



Prof. Dr. Helga Baum

Analysis 1 und 2

Skript zur Vorlesung Analysis 1 und 2

Bachelorstudiengang Mathematik mit Lehramtsoption
für die Studienanfänger des WS 2014/15

23. Juli 2015

Analysis 1: Kapitel 1 – 5

Analysis 2: Kapitel 6 – 9

Inhaltsverzeichnis

1	Grundlagen	1
1.1	Aussagen, logische Schlüsse und Beweisstrategien	1
1.2	Mengen und Abbildungen	7
1.2.1	Mengen	7
1.2.2	Abbildungen	11
1.3	Das Beweisprinzip der vollständigen Induktion	16
2	Reelle und komplexe Zahlen	21
2.1	Die reellen Zahlen	21
2.1.1	Die Körpereigenschaften von \mathbb{R}	21
2.1.2	Die Anordnungseigenschaften von \mathbb{R}	23
2.1.3	Vollständigkeitseigenschaft der reellen Zahlen	25
2.1.4	Die Überabzählbarkeit der Menge der reellen Zahlen	30
2.1.5	Wurzeln und Potenzen reeller Zahlen	33
2.2	Die komplexen Zahlen	38
2.2.1	Geometrische Motivation für die Definition der komplexen Zahlen ...	38
2.2.2	Definition und Eigenschaften der komplexen Zahlen	40
3	Folgen und Reihen	47
3.1	Folgen in \mathbb{R} und in \mathbb{C}	48
3.1.1	Konvergente Zahlenfolgen und ihre Eigenschaften	48
3.1.2	Spezielle Eigenschaften reeller Folgen	54
3.1.3	Cauchy-Folgen und Konvergenz	60
3.2	Reihen in \mathbb{R} und \mathbb{C}	61
3.2.1	Konvergente und divergente Reihen, Konvergenzkriterien	62
3.2.2	Komplexe Potenzreihen	72
3.2.3	Anwendung: Exponentialfunktion, Logarithmusfunktion und komplexe Potenzen	74
4	Stetige Funktionen	79
4.1	Definition stetiger Funktionen und Beispiele	80

4.2	Grenzwerte von Funktionen und stetige Fortsetzungen	87
4.3	Nullstellen, Fixpunkte und Extremwerte stetiger Funktionen	92
4.4	Anwendung: Der Fundamentalsatz der Algebra	95
4.5	Die trigonometrischen und die Hyperbelfunktionen und ihre Umkehrfunktionen	97
4.5.1	Die trigonometrischen Funktionen	97
4.5.2	Die Hyperbelfunktionen	103
5	Differentialrechnung für Funktionen einer reellen Variablen	107
5.1	Differenzierbare reelle Funktionen: Rechenregeln und Beispiele	107
5.2	Die Mittelwertsätze der Differentialrechnung und Anwendungen	115
5.3	Differenzierbarkeit von Potenzreihen	122
5.4	Taylorpolynome und Extremwertprobleme	128
5.5	Taylorreihen und reell-analytische Funktionen	133
6	Integralrechnung für Funktionen einer reellen Variablen	139
6.1	Das Riemann-Integral	139
6.2	Integration und Differentiation	148
6.3	Uneigentliche Riemann-Integrale	165
6.4	Geometrische Anwendungen des Riemann-Integrals	168
6.4.1	Parametrisierte Kurven und ihre Länge	169
6.4.2	Flächeninhalt ebener Gebiete	176
6.5	Analytische Anwendung: Differentialgleichungen mit getrennten Variablen .	181
7	Metrische Räume und stetige Abbildungen	189
7.1	Definition und Beispiele metrischer Räume	189
7.2	Konvergente Folgen und Cauchy-Folgen in metrischen Räumen	194
7.3	Spezielle Teilmengen in metrischen Räumen	197
7.4	Stetige Abbildungen und ihre Eigenschaften	202
8	Differentialrechnung für Funktionen mehrerer reeller Variablen	213
8.1	Differenzierbare Abbildungen, das Differential, die Richtungsableitungen und die partiellen Ableitungen	214
8.2	Kettenregeln und Mittelwertsätze	223
8.3	Die Jacobi-Matrix, der Gradient und die Hesse-Matrix	228
8.4	Die Taylorformel für Funktionen mehrerer reeller Variablen	233
8.5	Lokale Extrema für Funktionen mehrerer reeller Variablen	236
8.6	Koordinatentransformationen und der Satz über den lokalen Diffeomorphismus	239
8.7	Auflösen von Gleichungen. Der Satz über implizite Funktionen	246
8.8	Gleichungsdefinierte Flächen, ihre Tangentialebenen und Normalen	250

9 Integralrechnung für Funktionen mehrerer reeller Variablen (Ein Ausblick)	255
9.1 Das Problem der Volumendefinition für Teilmengen des \mathbb{R}^n	255
9.2 Das Jordan-Volumen	257
9.3 Das Riemann-Integral für Funktionen mehrerer reeller Variablen	259
9.4 Der Satz von Fubini und das Prinzip von Cavalieri	263
9.5 Die Transformationsformel für Riemann-Integrale	270
Sachverzeichnis	275

Grundlagen

1.1 Aussagen, logische Schlüsse und Beweisstrategien

Aussagen sind sprachliche oder schriftliche Gebilde, die im täglichen Leben zur Beschreibung und Mitteilung von Sachverhalten und Meinungen dienen. Die Mathematik befasst sich mit Aussagen, von denen man wissen will, ob sie *wahr* oder *falsch* sind. Insbesondere möchte man wahre Aussagen über vermutete Zusammenhänge machen. Wir präzisieren deshalb zunächst, was wir in der Mathematik unter *Aussagen* verstehen.

Eine mathematische Aussage ist eine Aussage, die in einem gegebenen Kontext immer entweder wahr oder falsch ist.

Je nachdem, ob eine Aussage wahr oder falsch ist, ordnen wir ihr einen *Wahrheitswert* $W(=wahr)$ oder $F(=falsch)$ zu.

Beispiele:

- "*Jedes Dreieck hat drei Ecken*" ist eine mathematische Aussage. Sie ist wahr.
- Die Aussage "*Die Innenwinkelsumme jedes Dreiecks beträgt 180°* " ist eine mathematische Aussage. Ob sie wahr oder falsch ist, hängt allerdings vom betrachteten Kontext ab. In der Euklidischen Geometrie ist sie wahr. Sie werden in der Vorlesung über Elementargeometrie die Lobatschewski-Geometrie kennenlernen, in der diese Aussage falsch ist (Hier ist die Aussage "*Die Innenwinkelsumme jedes Dreiecks ist kleiner als 180°* " wahr).
- Die Aussage "*Die Gleichung $x^2 + 2x + 2 = 0$ hat eine Lösung*" ist eine mathematische Aussage. Sucht man Lösungen x im Bereich der reellen Zahlen, so ist diese Aussage falsch. Im Bereich der komplexen Zahlen, die wir im Kapitel 2 dieser Vorlesung kennenlernen werden, ist diese Aussage wahr.
- Der Aussage "*In Berlin lebt man gern*" kann man weder den Wahrheitswert "wahr" noch den Wahrheitswert "falsch" zuordnen. Ob man diese Aussage bejaht oder nicht, hängt von der Auffassung des Betrachters ab. Sie hat deshalb nicht den Charakter einer mathematischen Aussage.

- Die Aussage "*Diese Aussage ist falsch*" ist keine mathematische Aussage. Wäre sie wahr, so wäre sie gleichzeitig falsch. Wäre sie falsch, so wäre sie gleichzeitig wahr.

Statt: *Die Aussage A wahr*, so sagt man auch: *A gilt* oder *A ist richtig*.

Im Folgenden werden wir unter *Aussagen* stets *mathematische Aussagen* verstehen. Die Formulierung von Aussagen kann sehr kompliziert sein, sie können durch Verknüpfung von mehreren Aussagen durch Worte wie "und", "oder", "nicht", "weder - noch", "sowohl als auch", "wenn - so", ... entstehen. Man kann sich fragen, wie der Wahrheitswert einer Aussage von diesen Teilaussagen abhängt. In der Umgangssprache wird das nicht immer einheitlich gehandhabt und führt dadurch mitunter zu Mißverständnissen. In der Mathematik fixiert man den Wahrheitswert der aus Teilaussagen verknüpften Aussage ein für alle mal. Wir stellen dies in den folgenden Tabellen zusammen, die man auch "Wahrheitstafeln" nennt.

1. Negation(Verneinung) einer Aussage A ("nicht A ", symbolisch: $\neg A$)

Beispiel: Die Aussage A lautet: *Alle Franzosen trinken gern Wein*.

Die Negation $\neg A$ lautet dann: *Nicht alle Franzosen trinken gern Wein* (oder auch: *Es gibt einen Franzosen, der nicht gern Wein trinkt*).

Wahrheitstafel:

A	$\neg A$
W	F
F	W

d.h., wenn A wahr ist, ist $\neg A$ falsch;
wenn A falsch ist, ist $\neg A$ wahr.

2. Konjunktion zweier Aussagen A, B : (" A und B ", symbolisch: $A \wedge B$).

Wahrheitstafel:

A	B	$A \wedge B$
W	W	W
W	F	F
F	W	F
F	F	F

d.h., $A \wedge B$ ist genau dann wahr, wenn sowohl A als auch B wahr sind.

3. Disjunktion zweier Aussagen A, B : (" A oder B ", symbolisch: $A \vee B$).

Wahrheitstafel:

A	B	$A \vee B$
W	W	W
W	F	W
F	W	W
F	F	F

d.h., $A \vee B$ ist genau dann wahr, wenn mindestens eine der Aussagen A oder B wahr ist.

Im Gegensatz zu mancher Verwendung des Wortes *oder* in der Umgangssprache als *ausschließendes oder* bedeutet das Wort *oder* in der Mathematik immer ein *nichtausschließendes oder*. A oder B ist wahr bedeutet also immer, dass eine der beiden Aussagen wahr ist, möglicherweise aber auch beide. Für ein *ausschließendes oder* benutzt man in der Mathematik immer die Worte *entweder ... oder*.

4. Implikation : ("Aus A folgt B ", symbolisch: $A \Rightarrow B$).

Wahrheitstafel:

A	B	$A \Rightarrow B$
W	W	W
W	F	F
F	W	W
F	F	W

d.h., wenn die Aussage A falsch ist, ist die Aussage $A \Rightarrow B$ immer wahr, unabhängig davon, ob B wahr oder falsch ist. Z.B. ist die Implikation *Aus n teilt 2 folgt n teilt 6* für alle natürlichen Zahlen n wahr.

Die Aussage $A \Rightarrow B$ ist genau dann wahr, wenn die Wahrheit von A stets die Wahrheit von B nach sich zieht. Für eine wahre Implikation $A \Rightarrow B$ sagt man deshalb auch:

wenn A gilt, so gilt auch B ;

B gilt, falls A gilt;

*(die Gültigkeit von) A ist **hinreichend** für (die Gültigkeit von) B ;*

*(die Gültigkeit von) B ist **notwendig** für (die Gültigkeit von) A .*

Beispiel: Es sei Q ein Viereck in der Euklidischen Ebene. Dann ist die Implikation:

$$Q \text{ ist ein Rechteck} \Rightarrow Q \text{ ist ein Trapez}$$

wahr. Die Eigenschaft, "Rechteck" ist hinreichend für die Eigenschaft "Trapez", die Eigenschaft "Trapez" ist notwendig für die Eigenschaft "Rechteck".

5. Äquivalenz zweier Aussagen

Definition 1.1. Zwei Aussagen A und B nennt man *äquivalent* (symbolisch: $A \Leftrightarrow B$), wenn sie die gleichen Wahrheitswerte haben, d.h., entweder beide Aussagen sind wahr oder beide Aussagen sind falsch.

Beispiel: Es sei Δ ein Dreieck in der Euklidischen Ebene. Die Aussagen

A : Das Dreieck Δ ist gleichseitig.

B : Das Dreieck Δ ist gleichwinklig.

sind äquivalent.

Sind A und B zwei Aussagen, so ist $A \Leftrightarrow B$ ebenfalls eine Aussage. Ihre aus der Definition der Äquivalenz folgenden Wahrheitswerte in Abhängigkeit von denen von A und B zeigt die folgende Wahrheitstafel:

A	B	$A \Leftrightarrow B$
W	W	W
F	F	W
W	F	F
F	W	F

Weitere in der Mathematik benutzte Symbole sind die Abkürzungen für die **Quantoren**:

- \exists es gibt ein¹
- $\exists!$ es gibt genau ein
- \nexists es gibt kein
- \forall für alle, für jedes

Die Wahrheitstabellen für die logischen Verknüpfungen $\neg, \wedge, \vee, \Rightarrow$ von Aussagen bilden die Grundannahmen (**Axiome**) für das logische Schließen in der Mathematik. Mit Hilfe der Wahrheitstabellen können wir weitere **logische Gesetze (Tautologien) beweisen**. Dabei versteht man unter einem logischen Gesetz eine Aussage, die bei *jeder* Belegung der Ausgangsaussagen mit Wahrheitswerten den Wahrheitswert W annimmt.

Satz 1.1 A, B und C bezeichnen Aussagen. Die folgenden Aussagen sind logische Gesetze:

1. *Doppelte Negation:* $(\neg(\neg A)) \Leftrightarrow A$.
2. *Transitivität:* $((A \Rightarrow B) \wedge (B \Rightarrow C)) \Rightarrow (A \Rightarrow C)$.
3. $(A \Rightarrow B) \Leftrightarrow ((\neg A) \vee B)$.
4. $(A \Leftrightarrow B) \Leftrightarrow ((A \Rightarrow B) \wedge (B \Rightarrow A))$,
 $(A \Leftrightarrow B) \Leftrightarrow (B \Leftrightarrow A)$.
5. *Regeln von de Morgan:*
 $(\neg(A \wedge B)) \Leftrightarrow ((\neg A) \vee (\neg B))$,
 $(\neg(A \vee B)) \Leftrightarrow ((\neg A) \wedge (\neg B))$.
6. *Kontrapositionsgesetz:*
 $(A \Rightarrow B) \Leftrightarrow ((\neg B) \Rightarrow (\neg A))$.
7. *Assoziativgesetze:*
 $(A \vee (B \vee C)) \Leftrightarrow ((A \vee B) \vee C)$,
 $(A \wedge (B \wedge C)) \Leftrightarrow ((A \wedge B) \wedge C)$.
8. *Distributivgesetze:*
 $(A \vee (B \wedge C)) \Leftrightarrow ((A \vee B) \wedge (A \vee C))$,
 $(A \wedge (B \vee C)) \Leftrightarrow ((A \wedge B) \vee (A \wedge C))$.
9. *Kommutativgesetze:*
 $(A \wedge B) \Leftrightarrow (B \wedge A)$,
 $(A \vee B) \Leftrightarrow (B \vee A)$.

¹ Die unbestimmten Artikel *ein, eine* werden in der Mathematik stets im Sinne von *mindestens ein* gebraucht.

Beweis. Um zu zeigen, dass die Aussagen 1. - 9. logische Gesetze sind, muß man nachweisen, dass die jeweilige Aussage für jede Belegung der Ausgangsaussagen mit Wahrheitswerten wahr ist. Dies verläuft ganz formal durch Benutzung der Wahrheitstabellen für die Verknüpfungsoperationen $\neg, \wedge, \vee, \Rightarrow, \Leftrightarrow$. Wir führen dies hier exemplarisch für 1., 3. und 6. aus und überlassen die anderen Fälle als Übungsaufgaben.

Zu 1.:

A	$\neg A$	$\neg(\neg A)$	$(\neg(\neg A)) \Leftrightarrow A$
W	F	W	W
F	W	F	W

Zu 3.:

A	B	$\neg A$	$(\neg A) \vee B$	$A \Rightarrow B$	$(A \Rightarrow B) \Leftrightarrow ((\neg A) \vee B)$
W	W	F	W	W	W
W	F	F	F	F	W
F	W	W	W	W	W
F	F	W	W	W	W

Zu 6.:

A	B	$A \Rightarrow B$	$(\neg B) \Rightarrow (\neg A)$	$(A \Rightarrow B) \Leftrightarrow ((\neg B) \Rightarrow (\neg A))$
W	W	W	W	W
W	F	F	F	W
F	W	W	W	W
F	F	W	W	W

□

Die Wahrheitstabellen und die logischen Gesetze bilden die Grundlage für den *Beweis der Wahrheit von Aussagen*. Wir merken uns dazu insbesondere die folgenden Spezialfälle:

Satz 1.2 A, B und C bezeichnen Aussagen. Dann gilt:

1. Sind A und $A \Rightarrow B$ wahr, so ist B wahr.
2. Sind A und $(\neg B) \Rightarrow (\neg A)$ wahr, so ist B wahr.
3. Ist $(\neg B) \Rightarrow C$ wahr und C falsch, so ist B wahr.

Beweis. Zu 1: Nach der Wahrheitstafel für die Aussage $A \Rightarrow B$ kann B für ein wahres A und ein wahres $A \Rightarrow B$ nur wahr sein.

Zu 2: Nach dem Kontrapositionsgesetz ist die Aussage $(\neg B) \Rightarrow (\neg A)$ genau dann wahr, wenn die Aussage $A \Rightarrow B$ wahr ist. Deshalb erhält man 2. aus 1..

Zu 3.: Dies erhält man aus 2., indem man für A die Aussage $\neg C$ einsetzt und das Gesetz der doppelten Negation benutzt. □

Dies liefert verschiedene Strategien, die Wahrheit (Richtigkeit, Gültigkeit) einer Aussage B zu beweisen:

Direkter Beweis:

Beim *direkten Beweis* geht man von einer wahren Aussage A aus und folgert durch bereits als wahr erkannte Implikationen (d.h. durch "korrekte/zulässige Schlüsse") $(A \Rightarrow C_1)$, $(C_1 \Rightarrow C_2), \dots, (C_{k-1} \Rightarrow C_k), (C_k \Rightarrow B)$, dass die Aussage B wahr ist. Es ist allerdings oft schwierig, eine *geeignete* wahre Aussage A für den Start des Beweises zu finden.

Indirekter Beweis/Widerspruchsbeweis:

Beim *indirekten Beweis/Widerspruchsbeweis* nimmt man an, dass die Aussage B falsch ist, d.h. dass die Negation $\neg B$ gilt. Daraus folgert man durch korrekte Schlüsse $(\neg B \Rightarrow D_1)$, $(D_1 \Rightarrow D_2), \dots, (D_{k-1} \Rightarrow D_k), (D_k \Rightarrow E)$ eine Aussage E , von der man bereits weiß, dass sie falsch ist. Man erhält also einen Widerspruch, denn E müßte nach unserer Annahme, dass $\neg B$ wahr ist, ebenfalls wahr sein. Die Annahme war also falsch, d.h. B ist wahr. Die Aussage $\neg E$ könnte man als Start für einen direkten Beweis der Wahrheit von B benutzen.

Beispiel: Aussage B : Für alle reellen Zahlen $a, b \geq 0$ gilt $\frac{a+b}{2} \geq \sqrt{a \cdot b}$.

Direkter Beweis: Es seien a und b reelle Zahlen mit $a, b \geq 0$.

Es gilt $(a - b)^2 \geq 0$ (wahre Aussage A). Daraus folgen durch korrekte Schlüsse, die man vom Rechnen mit reellen Zahlen kennt:

$$\begin{aligned} & (a - b)^2 \geq 0 && (A) \\ \Rightarrow & a^2 - 2ab + b^2 \geq 0 && (C_1) \\ \Rightarrow & a^2 + 2ab + b^2 \geq 4ab && (C_2) \\ \Rightarrow & (a + b)^2 \geq 4ab && (C_3) \\ \Rightarrow & (a + b) \geq 2\sqrt{ab} && (C_4) \text{ (korrekter Schluß, da } a, b \geq 0) \\ \Rightarrow & \frac{a+b}{2} \geq \sqrt{a \cdot b} && (B). \end{aligned}$$

Also gilt B . □

Indirekter Beweis: Wir nehmen an, dass die Aussage B falsch ist, d.h., dass es zwei reelle Zahlen $a, b \geq 0$ gibt mit $\frac{a+b}{2} < \sqrt{a \cdot b}$. (d.h. $\neg B$ ist wahr). Daraus folgen durch korrekte Schlüsse für diese Zahlen a, b :

$$\begin{aligned} & \frac{a+b}{2} < \sqrt{a \cdot b} && (\neg B) \\ \Rightarrow & \frac{(a+b)^2}{4} < a \cdot b && (D_1) \text{ (korrekter Schluß, da } a, b \geq 0) \\ \Rightarrow & (a + b)^2 < 4ab && (D_2) \\ \Rightarrow & a^2 + 2ab + b^2 < 4ab && (D_3) \\ \Rightarrow & a^2 - 2ab + b^2 < 0 && (D_4) \\ \Rightarrow & (a - b)^2 < 0 && (E). \end{aligned}$$

Die Aussage E ist falsch (Widerspruch), also gilt B . □

Der Ringschluß:

Ein weiteres oft benutztes Prinzip ist der **Ringschluß** beim Beweis von Äquivalenzen. Will man zeigen, dass die Aussagen $A_1, A_2, A_3, \dots, A_{k-1}, A_k$ zueinander äquivalent sind, so genügt es nach Satz 1.1, Punkt 2. und 4., zu zeigen, dass die folgenden Implikationen gelten:

$$(A_1 \Rightarrow A_2), (A_2 \Rightarrow A_3), \dots, (A_{k-1} \Rightarrow A_k), (A_k \Rightarrow A_1).$$

Angabe von Gegenbeispielen zum Widerlegen der Wahrheit einer Aussage:

Will man untersuchen, ob eine Aussage B wahr oder falsch ist, so kann man sich zunächst *Spezialfälle (Beispiele)* dieser Aussage ansehen und diese untersuchen. Findet man dabei *ein Beispiel*, für das die Aussage B nicht gilt, so ist B falsch. Z.B:

Aussage B : *Alle Hörer der Vorlesung Analysis 1 sind männlich.*

Diese Aussage ist falsch, denn z.B. Frau ... nimmt an der Vorlesung Analysis 1 teil.

1.2 Mengen und Abbildungen**1.2.1 Mengen**

In dieser Vorlesung benutzen wir den "naiven" Mengenbegriff, den Georg Cantor (1845-1918) geprägt hat².

Definition 1.2. *Unter einer Menge verstehen wir eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder des Denkens, welche Elemente der Menge heißen, zu einem Ganzen.*

Ist M eine Menge und x ein Element von M , so schreibt man $x \in M$. Ist x kein Element von M , so schreibt man $x \notin M$. Wir lassen auch Mengen zu, die kein Element enthalten. Eine solche Menge heißt *leere Menge* und wird mit \emptyset bezeichnet. Mengen kann man durch Aufzählen ihrer Elemente beschreiben, die man innerhalb von geschweiften Klammern auflistet (ggf. mit \dots , wenn der Sinn im Kontext klar ist), z.B.:

$$\begin{aligned} A &= \{1, 2, 3, \dots, 10\} && \text{Menge der natürlichen Zahlen von 1 bis 10,} \\ \mathbb{Z} &= \{0, \pm 1, \pm 2, \pm 3, \dots\} && \text{Menge der ganzen Zahlen.} \end{aligned}$$

Besteht eine Menge M aus allen Elementen x , die eine gewisse Eigenschaft $E(x)$ erfüllen, so schreiben wir dies in der Form

$$M = \{x \mid x \text{ erfüllt } E(x)\}.$$

² Mit dem axiomatischen Aufbau der Mengenlehre sind erhebliche Probleme verbunden, die wir in dieser Vorlesung nicht thematisieren wollen.

Definition 1.3. Seien A und B Mengen.

1. A heißt Teilmenge von B (symbolisch: $A \subset B$), wenn jedes Element von A auch Element von B ist (symbolisch: $x \in A \Rightarrow x \in B$).
2. Die Mengen A und B heißen gleich (symbolisch $A = B$), wenn $A \subset B$ und $B \subset A$.
3. A heißt echte Teilmenge von B , wenn $A \subset B$ und $A \neq B$.
4. Die Menge aller Teilmengen von A heißt Potenzmenge von A und wird mit $\mathcal{P}(A)$ bezeichnet:

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

Das Zeichen $:=$ wird in der Mathematik für die Definition von Begriffen benutzt. Links (auf der Seite von $:$) steht der zu definierende Begriff und rechts die ihn definierende Eigenschaft. Statt $A \subset B$ schreibt man auch $B \supset A$.

Aus der Definition folgt für Mengen M, N, P unmittelbar:

$$\begin{aligned} \emptyset &\subset M \quad \text{und} \quad M \subset M, \\ M \subset N \quad \text{und} \quad N \subset P &\Rightarrow M \subset P. \end{aligned}$$

Mengen, die nur aus einem Element x bestehen, bezeichnet man mit $\{x\}$. Für die Potenzmenge von A gilt:

$$\begin{aligned} A &\in \mathcal{P}(A) \quad \text{und} \quad \emptyset \in \mathcal{P}(A), \\ x \in A &\Leftrightarrow \{x\} \in \mathcal{P}(A), \\ M \subset A &\Leftrightarrow M \in \mathcal{P}(A). \end{aligned}$$

Die Potenzmenge von $A = \{1, 2, 3\}$ ist z.B.

$$\mathcal{P}(A) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

Definition 1.4. Seien X und Y Mengen. Dann heißt

$$\begin{aligned} X \cup Y &:= \{x \mid x \in X \text{ oder } x \in Y\} && \text{Vereinigung von } X \text{ und } Y, \\ X \cap Y &:= \{x \mid x \in X \text{ und } x \in Y\} && \text{Durchschnitt von } X \text{ und } Y. \end{aligned}$$

Zwei Mengen X und Y heißen *disjunkt*, wenn $X \cap Y = \emptyset$. Die Vereinigung zweier disjunkter Mengen X und Y nennt man auch *disjunkte Vereinigung*. Wir notieren dies mit dem Symbol: $X \dot{\cup} Y$.

Satz 1.3 (Rechenregeln für Vereinigung und Durchschnitt)

X, Y und Z seien Mengen³. Dann gilt:

1. $X \cap Y \subset Y \subset X \cup Y$.
2. *Kommutativgesetz:*
 $X \cup Y = Y \cup X$ und $X \cap Y = Y \cap X$.
3. *Assoziativgesetze:*
 $X \cup (Y \cup Z) = (X \cup Y) \cup Z$ und $X \cap (Y \cap Z) = (X \cap Y) \cap Z$.

³ Eine solche Formulierung bedeutet *immer*, dass X, Y und Z beliebige Mengen sind.

4. *Distributivgesetze:*

$$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z) \quad \text{und} \quad X \cup (Y \cap Z) = (X \cup Y) \cap (X \cup Z).$$

$$5. X \subset Y \Leftrightarrow X \cup Y = Y \Leftrightarrow X \cap Y = X.$$

Beweis. Die Behauptung 1 folgt direkt aus der Definition. Die Behauptungen 2. und 3. folgen direkt aus Satz 1.1 (Kommutativität und Assoziativität von *und* und *oder*).

Zu 4: Wir zeigen das erste Distributivgesetz durch Äquivalenzschlüsse:

$$\begin{aligned} x \in X \cap (Y \cup Z) &\Leftrightarrow x \in X \text{ und } x \in Y \cup Z && \text{(Def } \cap) \\ &\Leftrightarrow x \in X \text{ und } (x \in Y \text{ oder } x \in Z) && \text{(Def } \cup) \\ &\Leftrightarrow (x \in X \text{ und } x \in Y) \text{ oder } (x \in X \text{ und } x \in Z) && \text{(Distr. } \wedge, \vee) \\ &\Leftrightarrow x \in X \cap Y \text{ oder } x \in X \cap Z && \text{(Def } \cap) \\ &\Leftrightarrow x \in (X \cap Y) \cup (X \cap Z) && \text{(Def } \cup). \end{aligned}$$

Wenn die Äquivalenzschlüsse zu unübersichtlich sind, der kann den Beweis der Gleichheit von zwei Mengen M_1 und M_2 auch immer in 2 getrennten Schritten führen:

1. *Schritt:* Zeige: $x \in M_1 \Rightarrow x \in M_2$.

2. *Schritt:* Zeige: $x \in M_2 \Rightarrow x \in M_1$.

Wir zeigen das 2. Distributivgesetz zur Demonstration in dieser Form:

$$\begin{aligned} x \in X \cup (Y \cap Z) &\Rightarrow x \in X \text{ oder } x \in Y \cap Z \\ &\Rightarrow x \in X \text{ oder } (x \in Y \text{ und } x \in Z) \\ &\Rightarrow (x \in X \text{ oder } x \in Y) \text{ und } (x \in X \text{ oder } x \in Z) \\ &\Rightarrow x \in X \cup Y \text{ und } x \in X \cup Z \\ &\Rightarrow x \in (X \cup Y) \cap (X \cup Z). \end{aligned}$$

$$\begin{aligned} x \in (X \cup Y) \cap (X \cup Z) &\Rightarrow x \in X \cup Y \text{ und } x \in X \cup Z \\ &\Rightarrow (x \in X \text{ oder } x \in Y) \text{ und } (x \in X \text{ oder } x \in Z) \\ &\Rightarrow x \in X \text{ oder } (x \in Y \text{ und } x \in Z) \\ &\Rightarrow x \in X \cup (Y \cap Z). \end{aligned}$$

Zu 5: Wir zeigen diese Äquivalenzen durch einen Ringschluß:

$$X \subset Y \Rightarrow X \cup Y = Y \Rightarrow X \cap Y = X \Rightarrow X \subset Y.$$

Es gelte $X \subset Y$. Dann ist $X \cup Y \subset Y \subset X \cup Y$ und somit $X \cup Y = Y$. Nach dem Distributivgesetz folgt dann für den Durchschnitt mit X :

$$X \cap Y = X \cap (X \cup Y) = (X \cap X) \cup (X \cap Y) = X \cup (X \cap Y) = X.$$

Daraus folgt dann $X \subset X \cap Y \subset Y$. □

Definition 1.5. Seien X und A Mengen. Dann heißt die Differenz

$$X \setminus A := \{x \mid x \in X \text{ und } x \notin A\}$$

das Komplement von A in X (oder die Differenz von X und A).

Satz 1.4 (Rechenregeln für die Mengendifferenz)

Seien A, B und X Mengen. Dann gilt:

1. $X \setminus (X \setminus A) = X \cap A$.

Für $A \subset X$ gilt insbesondere $X \setminus (X \setminus A) = A$.

2. de Morgansche Regeln:

$$X \setminus (A \cup B) = (X \setminus A) \cap (X \setminus B),$$

$$X \setminus (A \cap B) = (X \setminus A) \cup (X \setminus B).$$

3. Distributivgesetze:

$$(A \cap B) \setminus X = (A \setminus X) \cap (B \setminus X),$$

$$(A \cup B) \setminus X = (A \setminus X) \cup (B \setminus X).$$

Beweis. Zu 1:

$$\begin{aligned} x \in X \setminus (X \setminus A) &\Leftrightarrow x \in X \text{ und } x \notin X \setminus A \\ &\Leftrightarrow x \in X \text{ und } x \in A \\ &\Leftrightarrow x \in A \cap X. \end{aligned}$$

Zu 2: Wir zeigen nur die erste Gleichheit, die zweite Gleichung beweist man analog.

$$\begin{aligned} x \in X \setminus (A \cup B) &\Leftrightarrow x \in X \text{ und } x \notin A \cup B \\ &\Leftrightarrow x \in X \text{ und } (x \notin A \text{ und } x \notin B) \\ &\Leftrightarrow (x \in X \text{ und } x \notin A) \text{ und } (x \in X \text{ und } x \notin B) \\ &\Leftrightarrow x \in X \setminus A \text{ und } x \in X \setminus B \\ &\Leftrightarrow x \in (X \setminus A) \cap (X \setminus B). \end{aligned}$$

Zu 3: Übungsaufgabe. □

Definition 1.6. Es seien eine Menge A sowie für jedes $\alpha \in A$ eine Menge X_α gegeben. Dann heißen die Mengen

$$\bigcup_{\alpha \in A} X_\alpha := \{x \mid \exists \alpha \in A \text{ mit } x \in X_\alpha\}$$

und

$$\bigcap_{\alpha \in A} X_\alpha := \{x \mid \forall \alpha \in A \text{ gilt } x \in X_\alpha\}$$

Vereinigung bzw. Durchschnitt der Mengenfamilie $\{X_\alpha\}_{\alpha \in A}$.

Die Vereinigung und der Durchschnitt von Mengenfamilien erfüllen die analogen Rechenregeln wie in den Sätzen 1.3 und 1.4 für zwei Mengen formuliert sind.

Definition 1.7. Seien X und Y Mengen. Das Produkt von X und Y ist die Menge der Tupel (geordneten Paare)

$$X \times Y := \{(x, y) \mid x \in X \text{ und } y \in Y\}.$$

Das Produkt von n Mengen X_1, X_2, \dots, X_n ist die Menge der n -Tupel

$$X_1 \times X_2 \times \dots \times X_n := \{(x_1, x_2, \dots, x_n) \mid x_i \in X_i, i = 1, \dots, n\}.$$

Satz 1.5 (Rechenregeln für das Mengenprodukt)

Seien A, B und M Mengen. Dann gilt:

1. $(A \cup B) \times M = (A \times M) \cup (B \times M)$.
2. $(A \cap B) \times M = (A \times M) \cap (B \times M)$.
3. $(A \setminus B) \times M = (A \times M) \setminus (B \times M)$.

Beweis. Übungsaufgabe. □

Abschließend listen wir die Bezeichnungen auf, die wir in dieser Vorlesung für die Zahlbereiche benutzen:

\mathbb{N}	Menge aller natürlichen Zahlen: 1, 2, 3, ...
\mathbb{N}_0	$\mathbb{N} \cup \{0\}$
\mathbb{Z}	Menge aller ganzen Zahlen
\mathbb{Q}	Menge der rationalen Zahlen
\mathbb{Q}^+	Menge der positiven rationalen Zahlen
\mathbb{R}	Menge der reellen Zahlen
$\mathbb{R} \setminus \mathbb{Q}$	Menge der irrationalen Zahlen
\mathbb{R}^+	Menge der positiven reellen Zahlen

1.2.2 Abbildungen

Definition 1.8. Es seien X und Y nichtleere Mengen.

1. Eine Vorschrift f , die jedem Element von X genau ein Element von Y zuordnet, heißt *Abbildung (oder Funktion⁴) von X in Y* . Symbolisch:

$$f : X \rightarrow Y$$

$$x \mapsto f(x).$$

$f(x) \in Y$ heißt *Bildpunkt* von x unter f . Ist $y \in Y$, so heißt jedes $x \in X$ mit $f(x) = y$ ein *Urbild* von y bei f .

2. X heißt *Definitionsbereich* von f und Y *Wertebereich* von f . Die Menge

$$\text{im}(f) := \{f(x) \mid x \in X\} = \{y \in Y \mid \exists x \in X \text{ mit } f(x) = y\}$$

heißt *Bild* von f .

3. Die Menge

$$\Gamma_f := \{(x, f(x)) \mid x \in X\} \subset X \times Y$$

heißt *der Graph* von f .

⁴ Wir unterscheiden in dieser Vorlesung nicht zwischen Abbildungen und Funktionen. Die Bezeichnung *Funktion* ziehen wir vor, wenn X und Y Zahlbereiche sind. Dies wird in der mathematischen Literatur nicht einheitlich gehandhabt!

4. Zwei Abbildungen $f, g : X \rightarrow Y$ heißen gleich (symbolisch $f = g$), wenn $f(x) = g(x)$ für alle $x \in X$.

Beispiele:

1. Die Identität auf X :

$$\begin{aligned} \text{Id}_X : X &\rightarrow X \\ x &\mapsto x. \end{aligned}$$

2. Sei $y_0 \in Y$ fixiert. Dann heißt

$$\begin{aligned} c_{y_0} : X &\rightarrow Y \\ x &\mapsto y_0 \end{aligned}$$

konstante Abbildung. Für ihren Graphen gilt $\Gamma_{c_{y_0}} = \{(x, y_0) \mid x \in X\} = X \times \{y_0\}$.

3. Ist $X \subset \mathbb{R}$, so heißt jede Abbildung $f : X \rightarrow \mathbb{R}$ *reelle Funktion*. Ein Beispiel ist die aus der Schule bekannte Logarithmusfunktion:

$$\begin{aligned} \ln : \mathbb{R}^+ &\rightarrow \mathbb{R} \\ x &\mapsto \ln x. \end{aligned}$$

4. Sei $f : X \rightarrow Y$ eine Abbildung und $A \subset X$ eine nichtleere Teilmenge.

Dann heißt die Abbildung $f|_A : A \rightarrow Y$, die jedem $a \in A$ den Wert $f(a)$ zuordnet, die *Einschränkung von f auf A* .

Ist $g : A \rightarrow Y$ eine Abbildung mit $g(a) = f(a)$ für alle $a \in A$, d.h. $g = f|_A$, so heißt f eine *Fortsetzung von g auf X* .

5. Ist $A \subset X$ eine nichtleere Teilmenge von X , so heißt die Abbildung

$$\begin{aligned} \iota_A : A &\rightarrow X \\ a &\mapsto a. \end{aligned}$$

Inklusion von A in X .

6. Sei $X_1 \times X_2$ das Produkt zweier nichtleerer Mengen und $i \in \{1, 2\}$. Die Abbildung

$$\begin{aligned} \text{pr}_i : X_1 \times X_2 &\rightarrow X_i \\ (x_1, x_2) &\mapsto x_i \end{aligned}$$

heißt *Projektion auf die i . Komponente von $X_1 \times X_2$* .

Man kann Abbildungen miteinander verknüpfen:

Definition 1.9. Seien X, Y, Z nichtleere Mengen und $f : X \rightarrow Y$ und $g : Y \rightarrow Z$ zwei Abbildungen. Dann heißt die Abbildung:

$$\begin{aligned} g \circ f : X &\rightarrow Z \\ x &\mapsto g(f(x)). \end{aligned}$$

die *Verknüpfung (oder Hintereinanderausführung) von f und g* .

Satz 1.6 Seien $f : X \rightarrow Y$, $g : Y \rightarrow Z$ und $h : Z \rightarrow W$ Abbildungen. Dann sind die Verknüpfungen $h \circ (g \circ f)$ und $(h \circ g) \circ f$ wohldefiniert und es gilt

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

Beweis. Die Betrachtung der Definitions- und Wertebereiche der jeweiligen Abbildungen zeigt die Wohldefiniertheit. Sei $x \in X$. Dann gilt nach Definition:

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

Somit sind beide Abbildungen gleich. □

Auf Grund dieses Satzes können wir beim Verknüpfen mehrerer Abbildungen die Klammern weglassen: $h \circ g \circ f := h \circ (g \circ f) = (h \circ g) \circ f$.

In der Analysis befassen wir uns mit Eigenschaften von Abbildungen. Zunächst betrachten wir die folgenden drei Eigenschaften:

Definition 1.10. Eine Abbildung $f : X \rightarrow Y$ heißt

1. *surjektiv*, falls jeder Punkt von Y (mindestens) ein Urbild bei f besitzt, d.h. falls gilt:

$$\forall y \in Y \exists x \in X \text{ mit } f(x) = y.$$

2. *injektiv*, falls jeder Punkt von Y höchstens ein Urbild besitzt, d.h. falls gilt:

$$x_1, x_2 \in X \wedge f(x_1) = f(x_2) \Rightarrow x_1 = x_2.$$

3. *bijektiv*, falls f surjektiv und injektiv ist, d.h. falls jeder Punkt von Y genau ein Urbild besitzt.

Ist $f : X \rightarrow Y$ bijektiv, so ist die *inverse Abbildung* (oder *Umkehrabbildung*) $f^{-1} : Y \rightarrow X$ zu f definiert durch:

$$\begin{aligned} f^{-1} : Y &\rightarrow X \\ y &\mapsto f^{-1}(y) := \text{das eindeutig bestimmte Urbild von } y \text{ bei } f. \end{aligned}$$

f^{-1} ist dann ebenfalls bijektiv und es gilt auf Grund der Definition:

$$f^{-1} \circ f = Id_X, \quad f \circ f^{-1} = Id_Y \quad \text{und} \quad (f^{-1})^{-1} = f.$$

Beispiele:

1. Die Identität $Id_X : X \rightarrow X$ ist bijektiv. Für ihre Umkehrfunktion gilt $(Id_X)^{-1} = Id_X$.
2. Die konstante Abbildung $c_{y_0} : X \rightarrow Y$ ist weder surjektiv noch injektiv, falls X und Y mehr als ein Element enthalten.

3. Die reelle Funktion $\ln : \mathbb{R}^+ \rightarrow \mathbb{R}$ ist bijektiv. Ihre Umkehrfunktion ist die Exponentialfunktion (Schulwissen).
4. Ist $A \subset X$ eine echte Teilmenge, so ist die Inklusion $\iota_A : A \rightarrow X$ injektiv, aber nicht surjektiv.
5. Die Projektion $\text{pr}_1 : X_1 \times X_2 \rightarrow X_1$ ist surjektiv, aber - falls X_2 mehr als ein Element enthält - nicht injektiv.

Satz 1.7 Sei $f : X \rightarrow Y$ eine Abbildung. Dann gilt:

1. f ist genau dann injektiv, wenn es eine Abbildung $g : Y \rightarrow X$ gibt, so dass $g \circ f = \text{Id}_X$. (Eine solche Abbildung g heißt Linksinverse von f).
2. f ist genau dann surjektiv, wenn es eine Abbildung $h : Y \rightarrow X$ gibt, so dass $f \circ h = \text{Id}_Y$. (Eine solche Abbildung h heißt Rechtsinverse von f).
3. f ist genau dann bijektiv, wenn es eine Abbildung $g : Y \rightarrow X$ gibt, die sowohl Rechts- als auch Linksinverse von f ist, d.h. so dass $g \circ f = \text{Id}_X$ und $f \circ g = \text{Id}_Y$. In diesem Fall ist g eindeutig bestimmt und es gilt $g = f^{-1}$.

Beweis. Zu 1:

(\implies): Sei f injektiv. Wir fixieren ein Element $x_0 \in X$ und definieren $g : Y \rightarrow X$ durch

$$g(y) := \begin{cases} x & \text{falls } y \in \text{im}(f) \text{ und } f(x) = y, \\ x_0 & \text{falls } y \notin \text{im}(f). \end{cases}$$

Dann ist g wohldefiniert und es gilt $g(f(x)) = x$ für alle $x \in X$. g ist also eine Linksinverse zu f .

(\impliedby): Sei $g : Y \rightarrow X$ eine Linksinverse von f . Seien $x_1, x_2 \in X$ zwei Punkte mit $f(x_1) = f(x_2)$. Dann gilt $x_1 = g(f(x_1)) = g(f(x_2)) = x_2$. Folglich ist f injektiv.

Zu 2:

(\implies): Sei f surjektiv. Für jedes $y \in Y$ wählen wir ein Urbild $x(y)$ von y bei f aus⁵. Dann ist die Abbildung $h : Y \rightarrow X$ mit $h(y) := x(y)$ wohldefiniert und es gilt $f(h(y)) = f(x(y)) = y$ für alle $y \in Y$. h ist also eine Rechtsinverse von f .

(\impliedby): Sei $h : Y \rightarrow X$ eine Rechtsinverse von f . Dann gilt $f(h(y)) = y$ für alle $y \in Y$. Also ist $h(y) \in X$ ein Urbild von y . Folglich ist f surjektiv.

Zu 3:

(\implies): Sei f bijektiv. Dann ist die Umkehrabbildung $f^{-1} : Y \rightarrow X$ eine Rechts- und Linksinverse von f (siehe oben).

(\impliedby): Sei $g : Y \rightarrow X$ eine Rechts- und Linksinverse von f . Dann ist f nach 1. injektiv und nach 2. surjektiv, also bijektiv. Die Abbildung g stimmt dann mit der Umkehrabbildung f^{-1} überein. Sie ist somit eindeutig bestimmt. □

⁵ Hierfür benötigt man das Auswahlaxiom der Mengenlehre. Dieses Axiom besagt, dass man für jede Familie nichtleerer Mengen aus jeder dieser Mengen jeweils ein Element auswählen kann. In der Analysis-Vorlesung werden wir dieses Auswahlaxiom immer voraussetzen.

Satz 1.8 *Es seien $f : X \rightarrow Y$ und $g : Y \rightarrow Z$ zwei bijektive Abbildungen. Dann ist auch die Verknüpfung $g \circ f : X \rightarrow Z$ bijektiv und für ihre inverse Abbildung gilt*

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}. \quad (1.1)$$

Beweis. Für die inversen Abbildungen $f^{-1} : Y \rightarrow X$ und $g^{-1} : Z \rightarrow Y$ zu f bzw. g gilt

$$f^{-1} \circ f = Id_X, \quad f \circ f^{-1} = Id_Y, \quad g^{-1} \circ g = Id_Y, \quad g \circ g^{-1} = Id_Z.$$

Daraus folgt wegen der Assoziativität der Verknüpfung \circ

$$\begin{aligned} (f^{-1} \circ g^{-1}) \circ (g \circ f) &= f^{-1} \circ (g^{-1} \circ g) \circ f = f^{-1} \circ Id_Y \circ f = f^{-1} \circ f = Id_X, \\ (g \circ f) \circ (f^{-1} \circ g^{-1}) &= g \circ (f \circ f^{-1}) \circ g^{-1} = g \circ Id_Y \circ g^{-1} = g \circ g^{-1} = Id_Z. \end{aligned}$$

Aus Satz 1.7 (3. Punkt) folgt dann, dass $g \circ f$ bijektiv ist und ihre Umkehrabbildung durch (1.1) gegeben ist. \square

Definition 1.11. *Sei $f : X \rightarrow Y$ eine Abbildung, $A \subset X$ und $B \subset Y$. Dann heißt*

$$\begin{aligned} f(A) &:= \{f(a) \in Y \mid a \in A\} \subset Y && \text{Bild von } A \text{ unter } f, \\ f^{-1}(B) &:= \{x \in X \mid f(x) \in B\} \subset X && \text{Urbild von } B \text{ unter } f. \end{aligned}$$

Ist $\{y\} \subset Y$ eine 1-elementige Menge, so schreiben wir auch kurz: $f^{-1}(y) := f^{-1}(\{y\})$. Offensichtlich gilt $\text{im}(f) = f(X)$.

Satz 1.9 *Sei $f : X \rightarrow Y$ eine Abbildung, $A, A_1, A_2 \subset X$ und $B, B_1, B_2 \subset Y$. Dann gilt:*

1. $f(A_1 \cup A_2) = f(A_1) \cup f(A_2)$.
2. $f(A_1 \cap A_2) \subset f(A_1) \cap f(A_2)$.
3. $f(X \setminus A) \supset f(X) \setminus f(A)$.
4. $f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$.
5. $f^{-1}(B_1 \cap B_2) = f^{-1}(B_1) \cap f^{-1}(B_2)$.
6. $f^{-1}(Y \setminus B) = X \setminus f^{-1}(B)$.

Beweis. Übungsaufgabe. (Beachte, dass in 2. und 3. keine Gleichheit steht!!) \square

1.3 Das Beweisprinzip der vollständigen Induktion

In der Mathematik trifft man viele Aussagen der Form:

Für jede natürliche Zahl $n \in \mathbb{N}$ gilt die Eigenschaft $E(n)$.

Um die Wahrheit solcher Aussagentypen zu beweisen, eignet sich in vielen Fällen das Beweisprinzip der vollständigen Induktion.

Die natürlichen Zahlen werden durch auf Peano zurückgehende Axiome (die *Peano-Axiome*) eingeführt⁶. Aus diesen Axiomen folgt die Induktionseigenschaft für die natürlichen Zahlen, die folgendes besagt:

Sei $M \subset \mathbb{N}_0$ eine Teilmenge der (um Null ergänzten) natürlichen Zahlen, die die folgenden beiden Eigenschaften erfüllt:

- (1) $n_0 \in M$,
- (2) $\forall n \geq n_0$ gilt: $n \in M \Rightarrow (n + 1) \in M$.

Dann gilt für diese Menge: $\{n \in \mathbb{N}_0 \mid n \geq n_0\} \subset M$.

Als Umformulierung dieser Induktionseigenschaft erhalten wir das Beweisprinzip der vollständigen Induktion.

Beweisprinzip der vollständigen Induktion

Sei $n_0 \in \mathbb{N}_0$ eine fixierte natürliche Zahl oder 0.

Für jede Zahl $n \in \mathbb{N}_0$ mit $n \geq n_0$ sei eine Aussage $A(n)$ gegeben.

Wir setzen voraus, dass die folgenden beiden Bedingungen erfüllt sind:

- (1) $A(n_0)$ ist richtig (Induktionsanfang).
- (2) $\forall n \geq n_0$ gilt: Ist $A(n)$ richtig, so ist auch $A(n + 1)$ richtig.
(Induktionsschritt).

Dann ist die Aussage $A(n)$ für alle Zahlen $n \in \mathbb{N}_0$ mit $n \geq n_0$ richtig.

Um einzusehen, dass das Prinzip der vollständigen Induktion aus der Induktionseigenschaft der natürlichen Zahlen folgt, setzen wir

$$M := \{n \in \mathbb{N}_0 \mid \text{Aussage } A(n) \text{ ist richtig}\}.$$

Dann gilt:

- $n_0 \in M$ (nach Induktionsanfang).
- $\forall n \geq n_0$ gilt: $n \in M \Rightarrow (n + 1) \in M$ (nach Induktionsschritt).

Aus der Induktionseigenschaft der natürlichen Zahlen folgt nun, dass $A(n)$ für alle $n \geq n_0$ richtig ist.

⁶ Zum Aufbau der Zahlbereiche gibt es im Lehramtsstudium eine extra Vorlesung. Wir zitieren deshalb hier nur die benötigten Eigenschaften. Literatur: J. Kramer, A-M. Pippich: *Von den natürlichen Zahlen zu den Quaternionen*, Kapitel 1, Springer-Spektrum 2013.

Eine typisches Anwendungsfeld für das Beweisprinzip der vollständigen Induktion sind Summenformeln. Wir demonstrieren dies an einem Beispiel:

Satz 1.10 Für jede natürliche Zahl $n \in \mathbb{N}$ gilt

$$\sum_{j=1}^n j := 1 + 2 + \dots + n = \frac{n(n+1)}{2}. \quad (\text{Aussage } A(n))$$

Beweis. Wir führen den Beweis durch vollständige Induktion.

Induktionsanfang: Die Aussage $A(1)$ ist richtig, denn $1 = \frac{1(1+1)}{2}$.

Induktionsschritt: Wir setzen voraus, dass die Aussage $A(n)$ für eine (beliebige, aber feste) natürliche Zahl n richtig ist (Induktionsvoraussetzung) und behaupten, dass dann auch die Aussage $A(n+1)$ richtig ist (Induktionsbehauptung).

Induktionsbeweis:

$$\begin{aligned} \sum_{j=1}^{n+1} j &= \left(\sum_{j=1}^n j \right) + (n+1) \stackrel{A(n)}{=} \frac{n(n+1)}{2} + (n+1) \\ &= (n+1) \left(\frac{n}{2} + 1 \right) = \frac{(n+1)(n+2)}{2}. \end{aligned}$$

□

Als weitere Anwendungen beweisen wir einige Eigenschaften der Fakultät einer natürlichen Zahl und der Binomialkoeffizienten.

Definition 1.12. Sei $n \in \mathbb{N}$. Die Zahl

$$n! := \prod_{j=1}^n j = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n$$

heißt n -Fakultät. Des Weiteren setzen wir $0! := 1$.

Satz 1.11 Die Anzahl a_n aller Anordnungen von n verschiedenen Objekten ist $n!$.

Beweis. Wir führen den Beweis durch vollständige Induktion.

Induktionsanfang: Für $n = 1$ gilt $a_1 = 1$ und $1! = 1$, also ist $a_1 = 1!$.

Induktionsschritt:

Induktionsvoraussetzung: Für eine natürliche Zahl $n \in \mathbb{N}$ gelte $a_n = n!$.

Induktionsbehauptung: Es gilt $a_{n+1} = (n+1)!$.

Induktionsbeweis: Wir betrachten $(n+1)$ Objekte O_1, \dots, O_{n+1} . Die möglichen Anordnungen dieser Objekte kann man in $(n+1)$ Klassen K_j mit $j \in \{1, \dots, n+1\}$ unterteilen: K_j sei die Menge derjenigen Anordnungen, in denen O_j als erstes Element steht, das heißt

$$K_j := \{(O_j, O_{i_1}, \dots, O_{i_n}) \mid \{i_1, i_2, \dots, i_n\} = \{1, \dots, n+1\} \setminus \{j\}\}.$$

Z_j sei die Anzahl der Elemente in K_j . Folglich ist Z_j gleich der Anzahl der Anordnungen der n Objekte $O_1, \dots, O_{j-1}, O_{j+1}, \dots, O_{n+1}$. Nach Induktionsvoraussetzung ist aber die Anzahl der Anordnungen von n Objekten gleich $a_n = n!$. Also gilt

$$a_{n+1} = \sum_{j=1}^{n+1} Z_j = \sum_{j=1}^{n+1} a_n = \sum_{j=1}^{n+1} n! = (n+1) \cdot n! = (n+1)! .$$

□

Definition 1.13. Sei $x \in \mathbb{R}$ und $k \in \mathbb{N}$. Die Zahl

$$\binom{x}{k} := \frac{x \cdot (x-1) \cdot (x-2) \cdot \dots \cdot (x-(k-1))}{k!}$$

heißt Binomialkoeffizient. (Sprich: “ x über k ”). Für $k=0$ setzt man $\binom{x}{0} := 1$.

Satz 1.12 Es seien $n \in \mathbb{N}$, $k \in \mathbb{N}_0$ und $x \in \mathbb{R}$. Dann gilt:

- a) $\binom{n}{k} = 0$, falls $k > n$.
 b) $\binom{n}{k} = \frac{n!}{k!(n-k)!} = \binom{n}{n-k}$, falls $0 \leq k \leq n$.
 c) $\binom{x}{k} + \binom{x}{k+1} = \binom{x+1}{k+1}$.

Beweis. Ist $k > n$, so tritt im Zähler von $\binom{n}{k}$ die Zahl 0 als Faktor auf. Folglich ist $\binom{n}{k} = 0$. Für $k=0$ und $k=n$ ist b) offensichtlich erfüllt. Für $0 < k < n$ gilt:

$$\begin{aligned} \binom{n}{k} &= \frac{n \cdot (n-1) \cdot \dots \cdot (n-(k-1))}{k!} = \frac{n \cdot (n-1) \cdot \dots \cdot (n-(k-1)) \cdot (n-k) \cdot \dots \cdot 2 \cdot 1}{k! \cdot (n-k)!} \\ &= \frac{n!}{k! \cdot (n-k)!} = \binom{n}{n-k}. \end{aligned}$$

Für $k=0$ ist c) offensichtlich erfüllt. Für $k > 0$ gilt:

$$\begin{aligned} \binom{x}{k} + \binom{x}{k+1} &= \frac{x \cdot (x-1) \cdot \dots \cdot (x-(k-1))}{k!} + \frac{x \cdot (x-1) \cdot \dots \cdot (x-k)}{(k+1)!} \\ &= \frac{x \cdot (x-1) \cdot \dots \cdot (x-(k-1))}{k!} \cdot \underbrace{\left(1 + \frac{x-k}{k+1}\right)}_{= \frac{k+1+x-k}{k+1} = \frac{x+1}{k+1}} \\ &= \frac{x \cdot (x-1) \cdot \dots \cdot ((x+1)-k)}{k!} \cdot \frac{(x+1)}{k+1} \\ &= \frac{(x+1) \cdot ((x+1)-1) \cdot ((x+1)-2) \cdot \dots \cdot ((x+1)-k)}{(k+1)!} \\ &= \binom{x+1}{k+1}. \end{aligned}$$

□

Satz 1.13 Seien k und n natürliche Zahlen und sei $1 \leq k \leq n$. Es bezeichne c_k^n die Anzahl aller k -elementigen Teilmengen einer n -elementigen Menge. Dann gilt $c_k^n = \binom{n}{k}$. Insbesondere ist $\binom{n}{k} \in \mathbb{N}$.

Beweis. Der Beweis von Satz 1.13 erfolgt durch vollständige Induktion über n .

Induktionsanfang: Es gilt $c_1^1 = 1 = \binom{1}{1}$, denn aus einer einelementigen Menge kann nur ein Element ausgewählt werden.

Induktionsschritt:

Induktionsvoraussetzung: Es gelte $c_k^n = \binom{n}{k}$ für alle $k \in \{1, \dots, n\}$.

Induktionsbehauptung: $c_k^{n+1} = \binom{n+1}{k}$ für alle $k \in \{1, \dots, n+1\}$.

Induktionsbeweis: Bei der Auswahl einer einelementigen Teilmenge aus einer $(n+1)$ -elementigen Menge hat man $(n+1)$ verschiedene Möglichkeiten. Es gilt somit:

$$c_1^{n+1} = n + 1 = \binom{n+1}{1}.$$

Betrachtet man die Anzahl aller $(n+1)$ -elementigen Teilmengen einer $(n+1)$ -elementigen Menge, so gilt offensichtlich

$$c_{n+1}^{n+1} = 1 = \binom{n+1}{n+1}.$$

Es genügt also, die Behauptung für $k \in \{2, \dots, n\}$ zu zeigen. Betrachten wir eine Menge $M = \{E_1, \dots, E_{n+1}\}$ mit $(n+1)$ Elementen. Dann zerfallen die k -elementigen Teilmengen von M in zwei disjunkte Klassen:

K_0 : alle Teilmengen, die E_{n+1} nicht enthalten, und

K_1 : alle Teilmengen, die E_{n+1} enthalten.

Die Anzahl der k -elementigen Teilmengen in Klasse K_0 ist gleich der Anzahl der k -elementigen Teilmengen von $\{E_1, \dots, E_n\}$, also entsprechend der Induktionsvoraussetzung gleich $c_k^n = \binom{n}{k}$. Die Anzahl der k -elementigen Teilmengen in Klasse K_1 ist gleich der Anzahl der $(k-1)$ -elementigen Teilmengen von $\{E_1, \dots, E_n\}$, also nach Induktionsvoraussetzung gleich $c_{k-1}^n = \binom{n}{k-1}$. Folglich gilt nach Satz 1.12

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

□

Reelle und komplexe Zahlen

Im Rahmen des Lehramtstudiums wird der Aufbau der Zahlbereiche ausführlich in der Vorlesung *Algebra/Zahlentheorie und ihre Didaktik* behandelt. Wir gehen in der Vorlesung *Analysis I* davon aus, daß der Aufbau der Zahlbereiche bis zu den reellen Zahlen bekannt ist. In Abschnitt 2.1 werden wir die grundlegenden Eigenschaften der reellen Zahlen zusammenstellen. In Abschnitt 2.2 führen wir die komplexen Zahlen ein.

2.1 Die reellen Zahlen

Im Folgenden setzen wir voraus, dass die reellen Zahlen existieren und dass sie dem Leser bereits bekannt sind. Das Ziel dieses Abschnittes besteht darin, noch einmal die grundlegenden, die reellen Zahlen eindeutig charakterisierenden Eigenschaften (ihre sogenannten “Axiome”) zusammenzustellen und daraus wichtige Rechenregeln abzuleiten. Diese grundlegenden Eigenschaften sind

- die Körperaxiome,
- die Anordnungsaxiome und
- das Vollständigkeitsaxiom.

Wir werden in dieser Vorlesung nicht darauf eingehen, ob überhaupt eine Menge existiert, die die obigen drei Axiome erfüllt, und wie und woraus man sie ggf. konstruieren kann. Wir werden auch nicht untersuchen, ob eine Menge, die die obigen Axiome erfüllt, eindeutig bestimmt ist. Für diese Fragen verweisen wir auf das Buch von Jürg Kramer und Anna-Maria von Pippich: *Von den natürlichen Zahlen zu den Quaternionen*, Springer-Spektrum 2013 sowie auf die Lehramts-Vorlesung *Algebra und Zahlentheorie*.

2.1.1 Die Körpereigenschaften von \mathbb{R}

Man kann reelle Zahlen addieren und multiplizieren:

$$\begin{aligned}(x, y) \in \mathbb{R} \times \mathbb{R} &\longmapsto x + y \in \mathbb{R} && \text{Addition,} \\(x, y) \in \mathbb{R} \times \mathbb{R} &\longmapsto x \cdot y \in \mathbb{R} && \text{Multiplikation.}\end{aligned}$$

Addition und Multiplikation haben folgende Eigenschaften K1 - K9 (Rechenregeln):

Addition:

$$\text{K1: } x + y = y + x \quad \forall x, y \in \mathbb{R} \quad (\text{Kommutativgesetz der Addition})$$

$$\text{K2: } (x + y) + z = x + (y + z) \quad \forall x, y, z \in \mathbb{R} \quad (\text{Assoziativgesetz der Addition})$$

$$\text{K3: } 0 + x = x \quad \forall x \in \mathbb{R} \quad (\text{Existenz eines neutralen Elementes})$$

$$\text{K4: Zu jedem } x \in \mathbb{R} \text{ gibt es ein } y \in \mathbb{R} \text{ mit } x + y = 0. \text{ } y \text{ heißt das Negative von } x \text{ und wird mit } y =: -x \text{ bezeichnet.} \quad (\text{Existenz des negativen Elementes})$$

Multiplikation:

$$\text{K5: } x \cdot y = y \cdot x \quad \forall x, y \in \mathbb{R} \quad (\text{Kommutativgesetz der Multiplikation})$$

$$\text{K6: } (x \cdot y) \cdot z = x \cdot (y \cdot z) \quad \forall x, y, z \in \mathbb{R} \quad (\text{Assoziativgesetz der Multiplikation})$$

$$\text{K7: } 1 \cdot x = x \cdot 1 = x \quad \forall x \in \mathbb{R} \quad (\text{Existenz eines neutralen Elementes})$$

$$\text{K8: Zu jedem } x \in \mathbb{R}, x \neq 0 \text{ existiert ein } z \in \mathbb{R} \text{ mit } x \cdot z = 1. \text{ } z \text{ heißt das inverse Element zu } x \text{ und wird mit } z =: \frac{1}{x} \text{ bezeichnet.} \quad (\text{Existenz des inversen Elementes})$$

$$\text{K9: } (x + y) \cdot z = x \cdot z + y \cdot z \quad \forall x, y, z \in \mathbb{R} \quad (\text{Distributivgesetz}).$$

Aus diesen neun grundlegenden Eigenschaften lassen sich die weiteren Rechenregeln für die reellen Zahlen ableiten. Beweisen Sie zur Übung, dass z.B. die folgenden Eigenschaften **allein aus K1 - K9** folgen:

- Die neutralen Elemente der Addition und der Multiplikation sind eindeutig bestimmt.
- Das Negative und das Inverse von $x \in \mathbb{R}$ sind eindeutig bestimmt.
- $0 \cdot x = 0$ für jedes $x \in \mathbb{R}$.
- Die Gleichung $a + x = b$, $a, b \in \mathbb{R}$, hat genau eine Lösung, nämlich $x = b + (-a) =: b - a$. ($b - a$ heißt *Differenz von b und a*).
- Die Gleichung $a \cdot x = b$, $a, b \in \mathbb{R}$, $a \neq 0$, hat genau eine Lösung, nämlich $x = b \cdot \frac{1}{a} =: \frac{b}{a}$. ($\frac{b}{a}$ heißt der *Quotient von b und a*).
- Für reelle Zahlen a, b, c, d mit $b \neq 0$ und $d \neq 0$ gilt:

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{a \cdot c}{b \cdot d} \quad \text{und} \quad \frac{a}{b} + \frac{c}{d} = \frac{a \cdot d + c \cdot b}{b \cdot d}.$$

Definition 2.1. Eine Menge \mathbb{K} mit mindestens zwei Elementen, auf der zwei Operationen $+$ und \cdot

$$\begin{aligned} + : \mathbb{K} \times \mathbb{K} &\longrightarrow \mathbb{K} & \cdot : \mathbb{K} \times \mathbb{K} &\longrightarrow \mathbb{K} \\ (x, y) &\longmapsto x + y & (x, y) &\longmapsto x \cdot y \end{aligned}$$

mit den Eigenschaften K1 bis K9 gegeben sind, heißt *Körper*¹.

¹ Wobei in K1 - K9 natürlich \mathbb{R} durch \mathbb{K} zu ersetzen ist, und in K3 und K7 die Existenz eines solchen neutralen Elementes 0 für die Addition bzw. 1 für die Multiplikation gefordert wird.

Wir schreiben den Körper \mathbb{K} mit seinen beiden Operationen $+$ und \cdot oft in der Form $[\mathbb{K}, +, \cdot]$. Der Begriff des Körpers ist ein zentraler algebraischer Begriff und wird in der Vorlesung *Lineare Algebra* ausführlich behandelt.

Körperaxiom der reellen Zahlen

$$\boxed{[\mathbb{R}, +, \cdot] \text{ ist ein Körper.}}$$

$[\mathbb{Q}, +, \cdot]$ ist ebenfalls ein Körper, während $[\mathbb{Z}, +, \cdot]$ kein Körper ist (zum Beispiel besitzt 2 kein multiplikativ inverses Element in \mathbb{Z}). Ein Körper mit zwei Elementen ist durch $\mathbb{K} := \{0, 1\}$ und die Operationen $0 + 0 := 0$, $0 + 1 = 1 + 0 := 1$, $1 + 1 := 0$, $0 \cdot 0 := 0$, $0 \cdot 1 = 1 \cdot 0 := 0$ und $1 \cdot 1 := 1$ gegeben.

Die Rechenregeln, die man aus den Eigenschaften K1-K9 herleiten kann, gelten in *jedem* Körper. Man braucht sie nur einmal zu beweisen. Dies ist der Vorteil dieses abstrakten Konzeptes.

Bezeichnungen: Für n reelle Zahlen x_1, \dots, x_n werden die Summe und das Produkt folgendermaßen abgekürzt:

$$\left. \begin{array}{l} \sum_{i=1}^n x_i := x_1 + x_2 + \dots + x_n \\ \prod_{i=1}^n x_i := x_1 \cdot x_2 \cdot \dots \cdot x_n \end{array} \right\} \begin{array}{l} \text{Klammern sind wegen K2 und K6} \\ \text{nicht nötig.} \end{array}$$

Für zwei Teilmengen $A, B \subset \mathbb{R}$ sei

$$\begin{aligned} A + B &:= \{a + b \mid a \in A, b \in B\} \subset \mathbb{R} \\ A \cdot B &:= \{a \cdot b \mid a \in A, b \in B\} \subset \mathbb{R} \\ -A &:= \{-a \mid a \in A\} \subset \mathbb{R}. \end{aligned}$$

Für das Produkt zweier reeller Zahlen a, b schreiben wir auch kurz: $ab := a \cdot b$.

2.1.2 Die Anordnungseigenschaften von \mathbb{R}

Außer $[\mathbb{R}, +, \cdot]$ gibt es noch viele andere Körper. Die Körperaxiome K1-K9 reichen also nicht aus, um \mathbb{R} eindeutig zu beschreiben. Auf dem Körper der reellen Zahlen kann man im Gegensatz zu einigen anderen Körpern zusätzlich eine Anordnung einführen.

Anordnungseigenschaften von \mathbb{R}

Der Körper der reellen Zahlen $[\mathbb{R}, +, \cdot]$ enthält eine Teilmenge von "positiven" reellen Zahlen \mathbb{R}^+ mit folgenden Eigenschaften:

A1: Für jede reelle Zahl x gilt entweder $x = 0$ oder $x \in \mathbb{R}^+$ oder $x \in -\mathbb{R}^+$, das heißt \mathbb{R} ist die disjunkte Vereinigung

$$\mathbb{R} = -\mathbb{R}^+ \dot{\cup} \{0\} \dot{\cup} \mathbb{R}^+.$$

A2: Ist $x, y \in \mathbb{R}^+$, so gilt $x + y \in \mathbb{R}^+$ und $x \cdot y \in \mathbb{R}^+$.

Definition 2.2. Ein Körper $[\mathbb{K}, +, \cdot]$, in dem eine Teilmenge “positiver Elemente” $\mathbb{K}^+ \subset \mathbb{K}$ existiert, so dass A1 und A2 gelten, heißt angeordneter Körper².

Anordnungsaxiom der reellen Zahlen

Die reellen Zahlen $[\mathbb{R}, +, \cdot]$ sind ein angeordneter Körper.

Mittels der Eigenschaften A1 und A2 kann man Elemente von \mathbb{R} vergleichen.

Definition 2.3. Man sagt “ x ist kleiner gleich y ” und schreibt $x \leq y$, falls $y - x \in \mathbb{R}^+ \cup \{0\}$.

Aus den Anordnungseigenschaften A1 und A2 erhält man die folgenden Eigenschaften der Relation \leq :

- | | |
|-----------------------------------------------------------------------|----------------------|
| O1 : Für alle $x, y \in \mathbb{R}$ gilt $x \leq y$ oder $y \leq x$. | |
| O2 : Für alle $x \in \mathbb{R}$ gilt $x \leq x$ | <i>Reflexivität</i> |
| O3 : Aus $x \leq y$ und $y \leq x$ folgt $x = y$ | <i>Antisymmetrie</i> |
| O4 : Aus $x \leq y$ und $y \leq z$ folgt $x \leq z$ | <i>Transitivität</i> |

Aus A1 und A2 folgen außerdem folgende Monotonieeigenschaften von \leq :

M1: Aus $x \leq y$ folgt $x + z \leq y + z$ für alle $z \in \mathbb{R}$.

M2: Aus $x \leq y$ folgt $x \cdot z \leq y \cdot z$ für alle $z \in \mathbb{R}^+$.

Bezeichnung:

- Gilt $x \leq y$ und $x \neq y$, so schreibt man auch $x < y$
(sprich: “ x kleiner als y ”).
- $x \geq y : \iff y \leq x$ bzw. $x > y : \iff y < x$.

Insbesondere gilt: Aus $x < y$ folgt $a \cdot x > a \cdot y$ für alle $a \in -\mathbb{R}^+$.

Mittels der Ordnungsrelation können wir Intervalle definieren:

Für $a \leq b$, $a, b \in \mathbb{R}$, sei

$$\begin{aligned} [a, b] &:= \{x \in \mathbb{R} \mid a \leq x \leq b\} && \text{(abgeschlossenes Intervall)} \\ [a, b) &:= \{x \in \mathbb{R} \mid a \leq x < b\} && \text{(halboffenes Intervall)} \\ (a, b] &:= \{x \in \mathbb{R} \mid a < x \leq b\} && \text{(halboffenes Intervall)} \\ (a, b) &:= \{x \in \mathbb{R} \mid a < x < b\} && \text{(offenes Intervall)} \end{aligned}$$

Des Weiteren seien

$$\begin{aligned} [a, \infty) &:= \{x \in \mathbb{R} \mid a \leq x\} \\ (a, \infty) &:= \{x \in \mathbb{R} \mid a < x\} \\ (-\infty, a) &:= \{x \in \mathbb{R} \mid x < a\} \\ (-\infty, a] &:= \{x \in \mathbb{R} \mid x \leq a\} \\ (-\infty, \infty) &:= \mathbb{R}. \end{aligned}$$

² Wobei hier in A1 und A2 natürlich \mathbb{R} durch \mathbb{K} und \mathbb{R}^+ durch \mathbb{K}^+ zu ersetzen ist.

Sei I eines der Intervalle $[a, b]$, (a, b) , $[a, b)$, oder $(a, b]$. Dann heißt die Zahl $L(I) := b - a$ *Länge des Intervalls I* .

Definition 2.4. *Unter dem Betrag einer reellen Zahl $x \in \mathbb{R}$ versteht man die Zahl*

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

Ist I ein Intervall der Länge L , so gilt für $x, y \in I$: $|x - y| \leq L$.

Satz 2.1 *Für den Betrag einer reellen Zahl gelten folgende Eigenschaften:*

- (1) $|x| \geq 0 \quad \forall x \in \mathbb{R}, \quad |x| = 0 \Leftrightarrow x = 0.$
- (2) $|x \cdot y| = |x| \cdot |y| \quad \forall x, y \in \mathbb{R}.$
- (3) $|x + y| \leq |x| + |y|.$ (Dreiecksungleichung)
- (4) $||x| - |y|| \leq |x + y|.$

Beweis. (1) und (2) folgen unmittelbar aus der Definition des Betrages $|\cdot|$.

Zum Beweis von (3) benutzen wir die Monotonieeigenschaften. Wegen $x \leq |x|$ und $-x \leq |x|$ bzw. $y \leq |y|$ und $-y \leq |y|$ folgt nach Addition dieser Gleichungen $x + y \leq |x| + |y|$ und $-(x + y) \leq |x| + |y|$ und folglich $|x + y| \leq |x| + |y|$.

Zum Beweis von (4) benutzen wir die Dreiecksungleichung und $|x| = |-x|$:

$$\begin{aligned} |x| &= |(x + y) - y| \leq |x + y| + |y|, \quad \text{und daher } |x| - |y| \leq |x + y|, \\ |y| &= |(x + y) - x| \leq |x + y| + |x|, \quad \text{und daher } |y| - |x| \leq |x + y|. \end{aligned}$$

Somit erhalten wir $||x| - |y|| \leq |x + y|$. □

Die bisherigen Eigenschaften (Körpereigenschaften K1-K9, Anordnungs-eigenschaften A1-A2) bestimmen $[\mathbb{R}, +, \cdot]$ noch immer nicht eindeutig. Sie gelten zum Beispiel auch für den Körper der rationalen Zahlen $[\mathbb{Q}, +, \cdot]$. Die reellen Zahlen \mathbb{R} haben aber eine grundsätzlich andere Eigenschaft als die rationalen Zahlen \mathbb{Q} : die Vollständigkeit.

2.1.3 Vollständigkeitseigenschaft der reellen Zahlen

Es gibt viele verschiedene Möglichkeiten die Vollständigkeitseigenschaft der reellen Zahlen zu beschreiben. Alle diese sind äquivalent. Wir benutzen hier die Existenz der Schnitzzahl von Dedekindschen Schnitten.

Definition 2.5. *Ein Dedekindscher Schnitt³ von \mathbb{R} ist eine Zerlegung $\mathbb{R} = A \dot{\cup} B$ der reellen Zahlen in zwei disjunkte, nichtleere Teilmengen A und B mit der Eigenschaft, dass jedes Element $a \in A$ kleiner als jedes Element $b \in B$ ist, das heißt,*

$$a < b \quad \forall a \in A, \quad \forall b \in B.$$

³ Richard Dedekind [1831-1916], deutscher Mathematiker, letzter Schüler von Carl Friedrich Gauss. Er hat 1872 die erste exakte Definition der reellen Zahlen mit Hilfe der jetzt sogenannten "Dedekindschen Schnitte" angegeben. Die Konstruktion der reellen Zahlen mit Hilfe von Dedekindschen Schnitten findet man im Buch von Oliver Deiser: *Analysis 1*.

Bezeichnung für Dedekindsche Schnitte: $(A | B)$

Die Definition eines Dedekindschen Schnittes ist in jedem angeordneten Körper möglich, da eine Relation “ $<$ ” definiert ist; zum Beispiel in $[\mathbb{Q}, +, \cdot]$.

Beispiel: Sei $a \in \mathbb{R}$ eine reelle Zahl.

$$\left. \begin{array}{l} A = (-\infty, a], \quad B = (a, \infty) \\ A = (-\infty, a), \quad B = [a, \infty) \end{array} \right\} (A | B) \text{ sind Dedekindsche Schnitte.}$$

Definition 2.6. Sei $(A | B)$ ein Dedekindscher Schnitt von \mathbb{R} . Eine Zahl $s \in \mathbb{R}$ heißt Schnitzzahl von $(A | B)$, falls $a \leq s \leq b$ für alle $a \in A$ und $b \in B$.

Wegen $\mathbb{R} = A \dot{\cup} B$, ist s entweder das größte Element von A (falls $s \in A$) oder das kleinste Element von B (falls $s \in B$).

Vollständigkeitsaxiom (V) der reellen Zahlen

Jeder Dedekindsche Schnitt $(A | B)$ von \mathbb{R} besitzt eine Schnitzzahl

Die Vollständigkeitseigenschaft gilt nicht in jedem angeordneten Körper, zum Beispiel nicht im Körper $[\mathbb{Q}, +, \cdot]$: Seien nämlich $A = (-\infty, \sqrt{2}] \cap \mathbb{Q}$ und $B = (\sqrt{2}, +\infty) \cap \mathbb{Q}$. Dann gilt $\mathbb{Q} = A \dot{\cup} B$. Somit bilden A und B einen Dedekindschen Schnitt von \mathbb{Q} . Dieser hat aber in \mathbb{Q} keine Schnitzzahl. (Wir werden noch sehen, dass die Zahl $\sqrt{2}$ nicht rational ist).

Definition 2.7. Ein angeordneter Körper $[\mathbb{K}, +, \cdot]$ mit der Eigenschaft (V), das heißt in dem jeder Dedekindsche Schnitt eine Schnitzzahl hat, heißt vollständig.

Zusammenfassung:

Die reellen Zahlen $[\mathbb{R}, +, \cdot]$ bilden einen vollständigen, angeordneten Körper.

Zwei vollständige, angeordnete Körper $[\mathbb{K}_1, +, \cdot]$ und $[\mathbb{K}_2, +, \cdot]$ sind isomorph⁴ (dies beweisen wir hier nicht). Somit sind die reellen Zahlen $[\mathbb{R}, +, \cdot]$ (bis auf Isomorphie) der einzige vollständige, angeordnete Körper. Die reellen Zahlen \mathbb{R} sind somit durch die Körpereigenschaften K1-K9, die Anordnungseigenschaften A1, A2 und die Vollständigkeitseigenschaft V (bis auf Isomorphie) eindeutig bestimmt.

Wir beweisen nun einige Eigenschaften der reellen Zahlen, die aus der Vollständigkeitseigenschaft (V) folgen.

⁴ Diesen Begriff werden Sie in der Vorlesung *Lineare Algebra* kennenlernen. Er bedeutet anschaulich, dass man solche Körper nicht unterscheiden kann, sie besitzen exakt die gleiche algebraische Struktur.

Definition 2.8.

1. Eine Teilmenge $A \subset \mathbb{R}$ heißt von oben beschränkt, falls eine Zahl $M \in \mathbb{R}$ existiert, so dass $a \leq M$ für alle $a \in A$ gilt. Eine solche Zahl M heißt obere Schranke von A .
2. Eine Teilmenge $A \subset \mathbb{R}$ heißt von unten beschränkt, falls eine Zahl $m \in \mathbb{R}$ existiert, so dass $m \leq a$ für alle $a \in A$ gilt. Eine solche Zahl m heißt untere Schranke von A .
3. Eine Teilmenge $A \subset \mathbb{R}$ heißt beschränkt, falls sie sowohl von unten als auch von oben beschränkt ist.

Definition 2.9. Sei $A \subset \mathbb{R}$ eine nichtleere Teilmenge.

1. Eine Zahl $M_0 \in \mathbb{R}$ heißt Supremum von A , falls sie die kleinste obere Schranke von A ist, das heißt falls
 - a) $a \leq M_0 \quad \forall a \in A$,
 - b) für jedes $\varepsilon > 0$ existiert ein $a \in A$, so dass $M_0 - \varepsilon < a$.
2. Eine Zahl $m_0 \in \mathbb{R}$ heißt Infimum von A , falls sie die größte untere Schranke von A ist, das heißt falls
 - a) $m_0 \leq a \quad \forall a \in A$,
 - b) für jedes $\varepsilon > 0$ existiert ein $a \in A$, so dass $a < m_0 + \varepsilon$.

Bezeichnung: Falls das Supremum bzw. das Infimum einer Menge $A \subset \mathbb{R}$ existiert, so bezeichnen wir es mit

$$\sup A := \text{Supremum von } A, \quad \inf A := \text{Infimum von } A.$$

Offensichtlich existiert höchstens ein Supremum und höchstens ein Infimum einer Menge $A \subset \mathbb{R}$. Aus der Vollständigkeitseigenschaft von \mathbb{R} erhält man die folgende Aussage über die Existenz von Supremum bzw. Infimum.

Satz 2.2 Jede nach oben beschränkte, nichtleere Menge $A \subset \mathbb{R}$ besitzt ein Supremum. Jede nach unten beschränkte, nichtleere Menge $A \subset \mathbb{R}$ besitzt ein Infimum.

Beweis. (1) Sei $A \subset \mathbb{R}$ von oben beschränkt. Wir betrachten die Menge

$$X := \{M \in \mathbb{R} \mid a \leq M \quad \forall a \in A\}.$$

Da A von oben beschränkt ist, ist $X \neq \emptyset$. Es sei $Y := \mathbb{R} \setminus X$. Dann gilt:

- a) $\mathbb{R} = Y \dot{\cup} X$.
- b) Sei $y \in Y$ und $x \in X$. Da $y \notin X$, existiert ein $a \in A$ mit $y < a$. Andererseits ist $a \leq x$ nach Definition von X . Folglich gilt $y < x$ für alle $y \in Y$ und $x \in X$.

Also ist $(Y \mid X)$ ein Dedekindscher Schnitt von \mathbb{R} . Nach dem Vollständigkeitsaxiom (V) von \mathbb{R} existiert eine Schnitzzahl M_0 dieses Dedekindschen Schnittes, also eine Zahl $M_0 \in \mathbb{R}$ mit

$$y \leq M_0 \leq x \quad \forall y \in Y, x \in X.$$

Wir zeigen, dass die Schnitzzahl M_0 in X liegt. Wir führen diesen Beweis indirekt. Wir nehmen an, dass $M_0 \notin X$ und führen dies zum Widerspruch. Ist $M_0 \notin X$, so ist M_0 das größte Element von Y . Nach Definition von Y gibt es ein $a_0 \in A$ mit $M_0 < a_0$. Dann ist wegen der Monotonieeigenschaft von $<$ aber auch

$$M_0 < \frac{M_0 + a_0}{2} < \frac{a_0}{2} + \frac{a_0}{2} = a_0$$

und folglich

$$\frac{M_0 + a_0}{2} \in Y.$$

Dann kann M_0 aber nicht das größte Element von Y sein, d.h. wir erhalten wir einen Widerspruch. Unsere Annahme war demnach falsch. Folglich ist $M_0 \in X$, also eine obere Schranke von A . Als Schnitzzahl von $(Y | X)$ ist M_0 das kleinste Element von X , also die kleinste obere Schranke von A . Das zeigt, dass $M_0 = \sup A$.

Der Beweis der 2. Aussage des Satzes wird analog geführt. \square

Definition 2.10.

1. Sei $A \subset \mathbb{R}$ eine nach oben beschränkte Menge. Liegt das Supremum von A in A , so nennt man es auch das Maximum von A und schreibt dafür $\max A$.
2. Sei $A \subset \mathbb{R}$ eine nach unten beschränkte Menge. Liegt das Infimum von A in A , so nennt man es auch das Minimum von A und schreibt dafür $\min A$.

Wir leiten aus Satz 2.2 einige Folgerungen ab.

Folgerung 2.1 (Archimedisches Axiom der reellen Zahlen)

Die Menge der natürlichen Zahlen $\mathbb{N} \subset \mathbb{R}$ ist nicht nach oben beschränkt, das heißt zu jedem $x \in \mathbb{R}$ existiert ein $n \in \mathbb{N}$ mit $x < n$. Das gleiche gilt auch für jede unendliche Teilmenge $N \subset \mathbb{N}$.

Beweis. Wir führen den Beweis indirekt. Angenommen \mathbb{N} ist nach oben beschränkt. Dann existiert nach Satz 2.2 das Supremum $M_0 = \sup \mathbb{N}$. Es sei $M := M_0 - \frac{1}{2}$. Da M_0 die kleinste obere Schranke ist, existiert ein $m \in \mathbb{N}$, mit $M_0 - \frac{1}{2} < m$. Folglich ist

$$M_0 < m + \frac{1}{2} < m + 1.$$

Da aber $m + 1 \in \mathbb{N}$ ist, kann M_0 keine obere Schranke sein. Dies ergibt den Widerspruch. Den Beweis für $N \subset \mathbb{N}$ führt man analog. \square

Folgerung 2.2

1. Zu jedem $\varepsilon \in \mathbb{R}^+$ existiert ein $n \in \mathbb{N}$ mit $\frac{1}{n} < \varepsilon$.
2. Zu jedem $q \in \mathbb{N}, q \neq 1$, und $\varepsilon \in \mathbb{R}^+$ existiert ein $n \in \mathbb{N}$ mit $\frac{1}{q^n} < \varepsilon$

Beweis. Zu 1) Zur Zahl $\frac{1}{\varepsilon} \in \mathbb{R}$ existiert nach dem Archimedischen Axiom ein $n \in \mathbb{N}$ mit $\frac{1}{\varepsilon} < n$. Folglich gilt $\frac{1}{n} < \varepsilon$. Zum Beweis der 2. Aussage setzen wir

$$N := \{q^n \mid n \in \mathbb{N}\}.$$

N ist eine unendliche Teilmenge von \mathbb{N} . Den Beweis kann man dann analog zu 1) führen. \square

Folgerung 2.3 *Sei $A \subset \mathbb{Z}$ eine nichtleere, nach oben (unten) beschränkte Menge ganzer Zahlen. Dann besitzt A ein Maximum (Minimum).*

Beweis. Sei A nach unten beschränkt und $d = \inf A$. Nach dem Archimedischen Axiom existiert ein $n_0 \in \mathbb{N}$ mit $|d| < n_0$. Also gilt $-n_0 < d$ und somit $0 < d + n_0 \leq a + n_0$ für alle $a \in A$. Betrachten wir nun die Menge $A_0 := \{a + n_0 \mid a \in A\} \subset \mathbb{N}$. Wir zeigen, dass diese Menge ein kleinstes Element besitzt. Sei $k \in A_0$ und bezeichne $(A_0)_k := \{x \in A_0 \mid x \leq k\}$. Die Menge $(A_0)_k$ ist endlich und besitzt deshalb ein kleinstes Element m_0 (siehe Übungsaufgaben). Dann ist m_0 auch das kleinste Element von A_0 und $m_0 - n_0$ das kleinste Element von A . Folglich gilt $m_0 - n_0 = \min A$.

Ist A von oben beschränkt, so ist $\max A = -\min(-A)$. \square

Satz 2.3 (Die Teilmenge $\mathbb{Q} \subset \mathbb{R}$ liegt dicht in \mathbb{R})

Seien $x, y \in \mathbb{R}$ und $x < y$. Dann existiert eine rationale Zahl $q \in \mathbb{Q}$ mit $x < q < y$.

Beweis. Wir wählen ein $n \in \mathbb{N}$ mit $\frac{1}{n} < y - x$ und setzen

$$A := \{z \in \mathbb{Z} \mid z > n \cdot x\}.$$

Wiederum nach dem Archimedischen Axiom ist A nicht leer und besitzt, da von unten beschränkt, ein Minimum (Folgerung 2.3). Sei $m_0 = \min A$. Dann gilt $m_0 \in A$ und $m_0 - 1 \notin A$. Folglich ist $\frac{m_0}{n} > x$ und $\frac{m_0 - 1}{n} \leq x$. Wir erhalten somit

$$x < \frac{m_0}{n} = \frac{m_0 - 1}{n} + \frac{1}{n} < x + (y - x) = y$$

und folglich liegt die rationale Zahl $q := \frac{m_0}{n}$ im Intervall (x, y) . \square

Definition 2.11. *Eine Familie abgeschlossener Intervalle $I_n \subset \mathbb{R}$, $n \in \mathbb{N}$, heißt Intervallschachtelung, wenn gilt:*

1. $I_n \subset I_m \quad \forall n > m$
2. Zu jeder positiven Zahl $\varepsilon \in \mathbb{R}^+$ existiert ein $n \in \mathbb{N}$ mit $L(I_n) < \varepsilon$.

Satz 2.4 (Prinzip der Intervallschachtelung)

Sei $I_1 \supset I_2 \supset I_3 \supset \dots$ eine Intervallschachtelung. Dann existiert genau eine reelle Zahl $x \in \mathbb{R}$, so dass $x \in I_n$ für alle $n \in \mathbb{N}$, das heißt

$$\bigcap_{n=1}^{\infty} I_n = \{x\}.$$

Beweis.

1. *Existenz:* Sei $I_n = [a_n, b_n]$. Da $I_n \subset I_m$ für alle $n \geq m$, folgt

$$a_m \leq a_n \leq b_n \leq b_m. \quad (*)$$

Wir betrachten die Menge der unteren Intervallgrenzen

$$A := \{a_n \mid n \in \mathbb{N}\} \subset \mathbb{R}.$$

A ist nach oben beschränkt, zum Beispiel durch b_1 , hat also nach Satz 2.2 ein Supremum. Sei $x = \sup A$. Wir zeigen, dass $x \in I_n$ für alle $n \in \mathbb{N}$. Nach Definition ist $a_n \leq x$. Es bleibt zu zeigen, dass $x \leq b_n$ für alle $n \in \mathbb{N}$. Angenommen $x > b_m$ für ein $m \in \mathbb{N}$. Da x die kleinste obere Schranke von A ist, kann b_m keine obere Schranke von A sein. Somit existiert ein $a_n \in A$, so dass $a_n > b_m$. Dies widerspricht aber der Schachtelungseigenschaft (*). Folglich war die Annahme falsch, das heißt $x \leq b_n$ für alle $n \in \mathbb{N}$ und somit gilt $a_n \leq x \leq b_n$, also $x \in I_n$ für alle $n \in \mathbb{N}$.

2. *Eindeutigkeit:* Angenommen es gäbe zwei Zahlen $x, y \in \mathbb{R}$ mit $x \neq y$ und $x, y \in I_n$ für alle $n \in \mathbb{N}$. Sei $\varepsilon = |x - y| > 0$. Dann existiert ein Intervall I_{n_0} mit $L(I_{n_0}) < \varepsilon$. Da $|x - y| > L(I_{n_0})$, können aber nicht beide Zahlen x, y in I_{n_0} liegen. Damit ist die Eindeutigkeit von x gezeigt. \square

Bemerkung: Die Vollständigkeitseigenschaft eines angeordneten Körpers kann man durch das Intervallschachtelungsprinzip oder die Existenz des Supremums ersetzen.

Es gilt: Sei $[\mathbb{K}, +, \cdot]$ ein angeordneter Körper. Dann sind folgende Aussagen äquivalent:

1. Jeder Dedekindsche Schnitt von \mathbb{K} besitzt eine Schnitzzahl.
2. Jede nach oben beschränkte Teilmenge von \mathbb{K} besitzt ein Supremum.
3. Es gilt das Intervallschachtelungsprinzip und das Archimedische Axiom.

Die Implikation: 1. \implies 2. \implies 3. haben wir bewiesen. Die Umkehrung werden wir hier nicht beweisen.

2.1.4 Die Überabzählbarkeit der Menge der reellen Zahlen

Wir beweisen mit Hilfe des Vollständigkeitsaxioms, dass die Menge der reellen Zahlen nicht abzählbar ist. Dazu zunächst einige Definitionen.

Definition 2.12. Eine Menge A heißt abzählbar, wenn es eine bijektive Abbildung $f: \mathbb{N} \rightarrow A$ von der Menge der natürlichen Zahlen auf die Menge A gibt.

Die bijektive Abbildung f liefert uns eine Abzählvorschrift für A : Mit der Bezeichnung $a_n := f(n)$ ist nämlich

$$A = \{a_1, a_2, a_3, \dots\} \quad \text{mit } a_i \neq a_j \text{ für } i \neq j.$$

Definition 2.13. Eine Menge A heißt überabzählbar, wenn sie weder leer, noch endlich oder abzählbar ist. Ist eine Menge A leer, endlich oder abzählbar, so sagen wir auch A ist höchstens abzählbar.

Beispiele:

1. Die Menge der natürlichen Zahlen \mathbb{N} und die Menge \mathbb{N}_0 sind abzählbar.
2. Die Menge der ganzen Zahlen \mathbb{Z} ist abzählbar, denn

$$\begin{aligned} f_{\mathbb{Z}} : \mathbb{N} &\longrightarrow \mathbb{Z} \\ 2k &\longmapsto k \\ 2k + 1 &\longmapsto -k \end{aligned}$$

ist eine bijektive Abbildung zwischen \mathbb{N} und \mathbb{Z} .

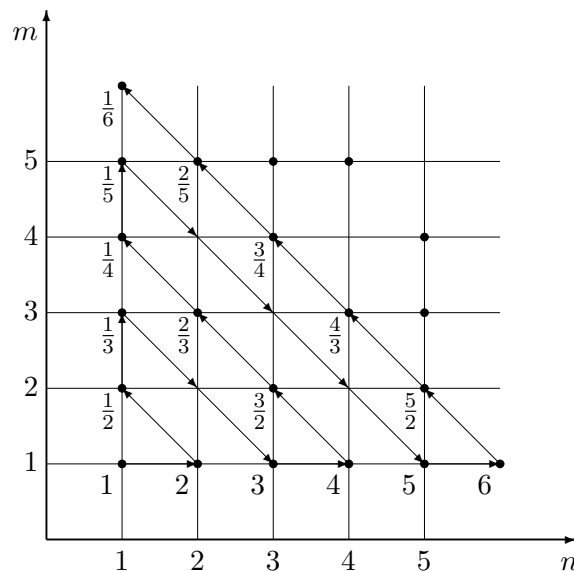
Satz 2.5 Die Menge \mathbb{Q} der rationalen Zahlen ist abzählbar.

Beweis. (1. Cantorsches Diagonalverfahren).

Wir geben zunächst eine Abzählvorschrift der Menge \mathbb{Q}^+ der positiven rationalen Zahlen an. Jede Zahl $q \in \mathbb{Q}^+$ sei als Bruch dargestellt:

$$q = \frac{n}{m}, \quad n, m \text{ teilerfremde natürliche Zahlen.}$$

Wir betrachten das folgende Schema, das die Paare (n, m) als Punkte eines ebenen Gitters darstellt. Dabei werden Punkte ausgelassen, für die m und n nicht teilerfremd sind.



Die Gitterpunkte werden nun längs des im Gitter gezeichneten Streckenzuges nummeriert. Dadurch erreicht man alle Punkte des konstruierten Gitters und erhält somit eine bijektive Abbildung $\varphi : \mathbb{N} \longrightarrow \mathbb{Q}^+$.

Diese Abzählung beginnt offensichtlich mit:

$$1, 2, \frac{1}{2}, \frac{1}{3}, 3, 4, \frac{3}{2}, \frac{2}{3}, \frac{1}{4}, \frac{1}{5}, 5, \dots$$

Wir erweitern nun φ zu einer bijektiven Abbildung $\phi : \mathbb{Z} \rightarrow \mathbb{Q}$ mittels

$$\phi(n) := \begin{cases} \varphi(n) & \text{falls } n \in \mathbb{N} \\ 0 & \text{falls } n = 0 \\ -\varphi(-n) & \text{falls } n \in \mathbb{Z}, n < 0. \end{cases}$$

Die Abbildung $\phi \circ f_{\mathbb{Z}} : \mathbb{N} \rightarrow \mathbb{Z} \rightarrow \mathbb{Q}$ bildet \mathbb{N} bijektiv auf \mathbb{Q} ab. Somit ist \mathbb{Q} abzählbar. \square

Satz 2.6 Die Menge \mathbb{R} der reellen Zahlen ist überabzählbar.

Beweis. Angenommen, es existiert eine Abzählung von \mathbb{R} , d.h. es gilt

$$\mathbb{R} = \{x_1, x_2, x_3, \dots\}.$$

Zu dieser Abzählung konstruieren wir induktiv eine Intervallschachtelung

$$I_1 \supset I_2 \supset I_3 \supset I_4 \supset I_5 \supset \dots$$

Es sei

$$I_1 := [x_1 + 1, x_1 + \frac{4}{3}].$$

Offensichtlich ist $x_1 \notin I_1$ und $L(I_1) = \frac{1}{3}$. Aus einem schon vorhandenen Intervall I_n konstruieren wir I_{n+1} wie folgt: Wir teilen I_n in drei gleichlange, abgeschlossene Intervalle und wählen als I_{n+1} eines dieser Teilintervalle, das x_{n+1} nicht enthält. Für die so konstruierte Folge von Intervallen gilt

$$\begin{aligned} I_1 \supset I_2 \supset I_3 \supset I_4 \supset \dots \\ x_n \notin I_n \\ L(I_1) = \frac{1}{3}, L(I_2) = \frac{1}{3^2}, \dots, L(I_n) = \frac{1}{3^n}, \dots \end{aligned}$$

Nach Folgerung 2.2 existiert für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass $L(I_{n_0}) < \varepsilon$. Somit ist $I_1 \supset I_2 \supset I_3 \supset \dots$ eine Intervallschachtelung von \mathbb{R} . Nach dem Intervallschachtelungsprinzip enthält der Durchschnitt $\bigcap_{n=1}^{\infty} I_n$ genau eine reelle Zahl. Sei nun $x \in \bigcap_{n=1}^{\infty} I_n$ diese Zahl. Nach Annahme war $\mathbb{R} = \{x_1, x_2, x_3, x_4, \dots\}$. Es muß also ein $k_0 \in \mathbb{N}$ mit $x = x_{k_0}$ geben. Dann ist $x_{k_0} \in I_{k_0}$. Dies widerspricht aber der Konstruktion der Intervalle. Somit war die Annahme der Abzählbarkeit von \mathbb{R} falsch. \square

Definition 2.14. Zwei Mengen A und B heißen gleichmächtig, falls eine bijektive Abbildung $f : A \rightarrow B$ existiert. Die Menge B hat eine größere Mächtigkeit als A , falls A zu einer Teilmenge von B gleichmächtig ist, aber B zu keiner Teilmenge von A .

Die Mengen \mathbb{N} , \mathbb{Z} und \mathbb{Q} sind gleichmächtig. Die Menge \mathbb{R} hat eine größere Mächtigkeit als diese drei Mengen. Es gibt Mengen, die eine größere Mächtigkeit als \mathbb{R} haben (z.B. die Potenzmengen $\mathcal{P}(\mathbb{R})$ – siehe Übung).

Kontinuumshypothese : *Es gibt keine Menge A , deren Mächtigkeit größer als die von \mathbb{N} und kleiner als die von \mathbb{R} ist.*

Diese Hypothese wurde 1878 von Georg Cantor [1845–1918] aufgestellt. Sie leitete die Entwicklung der Mengenlehre ein. Auf dem Internationalen Mathematikerkongreß 1900 in Paris hat David Hilbert [1862-1943] seine berühmte Liste von 23 ungelösten mathematischen Problemen vorgestellt. Die Kontinuumshypothese steht dabei an 1. Stelle. Inzwischen weiß man, dass sie auf der Basis der heute zugrundegelegten Axiome der Mengenlehre weder beweisbar noch widerlegbar. Kurt Gödel [1906-1978] hat 1939 gezeigt, dass sie nicht widerlegbar ist, Paul Cohen [1934-2007] hat 1963 gezeigt, dass sie auch nicht beweisbar ist (dafür hat er 1966 die Fields-Medaille bekommen). Solche Fragen werden in der *mathematischen Logik* behandelt.

2.1.5 Wurzeln und Potenzen reeller Zahlen

In diesem Abschnitt behandeln wir einige wichtige Gleichungen und Ungleichungen für Potenzen und Wurzeln reeller Zahlen.

Sei $x \in \mathbb{R}$ eine reelle Zahl. Die Potenz x^n für $n \in \mathbb{N}_0$ definieren wir induktiv durch:

$$x^0 := 1, \quad x^1 := x, \quad x^2 := x \cdot x, \quad \dots, \quad x^{n+1} := x^n \cdot x.$$

Für $x \neq 0$ setzen wir

$$x^{-n} := \frac{1}{x^n}.$$

Damit ist die k -te Potenz x^k für jede ganze Zahl $k \in \mathbb{Z}$ und jede reelle Zahl $x \in \mathbb{R}$, $x \neq 0$, definiert. Aus den Körper- und Anordnungseigenschaften der reellen Zahlen folgt sofort

1. Für $x \in \mathbb{R}$ mit $x \neq 0$ und $k, l \in \mathbb{Z}$ gilt

$$x^k \cdot x^l = x^{k+l}, \quad x^{k \cdot l} = (x^k)^l \quad \text{und} \quad (x \cdot y)^k = x^k \cdot y^k.$$

2. Ist $0 < x < y$, dann gilt $x^n < y^n$ für alle $n \in \mathbb{N}$.

Satz 2.7 (Binomischer Satz) *Seien $x, y \in \mathbb{R}$. Dann gilt für jedes $n \in \mathbb{N}$*

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

Beweis. Wir führen den Beweis durch vollständige Induktion über n :

Induktionsanfang: Für $n = 1$ gilt die Aussage, denn: $\binom{1}{0}x^0 \cdot y^1 + \binom{1}{1}x^1 \cdot y^0 = y + x = (x + y)^1$.

Induktionsschritt:

Induktionsvoraussetzung: Für ein $n \in \mathbb{N}$ gilt:

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

Induktionsbehauptung:

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

Induktionsbeweis:

$$\begin{aligned} (x + y)^{n+1} &= (x + y)^n \cdot (x + y) \\ &\stackrel{IV}{=} \left(\sum_{k=0}^n \binom{n}{k} x^k \cdot y^{n-k} \right) \cdot (x + y) \\ &= \sum_{k=0}^n \binom{n}{k} x^k \cdot x \cdot y^{n-k} + \sum_{k=0}^n \binom{n}{k} x^k \cdot y^{n-k} \cdot y \\ &= \sum_{k=0}^n \binom{n}{k} x^{k+1} \cdot y^{n+1-(k+1)} + \sum_{k=0}^n \binom{n}{k} x^k \cdot y^{n+1-k} \\ &= \sum_{l=1}^{n+1} \binom{n}{l-1} x^l \cdot y^{(n+1)-l} + \sum_{l=0}^n \binom{n}{l} x^l \cdot y^{(n+1)-l} \\ &= \sum_{l=1}^n \left(\binom{n}{l} + \binom{n}{l-1} \right) x^l \cdot y^{(n+1)-l} + \binom{n}{n} x^{n+1} y^0 + \binom{n}{0} x^0 y^{n+1} \\ &\stackrel{1.12}{=} \sum_{l=0}^{n+1} \binom{n+1}{l} x^l \cdot y^{(n+1)-l}. \end{aligned}$$

□

Aus dem Binomischen Satz 2.7 ergibt sich die

Folgerung 2.4

1. $(1 + x)^n = \sum_{k=0}^n \binom{n}{k} x^k,$
2. $\sum_{k=0}^n \binom{n}{k} = 2^n,$
3. $\sum_{k=0}^n (-1)^k \binom{n}{k} = 0.$

Beweis. (1) ist der Binomische Satz für $y = 1$, (2) ist der Binomische Satz für $x = y = 1$ und (3) ist der Binomische Satz für $x = -1$, $y = 1$. □

Satz 2.8 (Bernoullische Ungleichung) Für jede reelle Zahl $x \geq -1$ und für jedes $n \in \mathbb{N}$ gilt:

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

Beweis. Beweis durch vollständige Induktion über n .

Induktionsanfang: Die Aussage gilt offensichtlich für $n = 1$.

Induktionsschritt:

Induktionsvoraussetzung: Für ein $n \in \mathbb{N}$ gilt $(1 + x)^n \geq 1 + nx$.

Induktionsbehauptung: $(1 + x)^{n+1} \geq 1 + (n + 1)x$.

Induktionsbeweis: ´

$$(1 + x)^{n+1} = (1 + x)^n(1 + x) \stackrel{IV}{\geq} (1 + nx)(1 + x) = 1 + (n + 1)x + \underbrace{nx^2}_{\geq 0} \geq 1 + (n + 1)x.$$

□

Als Anwendung erhält man unmittelbar

Folgerung 2.5

1. Sei $y \in \mathbb{R}$, $y > 1$, und $r \in \mathbb{R}^+$. Dann existiert ein $n \in \mathbb{N}$, so daß $y^n > r$.
2. Sei $y \in \mathbb{R}$, $0 < y < 1$ und $r \in \mathbb{R}^+$. Dann existiert ein $n \in \mathbb{N}$ mit $y^n < r$.

Beweis. Sei $r \in \mathbb{R}^+$ und $y > 1$. Nach dem Archimedischen Axiom für reelle Zahlen existiert eine natürliche Zahl $n \in \mathbb{N}$, so daß $n > \frac{r}{y-1}$. Dann folgt mit der Bernoullischen Ungleichung

$$y^n = (1 + (y - 1))^n \geq 1 + n(y - 1) \geq n(y - 1) > r.$$

Ist $0 < y < 1$, so wenden wir das eben Bewiesene auf die reelle Zahl $\frac{1}{y} > 1$ und $\frac{1}{r}$ an und erhalten eine natürliche Zahl $n \in \mathbb{N}$ mit $(\frac{1}{y})^n > \frac{1}{r}$ und somit $y^n < r$. □

Satz 2.9 (Geometrische Summe) Für jede reelle Zahl $x \neq 1$ und jede natürliche Zahl n gilt:

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

Beweis. Beweis durch vollständige Induktion über n .

Induktionsanfang: $n = 1$:

$$\frac{1 - x^2}{1 - x} = \frac{(1 - x)(1 + x)}{1 - x} = 1 + x = x^0 + x^1.$$

Induktionsschritt:

Induktionsvoraussetzung: Die Behauptung ist für ein $n \in \mathbb{N}$ richtig.

Induktionsbehauptung:

$$\sum_{k=0}^{n+1} x^k = \frac{1 - x^{n+2}}{1 - x}.$$

Induktionsbeweis:

$$\begin{aligned}
\sum_{k=0}^{n+1} x^k &= \left(\sum_{k=0}^n x^k \right) + x^{n+1} \\
&\stackrel{IV}{=} \frac{1 - x^{n+1}}{1 - x} + x^{n+1} \\
&= \frac{1 - x^{n+1} + x^{n+1}(1 - x)}{1 - x} \\
&= \frac{1 - x^{n+2}}{1 - x}. \quad \square
\end{aligned}$$

Wir beweisen nun die Existenz der n -ten Wurzel einer positiven reellen Zahl.

Satz 2.10 Sei $x \in \mathbb{R}^+$ eine positive reelle Zahl und $n \in \mathbb{N}$. Dann existiert genau eine positive reelle Zahl $y \in \mathbb{R}^+$ mit $y^n = x$.

Bezeichnung: $y := \sqrt[n]{x}$ heißt die n -te Wurzel aus x .

Beweis. Zum Beweis benutzen wir das Intervallschachtelungsprinzip. Für $x = 1$ gilt die Behauptung offensichtlich. Es genügt, den Fall $x > 1$ zu behandeln. Den Fall $x < 1$ führt man durch Übergang zu $x' := \frac{1}{x}$ darauf zurück.

Wir definieren zunächst induktiv die folgende Familie abgeschlossener Intervalle $I_k := [a_k, b_k]$: Wir setzen $I_1 = [a_1, b_1] := [1, x]$. Dann gilt $a_1^n \leq x \leq b_1^n$, da $1 < x$. Sei das Intervall $I_k := [a_k, b_k]$ mit der Eigenschaft $a_k^n \leq x \leq b_k^n$ bereits konstruiert. Dann definieren wir I_{k+1} durch Halbierung von I_k : Sei $m = \frac{a_k + b_k}{2}$ der Mittelpunkt von I_k . Wir setzen dann

$$I_{k+1} = [a_{k+1}, b_{k+1}] := \begin{cases} [a_k, m] & \text{falls } m^n \geq x \\ [m, b_k] & \text{falls } m^n < x. \end{cases}$$

Dann gilt nach Konstruktion für diese Familie von abgeschlossenen Intervallen:

1. $I_1 \supset I_2 \supset I_3 \supset \dots$
2. $L(I_k) = (x - 1) \cdot \left(\frac{1}{2}\right)^{k-1}$ für alle $k \in \mathbb{N}$.
3. $a_k^n \leq x \leq b_k^n$ für alle $k \in \mathbb{N}$.

Wir erhalten also ineinander geschachtelte Intervalle, deren Längen nach Folgerung 2.2 beliebig klein werden. Nach dem Intervallschachtelungsprinzip existiert genau eine reelle Zahl $y \in \mathbb{R}$ mit $y \in I_k$ für jedes $k \in \mathbb{N}$. Wir zeigen nun, dass $y^n = x$ gilt.

Dazu betrachten wir die Intervalle

$$J_k := [a_k^n, b_k^n], \quad k \in \mathbb{N}.$$

Da $I_k \supset I_{k+1}$, gilt wegen der Monotonie der Potenzen auch $J_k \supset J_{k+1}$. Für die Länge von J_k erhalten wir

$$\begin{aligned}
L(J_k) &= b_k^n - a_k^n \\
&= (b_k - a_k)(b_k^{n-1} + b_k^{n-2}a_k + \dots + b_k a_k^{n-2} + a_k^{n-1}) \\
&= L(I_k) \cdot b_k^{n-1} \left(1 + \frac{a_k}{b_k} + \frac{a_k^2}{b_k^2} + \dots + \frac{a_k^{n-1}}{b_k^{n-1}} \right) \\
&\leq (x - 1) \cdot \frac{1}{2^{k-1}} \cdot b_1^{n-1} \cdot n.
\end{aligned}$$

Nach Folgerung 2.2 gibt es zu jedem $\varepsilon > 0$ ein $k \in \mathbb{N}$ mit $L(J_k) < \varepsilon$. Die Familie der Intervalle J_k ist also eine Intervallschachtelung von \mathbb{R} . Nach Konstruktion gilt aber sowohl $x \in J_k$ (Eigenschaft 3.) als auch $y^n \in J_k$ für alle $k \in \mathbb{N}$. Da der Durchschnitt $\bigcap_{k=1}^{\infty} J_k$ nur ein Element enthält, folgt $x = y^n$.

Die Eindeutigkeit der Zahl $y \in \mathbb{R}^+$ mit $y^n = x$ ist klar, denn ist z.B. $y_1 < y_2$, so folgt $y_1^n < y_2^n$. \square

Die Gleichung $y^n = x$ hat für gerade n zwei reelle Lösungen $y_1 = \sqrt[n]{x}$ und $y_2 = -\sqrt[n]{x}$. Die Eindeutigkeitsaussage von Satz 2.10 gilt also nur in \mathbb{R}^+ . Die Gleichung $y^n = x$ ist in \mathbb{Q} im Allgemeinen nicht lösbar.

Satz 2.11 *Seien n und k natürliche Zahlen. Dann ist $\sqrt[n]{k}$ genau dann eine rationale Zahl, falls k die n -te Potenz einer natürlichen Zahl ist, das heißt falls $k = m^n$ für ein $m \in \mathbb{N}$. Insbesondere gilt:*

- Für jede Primzahl p und jedes $n > 1$ ist die Zahl $\sqrt[n]{p}$ irrational.
- Wenn $\sqrt[n]{k}$ rational ist, so ist $\sqrt[n]{k}$ sogar eine natürliche Zahl.

Beweis.

