in \mathbb{R}^d \setminus \{0\}$ gilt, dass $v^T A v > 0$.
- (iv) *negativ definit* wenn für alle $v \in \mathbb{R}^d \setminus \{0\}$ gilt, dass $v^T A v < 0$.
- (v) *indefinit* falls A weder positiv semidefinit ist noch negativ semidefinit ist.^a

^aMan könnte äquivalent fordern: Es gibt ein $v_1 \in \mathbb{R}^d$ mit $v_1^T A v_1 < 0$ (d.h. A nicht positiv semidefinit) und es gibt ein $v_2 \in \mathbb{R}^d$ mit $v_2^T A v_2 > 0$ (d.h. A nicht negativ semidefinit)

Wie wir überprüfen ob eine konkret gegebene Matrix positiv/ negativ (semi)-definit ist, besprechen wir später. Jetzt sehen wir erstmal das erwartete hinreichende Kriterium für (strikte) lokale Extrema.

Satz 10.6.15 — HINREICHENDES KRITERIUM FÜR LOKALE EXTREMA

Es sei $D \subseteq \mathbb{R}^d$ offen und $f : D \rightarrow \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion.

- Ist $x_0 \in D$ so dass

(a) $\nabla f(x_0) = 0$ und

(b) $H_f(x_0)$ positiv definit ist,

so ist x_0 eine strikte lokale Minimumsstelle von f .

- Ist $x_0 \in D$ so dass

(a) $\nabla f(x_0) = 0$

(b) $H_f(x_0)$ negativ definit ist,

so ist x_0 eine strikte lokale Maximumsstelle von f .

Um den Beweis führen zu können benötigen wir zunächst ein vorbereitendes Lemma.

Lemma 10.6.16

Es sei $A \in \mathbb{R}^{d \times d}$ positiv definit. Dann gibt es ein $c > 0$ so dass

$$y^T A y \geq c \|y\|^2 \quad \forall y \in \mathbb{R}^d. \quad (634)$$

Beweis. Definiere

$$\mathbb{S} := \{v \in \mathbb{R}^d : \|v\| = 1\}.$$

Man sich leicht davon überzeugen, dass $\mathbb{S}$ kompakt ist^a. Die Funktion $q_A|_{\mathbb{S}} : \mathbb{S} \rightarrow \mathbb{R}, v \mapsto v^T A v$ ist dann stetig (vgl (628)) und auf einer kompakten Menge definiert. Nach dem Satz von Weierstraß 10.4.21 nimmt sie dann auf $\mathbb{S}$ ein Minimum an. Etwa sei $v_* \in \mathbb{S}$ so, dass

$$q_A(v_*) = \min_{v \in \mathbb{S}} q_A(v) = \min_{v \in \mathbb{S}} v^T A v$$

Definiere nun $c := q_A(v_*) = v_*^T A v_*$. Da A positiv definit ist (und $v_* \neq 0$, da $v_* \in \mathbb{S}$), gilt $v_*^T A v_* > 0$. Somit folgt auch $c > 0$. Wir zeigen nun (634) mit der obigen Wahl von c . Sei dazu $y \in \mathbb{R}^d$. Ist $y = 0$ so ist in (634) nichts zu zeigen. Ist $y \neq 0$ so kann man $w := \frac{y}{\|y\|}$ betrachten. Man stellt leicht mit der Homogenität der Norm fest, dass $\|w\| = \left\| \frac{y}{\|y\|} \right\| = \frac{1}{\|y\|} \|y\| = 1$, d.h. $w \in \mathbb{S}$. Man erhält

$$w^T A w \geq \min_{v \in \mathbb{S}} v^T A v = c.$$

Nun gilt aber auch

$$w^T A w = \left(\frac{y}{\|y\|} \right)^T A \left(\frac{y}{\|y\|} \right) = \frac{1}{\|y\|^2} y^T A y$$

Die vorigen beiden Gleichung implizieren

$$\frac{y^T A y}{\|y\|^2} \geq c$$

Das bedeutet, dass (wie behauptet) $y^T A y \geq c\|y\|^2$.

“Das haben wir im Beweis von Lemma 10.4.27 auch schon gemacht! Man kann aber auch leicht als Übungsaufgabe nochmal verifizieren, dass $\mathbb{S}$ abgeschlossen und beschränkt ist.

Nun sind wir in der Lage das hinreichende Kriterium aus Satz 10.6.15 zu beweisen

Beweis von Satz 10.6.15. Wir zeigen nur den ersten bullet point. Für einen Beweis des zweiten bullet point kann man $-f$ betrachten. Es sei $f : D \rightarrow \mathbb{R}$ wie zweimal stetig partiell differenzierbar und $x_0 \in D$ so, dass $\nabla f(x_0) = 0$ und $H_f(x_0)$ positiv definit ist. Nach Korollar 10.6.10 gibt es ein $r > 0$ und eine in x_0 stetige Funktion $\phi : B_r(x_0) \rightarrow \mathbb{R}$ mit $\phi(x_0) = 0$ so, dass für alle $x \in B_r(x_0)$ gilt

$$f(x) = f(x_0) + \nabla f(x_0)^T(x - x_0) + \frac{1}{2}(x - x_0)^T H_f(x_0)(x - x_0) + \|x - x_0\|^2 \phi(x).$$

Da $\nabla f(x_0) = 0$ hat man für alle $x \in B_r(x_0)$

$$f(x) = f(x_0) + \frac{1}{2}(x - x_0)^T H_f(x_0)(x - x_0) + \|x - x_0\|^2 \phi(x).$$

Da $H_f(x_0)$ positiv definit ist, gibt es nach Lemma 10.6.16 ein $c > 0$ so, dass $y^T H_f(x_0)y \geq c\|y\|^2$ für alle $y \in \mathbb{R}$. Wenden wir dies in obiger Gleichung für $y = x - x_0$ an so folgt

$$f(x) \geq f(x_0) + \frac{c}{2}\|x - x_0\|^2 + \|x - x_0\|^2 \phi(x) \quad \forall x \in B_r(x_0),$$

das bedeutet

$$f(x) \geq f(x_0) + \left[\frac{c}{2} + \phi(x)\right] \|x - x_0\|^2 \quad \forall x \in B_r(x_0) \quad (635)$$

Nun gilt (wegen der Stetigkeit von ϕ in x_0), dass $\lim_{x \rightarrow x_0} \phi(x) = \phi(x_0) = 0$. Also (nach Bemerkung 10.4.8)

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x \in D : \quad \|x - x_0\| \leq \delta \quad \text{gilt} \quad |\phi(x)| < \varepsilon.$$

Wenden wir dies mit $\varepsilon := \frac{c}{4} > 0$ an, so erhalten wir die Existenz eines $\delta > 0$ so, dass

$$\forall x \in D : \quad \|x - x_0\| \leq \delta \quad \text{gilt} \quad |\phi(x)| < \frac{c}{4}.$$

Wähle $\theta := \min\{r, \delta\}$. Für $x \in B_\theta(x_0)$ gilt dann wegen (635)

$$\begin{aligned} f(x) &\geq f(x_0) + \left[\frac{c}{2} + \phi(x)\right] \|x - x_0\|^2 \\ &\geq f(x_0) + \left[\frac{c}{2} - |\phi(x)|\right] \|x - x_0\|^2 \\ &\underset{x \in B_\delta(x_0)}{\geq} f(x_0) + \left[\frac{c}{2} - \frac{c}{4}\right] \|x - x_0\|^2 = f(x_0) + \frac{c}{4}\|x - x_0\|^2 \end{aligned}$$

Für $x \in B_\theta(x_0) \setminus \{x_0\}$ gilt dann also

$$f(x) \geq f(x_0) + \underbrace{\frac{c}{4} \|x - x_0\|^2}_{>0} > f(x_0).$$

Wir haben nachgewiesen, dass $f(x) > f(x_0)$ für alle $x \in D \cap B_\theta(x_0) \setminus \{x_0\}$. Nach Definition 10.6.11 folgt, dass x_0 ein striktes lokales Maximum von f ist.

06.07.2026

Bemerkung 10.6.17

Die Voraussetzung $\nabla f(x_0) = 0$ in Satz 10.6.15 kann nicht abgeschwächt werden, denn sie ist ja für das Vorliegen eines lokalen Extremums bereits notwendig (d.h. ohne diese Voraussetzung kann nie ein lokales Extremum vorliegen vgl. Satz 10.6.12)!

Die Voraussetzung an $H_f(x_0)$ kann aber auch nicht abgeschwächt werden! So kann zum Beispiel die Forderung „ $H_f(x_0)$ positiv definit“ im ersten bullet point von Satz 10.6.15 nicht durch „ $H_f(x_0)$ positiv semidefinit“ ersetzt werden. Betrachte dazu

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \quad f(x, y) = -(x^4 + y^4)$$

Es gilt

$$\nabla f(0, 0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad H_f(0, 0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Somit ist $(0, 0)^T$ ein kritischer Punkt, bei dem die Hessematrix $H_f(0, 0)$ positiv semidefinit ist.^a Jedoch ist $(0, 0)^T$ kein lokales Minimum von f . Im Gegenteil: $(0, 0)^T$ ist sogar ein striktes lokales Maximum von f ! In der Tat: Für $(x, y)^T \in \mathbb{R}^2 \setminus \{(0, 0)^T\}$ gilt

$$f(x, y) < 0 = f(0, 0).$$

Somit sind beide Voraussetzungen in Satz 10.6.15 essentiell.

^aDie positive Semidefinitheit der Nullmatrix sieht man leicht ein: Für $v = (v^{(1)}, v^{(2)})^T \in \mathbb{R}^2$ gilt

$$v^T \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} v = v^T \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} v^{(1)} \\ v^{(2)} \end{pmatrix} = v^T \begin{pmatrix} 0 \cdot v^{(1)} + 0 \cdot v^{(2)} \\ 0 \cdot v^{(1)} + 0 \cdot v^{(2)} \end{pmatrix} = v^T \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0.$$

Leider trifft der obige Satz nicht immer eine Aussage über das Vorliegen einer lokalen Extremstelle an einem kritischen Punkt x_0 . Vor allem dann nicht wenn die Hesse-Matrix indefinit oder nur semidefinit ist. In diesen Fällen kann man aber auch Informationen aus der Hesse-Matrix erhalten. Wir werden zum Beispiel im Folgenden zeigen, dass bei einer indefiniten Hesse-Matrix keine lokale Extremstelle vorliegen kann. Man spricht von einem *Sattelpunkt*. Bevor wir dieses Resultat besprechen hier eine vorbereitende Proposition:

Proposition 10.6.18

Es sei $D \subseteq \mathbb{R}^d$ offen, $f : D \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar und $x_0 \in D$. Dann gilt

- (i) Ist x_0 eine lokale Minimumsstelle von f so ist $H_f(x_0)$ positiv semidefinit.
- (ii) Ist x_0 eine lokale Maximumsstelle von f so ist $H_f(x_0)$ negativ semidefinit.

Beweis. Wir zeigen nur (i), denn für (ii) kann man wiederum $-f$ betrachten. Es sei $x_0 \in D$ eine lokale Minimumsstelle von f . Betrachte für $v \in \mathbb{R}^d \setminus \{0\}$ die Menge $D_v := \{t \in \mathbb{R} : x_0 + tv \in D\} \subseteq \mathbb{R}$ und die Sektorfunktion

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

Nach Bemerkung 10.5.2 ist 0 ein innerer Punkt von D_v . Man überzeugt sich dann leicht, dass 0 ein inneres lokales Minimum von γ_v in D_v ist (da eben x_0 ein lokales Minimum von f auf D ist)^a.

ZWISCHENBEHAUPTUNG. $\gamma_v''(0) \geq 0$.

In der Tat gilt allgemein: Hat eine zweimal stetig differenzierbare Funktion $f : (a, b) \rightarrow \mathbb{R}$ bei $z \in (a, b)$ ein lokales Minimum, so gilt $f''(z) \geq 0$. Das sieht man so: Beobachte zuerst dass nach Satz 7.2.6 notwendigerweise gilt $f'(z) = 0$. Wäre nun $f''(z) < 0$ so hätte man also $f'(z) = 0$ und $f''(z) < 0$. Nach Proposition 7.4.11 wäre z dann aber ein striktes lokales Maximum von f . Dies widerspricht der Voraussetzung dass z ein lokales Minimum von f ist.

Nach (627) ist die zweite Ableitung der Sektorfunktion $\gamma_v''(0)$ gegeben durch

$$\gamma_v''(0) = v^T H_f(x_0) v.$$

Wir folgern, dass $0 \leq v^T H_f(x_0) v$ für alle $v \in \mathbb{R}^d \setminus \{0\}$. Für den Nullvektor $v = 0$ ist diese Ungleichung klarerweise auch wahr, das $0^T H_f(x_0) 0 = 0 \geq 0$. Die positive Semidefinitheit von $H_f(x_0)$ folgt.

^aDas Argument, was hierfür benötigt wird ist ganz analog zu der Zwischenbehauptung im Beweis von Satz 10.6.12

Bemerkung 10.6.19

Wie bereits in Bemerkung 10.6.17 gesehen gilt bei obiger Proposition die Umkehrung nicht: Nur weil $H_f(x_0)$ positiv semidefinit ist, ist x_0 nicht zwingend ein lokales Minimum.

Als Korollar aus Proposition 10.6.18 fällt heraus: Bei einer indefiniten Hessematrix kann keine Extremstelle vorliegen!

Korollar 10.6.20

Es sei $D \subseteq \mathbb{R}^d$ offen, $f : D \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar und $x_0 \in D$. Ist $H_f(x_0)$ indefinit so ist x_0 keine lokale Extremstelle von f .

Beweis. Angenommen x_0 ist eine lokale Extremstelle von f .

FALL 1. x_0 ist lokale Minimumsstelle von f .

Dann wäre nach Proposition 10.6.18 (i) die Hesse-Matrix $H_f(x_0)$ positiv semidefinit. Ein Widerspruch zur Indefinitheit von $H_f(x_0)$ (vgl. Definition 10.6.14)

FALL 2. x_0 ist lokale Maximumsstelle von f .

Dann wäre nach Proposition 10.6.18 (ii) die Hesse-Matrix $H_f(x_0)$ negativ semidefinit. Ein Widerspruch zur Indefinitheit von $H_f(x_0)$.

Ein Beispiel für eine Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, die an einem kritischen Punkt eine indefinite Hesse-Matrix besitzt ist

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = x^2 - y^2 \quad (636)$$

In der Tat: Man berechnet für $(x, y)^T \in \mathbb{R}^2$

$$\nabla f(x, y) = \begin{pmatrix} 2x \\ -2y \end{pmatrix} \quad \text{und} \quad H_f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

Es gilt also $\nabla f(0, 0) = 0$ und $H_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$ was (wie wir später in Beispiel 10.6.22 noch diskutieren werden) indefinit ist. Somit liegt bei $(0, 0)^T$ kein lokales Extremum sondern ein Sattelpunkt vor.

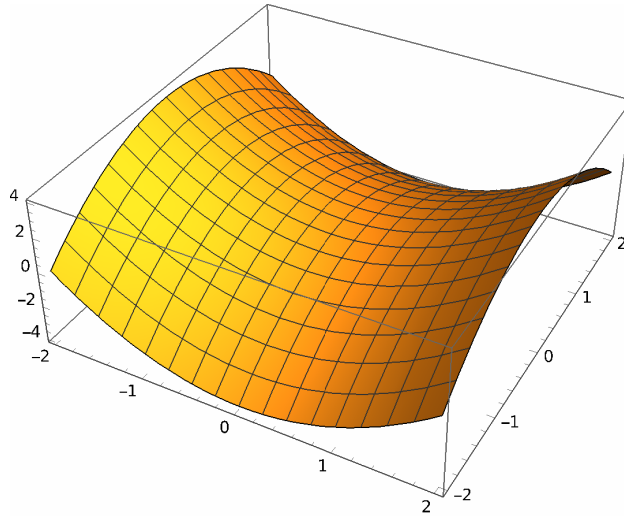


Abbildung 46: Ein Plot von f wie in (636). Man sieht, dass $(0, 0)^T$ ein Sattelpunkt ist.

ZUSAMMENFASSUNG.

Es sei $f : D \rightarrow \mathbb{R}$ zweimal stetig partiell differenzierbar und $x_0 \in D$ ein kritischer Punkt von f , d.h. $\nabla f(x_0) = 0$

- Ist $H_f(x_0)$ positiv definit, so ist x_0 eine (strikte) lokale Minimumsstelle (vgl. Satz 10.6.15).
- Ist $H_f(x_0)$ negativ definit, so ist x_0 eine (strikte) lokale Maximumsstelle (vgl. Satz 10.6.15).
- Ist $H_f(x_0)$ indefinit, so liegt in x_0 keine lokale Extremstelle vor (d.h. x_0 ist ein Sattelpunkt, vgl. Korollar 10.6.20).
- Ist $H_f(x_0)$ (positiv oder negativ) semidefinit, so kann mit den hier behandelten Kriterien keine Aussage getroffen werden.

Man beachte: Im letzten Fall trifft auch Proposition 10.6.18 keine Aussage, obwohl es auf den ersten Blick so scheinen könnte. Es ist nämlich so, dass die Implikation dort in die falsche Richtung geht!

Bemerkung 10.6.21

Dass im letzten Fall ohne weitere Bedingung keine Aussage getroffen werden kann, sieht man am besten, wenn man die Funktionen

$$f_1 : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f_1(x, y) = x^4 + y^4 \quad f_2 : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f_2(x, y) = -(x^4 + y^4)$$

betrachtet. Beide Funktionen haben einen kritischen Punkt bei $(0, 0)^T$ mit der (semidefini-

ten(!)) Hesse-Matrix $H_{f_1}(0,0) = H_{f_2}(0,0) = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$. Jedoch hat f_1 ein lokales Minimum bei $(0,0)^T$ und f_2 ein lokales Maximum bei $(0,0)^T$. Vgl. auch Bemerkung 10.6.17.

Überprüfung von Definitheit Die vorher kennengelernten Kriterien erlauben uns zumeist die *Klassifikation* kritischer Punkte (d.h. die Beantwortung der Frage ob es sich um lokale Maxima/Minima etc. handelt). Hierfür muss jedoch stets die *Definitheit* der Hesse-Matrix untersucht werden. Im Allgemeinen ist diese Untersuchung nicht ganz leicht — es müsste die *quadratische Form* gebildet werden und auf ein Vorzeichen untersucht werden. Wir werden im Laufe dieses Abschnittes das *Definitheitskriterium von Hurwitz-Sylvester* kennenlernen, welches diese Untersuchung erheblich erleichtert. Bevor wir dies tun wollen wir aber zunächst ein Beispiel studieren, welches uns eine Intuition für den Definitheitsbegriff geben soll.

Beispiel 10.6.22 — DEFINITHEIT BEI DIAGONALMATRIZEN

Es seien $\lambda_1, \lambda_2 \in \mathbb{R}$. Definiere

$$A := \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}.$$

Dann gilt für $v = (v^{(1)}, v^{(2)})^T \in \mathbb{R}^2$

$$v^T A v = v^T \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{pmatrix} v^{(1)} \\ v^{(2)} \end{pmatrix} = v^T \begin{pmatrix} \lambda_1 v^{(1)} \\ \lambda_2 v^{(2)} \end{pmatrix} = \begin{pmatrix} v^{(1)} \\ v^{(2)} \end{pmatrix}^T \begin{pmatrix} \lambda_1 v^{(1)} \\ \lambda_2 v^{(2)} \end{pmatrix} = \lambda_1 (v^{(1)})^2 + \lambda_2 (v^{(2)})^2.$$

FALL 1. $\lambda_1, \lambda_2 \geq 0$.

Dann gilt $v^T A v = \lambda_1 (v^{(1)})^2 + \lambda_2 (v^{(2)})^2 \geq 0$ für alle $v \in \mathbb{R}^2$. Somit ist nach Definition 10.6.14 A positiv semidefinit.

FALL 2. $\lambda_1, \lambda_2 \leq 0$.

Dann gilt $v^T A v = \lambda_1 (v^{(1)})^2 + \lambda_2 (v^{(2)})^2 \leq 0$ für alle $v \in \mathbb{R}^2$. Somit ist nach Definition 10.6.14 A negativ semidefinit.

FALL 3. $\lambda_1 > 0, \lambda_2 < 0$.

Man sieht leicht, dass A dann indefinit ist. In der Tat gilt

$$e_1^T A e_1 = \lambda_1 (e_1^{(1)})^2 + \lambda_2 (e_1^{(2)})^2 = \lambda_1 > 0$$

und

$$e_2^T A e_2 = \lambda_1 (e_2^{(1)})^2 + \lambda_2 (e_2^{(2)})^2 = \lambda_2 < 0$$

Mit Definition 10.6.14 (Fußnote von Punkt (v)) folgt die Indefinitheit von A .

FALL 4. $\lambda_1 < 0, \lambda_2 > 0$.

Mit derselben Technik wie in Fall 3 folgert man die Indefinitheit von A .

Man kann sich auch in Fall 1 und Fall 2 leicht davon überzeugen, dass

$$\lambda_1, \lambda_2 > 0 \quad \Rightarrow \quad A \text{ positiv definit}$$

und

$$\lambda_1, \lambda_2 < 0 \quad \Rightarrow \quad A \text{ negativ definit.}$$

Das Wissen aus dem vorigen Beispiel können wir auch einsetzen um Extrema zu klassifizieren.

Beispiel 10.6.23

Es sei für $\lambda_1, \lambda_2 \in \mathbb{R} \setminus \{0\}$

$$f: \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := \frac{1}{2}\lambda_1 x^2 + \frac{1}{2}\lambda_2 y^2$$

Behauptung. $(0, 0)^T$ ist der einzige kritische Punkt von f auf $\mathbb{R}^2$ und

$$H_f(0, 0) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Beweis.

$$\nabla f(x, y) = \begin{pmatrix} \partial_x f(x, y) \\ \partial_y f(x, y) \end{pmatrix} = \begin{pmatrix} \partial_x [\frac{1}{2}\lambda_1 x^2 + \frac{1}{2}\lambda_2 y^2] \\ \partial_y [\frac{1}{2}\lambda_1 x^2 + \frac{1}{2}\lambda_2 y^2] \end{pmatrix} = \begin{pmatrix} \lambda_1 x \\ \lambda_2 y \end{pmatrix}.$$

Somit hat man

$$\nabla f(x, y) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Longleftrightarrow \quad \lambda_1 x = 0 \quad \text{und} \quad \lambda_2 y = 0 \quad (637)$$

Da $\lambda_1, \lambda_2 \neq 0$ folgert man

$$\nabla f(x, y) = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \Longleftrightarrow \quad (x, y)^T = (0, 0)^T.$$

Nun zu $H_f(x, y)$. Es gilt nach (637)

$$\partial_x \partial_x f(x, y) = \partial_x (\lambda_1 x) = \lambda_1,$$

$$\partial_y \partial_x f(x, y) = \partial_y (\lambda_1 x) = 0,$$

$$\partial_x \partial_y f(x, y) \stackrel{\text{Satz 10.6.2}}{=} \partial_y \partial_x f(x, y) = 0$$

$$\partial_y \partial_y f(x, y) = \partial_y (\lambda_2 y) = \lambda_2.$$

Daher gilt

$$H_f(x, y) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \quad \forall (x, y)^T \in \mathbb{R}^2.$$

Insbesondere gilt auch $H_f(0, 0) = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$

Dies erlaubt uns eine Klassifikation des Typs des kritischen Punktes $(0, 0)^T$

FALL 1. $\lambda_1, \lambda_2 > 0$.

In diesem Fall ist $H_f(0, 0)$ nach Beispiel 10.6.22 positiv definit. Daher ist $(0, 0)^T$ ein striktes lokales Minimum.

FALL 2. $\lambda_1, \lambda_2 < 0$.

In diesem Fall ist $H_f(0, 0)$ nach Beispiel 10.6.22 negativ definit. Daher ist $(0, 0)^T$ ein striktes lokales Maximum.

FALL 3. λ_1, λ_2 haben verschiedene Vorzeichen.

In diesem Fall ist $H_f(0, 0)$ nach Beispiel 10.6.22 indefinit. Daher ist $(0, 0)^T$ keine lokale Extremstelle.

Nachfolgend finden Sie einige Plots der Funktionen aus Beispiel 10.6.23.

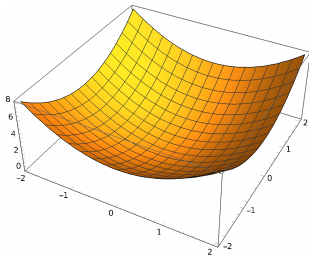


Abbildung 47: Plot für $\lambda_1 = \lambda_2 = 1$

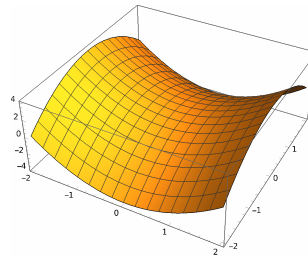


Abbildung 48: Plot für $\lambda_1 = 1, \lambda_2 = -1$

Damit wir nun nicht jedes Mal aufs Neue die quadratische Form bilden müssen, studieren wir das *Definitheitskriterium von Hurwitz-Sylvester*. Hierfür ist der Begriff der *Determinante* notwendig, den wir an dieser Stelle aus der linearen Algebra aufgreifen. Für $k \in \mathbb{N}$ sei $S_k := \{\sigma : \{1, \dots, k\} \rightarrow \{1, \dots, k\} : \sigma \text{ bijektive Abbildung}\}$. Für ein $\sigma \in S_k$ definieren wir

$$\text{sgn}(\sigma) := \prod_{\substack{i,j=1 \\ i \neq j}}^k \frac{\sigma(i) - \sigma(j)}{i - j}$$

Man kann zeigen, dass $\text{sgn}(\sigma) \in \{-1, 1\}$. Für eine $k \times k$ -Matrix $A \in \mathbb{R}^{k \times k}$ mit Einträgen

$(a_{ij})_{j=1,\dots,k}^{i=1,\dots,k}$ definiere die sogenannte *Determinante* von A durch

$$\det(A) := \sum_{\sigma \in S_k} \text{sgn}(\sigma) \prod_{i=1}^k a_{i\sigma(i)}. \quad (638)$$

Mit dieser Formel kann man die Determinante (mühsam, aber in endlicher Zeit) berechnen. Für $k = 1$ ist es einfach: $\det(a_{11}) = a_{11}$. Für $k = 2$ erhält man

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Für $k = 3$ erhält man die sogenannte *Sarrus-Regel*

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}.$$

Die Definitheit einer Matrix kann nun anhand der sogenannten *Hauptminoren* überprüft werden.

Definition 10.6.24 — HAUPTMINOREN

Es sei $A \in \mathbb{R}^{d \times d}$ eine Matrix, etwa $A = (a_{ij})_{j=1,\dots,d}^{i=1,\dots,d}$. Dann ist für $k \in \{1, \dots, d\}$ die k -te *Hauptminorenmatrix* $\text{HM}_k(A) \in \mathbb{R}^{k \times k}$ von A gegeben durch

$$\text{HM}_k(A) = (a_{ij})_{j=1,\dots,k}^{i=1,\dots,k}.$$

Die Determinante $\det(\text{HM}_k(A))$ heißt dann k -te *Hauptminore* von A .

In Worten ist $\text{HM}_k(A)$ die Matrix, die entsteht, wenn man nur den oberen $k \times k$ -Block der Matrix A betrachtet.

Satz 10.6.25 — DEFINITHEITSKRITERIUM VON HURWITZ-SYLVESTER

Es sei $A \in \mathbb{R}^{d \times d}$ eine symmetrische Matrix, d.h. $A^T = A$.^a Dann gilt

- (i) A ist positiv definit $\iff \det(\text{HM}_k(A)) > 0$ für alle $k \in \{1, \dots, d\}$
- (ii) A ist negativ definit $\iff (-1)^k \det(\text{HM}_k(A)) > 0$ für alle $k \in \{1, \dots, d\}$ ^b
- (iii) Ist $\det(A) \neq 0$ und treffen die Bedingungen auf der rechten Seite von (i) und (ii) beide nicht zu $\implies A$ ist indefinit.

^aDiese Voraussetzung ist für unsere Use-Cases mit der Hesse-Matrix zumeist kein Problem, vgl Bemerkung 10.6.6.

^bMit anderen Worten: $\det(\text{HM}_1(A)) < 0, \det(\text{HM}_2(A)) > 0, \det(\text{HM}_3(A)) < 0, \det(\text{HM}_4(A)) > 0, \det(\text{HM}_5(A)) < 0, \dots$ (die Vorzeichen alternieren!)

