Mathematik für Informatik 2

Einführung in die Analysis

Auf diesen Seiten werden die wesentlichen Inhalte der Vorlesung *Mathematik für Informatik 2* zusammengefasst, deren Schwerpunkt die Grundlagen der Analysis sind. Es gibt darüber hinaus zahlreiche Lehrbücher über Analysis, die im Wesentlichen ähnlich aufgebaut sind. Diese Bücher sind meist sehr umfangreich und wir haben in dieser Vorlesung versucht, uns auf die Kernaspekte der Analysis zu konzentrieren und dabei trotzdem den Großteil der vorgestellten mathematischen Aussagen zu beweisen. Neben den Grundlagen der Analysis soll Ihnen diese Vorlesung damit auch einen Einblick in die mathematische Beweisführung geben, die Sie im Laufe des Studiums auch in vielen anderen Kontexten anwenden werden.

Die stärksten Überschneidungen hat die Vorlesung mit dem Lehrbuch "Analysis 1" bzw. "Analysis 2" von Otto Forster. Das zweite Kapitel, in dem wir die reellen Zahlen einführen, orientiert sich am lesenswerten Skript von Dr. R. Busam (Universität Heidelberg). Außerdem basiert der gesamte Inhalt auf dem vorherigen MafI2-Skript von Prof. Dr. Peter Buchholz und Prof. Dr. Günter Rudolph. Ein weiteres zu empfehlendes Lehrbuch mit sehr vielen Anwendungsbeispielen und Programmieranwendungen in Python ist "Konkrete Mathematik – nicht nur für Informatiker" von Prof. Dr. Edmund Weitz (Universität Hamburg), von dem manche unserer Demo-Applikationen inspiriert wurden.

Das hier gewählte Onlineformat für das Skript bietet viele Vorteile und (hoffentlich) wenige Nachteile für Sie. So können wir zum Beispiel interaktive Demos einbinden, sowie Elemente interaktiv ein- und ausblenden. Die wichtigsten Funktionalitäten dieses Online-Skripts haben wir auf einer kleinen Tutorial-Seite zusammengefasst. Dies ist allerdings erst die erste Version des Skripts und sollten Ihnen Fehler auffallen, oder falls sie Ideen für sinnvolle Erweiterungen haben, geben Sie uns gerne Feedback.

1 Grundlagen

Sehr allgemein gefasst ist die Analysis die Wissenschaft der Beziehungen zwischen Zahlen. Die Zahlen bilden damit also das Fundament der Analysis und wir müssen uns zunächst einig sein, was diese überhaupt sind. Bevor wir jedoch in Kapitel 2 mit den Definitionen der Zahlenmengen starten können, führen wir ein paar grundlegende Definitionen über Aussagen und allgemeine Mengen in Kapitel 1.1 ein. Außerdem benötigen wir Techniken, die es uns erlauben, aus bestehenden Aussagen neue wahre Zusammenhänge/Eigenschaften zu erschließen. Die wichtigesten Beweistechniken werden dazu in Kapitel 1.2 vorgestellt. Die Inhalte dieses Grundlagenkapitels haben Sie vielleicht schon in der Veranstaltung *Mathematik für Informatik 1* kennen gelernt, da diese in fast jedem Gebiet der Mathematik essenziell sind. Sollten ihnen die Themen also noch gut im Gedächtnis sein, genügt es vermutlich, dieses Kapitel grob zu überfliegen.

1.1 Aussagen und Mengen

Definition 1.1 (Aussage)

Aussagen sind (schrift)sprachliche Gebilde, denen ein Wahrheitswert *wahr* (w) oder *falsch* (f) zugeordnet werden kann.

Wir starten mit unserer ersten Definition. Bei Definitionen geht es darum, dass wir einem Objekt, das eine Anzahl von Eigenschaften besitzt, einen Namen geben wollen. In diesem Fall wollen wir “(schrift)sprachlichen Gebilden, denen ein Wahrheitswert wahr oder falsch zugeordnet werden kann” den Namen “Aussage” geben, damit wir in Zukunft einfach dieses kurze Wort statt der langen Beschreibung verwenden können. Definitionen benötigen daher auch keinen Beweis, da wir uns dabei nur auf eine Kurzschreibweise festlegen aber noch keine Behauptungen damit aufstellen. Wir werden Definitionen immer mit der gleichen Farbe hinterlegen und durchnummerieren. Für unsere Definitionen werden wir in vielen Fällen auch direkt Beispiele angeben:

Beispiel 1.2 (Aussagen)

Ein paar Beispiele für Aussagen:

1. *Delfine sind Fische.* (f)
2. *Fünf ist eine ungerade Zahl.* (w)
3. *Es gibt unendlich viele Primzahlen.* (w)
4. *Jede gerade natürliche Zahl größer als zwei ist Summe zweier Primzahlen. (?)*

Der Wahrheitswert der letzten Aussage ist unbekannt. Sie wird auch als *Goldbachsche Vermutung* bezeichnet und stellt eine unbewiesene Aussage dar. Das heißt aber nicht, dass es keine Aussage ist, sie ist eindeutig entweder wahr oder falsch, wir wissen nur noch nicht, welcher der beiden Fälle zutrifft.

Die folgenden beiden Sätze stellen *keine* Aussagen dar:

1. *Guten Tag!*
2. *Diese Aussage ist falsch.*

Machen wir gleich mit einer zweiten Definition weiter:

Definition 1.3 (Operationen für Aussagen)

Seien A und B Aussagen. Wir definieren mithilfe folgender Operationen neue Aussagen:

1. **Negation** ("nicht A"):
   Die Aussage ist wahr, wenn A falsch ist.
   Symbolschreibweise: $(\neg A)$
2. **Disjunktion** ("A oder B"):
   Die Aussage ist wahr, wenn mindestens eine der beiden Aussagen wahr ist.
   Symbolschreibweise: $(A \vee B)$
3. **Konjunktion** ("A und B"):
   Die Aussage ist wahr, wenn beide Aussagen wahr sind.
   Symbolschreibweise: $(A \wedge B)$
4. **Implikation** ("aus A folgt B"):
   Die Aussage ist wahr, wenn A falsch oder B wahr ist.
   Symbolschreibweise: $(A \Rightarrow B)$
5. **Äquivalenz** ("A ist äquivalent zu B"):
   Die Aussage ist wahr, wenn beide Aussagen den gleichen Wahrheitswert besitzen.
   Symbolschreibweise: $(A \Leftrightarrow B)$

Wir sagen statt $(A \Leftrightarrow B)$ auch häufig: A gilt *genau dann, wenn* B gilt.

Wir können auch mehr als zwei Aussagen mit diesen Operationen verknüpfen und durch Klammern die Rangfolge angeben, wie z.B. $A \vee (B \wedge C)$, gelesen "A oder B und C". Bei der gesprochenen Variante, stellt sich bereits heraus, dass die Formelschreibweise eindeutiger ist, denn der gesprochene Satz könnte auch zu $(A \vee B) \wedge C$ passen.

Um festzustellen, wann so eine kombinierte Aussage wahr oder falsch ist, benutzt man häufig sogenannte Wahrheitstafeln. Dabei trägt man zunächst alle in der Gesamtaussage vorkommenden Teilaussagen in jeder möglichen Kombination von wahr und falsch ein. Anschließend arbeitet man sich zur Zielaussage vor. Bei zwei Aussagen A und B ergeben sich zum Beispiel folgende Kombinationsmöglichkeiten

A	B
f	f
f	w
w	f
w	w

Der letzten Definition folgend können wir die Wahrheitstafel für die Basisoperationen bestimmen

A	B	$\neg A$	$A \vee B$	$A \wedge B$	$A \Rightarrow B$	$A \Leftrightarrow B$
f	f	w	f	f	w	w
f	w	w	w	f	w	f
w	f	f	w	f	f	f
w	w	f	w	w	w	w

Mit den Wahrheitstafeln haben wir das erste Werkzeug, mit dessen Hilfe wir neue Zusammenhänge erschließen können. Nach Definition 1.3 sind zwei Aussagen äquivalent, wenn sie den gleichen Wahrheitswert besitzen. Das bedeutet, wenn für zwei Aussagen die jeweilige Spalte in der vollständigen Wahrheitstafel identisch ist, sind die Aussagen äquivalent. Damit können wir uns an unseren ersten Satz wagen. Sätze verwenden wir für alle wichtigen neuen Aussagen, die wir aus bekannten Aussagen oder Definitionen ableiten. Bisher können wir die Beweise nur über Wahrheitstafeln führen, aber wir werden in Kapitel 1.2 noch einige weitere Beweismethoden einführen.

Satz 1.4 (äquivalente Aussagen)

Seien A, B, C Aussagen, dann gelten folgende Äquivalenzen:

$$\neg(\neg A) \Leftrightarrow A \qquad \text{(doppelte Negation)}$$

$$A \wedge (A \vee B) \Leftrightarrow A$$
$$A \vee (A \wedge B) \Leftrightarrow A \qquad \text{(Absorption)}$$

$$A \wedge B \Leftrightarrow B \wedge A$$
$$A \vee B \Leftrightarrow B \vee A \qquad \text{(Kommutativität)}$$

$$\neg(A \wedge B) \Leftrightarrow \neg A \vee \neg B$$
$$\neg(A \vee B) \Leftrightarrow \neg A \wedge \neg B \qquad \text{(de-morgansche Regeln)}$$

$$A \vee (B \vee C) \Leftrightarrow (A \vee B) \vee C$$
$$A \wedge (B \wedge C) \Leftrightarrow (A \wedge B) \wedge C \qquad \text{(Assoziativität)}$$

$$A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$$
$$A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C) \qquad \text{(Distributivgesetze)}$$

▼ Beweis

Alle obigen Äquivalenzen sind recht einfach über Wahrheitstafeln zu beweisen, wir werden hier exemplarisch die Assoziativität von $\vee$ beweisen und die übrigen Beweise können Sie als Übung führen.

Im Fall der Assoziativität haben wir 3 Aussagen, also ergeben sich insgesamt $8 = 2^3$ Kombinationsmöglichkeiten:

A	B	C
w	w	w
w	w	f
w	f	w
w	f	f
f	w	w
f	w	f
f	f	w
f	f	f

Zur Erinnerung, wir wollen $A \vee (B \vee C) \Leftrightarrow (A \vee B) \vee C$ beweisen. Dafür wenden wir nun schrittweise (Klammern zuerst) die Definition 1.3 für $\vee$ an:

A	B	C	$B \vee C$	$A \vee (B \vee C)$	$A \vee B$	$(A \vee B) \vee C$
w	w	w	w	w	w	w

w	w	f	w	w	w	w
w	f	w	w	w	w	w
w	f	f	f	w	w	w
f	w	w	w	w	w	w
f	w	f	w	w	w	w
f	f	w	w	w	f	w
f	f	f	f	f	f	f

Da alle Einträge der drittletzten Spalte mit denen der letzte Spalte identisch sind, sind beide Aussagen Äquivalent.

In den Aussagen, welche wir im Laufe der Vorlesung beweisen wollen, kommen fast immer Mengen vor. Beispielsweise wollen wir eine Gleichung für alle ganzen Zahlen beweisen, oder für alle reellen Zahlen zwischen 0 und 1. Manchmal wollen wir zeigen, dass es eine bestimmte Zahl gibt, für die eine Aussage gilt. Wir werden die Zahlenmengen erst in Kapitel 2 detailliert einführen und starten hier zunächst mit dem allgemeinen Mengenbegriff.

Definition 1.5 (Menge, Elemente der Menge, leere Menge)

Unter einer *Menge* verstehen wir jede Zusammenfassung M von bestimmten wohlunterscheidbaren Objekten m unserer Anschauung oder unseres Denkens, welche *Elemente* von M genannt werden. Wir schreiben auch $m \in M$. Für ein Objekt n, das nicht in der Menge enthalten ist schreiben wir $n \notin M$

Die Menge, die keine Elemente enthält, bezeichnen wir als die *leere Menge* und verwenden das Symbol $\varnothing$

Beispiel 1.6 (Mengen, Mengenangabe)

Die obige Definition der Menge ist sehr allgemein gefasst und somit können wir z.B. die Menge der Wochentage W, oder der Farben der Olympiaringe O genauso angeben wie die Menge der Ziffern im Dezimalsystem Z:

$$\begin{aligned} W &= \{\text{Montag, Dienstag, Mittwoch, Donnerstag, Freitag, Samstag, Sonntag}\} \\ O &= \{\text{blau, gelb, schwarz, grün, rot}\} \\ Z &= \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\} \end{aligned}$$

Dabei ist die hier gewählte Reihenfolge willkürlich und unerheblich für unsere Definition der Menge, wir hätten also genauso gut schreiben können:

$$O = \{\text{gelb, blau, rot, grün, schwarz}\}$$

Wir werden im Laufe der Veranstaltung fast immer mit Mengen arbeiten, die aus Zahlen bestehen. Dabei ist es oft hilfreich, die Menge nicht über die Aufzählung ihrer Elemente anzugeben, sondern mit einer Aussage, die für alle Elemente der Menge wahr ist.

Wir können also die Menge M der Buchstaben im Wort “Mathe” entweder angeben als:

$$M = \{\text{M, a, t, h, e}\},$$

oder mit der folgenden Aussage über ihre Elemente:

$$M = \{m \mid m \text{ ist ein Buchstabe des Worts 'Mathe'}\}.$$

Nun können wir Aussagen mit Mengen kombinieren, indem wir eine (oder mehrere) Variable(n) in der Aussage vorkommen lassen.

Definition 1.7 (Aussage mit Variable)

Wie bezeichnen eine Aussage, welche von einer freien Variable n abhängt, mit $A(n)$.

Beispiel 1.8 (Aussagen mit Variable)

Einige Beispiele für Aussagen mit freier Variable:

- $n > 17$
- Mittwoch hat mehr als n Buchstaben
- n^2 ist eine ungerade Zahl

Wir erkennen an den obigen Beispielen, dass für eine eindeutige Zuweisung von wahr oder falsch noch angegeben werden muss, aus welcher Menge die Variable n stammt. Deswegen heißt es in der letzten Definition auch "freie Variable", weil diese bisher noch zu gar keiner Menge gehört. Die Mengenangabe wird sehr häufig von einem sogenannten Quantor begleitet:

Definition 1.9 (Allquantor und Existenzquantor)

Sei $A(n)$ eine Aussage mit freier Variable n und M eine Menge. Wir definieren die folgenden zwei Quantoren:

- **Allquantor**

  Wir sagen: "Für alle n aus M gilt: $A(n)$".
  Wir schreiben:

  $$\forall\, n \in M \;:\; A(n)$$

  Diese Aussage ist genau dann wahr, wenn $A(n)$ für alle n aus M gilt.

- **Existenzquantor**

  Wir sagen: "Es existiert ein n aus M, für das gilt: $A(n)$".
  Wir schreiben:

  $$\exists\, n \in M \;:\; A(n)$$

  Diese Aussage ist genau dann wahr, wenn $A(n)$ für mindestens ein n aus M gilt.

Beispiel 1.10 (Allquantor und Existenzquantor)

Wir betrachten wieder die Menge der Wochentagsnamen W vom vorletzten Beispiel. Beispiele für Aussagen sind:

$$\exists\, w \in W \ : \ w \text{ endet auf 'tag'}$$

Gesprochen: Es existiert ein w in der Menge W, für das gilt: w endet auf "tag" (wahr).

$$\forall\, w \in W \ : \ w \text{ endet auf 'tag'}$$

Gesprochen: Für alle w in der Menge W gilt: w endet auf "tag" (falsch).

▶ Weitere Beispiele

Ab sofort werden wir in allen Definitionen auch immer wieder mit Mengen verknüpfte Aussagen sehen, sodass Sie sich schnell an die Schreibweise gewöhnen werden. Als nächstes definieren wir, wie wir zwei Mengen vergleichen können.

Definition 1.11 (Mengenrelationen)

Seien A und B Mengen, dann gilt

- $A = B$, wenn

$$(\forall a \in A : a \in B) \wedge (\forall b \in B : b \in A)$$

  In Worten: Jedes Element von A ist auch ein Element von B und jedes Element von B ist auch ein Element von A.
  Wir sagen A und B sind *äquivalent* (oder gleich).

- $A \subseteq B$, wenn

$$\forall a \in A : a \in B$$

  In Worten: Iedes Element von A ist auch ein Element von B.
  Wir sagen: A ist eine *Teilmenge* von B.

- $A \subset B$, wenn

$$(A \subseteq B) \wedge \neg(A = B)$$

  In Worten: A ist eine Teilmenge von B und die Mengen sind nicht gleich.
  Wir sagen: A ist eine *echte Teilmenge* von B.

- Für $\neg(A = B)$ schreiben wir kurz $A \neq B$

Wir haben hier die echte Teilmenge definiert, indem wir die Definitionen von Gleichheit und Teilmenge kombiniert haben. Es wird noch häufiger vorkommen, dass wir aus vorherigen Definitionen wieder neue Definitionen ableiten. Sie können zur Übung einmal versuchen die Gleichheit über die Teilmenge zu definieren.

