



Prof. Dr. Helga Baum

Grundkurs Analysis

Skript zur Vorlesung Analysis I-III
Bachelor-Studiengang Mathematik
für die Studienanfänger des WS 2011/12

5. Oktober 2012

Vorwort

Dieses Skript dient als begleitendes Lehrmaterial für den 3-semesterigen Grundkurs Analysis im Pflichtbereich des Bachelorstudienganges "Mathematik" für die Studienanfänger des Wintersemesters 2011/12. Es soll Ihnen nicht den Gang in die Bibliothek ersetzen. Gehen Sie bei Gelegenheit dort hin und blättern Sie in den vielen dort stehenden Lehrbüchern zum Grundkurs Analysis. Sie werden dann selbst feststellen, welches dieser Bücher Ihnen am besten gefällt und Ihnen am meisten hilft. Dieses Skript enthält den Stoff des Grundkurses Analysis, so wie ich ihn in den nächsten drei Semestern lesen werde und soll Ihnen die Nacharbeit der Vorlesung erleichtern. Es soll Sie auf keinen Fall davon abhalten, während der Vorlesung mitzuschreiben. Erfahrungsgemäß ist das Mitschreiben einer Vorlesung (auch dann, wenn man ab einer bestimmten Stelle nicht mehr alles oder nichts mehr versteht) etwas, das vielen von Ihnen am Anfang schwer fällt. Es ist aber eine Fähigkeit, die Sie für Ihr Studium benötigen und lernen müssen.

Mathematik lernt man nur, wenn man sich selbst mit ihr beschäftigt. Es reicht also nicht, sich in die Vorlesung zu setzen. Die wenigsten von Ihnen werden nach den 90 Minuten rausgehen und alles verstanden haben. Das ist völlig normal. Arbeiten Sie die Vorlesung zu Hause an Hand Ihrer Mitschriften nach und versuchen Sie, die nicht verstandenen Stellen zu klären. Wenn Ihnen das allein nicht gelingt, nutzen Sie die Übung, eines der Tutorien und die Sprechstunden dazu. Auch die Diskussion mit Ihren Kommilitonen über den Vorlesungsstoff kann hilfreich sein. Den sicheren Umgang mit dem gelernten Stoff erwerben Sie nur durch das Lösen der Übungsaufgaben, die Sie jede Woche bekommen. Die Aufgaben sind nicht nur ein Selbsttest oder ein lästiges Übel, um den Übungsschein zu bekommen – sie sind das entscheidende Mittel, mit dem Sie zunehmend Routine im Umgang mit Mathematik bekommen.

Möchte man verstehen, 'was in der Welt vorgeht' und dies genauer analysieren, so stellt man schnell fest, daß man dazu die funktionalen Abhängigkeiten von Ursachen und Wirkungen geeignet modellieren muß. Die Analysis beschäftigt sich mit der Frage, wie man das Änderungsverhalten von Funktionen verstehen, beschreiben und beherrschen kann. Sie stellt Begriffe bereit, mit denen man die Änderung einer Funktion 'im Kleinen' (also bei geringen Änderungen ihrer unabhängigen Variablen) erfassen kann und untersucht, wann

und auf welche Weise man aus diesen Eigenschaften 'im Kleinen' globale Eigenschaften der Funktion bestimmen kann. Das wichtigste und unverzichtbare Hilfsmittel für solche Untersuchungen ist der Begriff des Grenzwertes. Man muß exakt formulieren können, was es in dem jeweils benutzten Modell bedeutet, daß man sich an einen Punkt 'annähert'.

Ich beginne deshalb den Grundkurs Analysis mit dem Studium einer Klasse von Räumen, in denen man einen solchen Grenzwertbegriff formulieren kann, mit den metrischen Räumen. Anschließend werden verschiedene Klassen von Funktionen zwischen allgemeinen und speziellen metrischen Räumen behandelt, insbesondere die stetigen, die differenzierbaren und die integrierbaren Funktionen. Als Anwendung der grundlegenden Eigenschaften verschiedener Funktionenklassen werden wir die Lösungstheorie gewöhnlicher Differentialgleichungen, die Maßtheorie und die Analysis auf Untermannigfaltigkeiten behandeln.

Im Einzelnen werden wir in den drei Semestern der Vorlesung folgende Schwerpunkte behandeln:

1. Konvergenz und Stetigkeit
 - Reelle und komplexe Zahlen
 - Metrische Räume und ihre topologischen Eigenschaften
 - Folgen und Reihen, Potenzreihen, elementare Funktionen
 - Stetige Abbildungen zwischen metrischen Räumen
2. Differential- und Integralrechnung einer und mehrerer reeller Variablen
 - Differentialrechnung von Funktionen einer reellen Variablen
 - Integralrechnung für Funktionen einer reellen Variablen (Riemann-Integral)
 - Differentialrechnung für Funktionen mehrerer reeller Variablen
 - Integralrechnung für Funktionen mehrerer reeller Variablen (Lebesgue-Integral als Teil der Maßtheorie)
3. Gewöhnlichen Differentialgleichungen
 - Elementare Lösungsverfahren
 - Allgemeine Aussagen über Differentialgleichungsprobleme (Existenz, Eindeutigkeit, Fortsetzbarkeit, Stabilität von Lösungen)
 - Lineare Differentialgleichungssysteme im \mathbb{R}^n
4. Maß- und Integrationstheorie
 - σ -Algebren und Maße (insbesondere Lebesgue-Maß im \mathbb{R}^n)
 - Der Integralbegriff über allgemeinen Maßräumen
 - Grenzwertsätze
 - Satz von Fubini
 - Transformationsformel für Lebesgue-Integrale
5. Differential- und Integralrechnung für Untermannigfaltigkeiten des \mathbb{R}^n
 - Untermannigfaltigkeiten des \mathbb{R}^n
 - Vektorfelder und Differentialformen
 - Integration über Untermannigfaltigkeiten
 - Der Satz von Stokes und die klassischen Integralsätze

Einige Lehrbücher zum Analysis-Grundkurs

- H. Amann, J. Escher: *Analysis I, II und III*, Birkhäuser-Verlag
- M. Barner, F. Flohr: *Analysis I und II*, de Gruyter-Verlag
- Th. Bröcker: *Analysis I, II und III*, Wissenschaftsverlag Mannheim
- J. Dieudonne: *Grundzüge der modernen Analysis*, Deutscher Verlag der Wissenschaften
- K. Endl, W. Luh: *Analysis I, II und III*, Aula-Verlag Wiesbaden
- O. Forster: *Analysis 1, 2 und 3*, Vieweg-Verlag
- H. Heuser: *Lehrbuch der Analysis, Teil 1 und 2*, Teubner-Verlag Stuttgart
- St. Hildebrandt: *Analysis 1 und 2*, Springer-Verlag
- K. Königsberger: *Analysis 1 und 2*, Springer-Verlag
- W. Rudin: *Analysis*, Oldenburg Verlag 2009
- W. Walter: *Analysis 1 und 2*, Springer-Verlag

und viele andere mehr . . . Sie finden in diesen Büchern viele interessante Beispiele und Anwendungen, historische Kommentare und weiterführende Kapitel, die wir aus Zeitgründen in der Vorlesung nicht behandeln können. Es lohnt sich deshalb, in diese Bücher hineinzuschauen.

Häufig benutzte Bezeichnungen und Abkürzungen

\forall	für alle
\exists	es existiert
$\exists!$	es existiert genau ein
\Leftrightarrow	genau dann, wenn
\Rightarrow	daraus folgt
$:=$	ist definiert als
$:\Leftrightarrow$	ist definiert durch
$x \in M$	x ist Element der Menge M
\emptyset	leere Menge (Menge, die kein Element enthält)
$A \subset M$	A ist Teilmenge von M , d.h. $x \in A \Rightarrow x \in M$
OBdA	Ohne Beschränkung der Allgemeinheit

A und B seien Mengen.

$A \cup B := \{x \mid x \in A \text{ oder } x \in B\}$ heißt die *Vereinigung* von A und B .

$A \cap B := \{x \mid x \in A \text{ und } x \in B\}$ heißt der *Durchschnitt* von A und B .

Falls $A \cap B = \emptyset$, heißen A und B *disjunkt*;

$A \dot{\cup} B$ bezeichnet die Vereinigung zweier *disjunkter* Mengen A und B . Man sagt dazu dann *disjunkte Vereinigung* von A und B .

$A \setminus B = \{x \mid x \in A \text{ und } x \notin B\}$ heißt *Differenzmenge*.

$A \times B = \{(x, y) \mid x \in A \text{ und } y \in B\}$ heißt das *Produkt* von A und B .

Für eine endliche Menge A bezeichnet $\sharp A$ die Anzahl ihrer Elemente.

Inhaltsverzeichnis

1	Reelle und komplexe Zahlen	1
1.1	Das Beweisprinzip der vollständigen Induktion	2
1.2	Die reellen Zahlen	5
1.2.1	Die Körpereigenschaften von \mathbb{R}	6
1.2.2	Die Anordnungseigenschaften von \mathbb{R}	7
1.2.3	Vollständigkeitseigenschaft der reellen Zahlen	9
1.2.4	Die Überabzählbarkeit der Menge der reellen Zahlen	14
1.2.5	Wurzeln und Potenzen reeller Zahlen	17
1.3	Die komplexen Zahlen	22
1.4	Die Vektorräume \mathbb{R}^n und \mathbb{C}^n	29
2	Metrische Räume	33
2.1	Definition und Beispiele metrischer Räume	33
2.2	Das Innere, der Abschluß und der Rand einer Menge	35
2.3	Offene und abgeschlossene Mengen in metrischen Räumen	41
2.4	Folgen in metrischen Räumen	46
2.4.1	Allgemeine Eigenschaften konvergenter Folgen	46
2.4.2	Spezielle Eigenschaften von konvergenten Folgen im Vektorraum \mathbb{C}^k bzw. \mathbb{R}^k	49
2.4.3	Spezielle Eigenschaften konvergenter Folgen in \mathbb{R}	53
2.5	Vollständige metrische Räume	59
2.6	Kompakte und folgenkompakte Teilmengen metrischer Räume	65
2.7	Zusammenhängende Teilmengen und Zusammenhangskomponenten eines metrischen Raumes	70
2.8	Banachräume und Hilberträume	73
3	Reihen in Banachräumen	79
3.1	Konvergente und divergente Reihen	79
3.2	Konvergenzkriterien für Reihen in Banachräumen	80
3.3	Das Cauchy-Produkt von Reihen komplexer Zahlen	89
3.4	Umordnung von Reihen	91

3.5	Komplexe Potenzreihen	94
3.6	Anwendung: Exponentialfunktion, Logarithmusfunktion und komplexe Potenzen	96
4	Stetige Abbildungen zwischen metrischen Räumen	101
4.1	Der Grenzwert einer Abbildung in einem Punkt	101
4.2	Stetige Abbildungen (Definition und Beispiele)	107
4.3	Eigenschaften stetiger Abbildungen	115
4.4	Folgen stetiger Abbildungen	122
4.5	Funktionenreihen	124
4.6	Die trigonometrischen und die Hyperbelfunktionen im Komplexen	127
4.6.1	Die trigonometrischen Funktionen	128
4.6.2	Die Hyperbelfunktionen	133
4.7	Der Fundamentalsatz der Algebra	135
4.8	Approximationssätze für stetige Abbildungen	137
5	Differentialrechnung für Funktionen einer reellen Variablen	149
5.1	Differenzierbare Abbildungen	149
5.2	Die Mittelwertsätze der Differentialrechnung und Anwendungen	157
5.3	Differentiation von Funktionenfolgen und -reihen	167
5.4	Potenzreihen mit reellem Zentrum	169
5.5	Reell-analytische Funktionen und Taylorreihen	174
5.6	Lokale Extrema für Funktionen einer reellen Variablen	181
6	Differentialrechnung für Funktionen mehrerer reeller Variablen	183
6.1	Ableitung, Richtungsableitung und Gradient einer Funktion	183
6.2	Die partiellen Ableitungen	194
6.3	Die Taylorformel für Funktionen mehrerer reeller Variablen	201
6.4	Lokale Extrema für Funktionen mehrerer reeller Variablen	203
6.5	Der Satz über den lokalen Diffeomorphismus	206
6.6	Der Satz über implizite Funktionen	213
6.7	Untermannigfaltigkeiten des \mathbb{R}^N und ihre Tangentialräume	217
6.8	Extrema unter Nebenbedingungen	221
7	Integralrechnung für Funktionen einer reellen Variablen	227
7.1	Stammfunktionen und ihre Berechnung	227
7.2	Das Riemann-Integral	239
7.3	Der Hauptsatz der Differential- und Integralrechnung	253
7.4	Die Mittelwertsätze der Integralrechnung	255
7.5	Parameterabhängige Integrale	257
7.6	Uneigentliche Riemann-Integrale	260
7.7	Die Länge von Kurven und der Flächeninhalt ebener Gebiete	264

Reelle und komplexe Zahlen

Wir gehen davon aus, daß der Aufbau der Zahlbereiche bis zu den reellen Zahlen bekannt ist. Man findet dies zum Beispiel in dem Buch von J. Kramer: *Zahlen für Einsteiger* (Vieweg-Verlag, 2006).

Wir benutzen in dieser Vorlesung die folgenden Bezeichnungen für die Zahlbereiche:

\mathbb{N}	Menge aller natürlichen Zahlen: $1, 2, 3, \dots$
\mathbb{N}_0	$\mathbb{N} \cup \{0\}$
\mathbb{Z}	Menge aller ganzen Zahlen: $0, \pm 1, \pm 2, \dots$
\mathbb{Q}	Menge aller rationalen Zahlen: $\{\frac{m}{n} \mid m \in \mathbb{Z}, n \in \mathbb{N}\}$
\mathbb{Q}^+	Menge der positiven rationalen Zahlen: $\{q \in \mathbb{Q} \mid q > 0\}$
\mathbb{R}	Menge der reellen Zahlen
$\mathbb{R} \setminus \mathbb{Q}$	Menge der irrationalen Zahlen
\mathbb{C}	Menge der komplexen Zahlen

Offensichtlich gilt: $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

Die Zahlbereiche werden bekanntlich aus folgendem Grund erweitert:

$\mathbb{N} \subset \mathbb{Z}$: Die Subtraktion ist durch die Erweiterung immer ausführbar. Seien $a, b \in \mathbb{N}$. Die Gleichung $x + a = b$ ist in \mathbb{Z} immer lösbar, aber nicht in \mathbb{N} .

$\mathbb{Z} \subset \mathbb{Q}$: Die Division ist durch die Erweiterung immer ausführbar. Seien $a, b \in \mathbb{Z}$, $a \neq 0$. Die Gleichung $x \cdot a = b$ ist in \mathbb{Q} immer lösbar, aber nicht in \mathbb{Z} .

$\mathbb{Q} \subset \mathbb{R}$: Die Erweiterung ist nötig, damit Wurzeln positiver Zahlen existieren. Sei $q \in \mathbb{Q}^+$, $n \in \mathbb{N}$. Die Gleichung $x^n = q$ ist in \mathbb{R} immer lösbar, aber nicht in \mathbb{Q} .

Im Abschnitt 1.1. werden wir zunächst ein wichtiges und oft benutztes Beweisprinzip wiederholen und an Beispielen demonstrieren: das der vollständigen Induktion. Im Abschnitt 1.2. fassen wir die grundlegenden, die Menge der reellen Zahlen charakterisierenden Eigenschaften (ihre *Axiome*) zusammen und leiten wesentliche, aus den Axiomen folgende Eigenschaften her. Der Abschnitt 1.3. enthält die Definition und die grundlegenden Eigenschaften der komplexen Zahlen. In Abschnitt 1.4. betrachten wir die Vektorräume \mathbb{R}^n und \mathbb{C}^n und definieren den Abstand von Vektoren in diesen Vektorräumen.

1.1 Das Beweisprinzip der vollständigen Induktion

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

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

- (1) $n_0 \in M$,
- (2) Ist $n \in M$ für eine Zahl $n \geq n_0$, so ist auch $(n + 1) \in M$.

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.

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) Falls $A(n)$ richtig ist für eine Zahl $n \geq n_0$, 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).
- Ist $n \in M$ für eine Zahl $n \geq n_0$, so ist $(n + 1) \in M$ (nach Induktionsschritt).

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

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

Satz 1.1 *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 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 1.1. 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 1.2 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 1.2. 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 1.3 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 1.4 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 1.4 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 1.3

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

□

1.2 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 eines der Bücher

- J. Kramer: *Zahlen für Einsteiger*, Vieweg-Verlag, 2006
- H.-D. Ebbinghaus: *Zahlen, Grundwissen Mathematik*, Springer, 2. Aufl. 1988
- A. Oberschelp: *Aufbau des Zahlensystems*, Vandenhoeck & Ruprecht, Göttingen 1976.

1.2.1 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:

- K1: $x + y = y + x \quad \forall x, y \in \mathbb{R}$ *(Kommutativgesetz der Addition)*
 K2: $(x + y) + z = x + (y + z) \quad \forall x, y, z \in \mathbb{R}$ *(Assoziativgesetz der Addition)*
 K3: $0 + x = x \quad \forall x \in \mathbb{R}$ *(Existenz eines neutralen Elementes)*
 K4: Zu jedem $x \in \mathbb{R}$ gibt es ein $y \in \mathbb{R}$ mit $x + y = 0$. y heißt das Negative von x und wird mit $y =: -x$ bezeichnet. *(Existenz des negativen Elements)*

Multiplikation:

- K5: $x \cdot y = y \cdot x \quad \forall x, y \in \mathbb{R}$ *(Kommutativgesetz der Multiplikation)*
 K6: $(x \cdot y) \cdot z = x \cdot (y \cdot z) \quad \forall x, y, z \in \mathbb{R}$ *(Assoziativgesetz der Multiplikation)*
 K7: $1 \cdot x = x \cdot 1 = x \quad \forall x \in \mathbb{R}$ *(Existenz eines neutralen Elementes)*
 K8: Zu jedem $x \in \mathbb{R}$, $x \neq 0$ existiert ein $z \in \mathbb{R}$ mit $x \cdot z = 1$. z heißt das inverse Element zu x und wird mit $z =: \frac{1}{x}$ bezeichnet. *(Existenz des inversen Elementes)*
 K9: $(x + y) \cdot z = x \cdot z + y \cdot z \quad \forall x, y, z \in \mathbb{R}$ *(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 1.3. 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 Körper¹.

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 Algebra-Vorlesung 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:

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

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

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

1.2.2 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.

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 gefordert wird.

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 1.4. 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 1.5. Man sagt " x ist kleiner gleich y " und schreibt $x \leq y$, falls $y - x \in \mathbb{R}^+ \dot{\cup} \{0\}$.

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 $x \leq x$

Reflexivität

O3: Aus $x \leq y$ und $y \leq x$ folgt $x = y$

Antisymmetrie

O4: Aus $x \leq y$ und $y \leq z$ folgt $x \leq z$

Transitivität

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. $x > y : \iff y < x$.

Mittels der Ordnungsrelation können wir Intervalle definieren:

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

$$[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\} \quad (\text{abgeschlossenes Intervall})$$

$$[a, b) := \{x \in \mathbb{R} \mid a \leq x < b\} \quad (\text{halboffenes Intervall})$$

$$(a, b] := \{x \in \mathbb{R} \mid a < x \leq b\} \quad (\text{halboffenes Intervall})$$

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

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

Des Weiteren seien

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

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 1.6. *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 1.5 *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, Anordnungseigenschaften 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.

1.2.3 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 1.7. Ein Dedekindscher Schnitt 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

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

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.

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

Definition 1.8. 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 1.9. 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 isomorph (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.

Definition 1.10.

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 1.11. Sei $A \subset \mathbb{R}$.

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 1.6 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 Schnittzahl 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 $(Y \mid X)$ 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 1.12.

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 1.6 einige Folgerungen ab.

Folgerung 1.1 (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 1.6 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 1.2

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 1.3 *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 1.7 (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 1.3). 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 1.13. *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 1.8 (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 1.6 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.

1.2.4 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 1.14. Seien X und Y zwei nichtleere Mengen. Eine Abbildung $f : X \rightarrow Y$ heißt

- *injektiv*, falls $f(x_1) \neq f(x_2)$ für alle $x_1, x_2 \in X$ mit $x_1 \neq x_2$.
- *surjektiv*, falls für jedes $y \in Y$ ein $x \in X$ mit $f(x) = y$ existiert.
- *bijektiv*, falls f injektiv und surjektiv ist, d.h. falls für jedes $y \in Y$ genau ein $x \in X$ mit $f(x) = y$ existiert.

Definition 1.15. 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

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

Eine Menge A heißt überabzählbar, wenn sie weder leer, noch endlich oder abzählbar ist. Wir sagen, die Menge A ist höchstens abzählbar, wenn sie leer, endlich oder abzählbar ist.

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

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

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

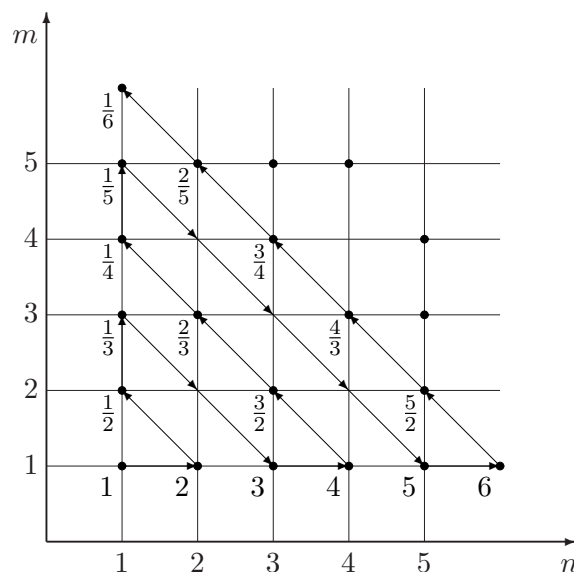
Satz 1.9 Die Menge \mathbb{Q} der rationalen Zahlen ist abzählbar.

Beweis. (1. Cantorsches Diagonalisierungsverfahren).

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 $\varphi : \mathbb{N} \rightarrow \mathbb{Q}^+$.

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 1.10 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

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

$$x_n \notin I_n$$

$$L(I_1) = \frac{1}{3}, L(I_2) = \frac{1}{3^2}, \dots, L(I_n) = \frac{1}{3^n}, \dots$$

Somit ist $I_1 \supset I_2 \supset I_3 \supset \dots$ eine Intervallschachtelung. Sei nun $x \in \bigcap_{n=1}^{\infty} I_n$. 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 1.16. 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.

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 den Vorlesungen über mathematische Logik behandelt.

1.2.5 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 1.11 (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:

$$\begin{aligned} (x+y)^{n+1} &= (x+y)^n \cdot (x+y) \\ &\stackrel{IV}{=} \left(\sum_{k=0}^n \binom{n}{k} x^k \cdot y^{n-k} \right) \cdot (x+y) \\ &= \sum_{k=0}^n \binom{n}{k} x^k \cdot x \cdot y^{n-k} + \sum_{k=0}^n \binom{n}{k} x^k \cdot y^{n-k} \cdot y \\ &= \sum_{k=0}^n \binom{n}{k} x^{k+1} \cdot y^{n+1-(k+1)} + \sum_{k=0}^n \binom{n}{k} x^k \cdot y^{n+1-k} \\ &= \sum_{l=1}^{n+1} \binom{n}{l-1} x^l \cdot y^{(n+1)-l} + \sum_{l=0}^n \binom{n}{l} x^l \cdot y^{(n+1)-l} \\ &= \sum_{l=1}^n \left(\binom{n}{l} + \binom{n}{l-1} \right) x^l \cdot y^{(n+1)-l} + \binom{n}{n} x^{n+1} y^0 + \binom{n}{0} x^0 y^{n+1} \\ &\stackrel{1,3}{=} \sum_{l=0}^{n+1} \binom{n+1}{l} x^l \cdot y^{(n+1)-l}. \end{aligned}$$

□

Aus dem Binomischen Satz 1.11 ergibt sich die

Folgerung 1.4

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 1.12 (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 1.5

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 1.13 (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:

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

□

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

Satz 1.14 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. 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 induktiv die folgende Folge abgeschlossener Intervalle: Wir setzen $I_1 := [1, x]$. Sei $I_k := [a_k, b_k]$ 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:

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 1.2 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

$$J_k := [a_k^n, b_k^n].$$

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 1.2 gibt es zu jedem $\varepsilon > 0$ ein $k \in \mathbb{N}$ mit $L(J_k) \leq \varepsilon$. Die Folge der Intervalle J_k ist also eine Intervallschachtelung. 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 1.14 gilt also nur in \mathbb{R}^+ . Die Gleichung $y^n = x$ ist in \mathbb{Q} im allgemeinen nicht lösbar.

Satz 1.15 *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 $k = \left(\frac{m}{l}\right)^n = \frac{m^n}{l^n}$ 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 1.17. *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. $x^q \cdot x^p = x^{p+q}$, $(x^q)^p = x^{p \cdot q}$, $x^q \cdot y^q = (xy)^q$.
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.

1.3 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.

Dazu betrachten wir die Menge der Paare reeller Zahlen

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

und führen auf dieser Menge 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 nach folgenden Regeln:

$$z_1 + z_2 := (a_1, b_1) + (a_2, b_2) = (a_1 + a_2, b_1 + b_2) \quad (1.1)$$

$$z_1 \cdot z_2 := (a_1, b_1) \cdot (a_2, b_2) = (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1). \quad (1.2)$$

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 (1.1) der reellen Paare auch noch die Multiplikation (1.2) festgelegt hat. Die Elemente von \mathbb{C} heißen *komplexe Zahlen*.

Satz 1.16 *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 $\frac{1}{w} := (\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2})$. \square

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 $z^{-n} := (\frac{1}{z})^n = \frac{1}{z^n}$. 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 1.11) die binomische Formel

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

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

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

$$\boxed{z = a + ib \quad a, b \in \mathbb{R}} \quad (1.3)$$

darstellbar. Dies ist die übliche Darstellung der komplexen Zahlen. Man kann dann mit den komplexen Zahlen wie mit den reellen rechnen, indem man $i^2 = -1$ berücksichtigt. Es gilt also für $z_1 = a_1 + i b_1$ und $z_2 = a_2 + i b_2$

$$z_1 + z_2 = (a_1 + i b_1) + (a_2 + i b_2) = (a_1 + a_2) + i (b_1 + b_2) \quad (1.4)$$

$$z_1 \cdot z_2 = (a_1 + i b_1) \cdot (a_2 + i b_2) = (a_1 a_2 - b_1 b_2) + i (a_1 b_2 + b_1 a_2) \quad (1.5)$$

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

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

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

also gilt

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

Definition 1.18. Ist $z = a + i b \in \mathbb{C}$ eine komplexe Zahl, so heißt $\bar{z} := a - i b$ die *konjugiert komplexe Zahl zu z* .

Es gelten folgende, leicht zu überprüfende Rechenregeln:

Satz 1.17 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. $z + \bar{z} = 2 \cdot Re(z)$, $z - \bar{z} = 2i \cdot Im(z)$.
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 1.19. 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 1.18 (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. $|z| = |\bar{z}|$
5. $|\operatorname{Re}(z)| \leq |z|$, $|\operatorname{Im}(z)| \leq |z|$.

Beweis. 1., 4. und 5. 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 folgt aus

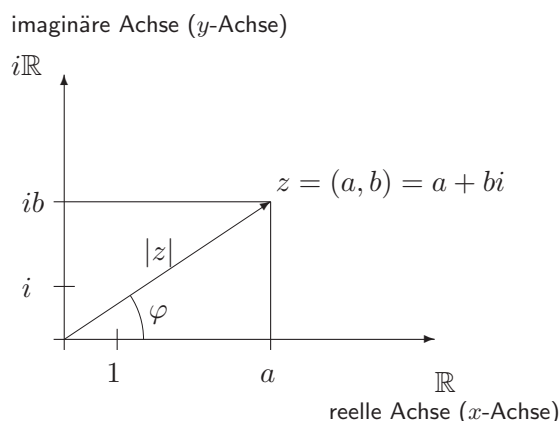
$$\begin{aligned} |z + w|^2 &= (z + w)(\overline{z + w}) \\ &= (z + w)(\bar{z} + \bar{w}) \\ &= z\bar{z} + w\bar{w} + w\bar{z} + z\bar{w} \\ &= z\bar{z} + w\bar{w} + w\bar{z} + \overline{w\bar{z}} \\ &= |z|^2 + |w|^2 + 2 \cdot \operatorname{Re}(w\bar{z}) \\ &\leq |z|^2 + |w|^2 + 2 \cdot |w\bar{z}| \\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

□

Die geometrische Interpretation der komplexen Zahlen

Der Darstellung der reellen Zahlen auf einer Geraden entspricht die Darstellung der komplexen Zahlen in der Ebene, die man dann oft Gaußsche Zahlenebene oder komplexe Zahlenebene nennt.

Wir wählen 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 $|z| = \sqrt{a^2 + b^2}$ 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 Winkel zwischen der x -Achse und dem Strahl vom Ursprung durch z , gemessen in positiver Richtung (entgegen dem Uhrzeigersinn). Dann gilt im rechtwinkligen Dreieck

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

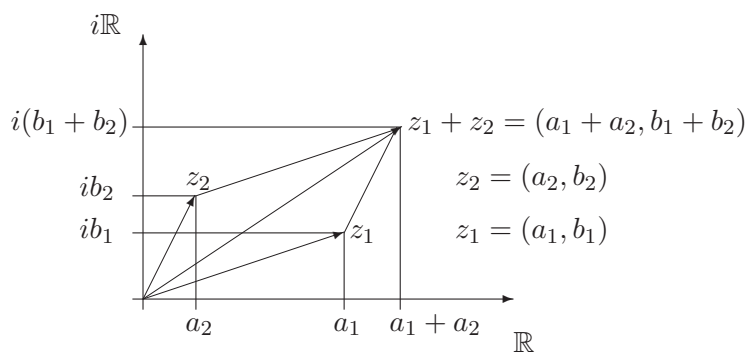
Die Darstellung

$$z = |z| (\cos \varphi + i \cdot \sin \varphi) \quad (1.6)$$

heißt *trigonometrische Darstellung* 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 2π 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 $z_1 \cdot z_2$:

Wir betrachten die trigonometrische Darstellung von z_1 und z_2

$$z_1 = |z_1|(\cos \varphi_1 + i \sin \varphi_1)$$

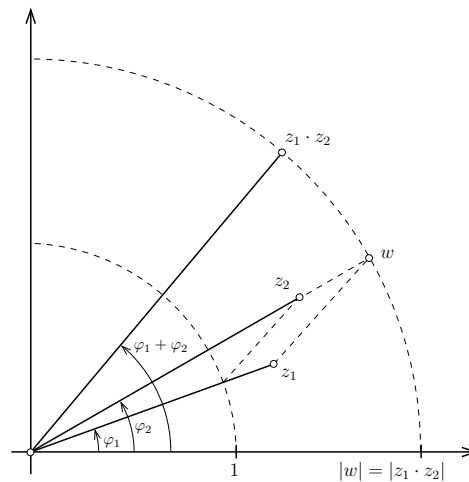
$$z_2 = |z_2|(\cos \varphi_2 + i \sin \varphi_2)$$

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 $|w| = |z_1 \cdot z_2|$. Man dreht diesen Punkt w um den Winkel φ_1 um den Ursprung und erhält den Punkt $z_1 \cdot z_2$.

Beispiele:

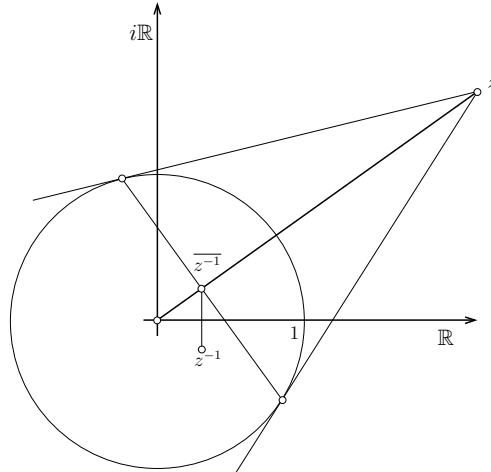
a) Die Abbildung $z \in \mathbb{C} \mapsto iz \in \mathbb{C}$ beschreibt die Drehung um den Ursprung um den Winkel $\frac{\pi}{2}$ (entgegen dem Uhrzeigersinn).

b) Die Abbildung $z \in \mathbb{C} \mapsto rz \in \mathbb{C}$, $r \in \mathbb{R}^+$, ist die Streckung von z um den Faktor r auf dem durch den Ursprung und z gehenden Strahl.

c) Was bedeutet die Inversion $z \in \mathbb{C} \mapsto z^{-1} = \frac{1}{z} \in \mathbb{C}$ geometrisch?

Betrachten wir den Kreis $K_1 = \{z \in \mathbb{C} \mid |z|^2 = 1\}$ vom Radius 1. Sei $z \in \mathbb{C}$ ein vom Ursprung verschiedener Punkt. Der Punkt $\tilde{z} \in \mathbb{C}$ heißt Spiegelpunkt von z an K_1 , falls

\tilde{z} auf dem von 0 ausgehenden Strahl durch z liegt und $|z| \cdot |\tilde{z}| = 1$ gilt. Dann existiert ein $c \in \mathbb{R}^+$ mit $\tilde{z} = cz$. Setzen wir das in $|z| \cdot |\tilde{z}| = 1$ ein, so erhalten wir $c = \frac{1}{|z|^2} = \frac{1}{z \cdot \bar{z}}$. Folglich ist der Spiegelpunkt $\tilde{z} = \frac{1}{z} = \overline{\left(\frac{1}{z}\right)}$. Die Inversionsabbildung $z \in \mathbb{C} \mapsto z^{-1} \in \mathbb{C}$ ist also die Hintereinanderausführung der Spiegelung am Kreis K_1 und der Spiegelung an der reellen Achse.



Wir erklären nun Wurzeln aus komplexen Zahlen: Wie wir gerade gesehen haben, gelten für eine komplexe Zahl z die Formeln

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

Damit erhalten wir

Satz 1.19 Sei $w \in \mathbb{C}$ eine von Null verschiedene komplexe Zahl mit dem Betrag $r := |w|$ und dem Argument $\varphi =: \arg(w) \in [0, 2\pi)$. Dann hat die Gleichung $z^n = w$ genau n verschiedene komplexe Lösungen, nämlich

$$z_k := \sqrt[n]{r} \cdot \left(\cos\left(\frac{\varphi}{n} + \frac{k \cdot 2\pi}{n}\right) + i \sin\left(\frac{\varphi}{n} + \frac{k \cdot 2\pi}{n}\right) \right)$$

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

Beweis. Für die komplexen Zahlen z_k gilt nach Definition

$$|z_k| = \sqrt[n]{r} \quad \text{und} \quad \arg(z_k) = \frac{\varphi}{n} + \frac{2\pi k}{n} =: \varphi_k.$$

Hieraus folgt

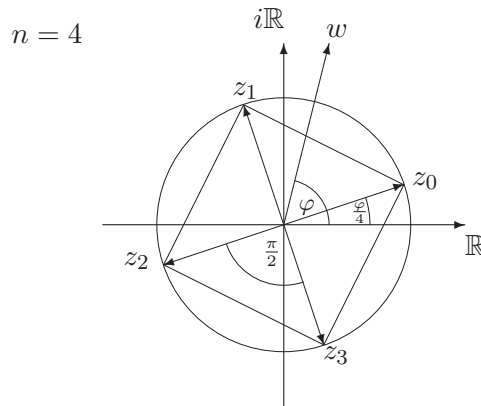
$$z_k^n = \underbrace{\left(\sqrt[n]{r}\right)^n}_{|z_k|^n} \underbrace{\left(\cos(\varphi + 2\pi k) + i \sin(\varphi + 2\pi k)\right)}_{n \cdot \arg(z_k)} = |w|(\cos \varphi + i \sin \varphi) = w.$$

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 $n \cdot \psi = \varphi + 2\pi l$, für ein $l \in \mathbb{Z}$ und somit $\psi = \frac{\varphi}{n} + \frac{2\pi l}{n}$. Wir teilen l durch n mit Rest: $l = k + rn$, $r \in \mathbb{Z}$ und $0 \leq k \leq n - 1$. Dann gilt $\psi = \varphi_k + r \cdot 2\pi$ und folglich $z = z_k$. □

Geometrische Deutung der Wurzeln:

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



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

Satz 1.20 (Fundamentalsatz der Algebra)

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

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

eine Lösung $z \in \mathbb{C}$. □

Der Beweis dieses Satzes für allgemeine n erfolgt später in Kapitel 4.7. 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:

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

Für $c \neq 0$ können wir die Gleichung $w^2 = c$ nach Satz 1.19 lösen und erhalten genau 2 verschiedene komplexe Lösungen von $z^2 + a_1z + a_0 = 0$. Für $c = 0$ gilt $a_0 = \frac{a_1^2}{4}$ und deshalb

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

In diesem Fall ist $z = -\frac{a_1}{2}$ eine 2-fache Lösung von $z^2 + a_1z + a_0 = 0$.

1.4 Die Vektorräume \mathbb{R}^n und \mathbb{C}^n

Definition 1.20. Mit \mathbb{R}^n bezeichnen wir die Menge aller n -Tupel reeller Zahlen. Mit \mathbb{C}^n bezeichnen wir die Menge aller n -Tupel komplexer Zahlen.

Wir addieren zwei n -Tupel $x = (x_1, \dots, x_n)$ und $y = (y_1, \dots, y_n)$ aus \mathbb{R}^n (\mathbb{C}^n) mittels:

$$x + y := (x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n).$$

Wir multiplizieren ein n -Tupel $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ (\mathbb{C}^n) mit einer Zahl $\lambda \in \mathbb{R}$ (\mathbb{C}) mittels:

$$\lambda \cdot x := \lambda \cdot (x_1, \dots, x_n) = (\lambda x_1, \dots, \lambda x_n).$$

Satz 1.21 \mathbb{R}^n ist ein n -dimensionaler Vektorraum über dem Körper der reellen Zahlen \mathbb{R} . \mathbb{C}^n ist ein n -dimensionaler Vektorraum über dem Körper der komplexen Zahlen \mathbb{C} .

Beweis. Dies wird in der Vorlesung über Lineare Algebra definiert und erklärt. \square

Die Elemente von \mathbb{R}^n und \mathbb{C}^n bezeichnet man deshalb auch als Vektoren. Der Vektor $0 := (0, 0, \dots, 0)$ heißt Nullvektor in \mathbb{R}^n bzw. \mathbb{C}^n .

Definition 1.21.

(1) Unter dem (kanonischen) Skalarprodukt zweier Vektoren $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ und $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ versteht man die Zahl

$$\langle x, y \rangle_{\mathbb{R}^n} := \sum_{j=1}^n x_j \cdot y_j \in \mathbb{R}.$$

(2) Unter dem (kanonischen) Skalarprodukt zweier Vektoren $z = (z_1, \dots, z_n) \in \mathbb{C}^n$ und $w = (w_1, \dots, w_n) \in \mathbb{C}^n$ versteht man die Zahl

$$\langle z, w \rangle_{\mathbb{C}^n} := \sum_{j=1}^n z_j \cdot \bar{w}_j \in \mathbb{C}.$$

(3) Die Norm eines Vektors $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ bzw. $z = (z_1, \dots, z_n) \in \mathbb{C}^n$ ist die Zahl

$$\|x\|_{\mathbb{R}^n} := \sqrt{\langle x, x \rangle_{\mathbb{R}^n}} = \sqrt{\sum_{j=1}^n x_j^2} \in \mathbb{R} \quad \text{bzw.} \quad \|z\|_{\mathbb{C}^n} := \sqrt{\langle z, z \rangle_{\mathbb{C}^n}} = \sqrt{\sum_{j=1}^n |z_j|^2} \in \mathbb{R}.$$

Offensichtlich gilt für $x, y \in \mathbb{R}^n \subset \mathbb{C}^n$

$$\langle x, y \rangle_{\mathbb{R}^n} = \langle x, y \rangle_{\mathbb{C}^n} \quad \text{und} \quad \|x\|_{\mathbb{R}^n} = \|x\|_{\mathbb{C}^n}.$$

Im Folgenden lassen wir zur Abkürzung die Indizes bei Skalarprodukt und Norm weg und bezeichnen beide kurz mit $\langle \cdot, \cdot \rangle$ bzw. $\| \cdot \|$.

Satz 1.22 (Eigenschaften des Skalarprodukts und der Norm)

Seien $z, \hat{z}, w \in \mathbb{C}^n$ und $\mu \in \mathbb{C}$. Dann gilt:

1. $\langle z, w \rangle = \overline{\langle w, z \rangle}$.
2. $\langle z + \hat{z}, w \rangle = \langle z, w \rangle + \langle \hat{z}, w \rangle$.
3. $\langle \mu z, w \rangle = \mu \langle z, w \rangle$ und $\langle z, \mu w \rangle = \bar{\mu} \langle z, w \rangle$.
4. $\|z\| \geq 0$, wobei $\|z\| = 0 \iff z = (0, \dots, 0)$.
5. $\|\mu z\| = |\mu| \cdot \|z\|$.
6. *Cauchy-Schwarzsche Ungleichung (CSU):*

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

Die Gleichheit gilt genau dann, wenn z und w linear abhängig sind.

7. *Dreiecksungleichung:*

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

Beweis. Die Aussagen 1.–5. folgen unmittelbar aus der Definition.

Wir beweisen nun die Cauchy-Schwarzsche Ungleichung:

Seien $z, w \in \mathbb{C}^n$ und $\lambda \in \mathbb{C}$. Da die Norm nicht-negativ ist, gilt

$$\begin{aligned} 0 &\leq \|z + \lambda w\|^2 && (*) \\ &= \langle z + \lambda w, z + \lambda w \rangle \\ &= \langle z, z \rangle + \lambda \langle w, z \rangle + \bar{\lambda} \langle z, w \rangle + \lambda \cdot \bar{\lambda} \langle w, w \rangle \\ &= \|z\|^2 + \lambda \overline{\langle z, w \rangle} + \bar{\lambda} \langle z, w \rangle + |\lambda|^2 \|w\|^2. \end{aligned}$$

Für $w = 0$ gilt die Cauchy-Schwarzsche Ungleichung offensichtlich. Es genügt also, den Fall $w \neq 0$ zu betrachten. Wir setzen jetzt in der obigen Ungleichung $\lambda := -\frac{\langle z, w \rangle}{\|w\|^2}$ und erhalten

$$0 \leq \|z\|^2 - \frac{\langle z, w \rangle \overline{\langle z, w \rangle}}{\|w\|^2} - \frac{\overline{\langle z, w \rangle} \langle z, w \rangle}{\|w\|^2} + \frac{|\langle z, w \rangle|^2}{\|w\|^2}.$$

Daraus folgt

$$0 \leq \|z\|^2 \|w\|^2 - |\langle z, w \rangle|^2$$

und somit

$$|\langle z, w \rangle| \leq \|z\| \|w\|.$$

Die Gleichheit kann dabei nur gelten, wenn in (*) die Gleichheit steht. Dies ist aber genau dann der Fall, wenn $z + \lambda w = 0$, d.h. wenn z und w linear abhängig sind.

Als letztes beweisen wir die Dreiecksungleichung: Es gilt

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

□

Definition 1.22. Seien x und y zwei Vektoren aus \mathbb{R}^n bzw. \mathbb{C}^n . Die Zahl

$$d(x, y) := \|x - y\|$$

heißt *Euklidischer Abstand* zwischen x und y .

Satz 1.23 Für alle Vektoren $x, y, u \in \mathbb{R}^n$ (\mathbb{C}^n) gilt

1. $d(x, y) \geq 0$, wobei $d(x, y) = 0 \iff x = y$.
2. $d(x, y) = d(y, x)$.
3. $d(x, y) \leq d(x, u) + d(u, y)$.

Beweis. Die Behauptungen folgen sofort aus Satz 1.22. □

Im folgenden Kapitel betrachten wir Eigenschaften von \mathbb{R}^n bzw. \mathbb{C}^n , die sich aus der Existenz des Euklidischen Abstandes ergeben. Dabei ist oft nicht wichtig, dass es sich bei der zugrundeliegenden Menge um \mathbb{R}^n oder \mathbb{C}^n handelt und dass der Abstand d der konkrete Euklidische Abstand ist. Man benötigt oft nur seine drei in Satz 1.23 formulierten Eigenschaften. Wir können deshalb die gleichen Untersuchungen für beliebige Mengen X machen, auf denen eine Abstandsfunktion, d.h. eine Funktion $d : X \times X \rightarrow \mathbb{R}$ mit den drei in Satz 1.23 formulierten Eigenschaften gegeben ist.

Eine solche Abstandsfunktion d gibt uns ein Maß dafür, wie weit zwei Punkte in X voneinander entfernt sind. Ein solches Abstandsmaß möchte man z.B. auf der Erdoberfläche haben, um anzugeben, wie weit zwei Orte voneinander entfernt sind. Man benötigt solche Abstandsmaße auch für Funktionenräume, z.B. in der Lösungstheorie von Differentialgleichungen, die viele Prozesse der Natur mathematisch beschreiben. Es gibt sehr viele weitere Gründe, sich bei den betrachteten Mengen X nicht auf die Vektorräume \mathbb{R}^n und \mathbb{C}^n zu beschränken.

Metrische Räume

In diesem Kapitel betrachten wir Mengen X , auf denen ein Abstand $d : X \times X \rightarrow \mathbb{R}$ mit den Eigenschaften aus Satz 1.23 gegeben ist. Mit Hilfe eines solchen Abstandes können wir erklären, was es bedeutet, dass eine Folge von Punkten $x_n \in X$, $n = 1, 2, \dots$, gegen einen Punkt $x \in X$ konvergiert. Der Konvergenzbegriff ist grundlegend für das Studium des lokalen Änderungsverhaltens von Funktionen, die auf der Menge X erklärt sind.

2.1 Definition und Beispiele metrischer Räume

Definition 2.1. Ein metrischer Raum ist ein Paar (X, d) bestehend aus einer nichtleeren Menge X und einer 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).

Die Elemente von X heißen *Punkte* des metrischen Raumes, $d(p, q)$ nennt man *Abstand* zwischen p und q . Die Abbildung d heißt *Metrik* (oder *Abstandsfunktion*) auf X .

Beispiel 1: Aus Kapitel 1 wissen wir, dass folgende Räume metrische Räume sind:

- \mathbb{R} mit $d(x, y) := |x - y|$ für $x, y \in \mathbb{R}$
- \mathbb{C} mit $d(z, w) := |z - w|$ für $z, w \in \mathbb{C}$
- \mathbb{R}^n mit $d(p, q) := \|p - q\|_{\mathbb{R}^n}$ für $p, q \in \mathbb{R}^n$
- \mathbb{C}^n mit $d(r, s) := \|r - s\|_{\mathbb{C}^n}$ für $r, s \in \mathbb{C}^n$,

wobei $|\cdot|$ den Betrag der reellen bzw. komplexen Zahl und $\|\cdot\|_{\mathbb{R}^n}$ und $\|\cdot\|_{\mathbb{C}^n}$ die Normen auf \mathbb{R}^n bzw. \mathbb{C}^n bezeichnen. Diese Metriken nennen wir die *Standardmetrik* auf \mathbb{R} , \mathbb{C} , \mathbb{R}^n bzw. \mathbb{C}^n . Ist nichts anderes vereinbart, so seien \mathbb{R} , \mathbb{C} , \mathbb{R}^n bzw. \mathbb{C}^n mit dieser Metrik versehen.

Beispiel 2: Auf einer Menge können verschiedene Metriken existieren:

Die folgenden 3 Abbildungen d_1, d_2, d_3 sind z.B. Metriken auf \mathbb{R}^2 :

- Der "Luftlinienabstand" (Standardmetrik):

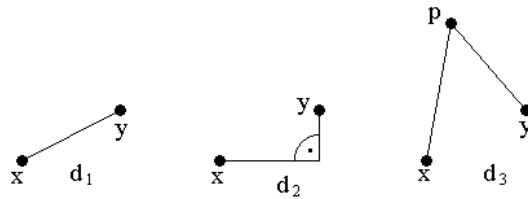
$$d_1(x, y) := \|x - y\| := \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}, \quad \text{wobei } x = (x_1, x_2), y = (y_1, y_2)$$

- Die "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 Metrik.

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: Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge. Dann ist $(A, d|_{A \times A})$ ebenfalls ein metrischer Raum. Hierbei bezeichnet $d|_{A \times A}$ die Einschränkung der Abbildung d auf die Menge $A \times A$. Sie heißt *durch d induzierte Metrik* auf A .

Beispiel 5: Seien $(X_1, d_1), \dots, (X_n, d_n)$ metrische Räume. Wir betrachten die Menge

$$X = X_1 \times X_2 \times \dots \times X_n = \{(p_1, \dots, p_n) \mid p_i \in X_i, i \in \{1, \dots, n\}\}$$

und definieren die *Produktmetrik*

$$d((p_1, \dots, p_n), (q_1, \dots, q_n)) := \sqrt{\sum_{j=1}^n d_j(p_j, q_j)^2}.$$

Dann ist (X, d) ein metrischer Raum und 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-Schwarzischen Ungleichung (CSU, siehe Kapitel 1.4) erhalten wir

$$\begin{aligned}
 d(p, q)^2 &= \sum_{j=1}^n d_j(p_j, q_j)^2 \\
 &\leq \sum_{j=1}^n \left(d_j(p_j, r_j) + d_j(r_j, q_j) \right)^2 \\
 &= \sum_{j=1}^n d_j(p_j, r_j)^2 + 2d_j(p_j, r_j) \cdot d_j(r_j, q_j) + d_j(r_j, q_j)^2 \\
 &= d(p, r)^2 + d(r, q)^2 + 2 \sum_{j=1}^n d_j(p_j, r_j) \cdot d_j(r_j, q_j) \\
 &\stackrel{\text{CSU}}{\leq} d(p, r)^2 + d(r, q)^2 + 2 \sqrt{\sum_{j=1}^n d(p_j, r_j)^2} \cdot \sqrt{\sum_{j=1}^n d(r_j, q_j)^2} \\
 &= d(p, r)^2 + d(r, q)^2 + 2 d(p, r) \cdot d(r, q) \\
 &= (d(p, r) + d(r, q))^2
 \end{aligned}$$

Da der Abstand nicht-negativ ist, können wir in der Ungleichung auf beiden Seiten die Wurzel ziehen und erhalten die Dreiecksungleichung für d . \square

Insbesondere gilt $\mathbb{R}^n = \mathbb{R} \times \dots \times \mathbb{R}$ und $\mathbb{C}^n = \mathbb{C} \times \dots \times \mathbb{C}$ (versehen mit der Standardmetrik bzw. der Produktmetrik).

Definition 2.2. Zwei metrische Räume (X, d) und (Y, \hat{d}) heißen *isometrisch*, wenn eine bijektive Abbildung $f : X \rightarrow Y$ existiert, so dass

$$d(x, y) = \hat{d}(f(x), f(y)) \quad \forall x, y \in X$$

gilt. Die Abbildung f heißt dann *Isometrie* zwischen (X, d) und (Y, \hat{d}) .

2.2 Das Innere, der Abschluß und der Rand einer Menge

Definition 2.3. Sei (X, d) ein metrischer Raum, $x \in X$ und $\varepsilon \in \mathbb{R}^+$.

Die Menge

$$K(x, \varepsilon) := \{y \in X \mid d(x, y) < \varepsilon\}$$

heißt ε -Kugel in X um x .

Beispiel 1: Sei $X = \mathbb{R}^n$ mit der Standardmetrik $d(x, y) = \|x - y\|$ versehen. Dann ist

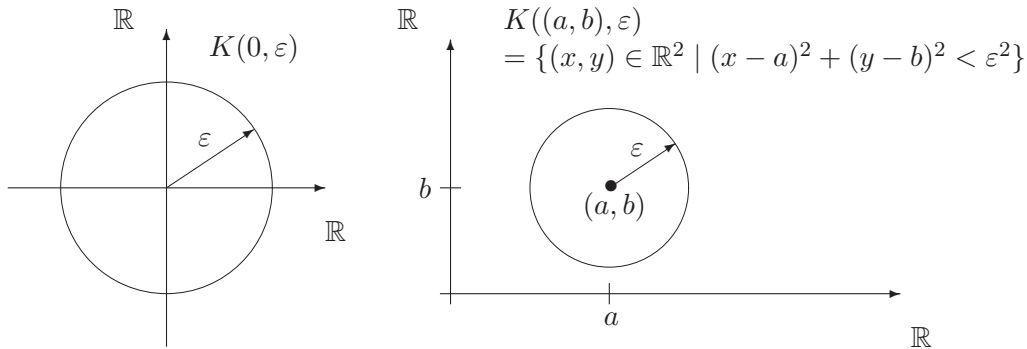
$$K(0, \varepsilon) = \{x \in \mathbb{R}^n \mid \|x\| < \varepsilon\} = \{x \in \mathbb{R}^n \mid x_1^2 + \dots + x_n^2 < \varepsilon^2\}$$

die Kugel vom Radius ε um den Nullpunkt im \mathbb{R}^n ohne ihren Rand.

Ist $n = 1$, so ist jedes offene Intervall eine ε -Kugel.



Für $n = 2$ sind die ε -Kugeln gerade die Kreisscheiben vom Radius ε ohne ihren Rand.

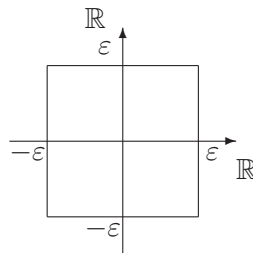


Beispiel 2: Sei $X = \mathbb{R}^2$ mit folgender Metrik versehen:

$$d((x_1, x_2), (y_1, y_2)) = \max(|x_1 - y_1|, |x_2 - y_2|).$$

(X, 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:

$$K(0, \varepsilon) = \{(x_1, x_2) \in \mathbb{R}^2 \mid |x_1| < \varepsilon, |x_2| < \varepsilon\}.$$

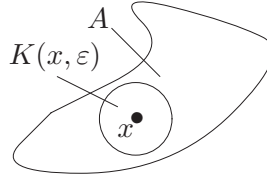


Beispiel 3: Sei (X, d) ein diskreter metrischer Raum. Dann gilt für die ε -Kugeln

$$K(x, \varepsilon) = \begin{cases} \{x\} & \text{falls } \varepsilon \leq 1 \\ X & \text{falls } \varepsilon > 1. \end{cases}$$

Definition 2.4. Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge von X . Ein Punkt $x \in A$ heißt innerer Punkt von A , falls eine ε -Kugel $K(x, \varepsilon)$ um x existiert, die vollständig in A liegt.

$\text{Int}(A)$ bezeichne die Menge aller inneren Punkte von $A \subset X$.



Satz 2.1 (Eigenschaften des Inneren einer Menge)

Sei (X, d) ein metrischer Raum und seien A und B Teilmengen von X . Dann gilt:

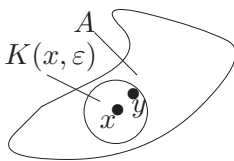
1. $\text{Int}(X) = X, \text{Int}(\emptyset) = \emptyset$.
2. $\text{Int}(A) \subset A$.
3. $\text{Int}(\text{Int}(A)) = \text{Int}(A)$.
4. $A \subset B \implies \text{Int}(A) \subset \text{Int}(B)$ (Monotonie).
5. $\text{Int}(A \cap B) = \text{Int}(A) \cap \text{Int}(B)$.
6. Für beliebig viele Teilmengen $A_i \subset X, i \in \Lambda$, gilt $\bigcup_{i \in \Lambda} \text{Int}(A_i) \subset \text{Int}(\bigcup_{i \in \Lambda} A_i)$.

Beweis. Die Eigenschaften 1. und 2. folgen unmittelbar aus der Definition des Inneren einer Menge.

Zu 3. Aus der 2. Eigenschaft des Inneren folgt $\text{Int}(\text{Int}(A)) \subset \text{Int}(A)$. Es bleibt somit noch

$$\text{Int}(A) \subset \text{Int}(\text{Int}(A))$$

zu zeigen. Sei $x \in \text{Int}(A)$. Nach Definition von $\text{Int}(A)$ existiert eine ε -Kugel um x mit $K(x, \varepsilon) \subset A$. Wir zeigen nun, dass $K(x, \varepsilon) \subset \text{Int}(A)$ gilt.



Sei $y \in K(x, \varepsilon)$ beliebig gewählt. Wir betrachten die Kugel um y mit dem Radius $\varepsilon - d(x, y)$. Dann gilt

$$K(y, \varepsilon - d(x, y)) \subset K(x, \varepsilon) \subset A.$$

Ist nämlich $z \in K(y, \varepsilon - d(x, y))$, so gilt

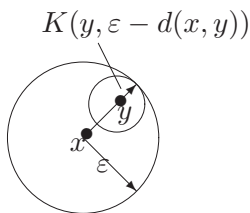
$$d(y, z) < \varepsilon - d(x, y)$$

und wir erhalten

$$d(x, z) \leq d(x, y) + d(y, z) < \varepsilon.$$

Folglich liegt z in $K(x, \varepsilon)$.

Nach Definition ist somit jeder Punkt y von $K(x, \varepsilon)$ ein innerer Punkt von A , also



$K(x, \varepsilon) \subset \text{Int}(A)$. Daraus folgt wiederum, dass x ein innerer Punkt der Menge $\text{Int}(A)$ ist. Damit haben wir $\text{Int}(A) \subset \text{Int}(\text{Int}(A))$ bewiesen.

Zu 4. Wir zeigen $A \subset B \implies \text{Int}(A) \subset \text{Int}(B)$:

Sei $x \in \text{Int}(A)$. Dann existiert eine ε -Kugel $K(x, \varepsilon)$ mit $K(x, \varepsilon) \subset A$. Da $A \subset B$ folgt $K(x, \varepsilon) \subset B$. Folglich ist $x \in \text{Int}(B)$.

Zu 5. Wir zeigen $\text{Int}(A \cap B) = \text{Int}(A) \cap \text{Int}(B)$:

Sei $x \in \text{Int}(A \cap B)$. Dann existiert eine ε -Kugel um x mit $K(x, \varepsilon) \subset A \cap B$ und es folgt, dass sowohl $K(x, \varepsilon) \subset A$ als auch $K(x, \varepsilon) \subset B$. Somit ist $x \in \text{Int}(A)$ und $x \in \text{Int}(B)$ und wir erhalten $x \in \text{Int}(A) \cap \text{Int}(B)$. Es gilt folglich $\text{Int}(A \cap B) \subset \text{Int}(A) \cap \text{Int}(B)$. Sei nun $x \in \text{Int}(A) \cap \text{Int}(B)$, das heißt $x \in \text{Int}(A)$ und $x \in \text{Int}(B)$. Folglich existieren $\varepsilon_1, \varepsilon_2 \in \mathbb{R}^+$, so dass

$$K(x, \varepsilon_1) \subset A \text{ und } K(x, \varepsilon_2) \subset B.$$

Wir betrachten $\varepsilon := \min(\varepsilon_1, \varepsilon_2)$. Dann ist

$$K(x, \varepsilon) \subset K(x, \varepsilon_1) \subset A, \text{ bzw. } K(x, \varepsilon) \subset K(x, \varepsilon_2) \subset B.$$

Also gilt $K(x, \varepsilon) \subset A \cap B$, somit ist $x \in \text{Int}(A \cap B)$. Dies zeigt, dass $\text{Int}(A) \cap \text{Int}(B) \subset \text{Int}(A \cap B)$.

Zu 6. Wir betrachten eine Familie $\{A_i\}_{i \in \Lambda}$ aus beliebig vielen Teilmengen $A_i \subset X$. Sei A_j eine fixierte Menge dieser Familie. Dann ist $A_j \subset \bigcup_{i \in \Lambda} A_i$. Aus der Monotonieeigenschaft des Inneren folgt

$$\text{Int}(A_j) \subset \text{Int}\left(\bigcup_{i \in \Lambda} A_i\right).$$

Da dies für alle $j \in \Lambda$ gilt, erhalten wir

$$\bigcup_{j \in \Lambda} \text{Int}(A_j) \subset \text{Int}\left(\bigcup_{i \in \Lambda} A_i\right).$$

□

Bemerkung: Im allgemeinen ist

$$\text{Int}(A) \cup \text{Int}(B) \neq \text{Int}(A \cup B).$$

Wir betrachten dazu als metrischen Raum $X = \mathbb{R}$ mit der Standardmetrik und die Teilmengen $A = \mathbb{Q}$ und $B = \mathbb{R} \setminus \mathbb{Q}$. Dann ist $A \cup B = \mathbb{R}$ und $\text{Int}(A \cup B) = \mathbb{R}$. Da in jedem Intervall sowohl eine rationale als auch eine irrationale Zahl liegt, gilt $K(x, \varepsilon) \not\subset \mathbb{Q}$, und $K(x, \varepsilon) \not\subset \mathbb{R} \setminus \mathbb{Q}$. Also ist $\text{Int}(\mathbb{Q}) = \text{Int}(\mathbb{R} \setminus \mathbb{Q}) = \emptyset$.

Definition 2.5. Sei (X, d) ein metrischer Raum und $A \subset X$ eine beliebige Teilmenge. Der Abschluss von A ist die Menge

$$\text{cl}(A) := \{x \in X \mid \forall \varepsilon > 0 \text{ gilt } K(x, \varepsilon) \cap A \neq \emptyset\}.$$

Beispiel: Sei $X = \mathbb{R}$ mit der Standardmetrik $d(x, y) = |x - y|$.
Für $A = [a, b)$ erhalten wir $\text{Int}(A) = (a, b)$ und $\text{cl}(A) = [a, b]$.

Satz 2.2 (Eigenschaften des Abschlusses)

Sei (X, d) ein metrischer Raum und A, B Teilmengen von X . Dann gilt:

1. $\text{cl}(A) = X \setminus \text{Int}(X \setminus A)$.
2. $\text{cl}(\emptyset) = \emptyset$ und $\text{cl}(X) = X$.
3. $A \subset \text{cl}(A)$.
4. $A \subset B \Rightarrow \text{cl}(A) \subset \text{cl}(B)$.
5. $\text{cl}(\text{cl}(A)) = \text{cl}(A)$.
6. $\text{cl}(A \cup B) = \text{cl}(A) \cup \text{cl}(B)$.
7. Seien $A_i, i \in I$, beliebig viele Teilmengen von X . Dann gilt

$$\text{cl}\left(\bigcap_{i \in I} A_i\right) \subset \bigcap_{i \in I} \text{cl}(A_i).$$

Beweis. Die erste Behauptung folgt aus folgenden Äquivalenzen:

$$\begin{aligned} x \in X \setminus \text{Int}(X \setminus A) &\iff x \in X \text{ und } x \notin \text{Int}(X \setminus A) \\ &\iff \forall \varepsilon > 0 \text{ gilt } K(x, \varepsilon) \not\subset X \setminus A \\ &\iff \forall \varepsilon > 0 \text{ gilt } K(x, \varepsilon) \cap A \neq \emptyset \\ &\iff x \in \text{cl}(A). \end{aligned}$$

Die anderen Aussagen kann man dann aus den Eigenschaften des Inneren (siehe Satz 2.1) ableiten. Wir überlassen dies dem Leser als Übungsaufgabe. \square

Beispiel: Im allgemeinen gilt nicht $\text{cl}(A \cap B) = \text{cl}(A) \cap \text{cl}(B)$.

Wir betrachten dazu den metrischen Raum $X = \mathbb{R}$ mit der Standardmetrik. Für $A = (0, 1)$ und $B = (1, 2)$ gilt $A \cap B = \emptyset$. Folglich ist $\text{cl}(A \cap B) = \emptyset$. Da $\text{cl}(A) = [0, 1]$ und $\text{cl}(B) = [1, 2]$, erhält man andererseits den nichtleeren Durchschnitt $\text{cl}(A) \cap \text{cl}(B) = \{1\}$.

Sei $A \subset X$ eine Teilmenge eines metrischen Raumes (X, d) . Dann gilt

$$\text{Int}(A) \subset A \subset \text{cl}(A).$$

Wir wollen nun die Punkte studieren, die in $\text{cl}(A) \setminus \text{Int}(A)$ liegen.

Definition 2.6. Sei $A \subset X$ eine Teilmenge eines metrischen Raumes (X, d) . Unter dem Rand von A versteht man die Menge

$$\partial A := \text{cl}(A) \setminus \text{Int}(A).$$

Ein Punkt $x \in \partial A$ heißt Randpunkt von A .

Satz 2.3 (Charakterisierung des Randes)

Sei (X, d) ein metrischer Raum und $A \subset X$. Dann gilt:

1. $\partial A = X \setminus \left(\text{Int}(A) \cup \text{Int}(X \setminus A) \right)$.
2. Ein Punkt $x \in X$ ist genau dann ein Randpunkt von A , wenn für alle $\varepsilon > 0$ gilt

$$K(x, \varepsilon) \cap A \neq \emptyset \quad \text{und} \quad K(x, \varepsilon) \cap (X \setminus A) \neq \emptyset.$$

Beweis. Sei $A \subset X$. Dann gilt

$$\partial A = \text{cl}(A) \setminus \text{Int}(A) = (X \setminus \text{Int}(X \setminus A)) \setminus \text{Int}(A) = X \setminus (\text{Int}(X \setminus A) \cup \text{Int}(A)).$$

Daraus erhalten wir

$$\begin{aligned} x \in \partial A &\iff x \notin \text{Int}(A) \quad \text{und} \quad x \notin \text{Int}(X \setminus A) \\ &\iff \forall \varepsilon > 0 \text{ gilt } K(x, \varepsilon) \cap (X \setminus A) \neq \emptyset \quad \text{und} \quad K(x, \varepsilon) \cap A \neq \emptyset. \end{aligned}$$

□

Bezeichnung: Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge. Dann bezeichnet $\text{Int}(A)$ die *inneren* Punkte von A . Die Punkte in $\text{Int}(X \setminus A)$ nennen wir auch die *äußeren* Punkte von A . Dann sagt uns Satz 2.3 insbesondere, dass sich X in die Menge der inneren Punkte von A , die Menge der äußeren Punkte von A und den Rand von A zerlegt, d.h. es gilt

$$X = \text{Int}(A) \dot{\cup} \text{Int}(X \setminus A) \dot{\cup} \partial A.$$

Beispiel 1: Sei $X = \mathbb{R}^2$ mit der Standardmetrik $d(x, y) = \|x - y\|$.

Für $A = K(x, \varepsilon)$ gilt:

- $\text{Int}(A) = K(x, \varepsilon)$ (siehe Beweis von Satz 2.1, Punkt 3),
- $\text{Int}(\mathbb{R}^2 \setminus A) = \{y \in \mathbb{R}^2 \mid d(x, y) > \varepsilon\}$,
- $\partial A = \{y \in \mathbb{R}^2 \mid d(x, y) = \varepsilon\}$ und
- $\text{cl}(A) = \{y \in \mathbb{R}^2 \mid d(x, y) \leq \varepsilon\}$.

Beispiel 2: Sei X eine nichtleere Menge mit der diskreten Metrik d .

Im metrischen Raum (X, d) gilt für die Kugeln von Radius $1/2$

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

Folglich gilt für eine beliebige Teilmenge $A \subset X$: $A = \text{Int}(A)$. Somit ist auch $X \setminus A = \text{Int}(X \setminus A)$ und wir erhalten für den Rand von A :

$$\partial A = X \setminus \left(\text{Int}(A) \cup \text{Int}(X \setminus A) \right) = X \setminus \left(A \cup (X \setminus A) \right) = X \setminus X = \emptyset.$$

Für den Abschluß von A ergibt sich: $\text{cl}(A) = \text{Int}(A) \cup \partial A = A$.

2.3 Offene und abgeschlossene Mengen in metrischen Räumen

Definition 2.7. Eine Teilmenge A eines metrischen Raumes (X, d) heißt *offen*, falls $\text{Int}(A) = A$ gilt, dh. falls zu jedem $x \in A$ eine ε -Kugel $K(x, \varepsilon)$ existiert, die vollständig in A liegt.

Beispiele:

1. Das Innere $\text{Int}(A)$ jeder Teilmenge A eines metrischen Raumes ist offen. (Nach Satz 2.1, Punkt 3.).
2. Die ε -Kugeln $K(x, \varepsilon)$ eines metrischen Raumes (X, d) sind offen. Ist nämlich $y \in K(x, \varepsilon)$, dann gilt $K(y, \varepsilon - d(x, y)) \subset K(x, \varepsilon)$ (siehe Beweis von 3. in Satz 2.1). Insbesondere sind die Intervalle (a, b) im metrischen Raum \mathbb{R} (mit der Standardmetrik) in diesem Sinne offen.
3. Jede Teilmenge in einem diskreten metrischen Raum (X, d) ist offen.

Satz 2.4 Die offenen Teilmengen eines metrischen Raumes (X, d) haben folgende Eigenschaften:

1. Die Mengen X und \emptyset sind offen.
2. Der Durchschnitt endlich vieler offener Mengen ist offen.
3. Die Vereinigung beliebig vieler offener Mengen ist offen.

Beweis. Zu 1: Siehe Satz 2.1, Punkt 1.

Zu 2: Seien U_1 und U_2 offen. Dann gilt $\text{Int}(U_1) = U_1$ und $\text{Int}(U_2) = U_2$. Aus Satz 2.1 folgt

$$\text{Int}(U_1 \cap U_2) = \text{Int}(U_1) \cap \text{Int}(U_2) = U_1 \cap U_2$$

und somit ist $U_1 \cap U_2$ offen. Induktiv schließt man auf die Behauptung für endlich viele Mengen.

Zu 3: Seien $U_i, i \in \Lambda$, beliebig viele offene Mengen. Dann gilt $\text{Int}(U_i) = U_i$ für alle $i \in \Lambda$. Wegen Satz 2.1 ist

$$\bigcup_{i \in \Lambda} U_i = \bigcup_{i \in \Lambda} \text{Int}(U_i) \subset \text{Int}\left(\bigcup_{i \in \Lambda} U_i\right) \subset \bigcup_{i \in \Lambda} U_i$$

und somit $\bigcup_{i \in \Lambda} U_i = \text{Int}\left(\bigcup_{i \in \Lambda} U_i\right)$. Also ist $\bigcup_{i \in \Lambda} U_i$ offen. □

Beispiel: Der Durchschnitt beliebig vieler offener Mengen ist im allgemeinen nicht offen: Sei $X = \mathbb{R}$ mit der Standardmetrik $d(x, y) = |x - y|$ versehen. Wir betrachten die offenen Intervalle $U_n = (-\frac{1}{n}, \frac{1}{n})$, $n \in \mathbb{N}$. Dann ist $\bigcap_{n \in \mathbb{N}} U_n = \{0\}$ nicht offen in \mathbb{R} .

Als nächstes wollen wir beschreiben, wie die offenen Mengen in Teilräumen aussehen. Ist (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge von X , so ist die Menge A selbst ein metrischer Raum mit der induzierten Metrik $d|_{A \times A} : A \times A \rightarrow \mathbb{R}$.

Satz 2.5 Sei (X, d) ein metrischer Raum und $A \subset X$ eine beliebige Teilmenge. Eine Menge $B \subset A$ ist im metrischen Raum $(A, d|_{A \times A})$ genau dann offen, wenn es in (X, d) eine offene Menge $U \subset X$ gibt, so daß $B = U \cap A$ gilt.

Beweis. Wir vergleichen zunächst die Kugeln des metrischen Raumes $(A, d|_{A \times A})$ mit denen des metrischen Raumes (X, d) . Es gilt

$$K_A(b, \varepsilon) = \{a \in A \mid d(b, a) < \varepsilon\} = K_X(b, \varepsilon) \cap A.$$

(1) Sei $B \subset A$ im metrischen Raum $(A, d|_{A \times A})$ offen. Nach Definition existiert für jedes $b \in B$ eine Zahl $\varepsilon(b) > 0$, so dass

$$K_A(b, \varepsilon(b)) = K_X(b, \varepsilon(b)) \cap A \subset B.$$

Wir betrachten die Menge

$$U := \bigcup_{b \in B} K_X(b, \varepsilon(b)) \subset X.$$

Da U die Vereinigung von Kugeln in X ist, ist U nach Satz 2.4, Punkt 3. offen in X . Da $B \subset A$ und $B \subset U$, ist $B \subset U \cap A$. Andererseits gilt

$$\begin{aligned} U \cap A &= \left(\bigcup_{b \in B} K_X(b, \varepsilon(b)) \right) \cap A \\ &= \bigcup_{b \in B} \left(K_X(b, \varepsilon(b)) \cap A \right) \\ &= \bigcup_{b \in B} K_A(b, \varepsilon(b)) \subset B, \end{aligned}$$

denn jede der Kugeln $K_A(b, \varepsilon(b))$ liegt in B . Folglich gilt $B = U \cap A$.

(2) Sei nun $B = U \cap A$, wobei $U \subset X$ eine offene Menge in (X, d) ist. Es ist zu zeigen, dass B offen im metrischen Raum $(A, d|_{A \times A})$ ist. Sei $b \in B$. Da $b \in U$ und U in (X, d) offen ist, existiert ein $\varepsilon(b) > 0$ mit $K_X(b, \varepsilon(b)) \subset U$. Folglich gilt

$$K_A(b, \varepsilon(b)) = K_X(b, \varepsilon(b)) \cap A \subset U \cap A = B.$$

Somit ist B offen in $(A, d|_{A \times A})$. □

Definition 2.8. Sei (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt abgeschlossen, falls $cl(A) = A$ gilt.

Satz 2.6 Sei (X, d) ein metrischer Raum und $A \subset X$. Dann gilt:

1. $A \subset X$ ist genau dann abgeschlossen, wenn $X \setminus A$ offen ist.
2. X und \emptyset sind abgeschlossen.
3. Der Durchschnitt beliebig vieler abgeschlossener Mengen ist abgeschlossen.
4. Die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen.

Beweis. (1) Wir benutzen Satz 2.2 und erhalten

$$\begin{aligned} A \subset X \text{ ist abgeschlossen} &\iff A = cl(A) \\ &\iff A = X \setminus Int(X \setminus A) \\ &\iff X \setminus A = Int(X \setminus A) \\ &\iff X \setminus A \text{ ist offen} \end{aligned}$$

(2)–(4) folgen aus (1), Satz 2.1 und den Beziehungen zwischen Vereinigung und Durchschnitt bei der Komplementbildung:

(2) $X = X \setminus \emptyset$. Da \emptyset offen ist, ist X abgeschlossen.

$\emptyset = X \setminus X$. Da X offen ist, ist \emptyset abgeschlossen.

(3) Seien $A_i, i \in A$, beliebig viele abgeschlossene Teilmengen von X . Dann gilt nach den bekannten Regeln der Mengenlehre

$$X \setminus \left(\bigcap_{i \in A} A_i \right) = \bigcup_{i \in A} (X \setminus A_i). \quad (*)$$

Da A_i abgeschlossen ist, ist $X \setminus A_i$ offen. Die Vereinigung beliebig vieler offener Mengen ist offen, folglich ist die rechte Seite von (*) ebenfalls offen. Wenden wir (1) auf die linke Seite von (*) an, so folgt, dass $\bigcap_{i \in A} A_i$ abgeschlossen ist.

(4) Es genügt, die Behauptung für zwei abgeschlossene Mengen zu beweisen. Die Gültigkeit für endlich viele Teilmengen erhält man dann durch Induktion. Seien A und B zwei abgeschlossene Teilmengen von X . Nach den bekannten Regeln der Mengenlehre gilt:

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

Da A und B abgeschlossen sind, sind $X \setminus A$ und $X \setminus B$ offen und somit deren Durchschnitt ebenfalls offen. Dann ist nach (1) aber $A \cup B$ abgeschlossen. □

Beispiel: Sei $X = \mathbb{R}^2$ mit der Standardmetrik $d(x, y) = \|x - y\|$.

Die Menge $A = \{x \in \mathbb{R}^2 \mid \|x\| < 1\} \cap \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 > 0\}$ ist offen. Folglich ist $\mathbb{R}^2 \setminus A = \{x \in \mathbb{R}^2 \mid \|x\| \geq 1\} \cup \{(x_1, x_2) \in \mathbb{R}^2 \mid x_2 \leq 0\}$ abgeschlossen.

Als nächstes definieren wir zwei spezielle Arten von Punkten in metrischen Räumen:

Definition 2.9. Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge. Ein Punkt $x \in X$ heißt *Häufungspunkt* von A , falls in jeder ε -Kugel $K(x, \varepsilon)$ ein von x verschiedener Punkt von A liegt, d.h. falls

$$(K(x, \varepsilon) \setminus \{x\}) \cap A \neq \emptyset.$$

Die Menge der Häufungspunkte von A wird mit $HP(A)$ bezeichnet.

Satz 2.7 Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge. Dann gilt:

1. $x \in X$ ist genau dann ein Häufungspunkt von A , wenn in jeder ε -Kugel $K(x, \varepsilon)$ unendlich viele Punkte von A liegen.
2. $cl(A) = A \cup HP(A)$.
3. A ist genau dann abgeschlossen, wenn $HP(A) \subset A$.

Beweis. (1) Sei $x \in X$ ein Häufungspunkt von A und $\varepsilon > 0$ eine beliebig fixierte Zahl. Wir müssen zeigen, dass die Kugel $K(x, \varepsilon)$ unendlich viele Punkte von A enthält. Aus der Definition des Häufungspunktes erhalten wir zunächst, dass ein von x verschiedener Punkt $a_1 \in K(x, \varepsilon)$ existiert mit $a_1 \in A$. Wir betrachten $\varepsilon_1 := d(x, a_1) > 0$. Dann ist $\varepsilon_1 < \varepsilon$. Wiederum aus der Häufungspunkt-Eigenschaft von x , erhalten wir einen weiteren von a_1 und x verschiedenen Punkt $a_2 \in K(x, \varepsilon_1)$ mit $a_2 \in A$. Wir betrachten dann $\varepsilon_2 := d(x, a_2)$ und erhalten einen von a_1, a_2 und x verschiedenen Punkt $a_3 \in K(x, \varepsilon_2)$ mit $a_3 \in A$. Fahren wir induktiv so fort, so erhalten wir unendlich viele voneinander verschiedene Punkte $a_1, a_2, a_3, a_4, \dots \in A \cap K(x, \varepsilon)$.

(2) Wir zeigen zuerst, dass $A \cup HP(A) \subset cl(A)$:

Da $A \subset cl(A)$, ist dazu nur zu zeigen, dass $HP(A) \subset cl(A)$. Sei $x \in HP(A)$. Dann gilt $(K(x, \varepsilon) \setminus \{x\}) \cap A \neq \emptyset$ für alle $\varepsilon > 0$. Folglich ist auch $K(x, \varepsilon) \cap A \neq \emptyset$ für alle $\varepsilon > 0$ und damit $x \in cl(A)$.

Wir zeigen nun $cl(A) \subset A \cup HP(A)$:

Sei $x \in cl(A) \setminus A$. Wir müssen zeigen, dass x ein Häufungspunkt von A ist. Da $x \in cl(A)$ gilt für alle $\varepsilon > 0$ dass $K(x, \varepsilon) \cap A \neq \emptyset$. Da $x \notin A$ folgt auch $K(x, \varepsilon) \setminus \{x\} \cap A \neq \emptyset$. Als ist $x \in HP(A)$.

(3) A ist genau dann abgeschlossen, wenn $A = cl(A)$, also wegen (2) genau dann wenn $A = A \cup HP(A)$. Dies ist aber äquivalent dazu, dass $HP(A) \subset A$. \square

Definition 2.10. Sei (X, d) metrischer Raum und $A \subset X$ Teilmenge. Ein Punkt $x \in A$ heißt isolierter Punkt von A , falls ein $\varepsilon > 0$ existiert, so daß $K(x, \varepsilon) \cap A = \{x\}$ gilt. Mit $Iso(A)$ bezeichnen wir die Menge der isolierten Punkte von A .

Beispiel: Sei $X = \mathbb{R}^1$ versehen mit der Standardmetrik.

Für die Menge $A = (0, 1) \cup \{\frac{3}{2}\} \cup (2, 3) \subset \mathbb{R}$ gilt:

$$\begin{aligned} Iso(A) &= \{\frac{3}{2}\} \\ HP(A) &= [0, 1] \cup [2, 3] \\ \partial A &= \{0, 1, \frac{3}{2}, 2, 3\} \\ Int(A) &= (0, 1) \cup (2, 3) \\ cl(A) &= [0, 1] \cup \{\frac{3}{2}\} \cup [2, 3]. \end{aligned}$$

Für die Menge $B = \{1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots\} \subset \mathbb{R}$ gilt:

$$Iso(B) = B, \quad HP(B) = \{0\}, \quad \partial B = cl(B) = B \cup \{0\} \quad \text{und} \quad Int(B) = \emptyset.$$

Definition 2.11. Sei (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt dicht in X , wenn $cl(A) = X$ gilt.

Satz 2.8 Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge von X . Dann sind die folgenden Aussagen äquivalent:

1. A ist dicht in X .
2. A besitzt keine äußeren Punkte, d.h. $Int(X \setminus A) = \emptyset$.
3. $X \setminus A$ enthält keine ε -Kugel.

Beweis. Es gelten die folgenden äquivalenten Aussagen:

$$\begin{aligned} A \subset X \text{ dicht} &\iff cl(A) = X \\ &\iff X \setminus Int(X \setminus A) = X \\ &\iff Int(X \setminus A) = \emptyset \\ &\iff X \setminus A \text{ enthält keine } \varepsilon\text{-Kugel.} \end{aligned}$$

□

Beispiel: Wir betrachten wieder $X = \mathbb{R}$ mit der Standardmetrik.

Die Menge der rationalen Zahlen \mathbb{Q} und die Menge der irrationalen Zahlen $\mathbb{R} \setminus \mathbb{Q}$ sind dicht in \mathbb{R} . Um das einzusehen, erinnern wir uns an Kapitel 1. Wir wissen, dass in jedem Intervall (a, b) eine rationale Zahl liegt. Folglich enthält $\mathbb{R} \setminus \mathbb{Q}$ keine offenen Intervalle, das heißt keine Kugeln des metrischen Raumes \mathbb{R} . Somit ist \mathbb{Q} nach Satz 2.8 dicht in \mathbb{R} . Analog schließen wir für $\mathbb{R} \setminus \mathbb{Q}$, da jedes offene Intervall auch eine irrationale Zahl enthält.

Definition 2.12. Sei (X, d) ein metrischer Raum und $x \in X$. Unter einer Umgebung von x verstehen wir eine offene Menge $U \subset X$, die x enthält.

Bemerkung:

- Jede Kugel $K(x, \varepsilon)$ ist eine Umgebung von x .
- Jede Umgebung U von x enthält eine ε -Kugel um x .
- Jede Umgebung U von x ist die Vereinigung von Kugeln.
Da U offen ist, existiert nämlich für jedes $y \in U$ ein $\varepsilon(y) > 0$, so daß $K(y, \varepsilon(y)) \subset U$.
Folglich gilt $U = \bigcup_{y \in U} K(y, \varepsilon(y))$.

Man kann deshalb in allen obigen Definitionen (innerer Punkt, Abschluss, Randpunkt, Häufungspunkt, isolierter Punkt, ...) den Begriff "Kugeln um x " durch "Umgebungen von x " ersetzen.

2.4 Folgen in metrischen Räumen

Definition 2.13. Sei X eine nichtleere Menge. Unter einer Folge in X versteht man eine Abbildung

$$F: \mathbb{N} \longrightarrow X \\ n \longmapsto F(n) =: x_n,$$

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:

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

Definition 2.14. Sei (X, d) ein metrischer Raum und (x_n) eine Folge in X . Wir sagen, dass (x_n) gegen $x \in X$ konvergiert, falls zu jedem $\varepsilon > 0$ ein (von ε -abhängiger) Index $n_0 \in \mathbb{N}$ existiert, so dass $x_n \in K(x, \varepsilon)$ 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 Folge (x_n) schreiben wir:

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

Es gilt also:

- $\lim_{n \rightarrow \infty} x_n = x \iff \forall \varepsilon > 0 \exists n_0 \text{ mit } d(x, x_n) < \varepsilon \quad \forall n \geq n_0.$
- Betrachten wir speziell $X = \mathbb{R}$ mit der Standardmetrik $d(x, y) = |x - y|$, so gilt

$$\lim_{n \rightarrow \infty} x_n = x \iff \forall \varepsilon > 0 \exists n_0 \text{ mit } |x - x_n| < \varepsilon \quad \forall n \geq n_0.$$

Wenn wir im Folgenden vom metrischen Raum \mathbb{R} reden, so meinen wir immer die mit der Standardmetrik versehenen reellen Zahlen. Wir erhalten daraus die folgende Charakterisierung der Konvergenz von Folgen in einem metrischen Raum:

Satz 2.9 Eine Folge (x_n) in einem metrischen Raum (X, d) konvergiert genau dann gegen $x \in X$, wenn die Folge der Abstände $(d(x, x_n))$ im metrischen Raum \mathbb{R} gegen Null konvergiert:

$$x_n \longrightarrow x \text{ in } (X, d) \iff d(x, x_n) \longrightarrow 0 \text{ in } \mathbb{R}.$$

2.4.1 Allgemeine Eigenschaften konvergenter Folgen

Satz 2.10 (Eindeutigkeit des Grenzwerts) Der Grenzwert einer konvergenten Folge eines metrischen Raumes ist eindeutig bestimmt.

Beweis. Sei (x_n) eine Folge des metrischen Raumes (X, d) , die gegen x und x^* konvergiert. Angenommen $x \neq x^*$. Dann ist $\varepsilon := d(x, x^*) > 0$. Also existieren n_0 und n_0^* mit

$$d(x, x_n) < \frac{\varepsilon}{2} \quad \forall n \geq n_0 \quad \text{und} \quad d(x^*, x_n) < \frac{\varepsilon}{2} \quad \forall n \geq n_0^*.$$

Somit gilt $d(x, x_n) < \frac{\varepsilon}{2}$ und $d(x^*, x_n) < \frac{\varepsilon}{2}$ für alle $n \geq \max(n_0, n_0^*)$. Nach Dreiecksungleichung folgt für ein solches n

$$\varepsilon = d(x, x^*) \leq d(x, x_n) + d(x_n, x^*) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Dies ist aber ein Widerspruch. □

Definition 2.15. Sei (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt beschränkt, falls es eine Kugel $K(x_0, M)$ 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 2.11 Jede konvergente Folge (x_n) eines metrischen Raumes (X, d) ist beschränkt.

Beweis. Sei (x_n) eine konvergente Folge und $x = \lim_{n \rightarrow \infty} x_n$. Nach Definition der Konvergenz existiert ein $n_0 \in \mathbb{N}$, so dass $d(x_n, x) < 1$ für alle $n \geq n_0$. Wir setzen nun $M := \max(d(x, x_1), d(x, x_2), \dots, d(x, x_{n_0-1})) + 1$. Dann gilt $d(x, x_n) < M$ für alle $n \in \mathbb{N}$. Damit liegt die Menge $\{x_1, x_2, \dots\}$ in der Kugel $K(x, M)$ und ist somit beschränkt. □

Als nächstes wollen wir ein Kriterium für die Konvergenz von Folgen in Produkträumen behandeln. Wir erinnern nochmal an die Definition der Produktmetrik. Seien (X_1, d_1) , (X_2, d_2) , \dots , (X_k, d_k) metrische Räume. Das Produkt dieser k metrischen Räume ist das Paar (X, d) mit

$$X := X_1 \times X_2 \times \dots \times X_k, \\ d(a, b) := \sqrt{\sum_{j=1}^k d_j(a_j, b_j)^2}, \quad \text{wobei } a = (a_1, \dots, a_k), b = (b_1, \dots, b_k).$$

Satz 2.12 Seien $(X_1, d_1), \dots, (X_k, d_k)$ metrische Räume und (X, d) das Produkt dieser Räume. Eine Folge $(x_n = (x_{n1}, \dots, x_{nk}))_{n=1}^\infty$ von Punkten im Produktraum X konvergiert genau dann gegen $y = (y_1, \dots, y_k) \in X$, wenn für jedes $j \in \{1, \dots, k\}$ die Komponentenfolgen $(x_{nj})_{n=1}^\infty$ in (X_j, d_j) gegen y_j konvergieren.

Beweis. Nach Definition der Produktmetrik ist

$$d(\underbrace{(y_1, \dots, y_k)}_y, \underbrace{(x_{n1}, \dots, x_{nk})}_{x_n}) = \sqrt{\sum_{i=1}^k d_i(y_i, x_{ni})^2} \geq d_j(y_j, x_{nj}) \quad (*)$$

für jedes $j \in \{1, \dots, k\}$. Sei nun (x_n) gegen y in (X, d) konvergent. Dann existiert für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ mit $d(y, x_n) < \varepsilon$ für alle $n \geq n_0$. Nach Abschätzung (*) folgt daraus

$d_j(y_j, x_{nj}) < \varepsilon$ für alle $n \geq n_0$ und jedes $j \in \{1, \dots, k\}$. Somit konvergiert x_{nj} gegen y_j in (X_j, d_j) für jedes $j \in \{1, \dots, k\}$.

Sei umgekehrt $\lim_{n \rightarrow \infty} x_{nj} = y_j$ für jedes $j \in \{1, \dots, k\}$ und $\varepsilon > 0$. Dann existieren $n_{0j} \in \mathbb{N}$ so dass $d_j(y_j, x_{nj}) < \frac{\varepsilon}{\sqrt{k}}$ für alle $n \geq n_{0j}$. Es folgt

$$d(y, x_n) = \sqrt{\sum_{j=1}^k d_j(y_j, x_{nj})^2} < \sqrt{\sum_{j=1}^k \frac{\varepsilon^2}{k}} = \varepsilon$$

für alle $n \geq m_0 = \max(n_{01}, \dots, n_{0k})$. Also konvergiert x_n gegen y im Produktraum (X, d) . \square

Wir charakterisieren nun den Abschluß einer Teilmenge eines metrischen Raumes durch konvergente Folgen.