***Beweisskizze.** Zu (i). Wir zeigen die Behauptung nur im Fall $d = 2$. Es sei

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

für gewisse $a, b, c \in \mathbb{R}$. Dann gilt

$$\text{HM}_1(A) = (a)$$

und

$$\text{HM}_2(A) = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

Insbesondere gilt

$$\det(\text{HM}_1(A)) = a, \quad \det(\text{HM}_2(A)) = ac - b^2. \quad (639)$$

Nun berechnen wir die quadratische Form. Sei dazu $v = (v^{(1)}, v^{(2)})^T \in \mathbb{R}^2$. Dann gilt

$$\begin{aligned} v^T A v &= v^T \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} v^{(1)} \\ v^{(2)} \end{pmatrix} = v^T \begin{pmatrix} av^{(1)} + bv^{(2)} \\ bv^{(1)} + cv^{(2)} \end{pmatrix} = v^{(1)}(av^{(1)} + bv^{(2)}) + v^{(2)}(bv^{(1)} + cv^{(2)}) \\ &= a(v^{(1)})^2 + bv^{(1)}v^{(2)} + bv^{(2)}v^{(1)} + c(v^{(2)})^2 = a(v^{(1)})^2 + c(v^{(2)})^2 + 2bv^{(1)}v^{(2)}. \end{aligned} \quad (640)$$

Jetzt zum Beweis. Zuerst zu „ $\Leftarrow$ “. Es seien alle Hauptminoren positiv, d.h. (vgl. (639)) $a > 0$ und $ac - b^2 > 0$. Dann kann man die obige quadratische Form umschreiben zu

$$\begin{aligned} v^T A v &= a(v^{(1)})^2 + c(v^{(2)})^2 + 2bv^{(1)}v^{(2)} \\ &= a \left((v^{(1)})^2 + 2\frac{b}{a}v^{(1)}v^{(2)} \right) + c(v^{(2)})^2 \\ &= a \left(\left(v^{(1)} + \frac{b}{a}v^{(2)} \right)^2 - \frac{b^2}{a^2}(v^{(2)})^2 \right) + c(v^{(2)})^2 \\ &= a \left(v^{(1)} + \frac{b}{a}v^{(2)} \right)^2 - \frac{b^2}{a}(v^{(2)})^2 + c(v^{(2)})^2 \\ &= a \left(v^{(1)} + \frac{b}{a}v^{(2)} \right)^2 + \left(c - \frac{b^2}{a} \right) (v^{(2)})^2 \\ &= a \left(v^{(1)} + \frac{b}{a}v^{(2)} \right)^2 + \frac{ac-b^2}{a} (v^{(2)})^2 \end{aligned}$$

Nun gilt, da $a > 0$ und $ac - b^2 > 0$ (d.h. auch $\frac{ac-b^2}{a} > 0$), dass

$$v^T A v = \underbrace{a}_{>0} \underbrace{\left(v^{(1)} + \frac{b}{a}v^{(2)} \right)^2}_{\geq 0} + \underbrace{\frac{ac-b^2}{a}}_{>0} \underbrace{(v^{(2)})^2}_{\geq 0} \geq 0.$$

Wir müssen noch zeigen, dass für $v \in \mathbb{R}^2 \setminus \{0\}$ sogar gilt, dass der obige Ausdruck strikt größer als 0 ist. Sei dazu $v \neq 0$. Wäre der Ausdruck gleich 0 so müsste gelten $v^{(1)} + \frac{b}{a}v^{(2)} = 0$ und $v^{(2)} = 0$. In dem Fall wäre jedoch auch

$$v^{(1)} = v^{(1)} + \frac{b}{a}v^{(2)} - \frac{b}{a}v^{(2)} = 0 - \frac{b}{a} \cdot 0 = 0.$$

Es würde folgen, dass $v^{(1)} = v^{(2)} = 0$ was der Annahme $v \neq 0$ widerspricht. Wir folgern also

$$v^T A v > 0 \quad \forall v \in \mathbb{R}^2 \setminus \{0\}.$$

Also ist A positiv definit.

Nun zu „ $\implies$ “. Es sei $A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$ positiv definit. Dann gilt

$$0 < e_1^T A e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}^T \begin{pmatrix} a \\ b \end{pmatrix} = a.$$

Wir folgern also $a > 0$ und somit nach (639) auch $\det(\text{HM}_1(A)) > 0$. Betrachte nun den Vektor $v = (-\frac{b}{a}, 1)^T$. Dann gilt (640)

$$\begin{aligned} 0 < v^T A v &\stackrel{(640)}{=} a(v^{(1)})^2 + c(v^{(2)})^2 + 2bv^{(1)}v^{(2)} = a\left(-\frac{b}{a}\right)^2 + c + 2b\left(-\frac{b}{a}\right) = \frac{b^2}{a} + c - 2\frac{b^2}{a} \\ &= c - \frac{b^2}{a} = \frac{ac-b^2}{a}. \end{aligned}$$

Da $a > 0$ können wir diese Ungleichung mit a multiplizieren und erhalten $ac - b^2 > 0$. Nach (639) haben wir dann auch $\det(\text{HM}_2(A)) > 0$ bewiesen.

Nun noch zu (ii). Man überzeugt sich leicht dass

$$A \text{ negativ definit} \quad \iff \quad -A \text{ positiv definit.} \quad (641)$$

Nach (i) gilt

$$-A \text{ positiv definit} \quad \iff \quad \det(\text{HM}_k(-A)) > 0 \text{ für alle } k \in \{1, \dots, d\}. \quad (642)$$

Nun folgert man für $k \in \{1, \dots, d\}$ leicht aus Definition 10.6.24, dass $\text{HM}_k(-A) = -\text{HM}_k(A)$, sprich: jeder Eintrag von $\text{HM}_k(A)$ wird mit -1 multipliziert. Man kann aus (638) ablesen, was mit der Determinante einer $k \times k$ -Matrix passiert, wenn man jeden Eintrag mit (-1) multipliziert. Da jeder Summand in (638) ein Produkt von k Matrixeinträgen enthält, entsteht immerzu der Faktor $(-1)^k$. Wir folgern für $k \in \{1, \dots, d\}$

$$\det(\text{HM}_k(-A)) = \det(-\text{HM}_k(A)) = (-1)^k \det(\text{HM}_k(A)).$$

Nutzen wir dies, (641) und (642) so erhalten wir

$$A \text{ negativ definit} \iff (-1)^k \det(\text{HM}_k(A)) > 0 \quad \forall k \in \{1, \dots, d\}.$$

Kriterium (iii) beweisen wir hier nicht.

Man könnte meinen, dass man die strikten Ungleichungen in Satz 10.6.25 durch schwache Ungleichungen ersetzen kann und so Kriterien für Semidefinitheit erhalten kann. Das ist leider nicht der Fall, wie wir im folgenden Beispiel sehen.

Beispiel 10.6.26

Betrachte die Matrix

$$A := \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix}$$

Behauptung. $\det(\text{HM}_1(A)) \geq 0, \det(\text{HM}_2(A)) \geq 0$, aber A ist nicht positiv semidefinit.

Somit gilt: (A positiv semidefinit $\iff \det(\text{HM}_k(A)) \geq 0 \quad \forall k = 1, \dots, d$). Die Definitheitskriterien aus Satz 10.6.25 übertragen sich also nicht ohne Weiteres auf Semidefinitheit!

Beweis. Es gilt $\text{HM}_1(A) = (0)$, weswegen $\det(\text{HM}_1(A)) = 0 \geq 0$. Ferner gilt

$$\det(\text{HM}_2(A)) = \det \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} = 0 \cdot (-1) - 0 \cdot 0 = 0 \geq 0.$$

Jedoch ist A nicht positiv semidefinit, denn

$$e_2^T A e_2 = e_2^T \begin{pmatrix} 0 \\ -1 \end{pmatrix} = -1 < 0.$$

Dasselbe Beispiel zeigt, dass man in Teil (iii) von Satz 10.6.25 nicht auf die Voraussetzung $\det(A) \neq 0$ verzichtet werden kann. In der Tat ist A negativ semidefinit, denn für $v = (v^{(1)}, v^{(2)})^T$ gilt

$$v^T A v = \begin{pmatrix} v^{(1)} \\ v^{(2)} \end{pmatrix}^T \begin{pmatrix} 0 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} v^{(1)} \\ v^{(2)} \end{pmatrix} = \begin{pmatrix} v^{(1)} \\ v^{(2)} \end{pmatrix}^T \begin{pmatrix} 0 \\ -v^{(2)} \end{pmatrix} = -(v^{(2)})^2 \leq 0.$$

Somit ist A insbesondere nicht indefinit. Allerdings erfüllt A alle Voraussetzungen von Satz 10.6.25 (iii) bis auf " $\det(A) = 0$ ".

Kurvendiskussionen im Mehrdimensionalen Wir werden nun anhand einiger Beispiele zusammenfassend besprechen, wie man Funktionen auf lokale und globale Extrema diskutieren kann.

Beispiel 10.6.1. Gegeben sei

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) := (x - 3)y^2 + x^2(x + 3)$$

Wir

- bestimmen alle kritischen Punkte von f und klassifizieren sie, d.h. wir untersuchen ob es sich tatsächlich um lokale Extrema handelt.
- untersuchen ob $\max f(\mathbb{R}^2)$, $\min f(\mathbb{R}^2)$ existiert.

Dazu

SCHRITT 1. Bestimmung der kritischen Punkte.

Es gilt

$$\nabla f(x, y) = \begin{pmatrix} \partial_x f(x, y) \\ \partial_y f(x, y) \end{pmatrix} = \begin{pmatrix} y^2 + 2x(x + 3) + x^2 \cdot 1 \\ 2(x - 3)y \end{pmatrix} = \begin{pmatrix} y^2 + 3x^2 + 6x \\ 2(x - 3)y \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Betrachte nun die zweite Komponente: $2(x - 3)y = 0$ gilt genau dann wenn $y = 0$ oder $x - 3 = 0$ (d.h. $x = 3$).

FALL 1. $y = 0$.

Setzt man $y = 0$ in die erste Komponente ein, so erhält man $3x^2 + 6x = 0$, d.h.

$$0 = 3x^2 + 6x = x(3x + 6).$$

Somit gilt entweder $x = 0$ oder $3x + 6 = 0$, d.h. $x = -2$. Somit hat man die kritischen Punkte $(0, 0)^T$ und $(-2, 0)^T$

FALL 2. $x = 3$.

Setzt man $x = 3$ in der ersten Komponente ein, so erhält man $y^2 + 3 \cdot 9 + 6 \cdot 3 = 0$, d.h. $y^2 + 45 = 0$. Diese Gleichung ist für kein einziges $y \in \mathbb{R}$ erfüllt. Deswegen fördert dieser Fall keine weiteren kritischen Punkte zutage.

SCHRITT 2. Klassifikation der kritischen Punkte.

Um die gefundenen kritischen Punkte $(0, 0)^T$ und $(-2, 0)^T$ zu klassifizieren berechnen wir die Hesse-Matrix an diesen Punkten. Dazu berechnen zunächst allgemein

$$H_f(x, y) = \begin{pmatrix} 6x + 6 & 2y \\ 2y & 2(x - 3) \end{pmatrix}.$$

Somit gilt

$$H_f(0,0) = \begin{pmatrix} 6 & 0 \\ 0 & -6 \end{pmatrix} \quad \text{und} \quad H_f(-2,0) = \begin{pmatrix} -6 & 0 \\ 0 & -10 \end{pmatrix}$$

Nach Beispiel 10.6.22 ist dann $H_f(0,0)$ indefinit und $H_f(-2,0)$ negativ definit. Wir folgern

- An der Stelle $(0,0)^T$ liegt kein lokales Extremum vor.
- An der Stelle $(-2,0)^T$ liegt ein lokales Maximum vor.

SCHRITT 3. Globale Extrema

Wir fragen uns nun, ob $\min f(\mathbb{R}^2)$ und $\max f(\mathbb{R}^2)$ existieren. Klarerweise kann f auf $\mathbb{R}^2$ kein globales Minimum annehmen, denn eine globale Minimumsstelle wäre ja auch eine lokale Minimumsstelle. Wir haben aber in Schritt 2 bereits herausgefunden, dass f keine lokalen Minimumsstelle besitzt. Somit existiert $\min f(\mathbb{R}^2)$ nicht. Beachte: $\max f(\mathbb{R}^2)$ existiert auch nicht! In der Tat: Es gilt für $y \in \mathbb{R}$

$$f(4,y) = (4-3)y^2 + 4^2(4+3) = y^2 + 16 \cdot 7 = y^2 + 112.$$

Somit ist $\{f(4,y) : y \in \mathbb{R}\} = [112, \infty)$ nach oben unbeschränkt. Dadurch ist auch $f(\mathbb{R}^2)$ nach oben unbeschränkt. Gäbe es nun $\max f(\mathbb{R}^2)$ so wäre dieses eine obere Schranke von $f(\mathbb{R}^2)$. Eine solche obere Schranke gibt es aber nicht!

In einem weiteren Beispiel untersuchen wir Extrema nicht auf ganz $\mathbb{R}^2$, sondern auf einem kompakten Bereich.

Beispiel 10.6.27

Es sei wie im vorigen Beispiel

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x,y) = (x-3)y^2 + x^2(x+3)$$

Ferner sei

$$K := \{(x,y)^T \in \mathbb{R}^2 : x^2 + y^2 \leq 9\}$$

Wir untersuchen ob $\max f(K)$ und $\min f(K)$ angenommen werden und berechnen ggf. deren Werte.

SCHRITT 1. Die globalen Extrema werden angenommen!

Dass $\max f(K)$, $\min f(K)$ angenommen werden liegt daran, dass K kompakt ist^a und f stetig ist (vgl. Satz von Weierstraß, Satz 10.4.21).

SCHRITT 2. Kandidaten aus $\overset{\circ}{K}$. (Hierbei nutzen wir $\overset{\circ}{K} = \{(x,y)^T \in \mathbb{R}^2 : x^2 + y^2 < 9\}$)^b

Wir suchen nun nach lokalen Extremstellen von f auf der offenen Menge $\overset{\circ}{K}$. Beachte: Aus Beispiel 10.6.1 wissen wir, dass

$$\nabla f(x, y) = (0, 0)^T \iff (x, y)^T = (-2, 0)^T \text{ oder } (x, y)^T = (0, 0)^T.$$

Bei $(0, 0)^T$ liegt aber laut Beispiel 10.6.1 kein lokales Extremum vor, da $H_f(0, 0)$ indefinit ist. Bei $(-2, 0)^T$ liegt ein lokales Maximum vor. Beachte: $(-2, 0)^T \in K$ da $(-2)^2 + 0^2 = 4 < 9$. Wir speichern nun $f(-2, 0)$ als möglichen Kandidaten für ein Maximum ab.

SCHRITT 3. Untersuchung des Randes ∂K .

Beachte: $\partial K = \overline{K} \setminus \overset{\circ}{K} = K \setminus \{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 < 9\} = \{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 = 9\}$. Somit gilt

$$\begin{aligned} \{f(x, y) \mid (x, y)^T \in \partial K\} &= \{f(x, y) \mid x^2 + y^2 = 9\} \\ &= \{(x-3)y^2 + x^2(x+3) \mid x^2 + y^2 = 9\} \\ &= \{(x-3)y^2 + x^2(x+3) \mid y^2 = 9 - x^2\} \\ &= \{(x-3)(9 - x^2) + x^2(x+3) \mid y^2 = 9 - x^2\} \end{aligned} \quad (643)$$

Eigentlich schreibt man die Menge am Ende dieser Gleichung besser auf als

$$\{(x-3)(9 - x^2) + x^2(x+3) \mid (x, y)^T \in \mathbb{R}^2, y^2 = 9 - x^2\}$$

Wir behaupten aufgrund von (643)

$$\{(x-3)(9 - x^2) + x^2(x+3) \mid (x, y)^T \in \mathbb{R}^2, y^2 = 9 - x^2\} = \{(x-3)(9 - x^2) + x^2(x+3) \mid x \in [-3, 3]\}$$

In der Tat: Für $x \in [-3, 3]$ kann man $y := \sqrt{9 - x^2}$ betrachten. Am Funktionswert $(x-3)(9 - x^2) + x^2(x+3)$ ändert das nichts. Für $(x, y) \in \mathbb{R}^2$ mit $y^2 = 9 - x^2$ gilt insbesondere $9 - x^2 = y^2 \geq 0$. Somit gilt $x^2 \leq 9$. Das zeigt, dass $x \in [-3, 3]$. Definiere nun $h : [-3, 3] \rightarrow \mathbb{R}$ durch

$$h(x) := (x-3)(9 - x^2) + x^2(x+3) = x^2[(x+3) - (x-3)] + 9(x-3) = 6x^2 + 9x - 27.$$

und nimm zur Kenntnis, dass

$$\begin{aligned} \max\{f(x, y) \mid (x, y)^T \in \partial K\} &= \max h([-3, 3]), \\ \min\{f(x, y) \mid (x, y)^T \in \partial K\} &= \min h([-3, 3]) \end{aligned}$$

Um Extrema von h zu finden bilden wir

$$h'(x) = 12x + 9 \begin{cases} > 0 & x > -\frac{3}{4} \\ < 0 & x < -\frac{3}{4} \end{cases}.$$

Nach dem Vorzeichenwechselkriterium (Satz 7.3.15) ist $h(-\frac{3}{4})$ ein globales Minimum von h auf $[-3, 3]$. Wegen dem Satz von Weierstraß muss h auf $[-3, 3]$ auch ein Maximum annehmen. Da es im Inneren keine lokalen Maxima gibt, müssen diese Maxima am Rand angenommen werden. Wir speichern also

- $h(-\frac{3}{4})$ als Kandidat für $\min f(K)$
- $h(3), h(-3)$ als Kandidaten für $\max f(K)$.

SCHRITT 4. Vergleich der Kandidaten.

Für $\min f(K)$ haben wir nur einen Kandidaten ($h(-\frac{3}{4})$) gespeichert. Da $\min f(K)$ nach Schritt 1 aber auch angenommen werden muss gilt

$$\min f(K) = h(-\frac{3}{4}) = 6(-\frac{3}{4})^2 + 9(-\frac{3}{4}) - 27 = \frac{54}{16} - \frac{27}{4} - 27 = -\frac{27}{8} - 27 = -\frac{9}{8} \cdot 27 = -\frac{243}{8}.$$

Für $\max f(K)$ haben wir drei Kandidaten.

$$f(-2, 0) = (-2 - 3) \cdot 0^2 + (-2)^2(-2 + 3) = 4$$

$$h(3) = 6 \cdot 3^2 + 9 \cdot 3 - 27 = 54, \quad h(-3) = 6 \cdot 3^2 - 9 \cdot 3 - 27 = 0.$$

Man erhält nach Vergleich der drei Werte 0, 4, 54, dass

$$\max f(K) = 54.$$

^aZur Kompaktheit von K . K kann umgeschrieben werden als

$$\begin{aligned} K &= \{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 \leq 9\} \\ &= \{(x, y)^T \in \mathbb{R}^2 : \|(x, y)^T - (0, 0)^T\|^2 \leq 9\} = \{(x, y) \in \mathbb{R}^2 : \|(x, y)^T - (0, 0)^T\| \leq 3\}. \end{aligned}$$

Nach Beispiel 10.3.4 gilt dann $K = \overline{B_3((0, 0)^T)}$. Man überzeugt sich damit leicht davon, dass K abgeschlossen ist. Auch die Beschränktheit liest man an der obigen Formel ab: Für $(x, y)^T \in K$ gilt stets $\|(x, y)^T\| \leq 3$.

^bWir beweisen in dieser Fußnote, dass $\overset{\circ}{K} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 9\}$. In der Tat: Definiere dazu $A := \{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 < 9\} = B_3(0)$. A ist nach Beispiel 10.3.7 offen. Deswegen ist jeder Punkt in A ein innerer Punkt von A . Das bedeutet: Für sämtliche $x_0 \in A$ gibt es ein $r > 0$ mit $B_r(x_0) \subseteq A$. Da $A \subseteq K$ gilt aber auch $B_r(x_0) \subseteq K$. Wir folgern: Für alle $x_0 \in A$ gibt es ein $r > 0$ mit $B_r(x_0) \subseteq K$. Dies zeigt $A \subseteq \overset{\circ}{K}$. Nun müssen wir noch zeigen, dass $\overset{\circ}{K} \subseteq A$. Dazu: Angenommen es gibt einen Punkt $(x_*, y_*)^T \in \overset{\circ}{K} \setminus A$. Da $\overset{\circ}{K} \subseteq K$ gilt $(x_*, y_*) \in K \setminus A$. Das bedeutet

$$x_*^2 + y_*^2 \leq 9 \quad \text{aber} \quad \neg(x_*^2 + y_*^2 < 9).$$

Wir folgern $x_*^2 + y_*^2 = 9$. Da $(x_*, y_*)^T \in \overset{\circ}{K}$ gibt es ein $r > 0$ mit $B_r((x_*, y_*)^T) \subseteq K$. Betrachte nun $(\tilde{x}, \tilde{y})^T := (1 + \frac{r}{6})(x_*, y_*)^T$. Dann gilt

$$\|(\tilde{x}, \tilde{y})^T - (x_*, y_*)^T\| = \|\frac{r}{6}(x_*, y_*)^T\| = \frac{r}{6}\|(x_*, y_*)^T\| = \frac{r}{6}\sqrt{x_*^2 + y_*^2} = \frac{r}{6}\sqrt{9} = \frac{r}{2} < r.$$

Somit hat man $(\tilde{x}, \tilde{y})^T \in B_r((x_*, y_*)^T) \subset K$. Allerdings gilt

$$\tilde{x}^2 + \tilde{y}^2 = (1 + \frac{r}{6})(x_*^2 + y_*^2) = (1 + \frac{r}{6}) \cdot 9 > 9,$$

weswegen $(\tilde{x}, \tilde{y})^T \notin K$. Ein Widerspruch.

08.07.2026

10.7 Differenzierbare Umkehrfunktionen

In der eindimensionalen Analysis (Satz 7.1.8) haben wir folgendes gesehen: Wenn wir eine differenzierbare bijektive Funktion $f : [a, b] \rightarrow Z$ gegeben haben und einen Punkt $x_0 \in [a, b]$ mit $f'(x_0) \neq 0$, so ist die Umkehrabbildung f^{-1} bei $y_0 := f(x_0)$ differenzierbar und es gilt

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} \left[= \frac{1}{f'(f^{-1}(y_0))} \right].$$

Wir wollen hier nun ein ähnliches Resultat in höherer Dimension besprechen. Den Beweis dieses Resultates vertagen wir in den Ferienkurs.

Satz 10.7.1 — LOKALER UMKEHRSATZ

Es sei $D \subseteq \mathbb{R}^d$ offen und $F : D \rightarrow \mathbb{R}^d$ auf D stetig partiell differenzierbar. Ferner sei $x_0 \in D$ so, dass

$$\det(J_F(x_0)) \neq 0.$$

Dann gibt es eine offene Menge $U_0 \subseteq D$ mit $x_0 \in U_0$ so, dass

- $V_0 := F(U_0)$ offen ist,
- $F|_{U_0} : U_0 \rightarrow V_0$ bijektiv ist,
- $(F|_{U_0})^{-1} : V_0 \rightarrow U_0$ stetig partiell differenzierbar ist.

^ad.h. notwendigerweise $m = d$, weswegen auch $J_F(x_0) \in \mathbb{R}^{d \times d}$ quadratisch ist.

Beweis. Später.

Der lokale Umkehrsatz heißt *lokal*, weil es sein kann dass die Umkehrfunktion nur auf einer sehr kleinen Menge U_0 definiert ist. Jedoch gilt: Weil U_0 offen ist und $x_0 \in U_0$ gibt es nach Bemerkung 10.3.6 stets ein $r > 0$ mit $B_r(x_0) \subseteq U_0$. Weil also ein kleiner Ball um den Punkt x_0 stets noch in U_0 liegt, nennen wir U_0 auch eine *offene Umgebung* von x_0 .

Bemerkung 10.7.2

Die Bedingung $\det(J_F(x_0)) \neq 0$ in Satz 10.7.1 kann mit ein wenig linearer Algebra umformuliert werden:

Wiederholung. (Determinanten und Invertierbarkeit). Es sei $A \in \mathbb{R}^{d \times d}$ eine Matrix. Nach

(584) besteht A aus den Spalten $Ae_1, \dots, Ae_d$. Die Zeilen von A erhält man indem man $e_1^T A, \dots, e_d^T A$ bildet. Dann sind folgende Aussagen äquivalent

(a) $\det(A) \neq 0$

(b) A ist *invertierbar*, d.h. es gibt eine Matrix $A^{-1} \in \mathbb{R}^{d \times d}$ ^(a) so, dass

$$AA^{-1} = A^{-1}A = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{pmatrix} \in \mathbb{R}^{d \times d} \quad (644)$$

(c) Die Menge der Spalten $\{Ae_1, \dots, Ae_d\}$ ist linear unabhängig.^b

(d) Die Menge der Zeilen $\{e_1^T A, \dots, e_d^T A\}$ sind linear unabhängig.

Anstatt der Bedingung $\det(J_F(x_0)) \neq 0$ kann man also eine der obigen Bedingungen (mit $A = J_F(x_0)$) fordern.

^aBeachte: Die Matrix A^{-1} ist eindeutig bestimmt durch die Gleichung (644) und wird von uns im Folgenden *inverse Matrix* genannt.

^bAn dieser Stelle erinnern wir an den Begriff der *linearen Unabhängigkeit*. Eine Menge von Vektoren $\{v_1, \dots, v_k\} \subseteq \mathbb{R}^d$ heißt linear unabhängig falls

$$\eta_1, \dots, \eta_k \in \mathbb{R} : \sum_{j=1}^k \eta_j v_j = 0 \quad \implies \quad \eta_1 = \dots = \eta_k = 0.$$

Als Korollar fällt eine Aussage heraus die ähnlich ist wie die Aussage von Satz 7.1.8 aus der eindimensionalen Analysis.

Korollar 10.7.3 — GLOBALER UMKEHRSATZ IM MEHRDIMENSIONALEN

Es seien $D_1, D_2 \subseteq \mathbb{R}^d$ offen und $F : D_1 \rightarrow D_2$ bijektiv und auf D_1 stetig partiell differenzierbar, mit Umkehrabbildung $F^{-1} : D_2 \rightarrow D_1$.^a Ferner gelte für ein $x_0 \in D_1$, dass $\det(J_F(x_0)) \neq 0$. Dann ist die Umkehrfunktion F^{-1} bei $y_0 := F(x_0)$ total differenzierbar und es gilt

$$J_{F^{-1}}(y_0) = [J_F(x_0)]^{-1} = [J_F(F^{-1}(y_0))]^{-1}. \quad (645)$$

^aBeachte: Die Umkehrabbildung existiert nach Proposition 1.3.18.

Beweis. Es sei $x_0 \in D_1$ wie in der Aussage. Nach Satz 10.7.1 gibt es eine offene Menge $U_0 \subseteq D_1$ mit $x_0 \in U_0$ so, dass $V_0 := F(U_0)$ offen ist und $F : U_0 \rightarrow V_0$ eine auf V_0 stetig partiell differenzierbare Umkehrabbildung besitzt. Da es nach Bemerkung 1.3.15 nur eine einzige Umkehrabbildung von $F : U_0 \rightarrow V_0$ geben darf, muss diese Umkehrabbildung mit $F^{-1}|_{V_0}$ übereinstimmen. Wir stellen also fest, dass $F^{-1}|_{V_0}$ auf ganz V_0 stetig partiell

differenzierbar ist. Da V_0 offen ist und $y_0 \in V_0$ (da $y_0 = F(x_0) \in F(U_0) = V_0$) folgt aus Satz 10.7.1 dass F^{-1} bei y_0 total differenzierbar ist. Jetzt muss nur noch (645) gezeigt werden. Dazu: Da F bei x_0 total differenzierbar ist und F^{-1} bei $y_0 = F(x_0)$ total differenzierbar ist kann man die Kettenregel (Satz 10.5.33) auf $F^{-1} \circ F$ anwenden und erhält

$$J_{F^{-1} \circ F}(x_0) = J_{F^{-1}}(F(x_0))J_F(x_0) = J_{F^{-1}}(y_0)J_F(x_0). \quad (646)$$

Beachte nun: Für alle $x \in D$ gilt

$$(F^{-1} \circ F)(x) = F^{-1}(F(x)) = x. \quad (647)$$

Somit kann $J_{F^{-1} \circ F}(x_0)$ auch auf eine andere Weise berechnet werden: Der Eintrag in der i -ten Zeile und j -ten Spalte von $J_{F^{-1} \circ F}(x)$ ist

$$\frac{\partial}{\partial x^{(i)}}(F^{-1} \circ F)^{(j)}(x) \stackrel{(647)}{=} \frac{\partial}{\partial x^{(i)}} x^{(j)} = \partial_{e_i} x^{(j)} \stackrel{\text{Bsp 10.5.7}}{\stackrel{(574)}}{=} e_i^{(j)} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}.$$

Somit sieht man

$$J_{F^{-1} \circ F}(x_0) = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{pmatrix}.$$

Wir erhalten gemeinsam mit (646)

$$\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & \cdots & 1 \end{pmatrix} = J_{F^{-1}}(y_0)J_F(x_0).$$

Die Matrix auf der linken Seite dieser Gleichung nennen wir $n \times n$ -Einheitsmatrix E_n . Man überzeugt sich leicht, dass für alle $A \in \mathbb{R}^{d \times d}$ gilt $E_n A = A E_n = A$. Multiplizieren wir mit $J_F(x_0)^{-1}$ von rechts so folgt

$$E_n J_F(x_0)^{-1} = J_{F^{-1}}(y_0)J_F(x_0)J_F(x_0)^{-1} \quad \text{und deswegen gilt} \quad J_F(x_0)^{-1} = J_{F^{-1}}(y_0).$$

Die Behauptung folgt.