Wir haben hier auf eine Schreibweise mit Quantoren zurückgegriffen, um Sie daran zu gewöhnen. Allerdings werden Sie in vielen Lehrbüchern eher eine Definition über Implikation und Äquivalenzpfeile finden, wobei man man den $\forall$-Quantor häufig weglässt. Sie können sich leicht selbst davon überzeugen, dass folgende Definitionen äquivalent zur obigen Definition sind:

- $A = B$, wenn $\quad a \in A \Leftrightarrow a \in B$
- $A \subseteq B$, wenn $\quad a \in A \Rightarrow a \in B$

Wir können Mengen auch kombinieren und damit neue Mengen erhalten:

Definition 1.12 (Verknüpfung von Mengen)

Seien A und B Mengen. Dann sind die folgenden Mengenverknüpfungen definiert:

- **Vereinigung:** $A \cup B := \{m \mid m \in A \vee m \in B\}$
- **Schnitt:** $A \cap B := \{m \mid m \in A \wedge m \in B\}$
- **Differenz:** $A \setminus B := \{m \mid m \in A \wedge m \notin B\}$
- **Kartesisches Produkt:** $A \times B := \{(m, n) \mid m \in A \wedge n \in B\}$

Der Vereinigungs- und Schnittoperator kann auch auf mehr als zwei Mengen verallgemeinert werden. Sei $\mathcal{N}$ eine Menge von Mengen, dann definieren wir

- $\bigcup_{M \in \mathcal{N}} M := \{m \mid \exists M \in \mathcal{N} \,:\, m \in M\}$
- $\bigcap_{M \in \mathcal{N}} M := \{m \mid \forall M \in \mathcal{N} \,:\, m \in M\}$

Definieren wir darüber hinaus eine Menge U (für Universum), als Vereinigung aller Mengen, können wir außerdem das Komplement definieren:

- **Komplement** $A^c = U \setminus A$

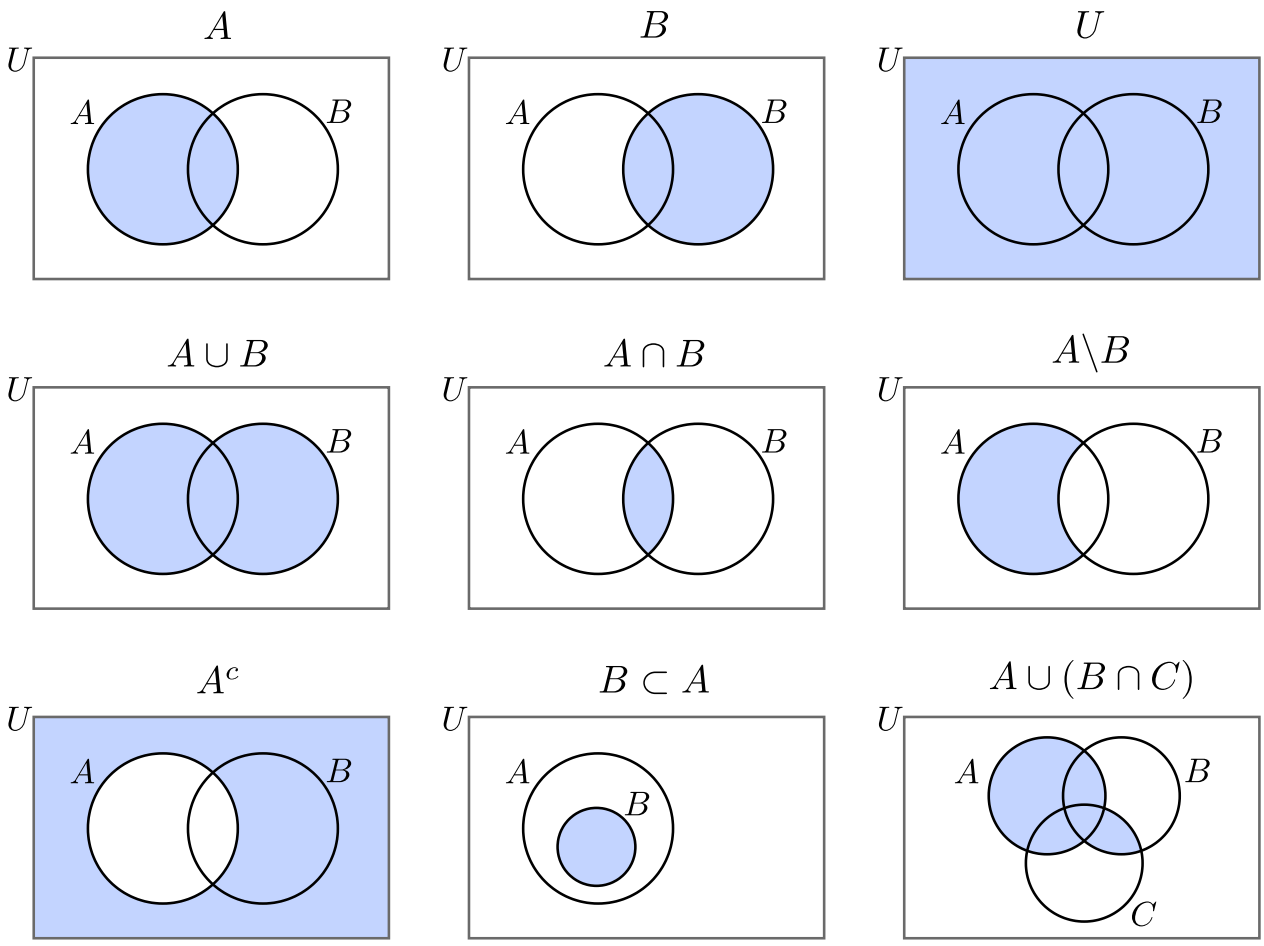
Beispiel 1.13 (Verknüpfung von Mengen)

Sei $M_1 = \{1, 3, 5, 7\}, M_2 = \{1, 2, 3\}, M_3 = \{2, 4\},\ \mathcal{N} = \{M_1, M_2, M_3\}$

Dann gilt

$$
\begin{aligned}
M_1 \cup M_2 &= \{1, 2, 3, 5, 7\} \\
M_1 \cap M_2 &= \{1, 3\} \\
M_1 \setminus M_2 &= \{5, 7\} \\
M_1 \cap M_3 &= \varnothing \\
M_1 \cap \mathbb{N} &= M_1 \\
M_2 \times M_3 &= \{(1, 2), (1, 4), (2, 2), (2, 4), (3, 2), (3, 4)\} \\
\bigcup_{M \in \mathcal{N}} M &= \{1, 2, 3, 4, 5, 7\}
\end{aligned}
$$

Die Mengenverknüpfungen lassen sich auch gut in sogenannten Venn-Diagrammen darstellen. Die Venn-Diagramme der Basisverknüpfungen sind in folgender Abbildung gezeigt.



Vermutlich ist Ihnen bereits aufgefallen, dass die Mengenverknüpfungen sehr ähnlich sind zu den Aussagenverknüpfungen aus Definition 1.3. Tatsächlich können Venn-Diagramme ebenfalls dabei helfen, Verknüpfungen von Aussagen graphisch darzustellen. Verknüpft man beispielsweise drei Aussagen A, B, C und stellt sich als Beispielaussagen $m \in M_A$ (für A), $m \in M_B$ (für B), $m \in M_C$ (für C) vor, dann ist eine Aussagenverknüpfung wie $A \wedge (B \vee C)$ äquivalent zu $m \in M_A \cap (M_B \cup M_C)$ und kann über ein Venn-Diagramm veranschaulicht werden. Versuchen Sie sich als Übung zu überlegen, wie die Venn-Diagramme der Aussagenverknüpfungen aus Definition 1.3 aussehen oder verdeutlichen Sie sich die Äquivalenzen aus Satz 1.4 mithilfe eines Venn-Diagramms. Zum Experimentieren mit Venn-Diagrammen, können Sie die folgende Demo benutzen:

▶ Demo: Venn-Diagramme

1.2 Beweistechniken

Den Nachweis der Richtigkeit einer Aussage auf Basis bekannter wahrer Aussagen nennt man *Beweis*. Bisher haben wir nur Beweise durch Wahrheitstafeln kennen gelernt. Diese Beweismethode wird allerdings nur in der grundlegenden Aussagenlogik verwendet. In diesem Kapitel stellen wir die gängisten Beweismethoden vor, welche wir in der gesamten Vorlesung immer wieder verwenden werden.

Wichtige neue Aussagen oder Gruppen von Aussagen nennen wir *Sätze*. In der Regel werden am Anfang eines Satzes eine Reihe von Voraussetzungen genannt, und anschließend eine Aussage getätigt, die unter diesen Voraussetzungen wahr sein soll. Der Beweis versucht nun, die Satzaussage auf *wahr* abzubilden. Zur Verfügung stehen dafür die Annahmen des Satzes und die sogenannten Axiome. Axiome sind Grundsätze, die man als wahr akzeptiert und damit keinen Beweis benötigen. In der Mathematik versucht man, mit möglichst wenigen dieser Axiome auszukommen. Wenn bereits andere Sätze bewiesen wurden, können auch diese bei der Beweisführung verwendet werden.

Am Ende des Beweises gönnt sich der Mathematiker ein kleines Quadrat (siehe rechts unten), um seine Freude über den gelungenen Beweis auszudrücken. Alternativ liest man auch öfter q.e.d. (quod erat demonstrandum) von Mathematikern, die mit ihren Lateinkenntnissen angeben wollen oder auch wzbw/wzzw (was zu beweisen/zeigen war).

□

Beispiele für die einzelnen Beweise werden erst im nächsten Kapitel folgen, da wir bisher noch nicht unsere Axiome eingeführt haben. Zusätzlich gibt es noch unseren kleinen Exkurs Beweistechniken. Hier haben wir separat für jede Kategorie einfache Beweise geführt, bei denen die natürlichen Zahlen als bekannt vorausgesetzt werden.

Direkter Beweis

Der Direkte Beweis ist vermutlich die am häufigsten eingesetzte Beweismethode. Hierbei nutzt man Implikationen, um von den Voraussetzungen des Satzes über die Axiome und bereits bewiesene Sätze auf die Behauptung des Satzes zu schließen. Dabei werden häufig mehrere Implikationen aneinandergereiht.

Dies ist zulässig, da für drei Aussagen A, B, C gilt:

$$((A \Rightarrow B) \wedge (B \Rightarrow C)) \quad \Leftrightarrow \quad (A \Rightarrow C),$$

wovon man sich leicht mit einer Wahrheitstafel überzeugen kann. Das heißt, eine Kette von Implikationen ist gleichbedeutend damit, dass aus der ersten Aussage, die letzte folgt.

Häufig sollen nicht nur Implikationen ($\Rightarrow$) sondern Äquivalenzen ($\Leftrightarrow$) gezeigt werden. Dafür verwendet man entweder analog zur Implikationskette eine Äquivalenzkette, also eine Reihe äquivalenter Aussagen, die von den Voraussetzungen des Satzes bis zur Behauptung des Satzes führt. Alternativ führt man zwei separate Beweise mit jeweils einer Implikation, denn es gilt:

$$(A \Leftrightarrow B) \quad \Leftrightarrow \quad ((A \Rightarrow B) \wedge (B \Rightarrow A))\,.$$

Eine Äquivalenz ist also gleichbedeutend mit zwei Implikationen, einmal von “links nach recht” und einmal von “rechts nach links”. Wir leiten die beiden Fälle dann auch oft mit ‘$\Rightarrow$ Richtung’ und ‘$\Leftarrow$ Richtung’ ein.

Indirekter Beweis - Kontraposition

Der Indirekte Beweis, oder der Beweis der Kontraposition ist eine Sonderform des direkten Beweises. Dabei wird folgende Äquivalenz ausgenutzt:

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

Ein kleines Beispiel:

- Ein Apfel ist eine Frucht ($A \Rightarrow F$ - direkt).
- Alles, was keine Frucht ist, kann auch kein Apfel sein ($\neg F \Rightarrow \neg A$ - Kontraposition).

Manchmal ist es einfacher die Kontraposition zu beweisen als die ursprüngliche Richtung der Implikation. Daher lohnt es sich, auch die Kontrapositon eines Satzes zu formulieren und dann zu überlegen, welcher Beweis einfacher zu führen ist. Andersherum ist es oft ratsam, sich zu einer bewiesenen Implikation auch noch die Kontraposition zu formulieren, da man dadurch häufig eine andere Sichtweise auf den Satz bekommt.

Widerspruchsbeweis

Sollen wir für einen Satz eine Aussage A beweisen, nehmen wir für den sogenannten *Widerspruchsbeweis* (oder auch *reductio ad absurdum*) an, dass A nicht gilt, Davon ausgehend versuchen wir auf eine andere Aussage C zu schließen, von der wir wissen, dass sie falsch ist. Damit kann unsere ursprüngliche Annahme, dass A nicht gilt, nicht richtig sein und somit folgt, dass A wahr sein muss. Der Widerspruch wird gerne mit einem Blitzsymbol ($\lightning$) markiert

Die aussagenlogische Grundlage für einen Widerspruchsbeweis ist ein wenig komplizierter, aber wir gehen es schrittweise noch einmal durch: Wir nehmen an, dass A nicht gilt, also $\neg A$ wahr ist. Daraus folgern wir eine Aussage C ($\neg A \Rightarrow C$), von der wir wissen, dass sie falsch ist, also $\neg C$ wahr ($(\neg A \Rightarrow C) \wedge \neg C$). Daraus folgt A, also insgesamt

$$((\neg A \Rightarrow C) \wedge \neg C) \quad \Rightarrow \quad A$$

Überzeugen Sie sich gerne mit einer Wahrheitstafel von der Richtigkeit dieser Aussage.

Ein Beispiel aus dem Alltag wäre:

- Ich finde meinen Schlüsselbund nicht.
- Ich nehme an, ich hätte ihn zuhause vergessen.
- Ich bin aber gerade in der Uni und mit dem Fahrrad hergekommen.
- Das Fahrrad ist angeschlossen. $\lightning$
- Also kann ich den Schlüssel nicht zuhause vergessen haben.

Vollständige Induktion

Die letzte Beweismethode ist die sogenannte *vollständige Induktion*. Diese ist besonders gut dazu geeignet, eine Aussage $A(n)$ zu beweisen, die für alle natürlichen Zahlen n gelten sollen. Wir können unmöglich die Aussage für jede Zahl ausprobieren, da es unendlich viele natürliche Zahlen gibt. Hier kommt die vollständige Induktion ins Spiel, diese erfolgt immer in den folgenden 3 Schritten:

1. Induktionsanfang: Wir beweisen $A(n = 1)$.
2. Induktionsvoraussetzung: Wir nehmen an, $A(n)$ gelte für ein $n \in \mathbb{N}$
3. Induktionsschritt: Wir beweisen die Implikation: $A(n) \Rightarrow A(n + 1)$

Wenn uns das gelingt, gilt die Aussage für alle $n \in \mathbb{N}$, denn wir haben gezeigt, dass sie für $n = 1$ gilt, daraus folgt, dass sie für $n = 2$ gilt, daraus folgt dass sie für $n = 3$ gilt, usw...

Oft vergleicht man die vollständige Induktion mit dem Versuch eine Leiter zu besteigen, von der man nicht weiß, wie hoch sie ist. Wenn man aber sicher weiß, dass man auf die erste Sprosse kommt (Induktionsanfang) und von egal welcher Sprosse ausgehend (Induktionsvoraussetzung) immer weiß, wie man die jeweils nächste erreicht (Induktionsschritt), dann kann man die komplette Leiter erklimmen.

Es gibt verschiedene Varianten der vollständigen Induktion. So muss man zum einen nicht bei $n = 1$ starten. Manche Aussagen gelten nur für alle $n \geq n_0$. Man kann eine Aussage auch für alle ganzen Zahlen beweisen, indem man bei $n = 0$ startet und dann sowohl in negative als auch in positive Richtung eine vollständige Induktion führt.

2 Zahlenmengen

2.1 Historie der Zahlen

Für einen etwas sanfteren Einstieg in das Thema wird in der Vorlesung zunächst die historische Entwicklung der Zahlen betrachtet. Ähnlich wie es auch in der Schule eingeführt wird, beginnt diese Entwicklung mit den natürlichen Zahlen und erst danach wurden nach und nach auch rationale, reelle und schließlich komplexe Zahlen von der Menschheit akzeptiert. Jeder Schritt war mit viel Widerstand verbunden, da Menschen sich bekanntermaßen nur äußerst ungern an Neues gewöhnen wollen. Da das Thema deutlich besser in Video- als in Schriftform konsumierbar ist, verweisen wir an diese Stelle auf die entsprechende Vorlesungsaufzeichnung.

2.2 Axiomatische Einführung der reellen Zahlen

Wir werden nun die Axiome einführen, welche die Basis für alle weiteren Sätze in dieser Vorlesung bilden. Ein Axiom ist ein Grundsatz, den man ohne Beweis als wahr anerkennt. Dabei versucht man in der Mathematik, so wenig Axiome wie mögliche zu verwenden und dagegen so viel wie möglich daraus abzuleiten, ohne weitere Axiome zu benötigen. Oft wird in der Schule größtenteils axiomatisch gearbeitet (ohne dieses Wort dafür zu verwenden): man lernt “Rechenregeln”, die man immer wieder anwendet und von denen man irgendwann “glaubt”, dass sie wahr sind. Im Studium dagegen versuchen wir, diesen Ansatz weitestgehend zu vermeiden und den überwiegenden Teil des Regelwerks, das wir uns über die Jahre erarbeiten, auch zu beweisen. Sie dürfen sich in dieser Veranstaltung also gerne eine Grundskepsis angewöhnen und immer wieder hinterfragen:

- Warum darf ich das so machen?
- Warum gilt das?

Da auch fast alle folgenden Veranstaltungen im Studium immer wieder Beweise für sehr abstrakte Konzepte durchführen werden, ist ein Ziel dieser Veranstaltung, Sie darauf entsprechend vorzubereiten.

Die Definition der reellen Zahlen $\mathbb{R}$ basiert auf drei Axiomklassen. Diese Axiomklassen kann man stark vereinfacht wie folgt auflisten:

1. Grundregeln der Addition und Multiplikation (Körperaxiome)
2. Grundregeln zum Größenvergleich zweier Zahlen (Ordnungsaxiome)
3. Jeder ‘Nachbar’ einer reellen Zahl ist wieder eine reelle Zahl (Vollständigkeitsaxiom)

Wir werden in diesem Kapitel die Axiome in dieser Reihenfolge einzeln durchgehen. Im nächsten Kapitel werden wir dann Folgerungen aus den Axiomen ableiten. Im Wesentlichen sind diese Folgerungen all die “Rechenregeln”, die Sie bereits aus der Schule kennen, nur werden wir diese hier beweisen, indem wir sie auf die Axiome zurückführen.

Beginnen wir also mit der Definition eines sogenannten Körpers.

Definition 2.1 (Körper)

Ein Körper ist eine Menge K, auf der zwei Verknüpfungen, die wir Addition ($+$) und Multiplikation ($\cdot$) nennen, definiert sind und für die folgende Eigenschaften erfüllt sind:

Eigenschaften der Addition:

(K1)	$\forall a, b, c \in K\colon\ a + (b + c) = (a + b) + c$	(Assoziativität $+$)
(K2)	$\forall a, b \in K\colon\ a + b = b + a$	(Kommutativität $+$)
(K3)	$\exists 0 \in K\ \forall a \in K\colon\ a + 0 = a$	(neutrales Element $+$)
(K4)	$\forall a \in K\ \exists (-a) \in K\colon\ (-a) + a = 0$	(inverses Element $+$)

Das neutrale Element der Addition nennen wir auch *Nullelement* und das inverse Element der Addition *negatives Element*.