1. (\Leftarrow): Sei $k = m^n$ mit $m \in \mathbb{N}$. Dann ist per Definition $m := \sqrt[n]{k} \in \mathbb{N} \subset \mathbb{Q}$.
2. (\Rightarrow): Sei $\sqrt[n]{k} \in \mathbb{Q}$. Dann existieren teilerfremde Zahlen $m, l \in \mathbb{N}$, so dass $\sqrt[n]{k} = \frac{m}{l}$. Nach Definition erhält man $k = (\frac{m}{l})^n = \frac{m^n}{l^n}$ und somit $kl^n = m^n$. Wir zeigen nun, dass $l = 1$ gilt. Angenommen $l > 1$. Dann existiert eine Primzahl $p > 1$, die l teilt. Folglich teilt p auch $kl^n = m^n$, das heißt p teilt auch m . Das ist aber ein Widerspruch dazu, dass l und m teilerfremd sind. Somit ist $l = 1$ und $k = m^n$ für $m \in \mathbb{N}$. \square

Wir können jetzt die Potenzen mit rationalen Exponenten definieren.

Definition 2.15. *Sei $x \in \mathbb{R}^+$ eine positive reelle Zahl und $q \in \mathbb{Q}$ eine rationale Zahl mit der Darstellung $q = \frac{n}{m}, n \in \mathbb{Z}, m \in \mathbb{N}$. Dann setzen wir:*

$$x^q := \left(\sqrt[m]{x} \right)^n.$$

Diese Definition ist korrekt, d.h. unabhängig von der Wahl der Darstellung von q .

Die folgenden Eigenschaften für die Potenzen mit rationalen Exponenten sind leicht nachzuprüfen: Seien $p, q \in \mathbb{Q}$ und $x, y \in \mathbb{R}^+$. Dann erhält man:

1. $x^p \cdot x^q = x^{p+q}$, $(x^q)^p = x^{p \cdot q}$, $x^q \cdot y^q = (xy)^q$.
2. Sei $p < q$. Dann gilt $x^p < x^q$ falls $x > 1$ und $x^p > x^q$ falls $0 < x < 1$.
3. Sei $0 < x < y$. Dann gilt $x^q < y^q$ falls $q > 0$ und $x^q > y^q$ falls $q < 0$.

Wir werden auf die Potenzen und ihre Eigenschaften später zurückkommen.

2.2 Die komplexen Zahlen

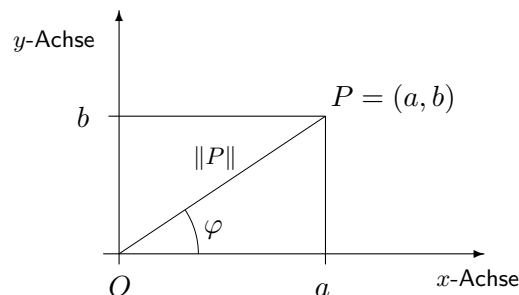
Für jede von Null verschiedene reelle Zahl x gilt $x^2 > 0$. Man kann im Zahlbereich der reellen Zahlen also keine Wurzeln aus negativen Zahlen ziehen. Insbesondere gibt es keine reelle Lösung der Gleichung $x^2 = -1$. Die komplexen Zahlen sind eine *Erweiterung der reellen Zahlen*, die es möglich macht, auch *Wurzeln aus negativen Zahlen zu ziehen*.

2.2.1 Geometrische Motivation für die Definition der komplexen Zahlen

Bevor wir die Definition der komplexen Zahlen angeben, wollen wir eine geometrische Motivation für diese Definition angeben. Wir suchen eine Erweiterung von \mathbb{R} , in der wir die Gleichung $x^2 = -1$ lösen können. Die reellen Zahlen können wir geometrisch als Punkte auf einer Geraden (reelle Zahlengerade) darstellen. Wir erweitern nun diese Gerade zu einer Euklidischen Ebene \mathcal{E} und führen auf dieser Ebene auf geometrische Weise eine Addition sowie eine Multiplikation von Punkten so ein, dass es einen Punkt gibt, für den die Gleichung $P \cdot P = -1$ gilt. Dazu fixieren wir auf der Ebene \mathcal{E} ein kartesisches Koordinatensystem, so dass die x -Achse der reellen Zahlengeraden entspricht. Dadurch können wir die Ebene \mathcal{E} mit der Menge der geordneten 2-Tupel reeller Zahlen \mathbb{R}^2 identifizieren: jedem Punkt $P \in \mathcal{E}$ werden seine Koordinaten $(a, b) \in \mathbb{R}^2$ in diesem kartesischen Koordinatensystem zugeordnet (Euklidische Koordinaten). Der Ursprung $O \in \mathcal{E}$ hat dabei die Koordinaten $(0, 0)$. Die reellen Zahlen entsprechen der Geraden $g = \{(x, 0) \mid x \in \mathbb{R}\} \subset \mathbb{R}^2$. Dies liefert eine naheliegende Möglichkeit, zwei Punkte der Ebene zu addieren: man addiert ihre Euklidischen Koordinaten: Für $P_1 = (a_1, b_1)$ und $P_2 = (a_2, b_2)$ sei

$$P_1 + P_2 := (a_1 + a_2, b_1 + b_2). \quad (2.1)$$

Das ist verträglich mit der Addition in den "reellen Zahlen" $\{(x, 0) \mid x \in \mathbb{R}\}$. Würde man die Multiplikation von Punkten auf die analoge Weise definieren, würde man nichts gewinnen. Man könnte die Gleichung $P \cdot P = (-1, 0)$ nicht lösen. Für eine geeignete Definition der Multiplikation benutzen wir die Euklidische Drehung von Punkten um den Ursprung O . Dazu erinnern wir an die Darstellung eines Punktes der Ebene durch Polarkoordinaten: Ist $P \neq O$, so ist P eindeutig durch die Länge der Strecke \overline{OP} , sowie den orientierten Winkel φ zwischen der x -Achse und dem Stahl \overrightarrow{OP} (gemessen entgegen dem Uhrzeigersinn) bestimmt.



Dabei gilt für $P = (a, b)$ nach dem Satz des Pythagoras sowie der elementargeometrischen Definition der Sinus- und Cosinus-Funktion im rechtwinkligen Dreieck⁵:

$$\begin{aligned} |\overline{OP}| &= \sqrt{a^2 + b^2} =: \|P\|, \\ P = (a, b) &= (\|P\| \cos \varphi, \|P\| \sin \varphi). \end{aligned}$$

Für die "reellen Zahlen" $P = (x, 0)$ gilt $\|P\| = |x|$ und $\varphi = 0$, falls x positiv, bzw. $\varphi = 180^\circ$, falls x negativ ist. Wir multiplizieren nun zwei Punkte

$$\begin{aligned} P_1 &= (a_1, b_1) = (\|P_1\| \cos \varphi_1, \|P_1\| \sin \varphi_1) \quad \text{und} \\ P_2 &= (a_2, b_2) = (\|P_2\| \cos \varphi_2, \|P_2\| \sin \varphi_2) \end{aligned} \quad (2.2)$$

durch die Forderung, dass sich ihre Abstände $\|P_1\|$ und $\|P_2\|$ multiplizieren sowie ihre Winkel φ_1 bzw. φ_2 addieren:

$$P_1 \cdot P_2 := (\|P_1\| \cdot \|P_2\| \cos(\varphi_1 + \varphi_2), \|P_1\| \cdot \|P_2\| \sin(\varphi_1 + \varphi_2)). \quad (2.3)$$

Dies ist verträglich mit der Multiplikation in den "reellen Zahlen" $\{(x, 0) \mid x \in \mathbb{R}\}$.

Nun "sieht" man aber sehr einfach einen Punkt P der Ebene mit $P \cdot P = (-1, 0)$, nämlich den Punkt P mit den Polarkoordinaten $\|P\| = 1$ und $\varphi = 90^\circ$. Für diesen Punkt gilt

$$P \cdot P = (\cos 180^\circ, \sin 180^\circ) = (-1, 0).$$

Benutzt man die Additionstheoreme für die Cosinus- und Sinusfunktion⁶:

$$\begin{aligned} \cos(\alpha + \beta) &= \cos \alpha \cos \beta - \sin \alpha \sin \beta, \\ \sin(\alpha + \beta) &= \sin \alpha \cos \beta + \cos \alpha \sin \beta, \end{aligned}$$

so erhält man mit (2.2) folgende Formel, die das Produkt $P_1 \cdot P_2$ in Euklidischen Koordinaten ausdrückt:

$$\begin{aligned} P_1 \cdot P_2 &= (\|P_1\| \|P_2\| (\cos \varphi_1 \cos \varphi_2 - \sin \varphi_1 \sin \varphi_2), \|P_1\| \|P_2\| (\sin \varphi_1 \cos \varphi_2 + \cos \varphi_1 \sin \varphi_2)) \\ &= (a_1 a_2 - b_1 b_2, b_1 a_2 + a_1 b_2). \end{aligned} \quad (2.4)$$

Diese geometrische Motivation führt uns zu der folgenden *Definition* der komplexen Zahlen, die vom irischen Mathematiker und Physiker William R. Hamilton [1805 – 1865] eingeführt wurde. Hamilton hat auf analoge Weise auch den Zahlbereich der komplexen Zahlen erweitert und den Schiefkörper der Quaternionen eingeführt, der wie die komplexen Zahlen nicht nur in der Mathematik, sondern auch für viele Belange der Physik nützlich ist.

⁵ Man setzt dazu die im rechtwinkligen Dreieck definierte Cosinus- und Sinusfunktion auf $[0^\circ, 360^\circ]$ fort durch: $\cos 0^\circ := 1$, $\sin 0^\circ := 0$, $\cos 90^\circ := 0$, $\sin 90^\circ := 1$ sowie $\cos(\varphi + 90^\circ) := -\sin \varphi$, $\sin(\varphi + 90^\circ) := \cos \varphi$. Anschließend setzt man sie 360° -periodisch auf \mathbb{R}° fort.

⁶ Bitte wiederholen Sie selbständig die elementargeometrischen Beweise der Additionstheoreme. Man findet sie z.B. im Buch von Helmut Koch: *Einführung in die Mathematik*, Springer 2004, Kapitel 8.

2.2.2 Definition und Eigenschaften der komplexen Zahlen

Auf der Menge der Paare reeller Zahlen

$$\mathbb{R}^2 := \mathbb{R} \times \mathbb{R} := \{(a, b) \mid a, b \in \mathbb{R}\}$$

führen wir eine Addition $+$: $\mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ und eine Multiplikation \cdot : $\mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ein. Zwei Paare $z_1 = (a_1, b_1)$ und $z_2 = (a_2, b_2)$ aus \mathbb{R}^2 addieren bzw. multiplizieren wir (motiviert durch (2.1) und (2.4)) nach folgenden Regeln:

$$z_1 + z_2 = (a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2), \quad (2.5)$$

$$z_1 \cdot z_2 = (a_1, b_1) \cdot (a_2, b_2) := (a_1 a_2 - b_1 b_2, b_1 a_2 + a_1 b_2). \quad (2.6)$$

Die mit dieser Addition und Multiplikation ausgestattete Menge \mathbb{R}^2 bezeichnet man mit dem neuen Symbol \mathbb{C} , d.h. $\mathbb{C} := \mathbb{R}^2$, um auszudrücken, dass man außer der üblichen Addition (2.5) der reellen Paare auch noch die Multiplikation (2.6) festgelegt hat. Die Elemente von \mathbb{C} heißen *komplexe Zahlen*.

Satz 2.12 *Die komplexen Zahlen $[\mathbb{C}, +, \cdot]$ bilden einen Körper. Es gelten also die Rechenregeln K1 – K9 für die Addition $+$ und die Multiplikation \cdot .*

Beweis. Diese Eigenschaften folgen direkt aus den Körpereigenschaften von \mathbb{R} und den Definitionen von $+$ und \cdot . Man erhält z.B. durch direktes Nachrechnen: $(0, 0)$ ist das neutrale Element der Addition, $(1, 0)$ das neutrale Element der Multiplikation. Das Negative zu $z = (a, b) \in \mathbb{C}$ ist $-z := (-a, -b)$. Das Inverse zu $w = (a, b) \neq (0, 0)$ ist $\frac{1}{w} := \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right)$, denn es gilt

$$(a, b) \cdot \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right) = \left(\frac{a^2+b^2}{a^2+b^2}, \frac{ba-ba}{a^2+b^2}\right) = (1, 0). \quad \square$$

Im Gegensatz zum Körper der reellen Zahlen ist der Körper der komplexen Zahlen nicht angeordnet (Übungsaufgabe).

Potenzen komplexer Zahlen

Für $z \in \mathbb{C}$ mit $z \neq 0$ sei $z^{-1} := \frac{1}{z}$. Die Potenzen z^n für $n \in \mathbb{N}$ seien induktiv durch $z^1 := z$, $z^{n+1} := z^n \cdot z$ erklärt. Weiterhin sei $z^{-n} := \left(\frac{1}{z}\right)^n = \frac{1}{z^n}$. Wir vereinbaren zusätzlich für jedes $z \in \mathbb{C}$, dass $z^0 = 1$. Für zwei komplexe Zahlen $z, w \in \mathbb{C}$ beweist man wie im Reellen (Satz 2.7) die **binomische Formel**

$$(z + w)^n = \sum_{k=0}^n \binom{n}{k} z^k w^{n-k} \quad \text{für alle } n \in \mathbb{N}.$$

Die algebraische Form der komplexen Zahlen

Für den bequemen Umgang mit den komplexen Zahlen eignen sich die nun folgenden Vereinbarungen:

Nach Definition gilt für die komplexen Zahlen $(a, 0)$ und $(b, 0)$

$$(a, 0) + (b, 0) = (a + b, 0) \quad \text{und} \quad (a, 0) \cdot (b, 0) = (a \cdot b, 0).$$

Die Zuordnung $a \in \mathbb{R} \mapsto (a, 0) \in \mathbb{C}$ ist also eine Einbettung der Menge der reellen Zahlen in die Menge der komplexen Zahlen, die mit den jeweiligen Körperoperationen $+$ und \cdot verträglich ist. Wir können deshalb \mathbb{R} als Teilkörper von \mathbb{C} auffassen. Dies werden wir in Zukunft tun und die komplexe Zahl $(a, 0)$ einfach mit a bezeichnen. Dies rechtfertigt auch die Bezeichnung $0 := (0, 0)$ für das neutrale Element der Addition und $1 := (1, 0)$ für das neutrale Element der Multiplikation. Die komplexe Zahl $(0, 1)$ bezeichnen wir mit i und nennen sie die *imaginäre Einheit*. Für $i = (0, 1)$ gilt

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

Die Gleichung $x^2 = -1$ ist also im Körper der komplexen Zahlen lösbar.

Ist $z = (a, b)$ eine beliebige komplexe Zahl, so gilt mit unseren Vereinbarungen

$$z = (a, b) = (a, 0) + (0, b) = (a, 0) + (0, 1)(b, 0) = a + i \cdot b.$$

Jede komplexe Zahl $z \in \mathbb{C}$ ist also in der Form

$$\boxed{z = a + ib \quad a, b \in \mathbb{R}} \quad (2.7)$$

darstellbar. Dies ist die übliche Darstellung der komplexen Zahlen. Man nennt die Darstellung (2.7) auch die *algebraische Form* der komplexen Zahl. Benutzt man diese Darstellung, dann kann man mit den komplexen Zahlen genauso wie mit den reellen Zahlen rechnen, indem man $i^2 = -1$ berücksichtigt. Es gilt also für $z_1 = a_1 + i b_1$ und $z_2 = a_2 + i b_2$

$$z_1 + z_2 = (a_1 + i b_1) + (a_2 + i b_2) = (a_1 + a_2) + i (b_1 + b_2) \quad (2.8)$$

$$z_1 \cdot z_2 = (a_1 + i b_1) \cdot (a_2 + i b_2) = (a_1 a_2 - b_1 b_2) + i (a_1 b_2 + b_1 a_2) \quad (2.9)$$

Definition 2.16. Ist $z = a + ib \in \mathbb{C}$, so heißt $\operatorname{Re}(z) := a$ Realteil von z und $\operatorname{Im}(z) := b$ Imaginärteil von z . Ist $\operatorname{Re}(z) = 0$, so heißt z rein imaginär. Ist $\operatorname{Im}(z) = 0$, so heißt z reell.

Beispiel: Sei $z = a + ib \neq 0$. Dann ist

$$\frac{1}{z} = \frac{1}{a + ib} = \frac{a - ib}{(a + ib)(a - ib)} = \frac{a - ib}{a^2 + b^2} = \frac{a}{a^2 + b^2} - i \frac{b}{a^2 + b^2},$$

also gilt

$$\operatorname{Re}\left(\frac{1}{z}\right) = \frac{a}{a^2 + b^2} \quad \text{bzw.} \quad \operatorname{Im}\left(\frac{1}{z}\right) = -\frac{b}{a^2 + b^2}.$$

Definition 2.17. Ist $z = a + ib \in \mathbb{C}$ eine komplexe Zahl, so heißt $\bar{z} := a - ib$ die konjugiert komplexe Zahl zu z .

Es gelten folgende, mit Hilfe der Definitionen leicht zu überprüfende Rechenregeln:

Satz 2.13 Für alle komplexen Zahlen z und w gilt:

1. $\overline{z+w} = \bar{z} + \bar{w}$, $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$, $\overline{\bar{z}} = z$.
2. $z + \bar{z} = 2 \cdot \operatorname{Re}(z)$, $z - \bar{z} = 2i \cdot \operatorname{Im}(z)$.
3. $\bar{z} = z \iff z \in \mathbb{R}$.
4. $z \cdot \bar{z} = \operatorname{Re}(z)^2 + \operatorname{Im}(z)^2$. Insbesondere ist $0 \leq z \cdot \bar{z} \in \mathbb{R}$. □

Definition 2.18. Sei $z = a + ib \in \mathbb{C}$ eine komplexe Zahl. Der Betrag von z ist die reelle Zahl

$$|z| := \sqrt{a^2 + b^2} = \sqrt{z \cdot \bar{z}}.$$

Satz 2.14 (Eigenschaften des Betrages komplexer Zahlen)

Seien z und w komplexe Zahlen. Dann gilt:

1. $|z| \geq 0$, wobei $|z| = 0$ genau dann, wenn $z = 0$.
2. $|z \cdot w| = |z| \cdot |w|$.
3. $|z + w| \leq |z| + |w|$. (Dreiecksungleichung)
4. $||z| - |w|| \leq |z + w|$.
5. $|z| = |\bar{z}|$
6. $|\operatorname{Re}(z)| \leq |z|$, $|\operatorname{Im}(z)| \leq |z|$.

Beweis. 1., 5. und 6. folgen trivialerweise aus der Definition. Formel 2. folgt aus

$$|z \cdot w|^2 = (zw)(\overline{zw}) = z\bar{z} \cdot w\bar{w} = |z|^2 \cdot |w|^2.$$

Die Dreiecksungleichung erhält man durch

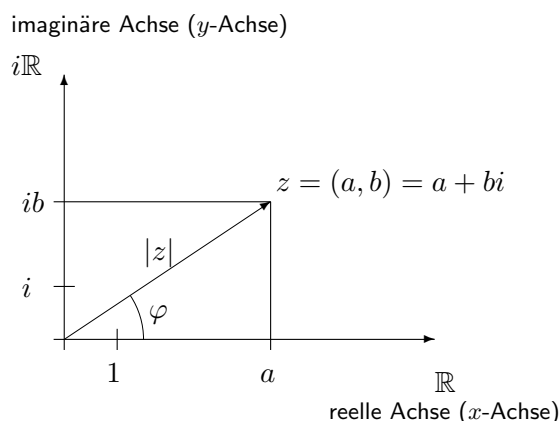
$$\begin{aligned} |z + w|^2 &= (z + w)(\overline{z + w}) \\ &= (z + w)(\bar{z} + \bar{w}) \\ &= z\bar{z} + w\bar{w} + w\bar{z} + z\bar{w} \\ &= z\bar{z} + w\bar{w} + w\bar{z} + \overline{w\bar{z}} \\ &= |z|^2 + |w|^2 + 2 \cdot \operatorname{Re}(w\bar{z}) \\ &\leq |z|^2 + |w|^2 + 2 \cdot |w\bar{z}| \\ &= |z|^2 + |w|^2 + 2 \cdot |w| \cdot |z| \\ &= (|z| + |w|)^2. \end{aligned}$$

Formel 4 wird wie in Satz 2.1 mit Hilfe der Dreiecksungleichung bewiesen. □

Die Polarform (trigonometrische Darstellung) der komplexen Zahlen

Wir kommen nochmal auf die geometrische Interpretation der komplexen Zahlen zurück. Der Darstellung der reellen Zahlen auf einer Geraden entspricht die Darstellung der komplexen Zahlen in der Ebene, die man deshalb oft Gaußsche Zahlenebene oder komplexe Zahlenebene nennt.

Wir fixieren ein kartesisches Koordinatensystem in der Ebene und stellen die komplexe Zahl $z = (a, b) = a + ib \in \mathbb{C}$ als Punkt der Ebene mit den Koordinaten (a, b) dar.



Die reellen Zahlen \mathbb{R} entsprechen der x -Achse, die rein imaginären Zahlen $i\mathbb{R}$ der y -Achse. Nach dem Satz von Pythagoras ist der Betrag $|z| = \sqrt{a^2 + b^2}$ gleich dem Abstand des Punktes $z = (a, b)$ zum Ursprung des Koordinatensystems. Die komplexe Zahl $\bar{z} = (a, -b) = a - ib$ entsteht durch Spiegelung von z an der reellen Achse. Für $z \neq 0$ sei φ der orientierte Winkel zwischen der x -Achse und dem Strahl vom Ursprung durch z , gemessen in positiver Richtung (entgegen dem Uhrzeigersinn). Dann gilt

$$\cos \varphi = \frac{a}{|z|} \quad \text{und} \quad \sin \varphi = \frac{b}{|z|}.$$

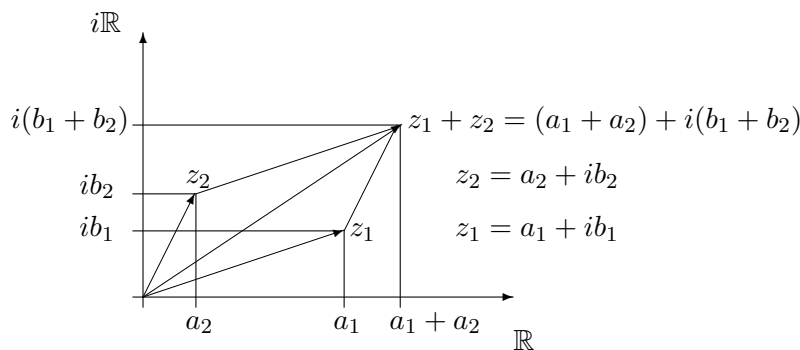
Die Darstellung

$$z = |z| (\cos \varphi + i \cdot \sin \varphi) \tag{2.10}$$

heißt *trigonometrische Darstellung* oder *Polarform* der komplexen Zahl $z \neq 0$. Der Winkel φ heißt *Argument* von z und wird mit $\arg(z)$ bezeichnet. Das Argument φ ist bis auf ganzzahlige Vielfache von 360° eindeutig bestimmt.

Geometrische Deutung von $z_1 + z_2$:

Die Summe $z_1 + z_2$ entspricht dem Endpunkt der vom Nullpunkt ausgehenden Diagonalen im von z_1 und z_2 gebildeten Parallelogramm.



Geometrische Deutung von $z_1 \cdot z_2$:

Seien $z_1, z_2 \neq 0$. Wir betrachten die trigonometrische Darstellung von z_1 und z_2

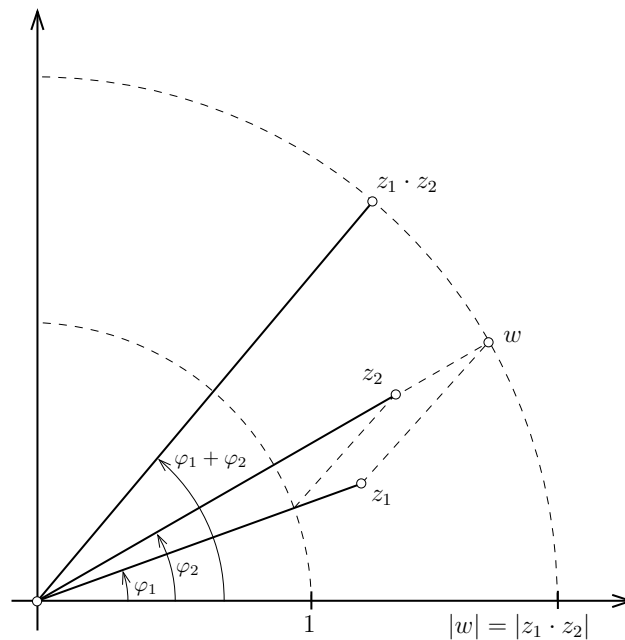
$$\begin{aligned} z_1 &= |z_1|(\cos \varphi_1 + i \sin \varphi_1), \\ z_2 &= |z_2|(\cos \varphi_2 + i \sin \varphi_2). \end{aligned}$$

Nach den Additionstheoremen für \cos und \sin gilt

$$\begin{aligned} z_1 \cdot z_2 &= |z_1||z_2| \cdot \{(\cos \varphi_1 \cdot \cos \varphi_2 - \sin \varphi_1 \cdot \sin \varphi_2) \\ &\quad + i \cdot (\sin \varphi_1 \cdot \cos \varphi_2 + \sin \varphi_2 \cdot \cos \varphi_1)\} \\ &= |z_1 \cdot z_2| \cdot (\cos(\varphi_1 + \varphi_2) + i \cdot \sin(\varphi_1 + \varphi_2)) \end{aligned}$$

und folglich $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$ und $\arg(z_1 \cdot z_2) = \arg(z_1) + \arg(z_2)$.

Nach diesen Formeln kann $z_1 \cdot z_2$ gezeichnet werden.



Für die Winkel gilt $\varphi_1 = \arg(z_1)$ und $\varphi_2 = \arg(z_2)$. Der Punkt $z_1 \cdot z_2$ liegt auf dem vom Ursprung ausgehenden Strahl, der mit der reellen Achse \mathbb{R} den Winkel $\varphi_1 + \varphi_2$ einnimmt. Mittels des Strahlensatzes erhält man einen Punkt w auf dem Strahl durch den Ursprung und z_2 mit $|w| = |z_1||z_2| = |z_1 \cdot z_2|$. Man dreht diesen Punkt w um den Winkel φ_1 um den Ursprung und erhält den Punkt $z_1 \cdot z_2$.

Wurzeln aus komplexen Zahlen

Wir definieren nun Wurzeln aus von Null verschiedenen komplexen Zahlen. Wie wir gerade gesehen haben, gelten für eine komplexe Zahl $z \neq 0$ die Formeln

$$|z^n| = |z|^n \quad \text{und} \quad \arg(z^n) = n \cdot \arg(z).$$

Damit erhalten wir

Satz 2.15 Sei $w \in \mathbb{C}$ eine von Null verschiedene komplexe Zahl mit dem Betrag $r := |w|$ und dem Argument $\varphi := \arg(w) \in [0, 360^\circ)$, d.h. $w = r(\cos \varphi + i \sin \varphi)$. Dann hat die Gleichung $z^n = w$ genau n verschiedene komplexe Lösungen, nämlich

$$z_k := \sqrt[n]{r} \cdot \left(\cos \left(\frac{\varphi}{n} + k \frac{360^\circ}{n} \right) + i \sin \left(\frac{\varphi}{n} + k \frac{360^\circ}{n} \right) \right),$$

wobei $k \in \{0, 1, 2, \dots, n-1\}$.

Beweis. Für die komplexen Zahlen z_k gilt nach Definition

$$|z_k| = \sqrt[n]{r} \quad \text{und} \quad \arg(z_k) = \frac{\varphi}{n} + k \frac{360^\circ}{n} =: \varphi_k.$$

Hieraus folgt

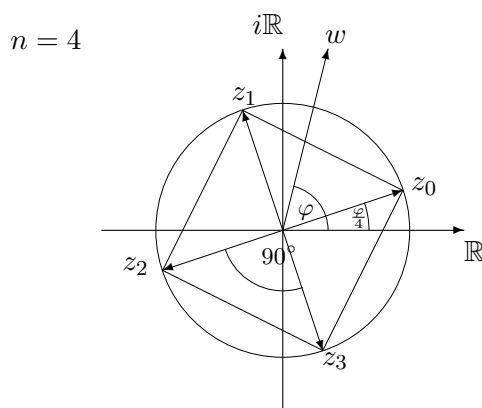
$$z_k^n = \underbrace{(\sqrt[n]{r})^n}_{|z_k|^n} \left(\underbrace{\cos(\varphi + k \cdot 360^\circ)}_{n \cdot \arg(z_k)} + i \sin(\varphi + k \cdot 360^\circ) \right) = |w|(\cos \varphi + i \sin \varphi) = w.$$

Wir haben also n verschiedene Lösungen der Gleichung $z^n = w$ gefunden. Wir zeigen, dass es keine weiteren Lösungen gibt. Sei z eine beliebige Lösung von $z^n = w$ und $z = |z|(\cos \psi + i \cdot \sin \psi)$ die trigonometrische Darstellung von z . Es gilt $|z|^n = |w|$ und folglich $|z| = \sqrt[n]{|w|}$. Weiterhin ist $n \cdot \psi = \varphi + 360^\circ \ell$, für ein $\ell \in \mathbb{Z}$ und somit $\psi = \frac{\varphi}{n} + \ell \frac{360^\circ}{n}$. Wir teilen ℓ durch n mit Rest: $\ell = k + rn$, $r \in \mathbb{Z}$ und $0 \leq k \leq n-1$. Dann gilt $\psi = \varphi_k + r \cdot 360^\circ$ und folglich $z = z_k$. \square

Die n verschiedenen Lösungen von $z^n = w$ mit $w \neq 0$, heißen *die n . Wurzeln von w .*

Geometrische Deutung der Wurzeln:

Die Lösungen z_k von $z^n = w$ bilden die Ecken eines regelmäßigen n -Ecks auf dem Kreis vom Radius $\sqrt[n]{|w|}$.



Die n verschiedenen Lösungen der Gleichung $z^n = 1$ heißen *n .-Einheitswurzeln*. Sie sind gegeben durch

$$z_k = \cos \left(k \frac{360^\circ}{n} \right) + i \sin \left(k \frac{360^\circ}{n} \right),$$

wobei $k \in \{0, 1, \dots, n-1\}$.

Wir formulieren abschließend den Fundamentalsatz der Algebra, der eine der wichtigsten Aussagen über komplexe Zahlen enthält.

Satz 2.16 (Fundamentalsatz der Algebra)

Es seien a_0, a_1, \dots, a_{n-1} komplexe Zahlen. Dann besitzt die Gleichung

$$P(z) := z^n + a_{n-1}z^{n-1} + \dots + a_1z + a_0 = 0$$

eine Lösung $\lambda \in \mathbb{C}$. Dabei gilt

$$P(z) = (z - \lambda)(z^{n-1} + b_{n-2}z^{n-2} + \dots + b_1z + b_0),$$

wobei b_0, \dots, b_{n-2} komplexe Zahlen sind. Ein komplexes Polynom $P(z)$ von Grad n hat also genau n Nullstellen $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ und es gilt

$$P(z) = (z - \lambda_1) \cdot (z - \lambda_2) \cdot \dots \cdot (z - \lambda_n),$$

wobei die Nullstellen nicht alle verschieden sein müssen.

Den Beweis dieses Satzes für allgemeine n werden wir später in Kapitel 4 führen. Wir beweisen den Fundamentalsatz der Algebra hier zunächst nur für $n = 2$. In diesem Fall erhält man alle komplexen Lösungen z der quadratischen Gleichung $z^2 + a_1z + a_0 = 0$ mittels quadratischer Ergänzung. Es gilt:

$$\begin{aligned} z^2 + a_1z + a_0 &= 0 \\ \iff \left(z + \frac{a_1}{2}\right)^2 + a_0 - \frac{a_1^2}{4} &= 0 \\ \iff v^2 = w \quad \text{wobei } v := z + \frac{a_1}{2} \quad \text{und} \quad w = \frac{a_1^2}{4} - a_0. \end{aligned}$$

1. Fall: $w \neq 0$. In diesem Fall können wir die Gleichung $v^2 = w$ nach Satz 2.15 lösen. Wir erhalten genau 2 verschiedene komplexe Lösungen $\pm v_0$ von $v^2 = w$. Daraus ergeben sich zwei komplexe Lösungen⁷ λ_1 und λ_2 von $z^2 + a_1z + a_0 = 0$ mit

$$\lambda_1 := -\frac{a_1}{2} + v_0 \quad \text{und} \quad \lambda_2 := -\frac{a_1}{2} - v_0, \quad \text{wobei} \quad v_0^2 = \frac{a_1^2}{4} - a_0.$$

Außerdem gilt:

$$\begin{aligned} (z - \lambda_1)(z - \lambda_2) &= \left(z - v_0 + \frac{a_1}{2}\right) \left(z + v_0 + \frac{a_1}{2}\right) = \left(z + \frac{a_1}{2}\right)^2 - v_0^2 \\ &= \left(z + \frac{a_1}{2}\right)^2 - w = \left(z + \frac{a_1}{2}\right)^2 - \frac{a_1^2}{4} + a_0 = z^2 + a_1z + a_0 = 0. \end{aligned}$$

2. Fall: $w = 0$. In diesem Fall gilt $a_0 = \frac{a_1^2}{4}$ und deshalb

$$z^2 + a_1z + a_0 = \left(z + \frac{a_1}{2}\right) \left(z + \frac{a_1}{2}\right).$$

Folglich ist $\lambda_1 = \lambda_2 = -\frac{a_1}{2}$ eine 2-fache komplexe Lösung von $z^2 + a_1z + a_0 = 0$. □

⁷ Vergleichen Sie diese Lösungen mit der Lösungsformel im reellen Fall!!

Folgen und Reihen

In diesem Kapitel beginnen wir mit der Behandlung von Grenzprozessen, die zu den wichtigsten Prinzipien der Analysis gehören. Viele Größen (spezielle Zahlen, Funktionswerte, Flächeninhalte, Volumen, ...) werden nicht durch eine geschlossene Formel angegeben, sondern als Grenzwert einer bestimmten Folge definiert und dann in der Regel durch Approximationsverfahren näherungsweise berechnet. Um dies zu tun, benötigt man den Begriff der Folge und ihres Grenzwertes.

Definition 3.1. *Sei X eine nichtleere Menge. Unter einer Folge in X versteht man eine Abbildung*

$$\begin{aligned} f : \mathbb{N} &\longrightarrow X \\ n &\longmapsto x_n, \end{aligned}$$

die jeder natürlichen Zahl n einen Punkt $x_n \in X$ zuordnet.

Eine Folge ist also eine durch die Abbildung f gegebene Aufzählung von Punkten in X , wobei Punkte auch mehrfach vorkommen können. Wir geben künftig lediglich die Bildwerte der Abbildung f an und benutzen für die Folge die nachstehenden Schreibweisen:

$$\boxed{x_1, x_2, x_3, \dots \quad \text{oder} \quad (x_n)_{n=1}^{\infty} \quad \text{oder} \quad (x_n)_{n \in \mathbb{N}} \quad \text{oder kurz} \quad (x_n).}$$

Um den Grenzwert einer Folge zu definieren, benötigt man ein Konzept, mit dem man den Abstand zwischen Punkten von X messen kann (wann sind sie weit von einander entfernt, wann liegen sie dicht beieinander?). Eines dieser Konzepte ist die Festlegung einer Abstandsfunktion d auf X , die den Abstand zwischen zwei Punkten mißt. Dies führt auf den Begriff des metrischen Raumes (X, d) , den wir im 2. Semester der Analysis-Vorlesung behandeln werden. In diesem Kapitel der Vorlesung werden wir zunächst den Spezialfall von Folgen reeller Zahlen und Folgen komplexer Zahlen behandeln ($X = \mathbb{R}$ bzw. $X = \mathbb{C}$). In diesem Fall haben wir durch den *Betrag* eine Abstandsfunktion gegeben:

Der Betrag $|x - y|$ beschreibt den Abstand der Zahlen x und y voneinander.

3.1 Folgen in \mathbb{R} und in \mathbb{C}

3.1.1 Konvergente Zahlenfolgen und ihre Eigenschaften

Definition 3.2. Sei (x_n) eine Folge reeller (bzw. komplexer) Zahlen. Wir sagen, dass (x_n) gegen die Zahl $x \in \mathbb{R}$ (bzw. $x \in \mathbb{C}$) konvergiert, falls zu jedem $\varepsilon > 0$ ein (von ε abhängender) Index $n_0 = n_0(\varepsilon) \in \mathbb{N}$ existiert, so dass $|x - x_n| < \varepsilon$ für alle $n \geq n_0$. Der Punkt x heißt Grenzwert (GW) der Folge (x_n) . Besitzt eine Folge (x_n) einen Grenzwert, so heißt sie konvergent. Besitzt die Folge (x_n) keinen Grenzwert, so heißt sie divergent.

Für eine gegen x konvergente Zahlenfolge (x_n) schreiben wir:

$$\boxed{\lim_{n \rightarrow \infty} x_n = x \quad \text{oder} \quad x_n \xrightarrow[n \rightarrow \infty]{} x \quad \text{oder kurz} \quad x_n \rightarrow x}.$$

Im Fall einer *komplexen* Zahlenfolge (x_n) bedeutet die Konvergenz $x_n \rightarrow x$, dass ab dem Index $n_0(\varepsilon)$ alle Folgenglieder x_n in der offenen Kreisscheibe $K_\varepsilon(x)$ um x mit dem Radius ε liegen:

$$x_n \in K_\varepsilon(x) := \{z \in \mathbb{C} \mid |x - z| < \varepsilon\} \quad \forall n \geq n_0(\varepsilon).$$

Im Fall einer *reellen* Zahlenfolge (x_n) bedeutet die Konvergenz $x_n \rightarrow x$, dass ab dem Index $n_0(\varepsilon)$ alle Folgenglieder x_n im offenen ε -Intervall $I_\varepsilon(x)$ um x liegen:

$$x_n \in I_\varepsilon(x) := (x - \varepsilon, x + \varepsilon) \subset \mathbb{R} \quad \forall n \geq n_0(\varepsilon).$$

Satz 3.1 (Eindeutigkeit des Grenzwertes) Der Grenzwert einer konvergenten Folge in \mathbb{R} und in \mathbb{C} ist eindeutig bestimmt.

Beweis. Sei (x_n) eine Folge in \mathbb{R} (bzw. in \mathbb{C}), die gegen die reellen (bzw. komplexen) Zahlen x und x^* konvergiert. Angenommen $x \neq x^*$. Dann ist $\varepsilon := |x - x^*| > 0$. Also existieren n_0 und n_0^* mit

$$|x - x_n| < \frac{\varepsilon}{2} \quad \forall n \geq n_0 \quad \text{und} \quad |x^* - x_n| < \frac{\varepsilon}{2} \quad \forall n \geq n_0^*.$$

Somit gilt $|x - x_n| < \frac{\varepsilon}{2}$ und $|x^* - x_n| < \frac{\varepsilon}{2}$ für alle $n \geq \max(n_0, n_0^*)$. Mit der Dreiecksungleichung für die Beträge (Satz 2.1 und Satz 2.14) folgt für ein solches n

$$\varepsilon = |x - x^*| \leq |x - x_n| + |x_n - x^*| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Dies ist aber ein Widerspruch. □

Satz 3.2 Ist (x_n) eine konvergente Folge reeller oder komplexer Zahlen, so ist die Menge $\{|x_n| \mid n \in \mathbb{N}\} \subset \mathbb{R}$ beschränkt.

Beweis. Sei (x_n) eine konvergente Folge in \mathbb{R} (bzw. in \mathbb{C}) und $x = \lim_{n \rightarrow \infty} x_n$. Nach Definition der Konvergenz existiert ein $n_0 \in \mathbb{N}$, so dass $|x_n - x| < 1$ für alle $n \geq n_0$. Aus der Dreiecksungleichung erhalten wir

$$|x_n| = |(x_n - x) + x| \leq |x_n - x| + |x| < 1 + |x| \quad \forall n \geq n_0.$$

Dann ist $M := \max(|x_1|, |x_2|, \dots, |x_{n_0-1}|, 1 + |x|)$ eine obere Schranke und $m = 0$ eine untere Schranke von $\{|x_n| \mid n \in \mathbb{N}\}$. \square

Ist (x_n) eine reelle Folge, dann liegen alle Folgenglieder im Intervall $(-M, M) \subset \mathbb{R}$. Ist (x_n) eine komplexe Folge, dann liegen alle Folgenglieder in der Kugel $K_M(0) \subset \mathbb{C}$.

Der nächste Satz zeigt, dass der Grenzwertbegriff mit der algebraischen Struktur von \mathbb{R} bzw. \mathbb{C} verträglich ist.

Satz 3.3 *Seien (z_n) und (w_n) konvergente Folgen reeller oder komplexer Zahlen mit den Grenzwerten $\lim_{n \rightarrow \infty} z_n = z$ und $\lim_{n \rightarrow \infty} w_n = w$. Dann konvergieren auch die Folgen $(z_n + w_n)$, $(z_n \cdot w_n)$, $(\frac{z_n}{w_n})$ falls $w \neq 0$, und $(|z_n|)$ und es gilt:*

a) $\lim_{n \rightarrow \infty} (z_n + w_n) = \lim_{n \rightarrow \infty} z_n + \lim_{n \rightarrow \infty} w_n = z + w$

b) $\lim_{n \rightarrow \infty} z_n \cdot w_n = \lim_{n \rightarrow \infty} z_n \cdot \lim_{n \rightarrow \infty} w_n = z \cdot w$.

c) *Ist $w \neq 0$, so ist auch $w_n \neq 0$ für alle n größer als ein $n_0 \in \mathbb{N}$, und es gilt*

$$\lim_{n \rightarrow \infty} \frac{z_n}{w_n} = \frac{\lim_{n \rightarrow \infty} z_n}{\lim_{n \rightarrow \infty} w_n} = \frac{z}{w}.$$

d) $\lim_{n \rightarrow \infty} |z_n| = |\lim_{n \rightarrow \infty} z_n| = |z|$.

Beweis. Zu a) Nach Dreiecksungleichung gilt

$$|(z_n + w_n) - (z + w)| = |(z_n - z) + (w_n - w)| \leq |z_n - z| + |w_n - w|.$$

Sei $\varepsilon > 0$. Da (z_n) gegen z und (w_n) gegen w konvergiert, existieren Indizes $n_0, n_0^* \in \mathbb{N}$, so dass

$$|z_n - z| < \frac{\varepsilon}{2} \quad \forall n \geq n_0 \quad \text{und} \quad |w_n - w| < \frac{\varepsilon}{2} \quad \forall n \geq n_0^*.$$

Daraus folgt

$$|(z_n + w_n) - (z + w)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \quad \forall n \geq \max(n_0, n_0^*)$$

und somit $\lim_{n \rightarrow \infty} (z_n + w_n) = (z + w)$.

Zu b) Nach Dreiecksungleichung gilt

$$|z_n w_n - z w| = |z_n(w_n - w) + w(z_n - z)| \leq |z_n| |w_n - w| + |w| |z_n - z|.$$

Da (z_n) konvergiert, existiert nach Satz 3.2 eine positive reelle Zahl M , so dass $|z_n| \leq M$ für alle $n \in \mathbb{N}$. Wir wählen M außerdem so groß, dass $M \geq |w|$. Dann folgt

$$|z_n w_n - zw| \leq M \cdot |w_n - w| + M \cdot |z_n - z| \quad \forall n \in \mathbb{N}.$$

Sei nun $\varepsilon > 0$. Da (z_n) gegen z und (w_n) gegen w konvergiert, existieren Indizes $n_0, n_0^* \in \mathbb{N}$, so dass

$$|z_n - z| < \frac{\varepsilon}{2M} \quad \forall n \geq n_0 \quad \text{und} \quad |w_n - w| < \frac{\varepsilon}{2M} \quad \forall n \geq n_0^*.$$

Folglich gilt

$$|z_n w_n - zw| < \varepsilon \quad \forall n \geq \max(n_0, n_0^*),$$

d.h. die Folge $(z_n w_n)$ konvergiert gegen zw .

Zu c) Da $w \neq 0$ und (w_n) gegen w konvergiert, existiert eine positive reelle Zahl η so dass $0 < \eta < |w_n|$ für alle n größer als ein n_0 . Man erhält

$$\begin{aligned} \left| \frac{z_n}{w_n} - \frac{z}{w} \right| &= \left| \frac{z_n \cdot w - w_n \cdot z}{w_n \cdot w} \right| \\ &= \frac{|(z_n - z)w - (w_n - w)z|}{|w_n| \cdot |w|} \\ &\leq \frac{|z_n - z||w| + |w_n - w||z|}{|w_n| \cdot |w|} \\ &\leq |z_n - z| \cdot \frac{|w|}{\eta|w|} + |w_n - w| \cdot \frac{|z|}{\eta|w|}. \end{aligned}$$

Sei nun $\varepsilon > 0$. Da (z_n) gegen z und (w_n) gegen w konvergiert, existiert ein Index $n_0^* \geq n_0$, so dass $|z - z_n| \cdot \frac{1}{\eta} < \frac{\varepsilon}{2}$ und $|w - w_n| \cdot \frac{|z|}{\eta|w|} < \frac{\varepsilon}{2}$ für alle $n \geq n_0^*$. Daraus folgt

$$\left| \frac{z_n}{w_n} - \frac{z}{w} \right| < \varepsilon \quad \forall n \geq n_0^*.$$

Also konvergiert die Folge $(\frac{z_n}{w_n})$ gegen $\frac{z}{w}$.

Zu d) Sei $z_n \rightarrow z$ und $\varepsilon > 0$. Aus den Betragseigenschaften folgt

$$||z_n| - |z|| \leq |z_n - z| < \varepsilon \quad \forall n \geq n_0(\varepsilon).$$

Also konvergiert $(|z_n|)$ gegen $|z|$. □

Der nächste Satz führt die Konvergenz von komplexen Folgen auf die Konvergenz von reellen Folgen zurück.

Satz 3.4 *Es sei (z_n) eine Folge komplexer Zahlen und $z \in \mathbb{C}$. Dann sind folgende Aussagen äquivalent:*

- a) (z_n) konvergiert gegen z .
- b) Die Folge der Realteile $(\operatorname{Re}(z_n))$ konvergiert gegen $\operatorname{Re}(z)$ und die Folge der Imaginärteile $(\operatorname{Im}(z_n))$ konvergiert gegen $\operatorname{Im}(z)$.
- c) Die Folge der konjugiert komplexen Zahlen (\bar{z}_n) konvergiert gegen \bar{z} .

Beweis. a) \iff b): Sei $z_n = a_n + ib_n$ und $z = a + ib$. Dann gilt

$$|z - z_n| = \sqrt{(a - a_n)^2 + (b - b_n)^2}. \quad (3.1)$$

Sei $\varepsilon > 0$.

a) \implies b): Konvergiert (z_n) gegen z , so existiert ein $n_0 \in \mathbb{N}$ mit $|z - z_n| < \varepsilon$ für alle $n \geq n_0$. Dann gilt wegen (3.1)

$$|a - a_n| \leq |z - z_n| < \varepsilon \quad \text{und} \quad |b - b_n| \leq |z - z_n| < \varepsilon \quad \forall n \geq n_0.$$

Folglich konvergiert (a_n) gegen a und (b_n) gegen b .

b) \implies a): Konvergiert (a_n) gegen a und (b_n) gegen b , so existiert ein $n_0 \in \mathbb{N}$ so dass

$$|a - a_n| < \frac{\varepsilon}{\sqrt{2}} \quad \text{und} \quad |b - b_n| < \frac{\varepsilon}{\sqrt{2}} \quad \forall n \geq n_0.$$

Dann folgt mit (3.1)

$$|z - z_n| < \varepsilon,$$

d.h. (z_n) konvergiert gegen z .

a) \iff c): Dies folgt aus $|z - z_n| = |\bar{z} - \bar{z}_n|$. □

In den *reellen* Zahlen haben wir außer den algebraischen Rechenregeln (Körpereigenschaften) auch die Anordnungseigenschaft ($x \leq y$). Auch diese Eigenschaft ist mit der Grenzwertbildung verträglich.

Satz 3.5 Seien (x_n) und (y_n) zwei konvergente Folgen reeller Zahlen mit den Grenzwerten $\lim_{n \rightarrow \infty} x_n = x$ und $\lim_{n \rightarrow \infty} y_n = y$.

a) Gilt $x_n \leq y_n$ für alle $n \geq n^*$, so folgt $x \leq y$.

b) Sei (u_n) eine weitere Folge reeller Zahlen mit $x_n \leq u_n \leq y_n$ für alle $n \geq n^*$ und sei $x = y$. Dann ist die Folge (u_n) ebenfalls konvergent und es gilt $\lim_{n \rightarrow \infty} u_n = x$.

[Sandwich-Lemma].

Beweis. Zu a) Angenommen, es wäre $x > y$. Wir setzen $\varepsilon = x - y > 0$. Dann existiert ein $n_0 \in \mathbb{N}$, so dass

$$|x_n - x| < \frac{\varepsilon}{2} \quad \text{und} \quad |y_n - y| < \frac{\varepsilon}{2} \quad \forall n \geq n_0.$$

Deshalb ist $x_n > x - \frac{\varepsilon}{2} = y + \frac{\varepsilon}{2} > y_n$ für alle $n \geq n_0$. Dies steht aber im Widerspruch zur Voraussetzung und somit war die Annahme $x > y$ falsch.

Zu b) Sei $x_n \leq u_n \leq y_n$ für alle $n \geq n^*$. Da $x_n \rightarrow x$ und $y_n \rightarrow x$, existiert für jedes $\varepsilon > 0$ ein n_0 so dass $|x_n - x| < \frac{\varepsilon}{4}$ und $|y_n - x| < \frac{\varepsilon}{4}$ für alle $n \geq n_0$. Aus der Dreiecks-Ungleichung folgt:

$$\begin{aligned} |x - u_n| &\leq |x - x_n| + |x_n - u_n| \\ &\leq |x - x_n| + |x_n - y_n| \\ &\leq |x - x_n| + |x_n - x| + |x - y_n| \\ &< \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < \varepsilon \quad \forall n \geq \max\{n^*, n_0\}. \end{aligned}$$

Folglich konvergiert die Folge (u_n) gegen x . □

Definition 3.3. Eine Folge (z_n) von reellen oder komplexen Zahlen, die gegen 0 konvergiert, heißt Nullfolge.

Aus den Sätzen 3.3 und Satz 3.5 folgt dann sofort:

- Eine Folge reeller oder komplexer Zahlen (z_n) ist genau dann eine Nullfolge, wenn die Folge der Beträge $(|z_n|)$ eine Nullfolge ist.
- Sind (z_n) und (w_n) Nullfolgen, so ist auch $(z_n + w_n)$ eine Nullfolge.
- Ist (z_n) eine Nullfolge und (w_n) eine konvergente Folge in \mathbb{R} oder in \mathbb{C} , so ist die Folge $(z_n w_n)$ ebenfalls eine Nullfolge.
- Ist (y_n) eine Nullfolge reeller Zahlen und (x_n) eine weitere Folge reeller Zahlen mit $0 \leq x_n \leq y_n$ für fast alle $n \in \mathbb{N}$, dann ist (x_n) ebenfalls eine Nullfolge.

Wichtige Beispiele konvergenter Folgen in \mathbb{R} bzw. \mathbb{C} :

1. Sei q eine positive rationale Zahl. Dann gilt $\boxed{\lim_{n \rightarrow \infty} \left(\frac{1}{n}\right)^q = 0.}$

Nach dem Archimedischen Axiom existiert zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass $\frac{1}{n_0} \leq \varepsilon^{\frac{1}{q}}$. Folglich ist $\left|\left(\frac{1}{n}\right)^q\right| < \varepsilon$ für alle $n \geq n_0$, woraus die Behauptung folgt.

2. Sei x eine positive reelle Zahl. Dann gilt $\boxed{\lim_{n \rightarrow \infty} \sqrt[n]{x} = 1.}$

Um dies einzusehen, betrachten wir zunächst $x > 1$. Sei $x_n := \sqrt[n]{x} - 1$. Dann gilt $x_n > 0$ und aus der Bernoullischen Ungleichung erhalten wir

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

Also gilt $0 < x_n \leq \frac{x-1}{n}$. Da $\left(\frac{x-1}{n}\right)$ eine Nullfolge ist, konvergiert (x_n) ebenfalls gegen Null, und somit $(\sqrt[n]{x})$ gegen 1.

Ist $0 < x < 1$, so folgt $\frac{1}{x} > 1$ und wir erhalten mittels Satz 3.3 und dem gerade Bewiesenen $\lim_{n \rightarrow \infty} \sqrt[n]{x} = \lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{\frac{1}{x}}} = 1$.

3. Es gilt $\boxed{\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1.}$

Zum Beweis betrachten wir die Folge $x_n := \sqrt[n]{n} - 1$ für $n \geq 2$. Es gilt $x_n > 0$. Aus der binomischen Formel folgt

$$n = (1 + x_n)^n > \binom{n}{2} x_n^2 = \frac{1}{2} n(n-1) x_n^2.$$

Folglich gilt $0 < x_n < \sqrt{\frac{2}{n-1}}$. Nach 1. ist $\left(\sqrt{\frac{2}{n-1}}\right)$ eine Nullfolge, somit ist (x_n) ebenfalls eine Nullfolge und folglich gilt $\lim_{n \rightarrow \infty} \sqrt[n]{n} = 1$.

4. Sei $z \in \mathbb{C}$ mit $|z| < 1$. Dann gilt $\boxed{\lim_{n \rightarrow \infty} z^n = 0.}$

Dies sieht man folgendermaßen: Da $0 \leq |z| < 1$, existiert für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ mit

$|z|^{n_0} < \varepsilon$. Da $|z| < 1$ ist, erhält man $|z|^n < |z|^{n_0}$ für alle $n \geq n_0$, das heißt $|z|^n = |z^n| < \varepsilon$ für alle $n \geq n_0$. Folglich gilt $\lim_{n \rightarrow \infty} z^n = 0$.

5. Sei $z \in \mathbb{C}$ mit $|z| > 1$ und $k \in \mathbb{N}$ eine fixierte natürliche Zahl. Dann gilt $\lim_{n \rightarrow \infty} \frac{n^k}{z^n} = 0$.

Dies bedeutet, dass für $|z| > 1$ die Folge der Potenzen $|z|^n$ schneller wächst als jede noch so große Potenz von n . Zum Beweis setzen wir $x := |z| - 1$ und wählen eine natürliche Zahl $p > k$. Für jedes $n > 2p$ folgt aus der binomischen Formel

$$(1+x)^n > \binom{n}{p} \cdot x^p = \frac{\overbrace{n(n-1) \cdots (n-(p-1))}^{p\text{-Faktoren}}}{p!} \cdot x^p.$$

Da $p < \frac{n}{2}$, ist jeder der Faktoren $n, (n-1), \dots, (n-(p-1))$ größer als $\frac{n}{2}$. Es folgt $(1+x)^n > (\frac{n}{2})^p \cdot \frac{x^p}{p!}$ und somit

$$0 < \frac{n^k}{|z|^n} < \frac{2^p \cdot p!}{x^p \cdot n^{p-k}} \leq \underbrace{\frac{2^p \cdot p!}{x^p}}_{\text{konstant}} \cdot \frac{1}{n}.$$

Auf der rechten Seite steht eine Nullfolge, also ist $\lim_{n \rightarrow \infty} \frac{n^k}{z^n} = 0$.

Definition 3.4. Sei (x_n) eine Folge reeller oder komplexer Zahlen. Unter einer Teilfolge von (x_n) verstehen wir eine unendliche Auswahl von Elementen dieser Folge, d.h. eine Folge $(x_{n_j})_{j=1}^{\infty}$, wobei $\{n_1, n_2, n_3, \dots\}$ eine Teilmenge von \mathbb{N} mit $n_1 < n_2 < n_3 < \dots$ ist.

Offensichtlich gilt:

Satz 3.6 Ist (x_n) eine gegen x konvergente Folge reeller oder komplexer Zahlen. Dann konvergiert jede Teilfolge von (x_n) ebenfalls gegen x .

Definition 3.5. Sei (x_n) eine Folge reeller (bzw. komplexer) Zahlen. Ein Punkt $x \in \mathbb{R}$ (bzw. $x \in \mathbb{C}$) heißt Häufungspunkt der Folge (x_n) , wenn es eine Teilfolge $(x_{n_j})_{j=1}^{\infty}$ von (x_n) gibt mit $\lim_{j \rightarrow \infty} x_{n_j} = x$.

Die Menge der Häufungspunkte von (x_n) bezeichnen wir mit $HP(x_n)$.

Beispiele:

1. Sei (x_n) eine gegen x konvergente Folge reeller oder komplexer Zahlen. Dann gilt nach Satz 3.6, dass $HP(x_n) = \{x\}$.
2. Wir betrachten die reelle Folge (x_n) mit

$$x_n := \begin{cases} \frac{1}{n} & \text{falls } n \text{ gerade} \\ 1 & \text{falls } n \text{ ungerade} \end{cases}$$

Dann gilt $HP(x_n) = \{0, 1\}$.

3. Es sei (z_n) die komplexe Folge mit $z_n := \cos(n \cdot 45^\circ) + i \sin(n \cdot 45^\circ)$. Dann gilt $z_n^8 = 1$. Für $n = 1, 2, \dots, 8$ erhalten wir die acht verschiedenen 8. Einheitswurzeln. Außerdem gilt $z_{n+8} = z_n$. Folglich gilt: $HP(z_n) = \{z_1, z_2, \dots, z_8\}$.

3.1.2 Spezielle Eigenschaften reeller Folgen

In diesem Abschnitt betrachten wir weitere, spezielle Eigenschaften von Folgen *reeller* Zahlen. Wir nutzen dabei aus, dass \mathbb{R} ein vollständiger angeordneter Körper ist.

Um später Formulierungen vereinheitlichen zu können, betrachten wir zunächst eine spezielle Sorte von *divergenten* Folgen reeller Zahlen und ordnen diesen den Grenzwert $+\infty$ oder $-\infty$ zu.

Definition 3.6. Sei (x_n) eine Folge reeller Zahlen.

Wir sagen, dass (x_n) gegen $+\infty$ strebt, falls zu jedem $M \in \mathbb{R}$ ein $n_0 \in \mathbb{N}$ existiert mit $x_n \geq M$ für alle $n \geq n_0$.

Wir sagen, dass (x_n) gegen $-\infty$ strebt, falls zu jedem $M \in \mathbb{R}$ ein $n_0 \in \mathbb{N}$ existiert mit $x_n \leq M$ für alle $n \geq n_0$.

Für Folgen reeller Zahlen, die gegen $+\infty$ bzw. $-\infty$ streben, benutzen wir die Schreibweise

$$\lim_{n \rightarrow \infty} x_n = +\infty \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} x_n = -\infty.$$

Man nennt diese Sorte divergenter Folgen reeller Zahlen oft auch *bestimmt divergent* oder *uneigentlich konvergent* (je nach Autor des benutzten Buches) und $\pm\infty$ den uneigentlichen Grenzwert.

Beispiele:

- Ist $x_n = n$ oder $x_n = n^2$, so gilt $\lim_{n \rightarrow \infty} x_n = +\infty$.
- Die Folge $1, 0, 2, 0, 3, 0, 4, 0, \dots$ ist in \mathbb{R} divergent, strebt aber *nicht* gegen $+\infty$, obwohl sie beliebig große Glieder enthält.
- Aus $x_n \rightarrow +\infty$ und $y_n \rightarrow +\infty$ folgt $x_n + y_n \rightarrow +\infty$ und $x_n \cdot y_n \rightarrow +\infty$.
- Aus $x_n \rightarrow +\infty$ und $y_n \rightarrow a$ für $a > 0$ folgt $x_n + y_n \rightarrow +\infty$ und $x_n \cdot y_n \rightarrow +\infty$.
- Wenn $x_n \rightarrow +\infty$, so $\frac{1}{x_n} \rightarrow 0$.
- Wenn $x_n \rightarrow 0$ und $x_n > 0$, so $\frac{1}{x_n} \rightarrow +\infty$.

Aus $x_n \rightarrow +\infty$ und $y_n \rightarrow 0$ kann man i.a. nichts über das Verhalten von $x_n \cdot y_n$ folgern, wie die folgenden Beispiele zeigen.

- Sei $x_n = n^2$ und $y_n = \frac{1}{n}$, so gilt $x_n \cdot y_n = n \rightarrow +\infty$.
- Sei $x_n = n^2$ und $y_n = \frac{1}{n^2}$, so gilt $x_n \cdot y_n = 1 \rightarrow 1$.
- Sei $x_n = n^2$ und $y_n = \frac{1}{n^3}$, so gilt $x_n \cdot y_n = \frac{1}{n} \rightarrow 0$.
- Für $x_n = n$ und $y_n = \begin{cases} \frac{1}{n} & \text{falls } n \text{ gerade} \\ \frac{1}{2n} & \text{falls } n \text{ ungerade} \end{cases}$ gilt $x_n \cdot y_n = \begin{cases} 1 & \text{falls } n \text{ gerade} \\ \frac{1}{2} & \text{falls } n \text{ ungerade} \end{cases}$.

Folglich konvergiert $x_n \cdot y_n$ überhaupt nicht.

Definition 3.7. Eine Folge reeller Zahlen (x_n) heißt *beschränkt* (von oben beschränkt bzw. von unten beschränkt), wenn die Menge der Folgenglieder $\{x_n \mid n \in \mathbb{N}\} \subset \mathbb{R}$ beschränkt (von oben beschränkt bzw. von unten beschränkt) ist.

Für eine beschränkte Folge (x_n) gibt es eine Zahl $M \in \mathbb{R}^+$ mit $|x_n| \leq M$ für alle $n \in \mathbb{N}$. Aus Satz 3.2 wissen wir, dass jede *konvergente* Folge reeller Zahlen (x_n) beschränkt ist. Die Umkehrung gilt nicht. Aus dem Vollständigkeitsaxiom der reellen Zahlen erhalten wir aber die folgende Eigenschaft beschränkter Folgen¹.

Satz 3.7 (Satz von Bolzano/Weierstraß, 1. Teil)

Jede beschränkte Folge reeller Zahlen besitzt eine konvergente Teilfolge, d.h. einen Häufungspunkt.

Beweis. Sei (x_n) eine beschränkte Folge reeller Zahlen. Dann existiert ein $M > 0$ mit $-M < x_n < M$ für alle $n \in \mathbb{N}$. Wir betrachten die folgende Menge $A \subset \mathbb{R}$:

$$A := \{x \in \mathbb{R} \mid x \leq x_n \text{ für unendlich viele } n\}.$$

1. A ist nicht leer, da $-M \in A$.
2. A ist nach oben beschränkt, da zum Beispiel M eine obere Schranke ist.

Nach dem Satz 2.2 existiert ein Supremum $g = \sup A$ der Menge A . Wir zeigen nun, dass g ein Häufungspunkt von (x_n) ist. Zunächst zeigen wir, dass für jedes $\varepsilon > 0$ im Intervall $(g - \varepsilon, g + \varepsilon)$ unendlich viele Folgenglieder liegen. Sei also $\varepsilon > 0$. Aus der Definition des Supremums erhalten wir

- a) $g + \varepsilon \notin A$. D.h. für höchstens endlich viele x_n gilt die Ungleichung $g + \varepsilon \leq x_n$.
- b) Es existiert ein $x \in A$ mit $g - \varepsilon < x$. Somit sind unendlich viele Folgenglieder x_n größer oder gleich x , also größer als $g - \varepsilon$.

Insgesamt folgt also, dass unendlich viele Glieder der Folge (x_n) im Intervall $(g - \varepsilon, g + \varepsilon)$ liegen. Wir konstruieren jetzt eine Teilfolge (x_{n_k}) von (x_n) auf folgende Weise.

$$\begin{aligned} \varepsilon = 1 &\implies \exists x_{n_1} : g - 1 < x_{n_1} < g + 1 \\ \varepsilon = \frac{1}{2} &\implies \exists x_{n_2} : g - \frac{1}{2} < x_{n_2} < g + \frac{1}{2}, \quad n_2 > n_1, \\ &\vdots \\ \varepsilon = \frac{1}{k} &\implies \exists x_{n_k} : g - \frac{1}{k} < x_{n_k} < g + \frac{1}{k}, \quad n_k > n_{k-1} \dots \end{aligned}$$

Damit haben wir eine Teilfolge (x_{n_k}) von (x_n) gefunden mit $|x_{n_k} - g| < \frac{1}{k}$ für alle $k \in \mathbb{N}$. Somit ist $\lim_{k \rightarrow \infty} x_{n_k} = g$ und g ein Häufungspunkt von (x_n) . \square

Definition 3.8. Eine Folge reeller Zahlen (x_n) heißt

- *monoton wachsend*, falls $x_1 \leq x_2 \leq x_3 \leq \dots$
- *monoton fallend*, falls $x_1 \geq x_2 \geq x_3 \geq \dots$
- *monoton*, wenn sie *monoton wachsend* oder *monoton fallend* ist.

¹ Karl Weierstraß (1815-1897), Bernhard Bolzano (1781-1848).

Satz 3.8 (Satz von Bolzano/Weierstraß, 2. Teil)

Jede monoton wachsende, nach oben beschränkte Folge reeller Zahlen (x_n) konvergiert, und zwar gegen $\sup \{x_n \mid n \in \mathbb{N}\}$.

Jede monoton fallende, nach unten beschränkte Folge reeller Zahlen (x_n) konvergiert, und zwar gegen $\inf \{x_n \mid n \in \mathbb{N}\}$.

Beweis. Sei (x_n) monoton wachsend und nach oben beschränkt. Nach Satz 2.2 existiert das Supremum $g = \sup \{x_n \mid n \in \mathbb{N}\}$. Wir zeigen, dass (x_n) gegen g konvergiert. Sei $\varepsilon > 0$. Nach Definition des Supremums existiert ein $m_0 \in \mathbb{N}$ mit $x_{m_0} > g - \varepsilon$, also mit $x_{m_0} \in (g - \varepsilon, g]$. Da (x_n) monoton wachsend ist, gilt $g - \varepsilon < x_{m_0} \leq x_n \leq g$ für alle $n \geq m_0$. Folglich konvergiert (x_n) gegen g .

Den Beweis für monoton fallende, nach unten beschränkte Folgen führt man analog. \square

Sei nun (x_n) eine beschränkte Folge reeller Zahlen. Dann ist die Menge der Häufungspunkte $HP(x_n)$ nicht leer (Satz 3.7) und ebenfalls beschränkt (Übungsaufgabe). Folglich existiert das Supremum und das Infimum² von $HP(x_n)$.

Definition 3.9. Sei (x_n) eine beschränkte Folge reeller Zahlen.

Das Supremum der Menge der Häufungspunkte von (x_n) nennt man auch *limes superior* von (x_n) und bezeichnet es mit

$$\limsup_{n \rightarrow \infty} x_n := \overline{\lim}_{n \rightarrow \infty} x_n := \sup HP(x_n).$$

Das Infimum der Menge der Häufungspunkte von (x_n) nennt man auch *limes inferior* von (x_n) und bezeichnet es mit

$$\liminf_{n \rightarrow \infty} x_n := \underline{\lim}_{n \rightarrow \infty} x_n := \inf HP(x_n).$$

Ist (x_n) nicht nach oben beschränkt, so setzen wir $\limsup_{n \rightarrow \infty} x_n := +\infty$.

Ist (x_n) nicht nach unten beschränkt, so setzen wir $\liminf_{n \rightarrow \infty} x_n := -\infty$.

Satz 3.9 Sei (x_n) eine beschränkte Folge reeller Zahlen. Dann gilt

$$\lim_{n \rightarrow \infty} x_n = g \iff \limsup_{n \rightarrow \infty} x_n = \liminf_{n \rightarrow \infty} x_n = g.$$

Beweis. (\implies) Sei (x_n) gegen $g \in \mathbb{R}$ konvergent. Dann gilt $HP(x_n) = \{g\}$ und somit $\liminf_{n \rightarrow \infty} x_n = \limsup_{n \rightarrow \infty} x_n = g$.

(\impliedby) Sei umgekehrt $\limsup_{n \rightarrow \infty} x_n = \liminf_{n \rightarrow \infty} x_n = g$. Dann gilt $HP(x_n) = \{g\}$, das heißt jede konvergente Teilfolge von (x_n) konvergiert gegen g . Angenommen (x_n) würde nicht gegen g konvergieren. Dann existiert ein $\varepsilon > 0$, so dass gilt

² Man kann auch zeigen, dass das Supremum und das Infimum in der Menge der Häufungspunkte liegen, d.h. dass es einen größten und einen kleinsten Häufungspunkt gibt.

$$\forall n_0 \exists n \geq n_0 \text{ mit } |x_n - g| \geq \varepsilon.$$

Wir konstruieren jetzt eine Teilfolge von (x_n) folgendermaßen:

$$\begin{aligned} n_0 = 1 &\implies \exists n_1 \geq 1 : |x_{n_1} - g| \geq \varepsilon, \\ n_0 = n_1 + 1 &\implies \exists n_2 > n_1 : |x_{n_2} - g| \geq \varepsilon, \dots, \\ n_0 = n_{k-1} + 1 &\implies \exists n_k > n_{k-1} : |x_{n_k} - g| \geq \varepsilon, \dots \end{aligned}$$

Dadurch erhalten wir eine Teilfolge (x_{n_k}) von (x_n) mit $|x_{n_k} - g| \geq \varepsilon$, die ebenfalls beschränkt ist. Nach Satz von Bolzano–Weierstraß enthält sie eine konvergente Teilfolge $(x_{n_{k_i}})$, deren Grenzwert g^* nach Konstruktion von g verschieden ist. Dies ist ein Widerspruch zu unserer Voraussetzung. \square

Anwendung: Die Eulerzahl e

Satz 3.10 Die Folge der reellen Zahlen (a_n) mit

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

ist in \mathbb{R} konvergent.

Der Grenzwert der Folge (a_n) heißt Eulerzahl e .

Beweis. Wir zeigen, dass (a_n) eine monoton wachsende, nach oben beschränkte Folge ist. Nach Satz 3.8 existiert dann ein Grenzwert für (a_n) .

1. *Beschränktheit von (a_n) :* Aus der binomischen Formel folgt

$$a_n = \left(1 + \frac{1}{n}\right)^n = \sum_{k=0}^n \binom{n}{k} \cdot \left(\frac{1}{n}\right)^k.$$

Wir schätzen den Term $\binom{n}{k} \cdot \left(\frac{1}{n}\right)^k$ für $1 \leq k \leq n$ ab:

$$\begin{aligned} \binom{n}{k} \cdot \left(\frac{1}{n}\right)^k &= \frac{\overbrace{n(n-1) \cdot \dots \cdot (n-(k-1))}^{k\text{-Faktoren}}}{k! \cdot \underbrace{n \cdot n \cdot \dots \cdot n}_{k\text{-mal}}} \\ &= \frac{1}{k!} \cdot 1 \cdot \underbrace{\left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdot \dots \cdot \left(1 - \frac{(k-1)}{n}\right)}_{\leq 1} \quad (*) \\ &\leq \frac{1}{k!}. \end{aligned}$$

Folglich gilt für alle $n \in \mathbb{N}$

$$\begin{aligned}
a_n &= \left(1 + \frac{1}{n}\right)^n \\
&\leq 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!} \\
&= 1 + \frac{1}{1} + \frac{1}{2} + \frac{1}{2 \cdot 3} + \frac{1}{2 \cdot 3 \cdot 4} + \dots + \frac{1}{2 \cdot 3 \cdot \dots \cdot n} \\
&\leq 1 + \frac{1}{2^0} + \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \dots + \frac{1}{2^{n-1}} \\
&= 1 + \sum_{k=0}^{n-1} \left(\frac{1}{2}\right)^k \\
&= 1 + \frac{1 - \left(\frac{1}{2}\right)^n}{1 - \frac{1}{2}} \quad (\text{geometrische Summe}) \\
&< 3.
\end{aligned}$$

2. *Monotonie von (a_n)* : Gleichung (*) zeigt, dass

$$\binom{n}{k} \cdot \left(\frac{1}{n}\right)^k < \binom{n+1}{k} \cdot \left(\frac{1}{n+1}\right)^k.$$

Somit gilt $a_n < a_{n+1}$ für alle $n \in \mathbb{N}$. □

Satz 3.11 Die Folge der reellen Zahlen (b_n) mit

$$b_n := 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!} = \sum_{k=0}^n \frac{1}{k!}$$

konvergiert in \mathbb{R} und es gilt

$$\lim_{n \rightarrow \infty} b_n = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e.$$

Beweis. Aus dem Beweis von Satz 3.10 folgt $a_n \leq b_n < 3$ für alle $n \in \mathbb{N}$. Folglich ist (b_n) eine nach oben beschränkte, monoton wachsende Folge. Nach Satz 3.8 existiert deshalb ein Grenzwert von (b_n) und es gilt

$$e = \lim_{n \rightarrow \infty} a_n \leq \lim_{n \rightarrow \infty} b_n.$$

Andererseits gilt für $m \leq n$ wegen Formel (*) aus dem Beweis von Satz 3.10

$$\begin{aligned}
a_n &= \left(1 + \frac{1}{n}\right)^n = \sum_{k=0}^n \binom{n}{k} \left(\frac{1}{n}\right)^k \\
&= 1 + 1 + \frac{(1 - \frac{1}{n})}{2!} + \frac{(1 - \frac{1}{n})(1 - \frac{2}{n})}{3!} + \dots + \frac{(1 - \frac{1}{n}) \cdot \dots \cdot (1 - \frac{n-1}{n})}{n!} \\
&\geq 1 + 1 + \frac{(1 - \frac{1}{n})}{2!} + \dots + \frac{(1 - \frac{1}{n}) \cdot \dots \cdot (1 - \frac{m-1}{n})}{m!}.
\end{aligned}$$

Wir halten m fest und gehen in dieser Ungleichung mit n gegen $+\infty$. Dann folgt

$$e = \lim_{n \rightarrow \infty} a_n \geq 1 + 1 + \frac{1}{2!} + \dots + \frac{1}{m!} = b_m \quad \forall m \in \mathbb{N}.$$

Deshalb gilt $\lim_{m \rightarrow \infty} b_m \leq e$ und wir erhalten zusammenfassend $e = \lim_{n \rightarrow \infty} b_n$. □

Um die Eulerzahl genauer berechnen zu können, beweisen wir die folgende Fehlerabschätzung.

Satz 3.12 *Es sei $b_n = \sum_{k=0}^n \frac{1}{k!}$. Dann gilt für jedes $n \in \mathbb{N}$ die folgende Abschätzung für die Eulerzahl e :*

$$b_n < e < b_n + \frac{1}{n \cdot n!}$$

Beweis. Für festes n gilt $e - b_n = \lim_{k \rightarrow \infty} (b_k - b_n)$. Für $k > n$ erhalten wir

$$\begin{aligned} b_k - b_n &= \frac{1}{(n+1)!} + \dots + \frac{1}{k!} \\ &= \frac{1}{(n+1)!} \left(1 + \frac{1}{(n+2)} + \frac{1}{(n+2)(n+3)} + \dots + \frac{1}{(n+2) \cdot \dots \cdot k} \right) \\ &< \frac{1}{(n+1)!} \left(1 + \frac{1}{(n+2)} + \frac{1}{(n+2)^2} + \dots + \frac{1}{(n+2)^{k-n-1}} \right) \\ &= \frac{1}{(n+1)!} \cdot \frac{1 - \left(\frac{1}{n+2}\right)^{k-n}}{1 - \left(\frac{1}{n+2}\right)} \quad (\text{geometrische Summe}) \\ &< \frac{1}{(n+1)!} \cdot \frac{1}{1 - \frac{1}{n+2}} = \frac{1}{(n+1)!} \cdot \frac{n+2}{n+1} = \frac{1}{n \cdot n!} \cdot \frac{n(n+2)}{(n+1)^2}. \end{aligned}$$

Der Grenzübergang in dieser Ungleichung für k gegen $+\infty$ liefert

$$e - b_n \leq \frac{1}{n \cdot n!} \cdot \frac{n(n+2)}{(n+1)^2} < \frac{1}{n \cdot n!} \quad \forall n \in \mathbb{N}.$$

□

Mittels der Fehlerabschätzung aus Satz 3.12 kann man Näherungswerte für e angeben. Man erhält zum Beispiel für $n = 10$:

$$b_{10} = \sum_{k=0}^{10} \frac{1}{k!} = 2.7182815\dots \quad \frac{1}{10 \cdot 10!} = 0.00000002\dots$$

Folglich ist $2.7182815 < e < 2.7182815\dots + 0.00000002\dots$ und somit

$$e \approx 2.7182815.$$

Als weitere Anwendung der Fehlerabschätzung beweisen wir

Satz 3.13 *Die Eulerzahl e ist irrational.*

Beweis. Angenommen e wäre eine rationale Zahl. Dann können wir e in der Form $e = \frac{p}{q}$ für $p, q \in \mathbb{N}$ darstellen. Für $q \in \mathbb{N}$ gilt nach Satz 3.12

$$0 < e - b_q < \frac{1}{q \cdot q!}.$$

Daraus folgt $0 < e \cdot q! - b_q \cdot q! < \frac{1}{q}$. Da $e = \frac{p}{q}$, ist $e \cdot q!$ ganzzahlig. Wegen

$$b_q = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \dots + \frac{1}{q!}$$

ist $b_q \cdot q!$ ganzzahlig. Somit ist auch $e \cdot q! - b_q \cdot q!$ ganzzahlig. Das ist aber ein Widerspruch zu $0 < e \cdot q! - b_q \cdot q! < \frac{1}{q} < 1$. □

3.1.3 Cauchy-Folgen und Konvergenz

Bei den bisherigen Konvergenzuntersuchungen von Folgen haben wir immer benutzt, dass wir den möglichen Grenzwert der Folge bereits kennen und dann das Konvergenzkriterium aus der Definition nachgeprüft. Wir fragen uns nun, ob man die Konvergenz einer Folge beschreiben kann *ohne* den möglichen Grenzwert zu kennen. Dazu definieren wir:

Definition 3.10. Eine Folge reeller oder komplexer Zahlen (z_n) heißt *Cauchy-Folge*, wenn zu jedem $\varepsilon > 0$ ein (von ε abhängiges) $n_0 \in \mathbb{N}$ existiert, so dass $|z_n - z_m| < \varepsilon$ für alle $n, m \geq n_0$ gilt.

Satz 3.14 (Konvergenzkriterium von Cauchy)

Eine Folge reeller oder komplexer Zahlen (z_n) ist genau dann konvergent, wenn sie eine Cauchy-Folge ist.

Beweis. 1. Wir zeigen, dass jede konvergente Folge eine Cauchy-Folge ist.

Sei (z_n) eine gegen z konvergente Folge reeller oder komplexer Zahlen und $\varepsilon > 0$. Dann existiert ein $n_0 \in \mathbb{N}$ mit $|z_n - z| < \frac{\varepsilon}{2}$ für alle $n \geq n_0$. Für $n, m \geq n_0$ folgt aus der Dreiecksungleichung

$$|z_n - z_m| \leq |z_n - z| + |z - z_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Somit ist (z_n) eine Cauchy-Folge.

2. Wir zeigen nun, dass jede Cauchy-Folge reeller oder komplexer Zahlen konvergiert.

1. Fall: Sei (z_n) eine Cauchy-Folge *reeller* Zahlen. Wir zeigen zunächst, dass (z_n) beschränkt ist. Nach Definition existiert ein $n_0 \in \mathbb{N}$ so dass $|z_n - z_m| < 1$ für alle $m, n \geq n_0$. Dann gilt

$$|z_n| \leq |z_n - z_{n_0}| + |z_{n_0}| < 1 + |z_{n_0}| \quad \forall n \geq n_0.$$

Dies zeigt

$$|z_n| \leq M := \max(1 + |z_{n_0}|, |z_1|, \dots, |z_{n_0-1}|) \quad \forall n \in \mathbb{N},$$

d.h. (z_n) ist beschränkt. Nach dem Satz von Bolzano-Weierstraß besitzt die beschränkte reelle Folge (z_n) eine konvergente Teilfolge (z_{n_j}) . Deren Grenzwert sei $z \in \mathbb{R}$. Wir zeigen, dass auch die gesamte Folge (z_n) gegen z konvergiert: Sei $\varepsilon > 0$. Dann existiert ein $n^* \in \mathbb{N}$, so dass $|z_n - z_m| < \frac{\varepsilon}{2}$ für alle $n, m \geq n^*$ sowie ein $n_k > n^*$ mit $|z_{n_k} - z| < \frac{\varepsilon}{2}$. Für alle $n \geq n^*$ folgt mit der Dreiecksungleichung

$$|z_n - z| \leq |z_n - z_{n_k}| + |z_{n_k} - z| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Somit gilt $\lim_{n \rightarrow \infty} z_n = z$.

2. Fall: Sei (z_n) eine Cauchy-Folge *komplexer* Zahlen und $z_n := a_n + ib_n$. Dann sind die Folge der Realteile (a_n) sowie die Folge der Imaginärteile (b_n) Cauchy-Folgen in \mathbb{R} , denn wegen

$$|a_n - a_m| \leq \sqrt{(a_n - a_m)^2 + (b_n - b_m)^2} = |z_n - z_m|,$$

$$|b_n - b_m| \leq \sqrt{(a_n - a_m)^2 + (b_n - b_m)^2} = |z_n - z_m|$$

überträgt sich die Cauchy-Folgen-Bedingung von (z_n) auf (a_n) und (b_n) . Aus dem 1. Fall wissen wir dann, dass (a_n) gegen eine reelle Zahl a und (b_n) gegen eine reelle Zahl b konvergiert. Nach Satz 3.4 konvergiert dann aber (z_n) gegen $z = a + ib$. \square

Achtung: Die Gültigkeit von Satz 3.14 basiert auf der Vollständigkeitseigenschaft der reellen Zahlen. Wir haben für den Beweis des Satzes von Bolzano-Weierstraß die Existenz des Supremums einer nach oben beschränkten Mengen benutzt! Satz 3.14 gilt für Folgen in **anderen** Mengen X nicht mehr !! Es kann Cauchy-Folgen in X geben, die in X **nicht** konvergieren.

Beispiel 1: Sei $X := (0, 1) \subset \mathbb{R}$. Auch in dieser Menge kann man den Abstand zwischen zwei Punkten x, y durch den Betrag $|x - y|$ messen. Wir betrachten die Folge (x_n) in $(0, 1)$ mit $x_n := \frac{1}{n}$. (x_n) ist eine Nullfolge in \mathbb{R} , folglich ist sie eine Cauchy-Folge in \mathbb{R} und damit auch in $X = (0, 1)$. Sie hat aber in $X = (0, 1)$ keinen Grenzwert, da $0 \notin (0, 1)$.

Beispiel 2: Sei $X = \mathbb{Q}$ die Menge der rationalen Zahlen, ebenfalls mit dem durch $|x - y|$ gegebenen Abstand.

a) Wir betrachten die Folge rationaler Zahlen (x_n) , definiert durch

$$x_1 := 1 \quad \text{und} \quad x_{n+1} := \frac{1}{2} \left(x_n + \frac{2}{x_n} \right) \quad \forall n \in \mathbb{N}.$$

Wir wissen, dass (x_n) in \mathbb{R} gegen $\sqrt{2}$ konvergiert (siehe Übungsaufgabe 21 a)). Die Folge (x_n) ist also eine Cauchy-Folge in \mathbb{R} und somit auch in \mathbb{Q} . Aber $\sqrt{2}$ ist keine rationale Zahl. Folglich hat die Cauchy-Folge (x_n) in \mathbb{Q} keinen Grenzwert.

b) Wir betrachten die Folge der rationalen Zahlen (y_n) , definiert durch $y_n := \left(1 + \frac{1}{n}\right)^n$. Die Folge (y_n) konvergiert in \mathbb{R} gegen die Eulerzahl e , sie ist also eine Cauchy-Folge in \mathbb{R} und somit auch in \mathbb{Q} . Da die Eulerzahl irrational ist, hat die Cauchy-Folge (y_n) in \mathbb{Q} keinen Grenzwert.

3.2 Reihen in \mathbb{R} und \mathbb{C}

Reihen sind spezielle Folgen. Sie werden z.B. oft benutzt, um Funktionen zu definieren oder Funktionen geeignet zu approximieren. Die Untersuchung der Konvergenz von Reihen ist deshalb von besonderem Interesse. In diesem Abschnitt werden wir Reihen reeller oder komplexer Zahlen behandeln und Kriterien für ihre Konvergenz kennenlernen.

3.2.1 Konvergente und divergente Reihen, Konvergenzkriterien

Da man Zahlen addieren kann, kann man jeder Folge reeller oder komplexer Zahlen (x_k) eine neue Folge zuordnen, die *Folge der Partialsummen* $(s_n)_{n=1}^\infty$:

$$s_n := x_1 + x_2 + \cdots + x_n = \sum_{k=1}^n x_k \quad \text{also}$$

$$s_1 := x_1$$

$$s_2 := x_1 + x_2$$

$$s_3 := x_1 + x_2 + x_3$$

$$s_4 := x_1 + x_2 + x_3 + x_4$$

...

Definition 3.11. Die Folge $(s_n)_{n=1}^\infty$ heißt *Reihe* mit den Gliedern x_k . Man schreibt für diese Reihe symbolisch

$$\sum_{k=1}^{\infty} x_k \quad \text{oder} \quad x_1 + x_2 + x_3 + \dots$$

Die Zahl $s_n := x_1 + \dots + x_n = \sum_{k=1}^n x_k$ nennt man die *n-te Partialsumme* der Reihe.

Eine Reihe $\sum_{k=1}^{\infty} x_k$ heißt *konvergent*, falls die Folge der Partialsummen (s_n) in \mathbb{R} bzw. in \mathbb{C} konvergiert. Ist (s_n) konvergent, so heißt $s := \lim_{n \rightarrow \infty} s_n$ *Wert der Reihe* und man schreibt

$$s = \sum_{k=1}^{\infty} x_k.$$

Eine Reihe, die in \mathbb{R} bzw. \mathbb{C} nicht konvergiert, heißt *divergent*.

Das Symbol $\sum_{k=1}^{\infty} x_k$ hat also zwei Bedeutungen: Es bezeichnet symbolisch die Folge (s_n) der Partialsummen und im Konvergenzfall auch ihren Grenzwert.

Aus den Grenzwertsätzen für Folgen erhält man, dass die Reihe $\sum_{k=1}^{\infty} x_k$ genau dann konvergiert, wenn die Reihe $\sum_{k=k_0}^{\infty} x_k$ für ein beliebig gewähltes $k_0 \in \mathbb{N}$ konvergiert.

Wenn für eine Reihe *reeller* Zahlen $\lim_{n \rightarrow \infty} s_n = +\infty$ oder $\lim_{n \rightarrow \infty} s_n = -\infty$ gilt, so schreibt man symbolisch

$$\sum_{k=1}^{\infty} x_k = +\infty \quad \text{oder} \quad \sum_{k=1}^{\infty} x_k = -\infty.$$

Die Reihe ist in diesem Fall in \mathbb{R} divergent. Gilt $x_k \geq 0$ für alle $k \in \mathbb{N}$, so bedeutet die Schreibweise

$$\sum_{k=1}^{\infty} x_k < +\infty,$$

dass die Reihe in \mathbb{R} konvergiert.

Definition 3.12. Eine Reihe $\sum_{k=1}^{\infty} x_k$ reeller oder komplexer Zahlen heißt *absolut-konvergent*, wenn die Reihe der Beträge $\sum_{k=1}^{\infty} |x_k|$ konvergiert.

Man kann Reihen addieren und mit Skalaren multiplizieren. Für die Grenzwerte gilt dabei:

Satz 3.15 Seien (x_k) und (y_k) Folgen in \mathbb{R} (bzw. in \mathbb{C}) und λ, μ reelle (bzw. komplexe) Zahlen. Konvergiert die Reihe $\sum_{k=1}^{\infty} x_k$ gegen x und die Reihe $\sum_{k=1}^{\infty} y_k$ gegen y , so konvergiert die Reihe

$$\lambda \cdot \sum_{k=1}^{\infty} x_k + \mu \cdot \sum_{k=1}^{\infty} y_k := \sum_{k=1}^{\infty} (\lambda x_k + \mu y_k)$$

gegen $\lambda x + \mu y$.

Beweis. Seien $s_n := \sum_{k=1}^n x_k$, $\hat{s}_n := \sum_{k=1}^n y_k$ und $s_n^* := \sum_{k=1}^n (\lambda x_k + \mu y_k)$. Dann gilt $s_n^* = \lambda s_n + \mu \hat{s}_n$. Die Behauptung des Satzes folgt aus Satz 3.3. \square

Aus Satz 3.8 erhalten wir das folgende Kriterium für die Konvergenz von Reihen reeller Zahlen:

Satz 3.16 Sind die Reihenglieder x_k reelle Zahlen mit $x_k \geq 0$, dann konvergiert die Reihe $\sum_{k=1}^{\infty} x_k$ genau dann, wenn die Folge der Partialsummen $(s_n = x_1 + \dots + x_n)$ beschränkt ist.

Beweis. Da die Reihenglieder x_k nicht negativ sind, ist die Folge der Partialsummen (s_n) monoton wachsend. Ist die Folge (s_n) außerdem beschränkt, so konvergiert sie nach Satz 3.8. Ist andererseits die Folge der Partialsummen (s_n) konvergent, so ist sie nach Satz 3.2 beschränkt. \square

Wir leiten jetzt einige wichtige **Konvergenzkriterien** für Reihen her.

Satz 3.17 (Cauchy-Kriterium für Reihen) Eine Reihe $\sum_{k=1}^{\infty} x_k$ reeller oder komplexer Zahlen ist genau dann konvergent, wenn zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass

$$|x_n + \dots + x_m| < \varepsilon \quad \forall m \geq n \geq n_0. \quad (*)$$

Beweis. Die Folge der Partialsummen $s_n = x_1 + \dots + x_n$ konvergiert genau dann, wenn sie eine Cauchy-Folge ist (Cauchy-Kriterium für Folgen, Satz 3.14). (s_n) ist genau dann eine Cauchy-Folge, wenn zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass für alle $m \geq n \geq n_0$ die Ungleichung $|s_m - s_{n-1}| < \varepsilon$ gilt, was äquivalent zum Cauchy-Kriterium (*) ist. \square

Daraus erhalten wir ein nützliches *notwendiges* Kriterium für die Konvergenz von Reihen:

Satz 3.18 Ist eine Reihe $\sum_{k=1}^{\infty} x_k$ reeller oder komplexer Zahlen konvergent, so ist die Folge der Reihenglieder (x_k) eine Nullfolge.

Beweis. Zum Beweis nutzen wir das Cauchy-Kriterium für $m = n$. Konvergiert $\sum_{k=1}^{\infty} x_k$, dann existiert für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass $|x_n| < \varepsilon$ für alle $n \geq n_0$. Daraus folgt $\lim_{n \rightarrow \infty} x_n = 0$. \square

Die Umkehrung dieses Satzes gilt im allgemeinen nicht.

Beispiel 1: Die harmonische Reihe

Wir betrachten die *harmonische Reihe*

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots$$

Behauptung: Die harmonische Reihe ist in \mathbb{R} divergent und es gilt

$$\sum_{k=1}^{\infty} \frac{1}{k} = +\infty.$$

Beweis. Sei $m \in \mathbb{N}$ fixiert. Wir wählen ein $n \in \mathbb{N}$ mit $n \geq 2^m$. Dann gilt

$$\begin{aligned} s_n &= 1 + \frac{1}{2} + \dots + \frac{1}{n} \\ &\geq 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \dots + \frac{1}{8}\right) + \dots + \underbrace{\left(\frac{1}{2^{m-1}+1} + \dots + \frac{1}{2^m}\right)}_{2^{m-1} \text{ Summanden}} \\ &\geq 1 + \underbrace{\frac{1}{2} + 2 \cdot \frac{1}{4} + 4 \cdot \frac{1}{8} + \dots + 2^{m-1} \cdot \frac{1}{2^m}}_{m \text{ Summanden}} \\ &\geq 1 + \frac{m}{2}. \end{aligned}$$

Folglich existiert zu jedem $M \in \mathbb{R}$ eine Zahl m , so dass $s_n \geq 1 + \frac{m}{2} \geq M$ für alle $n \geq 2^m$. Somit strebt die Folge der Partialsummen (s_n) gegen $+\infty$. \square

Beispiel 2: Die Riemannsche Zeta-Funktion

Sei $s \in \mathbb{Q}$ eine rationale Zahl. Die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k^s} = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \frac{1}{4^s} + \dots$$

ist konvergent, falls $s > 1$ und divergent, falls $s \leq 1$.

Beweis: 1. Fall: $s > 1$. Für $n \in \mathbb{N}$ wählen wir ein ν mit $2^\nu - 1 \geq n$ und schätzen die Partialsumme s_n mit Hilfe der Partialsumme $s_{2^\nu - 1}$ ab:

$$\begin{aligned} s_n \leq s_{2^\nu - 1} &= 1 + \left(\frac{1}{2^s} + \frac{1}{3^s}\right) + \left(\frac{1}{4^s} + \frac{1}{5^s} + \frac{1}{6^s} + \frac{1}{7^s}\right) + \dots + \left(\frac{1}{2^{(\nu-1)s}} + \dots + \frac{1}{(2^\nu - 1)^s}\right) \\ &\leq 1 + 2 \cdot \frac{1}{2^s} + 2^2 \cdot \frac{1}{2^{2s}} + \dots + 2^{\nu-1} \frac{1}{2^{(\nu-1)s}} \\ &= 1 + \frac{1}{2^{s-1}} + \left(\frac{1}{2^{s-1}}\right)^2 + \dots + \left(\frac{1}{2^{s-1}}\right)^{\nu-1} \\ &= \frac{1 - \left(\frac{1}{2^{s-1}}\right)^\nu}{1 - \frac{1}{2^{s-1}}} < \frac{1}{1 - \frac{1}{2^{s-1}}}. \quad (\text{geometrische Summe}) \end{aligned}$$

Die Folge der Partialsummen (s_n) ist somit beschränkt, also ist (s_n) konvergent.

2. Fall: $s \leq 1$. Dann gilt

$$s_n = 1 + \frac{1}{2^s} + \frac{1}{3^s} + \dots + \frac{1}{n^s} \geq 1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n}.$$

Da die Partialsummenfolge der harmonischen Reihe gegen $+\infty$ stebt (Beispiel 1), ist die Folge (s_n) unbeschränkt, also divergent. \square

Durch $\zeta(s) := \sum_{k=1}^{\infty} \frac{1}{k^s}$, $s > 1$, wird die sogenannte *Riemannsche Zeta-Funktion* definiert. Sie spielt z.B. bei der Untersuchung der Verteilung der Primzahlen eine wichtige Rolle.

Beispiel 3: Die geometrische Reihe

Sei $z \in \mathbb{C}$ eine fixierte komplexe Zahl. Wir betrachten die *geometrische Reihe*

$$\sum_{k=0}^{\infty} z^k = 1 + z + z^2 + z^3 + z^4 + \dots$$

Behauptung:

1. Ist $|z| < 1$, so konvergiert die geometrische Reihe und für ihren Wert gilt $\sum_{k=0}^{\infty} z^k = \frac{1}{1-z}$.
2. Ist $|z| \geq 1$, so divergiert die geometrische Reihe $\sum_{k=0}^{\infty} z^k$.

Beweis. Für die Partialsumme gilt nach Satz 2.9

$$s_n = 1 + z + \dots + z^n = \frac{1 - z^{n+1}}{1 - z}.$$

Für $|z| < 1$ ist (z^{n+1}) eine Nullfolge und somit gilt $\lim_{n \rightarrow \infty} s_n = \frac{1}{1-z}$.

Für $|z| \geq 1$ ist $|z^k| \geq 1$ und somit ist (z^k) keine Nullfolge. Deshalb ist $\sum_{k=0}^{\infty} z^k$ divergent (siehe Satz 3.18). \square

Satz 3.19 *Ist eine Reihe $\sum_{k=1}^{\infty} x_k$ reeller oder komplexer Zahlen absolut-konvergent, so ist sie auch konvergent und für die Werte der Reihen gilt*

$$\left| \sum_{k=1}^{\infty} x_k \right| \leq \sum_{k=1}^{\infty} |x_k|. \quad (3.2)$$

Beweis. Sei $\sum_{k=1}^{\infty} |x_k|$ konvergent. Entsprechend dem Cauchy-Kriterium gibt es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass

$$|x_n| + \dots + |x_m| < \varepsilon \quad \text{für alle } m \geq n \geq n_0.$$

Wegen der Dreiecksungleichung für den Betrag

$$|x_n + \dots + x_m| \leq |x_n| + \dots + |x_m|$$

gilt das Cauchy-Kriterium auch für die Reihe $\sum_{k=1}^{\infty} x_k$, die folglich ebenfalls konvergiert. Für die konvergente Folge (s_n) der Partialsummen $s_n := x_1 + \dots + x_n$ gilt $|\lim_{n \rightarrow \infty} s_n| = \lim_{n \rightarrow \infty} |s_n|$ (Satz 3.3 d)). Daraus folgt

$$\left| \sum_{k=1}^{\infty} x_k \right| = \left| \lim_{n \rightarrow \infty} s_n \right| = \lim_{n \rightarrow \infty} |s_n| \leq \lim_{n \rightarrow \infty} (|x_1| + \dots + |x_n|) = \sum_{k=1}^{\infty} |x_k|.$$

Dies zeigt die Abschätzung (3.2). \square

Satz 3.20 (Majorantenkriterium)

Sei $\sum_{k=1}^{\infty} x_k$ eine Reihe reeller oder komplexer Zahlen und (c_k) eine Folge reeller Zahlen mit $|x_k| \leq c_k$ für alle $k \in \mathbb{N}$. Konvergiert die Reihe $\sum_{k=1}^{\infty} c_k$ in \mathbb{R} , so ist die Reihe $\sum_{k=1}^{\infty} x_k$ absolut-konvergent und für die Werte der Reihen gilt

$$\left| \sum_{k=1}^{\infty} x_k \right| \leq \sum_{k=1}^{\infty} |x_k| \leq \sum_{k=1}^{\infty} c_k. \quad (3.3)$$

Beweis. Wir nutzen wiederum das Cauchy-Kriterium. Sei $\sum_{k=1}^{\infty} c_k$ konvergent. Dann gibt es für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass

$$c_n + \dots + c_m < \varepsilon \quad \forall m \geq n \geq n_0.$$

Nach Voraussetzung ist $|x_n| + \dots + |x_m| \leq c_n + \dots + c_m$. Folglich gilt das Cauchy-Kriterium auch für die Reihe $\sum_{k=1}^{\infty} |x_k|$. Somit ist die Reihe $\sum_{k=1}^{\infty} x_k$ absolut-konvergent, also auch konvergent. Die Ungleichung (3.3) folgt wie im Beweis von Satz 3.19. \square

Satz 3.21 (Wurzelkriterium)

Sei $\sum_{k=1}^{\infty} x_k$ eine Reihe reeller oder komplexer Zahlen und $\alpha := \limsup_{k \rightarrow \infty} \sqrt[k]{|x_k|}$.

1. Ist $\alpha < 1$, so ist die Reihe $\sum_{k=1}^{\infty} x_k$ absolut-konvergent und somit auch konvergent.
2. Ist $\alpha > 1$, so ist die Reihe $\sum_{k=1}^{\infty} x_k$ divergent.

Beweis. 1. Sei $\alpha = \limsup_{k \rightarrow \infty} \sqrt[k]{|x_k|} < 1$. Dann gilt $\alpha < \frac{1+\alpha}{2} < 1$. Da $\alpha = \sup HP(\sqrt[k]{|x_k|})$, sind höchstens endlich viele dieser Folgeglieder größer als $\frac{1+\alpha}{2}$. Es existiert folglich ein $k_0 \in \mathbb{N}$ so dass

$$\sqrt[k]{|x_k|} < \frac{1+\alpha}{2} \quad \forall k \geq k_0.$$

Somit gilt

$$|x_k| < \left(\frac{1+\alpha}{2}\right)^k \quad \forall k \geq k_0.$$

Da $\frac{1+\alpha}{2} < 1$, konvergiert die geometrische Reihe $\sum_{k=1}^{\infty} \left(\frac{1+\alpha}{2}\right)^k$. Aus dem Majorantenkriterium folgt dann die absolute Konvergenz der Reihe $\sum_{k=1}^{\infty} x_k$ und somit auch ihre Konvergenz.

2. Sei $\alpha = \limsup_{k \rightarrow \infty} \sqrt[k]{|x_k|} > 1$. Dann existiert eine Teilfolge (x_{k_j}) von (x_k) mit $\sqrt[k_j]{|x_{k_j}|} \geq 1$, also mit $|x_{k_j}| \geq 1$. Somit ist (x_k) keine Nullfolge, also konvergiert $\sum_{k=1}^{\infty} x_k$ nicht (Satz 3.18). \square

Satz 3.22 (Quotientenkriterium)

Sei $\sum_{k=1}^{\infty} x_k$ eine Reihe reeller oder komplexer Zahlen, deren Glieder x_k alle von Null verschieden sind.

1. Ist $\alpha := \limsup_{k \rightarrow \infty} \frac{|x_{k+1}|}{|x_k|} < 1$, so ist die Reihe $\sum_{k=1}^{\infty} x_k$ absolut-konvergent und somit auch konvergent.
2. Ist $\beta := \liminf_{k \rightarrow \infty} \frac{|x_{k+1}|}{|x_k|} > 1$, so ist die Reihe $\sum_{k=1}^{\infty} x_k$ divergent.

Beweis. 1. Sei $\alpha = \limsup_{k \rightarrow \infty} \frac{|x_{k+1}|}{|x_k|} < 1$. Dann existiert ein $k_0 \in \mathbb{N}$, so dass

$$\frac{|x_{k+1}|}{|x_k|} < \frac{1+\alpha}{2} < 1 \quad \forall k \geq k_0.$$

Folglich gilt

$$|x_{k+1}| < \left(\frac{1+\alpha}{2}\right) |x_k| \quad \forall k \geq k_0$$

und somit

$$|x_{k_0+j}| < \left(\frac{1+\alpha}{2}\right)^j |x_{k_0}| \quad \forall j \geq 0.$$

Da $\frac{1+\alpha}{2} < 1$, konvergiert die geometrische Reihe $\sum_{j=0}^{\infty} \left(\frac{1+\alpha}{2}\right)^j$. Aus dem Majorantenkriterium folgt dann, dass die Reihe $\sum_{k=1}^{\infty} x_k$ absolut konvergiert und somit auch konvergiert.

2. Sei nun $\beta = \liminf_{k \rightarrow \infty} \frac{|x_{k+1}|}{|x_k|} > 1$. Dann sind höchstens endlich viele der Zahlen $\frac{|x_{k+1}|}{|x_k|}$ kleiner als 1. Folglich existiert ein $k_0 \in \mathbb{N}$, so dass $0 < |x_k| \leq |x_{k+1}|$ für alle $k \geq k_0$. Also ist (x_k) keine Nullfolge. Nach Satz 3.18 ist deshalb die Reihe $\sum_{k=1}^{\infty} x_k$ divergent. \square

Beispiel 4: Ob man das Wurzel- oder das Quotientenkriterium anwendet, muß man anhand der Gestalt der Reihenglieder entscheiden. Das Wurzelkriterium ist leistungsfähiger als das Quotientenkriterium. Betrachten wir z.B. die Reihe $\sum_{k=1}^{\infty} x_k$, wobei $x_k := 2^{-k}$ für gerade k und $x_k := 8^{-k}$ für ungerades k sei. Das Wurzelkriterium zeigt Konvergenz an, da

$\limsup_{k \rightarrow \infty} \sqrt[k]{x_k} = \frac{1}{2}$, während das Quotientenkriterium keine Aussage liefert, da $\liminf_{k \rightarrow \infty} \frac{x_{k+1}}{x_k} = 0$ und $\limsup_{k \rightarrow \infty} \frac{x_{k+1}}{x_k} = +\infty$.

Beispiel 5: Für jede komplexe Zahl $z \in \mathbb{C}$ ist die Reihe

$$\sum_{k=0}^{\infty} \frac{z^k}{k!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \frac{z^4}{4!} + \dots$$

absolut-konvergent.

Beweis. Wir benutzen das Quotientenkriterium mit $x_k = \frac{z^k}{k!}$:

$$\left| \frac{x_{k+1}}{x_k} \right| = \left| \frac{z^{k+1} \cdot k!}{z^k \cdot (k+1)!} \right| = \frac{|z|}{k+1} \rightarrow 0.$$

□

Beispiel 6: Sei $z \in \mathbb{C}$. Die Reihe

$$\sum_{k=1}^{\infty} \frac{z^k}{k} = z + \frac{z^2}{2} + \frac{z^3}{3} + \frac{z^4}{4} + \dots +$$

ist absolut-konvergent, falls $|z| < 1$ und divergent, falls $|z| > 1$.

Beweis. Wir benutzen das Wurzelkriterium. Mit $x_k := \frac{z^k}{k}$ ist $\sqrt[k]{|x_k|} = \frac{|z|}{\sqrt[k]{k}}$. Da $\lim_{k \rightarrow \infty} \sqrt[k]{k} = 1$, konvergiert die Folge $(\sqrt[k]{|x_k|})$ gegen $|z|$. Damit ist $\limsup_{k \rightarrow \infty} \sqrt[k]{|x_k|} = |z|$ und das Wurzelkriterium liefert die Behauptung. □

Für $z \in \mathbb{C}$ mit $|z| = 1$ kann sowohl Konvergenz als auch Divergenz vorliegen. Für $z = 1$ ist die obige Reihe z.B. gerade die harmonische Reihe $\sum_{k=0}^{\infty} \frac{1}{k}$, also divergent. Wir zeigen mit dem nächsten Kriterium, dass im Gegensatz dazu die Reihe für $z = -1$, d.h.

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} \pm \dots$$

konvergiert.

Definition 3.13. Eine Reihe $\sum_{k=1}^{\infty} x_k$ reeller Zahlen heißt alternierend, wenn die Reihenglieder ihr Vorzeichen wechseln, d.h. wenn

$$x_{k+1} > 0 \iff x_k < 0 \quad \forall k \in \mathbb{N}.$$

Satz 3.23 (Leibniz–Kriterium für alternierende Reihen)

Sei $(b_k)_{k=0}^\infty$ eine monoton fallende Nullfolge positiver reeller Zahlen. Dann gilt:

1. Die alternierende Reihe $\sum_{k=0}^\infty (-1)^k b_k = b_0 - b_1 + b_2 - b_3 + \dots$ konvergiert.
2. Fehlerabschätzung für den Wert s der Reihe $\sum_{k=0}^\infty (-1)^k b_k$:

$$\left| s - \sum_{k=0}^n (-1)^k b_k \right| \leq b_{n+1}.$$

Beweis. Zu 1) Wir betrachten die n . Partialsumme $s_n := \sum_{k=0}^n (-1)^k b_k$. Da (b_k) monoton fallend ist, gilt:

$$\begin{aligned} s_{2n+2} - s_{2n} &= b_{2n+2} - b_{2n+1} \leq 0, \\ s_{2n+3} - s_{2n+1} &= -b_{2n+3} + b_{2n+2} \geq 0, \\ s_{2n+1} - s_{2n} &= -b_{2n+1} < 0, \end{aligned}$$

und folglich

$$\begin{aligned} s_0 &\geq s_2 \geq s_4 \geq s_6 \geq \dots \\ s_1 &\leq s_3 \leq s_5 \leq s_7 \leq \dots \\ s_1 &\leq s_{2n+1} < s_{2n} \leq s_0 \quad \forall n \in \mathbb{N}_0. \end{aligned} \tag{3.4}$$

Somit ist die Folge (s_{2n}) monoton fallend und von unten durch s_1 beschränkt; die Folge (s_{2n+1}) monoton wachsend und nach oben durch s_0 beschränkt. Nach Satz 3.8 existieren deshalb die Grenzwerte $g := \lim_{n \rightarrow \infty} s_{2n}$ und $u := \lim_{n \rightarrow \infty} s_{2n+1}$. Außerdem gilt:

$$g - u = \lim_{n \rightarrow \infty} (s_{2n} - s_{2n+1}) = \lim_{n \rightarrow \infty} b_{2n+1} = 0.$$

Wir zeigen nun, dass die gesamte Folge (s_n) gegen $s := g = u$ konvergiert. Sei $\varepsilon > 0$. Dann existieren $n_0, n_0^* \in \mathbb{N}$ so dass $|s - s_{2n}| < \varepsilon$ für alle $n \geq n_0$ und $|s - s_{2n+1}| < \varepsilon$ für alle $n \geq n_0^*$. Folglich gilt $|s - s_n| < \varepsilon$ für alle $n \geq \max(2n_0, 2n_0^* + 1)$, d.h. (s_n) konvergiert gegen s .

Zu 2) Die Abschätzung (3.4) zeigt, dass

$$s_{2m+1} \leq s \leq s_{2m} \quad \forall m \in \mathbb{N}$$

und somit

$$|s - s_n| \leq |s_{n+1} - s_n| = b_{n+1} \quad \forall n \in \mathbb{N}.$$

Dies zeigt die Fehlerabschätzung. □

Beispiel 7: Die alternierende harmonische Reihe und die Leibniz-Reihe

- Die alternierende harmonische Reihe

$$\sum_{k=1}^\infty (-1)^{k-1} \frac{1}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots$$

konvergiert. Wir werden in Kapitel 5 sehen, dass ihr Wert $\ln(2)$ ist.

- Die Leibniz-Reihe

$$\sum_{k=0}^{\infty} (-1)^k \frac{1}{2k+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \frac{1}{9} - \dots$$

konvergiert. Wir werden in Kapitel 5 sehen, dass ihr Wert $\frac{\pi}{4}$ ist.

Das Cauchy-Produkt von Reihen

Wir haben bereits gesehen, dass man konvergente Reihen von reellen oder komplexen Zahlen addieren kann. Wir überlegen uns jetzt, wie man Reihen *multiplizieren* kann und wann das Produkt konvergiert.

Definition 3.14. Seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ zwei Reihen in \mathbb{R} oder \mathbb{C} . Wir betrachten eine

neue Reihe $\sum_{k=0}^{\infty} c_k$ mit den Reihengliedern

$$c_k := \sum_{j=0}^k a_j \cdot b_{k-j} = a_0 b_k + a_1 b_{k-1} + \dots + a_{k-1} b_1 + a_k b_0$$

Die Reihe $\sum_{k=0}^{\infty} c_k$ heißt *Cauchy-Produkt* der Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$.

Wir wollen die Frage untersuchen, unter welchen Bedingungen aus der Konvergenz der Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ die Konvergenz des Cauchy-Produktes $\sum_{k=0}^{\infty} c_k$ folgt. Im allgemeinen folgt sie nicht.

Beispiel 8: Wir betrachten die alternierenden Reihen

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} b_k = \sum_{k=0}^{\infty} \frac{(-1)^k}{\sqrt{k+1}} = 1 - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{4}} \pm \dots$$

Diese Reihen konvergieren nach dem Leibniz-Kriterium für alternierende Reihen. Betrachten wir aber die Folge $c_k = \sum_{j=0}^k a_j b_{k-j}$, so erhalten wir

$$c_k = \sum_{j=0}^k \frac{(-1)^j}{\sqrt{j+1}} \cdot \frac{(-1)^{k-j}}{\sqrt{k-j+1}} = \sum_{j=0}^k \frac{(-1)^k}{\sqrt{(j+1)(k-j+1)}}.$$

Es gilt

$$(k-j+1)(j+1) = \left(\frac{1}{2}k+1\right)^2 - \left(\frac{1}{2}k-j\right)^2 \leq \left(\frac{1}{2}k+1\right)^2$$

und folglich

$$|c_k| \geq \frac{k+1}{\frac{1}{2}k+1} = \frac{2(k+1)}{k+2} \rightarrow 2.$$

Damit ist (c_k) keine Nullfolge und die Reihe $\sum_{k=0}^{\infty} c_k$ somit divergent.