Beispiel 10.7.4

Nach Satz 5.5.19 ist die *Polarabbildung*

$$f : (0, \infty) \times [0, 2\pi) \rightarrow \mathbb{R}^2 \setminus \{0\}, \quad f(r, \theta) := (r \cos \theta, r \sin \theta)^T \quad (648)$$

bijektiv. Da $\{f(r, 0) : r \in (0, \infty)\} = \{(x, 0) : x > 0\}$ ist auch

$$f : (0, \infty) \times (0, 2\pi) \rightarrow \mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\}, f(r, \theta) = (r \cos \theta, r \sin \theta)$$

bijektiv. Beachte: $D := (0, \infty) \times (0, 2\pi)$ ist offen. Man überzeugt sich leicht, dass f auch stetig partiell differenzierbar auf D ist. Außerdem gilt für alle $(r, \theta)^T \in (0, \infty) \times (0, 2\pi)$, dass

$$\det(J_f(r, \theta)) = \det \begin{pmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{pmatrix} = r \cos^2(\theta) + r \sin^2(\theta) = r \neq 0$$

Somit ist auch die Umkehrabbildung f^{-1} auf $\mathbb{R}^2 \setminus \{(x, 0) : x \geq 0\}$ total differenzierbar.

Man sieht an diesem Beispiel auch, dass auf die Offenheit der Menge D_1 in Korollar 10.7.3 nicht verzichtet werden kann. In der Tat: Wählt man $D_1 = (0, \infty) \times [0, 2\pi)$ und $f : D_1 \rightarrow \mathbb{R}^2 \setminus \{(0, 0)^T\}$ wie in (648). Nach wie vor gilt für alle $(r, \theta)^T \in D_1$

$$\det(J_f(r, \theta)) = \det \begin{pmatrix} \cos(\theta) & -r \sin(\theta) \\ \sin(\theta) & r \cos(\theta) \end{pmatrix} = r \cos^2(\theta) + r \sin^2(\theta) = r \neq 0.$$

Jedoch ist f^{-1} auf $\mathbb{R}^2 \setminus \{0\}$ nicht total differenzierbar. Noch weniger: f^{-1} ist auf $\mathbb{R}^2 \setminus \{0\}$ nicht einmal stetig! Dazu: Definiere für $n \in \mathbb{N}$ den Vektor $a_n := f(1, 2\pi - \frac{1}{n}) \in \mathbb{R}^2 \setminus \{0\}$. Dann gilt

$$a_n = f(1, 2\pi - \frac{1}{n}) = (\cos(2\pi - \frac{1}{n}), \sin(2\pi - \frac{1}{n}))^T \xrightarrow{n \rightarrow \infty} (\cos(2\pi), \sin(2\pi))^T = (1, 0)^T.$$

Jedoch gilt $f^{-1}(a_n) \xrightarrow{n \rightarrow \infty} f^{-1}(1, 0)$. In der Tat: Man hat $f^{-1}(1, 0) = (1, 0)^T$, da

$$f(1, 0) = (1 \cos(0), 1 \sin(0))^T = (1, 0)^T.$$

Aber

$$f^{-1}(a_n) = f^{-1}(f(1, 2\pi - \frac{1}{n})) = (1, 2\pi - \frac{1}{n})^T \xrightarrow{n \rightarrow \infty} (1, 2\pi)^T \neq (1, 0)^T.$$

Somit ist $(1, 0)^T$ eine Unstetigkeitsstelle von f^{-1} . Vergleiche auch mit Beispiel 10.4.16.

Eine bijektive und stetig partiell differenzierbare Abbildung zwischen zwei offenen Mengen D_1 und D_2 , deren Umkehrabbildung auch stetig partiell differenzierbar ist, nennt man übrigens *C^1 -Diffeomorphismus*.

10.8 Extrema unter Nebenbedingungen

In Beispiel 10.6.27 haben wir auf einer kompakten Menge $K \subseteq \mathbb{R}^2$ maximiert. Hierbei mussten wir den Rand ∂K auf Maxima und Minima untersuchen. Diese Untersuchung konnte auf ein eindimensionales Extremwertproblem zurückgeführt werden. Hierbei wurde massivst davon profitiert, dass ∂K durch die einfache Gleichung

$$\partial K = \{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 = 9\}$$

gegeben war. Diese Gleichung konnten wir nämlich (verhältnismäßig gut) umstellen und die sich daraus ergebende Bedingung in die Funktion einsetzen. Es gibt aber auch Mengen — gerade in höheren Dimensionen — wo diese Untersuchungen komplizierter sein können, weil der Rand z.B. durch mehrere Gleichungen ausgedrückt werden muss. Es gibt jedoch ein Theorem, mit dem sich solche Untersuchungen sehr elegant durchführen lassen — die *Multiplikatorenregel von Lagrange*. Wir wollen Sie im Folgenden besprechen.

Bevor wir einsteigen aber vielleicht ein ganz einfaches Anwendungsbeispiel für eine Extremwertaufgabe unter Nebenbedingungen. Haben wir ein Rechteck mit Seitenlängen $a, b > 0$ gegeben so ist der Flächeninhalt gleich $a \cdot b$ und der Umfang gleich $2a + 2b$. Will man nun das Rechteck mit Umfang 1 finden was den maximalen Flächeninhalt besitzt so maximiert man $f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(a, b) = a \cdot b$ unter der Nebenbedingung $2a + 2b = 1$. Im Folgenden würden wir diese Nebenbedingung wohl immer umschreiben als $2a + 2b - 1 = 0$.

Wir führen unsere Notation nun für allgemeinere Extremwertprobleme unter Nebenbedingungen ein.

Definition 10.8.1 — EXTREMWERTE UNTER NEBENBEDINGUNGEN

Es sei $D \subseteq \mathbb{R}^d$ offen und $f : D \rightarrow \mathbb{R}$ eine Funktion. Ferner seien $m \in \mathbb{N}$ und $b_1, \dots, b_m : D \rightarrow \mathbb{R}$ Funktionen. Sei nun $x_0 \in D$. Wir sagen

- (i) Wir sagen x_0 erfüllt die *Nebenbedingungen* $b_1, \dots, b_m$ falls $b_1(x_0) = b_2(x_0) = \dots = b_m(x_0) = 0$. Die Menge aller Punkte $x_0 \in D$, die die Nebenbedingungen $b_1, \dots, b_m$ erfüllen bezeichnen wir mit $D(b_1, \dots, b_m)$, d.h.

$$D(b_1, \dots, b_m) = \{x \in D : b_1(x) = \dots = b_m(x) = 0\}$$

- (ii) Wir sagen $x_0 \in D(b_1, \dots, b_m)$ ist eine *Maximumsstelle* unter den Nebenbedingungen $b_1, \dots, b_m$ falls

$$f(x) \leq f(x_0) \quad \forall x \in D(b_1, \dots, b_m).$$

- (iii) Wir sagen $x_0 \in D(b_1, \dots, b_m)$ ist eine *Minimumsstelle* unter den Nebenbedingungen $b_1, \dots, b_m$ falls

$$f(x) \geq f(x_0) \quad \forall x \in D(b_1, \dots, b_m).$$

- (iv) Wir sagen $x_0 \in D(b_1, \dots, b_m)$ ist eine *lokale Maximumsstelle* unter den Nebenbedingungen $b_1, \dots, b_m$ falls es ein $r > 0$ gibt mit

$$f(x) \leq f(x_0) \quad \forall x \in D(b_1, \dots, b_m) \cap B_r(x_0)$$

- (v) Wir sagen $x_0 \in D(b_1, \dots, b_m)$ ist eine *lokale Minimumsstelle* unter den Nebenbedingungen $b_1, \dots, b_m$ falls es ein $r > 0$ gibt mit

$$f(x) \geq f(x_0) \quad \forall x \in D(b_1, \dots, b_m) \cap B_r(x_0)$$

Ferner hat f eine (lokale) Extremstelle unter der Nebenbedingung $b_1, \dots, b_m$ wenn f eine (lokale) Maximumsstelle oder (lokale) Minimumsstelle von f ist.

Mit anderen Worten könnte man auch sagen: $f : D \rightarrow \mathbb{R}$ hat ein (lokales) Extremum unter den Nebenbedingungen $b_1, \dots, b_m$ wenn $f|_{D(b_1, \dots, b_m)}$ ein (lokales) Extremum hat.

Beispiel 10.8.2

Es sei $D = \mathbb{R}^2$ und

$$f : \mathbb{R}^2 \rightarrow \mathbb{R}, \quad f(x, y) = (x - 3)y^2 + x^2(x + 3)$$

Definiere auch $b_1 : \mathbb{R}^2 \rightarrow \mathbb{R}, b_1(x, y) := x^2 + y^2 - 9$. In Schritt 3 von Beispiel 10.6.27 haben wir nichts anderes gemacht also die Extremstellen von f unter der Nebenbedingung b_1 gesucht, denn für $D = \mathbb{R}^2$ gilt

$$\begin{aligned} D(b_1) &= \{(x, y)^T \in \mathbb{R}^2 : b_1(x, y) = 0\} = \{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 - 9 = 0\} \\ &= \{(x, y)^T \in \mathbb{R}^2 : x^2 + y^2 = 9\} \end{aligned}$$

Das ist genau die Menge über die wir in Schritt 3 von Beispiel 10.6.27 maximiert haben.

Wir können nun eine notwendige Bedingung für lokale Extrema formulieren, die *Multiplikatorenregel von Lagrange*. Die Formulierung wird zunächst ein wenig abschreckend wirken. Nachdem wir ein paar Korollare daraus formuliert haben, wird das ganze aber leicht zu überprüfen sein.

Satz 10.8.3 — BEOBACHTUNG VON LAGRANGE

Sei $D \subseteq \mathbb{R}^d$ offen. Es sei $f : D \rightarrow \mathbb{R}$ eine stetig partiell differenzierbare Funktion. Ferner sei $m < d$ und $b_1, \dots, b_m : D \rightarrow \mathbb{R}$ seien stetig partiell differenzierbare Funktionen. Ist $x_0 \in D(b_1, \dots, b_m)$ eine lokale Extremstelle von f unter den Nebenbedingungen $b_1, \dots, b_m$, so ist

$$\{\nabla f(x_0), \nabla b_1(x_0), \dots, \nabla b_m(x_0)\} \subseteq \mathbb{R}^d \quad \text{linear abhängig,}$$

d.h. es gibt $\theta_0, \dots, \theta_m \in \mathbb{R}$, nicht alle gleich 0, mit

$$\theta_0 \nabla f(x_0) + \theta_1 \nabla b_1(x_0) + \theta_2 \nabla b_2(x_0) + \dots + \theta_m \nabla b_m(x_0) = 0.$$

Beweis. Es sei $m < d$ (d.h. $m \leq d - 1$) und $b_1, \dots, b_m : D \rightarrow \mathbb{R}$ wie in der Aussage. Ferner sei $x_0 \in D(b_1, \dots, b_m)$ eine lokale Extremstelle von f unter den Nebenbedingungen $b_1, \dots, b_m$. ObdA sei x_0 ein lokales Maximum von f unter den Nebenbedingungen $b_1, \dots, b_m$, d.h.

$$\exists r > 0 : \quad f(x) \leq f(x_0) \quad \forall x \in B_r(x_0) \cap D(b_1, \dots, b_m). \quad (649)$$

Definiere $y_0 := f(x_0)$. Angenommen

$$\{\nabla f(x_0), \nabla b_1(x_0), \dots, \nabla b_m(x_0)\} \subseteq \mathbb{R}^d \quad \text{ist linear unabhängig} \quad (650)$$

Die obige Menge enthält dann $m + 1 \leq d$ linear unabhängige Vektoren. Man kann eine solche Menge nach dem *Basisergänzungssatz* aus der linearen Algebra zu einer *Basis*^a des $\mathbb{R}^d$ ergänzen. Das bedeutet man findet Vektoren $a_1, \dots, a_r \in \mathbb{R}^d$ ($r = d - (m + 1)$), leere Liste falls $d = m + 1$, so, dass die d -dimensionale Menge

$$\{\nabla f(x_0), \nabla b_1(x_0), \dots, \nabla b_m(x_0), a_1, \dots, a_r\} \subseteq \mathbb{R}^d \quad \text{linear unabhängig ist.} \quad (651)$$

Definiere nun

$$F : D \rightarrow \mathbb{R}^d, \quad F(x) := \begin{pmatrix} f(x) \\ b_1(x) \\ \vdots \\ b_m(x) \\ a_1^T(x - x_0) \\ \vdots \\ a_r^T(x - x_0) \end{pmatrix}. \quad (652)$$

ZWISCHENBEHAUPTUNG. F ist auf D stetig partiell differenzierbar und $\det(J_F(x_0)) \neq 0$.

Dass F stetig partiell differenzierbar ist folgt direkt daraus, dass jede Komponente von f nach Annahme stetig partiell differenzierbar ist. Ferner gilt

$$J_F(x_0) = \begin{pmatrix} \partial_1 f(x_0) & \partial_2 f(x_0) & \cdots & \partial_d f(x_0) \\ \partial_1 b_1(x_0) & \partial_2 b_1(x_0) & \cdots & \partial_d b_1(x_0) \\ \partial_1 b_2(x_0) & \partial_2 b_2(x_0) & \cdots & \partial_d b_2(x_0) \\ \vdots & & \ddots & \vdots \\ \partial_1 b_m(x_0) & \partial_2 b_m(x_0) & \cdots & \partial_d b_m(x_0) \\ & & a_1^T \\ & & a_2^T \\ & & \vdots \\ & & a_r^T \end{pmatrix}.$$

Man erhält

$$J_F(x_0) = \begin{pmatrix} \nabla f(x_0)^T \\ \nabla b_1(x_0)^T \\ \vdots \\ \nabla b_m(x_0)^T \\ a_1^T \\ \vdots \\ a_r^T \end{pmatrix}$$

Die Menge aller Zeilen von $J_F(x_0)$ ist also gegeben durch

$$\{\nabla f(x_0), \nabla b_1(x_0), \dots, \nabla b_m(x_0), a_1, \dots, a_r\}.$$

Nach (651) ist diese Menge linear unabhängig. Sind die Zeilen einer Matrix linear unabhängig, so ist sie invertierbar, vgl. 10.7.2. Somit ist $J_F(x_0)$ invertierbar und (wiederum nach Bemerkung 10.7.2) $\det(J_F(x_0)) \neq 0$.

ZWISCHENBEHAUPTUNG ENDE.

Nach dem lokalen Umkehrsatz (Satz 10.7.1) gibt es dann eine offene Menge $U_0 \subseteq D$ mit $x_0 \in U_0$ so, dass $V_0 = F(U_0)$ offen ist und $F : U_0 \rightarrow V_0$ eine stetig partiell differenzierbare Umkehrfunktion $(F|_{U_0})^{-1} : V_0 \rightarrow U_0$ besitzt. Beachte: Da $x_0 \in U_0$ gilt $F(x_0) \in F(U_0) = V_0$. Nun hat man nach (652)

$$F(x_0) = \begin{pmatrix} f(x_0) \\ b_1(x_0) \\ \vdots \\ b_m(x_0) \\ a_1^T(x_0 - x_0) \\ \vdots \\ a_r^T(x_0 - x_0) \end{pmatrix} = \begin{pmatrix} y_0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = y_0 e_1. \quad (653)$$

Daher ist $y_0 e_1 \in V_0$. Wegen der Offenheit von V_0 handelt es sich dabei auch um einen inneren Punkt von V_0 , d.h. es gibt ein $\theta > 0$ mit $B_\theta(y_0 e_1) \subseteq V_0$. Insbesondere gilt $(y_0 + \frac{1}{n})e_1 \in V_0$ für alle $n \in \mathbb{N}$ mit $\frac{1}{n} < \theta$ (d.h. $n > \frac{1}{\theta}$). Definiere nun für $n \in \mathbb{N}$ mit $n > \frac{1}{\theta}$

$$x_n := (F|_{U_0})^{-1}((y_0 + \frac{1}{n})e_1) \quad (b) \quad (654)$$

Da $(F|_{U_0})^{-1}$ auf V_0 stetig partiell diffbar ist, ist es nach Satz 10.5.30 auch total diffbar und deswegen nach Korollar 10.5.25 auch stetig. Wegen dieser Stetigkeit gilt

$$\lim_{n \rightarrow \infty} x_n = \lim_{(654) \ n \rightarrow \infty} (F|_{U_0})^{-1}((y_0 + \frac{1}{n})e_1) = (F|_{U_0})^{-1}(y_0 e_1).$$

Da nach (653) gilt $F(x_0) = y_0 e_1$ hat man $(F|_{U_0})^{-1}(y_0 e_1) = x_0$ und wir sehen, dass $\lim_{n \rightarrow \infty} x_n = x_0$. Wir folgern, dass es ein $n_0 \in \mathbb{N}$ gibt mit $x_{n_0} \in B_r(x_0)$, wobei r wie in (649) festgelegt ist (*). Nun gilt

$$F(x_{n_0}) = F((F|_{U_0})^{-1}((y_0 + \frac{1}{n})e_1)) = (y_0 + \frac{1}{n})e_1. \quad (655)$$

Daher hat man

$$\begin{pmatrix} y_0 + \frac{1}{n_0} \\ 0 \\ \vdots \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = (y_0 + \frac{1}{n})e_1 \stackrel{(655)}{=} F(x_{n_0}) \stackrel{(652)}{=} \begin{pmatrix} f(x_{n_0}) \\ b_1(x_{n_0}) \\ \vdots \\ b_m(x_{n_0}) \\ a_1^T(x_{n_0} - x_0) \\ \vdots \\ a_r^T(x_{n_0} - x_0) \end{pmatrix}. \quad (656)$$

Man folgert (durch Vergleich der Einträge in der 2-ten bis $m+1$ -sten Zeile), dass $b_1(x_{n_0}) = b_2(x_{n_0}) = \dots = b_m(x_{n_0}) = 0$. Mit anderen Worten hat man $x_{n_0} \in D(b_1, \dots, b_m)$. Wir folgern zusammen mit (*), dass $x_{n_0} \in B_r(x_0) \cap D(b_1, \dots, b_m)$. Wegen (649) folgt dann $f(x_{n_0}) \leq f(x_0)$. Vergleicht man aber die Einträge der ersten Zeile in (656) so erhält man

$$f(x_{n_0}) = y_0 + \frac{1}{n_0} = f(x_0) + \frac{1}{n_0} > f(x_0).$$

Ein Widerspruch zur Annahme (650). Somit folgt

$$\{\nabla f(x_0), \nabla b_1(x_0), \dots, \nabla b_m(x_0)\} \subseteq \mathbb{R}^d \quad \text{ist linear abhängig,}$$

was zu zeigen war.

^ad.h. einer d -elementigen linear unabhängigen Menge

^bDies ist zulässig weil $(y_0 + \frac{1}{n})e_1 \in V_0$ für alle $n \in \mathbb{N}$ mit $n > \frac{1}{\theta}$.

Um diesen Satz für Extremwertaufgaben nutzbar zu machen, benötigen wir zunächst ein Korollar.

Korollar 10.8.4 — LAGRANGE-MULTIPLIKATORENREGEL

Sei $D \subseteq \mathbb{R}^d$ offen. Es sei $f : D \rightarrow \mathbb{R}$ eine stetig partiell differenzierbare Funktion. Ferner sei $m < d$ und $b_1, \dots, b_m : D \rightarrow \mathbb{R}$ seien stetig partiell differenzierbare Funktionen. Ist $x_0 \in D(b_1, \dots, b_m)$ eine lokale Extremstelle von f unter den Nebenbedingungen $b_1, \dots, b_m$ und ist $\{\nabla b_1(x_0), \dots, \nabla b_m(x_0)\} \subseteq \mathbb{R}^d$ linear unabhängig, so gibt es $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ so, dass

$$\nabla f(x_0) = \sum_{i=1}^m \lambda_i \nabla b_i(x_0).$$

Beweis. Es sei $x_0 \in D$ wie in der Aussage. Nach Satz 10.8.3 gibt es $\theta_0, \dots, \theta_m \in \mathbb{R}$, nicht alle gleich 0, so, dass

$$\theta_0 \nabla f(x_0) + \theta_1 \nabla b_1(x_0) \dots + \theta_m \nabla b_m(x_0) = 0.$$

Das bedeutet

$$\theta_0 \nabla f(x_0) = -\theta_1 \nabla b_1(x_0) - \dots - \theta_m \nabla b_m(x_0) = 0.$$

FALL 1. $\theta_0 = 0$

Dann wäre $-\theta_1 \nabla b_1(x_0) - \dots - \theta_m \nabla b_m(x_0) = 0$. Das bedeutet (mit $\eta_j := -\theta_j$, $j = 1, \dots, m$), dass

$$\eta_1 \nabla b_1(x_0) + \dots + \eta_m \nabla b_m(x_0) = 0.$$

Wir folgern aus der Voraussetzung, dass $\{\nabla b_1(x_0), \dots, \nabla b_m(x_0)\}$ linear unabhängig ist, dass $\eta_1 = \dots = \eta_m = 0$. Da $\eta_j = -\theta_j$ für alle $j = 1, \dots, m$ folgt auch $\theta_1 = \dots = \theta_m = 0$. Wir erhalten in dem Fall also $\theta_0 = \theta_1 = \dots = \theta_m = 0$. Die erzeugt einen Widerspruch (denn die Zahlen $\theta_0, \dots, \theta_m$ waren nicht alle gleich 0). Fall 1 kann also nicht eintreten.

FALL 2. $\theta_0 \neq 0$

Dann kann man durch θ_0 teilen und erhält

$$\nabla f(x_0) = -\frac{\theta_1}{\theta_0} \nabla b_1(x_0) - \dots - \frac{\theta_m}{\theta_0} \nabla b_m(x_0).$$

Definieren wir nun $\lambda_1 := -\frac{\theta_1}{\theta_0}$, ..., $\lambda_m := -\frac{\theta_m}{\theta_0}$ so erhalten wir

$$\nabla f(x_0) = \lambda_1 \nabla b_1(x_0) + \dots + \lambda_m \nabla b_m(x_0) = \sum_{i=1}^m \lambda_i \nabla b_i(x_0).$$

Die Behauptung ist gezeigt.

Zum Abschluss ein Beispiel dazu, wie Korollar 10.8.4 konkret benutzt werden kann um Extremstellen zu berechnen.

Beispiel 10.8.5

Es sei $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ gegeben durch

$$f(x, y, z) = 4x + 4z - 5. \quad (657)$$

Ferner sei

$$M := \{(x, y, z)^T \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1, x + y + z = 1\}. \quad (658)$$

Siehe Abbildung 49 für eine Zeichnung von M . Man kann leicht zeigen, dass M kompakt ist. Da f auch stetig ist existieren (nach Satz 10.4.21) auch $\max f(M)$, $\min f(M)$.

$$\text{Behauptung.} \quad \max f(M) = \frac{1}{3}, \quad \min f(M) = -5.$$

Beweis. Definiere $b_1 : \mathbb{R}^3 \rightarrow \mathbb{R}$, $b_1(x, y, z) = x^2 + y^2 + z^2 - 1$, $b_2 : \mathbb{R}^3 \rightarrow \mathbb{R}$, $b_2(x, y, z) = x + y + z - 1$. Dann gilt

$$b_1(x, y, z) = b_2(x, y, z) = 0 \quad \Longleftrightarrow \quad x^2 + y^2 + z^2 = 1 \quad \text{und} \quad x + y + z = 1.$$

Man folgert daraus und aus (658): Extremstellen in M sind Extremstellen unter den Nebenbedingungen b_1, b_2 . Wir suchen also nun Extrema unter den Nebenbedingungen b_1, b_2 . Hierfür verwenden wir Korollar 10.8.4.^a

SCHRITT 1. Ausfindigmachen der Kandidaten.

Ist $(x_0, y_0, z_0)^T \in M$ eine lokale Extremstelle so gelten nach Korollar 10.8.4 folgende Gleichungen

$$\begin{cases} x_0^2 + y_0^2 + z_0^2 = 1 \\ x_0 + y_0 + z_0 = 1 \\ \nabla f(x_0, y_0, z_0) = \lambda_1 \nabla b_1(x_0, y_0, z_0) + \lambda_2 \nabla b_2(x_0, y_0, z_0) \end{cases} \quad \text{für gewisse } \lambda_1, \lambda_2 \in \mathbb{R}. \quad (b)$$

Nun gilt

$$\nabla f(x, y, z) = \begin{pmatrix} 4 \\ 0 \\ 4 \end{pmatrix}, \quad \nabla b_1(x, y, z) = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix}, \quad \nabla b_2(x, y, z) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Die letzte Gleichung im obigen Gleichungssystem schreibt sich daher um als

$$\begin{pmatrix} 4 \\ 0 \\ 4 \end{pmatrix} = 2\lambda_1 \begin{pmatrix} x_0 \\ y_0 \\ z_0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

Dies erlaubt uns das obige Gleichungssystem genauer aufzuschreiben: Wir suchen nach $x_0, y_0, z_0, \lambda_1, \lambda_2 \in \mathbb{R}$ so dass

$$\begin{cases} x_0^2 + y_0^2 + z_0^2 = 1 & \text{(I)} \\ x_0 + y_0 + z_0 = 1 & \text{(II)} \\ 4 = 2\lambda_1 x_0 + \lambda_2 & \text{(III)} \\ 0 = 2\lambda_1 y_0 + \lambda_2 & \text{(IV)} \\ 4 = 2\lambda_1 z_0 + \lambda_2 & \text{(V)} \end{cases}$$

Wir haben daher ein *nichtlineares^c Gleichungssystem* mit 5 Gleichungen für 5 Unbekannte. Dies lösen wir nun (zumindest so lange, bis wir x_0, y_0 und z_0 kennen. Die Werte von λ_1, λ_2 interessieren uns ja eigentlich nicht). Beachte: $\lambda_1 \neq 0$ denn wäre $\lambda_1 = 0$ so wäre nach (III) $4 = \lambda_2$ und nach (IV) $0 = \lambda_2$. Der Wert von λ_2 kann aber nicht gleichzeitig 4 und 0 sein. Wir können nun die Gleichng (V) von (III) abziehen und erhalten

$$4 - 4 = (2\lambda_1 x_0 + \lambda_2) - (2\lambda_1 z_0 + \lambda_2) \quad \text{d.h.} \quad 0 = 2\lambda_1(x_0 - z_0).$$

Da $\lambda_1 \neq 0$ folgt $x_0 - z_0 = 0$, d.h.

$$z_0 = x_0. \quad (660)$$

Aus Gleichung (II) lernen wir dann

$$y_0 = 1 - x_0 - z_0 = 1 - 2x_0. \quad (661)$$

Setzen wir (660) und (661) in (I) ein so erhalten wir

$$1 = x_0^2 + (1 - 2x_0)^2 + x_0^2 = 2x_0^2 + (1 - 4x_0 + 4x_0^2) = 6x_0^2 - 4x_0 + 1.$$

Ziehen wir auf beiden Seiten 1 ab so ergibt sich $0 = 6x_0^2 - 4x_0 = x_0(6x_0 - 4)$. Wir folgern

$$x_0 = 0 \quad \text{oder} \quad x_0 = \frac{2}{3}.$$

FALL 1. $x_0 = 0$.

Dann ist nach (660) und (661), d.h. $z_0 = x_0 = 0$ und $y_0 = 1 - 2x_0 = 1$. Somit gilt $(x_0, y_0, z_0)^T = (0, 1, 0)^T$.^d Wir checken kurz, dass der Punkt $(0, 1, 0)^T$ auch wirklich alle Nebenbedingungen erfüllt. In der Tat gilt $0^2 + 1^2 + 0^2 = 1$ und $0 + 1 + 0 = 1$.