Satz 2.13 *Sei A eine beliebige Teilmenge eines metrischen Raumes (X, d) . Dann gilt für den Abschluß von A :*

$$x \in cl(A) \iff \text{Es existiert eine Folge } (a_n) \text{ mit } a_n \in A, \text{ so dass } \lim_{n \rightarrow \infty} a_n = x.$$

Beweis. (\implies) Sei $x \in cl(A)$. Dann gilt $K(x, \varepsilon) \cap A \neq \emptyset$ für alle $\varepsilon > 0$. Sei $n \in \mathbb{N}$ und $\varepsilon = \frac{1}{n}$. Dann existiert ein Element $a_n \in K(x, \frac{1}{n}) \cap A$. Wir erhalten also eine Folge a_1, a_2, a_3, \dots in A mit $d(x, a_n) < \frac{1}{n}$. Die Folge der Abstände $(d(x, a_n))$ konvergiert in \mathbb{R} gegen Null (Archimedisches Axiom). Somit konvergiert die Folge (a_n) gegen x .

(\impliedby) Sei (a_n) eine Folge in A , die gegen x konvergiert und $\varepsilon > 0$. Dann existiert ein $n_0 \in \mathbb{N}$, so dass $a_n \in K(x, \varepsilon)$ für alle $n \geq n_0$. Folglich ist $K(x, \varepsilon) \cap A \neq \emptyset$, also $x \in cl(A)$. \square

Folgerung 2.1 *Sei (X, d) ein metrischer Raum. Dann gilt:*

1. *Eine Teilmenge $A \subset X$ ist genau dann abgeschlossen, wenn der Grenzwert jeder konvergenten, vollständig in A liegenden Folge gleichfalls in A liegt.*
2. *Eine Teilmenge $A \subset X$ ist genau dann dicht im metrischen Raum (X, d) , wenn für jedes $x \in X$ eine Folge (a_n) von Punkten aus A existiert, die gegen x konvergiert.*

Beweis. $A \subset X$ ist genau dann abgeschlossen, wenn $cl(A) = A$. $A \subset X$ ist genau dann dicht, wenn $cl(A) = X$. Die Charakterisierung von $cl(A)$ aus Satz 2.13 liefert dann die Behauptung. \square

Definition 2.16. *Sei (x_n) eine Folge in einer Menge X . 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 2.14 *Sei (X, d) ein metrischer Raum und (x_n) eine gegen $x \in X$ konvergente Folge. Dann konvergiert jede Teilfolge von (x_n) ebenfalls gegen x .*

Definition 2.17. Sei (x_n) eine Folge im metrischen Raum (X, d) . Ein Punkt $x \in X$ 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, d) ein metrischer Raum und (x_n) eine gegen $x \in X$ konvergente Folge. Dann gilt nach Satz 2.14, dass $HP(x_n) = \{x\}$.
2. Wir betrachten in \mathbb{R} die 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 $HP(x_n) = \{0, 1\}$.

2.4.2 Spezielle Eigenschaften von konvergenten Folgen im Vektorraum \mathbb{C}^k bzw. \mathbb{R}^k

In diesem Abschnitt betrachten wir spezielle Eigenschaften konvergenter Folgen im Vektorraum \mathbb{C}^k bzw. \mathbb{R}^k , den wir mit der in Kapitel 1.4 eingeführten Standardmetrik

$$d(z, w) := \|z - w\| := \sqrt{|z_1 - w_1|^2 + \dots + |z_k - w_k|^2}$$

$$z = (z_1, \dots, z_k), w = (w_1, \dots, w_k) \in \mathbb{C}^k \text{ bzw. } \mathbb{R}^k$$

versehen. Da der metrische Raum \mathbb{R}^k ein Teilraum des metrischen Raumes \mathbb{C}^k ist, gelten alle Eigenschaften, die wir im folgenden für Folgen in \mathbb{C}^k formulieren, auch für Folgen in \mathbb{R}^k . Aus Satz 2.12 folgt als Spezialfall sofort

Satz 2.15

1. Eine Folge von Vektoren $(z_n = (z_{n1}, \dots, z_{nk}))_{n=1}^{\infty}$ aus \mathbb{C}^k konvergiert genau dann gegen den Vektor $z^* = (z_1^*, \dots, z_k^*) \in \mathbb{C}^k$, wenn die Komponentenfolge $(z_{nj})_{n=1}^{\infty}$ für jedes $j \in \{1, \dots, k\}$ in \mathbb{C} gegen z_j^* konvergiert.
2. Eine Folge komplexer Zahlen (w_n) konvergiert genau dann gegen $w \in \mathbb{C}$, wenn die Folge der Realteile $(\operatorname{Re}(w_n))$ gegen $\operatorname{Re}(w)$ und die Folge der Imaginärteile $(\operatorname{Im}(w_n))$ gegen $\operatorname{Im}(w)$ konvergiert.
3. Eine Folge komplexer Zahlen (w_n) konvergiert genau dann gegen $w \in \mathbb{C}$, wenn die Folge (\overline{w}_n) gegen \overline{w} konvergiert.

Der folgende Satz fasst die wichtigsten Rechenregeln für Folgen in \mathbb{C}^k (bzw. in \mathbb{R}^k) zusammen.

Satz 2.16 (Rechenregeln für konvergente Folgen in \mathbb{C}^k)

1. Seien (z_n) und (w_n) konvergente Folgen in \mathbb{C}^k mit den Grenzwerten $\lim_{n \rightarrow \infty} z_n = z \in \mathbb{C}^k$ und $\lim_{n \rightarrow \infty} w_n = w \in \mathbb{C}^k$. Dann gilt:

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

$$(b) \lim_{n \rightarrow \infty} \mu \cdot z_n = \mu \cdot \lim_{n \rightarrow \infty} z_n = \mu \cdot z \quad \forall \mu \in \mathbb{C},$$

$$(c) \lim_{n \rightarrow \infty} \langle z_n, w_n \rangle = \langle \lim_{n \rightarrow \infty} z_n, \lim_{n \rightarrow \infty} w_n \rangle = \langle z, w \rangle,$$

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

2. Seien (z_n) und (w_n) zwei konvergente Folgen komplexer Zahlen mit den Grenzwerten $\lim_{n \rightarrow \infty} z_n = z$ und $\lim_{n \rightarrow \infty} w_n = w$ in \mathbb{C} . Dann gilt:

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

(b) 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

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

3. 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 fast alle¹ $n \in \mathbb{N}$, so folgt $x \leq y$.

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

Beweis. Zu 1. (a) Es gilt

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

und für alle $\varepsilon > 0$ existieren $n_0, n_0^* \in \mathbb{N}$, 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 $\|(z_n + w_n) - (z + w)\| < \varepsilon$ für alle $n \geq \max(n_0, n_0^*)$ und somit $\lim_{n \rightarrow \infty} (z_n + w_n) = (z + w)$.

(b) Ist $\mu = 0$, so gilt die Behauptung trivialerweise. Sei $\mu \neq 0$. Dann gilt

$$\|\mu z - \mu z_n\| = |\mu| \cdot \|z - z_n\|.$$

Da (z_n) gegen z konvergiert, existiert für $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass

$$\|z - z_n\| < \frac{\varepsilon}{|\mu|} \quad \forall n \geq n_0.$$

¹ für fast alle $n \in \mathbb{N}$ bedeutet, dass die Aussage mit eventueller Ausnahme von endlich vielen natürlichen Zahlen gilt.

Folglich gilt $\|\mu z - \mu z_n\| = |\mu| \cdot \|z - z_n\| < \varepsilon$ für alle $n \geq n_0$, und somit $\lim_{n \rightarrow \infty} \mu z_n = \mu z$.

(c) Mittels der Cauchy-Schwarzschen Ungleichung erhält man die folgende Abschätzung

$$\begin{aligned} |\langle z_n, w_n \rangle - \langle z, w \rangle| &= |\langle z_n - z, w_n \rangle + \langle z, w_n - w \rangle| \\ &\leq |\langle z_n - z, w_n \rangle| + |\langle z, w_n - w \rangle| \\ &\leq \|z_n - z\| \|w_n\| + \|z\| \|w_n - w\|. \end{aligned}$$

Da die Folge (w_n) konvergiert, ist die Menge $\{w_1, w_2, \dots\}$ beschränkt. Somit existiert ein $M \in \mathbb{R}^+$ mit $\|w_n\| \leq M$ für alle $n \in \mathbb{N}$. Wir wählen außerdem ein $C \in \mathbb{R}^+$ mit $\|z\| < C$. Sei nun $\varepsilon > 0$ gegeben. Dann existieren n_0 und n_0^* , so dass

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

Wir erhalten

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

Daraus folgt $\lim_{n \rightarrow \infty} \langle z_n, w_n \rangle = \langle z, w \rangle$.

(d) Sei $z_n \rightarrow z$ und $\varepsilon > 0$. Aus der Dreiecksungleichung für die Norm in \mathbb{C}^k folgt

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

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

Zu 2. (a) $z_n \cdot w_n = \langle z_n, \bar{w}_n \rangle \xrightarrow{(1c)} \langle z, \bar{w} \rangle = z \cdot w$.

(b) 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

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

Daraus folgt: Für alle $\varepsilon > 0$ existiert ein $n_0^* \geq n_0$, so dass

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

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

Zu 3. (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.

(b) Sei $x_n \leq u_n \leq y_n$ für alle $n \geq n_0$. Da $x_n \rightarrow x$ und $y_n \rightarrow x$, existiert für jedes $\varepsilon > 0$ ein n_1 so dass $|x_n - x| < \frac{\varepsilon}{4}$ und $|y_n - x| < \frac{\varepsilon}{4}$ für alle $n \geq n_1$. Aus der Dreiecks-Ungleichung folgt:

$$|x - u_n| \leq |x - x_n| + |x_n - u_n| \leq |x - x_n| + |x_n - y_n| \leq |x - x_n| + |x_n - x| + |x - y_n| < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < \varepsilon$$

für alle $n \geq \max\{n_0, n_1\}$. Folglich konvergiert die Folge (u_n) gegen x . \square

Definition 2.18. Eine Folge (z_n) in \mathbb{C}^k , die gegen den Nullvektor $0 = (0, \dots, 0)$ konvergiert, heißt Nullfolge.

Aus Satz 2.16 folgt dann:

- Eine Folge (z_n) von Vektoren aus \mathbb{C}^k ist genau dann eine Nullfolge, wenn die Folge der reellen Zahlen $(\|z_n\|)$ eine Nullfolge ist.
- Eine Folge von Vektoren (z_n) konvergiert genau dann gegen den Vektor z , wenn $(z_n - z)$ eine Nullfolge ist.
- Ist (z_n) eine Nullfolge von Vektoren und $\mu \in \mathbb{C}$, so ist auch $(\mu \cdot z_n)$ eine Nullfolge.
- Sind (z_n) und (w_n) Nullfolgen von Vektoren, so ist auch $(z_n + w_n)$ eine Nullfolge.
- Ist (z_n) eine Nullfolge und (w_n) eine konvergente Folge von Vektoren, so ist die Folge der Skalarprodukte $(\langle z_n, w_n \rangle)$ 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} :

Wir erinnern nochmal daran, dass \mathbb{R} und \mathbb{C} immer mit der durch den Betrag gegebenen Standardmetrik versehen sind.

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 2.16 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 $x_n := \sqrt[n]{n} - 1$. Es gilt $x_n \geq 0$. Aus der binomischen Formel folgt

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

Folglich gilt $0 \leq x_n \leq \sqrt{\frac{2}{n-1}}$. Nach 1. ist $(\sqrt{\frac{2}{n-1}})$ 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 $\boxed{\lim_{n \rightarrow \infty} \frac{n^k}{z^n} = 0}$.

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 $\lim_{n \rightarrow \infty} \frac{n^k}{z^n} = 0$.

2.4.3 Spezielle Eigenschaften konvergenter Folgen in \mathbb{R}

In diesem Abschnitt betrachten wir weitere, spezielle Eigenschaften von Folgen reeller Zahlen, wobei \mathbb{R} wieder mit der Standardmetrik $d(x, y) := |x - y|$ versehen sei. 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 2.19. 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 und strebt auch 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.

Eine wichtige Folgerung aus dem Vollständigkeitsaxiom der reellen Zahlen ist die folgende Eigenschaft beschränkter Folgen².

Satz 2.17 (Satz von Bolzano/Weierstraß)

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 1.6 existiert ein Supremum $g = \sup A$ der Menge A . Sei $\varepsilon > 0$. Aus der Definition des Supremums erhalten wir

² Karl Weierstraß (1815-1897), Bernhard Bolzano (1781-1848).

- 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 2.20. 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.

Satz 2.18 Jede *monoton wachsende, nach oben beschränkte Folge reeller Zahlen* (x_n) *konvergiert gegen* $\sup\{x_n \mid n \in \mathbb{N}\}$.

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

Beweis. Sei (x_n) *monoton wachsend* und *nach oben beschränkt*. Nach Satz 1.6 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

Das Supremum und das Infimum einer beschränkten Teilmenge $A \subset \mathbb{R}$ liegt im Abschluss $cl(A)$. Ist $A \subset \mathbb{R}$ *beschränkt* und *abgeschlossen*, so existiert folglich das Maximum $\max(A)$ und das Minimum $\min(A)$ von A .

Sei nun (x_n) eine *beschränkte Folge reeller Zahlen*. Dann ist die Menge der Häufungspunkte $HP(x_n)$ *nicht leer* (Satz 2.17) und *ebenfalls beschränkt*. Außerdem ist sie *abgeschlossen* (siehe Übungsaufgaben). Folglich existiert das Maximum $\max HP(x_n)$, d.h. ein *größter Häufungspunkt* von (x_n) , und das Minimum $\min HP(x_n)$, d.h. ein *kleinster Häufungspunkt* von (x_n) .

Definition 2.21. Sei (x_n) eine beschränkte Folge reeller Zahlen. Man nennt den größten Häufungspunkt von (x_n) auch *limes superior* von (x_n) und bezeichnet ihn mit

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

Den kleinsten Häufungspunkt von (x_n) nennt man auch *limes inferior* von (x_n) und bezeichnet ihn mit

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

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 2.19 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

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

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 2.20 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 2.18 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:

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

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 2.21 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 2.20 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 2.18 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 2.20

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

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

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

Satz 2.22 *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

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

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

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

\square

Mittels der Fehlerabschätzung aus Satz 2.22 kann man Näherungswerte für e angeben. Man erhält zum Beispiel für $n = 10$:

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

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

$$e \approx 2.7182815 .$$

Als weitere Anwendung der Fehlerabschätzung beweisen wir

Satz 2.23 *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 2.22

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

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

2.5 Vollständige metrische Räume

Um mittels der Definition nachzuweisen, dass eine Folge konvergiert, muß man ihren Grenzwert kennen. In diesem Abschnitt lernen wir eine Klasse von metrischen Räumen kennen, in denen man die Konvergenz einer *beliebigen* Folge überprüfen kann, ohne ihren Grenzwert zu kennen.

Definition 2.22. *Sei (X, d) ein metrischer Raum. Eine Folge (x_n) in (X, d) heißt Cauchy-Folge, wenn zu jedem $\varepsilon > 0$ ein (von ε abhängiges) $n_0 \in \mathbb{N}$ existiert, so dass $d(x_n, x_m) < \varepsilon$ für alle $n, m \geq n_0$.*

Satz 2.24 *Jede Cauchy-Folge in einem metrischen Raum ist beschränkt.*

Beweis. Sei $\varepsilon = 1$. Dann existiert ein $n_0 \in \mathbb{N}$ so dass $d(x_n, x_m) < 1$ für alle $n, m \geq n_0$. Insbesondere bedeutet das, dass $x_n \in K(x_{n_0}, 1)$ für alle $n \geq n_0$.

Wir setzen

$$r := \max \{ d(x_{n_0}, x_1), d(x_{n_0}, x_2), \dots, d(x_{n_0}, x_{n_0-1}) \} + 1.$$

Dann gilt $x_n \in K(x_{n_0}, r)$ für alle $n \in \mathbb{N}$. Also ist die Folge (x_n) beschränkt. \square

Satz 2.25 *Jede konvergente Folge in einem metrischen Raum ist eine Cauchy-Folge.*

Beweis. Sei (x_n) eine gegen x konvergente Folge und $\varepsilon > 0$. Dann existiert ein $n_0 \in \mathbb{N}$ mit $d(x_n, x) < \frac{\varepsilon}{2}$ für alle $n \geq n_0$. Für $n, m \geq n_0$ folgt dann aus der Dreiecksungleichung

$$d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Somit ist (x_n) eine Cauchy-Folge. □

Es gibt Cauchy-Folgen, die nicht konvergieren. Als ein Beispiel betrachten wir den metrischen Raum (X, d) mit $X = (0, 1] \subset \mathbb{R}$ und der Metrik $d(x, y) = |x - y|$. Sei $x_n := \frac{1}{n}$. Die Folge (x_n) ist eine Nullfolge in \mathbb{R} , folglich ist sie eine Cauchy-Folge in (\mathbb{R}, d) . Sie ist dann auch Cauchyfolge in (X, d) , hat aber in X keinen Grenzwert, da $0 \notin X$.

Definition 2.23. *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*

$$\boxed{\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \text{ mit } d(x_n, x_m) < \varepsilon \forall n, m \geq n_0.}$$

Wir sehen uns zunächst Beispiele für vollständige metrischer Räume an.

Satz 2.26 *Die reellen Zahlen \mathbb{R} mit der Standardmetrik $d(x, y) = |x - y|$ sind ein vollständiger metrischer Raum.*

Beweis. Sei (x_n) eine Cauchy-Folge reeller Zahlen. Dann ist (x_n) beschränkt. Nach dem Satz von Bolzano-Weierstraß existiert eine konvergente Teilfolge (x_{n_k}) von (x_n) . Sei $x = \lim_{k \rightarrow \infty} x_{n_k}$ und $\varepsilon > 0$. Dann existieren $k_0, n_0 \in \mathbb{N}$, so dass

$$\begin{aligned} |x - x_{n_k}| &< \frac{\varepsilon}{2} & \forall k \geq k_0 \\ |x_n - x_m| &< \frac{\varepsilon}{2} & \forall n, m \geq n_0. \end{aligned}$$

Sei nun $n \geq n_0$. Wir wählen ein $k \geq k_0$ mit $n_k \geq n_0$. Dann folgt aus der Dreiecksungleichung

$$|x_n - x| \leq |x_n - x_{n_k}| + |x_{n_k} - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} < \varepsilon.$$

Also konvergiert die Folge (x_n) gegen x . □

Beispiel 1: Die metrischen Räume $(0, +\infty)$, $(0, a)$, $(0, a]$ mit der Standardmetrik sind nicht vollständig.

Beispiel 2: Der metrische Raum der rationalen Zahlen \mathbb{Q} (mit der Standardmetrik) ist *nicht* vollständig.

- 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 Übungsaufgaben). 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 konvergiert (x_n) nicht in \mathbb{Q} .

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

Satz 2.27 Sind $(X_1, d_1), \dots, (X_k, d_k)$ vollständige metrische Räume, so ist auch das Produkt (X, d) dieser metrischen Räume vollständig.

Beweis. Sei $(x_n := (x_{n1}, \dots, x_{nk}))$ eine Cauchy-Folge im Produktraum $(X = X_1 \times \dots \times X_k, d)$ 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 $d_i(x_{ni}, x_{mi}) < \varepsilon$ für alle $n, m \geq n_0$ und jedes $i \in \{1, \dots, k\}$. Folglich sind die Folgen $(x_{ni})_{n=1}^\infty$ Cauchy-Folgen in (X_i, d_i) für jedes $i \in \{1, \dots, k\}$. Da (X_i, d_i) vollständige metrische Räume sind, konvergieren die Folgen $(x_{ni})_{n=1}^\infty$ gegen ein $y_i \in X_i$. Somit gilt nach Satz 2.12 $\lim_{n \rightarrow \infty} x_n = (y_1, \dots, y_k) \in X_1 \times \dots \times X_k$. \square

Folgerung 2.2 Die metrischen Räume \mathbb{C} , \mathbb{R}^k und \mathbb{C}^k (jeweils mit der Standardmetrik) sind vollständig.

Satz 2.28 Sei (X, d) ein vollständiger metrischer Raum und $A \subset X$ eine Teilmenge. Dann ist der metrische Raum $(A, d_A := d|_{A \times A})$ genau dann vollständig, wenn A abgeschlossen ist.

Beweis. (\Leftarrow): Sei A abgeschlossen und (a_n) eine Cauchy-Folge in (A, d_A) . Dann ist (a_n) auch Cauchy-Folge in (X, d) , die, da (X, d) vollständig ist, gegen ein $x \in X$ konvergiert. Da A abgeschlossen ist, liegt jeder Grenzwert einer Folge von Elementen aus A wieder in A (Folgerung 2.1). Somit ist $x \in A$, das heißt (a_n) konvergiert in A . Folglich ist (A, d_A) ein vollständiger metrischer Raum.

(\Rightarrow): Sei nun (A, d_A) vollständig. Wir zeigen, dass dann A abgeschlossen ist. Sei $x \in cl(A)$. Nach Satz 2.13 existiert eine Folge $a_n \in A$ mit $\lim_{n \rightarrow \infty} a_n = x$. Damit ist (a_n) eine Cauchy-Folge in (A, d_A) und somit in A konvergent. Das heißt, x liegt in A . Folglich gilt $cl(A) \subset A$. Die Teilmenge $A \subset X$ ist also abgeschlossen. \square

Folgerung 2.3 Jede abgeschlossene Teilmenge $A \subset \mathbb{R}^n$ bzw. $A \subset \mathbb{C}^n$ ist ein vollständiger metrischer Raum (mit Standardmetrik).

Wir wissen, dass nicht jeder metrische Raum vollständig ist. Man kann aber jeden metrischen Raum *vervollständigen*. Abschließend beschreiben wir diese *Vervollständigungsprozedur*, die uns zeigt, dass man jeden metrischen Raum als dichten Teilraum eines gewissen, bis auf Isometrie eindeutig bestimmten, vollständigen metrischen Raumes auffassen kann. Diese Prozedur wird in der Analysis, vor allem bei der Untersuchung partieller Differentialgleichungen, vielfältig angewendet³.

Satz 2.29 (Vervollständigung metrischer Räume)

Zu jedem metrischen Raum (X, d) gibt es einen vollständigen metrischen Raum (\tilde{X}, \tilde{d}) und eine Abbildung $\varphi : X \rightarrow \tilde{X}$, für die gilt:

1. $\varphi : (X, d) \rightarrow (\varphi(X), \tilde{d})$ ist eine Isometrie.
2. $\varphi(X) \subset \tilde{X}$ ist dicht.

Der metrische Raum (\tilde{X}, \tilde{d}) ist bis auf Isometrie eindeutig bestimmt, d.h. ist (X', d') ein weiterer vollständiger metrischer Raum und $\varphi' : X \rightarrow X'$ eine Abbildung mit den Eigenschaften 1. und 2., dann existiert eine Isometrie $F : (\tilde{X}, \tilde{d}) \rightarrow (X', d')$ so dass $F \circ \varphi = \varphi'$.

Definition 2.24. Der metrische Raum (\tilde{X}, \tilde{d}) aus Satz 2.29 heißt *Vervollständigung* von (X, d) .

Beweis. Den Beweis führen wir in mehreren Schritten.

1. Schritt: Definition der Menge \tilde{X} .

Dazu betrachten wir die Menge $\mathcal{CF}(X, d)$ der Cauchy-Folgen des metrischen Raumes (X, d) . Zwei Cauchy-Folgen $(x_n), (y_n) \in \mathcal{CF}(X, d)$ bezeichnen wir als äquivalent (symbolisch: $(x_n) \sim (y_n)$), wenn

$$d(x_n, y_n) \rightarrow 0 \quad \text{für } n \rightarrow +\infty.$$

Dies definiert eine Äquivalenzrelation auf dem Raum $\mathcal{CF}(X, d)$:

Aus $d(x_n, x_n) = 0$ und $d(x_n, y_n) = d(y_n, x_n)$ ergeben sich unmittelbar die Reflexivität und die Symmetrie der Relation \sim . Aus der Dreiecksungleichung

$$0 \leq d(x_n, z_n) \leq d(x_n, y_n) + d(y_n, z_n)$$

folgt, dass mit $d(x_n, y_n) \rightarrow 0$ und $d(y_n, z_n) \rightarrow 0$ stets $d(x_n, z_n) \rightarrow 0$ gilt. Dies zeigt die Transitivität von \sim . Die Menge \tilde{X} sei die Menge der Äquivalenzklassen von Cauchy-Folgen bzgl. der Relation \sim :

³ In der Vorlesung haben wir den Vervollständigungssatz als Fakt zitiert und aus Zeitgründen nur die Beweisidee angegeben. Es lohnt sich aber, den vollständigen Beweis anzusehen. Zum einen ist dies ein zentrales Resultat der Analysis. Zum anderen üben Sie beim Durcharbeiten des Beweises viele Begriffe und Techniken der bisherigen Vorlesungen Analysis I* und Lineare Algebra I*.

$$\tilde{X} := \mathcal{CF}(X, d) / \sim .$$

Die Äquivalenzklasse der Cauchy-Folge $(x_n) \in \mathcal{CF}(X, d)$ bezeichnen wir mit $[(x_n)]$.

2. Schritt: Definition des Abstandes \tilde{d} auf \tilde{X} .

Seien $(x_n), (y_n) \in \mathcal{CF}(X, d)$ zwei Cauchy-Folgen von (X, d) . Nach der Vierecksungleichung gilt

$$|d(x_m, y_m) - d(x_n, y_n)| \leq d(x_m, x_n) + d(y_m, y_n) .$$

Folglich ist die Folge der reellen Zahlen $(d(x_n, y_n))$ eine Cauchy-Folge im metrischen Raum \mathbb{R} . Da \mathbb{R} vollständig ist, konvergiert diese Folge, d.h. es existiert $\lim_{n \rightarrow \infty} d(x_n, y_n)$. Wir definieren für zwei Elemente $\xi = [(x_n)]$ und $\eta = [(y_n)]$ von \tilde{X} den Abstand

$$\tilde{d}(\xi, \eta) := \lim_{n \rightarrow \infty} d(x_n, y_n) .$$

Diese Definition ist korrekt, d.h. sie hängt nicht von der Wahl der Repräsentanten (x_n) bzw. (y_n) in den Äquivalenzklassen ξ bzw. η ab. Sind nämlich $(x_n) \sim (x'_n)$ und $(y_n) \sim (y'_n)$, so gilt nach Vierecksungleichung

$$|d(x'_n, y'_n) - d(x_n, y_n)| \leq d(x'_n, x_n) + d(y'_n, y_n) .$$

Benutzt man nun die Definition von \sim , so folgt

$$\lim_{n \rightarrow \infty} d(x_n, y_n) = \lim_{n \rightarrow \infty} d(x'_n, y'_n) .$$

Die Abbildung $\tilde{d} : \tilde{X} \times \tilde{X} \rightarrow \mathbb{R}$ ist ein Abstand auf \tilde{X} : Die Symmetrie und die Positivität von \tilde{d} sind offensichtlich wegen der analogen Eigenschaften von d . Die Äquivalenzrelation \sim auf $\mathcal{CF}(X, d)$ ist gerade so definiert, dass

$$\tilde{d}(\xi, \eta) = 0 \iff \lim_{n \rightarrow \infty} d(x_n, y_n) = 0 \iff (x_n) \sim (y_n) \iff \xi = \eta .$$

Die Dreiecksungleichung folgt ebenfalls aus derjenigen von d :

Seien $\xi = [(x_n)], \eta = [(y_n)], \rho := [(w_n)]$. Dann gilt

$$\begin{aligned} \tilde{d}(\xi, \eta) &= \lim_{n \rightarrow \infty} d(x_n, y_n) \\ &\leq \lim_{n \rightarrow \infty} (d(x_n, w_n) + d(w_n, y_n)) \\ &= \tilde{d}(\xi, \rho) + \tilde{d}(\rho, \eta) . \end{aligned}$$

Damit haben wir einen metrischen Raum (\tilde{X}, \tilde{d}) definiert.

3. Schritt: Definition der Abbildung φ :

Sei $x \in X$. Wir betrachten die konstante Folge $(x_n := x)$ in X und bezeichnen mit $\tilde{x} \in \tilde{X}$ ihre Äquivalenzklasse $\tilde{x} := [(x_n := x)]$ in (\tilde{X}, \tilde{d}) . Die Abbildung φ definieren wir durch

$$\begin{aligned} \varphi : X &\longrightarrow \tilde{X} \\ x &\longmapsto \tilde{x} \end{aligned}$$

Die Abbildung $\varphi : X \rightarrow \varphi(X)$ ist offensichtlich bijektiv und es gilt nach Definition von \tilde{d}

$$\tilde{d}(\varphi(x), \varphi(y)) = \tilde{d}(\tilde{x}, \tilde{y}) = \lim_{n \rightarrow \infty} d(x, y) = d(x, y).$$

(X, d) ist also isometrisch zum Teilraum $(\varphi(X), \tilde{d})$ von (\tilde{X}, \tilde{d}) .

4. Schritt: Der metrische Raum (\tilde{X}, \tilde{d}) ist vollständig.

Sei $(\xi_k)_{k=1}^{\infty}$ eine beliebige Cauchy-Folge aus (\tilde{X}, \tilde{d}) . Wir zeigen, dass (ξ_k) in (\tilde{X}, \tilde{d}) konvergiert. Zunächst wählen wir in jeder Äquivalenzklasse ξ_k einen Repräsentanten $(x_{kn})_{n=1}^{\infty} \in \mathcal{CF}(X, d)$. In jeder dieser Cauchy-Folgen wählen wir dann ein Folgenglied $x_{kn_k} \in X$ so, dass

$$d(x_{kn}, x_{kn_k}) < \frac{1}{k} \quad \forall n > n_k. \quad (*)$$

Wir zeigen nun, dass die Folge $(x_{kn_k})_{k=1}^{\infty}$ eine Cauchy-Folge in (X, d) ist. Aus der Isometrie zwischen (X, d) und $(\varphi(X), \tilde{d})$, der Dreiecksungleichung und $(*)$ folgt

$$\begin{aligned} d(x_{kn_k}, x_{k'n_{k'}}) &= \tilde{d}(\tilde{x}_{kn_k}, \tilde{x}_{k'n_{k'}}) \\ &\leq \tilde{d}(\tilde{x}_{kn_k}, \xi_k) + \tilde{d}(\xi_k, \xi_{k'}) + \tilde{d}(\xi_{k'}, \tilde{x}_{k'n_{k'}}) \\ &= \lim_{n \rightarrow \infty} d(x_{kn_k}, x_{kn}) + \tilde{d}(\xi_k, \xi_{k'}) + \lim_{n \rightarrow \infty} d(x_{k'n}, x_{k'n_{k'}}) \\ &< \frac{1}{k} + \tilde{d}(\xi_k, \xi_{k'}) + \frac{1}{k'}. \end{aligned}$$

Da $(\xi_k)_{k=1}^{\infty}$ eine Cauchy-Folge in (\tilde{X}, \tilde{d}) ist, existiert zu vorgegebenem $\varepsilon > 0$ ein Index $k_0 \in \mathbb{N}$, so dass

$$\tilde{d}(\xi_k, \xi_{k'}), \frac{1}{k}, \frac{1}{k'} < \frac{\varepsilon}{6} \quad \forall k, k' \geq k_0.$$

Folglich ist

$$d(x_{kn_k}, x_{k'n_{k'}}) < \frac{\varepsilon}{2} \quad \forall k, k' \geq k_0. \quad (**)$$

und somit $(x_{kn_k})_{k=1}^{\infty}$ eine Cauchy-Folge in (X, d) . Die Äquivalenzklasse dieser Folge sei $\xi := [(x_{kn_k})]$. Wir zeigen, dass (ξ_k) in (\tilde{X}, \tilde{d}) gegen ξ konvergiert.

Aus $(*)$ und $(**)$ folgt

$$\begin{aligned} \tilde{d}(\xi_k, \xi) &\leq \tilde{d}(\xi_k, \tilde{x}_{kn_k}) + \tilde{d}(\tilde{x}_{kn_k}, \xi) \\ &= \lim_{n \rightarrow \infty} d(x_{kn}, x_{kn_k}) + \lim_{k' \rightarrow \infty} d(x_{kn_k}, x_{k'n_{k'}}) \\ &< \varepsilon \quad \forall k > k_0. \end{aligned}$$

Folglich konvergiert die Cauchy-Folge $(\xi_k)_{k=1}^{\infty}$ im Raum (\tilde{X}, \tilde{d}) gegen $\xi \in \tilde{X}$.

5. Schritt: Wir zeigen, dass die Menge $\varphi(X)$ dicht in \tilde{X} ist.

Sei $\xi \in \tilde{X}$ ein beliebiges Element und $(x_n)_{n=1}^{\infty}$ eine Cauchy-Folge in der Äquivalenzklasse ξ . Wir betrachten die entsprechende Folge (\tilde{x}_n) in $\varphi(X)$. Für $\varepsilon > 0$ existiert ein $n_0 \in \mathbb{N}$ mit

$$\tilde{d}(\tilde{x}_n, \xi) = \lim_{m \rightarrow \infty} d(x_n, x_m) < \varepsilon \quad \forall n > n_0.$$

Folglich konvergiert (\tilde{x}_n) gegen ξ und die Dichtheit von $\varphi(X)$ in \tilde{X} ist bewiesen.

6. Schritt: (\tilde{X}, \tilde{d}) ist bis auf Isometrie eindeutig bestimmt:

Sei (X', d') ein weiterer vollständiger metrischer Raum und $\varphi' : X \rightarrow X'$ eine Abbildung, so dass $\varphi' : (X, d) \rightarrow (\varphi'(X), d')$ eine Isometrie und $\varphi'(X) \subset X'$ dicht ist. Dann definiert die Abbildung

$$\begin{aligned} f : \varphi(X) &\longrightarrow \varphi'(X) \\ \varphi(x) &\longmapsto \varphi'(x) \end{aligned}$$

offensichtlich eine Isometrie zwischen den metrischen Räumen $(\varphi(X), \tilde{d})$ und $(\varphi'(X), d')$. Wir erweitern f zu einer Isometrie F zwischen (\tilde{X}, \tilde{d}) und (X', d') : Sei $\xi \in \tilde{X}$ und $\xi = \lim_{n \rightarrow \infty} \tilde{x}_n$ für eine Folge (\tilde{x}_n) aus $\varphi(X)$. Da (\tilde{x}_n) eine Cauchy-Folge in $\varphi(X)$ ist, ist auch ihr Bild $(f(\tilde{x}_n))$ eine Cauchy-Folge in $\varphi'(X)$. Sei $\xi' := \lim_{n \rightarrow \infty} f(\tilde{x}_n) \in X'$. ξ' ist korrekt definiert, d.h. unabhängig von der Wahl der Folge (\tilde{x}_n) . Sei nämlich (\tilde{y}_n) eine weitere Folge mit $\xi = \lim_{n \rightarrow \infty} \tilde{y}_n$, so gilt $\tilde{d}(\tilde{x}_n, \tilde{y}_n) = d'(f(\tilde{x}_n), f(\tilde{y}_n)) \rightarrow 0$, also $\lim_{n \rightarrow \infty} f(\tilde{y}_n) = \xi'$. Wir erhalten somit eine Fortsetzung von f zu einer Abbildung

$$\begin{aligned} F : \tilde{X} &\longrightarrow X' \\ \xi = \lim_{n \rightarrow \infty} \tilde{x}_n &\longmapsto \xi' = \lim_{n \rightarrow \infty} f(\tilde{x}_n) . \end{aligned}$$

F ist offensichtlich bijektiv. Für $\xi = \lim_{n \rightarrow \infty} \tilde{x}_n \in \tilde{X}$ und $\eta = \lim_{n \rightarrow \infty} \tilde{y}_n \in \tilde{X}$ gilt

$$\tilde{d}(\xi, \eta) = \lim_{n \rightarrow \infty} \tilde{d}(\tilde{x}_n, \tilde{y}_n) = \lim_{n \rightarrow \infty} d'(f(\tilde{x}_n), f(\tilde{y}_n)) = d'(\xi', \eta') = d'(F(\xi), F(\eta)).$$

Die Abbildung F ist folglich eine Isometrie zwischen (\tilde{X}, \tilde{d}) und (X', d') . Aus der Definition von F folgt sofort, dass $F \circ \varphi = \varphi'$. Damit ist Satz 2.29 vollständig bewiesen. \square

Bemerkung: Eine Möglichkeit zur Konstruktion der reellen Zahlen aus den rationalen Zahlen besteht darin, die in Satz 2.29 beschriebene Vervollständigungsverfahren für die Menge der rationalen Zahlen \mathbb{Q} mit dem Abstand $d(x, y) = |x - y|$ auszuführen. Die entstehende Menge $\tilde{\mathbb{Q}}$ erfüllt dann die Axiome der reellen Zahlen aus Kapitel 1.

2.6 Kompakte und folgenkompakte Teilmengen metrischer Räume

Definition 2.25. Eine Teilmenge A eines metrischen Raumes (X, d) heißt folgenkompakt, falls jede Folge in A eine in A konvergente Teilfolge besitzt.

Ist die gesamte Menge X folgenkompakt, so nennt man den metrischen Raum (X, d) folgenkompakt.

Bemerkung: Sei (X, d) ein metrischer Raum und $A \subset X$. Dann ist $A \subset X$ genau dann folgenkompakt, wenn der metrische Raum $(A, d|_{A \times A})$ folgenkompakt ist.

Wir betrachten zunächst einige Beispiele für folgenkompakte Mengen. Im folgenden seien $\mathbb{R}, \mathbb{C}, \mathbb{R}^k$ und \mathbb{C}^k immer mit der Standardmetrik versehen.

Satz 2.30 *Jede beschränkte und abgeschlossene Menge in \mathbb{R} ist folgenkompakt.*

Beweis. Sei $A \subset \mathbb{R}$ eine abgeschlossene und beschränkte Menge und (a_n) eine Folge in A . Dann ist (a_n) ebenfalls beschränkt. Nach dem Satz von Bolzano-Weierstraß existiert eine konvergente Teilfolge (a_{n_k}) von (a_n) . Der Grenzwert x dieser Teilfolge liegt in X . Nun gilt aber $x = \lim_{k \rightarrow \infty} a_{n_k} \in cl(A)$ (siehe Satz 2.13). Da A abgeschlossen ist, ist $A = cl(A)$, also liegt x in A . Die Menge A ist demnach folgenkompakt. \square

Beispiel: Die abgeschlossenen Intervalle $[a, b] \subset \mathbb{R}$ sind folgenkompakt.

Satz 2.31 *Sei (X, d) das Produkt der metrischen Räume $(X_1, d_1), \dots, (X_k, d_k)$ und seien $A_j \subset X_j$ folgenkompakte Mengen in (X_j, d_j) , $j = 1, \dots, k$. Dann ist die Menge $A := A_1 \times A_2 \times \dots \times A_k$ folgenkompakt in (X, d) .*

Beweis. Den Beweis wird analog zum Beweis von Satz 2.27 geführt. Wir überlassen ihn dem Leser als Übungsaufgabe. \square

Beispiel: Die Quader $W := [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_k, b_k] \subset \mathbb{R}^k$ sind folgenkompakte Teilmengen des metrischen Raumes \mathbb{R}^k .

Satz 2.32 *Sei (X, d) ein metrischer Raum und $B \subset X$ folgenkompakt. Dann ist jede abgeschlossene Teilmenge $A \subset B$ ebenfalls folgenkompakt.*

Beweis. Sei (a_n) eine beliebige Folge in A . Dann ist (a_n) auch Folge in B und besitzt, da B folgenkompakt ist, eine in B konvergente Teilfolge (a_{n_j}) . Sei $b = \lim_{j \rightarrow \infty} a_{n_j} \in cl(A)$. Da A abgeschlossen ist, ist $cl(A) = A$ und somit $b \in A$. Also enthält (a_n) eine in A konvergente Teilfolge. Damit ist A folgenkompakt. \square

Wir erhalten daraus folgende Verallgemeinerung von Satz 2.30:

Satz 2.33 *Jede beschränkte und abgeschlossene Teilmenge von \mathbb{C} , \mathbb{C}^k und \mathbb{R}^k ist folgenkompakt.*

Beweis. Da \mathbb{C}^k isometrisch zu \mathbb{R}^{2k} ist, genügt es, die Behauptung für die reellen Vektorräume \mathbb{R}^k zu beweisen. Sei A eine beschränkte und abgeschlossene Teilmenge von \mathbb{R}^k . Da A beschränkt ist, gibt es einen Quader $W \subset \mathbb{R}^k$, der A enthält. Da W folgenkompakt und A abgeschlossen ist, ist A auch folgenkompakt. \square

Beispiele: 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 folgenkompakt.

Es gibt metrische Räume mit abgeschlossenen und beschränkten Teilmengen, die *nicht* folgenkompakt 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 folgenkompakt: 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.

Als nächstes beschäftigen wir uns mit speziellen Eigenschaften folgenkompakter Mengen. Dazu zunächst folgende Definition.

Definition 2.26. Sei (X, d) ein metrischer Raum und $A \subset X$. Die Menge A heißt *total beschränkt*, wenn es zu jedem $\varepsilon > 0$ endlich viele Punkte $a_1, a_2, \dots, a_k \in A$ gibt, so dass $A \subset \bigcup_{i=1}^k K(a_i, \varepsilon)$. Ist die gesamte Menge X total beschränkt, so sagt man, dass der metrische Raum (X, d) *total beschränkt* ist.

Eine total beschränkte Menge ist auch beschränkt, denn für endlich viele Kugeln gilt

$$K(a_1, \varepsilon) \cup K(a_2, \varepsilon) \cup \dots \cup K(a_k, \varepsilon) \subset K(a_1, \max\{d(a_1, a_j) \mid j = 2, \dots, k\} + \varepsilon).$$

Satz 2.34 Jede folgenkompakte Teilmenge eines metrischen Raumes ist abgeschlossen, total beschränkt und insbesondere auch beschränkt.

Beweis. Sei (X, d) ein metrischer Raum und $A \subset X$ folgenkompakt.

(1) Wir zeigen, dass A abgeschlossen ist. Sei $x \in cl(A)$. Dann gibt es eine Folge (a_n) in A mit $a_n \rightarrow x$. Da A folgenkompakt ist, besitzt (a_n) eine in A konvergente Teilfolge (a_{n_k}) . Dann gilt aber $\lim_{k \rightarrow \infty} a_{n_k} = \lim_{n \rightarrow \infty} a_n = x \in A$. Folglich gilt $cl(A) \subset A$, d.h. A ist abgeschlossen.

(2) Wir zeigen, dass A total beschränkt ist. Angenommen A wäre nicht total beschränkt. Dann existiert ein $\varepsilon_0 > 0$, so dass für alle $k \in \mathbb{N}$ und für k beliebige Punkte $a_1, \dots, a_k \in A$

$$A \not\subset K(a_1, \varepsilon_0) \cup \dots \cup K(a_k, \varepsilon_0).$$

Wir betrachten zunächst $k = 1$ und einen Punkt $a_1 \in A$. Dann ist $A \not\subset K(a_1, \varepsilon_0)$. Wir wählen einen Punkt $a_2 \in A \setminus K(a_1, \varepsilon_0)$. Dann gilt $d(a_1, a_2) \geq \varepsilon_0 > 0$. Da $A \not\subset K(a_1, \varepsilon_0) \cup K(a_2, \varepsilon_0)$, existiert ein $a_3 \in A \setminus (K(a_1, \varepsilon_0) \cup K(a_2, \varepsilon_0))$. Also gilt $d(a_3, a_1) \geq \varepsilon_0 > 0$ und $d(a_3, a_2) \geq \varepsilon_0 > 0$. Wir führen dieses Verfahren fort und erhalten eine Folge von Punkten a_1, a_2, a_3, \dots in A mit $d(a_i, a_j) \geq \varepsilon_0 > 0$ für alle $i \neq j$. Diese Folge kann aber keine konvergente Teilfolge enthalten. Dies widerspricht der Folgenkompaktheit von A . \square

Aus den Sätzen 2.33 und 2.34 erhalten wir eine spezielle Eigenschaft folgenkompakter Mengen in den metrischen Räumen \mathbb{R}^k und \mathbb{C}^k .

Satz 2.35 Eine Teilmenge in \mathbb{R}^k und in \mathbb{C}^k ist genau dann folgenkompakt, wenn sie beschränkt und abgeschlossen ist.

Wie wir oben an einem Beispiel gesehen hatten, ist diese Charakterisierung folgenkompakter Mengen in allgemeinen metrischen Räumen nicht möglich. Man hat aber für *vollständige* metrische Räume das folgende Kriterium für Folgenkompaktheit.

Satz 2.36 Sei (X, d) ein vollständiger metrischer Raum. Eine Teilmenge in (X, d) ist genau dann folgenkompakt, wenn sie abgeschlossen und total beschränkt ist.

Beweis. \implies ist die Aussage von Satz 2.34.

\impliedby findet man z.B. im Buch von Dieudonné: *Grundzüge der modernen Analysis*, Bd. 1. □

Nachdem wir bisher folgenkompakte Teilmengen betrachtet haben, die durch spezielle Eigenschaften von Folgen definiert wurden, betrachten wir als nächstes einen Kompaktheitsbegriff, der mit Hilfe von offenen Mengen definiert wird.

Definition 2.27. Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge. Eine Familie $\mathcal{U} = \{U_i\}_{i \in I}$ von Teilmengen von X heißt offene Überdeckung von A , falls die folgenden beiden Eigenschaften gelten:

- (1) $U_i \subset X$ ist offen für alle $i \in I$.
- (2) $A \subset \bigcup_{i \in I} U_i$.

Eine Teilmenge $\hat{\mathcal{U}} \subset \mathcal{U}$ der Überdeckung \mathcal{U} heißt Teilüberdeckung, wenn $A \subset \bigcup_{U \in \hat{\mathcal{U}}} U$.

Eine Teilüberdeckung heißt endlich, wenn sie endlich viele Elemente enthält.

Definition 2.28. Sei (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt kompakt, wenn man aus jeder offenen Überdeckung von A eine endliche Teilüberdeckung auswählen kann. Ist die gesamte Menge X kompakt, so sagt man, dass der metrische Raum (X, d) kompakt ist.

Es gilt wiederum, dass eine Teilmenge A eines metrischen Raumes (X, d) genau dann kompakt ist, wenn der metrische Raum $(A, d|_{A \times A})$ kompakt ist.

Im Folgenden beweisen wir, dass in metrischen Räumen beide Kompaktheitsbegriffe äquivalent sind.

Satz 2.37 Eine Teilmenge eines metrischen Raumes ist genau dann kompakt, wenn sie folgenkompakt ist.

Beweis. Sei (X, d) ein metrischer Raum und $A \subset X$ eine Teilmenge.

- (1) A kompakt $\implies A$ folgenkompakt:

Sei A kompakt. Angenommen, A wäre nicht folgenkompakt. Dann gibt es eine Folge (a_n) in A , die keine in A konvergente Teilfolge besitzt. Zu jedem $x \in A$ finden wir dann eine Kugel $K(x, \varepsilon(x))$, die höchstens endlich viele Folgenglieder von (a_n) enthält. Dann ist $\mathcal{U} := \{K(x, \varepsilon(x))\}_{x \in A}$ eine offene Überdeckung von A . Da A kompakt ist, enthält sie eine endliche Teilüberdeckung, d.h. es gilt

$$A \subset K(x_1, \varepsilon(x_1)) \cup \dots \cup K(x_k, \varepsilon(x_k)).$$

Damit könnte A aber nur endlich viele Folgenglieder von (a_n) enthalten. Dies ist ein Widerspruch zur Wahl von (a_n) . Also war unsere Annahme falsch und A ist folgenkompakt.

(2) A folgenkompakt $\implies A$ kompakt:

Sei A folgenkompakt und $\mathcal{U} := \{U_\alpha\}_{\alpha \in A}$ eine beliebige offene Überdeckung von A . Wir müssen zeigen, dass \mathcal{U} eine endliche Teilüberdeckung besitzt. Angenommen, dies wäre nicht so. Nach Satz 2.34 ist A total beschränkt. Zu jedem $n \in \mathbb{N}$ gibt es also endlich viele Kugeln mit Radius $\frac{1}{n}$ und Mittelpunkten in A , die A überdecken. Da \mathcal{U} keine endliche Teilüberdeckung besitzt, gibt es zu jedem $n \in \mathbb{N}$ eine dieser Kugeln, nennen wir ihren Mittelpunkt a_n , so dass $A \cap K(a_n, \frac{1}{n})$ nicht von endlich vielen Mengen aus \mathcal{U} überdeckt wird. Wir betrachten nun die Folge (a_n) in A . Da A folgenkompakt ist, hat (a_n) eine in A konvergente Teilfolge (a_{n_k}) . Sei $x := \lim_{k \rightarrow \infty} a_{n_k} \in A$. Da \mathcal{U} die Menge A überdeckt, gibt es eine Menge $U \in \mathcal{U}$, die x enthält. Da U offen ist, finden wir ein $\varepsilon > 0$ mit $K(x, \varepsilon) \subset U$. Da x Grenzwert der Teilfolge (a_{n_k}) ist, existiert ein $k_0 \in \mathbb{N}$, so dass $a_{n_k} \in K(x, \frac{\varepsilon}{2})$ für alle $k \geq k_0$. Wir wählen nun einen Index n_l der Teilfolge mit $l \geq k_0$ und $\frac{1}{n_l} < \frac{\varepsilon}{2}$. Dann gilt für jeden Punkt $y \in K(a_{n_l}, \frac{1}{n_l})$

$$d(y, x) \leq d(y, a_{n_l}) + d(a_{n_l}, x) \leq \frac{1}{n_l} + \frac{\varepsilon}{2} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Daraus folgt

$$K(a_{n_l}, \frac{1}{n_l}) \subset K(x, \varepsilon) \subset U \in \mathcal{U},$$

d.h. wir können $K(a_{n_l}, \frac{1}{n_l}) \cap A$ sogar durch eine einzige Menge von \mathcal{U} überdecken. Dies widerspricht der Wahl von $K(a_{n_l}, \frac{1}{n_l})$. Somit war unsere Annahme falsch, d.h. aus jeder offenen Überdeckung \mathcal{U} von A kann man eine endliche Teilüberdeckung auswählen. A ist also kompakt. \square

Wegen der gerade bewiesenen Äquivalenz von Folgenkompaktheit und Kompaktheit können wir die uns bereits bekannten Eigenschaften folgenkompakter Mengen auf kompakte Mengen übertragen.

Folgerung 2.4 *Kompakte Teilmengen metrischer Räume haben die folgenden Eigenschaften:*

1. *Jede kompakte Menge ist abgeschlossen.*
2. *Jede kompakte Menge ist total beschränkt, insbesondere beschränkt.*
3. *Jede abgeschlossene Teilmenge einer kompakten Menge ist selbst kompakt.*
4. *Das Produkt kompakter Mengen ist kompakt.*
5. *Eine Teilmenge im metrischen Raum \mathbb{R}^n ist genau dann kompakt, wenn sie beschränkt und abgeschlossen ist.*
6. *Eine Teilmenge in einem vollständigen metrischen Raum ist genau dann kompakt, wenn sie abgeschlossen und total beschränkt ist.*

Weitere Eigenschaften kompakter Mengen findet man in den Übungsaufgaben. Einen Spezialfall der 5. Aussage dieser Folgerung findet man häufig unter dem Namen *Überdeckungssatz von Heine/Borel*.

Bemerkung: Wenn man sich den Satz 2.37 ansieht, fragt man sich natürlich, warum man denn überhaupt zwei verschiedene Kompaktheitsbegriffe definiert, wenn diese dann doch übereinstimmen. Ein Grund dafür ist, dass man die kompakten Mengen dadurch auf zwei völlig verschiedene Weisen charakterisieren kann (einmal durch Eigenschaften von Folgen, zum anderen durch Eigenschaften offener Überdeckungen). Beide Varianten haben zum Nachweis von Kompaktheit in verschiedenen Situationen jeweils Vorteile.

Der tieferliegende Grund ist aber der folgende: Die metrischen Räume sind eine spezielle Klasse der sogenannten topologischen Räume, die wir hier kurz ergänzend einführen wollen:

Sei X eine nichtleere Menge. Eine *Topologie auf X* ist eine Familie \mathcal{T} von Teilmengen von X , die folgende Eigenschaften hat:

1. $X, \emptyset \in \mathcal{T}$.
2. Die Vereinigung beliebig vieler Teilmengen aus \mathcal{T} ist ebenfalls in \mathcal{T} .
3. Der Durchschnitt endlich vieler Teilmengen aus \mathcal{T} ist ebenfalls in \mathcal{T} .

Das Paar (X, \mathcal{T}) nennt man dann *topologischen Raum*. Die Elemente in \mathcal{T} heißen die *offenen* Mengen des topologischen Raumes (X, \mathcal{T}) .

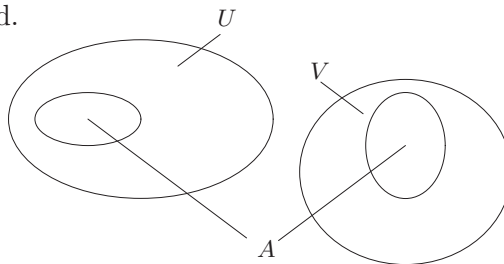
Ist (X, d) ein metrischer Raum, so bilden die offenen Mengen des metrischen Raumes nach Satz 2.4 eine Topologie auf X . Jeder metrische Raum ist folglich ein spezieller topologischer Raum. In topologischen Räumen kann man kompakte und folgenkompakte Mengen auf die gleiche Weise definieren, wie im metrischen Raum. Man kann dann zeigen, dass es sowohl topologische Räume gibt mit kompakten Mengen, die nicht folgenkompakt sind als auch topologische Räume mit folgenkompakten Mengen, die nicht kompakt sind !! Für topologische Räume, die zu den Grundstrukturen der Analysis gehören, fallen die beiden Kompaktheitsbegriffe also nicht mehr zusammen.

2.7 Zusammenhängende Teilmengen und Zusammenhangskomponenten eines metrischen Raumes

Definition 2.29. 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 $A \cap V \neq \emptyset$.

Ist die Menge X selbst zusammenhängend, so nennt man den metrischen Raum (X, d) *zusammenhängend*.

Beispiel: 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.



Bemerkung: Nach Definition ist ein metrischer Raum (X, d) zusammenhängend, wenn er sich nicht in die disjunkte Vereinigung zweier offener nichtleerer Mengen zerlegt. Man kann wiederum zeigen, dass eine Teilmenge A eines metrischen Raumes (X, d) genau dann zusammenhängend ist, wenn der metrische Raum $(A, d|_{A \times A})$ zusammenhängend ist.

Wir beschreiben nun die zusammenhängenden Mengen im metrischen Raum \mathbb{R} mit der Standardmetrik $d(x, y) = |x - y|$.

Satz 2.38 *Eine Teilmenge $A \subset \mathbb{R}$ ist genau dann zusammenhängend, wenn sie mit je zwei Punkten $a, b \in A$ auch das Intervall $[a, b]$ vollständig enthält. Die einzigen zusammenhängenden Teilmengen in \mathbb{R} sind folglich \mathbb{R} , $(-\infty, b)$, $(-\infty, b]$, (a, ∞) , $[a, \infty)$, (a, b) , $[a, b]$, $(a, b]$, $[a, b)$ und $\{a\}$.*

Beweis. (\implies) Sei $A \subset \mathbb{R}$ zusammenhängend. Angenommen es existieren $a, b \in A$ mit $[a, b] \not\subset A$. Dann gibt es 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.

(\impliedby) Sei $A \subset \mathbb{R}$ eine Teilmenge, die mit 2 Punkten $a, b \in A$, $a \leq b$, auch das Intervall $[a, b]$ enthält. Wir nehmen an, A sei nicht zusammenhängend. Dann existieren offene disjunkte Teilmengen $U, V \subset \mathbb{R}$ mit $A \subset U \cup V$, $A \cap U \neq \emptyset$ und $A \cap V \neq \emptyset$. Wir wählen $a \in A \cap U$ und $b \in A \cap V$. OBdA gelte $a < b$. Wir betrachten $z := \sup(U \cap [a, b])$. Dann gilt

$$\begin{aligned} z \in cl(U \cap [a, b]) &\subset cl(\mathbb{R} \setminus V) = \mathbb{R} \setminus V \quad (\text{da } V \text{ offen ist}), \\ z \in cl(U \cap [a, b]) &\subset cl([a, b]) = [a, b] \subset A. \end{aligned}$$

Da $U \cup V$ die Menge A überdeckt, liegt z in U . Da U offen ist, existiert ein $\varepsilon > 0$ mit $(z - \varepsilon, z + \varepsilon) \subset U$. Also gilt $(z - \varepsilon, z + \varepsilon) \cap [a, b] \subset U \cap [a, b]$. Es folgt $\min\{z + \varepsilon, b\} = \sup((z - \varepsilon, z + \varepsilon) \cap [a, b]) \leq \sup(U \cap [a, b]) = z$. Dann muss $b \leq z$ gelten. Da andererseits $z \in [a, b]$, folgt $b = z$ und somit $z \in V$. Da U und V disjunkt sind und $z \in U$, kann das nicht gelten. Damit war unsere Annahme, dass A nicht zusammenhängend ist, falsch. \square

Bemerkung: Der analoge Beweis liefert, dass jede Verbindungsstecke zwischen zwei Punkten im metrischen Raum \mathbb{R}^n zusammenhängend ist.

Weitere zusammenhängende Teilmengen erhält man durch die folgenden beiden Sätze.

Satz 2.39 *Sei (X, d) ein metrischer Raum und $A \subset X$ eine zusammenhängende Teilmenge. Gilt $A \subset B \subset cl(A)$, so ist B ebenfalls zusammenhängend.*

Inbesondere ist der Abschluss einer zusammenhängenden Menge selbst zusammenhängend.

Beweis. Angenommen B wäre nicht zusammenhängend. Dann existieren offene, disjunkte Mengen $U, V \subset X$ mit $B \subset U \cup V$, $B \cap U \neq \emptyset$ und $B \cap V \neq \emptyset$. Wegen $A \subset B$ gilt $A \subset U \cup V$. Da A zusammenhängend ist, muß $U \cap A = \emptyset$ oder $V \cap A = \emptyset$ gelten. Sei oBdA $U \cap A = \emptyset$. Dann gilt $A \subset X \setminus U$ und somit, da U offen ist, $B \subset cl(A) \subset cl(X \setminus U) = X \setminus U$. Dann wäre

$B \cap U = \emptyset$ im Gegensatz zur Wahl von U , d.h. wir erhalten einen Widerspruch. Folglich war unsere Annahme, dass B nicht zusammenhängend ist, falsch. \square

Satz 2.40 Sei (X, d) ein metrischer Raum und $A_i, i \in I$, beliebig viele zusammenhängende Teilmengen mit nichtleerem Durchschnitt $B := \bigcap_{i \in I} A_i \neq \emptyset$.

Dann ist auch die Vereinigung $A := \bigcup_{i \in I} A_i$ zusammenhängend.

Beweis. Wir nehmen an, A wäre nicht zusammenhängend. Dann existieren offene, disjunkte Mengen $U, V \subset X$ mit $A \subset U \cup V$, $A \cap U \neq \emptyset$ und $A \cap V \neq \emptyset$. Für $i \in I$ gilt dann auch $A_i \subset U \cup V$. Da A_i zusammenhängend ist, muss entweder $A_i \cap U = \emptyset$ oder $A_i \cap V = \emptyset$ gelten. Wir setzen

$$I_1 := \{i \in I \mid A_i \subset V\} \quad \text{und} \quad I_2 := \{i \in I \mid A_i \subset U\}.$$

Da $A \cap U \neq \emptyset$ und $A \cap V \neq \emptyset$ gilt $I_1 \neq \emptyset \neq I_2$. Wir erhalten

$$A = \bigcup_{i \in I} A_i = \left(\bigcup_{i \in I_1} A_i \right) \cup \left(\bigcup_{i \in I_2} A_i \right)$$

und

$$\left(\bigcup_{i \in I_1} A_i \right) \cap \left(\bigcup_{i \in I_2} A_i \right) \subset V \cap U = \emptyset.$$

Da für alle $i \in I$ gilt $\emptyset \neq B := \bigcap_{k \in I} A_k \subset A_i$, folgt

$$B \subset \left(\bigcup_{i \in I_1} A_i \right) \cap \left(\bigcup_{i \in I_2} A_i \right) = \emptyset.$$

Dies ist ein Widerspruch. Folglich war die Annahme falsch und A ist zusammenhängend. \square

Wir zeigen abschließend, dass sich jeder metrische Raum in disjunkte zusammenhängende Teilmengen zerlegt.

Definition 2.30. Sei (X, d) ein metrischer Raum und $x \in X$. Die Menge

$$\mathcal{C}(x) := \bigcup_{\substack{A \text{ zush.} \\ x \in A}} A$$

heißt die durch x bestimmte Zusammenhangskomponente von X .

Bemerkung:

(1) Nach Satz 2.40 ist $\mathcal{C}(x)$ zusammenhängend. Somit ist $\mathcal{C}(x)$ die größte zusammenhängende Teilmenge von X , die x enthält.

(2) Jede Zusammenhangskomponente von (X, d) ist abgeschlossen.

Nach Satz 2.39 ist nämlich der Abschluss $cl(\mathcal{C}(x))$ der Zusammenhangskomponente $\mathcal{C}(x)$ ebenfalls zusammenhängend. Folglich gilt $cl(\mathcal{C}(x)) \subset \mathcal{C}(x)$, also $cl(\mathcal{C}(x)) = \mathcal{C}(x)$. Somit ist

$\mathcal{C}(x)$ abgeschlossen.

(3) Sind $x, y \in X$, so gilt entweder $\mathcal{C}(x) = \mathcal{C}(y)$ oder $\mathcal{C}(x) \cap \mathcal{C}(y) = \emptyset$.

Wir betrachten den Fall, dass $\mathcal{C}(x) \cap \mathcal{C}(y) \neq \emptyset$. Dann ist $\mathcal{C}(x) \cup \mathcal{C}(y)$ zusammenhängend (nach Satz 2.40) und enthält x und y . Folglich gilt $\mathcal{C}(x) \cup \mathcal{C}(y) \subset \mathcal{C}(x)$ und $\mathcal{C}(x) \cup \mathcal{C}(y) \subset \mathcal{C}(y)$ und somit $\mathcal{C}(x) = \mathcal{C}(y) = \mathcal{C}(x) \cup \mathcal{C}(y)$.

Damit haben wir den folgenden Satz bewiesen:

Satz 2.41 *Sei (X, d) ein metrischer Raum. Dann zerlegt sich X in disjunkte Zusammenhangskomponenten, d.h.*

$$X = \bigcup_{i \in I} C_i,$$

wobei die Mengen C_i zusammenhängend, abgeschlossen und paarweise disjunkt sind, und $C_i = \mathcal{C}(x_i)$ für jeweils einen Punkt $x_i \in X$ gilt.

2.8 Banachräume und Hilberträume

In diesem Abschnitt betrachten wir spezielle vollständige metrische Räume, deren Basismenge Vektorräume sind und deren Metrik durch eine Norm bzw. durch ein Skalarprodukt gegeben ist.

Definition 2.31. *Ein Vektorraum über dem Körper \mathbb{K} (oder ein \mathbb{K} -Vektorraum) ist ein Tripel $[V, +, \cdot]$, wobei V eine nichtleere Menge und $+$ und \cdot zwei Operationen*

$$\begin{aligned} + : V \times V &\longrightarrow V && \text{(Addition)} \\ (v, w) &\longmapsto v + w \end{aligned}$$

$$\begin{aligned} \cdot : \mathbb{K} \times V &\longrightarrow V && \text{(Multiplikation mit Skalaren)} \\ (\lambda, v) &\longmapsto \lambda \cdot v \end{aligned}$$

mit den folgenden Eigenschaften sind:

1. $[V, +]$ ist eine abelsche Gruppe.
2. Es gelten die Distributivgesetze:
 $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$ und $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$
3. $\lambda \cdot (\mu \cdot v) = (\lambda \cdot \mu) \cdot v$ und $1 \cdot v = v$,

wobei $\lambda, \mu \in \mathbb{K}$ und $v, w \in V$.

Wir werden Vektorräume auch kurz mit V bezeichnen, wenn die Operationen $+$ und \cdot klar sind. Ist \mathbb{K} der Körper der reellen Zahlen, so heißt V *reeller Vektorraum*, ist \mathbb{K} der Körper der komplexen Zahlen, so heißt V *komplexer Vektorraum*. Die algebraische Theorie von Vektorräumen wird in der Vorlesung über lineare Algebra behandelt. Insbesondere wird dort erklärt werden, was die Dimension eines Vektorraumes ist. In der Analysis-Vorlesung interessieren wir uns für topologische Eigenschaften von Vektorräumen, insbesondere zunächst für die Definition eines geeigneten Konvergenzbegriffes für Folgen in Vektorräumen. Dazu benutzt man Normen.

Definition 2.32. Sei V ein reeller oder komplexer Vektorraum. Eine Norm auf V ist eine Abbildung $\|\cdot\| : V \rightarrow \mathbb{R}$ mit den folgenden Eigenschaften:

1. $\|x\| \geq 0$ für alle $x \in V$ und $\|x\| = 0 \iff x = 0$.
2. $\|\alpha \cdot x\| = |\alpha| \cdot \|x\|$ für alle $\alpha \in \mathbb{K}$ und $x \in V$.
3. $\|x + y\| \leq \|x\| + \|y\|$ (Dreiecksungleichung).

Das Paar $(V, \|\cdot\|)$ heißt normierter Vektorraum.

Beispiele für normierte Vektorräume:

1. Die Vektorräume \mathbb{R}^n bzw. \mathbb{C}^n mit den folgenden Normen

$$\|x\|_p := \left(\sum_{j=1}^n |x_j|^p \right)^{\frac{1}{p}}, \quad p \in \mathbb{N}, \quad \text{und} \quad \|x\|_\infty := \max_{j=1, \dots, n} |x_j|,$$

wobei $x = (x_1, \dots, x_n)$. Die Norm $\|\cdot\|_2$ auf dem \mathbb{R}^n bzw. \mathbb{C}^n ist die in Kapitel 1.4 betrachtete Euklidische Norm.

2. Sei X eine nichtleere Menge und bezeichne $B(X)$ den Vektorraum aller beschränkten reellwertigen Funktionen

$$B(X) := \{f : X \rightarrow \mathbb{R} \mid f(X) \subset \mathbb{R} \text{ beschränkt}\}.$$

Die Operationen $+$ und \cdot sind hierbei auf folgende Weise definiert:

$$\begin{aligned} B(X) \times B(X) &\longrightarrow B(X) \\ (f_1, f_2) &\longmapsto f_1 + f_2 \quad \text{mit} \quad (f_1 + f_2)(x) := f_1(x) + f_2(x) \\ \mathbb{R} \times B(X) &\longrightarrow B(X) \\ (\lambda, f) &\longmapsto \lambda \cdot f \quad \text{mit} \quad (\lambda \cdot f)(x) := \lambda \cdot f(x). \end{aligned}$$

Dann ist durch

$$\|f\|_\infty := \sup_{x \in X} |f(x)|, \quad f \in B(X)$$

eine Norm auf $B(X)$ gegeben.

3. Sei $\mathcal{BF}(\mathbb{R}^k)$ der Vektorraum aller beschränkten Folgen (x_n) von Vektoren in \mathbb{R}^k mit den Operationen

$$\begin{aligned} (x_n) + (y_n) &:= (x_n + y_n) \\ \lambda \cdot (x_n) &:= (\lambda \cdot x_n), \end{aligned}$$

wobei $\lambda \in \mathbb{R}$, $(x_n), (y_n) \in \mathcal{BF}(\mathbb{R}^k)$. Dann ist durch

$$\|(x_n)\|_\infty := \sup\{\|x_n\|_2 \mid n \in \mathbb{N}\}, \quad (x_n) \in \mathcal{BF}(\mathbb{R}^k)$$

eine Norm auf $\mathcal{BF}(\mathbb{R}^k)$ gegeben.

Auf jedem normierten Vektorraum wird durch die Norm auf kanonische Weise eine Metrik induziert, die wir im Folgenden immer auf normierten Vektorräumen festlegen werden:

Satz 2.42 Sei $(V, \|\cdot\|)$ ein normierter Raum. Dann ist die Abbildung $d : V \times V \rightarrow \mathbb{R}$,

$$d(x, y) := \|x - y\|, \quad x, y \in V,$$

eine Metrik auf V .

Beweis. Der Beweis wird genauso wie für die Euklidische Norm in Kapitel 1.4., Satz 1.23 geführt. Dort haben wir nur die Norm-Eigenschaften, aber nicht die konkrete Gestalt der Norm benutzt. \square

Definition 2.33. Ein normierter Vektorraum $(V, \|\cdot\|)$ heißt Banachraum, wenn er (als metrischer Raum) vollständig ist.

Beispiele für Banachräume:

1. \mathbb{R}^n und \mathbb{C}^n mit der Euklidischen Norm $\|\cdot\|_2$ sind Banachräume (siehe Kapitel 2.5).
2. Die Räume der beschränkten reellwertigen Funktionen $B(X)$ und der beschränkten Folgen $\mathcal{BF}(\mathbb{R}^n)$ sind mit der Supremum-Norm $\|\cdot\|_\infty$ Banachräume (Übungsaufgabe).

Wir werden nun zeigen, dass die Vektorräume \mathbb{R}^n und \mathbb{C}^n mit jeder Norm Banachräume sind. Dazu führen wir die folgende Äquivalenzrelation für Normen ein:

Definition 2.34. Sei V ein reeller oder komplexer Vektorraum. Zwei Normen $\|\cdot\|_1$ und $\|\cdot\|_2$ auf V heißen äquivalent, wenn es positive Konstanten $a, b \in \mathbb{R}^+$ gibt, so dass gilt

$$a \cdot \|x\|_1 \leq \|x\|_2 \leq b \cdot \|x\|_1 \quad \forall x \in V. \quad (*)$$

Satz 2.43 Seien $\|\cdot\|_1$ und $\|\cdot\|_2$ zwei äquivalente Normen auf dem Vektorraum V . Dann stimmen alle topologischen Eigenschaften von $(V, \|\cdot\|_1)$ und $(V, \|\cdot\|_2)$ überein. Insbesondere gilt für Folgen (x_n) in V und für Teilmengen $A \subset V$:

1. $(x_n) \rightarrow x \in V$ bzgl. $\|\cdot\|_1$ gdw. $(x_n) \rightarrow x \in V$ bzgl. $\|\cdot\|_2$.
2. (x_n) ist Cauchy-Folge bzgl. $\|\cdot\|_1$ gdw. (x_n) ist Cauchy-Folge bzgl. $\|\cdot\|_2$.
3. A ist offen bzgl. $\|\cdot\|_1$ gdw. A ist offen bzgl. $\|\cdot\|_2$.
4. A ist abgeschlossen bzgl. $\|\cdot\|_1$ gdw. A ist abgeschlossen bzgl. $\|\cdot\|_2$.
5. A ist kompakt bzgl. $\|\cdot\|_1$ gdw. A ist kompakt bzgl. $\|\cdot\|_2$.
6. A ist zusammenhängend bzgl. $\|\cdot\|_1$ gdw. A ist zusammenhängend bzgl. $\|\cdot\|_2$.

Beweis. Aus $(*)$ folgen für die Kugeln $K_i(x, \varepsilon)$ bzgl. der Norm $\|\cdot\|_i$ die folgenden Beziehungen

$$K_1(x, \varepsilon) \subset K_2(x, b \cdot \varepsilon) \quad , \quad K_2(x, \varepsilon) \subset K_1(x, \frac{\varepsilon}{a}) \quad \forall \varepsilon > 0.$$

Da Konvergenz von Folgen, Cauchy-Folgen und offene Mengen mittels ε -Kugeln definiert sind, folgt die Äquivalenz dieser Eigenschaften für $\|\cdot\|_1$ und $\|\cdot\|_2$ aus diesen Beziehungen der Kugeln zueinander. Eine Menge ist abgeschlossen, wenn ihr Komplement offen ist. Die Äquivalenz der Offenheit liefert somit auch die Äquivalenz der Abgeschlossenheit. Kompakte und zusammenhängende Mengen sind ebenfalls durch offene Mengen definiert, deshalb sind auch diese Eigenschaften für beide Normen äquivalent. \square

Satz 2.44 *Alle Normen auf den Vektorräumen \mathbb{R}^n bzw. \mathbb{C}^n sind zueinander äquivalent.*

Beweis. Wir beweisen die Behauptung für den reellen Vektorraum \mathbb{R}^n . Der Beweis für \mathbb{C}^n wird analog geführt.

Sei $\|\cdot\|_2 : \mathbb{R}^n \rightarrow \mathbb{R}$ die Euklidische Norm auf dem \mathbb{R}^n und $N : \mathbb{R}^n \rightarrow \mathbb{R}$ eine beliebige Norm. Da die Äquivalenz von Normen eine Äquivalenzrelation ist, genügt es zu zeigen, dass die Normen $\|\cdot\|_2$ und N äquivalent sind. Wir müssen also positive reelle Konstanten a und b finden, so dass

$$a \cdot \|x\|_2 \leq N(x) \leq b \cdot \|x\|_2 \quad \text{für alle } x \in \mathbb{R}^n.$$

(1) Es existiert ein $b \in \mathbb{R}^+$ mit $N(x) \leq b \cdot \|x\|_2$ für alle $x \in \mathbb{R}^n$:

Jeder Vektor $x \in \mathbb{R}^n$ hat die Form $x = (x_1, \dots, x_n) = \sum_{j=1}^n x_j e_j$, wobei $e_j = (0, \dots, 0, 1, 0, \dots, 0)$ mit 1 an der j -Stelle. Dann gilt wegen der Normeigenschaften und der Cauchy–Schwarzschen Ungleichung (siehe Satz 1.22)

$$\begin{aligned} N(x) &= N\left(\sum_{j=1}^n x_j \cdot e_j\right) \leq \sum_{j=1}^n N(x_j \cdot e_j) = \sum_{j=1}^n |x_j| \cdot N(e_j) \\ &\leq \sqrt{\sum_{j=1}^n |x_j|^2} \cdot \underbrace{\sqrt{\sum_{j=1}^n N(e_j)^2}}_{=: b > 0} \\ &= b \cdot \|x\|_2. \end{aligned}$$

(2) Es existiert ein $a \in \mathbb{R}^+$ mit $a \cdot \|x\|_2 \leq N(x)$ für alle $x \in \mathbb{R}^n$:

Sei $x \in \mathbb{R}^n$ ein vom Nullvektor verschiedener Vektor. Dann gilt

$$N(x) = N\left(\|x\|_2 \cdot \frac{x}{\|x\|_2}\right) = \|x\|_2 \cdot N\left(\frac{x}{\|x\|_2}\right).$$

Wir setzen $a := \inf\{N(y) \mid \|y\|_2 = 1\} \geq 0$. Dann gilt $N(x) \geq a \cdot \|x\|_2$ für alle $x \in \mathbb{R}^n$. Es bleibt zu zeigen, dass $a > 0$.

Angenommen $a = 0$. Sei $S^{n-1} := \{y \in \mathbb{R}^n \mid \|y\|_2 = 1\}$ die Sphäre vom Radius 1 im \mathbb{R}^n . Nach Definition von a existiert eine Folge (x_n) in S^{n-1} mit $N(x_n) \rightarrow a = 0$. Da die Sphäre S^{n-1} in \mathbb{R}^n beschränkt und abgeschlossen bzgl. der Euklidischen Norm $\|\cdot\|_2$ ist, ist sie folgenkompakt bzgl. $\|\cdot\|_2$ (siehe Satz 2.33). Folglich existiert eine Teilfolge (x_{n_k}) von (x_n) , die gegen einen Vektor $y_0 \in S^{n-1}$ bzgl. $\|\cdot\|_2$ konvergiert. Wir erhalten dann aus der Dreiecksungleichung für N und der in (1) bewiesenen Ungleichung

$$0 \leq N(y_0) \leq N(y_0 - x_{n_k}) + N(x_{n_k}) \leq b \cdot \|y_0 - x_{n_k}\|_2 + N(x_{n_k}).$$

Bei k gegen $+\infty$ konvergiert die rechte Seite dieser Ungleichung gegen 0. Folglich ist $N(y_0) = 0$ und somit $y_0 = 0$. Dies widerspricht $y_0 \in S^{n-1}$. Somit ist $a > 0$. \square

Da \mathbb{R}^n und \mathbb{C}^n mit der Euklidischen Norm vollständig sind, folgt aus den Sätzen 2.43 und 2.44, dass die Vektorräume \mathbb{R}^n und \mathbb{C}^n mit jeder Norm Banachräume sind. In der Vorlesung

über lineare Algebra wird beweisen, dass jeder reelle Vektorraum der Dimension $n < +\infty$ isomorph zu \mathbb{R}^n und jeder komplexe Vektorraum der Dimension $n < +\infty$ isomorph zu \mathbb{C}^n ist. Dies liefert

Folgerung 2.5 *Jeder endlich-dimensionale, reelle oder komplexe normierte Vektorraum ist ein Banachraum.*

Für ∞ -dimensionale Vektorräume gilt Satz 2.44 und die daraus abgeleitete Folgerung nicht mehr. Im Laufe der Analysis-Vorlesung werden wir weitere vollständige und unvollständige unendlich-dimensionale normierte Vektorräume, insbesondere Vektorräume von Funktionen, kennenlernen. Solche Vektorräume spielen bei der Untersuchung partieller Differentialgleichungen eine wichtige Rolle.

Abschließend befassen wir uns mit dem Begriff des Skalarproduktes auf einem Vektorraum, den wir bereits in einem Spezialfall in Kapitel 1.4. betrachtet haben.

Definition 2.35. *Sei H ein reeller oder komplexer Vektorraum und bezeichne \mathbb{K} den dazugehörigen Körper der reellen bzw. komplexen Zahlen. Ein Skalarprodukt auf H ist eine Abbildung $\langle \cdot, \cdot \rangle : H \times H \rightarrow \mathbb{K}$ mit folgenden Eigenschaften:*

- (1) $\langle x, y \rangle = \overline{\langle y, x \rangle}$ für $x, y \in H$,
- (2) $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$ für $x, y, z \in H$,
- (3) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle = \langle x, \bar{\lambda} y \rangle$ für $\lambda \in \mathbb{K}$ und $x, y \in H$,
- (4) $\langle x, x \rangle \geq 0$ und $\langle x, x \rangle = 0 \Leftrightarrow x = 0$.

Ein Beispiel für ein Skalarprodukt auf dem Vektorraum \mathbb{C}^n kennen wir bereits aus Kapitel 1.4:

$$\langle z, w \rangle_{\mathbb{C}^n} := \sum_{j=1}^n z_j \cdot \bar{w}_j, \quad \text{wobei } z = (z_1, \dots, z_n), w = (w_1, \dots, w_n).$$

Satz 2.45 *Sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf dem Vektorraum H . Dann ist durch*

$$\|x\| := \sqrt{\langle x, x \rangle}, \quad x \in H,$$

eine Norm auf H und durch

$$d(x, y) := \sqrt{\langle x - y, x - y \rangle} = \|x - y\|, \quad x, y \in H,$$

eine Metrik auf H definiert. Des Weiteren gilt die Cauchy-Schwarzsche Ungleichung:

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\| \quad \text{für alle } x, y \in H,$$

wobei die Gleichheit genau dann angenommen wird, wenn x und y linear abhängig sind.

Beweis. Der Beweis wird wie in Kapitel 1.4., Satz 1.22 geführt. Dort haben wir dies für das spezielle Skalarprodukt $\langle \cdot, \cdot \rangle_{\mathbb{C}^n}$ auf \mathbb{C}^n bewiesen. Wir haben dabei aber nicht die spezielle Form des Skalarproduktes, sondern nur seine allgemeinen Eigenschaften benutzt. \square

Bemerkung: Ein Vektorraum $(H, \langle \cdot, \cdot \rangle)$ mit Skalarprodukt ist im Folgenden immer mit der von $\langle \cdot, \cdot \rangle$ erzeugten Norm $\|x\| = \sqrt{\langle x, x \rangle}$ und der dadurch induzierten Metrik versehen.

Ist eine Norm gegeben, so möchte man gern wissen, ob diese durch ein Skalarprodukt wie oben beschrieben definiert ist. Dies ist nicht immer der Fall. Ein Kriterium dafür liefert der Satz

Satz 2.46 *Sei $(V, \|\cdot\|)$ ein normierter Vektorraum. Es existiert genau dann ein Skalarprodukt $\langle \cdot, \cdot \rangle$ mit $\|x\|^2 = \langle x, x \rangle$ für alle $x \in V$, wenn für die Norm das Parallelogrammgesetz gilt:*

$$\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2) \quad \forall x, y \in V.$$

Beweis. Übungsaufgabe. □

Definition 2.36. *Ein Hilbertraum ist ein Vektorraum H mit Skalarprodukt $\langle \cdot, \cdot \rangle$, dessen zugehöriger normierter Raum $(H, \|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle})$ ein Banachraum ist.*

Beispiele für Hilberträume:

1. \mathbb{R}^n mit dem Skalarprodukt $\langle x, y \rangle_{\mathbb{R}^n} := \sum_{j=1}^n x_j \cdot y_j$
2. \mathbb{C}^n mit dem Skalarprodukt $\langle z, w \rangle_{\mathbb{C}^n} := \sum_{j=1}^n z_j \cdot \bar{w}_j$

Reihen in Banachräumen

Reihen sind spezielle Folgen in Vektorräumen. 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 Kapitel werden wir Reihen in Banachräumen behandeln und Kriterien für ihre Konvergenz kennenlernen.

3.1 Konvergente und divergente Reihen

Im gesamten 3. Kapitel bezeichnet E einen Banachraum über dem Körper der reellen oder der komplexen Zahlen mit der Norm $\|\cdot\|$. Als Spezialfall kann man sich an Stelle von E zum Beispiel die reellen Zahlen \mathbb{R} bzw. die komplexen Zahlen \mathbb{C} mit dem Betrag der Zahlen als Norm oder die Vektorräume \mathbb{R}^n bzw. \mathbb{C}^n mit der Euklidischen Norm

$$\|x\| := \left(\sum_{k=1}^n |x_k|^2 \right)^{\frac{1}{2}}, \quad x = (x_1, \dots, x_n),$$

(oder mit einer beliebigen anderen Norm) vorstellen.

In einem Vektorraum kann man jeder Folge von Vektoren eine neue Folge zuordnen: die Folge der Partialsummen. Sei $(x_k)_{k=1}^{\infty}$ eine Folge von Vektoren aus E . Wir bilden daraus eine neue Folge

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

$$s_1 := x_1$$

$$s_2 := x_1 + x_2$$

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

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

...

Definition 3.1. Die Folge $(s_n)_{n=1}^{\infty}$ heißt Reihe in E mit den Gliedern x_k . Man schreibt für diese Reihe symbolisch

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

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

Eine Reihe $\sum_{k=1}^{\infty} x_k$ in E heißt konvergent, falls die Folge der Partialsummen (s_n) in E 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 E 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 kon-

vergiert, wenn die Reihe $\sum_{k=k_0}^{\infty} x_k$ für ein beliebig gewähltes $k_0 \in \mathbb{N}$ konvergiert.

Sei speziell $E = \mathbb{R}$. Wenn $\lim_{n \rightarrow \infty} s_n = +\infty$ oder $\lim_{n \rightarrow \infty} s_n = -\infty$, 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 3.2. Eine Reihe $\sum_{k=1}^{\infty} x_k$ im Banachraum E heißt absolut-konvergent, wenn die Reihe der reellen Zahlen $\sum_{k=1}^{\infty} \|x_k\|$ in \mathbb{R} konvergiert.

Im nächsten Abschnitt behandeln wir wichtige Konvergenzkriterien für Reihen in Banachräumen.

3.2 Konvergenzkriterien für Reihen in Banachräumen

Satz 3.1 (Cauchy-Kriterium) Eine Reihe $\sum_{k=1}^{\infty} x_k$ in einem Banachraum E ist genau dann konvergent, wenn zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass

$$\|x_n + \dots + x_m\| < \varepsilon \quad \forall m \geq n \geq n_0 \quad (*)$$

gilt.

Beweis. Da der Banachraum E vollständig ist, konvergiert die Folge der Partialsummen $s_n = x_1 + \dots + x_n$ genau dann, wenn sie eine Cauchy-Folge ist. Dies ist nach Definition der Cauchy-Folge genau dann erfüllt, 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

Der folgende Satz gibt ein nützliches notwendiges Kriterium für die Konvergenz einer Reihe an:

Satz 3.2 *Ist eine Reihe $\sum_{k=1}^{\infty} x_k$ im Banachraum E konvergent, so ist die Folge der Reihenglieder (x_k) eine Nullfolge in E .*

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

Sei $E = \mathbb{R}$. 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

$$\begin{aligned} s_n &= 1 + \frac{1}{2} + \dots + \frac{1}{n} \\ &\geq 1 + \frac{1}{2} + \underbrace{\left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \dots + \frac{1}{8}\right) + \dots + \left(\frac{1}{2^{m-1} + 1} + \dots + \frac{1}{2^m}\right)}_{2^{m-1} \text{ Summanden}} \\ &\geq 1 + \underbrace{\frac{1}{2} + 2 \cdot \frac{1}{4} + 4 \cdot \frac{1}{8} + \dots + 2^{m-1} \cdot \frac{1}{2^m}}_{m \text{ Summanden}} \\ &\geq 1 + \frac{m}{2}. \end{aligned}$$

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 geometrische Reihe

Sei $E = \mathbb{C}$ und $z \in \mathbb{C}$ 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 1.13

$$s_n = 1 + z + \dots + z^n = \frac{1 - z^{n+1}}{1 - z}.$$

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 3.2). \square

Satz 3.3 *Ist eine Reihe $\sum_{k=1}^{\infty} x_k$ im Banachraum E absolut-konvergent, so ist sie auch konvergent und für die Werte der Reihen gilt*

$$\left\| \sum_{k=1}^{\infty} x_k \right\| \leq \sum_{k=1}^{\infty} \|x_k\|. \quad (**)$$

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 die Norm

$$\|x_n + \dots + x_m\| \leq \|x_n\| + \dots + \|x_m\|$$

gilt das Cauchy-Kriterium auch für die Reihe $\sum_{k=1}^{\infty} x_k$, die folglich ebenfalls konvergiert. Aus den Grenzwertsätzen (analog zu Satz 2.16) erhält man dann für $s_n = x_1 + \dots + x_n$

$$\left\| \lim_{n \rightarrow \infty} s_n \right\| = \lim_{n \rightarrow \infty} \|s_n\| \leq \lim_{n \rightarrow \infty} (\|x_1\| + \dots + \|x_n\|) = \sum_{k=1}^{\infty} \|x_k\|.$$

Dies zeigt die Abschätzung (**). \square

Reihen im Banachraum E kann man addieren und mit reellen bzw. komplexen Zahlen multiplizieren. Im folgenden bezeichnet \mathbb{K} den Körper der reellen Zahlen, falls E ein reeller Banachraum ist, bzw. den Körper der komplexen Zahlen, falls E ein komplexer Banachraum ist.

Satz 3.4 Seien (x_k) und (y_k) Folgen im Banachraum E und $\lambda, \mu \in \mathbb{K}$. Konvergieren die Reihen $\sum_{k=1}^{\infty} x_k$ und $\sum_{k=1}^{\infty} y_k$ gegen x bzw. y , so konvergiert die Reihe $\sum_{k=1}^{\infty} (\lambda x_k + \mu y_k)$ 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 den Grenzwertsätzen für Folgen in Vektorräumen (analog zu Satz 2.16). \square

Satz 3.5 (Majorantenkriterium)

Sei $\sum_{k=1}^{\infty} x_k$ eine Reihe im Banachraum E 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 $\sum_{k=1}^{\infty} x_k$ in E absolutkonvergent und für die Werte der Reihen gilt

$$\left\| \sum_{k=1}^{\infty} x_k \right\| \leq \sum_{k=1}^{\infty} \|x_k\| \leq \sum_{k=1}^{\infty} c_k. \quad (***)$$

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 $\sum_{k=1}^{\infty} x_k$ absolutkonvergent, also auch konvergent. Die Ungleichung (***) folgt wie im Beweis von Satz 3.3. \square

Beispiel 3: Sei $E = \mathbb{C}$. Wir betrachten eine Folge komplexer Zahlen (a_k) mit $|a_k| \leq 1$ für alle $k \in \mathbb{N}$ und $z \in \mathbb{C}$ mit $|z| < 1$.

Behauptung: Die Reihe $\sum_{k=0}^{\infty} a_k z^k$ ist absolutkonvergent und für ihren Wert gilt:

$$\left| \sum_{k=0}^{\infty} a_k z^k \right| \leq \frac{1}{1 - |z|}.$$

Beweis. Dies folgt aus dem Majorantenkriterium und der Konvergenz der geometrischen Reihe, da

$$|a_k z^k| = |a_k| \cdot |z|^k \leq |z|^k \quad \text{und} \quad \sum_{k=0}^{\infty} |z|^k = \frac{1}{1 - |z|}.$$

\square

Satz 3.6 (Wurzelkriterium)

Sei $\sum_{k=1}^{\infty} x_k$ eine Reihe im Banachraum E 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.
2. Ist $\alpha > 1$, so ist die Reihe $\sum_{k=1}^{\infty} x_k$ divergent.