Satz 3.24 Seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ konvergente Reihen reeller oder komplexer Zahlen und sei mindestens eine der beiden Reihen absolut-konvergent. Dann konvergiert ihr Cauchy-Produkt $\sum_{k=0}^{\infty} c_k$ und für die Werte der Reihen gilt

$$\sum_{k=0}^{\infty} c_k = \left(\sum_{k=0}^{\infty} a_k \right) \cdot \left(\sum_{k=0}^{\infty} b_k \right).$$

*Beweis.*³ Sei oBdA $\sum_{k=0}^{\infty} a_k$ absolut-konvergent. Wir setzen $A_n := \sum_{k=0}^n a_k$, $B_n := \sum_{k=0}^n b_k$, $C_n := \sum_{k=0}^n c_k$. Weiterhin bezeichne $A := \lim_{n \rightarrow \infty} A_n$, $B = \lim_{n \rightarrow \infty} B_n$ und $\beta_n = B_n - B$. Dann erhalten wir

$$\begin{aligned} C_n &= c_0 + c_1 + \dots + c_n \\ &= a_0 b_0 + (a_0 b_1 + a_1 b_0) + \dots + (a_0 b_n + a_1 b_{n-1} + \dots + a_n b_0) \\ &= a_0 \cdot B_n + a_1 \cdot B_{n-1} + \dots + a_n \cdot B_0 \\ &= a_0(\beta_n + B) + a_1(\beta_{n-1} + B) + \dots + a_n(\beta_0 + B) \\ &= A_n \cdot B + \underbrace{a_0 \beta_n + a_1 \beta_{n-1} + \dots + a_n \beta_0}_{=:\gamma_n}. \end{aligned}$$

Da $A_n \cdot B$ gegen $A \cdot B$ konvergiert, bleibt $\gamma_n \rightarrow 0$ zu zeigen.

Sei $\varepsilon > 0$ gegeben. Da B_n gegen B konvergiert, existiert ein $n_0 \in \mathbb{N}$ so dass $|\beta_n| = |B_n - B| < \varepsilon$ für alle $n \geq n_0$. Damit schätzen wir $|\gamma_n|$ ab:

$$\begin{aligned} |\gamma_n| &= |a_0 \beta_n + a_1 \beta_{n-1} + \dots + a_n \beta_0| \\ &\leq |a_0 \beta_n + \dots + a_{n-n_0} \beta_{n_0}| + |a_{n-n_0+1} \beta_{n_0-1} + \dots + a_n \beta_0| \\ &\leq \varepsilon(|a_0| + \dots + |a_{n-n_0}|) + |a_{n-n_0+1} \beta_{n_0-1} + \dots + a_n \beta_0| \\ &\leq \varepsilon \cdot \underbrace{\left(\sum_{k=0}^{\infty} |a_k| \right)}_{=: A^*} + |a_{n-n_0+1}| |\beta_{n_0-1}| + \dots + |a_n| |\beta_0| \\ &\leq \varepsilon A^* + |a_{n-n_0+1}| |\beta_{n_0-1}| + \dots + |a_n| |\beta_0| \quad \forall n \geq n_0. \end{aligned}$$

A^* ist endlich, da die Reihe $\sum_{k=0}^{\infty} a_k$ nach Voraussetzung absolut-konvergiert. Da (a_n) eine Nullfolge ist, können wir die letzten n_0 -Summanden abschätzen: Es existiert ein $n_1 \in \mathbb{N}$, so dass

$$|\gamma_n| \leq \varepsilon A^* + \varepsilon \quad \forall n \geq \max\{n_0, n_1\}.$$

Somit konvergiert die Folge (γ_n) gegen 0 und die Folge $(C_n = A_n \cdot B + \gamma_n)$ gegen $A \cdot B$. \square

Satz 3.25 Das Cauchy-Produkt zweier absolut-konvergenter Reihen ist absolut-konvergent.

Beweis. Dies überlassen wir dem Leser als Übungsaufgabe. \square

³ Den Beweis haben wir in der Vorlesung aus Zeitgründen nicht geführt. Ich führe ihn hier für interessierte Studierende an.

3.2.2 Komplexe Potenzreihen

In diesem Abschnitt betrachten wir spezielle Reihen komplexer Zahlen, die sogenannten Potenzreihen.

Definition 3.15. Sei $z_0 \in \mathbb{C}$ und $(a_n)_{n=0}^{\infty}$ eine Folge komplexer Zahlen. Eine Potenzreihe mit dem Zentrum z_0 ist eine Reihe komplexer Zahlen der Form

$$P(z) := \sum_{n=0}^{\infty} a_n (z - z_0)^n, \quad z \in \mathbb{C}.$$

Wir wollen die Frage untersuchen, für welche $z \in \mathbb{C}$ die Potenzreihe $P(z)$ konvergiert. Offensichtlich konvergiert die Potenzreihe $P(z)$ für $z = z_0$ und es gilt $P(z_0) = a_0$.

Satz 3.26 Sei $P(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$ eine Potenzreihe mit dem Zentrum z_0 und sei $z_1 \neq z_0$.

1. Ist $P(z_1)$ konvergent, so ist $P(z)$ für jedes $z \in \mathbb{C}$ mit $|z - z_0| < |z_1 - z_0|$ absolut-konvergent.
2. Ist $P(z_1)$ divergent, so ist $P(z)$ für jedes $z \in \mathbb{C}$ mit $|z - z_0| > |z_1 - z_0|$ divergent.

Beweis. Zu 1: Da die Reihe $P(z_1) = \sum_{n=0}^{\infty} a_n (z_1 - z_0)^n$ konvergiert, ist die Folge $(a_n (z_1 - z_0)^n)$ eine Nullfolge, also insbesondere beschränkt. Sei $C \in \mathbb{R}$ so gewählt, dass $|a_n (z_1 - z_0)^n| \leq C$ für alle $n \in \mathbb{N}_0$. Dann gilt

$$|a_n (z - z_0)^n| = |a_n (z_1 - z_0)^n| \left| \frac{z - z_0}{z_1 - z_0} \right|^n \leq C \cdot \left| \frac{z - z_0}{z_1 - z_0} \right|^n.$$

Für $z \in \mathbb{C}$ mit $|z - z_0| < |z_1 - z_0|$ folgt $\left| \frac{z - z_0}{z_1 - z_0} \right| < 1$. Die Reihe $P(z)$ hat also eine konvergente Majorante. Mit dem Majorantenkriterium und dem Grenzwert der geometrischen Reihe erhält man

$$\sum_{n=0}^{\infty} |a_n (z - z_0)^n| \leq C \sum_{n=0}^{\infty} \left| \frac{z - z_0}{z_1 - z_0} \right|^n = C \cdot \frac{1}{1 - \left| \frac{z - z_0}{z_1 - z_0} \right|}.$$

Insbesondere ist $\sum_{n=0}^{\infty} a_n (z - z_0)^n$ absolut-konvergent.

Zu 2: Sei $P(z_1)$ divergent und $|z - z_0| > |z_1 - z_0|$. Wäre $P(z)$ konvergent, so würde aus 1. folgen, dass $P(z_1)$ absolut-konvergent wäre, was einen Widerspruch liefert. \square

Definition 3.16. Die Zahl

$$R := \sup\{|z - z_0| \mid P(z) \text{ ist konvergent}\} \in \mathbb{R} \cup \{+\infty\}$$

heißt *Konvergenzradius* der Potenzreihe $P(z)$.

Die offene Kreisscheibe $K(z_0, R) := \{z \in \mathbb{C} \mid |z - z_0| < R\} \subset \mathbb{C}$ heißt *Konvergenzkreis* von $P(z)$.

Die abgeschlossene Kreisscheibe um z_0 mit Radius R bezeichnen wir mit

$$\text{cl } K(z_0, R) := \{z \in \mathbb{C} \mid |z - z_0| \leq R\}.$$

Wir können die Aussage aus Satz 3.26 jetzt auch folgendermaßen formulieren:

Satz 3.27 Sei $P(z) := \sum_{n=0}^{\infty} a_n(z-z_0)^n$ eine komplexe Potenzreihe mit Zentrum $z_0 \in \mathbb{C}$ und dem Konvergenzradius R . Dann ist $P(z)$ für jeden Punkt $z \in K(z_0, R)$ absolut-konvergent und für jeden Punkt $z \in \mathbb{C} \setminus \text{cl } K(z_0, R)$ divergent.

Aus dem letzten Abschnitt kennen wir bereits folgende Beispiele:

- Für $P(z) = \sum_{n=0}^{\infty} z^n$ ist $R = 1$.
- Für $P(z) = \sum_{n=1}^{\infty} \frac{z^n}{n}$ ist $R = 1$.
- Für $P(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$ ist $R = +\infty$.

Die nächsten beiden Sätze zeigen, wie man den Konvergenzradius einer Potenzreihe bestimmen kann.

Satz 3.28 Sei $P(z) = \sum_{n=0}^{\infty} a_n(z-z_0)^n$ eine Potenzreihe und $\lambda := \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|}$. Dann gilt für den Konvergenzradius R von $P(z)$:

$$R = \begin{cases} \frac{1}{\lambda} & \text{falls } \lambda \in \mathbb{R}^+ \\ 0 & \text{falls } \lambda = +\infty \\ +\infty & \text{falls } \lambda = 0. \end{cases}$$

Beweis. Sei $\alpha := \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n(z-z_0)^n|} = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} \cdot |z-z_0| = \lambda \cdot |z-z_0|$. Nach dem Wurzelkriterium konvergiert $P(z)$ für $\alpha < 1$ und divergiert für $\alpha > 1$.

(1) Sei $0 < \lambda < +\infty$. Dann folgt sofort:

$$P(z) \begin{cases} \text{konvergiert für alle } z \text{ mit } |z - z_0| < \frac{1}{\lambda} \\ \text{divergiert für alle } z \text{ mit } |z - z_0| > \frac{1}{\lambda}. \end{cases}$$

Folglich ist $R = \frac{1}{\lambda}$.

(2) Sei $\lambda = 0$, dann ist auch $\alpha = 0$. Somit konvergiert $P(z)$ für alle $z \in \mathbb{C}$ und der Konvergenzradius R ist $+\infty$.

(3) Sei $\lambda = +\infty$. Dann gilt

$$\alpha = \begin{cases} 0 & z = z_0 \\ +\infty & z \neq z_0. \end{cases}$$

Die Reihe $P(z)$ divergiert also für alle $z \neq z_0$. Somit ist $R = 0$. □

Auf analoge Weise erhält man aus dem Quotientenkriterium:

Satz 3.29 Sei $P(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ eine Potenzreihe mit von Null verschiedenen Koeffizienten a_n und existiere der Grenzwert $\mu := \lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| \in \mathbb{R}^+ \cup \{+\infty\}$. Dann gilt für den Konvergenzradius R von $P(z)$:

$$R = \begin{cases} \frac{1}{\mu} & \text{falls } \mu \in \mathbb{R}^+ \\ 0 & \text{falls } \mu = +\infty \\ +\infty & \text{falls } \mu = 0. \end{cases}$$

3.2.3 Anwendung: Exponentialfunktion, Logarithmusfunktion und komplexe Potenzen

In Abschnitt 2.1.5 haben wir Potenzen a^q definiert, wobei a eine positive reelle Zahl und der Exponent q eine rationale Zahl bezeichnet. In diesem Abschnitt wollen wir Potenzen a^z für komplexe Exponenten $z \in \mathbb{C}$ erklären und die Eigenschaften dieser Potenzen untersuchen. Dazu betrachten wir zunächst die folgende komplexe Potenzreihe:

$$E(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots$$

Die Reihe $E(z)$ hat folgende Eigenschaften:

1. $E(z)$ ist für jedes $z \in \mathbb{C}$ absolut-konvergent.
2. Es gilt $E(0) = 1$ und $E(1) = \sum_{n=0}^{\infty} \frac{1}{n!} = e$, wobei e die Eulerzahl bezeichnet.
3. Es gilt $E(z_1) \cdot E(z_2) = E(z_1 + z_2)$ für alle $z_1, z_2 \in \mathbb{C}$.

Dies läßt sich mit der Formel für das Cauchy-Produkt aus Abschnitt 3.2.1 zeigen. Das Cauchy-Produkt der beiden absolut-konvergenten Reihen $E(z_1)$ und $E(z_2)$ ist

$$\begin{aligned} \sum_{n=0}^{\infty} \left(\sum_{k=0}^n \frac{z_1^k}{k!} \cdot \frac{z_2^{n-k}}{(n-k)!} \right) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(\sum_{k=0}^n \binom{n}{k} z_1^k \cdot z_2^{n-k} \right) \\ &= \sum_{n=0}^{\infty} \frac{(z_1 + z_2)^n}{n!} \\ &= E(z_1 + z_2). \end{aligned}$$

Folglich gilt nach Satz 3.24, dass $E(z_1 + z_2) = E(z_1) \cdot E(z_2)$.

Insbesondere ist $E(z) \neq 0$ und $E(z) \cdot E(-z) = E(0) = 1$ für alle $z \in \mathbb{C}$.

4. Es gilt $\overline{E(z)} = E(\bar{z})$.
5. Es gilt $|E(z) - 1| \leq \frac{|z|}{1 - |z|}$ für alle $z \in \mathbb{C}$ mit $|z| < 1$:

Wir benutzen dazu die Konvergenzeigenschaften der geometrischen Reihe und das Majorantenkriterium und erhalten für die Werte der Reihen:

$$|E(z) - 1| = \left| \sum_{n=1}^{\infty} \frac{z^n}{n!} \right| \leq \sum_{n=1}^{\infty} \frac{|z|^n}{n!} \leq \sum_{n=1}^{\infty} |z|^n = |z| \left(\sum_{n=0}^{\infty} |z|^n \right) = \frac{|z|}{1 - |z|}.$$

6. Es gilt $E(q) = e^q$ für alle $q \in \mathbb{Q}$:

Für $n \in \mathbb{N}$ ist

$$E(n) = E(\underbrace{1 + \dots + 1}_{n\text{-mal}}) = \underbrace{E(1) \cdot \dots \cdot E(1)}_{n\text{-mal}} = \underbrace{e \cdot \dots \cdot e}_{n\text{-mal}} = e^n$$

Für $n \in -\mathbb{N}$ gilt $E(n) = \frac{1}{E(-n)} = \frac{1}{e^{-n}} = e^n$. Folglich ist $E(n) = e^n$ für alle $n \in \mathbb{Z}$. Sei nun $q \in \mathbb{Q}$ und $q = \frac{n}{m}$, wobei $n \in \mathbb{Z}$ und $m \in \mathbb{N}$. Dann ist

$$\begin{aligned} (e^q)^m &= e^{q \cdot m} = e^n = E(n) = E(q \cdot m) = E(\underbrace{q + \dots + q}_{m\text{-mal}}) \\ &= \underbrace{E(q) \cdot \dots \cdot E(q)}_{m\text{-mal}} = E(q)^m. \end{aligned}$$

Da e^q und $E(q)$ positive reelle Zahlen sind, folgt $e^q = E(q)$.

Die letzte Eigenschaft von $E(z)$ rechtfertigt die folgende Definition:

Definition 3.17. Unter der komplexen Potenz e^z der Eulerzahl e verstehen wir den Wert der Potenzreihe $E(z)$, d.h.

$$e^z := E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

Die Funktion

$$\begin{aligned} \exp : \mathbb{C} &\longrightarrow \mathbb{C} \\ z &\longmapsto e^z \end{aligned}$$

heißt komplexe Exponentialfunktion.

Aus den Eigenschaften der Potenzreihe $E(z)$ folgt

Satz 3.30 Die komplexen Potenzen der Eulerzahl erfüllen

1. $e^{z+w} = e^z \cdot e^w$ für alle $z, w \in \mathbb{C}$.
2. $|e^z - 1| \leq \frac{|z|}{1-|z|}$ für alle $z \in \mathbb{C}$ mit $|z| < 1$.
3. Konvergiert (z_n) gegen $z \in \mathbb{C}$, so konvergiert (e^{z_n}) gegen e^z .

Beweis. Wir müssen nur noch 3. beweisen. Sei (z_n) eine konvergente Folge komplexer Zahlen mit dem Grenzwert z . Dann existiert ein $n_0 \in \mathbb{N}$, so dass $|z_n - z| < 1$ für alle $n \geq n_0$. Dann gilt wegen 2. für alle $n \geq n_0$

$$0 \leq |e^{z_n} - e^z| = |e^z(e^{z_n - z} - 1)| = |e^z| \cdot |e^{z_n - z} - 1| \leq |e^z| \frac{|z_n - z|}{1 - |z_n - z|}.$$

Die rechte Seite konvergiert gegen Null. Aus dem Sandwich-Lemma folgt $\lim_{n \rightarrow \infty} e^{z_n} = e^z$. \square

Wir schränken die Exponentialfunktion nun auf die reellen Zahlen ein und untersuchen die Eigenschaften der reellen Funktion $\exp|_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}$.

Definition 3.18. Eine reelle Funktion $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ heißt

monoton wachsend	$:\Leftrightarrow$	Für alle $x, y \in D$ mit $x < y$ gilt $f(x) \leq f(y)$,
streng monoton wachsend	$:\Leftrightarrow$	Für alle $x, y \in D$ mit $x < y$ gilt $f(x) < f(y)$,
monoton fallend	$:\Leftrightarrow$	Für alle $x, y \in D$ mit $x < y$ gilt $f(x) \geq f(y)$,
streng monoton fallend	$:\Leftrightarrow$	Für alle $x, y \in D$ mit $x < y$ gilt $f(x) > f(y)$.

Satz 3.31 1. Die Funktion $\exp|_{\mathbb{R}}$ nimmt nur Werte in \mathbb{R}^+ an und es gilt:

- a) $1 \leq e^x$ für alle $x \in \mathbb{R}$ mit $x \geq 0$.
 b) $0 < e^x \leq 1$ für alle $x \in \mathbb{R}$ mit $x \leq 0$.

2. Die Funktion $\exp|_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}^+$ ist streng monoton wachsend und bijektiv.

Beweis. Zu 1: Nach Definition ist $e^x = E(x) \geq 1$ für alle $x \geq 0$. Da $e^x \cdot e^{-x} = e^0 = 1$, folgt daraus $0 < e^x \leq 1$ für alle $x \leq 0$. Insondere nimmt $\exp|_{\mathbb{R}}$ nur Werte in \mathbb{R}^+ an.

Zu 2: Seien nun x und y zwei reelle Zahlen mit $x < y$. Dann ist

$$e^y = e^{x+(y-x)} = e^x \cdot e^{y-x} > e^x.$$

Somit ist die Funktion $\exp|_{\mathbb{R}}$ streng monoton wachsend. Insbesondere ist sie deshalb injektiv. Es bleibt zu zeigen, dass $\exp|_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}^+$ surjektiv ist.

Sei $y \in \mathbb{R}^+$. Wir suchen eine Zahl $s \in \mathbb{R}$ mit $e^s = y$. Dazu betrachten wir die folgenden Mengen

$$A := \{x \in \mathbb{R} \mid e^x < y\} \quad \text{und} \quad B := \{x \in \mathbb{R} \mid y \leq e^x\}.$$

Diese Mengen bilden einen Dedekindschen Schnitt $(A|B)$ von \mathbb{R} , denn

- \mathbb{R} ist die disjunkte Vereinigung von A und B .
- A und B sind nichtleer: Nach Definition gilt für $y \in \mathbb{R}^+$, dass $e^y > y$, also $y \in B$. Wir wählen ein $n \in \mathbb{N}$ mit $\frac{1}{n} < y$. Aus $n < e^n$ folgt $e^{-n} < \frac{1}{n} < y$ und somit $-n \in A$.
- Für $a \in A$ und $b \in B$ gilt $e^a < y \leq e^b$. Wegen der Monotonie der Exponentialfunktion folgt $a < b$.

Sei $s \in \mathbb{R}$ die auf Grund des Vollständigkeitsaxioms der reellen Zahlen existierende Schnittzahl des Dedekindschen Schnittes $(A|B)$, d.h. es gilt $a \leq s \leq b$ für alle $a \in A$ und $b \in B$.

Wir zeigen nun, dass $e^s = y$ gilt.

Wir betrachten dazu die Folgen $(a_n := s - \frac{1}{n})$ in A und $(b_n := s + \frac{1}{n})$ in B . Dann gilt:

$$e^s \cdot e^{-\frac{1}{n}} = e^{a_n} < y \leq e^{b_n} = e^s \cdot e^{\frac{1}{n}}.$$

Nach Satz 3.30 konvergieren die Folgen $(e^{\frac{1}{n}})$ und $(e^{-\frac{1}{n}})$ gegen $e^0 = 1$. Aus dem Sandwich-Lemma folgt somit $e^s \leq y \leq e^s$, d.h. $y = e^s$. \square

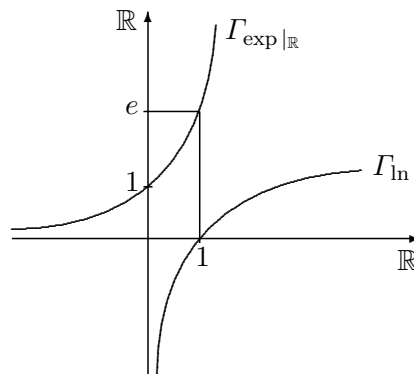
Definition 3.19. Die Umkehrfunktion der reellen Exponentialfunktion $\exp|_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}^+$ bezeichnen wir mit $\ln : \mathbb{R}^+ \rightarrow \mathbb{R}$ und nennen $\ln(y)$ den natürlichen Logarithmus von y .

Die Zahl $\ln(y) \in \mathbb{R}$ ist also eindeutig bestimmt durch die Bedingung $e^{\ln(y)} = y$. Daraus ergeben sich folgende Eigenschaften für den natürlichen Logarithmus:

Satz 3.32 *Es seien $y, y_1, y_2 \in \mathbb{R}^+$ und $q \in \mathbb{Q}$.*

1. Die Funktion $\ln : \mathbb{R}^+ \rightarrow \mathbb{R}$ ist bijektiv und streng monoton wachsend.
2. $\ln(y_1 \cdot y_2) = \ln(y_1) + \ln(y_2)$.
3. $\ln(y^q) = q \cdot \ln(y)$.
4. $\ln\left(\frac{y_1}{y_2}\right) = \ln(y_1) - \ln(y_2)$.
5. $\ln(1) = 0$, $\ln(e) = 1$.
6. $\ln(y) > 0$ für alle $y > 1$ und $\ln(y) < 0$ für alle $0 < y < 1$.

Das folgende Bild zeigt die Graphen der Funktionen $\exp|_{\mathbb{R}}$ und \ln .



Sei $a \in \mathbb{R}^+$. Dann gilt für die Potenz $a^q, q \in \mathbb{Q}$, die Formel

$$a^q = (e^{\ln(a)})^q = e^{\ln(a) \cdot q} = \sum_{n=0}^{\infty} \frac{(q \cdot \ln(a))^n}{n!}.$$

Dies rechtfertigt die folgende Definition der komplexen Potenzen einer positiven reellen Zahl a :

Definition 3.20. *Sei $a \in \mathbb{R}^+$ und $z \in \mathbb{C}$. Unter der Potenz a^z verstehen wir die komplexe Zahl*

$$a^z := e^{\ln(a) \cdot z} = \sum_{n=0}^{\infty} \frac{(z \cdot \ln(a))^n}{n!}.$$

Aus den Eigenschaften der Potenzreihe $E(z)$ ergibt sich unmittelbar:

Satz 3.33 *(Potenzgesetze für komplexe Potenzen)*

Es seien $a, b \in \mathbb{R}^+$, $x \in \mathbb{R}$ und $z, w \in \mathbb{C}$. Dann gilt:

1. $1^z = 1$ und $a^0 = 1$.
2. $a^z \cdot b^z = (ab)^z$.
3. $a^z \cdot a^w = a^{z+w}$.
4. $a^{-z} = \frac{1}{a^z}$.
5. $(a^x)^z = a^{x \cdot z}$.

Sei $a \in \mathbb{R}^+ \setminus \{1\}$. Dann ist die Funktion

$$\begin{aligned} \exp_a : \mathbb{R} &\longrightarrow \mathbb{R}^+ \\ x &\longmapsto a^x \end{aligned}$$

bijektiv, streng monoton wachsend, falls $a > 1$, und streng monoton fallend, falls $0 < a < 1$. Die Umkehrfunktion zu \exp_a bezeichnen wir mit

$$\log_a : \mathbb{R}^+ \longrightarrow \mathbb{R}$$

und nennen $\log_a(y)$ den Logarithmus von y zur Basis a . Aus den Potenzgesetzen von Satz 3.33 folgen dann unmittelbar die Aussagen des folgenden Satzes.

Satz 3.34 *Die Funktion $\log_a : \mathbb{R}^+ \longrightarrow \mathbb{R}$ ist bijektiv, streng monoton wachsend, falls $a > 1$, und streng monoton fallend, falls $0 < a < 1$. Für $x, y \in \mathbb{R}^+$, $\rho \in \mathbb{R}$ und $a \in \mathbb{R}^+$ gilt:*

1. $\log_a x = \frac{\ln x}{\ln a}$.
2. $\log_a(x \cdot y) = \log_a x + \log_a y$.
3. $\log_a(x^\rho) = \rho \cdot \log_a x$.
4. $\log_a\left(\frac{x}{y}\right) = \log_a(x) - \log_a(y)$.
5. $\log_a(1) = 0$, $\log_a(a) = 1$.

Stetige Funktionen

In diesem Kapitel beginnen wir mit der Untersuchung lokaler Eigenschaften von Funktionen. Wir beschränken uns zunächst auf Funktionen, deren Definitions- und Wertebereich in der Menge der komplexen Zahlen liegt.

Definition 4.1. Sei $D \subset \mathbb{R}$ oder $D \subset \mathbb{C}$.

Eine Funktion $f : D \rightarrow \mathbb{C}$ heißt *komplexwertige Funktion*.

Eine Funktion $f : D \rightarrow \mathbb{R}$ heißt *reellwertige Funktion*.

Eine Funktion $f : D \subset \mathbb{C} \rightarrow \mathbb{C}$ heißt *komplexe Funktion*.

Eine Funktion $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ heißt *reelle Funktion*.

In diesem Kapitel bezeichnet D immer eine Teilmenge von \mathbb{R} oder \mathbb{C} . Offensichtlich sind reelle (bzw. reellwertige) Funktionen Spezialfälle von komplexen (bzw. komplexwertigen) Funktionen. Alle Aussagen für komplexe (bzw. komplexwertige) Funktionen gelten deshalb auch für reelle (bzw. reellwertige) Funktionen.

Mit komplexwertigen Funktionen kann man die gleichen algebraischen Operationen ausführen wie in den komplexen Zahlen, man wendet sie jeweils auf die Werte der Funktionen an.

Definition 4.2. Seien $f, g : D \rightarrow \mathbb{C}$ zwei komplexwertige Funktionen. Dann sind die Summe $f + g : D \rightarrow \mathbb{C}$, das Produkt $f \cdot g : D \rightarrow \mathbb{C}$, der Quotient $\frac{f}{g} : D \rightarrow \mathbb{C}$, falls $g(z) \neq 0$ für alle $z \in D$, der Betrag $|f| : D \rightarrow \mathbb{R}$, die konjugiert-komplexe Funktion $\bar{f} : D \rightarrow \mathbb{C}$, der Realteil $\operatorname{Re}(f) : D \rightarrow \mathbb{R}$ sowie der Imaginärteil $\operatorname{Im}(f) : D \rightarrow \mathbb{R}$ die Funktionen, die jedem $z \in D$ jeweils den folgenden Wert zuordnen:

$$(f + g)(z) := f(z) + g(z),$$

$$(f \cdot g)(z) := f(z) \cdot g(z),$$

$$\frac{f}{g}(z) := \frac{f(z)}{g(z)},$$

$$|f|(z) := |f(z)|,$$

$$\bar{f}(z) := \overline{f(z)},$$

$$\operatorname{Re}(f)(z) := \operatorname{Re}(f(z)),$$

$$\operatorname{Im}(f)(z) := \operatorname{Im}(f(z)).$$

Insbesondere gilt $f = \operatorname{Re}(f) + i \operatorname{Im}(f)$ und $|f|^2 = f \cdot \bar{f} = \operatorname{Re}(f)^2 + \operatorname{Im}(f)^2$.

4.1 Definition stetiger Funktionen und Beispiele

Definition 4.3. Eine Funktion $f : D \rightarrow \mathbb{C}$ heißt in $x_0 \in D$ stetig, wenn es zu jedem $\varepsilon > 0$ ein (von ε und x_0 abhängiges) $\delta = \delta(\varepsilon, x_0) > 0$ gibt, so dass gilt:

$$\forall x \in D \text{ mit } |x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon.$$

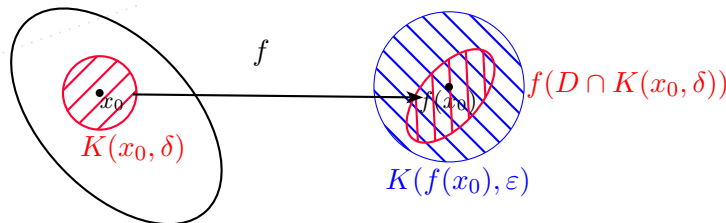
Eine Funktion $f : D \rightarrow \mathbb{C}$ heißt stetig auf D , wenn f in jedem Punkt $x_0 \in D$ stetig ist.

Geometrische Deutung der Stetigkeit:

1. Komplexe Funktionen:

Eine komplexe Funktion $f : D \subset \mathbb{C} \rightarrow \mathbb{C}$ ist genau dann in $x_0 \in D$ stetig, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass die offene Kreisscheibe um x_0 vom Radius δ durch f in die offene Kreisscheibe um $f(x_0)$ vom Radius ε abgebildet wird:

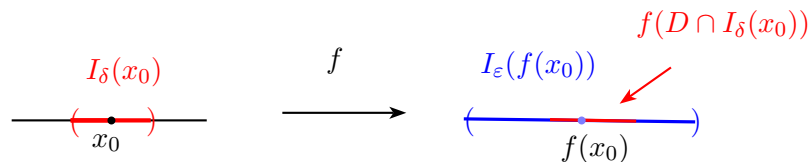
$$f(D \cap K(x_0, \delta)) \subset K(f(x_0), \varepsilon).$$



2. Reelle Funktionen:

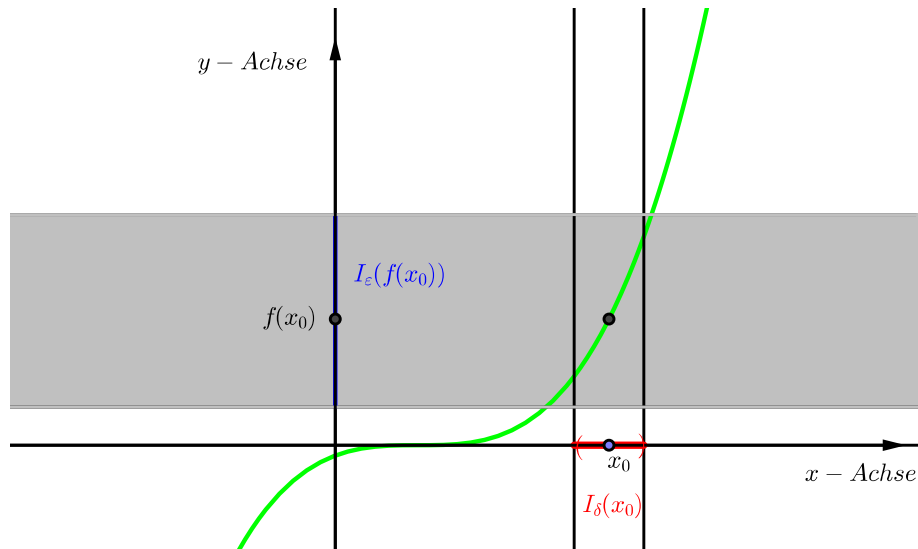
Eine reelle Funktion $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ ist genau dann in $x_0 \in D$ stetig, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass das offene Intervall $I_\delta(x_0) := (x_0 - \delta, x_0 + \delta)$ durch f in das offene Intervall $I_\varepsilon(f(x_0)) := (f(x_0) - \varepsilon, f(x_0) + \varepsilon)$ abgebildet wird:

$$f(D \cap (x_0 - \delta, x_0 + \delta)) \subset (f(x_0) - \varepsilon, f(x_0) + \varepsilon).$$



Dies kann man auch am Graphen der Funktion f ablesen: Der Graph von $f|_{I_\delta(x_0)}$ liegt im Streifen $D \times I_\varepsilon(f(x_0)) \subset D \times \mathbb{R}$:

$$\Gamma_{f|_{I_\delta(x_0)}} := \{(x, f(x)) \mid x \in I_\delta(x_0)\} \subset D \times I_\varepsilon(f(x_0)).$$



Satz 4.1 (Folgenkriterium für Stetigkeit)

Eine Funktion $f : D \rightarrow \mathbb{C}$ ist genau dann in $x_0 \in D$ stetig, wenn für jede Folge (x_n) in D , die gegen x_0 konvergiert, auch die Bildfolge $(f(x_n))$ gegen $f(x_0)$ konvergiert.

Beweis. 1. (\implies): Wir setzen voraus, dass $f : D \rightarrow \mathbb{C}$ in $x_0 \in D$ stetig ist und zeigen die Folgenbedingung des Satzes. Sei $\varepsilon > 0$. Nach Definition der Stetigkeit existiert ein $\delta > 0$, so dass $|f(x) - f(x_0)| < \varepsilon$ für alle $x \in D$ mit $|x - x_0| < \delta$. Sei nun (x_n) eine Folge in D , die gegen x_0 konvergiert. Nach Definition der Konvergenz gibt es ein $n_0 \in \mathbb{N}$, so dass $|x_n - x_0| < \delta$ für alle $n \geq n_0$. Nach Voraussetzung gilt dann $|f(x_n) - f(x_0)| < \varepsilon$ für alle $n \geq n_0$. Da ε beliebig war, bedeutet dies, dass $(f(x_n))$ gegen $f(x_0)$ konvergiert.

2. (\impliedby): Wir setzen voraus, dass die Folgenbedingung des Satzes gilt. Wir führen den Beweis der Stetigkeit von f in x_0 indirekt: Angenommen f ist in x_0 nicht stetig. Dann existiert ein $\varepsilon_0 > 0$, so dass für alle $\delta > 0$ ein $x_\delta \in D$ existiert mit $|x_\delta - x_0| < \delta$, aber $|f(x_\delta) - f(x_0)| \geq \varepsilon_0$. Wir betrachten nun speziell $\delta = \frac{1}{n}$ für $n \in \mathbb{N}$. Dann existiert ein $x_n \in D$ mit $|x_n - x_0| < \frac{1}{n}$, aber $|f(x_n) - f(x_0)| \geq \varepsilon_0$. Wir haben also eine Folge (x_n) in D gefunden, die gegen x_0 konvergiert, für die die Bildfolge $(f(x_n))$ aber nicht gegen $f(x_0)$ konvergieren kann. Dies ist ein Widerspruch zur Voraussetzung. Folglich war die Annahme falsch, f ist also in x_0 stetig. \square

Mit dem Folgenkriterium kann man die Stetigkeit einer Funktion leicht mit Hilfe der Grenzwertsätze für Folgen untersuchen. Wir zeigen dies an den folgenden Beispielen.

Beispiel 1: Seien $f, g : [0, 2] \subset \mathbb{R} \rightarrow \mathbb{R}$ die reellen Funktionen

$$f(x) := \begin{cases} 1 & \text{falls } x \in [0, 1], \\ 2 & \text{falls } x \in (1, 2]. \end{cases} \quad g(x) := \begin{cases} 1 & \text{falls } x \neq 1, \\ 2 & \text{falls } x = 1. \end{cases}$$

f und g sind in $x = 1$ nicht stetig. Wir betrachten dazu die spezielle Folge (x_n) mit $x_n := 1 + \frac{1}{n}$. Dann konvergiert (x_n) gegen 1, aber $f(x_n) = 2 \not\rightarrow f(1) = 1$ und $g(x_n) = 1 \not\rightarrow g(1) = 2$. Die Behauptung folgt aus Satz 4.1.

Beispiel 2: Die Identität $Id_{\mathbb{C}} : \mathbb{C} \rightarrow \mathbb{C}$ ist stetig.

Beweis: Für jede Folge (z_n) in \mathbb{C} mit $z_n \rightarrow z$ gilt $Id_{\mathbb{C}}(z_n) = z_n \rightarrow Id_{\mathbb{C}}(z) = z$.

Beispiel 3: Die konstante Abbildung $c_a : D \subset \mathbb{C} \rightarrow \mathbb{C}$, die jedem $z \in D$ die fixierte Zahl $a \in \mathbb{C}$ zuordnet, ist stetig. Ist nämlich (z_n) eine Folge in D , die gegen $z \in D$ konvergiert, so konvergiert $(c_a(z_n) = a)$ gegen $c_a(z) = a$.

Beispiel 4: *Polynome*

Seien $a_0, a_1, \dots, a_m \in \mathbb{C}$ und $P(z) := \sum_{k=0}^m a_k z^k = a_m z^m + a_{m-1} z^{m-1} + \dots + a_1 z + a_0$ ein Polynom. Dann ist die Funktion

$$P : \mathbb{C} \longrightarrow \mathbb{C} \\ z \longmapsto P(z) = a_m z^m + a_{m-1} z^{m-1} + \dots + a_1 z + a_0$$

auf \mathbb{C} stetig. Zum Beweis benutzen wir wieder die Grenzwertsätze für konvergente Folgen aus Kapitel 3.1 und das Folgenkriterium für die Stetigkeit. Sei (z_n) eine Folge komplexer Zahlen, die gegen $z \in \mathbb{C}$ konvergiert. Dann gilt $z_n^k \rightarrow z^k$ und somit $a_k z_n^k \rightarrow a_k z^k$ für alle $k \in \{0, 1, \dots, m\}$. Folglich gilt $P(z_n) \rightarrow P(z)$.

Beispiel 5: *Rationale Funktionen*

Es seien P und Q komplexe Polynome und $D \subset \mathbb{C}$ eine Teilmenge, die keine Nullstelle von Q enthält. Dann ist die *rationale Funktion*

$$R := \frac{P}{Q} : D \subset \mathbb{C} \longrightarrow \mathbb{C} \\ z \longmapsto \frac{P(z)}{Q(z)}$$

auf D stetig. Dies folgt wie in Beispiel 4 aus den Grenzwertsätzen für konvergente Folgen und dem Folgenkriterium für die Stetigkeit.

Beispiel 6: *Potenz- und Wurzelfunktionen*

Sei q eine rationale Zahl. Dann ist die Potenzfunktion

$$f : \mathbb{R}^+ \longrightarrow \mathbb{R} \\ x \longmapsto x^q$$

stetig. Insbesondere ist für jedes $m \in \mathbb{N}$ die Wurzelfunktion

$$\sqrt[m]{} : \mathbb{R}^+ \longrightarrow \mathbb{R} \\ x \longmapsto \sqrt[m]{x}$$

stetig. Zum Beweis benutzen wir die Übungsaufgabe 22: Ist (x_n) eine Folge in \mathbb{R}^+ , die gegen $x \in \mathbb{R}^+$ konvergiert, so konvergiert die Folge (x_n^q) gegen x^q .

Beispiel 7: Die komplexe Exponentialfunktion $\exp : \mathbb{C} \rightarrow \mathbb{C}$ ist stetig, denn aus $z_n \rightarrow z$ folgt $e^{z_n} \rightarrow e^z$ (Satz 3.30).

Beispiel 8: Die Betragsfunktion $|\cdot| : z \in \mathbb{C} \mapsto |z| \in \mathbb{R}$ ist stetig, denn aus $z_n \rightarrow z$ folgt $|z_n| \rightarrow |z|$ (Satz 3.3).

Satz 4.2 1. Seien $f, g : D \rightarrow \mathbb{C}$ in $x_0 \in D$ stetig. Dann sind auch die Funktionen $f + g$, $f \cdot g$, $\frac{f}{g}$ (falls g auf D keine Nullstelle hat), $|f|$ und \bar{f} in x_0 stetig.
 2. $f : D \rightarrow \mathbb{C}$ ist in $x_0 \in D$ genau dann stetig, wenn die beiden Funktionen $\operatorname{Re}(f)$ und $\operatorname{Im}(f)$ in $x_0 \in D$ stetig sind.
 3. Ist $f : D \rightarrow \mathbb{C}$ stetig und $\widehat{D} \subset D$ eine Teilmenge. Dann ist die Einschränkung $f|_{\widehat{D}} : \widehat{D} \rightarrow \mathbb{C}$ ebenfalls stetig.

Beweis. Für den Beweis aller drei Aussagen können wir das Folgenkriterium aus Satz 4.1 und die Grenzwertsätze für Folgen benutzen. Wir demonstrieren dies am Beweis der Stetigkeit von $f + g$ in x_0 . Die anderen Behauptungen zeigt man analog. Sei (x_n) eine Folge in D , die gegen x_0 konvergiert. Da f und g in x_0 stetig sind, konvergiert dann die Folge $(f(x_n))$ gegen $f(x_0)$ und die Folge $(g(x_n))$ gegen $g(x_0)$. Aus dem Grenzwertsatz für die Summe konvergenter Folgen erhält man

$$(f + g)(x_n) = f(x_n) + g(x_n) \rightarrow f(x_0) + g(x_0) = (f + g)(x_0).$$

Somit ist $f + g$ in x_0 stetig. □

Beispiel 9: Die reelle Exponentialfunktion $\exp|_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}$ ist stetig.

Satz 4.3 Seien $D, \widehat{D} \subset \mathbb{C}$ Teilmengen, $f : D \rightarrow \mathbb{C}$ und $g : \widehat{D} \rightarrow \mathbb{C}$ Funktionen mit $f(D) \subset \widehat{D}$ sowie $g \circ f : D \rightarrow \mathbb{C}$ die Verknüpfung von f und g :

$$(g \circ f)(x) := g(f(x)), \quad x \in D.$$

Dann gilt: Sind f in $x_0 \in D$ und g in $f(x_0)$ stetig, so ist die Verknüpfung $g \circ f$ in x_0 stetig.

Beweis. Wir benutzen zum Beweis das Folgenkriterium für die Stetigkeit. Sei (x_n) eine Folge in D , die gegen x_0 konvergiert. Da f in x_0 stetig ist, konvergiert die Folge $(f(x_n))$ gegen $f(x_0)$. Da g in $f(x_0)$ stetig ist, konvergiert die Folge $(g(f(x_n))) = (g \circ f)(x_n)$ gegen $g(f(x_0)) = (g \circ f)(x_0)$. Somit ist $g \circ f$ in x_0 stetig. □

Beispiel 10: Die Funktion $h : \mathbb{R} \setminus \{-1\} \rightarrow \mathbb{R}$, definiert durch $h(x) := e^{\frac{x^2}{x+1}}$, ist stetig, denn sie ist die Verknüpfung der stetigen Funktionen $h = \exp \circ f$ mit $f(x) = \frac{x^2}{x+1}$.

Als nächstes beweisen wir einen Satz über die Stetigkeit der Umkehrfunktion reeller Funktionen mit Hilfe der $\varepsilon - \delta$ - Definition.

Satz 4.4 (Stetigkeit der Umkehrfunktion)

Sei $I \subset \mathbb{R}$ ein beliebiges Intervall und $f : I \rightarrow \mathbb{R}$ streng monoton wachsend mit dem Bild $D = f(I)$. Dann ist $f : I \rightarrow D$ bijektiv und die Umkehrfunktion $f^{-1} : D \rightarrow I$ ist streng monoton wachsend und stetig.

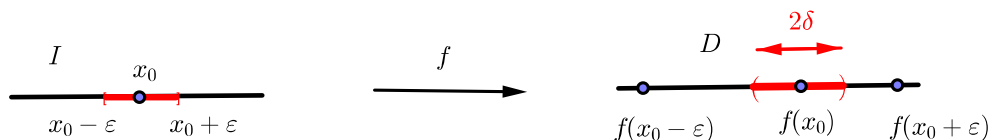
(Die analoge Aussage gilt auch für streng monoton fallende Funktionen f).

Beweis. a) Jede streng monoton wachsende Funktion ist injektiv. Da $D = f(I)$, ist $f : I \rightarrow D$ auch surjektiv, d.h. $f : I \rightarrow D$ ist bijektiv.

b) Es gelte $f(x) < f(y)$ für $x, y \in I$. Da f streng monoton wachsend ist, ist dies nur möglich, wenn $x < y$. Folglich gilt $x = f^{-1}(f(x)) < y = f^{-1}(f(y))$, d.h. f^{-1} ist ebenfalls streng monoton wachsend.

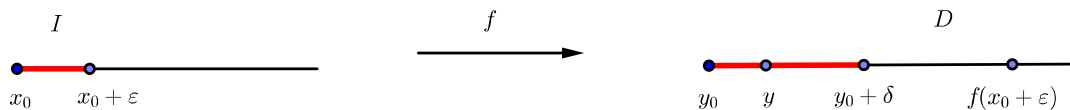
c) Wir zeigen, dass $f^{-1} : D \rightarrow I$ stetig ist: Sei $y_0 = f(x_0) \in D$ und $\varepsilon > 0$. Wir müssen ein $\delta > 0$ angeben, so dass $|f^{-1}(y) - f^{-1}(y_0)| = |f^{-1}(y) - x_0| < \varepsilon$ für alle $y \in D$ mit $|y - y_0| < \delta$.

1. *Fall:* x_0 ist kein Randpunkt des Intervalls I . Dann kann man oBdA annehmen, dass $[x_0 - \varepsilon, x_0 + \varepsilon] \subset I$ gilt (ggf. nach Verkleinern von ε). Wir wählen $\delta > 0$ so klein, dass $(y_0 - \delta, y_0 + \delta) \subset (f(x_0 - \varepsilon), f(x_0 + \varepsilon))$.



Für $y \in D$ mit $|y - y_0| < \delta$ gilt dann $f(x_0 - \varepsilon) < y < f(x_0 + \varepsilon)$. Da f^{-1} streng monoton wachsend ist, folgt $x_0 - \varepsilon < f^{-1}(y) < x_0 + \varepsilon$ und somit $|f^{-1}(y) - x_0| < \varepsilon$.

2. *Fall:* $x_0 \in I$ ist linker Randpunkt des Intervalls I , d.h. $x_0 = \min(I)$. Dann können wir oBdA annehmen, dass $[x_0, x_0 + \varepsilon] \subset I$ (ε ggf. verkleinern). Wir wählen nun $\delta > 0$ so klein, dass $[y_0, y_0 + \delta] \subset [y_0, f(x_0 + \varepsilon))$.



Für alle $y \in D$ mit $|y - y_0| < \delta$ gilt dann $y_0 = f(x_0) \leq y < f(x_0 + \varepsilon)$. Da f^{-1} streng monoton wachsend ist, folgt $x_0 \leq f^{-1}(y) < x_0 + \varepsilon$, d.h. $|f^{-1}(y) - x_0| < \varepsilon$.

3. *Fall:* x_0 ist rechter Randpunkt des Intervalls I , d.h. $x_0 = \max(I)$. Der Beweis in diesem Fall wird analog zu Fall 2 geführt. \square

Beispiel 11: Die Logarithmus-Funktion $\ln : \mathbb{R}^+ \rightarrow \mathbb{R}$ ist stetig.

Bemerkung zu Satz 4.4: Wir haben in Satz 4.4 nicht vorausgesetzt, dass f selbst stetig ist. Die Unstetigkeit von f führt zu einem unzusammenhängenden Definitionsbereich D

von f^{-1} , auf dem f^{-1} aber immer stetig ist. Satz 4.4 gilt nicht, wenn I kein Intervall ist. Als Beispiel betrachten wir die bijektive, streng monoton wachsende Funktion $f : [0, 1] \cup (2, 3] \rightarrow [0, 2]$ mit

$$f(x) := \begin{cases} x & \text{falls } x \in [0, 1] \\ x - 1 & \text{falls } x \in (2, 3]. \end{cases}$$

Die Umkehrfunktion $f^{-1} : [0, 2] \rightarrow [0, 1] \cup (2, 3]$ ist gegeben durch

$$f^{-1}(y) := \begin{cases} y & \text{falls } y \in [0, 1] \\ y + 1 & \text{falls } y \in (1, 2]. \end{cases}$$

f^{-1} ist in $y_0 = 1$ nicht stetig.

Die bisher definierte Stetigkeit beschreibt ein *lokales* Verhalten der Funktion f um jeden Punkt des Definitionsbereiches. Als nächstes definieren wir zwei Stetigkeitsbegriffe, die stärker als die gewöhnliche Stetigkeit sind und *globale* Eigenschaften von f beschreiben.

Definition 4.4.

1. Eine Funktion $f : D \rightarrow \mathbb{C}$ heißt *gleichmäßig stetig*, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass für alle $x, y \in D$ gilt:

$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

(Im Unterschied zur Definition der Stetigkeit hängt hier die Größe von δ nur von ε , aber nicht von x oder y ab.)

2. Eine Funktion $f : D \rightarrow \mathbb{C}$ heißt *lipschitzstetig*, wenn es eine positive Konstante $L \in \mathbb{R}^+$ gibt, so dass für alle $x, y \in D$ gilt:

$$|f(x) - f(y)| \leq L \cdot |x - y|.$$

L heißt *Lipschitz-Konstante* von f .

Satz 4.5 Sei $f : D \rightarrow \mathbb{C}$ eine Funktion. Dann gilt:

1. f ist lipschitzstetig $\implies f$ ist gleichmäßig stetig.
2. f ist gleichmäßig stetig $\implies f$ ist stetig.

Beweis. Stetigkeit folgt per Definition aus gleichmäßiger Stetigkeit. Wir müssen also nur zeigen, dass jede lipschitzstetige Funktion gleichmäßig stetig ist.

Sei f lipschitzstetig mit Lipschitz-Konstante L und $\varepsilon > 0$. Wir setzen $\delta := \frac{\varepsilon}{L}$. Seien nun $x, y \in D$ mit $|x - y| < \delta = \frac{\varepsilon}{L}$. Aus der Lipschitzstetigkeit folgt dann

$$|f(x) - f(y)| \leq L \cdot |x - y| < L \cdot \delta = \varepsilon.$$

Somit ist f gleichmäßig stetig. □

Die folgenden beiden Beispiele zeigen, dass die Umkehrungen der Aussagen von Satzes 4.5 nicht gelten.

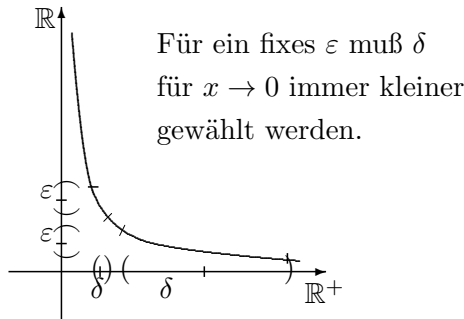
Beispiel 12:

Die Funktion

$$f : \mathbb{R}^+ \longrightarrow \mathbb{R}$$

$$x \longmapsto \frac{1}{x}$$

ist stetig, aber nicht gleichmäßig stetig.



Beweis. Sei $\delta > 0$ eine fixierte Zahl und $x \in \mathbb{R}^+$. Dann gilt

$$\left| f\left(x + \frac{\delta}{2}\right) - f(x) \right| = \left| \frac{1}{x + \frac{\delta}{2}} - \frac{1}{x} \right| = \frac{\delta}{2x\left(x + \frac{\delta}{2}\right)}. \quad (*)$$

Die rechte Seite von (*) konvergiert bei $x \rightarrow 0$ gegen $+\infty$. Man kann also für ein gegebenes $\varepsilon > 0$ kein $\delta > 0$ finden, so dass die rechte Seite von (*) für jedes $x > 0$ kleiner als ε bleibt. Folglich ist f nicht gleichmäßig stetig. □

Beispiel 13: Die Funktion $f : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$

$$x \longmapsto \sqrt{x}$$

ist gleichmäßig stetig, aber nicht lipschitzstetig.

Beweis. Nach Übungsaufgabe 15 gilt

$$|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x - y|} \quad \forall x, y \in \mathbb{R}^+.$$

Für $\varepsilon > 0$ setzen wir $\delta := \varepsilon^2$. Ist $|x - y| < \delta$, so folgt $|\sqrt{x} - \sqrt{y}| \leq \varepsilon$. Somit ist f gleichmäßig stetig.

Angenommen f wäre lipschitzstetig mit der Lipschitz-Konstanten L , das heißt

$$|\sqrt{x} - \sqrt{y}| \leq L|x - y| \quad \forall x, y \in \mathbb{R}^+.$$

Dann gilt

$$\left| \frac{\sqrt{x} - \sqrt{y}}{x - y} \right| = \frac{1}{\sqrt{x} + \sqrt{y}} \leq L \quad \forall x, y \in \mathbb{R}^+.$$

Für hinreichend kleine x und y kann man aber $\frac{1}{\sqrt{x} + \sqrt{y}}$ beliebig groß machen. Dies ergibt den Widerspruch. □

Beispiel 14: Die Betragsfunktion $|\cdot| : \mathbb{C} \rightarrow \mathbb{R}$ ist lipschitzstetig mit der Lipschitz-Konstanten $L = 1$, denn es gilt $||z| - |w|| \leq |z - w|$ für alle $z, w \in \mathbb{C}$ (Satz 2.14).

4.2 Grenzwerte von Funktionen und stetige Fortsetzungen

Wir wissen bereits, dass die Einschränkung einer stetigen Funktion auf eine Teilmenge wieder stetig ist. Wir stellen nun die umgekehrte Frage: Kann man eine stetige Funktion $f : D \rightarrow \mathbb{C}$ auf eine größere Teilmenge $\tilde{D} := D \cup \{a\}$ stetig fortsetzen, wobei a nicht zu D gehört. Dies geht offensichtlich nicht immer, wie die Funktionen f und g aus Beispiel 1 zeigen: Diese Funktionen sind auf $D = [0, 1) \cup (1, 2]$ stetig, aber nicht auf $\tilde{D} := [0, 2]$. Beim Fortsetzungsproblem in den Punkt a muß man zwei Fälle unterscheiden: a ist ein Häufungspunkt von D oder nicht.

Definition 4.5. Sei $D \subset \mathbb{C}$. Ein Punkt $a \in \mathbb{C}$ heißt Häufungspunkt von D , wenn es eine Folge (z_n) in $D \setminus \{a\}$ gibt, die gegen a konvergiert.

$HP(D)$ bezeichne der Menge der Häufungspunkte von D .

Beispiel 15: $HP((a, b)) = [a, b]$, $HP(K(z_0, R)) = cl K(z_0, R)$,
 $HP(\{\frac{1}{n} \mid n \in \mathbb{N}\}) = \{0\}$, $HP(\mathbb{Q}) = \mathbb{R}$.

1. Fall: Sei $a \in \mathbb{C}$ kein Häufungspunkt von D und $f : D \rightarrow \mathbb{C}$ stetig. Dann ist jede Fortsetzung $F : \tilde{D} := D \cup \{a\} \rightarrow \mathbb{C}$ von f stetig. Dies folgt aus dem Folgenkriterium für die Stetigkeit: Ist a kein Häufungspunkt von D , so sind für jede Folge (z_n) in \tilde{D} , die gegen a konvergiert, fast alle¹ Folgenglieder gleich a . Damit gilt $F(z_n) = F(a)$ für fast alle z_n . Die Folge $(F(z_n))$ konvergiert also gegen $F(a)$.

2. Fall: Sei $a \in \mathbb{C}$ ein Häufungspunkt von D und $F : \tilde{D} := D \cup \{a\} \rightarrow \mathbb{C}$ eine Fortsetzung der stetigen Funktion $f : D \rightarrow \mathbb{C}$. Dann ist F nach dem Folgenkriterium genau dann in a stetig, wenn für jede Folge (z_n) aus \tilde{D} , die gegen a konvergiert, die Bildfolge $(F(z_n))$ gegen den Wert $F(a)$ konvergiert. Dabei genügt es, Folgen (z_n) zu betrachten, die vollständig in $D \setminus \{a\}$ liegen, für die dann $(f(z_n) = F(z_n))$ gegen $F(a)$ konvergieren muß.

Dies führt auf die folgende Definition und das folgende Stetigkeitskriterium:

Definition 4.6. Eine Funktion $f : D \rightarrow \mathbb{C}$ hat im Häufungspunkt $a \in HP(D)$ den Grenzwert b , wenn für jede Folge (z_n) aus $D \setminus \{a\}$, die gegen a konvergiert, die Bildfolge $(f(z_n))$ gegen b konvergiert.

Bezeichnung: $\lim_{z \rightarrow a} f(z) = b$.

Aus dem Folgenkriterium für die Stetigkeit erhält man sofort:

Satz 4.6 Sei $f : D \rightarrow \mathbb{C}$ eine Funktion.

1. Ist $x_0 \in D$ ein Häufungspunkt von D , so ist $f : D \rightarrow \mathbb{C}$ genau dann in x_0 stetig, wenn f in x_0 einen Grenzwert besitzt und dieser gleich dem Funktionswert $f(x_0)$ ist:

$$f(x_0) = \lim_{z \rightarrow x_0} f(z).$$

¹ fast alle bedeutet: alle bis auf endlich viele.

2. Ist $a \in \text{HP}(D) \setminus D$, so besitzt f genau dann eine in a stetige Fortsetzung $F : D \cup \{a\} \rightarrow \mathbb{C}$, wenn f in a einen Grenzwert besitzt. In diesem Fall gilt:

$$F(z) := \begin{cases} f(z) & \text{falls } z \in D \\ \lim_{z \rightarrow a} f(z) & \text{falls } z = a. \end{cases}$$

Für Grenzwerte von Funktionen gelten die gleichen Rechenregeln wie für Grenzwerte konvergenter Folgen:

Satz 4.7 Sind $f, g : D \rightarrow \mathbb{C}$ Funktionen, $a \in \text{HP}(D)$ und existieren die Grenzwerte $\lim_{x \rightarrow a} f(x) = b$ und $\lim_{x \rightarrow a} g(x) = c$, dann existieren die Grenzwerte von $f + g$, $f \cdot g$, $\frac{f}{g}$ (falls g auf D keine Nullstelle hat und $c \neq 0$), $|f|$, \bar{f} , $\text{Re}(f)$ und $\text{Im}(f)$ im Punkt a und es gilt

$$\begin{aligned} \lim_{x \rightarrow a} (f(x) + g(x)) &= b + c, \\ \lim_{x \rightarrow a} f(x) \cdot g(x) &= b \cdot c, \\ \lim_{x \rightarrow a} \frac{f(x)}{g(x)} &= \frac{b}{c}, \\ \lim_{x \rightarrow a} |f(x)| &= |b|, \\ \lim_{x \rightarrow a} \overline{f(x)} &= \bar{b}, \\ \lim_{x \rightarrow a} \text{Re}(f(x)) &= \text{Re}(b), \\ \lim_{x \rightarrow a} \text{Im}(f(x)) &= \text{Im}(b). \end{aligned}$$

□

Wir betrachten nun den Spezialfall reeller Funktionen. In den reellen Zahlen hatten wir außer konvergenten Folgen auch Folgen betrachtet, die gegen $+\infty$ oder $-\infty$ streben und diesen den *uneigentlichen* Grenzwert $\pm\infty$ zugeordnet. Wir betrachten die analoge Situation für Grenzwerte von Funktionen $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ und vereinbaren analog zu Definition 4.6 die folgenden Bezeichnungen:

Definition 4.7. Sei $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ eine reelle Funktion, $a \in \text{HP}(D)$ und $b \in \mathbb{R} \cup \{\pm\infty\}$.

1. $\lim_{x \rightarrow a} f(x) = +\infty$: \iff für jede Folge (x_n) in $D \setminus \{a\}$ mit $x_n \rightarrow a$ gilt $f(x_n) \rightarrow +\infty$.
2. $\lim_{x \rightarrow a} f(x) = -\infty$: \iff für jede Folge (x_n) in $D \setminus \{a\}$ mit $x_n \rightarrow a$ gilt $f(x_n) \rightarrow -\infty$.
3. $\lim_{x \rightarrow +\infty} f(x) = b$: \iff für jede Folge (x_n) in D mit $x_n \rightarrow +\infty$ gilt $f(x_n) \rightarrow b$.
4. $\lim_{x \rightarrow -\infty} f(x) = b$: \iff für jede Folge (x_n) in D mit $x_n \rightarrow -\infty$ gilt $f(x_n) \rightarrow b$.

Dabei gelte in 3. $(c, +\infty) \subset D$ und in 4. $(-\infty, c) \subset D$ für ein $c \in \mathbb{R}$.

Beispiel 16: Seien $f_1, f_2 : (0, +\infty) \subset \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$f_1(x) := \frac{\cos x}{x}, \quad f_2(x) := \cos x$$

Dann gilt $\lim_{x \rightarrow +\infty} f_1(x) = 0$, während der Grenzwert von f_2 für $x \rightarrow +\infty$ nicht existiert.

Beweis. Die erste Behauptung folgt, da $|f_1(x)| \leq \frac{1}{x}$. Für die 2. Behauptung betrachte man die Folgen (x_n) mit $x_n := 2n\pi$ und (y_n) mit $y_n := (2n + 1)\pi$. Dann gilt $f_2(x_n) = 1 \rightarrow 1$ und $f_2(y_n) = -1 \rightarrow -1$. Folglich hat f_2 für $x \rightarrow +\infty$ keinen Grenzwert. \square

Beispiel 17: Es seien $\alpha \in \mathbb{R}$ und $\beta \in \mathbb{R}^+$. Dann gilt:

$$\lim_{x \rightarrow +\infty} \frac{e^x}{x^\alpha} = +\infty \quad \text{und} \quad \lim_{x \rightarrow +\infty} \frac{x^\beta}{\ln(x)} = +\infty.$$

D.h. die Exponentialfunktion wächst schneller als jede Potenz und jede Potenz mit positivem Exponenten wächst schneller als die Logarithmusfunktion.

Beweis. Sei (x_n) eine Folge in \mathbb{R}^+ mit $x_n \rightarrow +\infty$ und k eine fixierte natürliche Zahl mit $k \geq \alpha + 1$. Dann gilt:

$$e^{x_n} \geq \frac{x_n^k}{k!} \geq \frac{x_n^{\alpha+1}}{k!} = x_n^\alpha \cdot \frac{x_n}{k!}.$$

Folglich ist

$$\frac{e^{x_n}}{x_n^\alpha} \geq \frac{x_n}{k!}.$$

Da $x_n \rightarrow +\infty$ folgt $\frac{e^{x_n}}{x_n^\alpha} \rightarrow +\infty$.

Sei (y_n) eine Folge in \mathbb{R}^+ mit $y_n \rightarrow +\infty$. Da $\ln : \mathbb{R}^+ \rightarrow \mathbb{R}$ bijektiv und streng monoton wachsend ist, gilt $\ln(y_n) \rightarrow +\infty$. Da β positiv ist, folgt $x_n := \ln(y_n) \cdot \beta \rightarrow +\infty$. Wir erhalten aus der ersten Behauptung:

$$\frac{y_n^\beta}{\ln(y_n)} = \frac{e^{\ln(y_n) \cdot \beta}}{\ln(y_n)} = \beta \cdot \frac{e^{\ln(y_n) \cdot \beta}}{\ln(y_n) \cdot \beta} = \beta \cdot \frac{e^{x_n}}{x_n} \rightarrow +\infty.$$

\square

Im Fall von reellen Funktionen $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ kann man einseitige Grenzwerte definieren. Wir betrachten dazu die folgenden Teilmengen der Menge der Häufungspunkte von D :

$$HP(D)^- := \{a \in \mathbb{R} \mid \text{es existiert eine Folge } (x_n) \text{ in } D \text{ mit } x_n < a \text{ und } x_n \rightarrow a\}.$$

$$HP(D)^+ := \{a \in \mathbb{R} \mid \text{es existiert eine Folge } (x_n) \text{ in } D \text{ mit } x_n > a \text{ und } x_n \rightarrow a\}.$$

Definition 4.8. Sei $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ eine reelle Funktion und $b \in \mathbb{R} \cup \{\pm\infty\}$.

1. Man sagt, dass der linksseitige Grenzwert von f in $a \in HP(D)^-$ existiert und gleich b ist, falls für jede gegen a konvergente Folge (x_n) in D mit $x_n < a$ gilt: $\lim_{n \rightarrow \infty} f(x_n) = b$.
Wir bezeichnen den linksseitigen Grenzwert mit $\lim_{x \rightarrow a^-} f(x) = b$.
2. Man sagt, dass der rechtsseitige Grenzwert von f in $a \in HP(D)^+$ existiert und gleich b ist, falls für jede gegen a konvergente Folge (x_n) in D mit $x_n > a$ gilt: $\lim_{n \rightarrow \infty} f(x_n) = b$.
Wir bezeichnen den rechtsseitigen Grenzwert mit $\lim_{x \rightarrow a^+} f(x) = b$.

Wir können die Stetigkeit reeller Funktionen durch die einseitigen Grenzwerte charakterisieren:

Satz 4.8 Sei $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ eine reelle Funktion und $x_0 \in (c, d) \subset D$. Dann ist f in x_0 genau dann stetig, wenn die beiden einseitigen Grenzwerte von f im Punkt x_0 in \mathbb{R} existieren und gleich dem Funktionswert $f(x_0)$ sind:

$$\lim_{x \rightarrow x_0^-} f(x) = \lim_{x \rightarrow x_0^+} f(x) = f(x_0).$$

Beweis. (\implies) folgt aus Satz 4.1 als Spezialfall.

(\impliedby) Wir setzen voraus, dass die beiden einseitigen Grenzwerte existieren und dass gilt

$$\lim_{x \rightarrow x_0^-} f(x) = \lim_{x \rightarrow x_0^+} f(x) = f(x_0).$$

Angenommen, f sei in x_0 nicht stetig. Dann existiert nach dem Folgenkriterium eine Folge (x_n) in D , die gegen x_0 konvergiert, deren Bildfolge aber nicht gegen $f(x_0)$ konvergiert. Da dann fast alle Folgenglieder x_n von x_0 verschieden sind, existiert eine Teilfolge (x_{n_k}) von (x_n) mit $x_{n_k} < x_0$ oder eine Teilfolge (x_{n_k}) von (x_n) mit $x_{n_k} > x_0$, deren Bildfolge $(f(x_{n_k}))$ nicht gegen $f(x_0)$ konvergiert. Dies widerspricht der Voraussetzung $\lim_{x \rightarrow x_0^-} f(x) = \lim_{x \rightarrow x_0^+} f(x) = f(x_0)$. \square

Satz 4.8 zeigt, dass es genau drei verschiedene Typen von Unstetigkeitsstellen einer reellen Funktion gibt. Die Funktion $f : D \subset \mathbb{R} \rightarrow \mathbb{R}$ ist genau dann in $x_0 \in (c, d) \subset D$ *unstetig*, wenn einer der folgenden Fälle vorliegt:

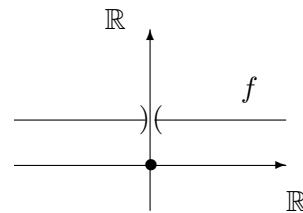
(1) *Hebbare Unstetigkeitsstelle*

Beide einseitigen Grenzwerte von f in x_0 existieren in \mathbb{R} und stimmen überein, sind aber ungleich $f(x_0)$. In diesem Fall kann man die Unstetigkeit von f in x_0 durch Abänderung von $f(x_0)$ beheben.

Beispiel:

$$f(x) := \begin{cases} 0 & x = 0 \\ 1 & x \neq 0 \end{cases}$$

Der Punkt $x_0 = 0$ ist eine hebbare Unstetigkeitsstelle.



(2) *Sprungstelle*

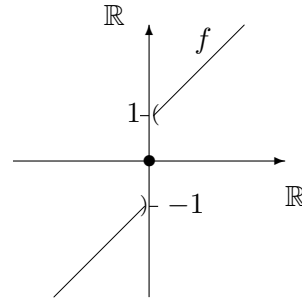
Die beiden einseitigen Grenzwerte von f in x_0 existieren in \mathbb{R} , sind aber voneinander verschieden. Unter dem *Sprung* $\sigma(f, x_0)$ von f in x_0 verstehen wir die Differenz der einseitigen Grenzwerte

$$\sigma(f, x_0) := \lim_{x \rightarrow x_0^+} f(x) - \lim_{x \rightarrow x_0^-} f(x).$$

Beispiel:

$$f(x) := \begin{cases} x + \frac{x}{|x|} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Dann ist der Punkt $x_0 = 0$ eine Sprungstelle und der Sprung $\sigma(f, x_0) = 2$.



(3) Unstetigkeitsstelle 2. Art²

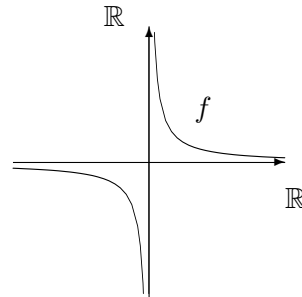
Mindestens einer der beiden einseitigen Grenzwerte von f in x_0 existiert nicht in \mathbb{R} . Die folgenden Beispiele zeigen die beiden typischen Fälle für diese Situation: Als erstes betrachten wir die Dirichlet-Funktion

$$h(x) := \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

Jeder Punkt $x_0 \in \mathbb{R}$ ist eine Unstetigkeitsstelle zweiter Art von h , da h in x_0 keinen rechtsseitigen und keinen linksseitigen Grenzwert besitzt. Das nächste Beispiel zeigt einen Fall mit uneigentlichen Grenzwerten.

Sei
$$f(x) := \begin{cases} \frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

Der Punkt $x_0 = 0$ ist eine Unstetigkeitsstelle zweiter Art, denn $\lim_{x \rightarrow 0^+} f(x) = +\infty$ und $\lim_{x \rightarrow 0^-} f(x) = -\infty$.



Als Spezialfall betrachten wir monotone Funktionen.

Satz 4.9 Sei $f : (\alpha, \beta) \subset \mathbb{R} \rightarrow \mathbb{R}$ eine monotone Funktion und $-\infty \leq \alpha < \beta \leq +\infty$. Dann gilt:

1. Für jedes $x_0 \in (\alpha, \beta)$ existieren die einseitigen Grenzwerte $\lim_{x \rightarrow x_0^-} f(x)$ und $\lim_{x \rightarrow x_0^+} f(x)$ in \mathbb{R} .
2. f ist in $x_0 \in (\alpha, \beta)$ genau dann stetig, wenn $\lim_{x \rightarrow x_0^+} f(x) = \lim_{x \rightarrow x_0^-} f(x)$.
3. f hat höchstens abzählbar viele Unstetigkeitsstellen. Jede Unstetigkeitsstelle ist eine Sprungstelle.

Beweis. Übungsaufgabe. □

² Hebbare Unstetigkeitsstellen und Sprungstellen nennt man auch Unstetigkeitsstellen 1. Art.

4.3 Nullstellen, Fixpunkte und Extremwerte stetiger Funktionen

In diesem Abschnitt wollen wir uns mit einigen speziellen Eigenschaften stetiger Funktionen befassen. Insbesondere wollen wir Bedingungen dafür angeben, dass eine stetige Funktion Nullstellen, Fixpunkte bzw. Extremwerte besitzt.

Definition 4.9. Sei $D \subset \mathbb{R}$ oder $D \subset \mathbb{C}$.

1. Ein Punkt $\xi \in D$ heißt Nullstelle einer Funktion $f : D \rightarrow \mathbb{C}$, wenn $f(\xi) = 0$.
2. Ein Punkt $\xi \in D$ heißt Fixpunkt einer Funktion $f : D \rightarrow \mathbb{C}$, wenn $f(\xi) = \xi$.
3. Eine Funktion $f : D \rightarrow \mathbb{R}$ nimmt auf D ein Minimum bzw. Maximum an, wenn es ein $\xi_1 \in D$ bzw. ein $\xi_2 \in D$ gibt, so dass

$$f(\xi_1) \leq f(x) \leq f(\xi_2) \quad \forall x \in D,$$

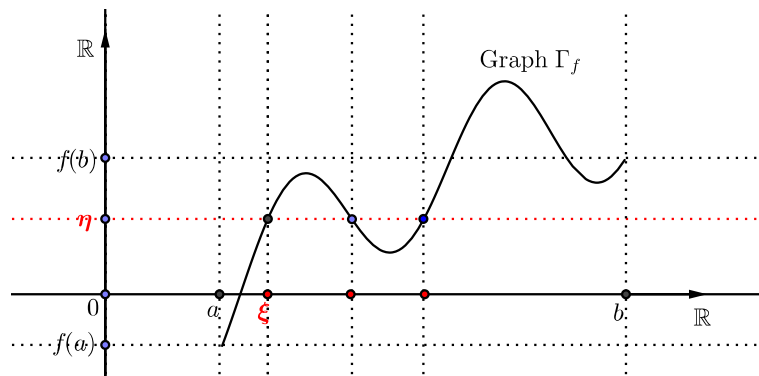
$$\text{d.h. } f(\xi_1) = \min\{f(x) \mid x \in D\} \text{ bzw. } f(\xi_2) = \max\{f(x) \mid x \in D\}.$$

Mit Hilfe des folgenden Satzes kann man in vielen Fällen die Existenz von Nullstellen oder Fixpunkten beweisen:

Satz 4.10 (Zwischenwertsatz)

Sei $f : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$ stetig und $\eta \in \mathbb{R}$ eine Zahl zwischen $f(a)$ und $f(b)$. Dann existiert ein $\xi \in [a, b]$ mit $f(\xi) = \eta$.

Beweis. OBdA. sei $f(a) \leq f(b)$.



Wir konstruieren induktiv eine Intervallschachtelung $\{I_n = [a_n, b_n]\}_{n \in \mathbb{N}_0}$ mit

$$f(a_n) \leq \eta \leq f(b_n) \quad \text{für alle } n \in \mathbb{N}_0.$$

Wir setzen $I_0 := [a, b]$. Dann gilt nach Voraussetzung $f(a_0) \leq \eta \leq f(b_0)$.

Sei das Intervall $I_n := [a_n, b_n]$ mit $f(a_n) \leq \eta \leq f(b_n)$ schon konstruiert. Wir bilden dann I_{n+1} durch Halbierung von I_n : Sei $M_n := \frac{a_n + b_n}{2}$ der Mittelpunkt von $[a_n, b_n]$. Dann gilt entweder $\eta \leq f(M_n)$ oder $\eta > f(M_n)$. Wir setzen

$$I_{n+1} := \begin{cases} [a_n, M_n], & \text{falls } \eta \leq f(M_n), \\ [M_n, b_n], & \text{falls } \eta > f(M_n), \end{cases}$$

und erhalten nach Konstruktion $f(a_{n+1}) \leq \eta \leq f(b_{n+1})$. Die Familie der abgeschlossenen Intervalle $\{I_n\}$ ist tatsächlich eine Intervallschachtelung, da nach Konstruktion

- a) $I_{n+1} \subset I_n$ für alle $n \in \mathbb{N}_0$.
 b) $L(I_n) = \frac{1}{2}L(I_{n-1}) = \frac{1}{2^n}L(I_0) = \frac{1}{2^n}(b-a)$ für alle $n \in \mathbb{N}_0$.

Nach dem Intervallschachtelungsprinzip (Satz 2.4) existiert genau eine reelle Zahl ξ mit $a_n \leq \xi \leq b_n$ für alle $n \in \mathbb{N}_0$. Da $|\xi - a_n| \leq |b_n - a_n| = L(I_n) \rightarrow 0$ und $|b_n - \xi| \leq |b_n - a_n| = L(I_n) \rightarrow 0$, konvergiert sowohl die Folge (a_n) als auch die Folge (b_n) gegen $\xi \in [a, b]$. Da f auf $[a, b]$ stetig ist, folgt mit dem Folgenkriterium der Stetigkeit $f(a_n) \rightarrow f(\xi)$ und $f(b_n) \rightarrow f(\xi)$. Da $f(a_n) \leq \eta \leq f(b_n)$ liefert das Sandwich-Lemma: $f(\xi) = \eta$. \square

Folgerung 4.1 1. Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und $f(a) \cdot f(b) < 0$, d.h. $f(a)$ und $f(b)$ haben verschiedene Vorzeichen. Dann besitzt f eine Nullstelle in $[a, b]$.

2. Sei $f : [a, b] \rightarrow [a, b]$ stetig. Dann besitzt f einen Fixpunkt in $[a, b]$.

Beweis. Die erste Behauptung folgt unmittelbar aus Satz 4.10 mit $\eta = 0$. Zum Beweis der zweiten Behauptung betrachten wir die Funktion $g : [a, b] \rightarrow \mathbb{R}$ definiert durch $g(x) := f(x) - x$ für alle $x \in [a, b]$. Dann ist g stetig und $g(a) = f(a) - a \geq 0$ und $g(b) = f(b) - b \leq 0$. Dann sind a oder b Fixpunkte von f oder es gilt $g(a) > 0$ und $g(b) < 0$. Im zweiten Fall hat g eine Nullstelle in (a, b) , d.h. es existiert ein $\xi \in (a, b)$ mit $f(\xi) = \xi$. \square

Eine stetige Funktion $f : D \rightarrow \mathbb{R}$ muß keine Extremwerte, d.h. Maxima oder Minima, besitzen. Auf einer bestimmten Sorte von Definitionsbereichen D , den kompakten Mengen, die wir jetzt definieren, ist dies aber immer der Fall.

Definition 4.10. Sei $A \subset \mathbb{R}$ oder $A \subset \mathbb{C}$.

1. A heißt abgeschlossen, wenn für jede konvergente Folge (a_n) aus A der Grenzwert ebenfalls in A liegt.
2. A heißt beschränkt, wenn es ein $C \in \mathbb{R}^+$ gibt mit $|a| \leq C$ für alle $a \in A$.
3. A heißt kompakt³, wenn A beschränkt und abgeschlossen ist.

Beispiele:

- Jedes abgeschlossene Intervall $[a, b] \subset \mathbb{R}$ ist kompakt.
- Die abgeschlossene Kreisscheibe $cl K(z_0, R) := \{z \in \mathbb{C} \mid |z - z_0| \leq R\} \subset \mathbb{C}$ ist kompakt.
- Die Kreislinie $S^1 := \{z \in \mathbb{C} \mid |z| = 1\} \subset \mathbb{C}$ ist kompakt.

Satz 4.11 (Folgenkriterium für Kompaktheit)

Eine Teilmenge $K \subset \mathbb{R}$ oder $K \subset \mathbb{C}$ ist genau dann kompakt, wenn jede Folge (z_n) aus K eine Teilfolge besitzt, die gegen ein Element aus K konvergiert.

³ Abgeschlossene, beschränkte und kompakte Mengen definiert man auch in metrischen Räumen, die die Räume $(\mathbb{R}, |\cdot|)$ und $(\mathbb{C}, |\cdot|)$ verallgemeinern. In metrischen Räumen ist jede kompakte Menge beschränkt und abgeschlossen, die Umkehrung gilt aber im Allgemeinen nicht!

Beweis. (\implies): Sei K kompakt und (z_n) eine Folge in K . Da K beschränkt ist, existiert ein $C \in \mathbb{R}^+$ mit $|z_n| \leq C$ für alle $n \in \mathbb{N}$. Ist $K \subset \mathbb{R}$, so besagt der Satz von Bolzano-Weierstraß, dass (z_n) eine konvergente Teilfolge (z_{n_k}) besitzt. Ist $K \subset \mathbb{C}$, aber $K \not\subset \mathbb{R}$, so betrachten wir die Folge der Realteile $(\operatorname{Re}(z_n))$ und der Imaginärteile $(\operatorname{Im}(z_n))$. Diese Folgen sind dann ebenfalls beschränkt und der Satz von Bolzano-Weierstraß liefert eine Teilfolge (z_{n_k}) von (z_n) , für die $(\operatorname{Re}(z_{n_k}))$ und $(\operatorname{Im}(z_{n_k}))$ konvergieren. Dann konvergiert aber auch die Teilfolge (z_{n_k}) in \mathbb{C} . Da K abgeschlossen ist, liegt der Grenzwert von (z_{n_k}) in K .

(\impliedby): Wir setzen voraus, dass jede Folge aus K eine Teilfolge besitzt, die gegen ein Element von K konvergiert und zeigen, dass K kompakt ist:

- a) K ist beschränkt: Angenommen, K wäre nicht beschränkt. Dann gibt es eine Folge (z_n) in K mit $|z_n| > n$ für alle $n \in \mathbb{N}$. Diese Folge (z_n) enthält keine konvergente Teilfolge. Dies ist ein Widerspruch zur Voraussetzung.
- b) K ist abgeschlossen: Sei (z_n) eine Folge in K mit $z_n \rightarrow z$. Nach Voraussetzung existiert eine Teilfolge (z_{n_k}) von (z_n) mit Grenzwert in K . Dann gilt $z = \lim_{n \rightarrow \infty} z_n = \lim_{k \rightarrow \infty} z_{n_k}$ und somit $z \in K$. \square

Satz 4.12 Sei $f : K \rightarrow \mathbb{C}$ eine stetige Funktion auf einer kompakten Menge K . Dann ist das Bild $f(K) \subset \mathbb{C}$ ebenfalls kompakt.

Beweis. Sei $(f(z_n))$ eine beliebige Folge in $f(K)$. Dann ist (z_n) eine Folge in K . Da K kompakt ist, existiert eine Teilfolge (z_{n_k}) von (z_n) , die gegen ein Element $\xi \in K$ konvergiert. Da f stetig ist, konvergiert $(f(z_{n_k}))$ gegen $f(\xi) \in f(K)$. Folglich ist $f(K)$ kompakt. \square

Satz 4.13 Sei $f : K \rightarrow \mathbb{R}$ eine stetige Funktion auf einer kompakten Menge K . Dann nimmt f auf K ein Minimum und ein Maximum an, d.h. es existieren $\xi_1, \xi_2 \in K$ mit

$$f(\xi_1) \leq f(x) \leq f(\xi_2) \quad \text{für alle } x \in K.$$

Beweis. Nach Satz 4.12 ist $f(K) \subset \mathbb{R}$ kompakt, also beschränkt, d.h. es existieren $\sup f(K)$ und $\inf f(K)$. Nach Definition von Supremum und Infimum existieren Folgen (x_n) und (y_n) in K mit $f(x_n) \rightarrow \inf f(K)$ und $f(y_n) \rightarrow \sup f(K)$. Da $f(K)$ abgeschlossen ist, folgt $\inf f(K) \in f(K)$ und $\sup f(K) \in f(K)$. Also existieren $\xi_1 \in K$ und $\xi_2 \in K$, so dass $f(\xi_1) = \inf f(K) = \min f(K)$ und $f(\xi_2) = \sup f(K) = \max f(K)$. \square

Satz 4.14 Sei $f : K \rightarrow \mathbb{C}$ eine stetige Abbildung auf einer kompakten Menge K . Dann ist f sogar gleichmäßig stetig.

Beweis. Angenommen, f wäre nicht gleichmäßig stetig. Dann existiert ein $\varepsilon_0 > 0$ ohne geeignetes δ (aus der Definition der gleichmäßigen Stetigkeit). Für alle $n \in \mathbb{N}$ gibt es also ein Paar von Zahlen $x_n, y_n \in K$, so dass

$$|x_n - y_n| < \frac{1}{n} \quad \text{und} \quad |f(x_n) - f(y_n)| \geq \varepsilon_0. \quad (4.1)$$

Da K kompakt ist, existiert eine Teilfolge (y_{n_k}) von (y_n) , die gegen ein Element $\xi \in K$ konvergiert. Dann gilt

$$|x_{n_k} - \xi| \leq |x_{n_k} - y_{n_k}| + |y_{n_k} - \xi| < \frac{1}{n_k} + |y_{n_k} - \xi|.$$

Folglich konvergiert (x_{n_k}) ebenfalls gegen ξ . Die Stetigkeit von f liefert $\lim_{k \rightarrow \infty} f(y_{n_k}) = f(\xi) = \lim_{k \rightarrow \infty} f(x_{n_k})$. Dann existiert aber ein $k_0 \in \mathbb{N}$, so dass

$$|f(x_{n_k}) - f(y_{n_k})| < \varepsilon_0 \quad \forall k \geq k_0.$$

Dies widerspricht (4.1). □

4.4 Anwendung: Der Fundamentalsatz der Algebra

Als Anwendung von Satz 4.13 wollen wir nun den in Kapitel 2.2 angekündigten Fundamentalsatz der Algebra beweisen, der weitreichende Konsequenzen sowohl in der Algebra als auch in vielen Bereichen der Analysis hat.

Satz 4.15 (Fundamentalsatz der Algebra)

Jedes nicht-konstante komplexe Polynom besitzt eine Nullstelle.

Beweis. Sei $P(z) := a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0$ ein komplexes Polynom mit $a_n, a_{n-1}, \dots, a_0 \in \mathbb{C}$, $n \geq 1$ und $a_n \neq 0$.

1. Schritt: Wir zeigen, dass die Abbildung $|P| : \mathbb{C} \rightarrow \mathbb{R}$ auf \mathbb{C} ein Minimum annimmt, d.h. es existiert ein $\xi \in \mathbb{C}$ mit $|P(\xi)| \leq |P(z)|$ für alle $z \in \mathbb{C}$.

Wir definieren dazu $Q(z)$ durch

$$P(z) = a_n z^n \left(1 + \underbrace{\frac{a_{n-1}}{a_n} \frac{1}{z} + \dots + \frac{a_0}{a_n} \frac{1}{z^n}}_{=: Q(z)} \right), \quad z \neq 0.$$

Sei nun $\alpha := \left| \frac{a_{n-1}}{a_n} \right| + \dots + \left| \frac{a_0}{a_n} \right|$ und $\beta := \max\{1, 2\alpha\}$. Für $|z| \geq \beta$ folgt dann

$$|Q(z)| \leq \left| \frac{a_{n-1}}{a_n} \right| \frac{1}{|z|} + \dots + \left| \frac{a_0}{a_n} \right| \frac{1}{|z|^n} \leq \left(\left| \frac{a_{n-1}}{a_n} \right| + \dots + \left| \frac{a_0}{a_n} \right| \right) \frac{1}{|z|} = \alpha \cdot \frac{1}{|z|} \leq \frac{1}{2}.$$

Dies zeigt, dass

$$|1 + Q(z)| \geq ||1| - |Q(z)|| = 1 - |Q(z)| \geq \frac{1}{2} \quad \text{für alle } |z| \geq \beta.$$

Wir setzen $r := \max\{\beta, (2\alpha)^{\frac{1}{n}}\}$. Dann folgt

$$|P(z)| = |a_n z^n| \cdot |1 + Q(z)| \geq |a_n| \cdot |z|^n \cdot \frac{1}{2} \geq |a_n| \cdot \alpha \quad \text{für alle } |z| \geq r.$$

Für alle $z \in \mathbb{C}$ außerhalb der abgeschlossenen Kreisscheibe $cl K(0, r)$ gilt also $|P(z)| \geq |a_n| \cdot \alpha$. Da abgeschlossene Kreisscheiben kompakt sind, nimmt $|P|$ auf $cl K(0, r)$ ein Minimum an. Da $|P(0)| = |a_0| \leq |a_n| \cdot \alpha$, ist dieses Minimum auch das Minimum von $|P|$ auf ganz \mathbb{C} . Damit ist gezeigt, dass ein $\xi \in \mathbb{C}$ existiert mit $|P(\xi)| = \min\{|P(z)| \mid z \in \mathbb{C}\}$.

2. Schritt: Wir zeigen, dass $P(\xi) = 0$.

Angenommen $P(\xi) \neq 0$. Wir betrachten das Polynom n -ten Grades

$$H(z) := \frac{P(z + \xi)}{P(\xi)}.$$

Da $|P|$ in ξ ein Minimum annimmt, gilt $|H(z)| \geq 1$ für alle $z \in \mathbb{C}$. Wegen $H(0) = 1$, hat H die Form

$$H(z) = b_n z^n + \dots + b_m z^m + 1,$$

wobei b_m, b_{m+1}, \dots, b_n komplexe Zahlen sind, $n \geq m \geq 1$ und $b_m \neq 0$. Die komplexe Zahl $-\frac{|b_m|}{b_m}$ hat den Betrag 1 und besitzt nach Satz 2.15 eine m -te Wurzel $w \in \mathbb{C}$, d.h. $w^m = -\frac{|b_m|}{b_m}$. Sei $w = \cos(\psi) + i \sin(\psi)$ die trigonometrische Darstellung von w . Dann ist $b_m (\cos(m\psi) + i \sin(m\psi)) = -|b_m|$. Wir betrachten nun H auf den komplexen Zahlen der Form $z = \varrho \cdot (\cos(\psi) + i \sin(\psi))$ mit $\varrho > 0$. Dann gilt

$$\left| H(\varrho \cdot (\cos(\psi) + i \sin(\psi))) \right| \leq |b_n| \cdot \varrho^n + \dots + |b_{m+1}| \cdot \varrho^{m+1} + \underbrace{|b_m \varrho^m \cdot (\cos(m\psi) + i \sin(m\psi)) + 1|}_{=|1 - |b_m| \cdot \varrho^m|}.$$

Sei nun ϱ so klein gewählt, dass $\varrho^m < \frac{1}{|b_m|}$ ist. Dann ist $1 - |b_m| \varrho^m > 0$ und damit

$$\begin{aligned} \left| H(\varrho \cdot (\cos(\psi) + i \sin(\psi))) \right| &\leq 1 - |b_m| \cdot \varrho^m + |b_{m+1}| \cdot \varrho^{m+1} + \dots + |b_n| \cdot \varrho^n \\ &= 1 - \varrho^m \underbrace{(|b_m| - |b_{m+1}| \cdot \varrho - \dots - |b_n| \cdot \varrho^{n-m})}_{>0 \text{ für } \varrho \text{ hinreichend klein}}. \end{aligned}$$

Somit ist $|H(\varrho \cdot (\cos(\psi) + i \sin(\psi)))| < 1$ für hinreichend kleine ϱ . Dies ist ein Widerspruch zu $|H(z)| \geq 1$ für alle $z \in \mathbb{C}$. Somit war unsere Annahme falsch und es gilt $P(\xi) = 0$. \square

Satz 4.16 (Zerlegungssatz für komplexe Polynome)

Sei $P(z) = a_n z^n + \dots + a_1 z + a_0$ ein komplexes Polynom vom Grad $n \geq 1$. Dann existieren n Nullstellen $\xi_1, \dots, \xi_n \in \mathbb{C}$ von P und es gilt

$$P(z) = a_n (z - \xi_1) \cdot (z - \xi_2) \cdot \dots \cdot (z - \xi_n),$$

wobei die Nullstellen ξ_1, \dots, ξ_n nicht alle verschieden sein müssen.

Beweis. Sei $\xi_1 \in \mathbb{C}$ eine Nullstelle von P , d.h. $P(\xi_1) = 0$. Wir wollen von P den Linearfaktor $(z - \xi_1)$ abspalten. Dazu betrachten wir die Polynome

$$Q_k(z) := z^{k-1} + z^{k-2} \cdot \xi_1 + z^{k-3} \cdot \xi_1^2 + \dots + z \cdot \xi_1^{k-2} + \xi_1^{k-1}.$$

Dann gilt $z^k - \xi_1^k = (z - \xi_1) \cdot Q_k(z)$ und folglich

$$\begin{aligned}
 P(z) &= P(z) - P(\xi_1) = a_1(z - \xi_1) + a_2(z^2 - \xi_1^2) + \dots + a_n(z^n - \xi_1^n) \\
 &= (z - \xi_1) \cdot (a_1Q_1(z) + a_2Q_2(z) + \dots + a_nQ_n(z)) \\
 &=: (z - \xi_1) \cdot P_1(z).
 \end{aligned}$$

$P_1(z)$ ist dabei ein Polynom vom Grad $n - 1$. Ist $n - 1 \geq 1$, so hat P_1 eine Nullstelle und wir können einen weiteren Linearfaktor von P_1 abspalten. Dieses Verfahren funktioniert n -mal. □

4.5 Die trigonometrischen und die Hyperbelfunktionen und ihre Umkehrfunktionen

Mit Hilfe der Exponentialfunktion definieren wir jetzt die trigonometrischen Funktionen sowie die Hyperbelfunktionen. Wir erinnern nochmal an die Definition der Exponentialfunktion:

$$\begin{aligned}
 \exp : \mathbb{C} &\longrightarrow \mathbb{C} \\
 z &\longmapsto e^z := \sum_{n=0}^{\infty} \frac{z^n}{n!}.
 \end{aligned}$$

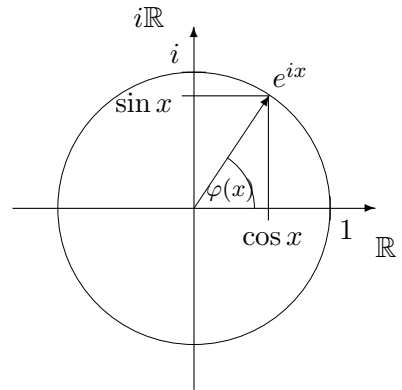
Diese Funktion ist auf \mathbb{C} stetig und erfüllt $e^0 = 1$ sowie $e^{z+w} = e^z \cdot e^w$.

4.5.1 Die trigonometrischen Funktionen

Für $x \in \mathbb{R}$ gilt

$$|e^{ix}|^2 = e^{ix} \cdot \overline{e^{ix}} = e^{ix} \cdot e^{-ix} = e^0 = 1.$$

Die komplexe Zahl e^{ix} liegt also auf der Kreislinie $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$. $\varphi(x)$ bezeichne den orientierten Winkel zwischen der reellen Achse und dem Strahl von 0 durch e^{ix} (im Gradmaß).



Definition 4.11 (Analytische Definition der Funktionen sin und cos).

Den Realteil von e^{ix} nennen wir *Cosinus* der reellen Zahl x , den Imaginärteil von e^{ix} nennen wir *Sinus* von x , d.h.

$$\begin{aligned}
 \cos(x) &:= \operatorname{Re}(e^{ix}) = \frac{e^{ix} + e^{-ix}}{2} = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \\
 \sin(x) &:= \operatorname{Im}(e^{ix}) = \frac{e^{ix} - e^{-ix}}{2i} = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}.
 \end{aligned}$$

Dann erhält man unmittelbar die **Euler-Formel**

$$e^{ix} = \cos(x) + i \sin(x) \quad \forall x \in \mathbb{R}.$$

Wir hatten in Kapitel 2 bereits die elementargeometrisch definierten trigonometrischen Funktionen **sin** und **cos** benutzt. Die Definition 4.11 zeigt, dass $\sin(\varphi(x)) = \sin(x)$ und $\cos(\varphi(x)) = \cos(x)$. (Man nennt x auch das Bogenmaß des orientierten Winkels zwischen der reellen Achse und dem Strahl von 0 durch e^{ix} , denn wie wir noch sehen werden, ist x die Länge der Kreisbogens von 1 nach e^{ix} . Im folgenden werden wir die analytische Definition der trigonometrischen Funktionen \sin und \cos benutzen.

Durch Einsetzen von e^{ix} in die Definitionen erhält man sofort die folgenden Eigenschaften und Rechenregeln für Sinus und Cosinus:

1. Die reellen Funktionen $\cos : \mathbb{R} \rightarrow \mathbb{R}$ und $\sin : \mathbb{R} \rightarrow \mathbb{R}$ sind stetig.
2. \cos ist eine *gerade Funktion*, d.h. es gilt $\cos(-x) = \cos(x)$ für alle $x \in \mathbb{R}$.
 \sin ist eine *ungerade Funktion*, d.h. es gilt $\sin(-x) = -\sin(x)$ für alle $x \in \mathbb{R}$.
 $\sin(0) = 0$ und $\cos(0) = 1$.
3. $\sin^2(x) + \cos^2(x) = 1$ für alle $x \in \mathbb{R}$.
(Da $\sin^2(x) + \cos^2(x) = |e^{ix}|^2 = 1$.)
4. *Additionstheoreme*: Für alle $x, y \in \mathbb{R}$ gilt:

$$\begin{aligned}\sin(x+y) &= \sin(x)\cos(y) + \sin(y)\cos(x) \\ \cos(x+y) &= \cos(x)\cos(y) - \sin(x)\sin(y).\end{aligned}$$

(Real- und Imaginärteil von $e^{i(x+y)} = e^{ix} \cdot e^{iy}$ vergleichen).

5. *Differenzformeln*: Für alle $x, y \in \mathbb{R}$ gilt:

$$\begin{aligned}\sin(x) - \sin(y) &= 2 \cdot \cos\left(\frac{x+y}{2}\right) \cdot \sin\left(\frac{x-y}{2}\right), \\ \cos(x) - \cos(y) &= -2 \cdot \sin\left(\frac{x+y}{2}\right) \cdot \sin\left(\frac{x-y}{2}\right).\end{aligned}$$

(Additionstheoreme auf $\frac{1}{2}(x+y) \pm \frac{1}{2}(x-y)$ anwenden.)

Einschließungslemma für sin und cos: Für alle reellen Zahlen $x \in (0, 2]$ gilt

$$\begin{aligned}1 - \frac{x^2}{2} < \cos(x) < 1 - \frac{x^2}{2} + \frac{x^4}{24}, \\ x - \frac{x^3}{6} < \sin(x) < x.\end{aligned}$$

Beweis. Die Reihen $\cos(x) := \sum_{k=0}^{\infty} (-1)^k \underbrace{\frac{x^{2k}}{(2k)!}}_{=: b_k}$ und $\sin(x) := \sum_{k=0}^{\infty} (-1)^k \underbrace{\frac{x^{2k+1}}{(2k+1)!}}_{=: c_k}$ sind

alternierende Reihen. Für $x \in (0, 2]$ sind $(b_k)_{k=1}^{\infty}$ und $(c_k)_{k=0}^{\infty}$ monoton fallende Nullfolgen positiver reeller Zahlen. Aus dem Leibniz-Kriterium für alternierende Reihen (siehe Beweis von Satz 3.23) kennen wir eine Abschätzung des Reihenwertes s einer alternierenden Reihe durch die Partialsummen s_n der Reihe:

$$s_{2m+1} < s < s_{2m} \quad \forall m \in \mathbb{N}_0.$$

Für $\cos(x)$ und $x \in (0, 2]$ ergibt sich bei $m = 1$:

$$1 - \frac{x^2}{2} < 1 - \frac{x^2}{2} + \underbrace{\frac{x^4}{24} - \frac{x^6}{6!}}_{>0} = s_3 < \cos(x) < s_2 = 1 - \frac{x^2}{2} + \frac{x^4}{24}.$$

Für $\sin(x)$ und $x \in (0, 2]$ ergibt sich bei $m = 0$:

$$x - \frac{x^3}{6} = s_1 < \sin(x) < s_0 = x.$$

□

Als nächstes diskutieren wir den Kurvenverlauf der Funktionen \cos und \sin . Zunächst sehen wir uns die Nullstellen an:

Satz 4.17 Die Funktion $\cos : \mathbb{R} \rightarrow \mathbb{R}$ hat auf dem Intervall $(0, 2)$ genau eine Nullstelle.

Beweis. 1. *Existenz der Nullstelle:* Wie wir bereits wissen gilt $\cos(0) = 1$. Andererseits zeigt das Einschließungslemma

$$\cos(2) < 1 - \frac{2^2}{2} + \frac{2^4}{24} = 1 - 2 + \frac{2}{3} = -\frac{1}{3}.$$

Da $\cos : \mathbb{R} \rightarrow \mathbb{R}$ stetig ist, folgt aus dem Zwischenwertsatz, dass ein $\xi \in (0, 2)$ mit $\cos(\xi) = 0$ existiert.

2. *Eindeutigkeit der Nullstelle:* Dazu zeigen wir, dass die Funktion \cos auf $(0, 2)$ streng monoton fallend ist:

a) Die Funktion \sin ist auf $(0, 2)$ positiv, da nach dem Einschließungslemma für $x \in (0, 2)$ gilt:

$$0 < x \left(1 - \frac{x^2}{6}\right) = x - \frac{x^3}{6} < \sin(x).$$

b) Seien nun $x, y \in (0, 2)$ mit $x > y$. Dann gilt $\frac{x \pm y}{2} \in (0, 2)$. Aus den Differenzformeln folgt

$$\cos(x) - \cos(y) = -2 \sin\left(\frac{x+y}{2}\right) \cdot \sin\left(\frac{x-y}{2}\right) < 0.$$

Folglich ist \cos auf dem Intervall $(0, 2)$ streng monoton fallend und somit ξ die einzige Nullstelle von \cos in $(0, 2)$. □

Definition 4.12 (Analytische Definition der Zahl π).

Die Nullstelle von \cos auf dem Intervall $(0, 2)$ heißt $\frac{\pi}{2}$.

In der Elementargeometrie definiert man die Zahl π als den Flächeninhalt der Kreisscheibe von Radius 1 und approximiert sie durch die Flächeninhalte von regelmäßigen n -Ecken mit Eckpunkten auf dem Rand der Kreisscheibe für $n \rightarrow \infty$. Dabei erhält man

$$\pi \approx 3,141592653\dots$$

Wir werden noch zeigen, dass beide Definitionen übereinstimmen. Die Zahl π ist irrational und transzendent, d.h. es gibt kein Polynom mit ganzzahligen Koeffizienten, das π als Nullstelle hat.

Da $\cos(\frac{\pi}{2}) = 0$, gilt $\sin(\frac{\pi}{2}) = 1$ und aus der Eulerformel folgt $e^{i\frac{\pi}{2}} = i$. Wegen $e^{ix+iy} = e^{ix} \cdot e^{iy}$ erhalten wir außerdem:

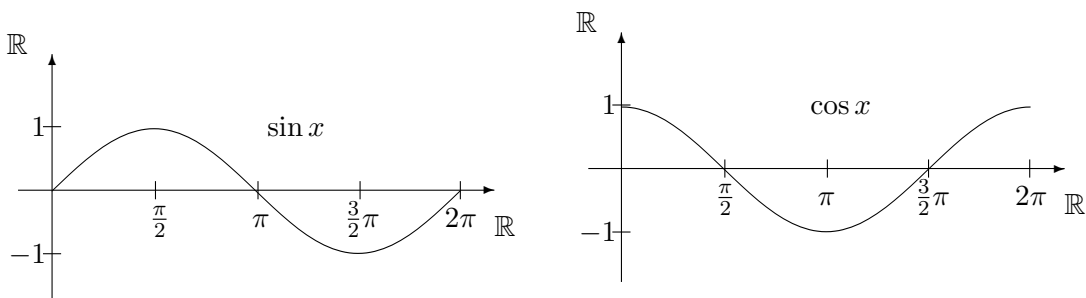
Folgerung 4.2 $e^{i\frac{\pi}{2}} = i$, $e^{i\pi} = -1$, $e^{i\frac{3}{2}\pi} = -i$, $e^{2\pi i} = 1$.
 $e^{ix+2\pi ik} = e^{ix}$ für alle $x \in \mathbb{R}$ und $k \in \mathbb{Z}$.

Der Vergleich von Real- und Imaginärteil in der Eulerschen Formel liefert dann die folgenden Beziehungen zwischen Sinus und Cosinus:

Folgerung 4.3 Für alle $x \in \mathbb{R}$ gilt

$$\begin{aligned} \sin(x + \frac{\pi}{2}) &= +\cos(x), & \cos(x + \frac{\pi}{2}) &= -\sin(x), \\ \sin(x + \pi) &= -\sin(x), & \cos(x + \pi) &= -\cos(x), \\ \sin(x + 2\pi) &= \sin(x), & \cos(x + 2\pi) &= \cos(x). \end{aligned}$$

Daraus erhalten wir die weiteren Nullstellen und die Monotoniebereiche von Sinus und Cosinus auf \mathbb{R} . Die folgenden Bilder zeigen den Kurvenverlauf der Cosinus- und Sinusfunktion auf $[0, 2\pi]$:



Die nächste Tabelle zeigt die wichtigsten konkreten Werte von \sin und \cos , die wir mit analytischen Mitteln herleiten werden⁴:

Satz 4.18 (Spezielle Werte von \sin und \cos)

Bogenmaß	x	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
Gradmaß	$\varphi(x)$	0°	30°	45°	60°	90°
	$\sin(x)$	0	$\frac{1}{2}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}\sqrt{3}$	1
	$\cos(x)$	1	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}$	0

⁴ Ich empfehle Ihnen, sich auch eine elementargeometrische Herleitung der speziellen Werte von \sin und \cos zu überlegen (Benutze Basiswinkelsatz und Satz des Pythagoras).

Beweis. a) Umrechnung vom Bogenmaß ins Gradmaß: $\varphi(0) = 0^\circ$ ist wegen $e^{i0} = 1$ offensichtlich. Da $e^{i\frac{\pi}{2}} = i$ folgt $\varphi(\frac{\pi}{2}) = 90^\circ$. Bei der Multiplikation der komplexen Zahlen e^{ix} und e^{iy} addieren sich die Winkel $\varphi(x)$ und $\varphi(y)$. Wegen $e^{ix} \cdot e^{iy} = e^{i(x+y)}$ folgt dann $\varphi(x) + \varphi(y) = \varphi(x+y)$. Damit erhält man die anderen Umrechnungen vom Bogen- ins Gradmaß.

b) $x = \frac{\pi}{4}$: Nach Folgerung 4.3 gilt:

$$\cos\left(\frac{\pi}{4}\right) = \cos\left(\frac{\pi}{2} - \frac{\pi}{4}\right) = -\sin\left(-\frac{\pi}{4}\right) = \sin\left(\frac{\pi}{4}\right).$$

Folglich ist

$$1 = \cos^2\left(\frac{\pi}{4}\right) + \sin^2\left(\frac{\pi}{4}\right) = 2\cos^2\left(\frac{\pi}{4}\right).$$

Da $\cos\left(\frac{\pi}{4}\right) > 0$ folgt

$$\cos\left(\frac{\pi}{4}\right) = \sqrt{\frac{1}{2}} = \frac{1}{2}\sqrt{2} = \sin\left(\frac{\pi}{4}\right).$$

c) $x = \frac{\pi}{3}$: Wir setzen $z := e^{i\frac{\pi}{3}}$. Dann gilt $z^3 = (e^{i\frac{\pi}{3}})^3 = e^{i\pi} = -1$ und wir erhalten $0 = z^3 + 1 = (z+1)(z^2 - z + 1)$. Da $z \neq -1$ folgt $z^2 - z + 1 = 0$ und somit $z - 1 + \frac{1}{z} = 0$. Deshalb gilt

$$1 = e^{i\frac{\pi}{3}} + \frac{1}{e^{i\frac{\pi}{3}}} = e^{i\frac{\pi}{3}} + e^{-i\frac{\pi}{3}} = 2\cos\left(\frac{\pi}{3}\right).$$

Dies zeigt, dass $\cos\left(\frac{\pi}{3}\right) = \frac{1}{2}$ und folglich $\sin\left(\frac{\pi}{3}\right) = \frac{1}{2}\sqrt{3}$.

d) $x = \frac{\pi}{6}$: Nach Folgerung 4.3 gilt

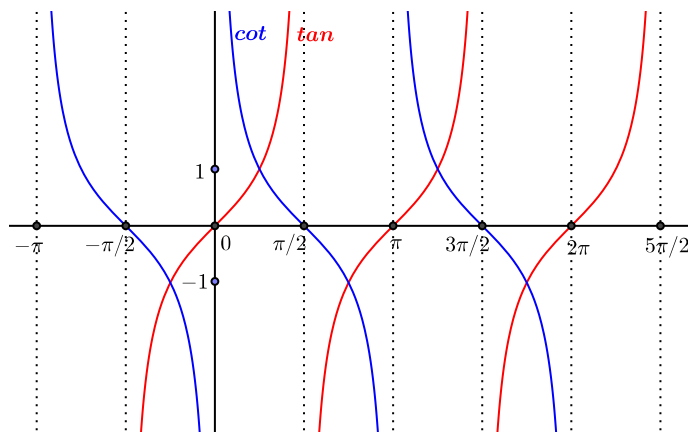
$$\sin\left(\frac{\pi}{6}\right) = \sin\left(\frac{\pi}{2} - \frac{\pi}{3}\right) = \cos\left(-\frac{\pi}{3}\right) = \cos\left(\frac{\pi}{3}\right) = \frac{1}{2}$$

und deshalb auch $\cos\left(\frac{\pi}{6}\right) = \frac{1}{2}\sqrt{3}$. □

Mittels der Sinus- und Cosinusfunktion definieren wir die trigonometrischen Funktionen Tangens (tan) und Cotangens (cot):

$$\begin{aligned} \tan : \mathbb{R} \setminus \left\{ \frac{\pi}{2} + k\pi \mid k \in \mathbb{Z} \right\} &\longrightarrow \mathbb{R} \\ x &\longmapsto \tan(x) := \frac{\sin(x)}{\cos(x)} \end{aligned}$$

$$\begin{aligned} \cot : \mathbb{R} \setminus \{k\pi \mid k \in \mathbb{Z}\} &\longrightarrow \mathbb{R} \\ x &\longmapsto \cot(x) := \frac{\cos(x)}{\sin(x)} \end{aligned}$$



Aus den Eigenschaften von Sinus und Cosinus erhält man die folgenden Eigenschaften von Tangens- und Cotangensfunktion:

1. \tan und \cot sind stetige, π -periodische, ungerade Funktionen.
2. Es gelten folgende *Additionstheoreme*:

$$\tan(x+y) = \frac{\tan(x) + \tan(y)}{1 - \tan(x) \cdot \tan(y)} \quad \text{und} \quad \cot(x+y) = \frac{\cot(x) \cdot \cot(y) - 1}{\cot(x) + \cot(y)}.$$

Die Umkehrfunktionen der trigonometrischen Funktionen

Da die trigonometrischen Funktionen periodisch sind, existieren keine globalen Umkehrfunktionen. Es existieren aber Umkehrfunktionen auf jedem Monotoniebereich der trigonometrischen Funktionen. Nach Satz 4.4 sind diese Umkehrfunktionen ebenfalls stetig.

Definition 4.13. Die Umkehrfunktion von $\sin|_{[-\frac{\pi}{2}, \frac{\pi}{2}]}$ heißt *Arcussinus* und wird mit \arcsin bezeichnet.

$\arcsin : [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$ ist bijektiv, stetig und streng monoton wachsend.

Die Umkehrfunktionen von $\sin|_{[-\frac{\pi}{2}+k\pi, \frac{\pi}{2}+k\pi]}$, $k \in \mathbb{Z}$, erfüllen nach den Additionstheoremen

$$(\sin|_{[-\frac{\pi}{2}+k\pi, \frac{\pi}{2}+k\pi]})^{-1}(x) = (-1)^k \cdot \arcsin(x) + k\pi.$$

Man nennt sie auch die Nebenzweige der Umkehrfunktion von Sinus.