FALL 2. $x_0 = \frac{2}{3}$.

Dann ist wiederum nach (660) und (661), d.h. $z_0 = x_0 = \frac{2}{3}$ und $y_0 = 1 - 2x_0 = 1 - \frac{4}{3} = -\frac{1}{3}$. Man erhält also $(x_0, y_0, z_0)^T = (\frac{2}{3}, -\frac{1}{3}, \frac{2}{3})^T$. Auch hier checken wir noch kurz, dass der Punkt $(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3})^T$ auch wirklich alle Nebenbedingungen erfüllt. Man sieht $(\frac{2}{3})^2 + (-\frac{1}{3})^2 + (\frac{2}{3})^2 = \frac{4}{9} + \frac{1}{9} + \frac{4}{9} = 1$. Auch gilt $\frac{2}{3} + (-\frac{1}{3}) + \frac{2}{3} = 1$.

SCHRITT 2. Vergleich der Kandidaten.

Wir haben die Kandidaten $(0, 1, 0)^T$ und $(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3})^T$ für lokale Extremstellen gefunden. Nun vergleichen wir die Funktionswerte. Nach (657) gilt

$$f(0, 1, 0) = 4 \cdot 0 + 4 \cdot 0 - 5 = -5,$$

$$f(\frac{2}{3}, -\frac{1}{3}, \frac{2}{3}) = 4 \cdot \frac{2}{3} + 4 \cdot \frac{2}{3} - 5 = \frac{16}{3} - 5 = \frac{1}{3}.$$

Da $\max f(M)$ und $\min f(M)$ angenommen werden müssen, können wir die Extremwerte unter den obigen Kandidaten suchen. Wir erhalten

$$\min f(M) = -5 \quad \text{und} \quad \max f(M) = \frac{1}{3},$$

was zu zeigen war.

^aVoraussetzung für die Anwendung der Multiplikatorenregel (Korollar 10.8.4) ist, dass $\{\nabla b_1(x_0, y_0, z_0), \nabla b_2(x_0, y_0, z_0)\}$ bei einer jeden Extremstelle $(x_0, y_0, z_0)^T \in M$ linear unabhängig sein müssen. Wir zeigen an dieser Stelle, dass $\{\nabla b_1(x, y, z), \nabla b_2(x, y, z)\}$ für alle $(x, y, z)^T \in M$ linear unabhängig ist. In der Tat: Die Vektoren

$$\nabla b_1(x, y, z) = \begin{pmatrix} 2x \\ 2y \\ 2z \end{pmatrix} \quad \text{und} \quad \nabla b_2(x, y, z) = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

sind linear abhängig genau dann wenn $x = y = z$. Ist $(x, y, z) \in M$ mit $x = y = z$ so gilt aber

$$1 = x^2 + y^2 + z^2 = 3x^2 \quad \text{und} \quad 1 = x + y + z = 3x. \quad (659)$$

Wir folgern also auch $3x^2 = 3x$, d.h. $0 = 3(x^2 - x) = 3x(x - 1)$. Also gilt entweder $x = 0$ oder $x = 1$. In beiden Fällen ist aber keine Gleichung aus (659) erfüllt. Ein Widerspruch.

^bNeben der Gleichung aus Korollar 10.8.4 sollte man an der Stelle nicht vergessen, die Nebenbedingungen mit aufzuschreiben. Sonst erfüllen die gefundenen Punkte eventuell nicht die Nebenbedingungen.

^cIch betone hier den Begriff *nichtlinear* um darauf hinzuweisen, dass man hier mit linearer Algebra nicht weiterkommt. Um ein solches nichtlineares Gleichungssystem analytisch zu lösen, muss man kreativ werden – einen allgemeinen Algorithmus zur Findung einer Lösung (wie bei linearen Gleichungssystemen) gibt es nicht. Es gibt aber gute Algorithmen zur Approximation solcher Lösungen. An dieser Stelle möchte ich auf die Numerik-Vorlesungen verweisen.

^dMan könnte an dieser Stelle noch herausfinden, welche Werte λ_1, λ_2 annehmen. Da das beim Funktionswert aber keine Rolle spielt, lassen wir es und geben uns mit dem Wissen über $(x_0, y_0, z_0)^T$ zufrieden.

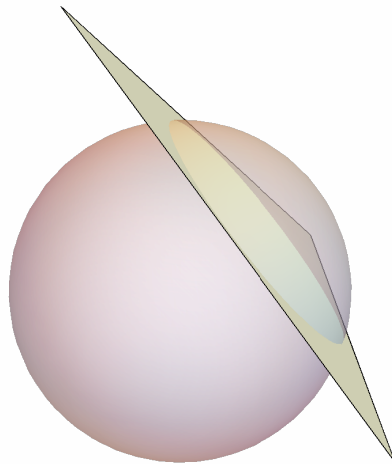


Abbildung 49: Die Menge M aus Beispiel 10.8.5 ist der Schnitt aus der rosa Sphäre mit der gelben Ebene.

11 Abstrakte Analysis

Am Anfang haben wir Analysis in den reellen Zahlen $\mathbb{R}$ oder komplexen Zahlen $\mathbb{C}$ betrieben. Im vorigen Kapitel haben wir uns überlegt, ob und wie wir die gelernten Konzepte (Folgen, Stetigkeit, Differentiation, ...) auch auf $\mathbb{R}^d$ verallgemeinern können. In unserer Realität müssen wir Analysis aber sogar auf noch allgemeineren Mengen betreiben: zum Beispiel auf *gekrümmten Flächen* wie der Erdoberfläche. Manchmal kommen auch *unendlichdimensionale Vektorräume* vor, auf denen wir Konzepte der Analysis besprechen müssen. Ziel dieses Kapitels ist, Ihnen Werkzeuge an die Hand zu geben, mit der sie die Analysis auf allgemeinere Mengen abstrahieren können. Anstatt $\mathbb{R}$, $\mathbb{C}$ oder $\mathbb{R}^d$ befinden wir uns hier also nun auf einer allgemeinen Menge X . Auf dieser Menge X haben wir zunächst keine *Rechengesetze* zur Verfügung. Trotzdem werden wir gewisse minimale Strukturanforderungen, die uns die Verallgemeinerung unserer Rechengesetze erlaubt.

11.1 Topologische Räume

Ein topologischer Raum ist eine Menge X , bei der wir bestimmte Teilmengen als *offen* bezeichnen dürfen. Für diese offenen Mengen müssen bestimmte Gesetze gelten, die wir bereits aus dem $\mathbb{R}^d$ -Fall kennen. Genauer:

Definition 11.1.1 — TOPOLOGIE

Es sei X eine nichtleere Menge und $\mathcal{O}$ eine Menge von Teilmengen von X (d.h. $\mathcal{O} \subseteq \mathcal{P}(X)$). Wir nennen $\mathcal{O}$ eine *Topologie* auf X falls

(T1) $X \in \mathcal{O}$,

(T2) Ist $(U_i)_{i \in I} \subseteq \mathcal{O}$ eine Familie von Mengen in $\mathcal{O}$ so ist die Vereinigungsmenge $\bigcup_{i \in I} U_i$ wieder ein Element von $\mathcal{O}$,

(T3) Ist $(U_i)_{i \in I_0} \subseteq \mathcal{O}$ eine Familie von Mengen in $\mathcal{O}$ und I_0 zusätzlich endlich, so ist die Schnittmenge $\bigcap_{i \in I_0} U_i$ wieder ein Element von $\mathcal{O}$.

Elemente von $\mathcal{O}$ nennt man auch *offene Mengen in X* ^a. Das Paar $(X, \mathcal{O})$ nennen wir *topologischer Raum*.

^aDie Bezeichnung *offen in X* ist eigentlich etwas irreführend, denn auf X kann es ja mehr als nur eine einzige Topologie geben. Besser würde man wohl schreiben „offen in $(X, \mathcal{O})$ “

Die Axiome (T1), (T2) und (T3) bilden gewissermaßen die Minimalanforderungen, die wir benötigen um Konzepte der Topologie (vgl. Sektion 10.3) auf beliebige Mengen X verallgemeinern zu können.

Bemerkung 11.1.2

Es sei $(X, \mathcal{O})$ ein topologischer Raum. Mit der Sprechweise, dass Elemente von $\mathcal{O}$ offene Mengen in X heißen lassen sich (T2) und (T3) wie folgt reformulieren

(T2') Die Vereinigung beliebig vieler offener Mengen in X ist wieder offen.

(T3') Der Schnitt endlich vieler offener Mengen in X ist wieder offen.

Die Aussagen (T2') und (T3') erinnern stark an Proposition 10.3.13. Aus dieser Proposition heraus erhalten wir sofort ein Beispiel für einen topologischen Raum.

Beispiel 11.1.3

Es sei $d \in \mathbb{N}$. Betrachte $X = \mathbb{R}^d$ und

$$\mathcal{O} := \{U \subseteq \mathbb{R}^d : U \text{ ist offen im Sinne von Definition 10.3.5}\}.$$

Auch könnte man schreiben

$$\mathcal{O} := \{U \subseteq \mathbb{R}^d : \forall x \in U \exists r > 0 : B_r(x) \subseteq U\}.$$

Behauptung. $(X, \mathcal{O})$ ist ein topologischer Raum.

Beweis. Wir prüfen (T1), (T2) und (T3) aus Definition 11.1.1 nach.

Zu (T1). $\mathbb{R}^d$ ist offen im Sinne von Definition 10.3.5, vgl. Beispiel 10.3.9.

Zu (T2). Dass die Vereinigung beliebig vieler offener Mengen in $\mathbb{R}^d$ wiederum offen ist folgt aus Proposition 10.3.13(i).

Zu (T3). Dass der Schnitt endlich vieler offener Mengen in $\mathbb{R}^d$ wiederum offen ist folgt aus Proposition 10.3.13.

Wir können uns auch noch mehr Beispiele topologischer Räume basteln

Beispiel 11.1.4 — RELATIVTOPOLOGIE

Es sei $S \subseteq \mathbb{R}^d$ beliebig. Definiere die sogenannte *Relativtopologie*^a auf S durch

$$\mathcal{O}_S := \{S \cap U : U \subseteq \mathbb{R}^d \text{ offen im Sinne von Definition 10.3.5}\} \quad (662)$$

Offene Mengen in S (d.h. Elemente von $\mathcal{O}_S$) nennen wir manchmal auch *relativoffen* in S . Solche bestehen also stets aus Durchschnitten von S mit einer offenen Menge $U \subseteq \mathbb{R}^d$.

Behauptung. $(S, \mathcal{O}_S)$ ist ein topologischer Raum.

Beweis. Wir prüfen wiederum (T1), (T2) und (T3) aus Definition 11.1.1 nach.

Zu (T1). Wir müssen zeigen dass $S \in \mathcal{O}_S$. Dazu beobachte, dass

$$S = S \cap \mathbb{R}^d.$$

Nun ist $\mathbb{R}^d$ offen im Sinne von Definition 10.3.5. Nach (662) liegt $S \cap \mathbb{R}^d$ dann in $\mathcal{O}_S$. Also folgt $S \in \mathcal{O}_S$ wie zu zeigen war.

Zu (T2). Es sei $(U_i)_{i \in I} \subseteq \mathcal{O}_S$ eine Familie von Mengen in $\mathcal{O}_S$. Das bedeutet, dass $U_i \in \mathcal{O}_S$ für alle $i \in I$. Somit gibt es zu jedem $i \in I$ eine Menge $V_i \subseteq \mathbb{R}^d$, die im Sinne von Definition 10.3.5 offen ist und

$$U_i = S \cap V_i \tag{663}$$

erfüllt. Zu zeigen ist, dass $\bigcup_{i \in I} U_i \in \mathcal{O}_S$. Dazu berechne

$$\bigcup_{i \in I} U_i \stackrel{(663)}{=} \bigcup_{i \in I} (S \cap V_i) \stackrel{\text{Übungsaufgabe}}{=} S \cap \left(\bigcup_{i \in I} V_i \right). \tag{664}$$

Nach Proposition 10.3.13(i) ist $\bigcup_{i \in I} V_i$ offen im Sinne von Definition 10.3.5. Nach (662) liegt dann $S \cap \left(\bigcup_{i \in I} V_i \right)$ in $\mathcal{O}_S$. Die Behauptung folgt dann aus (664).

Zu (T3). Es sei $(U_i)_{i \in I_0}$ eine endliche Familie von Mengen in $\mathcal{O}_S$ (d.h. I_0 ist endlich). Wiederum gilt dann $U_i \in \mathcal{O}_S$ für alle $i \in I_0$, weswegen es zu jedem $i \in I_0$ eine Menge $V_i \subseteq \mathbb{R}^d$ gibt, die im Sinne von Definition 10.3.5 offen ist und

$$U_i = S \cap V_i$$

erfüllt. Dann gilt

$$\bigcap_{i \in I_0} U_i = \bigcap_{i \in I_0} (S \cap V_i) = S \cap \left(\bigcap_{i \in I_0} V_i \right). \tag{665}$$

Da I_0 endlich ist, ist nach Proposition 10.3.13(ii) $\bigcap_{i \in I_0} V_i$ offen im Sinne von Definition 10.3.5.

Deswegen gilt nach (662), dass $S \cap \left(\bigcap_{i \in I_0} V_i \right) \in \mathcal{O}_S$. Mit (665) folgt die Behauptung.

^aDer Begriff ist zu dem Zeitpunkt ein wenig unangemessen, weil wir noch nicht gezeigt haben, dass es sich tatsächlich um eine Topologie auf S handelt. Das werden wir aber noch innerhalb dieses Beispiels nachholen.

Die Relativtopologie kann gut dafür verwendet werden, um die Topologie *gekrümmter Oberflächen in $\mathbb{R}^d$* zu verstehen. Ein Beispiel ist die Relativtopologie auf der *Sphäre*

$$S = \mathbb{S} := \{x \in \mathbb{R}^d : \|x\| = 1\}.$$

Was sind nun offene Mengen auf $\mathbb{S}$? Ein Beispiel ist die *obere Halbsphäre*

$$\mathbb{S}_+ := \{x \in \mathbb{R}^d : \|x\| = 1, x^{(d)} > 0\},$$

denn diese lässt sich schreiben als

$$\mathbb{S}_+ = \mathbb{S} \cap \{x \in \mathbb{R}^d : x^{(d)} > 0\}.$$

Da $U := \{x \in \mathbb{R}^d : x^{(d)} > 0\} \subseteq \mathbb{R}^d$ offen in $\mathbb{R}^d$ im Sinne von Definition 10.3.5 ist, ist nach (662)

$$\mathbb{S}_+ \in \mathcal{O}_{\mathbb{S}},$$

d.h. $\mathbb{S}_+$ ist (relativ-)offen in $\mathbb{S}$. Beachte: $\mathbb{S}^+$ ist aber nicht offen in $\mathbb{R}^d$! In der Tat: Es gilt $N = (0, \dots, 0, 1)^T \in \mathbb{S}_+$ aber man findet kein $r > 0$ mit $B_r(N) \subseteq \mathbb{S}_+$.

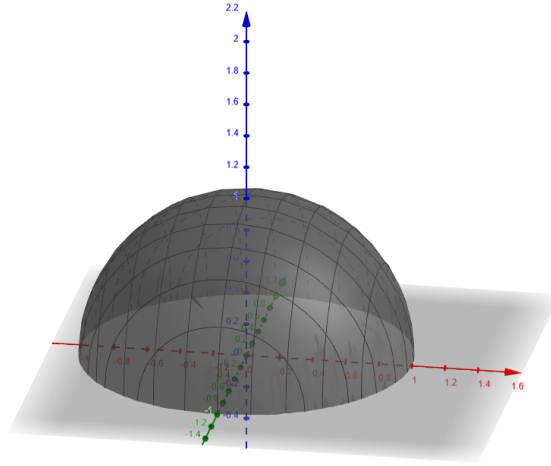


Abbildung 50: Die obere Halbsphäre $\mathbb{S}_+$ in $\mathbb{R}^3$.

Beispiel 11.1.5

Neben den vorhin besprochenen Topologien gibt es aber auch ganz pathologische topologische Räume. Ein Beispiel ist $X = \{1, 2, 3\}$ mit der Topologie

$$\mathcal{O} := \{\emptyset, \{1, 2\}, \{3\}, \{1, 2, 3\}\}. \quad (666)$$

Behauptung. $(X, \mathcal{O})$ ist ein topologischer Raum.

Beweis. Wir prüfen wiederum (T1), (T2) und (T3) aus Definition 11.1.1 nach. Zu (T1). Dass $X = \{1, 2, 3\} \in \mathcal{O}$ liest man direkt aus der Mengenbeschreibung von $\mathcal{O}$ heraus. Zu (T2). Es sei $(U_i)_{i \in I} \subseteq \mathcal{O}$ eine Familie von Mengen aus $\mathcal{O}$. Zu zeigen ist, dass $\bigcup_{i \in I} U_i \in \mathcal{O}$. Wir dürfen bei der Vereinigungsbildung alle $i \in I$ mit $U_i = \emptyset$ ignorieren. Also nehmen wir an, dass $U_i \neq \emptyset$ für alle $i \in I$. Sind alle Mengen U_i gleich (etwa $U_i = M$ für ein $M \in \mathcal{O}$), so ist die Aussage sicher zutreffend, denn dann gilt auch $\bigcup_{i \in I} U_i = M \in \mathcal{O}$. Kommen in der Familie $(U_i)_{i \in I}$ zwei verschiedene nichtleere Mengen vor, so muss wegen (666) bereits

$\bigcup_{i \in I} U_i = \{1, 2, 3\}$ gelten. Dies folgt aus

$$\begin{aligned}\{1, 2\} \cup \{3\} &= \{1, 2, 3\}, \\ \{1, 2\} \cup \{1, 2, 3\} &= \{1, 2, 3\}, \\ \{3\} \cup \{1, 2, 3\} &= \{1, 2, 3\}.\end{aligned}$$

Da $\{1, 2, 3\} \in \mathcal{O}$ folgt auch $\bigcup_{i \in I} U_i \in \mathcal{O}$. (T3) überlassen wir den Lesenden als Übungsaufgabe.

Stetigkeit

Wir erinnern uns aus Satz 10.4.14, dass sich Stetigkeit in $\mathbb{R}^d$ rein topologisch charakterisieren lässt. Diese Charakterisierung bildet die Grundlage für unsere Definition einer stetigen Abbildung zwischen zwei topologischen Räumen.

Definition 11.1.6

Definition 11.1.1. STETIGKEIT IN TOPOLOGISCHEN RÄUMEN. Es seien $(X_1, \mathcal{O}_1)$, $(X_2, \mathcal{O}_2)$ topologische Räume. Eine Abbildung $f : X_1 \rightarrow X_2$ heißt *stetig*, wenn für alle $U \in \mathcal{O}_2$ gilt, dass $f^{-1}(U) \in \mathcal{O}_1$. Mit anderen Worten: Ist U offen in X_2 so muss das Urbild $f^{-1}(U)$ offen in X_1 sein.

Beispiel 11.1.2. Es sei $X_1 = \mathbb{R}^d$ und $X_2 = \mathbb{R}^m$, beide jeweils ausgestattet mit der Topologie aus Beispiel 11.1.3. Ist $f : X_1 \rightarrow X_2$ stetig, so bedeutet das

$$\text{Für alle offenen Mengen } U \subseteq \mathbb{R}^m \text{ ist } f^{-1}(U) \subseteq \mathbb{R}^d \text{ offen in } \mathbb{R}^d.$$

Dass dies Äquivalent zur herkömmlichen Stetigkeit nach Definition 10.4.7 ist, haben wir in Satz 10.4.14 bewiesen.

Das obige Beispiel zeigt die *Konsistenz* der obigen Stetigkeitsdefinition mit Kapitel 10.4. Es gibt aber noch weitere Beispiele.

Beispiel 11.1.7

Beispiel 11.1.2. Wir betrachten wieder $X = \{1, 2, 3\}$ mit der Topologie aus Beispiel 11.1.5, d.h.

$$\mathcal{O} = \{\emptyset, \{1, 2\}, \{3\}, \{1, 2, 3\}\}$$

Definiere nun $f_1, f_2 : X \rightarrow \mathbb{R}^a$ durch

$$f_1(1) = 0, f_1(2) = -1, f_1(3) = 1 \quad \text{sowie} \quad f_2(1) = 0, f_2(2) = 0, f_2(3) = 1.$$

Behauptung. f_1 ist unstetig und f_2 ist stetig.

Beweis. Zuerst zur Unstetigkeit von f_1 . Wir weisen diese nach, indem wir eine offene Menge U_* finden, deren Urbild $f^{-1}(U_*)$ nicht offen ist. Dann kann die Bedingung aus Definition 11.1.1 nicht mehr erfüllt sein. Dazu: Das Intervall $(-\frac{1}{2}, \frac{1}{2}) = B_{\frac{1}{2}}(0)$ ist offen in $\mathbb{R}$ (vgl. Beispiel 10.3.7). Jedoch ist klarerweise

$$f_1^{-1}((-\frac{1}{2}, \frac{1}{2})) = \{x \in X : f_1(x) \in (-\frac{1}{2}, \frac{1}{2})\} = \{1\} \notin \mathcal{O}.$$

Das Urbild der offenen Menge $U_* = (-\frac{1}{2}, \frac{1}{2})$ ist also nicht offen in X . Die Unstetigkeit folgt.

Nun zur Stetigkeit von f_2 . Es sei $U \subseteq \mathbb{R}$ offen. Dann gilt

$$f_2^{-1}(U) = \begin{cases} \{1, 2\} & 0 \in U, 1 \notin U \\ \{3\} & 0 \notin U, 1 \in U \\ \{1, 2, 3\} & 0 \in U, 1 \in U \\ \emptyset & 0 \notin U, 1 \notin U \end{cases}.$$

In allen vier Fällen erhält man Mengen die in $\mathcal{O}$ liegen. Das zeigt dass $f_2^{-1}(U) \in \mathcal{O}$ für alle $U \subseteq \mathbb{R}$ offen. Nach Definition 11.1.1 ist f_2 also stetig.

^aAuf $\mathbb{R}$ betrachten wir hier die übliche Topologie aus Beispiel 11.1.3

Kompaktheit

Auch möchten wir das Konzept der Kompaktheit auf topologische Räume verallgemeinern. Hierbei lassen wir uns von der Überdeckungskompaktheit (Definition 10.3.25) inspirieren.

Definition 11.1.8 — (TOPOLOGISCHE) KOMPAKTHEIT

Es sei $(X, \mathcal{O})$ ein topologischer Raum und $K \subseteq X$. Wir sagen K ist (*topologisch*) *kompakt* in X wenn es zu jeder Familie $(U_i)_{i \in I} \subseteq \mathcal{O}$ offener Mengen in X mit

$$K \subseteq \bigcup_{i \in I} U_i$$

eine endliche Menge $I_0 \subseteq I$ gibt mit

$$K \subseteq \bigcup_{i \in I_0} U_i.$$

In Worten: Jede (in $(X, \mathcal{O})$) offene Überdeckung von K^a besitzt eine endliche Teilüberdeckung.

^aHierbei heißt eine Familie $(U_i)_{i \in I} \subseteq \mathcal{O}$ von offenen Mengen eine *offene Überdeckung* von K falls $K \subseteq \bigcup_{i \in I} U_i$.

Beispiel 11.1.9

Ist $X = \mathbb{R}^d$ ausgestattet mit der Topologie von Beispiel 11.1.3, so ist $K \subseteq \mathbb{R}^d$ topologisch kompakt genau dann wenn $K \subseteq \mathbb{R}^d$ überdeckungskompakt ist (vgl. Definition 10.3.25). Der Satz von Heine-Borel (Satz 10.3.27) besagt dann

$$K \subseteq \mathbb{R}^d \text{ topologisch kompakt} \iff K \subseteq \mathbb{R}^d \text{ abgeschlossen und beschränkt.}$$

Auf allgemeine topologische Räume kann dieses Resultat nicht übertragen werden. Denn in einem allgemeinen topologischen Raum $(X, \mathcal{O})$ macht zum Beispiel der Begriff der Beschränktheit überhaupt keinen Sinn!

Der Begriff der Kompaktheit ist nun auch etwas *rein topologisches*. Das heißt natürlich auch, dass verschiedene Topologien stets verschiedene kompakte Mengen liefern.

Beispiel 11.1.10

Betrachte die Menge $X = \mathbb{R}$. Es sei $\mathcal{O}_1$ die Topologie aus Beispiel 11.1.3 und $\mathcal{O}_2 := \{\emptyset, X\}^a$.

Behauptung. $\mathbb{N}$ ist nicht kompakt in $(X, \mathcal{O}_1)$ aber kompakt in $(X, \mathcal{O}_2)$.

Beweis. Mengen, die topologisch kompakt in $(X, \mathcal{O}_1)$ sind, sind nach Beispiel 11.1.9 insbesondere beschränkt in $\mathbb{R}$. Das trifft auf $\mathbb{N}$ nicht zu. Also ist $\mathbb{N}$ nicht kompakt in $(X, \mathcal{O}_1)$.

Um zu zeigen, dass $\mathbb{N}$ bezüglich $(X, \mathcal{O}_2)$ kompakt ist, gehen wir streng nach Definition 11.1.8 vor. Es sei $(U_i)_{i \in I} \subseteq \mathcal{O}_2$ eine offene Überdeckung von $\mathbb{N}$, d.h.

$$\mathbb{N} \subseteq \bigcup_{i \in I} U_i.$$

Da $\mathbb{N} \neq \emptyset$ muss mindestens eine der Mengen $U_i \in \mathcal{O}_2$ nichtleer sein. Sei etwa $i_0 \in I$ so, dass $U_{i_0} \neq \emptyset$. Da $\mathcal{O}_2$ aber nur aus $\emptyset$ und X besteht muss gelten, dass $U_{i_0} = X = \mathbb{R}$. Nun gilt aber

$$\mathbb{N} \subseteq \mathbb{R} = U_{i_0} = \bigcup_{i \in \{i_0\}} U_i.$$

Wählt man nun also $I_0 = \{i_0\}$ (was klarerweise endlich ist) so hat man eine endliche Teilüberdeckung gefunden.

^aMan überzeugt sich sehr leicht mit Definition 11.1.1 davon, dass $\mathcal{O}_2$ eine Topologie auf X ist

Wir haben gesehen: Das Konzept eines topologischen Raumes erlaubt eine Verallgemeinerung der Begriffe der Stetigkeit und der Kompaktheit. An dieser Stelle muss man sich auch fragen, ob sich Resultate über diese Begriffe auch auf das Setting topologischer Räume übertragen.

Ein wichtiges Resultat, was sich tatsächlich überträgt, ist die *Kompaktheitserhaltung unter stetigen Funktionen* (vgl. Satz 10.4.20) .

Satz 11.1.11 — ERHALTUNG VON KOMPAKTHEIT UNTER STETIGEN ABBILDUNGEN

Es seien $(X_1, \mathcal{O}_1)$ und $(X_2, \mathcal{O}_2)$ topologische Räume und $f : X_1 \rightarrow X_2$ eine stetige Funktion. Ist $K \subseteq X_1$ kompakt, so ist $f(K) \subseteq X_2$ kompakt.

Beweis. Wir können nicht wie im Beweis von Satz 10.4.20 argumentieren, denn dort wird die *folgenkompaktheit* verwendet. Auf topologischen Räumen haben wir aber das Konzept einer Folge gar nicht zur Verfügung (und schon gar nicht das Konzept der Konvergenz)^a. Deswegen müssen wir direkt mit Definition 11.1.8 arbeiten. Wir weisen mit dieser die topologische Kompaktheit von $f(K) \subseteq X_2$ nach. Dazu sei $(U_i)_{i \in I} \subseteq \mathcal{O}_2$ eine offene Überdeckung von $f(K)$, d.h.

$$f(K) \subseteq \bigcup_{i \in I} U_i.$$

Betrachtet man auf beiden Seiten das Urbild so erhält man

$$f^{-1}(f(K)) \subseteq f^{-1}\left(\bigcup_{i \in I} U_i\right) \stackrel{\text{Übungsaufgabe}}{=} \bigcup_{i \in I} f^{-1}(U_i) \quad (667)$$

Nun gilt nach Proposition 1.3.7

$$K \subseteq f^{-1}(f(K)) \stackrel{(667)}{\subseteq} \bigcup_{i \in I} f^{-1}(U_i). \quad (668)$$

Beachte: Da f stetig ist, ist nach Definition 11.1.1 $f^{-1}(U_i) \in \mathcal{O}_1$ für alle $i \in I$. Wegen (668) ist $(f^{-1}(U_i))_{i \in I} \subseteq \mathcal{O}_1$ also eine offene Überdeckung von K . Da nun $K \subseteq X_1$ kompakt ist, gibt es nach Definition 11.1.8 ein $I_0 \subseteq I$ endlich mit

$$K \subseteq \bigcup_{i \in I_0} f^{-1}(U_i).$$

Wir erhalten dann

$$f(K) \subseteq f\left(\bigcup_{i \in I_0} f^{-1}(U_i)\right) = \bigcup_{i \in I_0} f(f^{-1}(U_i))$$

Man überlegt sich nun leicht, dass $f(f^{-1}(U_i)) = U_i$ für alle $i \in I_0$. Das zeigt also

$$f(K) \subseteq \bigcup_{i \in I_0} U_i.$$

Wir haben somit die endliche Teilüberdeckung $(U_i)_{i \in I_0}$ gefunden.