Beweis. 1. Sei $\alpha = \limsup_{k \rightarrow \infty} \sqrt[k]{\|x_k\|} < 1$. Da α der größte Häufungspunkt der Folge $(\sqrt[k]{\|x_k\|})$ ist, sind höchstens endlich viele dieser Folgeglieder größer als $\frac{1+\alpha}{2}$. Es existiert folglich ein $k_0 \in \mathbb{N}$ so dass

$$\sqrt[k]{\|x_k\|} < \frac{1+\alpha}{2} < 1 \quad \forall k \geq k_0.$$

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

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 und $\sum_{k=1}^{\infty} x_k$ konvergiert nicht (Satz 3.2). \square

Satz 3.7 (Quotientenkriterium)

Sei $\sum_{k=1}^{\infty} x_k$ eine Reihe im Banachraum E , deren Glieder x_k alle vom Nullvektor 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.
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 $\frac{\alpha+1}{2} < 1$, konvergiert die geometrische Reihe $\sum_{j=0}^{\infty} (\frac{\alpha+1}{2})^j$. Aus dem Majorantenkriterium folgt dann, dass die Reihe $\sum_{k=1}^{\infty} x_k$ absolut 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 $k \geq k_0$ gilt. Also ist (x_k) keine Nullfolge. Nach Satz 3.2 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. $E = \mathbb{R}$ und 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 $\limsup_{k \rightarrow \infty} \sqrt[k]{x_k} = \frac{1}{2}$ gilt, während das Quotientenkriterium keine Aussage liefert, da $\liminf_{k \rightarrow \infty} \frac{x_{k+1}}{x_k} = 0$ und $\limsup_{k \rightarrow \infty} \frac{x_{k+1}}{x_k} = +\infty$ gilt.

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 $E = \mathbb{C}$, $z \in \mathbb{C}$ eine fixierte komplexe Zahl und $p \in \mathbb{Q}^+$ eine positive rationale Zahl. Die Reihe

$$\sum_{k=1}^{\infty} \frac{k^p}{z^k}$$

ist für $|z| > 1$ absolut-konvergent und für $|z| < 1$ divergent.

Beweis. Wir benutzen das Quotientenkriterium mit $x_k = \frac{k^p}{z^k}$.

$$\left| \frac{x_{k+1}}{x_k} \right| = \left| \frac{(k+1)^p \cdot z^k}{k^p \cdot z^{k+1}} \right| = \left(\frac{k+1}{k} \right)^p \cdot \frac{1}{|z|} \rightarrow \frac{1}{|z|}.$$

Es gilt also $\lim \left| \frac{x_{k+1}}{x_k} \right| = \limsup \left| \frac{x_{k+1}}{x_k} \right| = \liminf \left| \frac{x_{k+1}}{x_k} \right| = \frac{1}{|z|}$. Die Behauptung folgt dann aus dem Quotientenkriterium. \square

Beispiel 7: Sei $E = \mathbb{C}$ und $z \in \mathbb{C}$. Die Reihe

$$\sum_{k=1}^{\infty} \frac{z^k}{k}$$

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 $\lim_{k \rightarrow \infty} \sqrt[k]{k} = 1$ gilt, 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. \square

Wir interessieren uns natürlich auch dafür, was mit der Konvergenz für $z \in \mathbb{C}$ mit $|z| = 1$ passiert. Für $z = 1$ ist die obige Reihe gerade die harmonische Reihe, also divergent. Um die anderen Fälle mit $|z| = 1$ behandeln zu können, beweisen wir ein weiteres Kriterium.

Satz 3.8 (Abel–Dirichlet–Kriterium)

Sei (x_k) eine Folge im Banachraum E , deren Partialsummenfolge $(s_n = x_1 + x_2 + \dots + x_n)$ beschränkt ist. Sei weiterhin (a_k) eine monoton fallende Nullfolge positiver reeller Zahlen. Dann konvergiert die Reihe

$$\sum_{k=1}^{\infty} a_k x_k$$

im Banachraum E .

Beweis. Seien $s_n = x_1 + x_2 + \dots + x_n$ und $\sigma_n = a_1 x_1 + \dots + a_n x_n$. Nach Voraussetzung existiert eine reelle Zahl C , so dass $\|s_n\| \leq C$ für alle $n \in \mathbb{N}$. Für $m > n$ erhalten wir

$$\begin{aligned} \sigma_m - \sigma_n &= a_{n+1} x_{n+1} + a_{n+2} x_{n+2} + \dots + a_m x_m \\ &= a_{n+1}(s_{n+1} - s_n) + a_{n+2}(s_{n+2} - s_{n+1}) + \dots + a_m(s_m - s_{m-1}) \\ &= -a_{n+1} s_n + (a_{n+1} - a_{n+2}) s_{n+1} + \dots + (a_{m-1} - a_m) s_{m-1} + a_m s_m. \end{aligned}$$

Da nach Voraussetzung $a_k - a_{k+1} \geq 0$ gilt, folgt

$$\begin{aligned} \|\sigma_m - \sigma_n\| &\leq C \cdot \left(a_{n+1} + (a_{n+1} - a_{n+2}) + (a_{n+2} - a_{n+3}) + \dots \right. \\ &\quad \left. + (a_{m-1} - a_m) + a_m \right) \\ &= 2C \cdot a_{n+1}. \end{aligned}$$

Da (a_n) eine Nullfolge ist, existiert zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass $\|\sigma_m - \sigma_n\| < \varepsilon$ für alle $m > n \geq n_0$. Damit ist (σ_n) eine Cauchy-Folge und konvergiert im Banachraum E . \square

Beispiel 7 (Fortsetzung):

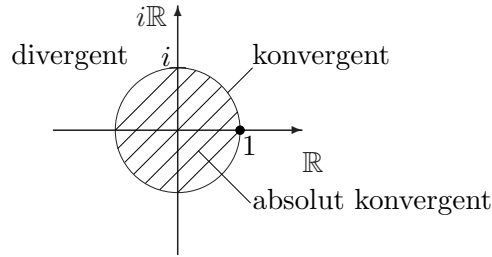
Die Reihe $\sum_{k=1}^{\infty} \frac{z^k}{k}$ ist konvergent für $z \in \mathbb{C}$ mit $|z| = 1$ und $z \neq 1$.

Beweis. Seien $a_k := \frac{1}{k}$, $x_k := z^k$ und $s_n := z + z^2 + \dots + z^n$. Es gilt

$$s_n = z + z^2 + \dots + z^n = z \left(\frac{1 - z^n}{1 - z} \right).$$

Da $|z| = 1$, folgt $|s_n| = |z| \cdot \frac{|1 - z^n|}{|1 - z|} \leq \frac{2}{|1 - z|}$. Somit ist die Folge der Partialsummen (s_n) beschränkt. Aus dem Abel-Dirichlet-Kriterium folgt dann, dass die Reihe $\sum_{k=1}^{\infty} \frac{z^k}{k}$ konvergiert. \square

Das folgende Bild zeigt das Konvergenzverhalten der Reihe $\sum_{k=1}^{\infty} \frac{z^k}{k}$.



Spezialfall ($z = -1$): Die alternierende harmonische Reihe

$$\sum_{k=1}^{\infty} (-1)^k \frac{1}{k} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} \pm \dots$$

ist konvergent. Wir werden später sehen, dass ihr Wert $-\ln 2$ ist.

Als nächstes betrachten wir Reihen reeller Zahlen, die sich so verhalten wie die alternierende harmonische Reihe.

Definition 3.3. 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 3.9 (Leibniz-Kriterium für alternierende Reihen)

Sei (b_k) eine monoton fallende Nullfolge positiver reeller Zahlen. Dann gilt:

1. Die alternierende Reihe $\sum_{k=1}^{\infty} (-1)^k b_k$ konvergiert.
2. Es gilt die folgende Fehlerabschätzung für den Wert s der Reihe $\sum_{k=1}^{\infty} (-1)^k b_k$:

$$\left| s - \sum_{k=1}^n (-1)^k b_k \right| \leq b_{n+1}$$

Beweis. Zum Beweis benutzen wir das Abel-Dirichlet-Kriterium. Wir setzen in Satz 3.8 $x_k := (-1)^k$ und $a_k := b_k$. Dann ist $s_n = x_1 + \dots + x_n \in \{0, -1\}$, also beschränkt, und (a_k) eine monoton fallende Nullfolge. Folglich konvergiert $\sum_{k=1}^n (-1)^k b_k$. Für den Wert s dieser Reihe gilt

$$\begin{aligned} s &= s_{2m} - b_{2m+1} + b_{2m+2} - b_{2m+3} + \dots \\ &= s_{2m+1} + b_{2m+2} - b_{2m+3} + b_{2m+4} - \dots \end{aligned}$$

Da die Folge (b_k) monoton fallend ist, gilt

$$s_{2m+1} < s < s_{2m}$$

und somit

$$|s - s_n| < |s_{n+1} - s_n| = b_{n+1}.$$

Dies zeigt die Fehlerabschätzung. □

Beispiel 8: Die Leibnitz-Reihe

Die 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 nach dem Leibnitz-Kriterium. Wir werden später sehen, dass der Grenzwert $\frac{\pi}{4}$ ist.

Satz 3.10 (Cauchysches Verdichtungskriterium)

Sei (x_k) eine monoton fallende Folge positiver reeller Zahlen. Dann gilt

$$\sum_{k=1}^{\infty} x_k \text{ konvergiert} \iff \sum_{i=0}^{\infty} 2^i x_{2^i} \text{ konvergiert.}$$

Beweis. Wir betrachten die Partialsummen beider Reihen

$$\begin{aligned} s_n &:= x_1 + \dots + x_n \\ t_m &:= x_1 + 2x_2 + 4x_4 + \dots + 2^m x_{2^m}. \end{aligned}$$

Zunächst schätzen wir diese beiden Partialsummen gegeneinander ab. Nach Voraussetzung ist die Folge (x_k) monoton fallend. Somit gilt für alle $n < 2^{m+1}$

$$\begin{aligned} s_n = x_1 + \dots + x_n &\leq x_1 + (x_2 + x_3) + \dots + \overbrace{(x_{2^m} + \dots + x_{2^{m+1}-1})}^{2^m \text{ Summanden}} \\ &\leq x_1 + 2x_2 + \dots + 2^m x_{2^m} = t_m. \end{aligned}$$

Folglich gilt

$$s_n \leq t_m \quad \forall n < 2^{m+1} \quad (*).$$

Andererseits folgt für $n \geq 2^m$

$$\begin{aligned} s_n = x_1 + \dots + x_{2^k} + \dots + x_n \\ &\geq x_1 + x_2 + (x_3 + x_4) + \dots + \overbrace{(x_{2^{m-1}+1} + \dots + x_{2^m})}^{2^{m-1} \text{ Summanden}} \\ &\geq \frac{1}{2} x_1 + x_2 + 2x_4 + \dots + 2^{m-1} x_{2^m} = \frac{1}{2} t_m. \end{aligned}$$

Folglich gilt

$$2s_n \geq t_m \quad \forall n \geq 2^m \quad (**).$$

Nun zum Beweis der behaupteten Äquivalenz:

(\implies): Sei $\sum_{k=1}^{\infty} x_k$ konvergent. Dann ist die Folge (s_n) konvergent und folglich beschränkt.

Wegen (***) ist dann auch die Folge (t_m) beschränkt. Außerdem ist (t_m) monoton wachsend.

Nach Satz 2.18 konvergiert dann die Folge (t_m) , also die Reihe $\sum_{j=0}^{\infty} 2^j x_{2^j}$.

(\Leftarrow): Sei $\sum_{j=0}^{\infty} 2^j x_{2^j}$ konvergent. Dann ist die Folge (t_m) konvergent und folglich beschränkt.

Wegen (*) ist dann auch die Folge (s_n) beschränkt. Außerdem ist (s_n) monoton wachsend.

Nach Satz 2.18 konvergiert dann die Folge (s_n) , also die Reihe $\sum_{k=1}^{\infty} x_k$. □

Beispiel 9: Sei p eine positive rationale Zahl. Die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k^p}$$

ist genau dann konvergent, wenn $p > 1$.

Beweis. Da $p > 0$, sind die Voraussetzungen von Satz 3.10 erfüllt. Dann gilt

$$\sum_{k=1}^{\infty} \frac{1}{k^p} \text{ konvergiert} \iff \sum_{j=0}^{\infty} 2^j \frac{1}{(2^j)^p} = \sum_{j=0}^{\infty} (2^{(1-p)})^j \text{ konvergiert.}$$

Die geometrische Reihe $\sum_{j=0}^{\infty} (2^{(1-p)})^j$ konvergiert genau dann, wenn $2^{(1-p)} < 1$ d.h. genau dann, wenn $p > 1$ gilt. □

3.3 Das Cauchy-Produkt von Reihen komplexer Zahlen

Wir haben bereits gesehen, dass man konvergente Reihen in Banachräumen addieren und mit Skalaren multiplizieren kann. Betrachtet man nur Reihen im Banachraum der komplexen Zahlen ($E = \mathbb{C}$), so kann man sie auch auf bestimmte Weise multiplizieren.

Definition 3.4. Seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ zwei Reihen in \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 10: 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 3.11 Seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ konvergente Reihen 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. 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 3.12 *Das Cauchy-Produkt zweier absolut-konvergenter Reihen ist absolut-konvergent.*

Beweis. Dies überlassen wir dem Leser als Übungsaufgabe. \square

Beispiel 11: $\sum_{k=1}^{\infty} k \cdot z^{k-1} = \frac{1}{(1-z)^2}$ für $|z| < 1$.

Beweis. Wir betrachten die Reihen $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ mit $a_k = b_k = z^k$. Dann ist

$$c_k = \sum_{j=0}^k z^j \cdot z^{k-j} = (k+1) \cdot z^k. \text{ Also ist das Cauchy-Produkt } \sum_{k=0}^{\infty} c_k = \sum_{k=1}^{\infty} k \cdot z^{k-1}. \text{ Da die}$$

geometrische Reihe $\sum_{k=0}^{\infty} z^k$ für $|z| < 1$ absolut konvergiert und ihr Wert $\frac{1}{1-z}$ ist, konvergiert das Cauchy-Produkt für $|z| < 1$ gegen den oben angegebenen Wert. (Die Konvergenz folgt zwar auch aus dem Quotientenkriterium, aber auf diese Weise kann man auch den Wert der Reihe bestimmen). \square

3.4 Umordnung von Reihen

In einer endlichen Summe kann man die Summanden beliebig umordnen, ohne den Wert der Summe zu verändern. Dies ist für Reihen ("unendliche Summen") nicht mehr der Fall. Abschließend untersuchen wir deshalb die Frage, unter welchen Bedingungen man Reihenglieder umordnen kann, ohne die Konvergenz zu verlieren.

Definition 3.5. Sei $\sum_{k=1}^{\infty} x_k$ eine Reihe im Banachraum E und $f: \mathbb{N} \rightarrow \mathbb{N}$ eine bijektive Abbildung. Die Reihe $\sum_{k=1}^{\infty} x_{f(k)}$ heißt Umordnung von $\sum_{k=1}^{\infty} x_k$.

Man kann die Glieder einer Reihe nicht beliebig umordnen. Betrachten wir z.B. die Reihe

$$\frac{1}{\sqrt{1}} - \frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} - \frac{1}{\sqrt{3}} \pm \dots \quad (*)$$

Die Reihe (*) ist konvergent, denn für die Partialsummen gilt

$$\left. \begin{array}{l} s_{2m} = 0 \\ s_{2m+1} = \frac{1}{\sqrt{m+1}} \end{array} \right\} \text{ dh. } s_m \rightarrow 0.$$

Betrachten wir nun folgende Umordnung der obigen Reihe (*)

$$\left(\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{1}} \right) + \left(\frac{1}{\sqrt{3}} + \frac{1}{\sqrt{4}} - \frac{1}{\sqrt{2}} \right) + \left(\frac{1}{\sqrt{5}} + \frac{1}{\sqrt{6}} - \frac{1}{\sqrt{3}} \right) + \dots \quad (**)$$

Für die $3n$ -ten Partialsummen s_{3n}^* dieser Reihe (**) gilt

$$s_{3n}^* = \underbrace{\frac{1}{\sqrt{n+1}} + \dots + \frac{1}{\sqrt{2n}}}_{n \text{ Summanden}} \geq \underbrace{\frac{1}{\sqrt{2n}} + \dots + \frac{1}{\sqrt{2n}}}_{n \text{ Summanden}} = \frac{n}{\sqrt{2n}} = \sqrt{\frac{n}{2}} \rightarrow \infty.$$

Somit ist diese Reihe (**) nicht konvergent.

Satz 3.13 (Umordnungssatz) Sei $\sum_{k=1}^{\infty} x_k$ eine absolut-konvergente Reihe im Banachraum E . Dann ist jede Umordnung von $\sum_{k=1}^{\infty} x_k$ konvergent und für den Wert der Reihen gilt:

$$\sum_{k=1}^{\infty} x_k = \sum_{k=1}^{\infty} x_{f(k)}.$$

Beweis. Sei $\sum_{k=1}^{\infty} x_{f(k)}$ eine Umordnung von $\sum_{k=1}^{\infty} x_k$. Nach Voraussetzung ist $\sum_{k=1}^{\infty} x_k$ absolut-konvergent, also auch konvergent, d.h. es existiert ein $s \in E$ mit $s = \sum_{k=1}^{\infty} x_k$. Es bezeichne

$$s_n := x_1 + \dots + x_n \quad \text{und} \quad s_n^* := x_{f(1)} + \dots + x_{f(n)}.$$

Sei nun eine Zahl $\varepsilon > 0$ gegeben. Da $\sum_{k=1}^{\infty} x_k$ absolut konvergiert, existiert nach dem Cauchy-Kriterium ein $n_0 \in \mathbb{N}$, so dass

$$\sum_{j=n}^m \|x_j\| < \frac{\varepsilon}{2}, \quad \forall m > n \geq n_0 \quad (*)$$

Da die Umordnung $f: \mathbb{N} \rightarrow \mathbb{N}$ bijektiv ist, gibt es ein $p_0 \in \mathbb{N}$, so dass $\{1, \dots, n_0\} \subset f(\{1, \dots, p_0\})$. Für $p \geq p_0$ ergibt sich dann

$$\begin{aligned} s_p - s_p^* &= x_1 + \dots + x_p - (x_{f(1)} + \dots + x_{f(p)}) \\ &= \text{Summe gewisser } x_i \text{ für } i > n_0. \end{aligned}$$

Aus (*) folgt $\|s_p - s_p^*\| < \frac{\varepsilon}{2}$ für alle $p \geq p_0$. Da die Folge (s_p) gegen s konvergiert, gibt es auch ein $n_0^* \in \mathbb{N}$ so dass $\|s_p - s\| < \frac{\varepsilon}{2}$ für alle $p > n_0^*$. Damit erhalten wir

$$\|s_n^* - s\| \leq \|s_n^* - s_n\| + \|s_n - s\| < \varepsilon \quad \forall n \geq \max\{p_0, n_0^*\}.$$

Folglich konvergiert die Folge (s_n^*) gegen s , also gilt $\sum_{k=1}^{\infty} x_{f(k)} = s$. \square

Ohne die Voraussetzung der absoluten Konvergenz gilt Satz 3.13 nicht mehr. Es gilt sogar der folgende ziemlich erstaunliche Satz:

Satz 3.14 (Riemannscher Umordnungssatz)

Sei $\sum_{k=1}^{\infty} a_k$ eine konvergente, aber nicht absolut-konvergente Reihe reeller Zahlen und sei $s \in \mathbb{R}$ eine beliebig (!) vorgegebene Zahl. Dann existiert eine bijektive Abbildung $f : \mathbb{N} \rightarrow \mathbb{N}$, so dass die Umordnung $\sum_{k=1}^{\infty} a_{f(k)}$ gegen s konvergiert.

Beweis. Sei oBdA $a_k \neq 0$ für alle $k \in \mathbb{N}$. Wir setzen

$$a_k^+ := \frac{|a_k| + a_k}{2} = \begin{cases} a_k & \text{falls } a_k > 0 \\ 0 & \text{falls } a_k < 0 \end{cases}$$

$$a_k^- := \frac{|a_k| - a_k}{2} = \begin{cases} 0 & \text{falls } a_k > 0 \\ -a_k & \text{falls } a_k < 0 \end{cases}$$

Dann gilt $a_k = a_k^+ - a_k^-$ und $|a_k| = a_k^+ + a_k^-$. Aus der Voraussetzung des Satzes folgt, dass beide Reihen $\sum_{k=1}^{\infty} a_k^+$ und $\sum_{k=1}^{\infty} a_k^-$ divergent sind. Wäre nämlich eine dieser beiden

Reihen konvergent, so würde wegen der Konvergenz von $\sum_{k=1}^{\infty} a_k$ auch die andere dieser

beiden Reihen konvergieren und die Konvergenz von $\sum_{k=1}^{\infty} |a_k|$ folgen (Satz 3.4).

Sei nun (p_m) die Teilfolge *aller* positiven Zahlen von (a_k) und (q_m) die Teilfolge *aller* negativen Zahlen von (a_k) . Offensichtlich entsteht die Folge (p_m) , wenn man aus der Folge (a_k^+) alle Nullen streicht und $(-q_m)$, wenn man aus der Folge (a_k^-) alle Nullen streicht.

Nach dem oben Gesagten sind die Reihen $\sum_{m=1}^{\infty} p_m$ und $\sum_{m=1}^{\infty} q_m$ divergent und es gilt

$$\sum_{m=1}^{\infty} p_m = +\infty \quad , \quad \sum_{m=1}^{\infty} q_m = -\infty \quad (*)$$

Wir konstruieren nun die gesuchte Umordnung der Reihe $\sum_{k=1}^{\infty} a_k$. Wegen (*) kann man schrittweise eine Folge jeweils kleinster Indizes n_0, n_1, n_2, \dots mit der folgenden Eigenschaft wählen:

$$\begin{array}{ll}
 n_0 \text{ so dass} & \sum_{m=1}^{n_0} p_m > s \\
 n_1 \text{ so dass} & \sum_{m=1}^{n_0} p_m + \sum_{m=1}^{n_1} q_m < s \\
 n_2 > n_0 \text{ so dass} & \sum_{m=1}^{n_0} p_m + \sum_{m=1}^{n_1} q_m + \sum_{m=n_0+1}^{n_2} p_m > s \\
 n_3 > n_1 \text{ so dass} & \sum_{m=1}^{n_0} p_m + \sum_{m=1}^{n_1} q_m + \sum_{m=n_0+1}^{n_2} p_m + \sum_{m=n_1+1}^{n_3} q_m < s \\
 \vdots &
 \end{array}$$

Die Reihe

$$p_1 + \dots + p_{n_0} + q_1 + \dots + q_{n_1} + p_{n_0+1} + \dots + p_{n_2} + q_{n_1+1} + \dots + q_{n_3} + \dots \quad (**)$$

ist eine Umordnung von $\sum_{k=1}^{\infty} a_k$. Aus der Minimalitätseigenschaft der Wahl der Indizes n_j erhält man folgende Abschätzung für die Partialsumme s_n^* der Reihe (**):

$$\begin{aligned}
 |s - s_n^*| &\leq p_{n_{2j}} & \forall \quad n_{2j-1} + n_{2j} \leq n < n_{2j} + n_{2j+1} \\
 |s - s_n^*| &\leq -q_{n_{2j+1}} & \forall \quad n_{2j} + n_{2j+1} \leq n < n_{2j+1} + n_{2j+2}
 \end{aligned}$$

Da die Reihe $\sum_{k=1}^{\infty} a_k$ konvergiert, sind die Teilfolgen $(p_{n_{2j}})_{j=0}^{\infty}$ und $(q_{n_{2j+1}})_{j=0}^{\infty}$ von (a_n) Nullfolgen. Folglich konvergiert die Folge der Partialsummen (s_n^*) der Umordnung (**) gegen s . □

3.5 Komplexe Potenzreihen

Als Anwendung der Konvergenzkriterien aus Kapitel 3.2 betrachten wir in diesem Abschnitt spezielle Reihen im Banachraum der komplexen Zahlen, die sogenannten Potenzreihen.

Definition 3.6. Sei $z_0 \in \mathbb{C}$ und (a_n) 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 3.15 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 $|z - z_0| < |z_1 - z_0|$ absolutkonvergent.
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 3.7. *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 $P(z)$.

Wir können die Aussage aus Satz 3.15 jetzt auch folgendermaßen formulieren:

Satz 3.16 *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 Kapitel 3.2 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 3.17 *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 3.18 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 $P(z)$:

$$R = \begin{cases} \frac{1}{\mu} & \text{falls } \mu \in \mathbb{R}^+ \\ 0 & \text{falls } \mu = +\infty \\ +\infty & \text{falls } \mu = 0. \end{cases}$$

3.6 Anwendung: Exponentialfunktion, Logarithmusfunktion und komplexe Potenzen

In Kapitel 1.2 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 3.3 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 3.11, 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 $E(q)$ positiv sind, folgt $e^q = E(q)$.

Die letzte Eigenschaft von $E(z)$ rechtfertigt die folgende Definition:

Definition 3.8. 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 *Exponentialfunktion*.

Aus den Eigenschaften der Potenzreihe $E(z)$ folgt

Satz 3.19 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 $|z| < 1$.

Satz 3.20 Die Funktion $\exp_{|\mathbb{R}} : \mathbb{R} \longrightarrow \mathbb{R}^+$ ist streng monoton wachsend und bijektiv. Außerdem gilt:

1. $1 \leq e^x$ für alle $x \in \mathbb{R}$ mit $x \geq 0$.
2. $0 < e^x \leq 1$ für alle $x \in \mathbb{R}$ mit $x \leq 0$.

Beweis. 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 $x \leq 0$. 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, denn $e^{\frac{y-1}{y}} < y \leq e^y$ (siehe Übung).
- 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 $(A|B)$. Wir zeigen nun, dass $e^s = y$ gilt.

Wir wissen, dass $a \leq s \leq b$ für alle $a \in A$, $b \in B$. Sei (a_n) eine Folge in A , die gegen s konvergiert. Dann gilt $e^{a_n} < y$ und

$$|e^s - e^{a_n}| = |e^{a_n}| |e^{(s-a_n)} - 1| < |y| \cdot \frac{|s - a_n|}{1 - |s - a_n|} \longrightarrow 0.$$

Daraus folgt $\lim_{n \rightarrow \infty} e^{a_n} = e^s \leq y$. Sei (b_n) eine Folge in B , die gegen s konvergiert. Dann gilt $e^{b_n} \geq y$ und

$$|e^{b_n} - e^s| = |e^s| |e^{(b_n-s)} - 1| \leq |e^s| \cdot \frac{|b_n - s|}{1 - |b_n - s|} \rightarrow 0.$$

Es folgt $\lim_{n \rightarrow \infty} e^{b_n} = e^s \geq y$. Also ist $y = e^s$. □

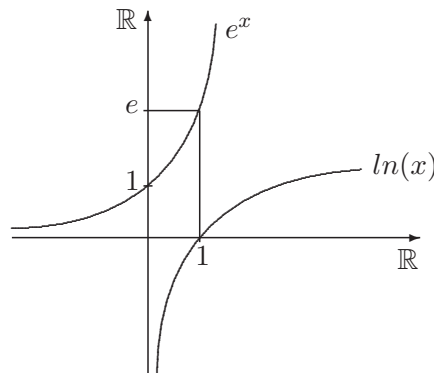
Definition 3.9. Die Umkehrfunktion von $\exp_{\mathbb{R}}$ bezeichnen wir mit $\ln : \mathbb{R}^+ \rightarrow \mathbb{R}$ und nennen $\ln(y)$ den natürlichen Logarithmus von y .

Die Zahl $\ln(y)$ ist also eindeutig bestimmt durch die Bedingung $e^{\ln(y)} = y$. Daraus ergeben sich folgende Eigenschaften für den natürlichen Logarithmus:

Satz 3.21 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 den Funktionsverlauf der reellen Funktionen $\exp_{\mathbb{R}}$ und \ln .



Sei $a \in \mathbb{R}^+$. Dann gilt für die Potenz $a^q, q \in \mathbb{Q}$, die Formel

$$a^q = e^{\ln(a) \cdot q} = \sum_{n=0}^{\infty} \frac{(q \cdot \ln(a))^n}{n!}.$$

Dies rechtfertigt die folgende Definition der komplexen Potenzen einer positiven reellen Zahl a :

Definition 3.10. 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 3.22 (*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 3.22 folgen dann unmittelbar die Aussagen des folgenden Satzes.

Satz 3.23 *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 Abbildungen zwischen metrischen Räumen

4.1 Der Grenzwert einer Abbildung in einem Punkt

In diesem Abschnitt bezeichnen (X, d_X) und (Y, d_Y) metrische Räume, $A \subset X$ eine Teilmenge von X und $f : A \subset X \rightarrow Y$ eine auf A definierte Abbildung. Die Vektorräume \mathbb{R}^n und \mathbb{C}^n , $n \in \mathbb{N}$, seien immer mit der Euklidischen Metrik versehen.

Definition 4.1. Sei $x_0 \in X$ ein Häufungspunkt von $A \subset X$. Ein Punkt $y_0 \in Y$ heißt Grenzwert der Abbildung $f : A \subset X \rightarrow Y$ im Punkt x_0 , wenn für jede Folge (a_n) aus $A \setminus \{x_0\}$, die gegen x_0 konvergiert, die Folge der Bildwerte $(f(a_n))$ gegen y_0 konvergiert. Wir schreiben dafür symbolisch:

$$y_0 = \lim_{a \rightarrow x_0} f(a).$$

Bemerkung: Der Punkt x_0 muss nicht unbedingt im Definitionsbereich A von f liegen. Wir benötigen lediglich, dass er der Grenzwert einer Folge von Punkten aus A ist.

Beispiel 1: Es sei $X = \mathbb{R}^2$, $A = \mathbb{R}^2 \setminus \{(0,0)\}$ und $Y = \mathbb{R}$. Wir betrachten die Abbildung $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) := \frac{x^3 - y^3}{x^2 + y^2}.$$

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

Beweis. Sei $(x, y) \neq (0, 0)$. Dann gilt

$$\frac{x^3 - y^3}{x^2 + y^2} = \frac{(x - y)(x^2 + xy + y^2)}{x^2 + y^2} = (x - y) \left(1 + \frac{xy}{x^2 + y^2} \right).$$

Wegen $\left| \frac{xy}{x^2 + y^2} \right| \leq \frac{1}{2}$, folgt

$$|f(x, y)| \leq \frac{3}{2}|x - y|.$$

Sei nun $((x_n, y_n))$ eine beliebige gegen $(0, 0)$ konvergente Folge in A . Dann gilt $x_n \rightarrow 0$ und $y_n \rightarrow 0$, somit $x_n - y_n \rightarrow 0$ und folglich $|f(x_n, y_n)| \rightarrow 0$. Wir erhalten also $\lim_{n \rightarrow \infty} f(x_n, y_n) = 0$ für alle Folgen $(x_n, y_n) \rightarrow (0, 0)$ und somit $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = 0$. \square

Beispiel 2: Es sei $X = \mathbb{R}^2$, $A = \mathbb{R}^2 \setminus \{(0,0)\}$ und $Y = \mathbb{R}$. Wir betrachten die Abbildung $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) = \frac{xy^2}{x^2 + y^4}$$

Behauptung: $\lim_{(x,y) \rightarrow (0,0)} f(x, y)$ existiert nicht.

Beweis. Wir betrachten die Folgen (a_n) und (b_n) in A mit $a_n := (\frac{1}{n}, \frac{1}{n})$ und $b_n := (\frac{1}{n^2}, \frac{1}{n})$. Dann gilt

$$\begin{aligned} a_n = \left(\frac{1}{n}, \frac{1}{n}\right) &\longrightarrow (0, 0), & f(a_n) &= \frac{\frac{1}{n^3}}{\frac{1}{n^2} + \frac{1}{n^4}} = \frac{n}{n^2+1} \longrightarrow 0, \\ b_n = \left(\frac{1}{n^2}, \frac{1}{n}\right) &\longrightarrow (0, 0), & f(b_n) &= \frac{\frac{1}{n^4}}{\frac{1}{n^4} + \frac{1}{n^4}} = \frac{1}{2} \longrightarrow \frac{1}{2}. \end{aligned}$$

Folglich hat f in $(0,0)$ keinen Grenzwert. \square

Satz 4.1 Sei $f : A \subset X \rightarrow Y$ eine Abbildung und $x_0 \in HP(A)$. Dann sind die folgenden Bedingungen äquivalent:

1. $\lim_{a \rightarrow x_0} f(a) = y_0$.
2. Für alle $\varepsilon > 0$ existiert ein $\delta > 0$, so dass

$$f\left(A \cap K_X(x_0, \delta) \setminus \{x_0\}\right) \subset K_Y(y_0, \varepsilon),$$

d.h. so dass für alle $a \in A$ mit $0 < d_X(x_0, a) < \delta$ gilt $d_Y(y_0, f(a)) < \varepsilon$.

Beweis. (1) \implies (2): Sei $\lim_{a \rightarrow x_0} f(a) = y_0$. Angenommen 2. gilt nicht, das heißt es existiert ein $\varepsilon_0 > 0$, so dass es zu jedem $\delta > 0$ ein $a \in A$ gibt mit $0 < d_X(x_0, a) < \delta$ und $d_Y(y_0, f(a)) \geq \varepsilon_0$. Wir wählen speziell $\delta = \frac{1}{n}$. Dann erhalten wir eine Folge (a_n) in $A \setminus \{x_0\}$ so dass $d_X(x_0, a_n) < \frac{1}{n}$ und $d_Y(y_0, f(a_n)) \geq \varepsilon_0$ gilt. Daraus folgt $a_n \rightarrow x_0$ und $f(a_n) \not\rightarrow y_0$. Dies steht aber im Widerspruch zu der Voraussetzung $\lim_{n \rightarrow \infty} f(a_n) = y_0$.

(2) \implies (1): Sei 2. erfüllt. Wir betrachten eine gegen x_0 konvergente Folge (a_n) in $A \setminus \{x_0\}$. Sei $\varepsilon > 0$. Nach Voraussetzung existiert ein $\delta > 0$, so dass

$$f\left(A \cap K_X(x_0, \delta) \setminus \{x_0\}\right) \subset K_Y(y_0, \varepsilon).$$

Da (a_n) gegen x_0 konvergiert, existiert ein $n_0 \in \mathbb{N}$, so dass $d_X(x_0, a_n) < \delta$ für alle $n \geq n_0$. Folglich ist $d_Y(y_0, f(a_n)) < \varepsilon$ für alle $n \geq n_0$. Also konvergiert $(f(a_n))$ gegen y_0 und zwar für jede Folge (a_n) in $A \setminus \{x_0\}$ mit $a_n \rightarrow x_0$. \square

Wir betrachten als nächstes den Fall, dass der metrische Raum (Y, d_Y) ein Produktraum ist.

Satz 4.2 (Grenzwerte von Abbildungen mit Werten in Produkträumen)

Sei (Y, d_Y) das Produkt der metrischen Räume $(Y_1, d_1), (Y_2, d_2), \dots, (Y_n, d_n)$ und sei $f : A \subset X \rightarrow Y = Y_1 \times Y_2 \times \dots \times Y_n$ eine Abbildung in den Produktraum. Wir bezeichnen

mit $f_j : A \rightarrow Y_j$, $j = 1, \dots, n$, die Komponenten der Abbildung f , definiert durch

$$f(a) =: (f_1(a), f_2(a), \dots, f_n(a)), \quad a \in A.$$

Sei $x_0 \in HP(A)$. Dann hat f in x_0 genau dann einen Grenzwert, wenn jede Komponente f_j von f in x_0 einen Grenzwert hat und für diese Grenzwerte gilt:

$$\lim_{a \rightarrow x_0} f(a) = \left(\lim_{a \rightarrow x_0} f_1(a), \dots, \lim_{a \rightarrow x_0} f_n(a) \right).$$

Beweis. Die Behauptung folgt aus dem Satz 2.12 über die Konvergenz von Folgen in Produkträumen. \square

Spezialfall: Für eine Abbildung $f : A \subset X \rightarrow \mathbb{C}$ gilt im Falle der Existenz der Grenzwerte:

$$\lim_{a \rightarrow x_0} f(a) = \lim_{a \rightarrow x_0} \operatorname{Re}(f(a)) + i \cdot \lim_{a \rightarrow x_0} \operatorname{Im}(f(a)).$$

Als nächstes betrachten wir Abbildungen mit Werten in normierten Vektorräumen.

Sei $(V, \|\cdot\|)$ ein normierter Vektorraum über dem Körper \mathbb{K} der reellen oder komplexen Zahlen. $(V, \|\cdot\|)$ sei mit der von der Norm induzierten Metrik versehen. Ist auf V ein Skalarprodukt $\langle \cdot, \cdot \rangle$ gegeben, so versehen wir V immer mit der vom Skalarprodukt induzierten Norm $\|v\| := \sqrt{\langle v, v \rangle}$. Sind $f, g : A \subset X \rightarrow V$ zwei Abbildungen in den Vektorraum V und $h : A \subset X \rightarrow \mathbb{K}$ eine Abbildung in den Körper \mathbb{K} , so bezeichnen $f + g : A \subset X \rightarrow V$, $h \cdot f : A \subset X \rightarrow V$, $\langle f, g \rangle : A \subset X \rightarrow \mathbb{K}$ und $\|f\| : A \subset X \rightarrow \mathbb{R}$ die Abbildungen

$$\begin{aligned} (f + g)(a) &:= f(a) + g(a), \\ (h \cdot f)(a) &:= h(a) \cdot f(a), \\ \langle f, g \rangle(a) &:= \langle f(a), g(a) \rangle, \\ \|f\|(a) &:= \|f(a)\|, \end{aligned} \quad \text{wobei } a \in A.$$

Die speziellen Rechenregeln für konvergente Folgen in Vektorräumen, siehe Satz 2.16, liefern unmittelbar die folgenden Aussagen über Grenzwerte von Abbildungen mit Werten in Vektorräumen¹:

Satz 4.3 (Grenzwerte von Abbildungen mit Werten in normierten Vektorräumen)

1. Seien $f, g : A \subset X \rightarrow V$ zwei Abbildungen, für die der Grenzwert im Punkt $x_0 \in HP(A)$ existiert. Dann gilt:

(a) Es existiert auch der Grenzwert von $f + g$ im Punkt x_0 und

$$\lim_{a \rightarrow x_0} (g + f)(a) = \lim_{a \rightarrow x_0} g(a) + \lim_{a \rightarrow x_0} f(a).$$

(b) Es existiert auch der Grenzwert von $\langle f, g \rangle$ im Punkt x_0 und

$$\lim_{a \rightarrow x_0} \langle f, g \rangle(a) = \langle \lim_{a \rightarrow x_0} f(a), \lim_{a \rightarrow x_0} g(a) \rangle.$$

¹ Im Beweis von Satz 2.16 hatten wir lediglich die Eigenschaften eines Vektorraumes und die Eigenschaften von Skalarprodukt und Norm benutzt und nicht die spezielle Situation für den \mathbb{K}^n .

(c) Es existiert auch der Grenzwert von $\|f\|$ im Punkt x_0 und

$$\lim_{a \rightarrow x_0} \|f\|(a) = \left\| \lim_{a \rightarrow x_0} f(a) \right\|.$$

2. Seien $f : A \subset X \rightarrow V$ und $h : A \subset X \rightarrow \mathbb{K}$ zwei Abbildungen, für die der Grenzwert im Punkt $x_0 \in HP(A)$ existiert. Dann existiert auch der Grenzwert von $h \cdot f$ im Punkt x_0 und es gilt

$$\lim_{a \rightarrow x_0} (h \cdot f)(a) = \lim_{a \rightarrow x_0} h(a) \cdot \lim_{a \rightarrow x_0} f(a).$$

3. Seien $h : A \subset X \rightarrow \mathbb{K}$ und $p : A \subset X \rightarrow \mathbb{K}$ zwei Abbildungen, für die der Grenzwert im Punkt $x_0 \in HP(A)$ existiert und sei $h(a) \neq 0$ für alle $a \in A$ und $\lim_{a \rightarrow x_0} h(a) \neq 0$. Dann existiert auch der Grenzwert von $\frac{p}{h}$ im Punkt x_0 und es gilt

$$\lim_{a \rightarrow x_0} \left(\frac{p}{h} \right)(a) = \frac{\lim_{a \rightarrow x_0} p(a)}{\lim_{a \rightarrow x_0} h(a)}.$$

□

Im metrischen Raum \mathbb{R} hatten wir außer den 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 Abbildungen $f : A \subset X \rightarrow Y$, wobei $X = \mathbb{R}$ oder $Y = \mathbb{R}$ ist. Dabei treten 3 Fälle auf:

1. Fall: Sei X ein beliebiger metrischer Raum, $Y = \mathbb{R}$ und $f : A \subset X \rightarrow \mathbb{R}$.

Man sagt, dass der (uneigentliche) Grenzwert von f in $x_0 \in HP(A)$ existiert und gleich $+\infty$ ist, falls $\lim_{n \rightarrow \infty} f(a_n) = +\infty$ für jede gegen x_0 konvergente Folge (a_n) in $A \setminus \{x_0\}$ ist. Wir schreiben in diesem Fall $\lim_{a \rightarrow x_0} f(a) = +\infty$.

Analog wird $\lim_{a \rightarrow x_0} f(a) = -\infty$ definiert.

Satz 4.4 Sei $f : A \subset X \rightarrow \mathbb{R}$ eine Abbildung von einer Teilmenge eines metrischen Raumes X in die reellen Zahlen und $x_0 \in HP(A)$. Dann sind folgende Bedingungen äquivalent:

1. $\lim_{a \rightarrow x_0} f(a) = +\infty$.
2. Für alle $M > 0$ existiert ein $\delta > 0$, so dass $f(a) > M$ für alle $a \in A$ mit $0 < d_X(a, x_0) < \delta$.

Beweis. Der Beweis wird analog zum Beweis von Satz 4.1 geführt. □

Beispiel 3: Wir betrachten die Funktion $f : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow \mathbb{R}$ definiert durch

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

Dann gilt $\lim_{(x,y) \rightarrow (0,0)} f(x, y) = +\infty$.

2. Fall: Sei $X = \mathbb{R}$, Y ein beliebiger metrischer Raum und $f : (c, +\infty) \subset \mathbb{R} \rightarrow Y$.

Man sagt, dass $\lim_{a \rightarrow +\infty} f(a)$ existiert und gleich $y_0 \in Y$ ist, falls $\lim_{n \rightarrow +\infty} f(a_n) = y_0$ für jede

Folge (a_n) in $(c, +\infty)$ mit $a_n \rightarrow +\infty$. Wir schreiben in diesem Fall $\lim_{a \rightarrow +\infty} f(a) = y_0$. Analog definiert man $\lim_{a \rightarrow -\infty} f(a) = y_0$.

Satz 4.5 *Mit den obigen Bezeichnungen sind folgende Bedingungen äquivalent:*

1. $\lim_{a \rightarrow +\infty} f(a) = y_0$.
2. Für alle $\varepsilon > 0$ existiert ein $M > 0$ so dass $d_Y(f(a), y_0) < \varepsilon$ für alle $a > M$.

Beweis. Der Beweis wird analog zum Beweis von Satz 4.1 geführt. □

Beispiel 4: Seien $f_1, f_2 : (0, +\infty) \subset \mathbb{R} \rightarrow \mathbb{C}$ gegeben durch

$$f_1(t) := \frac{1}{t}(\cos t + i \sin t) \quad f_2(t) := \cos t + i \sin t$$

Dann gilt $\lim_{t \rightarrow +\infty} f_1(t) = 0$ während der GW von f_2 für $t \rightarrow +\infty$ nicht existiert.

Beweis. Die erste Behauptung folgt, da $|f_1(t)| = \frac{1}{t}$. Für die 2. Behauptung betrachte man die Folge $(t_n = n\pi)$. Da $f_2(2k\pi) = 1$ und $f_2((2k+1)\pi) = -1$, hat die Bildfolge $(f_2(t_n))$ zwei verschiedene Häufungspunkte. □

3. Fall: Seien $X = Y = \mathbb{R}$ und $f : (c, +\infty) \subset \mathbb{R} \rightarrow \mathbb{R}$.

Man sagt, dass die Funktion f für a gegen $+\infty$ gegen $+\infty$ bzw. $-\infty$ strebt (symbolisch: $\lim_{a \rightarrow +\infty} f(a) = \pm\infty$), wenn für jede Folge (a_n) in $(c, +\infty)$, die gegen $+\infty$ strebt, die Folge $(f(a_n))$ gegen $+\infty$ bzw. $-\infty$ strebt.

Analog wird der uneigentliche Grenzwert $\pm\infty$ von f für $a \rightarrow -\infty$ definiert.

Im Fall von reellen Funktionen $f : A \subset \mathbb{R} \rightarrow \mathbb{R}$ kann man einseitige Grenzwerte definieren. Wir betrachten dazu die folgenden Teilmengen der Menge der Häufungspunkte von A :

$$HP(A)^- := \{x \in \mathbb{R} \mid \text{es existiert eine Folge } (a_n) \text{ in } A \text{ mit } a_n < x \text{ und } a_n \rightarrow x\}.$$

$$HP(A)^+ := \{x \in \mathbb{R} \mid \text{es existiert eine Folge } (a_n) \text{ in } A \text{ mit } a_n > x \text{ und } a_n \rightarrow x\}.$$

Definition 4.2. *Man sagt, dass der linksseitige Grenzwert von $f : A \subset \mathbb{R} \rightarrow \mathbb{R}$ in $x_0 \in HP(A)^-$ existiert und gleich $y_0 \in \mathbb{R} \cup \{\pm\infty\}$ ist, falls $\lim_{n \rightarrow \infty} f(a_n) = y_0$ für jede gegen x_0 konvergente Folge (a_n) in A mit $a_n < x_0$.*

Wir bezeichnen den linksseitigen Grenzwert mit $\lim_{a \rightarrow x_0^-} f(a) = y_0$.

Man sagt, dass der rechtsseitige Grenzwert von $f : A \subset \mathbb{R} \rightarrow \mathbb{R}$ in $x_0 \in HP(A)^+$ existiert und gleich $y_0 \in \mathbb{R} \cup \{\pm\infty\}$ ist, falls $\lim_{n \rightarrow \infty} f(a_n) = y_0$ für jede gegen x_0 konvergente Folge (a_n) in A mit $a_n > x_0$.

Wir bezeichnen den rechtsseitigen Grenzwert mit $\lim_{a \rightarrow x_0^+} f(a) = y_0$.

Satz 4.6 *Sei $f : (\alpha, \beta) \subset \mathbb{R} \rightarrow \mathbb{R}$ eine monoton wachsende Funktion. Dann existieren für jedes $x_0 \in (\alpha, \beta)$ die einseitigen Grenzwerte in \mathbb{R} und es gilt*

$$\lim_{a \rightarrow x_0^-} f(a) = \sup\{f(t) \mid \alpha < t < x_0\} \quad \text{und} \quad \lim_{a \rightarrow x_0^+} f(a) = \inf\{f(t) \mid x_0 < t < \beta\}.$$

Die analoge Aussage gilt für monoton fallende Funktionen.

Beweis. Da f monoton wächst, ist die Menge $\{f(t) \mid \alpha < t < x_0\}$ von oben durch $f(x_0)$ beschränkt. Folglich existiert das Supremum

$$y_0 := \sup\{f(t) \mid \alpha < t < x_0\} \leq f(x_0).$$

Nach Definition des Supremums gibt es zu jedem $\varepsilon > 0$ ein $t^* \in (\alpha, x_0)$, so dass $y_0 - \varepsilon < f(t^*) \leq y_0$. Da f monoton wachsend ist, gilt desweiteren

$$y_0 - \varepsilon < f(t^*) \leq f(t) \leq y_0 \quad \forall t \in (t^*, x_0) \quad (*)$$

Sei nun $\delta := x_0 - t^*$. Dann folgt aus (*), dass für alle $|x_0 - t| < \delta$ mit $t < x_0$ die Abschätzung $|y_0 - f(t)| < \varepsilon$ gilt. Somit ist $\lim_{t \rightarrow x_0^-} f(t) = y_0$.

Analog beweist man $\lim_{t \rightarrow x_0^+} f(t) = \inf\{f(t) \mid \beta > t > x_0\}$. □

Beispiel 5: Sei $a > 1$ gegeben. Dann gilt $\lim_{x \rightarrow +\infty} \frac{a^x}{x} = +\infty$.

Beweis. Es ist

$$a^x = e^{x \cdot \ln(a)} = 1 + x \cdot \ln(a) + \frac{x^2 \cdot \ln^2(a)}{2!} + \dots$$

Für $a > 1$ ist $\ln(a) > 0$. Für $x > 0$ erhalten wir somit

$$\frac{a^x}{x} = \frac{1}{x} + \underbrace{\ln(a)}_{>0} + x \cdot \frac{\ln^2(a)}{2!} + \underbrace{\dots}_{>0} > \frac{\ln^2(a)}{2} \cdot x.$$

Da offensichtlich $\lim_{x \rightarrow +\infty} x = +\infty$, folgt $\lim_{x \rightarrow +\infty} \frac{a^x}{x} = +\infty$. □

Beispiel 6: Es gilt $\lim_{x \rightarrow +\infty} \frac{x}{\ln(x)} = +\infty$.

Beweis. Wir wissen, dass $\frac{x}{\ln(x)} = \frac{e^{\ln(x)}}{\ln(x)}$. Sei (x_n) eine gegen $+\infty$ strebende Folge. Dann gilt $y_n := \ln(x_n) \rightarrow +\infty$, da \ln streng monoton wachsend und bijektiv ist. Aus Beispiel 5 erhalten wir $\lim_{n \rightarrow +\infty} \frac{e^{\ln(x_n)}}{\ln(x_n)} = +\infty$. □

Beispiel 7: Es gilt $\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = 1$.

Beweis. Für $x > -1$ gilt $\frac{x}{1+x} \leq \ln(1+x) \leq x$ (siehe Übung 10). Daraus folgt

$$\frac{1}{1+x} \leq \frac{\ln(1+x)}{x} \leq 1, \quad \text{falls } x > 0, \text{ und}$$

$$\frac{1}{1+x} \geq \frac{\ln(1+x)}{x} \geq 1, \quad \text{falls } -1 < x < 0.$$

Da $\lim_{x \rightarrow 0} \frac{1}{1+x} = 1$, folgt aus den Grenzwertsätzen $\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = 1$. □

Beispiel 8: Es gilt $\lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}} = e$ und $\lim_{x \rightarrow +\infty} (1+\frac{1}{x})^x = e$.

Beweis. Nach Satz 3.19 gilt $|e^x - 1| \leq \frac{|x|}{1-|x|}$ für alle $|x| < 1$. Folglich ist $\lim_{n \rightarrow \infty} e^{a_n} = 1$ für jede Nullfolge (a_n) . Sei nun (x_n) eine beliebige Nullfolge in \mathbb{R} mit $x_n \neq 0$ und $u_n := \frac{\ln(1+x_n)}{x_n} - 1$. Aus Beispiel 7 ergibt sich, dass (u_n) eine Nullfolge ist. Also gilt

$$1 = \lim_{n \rightarrow \infty} e^{u_n} = \lim_{n \rightarrow \infty} e^{\ln(1+x_n) \cdot \frac{1}{x_n} \cdot e^{-1}} = \lim_{n \rightarrow \infty} (1+x_n)^{\frac{1}{x_n}} \cdot e^{-1}$$

und somit

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

□

4.2 Stetige Abbildungen (Definition und Beispiele)

Definition 4.3. Es sei $f : X \rightarrow Y$ eine Abbildung zwischen metrischen Räumen.

1. f heißt stetig im Punkt $x_0 \in X$, falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass

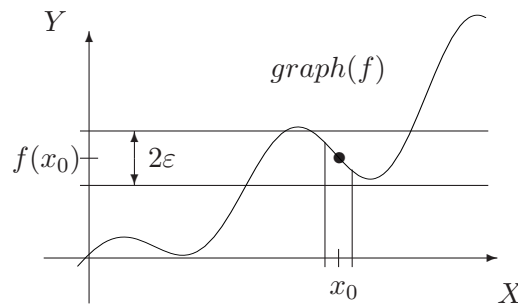
$$f(K_X(x_0, \delta)) \subset K_Y(f(x_0), \varepsilon),$$

d.h. so dass gilt:

$$x \in X \text{ und } d_X(x_0, x) < \delta \implies d_Y(f(x_0), f(x)) < \varepsilon.$$

(Die Größe von δ hängt von ε und von x_0 ab).

2. f heißt folgenstetig im Punkt $x_0 \in X$, falls für jede Folge (x_n) in X , die gegen x_0 konvergiert, die Bildfolge $(f(x_n))$ in Y gegen $f(x_0)$ konvergiert.
3. Die Abbildung f heißt stetig (folgenstetig), wenn sie in jedem Punkt $x_0 \in X$ stetig (folgenstetig) ist.



Satz 4.7 Sei $f : X \rightarrow Y$ eine Abbildung zwischen metrischen Räumen. Dann sind folgende Bedingungen äquivalent:

1. $f : X \rightarrow Y$ ist in $x_0 \in X$ stetig.

2. $f : X \rightarrow Y$ ist in $x_0 \in X$ folgenstetig.

3. x_0 ist ein isolierter Punkt von X oder x_0 ist ein Häufungspunkt von X und es gilt

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

Beweis. (1) \implies (2) : Sei $f : X \rightarrow Y$ stetig in $x_0 \in X$ und (x_n) eine (beliebige) gegen x_0 konvergente Folge in X . Wir müssen zeigen, dass die Folge $(f(x_n))$ gegen $f(x_0)$ konvergiert. Sei $\varepsilon > 0$. Nach Definition der Stetigkeit existiert ein $\delta > 0$, so dass

$$f(K_X(x_0, \delta)) \subset K_Y(f(x_0), \varepsilon).$$

Da die Folge (x_n) gegen x_0 konvergiert, existiert ein $n_0 \in \mathbb{N}$, so dass $x_n \in K_X(x_0, \delta)$ für alle $n \geq n_0$. Folglich ist $f(x_n) \in K_Y(f(x_0), \varepsilon)$ für alle $n \geq n_0$. Also konvergiert $(f(x_n))$ gegen $f(x_0)$.

(2) \implies (3) : Jeder Punkt $x_0 \in X$ ist entweder isoliert oder ein Häufungspunkt von X . Ist $x_0 \in \text{HP}(X)$, so existiert eine gegen x_0 konvergente Folge (x_n) in $X \setminus \{x_0\}$. Nach (2) konvergiert dann für jede dieser Folgen die Bildfolge $(f(x_n))$ gegen $f(x_0)$, das heißt es gilt

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

(3) \implies (1) : Wenn x_0 ein isolierter Punkt ist, so existiert ein $\delta > 0$, so dass $K_X(x_0, \delta) = \{x_0\}$. Dann ist

$$f(K_X(x_0, \delta)) = \{f(x_0)\} \subset K_Y(f(x_0), \varepsilon)$$

für alle $\varepsilon > 0$. Somit ist f in x_0 stetig.

Sei nun x_0 ein Häufungspunkt von X und $\lim_{x \rightarrow x_0} f(x) = f(x_0)$. Nach Satz 4.1 existiert für alle $\varepsilon > 0$ ein $\delta > 0$, so dass

$$f(K_X(x_0, \delta)) \subset K_Y(f(x_0), \varepsilon).$$

Folglich ist f in x_0 stetig. □

Als nächstes betrachten wir ein spezielles Kriterium für die Stetigkeit reeller Funktionen².

Satz 4.8 Sei $f : A \subset \mathbb{R} \rightarrow \mathbb{R}$ eine reelle Funktion und $x_0 \in (\alpha, \beta) \subset A$. Dann ist f in x_0 genau dann stetig, wenn die beiden einseitigen Grenzwerte in x_0 existieren und

$$\lim_{x \rightarrow x_0^-} f(x) = \lim_{x \rightarrow x_0^+} f(x) = f(x_0).$$

Beweis. (\implies) folgt aus Satz 4.7 als Spezialfall.

(\impliedby) Wir setzen voraus, dass die beiden einseitigen Grenzwerte existieren und dass

$$\lim_{x \rightarrow x_0^-} f(x) = \lim_{x \rightarrow x_0^+} f(x) = f(x_0).$$

Angenommen, f sei in x_0 nicht stetig. Dann ist f in x_0 nicht folgenstetig, d.h. es existiert eine Folge (a_n) in A , die gegen x_0 konvergiert, deren Bildfolge aber nicht gegen $f(x_0)$

² Auf Teilmengen $A \subset \mathbb{R}$ betrachten wir im Folgenden ebenfalls die durch den Betrag definierte Metrik.

konvergiert. Da dann fast alle Folgenglieder a_n von x_0 verschieden sind, existiert eine Teilfolge (a_{n_k}) von (a_n) mit $a_{n_k} < x_0$ oder eine Teilfolge (a_{n_k}) von (a_n) mit $a_{n_k} > x_0$, deren Bildfolge $(f(a_{n_k}))$ 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 $f : A \subset \mathbb{R} \rightarrow \mathbb{R}$ ist genau dann in $x_0 \in (\alpha, \beta) \subset A$ *unstetig*, wenn einer der folgenden Fälle vorliegt:

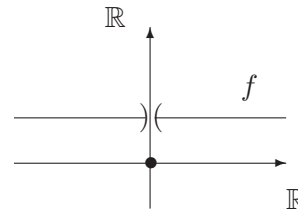
(1) *Hebbare Unstetigkeitsstelle*

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*

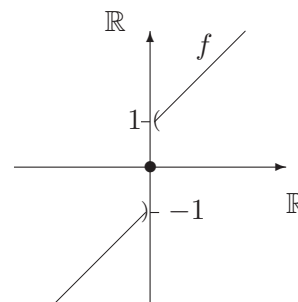
Die beiden einseitigen Grenzwerte von f in x_0 existieren in \mathbb{R} , sind aber voneinander verschieden. Unter dem *Sprung* $\sigma(f, x_0)$ von f in x_0 verstehen wir die Differenz der einseitigen Grenzwerte

$$\sigma(f, x_0) := \lim_{x \rightarrow x_0^+} f(x) - \lim_{x \rightarrow x_0^-} f(x).$$

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. Art*³

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

³ Hebbare Unstetigkeitsstellen und Sprungstellen nennt man auch Unstetigkeitsstellen 1. Art.

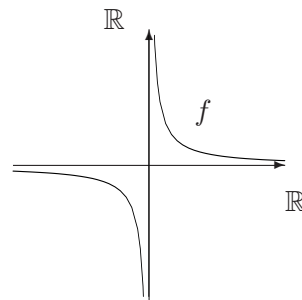
$$h(x) := \begin{cases} 1 & x \in \mathbb{Q} \\ 0 & x \notin \mathbb{Q} \end{cases}$$

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 $\lim_{x \rightarrow 0^-} f(x) = -\infty$.



Satz 4.9

1. Eine monotone Funktion $f : A \subset \mathbb{R} \rightarrow \mathbb{R}$ ist in $x_0 \in (\alpha, \beta) \subset A$ genau dann stetig, wenn

$$\lim_{x \rightarrow x_0^+} f(x) = \lim_{x \rightarrow x_0^-} f(x).$$

2. Eine monotone Funktion $f : (\alpha, \beta) \subset \mathbb{R} \rightarrow \mathbb{R}$, $-\infty \leq \alpha < \beta \leq +\infty$, hat höchstens abzählbar viele Unstetigkeitsstellen. Jede Unstetigkeitsstelle ist eine Sprungstelle.

3. Sei $f : (\alpha, \beta) \subset \mathbb{R} \rightarrow \mathbb{R}$ eine streng monotone Funktion, $-\infty \leq \alpha < \beta \leq +\infty$. Dann ist $f : (\alpha, \beta) \rightarrow f((\alpha, \beta))$ bijektiv und die Umkehrabbildung $f^{-1} : f((\alpha, \beta)) \rightarrow (\alpha, \beta)$ ist stetig.

Beweis. Übungsaufgabe (Benutze Satz 4.6). □

Als nächstes stellen wir einige Eigenschaften stetiger Abbildungen zwischen metrischen Räumen zusammen, die sich unmittelbar aus der Definition der Stetigkeit, der Äquivalenz zur Folgenstetigkeit und den Rechenregeln für konvergente Folgen ergeben.

Satz 4.10 Seien $f : X \rightarrow Y$ und $g : Y \rightarrow Z$ zwei Abbildungen zwischen metrischen Räumen, f in $x_0 \in X$ und g in $f(x_0) \in Y$ stetig. Dann ist $g \circ f$ in x_0 stetig. Insbesondere ist die Verknüpfung stetiger Abbildungen ebenfalls stetig.

Beweis. Sei (x_n) eine Folge in X , die gegen x_0 konvergiert. Da f in x_0 stetig und somit folgenstetig ist, konvergiert die Folge $(f(x_n))$ in Y gegen $f(x_0)$. Wegen der Stetigkeit von g in $f(x_0)$ konvergiert dann die Folge $(g(f(x_n)) = (g \circ f)(x_n))$ gegen $g(f(x_0)) = (g \circ f)(x_0)$. Folglich ist $g \circ f$ in x_0 folgenstetig und somit stetig. □

Satz 4.11 Seien (X, d_X) und (Y, d_Y) metrische Räume und A eine Teilmenge von X . Ist $f : X \rightarrow Y$ stetig, so ist auch die Einschränkung $f|_A : A \rightarrow Y$ stetig bzgl. der auf A durch d_X induzierten Metrik $d_A := (d_X)|_{A \times A}$.

Beweis. Sei (a_n) eine Folge in A , die im metrischen Raum (A, d_A) gegen $a \in A$ konvergiert. Dann konvergiert (a_n) auch bzgl. der Metrik d_X von X gegen a . Nach Voraussetzung konvergiert $(f(a_n))$ gegen $f(a)$ in Y , also ist $f|_A$ in a folgenstetig und somit stetig. \square

Satz 4.12

1. Sei $f : X \rightarrow Y_1 \times \dots \times Y_n$ eine Abbildung in das Produkt metrischer Räume und $f = (f_1, \dots, f_n)$ die Komponentendarstellung von f . Die Abbildung f ist genau dann in $x_0 \in X$ stetig, wenn jede Komponente $f_j : X \rightarrow Y_j$, $j = 1, \dots, n$, in x_0 stetig ist.
2. Sei $f : X_1 \times X_2 \rightarrow Y$ stetig in $(x_1, x_2) \in X_1 \times X_2$. Dann ist

$$\begin{aligned} h_1 : X_1 &\rightarrow Y && \text{in } x_1 \text{ stetig,} \\ x &\mapsto f(x, x_2) \\ h_2 : X_2 &\rightarrow Y && \text{in } x_2 \text{ stetig.} \\ x &\mapsto f(x_1, x) \end{aligned}$$

Satz 4.13 Sei X ein metrischer Raum und $(V, \|\cdot\|)$ ein normierter Vektorraum über dem Körper \mathbb{K} der reellen oder der komplexen Zahlen.

1. Seien $f, g : X \rightarrow V$ in $x_0 \in X$ stetig. Dann gilt:
 - (a) $f + g : X \rightarrow V$ ist in x_0 stetig.
 - (b) $\|f\| : X \rightarrow \mathbb{R}$ ist in x_0 stetig.
 - (c) Sei $\langle \cdot, \cdot \rangle$ ein Skalarprodukt auf V , das die Norm $\|\cdot\|$ induziert. Dann ist $\langle f, g \rangle : X \rightarrow \mathbb{K}$ in x_0 stetig.
2. Sind $f : X \rightarrow V$ und $h : X \rightarrow \mathbb{K}$ in x_0 stetig, so ist auch $h \cdot f : X \rightarrow V$ in x_0 stetig.
3. Sind $h, p : X \rightarrow \mathbb{K}$ in x_0 stetig und $p(x_0) \neq 0$, dann ist auch die Abbildung $\frac{h}{p} : A = \{x \in X \mid p(x) \neq 0\} \subset X \rightarrow \mathbb{K}$ in x_0 stetig.
4. Eine Abbildung $f : X \rightarrow \mathbb{C}$ ist genau dann in x_0 stetig, wenn sowohl der Realteil $\operatorname{Re}(f)$ als auch der Imaginarteil $\operatorname{Im}(f)$ in x_0 stetig sind.
Die Abbildung $f : X \rightarrow \mathbb{C}$ ist genau dann in x_0 stetig, wenn die konjugiert-komplexe Abbildung $\bar{f} : X \rightarrow \mathbb{C}$ in x_0 stetig ist.

Wir betrachten nun einige Beispiele für stetige Abbildungen:

Beispiel 1: Sei (X, d) ein metrischer Raum und $x_0 \in X$.

- (a) Die identische Abbildung $\operatorname{id}_X : X \rightarrow X$, $\operatorname{id}_X(x) := x$, ist stetig.
- (b) Die konstante Abbildung $c_{x_0} : X \rightarrow X$, $c_{x_0}(x) := x_0$, ist stetig.

Beispiel 2: *Rationale Funktionen*

Seien P_1 und P_2 zwei Polynome mit komplexen Koeffizienten vom Grad ≥ 0 , d.h.

$$\begin{aligned} P_1(z) &= a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 \\ P_2(z) &= b_m z^m + b_{m-1} z^{m-1} + \dots + b_1 z + b_0 \end{aligned}$$

wobei alle a_j und b_j komplexe Zahlen sind. Dann ist die *rationale* Funktion

$$f := \frac{P_1}{P_2} : A := \{z \in \mathbb{C} \mid P_2(z) \neq 0\} \subset \mathbb{C} \longrightarrow \mathbb{C}$$

$$f(z) := \frac{a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0}{b_m z^m + b_{m-1} z^{m-1} + \dots + b_1 z + b_0}$$

auf A stetig.

Beispiel 3: Die Abbildungen $\ln : \mathbb{R}^+ \longrightarrow \mathbb{R}$ und $\exp|_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}^+$ sind stetig.

Dies folgt aus Satz 4.9, denn beide Funktionen sind Umkehrfunktionen einer auf \mathbb{R} bzw. \mathbb{R}^+ definierten streng monotonen Funktion.

Beispiel 4: Die Abbildung

$$\begin{aligned} f : \mathbb{R}^+ \times \mathbb{C} &\longrightarrow \mathbb{C} \\ (a, z) &\longmapsto a^z \end{aligned}$$

ist stetig.

Beweis. Sei $(a_0, z_0) \in \mathbb{R}^+ \times \mathbb{C}$. Wir betrachten eine beliebige gegen (a_0, z_0) konvergente Folge $((a_n, z_n))$. Dann gilt $a_n \rightarrow a_0$ und $z_n \rightarrow z_0$. Da nach der natürliche Logarithmus stetig ist, folgt

$$\lim_{n \rightarrow \infty} \ln(a_n) \cdot z_n = \ln(a_0) \cdot z_0.$$

Aus Satz 3.19 erhalten wir dann für n hinreichend groß

$$\begin{aligned} \left| a_n^{z_n} - a_0^{z_0} \right| &= \left| e^{\ln(a_n) \cdot z_n} - e^{\ln(a_0) \cdot z_0} \right| \\ &= \left| e^{\ln(a_0) \cdot z_0} \left| e^{\ln(a_n) \cdot z_n - \ln(a_0) \cdot z_0} - 1 \right| \right| \\ &\leq \left| e^{\ln(a_0) \cdot z_0} \right| \frac{|\ln(a_n) \cdot z_n - \ln(a_0) \cdot z_0|}{1 - |\ln(a_n) \cdot z_n - \ln(a_0) \cdot z_0|} \quad (*) \end{aligned}$$

Die rechte Seite von (*) konvergiert bei $n \rightarrow +\infty$ gegen 0 und somit gilt $\lim_{n \rightarrow \infty} a_n^{z_n} = a_0^{z_0}$. Folglich ist f in (a_0, z_0) folgenstetig, also stetig. \square

Beispiel 5: Seien $a_0 \in \mathbb{R}^+$ und $z_0 \in \mathbb{C}$ fixiert. Aus Satz 4.12 und Beispiel 4 folgt die Stetigkeit der Abbildungen

$$\begin{aligned} \exp_{a_0} : \mathbb{C} &\longrightarrow \mathbb{C} & \text{und} & & p_{z_0} : \mathbb{R}^+ &\longrightarrow \mathbb{C} \\ z &\longmapsto a_0^z & & & a &\longmapsto a^{z_0}. \end{aligned}$$

Insbesondere sind die Wurzelfunktionen $\sqrt[n]{\cdot} : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$ stetig.

Beispiel 6: Sei (X, d) ein metrischer Raum. Die Abstandsfunktion

$$\begin{aligned} d : X \times X &\longrightarrow \mathbb{R} \\ (x, y) &\longmapsto d(x, y) \end{aligned}$$

ist stetig.

Beweis. Sei $(x, y) \in X \times X$ und $((x_n, y_n))$ eine Folge, die in $X \times X$ gegen (x, y) konvergiert. In einem metrischen Raum gilt die Vierecksungleichung

$$|d(x_n, y_n) - d(x, y)| \leq d(x_n, x) + d(y_n, y).$$

Da $x_n \rightarrow x$ und $y_n \rightarrow y$, folgt $d(x_n, x) \rightarrow 0$ und $d(y_n, y) \rightarrow 0$ und somit $d(x_n, y_n) \rightarrow d(x, y)$. \square

Beispiel 7: Sei V ein Vektorraum über dem Körper \mathbb{K} der reellen oder der komplexen Zahlen. Dann ist jede Norm $\|\cdot\| : V \rightarrow \mathbb{R}$ und jedes Skalarprodukt $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{K}$ stetig bezüglich der von der Norm bzw. vom Skalarprodukt auf V induzierten Metrik.

Beweis. Dies folgt aus dem Verhalten des Skalarproduktes und der Norm bei konvergenten Folgen (Satz 2.16), welches die Folgenstetigkeit der beiden Abbildungen zeigt. \square

Als nächstes definieren wir zwei Stetigkeitsbegriffe, die stärker als die gewöhnliche Stetigkeit sind.

Definition 4.4.

1. Eine Abbildung $f : X \rightarrow Y$ heißt *gleichmäßig stetig*, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass

$$f(K_X(x, \delta)) \subset K_Y(f(x), \varepsilon)$$

für alle $x \in X$, d.h. so dass gilt:

$$x, \tilde{x} \in X \text{ mit } d_X(x, \tilde{x}) < \delta \implies d_Y(f(x), f(\tilde{x})) < \varepsilon.$$

(Im Unterschied zur Definition der Stetigkeit hängt hier die Größe von δ nur von ε , aber nicht von x ab.)

2. Eine Abbildung $f : X \rightarrow Y$ heißt *lipschitzstetig*, wenn es eine positive Konstante $L \in \mathbb{R}^+$ gibt, so dass für alle $x_1, x_2 \in X$

$$d_Y(f(x_1), f(x_2)) \leq L \cdot d_X(x_1, x_2)$$

gilt. L heißt *Lipschitz-Konstante* von f .

Satz 4.14 Sei $f : X \rightarrow Y$ 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. Stetigkeit folgt per Definition aus gleichmäßiger Stetigkeit. Wir müssen also nur zeigen, dass jede Lipschitzstetige Abbildung gleichmäßig stetig ist.

Sei f Lipschitzstetig mit Lipschitz-Konstante L und $\varepsilon > 0$. Wir setzen $\delta := \frac{\varepsilon}{L}$. Seien nun $x_1, x_2 \in X$ mit $d_X(x_1, x_2) < \delta = \frac{\varepsilon}{L}$. Aus der Lipschitzstetigkeit folgt dann

$$d_Y(f(x_1), f(x_2)) \leq L \cdot d_X(x_1, x_2) < L \cdot \delta = \varepsilon.$$

Somit ist f gleichmäßig stetig. \square

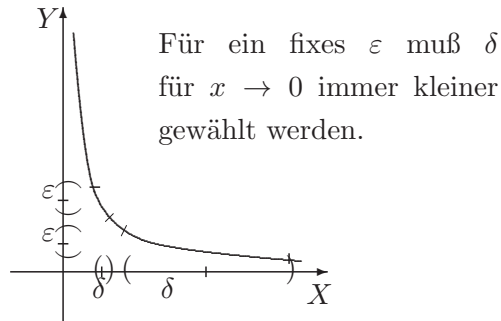
Die folgenden beiden Beispiele zeigen, dass die Umkehrungen der Aussagen des Satzes 4.14 nicht gelten.

Beispiel 1:

Die Abbildung

$$\begin{aligned} f : \mathbb{R}^+ &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{x} \end{aligned}$$

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. \square

Beispiel 2: Die Abbildung

$$\begin{aligned} f : \mathbb{R}^+ &\longrightarrow \mathbb{R}^+ \\ x &\longmapsto \sqrt{x} \end{aligned}$$

ist gleichmäßig stetig, aber nicht Lipschitzstetig.

Beweis. Es gilt

$$|\sqrt{x} - \sqrt{y}| \leq \sqrt{|x - y|} \quad \forall x, y \in \mathbb{R}^+.$$

(siehe Übungsaufgabe 10). 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. \square

4.3 Eigenschaften stetiger Abbildungen

In einem metrischen Raum sind Teilmengen mit speziellen Eigenschaften ausgezeichnet, wie offene, abgeschlossene, kompakte oder zusammenhängende Teilmengen. Wir untersuchen in diesem Abschnitt, was mit diesen Eigenschaften bei stetigen Abbildungen passiert.

Satz 4.15 *Sei $f : X \rightarrow Y$ eine Abbildung zwischen metrischen Räumen. Dann sind folgende Bedingungen äquivalent:*

1. f ist stetig.
2. Das Urbild jeder offenen Menge ist offen.
3. Das Urbild jeder abgeschlossenen Menge ist abgeschlossen.

Beweis. (1. \implies 2.) Sei $f : X \rightarrow Y$ stetig und $U \subset Y$ offen. Wir wollen 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. Sei also $f^{-1}(U) \neq \emptyset$ und $x_0 \in f^{-1}(U)$ ein beliebig gewählter Punkt. Dann gilt $f(x_0) \in U$, und da $U \subset Y$ offen ist, existiert ein $\varepsilon > 0$ mit $K_Y(f(x_0), \varepsilon) \subset U$. Da f in x_0 stetig ist, gibt es ein $\delta > 0$, so dass

$$f(K_X(x_0, \delta)) \subset K_Y(f(x_0), \varepsilon) \subset U.$$

Aus der Definition des Urbildes folgt dann:

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

Somit ist $f^{-1}(U)$ offen in X .

(2. \implies 1.) Sei $f^{-1}(U) \subset X$ offen für jede offene Menge $U \subset Y$. Wir wählen einen beliebigen Punkt $x_0 \in X$ und ein beliebiges $\varepsilon > 0$. Da die Kugel $K_Y(f(x_0), \varepsilon)$ in Y offen ist, ist nach Voraussetzung auch das Urbild $f^{-1}(K_Y(f(x_0), \varepsilon))$ in X offen. Es enthält den Punkt x_0 . Folglich existiert ein $\delta > 0$, so dass $K_X(x_0, \delta) \subset f^{-1}(K_Y(f(x_0), \varepsilon))$. Nach Definition des Urbildes folgt daraus $f(K_X(x_0, \delta)) \subset K_Y(f(x_0), \varepsilon)$. Das bedeutet aber, dass f in x_0 stetig ist. Da $x_0 \in X$ beliebig gewählt war, ist f stetig.

(2. \iff 3.) Um dies einzusehen, erinnern wir uns daran, dass die folgende Beziehung zwischen offenen und abgeschlossenen Mengen eines metrischen Raumes X gilt:

$B \subset X$ ist abgeschlossen $\iff X \setminus B \subset X$ ist offen .

Desweiteren gilt für das Urbild einer Komplementmenge

$$f^{-1}(Y \setminus M) = X \setminus f^{-1}(M).$$

Sei nun das Urbild $f^{-1}(U)$ jeder offenen Menge $U \subset Y$ offen. Wir betrachten eine beliebige abgeschlossene Menge $A \subset Y$. Dann ist $Y \setminus A$ offen in Y und folglich $f^{-1}(Y \setminus A) = X \setminus f^{-1}(A)$ offen in X . Dies bedeutet, dass $f^{-1}(A)$ abgeschlossen in X ist. Damit ist 2. \implies 3. bewiesen. Die Umkehrung zeigt man analog. \square

Wir betrachten als nächstes das Verhalten von zusammenhängenden Mengen bei stetigen Abbildungen:

Satz 4.16 *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 4.15 die Urbilder $f^{-1}(U)$ und $f^{-1}(V)$ ebenfalls offen. Desweiteren gilt für diese Urbilder

- $f^{-1}(U) \cap f^{-1}(V) = f^{-1}(U \cap V) = f^{-1}(\emptyset) = \emptyset$,
- $A \subset f^{-1}(f(A)) \subset f^{-1}(U \cup V) = f^{-1}(U) \cup f^{-1}(V)$,
- $A \cap f^{-1}(U) \neq \emptyset$ und $A \cap f^{-1}(V) \neq \emptyset$, da ein $a \in A$ mit $f(a) \in U$ und ein $b \in A$ mit $f(b) \in V$ existieren.

Folglich ist A nicht zusammenhängend, was ein Widerspruch zur Voraussetzung ist. \square

Als Anwendung erhalten wir für Abbildungen mit reellen Werten:

Satz 4.17 (Zwischenwertsatz)

Sei $f : X \rightarrow \mathbb{R}$ eine stetige Abbildung von einem metrischen Raum X in die reellen Zahlen und sei $A \subset X$ eine zusammenhängende Teilmenge. Seien desweiteren $a, b \in f(A)$ Bildpunkte mit $a < b$. Dann gilt $[a, b] \subset f(A)$.

Insbesondere hat jede reellwertige Abbildung, die auf A einen negativen und einen positiven Wert annimmt, auf A eine Nullstelle, d.h. es existiert ein $x_0 \in A$ mit $f(x_0) = 0$.

Beweis. Nach Satz 4.16 ist $f(A) \subset \mathbb{R}$ zusammenhängend und deshalb nach Satz 2.38 ein (verallgemeinertes) Intervall. Das heißt mit $a, b \in f(A)$ und $a < b$ ist auch $[a, b] \subset f(A)$. \square

Wir definieren nun noch einen weiteren Zusammenhangsbegriff.

Definition 4.5. Sei (X, d) ein metrischer Raum. Eine Teilmenge $A \subset X$ heißt *bogenzusammenhängend* (oder *wegzusammenhängend*), falls zu je zwei Punkten $a, b \in A$ eine stetige Abbildung $\omega : [0, 1] \subset \mathbb{R} \rightarrow A \subset X$ existiert, so dass $\omega(0) = a$ und $\omega(1) = b$.

Die Abbildung ω heißt *Weg in A von a nach b*.

Beispiele:

1. Konvexe Mengen im \mathbb{R}^n :

Eine Teilmenge $A \subset \mathbb{R}^n$ 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 bogenzusammenhängend, da die Abbildung $\omega : [0, 1] \rightarrow A$

$$\omega(t) := x + t(y - x)$$

stetig ist und die Punkte x und y verbindet.

Zum Beispiel ist jede Kugel $K(x_0, r)$ konvex.

2. Sternförmige Mengen im \mathbb{R}^n :

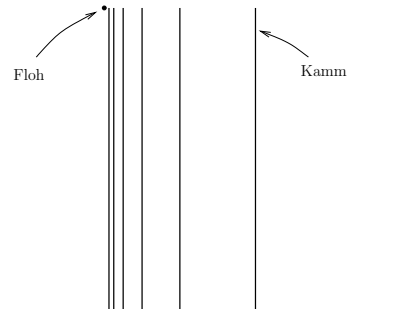
Eine Teilmenge $A \subset \mathbb{R}^n$ heißt *sternförmig*, wenn ein Punkt $x_0 \in A$ existiert, so dass für jeden Punkt $x \in A$ die Verbindungsstrecke $\overline{x_0x}$ vollständig in A liegt. Jede sternförmige Menge ist bogenzusammenhängend. Sind $x, y \in A$, so ist die Vereinigung der Strecken $\overline{x_0x} \cup \overline{x_0y}$ in A enthalten und man kann sie als Bild eines Weges von x nach y darstellen.

3. Jede bogenzusammenhängende Menge ist auch zusammenhängend (Übungsaufgabe).

Die Umkehrung gilt nicht. Ein Beispiel dafür ist die folgende Menge in der komplexen Ebene (*Floh und Kamm*):

$$A := [0, 1] \cup \left\{ \frac{1}{n} + iy \mid n \in \mathbb{N} \text{ und } y \in [0, 1] \right\} \cup \{i\}.$$

Diese Menge ist zusammenhängend, aber nicht bogenzusammenhängend.



Es gilt jedoch, dass jede *offene* zusammenhängende Menge des \mathbb{R}^n auch bogenzusammenhängend ist.

Satz 4.18 Sei $f : X \rightarrow Y$ eine stetige Abbildung. Dann ist das Bild jeder bogenzusammenhängenden Teilmenge ebenfalls bogenzusammenhängend.

Beweis. Sei $A \subset X$ bogenzusammenhängend. Wir wollen zeigen, dass das Bild $f(A)$ ebenfalls bogenzusammenhängend 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 bogenzusammenhängend ist, gibt es einen (stetigen) Weg $\omega : [0, 1] \rightarrow A$ 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 kompakten Mengen bei stetigen Abbildungen:

Satz 4.19 *Sei $f : X \rightarrow Y$ eine stetige Abbildung. Dann ist das Bild jeder kompakten Menge ebenfalls kompakt.*

Beweis. Sei $K \subset X$ eine kompakte Teilmenge. Wir wollen zeigen, dass dann auch das Bild $f(K) \subset Y$ kompakt ist. Sei $\mathcal{U} = \{U_i\}_{i \in I}$ eine beliebige offene Überdeckung von $f(K)$, das heißt die Mengen U_i sind offen in Y und es gilt

$$f(K) \subset \bigcup_{i \in I} U_i.$$

Da f stetig ist, sind die Urbilder $f^{-1}(U_i)$ in X offen (siehe Satz 4.15). Außerdem gilt

$$K \subset f^{-1}\left(\bigcup_{i \in I} U_i\right) = \bigcup_{i \in I} f^{-1}(U_i).$$

Folglich ist die Mengenfamilie

$$\mathcal{U}^* := \{f^{-1}(U_i)\}_{i \in I}$$

eine offene Überdeckung von K . Da $K \subset X$ kompakt ist, existiert eine endliche Teilüberdeckung von K aus \mathcal{U}^* , das heißt es existieren $i_1, \dots, i_n \in I$, so dass

$$K \subset f^{-1}(U_{i_1}) \cup \dots \cup f^{-1}(U_{i_n}) = f^{-1}(U_{i_1} \cup \dots \cup U_{i_n}).$$

Aus der Definition des Urbildes folgt

$$f(K) \subset U_{i_1} \cup \dots \cup U_{i_n}.$$

Somit ist $\{U_{i_1}, \dots, U_{i_n}\}$ eine endliche Teilüberdeckung von $f(K)$ aus \mathcal{U} . Folglich ist $f(K)$ kompakt. \square

Die folgenden beiden Sätze sind Anwendungen der gerade bewiesenen Eigenschaft stetiger Abbildungen.

Satz 4.20 (Satz von Weierstraß)

Sei $f : X \rightarrow \mathbb{R}$ eine stetige Abbildung von einem metrischen Raum X in die reellen Zahlen und sei $K \subset X$ eine kompakte Teilmenge. Dann nimmt f auf K ein Maximum und ein Minimum an, d.h. es existieren Punkte $a, b \in K$ mit $f(a) = \min\{f(x) \mid x \in K\}$ und $f(b) = \max\{f(x) \mid x \in K\}$.

Beweis. Sei $K \subset X$ kompakt. Dann ist $f(K) \subset \mathbb{R}$ ebenfalls kompakt, und somit insbesondere beschränkt und abgeschlossen (Folgerung 2.4). Da $f(K) \subset \mathbb{R}$ beschränkt ist, existieren das Infimum $S_* := \inf(f(K))$ und das Supremum $S^* := \sup(f(K))$. Da $f(K)$ abgeschlossen ist, liegen sowohl das Supremum S^* als auch das Infimum S_* in $f(K)$. D.h. es existieren Punkte $a, b \in K$ mit $S_* = f(a) = \min\{f(x) \mid x \in K\}$ und $S^* = f(b) = \max\{f(x) \mid x \in K\}$. \square

Satz 4.21 (Satz von Heine)

Sei $f : X \rightarrow Y$ eine stetige Abbildung zwischen metrischen Räumen und $K \subset X$ eine kompakte Teilmenge. Dann ist f auf K sogar gleichmäßig stetig.

Beweis. Angenommen $f|_K : K \rightarrow Y$ wäre nicht gleichmäßig stetig. Dann existiert ein $\varepsilon_0 > 0$, so dass es für alle $\delta > 0$ einen Punkt $x_\delta \in K$ gibt mit

$$f(K_X(x_\delta, \delta) \cap K) \not\subset K_Y(f(x_\delta), \varepsilon_0).$$

Wir setzen $\delta = \frac{1}{n}$. Dann erhalten wir für jedes $n \in \mathbb{N}$ zwei Punkte $x_n, x_n^* \in K$ mit

$$d_X(x_n, x_n^*) < \frac{1}{n} \quad \text{und} \quad d_Y(f(x_n), f(x_n^*)) \geq \varepsilon_0.$$

Betrachten wir nun die Folgen (x_n) und (x_n^*) in K genauer. Da K kompakt, also nach Satz 2.37 folgenkompakt ist, existiert eine Teilfolge (x_{n_k}) von (x_n) , die gegen einen Punkt $x_0 \in K$ konvergiert. Wir zeigen, dass die Teilfolge $(x_{n_k}^*)$ von (x_n^*) ebenfalls gegen x_0 konvergiert. Es gilt

$$d(x_0, x_{n_k}^*) \leq \underbrace{d(x_0, x_{n_k})}_{\rightarrow 0} + \underbrace{d(x_{n_k}, x_{n_k}^*)}_{< \frac{1}{n_k}}.$$

Daher gilt $x_{n_k}^* \rightarrow x_0 \in K$. Da f stetig, also auch folgenstetig ist, konvergieren dann die Bildfolgen $(f(x_{n_k}))$ und $(f(x_{n_k}^*))$ für k gegen $+\infty$ gegen $f(x_0)$. Somit ist

$$0 < \varepsilon_0 \leq d(f(x_{n_k}), f(x_{n_k}^*)) \leq \underbrace{d(f(x_{n_k}), f(x_0))}_{\rightarrow 0} + \underbrace{d(f(x_0), f(x_{n_k}^*))}_{\rightarrow 0}.$$

Dies ist aber ein Widerspruch. \square

Wir wenden uns jetzt der folgenden Frage zu:

Sei $f : X \rightarrow Y$ eine *bijektive* stetige Abbildung zwischen metrischen Räumen. Dann existiert die Umkehrabbildung (inverse Abbildung) $f^{-1} : Y \rightarrow X$. Ist f^{-1} dann auch stetig? Im Allgemeinen gilt das nicht.

Beispiel: Sei $X = (0, 1) \cup \{2\} \subset \mathbb{R}$ und $Y = (0, 1] \subset \mathbb{R}$. 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, da $f|_{(0,1)} = id_{(0,1)}$ und $2 \in X$ ein isolierter Punkt ist. Aber $f^{-1} : (0, 1] \rightarrow (0, 1) \cup \{2\}$ ist nicht stetig, da $(0, 1]$ zusammenhängend, aber $f^{-1}((0, 1]) = (0, 1) \cup \{2\}$ nicht zusammenhängend ist.

Definition 4.6. 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.

Bezeichnung: $X \simeq Y$.

Beispiel: Der \mathbb{R}^n ist homöomorph zur Kugel $K(x_0, r)$ um einen Punkt $x_0 \in \mathbb{R}^n$ vom Radius r . Die folgende Abbildung ist ein Homöomorphismus:

$$f : K(x_0, r) \rightarrow \mathbb{R}^n \\ x \mapsto \frac{x - x_0}{r - \|x - x_0\|}$$

Dabei bezeichnet $\|\cdot\|$ die Euklidische Norm auf \mathbb{R}^n (Übungsaufgabe).

Alle topologischen Eigenschaften von Teilmengen bleiben bei Homöomorphismen erhalten:

Satz 4.22 Sei $f : X \rightarrow Y$ ein Homöomorphismus zwischen metrischen Räumen. Dann gilt:

1. $A \subset X$ ist offen (abgeschlossen, kompakt, zusammenhängend bzw. bogenzusammenhängend) genau dann, wenn $f(A) \subset Y$ offen (abgeschlossen, kompakt, zusammenhängend bzw. bogenzusammenhängend) ist.
2. Eine Folge (x_n) konvergiert in X gegen x_0 genau dann, wenn die Bildfolge $(f(x_n))$ in Y gegen $f(x_0)$ konvergiert.

Beweis. Dies folgt aus den vorherigen Sätzen über das Verhalten von offenen, abgeschlossenen, kompakten, zusammenhängenden bzw. bogenzusammenhängenden Mengen bei stetigen Abbildungen und aus der Äquivalenz von Stetigkeit und Folgenstetigkeit. \square

Will man nachweisen, dass eine bijektive stetige Abbildung ein Homöomorphismus ist, so muß man die Stetigkeit der inversen Abbildung nachweisen. Ist der Urbildraum kompakt, so kann man sich diesen Nachweis sparen, denn es gilt der folgende Satz:

Satz 4.23 (Satz über die Stetigkeit der inversen Abbildung)

Sei $f : X \rightarrow Y$ eine bijektive stetige Abbildung und X ein kompakter metrischer Raum. Dann ist die inverse Abbildung $f^{-1} : Y \rightarrow X$ stetig.