Definition 4.14. Die Umkehrfunktion von $\cos|_{[0, \pi]}$ heißt *Arcuscosinus* und wird mit \arccos bezeichnet.

$\arccos : [-1, 1] \rightarrow [0, \pi]$ ist bijektiv, stetig und streng monoton fallend.

Die Umkehrfunktion von $\cos|_{[k\pi, (k+1)\pi]}$, $k \in \mathbb{Z}$, ist nach den Additionstheoremen

$$(\cos|_{[k\pi, (k+1)\pi]})^{-1}(x) = (-1)^k \cdot \arccos(x) + \frac{2k+1 - (-1)^k}{2} \cdot \pi.$$

Definition 4.15. Die Umkehrfunktion von $\tan|_{(-\frac{\pi}{2}, \frac{\pi}{2})}$ heißt *Arcustangens* und wird mit \arctan bezeichnet.

$\arctan : \mathbb{R} \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$ ist bijektiv, stetig und streng monoton wachsend.

Die Umkehrfunktion von $\tan|_{(-\frac{\pi}{2}+k\pi, \frac{\pi}{2}+k\pi)}$, $k \in \mathbb{Z}$, erfüllt nach den Additionstheoremen

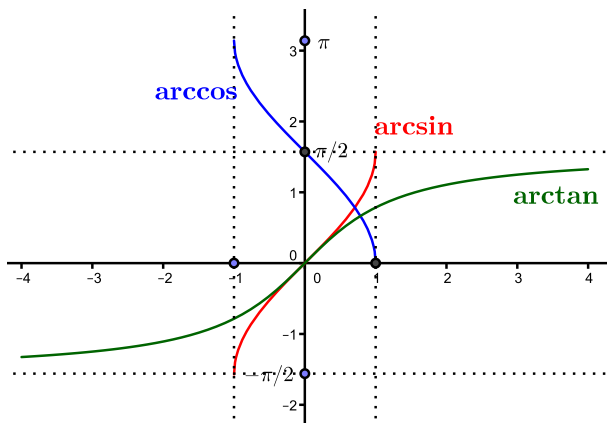
$$(\tan|_{(-\frac{\pi}{2}+k\pi, \frac{\pi}{2}+k\pi)})^{-1}(x) = \arctan(x) + k\pi.$$

Definition 4.16. Die Umkehrfunktion von $\cot|_{(0, \pi)}$ heißt *Arcuscotangens* und wird mit arccot bezeichnet.

$\operatorname{arccot} : \mathbb{R} \rightarrow (0, \pi)$ ist bijektiv, stetig und streng monoton fallend.

Die Umkehrfunktion von $\cot|_{(k\pi, (k+1)\pi)}$, $k \in \mathbb{Z}$, erfüllt nach den Additionstheoremen

$$(\cot|_{(k\pi, (k+1)\pi)})^{-1}(x) = \operatorname{arccot}(x) + k\pi.$$



4.5.2 Die Hyperbelfunktionen

Unter den Hyperbelfunktionen *Sinus hyperbolicus* $\sinh : \mathbb{R} \rightarrow \mathbb{R}$ und *Cosinus hyperbolicus* $\cosh : \mathbb{R} \rightarrow \mathbb{R}$ versteht man die folgenden reellen Funktionen:

$$\sinh(x) := \frac{e^x - e^{-x}}{2},$$

$$\cosh(x) := \frac{e^x + e^{-x}}{2}.$$

Aus der Reihendarstellung für e^x folgt

$$\sinh(x) = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \dots$$

$$\cosh(x) = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \dots$$

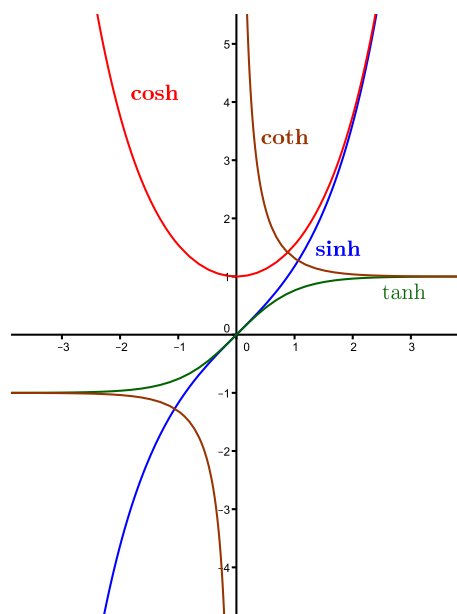
Wir bilden auch hier die Quotienten beider Funktionen und erhalten *Tangens hyperbolicus*

$\tanh : \mathbb{R} \rightarrow \mathbb{R}$:

$$\tanh(x) := \frac{\sinh(x)}{\cosh(x)} = \frac{e^{2x} - 1}{e^{2x} + 1}$$

und *Cotangens hyperbolicus* $\operatorname{coth} : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$:

$$\operatorname{coth}(x) := \frac{\cosh(x)}{\sinh(x)} = \frac{e^{2x} + 1}{e^{2x} - 1}.$$



Alle diese Hyperbelfunktionen sind offensichtlich stetig. Durch Einsetzen der Exponentialfunktion in die Definitionen bzw. Betrachtung der Potenzreihen erhält man folgende Rechenregeln und Eigenschaften für die Hyperbelfunktionen:

1. \sinh ist eine bijektive, streng monoton wachsende, ungerade Funktion mit $\sinh(0) = 0$.
2. \cosh ist eine gerade Funktion mit $\cosh(x) \geq 1$ für alle $x \in \mathbb{R}$. $\cosh|_{[0, +\infty)} : [0, +\infty) \rightarrow [1, +\infty)$ ist bijektiv und streng monoton wachsend mit $\cosh(0) = 1$.
3. $\cosh^2(x) - \sinh^2(x) = 1$ für alle $x \in \mathbb{R}$
(Definitionen einsetzen und ausrechnen).
4. *Additionstheoreme:*

$$\begin{aligned}\cosh(x + y) &= \cosh(x) \cosh(y) + \sinh(x) \sinh(y), \\ \sinh(x + y) &= \sinh(x) \cosh(y) + \cosh(x) \sinh(y), \\ \tanh(x + y) &= \frac{\tanh(x) + \tanh(y)}{1 + \tanh(x) \cdot \tanh(y)}.\end{aligned}$$

Bemerkung:

- a) Der Name *Hyperbelfunktionen* hat folgenden Grund:

So wie die trigonometrischen Funktionen \sin und \cos die Punkte der Kreislinie $S^1 := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ beschreiben:

$$S^1 = \{(\cos(t), \sin(t)) \mid t \in [0, 2\pi)\},$$

so beschreiben die Hyperbelfunktionen \sinh und \cosh die Punkte der Hyperbel $Hyp := \{(x, y) \in \mathbb{R}^2 \mid x^2 - y^2 = 1, x > 0\}$:

$$Hyp = \{(\cosh(t), \sinh(t)) \mid t \in \mathbb{R}\}.$$

- b) Hängt man ein ideales Seil zwischen zwei Punkten über der Erdoberfläche auf, so wird die entstehende *Seilkurve* durch die Funktion $f(x) = b \cosh(ax)$ beschrieben.

Die Umkehrfunktionen der Hyperbelfunktionen

Die Funktion $\sinh : \mathbb{R} \rightarrow \mathbb{R}$ ist bijektiv und streng monoton wachsend.

Die Funktion $\cosh : \mathbb{R} \rightarrow \mathbb{R}$ nimmt ihre Werte in $[1, +\infty)$ an, sie ist bijektiv und streng monoton wachsend von $[0, +\infty)$ auf $[1, +\infty)$.

Wir können deshalb die Umkehrfunktionen betrachten:

Definition 4.17. Die Umkehrfunktion von $\sinh : \mathbb{R} \rightarrow \mathbb{R}$ heißt *Areasinus hyperbolicus* und wird mit arsinh bezeichnet: $\operatorname{arsinh} : \mathbb{R} \rightarrow \mathbb{R}$.

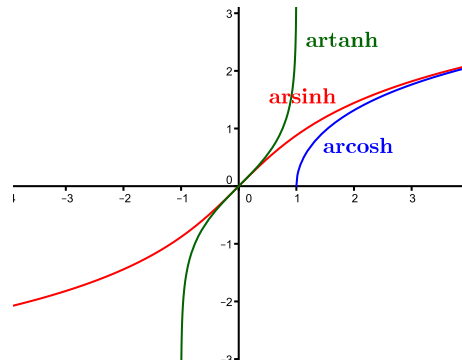
Die Umkehrfunktion von $\cosh|_{[0, +\infty)} : [0, +\infty) \rightarrow [1, +\infty)$ heißt *Areacosinus hyperbolicus* und wird mit arcosh bezeichnet: $\operatorname{arcosh} : [1, +\infty) \rightarrow [0, +\infty)$.

Die Funktion $\tanh : \mathbb{R} \rightarrow (-1, 1) \subset \mathbb{R}$ ist bijektiv und streng monoton wachsend. Die Funktion $\operatorname{coth}|_{\mathbb{R} \setminus \{0\}} : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R} \setminus [-1, 1]$ ist bijektiv und streng monoton fallend.

Definition 4.18. Die Umkehrfunktion von \tanh heißt *Areatangens hyperbolicus* :
 $\operatorname{artanh} : (-1, 1) \rightarrow \mathbb{R}$.

Die Umkehrfunktion von coth heißt *Areacotangens hyperbolicus* :
 $\operatorname{arcoth} : \mathbb{R} \setminus [-1, 1] \rightarrow \mathbb{R} \setminus \{0\}$.

Alle vier Umkehrfunktionen sind wegen Satz 4.4 stetig.



Satz 4.19 Es bestehen die folgenden Beziehungen zwischen den Umkehrfunktionen der Hyperbelfunktionen und der Logarithmusfunktion:

$$\begin{aligned} \operatorname{arsinh}(x) &= \ln(x + \sqrt{x^2 + 1}), & \forall x \in \mathbb{R}, \\ \operatorname{arcosh}(x) &= \ln(x + \sqrt{x^2 - 1}), & \forall x \geq 1, \\ \operatorname{artanh}(x) &= \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right), & \forall x \in \mathbb{R} \text{ mit } |x| < 1, \\ \operatorname{arcoth}(x) &= \frac{1}{2} \ln\left(\frac{x+1}{x-1}\right), & \forall x \in \mathbb{R} \text{ mit } |x| > 1. \end{aligned}$$

Beweis. Übungsaufgabe. □

Differentialrechnung für Funktionen einer reellen Variablen

In diesem Kapitel behandeln wir die Differentialrechnung für Funktionen einer reellen Variablen. Die Ideen zur Entwicklung der Differentialrechnung gehen bis ins 17. Jahrhundert zurück. Sie sind eng verbunden mit dem Versuch, Probleme der Geometrie durch Einsatz analytischer Methoden zu lösen (z.B. Bestimmung von Tangenten oder von Figuren und Körpern mit maximalen Flächeninhalten) sowie mit den Erfordernissen, die die Behandlung von Problemen aus der Mechanik (Verständnis von Geschwindigkeit und Beschleunigung sowie von Kraft und Trägheit) an die Mathematik stellte. Für einen kurzen historischen Abriss empfehle ich das Buch von W. Walter: *Analysis I*, § 10.

Die Grundidee der Differentialrechnung für Funktionen einer reellen Variable ist das Studium des lokalen Änderungsverhaltens der Funktion in der Nähe eines Punktes mit Hilfe der Approximation der Funktion durch lineare Abbildungen (Tangenten) bzw. durch Polynome. Die Eigenschaften dieser linearen bzw. polynomialen Approximation lassen viele Aussagen über das lokale Verhalten der Funktion zu.

Im Abschnitt über stetige Funktionen haben wir reelle und komplexe Funktionen gemeinsam behandelt. Die Definition der Stetigkeit ist in beiden Fällen die gleiche, die grundlegenden Eigenschaften stetiger Funktionen sind auch für beide Fälle analog. Bei differenzierbaren Funktionen ist die Situation grundsätzlich anders: Zwar kann man auch die Differenzierbarkeit für reelle und für komplexe Funktionen auf die gleiche Weise definieren, jedoch haben differenzierbare *komplexe* Funktionen ein grundsätzlich anderes Verhalten als differenzierbare *reelle* Funktionen. Deshalb hat sich ein eigenes Gebiet der Mathematik entwickelt, das sich mit differenzierbaren komplexen Funktionen beschäftigt, die *Funktionentheorie*. Zu diesem Themengebiet werden Vorlesungen im Wahlpflichtbereich des Lehramtsstudiums angeboten. Im Rahmen der Grundvorlesung *Analysis* befassen wir uns zunächst mit den Eigenschaften *reeller* differenzierbarer Funktionen.

5.1 Differenzierbare reelle Funktionen: Rechenregeln und Beispiele

In diesem Kapitel bezeichnet $I \subset \mathbb{R}$ ein reelles Intervall¹.

¹ Hierbei sind alle Sorten von Intervallen zugelassen, die mehr als einen Punkt enthalten, d.h. offene, halb-offene, abgeschlossene, beschränkte und unbeschränkte.

Definition 5.1. Eine Funktion $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ heißt in $x_0 \in I$ differenzierbar, falls der Grenzwert

$$f'(x_0) := \frac{df}{dx}(x_0) := \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

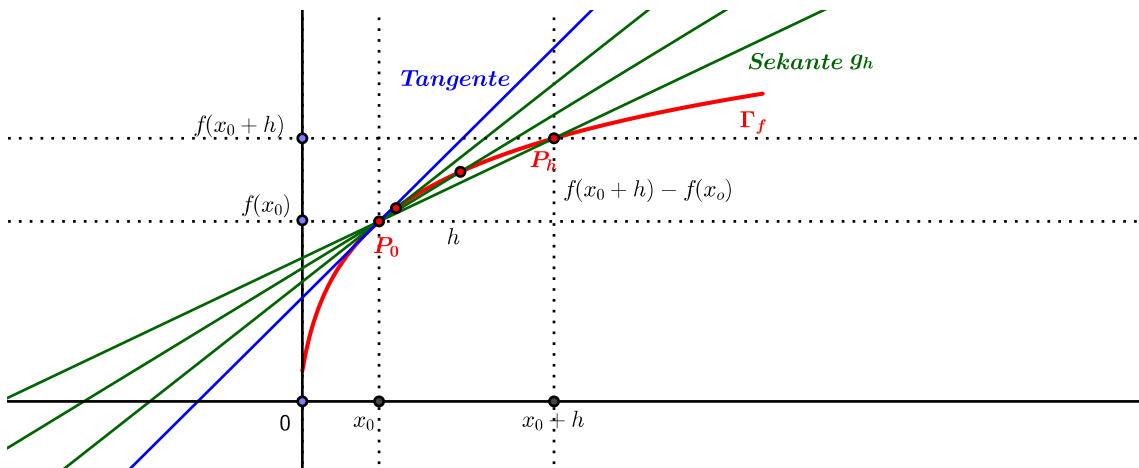
existiert. Dieser Grenzwert heißt Ableitung von f in x_0 .

Die Funktion $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ heißt differenzierbar, falls sie in jedem Punkt $x_0 \in I$ differenzierbar ist. In diesem Fall heißt die Funktion

$$\begin{aligned} f' : I \subset \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto f'(x) \end{aligned}$$

1. Ableitung von f .

Geometrische Interpretation der Ableitung:



Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ eine differenzierbare reelle Funktion. Dann beschreibt der *Differenzenquotient* $\frac{f(x_0+h)-f(x_0)}{h}$ den Anstieg der Sekante g_h durch die Punkte $P_0 := (x_0, f(x_0))$ und $P_h := (x_0 + h, f(x_0 + h))$ des Funktionsgraphen Γ_f . Die Gerade g_h wird durch die Funktion

$$S_h(x) = \underbrace{\frac{f(x_0 + h) - f(x_0)}{h}}_{\text{Anstieg}} \cdot x + c(h)$$

beschrieben (d.h. g_h ist der Graph von S_h). Die Tangente an die Kurve Γ_f im Punkt P_0 ist die Grenzgerade der Sekanten g_h für $h \rightarrow 0$. Ihre Funktionsgleichung lautet daher:

$$\begin{aligned} T_{x_0}(x) &:= \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \cdot x + \lim_{h \rightarrow 0} c(h) \\ &= f'(x_0) \cdot x + c(0). \end{aligned}$$

Da die Tangente die Kurve Γ_f im Punkt P_0 berührt, gilt $T_{x_0}(x_0) = f(x_0) = f'(x_0)x_0 + c(0)$. Damit kann man $c(0)$ bestimmen und erhält als Funktionsgleichung für die Tangente

$$T_{x_0}(x) = f'(x_0)(x - x_0) + f(x_0).$$

Folglich beschreibt die Ableitung $f'(x_0)$ den Anstieg der Tangenten an die Kurve Γ_f im Punkt P_0 .

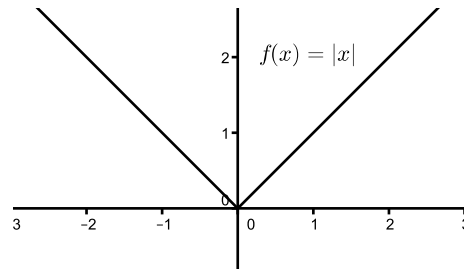
Satz 5.1 *Ist eine Funktion $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ in $x_0 \in I$ differenzierbar, so ist sie in x_0 auch stetig. Jede differenzierbare Funktion $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ ist somit auf I stetig.*

Beweis. Wir benutzen das Folgenkriterium für die Stetigkeit. Sei (x_n) eine Folge in I , die gegen x_0 konvergiert. OBdA. sei $x_n \neq x_0$ für alle $n \in \mathbb{N}$. Dann gilt:

$$f(x_n) - f(x_0) = (x_n - x_0) \left(\frac{f(x_n) - f(x_0)}{x_n - x_0} \right) \xrightarrow{n \rightarrow \infty} 0 \cdot f'(x_0) = 0.$$

Somit konvergiert die Bildfolge $(f(x_n))$ gegen $f(x_0)$. Nach dem Folgenkriterium für die Stetigkeit ist f deshalb in x_0 stetig. □

Bemerkung 1: Es gibt stetige Funktionen, die nicht differenzierbar sind. Z.B. ist die Betragsfunktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) := |x|$, in $x_0 = 0$ stetig, aber nicht differenzierbar.



Es gibt sogar stetige Funktionen, die in keinem Punkt differenzierbar sind².

Satz 5.2 (Ableitungsregeln)

Es seien $f, g : I \subset \mathbb{R} \rightarrow \mathbb{R}$ Funktionen, die in $x_0 \in I$ differenzierbar sind. Dann gilt:

1. **Summenregel:**

$$f + g : I \rightarrow \mathbb{R} \text{ ist in } x_0 \text{ differenzierbar und } (f + g)'(x_0) = f'(x_0) + g'(x_0).$$

2. **Produktregel:**

$$f \cdot g : I \rightarrow \mathbb{R} \text{ ist in } x_0 \text{ differenzierbar und}$$

$$(f \cdot g)'(x_0) = f'(x_0) \cdot g(x_0) + f(x_0) \cdot g'(x_0).$$

3. **Quotientenregel:**

g habe auf I keine Nullstelle. Dann ist $\frac{f}{g} : I \rightarrow \mathbb{R}$ in x_0 differenzierbar und

$$\left(\frac{f}{g} \right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}.$$

² Ein Beispiel können Sie sich im Buch von K. Königsberger: Analysis 1, 6. Auflage, Kapitel 9.11, oder im Buch von O. Deiser: Analysis 1, S. 283, ansehen (Takagi-Funktion).

Insbesondere gilt auch

$$\left(\frac{1}{g}\right)'(x_0) = -\frac{g'(x_0)}{g^2(x_0)}.$$

Beweis. 1. Die Behauptung folgt durch Limesbildung aus

$$\frac{f(x) + g(x) - (f(x_0) + g(x_0))}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x - x_0}.$$

2. Die Behauptung folgt auch hier wieder durch Limesbildung aus

$$\frac{(f \cdot g)(x) - (f \cdot g)(x_0)}{x - x_0} = f(x) \cdot \frac{g(x) - g(x_0)}{x - x_0} + \frac{f(x) - f(x_0)}{x - x_0} \cdot g(x_0),$$

wobei wir auch die Stetigkeit von f in x_0 benutzen.

3. Die letztere Behauptung folgt durch Limesbildung aus

$$\frac{\left(\frac{1}{g}\right)(x) - \left(\frac{1}{g}\right)(x_0)}{x - x_0} = \frac{\frac{1}{g(x)} - \frac{1}{g(x_0)}}{x - x_0} = \frac{g(x_0) - g(x)}{x - x_0} \cdot \frac{1}{g(x)g(x_0)} \xrightarrow{x \rightarrow x_0} -g'(x_0) \cdot \frac{1}{g^2(x_0)}.$$

Mit Hilfe der Produktregel erhält man dann sofort die Quotientenregel. \square

Satz 5.3 (Kettenregel)

Es seien $I, J \subset \mathbb{R}$ Intervalle, die Funktion $g : J \rightarrow I$ in $x_0 \in J$ differenzierbar und die Funktion $f : I \rightarrow \mathbb{R}$ in $g(x_0)$ differenzierbar. Dann ist die Verknüpfung $f \circ g : J \rightarrow \mathbb{R}$ in x_0 differenzierbar und es gilt

$$(f \circ g)'(x_0) = f'(g(x_0)) \cdot g'(x_0).$$

Beweis. Für $g(x) \neq g(x_0)$ gilt

$$\frac{(f \circ g)(x) - (f \circ g)(x_0)}{x - x_0} = \frac{g(x) - g(x_0)}{x - x_0} \cdot \frac{f(g(x)) - f(g(x_0))}{g(x) - g(x_0)}. \quad (*)$$

Sei (x_n) eine gegen x_0 konvergente Folge mit $x_n \neq x_0$ für alle $n \in \mathbb{N}$.

1. *Fall:* Es existiert ein $n_0 \in \mathbb{N}$ mit $g(x_n) \neq g(x_0)$ für alle $n \geq n_0$. Da g in x_0 stetig ist, konvergiert die Folge $(g(x_n))$ gegen $g(x_0)$ und aus (*) folgt wegen der Differenzierbarkeit von g in x_0 und f in $g(x_0)$:

$$\lim_{n \rightarrow \infty} \frac{f(g(x_n)) - f(g(x_0))}{x_n - x_0} = g'(x_0) \cdot f'(g(x_0)).$$

2. *Fall:* Es gelte $g(x_n) = g(x_0)$ für unendlich viele $n \in \mathbb{N}$. Wir betrachten die Teilfolge (x_{n_k}) dieser Folgenglieder von (x_n) . Da g in x_0 differenzierbar ist, gilt dann

$$g'(x_0) = \lim_{k \rightarrow \infty} \frac{g(x_{n_k}) - g(x_0)}{x_{n_k} - x_0} = 0.$$

Demnach ist

$$\lim_{k \rightarrow \infty} \frac{f(g(x_{n_k})) - f(g(x_0))}{x_{n_k} - x_0} = \lim_{k \rightarrow \infty} 0 = 0 = g'(x_0) \cdot f'(g(x_0)).$$

Unter Benutzung des 1. Falles erhalten wir die erforderliche Konvergenz für die gesamte Folge (x_n) . \square

Satz 5.4 (Ableitung der Umkehrfunktion)

Sei $f : I \rightarrow J$ eine bijektive Funktion zwischen zwei Intervallen, $x_0 \in I$ und es gelte:

- a) Die Funktion f ist in x_0 differenzierbar und $f'(x_0) \neq 0$,
 b) Die Umkehrfunktion $f^{-1} : J \rightarrow I$ ist in $y_0 := f(x_0) \in J$ stetig.

Dann ist die Umkehrfunktion $f^{-1} : J \rightarrow I$ in y_0 differenzierbar und für ihre Ableitung gilt:

$$(f^{-1})'(y_0) = \frac{1}{f'(f^{-1}(y_0))}.$$

Beweis. Für $y \in J$ mit $y_0 \neq y$ und $x := f^{-1}(y) \in I$ gilt:

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} = \frac{1}{\frac{f(x) - f(x_0)}{x - x_0}}. \quad (**)$$

Sei nun (y_n) eine Folge in J mit $y_n \neq y_0$, die gegen y_0 konvergiert. Da f^{-1} in y_0 stetig ist, konvergiert die Folge $(x_n := f^{-1}(y_n))$ gegen $x_0 = f^{-1}(y_0)$. Da f in x_0 differenzierbar ist, folgt dann mit (**)

$$\lim_{n \rightarrow \infty} \frac{f^{-1}(y_n) - f^{-1}(y_0)}{y_n - y_0} = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}.$$

□

Bemerkung 2: Da jede differenzierbare Funktion auch stetig ist, ist die Voraussetzung b) in Satz 5.4 notwendig. Ist f in x_0 differenzierbar, so ist die Voraussetzung $f'(x_0) \neq 0$ in a) ebenfalls notwendig. Sind nämlich f in x_0 und f^{-1} in $y_0 = f(x_0)$ differenzierbar, so gilt wegen $f^{-1} \circ f = \text{Id}_I$ nach Kettenregel:

$$(f^{-1})'(f(x_0)) \cdot f'(x_0) = 1,$$

und somit $f'(x_0) \neq 0$.

Die letzte Bedingung ist für die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^3$ verletzt. f ist bijektiv und differenzierbar, es gilt $f(0) = 0$ und $f'(0) = 0$. Die inverse Funktion $f^{-1} : \mathbb{R} \rightarrow \mathbb{R}$ ist gegeben durch

$$f^{-1}(y) := \begin{cases} \sqrt[3]{y} & \text{falls } y > 0, \\ 0 & \text{falls } y = 0, \\ -\sqrt[3]{|y|} & \text{falls } y < 0. \end{cases}$$

f^{-1} ist in 0 stetig, aber nicht differenzierbar. Z.B. existiert bereits der rechtsseitige Grenzwert des Differenzenquotienten von f^{-1} für y gegen 0^+ nicht in \mathbb{R} , da

$$\frac{f^{-1}(y) - f^{-1}(0)}{y} = \frac{\sqrt[3]{y}}{y} = \frac{1}{\sqrt[3]{y^2}} \xrightarrow{y \rightarrow 0^+} +\infty.$$

Dieses Beispiel zeigt auch, dass die Umkehrfunktion einer nicht differenzierbaren Funktion durchaus differenzierbar sein kann.

Ableitungen elementarer Funktionen

1. Polynome

Seien a_0, a_1, \dots, a_n reelle Zahlen und $P(x)$ das Polynom

$P(x) := a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$. Dann ist die Funktion $P : \mathbb{R} \rightarrow \mathbb{R}$ differenzierbar und für die Ableitung gilt:

$$P'(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \dots + 2 a_2 x + a_1.$$

Beweis. Wegen der Summen- und Produktregel für Ableitungen genügt es, die Polynome der Form $P(x) := x^k$ für $k \in \mathbb{N}$ zu betrachten. Für den Differenzenquotienten gilt:

$$\begin{aligned} \frac{x^k - x_0^k}{x - x_0} &= \frac{(x - x_0)(x^{k-1} + x^{k-2}x_0 + x^{k-3}x_0^2 + \dots + x \cdot x_0^{k-2} + x_0^{k-1})}{x - x_0} \\ &= x^{k-1} + x^{k-2}x_0 + x^{k-3}x_0^2 + \dots + x \cdot x_0^{k-2} + x_0^{k-1}. \end{aligned}$$

Für $x \rightarrow x_0$ konvergiert dies gegen $k \cdot x_0^{k-1}$. □

2. Die Exponentialfunktion

Die Exponentialfunktion $\exp : \mathbb{R} \rightarrow \mathbb{R}$, $\exp(x) := e^x$, ist differenzierbar und es gilt $\exp' = \exp$.

Beweis. Wir betrachten den Differenzenquotienten $\frac{e^{x+h} - e^x}{h} = e^x \cdot \frac{e^h - 1}{h}$. Nach Übungsaufgabe 32 gilt $\lim_{h \rightarrow 0} \frac{e^h - 1}{h} = 1$. Folglich ist \exp differenzierbar und für die Ableitung im Punkt $x \in \mathbb{R}$ gilt

$$\exp'(x) = \lim_{h \rightarrow 0} \frac{e^{x+h} - e^x}{h} = \lim_{h \rightarrow 0} e^x \cdot \frac{e^h - 1}{h} = e^x = \exp(x). \quad \square$$

3. Die Logarithmusfunktion

$\ln : \mathbb{R}^+ \rightarrow \mathbb{R}$ ist differenzierbar und es gilt $\ln'(x) = \frac{1}{x}$ für alle $x \in \mathbb{R}^+$.

Beweis. Die Logarithmusfunktion \ln ist die Umkehrfunktion von $\exp : \mathbb{R} \rightarrow \mathbb{R}^+$. Sie ist stetig, die Exponentialfunktion \exp ist differenzierbar mit $\exp' = \exp > 0$. Somit sind die Voraussetzungen des Satzes 5.4 erfüllt und es folgt für alle $x \in \mathbb{R}^+$:

$$\ln'(x) = \frac{1}{\exp'(\ln(x))} = \frac{1}{e^{\ln(x)}} = \frac{1}{x}. \quad \square$$

4. Potenzfunktionen und Wurzelfunktionen

Sei $a \in \mathbb{R}$ eine fixierte reelle Zahl. Wir betrachten die Potenzfunktion $p_a : \mathbb{R}^+ \rightarrow \mathbb{R}$ mit dem Exponenten a :

$$p_a(x) := x^a = e^{\ln(x) \cdot a}.$$

p_a ist differenzierbar und es gilt $p_a'(x) = a \cdot x^{a-1}$ für alle $x \in \mathbb{R}^+$.

(Der Spezialfall $a = \frac{1}{m}$ für $m \in \mathbb{N}$, $m \geq 2$, liefert die Wurzelfunktionen).

Beweis. Nach der Kettenregel (Satz 5.3) gilt:

$$p'_a(x) = \exp'(\ln(x) \cdot a) \cdot a \cdot \ln'(x) = e^{\ln(x) \cdot a} \cdot a \cdot \frac{1}{x} = \frac{x^a}{x} \cdot a = a \cdot x^{a-1}.$$

□

5. Die trigonometrischen Funktionen

$\sin : \mathbb{R} \rightarrow \mathbb{R}$ und $\cos : \mathbb{R} \rightarrow \mathbb{R}$ sind differenzierbar und es gilt $\cos' = -\sin$ und $\sin' = \cos$.

Beweis. Aus Übungsaufgabe 36 wissen wir, dass $\lim_{h \rightarrow 0} \frac{\sin(h)}{h} = 1$ und $\lim_{h \rightarrow 0} \frac{\cos(h)-1}{h} = 0$. Aus den Additionstheoremen folgt daher für die Differenzenquotienten:

$$\begin{aligned} \frac{\sin(x+h) - \sin(x)}{h} &= \frac{\sin(x)\cos(h) + \cos(x)\sin(h) - \sin(x)}{h} \\ &= \cos(x) \cdot \frac{\sin(h)}{h} + \sin(x) \cdot \frac{\cos(h) - 1}{h} \\ &\xrightarrow{h \rightarrow 0} \cos(x) \end{aligned}$$

und

$$\begin{aligned} \frac{\cos(x+h) - \cos(x)}{h} &= \frac{\cos(x)\cos(h) - \sin(x)\sin(h) - \cos(x)}{h} \\ &= \cos(x) \cdot \frac{\cos(h) - 1}{h} - \sin(x) \cdot \frac{\sin(h)}{h} \\ &\xrightarrow{h \rightarrow 0} -\sin(x). \end{aligned}$$

□

6. Die Hyperbelfunktionen

$\sinh : \mathbb{R} \rightarrow \mathbb{R}$ und $\cosh : \mathbb{R} \rightarrow \mathbb{R}$ sind differenzierbar und es gilt $\sinh' = \cosh$ und $\cosh' = \sinh$.

Beweis. Nach Definition gilt $\sinh(x) = \frac{1}{2}(e^x - e^{-x})$ und $\cosh(x) = \frac{1}{2}(e^x + e^{-x})$. Daraus folgt sofort

$$\begin{aligned} (\sinh)'(x) &= \frac{1}{2}(e^x + e^{-x}) = \cosh(x), \\ (\cosh)'(x) &= \frac{1}{2}(e^x - e^{-x}) = \sinh(x). \end{aligned}$$

□

7. Die Umkehrfunktionen der trigonometrischen Funktionen

$\arcsin|_{(-1,1)}$, $\arccos|_{(-1,1)}$, \arctan und arccot sind differenzierbar und es gilt:

$$\begin{aligned} \arcsin'(x) &= \frac{1}{\sqrt{1-x^2}}, & \arccos'(x) &= -\frac{1}{\sqrt{1-x^2}} & \text{für } x \in (-1, 1), \\ \arctan'(x) &= \frac{1}{1+x^2}, & \operatorname{arccot}'(x) &= -\frac{1}{1+x^2}. \end{aligned}$$

Beweis. Wir beweisen beispielhaft die Formel für \arctan :

Die Funktion $\tan := \frac{\sin}{\cos} : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}$ ist nach Quotientenregel differenzierbar und es gilt

$$\tan'(x) = \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = 1 + \tan^2(x) > 0.$$

Die Funktion \arctan ist stetig. Wir können also Satz 5.4 anwenden und erhalten

$$\arctan'(y) = \frac{1}{\tan'(\arctan(y))} = \frac{1}{1 + \tan^2(\arctan(y))} = \frac{1}{1 + y^2}. \quad \square$$

Die höheren Ableitungen einer Funktion

Wir definieren die höheren Ableitungen induktiv:

Definition 5.2. Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ differenzierbar und $f' : I \rightarrow \mathbb{R}$ ihre Ableitung. Ist f' in $x_0 \in I$ differenzierbar, so heißt

$$f''(x_0) := f^{(2)}(x_0) := \frac{d^2 f}{dx^2}(x_0) := (f')'(x_0)$$

die 2. Ableitung von f in x_0 . Ist f' in jedem Punkt von I differenzierbar, so heißt f 2-mal differenzierbar und die Funktion

$$\begin{aligned} f'' : I &\longrightarrow \mathbb{R} \\ x &\longmapsto f''(x) \end{aligned}$$

die 2. Ableitung von f .

Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ n -mal differenzierbar ($n \in \mathbb{N}$, $n \geq 2$) und $f^{(n)} : I \rightarrow \mathbb{R}$ ihre n -te Ableitung. Ist $f^{(n)}$ in $x_0 \in I$ differenzierbar, so nennt man

$$f^{(n+1)}(x_0) := \frac{d^{n+1} f}{dx^{n+1}}(x_0) := (f^{(n)})'(x_0)$$

die $(n+1)$ -te Ableitung von f in x_0 . Ist $f^{(n)}$ in jedem Punkt von I differenzierbar, so heißt f $(n+1)$ -mal differenzierbar und die Funktion

$$\begin{aligned} f^{(n+1)} : I &\longrightarrow \mathbb{R} \\ x &\longmapsto f^{(n+1)}(x) \end{aligned}$$

die $(n+1)$ -te Ableitung von f .

Definition 5.3. Eine Funktion $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ heißt stetig differenzierbar, wenn f differenzierbar und $f' : I \rightarrow \mathbb{R}$ stetig ist. $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ heißt n -mal stetig differenzierbar, falls alle Ableitungen $f^{(1)}, f^{(2)}, \dots, f^{(n)}$ existieren und stetig sind.

Eine Funktion $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ heißt glatt, wenn sie beliebig oft differenzierbar ist, das heißt, wenn für alle $n \in \mathbb{N}$ die Ableitungen $f^{(n)}$ existieren (und somit stetig sind).

Bezeichnungen:

$C^0(I, \mathbb{R}) :=$ Vektorraum aller stetigen Funktionen von I nach \mathbb{R}

$C^k(I, \mathbb{R}) :=$ Vektorraum aller k -fach stetig differenzierbaren Funktionen von I nach \mathbb{R} ,
 $k \in \mathbb{N}$.

$C^\infty(I, \mathbb{R}) :=$ Vektorraum aller glatten Funktionen von I nach \mathbb{R} .

Beispiel: Eine differenzierbare, aber nicht stetig differenzierbare Funktion.

Wir betrachten die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, definiert durch

$$f(x) := \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & \text{für } x \neq 0 \\ 0 & \text{für } x = 0. \end{cases}$$

Diese Funktion f ist differenzierbar, aber f' ist in 0 nicht stetig. Um dies einzusehen, betrachten wir zunächst $x \neq 0$. Dann ist

$$f'(x) = 2x \cdot \sin\left(\frac{1}{x}\right) + x^2 \cdot \cos\left(\frac{1}{x}\right) \cdot \left(-\frac{1}{x^2}\right) = 2x \cdot \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right).$$

Für den Differenzenquotienten von f in $x = 0$ gilt

$$\frac{f(h) - f(0)}{h} = \frac{h^2 \cdot \sin\left(\frac{1}{h}\right)}{h} = h \cdot \underbrace{\sin\left(\frac{1}{h}\right)}_{\text{beschränkt}} \xrightarrow{h \rightarrow 0} 0,$$

somit ist $f'(0) = 0$. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ ist also differenzierbar. Da $\cos\left(\frac{1}{x}\right)$ für $x \rightarrow 0$ keinen Grenzwert hat, hat auch f' keinen Grenzwert in $x = 0$. Somit ist die Ableitung f' in $x = 0$ nicht stetig. Insbesondere ist f auch nicht 2-mal differenzierbar in $x = 0$.

5.2 Die Mittelwertsätze der Differentialrechnung und Anwendungen

In diesem Abschnitt werden wir die Mittelwertsätze der Differentialrechnung für reellwertige differenzierbare Funktionen $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ beweisen und mit ihrer Hilfe Aussagen über den Kurvenverlauf des durch f definierten Funktionsgraphen

$$\Gamma_f := \{(x, f(x) \mid x \in I\} \subset \mathbb{R}^2 \text{ machen.}$$

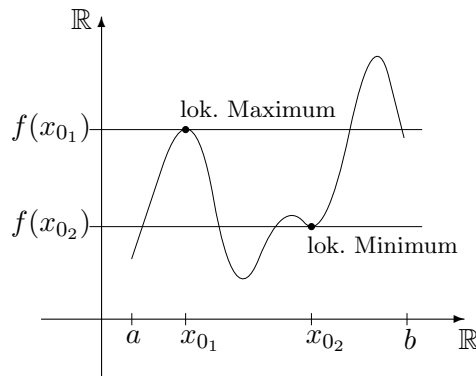
Definition 5.4. Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$.

a) Man sagt: f nimmt in $x_0 \in I$ ein lokales Maximum (bzw. lokales Minimum) an, falls ein $\varepsilon > 0$ existiert, so dass

$$f(y) \leq f(x_0) \quad (\text{bzw. } f(y) \geq f(x_0)) \quad \text{für alle } y \in I \text{ mit } |y - x_0| < \varepsilon. \quad (*)$$

b) Falls Gleichheit in (*) nur für $x = x_0$ eintritt, so heißt das lokale Maximum (Minimum) isoliert (oder strikt).

c) Ein lokales Maximum oder Minimum nennt man auch lokales Extremum.



Satz 5.5 Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$. Hat f in $x_0 \in (a, b) \subset I$ ein lokales Maximum oder lokales Minimum und ist f in x_0 differenzierbar, so gilt $f'(x_0) = 0$.

Beweis. Sei $f(x_0)$ ein lokales Maximum von f . Dann existiert ein $\varepsilon > 0$, so dass

$$\begin{aligned} \frac{f(x)-f(x_0)}{x-x_0} &\leq 0 && \text{für alle } x > x_0 \text{ mit } |x-x_0| < \varepsilon && \text{und} \\ \frac{f(x)-f(x_0)}{x-x_0} &\geq 0 && \text{für alle } x < x_0 \text{ mit } |x-x_0| < \varepsilon. \end{aligned}$$

Für die einseitigen Grenzwerte folgt daraus:

$$\lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \leq 0, \quad \text{und} \quad \lim_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \geq 0. \quad (*)$$

Da der Grenzwert $f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x)-f(x_0)}{x-x_0}$ nach Voraussetzung existiert, muss gelten

$$f'(x_0) = \lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0}.$$

Aus (*) folgt dann $f'(x_0) = 0$. □

Aus $f'(x_0) = 0$ folgt im Allgemeinen nicht, dass f in x_0 ein lokales Maximum oder Minimum hat. Wir betrachten zum Beispiel $f(x) = x^3$. Dann ist $f'(0) = 0$, aber 0 ist kein lokaler Extremwert von f . Zu beachten ist außerdem, dass die Aussage von Satz 5.5 nicht für die Randpunkte des Definitionsbereiches I von f gilt.

Satz 5.6 (Satz von Rolle) Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und auf (a, b) differenzierbar. Weiterhin gelte $f(a) = f(b)$. Dann existiert ein Punkt $x_0 \in (a, b)$ mit $f'(x_0) = 0$.

Beweis. Ist f konstant auf $[a, b]$, so ist die Behauptung trivial. Ist f nicht konstant, so existiert ein $x_1 \in (a, b)$, so dass

$$f(x_1) > f(a) = f(b) \quad \text{oder} \quad f(x_1) < f(a) = f(b).$$

Sei $f(x_1) > f(a) = f(b)$. Da $f : [a, b] \rightarrow \mathbb{R}$ stetig und $[a, b]$ kompakt ist, nimmt f auf $[a, b]$ ein globales Maximum in einem Punkt $x_0 \in [a, b]$ an. Das Maximum wird wegen obiger Annahme nicht auf den Randpunkten von $[a, b]$ angenommen, d.h. $x_0 \in (a, b)$. Aus Satz 5.5 folgt dann $f'(x_0) = 0$. Der Beweis für den Fall $f(x_1) < f(a) = f(b)$ verläuft analog mit Hilfe des globalen Minimums. □

Satz 5.7 (Verallgemeinerter Mittelwertsatz von Cauchy)

Seien $f, g : [a, b] \rightarrow \mathbb{R}$ stetig und auf (a, b) differenzierbar. Dann existiert ein $\xi \in (a, b)$ mit

$$(f(b) - f(a)) \cdot g'(\xi) = (g(b) - g(a)) \cdot f'(\xi). \quad (*)$$

Ist $g'(x) \neq 0$ für alle $x \in (a, b)$, so gilt auch $g(a) \neq g(b)$, d.h. (*) ist äquivalent zu

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

Beweis. Die 2. Behauptung folgt aus dem Satz von Rolle. Zum Beweis der ersten Behauptung betrachten wir die Funktion $\varphi : [a, b] \rightarrow \mathbb{R}$:

$$\varphi(x) := (f(b) - f(a)) \cdot g(x) - (g(b) - g(a)) \cdot f(x), \quad x \in [a, b].$$

Dann ist φ stetig in $[a, b]$ und differenzierbar in (a, b) und es gilt

$$\varphi(a) = f(b)g(a) - g(b)f(a) = \varphi(b).$$

Nach dem Satz von Rolle existiert ein $\xi \in (a, b)$ mit $\varphi'(\xi) = 0$. Wir erhalten somit

$$0 = \varphi'(\xi) = (f(b) - f(a))g'(\xi) - (g(b) - g(a))f'(\xi). \quad \square$$

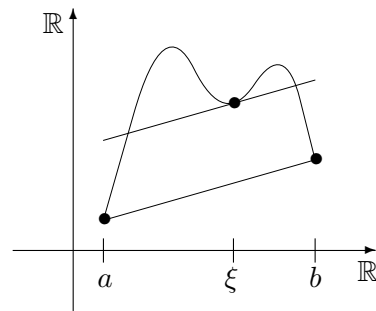
Satz 5.8 (Mittelwertsatz von Lagrange)

Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und auf (a, b) differenzierbar. Dann existiert ein $\xi \in (a, b)$, so dass

$$\frac{f(b) - f(a)}{b - a} = f'(\xi).$$

Beweis. Folgt aus Satz 5.7 mit $g(x) := x$. \square

Geometrisch besagt dieser Mittelwertsatz, dass ein $\xi \in (a, b)$ existiert, so dass der Anstieg der Tangente in $(\xi, f(\xi))$ gleich dem Anstieg der Sekante durch $(a, f(a))$ und $(b, f(b))$ ist.



Als erste Anwendung des Mittelwertsatzes erhalten wir

Satz 5.9 Jede differenzierbare Funktion $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ mit beschränkter Ableitung ist Lipschitzstetig. Insbesondere ist jede stetig differenzierbare Funktion $f : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$ Lipschitzstetig.

Beweis. Sei $f : I \rightarrow \mathbb{R}$ differenzierbar mit beschränkter Ableitung. Dann existiert eine Konstante $C \in \mathbb{R}^+$ mit $|f'(x)| \leq C$ für alle $x \in I$. Seien $x_1, x_2 \in I$ mit $x_1 < x_2$. Dann folgt aus dem Mittelwertsatz 5.8, dass ein $\xi \in (x_1, x_2)$ existiert, so dass

$$|f(x_2) - f(x_1)| = |f'(\xi)| \cdot |x_2 - x_1| \leq C \cdot |x_2 - x_1|.$$

Dies aber bedeutet, dass f Lipschitzstetig mit der Lipschitz-Konstanten C ist. Sei nun $f : [a, b] \rightarrow \mathbb{R}$ stetig differenzierbar. Dann ist f' stetig. Da $[a, b]$ kompakt ist, nimmt f' auf $[a, b]$ ein Maximum an, d.h. f' ist beschränkt. \square

Wir werden den Mittelwertsatz von Lagrange jetzt anwenden, um Aussagen über den Kurvenverlauf des durch $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ definierten Graphen $\Gamma = \Gamma_f \subset \mathbb{R}^2$ zu machen.

Satz 5.10 Sei $f : (a, b) \rightarrow \mathbb{R}$ eine differenzierbare Funktion³. Dann gilt:

1. Ist $f'(x) \geq 0$ für alle $x \in (a, b)$, so ist f monoton wachsend.
2. Ist $f'(x) > 0$ für alle $x \in (a, b)$, so ist f streng monoton wachsend.
3. Ist $f'(x) \leq 0$ für alle $x \in (a, b)$, so ist f monoton fallend.
4. Ist $f'(x) < 0$ für alle $x \in (a, b)$, so ist f streng monoton fallend.
5. Ist $f' \equiv 0$ auf (a, b) , so ist f konstant.

Beweis. Seien $x_1, x_2 \in (a, b)$ und $x_1 < x_2$. Wir wenden den Mittelwertsatz von Lagrange auf $f|_{[x_1, x_2]}$ an. Nach diesem Satz existiert ein $\xi \in (x_1, x_2)$, so dass

$$f'(\xi) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Da $x_2 - x_1 > 0$, folgen die Behauptungen. □

Satz 5.11 Sei $f : (a, b) \rightarrow \mathbb{R}$ eine differenzierbare Funktion, die in $x_0 \in (a, b)$ zweimal differenzierbar ist. Es gelte

1. $f'(x_0) = 0$ und
2. $f''(x_0) < 0$ (bzw. $f''(x_0) > 0$).

Dann hat f in x_0 ein isoliertes lokales Maximum (bzw. Minimum).

Beweis. Wir beweisen nur den Fall $f''(x_0) < 0$. Der Beweis für $f''(x_0) > 0$ verläuft analog. Aus

$$f''(x_0) = \lim_{x \rightarrow x_0} \frac{f'(x) - f'(x_0)}{x - x_0} < 0$$

folgt, dass ein $\delta > 0$ existiert, so dass

$$\frac{f'(x) - f'(x_0)}{x - x_0} < 0$$

für alle $x \in (a, b)$ mit $0 < |x - x_0| < \delta$. Da $f'(x_0) = 0$, erhält man

$$\begin{aligned} f'(x) < 0, & \quad \text{für alle } x > x_0 \text{ mit } x - x_0 < \delta & \quad \text{und} \\ f'(x) > 0, & \quad \text{für alle } x < x_0 \text{ mit } x_0 - x < \delta. \end{aligned}$$

Nach Satz 5.10 ist daher $f|_{(x_0-\delta, x_0)}$ streng monoton wachsend, während $f|_{(x_0, x_0+\delta)}$ streng monoton fallend ist. Somit hat f in x_0 ein isoliertes lokales Maximum. □

Definition 5.5. Eine Funktion $f : (a, b) \rightarrow \mathbb{R}$ heißt

- konvex, falls für alle $x_1, x_2 \in (a, b)$ und $\lambda \in (0, 1)$ gilt:

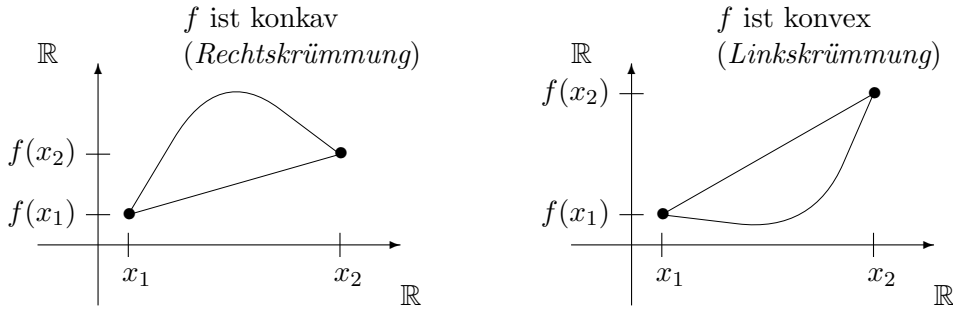
$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

³ Sofern nichts anderes gesagt wird, lassen wir im Folgenden bei offenen Intervallen (a, b) auch $a, b \in \{\pm\infty\}$ zu.

- *konkav, falls für alle $x_1, x_2 \in (a, b)$ und $\lambda \in (0, 1)$ gilt:*

$$f(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

Ist f konvex, so liegt der Graph von f , $\Gamma_f := \{(x, f(x)) \mid x \in (a, b)\} \subset \mathbb{R}^2$, für beliebige $x_1, x_2 \in (a, b)$ unterhalb der Geraden durch $(x_1, f(x_1))$ und $(x_2, f(x_2))$. Ist f konkav, so liegt der Graph Γ_f für beliebige $x_1, x_2 \in (a, b)$ oberhalb der Geraden durch $(x_1, f(x_1))$ und $(x_2, f(x_2))$.



Satz 5.12 Sei $f : (a, b) \rightarrow \mathbb{R}$ zweimal differenzierbar. Dann gilt:

1. f ist genau dann konvex, wenn $f''(x) \geq 0$ für alle $x \in (a, b)$.
2. f ist genau dann konkav, wenn $f''(x) \leq 0$ für alle $x \in (a, b)$.

Beweis. Wir zeigen nur die erste Behauptung. Der Beweis der zweiten Behauptung wird analog geführt.

(\Leftarrow) Sei $f''(x) \geq 0$ für alle $x \in (a, b)$. Nach Satz 5.10 ist f' auf (a, b) monoton wachsend. Seien $x_1, x_2 \in (a, b)$, $x_1 < x_2$ und $\lambda \in (0, 1)$. Wir setzen $x := \lambda x_1 + (1 - \lambda)x_2$. Dann ist $x_1 < x < x_2$. Nach dem Mittelwertsatz existieren ein $\xi_1 \in (x_1, x)$ und ein $\xi_2 \in (x, x_2)$, so dass

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1) \leq f'(\xi_2) = \frac{f(x_2) - f(x)}{x_2 - x}.$$

Da $x - x_1 = (1 - \lambda)(x_2 - x_1)$ und $x_2 - x = \lambda(x_2 - x_1)$, erhalten wir

$$\frac{f(x) - f(x_1)}{1 - \lambda} \leq \frac{f(x_2) - f(x)}{\lambda}.$$

Daraus folgt

$$f(x) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

Somit ist f konvex.

(\Rightarrow) Sei f konvex. Wir nehmen an, dass ein $x_0 \in (a, b)$ existiert mit $f''(x_0) < 0$. Sei $c := f'(x_0)$ und $\varphi : (a, b) \rightarrow \mathbb{R}$ die Funktion $\varphi(x) := f(x) - c(x - x_0)$. Dann ist φ zweimal differenzierbar und es gilt $\varphi'(x_0) = 0$ und $\varphi''(x_0) < 0$. Nach Satz 5.11 hat φ daher in x_0 ein isoliertes lokales Maximum, das heißt, es existiert ein $h > 0$ mit $[x_0 - h, x_0 + h] \subset (a, b)$, so dass $\varphi(x_0 - h) < \varphi(x_0)$ und $\varphi(x_0 + h) < \varphi(x_0)$. Deshalb gilt

$$f(x_0) = \varphi(x_0) > \frac{1}{2}(\varphi(x_0 - h) + \varphi(x_0 + h)) = \frac{1}{2}(f(x_0 - h) + f(x_0 + h)).$$

Mit $\lambda = \frac{1}{2}$, $x_1 = x_0 - h$ und $x_2 = x_0 + h$ folgt, dass f nicht konvex ist.

Definition 5.6. Sei $f : (a, b) \rightarrow \mathbb{R}$ eine stetige Funktion. Der Punkt $x_0 \in (a, b)$ heißt Wendepunkt von f , wenn ein Intervall $(x_0 - \varepsilon, x_0 + \varepsilon) \subset (a, b)$ existiert, so dass entweder $f|_{(x_0 - \varepsilon, x_0)}$ konkav und $f|_{(x_0, x_0 + \varepsilon)}$ konvex ist oder Umgekehrtes gilt.

Satz 5.13 Sei $f : (a, b) \subset \mathbb{R} \rightarrow \mathbb{R}$ zweimal stetig differenzierbar. Ist $x_0 \in (a, b)$ ein Wendepunkt von f , so gilt $f''(x_0) = 0$.

Beweis. Die Behauptung folgt aus Satz 5.12. □

Abschließend beweisen wir mit Hilfe des verallgemeinerten Mittelwertsatzes von Cauchy die *Regeln von L'Hospital*. Diese Regeln liefern ein einfaches Verfahren, Grenzwerte von Brüchen zweier Funktionen zu bestimmen, wenn bei Limesbildung Ausdrücke der Form $\frac{0}{0}$ oder $\frac{\infty}{\infty}$ auftreten.

Satz 5.14 (Die Regeln von L'Hospital)

Seien $f, g : (a, b) \rightarrow \mathbb{R}$ differenzierbare Funktionen mit den folgenden Eigenschaften:

- a) $\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^+} g(x) \in \{0, \pm\infty\}$,
- b) $g'(x) \neq 0$ für alle und $x \in (a, b)$,
- c) Es existiert $\lim_{x \rightarrow a^+} \frac{f'(x)}{g'(x)} = c \in \mathbb{R} \cup \{\pm\infty\}$.

Dann existiert auch $\lim_{x \rightarrow a^+} \frac{f(x)}{g(x)} \in \mathbb{R} \cup \{\pm\infty\}$ und es gilt $\lim_{x \rightarrow a^+} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a^+} \frac{f'(x)}{g'(x)} = c$. Entsprechendes gilt für $x \rightarrow b^-$.

Beweis. Wir zeigen die Behauptung für den Fall $x \rightarrow a^+ \in \mathbb{R}$ mit $a \in \mathbb{R}$. Die Behauptung für $x \rightarrow b^-$ mit $b \in \mathbb{R}$ wird analog geführt. Die Aussage für $x \rightarrow +\infty$ erhält man durch Substitution $x := \frac{1}{y}$ aus der Aussage für $y \rightarrow 0^+$. die Aussage für $x \rightarrow -\infty$ durch Substitution aus derjenigen für $y \rightarrow 0^-$.

1. Fall: $\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^+} g(x) = 0$.

Wir setzen f und g in a durch $f(a) := g(a) := 0$ fort. Dann sind $f, g : [a, b) \rightarrow \mathbb{R}$ stetig. Nach dem verallgemeinerten Mittelwertsatz (Satz 5.7) gibt es zu jedem $x \in (a, b)$ ein $\xi(x) \in (a, x)$, so dass

$$\frac{f(x) - f(a)}{g(x) - g(a)} = \frac{f(x)}{g(x)} = \frac{f'(\xi(x))}{g'(\xi(x))}.$$

$x \rightarrow a^+$ impliziert $\xi(x) \rightarrow a^+$, woraus sich die Behauptung ergibt.

2. Fall: $\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^+} g(x) = +\infty$ und $\lim_{x \rightarrow a^+} \frac{f'(x)}{g'(x)} = c \in \mathbb{R}$.

Zu $\varepsilon > 0$ wählen wir ein $\delta > 0$, so dass

$$\left| \frac{f'(t)}{g'(t)} - c \right| < \frac{\varepsilon}{2} \quad \text{für alle } t \in (a, a + \delta).$$

Nach dem verallgemeinerten Mittelwertsatz gilt dann für beliebige Punkte $x, y \in (a, a + \delta)$ mit $x \neq y$:

$$\left| \frac{f(x) - f(y)}{g(x) - g(y)} - c \right| < \frac{\varepsilon}{2}.$$

Durch Ausmultiplizieren prüft man folgende Formel nach:

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(y)}{g(x) - g(y)} \cdot \frac{1 - \frac{g(y)}{g(x)}}{1 - \frac{f(y)}{f(x)}}. \quad (*)$$

Da $f(x)$ und $g(x)$ für $x \rightarrow a^+$ gegen $+\infty$ streben, konvergiert der 2. Faktor auf der rechten Seite von (*) für $x \rightarrow a^+$ gegen 1. Insbesondere gibt es ein $\delta^* > 0$, so dass für alle $x \in (a, a + \delta^*)$ gilt:

$$\left| \frac{f(x)}{g(x)} - \frac{f(x) - f(y)}{g(x) - g(y)} \right| < \frac{\varepsilon}{2}.$$

Für $x \in (a, a + \min\{\delta, \delta^*\})$ ergibt sich mit der Dreiecks-Ungleichung:

$$\left| \frac{f(x)}{g(x)} - c \right| \leq \left| \frac{f(x)}{g(x)} - \frac{f(x) - f(y)}{g(x) - g(y)} \right| + \left| \frac{f(x) - f(y)}{g(x) - g(y)} - c \right| < \varepsilon.$$

Dies zeigt die Behauptung im 2. Fall.

3. Fall: $\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^+} g(x) = +\infty$ und $\lim_{x \rightarrow a^+} \frac{f'(x)}{g'(x)} = +\infty$.

Dann gilt $\lim_{x \rightarrow a^+} \frac{g'(x)}{f'(x)} = 0$ und aus dem 2. Fall folgt sofort $\lim_{x \rightarrow a^+} \frac{g(x)}{f(x)} = 0$. Nach Voraussetzung sind f und g in einer Umgebung von a positiv. Folglich gilt $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = +\infty$.

4. Fall: $\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^+} g(x) = +\infty$ und $\lim_{x \rightarrow a^+} \frac{f'(x)}{g'(x)} = -\infty$.

Dieser Fall kann nicht auftreten, da sonst f oder g in einer Umgebung von a monoton fallend wären und somit nicht gegen $+\infty$ konvergieren würden. \square

Mit Hilfe der Regeln von L'Hospital können wir Grenzwerte von Funktionen und damit auch von Folgen in vielen Fällen leichter ausrechnen, als es mit den Methoden der vorigen Kapitel möglich war.

Beispiele:

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

2. Für jede positive reelle Zahl α gilt:

$$\begin{aligned} \lim_{x \rightarrow \infty} \frac{\ln(x)}{x^\alpha} &= \lim_{x \rightarrow \infty} \frac{1}{x \cdot \alpha x^{\alpha-1}} = \lim_{x \rightarrow \infty} \frac{1}{\alpha x^\alpha} = 0, \\ \lim_{x \rightarrow \infty} \frac{e^x}{x^\alpha} &= \lim_{x \rightarrow \infty} \frac{e^x}{\alpha x^{\alpha-1}} = \dots = \lim_{x \rightarrow \infty} \frac{e^x}{\alpha(\alpha-1) \cdot \dots \cdot (\alpha - [\alpha] - 1) \cdot x^{\alpha - ([\alpha] + 1)}} = +\infty, \end{aligned}$$

wobei $[\alpha]$ die größte ganze Zahl k mit $k \leq \alpha$ bezeichnet.

$$3. \quad \lim_{x \rightarrow \infty} x^{\frac{1}{x}} = \lim_{x \rightarrow \infty} e^{\ln(x) \cdot \frac{1}{x}} = e^{\lim_{x \rightarrow \infty} \frac{\ln(x)}{x}} = e^0 = 1.$$

4. $\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = \lim_{x \rightarrow 0} \frac{\frac{1}{1+x}}{1} = 1.$
5. $\lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}} = \lim_{x \rightarrow 0} e^{\frac{\ln(1+x)}{x}} = e^{\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x}} = e^1 = e.$
6. $\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = \lim_{x \rightarrow 0} \frac{e^x}{1} = 1.$
7. $\lim_{x \rightarrow 0} (\cos x)^{\frac{1}{x}} = \lim_{x \rightarrow 0} e^{\frac{1}{x} \ln(\cos x)} = e^{\lim_{x \rightarrow 0} \frac{1}{x} \ln(\cos x)} = e^{\lim_{x \rightarrow 0} \frac{1}{\cos x} \cdot (-\sin x)} = e^0 = 1.$
8. $\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \cos(x) = 1.$

5.3 Differenzierbarkeit von Potenzreihen

Jedes reelle Polynom $P(x) = \sum_{k=0}^n a_k(x-x_0)^k$ ist differenzierbar mit der Ableitung

$P'(x) = \sum_{k=1}^n k a_k(x-x_0)^{k-1}$ für alle $x \in \mathbb{R}$. Wir verallgemeinern dies jetzt auf Potenzreihen.

Satz 5.15 Sei $P(x) := \sum_{k=0}^{\infty} a_k(x-x_0)^k$ eine reelle⁴ Potenzreihe mit positivem Konvergenzradius $R > 0$. Dann ist die Funktion $P : (x_0 - R, x_0 + R) \rightarrow \mathbb{R}$ mit

$$P(x) := \sum_{k=0}^{\infty} a_k(x-x_0)^k$$

differenzierbar und für die Ableitung $P' : (x_0 - R, x_0 + R) \rightarrow \mathbb{R}$ gilt

$$P'(x) = \sum_{k=1}^{\infty} k a_k(x-x_0)^{k-1}.$$

Beweis. Wir betrachten die Potenzreihe $Q(x) := \sum_{k=1}^{\infty} k a_k(x-x_0)^{k-1}$.

1. Zunächst überlegen wir uns, dass der Konvergenzradius R_Q von $Q(x)$ ebenfalls R ist: $Q(x)$ hat den gleichen Konvergenzradius wie die Potenzreihe $\hat{Q}(x) := \sum_{k=1}^{\infty} k a_k(x-x_0)^k$.

Da $\limsup_{k \rightarrow \infty} \sqrt[k]{k a_k} = \limsup_{k \rightarrow \infty} \sqrt[k]{a_k}$, stimmen die Konvergenzradien von $\hat{Q}(x)$ und $P(x)$ überein (siehe Satz 3.28). Also gilt $R_Q = R$.

2. Wir zeigen nun, dass P differenzierbar ist mit $P'(x) = Q(x)$ für alle $x \in (x_0 - R, x_0 + R)$.

Wir setzen $p_k(x) := a_k(x-x_0)^k$. Dann gilt $P(x) = \sum_{k=0}^{\infty} p_k(x)$ und $Q(x) = \sum_{k=0}^{\infty} p'_k(x)$.

Sei $\xi \in I_R(x_0) := (x_0 - R, x_0 + R)$ fixiert und $\varepsilon > 0$, so dass $J_\varepsilon(\xi) := [\xi - \varepsilon, \xi + \varepsilon] \subset I_R(x_0)$.

Die Polynome p_k sind auf $J_\varepsilon(\xi)$ stetig differenzierbar, also lipschitzstetig (siehe Satz 5.9), d.h. es gilt für alle $x \in J_\varepsilon(\xi)$:

$$\frac{|p_k(x) - p_k(\xi)|}{|x - \xi|} \leq L_k := \max \{ |p'_k(y)| \mid y \in J_\varepsilon(\xi) \} =: |p'_k(\eta)|,$$

⁴ d.h. $x_0 \in \mathbb{R}, a_0, a_1, \dots \in \mathbb{R}$.

wobei $\eta \in J_\varepsilon(\xi)$. Da $Q(x)$ auf $I_R(x_0)$ absolut konvergent ist, konvergiert die Reihe $\sum_{k=0}^{\infty} |p'_k(\eta)| = \sum_{k=0}^{\infty} L_k$ und es gilt

$$\left| \sum_{k=0}^{\infty} p'_k(x) \right| \leq \sum_{k=1}^{\infty} |p'_k(x)| \leq \sum_{k=0}^{\infty} L_k < \infty \quad \forall x \in J_\varepsilon(\xi).$$

Folglich existiert ein $n_0 \in \mathbb{N}$, so dass

$$\sum_{k=n_0+1}^{\infty} L_k < \frac{\varepsilon}{3} \quad \text{und} \quad \sum_{k=n_0+1}^{\infty} |p'_k(x)| < \frac{\varepsilon}{3} \quad \forall x \in J_\varepsilon(\xi).$$

Wir können jetzt den Differenzenquotienten für P abschätzen:

$$\begin{aligned} \left| \frac{P(x) - P(\xi)}{x - \xi} - Q(\xi) \right| &= \left| \sum_{k=0}^{\infty} \left(\frac{p_k(x) - p_k(\xi)}{x - \xi} - p'_k(\xi) \right) \right| \\ &\leq \sum_{k=0}^{n_0} \left| \frac{p_k(x) - p_k(\xi)}{x - \xi} - p'_k(\xi) \right| + \sum_{k=n_0+1}^{\infty} L_k + \sum_{k=n_0+1}^{\infty} |p'_k(\xi)|. \end{aligned}$$

Die letzten beiden Summanden sind jeweils kleiner als $\frac{\varepsilon}{3}$. Da $p'_k(\xi) = \lim_{x \rightarrow \xi} \frac{p_k(x) - p_k(\xi)}{x - \xi}$, existiert ein $\delta > 0$, so dass der 1. Summand ebenfalls kleiner als $\frac{\varepsilon}{3}$ ist für alle $x \in I_\delta(\xi)$. Folglich gilt

$$\left| \frac{P(x) - P(\xi)}{x - \xi} - Q(\xi) \right| < \varepsilon \quad \forall x \in I_\delta(\xi)$$

und somit $Q(\xi) = \lim_{x \rightarrow \xi} \frac{P(x) - P(\xi)}{x - \xi} = P'(\xi)$. Da ξ beliebig war, folgt $P' = Q$. □

Insbesondere folgt aus dem letzten Satz, dass jede Potenzreihe auf ihrem Konvergenzintervall auch stetig ist:

Folgerung 5.1 Sei $P(x) := \sum_{k=0}^{\infty} a_k(x - x_0)^k$ eine reelle Potenzreihe mit positivem Konvergenzradius $R > 0$. Dann ist die durch die Potenzreihe definierte Funktion $P : (x_0 - R, x_0 + R) \rightarrow \mathbb{R}$ stetig. □

Für reelle Potenzreihen können wir eine zusätzliche Feststellung über das Verhalten der Potenzreihe in den Randpunkten des Konvergenzintervalls machen.

Satz 5.16 (Abelscher Grenzwertsatz)

Sei $P(x) = \sum_{k=0}^{\infty} a_k(x - x_0)^k$ eine reelle Potenzreihe mit positivem Konvergenzradius $R > 0$.

1. Ist $P(x_0 + R)$ konvergent, so ist $P_{(x_0 - R, x_0 + R]}$ in $x_0 + R$ stetig, d.h. es gilt $\lim_{x \rightarrow (x_0 + R)^-} P(x) = P(x_0 + R) = \sum_{k=0}^{\infty} a_k R^k$.
2. Ist $P(x_0 - R)$ konvergent, so ist $P_{[x_0 - R, x_0 + R)}$ in $x_0 - R$ stetig, d.h. es gilt $\lim_{x \rightarrow (x_0 - R)^+} P(x) = P(x_0 - R) = \sum_{k=0}^{\infty} (-1)^k a_k R^k$.

Beweis. Wir beweisen die erste Behauptung⁵. Der Beweis der zweiten Behauptung erfolgt analog. Nach Voraussetzung konvergiert die Folge der Partialsummen $s_m = \sum_{k=0}^m a_k R^k$ gegen $s = P(x_0 + R)$. Wir müssen zeigen, dass für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass $|P(x) - s| < \varepsilon$ für alle $x \in (x_0 - R, x_0 + R)$ mit $x_0 + R - x < \delta$.

Wir setzen $y := x - x_0$. Aus den Eigenschaften der geometrischen Reihe und dem Cauchy-Produkt von Reihen folgt für alle y mit $|y| < R$:

$$\begin{aligned} P(y + x_0) &= \sum_{k=0}^{\infty} a_k y^k \\ &= \left(1 - \frac{y}{R}\right) \left(1 - \frac{y}{R}\right)^{-1} \cdot \sum_{k=0}^{\infty} a_k y^k \\ &= \left(1 - \frac{y}{R}\right) \cdot \underbrace{\sum_{k=0}^{\infty} \left(\frac{y}{R}\right)^k}_{\text{abs. konv.}} \cdot \underbrace{\sum_{k=0}^{\infty} a_k y^k}_{\text{abs. konv.}} \\ &= \left(1 - \frac{y}{R}\right) \sum_{m=0}^{\infty} \left\{ \sum_{\ell=0}^m \left(\frac{y}{R}\right)^{\ell} a_{m-\ell} y^{m-\ell} \right\} \\ &= \left(1 - \frac{y}{R}\right) \sum_{m=0}^{\infty} \left\{ \left(\frac{y}{R}\right)^m \cdot s_m \right\}. \end{aligned}$$

Außerdem gilt

$$s = \left(1 - \frac{y}{R}\right) \left(1 - \frac{y}{R}\right)^{-1} \cdot s = \left(1 - \frac{y}{R}\right) \sum_{m=0}^{\infty} s \cdot \left(\frac{y}{R}\right)^m.$$

Folglich ist für $|y| < R$

$$P(y + x_0) - s = \left(1 - \frac{y}{R}\right) \sum_{m=0}^{\infty} (s_m - s) \left(\frac{y}{R}\right)^m.$$

Sei nun $\varepsilon > 0$ gegeben. Da (s_m) gegen s konvergiert, existiert ein $m_0 \in \mathbb{N}$ mit

$$|s_m - s| < \frac{\varepsilon}{2} \quad \forall m \geq m_0.$$

Wir setzen $M := \sum_{m=0}^{m_0-1} |s_m - s|$ und $\delta := \min\{R, \frac{\varepsilon R}{2M}\}$. Für $|y| < R$ mit $R - y < \delta$ gilt

$$0 < 1 - \frac{y}{R} < \frac{\delta}{R}.$$

Wegen $\delta \leq R$ gilt außerdem $y > 0$. Daraus folgt

⁵ Den Beweis dieses Satzes haben wir in der Vorlesung nicht geführt.

$$\begin{aligned}
|P(y+x_0) - s| &\leq \left(1 - \frac{y}{R}\right) \cdot \left(\sum_{m=0}^{m_0-1} |s_m - s| \cdot \left(\frac{y}{R}\right)^m + \sum_{m=m_0}^{\infty} |s_m - s| \cdot \left(\frac{y}{R}\right)^m\right) \\
&< \frac{\delta}{R} \cdot \sum_{m=0}^{m_0-1} |s_m - s| + \left(1 - \frac{y}{R}\right) \cdot \frac{\varepsilon}{2} \sum_{m=m_0}^{\infty} \left(\frac{y}{R}\right)^m \\
&\leq \frac{\delta}{R} \cdot M + \frac{\varepsilon}{2} \\
&\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.
\end{aligned}$$

Somit ist die Funktion P im Punkt $x_0 + R$ stetig. \square

Für Potenzreihen gilt sogar mehr. Sie sind immer unendlich oft differenzierbar.

Satz 5.17 Sei $P(x) := \sum_{k=0}^{\infty} a_k(x-x_0)^k$ eine reelle Potenzreihe mit positivem Konvergenzradius $R > 0$. Dann ist die Funktion $P : (x_0 - R, x_0 + R) \rightarrow \mathbb{R}$ unendlich oft differenzierbar und es gilt $a_k = \frac{P^{(k)}(x_0)}{k!}$, d.h. :

$$P(x) = \sum_{k=0}^{\infty} \frac{P^{(k)}(x_0)}{k!} (x-x_0)^k.$$

Beweis. Nach Satz 5.15 ist P auf dem Konvergenzintervall $I_R(x_0) := (x_0 - R, x_0 + R)$ differenzierbar und es gilt

$$P'(x) = \sum_{k=1}^{\infty} k a_k (x-x_0)^{k-1} \quad \forall x \in I_R(x_0).$$

P und P' haben den gleichen Konvergenzradius. Wir können deshalb P' erneut ableiten und erhalten P'' mit

$$P''(x) = \sum_{k=2}^{\infty} k(k-1) a_k (x-x_0)^{k-2} \quad \forall x \in I_R(x_0).$$

Induktiv erhalten wir für jedes $n \in \mathbb{N}$: P ist n -mal differenzierbar und für die n -te Ableitung gilt:

$$P^{(n)}(x) = \sum_{k=n}^{\infty} k(k-1)\dots(k-(n-1)) a_k (x-x_0)^{k-n} \quad \forall x \in I_R(x_0).$$

Für das Zentrum x_0 der Potenzreihe folgt $P^{(n)}(x_0) = n! a_n$, also $a_n = \frac{P^{(n)}(x_0)}{n!}$ für alle $n \in \mathbb{N}$. \square

Folgerung 5.2 (Identitätssatz für Potenzreihen)

Es seien $P(x) = \sum_{k=0}^{\infty} a_k(x-x_0)^k$ und $Q(x) = \sum_{k=0}^{\infty} b_k(x-x_0)^k$ zwei reelle Potenzreihen mit dem gleichen Zentrum x_0 und positivem Konvergenzradius R_P bzw. R_Q .

Gibt es kleines offenes Intervall $I_\varepsilon(x_0) \subset I_{R_P}(x_0) \cap I_{R_Q}(x_0)$ um x_0 , auf dem die Funktionen P und Q übereinstimmen, dann gilt $a_k = b_k$ für alle $k \in \mathbb{N}_0$, d.h. $P = Q$. \square

Anwendung: Berechnung von π .

Als Anwendung berechnen wir die Zahl π . Dazu betrachten wir die Funktion $\arctan : \mathbb{R} \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$. Wegen $\tan(\frac{\pi}{4}) = 1$ ist $\arctan(1) = \frac{\pi}{4}$. Wir zeigen nun, dass man die Funktion \arctan auf $[-1, 1]$ als eine Potenzreihe darstellen kann.

$$\arctan x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} \quad \text{für alle } |x| \leq 1.$$

Beweis. Die Funktion \arctan ist differenzierbar und es gilt $\arctan'(x) = \frac{1}{x^2+1}$. Aus den Eigenschaften der geometrischen Reihe erhält man andererseits $\sum_{k=0}^{\infty} (-x^2)^k = \frac{1}{1+x^2}$ für alle $x \in (-1, 1)$. Folglich gilt $\arctan'(x) = \sum_{k=0}^{\infty} (-1)^k x^{2k}$ für alle $x \in (-1, 1)$.

Wir betrachten die Potenzreihe

$$Q(x) := \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}.$$

Der Konvergenzradius von Q ist 1 und es gilt

$$Q'(x) = \sum_{k=0}^{\infty} (-1)^k x^{2k} \quad \text{für alle } x \in (-1, 1).$$

Folglich ist $\arctan'(x) = Q'(x)$ für alle $x \in (-1, 1)$. Da $Q(0) = \arctan(0) = 0$, erhalten wir $\arctan(x) = Q(x)$ für alle $x \in (-1, 1)$ und somit

$$\arctan(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} \quad \text{für alle } |x| < 1.$$

Es bleibt, die Randpunkte $x = \pm 1$ des Konvergenzintervalls zu untersuchen. Nach dem Leibnizkriterium für alternierende Reihen ist die Reihe $Q(1) = \sum_{k=0}^{\infty} (-1)^k \frac{1}{2k+1}$ konvergent. Außerdem gilt $Q(-1) = -Q(1)$. Nach dem Abelschen Grenzwertsatz 5.16 ist damit Q in $x = \pm 1$ stetig und

$$Q(\pm 1) = \lim_{x \rightarrow \pm 1^\mp} Q(x) = \lim_{x \rightarrow \pm 1^\mp} \arctan(x) = \arctan(\pm 1).$$

Wir erhalten zusammenfassend

$$\arctan x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} \quad \text{für alle } |x| \leq 1. \quad \square$$

Da $\frac{\pi}{4} = \arctan(1)$, gilt für π die *Leibnizformel*

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \pm \dots = \sum_{k=0}^{\infty} (-1)^k \frac{1}{2k+1}.$$

Diese Reihe konvergiert sehr langsam gegen $\frac{\pi}{4}$. Man benötigt etwa 1000 Summanden, um π auf 3 Stellen genau zu erhalten. Wesentlich schneller konvergente Reihen mit dem Grenzwert $\frac{\pi}{4}$ erhält man durch den folgenden Trick:

Durch Umkehrung der Additionstheoreme folgt für $x \cdot y \neq 1$

$$\arctan(x) + \arctan(y) = \arctan\left(\frac{x+y}{1-x \cdot y}\right).$$

Wir setzen in diese Gleichung spezielle Werte für x und y ein und erhalten:

$$\begin{aligned} 2 \arctan\left(\frac{1}{5}\right) &= \arctan\left(\frac{5}{12}\right) & (x = y = \frac{1}{5}), \\ 2 \arctan\left(\frac{5}{12}\right) &= \arctan\left(\frac{120}{119}\right) & (x = y = \frac{5}{12}), \\ \arctan(1) + \arctan\left(\frac{1}{239}\right) &= \arctan\left(\frac{120}{119}\right) & (x = 1, y = \frac{1}{239}). \end{aligned}$$

Daraus folgt

$$\frac{\pi}{4} = \arctan(1) = 4 \cdot \arctan\frac{1}{5} - \arctan\frac{1}{239} \quad (*)$$

$$= 4 \cdot \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \left(\frac{1}{5}\right)^{2k+1} - \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} \left(\frac{1}{239}\right)^{2k+1}. \quad (**)$$

Die Formel (*) findet man auch unter dem Namen *Machinsche Formel*. Der englische Astronom John Machin (1680-1751) hat diese Formel 1706 gefunden und mit ihrer Hilfe die ersten 100 Nachkommastellen von π berechnet.

Ist $s = \sum_{k=0}^{\infty} (-1)^k a_k$ eine alternierende Reihe mit monoton fallender Nullfolge positiver reeller Zahlen (a_k) , so kann man den Fehler bei der Näherungsrechnung nach Satz 3.23 folgendermaßen abschätzen:

$$\left| s - \sum_{k=0}^n (-1)^k a_k \right| \leq a_{n+1}.$$

Für den Fehler F_1 bei Addition von 8 Reihengliedern der ersten Reihe in (**) gilt

$$F_1 \leq \underbrace{\frac{4}{17} \cdot \frac{1}{5^{17}}}_{=a_8} < 4 \cdot 10^{-13}.$$

Für den Fehler F_2 bei Addition von 2 Reihengliedern der zweiten Reihe in (**) gilt

$$F_2 \leq \underbrace{\frac{1}{5} \cdot \frac{1}{239^5}}_{a_2} < 3 \cdot 10^{-13}.$$

Wir erhalten also bereits durch die Addition sehr weniger Reihenglieder in (**) eine sehr gute Näherung von π , die weniger als $3 \cdot 10^{-12}$ vom wahren Wert abweicht. Die Addition dieser ersten Reihenglieder ergibt z.B. für die ersten 10 Nachkommastellen von π :

$$\pi = 3.1415926535 + Rest, \quad |Rest| < 10^{-11}.$$

5.4 Taylorpolynome und Extremwertprobleme

Für eine in $x_0 \in I$ differenzierbare Funktion $f : I \rightarrow \mathbb{R}$ existiert die Tangente an den Graphen Γ_f von f im Punkt $P_0 = (x_0, f(x_0))$. Sie wird durch die Funktionsgleichung

$$T_1(f, x_0)(x) := f(x_0) + f'(x_0)(x - x_0), \quad x \in \mathbb{R},$$

beschrieben. Die Funktion $T_1(f, x_0)$ kann man als lineare Approximation von f in der Nähe von x_0 verstehen. Je stärker der Funktionsgraph Γ_f in P_0 gekrümmt ist, desto schlechter ist diese lineare Approximation. Bessere Approximationen erhält man, wenn man statt linearer Funktionen Polynome benutzt.

Definition 5.7. Sei $f : I \rightarrow \mathbb{R}$ in $x_0 \in I$ n -mal differenzierbar. Dann heißt

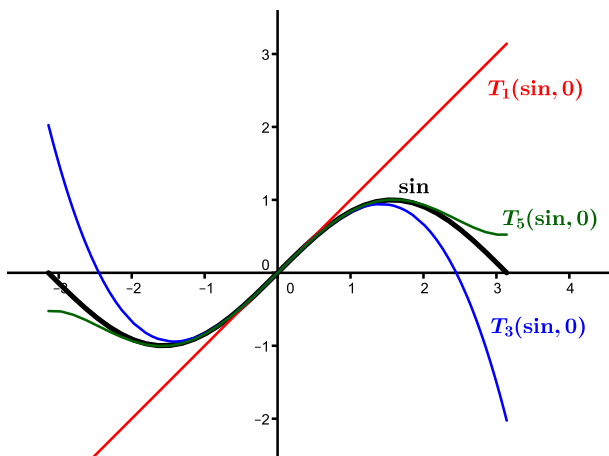
$$\begin{aligned} T_n(f, x_0)(x) &:= \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k \\ &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!} (x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \end{aligned}$$

das n -te Taylorpolynom von f in x_0 .

Das Taylorpolynom für $n = 1$ ist gerade die Tangentenfunktion $T_1(f, x_0)(x)$.

Ist $f''(x_0) \neq 0$, so beschreibt das Taylorpolynom für $n = 2$ eine Parabel, die sogenannte *Schmiegeparabel* an den Funktionsgraphen von f im Punkt $(x_0, f(x_0))$:

$$T_2(f, x_0)(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2.$$



Die Approximation von \sin durch Taylorpolynome nahe $x_0 = 0$:

$$T_1(\sin, 0)(x) = x$$

$$T_3(\sin, 0)(x) = x - \frac{1}{6}x^3$$

$$T_5(\sin, 0)(x) = x - \frac{1}{6}x^3 + \frac{1}{120}x^5.$$

Wir untersuchen jetzt, wie gut das n -te Taylorpolynom $T_n(f, x_0)$ die Funktion f in der Nähe von x_0 approximiert.

Definition 5.8. Sei $f : I \rightarrow \mathbb{R}$ in $x_0 \in I$ n -mal differenzierbar. Dann heißt $R_n(f, x_0)$ mit

$$R_n(f, x_0)(x) := f(x) - T(f, x_0)(x), \quad x \in I,$$

das n -te Restglied von f in x_0 .

Das n -te Restglied $R_n(f, x_0)$ beschreibt den Fehler, den wir bei der Approximation von f durch das n -te Taylorpolynom machen. Wir wollen jetzt Aussagen über die Größe dieses Fehlers machen. Für konkrete Abschätzungen des Fehlers $R_n(f, x_0)$ interessieren uns *explizite* Formeln für das Restglied $R_n(f, x_0)$.

Satz 5.18 (Explizite Formeln für das Restglied)

Sei $f : I \rightarrow \mathbb{R}$ $(n+1)$ -mal differenzierbar und $R_n(f, x_0) := f - T_n(f, x_0)$ das n -te Restglied von f in $x_0 \in I$. Dann existiert für jedes $x \in I$ ein $\xi \in I$ zwischen x_0 und x sowie eine Zahl $\theta \in (0, 1)$, so dass gilt:

$$R_n(f, x_0)(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1} \quad \text{Lagrange-Form des Restgliedes}$$

$$R_n(f, x_0)(x) = \frac{f^{(n+1)}(x_0 + \theta(x - x_0))}{n!} (1 - \theta)^n (x - x_0)^{n+1} \quad \text{Cauchy-Form des Restgliedes}$$

(ξ und θ hängen von x ab!!)

Eine weitere Darstellung für das Restglied $R_n(f, x_0)$, die *Integraldarstellung*, werden wir in Kapitel 6 kennenlernen.

Beweis. Sei $x \in I$ ein fixierter Punkt mit $x \neq x_0$.

1. Wir zeigen die Cauchy-Form für das Restglied.

Wir betrachten dazu die differenzierbare Funktion $g : I \rightarrow \mathbb{R}$ mit

$$g(y) := f(x) - T_n(f, y)(x).$$

Dann gilt $g(x) = 0$, $g(x_0) = R_n(f, x_0)(x)$ und

$$\begin{aligned} g'(y) &= -\frac{d}{dy} \left(\sum_{k=0}^n \frac{f^{(k)}(y)}{k!} (x - y)^k \right) \\ &= -\sum_{k=0}^n \frac{f^{(k+1)}(y)}{k!} (x - y)^k + \sum_{k=1}^n \frac{f^{(k)}(y)}{k!} k (x - y)^{k-1} \\ &= -\frac{(x - y)^n}{n!} f^{(n+1)}(y). \end{aligned}$$

Nach dem Mittelwertsatz der Differentialrechnung existiert ein $\theta \in (0, 1)$, so dass

$$\begin{aligned} g(x) - g(x_0) &= g'(\overbrace{x_0 + \theta(x - x_0)}^{=y}) \cdot (x - x_0) \\ &= -\frac{(x - x_0)^{n+1} (1 - \theta)^n}{n!} f^{(n+1)}(x_0 + \theta(x - x_0)). \end{aligned}$$

Folglich ist

$$R_n(f, x_0)(x) = -(g(x) - g(x_0)) = \frac{f^{(n+1)}(x_0 + \theta(x - x_0))}{n!} (1 - \theta)^n (x - x_0)^{n+1}.$$

Dies ist die Cauchy-Form des Restgliedes.

2. Wir zeigen die Lagrange-Form für das Restglied.

Wir betrachten dazu zusätzlich die differenzierbare Funktion $h : I \rightarrow \mathbb{R}$, definiert durch $h(y) := (x - y)^{n+1}$. Dann gilt $h(x) = 0$, $h(x_0) = (x - x_0)^{n+1}$ und

$$h'(y) = -(n+1)(x - y)^n.$$

Nach dem verallgemeinerten Mittelwertsatz existiert ein ξ zwischen x_0 und x , so dass

$$(g(x) - g(x_0)) \cdot h'(\xi) = (h(x) - h(x_0)) \cdot g'(\xi).$$

Durch Einsetzen von $R_n(f, x_0)(x) = -(g(x) - g(x_0))$ folgt

$$R_n(f, x_0)(x) \cdot (n+1)(x - \xi)^n = (x - x_0)^{n+1} \frac{(x - \xi)^n}{n!} f^{(n+1)}(\xi)$$

und somit

$$R_n(f, x_0)(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}.$$

Dies ist die Lagrange-Form des Restgliedes. □

Die Formel

$$f(x) = T_n(f, x_0)(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}$$

heißt *Taylorformel für f im Punkt x_0* . (Dabei könnte für das Restglied auch jede andere konkrete Darstellung stehen). Für viele Anwendungen reicht auch die folgende qualitative Beschreibung des Restgliedes:

Satz 5.19 (Qualitative Taylorformel)

Sei $f : I \rightarrow \mathbb{R}$ n -mal stetig differenzierbar und $x_0 \in I$. Dann gilt:

$$f(x) = T_n(f, x_0)(x) + \underbrace{(x - x_0)^n \cdot r(x)}_{=R_n(f, x_0)(x)}, \quad (*)$$

wobei $r : I \rightarrow \mathbb{R}$ eine stetige Funktion mit $r(x_0) = 0$ ist. Insbesondere gilt

$$\lim_{x \rightarrow x_0} \frac{R_n(f, x_0)(x)}{(x - x_0)^n} = 0,$$

d.h. das Restglied strebt für $x \rightarrow x_0$ schneller gegen Null als $(x - x_0)^n$.

Beweis. Für $x \neq x_0$ definieren wir

$$r(x) := \frac{R_n(f, x_0)(x)}{(x - x_0)^n}.$$

Dann ist $r : I \setminus \{x_0\} \rightarrow \mathbb{R}$ stetig. Um das Verhalten von r für $x \rightarrow x_0$ zu bestimmen, benutzen wir die Lagrange-Form des Restgliedes $R_{n-1}(f, x_0)(x)$: Es existiert ein ξ zwischen x_0 und x , so dass

$$\begin{aligned}
r(x) &= \frac{f(x) - T_n(f, x_0)(x)}{(x - x_0)^n} \\
&= \frac{f(x) - T_{n-1}(f, x_0)(x) - \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n}{(x - x_0)^n} \\
&= \frac{1}{n!}(f^{(n)}(\xi) - f^{(n)}(x_0)).
\end{aligned}$$

Für $x \rightarrow x_0$ gilt auch $\xi \rightarrow x_0$. Da $f^{(n)}$ stetig ist, folgt $\lim_{x \rightarrow x_0} r(x) = 0$. r ist also durch 0 stetig in x_0 fortsetzbar und es gilt (*). \square

Folgerung 5.3 Sei $f : I \rightarrow \mathbb{R}$ eine n -mal differenzierbare Funktion und $f^{(n)}(x) = 0$ für alle $x \in I$. Dann ist f ein Polynom vom Grad $\leq n - 1$.

Beweis. Wir fixieren ein $x_0 \in I$ und benutzen die Taylorformel für $n - 1$ mit der Lagrange-Form des Restgliedes:

$$f(x) = T_{n-1}(f, x_0)(x) + \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n.$$

Dabei ist ξ eine geeignete Zahl zwischen x_0 und x . Da nach Voraussetzung $f^{(n)}(\xi) = 0$, stimmt f auf I mit dem Taylorpolynom $T_{n-1}(f, x_0)$ überein. Dieses hat den Grad $\leq n - 1$. \square

Anwendung der Taylorformel auf Extremwertprobleme

Als weitere Anwendung der Taylorformel betrachten wir nochmal Extremwert-Probleme. Zunächst beweisen wir ein weiteres notwendiges Kriterium für die Existenz eines lokalen Extremwertes:

Satz 5.20 (Notwendige Bedingung für einen lokalen Extremwert)

Sei $f : (a, b) \rightarrow \mathbb{R}$ 2-mal stetig differenzierbar. Hat f in $x_0 \in (a, b)$ ein lokales Maximum (bzw. lokales Minimum), dann gilt:

- a) $f'(x_0) = 0$ und
- b) $f''(x_0) \leq 0$ (bzw. $f''(x_0) \geq 0$).

Beweis. Die erste Bedingung kennen wir bereits aus Satz 5.5. Für den Beweis der 2. Bedingung betrachten wir die qualitative Taylorformel für $n = 2$:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(x_0)(x - x_0)^2 + (x - x_0)^2 r(x),$$

mit $\lim_{x \rightarrow x_0} r(x) = 0$. Da $f'(x_0) = 0$, gilt

$$f(x) - f(x_0) = \left(\frac{1}{2}f''(x_0) + r(x)\right)(x - x_0)^2.$$

Besitzt f in x_0 ein lokales Maximum, so folgt

$$\frac{1}{2}f''(x_0) + r(x) \leq 0$$

für x hinreichend nahe an x_0 . Der Grenzübergang $x \rightarrow x_0$ liefert dann $f''(x_0) \leq 0$.

Analog behandelt man den Fall des lokalen Minimums. \square

In Satz 5.11 hatten wir bereits ein hinreichendes Kriterium für die Existenz eines lokalen Extremwertes einer 2-mal differenzierbaren Funktion kennengelernt. Dieses Kriterium können wir nun verallgemeinern.

Satz 5.21 (Hinreichende Bedingung für einen lokalen Extremwert)

Sei $f : (a, b) \rightarrow \mathbb{R}$ eine n -mal differenzierbare Funktion, $n \geq 2$, und $x_0 \in (a, b)$ ein Punkt mit $f^{(1)}(x_0) = f^{(2)}(x_0) = \dots = f^{(n-1)}(x_0) = 0$ und $f^{(n)}(x_0) \neq 0$. Dann gilt:

1. Ist n ungerade, so hat f in x_0 keinen lokalen Extremwert.
2. Ist n gerade und $f^{(n)}(x_0) > 0$, so hat f in x_0 ein isoliertes lokales Minimum.
Ist n gerade und $f^{(n)}(x_0) < 0$, so hat f in x_0 ein isoliertes lokales Maximum.

Beweis. Sei zunächst $f^{(n)}(x_0) > 0$. Dann gilt

$$f^{(n)}(x_0) = \lim_{x \rightarrow x_0} \frac{f^{(n-1)}(x) - f^{(n-1)}(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{f^{(n-1)}(x)}{x - x_0} > 0.$$

Deshalb gibt es ein $\delta > 0$, so dass $\frac{f^{(n-1)}(x)}{x - x_0} > 0$ für alle x mit $0 < |x - x_0| < \delta$. Daraus folgt

$$\begin{aligned} f^{(n-1)}(x) &< 0 && \text{auf } (x_0 - \delta, x_0), \\ f^{(n-1)}(x) &> 0 && \text{auf } (x_0, x_0 + \delta). \end{aligned}$$

Wir approximieren f durch das $(n-2)$ -te Taylorpolynom und benutzen die Lagrange-Form des Restgliedes:

$$\begin{aligned} f(x) &= T_{n-2}(f, x_0)(x) + R_{n-2}(f, x_0)(x) \\ &\stackrel{5.18}{=} f(x_0) + \frac{1}{(n-1)!} f^{(n-1)}(\xi)(x - x_0)^{n-1} \end{aligned}$$

für ein geeignetes ξ zwischen x_0 und x . Ist n gerade, so ist $f^{(n-1)}(\xi)(x - x_0)^{n-1} > 0$ für alle x mit $0 < |x - x_0| < \delta$. Somit ist $f(x) > f(x_0)$ für alle x mit $0 < |x - x_0| < \delta$, d.h. f hat in x_0 ein isoliertes lokales Minimum.

Ist n ungerade, so gilt

$$f^{(n-1)}(\xi)(x - x_0)^{n-1} \begin{cases} > 0 & \text{auf } (x_0, x_0 + \delta) \\ < 0 & \text{auf } (x_0 - \delta, x_0). \end{cases}$$

Folglich hat f in x_0 keinen lokalen Extremwert. Mit analogen Argumenten behandelt man den Fall $f^{(n)}(x_0) < 0$. \square

Ist f in $x_0 \in (a, b)$ unendlich oft differenzierbar und gilt $f^{(n)}(x_0) = 0$ für alle n , so kann man keine allgemeinen Aussagen über das Vorliegen eines lokalen Extremwertes machen. Sollen die Extrema von f auf einem *abgeschlossenen* Intervall $[a, b]$ bestimmt werden, so sind außer den Stellen $x_0 \in (a, b)$ mit $f'(x_0) = 0$ auch noch die Intervallenden zu untersuchen. In diesen Intervallenden gilt das Kriterium aus Satz 5.21 nicht!

5.5 Taylorreihen und reell-analytische Funktionen

Wir kennen aus Kapitel 4 bereits einige Funktionen, die durch Potenzreihen definiert sind. Wir wollen nun untersuchen, unter welchen Bedingungen man eine Funktion $f : I \rightarrow \mathbb{R}$ nahe $x_0 \in I$ als Potenzreihe darstellen kann. Nach Satz 5.17 muß f dazu in x_0 unendlich oft differenzierbar sein und als Potenzreihe kommt nur die Reihe $\sum_{k=0}^{\infty} a_k(x - x_0)^k$ mit $a_k = \frac{f^{(k)}(x_0)}{k!}$ in Frage.

Definition 5.9. Sei $f : I \rightarrow \mathbb{R}$ in $x_0 \in I$ unendlich oft differenzierbar. Dann heißt die Reihe

$$T(f, x_0)(x) := \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

Taylorreihe von f in x_0 .

Definition 5.10. Eine Funktion $f : I \rightarrow \mathbb{R}$ heißt in $x_0 \in I$ reell-analytisch, falls ein Intervall $(x_0 - r, x_0 + r) \subset I$ und eine reelle Potenzreihe $\sum_{k=0}^{\infty} a_k(x - x_0)^k$ mit Konvergenzradius $R > 0$ existieren, so dass

$$f(x) = \sum_{k=0}^{\infty} a_k(x - x_0)^k \quad \text{für alle } x \in (x_0 - \min\{r, R\}, x_0 + \min\{r, R\}).$$

Man sagt in diesem Fall auch, dass f in einer Umgebung von x_0 in eine Potenzreihe entwickelbar ist. x_0 heißt dann der Entwicklungspunkt.

Eine auf einem offenen Intervall (a, b) definierte Funktion $f : (a, b) \rightarrow \mathbb{R}$ heißt reell-analytisch, wenn f in jedem Punkt von (a, b) reell-analytisch ist.

$C^\omega((a, b), \mathbb{R})$ bezeichnet den Vektorraum der reell-analytischen Funktionen auf $(a, b) \subset \mathbb{R}$.

Beispiele für reell-analytische Funktionen:

1. Jede reelle Potenzreihe $f(x) := \sum_{k=0}^{\infty} a_k(x - x_0)^k$ mit positivem Konvergenzradius $R > 0$ definiert eine in x_0 reell-analytische Funktion f auf dem Konvergenzintervall.
2. Die Funktionen \exp , \sin , \cos , \sinh und \cosh sind reell-analytisch auf \mathbb{R} .
3. Die Funktion \arctan ist in $x_0 = 0$ reell-analytisch (siehe Kapitel 5.3).

Aus Satz 5.17 folgt

Satz 5.22 Sei $f : I \rightarrow \mathbb{R}$ reell-analytisch in $x_0 \in I$. Dann ist f in einem offenen Intervall (a, b) um x_0 unendlich oft differenzierbar und es gilt:

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} \cdot (x - x_0)^k \quad \forall x \in (a, b).$$

Nach Satz 5.22 hat die Taylorreihe $T(f, x_0)(x)$ einer in x_0 reell-analytischen Funktion f einen positiven Konvergenzradius. Des Weiteren stimmt f in einer Umgebung von x_0 mit ihrer Taylorreihe überein. Beide Eigenschaften sind für C^∞ -Funktionen i.a. nicht erfüllt, wie die folgenden beiden Beispiele zeigen. Es gibt also C^∞ -Funktionen, die nicht reell-analytisch sind:

Beispiel 1: Die Taylorreihe einer C^∞ -Funktion kann den Konvergenzradius $R = 0$ haben:

Wir zitieren dazu einen Satz von Borel: Seien c_0, c_1, c_2, \dots beliebig vorgegebene reelle Zahlen. Dann existiert eine C^∞ -Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f^{(k)}(0) = k! \cdot c_k$, d.h. so dass

$$T(f, 0)(x) = \sum_{k=0}^{\infty} c_k x^k.$$

Wir wählen nun die Koeffizienten c_k so, dass die Potenzreihe $\sum_{n=0}^{\infty} c_n x^n$ Konvergenzradius $R = 0$ hat. Den Beweis des Satzes von Borel kann man z.B. in *R. Narasimham: Analysis on real and complex manifolds, Nord Holland 1968* nachlesen.

Beispiel 2: Eine glatte Funktionen mit überall konvergenter Taylorreihe, die nicht mit ihrer Taylorreihe übereinstimmt:

Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ die Funktion

$$f(x) := \begin{cases} e^{-\frac{1}{x^2}} & \text{für } x > 0, \\ 0 & \text{für } x \leq 0. \end{cases}$$

f ist auf $\mathbb{R} \setminus \{0\}$ unendlich oft differenzierbar. Für die k -ten Ableitungen in $x > 0$ gilt:

$$f^{(k)}(x) = e^{-\frac{1}{x^2}} \cdot \frac{q_k(x)}{x^{3k}},$$

wobei $q_k(x)$ ein Polynom ist x ist (Beweis durch vollständige Induktion über $k \in \mathbb{N}$). Daraus kann man wiederum durch vollständige Induktion zeigen, dass f auch in $x_0 = 0$ unendlich oft differenzierbar ist, wobei für die k -ten Ableitungen in $x_0 = 0$ gilt

$$f^{(k)}(0) = 0 \quad \forall k \in \mathbb{N}.$$

Für die Taylorreihe von f in $x_0 = 0$ erhalten wir damit $T(f, 0)(x) = 0$ für alle $x \in \mathbb{R}$. Nach Definition ist aber $f(x) \neq 0$ für alle $x > 0$.

Satz 5.23 Sei $f : I \rightarrow \mathbb{R}$ in $x_0 \in I$ unendlich oft differenzierbar. Dann gilt:

Die Taylorreihe $T(f, x_0)(x)$ konvergiert für $x \in I$ genau dann gegen $f(x)$, wenn für die Restglieder in der Taylorformel gilt:

$$\lim_{n \rightarrow \infty} R_n(f, x_0)(x) = 0.$$

Beweis. Nach Definition ist das Taylorpolynom $T_n(f, x_0)(x)$ die n -te Partialsumme der Taylorreihe $T(f, x_0)(x)$, d.h. $T(f, x_0)(x)$ konvergiert genau dann, wenn der Grenzwert

$$T(f, x_0)(x) = \lim_{n \rightarrow \infty} T_n(f, x_0)(x)$$

existiert. Da $T_n(f, x_0)(x) = f(x) - R_n(f, x_0)(x)$ folgt

$$f(x) = T(f, x_0)(x) \iff \lim_{n \rightarrow \infty} R_n(f, x_0)(x) = 0.$$

□

Beispiel 3: Die Taylorentwicklung von $f(x) = \ln(1+x)$ in $x_0 = 0$.

Die Funktion $f(x) = \ln(x+1)$ ist in $x_0 = 0$ reell-analytisch und für ihre Taylorentwicklung gilt:

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \cdot x^k \quad \text{für alle } x \in (-1, 1].$$

Insbesondere gilt für die alternierende harmonische Reihe ($x = 1$)

$$\ln(2) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} \pm \dots$$

Beweis. Die Funktion $f(x) := \ln(x+1)$ ist auf $(-1, \infty)$ beliebig oft differenzierbar und es gilt $f(0) = 0$ sowie

$$f^{(k)}(x) = \frac{(k-1)!(-1)^{k+1}}{(1+x)^k} \quad \text{für alle } k \in \mathbb{N}.$$

Für die Taylorreihe von $f(x) = \ln(x+1)$ in $x_0 = 0$ folgt

$$T(f, 0)(x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \cdot x^k.$$

Wir zeigen, dass diese Reihe für alle $x \in (-1, 1)$ gegen $f(x)$ konvergiert. Dazu betrachten wir die Cauchy-Form des Restgliedes: Es existiert ein $\theta \in (0, 1)$ mit

$$R_n(f, 0)(x) = \frac{f^{(n+1)}(\theta x)}{n!} (1-\theta)^n x^{n+1} = \frac{(-1)^n}{(1+\theta x)^{n+1}} (1-\theta)^n x^{n+1}.$$

Ist $|x| < 1$, so gilt $1-\theta < 1-\theta|x|$ und $1+\theta x \geq 1-\theta|x| > 1-|x| > 0$. Daraus folgt für $x \in (-1, 1)$

$$\begin{aligned} |R_n(f, 0)(x)| &= |x|^{n+1} \frac{(1-\theta)^n}{(1+\theta x)^{n+1}} < |x|^{n+1} \frac{(1-\theta|x|)^n}{(1+\theta x)^{n+1}} < \frac{|x|^{n+1}}{1-|x|} \cdot \left(\frac{1-\theta|x|}{1+\theta x} \right)^n \\ &\leq \frac{|x|^{n+1}}{1-|x|}, \end{aligned}$$

und somit $\lim_{n \rightarrow \infty} R_n(f, 0)(x) = 0$. Also konvergiert die Taylorreihe $T(f, 0)(x)$ für $|x| < 1$ gegen $f(x)$:

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k \quad \text{für alle } |x| < 1.$$

Es bleibt die Konvergenz in $x = 1$ zu untersuchen. Nach dem Leibnizkriterium (Satz 3.23) konvergiert die alternierende harmonische Reihe $T(f, 0)(1) = 1 - \frac{1}{2} + \frac{1}{3} \pm \dots$. Wir wenden nun den Abelschen Grenzwertsatz (Satz 5.16) an und erhalten wegen der Stetigkeit von \ln

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

Somit gilt $\ln(1+x) = T(f, 0)(x)$ auch in $x = 1$. □

Beispiel 4: Die Taylorentwicklung von $f(x) := \ln\left(\frac{1+x}{1-x}\right)$.

Die Reihe $\ln(2) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}$ konvergiert sehr langsam, eignet sich also nicht zur Berechnung von $\ln(2)$. Für die numerische Berechnung von Werten der Logarithmusfunktion eignet sich die Taylorentwicklung von $\ln\left(\frac{1+x}{1-x}\right)$.

Die Funktion $\ln\left(\frac{1+x}{1-x}\right)$ ist in $x_0 = 0$ reell-analytisch und für ihre Taylorentwicklung gilt

$$\ln\left(\frac{1+x}{1-x}\right) = \sum_{k=0}^{\infty} \frac{2}{2k+1} x^{2k+1} \quad \text{für alle } x \in (-1, 1).$$

Für den Beweis können wir die Taylorentwicklung aus Beispiel 3 benutzen. Für alle $x \in \mathbb{R}$ mit $|x| < 1$ gilt

$$\begin{aligned} f(x) &= \ln\left(\frac{1+x}{1-x}\right) = \ln(1+x) - \ln(1-x) \\ &= \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k - \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} (-x)^k \\ &= \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} (1 - (-1)^k) x^k \\ &= \sum_{\ell=0}^{\infty} \frac{2}{2\ell+1} x^{2\ell+1}. \end{aligned}$$

Wir bestimmen damit jetzt $\ln(2)$ mit einer Genauigkeit von 4 Stellen nach dem Komma: Es gilt $\ln(2) = f\left(\frac{1}{3}\right) = T(f, 0)\left(\frac{1}{3}\right)$. Wir benutzen das Lagrange-Restglied, um den Fehler bei der Approximation von $\ln(2)$ durch das n -te Taylorpolynom abzuschätzen:

$$\ln(2) = T_n(f, 0)\left(\frac{1}{3}\right) + R_n(f, 0)\left(\frac{1}{3}\right) = T_n(f, 0)\left(\frac{1}{3}\right) + \frac{f^{(n+1)}(\xi)}{(n+1)!} \left(\frac{1}{3}\right)^{n+1},$$

wobei $0 < \xi < \frac{1}{3}$. Für die Ableitung berechnet man

$$f^{(n+1)}(\xi) = n! \left(\frac{(-1)^n}{(1+\xi)^{n+1}} + \frac{1}{(1-\xi)^{n+1}} \right).$$

Für $0 < \xi < \frac{1}{3}$ kann man das Restglied somit abschätzen durch

$$0 < R_n(f, 0)\left(\frac{1}{3}\right) < \frac{1}{n+1} \left(\frac{1}{(1+\xi)^{n+1}} + \frac{1}{(1-\xi)^{n+1}} \right) \left(\frac{1}{3}\right)^{n+1} < \frac{1}{n+1} \left(\frac{1}{3^{n+1}} + \frac{1}{2^{n+1}} \right).$$

Für $n = 10$ erhält man:

$$0 < R_{10}(f, 0)\left(\frac{1}{3}\right) < 5 \cdot 10^{-5}.$$

Für die Näherung von $\ln(2)$ durch das 10. Taylorpolynom $T_{10}(f, 0)\left(\frac{1}{3}\right)$ gilt:

$$\ln(2) \approx T_{10}(f, 0)\left(\frac{1}{3}\right) = 2 \left(\frac{1}{3} + \frac{1}{3 \cdot 3^3} + \frac{1}{5 \cdot 3^5} + \frac{1}{7 \cdot 3^7} + \frac{1}{9 \cdot 3^9} \right) = 0,693146047\dots$$

Somit approximiert das 10. Taylorpolynom die Zahl $\ln(2)$ auf vier Stellen genau.

Beispiel 5: Die Taylorentwicklung von $f(x) := (1+x)^\alpha$ in $x_0 = 0$.

Sei $\alpha \in \mathbb{R}$. Die Funktion $f(x) := (1+x)^\alpha$ ist in $x_0 = 0$ reell-analytisch und für ihre Taylorentwicklung gilt:

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \quad \text{für alle } x \in (-1, 1).$$

Für $\alpha = m \in \mathbb{N}$ ist diese Reihe endlich und beschreibt gerade die binomische Formel

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

Die Taylorentwicklung von $(1+x)^\alpha$ verallgemeinert also die binomische Formel. Die Reihe $B_\alpha(x) := \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$ heißt *Binomialreihe*. (Vergleiche auch Übungsaufgabe 26).

Beweis. Wir bestimmen zunächst wieder die Taylorreihe von f in $x_0 = 0$. $f(x) = (1+x)^\alpha$ ist auf $(-1, \infty)$ beliebig oft differenzierbar und es gilt $f(0) = 1$ sowie

$$f^{(k)}(x) = \alpha(\alpha-1) \cdot \dots \cdot (\alpha-k+1)(1+x)^{\alpha-k}$$

und daher

$$\frac{f^{(k)}(0)}{k!} = \frac{\alpha(\alpha-1) \cdot \dots \cdot (\alpha-k+1)}{k!} = \binom{\alpha}{k}.$$

Die Taylorreihe von $f(x) = (1+x)^\alpha$ in $x_0 = 0$ ist damit

$$T(f, 0)(x) = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k = B_\alpha(x).$$

Wir wissen bereits, dass die Binomialreihe $B_\alpha(x)$ für $|x| < 1$ konvergiert (siehe Übungsaufgabe 26). Wir zeigen nun, dass ihr Grenzwert $f(x) = (1+x)^\alpha$ ist. Wir betrachten dazu wieder die Cauchy-Form des Restgliedes

$$\begin{aligned} R_n(f, 0)(x) &= \frac{f^{(n+1)}(\theta x)}{n!} (1-\theta)^n x^{n+1} \\ &= \frac{\alpha(\alpha-1) \cdot \dots \cdot (\alpha-n)}{n!} (1+\theta x)^{\alpha-n-1} (1-\theta)^n x^{n+1} \\ &= \alpha \cdot \binom{\alpha-1}{n} (1-\theta)^n x^{n+1} (1+\theta x)^{\alpha-n-1} \\ &= \alpha \cdot \binom{\alpha-1}{n} \underbrace{\left(\frac{1-\theta}{1+\theta x} \right)^n}_{<1 \text{ für } |x|<1} x^{n+1} (1+\theta x)^{\alpha-1}. \end{aligned}$$

Daher gilt für $|x| < 1$

$$|R_n(f, 0)(x)| \leq \left| \binom{\alpha-1}{n} x^n \right| \cdot \underbrace{|\alpha x(1+\theta x)^{\alpha-1}|}_{=M \text{ unabh. von } n}.$$

Da die Binomialreihe $B_{\alpha-1}(x)$ für $|x| < 1$ konvergiert, ist $(|\binom{\alpha-1}{n} x^n|)_{n=0}^{\infty}$ eine Nullfolge und wir erhalten

$$\lim_{n \rightarrow \infty} R_n(f, 0)(x) = 0 \quad \text{für alle } |x| < 1.$$

Damit konvergiert die Binomialreihe $B_{\alpha}(x) = T(f, 0)(x)$ für $x \in (-1, 1)$ gegen $(1+x)^{\alpha}$.