Eigenschaften der Multiplikation:

(K5)	$\forall a, b, c \in K\colon\ a \cdot (b \cdot c) = (a \cdot b) \cdot c$	(Assoziativität $\cdot$)
(K6)	$\forall a, b \in K\colon\ a \cdot b = b \cdot a$	(Kommutativität $\cdot$)
(K7)	$\exists 1 \in K \setminus \{0\}\ \forall a \in K \setminus \{0\}\colon\ a \cdot 1 = a$	(neutrales Element $\cdot$)
(K8)	$\forall a \in K \setminus \{0\}\ \exists a^{-1} \in K\colon\ a^{-1} \cdot a = 1$	(inverses Element $\cdot$)

Das neutrale Element der Multiplikation nennen wir auch *Einselement* und das inverse Element der Multiplikation kurz *inverses Element*.

Kombination von Addition und Multiplikation:

(K9)	$\forall a, b, c \in K\colon\ a \cdot (b + c) = a \cdot b + a \cdot c$	
	$\forall a, b, c \in K\colon\ (a + b) \cdot c = a \cdot c + b \cdot c$	(Distributivgesetze)

Achtung: Dies ist bisher nur eine gewöhnliche Definition und noch kein Axiom. Dieses folgt erst jetzt:

Axiom 1

Die reellen Zahlen mit der üblichen Addition ($+$) und Multiplikation ($\cdot$) bilden einen Körper.

Da wir dies zum Axiom erhoben haben, müssen wir hier nichts beweisen oder die Definition nachprüfen. Das Axiom nehmen wir einfach für die Zukunft als wahr an. Es bildet damit einen der drei Steine des Fundaments der Analysis.

Es gibt verschiedene Körper in der Mathematik. Beispielsweise sind auch die rationalen Zahlen mit der üblichen Addition und Multiplikation ein Körper. Und auch die komplexen Zahlen, die wir später einführen werden, sind ein Körper. Die Operationen $+$ und $\cdot$ müssen auch nicht die übliche Addition und Multiplikation sein. Wichtig ist lediglich, dass die Eigenschaften K1 bis K9 erfüllt sind. Zur Vereinfachung treffen wir für die Zukunft noch die Vereinbarung, dass wir statt $a \cdot b$ auch kurz ab schreiben. Darüber hinaus vereinbaren wir, dass "Punkt vor Strich" gilt, wodurch wir z.B. kurz $a + bc$ schreiben können statt $a + (bc)$.

Beispiel 2.2 (Der kleinste Körper)

Nach der Definition muss ein Körper ein Nullelement (neutrales Element der Addition) und ein Einselement (neutrales Element der Multiplikation) enthalten. Wir werden später noch zeigen, dass diese in keinem Körper identisch sind. Man kann aber einen Körper aus nur diesen zwei Elementen bauen, indem man die Addition definiert als

$$\begin{aligned} 0 + 0 &= 0 \\ 0 + 1 &= 1 \\ 1 + 0 &= 1 \\ 1 + 1 &= 0 \end{aligned}$$

und die Multiplikation als

$$\begin{aligned} 0 \cdot 0 &= 0 \\ 0 \cdot 1 &= 0 \\ 1 \cdot 0 &= 0 \\ 1 \cdot 1 &= 1 \end{aligned}$$

Die Körpereigenschaften K1–K9 lassen sich hier sehr schnell zeigen, also ist $(\{0, 1\}, +, \cdot)$ ein Körper.

Als nächstes brauchen wir auf dem Körper der reellen Zahlen eine Ordnung, also eine Möglichkeit, verschiedene reelle Zahlen miteinander zu vergleichen. Dafür dient die nächste Definition.

Definition 2.3 (Geordneter Körper)

Ein Körper $(K, +, \cdot)$, in dem eine Teilmenge P – genannt die "positiven Zahlen" – mit den folgenden Eigenschaften existiert, nennt man einen (an)geordneten Körper.

(O1) $\forall\, a \in K$ gilt genau eine der drei Beziehungen $(a \in P), (a = 0), (-a \in P)$

(O2) $\forall\, a, b \in P :\ a + b \ \in P$

(O3) $\forall\, a, b \in P :\ ab \ \in P$

Warum diese Definition einen Vergleich möglich macht, sehen wir an der folgenden, darauf aufbauenden Definition.

Definition 2.4 (Vergleichsoperatoren)

Wir definieren für einen geordneten Körper K und zwei beliebige Elemente $a, b \in K$ die folgenden Vergleichsoperatoren:

$$\begin{aligned} a < b &:\Leftrightarrow (b - a) \in P \\ a > b &:\Leftrightarrow b < a \\ a \leq b &:\Leftrightarrow (a < b) \vee a = b \\ a \geq b &:\Leftrightarrow b \leq a \end{aligned}$$

Sobald wir also Zahlen als positiv oder nicht-positiv identifizieren können, sind die üblichen Größenvergleiche zwischen beliebigen Zahlen möglich. Um die Welt der Größenvergleiche für die reellen Zahlen zu eröffnen, erheben wir Folgendes zum Axiom:

Axiom 2

Der Körper der reellen Zahlen $(\mathbb{R}, +, \cdot)$ ist ein geordneter Körper.

Sobald ein Vergleich von Zahlen möglich ist, können wir auch das Maximum und das Minimum einer Menge bestimmen, also das größte oder kleinste Element. Die nächste Definition bietet sich hier also gut an:

Definition 2.5 (Minimum, Maximum)

Sei $(K, +, \cdot)$ ein geordneter Körper und $M \subseteq K$ eine Teilmenge des Körpers. Dann nennen wir ein Element $m \in M$

- Maximum von M, oder $\max(M)$, wenn $\forall n \in M : \ m \geq n$,
- Minimum von M, oder $\min(M)$, wenn $\forall n \in M : \ m \leq n$.

Hier schließt sich die Frage an, ob jede Menge ein Maximum/Minimum hat. Die reellen Zahlen selbst sind dafür direkt ein Gegenbeispiel, da man für jede reelle Zahl immer noch eine größere und kleinere reelle Zahl finden kann. Bei Mengen mit endlich vielen Elementen, wie zum Beispiel $M_1 = \{1, 2, 3\}$, ist es dagegen recht einfach, ein Minimum und ein Maximum zu finden. Aber auch unendliche Mengen, wie zum Beispiel die Menge $M_2 = \{m \in \mathbb{R} \mid m \leq 2 \wedge m \geq 1\}$, können ein klares Minimum und Maximum haben.

Betrachten wir nun die Menge

$$M_3 = \left\{ \frac{n}{n+1} \;\middle|\; \forall n \in \mathbb{N} \right\} = \left\{ \frac{1}{2}, \frac{2}{3}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{7}{8}, \dots \right\}.$$

Hier lässt sich recht einfach das Minimum $1/2$ bestimmen. Aber obwohl klar ist, dass jedes Element kleiner als Eins ist, da der Nenner des Bruchs kleiner ist als der Zähler, ist $\max(M_3) \neq 1$, denn 1 ist kein Element von M_3, und die Definition des Maximums fordert, dass das Maximum ein Element der Menge sein muss. Für solche Fälle, in denen wir ein "Maximum außerhalb der Menge" angeben können, ist die nächste Definition gedacht.

Definition 2.6 (obere/untere Schranke, Beschränktheit)

Sei $(K, +, \cdot)$ ein geordneter Körper und $M \subseteq K$ eine Teilmenge des Körpers. Dann nennen wir ein Element $m \in K$

- obere Schranke von M, wenn $\forall n \in M \; : \; m \geq n$,
- untere Schranke von M, wenn $\forall n \in M \; : \; m \leq n$.

Existiert für M eine obere und eine untere Schranke, sagen wir: M ist beschränkt.

Machen Sie sich den Unterschied zwischen den Definitionen 2.5 und 2.6 bewusst: Eine obere/untere Schranke muss kein Element der Menge sein, ein Minimum/Maximum schon. Für die Menge M_3 von oben wäre $m = 1$ eine obere Schranke, aber kein Maximum. Allerdings wäre auch $m = 2$, oder allgemein jedes $m \geq 1$ eine obere Schranke. Um diese Mehrdeutigkeit zu beseitigen, brauchen wir noch eine weitere Definition.

Definition 2.7 (Supremum/Infimum)

Sei $(K, +, \cdot)$ ein geordneter Körper und $M \subseteq K$ eine Teilmenge des Körpers. Dann nennen wir ein Element $m \in K$

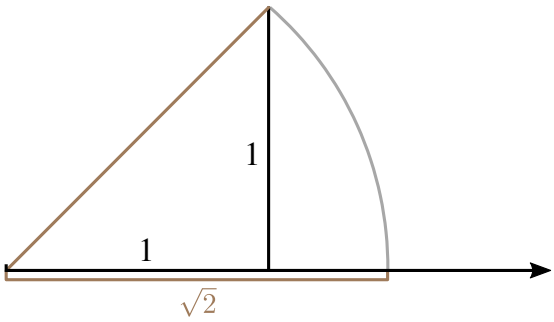
- Supremum von M, oder $\sup(M)$, wenn m die kleinste obere Schranke ist,
- Infimum von M, oder $\inf(M)$, wenn m die größte untere Schranke ist.

Hiermit wäre also $\sup(M_3) = 1$. Achtung, selbst wenn dies (wie hier) offensichtlich ist, muss eine solche Behauptung bewiesen werden. Da wir allerdings erst im nächsten Kapitel die wichtigsten Rechenregeln kennen lernen werden, lassen wir den Beweis hier ausnahmsweise aus. Führen Sie den Beweis aber gerne als Übung durch. Wir werden in den nächsten Kapiteln noch häufig mit beschränkten Mengen arbeiten, da diese viele nützliche Eigenschaften haben. Machen Sie sich bewusst, dass die letzten vier Definitionen alle auf der Definition des geordneten Körpers aufbauen.

Nun sind wir fast fertig mit der axiomatischen Einführung der reellen Zahlen. Den Großteil der Gesetze, die in der Schule benutzt werden, kann man aus den bisherigen Axiomen ableiten. Dies werden wir in Kapitel 2.3 auch ausführlich tun. Es gibt nur noch ein Problem: Sowohl die

Körperaxiome als auch die Ordnungsaxiome gelten auch für die rationalen Zahlen. Zur Zeit unterscheiden sich unsere reellen Zahlen also noch nicht von den rationalen Zahlen. Aber warum sind wir mit den rationalen Zahlen nicht einfach zufrieden? Die Antwort ist, dass wir die reellen Zahlen so definieren wollen, dass sie alle Zahlen enthalten, die wir auf einem Zahlenstrahl angeben/konstruieren können. Es soll also keine "Lücken" geben. Sie wissen aber vielleicht bereits aus der Schule, dass auch sogenannte irrationale (nicht-rationale) Zahlen existieren. Das prominenteste Beispiel ist $\sqrt{2}$, für das wir später noch beweisen werden, dass es nicht rational sein kann, also nicht als Bruch aus ganzen Zahlen darstellbar ist. Allerdings konnten bereits die alten Griechen eine Strecke konstruieren, deren Länge $x^2 = 2$ erfüllt. Bei den Bruchzahlen $\mathbb{Q}$ gibt es also noch "Lücken" auf der Zahlengerade, daher fehlt uns noch eine Eigenschaft der reellen Zahlen, um diese Lücken zu schließen.

Wir könnten jetzt einfach etwas fordern wie "$\mathbb{R}$ hat keine Lücken auf der Zahlengerade". Auch wenn das zunächst plausibel klingen mag, fehlt uns hier eine klare Definition der "Lücken" und der "Zahlengerade" und das, ohne dafür die reellen Zahlen zu nutzen. Genauso können wir auch nicht fordern, dass die reellen Zahlen die Lösung der Gleichung $x^2 = 2$ beinhalten sollen. Dann wären zwar die reellen und rationalen Zahlen unterschiedlich definiert, aber wir hätten damit noch lange nicht alle "Lücken" geschlossen. Wir suchen also eine möglichst Grundlegende Eigenschaft der reellen Zahlen, welche die rationalen Zahlen nicht erfüllen.



Genau darum geht es im letzten Axiom, dem sogenannten *Vollständigkeitsaxiom*. Genau genommen gibt es nicht *das* Vollständigkeitsaxiom, denn im Gegensatz zu den ersten beiden Axiomen, die in fast jedem Lehrbuch bis auf kleine Abwandlungen identisch sind, gibt es für die Vollständigkeit der reellen Zahlen sehr unterschiedliche Axiome, die sich allerdings alle jeweils auseinander herleiten lassen. Man wählt also *ein* Vollständigkeitsaxiom und die übrigen werden zu normalen Sätzen, die sich mithilfe des gewählten Axioms beweisen lassen. Geben wir also zunächst unsere Definition der Vollständigkeit an:

Definition 2.8 (vollständiger Körper)

Wir nennen einen geordneten Körper $(K, +, \cdot)$ vollständig, wenn für jede nach oben beschränkte Teilmenge $M \subseteq K$ gilt

$$\sup(M) \in K.$$

Axiom 3

Der Körper der reellen Zahlen $(\mathbb{R}, +, \cdot)$ ist vollständig.

Zunächst einmal fällt auf, dass dieses Axiom im Vergleich zu unseren eigenen Versuchen weiter oben mit unseren bisherigen Definitionen sehr klar formuliert ist. Wir könnten hier übrigens auch genauso gut das Infimum anstelle des Supremums verwenden.

Doch was sagt es aus? Die Aussage des Axioms ist uninteressant für Teilmengen, bei denen das Supremum ein Element der Teilmenge selbst ist (also das Maximum der Menge). Denn wenn das Element innerhalb einer Teilmenge von $\mathbb{R}$ liegt, liegt es natürlich auch in $\mathbb{R}$ selbst und Selbiges würde auch gelten, wenn wir $\mathbb{R}$ durch $\mathbb{Q}$ ersetzen. Interessanter sind Teilmengen, bei denen das Supremum nicht in der Menge liegt, wie zum Beispiel unsere zuvor definierte Menge M_3, bei denen Elemente der Menge beliebig nah an das Supremum herankommen, es aber nie ganz erreichen. Unser Axiom besagt, dass auch solche Suprema, die kein Element der Teilmenge sind, in $\mathbb{R}$ enthalten sein sollen. Oder anders formuliert: Wenn wir uns einer Zahl x durch andere Elemente aus $\mathbb{R}$ beliebig dicht nähern können, dann soll auch $x \in \mathbb{R}$ sein. Deswegen nennt man Körper mit dieser Eigenschaft *vollständig*, weil damit keine "benachbarten" Elemente des Körpers existieren können, die nicht auch zum Körper gehören.

Aber warum gilt das Axiom nicht auch für die rationalen Zahlen $\mathbb{Q}$? Um dies zu beweisen, müssen wir eine Teilmenge von $\mathbb{Q}$ finden, deren Supremum eine Zahl ist, die nicht in $\mathbb{Q}$ liegt. Diesen Beweis werden wir im folgenden Kapitel 2.4 führen. Machen Sie sich an dieser Stelle keine Sorgen, wenn Ihnen die Vollständigkeit der reellen Zahlen nicht ganz einleuchtend erscheint. Wir werden im Laufe der Vorlesung noch häufig darauf zurückkommen und viele Beispiele sehen.

Über diese drei Axiome besitzen die reellen Zahlen nun alle Eigenschaften, die wir in der Analysis benötigen. Alle nun folgenden Aussagen in dieser Vorlesung lassen sich aus diesen Axiomen herleiten. Als Nächstes können wir uns also endlich an die ersten Beweise wagen, anstatt immer nur zu definieren. Zur Erinnung: in Kapitel 1.2 finden Sie eine Übersicht über die wichtigsten Beweismethoden.

2.3 Folgerungen aus den Axiomen

Wie bereits angekündigt, werden wir alle Regeln/Aussagen, die über die eben eingeführen Axiome hinausgehen, daraus ableiten. Die Axiome sind damit die einzigen Grundsätze, was wir *glauben*, den Rest wollen wir *wissen*. Es ist Ihnen vielleicht aufgefallen, dass wir bisher noch keine Subtraktion oder Division eingeführt haben, oder einfache Regeln, wie Kürzungsregeln ($2/4 = 1/2$), oder binomische Formeln.

Hierzu werden wir nun die ersten Beweise durchführen, bei denen wir, ausgehend von den Axiomen, diese Rechenregeln herleiten. Die Beweise sind hier noch nicht besonders anspruchsvoll, aber gerade weil uns vieles als "ist doch logisch" vorkommt, fallen gerade diese Beweise am Anfang oft schwer.

Folgerungen aus den Körpereigenschaften der Addition

Wir starten mit den ersten vier Körpereigenschaften, die sich nur mit der Addition beschäftigen. Zur Wiederholung geben wir sie hier erneut an

(K1)	$\forall a, b, c \in K: \ a + (b + c) = (a + b) + c$	(Assoziativität $+$)
(K2)	$\forall a, b \in K: \ a + b = b + a$	(Kommutativität $+$)
(K3)	$\exists 0 \in K \ \forall a \in K \ : \ a + 0 = a$	(neutrales Element $+$)
(K4)	$\forall a \in K \ \exists (-a) \in K \ : \ (-a) + a = 0$	(inverses Element $+$)

Fangen wir mit ein paar ganz simplen Folgerungen an:

Satz 2.9 (Folgerungen der Addition 1)

Ist K ein Körper, dann gilt für alle $a, b, c \in K$:

a. $0 + a = a$ und $a + (-a) = 0$

b. Die Summe von a, b, c ist unabhängig von der Reihenfolge und wir schreiben daher einfach $a + b + c$

▼ Beweis

a. Folgt sofort aus (K3) bzw (K4), wenn man (K2) anwendet.

b. Wir führen den Beweis einmal am Beispiel $(b + a) + c = a + (b + c)$. Jede andere der 12 Vertauschungs- und Klammerungsvarianten funktioniert analog:

$$(b + a) + c \overset{\text{K2}}{=} (a + b) + c \overset{\text{K1}}{=} a + (b + c)$$

□

Satz 2.9(b) kann man ebenfalls für eine beliebige endliche Anzahl von Summanden zeigen, allerdings ist die Beweisführung immer eine Aneinaderreihung von K1 und K2 und daher nicht sonderlich spannend hier anzuführen. Für unendlich viele Summanden gilt die Regel übrigens nicht immer, wie wir in Kapitel 3 noch sehen werden.