Beweis. Sei $U \subset X$ offen. Wir zeigen, dass $(f^{-1})^{-1}(U) = f(U) \subset Y$ offen ist. Da $U \subset X$ offen ist, ist $X \setminus U \subset X$ abgeschlossen. Wir wissen aus Folgerung 2.4, dass jede abgeschlossene Teilmenge einer kompakten Menge selbst kompakt ist. Nach Voraussetzung ist X kompakt, folglich ist $X \setminus U \subset X$ kompakt. Wegen der Stetigkeit von f ist dann nach Satz

4.19 auch das Bild $f(X \setminus U)$ in Y kompakt und insbesondere abgeschlossen (Folgerung 2.4). Nun ist f bijektiv und folglich $f(X \setminus U) = Y \setminus f(U)$. Also ist die Teilmenge $f(U) \subset Y$ offen. Nach Satz 4.15 ist die Abbildung f^{-1} deshalb stetig. \square

Definition 4.7. Sei X 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

$$d(f(x), f(y)) \leq L \cdot d(x, y) \quad \text{für alle } x, y \in X.$$

Für *kontrahierende* Abbildungen in *vollständigen* metrischen Räumen gilt der folgende sehr nützliche Fixpunktsatz.

Satz 4.24 (Banachscher Fixpunktsatz)

Sei (X, d) ein vollständiger metrischer Raum und $f : X \rightarrow X$ eine kontrahierende Abbildung. Dann existiert genau ein Fixpunkt von f , d.h. genau ein Punkt $x^* \in X$ mit $f(x^*) = x^*$. Diesen Fixpunkt erhält man konstruktiv auf die folgende Weise:

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 x^* .

Beweis. (1) Eindeutigkeit des Fixpunktes: Angenommen es existieren zwei Fixpunkte x^* und y^* von f , dh. es gelte $f(x^*) = x^*$ und $f(y^*) = y^*$. Aus der Kontraktivität von f erhält man

$$d(x^*, y^*) = d(f(x^*), f(y^*)) \leq L \cdot d(x^*, y^*).$$

Da aber $0 < L < 1$ gilt, folgt $d(x^*, y^*) = 0$, also $x^* = y^*$.

(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

$$\begin{aligned}
d(x_n, x_m) &\leq d(x_n, x_{n-1}) + d(x_{n-1}, x_{n-2}) + \dots + d(x_{m+1}, x_m) \\
&\leq (L^{n-1} + \dots + L^m) \cdot d(x_0, x_1) \\
&= L^m (L^0 + L^1 + \dots + L^{n-1-m}) \cdot d(x_0, x_1) \\
&= L^m \cdot \frac{1 - L^{n-m}}{1 - L} \cdot d(x_0, x_1) \\
&< L^m \cdot \frac{1}{1 - L} \cdot d(x_0, x_1).
\end{aligned}$$

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 x^* aus X . Wir zeigen nun, dass dieser Grenzwert x^* der gesuchte Fixpunkt von f ist. Dazu betrachten wir

$$\begin{aligned}
d(f(x^*), x^*) &\leq d(f(x^*), x_n) + d(x_n, x^*) \\
&= d(f(x^*), f(x_{n-1})) + d(x_n, x^*) \\
&\leq L \underbrace{d(x^*, x_{n-1})}_{\rightarrow 0} + \underbrace{d(x_n, x^*)}_{\rightarrow 0}.
\end{aligned}$$

Daraus folgt $d(f(x^*), x^*) = 0$, also $f(x^*) = x^*$. \square

4.4 Folgen stetiger Abbildungen

In diesem Abschnitt bezeichnen X und Y wiederum metrische Räume. Wir betrachten jetzt Folgen von stetigen Abbildungen $f_n : X \rightarrow Y$, $n = 1, 2, 3, \dots$, und untersuchen, unter welchen Bedingungen deren Grenzfunktion ebenfalls stetig ist.

Definition 4.8. Die Folge der Abbildungen (f_n) heißt *punktweise konvergent*, falls für jedes $x \in X$ die Folge $(f_n(x))$ in Y konvergiert. Die durch

$$f(x) := \lim_{n \rightarrow \infty} f_n(x)$$

definierte Funktion $f : X \rightarrow Y$ heißt *Grenzfunktion* von (f_n) .

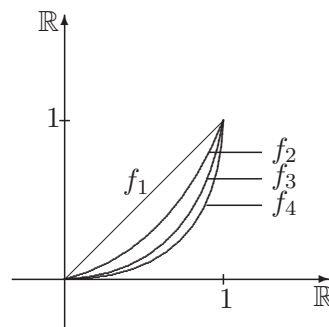
Schreibweise: $f_n \rightarrow f$.

Im allgemeinen ist die Grenzfunktion einer punktweise konvergenten Folge stetiger Abbildungen nicht notwendiger Weise stetig. Wir betrachten dazu das folgende Beispiel:

Seien $X = [0, 1]$, $Y = \mathbb{R}$ und $f_n : [0, 1] \rightarrow \mathbb{R}$ die stetigen Funktionen $f_n(x) := x^n$. Dann ist die Grenzfunktion

$$f(x) := \lim_{n \rightarrow \infty} f_n(x) = \begin{cases} 0 & \text{falls } 0 \leq x < 1 \\ 1 & \text{falls } x = 1 \end{cases}$$

offensichtlich nicht stetig.



Wir benötigen für die Stetigkeit der Grenzfunktion f eine stärkere Konvergenzeigenschaft der Folge (f_n) .

Definition 4.9. Die Folge der Abbildungen (f_n) konvergiert gleichmäßig gegen eine Grenzfunktion $f : X \rightarrow Y$, falls zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, so dass

$$d_Y(f_n(x), f(x)) < \varepsilon \quad \text{für alle } n \geq n_0 \text{ und für alle } x \in X.$$

Schreibweise: $f_n \rightrightarrows f$.

Bei gleichmäßiger Konvergenz, konvergiert die Folge $(f_n(x))$ für jeden Punkt $x \in X$ mit der "gleichen Geschwindigkeit" gegen $f(x)$.

Aus gleichmäßiger Konvergenz folgt offensichtlich die Konvergenz der Folge (f_n) .

Satz 4.25 Seien $f_n : X \rightarrow Y$ stetige Abbildungen, $n = 1, 2, 3, \dots$, und die Folge (f_n) gleichmäßig gegen $f : X \rightarrow Y$ konvergent. Dann ist die Grenzfunktion f stetig.

Beweis. Sei $x_0 \in X$ 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

$$d_Y(f_n(x), f(x)) < \frac{\varepsilon}{3} \quad \forall n \geq n_0, \forall x \in X.$$

Die Abbildung f_{n_0} ist in x_0 stetig, folglich gibt es ein $\delta > 0$, so dass

$$d_Y(f_{n_0}(x_0), f_{n_0}(x)) < \frac{\varepsilon}{3} \quad \forall x \in X \text{ mit } d_X(x_0, x) < \delta.$$

Für $x \in X$ mit $d_X(x_0, x) < \delta$ erhalten wir damit

$$\begin{aligned} d_Y(f(x), f(x_0)) &\leq d_Y(f(x), f_{n_0}(x)) + d_Y(f_{n_0}(x), f_{n_0}(x_0)) + d_Y(f_{n_0}(x_0), f(x_0)) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

Somit ist f in $x_0 \in X$ stetig. □

Satz 4.26 (Vertauschbarkeit von Grenzwerten)

Seien $f_n : X \rightarrow Y$, $n = 1, 2, 3, \dots$, stetige Abbildungen und sei die Folge (f_n) gleichmäßig konvergent. Dann gilt in jedem Häufungspunkt $x_0 \in X$

$$\lim_{x \rightarrow x_0} (\lim_{n \rightarrow \infty} f_n(x)) = \lim_{n \rightarrow \infty} (\lim_{x \rightarrow x_0} f_n(x)).$$

Beweis. Sei f die Grenzfunktion der Folge (f_n) , d.h. $f(x) = \lim_{n \rightarrow \infty} f_n(x)$. Nach Satz 4.25 ist f in x_0 stetig. Folglich existiert der Grenzwert

$$f(x_0) = \lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} \lim_{n \rightarrow \infty} f_n(x).$$

Da die Abbildungen f_n in x_0 stetig sind, gilt andererseits

$$f(x_0) = \lim_{n \rightarrow \infty} f_n(x_0) = \lim_{n \rightarrow \infty} \lim_{x \rightarrow x_0} f_n(x).$$

□

4.5 Funktionenreihen

In diesem Abschnitt wenden wir die im vorigen Abschnitt über Funktionenfolgen erhaltenen Resultate auf Funktionenreihen an. Insbesondere wollen wir weitere Aussagen über Potenzreihen machen.

Im folgenden bezeichnet (X, d) einen beliebigen metrischen Raum und $(E, \|\cdot\|)$ einen Banachraum. Seien $f_n : X \rightarrow E$ Abbildungen von X nach E , $n = 0, 1, 2, \dots$. Wir bilden aus der Folge (f_n) die Folge der Partialsummen (s_m) :

$$s_m := f_0 + f_1 + f_2 + \dots + f_m, \quad m \in \mathbb{N}_0.$$

Definition 4.10. Die Folge $(s_m := \sum_{n=0}^m f_n)$ heißt *Funktionenreihe* mit den Gliedern f_n . Wir bezeichnen sie symbolisch mit $\sum_{n=0}^{\infty} f_n$.

Die Funktionenreihe $\sum_{n=0}^{\infty} f_n$ heißt *punktweise* (bzw. *gleichmäßig*) *konvergent*, falls die Funktionenfolge (s_m) *punktweise* (bzw. *gleichmäßig*) *konvergiert*.

Aus den Sätzen 4.25 und 4.26 erhält man dann sofort

Satz 4.27 Sind $f_n : X \rightarrow E$ stetige Abbildungen, $n = 0, 1, 2, \dots$, und konvergiert die Funktionenreihe $\sum_{n=0}^{\infty} f_n$ gleichmäßig, so ist die durch

$$f(x) := \sum_{n=0}^{\infty} f_n(x) \quad \forall x \in X$$

definierte Grenzfunktion $f : X \rightarrow E$ stetig. Insbesondere gilt für jeden Häufungspunkt $x_0 \in X$

$$\lim_{x \rightarrow x_0} \sum_{n=0}^{\infty} f_n(x) = \sum_{n=0}^{\infty} \lim_{x \rightarrow x_0} f_n(x).$$

□

Als nächstes beweisen wir ein Kriterium für die gleichmäßige Konvergenz von Funktionenreihen.

Satz 4.28 (Weierstraßsches Majorantenkriterium)

Seien $f_n : X \rightarrow E$ beschränkte Abbildungen, $n \in \mathbb{N}_0$, M_n reelle Konstanten mit

$$\|f_n(x)\| \leq M_n \quad \forall n \in \mathbb{N}_0 \text{ und } \forall x \in X,$$

und sei die Reihe der Schranken $\sum_{n=0}^{\infty} M_n$ in \mathbb{R} konvergent. Dann ist die Funktionenreihe

$\sum_{n=0}^{\infty} f_n$ gleichmäßig konvergent. Sind alle Abbildungen $f_n : X \rightarrow E$ zusätzlich stetig, so ist

die durch $f(x) := \sum_{n=0}^{\infty} f_n(x)$ definierte Grenzfunktion $f : X \rightarrow E$ stetig.

Beweis. Sei $\varepsilon > 0$. Da die Reihe $\sum_{n=0}^{\infty} M_n$ konvergiert, gibt es nach dem Cauchy-Kriterium (Satz 3.1) ein $n_0 \in \mathbb{N}_0$, so dass

$$M_{m+1} + M_{m+2} + \cdots + M_k < \varepsilon \quad \text{für alle } k > m \geq n_0.$$

Sei $s_m := \sum_{n=0}^m f_n$ die m . Partialsumme der Folge (f_n) . Nach Voraussetzung gilt für alle $x \in X$

$$\begin{aligned} \|s_k(x) - s_m(x)\| &= \|f_{m+1}(x) + \cdots + f_k(x)\| \\ &\leq \|f_{m+1}(x)\| + \cdots + \|f_k(x)\| \\ &\leq M_{m+1} + \cdots + M_k. \end{aligned}$$

Wir erhalten daraus für jedes $x \in X$

$$\|s_k(x) - s_m(x)\| < \varepsilon \quad \forall k > m \geq n_0. \quad (*)$$

Also ist die Folge $(s_m(x))$ für jedes $x \in X$ eine Cauchy-Folge im Banachraum E . Da E vollständig ist, konvergiert die Folge $(s_m(x))$. Sei $f(x) := \lim_{m \rightarrow \infty} s_m(x)$. Die Funktionenreihe $\sum_{n=0}^{\infty} f_n$ konvergiert somit punktweise gegen die Grenzfunktion $f : X \rightarrow E$. Gehen wir in (*) mit k gegen ∞ , so folgt wegen der Stetigkeit der Norm

$$\lim_{k \rightarrow \infty} \|s_k(x) - s_m(x)\| = \left\| \lim_{k \rightarrow \infty} s_k(x) - s_m(x) \right\| = \|f(x) - s_m(x)\| \leq \varepsilon \quad \forall x \in X, \forall m > n_0.$$

Somit konvergiert die Funktionenreihe $\sum_{n=0}^{\infty} f_n$ sogar gleichmäßig gegen f . Die Stetigkeit der Grenzfunktion f unter der Voraussetzung der Stetigkeit der f_n folgt aus dem vorigen Satz. \square

Wir wenden diese Sätze nun auf komplexe Potenzreihen an. Sei (a_n) eine Folge komplexer Zahlen, $z_0 \in \mathbb{C}$ und

$$P(z) := \sum_{n=0}^{\infty} a_n (z - z_0)^n, \quad z \in \mathbb{C},$$

die Potenzreihe mit den Koeffizienten a_n und dem Zentrum z_0 . R bezeichne den Konvergenzradius der Potenzreihe P . In Kapitel 3.2. hatten wir Formeln zur Berechnung von R hergeleitet (siehe Sätze 3.17 und 3.18). Wir wissen bereits:

1. $P(z)$ konvergiert absolut für alle $z \in K(z_0, R)$.
2. $P(z)$ divergiert für $z \in \mathbb{C} \setminus cl(K(z_0, R))$.

Wir beweisen nun, dass die durch die Potenzreihe P auf dem Konvergenzkreis definierte Funktion stetig ist.

Satz 4.29 Sei $P(z) := \sum_{n=0}^{\infty} a_n(z - z_0)^n$ eine Potenzreihe mit dem Konvergenzradius R .

Dann gilt:

1. Die Potenzreihe $P(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ konvergiert gleichmäßig auf jeder kompakten Teilmenge $K \subset K(z_0, R)$.
2. Die Abbildung $P : z \in K(z_0, R) \mapsto P(z) \in \mathbb{C}$ ist stetig.

Beweis. (1) Sei $K \subset K(z_0, R)$ kompakt. Da die Funktion $z \in \mathbb{C} \mapsto |z - z_0| \in \mathbb{R}$ stetig ist, existiert

$$\eta(K) := \max\{|z - z_0| \mid z \in K\} < R.$$

Wir betrachten ein $z_1 \in K(z_0, R)$ mit $\eta(K) < |z_1 - z_0| < R$. Da die Reihe $P(z_1) = \sum_{n=0}^{\infty} a_n(z_1 - z_0)^n$ konvergiert, sind ihre Reihenglieder eine Nullfolge, also insbesondere beschränkt. Somit existiert eine Zahl $M > 0$, so dass $|a_n(z_1 - z_0)^n| < M$ für alle $n \in \mathbb{N}_0$.

Wir erhalten daraus

$$|a_n(z - z_0)^n| = \left| a_n(z_1 - z_0)^n \cdot \frac{(z - z_0)^n}{(z_1 - z_0)^n} \right| \leq M \cdot \underbrace{\left(\frac{\eta(K)}{|z_1 - z_0|} \right)^n}_{=: M_n} \quad \forall z \in K.$$

Nach Wahl von z_1 ist $\frac{\eta(K)}{|z_1 - z_0|} < 1$. Folglich ist die geometrische Reihe $\sum_{n=0}^{\infty} M_n$ konvergent.

Nach Satz 4.28 konvergiert dann $P(z)$ gleichmäßig auf K .

(2) Sei $z \in K(z_0, R)$ gegeben. Wir wählen ein $\varepsilon(z) > 0$, so dass

$$cl(K(z, \varepsilon(z))) \subset K(z_0, R).$$

Da die Menge $cl(K(z, \varepsilon(z)))$ sowohl abgeschlossen als auch beschränkt ist, ist sie kompakt. Folglich konvergiert die Funktionenreihe $P(z)$ auf $cl(K(z, \varepsilon(z)))$ gleichmäßig und die Grenzfunktion $P|_{cl(K(z, \varepsilon(z)))} : cl(K(z, \varepsilon(z))) \rightarrow \mathbb{C}$ ist stetig. Da dies für alle $z \in K(z_0, R)$ gilt, ist P auf dem gesamten Konvergenzkreis $K(z_0, R)$ stetig. \square

Als Anwendung der Stetigkeit der Grenzfunktion von Potenzreihen beweisen wir abschließend den folgenden Identitätssatz.

Satz 4.30 (Identitätssatz für Potenzreihen)

Seien $P(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ und $Q(z) = \sum_{n=0}^{\infty} b_n(z - z_0)^n$ zwei Potenzreihen mit dem Zentrum $z_0 \in \mathbb{C}$ und den Konvergenzradien $R_P > 0$ bzw. $R_Q > 0$ und $R := \min\{R_P, R_Q\}$. Existiert eine Folge komplexer Zahlen (w_m) in $K(z_0, R) \setminus \{z_0\}$ mit

1. $\lim_{m \rightarrow \infty} w_m = z_0$ und
2. $P(w_m) = Q(w_m)$ für alle $m \in \mathbb{N}$,

so gilt $a_n = b_n$ für alle $n \in \mathbb{N}_0$, das heißt die Potenzreihen P und Q sind identisch. Insbesondere sind P und Q identisch, wenn sie auf einer (beliebig kleinen) offenen Umgebung des Zentrums z_0 übereinstimmen.

Beweis. Sei (w_m) eine Folge in $K(z_0, R) \setminus \{z_0\}$, die gegen z_0 konvergiert und für die $P(w_m) = Q(w_m)$ gilt. Wir beweisen $a_n = b_n$ durch Induktion über den Index n .

Ind.-Anfang: Es ist zu zeigen, dass $a_0 = b_0$ gilt. Da (w_m) gegen z_0 konvergiert und P und Q auf $K(z_0, R)$ stetig sind, folgt

$$a_0 = P(z_0) = \lim_{m \rightarrow \infty} P(w_m) = \lim_{m \rightarrow \infty} Q(w_m) = Q(z_0) = b_0.$$

Ind.-Schritt: Es gelte $a_0 = b_0$, $a_1 = b_1, \dots, a_n = b_n$. Es ist zu zeigen, dass $a_{n+1} = b_{n+1}$ gilt. Dazu betrachten wir die Potenzreihen

$$P_1(z) := \sum_{k=n+1}^{\infty} a_k (z - z_0)^{k-(n+1)},$$

$$Q_1(z) := \sum_{k=n+1}^{\infty} b_k (z - z_0)^{k-(n+1)}.$$

Die Potenzreihen P_1 und Q_1 konvergieren auf $K(z_0, R)$, denn für $z \neq z_0$ gilt

$$P_1(z) = \frac{P(z) - \sum_{k=0}^n a_k (z - z_0)^k}{(z - z_0)^{n+1}},$$

$$Q_1(z) = \frac{Q(z) - \sum_{k=0}^n b_k (z - z_0)^k}{(z - z_0)^{n+1}}.$$

Nach Induktionsvoraussetzung gilt $Q_1(w_m) = P_1(w_m)$ für alle $m \in \mathbb{N}$. Da Q_1 und P_1 auf $K(z_0, R)$ stetig sind, folgt

$$a_{n+1} = P_1(z_0) = \lim_{m \rightarrow \infty} P_1(w_m) = \lim_{m \rightarrow \infty} Q_1(w_m) = Q_1(z_0) = b_{n+1}.$$

□

4.6 Die trigonometrischen und die Hyperbelfunktionen im Komplexen

In diesem Abschnitt werden wir die Ihnen vielleicht bereits aus der Schule bekannten trigonometrischen und Hyperbelfunktionen wiederholen und sie auf die komplexen Zahlen fortsetzen. Als Ausgangspunkt benutzen wir die komplexe Potenzreihe, die die Exponentialfunktion definiert:

$$\exp : \mathbb{C} \longrightarrow \mathbb{C},$$

$$z \longmapsto e^z := \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

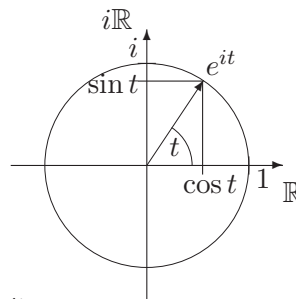
4.6.1 Die trigonometrischen Funktionen

Mit Hilfe der Exponentialfunktion definieren wir jetzt die trigonometrischen Funktionen. Wir betrachten zunächst die Exponentialfunktion auf der imaginären Achse.

Für e^{it} mit $t \in \mathbb{R}$ gilt

$$|e^{it}|^2 = e^{it} \cdot \overline{e^{it}} = e^{it} \cdot e^{-it} = e^0 = 1.$$

Die komplexe Zahl e^{it} liegt also auf der Kreislinie $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$. Den Realteil von e^{it} nennen wir *Cosinus* der reellen Zahl t , den Imaginärteil von e^{it} nennen wir *Sinus* von t :



$$\cos(t) := \operatorname{Re}(e^{it}) = \frac{e^{it} + e^{-it}}{2}, \quad \sin(t) := \operatorname{Im}(e^{it}) = \frac{e^{it} - e^{-it}}{2i}.$$

Dann erhält man unmittelbar

$$e^{it} = \cos(t) + i \sin(t) \quad \text{und} \quad \cos^2(t) + \sin^2(t) = 1 \quad \forall t \in \mathbb{R}. \quad (*)$$

Insbesondere ist $|\cos(t)| \leq 1$ und $|\sin(t)| \leq 1$ für jede reelle Zahl t .

Wir erweitern die Definition von Sinus und Cosinus nun auf die komplexen Zahlen und definieren $\cos : \mathbb{C} \rightarrow \mathbb{C}$ und $\sin : \mathbb{C} \rightarrow \mathbb{C}$ durch

$$\cos(z) := \frac{e^{iz} + e^{-iz}}{2}, \quad \sin(z) := \frac{e^{iz} - e^{-iz}}{2i} \quad \forall z \in \mathbb{C}.$$

Da die Exponentialfunktion stetig ist, sind die Sinus- und Cosinus-Funktion ebenfalls stetig. Außerdem gilt:

$$\begin{aligned} \sin(z) &= z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \dots = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}, \\ \cos(z) &= 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \dots = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}. \end{aligned}$$

Diese Potenzreihen haben Konvergenzradius $+\infty$, sind also für jedes $z \in \mathbb{C}$ absolut konvergent.

Durch Einsetzen von e^z in die Definitionen und Anwenden von $e^{z+w} = e^z \cdot e^w$ erhält man die folgenden Rechenregeln für Sinus und Cosinus:

1. $\boxed{\sin^2(z) + \cos^2(z) = 1 \quad \forall z \in \mathbb{C}}$:

$$\sin^2(z) + \cos^2(z) = -\frac{1}{4}(e^{2iz} + e^{-2iz} - 2) + \frac{1}{4}(e^{2iz} + e^{-2iz} + 2) = 1.$$

2. Die Eulersche Formel: $\boxed{e^{iz} = \cos(z) + i \cdot \sin(z) \quad \forall z \in \mathbb{C}}$:

$$\cos(z) + i \cdot \sin(z) = \frac{1}{2}(e^{iz} + e^{-iz}) + \frac{1}{2}(e^{iz} - e^{-iz}) = e^{iz}.$$

3. *Additionstheoreme:* Für alle $z_1, z_2 \in \mathbb{C}$ gilt:

$$\begin{aligned}\sin(z_1 + z_2) &= \sin(z_1) \cos(z_2) + \sin(z_2) \cos(z_1) \\ \cos(z_1 + z_2) &= \cos(z_1) \cos(z_2) - \sin(z_1) \sin(z_2).\end{aligned}$$

4. *Differenzformeln:* Für alle $z, w \in \mathbb{C}$ gilt:

$$\begin{aligned}\sin(z) - \sin(w) &= 2 \cdot \cos\left(\frac{z+w}{2}\right) \cdot \sin\left(\frac{z-w}{2}\right), \\ \cos(z) - \cos(w) &= -2 \cdot \sin\left(\frac{z+w}{2}\right) \cdot \sin\left(\frac{z-w}{2}\right).\end{aligned}$$

(Additionstheoreme auf $\frac{1}{2}(z+w) \pm \frac{1}{2}(z-w)$ anwenden.)

5. Die Sinusfunktion ist *ungerade*, die Cosinusfunktion ist *gerade*, d.h. es gilt:

$$\sin(-z) = -\sin(z) \quad \text{und} \quad \cos(-z) = \cos(z) \quad \forall z \in \mathbb{C}.$$

Als nächstes diskutieren wir nochmal den Kurvenverlauf der reellen Funktionen $\cos|_{\mathbb{R}}$ und $\sin|_{\mathbb{R}}$. Zunächst sehen wir uns die Nullstellen an:

Satz 4.31 Die Funktion $\cos|_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}$ hat auf dem Intervall $(0, 2)$ genau eine Nullstelle.

Beweis. 1. Existenz der Nullstelle: Aus der Definition von \cos folgt $\cos(0) = 1$. Andererseits ist

$$\begin{aligned}\cos(2) &= \sum_{n=0}^{\infty} (-1)^n \frac{2^{2n}}{(2n)!} \\ &= 1 - 2 + \frac{2}{3} + \sum_{n=3}^{\infty} (-1)^n \frac{2^{2n}}{(2n)!} \\ &= -\frac{1}{3} - \sum_{r=2}^{\infty} \left(\frac{2^{4r-2}}{(4r-2)!} - \frac{2^{4r}}{(4r)!} \right) \\ &= -\frac{1}{3} - \sum_{r=2}^{\infty} \frac{2^{4r-2}}{(4r-2)!} \cdot \underbrace{\left(1 - \frac{4}{4r(4r-1)} \right)}_{>0} \\ &< -\frac{1}{3}.\end{aligned}$$

(Die Umordnung der Reihe in der 3. Zeile konnten wir wegen ihrer absoluten Konvergenz vornehmen). Da $\cos(0) = 1 > 0$, $\cos(2) < -\frac{1}{3} < 0$ und $\cos|_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}$ 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 $\cos|_{\mathbb{R}}$ auf $(0, 2)$ streng monoton fallend ist. Die Funktion $\sin|_{\mathbb{R}}$ ist auf $(0, 2)$ positiv, da

$$\begin{aligned}
\sin(x) &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \\
&= \sum_{r=0}^{\infty} \left(\frac{x^{4r+1}}{(4r+1)!} - \frac{x^{4r+3}}{(4r+3)!} \right) \\
&= \sum_{r=0}^{\infty} \frac{x^{4r+1}}{(4r+1)!} \cdot \underbrace{\left(1 - \frac{x^2}{(4r+2)(4r+3)} \right)}_{>0} \\
&> 0 \quad \forall x \in (0, 2).
\end{aligned}$$

Seien nun $x, y \in (0, 2)$ mit $x < y$. Dann gilt $\frac{y \pm x}{2} \in (0, 2)$. Aus den Differenzformeln folgt

$$\cos(x) - \cos(y) = 2 \sin\left(\frac{y-x}{2}\right) \cdot \sin\left(\frac{y+x}{2}\right) > 0.$$

Folglich ist $\cos|_{\mathbb{R}}$ auf $(0, 2)$ streng monoton fallend und somit ξ die einzige Nullstelle in $(0, 2)$. \square

Definition 4.11. Die Nullstelle von $\cos|_{\mathbb{R}}$ auf dem Intervall $(0, 2)$ heißt $\frac{\pi}{2}$.

Wir werden später weitere Eigenschaften der Zahl π (sie ist irrational und transzendent) und ihre genauere Berechnung kennenlernen.

Nach Satz 4.31 gilt $\cos(\frac{\pi}{2}) = 0$ und $\sin(\frac{\pi}{2}) = 1$, woraus

$$e^{i\frac{\pi}{2}} = i, \quad e^{i\pi} = (e^{i\frac{\pi}{2}})^2 = -1 \quad \text{und} \quad e^{2\pi i} = (e^{i\pi})^2 = 1$$

folgt. Aus den Produkt-Eigenschaften der Exponentialfunktion erhalten wir damit

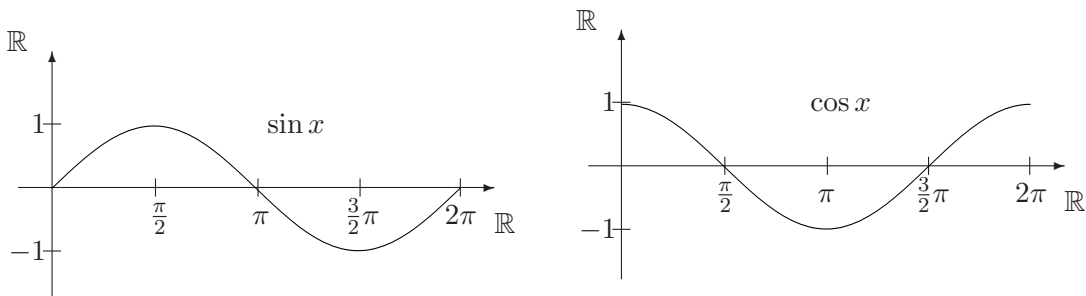
$$e^{i(x+\frac{\pi}{2})} = e^{ix} \cdot e^{i\frac{\pi}{2}} = i e^{ix}, \quad e^{i(x+\pi)} = e^{ix} \cdot e^{i\pi} = -e^{ix} \quad \text{für alle } x \in \mathbb{R}.$$

Der Vergleich von Real- und Imaginärteil in der Eulerschen Formel liefert dann die folgenden Beziehungen zwischen Sinus und Cosinus:

Folgerung 4.1 Für alle $x \in \mathbb{R}$ gilt

$$\sin\left(x + \frac{\pi}{2}\right) = +\cos(x), \quad \sin(x + \pi) = -\sin(x), \quad \cos(x + \pi) = -\cos(x).$$

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 $[0, 2\pi]$:



- Satz 4.32** 1. Die Exponentialfunktion $\exp : \mathbb{C} \rightarrow \mathbb{C}$ ist periodisch mit der Periode $2\pi i$, das heißt es gilt $e^z = e^{z+2\pi i}$ für alle $z \in \mathbb{C}$.
2. Die Funktionen $\sin : \mathbb{C} \rightarrow \mathbb{C}$ und $\cos : \mathbb{C} \rightarrow \mathbb{C}$ sind periodisch mit der Periode 2π , d.h. für alle $z \in \mathbb{C}$ gilt $\sin(z + 2\pi) = \sin(z)$ und $\cos(z + 2\pi) = \cos(z)$.
3. $e^z = 1 \iff z = 2\pi i k$ für ein $k \in \mathbb{Z}$.
4. Die Funktion $\sin : \mathbb{C} \rightarrow \mathbb{C}$ hat nur die (reellen) Nullstellen $x_k = \pi k$, $k \in \mathbb{Z}$. Die Funktion $\cos : \mathbb{C} \rightarrow \mathbb{C}$ hat nur die (reellen) Nullstellen $y_k = \frac{\pi}{2} + k\pi$, $k \in \mathbb{Z}$.
5. Zu jedem $z \in S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$ existiert genau eine reelle Zahl $t \in [0, 2\pi)$ mit $z = e^{it} = \cos(t) + i \sin(t)$.

Beweis.

- (1) Da $e^{2\pi i} = 1$, folgt $e^{z+2\pi i} = e^z \cdot e^{2\pi i} = e^z$.
- (2) Aus der Definition von Sinus erhält man:

$$\sin(z + 2\pi) = \frac{e^{i(2\pi+z)} - e^{-i(z+2\pi)}}{2i} = \frac{e^{iz} \cdot e^{2\pi i} - e^{-iz} \cdot e^{-2\pi i}}{2i} = \frac{e^{iz} - e^{-iz}}{2i} = \sin(z).$$

Analog zeigt man $\cos(z + 2\pi) = \cos(z)$. (Der Kurvenverlauf von $\cos|_{\mathbb{R}}$ und $\sin|_{\mathbb{R}}$ auf $[0, 2\pi]$ zeigt, dass 2π die kleinst mögliche Periode ist).

- (3) (\Leftarrow) Sei $z = 2\pi k i$ für ein $k \in \mathbb{Z}$. Dann gilt

$$e^z = e^{2\pi i k} = (e^{2\pi i})^k = 1^k = 1.$$

- (\Rightarrow) Sei $e^z = 1$ und $z = x + iy$ mit $x, y \in \mathbb{R}$. Dann gilt

$$\begin{aligned} 1 &= e^{x+iy} = e^x \cdot e^{iy} = e^x (\cos(y) + i \sin(y)) \quad \text{und somit} \\ 1 &= e^x |\cos(y) + i \sin(y)| = e^x \cdot \underbrace{\sqrt{\cos^2(y) + \sin^2(y)}}_{=1}. \end{aligned}$$

Da die Exponentialfunktion $\exp|_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}^+$ bijektiv ist, folgt $x = 0$. Somit erhalten wir $z = iy$ und aus $1 = e^{iy} = \cos(y) + i \sin(y)$ folgt $\cos(y) = 1$ und $\sin(y) = 0$. Damit ist $y = 2\pi k$, also $z = 2\pi k i$.

- (4) Aus der Definition von Sinus und Cosinus folgt

$$\begin{aligned} 0 = \sin(z) &= \frac{e^{iz} - e^{-iz}}{2i} \iff e^{iz} = e^{-iz} \iff e^{2iz} = 1 \iff z \in \mathbb{Z}\pi. \\ 0 = \cos(z) &= \frac{e^{iz} + e^{-iz}}{2} \iff e^{iz} = -e^{-iz} \iff e^{2iz} = -1 = e^{i\pi} \iff z \in \frac{\pi}{2} + \pi\mathbb{Z}. \end{aligned}$$

- (5) Sei $z \in S^1 = \{z \in \mathbb{C} \mid |z| = 1\}$. Wir wollen zeigen, dass es genau ein $t \in [0, 2\pi)$ gibt, mit $z = e^{it} = \cos(t) + i \sin(t)$. Sei $z = x + iy$, dann gilt $x^2 + y^2 = 1$ sowie $|x| \leq 1$. Wir beweisen nun durch Fallunterscheidung: (a) Sei $x = 1$. Dann ist $y = 0$ und $t = 0$ eindeutige Lösung von $x = \cos t$, $y = \sin t$ in $[0, 2\pi)$.

- (b) Sei $x = -1$. Dann ist $y = 0$ und $t = \pi$ eindeutige Lösung von $x = \cos t$, $y = \sin t$ in $[0, 2\pi)$.

(c) Bleibt also noch der Fall $x \in (-1, 1)$ zu betrachten. Sei $x \in (-1, 1)$. Dann existieren genau zwei Werte α, β mit

$$\alpha \in (0, 2\pi), \beta \in (0, 2\pi) \quad \text{und} \quad x = \cos \alpha = \cos \beta.$$

Dabei gilt $\sin \alpha > 0$ und $\sin \beta < 0$ und es folgt

$$|y| = \sqrt{1 - x^2} = \sqrt{1 - \cos^2 \alpha} = |\sin \alpha| = |\sin \beta|.$$

Ist $y > 0$, so ist $y = \sin \alpha$ und $t = \alpha$ die Lösung. Ist $y < 0$, so ist $y = \sin \beta$ und $t = \beta$ die Lösung. \square

Mittels der Sinus- und Cosinusfunktion definieren wir die Funktionen Tangens (tan) und Cotangens (cot):

$$\begin{aligned} \tan : \mathbb{C} \setminus \left(\frac{\pi}{2} + \mathbb{Z}\pi \right) &\longrightarrow \mathbb{C} \\ z &\longmapsto \tan(z) := \frac{\sin(z)}{\cos(z)} \\ \\ \cot : \mathbb{C} \setminus \mathbb{Z}\pi &\longrightarrow \mathbb{C} \\ z &\longmapsto \cot(z) := \frac{\cos(z)}{\sin(z)} \end{aligned}$$

Beide Funktionen sind offensichtlich stetig. Aus den Eigenschaften von Sinus und Cosinus erhält man die folgenden Eigenschaften von Tangens- und Cotangensfunktion:

1. tan und cot sind π -periodische, ungerade Funktionen.
2. Es gelten folgende *Additionstheoreme*:

$$\tan(z + w) = \frac{\tan(z) + \tan(w)}{1 - \tan(z) \cdot \tan(w)} \quad \text{und} \quad \cot(z + w) = \frac{\cot(z) \cdot \cot(w) - 1}{\cot(z) + \cot(w)}.$$

Die Umkehrfunktionen der reellen trigonometrischen Funktionen

Da die trigonometrischen Funktionen auf \mathbb{R} betrachtet periodisch sind, existieren keine globalen Umkehrfunktionen. Es existieren aber Umkehrfunktionen auf jedem Monotoniebereich der trigonometrischen Funktionen. Nach Satz 4.9 sind diese Umkehrfunktionen ebenfalls stetig.

Definition 4.12. Die Umkehrfunktion von $\sin|_{[-\frac{\pi}{2}, \frac{\pi}{2}]}$ heißt *Arcussinus* und wird mit *arcsin* bezeichnet.

$\arcsin : [-1, 1] \longrightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$ 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

$$\left(\sin|_{[-\frac{\pi}{2} + k\pi, \frac{\pi}{2} + k\pi]} \right)^{-1}(x) = (-1)^k \cdot \arcsin(x) + k\pi.$$

Man nennt sie auch die Zweige der Umkehrfunktion von Sinus.

Definition 4.13. 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 4.14. 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 4.15. 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

$$(\cot|_{[k\pi, (k+1)\pi]})^{-1}(x) = \operatorname{arccot}(x) + k\pi.$$

Den Kurvenverlauf der reellen trigonometrischen Funktionen und ihrer Umkehrfunktionen findet man in Tafelwerken.

4.6.2 Die Hyperbelfunktionen

Unter den Hyperbelfunktionen *Sinus hyperbolicus* $\sinh : \mathbb{C} \rightarrow \mathbb{C}$ und *Cosinus hyperbolicus* $\cosh : \mathbb{C} \rightarrow \mathbb{C}$ versteht man die folgenden komplexen Funktionen:

$$\begin{aligned} \sinh(z) &:= \frac{e^z - e^{-z}}{2}, \\ \cosh(z) &:= \frac{e^z + e^{-z}}{2}. \end{aligned}$$

Aus der Reihendarstellung für e^z folgt

$$\begin{aligned} \sinh(z) &= \sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!} = z + \frac{z^3}{3!} + \frac{z^5}{5!} + \dots \\ \cosh(z) &= \sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!} = 1 + \frac{z^2}{2!} + \frac{z^4}{4!} + \dots \end{aligned}$$

Wir bilden auch hier die Quotienten beider Funktionen und erhalten *Tangens hyperbolicus* $\tanh : \mathbb{C} \setminus \left(\frac{\pi}{2}i + \mathbb{Z}\pi i\right) \rightarrow \mathbb{C}$:

$$\tanh(z) := \frac{\sinh(z)}{\cosh(z)} = \frac{e^{2z} - 1}{e^{2z} + 1}$$

und *Cotangens hyperbolicus* $\coth : \mathbb{C} \setminus \mathbb{Z}\pi i \rightarrow \mathbb{C}$:

$$\coth(z) := \frac{\cosh(z)}{\sinh(z)} = \frac{e^{2z} + 1}{e^{2z} - 1}.$$

Alle diese Hyperbelfunktionen sind offensichtlich stetig. Durch Einsetzen der Exponentialabbildung in die Definitionen und die Rechenregeln für die trigonometrischen Funktionen erhält man folgende Rechenregeln für die Hyperbelfunktionen:

1. $\sinh(z) = -i \sin(iz)$ und $\cosh(z) = \cos(iz)$ für alle $z \in \mathbb{C}$.
2. $\cosh^2(z) - \sinh^2(z) = 1$ für alle $z \in \mathbb{C}$:

$$\cosh^2(z) - \sinh^2(z) = \cos^2(iz) - (-i \cdot \sin(iz))^2 = \cos^2(iz) + \sin^2(iz) = 1.$$

3. $\cosh(z) = \cosh(-z)$, $\sinh(z) = -\sinh(-z)$ für alle $z \in \mathbb{C}$
4. *Additionstheoreme*:

$$\begin{aligned} \cosh(z+w) &= \cosh(z)\cosh(w) + \sinh(z)\sinh(w), \\ \sinh(z+w) &= \sinh(z)\cosh(w) + \cosh(z)\sinh(w), \\ \tanh(z+w) &= \frac{\tanh(z) + \tanh(w)}{1 + \tanh(z) \cdot \tanh(w)}. \end{aligned}$$

Wir betrachten abschließend die Hyperbelfunktionen im Reellen: Die Kurvenverläufe dieser Funktionen und ihrer Umkehrfunktionen findet man in Tafelwerken.

So wie die reellen trigonometrischen Funktionen \sin und \cos die Punkte der Kreislinie $S^1 := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ beschreiben:

$$S^1 = \{(\cos(t), \sin(t)) \mid t \in \mathbb{R}\},$$

so beschreiben die reellen hyperbolischen Funktionen die Punkte der Hyperbel $Hyp := \{(x, y) \in \mathbb{R}^2 \mid x^2 - y^2 = 1, x > 0\}$ (dies motiviert den Namen dieser Funktionenklasse):

$$Hyp = \{(\cosh(t), \sinh(t)) \mid t \in \mathbb{R}\}.$$

Hängt man ein ideales Seil zwischen zwei Punkten über der Erdoberfläche auf, so wird die entstehende *Seilkurve* durch die Funktion $b \cosh(ax)$ beschrieben. Dies werden wir später im Abschnitt über Differentialgleichungen beweisen.

Die Funktion $\sinh|_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}$ ist bijektiv, streng monoton wachsend und hat die Nullstelle $x = 0$. Die Funktion $\cosh|_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}$ nimmt ihre Werte in $[1, +\infty)$ an, sie ist streng monoton wachsend und bijektiv von $[0, +\infty)$ auf $[1, +\infty)$. Dabei ist $\cosh(0) = 1$. Wir können deshalb die Umkehrfunktionen betrachten:

Definition 4.16. Die Umkehrfunktion von $\sinh|_{\mathbb{R}} : \mathbb{R} \rightarrow \mathbb{R}$ 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 $\tanh|_{\mathbb{R}} : \mathbb{R} \rightarrow (-1, 1) \subset \mathbb{R}$ ist streng monoton wachsend und bijektiv. Die Funktion $\coth|_{\mathbb{R} \setminus \{0\}} : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R} \setminus [-1, 1]$ ist streng monoton fallend und bijektiv.

Definition 4.17. Die Umkehrfunktion von $\tanh|_{\mathbb{R}}$ heißt *Areatangens hyperbolicus* :

$$\operatorname{artanh} : (-1, 1) \rightarrow \mathbb{R}.$$

Die Umkehrfunktion von $\coth|_{\mathbb{R} \setminus \{0\}}$ heißt *Areacotangens hyperbolicus* :

$$\operatorname{arcoth} : \mathbb{R} \setminus [-1, 1] \rightarrow \mathbb{R} \setminus \{0\}.$$

Satz 4.33 Es bestehen die folgenden Beziehungen zwischen den Umkehrfunktionen der Hyperbelfunktionen und der Logarithmusfunktion:

$$\begin{aligned} \operatorname{arsinh}(x) &= \ln(x + \sqrt{x^2 + 1}), & x \in \mathbb{R}, \\ \operatorname{arcosh}(x) &= \ln(x + \sqrt{x^2 - 1}), & x \geq 1, \\ \operatorname{artanh}(x) &= \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right), & |x| < 1, \\ \operatorname{arcoth}(x) &= \frac{1}{2} \ln\left(\frac{x+1}{x-1}\right), & |x| > 1. \end{aligned}$$

Beweis. Übungsaufgabe. □

4.7 Der Fundamentalsatz der Algebra

Als Anwendung unserer bisherigen Erkenntnisse über stetige Funktionen beweisen wir den Fundamentalsatz der Algebra, der weitreichende Konsequenzen sowohl in der Algebra als auch in vielen Bereichen der Analysis hat.

Satz 4.34 (Fundamentalsatz der Algebra)

Jedes nicht-konstante komplexe Polynom hat eine Nullstelle.

Beweis. Sei $Q \in \mathbb{C}[z]$ ein komplexes Polynom vom Grad $n \geq 1$. Dann gibt es komplexe Zahlen a_0, a_1, \dots, a_n mit $a_n \neq 0$, so dass

$$Q(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0.$$

(1) Zuerst zeigen wir, dass für ein beliebig vorgegebenes $\alpha > 0$ ein $r > 0$ existiert, so dass $|Q(z)| \geq \alpha$ für alle $|z| \geq r$.

Wir definieren dazu $\tilde{Q}(z)$ durch

$$Q(z) = a_n z^n \left(1 + \underbrace{\frac{a_{n-1}}{a_n} \frac{1}{z} + \dots + \frac{a_0}{a_n} \frac{1}{z^n}}_{=: \tilde{Q}(z)} \right), \quad z \neq 0.$$

Sei nun $\beta := 1 + \left| \frac{a_{n-1}}{a_n} \right| + \dots + \left| \frac{a_0}{a_n} \right| \geq 1$. Für $|z| \geq 2\beta$ folgt dann

$$\begin{aligned} |\tilde{Q}(z)| &\leq \left| \frac{a_{n-1}}{a_n} \right| \frac{1}{|z|} + \dots + \left| \frac{a_0}{a_n} \right| \frac{1}{|z|^n} \leq \left(\left| \frac{a_{n-1}}{a_n} \right| + \dots + \left| \frac{a_0}{a_n} \right| \right) \frac{1}{|z|} \\ &\leq \beta \cdot \frac{1}{|z|} \leq \frac{1}{2}. \end{aligned}$$

Dies zeigt, dass $|1 + \tilde{Q}(z)| \geq \frac{1}{2}$ und

$$|Q(z)| = |a_n z^n| \cdot |1 + \tilde{Q}(z)| \geq \frac{1}{2} |a_n| \cdot |z|^n \quad \text{für alle } |z| \geq 2\beta.$$

Setzen wir $r := \max\{2\beta, \left(\frac{2\alpha}{|a_n|}\right)^{\frac{1}{n}}\}$, so erhalten wir

$$|Q(z)| \geq \alpha \quad \text{für alle } z \in \mathbb{C} \text{ mit } |z| \geq r.$$

(2) Wir zeigen nun, dass ein $\xi \in \mathbb{C}$ existiert, so dass $Q(\xi) = 0$.

Sei $A := \{|Q(z)| \mid z \in \mathbb{C}\} \subset \mathbb{R}$. Da $|Q(z)| \geq 0$ gilt, ist A von unten beschränkt. Somit existiert das Infimum $\mu := \inf A \geq 0$. Nach (1) gibt es ein $r > 0$, so dass

$$|Q(z)| \geq \mu + 1 \quad \forall |z| \geq r.$$

Also ist $\mu = \inf\{|Q(z)| \mid z \in \mathbb{C}\} = \inf\{|Q(z)| \mid |z| \leq r\}$. Da die abgeschlossene Kugel $\text{cl } K(0, r) \subset \mathbb{C}$ kompakt ist und die Funktion

$$\begin{aligned} q : \mathbb{C} &\longrightarrow \mathbb{R} \\ z &\mapsto |Q(z)| \end{aligned}$$

stetig ist, existiert ein Minimum von q auf $\text{cl } K(0, r)$, das heißt, es existiert ein $\xi \in \text{cl } K(0, r)$ mit $|Q(\xi)| = \mu$. Wir zeigen, dass $Q(\xi) = 0$ gilt. Wir nehmen an, dass $Q(\xi) \neq 0$ ist und betrachten

$$H(z) := \frac{Q(z + \xi)}{Q(\xi)}.$$

Da $|Q(\xi)| = \min\{|Q(z)| \mid z \in \mathbb{C}\}$, gilt $|H(z)| \geq 1$ für alle $z \in \mathbb{C}$. Da $H(0) = 1$, 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 $b_m \neq 0$. Für die komplexe Zahl $w := -\frac{|b_m|}{b_m}$ der Norm 1 existiert ein $\psi \in \mathbb{R}$, so dass $w = e^{im\psi}$ (siehe Satz 4.32). Dann ist $b_m e^{im\psi} = -|b_m|$. Wir betrachten nun H auf den komplexen Zahlen der Form $z = \varrho \cdot e^{i\psi}$ mit $\varrho > 0$. Dann gilt

$$|H(\varrho \cdot e^{i\psi})| \leq |b_n| \cdot \varrho^n + \dots + |b_{m+1}| \cdot \varrho^{m+1} + \underbrace{|b_m \varrho^m \cdot e^{im\psi} + 1|}_{=|1 - |b_m| \cdot \varrho^m|}.$$

Sei nun ϱ so klein gewählt, dass $\varrho^m < \frac{1}{|b_m|}$ ist. Dann ist $1 - |b_m| \varrho^m > 0$ und damit

$$\begin{aligned} |H(\varrho \cdot e^{i\psi})| &\leq 1 - |b_m| \cdot \varrho^m + |b_{m+1}| \cdot \varrho^{m+1} + \dots + |b_n| \cdot \varrho^n \\ &= 1 - \varrho^m \underbrace{(|b_m| - |b_{m+1}| \cdot \varrho - \dots - |b_n| \cdot \varrho^{n-m})}_{>0 \text{ für } \varrho \text{ hinreichend klein}}. \end{aligned}$$

Somit ist $|H(\varrho \cdot e^{i\psi})| < 1$ 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 $Q(\xi) = 0$. \square

Satz 4.35 (Zerlegungssatz für komplexe Polynome)

Sei $Q \in \mathbb{C}[z]$ ein Polynom vom Grad $n \geq 1$ und seien $\xi_1, \dots, \xi_m \in \mathbb{C}$ die verschiedenen Nullstellen von Q . Dann gilt

$$Q(z) = a_n z^n + \dots + a_1 z + a_0 = a_n (z - \xi_1)^{\nu_1} \cdot (z - \xi_2)^{\nu_2} \cdot \dots \cdot (z - \xi_m)^{\nu_m},$$

wobei $\nu_j \in \mathbb{N}$ eindeutig bestimmte natürliche Zahlen sind und $n = \nu_1 + \dots + \nu_m$.

Beweis. Sei $Q(\xi_1) = 0$. Wir betrachten das Polynom

$$q_k(z) := z^{k-1} + z^{k-2} \cdot \xi_1 + z^{k-3} \cdot \xi_1^2 + \dots + z \cdot \xi_1^{k-2} + \xi_1^{k-1}.$$

Dann gilt $z^k - \xi_1^k = (z - \xi_1) \cdot q_k(z)$ und folglich

$$\begin{aligned} Q(z) &= Q(z) - Q(\xi_1) = a_1(z - \xi_1) + a_2(z^2 - \xi_1^2) + \dots + a_n(z^n - \xi_1^n) \\ &= (z - \xi_1) \cdot (a_1 q_1(z) + a_2 q_2(z) + \dots + a_n q_n(z)) \\ &=: (z - \xi_1) \cdot Q_1(z). \end{aligned}$$

$Q_1(z)$ ist dabei ein Polynom vom Grad $n - 1$. Ist $n - 1 \geq 1$, so hat Q_1 eine Nullstelle und wir können einen weiteren Linearfaktor von Q_1 abspalten. Dieses Verfahren funktioniert n -mal. \square

Bemerkung: ν_j heißt die *algebraische Vielfachheit* der Nullstelle ξ_j . Ein nicht-konstantes komplexes Polynom Q vom Grad n hat als genau n komplexe Nullstellen (gezählt mit Vielfachheit).

4.8 Approximationssätze für stetige Abbildungen

In diesem Abschnitt werden wir untersuchen, wie man stetige Abbildungen durch gewisse "einfache" stetige Abbildungen approximieren kann. Dazu ist zunächst zu klären, was wir unter "approximieren" verstehen wollen. Wir benötigen dazu eine Metrik auf dem Raum der stetigen Abbildungen, mit der wir Konvergenz beschreiben können.

Wir betrachten zwei metrische Räume (X, d_X) und (Y, d_Y) und bezeichnen mit $C(X, Y)$ die Menge aller stetigen Abbildungen von X nach Y :

$$C(X, Y) := \{f : X \longrightarrow Y \mid f \text{ stetig} \}.$$

Um auf $C(X, Y)$ eine geeignete Metrik zu definieren, setzen wir nun voraus, dass der metrische Raum X *kompakt* ist. Wir definieren für $f, h \in C(X, Y)$

$$d_\infty(f, h) := \max\{d_Y(f(x), h(x)) \mid x \in X\}.$$

Da die Metrik d_Y stetig ist, ist die Abbildung $x \in X \mapsto d_Y(f(x), h(x)) \in \mathbb{R}$ stetig, sie besitzt also tatsächlich ein Maximum auf dem kompakten metrischen Raum X . Aus den

Eigenschaften der Metrik d_Y erhält man ohne Probleme, dass d_∞ eine Metrik auf dem Raum der stetigen Abbildungen $C(X, Y)$ ist.

Ist der Bildraum ein normierter \mathbb{K} -Vektorraum $(V, \|\cdot\|)$, so trägt auch die Menge der stetigen Abbildungen $C(X, V)$ eine \mathbb{K} -Vektorraumstruktur (\mathbb{K} ist hier der Körper der reellen oder der komplexen Zahlen). Zur Erinnerung: Für $f, h \in C(X, V)$ und $\lambda \in \mathbb{K}$ definieren wir

$$(f + h)(x) := f(x) + h(x) \quad \text{und} \quad (\lambda \cdot f)(x) := \lambda \cdot f(x) \quad \forall x \in X.$$

Die Norm $\|\cdot\|$ auf V liefert dann eine Norm $\|\cdot\|_\infty$ auf dem Vektorraum $C(X, V)$:

$$\|f\|_\infty := \max\{\|f(x)\| \mid x \in X\}, \quad f \in C(X, V).$$

Die durch $\|\cdot\|_\infty$ gegebene Metrik auf $C(X, V)$ stimmt mit der Metrik d_∞ überein, die auf $C(X, V)$ durch die durch $\|\cdot\|$ gegebene Metrik $d_{\|\cdot\|}$ von V induziert wird:

$$\begin{aligned} d_\infty(f, h) &= \max\{d_{\|\cdot\|}(f(x), h(x)) \mid x \in X\} \\ &= \max\{\|f(x) - h(x)\| \mid x \in X\} \\ &= \|f - h\|_\infty. \end{aligned}$$

Im folgenden sind die Mengen $C(X, Y)$ bzw. $C(X, V)$ immer mit der Metrik d_∞ bzw. der Norm $\|\cdot\|_\infty$ versehen.

Satz 4.36 *Sei X ein kompakter metrischer Raum und Y ein beliebiger metrischer Raum.*

1. *Sei (f_n) eine Folge stetiger Abbildungen $f_n \in C(X, Y)$. Die Folge (f_n) konvergiert im metrischen Raum $C(X, Y)$ genau dann gegen $f \in C(X, Y)$, wenn die Funktionenfolge (f_n) gleichmäßig gegen f konvergiert.*
2. *Ist Y vollständig, so ist $C(X, Y)$ vollständig.*
3. *Ist V ein Banachraum, so ist $C(X, V)$ ebenfalls ein Banachraum.*

Beweis. (1) Die erste Behauptung folgt aus der Äquivalenz der folgenden Bedingungen:

$$\begin{aligned} &f_n \longrightarrow f \quad \text{in } C(X, Y) \\ \iff &\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \text{ so dass } d_\infty(f_n, f) < \varepsilon \quad \forall n \geq n_0 \\ \iff &\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \text{ so dass } \max\{d_Y(f_n(x), f(x)) \mid x \in X\} < \varepsilon \quad \forall n \geq n_0 \\ \iff &\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \text{ so dass } d_Y(f_n(x), f(x)) < \varepsilon \quad \forall n \geq n_0 \quad \forall x \in X. \\ \iff &(f_n) \text{ konvergiert gleichmäßig gegen } f. \end{aligned}$$

(2) Sei der metrische Raum (Y, d_Y) vollständig. Wir zeigen, dass dann auch $(C(X, Y), d_\infty)$ vollständig ist. Sei (f_n) eine Cauchy-Folge in $C(X, Y)$ und $\varepsilon > 0$. Dann existiert ein $n_0 \in \mathbb{N}$, so dass für alle $x \in X$

$$d_Y(f_n(x), f_m(x)) \leq d_\infty(f_n, f_m) < \varepsilon \quad \forall n, m \geq n_0. \quad (*)$$

Folglich ist die Folge $(f_n(x))$ für jedes $x \in X$ eine Cauchy-Folge in Y . Da Y vollständig ist, existiert $f(x) := \lim_{n \rightarrow \infty} f_n(x)$ auf X . Da die Metrik d_Y stetig ist, folgt aus (*) für jedes $x \in X$

$$\lim_{m \rightarrow \infty} d_Y(f_n(x), f_m(x)) = d_Y(f_n(x), \lim_{m \rightarrow \infty} f_m(x)) = d_Y(f_n(x), f(x)) \leq \varepsilon \quad \forall n \geq n_0.$$

Dies zeigt, dass die Funktionenfolge (f_n) gleichmäßig gegen f konvergiert. Nach Satz 4.25 ist f stetig und nach (1) konvergiert die Folge (f_n) in $C(X, Y)$ gegen f .

(3) ist ein Spezialfall von (2). □

In verschiedenen Situationen kann man stetige Abbildungen durch "einfache" stetige Abbildungen approximieren. Ein Beispiel kennen wir bereits:

Sei $\sum_{k=0}^{\infty} a_k x^k$ eine komplexe Potenzreihe mit dem Konvergenzradius $R > 0$ und bezeichne $p_n(x) := \sum_{k=0}^n a_k x^k$ ihre n -te Partialsumme. Dann ist die Funktion

$$\begin{aligned} f : (-R, R) \subset \mathbb{R} &\longrightarrow \mathbb{C} \\ t &\longmapsto \sum_{k=0}^{\infty} a_k t^k \end{aligned}$$

stetig und die Folge der Polynome (p_n) konvergiert gleichmäßig gegen f auf jedem kompakten Intervall $[a, b] \subset (-R, R)$.

Wir wollen nun *beliebige* stetige Funktionen aus $C([a, b], \mathbb{K})$ durch Polynome approximieren. Wir bezeichnen mit $P([a, b], \mathbb{K})$ diejenigen stetigen Funktionen, die mittels Polynomen definiert sind:

$$P([a, b], \mathbb{K}) := \{ p : t \in [a, b] \mapsto p(t) \in \mathbb{K} \mid p \in \mathbb{K}[x] \} \subset C([a, b], \mathbb{K}).$$

Satz 4.37 (Weierstraßscher Approximationssatz)

Die Menge der Polynome $P([a, b], \mathbb{K})$ liegt dicht in $C([a, b], \mathbb{K})$, das heißt, für jede stetige Funktion $f : [a, b] \rightarrow \mathbb{K}$ existiert eine Folge von Polynomen $p_n : [a, b] \rightarrow \mathbb{K}$, $n \in \mathbb{N}$, die auf $[a, b]$ gleichmäßig gegen f konvergiert.

Beweis. (1) Es genügt, die Behauptung für das Intervall $[0, 1]$ zu zeigen:

Sei die Behauptung für das Intervall $[0, 1]$ bewiesen und $[a, b]$ ein beliebiges Intervall. Wir betrachten den Homöomorphismus

$$\begin{aligned} \phi : [0, 1] &\longrightarrow [a, b] \\ t &\longmapsto a + (b - a)t. \end{aligned}$$

Die inverse Abbildung ϕ^{-1} ist durch

$$\phi^{-1}(x) = \frac{x - a}{b - a}$$

gegeben. Sei nun $f \in C([a, b], \mathbb{K})$ und $h := f \circ \phi \in C([0, 1], \mathbb{K})$. Nach unserer Annahme existiert eine Polynomfolge (p_n) , die auf $[0, 1]$ gleichmäßig gegen h konvergiert. Dann ist $q_n := p_n \circ \phi^{-1}$ eine Polynomfolge, die auf $[a, b]$ gleichmäßig gegen $f = h \circ \phi^{-1}$ konvergiert.

(2) Sei $f \in C([0, 1], \mathbb{K})$. Unter dem n -ten *Bernstein-Polynom* zu f verstehen wir das folgende Polynom n -ten Grades:

$$B[f]_n(x) := \sum_{k=0}^n f\left(\frac{k}{n}\right) \underbrace{\binom{n}{k} (1-x)^{n-k} \cdot x^k}_{=:\beta_{k,n}(x)}.$$

Die Polynome $\beta_{k,n}(x)$ haben folgende Eigenschaften:

- (1) $\beta_{k,n}(x) \geq 0 \quad \forall x \in [0, 1]$,
- (2) $\sum_{k=0}^n \beta_{k,n}(x) = (1-x+x)^n = 1 \quad \forall x \in [0, 1]$,
- (3) $\sum_{k=0}^n (k-nx)^2 \cdot \beta_{k,n}(x) = nx(1-x) \quad \forall x \in [0, 1]$.

Die ersten beiden Eigenschaften folgen direkt aus der Definition, die 3. Eigenschaft ist eine Übungsaufgabe zum Rechnen mit Binomialkoeffizienten.

Im folgenden zeigen wir, dass die Bernstein-Polynome $B[f]_n$ auf $[0, 1]$ gleichmäßig gegen f konvergieren.

Da f stetig und $[0, 1]$ kompakt ist, ist $f([0, 1]) \subset \mathbb{K}$ ebenfalls kompakt, also insbesondere beschränkt. Es existiert also eine Konstante $C > 0$ so dass

$$|f(x)| \leq C \quad \text{für alle } x \in [0, 1]. \quad (*)$$

Da $[0, 1]$ kompakt ist, ist die Funktion f sogar gleichmäßig stetig. Folglich existiert zu jedem $\varepsilon > 0$ ein $\delta > 0$ so dass

$$|f(x_1) - f(x_2)| < \frac{\varepsilon}{2} \quad \forall x_1, x_2 \in [0, 1] \text{ mit } |x_1 - x_2| < \delta. \quad (**)$$

Wir fixieren nun ein $n_0 \in \mathbb{N}$ mit $n_0 > \frac{C}{\varepsilon \cdot \delta^2}$ und zeigen, dass

$$|B[f]_n(x) - f(x)| < \varepsilon \quad \forall n \geq n_0, \forall x \in [0, 1].$$

Für $x \in [0, 1]$ und $n \in \mathbb{N}$ betrachten wir die folgenden Mengen

$$I_n(x) := \left\{ k \in \{0, 1, \dots, n\} \mid \left| \frac{k}{n} - x \right| < \delta \right\}$$

$$J_n(x) := \left\{ k \in \{0, 1, \dots, n\} \mid \left| \frac{k}{n} - x \right| \geq \delta \right\}.$$

Dann erhalten wir

$$\begin{aligned} |B[f]_n(x) - f(x)| &\stackrel{(2)}{=} \left| \sum_{k=0}^n \left(f\left(\frac{k}{n}\right) - f(x) \right) \cdot \beta_{k,n}(x) \right| \\ &\leq \underbrace{\left| \sum_{k \in I_n(x)} \left(f\left(\frac{k}{n}\right) - f(x) \right) \beta_{k,n}(x) \right|}_{:=S_n(x)} + \underbrace{\left| \sum_{k \in J_n(x)} \left(f\left(\frac{k}{n}\right) - f(x) \right) \cdot \beta_{k,n}(x) \right|}_{:=R_n(x)}. \end{aligned}$$

Für den Summanden $S_n(x)$ ergibt sich

$$S_n(x) \stackrel{(1)}{\leq} \sum_{I_n(x)} \left| f\left(\frac{k}{n}\right) - f(x) \right| \cdot \beta_{k,n}(x) \stackrel{(**)}{\leq} \frac{\varepsilon}{2} \cdot \sum_{I_n(x)} \beta_{k,n}(x) \leq \frac{\varepsilon}{2} \sum_{k=0}^n \beta_{k,n}(x) \stackrel{(2)}{=} \frac{\varepsilon}{2}.$$

Ist $k \in J_n(x)$, so gilt $(\frac{k}{n} - x)^2 \geq \delta^2$, also $\frac{(k-nx)^2}{n^2\delta^2} \geq 1$. Für den Summanden $R_n(x)$ ergibt sich

$$\begin{aligned} R_n(x) &\leq \sum_{J_n(x)} \left\{ \left| f\left(\frac{k}{n}\right) \right| + |f(x)| \right\} \cdot \beta_{k,n}(x) \cdot \underbrace{\frac{(k-nx)^2}{n^2\delta^2}}_{\geq 1} \\ &\stackrel{(*),(3)}{\leq} \frac{2C}{n^2\delta^2} \cdot \underbrace{\sum_{k=0}^n (k-nx)^2 \beta_{k,n}(x)}_{=nx(1-x)} \\ &= \frac{2C}{n\delta^2} x(1-x). \end{aligned}$$

Sei nun $n \geq n_0 > \frac{C}{\varepsilon\delta^2}$. Dann ist $\frac{2C}{n\delta^2} < 2\varepsilon$. Für $x \in [0, 1]$ gilt außerdem

$$x(1-x) = x - x^2 = -\left(x - \frac{1}{2}\right)^2 + \frac{1}{4} < \frac{1}{4}.$$

Folglich ist $R_n(x) < \frac{\varepsilon}{2}$. Insgesamt erhalten wir

$$|B[f]_n(x) - f(x)| < \varepsilon \quad \text{für alle } n \geq n_0 \text{ und } x \in [0, 1].$$

Somit konvergiert die Folge der Bernstein-Polynome $(B[f]_n)$ auf $[0, 1]$ gleichmäßig gegen die stetige Funktion f . □

Wir wollen den Weierstraßschen Approximationssatz nun auf Funktionen mit beliebigem kompakten Definitionsbereich verallgemeinern.

Wir betrachten dazu einen beliebigen kompakten metrischen Raum X und den Banachraum $C(X, \mathbb{K})$ der stetigen reell- bzw. komplexwertigen Funktionen auf X mit der Norm $\|\cdot\|_\infty$ und fragen uns, wann eine Teilmenge $\mathcal{R} \subset C(X, \mathbb{K})$ dicht in $C(X, \mathbb{K})$ liegt.

Definition 4.18. *Unter einer Unteralgebra von $C(X, \mathbb{K})$ versteht man einen linearen Unterraum⁴ $\mathcal{R} \subset C(X, \mathbb{K})$ mit der zusätzlichen Eigenschaft, dass mit zwei Funktionen $f, g \in \mathcal{R}$ auch das Produkt $f \cdot g$, definiert durch $(f \cdot g)(x) = f(x) \cdot g(x)$, in \mathcal{R} liegt.*

Ein Beispiel für eine Unteralgebra ist die im Weierstraßschen Approximationssatz betrachtete Menge der Polynome $P([a, b], \mathbb{K}) \subset C([a, b], \mathbb{K})$.

Wir wollen nun Bedingungen dafür angeben, dass eine *Unteralgebra* $\mathcal{R} \subset C(X, \mathbb{K})$ dicht in $C(X, \mathbb{K})$ liegt. Als Hilfsmittel beweisen wir zunächst folgende Lemmata.

⁴ Ein linearer Unterraum ist eine Teilmenge von $C(X, \mathbb{K})$, die abgeschlossen gegenüber den Vektorraum-Operationen ist.

Lemma 4.19. *Sei X ein kompakter metrischer Raum. Ist $\mathcal{R} \subset C(X, \mathbb{K})$ eine Unteralgebra im Banachraum der stetigen \mathbb{K} -wertigen Funktionen auf X , so ist auch der Abschluss $cl(\mathcal{R}) \subset C(X, \mathbb{K})$ eine Unteralgebra.*

Beweis. Seien $f, g \in cl(\mathcal{R}) \subset C(X, \mathbb{K})$ stetige Funktionen, die im Abschluss der Unteralgebra \mathcal{R} liegen. Dann existieren Folgen stetiger Funktionen (f_n) und (g_n) mit $f_n, g_n \in \mathcal{R}$, die bzgl. $\|\cdot\|_\infty$ gegen f bzw. g konvergieren. Da \mathcal{R} eine Unteralgebra ist, liegen die stetigen Funktionen $f_n + g_n$, $\lambda \cdot f_n$ und $f_n \cdot g_n$ ebenfalls in \mathcal{R} , wobei $\lambda \in \mathbb{K}$. Mit den Normeigenschaften sieht man sofort, dass die Folgen $(f_n + g_n)$ und $(\lambda \cdot f_n)$ bzgl. $\|\cdot\|_\infty$ gegen $f + g$ bzw. $\lambda \cdot f$ konvergieren. Somit liegen $f + g$ und $\lambda \cdot f$ in $cl(\mathcal{R})$. Es bleibt zu zeigen, dass auch $f \cdot g$ in $cl(\mathcal{R})$ liegt.

Da die Folge (f_n) konvergiert, ist sie beschränkt, d.h. es existiert eine Konstante $C \in \mathbb{R}^+$ mit

$$|f_n(x)| \leq \|f_n\|_\infty \leq C \quad \forall n \in \mathbb{N} \quad \forall x \in X.$$

Nach Definition der Norm $\|\cdot\|_\infty$ gilt außerdem

$$|g(x)| \leq \|g\|_\infty \quad \forall x \in X.$$

Damit erhalten wir

$$\begin{aligned} \|f \cdot g - f_n \cdot g_n\|_\infty &= \|f \cdot g - f_n \cdot g + f_n \cdot g - f_n \cdot g_n\|_\infty \\ &= \|g(f - f_n) + f_n(g - g_n)\|_\infty \\ &\leq \|g(f - f_n)\|_\infty + \|f_n(g - g_n)\|_\infty \\ &= \max_{x \in X} \{|g(x)| |f(x) - f_n(x)|\} + \max_{x \in X} \{|f_n(x)| |g(x) - g_n(x)|\} \\ &\leq \|g\|_\infty \cdot \|f - f_n\|_\infty + C \cdot \|g - g_n\|_\infty \end{aligned}$$

Da (f_n) gegen f und (g_n) gegen g konvergiert, folgt aus dieser Ungleichung, dass $f_n \cdot g_n$ gegen $f \cdot g$ konvergiert. Somit ist $f \cdot g \in cl(\mathcal{R})$. \square

Lemma 4.20. *Sei X ein kompakter metrischer Raum und $\mathcal{R} \subset C(X, \mathbb{R})$ eine Unteralgebra, die alle konstanten Funktionen enthält. Dann liegen für zwei Funktionen $f, g \in \mathcal{R}$ die Funktionen $|f|$, $\max\{f, g\}$ und $\min\{f, g\}$ im Abschluss $cl(\mathcal{R})$.*

Beweis. (1) Sei $f \in \mathcal{R}$. Wir wissen, dass

$$|f(x)| \leq \|f\|_\infty =: C \in \mathbb{R}^+ \quad \forall x \in X.$$

Wir betrachten die stetige Funktion $w : x \in [0, C^2] \mapsto \sqrt{x} \in \mathbb{R}$. Sei (p_n) eine Folge von Polynomen, die auf $[0, C^2]$ gleichmäßig gegen w konvergiert. Eine solche Folge existiert nach Satz 4.37. Dann konvergiert die Funktionenfolge $(p_n \circ f^2)$ auf X gleichmäßig gegen $w \circ f^2 = |f|$. Folglich konvergiert die Folge $(p_n \circ f^2)$ im metrischen Raum $(C(X, \mathbb{R}), \|\cdot\|_\infty)$ gegen $|f|$. Da \mathcal{R} eine Algebra ist, ist mit $f \in \mathcal{R}$ auch jede Potenz $f^i \in \mathcal{R}$. Außerdem liegen

nach Voraussetzung alle konstanten Funktionen in \mathcal{R} . Seien $p_n(x) = \sum_{i=0}^n a_{ni}x^i$ die obigen Polynome. Dann folgt:

$$p_n \circ f^2 = \sum_{i=1}^n a_{in}(f^2)^i \in \mathcal{R}.$$

Somit ist $|f|$ der Grenzwert einer Folge von Elementen aus \mathcal{R} und folglich gilt $|f| \in cl(\mathcal{R})$. (2) Seien $f, g \in \mathcal{R}$. Dann sind $f + g$ und $f - g$ Elemente in \mathcal{R} , weil \mathcal{R} ein Unterraum von $C(X, \mathbb{R})$ ist. Aus (1) folgt $|f - g| \in cl(\mathcal{R})$ und $|f + g| \in cl(\mathcal{R})$. Da nach Lemma 4.19 der Abschluß $cl(\mathcal{R})$ ebenfalls eine Unteralgebra ist, erhalten wir

$$\left. \begin{aligned} \max\{f, g\} &= \frac{1}{2}(f + g + |f - g|) \\ \min\{f, g\} &= \frac{1}{2}(f + g - |f - g|) \end{aligned} \right\} \in cl(\mathcal{R}).$$

□

Mit Hilfe dieser Lemmata beweisen wir den folgenden Approximationssatz für reellwertige Funktionen:

Satz 4.38 (Approximationssatz von Stone–Weierstraß (1))

Sei X ein kompakter metrischer Raum und $\mathcal{R} \subset C(X, \mathbb{R})$ eine Unteralgebra im Banachraum der reellwertigen Funktionen auf X mit folgenden Eigenschaften:

1. \mathcal{R} enthält alle konstanten Funktionen,
2. \mathcal{R} trennt Punkte, d.h. für beliebige voneinander verschiedene Punkte $x, y \in X$ existiert eine Funktion $f \in \mathcal{R}$ mit $f(x) \neq f(y)$.

Dann ist \mathcal{R} dicht in $(C(X, \mathbb{R}), \|\cdot\|_\infty)$, d.h. zu jedem $f \in C(X, \mathbb{R})$ existiert eine Folge von Funktionen (f_n) aus \mathcal{R} , die auf X gleichmäßig gegen f konvergiert.

Beweis. (1) Wir zeigen zunächst die folgende Eigenschaft von \mathcal{R} :

Sind x_0 und y_0 zwei verschiedene Punkte in X und r und s zwei reelle Zahlen, so existiert eine Funktion $h \in \mathcal{R}$ mit $h(x_0) = r$ und $h(y_0) = s$:

Da die Algebra \mathcal{R} Punkte trennt, existiert eine Funktion $f \in \mathcal{R}$ mit $f(x_0) \neq f(y_0)$. Sei $g : X \rightarrow \mathbb{R}$ die stetige Funktion

$$g(x) := \frac{f(x) - f(x_0)}{f(y_0) - f(x_0)}, \quad x \in X.$$

Da \mathcal{R} alle konstanten Funktionen enthält, liegt g in \mathcal{R} . Wir betrachten nun die Funktion

$$h := (s - r) \cdot g + r \in \mathcal{R}.$$

Dann gilt offensichtlich

$$h(x_0) = (s - r)g(x_0) + r = r \quad \text{und} \quad h(y_0) = (s - r)g(y_0) + r = s.$$

(2) Wir zeigen nun, dass \mathcal{R} dicht in $C(X, \mathbb{R})$ liegt.

Sei $f \in C(X, \mathbb{R})$. Wir müssen zeigen, dass für jedes $\varepsilon > 0$ ein $f_\varepsilon \in \mathcal{R}$ existiert, so dass

$\|f - f_\varepsilon\|_\infty < \varepsilon$:

Seien $a, b \in X$ zwei beliebige Punkte. Nach (1) existiert eine Funktion $h_{a,b} \in \mathcal{R}$ mit $h_{a,b}(a) = f(a)$ und $h_{a,b}(b) = f(b)$. (Falls $a = b$, so setzen wir in (1) $x_0 = a = b$, $r = f(a)$ und y_0, s beliebig). Da f und $h_{a,b}$ stetig sind, ist die Teilmenge

$$U_{a,b} := \left\{ x \in X \mid h_{a,b}(x) > f(x) - \frac{\varepsilon}{2} \right\} = (h_{a,b} - f)^{-1} \left(\left(-\frac{\varepsilon}{2}, +\infty \right) \right) \subset X$$

offen. Nach Definition ist $a \in U_{a,b}$. Sei $b \in X$ ein fixiertes Element. Dann ist $\mathcal{U}_b := \{U_{a,b}\}_{a \in X}$ eine offene Überdeckung von X . Da X kompakt ist, existiert eine endliche Teilüberdeckung $\mathcal{U}_b^* \subset \mathcal{U}_b$, d.h. es existieren endlich viele Punkte $a_1, a_2, \dots, a_n \in X$, so dass

$$X = U_{a_1,b} \cup U_{a_2,b} \cup \dots \cup U_{a_n,b}.$$

Wir betrachten die Funktion

$$f_b := \max\{h_{a_1,b}, \dots, h_{a_n,b}\}.$$

Nach Lemma 4.20 ist $f_b \in cl(\mathcal{R})$. Ist $x \in X$, so gilt $x \in U_{a_j,b}$ für einen Index $j \in \{1, \dots, n\}$. Damit erhalten wir

$$f_b(x) \geq h_{a_j,b}(x) > f(x) - \frac{\varepsilon}{2}.$$

Wir haben zur Funktion $f \in C(X, \mathbb{R})$ also zunächst für jedes $b \in X$ eine Funktion $f_b \in cl(\mathcal{R})$ gefunden, so dass

$$f_b(x) > f(x) - \frac{\varepsilon}{2} \quad \forall x \in X. \quad (*)$$

Da f und f_b stetig sind, ist die Teilmenge

$$V_b := \left\{ x \in X \mid f_b(x) < f(x) + \frac{\varepsilon}{2} \right\} = (f_b - f)^{-1} \left(\left(-\infty, \frac{\varepsilon}{2} \right) \right) \subset X$$

offen. Aus der Definition der Funktionen $h_{a,b}$ folgt, dass $f_b(b) = f(b)$, d.h. $b \in V_b$. Folglich ist $\mathcal{V} := \{V_b\}_{b \in X}$ eine offene Überdeckung von X . Da X kompakt ist, existiert eine endliche Teilüberdeckung $\mathcal{V}^* \subset \mathcal{V}$, d.h. es existieren endlich viele Punkte $b_1, b_2, \dots, b_k \in X$, so dass

$$X = V_{b_1} \cup V_{b_2} \cup \dots \cup V_{b_k}.$$

Wir betrachten die stetige Funktion

$$g := \min\{f_{b_1}, \dots, f_{b_k}\}.$$

Aus den Lemmata 4.19 und 4.20 folgt $g \in cl(\mathcal{R})$. Ist $x \in X$, so gilt $x \in V_{b_i}$ für einen Index $i \in \{1, \dots, k\}$. Damit erhalten wir

$$g(x) \leq f_{b_i}(x) < f(x) + \frac{\varepsilon}{2}.$$

Andererseits existiert für $x \in X$ ein $l \in \{1, \dots, k\}$ mit $g(x) = f_{b_l}(x)$ und folglich gilt wegen (*)

$$g(x) = f_{b_i}(x) > f(x) - \frac{\varepsilon}{2}.$$

Wir haben also eine stetige Funktion $g \in cl(\mathcal{R})$ gefunden, so dass

$$f(x) - \frac{\varepsilon}{2} < g(x) < f(x) - \frac{\varepsilon}{2} \quad \forall x \in X,$$

also

$$\|f - g\|_\infty < \frac{\varepsilon}{2}.$$

Da g im Abschluss von \mathcal{R} liegt, gibt es eine Funktion $f_\varepsilon \in \mathcal{R}$ mit

$$\|g - f_\varepsilon\|_\infty < \frac{\varepsilon}{2}.$$

Wir erhalten somit

$$\|f - f_\varepsilon\|_\infty \leq \|f - g\|_\infty + \|g - f_\varepsilon\|_\infty < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Damit ist bewiesen, dass die Unteralgebra \mathcal{R} dicht in $C(X, \mathbb{R})$ liegt. \square

Als nächstes beweisen wir einen analogen Approximationssatz für *komplexwertige* Funktionen.

Satz 4.39 (Approximationssatz von Stone–Weierstraß (2))

Sei X ein kompakter, metrischer Raum und $\mathcal{R} \subset C(X, \mathbb{C})$ eine Unteralgebra im Banachraum der stetigen, komplexwertigen Funktionen mit folgenden Eigenschaften:

1. \mathcal{R} enthält alle konstanten Funktionen,
2. \mathcal{R} trennt Punkte,
3. Wenn $f \in \mathcal{R}$, so ist $\bar{f} \in \mathcal{R}$.

Dann ist \mathcal{R} dicht in $C(X, \mathbb{C})$, d.h., für jede Funktion $f \in C(X, \mathbb{C})$ existiert eine Folge von Funktionen (f_n) in \mathcal{R} , die auf X gleichmäßig gegen f konvergiert.

Beweis. Wir betrachten die folgende Teilmenge im Raum der reellwertigen stetigen Funktionen:

$$\mathcal{R}^* := \{f \in \mathcal{R} \mid \operatorname{Im}(f) = 0\} \subset C(X, \mathbb{R}).$$

Dann gilt

1. Da \mathcal{R} eine komplexe Unteralgebra in $C(X, \mathbb{C})$ ist, ist \mathcal{R}^* eine reelle Unteralgebra in $C(X, \mathbb{R})$.
2. Da \mathcal{R} alle komplexen Konstanten enthält, enthält \mathcal{R}^* alle reellen Konstanten.
3. \mathcal{R}^* trennt ebenfalls Punkte: Seien $x, y \in X$ zwei verschiedene Punkte. Wie im 1. Teil des Beweises von Satz 4.38 zeigt man, dass eine Funktion $f \in \mathcal{R}$ existiert, so dass $f(x) = 0$ und $f(y) = 1$ gilt. Dann gilt für die Funktion $f^* = f + \bar{f} \in \mathcal{R}^*$, dass $f^*(x) = 0$ und $f^*(y) = 2$. f^* trennt also die Punkte x und y .

Nach dem Approximationssatz von Stone-Weierstrass für reellwertige Funktionen (Satz 4.38), liegt dann \mathcal{R}^* dicht im Banachraum $C(X, \mathbb{R})$. Folglich ist die Menge der Funktionen $\mathcal{R}^* + i\mathcal{R}^*$ dicht im Banachraum $C(X, \mathbb{C})$. Da \mathcal{R} eine komplexe Unteralgebra ist und $\mathcal{R}^* \subset \mathcal{R}$, gilt $\mathcal{R}^* + i\mathcal{R}^* \subset \mathcal{R}$. Daraus erhält man

$$C(X, \mathbb{C}) = cl(\mathcal{R}^* + i\mathcal{R}^*) \subset cl(\mathcal{R}) \subset C(X, \mathbb{C}),$$

also $cl(\mathcal{R}) = C(X, \mathbb{C})$. □

Ist $X = [a, b] \subset \mathbb{R}$ und $P([a, b], \mathbb{R}) \subset C([a, b], \mathbb{R})$ die Menge der reellen Polynome, so erfüllt $P([a, b], \mathbb{R})$ die Bedingungen an die Unteralgebra \mathcal{R} im Satz 4.38. Die Menge der komplexen Polynome $P([a, b], \mathbb{C}) \subset C([a, b], \mathbb{C})$ erfüllt die Bedingungen der Unteralgebra \mathcal{R} im Satz 4.39.

Wir wenden nun die Approximationssätze von Stone-Weierstraß auf eine weitere Klasse von Funktionen an und beweisen Approximationssätze für periodische Funktionen.

Eine Funktion $f : \mathbb{R} \rightarrow \mathbb{K}$ heißt 2π -periodisch, wenn $f(x + 2\pi) = f(x)$ für alle $x \in \mathbb{R}$. Wir bezeichnen den \mathbb{K} -Vektorraum der 2π -periodischen Funktionen mit

$$C(\mathbb{R}, \mathbb{K})^{per} := \{f \in C(\mathbb{R}, \mathbb{K}) \mid f \text{ ist } 2\pi\text{-periodisch}\}.$$

Obwohl \mathbb{R} nicht kompakt ist, können wir wegen der Periodizität die folgende Maximum-Norm auf $C(\mathbb{R}, \mathbb{K})^{per}$ definieren: Für $f \in C(\mathbb{R}, \mathbb{K})^{per}$ ist

$$\|f\|_\infty := \max\{|f(x)| \mid x \in \mathbb{R}\} = \max\{|f(x)| \mid x \in [0, 2\pi]\}.$$

Wir werden nun die periodischen Funktionen $C(\mathbb{R}, \mathbb{K})^{per}$ mit den stetigen Funktionen auf der Kreislinie $S^1 := \{z \in \mathbb{C} \mid |z| = 1\}$ identifizieren.

Lemma 4.21. *Die normierten Vektorräume $(C(\mathbb{R}, \mathbb{K})^{per}, \|\cdot\|_\infty)$ und $(C(S^1, \mathbb{K}), \|\cdot\|_\infty)$ sind isometrisch.*

Beweis. Sei $z \in S^1$. Nach Satz 4.32 wissen wir, dass es genau eine reelle Zahl $t(z) \in [0, 2\pi)$ gibt mit $z = e^{it(z)} = \cos(t(z)) + i \sin(t(z))$. Für $f \in C(\mathbb{R}, \mathbb{K})^{per}$ definieren wir $f^* : S^1 \rightarrow \mathbb{K}$ durch

$$f^*(z) := f(t(z)).$$

Da die Umkehrfunktionen von Sinus und Cosinus stetig sind und f 2π -periodisch, ist f^* stetig (Übungsaufgabe). Die Abbildung

$$\begin{array}{ccc} \phi : C(\mathbb{R}, \mathbb{K})^{per} & \longrightarrow & C(S^1, \mathbb{K}) \\ f & \longmapsto & f^* \end{array}$$

ist ein Vektorraum-Isomorphismus, der auch mit dem Produkt von Funktionen verträglich ist (d.h. ein Algebren-Isomorphismus). Die Umkehrabbildung von ϕ ist durch

$$\phi^{-1}(h)(t) := h(e^{it}), \quad t \in \mathbb{R}$$

gegeben, wobei hier $h \in C(S^1, \mathbb{K})$. ϕ ist isometrisch, da für alle $f \in C(\mathbb{R}, \mathbb{K})^{per}$

$$\begin{aligned} \|\phi(f)\|_\infty &= \|f^*\|_\infty = \max\{|f^*(z)| \mid z \in S^1\} \\ &= \max\{|f(t(z))| \mid z \in S^1\} = \max\{|f(t)| \mid t \in \mathbb{R}\} \\ &= \|f\|_\infty. \end{aligned}$$

□

Definition 4.22. Ein reelles trigonometrisches Polynom ist eine 2π -periodische Funktion $f \in C(\mathbb{R}, \mathbb{R})^{per}$ mit

$$f(t) := a_0 + \sum_{k=1}^m \left(a_k \cos(kt) + b_k \sin(kt) \right) \quad \forall t \in \mathbb{R},$$

wobei $a_0, a_1, \dots, a_m, b_1, \dots, b_m \in \mathbb{R}$ und $m \in \mathbb{N}$.

Satz 4.40 Sei $f \in C(\mathbb{R}, \mathbb{R})^{per}$ eine stetige, 2π -periodische, reellwertige Funktion. Dann existiert eine Folge reeller trigonometrischer Polynome (f_n) , die auf \mathbb{R} gleichmäßig gegen f konvergiert.

Beweis. Wir bezeichnen mit \mathcal{P} die Menge der reellen trigonometrischen Polynome. Wegen der Additionstheoreme

$$\begin{aligned} \cos(kt) \cdot \cos(lt) &= \frac{1}{2} \{ \cos((k+l)t) + \cos((k-l)t) \} \\ \cos(kt) \cdot \sin(lt) &= \frac{1}{2} \{ \sin((k+l)t) - \sin((k-l)t) \} \\ \sin(kt) \cdot \sin(lt) &= \frac{1}{2} \{ \cos((k-l)t) - \cos((l+k)t) \} \end{aligned}$$

ist \mathcal{P} eine Unteralgebra von $C(\mathbb{R}, \mathbb{R})^{per}$. Wir betrachten die Unteralgebra $\mathcal{P}^* := \phi(\mathcal{P}) \subset C(S^1, \mathbb{R})$ und wenden auf sie den Satz von Stone-Weierstraß 4.38 an. \mathcal{P} enthält per Definition alle konstanten Funktionen. Für $f(t) := \cos(t)$ gilt $f^*(z) = \operatorname{Re}(z)$, für $f(t) := \sin(t)$ gilt $f^*(z) = \operatorname{Im}(z)$. Also trennt \mathcal{P}^* Punkte. Somit liegt \mathcal{P}^* dicht in $C(S^1, \mathbb{R})$ und aus Lemma 4.21 folgt, dass \mathcal{P} dicht in $C(\mathbb{R}, \mathbb{R})^{per}$ ist. □

Definition 4.23. Ein komplexes, trigonometrisches Polynom ist eine 2π -periodische Funktion $f \in C(\mathbb{R}, \mathbb{C})^{per}$ mit

$$f(t) = \sum_{k=-m}^M c_k e^{ikt},$$

wobei $m, M \in \mathbb{N}_0$ und $c_k \in \mathbb{C}$.

Satz 4.41 Sei $f \in C(\mathbb{R}, \mathbb{C})^{per}$ eine stetige, 2π -periodische, komplexwertige Funktion. Dann existiert eine Folge komplexer trigonometrischer Polynome (f_n) , die auf \mathbb{R} gleichmäßig gegen f konvergiert.

Beweis. Wir gehen analog zu Satz 4.40 vor und betrachten die Menge der komplexen trigonometrischen Polynome $\mathcal{Q} \subset C(\mathbb{R}, \mathbb{C})^{per} \simeq C(S^1, \mathbb{C})$. Man zeigt leicht, dass \mathcal{Q} die Eigenschaften der Unteralgebra aus Satz 4.39 erfüllt. □

Differentialrechnung für Funktionen einer reellen Variablen

In diesem Kapitel behandeln wir die Differentialrechnung für Funktionen einer reellen Variablen mit Werten in normierten Vektorräumen. 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.

5.1 Differenzierbare Abbildungen

In diesem Kapitel bezeichnet E einen normierten Vektorraum über dem Körper \mathbb{K} der reellen oder komplexen Zahlen mit der Norm $\|\cdot\|$ und $I \subset \mathbb{R}$ ein reelles Intervall¹.