□

Integralrechnung für Funktionen einer reellen Variablen

In diesem Kapitel beginnen wir mit der Integralrechnung. Die Integralrechnung wird durch zwei verschiedene Problemstellungen motiviert:

1. *Bestimmung von Längen, Flächeninhalten und Volumen.*

Die Integralrechnung stellt Methoden bereit, mit denen man geometrische Größen wie Längen, Flächeninhalte und Volumen berechnen kann.

2. *Lösung von Differentialgleichungen.*

Viele Prozesse in der Natur werden durch Differentialgleichungen modelliert. Die Aufgabe besteht dann immer darin, aus gegebenen Ableitungen einer Funktion die Funktion selbst zu bestimmen. Dazu muß man den Prozeß des Differenzierens "umkehren". Auch dabei hilft die Integralrechnung.

6.1 Das Riemann-Integral

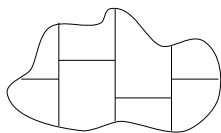
Die erste Motivation für die Entwicklung der Integralrechnung entstand aus dem Versuch, Flächeninhalte und Volumen zu berechnen. Wir werden hier zunächst das Riemann-Integral für reelle Funktionen einer Variablen behandeln, mit dessen Hilfe man Flächeninhalte für Teilmengen des \mathbb{R}^2 berechnen kann. In Kapitel ?? werden wir dann die Integralrechnung für Funktionen mehrerer reeller Variablen behandeln, mit deren Hilfe man Volumen für Teilmengen des \mathbb{R}^3 berechnen kann. Das heute so genannte Riemann-Integral wurde in rigoroser Form von Bernard Riemann in seiner Habilitationsschrift (1854) definiert. Für einen kurzen historischen Überblick zur Entwicklung der Integralrechnung verweisen wir auf das 9. Kapitel des Buches *Analysis I* von W. Walter.

Zunächst formulieren wir einige Forderungen, die ein vernünftig definierter geometrischer Flächeninhalt $\mu(A)$ für Teilmengen $A \subset \mathbb{R}^2$ erfüllen sollte:

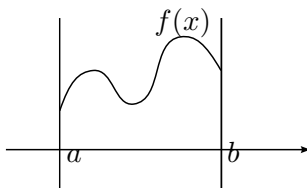
1. $\mu(A) \geq 0$.
2. $A \subset B \implies \mu(A) \leq \mu(B)$.
3. $\mu([a, b] \times [c, d]) = (b - a) \cdot (d - c)$.
4. Haben zwei Teilmengen $A, B \subset \mathbb{R}^2$ höchstens Randpunkte gemeinsam, so gilt $\mu(A \cup B) = \mu(A) + \mu(B)$.

5. $\mu(A)$ ist invariant gegen Euklidische Bewegungen (z.B. Drehungen, Verschiebungen, Spiegelungen).

Legt man diese Eigenschaften zugrunde, so kann man den Flächeninhalt einer Teilmenge $A \subset \mathbb{R}^2$ bestimmen, in dem man sie in Mengen "einfacher" Form zerlegt. Z.B. kann man Mengen der Form



in Teile der Form



zerlegen, wobei die obere Begrenzungskurve durch einen Funktionsgraphen gegeben wird. Es genügt deshalb, den Flächeninhalt von Mengen der Form

$$A = \{(x, y) \in \mathbb{R}^2 \mid x \in [a, b], 0 \leq y \leq f(x)\},$$

zu kennen, wobei f eine nicht-negative Funktion auf dem Intervall $[a, b]$ ist. Dies motiviert die jetzt folgende Definition des so-geannten Riemann-Integrals.

Definition 6.1. Sei $I = [a, b] \subset \mathbb{R}$ ein abgeschlossenes Intervall mit $a < b$.

Eine Menge von Punkten $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ mit $a = x_0 < \dots < x_n = b$ heißt Zerlegung von I . $I_k := [x_{k-1}, x_k]$ bezeichnet das k -te Teilintervall von \mathcal{P} und $L(I_k) := x_k - x_{k-1}$ seine Länge. $\|\mathcal{P}\| := \max\{L(I_k) \mid k = 1, \dots, n\}$ heißt die Feinheit der Zerlegung \mathcal{P} .

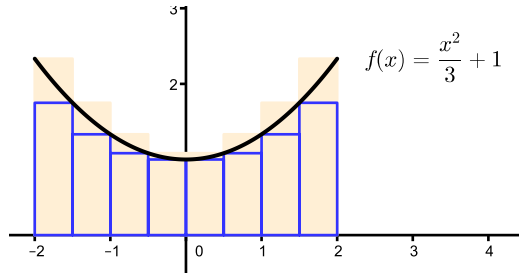
Eine Zerlegung $\tilde{\mathcal{P}}$ heißt Verfeinerung der Zerlegung \mathcal{P} (symbolisch: $\tilde{\mathcal{P}} \geq \mathcal{P}$), falls $\tilde{\mathcal{P}} = \mathcal{P}$ oder $\tilde{\mathcal{P}}$ aus \mathcal{P} durch Hinzunahme weiterer Teilungspunkte entsteht.

Definition 6.2. Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion, $\mathcal{P} = \{x_0, \dots, x_n\}$ eine Zerlegung von $[a, b]$ und $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ ein Tupel von Zahlen mit $\xi_k \in I_k$ für $k = 1, \dots, n$. Des Weiteren bezeichne

$$m_k := \inf\{f(x) \mid x \in I_k\} \quad \text{und} \quad M_k := \sup\{f(x) \mid x \in I_k\}.$$

Dann heißt

$$\begin{aligned} \underline{S}(f, \mathcal{P}) &:= \sum_{k=1}^n m_k \cdot L(I_k) && \text{Untersumme von } f \text{ bzgl. } \mathcal{P}, \\ \overline{S}(f, \mathcal{P}) &:= \sum_{k=1}^n M_k \cdot L(I_k) && \text{Obersumme von } f \text{ bzgl. } \mathcal{P}, \\ S(f, \mathcal{P}, \xi) &:= \sum_{k=1}^n f(\xi_k) \cdot L(I_k) && \text{Riemannsche Summe für } f \text{ bzgl der Zerlegung } \mathcal{P} \\ &&& \text{und der Stützstellen } \xi = (\xi_1, \dots, \xi_n). \end{aligned}$$



Satz 6.1 *Mit den obigen Bezeichnungen gilt:*

1. Sind ξ beliebige Stützstellen von \mathcal{P} , so gilt $\underline{S}(f, \mathcal{P}) \leq S(f, \mathcal{P}, \xi) \leq \overline{S}(f, \mathcal{P})$.
2. Für $\tilde{\mathcal{P}} \geq \mathcal{P}$ gilt

$$\inf_{x \in [a, b]} f(x) \cdot (b-a) \leq \underline{S}(f, \mathcal{P}) \leq \underline{S}(f, \tilde{\mathcal{P}}) \leq \overline{S}(f, \tilde{\mathcal{P}}) \leq \overline{S}(f, \mathcal{P}) \leq \sup_{x \in [a, b]} f(x) \cdot (b-a).$$

3. Für beliebige Zerlegungen \mathcal{P} und \mathcal{P}^* gilt $\underline{S}(f, \mathcal{P}) \leq \overline{S}(f, \mathcal{P}^*)$.

Beweis. Die ersten beiden Aussagen folgen unmittelbar aus der Definition. Für die 3. Aussage betrachten wir die gemeinsame Unterteilung $\tilde{\mathcal{P}} := \mathcal{P} \cup \mathcal{P}^*$, die eine Verfeinerung von \mathcal{P} sowie von \mathcal{P}^* ist, und wenden 2. darauf an. □

Definition 6.3. *Wir nennen*

$$\int_a^b f(x) dx := \sup \{ \underline{S}(f, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } [a, b] \} \quad \begin{array}{l} \text{unteres Riemann Integral} \\ \text{von } f \text{ über } [a, b], \end{array}$$

$$\int_a^b f(x) dx := \inf \{ \overline{S}(f, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } [a, b] \} \quad \begin{array}{l} \text{oberes Riemann-Integral} \\ \text{von } f \text{ über } [a, b]. \end{array}$$

Aus Satz 6.1 folgt, dass das obere und das untere Riemann-Integral existieren und dass gilt:

$$\int_a^b f(x) dx \leq \int_a^b f(x) dx.$$

Definition 6.4. *Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ heißt über $[a, b]$ Riemann-integrierbar, wenn sie beschränkt ist und das obere und das untere Riemann-Integral von f über $[a, b]$ übereinstimmen. In diesem Fall nennt man*

$$\int_a^b f(x) dx := \int_a^b f(x) dx = \int_a^b f(x) dx$$

das Riemann-Integral von f über $[a, b]$.

$\mathcal{R}([a, b], \mathbb{R})$ bezeichnet die Menge der Riemann-integrierbaren Funktionen von $[a, b]$ in \mathbb{R} .

Wir setzen außerdem:

$$\int_a^a f(x) dx := 0 \quad \text{und}$$

$$\int_b^a f(x) dx := - \int_a^b f(x) dx \quad \text{falls } f \in \mathcal{R}([a, b], \mathbb{R}).$$

Satz 6.2 (Riemannsches Integrierbarkeitskriterium)

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion. Dann gilt:

f ist genau dann über $[a, b]$ Riemann-integrierbar, wenn zu jedem $\varepsilon > 0$ eine Zerlegung $\mathcal{P}(\varepsilon)$ von $[a, b]$ existiert, so dass

$$\overline{S}(f, \mathcal{P}(\varepsilon)) - \underline{S}(f, \mathcal{P}(\varepsilon)) < \varepsilon.$$

Beweis. (\Leftarrow): Sei $n \in \mathbb{N}$ und \mathcal{P}_n eine Zerlegung von $[a, b]$ mit $\overline{S}(f, \mathcal{P}_n) - \underline{S}(f, \mathcal{P}_n) < \frac{1}{n}$. Dann gilt

$$0 \leq \int_a^b f(x) dx - \int_a^b f(x) dx \leq \overline{S}(f, \mathcal{P}_n) - \underline{S}(f, \mathcal{P}_n) \leq \frac{1}{n}.$$

Der Grenzübergang $n \rightarrow \infty$ liefert die Behauptung (Sandwich-Lemma).

(\Rightarrow): Sei $\varepsilon > 0$ gegeben. Nach Definition des Supremums und des Infimums gibt es Zerlegungen \mathcal{P} und $\tilde{\mathcal{P}}$ von $[a, b]$, so dass

$$\int_a^b f(x) dx - \underline{S}(f, \mathcal{P}) < \frac{\varepsilon}{2} \quad \text{und} \quad \overline{S}(f, \tilde{\mathcal{P}}) - \int_a^b f(x) dx < \frac{\varepsilon}{2}. \quad (6.1)$$

Für die Zerlegung $\mathcal{P}^* := \mathcal{P} \cup \tilde{\mathcal{P}}$ gilt dann wegen der Riemann-Integrierbarkeit von f

$$\begin{aligned} \overline{S}(f, \mathcal{P}^*) - \underline{S}(f, \mathcal{P}^*) &\leq \overline{S}(f, \tilde{\mathcal{P}}) - \underline{S}(f, \mathcal{P}) \\ &\leq \left(\overline{S}(f, \tilde{\mathcal{P}}) - \int_a^b f(x) dx \right) + \left(\int_a^b f(x) dx - \underline{S}(f, \mathcal{P}) \right) \\ &\stackrel{(6.1)}{<} \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

□

Satz 6.3 Jede monotone Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist Riemann-integrierbar.

Beweis. Wir beweisen die Behauptung für monoton wachsende Funktionen. Den Beweis für monoton fallende Funktionen führt man analog.

Da f monoton wachsend ist, gilt $f(a) \leq f(x) \leq f(b)$ für alle $x \in [a, b]$, d.h. f ist beschränkt. Wir prüfen nun das Riemannsches Integrierbarkeitskriterium. Sei $n \in \mathbb{N}$. Wir wählen die äquidistante Zerlegung $\mathcal{P}_n := \{x_0, \dots, x_n\}$ von $[a, b]$ mit $x_k := a + k \cdot \frac{b-a}{n}$. Dann gilt

$L(I_k) = x_k - x_{k-1} = \frac{b-a}{n} = \|\mathcal{P}_n\|$ für alle $k = 1, \dots, n$. Da f monoton wachsend ist, erhalten wir

$$\begin{aligned}\bar{S}(f, \mathcal{P}_n) - \underline{S}(f, \mathcal{P}_n) &= \sum_{k=1}^n f(x_k)L(I_k) - \sum_{k=1}^n f(x_{k-1})L(I_k) \\ &= \left(\sum_{k=1}^n f(x_k) - f(x_{k-1}) \right) \frac{b-a}{n} \\ &= (f(b) - f(a)) \cdot \frac{b-a}{n}.\end{aligned}$$

Für $\varepsilon > 0$ existiert ein $n \in \mathbb{N}$ so dass $(f(b) - f(a))\frac{b-a}{n} < \varepsilon$. Somit folgt die Behauptung aus dem Riemannschem Integrierbarkeitskriterium. \square

Satz 6.4 *Jede stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist Riemann-integrierbar.*

Beweis. Da f stetig und $[a, b]$ kompakt ist, ist f beschränkt. Für $n \in \mathbb{N}$ sei \mathcal{P}_n wieder die äquidistante Zerlegung von $[a, b]$ in n Teilintervalle. Da f stetig ist, nimmt f auf jedem Intervall $I_k = [x_{k-1}, x_k]$ ein Maximum und ein Minimum an (Satz 4.13). Seien $c_k, d_k \in I_k$ mit $f(c_k) = \min\{f(x) \mid x \in I_k\}$ und $f(d_k) = \max\{f(x) \mid x \in I_k\}$. Dann gilt

$$\bar{S}(f, \mathcal{P}_n) - \underline{S}(f, \mathcal{P}_n) = \sum_{k=1}^n (f(d_k) - f(c_k))L(I_k) = \left(\sum_{k=1}^n f(d_k) - f(c_k) \right) \frac{b-a}{n}. \quad (6.2)$$

Sei nun $\varepsilon > 0$. Die stetige Abbildung f ist auf dem kompakten Intervall $[a, b]$ gleichmäßig stetig (Satz 4.14). Folglich existiert ein $\delta > 0$, so dass für alle $v, w \in [a, b]$ mit $|v - w| < \delta$ gilt:

$$|f(v) - f(w)| < \frac{\varepsilon}{b-a}.$$

Wir wählen jetzt $n \in \mathbb{N}$ so groß, dass $\frac{b-a}{n} < \delta$. Dann gilt $|d_k - c_k| < \delta$ für $k = 1, \dots, n$ und aus (6.2) folgt

$$\bar{S}(f, \mathcal{P}_n) - \underline{S}(f, \mathcal{P}_n) < \left(\sum_{k=1}^n \frac{\varepsilon}{b-a} \right) \frac{b-a}{n} = \varepsilon.$$

Das Riemannschem Integrierbarkeitskriterium liefert dann wieder die Behauptung. \square

Der nächste Satz zeigt, dass man das Riemann-Integral für eine Riemann-integrierbare Funktion als Grenzwert einer *beliebigen Folge* von Riemannschem Summen ausrechnen kann.

Satz 6.5 *Sei $f : [a, b] \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion. Dann gilt für jede Folge (\mathcal{P}_m) von Zerlegungen von $[a, b]$ mit $\lim_{m \rightarrow \infty} \|\mathcal{P}_m\| = 0$ und für beliebige Stützstellen ξ_m von \mathcal{P}_m :*

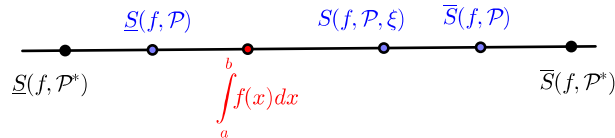
$$\int_a^b f(x) dx = \lim_{m \rightarrow \infty} S(f, \mathcal{P}_m, \xi_m).$$

Beweis. Sei $\varepsilon > 0$ gegeben und \mathcal{P}^* eine Zerlegung von $[a, b]$ mit

$$\bar{S}(f, \mathcal{P}^*) - \underline{S}(f, \mathcal{P}^*) < \frac{\varepsilon}{2},$$

die nach dem Riemannschem Integrierbarkeitskriterium (Satz 6.2) existiert.

Aus Satz 6.1 erhält man dann für alle Verfeinerungen $\mathcal{P} \geq \mathcal{P}^*$ und alle Stützstellen ξ von \mathcal{P}



$$\left| \int_a^b f(x) dx - S(f, \mathcal{P}, \xi) \right| < \frac{\varepsilon}{2}. \tag{6.3}$$

Da f beschränkt ist, existiert eine Konstante $C > 0$ so dass $|f(x)| \leq C$ für alle $x \in [a, b]$. Die Zerlegung \mathcal{P}^* habe r Teilintervalle. Da $\|\mathcal{P}_m\| \rightarrow 0$, existiert ein $m_0 \in \mathbb{N}$, so dass

$$2Cr\|\mathcal{P}_m\| < \frac{\varepsilon}{2} \quad \forall m \geq m_0.$$

Wir betrachten nun \mathcal{P}_m mit $m \geq m_0$. Seien J_1, J_2, \dots, J_s diejenigen Teilintervalle von \mathcal{P}_m , die durch Punkte von \mathcal{P}^* echt zerlegt werden. Dann ist $s \leq r$. Wir fügen nun zu den Stützstellen ξ_m von \mathcal{P}_m neue Stützstellen hinzu, so dass Stützstellen $\tilde{\xi}_m$ von $\mathcal{P}^* \cup \mathcal{P}_m$ entstehen. Damit erhalten wir mit Hilfe der Dreiecksungleichung

$$\begin{aligned} |S(f, \mathcal{P}_m, \xi_m) - S(f, \mathcal{P}_m \cup \mathcal{P}^*, \tilde{\xi}_m)| &\leq 2C \sum_{\ell=1}^s L(J_\ell) \\ &\leq 2Cs\|\mathcal{P}_m\| \leq 2Cr\|\mathcal{P}_m\| < \frac{\varepsilon}{2}. \end{aligned} \tag{6.4}$$

Da $\mathcal{P}_m \cup \mathcal{P}^* \geq \mathcal{P}^*$, folgt aus (6.3) und (6.4)

$$\begin{aligned} \left| S(f, \mathcal{P}_m, \xi_m) - \int_a^b f(x) dx \right| &\leq |S(f, \mathcal{P}_m, \xi_m) - S(f, \mathcal{P}_m \cup \mathcal{P}^*, \tilde{\xi}_m)| + \\ &\quad + \left| S(f, \mathcal{P}_m \cup \mathcal{P}^*, \tilde{\xi}_m) - \int_a^b f(x) dx \right| \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Dies gilt für alle $m \geq m_0$. Somit folgt $\lim_{m \rightarrow \infty} S(f, \mathcal{P}_m, \xi_m) = \int_a^b f(x) dx$. □

Satz 6.5 ist gut geeignet, um Rechenregeln für Riemann-Integrale zu beweisen und Abschätzungen für Summen auf Abschätzungen für Integrale zu übertragen.

Satz 6.6 1. Sei $f : [a, c] \rightarrow \mathbb{R}$ und $a < b < c$. Ist $f|_{[a,b]} \in \mathcal{R}([a, b], \mathbb{R})$ und $f|_{[b,c]} \in \mathcal{R}([b, c], \mathbb{R})$, so ist $f \in \mathcal{R}([a, c], \mathbb{R})$ und es gilt

$$\int_a^c f(x) dx = \int_a^b f(x) dx + \int_b^c f(x) dx.$$

2. Ist $f \in \mathcal{R}([a, b], \mathbb{R})$ und $[\alpha, \beta] \subset [a, b]$, so ist $f|_{[\alpha, \beta]} \in \mathcal{R}([\alpha, \beta], \mathbb{R})$.

Beweis. Übungsaufgabe. □

Satz 6.7 (Rechenregeln für das Riemann-Integral) Es seien $f, g : [a, b] \rightarrow \mathbb{R}$ zwei Riemann-integrierbare Funktionen. Dann gilt:

1. Für alle $\mu, \lambda \in \mathbb{R}$ ist die Funktion $\mu f + \lambda g$ Riemann-integrierbar und

$$\int_a^b (\mu f(x) + \lambda g(x)) dx = \mu \int_a^b f(x) dx + \lambda \int_a^b g(x) dx.$$

2. Sei $f \leq g$. Dann gilt

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

3. $|f|$ ist Riemann-integrierbar und

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

4. Für jedes $p \in [1, \infty)$ ist $|f|^p$ Riemann-integrierbar.

5. Das Produkt $f \cdot g$ ist Riemann-integrierbar.

Beweis. Den Beweis lassen wir als Übungsaufgabe. Man kann z.B. das Riemannsche Integrierbarkeitskriterium, Satz 6.5 und die entsprechenden Eigenschaften für Summen benutzen. □

Die Definition des Riemann-Integrals rechtfertigt folgende Definition:

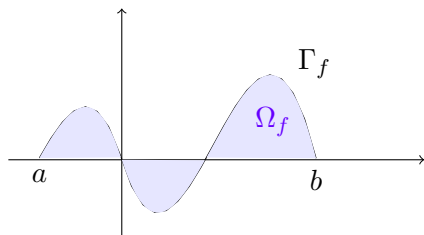
Definition 6.5. Sei $f : [a, b] \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion und $\Omega_f \subset \mathbb{R}^2$ das Gebiet zwischen dem Graphen Γ_f von f und der x -Achse, d.h.

$$\Omega_f := \{(x, y) \in \mathbb{R}^2 \mid 0 \leq y \leq f(x) \text{ falls } f(x) \geq 0, \quad f(x) \leq y \leq 0 \text{ falls } f(x) \leq 0\}.$$

Dann heißt

$$\text{Area}(\Omega_f) := \int_a^b |f(x)| dx$$

Flächeninhalt des Gebietes Ω_f .



Satz 6.8 (Mittelwertsatz der Integralrechnung)

Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und $g : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar mit $g \geq 0$. Dann existiert ein $\eta \in [a, b]$ so dass

$$\int_a^b f(x)g(x) dx = f(\eta) \cdot \int_a^b g(x) dx.$$

Beweis. Da f stetig und $[a, b]$ kompakt ist, existieren $m := \min\{f(x) \mid x \in [a, b]\}$ und $M := \max\{f(x) \mid x \in [a, b]\}$. Da $g \geq 0$ folgt

$$mg(x) \leq f(x)g(x) \leq Mg(x) \quad \forall x \in [a, b].$$

und mit den Rechenregeln aus Satz 6.7 somit

$$m \int_a^b g(x) dx \leq \int_a^b f(x)g(x) dx \leq M \int_a^b g(x) dx.$$

Sei $\int_a^b g(x) dx \neq 0$. Dann ist dieses Integral positiv und wir erhalten

$$m \leq \frac{\int_a^b f(x)g(x) dx}{\int_a^b g(x) dx} \leq M.$$

Wir setzen $\mu := \frac{\int_a^b f(x)g(x) dx}{\int_a^b g(x) dx}$. Dann gilt $\mu \in [m, M]$ und

$$\int_a^b f(x)g(x) dx = \mu \cdot \int_a^b g(x) dx.$$

Im Fall $\int_a^b g(x) dx = 0$ gilt die letzte Gleichung für ein beliebiges $\mu \in [m, M]$. Da f stetig ist, existiert nach dem Zwischenwertsatz für stetige Funktionen (Satz 4.10) ein $\eta \in [a, b]$ mit $f(\eta) = \mu$. Dies zeigt die Behauptung des Satzes. \square

Folgerung 6.1 Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig, dann existiert ein $\eta \in [a, b]$, so dass gilt

$$\int_a^b f(x) dx = f(\eta) \cdot (b - a).$$

Beweis. Dies folgt aus Satz 6.8 mit $g \equiv 1$. \square

Für stetige Funktionen $f \geq 0$ bedeutet das, dass der Flächeninhalt zwischen der x -Achse und dem Graphen von f gleich der Fläche des Rechtecks mit den Kantenlängen $[a, b]$ und $f(\eta)$ ist.

Folgerung 6.2 Sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetige Funktion mit $f \geq 0$ und $\int_a^b f(x) dx = 0$. Dann gilt $f = 0$.

Beweis. Sei $c \in (a, b)$. Wir betrachten Folgen (x_n) und (y_n) mit $a \leq x_n < c < y_n \leq b$ und $x_n \rightarrow c, y_n \rightarrow c$. Da $f \geq 0$, gilt

$$0 \leq \int_{x_n}^{y_n} f(x) dx \leq \int_a^b f(x) dx = 0.$$

Nach dem Mittelwertsatz existiert ein $\eta_n \in [x_n, y_n]$ mit

$$0 = \int_{x_n}^{y_n} f(x) dx = f(\eta_n)(y_n - x_n).$$

Folglich ist $f(\eta_n) = 0$. Nach Konstruktion konvergiert die Folge (η_n) gegen c . Da f stetig ist, konvergiert dann auch $f(\eta_n)$ gegen $f(c)$. Folglich ist $f(c) = 0$ für jedes $c \in (a, b)$. Für die Randpunkte folgt dies dann wegen der Stetigkeit von f . \square

Man kann mit Hilfe von Satz 6.5 auch Riemann-Integrale berechnen:

Einfaches Beispiel: Berechne das Riemann-Integral $\int_a^b x^2 dx$:

Die Funktion $f : [a, b] \rightarrow \mathbb{R}$, definiert durch $f(x) := x^2$, ist stetig, also Riemann-integrierbar. Sei $\mathcal{P}_n = \{x_0, \dots, x_n\}$ die äquidistante Unterteilung von $[a, b]$ mit $x_k = a + k \cdot \frac{b-a}{n}$. Wir wählen die Stützstellen $\xi = (x_1, \dots, x_n)$. Dann gilt

$$\begin{aligned} S(f, \mathcal{P}_n, \xi) &= \sum_{k=1}^n f(x_k) L(I_k) \\ &= \sum_{k=1}^n \left(a + k \cdot \frac{b-a}{n} \right)^2 \cdot \frac{b-a}{n} \\ &= \frac{b-a}{n} \cdot \sum_{k=1}^n \left(a^2 + 2ka \cdot \frac{b-a}{n} + k^2 \cdot \frac{(b-a)^2}{n^2} \right) \\ &= a^2(b-a) + 2a \frac{(b-a)^2}{n^2} \cdot \sum_{k=1}^n k + \frac{(b-a)^3}{n^3} \cdot \sum_{k=1}^n k^2 \\ &= a^2(b-a) + 2a \frac{(b-a)^2}{n^2} \cdot \frac{n(n+1)}{2} + \frac{(b-a)^3}{n^3} \cdot \frac{n(n+1)(2n+1)}{6}. \end{aligned}$$

Es folgt

$$\begin{aligned}
 \int_a^b x^2 dx &= \lim_{n \rightarrow \infty} S(f, \mathcal{P}_n, \xi) \\
 &= a^2(b-a) + a(b-a)^2 + \frac{1}{3}(b-a)^3 \\
 &= a^2b - a^3 + ab^2 + a^3 - 2a^2b + \frac{1}{3}(b^3 - 3b^2a + 3ba^2 - a^3) \\
 &= \frac{1}{3}(b^3 - a^3).
 \end{aligned}$$

Dies ist eine etwas umständliche Methode zur Berechnung des Riemann-Integrals. Im nächsten Abschnitt lernen wir einfachere Verfahren zur Berechnung des Riemann-Integrals kennen.

6.2 Integration und Differentiation

Wir wollen jetzt nachweisen, dass die Integration für gewisse Klassen von Funktionen die Umkehrung der Differentiation ist.

Im Folgenden bezeichne $I \subset \mathbb{R}$ immer ein *beliebiges* beschränktes oder unbeschränktes Intervall, das aus mehr als einem Punkt besteht.

Definition 6.6. Eine differenzierbare Funktion $F : I \rightarrow \mathbb{R}$ heißt *Stammfunktion* einer Funktion $f : I \rightarrow \mathbb{R}$, wenn $F' = f$.

Satz 6.9 Sind $F : I \rightarrow \mathbb{R}$ und $G : I \rightarrow \mathbb{R}$ zwei Stammfunktionen von $f : I \rightarrow \mathbb{R}$, dann gilt $G = F + c$, wobei c eine reelle Konstante ist.

Beweis. Wir betrachten die Funktion $G - F : I \rightarrow \mathbb{R}$. Nach Voraussetzung gilt $(G - F)' = G' - F' = f - f = 0$. Folglich ist $G - F$ eine konstante Funktion (siehe Satz 5.10), also $G = F + c$ für ein $c \in \mathbb{R}$. \square

Wenn eine Stammfunktion von f existiert, so ist sie also bis auf eine additive Konstante eindeutig bestimmt. Wir interessieren uns nun für die *Existenz* von Stammfunktionen.

Satz 6.10 (Hauptsatz der Differential- und Integralrechnung - Teil 1)

Sei $f : I \rightarrow \mathbb{R}$ eine stetige Funktion und $a \in I$. Wir definieren die Funktion $F : I \rightarrow \mathbb{R}$ durch das Riemann-Integral

$$F(x) := \int_a^x f(t) dt.$$

Dann ist F eine Stammfunktion von f .

Beweis. Wir betrachten den Differenzenquotienten von F in $x \in I$:

$$\frac{F(x+h) - F(x)}{h} = \frac{1}{h} \left(\int_a^{x+h} f(t) dt - \int_a^x f(t) dt \right) = \frac{1}{h} \int_x^{x+h} f(t) dt.$$

Nach dem Mittelwertsatz der Integralrechnung (Satz 6.8) existiert ein $\eta(h) \in [x, x+h]$ (bzw. $\eta(h) \in [x+h, x]$, falls $h < 0$) mit

$$\int_x^{x+h} f(t) dt = f(\eta(h)) \cdot h.$$

Wegen $\lim_{h \rightarrow 0} \eta(h) = x$ und der Stetigkeit von f existiert der Grenzwert

$$F'(x) = \lim_{h \rightarrow 0} \frac{F(x+h) - F(x)}{h} = \lim_{h \rightarrow 0} f(\eta(h)) = f(x). \quad \square$$

Bemerkung: *Stetige* Funktionen besitzen also immer eine Stammfunktionen, die durch ihr Riemann-Integral definiert ist. Man beachte aber:

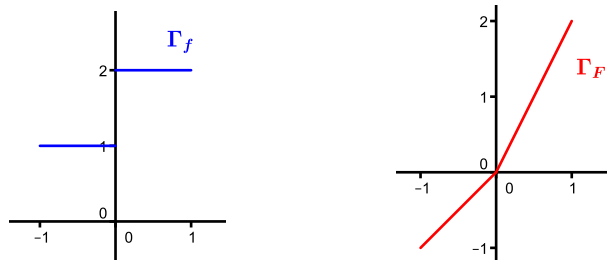
- Nicht jede Riemann-integrierbare Funktion besitzt eine Stammfunktion. Die Funktion $f : [-1, 1] \rightarrow \mathbb{R}$ mit

$$f(x) := \begin{cases} 1 & \text{falls } x \in [-1, 0], \\ 2 & \text{falls } x \in (0, 1]. \end{cases}$$

ist Riemann-integrierbar, besitzt aber keine Stammfunktion. Eine stetige Funktion $F : [-1, 1] \rightarrow \mathbb{R}$ mit $F'(x) = f(x)$ für alle $x \neq 0$ hätte nämlich die Form

$$F(x) = \begin{cases} x + c & \text{falls } x \in [-1, 0], \\ 2x + c & \text{falls } x \in (0, 1], \end{cases}$$

für eine Konstante $c \in \mathbb{R}$. Eine solche Funktion F ist aber in $x = 0$ nicht differenzierbar.



- Es gibt auch Funktionen, die Stammfunktionen besitzen, aber nicht Riemann-integrierbar sind. Wir betrachten dazu die Funktion $F : [0, 1] \subset \mathbb{R} \rightarrow \mathbb{R}$

$$F(x) := \begin{cases} x^2 \sin \frac{1}{x^2} & x \in (0, 1], \\ 0 & x = 0. \end{cases}$$

F ist auf $[0, 1]$ differenzierbar, aber $f := F'$ ist auf $[0, 1]$ nicht beschränkt (Übungsaufgabe).

Satz 6.11 (Hauptsatz der Differential- und Integralrechnung - Teil 2)

Sei $f : I \rightarrow \mathbb{R}$ stetig und $F : I \rightarrow \mathbb{R}$ eine Stammfunktion von f . Dann gilt für alle $a, b \in I$:

$$\int_a^b f(x) dx = F(b) - F(a).$$

Beweis. Wir betrachten die Funktion $F_0 : I \rightarrow \mathbb{R}$ mit $F_0(t) := \int_a^t f(x) dx$. F_0 ist eine Stammfunktion von f mit $F_0(a) = 0$ und $F_0(b) = \int_a^b f(x) dx$. Für die Stammfunktion F gilt dann $F = F_0 + c$ für eine Konstante $c \in \mathbb{R}$ und folglich

$$F(b) - F(a) = F_0(b) - F_0(a) = F_0(b) = \int_a^b f(x) dx. \quad \square$$

Bezeichnung: Man setzt: $F(x) \Big|_a^b := F(b) - F(a)$. Die Formel aus Satz 6.11 schreibt sich dann in der Form: $\int_a^b f(x) dx = F(x) \Big|_a^b$. Deshalb wird häufig die folgende Symbolik für die Stammfunktion F von f benutzt¹:

$$\int f(x) dx = F(x) \quad \text{auf } I.$$

Man nennt

$$\int f(x) dx \quad \text{unbestimmtes Integral von } f,$$

$$\int_a^b f(x) dx \quad \text{bestimmtes Integral von } f.$$

f heißt *Integrand*, x *Integrationsvariable*, $[a, b]$ das *Integrationsintervall*, a und b die *untere* bzw. *obere Integrationsgrenze*.

In der folgenden Liste stellen wir einige wichtige Grundintegrale zusammen. Der Beweis erfolgt durch Ableiten der Stammfunktion. Die Beziehungen zwischen den Umkehrfunktionen der Hyperbelfunktionen und der Logarithmusfunktion folgen aus Satz 4.19.

¹ Diese Bezeichnung ist etwas problematisch, da die Stammfunktion nur bis auf eine Konstante eindeutig bestimmt ist. Man merke sich deshalb, was mit dieser Symbolik gemeint ist.

Wichtige Grundintegrale:

$$(1) \int x^\alpha dx = \frac{1}{\alpha+1} x^{\alpha+1} \quad \text{für } \alpha \in \mathbb{R} \text{ mit } \alpha \neq -1.$$

Dies gilt für nicht ganzzahlige $\alpha \in \mathbb{R}$ auf $I = (0, \infty)$, für $\alpha \in \mathbb{N}_0$ auf \mathbb{R} und für $\alpha \in \mathbb{Z}$ mit $\alpha \leq -1$ auf $I = (-\infty, 0) \cup (0, \infty)$.

$$(2) \int \frac{1}{x} dx = \ln|x| \quad \text{auf } (0, \infty) \text{ und } (-\infty, 0).$$

$$(3) \int e^{ax} dx = \frac{1}{a} \cdot e^{ax} \quad \text{für } a \in \mathbb{R} \text{ mit } a \neq 0, \quad \text{auf } \mathbb{R}.$$

$$(4) \int \cos x dx = \sin x, \quad \int \sin x dx = -\cos x \quad \text{auf } \mathbb{R}.$$

$$(5) \int \cosh x dx = \sinh x, \quad \int \sinh x dx = \cosh x \quad \text{auf } \mathbb{R}.$$

$$(6) \int \frac{1}{1+x^2} dx = \arctan x \quad \text{auf } \mathbb{R}.$$

$$(7) \int \frac{1}{1-x^2} dx = \frac{1}{2} \ln \left| \frac{1+x}{1-x} \right| = \begin{cases} \operatorname{artanh}(x) & \text{auf } (-1, 1) \\ \operatorname{arcoth}(x) & \text{auf } (-\infty, -1) \cup (1, \infty) \end{cases}.$$

$$(8) \int \frac{1}{\sqrt{1+x^2}} dx = \ln(x + \sqrt{x^2+1}) = \operatorname{arsinh}(x) \quad \text{auf } \mathbb{R}.$$

$$(9) \int \frac{1}{\sqrt{1-x^2}} dx = \arcsin(x) \quad \text{auf } (-1, 1).$$

$$(10) \int \frac{1}{\sqrt{x^2-1}} dx = \ln|x + \sqrt{x^2-1}| = \begin{cases} \operatorname{arcosh}(x) & \text{auf } (1, \infty) \\ -\operatorname{arcosh}(-x) & \text{auf } (-\infty, -1) \end{cases}.$$

Wir beweisen jetzt die beiden wichtigsten Rechenregeln für Integrale, die *partielle Integration* und die *Substitutionsregel*.

Satz 6.12 (Partielle Integration)

Seien $g, f : I \rightarrow \mathbb{R}$ stetig differenzierbare Funktionen. Dann gilt für alle $a, b \in I$:

$$\int_a^b f'(x)g(x) dx = - \int_a^b f(x)g'(x) dx + f(x)g(x) \Big|_a^b,$$

bzw. für die unbestimmten Integrale

$$\int f'(x)g(x) dx = - \int f(x)g'(x) dx + f(x)g(x) \quad \text{auf } I.$$

Beweis. Für $F := f \cdot g$ gilt nach der Produktregel für Ableitungen $F' = f' \cdot g + f \cdot g'$. Die Linearität des Integrals und der Hauptsatz der Differential- und Integralrechnung liefert dann für alle $a, b \in I$

$$\begin{aligned} \int_a^b f'(x)g(x) dx + \int_a^b f(x)g'(x) dx &= \int_a^b (f'(x)g(x) + f(x)g'(x)) dx \\ &= \int_a^b (f \cdot g)'(x) dx = f(x) \cdot g(x) \Big|_a^b. \end{aligned} \quad \square$$

Satz 6.13 (Substitutionsregel)

Sei $\phi : I \rightarrow \mathbb{R}$ stetig differenzierbar und $f : J := \phi(I) \rightarrow \mathbb{R}$ stetig. Dann gilt für alle $c, d \in I$

$$\int_c^d f(\phi(t))\phi'(t) dt = \int_{\phi(c)}^{\phi(d)} f(x) dx.$$

Ist ϕ zusätzlich streng monoton, so existiert die Umkehrfunktion $\phi^{-1} : J \rightarrow I$ und es gilt für alle $a, b \in J$

$$\int_a^b f(x) dx = \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f(\phi(t))\phi'(t) dt.$$

(Formale Regel: Man substituiert $x = \phi(t)$, setzt formal $dx = \phi'(t)dt$ und ändert die Integrationsgrenzen entsprechend).

Beweis. Wir bemerken zunächst, dass wegen der Stetigkeit von ϕ mit I auch $J = \phi(I) \subset \mathbb{R}$ ein Intervall ist (Zwischenwertsatz). Sei nun $F : J \rightarrow \mathbb{R}$ eine Stammfunktion von f . Dann gilt für $F \circ \phi : I \rightarrow \mathbb{R}$ nach der Kettenregel für Ableitungen

$$(F \circ \phi)'(t) = F'(\phi(t)) \cdot \phi'(t) = f(\phi(t)) \cdot \phi'(t) \quad \forall t \in I.$$

Nach Satz 6.11 folgt für $c, d \in I$:

$$\int_c^d f(\phi(t))\phi'(t) dt = (F \circ \phi)(d) - (F \circ \phi)(c) = F(\phi(d)) - F(\phi(c)) = \int_{\phi(c)}^{\phi(d)} f(x) dx.$$

Wenn ϕ streng monoton ist, so ist ϕ injektiv und $\phi : I \rightarrow J$ somit bijektiv. Folglich existiert die Umkehrfunktion $\phi^{-1} : J \rightarrow I$. Für gegebene $a, b \in J$ kann man also eindeutige bestimmte $c, d \in I$ finden mit $\phi(c) = a$ und $\phi(d) = b$. Die 2. Behauptung folgt dann aus der ersten. \square

Bemerkung: In der unbestimmten Form schreibt sich die Substitutionsregel als

$$\int f(\phi(t))\phi'(t) dt = \int f(x) dx \Big|_{x=\phi(t)} \quad \text{auf } I$$

bzw. im Fall, dass die Umkehrfunktion ϕ^{-1} existiert als

$$\int f(x) dx = \int f(\phi(t))\phi'(t) dt \Big|_{t=\phi^{-1}(x)} \quad \text{auf } J.$$

Beispiel 1: Sei $f : I \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion ohne Nullstellen auf I . Dann gilt

$$\int \frac{f'(x)}{f(x)} dx = \ln |f(x)| \quad \text{auf } I.$$

Dies sieht man durch Ableiten der Funktion auf der rechten Seite oder durch die Substitution $y = f(x)$, $dy = f'(x) dx$:

$$\int \frac{f'(x)}{f(x)} dx = \int \frac{dy}{y} \Big|_{y=f(x)} = \ln |y| \Big|_{y=f(x)} = \ln |f(x)| \quad \text{auf } I.$$

So ist zum Beispiel

$$\int \tan x dx = \int \frac{\sin x}{\cos x} dx = - \int \frac{\cos' x}{\cos x} dx = - \ln |\cos x| \quad \text{auf } \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Beispiel 2: Berechnung von $I_n(x) := \int \sin^n(x)$, $n \in \mathbb{N}_0$.

Für $n = 0, 1$ gilt:

$$I_0(x) = x \quad \text{und} \quad I_1(x) = -\cos x.$$

Für $n \geq 2$ leiten wir mittels partieller Integration eine Rekursionsformel her:

$$\begin{aligned} I_n(x) &= - \int \sin^{n-1}(x) \cdot \cos'(x) dx \\ &\stackrel{\text{part. Int.}}{=} - \sin^{n-1}(x) \cdot \cos(x) + \int \frac{d}{dx} \sin^{n-1}(x) \cdot \cos(x) dx \\ &= - \sin^{n-1}(x) \cdot \cos(x) + (n-1) \int \sin^{n-2}(x) (1 - \sin^2(x)) dx \\ &= - \sin^{n-1}(x) \cdot \cos(x) + (n-1)I_{n-2}(x) - (n-1)I_n(x). \end{aligned}$$

Daraus erhält man die Rekursionsformel:

$$I_n(x) = -\frac{1}{n} \cdot \sin^{n-1}(x) \cdot \cos(x) + \frac{n-1}{n} I_{n-2}(x) \quad \text{für } n \geq 2.$$

Für $J_n(x) := \int \cos^n(x) dx$ berechnet man analog

$$J_n(x) = \frac{1}{n} \cos^{n-1}(x) \cdot \sin(x) + \frac{n-1}{n} J_{n-2}(x) \quad \text{für } n \geq 2$$

mit $J_0(x) = x$ und $J_1(x) = \sin(x)$.

Beispiel 3: Integrale rationaler Funktionen $\int \frac{a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0}{b_n x^n + b_{n-1} x^{n-1} + \dots + b_1 x + b_0} dx$:

Solche Integrale werden mit Hilfe der *Partialbruchzerlegung*² berechnet. Dabei stellt man die rationale Funktion $R(x) := \frac{a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0}{b_n x^n + b_{n-1} x^{n-1} + \dots + b_1 x + b_0}$ als Summe *einfacher* Funktionen, die sogenannten *Partialbrüche*, dar. Im folgenden beschreiben wir dieses Verfahren zunächst für komplexe rationale Funktionen und spezialisieren es anschließend auf den reellen Fall.

1. Komplexe Partialbruchzerlegung

Sei $Q \in \mathbb{C}[z]$ ein komplexes Polynom vom Grad $n \geq 1$, d.h.

$$Q(z) := b_n z^n + b_{n-1} z^{n-1} + \dots + b_1 z + b_0$$

mit $b_0, b_1, \dots, b_n \in \mathbb{C}$ und $b_n \neq 0$. Nach dem Fundamentalsatz der Algebra hat Q n komplexe Nullstellen und läßt sich in komplexe Linearfaktoren zerlegen (siehe Satz 4.15 und Satz 4.16):

Sind ξ_1, \dots, ξ_k die verschiedenen komplexen Nullstellen von Q und ν_1, \dots, ν_k ihre jeweilige Vielfachheit, dann gilt

$$Q(z) = b_n (z - \xi_1)^{\nu_1} \cdot (z - \xi_2)^{\nu_2} \cdot \dots \cdot (z - \xi_k)^{\nu_k}.$$

Satz 6.14 (Partialbruchzerlegung komplexer rationaler Funktionen)

Seien $P, Q \in \mathbb{C}[z]$ komplexe Polynome mit $\deg(P) < \deg(Q)$. Q habe die verschiedenen Nullstellen ξ_1, \dots, ξ_k mit der jeweiligen Vielfachheit ν_1, \dots, ν_k . Dann existieren eindeutig bestimmte Konstanten $A_{j1}, A_{j2}, \dots, A_{j\nu_j} \in \mathbb{C}$ für $j \in \{1, \dots, k\}$, so dass

$$\frac{P(z)}{Q(z)} = \sum_{j=1}^k \left(\frac{A_{j1}}{(z - \xi_j)} + \frac{A_{j2}}{(z - \xi_j)^2} + \dots + \frac{A_{j\nu_j}}{(z - \xi_j)^{\nu_j}} \right) \quad \forall z \notin \{\xi_1, \dots, \xi_k\}. \quad (6.5)$$

Beweis. Wir führen den Beweis durch Induktion über $\deg(Q)$.

Ind.-Anfang: Sei $\deg(Q) = 1$. Da nach Voraussetzung $\deg(P) < \deg(Q)$, hat P den Grad Null. Es gilt also $P(z) = a_0$ und $Q(z) = b_1 z + b_0 = b_1(z - \xi_1)$ mit $\xi_1 = -\frac{b_0}{b_1}$. Daraus folgt

$$\frac{P(z)}{Q(z)} = \frac{a_0}{b_1(z - \xi_1)} = \frac{A_{11}}{z - \xi_1}$$

für die eindeutig bestimmte Zahl $A_{11} := \frac{a_0}{b_1}$.

Ind.-Schritt: Wir setzen voraus, dass die Behauptung für alle rationalen Funktionen $\tilde{R} = \frac{\tilde{P}}{\tilde{Q}}$ mit $\deg(\tilde{Q}) \leq n - 1$ gilt und zeigen sie dann für rationale Funktionen $R = \frac{P}{Q}$ mit $\deg(Q) = n$:

Sei ξ_1 eine Nullstelle von Q mit Vielfachheit ν_1 . Dann kann man die Linearfaktoren $(z - \xi_1)^{\nu_1}$ abspalten und erhält $Q(z) = (z - \xi_1)^{\nu_1} \cdot S(z)$, wobei S ein komplexes Polynom vom Grad $n - \nu_1$ mit $S(\xi_1) \neq 0$ ist. Sei $a := \frac{P(\xi_1)}{S(\xi_1)}$. Dann gilt

² Den Beweis der Sätze 6.14 und 6.15 haben wir in der Vorlesung nicht vorgeführt.

$$\frac{P(z)}{Q(z)} - \frac{a}{(z - \xi_1)^{\nu_1}} = \frac{P(z) - a \cdot S(z)}{(z - \xi_1)^{\nu_1} \cdot S(z)}. \quad (*)$$

Nach Definition von a ist gilt $P(\xi_1) - aS(\xi_1) = 0$, d.h. ξ_1 ist eine Nullstelle des Polynoms $P - aS$. Ist $P - aS = 0$, so folgt aus $(*)$ sofort die Induktionsbehauptung. Ist $P - aS \neq 0$, so kann man den Linearfaktor $(z - \xi_1)$ abspalten, d.h. es gilt $P(z) - aS(z) = (z - \xi_1)\tilde{S}(z)$ für ein Polynom \tilde{S} . Es folgt

$$\frac{P(z)}{Q(z)} \stackrel{(*)}{=} \frac{a}{(z - \xi_1)^{\nu_1}} + \frac{\tilde{S}(z)}{(z - \xi_1)^{\nu_1-1} \cdot S(z)}.$$

Das Nennerpolynom im zweiten Summanden hat den Grad $n - 1$ und als seine Nullstellen kommen alle Nullstellen von Q vor, wobei ξ_1 mit Vielfachheit $\nu_1 - 1$ auftritt. Das Zählerpolynom \tilde{S} im zweiten Summanden hat den Grad $< n - 1$. Nach Ind.-Voraussetzung kann man die rationale Funktion im zweiten Summanden entsprechend der Behauptung des Satzes zerlegen mit eindeutig bestimmten Koeffizienten $A_{j\ell}$. $A_{1\nu_1} := a$ ist ebenfalls eindeutig bestimmt. \square

2. Reelle Partialbruchzerlegung

Sei $Q \in \mathbb{R}[x]$ nun ein reelles Polynom vom Grad $n \geq 1$, d.h.

$$Q(x) = b_n x^n + b_{n-1} x^{n-1} + \dots + b_1 x^1 + b_0$$

mit $b_0, b_1, \dots, b_n \in \mathbb{R}$ und $b_n \neq 0$. Dann tritt mit jeder echt komplexen Nullstelle ξ von Q auch die konjugiert-komplexe Zahl $\bar{\xi}$ als Nullstelle mit der gleichen Vielfachheit auf. Aus

$$0 = Q(\xi) = b_n \xi^n + b_{n-1} \xi^{n-1} + \dots + b_1 \xi^1 + b_0$$

folgt nämlich durch Konjugieren

$$0 = \overline{Q(\xi)} = \overline{b_n \xi^n + b_{n-1} \xi^{n-1} + \dots + b_1 \xi^1 + b_0} = b_n \bar{\xi}^n + b_{n-1} \bar{\xi}^{n-1} + \dots + b_1 \bar{\xi}^1 + b_0 = Q(\bar{\xi}).$$

Das quadratische Polynom $(x - \xi)(x - \bar{\xi})$ ist dann reell und reell-irreduzibel, da

$$\begin{aligned} (x - \xi)(x - \bar{\xi}) &= x^2 - (\xi + \bar{\xi})x + \xi\bar{\xi} = x^2 - 2 \operatorname{Re}(\xi)x + |\xi|^2 \\ &=: x^2 + px + q, \end{aligned}$$

wobei für die Koeffizienten p und q gilt $\frac{p^2}{4} - q = \operatorname{Re}(\xi)^2 - |\xi|^2 = -\operatorname{Im}(\xi)^2 < 0$.

Sind $\lambda_1, \dots, \lambda_s$ die reellen Nullstellen von Q mit der jeweiligen Vielfachheit μ_1, \dots, μ_s und $\xi_1, \bar{\xi}_1, \dots, \xi_r, \bar{\xi}_r$ die echt komplexen Nullstellen von Q mit der Vielfachheit $\nu_1, \nu_1, \dots, \nu_r, \nu_r$, so hat die (komplexe) Linearfaktorzerlegung von $Q(x)$ die Form

$$Q(x) = b_n (x - \lambda_1)^{\mu_1} \cdot \dots \cdot (x - \lambda_s)^{\mu_s} \cdot (x - \xi_1)^{\nu_1} (x - \bar{\xi}_1)^{\nu_1} \cdot \dots \cdot (x - \xi_r)^{\nu_r} (x - \bar{\xi}_r)^{\nu_r}.$$

Wir fassen die Produkte $(x - \xi_j)(x - \bar{\xi}_j)$ zu einem reellen Polynom zusammen und erhalten die *reelle Faktorzerlegung* von $Q(x)$:

$$Q(x) = b_n (x - \lambda_1)^{\mu_1} \cdot \dots \cdot (x - \lambda_s)^{\mu_s} \cdot \underbrace{(x^2 + p_1 x + q_1)^{\nu_1}}_{\text{reell-irreduzibel}} \cdot \dots \cdot \underbrace{(x^2 + p_r x + q_r)^{\nu_r}}_{\text{reell-irreduzibel}}.$$

Satz 6.15 (Partialbruchzerlegung reeller rationaler Funktionen)

Seien $P, Q \in \mathbb{R}[x]$ zwei reelle Polynome mit $\deg(P) < \deg(Q)$ und sei

$$Q(x) = b_n(x - \lambda_1)^{\mu_1} \cdot \dots \cdot (x - \lambda_s)^{\mu_s} \cdot \underbrace{(x^2 + p_1x + q_1)^{\nu_1}}_{\text{reell-irreduzibel}} \cdot \dots \cdot \underbrace{(x^2 + p_rx + q_r)^{\nu_r}}_{\text{reell-irreduzibel}}$$

die reelle Faktorzerlegung von $Q(x)$. Dann existieren eindeutig bestimmte reelle Zahlen $A_{j1}, A_{j2}, \dots, A_{j\mu_j}$ für $j = 1, \dots, s$, und $B_{\ell 1}, B_{\ell 2}, \dots, B_{\ell\nu_\ell}, C_{\ell 1}, C_{\ell 2}, \dots, C_{\ell\nu_\ell}$ für $\ell = 1, \dots, r$, so dass für alle $x \in \mathbb{R}$ mit $x \notin \{\lambda_1, \dots, \lambda_s\}$ gilt:

$$\begin{aligned} \frac{P(x)}{Q(x)} &= \sum_{j=1}^s \left(\frac{A_{j1}}{(x - \lambda_j)} + \frac{A_{j2}}{(x - \lambda_j)^2} + \dots + \frac{A_{j\mu_j}}{(x - \lambda_j)^{\mu_j}} \right) \\ &+ \sum_{\ell=1}^r \left(\frac{B_{\ell 1}x + C_{\ell 1}}{(x^2 + p_\ell x + q_\ell)} + \frac{B_{\ell 2}x + C_{\ell 2}}{(x^2 + p_\ell x + q_\ell)^2} + \dots + \frac{B_{\ell\nu_\ell}x + C_{\ell\nu_\ell}}{(x^2 + p_\ell x + q_\ell)^{\nu_\ell}} \right). \end{aligned} \quad (6.6)$$

Beweis. Wir machen zunächst die komplexe Partialbruchzerlegung wie in Satz 6.14 und stellen $\frac{P(x)}{Q(x)}$ als Summe komplexer Partialbrüche nach Formel (6.5) dar. Es existieren also eindeutig bestimmte komplexe Zahlen $A_{j1}, \dots, A_{j\mu_j}$ und $a_{\ell 1}, \dots, a_{\ell\nu_\ell}, b_{\ell 1}, \dots, b_{\ell\nu_\ell}$, so dass für alle $x \in \mathbb{R}$ mit $x \notin \{\lambda_1, \dots, \lambda_s\}$ gilt

$$\begin{aligned} \frac{P(x)}{Q(x)} &= \sum_{j=1}^s \left(\frac{A_{j1}}{(x - \lambda_j)} + \frac{A_{j2}}{(x - \lambda_j)^2} + \dots + \frac{A_{j\mu_j}}{(x - \lambda_j)^{\mu_j}} \right) \\ &+ \sum_{\ell=1}^r \left(\frac{a_{\ell 1}}{(x - \xi_\ell)} + \frac{a_{\ell 2}}{(x - \xi_\ell)^2} + \dots + \frac{a_{\ell\nu_\ell}}{(x - \xi_\ell)^{\nu_\ell}} \right) \\ &+ \sum_{\ell=1}^r \left(\frac{b_{\ell 1}}{(x - \bar{\xi}_\ell)} + \frac{b_{\ell 2}}{(x - \bar{\xi}_\ell)^2} + \dots + \frac{b_{\ell\nu_\ell}}{(x - \bar{\xi}_\ell)^{\nu_\ell}} \right). \end{aligned} \quad (6.7)$$

Da $P(x)$ und $Q(x)$ reelle Polynome sind, gilt $\overline{\frac{P(x)}{Q(x)}} = \frac{P(x)}{Q(x)}$ für alle $x \in \mathbb{R}$ mit $Q(x) \neq 0$. Dies liefert

$$0 = \frac{P(x)}{Q(x)} - \overline{\frac{P(x)}{Q(x)}} = \sum_{j=1}^s \sum_{k=1}^{\mu_j} \frac{A_{jk} - \overline{A_{jk}}}{(x - \lambda_j)^k} + \sum_{\ell=1}^r \sum_{k=1}^{\nu_\ell} \left(\frac{a_{\ell k} - \overline{b_{\ell k}}}{(x - \xi_\ell)^k} + \frac{b_{\ell k} - \overline{a_{\ell k}}}{(x - \bar{\xi}_\ell)^k} \right)$$

für alle $x \in \mathbb{R}$ mit $Q(x) \neq 0$. Wegen der Eindeutigkeit der Konstanten in der komplexen Partialbruchzerlegung (der linken Seite $0 = \frac{0}{Q(x)}$) folgt $A_{jk} = \overline{A_{jk}}$, also $A_{jk} \in \mathbb{R}$, und $b_{\ell k} = \overline{a_{\ell k}}$.

Wir fassen in (6.7) nun noch die komplexen Partialbrüche vom Typ $\frac{a}{(x - \xi)^k} + \frac{\bar{a}}{(x - \bar{\xi})^k}$ zu reellen Partialbrüchen zusammen:

$$\frac{a}{(x - \xi)^k} + \frac{\bar{a}}{(x - \bar{\xi})^k} = \frac{a(x - \bar{\xi})^k + \bar{a}(x - \xi)^k}{(x^2 + px + q)^k} =: \frac{S(x)}{(x^2 + px + q)^k},$$

wobei $S(x)$ ein reelles Polynom von Grad $\leq k$ ist. Ist $k = 1$, so hat dieser Bruch bereits die gewünschte Form. Ist $k > 1$, so teilen wir $S(x)$ durch $(x^2 + px + q)$ mit Rest:

$$S(x) = S_1(x)(x^2 + px + q) + (D_1x + E_1)$$

und wiederholen dies ggf., d.h.,

$$\begin{aligned} \frac{a}{(x-\xi)^k} + \frac{\bar{a}}{(x-\bar{\xi})^k} &= \frac{S_1(x)}{(x^2+px+q)^{k-1}} + \frac{D_1x+E_1}{(x^2+px+q)^k} \\ &= \frac{S_2(x)}{(x^2+px+q)^{k-2}} + \frac{D_2x+E_2}{(x^2+px+q)^{k-1}} + \frac{D_1x+E_1}{(x^2+px+q)^k} \\ &\vdots \\ &= \frac{D_mx+E_m}{(x^2+px+q)} + \dots + \frac{D_2x+E_2}{(x^2+px+q)^{k-1}} + \frac{D_1x+E_1}{(x^2+px+q)^k}. \end{aligned}$$

Damit erhalten wir die Behauptung des Satzes. \square

Methode zur Berechnung der Koeffizienten in der Partialbruchzerlegung:

Sei $R(x) := \frac{P(x)}{Q(x)}$ eine reelle rationale Funktion mit der Partialbruchzerlegung (6.6) aus Satz 6.15. Wir multiplizieren Gleichung (6.6) mit dem Nennerpolynom $Q(x)$ und erhalten daraus eine Polynom-Gleichung:

$$P(x) = \text{Polynom} \dots$$

Das Polynom auf der rechten Seite enthält die gesuchten reellen Zahlen A_{jk} , B_{ℓ} und C_{ℓ} in seinen Koeffizienten. Der Koeffizientenvergleich für beide Polynome liefert ein Gleichungssystem für diese Zahlen, das man eindeutig lösen kann.

Beispiel: Partialbruchzerlegung von $R(x) := \frac{3x^3-9x^2+14x+10}{x^4-6x^3+10x^2}$

Es gilt $Q(x) := x^4 - 6x^3 + 10x^2 = x^2(x^2 - 6x + 10)$. Q hat $\lambda_1 = 0$ als doppelte reelle Nullstelle. Der letzte Faktor ist ein reell-irreduzibles quadratisches Polynom. Der Ansatz für die Partialbruchzerlegung von $R(x)$ lautet somit:

$$R(x) = \frac{A}{x} + \frac{B}{x^2} + \frac{Cx+D}{x^2-6x+10}$$

für eindeutig bestimmte reelle Zahlen A, B, C, D . Multiplizieren wir diese Gleichung mit $Q(x)$, so erhalten wir die Polynom-Gleichung

$$\begin{aligned} 3x^3 - 9x^2 + 14x + 10 &= Ax(x^2 - 6x + 10) + B(x^2 - 6x + 10) + Cx^3 + Dx^2 \\ &= (A+C)x^3 + (-6A+B+D)x^2 + (10A-6B)x + 10B. \end{aligned}$$

Der Koeffizientenvergleich liefert das Gleichungssystem

$$3 = A + C, \quad -9 = -6A + B + D, \quad 14 = 10A - 6B, \quad 10 = 10B$$

mit der Lösung $A = 2$, $B = 1$, $C = 1$, $D = 2$. Die Partialbruchzerlegung lautet somit:

$$R(x) = \frac{2}{x} + \frac{1}{x^2} + \frac{x+2}{x^2-6x+10}.$$

Anwendung auf die Berechnung der Integrale rationaler Funktionen

Seien $P, Q \in \mathbb{R}[x]$ reelle Polynome und $I \subset \mathbb{R}$ ein Intervall, das keine Nullstellen von Q enthält. Wir wollen das Integral

$$\int \frac{P(x)}{Q(x)} dx \quad \text{auf } I$$

berechnen. Dazu gehen wir folgendermaßen vor:

a) Ist $\deg(P) \geq \deg Q$, so dividieren wir P durch Q und erhalten

$$\frac{P(x)}{Q(x)} = P_1(x) + \frac{P_2(x)}{Q(x)},$$

wobei $P_1, P_2 \in \mathbb{R}[x]$ reelle Polynome sind und $\deg(P_2) < \deg(Q)$ gilt.

b) Wir zerlegen $\frac{P_2(x)}{Q(x)}$ in Partialbrüche nach Satz 6.15.

c) Es bleibt nun, die Integrale der auftretenden Partialbrüche auf I zu berechnen.

Wir geben die relevanten Stammfunktionen an, die man aus den Grundintegralen nach geeigneter Substitution erhält. In den folgenden Formeln sei $k \in \mathbb{N}$ mit $k > 1$, $\lambda, p, q \in \mathbb{R}$ und $q - \frac{p^2}{4} > 0$.

$$\int \frac{1}{(x - \lambda)} dx = \ln|x - \lambda| \quad (\text{Subst. } y = x - \lambda).$$

$$\int \frac{1}{(x - \lambda)^k} dx = -\frac{1}{k-1} \cdot \frac{1}{(x - \lambda)^{k-1}} \quad (\text{Subst. } y = x - \lambda).$$

$$\int \frac{1}{x^2 + px + q} dx = \frac{1}{\sqrt{q - p^2/4}} \cdot \arctan \frac{x + p/2}{\sqrt{q - p^2/4}} \quad (\text{Subst. } y = \frac{x + p/2}{\sqrt{q - p^2/4}}).$$

$$\int \frac{2x + p}{(x^2 + px + q)} dx = \ln|x^2 + px + q| \quad (\text{nach Beispiel 1}).$$

$$\int \frac{2x + p}{(x^2 + px + q)^k} dx = -\frac{1}{k-1} \cdot \frac{1}{(x^2 + px + q)^{k-1}} \quad (\text{Subst. } y = x^2 + px + q).$$

Das noch fehlende Integral

$$\int \frac{1}{(x^2 + px + q)^k} dx$$

führt man mit Hilfe der Rekursionsformel

$$\int \frac{1}{(y^2 + 1)^k} dy = \frac{1}{(2k-2)} \cdot \frac{y}{(y^2 + 1)^{k-1}} + \frac{2k-3}{2k-2} \cdot \int \frac{1}{(y^2 + 1)^{k-1}} dy$$

schrittweise auf den arctan zurück.

Für das obige Beispiel erhalten wir mit diesen Formeln:

$$\begin{aligned} \int \frac{3x^3 - 9x^2 + 14x + 10}{x^4 - 6x^3 + 10x^2} dx &= \int \frac{2}{x} dx + \int \frac{1}{x^2} dx + \int \frac{x+2}{x^2 - 6x + 10} dx \\ &= 2 \ln|x| - \frac{1}{x} + \frac{1}{2} \int \frac{2x-6}{x^2 - 6x + 10} dx + 5 \int \frac{1}{x^2 - 6x + 10} dx \\ &= 2 \ln|x| - \frac{1}{x} + \frac{1}{2} \ln|x^2 - 6x + 10| + 5 \arctan(x-3). \end{aligned}$$

Wir betrachten nun noch zwei weitere Anwendungen des Hauptsatzes der Differential- und Integralrechnung, die man zur Approximation von Funktionen bzw. von Integralen benutzen kann. Zunächst erhalten wir eine weitere Formel für das Restglied in der Taylorentwicklung einer $(n + 1)$ -mal stetig differenzierbaren Funktion, die wir in Abschnitt 5.4. bereits angekündigt hatten.

Satz 6.16 (Integralform des Restgliedes in der Taylorentwicklung)

Sei $f : I \rightarrow \mathbb{R}$ eine $(n + 1)$ -mal stetig differenzierbare Funktion, $x_0 \in I$ und

$$f(x) = \underbrace{f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2!}(x - x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n}_{=: T_n(f, x_0)(x) \text{ } n\text{-tes Taylorpolynom von } f \text{ in } x_0} + R_n(f, x_0)(x)$$

die Taylorformel n -ter Ordnung von f in x_0 auf I . Dann gilt für das Restglied $R_n(f, x_0)(x)$

$$R_n(f, x_0)(x) = \frac{1}{n!} \int_{x_0}^x (x - t)^n f^{(n+1)}(t) dt.$$

Beweis. Wir zeigen die Behauptung durch vollständige Induktion über $n \in \mathbb{N}_0$.

Induktionsanfang: Für $n = 0$ ist die zu beweisende Formel

$$R_0(f, x_0)(x) = f(x) - f(x_0) = \int_{x_0}^x f'(t) dt$$

gerade der Hauptsatz der Differential- und Integralrechnung.

Induktionsschritt: Wir setzen voraus, dass die Behauptung für ein beliebiges aber festes $n \in \mathbb{N}_0$ gilt und beweisen sie für $(n + 1)$:

$$\begin{aligned} R_{n+1}(f, x_0)(x) &= f(x) - T_{n+1}(f, x_0)(x) \\ &= f(x) - T_n(f, x_0)(x) - \frac{f^{(n+1)}(x_0)}{(n+1)!}(x - x_0)^{n+1} \\ &= R_n(f, x_0)(x) - \frac{f^{(n+1)}(x_0)}{(n+1)!}(x - x_0)^{n+1}. \end{aligned} \quad (6.8)$$

Für das n -te Restglied $R_n(f, x_0)(x)$ erhalten wir aus der Induktionsvoraussetzung

$$\begin{aligned} R_n(f, x_0)(x) &= \frac{1}{n!} \int_{x_0}^x (x - t)^n f^{(n+1)}(t) dt \\ &= - \int_{x_0}^x \frac{d}{dt} \left(\frac{(x - t)^{n+1}}{(n+1)!} \right) f^{(n+1)}(t) dt \\ &\stackrel{\text{part. Int.}}{=} - \frac{(x - t)^{n+1}}{(n+1)!} f^{(n+1)}(t) \Big|_{x_0}^x + \frac{1}{(n+1)!} \int_{x_0}^x (x - t)^{n+1} f^{(n+2)}(t) dt \\ &= \frac{1}{(n+1)!} \int_{x_0}^x (x - t)^{n+1} f^{(n+2)}(t) dt + \frac{f^{(n+1)}(x_0)}{(n+1)!} (x - x_0)^{n+1}. \end{aligned}$$

Mit (6.8) folgt dann die Induktionsbehauptung. □

In den meisten Fällen lässt sich das Integral einer Funktion nicht durch elementare Funktionen ausdrücken. Das tritt bereits auf, wenn man z.B. die Länge einer Ellipse berechnen will. In diesen Fällen hilft Satz 6.5, das Integral zu approximieren. Der folgende Satz erlaubt es, den Fehler bei Approximation des Riemann-Integrals durch Trapezsummen abzuschätzen.

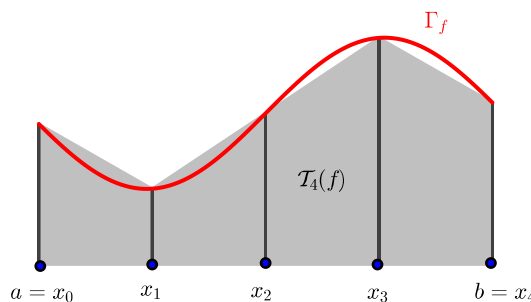
Satz 6.17 (Trapez-Regel)

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine 2-mal stetig differenzierbare Funktion, $K := \max\{f''(\xi) \mid \xi \in [a, b]\}$ das Maximum von f'' auf $[a, b]$, $\mathcal{P}_n = \{x_0, x_1, \dots, x_n\}$ die äquidistante Zerlegung von $[a, b]$ in n Teilintervalle und

$$\mathcal{T}_n(f) = \frac{b-a}{n} \cdot \sum_{k=1}^n \frac{f(x_k) + f(x_{k-1})}{2}$$

die Trapezsumme von f zur Schrittweite $\frac{b-a}{n}$.
Dann gilt

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \mathcal{T}_n(f)$$



mit der Fehlerabschätzung

$$\left| \int_a^b f(x) dx - \mathcal{T}_n(f) \right| \leq \frac{(b-a)^3 \cdot K}{12} \cdot \frac{1}{n^2}.$$

Beweis. 1. Die Darstellung des Riemann-Integrals als Grenzwert der Trapez-Summen $\mathcal{T}_n(f)$ wurde in Übungsaufgabe 6 (Analysis 2) bewiesen.

2. Für die Fehlerabschätzung betrachten wir zunächst einen Spezialfall und zeigen:

Ist $g : [0, 1] \rightarrow \mathbb{R}$ eine 2-mal stetig differenzierbare Funktion, so existiert ein $\xi \in [0, 1]$ mit

$$\int_0^1 g(x) dx = \frac{1}{2}(g(0) + g(1)) - \frac{1}{12}g''(\xi) = \mathcal{T}_1(g) - \frac{1}{12}g''(\xi). \tag{6.9}$$

Sei dazu $\psi : [0, 1] \rightarrow \mathbb{R}$ die Funktion $\psi(x) := \frac{1}{2}(x - x^2)$. Dann gilt $\psi \geq 0$, $\psi(0) = \psi(1) = 0$, $\psi'(x) = \frac{1}{2} - x$ und $\psi''(x) = -1$. Mit partieller Integration folgt

$$\begin{aligned} \int_0^1 \psi(x)g''(x) dx &\stackrel{p.Int.}{=} - \int_0^1 \psi'(x)g'(x) dx + \underbrace{\psi(x)g'(x)}_{=0} \Big|_0^1 \\ &\stackrel{p.Int.}{=} \int_0^1 \psi''(x)g(x) dx - \psi'(x)g(x) \Big|_0^1 \\ &= - \int_0^1 g(x) dx + \frac{1}{2}(g(1) + g(0)). \end{aligned} \tag{6.10}$$

Aus dem Mittelwertsatz der Integralrechnung (anwendbar wegen $\psi \geq 0$) folgt andererseits die Existenz eines $\xi \in [0, 1]$ mit

$$\int_0^1 \psi(x)g''(x) dx = g''(\xi) \int_0^1 \psi(x) dx = \frac{1}{2} \cdot g''(\xi) \int_0^1 (x - x^2) dx = \frac{1}{12}g''(\xi). \quad (6.11)$$

Aus (6.10) und (6.11) folgt die Behauptung (6.9).

3. Wir beweisen nun die Fehlerabschätzung des Satzes. Für die Teilungspunkte der äquidistanten Zerlegung $\mathcal{P}_n = \{x_0, x_1, \dots, x_n\}$ gilt $x_k = a + k \cdot \frac{b-a}{n}$. Sei $h := \frac{b-a}{n}$. Dann folgt mit der Substitution $\phi_k : [0, 1] \rightarrow [x_{k-1}, x_k]$,

$$\phi_k(t) := x_{k-1} + t \cdot h,$$

unter Benutzung von Schritt 2.:

$$\begin{aligned} \int_a^b f(x) dx &= \sum_{k=1}^n \int_{x_{k-1}}^{x_k} f(x) dx \stackrel{\text{Subst.}}{=} \sum_{k=1}^n \int_0^1 f(\phi_k(t)) \cdot h dt = h \cdot \sum_{k=1}^n \int_0^1 \underbrace{(f \circ \phi_k)}_{=g}(t) dt \\ &\stackrel{(6.9)}{=} h \cdot \sum_{k=1}^n \left(\frac{f(x_k) + f(x_{k-1})}{2} - \frac{1}{12}(f \circ \phi_k)''(\xi_k) \right) \\ &= \mathcal{T}_n(f) - \frac{h}{12} \cdot \sum_{k=1}^n (f \circ \phi_k)''(\xi_k) \end{aligned}$$

für gewisse $\xi_k \in [0, 1]$. Aus der Kettenregel für die Ableitungen folgt

$$\begin{aligned} (f \circ \phi_k)'(t) &= f'(\phi_k(t)) \cdot \phi_k'(t) = f'(\phi_k(t)) \cdot h, \\ (f \circ \phi_k)''(t) &= f''(\phi_k(t)) \cdot h^2. \end{aligned}$$

Folglich gilt:

$$\begin{aligned} \left| \int_a^b f(x) dx - \mathcal{T}_n(f) \right| &= \frac{h^3}{12} \cdot \left| \sum_{k=1}^n f''(\phi_k(\xi_k)) \right| \leq \frac{h^3}{12} \cdot \sum_{k=1}^n |f''(\phi_k(\xi_k))| \\ &\leq \frac{h^3}{12} \cdot nK = \frac{(b-a)^3 \cdot K}{12} \cdot \frac{1}{n^2}. \end{aligned}$$

□

Eine andere Möglichkeit, Integrale näherungsweise zu berechnen, besteht darin, den Integranden zu approximieren.

Definition 6.7. Sei (f_n) eine Folge von Funktionen $f_n : I \rightarrow \mathbb{R}$ und $f : I \rightarrow \mathbb{R}$ eine weitere Funktion. Man sagt:

i) (f_n) konvergiert punktweise gegen f (Bez: $f_n \rightarrow f$), wenn $\lim_{n \rightarrow \infty} f_n(x) = f(x)$ für alle $x \in I$,

ii) (f_n) konvergiert gleichmäßig gegen f (Bez: $f_n \rightrightarrows f$), wenn für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass

$$|f(x) - f_n(x)| < \varepsilon \quad \forall n \geq n_0 \quad \text{und} \quad \forall x \in I.$$

Die gleichmäßige Konvergenz von (f_n) ist stärker als die punktweise Konvergenz. Sie bedeutet, dass die Zahlenfolge $(f_n(x))$ für jedes $x \in I$ mit der "gleichen Geschwindigkeit" gegen $f(x)$ konvergiert. Dies hat wichtige Eigenschaften der Grenzfunktion f zur Folge.

Satz 6.18 Sei (f_n) eine Folge stetiger Funktionen $f_n : I \rightarrow \mathbb{R}$, die gleichmäßig gegen $f : I \rightarrow \mathbb{R}$ konvergent. Dann ist die Grenzfunktion f ebenfalls stetig.

Beweis. Sei $x_0 \in I$ beliebig fixiert. Wir zeigen, dass f in x_0 stetig ist. Sei $\varepsilon > 0$. Da (f_n) gleichmäßig gegen f konvergiert, existiert ein $n_0 \in \mathbb{N}$ mit

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3} \quad \forall n \geq n_0, \forall x \in I.$$

Die Abbildung f_{n_0} ist in x_0 stetig, folglich gibt es ein $\delta > 0$, so dass

$$|f_{n_0}(x_0) - f_{n_0}(x)| < \frac{\varepsilon}{3} \quad \forall x \in I \text{ mit } |x_0 - x| < \delta.$$

Für alle $x \in I$ mit $|x_0 - x| < \delta$ erhalten wir damit

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_{n_0}(x)| + |f_{n_0}(x) - f_{n_0}(x_0)| + |f_{n_0}(x_0) - f(x_0)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Somit ist f in $x_0 \in I$ stetig. □

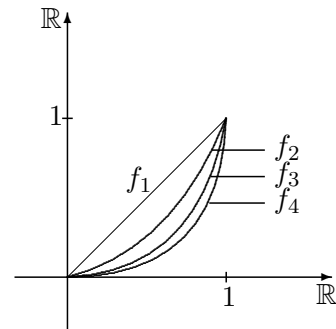
Beispiel 1: Eine nicht gleichmäßig konvergente Funktionenfolge.

Seien $f_n : [0, 1] \rightarrow \mathbb{R}$ die stetigen Funktionen $f_n(x) := x^n$.

(f_n) konvergiert punktweise gegen die Funktion

$$f(x) := \lim_{n \rightarrow \infty} f_n(x) = \begin{cases} 0 & \text{falls } 0 \leq x < 1 \\ 1 & \text{falls } x = 1 \end{cases}$$

f ist in 1 nicht stetig (insbesondere ist (f_n) nicht gleichmäßig gegen f konvergent).



Beispiel 2: Gleichmäßige Konvergenz von Potenzreihen.

Sei $f(x) := \sum_{n=1}^{\infty} a_n(x-x_0)^n$ eine reelle Potenzreihe mit positivem Konvergenzradius $R > 0$

und $p_N(x) := \sum_{n=0}^N a_n(x-x_0)^n$ ihre N -te Partialsumme. Dann konvergiert die Folge (p_N) auf jedem abgeschlossenen Intervall $[x_0 - r, x_0 + r]$ mit $0 < r < R$ gleichmäßig gegen f .

(Man sagt dann, dass die Potenzreihe $f(x)$ auf $[x_0 - r, x_0 + r]$ gleichmäßig konvergiert.)

Beweis. Sei r_1 eine reelle Zahl mit $0 < r < r_1 < R$. Dann ist $f(x_0 + r_1) = \sum_{n=0}^{\infty} a_n r_1^n$ konvergent, die Folge $(a_n r_1^n)$ somit eine Nullfolge und insbesondere beschränkt. Sei $C \in \mathbb{R}^+$ mit $|a_n r_1^n| \leq C$ für alle $n \in \mathbb{N}_0$. Dann gilt für alle $x \in [x_0 - r, x_0 + r]$:

$$|a_n(x - x_0)^n| = |a_n r_1^n| \cdot \frac{|x - x_0|^n}{r_1^n} \leq C \left(\frac{r}{r_1}\right)^n = C\theta^n \quad \text{mit } \theta := \frac{r}{r_1} \in (0, 1).$$

Daraus folgt für alle $x \in [x_0 - r, x_0 + r]$:

$$\begin{aligned} |f(x) - p_N(x)| &= \left| \sum_{n=N+1}^{\infty} a_n(x - x_0)^n \right| \leq \sum_{n=N+1}^{\infty} |a_n(x - x_0)^n| \\ &\leq C \cdot \sum_{n=N+1}^{\infty} \theta^n = C \cdot \theta^{N+1} \cdot \frac{1}{1 - \theta} \quad (\text{GW geometr. Reihe}). \end{aligned}$$

Da (θ^{N+1}) eine Nullfolge ist, existiert für jedes $\varepsilon > 0$ ein $N_0 \in \mathbb{N}_0$, so dass

$$|f(x) - p_N(x)| < \varepsilon \quad \forall N \geq N_0 \quad \text{und } x \in [x_0 - r, x_0 + r],$$

(p_N) konvergiert somit auf $[x_0 - r, x_0 + r]$ gleichmäßig gegen f . \square

Satz 6.19 (Vertauschbarkeit von Limes und Integral)

Sei (f_n) eine Folge stetiger³ Funktionen $f_n : [a, b] \rightarrow \mathbb{R}$, die gleichmäßig gegen eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ konvergiert. Dann ist f Riemann-integrierbar und für das Riemann-Integral gilt:

$$\int_a^b f(x) dx = \int_a^b \lim_{n \rightarrow \infty} f_n(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

Beweis. Da (f_n) gleichmäßig gegen f konvergiert, ist f stetig und somit Riemann-integrierbar. Nach Definition existiert zu beliebig gegebenem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass

$$|f(x) - f_n(x)| < \frac{\varepsilon}{3(b-a)} \quad \forall n \geq n_0 \quad \text{und } \forall x \in [a, b]. \quad (6.12)$$

Wir approximieren die Riemann-Integrale durch Riemannsche Summen.

Sei $\mathcal{P}_m = \{x_0, x_1, \dots, x_m\}$ die äquidistante Zerlegung von $[a, b]$ in m Teilintervalle und ξ_m das Stützstellentupel $\xi_m = (x_1, x_2, \dots, x_m)$. Dann gilt für die Riemanschen Summen

$$\begin{aligned} |S(f, \mathcal{P}_m, \xi_m) - S(f_n, \mathcal{P}_m, \xi_m)| &= \left| \sum_{k=1}^m (f(x_k) - f_n(x_k)) \cdot \frac{b-a}{m} \right| \\ &\leq \sum_{k=1}^m |f(x_k) - f_n(x_k)| \cdot \frac{b-a}{m} \\ &\stackrel{(6.12)}{<} \frac{\varepsilon}{3(b-a)} \cdot (b-a) = \frac{\varepsilon}{3} \end{aligned} \quad (6.13)$$

³ Dieser Satz gilt auch allgemeiner für eine Folge gleichmäßig konvergenter Riemann-integrierbarer Funktionen. Die Grenzfunktion ist dann auch Riemann-integrierbar.

für alle $n \geq n_0$ und $m \in \mathbb{N}$. Da die Riemannschen Summen von f bzw. f_n gegen die jeweiligen Riemann-Integrale konvergieren (Satz 6.5), existiert ein $m_0 \in \mathbb{N}$ und ein $m_1(n) \in \mathbb{N}$, so dass

$$\left| \int_a^b f(x) dx - S(f, \mathcal{P}_m, \xi_m) \right| < \frac{\varepsilon}{3} \quad \forall m \geq m_0 \quad (6.14)$$

und

$$\left| \int_a^b f_n(x) dx - S(f_n, \mathcal{P}_m, \xi_m) \right| < \frac{\varepsilon}{3} \quad \forall m \geq m_1(n). \quad (6.15)$$

Für alle $n \geq n_0$ folgt aus (6.13), (6.14) und (6.15) mit einem beliebig gewählten $m \geq \max(m_0, m_1(n))$:

$$\begin{aligned} \left| \int_a^b f(x) dx - \int_a^b f_n(x) dx \right| &\leq \left| \int_a^b f(x) dx - S(f, \mathcal{P}_m, \xi_m) \right| + \left| S(f, \mathcal{P}_m, \xi_m) - S(f_n, \mathcal{P}_m, \xi_m) \right| + \\ &\quad + \left| S(f_n, \mathcal{P}_m, \xi_m) - \int_a^b f_n(x) dx \right| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Dies zeigt die Behauptung des Satzes. \square

Als Spezialfall erhalten wir

Satz 6.20 Ist $f(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$ eine reelle Potenzreihe mit positivem Konvergenzradius $R > 0$ und $\psi : [c, d] \rightarrow \mathbb{R}$ eine stetige Funktion mit Bild im Konvergenzintervall $(x_0 - R, x_0 + R)$ von f . Dann gilt:

$$\int_c^d f(\psi(t)) dt = \int_c^d \sum_{n=0}^{\infty} a_n(\psi(t) - x_0)^n dt = \sum_{n=0}^{\infty} a_n \int_c^d (\psi(t) - x_0)^n dt.$$

Beweis. Da ψ stetig ist, ist $\psi([c, d])$ ein abgeschlossenes Intervall in Konvergenzintervall von f , d.h. es existiert ein $0 < r < R$ so dass $\psi([c, d]) \subset [x_0 - r, x_0 + r]$. Bezeichne (p_N) die Partialsummenfolge von f . Da (p_N) auf $[x_0 - r, x_0 + r]$ gleichmäßig gegen f konvergiert, konvergiert auch $p_n \circ \psi$ auf $[c, d]$ gleichmäßig gegen $f \circ \psi$. Dann folgt die Behauptung aus Satz 6.19. \square

Beispiel: Berechne das Integral $I = \int_0^{2\pi} \sqrt{1 - k^2 \cos^2(t)} dt$ für $0 < k < 1$.

Dieses Integral können wir nicht durch elementare Funktionen ausdrücken (es gehört zu den sogenannten *elliptischen Integralen*). Wir werden es bei der Berechnung der Länge einer Ellipse wiedertreffen.

Wir geben I durch Reihenentwicklung des Integranden an. Wir betrachten die Funktion $f(x) = \sqrt{1+x}$. Dann ist der Integrand von I gegeben durch $f \circ \psi$ mit $\psi(t) = -k^2 \cos^2(t)$. In Kapitel 5 haben wir die Taylorentwicklung von f um $x_0 = 0$ bestimmt (Beispiel 5, Abschnitt 5.5) und dabei die Binomialreihe $B_{\frac{1}{2}}(x)$ erhalten: Es gilt

$$f(x) = \sqrt{1+x} = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} x^n = B_{\frac{1}{2}}(x) \quad \forall x \in (-1, 1).$$

Nach Definition ist $\binom{\frac{1}{2}}{0} = 1$ und $\binom{\frac{1}{2}}{1} = \frac{1}{2}$. Für $n \geq 2$ berechnet man

$$\begin{aligned} \binom{\frac{1}{2}}{n} &= \frac{\frac{1}{2} \cdot (\frac{1}{2} - 1) \cdot (\frac{1}{2} - 2) \cdot \dots \cdot (\frac{1}{2} - (n-1))}{n!} = \frac{1}{2^n} \cdot \frac{1 \cdot (1-2) \cdot (1-4) \cdot \dots \cdot (1-2(n-1))}{1 \cdot 2 \cdot \dots \cdot n} \\ &= (-1)^{n-1} \cdot \frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-3)}{2 \cdot 4 \cdot 6 \cdot \dots \cdot (2n)}. \end{aligned}$$

Aus Satz 6.20 folgt damit:

$$\begin{aligned} I &= \int_0^{2\pi} \sqrt{1 - k^2 \cos^2(t)} dt = \int_0^{2\pi} \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-1)^n k^{2n} \cos^{2n}(t) dt \\ &= \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} (-1)^n \cdot \int_0^{2\pi} \cos^{2n}(t) dt \cdot k^{2n} \\ &\stackrel{\text{Aufg. 9}}{=} 2\pi \left(1 - \sum_{n=1}^{\infty} \left[\left(\frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1)}{2 \cdot 4 \cdot 6 \cdot \dots \cdot (2n)} \right)^2 \cdot \frac{1}{2n-1} \cdot k^{2n} \right] \right) \\ &= 2\pi \left(1 - \frac{1}{4}k^2 - \frac{3}{64}k^4 - \frac{5}{256}k^6 - \dots \right). \end{aligned}$$

6.3 Uneigentliche Riemann-Integrale

Bisher haben wir das Riemann-Integral für Funktionen über *abgeschlossenen* Intervallen definiert. Wir wollen die Integraldefinition jetzt auf offene, halb-offene bzw. unendliche Intervalle ausweiten.

Definition 6.8. Eine Funktion $f : [a, b) \subset \mathbb{R} \rightarrow \mathbb{R}$ mit $a < b \leq +\infty$ heißt *uneigentlich Riemann-integrierbar*, wenn $f|_{[a,s]}$ für alle abgeschlossenen Intervalle $[a, s] \subset [a, b)$ Riemann-integrierbar ist und der Grenzwert

$$\int_a^b f(x) dx := \lim_{s \rightarrow b^-} \int_a^s f(x) dx$$

existiert.

Eine Funktion $f : (a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$ mit $-\infty \leq a < b$ heißt *uneigentlich Riemann-integrierbar*, wenn $f|_{[s,b]}$ für alle abgeschlossenen Intervalle $[s, b] \subset (a, b]$ Riemann-integrierbar ist und der Grenzwert

$$\int_a^b f(x) dx := \lim_{s \rightarrow a^+} \int_s^b f(x) dx$$

existiert.

Eine Funktion $f : (a, b) \subset \mathbb{R} \rightarrow \mathbb{R}$ mit $-\infty \leq a < b \leq +\infty$ heißt *uneigentlich Riemann-integrierbar*, wenn $f|_{[s,t]}$ für jedes abgeschlossene Intervall $[s, t] \subset (a, b)$

Riemann-integrierbar ist und die *uneigentlichen Riemann-Integrale* $\int_{x_0}^b f(x) dx$ und $\int_a^{x_0} f(x) dx$ für ein $x_0 \in (a, b)$ existieren. Wir definieren dann

$$\int_a^b f(x) dx := \int_a^{x_0} f(x) dx + \int_{x_0}^b f(x) dx.$$

(Dies ist unabhängig von der Wahl von x_0 .)

Die Rechenregeln für Riemann-Integrale (Linearität, Monotonie, Aufspaltungseigenschaft, ...) übertragen sich mit Hilfe der Grenzwertsätze für Folgen auf die uneigentlichen Integrale.

Beispiel 1: Die Gamma-Funktion

Die Gamma-Funktion $\Gamma : \mathbb{R}^+ \rightarrow \mathbb{R}$ ist definiert durch

$$\Gamma(x) := \int_0^\infty t^{x-1} e^{-t} dt.$$

Wir zeigen, dass das uneigentliche Integral $\Gamma(x)$ für jedes $x \in \mathbb{R}^+$ existiert:

(1) Für $t > 0$ gilt $0 < t^{x-1} e^{-t} < t^{x-1}$. Für $0 < \alpha < 1$ erhalten wir damit

$$0 < \int_\alpha^1 t^{x-1} e^{-t} dt < \int_\alpha^1 t^{x-1} dt = \frac{1}{x} t^x \Big|_\alpha^1 = \frac{1}{x} - \frac{\alpha^x}{x} < \frac{1}{x}.$$

Da die Funktion $g(\alpha) = \int_\alpha^1 t^{x-1} e^{-t} dt$ für $\alpha \rightarrow 0^+$ monoton wachsend und durch $\frac{1}{x}$ beschränkt ist, existiert der Grenzwert

$$\lim_{\alpha \rightarrow 0^+} \int_\alpha^1 t^{x-1} e^{-t} dt = \int_0^1 t^{x-1} e^{-t} dt.$$

(2) Es gilt $t^{x-1} e^{-t} = \frac{t^{x+1}}{e^t} \cdot \frac{1}{t^2}$. Da $\lim_{t \rightarrow \infty} \frac{t^{x+1}}{e^t} = 0$, existiert ein $M > 0$, so dass

$$t^{x-1} e^{-t} \leq \frac{1}{t^2} \quad \forall t \geq M.$$

Daraus folgt für alle $\beta > M$

$$\int_M^\beta t^{x-1} e^{-t} dt \leq \int_M^\beta \frac{1}{t^2} dt = -\frac{1}{t} \Big|_M^\beta = \frac{1}{M} - \frac{1}{\beta} < \frac{1}{M}.$$

Da $h(\beta) = \int_M^\beta t^{x-1} e^{-t} dt$ für $\beta \rightarrow \infty$ monoton wachsend und durch $\frac{1}{M}$ beschränkt ist, existiert der Grenzwert

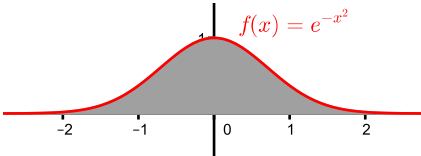
$$\lim_{\beta \rightarrow \infty} \int_M^\beta t^{x-1} e^{-t} dt = \int_M^\infty t^{x-1} e^{-t} dt.$$

Folglich existiert

$$\int_0^\infty t^{x-1} e^{-t} dt = \lim_{\alpha \rightarrow 0^+} \int_\alpha^1 t^{x-1} e^{-t} dt + \int_1^M t^{x-1} e^{-t} dt + \lim_{\beta \rightarrow \infty} \int_M^\beta t^{x-1} e^{-t} dt = \Gamma(x).$$

Die Gamma-Funktion spielt eine wichtige Rolle in der Zahlentheorie und der komplexen Analysis⁴. Es gilt z.B. $\Gamma(n) = (n - 1)!$ für alle natürlichen Zahlen (Übungsaufgabe), d.h. die Gamma-Funktion setzt die Fakultäten auf \mathbb{R}^+ fort. Man findet auch π in der Gamma-Funktion wieder. Es gilt: $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. Wir benutzen dies, um das in der Stochastik wichtige Integral über die Gaußsche Glockenfunktion $f(x) := e^{-x^2}$ zu berechnen.

Beispiel 2: Das Gauß-Integral $\int_{-\infty}^\infty e^{-x^2} dx = \sqrt{\pi}$:



Mit Hilfe der Gamma-Funktion berechnen wir mit der Substitution $t = x^2$, $dt = 2x dx$:

$$\int_0^\infty e^{-x^2} dx = \int_0^\infty e^{-t} \frac{1}{2\sqrt{t}} dt = \frac{1}{2} \int_0^\infty t^{-\frac{1}{2}} e^{-t} dt = \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}.$$

Da die Gaußsche Glockenfunktion eine gerade Funktion ist, folgt für das Gauß-Integral

$$\int_{-\infty}^\infty e^{-x^2} dx = 2 \cdot \int_0^\infty e^{-x^2} dx = \sqrt{\pi}.$$

Als Anwendung des uneigentlichen Riemann-Integrals beweisen noch wir ein Kriterium für die Konvergenz von Reihen:

Satz 6.21 (Integral-Vergleichskriterium für Reihen)

Sei $f : [n_0, \infty) \rightarrow \mathbb{R}$ eine monoton-fallende, positive Funktion, wobei $n_0 \in \mathbb{N}_0$. Dann gilt:

$$\sum_{n=n_0}^\infty f(n) \text{ konvergiert} \iff \int_{n_0}^\infty f(x) dx \text{ existiert.}$$

Im Fall der Konvergenz gilt

$$0 \leq \sum_{n=n_0}^\infty f(n) - \int_{n_0}^\infty f(x) dx \leq f(n_0).$$

⁴ Wichtige Eigenschaften der Gamma-Funktion werden z.B. im Buch von O. Forster: *Analysis 1*, 11. Auflage, § 20, bewiesen.

Beweis. Sei $n \in \mathbb{N}$ beliebig mit $n \geq n_0$. Da f monoton fallend ist, gilt $f(n) \geq f(x) \geq f(n+1)$ für alle $x \in [n, n+1]$. Aus der Monotonie des Integrals folgt

$$f(n) = \int_n^{n+1} f(n) dx \geq \int_n^{n+1} f(x) dx \geq \int_n^{n+1} f(n+1) dx = f(n+1).$$

Daraus folgt für die N -te Partialsumme der Reihe

$$\sum_{n=n_0}^N f(n) \geq \int_{n_0}^{N+1} f(x) dx \geq \sum_{n=n_0}^N f(n+1) = \sum_{n=n_0+1}^{N+1} f(n).$$

Die erste Ungleichung zeigt Behauptung \implies , die 2. Ungleichung zeigt Behauptung \impliedby . Außerdem gilt im Fall der Konvergenz

$$\sum_{n=n_0}^{\infty} f(n) - f(n_0) \leq \int_{n_0}^{\infty} f(x) dx \leq \sum_{n=n_0}^{\infty} f(n)$$

und damit die Abschätzung. □

6.4 Geometrische Anwendungen des Riemann-Integrals

Wir werden die Integralrechnung jetzt benutzen, um die Länge von Kurven sowie Flächeninhalte "komplizierterer" ebener Gebiete zu bestimmen.

Zunächst erinnern wir an einige Kenntnisse aus der Vorlesung *Lineare Algebra und analytische Geometrie I*. \mathbb{R}^n bezeichnet wie üblich die Menge der n -Tupel reeller Zahlen

$$\mathbb{R}^n := \{(x_1, x_2, \dots, x_n) \mid x_k \in \mathbb{R} \text{ für } k = 1, \dots, n\}.$$

Auf dem Vektorraum \mathbb{R}^n ist das Euklidische Skalarprodukt $\langle \cdot, \cdot \rangle$ und die Euklidische Norm $\|\cdot\|$ definiert. Für Vektoren $x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n) \in \mathbb{R}^n$ ist

$$\begin{aligned} \langle x, y \rangle &:= x_1 y_1 + x_2 y_2 + \dots + x_n y_n, \\ \|x\| &:= \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}. \end{aligned}$$

Es gilt die *Dreiecksungleichung* (Δ)

$$\|x + y\| \leq \|x\| + \|y\| \quad \forall x, y \in \mathbb{R}^n,$$

sowie die *Cauchy-Schwarzsche Ungleichung* (CSU)

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| \quad \forall x, y \in \mathbb{R}^n.$$

Der Winkel $\angle(x, y)$ zwischen zwei Vektoren $x, y \in \mathbb{R}^n \setminus \{0\}$ ist die eindeutig bestimmte Zahl $\alpha \in [0, \pi]$ mit

$$\cos \alpha = \frac{\langle x, y \rangle}{\|x\| \cdot \|y\|}.$$

Legt man in der Ebene, dem 3-dimensionalen (oder höher-dimensionalen Raum) ein kartesisches Koordinatensystem fest, so kann man sie ebenfalls durch \mathbb{R}^2 , \mathbb{R}^3 (bzw. \mathbb{R}^n) beschreiben, indem man jedem Punkt seine kartesischen Koordinaten zuordnet.

In gesamten Abschnitt 6.4 bezeichnet I wieder ein Intervall, das aus mehr als einem Punkt besteht.

6.4.1 Parametrisierte Kurven und ihre Länge

Wir verwenden einen Kurvenbegriff, der in der Kinematik wurzelt. Er ist die mathematische Abstraktion der Bewegung eines Punktes in der Ebene \mathbb{R}^2 oder im Raum \mathbb{R}^3 , die durch die Angabe des Ortes $\gamma(t)$ zum Zeitpunkt t beschrieben wird.

Definition 6.9. Eine parametrisierte Kurve im \mathbb{R}^n ist eine Abbildung $\gamma : I \rightarrow \mathbb{R}^n$ mit

$$t \in I \mapsto \gamma(t) =: (\gamma_1(t), \gamma_2(t), \dots, \gamma_n(t)) \in \mathbb{R}^n,$$

deren Komponentenfunktionen $\gamma_k : I \rightarrow \mathbb{R}$, $k = 1, \dots, n$, stetig sind.

$\gamma : I \rightarrow \mathbb{R}^n$ heißt differenzierbar (stetig differenzierbar, ...), wenn jede der Komponentenfunktionen γ_k die entsprechende Eigenschaft hat.

Die Elemente von I heißen die Parameter von γ . Die Bildmenge $\Gamma := \gamma(I) \subset \mathbb{R}^n$ nennt man auch die Spur von γ . γ heißt dann eine Parametrisierung der Menge $\Gamma \subset \mathbb{R}^n$.

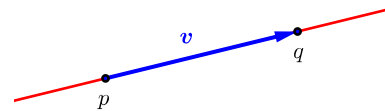
Statt parametrisierter Kurve sagen wir auch kurz nur Kurve.

Bei unserem Kurvenbegriff interessieren wir uns also nicht nur für die Spur der Kurve im \mathbb{R}^n , sondern auch dafür, wie die Kurve durchlaufen wird (z.B. mit welchem Zeitplan).

Beispiel 1 Geraden im \mathbb{R}^n

Sei $L \subset \mathbb{R}^n$ eine Gerade, $p, q \in L$ zwei verschiedene Punkte und $v := q - p$ der Verbindungsvektor von p nach q . Eine Parametrisierung von L ist gegeben durch $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ mit

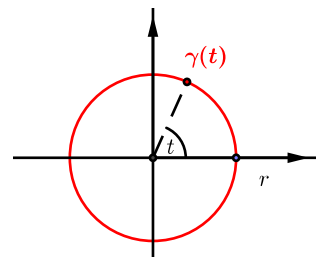
$$\gamma(t) := p + t \cdot v.$$



Beispiel 2 Der Kreis vom Radius r

Sei $K_r := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2\}$ der Kreis vom Radius r . Man kann ihn auf verschiedene Weise parametrisieren:

$$\begin{aligned} \gamma(t) &:= (r \cos t, r \sin t), & t \in [0, 2\pi], \\ \delta(t) &:= (r \cos t, -r \sin t), & t \in [0, 2\pi], \\ \eta(t) &:= (r \cos 2t, r \sin 2t), & t \in [0, 2\pi], \end{aligned}$$



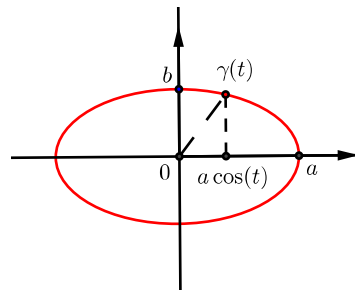
sind drei verschiedene Parametrisierungen des Kreises, die verschiedene Durchlaufesigenschaften beschreiben.

Beispiel 3 Die Ellipse mit den Halbachsen a und b

Sei $E := \{(x, y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\}$ die Ellipse mit den Halbachsen $a, b > 0$. Die Kurve

$$\gamma(t) := (a \cos t, b \sin t), \quad t \in [0, 2\pi],$$

ist eine Parametrisierung von E .



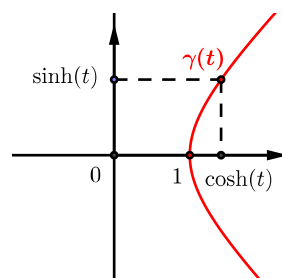
Beispiel 4 Die Hyperbel

Sei $H := \{(x, y) \in \mathbb{R}^2 \mid x^2 - y^2 = 1, x > 0\}$ die Hyperbel.

Die Kurve

$$\gamma(t) := (\cosh t, \sinh t), \quad t \in \mathbb{R},$$

ist eine Parametrisierung von H .

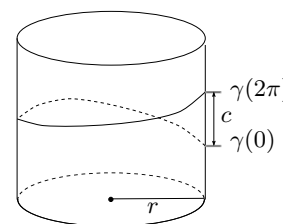


Beispiel 5: Die Schraubenlinie

Wir betrachten die Schraubenlinie $\gamma : \mathbb{R} \rightarrow \mathbb{R}^3$, gegeben durch

$$\gamma(t) := (r \cos t, r \sin t, \frac{c}{2\pi}t), \quad t \in \mathbb{R},$$

wobei $r \in \mathbb{R}^+$ und $c \in \mathbb{R}$ Konstanten sind. Die Spur von γ schraubt sich mit konstanter Ganghöhe c um den Zylinder vom Radius r .

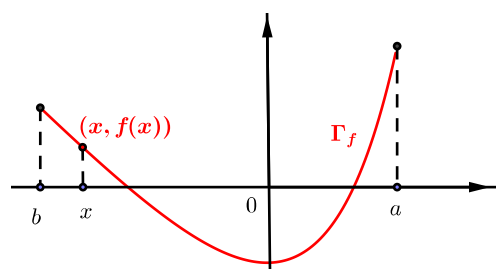


Beispiel 6: Ein Funktionsgraph

Sei $f : I \rightarrow \mathbb{R}$ eine stetige Funktion und $\Gamma_f := \{(x, f(x)) \in \mathbb{R}^2 \mid x \in I\}$ ihr Funktionsgraph. Γ_f kann man offensichtlich durch die Kurve

$$\gamma(x) := (x, f(x)), \quad x \in I,$$

parametrisieren.



Das Bild einer Kurve $\gamma : [0, 1] \rightarrow \mathbb{R}^2$ kann ein ganzes Quadrat ausfüllen (*Peano-Kurven*). Beispiele dafür findet man in K. Königsberger: *Analysis I*, Kapitel 12, Aufgabe 14, oder in H. Sagan: *Space-Filling Curves*. Solche (pathologischen) Fälle wollen wir ausschließen. Wir setzen deshalb im Folgenden voraus, dass unsere Kurven $\gamma : I \rightarrow \mathbb{R}^n$ nicht nur stetig, sondern differenzierbar sind. In diesem Fall können wir auch Tangenten an Kurven und Winkel zwischen Kurven analytisch beschreiben.

Definition 6.10. Sei $\gamma : I \rightarrow \mathbb{R}^n$ eine differenzierbare Kurve. Dann heißt

$$\gamma'(t) := (\gamma'_1(t), \gamma'_2(t), \dots, \gamma'_n(t))$$

Tangentialvektor von γ im Parameter $t \in I$. Die Kurve γ heißt im Parameter $t \in I$ regulär, wenn $\gamma'(t) \neq 0$. γ heißt regulär, wenn sie in jedem Parameter regulär ist.

Wir definieren für Folgen im \mathbb{R}^n den folgenden Konvergenzbegriff:

Definition 6.11. Eine Folge $(x_m = (x_{m1}, x_{m2}, \dots, x_{mn}))_{m=1}^\infty$ konvergiert gegen $z = (z_1, z_2, \dots, z_n)$, wenn die Folge der Komponenten $(x_{mk})_{m=1}^\infty$ gegen z_k konvergiert, d.h.

$$\lim_{m \rightarrow \infty} x_m = z \quad :\iff \quad \lim_{m \rightarrow \infty} x_{mk} = z_k \quad \forall k \in \{1, \dots, n\}.$$

Dann gilt für den Tangentialvektor:

$$\gamma'(t) = \lim_{h \rightarrow 0} \frac{\gamma(t+h) - \gamma(t)}{h}.$$

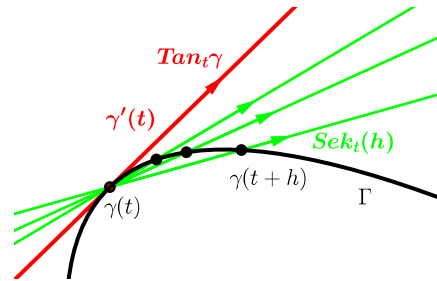
Ist $\gamma'(t) \neq 0$, so beschreibt die Gerade

$$\text{Tan}_t \gamma := \gamma(t) + \mathbb{R} \cdot \gamma'(t)$$

die Tangente von γ im Parameter t .

Sie entsteht als Grenzgerade der Sekanten $\text{Sek}_t(h)$ durch $\gamma(t)$ und $\gamma(t+h)$ bei $h \rightarrow 0$:

$$\text{Sek}_t(h) := \gamma(t) + \mathbb{R} \cdot \frac{\gamma(t+h) - \gamma(t)}{h}.$$



Sind γ und δ zwei reguläre differenzierbare Kurven, die sich im Punkt $p \in \mathbb{R}^n$ schneiden, d.h. $p = \gamma(t_0) = \delta(s_0)$, dann versteht man unter dem Schnittwinkel von γ und δ im Punkt p den Winkel zwischen den Tangentialvektoren $\gamma'(t_0)$ und $\delta'(s_0)$:

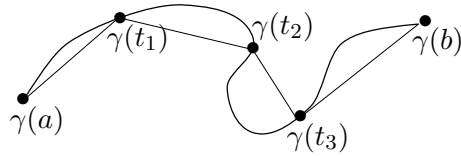
$$\angle_p(\gamma, \delta) := \angle(\gamma'(t_0), \delta'(s_0)) = \arccos \frac{\langle \gamma'(t_0), \delta'(s_0) \rangle}{\|\gamma'(t_0)\| \cdot \|\delta'(s_0)\|}.$$

Um die Länge einer Kurve $\gamma : I \rightarrow \mathbb{R}^n$ zu definieren, benutzen wir die geometrische Intuition. Bereits im Altertum haben Mathematiker den Kreisumfang berechnet, in dem sie ihn durch einbeschriebene reguläre n -Ecke approximiert haben. Für allgemeine Kurven knüpft man an dieses Verfahren an.

Sei $\mathcal{P} = \{t_0, t_1, \dots, t_m\}$ eine endliche Menge von Teilungspunkten des Intervalls I mit $t_0 < t_1 < \dots < t_m$. Ist $I = [a, b]$, so setzt man $t_0 = a$ und $t_m = b$.

Dann beschreibt

$$L(\gamma, \mathcal{P}) := \sum_{k=1}^m \|\gamma(t_k) - \gamma(t_{k-1})\|$$



die Länge des durch die Zerlegung \mathcal{P} definierten Sehnepolygons durch die Punkte $\gamma(t_0), \gamma(t_1), \dots, \gamma(t_m)$.

Ist $\tilde{\mathcal{P}} \geq \mathcal{P}$, so folgt aus der Dreiecksungleichung für die Norm $\|\cdot\|$, dass $L(\gamma, \tilde{\mathcal{P}}) \geq L(\gamma, \mathcal{P})$.

Definition 6.12. Eine Kurve $\gamma : I \rightarrow \mathbb{R}^n$ heißt rektifizierbar, wenn die Menge der Längen aller einbeschriebenen Sehnepolygone beschränkt ist. In diesem Fall nennt man

$$L(\gamma) := \sup\{L(\gamma, \mathcal{P}) \mid \mathcal{P} \text{ endliche Zerlegung von } I\}$$

die Länge von γ .

Ist γ rektifizierbar und $I = I_1 \cup I_2$ eine Zerlegung von I in zwei Teilintervalle, so sind $\gamma|_{I_1}$ und $\gamma|_{I_2}$ ebenfalls rektifizierbar und es gilt

$$L(\gamma) = L(\gamma|_{I_1}) + L(\gamma|_{I_2}).$$

Nicht jede differenzierbare Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^n$ ist rektifizierbar (ein Beispiel werden wir in den Übungen sehen), aber stetig differenzierbare Kurven sind dies immer und ihre Länge kann man mit Hilfe des Riemann-Integrals berechnen. Um dies zu beweisen, machen wir zunächst eine Vorbetrachtung.

Sei $f : [a, b] \rightarrow \mathbb{R}^n$ eine stetige⁵ Abbildung mit den Komponenten $f = (f_1, \dots, f_n)$. Dann ist $\|f\| : [a, b] \rightarrow \mathbb{R}$ ebenfalls stetig. Wir definieren das Riemann-Integral von f komponentenweise durch

$$\int_a^b f(x) dx := \left(\int_a^b f_1(x) dx, \dots, \int_a^b f_n(x) dx \right).$$

Satz 6.22 Sei $f : [a, b] \rightarrow \mathbb{R}^n$ stetig. Dann gilt

$$\left\| \int_a^b f(x) dx \right\| \leq \int_a^b \|f(x)\| dx.$$

Beweis. Für den Beweis benutzen wir die Cauchy-Schwarzsche Ungleichung.

Sei $v := \int_a^b f(x) dx =: (v_1, \dots, v_n) \in \mathbb{R}^n$. Dann gilt:

$$\begin{aligned} \|v\|^2 &= \langle v, v \rangle = \sum_{k=1}^n v_k \cdot \int_a^b f_k(x) dx = \int_a^b \sum_{k=1}^n v_k \cdot f_k(x) dx = \int_a^b \langle v, f(x) \rangle dx \\ &\stackrel{CSU}{\leq} \int_a^b \|v\| \cdot \|f(x)\| dx = \|v\| \cdot \int_a^b \|f(x)\| dx. \end{aligned}$$

⁵ d.h. alle Komponentenfunktionen seien stetig

Ist $v = 0$, so ist die Behauptung des Satzes offensichtlich erfüllt. Ist $v \neq 0$, so teilen wir durch $\|v\| \neq 0$ und erhalten die Behauptung. \square

Satz 6.23 Sei $\gamma : [a, b] \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Kurve. Dann ist γ rektifizierbar und es gilt

$$L(\gamma) = \int_a^b \|\gamma'(t)\| dt.$$

Beweis. (1) Wir zeigen zuerst, dass $\int_a^b \|\gamma'(t)\| dt$ eine obere Schranke für die Längen aller einbeschriebenen Sehnenpolygone ist. Insbesondere ist γ dann rektifizierbar.