Im nächsten Satz sehen wir zum ersten Mal einen Eindeutigkeitsbeweis, in diesem Fall die Eindeutigkeit des 0-Elements und des $(-a)$-Elements. Eindeutigkeitsbeweise sollen stets zeigen, dass es kein zweites Element mit der gleichen Eigenschaft geben kann. Ein Eindeutigkeitsbeweis ist kein Existenzbeweis, also nur weil wir beweisen, dass es *höchstens* ein Element mit dieser Eigenschaft gibt, heißt es noch nicht, dass es auch *genau* eins gibt. Daher führt man häufig sowohl einen Existenzbeweis als auch einen Eindeutigkeitsbeweis durch. Letzterer wird fast immer per Widerspruchsbeweis geführt. In diesem Fall ist die Existenz des Nullelements und des negativen Elements bereits über K3 bzw K4 gegeben, das heißt wir müssen nur noch die Eindeutigkeit zeigen.

Satz 2.10 (Folgerungen der Addition 2)

Für einen Körper K gilt:

a. Das neutrale Element der Addition (Nullelement) ist eindeutig

b. Das inverse Element der Addition (negative Element) ist eindeutig.

c. Für alle $a \in K$ ist das negative Element des negativen Elements wieder a, oder kurz: $\forall a \in K : -(-a) = a$

▼ Beweis

a. Wir nehmen an, es gäbe ein zweites Element $0' \neq 0 \in K$ für das gilt $a + 0' = a$.
Dann folgt aus $a + 0 = a$ mit $a = 0'$:

$$0' + 0 = 0'$$

und aus $a + 0' = a$ mit $a = 0$:

$$0' + 0 = 0$$

zusammen ergeben die beiden Gleichungen also

$$0' = 0' + 0 \overset{\text{K2}}{=} 0 + 0' = 0.$$

Dies stellt einen Widerspruch zur Annahme $0' \neq 0$ dar. (↯)
Daher ist das Nullelement eindeutig.

b. Wir nehmen an, es gäbe ein zweites Element $a' \neq -a \in K$ für das gilt $a + a' = 0$ (*).
Dann folgt

$$\begin{aligned} -a &= -a + 0 && \text{(K3)} \\ &= -a + (a + a') && \text{(*)} \\ &= (-a + a) + a' && \text{(K1)} \\ &= 0 + a' && \text{(K4)} \\ &= a' && \text{(S.2.9(a))}. \end{aligned}$$

Dies stellt einen Widerspruch zur Annahme $a' \neq -a$ dar. (↯)
Daher ist das negative Element $(-a)$ eindeutig.

c. Es gilt nach K4 für (-a): $(-a) + (-(-a)) = 0$.
Allerdings gilt auch nach K4 für a: $0 = a + (-a) \overset{\text{K2}}{=} (-a) + a$
Wegen der Eindeutigkeit des negativen Elements (b), kann es für $(-a)$ keine zwei verschiedenen negativen Elemente geben. Daher folgt: $(-(-a)) = a$.

□

Ihnen fällt vielleicht auf, wie akribisch wir jeden Schritt mit einem unserer Axiome oder den Satzannahmen begründen. Dies ist besonders an dieser Stelle sehr wichtig, da man sonst leicht vergisst, warum wir eine Regel anwenden dürfen und stattdessen schnell wieder in unseren "Schulregelglauben" verfallen. Später werden wir natürlich nicht mehr für jede Vertauschung von zwei Summanden auf K2 verweisen. Aber für Ihre ersten Übungen sollten Sie sich dieses akribische Vorgehen angewöhnen.

Wir schließen die Additionsregeln mit einem sehr wichtigen Satz, in dem wir beweisen, dass wir eine Gleichung mit einer Addition wie gewohnt umstellen dürfen:

Satz 2.11 (Folgerungen der Addition 3)

In einem Körper K gilt für beliebige Elemente $a, b, x \in K$:

a. $a + x = b \Leftrightarrow x = b + (-a)$

b. $a = b \Leftrightarrow a + x = b + x$

c. Für die Summe $a + b$ ist das negative Element $-(a + b) = (-a) + (-b)$

▼ Beweis

a. Hier muss eine Äquivalenz gezeigt werden, also führen wir den Beweis in zwei Richtungen:
⇐-Richtung (Existenz) $a + x = b \Leftarrow x = b + (-a)$:

$$a + x \overset{\text{(Def.x)}}{=} a + b + (-a) \overset{\text{K2}}{=} a + (-a) + b \overset{\text{K4}}{=} b.$$

⇒-Richtung (Eindeutigkeit) $a + x = b \Rightarrow x = b + (-a)$:

$$x \overset{\text{K3}}{=} x + 0 \overset{\text{K4}}{=} x + a + (-a) \overset{\text{K2}}{=} a + x + (-a) \overset{\text{(Def.b)}}{=} b + (-a).$$

b. Zum Beweis nutzen wir (a), um die rechte Seite nach a aufzulösen:
$a + x = b + x \Leftrightarrow a = b + x + (-x) \overset{\text{K4}}{=} b$

c. Zum Einen ist $(-a) + (-b)$ eine Lösung der Gleichung $a + b + x = 0$ nach $2\times$(a).
Zum Anderen ist $-(a + b)$ eine Lösung der Gleichung $a + b + x = 0$ nach (K4).
Wegen der Eindeutigkeit der Lösung (a) muss gelten $-(a + b) = (-a) + (-b)$.

□

Wir benutzen in Zukunft meist die Kurzschreibweise $a + (-b) = a - b$, wodurch die Subtraktion als Addition einer negativen Zahl definiert ist.

Folgerungen aus den Körpereigenschaften der Multiplikation

Fahren wir fort mit den 4 Eigenschaften der Multiplikation:

(K5)	$\forall a, b, c \in K : \ a \cdot (b \cdot c) = (a \cdot b) \cdot c$	(Assoziativität $\cdot$)
(K6)	$\forall a, b \in K : \ a \cdot b = b \cdot a$	(Kommutativität $\cdot$)
(K7)	$\exists 1 \in K \setminus \{0\} \ \forall a \in K \setminus \{0\} : \ a \cdot 1 = a$	(neutrales Element $\cdot$)
(K8)	$\forall a \in K \setminus \{0\} \ \exists a^{-1} \in K : \ a^{-1} \cdot a = 1$	(inverses Element $\cdot$)

Vergleichen wir diese mit K1 bis K4 fällt auf, dass die Additionseigenschaften sehr ähnlich sind zu K5 bis K8, es muss lediglich überall $+$ durch $\cdot$ ersetzt werden und das neutrale und inverse Element haben ein eigenes Symbol. Man nennt Mengen, die diese 4 Eigenschaften bezüglich einer Operation erfüllen, eine Gruppe. Wir hätten also auch den Körper definieren können, indem wir sagen, dass ein Körper $(K, +, \cdot)$ die 4 Gruppeneigenschaften bezüglich $(K, +)$ und $(K \setminus \{0\}, \cdot)$ erfüllen muss. Dann hätte nur noch K9 gefehlt. Wir können also alle Folgerungen aus dem letzten Kapitel einfach auf die Multiplikation übertragen, ohne auch nur einen weiteren Beweis zu führen.

Satz 2.12 (Folgerungen der Multiplikation)

In einem Körper K gilt für beliebige Elemente $a, b, c, x \in K$:

a. $1 \cdot a = a$ und für $a \neq 0$ gilt $a \cdot a^{-1} = 1$

b. Das Produkt von a, b, c ist unabhängig von der Reihenfolge und wir schreiben daher einfach $a \cdot b \cdot c$.

c. Das neutrale Element der Multiplikation (Einselement) ist eindeutig

d. Das inverse Element der Multiplikation (inverses Element) ist eindeutig.

e. Das inverse Element des inversen Elements von a ist wieder a, oder kurz $(a^{-1})^{-1} = a$.

f. $\forall a \neq 0 \,:\, a \cdot x = b \Leftrightarrow x = b \cdot a^{-1}$

g. $\forall x \neq 0 \,:\, a = b \Leftrightarrow a \cdot x = b \cdot x$

h. Für das Produkt $a \cdot b \neq 0$ ist das inverse Element $(a \cdot b)^{-1} = a^{-1} \cdot b^{-1}$.

▼ Beweis

Ersetze in den letzten drei Sätzen jeweils:

- $+ \to \cdot$
- $0 \to 1$
- $-a \to a^{-1}$

□

Wie bereits geschrieben, benutzen wir in Zukunft meist die Kurzschreibweise $ab = a \cdot b$. Außerdem definieren wir die Division $\frac{a}{b} := ab^{-1}$. Die typischen Bruchrechnungsregeln ergeben sich ebenfalls aus den Körpereigenschaften:

Satz 2.13 (Regeln der Bruchrechnung)

In einem Körper K gilt für beliebige Elemente $a, c \in K$ und $b, d, e \in K \setminus \{0\}$:

a. $\frac{a}{b} = \frac{c}{d} \quad \Leftrightarrow \quad ad = bc$

b. $\frac{ae}{be} = \frac{a}{b}$

c. $\frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}$

d. $\frac{\frac{a}{b}}{\frac{e}{d}} = \frac{ad}{be}$

▼ Beweis

a. Da $b, d \neq 0$, ist nach (kleiner Vorgriff) Satz 2.14(b) auch $bd \neq 0$ und wir können mit bd multiplizieren:
$ab^{-1}bd = cd^{-1}bd$ nach Anwendung von K6 und K8 folgt die Behauptung.

b. $\frac{a}{b} \overset{\text{K7}}{=} \frac{a}{b} \cdot 1 \overset{\text{K8}}{=} \frac{a}{b} \cdot ee^{-1} = ab^{-1} \cdot ee^{-1} \overset{\text{K6}}{=} (ae)(b^{-1}e^{-1}) \overset{\text{S.2.12(h)}}{=} (ae)(be)^{-1} = \frac{ae}{be}$

c. Folgt nach Multiplikation mit bd.

d. Folgt nach Anwendung von Satz 2.12(h).

□

Folgerungen aus den Distributivgesetzen

Es fehlt noch eine letzte Körpereigenschaft, bei der sowohl Addition als auch Multiplikation zum Einsatz kommen. Auch diese hier noch einmal zur Wiederholung:

(K9) $\quad \forall a, b, c \in K : \quad a \cdot (b + c) = a \cdot b + a \cdot c$
$\quad \forall a, b, c \in K : \quad (a + b) \cdot c = a \cdot c + b \cdot c$ (Distributivgesetze)

Daraus können wir folgern

Satz 2.14 (Folgerungen der Distributivgesetze)

In einem Körper K gilt für beliebige Elemente $a, b \in K$:

a. $0 \cdot a = a \cdot 0 = 0$
b. $ab = 0 \quad \Leftrightarrow \quad (a = 0) \vee (b = 0)$
c. $(-a)b = a(-b) = -(ab)$
d. $(-a)(-b) = ab$, insbesondere $(-1)(-1) = 1$

▼ Beweis

a. $a \cdot 0 \overset{\text{K3}}{=} a(0+0) \overset{\text{K9}}{=} a \cdot 0 + a \cdot 0$
also $a \cdot 0 = a \cdot 0 + a \cdot 0$
Wegen der Eindeutigkeit des neutralen Elements der Addition (Satz 2.10(a)) muss gelten
$a \cdot 0 = 0$.

b. Die Rückwärtsrichtung $\Leftarrow$ folgt direkt aus (a).
$\Rightarrow$-Richtung: wenn $b = 0$ gilt die Aussage, wenn $b \neq 0$ dann exisiert ein b^{-1} (K8)
und $ab = 0$ wird gelöst durch (Satz 2.12(f)) $a = 0 \cdot b^{-1} = 0$. Damit folgt die Behauptung.

c. Wir beweisen nur $a(-b) = -(ab)$, der andere Fall ist analog zu führen:
$ab + a(-b) \overset{\text{K9}}{=} a(b-b) \overset{\text{K4}}{=} a \cdot 0 \overset{(a)}{=} 0$.
Damit folgt insgesamt wegen der Eindeutigkeit des negativen Elements von ab, dass
$a(-b) = -(ab)$.

d. $(-a)(-b) \overset{(c)}{=} -(a(-b)) \overset{(c)}{=} -(-(ab)) \overset{\text{(S.2.10(c))}}{=} ab$.

□

Damit haben wir bereits einen Großteil der Regeln, die Sie vermutlich schon seit Jahren im Schlaf anwenden, aus den neun Eigenschaften eines Körpers abgeleitet. Wie bereits gesagt, gibt es auch andere Körper (wie $\mathbb{Q}$). Da wir nur die Körpereigenschaften benutzt haben, gelten die Regeln für beliebige Körper. Gehen wir nun einen Schritt weiter zu geordneten Körpern.

Folgerungen aus den Eigenschaften eines geordneten Körpers

Legen wir direkt los, indem wir die Eigenschaften eines geordneten Körpers noch einmal wiederholen:

(O1) $\forall\, a \in K$ gilt genau eine der drei Beziehungen $(a \in P), (a = 0), (-a \in P)$

(O2) $\forall\, a, b \in P :\ a + b \ \in P$

(O3) $\forall\, a, b \in P :\ ab \ \in P$

Darüber hatten wir die Vergleichsoperatoren definiert:

$$
\begin{aligned}
&a < b \Leftrightarrow (b - a) \in P \\
&a > b \Leftrightarrow b < a \\
&a \leq b \Leftrightarrow (a < b) \vee a = b \\
&a \geq b \Leftrightarrow b \leq a.
\end{aligned}
$$

Mit diesen Eigenschaften können wir die üblichen Regeln für das Arbeiten mit Ungleichungen herleiten. Ab hier werden wir die Körpereigenschaften nicht mehr jedes Mal explizit in den Beweisen mit angeben, aber prüfen Sie gerne selbst nach, warum die Umformungen jeweils gelten.

Satz 2.15 (Folgerungen für Ungleichungen)

In einem geordneten Körper K gilt für beliebige Elemente $a, b, c, d \in K$ und $x \in K \setminus \{0\}$:

a. Es gilt *genau eine* der drei Aussagen: $(a < b)$, $(a > b)$, $(a = b)$

b. $(a < b) \wedge (b < c) \quad \Rightarrow \quad a < c$

c. $(a < b) \wedge (c \leq d) \quad \Rightarrow \quad a + c < b + d$

d. $(a < b) \wedge (x > 0) \quad \Rightarrow \quad ax < bx$
$(a < b) \wedge (x < 0) \quad \Rightarrow \quad ax > bx$

e. $a < b \quad \Leftrightarrow \quad -a > -b$

f. $x^2 := x \cdot x > 0$, insbesondere $1^2 = 1 > 0$

g. $0 < a < b \quad \Leftrightarrow \quad 0 < b^{-1} < a^{-1}$

▼ Beweis

a. Folgt sofort aus O1, wenn man die Definition der Vergleichsoperatoren einsetzt.

b. $a < b \Leftrightarrow (b - a) \in P$ und $b < c \Leftrightarrow (c - b) \in P$.
Daraus folgt mit O2: $(b - a) + (c - b) \in P$.
Es gilt also $(b - a) + (c - b) = c - a \in P$.
Daraus folgt die Behauptung $a < c$.

c. Betrachten wir zunächst den Fall $c < d$: $a < b \Leftrightarrow (b - a) \in P$ und $c < d \Leftrightarrow (d - c) \in P$.
Daraus folgt mit O2: $(b - a) + (d - c) \in P$.
Es gilt weiter: $(b - a) + (d - c) = b + d - (a + c) \in P$.
Hieraus folgt die Behauptung: $a + c < b + d$.

Nun fehlt noch der Fall c = d: $a < b \Leftrightarrow (b - a) \in P$.
Nach K3 und K4 gilt:
$b - a = (b + 0 - a) = b + c - c - a = (b + c) - (c + a)$.
Hieraus folgt die Behauptung: $a + c < b + d$.

d. Betrachten wir zunächst den Fall $x > 0 \Leftrightarrow x \in P$:
mit $(b - a) \in P$ folgt aus O3 sofort
$(b - a)x = bx - ax \in P \Leftrightarrow ax < bx$.

Betrachten wir nun den Fall $x < 0 \Leftrightarrow -x \in P$:
mit $(b - a) \in P$ folgt aus O3 sofort
$(b - a)(-x) = -bx - (-ax) = ax - bx \in P \Leftrightarrow ax > bx$.

e. Folgt aus (c) mit $c = d = -a - b$.

f. Ist $x > 0$ so ist nach O3 $\quad x \cdot x = x^2 > 0$.
Ist $x < 0$ so ist $-x > 0$ und damit nach O3: $\quad (-x)(-x) \overset{\text{S.2.14(d)}}{=} x^2 > 0$.

g. Wir zeigen zunächst $a > 0 \Leftrightarrow a^{-1} > 0$: Ist $a > 0$, dann folgt mit $aa^{-1} = 1$ und Satz 2.14(b), dass auch $a^{-1} \neq 0$ gelten muss.
Wäre $a^{-1} < 0$ so würde aus (d) folgen $aa^{-1} > 0$.
Dies steht aber im Widerspruch zu $1 > 0$ (f) (↯). Also muss gelten $a^{-1} > 0$.
Die umgekehrte Richtung gilt analog, da $(a^{-1})^{-1} = a$.

Also gilt $(0 < a < b) \Rightarrow (a > 0 \wedge b > 0) \Leftrightarrow (a^{-1} > 0 \wedge b^{-1} > 0 \Rightarrow a^{-1}b^{-1} > 0)$.
Für $0 < b^{-1} < a^{-1}$ ist sofort klar, dass hier ebenfalls gilt $(a^{-1}b^{-1} > 0)$.

Also können wir die Ausgangsgleichung wie in (d) mit $x = a^{-1}b^{-1} > 0$ multiplizieren, woraus folgt:
$(0 < a < b) \Leftrightarrow (0 \cdot a^{-1}b^{-1} < aa^{-1}b^{-1} < ba^{-1}b^{-1}) \Leftrightarrow 0 < b^{-1} < a^{-1}$.