^aAn dieser Stelle sei gesagt, dass es Möglichkeiten gibt, Konvergenz einer Folge in einem topologischen Raum zu definieren. Dort treten dann aber einige kontraintuitive Eigenschaften der Folgenkonvergenz auf,

weswegen der Begriff keine große Bedeutung in der Topologie hat. Auch der Begriff der Stetigkeit kann dann zum Beispiel nicht mithilfe von Folgen charakterisiert werden. Daher ist der Beweis von Satz 10.4.20 wirklich nicht kopierbar!

Diese Einführung in topologische Räume war sehr kurz. Mehr dazu lernen Sie aber sicherlich in der Vorlesung *Einführung in die Topologie*.

11.2 Metrische Räume

Nicht alle Konzepte, die wir in der Analysis besprochen haben übertragen sich gut auf topologische Räume. Insbesondere bei Folgenkonvergenz bzw. Cauchy-Folgen gibt es Ungeheimheiten. Deswegen untersuchen wir im folgenden *metrische Räume*, welche mehr Struktur haben. Hier wird die Topologie durch eine *Metrik* beschrieben. Eine Metrik ist hierbei ein Tool, um *Abstände* zu messen. Auf $\mathbb{R}^d$ haben wir bereits eine *Norm* kennengelernt, mithilfe derer man auch Abstände messen kann. Daher müssen wir nach diesem Kapitel auch den Zusammenhang zwischen Norm und Metrik diskutieren.

Definition 11.2.1 — METRISCHER RAUM

Es sei X eine nichtleere Menge. Eine Abbildung $d : X \times X \rightarrow \mathbb{R}$ heißt *Metrik auf X* falls gilt

(M1) *Definitheit*. $d(x, y) \geq 0$ für alle $x, y \in X$ und $(d(x, y) = 0 \iff x = y)$.

(M2) *Symmetrie*. Für alle $x, y \in X$ gilt $d(x, y) = d(y, x)$.

(M3) *Dreiecksungleichung*. Für alle $x, y, z \in X$ gilt $d(x, z) \leq d(x, y) + d(y, z)$.

Das Paar (X, d) ist nennen wir einen metrischen Raum. Die Zahl $d(x, y)$ wird auch d -Abstand zwischen x und y genannt.

Bemerkung 11.2.2

Wir interpretieren die Axiome (M1), (M2), (M3) aus Definition 11.2.1 mit dem Hintergrund, dass $d(x, y)$ eine Art *Abstand* zwischen x und y ist.

- (M1) besagt, dass ein Abstand zwischen zwei Punkten immer echt positiv sein soll, sofern die beiden Punkte nicht gleich sind.
- (M2) besagt dass der Abstand zwischen x und y derselbe sein soll wie der Abstand zwischen y und x .
- Schließlich besagt (M3), dass der Abstand von zwei Punkten x und z nicht kürzer gemacht werden kann, indem man erst den Abstand von Punkt x nach y misst und dann den Abstand von y nach z addiert.

Bemerkung 11.2.3

Auf die Forderung in (M1), dass „ $d(x, y) \geq 0$ für alle $x, y \in X$ “ kann verzichtet werden. Präziser:

(M1) kann durch die schwächere Voraussetzung

$$(M1') \quad d(x, y) = 0 \iff x = y$$

ersetzt werden.

Das beweisen wir kurz. Es erfülle eine Abbildung $\tilde{d} : X \times X$ die Axiome (M2) und (M3) aus Definition 11.2.1 und zusätzlich (M1')

$\mathbb{Z}$: $\tilde{d}$ ist eine Metrik auf X .

Hierzu fehlt nur noch ein Teil von (M1), nämlich eben $\tilde{d}(x, y) \geq 0$ für alle $x, y \in X$. Dazu berechnen wir für beliebige $x, y \in X$

$$0 \underset{(M1')}{=} \tilde{d}(x, x) \underset{(M3)}{\leq} \tilde{d}(x, y) + \tilde{d}(y, x) \underset{(M2)}{=} 2\tilde{d}(x, y).$$

Teilen wir durch 2 so ergibt sich $\tilde{d}(x, y) \geq 0$, was zu zeigen war.

Nun sehen wir einige Beispiele von metrischen Räumen.

Beispiel 11.2.4 — EUKLIDISCHE METRIK

Betrachte $X = \mathbb{R}^d$ und $d_{\text{euc}} : X \times X \rightarrow \mathbb{R}$ gegeben durch

$$d_{\text{euc}}(x, y) := \|x - y\|. \quad (669)$$

Behauptung. d_{euc} ist eine Metrik auf X (genannt *euklidische Metrik*)

Beweis. Wir prüfen die Axiome (M1), (M2), (M3) aus Definition 11.2.1 nach. Hierbei ist Satz 10.1.2 eine große Hilfe. Zu (M1). Seien $x, y \in \mathbb{R}^d$

$$d_{\text{euc}}(x, y) = \|x - y\| \geq 0 \quad \text{gilt nach Satz 10.1.2 (i).}$$

Außerdem gilt

$$d_{\text{euc}}(x, y) = 0 \iff \|x - y\| = 0 \underset{\text{Satz 10.1.2(i)}}{\iff} x - y = 0 \iff x = y.$$

Nun zu (M2). Für $x, y \in X$ gilt

$$d_{\text{euc}}(x, y) \stackrel{(669)}{=} \|x - y\| = \|(-1)(y - x)\| \stackrel{\text{Satz 10.1.2(ii)}}{=} |(-1)| \|y - x\| = \|y - x\| \stackrel{(669)}{=} d_{\text{euc}}(y, x).$$

Auch (M3) kann aus Satz 10.1.2 gefolgert werden. Für $x, y, z \in X$ gilt

$$\begin{aligned} d_{\text{euc}}(x, z) &= \|x - z\| = \|x - y + y - z\| = \|(x - y) + (y - z)\| \\ &\stackrel{\text{Satz 10.1.2(iii)}}{\leq} \|x - y\| + \|y - z\| = d_{\text{euc}}(x, y) + d_{\text{euc}}(y, z). \end{aligned}$$

Die Behauptung ist gezeigt.

Wiederum findet man auch pathologischere Beispiele von Metriken.

Beispiel 11.2.5 — DISKRETE METRIK

Es sei X eine beliebige nichtleere Menge und $d : X \times X \rightarrow \mathbb{R}$ gegeben durch

$$d(x, y) := \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}. \quad (670)$$

Behauptung. d ist eine Metrik auf X .

Beweis. (M1) und (M2) sind den Lesenden als Übungsaufgabe überlassen. Zu (M3). Seien $x, y, z \in X$

$$\mathbb{Z}: \quad d(x, z) \leq d(x, y) + d(y, z).$$

Falls $x = z$ ist die linke Seite gleich 0 und daher ist die Aussage sicher erfüllt. Falls hingegen $x \neq z$ so ist nach (670) $d(x, z) = 1$. Gleichsam kann aber auch y nicht gleichzeitig mit x und z übereinstimmen. Das bedeutet, dass mindestens einer der Ausdrücke $d(x, y)$ oder $d(y, z)$ gleich 1 sein muss. Wir folgern, dass $d(x, y) + d(y, z) \geq 1$ und somit die Behauptung.

Beispiel 11.2.6

Es sei (X, d) ein metrischer Raum und $A \subseteq X$. Dann kann man auf A ganz leicht eine Metrik $d_A : A \times A \rightarrow \mathbb{R}$ definieren, nämlich durch

$$d_A(x, y) := d(x, y) \quad \forall x, y \in A.$$

Es ist nicht schwer nachzuweisen, dass d_A auch die Axiome (M1), (M2), (M3) aus 11.2.1 erfüllt. (Die Eigenschaften *erben sich* von den Eigenschaften von d). Anstatt d_A schreiben wir im Folgenden häufig auch einfach d .

Wichtig in der höheren Analysis sind *Funktionenräume*, welche wir im Folgenden auch mit einer Metrik ausstatten können.

Definition 11.2.7 — FUNKTIONENRÄUME

Es sei $D \subseteq \mathbb{R}^d$ eine Menge und $k \in \mathbb{N}$. Wir definieren die folgenden Mengen

- (i) $B(D, \mathbb{R}^m) := \{f : D \rightarrow \mathbb{R}^m \mid f \text{ ist beschränkt, d.h. } \exists L > 0: \|f(x)\| \leq L \forall x \in D\}.$
- (ii) $C^0(D, \mathbb{R}^m) := \{f : D \rightarrow \mathbb{R}^m \mid f \text{ ist auf } D \text{ stetig}\}.$
- (iii) $C_b^0(D, \mathbb{R}^m) := C^0(D, \mathbb{R}^m) \cap B(D, \mathbb{R}^m).$
- (iv) $C^1(D, \mathbb{R}^m) := \{f : D \rightarrow \mathbb{R}^m \mid f \text{ ist auf } D \text{ stetig partiell differenzierbar}\}^a$
- (v) $C_b^1(D, \mathbb{R}^m) := \{f \in C^1(D, \mathbb{R}^m) \mid f, \partial_1 f, \dots, \partial_d f \in B(D, \mathbb{R}^m)\}$
- (vi) $C^k(D, \mathbb{R}^m) := \{f : D \rightarrow \mathbb{R}^m \mid f \text{ ist auf } D \text{ } k\text{-mal stetig partiell differenzierbar}\}$
- (vii) $C_b^k(D, \mathbb{R}^m) := \{f \in C^k(D, \mathbb{R}^m) \mid \forall \ell \in \mathbb{N}_0 \forall i_1, \dots, i_\ell \in \{1, \dots, d\} : \partial_{i_1} \dots \partial_{i_\ell} f \in B(D, \mathbb{R}^m)\}^b$

^aSie erinnern sich, dass wir beim partiellen Ableiten oft angenommen haben, dass D offen ist. Wir müssen dies aber per Definition nicht tun, siehe Definition 10.5.11. Falls wir es nicht tun so muss aber Bemerkung 10.5.4 beachtet werden.

^bHierbei ist für $\ell = 0$ die Liste $i_1, \dots, i_\ell$ die leere Liste. Wir einigen uns dann per Konvention darauf, dass für $\ell = 0$ gemeint ist, dass $f \in B(D, \mathbb{R}^m)$.

Bemerkung 11.2.8

Ein paar Worte zu den Relationen zwischen den Räumen. Ist $D \subseteq \mathbb{R}^d$ offen so gilt

$$C^1(D, \mathbb{R}^m) \subseteq C^0(D, \mathbb{R}^m). \quad (671)$$

In der Tat: Ist $f \in C^1(D, \mathbb{R}^m)$ so ist f auf D stetig partiell differenzierbar. Dann ist f nach Satz 10.5.30 auf D auch total differenzierbar. Aus Korollar 10.5.25 folgt dann auch die Stetigkeit von f auf D , d.h. $f \in C^0(D, \mathbb{R}^m)$. Gleichung (671) ist bewiesen. Somit erhält man auch klarerweise

$$C_b^1(D, \mathbb{R}^m) \subseteq C_b^0(D, \mathbb{R}^m).$$

Das zeigt nun auch, dass für alle $k \in \mathbb{N}$ gilt

$$C^k(D, \mathbb{R}^m) \subseteq C^{k-1}(D, \mathbb{R}^m) \subseteq \dots \subseteq C^0(D, \mathbb{R}^m).$$

und

$$C_b^k(D, \mathbb{R}^m) \subseteq C_b^{k-1}(D, \mathbb{R}^m) \subseteq \dots \subseteq C_b^0(D, \mathbb{R}^m).$$

Ein weiterer beobachtenswerter Zusammenhang ist der Folgende:

$$\text{Ist } D \subseteq \mathbb{R}^d \text{ kompakt so gilt } C^0(D, \mathbb{R}^m) \subseteq B(D, \mathbb{R}^m). \quad (672)$$

Dazu: Ist $f \in C^0(D, \mathbb{R}^m)$, so ist f auf der kompakten Menge D stetig. Nach Satz 10.4.20 ist dann auch $f(D) \subseteq \mathbb{R}^m$ kompakt in $\mathbb{R}^m$. Das bedeutet nach Definition 10.3.15 auch, dass $f(D)$ beschränkt ist, d.h.

$$\exists L > 0 : \quad \|y\| \leq L \quad \forall y \in f(D).$$

Für alle $x \in D$ gilt $f(x) \in f(D)$. Daher folgern wir

$$\exists L > 0 : \quad \|f(x)\| \leq L \quad \forall x \in D.$$

Wir folgern, dass f beschränkt ist, d.h. $f \in B(D, \mathbb{R}^m)$. Wegen (672) gilt dann klarerweise auch

$$C_b^0(D, \mathbb{R}^m) = C^0(D, \mathbb{R}^m) \quad \text{sofern } D \subseteq \mathbb{R}^d \text{ kompakt ist.}$$

Nun wollen wir die obigen Funktionenräume mit einer Metrik ausstatten.

Proposition 11.2.9 — EINE METRIK AUF $B(D, \mathbb{R}^m)$

Wir definieren für $f, g \in B(D, \mathbb{R}^m)$

$$d_\infty(f, g) := \sup_{y \in D} \|f(y) - g(y)\|. \quad (673)$$

Dann ist die Abbildung $d_\infty : B(D, \mathbb{R}^m) \times B(D, \mathbb{R}^m) \rightarrow \mathbb{R}$ eine Metrik auf $B(D, \mathbb{R}^m)$. Mit anderen Worten ist $(B(D, \mathbb{R}^m), d_\infty)$ ein metrischer Raum.

Beweis. Bevor wir hier die Axiome (M1),(M2),(M3) aus Definition 11.2.1 nachprüfen müssen wir nachweisen, dass d_∞ eine wohldefinierte Abbildung ist. Das bedeutet in unserem Fall, dass $d_\infty(f, g)$ auch wirklich immer eine reelle Zahl liefert (d.h. das Supremum aus (673) auch wirklich existiert). Dazu: Wegen der Ordnungsvollständigkeit von $\mathbb{R}$ (vgl. Definition 3.4.1) genügt es zu zeigen, dass die Menge

$$\{\|f(y) - g(y)\| : y \in D\} \quad (674)$$

nach oben beschränkt ist. Da $f, g \in B(D; \mathbb{R}^m)$ gibt es $L_1, L_2 > 0$ mit

$$\|f(x)\| \leq L_1 \quad \forall y \in D \quad \text{und} \quad \|g(x)\| \leq L_2 \quad \forall y \in D.$$

Es folgt, dass für beliebiges $y \in D$ gilt

$$\|f(y) - g(y)\| = \|f(y) + (-g(y))\| \leq \|f(y)\| + \|-g(y)\| = \|f(y)\| + \|g(y)\| \leq L_1 + L_2$$

Das beweist, dass die Menge in (674) durch $L_1 + L_2$ nach oben beschränkt ist. Wir folgern, dass das Supremum in (673) existiert und

$$\sup_{y \in D} \|f(y) - g(y)\| \leq L_1 + L_2$$

erfüllt.

Nun müssen wir noch die Axiome (M1),(M2),(M3) nachweisen. (M1) und (M2) überlassen wir den Lesenden als Übungsaufgabe. Zu (M3). Seien $f, g, h \in B(D, \mathbb{R}^m)$

$$\mathbb{Z} : \quad d_\infty(f, h) \leq d_\infty(f, g) + d_\infty(g, h).$$

Dazu nimm ein beliebiges $y \in D$. Wir schätzen ab

$$\|f(y) - h(y)\| = \|f(y) - g(y) + g(y) - h(y)\| \leq \|f(y) - g(y)\| + \|g(y) - h(y)\|.$$

Nun gilt $\|f(y) - g(y)\| \leq \sup_{z \in D} \|f(z) - g(z)\| = d_\infty(f, g)$ und analog gilt $\|g(y) - h(y)\| \leq d_\infty(g, h)$. Wir erhalten daher

$$\|f(y) - h(y)\| \leq d_\infty(f, g) + d_\infty(g, h) \quad \forall y \in D$$

Insbesondere ist $d_\infty(f, g) + d_\infty(g, h)$ eine obere Schranke von $\{\|f(y) - h(y)\| : y \in D\}$. Wir erhalten daher nach Definition 3.2.10

$$\sup_{y \in D} \|f(y) - h(y)\| \leq d_\infty(f, g) + d_\infty(g, h).$$

Zusammen mit (673) ergibt sich

$$d_\infty(f, h) \leq d_\infty(f, g) + d_\infty(g, h).$$

Nach Beispiel 11.2.6 können nun auch kleinere Funktionenräume mit derselben Metrik ausgestattet werden.

Korollar 11.2.10

- (i) Es sei $D \subseteq \mathbb{R}^d$ offen und $k \in \mathbb{N}_0$. Dann ist $(C_b^k(D, \mathbb{R}^m), d_\infty)$ ein metrischer Raum (wobei hier d_∞ gegeben ist wie in (673))
- (ii) Es sei $D \subseteq \mathbb{R}^d$ kompakt. Dann ist $(C^0(D, \mathbb{R}^m), d_\infty)$ ein metrischer Raum.

Beweis. Zu (i). Es sei $D \subseteq \mathbb{R}^d$ offen. Aus Bemerkung 11.2.8 lernt man, dass

$$C_b^k(D, \mathbb{R}^m) \subseteq C_b^0(D, \mathbb{R}^m) \subseteq B(D, \mathbb{R}^m).$$

Da $(B(D, \mathbb{R}^m), d_\infty)$ ein metrischer Raum ist folgt die Behauptung aus Beispiel 11.2.6.

Zu (ii). Es sei $D \subseteq \mathbb{R}^d$ kompakt. Aus Bemerkung 11.2.8 lernt man, dass

$$C(D, \mathbb{R}^m) \subseteq B(D, \mathbb{R}^m).$$

Da $(B(D, \mathbb{R}^m), d_\infty)$ ein metrischer Raum ist folgt die Behauptung aus Beispiel 11.2.6.

Für $k \in \mathbb{N}$ und $D \subseteq \mathbb{R}^d$ offen gibt es auf $C_b^k(D, \mathbb{R}^m)$ sogar noch eine populärere Metrik (welche häufiger benutzt wird). Für $f, g \in C_b^1(D, \mathbb{R}^m)$ definieren wir

$$d_\infty^1(f, g) := d_\infty(f, g) + \sum_{i=1}^d d_\infty(\partial_i f, \partial_i g). \quad (675)$$

Allgemein, für $k \in \mathbb{N}$ und $f, g \in C_b^k(D, \mathbb{R}^m)$ definieren wir

$$d_\infty^k(f, g) := \sum_{\ell=0}^k \sum_{i_1, \dots, i_\ell=1}^d d_\infty(\partial_{i_1} \dots \partial_{i_\ell} f, \partial_{i_1} \dots \partial_{i_\ell} g).$$

Nun haben wir einige Beispiel für metrische Räume gesehen. Wir würden nun gerne noch diskutieren, was metrische Räume mit topologischen Räumen zu tun haben. Dabei wird sich zeigen, dass metrische Räume eigentlich nur ein Spezialfall topologischer Räume sind. Das liegt daran, dass sich aus jeder Metrik eine Topologie *basteln* lässt. Um diesen Sachverhalt aufzuklären benötigen wir zunächst einige Vorbereitungen. Die erste Vorbereitung ist die Verallgemeinerung des Begriffes der r -Umgebung/des r -Balls auf metrische Räume (vgl. Definition 10.1.4).

Definition 11.2.11 — BÄLLE IN METRISCHEN RÄUMEN

Es sei (X, d) ein metrischer Raum, $x_0 \in X$ und $r > 0$. Dann definieren wir

$$B_r(x_0) := \{x \in X : d(x, x_0) < r\}.$$

Wir nennen $B_r(x_0)$ den r -Ball oder die r -Umgebung um x_0 .

Jetzt wo wir das Konzept eines Balles verallgemeinert haben, können wir (inspiriert von der Charakterisierung in Bemerkung 10.3.6) auch verallgemeinern was offene Mengen sind.

Definition 11.2.12

Es sei (X, d) ein metrischer Raum und $U \subseteq X$. Wir nennen U *offen in (X, d)* falls es zu jedem $x \in U$ ein $r > 0$ gibt so, dass $B_r(x) \subseteq U$.

Wer von offenen Mengen sprechen kann, kann auch eine Topologie definieren. Jedenfalls dann, wenn diese offenen Mengen die Axiome aus Definition 11.1.1 erfüllen.

Proposition 11.2.13

Es sei (X, d) ein metrischer Raum. Definiere

$$\begin{aligned} \mathcal{O}_d &:= \{U \subseteq X : U \text{ offen in } (X, d) \text{ im Sinne von Definition 11.2.12}\} \\ &= \{U \subseteq X : \forall x \in U \exists r > 0 : B_r(x) \subseteq U\}. \end{aligned} \quad (676)$$

Dann ist $\mathcal{O}_d$ eine Topologie auf X , d.h. $(X, \mathcal{O}_d)$ ist auch ein topologischer Raum.

Beweis. Es muss gezeigt werden, dass die Axiome (T1),(T2),(T3) aus Definition 11.1.1 erfüllt sind.

Zu (T1). Wir müssen zeigen, dass $X \in \mathcal{O}_d$. Dazu: Für $x \in X$ kann man $r > 0$ beliebig wählen und erhält immer $B_r(x) \subseteq X$. Aus (676) folgt dann $X \in \mathcal{O}_d$.

Zu (T2). Es sei $(U_i)_{i \in I} \subseteq \mathcal{O}_d$ eine Familie offener Mengen.

$$\mathbb{Z}: \quad \bigcup_{i \in I} U_i \in \mathcal{O}_d$$

Dazu: Es sei $x \in \bigcup_{i \in I} U_i$ beliebig. Dann gibt es ein $i_0 \in I$ mit $x \in U_{i_0}$. Da $U_{i_0} \in \mathcal{O}_d$, finden wir nach Definition 11.2.12 ein $r > 0$ mit $B_r(x) \subseteq U_{i_0}$. Dann gilt aber auch

$$B_r(x) \subseteq U_{i_0} \subseteq \bigcup_{i \in I} U_i.$$

Wir haben also für jedes $x \in \bigcup_{i \in I} U_i$ ein $r > 0$ gefunden, sodass $B_r(x) \subseteq \bigcup_{i \in I} U_i$.

Zu (T3). Es sei $(U_i)_{i \in I_0} \subseteq \mathcal{O}_d$ eine Familie von Mengen in $\mathcal{O}_d$ und I_0 endlich, etwa gelte $I_0 = \{i_1, \dots, i_n\}$ für ein $n \in \mathbb{N}$.

$$\mathbb{Z}: \quad \bigcap_{i \in I_0} U_i = U_{i_1} \cap \dots \cap U_{i_n} \in \mathcal{O}_d$$

Dazu sei wieder $x \in U_{i_1} \cap \dots \cap U_{i_n}$ beliebig. Dann gilt $x \in U_{i_k}$ für alle $k = 1, \dots, n$. Sei nun $k \in \{1, \dots, n\}$. Da U_{i_k} in (X, d) offen ist, gibt es auch ein $r_k > 0$ so, dass

$$B_{r_k}(x) \subseteq U_{i_k} \tag{677}$$

Definiere nun

$$r := \min\{r_1, \dots, r_n\} > 0.$$

Wir behaupten, dass $B_r(x) \subseteq U_{i_1} \cap \dots \cap U_{i_n}$. Dazu sei $y \in B_r(x)$. Dann gilt

$$d(y, x) < r \quad \implies \quad d(y, x) < r_1 \quad \implies \quad y \in B_{r_1}(x) \quad \xRightarrow{(677)} \quad y \in U_{i_1}.$$

Analog gilt

$$d(y, x) < r \quad \implies \quad d(y, x) < r_2 \quad \implies \quad y \in B_{r_2}(x) \quad \xRightarrow{(677)} \quad y \in U_{i_2}.$$

Machen wir so weiter für alle $k = 1, \dots, n$ so folgern wir

$$(y \in U_{i_1}) \wedge (y \in U_{i_2}) \wedge \dots \wedge (y \in U_{i_n})$$

Das heißt nichts anderes als $y \in U_{i_1} \cap \dots \cap U_{i_n}$. Da $y \in B_r(x)$ beliebig war folgt schließlich

$$B_r(x) \subseteq U_{i_1} \cap \dots \cap U_{i_n}$$

Definition 11.2.14 — DURCH METRIK INDUZIERTE TOPOLOGIE

Es sei (X, d) ein metrischer Raum und $\mathcal{O}_d$ wie in Proposition 11.2.13. Dann nennen wir $\mathcal{O}_d$ die *durch d induzierte Topologie*.

Mit jedem metrischen Raum (X, d) haben wir also auch einen topologischen Raum $(X, \mathcal{O}_d)$ vorliegen. Daher können wir natürlich von topologischen Konzepten wie Kompaktheit in (X, d) sprechen, vgl. Definition 11.1.8. Außerdem können wir Stetigkeit von Abbildungen mit Definitionsbereich X oder Bildbereich X verstehen, vgl. Definition 11.1.1.

Konvergenz in metrischen Räumen

Ein weiteres Konzept was wir nun in metrischen Räumen verstehen können ist die *Konvergenz*. Eine Metrik $d : X \times X \rightarrow \mathbb{R}$ misst den Abstand zweier Punkte in X . Dies kann auf ganz natürliche Weise benutzt werden, um Konvergenz in metrischen Räumen zu definieren.

Definition 11.2.15 — KONVERGENZ IN METRISCHEN RÄUMEN

Es sei (X, d) ein metrischer Raum und $(a_n)_{n \in \mathbb{N}} \subseteq X$ eine Folge in X . Wir nennen $(a_n)_{n \in \mathbb{N}}$ gegen ein $a \in X$ konvergent falls

$$\lim_{n \rightarrow \infty} d(a_n, a) = 0. \quad (678)$$

Das bedeutet, die reelle Zahlenfolge $(d(a_n, a))_{n \in \mathbb{N}} \subseteq \mathbb{R}$ konvergiert im Sinne von Definition 2.3.1 gegen 0. Wir schreiben in dem Fall auch $a_n \xrightarrow[d]{n \rightarrow \infty} a$ oder $d\text{-}\lim_{n \rightarrow \infty} a_n = a$. Ist die verwendete Metrik d aus dem Zusammenhang klar so kann man das d bei den Schreibweisen auch weglassen und auch $\lim_{n \rightarrow \infty} a_n = a$ schreiben.

Wie gewohnt nennen wir eine Folge $(a_n)_{n \in \mathbb{N}} \subseteq X$ in (X, d) *konvergent*, falls es ein $a \in X$ gibt mit $a_n \xrightarrow[d]{n \rightarrow \infty} a$.

Wir haben also die Frage, ob eine Folge $(a_n)_{n \in \mathbb{N}} \subseteq X$ gegen ein $a \in X$ konvergiert, auf die Konvergenzfrage für die reelle Folge $(d(a_n, a))_{n \in \mathbb{N}} \subseteq \mathbb{R}$ zurückgeführt.

Bemerkung 11.2.1. Es sei (X, d) ein metrischer Raum, $a \in X$ und $(a_n)_{n \in \mathbb{N}} \subseteq X$ eine Folge. Würde man noch mit Definition 2.3.1 ausschreiben was (678) eigentlich bedeutet so hätte man

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad d(a_n, a) \leq \varepsilon \quad \forall n \geq n_0.$$

Grenzwertregeln können wir an dieser Stelle nicht formulieren, denn wir haben auf einem allgemeinen metrischen Raum X ja gar keine Rechenregeln! Eine wichtige Eigenschaft der Konvergenz sollten wir hier aber doch noch formulieren: die *Eindeutigkeit von Grenzwerten*.

Proposition 11.2.16

[Eindeutigkeit von Grenzwerten] Es sei (X, d) ein metrischer Raum und $(a_n)_{n \in \mathbb{N}} \subseteq X$ eine Folge sowie $a, b \in X$. Gilt $a_n \xrightarrow[d]{n \rightarrow \infty} a$ und $a_n \xrightarrow[d]{n \rightarrow \infty} b$ so folgt $a = b$.