Sei $\mathcal{P} = \{t_0, t_1, \dots, t_m\}$ eine beliebige Zerlegung von $[a, b]$. Dann gilt

$$\begin{aligned} L(\gamma, \mathcal{P}) &= \sum_{k=1}^m \|\gamma(t_k) - \gamma(t_{k-1})\| = \sum_{k=1}^m \left\| \int_{t_{k-1}}^{t_k} \gamma'(t) dt \right\| \\ &\stackrel{S.6.22}{\leq} \sum_{k=1}^m \int_{t_{k-1}}^{t_k} \|\gamma'(t)\| dt = \int_a^b \|\gamma'(t)\| dt. \end{aligned}$$

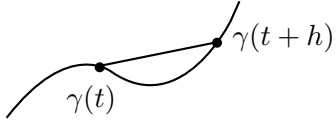
Insbesondere gilt

$$L(\gamma) := \sup\{L(\gamma, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } [a, b]\} \leq \int_a^b \|\gamma'(t)\| dt. \tag{6.16}$$

(2) Wir zeigen nun, dass in (6.16) Gleichheit gilt. Wir betrachten dazu die Hilfsfunktion $\ell : [a, b] \rightarrow [0, L(\gamma)]$,

$$\ell(t) := L(\gamma|_{[a,t]}) = \sup\{L(\gamma|_{[a,t]}, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } [a, t]\}.$$

Sei $t \in [a, b]$, $h > 0$ und $t + h \in [a, b]$. Wir wenden (6.16) auf $\gamma|_{[t,t+h]}$ an und erhalten

$$\|\gamma(t+h) - \gamma(t)\| \leq L(\gamma|_{[t,t+h]}) \leq \int_t^{t+h} \|\gamma'(s)\| ds.$$


Wegen

$$L(\gamma|_{[t,t+h]}) = L(\gamma|_{[a,t+h]}) - L(\gamma|_{[a,t]}) = \ell(t+h) - \ell(t)$$

folgt dann

$$\left\| \frac{\gamma(t+h) - \gamma(t)}{h} \right\| \leq \frac{\ell(t+h) - \ell(t)}{h} \leq \frac{1}{h} \int_t^{t+h} \|\gamma'(s)\| ds = \frac{1}{h} (F(t+h) - F(t)),$$

wobei F die Stammfunktion von $\|\gamma'\|$ bezeichnet. Durch Limesbildung $\lim_{h \rightarrow 0^+}$ ergibt sich

$$\|\gamma'(t)\| \leq \lim_{h \rightarrow 0^+} \frac{\ell(t+h) - \ell(t)}{h} \leq F'(t) = \|\gamma'(t)\|.$$

Also gilt

$$\lim_{h \rightarrow 0^+} \frac{\ell(t+h) - \ell(t)}{h} = \|\gamma'(t)\|.$$

Auf analoge Weise zeigt man

$$\lim_{h \rightarrow 0^-} \frac{\ell(t+h) - \ell(t)}{h} = \|\gamma'(t)\|.$$

Also ist ℓ differenzierbar und es gilt $\ell'(t) = \|\gamma'(t)\|$. Damit erhalten wir

$$L(\gamma) = \ell(b) - \ell(a) = \int_a^b \ell'(t) dt = \int_a^b \|\gamma'(t)\| dt. \quad \square$$

Bemerkung:

1. Sei $I \subset \mathbb{R}$ ein offenes, halboffenes oder unbeschränktes Intervall mit den Grenzen a und b , $-\infty \leq a < b \leq +\infty$, und $\gamma : I \rightarrow \mathbb{R}^n$ eine stetig differenzierbare Kurve. Dann ist γ genau dann rektifizierbar, wenn das uneigentliche Riemann-Integral $\int_a^b \|\gamma'(t)\| dt$ existiert. In diesem Fall gilt wie für abgeschlossene Intervalle $L(\gamma) = \int_a^b \|\gamma'(t)\| dt$. Ist der Grenzwert des uneigentlichen Riemann-Integrals $+\infty$, so sagen wir γ hat unendliche Länge.
2. Man nennt $\gamma : I \rightarrow \mathbb{R}^n$ stückweise stetig differenzierbar, wenn γ stetig ist und es endlich viele Teilungspunkte $a = t_0 < t_1 < t_2 < \dots < t_m = b$ des Definitionsbereiches I gibt, so dass $\gamma|_{(t_{k-1}, t_k)}$ stetig differenzierbar ist. Dann gilt

$$L(\gamma) = \sum_{k=1}^m L(\gamma|_{(t_{k-1}, t_k)}) = \sum_{k=1}^m \int_{t_{k-1}}^{t_k} \|\gamma'(t)\| dt.$$

Beispiel 1: *Kreisumfang, Länge eines Kreisbogens*

Wir betrachten den Kreis vom Radius r , parametrisiert durch $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$ mit $\gamma(t) := (r \cos t, r \sin t)$. Dann ist

$$\gamma'(t) = (-r \sin t, r \cos t) \quad \text{und} \quad \|\gamma'(t)\| = r.$$

Folglich erhalten wir für die Länge

$$L(\gamma) = \int_0^{2\pi} r dt = 2\pi r \quad \text{sowie} \quad L(\gamma|_{[0, \varphi]}) = \varphi \cdot r.$$

Beispiel 2: *Länge der Schraubenlinie*

Wir betrachten die Schraubenlinie $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^3$, gegeben durch $\gamma(t) = (r \cos t, r \sin t, ct)$.

Dann ist

$$\gamma'(t) = (-r \sin t, r \cos t, c) \quad \text{und} \quad \|\gamma'(t)\| = \sqrt{r^2 + c^2}.$$

Folglich erhalten wir für die Länge der Schraubenlinie bei einer Umdrehung

$$L(\gamma) = 2\pi\sqrt{r^2 + c^2}.$$

Beispiel 3: *Umfang der Ellipse*

Wir betrachten die Ellipse $E := \{(x, y) \in \mathbb{R}^2 \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\}$, wobei $a, b \in \mathbb{R}^+$. Es gelte $a > b$. Sei $k^2 := 1 - \frac{b^2}{a^2} < 1$. Wir parametrisieren die Ellipse E durch die Kurve $\gamma : [0, 2\pi) \rightarrow \mathbb{R}^2$ mit $\gamma(t) = (a \cos t, b \sin t)$. Dann gilt

$$\gamma'(t) = (-a \sin t, b \cos t) \quad \text{und} \quad \|\gamma'(t)\| = \sqrt{a^2 \sin^2 t + b^2 \cos^2 t}$$

und wir erhalten

$$L(\gamma) = a \int_0^{2\pi} \sqrt{\sin^2 t + \frac{b^2}{a^2} \cos^2 t} dt = a \int_0^{2\pi} \sqrt{1 - k^2 \cos^2 t} dt.$$

Das Integral hatten wir in Abschnitt 6.2. mittels Reihenentwicklung berechnet. Es folgt

$$L(\gamma) = 2\pi a \cdot \left(1 - \sum_{n=1}^{\infty} \left[\left(\frac{1 \cdot 3 \cdot 5 \cdot \dots \cdot (2n-1)}{2 \cdot 4 \cdot 6 \cdot \dots \cdot (2n)} \right)^2 \cdot \frac{1}{2n-1} \cdot k^{2n} \right] \right)$$

Beispiel 4: *Länge eines Graphen*

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion und $\Gamma_f := \{(x, f(x)) \mid x \in [a, b]\}$ ihr Graph. Wir parametrisieren Γ_f durch $\gamma : [a, b] \rightarrow \mathbb{R}^2$ mit $\gamma(x) := (x, f(x))$. Dann gilt

$$\gamma'(x) = (1, f'(x)) \quad \text{und} \quad \|\gamma'(x)\| = \sqrt{1 + f'(x)^2}.$$

Folglich ist die Länge

$$L(\gamma) = \int_a^b \sqrt{1 + f'(x)^2} dx.$$

Wir haben die geometrischen Eigenschaften Tangente, Schnittwinkel, Länge einer Kurve mit Hilfe ihrer Parametrisierung erklärt. Man kann eine Spurkurve Γ natürlich auf verschiedene Weisen (z.B. mit unterschiedlicher Geschwindigkeit) durchlaufen. Sinnvollerweise sollten sich ihre geometrischen Eigenschaften bei Umparametrisierung nicht ändern. Wir werden uns in einer Übungsaufgabe davon überzeugen, dass dies tatsächlich so ist.

6.4.2 Flächeninhalt ebener Gebiete

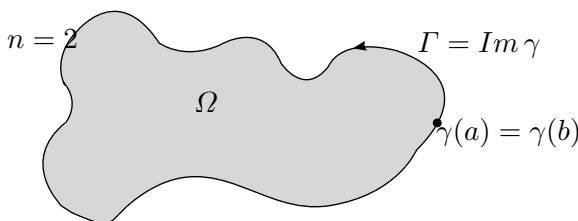
In diesem Abschnitt wollen wir Methoden herleiten, mit den wir den Flächeninhalt "komplizierterer" ebener Gebiete berechnen können.

Definition 6.13. Eine Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^2$ heißt

- geschlossen, falls $\gamma(a) = \gamma(b)$.
- einfach, falls $\gamma : (a, b) \rightarrow \mathbb{R}^2$ injektiv ist.

Eine einfache, geschlossene Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^2$ heißt positiv-orientiert, falls sie entgegen dem Uhrzeigersinn durchlaufen wird.

Sei nun $\gamma : [a, b] \rightarrow \mathbb{R}^2$ eine einfache, geschlossene und positiv-orientierte Kurve:

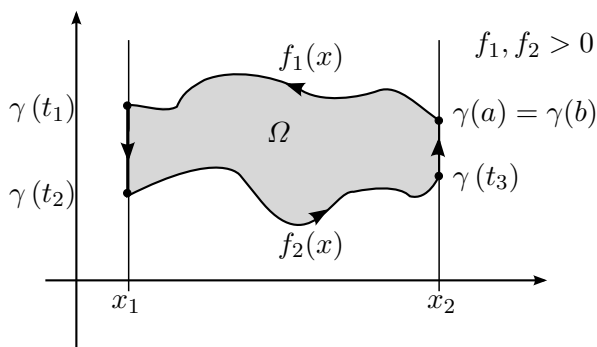


Ω bezeichne das von $\Gamma := \text{Im}(\gamma)$ umschlossene Gebiet. Wir geben nun eine Formel an, mit der man den Flächeninhalt von Ω mit Hilfe der Randkurve berechnen kann.

Satz 6.24 Sei $\gamma : [a, b] \rightarrow \mathbb{R}^2$ eine einfache, geschlossene, positiv-orientierte, (stückweise) stetig differenzierbare Kurve und sei $\Omega \subset \mathbb{R}^2$ das von $\Gamma := \text{Im}(\gamma)$ umschlossene Gebiet. Weiterhin seien $x, y : [a, b] \rightarrow \mathbb{R}$ die Koordinatenfunktionen von γ , d.h., $\gamma(t) =: (x(t), y(t))$. Dann gilt für den Flächeninhalt von Ω :

$$\begin{aligned} \text{Area}(\Omega) &= - \int_a^b y(t)x'(t) dt \stackrel{\text{p.I.}}{=} \int_a^b y'(t)x(t) dt \\ &= \frac{1}{2} \int_a^b (x(t)y'(t) - y(t)x'(t)) dt. \end{aligned}$$

Beweis. 1) Wir betrachten zunächst Kurven γ , deren Spur aus zwei zur y -Achse parallelen Strecken und zwei Bögen, die Graphen von Funktionen f_1 und f_2 mit $0 < f_1 < f_2$ sind, besteht.



Nach Definition des Riemann-Integrals ist

$$Area(\Omega) = \int_{x_1}^{x_2} f_1(x) dx - \int_{x_1}^{x_2} f_2(x) dx. \quad (*)$$

Für die obige Kurve $\gamma(t) = (x(t), y(t))$ gilt

$$\gamma(t) = \begin{cases} (x(t), f_1(x(t))), & t \in [a, t_1] \\ (x_1, y(t)), & t \in [t_1, t_2] \\ (x(t), f_2(x(t))), & t \in [t_2, t_3] \\ (x_2, y(t)), & t \in [t_3, b]. \end{cases}$$

Die Substitution $x = x(t)$, $dx = x'(t)dt$ mit $x_1 = x(t_1)$ und $x_2 = x(a)$ liefert für das erste Integral in (*)

$$\int_{x_1}^{x_2} f_1(x) dx = \int_{t_1}^a f_1(x(t)) \cdot x'(t) dt = - \int_a^{t_1} y(t)x'(t) dt.$$

Analog ergibt sich mit $x = x(t)$, $dx = x'(t)dt$ sowie $x_1 = x(t_2)$ und $x_2 = x(t_3)$ für das zweite Integral in (*)

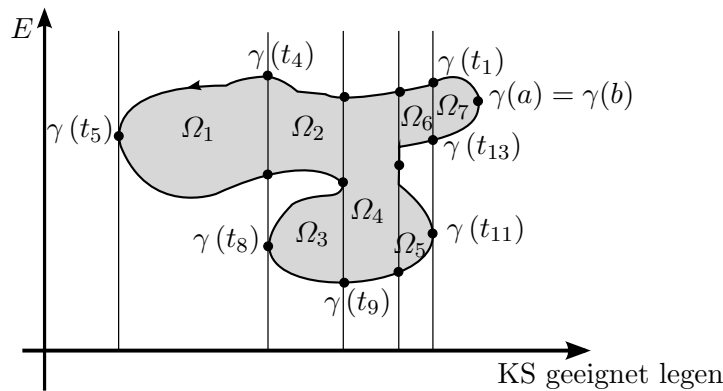
$$\int_{x_1}^{x_2} f_2(x) dx = \int_{t_2}^{t_3} y(t)x'(t) dt.$$

Damit ergibt sich insgesamt

$$Area(\Omega) = - \left(\int_a^{t_1} y(t)x'(t) dt + \int_{t_2}^{t_3} y(t)x'(t) dt \right) = - \int_a^b y(t)x'(t) dt,$$

da $x(t) = const$ für $t \in [t_1, t_2]$ und $t \in [t_3, b]$.

2) Wir betrachten nun den allgemeinen Fall. Zunächst sei die Kurve γ stetig differenzierbar.



Beh.: Man kann Ω in eine endliche Zahl von Gebieten zerlegen, die die Form aus 1) haben. Um das einzusehen, legen wir das Koordinatensystem so, dass Γ im positiven Quadranten

liegt. Sei E die y -Achse. Der Abstand des Punktes $\gamma(t)$ zur y -Achse E ist dann gegeben durch

$$x(t) = \text{dist}(E, \gamma(t)).$$

Da das Intervall $[a, b]$ kompakt ist, hat $x : [a, b] \rightarrow \mathbb{R}^+$ nur endlich viele kritische Werte. Wir zeichnen in diesen Punkten, wie im Bild dargestellt, die zu E parallelen Geraden. Dann zerlegt sich Ω in Gebiete der Form 1). Ist nämlich $x'(t) \neq 0$, so ist $x' > 0$ oder $x' < 0$. Daher ist $x(t)$ zwischen den kritischen Punkten von $x : [a, b] \rightarrow \mathbb{R}^+$ streng monoton wachsend oder streng monoton fallend. Deshalb existiert dort eine Umkehrfunktion $t = t(x)$ und das entsprechende Teilstück von γ kann man als Graph der Funktion $f(x) := y(t(x))$ darstellen:

$$\gamma(t) = (x(t), y(t)) = (x(t), f(x(t))).$$

Nun können wir die Fläche von Ω leicht berechnen:

$$\begin{aligned} \text{Area}(\Omega) &= \sum_{i=1}^m \text{Area}(\Omega_i) \\ &= -\left(\int_a^{t_1} y(t)x'(t) dt + \int_{t_{13}}^b y(t)x'(t) dt + \int_{t_1}^{t_2} y(t)x'(t) dt + \int_{t_{12}}^{t_3} y(t)x'(t) dt + \dots \right) \\ &= -\int_a^b y(t)x'(t) dt, \end{aligned}$$

wobei wir im letzten Schritt benutzt haben, dass alle Parameterabschnitte $[t_i, t_{i+1}]$ genau einmal auftreten. Ist γ nur stückweise stetig differenzierbar, so zeichnen wir durch die "Ecken" von γ zusätzliche parallele Geraden zu E und schließen dann analog. \square

Folgerung 6.3 Ist die Randkurve γ in Polarkoordinaten gegeben, d.h., gilt

$$\gamma(t) = r(t)e^{i\varphi(t)} = (r(t) \cos \varphi(t), r(t) \sin \varphi(t)),$$

so ist

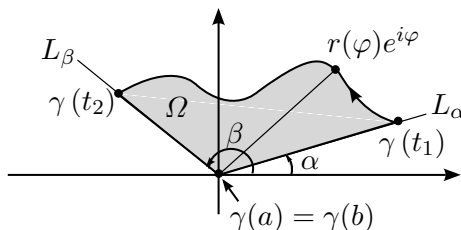
$$\text{Area}(\Omega) = \frac{1}{2} \int_a^b r^2(t)\varphi'(t) dt.$$

Folgerung 6.4 (Leibnizsche Sektorformel)

Das Gebiet Ω sei begrenzt durch die Strahlen L_α und L_β mit den Winkeln α bzw. β zur x -Achse, $0 \leq \alpha < \beta \leq 2\pi$, und eine durch Polarkoordinaten beschriebene Kurve $r = r(\varphi)$, wobei der Winkel $\varphi \in [\alpha, \beta]$ beim Durchlauf durch die Kurve streng monoton wächst.

Dann gilt

$$\text{Area}(\Omega) = \frac{1}{2} \int_\alpha^\beta r^2(\varphi) d\varphi.$$



Beweis. Wir parametrisieren die Randkurve von Ω durch $\gamma(t) = r(t)e^{i\varphi(t)}$, $t \in [a, b]$,

$$\varphi(t) = \begin{cases} \alpha, & t \in [a, t_1] \\ \varphi(t), & t \in [t_1, t_2] \\ \beta, & t \in [t_2, b] \end{cases}$$

Aus Folgerung 6.3 erhalten wir

$$Area(\Omega) = \frac{1}{2} \int_{t_1}^{t_2} r^2(t) \varphi'(t) dt.$$

Wir substituieren $\varphi = \varphi(t)$, $d\varphi = \varphi'(t)dt$. Da $\varphi(t)$ auf $[t_1, t_2]$ streng monoton wachsend ist, existiert eine Umkehrfunktion $t = t(\varphi)$. Somit können wir r in Abhängigkeit von φ darstellen, $r = r(\varphi)$, und erhalten insgesamt

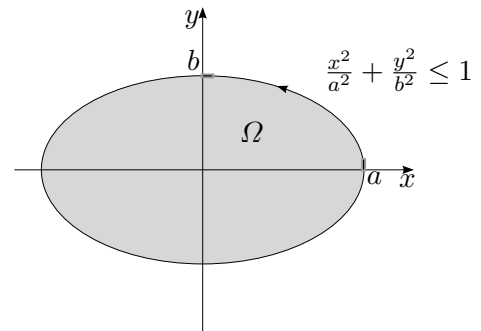
$$Area(\Omega) = \frac{1}{2} \int_{\alpha}^{\beta} r^2(\varphi) d\varphi. \quad \square$$

Beispiel 1: *Flächeninhalt einer Ellipse*

Sei Ω die von einer Ellipse mit den Halbachsen a und b eingeschlossene Fläche.

Wir parametrisieren die Ellipse durch die Kurve $\gamma(t) = (a \cos t, b \sin t) =: (x(t), y(t))$ mit $t \in [0, 2\pi]$.

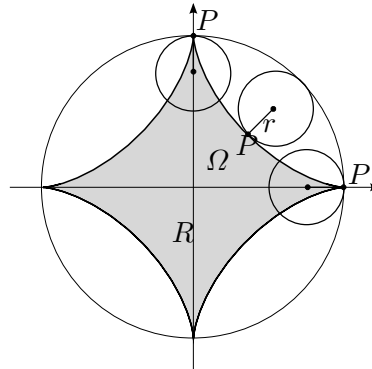
Dann ist nach Satz 6.24



$$Area(\Omega) = \frac{1}{2} \int_0^{2\pi} (x(t)y'(t) - y(t)x'(t)) dt = \frac{1}{2} \int_0^{2\pi} ab \cdot (\cos^2 t + \sin^2 t) dt = \pi ab.$$

Beispiel 2: *Flächeninhalt der Sternkurve (Astroide)*

Ein Kreis vom Radius $r = \frac{R}{4}$ rolle auf der Innenseite eines Kreises vom Radius R entlang.



Der Weg, den der feste Punkt P während des Rollens zurücklegt, wird Astroide genannt und ist bestimmt durch die Gleichung $x^{\frac{2}{3}} + y^{\frac{2}{3}} = R^{\frac{2}{3}}$. Wir können die Astroide parametrisieren durch $\gamma : [0, 2\pi] \rightarrow \mathbb{R}$ mit $\gamma(t) := R(\cos^3 t, \sin^3 t)$. Dann folgt für den Flächeninhalt des Gebietes Ω , das durch die Astroide begrenzt wird

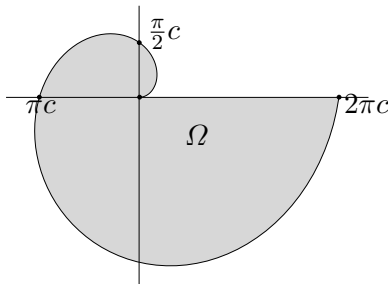
$$\begin{aligned} \text{Area}(\Omega) &= \frac{1}{2} R^2 \int_0^{2\pi} (\cos^3 t \cdot 3 \sin^2 t \cdot \cos t + 3 \cos^2 t \cdot \sin t \cdot \sin^3 t) dt \\ &= \frac{3}{2} R^2 \int_0^{2\pi} \sin^2 t \cos^2 t (\cos^2 t + \sin^2 t) dt \\ &= \frac{3}{2} R^2 \int_0^{2\pi} (\sin t \cos t)^2 dt = \frac{3}{2} R^2 \int_0^{2\pi} \left(\frac{1}{2} \sin 2t\right)^2 dt \\ &= \frac{3}{8} R^2 \int_0^{2\pi} \sin^2 2t dt, \quad x := 2t, dx = 2dt \\ &= \frac{3}{16} R^2 \int_0^{4\pi} \sin^2 x dx = \frac{3}{16} R^2 \cdot 8 \underbrace{\int_0^{\frac{\pi}{2}} \sin^2 x dx}_{=\frac{\pi}{4}} = \frac{3}{8} \pi R^2. \end{aligned}$$

Beispiel 3: Die Archimedische Spirale

Die archimedische Spirale ist die in Polarkoordinaten durch $r(\varphi) = c\varphi$, mit einer Konstanten $c \in \mathbb{R}$, beschriebene Kurve. In Euklidischen Koordinaten ist sie also durch

$$\gamma(\varphi) = r(\varphi)e^{i\varphi} = (c\varphi \cos \varphi, c\varphi \sin \varphi), \quad \varphi \in [0, 2\pi],$$

gegeben.



Nach der Leibnizschen Sektorformel ist der Flächeninhalt des Gebietes Ω

$$\text{Area}(\Omega) = \frac{1}{2} \int_0^{2\pi} r^2(\varphi) d\varphi = \frac{1}{2} c^2 \int_0^{2\pi} \varphi^2 d\varphi = \frac{1}{2} c^2 \cdot \frac{1}{3} 8\pi^3 = \frac{4}{3} \pi^3 c^2.$$

6.5 Analytische Anwendung: Differentialgleichungen mit getrennten Variablen

Eine Differentialgleichung ist eine Gleichung, die eine Beziehung zwischen einer Funktion und ihren Ableitungen beschreibt. Auf solche Gleichungen trifft man immer, wenn man dynamische Prozesse in der Natur mathematisch modellieren möchte.

Definition 6.14. Sei $F : U \subset \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$ eine Funktion. Dann heißt die Gleichung

$$y^{(n)} = F(x, y, y', \dots, y^{(n-1)}) \quad (6.17)$$

(gewöhnliche) Differentialgleichung n -ter Ordnung.

Eine Lösung von (6.17) ist eine n -mal differenzierbare Funktion $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ mit

$$f^{(n)}(x) = F(x, f(x), f'(x), \dots, f^{(n-1)}(x)) \quad \forall x \in I.$$

Ist zusätzlich $(x_0, y_0, y_1, \dots, y_{n-1}) \in U$, dann heißt die Gleichung (6.17) zusammen mit den Bedingungen

$$y(x_0) = y_0, y'(x_0) = y_1, \dots, y^{(n-1)}(x_0) = y_{n-1} \quad (6.18)$$

Anfangswertproblem (AWP) n -ter Ordnung.

Eine Lösung des Anfangswertproblems ist eine n -mal differenzierbare Funktion $f : I \rightarrow \mathbb{R}$ mit $x_0 \in I$, die die Differentialgleichung (6.17) erfüllt und zusätzlich den Anfangsbedingungen

$$f(x_0) = y_0, f'(x_0) = y_1, \dots, f^{(n-1)}(x_0) = y_{n-1}$$

genügt.

Man möchte nun folgende Fragen beantworten:

- Unter welchen Bedingungen an F gibt es Lösungen der Differentialgleichung (6.17) bzw. des Anfangswertproblems (6.17) und (6.18)?
- Unter welchen Bedingungen an F ist die Lösung eindeutig bestimmt?
- Wie lange "lebt" die Lösung (d.h. wie groß ist der Definitionsbereich I)?
- Wie kann man Lösungen bestimmen?

Zur Antwort auf diese Fragen gibt es eine weit entwickelte Theorie der *gewöhnlichen Differentialgleichungen*, die man während der Masterphase des Lehramtsstudiums in einer Wahlpflicht-Vorlesung kennenlernen kann.

Beispiel 1: Die Differentialgleichung $y^{(n)} = 0$:

In diesem Fall ist $F = 0$. Wir kennen die Lösungen bereits aus Kapitel 5. Jede Lösung $f : \mathbb{R} \rightarrow \mathbb{R}$ ist ein Polynom höchstens $(n - 1)$ -ten Grades, d.h.

$$f(x) = a_{n-1}x^{n-1} + a_{n-2}x^{n-2} + \dots + a_1x + a_0, \quad x \in \mathbb{R}.$$

Aus den Anfangsbedingungen $f(x_0) = y_0, f'(x_0) = y_1, \dots, f^{(n-1)}(x_0) = y_{n-1}$ kann man die Koeffizienten a_0, a_1, \dots, a_{n-1} eindeutig bestimmen.

Beispiel 2: Die Differentialgleichung $y' = c \cdot y$, mit $c \in \mathbb{R}$ konstant:

In diesem Fall ist $F(x, y) = c \cdot y$. Die Funktionen der Gestalt

$$f(x) = k \cdot e^{cx}, \quad x \in \mathbb{R},$$

mit einer Konstante $k \in \mathbb{R}$, lösen offensichtlich die Differentialgleichung. Es gibt keine anderen Lösungen: Ist nämlich $g : \mathbb{R} \rightarrow \mathbb{R}$ eine weitere Lösung, dann gilt für die Funktion $h : \mathbb{R} \rightarrow \mathbb{R}$ mit $h(x) := g(x) \cdot e^{-cx}$:

$$h'(x) = g'(x)e^{-cx} - cg(x)e^{-cx} = cg(x)e^{-cx} - cg(x)e^{-cx} = 0 \quad \forall x \in \mathbb{R}.$$

Folglich ist die Funktion h konstant, d.h. es existiert eine Konstante $k \in \mathbb{R}$ mit $g(x) = ke^{cx}$ für alle $x \in \mathbb{R}$. Aus der Anfangsbedingung $f(x_0) = y_0$ kann man die Konstante k eindeutig bestimmen.

Wir wollen uns hier mit einem einfachen Typ einer Differentialgleichung 1. Ordnung befassen, der *Differentialgleichung mit getrennten Variablen*.

Definition 6.15. Eine Differentialgleichung 1. Ordnung $y' = F(x, y)$ heißt *Differentialgleichung mit getrennten Variablen*, wenn die Funktion F die Gestalt

$$F(x, y) = p(x) \cdot q(y) \tag{6.19}$$

für stetige Funktionen $p : J_1 \rightarrow \mathbb{R}$ und $q : J_2 \rightarrow \mathbb{R}$ hat, wobei J_1 und J_2 offene Intervalle sind und q keine Nullstelle auf J_2 besitzt.

Satz 6.25 (Differentialgleichung mit getrennten Variablen)

Wir betrachten ein Anfangswertproblem mit getrennten Variablen

$$y' = p(x) \cdot q(y), \quad y(x_0) = y_0, \tag{6.20}$$

mit Bedingungen wie in (6.19). Sei G eine Stammfunktion von $\frac{1}{q}$ auf J_2 . Dann ist G auf J_2 injektiv und das Anfangswertproblem (6.20) besitzt auf einem hinreichend kleinen Intervall $I \subset J_1$ um x_0 eine eindeutig bestimmte Lösung. Diese ist gegeben durch

$$f(t) = G^{-1}\left(G(y_0) + \int_{x_0}^t p(x) dx\right). \tag{6.21}$$

Beweis. 1) Wir nehmen zuerst an, dass die Differentialgleichung (6.20) eine Lösung f besitzt und bestimmen ihre Form. Sei also $f : I \rightarrow \mathbb{R}$ eine differenzierbare Funktion mit $f'(x) = p(x) \cdot q(f(x))$ für alle $x \in I$ und $f(x_0) = y_0$. Dann ist insbesondere $f(I) \subset J_2$. Wir integrieren die Gleichung $\frac{f'(x)}{q(f(x))} = p(x)$ von x_0 bis $t \in I$:

$$\int_{x_0}^t \frac{f'(x)}{q(f(x))} dx = \int_{x_0}^t p(x) dx.$$

Im ersten Integral führen wir die Substitution $y = f(x)$ und $dy = f'(x) dx$ durch und erhalten mit der Stammfunktion G von $\frac{1}{q}$

$$G(f(t)) - G(y_0) = \int_{y_0}^{f(t)} \frac{1}{q(y)} dy = \int_{x_0}^t p(x) dx.$$

Dies liefert die folgende *implizite* Beschreibung von f :

$$G(f(t)) = G(y_0) + \int_{x_0}^t p(x) dx \quad \forall t \in I. \quad (6.22)$$

Insbesondere liegt die Zahl auf der rechten Seite von (6.22) für alle $t \in I$ im Bild von G . Da $\frac{1}{q}$ stetig ist und auf J_2 keine Nullstellen besitzt, gilt $G' = \frac{1}{q} > 0$ auf J_2 oder $G' = \frac{1}{q} < 0$ auf J_2 . D.h. G ist streng monoton und somit injektiv. Folglich existiert die Umkehrfunktion $G^{-1} : \text{Im}(G) \rightarrow J_2$. Somit kann man die Gleichung (6.22) nach $f(t)$ auflösen und erhält eine *explizite* Formel für f :

$$f(t) = G^{-1}\left(G(y_0) + \int_{x_0}^t p(x) dx\right) \quad \forall t \in I. \quad (6.23)$$

Wenn es also eine Lösung $f : I \rightarrow \mathbb{R}$ des AWP (6.20) gibt, so ist sie durch die Gleichung (6.23) eindeutig bestimmt.

2) Wir zeigen nun die Existenz einer Lösung. Dazu *definieren* wir uns eine Funktion durch die Gleichung (6.23): Wir betrachten ein Intervall $I \subset J_1$ um x_0 , das so klein gewählt ist, dass $G(y_0) + \int_{x_0}^t p(x) dx$ für alle $t \in I$ im Bild von G liegt und definieren $f : I \rightarrow \mathbb{R}$ durch

$$f(t) := G^{-1}\left(G(y_0) + \int_{x_0}^t p(x) dx\right) \quad \forall t \in I.$$

f löst tatsächlich das Anfangswertproblem (6.20), denn nach Kettenregel gilt für alle $t \in I$:

$$f'(t) = \underbrace{(G^{-1})' \left(G(y_0) + \int_{x_0}^t p(x) dx \right)}_{= G(f(t))} \cdot p(t) = \frac{1}{G'(f(t))} \cdot p(t) = q(f(t)) \cdot p(t)$$

und $f(x_0) = G^{-1}(G(y_0)) = y_0$. □

Satz 6.25 rechtfertigt das folgende **formale Vorgehen** zur Lösung einer Differentialgleichung mit getrennten Variablen $y' = p(x) \cdot q(y)$:

Wir setzen $y' = \frac{dy}{dx}$ und bringen alle Ausdrücke, die y enthalten auf die linke Seite und alle Ausdrücke, die x enthalten auf die rechte Seite ("Trennung der Variablen"). Danach integrieren wir links über y und rechts über x :

$$y' = \frac{dy}{dx} \implies \frac{dy}{dx} = p(x) \cdot q(y) \implies \frac{1}{q(y)} dy = p(x) dx. \quad (6.24)$$

a) Bei gegebener Anfangsbedingung $y(x_0) = y_0$ integriert man links über $[y_0, y(t)]$ und rechts über $[x_0, t]$:

$$\stackrel{(6.24)}{\implies} \int_{y_0}^{y(t)} \frac{1}{q(y)} dy = \int_{x_0}^t p(x) dx.$$

Ist G eine Stammfunktion von $\frac{1}{q}$, so erhält man daraus die implizite Form für die Lösung $y(t)$

$$G(y(t)) = G(y_0) + \int_{x_0}^t p(x) dx,$$

und durch Anwenden der Umkehrfunktion G^{-1} (falls man sie berechnen kann) die explizite Form der Lösung

$$y(t) = G^{-1}\left(G(y_0) + \int_{x_0}^t p(x) dx\right).$$

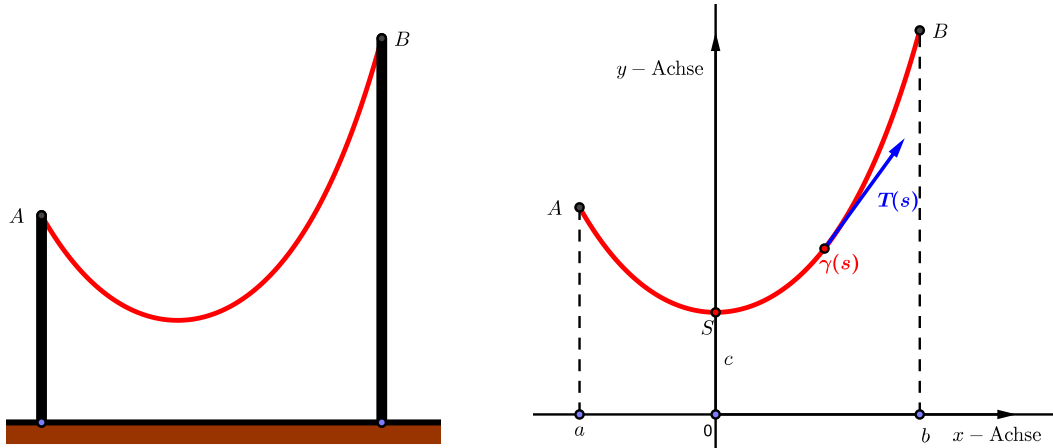
b) Ist keine Anfangsbedingung vorgegeben, so benutzt man bei der Integration die Stammfunktionen (unbestimmte Integration):

$$\begin{aligned} \stackrel{(6.24)}{\implies} \int \frac{1}{q(y)} dy \Big|_{y=y(x)} &= \int p(x) dx \\ \implies G(y(x)) &= \int p(x) dx + C, \end{aligned}$$

wobei G eine Stammfunktion von $\frac{1}{q}$ und C eine reelle Konstante ist. Auch diese Gleichung kann man ggf. nach $y(x)$ auflösen. Die Konstante C ist durch Festlegung eines Anfangswertes $y(x_0) = y_0$ zu bestimmen.

Beispiel: Die Kettenlinie (Seilkurve)

Wir wollen dieses Verfahren jetzt benutzen, um die Gleichung für die Kettenlinie herzuleiten. Wir hängen ein Seil (oder eine Kette) zwischen zwei Pfosten so hoch auf, dass es nicht auf der Erde schleift. Wie können wir die entstehende Form des Seils mathematisch beschreiben? Wir leiten eine Gleichung unter idealisierten Annahmen her (Seil ist vollkommen biegsam, das Gewicht des Seils pro Längeneinheit ist überall gleich).



Zur mathematischen Beschreibung legen wir ein Koordinatensystem so fest, dass das Seil in der xy -Ebene hängt, die x -Achse auf dem Erdboden liegt und die y -Achse durch den tiefsten Punkt S des Seils läuft. c sei die Höhe von S über dem Boden und a und b die x -Koordinate der Aufhängungspunkte A bzw. B . Wir wollen das Seil als Graph einer Funktion $f : [a, b] \rightarrow \mathbb{R}$ beschreiben.

Wir betrachten die rechte Hälfte des Seils. Das Seil habe zwischen S und B die Länge L . Jedem Punkt P auf der rechten Seilhälfte ordnen wir als Parameter die Länge s des Seils zwischen S und P zu, d.h. die rechte Seilhälfte ist beschrieben als Spur der Kurve $\gamma : [0, L] \rightarrow \mathbb{R}^2$, wobei das Seil vom Seilpunkt S bis zum Seilpunkt $\gamma(s)$ die Länge s hat. Auf den Seilpunkt $\gamma(s)$ wirkt eine Spannungskraft $T(s)$ tangential an die Kurve γ (durch die Aufhängung an den Pfosten). Gleichzeitig wirkt auf jeden Seilabschnitt die Gewichtskraft durch das Gewicht des Seils senkrecht nach unten. Wir nehmen (idealisiert) an, dass keine weiteren Kräfte auf das Seil wirken.

Wir leiten zunächst eine Formel für die Spannungskraft her. Dazu benutzen wir das folgende physikalische Gesetz: Das hängende Seil befindet sich im Gleichgewichtszustand, wenn sich in jedem beliebig kleinen Seilabschnitt die wirkenden Kräfte aufheben. D.h. für zwei beliebig nahe Seilpunkte $\gamma(s)$ und $\gamma(s+h)$ mit $h > 0$ gilt

$$T(s+h) - T(s) + G_s(h) = 0,$$

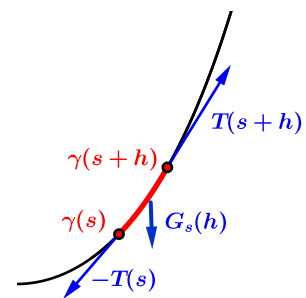
wobei $T(s+h)$ die Spannungskraft im Punkt $\gamma(s+h)$, $-T(s)$ die Gegenspannung im Punkt $\gamma(s)$ und $G_s(h)$ die Gewichtskraft ist, die durch das Gewicht des Seils zwischen $\gamma(s)$ und $\gamma(s+h)$ entsteht. Sei $g > 0$ das konstante Gewicht des Seils pro Längeneinheit.

Da h die Länge des Seils zwischen $\gamma(s)$ und $\gamma(s+h)$ ist, folgt für den Vektor der Gewichtskraft:

$$G_s(h) = (0, -h \cdot g).$$

Wir erhalten somit für den Spannungsvektor

$$\frac{T(s+h) - T(s)}{h} = (0, g)$$



und beim Grenzübergang $h \rightarrow 0$ für seine Ableitung

$$T'(s) = (0, g).$$

Bezeichnet $T(s) =: (T_1(s), T_2(s))$, so folgt für die Komponenten des Spannungsvektors:

$$T_1(s) = c_1 \quad \text{und} \quad T_2(s) = gs + c_2$$

wobei c_1 und c_2 reelle Konstanten sind. c_1 ist positiv (da wir auf der rechten Seilhälfte sind) und $c_2 = 0$, da im Punkt $S = \gamma(0)$ der Spannungsvektor parallel zur x -Achse ist, d.h. $T_2(0) = 0$ gilt. Wir erhalten für die Spannungskraft im Seilpunkt $\gamma(s)$ somit

$$T(s) = (c_1, sg) \quad \text{mit } c_1 \in \mathbb{R}^+.$$

Dies genügt nun, um die gesuchte Funktionsgleichung für das Seil herzuleiten. Wir suchen eine Funktion $f : [0, b] \rightarrow \mathbb{R}$, so dass das Seil zwischen S und B der Graph von f ist. Mit unserer Parametrisierung γ gilt für einen Seilpunkt $(x, f(x))$:

$$\gamma(s) = (x, f(x)), \quad \text{wobei } s \text{ die Länge des Seil von } S \text{ bis } (x, f(x)) \text{ ist.}$$

Der Tangentialvektor an den Graphen der Funktion f im Punkt $(x, f(x))$ ist gegeben durch $(1, f'(x))$, d.h. die Vektoren $T(s)$ und $(1, f'(x))$ sind parallel und zeigen in die gleiche Richtung. Folglich gilt für ihren Anstieg

$$\frac{f'(x)}{1} = \frac{T_2(s)}{T_1(s)} = \underbrace{\frac{g}{c_1}}_{=: k} \cdot s.$$

Die Länge s des Seils zwischen $S = (0, f(0))$ und $(x, f(x))$ berechnet sich in der Parametrisierung als Graph durch

$$s = \int_0^x \sqrt{1 + f'(t)^2} dt.$$

Daraus erhalten wir

$$f'(x) = k \cdot \int_0^x \sqrt{1 + f'(t)^2} dt.$$

und durch Ableiten nach x

$$f''(x) = k \cdot \sqrt{1 + f'(x)^2}.$$

Betrachtet man analog die linke Seilhälfte, so erhält man die gleiche Differentialgleichung für f auf dem Intervall $[a, 0]$.

Unsere physikalischen Annahmen führen also auf die folgende Differentialgleichung 2. Ordnung für die gesuchte Funktion $f : [a, b] \rightarrow \mathbb{R}$

$f''(x) = k \cdot \sqrt{1 + f'(x)^2}$	<i>Differentialgleichung der Kettenlinie (Seilkurve)</i>
---------------------------------------	----------------------------------------------------------

mit den Anfangsbedingungen $f(0) = c$ und $f'(0) = 0$, wobei $k > 0$ eine Materialkonstante des Seils und c die Höhe des tiefsten Punktes des Seils über dem Boden ist.

Lösung der Differentialgleichung der Kettenlinie

Wir setzen $z(x) := f'(x)$. Dann hat die Differentialgleichung der Kettenlinie, betrachtet als Differentialgleichung für z , die Form einer Differentialgleichung mit getrennten Variablen:

$$z' = k \cdot \sqrt{1 + z^2} \quad \text{mit der Anfangsbedingung } z(0) = 0.$$

Wir setzen $z' = \frac{dz}{dx}$ und trennen die Variablen:

$$\frac{1}{\sqrt{1 + z^2}} dz = k \cdot dx.$$

Nach Integration folgt

$$\int_0^{z(t)} \frac{1}{\sqrt{1 + z^2}} dz = \int_0^t k dx$$

und daraus mit der Stammfunktion arsinh von $\frac{1}{\sqrt{1+z^2}}$

$$\operatorname{arsinh}(z(t)) - \operatorname{arsinh}(0) = kt.$$

Durch Anwenden von \sinh folgt

$$f'(t) = z(t) = \sinh(kt) \quad \forall t \in \mathbb{R}.$$

f erhalten wir dann durch nochmaliges Integrieren:

$$f(x) = f(0) + \int_0^x f'(t) dt = c + \frac{1}{k} \cdot \cosh(kt) \Big|_{t=0}^{t=x},$$

d.h.

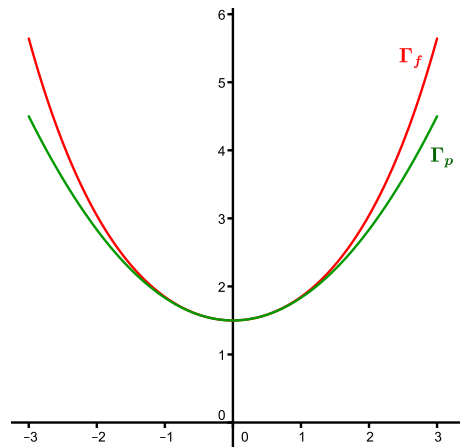
$$f(x) = c - \frac{1}{k} + \frac{1}{k} \cosh(kx) \quad \forall x \in \mathbb{R}$$

Legt man das Koordinatensystem jetzt noch so, dass $c = \frac{1}{k}$ gilt, so erhält man *in diesem* Koordinatensystem

$$f(x) = \frac{1}{k} \cosh(kx) \quad \forall x \in \mathbb{R}.$$

Die Reihenentwicklung von \cosh zeigt, inwieweit diese Funktion von einer Parabel, die ein Schüler vielleicht als Lösung vermuten würde, abweicht:

$$\begin{aligned}
 f(x) &= \frac{1}{k} \cosh(kx) = \frac{1}{k} \cdot \frac{e^{kx} + e^{-kx}}{2} \\
 &= \frac{1}{k} \cdot \sum_{n=0}^{\infty} \frac{(kx)^{2n}}{(2n)!} \\
 &= \underbrace{\frac{1}{k} + \frac{k}{2}x^2}_{=: p(x) \text{ Parabel}} + \frac{k^3}{24}x^4 + \dots
 \end{aligned}$$



Metrische Räume und stetige Abbildungen

In der bisherigen Vorlesung haben wir reelle Funktionen $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ und ihre Eigenschaften (Stetigkeit, Differenzierbarkeit, Integrierbarkeit) behandelt. Diese drei Eigenschaften haben wir mit Hilfe des Grenzwertbegriffes für reelle Zahlenfolgen definiert. Das zentrale Hilfsmittel zur Definition des Grenzwertes war der Abstand $|x - y|$ zwischen zwei reellen Zahlen x und y . Die bisher betrachtete Funktionenklasse reicht für die Beschreibung vieler Prozesse in der Natur nicht aus. Bereits wenn wir z.B. den Luftdruck oder die Temperatur in den Punkten der Erdoberfläche E beschreiben wollen, benötigen wir Funktionen vom Typ $f : E \rightarrow \mathbb{R}$, die von mehr als einer reellen Variablen abhängen. Wir werden deshalb in diesem Kapitel eine allgemeinere Klasse von Abbildungen $f : X \rightarrow Y$ betrachten und zunächst nur voraussetzen, dass im Definitionsbereich X und im Wertebereich Y Abstandsfunktionen gegeben sind, die die Definition von Grenzwerten ermöglichen. Dies genügt, um den Begriff der Stetigkeit für solche Abbildungen zu definieren. Wir werden sehen, dass sich zentrale Sätze über stetige reelle Funktionen $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$, die wir in Kapitel 4 behandelt haben, auf diesen allgemeineren Fall übertragen lassen.

7.1 Definition und Beispiele metrischer Räume

Definition 7.1. Sei X eine beliebige nichtleere Menge. Eine Abstandsfunktion (oder Metrik) auf X ist eine Abbildung $d : X \times X \rightarrow \mathbb{R}$ mit folgenden Eigenschaften

1. $d(p, q) \geq 0 \quad \forall p, q \in X$ und $d(p, q) = 0 \iff p = q$ (Positivität),
2. $d(p, q) = d(q, p) \quad \forall p, q \in X$ (Symmetrie),
3. $d(p, q) \leq d(p, r) + d(r, q) \quad \forall p, q, r \in X$ (Dreiecksungleichung).

Das Paar (X, d) heißt metrischer Raum. Die Elemente von X heißen Punkte des metrischen Raumes, die Zahl $d(p, q)$ nennt man den Abstand zwischen p und q .

Beispiel 1: Die Standardmetriken auf \mathbb{R} und \mathbb{C} .

Sei $|x|$ der Betrag einer reellen Zahl x . Dann ist $d : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ mit

$$d(x, y) := |x - y| \quad \text{für } x, y \in \mathbb{R},$$

eine Abstandsfunktion auf \mathbb{R} .

Analog erhält man aus dem Betrag von komplexen Zahlen eine Abstandsfunktion auf \mathbb{C} :

$$d(z, w) := |z - w| \quad \text{für } z, w \in \mathbb{C}.$$

Diese durch den Betrag definierte Abstandsfunktion d nennen wir die *Standardmetrik* auf \mathbb{R} bzw. \mathbb{C} . Ist nichts anderes vereinbart, so seien \mathbb{R} und \mathbb{C} immer mit dieser Abstandsfunktion versehen.

Beispiel 2: Auf einer Menge X können verschiedene Abstandsfunktionen existieren.

Die folgenden drei Abbildungen d_1, d_2, d_3 sind z.B. Abstandsfunktionen auf \mathbb{R}^2 :

- Der "Luftlinienabstand" (Standardmetrik oder Euklidischer Abstand):

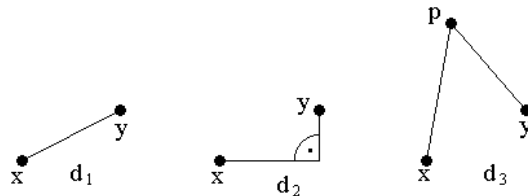
$$d_1(x, y) := \|x - y\| := \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}, \quad \text{wobei } x = (x_1, x_2), y = (y_1, y_2).$$

- Die "Taxifahrer-Metrik" (oder "Mannheimer-Metrik" – in Mannheim muß man rechtwinklige Straßen langlaufen, um von einem Punkt zum anderen zu kommen) :

$$d_2(x, y) := |x_1 - y_1| + |x_2 - y_2|.$$

- Die "Metrik der französischen Eisenbahn" (um von einer Stadt zur anderen zu kommen, muß man über Paris fahren): Sei p ein fixierter Punkt.

$$d_3(x, y) := \begin{cases} d_1(x, p) + d_1(p, y) & \text{falls } x \neq y, \\ 0 & \text{falls } x = y. \end{cases}$$



Beispiel 3: Auf jeder nichtleeren Menge X existiert eine Abstandsfunktion.

Wir definieren $d: X \times X \rightarrow \mathbb{R}$ durch

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

Dann ist (X, d) ein metrischer Raum. (X, d) heißt *diskreter metrischer Raum* und d die *diskrete Metrik*.

Beispiel 4: *Das kartesische Produkt metrischer Räume.*

Seien $(X_1, d_1), \dots, (X_n, d_n)$ metrische Räume. Wir betrachten die Produktmenge

$$X := X_1 \times X_2 \times \dots \times X_n := \{(p_1, \dots, p_n) \mid p_i \in X_i, i \in \{1, \dots, n\}\}$$

und definieren die *Produktmetrik* auf X durch

$$d((p_1, \dots, p_n), (q_1, \dots, q_n)) := \sqrt{d_1(p_1, q_1)^2 + \dots + d_n(p_n, q_n)^2} = \sqrt{\sum_{j=1}^n d_j(p_j, q_j)^2}.$$

Dann ist (X, d) ein metrischer Raum. Er heißt das *kartesische Produkt* der metrischen Räume $(X_1, d_1), \dots, (X_n, d_n)$.

Beweis. Die Positivität und Symmetrie sind aus der Definition sofort ersichtlich. Wir zeigen die Dreiecksungleichung. Seien dazu $p = (p_1, \dots, p_n)$, $q = (q_1, \dots, q_n)$ und $r = (r_1, \dots, r_n)$ drei beliebige Punkte in X . Unter Benutzung der Dreiecksungleichung für die Metriken d_j und der Cauchy-Schwarzschen Ungleichung (CSU) für das Euklidische Skalarprodukt (siehe Kapitel 6.4 oder Vorlesung *Lineare Algebra I*) erhalten wir

$$\begin{aligned} d(p, q)^2 &= \sum_{j=1}^n d_j(p_j, q_j)^2 \\ &\stackrel{\Delta}{\leq} \sum_{j=1}^n \left(d_j(p_j, r_j) + d_j(r_j, q_j) \right)^2 \\ &= \sum_{j=1}^n d_j(p_j, r_j)^2 + 2d_j(p_j, r_j) \cdot d_j(r_j, q_j) + d_j(r_j, q_j)^2 \\ &= d(p, r)^2 + d(r, q)^2 + 2 \sum_{j=1}^n d_j(p_j, r_j) \cdot d_j(r_j, q_j) \\ &\stackrel{\text{CSU}}{\leq} d(p, r)^2 + d(r, q)^2 + 2 \sqrt{\sum_{j=1}^n d_j(p_j, r_j)^2} \cdot \sqrt{\sum_{j=1}^n d_j(r_j, q_j)^2} \\ &= d(p, r)^2 + d(r, q)^2 + 2 d(p, r) \cdot d(r, q) \\ &= (d(p, r) + d(r, q))^2 \end{aligned}$$

Da der Abstand nicht-negativ ist, können wir in der Ungleichung auf beiden Seiten die Wurzel ziehen und erhalten die Dreiecksungleichung für d . \square

Beispiel 5: *Zwei Abstandsfunktionen auf der Menge der stetigen Funktionen $C^0([a, b], \mathbb{R})$.*

Sei $C^0([a, b], \mathbb{R})$ der Vektorraum der stetigen Funktionen von $[a, b]$ in \mathbb{R} . Dann definieren

$$\begin{aligned} d_\infty(f, g) &:= \max\{|f(x) - g(x)| \mid x \in [a, b]\}, & f, g \in C^0([a, b], \mathbb{R}), \\ d_1(f, g) &:= \int_a^b |f(x) - g(x)| dx, \end{aligned}$$

Abstandsfunktionen auf $C^0([a, b], \mathbb{R})$ (Übungsaufgabe).

Beispiel 6: *Metriken auf normierten Vektorräumen.*

Sei V ein reeller Vektorraum. Eine *Norm auf V* ist eine Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}$ mit

1. $\|v\| \geq 0$ für alle $v \in V$ und $\|v\| = 0 \iff v = 0$ (Positivität).
2. $\|\lambda \cdot v\| = |\lambda| \cdot \|v\|$ für alle $v \in V$ und $\lambda \in \mathbb{R}$ (Homogenität).
3. $\|v + w\| \leq \|v\| + \|w\|$ für alle $v, w \in V$ (Dreiecksungleichung).

Das Paar $(V, \|\cdot\|)$ heißt *normierter Vektorraum*. Dann ist $d : V \times V \rightarrow \mathbb{R}$, definiert durch

$$d(v, w) := \|v - w\| \quad v, w \in V,$$

eine Abstandsfunktion auf V . Sie heißt die *durch $\|\cdot\|$ induzierte Metrik*. Ein normierter Vektorraum $(V, \|\cdot\|)$ ist immer mit dieser Metrik versehen (falls nicht ausdrücklich anders gesagt).

Beweis. Die Positivität und Symmetrie von d folgt aus der 1. und 2. Eigenschaft für die Norm. Die Dreiecksungleichung für d erhält man aus der Dreiecksungleichung für die Norm. Für $v, w, u \in V$ gilt nämlich

$$d(v, u) = \|v - u\| = \|(v - w) + (w - u)\| \leq \|v - w\| + \|w - u\| = d(v, w) + d(w, u). \quad \square$$

Dies verallgemeinert die Situation in Beispiel 1. Dort ist $V = \mathbb{R}$ und $\|\cdot\| = |\cdot|$.

In der analytischen Geometrie wird die *Euklidische Norm* auf dem \mathbb{R}^n benutzt

$$\|v\| := \sqrt{v_1^2 + v_2^2 + \dots + v_n^2} \quad \text{für } v = (v_1, v_2, \dots, v_n) \in \mathbb{R}^n,$$

und der davon induzierte *Euklidische Abstand* zwischen zwei Punkten $p = (p_1, \dots, p_n)$ und $q = (q_1, \dots, q_n)$:

$$d(p, q) := \|p - q\| = \sqrt{(p_1 - q_1)^2 + \dots + (p_n - q_n)^2}.$$

(siehe Vorlesung *Lineare Algebra I* und Kapitel 6.4. dieser Vorlesung). Ist der \mathbb{R}^n mit diesem Abstand versehen, so nennt man ihn den *Euklidischen Raum*. Er ist das n -fache kartesische Produkt von \mathbb{R} mit der Standardmetrik. In der Vorlesung *Elementargeometrie* des Lehramtsstudiums werden Sie andere Geometrien kennenlernen (hyperbolische Geometrie oder sphärische Geometrie), in denen die Abstände zweier Punkte mit anderen Abstandsfunktionen gemessen werden.

Definition 7.2. *Zwei metrische Räume (X, d_X) und (Y, d_Y) heißen isometrisch, wenn eine bijektive Abbildung $\phi : X \rightarrow Y$ existiert, so dass*

$$d_X(p, q) = d_Y(\phi(p), \phi(q)) \quad \text{für alle } p, q \in X.$$

Die Abbildung ϕ heißt dann *Isometrie* zwischen (X, d_X) und (Y, d_Y) .

In isometrischen metrischen Räumen stimmen alle Eigenschaften, die mittels der Abstandsfunktion charakterisiert werden, überein.

Für spätere Zwecke definieren wir das Analogon von offenen Intervallen im \mathbb{R} , offenen Kreisscheiben im \mathbb{R}^2 (bzw. in \mathbb{C}) und offenen Kugeln im Euklidischen Raum \mathbb{R}^3 für beliebige metrische Räume:

Definition 7.3. Sei (X, d) ein metrischer Raum, $p \in X$ und $\varepsilon \in \mathbb{R}^+$. Die Menge

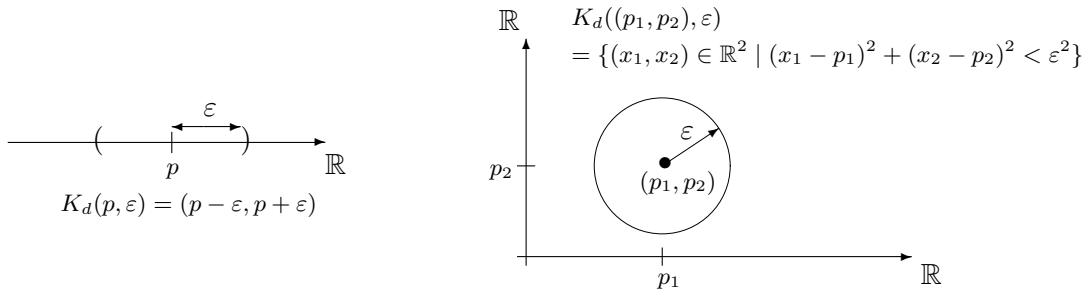
$$K(p, \varepsilon) := \{x \in X \mid d(p, x) < \varepsilon\}$$

heißt ε -Kugel in (X, d) um p .

Beispiel 1: Sei \mathbb{R}^n der Euklidische Raum. Dann ist

$$K(p, \varepsilon) = \{x \in \mathbb{R}^n \mid \|x - p\| < \varepsilon\} = \{x \in \mathbb{R}^n \mid (x_1 - p_1)^2 + \dots + (x_n - p_n)^2 < \varepsilon^2\}$$

die übliche n -dimensionale Kugel vom Radius ε um den Punkt $p = (p_1, \dots, p_n) \in \mathbb{R}^n$ ohne ihren Rand. Für $n = 1$ sind die ε -Kugeln offene Intervalle, für $n = 2$ Kreisscheiben ohne ihren Rand.

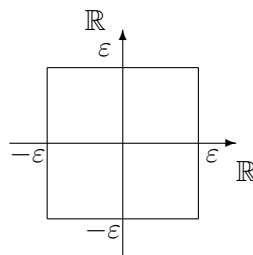


Beispiel 2: Sei \mathbb{R}^2 mit folgender Metrik versehen:

$$d_\infty((x_1, x_2), (y_1, y_2)) = \max(|x_1 - y_1|, |x_2 - y_2|).$$

(\mathbb{R}^2, d_∞) ist tatsächlich ein metrischer Raum (Übungsaufgabe). Die ε -Kugel um den Nullpunkt ist für diese Metrik ein Quadrat mit der Seitenlänge 2ε ohne seinen Rand:

$$K(0, \varepsilon) = \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| < \varepsilon, |x_2| < \varepsilon\}.$$



Beispiel 3: Sei (X, d) ein diskreter metrischer Raum. Dann gilt für die ε -Kugeln

$$K(x, \varepsilon) = \begin{cases} \{x\} & \text{falls } \varepsilon \leq 1 \\ X & \text{falls } \varepsilon > 1. \end{cases}$$

7.2 Konvergente Folgen und Cauchy-Folgen in metrischen Räumen

Definition 7.4. Sei (X, d) ein metrischer Raum und (x_n) eine Folge in X .

Wir sagen, dass (x_n) gegen $p \in X$ konvergiert, falls zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass $d(p, x_n) < \varepsilon$ für alle $n \geq n_0$, d.h. falls $x_n \in K(p, \varepsilon)$ für alle $n \geq n_0$.

Der Punkt p heißt Grenzwert (GW) der Folge (x_n) . Besitzt eine Folge (x_n) einen Grenzwert, so heißt sie konvergent. Besitzt die Folge (x_n) keinen Grenzwert, so heißt sie divergent.

Für eine gegen p konvergente Folge (x_n) schreiben wir:

$$\lim_{n \rightarrow \infty} x_n = p \quad \text{oder} \quad x_n \xrightarrow[n \rightarrow \infty]{} p \quad \text{oder kurz} \quad x_n \longrightarrow p.$$

Aus der Definition folgt unmittelbar:

Satz 7.1 Eine Folge (x_n) in einem metrischen Raum (X, d) konvergiert genau dann gegen $p \in X$, wenn die Folge der Abstände $(d(p, x_n))$ in \mathbb{R} gegen Null konvergiert:

$$x_n \longrightarrow p \text{ in } (X, d) \iff d(p, x_n) \longrightarrow 0 \text{ in } \mathbb{R}.$$

Satz 7.2 Der Grenzwert einer konvergenten Folge eines metrischen Raumes ist eindeutig bestimmt.

Beweis. Der Beweis wird analog zum Beweis des entsprechenden Satzes für konvergente Folgen in \mathbb{R} geführt (Satz 3.1). Wir überlassen das Übertragen als Übungsaufgabe. \square

Definition 7.5. Sei (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt beschränkt, falls es eine Kugel $K(p, M)$ des metrischen Raumes gibt, die A enthält. Eine Folge (x_n) in (X, d) heißt beschränkt, wenn die Menge der Folgenglieder $\{x_1, x_2, x_3, \dots\}$ beschränkt ist.

Satz 7.3 Jede konvergente Folge (x_n) eines metrischen Raumes (X, d) ist beschränkt.

Beweis. Sei (x_n) eine konvergente Folge in (X, d) und $p \in X$ ihr Grenzwert. Wir betrachten die Kugel $K(p, 1)$ um p vom Radius 1. Dann existiert ein $n_0 \in \mathbb{N}$, so dass $x_n \in K(p, 1)$ für alle $n \geq n_0$. Wir setzen nun

$$M := \max(d(p, x_1), d(p, x_2), \dots, d(p, x_{n_0-1})).$$

Dann gilt $x_n \in K(p, M + 1)$ für alle $n \in \mathbb{N}$. Folglich ist (x_n) beschränkt. \square

Als nächstes charakterisieren wir die Konvergenz von Folgen im kartesischen Produkt metrischer Räume.

Satz 7.4 Seien $(X_1, d_1), \dots, (X_k, d_k)$ metrische Räume und (X, d) das kartesische Produkt dieser Räume. Eine Folge (x_n) in X mit $x_n = (x_{n1}, x_{n2}, \dots, x_{nk})$ konvergiert genau dann gegen $p = (p_1, p_2, \dots, p_k) \in X$, wenn für jedes $j \in \{1, \dots, k\}$ die Komponentenfolge $(x_{nj})_{n=1}^\infty$ in (X_j, d_j) gegen p_j konvergiert.

Beweis. Nach Definition der Produktmetrik ist

$$d(\underbrace{(p_1, \dots, p_k)}_p, \underbrace{(x_{n1}, \dots, x_{nk})}_{x_n}) = \sqrt{\sum_{i=1}^k d_i(p_i, x_{ni})^2} \geq d_j(p_j, x_{nj}) \quad (*)$$

für jedes $j \in \{1, \dots, k\}$. Sei nun (x_n) gegen p in (X, d) konvergent. Dann existiert für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ mit $d(p, x_n) < \varepsilon$ für alle $n \geq n_0$. Nach Abschätzung (*) folgt daraus $d_j(p_j, x_{nj}) < \varepsilon$ für alle $n \geq n_0$ und jedes $j \in \{1, \dots, k\}$. Somit konvergiert (x_{nj}) gegen p_j in (X_j, d_j) für jedes $j \in \{1, \dots, k\}$.

Sei umgekehrt $\lim_{n \rightarrow \infty} x_{nj} = p_j$ für jedes $j \in \{1, \dots, k\}$ und $\varepsilon > 0$. Dann existieren $n_{0j} \in \mathbb{N}$ so dass $d_j(p_j, x_{nj}) < \frac{\varepsilon}{\sqrt{k}}$ für alle $n \geq n_{0j}$. Es folgt

$$d(p, x_n) = \sqrt{\sum_{j=1}^k d_j(p_j, x_{nj})^2} < \sqrt{\sum_{j=1}^k \frac{\varepsilon^2}{k}} = \varepsilon$$

für alle $n \geq m_0 = \max(n_{01}, \dots, n_{0k})$. Also konvergiert x_n gegen p im Produktraum (X, d) . \square

Wir betrachten als Beispiel den Euklidischen Raum \mathbb{R}^k . Er ist das k -fache kartesische Produkt von \mathbb{R} mit der Standardmetrik. Die komponentenweise Konvergenz aus Satz 7.4 entspricht der Konvergenz von Folgen im Euklidischen Raum, die wir in Abschnitt 6.4 bereits benutzt hatten. Als Konsequenz erhält man aus den Grenzwertsätzen für Zahlenfolgen, dass auch die algebraischen Strukturen auf dem Euklidischen Vektorraum \mathbb{R}^k mit der Grenzwertbildung verträglich sind:

Satz 7.5 Seien (x_n) und (y_n) zwei Folgen im Euklidischen Raum \mathbb{R}^k , die gegen p bzw. q konvergieren, $\langle \cdot, \cdot \rangle$ das Euklidische Skalarprodukt und $\|\cdot\|$ die Euklidische Norm. Dann gilt:

1. $\lim_{n \rightarrow \infty} (\lambda \cdot x_n + \mu \cdot y_n) = \lambda \cdot \lim_{n \rightarrow \infty} x_n + \mu \cdot \lim_{n \rightarrow \infty} y_n = \lambda \cdot p + \mu \cdot q$ für alle $\lambda, \mu \in \mathbb{R}$.
2. $\lim_{n \rightarrow \infty} \langle x_n, y_n \rangle = \langle \lim_{n \rightarrow \infty} x_n, \lim_{n \rightarrow \infty} y_n \rangle = \langle p, q \rangle$.
3. $\lim_{n \rightarrow \infty} \|x_n\| = \|\lim_{n \rightarrow \infty} x_n\| = \|p\|$.

Wie bei Zahlenfolgen möchte man auch für Folgen in metrischen Räumen die Konvergenz untersuchen, ohne den Grenzwert bereits zu kennen. Dies geht in den *vollständigen* metrischen Räumen, die wir jetzt definieren werden.

Definition 7.6. Sei (X, d) ein metrischer Raum. Eine Folge (x_n) in (X, d) heißt *Cauchy-Folge*, wenn zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass $d(x_n, x_m) < \varepsilon$ für alle $n, m \geq n_0$.

Satz 7.6 Jede konvergente Folge in einem metrischen Raum ist eine Cauchy-Folge.

Beweis. Sei (x_n) eine gegen p konvergente Folge und $\varepsilon > 0$. Dann existiert ein $n_0 \in \mathbb{N}$ mit $d(p, x_n) < \frac{\varepsilon}{2}$ für alle $n \geq n_0$. Für $n, m \geq n_0$ folgt aus der Dreiecksungleichung

$$d(x_n, x_m) \leq d(x_n, p) + d(p, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Somit ist (x_n) eine Cauchy-Folge. □

Wir wissen bereits aus Kapitel 3.1.3, dass die Umkehrung von Satz 7.6 nicht gilt. Es gibt metrische Räume mit Cauchy-Folgen, die nicht konvergieren.

Definition 7.7. Ein metrischer Raum (X, d) heißt *vollständig*, wenn jede Cauchy-Folge in (X, d) konvergiert.

In einem *vollständigen* metrischen Raum kann man also die Konvergenz einer Folge untersuchen, ohne ihren Grenzwert zu kennen. Man prüft dafür die *Cauchy-Bedingung*

$$\boxed{\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \text{ mit } d(x_n, x_m) < \varepsilon \forall n, m \geq n_0.}$$

Beispiel 1: Die reellen Zahlen \mathbb{R} mit der Standardmetrik $d(x, y) = |x - y|$, sowie die komplexen Zahlen \mathbb{C} mit der Standardmetrik $d(z, w) := |z - w|$ sind vollständige metrische Räume (siehe Satz 3.14).

Beispiel 2: Der metrische Raum der rationalen Zahlen \mathbb{Q} mit der Standardmetrik $d(x, y) = |x - y|$ ist *nicht* vollständig (siehe Kapitel 3.1.3).

Beispiel 3: Wir betrachten $X = (-\frac{\pi}{2}, \frac{\pi}{2}) \subset \mathbb{R}$ mit den Metriken d und \widehat{d} :

$$\begin{aligned} d(x, y) &= |x - y|, \\ \widehat{d}(x, y) &= |\tan(x) - \tan(y)|. \end{aligned}$$

Dann ist (X, d) nicht vollständig, aber (X, \widehat{d}) vollständig (Übungsaufgabe).

Beispiel 4: Sei $X = C^0([a, b], \mathbb{R})$ der Vektorraum der stetigen Funktionen von $[a, b]$ in \mathbb{R} mit den Metriken d_∞ und d_1 :

$$\begin{aligned} d_\infty(f, g) &= \max\{|f(x) - g(x)| \mid x \in [a, b]\}, \\ d_1(f, g) &= \int_a^b |f(x) - g(x)| dx. \end{aligned}$$

Dann ist (X, d_∞) vollständig, aber (X, d_1) nicht vollständig (Übungsaufgabe).

Die Vollständigkeit eines metrischen Raumes (X, d) ist also von der auf X gewählten Metrik abhängig.

Satz 7.7 Sind $(X_1, d_1), \dots, (X_k, d_k)$ vollständige metrische Räume, so ist auch das kartesische Produkt (X, d) dieser metrischen Räume vollständig.

Beweis. Sei (x_n) mit $x_n = (x_{n1}, \dots, x_{nk})$, eine Cauchy-Folge im kartesischen Produkt $X = X_1 \times \dots \times X_k$ und $\varepsilon > 0$. Dann existiert ein $n_0 \in \mathbb{N}$, so dass

$$d(x_n, x_m) = \sqrt{\sum_{i=1}^k d_i(x_{ni}, x_{mi})^2} < \varepsilon \quad \forall n, m \geq n_0.$$

Daraus erhalten wir $d_j(x_{nj}, x_{mj}) < \varepsilon$ für alle $n, m \geq n_0$ und jedes $j \in \{1, \dots, k\}$. Folglich sind die Folgen $(x_{nj})_{n=1}^\infty$ Cauchy-Folgen in (X_j, d_j) für jedes $j \in \{1, \dots, k\}$. Da (X_j, d_j) vollständige metrische Räume sind, konvergieren die Folgen $(x_{nj})_{n=1}^\infty$ gegen ein $p_j \in X_j$. Somit gilt nach Satz 7.4 $\lim_{n \rightarrow \infty} x_n = (p_1, \dots, p_k) \in X_1 \times \dots \times X_k$. □

Folgerung 7.1 Der Euklidische Raum \mathbb{R}^k ist vollständig.

Bemerkung: Am Ende von Kapitel 7.4 werden wir zeigen, dass jede Norm auf dem Vektorraum \mathbb{R}^k eine vollständige Metrik auf \mathbb{R}^k induziert. Insbesondere sind die normierten Vektorräume $(\mathbb{R}^k, \|\cdot\|_p)$ für $p \in [1, \infty)$ und $(\mathbb{R}^k, \|\cdot\|_\infty)$ mit

$$\|x\|_p := \left(\sum_{j=1}^k |x_j|^p \right)^{\frac{1}{p}},$$

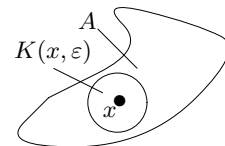
$$\|x\|_\infty := \max\{|x_1|, \dots, |x_k|\},$$

vollständig (siehe auch Übungsaufgabe 18). $\|\cdot\|_2$ ist die Euklidische Norm.

7.3 Spezielle Teilmengen in metrischen Räumen

In diesem Abschnitt werden wir spezielle Sorten von Teilmengen metrischer Räume definieren, die die von uns bisher im "naiven" Sinne benutzten Bezeichnungen *offene* und *abgeschlossene* Intervalle, Randpunkte von Intervallen bzw. Rand von Gebieten in der Ebene verallgemeinern.

Definition 7.8. Eine Teilmenge A eines metrischen Raumes (X, d) heißt *offen*, wenn zu jedem $p \in A$ eine ε -Kugel $K(p, \varepsilon)$ existiert mit $K(p, \varepsilon) \subset A$.



Beispiel 1: Die ε -Kugeln $K(p, \varepsilon)$ eines metrischen Raumes (X, d) sind offen.

Um dies zu zeigen, betrachten wir einen beliebigen Punkt $x \in K(p, \varepsilon)$ und zeigen, dass $K(x, \varepsilon - d(p, x)) \subset K(p, \varepsilon)$: Sei $z \in K(x, \varepsilon - d(p, x))$. Dann gilt $d(x, z) < \varepsilon - d(p, x)$ und folglich $d(z, x) + d(x, p) < \varepsilon$. Aus der Dreiecksungleichung folgt dann $d(p, z) \leq d(p, x) + d(x, z) < \varepsilon$. Folglich liegt z in $K(p, \varepsilon)$.

Insbesondere sind die Intervalle (a, b) im metrischen Raum \mathbb{R} (mit der Standardmetrik) in diesem Sinne offen. Die Intervalle der Form $(-\infty, a)$ und (a, ∞) sind ebenfalls offen.

Definition 7.9. Sei (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt abgeschlossen, wenn ihr Komplement $X \setminus A$ offen ist.

Beispiel 2: Es gilt $[a, b] = \mathbb{R} \setminus ((-\infty, a) \cup (b, \infty))$, $[a, \infty) = \mathbb{R} \setminus (-\infty, a)$ und $(-\infty, a] = \mathbb{R} \setminus (a, \infty)$. Folglich sind die Intervalle $[a, b]$, $[a, \infty)$, $(-\infty, a]$ abgeschlossen.

Beispiel 3: Sei (X, d) der diskrete metrische Raum.

Dann ist jede Teilmenge $A \subset X$ offen, denn für jedes $p \in A$ gilt $K(p, \frac{1}{2}) = \{p\} \subset A$. Folglich ist auch jede Teilmenge von X abgeschlossen.

Die Eigenschaft einer Menge, abgeschlossen zu sein, kann man auch durch konvergente Folgen charakterisieren. In Kapitel 4.3. hatten wir abgeschlossene Teilmengen in \mathbb{R} und \mathbb{C} auf diese Weise definiert.

Satz 7.8 Sei (X, d) ein metrischer Raum. Dann gilt:

Eine Teilmenge $A \subset X$ ist genau dann abgeschlossen, wenn der Grenzwert jeder konvergenten Folge von Punkten aus A selbst in A liegt.

Beweis. (\implies) Sei $A \subset X$ abgeschlossen und (a_n) eine Folge in A , die gegen $x \in X$ konvergiert. Angenommen $x \notin A$. Dann ist $x \in X \setminus A$. Da $X \setminus A$ offen ist, existiert eine ε -Kugel $K(x, \varepsilon) \subset X \setminus A$. Da (a_n) gegen x konvergiert, existiert ein $n_0 \in \mathbb{N}$, so dass $a_n \in K(x, \varepsilon)$ für alle $n \geq n_0$. Dann können diese a_n nicht in A liegen (Widerspruch).

(\impliedby) Angenommen A ist nicht abgeschlossen, d.h. $X \setminus A$ nicht offen. Dann existiert ein $x \in X \setminus A$, so dass $K(x, \varepsilon) \cap A \neq \emptyset$ für alle $\varepsilon > 0$. Wir wählen nun $a_n \in K(x, \frac{1}{n}) \cap A$. Dann ist (a_n) eine Folge in A mit $d(x, a_n) \leq \frac{1}{n} \rightarrow 0$. Folglich konvergiert (a_n) gegen x . Nach Voraussetzung liegt aber x nicht in A (Widerspruch). \square

Man kann jede Teilmenge $A \subset X$ in eine offene und eine abgeschlossene Menge einschließen. Dazu definieren wir das Innere $\text{Int}(A)$ und den Abschluß $\text{cl}(A)$ von A , so dass

$$\text{Int}(A) \subset A \subset \text{cl}(A).$$

Definition 7.10. Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge von X .

1. Das Innere von A ist die Menge

$$\text{Int}(A) := \{x \in A \mid \exists \varepsilon > 0 \text{ mit } K(x, \varepsilon) \subset A\}.$$

Ein Punkt $x \in \text{Int}(A)$ heißt innerer Punkt von A .

2. Der Abschluß von A ist die Menge

$$\text{cl}(A) := \{x \in X \mid \forall \varepsilon > 0 \text{ gilt } K(x, \varepsilon) \cap A \neq \emptyset\}.$$

3. Der Rand von A ist die Menge

$$\partial A := \text{cl}(A) \setminus \text{Int}(A).$$

Ein Punkt $x \in \partial A$ heißt Randpunkt von A .

Satz 7.9 Für jede Teilmenge A eines metrischen Raumes gilt:

1. $\text{Int}(A)$ ist offen.
2. $\text{cl}(A) = X \setminus \text{Int}(X \setminus A)$. Insbesondere ist $\text{cl}(A)$ abgeschlossen.
3. $\partial A = X \setminus \left(\text{Int}(A) \cup \text{Int}(X \setminus A) \right)$. Insbesondere ist ∂A abgeschlossen.
4. Ein Punkt $x \in X$ ist genau dann ein Randpunkt von A , wenn für alle $\varepsilon > 0$ gilt

$$K(x, \varepsilon) \cap A \neq \emptyset \text{ und } K(x, \varepsilon) \cap (X \setminus A) \neq \emptyset.$$

Beweis. Zu 1: Sei $x \in \text{Int}(A)$. Dann existiert eine ε -Kugel $K(x, \varepsilon) \subset A$. Für jeden Punkt $z \in K(x, \varepsilon)$ gilt $K(z, \varepsilon - d(x, z)) \subset K(x, \varepsilon) \subset A$. Folglich ist jeder Punkt von $K(x, \varepsilon)$ ein innerer Punkt von A , d.h. $K(x, \varepsilon) \subset \text{Int}(A)$. Somit ist $\text{Int}(A)$ offen.

Zu 2: Es gilt:

$$\begin{aligned} x \in X \setminus \text{Int}(X \setminus A) &\iff x \in X \text{ und } x \notin \text{Int}(X \setminus A) \\ &\iff \forall \varepsilon > 0 \text{ gilt } K(x, \varepsilon) \not\subset X \setminus A \\ &\iff \forall \varepsilon > 0 \text{ gilt } K(x, \varepsilon) \cap A \neq \emptyset \\ &\iff x \in \text{cl}(A). \end{aligned}$$

Zu 3: Aus $\text{cl}(A) = X \setminus \text{Int}(X \setminus A)$ folgt

$$\partial A = \text{cl}(A) \setminus \text{Int}(A) = (X \setminus \text{Int}(X \setminus A)) \setminus \text{Int}(A) = X \setminus \underbrace{(\text{Int}(X \setminus A) \cup \text{Int}(A))}_{\text{offen}}.$$

Zu 4: Daraus erhalten wir

$$\begin{aligned} x \in \partial A &\iff x \notin \text{Int}(A) \text{ und } x \notin \text{Int}(X \setminus A) \\ &\iff \forall \varepsilon > 0 \text{ gilt } K(x, \varepsilon) \cap (X \setminus A) \neq \emptyset \text{ und } K(x, \varepsilon) \cap A \neq \emptyset. \end{aligned}$$

□

Beispiel 4: Wir betrachten \mathbb{R} mit der Standardmetrik $d(x, y) = |x - y|$. Für $A = [a, b)$ erhalten wir $\text{Int}(A) = (a, b)$, $\text{cl}(A) = [a, b]$ und $\partial A = \{a, b\}$.

Beispiel 5: Sei \mathbb{R}^2 der Euklidische Raum mit der Euklidischen Norm $\|\cdot\|$. Für $A = K(x, \varepsilon)$ gilt:

- $\text{Int}(A) = K(x, \varepsilon)$,
- $\partial A = \{y \in \mathbb{R}^2 \mid \|x - y\| = \varepsilon\}$,
- $\text{cl}(A) = \{y \in \mathbb{R}^2 \mid \|x - y\| \leq \varepsilon\}$.

Als nächstes definieren wir die kompakten Teilmengen eines metrischen Raumes, die wir für die metrischen Räume \mathbb{R} und \mathbb{C} bereits aus Kapitel 4.3 kennen.

Ist (x_n) eine Folge in (X, d) . Eine Folge der Form (x_{n_k}) mit $n_1 < n_2 < n_3 < \dots$ heißt *Teilfolge von (x_n)* .

Definition 7.11. Eine Teilmenge A eines metrischen Raumes (X, d) heißt *kompakt*, falls jede Folge in A eine konvergente Teilfolge besitzt, deren Grenzwert in A liegt¹.

Beispiel 6: Eine Teilmenge $A \subset \mathbb{R}$ bzw. $B \subset \mathbb{C}$ ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist (siehe Satz 4.11).

Satz 7.10 Jede kompakte Teilmenge eines metrischen Raumes ist abgeschlossen und beschränkt.

Beweis. Sei A eine kompakte Teilmenge des metrischen Raumes (X, d) .

1. Wir zeigen mit dem Folgenkriterium aus Satz 7.8, dass A abgeschlossen ist:

Sei (a_n) eine Folge in A , die gegen $x \in X$ konvergiert. Wir müssen zeigen, dass x in A liegt. Da A kompakt ist, besitzt (a_n) eine Teilfolge (a_{n_k}) , die gegen einen Punkt $a \in A$ konvergiert. Es gilt aber $a = \lim_{k \rightarrow \infty} a_{n_k} = \lim_{n \rightarrow \infty} a_n = x$. Folglich ist $x = a \in A$.

2. Wir zeigen, dass A beschränkt ist:

Sei $p \in X$. Angenommen, A wäre nicht beschränkt. Dann gilt $A \not\subset K(p, n)$ für jedes $n \in \mathbb{N}$. Wir wählen jeweils einen Punkt $a_n \in A \setminus K(p, n)$. Dann ist (a_n) eine Folge in A mit $d(a_n, p) \geq n$. Da A kompakt ist, besitzt (a_n) eine Teilfolge (a_{n_k}) , die gegen ein $a \in A$ konvergiert. Dann existiert ein $k_0 \in \mathbb{N}$, so dass $d(a, a_{n_k}) < 1$ für alle $k \geq k_0$. Daraus folgt für $k \geq k_0$:

$$d(p, a_{n_k}) \leq d(p, a) + d(a, a_{n_k}) < d(p, a) + 1.$$

Dies ist ein Widerspruch, da $d(p, a_{n_k}) \geq n_k \rightarrow \infty$. □

Satz 7.11 Sei (X, d) ein metrischer Raum und $B \subset X$ kompakt. Dann ist jede abgeschlossene Teilmenge $A \subset B$ ebenfalls kompakt.

Beweis. Sei (a_n) eine beliebige Folge in A . Dann ist (a_n) auch Folge in B und besitzt, da B kompakt ist, eine in B konvergente Teilfolge (a_{n_j}) . Sei $b = \lim_{j \rightarrow \infty} a_{n_j}$. Da A abgeschlossen ist, liegt der Grenzwert b in A . Also enthält (a_n) eine konvergente Teilfolge mit Grenzwert in A . Damit ist A kompakt. □

Satz 7.12 Sei (X, d) das kartesische Produkt der metrischen Räume $(X_1, d_1), \dots, (X_k, d_k)$ und seien $A_j \subset X_j$ kompakte Mengen in (X_j, d_j) , $j = 1, \dots, k$. Dann ist die Menge $A := A_1 \times A_2 \times \dots \times A_k$ ebenfalls kompakt in (X, d) .

Beweis. Den Beweis wird analog zum Beweis von Satz 7.7 geführt. Wir überlassen ihn deshalb als Übungsaufgabe. □

Beispiel 7: Die Quader $W := [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_k, b_k] \subset \mathbb{R}^k$ sind kompakte Teilmengen des Euklidischen Raumes \mathbb{R}^k .

Wir erhalten daraus folgende Verallgemeinerung von Satz 4.11:

¹ Man kann kompakte Mengen äquivalent auch durch Eigenschaften offener Überdeckungen definieren. Wir wollen dies hier aus Zeitgründen nicht tun, sondern verweisen dazu auf O. Deiser. Analysis 2.

Satz 7.13 Eine Teilmenge des Euklidischen Raumes \mathbb{R}^k ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.

Beweis. Nach Satz 7.10 ist jede kompakte Menge abgeschlossen und beschränkt. Wir müssen die Umkehrung zeigen. Sei $A \subset \mathbb{R}^k$ beschränkt und abgeschlossen. Da A beschränkt ist, gibt es eine Kugel, die A enthält. Jede Kugel des Euklidischen Raumes ist in einem Quader W enthalten. Folglich gilt $A \subset W$. Da W kompakt und A abgeschlossen ist, ist A nach Satz 7.11 kompakt. \square

Beispiel 8: Die Sphäre $S_r^{n-1} := \{x \in \mathbb{R}^n \mid \|x\| = r\} \subset \mathbb{R}^n$ und die abgeschlossene Kugel $D_r^n := \{x \in \mathbb{R}^n \mid \|x\| \leq r\} \subset \mathbb{R}^n$ sind kompakt.

Beispiel 9: Es gibt metrische Räume mit abgeschlossenen und beschränkten Teilmengen, die *nicht* kompakt sind.

Sei z.B. X eine unendliche Menge mit der diskreten Metrik

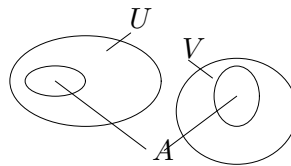
$$d(x, y) = \begin{cases} 0 & x = y \\ 1 & x \neq y. \end{cases}$$

In diesem metrischen Raum ist *jede* Teilmenge abgeschlossen und beschränkt. Eine abzählbare Teilmenge $A := \{a_1, a_2, \dots\} \subset X$ ist aber nicht kompakt: Eine Folge in (X, d) ist genau dann konvergent, wenn sie ab einem bestimmten Index konstant ist. Folglich besitzt die Folge (a_n) keine konvergente Teilfolge.

Abschließend verallgemeinern wir eine weitere Eigenschaft, die Intervalle in \mathbb{R} haben: Sie zerfallen nicht in zwei disjunkte offene Mengen.

Definition 7.12. Sei (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt *zusammenhängend*, wenn es keine offenen und zueinander disjunkten Teilmengen $U, V \subset X$ gibt, so dass $A \subset U \cup V$, $A \cap U \neq \emptyset$ und $A \cap V \neq \emptyset$.

Beispiel 10: Sei $X = \mathbb{R}^2$, $U, V \subset X$ offene Teilmengen von X und $A \subset X$ wie im Bild. Dann ist A nicht zusammenhängend.



Wir beschreiben nun die zusammenhängenden Mengen im metrischen Raum \mathbb{R} mit der Standardmetrik $d(x, y) = |x - y|$.

Satz 7.14 Eine Teilmenge $A \subset \mathbb{R}$ ist genau dann zusammenhängend, wenn sie ein Intervall ist.

Beweis. (\implies) Sei $A \subset \mathbb{R}$ zusammenhängend. Angenommen A wäre kein Intervall. Dann existieren $a, b \in A$ mit $[a, b] \not\subset A$, d.h. es gibt ein $x \in (a, b)$ mit $x \notin A$. Wir betrachten die

Mengen $U := (-\infty, x)$ und $V := (x, \infty)$. U und V sind disjunkte offene Teilmengen in \mathbb{R} mit $A \subset U \cup V = \mathbb{R} \setminus \{x\}$. Außerdem ist $A \cap U \neq \emptyset$, da $a \in A \cap U$ und $A \cap V \neq \emptyset$, da $b \in A \cap V$. Dies widerspricht aber der Voraussetzung, dass A zusammenhängend ist.

(\Leftarrow) Sei A ein Intervall in \mathbb{R} . Angenommen, A wäre nicht zusammenhängend. Dann existieren offene, disjunkte Mengen $U, V \subset \mathbb{R}$, so dass $A \subset U \cup V$ und $a \in A \cap U$, $b \in A \cap V$. Da A ein Intervall ist, gilt $[a, b] \subset A \subset U \cup V$. Wir betrachten die Funktion $f : U \cup V \rightarrow \mathbb{R}$, definiert durch

$$f(x) := \begin{cases} 1 & \text{falls } x \in U, \\ -1 & \text{falls } x \in V. \end{cases}$$

f ist stetig: Sei $x \in U$. Da U offen ist, gibt es eine Kugel $K(x, \varepsilon) \subset U$. Folglich gilt für jede Folge (x_n) in $U \cup V$, die gegen x konvergiert, dass $x_n \in K(x, \varepsilon) \subset U$ für alle $n \geq n_0$. Somit ist $f(x_n) = 1$ für alle $n \geq n_0$, d.h. $f(x_n)$ konvergiert gegen $f(x) = 1$. Nach dem Folgenkriterium für stetige reelle Funktionen (Satz 4.1) ist f in x stetig. Analog zeigt man, dass f in jedem Punkt $x \in V$ stetig ist. Folglich ist $f|_{[a,b]}$ stetig. Es gilt $f(a) = 1$ und $f(b) = -1$, f nimmt aber nur die beiden Werte -1 und 1 an. Dies widerspricht dem Zwischenwertsatz für stetige reelle Funktionen (Satz 4.10). \square

7.4 Stetige Abbildungen und ihre Eigenschaften

Wir definieren die Stetigkeit von Abbildungen zwischen beliebigen metrischen Räumen völlig analog zum Fall reeller Funktionen. Wir werden sehen, dass sich wichtige Eigenschaften, die wir für stetige reelle Funktionen aus Kapitel 4 kennen, auf den allgemeinen Fall übertragen.

In diesem Abschnitt bezeichnen (X, d_X) und (Y, d_Y) stets metrische Räume.

Definition 7.13. Sei $f : A \subset X \rightarrow Y$ eine Abbildung.

1. f heißt im Punkt $p \in A$ stetig, falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass gilt:

$$\forall x \in A \text{ mit } d_X(p, x) < \delta \implies d_Y(f(p), f(x)) < \varepsilon,$$

bzw., so dass für die Kugeln gilt:

$$f(K_X(p, \delta) \cap A) \subset K_Y(f(p), \varepsilon).$$

2. Die Abbildung f heißt stetig, wenn sie in jedem Punkt $p \in A$ stetig ist.

Satz 7.15 (Folgenkriterium für Stetigkeit)

Eine Abbildung $f : A \subset X \rightarrow Y$ ist genau dann in $p \in A$ stetig, wenn für jede Folge (x_n) in A , die gegen p konvergiert, die Bildfolge $(f(x_n))$ gegen $f(p)$ konvergiert.

Beweis. Der Beweis ist völlig analog zum Beweis von Satz 4.1. Wir überlassen das Übertragen deshalb als Übungsaufgabe. \square

Wie wir bereits für reelle Funktionen gesehen hatten, ist dieses Kriterium ein äußerst nützliches Hilfsmittel, um weitere Eigenschaften stetiger Abbildungen mit Hilfe von Grenzwertsätzen für konvergente Folgen zu beweisen. Zunächst betrachten wir einige Beispiele, die sofort aus dem Folgenkriterium folgen.

Beispiel 1:

- a) Die identische Abbildung $Id_X : X \rightarrow X$, $Id_X(x) := x$, ist stetig.
- b) Die konstante Abbildung $c_p : X \rightarrow X$, $c_p(x) := p$, ist stetig.

Beispiel 2: Die Projektionen auf die Faktoren eines Produktraumes sind stetig:

Seien $(X_1, d_1), \dots, (X_k, d_k)$ metrische Räume und (X, d) ihr kartesisches Produkt. Dann ist jede Projektion

$$\begin{aligned} \pi_j : X = X_1 \times \dots \times X_k &\longrightarrow X_j \\ (x_1, \dots, x_k) &\longmapsto x_j \end{aligned}$$

stetig.

Beispiel 3: Alle algebraischen Operationen auf dem Euklidischen Vektorraum \mathbb{R}^k sind stetig, d.h.

die Summe	$+$: $\mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}^k$,
die skalare Multiplikation	\cdot : $\mathbb{R} \times \mathbb{R}^k \rightarrow \mathbb{R}^k$,
das Euklidische Skalarprodukt	$\langle \cdot, \cdot \rangle$: $\mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}$,
die Euklidische Norm	$\ \cdot \ $: $\mathbb{R}^k \rightarrow \mathbb{R}$

sind stetige Abbildungen. Dies folgt mit dem Folgenkriterium aus den Grenzwertsätzen (Satz 7.5) für Folgen im \mathbb{R}^k .

Beispiel 4: Jede lineare Abbildung $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ zwischen den Euklidischen Vektorräumen \mathbb{R}^k und \mathbb{R}^m ist stetig.

Beweis: Sei (v_n) eine Folge von Vektoren im \mathbb{R}^k mit $v_n = (v_{n1}, \dots, v_{nk})$, die gegen einen Vektor $w = (w_1, \dots, w_k) \in \mathbb{R}^k$ konvergiert. Dann konvergiert nach Satz 7.4 für jedes $j = 1, \dots, k$ die Komponentenfolge (v_{nj}) gegen w_j . Sei $e_j := (0, \dots, 0, \underbrace{1}_{j.\text{Stelle}}, 0, \dots, 0)$ der j -te kanonische Basisvektor des \mathbb{R}^k . Dann folgt aus der Linearität von L und der Stetigkeit der linearen Operationen auf \mathbb{R}^m :

$$L(v_n) = L\left(\sum_{j=1}^k v_{nj}e_j\right) = \sum_{j=1}^k v_{nj}L(e_j) \longrightarrow \sum_{j=1}^k w_jL(e_j) = L\left(\sum_{j=1}^k w_je_j\right) = L(w).$$

Nach dem Folgenkriterium ist L somit stetig.

Satz 7.16 Seien $f : A \subset X \rightarrow Y$ und $g : B \subset Y \rightarrow Z$ zwei Abbildungen zwischen metrischen Räumen, $f(A) \subset B$ und $g \circ f : A \subset X \rightarrow Z$ die Verknüpfung von f und g . Dann gilt: Ist f in $p \in A$ und g in $f(p) \in B$ stetig, so ist $g \circ f$ in $p \in A$ stetig. Insbesondere ist die Verknüpfung stetiger Abbildungen ebenfalls stetig.

Beweis. Analog zum Beweis von Satz 4.3. □

Satz 7.17 Sei $f : A \subset X \rightarrow Y_1 \times \dots \times Y_k$ eine Abbildung in das kartesische Produkt metrischer Räume und $f = (f_1, \dots, f_k)$ die Komponentendarstellung von f :

$$f(x) =: (f_1(x), \dots, f_k(x)) \quad \forall x \in X.$$

Dann gilt: Die Abbildung f ist genau dann in $p \in A$ stetig, wenn jede Komponentenfunktion $f_j : A \subset X \rightarrow Y_j$, $j = 1, \dots, k$, in p stetig ist.

Beweis. (\implies) Da $f_j = \pi_j \circ f$, folgt dies aus Beispiel 2 und Satz 7.16.

(\impliedby) Sei (x_n) eine Folge in A , die gegen $p \in A$ konvergiert. Da alle f_j in p stetig sind, konvergiert $(f_j(x_n))$ gegen $f_j(p) \in Y_j$ für jedes $j = 1, \dots, k$. Somit konvergiert $(f(x_n)) = ((f_1(x_n), \dots, f_k(x_n)))$ gegen $(f_1(p), \dots, f_k(p)) = f(p)$, d.h. f ist in p stetig. □

Dies entspricht dem von uns in Kapitel 6.4. für parametrisierte Kurven $\gamma : I \subset \mathbb{R} \rightarrow \mathbb{R}^k$ benutzten Stetigkeitsbegriff.

Neben dem Folgenkriterium ist auch das folgende topologische Kriterium für die Stetigkeit von Abbildungen nützlich:

Satz 7.18 (Topologisches Kriterium für Stetigkeit)

Eine Abbildung $f : X \rightarrow Y$ zwischen metrischen Räumen ist genau dann stetig, wenn das Urbild $f^{-1}(U) \subset X$ jeder offenen Menge $U \subset Y$ offen ist.

Beweis. (\implies) Sei $f : X \rightarrow Y$ stetig und $U \subset Y$ offen. Wir zeigen, dass das Urbild

$$f^{-1}(U) := \{x \in X \mid f(x) \in U\} \subset X$$

dann ebenfalls offen ist.

Ist $f^{-1}(U) = \emptyset$, so ist die Behauptung erfüllt, da \emptyset offen ist. Sei also $f^{-1}(U) \neq \emptyset$ und $p \in f^{-1}(U)$ ein beliebig gewählter Punkt. Dann gilt $f(p) \in U$, und da $U \subset Y$ offen ist, existiert ein $\varepsilon > 0$ mit $K_Y(f(p), \varepsilon) \subset U$. Da f in p stetig ist, gibt es ein $\delta > 0$, so dass

$$f(K_X(p, \delta)) \subset K_Y(f(p), \varepsilon) \subset U.$$

Aus der Definition des Urbildes folgt dann:

$$K_X(p, \delta) \subset f^{-1}(U).$$

Somit ist $f^{-1}(U)$ offen in X .

(\impliedby) Sei $f^{-1}(U) \subset X$ offen für jede offene Menge $U \subset Y$. Wir zeigen, dass f stetig ist. Wir wählen einen beliebigen Punkt $p \in X$ und ein beliebiges $\varepsilon > 0$. Da die Kugel $K_Y(f(p), \varepsilon)$ in Y offen ist, ist nach Voraussetzung auch das Urbild $f^{-1}(K_Y(f(p), \varepsilon))$ in X offen. Es enthält den Punkt p . Folglich existiert ein $\delta > 0$, so dass $K_X(p, \delta) \subset f^{-1}(K_Y(f(p), \varepsilon))$. Nach Definition des Urbildes folgt daraus $f(K_X(p, \delta)) \subset K_Y(f(p), \varepsilon)$. Das bedeutet aber, dass f in p stetig ist. Da $p \in X$ beliebig gewählt war, ist f stetig. □

Genauso wie für reelle Funktionen hat man für Abbildungen zwischen metrischen Räumen zwei stärkere Stetigkeitsbegriffe.

Definition 7.14. Sei $f : A \subset X \rightarrow Y$ eine Abbildung zwischen metrischen Räumen.

1. f heißt *gleichmäßig stetig*, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass gilt:

$$\forall p, q \in A \text{ mit } d_X(p, q) < \delta \implies d_Y(f(p), f(q)) < \varepsilon.$$

(Im Unterschied zur Definition der Stetigkeit hängt hier die Größe von δ nur von ε , aber nicht von p oder q ab.)

2. f heißt *lipschitzstetig*, wenn es eine positive Konstante $L \in \mathbb{R}^+$ gibt, so dass gilt:

$$d_Y(f(p), f(q)) \leq L \cdot d_X(p, q) \quad \forall p, q \in A.$$

L heißt *Lipschitz-Konstante* von f .

Satz 7.19 Sei $f : A \subset X \rightarrow Y$ eine Abbildung zwischen metrischen Räumen. Dann gilt:

1. f ist lipschitzstetig $\implies f$ ist gleichmäßig stetig.
2. f ist gleichmäßig stetig $\implies f$ ist stetig.

Beweis. Analog zum Beweis von Satz 4.5. □

Beispiel 5: Jede lineare Abbildung $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ zwischen den Euklidischen Vektorräumen \mathbb{R}^k und \mathbb{R}^m ist sogar lipschitzstetig (Übungsaufgabe 23 (Analysis 2)).

Wir betrachten nun das Verhalten von stetigen Abbildungen auf zusammenhängenden Mengen und verallgemeinern den Zwischenwertsatz für reelle Funktionen (Satz 4.10).

Satz 7.20 Sei $f : X \rightarrow Y$ eine stetige Abbildung zwischen metrischen Räumen. Dann ist das Bild jeder zusammenhängenden Menge ebenfalls zusammenhängend.

Beweis. Sei $A \subset X$ eine zusammenhängende Menge. Wir wollen zeigen, dass dann auch das Bild $f(A) \subset Y$ zusammenhängend ist. Angenommen $f(A) \subset Y$ wäre nicht zusammenhängend. Dann existieren offene Mengen $U, V \subset Y$ mit

- $U \cap V = \emptyset$,
- $f(A) \subset U \cup V$,
- $f(A) \cap U \neq \emptyset$ und $f(A) \cap V \neq \emptyset$.

Da f stetig ist, sind nach Satz 7.18 die Urbilder $f^{-1}(U)$ und $f^{-1}(V)$ ebenfalls offen. Desweiteren gilt für diese Urbilder

- $f^{-1}(U) \cap f^{-1}(V) = f^{-1}(U \cap V) = f^{-1}(\emptyset) = \emptyset$.
- $A \subset f^{-1}(f(A)) \subset f^{-1}(U \cup V) = f^{-1}(U) \cup f^{-1}(V)$.
- Nach Voraussetzung existiert ein $a \in A$ mit $f(a) \in U$ und ein $b \in A$ mit $f(b) \in V$.
Dann gilt: $a \in A \cap f^{-1}(U) \neq \emptyset$ und $b \in A \cap f^{-1}(V) \neq \emptyset$.

Folglich ist A nicht zusammenhängend, was ein Widerspruch zur Voraussetzung ist. \square

Wir definieren noch einen weiteren Zusammenhangsbegriff für Teilmengen metrischer Räume (X, d) .

Definition 7.15. Eine Teilmenge $A \subset X$ heißt *wegzusammenhängend*, falls zu je zwei Punkten $a, b \in A$ eine stetige Abbildung $\omega : [0, 1] \subset \mathbb{R} \rightarrow X$ existiert mit $\omega([0, 1]) \subset A$, $\omega(0) = a$ und $\omega(1) = b$.

Die Abbildung ω heißt *Weg in A von a nach b* .

Beispiel 6: Beziehung zwischen den beiden Zusammenhangsbegriffen

a) Für Teilmengen $A \subset \mathbb{R}$ sind die Zusammenhangsbegriffe äquivalent:

$$A \text{ zusammenhängend} \stackrel{S. 7.14}{\iff} A \text{ Intervall} \iff A \text{ wegzusammenhängend.}$$

b) In metrischen Räumen ist jede wegzusammenhängende Menge auch zusammenhängend, aber die Umkehrung gilt i.a. nicht (für ein Beispiel siehe O. Deiser: Analysis 2, Seite 198).

c) Konvexe Mengen im Euklidischen Raum \mathbb{R}^k :

Eine Teilmenge $A \subset \mathbb{R}^k$ heißt *konvex*, wenn mit je zwei Punkten $x, y \in A$ auch die Verbindungsstrecke $\overline{xy} := \{x + t(y - x) \mid t \in [0, 1]\}$ vollständig in A liegt.

Jede konvexe Menge ist wegzusammenhängend, da die Abbildung $\omega : [0, 1] \rightarrow A \subset \mathbb{R}^k$,

$$\omega(t) := x + t(y - x),$$

stetig ist und die Punkte x und y verbindet.

Zum Beispiel ist jede Kugel $K(x_0, r)$ und ihr Abschluß konvex.

Der nächste Satz ist eine weitere Verallgemeinerung des Zwischenwertsatzes für reelle Funktionen.

Satz 7.21 Sei $f : X \rightarrow Y$ eine stetige Abbildung. Dann ist das Bild jeder wegzusammenhängenden Teilmenge ebenfalls wegzusammenhängend.

Beweis. Sei $A \subset X$ wegzusammenhängend. Wir wollen zeigen, dass das Bild $f(A)$ ebenfalls wegzusammenhängend ist. Seien $x, y \in f(A)$ zwei Punkte in $f(A)$. Dann existieren $a, b \in A$ mit $x = f(a)$, $y = f(b)$. Da A wegzusammenhängend ist, gibt es einen (stetigen) Weg $\omega : [0, 1] \rightarrow A \subset X$ von a nach b . Wir betrachten die Abbildung $f \circ \omega : [0, 1] \rightarrow f(A) \subset Y$. Da f und ω stetig sind, ist $f \circ \omega$ ebenfalls stetig. Weiterhin gilt $(f \circ \omega)(0) = f(\omega(0)) = f(a) = x$ und $(f \circ \omega)(1) = f(\omega(1)) = f(b) = y$, d.h. $f \circ \omega$ ist ein Weg in $f(A)$ von x nach y . \square

Als nächstes beschäftigen wir uns mit dem Verhalten von stetigen Abbildungen auf kompakten Mengen. Die für stetige reelle Funktionen bekannten Aussagen kann man durch Ersetzen der Metrik direkt auf den Fall von Abbildungen zwischen metrischen Räumen übertragen:

Satz 7.22 Sei $f : X \rightarrow Y$ eine stetige Abbildung. Dann gilt:

1. Das Bild jeder kompakten Menge ist kompakt.
2. f ist auf jeder kompakten Menge $K \subset X$ gleichmäßig stetig.
3. Ist f insbesondere reell-wertig, so nimmt f auf jeder kompakten Menge $K \subset X$ ein Maximum und ein Minimum an, d.h. es existieren Punkte $\xi_1, \xi_2 \in K$ mit

$$f(\xi_1) \leq f(x) \leq f(\xi_2) \quad \forall x \in K.$$

Beweis. Beweis von 1. ist analog zum Beweis von Satz 4.12, Beweis von 2. analog zum Beweis von Satz 4.14, Beweis von 3. analog zum Beweis von Satz 4.13. \square

Definition 7.16. Eine Abbildung $f : X \rightarrow Y$ zwischen zwei metrischen Räumen heißt Homöomorphismus, wenn f bijektiv ist und f und f^{-1} stetig sind.

Zwei metrische Räume X und Y heißen homöomorph, falls es einen Homöomorphismus $f : X \rightarrow Y$ gibt.

Bei homöomorphen metrischen Räumen bleibt zwar die Abstandsmessung nicht erhalten, aber alle topologischen Eigenschaften von Teilmengen (offen, abgeschlossen, kompakt, zusammenhängend, wegzusammenhängend, ...) übertragen sich auf die Bildmenge. Die Konvergenz einer Folge überträgt sich ebenfalls auf die Konvergenz der Bildfolge.

Beispiel 7: Die inverse Abbildung einer bijektiven, stetigen Abbildung ist im Allgemeinen nicht stetig, wie das folgende Beispiel zeigt:

Seien $X = (0, 1) \cup \{2\} \subset \mathbb{R}$ und $Y = (0, 1] \subset \mathbb{R}$, versehen mit der Standardmetrik $d(x, y) = |x - y|$. Wir betrachten die Abbildung $f : X \rightarrow Y$ gegeben durch

$$f(t) := \begin{cases} t & \text{falls } t \in (0, 1) \\ 1 & \text{falls } t = 2. \end{cases}$$

f ist stetig und bijektiv. Aber $f^{-1} : (0, 1] \rightarrow (0, 1) \cup \{2\}$ ist nach Zwischenwertsatz nicht stetig.

Für kompakte metrische Räume gilt aber folgender nützliche Satz:

Satz 7.23 (Satz über die Stetigkeit der inversen Abbildung)

Sei $f : X \rightarrow Y$ eine bijektive stetige Abbildung zwischen metrischen Räumen und sei X kompakt. Dann ist die inverse Abbildung $f^{-1} : Y \rightarrow X$ stetig.

Beweis. Wir benutzen das topologische Kriterium für Stetigkeit (Satz 7.18). Sei $U \subset X$ offen. Wir zeigen, dass das Urbild $(f^{-1})^{-1}(U) = f(U) \subset Y$ offen ist.

Da $U \subset X$ offen ist, ist $X \setminus U \subset X$ abgeschlossen. Da X kompakt ist, ist $X \setminus U \subset X$ ebenfalls kompakt (Satz 7.11). Dann ist nach Satz 7.22 auch das Bild $f(X \setminus U)$ in Y kompakt und insbesondere abgeschlossen (Satz 7.10). Da f bijektiv ist, gilt $f(X \setminus U) = Y \setminus f(U)$. Also ist die Teilmenge $f(U) \subset Y$ offen. \square

Definition 7.17. Sei (X, d) ein metrischer Raum. Eine Abbildung $f : X \rightarrow X$ heißt kontrahierend (oder Kontraktion), wenn sie Lipschitzstetig mit einer Lipschitzkonstanten $0 < L < 1$ ist, das heißt wenn eine Konstante $0 < L < 1$ existiert, so dass

$$d(f(x), f(y)) \leq L \cdot d(x, y) \quad \text{für alle } x, y \in X.$$

Für kontrahierende Abbildungen in vollständigen metrischen Räumen gilt der folgende sehr nützliche Fixpunktsatz, den wir für reelle Funktionen bereits kennen (Übungsaufgabe 35, Analysis 1).

Satz 7.24 (Banachscher Fixpunktsatz)

Sei (X, d) ein vollständiger metrischer Raum und $f : X \rightarrow X$ eine kontrahierende Abbildung. Dann hat f genau einen Fixpunkt. Diesen Fixpunkt erhält man konstruktiv:

Sei $x_0 \in X$ ein beliebiger Punkt und $x_n := \underbrace{f \circ \dots \circ f}_{n\text{-mal}}(x_0) =: f^n(x_0)$. Dann konvergiert die Folge (x_n) gegen den Fixpunkt von f .

Beweis. (1) Eindeutigkeit des Fixpunktes: Angenommen es existieren zwei Fixpunkte ξ und η von f , dh. es gelte $f(\xi) = \xi$ und $f(\eta) = \eta$. Aus der Kontraktivität von f erhält man

$$d(\xi, \eta) = d(f(\xi), f(\eta)) \leq L \cdot d(\xi, \eta).$$

Da aber $0 < L < 1$ gilt, folgt $d(\xi, \eta) = 0$, also $\xi = \eta$.

(2) Existenz des Fixpunktes: Sei $x_0 \in X$ ein beliebig gewählter Punkt. Wir definieren eine Folge von Punkten in X durch

$$\begin{aligned} x_1 &:= f(x_0), \\ x_2 &:= f(x_1) = f^2(x_0), \\ &\vdots \\ x_n &:= f(x_{n-1}) = f^2(x_{n-2}) = \dots = f^n(x_0). \end{aligned}$$

Dann gilt

$$\begin{aligned} d(x_{n+1}, x_n) &= d(f(x_n), f(x_{n-1})) \\ &\leq L \cdot d(x_n, x_{n-1}) = L \cdot d(f(x_{n-1}), f(x_{n-2})) \\ &\leq L^2 \cdot d(x_{n-1}, x_{n-2}) = L^2 \cdot d(f(x_{n-2}), f(x_{n-3})) \\ &\vdots \\ &\leq L^n \cdot d(x_1, x_0). \end{aligned}$$

Für $n > m$ folgt mit der Dreiecksungleichung und der Formel für die geometrische Summe

$$\begin{aligned}
d(x_n, x_m) &\leq d(x_n, x_{n-1}) + d(x_{n-1}, x_{n-2}) + \dots + d(x_{m+1}, x_m) \\
&\leq (L^{n-1} + \dots + L^m) \cdot d(x_0, x_1) \\
&= L^m (L^0 + L^1 + \dots + L^{n-1-m}) \cdot d(x_0, x_1) \\
&= L^m \cdot \frac{1 - L^{n-m}}{1 - L} \cdot d(x_0, x_1) \\
&\leq L^m \cdot \frac{1}{1 - L} \cdot d(x_0, x_1).
\end{aligned}$$

Da $0 < L < 1$, ist (L^m) eine Nullfolge. Somit ist (x_n) eine Cauchyfolge in X . Da der metrische Raum X vollständig ist, konvergiert diese Cauchyfolge gegen einen Punkt ξ aus X . Wir zeigen nun, dass dieser Grenzwert ξ der gesuchte Fixpunkt von f ist. Dazu betrachten wir

$$\begin{aligned}
d(f(\xi), \xi) &\leq d(f(\xi), x_n) + d(x_n, \xi) \\
&= d(f(\xi), f(x_{n-1})) + d(x_n, \xi) \\
&\leq L \underbrace{d(\xi, x_{n-1})}_{\rightarrow 0} + \underbrace{d(x_n, \xi)}_{\rightarrow 0}.
\end{aligned}$$

Daraus folgt $d(f(\xi), \xi) = 0$, also $f(\xi) = \xi$. \square

Abschließend beweisen wir einen Satz, der zeigt, dass alle Eigenschaften, die wir in diesem Kapitel für den Euklidischen Raum \mathbb{R}^k bewiesen haben, auch gelten, wenn wir den \mathbb{R}^k mit einer beliebigen Norm versehen.