□

In vielen Fällen ist es sinnvoll, reelle Zahlen zu betrachten, die einen gewissen maximalen Abstand zu 0 besitzen. Zum Beispiel haben alle $x \in \mathbb{R}$ mit $-2 \leq x \leq 2$ einen maximalen Abstand von 2. Um so etwas kompakt angeben zu können, definieren wir uns den (Absolut-)Betrag einer reellen Zahl.

Definition 2.16 (Betrag)

Für $x \in \mathbb{R}$ heißt

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

der (Absolut-)Betrag von x.

Diese Betragsdefinition setzt einen geordneten Körper voraus. Wir werden später noch andere Betragsdefinitionen für ungeordnete Körper kennen lernen. Für Beträge lassen sich die folgenden Eigenschaften zeigen:

Satz 2.17 (Folgerungen für Beträge)

Für alle $x, y \in \mathbb{R}$ und $\varepsilon \in \mathbb{R}$ mit $\varepsilon > 0$ gilt

a. $|x| \geq 0 \ \wedge \ (|x| = 0 \ \Leftrightarrow \ x = 0)$

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

c. $(|x| < \varepsilon) \ \Leftrightarrow \ (x < \varepsilon) \ \wedge \ (-\varepsilon < x) \ \Leftrightarrow \ (-\varepsilon < x < \varepsilon)$
$(|x| \leq \varepsilon) \ \Leftrightarrow \ (x \leq \varepsilon) \ \wedge \ (-\varepsilon \leq x) \ \Leftrightarrow \ (-\varepsilon \leq x \leq \varepsilon)$

d. $|x + y| \leq |x| + |y|$ (Dreiecksungleichung)

e. $||x| - |y|| \leq |x - y|$ (umgekehrte Dreiecksungleichung)

▼ Beweis

a. Folgt unmittelbar aus der Definition

b. Folgt durch sorgsame Fallunterscheidung mit allen Kombinationen aus $x, y \geq 0$ und $x, y < 0$ (als Übung).

c. Die zweite Äquivalenz jeder Zeile ist jeweils nur eine andere Schreibweise. Zu zeigen ist die erste. Wir führen nur den $<$-Fall, da der $\leq$-Fall völlig analog geführt werden kann.

$\Rightarrow$-Richtung: Sei $|x| < \varepsilon$.
Für $x \geq 0$ folgt dann $-\varepsilon < 0 < x = |x| < \varepsilon$
Für $x < 0$ folgt $-x > 0$ und damit $-\varepsilon < 0 < -x = |x| < \varepsilon$.

$\Leftarrow$-Richtung: Sei $-\varepsilon < x < \varepsilon$.
Für $x \geq 0$ folgt dann $|x| = x < \varepsilon$.
Für $x < 0$ folgt $|x| = -x < \varepsilon$, da $x > -\varepsilon \Leftrightarrow -x < \varepsilon$.

d. Da $x \leq |x|$ und $y \leq |y|$ folgt aus Satz 2.15(c)

$$x + y \leq |x| + |y|$$

und da ebenfalls $-x \leq |x|$ und $-y \leq |y|$ gilt, folgt

$$-(x + y) = -x - y \leq |x| + |y| \, .$$

Damit gilt nach (c)

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

e. Beweis als Übung.

□

Über Beträge kann nicht nur der Abstand zur Null, sondern zu einer beliebigen Zahl angegeben werden. So haben beispielsweise alle $x \in \mathbb{R}$, für die gilt $|x-2| < 3$ einen kleineren Abstand als 3 zur Zahl 2. Dies definiert eine sogenannte Metrik:

Definition 2.18 (Metrik)

Sei A eine Menge. Wir nennen eine Abbildung $d : A \times A \to \mathbb{R}$ eine Metrik auf A, wenn für alle $x, y, z \in A$ die folgenden drei Eigenschaften erfüllt sind:

a. Positive Definitheit:

$$d(x,y) > 0 \text{ für } x \neq y$$
$$d(x,y) = 0 \text{ für } x = y$$

b. Symmetrie:

$$d(x,y) = d(y,x)$$

c. Dreiecksungleichung:

$$d(x,y) \leq d(x,z) + d(z,y)$$

Es lässt sich leicht zeigen, dass wir über den Betrag eine Metrik auf $\mathbb{R}$ definieren können:

Satz 2.19

Die Abbildung $d : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ mit

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

definiert eine Metrik auf $\mathbb{R}$.

▼ Beweis

a. Folgt direkt aus Satz 2.17(a).

b. Folgt aus Satz 2.17(b) mit

$$d(x,y) = |x-y| = |-1(y-x)| = |-1|\,|y-x| = |y-x| = d(y,x)$$

c. Folgt aus Satz 2.17(d) mit

$$d(x,y) = |x-y| = |x-z+z-y| \leq |x-z| + |z-y| = d(x,z) + d(z,y)$$

□

Auch wenn wir für ungeordnete Körper verschiedene Elemente nicht der Größe nach ordnen können, so lässt sich zeigen, dass man in einigen Fällen eine Metrik definieren kann. Damit können wir trotzdem Sätze anwenden, in denen wir Abstände vergleichen.

Folgerungen aus der Vollständigkeit

Bisher haben wir unser Vollständigkeitsaxiom noch nicht verwendet und daher gelten alle bisherigen Regeln für beliebige (geordnete) Körper (z.B. auch $\mathbb{Q}$). Die Folgerungen aus der Vollständigkeit gelten dagegen nur für vollständige Körper (wie $\mathbb{R}$).

Wir werden im Laufe der Vorlesung immer wieder auf die Vollständigkeit zurückkommen, wenn wir uns wichtige Hilfsmittel der Analysis (wie Folgen und Reihen) erarbeitet haben. Schauen wir uns zunächst eine "kleine" Folgerung an. Zur Wiederholung, noch einmal unsere Definition der Vollständigkeit:

- Wir nennen einen geordneten Körper $(K, +, \cdot)$ vollständig, wenn für jede nach oben beschränkte Teilmenge $M \subseteq K$ gilt: $\sup(M) \in K$.

Die Existenz eines Supremums für abgeschlossene Teilmengen reeller Zahlen ist damit schon durch das Axiom 3 gegeben. Es fehlt noch ein Beweis für die Eindeutigkeit von Suprema, sowie für die Existenz und Eindeutigkeit von Infima.

Satz 2.20 (Existenz und Eindeutigkeit von Suprema und Infima)

In einem vollständigen Körper $(K, +, \cdot)$

- existiert für jede nach oben beschränkte Teilmenge $\overline{M} \subseteq K$ ein eindeutiges Supremum $s = \sup(\overline{M}) \in K$.
- existiert für jede nach unten beschränkte Teilmenge $\underline{M} \subseteq K$ ein eindeutiges Infimum $i = \inf(\underline{M}) \in K$.

▼ Beweis

Das Supremum existiert nach Axiom 3.

Angenommen das Supremum s wäre nicht eindeutig, dann existiert ein weiteres $s' = \sup(\overline{M})$.
s' kann nicht größer als s sein, da sonst s' nicht die kleinste obere Schranke wäre.
s' kann nicht kleiner als s sein, da sonst s nicht die kleinste obere Schranke wäre.
Nach O1 folgt dann dass $s = s'$ ist. Das Supremum ist also eindeutig.

Für das Infimum betrachten wir die am Nullpunkt gespiegelte Menge $-\underline{M} = \{x \mid -x \in M\}$. Da $\underline{M}$ nach unten beschränkt ist, muss $-\underline{M}$ nach oben beschränkt sein. Daher muss nach Axiom 3 für $-\underline{M}$ ein Supremum existieren, das, wie wir oben gezeigt haben, eindeutig ist. Dieses ist per Konstruktion das Infimum von $\underline{M}$.

□

Da alle anderen Axiome auch für $\mathbb{Q}$ gelten, können wir aus der Vollständigkeit der reellen Zahlen nur Folgerungen ziehen, die nicht für $\mathbb{Q}$ gelten. Also z.B. die Existenz einer Zahl $\sqrt{a} \in \mathbb{R}$, für die $(\sqrt{a})^2 = a$ gilt. Natürlich gibt es einzelne rationale Zahlen, für die das ebenfalls gilt, wie $2 = \sqrt{4}$. Aber allgemein kann man die Existenz von Wurzeln nur für vollständige Körper zeigen.

Satz 2.21 (Existenz von Quadratwurzeln)

Für alle reellen Zahlen $a \geq 0$ existiert eine eindeutige nicht-negative reelle Zahl x, für die gilt $x^2 := x \cdot x = a$. Wir nennen diese Zahl Quadratwurzel von a, oder $\sqrt{a}$.

▼ Beweis

Wir haben die 2 bisher nicht definiert. Darauf gehen wir später noch ein, aber für diesen Beweis definieren wir $2 := 1 + 1$. Dies wird unser erster etwas umfangreicherer Beweis.

Fall 1: $a = 0$
Nach Satz 2.14(b) kann ein Produkt $x \cdot x$ nur dann 0 sein, wenn mindestens einer der beiden Faktoren 0 ist, daraus folgt $\sqrt{0} = 0$.

Fall 2: $a > 0$

Existenz:

Wir starten zunächst mit dem Beweis der Existenz einer Lösung. Dazu betrachten wir die Menge M aller positiven reellen Zahlen, für die $y^2 \leq a$ gilt, also:

$$M = \{y \in \mathbb{R} \mid y \geq 0 \wedge y^2 \leq a\}$$

Da $0^2 = 0 < a$, muss 0 eine Element von M sein, also ist M nicht leer.

Außerdem gilt für $a + 1$:

$$(a+1)^2 = a^2 + 2a + 1 > a^2$$

Also ist $a + 1$ eine obere Schranke von M und M ist nach oben beschränkt. Damit sind die Bedingungen für unser Vollständigkeitsaxiom für M erfüllt, und es existiert ein $x = \sup(M) \in \mathbb{R}$. Es fehlt also noch zu zeigen, dass $x^2 = a$ gilt. Das beweisen wir per Widerspruch. Wir nehmen also an $x^2 \neq a$, das bedeutet, dass entweder $x^2 > a$ oder $x^2 < a$ gelten muss (Satz 2.15).

Widerspruch Fall 2.1: ($x^2 > a$)

Wir betrachten $s = x - \frac{x^2 - a}{2x}$.

Es gilt $x^2 > a \Leftrightarrow x^2 - a > 0 \Leftrightarrow \frac{x^2-a}{2x} > 0$
Gleichzeitig gilt auch: $x^2 > -a \Leftrightarrow 2x^2 > x^2 - a \Leftrightarrow x > \frac{x^2-a}{2x}$

Daraus folgt $x > x - \frac{x^2-a}{2x} = s > 0$.

Betrachten wir s^2, dann folgt:

$$
\begin{aligned}
s^2 &= \left(x - \frac{x^2 - a}{2x}\right)^2 \\
&= x^2 - 2x\frac{x^2 - a}{2x} + \underbrace{\left(\frac{x^2-a}{2x}\right)^2}_{>0} \\
&> x^2 - 2x\frac{x^2 - a}{2x} \\
&= x^2 - x^2 + a \\
&= a
\end{aligned}
$$

Also $s^2 > a$. Damit gilt für jedes $y \in M$: $y^2 \leq a < s^2$. Daraus folgt auch $y < s$ (Beweisen Sie zur Übung diese Folgerung per Kontraposition).

Also ist s eine obere Schranke und $s < x$ steht im Widerspruch dazu, dass x die kleinste obere Schranke ist (↯). Somit muss unsere Annahme ($x^2 > a$) falsch sein.

Widerspruch Fall 2.2: ($x^2 < a$)

Hier betrachten wir zunächst $z = \min\left\{1, \frac{a-x^2}{2x+1}\right\}$

Es lässt sich wieder leicht zeigen, dass $\frac{a-x^2}{2x+1} > 0$ gilt.
Dadurch, dass wir das Minimum mit 1 bilden, gilt insgesamt $0 < z \leq 1$, woraus durch Multiplikation mit z folgt $0 < z^2 \leq z$.

Dann gilt:

$$
\begin{aligned}
(x+z)^2 &= x^2 + 2xz + z^2 \\
&\leq x^2 + 2xz + z \\
&= x^2 + (2x+1)z \\
&\leq x^2 + (2x+1)\frac{a - x^2}{2x+1} \\
&= a
\end{aligned}
$$

Also insgesamt $(x+z)^2 \leq a$. Nach der Definition von M muss also gelten $x + z \in M$. Dies steht aber im Wiederspruch dazu, dass x die kleinste obere Schranke ist, da $x < x + z$. Somit muss unsere Annahme ($x^2 < a$) falsch sein.

Aus den beiden Wiedersprüchen 2.1 und 2.2 folgt, dass für unser $x = \sup(M)$ gelten muss: $x^2 = a$.

Eindeutigkeit

Noch zu zeigen ist, dass es kein zweites positives x' gibt, für das ebenfalls gilt $x'^2 = a$.

Dies ist einfach: Nehmen wir an, es gäbe so ein x', dann folgt:

$$
(x - x')(x + x') = x^2 - x'^2 = a - a = 0
$$

Ein Produkt aus zwei zwei Faktoren kann aber nur dann 0 sein, wenn einer der Faktoren 0 ist (Satz 2.14(b)).

$x + x' = 0$ kann nicht gelten, da daraus folgt $x' = -x < 0$, was $x' > 0$ widerspricht (↯).

$x - x' = 0$ führt zu $x = x'$.

Es gibt also genau eine positive reelle Lösung zu $x^2 = a$.

□

Ganz analog lässt sich auch die Existenz von k-ten Wurzeln mit $k \in \mathbb{N}$, also $\sqrt[k]{a}$ für $a > 0$, beweisen. Hierfür werden wir aber später noch einen anderen Beweis sehen.

2.4 Definition der natürlichen, ganzen und rationalen Zahlen

Ausgestattet mit unserer Definition der reellen Zahlen können wir nun die natürlichen Zahlen und darauf aufbauend die ganzen und rationalen Zahlen jeweils als Teilmengen der reellen Zahlen definieren. Da wir Regeln wie $x^2 > 0$ aus dem letzten Abschnitt für *allgemeine* reelle Zahlen bewiesen haben, gelten diese auch für *bestimmte* reelle Zahlen, wie zum Beispiel natürliche Zahlen. Wir werden also nun die einzelnen Zahlenmengen über die reellen Zahlen definieren.

Natürliche Zahlen

Wie Ihnen vielleicht aufgefallen ist, haben wir bisher nur zwei Zahlen in den reellen Zahlen wirklich klar festgelegt, nämlich 0 und 1 (und evtl. noch -1 durch die Existenz des Inversen). Wissen wir also überhaupt, dass nach unseren bisherigen Axiomen $2 \in \mathbb{R}$ sein muss, oder $1 + 1 = 2$ gilt? Dies kann zumindest nicht aus den Körpereigenschaften folgen, da wir bereits in Beispiel 2.2 einen Körper gesehen haben, der nur aus $\{0, 1\}$ besteht und für den $1 + 1 = 0$ gilt. Genau wie die Symbole 0 und 1, die wir für das Null- und Einselement gewählt haben, ist 2 lediglich ein Symbol des Dezimalsystems. Und dieses Symbol ist auch nicht überall gleich. Ein Mensch aus Japan oder China würde vielleicht eher 二 statt 1 verwenden. Die 2 ist ein gewähltes Zeichen, wenn wir (im Dezimalsystem) nicht umständlich $1 + 1$ sagen wollen. Im Japanischen/Chinesischen sieht man das dem gewählten Zeichen (二) sogar an, genauso wie in vielen anderen Schriftsystemen (z.B. die römische Zahl II). Als Informatiker fällt uns natürlich sofort eine weitere Art ein, $1 + 1$ darzustellen, nämlich im Binärsystem als 10. Wir werden uns später noch mit verschiedenen Stellenwertsystemen beschäftigen, aber bis hierhin sollte Ihnen klar geworden sein, dass die Zahlzeichen $(2, 3, 4, \ldots)$ natürlich nicht aus unseren Axiomen folgen.

Die natürlichen Zahlen definieren wir in zwei Schritten. Wir beginnen mit der induktiven Menge

Definition 2.22 (Induktive Menge)

Wir nennen eine Menge $M \subseteq \mathbb{R}$ induktive Menge, wenn gilt:

1. $1 \in M$
2. $\forall x \in M : x + 1 \in M$

Auf den ersten Blick sieht es so aus, als wären wir hier schon fertig mit der Definition von $\mathbb{N}$, aber nach kurzer Überlegung fällt auf, dass auch andere Mengen existieren, wie $M = \{1, 1.5, 2, 2.5, 3, 3.5, ...\}$, auf die die Definition zutrifft. Sogar $\mathbb{R}$ selbst ist eine induktive Menge. Allerdings enthalten alle diese Mengen die natürlichen Zahlen, also $\mathbb{N} \subseteq M$. $\mathbb{N}$ ist also die 'kleinste' induktive Menge. Nun ist es aber nicht so einfach definierbar, was 'kleinste' hier bedeutet, da $\mathbb{N}$ bekanntlich aus unendlich vielen Elementen besteht. Daher benutzen wir stattdessen den $\cap$-Operator aus Definition 1.12:

Definition 2.23 (Natürliche Zahlen)

Sei $\mathcal{N}$ die Menge aller induktiven Mengen. Wir definieren die Menge der natürlichen Zahlen $\mathbb{N}$ als:

$$\mathbb{N} = \bigcap_{M \in \mathcal{N}} M$$

Darüber hinaus definieren wir

$$\mathbb{N}_0 = \mathbb{N} \cup \{0\}$$

Anders gesagt besteht $\mathbb{N}$ aus allen Elementen, die in jeder induktiven Menge enthalten sind. Wenn wir uns nun noch, wie oben bereits beschrieben, unsere typischen Zahlsymbole definieren ($2 := 1 + 1, 3 := 2 + 1, 4 := 3 + 1, ...$) wird klar, dass demnach $\mathbb{N} = \{1, 2, 3, 4, ...\}$ gilt. Es lässt sich ebenfalls sofort zeigen, dass $\mathbb{N}$ selbst eine induktive Menge ist. Wie bereits gesagt, gelten die meisten Regeln, die für Elemente von $\mathbb{R}$ gelten, auch für natürliche Zahlen, da diese per Definition auch reell sind.