Definition 5.1. Eine Abbildung $f : I \subset \mathbb{R} \rightarrow E$ heißt differenzierbar in $x_0 \in I$, 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 Abbildung $f : I \subset \mathbb{R} \rightarrow E$ heißt differenzierbar, falls sie in jedem Punkt $x_0 \in I$ differenzierbar ist. In diesem Fall heißt die Abbildung

$$\begin{aligned} f' : I \subset \mathbb{R} &\longrightarrow E \\ x &\longmapsto f'(x) \end{aligned}$$

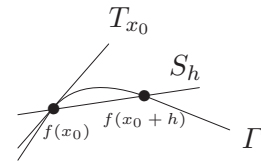
1. Ableitung von f .

¹ Hier sind auch alle Sorten verallgemeinerter Intervalle zugelassen, die mehr als einen Punkt enthalten.

Geometrische Interpretation: Wir betrachten das Bild der Abbildung $f : I \subset \mathbb{R} \rightarrow E$ als Kurve im Raum E . Für $E = \mathbb{R}^2$ wäre $\Gamma := f(I) \subset \mathbb{R}^2$ z.B. eine Kurve in der Ebene. Wir setzen voraus, dass f differenzierbar ist und $f'(x_0) \neq 0$ für alle $x_0 \in I$ gilt.

Wir betrachten die Sekante durch die Kurvenpunkte $f(x_0)$ und $f(x_0 + h)$ von Γ . Sie wird durch die Geradengleichung

$$S_h(x) := f(x_0) + \underbrace{\frac{f(x_0 + h) - f(x_0)}{h}}_{\text{Richtungsvektor}} \cdot (x - x_0), \quad x \in \mathbb{R},$$



beschrieben. Der Differenzenquotient $\frac{f(x_0+h)-f(x_0)}{h} \in E$ beschreibt also den Richtungsvektor der Sekante durch $f(x_0)$ und $f(x_0 + h)$. Da

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h},$$

konvergieren die die Sekanten beschreibenden Richtungsvektoren bei $h \rightarrow 0$ gegen den Vektor $f'(x_0) \in E$. Somit konvergieren die Sekanten S_h bei $h \rightarrow 0$ gegen die durch

$$T_{x_0}(x) := f(x_0) + f'(x_0)(x - x_0), \quad x \in \mathbb{R},$$

beschriebene Gerade in E . Diese Gerade heißt *Tangente* an die Kurve Γ im Punkt $f(x_0)$. Die Tangente T_{x_0} approximiert die Abbildung f bzw. die durch sie beschriebene Kurve $\Gamma \subset E$ nahe x_0 *in erster Näherung*, das heißt, es gilt

$$f(x_0) = T_{x_0}(x_0) \quad \text{und} \quad \lim_{x \rightarrow x_0} \frac{f(x) - T_{x_0}(x)}{x - x_0} = 0.$$

Interpretation in der Mechanik: In der Mechanik beschreibt man die Bewegung eines Massenpunktes bzw. eines Objektes P im Raum durch eine differenzierbare Abbildung $s : I \subset \mathbb{R} \rightarrow \mathbb{R}^3$. Dabei ist I das Zeitintervall, in dem man die Bewegung beschreiben will und $s(t) \in \mathbb{R}^3$ ist der Ort, an dem sich der Massenpunkt bzw. das Objekt P zum Zeitpunkt t befindet. Dann ist die Ortsänderung pro Zeiteinheit gleich dem Differenzenquotienten $\frac{s(t)-s(t_0)}{t-t_0}$. Folglich gibt der Vektor $s'(t_0) = \lim_{t \rightarrow t_0} \frac{s(t)-s(t_0)}{t-t_0} \in \mathbb{R}^3$ die Geschwindigkeit von P zur Zeit $t = t_0$ an und $\|s'(t_0)\|$ ihren Betrag.

Interpretation als Anstieg der Tangente an den Funktionsgraphen:

Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ eine reellwertige Funktion. Dann beschreibt der Differenzenquotient $\frac{f(x_0+h)-f(x_0)}{h}$ den Anstieg der Sekante durch die Punkte $(x_0, f(x_0))$ und $(x_0 + h, f(x_0 + h))$ des Funktionsgraphen. $f'(x_0)$ ist dann der Anstieg der Tangente an den Funktionsgraphen im Punkt $(x_0, f(x_0))$.

Satz 5.1 *Ist eine Abbildung $f : I \subset \mathbb{R} \rightarrow E$ in $x_0 \in I$ differenzierbar, so ist sie in x_0 auch stetig. Jede differenzierbare Abbildung $f : I \subset \mathbb{R} \rightarrow E$ ist somit stetig.*

Beweis. Sei (x_n) eine Folge in I , die gegen x_0 konvergiert. OBdA. sei $x_n \neq x_0$ für alle $n \in \mathbb{N}$. Die Bildfolge $(f(x_n))$ konvergiert gegen $f(x_0)$, denn unter Ausnutzung der Stetigkeit der Norm $\|\cdot\|$ erhalten wir:

$$\|f(x_n) - f(x_0)\| = |x_n - x_0| \left\| \frac{f(x_n) - f(x_0)}{x_n - x_0} \right\| \xrightarrow{n \rightarrow \infty} 0 \cdot \|f'(x_0)\| = 0.$$

Somit ist f in $x_0 \in I$ folgenstetig, also stetig. \square

Beispiel: Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) := |x|$ ist stetig, aber in $x = 0$ nicht differenzierbar.

Satz 5.2 *Es seien $f, g : I \subset \mathbb{R} \rightarrow E$ sowie $h : I \subset \mathbb{R} \rightarrow \mathbb{K}$ Abbildungen, die in $x_0 \in I$ differenzierbar sind. Dann gilt:*

1. $f + g : I \rightarrow E$ ist in x_0 differenzierbar und $(f + g)'(x_0) = f'(x_0) + g'(x_0)$.
2. $h \cdot f : I \rightarrow E$ ist in x_0 differenzierbar und

$$(h \cdot f)'(x_0) = h'(x_0) \cdot f(x_0) + h(x_0) \cdot f'(x_0).$$

3. Sei $h(x_0) \neq 0$. Dann ist $\frac{f}{h} : \{x \in I \mid h(x) \neq 0\} \rightarrow E$ in x_0 differenzierbar und

$$\left(\frac{f}{h}\right)'(x_0) = \frac{h(x_0)f'(x_0) - h'(x_0)f(x_0)}{h^2(x_0)}.$$

Insbesondere gilt auch

$$\left(\frac{1}{h}\right)'(x_0) = -\frac{h'(x_0)}{h^2(x_0)}.$$

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

$$\frac{(h \cdot f)(x) - (h \cdot f)(x_0)}{x - x_0} = h(x) \cdot \frac{f(x) - f(x_0)}{x - x_0} + \frac{h(x) - h(x_0)}{x - x_0} \cdot f(x_0),$$

wobei wir auch die Stetigkeit von h in x_0 benutzen.

3. Die letztere Behauptung folgt durch Limesbildung aus

$$\frac{\left(\frac{1}{h}\right)(x) - \left(\frac{1}{h}\right)(x_0)}{x - x_0} = \frac{\frac{1}{h(x)} - \frac{1}{h(x_0)}}{x - x_0} = \frac{h(x_0) - h(x)}{x - x_0} \cdot \frac{1}{h(x)h(x_0)} \xrightarrow{x \rightarrow x_0} -h'(x_0) \cdot \frac{1}{h^2(x_0)}.$$

Mit Hilfe der Produktregel erhält man dann sofort die Quotientenregel. \square

Aus diesen Rechenregeln erhält man sofort das folgende Beispiel:

Beispiel: Seien a_0, a_1, \dots, a_n Vektoren aus E und $p : \mathbb{R} \rightarrow E$ die Abbildung

$$p(x) := x^n a_n + x^{n-1} a_{n-1} + \dots + x a_1 + a_0.$$

Dann ist p auf \mathbb{R} differenzierbar und für die Ableitung gilt:

$$p'(x) = n x^{n-1} a_n + (n-1) x^{n-2} a_{n-1} + \dots + 2 x a_2 + a_1.$$

Satz 5.3 (Kettenregel) *Es seien $I, J \subset \mathbb{R}$ Intervalle, die Abbildung $g : J \rightarrow I$ in $x_0 \in J$ differenzierbar und die Abbildung $f : I \rightarrow E$ in $g(x_0)$ differenzierbar. Dann ist $f \circ g : J \rightarrow E$ in x_0 differenzierbar und es gilt²*

$$(f \circ g)'(x_0) = g'(x_0) \cdot f'(g(x_0)).$$

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) . □

Satz 5.4 (Ableitung der inversen Abbildung) *Sei $f : I \rightarrow J$ eine bijektive Abbildung zwischen zwei Intervallen und $x_0 \in I$. f habe folgende Eigenschaften:*

- (1) *Die Umkehrabbildung $f^{-1} : J \rightarrow I$ ist in $f(x_0)$ stetig,*
- (2) *Die Abbildung f ist in x_0 differenzierbar und $f'(x_0) \neq 0$.*

Dann ist die Umkehrabbildung $f^{-1} : J \rightarrow I$ in $f(x_0)$ differenzierbar und für ihre Ableitung gilt:

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}.$$

Beweis. Für $y \in J$ mit $f(x_0) \neq y$ und $x := f^{-1}(y) \in I$ gilt:

$$\frac{f^{-1}(y) - f^{-1}(f(x_0))}{y - f(x_0)} = \frac{x - x_0}{f(x) - f(x_0)} = \frac{1}{\frac{f(x) - f(x_0)}{x - x_0}}. \quad (**)$$

² Die Reihenfolge der Faktoren wurde so gewählt, weil man die Vektoren aus E nach Vereinbarung von links mit den Skalaren multipliziert.

Sei nun (y_n) eine Folge in J mit $y_n \neq f(x_0)$, die gegen $f(x_0)$ konvergiert. Da f^{-1} in $f(x_0)$ stetig ist, konvergiert die Folge $(x_n := f^{-1}(y_n))$ gegen $x_0 = f^{-1}(f(x_0))$. Da f in x_0 differenzierbar ist, folgt dann mit (**)

$$\lim_{n \rightarrow \infty} \frac{f^{-1}(y_n) - f^{-1}(f(x_0))}{y_n - f(x_0)} = \frac{1}{f'(x_0)}.$$

□

Ohne die Voraussetzung $f'(x_0) \neq 0$ ist die inverse Abbildung in $f(x_0)$ i.a. nicht differenzierbar. Wir betrachten dazu als Beispiel die bijektive und differenzierbare Abbildung $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) = x^3$. f ist bijektiv und differenzierbar, es gilt $f(0) = 0$ und $f'(0) = 0$. Die inverse Abbildung $f^{-1} : \mathbb{R} \rightarrow \mathbb{R}$ 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 x gegen 0^+ nicht, da

$$\frac{f^{-1}(x) - f^{-1}(0)}{x} = \frac{\sqrt[3]{x}}{x} = \frac{1}{\sqrt[3]{x^2}}.$$

Ableitungen elementarer Funktionen

1. Die Exponentialfunktion $\exp : \mathbb{R} \rightarrow \mathbb{R}$

$$\exp(x) := e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

ist differenzierbar und es gilt $\exp' = \exp$.

Beweis. Die Potenzreihe

$$\frac{e^h - 1}{h} = \sum_{n=1}^{\infty} \frac{h^{n-1}}{n!} = 1 + \frac{h}{2!} + \frac{h^2}{3!} + \dots$$

ist auf dem kompakten Intervall $[-1, 1] \subset \mathbb{R}$ gleichmäßig konvergent. Deswegen sind \lim und \sum auf diesem Intervall vertauschbar und wir erhalten

$$\lim_{h \rightarrow 0} \frac{e^h - 1}{h} = \sum_{n=1}^{\infty} \lim_{h \rightarrow 0} \frac{h^{n-1}}{n!} = 1.$$

Daraus folgt für $x \in \mathbb{R}$

$$\exp'(x) = \lim_{h \rightarrow 0} \frac{e^{x+h} - e^x}{h} = \lim_{h \rightarrow 0} e^x \cdot \frac{e^h - 1}{h} = e^x = \exp(x).$$

2. 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 und $\exp'(y) = \exp(y) \neq 0$ für alle $y \in \mathbb{R}$. Somit sind die Voraussetzungen des Satzes 5.4 erfüllt und es folgt

$$\ln'(e^y) = \frac{1}{(e^y)'} = \frac{1}{e^y}.$$

Somit ist

$$\ln'(x) = \frac{1}{x} \quad \text{für alle } x \in \mathbb{R}^+.$$

3. Die Potenzfunktion $p_a : \mathbb{R}^+ \rightarrow \mathbb{R}$ mit dem Exponenten $a \in \mathbb{R}$:

$$p_a(x) := x^a = e^{\ln(x) \cdot a}$$

ist differenzierbar und es gilt $p_a'(x) = a \cdot x^{a-1}$ für alle $x \in \mathbb{R}^+$.

Beweis. Nach der Kettenregel (Satz 5.3) gilt:

$$p_a'(x) = \exp'(\ln(x) \cdot a) \cdot a \cdot \ln'(x) = e^{\ln(x) \cdot a} \cdot a \cdot \frac{1}{x} = \frac{x^a}{x} \cdot a = a \cdot x^{a-1}.$$

4. Die Exponentialfunktion $\exp_b : \mathbb{R} \rightarrow \mathbb{R}^+$ zur Basis $b \in \mathbb{R}^+$:

$$\exp_b(x) := b^x := e^{\ln(b) \cdot x}$$

ist differenzierbar und es gilt nach Kettenregel

$$(\exp_b)'(x) = e^{\ln(b) \cdot x} \cdot \ln(b) = \ln(b) \cdot b^x = \ln(b) \cdot \exp_b(x) \quad \text{für alle } x \in \mathbb{R}.$$

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. Die Reihen $\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$ und $\cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$ sind gleichmäßig konvergent auf dem kompakten Intervall $[-1, 1] \subset \mathbb{R}$. Deshalb können wir \lim und \sum auf diesem Intervall vertauschen und erhalten:

$$\lim_{h \rightarrow 0} \frac{\sin(h)}{h} = \sum_{n=0}^{\infty} (-1)^n \frac{1}{(2n+1)!} \cdot \lim_{h \rightarrow 0} h^{2n} = 1$$

und

$$\lim_{h \rightarrow 0} \frac{\cos(h) - 1}{h} = \sum_{n=1}^{\infty} (-1)^n \lim_{h \rightarrow 0} \frac{h^{2n-1}}{(2n)!} = 0.$$

Daher gilt 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

$$\sinh(x) = \frac{1}{2}(e^x - e^{-x}) \quad \text{und} \quad \cosh(x) = \frac{1}{2}(e^x + e^{-x}).$$

Daraus folgt sofort

$$\begin{aligned} (\sinh)'(x) &= \frac{1}{2}(e^x + e^{-x}) = \cosh(x), \\ (\cosh)'(x) &= \frac{1}{2}(e^x - e^{-x}) = \sinh(x). \end{aligned}$$

Wir definieren jetzt die höheren Ableitungen einer Abbildung $f : I \subset \mathbb{R} \rightarrow E$.

Definition 5.2. Sei $f : I \subset \mathbb{R} \rightarrow E$ eine differenzierbare Abbildung. f heißt in $x_0 \in I$ 2-mal differenzierbar, falls die Ableitung von $f' : I \rightarrow E$ in x_0 existiert.

$f''(x_0) := \frac{d^2 f}{dx^2}(x_0) := (f')'(x_0)$ heißt die 2.-te Ableitung von f in x_0 .

Existiert die $(n-1)$ -te Ableitung $f^{(n-1)} : I \subset \mathbb{R} \rightarrow E$ von f und ist sie in x_0 differenzierbar, so nennt man f in x_0 n -mal differenzierbar ($n \in \mathbb{N}$, $n \geq 2$).

$f^{(n)}(x_0) := \frac{d^n f}{dx^n}(x_0) := (f^{(n-1)})'(x_0)$ heißt die n -te Ableitung von f in x_0 .

Definition 5.3. Eine Abbildung $f : I \subset \mathbb{R} \rightarrow E$ heißt stetig differenzierbar, wenn f differenzierbar und $f' : I \rightarrow E$ stetig ist. $f : I \subset \mathbb{R} \rightarrow E$ heißt n -mal stetig differenzierbar, falls alle Ableitungen $f^{(1)}, f^{(2)}, \dots, f^{(n)}$ existieren und stetig sind.

Eine Abbildung $f : I \subset \mathbb{R} \rightarrow E$ 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).

Bezeichnung: $C^{(k)}(I, E)$ bezeichne den Vektorraum aller k -fach stetig differenzierbaren Abbildungen von I nach E . $C^\infty(I, E)$ bezeichne den Vektorraum aller glatten Abbildungen von I nach E .

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

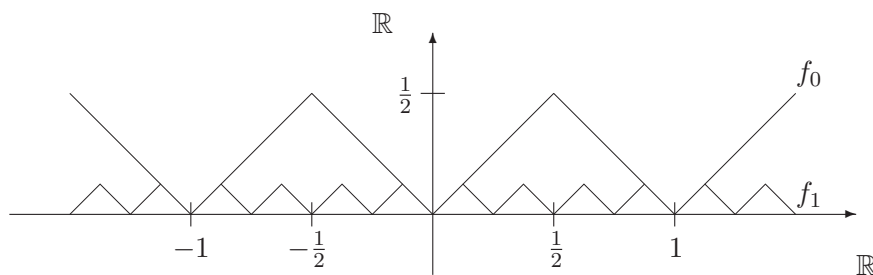
$$\frac{f(h) - f(0)}{h} = \frac{h^2 \cdot \sin\left(\frac{1}{h}\right)}{h} = h \cdot \underbrace{\sin\left(\frac{1}{h}\right)}_{\text{beschränkt}} \xrightarrow{h \rightarrow 0} 0,$$

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 $x = 0$.

Beispiel: Eine stetige, nirgends differenzierbare Funktion.

Folgende Betrachtungen (Tagaki, 1903) liefern uns eine auf ganz \mathbb{R} stetige, aber nirgends differenzierbare Funktion. Wir betrachten die stückweise lineare Funktion $f_0 : \mathbb{R} \rightarrow \mathbb{R}$, definiert durch

$$f_0(x) := \begin{cases} x & \text{falls } x \in [0, \frac{1}{2}] \\ 1 - x & \text{falls } x \in [\frac{1}{2}, 1] \end{cases} \quad \text{und} \quad f_0(x+1) = f_0(x) \quad \text{für alle } x \in \mathbb{R}.$$



Die Funktion f_0 hat also die Periode 1. Wir definieren nun für $n \in \mathbb{N}$ die Funktionen

$$f_n(x) := 4^{-n} f_0(4^n \cdot x) \quad x \in \mathbb{R}.$$

Dann hat f_n die Periode 4^{-n} . Wir betrachten nun die Funktionenreihe $\sum_{n=0}^{\infty} f_n$. Dann gilt:

1. Die Funktionen $f_n : \mathbb{R} \rightarrow \mathbb{R}$ sind stetig.

2. Die Funktionen f_n sind auf \mathbb{R} beschränkt, da $|f_n(x)| \leq \frac{1}{2}4^{-n} =: M_n$ für alle $x \in \mathbb{R}$. Aus den Eigenschaften der geometrischen Reihe folgt für die Reihe der Schranken:

$$\sum_{n=0}^{\infty} M_n = \frac{1}{2} \sum_{n=0}^{\infty} \left(\frac{1}{4}\right)^n = \frac{2}{3} < \infty.$$

Somit ist die Funktionenreihe $\sum_{n=0}^{\infty} f_n$ gleichmäßig konvergent auf \mathbb{R} . Wir bezeichnen mit f ihre Grenzfunktion.

Aus (1) und (2) folgt, dass $f : \mathbb{R} \rightarrow \mathbb{R}$ stetig ist. Trotzdem ist f nirgends differenzierbar, wie wir jetzt sehen werden:

Sei $x \in \mathbb{R}$ beliebig gegeben. Für $n \in \mathbb{N}$ wählen wir $h_n := \frac{1}{4} \cdot 4^{-n}$ oder $h_n := -\frac{1}{4} \cdot 4^{-n}$ derart, dass f_n zwischen x und $x + h_n$ linear ist. Nach Konstruktion sind dann f_k für alle $k \leq n$ zwischen x und $x + h_n$ linear. Da die linearen Abschnitte von f_n den Anstieg ± 1 haben, folgt

$$\frac{f_k(x + h_n) - f_k(x)}{h_n} = \pm 1 \quad \forall k \leq n.$$

Für $k > n$ ist $h_n = 4^{-(n+1)}$ eine Periode von f_k . Folglich ist $\frac{f_k(x+h_n)-f_k(x)}{h_n} = 0$ für alle $k > n$ und daher gilt

$$\frac{f(x + h_n) - f(x)}{h_n} = \sum_{k=0}^n \frac{f_k(x + h_n) - f_k(x)}{h_n} = \sum_{k=0}^n \pm 1.$$

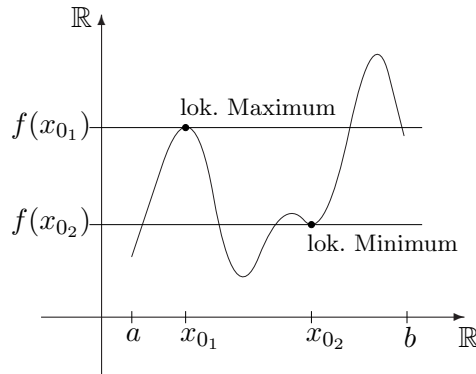
Diese Partialsummenfolge kann für $n \rightarrow \infty$ nicht konvergieren, da die Reihenglieder keine Nullfolge sind. Somit existiert der Grenzwert $f'(x) = \lim_{n \rightarrow \infty} \frac{f(x+h_n)-f(x)}{h_n}$ nicht.

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 der durch f definierten Kurve $graph(f) := \{(x, f(x) \mid x \in I\}$ machen.

Definition 5.4. Man sagt: Eine Funktion $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ nimmt in $x_0 \in I$ ein lokales Maximum (bzw. lokales Minimum) an, falls ein $\varepsilon > 0$ existiert, so dass

$$f(y) \leq f(x_0) \quad (\text{bzw. } f(y) \geq f(x_0)) \quad \text{für alle } y \in I \text{ mit } |y - x_0| < \varepsilon.$$



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 (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:

$$\lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \leq 0, \quad \text{und} \quad \lim_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \geq 0. \quad (\star)$$

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 (\star) 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 entweder

$$f(x_1) > f(a) = f(b) \quad \text{oder} \quad f(x_1) < f(a) = f(b).$$

Wir betrachten zunächst den ersten Fall. Da $f : [a, b] \rightarrow \mathbb{R}$ stetig und $[a, b]$ kompakt ist, nimmt f auf $[a, b]$ ein globales Maximum an. Das Maximum liegt wegen obiger Annahme nicht auf den Randpunkten. Folglich existiert ein $x_0 \in (a, b)$ mit

$$f(x_0) = \max\{f(x) \mid x \in [a, b]\}.$$

Nach Satz 5.5 folgt dann $f'(x_0) = 0$. Der Beweis für den zweiten Fall 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

$$(f(b) - f(a)) \cdot g'(\xi) = (g(b) - g(a)) \cdot f'(\xi).$$

Beweis. Wir betrachten 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

$$0 = \varphi'(\xi) = (f(b) - f(a))g'(\xi) - (g(b) - g(a))f'(\xi).$$

□

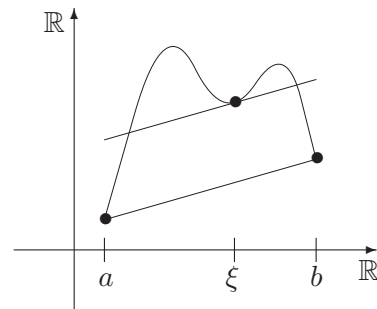
Satz 5.8 (Mittelwertsatz von Lagrange)

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 $g(x) := x$. □

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.



Wir werden den Mittelwertsatz von Lagrange jetzt anwenden, um Aussagen über den Kurvenverlauf der durch $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ definierten Kurve $\Gamma = \text{graph}(f)$ im \mathbb{R}^2 zu machen.

Satz 5.9 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.
Ist $f'(x) > 0$ für alle $x \in (a, b)$, so ist f streng monoton wachsend.
2. Ist $f'(x) \leq 0$ (< 0) für alle $x \in (a, b)$, so ist f (streng) monoton fallend.
3. 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 aus Satz 5.8. □

³ Sofern nichts anderes gesagt wird, lassen wir im Folgenden bei offenen Intervallen (a, b) auch $a, b \in \{\pm\infty\}$ zu.

Als Anwendung beweisen wir zunächst die folgende Charakterisierung der Exponentialfunktion.

Folgerung 5.1 *Seien $a, c \in \mathbb{R}$ gegebene Konstanten. Dann existiert genau eine differenzierbare Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f' = cf$ und $f(0) = a$. Diese Funktion ist gegeben durch $f(x) = ae^{cx}$ für alle $x \in \mathbb{R}$.*

Beweis. Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x) := ae^{cx}$ erfüllt $f'(x) = cae^{cx} = cf(x)$ und $f(0) = a$. Sei nun $f : \mathbb{R} \rightarrow \mathbb{R}$ eine beliebige Funktion mit $f' = cf$ und $f(0) = a$. Wir setzen $F(x) := f(x) \cdot e^{-cx}$. Dann gilt

$$F'(x) = f'(x) \cdot e^{-cx} - c \cdot f(x) \cdot e^{-cx} = (cf - cf)e^{-cx} = 0.$$

Somit ist F konstant. Da $F(0) = a$, gilt $a = f(x) \cdot e^{-cx}$ und folglich $f(x) = ae^{cx}$ für alle $x \in \mathbb{R}$. \square

Satz 5.10 *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), das heißt, es existiert ein $\varepsilon > 0$, so dass $f(x) < f(x_0)$ (bzw. $f(x) > f(x_0)$) für alle $x \in (a, b)$ mit $0 < |x - x_0| < \varepsilon$.

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 5.9 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. \square

Definition 5.5. Eine Funktion $f : (a, b) \rightarrow \mathbb{R}$ heißt

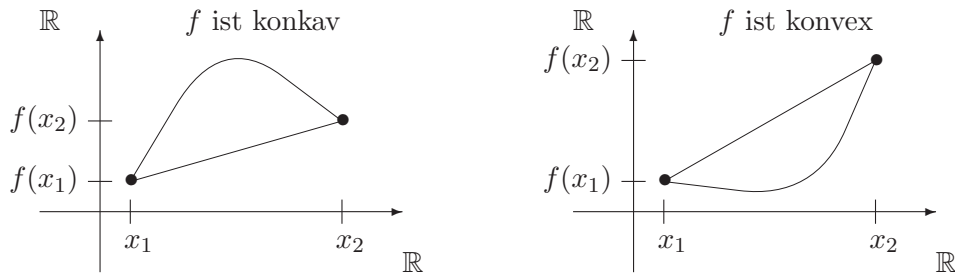
- *konvex*, falls für alle $x_1, x_2 \in (a, b)$ und $\lambda \in (0, 1)$

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

- *konkav*, falls für alle $x_1, x_2 \in (a, b)$ und $\lambda \in (0, 1)$

$$f(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

Ist f konvex, so liegt die Kurve $\text{graph}(f) := \{(x, f(x)) \mid x \in (a, b)\} \subset \mathbb{R}^2$ 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 $\text{graph}(f)$ 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 5.11 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 5.9 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 ?? 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 5.12 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 5.11. □

Wir haben den Mittelwertsatz der Differentialrechnung für reellwertige Funktionen bewiesen. Für komplexwertige oder vektorwertige Funktionen gilt der Mittelwertsatz in dieser Form nicht mehr. Wir betrachten als Beispiel die Funktion $f : [0, 2\pi] \rightarrow \mathbb{R}^2$:

$$f(t) := (\cos(t), \sin(t)).$$

Es gilt $f(2\pi) - f(0) = (0, 0)$, also $\frac{f(2\pi) - f(0)}{2\pi - 0} = (0, 0)$. Für die Ableitung erhalten wir $f'(t) = (-\sin(t), \cos(t))$. Somit ist $\|f'(t)\| = 1$, d.h. $f'(t) \neq (0, 0)$ für alle $t \in [0, 2\pi]$.

Es gilt aber der folgende Mittelwertsatz für vektorwertige (und damit auch für komplexwertige) Funktionen:

Satz 5.13 (Mittelwertsatz für vektorwertige Funktionen)

Sei $(E, \|\cdot\|)$ ein normierter Vektorraum, $f : [a, b] \subset \mathbb{R} \rightarrow E$ stetig und auf (a, b) differenzierbar. Dann existiert ein Punkt $\xi \in (a, b)$, so dass

$$\|f(b) - f(a)\| \leq (b - a)\|f'(\xi)\|.$$

Beweis. 1. Zum Beweis des Satzes benötigen wir das folgende Lemma:

Lemma 5.7. Sei $f : [a, b] \rightarrow E$ in $x_0 \in (a, b)$ differenzierbar, und seien (α_k) und (β_k) Folgen in $[a, b]$ mit $\alpha_k < x_0 < \beta_k$, die beide gegen x_0 konvergieren. Dann gilt

$$\lim_{k \rightarrow \infty} \frac{f(\beta_k) - f(\alpha_k)}{\beta_k - \alpha_k} = f'(x_0).$$

Beweis. Wir betrachten $y_k := \frac{\beta_k - x_0}{\beta_k - \alpha_k} \in (0, 1)$. Dann gilt:

$$\begin{aligned} & \frac{f(\beta_k) - f(\alpha_k)}{\beta_k - \alpha_k} - f'(x_0) \\ &= \underbrace{y_k}_{\text{beschr.}} \cdot \left(\frac{f(\beta_k) - f(x_0)}{\beta_k - x_0} - f'(x_0) \right) + \underbrace{(1 - y_k)}_{\text{beschr.}} \cdot \left(\frac{f(\alpha_k) - f(x_0)}{\alpha_k - x_0} - f'(x_0) \right). \end{aligned}$$

Für $k \rightarrow \infty$ konvergieren beide Summanden gegen Null. □

2. Wir beweisen nun die Aussage des Satzes. Sei $L := \frac{b-a}{3}$ und $M := \|f(b) - f(a)\|$. Wir betrachten die Funktion $g : [a, a + 2L] \subset \mathbb{R} \rightarrow E$, definiert durch $g(s) := f(s + L) - f(s)$. Dann gilt:

$$g(a) + g(a + L) + g(a + 2L) = f(b) - f(a)$$

und somit

$$M \leq \|g(a)\| + \|g(a + L)\| + \|g(a + 2L)\|. \quad (\star)$$

Wir überlegen uns zunächst, dass ein $s_1 \in (a, a + 2L)$ existiert mit $\|g(s_1)\| \geq \frac{1}{3}M$. Nehmen wir an, dass für alle $s \in (a, a + 2L)$ die Ungleichung $\|g(s)\| < \frac{1}{3}M$ gilt. Da die Funktion $g : [a, a + 2L] \rightarrow E$ und die Norm $\|\cdot\|$ stetig sind, erhalten wir auch

$$\|g(a)\| \leq \frac{1}{3}M \quad \text{und} \quad \|g(a + 2L)\| \leq \frac{1}{3}M.$$

Dies widerspricht (\star) , es gibt also ein $s_1 \in (a, a + 2L)$ mit $\|g(s_1)\| \geq \frac{1}{3}M$.

Sei nun $t_1 := s_1 + L \in (a, b)$. Dann gilt:

- a) $a < s_1 < t_1 < b$,
- b) $t_1 - s_1 = L = \frac{b-a}{3}$,
- c) $\|g(s_1)\| = \|f(t_1) - f(s_1)\| \geq \frac{1}{3}M$.

Wir wiederholen nun diese Konstruktion, indem wir statt des Intervalls $[a, b]$ das Intervall $[s_1, t_1]$ betrachten. Dadurch erhalten wir Folgen (s_n) und (t_n) , so dass gilt

$$[s_n, t_n] \subset (s_{n-1}, t_{n-1}), \quad t_n - s_n = \frac{1}{3^n}(b - a), \quad \|f(t_n) - f(s_n)\| \geq \frac{1}{3^n} \cdot M.$$

Wir haben also eine Intervallschachtelung $(I_n := [s_n, t_n])$, bei der die Länge der Intervalle I_n gegen 0 konvergiert. Nach dem Satz über die Intervallschachtelung existiert dann ein $\xi \in \mathbb{R}$ mit $\bigcap_{n=1}^{\infty} I_n = \{\xi\}$. Es gilt somit

$$s_n \rightarrow \xi, \quad t_n \rightarrow \xi, \quad s_n < \xi < t_n$$

und

$$\frac{\|f(t_n) - f(s_n)\|}{t_n - s_n} \geq \frac{M}{(t_n - s_n) \cdot 3^n} = \frac{M}{b - a}.$$

Aus Lemma 5.7 folgt nun

$$f'(\xi) = \lim_{n \rightarrow \infty} \frac{f(t_n) - f(s_n)}{t_n - s_n}$$

und damit

$$\begin{aligned} \|f'(\xi)\| &= \left\| \lim_{n \rightarrow \infty} \frac{f(t_n) - f(s_n)}{t_n - s_n} \right\| = \lim_{n \rightarrow \infty} \frac{\|f(t_n) - f(s_n)\|}{t_n - s_n} \\ &\geq \frac{M}{b - a} = \frac{\|f(b) - f(a)\|}{b - a}. \end{aligned}$$

□

Satz 5.14 Sei $f : I \subset \mathbb{R} \rightarrow E$ eine differenzierbare Funktion mit beschränkter Ableitung. Dann ist die Funktion f Lipschitzstetig. Insbesondere ist jede stetig differenzierbare Funktion $f : [a, b] \subset \mathbb{R} \rightarrow E$ Lipschitzstetig.

Beweis. Sei $f : I \rightarrow E$ 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 5.13, dass ein $\xi \in (x_1, x_2)$ existiert, so dass

$$\|f(x_2) - f(x_1)\| \leq \|f'(\xi)\| \cdot (x_2 - x_1) \leq C \cdot |x_2 - x_1|.$$

Dies aber bedeutet, dass f Lipschitzstetig mit der Lipschitz-Konstanten C ist. Sei nun $f : [a, b] \rightarrow E$ stetig differenzierbar. Dann ist f' stetig, d.h. $f' \in C([a, b], E)$. Da $[a, b]$ kompakt ist, existiert

$$\|f'\|_\infty := \max\{\|f'(x)\| \mid x \in [a, b]\} \in \mathbb{R}.$$

In diesem Fall können wir $C := \|f'\|_\infty$ setzen. □

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 5.15 (Die Regeln von L'Hospital - (1))

Sei $x_0 \in [a, b] \subset \mathbb{R}$ und seien $f, g : (a, b) \setminus \{x_0\} \rightarrow \mathbb{R}$ differenzierbare Funktionen mit

1. $\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) \in \{0, \pm\infty\}$,
2. $g'(x) \neq 0$ für alle $x \in (a, b) \setminus \{x_0\}$,
3. $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} = c \in \mathbb{R} \cup \{\pm\infty\}$.

Dann gilt $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = c$.

Beweis. 1. Fall: $\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = 0$.

Wir setzen f und g in x_0 durch $f(x_0) := g(x_0) := 0$ fort. Dann sind $f, g : (a, b) \rightarrow \mathbb{R}$ stetig. Wir betrachten nun eine gegen x_0 konvergente Folge (x_n) in $(a, b) \setminus \{x_0\}$. Dann sind $f, g : [x_n, x_0] \rightarrow \mathbb{R}$ stetig und auf (x_n, x_0) differenzierbar (falls $x_n < x_0$, ansonsten betrachten wir $[x_0, x_n]$). Nach dem verallgemeinerten Mittelwertsatz 5.7 existiert ein $\xi_n \in (x_n, x_0)$ (bzw. $\xi_n \in (x_0, x_n)$), so dass

$$f(x_n)g'(\xi_n) = g(x_n)f'(\xi_n).$$

Da $g'(x) \neq 0$ für $x \in (x_n, x_0)$ (bzw. (x_0, x_n)), ist $g(x_n) \neq 0$ (Satz von Rolle). Damit folgt

$$\frac{f(x_n)}{g(x_n)} = \frac{f'(\xi_n)}{g'(\xi_n)}.$$

Da (x_n) gegen x_0 konvergiert, ist auch (ξ_n) gegen x_0 konvergent und wir erhalten

$$\lim_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} = \lim_{n \rightarrow \infty} \frac{f'(\xi_n)}{g'(\xi_n)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} = c.$$

2. Fall: $\lim_{x \rightarrow x_0} g(x) = +\infty$ und $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} = 0$.

Sei (x_n) eine gegen x_0 konvergente Folge in (a, b) mit $x_n < x_0$ und $\varepsilon > 0$. Nach Voraussetzung existiert ein $x^* \in (a, x_0)$, so dass

$$g(x) > 0 \quad \text{und} \quad \left| \frac{f'(x)}{g'(x)} \right| < \varepsilon \quad \text{für alle } x \in (x^*, x_0).$$

Durch Ausmultiplizieren prüft man folgende Formel nach:

$$\frac{f(x_n)}{g(x_n)} = \frac{f(x^*)}{g(x_n)} + \frac{f(x_n) - f(x^*)}{g(x_n) - g(x^*)} \cdot \left(1 - \frac{g(x^*)}{g(x_n)} \right). \quad (*)$$

Da (x_n) gegen x_0 konvergiert, gibt es ein $n_0 \in \mathbb{N}$, so dass $x_n \in (x^*, x_0)$ für alle $n \geq n_0$. Nach verallgemeinertem Mittelwertsatz existiert ein $\xi_n \in (x^*, x_n)$ mit

$$\frac{f(x_n) - f(x^*)}{g(x_n) - g(x^*)} = \frac{f'(\xi_n)}{g'(\xi_n)}.$$

Eingesetzt in Gleichung $(*)$ ergibt dies

$$\left| \frac{f(x_n)}{g(x_n)} \right| \leq \left| \frac{f(x^*)}{g(x_n)} \right| + \underbrace{\left| \frac{f'(\xi_n)}{g'(\xi_n)} \right|}_{< \varepsilon} \left| 1 - \frac{g(x^*)}{g(x_n)} \right|$$

und damit

$$\left| \frac{f(x_n)}{g(x_n)} \right| \leq \left| \frac{f(x^*)}{g(x_n)} \right| + \varepsilon \left| 1 - \frac{g(x^*)}{g(x_n)} \right|.$$

Da $(g(x_n))$ gegen $+\infty$ konvergiert, folgt

$$\lim_{n \rightarrow \infty} \left| \frac{f(x^*)}{g(x_n)} \right| = 0 \quad \text{und} \quad \lim_{n \rightarrow \infty} \frac{g(x^*)}{g(x_n)} = 0.$$

Wir erhalten:

$$\limsup_{n \rightarrow \infty} \left| \frac{f(x_n)}{g(x_n)} \right| \leq \varepsilon.$$

Da ε beliebige Werte annehmen kann, folgt daraus

$$\lim_{n \rightarrow \infty} \frac{f(x_n)}{g(x_n)} = 0$$

für alle gegen x_0 konvergente Folgen (x_n) mit $x_n < x_0$.

Analog behandelt man gegen x_0 konvergente Folgen (x_n) mit $x_n > x_0$ und erhält das gleiche Resultat. Dies zeigt, dass

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = 0.$$

3. Fall: $\lim_{x \rightarrow x_0} g(x) = +\infty$ und $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} = c \in \mathbb{R}$ beliebig.

Wir betrachten die Funktion $f_1(x) := f(x) - c \cdot g(x)$. Dann ist

$$\lim_{x \rightarrow x_0} \frac{f_1'(x)}{g'(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} - c = 0.$$

Wir wenden den 2. Fall an und erhalten

$$\lim_{x \rightarrow x_0} \frac{f_1(x)}{g(x)} = 0.$$

Daraus folgt

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = c + \lim_{x \rightarrow x_0} \frac{f_1(x)}{g(x)} = c.$$

4. Fall: $\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = +\infty$ und $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} = +\infty$.

Dann gilt $\lim_{x \rightarrow x_0} \frac{g'(x)}{f'(x)} = 0$ und aus dem 2. Fall folgt sofort $\lim_{x \rightarrow x_0} \frac{g(x)}{f(x)} = 0$. Nach Voraussetzung sind f und g in einer Umgebung von x_0 positiv. Folglich gilt $\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = +\infty$.

5. Fall: $\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = +\infty$ und $\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)} = -\infty$.

Dieser Fall kann nicht auftreten, da sonst f oder g in einer Umgebung von x_0 monoton fallend wären und somit nicht gegen $+\infty$ konvergieren würden. \square

Satz 5.16 (Die Regeln von L'Hospital - (2))

Seien $f, g : (a, \infty) \rightarrow \mathbb{R}$ differenzierbare Funktionen mit $g'(x) \neq 0$ für alle $x \in (a, \infty)$. Ist

$\lim_{x \rightarrow \infty} f(x) = \lim_{x \rightarrow \infty} g(x) \in \{0, \pm\infty\}$ und $\lim_{x \rightarrow \infty} \frac{f'(x)}{g'(x)} = c \in \mathbb{R} \cup \{\pm\infty\}$, so gilt

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = c.$$

Beweis. OBdA. sei $a > 0$. Wir betrachten die Funktionen $f_1(x) = f\left(\frac{1}{x}\right)$ und $g_1(x) = g\left(\frac{1}{x}\right)$ auf dem Intervall $(0, \frac{1}{a})$. Dann gilt

$$f_1'(x) = -\frac{1}{x^2} f'\left(\frac{1}{x}\right) \quad \text{und} \quad g_1'(x) = -\frac{1}{x^2} g'\left(\frac{1}{x}\right) \neq 0$$

und daher

$$\lim_{x \rightarrow 0^+} \frac{f_1'(x)}{g_1'(x)} = \lim_{x \rightarrow 0^+} \frac{f'\left(\frac{1}{x}\right)}{g'\left(\frac{1}{x}\right)} \stackrel{\text{Vor.}}{=} c.$$

Mit Satz 5.15 folgt dann

$$\lim_{x \rightarrow \infty} \frac{f(x)}{g(x)} = \lim_{x \rightarrow 0^+} \frac{f_1(x)}{g_1(x)} = c.$$

\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:

1. Für jede positive reelle Zahl α gilt:

$$\lim_{x \rightarrow \infty} \frac{\ln(x)}{x^\alpha} = \lim_{x \rightarrow \infty} \frac{1}{x \cdot \alpha x^{\alpha-1}} = \lim_{x \rightarrow \infty} \frac{1}{\alpha x^\alpha} = 0,$$

$$\lim_{x \rightarrow \infty} \frac{e^x}{x^\alpha} = \lim_{x \rightarrow \infty} \frac{e^x}{\alpha x^{\alpha-1}} = \dots = \lim_{x \rightarrow \infty} \frac{e^x}{\alpha(\alpha-1) \cdot \dots \cdot (\alpha - [\alpha] - 1) \cdot x^{\alpha - ([\alpha] + 1)}} = +\infty,$$

wobei $[\alpha]$ die größte ganze Zahl k mit $k \leq \alpha$ bezeichnet.

2. $\lim_{x \rightarrow \infty} x^{\frac{1}{x}} = \lim_{x \rightarrow \infty} e^{\ln(x) \cdot \frac{1}{x}} = e^{\lim_{x \rightarrow \infty} \frac{\ln(x)}{x}} = e^0 = 1.$
3. $\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x} = \lim_{x \rightarrow 0} \frac{\frac{1}{1+x}}{1} = 1.$
4. $\lim_{x \rightarrow 0} (1+x)^{\frac{1}{x}} = \lim_{x \rightarrow 0} e^{\frac{\ln(1+x)}{x}} = e^{\lim_{x \rightarrow 0} \frac{\ln(1+x)}{x}} = e^1 = e.$
5. $\lim_{x \rightarrow 0} \frac{e^x - 1}{x} = \lim_{x \rightarrow 0} \frac{e^x}{1} = 1.$
6. $\lim_{x \rightarrow 0} (\cos x)^{\frac{1}{x}} = \lim_{x \rightarrow 0} e^{\frac{1}{x} \ln(\cos x)} = e^{\lim_{x \rightarrow 0} \frac{1}{x} \ln(\cos x)} = e^{\lim_{x \rightarrow 0} \frac{1}{\cos x} \cdot (-\sin x)} = e^0 = 1.$
7. $\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \cos(x) = 1.$

5.3 Differentiation von Funktionenfolgen und -reihen

Eine *gleichmäßig konvergente* Folge *stetiger* Funktionen ($f_n : I \subset \mathbb{R} \rightarrow E$) konvergiert gegen eine *stetige* Grenzfunktion. Sind alle Funktionen f_n differenzierbar, so muß die Grenzfunktion i.a. nicht differenzierbar sein. Als Beispiel betrachten wir $f_n : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$f_n(x) := \begin{cases} \frac{n}{2}x^2 + \frac{1}{2n} & \text{für } |x| \leq \frac{1}{n}, \\ |x| & \text{sonst.} \end{cases}$$

Die Funktionen f_n sind differenzierbar, die Funktionenfolge (f_n) konvergiert gleichmäßig gegen die Funktion $f(x) = |x|$. Die Grenzfunktion ist also nicht differenzierbar (Übungsaufgabe). Der folgende Satz gibt ein Kriterium dafür an, dass eine Folge *differenzierbarer* Funktionen gegen eine *differenzierbare* Grenzfunktion konvergiert und der Grenzwert der Funktionenfolge mit den Ableitungen vertauschbar ist.

Satz 5.17 Sei $f_n : [a, b] \subset \mathbb{R} \rightarrow E$, $n \in \mathbb{N}$, eine Folge differenzierbarer Funktionen mit Werten in einem Banachraum E und gelte:

1. Die Folge ($f_n(x_0)$) konvergiert für mindestens einen Punkt $x_0 \in [a, b]$.
2. Die Folge der Ableitungen (f'_n) konvergiert gleichmäßig auf $[a, b]$.

Dann konvergiert die Funktionenfolge (f_n) auf $[a, b]$ gleichmäßig gegen eine differenzierbare Funktion $f : [a, b] \rightarrow E$ und für die Ableitung gilt

$$f'(x) = \lim_{n \rightarrow \infty} f'_n(x).$$

Beweis. (1) Wir zeigen als erstes, dass die Funktionenfolge (f_n) gleichmäßig gegen eine Grenzfunktion $f : [a, b] \rightarrow E$ konvergiert.

Sei $\varepsilon > 0$. Da $(f_n(x_0))$ konvergent ist, existiert ein $n_0 \in \mathbb{N}$, so dass für alle $m \geq n \geq n_0$

$$\|f_n(x_0) - f_m(x_0)\| < \frac{\varepsilon}{2}. \quad (\text{A})$$

Weil (f'_n) gleichmäßig konvergiert, gibt es ein $n_1 \in \mathbb{N}$, so dass für alle $t \in [a, b]$ und für alle $m \geq n \geq n_1$

$$\|f'_n(t) - f'_m(t)\| < \frac{\varepsilon}{2(b-a)}. \quad (\text{B})$$

Wir wenden den Mittelwertsatz für Vektorfunktionen auf $f_n - f_m : [a, b] \rightarrow E$ an: Für zwei Werte $x, t \in [a, b]$ existiert ein $\xi_{m,n}$ zwischen x und t mit

$$\|(f_n(x) - f_m(x)) - (f_n(t) - f_m(t))\| \leq |x - t| \cdot \|f'_n(\xi_{m,n}) - f'_m(\xi_{m,n})\|.$$

Mit (B) folgt dann

$$\|(f_n(x) - f_m(x)) - (f_n(t) - f_m(t))\| < |x - t| \cdot \frac{\varepsilon}{2(b-a)} \leq \frac{\varepsilon}{2} \quad (\text{C})$$

für alle $m \geq n \geq n_1$ und für alle $x, t \in [a, b]$. Daher ist für $t = x_0$ mit (C) und (A)

$$\begin{aligned} \|f_n(x) - f_m(x)\| &\leq \|(f_n(x) - f_m(x)) - (f_n(x_0) - f_m(x_0))\| + \|f_n(x_0) - f_m(x_0)\| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \end{aligned} \quad (\text{D})$$

für alle $m \geq n \geq n^* := \max\{n_0, n_1\}$ und für alle $x \in [a, b]$. (f_n) ist also eine "gleichmäßige" Cauchy-Folge. Insbesondere ist $(f_n(x))$ für jedes $x \in [a, b]$ eine Cauchyfolge in E . Da E vollständig ist, konvergiert diese Folge. Wir definieren $f : [a, b] \rightarrow \mathbb{R}$

$$f(x) := \lim_{n \rightarrow \infty} f_n(x), \quad x \in [a, b].$$

Gehen wir in (D) mit m gegen ∞ , so folgt wegen der Stetigkeit der Norm

$$\|f_n(x) - f(x)\| \leq \varepsilon \quad \text{für alle } x \in [a, b] \text{ und } n \geq n^*.$$

Daher konvergiert (f_n) auf $[a, b]$ gleichmäßig gegen f .

(2) Wir zeigen nun, dass die Grenzfunktion f differenzierbar ist und ihre Ableitung durch $f'(x) = \lim_{n \rightarrow \infty} f'_n(x)$ gegeben ist.

Sei $x \in [a, b]$ fixiert. Wir definieren $\varphi_n : [a, b] \rightarrow E$ durch

$$\varphi_n(t) := \begin{cases} \frac{f_n(t) - f_n(x)}{t - x} & \text{für } x \neq t, \\ f'_n(x) & \text{für } x = t. \end{cases}$$

Da die Funktionen f_n stetig sind und der Grenzwert $\lim_{t \rightarrow x} \frac{f_n(t) - f_n(x)}{t - x} = f'_n(x)$ existiert, sind die Funktionen $\varphi_n : [a, b] \rightarrow E$ stetig. Für $t \neq x$ gilt mit (C)

$$\begin{aligned} \|\varphi_n(t) - \varphi_m(t)\| &= \frac{1}{|t - x|} \|(f_n(t) - f_m(t)) - (f_n(x) - f_m(x))\| \\ &\leq \frac{\varepsilon}{2(b-a)} \quad \forall n, m \geq n_1. \end{aligned}$$

Da $\varphi_n - \varphi_m$ stetig ist, gilt diese Abschätzung auch für $t = x$. (φ_n) ist also eine "gleichmäßige" Cauchy-Folge. Mit den gleichen Argumenten wie im Schritt (1) erhalten wir, dass die Grenzwerte $\varphi(t) := \lim_{n \rightarrow \infty} \varphi_n(t)$ für jedes $t \in [a, b]$ existieren und dass die Folge (φ_n) auf $[a, b]$ gleichmäßig auf gegen φ konvergiert. Da die Funktionen φ_n stetig sind, ist auch φ stetig und wir erhalten

$$\begin{aligned} \varphi(x) &= \lim_{t \rightarrow x} \varphi(t) = \lim_{t \rightarrow x} \lim_{n \rightarrow \infty} \varphi_n(t) \\ &= \lim_{t \rightarrow x} \left(\lim_{n \rightarrow \infty} \frac{f_n(t) - f_n(x)}{t - x} \right) = \lim_{t \rightarrow x} \frac{f(t) - f(x)}{t - x}. \end{aligned}$$

Folglich ist f in $x \in [a, b]$ differenzierbar und es gilt

$$f'(x) = \varphi'(x) = \lim_{n \rightarrow \infty} \varphi_n'(x) = \lim_{n \rightarrow \infty} f_n'(x).$$

□

Wir wenden den letzten Satz auf Funktionenreihen an und erhalten

Satz 5.18 Sei $f_n : [a, b] \rightarrow E$, $n \in \mathbb{N}$, eine Folge differenzierbarer Funktionen mit Werten in einem Banachraum E und gelte

1. $\sum_{n=1}^{\infty} f_n(x_0)$ konvergiert für mindestens ein $x_0 \in [a, b]$.
2. $\sum_{n=1}^{\infty} f_n'$ konvergiert gleichmäßig auf $[a, b]$.

Dann konvergiert die Funktionenreihe $\sum_{n=1}^{\infty} f_n$ auf $[a, b]$ gleichmäßig gegen eine differenzierbare Grenzfunktion $f : [a, b] \rightarrow E$,

$$f(x) := \sum_{n=1}^{\infty} f_n(x),$$

und für die Ableitung gilt

$$f'(x) = \sum_{n=1}^{\infty} f_n'(x).$$

5.4 Potenzreihen mit reellem Zentrum

Wir wenden nun die im letzten Abschnitt bewiesenen Differenzierbarkeitskriterien von Funktionenreihen auf eine spezielle Sorte von Potenzreihen an. Im Folgenden bezeichne \mathbb{K} wieder die reellen oder die komplexen Zahlen mit der Betragsfunktion als Norm. Für reelle Zahlen $x \in \mathbb{R}$ betrachten wir eine Potenzreihe

$$P(x) := \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

mit reellem Zentrum $x_0 \in \mathbb{R}$, Koeffizienten $a_n \in \mathbb{K}$ und Konvergenzradius $\rho > 0$. Aus den Kapiteln 3.5 und 4.5. wissen wir bereits das Folgende:

1. $P(x)$ ist für alle $x \in (x_0 - \rho, x_0 + \rho)$ absolut konvergent.
2. $P(x)$ divergiert für alle $x \in \mathbb{R}$ mit $|x - x_0| > \rho$.
3. P konvergiert gleichmäßig auf jedem kompakten Teilintervall von $(x_0 - \rho, x_0 + \rho)$.
4. Die Potenzreihe P definiert eine stetige Funktion $P : (x_0 - \rho, x_0 + \rho) \rightarrow \mathbb{K}$:

$$P(x) := \sum_{n=0}^{\infty} a_n(x - x_0)^n.$$

Das Intervall $(x_0 - \rho, x_0 + \rho)$ nennt man auch das *Konvergenzintervall* von P . Wir untersuchen zunächst das Verhalten von P in den Randpunkten des Konvergenzintervalls.

Satz 5.19 (Abelscher Grenzwertsatz)

Sei $P(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$ eine Potenzreihe mit reellem Zentrum $x_0 \in \mathbb{R}$ und Konvergenzradius $\rho > 0$.

1. Ist $P(x_0 + \rho)$ konvergent, so ist $P|_{(x_0 - \rho, x_0 + \rho]}$ in $x_0 + \rho$ linksseitig stetig, d.h. es gilt

$$\lim_{x \rightarrow (x_0 + \rho)^-} P(x) = P(x_0 + \rho) = \sum_{n=0}^{\infty} a_n \rho^n.$$

2. Ist $P(x_0 - \rho)$ konvergent, so ist $P|_{[x_0 - \rho, x_0 + \rho)}$ in $x_0 - \rho$ rechtsseitig stetig, d.h. es gilt

$$\lim_{x \rightarrow (x_0 - \rho)^+} P(x) = P(x_0 - \rho) = \sum_{n=0}^{\infty} (-1)^n a_n \rho^n.$$

Beweis. Wir beweisen die erste Behauptung. Der Beweis der zweiten Behauptung erfolgt analog. Nach Voraussetzung konvergiert die Folge der Partialsummen $s_m = \sum_{n=0}^m a_n \rho^n$ gegen $s = P(x_0 + \rho)$. Wir müssen zeigen, dass für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass $|P(x) - s| < \varepsilon$ für alle $x \in (x_0 - \rho, x_0 + \rho)$ mit $x_0 + \rho - x < \delta$.

Wir setzen $y := x - x_0$. Aus den Eigenschaften der geometrischen Reihe und dem Cauchy-Produkt von Reihen folgt für alle y mit $|y| < \rho$:

$$\begin{aligned} P(y + x_0) &= \sum_{n=0}^{\infty} a_n y^n \\ &= \left(1 - \frac{y}{\rho}\right) \left(1 - \frac{y}{\rho}\right)^{-1} \cdot \sum_{n=0}^{\infty} a_n y^n \\ &= \left(1 - \frac{y}{\rho}\right) \cdot \underbrace{\sum_{n=0}^{\infty} \left(\frac{y}{\rho}\right)^n}_{\text{abs. konv.}} \cdot \underbrace{\sum_{n=0}^{\infty} a_n y^n}_{\text{abs. konv.}} \\ &= \left(1 - \frac{y}{\rho}\right) \sum_{m=0}^{\infty} \left\{ \sum_{k=0}^m \left(\frac{y}{\rho}\right)^k a_{m-k} y^{m-k} \right\} \\ &= \left(1 - \frac{y}{\rho}\right) \sum_{m=0}^{\infty} \left\{ \left(\frac{y}{\rho}\right)^m \cdot s_m \right\}. \end{aligned}$$

Außerdem gilt

$$s = \left(1 - \frac{y}{\rho}\right) \left(1 - \frac{y}{\rho}\right)^{-1} \cdot s = \left(1 - \frac{y}{\rho}\right) \sum_{m=0}^{\infty} s \cdot \left(\frac{y}{\rho}\right)^m.$$

Folglich ist für $|y| < \rho$

$$P(y + x_0) - s = \left(1 - \frac{y}{\rho}\right) \sum_{m=0}^{\infty} (s_m - s) \left(\frac{y}{\rho}\right)^m.$$

Sei nun $\varepsilon > 0$ gegeben. Da (s_m) gegen s konvergiert, existiert ein $n_0 \in \mathbb{N}$ mit

$$|s_m - s| < \frac{\varepsilon}{2} \quad \forall m \geq n_0.$$

Wir setzen $M := \sum_{m=0}^{n_0-1} |s_m - s|$ und $\delta := \min\{\rho, \frac{\varepsilon \rho}{2M}\}$. Für $|y| < \rho$ mit $\rho - y < \delta$ gilt

$$0 < 1 - \frac{y}{\rho} < \frac{\delta}{\rho}.$$

Wegen $\delta \leq \rho$ gilt außerdem $y > 0$. Daraus folgt

$$\begin{aligned} |P(y + x_0) - s| &\leq \left(1 - \frac{y}{\rho}\right) \cdot \left(\sum_{m=0}^{n_0-1} |s_m - s| \cdot \left(\frac{y}{\rho}\right)^m + \sum_{m=n_0}^{\infty} |s_m - s| \cdot \left(\frac{y}{\rho}\right)^m\right) \\ &< \frac{\delta}{\rho} \cdot \sum_{m=0}^{n_0-1} |s_m - s| + \left(1 - \frac{y}{\rho}\right) \cdot \frac{\varepsilon}{2} \sum_{m=n_0}^{\infty} \left(\frac{y}{\rho}\right)^m \\ &\leq \frac{\delta}{\rho} \cdot M + \frac{\varepsilon}{2} \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Somit ist die Funktion P im Punkt $x_0 + \rho$ linksseitig stetig. \square

Als nächstes zeigen wir, dass die Grenzfunktion einer Potenzreihe mit reellem Zentrum auf dem Konvergenzintervall auch differenzierbar ist.

Satz 5.20 Sei $P(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n$ eine Potenzreihe mit reellem Zentrum $x_0 \in \mathbb{R}$ und positivem Konvergenzradius $\rho > 0$. Dann ist $P : (x_0 - \rho, x_0 + \rho) \subset \mathbb{R} \rightarrow \mathbb{K}$ differenzierbar und es gilt

$$P'(x) = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1}.$$

Beweis. Die Funktionen $f_n(x) := a_n(x - x_0)^n$ sind differenzierbar und es gilt

$$f'_0(x) = 0 \quad \text{und} \quad f'_n(x) = n a_n (x - x_0)^{n-1} \quad \text{für } n > 0.$$

Offensichtlich konvergiert die Potenzreihe $P(x_0)$. Wir betrachten nun die Reihe der Ableitungen

$$Q(x) := \sum_{n=0}^{\infty} f'_n(x) = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1}.$$

Da

$$\limsup_{n \rightarrow \infty} \sqrt[n]{n|a_n|} = \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|},$$

stimmen die Konvergenzradien von P und Q überein. Q konvergiert also gleichmäßig auf jedem kompakten Teilintervall des Konvergenzintervalls $(x_0 - \rho, x_0 + \rho)$. Nach Satz 5.18 ist P deshalb für jedes $0 < r < \rho$ auf dem Intervall $I_r := [x_0 - r, x_0 + r]$ differenzierbar und es gilt

$$P'(x) = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1} \quad \text{für alle } x \in I_r.$$

Mit $r \rightarrow \rho$ folgt die Behauptung. □

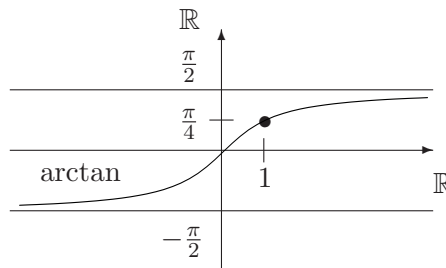
Anwendung: Berechnung von π .

Als Anwendung berechnen wir die Zahl π . Dazu betrachten wir die Funktion

$$\arctan : \mathbb{R} \longrightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Aus dem Additionstheorem für \cos erhält man $\tan\left(\frac{\pi}{4}\right) = 1$, also $\arctan(1) = \frac{\pi}{4}$. Wir zeigen nun

$$\arctan x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} \quad \text{für alle } |x| \leq 1.$$



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 5.19 ist damit Q in $x = \pm 1$ (einseitig) 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

$$\arctan x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} \quad \text{für alle } |x| \leq 1.$$

□

Da $\frac{\pi}{4} = \arctan(1)$, gilt für π die *Leibnizformel*

$$\boxed{\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \pm \dots = \sum_{k=0}^{\infty} (-1)^k \frac{1}{2k+1}.}$$

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 3.9 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 :

$$\pi = 3.1415926535 + Rest, \quad |Rest| < 10^{-11}.$$

5.5 Reell-analytische Funktionen und Taylorreihen

Wir kennen aus Kapitel 4 bereits einige Funktionen, die durch Potenzreihen definiert sind. In diesem Abschnitt wollen wir untersuchen, unter welchen Bedingungen man eine Funktion $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ als Potenzreihe darstellen kann.

Bevor wir beginnen, vereinbaren wir zunächst einige nützliche, in der Literatur gebräuchliche Abkürzungen.

Die Landau-Symbole o und O :

Seien $h, g : I \subset \mathbb{R} \rightarrow \mathbb{R}$ reellwertige Funktionen und $x_0 \in I$.

- $h(x) = o(g(x))$ für $x \rightarrow x_0$, ist die Abkürzung für die Eigenschaft

$$\lim_{x \rightarrow x_0} \frac{h(x)}{g(x)} = 0.$$

- $h(x) = O(g(x))$ für $x \rightarrow x_0$, ist die Abkürzung dafür, dass es eine Konstante $C > 0$ und ein Intervall $(x_0 - \varepsilon, x_0 + \varepsilon)$ gibt mit

$$|h(x)| \leq C |g(x)| \quad \text{für alle } x \in (x_0 - \varepsilon, x_0 + \varepsilon) \cap I.$$

- $h_1(x) = h_2(x) + o(g(x))$ für $x \rightarrow x_0$, ist die Abkürzung für $h_1(x) - h_2(x) = o(g(x))$ für $x \rightarrow x_0$. Analog für $O(g(x))$.
- Insbesondere heißt $h(x) = o((x - x_0)^n)$ für $x \rightarrow x_0$, dass

$$\lim_{x \rightarrow x_0} \frac{h(x)}{(x - x_0)^n} = 0.$$

$h(x)$ geht für $x \rightarrow x_0$ also schneller gegen 0 als $(x - x_0)^n$. Man sagt auch h *verschwindet in x_0 von höherer als n -ter Ordnung*. Für $h_1(x) = h_2(x) + o((x - x_0)^n)$ für $x \rightarrow x_0$, sagt man auch h_2 *approximiert h_1 in x_0 von höherer als n -ter Ordnung*.

- Zur Eigenschaft $h(x) = O((x - x_0)^n)$ für $x \rightarrow x_0$, d. h.

$$|h(x)| \leq C |x - x_0|^n \quad \text{für alle } x \in (x_0 - \varepsilon, x_0 + \varepsilon) \cap I,$$

sagt man auch h wächst in x_0 von höchstens n -ter Ordnung.

Definition 5.8. Eine Funktion $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ heißt reell-analytisch in $x_0 \in I$, falls ein Intervall $(x_0 - r, x_0 + r) \subset I$ und eine Potenzreihe $\sum_{k=0}^{\infty} a_k(x - x_0)^k$ mit Zentrum $x_0 \in \mathbb{R}$ und Konvergenzradius $\rho > 0$ existieren, so dass

$$f(x) = \sum_{k=0}^{\infty} a_k(x - x_0)^k \quad \text{für alle } x \in (x_0 - \min\{r, \rho\}, x_0 + \min\{r, \rho\}).$$

Man sagt in diesem Fall auch, dass f in einer Umgebung von x_0 in eine Potenzreihe entwickelbar ist. Ist $U \subset \mathbb{R}$ ein offenes Intervall, so heißt $f : U \rightarrow \mathbb{R}$ reell-analytisch, wenn f in jedem Punkt von U reell-analytisch ist.

$C^\omega(U, \mathbb{R})$ bezeichne den Vektorraum der reell-analytischen Funktionen auf $U \subset \mathbb{R}$.

Beispiele für reell-analytische Funktionen:

1. Jede Potenzreihe $f(x) := \sum_{k=0}^{\infty} a_k(x - x_0)^k$ mit reellem Zentrum, reellen Koeffizienten a_k und positivem Konvergenzradius $\rho > 0$ 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} .

Satz 5.21 Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ reell-analytisch in $x_0 \in I$. Dann ist f in einer Umgebung U von x_0 unendlich oft differenzierbar und es gilt:

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} \cdot (x - x_0)^k \quad \forall x \in U.$$

Beweis. Nach Voraussetzung existiert eine Potenzreihe $Q(x) := \sum_{k=0}^{\infty} a_k(x - x_0)^k$ mit positivem Konvergenzradius $\rho > 0$ und ein $r > 0$, so dass $f(x) = Q(x)$ für alle $x \in U := (x_0 - \min\{r, \rho\}, x_0 + \min\{r, \rho\})$. Nach Satz 5.20 ist die Funktion f auf U differenzierbar und für ihre Ableitung gilt

$$f'(x) = \sum_{k=1}^{\infty} k \cdot a_k(x - x_0)^{k-1} \quad \forall x \in U.$$

Diese Potenzreihe kann man wiederum auf U gliedweise ableiten ohne den Konvergenzbe-
reich zu verkleinern. Nach Satz 5.20 existiert wiederum $f''(x)$ mit

$$f''(x) = \sum_{k=2}^{\infty} k(k-1)a_k(x - x_0)^{k-2} \quad \forall x \in U.$$

Durch weiteres Anwenden des Satzes 5.20 folgt, dass f auf U beliebig oft differenzierbar ist mit der n -ten Ableitung

$$f^{(n)}(x) = \sum_{k=n}^{\infty} k(k-1) \cdot \dots \cdot (k-n+1) a_k (x-x_0)^{k-n} \quad \forall x \in U.$$

Daraus erhält man

$$f^{(n)}(x_0) = n! \cdot a_n$$

für jedes $n \in \mathbb{N}_0$. Dies zeigt die Behauptung. \square

Definition 5.9. Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ 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 .

Nach Satz 5.21 hat die Taylorreihe $T(f, x_0)(x)$ einer in x_0 reell-analytischen Funktion f 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 Beispiele zeigen. Es gibt also C^∞ -Funktionen, die nicht reell-analytisch sind:

Beispiel 1: Die Taylorreihe einer C^∞ -Funktion kann den Konvergenzradius $\rho = 0$ 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 $\rho = 0$ 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 unendlich oft differenzierbar und für die Ableitungen in $x_0 = 0$ gilt $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}_0$ (Übungsaufgabe). 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 $x > 0$.

Wir wollen nun Bedingungen finden, unter denen eine C^∞ -Funktion reell-analytisch ist. Dazu betrachten wir die Taylorpolynome von f , die als Partialsummen der Taylorreihe auftreten und untersuchen ihre Abweichung von f :

Definition 5.10. Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ in $x_0 \in I$ n -mal differenzierbar. Dann heißt

$$T_n(f, x_0)(x) := \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

das n -te Taylorpolynom von f in x_0 .

Das Taylorpolynom für $n = 1$ beschreibt die Tangente an den Funktionsgraphen von f im Punkt $(x_0, f(x_0))$:

$$T_1(f, x_0)(x) = f(x_0) + f'(x_0)(x - x_0).$$

Das Taylorpolynom für $n = 2$ beschreibt 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 Ableitungen von f und $T_n(f, x_0)$ stimmen in $x = x_0$ überein:

$$\frac{d^j}{dx^j} T_n(f, x_0)(x_0) = f^{(j)}(x_0) \quad \text{für } j = 0, \dots, n.$$

Satz 5.22 Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ n -mal differenzierbar und $x_0 \in I$. Dann gilt:

$$f(x) = T_n(f, x_0)(x) + o((x - x_0)^n) \quad \text{für } x \rightarrow x_0,$$

d.h. das n -te Taylorpolynom $T_n(f, x_0)$ approximiert f in x_0 von höherer als n -ter Ordnung.

Beweis. Wir beweisen die Behauptung durch Induktion über n :

Induktionsanfang: Sei $n = 1$. Für die Tangente $T_1(f, x_0)(x) = f(x_0) + f'(x_0)(x - x_0)$ an den Graphen der Funktion f im Punkt $(x_0, f(x_0))$ gilt

$$\frac{f(x) - T_1(f, x_0)(x)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0).$$

Da f in x_0 differenzierbar ist, folgt

$$\lim_{x \rightarrow x_0} \frac{f(x) - T_1(f, x_0)(x)}{x - x_0} = 0.$$

Dies zeigt die Behauptung für $n = 1$.

Induktionsschritt: Wir setzen voraus, dass die Behauptung für ein $n \in \mathbb{N}$ richtig ist und zeigen, dass sie dann auch für den Nachfolger $n + 1$ gilt.

Induktionsbeweis: Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ $(n + 1)$ -mal differenzierbar. Dann ist die Funktion $R_{n+1} := f - T_{n+1}(f, x_0)$ $(n + 1)$ -mal differenzierbar und es gilt

$$R'_{n+1} = f' - \frac{d}{dx} T_{n+1}(f, x_0) = f' - T_n(f', x_0).$$

Wir wenden nun die Induktionsvoraussetzung auf die n -mal differenzierbare Funktion f' an und erhalten $f'(x) = T_n(f', x_0)(x) + o((x - x_0)^n)$. Folglich existiert für alle $\varepsilon > 0$ ein $\delta > 0$, so dass

$$\frac{|R'_{n+1}(x)|}{|x - x_0|^n} = \frac{|f'(x) - T_n(f', x_0)(x)|}{|x - x_0|^n} < \varepsilon \quad \forall x \in I \text{ mit } 0 < |x - x_0| < \delta. \quad (\star)$$

Sei $x \in I$ mit $0 < |x - x_0| < \delta$. Die Funktion R_{n+1} ist zwischen x_0 und x differenzierbar. Nach dem Mittelwertsatz existiert ein ξ zwischen x und x_0 , so dass

$$|R_{n+1}(x) - \underbrace{R_{n+1}(x_0)}_{=0}| \leq |R'_{n+1}(\xi)| |x - x_0| \stackrel{(\star)}{<} \varepsilon \cdot |\xi - x_0|^n |x - x_0| < \varepsilon |x - x_0|^{n+1}.$$

Wir erhalten $\left| \frac{R_{n+1}(x)}{(x - x_0)^{n+1}} \right| < \varepsilon$ für alle $x \in I$ mit $0 < |x - x_0| < \delta$. Daraus folgt

$$\lim_{x \rightarrow x_0} \frac{R_{n+1}(x)}{(x - x_0)^{n+1}} = \lim_{x \rightarrow x_0} \frac{f(x) - T_{n+1}(f, x_0)(x)}{(x - x_0)^{n+1}} = 0,$$

also $f(x) = T_{n+1}(f, x_0)(x) + o((x - x_0)^{n+1})$ für $x \rightarrow x_0$. □

Definition 5.11. Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ in $x_0 \in I$ n -mal differenzierbar. Dann heißt die Differenz $R_n(f, x_0) := f - T_n(f, x_0)$ das n -te Restglied von f in x_0 .

Ist f unendlich oft differenzierbar, so stimmt $f(x)$ genau dann mit der Taylorreihe $T(f, x_0)(x)$ überein, wenn

$$\lim_{n \rightarrow \infty} R_n(f, x_0)(x) = 0.$$

Wir interessieren uns deshalb für explizite Formeln für das Restglied $R_n(f, x_0)(x)$, die es ermöglichen, dieses Kriterium zu überprüfen bzw. den Fehler bei der Approximation von $f(x)$ durch das n -te Taylorpolynom $T_n(f, x_0)(x)$ zu beschreiben.

Satz 5.23 Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ $(n + 1)$ -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 existieren $\vartheta, \theta \in (0, 1)$, so dass gilt

$$R_n(f, x_0)(x) = \frac{f^{(n+1)}(x_0 + \vartheta(x - x_0))}{(n + 1)!} (x - x_0)^{n+1} \quad \text{Lagrange-Form des Restgliedes}$$

$$R_n(f, x_0)(x) = \frac{f^{(n+1)}(x_0 + \theta(x - x_0))}{n!} (1 - \theta)^n (x - x_0)^{n+1} \quad \text{Cauchy-Form des Restgliedes}$$

Beweis. Sei $x \in I$ ein fixierter Punkt mit $x \neq x_0$. Wir betrachten 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

$$R_n(f, x_0)(x) = \frac{(1 - \theta)^n (x - x_0)^{n+1}}{n!} f^{(n+1)}(x_0 + \theta(x - x_0)).$$

Dies ist die Cauchy-Form des Restgliedes.

Wir betrachten nun 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 $\vartheta \in (0, 1)$, so dass

$$(g(x) - g(x_0)) \cdot h'(x_0 + \vartheta(x - x_0)) = (h(x) - h(x_0)) \cdot g'(x_0 + \vartheta(x - x_0)).$$

Für das Restglied folgt

$$\begin{aligned} R_n(f, x_0)(x) \cdot (n + 1)(x - x_0)^n (1 - \vartheta)^n \\ = (x - x_0)^{n+1} \frac{(x - x_0)^n (1 - \vartheta)^n}{n!} f^{(n+1)}(x_0 + \vartheta(x - x_0)) \end{aligned}$$

und somit

$$R_n(f, x_0)(x) = \frac{(x - x_0)^{n+1}}{(n + 1)!} f^{(n+1)}(x_0 + \vartheta(x - x_0)).$$

Dies ist die Lagrange-Form des Restgliedes. □

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:

$$\ln(1 + x) = \sum_{n=1}^{\infty} (-1)^{n+1} \cdot \frac{x^n}{n} \quad \text{für alle } x \in (-1, 1].$$

Insbesondere gilt für die alternierende harmonische Reihe

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \ln(2).$$

Beweis. Die Funktion $f(x) := \ln(x + 1)$ ist auf $(-1, \infty)$ beliebig oft differenzierbar und es gilt $f(0) = 0$ sowie

$$f^{(n)}(x) = \frac{(n - 1)!(-1)^{n+1}}{(1 + x)^n} \quad \text{für alle } n \in \mathbb{N}.$$

Für die Taylorreihe von $f(x) = \ln(x+1)$ in $x_0 = 0$ folgt

$$T(f, 0)(x) = \sum_{n=1}^{\infty} (-1)^{n+1} \cdot \frac{x^n}{n}.$$

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 $f(x)$:

$$\ln(1+x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} x^n \quad \text{für alle } |x| < 1.$$

Es bleibt die Konvergenz in $x = 1$ zu untersuchen. Nach dem Leibnizkriterium (Satz 3.9) 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 5.19) 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 $f(x) := (1+x)^\alpha$ 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:

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \quad \text{für alle } x \in (-1, 1).$$

Die Reihe $B_\alpha(x) := \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k$ heißt *Binomialreihe*.

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 36). Wir zeigen nun, dass ihr Grenzwert $f(x) = (1+x)^\alpha$ ist. (Für $\alpha \in \mathbb{N}$ ist die Reihe endlich und wir erhalten die Binomische Formel). 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 $(|\binom{\alpha-1}{n} x^n|)$ 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 $f(x)$. \square

5.6 Lokale Extrema für Funktionen einer reellen Variablen

In Satz 5.10 hatten wir bereits ein Kriterium für die Existenz eines lokalen Extremwertes einer 2-mal differenzierbaren Funktion kennengelernt. Mit Hilfe der Ergebnisse des letzten Abschnittes können wir dieses Kriterium nun verallgemeinern.

Satz 5.24 (Hinreichende Bedingung für einen lokalen Extremwert)

Sei $f : (a, b) \subset \mathbb{R} \rightarrow \mathbb{R}$ 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:

$$\begin{aligned} f(x) &= T_{n-2}(f, x_0)(x) + R_{n-2}(f, x_0)(x) \\ &\stackrel{5.23}{=} f(x_0) + \frac{1}{(n-1)!} f^{(n-1)}(\overbrace{x_0 + \theta(x - x_0)}^{\xi})(x - x_0)^{n-1} \quad \text{für ein } \theta \in (0, 1). \end{aligned}$$

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 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 $f^{(n)}(x_0) < 0$. □

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 5.24 nicht!

Differentialrechnung für Funktionen mehrerer reeller Variablen

Nachdem wir in Kapitel 5 die Differentialrechnung für Funktionen *einer* reeller Variablen behandelt haben, wollen wir uns jetzt mit Funktionen befassen, die von *mehreren* reellen Variablen abhängen. Auch für solche Funktionen werden wir einen Differenzierbarkeitsbegriff einführen und untersuchen, welche Rückschlüsse die Eigenschaften der Ableitungen auf das Verhalten der Funktion selbst zulassen.

In dem gesamten Kapitel studieren wir Funktionen $f : U \subset \mathbb{R}^n \rightarrow E$, die von einer *offenen* Teilmenge $U \subset \mathbb{R}^n$ in einen reellen normierten Vektorraum $(E, \|\cdot\|_E)$ abbilden. Den \mathbb{R}^n betrachten wir ebenfalls als normierten Vektorraum mit der Euklidischen Norm $\|\cdot\|$. Wie wir aus Kapitel 4 wissen, ist es für Konvergenzfragen allerdings egal, welche Norm auf dem \mathbb{R}^n wir benutzen, denn sie sind alle zueinander äquivalent.

6.1 Ableitung, Richtungsableitung und Gradient einer Funktion

Zunächst erinnern wir uns nochmals daran, wie wir die Differenzierbarkeit für Funktionen *einer* Variablen definiert hatten: Eine Funktion $f : U \subset \mathbb{R} \rightarrow E$ heißt in $x_0 \in U$ differenzierbar, wenn der Grenzwert

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} =: f'(x_0) \in E$$

existiert. Dies lässt sich auch schreiben als

$$f(x_0 + h) - f(x_0) = L(h) + o(h) \quad \text{für } h \rightarrow 0,$$

wobei $L : \mathbb{R} \rightarrow E$ die durch $L(h) := h \cdot f'(x_0)$ definierte lineare Abbildung bezeichnet. Das motiviert die folgende Definition für Funktionen *mehrerer* reeller Variablen:

Definition 6.1. Eine Funktion $f : U \subset \mathbb{R}^n \rightarrow E$ heißt in $x_0 \in U$ differenzierbar, falls eine lineare Abbildung $L : \mathbb{R}^n \rightarrow E$ existiert, so dass

$$f(x_0 + h) - f(x_0) = L(h) + o(\|h\|) \quad \text{für } h \rightarrow 0. \quad (\star)$$

Die lineare Abbildung $L : \mathbb{R}^n \rightarrow E$ heißt *Ableitung von f in x_0* (auch: *Differential von f in x_0* oder *Linearisierung von f in x_0*). Wir bezeichnen sie mit $L =: Df(x_0)$.

$f : U \subset \mathbb{R}^n \rightarrow E$ heißt *differenzierbar*, falls f in jedem $x_0 \in U$ differenzierbar ist.

Die Bedingung (\star) ist zu den folgenden beiden Forderungen äquivalent:

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - L(h)}{\|h\|} = 0 \quad \text{in } E \quad \text{bzw.}$$

$$\lim_{h \rightarrow 0} \frac{\|f(x_0 + h) - f(x_0) - L(h)\|_E}{\|h\|} = 0 \quad \text{in } \mathbb{R}.$$

Für $n = 1$ haben wir jetzt zwei verschiedene Ableitungsbegriffe definiert, die folgendermaßen zusammenhängen: Ist $f : U \subset \mathbb{R} \rightarrow E$ in $x_0 \in U$ differenzierbar, so gilt $f'(x_0) = Df(x_0)(1)$.

Wir werden in diesem Kapitel häufig Kenntnisse über lineare Abbildungen aus der Vorlesung *Lineare Algebra* verwenden. Zunächst vermerken wir die folgende Stetigkeitseigenschaft linearer Abbildungen:

Lemma 6.2. *Sei $L : \mathbb{R}^n \rightarrow E$ eine lineare Abbildung. Dann existiert eine positive Konstante $C_L \in \mathbb{R}$, so dass*

$$\|L(x)\|_E \leq C_L \|x\| \quad \text{für alle } x \in \mathbb{R}^n.$$

Insbesondere ist L stetig.

Beweis. Wir betrachten die kanonische Basis (e_1, \dots, e_n) im Vektorraum \mathbb{R}^n . Dann gilt für alle $x = (x_1, \dots, x_n) \in \mathbb{R}^n$:

$$L(x) = L\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i L(e_i) \quad \text{und somit}$$

$$\|L(x)\|_E \leq \sum_{i=1}^n |x_i| \cdot \|L(e_i)\|_E \leq \max\{\|L(e_i)\|_E \mid i = 1, \dots, n\} \cdot \sum_{i=1}^n |x_i| \leq C_L \|x\|,$$

wobei C_L eine geeignete Konstante ist, die man aus der Äquivalenz der Normen im \mathbb{R}^n erhält. Sei nun (y_k) eine Folge im \mathbb{R}^n , die gegen x_0 konvergiert. Dann gilt

$$\|L(y_k) - L(x_0)\|_E \leq C_L \|y_k - x_0\| \xrightarrow{k \rightarrow \infty} 0.$$

Folglich ist L folgenstetig und somit auch stetig. □

Satz 6.1 *Sei $f : U \subset \mathbb{R}^n \rightarrow E$ in $x_0 \in U$ differenzierbar. Dann gilt:*

1. *Das Differential $Df(x_0)$ ist eindeutig bestimmt.*
2. *f ist in x_0 stetig.*

Beweis. Zu 1) Seien $L, \tilde{L} : \mathbb{R}^n \rightarrow E$ zwei lineare Abbildungen mit

$$\begin{aligned} f(x_0 + h) - f(x_0) &= L(h) + o(\|h\|) \\ &= \tilde{L}(h) + o(\|h\|) \quad \text{für } h \rightarrow 0. \end{aligned}$$

Dann folgt

$$\lim_{h \rightarrow 0} \frac{L(h) - \tilde{L}(h)}{\|h\|} = 0.$$

Da beide Abbildungen linear sind, gilt $L(0) = \tilde{L}(0)$. Sei nun $x \in \mathbb{R}^n \setminus \{0\}$ und $t \in \mathbb{R}$. Betrachten wir die Folgen $h := tx$ für $t \rightarrow 0^+$, so erhalten wir wegen der Linearität von L und \tilde{L}

$$0 = \lim_{t \rightarrow 0^+} \frac{L(tx) - \tilde{L}(tx)}{\|tx\|} = \lim_{t \rightarrow 0^+} \frac{|t| L(x) - \tilde{L}(x)}{|t| \|x\|} = \frac{L(x) - \tilde{L}(x)}{\|x\|}.$$

Also ist $L(x) = \tilde{L}(x)$ für alle $x \in \mathbb{R}^n$.

Zu 2) Sei $f : U \subset \mathbb{R}^n \rightarrow E$ in $x_0 \in U$ differenzierbar und (x_k) eine gegen x_0 konvergente Folge in U . Dann gilt nach Dreiecksungleichung

$$\|f(x_k) - f(x_0)\|_E \leq \underbrace{\frac{\|f(x_k) - f(x_0) - L(x_k - x_0)\|_E}{\|x_k - x_0\|}}_{\rightarrow 0} \cdot \underbrace{\|x_k - x_0\|}_{\rightarrow 0} + \underbrace{\|L(x_k - x_0)\|_E}_{\rightarrow 0}.$$

Die ersten beiden Konvergenzen folgen aus der Definition der Differenzierbarkeit und der Stetigkeit der Norm. Für die dritte Konvergenz benutzen wir die Stetigkeit der linearen Abbildung L . Folglich konvergiert die Folge $(f(x_k))$ gegen $f(x_0)$. Somit ist f in x_0 folgenstetig, also auch stetig. \square

Definition 6.3. Sei $f : U \subset \mathbb{R}^n \rightarrow E$ eine Abbildung und $\mathbf{a} \in \mathbb{R}^n$ ein fixierter Vektor. Man sagt: f besitzt in $x_0 \in U$ eine Ableitung in Richtung $\mathbf{a} \in \mathbb{R}^n$, falls der Grenzwert

$$\nabla_{\mathbf{a}} f(x_0) := \lim_{t \rightarrow 0} \frac{f(x_0 + t\mathbf{a}) - f(x_0)}{t} \in E \quad (\text{Hier ist } t \in \mathbb{R})$$

existiert. Dieser Grenzwert heißt die Richtungsableitung von f in Richtung \mathbf{a} an der Stelle x_0 .

Bei der Richtungsableitung wird f also nur entlang des Geradenstückes

$$\{x_0 + t\mathbf{a} \mid t \in (-\varepsilon, \varepsilon)\} \subset U$$

betrachtet, das heißt, es gilt $\nabla_{\mathbf{a}} f(x_0) = h'(0)$, wobei h die Funktion $h(t) := f(x_0 + t\mathbf{a})$ ist.

Satz 6.2 Ist $f : U \subset \mathbb{R}^n \rightarrow E$ in x_0 differenzierbar, so besitzt f in x_0 in jeder Richtung \mathbf{a} die Richtungsableitung und es gilt

$$Df(x_0)(\mathbf{a}) = \nabla_{\mathbf{a}} f(x_0).$$

Beweis. Sei f in x_0 differenzierbar, $L := Df(x_0) : \mathbb{R}^n \rightarrow E$ und $\mathbf{a} \in \mathbb{R}^n$. Für $\mathbf{a} = 0$ folgt die Behauptung aus der Linearität von $Df(x_0)$. Für $\mathbf{a} \neq 0$ betrachten wir $h = t\mathbf{a}$. Dann gilt nach Definition der Differenzierbarkeit

$$\lim_{t \rightarrow 0} \frac{\|f(x_0 + t\mathbf{a}) - f(x_0) - L(t\mathbf{a})\|_E}{|t| \cdot \|\mathbf{a}\|} = 0.$$

Da L linear ist, folgt

$$\lim_{t \rightarrow 0} \left\| \frac{f(x_0 + t\mathbf{a}) - f(x_0)}{t} - L(\mathbf{a}) \right\|_E = 0.$$

Damit existiert

$$\nabla_{\mathbf{a}} f(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + t\mathbf{a}) - f(x_0)}{t} = L(\mathbf{a}) = Df(x_0)(\mathbf{a}). \quad \square$$

Beispiel 1: Eine Funktion kann in einem Punkt x_0 alle Richtungsableitungen besitzen, ohne in diesem Punkt differenzierbar zu sein.

Wir betrachten dazu die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}^1$, definiert durch

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2+y^4} & \text{falls } (x, y) \neq (0, 0), \\ 0 & \text{falls } (x, y) = (0, 0). \end{cases}$$

Aus Kapitel 4.1 (Beispiel 2) wissen wir, dass f in $x_0 := (0, 0)$ nicht stetig, also auch nicht differenzierbar ist. 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:

$$\frac{f(x_0 + t\mathbf{a}) - f(x_0)}{t} = \frac{f(t\mathbf{a})}{t} = \frac{t^3 a_1 a_2^2}{t(t^2 a_1^2 + t^4 a_2^4)} = \frac{a_1 a_2^2}{a_1^2 + t^2 a_2^4} \xrightarrow{t \rightarrow 0} \begin{cases} \frac{a_2^2}{a_1} & a_1 \neq 0, \\ 0 & a_1 = 0. \end{cases}$$

Somit existiert $\nabla_{\mathbf{a}}f(x_0)$ für alle $\mathbf{a} = (a_1, a_2) \in \mathbb{R}^2$.