Satz 7.25 Sei \mathbb{R}^k der Euklidische Vektorraum mit der Euklidischen Norm $\|\cdot\|$ und sei $N : \mathbb{R}^k \rightarrow \mathbb{R}$ eine weitere Norm auf \mathbb{R}^k . Dann gilt:

1. $N : \mathbb{R}^k \rightarrow \mathbb{R}$ ist Lipschitzstetig.
2. Es existieren Konstanten $a, b \in \mathbb{R}^+$, so dass

$$a \cdot \|x\| \leq N(x) \leq b \cdot \|x\| \quad \forall x \in \mathbb{R}^k.$$

Beweis. 1) Wir zeigen mittels der Cauchy-Schwarzschen Ungleichung und der Normeigenschaften die Existenz von $b \in \mathbb{R}^+$:

Sei $x \in \mathbb{R}^k$ und $x = \sum_{j=1}^k x_j e_j$ die Basisdarstellung von x bezüglich der kanonischen Basis (e_1, \dots, e_k) im \mathbb{R}^k . Dann gilt:

$$\begin{aligned}
N(x) &= N\left(\sum_{j=1}^k x_j e_j\right) \stackrel{\Delta}{\leq} \sum_{j=1}^k N(x_j e_j) = \sum_{j=1}^k |x_j| N(e_j) \\
&\stackrel{CSU}{\leq} \sqrt{\sum_{j=1}^k |x_j|^2} \cdot \sqrt{\sum_{j=1}^k N(e_j)^2} \\
&= \|x\| \cdot \underbrace{\sqrt{\sum_{j=1}^k N(e_j)^2}}_{=: b > 0}.
\end{aligned}$$

2) Für eine Norm N gilt die umgekehrte Dreiecksungleichung

$$|N(x) - N(y)| \leq N(x - y) \quad \forall x, y \in \mathbb{R}^k. \quad (7.1)$$

Dies folgt aus zweifacher Anwendung der Dreiecksungleichung für N :

$$\begin{aligned} N(x) &= N(y + (x - y)) \stackrel{\Delta}{\leq} N(y) + N(x - y), \\ N(y) &= N(x + (y - x)) \stackrel{\Delta}{\leq} N(x) + N(y - x) = N(x) + N(x - y). \end{aligned}$$

Durch Umstellen erhalten wir

$$\pm(N(x) - N(y)) \leq N(x - y).$$

Also gilt (7.1).

3) N ist Lipschitzstetig mit der Lipschitz-Konstanten $b \in \mathbb{R}^+$, da

$$|N(x) - N(y)| \stackrel{2)}{\leq} N(x - y) \stackrel{1)}{\leq} b \cdot \|x - y\| \quad \forall x, y \in \mathbb{R}^k.$$

4) Wir zeigen die Existenz von $a \in \mathbb{R}^+$:

Für $x = 0$ gilt die im Satz behauptete Ungleichung immer. Sei also $x \neq 0$. Dann gilt

$$N(x) = N\left(\frac{x}{\|x\|}\right) = \|x\| \cdot N\left(\frac{x}{\|x\|}\right). \quad (7.2)$$

Da $\left\|\frac{x}{\|x\|}\right\| = \frac{\|x\|}{\|x\|} = 1$, liegt $\frac{x}{\|x\|}$ in der Euklidischen Sphäre $S^{k-1} := \{y \in \mathbb{R}^k \mid \|y\| = 1\}$. Die Abbildung $N : \mathbb{R}^k \rightarrow \mathbb{R}$ ist stetig und die Menge $S^{k-1} \subset \mathbb{R}^k$ ist kompakt. Folglich existiert das Minimum

$$a := \min\{N(y) \mid y \in S^{k-1}\}$$

(Satz 7.22). Dann ist $a = N(p)$ für ein $p \in S^{k-1}$, also $a > 0$. Aus (7.2) folgt dann

$$N(x) \geq \|x\| \cdot a \quad \forall x \in \mathbb{R}^k. \quad \square$$

Als Konsequenz erhalten wir

Satz 7.26 Sei $\|\cdot\|$ die Euklidische Norm und N eine weitere Norm auf dem Vektorraum \mathbb{R}^k . Dann stimmen alle topologischen Eigenschaften von $(\mathbb{R}^k, \|\cdot\|)$ und (\mathbb{R}^k, N) überein. D.h. für Folgen (x_n) im \mathbb{R}^k und für Teilmengen $A \subset \mathbb{R}^k$ gilt:

1. (x_n) konvergiert gegen p in $(\mathbb{R}^k, \|\cdot\|) \iff (x_n)$ konvergiert gegen p in (\mathbb{R}^k, N) .
2. (x_n) ist Cauchy-Folge in $(\mathbb{R}^k, \|\cdot\|) \iff (x_n)$ ist Cauchy-Folge in (\mathbb{R}^k, N) .
3. A ist offen (abgeschlossen, beschränkt, kompakt, zusammenhängend bzw. wegzusammenhängend) bzgl. $\|\cdot\|$ genau dann, wenn A offen (abgeschlossen, beschränkt, kompakt, zusammenhängend bzw. wegzusammenhängend) bzgl. N ist.

Insbesondere ist der normierte Vektorraum (\mathbb{R}^k, N) für jede Norm N vollständig.

Beweis. Aus der Abschätzung in Satz 7.25 folgt für die Kugeln $K_{eu}(x, \varepsilon)$ bzgl. der Norm $\|\cdot\|$ und $K_N(x, \varepsilon)$ bzgl. der Norm N :

$$K_{eu}(x, \varepsilon) \subset K_N(x, b \cdot \varepsilon) \quad \text{und} \quad K_N(x, \varepsilon) \subset K_{eu}\left(x, \frac{\varepsilon}{a}\right) \quad \forall \varepsilon > 0.$$

Da Konvergenz von Folgen, Cauchy-Folgen, offene und beschränkte Mengen mittels Abständen bzw. ε -Kugeln definiert sind, folgt die Äquivalenz dieser Eigenschaften für $\|\cdot\|$ und N . Abgeschlossene und kompakte Mengen sind durch Konvergenzeigenschaften von Folgen charakterisiert. Folglich stimmen diese Eigenschaften für beide Normen überein. Der Zusammenhang ist durch offene Mengen, der Wegzusammenhang durch stetige Abbildungen, also ebenfalls durch Konvergenzeigenschaften charakterisiert, folglich stimmen auch diese Eigenschaften für beide Normen überein. \square

Differentialrechnung für Funktionen mehrerer reeller Variablen

Nachdem wir in Kapitel 5 die Differentialrechnung für Funktionen *einer* reellen Variablen behandelt haben, wollen wir uns jetzt mit Funktionen bzw. Abbildungen befassen, die von *mehreren* reellen Variablen abhängen, d.h. mit Abbildungen der Form $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$. Auch für solche Abbildungen werden wir einen Differenzierbarkeitsbegriff einführen, der der Abbildung f in jedem Punkt $p \in U$ eine Ableitung zuordnet. Aus den Eigenschaften der Ableitung möchte man wieder Schlüsse für die Eigenschaften von f selbst ziehen. Für eine Abbildung $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ ist die Ableitung im Punkt $p \in U$ eine *lineare Abbildung* vom \mathbb{R}^k in den \mathbb{R}^m . *Lineare* Abbildungen sind im Gegensatz zu beliebigen Abbildungen sehr "einfach". Durch die lineare Struktur kann man mit ihnen sehr gut umgehen (siehe die Vorlesungen über lineare Algebra). Wir können also Methoden der linearen Algebra benutzen, um mit Hilfe der Ableitungen analytische Eigenschaften von beliebigen Abbildungen $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ zu untersuchen. Uns interessieren dabei z.B. folgende Fragen:

- Wie kann man lokale Extrema von Funktionen $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ finden?
- Wie kann man Tangentialebenen an Flächen im \mathbb{R}^3 berechnen.
- Wann gibt es Lösungen der Gleichung $f(x_1, \dots, x_k) = 0$. Wie sieht Lösungsmenge aus?
- Wie kann man einer Abbildung $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^k$ ansehen, ob sie (wenigstens lokal) umkehrbar ist? Wann ist sie eine Koordinatentransformation?

Für die Differentialrechnung spielt neben der linearen Struktur der benutzen Räume wieder die Konvergenz von Folgen eine entscheidende Rolle. Wir versehen deshalb die Räume \mathbb{R}^k bzw. \mathbb{R}^m mit der Euklidischen Norm $\|\cdot\|$ und betrachten den dadurch induzierten Konvergenzbegriff. Wie wir aus Kapitel 7.4 wissen, ist es für Konvergenzfragen allerdings egal, welche Norm auf dem \mathbb{R}^k bzw. \mathbb{R}^m wir benutzen.

In den Vektorräumen \mathbb{R}^k benutzen wir oft die kanonische Basis (e_1, \dots, e_k) , wobei

$$e_i := (0, \dots, 0, \underbrace{1}_{i.\text{Stelle}}, 0, \dots, 0) \in \mathbb{R}^k.$$

Im folgenden seien die Definitionsbereiche $U \subset \mathbb{R}^k$ der Abbildungen immer *offene* Teilmengen, d.h. jeder Punkt $p \in U$ ist ein innerer Punkt.

8.1 Differenzierbare Abbildungen, das Differential, die Richtungsableitungen und die partiellen Ableitungen

Bevor wir uns mit der Differenzierbarkeit von Abbildungen mehrerer reeller Variablen befassen, erinnern wir uns nochmal daran, wie wir die Differenzierbarkeit für Funktionen *einer* Variablen in Kapitel 5 definiert hatten:

Eine Funktion $f : U \subset \mathbb{R} \rightarrow \mathbb{R}$ heißt in $p \in U$ differenzierbar, wenn der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(p+h) - f(p)}{h} =: f'(p) \in \mathbb{R} \quad (8.1)$$

existiert. Den Begriff des Grenzwertes kann man für Abbildungen mehrerer reeller Variablen genauso definieren wie für reelle Funktionen einer Variablen.

Definition 8.1. Sei $\phi : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine Abbildung.

Man sagt: ϕ besitzt in $q \in U$ den Grenzwert $y \in \mathbb{R}^m$, wenn für jede Folge (x_n) in $U \setminus \{q\}$, die gegen q konvergiert, die Bildfolge $(\phi(x_n))$ gegen y konvergiert.

Bezeichnung: $\lim_{x \rightarrow q} \phi(x) = y$.

Die Bedingung (8.1) lässt sich auch schreiben als

$$\lim_{h \rightarrow 0} \frac{f(p+h) - f(p) - f'(p) \cdot h}{h} = 0$$

oder in der Form

$$f(p+h) = f(p) + f'(p) \cdot h + r_p(h),$$

wobei $r_p(h) \in \mathbb{R}$ ein Fehlerterm ist, für den

$$\lim_{h \rightarrow 0} \frac{r_p(h)}{h} = 0$$

gilt. Die Zuordnung $L : h \in \mathbb{R} \mapsto f'(p) \cdot h \in \mathbb{R}$ ist dabei eine lineare Abbildung.

Das motiviert die folgende Definition für Abbildungen, die von *mehreren* reellen Variablen abhängen:

Definition 8.2. Eine Abbildung $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ heißt in $p \in U$ differenzierbar, wenn eine lineare Abbildung $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ existiert, so dass gilt

$$\lim_{h \rightarrow 0} \frac{f(p+h) - f(p) - L(h)}{\|h\|} = 0. \quad (8.2)$$

Die lineare Abbildung $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ heißt Ableitung von f in p oder auch Differential von f in p . Wir bezeichnen die Ableitung bzw. das Differential mit $L =: df_p$.

$f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ heißt differenzierbar, wenn f in jedem $p \in U$ differenzierbar ist.¹

¹ In den Ingenieurwissenschaften sagt man auch f ist in p total differenzierbar und nennt df_p das totale Differential von f in p .

Die Bedingung (8.2) ist äquivalent zu

$$\lim_{h \rightarrow 0} \frac{\|f(p+h) - f(p) - L(h)\|}{\|h\|} = 0. \quad (8.3)$$

Wir können (8.2) auch in der folgenden Form schreiben:

$$f(p+h) = f(p) + L(h) + r_p(h) \quad \text{mit} \quad \lim_{h \rightarrow 0} \frac{r_p(h)}{\|h\|} = 0. \quad (8.4)$$

Satz 8.1 Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ in $p \in U$ differenzierbar. Dann gilt:

1. Das Differential $df_p : \mathbb{R}^k \rightarrow \mathbb{R}^m$ ist eindeutig bestimmt.
2. f ist in p stetig.
3. Sei $f =: (f_1, \dots, f_m) : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ die Komponentendarstellung von f . Die Abbildung f ist genau dann in $p \in U$ differenzierbar, wenn jede Komponente f_1, \dots, f_m in $p \in U$ differenzierbar ist. In diesem Fall gilt

$$df_p = ((df_1)_p, \dots, (df_m)_p).$$

Beweis. Zu 1) Seien $L, \tilde{L} : \mathbb{R}^k \rightarrow \mathbb{R}^m$ zwei lineare Abbildungen mit

$$f(p+h) - f(p) = L(h) + r_p(h) = \tilde{L}(h) + \tilde{r}_p(h)$$

und $\lim_{h \rightarrow 0} \frac{r_p(h)}{\|h\|} = \lim_{h \rightarrow 0} \frac{\tilde{r}_p(h)}{\|h\|} = 0$. Dann folgt

$$\lim_{h \rightarrow 0} \frac{L(h) - \tilde{L}(h)}{\|h\|} = 0. \quad (8.5)$$

Da beide Abbildungen linear sind, gilt $L(0) = \tilde{L}(0) = 0$. Sei nun $x \in \mathbb{R}^k$, $x \neq 0$ und $t \in \mathbb{R}^+$. Wir setzen in (8.5) $h := tx$ mit $t \rightarrow 0^+$ ein. Wegen der Linearität von L und \tilde{L} erhalten wir

$$0 = \lim_{t \rightarrow 0^+} \frac{L(tx) - \tilde{L}(tx)}{\|tx\|} = \lim_{t \rightarrow 0^+} \frac{L(x) - \tilde{L}(x)}{\|x\|} = \frac{L(x) - \tilde{L}(x)}{\|x\|}.$$

Also ist $L(x) = \tilde{L}(x)$ für alle $x \in \mathbb{R}^k$.

Zu 2) Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ in $p \in U$ differenzierbar. Wir benutzen das Folgenkriterium, um die Stetigkeit von f in $p \in U$ zu zeigen. Sei (x_n) eine gegen p konvergente Folge in U und $x_n \neq p$. Dann gilt nach Dreiecksungleichung

$$\begin{aligned} \|f(x_n) - f(p)\| &\leq \|f(x_n) - f(p) - L(x_n - p)\| + \|L(x_n - p)\| \\ &= \underbrace{\frac{\|f(x_n) - f(p) - L(x_n - p)\|}{\|x_n - p\|}}_{\rightarrow 0} \cdot \underbrace{\|x_n - p\|}_{\rightarrow 0} + \underbrace{\|L(x_n - p)\|}_{\rightarrow 0}. \end{aligned}$$

Die ersten beiden Konvergenzen folgen aus der Definition der Differenzierbarkeit und der Stetigkeit der Norm. Für die dritte Konvergenz benutzen wir außerdem die Stetigkeit der

linearen Abbildung L . Folglich konvergiert die Folge $(f(x_n))$ gegen $f(p)$, d.h. f ist in $p \in U$ stetig.

Zu 3) Die Bedingung

$$\lim_{h \rightarrow 0} \frac{f(p+h) - f(p) - L(h)}{\|h\|} = 0$$

für eine Abbildung $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ ist äquivalent zur Bedingung an die Komponenten

$$\lim_{h \rightarrow 0} \frac{f_j(p+h) - f_j(p) - L_j(h)}{\|h\|} = 0 \quad \forall j = 1, \dots, m,$$

wobei $L =: (L_1, \dots, L_m)$ die Komponenten von L bezeichnen. $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ ist genau dann linear, wenn alle Komponenten $L_j : \mathbb{R}^k \rightarrow \mathbb{R}$ linear sind. Somit ist f genau dann in p differenzierbar, wenn alle f_1, \dots, f_m in p differenzierbar sind und es gilt $df_p = L = (L_1, \dots, L_m) = ((df_1)_p, \dots, (df_m)_p)$. \square

Beispiel 1: Für Abbildungen $f : U \subset \mathbb{R} \rightarrow \mathbb{R}^m$, die nur von einer reellen Variablen abhängen, stimmen die bisherigen Ableitungsbegriffe überein, wenn man den Vektorraum \mathbb{R}^m mit den linearen Abbildungen $L(\mathbb{R}, \mathbb{R}^m)$ identifiziert. Es gilt:

$$f'(p) = f'(p) \cdot 1 = df_p(1) \in \mathbb{R}^m.$$

Beispiel 2: *Das Differential einer linearen Abbildung.*

Eine lineare Abbildung $L : \mathbb{R}^k \rightarrow \mathbb{R}^m$ ist in jedem Punkt $p \in \mathbb{R}^k$ differenzierbar und es gilt $dL_p = L$ für alle $p \in \mathbb{R}^k$:

Beweis: Wegen der Linearität von L gilt

$$\lim_{h \rightarrow 0} \frac{L(p+h) - L(p) - L(h)}{\|h\|} = \lim_{h \rightarrow 0} \frac{L(p) + L(h) - L(p) - L(h)}{\|h\|} = \lim_{h \rightarrow 0} \underbrace{\frac{0}{\|h\|}}_{=0} = 0.$$

L erfüllt also die Bedingungen für das Differential, folglich ist $L = dL_p$.

Definition 8.3. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine Abbildung und $\mathbf{a} \in \mathbb{R}^k$ ein fixierter Vektor.

Man sagt: f besitzt in $p \in U$ eine Ableitung in Richtung \mathbf{a} , falls der Grenzwert

$$\nabla_{\mathbf{a}} f(p) := \lim_{t \rightarrow 0} \frac{f(p + t\mathbf{a}) - f(p)}{t} \in \mathbb{R}^m \quad (\text{hier ist } t \in \mathbb{R})$$

existiert. Dieser Grenzwert heißt die Richtungsableitung von f in Richtung \mathbf{a} an der Stelle p .

Bei der Richtungsableitung wird die Abbildung f also nur entlang des Geradenstückes $\{p + t\mathbf{a} \mid t \in (-\varepsilon, \varepsilon)\} \subset U$ betrachtet. Bezeichnet $\varphi_{\mathbf{a}} : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^m$ die Abbildung

$$\varphi_{\mathbf{a}}(t) := f(p + t\mathbf{a}),$$

so gilt

$$\nabla_{\mathbf{a}} f(p) = \lim_{t \rightarrow 0} \frac{\varphi_{\mathbf{a}}(t) - \varphi_{\mathbf{a}}(0)}{t} = \varphi'_{\mathbf{a}}(0).$$

Satz 8.2 Ist $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ in $p \in U$ differenzierbar, so besitzt f in p in jeder Richtung $\mathbf{a} \in \mathbb{R}^k$ die Richtungsableitung und es gilt

$$df_p(\mathbf{a}) = \nabla_{\mathbf{a}}f(p).$$

Beweis. Sei f in p differenzierbar, $df_p : \mathbb{R}^k \rightarrow \mathbb{R}^m$ das Differential und $\mathbf{a} \in \mathbb{R}^k$. Für $\mathbf{a} = 0$ folgt die Behauptung aus der Linearität von df_p . Für $\mathbf{a} \neq 0$ betrachten wir $h = t\mathbf{a}$. Dann gilt nach Definition der Differenzierbarkeit und der Linearität von df_p :

$$0 = \lim_{t \rightarrow 0} \frac{\|f(p + t\mathbf{a}) - f(p) - df_p(t\mathbf{a})\|}{\|t\mathbf{a}\|} = \frac{1}{\|\mathbf{a}\|} \cdot \lim_{t \rightarrow 0} \left\| \frac{f(p + t\mathbf{a}) - f(p)}{t} - df_p(\mathbf{a}) \right\|.$$

Damit existiert

$$\nabla_{\mathbf{a}}f(p) = \lim_{t \rightarrow 0} \frac{f(p + t\mathbf{a}) - f(p)}{t} = df_p(\mathbf{a}). \quad \square$$

Beispiel 3: Eine Abbildung kann in einem Punkt p alle Richtungsableitungen besitzen, ohne in diesem Punkt differenzierbar zu sein.

Wir betrachten dazu die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, definiert durch

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2+y^4} & \text{falls } (x, y) \neq (0, 0), \\ 0 & \text{falls } (x, y) = (0, 0). \end{cases}$$

Die Funktion f ist im Punkt $p := (0, 0)$ nicht stetig, also auch nicht differenzierbar. Um dies einzusehen, betrachten wir die Folge (a_n) mit $a_n := (\frac{1}{n^2}, \frac{1}{n})$. Dann gilt

$$a_n = (\frac{1}{n^2}, \frac{1}{n}) \longrightarrow (0, 0), \quad f(a_n) = \frac{\frac{1}{n^4}}{\frac{1}{n^4} + \frac{1}{n^4}} = \frac{1}{2} \longrightarrow \frac{1}{2} \neq f((0, 0)).$$

Folglich ist f in $p = (0, 0)$ nicht stetig.

Es existieren aber alle Richtungsableitungen $\nabla_{\mathbf{a}}f(p)$. Ist $\mathbf{a} = (0, 0)$, so folgt aus der Definition sofort $\nabla_{\mathbf{a}}f(x_0) = 0$. Für einen beliebigen Vektor $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$ mit $\mathbf{a} \neq (0, 0)$ gilt:

$$\frac{f(p + t\mathbf{a}) - f(p)}{t} = \frac{f(t\mathbf{a})}{t} = \frac{t^3 a_1 a_2^2}{t(t^2 a_1^2 + t^4 a_2^4)} = \frac{a_1 a_2^2}{a_1^2 + t^2 a_2^4} \xrightarrow{t \rightarrow 0} \begin{cases} \frac{a_2^2}{a_1} & a_1 \neq 0, \\ 0 & a_1 = 0. \end{cases}$$

Somit existiert $\nabla_{\mathbf{a}}f(p)$ für alle $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$.

Wichtige Richtungsableitungen sind die Richtungsableitungen in Richtung der Koordinatenachsen, d.h. in Richtung der kanonischen Basisvektoren e_1, \dots, e_k des \mathbb{R}^k .

Definition 8.4. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine Abbildung und $p \in U$.

Man sagt: f besitzt in $p \in U$ die partielle Ableitung nach der i -ten Koordinate, wenn die Richtungsableitung $\nabla_{e_i}f(p)$ von f in p in Richtung e_i existiert.

Bezeichnung:

$$\frac{\partial f}{\partial x_i}(p) := \nabla_{e_i}f(p) = \lim_{t \rightarrow 0} \frac{f(p + te_i) - f(p)}{t} \in \mathbb{R}^m.$$

f heißt partiell differenzierbar, wenn alle partiellen Ableitungen

$\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_k} : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ existieren.

Sei $p = (p_1, \dots, p_k)$ und $\varphi_i : (p_i - \varepsilon, p_i + \varepsilon) \subset \mathbb{R} \rightarrow \mathbb{R}^m$ die Funktion

$$\varphi_i(x) := f(p_1, \dots, p_{i-1}, x, p_{i+1}, \dots, p_k).$$

Dann gilt nach Definition

$$\frac{\partial f}{\partial x_i}(p) = \varphi_i'(p_i).$$

Dies motiviert den Namen *partielle Ableitung nach der i -ten Koordinate*. Man hält alle Koordinaten außer der i -ten fest und leitet im üblichen Sinne nach der i -ten Koordinate ab.

Beispiel 4: Wir betrachten die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, definiert durch $f(x, y) = \sin(xy^2)$. Die partiellen Ableitungen sind dann

$$\frac{\partial f}{\partial x}(x, y) = y^2 \cos(xy^2) \quad \text{und} \quad \frac{\partial f}{\partial y}(x, y) = 2xy \cos(xy^2).$$

Satz 8.3 Ist $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ in $p \in U$ differenzierbar, so existieren die partiellen Ableitungen $\frac{\partial f}{\partial x_1}(p), \dots, \frac{\partial f}{\partial x_k}(p)$ und es gilt

$$df_p(h) = \sum_{i=1}^k h_i \cdot \frac{\partial f}{\partial x_i}(p), \quad \text{wobei } h = (h_1, \dots, h_k) \in \mathbb{R}^k. \quad (8.6)$$

Beweis. Nach Satz 8.2 wissen wir, dass die Richtungsableitungen $\nabla_{e_i} f(p)$, und somit die partiellen Ableitungen von f in $p \in U$ existieren. Außerdem folgt aus der Definition der partiellen Ableitungen für $h = (h_1, \dots, h_k) = \sum_{i=1}^k h_i e_i$:

$$df_p(h) = df_p\left(\sum_{i=1}^k h_i e_i\right) = \sum_{i=1}^k h_i \cdot df_p(e_i) = \sum_{i=1}^k h_i \cdot \frac{\partial f}{\partial x_i}(p). \quad \square$$

Mit der Formel (8.6) kann man die Differentiale einer *differenzierbaren* Abbildung mit Hilfe der partiellen Ableitungen ausrechnen. Es gilt aber noch mehr! Die partiellen Ableitungen sind ein wichtiges Hilfsmittel, um festzustellen, ob eine Abbildung $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ überhaupt differenzierbar ist. Wir wissen aus Beispiel 3, dass die partiellen Ableitungen existieren können, ohne dass die Abbildung differenzierbar ist. Es gilt aber folgender starker Satz.

Satz 8.4 (Hauptkriterium für Differenzierbarkeit)

Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine Abbildung und $p \in U$.

Wenn die partiellen Ableitungen $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ (auf der gesamten Menge U !) existieren und im Punkt $p \in U$ stetig sind, dann ist f in $p \in U$ differenzierbar.

Beweis. Auf Grund von Satz 8.1, Punkt 3) genügt es, die Behauptung für reellwertige Abbildungen zu beweisen. Sei also nun $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ reellwertig, $p = (p_1, \dots, p_k)$ und $h = (h_1, \dots, h_k) \in \mathbb{R}^k$ so klein, dass $h \in K(p, \varepsilon) \subset U$. Dann gilt

$$\begin{aligned} f(p+h) - f(p) &= f(p_1+h_1, p_2, \dots, p_k) - f(p_1, \dots, p_k) \\ &\quad + f(p_1+h_1, p_2+h_2, p_3, \dots, p_k) - f(p_1+h_1, p_2, \dots, p_k) \\ &\quad + f(p_1+h_1, p_2+h_2, p_3+h_3, p_4, \dots, p_k) - f(p_1+h_1, p_2+h_2, p_3, \dots, p_k) \\ &\quad \vdots \\ &\quad + f(p_1+h_1, \dots, p_k+h_k) - f(p_1+h_1, \dots, p_{k-1}+h_{k-1}, p_k). \end{aligned}$$

Wir betrachten die Funktion $\varphi_1(x) := f(x, p_2, \dots, p_k)$. Nach Voraussetzung existiert die Ableitung $\varphi_1'(x) = \frac{\partial f}{\partial x_1}(x, p_2, \dots, p_k)$ für alle x zwischen p_1 und p_1+h_1 . Wir wenden den Mittelwertsatz der Differentialrechnung auf φ_1 an, und erhalten eine Zahl ξ_1 zwischen p_1 und p_1+h_1 , so dass gilt

$$f(p_1+h_1, p_2, \dots, p_k) - f(p_1, \dots, p_k) = h_1 \cdot \frac{\partial f}{\partial x_1}(\xi_1, p_2, \dots, p_k).$$

Analog existiert für alle $j \in \{2, \dots, k\}$ ein ξ_j zwischen p_j und p_j+h_j , so dass gilt

$$\begin{aligned} f(p_1+h_1, \dots, p_j+h_j, p_{j+1}, \dots, p_k) - f(p_1+h_1, \dots, p_{j-1}+h_{j-1}, p_j, p_{j+1}, \dots, p_k) \\ = h_j \cdot \frac{\partial f}{\partial x_j}(\underbrace{p_1+h_1, \dots, p_{j-1}+h_{j-1}, \xi_j, p_{j+1}, \dots, p_k}_{=: c_j=c_j(h)}). \end{aligned}$$

Folglich gilt

$$f(p+h) - f(p) = \sum_{j=1}^k h_j \cdot \frac{\partial f}{\partial x_j}(c_j) = \underbrace{\sum_{j=1}^k h_j \cdot \frac{\partial f}{\partial x_j}(p)}_{=: L(h), L \text{ linear}} + \underbrace{\sum_{j=1}^k h_j \cdot \left(\frac{\partial f}{\partial x_j}(c_j) - \frac{\partial f}{\partial x_j}(p) \right)}_{=: r_p(h)}.$$

Es bleibt zu zeigen, dass $\lim_{h \rightarrow 0} \frac{r_p(h)}{\|h\|} = 0$. Mit der Cauchy-Schwarzschen Ungleichung erhalten wir:

$$|r_p(h)| = \left| \sum_{j=1}^k h_j \cdot \underbrace{\left(\frac{\partial f}{\partial x_j}(c_j) - \frac{\partial f}{\partial x_j}(p) \right)}_{=: b_j=b_j(h)} \right| \stackrel{CSU}{\leq} \|h\| \cdot \|b\|,$$

wobei $b := (b_1, \dots, b_k)$. Folglich gilt

$$0 \leq \frac{|r_p(h)|}{\|h\|} \leq \|b(h)\|.$$

Für $h \rightarrow 0$ gilt $c_j(h) \rightarrow p$ und somit wegen der Stetigkeit der partiellen Ableitungen in p auch $b_j(h) \rightarrow 0$, also $\|b(h)\| \rightarrow 0$. Daraus folgt

$$\lim_{h \rightarrow 0} \frac{|r_p(h)|}{\|h\|} = 0.$$

Folglich ist f in $p \in U$ differenzierbar und es gilt

$$df_p(h) = \sum_{j=1}^k h_j \cdot \frac{\partial f}{\partial x_j}(p). \quad \square$$

Geometrische Bedeutung der Ableitung

1. Zur Erinnerung: Für eine reelle differenzierbare Funktion, die von einer Variablen abhängt, beschreibt man mit Hilfe der Ableitung die Tangente an den Graphen von f .

Sei $f : U \subset \mathbb{R} \rightarrow \mathbb{R}$ eine differenzierbare Funktion. Dann ist ihr Graph eine Kurve im \mathbb{R}^2 :

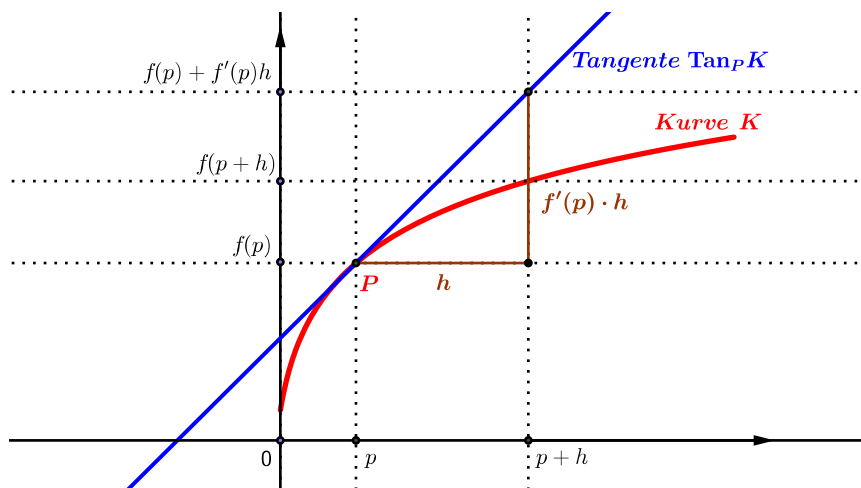
$$K := \{(x, f(x)) \mid x \in U\} \subset \mathbb{R}^2.$$

Sei $p \in U$. Die Tangente an K im Punkt $P := (p, f(p))$ ist der Graph der Funktion $T_1(f, p) : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$T_1(f, p)(x) := f(p) + f'(p)(x - p) \quad (1. \text{ Taylorpolynom} = \text{Tangentenfunktion}),$$

d.h.

$$\begin{aligned} \text{Tan}_P K &= \{(x, f(p) + f'(p)(x - p)) \mid x \in \mathbb{R}\} \\ &\stackrel{h=x-p}{=} \{(p, f(p)) + h \cdot (1, f'(p)) \mid h \in \mathbb{R}\} \\ &= P + \mathbb{R} \cdot (1, f'(p)) \subset \mathbb{R}^2. \end{aligned}$$



2. Für eine reelle differenzierbare Funktion, die von zwei Variablen abhängt, beschreibt man mit Hilfe der Ableitung die Tangentialebene an den Graphen von f .

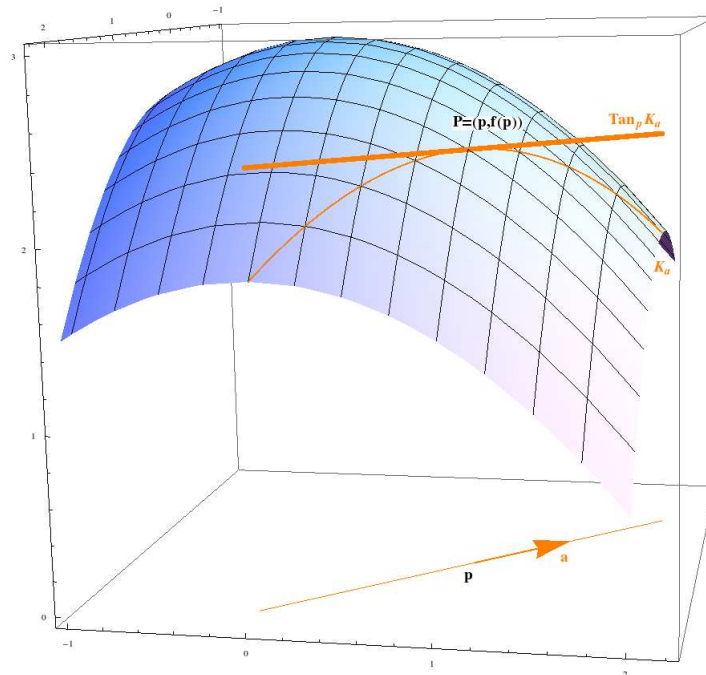
Sei $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ eine differenzierbare Funktion. Der Graph von f ist in diesem Fall eine Fläche im \mathbb{R}^3 :

$$F := \{(x, y, f(x, y)) \mid (x, y) \in U\} \subset \mathbb{R}^3.$$

Sei $p \in U$ und $df_p : \mathbb{R}^2 \rightarrow \mathbb{R}$ das Differential von f in p . Für jeden Vektor $\mathbf{a} \in \mathbb{R}^2$ mit $\mathbf{a} \neq 0$ betrachten wir die Funktion

$$\varphi_{\mathbf{a}} : t \in (-\varepsilon, \varepsilon) \mapsto f(p + t\mathbf{a}),$$

die das Verhalten von f entlang des Geradenstückes $\{p + t\mathbf{a} \mid t \in (-\varepsilon, \varepsilon)\} \subset U$ beschreibt.

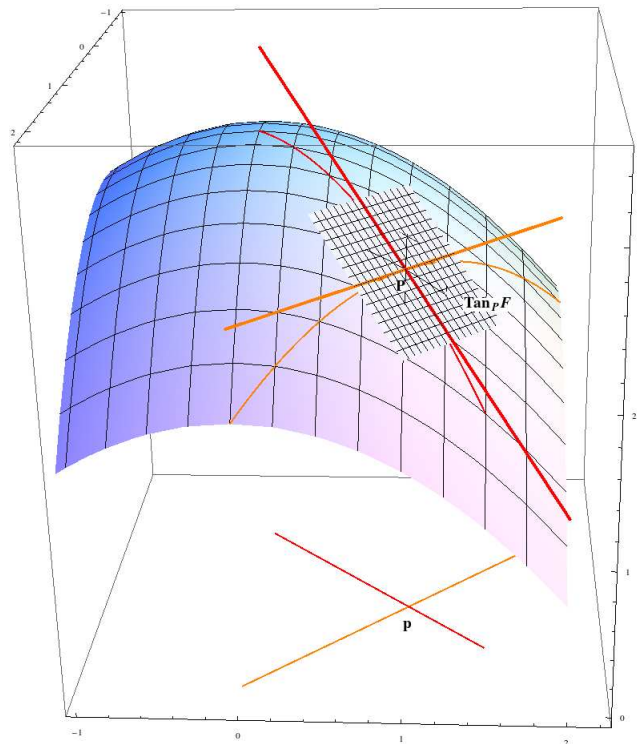


Der Graph $K_{\mathbf{a}}$ von $\varphi_{\mathbf{a}}$ ist eine Kurve auf der Fläche F , die durch den Punkt $P = (p, f(p)) = (p, \varphi_{\mathbf{a}}(0))$ läuft. Wir wenden 1) auf die Funktion $\varphi_{\mathbf{a}}$ an und erhalten für die Tangente an $K_{\mathbf{a}}$ im Punkt $P = (p, f(p))$

$$\begin{aligned} \text{Tan}_P K_{\mathbf{a}} &= \{(p + t\mathbf{a}, \varphi_{\mathbf{a}}(0) + \varphi'_{\mathbf{a}}(0) \cdot t \mid t \in \mathbb{R}\} \\ &= \{(p + t\mathbf{a}, f(p) + \nabla_{\mathbf{a}} f(p) \cdot t \mid t \in \mathbb{R}\} \\ &= \{(p, f(p)) + t \cdot (\mathbf{a}, df_p(\mathbf{a})) \mid t \in \mathbb{R}\} \\ &= P + \mathbb{R} \cdot (\mathbf{a}, df_p(\mathbf{a})). \end{aligned}$$

Die Ebene, die von allen diesen Tangenten $\text{Tan}_P K_{\mathbf{a}}$, für $\mathbf{a} \in \mathbb{R}^2$, gebildet wird, nennt man die *Tangentialebene an die Fläche F im Punkt P* und bezeichnet sie mit $\text{Tan}_P F$. Für die Tangentialebene gilt also:

$$\begin{aligned} \text{Tan}_P F &= P + \{(\mathbf{a}, df_p(\mathbf{a})) \mid \mathbf{a} \in \mathbb{R}^2\} \\ &\stackrel{\text{Basis } e_1, e_2}{=} P + \mathbb{R} \cdot (1, 0, \frac{\partial f}{\partial x}(p)) + \mathbb{R} \cdot (0, 1, \frac{\partial f}{\partial y}(p)). \end{aligned} \quad (8.7)$$



Die Tangentialebene an die Fläche F im Punkt P kann man ebenfalls als Graph des 1. Taylorpolynoms $T_1(f, p) : \mathbb{R}^2 \rightarrow \mathbb{R}$ beschreiben:

Das 1. Taylorpolynom ist in diesem Fall ein Polynom vom Grad 1 in den zwei Variablen x und y :

$$T_1(f, p)(x, y) := f(p) + \frac{\partial f}{\partial x}(p)(x - p_1) + \frac{\partial f}{\partial y}(p)(y - p_2) \quad \begin{array}{l} \text{(1. Taylorpolynom} \\ \text{= Tangentialebenenfunktion).} \end{array}$$

Aus (8.7) folgt für die Tangentialebene durch Einsetzen der Parameter $x - p_1$ und $y - p_2$

$$\text{Tan}_P F = \{(x, y, f(p) + \frac{\partial f}{\partial x}(p)(x - p_1) + \frac{\partial f}{\partial y}(p)(y - p_2) \mid (x, y) \in \mathbb{R}^2\}.$$

Wir werden auf die Taylorentwicklung für Abbildungen, die von mehreren Variablen abhängen, in Abschnitt 8.4 zurückkommen.

8.2 Kettenregeln und Mittelwertsätze

Für die Differentiale gelten die gleichen Rechenregeln wie für die Ableitungen von Funktionen einer reellen Variablen.

Satz 8.5 (Rechenregeln für Differentiale)

Seien $f, g : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$, $\varphi, \psi : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ und $p \in U$.

1. Sind f und g in p differenzierbar, so ist $f + g$ in p differenzierbar und es gilt

$$d(f + g)_p = df_p + dg_p.$$

2. Sind f und φ in p differenzierbar, so ist $\varphi \cdot f$ in p differenzierbar und es gilt

$$d(\varphi \cdot f)_p = d\varphi_p \cdot f(p) + \varphi(p) \cdot df_p.$$

3. Sind f und ψ in p differenzierbar und $\psi(p) \neq 0$, so ist $\frac{f}{\psi}$ in einer Umgebung von p definiert, in p differenzierbar und es gilt

$$d\left(\frac{f}{\psi}\right)_p = \frac{df_p \cdot \psi(p) - f(p) \cdot d\psi_p}{\psi(p)^2}.$$

Beweis. Die Aussagen 1) und 2) beweist man durch Nachprüfen der Bedingung (8.4). Die Aussage 3) folgt aus der Produktregel 2) und der bereits bekannten Quotientenregel für Funktionen einer Variablen:

$$d\left(\frac{1}{\psi}\right)_p(h) = \left(\frac{1}{\psi}\right)'(p) \cdot h = -\frac{\psi'(p)}{\psi(p)^2} \cdot h = -\frac{d\psi_p(h)}{\psi(p)^2}.$$

Wir lassen die Ausführung als Übungsaufgabe. □

Satz 8.6 (Kettenregel für Funktionen mehrerer Variablen)

Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ in p differenzierbar, $g : V \subset \mathbb{R}^m \rightarrow \mathbb{R}^n$ in $q := f(p)$ differenzierbar und $f(U) \subset V$. Dann ist die Verknüpfung $g \circ f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^n$ in p differenzierbar und es gilt

$$d(g \circ f)_p = dg_{f(p)} \circ df_p.$$

Für $k = 1$ gilt insbesondere: $(g \circ f)'(p) = dg_{f(p)}(f'(p))$.

Beweis. Wir setzen $L := df_p$ und $\tilde{L} := dg_q$ und betrachten die Abbildungen

$$\varphi(x) := f(x) - f(p) - L(x - p),$$

$$\psi(y) := g(y) - g(q) - \tilde{L}(y - q),$$

$$\varrho(x) := (g \circ f)(x) - (g \circ f)(p) - (\tilde{L} \circ L)(x - p).$$

Nach Voraussetzung gilt

$$\lim_{x \rightarrow p} \frac{\varphi(x)}{\|x - p\|} = 0 \quad \text{und} \quad \lim_{y \rightarrow q} \frac{\psi(y)}{\|y - q\|} = 0. \quad (8.8)$$

Es ist zu zeigen, dass $\lim_{x \rightarrow p} \frac{\varrho(x)}{\|x-p\|} = 0$ gilt. Da

$$\begin{aligned}\varrho(x) &= g(f(x)) - g(f(p)) - \tilde{L}(L(x-p)) \\ &= g(f(x)) - g(q) - \tilde{L}(f(x) - f(p)) - \varphi(x) \\ &= g(f(x)) - g(q) - \tilde{L}(f(x) - q) + \tilde{L}(\varphi(x)) \\ &= \psi(f(x)) + \tilde{L}(\varphi(x)),\end{aligned}$$

genügt es dazu $\lim_{x \rightarrow p} \frac{\|\tilde{L}(\varphi(x))\|}{\|x-p\|} = 0$ und $\lim_{x \rightarrow p} \frac{\|\psi(f(x))\|}{\|x-p\|} = 0$ zu zeigen. Wir benutzen dazu die Abschätzung von $\|L(x)\|$ und $\|\tilde{L}(\varphi(x))\|$ durch die Operatornormen. Da $L: \mathbb{R}^k \rightarrow \mathbb{R}^m$ und $\tilde{L}: \mathbb{R}^m \rightarrow \mathbb{R}^n$ linear sind, gilt für alle $x \in U$

$$\begin{aligned}\|L(x)\| &\leq \|L\| \cdot \|x\| \quad \text{und} \\ \|\tilde{L}(\varphi(x))\| &\leq \|\tilde{L}\| \cdot \|\varphi(x)\|,\end{aligned}$$

wobei $\|L\|$ und $\|\tilde{L}\|$ die Operatornormen von L bzw. \tilde{L} sind (siehe Übungsaufgabe 23, Analysis 2).

1) $\lim_{x \rightarrow p} \frac{\|\tilde{L}(\varphi(x))\|}{\|x-p\|} = 0$, denn

$$0 \leq \lim_{x \rightarrow p} \frac{\|\tilde{L}(\varphi(x))\|}{\|x-p\|} \leq \|\tilde{L}\| \cdot \underbrace{\lim_{x \rightarrow p} \frac{\|\varphi(x)\|}{\|x-p\|}}_{=0 \text{ nach (8.8)}} = 0.$$

2) $\lim_{x \rightarrow p} \frac{\|\psi(f(x))\|}{\|x-p\|} = 0$:

Nach (8.8) gilt $\lim_{y \rightarrow q} \frac{\|\psi(y)\|}{\|y-q\|} = 0$. Folglich existiert für jedes $\varepsilon > 0$ ein $\delta > 0$, so dass

$$\|\psi(f(x))\| < \varepsilon \cdot \|f(x) - q\| \quad \text{für alle } x \in U \text{ mit } \|f(x) - q\| < \delta.$$

Da f in p stetig ist, existiert ein $\delta_1 > 0$, so dass

$$\|f(x) - q\| < \delta \quad \text{für alle } x \in U \text{ mit } \|x - p\| < \delta_1.$$

Für $x \in U$ mit $\|x - p\| < \delta_1$ gilt also

$$\begin{aligned}\|\psi(f(x))\| &< \varepsilon \cdot \|f(x) - q\| = \varepsilon \cdot \|\varphi(x) + L(x-p)\| \\ &< \varepsilon \cdot \|\varphi(x)\| + \varepsilon \cdot \|L\| \cdot \|x-p\|\end{aligned}$$

und damit

$$\frac{\|\psi(f(x))\|}{\|x-p\|} < \varepsilon \cdot \underbrace{\frac{\|\varphi(x)\|}{\|x-p\|}}_{\rightarrow 0 \text{ für } x \rightarrow p} + \varepsilon \cdot \|L\| \quad \forall \|x-p\| < \delta_1.$$

Wir erhalten daraus

$$0 \leq \limsup_{x \rightarrow p} \frac{\|\psi(f(x))\|}{\|x-p\|} \leq \varepsilon \cdot \|L\| \quad \forall \varepsilon > 0$$

und somit

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

Damit ist die Kettenregel bewiesen.

Ist $k = 1$, so ist $g \circ f : U \subset \mathbb{R} \rightarrow \mathbb{R}^n$ eine Funktion, die nur von *einer* reellen Variablen abhängt. Dann erhalten wir die folgende spezielle Form der Kettenregel für die Ableitung in p :

$$(g \circ f)'(p) = (g \circ f)'(p) \cdot 1 = d(g \circ f)_p(1) = dg_{f(p)}(df_p(1)) = dg_{f(p)}(f'(p)).$$

□

Satz 8.7 (Kettenregel für partielle Ableitungen)

Seien $g_1, \dots, g_n : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ reellwertige Funktionen, die in $p \in U$ differenzierbar sind. Sei $V \subset \mathbb{R}^n$ eine offene Menge mit $(g_1, \dots, g_n)(U) \subset V$ und $f : V \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine in $(g_1(p), \dots, g_n(p))$ differenzierbare Abbildung. Dann ist die Abbildung $F : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$,

$$F(x) := f(g_1(x), \dots, g_n(x)),$$

in $p \in U$ differenzierbar und für ihre partiellen Ableitungen gilt

$$\frac{\partial F}{\partial x_i}(p) = \sum_{j=1}^n \frac{\partial f}{\partial y_j}(g_1(p), \dots, g_n(p)) \cdot \frac{\partial g_j}{\partial x_i}(p), \quad i = 1, \dots, k.$$

Ist $k = 1$, so gilt insbesondere:

$$F'(p) = \sum_{j=1}^n \frac{\partial f}{\partial y_j}(g_1(p), \dots, g_n(p)) \cdot g'_j(p).$$

Beweis. Wir betrachten $g : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^n$, definiert durch $g(x) := (g_1(x), \dots, g_n(x))$. Dann gilt $F = f \circ g$ und nach Kettenregel für die Differentiale $dF_p = df_{g(p)} \circ dg_p$. Für die partiellen Ableitungen folgt daraus

$$\begin{aligned} \frac{\partial F}{\partial x_i}(p) &= dF_p(e_i) = df_{g(p)}(dg_p(e_i)) = df_{g(p)}\left(\sum_{j=1}^n (dg_j)_p(e_i) \cdot e_j\right) \\ &= df_{g(p)}\left(\sum_{j=1}^n \frac{\partial g_j}{\partial x_i}(p) \cdot e_j\right) = \sum_{j=1}^n \frac{\partial g_j}{\partial x_i}(p) \cdot df_{g(p)}(e_j) \\ &= \sum_{j=1}^n \frac{\partial g_j}{\partial x_i}(p) \cdot \frac{\partial f}{\partial y_j}(g(p)). \end{aligned}$$

□

Mit Hilfe der Kettenregeln kann man den Mittelwertsatz für reelle Funktionen (Satz 5.8) verallgemeinern.

Satz 8.8 (Mittelwertsätze für Funktionen mehrerer reeller Variablen)

1. Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ eine differenzierbare reellwertige Funktion und $x, y \in U$ zwei Punkte in U , für die die gesamte Strecke \overline{xy} zwischen x und y ebenfalls in U liegt. Dann existiert ein $\xi \in \overline{xy}$ mit $\xi \neq x, \xi \neq y$, so dass

$$f(y) - f(x) = df_\xi(y - x).$$

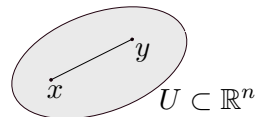
2. Sei $f : [a, b] \rightarrow \mathbb{R}^n$ stetig und auf (a, b) differenzierbar, dann existiert ein $\eta \in (a, b)$, so dass

$$\|f(b) - f(a)\| \leq \|f'(\eta)\| \cdot (b - a).$$

Beweis. Zu 1) Wir betrachten die reelle Funktion $\varphi : [0, 1] \rightarrow \mathbb{R}$ definiert durch $\varphi(t) := f(x + t(y - x))$.

Nach Voraussetzung ist φ auf $[0, 1]$ differenzierbar. Daher existiert nach dem Mittelwertsatz von Lagrange ein $\theta \in (0, 1)$ mit

$$\varphi(1) - \varphi(0) = \varphi'(\theta) \cdot (1 - 0) = \varphi'(\theta).$$



Nach Definition von h und aufgrund der Kettenregel ist dies äquivalent zu

$$f(y) - f(x) = h(1) - h(0) = h'(\theta) = \underbrace{df_{x + \theta(y - x)}}_{=: \xi} (y - x) = df(\xi)(y - x).$$

Zu 2) Wir betrachten die reelle Funktion $\phi : [a, b] \rightarrow \mathbb{R}$, definiert durch das Skalarprodukt

$$\phi(t) := \langle f(b) - f(a), f(t) \rangle.$$

ϕ ist auf $[a, b]$ stetig und auf (a, b) differenzierbar. Nach dem Mittelwertsatz von Lagrange existiert ein $\eta \in (a, b)$ mit

$$\phi(b) - \phi(a) = (b - a) \cdot \phi'(\eta) = (b - a) \cdot \langle f(b) - f(a), f'(\eta) \rangle.$$

Aus der Cauchy-Schwarzschen Ungleichung folgt dann

$$|\phi(b) - \phi(a)| \leq (b - a) \cdot |\langle f(b) - f(a), f'(\eta) \rangle| \leq (b - a) \cdot \|f(b) - f(a)\| \cdot \|f'(\eta)\|.$$

Andererseits ist

$$\phi(b) - \phi(a) = \langle f(b) - f(a), f(b) - f(a) \rangle = \|f(b) - f(a)\|^2.$$

Daraus folgt die Behauptung

$$\|f(b) - f(a)\| \leq (b - a) \|f'(\eta)\|.$$

□

Wir betrachten zwei Anwendungen der Mittelwertsätze.

Satz 8.9 Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ differenzierbar und U offen und wegzusammenhängend. Dann ist die Abbildung f genau dann konstant, wenn $df_p = 0$ für alle $p \in U$.

Beweis. Wegen Satz 8.1 3) genügt es, die Behauptung für jede Komponente von f , d.h. für reellwertige Funktionen zu beweisen. Sei also oBdA $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$.

(\implies) Wenn die Abbildung f konstant ist, so gilt für die Richtungsableitung in jedem Punkt $p \in U$ und in jeder Richtung $\mathbf{a} \in \mathbb{R}^k$

$$df_p(\mathbf{a}) = \nabla_{\mathbf{a}} f(p) = \lim_{t \rightarrow 0} \frac{f(p + t\mathbf{a}) - f(p)}{t} = \lim_{t \rightarrow 0} \frac{0}{t} = 0.$$

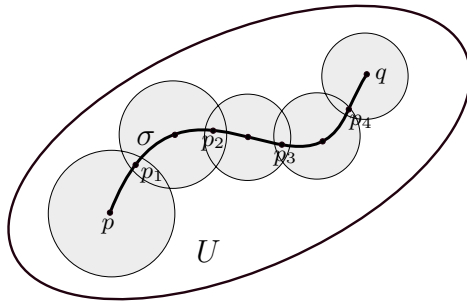
Somit ist $df_p = 0$ für alle $p \in U$.

(\impliedby) Sei $df_p = 0$ für alle $p \in U$ und sei $x_0 \in U$ ein fixierter Punkt. Da U offen ist, existiert ein $\varepsilon(x_0) > 0$, so dass $K(x_0, \varepsilon(x_0)) \subset U$. Die Kugel ist konvex, folglich liegt für jedes $z \in K(x_0, \varepsilon(x_0))$ mit $z \neq x_0$ die Strecke $\overline{x_0 z}$ ebenfalls in $K(x_0, \varepsilon(x_0))$. Nach Satz 8.8 existiert ein $\xi \in \overline{x_0 z}$ mit

$$f(z) - f(x_0) = df_{\xi}(z - x_0) \stackrel{Vor.}{=} 0.$$

Folglich ist f auf der Kugel $K(x_0, \varepsilon(x_0))$ konstant mit dem Wert $f(x_0)$.

Seien nun $p, q \in U$ beliebige, aber fixierte Punkte. Da U wegzusammenhängend ist, existiert eine stetige Abbildung $\sigma : [0, 1] \rightarrow U \subset \mathbb{R}^k$ mit $\sigma(0) = p$ und $\sigma(1) = q$. Die Menge der Kugeln $\{K(\sigma(t), \varepsilon(\sigma(t)))\}_{t \in [0, 1]}$ überdeckt $\sigma([0, 1])$. Da $[0, 1]$ kompakt und σ stetig ist, ist auch $\sigma([0, 1])$ kompakt.



Folglich findet man endlich viele dieser Kugeln $K(\sigma(t_1), \varepsilon_1), \dots, K(\sigma(t_r), \varepsilon_r)$, die die Menge $\sigma([0, 1])$ bereits überdecken (Übungsaufgabe). Auf jeder dieser Kugeln gilt aber

$$f|_{K(\sigma(t_j), \varepsilon_j)} = \text{konst.} = f(\sigma(t_j)).$$

Daraus folgt

$$f(p) = f(\sigma(t_1)) = f(p_1) = f(\sigma(t_2)) = f(p_2) = \dots = f(\sigma(t_r)) = f(q),$$

mit $p_i \in K(\sigma(t_i), \varepsilon_i) \cap K(\sigma(t_{i+1}), \varepsilon_{i+1})$. Folglich ist f auf U konstant. □

Man kann zeigen, dass offene und zusammenhängende Teilmengen $U \subset \mathbb{R}^k$ auch wegzusammenhängend sind. Satz 8.9 gilt deshalb auch für offene, zusammenhängende Mengen $U \subset \mathbb{R}^k$.

Satz 8.10 Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine differenzierbare Abbildung und U eine offene und konvexe Menge. Es existiere eine Konstante $C \in \mathbb{R}^+$, so dass

$$\|df_p\| \leq C \quad \forall p \in U,$$

wobei $\|df_p\|$ die Operatornorm von df_p bezeichnet. Dann gilt

$$\|f(y) - f(x)\| \leq C \cdot \|y - x\| \quad \forall x, y \in U.$$

Insbesondere ist f auf U Lipschitzstetig.

Beweis. Seien $x, y \in U$ fixiert. Wir parametrisieren die Strecke von x nach y durch $\gamma : [0, 1] \rightarrow U$ mit $\gamma(t) := x + t(y - x)$ und betrachten die Abbildung $g : [0, 1] \rightarrow \mathbb{R}^m$ mit $g(t) := f(\gamma(t))$. Dann gilt nach Kettenregel

$$g'(t) = df_{x+t(y-x)}(y - x)$$

und daher

$$\|g'(t)\| \leq \|df_{x+t(y-x)}\| \cdot \|y - x\| \leq C \cdot \|y - x\| \quad \forall t \in [0, 1].$$

Aus dem Mittelwertsatz (Satz 8.8, 2) folgt die Existenz einer Zahl $\eta \in (0, 1)$ mit

$$\|f(y) - f(x)\| = \|g(1) - g(0)\| \leq \|g'(\eta)\| \leq C \cdot \|y - x\|.$$

□

8.3 Die Jacobi-Matrix, der Gradient und die Hesse-Matrix

Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine in $p \in U$ differenzierbare Abbildung und $df_p : \mathbb{R}^k \rightarrow \mathbb{R}^m$ ihr Differential in p . Aus der Vorlesung *Lineare Algebra* ist bekannt, dass man lineare Abbildungen durch Matrizen beschreiben kann, wenn man im Urbild- und im Bildvektorraum jeweils eine Basis fixiert. Wir können das Differential df_p also auch durch eine Matrix angeben.

Wir fixieren dazu in den Vektorräumen \mathbb{R}^k und \mathbb{R}^m jeweils die kanonischen Basen (e_1, \dots, e_k) bzw. $(\hat{e}_1, \dots, \hat{e}_m)$. Dann entspricht jeder linearen Abbildung $A : \mathbb{R}^k \rightarrow \mathbb{R}^m$ die $(m \times k)$ -Matrix $[A] := (A_{ij})$, definiert durch

$$A(e_j) =: \sum_{i=1}^m A_{ij} \hat{e}_i \quad j = 1, \dots, k.$$

In der j -ten Spalte der Matrix $[A]$ stehen die Komponenten des Vektors $A(e_j) \in \mathbb{R}^m$ bzgl. der kanonischen Basis $(\hat{e}_1, \dots, \hat{e}_m)$ im \mathbb{R}^m . Der Anwendung der linearen Abbildung A auf den Vektor $x = \sum_{j=1}^k x_j e_j = (x_1, \dots, x_k) \in \mathbb{R}^k$ entspricht dabei die Anwendung der Matrix $[A] = (A_{ij})$ auf den Spaltenvektor $(x_1, \dots, x_k)^t$:

$$\begin{array}{ccc} \mathbb{R}^k & \xrightarrow{A} & \mathbb{R}^m \\ \text{transponieren} = t \downarrow & & \downarrow t = \text{transponieren} \\ \mathbb{R}^k & \xrightarrow{[A]} & \mathbb{R}^m \end{array} .$$

Sei $f = (f_1, \dots, f_m)$ die Komponentendarstellung von f . Dann gilt für das Differential $df_p : \mathbb{R}^k \rightarrow \mathbb{R}^m$:

$$df_p(e_j) = \frac{\partial f}{\partial x_j}(p) = \sum_{i=1}^m \frac{\partial f_i}{\partial x_j}(p) \cdot \hat{e}_i \quad j = 1, \dots, k.$$

Die zu df_p gehörende Matrix $[df_p]$ hat also die Form

$$[df_p] = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \frac{\partial f_1}{\partial x_2}(p) & \cdots & \frac{\partial f_1}{\partial x_k}(p) \\ \frac{\partial f_2}{\partial x_1}(p) & \frac{\partial f_2}{\partial x_2}(p) & \cdots & \frac{\partial f_2}{\partial x_k}(p) \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(p) & \frac{\partial f_m}{\partial x_2}(p) & \cdots & \frac{\partial f_m}{\partial x_k}(p) \end{pmatrix} = \begin{pmatrix} \frac{\partial f^t}{\partial x_1}(p) & \frac{\partial f^t}{\partial x_2}(p) & \cdots & \frac{\partial f^t}{\partial x_k}(p) \end{pmatrix}.$$

Definition 8.5. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine Abbildung, die in $p \in U$ partiell differenzierbar ist und seien $f = (f_1, \dots, f_m)$ ihre Komponenten. Dann nennt man die Matrix

$$Jf_p := \begin{pmatrix} \frac{\partial f_i}{\partial x_j}(p) \end{pmatrix}_{\substack{i=1, \dots, m \text{ Zeilen} \\ j=1, \dots, k \text{ Spalten}}}$$

Jacobi-Matrix von f im Punkt p .

Ist $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ in $p \in U$ differenzierbar, so gilt also $[df_p] = Jf_p$.

Eine besondere Bedeutung haben die Vektoren, die in den Zeilen der Jacobi-Matrix auftreten.

Definition 8.6. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ eine reellwertige Funktion, die im Punkt $p \in U$ partiell differenzierbar ist. Dann heißt der Vektor

$$\text{grad}f(p) := \left(\frac{\partial f}{\partial x_1}(p), \frac{\partial f}{\partial x_2}(p), \dots, \frac{\partial f}{\partial x_k}(p) \right) \in \mathbb{R}^k$$

Gradient von f im Punkt $p \in U$.

Für reellwertige partiell differenzierbare Funktionen $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ stimmt die Jacobi-Matrix also mit dem Gradienten überein:

$$Jf_p = \text{grad}f(p) = \left(\frac{\partial f}{\partial x_1}(p), \dots, \frac{\partial f}{\partial x_k}(p) \right).$$

Die Jacobi-Matrix einer partiell differenzierbaren Abbildung $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ mit den Komponenten $f = (f_1, \dots, f_m)$ schreibt sich durch die Gradienten der Komponentenfunktionen in der Form:

$$Jf_p = \begin{pmatrix} \text{grad}f_1(p) \\ \text{grad}f_2(p) \\ \vdots \\ \text{grad}f_m(p) \end{pmatrix}.$$

Der folgende Satz erklärt die geometrische Bedeutung des Gradienten einer Funktion. Ist $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ in p differenzierbar, so gilt für alle $\mathbf{a} = (a_1, \dots, a_k) \in \mathbb{R}^k$ mit dem Euklidischen Skalarprodukt $\langle \cdot, \cdot \rangle$

$$df_p(\mathbf{a}) = \sum_{j=1}^k \frac{\partial f}{\partial x_j}(p) \cdot a_j = \langle \text{grad} f(p), \mathbf{a} \rangle. \quad (8.9)$$

Wir erhalten daraus folgende geometrische Interpretation von $\text{grad} f(p)$:

Satz 8.11 Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ in $p \in U$ differenzierbar.

1. $\text{grad} f(p) = 0$ genau dann, wenn $df_p = 0$.
2. Ist $\text{grad} f(p) \neq 0$, so gibt $\text{grad} f(p) \in \mathbb{R}^k$ diejenige Richtung an, in der die Funktion f im Punkt p am schnellsten wächst, d.h. diejenige Richtung mit dem größten Funktionsanstieg.
3. Sei $c \in \mathbb{R}$ im Bild von f und bezeichne $M_c := \{x \in U \mid f(x) = c\}$ die Niveaufläche von f zum Niveau c . Ist $p \in M_c$ ein Punkt auf dieser Niveaufläche und $\gamma : I \subset \mathbb{R} \rightarrow M_c$ eine differenzierbare Kurve mit $\gamma(t_0) = p$ und Spur in M_c , dann gilt:

$$\gamma'(t_0) \perp \text{grad} f(p),$$

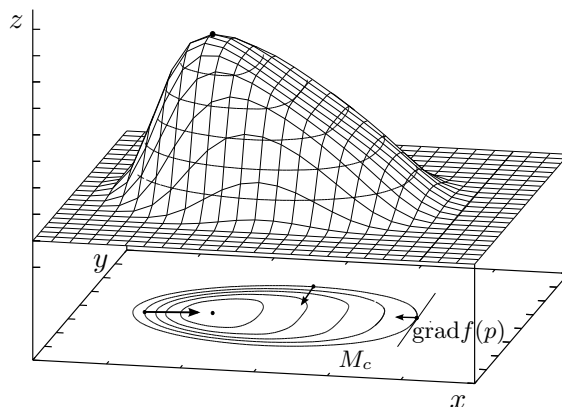
d.h. der Vektor $\text{grad} f(p)$ steht senkrecht auf allen Tangenten an die Niveaufläche M_c im Punkt p .

Beispiel: Höhenlinien

Wir betrachten zur Illustration von Satz 8.11 die Abbildung, die die Höhe eines Ortes über dem Meeresspiegel angibt:

$$\begin{aligned} f : U \subset \mathbb{R}^2 &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto \text{Höhe des Ortes über dem Meeresspiegel.} \end{aligned}$$

Dann sieht man die Niveauflächen von f als Höhenlinien auf der Landkarte.



Beweis von Satz 8.11 Zu 1) Das folgt unmittelbar aus (8.9), da das Euklidische Skalarprodukt nicht ausgeartet ist.

Zu 2) Sei $\mathbf{a} \in \mathbb{R}^n$ mit $\|\mathbf{a}\| = 1$. Dann folgt aus der Cauchy-Schwarzschen Ungleichung

$$\nabla_{\mathbf{a}}f(p) = \langle \text{grad}f(p), \mathbf{a} \rangle \stackrel{CSU}{\leq} \|\text{grad}f(p)\| \cdot \|\mathbf{a}\| = \|\text{grad}f(p)\|. \quad (*)$$

Der maximal mögliche Wert von $\nabla_{\mathbf{a}}f(p)$, wenn man die Vektoren \mathbf{a} mit $\|\mathbf{a}\| = 1$ durchläuft, kann also höchstens gleich $\|\text{grad}f(p)\|$ sein. In der Cauchy-Schwarzschen Ungleichung (*) wird aber die Gleichheit angenommen, und zwar genau dann, wenn die auftretenden Vektoren linear abhängig sind und in die gleiche Richtung zeigen, d.h. genau dann, wenn $\mathbf{a} = \frac{\text{grad}f(p)}{\|\text{grad}f(p)\|}$. Die Richtung mit dem größten Funktionsanstieg ist also diejenige, in die der Vektor $\text{grad}f(p)$ zeigt.

Zu 3) Sei $\gamma : I \rightarrow M_c$ eine differenzierbare Kurve mit $\gamma(t_0) = p$ und Spur in M_c . Dann gilt $f(\gamma(t)) = c$ für alle $t \in I$. Nach der Kettenregel ist dann

$$0 = (f \circ \gamma)'(t_0) = df_p(\gamma'(t_0)) = \langle \text{grad}f(p), \gamma'(t_0) \rangle.$$

Also steht der Gradient $\text{grad}f(p)$ senkrecht auf dem Tangentialvektor $\gamma'(t_0)$. \square

Wir definieren als nächstes die partiellen Ableitungen höherer Ordnung.

Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine Funktion, für die die partielle Ableitung $\frac{\partial f}{\partial x_i} : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ existiert. Wenn die partielle Ableitung der Funktion $\frac{\partial f}{\partial x_i} : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ nach der j -ten Koordinate existiert, so bezeichnet man sie mit

$$\frac{\partial^2 f}{\partial x_j \partial x_i} := \frac{\partial}{\partial x_j} \left(\frac{\partial f}{\partial x_i} \right).$$

Auf diese Weise entstehen partielle Ableitungen höherer Ordnung. Die Funktion

$$\frac{\partial^n f}{\partial x_{i_n} \cdots \partial x_{i_1}} := \frac{\partial}{\partial x_{i_n}} \left(\frac{\partial^{n-1} f}{\partial x_{i_{n-1}} \cdots \partial x_{i_1}} \right) : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$$

heißt (falls sie existiert) n -te partielle Ableitung von f nach den Variablen x_{i_1}, \dots, x_{i_n} .

Leitet man n -mal nach der gleichen Variablen ab, so schreibt man zur Abkürzung

$$\frac{\partial^n f}{\partial x_i \cdots \partial x_i} =: \frac{\partial^n f}{\partial x_i^n}.$$

Definition 8.7. Eine Funktion $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ heißt n -mal stetig differenzierbar (C^n -Funktion oder von der Klasse C^n) mit $n \in \mathbb{N}$, wenn alle partiellen Ableitungen von f der Ordnung $\leq n$ existieren und stetig sind.

Bezeichnung: $C^n(U, \mathbb{R}^m) := \{f : U \rightarrow \mathbb{R}^m \mid f \text{ } C^n\text{-Funktion}\}.$

Offensichtlich gilt für $n \in \mathbb{N}$

$$C^n(U, \mathbb{R}^m) \subset C^{n-1}(U, \mathbb{R}^m) \subset C^{n-2}(U, \mathbb{R}^m) \subset \dots \subset C^0(U, \mathbb{R}^m).$$

Beispiel: Sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ die Funktion $f(x, y) := x^3y^2$.

Dann erhalten wir für die partiellen Ableitungen erster und zweiter Ordnung

$$\begin{aligned}\frac{\partial f}{\partial x}(x, y) &= 3x^2y^2, & \frac{\partial f}{\partial y}(x, y) &= 2x^3y \\ \frac{\partial^2 f}{\partial x^2}(x, y) &= 6xy^2, & \frac{\partial^2 f}{\partial y^2}(x, y) &= 2x^3, & \frac{\partial^2 f}{\partial y\partial x}(x, y) &= 6x^2y = \frac{\partial f}{\partial x\partial y}(x, y).\end{aligned}$$

f ist also 2-fach stetig differenzierbar. In diesem Beispiel sind die gemischten 2. partiellen Ableitungen unabhängig von der Reihenfolge der Differentiation.

Im Allgemeinen hängen die partiellen Ableitungen von der Reihenfolge der Variablen ab, nach denen abgeleitet wird (siehe Übungsaufgabe 29). Der folgende Satz gibt eine hinreichende Bedingung dafür an, dass man die Reihenfolge der partiellen Ableitungen vertauschen kann.

Satz 8.12 (Lemma von Schwarz)

Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}^m$ eine n -mal stetig differenzierbare Abbildung. Dann sind alle partiellen Ableitungen der Ordnung $\leq n$ unabhängig von der Reihenfolge des Differenzierens. Insbesondere gilt für $f \in C^2(U, \mathbb{R}^m)$

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

Beweis. Es genügt wieder, die Behauptung für reellwertige Funktionen zu zeigen. Außerdem genügt es, die Behauptung für die 2. partiellen Ableitungen zu zeigen. Die Behauptung für die partiellen Ableitungen höherer Ordnung beweist man dann durch Induktion. Bei den 2. partiellen Ableitungen läßt man bis auf zwei Variablen alle anderen fest. Es genügt also für jede C^2 -Funktion $f : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial^2 f}{\partial x_2 \partial x_1}$$

zu zeigen. Das werden wir jetzt tun. Sei $p = (p_1, p_2) \in U$ und seien $h = (h_1, h_2) \in \mathbb{R}^2$ so gewählt, dass $h_1, h_2 \neq 0$ und $h \in K(p, \varepsilon) \subset U$.

Wir betrachten die Funktion $\varphi(x) := f(x, p_2 + h_2) - f(x, p_2)$. Dann ist φ auf dem Intervall $[p_1 - |h_1|, p_1 + |h_1|]$ differenzierbar. Nach dem Mittelwertsatz existiert ein ξ_1 zwischen p_1 und $p_1 + h_1$, so dass gilt

$$\varphi(p_1 + h_1) - \varphi(p_1) = h_1 \cdot \varphi'(\xi_1).$$

Sei nun

$$F(h_1, h_2) := f(p_1 + h_1, p_2 + h_2) - f(p_1 + h_1, p_2) - f(p_1, p_2 + h_2) + f(p_1, p_2).$$

Dann gilt

$$\begin{aligned}F(h_1, h_2) &= \varphi(p_1 + h_1) - \varphi(p_1) = h_1 \cdot \varphi'(\xi_1) \\ &= h_1 \left(\frac{\partial f}{\partial x_1}(\xi_1, p_2 + h_2) - \frac{\partial f}{\partial x_1}(\xi_1, p_2) \right).\end{aligned}$$

Da die Funktion $\frac{\partial f}{\partial x_1}$ auf U nach der 2. Variable differenzierbar ist, können wir den Mittelwertsatz auch auf die Funktion $\frac{\partial f}{\partial x_1}(\xi_1, \cdot)$ anwenden. Es existiert also ein ξ_2 zwischen p_2 und $p_2 + h_2$, so dass

$$F(h_1, h_2) = h_1 \cdot h_2 \cdot \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_1, \xi_2). \quad (\star)$$

Verfährt man analog mit der Funktion $\psi(x) := f(p_1 + h_1, x) - f(p_1, x)$, so existieren nach dem Mittelwertsatz ein $\tilde{\xi}_1$ zwischen p_1 und $p_1 + h_1$ und ein $\tilde{\xi}_2$ zwischen p_2 und $p_2 + h_2$, so dass

$$F(h_1, h_2) = h_1 \cdot h_2 \cdot \frac{\partial^2 f}{\partial x_1 \partial x_2}(\tilde{\xi}_1, \tilde{\xi}_2). \quad (\star\star)$$

Da $h_1, h_2 \neq 0$, folgt aus (\star) und $(\star\star)$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(\tilde{\xi}_1, \tilde{\xi}_2) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_1, \xi_2).$$

Bei $h = (h_1, h_2) \rightarrow (0, 0)$ konvergieren sowohl (ξ_1, ξ_2) als auch $(\tilde{\xi}_1, \tilde{\xi}_2)$ gegen $p = (p_1, p_2)$. Da die partiellen Ableitungen $\frac{\partial^2 f}{\partial x_1 \partial x_2}$ und $\frac{\partial^2 f}{\partial x_2 \partial x_1}$ in (p_1, p_2) nach Voraussetzung stetig sind, folgt mit $h = (h_1, h_2) \rightarrow (0, 0)$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(p) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(p). \quad \square$$

Definition 8.8. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ eine 2-mal stetig differenzierbare reellwertige Funktion. Die symmetrische $(k \times k)$ -Matrix

$$\text{Hess}f(p) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(p) \right)_{i,j=1,\dots,k}$$

heißt Hesse-Matrix von f im Punkt $p \in U$.

8.4 Die Taylorformel für Funktionen mehrerer reeller Variablen

Funktionen einer reellen Variablen $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$, die auf einem Intervall I definiert sind, haben wir durch ihre Taylorpolynome approximiert. Ist f $(n+1)$ -mal stetig differenzierbar und $p \in I$, so gilt auf I die Taylorformel

$$f(x) = \underbrace{f(p) + f'(p)(x-p) + \frac{f''(p)}{2!}(x-p)^2 + \dots + \frac{f^{(n)}(p)}{n!}(x-p)^n}_{=: T_n(f,p)(x) \text{ } n\text{-tes Taylorpolynom}} + \underbrace{R_n(f,p)(x)}_{n\text{-tes Restglied}$$

wobei das n -te Restglied $R_n(f,p)(x)$ z.B. in der Lagrange-Form

$$R_n(f,p)(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-p)^{n+1}$$

für ein geeignetes ξ zwischen p und x zu beschreiben ist.

Man kann auch Funktionen mehrerer reeller Variablen durch Taylorpolynome approximieren. Wie im Fall einer reellen Variablen wird sich dies als wichtiges Hilfsmittel bei der Untersuchung von Extremwerten erweisen. Um die Taylorformel für Funktionen mehrerer Variablen beschreiben zu können, führen wir zunächst einige Bezeichnungen ein.

Definition 8.9. Ein Multiindex der Länge k ist ein k -Tupel von natürlichen Zahlen $\alpha = (\alpha_1, \dots, \alpha_k)$, $\alpha_j \in \mathbb{N}_0$.

Die Zahl $|\alpha| := \alpha_1 + \dots + \alpha_k$ heißt die Ordnung des Multiindex α und die Zahl $\alpha! := (\alpha_1!) \cdot \dots \cdot (\alpha_k!)$ die Fakultät von α . Für $y = (y_1, \dots, y_k) \in \mathbb{R}^k$ sei

$$y^\alpha := y_1^{\alpha_1} \cdot \dots \cdot y_k^{\alpha_k}.$$

Für eine reellwertige Funktion $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ bezeichne (sofern existent)

$$\frac{\partial^{|\alpha|} f}{\partial x^\alpha} := \frac{\partial^{|\alpha|} f}{(\partial x_1)^{\alpha_1} \dots (\partial x_k)^{\alpha_k}}.$$

Definition 8.10. Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ eine n -mal stetig differenzierbare Funktion. Das Taylorpolynom n -ter Ordnung von f im Entwicklungspunkt $p \in U$ ist das Polynom n -ten Grades in den Unbestimmten x_1, \dots, x_k gegeben durch

$$\begin{aligned} T_n(f, p)(x) &:= \sum_{|\alpha| \leq n} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(p) \cdot (x - p)^\alpha \quad (\alpha \text{ sind hier Multiindizes der Länge } k) \\ &= f(p) + \sum_{|\alpha|=1} \frac{1}{\alpha!} \frac{\partial f}{\partial x^\alpha}(p) \cdot (x - p)^\alpha + \sum_{|\alpha|=2} \frac{1}{\alpha!} \frac{\partial^2 f}{\partial x^\alpha}(p) \cdot (x - p)^\alpha + \dots + \\ &\quad + \sum_{|\alpha|=n} \frac{1}{\alpha!} \frac{\partial^n f}{\partial x^\alpha}(p) \cdot (x - p)^\alpha. \end{aligned}$$

Es gilt:

$$\begin{aligned} |\alpha| = 1 &\iff \alpha = (0, \dots, 0, \overset{i}{1}, 0, \dots, 0) \quad \text{für } i = 1, \dots, k \\ |\alpha| = 2 &\iff \alpha = (0, \dots, 0, \overset{i}{2}, 0, \dots, 0) \quad \text{für } 1 \leq i \leq k \text{ oder} \\ &\quad \alpha = (0, \dots, 0, \overset{i}{1}, 0, \dots, 0, \overset{j}{1}, 0, \dots, 0) \quad \text{für } 1 \leq i < j \leq k \end{aligned}$$

Daraus folgt für das erste und zweite Taylorpolynom:

$$\begin{aligned} T_1(f, p)(x) &= f(p) + \sum_{i=1}^k \frac{\partial f}{\partial x_i}(p) \cdot (x_i - p_i) \\ &= f(p) + \langle \text{grad} f(p), x - p \rangle. \end{aligned} \tag{8.10}$$

$$\begin{aligned} T_2(f, p)(x) &= T_1(f, p)(x) + \frac{1}{2} \sum_{i,j=1}^k \frac{\partial^2 f}{\partial x_i \partial x_j}(p) \cdot (x_i - p_i) \cdot (x_j - p_j) \\ &= f(p) + \langle \text{grad} f(p), x - p \rangle + \frac{1}{2} (x - p) \cdot \text{Hess} f(p) \cdot (x - p)^t. \end{aligned} \tag{8.11}$$

Der Graph des ersten Taylorpolynoms beschreibt die Tangentialebene an die Hyperfläche $F := \{(x, f(x)) \mid x \in U\} \subset \mathbb{R}^{k+1}$ im Punkt $P = (p, f(p))$ (siehe Abschnitt 8.1). Der Graph des zweiten Taylorpolynoms ist die sogenannte *Schmiegequadrik an F im Punkt P* . Die Normalformen solcher Quadriken und ihre geometrische Gestalt lernen Sie in der Vorlesung *Lineare Algebra 2* kennen.

Satz 8.13 (Taylorformel n -ter Ordnung)

Sei $f : U \subset \mathbb{R}^k \rightarrow \mathbb{R}$ eine $(n + 1)$ -mal stetig differenzierbare reellwertige Funktion und $p, x \in U$ zwei Punkte, für die die Strecke \overline{px} in U liegt. Dann existiert ein $\xi \in \overline{px}$ mit $\xi \neq p, \xi \neq x$, so dass

$$f(x) = T_n(f, p)(x) + \underbrace{\sum_{|\alpha|=n+1} \frac{1}{\alpha!} \frac{\partial^{n+1} f}{\partial x^\alpha}(\xi) \cdot (x - p)^\alpha}_{=: R_n(f, p)(x) \text{ } n\text{-tes Restglied}}$$

Beweis. Wir parametrisieren die Strecke \overline{px} durch $\sigma : [0, 1] \rightarrow U$ mit $\sigma(t) := p + t(x - p)$ und betrachten die Funktion $g : [0, 1] \rightarrow \mathbb{R}$ mit $g(t) := f(\sigma(t)) = f(p + t(x - p))$. Dann ist $g \in C^{n+1}([0, 1], \mathbb{R})$. Auf g wenden wir die Taylorformel für Funktionen einer reellen Variablen an und erhalten: Es existiert ein $\theta \in (0, 1)$, so dass

$$g(1) = \sum_{j=0}^n \frac{1}{j!} g^{(j)}(0) + \underbrace{\frac{1}{(n+1)!} g^{(n+1)}(\theta)}_{\text{Lagrange-Restglied}}$$

Mit der Kettenregel für partielle Ableitungen folgt:

$$\begin{aligned} g^{(0)}(t) &= g(t) = f(\sigma(t)), \\ g^{(1)}(t) &= \sum_{i=1}^k \frac{\partial f}{\partial x_i}(\sigma(t)) \cdot (x_i - p_i) = \sum_{|\alpha|=1} \frac{\partial^\alpha f}{\partial x^\alpha}(\sigma(t)) \cdot (x - p)^\alpha, \\ g^{(2)}(t) &= \sum_{i,j=1}^k \frac{\partial^2 f}{\partial x_i \partial x_j}(\sigma(t)) \cdot (x_i - p_i) \cdot (x_j - p_j) \quad (*) \\ &= \sum_{|\alpha|=2} \frac{2!}{\alpha!} \frac{\partial^2 f}{\partial x^\alpha}(\sigma(t)) \cdot (x - p)^\alpha. \quad (**) \end{aligned}$$

Der Faktor $\frac{2!}{\alpha!}$ in $(**)$ ist notwendig, da in $(*)$ $\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}$ zweimal auftritt, während $\frac{\partial^2 f}{\partial x_i^2}$ in $(**)$ nur einmal vorhanden ist.

Analog beweist man mit der Kettenregel und etwas Kombinatorik mittels Induktion für die j -te Ableitung

$$g^{(j)}(t) = \sum_{|\alpha|=j} \frac{j!}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(\sigma(t)) \cdot (x - p)^\alpha.$$

Daraus folgt mit $\xi := p + \theta(x - p) = \sigma(\theta)$ die Taylorformel

$$f(x) = \sum_{|\alpha| \leq n} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(p) \cdot (x - p)^\alpha + \sum_{|\alpha|=n+1} \frac{1}{\alpha!} \frac{\partial^{n+1} f}{\partial x^\alpha}(\xi) \cdot (x - p)^\alpha.$$

□

8.5 Lokale Extrema für Funktionen mehrerer reeller Variablen

Für 2-mal stetig differenzierbare Funktionen $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ kennen wir Kriterien für das Vorliegen lokaler Extremwerte (siehe Satz 5.20 und Satz 5.21). Wir wollen dies jetzt auf den Fall von Funktionen mehrerer Variablen verallgemeinern.

Dazu erinnern wir zunächst an einige Kenntnisse aus der Vorlesung *Lineare Algebra*.

Definition 8.11. Sei $A = (A_{ij})$ eine symmetrische reelle $(n \times n)$ -Matrix. A heißt

- positiv definit (symbolisch $A > 0$), falls $xAx^t > 0 \quad \forall x = (x_1, \dots, x_n) \in \mathbb{R}^n, x \neq 0$,
- negativ definit (symbolisch $A < 0$), falls $xAx^t < 0 \quad \forall x = (x_1, \dots, x_n) \in \mathbb{R}^n, x \neq 0$,
- positiv semidefinit (symbolisch $A \geq 0$), falls $xAx^t \geq 0 \quad \forall x = (x_1, \dots, x_n) \in \mathbb{R}^n$,
- negativ semidefinit (symbolisch $A \leq 0$), falls $xAx^t \leq 0 \quad \forall x = (x_1, \dots, x_n) \in \mathbb{R}^n$,
- indefinit, falls A weder positiv semidefinit noch negativ semidefinit ist.

Algebraische Fakten:

1. Jede symmetrische reelle $(n \times n)$ -Matrix ist diagonalisierbar, d.h. sie hat n reelle Eigenwerte.
2. *Eigenwert-Kriterium:* Für eine symmetrische reelle $(n \times n)$ -Matrix A gilt:

$$A > 0 \iff \text{alle Eigenwerte von } A \text{ sind positiv.}$$

$$A < 0 \iff \text{alle Eigenwerte von } A \text{ sind negativ.}$$

$$A \geq 0 \iff \text{alle Eigenwerte von } A \text{ sind } \geq 0.$$

$$A \leq 0 \iff \text{alle Eigenwerte von } A \text{ sind } \leq 0.$$

3. *Determinanten-Kriterium:* Für eine symmetrische reelle $(n \times n)$ -Matrix (A_{ij}) bezeichne

$$A[k] := \begin{pmatrix} A_{11} & \dots & A_{1k} \\ \vdots & \dots & \vdots \\ A_{k1} & \dots & A_{kk} \end{pmatrix}$$

ihren k -ten Hauptminor. Dann gilt:

$$A > 0 \iff \det A[k] > 0 \quad \forall k = 1, \dots, n.$$

$$A < 0 \iff (-1)^k \cdot \det A[k] > 0 \quad \forall k = 1, \dots, n.$$

Definition 8.12. Sei $B \subset \mathbb{R}^n$ eine beliebige Teilmenge und $f : B \subset \mathbb{R}^n \rightarrow \mathbb{R}$ eine reellwertige Funktion. Man sagt:

1. f nimmt in $p \in B$ ein lokales Minimum an, falls ein $\varepsilon > 0$ existiert, so dass $f(x) \geq f(p)$ für alle $x \in B$ mit $\|x - p\| < \varepsilon$.
2. f nimmt in $p \in B$ ein isoliertes (oder striktes) lokales Minimum an, falls ein $\varepsilon > 0$ existiert, so dass $f(x) > f(p)$ für alle $x \in B$ mit $0 < \|x - p\| < \varepsilon$.
3. f nimmt in $p \in B$ ein lokales Maximum an, falls ein $\varepsilon > 0$ existiert, so dass $f(x) \leq f(p)$ für alle $x \in B$ mit $\|x - p\| < \varepsilon$.

4. f nimmt in $p \in B$ ein isoliertes (oder striktes) lokales Maximum an, falls ein $\varepsilon > 0$ existiert, so dass $f(x) < f(p)$ für alle $x \in B$ mit $0 < \|x - p\| < \varepsilon$.

Definition 8.13. Sei $U \subset \mathbb{R}^n$ offen und $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar. $p \in U$ heißt kritischer Punkt von f , wenn $df_p = 0$ (bzw. dazu äquivalent, wenn $\text{grad}f(p) = 0$). Andernfalls heißt $p \in U$ regulärer Punkt von f .

Satz 8.14 Sei $U \subset \mathbb{R}^n$ eine offene Menge und $f : U \rightarrow \mathbb{R}$ 2-mal stetig differenzierbar.

1. Hat f in $p \in U$ ein lokales Maximum (lokales Minimum), so ist p ein kritischer Punkt von f und die Hesse-Matrix $\text{Hess}f(p)$ ist negativ semidefinit (positiv semidefinit).
2. Ist $p \in U$ ein kritischer Punkt von f und die Hesse-Matrix $\text{Hess}f(p)$ negativ definit (positiv definit), so nimmt f in p ein isoliertes lokales Maximum (isoliertes lokales Minimum) an.

Beweis. Wir zeigen die Behauptungen nur für den Fall eines lokalen Maximums. Die Aussagen für das lokale Minimum folgt dann durch Übergang von f zu $-f$.

(1) f habe in $p \in U$ ein lokales Maximum. Sei $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$. Da U offen ist, existiert ein $\varepsilon > 0$, so dass $K(p, \varepsilon) \subset U$. Folglich ist die Funktion $g_{\mathbf{a}} : \left(-\frac{\varepsilon}{\|\mathbf{a}\|}, \frac{\varepsilon}{\|\mathbf{a}\|}\right) \rightarrow \mathbb{R}$ mit $g_{\mathbf{a}}(t) := f(p + t\mathbf{a})$ korrekt definiert. $g_{\mathbf{a}}$ ist 2-mal stetig differenzierbar hat in $t = 0$ ein lokales Maximum. Somit gilt $g'_{\mathbf{a}}(0) = 0$ und $g''_{\mathbf{a}}(0) \leq 0$ (siehe Satz 5.20). Aus der Kettenregel für Differentiale folgt:

$$0 = g'_{\mathbf{a}}(0) = df_p(\mathbf{a}).$$

Dies gilt für alle $\mathbf{a} \in \mathbb{R}^n$. Folglich ist $df_p = 0$, also $p \in U$ ein kritischer Punkt von f .

Aus der Kettenregel für partielle Ableitungen folgt mit $\gamma(t) := p + t\mathbf{a}$:

$$g'_{\mathbf{a}}(t) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\gamma(t)) \cdot a_i,$$

$$g''_{\mathbf{a}}(t) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\gamma(t)) \cdot a_i a_j.$$

Folglich ist

$$0 \geq g''_{\mathbf{a}}(0) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(p) \cdot a_i a_j = \mathbf{a} \cdot \text{Hess}f(p) \cdot \mathbf{a}^t.$$

Dies gilt für alle $\mathbf{a} \in \mathbb{R}^n$, somit ist die Hesse-Matrix von f in p negativ semidefinit.

(2) Sei $p \in U$ ein kritischer Punkt von f und $\text{Hess}f(p) < 0$. Wir zeigen zunächst, dass es eine Kugel $K(p, r) \subset U$ um p gibt, so dass $\text{Hess}f(\xi) < 0$ für alle $\xi \in K(p, r)$. Dazu benutzen wir das Determinanten-Kriterium für negative Definitheit. Für $k = 1, \dots, n$ bezeichne $H(\xi)[k]$ den k -ten Hauptminor der Hesse-Matrix $\text{Hess}f(\xi)$ und $\phi_k : U \rightarrow \mathbb{R}$ die Funktion

$$\phi_k(\xi) := (-1)^k \cdot \det H(\xi)[k].$$

Da f 2-mal stetig differenzierbar und die Determinanten-Funktion stetig ist, sind die Funktionen ϕ_k stetig. Da nach Voraussetzung $\text{Hess}f(p) < 0$, gilt $\phi_k(p) > 0$ für alle $k = 1, \dots, n$. Wegen der Stetigkeit der Funktionen ϕ_k existiert dann eine Kugel $K(p, r) \subset U$, so dass $\phi_k(\xi) > 0$ für alle $\xi \in K(p, r)$ und alle $k = 1, \dots, n$. Folglich ist $\text{Hess}f(\xi) < 0$ für alle $\xi \in K(p, r)$.

Wir approximieren die Funktion f nun bei p durch das Taylorpolynom 1. Grades und erhalten für alle $x \in K(p, r)$ ein $\xi \in \overline{px} \subset K(p, r)$, so dass

$$f(x) = f(p) + \langle \text{grad}f(p), x - p \rangle + \frac{1}{2}(x - p) \cdot \text{Hess}f(\xi) \cdot (x - p)^t$$

(siehe Satz 8.13). Da p ein kritischer Punkt von f ist, gilt $\text{grad}f(p) = 0$. Außerdem wissen wir aus dem oben Bewiesenen, dass die Hesse-Matrix $\text{Hess}f(\xi)$ negativ definit ist. Daraus folgt

$$f(x) < f(p)$$

für alle $x \in K(p, r)$ mit $x \neq p$. Dies zeigt, dass f in p ein isoliertes lokales Maximum annimmt. \square

Beispiel: Wir betrachten die Funktion $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ mit

$$f(x, y, z) := 35 - 6x + 2z + x^2 - 2xy + 2y^2 + 2yz + 3z^2.$$

Wir wollen untersuchen, in welchen Punkten f lokale Extremwerte annimmt. Dazu gehen wir folgendermaßen vor:

1. *Bestimme die kritischen Punkte von f :*

Sei $p = (x, y, z)$. Es gilt

$$\text{grad}f(p) = (-6 + 2x - 2y, -2x + 4y + 2z, 2 + 2y + 6z).$$

Folglich ist $p = (x, y, z)$ genau dann ein kritischer Punkt von f , wenn das Gleichungssystem

$$\begin{pmatrix} 2 & -2 & 0 \\ -2 & 4 & 2 \\ 0 & 2 & 6 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 6 \\ 0 \\ -2 \end{pmatrix}$$

erfüllt ist. Dieses hat genau eine Lösung, nämlich $p_0 = (8, 5, -2)$.

2. *Bestimme die Hesse-Matrix von f in den kritischen Punkten, also hier im Punkt p_0 :*

$$\text{Hess}f(p_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x \partial x} & \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial x \partial z} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y \partial y} & \frac{\partial^2 f}{\partial y \partial z} \\ \frac{\partial^2 f}{\partial z \partial x} & \frac{\partial^2 f}{\partial z \partial y} & \frac{\partial^2 f}{\partial z \partial z} \end{pmatrix} (p_0) = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 4 & 2 \\ 0 & 2 & 6 \end{pmatrix} =: H.$$

3. *Untersuche $\text{Hess}f(p_0)$ auf Definitheit:*

Dazu kann man die Eigenwerte von H bestimmen oder die Determinanten der Hauptminoren von H ausrechnen. Wir betrachten letzteres:

$$\det H[1] = H_{11} = 2, \quad \det H[2] = \det \begin{pmatrix} 2 & -2 \\ -2 & 4 \end{pmatrix} = 4, \quad \det H[3] = \det H = 16.$$

Alle diese Determinanten sind positiv, folglich ist $\text{Hess}f(p_0)$ positiv definit. f hat in p_0 somit ein striktes lokales Minimum. Weitere lokale Extrema existieren nicht.

Achtung: Die Kriterien von Satz 8.14 gelten nur auf *offenen* Teilmengen. Sucht man lokale bzw. globale Extremwerte einer Funktion f auf Mengen $B \subset \mathbb{R}^n$, die nicht offen sind, so wendet man die Kriterien zunächst auf die Punkte im Inneren von B an (die Menge $\text{Int}(B)$ ist offen). Anschließend muß man das Verhalten von f in den restlichen Punkten von B mit anderen Methoden untersuchen.

8.6 Koordinatentransformationen und der Satz über den lokalen Diffeomorphismus

Definition 8.14. *Es seien $U, V \subset \mathbb{R}^n$ offene Teilmengen des \mathbb{R}^n .*

Eine Abbildung $f : U \rightarrow V$ heißt Diffeomorphismus (oder Koordinatentransformation), wenn f bijektiv ist und sowohl f als auch f^{-1} stetig differenzierbar sind.

Offensichtlich ist jeder Diffeomorphismus auch ein Homöomorphismus. Die Verknüpfung zweier Diffeomorphismen ist wieder ein Diffeomorphismus.

Mit Hilfe eines Diffeomorphismus $f : U \rightarrow V$ ordnet man jedem Punkt p aus V neue Koordinaten zu, nämlich diejenigen $(x_1, \dots, x_n) \in U$ mit $f(x_1, \dots, x_n) = p$. Wir werden sehen, wie nützlich solche Koordinatentransformationen sind. Wie im eindimensionalen Fall vereinfachen sie z.B. oft die Berechnung von Integralen. In der *linearen Algebra* werden sie benutzt, um die geometrische Gestalt von Quadriken zu bestimmen (Hauptachsentransformation). Deshalb werden wir uns in diesem Abschnitt damit beschäftigen, wie man feststellen kann, ob eine Abbildung $f : U \rightarrow V$ ein Diffeomorphismus (bzw. eine Koordinatentransformation) ist. Zunächst sehen wir uns einige Beispiele an.

Beispiel 1: *Umsortierung der Koordinaten des \mathbb{R}^n*

Sei ρ eine Permutation der Zahlen $\{1, \dots, n\}$. Die Abbildung $f_\rho : \mathbb{R}^n \rightarrow \mathbb{R}^n$, die die Reihenfolge der Koordinaten des \mathbb{R}^n entsprechend der Permutation ρ vertauscht, d.h.

$$f_\rho(x_1, x_2, \dots, x_n) := (x_{\rho(1)}, x_{\rho(2)}, \dots, x_{\rho(n)}),$$

ist eine Koordinatentransformation.

Beispiel 2: *Affine Koordinatentransformation*

Sei $A \in L(\mathbb{R}^n, \mathbb{R}^n)$ ein (linearer) Isomorphismus und $\mathbf{b} \in \mathbb{R}^n$ ein Vektor. Dann heißt die Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ mit

$$f(x) := A(x) + \mathbf{b}$$

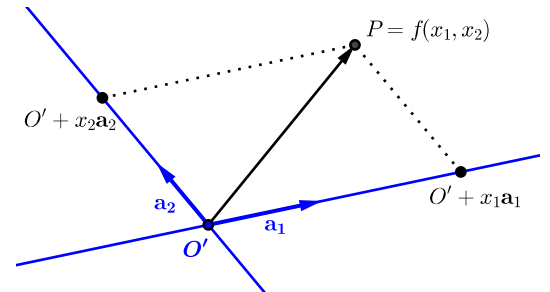
affine Koordinatentransformation. f ist bijektiv, die Umkehrabbildung ist gegeben durch $f^{-1}(y) = A^{-1}(y) - A^{-1}(\mathbf{b})$. f ist stetig differenzierbar, da

$$\frac{\partial f}{\partial x_i}(x) = df_x(e_i) = dA_x(e_i) = A(e_i) = \textit{konstant}.$$

Analog zeigt man, dass f^{-1} stetig differenzierbar ist.

Die Koordinaten (x_1, \dots, x_n) des Punktes $P = f(x_1, \dots, x_n)$ sind die Koordinaten von P im Koordinatensystem $(O'; \mathbf{a}_1, \dots, \mathbf{a}_n)$ mit dem Ursprung $O' := O + \mathbf{b}$ und den Basisvektoren $\mathbf{a}_i := A e_i$, denn

$$P = f(x_1, \dots, x_n) = \mathbf{b} + A \left(\sum_{i=1}^n x_i e_i \right) = O' + \sum_{i=1}^n x_i \mathbf{a}_i.$$



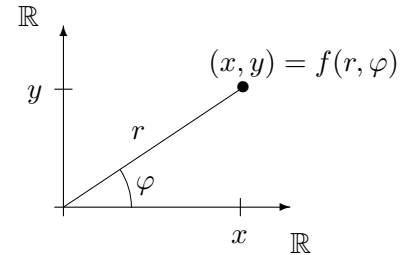
Beispiel 3: Polarkoordinaten in der Ebene

Wir betrachten die Abbildung

$$f : (0, \infty) \times (0, 2\pi) \longrightarrow \mathbb{R}^2 \setminus ([0, \infty) \times \{0\})$$

$$f(r, \varphi) := (r \cos(\varphi), r \sin(\varphi)).$$

(r, φ) heißen *Polarkoordinaten des Punktes* $(x, y) = f(r, \varphi)$.



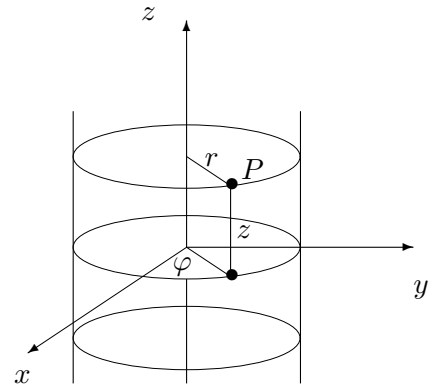
Beispiel 4: Zylinderkoordinaten im \mathbb{R}^3

Wir betrachten die Abbildung

$$f : \underbrace{(0, \infty) \times (0, 2\pi)}_{\text{Polarkoordinaten}} \times \mathbb{R} \longrightarrow (\mathbb{R}^2 \setminus ([0, \infty) \times \{0\})) \times \mathbb{R}$$

$$f(r, \varphi, z) := (r \cos(\varphi), r \sin(\varphi), z).$$

(r, φ, z) heißen *Zylinderkoordinaten des Punktes* $P = f(r, \varphi, z)$.



Beispiel 5: Kugelkoordinaten im \mathbb{R}^3

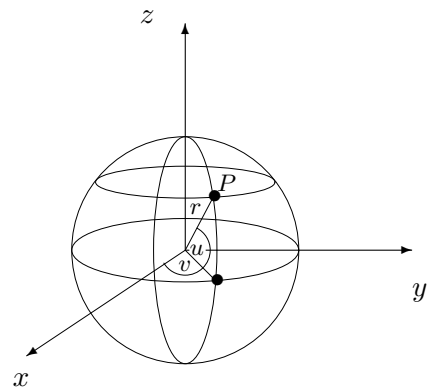
Wir betrachten die Abbildung

$$f : (0, \infty) \times (-\frac{\pi}{2}, \frac{\pi}{2}) \times (0, 2\pi) \longrightarrow \mathbb{R}^3 \setminus ([0, \infty) \times \{0\} \times \mathbb{R})$$

$$f(r, u, v) := (r \cos(u) \cos(v), r \cos(u) \sin(v), r \sin(u)).$$

(r, u, v) heißen *Kugelkoordinaten des Punktes* $P = f(r, u, v)$.

$(r \cos(u), v)$ sind dabei die Polarkoordinaten der Projektion (P_1, P_2) von P auf die xy -Ebene.



Die Bijektivität der Abbildungen in den letzten drei Beispielen folgt aus ihrer geometrischen Konstruktion. Die Abbildungen sind offensichtlich stetig differenzierbar. Um zu überprüfen, ob die inverse Abbildung ebenfalls stetig differenzierbar ist, müßte man sie ausrechnen. Wir werden im folgenden eine Möglichkeit angeben, mit denen man sich die explizite Berechnung der inversen Abbildung ersparen kann.

Zunächst hat man die folgende *notwendige* Bedingung dafür, dass $f : U \rightarrow V$ ein Diffeomorphismus ist.

Satz 8.15 *Sei $f : U \rightarrow V$ ein Diffeomorphismus zwischen offenen Mengen des \mathbb{R}^n . Dann ist das Differential $df_p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ für alle $p \in U$ ein Isomorphismus und seine inverse Abbildung ist gegeben durch*

$$(df_p)^{-1} = (df^{-1})_{f(p)}.$$

Beweis. Da f stetig differenzierbar ist, ist f auch differenzierbar.

Wir wenden die Kettenregel für Differentiale auf die Abbildungen $f \circ f^{-1} = \text{Id}_V$ und $f^{-1} \circ f = \text{Id}_U$ an. Für alle $q \in V$ und $p \in U$ gilt dann

$$\begin{aligned} d(f \circ f^{-1})_q &= df_{f^{-1}(q)} \circ (df^{-1})_q = (d\text{Id}_{\mathbb{R}^n})_q = \text{Id}_{\mathbb{R}^n} \quad \text{und} \\ d(f^{-1} \circ f)_p &= (df^{-1})_{f(p)} \circ df_p = (d\text{Id}_{\mathbb{R}^n})_p = \text{Id}_{\mathbb{R}^n}. \end{aligned}$$

Betrachten wir $q = f(p)$, so folgt $(df^{-1})_{f(p)} = (df_p)^{-1}$. □

Bemerkung 1: Die Jacobi-Matrix Jf_p ist die Darstellungsmatrix des Differentials df_p . Folglich gilt:

$$df_p : \mathbb{R}^n \rightarrow \mathbb{R}^n \text{ ist ein Isomorphismus} \iff Jf_p \text{ ist invertierbar} \iff \det Jf_p \neq 0.$$

Bemerkung 2: Im 1-dimensionalen gilt auch die Umkehrung von Satz 8.15.

Sei $U \subset \mathbb{R}$ ein offenes Intervall und $f : U \subset \mathbb{R} \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion mit $f'(x) \neq 0$ für alle $x \in U$. Dann ist f streng monoton. Insbesondere folgt dann bereits, dass $f : U \rightarrow f(U)$ bijektiv ist. Außerdem ist die Umkehrfunktion stetig differenzierbar (siehe Satz 4.4 und Satz 5.4). $f : U \rightarrow f(U)$ ist also ein Diffeomorphismus.

Im \mathbb{R}^n mit $n > 1$ gilt diese Aussage nicht mehr. Eine stetig differenzierbare Abbildung $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$, deren Jacobi-Matrix in jedem Punkt invertierbar ist, muß nicht mal injektiv sein. Als Beispiel betrachten wir die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$, definiert durch

$$f(x, y) := (e^x \cos y, e^x \sin y).$$

Diese Abbildung ist stetig differenzierbar und für die Determinante der Jacobi-Matrix von f gilt:

$$\det Jf_{(x,y)} = \det \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} = e^{2x} > 0.$$

f ist aber offensichtlich nicht injektiv. Es gibt also keine globale Umkehrabbildung von $f : U \rightarrow f(U)$.

Wir wollen nun untersuchen, unter welchen Bedingungen eine stetig differenzierbare Abbildung $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ wenigstens lokal eine (stetig differenzierbare) Umkehrfunktion besitzt.

Definition 8.15. Eine Abbildung $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt lokaler Diffeomorphismus um $p \in U$, wenn es offene Umgebungen² $\tilde{U} \subset U$ von p und $\tilde{V} \subset \mathbb{R}^n$ von $f(p)$ gibt, so dass $f|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ein Diffeomorphismus ist.

Satz 8.16 (Satz über den lokalen Diffeomorphismus)

Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar, $p \in U$ und df_p ein Isomorphismus. Dann ist f ein lokaler Diffeomorphismus um p .

Beweis. 1. Schritt: Reduktion auf eine einfache Situation:

a) Wir können annehmen, dass $df_p = \text{Id}_{\mathbb{R}^n}$ gilt.

Anderenfalls betrachten wir die Abbildung $\tilde{f} := (df_p)^{-1} \circ f$. Da $(df_p)^{-1}$ linear ist, folgt aus der Kettenregel für Differentiale $d\tilde{f}_p = (df_p)^{-1} \circ df_p = \text{Id}_{\mathbb{R}^n}$. Ist $\tilde{f}|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ein Diffeomorphismus, so ist $f|_{\tilde{U}} : \tilde{U} \rightarrow df_p(\tilde{V})$ ein Diffeomorphismus.

b) Wir können außerdem annehmen, dass $p = 0$ und $f(p) = 0$ gilt.

Anderenfalls betrachten wir die Abbildung \hat{f} , definiert durch $\hat{f}(x) := f(x + p) - f(p)$. Dann gilt $\hat{f}(0) = 0$ und $d\hat{f}_0 = df_p$. Ist $\hat{f}|_{U_0} : U_0 \rightarrow V_0$ ein Diffeomorphismus um 0 und $\tilde{U} := U_0 + p$, $\tilde{V} := V_0 + f(p)$, dann ist $f|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ein Diffeomorphismus um p .

Es genügt also zu zeigen: Ist $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar, $0 \in U$, $f(0) = 0$ und $df_0 = \text{Id}_{\mathbb{R}^n}$, so ist f ein lokaler Diffeomorphismus um 0.

2. Schritt: Wir konstruieren offene Mengen U_0 und V_0 um 0, so dass $f|_{U_0} : U_0 \rightarrow V_0$ bijektiv ist: Da f stetig differenzierbar ist, ist die Abbildung

$$x \in U \mapsto Jf_x \in M(n, n, \mathbb{R})$$

stetig (bzgl. der Euklidischen Norm auf den Matrizen). Da die gewählte Norm für Konvergenzfragen keine Rolle spielt (siehe Satz 7.26) ist auch die Abbildung

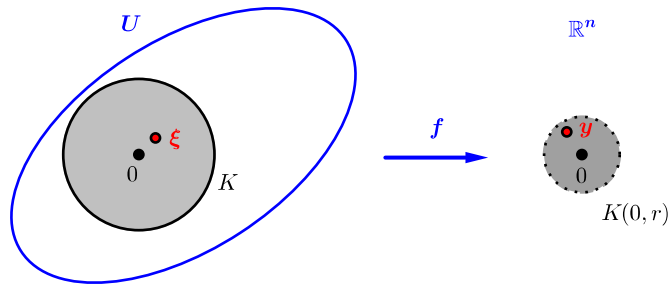
$$x \in U \mapsto df_x \in L(\mathbb{R}^n, \mathbb{R}^n)$$

stetig bzgl. der Operatornorm auf $L(\mathbb{R}^n, \mathbb{R}^n)$. Folglich existiert ein $r > 0$, so dass die abgeschlossene Kugel $K := \{x \in \mathbb{R}^n \mid \|x\| \leq 2r\}$ in U liegt und für alle $x \in K$ gilt

$$\|df_x - df_0\| = \|df_x - \text{Id}_{\mathbb{R}^n}\| < \frac{1}{2}. \quad (8.12)$$

Sei $y \in K(0, r)$ fixiert. Wir zeigen nun mit Hilfe des Banachschen Fixpunktsatzes, dass es genau ein $\xi \in K(0, 2r) = \text{Int}(K)$ gibt mit $f(\xi) = y$.

² Mit *Umgebung eines Punktes* p meinen wir eine Menge, die den Punkt p enthält.