Satz 2.24 (Eigenschaften der natürlichen Zahlen)

a. Für jede Zahl $n \in \mathbb{N}$ gilt $n \geq 1$.

b. Summen und Produkte natürlicher Zahlen sind wieder natürliche Zahlen.

c. Für $n, m \in \mathbb{N}$ ist die Differenz $n - m$ genau dann eine natürliche Zahl, wenn $n > m$.

d. Für alle $n \in \mathbb{N}$ existiert kein $m \in \mathbb{N}$ für das gilt $n < m < n + 1$.

▼ Beweise per vollständiger Induktion als Übung

Falls Sie im Kapitel der Beweistechniken unseren Exkurs in die geraden und ungeraden Zahlen gelesen haben, wird Ihnen hier auffallen, dass diese Eigenschaften, besonders (b), dort wichtig waren.

Zuletzt beweisen wir noch einen nützlichen kleinen Satz.

Satz 2.25 (Archimedische Eigenschaft von $\mathbb{R}$)

$$\forall x \in \mathbb{R}\, \exists n \in \mathbb{N} : x < n,$$

oder in anderen Worten: $\mathbb{N}$ besitzt keine obere Schranke.

▼ Beweis

Wäre $\mathbb{N}$ nach oben beschränkt, folgt nach dem Vollständigkeitsaxiom, dass für $\mathbb{N}$ ein Supremum $s \in \mathbb{R}$ existiert. Es muss daher ein $n \in \mathbb{N}$ existieren, mit $n > s - 1$, da sonst eine kleinere obere Schranke $s' = s - 1$ existieren würde.

Aus $n > s - 1$ folgt $n + 1 > s$. Wenn $n \in \mathbb{N}$, muss nach der Definition der induktiven Menge auch $n + 1 \in \mathbb{N}$ sein. Demnach kann s aber nicht das Supremum sein, was einen Widerspruch darstellt (↯). Somit muss $\mathbb{N}$ nach oben unbeschränkt sein.

□

Die natürliche Zahlen bilden die Basis für die folgenden Zahlenmengen.

Ganze Zahlen

Definition 2.26 (Ganze Zahlen)

Wir definieren die Menge der ganzen Zahlen $\mathbb{Z}$ als

$$\mathbb{Z} = \{0\} \cup \{n \mid n \in \mathbb{N} \vee -n \in \mathbb{N}\}.$$

Also anders gesagt: $\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, ...\}$. Man kann zeigen, dass die ganzen Zahlen die ersten vier Körpereigenschaften erfüllen und damit mit der Addition eine Gruppe bilden. Die Eigenschaften (b) und (d) von Satz 2.24 gelten auch wenn man im Satz $\mathbb{N}$ durch $\mathbb{Z}$ ersetzt. Außerdem gilt folgender nützlicher Satz:

Satz 2.27 (Jede reelle Zahl liegt zwischen zwei ganzen Zahlen)

Für jedes $x \in \mathbb{R}$ gibt es zwei eindeutige Zahlen $m, n \in \mathbb{Z}$, sodass gilt:

a. $m \leq x < m + 1$
b. $n - 1 < x \leq n$

Somit können wir jede reelle Zahl schreiben als $x = m + \rho$, wobei $\rho = x - m$.

Wir definieren darüber:

$\lfloor x \rfloor := m$ (untere Gaußklammer)
$\lceil x \rceil := n$ (obere Gaußklammer)

▼ Beweis

Hier nur exemplarisch für (a), (b) beweist man analog. Zusätzlich beschränken wir uns auf $x > 0$ für $x < 0$ wendet man die selbe Strategie auf $-x$ an:

Nach der archimedischen Eigenschaft von $\mathbb{R}$ gibt es mindestens ein $n \in \mathbb{N} : n > x$. Wir wählen aus der Menge dieser größeren natürlichen Zahlen das Minimum:

$$m' = \min\{n \in \mathbb{N} : n > x\}$$

dann ist $m = m' - 1$ und es gilt $m < x < m + 1$. Wir müssen also nur noch zeigen, dass kein zweites $k \in \mathbb{Z} \setminus \{m\}$ mit $k \leq x < k + 1$ existieren kann:

Würde es existieren, müsste $k < m$ oder $k > m$ gelten, was für ganze Zahlen identisch ist mit $k + 1 \leq m$ oder $k - 1 \geq m$. Sei also $k + 1 \leq m$, dann folgt: $k + 1 \leq m \leq x < k + 1$, also der Widerspruch $k + 1 < k + 1$ (↯). Der zweite Fall führt analog zum Widerspruch.

Damit muss m eindeutig sein.

□

Beispiel 2.28 (Gaußklammern)

$$\begin{aligned}\lfloor 2.1 \rfloor &= 2\\ \lceil 2.1 \rceil &= 3\\ \lfloor -2.1 \rfloor &= -3\\ \lceil -2.1 \rceil &= -2\\ \lfloor 2 \rfloor &= 2\\ \lceil 2 \rceil &= 2\end{aligned}$$

Das letzte Beispiel zeigt, warum wir m und n brauchen, denn wenn wir $\lceil x \rceil = m + 1$ definiert hätten, wäre $\lceil 2 \rceil = 3$. Wenn x keine ganze Zahl ist, gilt aber $m + 1 = n$.

Mit dem Satz 2.27 können wir den Modulo-Operator einführen, die Basis für die Arbeit mit sogenannten Restklassen, die Ihnen noch häufiger im Studium begegnen werden:

Satz 2.29 (Division mit Rest, der Modulo Operator)

Für alle $z \in \mathbb{Z}$ und $n \in \mathbb{N}$ gibt es eindeutige $m, r \in \mathbb{Z}$, $r < n$ mit

$$z = mn + r$$

Man schreibt auch

$$r = z \,(\text{modulo}\, n)$$

als Rest der (ganzzahligen) Division von z mit n.

▼ Beweis

Der Beweis folgt aus Satz 2.27 mit $x = \frac{z}{n}$. Demnach gibt es ein eindeutiges m mit $m \leq x < m + 1$.

Damit ergibt sich

$$x = \frac{z}{n} = m + (\frac{z}{n} - m) \quad \Leftrightarrow \quad z = mn + r$$

mit $r = z - mn$.

□

Rationale Zahlen

Definition 2.30 (Rationale Zahlen)

Wir definieren die Menge der rationalen Zahlen als

$$\mathbb{Q} := \left\{ \frac{p}{q} \;\middle|\; p, q \in \mathbb{Z} \wedge q \neq 0 \right\}$$

Wir bezeichnen p als Zähler und q als Nenner.

Wir nennen alle $x \in \mathbb{R} \setminus \mathbb{Q}$ irrational.

Wir wollen hier auch schon zwei Beispiele für Zahlen geben, die irrational sind. Genau genommen zeigen wir hier nur, dass sie nicht rational sind. Es fehlt noch der Beweis, dass es sie wirklich in $\mathbb{R}$ gibt. Für Quadratwurzeln wurde dies bereits in Satz 2.21 gezeigt.

Beispiel 2.31 (Irrationale Zahlen)

Als Erstes zeigen wir das typische Lehrbuchbeispiel für eine irrationale Zahl: $\sqrt{2}$. Im Anschluss führen wir einen geometrischen Beweis für die Irrationalität des sogeannten goldenen Schnitts.

▼ Beweis Wurzel aus 2 ist nicht rational

Der Beweis wird klassisch über einen Widerspruchsbeweis geführt. Wir nehmen also an, es gäbe eine Lösung von $x^2 = 2$ mit $x \in \mathbb{Q}$, also $x = p/q$ für ein $p \in \mathbb{Z}$ und $q \in \mathbb{N}$.

Wir nehmen außerdem an, dass p und q teilerfremd sind (also x maximal gekürzt ist). Das ist immer möglich (überlegen Sie, warum?).

Daraus folgt

$$2 = \left(\frac{p}{q}\right)^2 \Leftrightarrow p^2 = 2q^2.$$

Aus unserem Exkurs der Beweistechniken wissen wir, dass $2q^2$ eine gerade Zahl ist. Also muss p^2 ebenfalls gerade sein. Ebenfalls aus dem Exkurs (Satz E1.3) wissen wir, dass dann p ebenfalls gerade ist.

Wenn p aber gerade ist, gibt es ein $k \in \mathbb{Z}$ mit $p = 2k$. Daraus folgt

$$p^2 = (2k)^2 = 4k^2 = 2q^2 \Leftrightarrow 2k^2 = q^2.$$

Was bedeutet, dass q^2 und somit auch q gerade sind. Wenn aber sowohl p als auch q gerade sind, steht dass im Widerspruch zur Voraussetzung, dass p und q teilerfremd sind (↯). Daher kann keine rationale Zahl exisitieren, die $x^2 = 2$ löst, oder anders gesagt, wenn es eine reelle Lösung gibt, dann ist sie irrational.

□

▼ Beweis goldener Schnitt ist nicht rational

Diesen Beweis führen wir hier einmal ganz anders, ähnlich wie ihn die alten Griechen geführt hätten. Damals wurden alle Beweise geometrisch geführt, also über Längen von Strecken, Winkel, Flächen, usw. Dadurch sind diese Beweise manchmal oft sehr anschaulich.

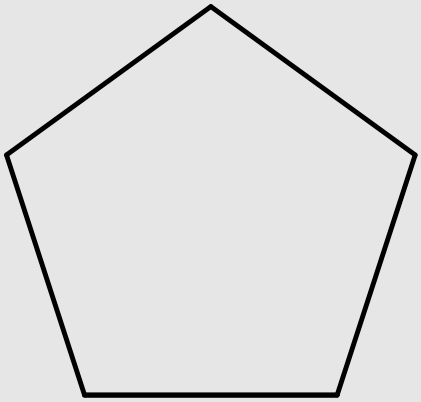
Der goldene Schnitt ist das Verhältnis (der Bruch) q aus zwei Zahlen a, b für die gilt:

$$q = \frac{a}{b} = \frac{a+b}{a}$$

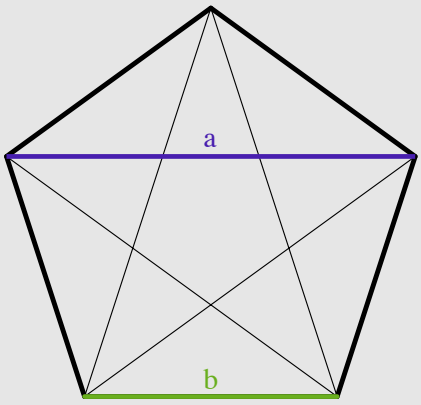
Durch ein wenig Umformen lässt sich diese Gleichung lösen:

$$
\begin{aligned}
q = \frac{a+b}{a} &= 1 + \frac{1}{q} \\
\Leftrightarrow 0 &= q^2 - q - 1 \\
\Leftrightarrow q &= \frac{1+\sqrt{5}}{2}
\end{aligned}
$$

Wir könnten analog zum Beiweis von $\sqrt{2}$ zeigen, dass $\sqrt{5}$ irrational ist, aber das wäre nicht die Art und Weise, wie die alten Griechen den Beweis geführt hätten. Dort hätte man sich den goldenen Schnitt beispielsweise durch ein regelmäßiges Fünfeck visualisiert:



Verbindet man hier die Eckpunkte miteinander, ergibt sich in der Mitte der Figur erneut ein regelmäßiges Fünfeck. Wenn Sie ihre Schulgeometriekenntnisse nutzen, können Sie mit dem Strahlensatz zeigen, dass das Verhältnis den beiden farbigen Seitenlängen genau dem goldenen Schnitt entspricht.



Die Frage nach der Rationalität würden die alten Griechen folgendermaßen stellen: Gibt es einen Stab einer bestimmten Länge x, der genau genau n-mal in die große Fünfeckseite passt und genau m-mal in die kleine (man sagt auch ‘mit dem Stab kann die Länge der Seiten *abgetragen* werden’)? Dabei müssen n und m natürliche Zahlen sein müssen und für das Verhältnis ergäbe sich eine rationale Zahl:

$$
q = \frac{a}{b} = \frac{nx}{mx} = \frac{n}{m} \in \mathbb{Q}
$$

Doch wie prüft man das geometrisch? Die Idee war damals folgende: Wenn man zwei natürliche Zahlen, z.B. $6, 2$ betrachtet, dann sind diese entweder gleich groß oder unterschiedlich groß. Im zweiten Fall können wir die Differenz aus der größeren und der kleineren Zahl bilden:

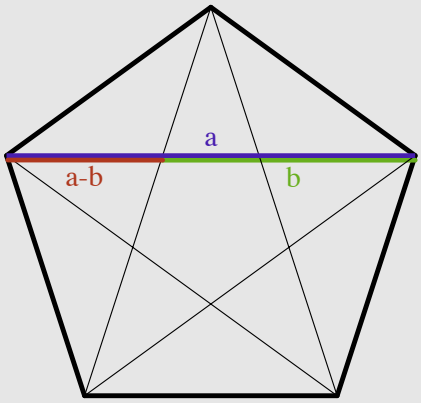
$$6 - 2 = 4$$

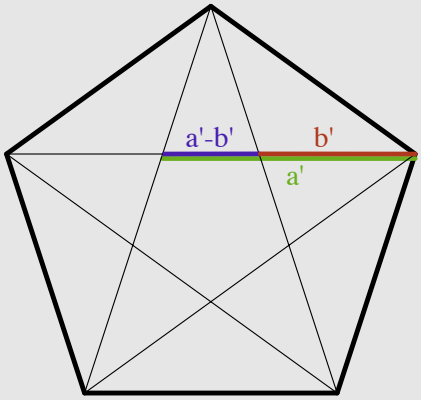
und dies ergibt wieder eine natürlich Zahl. Von der Differenz und den beiden ursprünglichen Zahlen wählen wir die beiden kleinsten aus $(2, 4)$ und bilden erneut die Differenz aus der größeren und der kleineren:

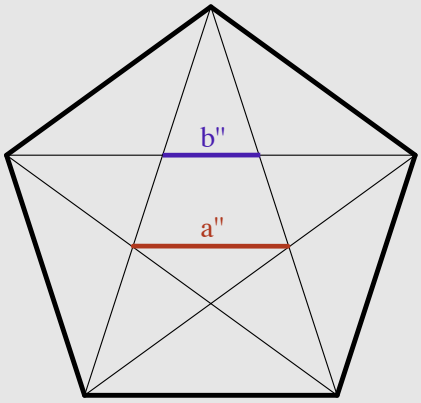
$$4 - 2 = 2$$

Dieser Algorithmus führt auf immer kleinere Zahlenpaare und muss früher oder später auf zwei gleiche natürliche Zahlen führen (spätestens bei $(1, 1)$). Die alten Griechen würden also nach und nach die kleinere Länge von der Größeren abziehen, bis sich zwei gleiche Längen ergäben. Der Algorithmus nennt sich übrigens Euklidischer Algorithmus und führt auf den größten gemeinsamen Teiler der beiden Zahlen.

Führen wir dies für das Fünfeck aus:







Nach zwei Schritten landen wir wieder in der Ausgangssituation, nur mit einem kleineren Fünfeck. Man kann sich leicht klar machen, dass auch die weiteren Schritte immer wieder auf kleinere Fünfecke führen würden und man somit niemals auf zwei gleiche Längen kommt. Dies zeigt, dass wir keinen Stab finden können, mit dem wir beide Längen abtragen können, oder mit unseren heutigen Worten: das Verhältnis der beiden Zahlen kann niemals aus einem Bruch natürlicher Zahlen bestehen. Der goldene Schnitt ist also nicht rational.

□

2.5 Abzählbarkeit von Zahlenmengen

In diesem Kapitel werden wir uns mit der Mächtigkeit von Zahlenmengen beschäftigen. Wir werden den Begriff der Mächtigkeit im Verlauf der Vorlesung nicht wieder benötigen, aber trotzdem gibt gibt es uns ein paar nützliche Einsichten über die verschiedenen Zahlenmengen. Daher werden wir hier ausnahmsweise keine strikten Beweise führen und eher auf sehr anschauliche Begründungen zurückgreifen.

Die Mächtigkeit einer Menge M schreibt man als $|M|$ und für abzählbare Mengen entspricht diese der Anzahl der Elemente. Für Mengen mit unendlich vielen Elementen vergleicht man die Mächtigkeit der Menge mit der Mächtigkeit der natürlichen Zahlen $\mathbb{N}$. Allerdings ist auch hier auf den ersten Blick noch nicht klar, wie dieser Vergleich funktioniert, da es unendlich viele Elemente in $\mathbb{N}$ gibt. Die Lösung ist eine Zuordnung: Wir bestimmen also für jedes Element m aus der zu vergleichenden Menge M eine zugeordnete natürliche Zahle $n \in \mathbb{N}$. Können wir eine Zuordnung so angeben, bei der jedem $m \in M$ ein *unterschiedliches* $n \in \mathbb{N}$ zugeordnet wird, dann nennen wir die Mächtigkeit der Menge M *gleichmächtig* zu $\mathbb{N}$.

Wir nennen außerdem Mengen, die gleichmächtig zu $\mathbb{N}$ sind *abzählbar unendlich*, weil die Menge unendlich viele Elemente enthält, aber wir diese nacheinander (mit unendlich viel Zeit) aufzählen könnten: dies ist das 1. Element, dies ist das 2., dies ist das 3. usw.. Manchmal hört man auch die Aussage, dass so eine abzählbar unendliche Menge “gleich viele” Elemente wie $\mathbb{N}$ besitzt. Allerdings versagt die Vorstellung von “gleich viel” bei unendlichen Anzahlen, wodurch diese Aussage immer sehr

unintuitiv wirkt, wie wir noch sehen werden. Mengen, für die man zeigen kann, dass keine solche Zuordnung existiert, nennt man *überabzählbar unendlich*.