Beweis. Es sei $(a_n)_{n \in \mathbb{N}}$ wie in der Aussage. Dann gilt für $n \in \mathbb{N}$

$$0 \underset{(M1)}{\leq} d(a, b) \underset{(M3)}{\leq} d(a, a_n) + d(a_n, b). \quad (679)$$

Da $a_n \xrightarrow[d]{n \rightarrow \infty} a$ und $a_n \xrightarrow[d]{n \rightarrow \infty} b$ gilt $\lim_{n \rightarrow \infty} d(a_n, a) = \lim_{n \rightarrow \infty} d(a_n, b) = 0$. Wir folgern aus (679) und (M2),

$$0 \leq d(a, b) \leq d(a_n, a) + d(a_n, b) \xrightarrow{n \rightarrow \infty} 0.$$

Also gilt $d(a, b) = 0$ und deswegen $a = b$.

Als nächstes versuchen wir, die Konvergenz von Folgen in verschiedenen metrischen Räumen zu charakterisieren.

Beispiel 11.2.2. Betrachte $X = \mathbb{R}^d$ ausgestattet mit der Metrik d_{euc} . Ist $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ im Sinne von Definition 11.2.15 gegen ein $a \in \mathbb{R}^d$ konvergent so gilt

$$0 = \lim_{n \rightarrow \infty} d_{\text{euc}}(a_n, a) = \lim_{n \rightarrow \infty} \|a_n - a\|.$$

Nach Satz 10.2.6 wissen wir, dass dies äquivalent zur Konvergenz im Sinne von Definition 10.2.2 ist. Somit sind die Konvergenzbegriffe, die wir auf $\mathbb{R}^d$ eingeführt haben, konsistent.

Auch die gleichmäßige Konvergenz von Funktionenfolgen $(f_n : D \rightarrow \mathbb{R})_{n \in \mathbb{N}}$ ist ein Konvergenzbegriff, der von einer Metrik kommt.

Beispiel 11.2.17 — METRISIERUNG DER GLEICHMÄSSIGEN KONVERGENZ

Betrachte für $D \subseteq \mathbb{R}$ den Raum $X = B(D, \mathbb{R})$ (d.h. $m = 1$) ausgestattet mit der Metrik d_∞ , vgl. Proposition 11.2.9. Ferner sei $(f_n)_{n \in \mathbb{N}} \subseteq B(D, \mathbb{R})$ (eine Funktionenfolge) und $f \in B(D, \mathbb{R})$.

$$\text{Behauptung.} \quad f_n \xrightarrow[d_\infty]{n \rightarrow \infty} f \quad \Longleftrightarrow \quad (f_n)_{n \in \mathbb{N}} \text{ ist gleichmäßig gegen } f \text{ konvergent.}$$

Man sagt: Die Konvergenz, die von der Metrik d_∞ kommt ist gerade die gleichmäßige Konvergenz.

Beweis. Es gilt

$$\begin{aligned}
 & f_n \xrightarrow[d_\infty]{n \rightarrow \infty} f \\
 \iff_{\text{Def 11.2.15}} & d_\infty(f_n, f) \xrightarrow{n \rightarrow \infty} 0 \\
 \iff_{(673)} & \sup_{y \in D} \|f_n(y) - f(y)\| \xrightarrow{n \rightarrow \infty} 0 \\
 \iff_{m=1} & \sup_{y \in D} |f_n(y) - f(y)| \xrightarrow{n \rightarrow \infty} 0.
 \end{aligned}$$

In der letzten Zeile steht nun gerade die Definition der gleichmäßigen Konvergenz, vgl. Definition 9.1.1.

Bei dem Konvergenzbegriff muss man stets sehr genau auf den zugrundeliegenden metrischen Raum achten.

Beispiel 11.2.18

Es sei $X = (0, 1) \subseteq \mathbb{R}$ ausgestattet mit d_{euc} (d.h. für $x, y \in X$ gilt $d_{\text{euc}}(x, y) = |x - y|$).

Behauptung. Die Folge $(\frac{1}{n})_{n \in \mathbb{N}} \subseteq X$ ist in (X, d_{euc}) nicht konvergent.

Beweis. Angenommen $(\frac{1}{n})_{n \in \mathbb{N}}$ wäre in (X, d_{euc}) konvergent. Dann müsste es ein $a \in X$ (d.h. $a \in (0, 1)$) geben mit

$$0 = \lim_{n \rightarrow \infty} d_{\text{euc}}(\frac{1}{n}, a) = \lim_{n \rightarrow \infty} |\frac{1}{n} - a| \quad (680)$$

Jedoch wissen wir nach den Grenzwertregeln für reelle Zahlenfolgen, dass $\lim_{n \rightarrow \infty} |\frac{1}{n} - a| = |a|$. Aus (680) folgt also $0 = |a|$, d.h. $a = 0$, was $a \in (0, 1)$ widerspricht.

Jetzt wo wir einen Konvergenzbegriff in einem metrischen Raum (X, d) haben können wir auch den Begriff der *Abgeschlossenheit* verallgemeinern.

Definition 11.2.19 — ABGESCHLOSSENHEIT IN METRISCHEN RÄUMEN

Es sei (X, d) ein metrischer Raum. Eine Menge $A \subseteq X$ heißt *abgeschlossen in (X, d)* falls jede in (X, d) konvergente Folge $(a_n)_{n \in \mathbb{N}} \subseteq A$ ihren Grenzwert auch in A besitzt. Präziser:

$$(a_n)_{n \in \mathbb{N}} \subseteq A, \quad a_n \xrightarrow[d]{n \rightarrow \infty} a \quad \implies \quad a \in A.$$

Nach Bemerkung 10.3.6 (ii) ist dieser Abgeschlossenheitsbegriff konsistent mit der Abgeschlossenheit in $(\mathbb{R}^d, d_{\text{euc}})$, d.h. im Falle von $(X, d) = (\mathbb{R}^d, d_{\text{euc}})$ stimmen die beiden kennengelernten

Abgeschlossenheitsbegriffe überein. Wir wollen den Begriff aber auch in anderen metrischen Räumen studieren. Ein Beispiel dazu

Beispiel 11.2.20

Es sei $D \subseteq \mathbb{R}$. Betrachte $(X, d) = (B(D, \mathbb{R}), d_\infty)$ und $A = C_b(D, \mathbb{R}) \subseteq X$.

Behauptung. $A = C_b^0(D, \mathbb{R})$ ist abgeschlossen in (X, d) .

Beweis. Um dies zu zeigen prüfen wir Definition 11.2.19 nach. Es sei $(f_n)_{n \in \mathbb{N}} \subseteq C_b(D, \mathbb{R})$ und $f \in B(D, \mathbb{R})$ so, dass $f_n \xrightarrow[d_\infty]{n \rightarrow \infty} f$

$$\mathbb{Z} : \quad f \in C_b^0(D, \mathbb{R}).$$

Dazu muss nur gezeigt werden, dass f stetig ist (denn beschränkt ist es ja schon). Dass $f_n \xrightarrow[d_\infty]{n \rightarrow \infty} f$ ist nach Beispiel 11.2.17 gleichbedeutend damit, dass $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen f konvergiert. Da $(f_n)_{n \in \mathbb{N}} \subseteq C_b^0(D, \mathbb{R})$ ist f_n stetig für alle $n \in \mathbb{N}$. Dies und die gleichmäßige Konvergenz implizieren, dass auch f stetig ist (vgl. Satz 9.2.2). Somit gilt also $f \in C_b^0(D, \mathbb{R})$, wie zu zeigen war.

Allerdings sind nicht alle Funktionenräume bezüglich d_∞ abgeschlossen.

Beispiel 11.2.21

Es sei $(X, d) = (B((-1, 1), \mathbb{R}), d_\infty)$

Behauptung. $C_b^1((-1, 1), \mathbb{R})$ ist nicht abgeschlossen in (X, d) .

Beweis. Wir geben eine Folge $(f_n)_{n \in \mathbb{N}} \subseteq C_b^1((-1, 1), \mathbb{R})$ an, die in (X, d) konvergiert, aber deren Grenzwert nicht in $C_b^1((-1, 1), \mathbb{R})$. Betrachte dazu

$$f_n : (-1, 1) \rightarrow \mathbb{R} \quad f_n(x) = \sqrt{x^2 + \frac{1}{n}}$$

In Beispiel 9.1.5 haben wir bereits gesehen, dass $(f_n)_{n \in \mathbb{N}}$ gleichmäßig gegen die Betragsfunktion $f : (-1, 1) \rightarrow \mathbb{R}, x \mapsto |x|$ konvergiert. Also gilt nach Beispiel 11.2.17 auch, dass $f_n \xrightarrow[d_\infty]{n \rightarrow \infty} f$. Nun ist $f_n \in C_b^1((-1, 1), \mathbb{R})$ aber $f \notin C_b^1((-1, 1), \mathbb{R})$. Das zeigt, dass $C_b^1((-1, 1), \mathbb{R})$ nicht abgeschlossen ist, vgl. Definition 11.2.19.

Zum Ende dieses Abschnittes werden wir noch die (rein topologische!) Charakterisierung der Abgeschlossenheit aus Proposition 10.3.10 verallgemeinern.

Proposition 11.2.22 — TOPOLOGISCHE CHARAKTERISIERUNG VON ABGESCHLOSSENHEIT

Es sei (X, d) ein metrischer Raum und $A \subseteq X$. Dann gilt

$$A \text{ ist abgeschlossen in } (X, d) \quad \Longleftrightarrow \quad A^C := X \setminus A \text{ ist offen in } (X, d)$$

Beweis. Zuerst zu „ $\Rightarrow$ “. Es sei $A \subseteq X$ abgeschlossen in (X, d) . Zu zeigen ist, dass A^C offen in (X, d) ist. Das heißt nach Definition 11.2.12

$$\mathbb{Z}: \quad \forall x \in A^C \quad \exists r > 0 : \quad B_r(x) \subseteq A^C.$$

Fixieren wir dazu ein $x \in A^C$. Angenommen es gibt kein $r > 0$ mit $B_r(x) \subseteq A^C$. Dann muss also für jedes $r > 0$ der Ball $B_r(x)$ einen Punkt von A enthalten. Insbesondere gilt

$$B_{\frac{1}{n}}(x) \cap A \neq \emptyset \quad \forall n \in \mathbb{N}.$$

Wähle nun für $n \in \mathbb{N}$ ein Element $a_n \in B_{\frac{1}{n}}(x) \cap A$. Wir können alle diese Elemente in eine Folge $(a_n)_{n \in \mathbb{N}}$ zusammenfassen. Insbesondere gilt $(a_n)_{n \in \mathbb{N}} \subseteq A$. Außerdem gilt auch $a_n \xrightarrow[d]{n \rightarrow \infty} x$, weil

$$0 \leq d(a_n, x) \underset{a_n \in B_{\frac{1}{n}}(x)}{\leq} \frac{1}{n} \xrightarrow{n \rightarrow \infty} 0.$$

Also gilt $\lim_{n \rightarrow \infty} d(a_n, x) = 0$ und nach Definition 11.2.15 folgt $a_n \xrightarrow[d]{n \rightarrow \infty} x$. Aus der Abgeschlossenheit von A (vgl. Definition 11.2.19) würde dann $x \in A$ folgen. Ein Widerspruch zu $x \in A^C$.

Nun noch zu „ $\Leftarrow$ “. Es sei $A \subseteq X$ so, dass A^C offen ist. Zu zeigen ist, dass A abgeschlossen ist. Dazu prüfen wir Definition 11.2.19 nach. Sei $(a_n)_{n \in \mathbb{N}} \subseteq A$ eine konvergente Folge, etwa gelte $a_n \xrightarrow[d]{n \rightarrow \infty} a$ für ein $a \in X$.

$$\mathbb{Z}: \quad a \in A.$$

Wir beweisen dies abermals durch Widerspruch. Angenommen $a \in A^C$. Weil A^C offen ist, gibt es $r > 0$ so, dass $B_r(a) \subseteq A^C$. Wir erinnern uns nun, dass $a_n \xrightarrow[d]{n \rightarrow \infty} a$. Wenden wir Bemerkung 11.2.1 mit $\varepsilon = r$ an, finden wir

$$\exists n_0 \in \mathbb{N} : \quad d(a_n, a) < r \quad \forall n \geq n_0.$$

Das heißt gerade, dass für alle $n \geq n_0$ gilt, dass $a_n \in B_r(a) \subseteq A^C$. Insbesondere gilt $a_{n_0} \in A^C$ was der Voraussetzung $(a_n)_{n \in \mathbb{N}} \subseteq A$ widerspricht.

Das bedeutet: Die Abgeschlossenheit einer Menge $A \subseteq X$ lässt sich *rein topologisch* (d.h. nur mit dem Begriff einer offenen Menge) charakterisieren. Insbesondere kann man dem Begriff der Abgeschlossenheit auch in einem beliebigen topologischen Raum $(X, \mathcal{O})$ einen Sinn geben:

Man würde eine Teilmenge $A \subseteq X$ abgeschlossen nennen, wenn $A^C := X \setminus A$ offen ist, d.h. $X \setminus A \in \mathcal{O}$.

Cauchy-Folgen und Vollständigkeit

Parallel zum Konvergenzbegriff haben wir immer den Begriff der *Cauchy-Folgen* behandelt. Auch diesen Begriff können wir auf natürliche Weise auf metrische Räume verallgemeinern. Es wird sich aber herausstellen, dass in allgemeinen metrischen Räumen (X, d) Cauchy-Folgen nicht zwingend konvergent sind!

Definition 11.2.23 — CAUCHY-FOLGEN IN METRISCHEN RÄUMEN

Es sei (X, d) ein metrischer Raum. Eine Folge $(a_n)_{n \in \mathbb{N}} \subseteq X$ heißt *Cauchy-Folge in (X, d)* , wenn

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad d(a_n, a_m) \leq \varepsilon \quad \forall n, m \geq n_0.$$

In $(\mathbb{R}, d_{\text{euc}})$ haben wir den Begriff der Cauchy-Folge als Konvergenzkriterium benutzt. In allgemeinen metrischen Räumen konvergiert aber nicht jede Cauchy-Folge, wie das nachfolgende (eigentlich auf der Hand liegende) Beispiel zeigt.

Beispiel 11.2.24

Betrachte wie in Beispiel 11.2.18 den metrischen Raum $X = (0, 1)$ ausgestattet mit d_{euc} . Wir haben dort bereits gesehen, dass die Folge $(\frac{1}{n})_{n \in \mathbb{N}} \subseteq X$ in (X, d) nicht konvergent ist.

Behauptung. Allerdings ist $(\frac{1}{n})_{n \in \mathbb{N}} \subseteq X$ eine Cauchy-Folge in (X, d) .

Beweis. Wir weisen Definition 11.2.23 nach. Es sei dazu $\varepsilon > 0$ beliebig vorgegeben.

$$\mathbb{Z} : \quad \exists n_0 \in \mathbb{N} : \quad d_{\text{euc}}\left(\frac{1}{n}, \frac{1}{m}\right) \leq \varepsilon \quad \forall n, m \geq n_0.$$

Wähle $n_0 := \lceil \frac{1}{\varepsilon} \rceil + 1$. Dann gilt für $n, m \geq n_0$

$$d_{\text{euc}}\left(\frac{1}{n}, \frac{1}{m}\right) = \left| \frac{1}{n} - \frac{1}{m} \right| = \frac{1}{\min\{n, m\}} - \frac{1}{\max\{n, m\}} \leq \frac{1}{\min\{n, m\}} \leq \frac{1}{n_0} = \frac{1}{\lceil \frac{1}{\varepsilon} \rceil + 1} \leq \frac{1}{\frac{1}{\varepsilon}} = \varepsilon.$$

Das zeigt, dass $(\frac{1}{n})_{n \in \mathbb{N}}$ eine Cauchy-Folge ist.

Somit sind Cauchy-Folgen in der abstrakten Analysis nicht zwingend konvergent. Immerhin funktioniert es noch andersherum...

Proposition 11.2.25

Es sei (X, d) ein metrischer Raum und $(a_n)_{n \in \mathbb{N}} \subseteq X$ eine konvergente Folge. Dann ist $(a_n)_{n \in \mathbb{N}}$ auch eine Cauchy-Folge.

Beweis. Es sei $(a_n)_{n \in \mathbb{N}}$ eine konvergente Folge. Etwa gelte

$$a_n \xrightarrow[d]{n \rightarrow \infty} a \quad (681)$$

Wir weisen Definition 11.2.23 nach. Es sei dazu $\varepsilon > 0$ beliebig vorgegeben. Wegen (681) und Bemerkung 11.2.1 gilt

$$\exists n_0 \in \mathbb{N} : \quad d(a_k, a) \leq \frac{\varepsilon}{2} \quad \forall k \geq n_0.$$

Für beliebige $m, n \geq n_0$ gilt dann

$$d(a_n, a_m) \leq d(a_n, a) + d(a, a_m) = d(a_n, a) + d(a_m, a) \underset{n, m \geq n_0}{\leq} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Das zeigt: Für jedes $\varepsilon > 0$ konnten wir ein $n_0 \in \mathbb{N}$ finden mit

$$d(a_n, a_m) \leq \varepsilon \quad \forall n, m \geq n_0.$$

Die Behauptung ist gezeigt.

Wir wissen nun also, dass das Cauchy-Kriterium notwendig für Konvergenz ist. Außerdem haben wir gesehen, dass es im Allgemeinen kein hinreichendes Kriterium ist. In ein paar speziellen metrischen Räumen [zum Beispiel in $(\mathbb{R}, d_{\text{euc}})$] ist das Cauchy-Kriterium allerdings doch hinreichend! Solche metrische Räume nennen wir *vollständig*.

Definition 11.2.26 — VOLLSTÄNDIGKEIT

Es sei (X, d) ein metrischer Raum. Wir nennen (X, d) *vollständig*, wenn jede Cauchy-Folge in (X, d) auch konvergent ist.

Beispiel 11.2.27

$(\mathbb{R}, d_{\text{euc}})$ ist vollständig. Dies zeigt Bemerkung 3.5.35. $((0, 1), d_{\text{euc}})$ ist nicht vollständig, wie Beispiel 11.2.24 zeigt.

Beispiel 11.2.28

Behauptung. $(\mathbb{R}^d, d_{\text{euc}})$ ist vollständig.

Beweis. Es sei $(a_n)_{n \in \mathbb{N}} \subseteq \mathbb{R}^d$ eine Cauchy-Folge in $(\mathbb{R}^d, d_{\text{euc}})$, d.h. nach Definition 11.2.23

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad d_{\text{euc}}(a_n, a_m) \leq \varepsilon \quad \forall n, m \geq n_0.$$

Schreibt man die Definition von d_{euc} aus so ergibt sich

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad \|a_n - a_m\| \leq \varepsilon \quad \forall n, m \geq n_0. \quad (682)$$

Es sei nun $i \in \{1, \dots, d\}$ beliebig. Für $m, n \in \mathbb{N}$

$$|a_n^{(i)} - a_m^{(i)}| = \sqrt{|a_n^{(i)} - a_m^{(i)}|^2} \leq \sqrt{\sum_{j=1}^d |a_n^{(j)} - a_m^{(j)}|^2} = \|a_n - a_m\|.$$

Nutzt man dies und (682) so erhält man

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad |a_n^{(i)} - a_m^{(i)}| \leq \varepsilon \quad \forall n, m \geq n_0.$$

Somit ist $(a_n^{(i)})_{n \in \mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$. Nach Bemerkung 3.5.35 ist $(a_n^{(i)})_{n \in \mathbb{N}}$ auch konvergent. Da $i \in \{1, \dots, d\}$ beliebig war konvergiert nun jede Komponentenfolge der Folge $(a_n)_{n \in \mathbb{N}}$. Nach Definition 10.2.2 ist $(a_n)_{n \in \mathbb{N}}$ in $\mathbb{R}^d$ konvergent. Sei etwa $a = \lim_{n \rightarrow \infty} a_n$. Dann sagt Satz 10.2.6

$$0 = \lim_{n \rightarrow \infty} \|a_n - a\| = \lim_{n \rightarrow \infty} d_{\text{euc}}(a_n, a).$$

Mit Definition 11.2.15 folgt, dass $(a_n)_{n \in \mathbb{N}}$ in $(\mathbb{R}^d, d_{\text{euc}})$ konvergent ist.

Nun diskutieren wir auch die Vollständigkeit von Funktionenräumen.

Proposition 11.2.29

Es sei $D \subseteq \mathbb{R}$. Dann ist $(B(D, \mathbb{R}), d_\infty)$ ein vollständiger metrischer Raum.

Beweis. Es sei $(f_n)_{n \in \mathbb{N}} \subseteq B(D, \mathbb{R})$ eine Cauchy-Folge. Das heißt

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad d_\infty(f_n, f_m) \leq \varepsilon \quad \forall n, m \geq n_0. \quad (683)$$

$$\mathbb{Z}: \quad (f_n)_{n \in \mathbb{N}} \text{ ist konvergent in } (B(D, \mathbb{R}), d_\infty)$$

Dazu: Schreiben wir mit (673) aus was $d_\infty(f_n, f_m)$ bedeutet so erhalten wir aus (683)

$$\forall \varepsilon > 0 \quad \exists n_0 \in \mathbb{N} : \quad \sup_{y \in D} |f_n(y) - f_m(y)| \leq \varepsilon \quad \forall n, m \geq n_0.$$

Dies ist gerade genau das Cauchy-Kriterium für gleichmäßige Konvergenz aus Satz 9.1.8. Aus diesem schließen wir, dass es ein $f : D \rightarrow \mathbb{R}$ gibt mit $f_n \xrightarrow[\text{glm}]{n \rightarrow \infty} f$. Man könnte jetzt denken wir sind fertig. Jedoch fehlt noch eine Beobachtung

$$\text{Noch } \mathbb{Z}: \quad f \in B(D, \mathbb{R}).$$

Wenn das gezeigt ist, dann folgt nämlich aus Beispiel 11.2.17 dass $f_n \xrightarrow[d_\infty]{n \rightarrow \infty} f$ und damit wäre dann die Konvergenz von $(f_n)_{n \in \mathbb{N}}$ in $(B(D, \mathbb{R}), d_\infty)$ bewiesen. Nun also zur Beschränktheit

von f . Da $(f_n)_{n \in \mathbb{N}}$ bereits gleichmäßig gegen f konvergiert gibt es (vgl. Definition 9.1.1) ein $n_0 \in \mathbb{N}$ mit

$$\sup_{y \in D} |f_n(y) - f(y)| \leq 1 \quad \forall n \geq n_0. \quad (684)$$

Dann gilt für $x \in D$

$$|f(x)| = |f(x) - f_{n_0}(x) + f_{n_0}(x)| \leq |f(x) - f_{n_0}(x)| + |f_{n_0}(x)| \leq \sup_{y \in D} |f(y) - f_{n_0}(y)| + |f_{n_0}(x)|. \quad (685)$$

Da $f_{n_0} \in B(D, \mathbb{R})$ gibt es ein $L > 0$ mit $|f_{n_0}(x)| \leq L$ für alle $x \in D$. Zusammen mit (684) und (685) erhalten wir

$$|f(x)| \leq 1 + L \quad \forall x \in D.$$

Mit anderen Worten haben wir $f \in B(D, \mathbb{R})$. Die Behauptung ist gezeigt.

Nun wollen wir auch die Frage beantworten, unter welchen Umständen $(C_b^0(D, \mathbb{R}), d_\infty)$ und $(C_b^1(D, \mathbb{R}), d_\infty)$ vollständig sind. Dabei hilft uns folgende Beobachtung

Proposition 11.2.30

Es sei (X, d) ein vollständiger metrischer Raum und $A \subseteq X$ eine Teilmenge. Dann gilt

$$(A, d) \text{ ist ein vollständiger metrischer Raum}^{(a)} \iff A \text{ ist abgeschlossen in } (X, d).$$

^aHier würde man wohl streng nach Beispiel 11.2.6 wohl eigentlich (A, d_A) schreiben

Beweis. Zuerst zu „ $\Rightarrow$ “. Es sei (A, d) ein vollständiger metrischer Raum. Wir müssen nun zeigen, dass A abgeschlossen in (X, d) ist. Nimm dazu eine in (X, d) konvergente Folge $(a_n)_{n \in \mathbb{N}} \subseteq A$, etwa $a_n \xrightarrow[d]{n \rightarrow \infty} a$ für ein $a \in X$.

$$\mathbb{Z}: \quad a \in A.$$

Dazu: Da $(a_n)_{n \in \mathbb{N}}$ in (X, d) konvergent ist, ist nach Proposition 11.2.25 $(a_n)_{n \in \mathbb{N}}$ auch eine Cauchy-Folge in (X, d) . Da $(a_n)_{n \in \mathbb{N}} \subseteq A$ ist $(a_n)_{n \in \mathbb{N}}$ aber auch eine Cauchy-Folge in (A, d) . Da (A, d) vollständig ist, gibt es ein $b \in A$ mit $a_n \xrightarrow[d]{n \rightarrow \infty} b$. Betrachten wir nun b als ein Element von X , so haben wir in (X, d) folgende Situation

$$a_n \xrightarrow[d]{n \rightarrow \infty} a \quad \text{und} \quad a_n \xrightarrow[d]{n \rightarrow \infty} b.$$

Aus Proposition 11.2.16 folgt $a = b$. Insbesondere gilt $a \in A$ (da ja b ein Element aus A war). Die Behauptung ist gezeigt.

Nun zu „ $\Leftarrow$ “. Es sei $A \subseteq X$ abgeschlossen in (X, d) . Zu zeigen ist, dass (A, d) vollständig ist. Dazu sei $(a_n)_{n \in \mathbb{N}} \subseteq A$ eine Cauchy-Folge in (A, d) . Da $A \subseteq X$ können wir $(a_n)_{n \in \mathbb{N}}$ auch als Cauchy-Folge in (X, d) sehen (die Metrik ist ja schließlich bis auf Einschränkung

dieselbe, vgl. 11.2.23). Da (X, d) vollständig ist, ist nach Definition 11.2.26 $(a_n)_{n \in \mathbb{N}}$ in (X, d) konvergent. Also gibt es $x \in X$ mit $a_n \xrightarrow[d]{n \rightarrow \infty} x$. Da $(a_n)_{n \in \mathbb{N}} \subseteq A$ und A abgeschlossen ist, folgt nach Definition 11.2.19, dass $x \in A$. Wir haben also ein $x \in A$ gefunden mit

$$\lim_{n \rightarrow \infty} d(a_n, x) = 0.$$

Das heißt nichts anderes als dass $(a_n)_{n \in \mathbb{N}}$ in (A, d) konvergent ist.

Als Korollar erhalten wir

Korollar 11.2.31 — VOLLSTÄNDIGKEIT VON FUNKTIONENRÄUMEN

- (i) Für beliebige $D \subseteq \mathbb{R}$ ist der metrische Raum $(C_b^0(D, \mathbb{R}), d_\infty)$ vollständig.
- (ii) Der metrische Raum $(C_b^1((-1, 1), \mathbb{R}), d_\infty)$ ist nicht vollständig.

Beweis. Zu (i). Da $C_b^0(D, \mathbb{R}) \subseteq B(D, \mathbb{R})$ abgeschlossen ist (vgl. Beispiel 11.2.20) und $(B(D, \mathbb{R}), d_\infty)$ ein vollständiger metrischer Raum ist (vgl. Proposition 11.2.29) ist nach Proposition 11.2.30 auch $(C_b^0(D, \mathbb{R}), d_\infty)$ vollständig.

Zu (ii). Wäre $(C_b^1((-1, 1), \mathbb{R}), d_\infty)$ vollständig, so wäre nach der vorigen Proposition auch $C_b^1((-1, 1), \mathbb{R}) \subseteq B((-1, 1), \mathbb{R})$ abgeschlossen. Das ist aber nicht der Fall, vgl. Beispiel 11.2.21.

Spannenderweise wird $C_b^1((-1, 1), \mathbb{R})$ mit der anderen Metrik d_∞^1 aus (675) vollständig! Das folgert man leicht aus Satz 9.3.5. Details dazu überlassen wir den Lesenden als Übungsaufgabe! Daher wird die Metrik d_∞^1 auf $C_b^1((-1, 1), \mathbb{R})$ auch viel häufiger genutzt als d_∞ !

Der Banach'sche Fixpunktsatz

Im folgenden Abschnitt wollen wir anfangen, die Früchte unserer abstrakten Analysis zu ernten. Ein erstes wichtiges Resultat hierbei ist der *Banach'sche Fixpunktsatz*, welchen wir hier diskutieren. Bei diesem Satz geht es um Fixpunkte von Abbildungen, definiert wie folgt

Definition 11.2.32 — FIXPUNKT

Es sei M eine Menge und $f : M \rightarrow M$ eine Abbildung.^a Wir nennen einen Punkt $\bar{x} \in M$ Fixpunkt von f falls $f(\bar{x}) = \bar{x}$.