Beispiel 2: *Das Differential einer linearen Abbildung.*

Eine lineare Abbildung $L : \mathbb{R}^n \rightarrow E$ ist in jedem Punkt $x_0 \in \mathbb{R}^n$ differenzierbar und es gilt $DL(x_0) = L$, da aufgrund der Linearität von L

$$\lim_{h \rightarrow 0} \frac{\|L(x_0 + h) - L(x_0) - L(h)\|_E}{\|h\|} = \lim_{h \rightarrow 0} \underbrace{\frac{\|0\|_E}{\|h\|}}_{=0} = 0.$$

Beispiel 3: *Das Differential einer multilinearen Abbildung.*

Sei $f : \mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_k} \rightarrow E$ eine multilineare Abbildung. Dann ist f in jedem Punkt x_0 differenzierbar und es gilt

$$Df(x_1, \dots, x_k)(\mathbf{a}_1, \dots, \mathbf{a}_k) = \sum_{j=1}^k f(x_1, \dots, x_{j-1}, \mathbf{a}_j, x_{j+1}, \dots, x_k).$$

Beweis. (1) Wir schätzen zunächst die Norm der Abbildung f ab. Jede multilineare Abbildung $f : \mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_k} \rightarrow E$ ist stetig. Insbesondere existiert ein $\delta > 0$, so dass für alle $\|x_1\|, \dots, \|x_k\| \leq \delta$ gilt

$$\|f(x_1, \dots, x_k)\|_E \leq 1. \tag{*}$$

Folglich ist für alle $x_j \in \mathbb{R}^{n_j}$ mit $x_j \neq 0$

$$\|f(x_1, \dots, x_k)\|_E = \frac{\|x_1\|}{\delta} \cdot \dots \cdot \frac{\|x_k\|}{\delta} \left\| f \left(\frac{\delta x_1}{\|x_1\|}, \dots, \frac{\delta x_k}{\|x_k\|} \right) \right\|_E \leq \frac{1}{\delta^k} \|x_1\| \cdot \dots \cdot \|x_k\|.$$

Somit existiert für jede multilineare Abbildung $f : \mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_k} \rightarrow E$ eine Konstante $C > 0$ mit

$$\|f(x_1, \dots, x_k)\|_E \leq C \cdot \|x_1\| \cdot \dots \cdot \|x_k\|.$$

(2) Wir zeigen nun die Differenzierbarkeit von f . Es gilt:

$$\begin{aligned}
 & f(x_1 + \mathbf{a}_1, \dots, x_k + \mathbf{a}_k) - f(x_1, \dots, x_k) - \underbrace{\sum_{j=1}^k f(x_1, \dots, x_{j-1}, \mathbf{a}_j, x_{j+1}, \dots, x_k)}_{=:L(\mathbf{a})} \\
 &= \sum_{j_1 < j_2} f(x_1, \dots, \mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_2}, \dots, x_k) \\
 & \quad + \sum_{j_1 < j_2 < j_3} f(x_1, \dots, \mathbf{a}_{j_1}, \dots, \mathbf{a}_{j_2}, \dots, \mathbf{a}_{j_3}, \dots, x_k) + \dots + f(\mathbf{a}_1, \dots, \mathbf{a}_k).
 \end{aligned}$$

Folglich gilt wegen (*)

$$\begin{aligned}
 & \|f(x_1 + \mathbf{a}_1, \dots, x_k + \mathbf{a}_k) - f(x_1, \dots, x_k) - L(\mathbf{a})\|_E \\
 & \leq C \left(\sum_{j_1 < j_2} \|x_1\| \cdot \dots \cdot \|\mathbf{a}_{j_1}\| \cdot \dots \cdot \|\mathbf{a}_{j_2}\| \cdot \dots \cdot \|x_k\| + \right. \\
 & \quad \left. + \sum_{j_1 < j_2 < j_3} (\dots) + \dots + \|\mathbf{a}_1\| \cdot \dots \cdot \|\mathbf{a}_k\| \right).
 \end{aligned}$$

Jeder Summand auf der rechten Seite enthält mindestens 2 Faktoren der Form $\|\mathbf{a}_j\|$. Da $\|a_{j_1}\| \cdot \|a_{j_2}\| \leq \frac{1}{2}(\|a_{j_1}\|^2 + \|a_{j_2}\|^2) \leq \frac{1}{2}\|\mathbf{a}\|^2$, kann man alle Summanden auf der rechten Seite durch $\|\mathbf{a}\|^2$ abschätzen. Daraus folgt

$$\lim_{\mathbf{a} \rightarrow 0} \frac{\|f(\mathbf{x} + \mathbf{a}) - f(\mathbf{x}) - L(\mathbf{a})\|_E}{\|\mathbf{a}\|} \leq \lim_{\mathbf{a} \rightarrow 0} C \cdot (\dots) \cdot \|\mathbf{a}\| = 0. \quad \square$$

Satz 6.3 (Rechenregeln für Ableitungen und Richtungsableitungen)

1. Seien $f, g : U \subset \mathbb{R}^n \rightarrow E$ und $h, p : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$.

a) Sind f und g in x_0 differenzierbar, so ist $f + g$ in x_0 differenzierbar und es gilt

$$D(f + g)(x_0) = Df(x_0) + Dg(x_0).$$

Existieren $\nabla_{\mathbf{a}}f(x_0)$ und $\nabla_{\mathbf{a}}g(x_0)$, so existiert $\nabla_{\mathbf{a}}(f + g)(x_0)$ und es gilt

$$\nabla_{\mathbf{a}}(f + g)(x_0) = \nabla_{\mathbf{a}}f(x_0) + \nabla_{\mathbf{a}}g(x_0).$$

b) Sind f und h in x_0 differenzierbar, so ist $h \cdot f$ in x_0 differenzierbar und es gilt

$$D(h \cdot f)(x_0) = Dh(x_0) \cdot f(x_0) + h(x_0) \cdot Df(x_0).$$

Existieren $\nabla_{\mathbf{a}}f(x_0)$ und $\nabla_{\mathbf{a}}h(x_0)$, so existiert $\nabla_{\mathbf{a}}(h \cdot f)(x_0)$ und es gilt

$$\nabla_{\mathbf{a}}(h \cdot f)(x_0) = \nabla_{\mathbf{a}}h(x_0) \cdot f(x_0) + h(x_0) \cdot \nabla_{\mathbf{a}}f(x_0).$$

c) Sind h und p in x_0 differenzierbar und $h(x_0) \neq 0$, so ist $\frac{p}{h}$ in einer Umgebung von x_0 definiert, in x_0 differenzierbar und es gilt

$$D\left(\frac{p}{h}\right)(x_0) = \frac{h(x_0)Dp(x_0) - p(x_0)Dh(x_0)}{h(x_0)^2}.$$

2. Kettenregel für Funktionen mehrerer Variablen:

Ist $f : U \subset \mathbb{R}^n \rightarrow V \subset \mathbb{R}^m$ in x_0 differenzierbar und $g : V \subset \mathbb{R}^m \rightarrow E$ in $f(x_0)$ differenzierbar, so ist $g \circ f : U \subset \mathbb{R}^n \rightarrow E$ in x_0 differenzierbar und es gilt

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

Für $n = 1$ gilt insbesondere: $(g \circ f)'(x_0) = Dg(f(x_0))(f'(x_0))$.

3. Ableitung von Abbildungen mit Werten in Produkträumen:

Sei $f = (f_1, f_2) : U \subset \mathbb{R}^n \rightarrow E_1 \times E_2$ eine Abbildung in das Produkt der normierten Vektorräume E_1 und E_2 . Die Abbildung f ist genau dann in x_0 differenzierbar, wenn f_1 und f_2 in x_0 differenzierbar sind. In diesem Fall gilt:

$$Df(x_0) = (Df_1(x_0), Df_2(x_0)).$$

Beweis. Die Aussagen 1a) und 3. folgen direkt aus der Definition und den Grenzwertsätzen, die Aussagen 1b) und 1c) folgen mit analogen Argumenten wie für Funktionen einer Variablen. Wir lassen dies deshalb als Übungsaufgabe.

Wir beweisen hier nur die *Kettenregel*. Wir setzen dazu $y_0 := f(x_0)$, $L := Df(x_0)$ und $\tilde{L} = Dg(f(x_0))$ und betrachten die Abbildungen

$$\begin{aligned}\varphi(x) &:= f(x) - f(x_0) - L(x - x_0), \\ \psi(y) &:= g(y) - g(y_0) - \tilde{L}(y - y_0), \\ \varrho(x) &:= (g \circ f)(x) - (g \circ f)(x_0) - (\tilde{L} \circ L)(x - x_0).\end{aligned}$$

Nach Voraussetzung gilt

$$\varphi(x) = o(\|x - x_0\|) \quad \text{für } x \rightarrow x_0 \quad \text{und} \quad \psi(y) = o(\|y - y_0\|) \quad \text{für } y \rightarrow y_0.$$

Es ist zu zeigen, dass $\varrho(x) = o(\|x - x_0\|)$ für $x \rightarrow x_0$ gilt. Da

$$\begin{aligned}\varrho(x) &= g(f(x)) - g(f(x_0)) - \tilde{L}(L(x - x_0)) \\ &= g(f(x)) - g(y_0) - \tilde{L}(f(x) - f(x_0) - \varphi(x)) \\ &= (g(f(x)) - g(y_0) - \tilde{L}(f(x) - y_0)) + \tilde{L}(\varphi(x)) \\ &= \psi(f(x)) + \tilde{L}(\varphi(x)),\end{aligned}$$

genügt es dazu $\lim_{x \rightarrow x_0} \frac{\|\tilde{L}(\varphi(x))\|_E}{\|x - x_0\|} = 0$ und $\lim_{x \rightarrow x_0} \frac{\|\psi(f(x))\|_E}{\|x - x_0\|} = 0$ zu zeigen.

Da $\tilde{L} : \mathbb{R}^m \rightarrow E$ linear ist, existiert eine positive Konstante $C_{\tilde{L}} \in \mathbb{R}$, so dass

$$\|\tilde{L}(\varphi(x))\|_E \leq C_{\tilde{L}} \|\varphi(x)\|_{\mathbb{R}^m} \quad \text{für alle } x \in U.$$

Folglich ist

$$0 \leq \lim_{x \rightarrow x_0} \frac{\|\tilde{L}(\varphi(x))\|_E}{\|x - x_0\|} \leq C_{\tilde{L}} \cdot \underbrace{\lim_{x \rightarrow x_0} \frac{\|\varphi(x)\|_{\mathbb{R}^m}}{\|x - x_0\|}}_{=0} = 0.$$

Dies zeigt den 1. Grenzwert. Sei $\varepsilon > 0$. Da $\lim_{y \rightarrow y_0} \frac{\|\psi(y)\|_E}{\|y - y_0\|} = 0$, existiert ein $\delta > 0$, so dass

$$\|\psi(f(x))\|_E < \varepsilon \|f(x) - y_0\|_{\mathbb{R}^m} \quad \text{für alle } x \in U \text{ mit } \|f(x) - y_0\|_{\mathbb{R}^m} < \delta.$$

Da f in x_0 stetig ist, existiert ein $\delta_1 > 0$, so dass

$$\|f(x) - y_0\| < \delta \quad \text{für alle } x \in U \text{ mit } \|x - x_0\| < \delta_1.$$

Für $x \in U$ mit $\|x - x_0\| < \delta_1$ gilt also

$$\begin{aligned} \|\psi(f(x))\|_E &\leq \varepsilon \|f(x) - y_0\|_{\mathbb{R}^m} = \varepsilon \|\varphi(x) + L(x - x_0)\|_{\mathbb{R}^m} \\ &\leq \varepsilon \|\varphi(x)\|_{\mathbb{R}^m} + \varepsilon \cdot C_L \|x - x_0\|_{\mathbb{R}^m} \end{aligned}$$

und damit

$$\frac{\|\psi(f(x))\|}{\|x - x_0\|} \leq \varepsilon \underbrace{\frac{\|\varphi(x)\|}{\|x - x_0\|}}_{x \rightarrow x_0} + \varepsilon \cdot C_L \quad \forall \|x - x_0\| < \delta_1.$$

Wir erhalten daraus

$$0 \leq \limsup_{x \rightarrow x_0} \frac{\|\psi(f(x))\|_E}{\|x - x_0\|} \leq \varepsilon \cdot C_L \quad \forall \varepsilon > 0.$$

Daraus folgt

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

womit die Kettenregel bewiesen ist.

Ist $n = 1$, so ist $g \circ f : U \subset \mathbb{R} \rightarrow E$ eine Funktion, die nur von *einer* reellen Variablen abhängt. Dann erhalten wir die folgende spezielle Form der Kettenregel für die Ableitung in x_0 :

$$(g \circ f)'(x_0) = D(g \circ f)(x_0)(1) = Dg(f(x_0)) (Df(x_0)(1)) = Dg(f(x_0))(f'(x_0)).$$

□

Wir betrachten nun reellwertige Funktionen $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Aus der *Linearen Algebra* ist bekannt, dass es zu jeder linearen Abbildung $L : \mathbb{R}^n \rightarrow \mathbb{R}$ einen eindeutig bestimmten Vektor $v_L \in \mathbb{R}^n$ gibt, so dass

$$L(\mathbf{a}) = \langle v_L, \mathbf{a} \rangle_{\mathbb{R}^n} \quad \forall \mathbf{a} \in \mathbb{R}^n.$$

Definition 6.4. Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ eine in $x_0 \in U$ differenzierbare Funktion. Der Gradient von f in x_0 ist der Vektor $\text{grad}f(x_0) \in \mathbb{R}^n$, definiert durch

$$Df(x_0)(\mathbf{a}) = \nabla_{\mathbf{a}} f(x_0) = \langle \text{grad}f(x_0), \mathbf{a} \rangle \quad \forall \mathbf{a} \in \mathbb{R}^n.$$

Definition 6.5. Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ in $x_0 \in U$ differenzierbar. Dann heißt x_0 regulärer Punkt von f , wenn $\text{grad}f(x_0) \neq 0$.

Ist $\text{grad}f(x_0) = 0$, so sind alle Richtungsableitungen von f im Punkt x_0 Null. In den regulären Punkten $x_0 \in U$ gilt:

Satz 6.4 Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ in $x_0 \in \mathbb{R}^n$ differenzierbar und x_0 ein regulärer Punkt von f . Der Gradient $\text{grad}f(x_0) \in \mathbb{R}^n$ gibt diejenige Richtung an, in der die Funktion f im Punkt x_0 am schnellsten wächst, also die Richtung mit der größten Richtungsableitung.

Beweis. Sei $\mathbf{a} \in \mathbb{R}^n$ mit $\|\mathbf{a}\| = 1$. Dann folgt aus der Cauchy-Schwarzschen Ungleichung

$$\nabla_{\mathbf{a}}f(x_0) = \langle \text{grad}f(x_0), \mathbf{a} \rangle \stackrel{CSU}{\leq} \|\text{grad}f(x_0)\| \cdot \|\mathbf{a}\| = \|\text{grad}f(x_0)\|. \quad (*)$$

Für $\mathbf{a} := \frac{\text{grad}f(x_0)}{\|\text{grad}f(x_0)\|}$ gilt in (*) die Gleichheit. Dies ist die einzige Richtung \mathbf{a} für die in (*) Gleichheit gilt, denn wie wir wissen, gilt in der Cauchy-Schwarzschen Ungleichung genau dann die Gleichheit, wenn beide Vektoren linear abhängig sind. Die Gleichheit in (*), also der maximal mögliche Wert von $\nabla_{\mathbf{a}}f(x_0)$, bei $\|\mathbf{a}\| = 1$, wird somit genau dann angenommen wenn \mathbf{a} in die Richtung von $\text{grad}f(x_0)$ zeigt. \square

Um eine weitere geometrische Bedeutung des Gradienten kennenzulernen, betrachten wir die Niveauflächen einer Funktion.

Definition 6.6. Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ eine differenzierbare Funktion und $c \in \text{Im}(f)$. Dann heißt

$$M_c := \{x \in U \mid f(x) = c\} \subset \mathbb{R}^n$$

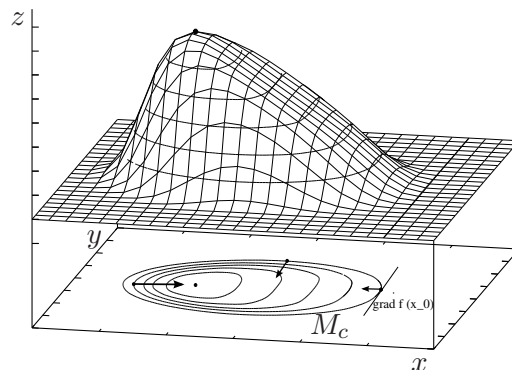
Niveaufläche von f zum Niveau c .

Beispiel: Höhenlinien

Wir betrachten die Abbildung, die die Höhe eines Ortes über dem Meeresspiegel angibt:

$$\begin{aligned} f : U \subset \mathbb{R}^2 &\rightarrow \mathbb{R} \\ (x, y) &\mapsto \text{Höhe des Ortes über dem Meeresspiegel.} \end{aligned}$$

Dann sieht man die Niveauflächen von f als Höhenlinien auf der Landkarte.



Definition 6.7. Sei $p \in M_c$ ein regulärer Punkt auf einer Niveaufläche von f . Der Tangentialraum an die Niveaufläche M_c im Punkt $p \in M_c$ ist die Menge der Vektoren

$$T_p M_c := \left\{ X \in \mathbb{R}^n \mid \exists \text{ differenzierbare Kurve } \gamma : (-\varepsilon, \varepsilon) \longrightarrow M_c \subset \mathbb{R}^n \text{ mit } \begin{array}{l} \gamma(0) = p, \\ \gamma'(0) = X \end{array} \right\}.$$

Die Vektoren X heißen Tangentialvektoren an M_c im Punkt p . Die Ebene $p + T_p M_c$ heißt Tangentialebene in $p \in M_c$.

Satz 6.5 Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar und $p \in M_c$ ein regulärer Punkt von f . Dann gilt für jeden Tangentialvektor $X \in T_p M_c$

$$\text{grad} f(p) \perp X.$$

Beweis. Sei $\gamma : (-\varepsilon, \varepsilon) \rightarrow M_c$ eine differenzierbare Kurve mit $\gamma(0) = p$ und $\gamma'(0) = X$. Da M_c die Niveauläche zum Funktionswert c ist, gilt $f(\gamma(t)) = c$ für alle $t \in (-\varepsilon, \varepsilon)$. Nach der Kettenregel ist dann

$$0 = (f \circ \gamma)'(0) = Df(\gamma(0))(\gamma'(0)) = Df(p)(X) = \langle \text{grad} f(p), X \rangle.$$

Also steht der Gradient $\text{grad} f(p)$ senkrecht auf X , d.h. $\text{grad} f(p) \perp X$. □

Wir werden später sogar noch mehr zeigen: Ist $p \in U$ ein regulärer Punkt von f , so gilt für die Niveauläche M_c , die p enthält

$$T_p M_c = (\text{grad} f(p))^\perp := \{X \in \mathbb{R}^n \mid \langle \text{grad} f(p), X \rangle = 0\}.$$

Somit ist der Tangentialraum $T_p M_c \subset \mathbb{R}^n$ ein $(n - 1)$ -dimensionaler Unterraum, den man mit Hilfe des Gradienten von f berechnen kann.

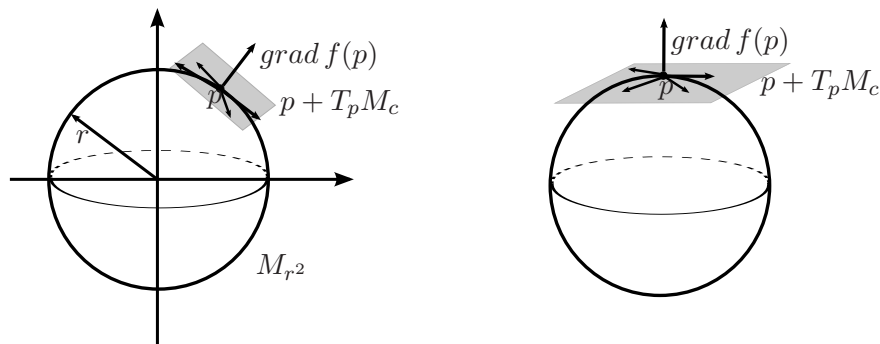
Beispiel: Der Tangentialraum an die Sphäre.

Sei $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ die Funktion $f(x, y, z) := x^2 + y^2 + z^2$ und r eine positive Zahl. Dann ist die Niveauläche von f zum Niveau r^2 gleich der Sphäre vom Radius r im \mathbb{R}^3 :

$$S_r^2 := \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = r^2\} = M_{r^2}.$$

Behauptung: Für den Tangentialraum an die Sphäre gilt:

$$T_p S_r^2 = \{X \in \mathbb{R}^3 \mid \langle X, p \rangle = 0\}.$$



Um das zu beweisen, bestimmen wir zunächst den Gradienten von f . Sei $p = (p_1, p_2, p_3)$ und $h = (h_1, h_2, h_3)$. Dann gilt:

$$\begin{aligned} f(p+h) - f(p) &= (2p_1h_1 + 2p_2h_2 + 2p_3h_3) + (h_1^2 + h_2^2 + h_3^2) \\ &= \underbrace{\langle 2p, h \rangle}_{=Df(p)(h)=\langle \text{grad}f(p), h \rangle} + \underbrace{\|h\|^2}_{o(\|h\|)}. \end{aligned}$$

Folglich ist $\text{grad}f(p) = 2p$ und wir erhalten für den Tangentialraum an die Sphäre nach Satz 6.5:

$$T_p S_r^2 \subset \{X \in \mathbb{R}^3 \mid \langle X, p \rangle = 0\}.$$

Andererseits: Sei $X \in \mathbb{R}^3$ ein Vektor mit $\langle X, p \rangle = 0$. Wir betrachten die differenzierbare Kurve $\gamma : (-\varepsilon, \varepsilon) \rightarrow S_r^2$,

$$\gamma(t) := \cos\left(\frac{\|X\|}{r}t\right) \cdot p + \sin\left(\frac{\|X\|}{r}t\right) \cdot \frac{r}{\|X\|} \cdot X.$$

Das Bild von γ liegt auf der Sphäre, und es gilt $\gamma(0) = p$ und $\gamma'(0) = X$. Wir erhalten folglich $X \in T_p S_r^2$ und somit

$$\{X \in \mathbb{R}^3 \mid \langle X, p \rangle = 0\} \subset T_p S_r^2.$$

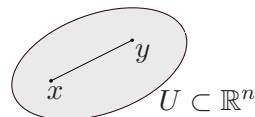
Satz 6.6 (Mittelwertsatz für Funktionen mehrerer reeller Variablen)

Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ eine differenzierbare 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

$$f(y) - f(x) = \langle \text{grad}f(\xi), y - x \rangle_{\mathbb{R}^n} = Df(\xi)(y - x).$$

Beweis. Wir betrachten die Abbildung $h : [0, 1] \rightarrow \mathbb{R}$ definiert durch $h(t) := f(x + t(y - x))$. Nach Voraussetzung ist h auf $[0, 1]$ differenzierbar. Daher existiert nach dem Mittelwertsatz von Lagrange ein $\theta \in (0, 1)$ mit

$$h(1) - h(0) = h'(\theta) \cdot (1 - 0) = h'(\theta).$$



Nach Definition von h und aufgrund der Kettenregel ist dies äquivalent zu

$$f(y) - f(x) = Df(x + \theta(y - x))(y - x) =: Df(\xi)(y - x). \quad \square$$

Satz 6.7 Sei $f : U \rightarrow E$ differenzierbar und U offen und bogenzusammenhängend.

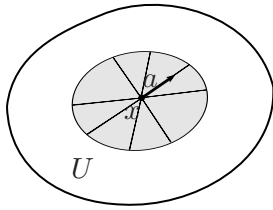
1. $Df(x) = 0$ für alle $x \in U \iff f$ ist konstant.
2. Ist $E = \mathbb{R}$, so gilt: $\text{grad}f(x) = 0$ für alle $x \in U \iff f$ ist konstant.

Beweis. (1) Wenn die Abbildung f konstant ist, so gilt für jede Richtungsableitung

$$\nabla_{\mathbf{a}} f(x) = \lim_{t \rightarrow 0} \frac{f(x + t\mathbf{a}) - f(x)}{t} = \lim_{t \rightarrow 0} \frac{0}{t} = 0,$$

und somit $Df(x) = 0$.

(2) Sei $Df(x) = 0$ für alle $x \in U$ und sei $x_0 \in U$ ein fixierter Punkt. Da U offen ist, existiert ein $\varepsilon(x_0) > 0$, so dass $K(x_0, \varepsilon(x_0)) \subset U$. Für $a \in \mathbb{R}^n$ mit $\|a\| = 1$ ist die Abbildung

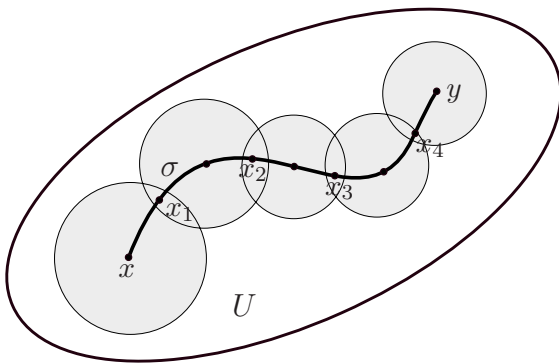


$$\begin{aligned} \varphi_a : (-\varepsilon(x_0), \varepsilon(x_0)) &\longrightarrow \mathbb{R} \\ t &\longmapsto \varphi_a(t) := f(x_0 + ta) \end{aligned}$$

wohldefiniert und differenzierbar und es gilt nach Kettenregel

$$\varphi'_a(t) = Df(x_0 + ta)(a) = 0 \quad \forall t \in (-\varepsilon(x_0), \varepsilon(x_0)).$$

Folglich ist φ_a auf $(-\varepsilon(x_0), \varepsilon(x_0))$ konstant. Da aber $\varphi_a(0) = f(x_0)$ für alle $a \in \mathbb{R}^n$ mit $\|a\| = 1$, ist f auf $K(x_0, \varepsilon(x_0))$ konstant. Seien nun $x, y \in U$ beliebige, aber fixierte Punkte. Da U bogenzusammenhängend ist, existiert eine stetige Abbildung $\sigma : [0, 1] \rightarrow U \subset \mathbb{R}^n$ mit $\sigma(0) = x$ und $\sigma(1) = y$. Wir betrachten die offene Überdeckung $\{K(\sigma(t), \varepsilon(\sigma(t)))\}_{t \in [0, 1]}$ von $\sigma([0, 1])$. Da $[0, 1]$ kompakt und σ stetig ist, ist auch $\sigma([0, 1])$ kompakt.



Folglich existiert eine endliche Teilüberdeckung $K(\sigma(t_1), \varepsilon_1), \dots, K(\sigma(t_r), \varepsilon_r)$ von $\sigma([0, 1])$. Auf jeder dieser offenen Kugeln gilt aber

$$f|_{K(\sigma(t_j), \varepsilon_j)} = \text{konst.} = f(\sigma(t_j)).$$

Daraus folgt

$$f(x) = f(\sigma(t_1)) = f(x_1) = f(\sigma(t_2)) = f(x_2) = \dots = f(\sigma(t_2)) = f(y),$$

mit $x_i \in K(\sigma(t_i), \varepsilon_i) \cap K(\sigma(t_{i+1}), \varepsilon_{i+1})$. Folglich ist f auf U konstant. □

Wir wissen aus Kapitel 2, dass jede bogenzusammenhängende Teilmenge eines metrischen Raumes auch zusammenhängend ist, die Umkehrung gilt im Allgemeinen nicht. Eine *offene* Teilmenge $U \subset \mathbb{R}^n$ ist aber genau dann zusammenhängend, wenn sie bogenzusammenhängend ist (siehe Übungsaufgaben). Satz 6.7 gilt somit auch für offene, zusammenhängende Mengen $U \subset \mathbb{R}^n$.

Wir haben bisher einer in $x_0 \in U$ differenzierbaren Funktion $f : U \subset \mathbb{R}^n \rightarrow E$ folgende Daten zugeordnet:

- Eine lineare Abbildung $Df(x_0) : \mathbb{R}^n \rightarrow E$ (das Differential von f)
- Die Vektoren $\nabla_a f(x_0) \in E$ (die Richtungsableitungen von f).
- Im Falle $E = \mathbb{R}$, den Vektor $\text{grad}f(x_0) \in \mathbb{R}^n$ (den Gradienten von f).

Wie wir aus der Vorlesung über Lineare Algebra wissen, ist jede lineare Abbildung $L : \mathbb{R}^n \rightarrow E$ durch ihre Werte auf den Vektoren einer Basis eindeutig bestimmt. Für den Vektorraum \mathbb{R}^n bietet sich hier die kanonische Basis $e = (e_1, \dots, e_n)$ mit

$$e_i := (0, \dots, 0, 1, 0, \dots, 0), \quad \text{mit 1 an der } i\text{-ten Stelle,}$$

besonders an. Im nächsten Abschnitt werden wir die Differentiale von f durch die Vektoren $Df(x_0)(e_i) = \nabla_{e_i} f(x_0) \in E$ beschreiben.

6.2 Die partiellen Ableitungen

Definition 6.8. Sei $f : U \subset \mathbb{R}^n \rightarrow E$ eine Abbildung und $u \in U$. Man sagt: f besitzt in $u \in U$ die partielle Ableitung nach der i -ten Koordinate, wenn die Richtungsableitung $\nabla_{e_i} f(u)$ von f in x_0 in Richtung e_i existiert.

Bezeichnung:

$$\frac{\partial f}{\partial x_i}(u) := \nabla_{e_i} f(u) = \lim_{t \rightarrow 0} \frac{f(u + te_i) - f(u)}{t} \in E.$$

Sei $u = (u_1, \dots, u_n) \in U$ ein fixierter Punkt und $\varphi_i : (u_i - \varepsilon, u_i + \varepsilon) \subset \mathbb{R} \rightarrow E$ die Funktion

$$\varphi_i(x) := f(u_1, \dots, u_{i-1}, x, u_{i+1}, \dots, u_n).$$

Dann gilt nach Definition

$$\frac{\partial f}{\partial x_i}(u) = \varphi_i'(u_i).$$

Dies motiviert den Namen *partielle Ableitung nach der i -ten Koordinate*.

Beispiel: 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 6.8 Ist $f : U \subset \mathbb{R}^n \rightarrow E$ in $u \in U$ differenzierbar, so existieren die partiellen Ableitungen $\frac{\partial f}{\partial x_1}(u), \dots, \frac{\partial f}{\partial x_n}(u)$ und es gilt

$$Df(u)(h) = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(u), \quad \text{wobei } h = (h_1, \dots, h_n) \in \mathbb{R}^n.$$

Beweis. Nach Satz 6.2 wissen wir, dass die Richtungsableitungen $\nabla_{e_i} f(u)$, und somit die partiellen Ableitungen von f in $u \in U$ existieren. Außerdem folgt aus der Definition der partiellen Ableitungen für $h = (h_1, \dots, h_n) = \sum_{i=1}^n h_i e_i$:

$$Df(u)(h) = Df(u)\left(\sum_{i=1}^n h_i e_i\right) = \sum_{i=1}^n h_i \cdot Df(u)(e_i) = \sum_{i=1}^n h_i \frac{\partial f}{\partial x_i}(u). \quad \square$$

Aus Beispiel 1 in Abschnitt 6.1. wissen wir, dass die partiellen Ableitungen existieren können, ohne dass die Funktion differenzierbar ist. Wir fragen wir uns deshalb, wie man den partiellen Ableitungen ansehen kann, ob die Funktion differenzierbar ist. Dazu geben wir ein hinreichendes Kriterium für den Fall von Funktionen mit Werten in einem *endlich-dimensionalen* Vektorraum $E = \mathbb{R}^m$ an.

Satz 6.9 (Hinreichende Bedingung für Differenzierbarkeit)

Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine Abbildung und $u \in U$. Falls die partiellen Ableitungen $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ existieren und im Punkt $u \in U$ stetig sind, so ist f in $u \in U$ differenzierbar.

Beweis. (1) Es genügt, die Behauptung für reellwertige Funktionen zu beweisen:

Sei $f = (f_1, \dots, f_m)$ die Komponentendarstellung von f . f ist genau dann in $u \in U$ differenzierbar (bzw. stetig), wenn alle Komponenten f_1, \dots, f_m in $u \in U$ differenzierbar (bzw. stetig) sind. In diesem Fall gilt dann auch $Df(u)(h) = (Df_1(u)(h), \dots, Df_m(u)(h))$. Die Behauptung des Satzes gilt somit für f , wenn sie für jede Komponente f_j gilt.

(2) Sei nun $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ reellwertig, $u = (u_1, \dots, u_n)$ und $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ so klein, dass $u + h \in U$. Dann gilt

$$\begin{aligned} f(u+h) - f(u) &= f(u_1 + h_1, u_2, \dots, u_n) - f(u_1, \dots, u_n) \\ &\quad + f(u_1 + h_1, u_2 + h_2, u_3, \dots, u_n) - f(u_1 + h_1, u_2, \dots, u_n) \\ &\quad + f(u_1 + h_1, u_2 + h_2, u_3 + h_3, u_4, \dots, u_n) - f(u_1 + h_1, u_2 + h_2, u_3, \dots, u_n) \\ &\quad \vdots \\ &\quad + f(u_1 + h_1, \dots, u_n + h_n) - f(u_1 + h_1, \dots, u_{n-1} + h_{n-1}, u_n). \end{aligned}$$

Wir betrachten die Funktion $g_1(x) := f(x, u_2, \dots, u_n)$. Nach Voraussetzung existiert die Ableitung $g_1'(x) = \frac{\partial f}{\partial x_1}(x, u_2, \dots, u_n)$. Wenden wir den Mittelwertsatz auf g_1 an, so erhalten wir eine Zahl ξ_1 zwischen u_1 und $u_1 + h_1$, so dass gilt

$$f(u_1 + h_1, u_2, \dots, u_n) - f(u_1, \dots, u_n) = h_1 \cdot \frac{\partial f}{\partial x_1}(\xi_1, u_2, \dots, u_n).$$

Analog existiert für alle $j \in \{2, \dots, n\}$ ein ξ_j zwischen u_j und $u_j + h_j$, so dass gilt

$$\begin{aligned} &f(u_1 + h_1, \dots, u_j + h_j, u_{j+1}, \dots, u_n) - f(u_1 + h_1, \dots, u_{j-1} + h_{j-1}, u_j, \dots, u_j) \\ &= h_j \cdot \frac{\partial f}{\partial x_j}(\underbrace{u_1 + h_1, \dots, u_{j-1} + h_{j-1}, \xi_j, u_{j+1}, \dots, u_n}_{=: c_j = c_j(h)}). \end{aligned}$$

Folglich gilt

$$f(u+h) - f(u) = \sum_{j=1}^n h_j \frac{\partial f}{\partial x_j}(c_j) = \underbrace{\sum_{j=1}^n h_j \frac{\partial f}{\partial x_j}(u)}_{=: L(h), L \text{ linear}} + \underbrace{\sum_{j=1}^n h_j \left(\frac{\partial f}{\partial x_j}(c_j) - \frac{\partial f}{\partial x_j}(u) \right)}_{=: R(h)}.$$

Es bleibt zu zeigen, dass $R(h) = o(\|h\|)$ für $h \rightarrow 0$. Mit der Cauchy-Schwarzschen Ungleichung erhalten wir:

$$|R(h)| = \left| \sum_{j=1}^n h_j \cdot \underbrace{\left(\frac{\partial f}{\partial x_j}(c_j) - \frac{\partial f}{\partial x_j}(u) \right)}_{=: b_j = b_j(h)} \right| \stackrel{CSU}{\leq} \|h\| \cdot \|b\|,$$

wobei $b := (b_1, \dots, b_n)$. Folglich gilt

$$0 \leq \frac{|R(h)|}{\|h\|} \leq \|b(h)\|.$$

Für $h \rightarrow 0$ gilt $c_j(h) \rightarrow u$ und somit wegen der Stetigkeit der partiellen Ableitungen in u auch $b_j(h) \rightarrow 0$, also $\|b(h)\| \rightarrow 0$. Daraus folgt

$$\lim_{h \rightarrow 0} \frac{|R(h)|}{\|h\|} = 0.$$

Folglich ist f in $u \in U$ differenzierbar und es gilt

$$Df(u)(h) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(u) h_j. \quad \square$$

Bemerkung:

1. Die Stetigkeit der partiellen Ableitungen ist nicht notwendig für die Differenzierbarkeit.

Als Beispiel betrachten wir die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) := \begin{cases} (x^2 + y^2) \sin\left(\frac{1}{\sqrt{x^2 + y^2}}\right) & \text{für } x^2 + y^2 > 0, \\ 0 & \text{für } x = y = 0. \end{cases}$$

f ist in $(0, 0)$ differenzierbar, aber $\frac{\partial f}{\partial x}$ und $\frac{\partial f}{\partial y}$ sind in $(0, 0)$ nicht stetig (Übungsaufgabe).

2. Eine Funktion f kann in $u \in U$ differenzierbar sein, ohne dass die partiellen Ableitungen in einer Umgebung von u existieren. Als Beispiel betrachten wir $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ definiert durch $f(x, y) = |xy|$. f ist in $(0, 0)$ differenzierbar, es existiert aber keine Umgebung von $(0, 0)$, auf der die partiellen Ableitungen existieren.

Satz 6.10 (Kettenregel für partielle Ableitungen)

Seien $g_1, \dots, g_m : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ in $u \in U$ und $f : \mathbb{R}^m \rightarrow E$ in $(g_1(u), \dots, g_m(u))$ differenzierbar. Dann ist die Abbildung $F : U \subset \mathbb{R}^n \rightarrow E$,

$$F(x) := f(g_1(x), \dots, g_m(x)),$$

in $u \in U$ differenzierbar und für die partiellen Ableitungen gilt

$$\frac{\partial F}{\partial x_i}(u) = \sum_{j=1}^m \frac{\partial g_j}{\partial x_i}(u) \cdot \frac{\partial f}{\partial y_j}(g_1(u), \dots, g_m(u)).$$

Ist $n = 1$, so gilt insbesondere:

$$F'(u) = \sum_{j=1}^m g'_j(u) \cdot \frac{\partial f}{\partial y_j}(g_1(u), \dots, g_m(u)).$$

Beweis. Wir betrachten $g : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, definiert durch $g(x) := (g_1(x), \dots, g_m(x))$. Dann gilt $F = f \circ g$ und nach Kettenregel für die Differentiale $DF(u) = Df(g(u)) \circ Dg(u)$. Für die partiellen Ableitungen folgt daraus

$$\begin{aligned} \frac{\partial F}{\partial x_i}(u) &= DF(u)(e_i) = Df(g(u))(Dg(u)(e_i)) = Df(g(u))\left(\sum_{j=1}^m Dg_j(u)(e_i) \cdot e_j\right) \\ &= Df(g(u))\left(\sum_{j=1}^m \frac{\partial g_j}{\partial x_i}(u) \cdot e_j\right) = \sum_{j=1}^m \frac{\partial g_j}{\partial x_i}(u) \cdot Df(g(u))(e_j) \\ &= \sum_{j=1}^m \frac{\partial g_j}{\partial x_i}(u) \cdot \frac{\partial f}{\partial y_j}(g(u)). \end{aligned} \quad \square$$

Aus der Vorlesung über lineare Algebra ist bekannt, dass man lineare Abbildungen durch Matrizen beschreiben kann. Wir erinnern nochmal an das Verfahren:

Sei V^n ein endlich-dimensionaler reeller Vektorraum und $a = (a_1, \dots, a_n)$ eine Basis von V^n . Dann ist die Abbildung $\phi_a : V \rightarrow \mathbb{R}^n$,

$$\phi_a(v) = \phi_a\left(\underbrace{\sum_{i=1}^n x_i a_i}_{=v}\right) := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

ein Isomorphismus der Vektorräume. Sei nun W^m ein weiterer endlich-dimensionaler reeller Vektorraum, $b = (b_1, \dots, b_m)$ eine Basis in W^m und $L : V^n \rightarrow W^m$ eine lineare Abbildung. Dann kann man L die durch das folgende Diagramm definierte $(m \times n)$ -Matrix $M_b^a(L)$ zuordnen:

$$\begin{array}{ccc} V & \xrightarrow{L} & W \\ \phi_a \downarrow & & \downarrow \phi_b \\ \mathbb{R}^n & \xrightarrow{M_b^a(L)} & \mathbb{R}^m \end{array} .$$

Dann gilt

$$M_b^a(L) = \left(\phi_b(L(a_1)) \quad \phi_b(L(a_2)) \quad \dots \quad \phi_b(L(a_n)) \right).$$

In der i -ten Spalte dieser Matrix stehen also die Komponenten des Vektors $L(a_i)$ bezüglich der Basis b .

In diesem Sinne ordnen wir nun dem Differential einer differenzierbaren Abbildung $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine $(m \times n)$ -Matrix zu:

Definition 6.9. Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine in $u \in U$ differenzierbare Funktion. Die Jacobi-Matrix von f im Punkt $u \in U$ ist die $(m \times n)$ -Matrix $M(Df(u))$, die der linearen Abbildung $Df(u) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ bezüglich der kanonischen Basen in \mathbb{R}^n bzw. \mathbb{R}^m entspricht.

Wir erhalten die folgende Formel für die Jacobi-Matrix:

Sei $f = (f_1, \dots, f_m)$ die Komponentendarstellung von f . Dann gilt:

$$\begin{aligned}
M(Df(u)) &= \left(\phi_e(Df(u)(e_1)) \quad \dots \quad \phi_e(Df(u)(e_n)) \right) \\
&= \left(\phi_e\left(\frac{\partial f}{\partial x_1}(u)\right) \quad \dots \quad \phi_e\left(\frac{\partial f}{\partial x_n}(u)\right) \right) \\
&= \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(u) & \frac{\partial f_1}{\partial x_2}(u) & \dots & \frac{\partial f_1}{\partial x_n}(u) \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial f_m}{\partial x_1}(u) & \frac{\partial f_m}{\partial x_2}(u) & \dots & \frac{\partial f_m}{\partial x_n}(u) \end{pmatrix}
\end{aligned}$$

Für $m = 1$ stimmt die Jacobi-Matrix der reellwertigen Funktion $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ mit ihrem Gradienten überein:

$$\begin{aligned}
\text{grad} f(u) &= \sum_{i=1}^n \langle \text{grad} f(u), e_i \rangle \cdot e_i = \sum_{i=1}^n Df(u)(e_i) \cdot e_i = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(u) \cdot e_i \\
&= \left(\frac{\partial f}{\partial x_1}(u), \dots, \frac{\partial f}{\partial x_n}(u) \right).
\end{aligned}$$

Insbesondere kann man die Jacobi-Matrix von $f = (f_1, \dots, f_m) : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ auch durch die Gradienten der Komponentenfunktionen ausdrücken:

$$M(Df(u)) = \begin{pmatrix} \text{grad} f_1(u) \\ \text{grad} f_2(u) \\ \vdots \\ \text{grad} f_m(u) \end{pmatrix}.$$

Wir definieren als nächstes die partiellen Ableitungen höherer Ordnung.

Sei $f : U \subset \mathbb{R}^n \rightarrow E$ eine Funktion, für die die partielle Ableitung $\frac{\partial f}{\partial x_i} : U \subset \mathbb{R}^n \rightarrow E$ existiert. Wenn die partielle Ableitung der Funktion $\frac{\partial f}{\partial x_i} : U \subset \mathbb{R}^n \rightarrow E$ 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

$$\frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}} := \frac{\partial}{\partial x_{i_k}} \left(\frac{\partial^{k-1} f}{\partial x_{i_{k-1}} \dots \partial x_{i_1}} \right) : U \subset \mathbb{R}^n \rightarrow E$$

heißt (falls sie existiert) k -te partielle Ableitung von f nach den Variablen x_{i_1}, \dots, x_{i_k} .

Insbesondere schreibt man zur Abkürzung

$$\frac{\partial^k f}{\partial x_i \dots \partial x_i} =: \frac{\partial^k f}{\partial x_i^k}.$$

Definition 6.10. Eine Funktion $f : U \subset \mathbb{R}^n \rightarrow E$ heißt k -fach stetig differenzierbar, C^k -Funktion oder von der Klasse C^k , $k \in \mathbb{N}$, wenn alle partiellen Ableitungen von f der Ordnung $\leq k$ existieren und stetig sind.

Bezeichnung: $C^k(U, E) := \{f : U \rightarrow E \mid f \text{ } C^k\text{-Funktion}\}.$

Offensichtlich gilt für $k \in \mathbb{N}$

$$C^k(U, E) \subset C^{k-1}(U, E) \subset C^{k-2}(U, E) \subset \dots \subset C^0(U, E).$$

Wir fragen uns nun, ob man bei den höheren partiellen Ableitungen die Reihenfolge der Ableitungen vertauschen kann. Wir betrachten zunächst zwei Beispiele:

Beispiel 1: 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

$$\frac{\partial f}{\partial x}(x, y) = 3x^2y^2, \quad \frac{\partial f}{\partial y}(x, y) = 2x^3y \quad \text{sowie} \quad \frac{\partial^2 f}{\partial y \partial x}(x, y) = 6x^2y = \frac{\partial f}{\partial x \partial y}(x, y).$$

In diesem Fall kann man die Reihenfolge der partiellen Ableitungen nach x und y vertauschen.

Beispiel 2: Wir betrachten die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, definiert durch

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

Für die partiellen Ableitungen erhalten wir in $(x, y) \neq (0, 0)$:

$$\frac{\partial f}{\partial x}(x, y) = y \frac{x^4 + 4x^2y^2 - y^4}{(x^2 + y^2)^2}, \quad \frac{\partial f}{\partial y}(x, y) = x \frac{x^4 - 4x^2y^2 - y^4}{(x^2 + y^2)^2}.$$

In $(x, y) = (0, 0)$ gilt:

$$\frac{\partial f}{\partial x}(0, 0) = \lim_{t \rightarrow 0} \frac{f(t, 0) - f(0, 0)}{t} = 0, \quad \frac{\partial f}{\partial y}(0, 0) = \lim_{t \rightarrow 0} \frac{f(0, t) - f(0, 0)}{t} = 0.$$

Für die 2-fachen partiellen Ableitungen in $(0, 0)$ folgt:

$$\frac{\partial^2 f}{\partial y \partial x}(0, 0) = \lim_{t \rightarrow 0} \frac{\frac{\partial f}{\partial x}(0, t) - \frac{\partial f}{\partial x}(0, 0)}{t} = -1, \quad \frac{\partial^2 f}{\partial x \partial y}(0, 0) = \lim_{t \rightarrow 0} \frac{\frac{\partial f}{\partial y}(t, 0) - \frac{\partial f}{\partial y}(0, 0)}{t} = +1.$$

In diesem Fall kann man die Reihenfolge der partiellen Ableitungen nach x und y also nicht vertauschen.

Der folgende Satz gibt eine hinreichende Bedingung dafür an, dass man die Reihenfolge der partiellen Ableitungen vertauschen kann.

Satz 6.11 (Lemma von Schwarz) Sei $f \in C^k(U, \mathbb{R}^m)$ eine k -fach stetig differenzierbare Funktion. Dann sind alle partiellen Ableitungen der Ordnung $\leq k$ 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, 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. Die Behauptung die für partiellen Ableitungen höherer Ordnung beweist man dann durch Induktion. Sei $u = (u_1, u_2) \in U$ und seien $h = (h_1, h_2)$ so gewählt, dass $h_1, h_2 \neq 0$ und

$$[u_1 - |h_1|, u_1 + |h_1|] \times [u_2 - |h_2|, u_2 + |h_2|] \subset U.$$

Wir betrachten die Funktion $\varphi(x) := f(x, u_2 + h_2) - f(x, u_2)$. Dann ist φ auf dem Intervall $[u_1 - |h_1|, u_1 + |h_1|]$ differenzierbar. Nach dem Mittelwertsatz existiert ein ξ_1 zwischen u_1 und $u_1 + h_1$, so dass gilt

$$\varphi(u_1 + h_1) - \varphi(u_1) = h_1 \cdot \varphi'(\xi_1).$$

Sei nun

$$F(h_1, h_2) := f(u_1 + h_1, u_2 + h_2) - f(u_1 + h_1, u_2) - f(u_1, u_2 + h_2) + f(u_1, u_2).$$

Dann gilt

$$\begin{aligned} F(h_1, h_2) &= \varphi(u_1 + h_1) - \varphi(u_1) = h_1 \cdot \varphi'(\xi_1) \\ &= h_1 \left(\frac{\partial f}{\partial x_1}(\xi_1, u_2 + h_2) - \frac{\partial f}{\partial x_1}(\xi_1, u_2) \right). \end{aligned}$$

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 u_2 und $u_2 + h_2$, 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 $\psi(x) := f(u_1 + h_1, x) - f(u_1, x)$, so existieren nach dem Mittelwertsatz ein $\tilde{\xi}_1$ zwischen u_1 und $u_1 + h_1$ und ein $\tilde{\xi}_2$ zwischen u_2 und $u_2 + h_2$, 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 $u = (u_1, u_2)$. 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 (u_1, u_2) nach Voraussetzung stetig sind, folgt mit $h = (h_1, h_2) \rightarrow (0, 0)$

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(u) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(u).$$

□

Bemerkung: In der Literatur wird oft die folgende Bezeichnung für die partiellen Ableitungen benutzt:

$$f_x := \frac{\partial f}{\partial x}, \quad f_{xy} := \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) = \frac{\partial^2 f}{\partial y \partial x}.$$

6.3 Die Taylorformel für Funktionen mehrerer reeller Variablen

Wie Funktionen einer reeller Variablen, kann man auch Funktionen mehrerer reeller Variablen durch ihre Taylorpolynome approximieren. Um dies zu beschreiben führen wir zunächst einige Bezeichnungen ein.

Definition 6.11. Ein Multiindex ist ein n -Tupel von natürlichen Zahlen $\alpha = (\alpha_1, \dots, \alpha_n)$, $\alpha_j \in \mathbb{N}_0$. Die Zahl $|\alpha| := \alpha_1 + \dots + \alpha_n$ heißt die Ordnung des Multiindex und wir definieren $\alpha! := (\alpha_1!) \cdot \dots \cdot (\alpha_n!)$. Für $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ sei

$$y^\alpha := y_1^{\alpha_1} \cdot \dots \cdot y_n^{\alpha_n}.$$

Für eine Funktion $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ bezeichne (sofern existent)

$$\frac{\partial^{|\alpha|} f}{\partial x^\alpha} := \frac{\partial^{|\alpha|} f}{(\partial x_1)^{\alpha_1} \dots (\partial x_n)^{\alpha_n}}.$$

Satz 6.12 (Taylorformel)

Sei $U \subset \mathbb{R}^n$ offen und $f \in C^{k+1}(U, \mathbb{R})$. Seien $x_0, x \in U$ zwei Punkte, für die die Strecke $\overline{x_0 x}$ in U liegt. Dann existiert ein $\xi \in \overline{x_0 x}$ mit $\xi \neq x_0, \xi \neq x$, so dass

$$f(x) = \underbrace{\sum_{|\alpha| \leq k} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x_0) (x - x_0)^\alpha}_{=: T_k(f, x_0)(x)} + \underbrace{\sum_{|\alpha|=k+1} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(\xi) (x - x_0)^\alpha}_{=: R_k(f, x_0)(x)}.$$

k -tes Taylorpolynom
 k -tes Restglied

Beweis. Wir betrachten die Funktion $g : [0, 1] \rightarrow \mathbb{R}$ mit $g(t) := f(x_0 + t(x - x_0))$. Dann ist $g \in C^{k+1}([0, 1], \mathbb{R})$. Wir wenden auf g die Taylorformel für Funktionen einer reellen Variablen an und erhalten: Es existiert ein $\theta \in (0, 1)$, so dass

$$g(1) = \sum_{j=0}^k \frac{1}{j!} g^{(j)}(0) + \underbrace{\frac{1}{(k+1)!} g^{(k+1)}(\theta)}_{\text{Lagrange-Restglied}}.$$

Mit der Kettenregel für partielle Ableitungen erhalten wir:

$$\begin{aligned}
 g^{(0)}(0) &= g(0) = f(x_0). \\
 g^{(1)}(0) &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0) \cdot (x - x_0)_i = \sum_{|\alpha|=1} \frac{\partial^\alpha f}{\partial x^\alpha}(x_0) \cdot (x - x_0)^\alpha \\
 g^{(2)}(0) &= \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) \cdot (x - x_0)_i \cdot (x - x_0)_j \quad (*) \\
 &= \sum_{|\alpha|=2} \frac{2!}{\alpha!} \frac{\partial^2 f}{\partial x^\alpha}(x) \cdot (x - x_0)^\alpha. \quad (**)
 \end{aligned}$$

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 erhält man durch mehrfaches Anwenden der Kettenregel und etwas Kombinatorik

$$g^{(j)}(0) = \sum_{|\alpha|=j} \frac{j!}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x_0) \cdot (x - x_0)^\alpha.$$

Damit folgt,

$$f(x) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} \frac{\partial^{|\alpha|} f}{\partial x^\alpha}(x_0) \cdot (x - x_0)^\alpha + \sum_{|\alpha|=k+1} \frac{1}{\alpha!} \frac{\partial^{k+1} f}{\partial x^\alpha}(\xi) \cdot (x - x_0)^\alpha$$

mit $\xi := x_0 + \theta(x - x_0)$. □

Definition 6.12. Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ 2 mal stetig differenzierbar. Die symmetrische $(n \times n)$ -Matrix

$$\text{Hess}f(u) := \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(u) \right)_{i,j=1,\dots,n}$$

heißt Hesse-Matrix von f im Punkt $u \in U$.

Wir betrachten die Spezialfälle der Taylorformel für $k = 0, 1, 2$ genauer.

Für $k = 0$ erhalten wir aus der Taylorformel den Mittelwertsatz:

$$\begin{aligned}
 f(x) &= f(x_0) + \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\xi) \cdot (x - x_0)_j \\
 &= f(x_0) + \langle \text{grad}f(\xi), x - x_0 \rangle.
 \end{aligned}$$

Für $k = 1$ besagt die Taylorformel

$$\begin{aligned}
 f(x) &= f(x_0) + \langle \text{grad}f(x_0), x - x_0 \rangle + \sum_{j=1}^n \frac{1}{2!} \cdot \frac{\partial^2 f}{\partial x_j^2}(\xi) \cdot (x - x_0)_j^2 \\
 &\quad + \sum_{i < j} \frac{1}{1!1!} \cdot \frac{\partial^2 f}{\partial x_i \partial x_j}(\xi) \cdot (x - x_0)_i \cdot (x - x_0)_j \\
 &= f(x_0) + \langle \text{grad}f(x_0), x - x_0 \rangle + \frac{1}{2} (x - x_0) \text{Hess}f(\xi) (x - x_0)^t.
 \end{aligned}$$

Für $k = 2$ erhalten wir

$$f(x) = f(x_0) + \langle \operatorname{grad} f(x_0), x - x_0 \rangle + \frac{1}{2}(x - x_0) \operatorname{Hess} f(x_0) (x - x_0)^t + \sum_{|\alpha|=3} \frac{1}{\alpha!} \cdot \frac{\partial^3 f}{\partial x^\alpha}(\xi) \cdot (x - x_0)^\alpha.$$

6.4 Lokale Extrema für Funktionen mehrerer reeller Variablen

In diesem Abschnitt wollen wir mit Hilfe der Differentialrechnung Kriterien für das Vorliegen lokaler Extremwerte für reellwertige Funktionen, die von mehreren reellen Variablen abhängen, herleiten. Dazu erinnern wir zunächst an einige Kenntnisse aus der Vorlesung *Lineare Algebra*.

Definition 6.13. 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 $xAx^t \leq 0 \quad \forall x = (x_1, \dots, x_n) \in \mathbb{R}^n$.

Algebraische Fakten:

1. Jede symmetrische reelle $(n \times n)$ -Matrix ist diagonalisierbar, d.h. sie hat n reelle Eigenwerte.
2. Für eine symmetrische reelle $(n \times n)$ -Matrix A gilt:
 - $A > 0 \iff$ alle Eigenwerte von A sind positiv.
 - $A < 0 \iff$ alle Eigenwerte von A sind negativ.
 - $A \geq 0 \iff$ alle Eigenwerte von A sind ≥ 0 .
 - $A \leq 0 \iff$ alle Eigenwerte von A sind ≤ 0 .
3. 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:

$$\begin{aligned} A > 0 &\iff \det A(k) > 0 \quad \forall k = 1, \dots, n. \\ A < 0 &\iff (-1)^k \cdot \det A(k) > 0 \quad \forall k = 1, \dots, n. \end{aligned}$$

Definition 6.14. Sei $U \subset \mathbb{R}^n$ eine offene Menge und $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ eine reellwertige Funktion. Man sagt:

1. f nimmt in $x_0 \in U$ ein lokales Minimum an, falls ein $\varepsilon > 0$ existiert, so dass $f(x) \geq f(x_0)$ für alle $x \in U$ mit $\|x - x_0\| < \varepsilon$.
2. f nimmt in $x_0 \in U$ ein striktes (oder isoliertes) lokales Minimum an, falls ein $\varepsilon > 0$ existiert, so dass $f(x) > f(x_0)$ für alle $x \in U$ mit $0 < \|x - x_0\| < \varepsilon$.
3. f nimmt in $x_0 \in U$ ein lokales Maximum an, falls ein $\varepsilon > 0$ existiert, so dass $f(x) \leq f(x_0)$ für alle $x \in U$ mit $\|x - x_0\| < \varepsilon$.
4. f nimmt in $x_0 \in U$ ein striktes (oder isoliertes) lokales Maximum an, falls ein $\varepsilon > 0$ existiert, so dass $f(x) < f(x_0)$ für alle $x \in U$ mit $0 < \|x - x_0\| < \varepsilon$.

Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar. Dann heißt $x_0 \in U$ kritischer Punkt von f , wenn $Df(x_0) = 0$ (bzw. dazu äquivalent wenn $\text{grad}f(x_0) = 0$). Anderenfalls heißt $x_0 \in U$ regulärer Punkt von f .

Satz 6.13 Sei $U \subset \mathbb{R}^n$ eine offene Menge und $f \in C^2(U, \mathbb{R})$.

1. Hat f in $x_0 \in U$ ein lokales Maximum (lokales Minimum), so ist x_0 ein kritischer Punkt von f und die Hesse-Matrix $\text{Hess}f(x_0)$ ist negativ semidefinit (positiv semidefinit).
2. Ist $x_0 \in U$ ein kritischer Punkt von f und die Hesse-Matrix $\text{Hess}f(x_0)$ negativ definit (positiv definit), so nimmt f in x_0 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 $x_0 \in U$ ein lokales Maximum. Sei $a = (a_1, \dots, a_n) \in \mathbb{R}^n$. Da x_0 ein innerer Punkt von U ist, existiert ein $\varepsilon > 0$, so dass die Funktion $g : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ mit $g(t) := f(x_0 + ta)$ definiert ist. g ist 2-mal differenzierbar und hat in $t = 0$ ein lokales Maximum. Folglich gilt $g'(0) = 0$ und $g''(0) \leq 0$. Aus der Kettenregel für Differentiale folgt:

$$0 = g'(0) = Df(x_0)(a).$$

Dies gilt für alle $a \in \mathbb{R}^n$. Folglich ist $Df(x_0) = 0$, also $x_0 \in U$ ein kritischer Punkt von f . Aus der Kettenregel für partielle Ableitungen folgt:

$$g'(t) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\gamma(t)) \cdot a_i,$$

$$g''(t) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\gamma(t)) \cdot a_i a_j.$$

Folglich ist

$$0 \geq g''(0) = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) \cdot a_i a_j = a \text{Hess}f(x_0) a^t.$$

Dies gilt für alle $a \in \mathbb{R}^n$, somit ist die Hesse-Matrix von f in x_0 negativ semidefinit.

(2) Sei $x_0 \in U$ ein kritischer Punkt von f und $\text{Hess}f(x_0) < 0$. Wir zeigen zunächst, dass

es eine Kugel $K(x_0, r) \subset U$ um x_0 gibt, so dass $\text{Hess}f(\xi) < 0$ für alle $\xi \in K(x_0, r)$. Dazu betrachten wir die Abbildung $F: \mathbb{R}^n \times U \rightarrow \mathbb{R}$

$$F(a, x) := a \text{Hess}f(x) a^t = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(x) \cdot a_i a_j.$$

Da f zweifach stetig-differenzierbar ist, ist F stetig. Da $F(a, x_0) < 0$ für alle $a \in \mathbb{R}^n$, existieren wegen der Stetigkeit von F für jedes $a \in \mathbb{R}^n$ Kugeln $K(a, \varepsilon_a) \subset \mathbb{R}^n$ und $K(x_0, \mu_a) \subset U$, so dass $F|_{K(a, \varepsilon_a) \times K(x_0, \mu_a)} < 0$. Wir betrachten nun die folgende offene Überdeckung \mathcal{U} der Sphäre $S^{n-1} := \{a \in \mathbb{R}^n \mid \|a\| = 1\}$:

$$\mathcal{U} := \bigcup_{a \in S^{n-1}} K(a, \varepsilon_a).$$

Da die Sphäre S^{n-1} kompakt ist, kann man aus \mathcal{U} eine endliche Teilüberdeckung auswählen. Es gibt also $p_1, \dots, p_N \in S^{n-1}$ so dass

$$S^{n-1} \subset K(p_1, \varepsilon_{p_1}) \cup \dots \cup K(p_N, \varepsilon_{p_N}).$$

Wir betrachten nun die Kugel $K(x_0, r)$ um $x_0 \in U$ mit $r := \min(\mu_{p_1}, \dots, \mu_{p_N})$. Für diese Kugel gilt dann $F|_{S^{n-1} \times K(x_0, r)} < 0$ und somit auch

$$F|_{\mathbb{R}^n \times K(x_0, r)} < 0.$$

Folglich ist die Hesse-Matrix $\text{Hess}f(\xi)$ negativ definit für alle $\xi \in K(x_0, r)$.

Wir approximieren die Funktion f nun bei x_0 durch das Taylorpolynom 1. Grades und erhalten für alle $x \in K(x_0, r)$ ein $\xi \in \overline{x_0 x} \subset K(x_0, r)$, so dass

$$f(x) = f(x_0) + \langle \text{grad}f(x_0), x - x_0 \rangle + \frac{1}{2}(x - x_0) \text{Hess}f(\xi)(x - x_0)^t.$$

Da x_0 ein kritischer Punkt von f ist, gilt $\text{grad}f(x_0) = 0$. Außerdem wissen wir aus dem oben Bewiesenen, dass die Hesse-Matrix $\text{Hess}f(\xi)$ negativ definit ist. Daraus folgt

$$f(x) - f(x_0) < 0$$

für alle $x \in K(x_0, r)$ mit $x \neq x_0$. Dies zeigt, dass f in x_0 ein isoliertes lokales Maximum hat. □

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. Bestimmen 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 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.

6.5 Der Satz über den lokalen Diffeomorphismus

Definition 6.15. *Es seien $U, V \subset \mathbb{R}^n$ offene Teilmengen des \mathbb{R}^n . Eine Abbildung $f : U \rightarrow V$ heißt C^k -Diffeomorphismus, $k \geq 1$, wenn Folgendes gilt:*

1. f ist bijektiv und
2. f und f^{-1} sind C^k -Abbildungen.

Einen C^1 -Diffeomorphismus nennen wir auch kurz Diffeomorphismus. Ein Diffeomorphismus heißt auch Koordinatentransformation, denn mittels der Abbildung f ordnet man jedem Punkt aus V neue Koordinaten zu. Wir werden später sehen, wie nützlich solche Koordinatentransformationen sind. Sie vereinfachen oft die Berechnung von Integralen oder das Lösen von Differentialgleichungen. Deshalb werden wir uns in diesem Abschnitt damit beschäftigen, wie man feststellen kann, ob eine Abbildung f ein C^k -Diffeomorphismus (bzw. eine Koordinatentransformation) ist. Zunächst sehen wir uns drei Beispiele an.

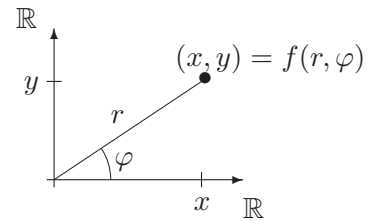
Beispiel 1: *Polarkoordinaten in der Ebene*

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*
 $(x, y) = f(r, \varphi)$.



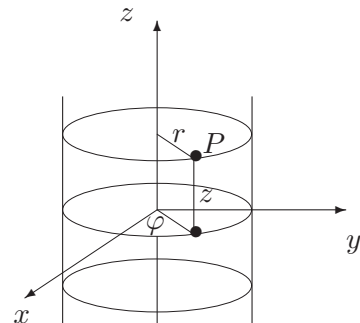
Beispiel 2: *Zylinderkoordinaten im \mathbb{R}^3*

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*
 $P = f(r, \varphi, z)$.



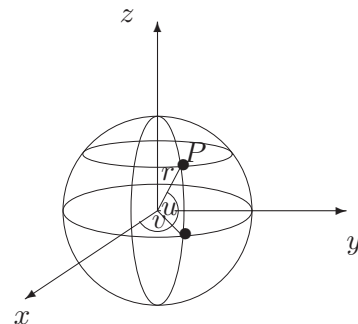
Beispiel 3: *Kugelkoordinaten im \mathbb{R}^3*

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*
 $P = f(r, u, v)$.



Die Bijektivität der Abbildungen in den letzten drei Beispielen folgt aus ihrer geometrischen Konstruktion. Die Abbildungen sind offensichtlich C^∞ . Um zu überprüfen, ob die inverse Abbildung ebenfalls C^∞ 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 folgende notwendige Bedingung dafür, dass $f : U \longrightarrow V$ ein Diffeomorphismus ist.

Satz 6.14 *Sei $f : U \longrightarrow V$ ein Diffeomorphismus zwischen offenen Mengen des \mathbb{R}^n . Dann ist das Differential $Df(x) : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ für alle $x \in U$ ein Isomorphismus, und es gilt*

$$(Df(x))^{-1} = Df^{-1}(f(x)).$$

Beweis. Da f stetig differenzierbar ist, ist f in jedem Punkt von U differenzierbar.

Wir wenden die Kettenregel für Differentiale auf die Abbildungen $f \circ f^{-1} = Id_V$ und $f^{-1} \circ f = Id_U$ an und erhalten:

$$\begin{aligned} D(f \circ f^{-1})(y) &= Df(f^{-1}(y)) \circ Df^{-1}(y) = D(Id_{\mathbb{R}^n})(y) = Id_{\mathbb{R}^n} \quad \text{und} \\ D(f^{-1} \circ f)(x) &= Df^{-1}(f(x)) \circ Df(x) = D(Id_{\mathbb{R}^n})(x) = Id_{\mathbb{R}^n}. \end{aligned}$$

Betrachten wir $y = f(x)$, so folgt $Df^{-1}(f(x)) = (Df(x))^{-1}$. \square

Bemerkung: Für Funktionen einer Variablen gilt auch die Umkehrung von Satz 6.14: Sei $f : U \subset \mathbb{R} \rightarrow \mathbb{R}$ eine C^1 -Abbildung von einer offenen Teilmenge $U \subset \mathbb{R}$ mit $f'(x) \neq 0$ für alle $x \in U$. Dann gilt:

1. $V = f(U) \subset \mathbb{R}$ ist offen und
2. $f : U \rightarrow V$ ist ein Diffeomorphismus.

In höheren Dimensionen gilt diese Aussage nicht mehr! 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 (sogar unendlich oft) und für die Determinante der Jacobi-Matrix von f gilt:

$$\det(Df(x, y)) = \det \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix} = e^{2x} > 0.$$

Folglich ist das Differential $Df(x, y) : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ein Isomorphismus für alle $(x, y) \in \mathbb{R}^2$. Aber f ist offensichtlich nicht injektiv, f also kein Diffeomorphismus, da f^{-1} nicht global existiert.

Wir wollen nun untersuchen, unter welchen Bedingungen eine C^1 -Funktion f wenigstens lokal eine C^1 -Umkehrfunktion besitzt.

Definition 6.16. Eine Abbildung $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ heißt lokaler C^k -Diffeomorphismus um $x_0 \in U$, wenn es (offene) Umgebungen $\tilde{U} \subset U$ von x_0 und $\tilde{V} \subset \mathbb{R}^n$ von $f(x_0)$ gibt, so dass $f|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ein C^k -Diffeomorphismus ist.

Wir beweisen zunächst einen nützlichen Hilfssatz.

Lemma 6.17. Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine C^1 -Funktion, $f = (f_1, \dots, f_n)$, $Q \subset U$ ein Quader und es gelte

$$\left| \frac{\partial f_i}{\partial x_j}(u) \right| \leq M \quad \forall u \in Q.$$

Dann folgt

$$\|f(b) - f(a)\| \leq n^2 M \|b - a\| \quad \forall a, b \in Q.$$

Inbesondere ist $f|_Q$ für jede C^1 -Funktion f Lipschitzstetig.

Beweis. Seien $a, b \in Q$. Für jede Komponentenfunktion f_i gilt

$$f_i(b) - f_i(a) = \sum_{j=1}^n \left(f_i(b_1, \dots, b_j, a_{j+1}, \dots, a_n) - f_i(b_1, \dots, b_{j-1}, a_j, \dots, a_n) \right).$$

Wenden wir den Mittelwertsatz von Lagrange auf die Funktion f_i in der j -ten Variablen an, so erhalten wir ein ξ_{ij} zwischen a_j und b_j mit

$$f_i(b) - f_i(a) = \sum_{j=1}^n (b_j - a_j) \frac{\partial f_i}{\partial x_j} \underbrace{(b_1, \dots, b_{j-1}, \xi_{ij}, a_{j+1}, \dots, a_n)}_{\in Q}.$$

Daraus folgt

$$|f_i(b) - f_i(a)| \leq \sum_{j=1}^n |b_j - a_j| \cdot M \leq n \cdot M \cdot \|b - a\|$$

und damit

$$\|f(b) - f(a)\| = \left(\sum_{i=1}^n |f_i(b) - f_i(a)|^2 \right)^{\frac{1}{2}} \leq \sqrt{n \cdot n^2 \cdot M^2 \cdot \|b - a\|^2} \leq n^2 M \|b - a\|. \quad \square$$

Satz 6.15 (Satz über den lokalen Diffeomorphismus)

Es sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine C^k -Abbildung von einer offenen Menge $U \subset \mathbb{R}^n$, $k \geq 1$, und $x_0 \in U$ ein Punkt mit $\det(Df(x_0)) \neq 0$ (das heißt $Df(x_0)$ ist ein Isomorphismus). Dann ist f um x_0 ein lokaler C^k -Diffeomorphismus, d.h. es existieren offene Umgebungen \tilde{U} von x_0 und \tilde{V} von $f(x_0)$, so dass $f|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ein C^k -Diffeomorphismus ist.

Beweis. 1. Schritt: Wir zeigen die Behauptung zunächst unter der Zusatzbedingung, dass $Df(x_0) = Id_{\mathbb{R}^n}$. Dann ist die Jacobi-Matrix von f im Punkt x_0 die Einheitsmatrix, d.h. es gilt $\frac{\partial f_i}{\partial x_j}(x_0) = \delta_{ij}$. Wir wählen nun einen Quader $K \subset U$, der x_0 als inneren Punkt enthält, so dass gilt

- (i) $\det(Df(x)) \neq 0 \quad \forall x \in K$,
- (ii) $\left| \frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(x_0) \right| < \frac{1}{2n^2} \quad \forall x \in K, \forall i, j \in \{1, \dots, n\}$.

(1) Im Folgenden zeigen wir, dass f auf K injektiv ist. Wir betrachten dazu die Abbildung $g(x) := f(x) - x$. Dann ist g eine C^1 -Abbildung und es gilt

$$\left| \frac{\partial g_i}{\partial x_j}(x) \right| = \left| \frac{\partial f_i}{\partial x_j}(x) - \delta_{ij} \right| = \left| \frac{\partial f_i}{\partial x_j}(x) - \frac{\partial f_i}{\partial x_j}(x_0) \right| < \frac{1}{2n^2} \quad \forall x \in K.$$

Aus dem Lemma 6.17 folgt dann

$$\|g(b) - g(a)\| < n^2 \cdot \frac{1}{2n^2} \cdot \|b - a\| = \frac{1}{2} \|b - a\| \quad \forall a, b \in K.$$

Mit Hilfe der Dreiecksungleichung erhalten wir daraus

$$\|b - a\| - \|f(b) - f(a)\| \leq \|(f(b) - b) - (f(a) - a)\| = \|g(b) - g(a)\| \leq \frac{1}{2} \|b - a\|$$

und nach Umstellen

$$\|b - a\| \leq 2\|f(b) - f(a)\| \quad \forall a, b \in K. \quad (A)$$

Somit ist f auf K injektiv. Da $x_0 \in \text{Int}(K)$, gilt außerdem $f(x_0) \notin f(\partial K)$.

(2) Als nächstes definieren wir die Umgebung \tilde{V} von $f(x_0)$. Da K kompakt ist, ist auch der Rand ∂K kompakt. Da f stetig ist, ist $f(\partial K)$ ebenfalls kompakt. Der Punkt $f(x_0)$ liegt nicht auf $f(\partial K)$, deshalb hat er einen positiven Abstand zur kompakten Menge $f(\partial K)$, d.h.

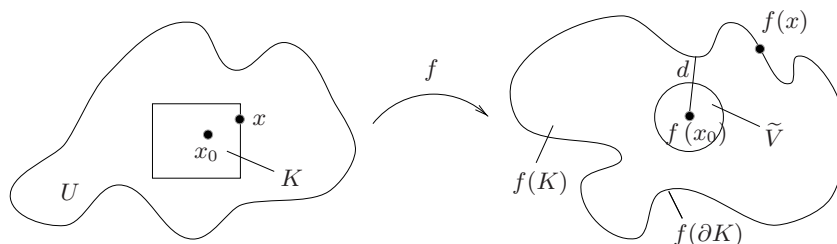
$$\text{dist}(f(x_0), f(\partial K)) := \inf\{\|f(x_0) - f(x)\| \mid x \in \partial K\} =: d > 0.$$

Wir definieren nun die Umgebung \tilde{V} von $f(x_0)$ durch

$$\tilde{V} := K(f(x_0), \frac{d}{2}) = \{y \in \mathbb{R}^n \mid \|y - f(x_0)\| < \frac{d}{2}\}.$$

Damit erhalten wir insbesondere

$$\|y - f(x_0)\| < \|y - f(x)\| \quad \forall y \in \tilde{V}, \forall x \in \partial K. \quad (B)$$



(3) Wir wollen nun die Umgebung \tilde{U} von x_0 definieren. Dazu beweisen wir zunächst, dass für jedes $y \in \tilde{V}$ genau ein $x \in \text{Int}(K)$ existiert mit $f(x) = y$. Die Eindeutigkeit wurde bereits in (1) gezeigt. Wir müssen also nur die Existenz nachweisen. Sei $y \in \tilde{V}$ gegeben. Wir betrachten die Abbildung

$$g_y : K \longrightarrow \mathbb{R} \\ x \longmapsto g_y(x) := \|y - f(x)\|^2 = \langle y - f(x), y - f(x) \rangle.$$

Da K kompakt und g_y stetig ist, nimmt g_y auf K ein absolutes Minimum an. Die Abschätzung (B) zeigt, dass dieses absolute Minimum nicht auf dem Rand von K angenommen wird, sondern im Inneren von K . Das heißt, es existiert ein $u_0 \in \text{Int}(K)$, in dem g_y sein absolutes Minimum annimmt. Wir zeigen nun, dass $f(u_0) = y$ gilt. Da u_0 ein kritischer Punkt von g_y ist, erhalten wir mit der Produktregel für Ableitungen von Skalarprodukten

$$0 = \frac{\partial g_y}{\partial x_j}(u_0) = -2 \left\langle \frac{\partial f}{\partial x_j}(u_0), y - f(u_0) \right\rangle.$$

Damit stehen alle Spaltenvektoren der Jacobi-Matrix von g_y im Punkt u_0 senkrecht auf dem Vektor $y - f(u_0)$. Da nach Wahl des Quaders K die Jacobi-Matrix in allen Punkten

von K regulär ist, bilden ihre Spaltenvektoren eine Basis des \mathbb{R}^n . Folglich ist $y - f(u_0)$ der Nullvektor, also $f(u_0) = y$.

Wir setzen nun $\tilde{U} := \text{Int}(K) \cap f^{-1}(\tilde{V})$. Dann ist \tilde{U} eine offene Umgebung von x_0 in U und $f|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ist bijektiv.

(4) Wir zeigen nun, dass $(f|_{\tilde{U}})^{-1} : \tilde{V} \rightarrow \tilde{U}$ stetig ist. Aus (A) erhalten wir die Ungleichung

$$\|(f|_{\tilde{U}})^{-1}(x) - (f|_{\tilde{U}})^{-1}(y)\| \leq 2\|x - y\| \quad \forall x, y \in \tilde{V}.$$

Folglich ist $(f|_{\tilde{U}})^{-1}$ auf \tilde{V} Lipschitzstetig, also stetig.

(5) Wir zeigen als nächstes, dass $(f|_{\tilde{U}})^{-1} : \tilde{V} \rightarrow \tilde{U}$ differenzierbar ist. Sei dazu $y = f(x) \in \tilde{V}$ mit $x \in \tilde{U}$ und $L_x := Df(x)$. Wir zeigen, dass $(f|_{\tilde{U}})^{-1}$ in y differenzierbar ist. Da f in x differenzierbar ist, gilt für $x + h \in \tilde{U}$

$$f(x + h) - f(x) = L_x(h) + o(\|h\|) \quad \text{für } h \rightarrow 0.$$

Daraus folgt

$$L_x^{-1}(f(x + h) - f(x)) = h + L_x^{-1}(o(\|h\|)).$$

Setzen wir $y_1 := f(x + h)$, so ist $(f|_{\tilde{U}})^{-1}(y_1) - (f|_{\tilde{U}})^{-1}(y) = x + h - x = h$ und es folgt

$$(f|_{\tilde{U}})^{-1}(y_1) - (f|_{\tilde{U}})^{-1}(y) = L_x^{-1}(y_1 - y) - L_x^{-1}(o(\|h\|)).$$

Da L_x^{-1} linear ist, müssen wir für die Differenzierbarkeit von $(f|_{\tilde{U}})^{-1}$ in y nur noch das Verhalten des Restgliedes untersuchen und $-L_x^{-1}(o(\|h\|)) = o(\|y_1 - y\|)$ für $y_1 \rightarrow y$ zeigen. Da L_x^{-1} linear ist, existiert eine Konstante $C \in \mathbb{R}^+$, so dass $\|L_x^{-1}(v)\| \leq C\|v\|$ für alle $v \in \mathbb{R}^n$. Daraus erhalten wir

$$0 \leq \frac{\| -L_x^{-1}(o(\|h\|)) \|}{\|y_1 - y\|} \leq C \cdot \frac{\|o(\|h\|)\|}{\|y_1 - y\|} = C \cdot \frac{\|o(\|h\|)\|}{\|h\|} \cdot \frac{\|h\|}{\|f(x + h) - f(x)\|}$$

Die Formel (A) zeigt uns, dass $\|h\| \leq 2\|f(x + h) - f(x)\|$, folglich gilt

$$0 \leq \frac{\| -L_x^{-1}(o(\|h\|)) \|}{\|y_1 - y\|} \leq 2C \cdot \frac{\|o(\|h\|)\|}{\|h\|}. \quad (C)$$

Da $(f|_{\tilde{U}})^{-1}$ stetig ist, folgt aus $y_1 \rightarrow y$, dass $h \rightarrow 0$ und die Abschätzung (C) liefert

$$\lim_{y_1 \rightarrow y} \frac{\| -L_x^{-1}(o(\|h\|)) \|}{\|y_1 - y\|} = 0.$$

Somit ist $(f|_{\tilde{U}})^{-1}$ in y differenzierbar und es gilt $Df^{-1}(f(x)) = (Df(x))^{-1}$.

(6) Als letztes zeigen wir, dass aus $f \in C^k$ auch $(f|_{\tilde{U}})^{-1} \in C^k$ folgt. Bezeichne zur Abkürzung $g := (f|_{\tilde{U}})^{-1}$. Da die Jacobi-Matrix von g in $y \in \tilde{V}$ die inverse Matrix der Jacobi-Matrix von f in $x = g(y)$ ist, zeigt uns die Vorschrift zur Berechnung inverser Matrizen, dass die partiellen Ableitungen von g in y rationale Funktionen der partiellen Ableitungen von f in x sind, d.h. es existieren rationale Funktionen R_{kl} in n^2 Variablen, so dass

$$\frac{\partial g_k}{\partial y_l}(y) = R_{kl} \left(\frac{\partial f_i}{\partial x_j}(g(y)), \quad i, j = 1, \dots, n \right). \quad (D)$$

Ist f eine C^1 -Funktion, so sind die partiellen Ableitungen von f stetig, g ist ebenfalls stetig, somit sind die partiellen Ableitungen von g ebenfalls stetig. Successives Ableiten der Formel (D) zeigt, dass g k -fach stetig differenzierbar ist, wenn f k -fach stetig differenzierbar ist.

Damit ist die Behauptung des Satzes für den Fall bewiesen, dass $Df(x_0) = Id_{\mathbb{R}^n}$ gilt.

2. *Schritt:* Wir zeigen nun die Behauptung des Satzes im allgemeinen Fall. Sei f eine beliebige Abbildung mit den Voraussetzungen des Satzes. Wir setzen $L := Df(x_0)^{-1}$ und betrachten die Abbildung $\tilde{f} := L \circ f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$. Dann ist \tilde{f} eine C^k -Abbildung und es gilt nach Kettenregel

$$D\tilde{f}(x_0) = DL(f(x_0)) \circ Df(x_0) = L \circ L^{-1} = Id_{\mathbb{R}^n}.$$

Wir können also auf \tilde{f} unsere Erkenntnisse aus dem 1. Schritt anwenden und erhalten Umgebungen \tilde{U} von x_0 und \tilde{V} von $\tilde{f}(x_0)$, so dass $\tilde{f}|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}$ ein C^k -Diffeomorphismus ist. Dann ist $f|_{\tilde{U}} := L^{-1} \circ \tilde{f}|_{\tilde{U}} : \tilde{U} \rightarrow \tilde{V}' := L^{-1}(\tilde{V})$ ebenfalls ein C^k -Diffeomorphismus. \square

Satz 6.16 (Globaler Umkehrsatz)

Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine C^k -Abbildung von einer offenen Menge U und sei für alle $x \in U$ das Differential $Df(x)$ 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 C^k -Diffeomorphismus.

Beweis. (1) Nach Satz 6.15 existiert für alle $x \in U$ eine Umgebung U_x von x und V_x von $f(x)$, so dass $f|_{U_x} : U_x \rightarrow V_x$ ein C^k -Diffeomorphismus ist. Dann gilt $U = \bigcup_{x \in U} U_x$ und

$$f(U) = \bigcup_{x \in U} f|_{U_x}(U_x) = \bigcup_{x \in U} V_x =: V.$$

Die Menge V ist insbesondere offen.

(2) Ist f injektiv, so ist $f : U \rightarrow V = f(U)$ bijektiv. Da $f|_{U_x} : U_x \rightarrow V_x$ ein C^k -Diffeomorphismus für alle $U_x \subset U$ ist und Differenzierbarkeit eine lokale Eigenschaft ist, ist $f : U \rightarrow V = f(U)$ ein (globaler) C^k -Diffeomorphismus. \square

6.6 Der Satz über implizite Funktionen

Sei $f = (f_1, \dots, f_m) : U \subset \mathbb{R}^N \rightarrow \mathbb{R}^m$ eine Funktion mit m -Komponenten, die von N Variablen abhängt, wobei $m \leq N$. Wir betrachten 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: $f(x) = 0$. Dann stellen sich folgende Fragen:

- Hat dieses Gleichungssystem eine Lösung?
- Von wievielen Parametern hängt die Lösungsmenge ab? Wie ist die Abhängigkeit (stetig, differenzierbar, C^k)?
- Kann man die Lösungsmenge explizit beschreiben?

Um diese Fragen zu beantworten, versucht man, die Gleichungen nach einer oder mehreren Variablen aufzulösen.

In der Vorlesung über Lineare Algebra haben Sie sich mit dieser Frage für den Spezialfall *linearer* Gleichungssysteme beschäftigt. In diesem Fall ist f durch eine $(m \times N)$ -Matrix $A = (A_{ki})$ und einen Vektor $b = (b_1, \dots, b_m)^t$ definiert:

$$f_k(x_1, \dots, x_N) := A_{k1}x_1 + A_{k2}x_2 + \dots + A_{kN}x_N - b_k = 0 \quad \text{für } k = 1, \dots, m.$$

In diesem Spezialfall gilt das folgende: Ist A eine Matrix von maximalem Rang, d.h. vom Rang m , und seien oBdA. die letzten m Spalten von A linear unabhängig. Dann kann man das lineare Gleichungssystem $Ax = b$ nach den letzten m Variablen auflösen, d.h. es existieren *eindeutig bestimmte* Funktionen $\varphi = (\varphi_1, \dots, \varphi_m) : \mathbb{R}^{N-m} \rightarrow \mathbb{R}^m$, so dass

$$Ax = b \iff x = \left(x_1, \dots, x_{N-m}, \varphi_1(x_1, \dots, x_{N-m}), \dots, \varphi_m(x_1, \dots, x_{N-m}) \right)^t.$$

Mit den Methoden der Differentialrechnung können wir nun auch Aussagen über die Auflösbarkeit von *nicht-linearen* Gleichungssystemen machen. Wir möchten auch in diesem Fall Funktionen $\varphi : W \subset \mathbb{R}^{N-m} \rightarrow \mathbb{R}^m$ finden, so dass

$$f(x) = 0 \iff x = (z, \varphi(z)), \quad z \in W.$$

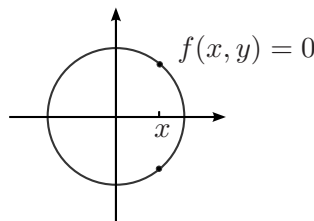
Mit anderen Worten, wir wollen die Lösungsmenge von $f(x) = 0$ explizit als Graph einer Funktion φ beschreiben. Zunächst sehen wir uns drei Beispiele an, die Unterschiede zum linearen Fall zeigen.

Beispiel 1: Wir betrachten die Abbildung $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) := x^2 + y^2 - 1.$$

Die Lösungsmenge dieser Gleichung ist die Kreislinie S^1 . 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 nur einen Teil der Lösungsmenge.

Beispiel 2: Wir betrachten die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) = x - y^3.$$

Man kann die Gleichung $x - y^3 = 0$ nach y auflösen und erhält $y = \sqrt[3]{x} =: \varphi(x)$. Die Lösungsmenge der Gleichung $x - y^3 = 0$ ist der Graph der Funktion φ . Diese Funktion ist im Punkt $x = 0$ nicht differenzierbar.

Beispiel 3: Wir betrachten die Funktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ mit

$$f(x, y) = x^2.$$

Die Lösungsmenge der Gleichung $f(x, y) = x^2 = 0$ ist die y -Achse. Man kann diese Gleichung nicht nach y auflösen, es gibt aber eine Auflösung nach x : Der Graph von $x = \psi(y) := 0$ beschreibt die y -Achse.

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 hat aber Formeln für ihre Ableitungen.

Satz 6.17 (Satz über implizite Funktionen)

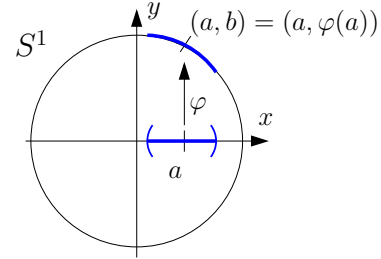
Sei $U \subset \mathbb{R}^n \times \mathbb{R}^m$ offen und $f : U \rightarrow \mathbb{R}^m$ eine C^1 -Funktion. Wir betrachten einen Punkt $(a, b) \in U$ mit

1. $f(a, b) = 0$, d.h. (a, b) löst die Gleichung $f(x, y) = 0$, und
2. Die letzten m Spalten der Jacobi-Matrix von f im Punkt (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.



Darüber hinaus gilt:

4. $\varphi : A(a) \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ ist differenzierbar.
5. Falls $f \in C^k$, so ist auch $\varphi \in C^k$.
6. Für die partiellen Ableitungen der Funktion φ gilt:

$$\underbrace{\left(\frac{\partial \varphi_l}{\partial x_i}(x)\right)}_{\in \mathbb{R}^{m \times n}} = - \underbrace{\left(\frac{\partial f_k}{\partial y_l}(x, \varphi(x))\right)^{-1}}_{\in \mathbb{R}^{m \times m}} \circ \underbrace{\left(\frac{\partial f_k}{\partial x_i}(x, \varphi(x))\right)}_{\in \mathbb{R}^{m \times n}} \quad \forall x \in A(a).$$

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

$$F(x, y) := (x, f(x, y)).$$

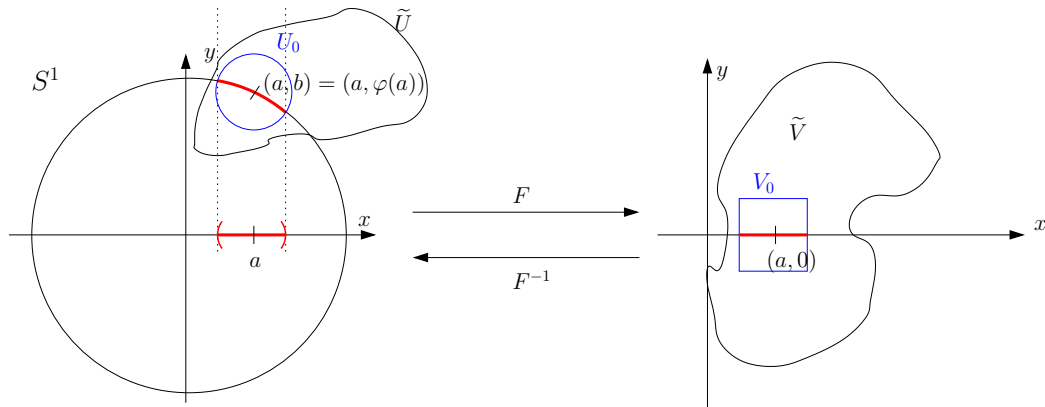
F ist eine C^1 -Funktion und nach Voraussetzung gilt:

$$\det DF(a, b) = \det \begin{pmatrix} E_n & & 0 \\ & \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_m} \\ * & \vdots & & \vdots \\ & \frac{\partial f_m}{\partial y_1} & \cdots & \frac{\partial f_m}{\partial y_m} \end{pmatrix} (a, b) \neq 0.$$

Nach Satz 6.15 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 C^1 -Diffeomorphismus ist. Da \tilde{V} offen ist, können wir eine offene Umgebung $A(a) \subset \mathbb{R}^n$ von a und einen 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 C^1 -Diffeomorphismus. Wir betrachten die inverse Abbildung $(F|_{U_0})^{-1} : V_0 \rightarrow U_0$ mit den Komponenten

$$(F|_{U_0})^{-1} =: (F_1^-, F_2^-).$$

$\begin{matrix} \uparrow & \uparrow \\ \mathbb{R}^n & \mathbb{R}^m \end{matrix}$



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

$$(F|_{U_0})^{-1}(A(a) \times \{0\}) = f^{-1}(0) \cap U_0.$$

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. $f^{-1}(0) \cap U_0 = (F|_{U_0})^{-1}(A(a) \times \{0\}) = \{(x, F_2^-(x, 0)) \mid x \in A(a)\} = \{(x, \varphi(x)) \mid x \in A(a)\} = \text{graph}(\varphi)$.