Dazu betrachten wir die Abbildung $\phi : U \rightarrow \mathbb{R}^n$ mit

$$\phi(x) := y + x - f(x).$$

Dann gilt $\phi(\xi) = \xi$ genau dann, wenn $f(\xi) = y$. Die Urbilder von y bei f sind also die Fixpunkte von ϕ . Außerdem gilt $\phi(0) = y$ und $d\phi_x = \text{Id}_{\mathbb{R}^n} - df_x$ für alle $x \in U$. Aus (8.12) folgt dann

$$\|d\phi_x\| = \|\text{Id}_{\mathbb{R}^n} - df_x\| < \frac{1}{2} \quad \forall x \in K.$$

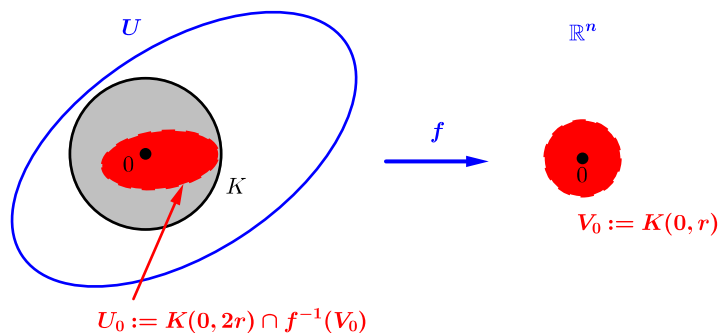
Somit ist ϕ auf K Lipschitzstetig mit der Lipschitzkonstanten $\frac{1}{2}$ (siehe Satz 8.10), d.h. es gilt

$$\|\phi(b) - \phi(a)\| \leq \frac{1}{2}\|b - a\| \quad \forall a, b \in K. \tag{8.13}$$

Insbesondere ist

$$\begin{aligned} \|\phi(x)\| &\leq \|\phi(x) - \phi(0)\| + \|\phi(0)\| \\ &\stackrel{(8.13)}{\leq} \frac{1}{2}\|x\| + \|y\| < r + r = 2r \quad \forall x \in K, \end{aligned}$$

d.h., ϕ bildet die abgeschlossene Kugel K in die offene Kugel $K(0, 2r) \subset K$ ab. Da $K \subset \mathbb{R}^n$ abgeschlossen und \mathbb{R}^n ein vollständiger metrischer Raum ist, ist K ebenfalls vollständig. $\phi : K \rightarrow K$ ist also eine kontrahierende Abbildung auf einem vollständigen metrischen Raum. Der Banachsche Fixpunktsatz (Satz 7.24) liefert dann die Existenz eines Fixpunktes ξ von ϕ auf K und seine Eindeutigkeit. Da ϕ in $K(0, 2r)$ abbildet, gilt $\xi \in K(0, 2r)$. Es gibt also genau ein $\xi \in K(0, 2r)$ mit $f(\xi) = y$. Die Kugel $V_0 := K(0, r)$ ist eine offene Umgebung von $f(0) = 0$, die Menge $U_0 := K(0, 2r) \cap f^{-1}(V_0)$ ist eine offene Umgebung von 0 (topologisches Stetigkeitskriterium) und nach Konstruktion ist die Abbildung $f|_{U_0} : U_0 \rightarrow V_0$ bijektiv.



3. *Schritt:* Wir zeigen nun, dass die Umkehrabbildung $g := (f|_{U_0})^{-1} : V_0 \rightarrow U_0$ stetig ist. Dazu betrachten wir die Abbildung $\varphi : U \rightarrow \mathbb{R}^n$ mit $\varphi(x) = x - f(x)$ für alle $x \in U$. φ stimmt mit der Abbildung ϕ (für $y = 0$) aus dem 2. Schritt überein. Für $a, b \in U_0$ folgt damit

$$b - a = \varphi(b) + f(b) - (\varphi(a) + f(a))$$

und somit

$$\|b - a\| \leq \|\varphi(b) - \varphi(a)\| + \|f(b) - f(a)\| \stackrel{(8.13)}{\leq} \frac{1}{2}\|b - a\| + \|f(b) - f(a)\|.$$

Also gilt

$$\|b - a\| \leq 2 \cdot \|f(b) - f(a)\| \quad \forall b, a \in U_0. \tag{8.14}$$

Für $u, w \in V_0$ setzen wir $a := g(u)$, $b := g(w)$. Dann folgt aus (8.14)

$$\|g(w) - g(u)\| \leq 2 \cdot \|w - u\| \quad \forall w, u \in V_0.$$

g ist also Lipschitzstetig und somit stetig.

4. *Schritt:* Das Differential $df_p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist für jedes $p \in U_0$ ein Isomorphismus.

Angenommen, es gäbe ein $v \in \mathbb{R}^n$ mit $v \neq 0$ und $df_p(v) = 0$. Dann ergibt

$$\|v\| = \|v - df_p(v)\| = \|(\text{Id}_{\mathbb{R}^n} - df_p)(v)\| \leq \|\text{Id}_{\mathbb{R}^n} - df_p\| \cdot \|v\| \stackrel{(8.12)}{\leq} \frac{1}{2}\|v\|$$

einen Widerspruch.

5. *Schritt:* Die Umkehrabbildung $g := (f|_{U_0})^{-1} : V_0 \rightarrow U_0$ ist differenzierbar.

Sei $q \in V_0$ ein beliebiger Punkt und $p = g(q) \in U_0$. Wir zeigen, dass g in q differenzierbar ist. Da f in p differenzierbar ist, gilt für $p + h \in U_0$

$$f(p + h) - f(p) = df_p(h) + r_p(h) \quad \text{mit} \quad \lim_{h \rightarrow 0} \frac{r_p(h)}{\|h\|} = 0. \tag{8.15}$$

Sei $y := f(p + h)$. Dann gilt $g(y) = p + h = g(q) + h$. Aus der Gleichung (8.15) erhalten wir h durch Anwendung von $(df_p)^{-1}$:

$$h = g(y) - g(q) = \underbrace{(df_p)^{-1}(y - q)}_{\text{linear}} - \underbrace{(df_p)^{-1}(r_p(h))}_{\text{Restglied}}.$$

Wir müssen nun für das Restglied zeigen, dass

$$\lim_{y \rightarrow q} \frac{\|(df_p)^{-1}(r_p(h))\|}{\|y - q\|} = 0. \tag{8.16}$$

Es gilt

$$\frac{\|(df_p)^{-1}(r_p(h))\|}{\|y - q\|} \leq \|(df_p)^{-1}\| \cdot \frac{\|r_p(h)\|}{\|y - q\|} = \|(df_p)^{-1}\| \cdot \frac{\|r_p(h)\|}{\|h\|} \cdot \frac{\|h\|}{\|f(p + h) - f(p)\|}.$$

Aus (8.14) folgt

$$\|h\| = \|(p+h) - p\| \leq 2 \cdot \|f(p+h) - f(p)\|.$$

Wir erhalten also die Abschätzung

$$\frac{\|(df_p)^{-1}(r_p(h))\|}{\|y - q\|} \leq 2 \cdot \|(df_p)^{-1}\| \cdot \frac{\|r_p(h)\|}{\|h\|}.$$

Da g stetig ist, folgt aus $y \rightarrow q$, dass $h = g(y) - g(q) \rightarrow 0$. Dies liefert mit (8.15) die Restgliedeigenschaft (8.16). Somit ist g in $q \in V_0$ differenzierbar und es gilt $dg_q = (df_p)^{-1}$.

6. *Schritt:* Als letztes zeigen wir, dass $g := (f|_{U_0})^{-1} : V_0 \rightarrow U_0$ stetig differenzierbar ist. Die Jacobi-Matrix von g in $q \in V_0$ ist die inverse Matrix der Jacobi-Matrix von f in $p = g(q)$. Die Vorschrift zur Berechnung inverser Matrizen zeigt, dass die partiellen Ableitungen von g in q rationale Funktionen der partiellen Ableitungen von f in p sind, d.h. es existieren rationale Funktionen R_{kl} in n^2 Variablen, so dass

$$\frac{\partial g_k}{\partial y_l}(q) = R_{kl} \left(\frac{\partial f_i}{\partial x_j}(g(q)), \quad i, j = 1, \dots, n \right).$$

Da f stetig differenzierbar ist, sind die partiellen Ableitungen von f stetig. g ist stetig (Schritt 3). Rationale Funktionen sind auch stetig. Somit sind die partiellen Ableitungen von g als Verknüpfung stetiger Funktionen ebenfalls stetig. \square

Abschließend beweisen wir ein Kriterium für das Vorliegen eines globalen Diffeomorphismus (bzw. einer Koordinatentransformation).

Satz 8.17 (Globaler Umkehrsatz)

Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar und sei das Differential $df_p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ für alle $p \in U$ ein Isomorphismus. Dann gilt:

1. $V := f(U) \subset \mathbb{R}^n$ ist offen.
2. Ist f zusätzlich injektiv, so ist $f : U \rightarrow V$ ein Diffeomorphismus.

Beweis. (1) Nach Satz 8.16 existiert für alle $p \in U$ eine offene Umgebung $U_p \subset U$ von p und V_p von $f(p)$, so dass $f|_{U_p} : U_p \rightarrow V_p$ ein Diffeomorphismus ist. Dann gilt $U = \bigcup_{p \in U} U_p$ und

$$f(U) = \bigcup_{p \in U} f|_{U_p}(U_p) = \bigcup_{p \in U} V_p =: V.$$

Die Menge V ist insbesondere offen (Siehe Übungsaufgabe 20).

(2) Ist f injektiv, so ist $f : U \rightarrow V = f(U)$ bijektiv. Da $f|_{U_p} : U_p \rightarrow V_p$ für alle $p \in U$ ein Diffeomorphismus ist und Differenzierbarkeit eine lokale Eigenschaft ist, ist $f : U \rightarrow V = f(U)$ ein (globaler) Diffeomorphismus. \square

Mit diesem Satz kann man nun leicht zeigen, dass Abbildungen aus den Beispielen 3, 4 und 5, die die Polarkoordinaten, die Zylinderkoordinaten und die Kugelkoordinaten definieren, tatsächlich Koordinatentransformationen sind. Man berechnet dazu die Jacobi-Matrizen und zeigt, dass ihre Determinate in jedem Punkt ungleich Null ist. (Übungsaufgabe).

8.7 Auflösen von Gleichungen. Der Satz über implizite Funktionen

Eine Gerade in der Euklidischen Ebene \mathbb{R}^2 kann man auf verschiedene Weise beschreiben:

a) *Implizite Beschreibung* als Lösungsmenge der Gleichung

$$f(x, y) := \alpha x + \beta y + \gamma = 0, \quad (*)$$

wobei α, β, γ reelle Konstanten mit $(\alpha, \beta) \neq (0, 0)$ sind.

b) *Explizite Beschreibung* als Graph einer Funktion $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ über der x -Achse oder einer Funktion $\psi : \mathbb{R} \rightarrow \mathbb{R}$ über der y -Achse mit

$$\varphi(x) := mx + d \quad \text{bzw.} \quad \psi(y) = \tilde{m}y + \tilde{d}.$$

Ist $\beta \neq 0$, so erhält man φ durch Auflösen der Gleichung $(*)$ nach der y -Variablen:

$$y = \varphi(x) := -\frac{\alpha}{\beta}x - \frac{\gamma}{\beta}.$$

Ist $\alpha \neq 0$ so kann man $(*)$ nach x auflösen und erhält

$$x = \psi(y) := -\frac{\beta}{\alpha}y - \frac{\gamma}{\alpha}.$$

In der expliziten Form erkennt man die Gestalt der Geraden sofort.

In der *Lineare Algebra* wird dieses Verfahren auf Systeme von m *linearen* Gleichungen mit $N \geq m$ Unbestimmten verallgemeinert. Auch hier löst man die Gleichungen nach einzelnen Unbestimmten auf (Gauß-Algorithmus).

Wir wollen jetzt die analoge Frage für *nicht-lineare* Gleichungen stellen. Wir betrachten m Funktionen f_1, \dots, f_m , die von $N \geq m$ Variablen (x_1, \dots, x_N) abhängen und das System der Gleichungen

$$\begin{aligned} f_1(x_1, \dots, x_N) &= 0, \\ f_2(x_1, \dots, x_N) &= 0, \\ &\vdots \\ f_m(x_1, \dots, x_N) &= 0, \end{aligned}$$

oder in Kurzfassung: $f(x) = 0$ für $f = (f_1, \dots, f_m)$ und $x = (x_1, \dots, x_N)$. Dann stellen sich folgende Fragen:

- Hat dieses Gleichungssystem eine Lösung?
- Von wievielen Parametern hängt die Lösungsmenge ab?
(D.h. wie "groß" ist die Lösungsmenge?)
- Wann kann man die Lösungsmenge *explizit* als Graph einer Funktion φ beschreiben?

Um diese Fragen zu beantworten, versucht man wieder, die Gleichungen nach einer oder mehreren Variablen aufzulösen.

Wir sehen uns zunächst zwei Beispiele in der Ebene an, die die typische Effekte zeigen, die auftreten, wenn man eine *nicht-lineare* Gleichung nach einer Variablen aufzulösen versucht.

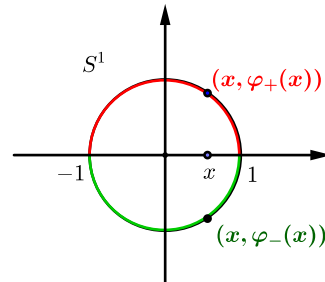
Beispiel 1: *Der Kreis S^1*

Wir betrachten in der Ebene den Kreis S^1 um $(0,0)$ vom Radius 1. S^1 ist die Lösungsmenge der Gleichung

$$f(x, y) := x^2 + y^2 - 1 = 0.$$

Man kann die Gleichung $x^2 + y^2 - 1 = 0$ nach y auflösen und erhält

$$y = \pm\sqrt{1 - x^2} =: \varphi_{\pm}(x).$$



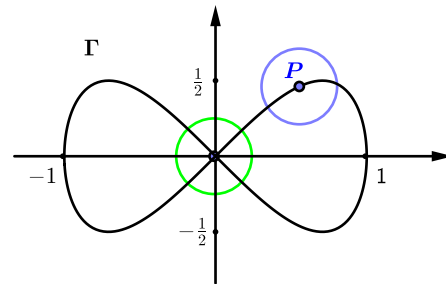
Es gibt hier also zwei Auflösungen φ_{\pm} der Gleichung. Beide sind nur auf dem Intervall $[-1, 1]$ definiert und beschreiben jeweils nur einen Teil der Lösungsmenge S^1 . Eine *globale* Beschreibung von S^1 als Graph einer Funktion gibt es nicht.

Beispiel 2: *Die Lemniskate von Geron*

Sei $\Gamma \subset \mathbb{R}^2$ die Lösungsmenge der Gleichung

$$f(x, y) := x^2(x^2 - 1) + y^2 = 0.$$

Betrachtet man einen beliebigen Punkt $P \in \Gamma$ mit $P \neq (0,0)$, so gibt es eine kleine Umgebung U von P , für die man $\Gamma \cap U$ als Graph einer Funktion über der x -Achse oder der y -Achse schreiben kann. Im Punkt $(0,0)$ geht das allerdings nicht, da Γ mit zwei Zweigen durch $(0,0)$ läuft. In keiner noch so kleinen Umgebung von $(0,0)$ kann man Γ nach x oder y auflösen.



Lemniskate von Geron

Im nächsten Satz geben wir Bedingungen dafür an, wann man die Lösungsmenge einer Gleichung $f(x) = 0$ wenigstens *lokal* nach einer bestimmten Zahl von Variablen auflösen kann, d.h. wann man sie *lokal* als Graph einer Funktion φ beschreiben kann. Dabei kann man in der Regel zwar φ nicht konkret angeben, man erhält aber nützliche konkrete Formeln für ihre partiellen Ableitungen.

Satz 8.18 (Satz über implizite Funktionen)

Sei $U \subset \mathbb{R}^n \times \mathbb{R}^m$ offen und $f : U \rightarrow \mathbb{R}^m$ eine stetig differenzierbare Abbildung.

Wir betrachten einen Punkt $(a, b) \in U$ mit

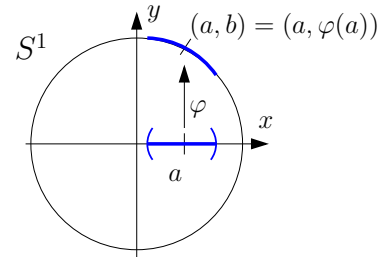
1. $f(a, b) = 0$, d.h. (a, b) löst die Gleichung $f(x, y) = 0$.
2. Die letzten m Spalten der Jacobi-Matrix $Jf_{(a,b)}$ sind linear unabhängig, d.h.,

$$\det \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(a, b) & \frac{\partial f_1}{\partial y_2}(a, b) & \dots & \frac{\partial f_1}{\partial y_m}(a, b) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m}{\partial y_1}(a, b) & \frac{\partial f_m}{\partial y_2}(a, b) & \dots & \frac{\partial f_m}{\partial y_m}(a, b) \end{pmatrix} \neq 0.$$

Dann existieren offene Umgebungen $U_0 \subset U$ von (a, b) und $A(a) \subset \mathbb{R}^n$ von a , sowie eine eindeutig bestimmte Funktion $\varphi : A(a) \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ mit

1. $\varphi(a) = b$,
2. $f(x, \varphi(x)) = 0$ für alle $x \in A(a)$.
3. $f^{-1}(0) \cap U_0 = \text{graph}(\varphi)$,

d.h., in einer Umgebung des Punktes (a, b) ist die Lösungsmenge der Gleichung $f(x, y) = 0$ durch den Graphen der Funktion φ beschrieben.



4. $\varphi : A(a) \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist stetig differenzierbar und die partiellen Ableitungen sind gegeben durch

$$\underbrace{\left(\frac{\partial \varphi_l}{\partial x_i}(x) \right)}_{\in M(m, n, \mathbb{R})} = - \underbrace{\left(\frac{\partial f_k}{\partial y_l}(x, \varphi(x)) \right)^{-1}}_{\in M(m, m, \mathbb{R})} \cdot \underbrace{\left(\frac{\partial f_k}{\partial x_i}(x, \varphi(x)) \right)}_{\in M(m, n, \mathbb{R})} \quad \forall x \in A(a).$$

Beweis. 1) Konstruktion von φ (d.h. Auflösung von $f(x, y) = 0$ nach $y = (y_1, \dots, y_m)$):
Wir wenden dazu den Satz über den lokalen Diffeomorphismus an. Wir betrachten die Abbildung $F : U \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n \times \mathbb{R}^m$, definiert durch

$$F(x, y) := (x, f(x, y)).$$

F ist stetig differenzierbar und nach Voraussetzung gilt:

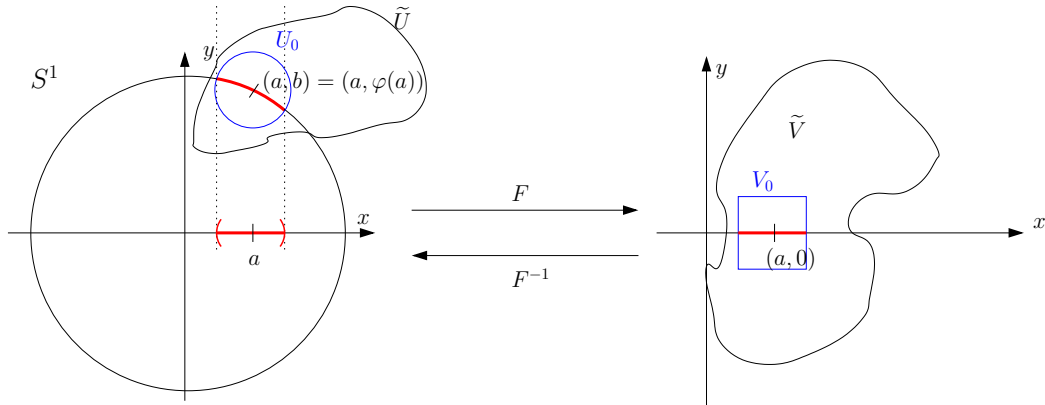
$$\det JF_{(a,b)} = \det \begin{pmatrix} E_n & & 0 \\ & \frac{\partial f_1}{\partial y_1}(a, b) & \dots & \frac{\partial f_1}{\partial y_m}(a, b) \\ * & \vdots & & \vdots \\ & \frac{\partial f_m}{\partial y_1}(a, b) & \dots & \frac{\partial f_m}{\partial y_m}(a, b) \end{pmatrix} \neq 0.$$

Nach dem Satz über den lokalen Diffeomorphismus (Satz 8.16) existieren offene Umgebungen $\tilde{U} \subset U$ um (a, b) und $\tilde{V} \subset \mathbb{R}^n \times \mathbb{R}^m$ um $F(a, b) = (a, 0)$, so dass $F|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ein Diffeomorphismus ist. Da \tilde{V} offen ist, können wir eine offene Umgebung $A(a) \subset \mathbb{R}^n$ von a und einen offenen Würfel $(-\varepsilon, \varepsilon)^m \subset \mathbb{R}^m$ wählen, so dass $V_0 := A(a) \times (-\varepsilon, \varepsilon)^m \subset \tilde{V}$. Sei nun $U_0 := F^{-1}(V_0) \subset \tilde{U}$. Dann ist $F|_{U_0} : U_0 \rightarrow V_0$ ein Diffeomorphismus. Wir betrachten die inverse Abbildung $(F|_{U_0})^{-1} : V_0 \rightarrow U_0$ mit den Komponenten

$$(F|_{U_0})^{-1} =: (F_1^-, F_2^-).$$

\uparrow
 \mathbb{R}^n

\uparrow
 \mathbb{R}^m



Dann erhalten wir

$$(x, y) = F(F|_{U_0})^{-1}(x, y) = F(F_1^-(x, y), F_2^-(x, y)) \stackrel{\text{Def.}}{=} (F_1^-(x, y), f(F_1^-(x, y), F_2^-(x, y))).$$

Folglich gilt

$$\begin{aligned} F_1^-(x, y) &= x, \\ f(x, F_2^-(x, y)) &= y \quad \forall (x, y) \in V_0. \end{aligned}$$

Dies zeigt insbesondere, dass

$$f^{-1}(0) \cap U_0 = \{(x, F_2^-(x, 0)) \mid x \in A(a)\}.$$

Wir definieren nun $\varphi : A(a) \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ durch

$$\varphi(x) := F_2^-(x, 0).$$

Dann gilt:

1. $\varphi(a) = F_2^-(a, 0) = b$.
2. $f(x, \varphi(x)) = f(x, F_2^-(x, 0)) = 0$ für alle $x \in A(a)$.
3. $f^{-1}(0) \cap U_0 = \{(x, \varphi(x)) \mid x \in A(a)\} = \text{graph}(\varphi)$.

Die Eindeutigkeit der Abbildung φ mit den Eigenschaften 1) - 3) folgt aus der Konstruktion.

2) *Differenzierbarkeit von φ :*

$F|_{U_0} : U_0 \rightarrow V_0$ ist ein Diffeomorphismus, folglich ist die inverse Abbildung $(F|_{U_0})^{-1}$ stetig differenzierbar. Das gleiche gilt dann auch für ihre Komponente F_2^- . Da $\varphi(x) = F_2^-(x, 0)$, ist auch φ stetig differenzierbar.

3) *Die Formel für die partiellen Ableitungen von φ :*

Wir wenden die Kettenregel für partielle Ableitungen auf die Funktion

$$g : A(a) \subset \mathbb{R}^n \rightarrow \mathbb{R}^m,$$

$$g(x) := f(x, \varphi(x)) = f(x_1, \dots, x_n, \varphi_1(x_1, \dots, x_n), \dots, \varphi_m(x_1, \dots, x_n)) = 0,$$

an und erhalten für $x \in A(a)$:

$$\begin{aligned}
0 &= \frac{\partial g}{\partial x_i}(x) \\
&= \sum_{j=1}^n \frac{\partial f}{\partial x_j}(x, \varphi(x)) \cdot \underbrace{\frac{\partial x_j}{\partial x_i}}_{=\delta_{ij}} + \sum_{l=1}^m \frac{\partial f}{\partial y_l}(x, \varphi(x)) \cdot \frac{\partial \varphi_l}{\partial x_i}(x) \\
&= \frac{\partial f}{\partial x_i}(x, \varphi(x)) + \sum_{l=1}^m \frac{\partial f}{\partial y_l}(x, \varphi(x)) \cdot \frac{\partial \varphi_l}{\partial x_i}(x).
\end{aligned}$$

Für die Komponenten (f_1, \dots, f_m) der Funktion f bedeutet dies in Matrixschreibweise

$$-\left(\frac{\partial f_k}{\partial x_i}(x, \varphi(x))\right) = \left(\frac{\partial f_k}{\partial y_l}(x, \varphi(x))\right) \cdot \left(\frac{\partial \varphi_l}{\partial x_i}(x)\right). \quad (*)$$

Da die Abbildung $F|_{U_0} : U_0 \rightarrow V_0$ ein Diffeomorphismus ist, ist die Matrix $\left(\frac{\partial f_k}{\partial y_l}(x, \varphi(x))\right)$ für jedes $x \in A(a)$ invertierbar. Nach Multiplikation der Gleichung $(*)$ mit der Inversen der Matrix $\left(\frac{\partial f_k}{\partial y_l}(x, \varphi(x))\right)$ erhalten wir für die Ableitungen von φ

$$\left(\frac{\partial \varphi_l}{\partial x_i}(x)\right) = -\left(\frac{\partial f_k}{\partial y_l}(x, \varphi(x))\right)^{-1} \cdot \left(\frac{\partial f_k}{\partial x_i}(x, \varphi(x))\right).$$

□

8.8 Gleichungsdefinierte Flächen, ihre Tangentialebenen und Normalen

Am Ende von Abschnitt 8.1 hatten wir eine Formel für die Tangentialebenen an Flächen angegeben, die als Graph einer reellen Funktion in 2 Variablen gegeben sind:

Sei $\varphi : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion und F der Graph von φ , d.h.

$$F := \{(x, \varphi(x)) \mid x \in U\}.$$

Dann gilt für die Tangentialebene an F im Punkt $p = (x, y, \varphi(x, y)) \in F$

$$\text{Tan}_p F = p + \mathbb{R} \cdot (1, 0, \frac{\partial \varphi}{\partial x}(x, y)) + \mathbb{R} \cdot (0, 1, \frac{\partial \varphi}{\partial y}(x, y)).$$

Wir wollen nun mit Hilfe des Satzes über implizite Funktionen eine Formel für die Tangentialebenen an gleichungsdefinierte Flächen in Termen der definierenden Gleichung herleiten.

Definition 8.16. Sei $f : U \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion und

$$M := \{(x, y, z) \in U \mid f(x, y, z) = 0\} = f^{-1}(0) \subset \mathbb{R}^3$$

die Lösungsmenge der Gleichung $f(x, y, z) = 0$.

M heißt reguläre Fläche, wenn jeder Punkt $p \in M$ ein regulärer Punkt von f ist, d.h. wenn $\text{grad} f(p) \neq 0$ für alle $p \in M$.

Der Tangentialraum an M im Punkt $p \in M$ ist die Menge der Vektoren

$$T_p M := \{v \in \mathbb{R}^3 \mid \exists \text{ diffb. Kurve } \gamma : (-\varepsilon, \varepsilon) \rightarrow M \text{ mit } \gamma(0) = p \text{ und } \gamma'(0) = v\}.$$

Beispiel: Der Graph $F = \text{graph}(\varphi)$ ist eine reguläre Fläche.

Wir betrachten die Funktion $f : U \times \mathbb{R} \rightarrow \mathbb{R}$ mit

$$f(x, y, z) := z - \varphi(x, y).$$

Dann gilt $F = f^{-1}(0) = \{(x, y, z) \in U \times \mathbb{R} \mid z = \varphi(x, y)\}$. Für den Gradienten von f erhalten wir

$$\text{grad}f(x, y, z) = \left(-\frac{\partial\varphi}{\partial x}(x, y), -\frac{\partial\varphi}{\partial y}(x, y), 1 \right) \neq 0.$$

Folglich ist F eine reguläre Fläche.

Satz 8.19 Sei $f : U \subset \mathbb{R}^3$ stetig differenzierbar, $M = f^{-1}(0) \subset \mathbb{R}^3$ eine reguläre Fläche und $p \in M$. Dann ist der Tangentialraum T_pM ein 2-dimensionaler Vektorraum und es gilt

$$T_pM = \{v \in \mathbb{R}^3 \mid v \perp \text{grad}f(p)\} = \{v \in \mathbb{R}^3 \mid \langle v, \text{grad}f(p) \rangle = 0\}.$$

Beweis. Sei $p = (p_1, p_2, p_3) \in M$. Da p ein regulärer Punkt von f ist, gilt $\text{grad}f(p) \neq 0$, d.h. mindestens eine der partiellen Ableitungen $\frac{\partial f}{\partial x}(p)$, $\frac{\partial f}{\partial y}(p)$, $\frac{\partial f}{\partial z}(p)$ ist von Null verschieden. Wir nehmen oBdA. den 3. Fall an (sonst sortieren wir die Koordinaten um). Nach dem Satz über implizite Funktionen kann man die Gleichung $f(x, y, z) = 0$ in der Umgebung von p nach z auflösen, d.h. existiert eine offene Umgebung $A \subset \mathbb{R}^2$ von (p_1, p_2) und eine stetig differenzierbare Funktion $\varphi : A \rightarrow \mathbb{R}$, so dass $\varphi(p_1, p_2) = p_3$ und $f(x, y, \varphi(x, y)) = 0$ für alle $(x, y) \in A$.³ Es folgt mit der Kettenregel durch partielles Ableiten nach x und y :

$$0 = \frac{\partial f}{\partial x}(p) + \frac{\partial f}{\partial z}(p) \cdot \frac{\partial \varphi}{\partial x}(p_1, p_2), \quad (8.17)$$

$$0 = \frac{\partial f}{\partial y}(p) + \frac{\partial f}{\partial z}(p) \cdot \frac{\partial \varphi}{\partial y}(p_1, p_2). \quad (8.18)$$

Wir zeigen nun die Formel für den Tangentialraum.

1) $T_pM \subset \{v \in \mathbb{R}^3 \mid v \perp \text{grad}f(p)\}$:

Sei $v \in T_pM$. Dann existiert eine differenzierbare Kurve $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ mit $\gamma(0) = p$ und $\gamma'(0) = v$. Da $f(\gamma(t)) = 0$ für alle $t \in (-\varepsilon, \varepsilon)$, gilt nach Kettenregel für Differentiale

$$0 = df_{\gamma(t)}(\gamma'(t)) = \langle \text{grad}f(\gamma(t)), \gamma'(t) \rangle = 0,$$

und somit für $t = 0$:

$$0 = \langle \text{grad}f(p), v \rangle = 0.$$

Also ist v senkrecht zum Vektor $\text{grad}f(p)$.

2) $\{v \in \mathbb{R}^3 \mid v \perp \text{grad}f(p)\} \subset T_pM$:

Sei $v = (v_1, v_2, v_3) \in \mathbb{R}^3$ ein Vektor, der senkrecht zu $\text{grad}f(p)$ steht. Dann gilt

$$0 = \langle \text{grad}f(p), v \rangle = \frac{\partial f}{\partial x}(p) \cdot v_1 + \frac{\partial f}{\partial y}(p) \cdot v_2 + \frac{\partial f}{\partial z}(p) \cdot v_3.$$

³ Man könnte jetzt die Formeln für die Tangentialebenen eines Graphen aus Abschnitt 8.1 benutzen. Wir werden die Behauptung aber nochmal direkt beweisen.

Setzen wir in diese Gleichung (8.17) und (8.18) ein, so erhalten wir

$$0 = -\frac{\partial f}{\partial z}(p) \cdot \frac{\partial \varphi}{\partial x}(p_1, p_2) \cdot v_1 - \frac{\partial f}{\partial z}(p) \cdot \frac{\partial \varphi}{\partial y}(p_1, p_2) \cdot v_2 + \frac{\partial f}{\partial z}(p) \cdot v_3,$$

und nach Division durch $\frac{\partial f}{\partial z}(p) \neq 0$:

$$v_3 = \frac{\partial \varphi}{\partial x}(p_1, p_2) \cdot v_1 + \frac{\partial \varphi}{\partial y}(p_1, p_2) \cdot v_2. \tag{8.19}$$

Wir betrachten nun das parametrisierte Geradenstück $\sigma : (-\varepsilon, \varepsilon) \rightarrow A$ mit $\sigma(t) = (p_1, p_2) + t(v_1, v_2)$ und die differenzierbare Kurve $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ mit $\gamma(t) = (\sigma(t), \varphi(\sigma(t)))$. Dann gilt $\gamma(0) = (p_1, p_2, \varphi(p_1, p_2)) = p$ und nach Kettenregel

$$\begin{aligned} \gamma'(0) &= \left(v_1, v_2, d\varphi_{(p_1, p_2)}(v_1, v_2) \right) = \left(v_1, v_2, \frac{\partial \varphi}{\partial x}(p_1, p_2) \cdot v_1 + \frac{\partial \varphi}{\partial y}(p_1, p_2) \cdot v_2 \right) \\ &\stackrel{(8.19)}{=} (v_1, v_2, v_3) = v. \end{aligned}$$

Somit gilt $v \in T_p M$. □

Definition 8.17. Sei $f : U \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ eine stetig differenzierbare Funktion, $M = f^{-1}(0) \subset \mathbb{R}^3$ eine reguläre Fläche und $p \in M$. Dann heißt die Ebene

$$Tan_p M := p + T_p M \subset \mathbb{R}^3$$

Tangentialebene an M im Punkt p und die Gerade

$$Nor_p M := p + \mathbb{R} \cdot \text{grad} f(p) \subset \mathbb{R}^3$$

die Normale an M im Punkt p .

Beispiel 1: Das einschalige Hyperboloid

Wir betrachten das einschalige Hyperboloid⁴

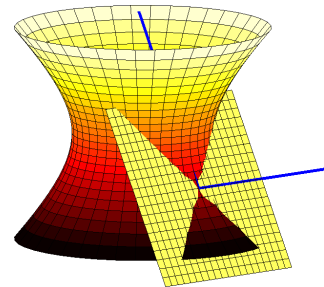
$$M := \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = z^2 + 1\}.$$

M ist die Nullstellenmenge der Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ mit

$$f(x, y, z) = x^2 + y^2 - z^2 - 1.$$

Da $\text{grad} f(x, y, z) = (2x, 2y, -2z) \neq 0$ für alle Punkte $p = (x, y, z) \in M$, ist M eine reguläre Fläche im \mathbb{R}^3 . Für den Tangentialraum in einem Punkt $p = (p_1, p_2, p_3) \in M$ erhalten wir:

$$T_p M = \{v \in \mathbb{R}^3 \mid \langle \text{grad} f(p), v \rangle = 0\} = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid p_1 v_1 + p_2 v_2 = p_3 v_3\}.$$



⁴ Vielen Dank an Christoph Stadtmüller und Thomas Neukirchner für die Bilder

Beispiel 2: *Das Paraboloid*

Wir betrachten das Paraboloid

$$M := \{(x, y, z) \in \mathbb{R}^3 \mid z = x^2 + y^2 + 1\}.$$

M ist der Graph der Funktion $\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}$, gegeben durch

$$\varphi(x, y) := x^2 + y^2 + 1.$$

Wir können M als Nullstellenmenge der Funktion $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ mit

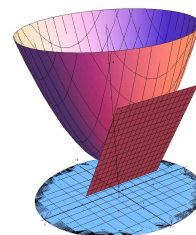
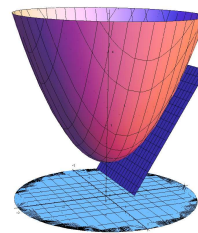
$$f(x, y, z) = x^2 + y^2 - z + 1$$

beschreiben. Für den Gradienten von f gilt

$$\operatorname{grad} f(x, y, z) = (2x, 2y, -1) \neq 0.$$

M ist also eine reguläre Fläche im \mathbb{R}^3 und für den Tangentialraum im Punkt $p = (x, y, x^2 + y^2 + 1) \in M$ erhalten wir:

$$\begin{aligned} T_p M &= \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid 2xv_1 + 2yv_2 - v_3 = 0\} \\ &= \{(v_1, v_2, 2xv_1 + 2yv_2) \mid v_1, v_2 \in \mathbb{R}\} \\ &= \mathbb{R} \cdot (1, 0, 2x) + \mathbb{R} \cdot (0, 1, 2y). \end{aligned}$$



Integralrechnung für Funktionen mehrerer reeller Variablen (Ein Ausblick)

Wir wollen nun auch die Integralrechnung auf Funktionen mehrerer reeller Variablen verallgemeinern. Die Motivation zur Einführung des Integrals ist auch hier Frage, wie man Volumen von Teilmengen im \mathbb{R}^3 (oder im \mathbb{R}^n) vernünftig definieren kann und wie man dieses Volumen auch möglichst effektiv berechnen kann.

9.1 Das Problem der Volumendefinition für Teilmengen des \mathbb{R}^n

In Kapitel 6 hatten wir das Riemann-Integral von Funktionen einer Variablen benutzt, um den Flächeninhalt *gewisser* Teilmengen des \mathbb{R}^2 zu definieren und zu berechnen. Wir wollen uns nun zunächst mit der Frage beschäftigen, ob man für *jede* Teilmenge A des \mathbb{R}^n ein Volumen $\mu_n(A)$ definieren kann, so dass eine Reihe geometrisch vernünftiger Eigenschaften erfüllt sind? Für $n = 2$ soll μ_2 alle Flächeninhaltsformeln aus Kapitel 6 erfüllen. Insbesondere soll das Volumen $\mu_2(A)$ für eine klassische ebene Figur $A \subset \mathbb{R}^2$ (Dreieck, Trapez, Rechteck, ...) der aus der Schule bekannte elementargeometrische Flächeninhalt sein. Das Volumen $\mu_3(K)$ eines Körpers $K \subset \mathbb{R}^3$ (Kegel, Zylinder, Kugel, ...) soll mit dem aus der Schule bekannten elementargeometrischen Volumen übereinstimmen. Wir stellen an eine Volumenfunktion μ_n außerdem die folgenden vernünftigen Forderungen:

μ_n sei eine Funktion

$$\mu_n : A \subset \mathbb{R}^n \longmapsto \mu_n(A) =: \text{''Volumen von } A\text{''} \in [0, \infty) \cup \{\infty\} =: [0, \infty]$$

mit den folgenden Eigenschaften:

- 1) Ist $A \subset B$, so gilt $\mu_n(A) \leq \mu_n(B)$ (*Monotonie*).
- 2) μ_n ist *translationsinvariant*, das heißt für $p \in \mathbb{R}^n$ gilt $\mu_n(A + p) = \mu_n(A)$.
- 3) Ist $W = [a_1, b_1] \times \dots \times [a_n, b_n]$ ein Quader im \mathbb{R}^n , so gilt $\mu_n(W) = \prod_{i=1}^n (b_i - a_i)$.
- 4) Sind A_1, A_2, \dots paarweise disjunkte Teilmengen, so gilt $\mu_n\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu_n(A_n)$ (*σ -Additivität*).

Satz 9.1 *Es existiert keine Funktion $\mu_n : \mathcal{P}(\mathbb{R}^n) \rightarrow [0, \infty]$ mit den Eigenschaften 1)–4).*

Beweis. Angenommen $\mu_n : \mathcal{P}(\mathbb{R}^n) \rightarrow [0, \infty]$ ist eine Funktion mit den Eigenschaften 1) - 4). Wir konstruieren uns eine (etwas merkwürdige) Menge $B \subset \mathbb{R}^n$, die dies zum Widerspruch führt. Dazu betrachten wir auf dem Würfel $[0, 1]^n \subset \mathbb{R}^n$ folgende Äquivalenzrelation:

$$x \sim y : \iff x - y \in \mathbb{Q}^n.$$

Sei $A \subset [0, 1]^n$ eine Teilmenge des Würfels, die aus jeder dieser Äquivalenzklassen genau ein Element enthält. Wir definieren die Menge B nun durch

$$B := \bigcup_{r \in [-1, 1]^n \cap \mathbb{Q}^n} A + r.$$

Dann gilt:

a) B ist die Vereinigung abzählbar vieler Mengen, da \mathbb{Q}^n abzählbar ist.

b) B ist eine disjunkte Vereinigung:

Seien $r, s \in [-1, 1]^n \cap \mathbb{Q}^n$ mit $r \neq s$. Angenommen es gäbe ein $b \in (A + r) \cap (A + s)$. Dann ist $b = a + r = a' + s$ für $a, a' \in A$. Daraus folgt $a - a' = s - r \in \mathbb{Q}^n$, folglich sind a und a' äquivalent und somit gleich (denn A enthält aus jeder Äquivalenzklasse genau ein Element). Dann wäre aber $r = s$ (Widerspruch). Für verschiedene r, s gilt also $(A + r) \cap (A + s) = \emptyset$.

c) Nach Konstruktion gilt $[0, 1]^n \subset B \subset [-1, 2]^n$.

Wegen der Monotonie von μ_n und der Festlegung des Würfelvolumens gilt dann

$$1 = \mu_n([0, 1]^n) \leq \mu_n(B) \leq \mu_n([-1, 2]^n) = 3^n.$$

Wegen der σ -Additivität und der Translationsinvarianz erhält man

$$\mu_n(B) = \sum_{r \in \mathbb{Q}^n \cap [-1, 1]^n} \mu_n(A + r) = \sum_{r \in \mathbb{Q}^n \cap [-1, 1]^n} \mu_n(A).$$

Es folgt also

$$1 \leq \sum_{r \in \mathbb{Q}^n \cap [-1, 1]^n} \mu_n(A) \leq 3^n.$$

$\mu_n(A)$ wird dabei unendlich oft summiert. Dies geht aber nicht für eine Menge A mit $0 \leq \mu_n(A) \leq \infty$. Dies liefert den Widerspruch. \square

Ausweg: Man definiert ein *n-dimensionales Volumen* nicht für alle Teilmengen, sondern nur für eine bestimmte Auswahl von Teilmengen.

- 1) Das Jordan-Volumen: $\mu_n(A)$ wird für bestimmte *beschränkte* Mengen $A \subset \mathbb{R}^n$ definiert. Dies ist die historisch erste Variante zur Präzisierung des Volumenbegriffes und geometrisch leicht zugänglich. Alle uns in dieser Vorlesung interessierenden geometrisch 'vernünftigen' beschränkten Mengen gehören dazu.
- 2) Das Lebesgue-Maß: $\mu_n(A)$ wird für eine viel größere Klasse von Mengen definiert, z.B. auch für unbeschränkte Mengen. Das ist die modernere Variante. Die Definition ist technisch viel aufwändiger. Die zugehörige Integrationstheorie (Lebesgue-Integral) hat aber wesentlich bessere Eigenschaften als das Riemann-Integral).

9.2 Das Jordan-Volumen

Ähnlich wie beim Riemann-Integral in Kapitel 6, wollen wir die 'Größe' einer beschränkten Menge $A \subset \mathbb{R}^n$ dadurch messen, dass wir sie von Innen und von Außen durch kleine Quader approximieren und die 'Größe' von A durch das geometrische Volumen dieser approximierenden Quader angeben.

Das Volumen des Quaders $W = [a_1, b_1] \times \dots \times [a_n, b_n] \subset \mathbb{R}^n$ ist die Zahl

$$vol(W) := \prod_{i=1}^n (b_i - a_i).$$

Definition 9.1. Eine Zerlegung des Quaders W ist ein n -Tupel $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_n)$ von Zerlegungen \mathcal{P}_k der Kanten $[a_k, b_k]$. Die Zerlegung $\mathcal{P}' = (\mathcal{P}'_1, \dots, \mathcal{P}'_n)$ von W heißt Verfeinerung der Zerlegung $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_n)$, wenn \mathcal{P}'_k für jedes $k = 1, \dots, n$ eine Verfeinerung von \mathcal{P}_k ist (Bezeichnung: $\mathcal{P}' \geq \mathcal{P}$).

Sei $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_n)$ eine Zerlegung von W mit $\mathcal{P}_k = \{x_{k0}, x_{k1}, \dots, x_{kr_k}\}$ wobei

$$a_k = x_{k0} < x_{k1} < \dots < x_{kr_k} = b_k.$$

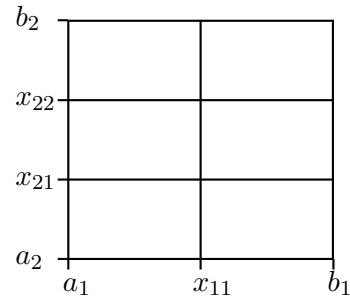
Durch

$$W_{\alpha_1 \dots \alpha_n} := [x_{1\alpha_1-1}, x_{1\alpha_1}] \times \dots \times [x_{n\alpha_n-1}, x_{n\alpha_n}]$$

für $1 \leq \alpha_i \leq r_i$ werden kleinere Quader definiert, die W zerlegen. Somit ergibt sich für den Quader W

$$W = \bigcup_{\alpha_1=1}^{r_1} \dots \bigcup_{\alpha_n=1}^{r_n} W_{\alpha_1 \dots \alpha_n},$$

$$vol(W) = \sum_{\alpha_1 \dots \alpha_n} vol(W_{\alpha_1 \dots \alpha_n}).$$



Definition 9.2. Sei $A \subset \mathbb{R}^n$ eine beschränkte Teilmenge, $Int(A)$ ihr Inneres und $cl(A)$ ihr Abschluß. Wir wählen einen Quader $W \subset \mathbb{R}^n$ mit $A \subset W$. Sei \mathcal{P} eine Zerlegung von W . Dann heißt

$$\overline{S}(A, W, \mathcal{P}) := \sum_{W_{\alpha_1 \dots \alpha_n} \cap cl(A) \neq \emptyset} vol(W_{\alpha_1 \dots \alpha_n}) \quad \text{Obersumme von } A \text{ bzgl. } \mathcal{P},$$

$$\underline{S}(A, W, \mathcal{P}) := \sum_{W_{\alpha_1 \dots \alpha_n} \subset Int(A)} vol(W_{\alpha_1 \dots \alpha_n}) \quad \text{Untersumme von } A \text{ bzgl. } \mathcal{P}.$$

Für zwei beliebige Zerlegungen \mathcal{P} und \mathcal{P}^* von W gilt

$$\underline{S}(A, W, \mathcal{P}) \leq \overline{S}(A, W, \mathcal{P}^*).$$

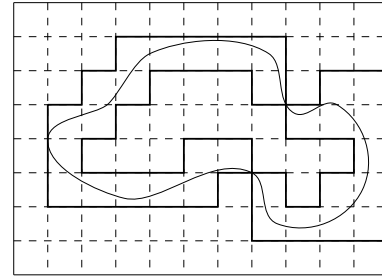
Definition 9.3. Die Zahl

$$\underline{vol}(A) := \sup \{ \underline{S}(A, W, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } W \}$$

heißt inneres Volumen von A , die Zahl

$$\overline{vol}(A) := \inf \{ \overline{S}(A, W, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } W \}$$

heißt äußeres Volumen von A .



Diese Definition ist unabhängig von der Wahl des Würfels $W \supset A$ und es gilt

$$0 \leq \underline{vol}(A) \leq \overline{vol}(A) < +\infty.$$

Definition 9.4. Eine Teilmenge $A \subset \mathbb{R}^n$ heißt Jordan-meißbar, wenn sie beschränkt ist und $\underline{vol}(A) = \overline{vol}(A)$ gilt. In diesem Fall heißt

$$vol(A) := \underline{vol}(A) = \overline{vol}(A)$$

das n -dimensionale Jordan-Volumen (oder der n -dimensionale Jordan-Inhalt) von A . Eine Teilmenge $A \subset \mathbb{R}^n$ heißt Jordansche Nullmenge, wenn sie beschränkt ist und wenn $\overline{vol}(A) = 0$ gilt.

$\mathcal{J}(\mathbb{R}^n)$ bezeichnet die Menge der Jordan-meißbaren Teilmengen des \mathbb{R}^n .

Im Fall $n = 2$ nennt man das Jordan-Volumen auch Flächeninhalt von A und bezeichnet es mit $Area(A)$. Im Fall $n = 3$ nennt man das Jordan-Volumen kurz Volumen von A .

Jordansche Nullmengen sind Jordan-meißbar und haben das Jordan-Volumen 0. Jede Teilmenge einer Jordanschen Nullmenge ist selbst eine Jordansche Nullmenge.

Der folgende Satz bietet ein gut handhabbares Kriterium, um Jordan-Meißbarkeit nachzuweisen:

Satz 9.2 (Kriterium für Jordan-Meißbarkeit)

Sei $A \subset \mathbb{R}^n$ eine beschränkte Menge und ∂A ihr Rand. Dann gilt

$$\overline{vol}(A) - \underline{vol}(A) = \overline{vol}(\partial A).$$

Insbesondere ist eine beschränkte Menge $A \subset \mathbb{R}^n$ genau dann Jordan-meißbar, wenn ihr Rand ∂A eine Jordansche Nullmenge ist.

Beispiele für Jordan-meßbare Mengen

- a) Jeder Quader $W = [a_1, b_1] \times \dots \times [a_n, b_n] \subset \mathbb{R}^n$ ist Jordan-meßbar mit den Jordan-Volumen $vol(W) = (b_1 - a_1) \cdot \dots \cdot (b_n - a_n)$.
- b) Die Menge $[0, 1]^2 \cap \mathbb{Q}^2 \subset \mathbb{R}^2$ ist nicht Jordan-meßbar.
- c) Sei $\varphi : K \subset \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ eine stetige Funktion auf einer kompakten Menge $K \subset \mathbb{R}^{n-1}$. Dann ist der Graph $\Gamma_\varphi := \{(x, \varphi(x)) \mid x \in K\} \subset \mathbb{R}^n$ eine Jordansche Nullmenge.
- d) Jedes beschränkte Gebiet $A \subset \mathbb{R}^2$, dessen Rand ∂A aus Stücken besteht, die man als Graphen stetiger Funktionen $\varphi : [a, b] \rightarrow \mathbb{R}$ darstellen kann, ist Jordan-meßbar. Z.B. ist jedes Gebiet Ω_f unter dem Graph einer stetigen nicht-negativen Funktion $f : [a, b] \rightarrow \mathbb{R}$ Jordan-meßbar (solche Gebiete haben wir in Kapitel 6.1 betrachtet).
- e) Jedes beschränkte Gebiet $\Omega \subset \mathbb{R}^2$, dessen Rand $\Gamma = \partial\Omega$ man durch eine einfache stückweise stetig differenzierbare Kurve parametrisieren kann, ist Jordan-meßbar (solche Gebiete haben wir in Kapitel 6.4.2 betrachtet).
- f) Jede beschränkte Teilmenge $B \subset \mathbb{R}^3$, deren Rand ∂B aus Graphen von stetigen Funktionen $\varphi : K \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ auf kompakten Mengen K besteht, ist Jordan-meßbar. Solche Ränder liegen z.B. oft als gleichungsdefinierte Flächen vor (siehe Kapitel 8.7.).

Mit dem Kriterium aus Satz 9.2 kann man auch die folgenden Eigenschaften Jordan-meßbarer Mengen beweisen:

Satz 9.3 1. Ist $A \subset \mathbb{R}^n$ Jordan-meßbar, so sind auch das Innere $Int(A)$, der Abschluß $cl(A)$ und der Rand ∂A Jordan-meßbar und es gilt

$$vol(Int(A)) = vol(A) = vol(cl(A)) \quad \text{und} \quad vol(\partial A) = 0.$$

2. Sind A und B Jordan-meßbar, so sind auch $A \cup B$, $A \cap B$ und $A \setminus B$ Jordan-meßbar.

3. Das Jordan-Volumen definiert eine Volumenfunktion

$$\mu_n : A \in \mathcal{J}(\mathbb{R}^n) \mapsto vol(A) \in [0, \infty)$$

mit den Eigenschaften 1)–4) aus Abschnitt 9.1.

Im nächsten Abschnitt werden wir das Riemann-Integral für Funktionen mehrerer reeller Variablen definieren. Es wird uns Methoden liefern, um das Jordan-Volumen zu berechnen.

9.3 Das Riemann-Integral für Funktionen mehrerer reeller Variablen

Die Definition des Riemann-Integrals im Mehrdimensionalen verläuft völlig analog zum 1-dimensionalen Fall (siehe Kapitel 6).

Sei $W \subset \mathbb{R}^n$ ein Quader und $f : W \rightarrow \mathbb{R}$ eine beschränkte Funktion. Wir betrachten eine Zerlegung \mathcal{P} von W mit den Teilquadern $W_{\alpha_1 \dots \alpha_n}$.

Als *Obersumme* von f bezüglich \mathcal{P} bezeichnet man die Zahl

$$\bar{S}(f, \mathcal{P}) := \sum_{W_{\alpha_1 \dots \alpha_n}} \sup(f|_{W_{\alpha_1 \dots \alpha_n}}) \cdot \text{vol}(W_{\alpha_1 \dots \alpha_n}).$$

Als *Untersumme* von f bezüglich \mathcal{P} bezeichnet man

$$\underline{S}(f, \mathcal{P}) := \sum_{W_{\alpha_1 \dots \alpha_n}} \inf(f|_{W_{\alpha_1 \dots \alpha_n}}) \cdot \text{vol}(W_{\alpha_1 \dots \alpha_n}).$$

Für zwei beliebige Zerlegungen \mathcal{P} und \mathcal{P}^* von W gilt wieder

$$\underline{S}(f, \mathcal{P}) \leq \bar{S}(f, \mathcal{P}^*).$$

Wir erhalten also

$$\begin{aligned} \sup_{\mathcal{P}} \underline{S}(f, \mathcal{P}) &:= \sup \{ \underline{S}(f, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } W \} \\ &\leq \inf \{ \bar{S}(f, \mathcal{P}) \mid \mathcal{P} \text{ Zerlegung von } W \} \\ &=: \inf_{\mathcal{P}} \bar{S}(f, \mathcal{P}). \end{aligned}$$

Definition 9.5. Eine Funktion $f : W \rightarrow \mathbb{R}$ heißt *Riemann-integrierbar*, wenn sie beschränkt ist und $\sup_{\mathcal{P}} \underline{S}(f, \mathcal{P}) = \inf_{\mathcal{P}} \bar{S}(f, \mathcal{P})$ gilt.

In diesem Fall heißt die Zahl

$$\int_W f(x) dx := \sup_{\mathcal{P}} \underline{S}(f, \mathcal{P}) = \inf_{\mathcal{P}} \bar{S}(f, \mathcal{P})$$

das *Riemann-Integral* von f über W . (Dabei ist $x = (x_1, \dots, x_n) \in \mathbb{R}^n$).

Auch im Mehrdimensionalen gilt:

Satz 9.4 (Riemannsches Integrierbarkeitskriterium)

Sei W ein Quader im \mathbb{R}^n . Eine Funktion $f : W \rightarrow \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn sie beschränkt ist und für jedes $\varepsilon > 0$ eine Zerlegung \mathcal{P} von W existiert, so daß

$$\bar{S}(f, \mathcal{P}) - \underline{S}(f, \mathcal{P}) < \varepsilon.$$

Wie im 1-dimensionalen Fall beweist man folgende Eigenschaften des Riemann-Integrals:

Satz 9.5 Seien $f, g : W \subset \mathbb{R}^n \rightarrow \mathbb{R}$ Riemann-integrierbare Funktionen auf einem Quader W . Dann gilt

1. Für reelle Zahlen λ und μ ist die Funktion $\lambda f + \mu g$ ebenfalls Riemann-integrierbar und

$$\int_W (\lambda \cdot f(x) + \mu \cdot g(x)) dx = \lambda \cdot \int_W f(x) dx + \mu \cdot \int_W g(x) dx.$$

2. Das Produkt $f \cdot g : W \rightarrow \mathbb{R}$ ist Riemann-integrierbar.
3. Der Betrag $|f| : W \rightarrow [0, \infty)$ ist Riemann-integrierbar und

$$\left| \int_W f(x) dx \right| \leq \int_W |f(x)| dx.$$

4. Ist $f \leq g$, so gilt $\int_W f(x) dx \leq \int_W g(x) dx$.

Im folgenden Satz sind die drei wichtigsten Beispiele für Riemann-integrierbare Funktionen zusammengestellt, die wir im Folgenden benutzen werden.

Satz 9.6 Sei $W \subset \mathbb{R}^n$ ein Quader.

1. Jede stetige Funktion $f : W \rightarrow \mathbb{R}$ ist Riemann-integrierbar.
2. Sei $f : W \rightarrow \mathbb{R}$ beschränkt und die Menge der Unstetigkeitsstellen von f eine Jordansche Nullmenge. Dann ist f Riemann-integrierbar.
3. Sei $A \subset \mathbb{R}^n$ eine Jordan-messbare Menge und W ein Quader im \mathbb{R}^n , der A enthält. Mit $\chi_A : W \rightarrow \mathbb{R}$ bezeichnen wir die charakteristische Funktion von A , das heißt die Funktion

$$\chi_A(x) := \begin{cases} 1 & \text{falls } x \in A, \\ 0 & \text{falls } x \in W \setminus A. \end{cases}$$

Dann ist χ_A Riemann-integrierbar und für das Jordan-Volumen von A gilt

$$\text{vol}(A) = \int_W \chi_A(x) dx.$$

Beweis. Wir beweisen 3). Sei \mathcal{P} eine Zerlegung von W . Dann gilt

$$\begin{aligned} \overline{S}(\chi_A, \mathcal{P}) &= \sum_{W^* \in \mathcal{P}} \sup(\chi_A|_{W^*}) \cdot \text{vol}(W^*) = \sum_{W^* \cap A \neq \emptyset} \text{vol}(W^*) \\ &\leq \sum_{W^* \cap \text{cl}(A) \neq \emptyset} \text{vol}(W^*) = \overline{S}(A, W, \mathcal{P}), \end{aligned}$$

$$\begin{aligned} \underline{S}(\chi_A, \mathcal{P}) &= \sum_{W^* \in \mathcal{P}} \inf(\chi_A|_{W^*}) \cdot \text{vol}(W^*) = \sum_{W^* \subset A} \text{vol}(W^*) \\ &\geq \sum_{W^* \subset \text{Int}(A)} \text{vol}(W^*) = \underline{S}(A, W, \mathcal{P}). \end{aligned}$$

Somit gilt

$$\underline{S}(A, W, \mathcal{P}) \leq \underline{S}(\chi_A, \mathcal{P}) \leq \overline{S}(\chi_A, \mathcal{P}) \leq \overline{S}(A, W, \mathcal{P})$$

und folglich

$$\underline{\text{vol}}(A) = \sup_{\mathcal{P}} \underline{S}(A, W, \mathcal{P}) \leq \sup_{\mathcal{P}} \underline{S}(\chi_A, \mathcal{P}) \leq \inf_{\mathcal{P}} \overline{S}(\chi_A, \mathcal{P}) \leq \inf_{\mathcal{P}} \overline{S}(A, W, \mathcal{P}) = \overline{\text{vol}}(A). \quad (9.1)$$

Da A Jordan-meßbar ist, gilt $\text{vol}(A) = \overline{\text{vol}}(A) = \text{vol}(A)$. Folglich muß in (9.1) überall Gleichheit stehen. D.h. χ_A ist Riemann-integrierbar und es gilt

$$\text{vol}(A) = \int_W \chi_A(x) dx.$$

□

Wir wollen jetzt Funktionen über Jordan-meßbaren Mengen $A \subset \mathbb{R}^n$ integrieren.

Ist W ein Quader, der A enthält, und $f : W \rightarrow \mathbb{R}$ Riemann-integrierbar, so ist auch das Produkt $f \cdot \chi_A : W \rightarrow \mathbb{R}$ Riemann-integrierbar (Satz 9.5 und Satz 9.6). Wir schneiden f hierbei außerhalb von A ab und ersetzen die Funktionswerte dort durch Null. Auf A stimmen f und $f \cdot \chi_A$ überein.

Haben wir eine nur auf A definierte Funktion $f : A \rightarrow \mathbb{R}$ gegeben, so setzen wir f zunächst auf W durch Null fort, d.h. wir definieren zusätzlich $f(x) := 0$ für alle $x \in W \setminus A$. Ist f z.B. auf A stetig, so liegen die Unstetigkeitsstellen der Fortsetzung auf dem Rand ∂A . Da A Jordan-meßbar ist, ist der Rand ∂A eine Jordansche Nullmenge und deshalb 'klein' genug, um die Riemann-Integrierbarkeit der Fortsetzung zu sichern (siehe Satz 9.6).

Definition 9.6. Sei $A \subset W$ Jordan-meßbar und $f : W \rightarrow \mathbb{R}$ Riemann-integrierbar. Unter dem Riemann-Integral von f über A verstehen wir die Zahl

$$\int_A f(x) dx := \int_W (f \cdot \chi_A)(x) dx.$$

Insbesondere gilt für das Jordan-Volumen einer Jordan-meßbaren Menge $A \subset \mathbb{R}^n$

$$\text{vol}(A) = \int_A 1 dx.$$

Die Rechenregeln von Satz 9.5 gelten auch für das Integral über eine Jordan-meßbare Menge A . Außerdem erhält man durch Einsetzen der Definition:

Satz 9.7 Sei $f : W \rightarrow \mathbb{R}$ Riemann-integrierbar, $A, B \subset W$ Jordan-meßbar und disjunkt und $N \subset W$ eine Jordansche Nullmenge. Dann gilt

$$\begin{aligned} \int_{A \cup B} f(x) dx &= \int_A f(x) dx + \int_B f(x) dx, \\ \int_A f(x) dx &= \int_{A \setminus N} f(x) dx, \\ \int_N f(x) dx &= 0. \end{aligned}$$

Völlig analog man zum 1-dimensionalen Fall kann man den Mittelwertsatz beweisen.

Satz 9.8 (Mittelwertsatz)

Sei $f : W \rightarrow \mathbb{R}$ stetig, $g : W \rightarrow \mathbb{R}$ Riemann-integrierbar und $g \geq 0$. Dann existiert ein $\xi \in W$ mit

$$\int_W f(x)g(x) dx = f(\xi) \cdot \int_W g(x) dx.$$

Insbesondere existiert ein $\xi_0 \in W$ mit

$$\int_W f(x) dx = f(\xi_0) \cdot \text{vol}(W).$$

In den folgenden beiden Abschnitten werden wir Methoden kennenlernen, mit denen man die Berechnung des mehrdimensionalen Riemann-Integrals auf die Berechnung des Integrals im 1-dimensionalen Fall zurückführen kann. Die Methoden für Letzteres haben wir in Kapitel 6 behandelt.

9.4 Der Satz von Fubini und das Prinzip von Cavalieri

Der zentrale Satz der mehr-dimensionalen Integralrechnung ist der Satz von Fubini. Mit diesem Satz führt man die Integration über einem Quader $W_1 \times W_2 \subset \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$ auf die Integration über die Faktoren W_1 und W_2 zurück. Dies erlaubt es dann, die Integration auf die 1-dimensionalen Riemann-Integrale zurückzuführen, die wir aus Kapitel 6 kennen.

Satz 9.9 (Satz von Fubini)

Seien $W_1 \subset \mathbb{R}^n$ und $W_2 \subset \mathbb{R}^m$ zwei Quader, $W := W_1 \times W_2$ ihr Produkt im \mathbb{R}^{n+m} und $f : W_1 \times W_2 \rightarrow \mathbb{R}$ eine Riemann-integrierbare Funktion.

1. Sei $x \in W_1$ fixiert und $g_x : W_2 \rightarrow \mathbb{R}$ definiert durch $g_x(y) := f(x, y)$. Wir setzen voraus, dass die Funktion g_x für alle $x \in W_1$ Riemann-integrierbar ist. Dann ist die Funktion $G : W_1 \rightarrow \mathbb{R}$, definiert durch die partielle Integration über W_2

$$G(x) := \int_{W_2} g_x(y) dy = \int_{W_2} f(x, y) dy,$$

Riemann-integrierbar und es gilt

$$\int_{W_1 \times W_2} f(x, y) d(x, y) = \int_{W_1} G(x) dx = \int_{W_1} \left(\int_{W_2} f(x, y) dy \right) dx.$$

2. Sei $y \in W_2$ fixiert und $h_y : W_1 \rightarrow \mathbb{R}$ definiert durch $h_y(x) := f(x, y)$. Wir setzen voraus, dass die Funktion h_y für alle $y \in W_2$ Riemann-integrierbar ist. Dann ist die Funktion $H : W_2 \rightarrow \mathbb{R}$, definiert durch die partielle Integration über W_1

$$H(y) := \int_{W_1} h_x(y) dx = \int_{W_1} f(x, y) dx,$$

Riemann-integrierbar und es gilt

$$\int_{W_1 \times W_2} f(x, y) d(x, y) = \int_{W_2} H(y) dy = \int_{W_2} \left(\int_{W_1} f(x, y) dx \right) dy.$$

Bemerkung: Ist $f : W_1 \times W_2 \rightarrow \mathbb{R}$ Riemann-integrierbar, so müssen die Funktionen $g_x : W_2 \rightarrow \mathbb{R}$ und $h_y : W_1 \rightarrow \mathbb{R}$ nicht unbedingt Riemann-integrierbar sein. Dies muß man deshalb zusätzlich voraussetzen.

Ist $f : W_1 \times W_2 \rightarrow \mathbb{R}$ stetig, so sind $g_x : W_2 \rightarrow \mathbb{R}$ und $h_y : W_1 \rightarrow \mathbb{R}$ ebenfalls stetig und folglich auch Riemann-integrierbar. In diesem Fall können wir den Satz von Fubini ohne zusätzliche Voraussetzungen über g_x und h_y anwenden und erhalten die Formel für die iterierte Integration

$$\int_{W_1 \times W_2} f(x, y) d(x, y) = \int_{W_1} \left(\int_{W_2} f(x, y) dy \right) dx = \int_{W_2} \left(\int_{W_1} f(x, y) dx \right) dy.$$

Beweis des Satzes von Fubini. Sei $\mathcal{P} = (\mathcal{P}_1, \dots, \mathcal{P}_n)$ eine Zerlegung von W_1 mit den Teilwürfeln W^* , $\tilde{\mathcal{P}} = (\tilde{\mathcal{P}}_1, \dots, \tilde{\mathcal{P}}_m)$ eine Zerlegung von W_2 mit den Teilwürfeln W^{**} . Dann ist $\mathcal{Q} := (\mathcal{P}, \tilde{\mathcal{P}})$ eine Zerlegung von $W_1 \times W_2$ mit den Teilwürfeln $W^* \times W^{**}$ und es gilt

$$\underline{S}(f, W_1 \times W_2, \mathcal{Q}) = \sum_{W^* \in \mathcal{P}, W^{**} \in \tilde{\mathcal{P}}} \inf(f|_{W^* \times W^{**}}) \cdot \text{vol}(W^*) \cdot \text{vol}(W^{**}).$$

Ist $x \in W^*$, so gilt $\inf(f|_{W^* \times W^{**}}) \leq \inf(g_x|_{W^{**}})$. Für festes W^* und $x \in W^*$ erhalten wir also

$$\begin{aligned} \sum_{W^{**} \in \tilde{\mathcal{P}}} \inf(f|_{W^* \times W^{**}}) \cdot \text{vol}(W^{**}) &\leq \sum_{W^{**} \in \tilde{\mathcal{P}}} \inf(g_x|_{W^{**}}) \cdot \text{vol}(W^{**}) = \underline{S}(g_x, W_2, \tilde{\mathcal{P}}) \\ &\leq \int_{W_2} g_x(y) dy = G(x). \end{aligned}$$

Daraus folgt

$$\sum_{W^{**} \in \tilde{\mathcal{P}}} \inf(f|_{W^* \times W^{**}}) \cdot \text{vol}(W^{**}) \leq \inf\{G(x) \mid x \in W^*\} \cdot \text{vol}(W^*) = \inf(G|_{W^*}).$$

Somit ist

$$\underline{S}(f, W_1 \times W_2, \mathcal{Q}) \leq \sum_{W^* \in \mathcal{P}} \inf(G|_{W^*}) \cdot \text{vol}(W^*) = \underline{S}(G, W_1, \mathcal{P}). \quad (*)$$

Analog zeigt man,

$$\bar{S}(f, W_1 \times W_2, \mathcal{Q}) \geq \bar{S}(G, W_1, \mathcal{P}). \quad (**)$$

Daraus erhalten wir

$$\underline{S}(f, W_1 \times W_2, \mathcal{Q}) \stackrel{(*)}{\leq} \underline{S}(G, W_1, \mathcal{P}) \leq \overline{S}(G, W_1, \mathcal{P}) \stackrel{(**)}{\leq} \overline{S}(f, W_1 \times W_2, \mathcal{Q}).$$

Da f auf $W_1 \times W_2$ integrierbar ist, folgt

$$\int_{W_1 \times W_2} f(x, y) d(x, y) = \sup_{\mathcal{Q}} \underline{S}(f, W_1 \times W_2, \mathcal{Q}) = \inf_{\mathcal{Q}} \overline{S}(f, W_1 \times W_2, \mathcal{Q}).$$

Folglich gilt

$$\sup \underline{S}(G, W_1, \mathcal{P}) = \inf \overline{S}(G, W_1, \mathcal{P}) = \int_{W_1 \times W_2} f(x, y) d(x, y).$$

$G : W_1 \rightarrow \mathbb{R}$ ist also Riemann-integrierbar und es gilt

$$\int_{W_1} \left(\int_{W_2} f(x, y) dy \right) dx = \int_{W_1 \times W_2} f(x, y) d(x, y).$$

Offensichtlich gilt der gleiche Beweis, wenn man die Rollen von W_1 und W_2 vertauscht. Dies beweist dann den 2. Teil von Satz 9.9. □

Folgerung 9.1 (Iterierte Integration stetiger Funktionen auf Quadern)

Sei $f : W = [a_1, b_1] \times \dots \times [a_n, b_n] \rightarrow \mathbb{R}$ eine stetige Funktion. Dann kann man das Riemann-Integral von f über W durch schrittweise Integration nach jeweils einer Variablen berechnen und das Ergebnis hängt nicht von der gewählten Reihenfolge der Integration ab. D.h. für jede Permutation σ von $\{1, 2, \dots, n\}$ gilt:

$$\begin{aligned} \int_W f(x) dx &= \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} \dots \int_{a_{n-1}}^{b_{n-1}} \left(\int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n \right) dx_{n-1} \dots \right) dx_1 \\ &= \int_{a_{\sigma(1)}}^{b_{\sigma(1)}} \left(\int_{a_{\sigma(2)}}^{b_{\sigma(2)}} \dots \int_{a_{\sigma(n-1)}}^{b_{\sigma(n-1)}} \left(\int_{a_{\sigma(n)}}^{b_{\sigma(n)}} f(x_1, \dots, x_n) dx_{\sigma(n)} \right) dx_{\sigma(n-1)} \dots \right) dx_{\sigma(1)}. \end{aligned}$$

Diese Rechenregel kann man auf die Integration über meßbaren Mengen $A \subset \mathbb{R}^n$ übertragen. Dazu schneidet man A mit Hyperebenen, die parallel zu den Koordinatenebenen sind.

Definition 9.7. Sei $A \subset \mathbb{R}^n$ und $c \in \mathbb{R}$ eine fixierte Zahl. Die Menge

$$A_c := \{(x_1, \dots, x_{n-1}) \mid (x_1, \dots, x_{n-1}, c) \in A\} = \text{proj}_{\mathbb{R}^{n-1}} (A \cap \{x_n = c\}) \subset \mathbb{R}^{n-1}.$$

heißt c -Schnitt von A entlang der n -ten Koordinate. Analog definiert man den c -Schnitt entlang der anderen Koordinaten.

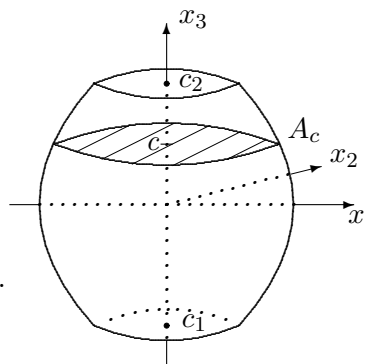
Folgerung 9.2 (Iterierte Integration stetiger Funktionen auf Jordan-Mengen)

Sei $A \subset \mathbb{R}^n$ eine Jordan-messbare Menge, $c \in \mathbb{R}$ und sei der c -Schnitt

$$A_c := \{(x_1, \dots, x_{n-1} \mid (x_1, \dots, x_{n-1}, c) \in A\} \subset \mathbb{R}^{n-1}$$

Jordan-messbar für alle $c \in \mathbb{R}$. Wir wählen ein Intervall $[c_1, c_2]$ mit $A_c = \emptyset$ für alle $c \notin [c_1, c_2]$. Sei $f : A \rightarrow \mathbb{R}$ eine stetige Funktion. Dann gilt

$$\int_A f(x) dx = \int_{c_1}^{c_2} \left(\int_{A_c} f(x_1, \dots, x_{n-1}, c) d(x_1, \dots, x_{n-1}) \right) dc.$$



Die analoge Formel gilt für die Schnitte entlang der anderen Koordinaten.

Beweis. Da A beschränkt ist, gibt es einen Quader $W := W_1 \times [c_1, c_2]$, der A enthält. Nach Definition des Riemann-Integrals von f über A gilt

$$\int_A f(x) dx = \int_W \chi_A(x) f(x) dx.$$

Mit Satz dem Satz von Fubini folgt dann

$$\int_A f(x) dx = \int_{W_1 \times [c_1, c_2]} \chi_A(x) f(x) dx = \int_{c_1}^{c_2} \left(\int_{W_1} (\chi_A(y, c) \cdot f(y, c)) dy \right) dc,$$

wobei $y = (x_1, \dots, x_{n-1})$. Es gilt $y \in A_c$ genau dann, wenn $(y, c) \in A$. Folglich ist $\chi_A(y, c) = \chi_{A_c}(y)$ und wir erhalten

$$\int_A f(x) dx = \int_{c_1}^{c_2} \left(\int_{W_1} \chi_{A_c}(y) \cdot f(y, c) dy \right) dc = \int_{c_1}^{c_2} \left(\int_{A_c} f(y, c) dy \right) dc$$

□

Setzt man der letzten Folgerung $f = 1$ ein, so erhält man eine Methode zur Berechnung des n -dimensionalen Jordan-Volumens $vol_n(A)$ von $A \in \mathcal{J}(\mathbb{R}^n)$ durch die $(n - 1)$ -dimensionalen Jordan-Volumen $vol_{n-1}(A_c)$ der Schnittmengen $A_c \subset \mathbb{R}^{n-1}$.

Folgerung 9.3 (Das Prinzip von Cavalieri zur Volumenberechnung)

Sei $A \subset \mathbb{R}^n$ eine Jordan-messbare Menge und seien alle c -Schnitte $A_c \subset \mathbb{R}^{n-1}$ nach einer fixierten Variablen ebenfalls Jordan-messbar. Dann gilt

$$vol_n(A) = \int_{c_1}^{c_2} vol_{n-1}(A_c) dc,$$

wobei $[c_1, c_2]$ so gewählt ist, dass $A_c = \emptyset$ für alle $c \notin [c_1, c_2]$.

Beispiel 1: Sei $f : W = [a_1, b_1] \times [a_2, b_2] \rightarrow \mathbb{R}$ eine stetige Funktion der Form

$$f(x, y) = h_1(x) \cdot h_2(y).$$

Dann gilt

$$\int_W f(x, y) d(x, y) = \int_{a_1}^{b_1} h_1(x) dx \cdot \int_{a_2}^{b_2} h_2(y) dy.$$

Wir benutzen zum Beweis die Iterationsformel und erhalten

$$\begin{aligned} \int_W f(x, y) d(x, y) &= \int_{a_1}^{b_1} \left(\int_{a_2}^{b_2} h_1(x) \cdot h_2(y) dy \right) dx = \int_{a_1}^{b_1} h_1(x) \cdot \left(\int_{a_2}^{b_2} h_2(y) dy \right) dx \\ &= \int_{a_1}^{b_1} h_1(x) dx \cdot \int_{a_2}^{b_2} h_2(y) dy. \end{aligned}$$

Eine analoge Formel gilt für drei oder mehr Variablen.

Beispiel 2: Wir betrachten die Funktion $f : [0, 1] \times [0, 1] \rightarrow \mathbb{R}$,

$$f(x, y) := x^2 + y^4,$$

und berechnen das Integral von f über $[0, 1] \times [0, 1]$:

$$\begin{aligned} \int_{[0,1] \times [0,1]} f(x, y) d(x, y) &= \int_0^1 \left(\int_0^1 (x^2 + y^4) dx \right) dy = \int_0^1 \left(\left. \left(\frac{1}{3}x^3 + y^4x \right) \right|_0^1 \right) dy \\ &= \int_0^1 \left(\frac{1}{3} + y^4 \right) dy = \frac{1}{3} + \frac{1}{5} = \frac{8}{15}. \end{aligned}$$

Beispiel 3: Wir betrachten die Kreisscheibe $D := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\}$ und berechnen das Integral der Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, mit $f(x, y) := x + y$, über D .

Der Rand von D , d.h. die Menge $\partial D = \{(x, y) \mid f(x, y) = x^2 + y^2 = 1\} = S^1$, ist eine Jordansche Nullmenge (man kann sie als Graphen von Funktionen darstellen). Also ist die Kreisscheibe D Jordan-messbar und wir können rechnen

$$\int_D f(x, y) d(x, y) = \int_{-1}^1 \left(\int_{D_c} f(x, c) dx \right) dc.$$

Da

$$D_c = \{x \in \mathbb{R} \mid x^2 + c^2 \leq 1\} = [-\sqrt{1 - c^2}, \sqrt{1 - c^2}]$$

folgt

$$\begin{aligned} \int_D f(x, y) d(x, y) &= \int_{-1}^1 \left(\int_{-\sqrt{1-c^2}}^{\sqrt{1-c^2}} (x+c) dx \right) dc = \int_{-1}^1 \left(\frac{1}{2}x^2 + cx \right) \Big|_{-\sqrt{1-c^2}}^{\sqrt{1-c^2}} dc \\ &= \int_{-1}^1 2c\sqrt{1-c^2} dc = 0. \quad (\text{ungerade Funktion}) \end{aligned}$$

Beispiel 4: Das Volumen von Mengen unter Funktionsgraphen.

Sei $A \subset \mathbb{R}^n$ Jordan-meßbar, $f : A \rightarrow [0, \infty)$ stetig und

$$\Omega_f := \{(x, y) \in A \times \mathbb{R} \mid 0 \leq y \leq f(x)\}$$

die Menge unter dem Funktionsgraphen von f . Die Menge Ω_f ist Jordan-meßbar, da der Rand von Ω_f aus Graphen stetiger Funktionen besteht. Für das Jordan-Volumen von Ω_f gilt:

$$\text{vol}(\Omega_f) = \int_A f(x) dx.$$

Ist $A = [a, b] \subset \mathbb{R}$, so entspricht dies dem Flächeninhalt von Ω_f , den wir in Kapitel 6 definiert haben.

Zur Berechnung des Integrals wählen wir einen Quader $W_1 \subset \mathbb{R}^n$, der A enthält. Nach Definition ist dann

$$\text{vol}(\Omega_f) = \int_{W_1 \times [0, \sup f]} \chi_{\Omega_f}(x, y) d(x, y) = \int_{W_1} \left(\int_0^{\sup f} \chi_{\Omega_f}(x, y) dy \right) dx.$$

Es gilt

$$\chi_{\Omega_f}(x, y) = \begin{cases} \chi_A(x) & \text{falls } y \in [0, f(x)], \\ 0 & \text{sonst.} \end{cases}$$

Daraus folgt

$$\text{vol}(\Omega_f) = \int_{W_1} \left(\int_0^{f(x)} \underbrace{\chi_A(x)}_{\text{konst.}} dy \right) dx = \int_{W_1} \chi_A(x) \cdot f(x) dx = \int_A f(x) dx.$$

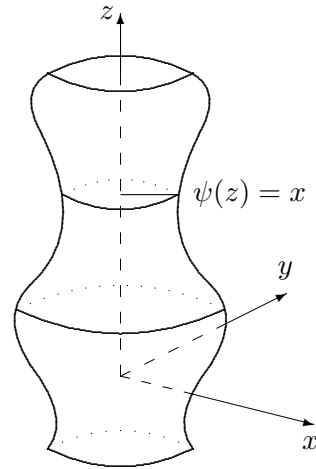
Beispiel 5: *Das Volumen von Rotationskörpern im \mathbb{R}^3 .*

Sei $\psi : [a, b] \rightarrow \mathbb{R}$ eine nicht-negative stetige Funktion. Wir betrachten in der zx -Ebene den Graphen der Funktion $x = \psi(z)$ und rotieren diese Kurve um die z -Achse. Die Menge V_ψ , die von der dabei entstehenden Fläche eingeschlossen wird, heißt der von ψ erzeugte *Rotationskörper*. Sie ist beschrieben durch

$$V_\psi := \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 \leq \psi(z)^2\}.$$

Der Rotationskörper V_ψ ist Jordan-messbar, da sein Rand eine reguläre Fläche, d.h. eine Jordansche Nullmenge ist. Dann gilt für das Volumen von V_ψ

$$\text{vol}(V_\psi) = \pi \cdot \int_a^b \psi^2(z) dz.$$



Zum Beweis betrachten wir die Schnitte von V_ψ nach der 3. Koordinate. Der z -Schnitt von V_ψ ist die Kreisscheibe $D_{\psi(z)}^2$ vom Radius $\psi(z)$. Eine Kreisscheibe D_r^2 vom Radius r hat das Jordan-Volumen (d.h. in diesem Fall den Flächeninhalt) $\text{Area}(D_r^2) = \pi r^2$. Aus dem Prinzip von Cavalieri folgt damit

$$\text{vol}(V_\psi) = \int_a^b \text{Area}(D_{\psi(z)}^2) dz = \pi \cdot \int_a^b \psi^2(z) dz.$$

Beispiel 6: *Das Volumen der Kugel im \mathbb{R}^3 vom Radius R .*

Wir bezeichnen die abgeschlossene Kugel im \mathbb{R}^3 vom Radius R mit D_R^3 , d.h.

$$D_R^3 := \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 \leq R^2\}.$$

Für das Volumen von D_R^3 gilt

$$\text{vol}(D_R^3) = \frac{4}{3} \pi R^3.$$

Zum Beweis betrachten wir die Kugel als Rotationskörper. Die rotierende Funktion in der zx -Ebene ist hierbei $\psi : [-R, R] \rightarrow \mathbb{R}$ mit

$$\psi(z) = \sqrt{R^2 - z^2}.$$

Folglich gilt nach Beispiel 5:

$$\text{vol}(D_R^3) = \pi \int_{-R}^R (R^2 - z^2) dz = \pi \left(R^2 z \Big|_{-R}^R - \frac{1}{3} z^3 \Big|_{-R}^R \right) = \pi \left(2R^3 - \frac{2}{3} R^3 \right) = \frac{4}{3} \pi R^3.$$

9.5 Die Transformationsformel für Riemann-Integrale

Die zweite wichtige Formel zur Integralberechnung ist die Transformationsformel. Mit dieser Formel kann man das Riemann-Integral über einer Jordan-meßbaren Menge B , die "kompliziert" aussieht, auf ein Riemann-Integral über einer Menge A zurückführen, die "einfacher" aussieht (die z.B. ein Quader ist). Dazu benutzt man eine Koordinatentransformation, die die Menge B durch Koordinaten aus der einfacheren Menge A beschreibt.

Satz 9.10 Sei $\phi : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^n$ ein Diffeomorphismus (d.h. eine Koordinatentransformation) zwischen den offenen Mengen U und V , $f : V \rightarrow \mathbb{R}$ eine stetige Funktion und $A \subset U$ Jordan-meßbar. Dann ist $\phi(A)$ ebenfalls Jordan-meßbar und es gilt

$$\int_{\phi(A)} f(y) dy = \int_A f(\phi(x)) \cdot |\det J\phi_x| dx.$$

Diese Formel verallgemeinert die Substitutionsregel für Riemann-Integrale von Funktionen einer Variablen: Sei $\phi : [a, b] \rightarrow [c, d]$ ein Diffeomorphismus. In diesem Fall ist die Jacobi-Matrix $J\phi_x = \phi'(x)$. Wir führen die Substitution $y = \phi(x)$, $dy = \phi'(x) dx$ durch. Dann gilt

$$\int_c^d f(y) dy = \int_{\phi^{-1}(c)}^{\phi^{-1}(d)} f(\phi(x)) \phi'(x) dx.$$

Es ist entweder $\phi'(x) > 0$ für alle $x \in [a, b]$ oder $\phi'(x) < 0$ für alle $x \in [a, b]$, denn ϕ' ist stetig und $\phi'(x) \neq 0$, da ϕ ein Diffeomorphismus ist. Ist $\phi'(x) > 0$ für alle $x \in [a, b]$, so ist ϕ streng monoton wachsend und es gilt $\phi^{-1}(c) = a$ und $\phi^{-1}(d) = b$. Ist $\phi'(x) < 0$ für alle $x \in [a, b]$, so ist ϕ streng monoton fallend und es gilt $\phi^{-1}(c) = b$ und $\phi^{-1}(d) = a$. Folglich erhalten wir

$$\int_c^d f(y) dy = \int_a^b f(\phi(x)) \cdot |\phi'(x)| dx = \int_a^b f(\phi(x)) \cdot |\det J\phi_x| dx.$$

In den folgenden Beispielen demonstrieren wir die Anwendung der Transformationsformel.

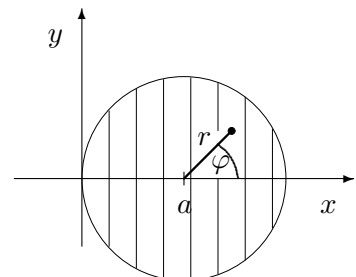
Beispiel 7: Wir berechnen das Integral $\int_{(x-a)^2+y^2 \leq a^2} (x^2 + y^2) d(x, y)$.

Der Integrationsbereich ist die Kreisscheibe um $(a, 0)$ vom Radius a , d.h.

$$B := \{(x, y) \in \mathbb{R}^2 \mid (x - a)^2 + y^2 \leq a^2\}.$$

Um das Integral zu berechnen, beschreiben wir B (bis auf eine Nullmenge) durch Polarkoordinaten um $(a, 0)$:

Jeden Punkt (x, y) von B schreiben wir als $x = a + r \cos \varphi$ und $y = r \sin \varphi$. Die Abbildung



$$\begin{aligned} \phi : (0, \infty) \times (0, 2\pi) &\longrightarrow \mathbb{R}^2 \setminus \underbrace{([0, \infty) \times \{0\})}_{:=N \text{ Nullmenge}} \\ (r, \varphi) &\longmapsto (a + r \cos \varphi, r \sin \varphi). \end{aligned}$$

ein Diffeomorphismus. Es gilt $B \setminus N = \phi((0, a) \times (0, 2\pi))$. Die Determinante der Jacobi-Matrix von ϕ ist

$$\det J\phi_{(r,\varphi)} = \det \begin{pmatrix} \cos \varphi & -r \sin \varphi \\ \sin \varphi & r \cos \varphi \end{pmatrix} = r > 0.$$

Wir erinnern uns daran, dass man zum Integrationsbereich Nullmengen beliebig hinzunehmen oder herausnehmen kann, ohne das Integral zu verändern (siehe Satz 9.7). Aus der Transformationsformel folgt dann

$$\begin{aligned} \int_B (x^2 + y^2) d(x, y) &= \int_{B \setminus N} (x^2 + y^2) d(x, y) \\ &\stackrel{\text{Trafo}}{=} \int_{(0,a) \times (0,2\pi)} (a + r \cos \varphi)^2 + r^2 \sin^2 \varphi \cdot |\det J\phi_{(r,\varphi)}| d(r, \varphi) \\ &= \int_{[0,a] \times [0,2\pi]} (a^2 r + 2r^2 a \cos \varphi + r^3 \cos^2 \varphi + r^3 \sin^2 \varphi) d(r, \varphi) \\ &\stackrel{\text{Fubini}}{=} \int_0^a \left(\int_0^{2\pi} (a^2 r + 2r^2 a \cos \varphi + r^3) d\varphi \right) dr \\ &= 2\pi \int_0^a (a^2 r + r^3) dr = 2\pi \left(a^2 \frac{r^2}{2} \Big|_0^a + \frac{r^4}{4} \Big|_0^a \right) \\ &= \frac{3}{2} \pi a^4. \end{aligned}$$

Beispiel 8: Wir berechnen das Integral $\int_{x^2+y^2+z^2 \leq 1} |xy| d(x, y, z)$.

Der Integrationsbereich ist die Kugel D_1^3 vom Radius 1. Wir betrachten die Kugelkoordinaten im \mathbb{R}^3

$$\phi : (0, \infty) \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \times (0, 2\pi) \longrightarrow \mathbb{R}^3 \setminus \underbrace{([0, \infty) \times \{0\} \times \mathbb{R})}_{:=L \text{ Nullmenge}}$$

mit

$$\phi(r, u, v) := (r \cos u \cos v, r \cos u \sin v, r \sin u).$$

Die Menge $D_1^3 \setminus L$ wird durch die Kugelkoordinaten aus der Menge

$$A := (0, 1] \times \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \times (0, 2\pi)$$

beschrieben, d.h. es gilt $\phi(A) = D_1^3 \setminus L$. ϕ ist ein Diffeomorphismus mit der Jacobi-Matrix

$$J\phi_{(r,u,v)} = \begin{pmatrix} \cos u \cdot \cos v & -r \cdot \sin u \cdot \cos v & -r \cdot \cos u \cdot \sin v \\ \cos u \cdot \sin v & -r \cdot \sin u \cdot \sin v & r \cdot \cos u \cdot \cos v \\ \sin u & r \cos u & 0 \end{pmatrix}.$$

Für die Determinante der Jacobi-Matrix berechnet man

$$|\det J\phi_{(r,u,v)}| = r^2 \cdot \cos u.$$

Wir benutzen wieder, dass man zu Integrationsgebieten Nullmenge beliebig hinzufügen oder wegnehmen kann, ohne das Integral zu verändern und erhalten aus der Transformationsformel

$$\begin{aligned} \int_{D_1^3} |xy| d(x, y, z) &= \int_{D_1^3 \setminus L} |xy| d(x, y, z) \\ &\stackrel{\text{Trafo}}{=} \int_A |r \cos u \cos v| \cdot |r \cos u \sin v| \cdot |\det J\phi_{(r,u,v)}| d(r, u, v) \\ &= \int_{[0,1] \times [-\frac{\pi}{2}, \frac{\pi}{2}] \times [0, 2\pi]} r^4 \cos^3 u \cdot |\sin v \cos v| d(r, u, v) \\ &\stackrel{\text{Bsp.1}}{=} \int_0^1 r^4 dr \cdot \int_{-\pi/2}^{\pi/2} \cos^3 u du \cdot \int_0^{2\pi} \frac{1}{2} |\sin 2v| dv. \end{aligned}$$

Da

$$\frac{1}{2} \int_0^{2\pi} |\sin 2v| dv = \frac{1}{4} \int_0^{4\pi} |\sin w| dw = \int_0^{\pi} \sin w dw = 2,$$

und

$$\int_{-\pi/2}^{\pi/2} \cos^3 u du \stackrel{\text{P.Int}}{=} \frac{\cos^2 u \sin u}{3} \Big|_{\pi/2}^{\pi/2} + \frac{2}{3} \int_{-\pi/2}^{\pi/2} \cos u du = \frac{4}{3}$$

folgt

$$\int_{D_1^3} |xy| d(x, y, z) = \frac{8}{15}.$$

Beispiel 9: Geometrische Interpretation der Determinante

Sei $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein linearer Isomorphismus. Dann gilt für jede Jordan-meßbare Teilmenge $A \subset \mathbb{R}^n$

$$\text{vol}(L(A)) = |\det(L)| \cdot \text{vol}(A).$$

Zum Beweis benutzen wir, dass $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ein Diffeomorphismus ist und wenden die Transformationsformel an:

$$\begin{aligned} \text{vol}(L(A)) &= \int_{L(A)} 1 \, dx = \int_A |\det dL_x| \, dx = \int_A |\det L| \, dx \\ &= |\det L| \int_A 1 \, dx = |\det L| \cdot \text{vol}(A). \end{aligned}$$

Daraus erhalten wir die aus der Vorlesung *Lineare Algebra* bekannte geometrische Interpretation der Determinante einer regulären Matrix:

Sei $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ eine $(n \times n)$ -Matrix, deren Spaltenvektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ linear unabhängig sind. Wir betrachten den linearen Isomorphismus $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$, der die kanonischen Basisvektoren e_i auf \mathbf{a}_i abbildet ($i = 1, \dots, n$). Dann ist $(\mathbf{a}_1, \dots, \mathbf{a}_n)$ gerade die Darstellungsmatrix $[L]$ von L bzgl. der kanonischen Basis des \mathbb{R}^n . Die Vektoren $\mathbf{a}_1, \dots, \mathbf{a}_n$ spannen das Parallelotop

$$\mathcal{P}(\mathbf{a}_1, \dots, \mathbf{a}_n) := \left\{ \sum_{i=1}^n \lambda_i \mathbf{a}_i \mid 0 \leq \lambda_i \leq 1 \right\} \subset \mathbb{R}^n$$

auf. Dann gilt

$$\mathcal{P}(\mathbf{a}_1, \dots, \mathbf{a}_n) = L([0, 1]^n)$$

und wir erhalten

$$\text{vol}(\mathcal{P}(\mathbf{a}_1, \dots, \mathbf{a}_n)) = \text{vol}(L([0, 1]^n)) = |\det L| \cdot \text{vol}([0, 1]^n) = |\det[L]| = |\det(\mathbf{a}_1 \ \dots \ \mathbf{a}_n)|.$$

Dies liefert

$$|\det(\mathbf{a}_1 \ \dots \ \mathbf{a}_n)| = \text{vol}(\mathcal{P}(\mathbf{a}_1, \dots, \mathbf{a}_n)).$$

Sachverzeichnis

- $-A$, 23
- $A + B$, 23
- C^n , 231
- $E(z)$, 74
- $HP(x_n)$, 53
- $K_d(p, \varepsilon)$, 193
- $L(I)$, 25
- $L(\gamma)$, 172, 173
- R , 72
- \mathbb{R} , 24
- $T(f, x_0)$, 133
- $T_n(f, x_0)$, 128
- a^z , 77
- arccos, 102
- arctan, 102
- $arg(z)$, 43
- $arsinh$, 104
- $artanh$, 105
- cosh, 103
- cot, 101
- d , 189
- $d(p, q)$, 189
- df_p , 214
- e , 57, 75
- e^z , 75
- exp, 75
- f' , 108
- $f'(x_0)$, 108
- i , 41
- $\lim_{x \rightarrow a^-} f(x)$, 89
- lim inf, 56
- lim sup, 56
- ln, 76
- max A , 28
- min A , 28
- $n!$, 17
- sinh, 103
- sup A , 27
- tan, 101
- x über k , 18
- x^q , 37
- x^n , 33
- $x_n \rightarrow x$, 48
- $x_n \rightarrow p$, 194
- y^α , 234
- \bar{z} , 41
- $\alpha!$, 234
- ∂A , 198
- π , 99
- $(-\infty, a)$, 24
- $(-\infty, a]$, 24
- $(A \mid B)$, 26
- (a, ∞) , 24
- (a, b) , 24
- $(a, b]$, 24
- $(x_{n_j})_{j=1}^\infty$, 53
- (x_n) , 47
- $(x_n)_{n=1}^\infty$, 47
- $A \cdot B$, 23
- \mathbb{C} , 40
- $C^{(k)}(I, \mathbb{R})$, 114
- $C^m(U, \mathbb{R})$, 231
- $C^\infty(I; \mathbb{R})$, 114
- $Int(A)$, 198
- $K(z_0, R)$, 72
- $[\mathbb{K}, +, \cdot]$, 23
- arccot, 102
- $arcosh$, 104
- $arcoth$, 105
- arcsin, 102
- $cl(A)$, 198
- $f^{(n)}$, 114
- inf A , 27
- $\lim_{n \rightarrow \infty} x_n$, 48, 194

- $\lim_{x \rightarrow a^+} f(x)$, 89
 $x_n \xrightarrow{n \rightarrow \infty} p$, 194
 $x_n \xrightarrow{n \rightarrow \infty} x$, 48
 $[a, \infty)$, 24
 $[a, b)$, 24
 $[a, b]$, 24
 $\binom{x}{k}$, 18
 $\binom{x}{0}$, 18
 $\frac{\partial f}{\partial x_i}(p)$, 217
 $\frac{\partial^n f}{\partial x_{i_1} \cdots \partial x_{i_n}}$, 231
 $\frac{df}{dx}(x_0)$, 108
 \int , 165
 $\frac{a}{\leq}$, 24
 $\nabla_{\mathbf{a}} f(p)$, 216
 $\prod_{i=1}^n x_i$, 23
 $\sqrt[n]{x}$, 36
 $\sum_{i=1}^n$, 23
 $\sum_{k=1}^{\infty} x_k$, 62
 f'' , 114
- Abbildung, 11
 Abelscher Grenzwertsatz, 123
 abgeschlossen, 198
 Ableitung, 108, 214
 der inversen Funktion, 111
 erste, 108
 n-te, 114
 partielle, 217
 Abschluss einer Menge, 198
 absolut-konvergente Reihe, 63
 Abstand, 189
 Abstandsfunktion, 189
 abzählbar, 30
 Additionstheoreme, 98, 102, 104
 Äquivalenz zweier Aussagen, 3
 äußeres Volumen von A , 258
 alternierende harmonische Reihe, 69
 angeordneter Körper, 24
 Anordnungsaxiom der reellen Zahlen, 24
 archimedische Spirale, 180
 Archimedisches Axiom der reellen Zahlen, 28
 Arcuscosinus, 102
 Arcuscotangens, 102
 Arcussinus, 102
 Arcustangens, 102
 Areacosinus hyperbolicus, 104
 Areacotangens hyperbolicus, 105
 Areasinus hyperbolicus, 104
 Areatangens hyperbolicus, 105
 Astroide, 179
 Aussage, 1
 Banachscher Fixpunktsatz, 208
 Bernoullische Ungleichung, 34
 beschränkt, 27
 von oben, 27
 von unten, 27
 beschränkte Folge, 194
 beschränkte Teilmenge, 194
 bestimmt divergent, 54
 Betrag einer komplexen Zahl, 42
 Betrag einer reellen Zahl, 25
 bijektive Abbildung, 13
 Bild einer Abbildung, 11
 Binomialkoeffizient, 18
 Binomialreihe, 137
 Binomischer Satz, 33
 Cantorsches Diagonalverfahren, 31
 Cauchy-Bedingung, 196
 Cauchy-Folge, 60, 196
 Cauchy-Kriterium für Reihen, 63
 Cauchy-Produkt von Reihen, 70
 C^m -Funktion, 231
 Cosinus, 97
 Cosinus hyperbolicus, 103
 Cotangens, 101
 Dedekindscher Schnitt von \mathbb{R} , 25
 Definitionsbereich, 11
 Diagonalverfahren von Cantor, 31
 Diffeomorphismus, 239
 Differential, 214
 Differenz von Mengen, 9
 differenzierbar, 108, 214
 differenzierbare Funktion, 108
 Differenzierbarkeit
 hinreichende Bedingung für, 218
 direkter Beweis, 6
 Disjunktion zweier Aussagen, 2
 diskrete Metrik, 190
 diskreter metrischer Raum, 190
 divergente Folge, 48, 194
 divergente Reihe, 62
 Dreiecksungleichung, 25, 42, 189
 Durchschnitt von Mengen, 8
 echte Teilmenge, 8
 einfache Kurve, 176
 Einschränkung einer Abbildung, 12

- Element einer Menge, 7
- Ellipse, 170
 - Flächeninhalt, 179
 - Umfang, 175
- ε -Kugel, 193
- Eulerzahl, 57
- Exponentialfunktion, komplexe, 75

- Fakultät, 17
- Feinheit der Zerlegung eines Intervalls, 140
- Fixpunkt einer Funktion, 92
- Fixpunktsatz, Banachscher, 208
- Folge, 47
 - bestimmt divergente, 54
 - der Partialsummen, 62
 - divergente, 48, 194
 - konvergente, 48, 194
 - monoton fallende, 55
 - monoton wachsende, 55
 - monotone, 55
 - uneigentlich konvergente, 54
- Folge, nach $\pm\infty$ strebend, 54
- Folgenkriterium für Stetigkeit, 81
- Fortsetzung einer Abbildung, 12
- Fundamentalsatz der Algebra, 46, 95
- Funktion, 11
 - Gaußsche Zahlenebene, 42
 - geometrische Reihe, 35, 65
 - geometrische Summe, 35
 - geschlossene Kurve, 176
 - glatte Abbildung, 114
 - gleichmäßig konvergente
 - Folge von Funktionen, 162
 - gleichmäßig stetig, 85, 205
 - gleichmächtige Mengen, 32
 - Graph einer Abbildung, 11
 - Grenzwert
 - Eindeutigkeit, 48, 194
 - einer Folge, 48, 194
 - linksseitiger, 89
 - rechtsseitiger, 89
 - uneigentlicher, 54
 - Grenzwertsatz, Abelscher, 123

 - Häufungspunkt einer Folge, 53
 - höchstens abzählbar, 31
 - harmonische Reihe, 64
 - alternierende, 69, 135
 - Hebbare Unstetigkeitsstelle, 90
 - hinreichend, 3
 - homöomorph, 207
 - Homöomorphismus, 207
 - Hyperbel, 170
 - Hyperbelfunktionen, 103

 - Identität, 12
 - imaginäre Achse, 43
 - imaginäre Einheit, 41
 - Implikation, 3
 - implizite Funktionen
 - Satz über, 247
 - indirekter Beweis, 6
 - Induktion, vollständige, 16
 - Infimum, 27
 - injektive Abbildung, 13
 - Inklusionsabbildung, 12
 - innerer Punkt, 198
 - Inneres einer Menge, 198
 - inneres Volumen von A , 258
 - Integration
 - partielle, 151
 - Integrierbarkeitskriterium
 - Riemannsches, 142
 - Intervall
 - abgeschlossenes, 24
 - halboffenes, 24
 - Länge, 25
 - offenes, 24
 - Intervallschachtelung, 29
 - inverse Abbildung, 13
 - isoliertes lokales Maximum, 237
 - isoliertes lokales Minimum, 236
 - Isometrie, 192
 - isometrische metrische Räume, 192
 - Iterierte Integration, 265

 - Jordan-meißbare Menge, 258
 - Jordan-Volumen, 258
 - Jordansche Nullmenge, 258

 - Körper, 22
 - angeordneter, 24
 - vollständiger angeordneter, 26
 - Körperaxiom der reellen Zahlen, 23
 - kartesisches Produkt metrischer Räume, 191
 - Kettenregel, 110, 223
 - für partielle Ableitungen, 225
 - kleiner gleich, 24
 - kompakt, 200
 - Komplement von Mengen, 9
 - komplex-wertige Funktion, 79

- komplexe Funktion, 79
- komplexe Potenz der Eulerzahl, 75
- komplexe Zahl, 40
 - n -te Wurzel
 - geometrische Bedeutung, 45
 - algebraische Form, 41
 - Argument, 43
 - Betrag, 42
 - n -te Wurzel, 45
 - Polarform, 43
 - trigonometrische Darstellung, 43
- komplexe Zahlen
 - Dreiecksungleichung, 42
 - geometrische Deutung der Addition, 43
 - geometrische Deutung der Multiplikation, 44
- komplexe Zahlenebene, 42
- konjugiert komplex, 41
- Konjunktion zweier Aussagen, 2
- konkave Funktion, 119
- konstante Abbildung, 12
- Kontinuumshypothese, 33
- kontrahierende Abbildung, 208
- Kontraktion, 208
- konvergente Folge, 48, 194
- konvergente Reihe, 62
- Konvergenzkreis einer Potenzreihe, 72
- Konvergenzradius, 72–74
- konvexe Funktion, 118
- konvexe Menge, 206
- Koordinatentransformation, 239
- Kugelkoordinaten, 240
- Kurve
 - einfache, 176
 - geschlossene, 176
 - parametrisierte, 169
 - positiv-orientierte, 176
- L'Hospitalsche Regeln, 120
- Länge
 - einer rektifizierbaren Kurve, 172
 - einer stückweisen C^1 -Kurve, 174
 - eines Intervalls, 140
- leere Menge, 7
- Leibnizsche Sektorformel, 178
- Leibniz-Kriterium für alternierende Reihen, 69
- Leibniz-Reihe, 70
- Leibnizformel, 126
- limes inferior, 56
- limes superior, 56
- Linksinverse einer Abbildung, 14
- linksseitiger Grenzwert, 89
- Lipschitz-Konstante, 85, 205
- lipschitzstetig, 85, 205
- Logarithmus
 - natürlicher, 76
- logisches Gesetz, 4
- lokaler Extremwert
 - hinreichende Bedingung, 132
- lokales Maximum, 115, 236
- lokales Minimum, 115, 236
- Luftlinienabstand, 190
- Mächtigkeit, größere, 32
- Machinsche Formel, 127
- Majorantenkriterium, 66
- Mannheimer-Metrik, 190
- Maximum, 28
 - lokales, 236
 - striktes lokales, 237
- Maximum einer Funktion, 92
- Maximum,lokales, 115
- Menge, 7
- Metrik, 189
 - der französischen Eisenbahn, 190
 - diskrete, 190
- metrische Räume
 - isometrische, 192
 - kartesische Produkt, 191
- metrischer Raum, 189
 - abgeschlossene Teilmenge, 198
 - diskreter, 190
 - innerer Punkt, 198
 - offene Teilmenge, 197
 - Rand einer Teilmenge, 198
 - vollständiger, 196
- Minimum, 28
 - lokales, 115, 236
 - striktes lokales, 236
- Minimum einer Funktion, 92
- Mittelwertsatz
 - für Funktionen mehrerer Veränderlicher, 226
 - von Lagrange, 117
- Mittelwertsatz für Riemann-Integrale, 262
- monoton fallende Folge, 55
- monoton wachsende Folge, 55
- monotone Folge, 55
- Multiindex, 234
 - Ordnung eines, 234
- n -te partielle Ableitung, 231
- Negation einer Aussage, 2
- Norm, 192

- Normale an eine Fläche, 252
- normierter Vektorraum, 192
- notwendig, 3
- n -te Ableitung, 114
- Nullfolge, 52
- Nullstelle einer Funktion, 92
- oberes Riemann-Integral, 141
- Obersumme von A bzgl. \mathcal{P} , 257
- Obersumme von f bzgl. \mathcal{P} , 260
- offen, 197
- parametrisierte Kurve, 169
- Parametrisierung einer Kurve, 169
- Partialbruchzerlegung
 - komplexer rationaler Funktionen, 154
 - reeller rationaler Funktionen, 156
- Partialsomme, 62
- Partialsommen, Folge der, 62
- partielle Ableitung, 217
 - der Ordnung n , 231
 - höherer Ordnung, 231
- partielle Ableitungen
 - Kettenregel, 225
 - Vertauschbarkeit, 232
- partielle Integration, 151
- pi
 - Berechnung, 126
- π , 99
- positiv-orientierte Kurve, 176
- Positivität, 189
- Potenz, 33
 - mit rationalem Exponent, 37
- Potenz, komplexe, 77
 - der Eulerzahl, 75
- Potenzmenge, 8
- Potenzreihe, 72
- Prinzip der Intervallschachtelung, 29
- Prinzip von Cavalieri, 266
- Produkt von Mengen, 10
- Produktmetrik, 191
- Produktregel, 109
- Projektion, 12
- punktweise konvergente
 - Folge von Funktionen, 161
- Quantoren, 4
- Quotientenkriterium, 67
- Quotientenregel, 109
- Rand einer Menge, 198
- Randpunkt einer Menge, 198
- Rechtsinverse einer Abbildung, 14
- rechtsseitiger Grenzwert, 89
- reell-analytisch, 133
- reell-wertige Funktion, 79
- reelle Funktion, 79
- reelle Zahlen, 21
 - Anordnungsaxiom, 24
 - Körperaxiom, 23
 - Vollständigkeitsaxiom, 26
- reguläre Fläche, 250
- Reihe, 62
 - absolut-konvergente, 63
 - divergente, 62
 - geometrische, 65
 - harmonische, 64
 - konvergente, 62
 - Partialsomme, 62
 - Wert einer Reihe, 62
- rektifizierbare Kurve, 172
- Restglied, 128
- Richtungsableitung, 216
- Riemann-Integral für Funktionen mehrerer Variabler, 260
- Riemann-integrierbar
 - uneigentlich, 165
- Riemann-Integral über eine Jordan-meßbare Menge, 262
- Riemann-integrierbare Funktion mehrerer Variablen, 260
- Riemannsches Integrierbarkeitskriterium, 142
- Ringschluß, 7
- Rotationskörper, 269
- Satz über
 - die Stetigkeit der inversen Abbildung, 207
 - die Umkehrabbildung, 242
 - implizite Funktionen, 247
- Satz von
 - Bolzano/Weierstraß, 55
 - Borel, 134
 - Rolle, 116
 - Schwarz, 232
- Satz von Fubini, 263
- Schnitte Jordan-meßbarer Mengen, 265
- Schnittzahl von $(A \mid B)$, 26
- Schranke
 - obere, 27
 - untere, 27
- Schraubenlinie, 170, 175
- Sektorformel, Leibnitzsche, 178
- Sinus, 97

- Sinus hyperbolicus, 103
- Spirale, archimedische, 180
- Sprungstelle, 90
- Standardmetrik, 190
- Sternkurve, 179
- stetig differenzierbar, 114
 - n -mal, 114
- stetige Abbildung, 202
- stetige Abbildung in einem Punkt, 202
- stetige Funktion, 80
- stetige Funktion in x_0 , 80
- Stetigkeit
 - Folgenkriterium, 81
- striktes lokales Maximum, 237
- striktes lokales Minimum, 236
- Substitutionsregel, 152
- Summenregel, 109
- Supremum, 27
- surjektive Abbildung, 13
- Symmetrie, 189

- Tangens, 101
- Tangentialebene an eine Fläche, 252
- Tangentialraum an eine Fläche, 250
- Tautologie, 4
- Taylorformel
 - mehrdimensionale, 235
- Taylorpolynom, 128
- Taylorreihe, 133
- Teilfolge, 53
- Teilintervall, 140
- Teilmenge, 8
- Transformationsformel für Riemann-Integrale, 270
- trigonometrische Funktionen, 97

- überabzählbar, 31
- Umkehrabbildung, 13
- Umkehrfunktion
 - Ableitung der, 241
- Umkehrsatz, 242

- uneigentlich konvergente Folge, 54
- uneigentlich Riemann-integrierbar, 165
- uneigentlicher Grenzwert, 54
- Unstetigkeitsstelle
 - 2. Art, 91
 - hebbare, 90
 - Sprungstelle, 90
- unteres Riemann Integral, 141
- Untersumme von A bzgl. \mathcal{P} , 257
- Untersumme von f bzgl. \mathcal{P} , 260
- Urbild eines Punktes, 11

- Verallgemeinerter Mittelwertsatz der Differentialrechnung von Cauchy, 116
- Vereinigung von Mengen, 8
- Verfeinerung der Zerlegung eines Intervalls, 140
- Verknüpfung von Abbildungen, 12
- vollständige Induktion, 16
- vollständiger angeordneter Körper, 26
- vollständiger metrischer Raum, 196
- Vollständigkeitsaxiom der reellen Zahlen, 26

- Wahrheitstafel, 2
- Wahrheitswert, 2
- Weg, 206
- wegzusammenhängend, 206
- Wendepunkt, 120
- Wert einer Reihe, 62
- Wertebereich, 11
- Wurzel, n -te, 36
- Wurzelkriterium, 66

- Zentrum einer Potenzreihe, 72
- Zerlegung eines Intervalls, 140
- Zerlegung eines Quaders, 257
- Zerlegungssatz für komplexe Polynome, 96
- zusammenhängende Teilmenge, 201
- zweimal differenzierbar, 114
- zweite Ableitung, 114
- Zylinderkoordinaten, 240