^aWenn Definitionsbereich und Wertebereich gleich sind, sprechen wir auch häufig von einer *Selbstabbildung*.

Der Banach'sche Fixpunktsatz gibt einen Algorithmus an, mit dem man Fixpunkte bei sogenannten *strikten Kontraktionsabbildungen* finden kann, d.h.

Definition 11.2.33 — (STRIKTE) KONTRAKTIONSABBILDUNG

Es seien (X, d) ein metrischer Raum und $f : X \rightarrow X$ eine (Selbst-)Abbildung.

(i) f heißt *Kontraktionsabbildung* wenn

$$d(f(x), f(y)) \leq d(x, y) \quad \forall x, y \in X.$$

(ii) f heißt *strikte Kontraktionsabbildung* wenn es ein $q \in [0, 1)$ gibt mit

$$d(f(x), f(y)) \leq q \cdot d(x, y) \quad \forall x, y \in X. \quad (686)$$

Beispiel 11.2.34

Es sei (X, d) ein beliebiger metrischer Raum und $f : X \rightarrow X$ die *identische Abbildung*, d.h. $f(x) = x$ für alle $x \in X$. Dann ist f eine Kontraktionsabbildung, denn

$$d(f(x), f(y)) = d(x, y).$$

Jedoch handelt es sich niemals um eine strikte Kontraktionsabbildung, es sei denn X besteht aus einem einzigen Punkt. In der Tat: Wäre für ein $q \in [0, 1)$ (686) wahr so wäre für alle $x, y \in X$

$$d(f(x), f(y)) \leq qd(x, y) \quad \text{d.h. in unserem Fall} \quad d(x, y) \leq qd(x, y),$$

Das würde klarerweise bedeuten, dass $d(x, y) = 0$ für alle $x, y \in X$. Aus (M1) von Definition 11.2.1 folgt, dass $x = y$ für alle $x, y \in X$. Somit ist X einelementig.

Beispiel 11.2.35

Es sei $X = \mathbb{R}$ und $d = d_{\text{euc}}$. Definiere $f : X \rightarrow X$ durch $f(x) := \frac{1}{2}x + 1$.

Behauptung. Dann ist f eine strikte Kontraktionsabbildung.

Beweis. Es seien $x, y \in X$. Dann gilt

$$d(f(x), f(y)) = d_{\text{euc}}(\tfrac{1}{2}x + 1, \tfrac{1}{2}y + 1) = |\tfrac{1}{2}x + 1 - (\tfrac{1}{2}y + 1)| = |\tfrac{1}{2}x - \tfrac{1}{2}y| = \tfrac{1}{2}|x - y| = \tfrac{1}{2}d(x, y).$$

Gleichung (686) ist also mit $q = \frac{1}{2} \in [0, 1)$ erfüllt. Somit ist f eine strikte Kontraktionsabbildung.

Nun sind wir in der Lage, den Banach'schen Fixpunktsatz zu formulieren und zu beweisen

Satz 11.2.36 — BANACH'SCHER FIXPUNKTSATZ

Es sei (X, d) ein vollständiger metrischer Raum und $f : X \rightarrow X$ eine strikte Kontraktionsabbildung. Dann hat f genau einen Fixpunkt, d.h. es gibt genau ein $\bar{x} \in X$ mit $f(\bar{x}) = \bar{x}$.

Beweis. Es sei $f : X \rightarrow X$ eine strikte Kontraktion, etwa sei $q \in [0, 1)$ so dass

$$d(f(x), f(y)) \leq qd(x, y) \quad \forall x, y \in X. \quad (687)$$

SCHRITT 1. Existenz eines Fixpunktes.

Wir geben einen Algorithmus an, der einen Fixpunkt findet. Dazu wähle einen beliebigen Punkt $x_0 \in X$ und definiere für $n \in \mathbb{N}_0$ rekursiv

$$x_{n+1} := f(x_n) \quad (688)$$

(d.h. $x_1 = f(x_0), x_2 = f(x_1), x_3 = f(x_2), \dots$, vgl. Bemerkung 3.3.5).

ZWISCHENBEHAUPTUNG. Die Folge $(x_n)_{n \in \mathbb{N}}$ ist konvergent

Hierzu genügt es wegen der Vollständigkeit zu zeigen, dass $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge ist. Es seien dazu $n, m \in \mathbb{N}$, oBdA $n < m$. Dann schätzen wir ab

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_m) \\ &\leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + d(x_{n+2}, x_m) \\ &\leq \dots \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + d(x_{n+2}, x_{n+3}) + \dots + d(x_{m-1}, x_m) \\ &\leq \sum_{k=n}^{m-1} d(x_k, x_{k+1}). \end{aligned} \quad (689)$$

Nun können wir auch für $k \in \mathbb{N}$ beliebig abschätzen

$$d(x_k, x_{k+1}) \stackrel{(688)}{=} d(f(x_{k-1}), f(x_k)) \stackrel{(687)}{\leq} qd(x_{k-1}, x_k). \quad (690)$$

Diese Ungleichung kann man auch iterieren:

$$d(x_k, x_{k+1}) \stackrel{(690)}{\leq} qd(x_{k-1}, x_k) \stackrel{(690)}{\leq} q^2 d(x_{k-2}, x_{k-1}) \stackrel{(690)}{\leq} \dots \stackrel{(690)}{\leq} q^k d(x_0, x_1) = q^k d(x_0, f(x_0)).$$

Setzen wir diese Abschätzung nun in (689) ein so erhalten wir

$$d(x_n, x_m) \leq \sum_{k=n}^{m-1} q^k d(x_0, f(x_0)) = q^n \sum_{k=n}^{m-1} q^{k-n} d(x_0, f(x_0)) \stackrel{\text{Indexversch.}}{=} q^n d(x_0, f(x_0)) \sum_{k=0}^{m-1-n} q^k.$$

Wir erhalten mit der geometrischen Summenformel

$$d(x_n, x_m) \leq q^n d(x_0, f(x_0)) \frac{1 - q^{m-n}}{1 - q}.$$

Da $q \in [0, 1)$ und $m - n > 0$ gilt $\frac{1 - q^{m-n}}{1 - q} \leq \frac{1}{1 - q}$ und daher

$$d(x_n, x_m) \leq q^n \frac{d(x_0, f(x_0))}{1 - q}.$$

Will man die obige Formel auch ohne die Annahme $m > n$ aufschreiben so kann man schreiben

$$d(x_n, x_m) \leq \frac{d(x_0, f(x_0))}{1 - q} q^{\min\{n, m\}} \quad \forall n, m \in \mathbb{N}_0.$$

Nun sei $\varepsilon > 0$ beliebig vorgegeben. Da $\lim_{n \rightarrow \infty} q^n \frac{d(x_0, f(x_0))}{1 - q} = 0$ gibt es ein $n_0 \in \mathbb{N}$ mit

$$\frac{d(x_0, f(x_0))}{1 - q} q^n \leq \varepsilon \quad \forall n \geq n_0.$$

Für $n, m \geq n_0$ gilt dann

$$d(x_n, x_m) \leq \frac{d(x_0, f(x_0))}{1 - q} q^{\min\{n, m\}} \leq_{q < 1} \frac{d(x_0, f(x_0))}{1 - q} q^{n_0} \leq \varepsilon.$$

Das zeigt, dass $(x_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge in (X, d) ist. Da (X, d) vollständig ist, gibt es ein $\hat{x} \in X$ mit $x_n \xrightarrow{n \rightarrow \infty} \hat{x}$

ZWISCHENBEHAUPTUNG ENDE.

Nun werden wir zeigen dass der konstruierte Grenzwert $\hat{x} = \lim_{n \rightarrow \infty} x_n$ ein Fixpunkt von f ist. Dazu schätzen wir für beliebiges $n \in \mathbb{N}$ ab

$$\begin{aligned} 0 \leq d(f(\hat{x}), \hat{x}) &\leq d(f(\hat{x}), f(x_n)) + d(f(x_n), \hat{x}) \stackrel{(688)}{=} d(f(\hat{x}), f(x_n)) + d(x_{n+1}, \hat{x}) \\ &\leq qd(\hat{x}, x_n) + d(x_{n+1}, \hat{x}). \end{aligned}$$

Da $x_n \xrightarrow[n \rightarrow \infty]{d} \hat{x}$ gilt $\lim_{n \rightarrow \infty} d(x_n, \hat{x}) = 0$. Somit gilt natürlich auch $\lim_{n \rightarrow \infty} d(x_{n+1}, \hat{x}) = 0$. Man erhält

$$0 \leq d(f(\hat{x}), \hat{x}) \leq qd(\hat{x}, x_n) + d(x_{n+1}, \hat{x}) \xrightarrow{n \rightarrow \infty} 0$$

und folgert, dass $d(f(\hat{x}), \hat{x}) = 0$ gelten muss. Also gilt nach (M1) auch $f(\hat{x}) = \hat{x}$. Die Behauptung folgt.

SCHRITT 2. Eindeutigkeit des Fixpunktes.

Im Folgenden zeigen wir noch, dass es nur einen Fixpunkt von f geben kann. In der Tat: Sind x_1, x_2 Fixpunkte von f so gilt

$$d(x_1, x_2) \stackrel{\text{Def 11.2.32}}{=} d(f(x_1), f(x_2)) \stackrel{(687)}{\leq} qd(x_1, x_2).$$

Formen wir um, so erhalten wir $(1 - q)d(x_1, x_2) \leq 0$. Da $q < 1$ ist dies aber nur möglich wenn $d(x_1, x_2) = 0$. Es folgt, dass $x_1 = x_2$.

^aDies ist machbar, da nach Definition 11.2.1 $X \neq \emptyset$

Bemerkung 11.2.37

In Satz 11.2.36 wurde vorausgesetzt, dass f eine strikte Kontraktionsabbildung ist. Diese Voraussetzung kann nicht angeschwächt werden: Der Satz gilt nicht, wenn man nur eine Kontraktionsabbildung fordert. Hierbei können sowohl Existenz als auch Eindeutigkeit schief gehen.

BEISPIEL 1. *Eine Kontraktionsabbildung ohne Fixpunkte.* Betrachte den vollständigen metrischen Raum $(X, d) = (\mathbb{R}, d_{\text{euc}})$ und die Abbildung

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) := x + 1$$

Man überzeugt sich leicht, dass f eine Kontraktionsabbildung ist, denn für $x, y \in \mathbb{R}$ gilt

$$d_{\text{euc}}(f(x), f(y)) = |f(x) - f(y)| = |x + 1 - (y + 1)| = |x - y| = d_{\text{euc}}(x, y) \leq d_{\text{euc}}(x, y).$$

Eine strikte Kontraktion liegt allerdings klarerweise nicht vor. Einen Fixpunkt hat f auch nicht, denn

$$f(\bar{x}) = \bar{x} \iff \bar{x} + 1 = \bar{x} \iff 1 = 0 \quad \nexists.$$

BEISPIEL 2. *Eine Kontraktionsabbildung mit mehr als einem Fixpunkt.* Wiederum könnte man $(X, d) = (\mathbb{R}, d_{\text{euc}})$ betrachten und f als die identische Abbildung wählen. Diese ist nach Beispiel 11.2.34 eine Kontraktionsabbildung, aber keine strikte Kontraktionsabbildung. Außerdem hat die Abbildung unendlich viele Fixpunkte, denn jedes Element aus $\mathbb{R}$ ist klarerweise ein Fixpunkt von f .

Anwendung: Beweis des lokalen Umkehrsatzes (Satz 10.7.1).

Der Banach'sche Fixpunktsatz versetzt uns in die Lage, Satz 10.7.1 zu beweisen. Für die Lesefreundlichkeit wiederholen wir nun die Aussage von Satz 10.7.1

Wiederholung Satz 10.7.1. LOKALER UMKEHRSATZ.

Es sei $D \subseteq \mathbb{R}^d$ offen und $F : D \rightarrow \mathbb{R}^d$ auf D stetig partiell differenzierbar. Ferner sei $x_0 \in D$ so, dass

$$\det(J_F(x_0)) \neq 0.$$

Dann gibt es eine offene Menge $U_0 \subseteq D$ mit $x_0 \in U_0$ so, dass

- $V_0 := F(U_0)$ offen ist,
- $F|_{U_0} : U_0 \rightarrow V_0$ bijektiv ist,
- $(F|_{U_0})^{-1} : V_0 \rightarrow U_0$ stetig partiell differenzierbar ist.

^ad.h. notwendigerweise $m = d$, weswegen auch $J_F(x_0) \in \mathbb{R}^{d \times d}$ quadratisch ist.

Vor dem eigentlichen Beweis benötigen wir allerdings zuerst ein vorbereitendes Lemma:

Lemma 11.2.38

Es sei $A \in \mathbb{R}^{m \times d}$ eine Matrix, etwa $A = (a_{ij})_{j=1, \dots, d}^{i=1, \dots, m}$. Definiere die sogenannte *Hilbert-Schmidt-Norm* von A durch

$$\|A\|_{HS} := \sqrt{\sum_{i=1}^m \sum_{j=1}^d a_{ij}^2}. \quad (691)$$

Dann gilt für alle $x \in \mathbb{R}^d$ die Ungleichung

$$\|Ax\| \leq \|A\|_{HS} \|x\|.$$

Beweis. Es sei $x = (x^{(1)}, \dots, x^{(d)})^T \in \mathbb{R}^d$ beliebig. Wir wissen aus (583), dass die i -te Komponente des Vektors Ax durch

$$\sum_{j=1}^d a_{ij} x^{(j)}$$

gegeben ist. Daher gilt

$$\|Ax\|^2 = \sum_{i=1}^m \left(\sum_{j=1}^d a_{ij} x^{(j)} \right)^2 = \sum_{i=1}^m \left| \sum_{j=1}^d a_{ij} x^{(j)} \right|^2.$$

In der inneren Summe können wir nun für festes i die Cauchy-Schwarz-Ungleichung (Lemma 10.1.3) anwenden und erhalten

$$\|Ax\|^2 \leq \sum_{i=1}^m \left(\sum_{j=1}^d a_{ij}^2 \right) \left(\sum_{j=1}^d (x^{(j)})^2 \right).$$

Nun gilt $\sum_{j=1}^d (x^{(j)})^2 = \|x\|^2$ und daher hat man

$$\|Ax\|^2 \leq \left(\sum_{i=1}^m \sum_{j=1}^d a_{ij}^2 \right) \|x\|^2 \stackrel{(691)}{=} \|A\|_{HS}^2 \|x\|^2.$$

Nutzt man nun die Monotonie der Wurzelfunktion auf $[0, \infty)$ so erhält man

$$\|Ax\| \leq \|A\|_{HS} \|x\|.$$

Beweis (von Satz 10.7.1.) Es sei $F : D \rightarrow \mathbb{R}^d$ und $x_0 \in D$ wie in der Aussage. Insbesondere ist $\det(J_F(x_0)) \neq 0$. Also ist $J_F(x_0)$ invertierbar, vgl. Bemerkung 10.7.2.

SCHRITT 1. Reduktion auf eine Modellsituation.

Wir überlegen uns zunächst, dass wir ohne Einschränkung $x_0 = 0$, $F(x_0) = 0$ und $J_F(x_0) = I_{d \times d}$ (die $d \times d$ -Einheitsmatrix) annehmen können. In der Tat: Falls obiges nicht der Fall ist betrachte $\tilde{D} := \{y - x_0 : y \in D\}$ und definiere $\tilde{F} : D \rightarrow \mathbb{R}^d$ durch

$$\tilde{F}(x) = J_F(x_0)^{-1}(F(x + x_0) - F(x_0)).$$

Dann gilt $\tilde{F}(0) = J_F(x_0)^{-1}(F(x_0) - F(x_0)) = J_F(x_0)^{-1} \cdot 0 = 0$ und (Übungsaufgabe)

$$J_{\tilde{F}}(x) = J_F(x_0)^{-1} J_F(x + x_0) \quad \text{d.h.} \quad J_{\tilde{F}}(x_0) = J_F(x_0)^{-1} J_F(x_0) = I_{d \times d}.$$

Man überzeugt sich auch leicht davon, aus der Umkehrbarkeit von $\tilde{F}$ auf einer offenen Menge $\tilde{U}_0 \subseteq \tilde{D}$ auch die Umkehrbarkeit von F auf $U_0 := \{x + x_0 : x \in \tilde{U}_0\}$ folgen würde. Somit muss der Satz also nur im Spezialfall

$$x_0 = F(x_0) = 0, \quad J_F(0) = I_{d \times d} \tag{692}$$

gezeigt werden. Ab sofort gelte dieser Spezialfall ohne Einschränkung.

SCHRITT 2. Bestimmung von U_0, V_0 über Umformulierung als Fixpunktproblem.

Es sei $y \in \mathbb{R}^n$. Definiere nun die Abbildung $T_y : D \rightarrow \mathbb{R}^d$ durch

$$T_y(x) := y + x - F(x) \quad \text{für } x \in D. \tag{693}$$

Desweiteren bezeichne $T_y^{(j)}(x) := T_y(x)^{(j)}$ die j -te Koordinate von $T_y(x)$. Beachte: T_y ist stetig partiell differenzierbar und für $x \in D$ gilt

$$J_{T_y}(x) = I_{d \times d} - J_F(x). \tag{694}$$

Außerdem hat man

$$F(x) = y \iff 0 = y - F(x) \iff x = y + x - F(x) \iff T_y(x) = x. \quad (695)$$

Mit anderen Worten gilt $F(x) = y$ genau dann wenn x ein Fixpunkt von T_y ist. Wir werden nun also die Fixpunkte von T_y untersuchen. Hierfür werden wir den Banach'schen Fixpunktsatz (Satz 11.2.36) benutzen. Dazu müssen wir zuerst ein Setup finden, um T_y zu einer strikten Kontraktionsabbildung zwischen vollständigen metrischen Räumen zu machen. Für $x \in D$ gilt

$$\|J_F(x) - I_{d \times d}\|_{HS} \stackrel{(692)}{=} \|J_F(x) - J_F(0)\|_{HS} = \sqrt{\sum_{i=1}^d \sum_{j=1}^d (\partial_j F^{(i)}(x) - \partial_j F^{(i)}(0))^2} \quad (696)$$

Da F stetig partiell differenzierbar ist gilt

$$\lim_{x \rightarrow 0} \sqrt{\sum_{i=1}^d \sum_{j=1}^d (\partial_j F^{(i)}(x) - \partial_j F^{(i)}(0))^2} = 0$$

Also auch (nach (696)) $\lim_{x \rightarrow 0} \|J_F(x) - I_{d \times d}\|_{HS} = 0$. Wendet man Satz 10.4.5 mit $\varepsilon = \frac{1}{2\sqrt{d}}$ an so findet man

$$\exists \delta > 0 : \quad \forall x \in D \text{ mit } \|x - 0\| < \delta \quad \text{gilt} \quad \|J_F(x) - I_{d \times d}\|_{HS} \leq \frac{1}{2\sqrt{d}}. \quad (697)$$

Da $D \subseteq \mathbb{R}^d$ offen ist, können wir ein $r \in (0, \delta)$ wählen mit $\overline{B_r(0)} = \{x \in \mathbb{R}^d : \|x\| \leq r\} \subseteq D$. Fixiere nun $x, x' \in \overline{B_r(0)}$ und betrachte die (stetig differenzierbare) Kurve $\gamma : [0, 1] \rightarrow \mathbb{R}, \gamma(t) := tx + (1 - t)x'$. Nun gilt für $j \in \{1, \dots, d\}$

$$\begin{aligned} |T_y^{(j)}(x) - T_y^{(j)}(x')| &= |T_y^{(j)}(\gamma(1)) - T_y^{(j)}(\gamma(0))| = \left| \int_0^1 \frac{d}{dt} (T_y^{(j)} \circ \gamma)(t) \, dt \right| \\ &\stackrel{\text{Kor 10.5.34}}{=} \left| \int_0^1 J_{T_y^{(j)}}(\gamma(t)) \gamma'(t) \, dt \right| = \left| \int_0^1 J_{T_y^{(j)}}(tx + (1 - t)x')(x - x') \, dt \right| \\ &\stackrel{\text{Kor 8.2.6}}{\leq} \int_0^1 |J_{T_y^{(j)}}(tx + (1 - t)x')(x - x')| \, dt \\ &\leq \int_0^1 \|J_{T_y^{(j)}}(tx + (1 - t)x')\|_{HS} \|x - x'\| \, dt \end{aligned} \quad (698)$$

Man hat außerdem (weil $J_{T_y^{(j)}}$ eine $1 \times d$ -Matrix ist)

$$\begin{aligned} \|J_{T_y^{(j)}}(tx + (1-t)x')\|_{HS} &= \sqrt{\sum_{i=1}^d |\partial_i T_y^{(j)}(tx + (1-t)x')|^2} \\ &\leq \sqrt{\sum_{i=1}^d \sum_{j=1}^d |\partial_i T_y^{(j)}(tx + (1-t)x')|^2} \\ &= \|J_{T_y}(tx + (1-t)x')\|_{HS} \stackrel{(694)}{=} \|J_F(tx + (1-t)x') - I_{d \times d}\| \leq \frac{1}{2\sqrt{d}}, \end{aligned}$$

da für alle $x, x' \in \overline{B_r(0)}$ und $t \in [0, 1]$ auch gilt, dass $tx + (1-t)x' \in \overline{B_r(0)}$ (Übungsaufgabe). Zusammen mit (698) folgern wir

$$|T_y^{(j)}(x) - T_y^{(j)}(x')| \leq \int_0^1 \frac{1}{2} \|x - x'\| \, dt = \frac{1}{2\sqrt{d}} \|x - x'\|.$$

Daher gilt für $x, x' \in \overline{B_r(0)}$

$$\|T_y(x) - T_y(x')\| \leq \sqrt{\sum_{j=1}^d |T_y^{(j)}(x) - T_y^{(j)}(x')|^2} \leq \sqrt{\sum_{j=1}^d \frac{1}{4d} \|x - x'\|^2} = \frac{1}{2} \|x - x'\|. \quad (699)$$

Daraus folgt auch für $x \in \overline{B_r(0)}$

$$\begin{aligned} \|T_y(x)\| &= \|T_y(x) - T_y(0) + T_y(0)\| \leq \|T_y(x) - T_y(0)\| + \|T_y(0)\| \\ &\leq \frac{1}{2} \|x - 0\| + \|y + 0 - F(0)\| = \frac{1}{2} \|x\| + \|y\|. \end{aligned}$$

Das zeigt: Für $y \in B_{\frac{r}{2}}(0)$ und $x \in \overline{B_r(0)}$ gilt

$$\|T_y(x)\| \leq \frac{1}{2} \underbrace{\|x\|}_{\leq r} + \underbrace{\|y\|}_{< r} < \frac{1}{2}r + \frac{r}{2} = r. \quad (700)$$

Mit anderen Worten gilt $T_y(x) \in B_r(0) \subseteq \overline{B_r(0)}$. Also: Für $y \in B_{\frac{r}{2}}(0)$ kann $T_y : \overline{B_r(0)} \rightarrow \overline{B_r(0)}$ als Selbstabbildung betrachtet werden. Die Menge $\overline{B_r(0)}$ kann hier als metrischer Raum mit Metrik d_{euc} aufgefasst werden. Dank Proposition 11.2.30 bildet $(\overline{B_r(0)}, d_{\text{euc}})$ sogar einen vollständigen metrischen Raum (weil $\overline{B_r(0)} \subseteq \mathbb{R}^d$ abgeschlossen ist). Nach (699) handelt es sich bei $T_y : \overline{B_r(0)} \rightarrow \overline{B_r(0)}$ auch um eine strikte Kontraktionsabbildung, denn

$$d_{\text{euc}}(T_y(x), T_y(x')) \stackrel{(699)}{\leq} \frac{1}{2} d_{\text{euc}}(x, x').$$

Nach dem Banach'schen Fixpunktsatz gibt es dann ein eindeutiges $x_y \in \overline{B_r(0)}$ mit $T_y(x_y) = x_y$. Übrigens gilt sogar, dass $x_y \in B_r(0)$, denn $\|x_y\| = \|T_y(x_y)\| \underset{(700)}{<} r$. Wegen (695) folgt aus $T_y(x_y) = x_y$ auch $F(x_y) = y$. Wir lernen also

(*) Für alle $y \in B_{\frac{r}{2}}(0)$ gibt es ein eindeutiges $x_y \in B_r(0)$ mit $F(x_y) = y$.

Wir definieren nun die offene Menge $V_0 = B_{\frac{r}{2}}(0)$ und $U_0 := F^{-1}(V_0) \cap B_r(0)$. Beachte: Da F stetig ist^a und V_0 offen ist (vgl. Beispiel 10.3.7) ist nach Satz 10.4.14 auch $F^{-1}(V_0)$ offen. Mit Proposition 10.3.13 folgt dann die Offenheit von U_0 . Außerdem gilt $0 \in U_0$, denn nach (692) $F(0) = 0 \in V_0$ d.h. $0 \in F^{-1}(V_0)$.

SCHRITT 3. $V_0 = F(U_0)$ und $F|_{U_0} : U_0 \rightarrow V_0$ ist bijektiv.

Wir zeigen, dass es zu jedem $y \in V_0$ genau ein $x \in U_0$ gibt mit $F(x) = y$. Sei dazu $y \in V_0$. Nach (*) gibt es genau ein $x_y \in B_r(0)$ mit $F(x_y) = y$. Es gilt auch $x_y \in U_0$ (denn $F(x_y) = y \in V_0$ zeigt auch, dass $x_y \in F^{-1}(U_0)$, d.h. $x_y \in B_r(0) \cap F^{-1}(U_0)$). Das zeigt die Bijektivität vgl. Definition 1.3.9.

SCHRITT 4. Stetigkeit von $(F|_{U_0})^{-1}$ auf V_0 .

Bevor wir zur Stetigkeit kommen, zeigen wir zunächst eine hilfreiche Abschätzung. Es seien $x, x' \in U_0$. Beachte $x = x - F(x) + F(x) = T_0(x) + F(x)$ für T_0 definiert wie in (693) mit $y = 0$. Nun gilt

$$\begin{aligned} \|x - x'\| &= \|T_0(x) + F(x) - (T_0(x') + F(x'))\| = \|T_0(x) - T_0(x') + (F(x) - F(x'))\| \\ &\leq \|T_0(x) - T_0(x')\| + \|F(x) - F(x')\| \underset{(699)}{\leq} \frac{1}{2}\|x - x'\| + \|F(x) - F(x')\|. \end{aligned}$$

Wir erhalten durch Subtrahieren von $\frac{1}{2}\|x - x'\|$ auf beiden Seiten $\frac{1}{2}\|x - x'\| \leq \|F(x) - F(x')\|$ und daher

$$\|x - x'\| \leq 2\|F(x) - F(x')\|. \quad (701)$$

Nun können wir endlich die Stetigkeit von $(F|_{U_0})^{-1}$ auf V_0 zeigen. Dazu seien $y, y' \in V_0$. Definiere $x := (F|_{U_0})^{-1}(y) \in U_0$ und $x' := (F|_{U_0})^{-1}(y') \in U_0$. Dann gilt

$$\begin{aligned} \|(F|_{U_0})^{-1}(y) - (F|_{U_0})^{-1}(y')\| &= \|x - x'\| \underset{(701)}{\leq} 2\|F(x) - F(x')\| \\ &= 2\|F((F|_{U_0})^{-1}(y)) - F((F|_{U_0})^{-1}(y'))\| = 2\|y - y'\| \quad (702) \end{aligned}$$

Wegen dieser Abschätzung sprechen wir davon, dass $(F|_{U_0})^{-1}$ Lipschitzstetig mit Lipschitzkonstante 2 ist. Wie man sich leicht überlegt ist $(F|_{U_0})^{-1}$ als Konsequenz auch stetig.
b

SCHRITT 5. Für alle $x \in U_0$ ist $J_F(x) \in \mathbb{R}^{d \times d}$ eine invertierbare Matrix.

Wäre für ein $x_0 \in U_0$ die Jacobimatrix $J_F(x_0)$ nicht invertierbar so gibt es (laut linearer Algebra) einen Vektor $v \in \mathbb{R}^d, v \neq 0$, der im *Kern* von $J_F(x_0)$ liegt, d.h. $J_F(x_0)v = 0$ genügt. Dann gilt

$$\begin{aligned} \|v\| &= \|I_{d \times d} v\| = \|I_{d \times d} v - J_F(x_0)v\| = \|(I_{d \times d} - J_F(x_0))v\| \\ &\stackrel{\text{Lem 11.2.38}}{\leq} \|I_{d \times d} - J_F(x_0)\|_{HS} \|v\| \stackrel{(697)}{\leq} \frac{1}{2\sqrt{d}} \|v\|. \end{aligned}$$

Da $\frac{1}{2\sqrt{d}} < 1$ kann die obige Ungleichung nur erfüllt sein wenn $\|v\| = 0$. Dies widerspricht $v \neq 0$.