Die Eindeutigkeit der Funktion φ mit den Eigenschaften 1) - 3) folgt aus der Konstruktion.

2) *Differenzierbarkeit von φ :*

Wenn f eine C^k -Funktion ist, dann ist nach Definition auch F eine C^k -Funktion. Nach dem Satz über den lokalen Diffeomorphismus ist die inverse Funktion $(F|_{U_0})^{-1}$ ebenfalls C^k , das gleiche gilt dann auch für ihre Komponente F_2^- . Da $\varphi(x) = F_2^-(x, 0)$, ist auch φ eine C^k -Funktion. Nach Voraussetzung ist f mindestens C^1 . Somit ist φ auf jeden Fall eine C^1 -Funktion, also insbesondere differenzierbar (siehe Satz 6.9).

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

$$- \left(\frac{\partial f_k}{\partial x_i}(x, \varphi(x)) \right) = \left(\frac{\partial f_k}{\partial y_l}(x, \varphi(x)) \right) \circ \left(\frac{\partial \varphi_l}{\partial x_i}(x) \right). \quad (*)$$

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

$$\left(\frac{\partial \varphi_l}{\partial x_i}(x) \right) = - \left(\frac{\partial f_k}{\partial y_l}(x, \varphi(x)) \right)^{-1} \circ \left(\frac{\partial f_k}{\partial x_i}(x, \varphi(x)) \right).$$

□

6.7 Untermannigfaltigkeiten des \mathbb{R}^N und ihre Tangentialräume

In diesem Abschnitt wollen wir den Satz über implizite Funktionen benutzen, um eine Formel für die Tangentialebenen an gleichungsdefinierte Teilmengen des \mathbb{R}^N herzuleiten.

Definition 6.18. Sei $\phi : U \subset \mathbb{R}^N \rightarrow \mathbb{R}^m$ eine C^k -Funktion und $N \geq m$.

1. Ein Punkt $p \in U$ heißt *regulärer Punkt* von ϕ , wenn das Differential von ϕ im Punkt p , d.h. $D\phi(p) : \mathbb{R}^N \rightarrow \mathbb{R}^m$, surjektiv ist.
2. Ein Punkt $c \in \phi(U) \subset \mathbb{R}^m$ heißt *regulärer Wert* von ϕ , wenn das Urbild $\phi^{-1}(c) \subset U$ nur aus regulären Punkten von ϕ besteht.
3. Ist $c \in \phi(U) \subset \mathbb{R}^m$ ein regulärer Wert von ϕ , so nennt man die Niveaumenge

$$M := \{x \in U \mid \phi(x) = c\} \subset \mathbb{R}^N$$

gleichungsdefinierte (C^k) -Untermannigfaltigkeit des \mathbb{R}^N der Dimension $(N - m)$. Eine 1-dimensionale Untermannigfaltigkeit nennen wir auch kurz *Kurve*, eine 2-dimensionale kurz *Fläche*.

Bemerkung:

- Ist $m = 1$, also ϕ reellwertig, so ist $D\phi(p)$ genau dann surjektiv, wenn $\text{grad}\phi(p) \neq 0$. Unsere Definition stimmt also mit der in Abschnitt 6.1. für reellwertige Funktionen gegebenen überein.

- Das Differential $D\phi(p) : \mathbb{R}^N \rightarrow \mathbb{R}^m$ ist genau dann surjektiv, wenn die Jacobi-Matrix $\left(\frac{\partial\phi_k}{\partial x_i}(p)\right)$ maximalen Rang hat, also m linear unabhängige Spalten besitzt. Wir können in diesem Fall den Satz über implizite Funktionen auf den Punkt $p \in M$ anwenden. Somit besitzt jeder Punkt p einer $(N - m)$ -dimensionalen Untermannigfaltigkeit M eine Umgebung $U(p) \subset M$, die der Graph einer C^k -Funktion $\varphi_p : A_p \subset \mathbb{R}^{N-m} \rightarrow \mathbb{R}^m$ ist. Man kann $U(p)$ also durch $(N - m)$ reelle Parameter aus A_p eindeutig beschreiben. (Dies ist der Grund dafür, der Menge M die Dimension $(N - m)$ zuzuordnen).

Wir erinnern nochmal an die Definition des Tangentialraums und der Tangentialebene an eine Untermannigfaltigkeit aus Kapitel 6.1:

Definition 6.19. Sei $M \subset \mathbb{R}^N$ eine gleichungsdefinierte Untermannigfaltigkeit des \mathbb{R}^N . Der Tangentialraum an M im Punkt $p \in M$ ist die Menge der Vektoren

$$T_p M := \{v \in \mathbb{R}^N \mid \exists \text{ diffb. Kurve } \gamma : (-\varepsilon, \varepsilon) \rightarrow M \text{ mit } \gamma(0) = p \text{ und } \gamma'(0) = v.\}$$

Die Tangentialebene an M im Punkt $p \in M$ ist die Ebene

$$\text{Tan}_p M = p + T_p M.$$

Der folgende Satz rechtfertigt diese Bezeichnungen. Er zeigt, dass $T_p M$ tatsächlich ein $(N - m)$ -dimensionaler Unterraum des Vektorraumes \mathbb{R}^N ist, und $\text{Tan}_p M$ somit die durch p und $T_p M$ aufgespannte Ebene.

Satz 6.18 Sei $\phi = (\phi_1, \dots, \phi_m) : U \subset \mathbb{R}^N \rightarrow \mathbb{R}^m$ eine C^k -Funktion, $c \in \phi(U)$ ein regulärer Wert von ϕ und $M := \{x \in U \mid \phi(x) = c\} \subset \mathbb{R}^N$ eine gleichungsdefinierte Untermannigfaltigkeit. Dann gilt für den Tangentialraum an M im Punkt $p \in M$

$$\begin{aligned} T_p M &= \text{Ker } D\phi(p) \\ &= \{v \in \mathbb{R}^N \mid D\phi(p)(v) = 0\} \\ &= \{v \in \mathbb{R}^N \mid \langle \text{grad}\phi_1(p), v \rangle = \dots = \langle \text{grad}\phi_m(p), v \rangle = 0\}. \end{aligned}$$

Ist ϕ reellwertig ($m=1$), so folgt insbesondere

$$T_p M = (\text{grad}\phi(p))^\perp.$$

Beweis. 1) Wir zeigen: $T_p M \subset \text{Ker } D\phi(p)$:

Sei $v \in T_p M$. Dann existiert eine differenzierbare Kurve $\gamma : (-\varepsilon, \varepsilon) \rightarrow M$ mit $\gamma(0) = p$ und $\gamma'(0) = v$. Da das Bild von γ auf M liegt, gilt $\phi(\gamma(t)) = c$ für alle $t \in (-\varepsilon, \varepsilon)$. Dann folgt aus der Kettenregel $0 = \frac{d}{dt}\phi(\gamma(t)) = D\phi(\gamma(t))(\gamma'(t))$. Für $t = 0$ erhalten wir $D\phi(p)(v) = 0$, also liegt v im Kern von $D\phi(p)$.

2) Wir zeigen: $\text{Ker } D\phi(p) \subset T_p M$:

Nach Voraussetzung ist $D\phi(p) : \mathbb{R}^N \rightarrow \mathbb{R}^m$ surjektiv. Wir können oBdA annehmen, dass die letzten m -Spalten der Jacobi-Matrix von ϕ im Punkt p linear unabhängig sind (anderenfalls sortieren wir die Variablen des \mathbb{R}^N um). Wir bezeichnen die Variablen des \mathbb{R}^N

mit (u_1, \dots, u_N) . Wollen wir die ersten $(N-m)$ von den letzten m -Variablen unterscheiden, so benutzen wir wie im Satz über implizite Funktionen die Bezeichnungen

$$\underbrace{(u_1, \dots, u_{N-m})}_{=:(x_1, \dots, x_{N-m})}, \underbrace{(u_{N-m+1}, \dots, u_N)}_{=:(y_1, \dots, y_m)}.$$

Sei $p = (p', p'') \in \mathbb{R}^{N-m} \times \mathbb{R}^m$. Nach dem Satz über implizite Funktionen gibt es offene Umgebungen $U(p) \subset M$ von p und $A(p') \subset \mathbb{R}^{N-m}$ von p' , sowie eine C^k -Funktion $\varphi : A(p') \rightarrow \mathbb{R}^m$ so dass $\text{graph}(\varphi) = U(p)$. Wir betrachten nun einen Vektor

$$v = \underbrace{(v_1, \dots, v_{N-m})}_{=:v'} \underbrace{(v_{N-m+1}, \dots, v_N)}_{=:v''} \in \mathbb{R}^{N-m} \times \mathbb{R}^m.$$

Dann gilt:

$$\begin{aligned} D\phi(p)(v) = 0 &\iff \left(\frac{\partial \phi_k}{\partial u_j}(p) \right) v^t = 0 \\ &\iff \sum_{j=1}^N \frac{\partial \phi_k}{\partial u_j}(p) v_j = 0 \quad \forall k = 1, \dots, m \\ &\iff \sum_{i=1}^{N-m} \frac{\partial \phi_k}{\partial x_i}(p) v_i = - \sum_{l=1}^m \frac{\partial \phi_k}{\partial y_l}(p) v_{N-m+l} \quad \forall k = 1, \dots, m \\ &\iff \left(\frac{\partial \phi_k}{\partial x_i}(p) \right) (v')^t = - \left(\frac{\partial \phi_k}{\partial y_l}(p) \right) (v'')^t \\ &\iff - \left(\frac{\partial \phi_k}{\partial y_l}(p) \right)^{-1} \circ \left(\frac{\partial \phi_k}{\partial x_i}(p) \right) (v')^t = (v'')^t \\ &\stackrel{\text{Satz 6.17}}{\iff} \left(\frac{\partial \varphi_k}{\partial x_i}(p') \right) (v')^t = (v'')^t \\ &\iff D\varphi(p')(v') = v''. \end{aligned}$$

Für einen Vektor $v = (v', v'') \in \text{Ker } D\phi(p)$ betrachten wir nun das Geradenstück im Definitionsbereich von φ , das in Richtung v' durch p' geht. Dieses ist durch $\sigma(t) := p' + tv'$ für $t \in (-\varepsilon, \varepsilon)$, ε hinreichend klein, parametrisiert. Dann ist die Kurve $\gamma : (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}^N$,

$$\gamma(t) := (\sigma(t), \varphi(\sigma(t))),$$

differenzierbar, ihre Werte liegen auf M und es gilt

$$\begin{aligned} \gamma(0) &= (\sigma(0), \varphi(\sigma(0))) = (p', p'') = p, \\ \gamma'(0) &= (\sigma'(0), D\varphi(p')(\sigma'(0))) = (v', D\varphi(p')(v')) = (v', v'') = v. \end{aligned}$$

Folglich gilt $v \in T_p M$. □

Beispiel: *Das einschalige Hyperboloid*

Wir betrachten das einschalige Hyperboloid¹

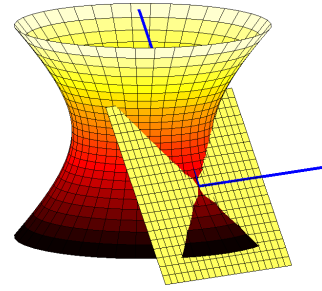
$$M := \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 = z^2 + 1\}.$$

M ist die Nullstellenmenge der Funktion $\phi : \mathbb{R}^3 \rightarrow \mathbb{R}$ mit

$$\phi(x, y, z) = x^2 + y^2 - z^2 - 1.$$

Da $\text{grad}\phi(x, y, z) = (2x, 2y, -2z) \neq 0$ für alle Punkte $p = (x, y, z) \in M$, ist M eine 2-dimensionale Untermannigfaltigkeit im \mathbb{R}^3 . Für den Tangentialraum in einem Punkt $p = (p_1, p_2, p_3) \in M$ erhalten wir:

$$T_p M = \{v \in \mathbb{R}^3 \mid \langle \text{grad}\phi(p), v \rangle = 0\} = \{(v_1, v_2, v_3) \in \mathbb{R}^3 \mid p_1 v_1 + p_2 v_2 = p_3 v_3\}.$$



Der Beweis von Satz 6.18 zeigt, dass man den Tangentialraum einer gleichungsdefinierten Untermannigfaltigkeit auch mit Hilfe der sie lokal explizit als Graph beschreibenden Funktion φ ausdrücken kann. Wir formulieren dies abschließend nochmal als Satz:

Satz 6.19 *Sei $\varphi : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ eine C^k -Funktion, A offen, und bezeichne*

$$M := \text{graph}(\varphi) = \{(x, \varphi(x)) \mid x \in A\} \subset \mathbb{R}^{n+m}$$

den Graphen der Funktion φ . Dann ist M eine n -dimensionale (gleichungsdefinierte) Untermannigfaltigkeit des \mathbb{R}^{n+m} und für den Tangentialraum gilt

$$\begin{aligned} T_{(x, \varphi(x))} M &= \left\{ \underbrace{(v_1, \dots, v_n)}_{=:v}, D\varphi(x)(v) \mid v \in \mathbb{R}^n \right\} \\ &= \text{span} \left\{ \left(1, 0, \dots, 0, \frac{\partial \varphi}{\partial x_1}(x) \right), \left(0, 1, \dots, 0, \frac{\partial \varphi}{\partial x_2}(x) \right), \dots, \left(0, \dots, 0, 1, \frac{\partial \varphi}{\partial x_n}(x) \right) \right\}. \end{aligned}$$

Beweis. Wir betrachten die Abbildung $\phi : A \times \mathbb{R}^m \rightarrow \mathbb{R}^m$, gegeben durch

$$\phi(x, y) := \varphi(x) - y.$$

ϕ ist von der Klasse C^k und es gilt

$$\phi^{-1}(0) = \{(x, \varphi(x)) \mid x \in A\} = \text{graph}(\varphi) = M.$$

Für die Jacobi-Matrix von ϕ im Punkt $(x, \varphi(x))$ gilt

$$\left(\frac{\partial \phi_k}{\partial u_j}(x, \varphi(x)) \right) = \left(\frac{\partial \varphi_k}{\partial x_i}(x) \quad - E_m \right).$$

Folglich ist $D\phi(x, \varphi(x))$ surjektiv für alle $p = (x, \varphi(x)) \in M$. Somit ist M eine gleichungsdefinierte Untermannigfaltigkeit. Aus dem Beweis von Satz 6.18 folgt für den Tangentialraum in $p = (x, \varphi(x)) \in M$

$$T_p M = \left\{ \underbrace{(v_1, \dots, v_n)}_{=:v}, D\varphi(x)(v) \mid v \in \mathbb{R}^n \right\}$$

□

¹ Vielen Dank an Christoph Stadtmüller und Thomas Neukirchner für die Bilder in §6.6 und §6.7.

Beispiel: *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

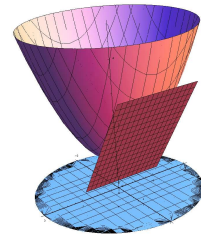
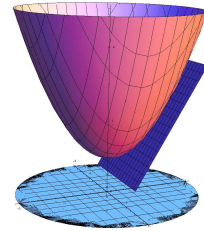
$$\varphi(x, y) := x^2 + y^2 + 1.$$

Für die Jacobi-Matrix von φ gilt:

$$\left(\frac{\partial \varphi}{\partial x}(x, y) \quad \frac{\partial \varphi}{\partial y}(x, y) \right) = (2x \quad 2y).$$

Für den Tangentialraum im Punkt $p = (x, y, \varphi(x, y)) \in M$ erhalten wir

$$\begin{aligned} T_p M &= \{(v_1, v_2, 2xv_1 + 2yv_2) \mid (v_1, v_2) \in \mathbb{R}^2\} \\ &= \text{span}\{(1, 0, 2x), (0, 1, 2y)\}. \end{aligned}$$



6.8 Extrema unter Nebenbedingungen

In Abschnitt 6.4 hatten wir Bedingungen für das Vorliegen lokaler Extrema einer C^2 -Funktion $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ hergeleitet. Oft benötigt man das Extremwertverhalten einer Funktion, wenn die Variablen einer zusätzlichen Nebenbedingung unterworfen sind. Sei zum Beispiel $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ eine Temperaturverteilung. Wir wollen dann z.B. wissen, was die maximale Temperatur auf der Oberfläche eines Gegenstandes im \mathbb{R}^3 ist, und wo dieses Maximum angenommen wird. Auch solche Fragen kann man in speziellen Fällen mit den Methoden der Differentialrechnung behandeln.

Wir betrachten hier den Fall, dass die Nebenbedingungen in der Form eines Gleichungssystems

$$\begin{aligned} g_1(x_1, \dots, x_n) &= 0, \\ g_2(x_1, \dots, x_n) &= 0, \\ &\vdots \\ g_m(x_1, \dots, x_n) &= 0. \end{aligned}$$

gegeben sind, d.h. als Nullstellenmenge einer Funktion $g := (g_1, \dots, g_m) : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$.

Definition 6.20. *Man sagt, f nimmt in einem Punkt $p \in \mathcal{N}_g := \{x \in U \mid g(x) = 0\}$ ein lokales Maximum (lokales Minimum) unter der Nebenbedingung $g(x) = 0$ an, falls ein $\varepsilon > 0$ existiert, so dass*

$$f(x) \leq f(p) \quad (f(x) \geq f(p)) \quad \forall x \in K(p, \varepsilon) \cap \mathcal{N}_g.$$

Wir erinnern nochmal an folgendes Kriterium für lokale Extrema ohne Nebenbedingungen²: Hat eine C^1 -Funktion $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ein lokales Extremum in $p \in U$, so ist $\text{grad}f(p) = 0$. Der folgende Satz verallgemeinert dieses Kriterium auf den Fall von Extrema mit Nebenbedingungen und ist insbesondere anwendbar, wenn die Nullstellenmenge $\mathcal{N}_g = g^{-1}(0)$ eine $(n - m)$ -dimensionale Untermannigfaltigkeit des \mathbb{R}^n ist.

Satz 6.20 (Notwendige Bedingungen für Extrema unter Nebenbedingungen)

Seien $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ und $g : U \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, $m \leq n$, C^1 -Funktionen und $p \in U$ ein Punkt mit $g(p) = 0$ und

$$\text{rang}(Dg(p)) = \text{rang}\left(\frac{\partial g_k}{\partial x_i}(p)\right) = m.$$

Hat f in p ein lokales Extremum unter der Nebenbedingung $g(x) = 0$, dann existieren eindeutig bestimmte Konstanten $\lambda_1, \dots, \lambda_m \in \mathbb{R}$, so dass

$$\text{grad}f(p) = \lambda_1 \cdot \text{grad}g_1(p) + \dots + \lambda_m \cdot \text{grad}g_m(p). \quad (1)$$

Bezeichnung: Die Zahlen $\lambda_1, \dots, \lambda_m$ heißen auch *Lagrange-Multiplikatoren*.

Beweis von Satz 6.20. Da $\text{rang}(Dg(p)) = m$, können wir OBdA voraussetzen, dass die ersten m Zeilen der Jacobi-Matrix von g im Punkt p linear unabhängig sind:

$$\det \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(p) & \dots & \frac{\partial g_1}{\partial x_m}(p) \\ \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1}(p) & \dots & \frac{\partial g_m}{\partial x_m}(p) \end{pmatrix} \neq 0,$$

(Anderenfalls ordnen wir die (x_1, \dots, x_n) entsprechend um). Dann hat das lineare Gleichungssystem

$$\sum_{v=1}^m \lambda_v \cdot \frac{\partial g_v}{\partial x_i}(p) = \frac{\partial f}{\partial x_i}(p), \quad i = 1, \dots, m, \quad (2)$$

genau eine Lösung $(\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$. Dies ist *ein Teil* der gesuchten Gleichung (1), er enthält die ersten m der insgesamt n Komponenten von $\text{grad}f$. Es bleibt zu zeigen, dass (1) auch für die Komponenten $i = m + 1, \dots, n$ erfüllt ist.

Wir bezeichnen $p =: (\hat{p}, p')$ mit $\hat{p} = (p_1, \dots, p_m)$ und $p' = (p_{m+1}, \dots, p_n)$ und analog $x =: (\hat{x}, x')$ mit $\hat{x} = (x_1, \dots, x_m)$ und $x' = (x_{m+1}, \dots, x_n)$. Nach dem Satz über implizite Funktionen kann man

$$g(\underbrace{x_1, \dots, x_m}_{\hat{x}}, \underbrace{x_{m+1}, \dots, x_n}_{x'}) = 0$$

in einer Umgebung $A(p') \subset \mathbb{R}^{n-m}$ nach x_1, \dots, x_m auflösen, d.h., es existieren C^1 -Funktionen

$$\varphi = (\varphi_1, \dots, \varphi_m) : A(p') \rightarrow \mathbb{R}^m$$

² Zur Erinnerung: Der Definitionsbereich U von f bezeichnet in diesem Kapitel immer eine *offene* Menge.

so dass $\varphi(p') = \hat{p}$ und

$$g(\varphi_1(x_{m+1}, \dots, x_n), \dots, \varphi_m(x_{m+1}, \dots, x_n), x_{m+1}, \dots, x_n) = 0.$$

Wir leiten diese Gleichung nach den Variablen x_μ für $\mu = m+1, \dots, n$ ab und erhalten mit der Kettenregel für partielle Ableitungen:

$$0 = \sum_{i=1}^m \frac{\partial g_v}{\partial x_i}(p) \cdot \frac{\partial \varphi_i}{\partial x_\mu}(p') + \frac{\partial g_v}{\partial x_\mu}(p) \quad \forall \nu = 1, \dots, m,$$

bzw.

$$\frac{\partial g_v}{\partial x_\mu}(p) = - \sum_{i=1}^m \frac{\partial g_v}{\partial x_i}(p) \cdot \frac{\partial \varphi_i}{\partial x_\mu}(p') \quad \forall \nu = 1, \dots, m. \quad (3)$$

Wir betrachten nun die folgende C^1 -Funktion $\psi : A(p') \rightarrow \mathbb{R}$:

$$\psi(x_{m+1}, \dots, x_n) := f(\varphi_1(x'), \dots, \varphi_m(x'), \underbrace{x_{m+1}, \dots, x_n}_{x'}).$$

Nach Voraussetzung hat f in $p = (\varphi(p'), p')$ einen lokalen Extremwert unter der Nebenbedingung $g(x) = 0$, daher hat ψ in p' ein lokales Extremum. Folglich ist $\text{grad}\psi(p') = 0$ und mit der Kettenregel folgt

$$0 = \frac{\partial \psi}{\partial x_\mu}(p') = \sum_{i=1}^m \frac{\partial f}{\partial x_i}(p) \cdot \frac{\partial \varphi_i}{\partial x_\mu}(p') + \frac{\partial f}{\partial x_\mu}(p) = 0 \quad \forall \mu = m+1, \dots, n$$

bzw.

$$\frac{\partial f}{\partial x_\mu}(p) = - \sum_{i=1}^m \frac{\partial f}{\partial x_i}(p) \cdot \frac{\partial \varphi_i}{\partial x_\mu}(p'), \quad \forall \mu = m+1, \dots, n. \quad (4)$$

Setzen wir (2) und (3) nacheinander in (4) ein, so ergibt sich

$$\begin{aligned} \frac{\partial f}{\partial x_\mu}(p) &\stackrel{(2)}{=} - \sum_{i=1}^m \sum_{v=1}^m \lambda_v \cdot \frac{\partial g_v}{\partial x_i}(p) \cdot \frac{\partial \varphi_i}{\partial x_\mu}(p') \\ &\stackrel{(3)}{=} \sum_{v=1}^m \lambda_v \cdot \frac{\partial g_v}{\partial x_\mu}(p) \quad \forall \mu = m+1, \dots, n. \end{aligned} \quad (5)$$

Die Gleichungen (2) und (5) liefern

$$\text{grad}f(p) = \sum_{v=1}^m \lambda_v \cdot \text{grad}g_v(p).$$

□

Wir setzen nun voraus, dass die Nullstellenmenge $\mathcal{N}_g = g^{-1}(0)$ eine $n - m$ -dimensionale Untermannigfaltigkeit des \mathbb{R}^n ist. Dann ist die Voraussetzung an g in Satz 6.20 in jedem Punkt $p \in \mathcal{N}_g$ erfüllt. Zur Bestimmung der lokalen Extrema von $f|_{\mathcal{N}_g} : \mathcal{N}_g \rightarrow \mathbb{R}$ können wir folgendermaßen vorgehen:

Wir definieren die Funktion $F : U \times \mathbb{R}^m \rightarrow \mathbb{R}$ durch

$$F(x_1, \dots, x_n, c_1, \dots, c_m) := f(x) - \sum_{\nu=1}^m c_\nu g_\nu(x)$$

und betrachten ihre kritischen Punkte (p, λ) . Dann gilt $\text{grad}F(p, \lambda) = 0$, d.h. die kritischen Punkte von F erfüllen das sogenannte *Lagrangesche Gleichungssystem*

$$\begin{aligned} \frac{\partial F}{\partial x_i}(p, \lambda) &= 0 & \forall i = 1, \dots, n, \\ \frac{\partial F}{\partial c_\nu}(p, \lambda) &= 0, & \forall \nu = 1, \dots, m. \end{aligned}$$

Offensichtlich sind die kritischen Punkte (p, λ) von F gerade diejenigen Punkte $p \in \mathcal{N}_g$, in denen die notwendige Bedingung (1) aus Satz 6.20 erfüllt ist. Wir können also Satz 6.20 folgendermaßen umformulieren:

Folgerung 6.1 Sei $f : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ eine C^1 -Funktion und $\mathcal{N}_g \subset U$ eine $(n - m)$ -dimensionale gleichungsdefinierte Untermannigfaltigkeit, wobei $g = (g_1, \dots, g_m)$. Wir bezeichnen mit $F : U \times \mathbb{R}^m \rightarrow \mathbb{R}$ die Funktion

$$F(x_1, \dots, x_n, c_1, \dots, c_m) := f(x) - \sum_{\nu=1}^m c_\nu g_\nu(x).$$

Dann gilt: Hat $f|_{\mathcal{N}_g} : \mathcal{N}_g \rightarrow \mathbb{R}$ in $p \in \mathcal{N}_g$ einen lokalen Extremwert, so existiert ein (eindeutig bestimmtes) $\lambda = (\lambda_1, \dots, \lambda_m) \in \mathbb{R}^m$, so dass (p, λ) kritischer Punkt von F ist, d.h. es gilt das Lagrangesche Gleichungssystem

$$\frac{\partial F}{\partial x_i}(p, \lambda) = 0 \quad \text{und} \quad \frac{\partial F}{\partial c_\nu}(p, \lambda) = 0 \quad \forall i = 1, \dots, n, \nu = 1, \dots, m.$$

Man findet also die lokalen Extrema von $f|_{\mathcal{N}_g}$ unter den kritischen Punkten von F , d.h. unter den Lösungen des Lagrangeschen Gleichungssystems. Natürlich liefert nicht jeder kritische Punkt (p, λ) von F ein lokales Extremum p von $f|_{\mathcal{N}_g}$. Man muss für jeden kritischen Punkt (p, λ) extra untersuchen, ob p tatsächlich ein lokales Extremum von $f|_{\mathcal{N}_g}$ ist. Ist $\mathcal{N}_g \subset \mathbb{R}^n$ kompakt, so existieren $\max f|_{\mathcal{N}_g}$ und $\min f|_{\mathcal{N}_g}$. Insbesondere müssen zwei der Lösungen des Lagrangeschen Gleichungssystems diese Punkte liefern.

Beispiel 1: Wir bestimmen Maxima und Minima der Funktion $f(x, y) = x \cdot y$ auf der Sphäre $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$:

1) Wir wissen bereits aus Abschnitt 6.1, dass S^1 eine 1-dimensionale Untermannigfaltigkeit des \mathbb{R}^2 ist. Die definierende Gleichung lautet $g(x, y) := x^2 + y^2 - 1 = 0$.

2) Wir betrachten die Funktion

$$F(x, y, c) := f(x, y) - c g(x, y) = x \cdot y - cx^2 - cy^2 + c.$$

und lösen das Lagrangesche Gleichungssystem

$$\begin{aligned}\frac{\partial F}{\partial x}(x, y, c) &= y - 2cx = 0 \\ \frac{\partial F}{\partial y}(x, y, c) &= x - 2cy = 0, \\ \frac{\partial F}{\partial c}(x, y, c) &= -x^2 - y^2 + 1 = 0.\end{aligned}$$

Aus den ersten beiden Gleichungen folgt $cx^2 = cy^2$. $c = 0$ können wir ausschließen, sonst wäre $(x, y) = (0, 0) \notin S^1$. Somit gilt $x^2 = y^2$ und daher $|x| = |y|$. Da außerdem $x^2 + y^2 = 1$, folgt $x = \pm\sqrt{\frac{1}{2}}$ und $y = \pm\sqrt{\frac{1}{2}}$. Damit sind die einzig *möglichen* Punkte $p \in S^1$, in denen $f|_{S^1}$ einen lokalen Extremwert annehmen könnte

$$p \in \left\{ \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right), \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \right\}$$

mit den Funktionswerten

$$f(p) = \begin{cases} \frac{1}{2} & \text{für } p = \pm \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \\ -\frac{1}{2} & \text{für } p = \pm \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right). \end{cases}$$

Da S^1 kompakt ist, existieren $\max f|_{S^1}$ und $\min f|_{S^1}$. Somit ist

$$\begin{aligned}\max(f|_{S^1}) &= \frac{1}{2} & \text{und wird angenommen in } & \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right), \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right), \\ \min(f|_{S^1}) &= -\frac{1}{2} & \text{und wird angenommen in } & \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right), \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right).\end{aligned}$$

Weitere lokale Extremwerte treten nicht auf.

Beispiel 2: Wir bestimmen den Lotpunkt des Punktes $P = (1, 0, 0)$ auf die Ebene $E := \{(x, y, z) \in \mathbb{R}^3 \mid z = x + y\}$, also den Punkt $P_0 \in E$, der den kleinsten Abstand zu P hat: Der Abstand von P zu E ist definiert als

$$\text{dist}(P, E) := \inf\{d(P, Q) \mid Q \in E\}.$$

Außerhalb eines kompakten Bereiches der Ebene überschreiten die Abstände zu P einen gegebenen Abstand. Folglich existiert ein $P_0 \in E$ mit $\text{dist}(P, E) = d(P, P_0)$. Um P_0 zu finden, betrachten wir das Quadrat der Abstandsfunktion

$$f(x, y, z) := d((x, y, z), P)^2 = \|(x - 1, y, z)\|^2 = (x - 1)^2 + y^2 + z^2$$

und suchen einen Punkt $P_0 \in E$, in dem $f|_E$ minimal wird.

1) Die Ebene E ist der Graph einer Funktion, also eine 2-dimensionale Untermannigfaltigkeit von \mathbb{R}^3 .

2) E ist die Nullstellenmenge von $g(x, y, z) = z - (x + y)$. Wir betrachten die Funktion

$$F(x, y, z, c) := f(x, y, z) - cg(x, y, z) = (x - 1)^2 + y^2 + z^2 - cz + cx + cy$$

und lösen das Lagrangesche Gleichungssystem

$$\begin{aligned}\frac{\partial F}{\partial x}(x, y, z, c) &= 2(x - 1) + c = 0, \\ \frac{\partial F}{\partial y}(x, y, z, c) &= 2y + c = 0, \\ \frac{\partial F}{\partial z}(x, y, z, c) &= 2z - c = 0, \\ \frac{\partial F}{\partial c}(x, y, z, c) &= x + y - z = 0.\end{aligned}$$

Die einzige Lösung dieses Gleichungssystems ist $(x, y, z, c) = (\frac{2}{3}, -\frac{1}{3}, \frac{1}{3}, \frac{2}{3})$. Somit haben wir unseren gesuchten Punkt $P_0 \in E$ gefunden: $P_0 = (\frac{2}{3}, -\frac{1}{3}, \frac{1}{3})$.

Bemerkung: In diesem Beispiel ist die Nebenbedingung schon in expliziter Form durch $z = x + y$ gegeben. Dadurch könnten wir hier auch gleich so vorgehen wie im zweiten Teil des Beweises von Satz 6.20. Wir setzen die Nebenbedingung $z = x + y$ in die zu minimierende Funktion $f(x, y, z) = (x - 1)^2 + y^2 + z^2$ ein und bestimmen den Punkt, in dem die entstehende Funktion

$$\psi(x, y) := f(x, y, x + y) = (x - 1)^2 + y^2 + (x + y)^2.$$

ihr globales Minimum annimmt.

Integralrechnung für Funktionen einer reellen Variablen

In diesem Kapitel beginnen wir mit der Integralrechnung. Die Integralrechnung wird durch zwei verschiedene Problemstellungen motiviert:

1. *Lösung von Differentialgleichungen.*

Wir wollen z.B. die Bahnkurve eines Teilchens aus der Kenntnis seiner Geschwindigkeit bestimmen. Hier ist also die Geschwindigkeitsfunktion $v := \gamma' : I \subset \mathbb{R} \rightarrow \mathbb{R}^3$ gegeben und wir wollen γ bestimmen. Viele Prozesse in der Natur werden durch Differentialgleichungen modelliert. Die Aufgabe besteht dann immer darin, aus gegebenen Ableitungen (gewöhnliche oder partielle) einer Funktion die Funktion selbst zu bestimmen. Dazu muß man den Prozeß des Differenzierens "umkehren".

2. *Bestimmung von Flächeninhalten und Volumen.*

Die Integralrechnung stellt Methoden bereit, mit denen man Flächeninhalte und Volumen von Teilmengen des \mathbb{R}^n berechnen kann. In Analysis III werden wir die Integralrechnung für allgemeine Maßräume behandeln. Dabei messen dann Integrale nicht mehr das geometrische Volumen einer Menge, sondern andere Eigenschaften dieser Menge.

In diesem Abschnitt behandeln wir zunächst die Integralrechnung für Funktionen *einer* reellen Variablen. Wie in Kapitel 5 betrachten wir im gesamten Kapitel 7 Funktionen des Typs $f : I \subset \mathbb{R} \rightarrow E$, wobei $I \subset \mathbb{R}$ ein Intervall¹ und $(E, \|\cdot\|)$ ein normierter Vektorraum über dem Körper \mathbb{K} der reellen oder komplexen Zahlen ist.

7.1 Stammfunktionen und ihre Berechnung

Wir beginnen mit der Integralrechnung als Umkehrprozeß des Differenzierens.

Definition 7.1. *Eine Funktion $F : I \subset \mathbb{R} \rightarrow E$ heißt Stammfunktion der Funktion $f : I \subset \mathbb{R} \rightarrow E$, wenn F differenzierbar ist und $F' = f$ gilt.*

$S(I, E)$ bezeichnet die Menge der Funktionen von I nach E , die eine Stammfunktionen besitzen.

¹ I ist ein Intervall beliebiger Sorte mit mehr als einen Punkt.

Nicht jede Funktion besitzt eine Stammfunktion. Insbesondere kann diese Eigenschaft vom gewählten Definitionsbereich abhängen (Übungsaufgabe). In Abschnitt 7.3. werden wir zeigen, dass jede *stetige* Funktion $f : [a, b] \rightarrow E$ mit Werten in einem Banachraum E eine Stammfunktion besitzt.

Satz 7.1 Sei $F : I \subset \mathbb{R} \rightarrow E$ eine Stammfunktion von $f : I \subset \mathbb{R} \rightarrow E$. Dann ist $\tilde{F} : I \subset \mathbb{R} \rightarrow E$ genau dann eine Stammfunktion von f , wenn ein konstanter Vektor $c \in E$ existiert, so dass $\tilde{F} = F + c$.

Beweis. (\Leftarrow) Ist $\tilde{F} = F + c$, so ist \tilde{F} differenzierbar und es gilt $\tilde{F}' = F' = f$. (\Rightarrow) Sei \tilde{F} eine Stammfunktion von f . Dann ist $\tilde{F} - F$ differenzierbar und $(\tilde{F} - F)' = 0$. Aus dem Mittelwertsatz der Differentialrechnung folgt dann, dass $\tilde{F} - F$ eine konstante Funktion ist. \square

Definition 7.2. Sei $f \in S(I, E)$. Unter dem *unbestimmten Integral* von f versteht man die Menge aller Stammfunktionen von f . Wir bezeichnen diese Menge mit dem Symbol

$$\int f(x) dx := \{F : I \subset \mathbb{R} \rightarrow E \mid F \text{ Stammfunktion von } f\}$$

und schreiben dafür auch

$$\int f(x) dx = F(x) + c \quad \text{auf } I.$$

f heißt der *Integrand* und x die *Integrationsvariable*.

In der folgenden Liste stellen wir einige wichtige Grundintegrale zusammen. Der Beweis erfolgt durch Ableiten der Stammfunktion.

Wichtige Grundintegrale:

- (1) $\int x^\alpha dx = \frac{1}{\alpha + 1} x^{\alpha+1} + c$. Dies gilt für $\alpha \in \mathbb{N}$ auf \mathbb{R} , für $\alpha \in -\mathbb{N} \setminus \{-1\}$ auf $(-\infty, 0)$ und $(0, \infty)$ sowie für $\alpha \in \mathbb{R} \setminus \{\mathbb{Z}\}$ auf $(0, \infty)$.
- (2) $\int \frac{1}{x} dx = \ln |x| + c$ auf $(0, \infty)$ und $(-\infty, 0)$.
- (3) $\int e^x dx = e^x + c$ auf \mathbb{R} .
- (4) $\int \cos x dx = \sin x + c$, $\int \sin x dx = -\cos x + c$ auf \mathbb{R} .
- (5) $\int \cosh x dx = \sinh x + c$, $\int \sinh x dx = \cosh x + c$ auf \mathbb{R} .
- (6) $\int \frac{1}{1+x^2} dx = \arctan x + c$ auf \mathbb{R} .
- (7) $\int \frac{1}{1-x^2} dx = \frac{1}{2} \ln \left| \frac{1+x}{1-x} \right| + c = \begin{cases} \operatorname{artanh}(x) + c & \text{auf } (-1, 1) \\ \operatorname{arcoth}(x) + c & \text{auf } (-\infty, -1) \cup (1, \infty) \end{cases}$.

$$(8) \int \frac{1}{\sqrt{1+x^2}} dx = \operatorname{arsinh}(x) + c \quad \text{auf } \mathbb{R}.$$

$$(9) \int \frac{1}{\sqrt{1-x^2}} dx = \arcsin x + c \quad \text{auf } (-1, 1).$$

$$(10) \int \frac{1}{\sqrt{x^2-1}} dx = \ln|x + \sqrt{x^2-1}| + c = \begin{cases} \operatorname{arcosh}(x) + c & \text{auf } (1, \infty) \\ -\operatorname{arcosh}(-x) + c & \text{auf } (-\infty, -1) \end{cases}.$$

$$(11) \int e^{(a+ib)x} dx = \frac{e^{(a+ib)x}}{(a+ib)} + c \quad \text{auf } \mathbb{R} \text{ für } a+ib \neq 0.$$

$$(12) \int \frac{dx}{(x-(a+ib))} = \ln|x-(a+ib)| + i \cdot \arctan\left(\frac{x-a}{b}\right) + c \quad \text{für } b \neq 0.$$

$$(13) \int \frac{dx}{(x-(a+ib))^n} = -\frac{1}{n-1} \cdot \frac{1}{(x-(a+ib))^{n-1}} + c \quad \text{für } n \in \mathbb{N}, n > 1.$$

Satz 7.2 (Rechenregeln für unbestimmte Integrale)

1. Seien $f, g \in S(I, E)$ und $\mu, \lambda \in \mathbb{K}$. Dann ist auch $\lambda f + \mu g \in S(I, E)$ und es gilt

$$\int (\lambda f(x) + \mu g(x)) dx = \lambda \int f(x) dx + \mu \int g(x) dx.$$

2. Partielle Integration:

Seien $f, g : I \subset \mathbb{R} \rightarrow E$ differenzierbar und $f \cdot g' \in S(I, E)$. Dann gilt $f' \cdot g \in S(I, E)$ und

$$\int f'(x)g(x) dx = - \int f(x)g'(x) dx + f(x)g(x) + c.$$

3. Substitutionsregel:

$g \in S(I, E)$ habe die Stammfunktion G , $f : J \subset \mathbb{R} \rightarrow I \subset \mathbb{R}$ sei differenzierbar. Dann hat die Funktion $g^* := f' \cdot (g \circ f) : J \subset \mathbb{R} \rightarrow E$ die Stammfunktion $G^* := G \circ f$, das heißt

$$\int g(y) dy \Big|_{y=f(x)} = \int f'(x) \cdot g(f(x)) dx.$$

(Formale Regel: Man substituiert $y := f(x)$ und schreibt $dy = f'(x) dx$.)

4. Integrale in Produkträumen:

Sei $f = (f_1, f_2) : I \subset \mathbb{R} \rightarrow E_1 \times E_2$. Dann gilt $f \in S(I, E_1 \times E_2)$ genau dann, wenn $f_i \in S(I, E_i)$ für $i \in \{1, 2\}$ und in diesem Fall

$$\int f(x) dx = \left(\int f_1(x) dx, \int f_2(x) dx \right).$$

Beweis. Der Beweis folgt aus den Rechenregeln für die Ableitung (Summen- und Produktregel, Kettenregel, komponentenweises Differenzieren). \square

Beispiel 1: Sei $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ eine differenzierbare Funktion ohne Nullstellen auf I . Dann gilt

$$\int \frac{f'(x)}{f(x)} dx = \ln |f(x)| + c \quad \text{auf } I.$$

Dies sieht man durch Ableiten der Funktion auf der rechten Seite oder durch die Substitution $y = f(x)$, $dy = f'(x) dx$:

$$\int \frac{f'(x)}{f(x)} dx = \int \frac{dy}{y} \Big|_{y=f(x)} = \ln |y| \Big|_{y=f(x)} + c = \ln |f(x)| + c \quad \text{auf } I.$$

So ist zum Beispiel

$$\int \tan x dx = \int \frac{\sin x}{\cos x} dx = -\ln |\cos x| + c \quad \text{auf } \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Beispiel 2: Berechnung von $I_n := \int \sin^n(x)$, $n \in \mathbb{N}_0$.

Wir nehmen an, dass $\sin^n(x)$ eine Stammfunktion besitzt und benutzen partielle Integration:

$$\begin{aligned} I_n &= - \int \sin^{n-1}(x) \cdot \cos'(x) dx \\ &\stackrel{\text{part. Int.}}{=} - \sin^{n-1}(x) \cdot \cos(x) + \int \frac{d}{dx} \sin^{n-1}(x) \cdot \cos(x) dx \\ &= - \sin^{n-1}(x) \cdot \cos(x) + (n-1) \int \sin^{n-2}(x) (1 - \sin^2(x)) dx \\ &= - \sin^{n-1}(x) \cdot \cos(x) + (n-1)I_{n-2} - (n-1)I_n. \end{aligned}$$

Daraus erhält man die Rekursionsformel:

$$I_n = -\frac{1}{n} \cdot \sin^{n-1}(x) \cdot \cos(x) + \frac{n-1}{n} I_{n-2} \quad \text{für } n \geq 2,$$

wobei $I_0 = x + c$ und $I_1 = -\cos x + c$ ist.

(Da I_0 und I_1 existieren, existiert nach Induktion auch I_n und läßt sich rekursiv berechnen).

Für $J_n := \int \cos^n(x) dx$ berechnet man analog

$$J_n = \frac{1}{n} \cos^{n-1}(x) \cdot \sin(x) + \frac{n-1}{n} J_{n-2} \quad \text{für } n \geq 2$$

mit $J_0 = x + c$ und $J_1 = \sin(x) + c$.

Beispiel 3: Integrale rationaler Funktionen.

Seien $P, Q \in \mathbb{K}[z]$ Polynome über dem Körper \mathbb{K} , $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$, und $Q \neq 0$. Wir betrachten eine Teilmenge $A \subset \mathbb{K}$, auf der Q keine Nullstelle hat. Dann heißt die Funktion $f : A \subset \mathbb{K} \rightarrow \mathbb{K}$ mit

$$f(x) := \frac{P(x)}{Q(x)}$$

komplexe ($\mathbb{K} = \mathbb{C}$) bzw. reelle ($\mathbb{K} = \mathbb{R}$) rationale Funktion auf A .

Unser Ziel ist die Berechnung des Integrals $\int \frac{P(x)}{Q(x)} dx$ auf $I \subset \mathbb{R}$. Als Methode benutzen wir die *Partialbruchzerlegung rationaler Funktionen*. Wir erinnern dazu an den Fundamentalsatz der Algebra (siehe Kapitel 4.7):

Sei $Q \in \mathbb{C}[z]$ ein komplexes Polynom vom Grad $n \geq 1$. Dann hat Q n komplexe Nullstellen. Sind ξ_1, \dots, ξ_m die verschiedenen Nullstellen von Q mit der Vielfachheit jeweils ν_j , so kann man Q in Linearfaktoren zerlegen:

$$Q(z) = a_n(z - \xi_1)^{\nu_1} \cdot (z - \xi_2)^{\nu_2} \cdot \dots \cdot (z - \xi_m)^{\nu_m}.$$

Satz 7.3 (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_m mit der Vielfachheit ν_1, \dots, ν_m . Dann existieren eindeutig bestimmte Konstanten $C_{jk} \in \mathbb{C}$ für $j \in \{1, \dots, m\}$, $k \in \{1, \dots, \nu_j\}$, so dass

$$\begin{aligned} f(z) := \frac{P(z)}{Q(z)} &= \sum_{j=1}^m \sum_{k=1}^{\nu_j} \frac{C_{jk}}{(z - \xi_j)^k} \\ &= \sum_{j=1}^m \left(\frac{C_{j1}}{z - \xi_j} + \frac{C_{j2}}{(z - \xi_j)^2} + \dots + \frac{C_{j\nu_j}}{(z - \xi_j)^{\nu_j}} \right) \quad \forall z \neq \xi_j. \end{aligned}$$

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 $Q(z) = b_1 z + b_0 = b_1(z - \xi_1)$. Daraus folgt $f(z) = \frac{a_0}{b_1(z - \xi_1)}$, das heißt die Behauptung folgt mit $C_{11} := \frac{a_0}{b_1}$. C_{11} ist offensichtlich eindeutig bestimmt.

Ind.-Schritt: Wir setzen voraus, dass die Behauptung für alle rationalen Funktionen $\tilde{f} = \frac{\tilde{P}}{\tilde{Q}}$ mit $\deg(\tilde{Q}) \leq n - 1$ gilt und zeigen sie dann für rationale Funktionen $f = \frac{P}{Q}$ 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 $S(z) \in \mathbb{C}[z]$ mit $\deg(S) = n - \nu_1 < n$ und $S(\xi_1) \neq 0$. Sei $a := \frac{P(\xi_1)}{S(\xi_1)}$. Dann gilt

$$\frac{P(z)}{Q(z)} - \frac{a}{(z - \xi_1)^{\nu_1}} = \frac{P(z) - a \cdot S(z)}{(z - \xi_1)^{\nu_1} \cdot S(z)}. \quad (*)$$

Nach Definition von a gilt $P(\xi_1) - aS(\xi_1) = 0$. Ist $P - aS = 0$, so gilt die Behauptung. Sei nun $P - aS \neq 0$. Dann kann man den Linearfaktor $(z - \xi_1)$ abspalten, das heißt es gilt $P(z) - aS(z) = (z - \xi_1)\tilde{S}(z)$. Es folgt

$$\frac{P(z)}{Q(z)} \stackrel{(*)}{=} \frac{a}{(z - \xi_1)^{\nu_1}} + \frac{\tilde{S}(z)}{(z - \xi_1)^{\nu_1-1} \cdot S(z)}.$$

Da $\deg(S(z) \cdot (z - \xi_1)^{\nu_1-1}) = \deg(S) + \nu_1 - 1 = n - 1$, kann man nach Ind.-Voraussetzung die rationale Funktion im zweiten Summanden entsprechend der Behauptung zerlegen. Wegen $Q(z) = (z - \xi_1)^{\nu_1} S(z)$ kommen im Nenner dieser rationalen Funktion alle Nullstellen von

Q vor, wobei ξ_1 mit Vielfachheit $\nu_1 - 1$ auftritt. Die Koeffizienten in der Zerlegung des zweiten Summanden sind nach Ind.-Voraussetzung eindeutig bestimmt, somit ist auch $C_{1\nu_1} := a$ eindeutig bestimmt. \square

Ist $Q \in \mathbb{R}[x]$ ein reelles Polynom, so tritt mit jeder echt komplexen Nullstelle $\xi = a + ib$ von Q auch die konjugiert-komplexe Zahl $\bar{\xi} = a - ib$ als Nullstelle mit der gleichen Vielfachheit auf. Dies liefert die folgende speziellere Partialbruchzerlegung für reelle rationale Funktionen:

Satz 7.4 (Partialbruchzerlegung reeller rationaler Funktionen)

Seien $P, Q \in \mathbb{R}[x]$ zwei reelle Polynome mit $\deg(P) < \deg(Q)$. Bezeichne $\lambda_1, \dots, \lambda_s$ die verschiedenen reellen Nullstellen von Q mit der Vielfachheit μ_1, \dots, μ_s und ξ_1, \dots, ξ_r sowie $\bar{\xi}_1, \dots, \bar{\xi}_r$ die verschiedenen echt komplexen Nullstellen von Q mit der Vielfachheit ν_1, \dots, ν_r . Dann existieren eindeutig bestimmte reelle Zahlen $C_{jk} \in \mathbb{R}$ und komplexe Zahlen $A_{lk} \in \mathbb{C}$ so dass

$$f(x) := \frac{P(x)}{Q(x)} = \sum_{j=1}^s \sum_{k=1}^{\mu_j} \frac{C_{jk}}{(x - \lambda_j)^k} + \sum_{l=1}^r \sum_{k=1}^{\nu_l} \left(\frac{A_{lk}}{(x - \xi_l)^k} + \frac{\bar{A}_{lk}}{(x - \bar{\xi}_l)^k} \right) \quad \forall x \in \mathbb{R}, x \neq \lambda_j.$$

Beweis. Wir betrachten f als komplexe rationale Funktion und machen ihre Partialbruchzerlegung wie in Satz 7.3. Es existieren also komplexe Zahlen A_{lk}, B_{lk}, C_{jk} , so dass

$$f(x) = \frac{P(x)}{Q(x)} = \sum_{j=1}^s \sum_{k=1}^{\mu_j} \frac{C_{jk}}{(x - \lambda_j)^k} + \sum_{l=1}^r \sum_{k=1}^{\nu_l} \left(\frac{A_{lk}}{(x - \xi_l)^k} + \frac{B_{lk}}{(x - \bar{\xi}_l)^k} \right) \quad \forall x \in \mathbb{R}, x \neq \lambda_j.$$

Für $x \in \mathbb{R}$ gilt $\overline{f(x)} = f(x)$. Dies liefert

$$\begin{aligned} 0 &= f(x) - \overline{f(x)} \\ &= \sum_{j=1}^s \sum_{k=1}^{\mu_j} \frac{C_{jk} - \bar{C}_{jk}}{(x - \lambda_j)^k} + \sum_{l=1}^r \sum_{k=1}^{\nu_l} \left(\frac{A_{lk} - \bar{B}_{lk}}{(x - \xi_l)^k} + \frac{B_{lk} - \bar{A}_{lk}}{(x - \bar{\xi}_l)^k} \right) \quad \forall x \in \mathbb{R}, x \neq \lambda_j. \end{aligned}$$

Wegen der Eindeutigkeit der Konstanten in der Partialbruchzerlegung folgt $C_{jk} = \bar{C}_{jk} \in \mathbb{R}$ und $A_{lk} = \bar{B}_{lk}$. \square

Methoden zur Berechnung der Koeffizienten in der Partialbruchzerlegung:

Sei f eine komplexe rationale Funktion und

$$f(z) = \frac{P(z)}{Q(z)} = \sum_{j=1}^m \sum_{k=1}^{\nu_j} \frac{C_{jk}}{(z - \xi_j)^k} \quad (*)$$

ihre Partialbruchzerlegung. Es gibt drei verschiedene Methoden, die Koeffizienten C_{jk} in der Partialbruchzerlegung von $f(z)$ zu bestimmen.

- a) Wir multiplizieren $f(z)$ mit $Q(z)$. Dann erhalten wir eine Gleichung für zwei Polynome und können die Konstanten C_{jk} durch Koeffizientenvergleich bestimmen.

- b) Wir setzen n spezielle Werte $z_1, \dots, z_n \in \mathbb{C}$ in die Gleichung (*) ein und erhalten ein lineares Gleichungssystem für die Konstanten C_{jk} , das wir lösen können.
- c) Wir multiplizieren (*) mit $(z - \xi_j)^{\nu_j}$ und erhalten für den Koeffizienten $C_{j\nu_j}$

$$C_{j\nu_j} = \lim_{z \rightarrow \xi_j} \left(\frac{P(z)}{Q(z)} \cdot (z - \xi_j)^{\nu_j} \right) = \lim_{z \rightarrow \xi_j} \frac{P(z)}{\prod_{i \neq j} (z - \xi_i)^{\nu_i}}.$$

Die restlichen Koeffizienten lassen sich wie in a), b) oder c) bestimmen.

Beispiel 3a: Partialbruchzerlegung der komplexen rationalen Funktion

$$f(z) = \frac{z + 1}{z^4 - z^3 + z^2 - z} = \frac{P(z)}{Q(z)}.$$

Dann sind offensichtlich $\xi_1 = 0$, $\xi_2 = 1$ Nullstellen von Q und es folgt mit Polynomdivision

$$\frac{Q(z)}{z(z-1)} = \frac{z^4 - z^3 + z^2 + z}{z^2 - z} = z^2 + 1.$$

Somit sind die weiteren Nullstellen von Q gleich $\pm i$. Wir erhalten für die Partialbruchzerlegung

$$f(z) = \frac{C_1}{z} + \frac{C_2}{z-1} + \frac{C_3}{z-i} + \frac{C_4}{z+i}.$$

Zur Berechnung der Koeffizienten wenden wir Methode c) an, das heißt

$$C_1 = \lim_{z \rightarrow 0} \left(\frac{z+1}{(z-1)(z-i)(z+i)} \right) = \frac{1}{-1 \cdot (-i) \cdot i} = -1,$$

$$C_2 = \lim_{z \rightarrow 1} \left(\frac{z+1}{z(z-i)(z+i)} \right) = 1.$$

Durch analoges Rechnen erhält man $C_3 = \frac{i}{2}$ und $C_4 = -\frac{i}{2}$. Somit gilt

$$f(z) = -\frac{1}{z} + \frac{1}{z-1} + \frac{i}{2(z-i)} - \frac{i}{2(z+i)}.$$

Beispiel 3b: Partialbruchzerlegung der reellen rationalen Funktion

$$f(x) = \frac{2x^3 + 2x^2 + 2x - 2}{(x^2 + 1)^2} = \frac{P(x)}{Q(x)}.$$

Offensichtlich gilt

$$Q(x) = (x^2 + 1)^2 = (x-i)^2(x+i)^2.$$

Folglich sind $\xi_1 = i$ und $\xi_2 = -i$ zweifache Nullstellen von Q . Wir erhalten für die Partialbruchzerlegung:

$$f(z) = \frac{A}{z-i} + \frac{B}{(z-i)^2} + \frac{\bar{A}}{z+i} + \frac{\bar{B}}{(z+i)^2}.$$

Den Koeffizienten B bestimmen wir mit Methode c):

$$B = \lim_{z \rightarrow i} \frac{2z^3 + 2z^2 + 2z - 2}{(z+i)^2} = \frac{-2i - 2 + 2i - 2}{4i^2} = 1.$$

Setzen wir dann zum Beispiel die speziellen Werte $x = 0$ und $x = 1$ ein, so erhalten wir das folgende Gleichungssystem für A und \bar{A} :

$$\begin{aligned} 0 &= A + \bar{A} \\ 1 &= \frac{A}{1-i} + \frac{1}{(1-i)^2} + \frac{\bar{A}}{1+i} + \frac{1}{(1+i)^2}. \end{aligned}$$

Die Lösung liefert $A = \bar{A} = 1$.

Oder wir multiplizieren mit $Q(x)$ und nehmen einen Koeffizientenvergleich vor:

$$\begin{aligned} 2x^3 + 2x^2 + 2x - 2 &= A(x-i)(x+i)^2 + 1 \cdot (x+i)^2 + \bar{A}(x-i)^2(x+i) + 1 \cdot (x-i)^2 \\ &= x^3(A + \bar{A}) + x^2(2 + Ai - \bar{A}i) + x(A + \bar{A}) + (-2 + Ai - \bar{A}i) \end{aligned}$$

Daraus erhalten wir ebenfalls $A = \bar{A} = 1$. Somit gilt

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

Anwendung auf die Berechnung der Integrale rationaler Funktionen

Seien $P, Q \in \mathbb{C}[z]$ komplexe 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 $P_1, P_2 \in \mathbb{C}[z]$ und $\deg(P_2) < \deg(Q)$.

b) Wir zerlegen $\frac{P_2(x)}{Q(x)}$ in Partialbrüche.

c) Es bleibt nun, die Integrale der einzelnen Summanden auf I zu berechnen. Wir geben die relevanten Stammfunktionen an. Der Beweis erfolgt durch Ableiten der Stammfunktion:

$$\begin{aligned} \int x^k dx &= \frac{1}{k+1} \cdot x^{k+1} + c \quad \text{für } k \geq 0, \\ \int \frac{1}{x-\xi} dx &= i \arctan\left(\frac{x-a}{b}\right) + \ln|x-\xi| + c \quad \text{für } \xi = a + ib \text{ mit } b \neq 0, \\ \int \frac{1}{x-\lambda} dx &= \ln|x-\xi| + c \quad \text{für } \lambda \in \mathbb{R}, \\ \int \frac{1}{(x-\xi)^k} dx &= -\frac{1}{k-1} \cdot \frac{1}{(x-\xi)^{k-1}} + c \quad \text{für } k \in \mathbb{N}, k > 1, \xi \in \mathbb{C}. \end{aligned}$$

Dadurch kann man $\int \frac{P(x)}{Q(x)} dx$ vollständig berechnen.

- d) Sind die Polynome P und Q reell, so beschreibt man die Stammfunktion ebenfalls als reelle Funktion. Wir gehen wie oben vor und fassen am Schluß die auftretenden komplexen und konjugiert komplexen Summanden zu reellen Funktionen zusammen. Für echt komplexe Zahlen $\xi = a + bi$, $b \neq 0$ erhält man:

$$\int \left(\frac{A}{x - \xi} + \frac{\bar{A}}{x - \bar{\xi}} \right) dx = 2 \cdot \operatorname{Re}(A) \cdot \ln \sqrt{(x - a)^2 + b^2} - 2 \cdot \operatorname{Im}(A) \cdot \arctan \left(\frac{x - a}{b} \right) + c$$

und

$$\begin{aligned} \int \left(\frac{A}{(x - \xi)^k} + \frac{\bar{A}}{(x - \bar{\xi})^k} \right) dx &= -\frac{1}{k-1} \left(\frac{A}{(x - \xi)^{k-1}} + \frac{\bar{A}}{(x - \bar{\xi})^{k-1}} \right) + c \\ &= -\frac{1}{k-1} \left(\frac{A(x - \bar{\xi})^{k-1} + \bar{A}(x - \xi)^{k-1}}{(x^2 - 2x \cdot a + |\xi|^2)^{k-1}} \right) + c. \end{aligned}$$

Beispiel 3c. Wir berechnen das Integral

$$I := \int \frac{2x^3 + 2x^2 + 2x - 2}{(x^2 + 1)^2} dx.$$

In Beispiel 3b hatten wir die Partialbruchzerlegung des Integranden bestimmt. Wir erhalten

$$\begin{aligned} I &= \int \left(\frac{1}{(x - i)} + \frac{1}{(x + i)} \right) + \int \left(\frac{1}{(x - i)^2} + \frac{1}{(x + i)^2} \right) dx \\ &= 2 \ln \sqrt{x^2 + 1} - \frac{1}{(x - i)} - \frac{1}{(x + i)} + c \\ &= 2 \ln \sqrt{x^2 + 1} - \frac{2x}{x^2 + 1} + c \quad \text{auf } \mathbb{R}. \end{aligned}$$

Beispiel 4: Standardsubstitutionen für Integrale der Form

$$\int R(f_1(x), f_2(x), \dots, f_N(x)) dx.$$

Hier bezeichnet R eine rationale Funktion in N Variablen, d.h.

$$R(u_1, \dots, u_N) := \frac{\sum_{\alpha} \alpha_{i_1 \dots i_N} u_1^{\alpha_1} \cdot \dots \cdot u_N^{\alpha_N}}{\sum_{\beta} \beta_{i_1 \dots i_N} u_1^{\beta_1} \cdot \dots \cdot u_N^{\beta_N}}, \quad \alpha_I, \beta_I \in \mathbb{C}.$$

und $f_1, \dots, f_N : I \subset \mathbb{R} \rightarrow \mathbb{C}$ Funktionen. Solche Integrale lassen sich in speziellen Fällen (für sehr spezielle f_i) mittels geeigneter Substitutionen auf Integrale rationaler Funktionen in *einer* Variablen zurückführen. Wir geben einige dieser Standardsubstitutionen an. Das Intervall I , für das das Integral berechnet wird, sei im Folgenden immer so gewählt, dass alles wohldefiniert ist.

1. Typ: $\int R(e^{ax}) dx$ mit $a \in \mathbb{R}, a \neq 0$:

Substituiere $y := e^{ax}, dy = ae^{ax} dx$.

2. Typ: $\int R(e^{ax}, \sinh ax, \cosh ax) dx$ mit $a \in \mathbb{R}, a \neq 0$:

Substituiere $y := e^{ax}, dy = ae^{ax} dx$ und benutze

$$\cosh ax = \frac{e^{ax} + e^{-ax}}{2} \quad \text{und} \quad \sinh ax = \frac{e^{ax} - e^{-ax}}{2}.$$

3. Typ: $\int R(\cos x) \cdot \sin x dx$

Substituiere $y := \cos x, dy = -\sin x dx$.

4. Typ: $\int R(\sin x) \cdot \cos x dx$

Substituiere $y := \sin x, dy = \cos x dx$.

5. Typ: $\int R(\sin x, \cos x, \tan x, \cot x) dx$

Substituiere $y := \tan \frac{x}{2}$ und benutze

$$dx = \frac{2dy}{1+y^2}, \quad \cos x = \frac{1-y^2}{1+y^2}, \quad \sin x = \frac{2y}{1+y^2}, \quad \tan x = \frac{2y}{1-y^2}, \quad \cot x = \frac{1-y^2}{2y}.$$

6. Typ: $\int R(\sinh x, \cosh x, \tanh x, \coth x) dx$

Substituiere $y := \tanh \frac{x}{2}$ oder drücke die Hyperbelfunktionen durch die Exponentialfunktion aus und gehe wie für Typ 1 vor.

7. Typ: $\int R\left(x, \sqrt[n]{\frac{ax+b}{cx+d}}\right) dx$ mit $a, b, c, d \in \mathbb{R}$ und $ad - bc \neq 0$:

Substituiere $y := \sqrt[n]{\frac{ax+b}{cx+d}}$

8. Typ: $\int R\left(x^{\frac{1}{n}}, x^{\frac{1}{m}}, x^{\frac{1}{p}}, \dots\right) dx$

Substituiere $y := x^{\frac{1}{r}}, dx = ry^{r-1} dy$, wobei $r = \text{kgV}(n, m, p, \dots)$.

9. Typ: $\int R(x, \sqrt{x^2 + a^2}) dx$ mit $a \in \mathbb{R}, a > 0$:

Substituiere $x := a \sinh y, dx = a \cosh y dy, \sqrt{x^2 + a^2} = a \cosh y$.

10. Typ: $\int R(x, \sqrt{x^2 - a^2}) dx$ mit $a \in \mathbb{R}, a > 0$:

Substituiere $x := a \cosh y, dx = a \sinh y dy, \sqrt{x^2 - a^2} = |a \sinh y|$.

11. Typ: $\int R(x, \sqrt{a^2 - x^2}) dx$ mit $a \in \mathbb{R}, a > 0$:

Substituiere $x = a \sin y, dx = a \cos y dy, \sqrt{a^2 - x^2} = |a \cos y|$.

12. Typ: $\int R(x, \sqrt{ax^2 + bx + c}) dx :$

Substituiere $y = \frac{ax+b}{\sqrt{|b^2-4ax|}}$. Dann weiter substituieren $t = \sinh y, t = \cosh y$ oder $t = \cos y$.

13. Typ: $\int R(x, \sqrt{ax+b}, \sqrt{cx+d}) dx$

Kann durch Substitution $y := \sqrt{ax+b}, dx = \frac{1}{a}y dy$ auf Typ 12 zurückgeführt werden.

Nicht jedes Integral man durch elementare Funktionen ausdrücken, auch wenn die Stammfunktion existiert. Zum Beispiel kann man Integrale der Form $\int R(x, \sqrt{P(x)}) dx$, wobei $P(x)$ ein Polynom vom Grad 3 oder 4 mit ausschließlich einfachen Nullstellen ist, nicht durch elementare Funktionen beschreiben (elliptische Integrale). Man benötigt dann andere Methoden, z.B. die Entwicklung von Funktionen in Reihen.

Definition 7.3. Sei $f \in S(I, E)$ und $a, b \in I$. Unter dem bestimmten Integral von f von a nach b versteht man den Vektor (bzw. für $E = \mathbb{K}$ die Zahl):

$$\int_a^b f(x) dx := F(b) - F(a) \in E,$$

wobei $F : I \subset \mathbb{R} \rightarrow E$ eine Stammfunktion von f ist.

Wir werden auch die folgende Bezeichnung benutzen: $F(b) - F(a) =: F(x)|_a^b$

Satz 7.5 (Rechenregeln für das bestimmte Integral)

1. Sei $f \in S(I, E)$ und $a, b, c \in I$. Dann gilt

$$\int_a^a f(x) dx = 0, \quad \int_a^b f(x) dx = - \int_b^a f(x) dx, \quad \int_a^b f(x) dx + \int_b^c f(x) dx = \int_a^c f(x) dx.$$

2. Seien $f, g \in S(I, E)$, $\lambda, \mu \in \mathbb{K}$ und $a, b \in I$. Dann gilt

$$\lambda \int_a^b f(x) dx + \mu \int_a^b g(x) dx = \int_a^b (\lambda f(x) + \mu g(x)) dx.$$

3. Sei $f \in S(I, E_1 \times E_2)$ und $f = (f_1, f_2)$. Dann gilt

$$\int_a^b f(x) dx = \left(\int_a^b f_1(x) dx, \int_a^b f_2(x) dx \right).$$

4. Partielle Integration:

Seien $f, g : I \subset \mathbb{R} \rightarrow E$ differenzierbar, $a, b \in I$, $f \cdot g' \in S(I, E)$. Dann gilt

$$\int_a^b f'(x)g(x) dx = - \int_a^b f(x)g'(x) dx + f(b)g(b) - f(a)g(a).$$

5. Substitutionsregel:

Sei $g \in S(I, E)$ und $f : J \subset \mathbb{R} \rightarrow I \subset \mathbb{R}$ differenzierbar. Dann gilt

$$\int_a^b g(f(x))f'(x) dx = \int_{f(a)}^{f(b)} g(y) dy.$$

Beweis. Dies folgt unmittelbar aus Satz 7.2. □

Beispiel 2': Für $n \in \mathbb{N}$, $n > 1$ gilt

$$\int_0^{\frac{\pi}{2}} \sin^n(x) dx = \begin{cases} \frac{1 \cdot 3 \cdot \dots \cdot (2k-1)}{2 \cdot 4 \cdot \dots \cdot 2k} \cdot \frac{\pi}{2} & \text{für } n = 2k \\ \frac{2 \cdot 4 \cdot \dots \cdot 2k}{3 \cdot 5 \cdot \dots \cdot (2k+1)} & \text{für } n = 2k + 1. \end{cases}$$

Um diese Formel nachzuweisen, benutzen wir die Rekursionsformel für die Stammfunktion $I_n(x)$ von $\sin^n(x)$ aus Beispiel 2. Es gilt $I_0(x) = x + c$, $I_1(x) = -\cos x + c$, und

$$I_n(x) = -\frac{1}{n} \sin^{n-1} x \cdot \cos x + \frac{n-1}{n} \cdot I_{n-2}(x) \quad \text{für } n \in \mathbb{N}, n > 1.$$

Somit erhalten wir

$$I_n(0) = \frac{n-1}{n} \cdot I_{n-2}(0) \quad \text{und} \quad I_n\left(\frac{\pi}{2}\right) = \frac{n-1}{n} \cdot I_{n-2}\left(\frac{\pi}{2}\right),$$

sowie

$$I_0(0) = 0, \quad I_1(0) = -1 \quad \text{und} \quad I_0\left(\frac{\pi}{2}\right) = \frac{\pi}{2}, \quad I_1\left(\frac{\pi}{2}\right) = 0.$$

Daraus folgt

$$\begin{aligned} I_{2k}(0) &= \frac{2k-1}{2k} \cdot I_{2k-2}(0) = \frac{(2k-1)(2k-3)}{(2k)(2k-2)} \cdot I_{2k-4}(0) \\ &= \dots = \frac{(2k-1)(2k-3) \cdot \dots \cdot 3 \cdot 1}{2k(2k-2) \cdot \dots \cdot 2} \cdot I_0(0) = 0. \end{aligned}$$

Analog findet man

$$I_{2k}\left(\frac{\pi}{2}\right) = \frac{(2k-1) \cdot \dots \cdot 3 \cdot 1}{2k \cdot \dots \cdot 4 \cdot 2} \cdot I_0\left(\frac{\pi}{2}\right) = \frac{(2k-1) \cdot \dots \cdot 3 \cdot 1}{2k \cdot \dots \cdot 4 \cdot 2} \cdot \frac{\pi}{2},$$

sowie

$$\begin{aligned} I_{2k+1}(0) &= \frac{(2k) \cdot \dots \cdot 4 \cdot 2}{(2k+1)(2k-1) \cdot \dots \cdot 3 \cdot 1} \cdot \underbrace{I_1(0)}_{=-1} \\ I_{2k+1}\left(\frac{\pi}{2}\right) &= \frac{(2k) \cdot \dots \cdot 4 \cdot 2}{(2k+1)(2k-1) \cdot \dots \cdot 3 \cdot 1} \cdot \underbrace{I_1\left(\frac{\pi}{2}\right)}_{=0} = 0. \end{aligned}$$

Dies ergibt die behauptete Formel. □

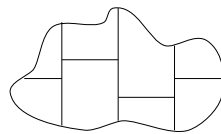
7.2 Das Riemann-Integral

Die zweite Motivation für die Entwicklung der Integralrechnung entstand aus dem Versuch, Längen, Flächeninhalte und Volumen zu berechnen. Wir werden hier zunächst das Riemann-Integral behandeln, das in rigoroser Form von Bernard Riemann in seiner Habilitationsschrift (1854) definiert wurde. Im Rahmen der Maß- und Integrationstheorie im 3. Semester werden wir später das Lebesgue-Integral kennenlernen, das etwas bessere Eigenschaften als das Riemann-Integral hat. Für einen kurzen historischen Überblick zur Entwicklung der Integralrechnung verweisen wir auf das 9. Kapitel des Buches *Analysis I* von W. Walter.

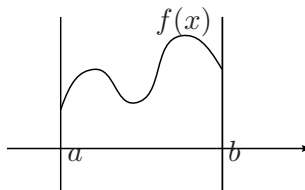
Zunächst formulieren wir einige Forderungen, die ein vernünftig definierter geometrischer Flächeninhalt $\mu(A)$ für Teilmengen $A \subset \mathbb{R}^2$ erfüllen sollte:

1. $\mu(A) \geq 0$,
2. $\mu([a, b] \times [c, d]) = (b - a) \cdot (d - c)$,
3. $\text{Int}(A) \cap \text{Int}(B) = \emptyset \implies \mu(A \cup B) = \mu(A) + \mu(B)$,
4. $A \subset B \implies \mu(A) \leq \mu(B)$.
5. $\mu(A)$ ist invariant gegen Euklidische Bewegungen (z.B. Drehungen, Verschiebungen).

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. Mit solchen Mengen werden wir uns nun beschäftigen. Ihr Flächeninhalt wird durch das Riemann-Integral von f (falls es existiert) beschrieben.

Wir definieren im Folgenden das Riemannsche Integral für Funktionen $f : [a, b] \subset \mathbb{R} \rightarrow E$, wobei E ein *Banachraum* über \mathbb{R} oder \mathbb{C} ist.

Definition 7.4. Sei $I = [a, b] \subset \mathbb{R}$ ein kompaktes Intervall. Unter einer Zerlegung (oder Unterteilung) von I versteht man eine Menge von Punkten $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ mit $a = x_0 < \dots < x_n = b$. $I_k := [x_{k-1}, x_k]$ heißt das k -te Teilintervall von \mathcal{P} , $L(I_k) := x_k - x_{k-1}$ die Länge von I_k und $\|\mathcal{P}\| := \max\{L(I_k) \mid k = 1, \dots, n\}$ 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 7.5. Sei $f : [a, b] \rightarrow E$ eine 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 $k = 1, \dots, n$. Dann heißt

$$S(f, \mathcal{P}, \xi) := \sum_{k=1}^n L(I_k) \cdot f(\xi_k)$$

Riemannsche Summe für f bzgl. der Zerlegung \mathcal{P} und der Stützstellen $\xi = (\xi_1, \dots, \xi_n)$.

Definition 7.6. Sei $f : [a, b] \rightarrow E$. Der Vektor $e \in E$ heißt Grenzwert der Riemannschen Summen für f , falls für alle $\varepsilon > 0$ eine Zerlegung \mathcal{P}_ε existiert, so dass

$$\|S(f, \mathcal{P}, \xi) - e\| < \varepsilon \quad \forall \mathcal{P} \geq \mathcal{P}_\varepsilon \text{ und für alle Stützstellen } \xi \text{ von } \mathcal{P}.$$

Bezeichnung: $e =: \lim_{\mathcal{P}} S(f, \mathcal{P}, \xi)$.

Der Grenzwert der Riemannschen Summen ist, falls er existiert, eindeutig bestimmt. Wir überlassen dem Leser den Beweis als Übungsaufgabe.

Definition 7.7. Eine Abbildung $f : [a, b] \rightarrow E$ heißt Riemann-integrierbar, falls die Riemannschen Summen für f einen Grenzwert besitzen. Den Grenzwert

$$\int_a^b \mathcal{R} f(x) dx := \lim_{\mathcal{P}} S(f, \mathcal{P}, \xi)$$

nennen wir dann das Riemann-Integral von f über $[a, b]$.

$\mathcal{R}([a, b], E)$ bezeichnet die Menge der Riemann-integrierbaren Funktionen von $[a, b]$ in E .

Bemerkung:

1. Für eine Funktion $f : [a, b] \rightarrow E$, die sowohl eine Stammfunktion besitzt als auch Riemann-integrierbar ist, stimmen das Riemann-Integral und das bestimmte Integral überein, d.h.,

$$\int_a^b \mathcal{R} f(x) dx = \int_a^b f(x) dx.$$

(siehe Übungsaufgabe 95). Wir lassen deshalb im Folgenden \mathcal{R} im Integralzeichen weg. Im Abschnitt 7.2 sei mit \int immer das Riemann-Integral gemeint!

2. Ist $f : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar, $f \geq 0$ und A die Menge

$$A := \{(x, y) \in \mathbb{R}^2 \mid a \leq x \leq b, 0 \leq y \leq f(x)\}.$$

Dann ist $\mu(A) := \int_a^b f(x) dx$ der Flächeninhalt von A . (Man benutzt dies als Definition für den Flächeninhalt. Dies entspricht der geometrischen Vorstellung und Motivation).

3. Es gibt Funktionen, die Stammfunktionen besitzen, aber nicht Riemann-integrierbar sind. Wir betrachten als Beispiel 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 Riemann-integrierbar (Übungsaufgabe 94).

4. Es gibt Funktionen, die Riemann-integrierbar sind, aber keine Stammfunktion besitzen. Sei $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ eine Zerlegung des Intervalls $[a, b]$ und seien $e_1, \dots, e_n \in E$ Vektoren eines Banachraumes E . Eine Funktion $f : [a, b] \rightarrow E$ mit der Eigenschaft

$$f|_{(x_{k-1}, x_k)} = e_k \quad k = 1, 2, \dots, n \quad (*)$$

heißt Treppenfunktion (die Werte $f(x_k)$ können beliebig sein).

Nicht-konstante Treppenfunktionen besitzen keine Stammfunktion, sind aber Riemann-integrierbar.

Als nächstes wollen wir untersuchen, welche Funktionen $f : [a, b] \rightarrow E$ Riemann-integrierbar sind. Zunächst folgt eine notwendige Bedingung.

Satz 7.6 *Ist $f : [a, b] \rightarrow E$ Riemann-integrierbar, so ist f beschränkt.*

Beweis. Sei $e := \int_a^b f(x) dx$. Nach Definition existiert eine Zerlegung \mathcal{P} von $[a, b]$, so dass für beliebige Stützstellen ξ von \mathcal{P} gilt

$$\|S(f, \mathcal{P}, \xi) - e\| < 1.$$

Wir betrachten ein Teilintervall I_ν von \mathcal{P} . Für $\mu \neq \nu$ fixieren wir Zahlen $\xi_\mu \in I_\mu$ und definieren

$$M_\nu := \sum_{\mu \neq \nu} L(I_\mu) \cdot f(\xi_\mu) \in E.$$

Für jedes $x \in I_\nu$ gilt dann $\|L(I_\nu) \cdot f(x) + M_\nu - e\| < 1$. Mit der Dreiecksungleichung folgt

$$\|f(x)\| \leq \left\| f(x) + \frac{M_\nu - e}{L(I_\nu)} \right\| + \left\| \frac{M_\nu - e}{L(I_\nu)} \right\| < \frac{1}{L(I_\nu)} + \left\| \frac{M_\nu - e}{L(I_\nu)} \right\|$$

für alle $x \in I_\nu$. Sei nun $M := \max \left\{ \frac{1}{L(I_\nu)} (1 + \|M_\nu - e\|) \mid \nu = 1, \dots, n \right\}$. Dann ist $\|f(x)\| \leq M$ für alle $x \in [a, b]$. \square

Satz 7.7 (Cauchysches Integrierbarkeitskriterium)

Sei $f : [a, b] \subset \mathbb{R} \rightarrow E$, eine Funktion mit Werten in einem Banachraum E . Dann gilt: f ist genau dann Riemann-integrierbar, wenn zu jedem $\varepsilon > 0$ eine Zerlegung \mathcal{P}_ε von $[a, b]$ existiert, so dass für beliebige Verfeinerungen $\mathcal{P}, \tilde{\mathcal{P}}$ von \mathcal{P}_ε und Stützstellen $\xi, \tilde{\xi}$ von \mathcal{P} bzw. $\tilde{\mathcal{P}}$ gilt

$$\|S(f, \mathcal{P}, \xi) - S(f, \tilde{\mathcal{P}}, \tilde{\xi})\| < \varepsilon.$$

Beweis. (\implies) Sei $f : [a, b] \rightarrow E$ Riemann-integrierbar und $\varepsilon > 0$ gegeben. Dann existieren $e \in E$ und \mathcal{P}_ε , so dass gilt

$$\|S(f, \mathcal{P}, \xi) - e\| < \frac{\varepsilon}{2} \quad \forall \mathcal{P} \geq \mathcal{P}_\varepsilon.$$

Für $\mathcal{P}, \tilde{\mathcal{P}} \geq \mathcal{P}_\varepsilon$ folgt dann

$$\|S(f, \mathcal{P}, \xi) - S(f, \tilde{\mathcal{P}}, \tilde{\xi})\| \leq \|S(f, \mathcal{P}, \xi) - e\| + \|S(f, \tilde{\mathcal{P}}, \tilde{\xi}) - e\| \leq \varepsilon.$$

(\impliedby) Sei (ε_n) eine Nullfolge positiver Zahlen. Nach Voraussetzung existiert für jedes $n \in \mathbb{N}$ eine Zerlegung \mathcal{P}_n von $[a, b]$ so dass

$$\|S(f, \mathcal{P}, \xi) - S(f, \tilde{\mathcal{P}}, \tilde{\xi})\| < \varepsilon_n \quad \forall \mathcal{P}, \tilde{\mathcal{P}} \geq \mathcal{P}_n. \quad (*)$$

Wir bilden nun sukzessive die folgenden Zerlegungen \mathcal{U}_n von $[a, b]$:

$$\begin{aligned} \mathcal{U}_1 &:= \mathcal{P}_1, \\ \mathcal{U}_2 &:= \mathcal{P}_1 \cup \mathcal{P}_2 = \mathcal{U}_1 \cup \mathcal{P}_2, \\ \mathcal{U}_3 &:= \mathcal{P}_1 \cup \mathcal{P}_2 \cup \mathcal{P}_3 = \mathcal{U}_2 \cup \mathcal{P}_3, \quad \dots \\ \mathcal{U}_n &:= \mathcal{P}_1 \cup \dots \cup \mathcal{P}_n = \mathcal{U}_{n-1} \cup \mathcal{P}_n, \quad \dots \end{aligned}$$

Dann ist $\mathcal{P}_n \leq \mathcal{U}_n \leq \mathcal{U}_{n+1} \leq \dots$. Zu jeder Zerlegung \mathcal{U}_n wählen wir ein Tupel von Stützstellen ξ_n . Sei $\varepsilon > 0$ gegeben. Dann existiert ein $n_0 \in \mathbb{N}$ so dass $\varepsilon_n < \frac{\varepsilon}{2}$ für alle $n \geq n_0$. Aus (*) folgt dann für alle $m \geq n \geq n_0$

$$\|S(f, \mathcal{U}_n, \xi_n) - S(f, \mathcal{U}_m, \xi_m)\| < \varepsilon_n < \frac{\varepsilon}{2}. \quad (**)$$

Folglich bilden die Vektoren $(S(f, \mathcal{U}_n, \xi_n))_{n=1}^\infty$ eine Cauchy-Folge im Banachraum E . Da E vollständig ist, existiert $e := \lim_{n \rightarrow \infty} S(f, \mathcal{U}_n, \xi_n)$. Gehen wir in (**) mit m gegen ∞ , so erhalten wir

$$\|S(f, \mathcal{U}_n, \xi_n) - e\| \leq \frac{\varepsilon}{2} \quad \forall n \geq n_0.$$

Sei nun \mathcal{P} eine beliebige Zerlegung mit $\mathcal{P} \geq \mathcal{U}_{n_0}$. Da $\mathcal{P}, \mathcal{U}_{n_0} \geq \mathcal{P}_{n_0}$, erhalten wir mit (*)

$$\begin{aligned} \|S(f, \mathcal{P}, \xi) - e\| &\leq \|S(f, \mathcal{P}, \xi) - S(f, \mathcal{U}_{n_0}, \xi_{n_0})\| + \|S(f, \mathcal{U}_{n_0}, \xi_{n_0}) - e\| \\ &\stackrel{(*)}{<} \varepsilon_{n_0} + \frac{\varepsilon}{2} < \varepsilon. \end{aligned}$$

Folglich ist die Funktion f Riemann-integrierbar. □

Satz 7.8 Jede stetige Abbildung $f : [a, b] \subset \mathbb{R} \rightarrow E$ mit Werten in einem Banachraum E ist Riemann-integrierbar.

Beweis. Wir werden das Cauchysche Integrierbarkeitskriterium anwenden. Da f stetig und $[a, b]$ kompakt ist, ist f gleichmäßig stetig, das heißt, für alle $\varepsilon > 0$ existiert ein $\delta > 0$, so dass für alle $x_1, x_2 \in [a, b]$ gilt:

$$|x_1 - x_2| < \delta \implies \|f(x_1) - f(x_2)\| < \frac{\varepsilon}{2(b-a)}.$$

Sei \mathcal{P}_0 eine Zerlegung von $[a, b]$ mit $\|\mathcal{P}_0\| < \delta$. Dann gilt für alle $\tilde{\mathcal{P}} \geq \mathcal{P}_0 = \{x_0, x_1, \dots, x_n\}$, wobei $\tilde{\mathcal{P}} = \{y_0, y_1, \dots, y_N\}$ die Punkte x_i enthält

$$\begin{aligned} \|S(f, \tilde{\mathcal{P}}, \tilde{\xi}) - S(f, \mathcal{P}_0, \xi_0)\| &= \left\| \sum_{l=1}^N (y_l - y_{l-1}) \cdot f(\tilde{\xi}_l) - \sum_{k=1}^n (x_k - x_{k-1}) \cdot f(\xi_k) \right\| \\ &= \left\| \sum_{l=1}^N (y_l - y_{l-1}) \cdot (f(\tilde{\xi}_l) - f(\xi_{k_l})) \right\| \quad \text{wobei } \tilde{I}_l \subset I_{k_l} \\ &\leq \sum_{l=1}^N \underbrace{\|f(\tilde{\xi}_l) - f(\xi_{k_l})\|}_{< \frac{\varepsilon}{2(b-a)}} \cdot (y_l - y_{l-1}) \\ &< \frac{\varepsilon}{2(b-a)} \cdot (b-a) = \frac{\varepsilon}{2}. \end{aligned}$$

Seien nun $\tilde{\mathcal{P}}, \tilde{\tilde{\mathcal{P}}} \geq \mathcal{P}_0$. Die Dreiecksungleichung liefert

$$\begin{aligned} \|S(f, \tilde{\tilde{\mathcal{P}}}, \tilde{\tilde{\xi}}) - S(f, \tilde{\mathcal{P}}, \tilde{\xi})\| &\leq \|S(f, \tilde{\tilde{\mathcal{P}}}, \tilde{\tilde{\xi}}) - S(f, \mathcal{P}_0, \xi)\| + \|S(f, \tilde{\mathcal{P}}, \tilde{\xi}) - S(f, \mathcal{P}_0, \xi)\| \\ &< \varepsilon. \end{aligned}$$

Nach dem Cauchyschen Integrierbarkeitskriterium ist f Riemann-integrierbar. \square

Der nächste Satz zeigt, dass man das Riemann-Integral für eine Riemann-integrierbare Funktion als Grenzwert einer beliebigen Folge von Riemannschen Summen ausrechnen kann.

Satz 7.9 Sei $f : [a, b] \rightarrow E$ eine Riemann-integrierbare Funktion, (\mathcal{P}_n) eine Folge von Zerlegungen von $[a, b]$ mit $\|\mathcal{P}_n\| \rightarrow 0$ und ξ_n Stützstellen von \mathcal{P}_n . Dann gilt

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} S(f, \mathcal{P}_n, \xi_n).$$

Beweis. Nach Satz 7.6 ist f beschränkt, also existiert ein $M > 0$ so dass $\|f(x)\| \leq M$ für alle $x \in [a, b]$. Sei $\varepsilon > 0$. Da f Riemann-integrierbar ist, existiert eine Zerlegung \mathcal{P}^* von $[a, b]$, so dass

$$\left\| S(f, \mathcal{P}, \xi) - \int_a^b f(x) dx \right\| < \frac{\varepsilon}{2}$$

für alle Verfeinerungen $\mathcal{P} \geq \mathcal{P}^*$. Die Zerlegung \mathcal{P}^* habe r Teilintervalle. Da $\|\mathcal{P}_n\| \rightarrow 0$, existiert ein $n_0 \in \mathbb{N}$, so dass

$$2Mr\|\mathcal{P}_n\| < \frac{\varepsilon}{2} \quad \forall n \geq n_0.$$

Wir betrachten nun \mathcal{P}_n mit $n \geq n_0$. Seien J_1, J_2, \dots, J_l diejenigen Teilintervalle von \mathcal{P}_n , die nicht in $\mathcal{P}^* \cup \mathcal{P}_n$ vorkommen, das heißt diejenigen, die durch Punkte von \mathcal{P}^* zerlegt werden. Dann ist $l \leq r$. Wir fügen nun zu den Stützstellen ξ_n von \mathcal{P}_n neue Stützstellen hinzu, so dass Stützstellen $\tilde{\xi}_n$ von $\mathcal{P}^* \cup \mathcal{P}_n$ entstehen. Damit erhalten wir

$$\|S(f, \mathcal{P}_n, \xi_n) - S(f, \mathcal{P}_n \cup \mathcal{P}^*, \tilde{\xi}_n)\| \leq 2M \sum_{v=1}^l L(J_v) \leq 2Ml\|\mathcal{P}_n\| \leq 2Mr\|\mathcal{P}_n\| < \frac{\varepsilon}{2}.$$

Da $\mathcal{P}_n \cup \mathcal{P}^* \geq \mathcal{P}^*$, folgt

$$\begin{aligned} \left\| S(f, \mathcal{P}_n, \xi_n) - \int_a^b f(x) dx \right\| &\leq \|S(f, \mathcal{P}_n, \xi_n) - S(f, \mathcal{P}_n \cup \mathcal{P}^*, \tilde{\xi}_n)\| + \\ &\quad + \left\| S(f, \mathcal{P}_n \cup \mathcal{P}^*, \tilde{\xi}_n) - \int_a^b f(x) dx \right\| \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Dies gilt für alle $n \geq n_0$. Somit folgt $\lim_{n \rightarrow \infty} S(f, \mathcal{P}_n, \xi_n) = \int_a^b f(x) dx$. \square

Satz 7.9 ist gut geeignet, um

1. Rechenregeln für Riemann-Integrale zu beweisen,
2. Abschätzungen für Summen auf Abschätzungen für Integrale zu übertragen,
3. Grenzwerte zu berechnen.