Betrachten wir zwei Beispiele:

1. Die Menge der geraden Zahlen

$$G = \{2k \mid \ k \in \mathbb{N}\}$$

   Es ist schnell klar, dass es unendlich viele gerade Zahlen gibt. Vom Gefühl her würden wir denken, es gibt halb so viele gerade Zahlen wie natürliche Zahlen, weil jede zweite natürliche Zahl gerade ist. Mit dem Begriff der Mächtigkeit ist G allerdings gleich mächtig zu $\mathbb{N}$ ist (abzählbar unendlich), da wir jedem $k \in \mathbb{N}$ genau ein $g \in G$ per $g = 2k$ zuordnen können. Hier wird klar, dass man gleiche Mächtigkeit nicht mit "gleich viel" übersetzen sollte.

2. Die Menge der Ganzen Zahlen

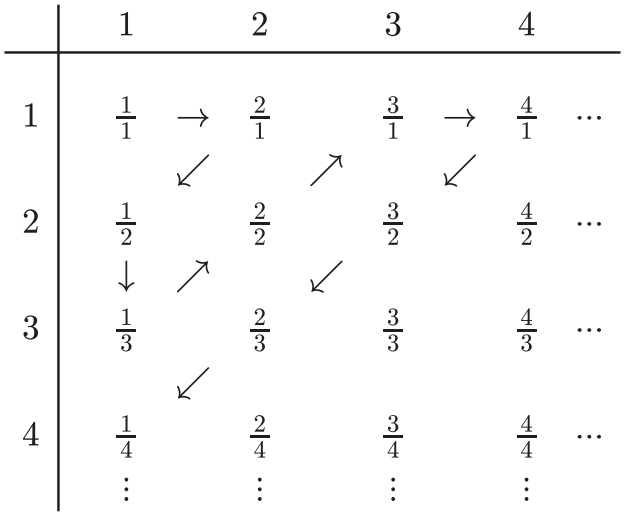
$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$$

   Auch hier ist klar, dass es unendlich viele ganze Zahlen gibt. Vom Gefühl her würden wir denken, es gibt mindestens doppelt so viele gerade Zahlen wie natürliche Zahlen, da es für jedes $n \in \mathbb{N}$ sowohl n als auch $-n$ in $\mathbb{Z}$ gibt (und zusätzlich noch das Nullelement). Allerdings können wir wie folgt zuordnen: $(1 \to 0),\ (2 \to 1),\ (3 \to -1),\ (4 \to 2),\ (5 \to -2),\ (6 \to 3),\ (7 \to -3),\ (8 \to 4)$, u.s.w. Also jeder gerade natürlichen Zahl nacheinander die positiven ganzen Zahlen und jeder ungeraden natürlichen Zahl > 1 die negativen (und $1 \to 0$). Damit gibt es also eine Zuordnung von $\mathbb{N}$ zu $\mathbb{Z}$, die alle ganzen Zahlen abdeckt und somit sind $\mathbb{Z}$ und $\mathbb{N}$ gleichmächtig. Auch hier klingt die Aussage $\mathbb{Z}$ ist abzählbar unendlich akzeptierbarer als "es gibt gleich viele ganze und natürliche Zahlen".

Interessant wird es, wenn wir die "nächstgrößere" Zahlenmenge betrachten: die rationalen Zahlen $\mathbb{Q}$. Also alle Zahlen, die sich als Bruch aus ganzen Zahlen schreiben lassen. Hier gilt sogar, dass es unendlich viele rationale Zahlen alleine zwischen 0 und 1 gibt, z.B.

$$\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \frac{1}{8}, \frac{1}{9}, \dots$$

Aber auch hier lässt sich zeigen, dass eine Zuordnung aus den natürlichen Zahlen in die rationalen Zahlen existiert, die alle rationalen Zahlen abdeckt. Dafür ordnet man die rationalen Zahlen in einem Gitter an: die Spalte bestimmt den Zähler und die Reihe den Nenner:



Die Pfeilrichtung gibt dabei die Reihenfolge der Nummerierung an. Es ist leicht erkennbar, dass man dies bis zu jeder (positiven) rationalen Zahl fortführen könnte. Anschließend verwendet man wieder den gleichen Trick, wie bei den ganzen Zahlen: die 0 ordnet man der 1 zu und anschließend nimmt man abwechselnd immer den positiven und dann den negativen Eintrag des Tabellenwerts in Pfeilrichtung. Somit kann man also auch für $\mathbb{Q}$ die Abzählbarkeit zeigen. Diesen graphischen Beweis nennt man auch "Cantors erstes Diagonalargument" nach Gregor Cantor, einem Mathematiker aus dem 19. Jahrhundert, der sich viel mit den Zahlenmengen und Unendlichkeiten beschäftigt hat.

Auf ihn geht auch ein sehr anschaulicher Beweis zurück, der zeigt, warum die reellen Zahlen $\mathbb{R}$ nicht abzählbar (überabzählbar) sind. Dieser ist "Cantors zweites Diagonalargument". Um den Beweis zu verstehen, benötigen wir die Dezimaldarstellung reeller Zahlen. Diese können wir erst in Kapitel 3 einführen, da wir hierzu ein Grundverständnis von unendlichen Summen benötigen. Aber Sie kennen natürlich bereits aus der Schule die Dezimalschreibweise reeller Zahlen:

- $7/5 = 1.4$
- $1/1000 = 0.001$
- $1/3 = 0.33333333 ... = 0.\overline{3}$
- $1/7 = 0.142857142857142857142857142857 ... = 0.\overline{142857}$
- $\pi = 3.14159265358979323846264338327950288419716939937510582097494 4 ...$

Dabei gibt es (wie wir später noch beweisen werden) endliche Dezimalzahlen, bei denen ab einer gewissen Dezimalstelle nur noch Nullen folgen (wie $7/5$), periodisch unendliche Dezimalzahlen, bei denen sich eine bestimmte Zahlenfolge unendlich oft wiederholt (wie $1/7$) und nicht-periodische unendliche Dezimalzahlen, bei denen keines der beiden ersten Kriterien zutrifft. Auf rationale Zahlen trifft immer der erste oder der zweite Fall zu. Alle irrationale Zahlen, also reelle Zahlen, die nicht rational sind, fallen in die dritte Kategorie.

Jetzt könnten wir uns zunächst fragen, wie "häufig" irrationale Zahlen unter den reellen Zahlen sind. Also etwas unmathematisch formuliert: wenn ich zufällig eine Zahl unter allen reellen auswählen würde, wie wahrscheinlich ist es, dass diese irrational ist? Da man in der Regel nur wenige Beispiele für irrationale Zahlen in der Schule (und selbst im Studium) kennenlernt und wir wissen, dass unendlich viele rationale Zahlen auf jeden noch so kleinen Teilabschnitt der Zahlengerade passen, liegt die Vermutung nahe, dass die irrationalen Zahlen eher die Ausnahme unter den reellen Zahlen sind.

Betrachten wir allerdings die unterschiedlichen Fälle reeller Dezimalzahlen wird schnell klar, dass rationale Zahlen eher die Ausnahme sind: Rationale Zahlen benötigen eine ganz bestimmte Struktur in der Dezimaldarstellung, wohingegen die irrationalen komplett chaotisch sein dürfen. Stellt man sich vor, man würfelt die einzelnen Dezimalstellen jeweils mit einem Zehnerwürfel (0–9), so wird offensichtlich, dass es viel wahrscheinlicher ist, dabei keine sich unendlich oft wiederholende Ziffernfolge zu würfeln, sondern eine eher chaotische Folge.

Es ist gar nicht so einfach zu zeigen, dass eine Zahl nicht rational ist. Beipielsweise existieren Beweise dafür, dass

- $\sqrt{\frac{m+1}{m}},\ \forall m \in \mathbb{N}$
- die eulersche Zahl e
- die Kreizahl π
- e^{π}
- $\sqrt{2}^{\sqrt{2}}$

irrational sind. Aber man konnte bisher nicht beweisen, dass auch

- $\pi + e$
- $\pi - e$
- πe
- π / e

irrational sind, auch wenn dies vermutet wird. Es ist bisher sogar unklar, ob $(\pi^{\pi})^{\pi}$ eine natürliche Zahl ist oder nicht, auch wenn man vermuten würde, dass das sehr offensichtlich nicht sein kann.

Wir haben also ein erstes Gefühl dafür bekommen, dass es deutlich mehr irrationale als rationale Zahlen gibt. Kommen wir also nun zur Mächtigkeit der reellen Zahlen $\mathbb{R}$ und Cantors zweitem Diagonalargument. Hierbei wird die Überabzählbarkeit der reellen Zahlen zwischen 0 und 1 anschaulich bewiesen. Die Argumentation erfolgt als ein Widerspruchsbeweis: Wir nehmen also an, es gäbe eine Zuordnung von $\mathbb{N}$ zu allen reellen Zahlen zwischen 0 und 1. Dann könnten wir deren Dezimaldarstellungen der Reihe nach hinschreiben. In dieser unendlich langen Aufzählung kämen alle reellen Zahlen zwischen 0 und 1 vor. Der Widerspruch kommt dadurch zustande, dass Cantor zeigt, wie man aus dieser Aufzählung eine neue reelle Zahl z zwischen 0 und 1 konstruieren kann, die nicht in der Aufzählung vorkommt. Dabei geht er wie folgt vor:

1. Betrachte die erste Dezimalstelle der ersten Zahl der Aufzählung
   - Ist die Ziffer eine 9, setze die erste Dezimalstelle unserer neuen Zahl z auf 8.
   - Ansonsten setze die erste Dezimalstelle von z auf diese Ziffer **plus 1**.
2. Wiederhole das Vorgehen für die zweite Dezimalstelle der zweiten Zahl in der Aufzählung und setze damit die zweite Dezimalstelle von z.
3. Wiederhole das Vorgehen für die dritte Dezimalstelle der dritten Zahl in der Aufzählung und setze damit die dritte Dezimalstelle von z.
4. ...

Beispiel 2.32 (Beispielkonstruktion für Cantors 2. Diagonalargument)

$$
\begin{aligned}
x_1 &= 0.\,3\,2\,9\,4\,2\,5\,7\,5\,6\,6\,9\,1\,7\,4\,0\,1\,0\,4\,7\,3\,7\,2\,4\,9\,2\,1\,3\,7\,1\,4\,3\,3\,2\,4\ \ldots\\
x_2 &= 0.\,3\,7\,2\,8\,4\,6\,2\,0\,9\,4\,7\,5\,3\,9\,4\,8\,5\,6\,9\,3\,6\,4\,8\,2\,3\,4\,7\,3\,9\,8\,5\,6\,4\,7\ \ldots\\
x_3 &= 0.\,9\,0\,8\,4\,3\,6\,5\,8\,3\,7\,6\,4\,2\,3\,4\,6\,3\,8\,2\,4\,6\,8\,3\,2\,7\,5\,9\,4\,8\,5\,6\,3\,9\,4\ \ldots\\
x_4 &= 0.\,2\,9\,8\,4\,7\,5\,6\,9\,4\,3\,8\,6\,5\,9\,1\,2\,3\,8\,4\,7\,3\,9\,4\,5\,3\,4\,9\,0\,0\,0\,0\,0\,0\,3\ \ldots\\
x_5 &= 0.\,4\,6\,3\,4\,7\,5\,9\,0\,8\,8\,9\,8\,9\,8\,9\,8\,9\,8\,9\,8\,9\,8\,9\,8\,2\,3\,6\,3\,7\,4\,5\,2\,8\,3\ \ldots\\
x_6 &= 0.\,0\,0\,0\,0\,0\,0\,0\,0\,3\,2\,8\,4\,5\,7\,4\,5\,8\,4\,3\,7\,5\,6\,6\,6\,6\,4\,9\,0\,0\,0\,0\,0\,4\,5\ \ldots\\
x_7 &= 0.\,4\,4\,4\,4\,4\,4\,4\,4\,4\,4\,4\,4\,4\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\,0\ \ldots\\
x_8 &= 0.\,8\,3\,9\,2\,8\,5\,4\,9\,3\,0\,4\,9\,5\,7\,8\,4\,3\,6\,9\,4\,5\,6\,7\,4\,5\,3\,8\,5\,7\,9\,0\,3\,4\,7\ \ldots\\
x_9 &= 0.\,2\,3\,9\,4\,6\,9\,8\,4\,9\,3\,9\,8\,2\,4\,6\,3\,9\,7\,5\,6\,4\,3\,9\,5\,7\,2\,0\,4\,3\,7\,9\,2\,3\,8\ \ldots\\
x_{10} &= 0.\,8\,3\,9\,4\,4\,2\,8\,5\,4\,0\,7\,7\,7\,7\,7\,7\,7\,7\,7\,7\,7\,7\,7\,7\,7\,7\,7\,7\,7\,7\,7\,7\ \ldots\\
x_{11} &= 0.\,8\,3\,9\,2\,6\,5\,6\,8\,5\,4\,4\,0\,0\,1\,0\,0\,0\,1\,0\,0\,0\,0\,0\,1\,0\,0\,0\,0\,1\,0\,0\,0\,0\,0\ \ldots\\
x_{12} &= 0.\,1\,8\,3\,3\,9\,2\,3\,8\,7\,5\,4\,4\,3\,1\,4\,1\,5\,3\,7\,2\,8\,4\,3\,9\,2\,8\,5\,7\,5\,4\,9\,3\,5\,8\ \ldots\\
x_{13} &= 0.\,3\,2\,9\,6\,3\,4\,8\,7\,4\,8\,3\,9\,2\,9\,2\,3\,8\,4\,7\,8\,9\,8\,9\,8\,9\,8\,9\,8\,9\,8\,9\,8\,9\,8\ \ldots\\
&\ \ \vdots\\
z &= 0.\,4\,8\,9\,5\,8\,1\,5\,8\,8\,1\,5\,5\,3\ \ldots
\end{aligned}
$$

Die so konstruierte neue Zahl z kann nicht gleich x_1 sein, da deren 1. Dezimalstelle per Konstruktion unterschiedlich zu der von z ist. Sie kann aber auch nicht gleich x_2 sein, da die 2. Dezimalstelle verschieden ist. Allgemein kann sie nicht gleich x_n sein, weil sie sich in der n-ten Dezimalstelle unterscheidet.

Somit ist z ungleich aller reellen Zahlen in der Aufzählung. z, liegt aber ebenfalls zwischen 0 und 1. Dies steht im Widerspruch zu der Annahme, dass die Reihe alle reellen Zahlen enthält. Damit muss die Annahme, dass so eine Aufzählung existiert, falsch sein. Die reellen Zahlen zwischen 0 und 1 und somit natürlich auch $\mathbb{R}$ im allgemeinen ist damit überabzählbar unendlich. Wie schon erwähnt, ist dies an dieser Stelle eigentlich noch kein gültiger Beweis, da wir die Grundlage (die Dezimaldarstellung reeller Zahlen) erst später einführen werden.

2.6 Summen, Produkte und Kombinatorik

Zur vereinfachten Schreibweise führen wir folgende Definition für Summen und Produkte ein:

Definition 2.33 (Summen- und Produktzeichen)

Für $n, m \in \mathbb{Z}$ und $a_k \in \mathbb{R}$ definieren wir

$$\sum_{k=m}^{n} a_k := \begin{cases} a_m + a_{m+1} + \ldots + a_n, & \text{falls } m \leq n \\ 0, & \text{falls } m > n \end{cases}$$

$$\prod_{k=m}^{n} a_k := \begin{cases} a_m \cdot a_{m+1} \cdot \ldots \cdot a_n, & \text{falls } m \leq n \\ 1, & \text{falls } m > n \end{cases}$$

Bisher sind wir noch größtenteils ohne Potenzen ausgekommen, bis auf ein vereinzeltes x^2. Mit den natürlichen und ganzen Zahlen können wir endlich die ganzzahligen Potenzen definieren.

Definition 2.34 (Ganzzahlige Potenzen)

Für $x \in \mathbb{R}$ und $n \in \mathbb{N}$ definieren wir

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

Für $x \neq 0$ definieren wir außerdem

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

$$x^0 := 1.$$

Damit ergeben sich die folgenden Regeln für das Rechnen mit Potenzen, die wir hier nicht beweisen werden. Sie ergeben sich alle direkt aus der Definition.

Satz 2.35 (Rechenregeln für ganzzahlige Potenzen)

Für $a, b \in \mathbb{R}_{\neq 0}$ und $n, m \in \mathbb{Z}$ gilt:

$$\begin{aligned} a^n a^m &= a^{n+m} \\ (a^n)^m &= a^{nm} \\ a^n b^n &= (ab)^n \end{aligned}$$

▼ Beweis zur Übung

Satz 2.36 (Ordnung von Potenzen)

Für alle $c, d \in \mathbb{R}_{>0}$ und $k \in \mathbb{N}$ gilt:

$$c < d \Leftrightarrow c^k < d^k$$

bzw.

$$c \leq d \Leftrightarrow c^k \leq d^k$$

▼ Beweis

Folgt durch mehrmalige Anwendung von Satz 2.15 (d) und (b):

$\Rightarrow$-Richtung:

$$\begin{aligned} c < d &\Rightarrow (c^2 < cd) \wedge (cd < d^2) && \Rightarrow c^2 < d^2 \\ &\Rightarrow (c^3 < c^2 d) \wedge (c^2 d < d^3) && \Rightarrow c^3 < d^3 \\ &\Rightarrow (c^4 < c^3 d) \wedge (c^3 d < d^4) && \Rightarrow c^4 < d^4 \\ &\Rightarrow \ \ldots \end{aligned}$$

$\Leftarrow$-Richtung:
Beweis funktioniert analog zur $\Rightarrow$-Richtung, wenn mit jeweils mit c^{-1} statt c und d^{-1} statt d multipliziert wird.

Der $\leq$-Fall folgt analog.

□

Es gibt noch ein weiteres spezielles Produkt, die sogenannte Fakultät, die wir im nächsten Kapitel sehr häufig nutzen werden.

Definition 2.37 (Fakultät)

Für $n \in \mathbb{N}_0$ definieren wir die Fakultät von n als

$$n! := \prod_{k=1}^{n} k = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n$$

Nach Definition 2.33 ergibt sich damit insbesondere $0! = 1$.