SCHRITT 6. $(F|_{U_0})^{-1}$ ist auf V_0 total differenzierbar.

Es sei dazu $y_0 \in V_0$ beliebig. Sei nun $x_0 \in U_0$ mit $F|_{U_0}(x_0) = y_0$. Wir betrachten die lineare Abbildung $L_{y_0} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ gegeben durch $L_{y_0}(v) := J_F(x_0)^{-1}v$ und rechnen mit dieser Wahl Definition 10.5.23 nach. Dazu: Für $y \in V_0 \setminus \{y_0\}$ gilt

$$\begin{aligned} &\frac{(F|_{U_0})^{-1}(y) - (F|_{U_0})^{-1}(y_0) - L_{y_0}(y - y_0)}{\|y - y_0\|} \\ &= \frac{(F|_{U_0})^{-1}(y) - (F|_{U_0})^{-1}(y_0) - J_F(x_0)^{-1}(y - y_0)}{\|y - y_0\|} \\ &= J_F(x_0)^{-1} \left(\frac{J_F(x_0)((F|_{U_0})^{-1}(y) - (F|_{U_0})^{-1}(y_0)) - (y - y_0)}{\|y - y_0\|} \right) \\ &= -J_F(x_0)^{-1} \left(\frac{(y - y_0) - J_F(x_0)((F|_{U_0})^{-1}(y) - (F|_{U_0})^{-1}(y_0))}{\|y - y_0\|} \right) \end{aligned}$$

Setzt man nun $x = (F|_{U_0})^{-1}(y)$ und verwendet $x_0 = (F|_{U_0})^{-1}(y_0)$ so erhält man

$$\begin{aligned} &\frac{(F|_{U_0})^{-1}(y) - (F|_{U_0})^{-1}(y_0) - L_{y_0}(y - y_0)}{\|y - y_0\|} = -J_F(x_0)^{-1} \left(\frac{F(x) - F(x_0) - J_F(x_0)(x - x_0)}{\|F(x) - F(x_0)\|} \right) \\ &= -\frac{\|x - x_0\|}{\|F(x) - F(x_0)\|} J_F(x_0)^{-1} \left(\frac{F(x) - F(x_0) - J_F(x_0)(x - x_0)}{\|x - x_0\|} \right) \end{aligned}$$

Es sei nun $(y_n)_{n \in \mathbb{N}} \subseteq V_0$ eine Folge mit $\lim_{n \rightarrow \infty} y_n = y_0$. Definiert man $x_n = (F|_{U_0})^{-1}(y_n)$ so hat man

$$\begin{aligned} &\frac{(F|_{U_0})^{-1}(y_n) - (F|_{U_0})^{-1}(y_0) - L_{y_0}(y_n - y_0)}{\|y_n - y_0\|} \\ &= -\frac{\|x_n - x_0\|}{\|F(x_n) - F(x_0)\|} J_F(x_0)^{-1} \left(\frac{F(x_n) - F(x_0) - J_F(x_0)(x_n - x_0)}{\|x_n - x_0\|} \right). \quad (703) \end{aligned}$$

Da F in x_0 (als stetig partiell differenzierbare Funktion) total diffbar ist gilt

$$\frac{F(x_n) - F(x_0) - J_F(x_0)(x_n - x_0)}{\|x_n - x_0\|} \xrightarrow{n \rightarrow \infty} 0$$

. Nach Bemerkung 10.5.22 gilt dann auch $J_F(x_0)^{-1} \left(\frac{F(x_n) - F(x_0) - J_F(x_0)(x_n - x_0)}{\|x_n - x_0\|} \right) \xrightarrow{n \rightarrow \infty} 0$.
Nach (701) hat man ferner

$$0 \leq \frac{\|x_n - x_0\|}{\|F(x_n) - F(x_0)\|} \leq 2,$$

sodass der Term auf der rechten Seite von (703) das Produkt einer beschränkten Folge und einer Nullfolge ist. Wir folgern aus (703) und Proposition 10.2.3 (iii), dass

$$\lim_{n \rightarrow \infty} \frac{(F|_{U_0})^{-1}(y_n) - (F|_{U_0})^{-1}(y_0) - L_{y_0}(y_n - y_0)}{\|y_n - y_0\|} = 0.$$

Da $(y_n)_{n \in \mathbb{N}} \subseteq V_0$ mit $\lim_{n \rightarrow \infty} y_n = y_0$ beliebig war folgt

$$\lim_{y \rightarrow y_0} \frac{(F|_{U_0})^{-1}(y) - (F|_{U_0})^{-1}(y_0) - L_{y_0}(y - y_0)}{\|y - y_0\|} = 0.$$

Es folgt dass $(F|_{U_0})^{-1}$ bei y_0 total differenzierbar ist mit totaler Ableitung L_{y_0} . Dadurch ist $(F|_{U_0})^{-1}$ übrigens auch partiell differenzierbar und es gilt nach Satz 10.5.26(ii)

$$J_{(F|_{U_0})^{-1}}(y_0)v = L_{y_0}(v) = J_F(x_0)^{-1}v \quad \forall v \in \mathbb{R}^d$$

Betrachtet man nacheinandere $v = e_1, \dots, v = e_d$ so sieht man, dass alle Spalten von $J_{(F|_{U_0})^{-1}}(y_0)$ und $J_F(x_0)^{-1}$ übereinstimmen. Das bedeutet

$$J_{(F|_{U_0})^{-1}}(y_0) = J_F(x_0)^{-1} = J_F((F|_{U_0})^{-1}(y_0))^{-1} \quad (704)$$

SCHRITT 7. Die partiellen Ableitungen von $(F|_{U_0})^{-1}$ sind stetig auf V_0 .

Es nun noch gezeigt werden, dass jeder Eintrag von $y \mapsto J_{(F|_{U_0})^{-1}}(y)$ stetig auf V_0 ist. Verwendet man (704) so sieht man

$$J_{(F|_{U_0})^{-1}}(y) = J_F((F|_{U_0})^{-1}(y))^{-1}.$$

Aus der linearen Algebra wissen wir, dass für eine invertierbare Matrix $A \in \mathbb{R}^{d \times d}$ gilt, dass

$$A^{-1} = \frac{1}{\det(A)} \tilde{A} \quad \text{wobei} \quad \tilde{A} = ((-1)^{i+j} \det(A^{(j,i)}))_{j=1, \dots, d}^{i=1, \dots, d}.$$

Hierbei ist $A^{(j,i)}$ die Matrix, die durch Streichen der i -ten Zeile und der j -ten Spalte entsteht. Da $\det(A)$ und $\det(A^{(j,i)})$ stets Polynome in den Einträgen von A sind (vgl. (638)), ist jeder Eintrag von A^{-1} als Quotient von Polynomen aus Einträgen von A darstellbar (mit nichtverschwindendem Nenner, weil A ja invertierbar ist). Somit ist $J_F((F|_{U_0})^{-1}(y))^{-1}$ ein Quotient von Polynomen aus den Einträgen von $J_F((F|_{U_0})^{-1}(y))$. Da die Einträge von $z \mapsto J_F(z)^c$ und $(F|_{U_0})^{-1}$ stetig sind (vgl. Schritt 4), ist jeder Eintrag von $y \mapsto J_{(F|_{U_0})^{-1}}(y)$ als Komposition stetiger Funktionen (Quotient, Polynome, partielle Ableitungen von F und $(F|_{U_0})^{-1}$) stetig. Die Behauptung folgt.

^aDie Stetigkeit von F folgt, weil F als stetig partiell diffbare Funktion auch total diffbar ist (vgl. Satz 10.5.30). Deswegen ist F auch stetig!

^bIn der Tat: Ist $y_0 \in V_0$ und $(y_n)_{n \in \mathbb{N}} \subseteq V_0$ eine beliebige gegen y_0 konvergente Folge so gilt nach (702)

$$0 \leq \|(F|_{U_0})^{-1}(y_n) - (F|_{U_0})^{-1}(y_0)\| \leq 2\|y_n - y_0\| \xrightarrow{n \rightarrow \infty} 0.$$

Das zeigt, dass $\lim_{n \rightarrow \infty} (F|_{U_0})^{-1}(y_n) = (F|_{U_0})^{-1}(y_0)$.

^cd.h. die partiellen Ableitungen von F

Stetigkeit und Kompaktheit in metrischen Räumen

Aus dem Studium topologischer Räume kennen wir die Begriffe der Stetigkeit und der Kompaktheit. In metrischen Räumen, wo wir das Konzept der Folgenkonvergenz haben, können wir aber auch andere Konzepte dieser Begriffe einführen, nämlich die *Folgenstetigkeit* und die *Folgenkompaktheit*. Wir müssen uns danach über die Relation dieser Konzepte unterhalten. Zunächst sprechen wir über die Folgenstetigkeit.

Definition 11.2.39 — FOLGENSTETIGKEIT

Es seien (X_1, d_1) und (X_2, d_2) metrische Räume sowie $x_0 \in X_1$. Eine Abbildung $f : X_1 \rightarrow X_2$ heißt *folgenstetig in x_0* wenn für jede Folge $(x_n)_{n \in \mathbb{N}} \subseteq X_1$ mit $x_n \xrightarrow[d_1]{n \rightarrow \infty} x_0$ gilt, dass $f(x_n) \xrightarrow[d_2]{n \rightarrow \infty} f(x_0)$.

Wir nennen f auf X_1 folgenstetig falls f in jedem Punkt $x_0 \in X_1$ folgenstetig ist.

Beispiel 11.2.40

Es sei $X_1 = X_2 = X$ und $d_1 = d_2 = d$ für einen metrischen Raum (X, d) . Ferner sei $f : X \rightarrow X$ eine Kontraktionsabbildung, d.h. nach Definition 11.2.33, dass

$$d(f(x), f(y)) \leq d(x, y) \quad \forall x, y \in X.$$

BEHAUPTUNG. f ist auf X folgenstetig.

Beweis. Es sei $x_0 \in X$ beliebig. Wir zeigen nun mit Definition 11.2.39, dass f bei x_0

folgenstetig ist. Dazu sei $(x_n)_{n \in \mathbb{N}}$ eine beliebige Folge mit $x_n \xrightarrow[d]{n \rightarrow \infty} x_0$. Dann gilt

$$0 \leq d(f(x_n), f(x_0)) \leq d(x_n, x_0) \xrightarrow{n \rightarrow \infty} 0$$

Mit dem Sandwichlemma folgern wir, dass $d(f(x_n), f(x_0)) \xrightarrow[d]{n \rightarrow \infty} 0$, weswegen $f(x_n) \xrightarrow[d]{n \rightarrow \infty} f(x_0)$.

Wir klären nun den Zusammenhang zwischen Folgenstetigkeit und topologischer Stetigkeit auf.

Satz 11.2.41

Es seien (X_1, d_1) und (X_2, d_2) metrische Räume und $f : X_1 \rightarrow X_2$ eine Abbildung. Ferner seien $\mathcal{O}_{d_1}$ und $\mathcal{O}_{d_2}$ die induzierten Topologien auf X_1 bzw. X_2 , vgl. Definition 11.2.14. Dann sind folgende Aussagen äquivalent

- (a) $f : X_1 \rightarrow X_2$ ist folgenstetig
- (b) $f : X_1 \rightarrow X_2$ ist eine stetige Abbildung zwischen den topologischen Räumen $(X_1, \mathcal{O}_{d_1})$ und $(X_2, \mathcal{O}_{d_2})$, vgl. Definition 11.1.1.

Beweisidee. Elemente von $\mathcal{O}_{d_i}$, $i = 1, 2$ nennen wir nach Definition 11.2.12 auch offene Mengen in (X_i, d_i) . Schreiben wir zunächst (b) mit Definition 11.1.1 um so haben wir die äquivalente Bedingung

(b') Für alle offenen Mengen U in (X_2, d_2) ist $f^{-1}(U)$ offen in (X_1, d_1) .

Nun zeigt man zuerst wie im Beweis von Satz 6.1.23, dass die Folgenstetigkeit von $f : X_1 \rightarrow X_2$ in $x_0 \in X_1$ äquivalent zum folgenden *Epsilon-Delta-Kriterium* ist:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x \in X_1 \quad \text{mit} \quad d_1(x, x_0) < \delta \quad \text{gilt} \quad d_2(f(x), f(x_0)) < \varepsilon.$$

Folgenstetigkeit auf ganz X_1 ist also äquivalent zur Aussage

(a') $\forall x_0 \in X_1 \quad \forall \varepsilon > 0 \quad \exists \delta > 0 : \quad \forall x \in X_1 \quad \text{mit} \quad d_1(x, x_0) < \delta \quad \text{gilt} \quad d_2(f(x), f(x_0)) < \varepsilon.$

Dass (a') und (b') äquivalent sind kann genau wie im Beweis von Satz 10.4.14 diskutiert werden. Es muss lediglich der Ausdruck $\|x - y\|$ stets durch $d(x, y)$ ersetzt werden.

Folgenstetigkeit in metrischen Räumen ist also nichts weiter als eine weitere Charakterisierung der Stetigkeit im Sinne von Definition 11.1.1.

Auch den Begriff der Kompaktheit können wir mit dem Konzept der Folgenkompaktheit charakterisieren. Dies werden wir nun am Ende dieses Abschnittes noch studieren.

Definition 11.2.42 — FOLGENKOMPAKTHEIT

Es sei (X, d) ein metrischer Raum. Eine Teilmenge $M \subseteq X$ heißt *folgenkompakt* wenn für jede Folge $(x_n)_{n \in \mathbb{N}} \subseteq M$ eine konvergente Teilfolge $(x_{\ell_n})_{n \in \mathbb{N}}$ existiert, deren Grenzwert

wiederum in M liegt.

Bemerkung 11.2.43

Wir wissen bereits aus Bemerkung 10.3.21, dass im Spezialfall $(X, d) = (\mathbb{R}^d, d_{\text{euc}})$ die Folgenkompaktheit äquivalent zur Kompaktheit ist. Diesen Erkenntnis wollen wir nun auf den allgemeinen Fall ausweiten.

Wir lernen nun den sogenannten *abstrakten Satz von Heine-Borel* kennen, welcher die kennengelernten Kompaktheitskonzepte in metrischen Räumen miteinander in Beziehung setzt. Beweisen werden wir diesen Satz nicht. Wir verweisen auf die Vorlesung *Funktionalanalysis*.

Satz 11.2.44 — ABSTRAKTER SATZ VON HEINE-BOREL

Es sei (X, d) ein metrischer Raum mit induzierter Topologie $\mathcal{O}_d$ und $M \subseteq X$. Dann sind folgende Aussagen äquivalent

- (a) M ist folgenkompakt
- (b) M ist eine kompakte Teilmenge von $(X, \mathcal{O}_d)$ im Sinne von Definition 11.1.8.

Beweis. Ohne Beweis.

11.3 Normierte Räume

Wir haben nun topologische Räume und metrische Räume studiert. Jetzt können wir noch konkreter werden und sogenannte *normierte Räume* studieren. Normierte Räume benötigen zum ersten Mal eine *algebraische Struktur*, nämlich die eines *Vektorraumes*. Wir verweisen auf die lineare Algebra für die Definition eines solchen Vektorraumes. Eine *Norm* auf einem Vektorraum ist dann eine Abbildung, die Abstände misst – und zwar so, dass diese Abstandsmessung *verträglich mit den algebraischen Operationen* (Vektorsumme, Multiplikation mit einem Skalar) ist. Den Skalkörper bezeichnen wir hier wieder mit $\mathbb{K}$ (wobei in unserer Betrachtung $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$).

Definition 11.3.1 — NORM

Es sei X ein Vektorraum über $\mathbb{K}$. Eine Abbildung $\|\cdot\|_X : X \rightarrow \mathbb{R}$ (d.h. $x \mapsto \|x\|_X$ heißt *Norm auf X* falls

- (N1) *Definitheit*. Für alle $x \in X$ gilt $\|x\|_X \geq 0$ und $(\|x\|_X = 0 \iff x = 0)^a$
- (N2) *Homogenität*. Für alle $x \in X$ und $\lambda \in \mathbb{K}$ gilt $\|\lambda x\|_X = |\lambda| \|x\|_X$.^(b)
- (N3) *Dreiecksungleichung*. Für alle $x, y \in X$ gilt $\|x + y\|_X \leq \|x\|_X + \|y\|_X$.

Das Paar $(X, \|\cdot\|_X)$ heißt dann *normierter Raum*.

^awobei hier $x = 0$ bedeutet, dass X das Nullelement des vorliegenden Vektorraumes ist.

^bMit $|\cdot|$ ist der Absolutbetrag auf $\mathbb{K}$ (also entweder der reelle oder der komplexe Absolutbetrag) gemeint.

Schon in Definition 10.4.22 haben wir Normen auf $\mathbb{R}^d$ besprochen.

Beispiel 11.3.2

Es sei $N : \mathbb{R}^d \rightarrow \mathbb{R}$ eine Norm im Sinne von Definition 10.4.22. Dann kann man durch

$$\|x\|_{\mathbb{R}^d} := N(x)$$

eine Norm auf $\mathbb{R}^d$ definieren (wie man leicht sieht wenn man (N1), (N2) und (N3) aus der obigen Definition mit Definition 10.4.22 vergleicht). Ein Spezialfall ist die Euklidische Norm $\|\cdot\|$.

Es gibt aber auch noch weitere Beispiele für Normen. So zum Beispiel kann man auch zeigen, dass auf $X = B(D, \mathbb{R}^m)$ ($D \subseteq \mathbb{R}^d$) die Abbildung

$$\|f\|_{B(D, \mathbb{R}^m)} := \sup_{x \in D} \|f(x)\|$$

eine Norm definiert. Anstatt $\|f\|_{B(D, \mathbb{R}^m)}$ schreiben wir manchmal auch kurz $\|f\|_\infty$. Normierte Räume sind Spezialfälle von metrischen Räumen, wie wir im Folgenden noch besprechen wollen.

Satz 11.3.3

Es sei $(X, \|\cdot\|_X)$ ein normierter Raum. Definiere $d : X \times X \rightarrow \mathbb{R}$ durch

$$d(x, y) := \|x - y\|_X \quad \forall x, y \in \mathbb{R}^d.$$

Dann ist d eine Metrik auf X , d.h. (X, d) ist ein metrischer Raum.

Beweis. Wir müssen die Axiome (M1), (M2), (M3) aus Definition 11.2.1 überprüfen und dürfen dabei die Eigenschaften (N1), (N2), (N3) von $\|\cdot\|_X$ aus Definition 11.3.1 verwenden.

Zu (M1). Seien $x, y \in X$. Dann gilt

$$d(x, y) = \|x - y\|_X \underset{(N1)}{\geq} 0$$

und

$$d(x, y) = 0 \iff \|x - y\|_X = 0 \underset{(N1)}{\iff} x - y = 0 \iff x = y.$$

Zu (M2). Seien $x, y \in X$. Dann gilt

$$d(x, y) = \|x - y\|_X = \|(-1)(y - x)\|_X \stackrel{(N2)}{=} |(-1)| \|y - x\|_X = \|y - x\|_X = d(y, x).$$

Somit folgt die Symmetrie von d .

Zu (M3). Seien $x, y, z \in X$. Dann gilt

$$d(x, z) = \|x - z\|_X = \|(x - y) + (y - z)\|_X \stackrel{(N3)}{\leq} \|x - y\|_X + \|y - z\|_X = d(x, y) + d(y, z).$$

Somit folgt die Dreiecksungleichung für d .

Definition 11.3.4

Die Metrik d aus Satz 11.3.3 nennen wir auch häufig die *durch $\|\cdot\|_X$ induzierte Metrik*.

Jetzt, wo normierte Räume nichts weiter als ein Spezialfall von metrischen Räumen sind, haben wir die bekannten Konzepte der Analysis auch in normierten Räumen zur Verfügung. Konkreter könnten wir folgende Begriffe definieren

- Folgenkonvergenz
- Cauchy-Folgen
- Vollständigkeit
- Stetigkeit
- Offenheit (bzw eine Topologie)
- Abgeschlossenheit
- Kompaktheit
- ...

Das zu tun überlasse ich Ihnen. Zum Abschluss dieses Abschnittes definieren wir noch den Begriff des *Banachraums*, den Sie in der Funktionalanalysis häufig studieren werden.

Definition 11.3.5 — BANACHRAUM

Ein normierter Raum $(X, \|\cdot\|_X)$ heißt *Banachraum*, wenn für die durch $\|\cdot\|_X$ induzierte Metrik d gilt, dass (X, d) vollständig ist.

Kurz gesagt: Ein Banachraum ist ein *vollständiger normierter Raum*.

Index

- r -Umgebung, 289
- Ableitung, 328, 331
 - Höherer Ordnung, 392
 - Totale Ableitung, 588, 592
- Ableitungsregel
 - Kettenregel, 603
- Ableitungsregeln
 - Ableitung der Umkehrfunktion, 348
 - Kettenregel, 345
 - Produktregel, 341
 - Quotientenregel, 341
 - Summenregel, 341
- Absolutbetrag, 27
 - Eigenschaften, 28
- Abzählbarkeit, 91
- Allgemeine Potenz, 315
- Aussage, 6
- Banach'scher Fixpunktsatz, 685
- Bijektiv, 20
- Bild, 19
- Binomialkoeffizient, 53, 141
- Binomischer Lehrsatz, 58
- Bolzano-Weierstrass, 523
- Bolzano-Weierstraß, 165
- Cauchy-Folgen
 - in Metrischen Räumen, 679
- Cauchy-Kriterium
 - Für gleichmäßige Konvergenz, 500
- Cauchy-Schwarz-Ungleichung, 515
- Definitheit, 627
- Definitheitskriterium von
 - Hurwitz-Sylvester, 637
- Die Kreiszahl π , 252
- Endlichkeit, 30
- Euler'sche Zahl, 162
- Exponentialfunktion, 238
 - Funktionalgleichung, 239
- Exponentialreihe, 238
- Extrema, 352
- Exytrema
 - im Mehrdimensionalen, 625
- Fakultät, 141
- Folge, 34
 - beschränkt, 35
 - Bestimmte Divergenz, 89
 - Cauchy, 175
 - Cauchy-Eigenschaft, 192
 - Divergenz, 71
 - Häufungswert, 166
 - Konvergenz, 65
 - in Metrischen Räumen, 674
 - Limes Inferior, 170
 - Limes Superior, 170
 - Majorisierung, 40
 - Minorisierung, 40
 - monoton, 38
 - Nullfolge, 86
 - Teilfolgen, 162, 522
- Fundamentalsatz der Analysis, 459
- Funktion, 17
 - $(\mathbb{K}_1, \mathbb{K}_2)$ -Funktion, 266
 - differenzierbar, 328
 - Graph, 267
 - Imaginärteil, 267
 - Komposition, 22, 267
 - konvex/konkav, 414
 - Maximum, 298
 - Minimum, 298
 - Monotonie, 307
 - Realteil, 267
 - Reell-Analytisch, 406
 - Riemann-integrierbar, 431
 - Stetig Differenzierbar, 392
- Funktionenfolge, 494
 - Gleichmäßige Konvergenz, 494
 - Stetigkeitsvererbung, 505
 - Punktweise Konvergenz, 494
- Funktionenräume, 669
- Funktionsgrenzwert, 269
 - Bestimmte Divergenz, 284
 - einseitig, 277
 - Epsilon-Delta-Kriterium, 285, 551

- im Unendlichen, 283
- Mehrdimensional, 548
- Ganze Zahlen, 132
- Gauß-Klammer, 155
- Geordnetes Paar, 14
- Gleichmäßige Stetigkeit, 325, 560
- Gradient, 582
- Grenzwertregeln, 75, 77, 78, 82
 - In $\mathbb{C}$, 192
- Hauptsatz der Differential -und
Integralrechnung, 466
- Hesse-Matrix, 618
- Häufungspunkt, 268
- Identische Abbildung, 18
- Imaginäre Einheit, 180
- Imaginärteil, 182
- Infimum, 107
- Injektiv, 20
- Innere Verknüpfung, 96
- Integral, 431
 - Uneigentliches Integral, 486
 - Integral über komplexwertige
Funktionen, 491
 - Linearität, 445
 - Monotonie des Integrals, 447
 - Orientiertes Integral, 457
 - Teilungsformel, 455
- Intervall, 117
- Irrationale Zahlen, 153
- Jacobi-Matrix, 581
- Junktor, 6
- Kartesisches Produkt, 14
- Komplexe Konjugation, 183
- Komplexe Zahlen, 180
- Komplexer Absolutbetrag, 186
 - Eigenschaften, 189
- Konstanzsatz, 369, 608
- Konvergenzradius, 230
- Kosinus, 246
- Kurven, 582
- Körper, 96
- Angeordneter Körper, 110
- Ordnungsvollständigkeit, 145
- L'Hospital
 - Fall $\frac{0}{0}$, 380
 - Fall $\frac{\infty}{\infty}$, 388
- Lagrange-Multiplikatorenregel, 653
- Lokale Extremstellen
 - Vorzeichenwechselkriterium, 375
 - Kriterium Gerader Ordnung, 402
 - Kriterium Zweiter Ordnung, 401
 - Kriterium zweiter Ordnung,
mehrdimensional, 628
 - Notwendiges Kriterium, 355
 - Notwendiges Kriterium,
mehrdimensional, 626
- Lokaler Umkehrsatz, 645
- Maximum, 105
- Menge, 10
 - (Topologisch) Kompakt, 663
 - Abgeschlossen, 527
 - Folgenkompakt, 537
 - Gleichheit, 11
 - Kompakt, 536
 - Leere Menge, 11
 - Offen, 527
 - Operationen, 12
 - Teilmenge, 11
 - Überdeckungskompakt, 539
- Metrik
 - Diskrete, 668
 - Euklidische, 667
- Metrischer Raum, 666
- Minimum, 105
- Mittelwertsatz
 - der Differentialrechnung, 362
 - der Integralrechnung, 459
- Natürliche Zahlen, 120
- Natürlicher Logarithmus, 312
- Norm, 521, 562
 - Euklidische, 515
 - Normäquivalenz, 567
- Ober/Unterintegral, 430

Ober/Untersumme, 427
 Ordnungsrelation, 103

 Partialbruchzerlegung, 479
 Partialsumme, 197
 partielle Ableitungen, 579
 Partielle Integration, 474
 Phönix, 482
 Polarkoordinaten, 261
 Potenz, 127
 Potenzmenge, 15
 Potenzreihen, 226
 Produktzeichen, 140
 Punkt
 Berührungspunkt, 524
 Innerer Punkt, 524
 Randpunkt, 524

 Quaderteilung, 542
 Quadratische Form, 621

 r-Umgebung, 518
 Rationale Zahlen, 144
 Realteil, 182
 Reelle Zahlen, 95, 148
 Reihe, 197
 Abkling-Kriterium, 206
 Absolut konvergent, 211
 Cauchy-Produkt, 234
 Harmonische Reihe, 200
 Leibniz-Kriterium, 208
 Majorantenkriterium, 217
 Quotientenkriterium, 223
 Umordnungssatz, 215
 Wurzelkriterium, 220
 Relation, 103
 Relativtopologie, 659
 Richtungsableitung, 574

 Satz von Heine, 325
 Satz von Heine-Borel, 541
 Satz von Schwarz, 612

 Satz von Weierstraß, 299
 im Mehrdimensionalen, 561
 Schachtelungsprinzip, 538
 Schranke, 105
 Sektorfunktion, 573
 Sinus, 246
 Stammfunktion, 464
 Stetigkeit, 291
 Folgenstetigkeit, 695
 im Mehrdimensionalen, 552
 Topologische Charakterisierung, 555
 Stetigkeit der Umkehrfunktion, 307
 Substitutionsregel, 471
 Summenzeichen, 133
 Supremum, 107
 Surjektiv, 20

 Taylorapproximation, 397
 im Mehrdimensionalen, 622
 Taylorformel, 394
 im Mehrdimensionalen, 622
 Taylorpolynom, 399
 Taylorreihe, 405
 Topologischer Raum
 Induzierte Topologie, 674
 topologischer Raum, 658

 Umkehrfunktion, 23
 Umordnung, 213
 Unendlichkeit, 30
 Urbild, 19

 Vertauschung
 Ableitung und Grenzfunktion, 512
 Integral und Funktionsgrenzwert, 506
 Vollständig, 680

 Weierstraß-M-Test, 502
 Weierstraß-Substitution, 484
 Wurzel, 151

 Zwischenwertsatz, 301