Beispiel: Sei $\alpha > 0$ eine positive reelle Zahl. Wir wollen den Grenzwert

$$\lim_{n \rightarrow \infty} \frac{1^\alpha + 2^\alpha + \dots + n^\alpha}{n^{\alpha+1}} = \lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{k^\alpha}{n^\alpha} \cdot \frac{1}{n}$$

berechnen. Dazu betrachten wir die Funktion $f(x) := x^\alpha$ auf $[0, 1]$ und die Zerlegung $\mathcal{P}_n = \{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\}$. f ist stetig, also Riemann-integrierbar und $\|\mathcal{P}_n\| = \frac{1}{n}$. Dann ist $f(\frac{k}{n}) = \frac{k^\alpha}{n^\alpha}$ und $\|\mathcal{P}_n\| \rightarrow 0$. Wir können also Satz 7.9 anwenden. f hat auch eine Stammfunktion, deshalb stimmt das Riemann-Integral mit dem bestimmten Integral überein (siehe Übungsaufgabe 95). Daraus folgt $\int_0^1 f(x) dx = \int_0^1 x^\alpha dx = \frac{1}{\alpha+1}$ und somit

$$\frac{1}{\alpha+1} = \int_0^1 f(x) dx \stackrel{\text{Satz 7.9}}{=} \lim_{n \rightarrow \infty} \sum_{k=1}^n f\left(\frac{k}{n}\right) \cdot \frac{1}{n} = \lim_{n \rightarrow \infty} \sum_{k=1}^n \frac{k^\alpha}{n^\alpha} \cdot \frac{1}{n}.$$

Satz 7.10 (Rechenregeln für das Riemann-Integral)

1. Seien $f, g \in \mathcal{R}([a, b], E)$ und $\mu, \lambda \in \mathbb{K}$. Dann ist $\mu f + \lambda g \in \mathcal{R}([a, b], E)$ und es gilt

$$\int_a^b (\mu f(x) + \lambda g(x)) dx = \mu \int_a^b f(x) dx + \lambda \int_a^b g(x) dx.$$

2. Ist $f \in \mathcal{R}([a, c], E) \cup \mathcal{R}([c, b], E)$, so ist $f \in \mathcal{R}([a, b], E)$ und es gilt

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

3. Ist $f \in \mathcal{R}([a, b], E)$, $[\alpha, \beta] \subset [a, b]$, so ist $f \in \mathcal{R}([\alpha, \beta], E)$.

4. Sei $f = (f_1, f_2) : [a, b] \rightarrow E := E_1 \times E_2$ eine Abbildung in einen Produktraum. Dann ist $f \in \mathcal{R}([a, b], E)$ genau dann, wenn $f_i \in \mathcal{R}([a, b], E_i)$ für $i = 1, 2$. In diesem Fall gilt

$$\int_a^b f(x) dx = \left(\int_a^b f_1(x) dx, \int_a^b f_2(x) dx \right).$$

5. Ist $E = \mathbb{C}$ und $f \in \mathcal{R}([a, b], \mathbb{C})$, so ist auch $\bar{f} \in \mathcal{R}([a, b], \mathbb{C})$ und es gilt

$$\overline{\int_a^b f(x) dx} = \int_a^b \overline{f(x)} dx.$$

6. Sei $E = \mathbb{R}$, $f, g \in \mathcal{R}([a, b], \mathbb{R})$ und $f \leq g$. Dann gilt

$$\int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

7. Seien $f, g \in \mathcal{R}([a, b], E)$. Stimmen f und g auf einer dichten Teilmenge von $[a, b]$ überein, so gilt

$$\int_a^b f(x) dx = \int_a^b g(x) dx.$$

8. Ist $f \in C([a, b], E)$ stetig, dann gilt

$$\left\| \int_a^b f(x) dx \right\| \leq \int_a^b \|f(x)\| dx \leq \|f\|_\infty \cdot (b - a).$$

Dies ist die Fundamentalungleichung für Riemann-Integrale.

Beweis. Den Beweis lassen wir als Übungsaufgabe. Man benutzt den Satz 7.9 und die entsprechenden Eigenschaften für Summen. \square

Als nächstes befassen wir uns mit Funktionenfolgen und untersuchen, unter welchen Bedingungen die Grenzfunktion einer Folge Riemann-integrierbarer Funktionen wieder Riemann-integrierbar ist.

Satz 7.11 (Vertauschbarkeit von Limes und Integral)

Sei (f_n) mit $f_n : [a, b] \rightarrow E$ eine Folge Riemann-integrierbarer Funktionen und konvergiere (f_n) gleichmäßig gegen f . Dann ist f Riemann-integrierbar und es gilt

$$\int_a^b f(x) dx = \lim_{n \rightarrow \infty} \int_a^b f_n(x) dx.$$

Beweis. Da (f_n) gleichmäßig gegen f konvergiert, existiert für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ (von ε abhängig), so dass

$$\|f(x) - f_n(x)\| < \frac{\varepsilon}{4} \cdot \frac{1}{b-a} \quad (*)$$

für alle $n \geq n_0$ und für alle $x \in [a, b]$.

(1) Wir zeigen zunächst, dass f Riemann-integrierbar ist. Da f_{n_0} Riemann-integrierbar ist, existiert eine Zerlegung \mathcal{P}_0 , so dass

$$\|S(f_{n_0}, \mathcal{P}, \xi) - S(f_{n_0}, \tilde{\mathcal{P}}, \tilde{\xi})\| < \frac{\varepsilon}{2}$$

für alle Verfeinerungen $\mathcal{P}, \tilde{\mathcal{P}} \geq \mathcal{P}_0$. Damit folgt

$$\begin{aligned} & \|S(f, \mathcal{P}, \xi) - S(f, \tilde{\mathcal{P}}, \tilde{\xi})\| \\ & \leq \|S(f, \mathcal{P}, \xi) - S(f_{n_0}, \mathcal{P}, \xi)\| + \|S(f_{n_0}, \mathcal{P}, \xi) - S(f_{n_0}, \tilde{\mathcal{P}}, \tilde{\xi})\| \\ & \quad + \|S(f_{n_0}, \tilde{\mathcal{P}}, \tilde{\xi}) - S(f, \tilde{\mathcal{P}}, \tilde{\xi})\| \\ & \leq \left\| \sum_k L(I_k)(f(\xi_k) - f_{n_0}(\xi_k)) \right\| + \frac{\varepsilon}{2} + \left\| \sum_l L(\tilde{I}_l)(f(\tilde{\xi}_l) - f_{n_0}(\tilde{\xi}_l)) \right\| \\ & \stackrel{(*)}{<} 2 \cdot \frac{\varepsilon}{4} \cdot \frac{1}{b-a} \cdot (b-a) + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Folglich ist f Riemann-integrierbar.

(2) Als zweites beweisen wir die Vertauschbarkeit von Limes und Integral. Wir benutzen wieder den Satz 7.9. Sei (\mathcal{P}_m) eine Folge von Zerlegungen von $[a, b]$ mit $\|\mathcal{P}_m\| \rightarrow 0$ und seien ξ_m Stützstellen von \mathcal{P}_m . Sei $\varepsilon > 0$. Nach Satz 7.9 existieren m_0 und $m_0(n)$, so dass

$$\left\| \int_a^b f(x) dx - S(f, \mathcal{P}_m, \xi_m) \right\| < \frac{\varepsilon}{4} \quad \text{und} \quad \left\| \int_a^b f_n(x) dx - S(f_n, \mathcal{P}_m, \xi_m) \right\| < \frac{\varepsilon}{4}$$

für alle $m \geq \max(m_0, m_0(n))$. Sei nun $n \geq n_0$ und $m \geq \max\{m_0, m_0(n)\}$. Dann folgt

$$\begin{aligned} & \left\| \int_a^b f(x) dx - \int_a^b f_n(x) dx \right\| \\ & \leq \left\| \int_a^b f(x) dx - S(f, \mathcal{P}_m, \xi_m) \right\| + \|S(f, \mathcal{P}_m, \xi_m) - S(f_n, \mathcal{P}_m, \xi_m)\| \\ & \quad + \left\| S(f_n, \mathcal{P}_m, \xi_m) - \int_a^b f_n(x) dx \right\| \\ & < \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} < \varepsilon. \end{aligned}$$

Somit gilt $\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx$. □

Als Spezialfall erhalten wir

Satz 7.12 Sei (f_n) mit $f_n : [a, b] \rightarrow \mathbb{K}$ eine Folge Riemann-integrierbarer Funktionen und konvergiere die Funktionenreihe $\sum_{n=1}^{\infty} f_n(x)$ auf $[a, b]$ gleichmäßig gegen $f : [a, b] \rightarrow \mathbb{K}$. Dann ist f Riemann-integrierbar und es gilt

$$\int_a^b f(x) dx = \sum_{n=1}^{\infty} \int_a^b f_n(x) dx. \quad \square$$

Die gleichmäßige Konvergenz einer Funktionenfolge ist eine sehr starke Forderung, die oft nicht erfüllt ist. Für das allgemeinere Lebesgue-Integral gilt die Vertauschbarkeit von Limes und Integral unter wesentlich schwächeren Bedingungen (siehe Vorlesung Analysis III).

Anwendung: *Integralberechnung für elliptische Integrale*

Oft sind Integrale nicht durch elementare Funktionen auszudrücken, wie zum Beispiel elliptische Integrale. Dann kann man sie durch eine Reihendarstellung näherungsweise berechnen.

Wir demonstrieren dies am Beispiel des (Riemann-)Integrals

$$I_k := \int_0^{\frac{\pi}{2}} \frac{1}{\sqrt{1 - k^2 \sin^2 x}} dx,$$

wobei $|k| < 1$. Da $f : [0, \frac{\pi}{2}] \rightarrow \mathbb{R}$ stetig ist, existiert dieses Integral.

Wir betrachten die Taylorreihe der Funktion $h(y) := \frac{1}{\sqrt{1+y}}$ (siehe Beispiel 4 in Kapitel 5.5). Für $|y| < 1$ gilt

$$\frac{1}{\sqrt{1+y}} = 1 - \frac{1}{2}y + \frac{1 \cdot 3}{2 \cdot 4}y^2 - \frac{1 \cdot 3 \cdot 5}{2 \cdot 4 \cdot 6}y^3 \pm \dots$$

Daraus folgt mit $y = -k^2 \sin^2 x$ für $x \in [0, \frac{\pi}{2}]$

$$\frac{1}{\sqrt{1 - k^2 \sin^2 x}} = 1 + \frac{1}{2}k^2 \sin^2 x + \frac{1 \cdot 3}{2 \cdot 4}k^4 \sin^4 x + \dots \quad (*)$$

Diese Reihe hat für $|k| < 1$ eine konvergente Majorante (die Reihe mit dem Grenzwert $\frac{1}{\sqrt{1-k^2}}$). Nach dem Weierstraßschen Majorantenkriterium für Funktionenreihen ist die Reihe (*) auf $[0, \frac{\pi}{2}]$ gleichmäßig konvergent. Mit Satz 7.12 folgt daher

$$\begin{aligned}
 I_k &= \int_0^{\frac{\pi}{2}} \frac{dx}{\sqrt{1 - k^2 \sin^2 x}} \\
 &= \int_0^{\frac{\pi}{2}} dx + \sum_{n=1}^{\infty} \int_0^{\frac{\pi}{2}} \frac{1 \cdot 3 \cdot \dots \cdot (2n-1)}{2 \cdot 4 \cdot \dots \cdot (2n)} k^{2n} \sin^{2n} x \, dx \\
 &= \frac{\pi}{2} + \sum_{n=1}^{\infty} \frac{1 \cdot 3 \cdot \dots \cdot (2n-1)}{2 \cdot 4 \cdot \dots \cdot (2n)} k^{2n} \int_0^{\frac{\pi}{2}} \sin^{2n} x \, dx \\
 &= \frac{\pi}{2} \left(1 + \sum_{n=1}^{\infty} \left(\frac{1 \cdot 3 \cdot \dots \cdot (2n-1)}{2 \cdot 4 \cdot \dots \cdot (2n)} \right)^2 k^{2n} \right) \quad (\text{siehe Bsp. 2' aus Kap. 7.1}).
 \end{aligned}$$

Im letzten Schritt haben wir benutzt, dass das Riemann-Integral für die Funktion $h(x) = \sin^{2n}(x)$ auf $[0, \pi/2]$ gleich dem bestimmten Integral ist. Dies ist möglich, da h sowohl Riemann-integrierbar ist als auch eine Stammfunktion besitzt (siehe Übungsaufgaben).

Abschließend leiten wir spezielle Integrierbarkeitskriterien für *reellwertige* Funktionen her. Sei $f : [a, b] \rightarrow \mathbb{R}$ eine *beschränkte* Funktion und $\mathcal{P} = \{x_0, \dots, x_n\}$ eine Zerlegung von $[a, b]$. Wir definieren

$$m_k := \inf\{f(x) \mid x \in I_k\} \quad \text{und} \quad M_k := \sup\{f(x) \mid x \in I_k\}$$

sowie

$$\begin{aligned}
 \underline{S}(f, \mathcal{P}) &:= \sum_{k=1}^n m_k \cdot L(I_k) \quad \text{Untersumme von } f \text{ bzgl. } \mathcal{P}, \\
 \overline{S}(f, \mathcal{P}) &:= \sum_{k=1}^n M_k \cdot L(I_k) \quad \text{Obersumme von } f \text{ bzgl. } \mathcal{P}.
 \end{aligned}$$

Dann folgt:

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 $\tilde{\mathcal{P}}$ gilt $\underline{S}(f, \mathcal{P}) \leq \overline{S}(f, \tilde{\mathcal{P}})$.

Dies zeigt die Existenz von

$$\underbrace{\int_a^b f(x) \, dx := \sup_{\mathcal{P}} \underline{S}(f, \mathcal{P})}_{=: \text{unteres Riemann Integral}} \quad \text{und} \quad \underbrace{\int_a^b f(x) \, dx := \inf_{\mathcal{P}} \overline{S}(f, \mathcal{P})}_{=: \text{oberes Riemann-Integral}}$$

wobei

$$\int_a^b f(x) \, dx \leq \int_a^b f(x) \, dx.$$

Satz 7.13 (Riemannsches Integrierbarkeitskriterium)

Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion. Dann sind die folgenden Bedingungen äquivalent:

1. f ist Riemann-integrierbar.
2. Für alle $\varepsilon > 0$ existiert eine Zerlegung \mathcal{P}_ε mit $\overline{S}(f, \mathcal{P}_\varepsilon) - \underline{S}(f, \mathcal{P}_\varepsilon) < \varepsilon$.
3. $\int_{\overline{a}}^b f(x) dx = \int_a^b f(x) dx$.

In diesem Fall gilt $\int_a^b f(x) dx = \int_a^b f(x) dx = \int_{\overline{a}}^b f(x) dx$.

Beweis. (1) \implies (2): Sei f Riemann-integrierbar. Nach dem Cauchy-Kriterium aus Satz 7.7 existiert zu jedem $\varepsilon > 0$ eine Zerlegung \mathcal{P}_ε mit

$$S(f, \mathcal{P}_\varepsilon, \xi) - S(f, \mathcal{P}_\varepsilon, \tilde{\xi}) < \frac{\varepsilon}{2}, \quad (*)$$

wobei $\xi = (\xi_1, \dots, \xi_r)$ und $\tilde{\xi} = (\tilde{\xi}_1, \dots, \tilde{\xi}_r)$ beliebige Stützstellen von \mathcal{P}_ε mit $f(\xi_k) \geq f(\tilde{\xi}_k)$ für $k = 1, \dots, r$ sind. Wir wählen nun Folgen von Stützstellen (ξ_n) und $(\tilde{\xi}_n)$, für die $f(\xi_{kn}) \rightarrow M_k$ und $f(\tilde{\xi}_{kn}) \rightarrow m_k$ für $k = 1, \dots, r$ gilt, setzen diese in (*) ein und gehen mit $n \rightarrow \infty$. Dann folgt

$$\overline{S}(f, \mathcal{P}_\varepsilon) - \underline{S}(f, \mathcal{P}_\varepsilon) \leq \frac{\varepsilon}{2} < \varepsilon.$$

(2) \implies (3): Für alle $\varepsilon > 0$ existiere eine Zerlegung \mathcal{P}_ε , so dass $\overline{S}(f, \mathcal{P}_\varepsilon) - \underline{S}(f, \mathcal{P}_\varepsilon) < \varepsilon$. Es gilt

$$\overline{S}(f, \mathcal{P}_\varepsilon) \geq \int_a^b f(x) dx \geq \int_{\overline{a}}^b f(x) dx \geq \underline{S}(f, \mathcal{P}_\varepsilon)$$

und daher

$$0 \leq \int_a^b f(x) dx - \int_{\overline{a}}^b f(x) dx < \varepsilon$$

für alle $\varepsilon > 0$. Dies zeigt die Gleichheit von unterem und oberem Riemann-Integral.

(3) \implies (1): Sei $\int_a^b f(x) dx = \int_{\overline{a}}^b f(x) dx =: e$. Dann existieren für jedes $\varepsilon > 0$ Zerlegungen \mathcal{P} und $\tilde{\mathcal{P}}$ mit

$$\underline{S}(f, \mathcal{P}) > e - \frac{\varepsilon}{2} \quad \text{und} \quad \overline{S}(f, \tilde{\mathcal{P}}) < e + \frac{\varepsilon}{2}$$

Folglich ist $\overline{S}(f, \tilde{\mathcal{P}}) - \underline{S}(f, \mathcal{P}) < \varepsilon$. Sei nun $\mathcal{P}_\varepsilon = \mathcal{P} \cup \tilde{\mathcal{P}}$. Dann gilt

$$\overline{S}(f, \mathcal{P}_\varepsilon) - \underline{S}(f, \mathcal{P}_\varepsilon) \leq \overline{S}(f, \tilde{\mathcal{P}}) - \underline{S}(f, \mathcal{P}) < \varepsilon.$$

Ist nun \mathcal{P} eine Zerlegung von $[a, b]$ mit $\mathcal{P} \geq \mathcal{P}_\varepsilon$. Dann liegen sowohl e als auch die Riemannschen Summen $S(f, \mathcal{P}, \xi)$ für beliebige Stützstellen im Intervall $(\underline{S}(f, \mathcal{P}_\varepsilon), \overline{S}(f, \mathcal{P}_\varepsilon))$.

Folglich gilt $\|S(f, \mathcal{P}, \xi) - e\| < \varepsilon$. e ist also der Grenzwert der Riemannschen Summen für f , d.h. f ist Riemann-integrierbar und

$$\int_a^b f(x) dx = \int_a^b f(x) dx = \int_{\bar{a}}^b f(x) dx.$$

□

Unser Ziel ist es nun zu zeigen, dass $f : [a, b] \rightarrow \mathbb{R}$ genau dann Riemann-integrierbar ist, wenn f beschränkt und "fast überall" stetig ist. Zunächst klären wir, was mit "fast überall" gemeint ist.

Definition 7.8. Eine Teilmenge $A \subset \mathbb{R}$ hat das Lebesgue-Maß Null (wir sagen auch kurz: A ist eine Nullmenge), wenn für alle $\varepsilon > 0$ eine Folge von Intervallen I_1, I_2, \dots existiert, so dass

$$A \subset \bigcup_{n=1}^{\infty} \text{Int}(I_n) \quad \text{und} \quad \sum_{n=1}^{\infty} L(I_n) < \varepsilon.$$

Man sagt: Eine Eigenschaft für Punkte einer Menge $B \subset \mathbb{R}$ gilt "fast überall" wenn sie für alle Punkte von $(B \setminus \text{Nullmenge})$ gilt.

Offensichtlich gilt:

1. Die einpunktigen Mengen $\{x\} \subset \mathbb{R}$ sind Nullmengen.
2. $B \subset A$ und A Nullmenge $\implies B$ Nullmenge.
3. Sei $A \subset [a, b]$ eine Nullmenge. Dann ist $[a, b] \setminus A \subset [a, b]$ eine dichte Teilmenge.

Satz 7.14

1. Seien $A_1, A_2, \dots \subset \mathbb{R}$ Nullmengen. Dann ist $\bigcup_{n=1}^{\infty} A_n$ ebenfalls eine Nullmenge.
2. Jede abzählbare Menge $A \subset \mathbb{R}$ ist eine Nullmenge.

Beweis. (1) Sei $\varepsilon > 0$. Da A_n Nullmengen für alle $n \in \mathbb{N}$, existieren Intervalle I_{n1}, I_{n2}, \dots mit

$$A_n \subset \bigcup_{i=1}^{\infty} \text{Int}(I_{ni}) \quad \text{und} \quad \sum_{i=1}^{\infty} L(I_{ni}) < \frac{\varepsilon}{2^n}.$$

Daraus folgt

$$A = \bigcup_{n=1}^{\infty} A_n \subset \bigcup_{n=1}^{\infty} \bigcup_{i=1}^{\infty} \text{Int}(I_{ni}) \quad (\text{abzählbar viele Mengen})$$

und

$$\sum_{n=1}^{\infty} \sum_{i=1}^{\infty} L(I_{ni}) < \varepsilon \cdot \sum_{n=1}^{\infty} \frac{1}{2^n} = \varepsilon \cdot \frac{1}{2} \cdot \frac{1}{1 - \frac{1}{2}} = \varepsilon.$$

(2) Folgt aus (1), da einpunktige Mengen Nullmengen sind. □

Satz 7.15 (Lebesguesches Integrierbarkeitskriterium)

Eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn sie beschränkt und fast überall stetig ist.

Beweis. (\Leftarrow) Sei $f : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}$ beschränkt und fast überall stetig. Dann existiert ein $C \in \mathbb{R}^+$ mit $|f(x)| < C$ für alle $x \in [a, b]$. Wir bezeichnen die Menge der Unstetigkeitsstellen von f mit A , d.h.

$$A := \{x_0 \in [a, b] \mid f \text{ in } x_0 \text{ nicht stetig}\}.$$

Sei $\varepsilon > 0$. Da A eine Nullmenge ist, existieren Intervalle J_1, J_2, \dots , so dass

$$A \subset \bigcup_{n=1}^{\infty} \text{Int}(J_n) \quad \text{und} \quad \sum_{n=1}^{\infty} L(J_n) < \varepsilon.$$

Wir betrachten nun die sogenannte *Oszillation von f in $x_0 \in [a, b]$* :

$$\text{osz}(f, x_0) := \lim_{\delta \rightarrow 0^+} (\sup\{f(x) \mid |x - x_0| < \delta\} - \inf\{f(x) \mid |x - x_0| < \delta\}).$$

Mit Hilfe der Oszillation kann man die Stetigkeit folgendermaßen charakterisieren: f ist in x_0 genau dann stetig, wenn $\text{osz}(f, x_0) = 0$. Für die Punkte $x_0 \in [a, b] \setminus A$ (in denen f ja stetig ist) existiert somit ein $\delta(x_0) > 0$, so dass

$$\sup\{f(x) \mid |x - x_0| < \delta(x_0)\} - \inf\{f(x) \mid |x - x_0| < \delta(x_0)\} < \varepsilon.$$

Die Familie $\mathcal{U} := \{\text{Int}(J_n) \mid n \in \mathbb{N}\} \cup \{K(x_0, \delta(x_0)) \mid x_0 \in [a, b] \setminus A\}$ ist eine offene Überdeckung der kompakten Menge $[a, b]$. Dann existiert eine endliche Teilüberdeckung

$$\mathcal{U}^* := \{\text{Int}(J_{n_1}), \dots, \text{Int}(J_{n_N}), K(x_1, \delta(x_1)), \dots, K(x_r, \delta(x_r))\}$$

von $[a, b]$. Wir wählen nun eine Zerlegung $\mathcal{P} = \{x_0, x_1, \dots, x_n\}$ von $[a, b]$, deren Teilintervalle I_k in $\text{Int}(J_{n_j})$ oder in $K(x_i, \delta(x_i))$ liegen, wobei $i = 1, \dots, r$ und $j = 1, \dots, N$. Wir bezeichnen mit J^* die Menge derjenigen Teilintervalle von \mathcal{P} , die in einem J_{n_j} liegen und mit J^{**} die Menge der restlichen Teilintervallen von \mathcal{P} . Dann gilt wegen $|f(x)| \leq C$ die folgende Ungleichung

$$\begin{aligned} \overline{S}(f, \mathcal{P}) - \underline{S}(f, \mathcal{P}) &= \sum_{I_k \in J^*} (M_k - m_k) L(I_k) + \sum_{I_k \in J^{**}} (M_k - m_k) L(I_k) \\ &\leq 2C \underbrace{\sum_{I_k \in J^*} L(I_k)}_{< \varepsilon} + \varepsilon \cdot \underbrace{\sum_{I_k \in J^{**}} L(I_k)}_{< b-a} = (2C + (b-a)) \cdot \varepsilon. \end{aligned}$$

Nach Satz 7.13 ist f dann Riemann-integrierbar.

(\Rightarrow) Sei f Riemann-integrierbar. Dann ist f nach Satz 7.6 beschränkt. Es bleibt also zu zeigen, dass die Menge A der Unstetigkeitsstellen von f eine Nullmenge ist.

Sei $A_\alpha := \{x \in [a, b] \mid \text{osz}(f, x) \geq \alpha\}$. Da f genau dann in x stetig ist, wenn $\text{osz}(f, x) = 0$, folgt

$$A = A_1 \cup A_{\frac{1}{2}} \cup A_{\frac{1}{3}} \cup \dots$$

Nach Satz 7.14 genügt es nun zu zeigen, dass die Mengen $A_{\frac{1}{n}}$ für jedes $n \in \mathbb{N}$ Nullmengen sind. Sei $n \in \mathbb{N}$ fixiert und $\varepsilon > 0$ vorgegeben. Nach Voraussetzung existiert dann eine Zerlegung \mathcal{P} mit

$$\overline{S}(f, \mathcal{P}) - \underline{S}(f, \mathcal{P}) < \frac{\varepsilon}{2n}.$$

Sei J^* die Menge der Teilintervalle I_k von \mathcal{P} mit $\text{Int}(I_k) \cap A_{\frac{1}{n}} \neq \emptyset$. Nach Definition der Oszillation folgt dann für $x_0 \in \text{Int}(I_k) \cap A_{\frac{1}{n}}$

$$M_k - m_k = \sup f|_{I_k} - \inf f|_{I_k} \geq \text{osz}(f, x_0) \geq \frac{1}{n}.$$

Daraus ergibt sich

$$\frac{1}{n} \sum_{I_k \in J^*} L(I_k) \leq \sum_{I_k \in J^*} (M_k - m_k) L(I_k) \leq \overline{S}(f, \mathcal{P}) - \underline{S}(f, \mathcal{P}) < \frac{\varepsilon}{2n}.$$

Außerdem wählen wir Intervalle \tilde{I}_k um die Teilungspunkte von \mathcal{P} mit $\sum_{k=1}^m L(\tilde{I}_k) < \frac{\varepsilon}{2}$. Nach

Wahl von J^* folgt dann

$$A_{\frac{1}{n}} \subset \bigcup_{I_k \in J^*} \text{Int}(I_k) \cup \bigcup_{k=1}^m \text{Int}(\tilde{I}_k), \quad \text{und} \quad \sum_{I_k \in J^*} L(I_k) + \sum_{k=1}^m L(\tilde{I}_k) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Somit ist $A_{\frac{1}{n}}$ eine Nullmenge. □

Das Lebesguesche Integrierbarkeitskriterium liefert uns nun die folgenden Beispielklassen für Riemann-integrierbare Funktionen:

1. Jede beschränkte Funktion $f : [a, b] \rightarrow \mathbb{R}$ mit abzählbar vielen Unstetigkeitsstellen ist Riemann-integrierbar.
2. Jede monotone Funktion $f : [a, b] \rightarrow \mathbb{R}$ ist Riemann-integrierbar.
3. Eine Funktion $f : [a, b] \rightarrow \mathbb{K}^n$ ist genau dann Riemann-integrierbar, wenn f beschränkt und fast überall stetig ist (wobei $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$).
4. Ist $f : [a, b] \rightarrow \mathbb{R}^n$ Riemann-integrierbar und $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ stetig, so ist $h \circ f : [a, b] \rightarrow \mathbb{R}^m$ Riemann-integrierbar. Insbesondere ist $\|f\| : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar und es gilt

$$\left\| \int_a^b f(x) dx \right\| \leq \int_a^b \|f(x)\| dx.$$

5. Sind $f, g : [a, b] \rightarrow \mathbb{C}$ Riemann-integrierbar, so ist auch $f \cdot g : [a, b] \rightarrow \mathbb{C}$ Riemann-integrierbar.

Bemerkung: Die *Dirichlet-Funktion* $\delta : [0, 1] \rightarrow \mathbb{R}$, gegeben durch

$$\delta(x) = \begin{cases} 1 & x \text{ rational} \\ 0 & x \text{ irrational,} \end{cases}$$

ist in *keinem* Punkt aus $[0, 1]$ stetig und folglich nicht Riemann-integrierbar. In Analysis 3 führen wir das Lebesgue-Integral ein und werden sehen, dass δ Lebesgue-integrierbar ist.

7.3 Der Hauptsatz der Differential- und Integralrechnung

Wir wissen aus dem letzten Abschnitt, dass jede stetige Funktion auf einem kompakten Intervall Riemann-integrierbar ist. In diesem Abschnitt zeigen wir, dass jede stetige Funktion auch eine Stammfunktion besitzt. Aus Übungsaufgabe 95 wissen wir, dass für Riemann-integrierbare Funktionen, die eine Stammfunktion besitzen, das Riemann- und das unbestimmte Integral übereinstimmen. Dies legt den folgenden Satz nahe:

Satz 7.16 (Hauptsatz der Differential- und Integralrechnung)

Sei $f : [a, b] \rightarrow E$ stetig und die Funktion $F : [a, b] \rightarrow E$ definiert durch

$$F(x) := \int_a^x f(t) dt.$$

Dann gilt:

1. $F : [a, b] \rightarrow E$ ist Lipschitzstetig,
2. F ist differenzierbar und $F' = f$, d.h. F ist eine Stammfunktion von f .
3. Das Riemann-Integral von f über $[a, b]$ stimmt mit dem bestimmten Integral von f über $[a, b]$ überein.

Beweis. (1) Aus den Rechenregeln für Riemann-Integrale (Satz 7.10) folgt

$$\begin{aligned} \|F(x) - F(y)\| &= \left\| \int_a^x f(t) dt - \int_a^y f(t) dt \right\| = \left\| \int_x^y f(t) dt \right\| \\ &\leq \left| \int_x^y \|f(t)\| dt \right| \leq |y - x| \cdot \|f\|_\infty. \end{aligned}$$

(Hier benutzen wir die Vereinbarung: $\int_y^x f(t) dt := -\int_x^y f(t) dt$ falls $x \leq y$.)

(2) Sei $\varepsilon > 0$. Da f in $x_0 \in [a, b]$ stetig ist, existiert ein $\delta > 0$, so dass $\|f(t) - f(x_0)\| < \varepsilon$ für alle $t \in [a, b]$ mit $|t - x_0| < \delta$. Also gilt für alle diese t :

$$\begin{aligned} \left\| \frac{F(t) - F(x_0)}{t - x_0} - f(x_0) \right\| &= \left\| \frac{1}{|t - x_0|} \cdot \int_{x_0}^t f(u) du - f(x_0) \right\| \\ &= \left\| \frac{1}{|t - x_0|} \cdot \left(\int_{x_0}^t (f(u) - f(x_0)) du \right) \right\| \\ &\leq \frac{1}{|t - x_0|} \cdot \int_{x_0}^t \|f(u) - f(x_0)\| du \\ &\leq \frac{1}{|t - x_0|} \cdot \varepsilon |t - x_0| = \varepsilon. \end{aligned}$$

Folglich ist

$$\lim_{t \rightarrow x_0} \frac{F(t) - F(x_0)}{t - x_0} = F'(x_0) = f(x_0).$$

(3) F ist eine Stammfunktion von f mit $F(a) = 0$. Somit gilt für das bestimmte Integral von f über $[a, b]$:

$$\int_a^b f(t) dt = F(b) - F(a) = F(b) = \int_a^b f(t) dt. \quad \square$$

Für stetige Funktionen werden wir für das Riemann-Integral jetzt immer $\int_a^b f(x) dx$ schreiben. Wir können für dieses Integral sowohl die Eigenschaften des bestimmten Integrals (z.B. Partielle Integration, Substitutionsregel) als auch die Eigenschaften des Riemann-Integrals (z.B. Monotonieverhalten, Normabschätzung) benutzen.

Anwendung: Als Anwendung beweisen wir zwei Eigenschaften der Zahl π .

1. *Die Zahl π ist irrational.*

Es genügt zu zeigen, dass π^2 irrational ist. Angenommen π^2 ist rational, d.h. $\pi^2 = \frac{a}{b}$ mit $a, b \in \mathbb{N}$. Wir wählen ein $n \in \mathbb{N}$ mit $\frac{2a^n}{n!} < 1$ und betrachten die Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) := \frac{1}{n!} x^n (1-x)^n = \frac{1}{n!} \cdot \sum_{k=0}^{2n} c_k x^k, \quad \text{wobei } c_k \in \mathbb{Z}.$$

Dann gilt für das Integral

$$I := \pi \int_0^1 a^n f(x) \sin \pi x dx$$

wegen $0 < f|_{(0,1)} < \frac{1}{n!}$ die Ungleichung

$$0 < I < \frac{\pi a^n}{n!} \int_0^1 \sin \pi x dx = \frac{2a^n}{n!} < 1.$$

Wir zeigen nun, dass I ganzzahlig ist: Für die k -ten Ableitungen von f gilt $f^{(k)}(0) \in \mathbb{Z}$ und $f^{(k)}(1) \in \mathbb{Z}$ für alle $k \in \mathbb{N}_0$. Sei $F: \mathbb{R} \rightarrow \mathbb{R}$ die Funktion

$$F(x) := b^n \left(\pi^{2n} f(x) - \pi^{2n-2} f^{(2)}(x) + \pi^{2n-4} f^{(4)}(x) - \dots + (-1)^n f^{(2n)}(x) \right).$$

Nach unserer Annahme ist $\pi^{2(n-k)} = \frac{a^{n-k}}{b^{n-k}}$ für $0 \leq k \leq n$. Folglich gilt $F(0), F(1) \in \mathbb{Z}$. Außerdem erhalten wir

$$\begin{aligned} \frac{d}{dx} (F'(x) \sin \pi x - \pi F(x) \cos \pi x) &= (F^{(2)}(x) + \pi^2 F(x)) \sin \pi x = b^n \pi^{2n+2} f(x) \sin \pi x \\ &= a^n \pi^2 f(x) \sin \pi x \end{aligned}$$

und somit

$$I = \frac{1}{\pi} (F'(1) \sin \pi - \pi F(1) \cos \pi - F'(0) \sin 0 + \pi F(0) \cos 0) = F(1) + F(0) \in \mathbb{Z}.$$

Also müßte I eine ganze Zahl in $(0, 1)$ sein, was nicht möglich ist. Somit war unsere Annahme falsch, d.h. π ist irrational.

2. Es gilt die *Wallissche Formel*:
$$\frac{\pi}{2} = \lim_{n \rightarrow \infty} \prod_{k=1}^n \frac{4k^2}{4k^2 - 1}.$$

Zum Beweis betrachten wir das uns bereits bekannte Integral

$$A_n := \int_0^{\frac{\pi}{2}} \sin^n x \, dx = \begin{cases} \frac{1 \cdot 3 \cdot \dots \cdot (n-1)}{2 \cdot 4 \cdot \dots \cdot n} \cdot \frac{\pi}{2} & \text{falls } n \text{ gerade und } n \geq 2 \\ \frac{2 \cdot 4 \cdot \dots \cdot (n-1)}{1 \cdot 3 \cdot \dots \cdot n} & \text{falls } n \text{ ungerade und } n \geq 3. \end{cases}$$

Für $x \in [0, \frac{\pi}{2}]$ gilt $\sin^{2n+2} x \leq \sin^{2n+1} x \leq \sin^{2n} x$. Aus der Monotonie des Riemann-Integrals folgt somit $A_{2n+2} \leq A_{2n+1} \leq A_{2n}$. Wir erhalten also

$$1 = \lim_{n \rightarrow \infty} \frac{A_{2n}}{A_{2n}} \geq \lim_{n \rightarrow \infty} \frac{A_{2n+1}}{A_{2n}} \geq \lim_{n \rightarrow \infty} \frac{A_{2n+2}}{A_{2n}} = \lim_{n \rightarrow \infty} \frac{2n+1}{2n+2} = 1.$$

Wir haben also in der letzten Ungleichung überall ein Gleichheitszeichen. Damit folgt

$$\begin{aligned} 1 &= \lim_{n \rightarrow \infty} \frac{A_{2n+1}}{A_{2n}} = \lim_{n \rightarrow \infty} \frac{2 \cdot 4 \cdot \dots \cdot (2n)}{1 \cdot 3 \cdot \dots \cdot (2n+1)} \cdot \frac{2 \cdot 4 \cdot \dots \cdot (2n)}{1 \cdot 3 \cdot \dots \cdot (2n-1)} \cdot \frac{2}{\pi} \\ &= \lim_{n \rightarrow \infty} \prod_{k=1}^n \frac{(2k)^2}{(2k+1)(2k-1)} \cdot \frac{2}{\pi}. \end{aligned}$$

Dies zeigt die Wallissche Formel.

7.4 Die Mittelwertsätze der Integralrechnung

Satz 7.17 (1. Mittelwertsatz der Integralrechnung)

Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, $g : [a, b] \rightarrow \mathbb{R}$ Riemann-integrierbar² und $g \geq 0$. Dann existiert ein $\xi \in [a, b]$ mit

$$\int_a^b f(x)g(x) \, dx = f(\xi) \cdot \int_a^b g(x) \, dx.$$

Beweis. Da $f : [a, b] \rightarrow \mathbb{R}$ stetig und $[a, b]$ kompakt, existieren $m := \min\{f(x) \mid x \in [a, b]\}$ und $M := \max\{f(x) \mid x \in [a, b]\}$. Es folgt

$$m \cdot g(x) \leq f(x) \cdot g(x) \leq M \cdot g(x) \quad \forall x \in [a, b]. \quad (*)$$

1. Fall: Sei $\int_a^b g(x) \, dx = 0$. Nach Integration von (*) ergibt sich $0 = \int_a^b f(x)g(x) \, dx$ und damit die Behauptung.

2. Fall: Sei $\int_a^b g(x) \, dx > 0$. Nach Integration von (*) ergibt sich

$$m = f(x_{\min}) \leq \frac{\int_a^b f(x)g(x) \, dx}{\int_a^b g(x) \, dx} \leq M = f(x_{\max}).$$

² Mit den hier auftretenden Integralen meinen wir immer die Riemann-Integrale.

Da f stetig ist, existiert nach dem Zwischenwertsatz ein $\xi \in [x_{\min}, x_{\max}] \subset [a, b]$, so dass

$$f(\xi) = \frac{\int_a^b f(x)g(x) dx}{\int_a^b g(x) dx}. \quad \square$$

Folgerung 7.1 Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig, dann existiert ein $\xi \in [a, b]$, so dass gilt

$$\int_a^b f(x) dx = f(\xi) \cdot (b - a).$$

Beweis. Dies folgt aus Satz 7.17 mit $g \equiv 1$. □

Folgerung 7.2 Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, $f \geq 0$ 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 $\xi_n \in [x_n, y_n]$ mit

$$0 = \int_{x_n}^{y_n} f(x) dx = f(\xi_n)(y_n - x_n).$$

Folglich ist $f(\xi_n) = 0$. Nach Konstruktion konvergiert die Folge (ξ_n) gegen c . Da f stetig ist, konvergiert dann auch $f(\xi_n)$ 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 . □

Satz 7.18 (2. Mittelwertsatz der Integralrechnung)

Sei $f : [a, b] \rightarrow \mathbb{R}$ stetig, $g : [a, b] \rightarrow \mathbb{R}$ monoton wachsend und stetig differenzierbar.

Dann existiert ein $\xi \in [a, b]$, so dass

$$\int_a^b f(x)g(x) dx = g(a) \int_a^\xi f(x) dx + g(b) \int_\xi^b f(x) dx.$$

Beweis. Sei F eine Stammfunktion von f . Da g differenzierbar und monoton wachsend ist, gilt $g' \geq 0$. Wenden wir den 1. Mittelwertsatz der Integralrechnung auf $F \cdot g'$ an, so erhalten wir ein $\xi \in [a, b]$, so dass

$$\int_a^b F(x)g'(x) dx = F(\xi) \cdot \int_a^b g'(x) dx = F(\xi) \cdot (g(b) - g(a)). \quad (*)$$

Mit partieller Integration folgt

$$\begin{aligned}
 \int_a^b f(x)g(x) dx &= \int_a^b F'(x)g(x) dx = - \int_a^b F(x)g'(x) dx + F(x)g(x) \Big|_a^b \\
 &\stackrel{(*)}{=} -F(\xi)(g(b) - g(a)) + F(b)g(b) - F(a)g(a) \\
 &= g(a)(F(\xi) - F(a)) + g(b)(F(b) - F(\xi)) \\
 &= g(a) \int_a^\xi f(x) dx + g(b) \int_\xi^b f(x) dx.
 \end{aligned}$$

□

7.5 Parameterabhängige Integrale

In diesem Abschnitt integrieren wir Funktionen, die von zusätzlichen Parametern abhängen und untersuchen, wann Integral und Ableitung vertauschbar sind.

Satz 7.19 Sei $U \subset \mathbb{R}^n$ offen und $f : [a, b] \times U \subset \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ eine stetige Funktion. Sei $F : U \rightarrow \mathbb{R}$ definiert durch

$$F(x) := \int_a^b f(t, x) dt.$$

Dann gilt:

1. $F : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ ist stetig.
2. Existiert $\frac{\partial f}{\partial x_i}$ und ist $\frac{\partial f}{\partial x_i} : [a, b] \times U \rightarrow \mathbb{R}$ stetig, so ist $F : U \rightarrow \mathbb{R}$ nach x_i stetig differenzierbar und es gilt

$$\frac{\partial F}{\partial x_i}(x) = \int_a^b \frac{\partial f}{\partial x_i}(t, x) dt.$$

Beweis. (1) Sei $x \in U$. Wir wählen eine kompakte Umgebung $W(x) \subset U$ von x . Nach Voraussetzung ist $f|_{[a,b] \times W}$ stetig. Da $[a, b] \times W$ kompakt ist, ist $f|_{[a,b] \times W}$ sogar gleichmäßig stetig, d.h., für alle $\varepsilon > 0$ existiert ein $\delta > 0$, so daß

$$|f(t, x+h) - f(t, x)| < \frac{\varepsilon}{b-a}$$

für alle $t \in [a, b]$ und alle $h \in \mathbb{R}^n$ mit $\|h\| < \delta$. Es folgt

$$\begin{aligned}
 |F(x+h) - F(x)| &= \left| \int_a^b f(t, x+h) - f(t, x) dt \right| \leq \int_a^b |f(t, x+h) - f(t, x)| dt \\
 &\leq \int_a^b \frac{\varepsilon}{b-a} dt = \varepsilon.
 \end{aligned}$$

Folglich ist F in x stetig. Außerdem gilt

$$\int_a^b \lim_{h \rightarrow 0} f(t, x+h) dt = \int_a^b f(t, x) dt = F(x) = \lim_{h \rightarrow 0} F(x+h) = \lim_{h \rightarrow 0} \int_a^b f(t, x+h) dt. \quad (*)$$

(2) Um die Formel für die partielle Ableitung von F nach der Variablen x_i zu beweisen, betrachten wir die Funktion $g : [a, b] \times U \times (-\varepsilon, \varepsilon) \rightarrow \mathbb{R}$ mit

$$g(t, x, \lambda) := \begin{cases} \frac{f(t, x+\lambda e_i) - f(t, x)}{\lambda} & \text{falls } \lambda \neq 0, \\ \frac{\partial f}{\partial x_i}(t, x) & \text{falls } \lambda = 0. \end{cases}$$

Da f stetig ist, ist g in allen Punkten (t, x, λ) des Definitionsbereiches mit $\lambda \neq 0$ stetig. Wir zeigen nun die Stetigkeit von g in $(t, x, 0)$. Dazu wenden wir den Mittelwertsatz der Differentialrechnung auf die Funktion f in der x_i -Komponente an. Ist $\lambda \neq 0$, so existiert ein $\theta_\lambda \in (0, 1)$, so dass

$$f(t, x + \lambda e_i) - f(t, x) = \frac{\partial f}{\partial x_i}(t, x + \theta_\lambda \cdot \lambda e_i) \cdot \lambda.$$

Folglich gilt

$$g(t, x, \lambda) = \frac{\partial f}{\partial x_i}(t, x + \theta_\lambda \cdot \lambda e_i).$$

Sei nun $((t_k, x_k, \lambda_k))$ eine Folge mit $\lambda_k \neq 0$, die gegen $(t, x, 0)$ konvergiert. Dann konvergiert die Folge $(t_k, x_k + \theta_{\lambda_k} \cdot \lambda_k e_i)$ gegen (t, x) . Nach Voraussetzung ist $\frac{\partial f}{\partial x_i}$ stetig. Somit gilt

$$\lim_{k \rightarrow \infty} g(t_k, x_k, \lambda_k) = \lim_{k \rightarrow \infty} \frac{\partial f}{\partial x_i}(t_k, x_k + \theta_{\lambda_k} \cdot \lambda_k e_i) = \frac{\partial f}{\partial x_i}(t, x) = g(t, x, 0).$$

Folglich ist g auch in $(t, x, 0)$ stetig. Wir wenden nun die Formel (*) aus (1) auf g an und erhalten

$$\begin{aligned} \frac{\partial F}{\partial x_i}(x) &= \lim_{\lambda \rightarrow 0} \frac{F(x + \lambda e_i) - F(x)}{\lambda} = \lim_{\lambda \rightarrow 0} \int_a^b \frac{f(t, x + \lambda e_i) - f(t, x)}{\lambda} dt \\ &= \lim_{\lambda \rightarrow 0} \int_a^b g(t, x, \lambda) dt \stackrel{(*)}{=} \int_a^b \lim_{\lambda \rightarrow 0} g(t, x, \lambda) dt = \int_a^b g(t, x, 0) dt \\ &= \int_a^b \frac{\partial f}{\partial x_i}(t, x) dt. \end{aligned}$$

Da $\frac{\partial f}{\partial x_i}$ auf $[a, b] \times U$ stetig ist, ist $\frac{\partial F}{\partial x_i}$ nach (1) auf U stetig. \square

Wir betrachten nun die analoge Situation, wenn zusätzlich die Integralgrenzen von Parametern abhängen.

Satz 7.20 Sei $U \subset \mathbb{R}^n$ offen, $f : [a, b] \times U \subset \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ stetig differenzierbar und $g_0, g_1 : U \subset \mathbb{R}^n \rightarrow [a, b] \subset \mathbb{R}$ zwei differenzierbare Funktionen. Dann ist die Abbildung $G : U \subset \mathbb{R}^n \rightarrow \mathbb{R}$ mit

$$G(x) := \int_{g_0(x)}^{g_1(x)} f(t, x) dt$$

partiell differenzierbar und es gilt

$$\frac{\partial G}{\partial x_i}(x) = \int_{g_0(x)}^{g_1(x)} \frac{\partial f}{\partial x_i}(t, x) dt + \frac{\partial g_1}{\partial x_i}(x) \cdot f(g_1(x), x) - \frac{\partial g_0}{\partial x_i}(x) \cdot f(g_0(x), x).$$

Beweis. Wir betrachten die Abbildungen

$$F(x) := \int_a^b f(t, x) dt, \quad F_0(y, x) := \int_a^y f(t, x) dt, \quad F_1(y, x) := \int_y^b f(t, x) dt.$$

Dann gilt $G(x) = F(x) - F_0(g_0(x), x) - F_1(g_1(x), x)$.

(1) Nach Satz 7.19 existiert $\frac{\partial F}{\partial x_i}(x)$ und es gilt

$$\frac{\partial F}{\partial x_i}(x) = \int_a^b \frac{\partial f}{\partial x_i}(t, x) dt.$$

(2) Nach dem Hauptsatz der Differential- und Integralrechnung ist $F_0(\cdot, x)$ für festes x eine Stammfunktion von $f(\cdot, x)$. Das heißt, es gilt $\frac{\partial F_0}{\partial y}(y, x) = f(y, x)$ für alle (y, x) . Nach Satz 7.19 gilt außerdem für festes y

$$\frac{\partial F_0}{\partial x_i}(y, x) = \int_a^y \frac{\partial f}{\partial x_i}(t, x) dt.$$

Aus der Kettenregel für partielle Ableitungen folgt dann

$$\begin{aligned} \frac{\partial}{\partial x_i} (F_0(g_0(x), x)) &= \frac{\partial F_0}{\partial y}(g_0(x), x) \cdot \frac{\partial g_0}{\partial x_i}(x) + \frac{\partial F_0}{\partial x_i}(g_0(x), x) \\ &= f(g_0(x), x) \cdot \frac{\partial g_0}{\partial x_i}(x) + \int_a^{g_0(x)} \frac{\partial f}{\partial x_i}(t, x) dt. \end{aligned}$$

(3) Analog gilt für F_1

$$\frac{\partial}{\partial x_i} (F_1(g_1(x), x)) = -f(g_1(x), x) \cdot \frac{\partial g_1}{\partial x_i}(x) + \int_{g_1(x)}^b \frac{\partial f}{\partial x_i}(t, x) dt.$$

Insgesamt erhalten wir

$$\frac{\partial G}{\partial x_i}(x) = \int_{g_0(x)}^{g_1(x)} \frac{\partial f}{\partial x_i}(t, x) dt + \frac{\partial g_1}{\partial x_i}(x) \cdot f(g_1(x), x) - \frac{\partial g_0}{\partial x_i}(x) \cdot f(g_0(x), x).$$

□

7.6 Uneigentliche Riemann-Integrale

Wir wollen nun auch Funktionen integrieren, die auf nicht-kompakten Intervallen definiert sind.

Definition 7.9. Eine Funktion $f : [a, c) \rightarrow E$ mit $a < c \leq +\infty$ heißt *uneigentlich Riemann-integrierbar*, wenn $f|_{[a,b]}$ für alle kompakten Intervalle $[a, b] \subset [a, c)$ Riemann-integrierbar ist und der Grenzwert

$$\int_a^c f(t) dt := \lim_{x \rightarrow c^-} \int_a^x f(t) dt \in E$$

existiert³.

Analog definiert man $\int_c^a f(t) dt$ für eine Funktion $f : (c, a] \rightarrow \mathbb{R}$, mit $-\infty \leq c < a$.

Eine Funktion $f : (c, d) \rightarrow E$ mit $-\infty \leq c < d \leq +\infty$ heißt *uneigentlich Riemann-integrierbar*, wenn $f|_{[a,b]}$ für jedes kompakte Intervall $[a, b] \subset (c, d)$ Riemann-integrierbar ist und die uneigentlichen Riemann-Integrale $\int_{x_0}^d f(t) dt$ und $\int_c^{x_0} f(t) dt$ für ein $x_0 \in (c, d)$ existieren. Wir definieren dann

$$\int_c^d f(t) dt := \int_c^{x_0} f(t) dt + \int_{x_0}^d f(t) dt.$$

(Dies ist unabhängig von der Wahl von x_0)

Die Rechenregeln für Riemann-Integrale übertragen sich mit Hilfe der Grenzwertsätze 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

³ Auch hier sind mit dem Integralzeichen immer die Riemann-Integrale gemeint.

$$\lim_{\alpha \rightarrow 0} \int_{\alpha}^1 t^{x-1} e^{-t} dt = \int_0^1 t^{x-1} e^{-t} dt.$$

(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

$$\int_0^{\infty} t^{x-1} e^{-t} dt = \lim_{\alpha \rightarrow 0^+} \int_{\alpha}^1 t^{x-1} e^{-t} dt + \int_1^M t^{x-1} e^{-t} dt + \lim_{\beta \rightarrow \infty} \int_M^{\beta} t^{x-1} e^{-t} dt = \Gamma(x).$$

Satz 7.21 (Eigenschaften der Γ -Funktion)

Die Gamma-Funktion $\Gamma: \mathbb{R}^+ \rightarrow \mathbb{R}$ hat folgende Eigenschaften:

1. $\Gamma(x+1) = x \cdot \Gamma(x) \quad \forall x > 0$.
2. $\Gamma(n) = (n-1)! \quad \forall n \in \mathbb{N}$.
3. $\Gamma(x \cdot a + (1-x) \cdot b) \leq \Gamma(a)^x \cdot \Gamma(b)^{1-x}$ für alle $a, b \in \mathbb{R}^+$ und $x \in (0, 1)$
(d.h. die Funktion $\ln \Gamma$ ist konvex).
4. $\Gamma(x) = \lim_{n \rightarrow \infty} \frac{n! \cdot n^x}{x(x+1) \cdots (x+n)}$ (Gauß-Formel für die Γ -Funktion).
5. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Beweis. 1. Für $x > 0$ und $0 < \varepsilon < r$ folgt mit partieller Integration

$$\int_{\varepsilon}^r t^x e^{-t} dt = \int_{\varepsilon}^r -t^x (e^{-t})' dt \stackrel{p.I.}{=} \int_{\varepsilon}^r x \cdot t^{x-1} \cdot e^{-t} dt - t^x e^{-t} \Big|_{\varepsilon}^r.$$

Geht man mit $\varepsilon \rightarrow 0$ und dann mit $r \rightarrow \infty$, so folgt:

$$\Gamma(x+1) = \int_0^{\infty} t^x e^{-t} dt = x \cdot \int_0^{\infty} t^{x-1} \cdot e^{-t} dt = x \cdot \Gamma(x).$$

2. Wir beweisen die Behauptung durch Induktion.

Induktionsanfang: Für $n = 1$ gilt

$$\Gamma(1) = \int_0^{\infty} e^{-t} dt = -e^{-t} \Big|_0^{\infty} = -\lim_{t \rightarrow \infty} \frac{1}{e^t} + e^0 = 1 = 0!.$$

Induktionsschritt: Die Behauptung gelte für n . Mit 1. folgt dann $\Gamma(n+1) = n\Gamma(n) = n!$.

3. Wir benutzen die Hölder-Ungleichung für Riemann-Integrale (siehe Übungsaufgaben):

$$\int_{\varepsilon}^r |f(t)g(t)| dt \leq \left(\int_{\varepsilon}^r |f(t)|^p dt \right)^{\frac{1}{p}} \cdot \left(\int_{\varepsilon}^r |g(t)|^q dt \right)^{\frac{1}{q}},$$

wobei $p, q > 0$ und $\frac{1}{p} + \frac{1}{q} = 1$ gelten. Seien $x \in (0, 1)$ und $a, b \in \mathbb{R}^+$. Wir setzen in der Hölder-Ungleichung $p = \frac{1}{x}$ und $q = \frac{1}{1-x}$ und betrachten die Funktionen $f(t) := t^{\frac{a-1}{p}} e^{-\frac{t}{p}}$ und $g(t) := t^{\frac{b-1}{q}} e^{-\frac{t}{q}}$. Dann folgt

$$\int_{\varepsilon}^r t^{\frac{a}{p} + \frac{b}{q} - 1} e^{-t} dt \leq \left(\int_{\varepsilon}^r t^{a-1} e^{-t} dt \right)^x \left(\int_{\varepsilon}^r t^{b-1} e^{-t} dt \right)^{1-x}.$$

Lassen wir nun ε gegen 0 und r gegen ∞ konvergieren, so erhalten wir

$$\Gamma(x \cdot a + (1-x) \cdot b) \leq \Gamma(a)^x \cdot \Gamma(b)^{1-x}.$$

4. Für $x = 1$ ist die angegebene Formel offensichtlich erfüllt. Sei nun $x \in (0, 1)$. Für alle natürlichen Zahlen n gilt die Formel $n+x = (1-x)n + x(n+1)$. Wir benutzen, dass Γ logarithmisch konvex ist und erhalten mit 1. und 2.

$$\Gamma(n+x) \leq \Gamma(n)^{1-x} \cdot \Gamma(n+1)^x = \Gamma(n)^{1-x} \cdot \Gamma(n)^x \cdot n^x = (n-1)! \cdot n^x. \quad (*)$$

Des Weiteren gilt für $x \in (0, 1)$ und $n \in \mathbb{N}$ die Formel $n+1 = x(n+x) + (1-x)(n+1+x)$ und wir erhalten analog mit 1.

$$n! = \Gamma(n+1) \leq \Gamma(n+x)^x \cdot \Gamma(n+1+x)^{1-x} = \Gamma(n+x) \cdot (n+x)^{1-x}.$$

Es folgt

$$n!(n+x)^{x-1} \leq \Gamma(n+x). \quad (**)$$

Aus 1. folgt $\Gamma(n+x) = \Gamma(x) \cdot x \cdot (x+1) \cdot \dots \cdot (x+n-1)$ und (*) und (**) liefern dann

$$\frac{n!(n+x)^{x-1}}{\underbrace{x(x+1) \cdot \dots \cdot (x+n-1)}_{=: a_n(x)}} \leq \Gamma(x) \leq \frac{(n-1)! \cdot n^x}{\underbrace{x(x+1) \cdot \dots \cdot (x+n-1)}_{=: b_n(x)}}.$$

Also gilt

$$\frac{a_n(x)}{b_n(x)} \leq \frac{\Gamma(x)}{b_n(x)} \leq 1.$$

Für den Grenzwert auf der linken Seite erhalten wir

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

Daher ist

$$\begin{aligned} \Gamma(x) &= \lim_{n \rightarrow \infty} b_n(x) = \lim_{n \rightarrow \infty} \frac{(n-1)! \cdot n^x}{x(x+1) \cdot \dots \cdot (x+n-1)} \\ &= \lim_{n \rightarrow \infty} \frac{n! \cdot n^x}{x(x+1) \cdot \dots \cdot (x+n)} \underbrace{\left(\frac{x+n}{n} \right)}_{\rightarrow 1} \\ &= \lim_{n \rightarrow \infty} \frac{n! \cdot n^x}{x(x+1) \cdot \dots \cdot (x+n)}. \end{aligned}$$

Somit gilt die Gauß-Formel für Γ für $x \in (0, 1]$. Es bleibt zu zeigen, dass sie mit x auch für $x+1$ gilt. Dies folgt mit 1.:

$$\begin{aligned} \Gamma(x+1) &= x \cdot \Gamma(x) = \lim_{n \rightarrow \infty} \frac{n! n^x}{x(x+1) \cdot \dots \cdot (x+n)} \cdot x \\ &= \lim_{n \rightarrow \infty} \frac{n! n^{x+1}}{(x+1) \cdot \dots \cdot (x+1+n)} \cdot \underbrace{\frac{x+1+n}{n}}_{\rightarrow 1} \\ &= \lim_{n \rightarrow \infty} \frac{n! n^{x+1}}{(x+1) \cdot \dots \cdot (x+1+n)}. \end{aligned}$$

5. Wir benutzen die Gauß-Formel für $x = \frac{1}{2}$ und erhalten

$$\Gamma\left(\frac{1}{2}\right) = \lim_{n \rightarrow \infty} \frac{n! \sqrt{n}}{\frac{1}{2}(\frac{1}{2}+1) \cdot \dots \cdot (\frac{1}{2}+n)} = \lim_{n \rightarrow \infty} \frac{n! \sqrt{n}}{(1-\frac{1}{2})(2-\frac{1}{2}) \cdot \dots \cdot (n-\frac{1}{2})(n+\frac{1}{2})}.$$

Mit der Wallischen Formel folgt dann

$$\begin{aligned} \Gamma\left(\frac{1}{2}\right)^2 &= \lim_{n \rightarrow \infty} \frac{n \cdot (n!)^2}{\frac{1}{2}(n+\frac{1}{2}) \cdot \prod_{k=1}^n (k-\frac{1}{2})(k+\frac{1}{2})} = 2 \cdot \lim_{n \rightarrow \infty} \prod_{k=1}^n \frac{k^2}{k^2 - \frac{1}{4}} \\ &= 2 \cdot \frac{\pi}{2} = \pi. \end{aligned}$$

Somit gilt $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. □

Beispiel 2: Die Integrale $\int_0^{\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}$ und $\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}$.

Wir berechnen das erste Integral mit der Substitution $t = x^2$, $dt = 2x dx$:

$$\int_0^{\infty} e^{-x^2} dx = \int_0^{\infty} e^{-t} \frac{1}{2x} dt = \frac{1}{2} \int_0^{\infty} t^{-\frac{1}{2}} e^{-t} dt = \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = \frac{\sqrt{\pi}}{2}.$$

Das zweite Integral folgt aus dem ersten, denn mit der Substitution $x = -y$, $dx = -dy$ folgt

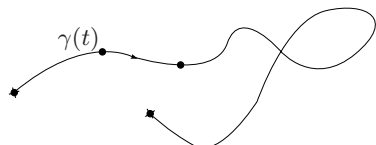
$$\int_{-\infty}^0 e^{-x^2} dx = \int_{\infty}^0 -e^{-y^2} dy = \int_0^{\infty} e^{-y^2} dy. \quad \square$$

7.7 Die Länge von Kurven und der Flächeninhalt ebener Gebiete

In diesem Abschnitt werden wir die Integralrechnung benutzen, um Längen von Kurven sowie Flächeninhalte ebener Gebiet zu bestimmen.

Definition 7.10. Eine parametrisierte Kurve im \mathbb{R}^n ist eine stetige Abbildung $\gamma : I \rightarrow \mathbb{R}^n$, wobei $I \subset \mathbb{R}$ ein beliebiges Intervall bezeichnet.

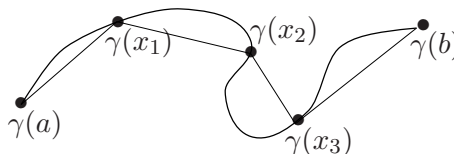
Sei $\Gamma \subset \mathbb{R}^n$ eine Teilmenge. Unter einer Parametrisierung von Γ versteht man eine stetige Abbildung $\gamma : I \rightarrow \mathbb{R}^n$ mit $\gamma(I) = \Gamma$.



Wir interessieren uns also nicht nur für das Bild der Kurve $\Gamma := \gamma(I)$ im \mathbb{R}^n , sondern auch dafür, wie die Kurve durchlaufen wird (z.B. mit welchem Zeitplan). Um die Länge einer parametrisierten Kurve $\gamma : I \rightarrow \mathbb{R}^n$ zu definieren, benutzen wir die geometrische Intuition.

Sei $\mathcal{P} = \{x_0 < x_1 < \dots < x_m\}$ eine endliche Menge von Teilungspunkten des Intervalls I . Dann beschreibt

$$L(\gamma, \mathcal{P}) := \sum_{k=1}^m \|\gamma(x_k) - \gamma(x_{k-1})\|$$



die Länge des durch die Zerlegung \mathcal{P} definierten Sehnepolygons durch die Punkte $\gamma(x_0), \gamma(x_1), \dots, \gamma(x_m)$. (Ist $I = [a, b]$, so setzt man $x_0 = a$ und $x_m = b$).

Ist $\tilde{\mathcal{P}} \geq \mathcal{P}$, so folgt aus der Dreiecksungleichung $L(\gamma, \tilde{\mathcal{P}}) \geq L(\gamma, \mathcal{P})$.

Definition 7.11. Eine parametrisierte Kurve $\gamma : I \rightarrow \mathbb{R}^n$ heißt rektifizierbar, falls

$$L(\gamma) := \sup\{L(\gamma, \mathcal{P}) \mid \mathcal{P} \text{ endliche Zerlegung von } I\}$$

existiert. $L(\gamma)$ heißt dann Länge von γ .

Bemerkung:

1. Ist γ rektifizierbar und $I = I_1 \cup I_2$ eine Zerlegung von I in zwei Teilintervalle, so gilt $L(\gamma) = L(\gamma|_{I_1}) + L(\gamma|_{I_2})$.
2. Jede lipschitzstetige Kurve $\gamma : I \rightarrow \mathbb{R}^n$, die auf einem beschränkten Intervall definiert ist, ist rektifizierbar, denn aus der Lipschitz-Bedingung $\|\gamma(t) - \gamma(s)\| \leq L(t - s)$ für alle $s, t \in I$ folgt für die Längen aller einbeschriebenen Sehnepolygone

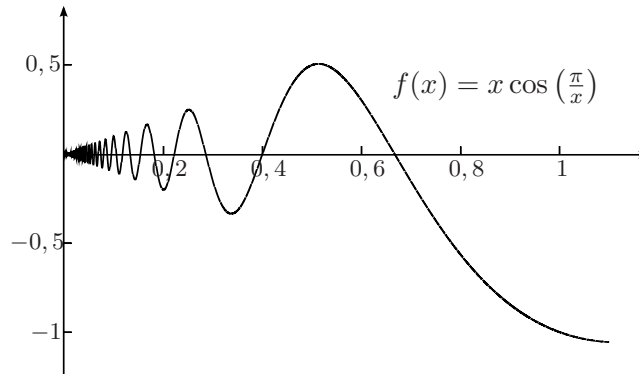
$$L(\gamma, \mathcal{P}) = \sum_{k=1}^m \|\gamma(x_k) - \gamma(x_{k-1})\| \leq L \sum_{k=1}^m (x_k - x_{k-1}) \leq L \cdot \text{Länge}(I).$$

3. Nicht jede (stetige) parametrisierte Kurve ist rektifizierbar.

Zum Beispiel ist die Kurve $\gamma : [0, 1] \rightarrow \mathbb{R}^2$,

$$\gamma(t) := \begin{cases} (t, t \cdot \cos(\frac{\pi}{t})) & t \in (0, 1] \\ 0 & t = 0 \end{cases}$$

stetig, schwankt aber bei $t \rightarrow 0$ zu oft hin und her und ist deshalb nicht rektifizierbar (Übungsaufgabe).



4. Das Bild einer *stetigen* parametrisierten Kurve $[0, 1] \rightarrow \mathbb{R}^2$ kann ein ganzes Quadrat ausfüllen (Peano-Kurven) (für ein Beispiel siehe K. Königsberger: Analysis I, Kapitel 12, oder H. Sagan: Space-Filling Curves. Solche (pathologischen) Fälle wollen wir ausschließen. Wir setzen deshalb im Folgenden voraus, dass unsere parametrisierten Kurven $\gamma : I \rightarrow \mathbb{R}^n$ nicht nur stetig, sondern differenzierbar sind.

Definition 7.12. Eine differenzierbare parametrisierte Kurve $\gamma : I \rightarrow \mathbb{R}^n$ heißt *regulär*, wenn $\gamma'(t) \neq 0$ für alle $t \in I$.

Reguläre parametrisierte Kurven besitzen in jedem Parameter $t \in I$ eine Tangente, die durch die Gerade $Tan_t \gamma := \gamma(t) + \mathbb{R}\gamma'(t)$ beschrieben wird.

Die Länge von stetig differenzierbaren parametrisierten Kurven kann man mit Hilfe des Riemann-Integrals berechnen.

Satz 7.22 Sei $\gamma : [a, b] \rightarrow \mathbb{R}^n$ eine stetig differenzierbare parametrisierte Kurve. Dann ist γ rektifizierbar und es gilt

$$L(\gamma) = \int_a^b \|\gamma'(t)\| dt.$$

Beweis. Da γ stetig differenzierbar ist, ist γ auch Lipschitzstetig und damit rektifizierbar.

(1) Wir zeigen zuerst: $L(\gamma) \leq \int_a^b \|\gamma'(t)\| dt$:

Sei $\mathcal{P} = \{a = x_0 < x_1 < \dots < x_m = b\}$ eine beliebige Zerlegung von $[a, b]$. Da γ differenzierbar ist, gilt

$$L(\gamma, \mathcal{P}) = \sum_{k=1}^m \|\gamma(x_k) - \gamma(x_{k-1})\| = \sum_{k=1}^m \left\| \int_{x_{k-1}}^{x_k} \gamma'(t) dt \right\| \leq \sum_{k=1}^m \int_{x_{k-1}}^{x_k} \|\gamma'(t)\| dt = \int_a^b \|\gamma'(t)\| dt$$

und wir erhalten

$$L(\gamma) = \sup_{\mathcal{P}} L(\gamma, \mathcal{P}) \leq \int_a^b \|\gamma'(t)\| dt. \tag{*}$$

(2) Wir zeigen nun, dass in (*) 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 $t + h \in [a, b]$. Dann ist

$$L(\gamma|_{[t,t+h]}) \geq \underbrace{\|\gamma(t+h) - \gamma(t)\|}_{\text{kürzeste Länge zw. den EP}}$$


Wir wenden (*) 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

$$L(\gamma) = \ell(b) - \ell(a) = \int_a^b \ell'(t) dt = \int_a^b \|\gamma'(t)\| dt.$$

□

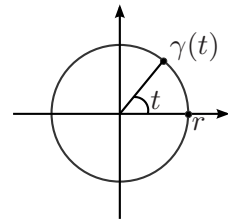
Bemerkung:

1. Sei $I \subset \mathbb{R}$ ein nicht-kompaktes Intervall mit den Grenzen a und b , $-\infty \leq a < b \leq +\infty$, und $\gamma : I \rightarrow \mathbb{R}^n$ eine parametrisierte C^1 -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 kompakte 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. Sei $\gamma : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ eine stückweise C^1 -Kurve, d.h., γ ist stetig und es existieren endlich viele Teilungspunkte $a = t_0 < t_1 < t_2 < \dots < t_m = b$ des Definitionsbereiches I , so dass $\gamma|_{(t_i, t_{i+1})}$ 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: Die Kreislinie

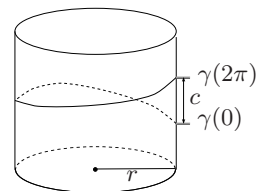
Wir betrachten die Kreislinie vom Radius r parametrisiert durch $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$ mit $\gamma(t) := (r \cos t, r \sin t)$. Dann ist $\gamma'(t) = (-r \sin t, r \cos t)$ und $\|\gamma'(t)\| = r$. Folglich erhalten wir für die Länge



$$L(\gamma) = \int_0^{2\pi} r dt = 2\pi r \quad \text{und} \quad L(\gamma|_{[0, \varphi]}) = \varphi \cdot r.$$

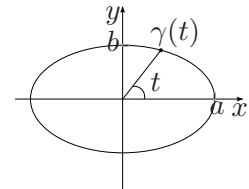
Beispiel 2: Die Schraubenlinie

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 $\gamma'(t) = (-r \sin t, r \cos t, c)$ und $\|\gamma'(t)\| = \sqrt{r^2 + c^2}$. 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 $a, b \in \mathbb{R} \setminus \{0\}$. Es gelte $a > b$ und es 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

$$L(\gamma) = a \int_0^{2\pi} \sqrt{\sin^2 t + \frac{b^2}{a^2} \cos^2 t} dt = a \int_0^{2\pi} \sqrt{1 - k^2 \cos^2 t} dt.$$

Dies ist ein elliptisches Integral, es ist nicht durch elementare Funktionen ausdrückbar. Durch Reihenentwicklung von $\sqrt{1-x}$ (wie in Kapitel 7.2) erhalten wir

$$L(\gamma) = 2a\pi \left(1 + \sum_{j=1}^{\infty} \binom{\frac{1}{2}}{j} (-1)^j k^{2j} \frac{1 \cdot 3 \cdot \dots \cdot (2j-1)}{2 \cdot 4 \cdot \dots \cdot (2j)} \right).$$

Man kann eine Kurve Γ auf verschiedene Weisen (z.B. mit unterschiedlicher Geschwindigkeit) durchlaufen.

Definition 7.13. Sei $\gamma : I \subset \mathbb{R} \rightarrow \mathbb{R}^n$ eine parametrisierte C^1 -Kurve und $\tau : \tilde{I} \rightarrow I$ eine Parametertransformation, d.h. ein C^1 -Diffeomorphismus. Dann heißt $\tilde{\gamma} := \gamma \circ \tau : \tilde{I} \rightarrow \mathbb{R}^n$ Umparametrisierung von γ .

Es gilt also $\Gamma = \text{Im}(\gamma) = \text{Im}(\tilde{\gamma})$. Sinnvollerweise sollten sich geometrische Eigenschaften, wie die Länge der Kurve, bei Umparametrisierung nicht ändern.

Satz 7.23 Sei $\gamma : [a, b] \rightarrow \mathbb{R}^n$ eine parametrisierte C^1 -Kurve.

1. Ist $\tilde{\gamma} = \gamma \circ \tau : [\tilde{a}, \tilde{b}] \rightarrow \mathbb{R}^n$ eine Umparametrisierung von γ , so gilt $L(\gamma) = L(\tilde{\gamma})$.
2. Ist $R : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine Euklidische Bewegung, d.h., $R(x) = Ax + x_0$ mit $A \in O(n)$, $x_0 \in \mathbb{R}^n$, so gilt $L(\gamma) = L(R \circ \gamma)$.

Beweis. 1) Aus $\tilde{\gamma}(s) = \gamma(\tau(s))$ folgt $\tilde{\gamma}'(s) = \gamma'(\tau(s)) \cdot \tau'(s)$ und damit

$$L(\tilde{\gamma}) = \int_{\tilde{a}}^{\tilde{b}} \|\tilde{\gamma}'(s)\| ds = \int_{\tilde{a}}^{\tilde{b}} \|\gamma'(\tau(s))\| \cdot |\tau'(s)| ds.$$

Wir substituieren $t := \tau(s)$. Dann ist $dt = \tau'(s)ds$.

a) Ist $\tau'(s) > 0$, so ist τ monoton wachsend, d.h., es gilt $\tau(\tilde{a}) = a$ und $\tau(\tilde{b}) = b$. Weiterhin ist $|\tau'(s)| = \tau'(s)$ und es folgt

$$L(\tilde{\gamma}) = \int_a^b \|\gamma'(t)\| dt = L(\gamma).$$

b) Ist $\tau'(s) < 0$, so ist τ monoton fallend, d.h., es gilt $\tau(\tilde{a}) = b$ und $\tau(\tilde{b}) = a$. Weiterhin ist $|\tau'(s)| = -\tau'(s)$ und es folgt

$$L(\tilde{\gamma}) = - \int_b^a \|\gamma'(t)\| dt = \int_a^b \|\gamma'(t)\| dt = L(\gamma).$$

2) Sei die Kurve $\delta : [a, b] \rightarrow \mathbb{R}^n$ definiert durch $\delta(t) := R(\gamma(t)) = A\gamma(t) + x_0$. Dann ist $\delta'(t) = A\gamma'(t)$ und da A eine orthogonale Matrix ist, gilt

$$\|\delta'(t)\| = \|A\gamma'(t)\| = \|\gamma'(t)\|.$$

Also ist $L(\delta) = L(\gamma)$. □

Oft wählt man Parametrisierungen, bei denen die Kurve mit konstanter Geschwindigkeit durchlaufen wird, z.B. mit der Geschwindigkeit 1, d.h. so, dass $\|\gamma'(t)\| = 1$. In diesem Fall stimmt die Länge der Kurve mit der Länge des durchlaufenen Parameterintervalls überein.

Definition 7.14. Eine parametrisierte Kurve $\gamma : I \subset \mathbb{R} \rightarrow \mathbb{R}^{n+1}$ heißt auf Bogenlänge (oder durch ihre Bogenlänge) parametrisiert, falls $L(\gamma|_{[\alpha,\beta]}) = \beta - \alpha$ für alle $[\alpha, \beta] \subset I$.

Satz 7.24 (Umparametrisierung regulärer Kurven auf Bogenlänge)

1. Eine stetig differenzierbare, parametrisierte Kurve $\gamma : I \rightarrow \mathbb{R}^n$ ist genau dann auf Bogenlänge parametrisiert, wenn $\|\gamma'(t)\| = 1$ für alle $t \in I$.
2. Ist $\gamma : I \rightarrow \mathbb{R}^n$ stetig differenzierbar und regulär, so existiert eine Umparametrisierung von γ auf Bogenlänge.

Beweis. 1) (\implies) Sei $t_0 \in I$. Wir wählen ein kompaktes Intervall $[a, b] \subset I$ mit $t_0 \in [a, b]$ und betrachten die Längenfunktion $\ell : [a, b] \rightarrow [0, L(\gamma)]$,

$$\ell(t) := L(\gamma|_{[a,t]}) = \int_a^t \|\gamma'(s)\| ds.$$

Dann gilt nach Voraussetzung $\ell(t) = t - a$ und somit $\ell'(t) = 1$. Andererseits ist nach dem Hauptsatz der Differential- und Integralrechnung $\ell'(t) = \|\gamma'(t)\|$. Daraus folgt $\|\gamma'(t)\| = 1$ für alle $t \in [a, b]$ und somit für alle $t \in I$.

(\impliedby) Wenn $\|\gamma'(t)\| = 1$ für alle $t \in I$, so gilt $L(\gamma|_{[\alpha,\beta]}) = \int_\alpha^\beta \|\gamma'(t)\| dt = \beta - \alpha$ für alle $[\alpha, \beta] \subset I$, d.h., γ ist auf Bogenlänge parametrisiert.

2) Wir betrachten wieder die Längenfunktion $\ell : I \rightarrow \tilde{I}$. Wenn γ regulär ist, gilt $\ell'(t) = \|\gamma'(t)\| > 0$. Folglich ist ℓ stetig differenzierbar und streng monoton wachsend. Deshalb existiert eine differenzierbare Umkehrfunktion $\tau : \tilde{I} \rightarrow I$ von ℓ . Für die Ableitung von τ gilt dann $\tau'(s) = \frac{1}{\ell'(\tau(s))} = \frac{1}{\|\gamma'(\tau(s))\|} > 0$. Für $\tilde{\gamma} := \gamma \circ \tau$ folgt mit Kettenregel

$$\tilde{\gamma}'(s) = \gamma'(\tau(s)) \cdot \tau'(s) = \frac{\gamma'(\tau(s))}{\|\gamma'(\tau(s))\|}.$$

Also ist $\|\tilde{\gamma}'(s)\| = 1$ für alle $s \in \tilde{I}$. □

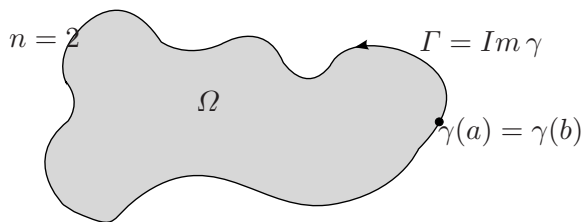
Als nächstes befassen wir uns mit dem Flächeninhalt ebener Gebiete.

Definition 7.15. Eine stetige Kurve $\gamma : [a, b] \rightarrow \mathbb{R}^n$ heißt

- geschlossen, falls $\gamma(a) = \gamma(b)$.
- einfach, falls $\gamma : (a, b) \rightarrow \mathbb{R}^n$ 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:

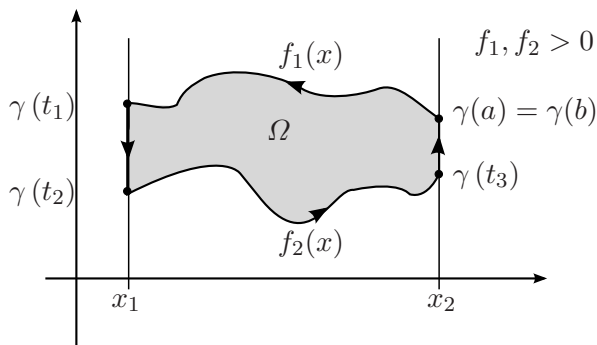


Ω bezeichne das von $\Gamma := \text{Im}(\gamma)$ umschlossene Gebiet. Wir geben nun eine Formel an, mit der man den Flächeninhalt von Ω mit Hilfe der Randkurve berechnen kann.

Satz 7.25 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 parametrisierte Kurven γ , bei denen das Bild der Kurve aus zwei zur y -Achse parallelen Geraden 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

$$\text{Area}(\Omega) = \int_{x_1}^{x_2} f_1(x) dx - \int_{x_1}^{x_2} f_2(x) dx. \quad (*)$$

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 Koordinatentransformation $x = x(t)$, $dx = x'(t)dt$ 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 (*)

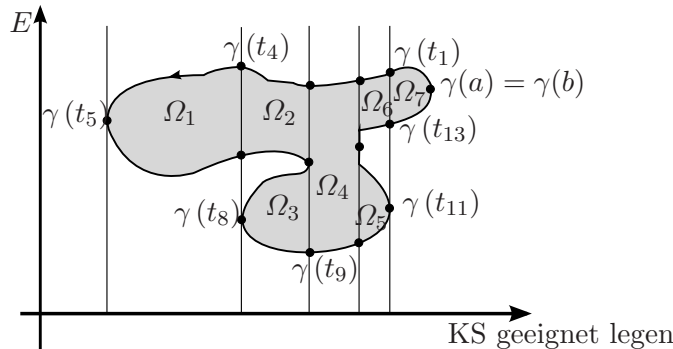
$$\int_{x_1}^{x_2} f_2(x) dx = \int_{t_2}^{t_3} y(t)x'(t) dt.$$

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.



Beh.: Man kann Ω in eine endliche Zahl von Gebieten zerlegen, die die Form aus 1) haben. Wir legen das Koordinatensystem so, dass Γ im positiven Quadranten liegt. Sei E die y -Achse. Wir betrachten die stetig differenzierbare Abbildung

$$\varrho(t) = dist(E, \gamma(t)) = x(t).$$

Da das Intervall $[a, b]$ kompakt ist, hat ϱ nur endlich viele lokale Maxima und Minima. Wir zeichnen in diesen Punkten, wie im Bild dargestellt, die zu E parallelen Geraden. Dann zerlegt sich Ω in Gebiete der Form 1).

Wenn $x'(t) \neq 0$, so ist $x' > 0$ oder $x' < 0$. Daher ist $x(t)$ zwischen den kritischen Punkten von $x(t)$ monoton wachsend oder monoton fallend. Deshalb existiert eine Umkehrfunktion $t = t(x)$ und die Teilstücke von γ sind Graphen einer Funktion, d.h., $\gamma(t) = (x(t), y(t))$ mit $y = y(t) = y(t(x)) =: f(x)$. 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. \square

Folgerung 7.3 Ist die Randkurve γ in Polarkoordinaten gegeben, d.h., gilt

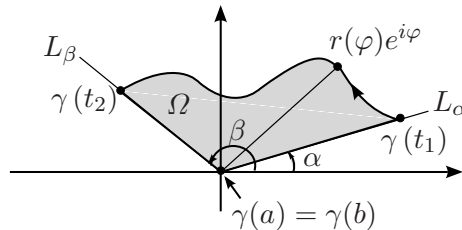
$$\gamma(t) = r(t)e^{i\varphi(t)} = (r(t) \cos \varphi(t), r(t) \sin \varphi(t)),$$

so ist

$$\text{Area}(\Omega) = \frac{1}{2} \int_a^b r^2(t) \varphi'(t) dt.$$

Folgerung 7.4 (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 7.3 erhalten wir

$$\text{Area}(\Omega) = \frac{1}{2} \int_{t_1}^{t_2} r^2(t) \varphi'(t) dt.$$

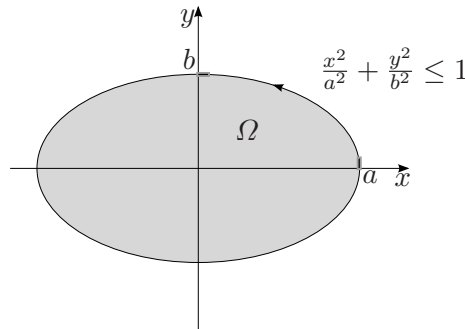
Wir substituieren $\varphi = \varphi(t)$, $d\varphi = \varphi'(t)dt$. Da $\varphi(t)$ auf $[t_1, t_2]$ streng monoton wachsend ist, existiert eine Umkehrfunktion $t = t(\varphi)$. Somit können wir r in Abhängigkeit von φ darstellen, $r = r(\varphi)$, und erhalten insgesamt

$$Area(\Omega) = \frac{1}{2} \int_{\alpha}^{\beta} r^2(\varphi) d\varphi.$$

□

Beispiel 1: *Flächeninhalt einer Ellipse*

Sei Ω die von einer Ellipse mit den Halbachsen a und b eingeschlossene Fläche.

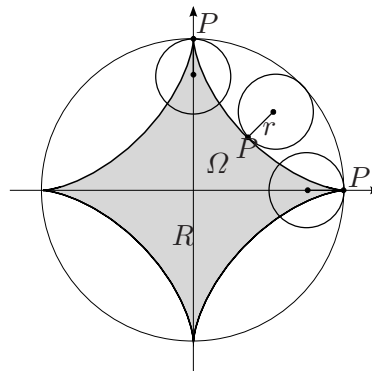


Wir parametrisieren die Ellipse durch die Kurve $\gamma(t) = (a \cos t, b \sin t) =: (x(t), y(t))$ mit $t \in [0, 2\pi]$. Dann ist nach Satz 7.25

$$Area(\Omega) = \frac{1}{2} \int_0^{2\pi} (x(t)y'(t) - y(t)x'(t)) dt = \frac{1}{2} \int_0^{2\pi} ab \cdot (\cos^2 t + \sin^2 t) dt = \pi ab.$$

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 $\gamma(t) = R (\cos^3 t, \sin^3 t)$. Dann folgt für den Flächeninhalt des Gebietes Ω , das durch die Astroide begrenzt wird

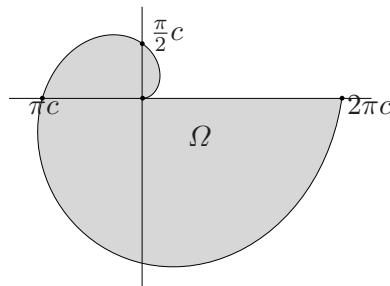
$$\begin{aligned}
Area(\Omega) &= \frac{1}{2} R^2 \int_0^{2\pi} (\cos^3 t \cdot 3 \sin^2 t \cdot \cos t + 3 \cos^2 t \cdot \sin t \cdot \sin^3 t) dt \\
&= \frac{3}{2} R^2 \int_0^{2\pi} \sin^2 t \cos^2 t (\cos^2 t + \sin^2 t) dt \\
&= \frac{3}{2} R^2 \int_0^{2\pi} (\sin t \cos t)^2 dt = \frac{3}{2} R^2 \int_0^{2\pi} \left(\frac{1}{2} \sin 2t\right)^2 dt \\
&= \frac{3}{8} R^2 \int_0^{2\pi} \sin^2 2t dt, \quad x := 2t, dx = 2dt \\
&= \frac{3}{16} R^2 \int_0^{4\pi} \sin^2 x dx = \frac{3}{16} R^2 \cdot 8 \int_0^{\frac{\pi}{2}} \sin^2 x dx = \frac{\pi}{4} = \frac{3}{8} \pi R^2.
\end{aligned}$$

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

$$\gamma(\varphi) = r(\varphi)e^{i\varphi} = (c\varphi \cos \varphi, c\varphi \sin \varphi)$$

gegeben.



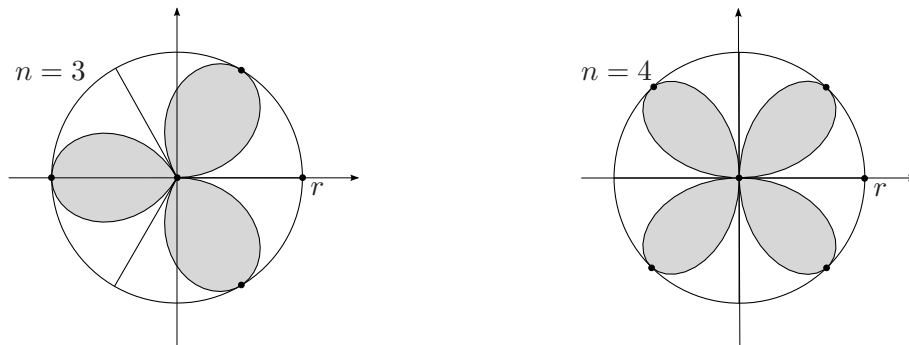
Nach der Leibnizschen Sektorformel ist der Flächeninhalt des Gebietes Ω

$$Area(\Omega) = \frac{1}{2} \int_0^{2\pi} r^2(\varphi) d\varphi = \frac{1}{2} c^2 \int_0^{2\pi} \varphi^2 d\varphi = \frac{1}{2} c^2 \cdot \frac{1}{3} 8\pi^3 = \frac{4}{3} \pi^3 c^2.$$

Beispiel 4: Das Kleeblatt

Den Rand des n -blättrigen Kleeblatts beschreiben wir durch eine Kurve in Polarkoordinaten:

$$r(\varphi) = r \cdot \left| \sin \left(\frac{n}{2} \varphi \right) \right|, \quad \text{wobei } \varphi \in [0, 2\pi] \text{ und } r \in \mathbb{R}.$$



Es gilt $r(\varphi) = 0 \iff \frac{n}{2}\varphi = k\pi$ und $r(\varphi) = r \iff \frac{n}{2}\varphi = \frac{\pi}{2} + k\pi$, wobei $k \in \mathbb{Z}$.

Unabhängig von der Anzahl n der einzelnen Blätter ist die Kleeblattfläche K_n halb so groß wie die Fläche der Kreisscheibe vom Radius r , d.h.,

$$\text{Area}(K_n) = \frac{1}{2}\pi r^2 \quad \text{für alle } n \in \mathbb{N}.$$

Aus der Leibnizschen Sektorformel folgt nämlich mit der Substitution $t := \frac{n}{2}\varphi$, $dt = \frac{n}{2}d\varphi$

$$\text{Area}(K_n) = \frac{1}{2}r^2 \int_0^{2\pi} \sin^2\left(\frac{n}{2}\varphi\right) d\varphi = \frac{1}{2}r^2 \cdot \frac{2}{n} \int_0^{n\pi} \sin^2 t dt = \frac{r^2}{n} \cdot n \int_0^{\pi} \sin^2 t dt = \frac{1}{2}\pi r^2.$$