Beispiel 2.38 (Fakultät)

Die Fakultät hat in der Kombinatorik eine wichtige Bedeutung: Sie gibt die Anzahl an Reihenfolgen an, in die wir eine n-elementige Menge anordnen können. Treten zum Beispiel beim Pferderennen 6 Pferde an, gibt es $6! = 720$ Möglichkeiten, in denen die Pferde ins Ziel einlaufen können (wenn wir ausschließen, dass zwei Pferde gleichzeitig ankommen können): Für den Sieger gibt es 6 Varianten, für den Zweitplatzierten bleiben anschließend noch 5 Möglichkeiten, das ergibt $6 \cdot 5 = 30$ Möglichkeiten für die ersten 2 Plätze. Für den Drittplatzierten bleiben, nachdem die ersten zwei feststehen, noch 4 Kandidaten übrig, was insgesamt $6 \cdot 5 \cdot 4 = 120$ Möglichkeiten für die ersten 3 Plätze ergibt. Führt man die Überlegung fort, kommt man insgesamt auch $6! = 6 \cdot 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 720$ Möglichkeiten.

Die Werte der Fakultät wachsen relativ schnell. Wir werden später noch das Wachstum von $n!$ mit dem von anderen Funktionen wir n^2 oder 2^n vergleichen.

$$\begin{aligned}
1! &= 1 \\
2! &= 2 \\
3! &= 6 \\
4! &= 24 \\
5! &= 120 \\
6! &= 720 \\
7! &= 5040 \\
8! &= 40320 \\
9! &= 362880 \\
10! &= 3628800 \\
20! &= 2432902008176640000
\end{aligned}$$

Definition 2.39 (Binomialkoeffizient)

Für $n, k \in \mathbb{N}_0$ definieren wir den Binomialkoeffizienten

$$\binom{n}{k} := \begin{cases} \frac{n!}{(n-k)!\cdot k!} = \frac{n\cdot(n-1)\cdot\ldots\cdot(n-k+1)}{k\cdot(k-1)\cdot\ldots\cdot 1} & \text{für } n \geq k \\ 0 & \text{für } n < k \end{cases}$$

Beispiel 2.40 (Binomialkoeffizient)

Der Binomialkoeffizient hat in der Kombinatorik genauso wie die Fakultät eine wichtige Bedeutung. So gibt $\binom{n}{k}$ die Anzahl an Möglichkeiten an, wie man k Elemente aus einer n-elementigen Menge auswählen kann.

Nehmen wir beispielsweise an, wir hätten in einer Übungsgruppe 10 Studierende, die wir mit 1 bis 10 durchnummerieren. Nach Abgabe des ersten Zettels sollen zwei von ihnen eine Aufgabe vorrechnen. Der Tutor wählt diese zwei zufällig aus. Dann gibt es $\binom{10}{2} = 45$ Möglichkeiten, welche Studierenden drankommen könnten. Diese kann man noch überschaubar aufzählen:

▶ Möglichkeiten

Die Formel begründet sich wie folgt: $n!$ sind die möglichen Reihenfolgen, alle Elemente der Menge anzuordnen. Für die Auswahl von k Elementen interessiert uns aber nicht die Reihenfolge der k Elemente, also teilt man durch die Anzahl an Möglichkeiten, wie man k anordnen kann ($k!$). Außerdem interessiert uns auch nicht die Reihenfolge der $n-k$ nicht ausgewählten Elemente, also teilt man ebenfalls durch deren Anordnungsmöglichkeiten ($(n-k)!$). Damit ergibt sich also

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

Alternativ kann man die Erklärung auch mit den angeordneten Möglichkeiten der k Elemente starten: Für das erste Element gibt es n Möglichkeiten zur Auswahl, für das zweite dann noch $(n-1)$, usw. Also insgesamt $n\cdot(n-1)\cdot\ldots\cdot(n-k+1)$ angeordnete Möglichkeiten. Nun interessiert uns wieder nicht die Reihenfolge der ausgewählten k Elemente, deswegen sind für uns immer jeweils $k!$ dieser Möglichkeiten identisch, da sie nur anders angeordnet sind. Daher wird noch durch $k!$ geteilt:

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

Es lässt sich leicht zeigen, dass beide Formeln identisch sind.

Das vorherige Beispiel zeigt die Bedeutung des Binomialkoeffizienten in der *abzählenden Kombinatorik*, bei der es darum geht, die Anzahl günstiger Ergebnisse aus einer endlichen Menge möglicher Ergebnisse zu bestimmen. Abzählende Kombinatorik ist ein Vorläufer der Wahrscheinlichkeitsrechnung und wurde im 17 Jahrhundert entwickelt, um den Ausgang von Glücksspielen vorherzusagen. Der folgende Satz fasst einige Ergebnisse zusammen.

Satz 2.41 (Kombinationen)

Die Anzahl der Möglichkeiten für die Auswahl von k Elementen einer n-elementigen Menge für $k, n \in \mathbb{N}_0$ ist:

	Anordnung im $k-$Tupel	keine Anordnung
mit Zurücklegen	n^k	$\binom{n+k-1}{k}$
ohne Zurücklegen	$\frac{n!}{(n-k)!}$	$\binom{n}{k}$

▼ Beweis

Zunächst sollen die gezogenen Elemente von links nach rechts in einem k-Tupel abgelegt werden.

Wenn die Elemente nach jedem Zug wieder in die Menge zurückgelegt werden können, dann hat man

$$\left.\begin{array}{ll} \text{bei der 1. Komponente} & : n \text{ Möglichkeiten} \\ \text{bei der 2. Komponente} & : n \text{ Möglichkeiten} \\ \vdots & \vdots \\ \text{bei der } k\text{. Komponente} & : n \text{ Möglichkeiten} \end{array}\right\} \underbrace{n \cdot n \cdot \ldots \cdot n}_{k \text{ mal}} = n^k$$

Wenn die Elemente nach jedem Zug nicht wieder in die Menge zurückgelegt werden können, dann hat man

$$\left.\begin{array}{ll} \text{bei der 1. Komponente} & : n \text{ Möglichkeiten} \\ \text{bei der 2. Komponente} & : n-1 \text{ Möglichkeiten} \\ \vdots & \vdots \\ \text{bei der } k\text{. Komponente} & : n-(k-1) \text{ Möglichkeiten} \end{array}\right\} = \frac{n!}{(n-k)!}$$

Wenn die Reihenfolge der gezogenen Elemente keine Rolle spielt, können wir die Argumentation aus dem verherigen Beispiel nutzen. Für den Fall ohne Zurücklegen haben wir im Beispiel bereits zwei Beweisansätze gesehen. Damit bleibt noch der Fall mit Zurücklegen. In diesem Fall nehmen wir an, dass wir n Behälter, man spricht in der Kombinatorik oft von Urnen, in denen sich jeweils mindestens k identische Kugeln befinden. Wir können nun entscheiden, wie wir unsere k Elemente ziehen. Dazu ziehen wir i_1 aus der ersten Urne, i_2 aus der zweiten, usw. Es muss dann gelten $i_1 + i_2 + \ldots + i_n = k$. Man könnte das Problem auch so formulieren, dass man zwischen zwei Urnen, also z.B. den Urnen j und $j+1$ i_j identische Symbole einfügt. Dies entspricht gerade der Verteilung von k identischen Symbolen auf $n-1$ Lücken. Wenn man dies als eine Sequenz darstellt, bei der die Reihenfolge der Urnen fest vorgegeben ist und die k identischen Symbole dazwischen zugeordnet werdne können, so haben wir

insgesamt $n-1+k$ Elemente, aus denen k ausgewählt werden. Diese Anzahl ist aber gerade $\binom{n+k-1}{k}$.

Beispiel 2.42 (Glücksspiele)

Beim Fußballtoto müssen 13 Spiele getippt werden. Pro Spiel gibt es 3 Möglichkeiten zu tippen, 1 Heimsieg, 0 unentschieden und 2 Auswärtssieg. Damit gibt es $3^{13} = 1594323$ mögliche Tipps.

Beim Lottospiel müssen 6 aus 49 Zahlen ausgewählt werden. Die Reihenfolge der Auswahl spielt natürlich keine Rolle. Damit gibt es $\binom{49}{6} = 13983816$ Möglichkeiten

Wenn zusätzlich noch die Superzahl zum Gewinn des Jackpots richtig sein soll, so gibt es $\binom{49}{6}\binom{10}{1} = 139838160$ Möglichkeiten

Daraus folgt auch, dass die Wahrscheinlichkeit für 6 richtge Zahlen und eine falsche Superzahl ungefähr $\left(\binom{49}{6}\right)^{-1}\frac{9}{10} \approx 1/15537573 \approx 0.000000064360$ ist.

Wenn wir beim Würfeln mit 3 Würfeln eine möglichst hohe dreistellige Zahl bilden sollen, so gibt es $\binom{6+3-1}{3} = \binom{8}{3} = 56$ Möglichkeiten.

Wenn man bei einem Autorennen mit 20 Teilnehmenden die ersten 3 Plätze tippen möchte, so gibt es $\frac{20!}{(20-3)!} = \frac{20!}{17!} = 6840$ Möglichkeiten.

Für Binomialkoeffizienten gilt eine nützliche Gleichheit:

Satz 2.43 (Summe von Binomialkoeffizienten)

$$\forall n, k \in \mathbb{N}_0 : \binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$$

▼ Beweis als Übung

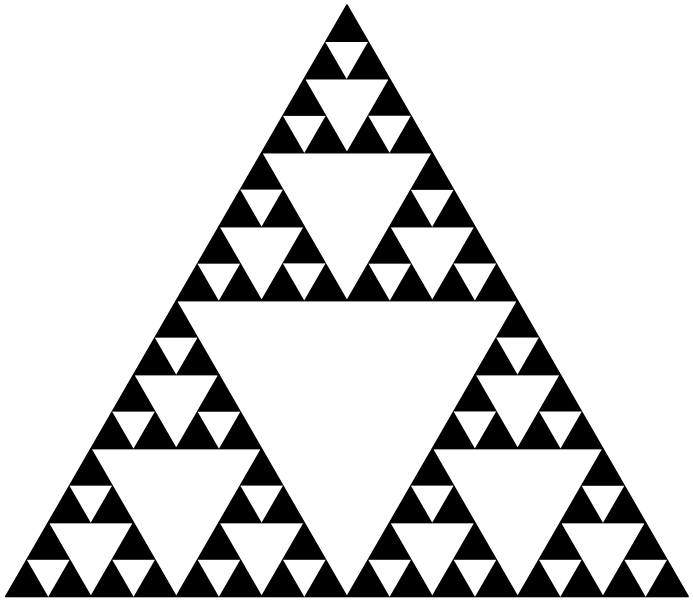
Anschaulich ergibt sich mit dem obigen Satz das sogenannte *Pascalsche Dreieck* nach Blaise Pascal:

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

Für den rechten und linken Schenkel des Dreiecks ist jeder Binomialkoeffizient jeweils 1 und jeder Eintrag in der Mitte des Dreiecks ergibt sich aus der Summe der beiden darüber liegenden. Somit ergeben sich insgesamt die Einträge:

$$
\begin{array}{ccccccccc}
 & & & 1 & & & & & \\
 & & 1 & & 1 & & & & \\
 & 1 & & 2 & & 1 & & & \\
1 & & 3 & & 3 & & 1 & & \\
1 & & 4 & & 6 & & 4 & & 1
\end{array}
$$

Das Dreieck lässt sich hiermit sehr einfach nach unten für höhere Binomialkoeffizienten erweitern. Wenn man ein kleines schwarzes Dreieck für jeden ungeraden Eintrag im Pascalschen Dreieck zeichnet, ergibt sich eine Struktur, die auch als Sierpinski-Dreieck bekannt ist, hier gezeigt bis $n = 15$:



Mithilfe von Binomialkoeffizienten können wir eine Verallgemeinerung der binomischen Formel $(a + b)^2 = a^2 + 2ab + b^2$ beweisen, die wir im Laufe der Vorlesung häufiger benötigen werden:

Satz 2.44 (Binomischer Lehrsatz)

Für $a, b \in \mathbb{R}$ und $n \in \mathbb{N}$ gilt:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

▼ Beweis

Beweis per vollständiger Induktion:

Induktionsanfang ($n = 1$):

$$(a+b)^1 = \binom{1}{0} a + \binom{1}{1} b = a + b.$$

Induktionsvoraussetzung: Wir nehmen an, folgende Aussage gilt für ein $n \in \mathbb{N}$:

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k.$$

Induktionsschritt ($n \to n+1$):

$$(a+b)^{n+1} = (a+b)(a+b)^n \overset{IV}{=} (a+b) \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k = \sum_{k=0}^{n} \binom{n}{k} a^{n-k+1} b^k + \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^{k+1}.$$

Die beiden Summen betrachten wir einmal in ausführlicher Form:

$$\begin{array}{lllllll} \sum_{k=0}^{n} \binom{n}{k} a^{n-k+1} b^k & = \binom{n}{0} a^{n+1} + & \binom{n}{1} a^n b + & \binom{n}{2} a^{n-1} b^2 + \ldots + & \binom{n}{n} a b^n & & \\ \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^{k+1} & = & \binom{n}{0} a^n b + & \binom{n}{1} a^{n-1} b^2 + \ldots + & \binom{n}{n-1} a b^n & + \binom{n}{n} b^{n+1} \end{array}$$

Es gilt

$$\binom{n}{0} = 1 = \binom{n+1}{0}$$

und genauso

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

Nach Satz 2.43 können wir außerdem die Terme, die jeweils untereinander stehen zusammenfassen, da z.B.:

$$\binom{n}{1} + \binom{n}{0} = \binom{n+1}{1}$$

$$\binom{n}{2} + \binom{n}{1} = \binom{n+1}{2}.$$

Somit folgt zusammengefasst:

$$(a+b)^{n+1} = \sum_{k=0}^{n} \binom{n}{k} a^{n-k+1} b^k + \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^{k+1} = \sum_{k=0}^{n+1} \binom{n+1}{k} a^{(n+1)-k} b^k.$$

Damit gilt die Aussage für $n+1$ und somit für alle $n \in \mathbb{N}$.

□

Damit kann man die Vorfaktoren einer allgemeinen binomischen Formel einfach aus dem Pascalschen Dreieck bestimmen. Für $n = 4$ ist die Zeile im Dreieck: $(1, 4, 6, 4, 1)$, damit ergibt sich also: $(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$.

Es folgen zwei nützliche Ungleichungen, die wir häufig nutzen werden, wenn wir es mit Potenzen zu tun haben, bei denen der Exponent variabel ist:

Satz 2.45 (Bernoullische Ungleichung)

Für alle $x \in \mathbb{R}$ mit $x \geq -1$ und alle $n \in \mathbb{N}_0$ gilt

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

▼ Beweis

Per Induktion über n.

Induktionsanfang $n = 0$**:**
$(1+x)^0 = 1 = 1 + 0 \cdot x$ und damit auch $(1+x)^0 \geq 1 + 0 \cdot x$.

Induktionsvoraussetzung:
Wir nehmen an, dass $(1+x)^n \geq 1 + nx$ für ein beliebiges $n \in \mathbb{N}_0$ gilt.

Induktionsschritt $n \to n+1$**:**
Mit $1 + x \geq 0$ folgt

$$\begin{aligned}
(1+x)^{n+1} &= (1+x)(1+x)^n \\
&\overset{IV}{\geq} (1+x)(1+nx) \\
&= 1 + nx + x + nx^2 \\
&= 1 + (n+1)x + nx^2 \\
&\geq 1 + (n+1)x \qquad \text{(da } n \geq 0 \text{ und } x^2 \geq 0\text{)}
\end{aligned}$$

□

Sie können zur Übung einmal versuchen, die Bernoullische Ungleichung für $x \geq 0$ mit dem Binomischen Lehrsatz zu beweisen. Die folgende zweite nützliche Abschätzung folgt ebenfalls direkt aus dem Binomischen Lehrsatz. Den Beweis überlassen wir Ihnen zur Übung.

Satz 2.46

Für alle $x \in \mathbb{R}$ mit $x \geq 0$ und alle $n \in \mathbb{N}$ mit $n \geq 2$ gilt

$$(1+x)^n \geq \frac{n^2 x^2}{4}.$$

▼ Beweis als Übung

2.7 Komplexe Zahlen

Zuletzt wollen wir noch eine Zahlenmenge vorstellen, die man in der Regel noch nicht aus der Schule kennt. Die *komplexen Zahlen* $\mathbb{C}$ sind eine sehr nützliche Erweiterung der reellen Zahlen (also $\mathbb{R} \subset \mathbb{C}$), die in vielen Bereichen Anwendung findet, wie zum Beispiel in der Signalverarbeitung, der Computergraphik oder der Physik. Außerdem werden wir sehen, dass manche Gesetzmäßigkeiten in der Analysis erst ihre volle Aussagekraft erhalten, wenn wir sie im komplexen Zahlenbereich betrachten. Beginnen wir also mit der Definition von $\mathbb{C}$.

Definition 2.47 (Komplexe Zahlen)

Wir definieren die Menge der komplexen Zahlen $\mathbb{C}$ als

$$\mathbb{C} := \{(a, b) \in \mathbb{R} \times \mathbb{R}\}$$

Für alle $z = (a, b) \in \mathbb{C}$ nennen wir $\mathrm{Re}(z) := a$ den *Realteil* von z und $\mathrm{Im}(z) := b$ den *Imaginärteil* von z.

Wir können $(a, 0) \in \mathbb{C}$ mit $a \in \mathbb{R}$ identifizieren, so dass $\mathbb{R}$ als Teilkörper von $\mathbb{C}$ angesehen werden kann.

Satz 2.48 (Der Körper der komplexen Zahlen)

Die Menge $\mathbb{C}$ bildet mit folgender Addition und Multiplikation (für $z_1 = (a_1, b_1), z_2 = (a_2, b_2)$)

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

einen Körper (auf der rechten Seite der Definition wurde die übliche Addition und Multiplikation der reellen Zahlen genutzt).

Das Null-Element ist $(0, 0)$, das Eins-Element ist $(1, 0)$.

Das negative Element zu (a, b) und das inverse Element zu $(a, b) \neq (0, 0)$ sind

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

▼ Beweis als Übung

Da wir bereits gezeigt haben, dass das Nullelement, Einselement, negative Element und inverse Element in jedem Körper eindeutig sind, genügt es diese anzugeben und Sie müssen nun nur noch die Körpereigenschaften von $\mathbb{C}$ mithilfe der Körpereigenschaften von $\mathbb{R}$ nachweisen.

Häufig sieht man auch die Schreibweise $(a, b) = a + \mathrm{i}b$, wobei i die sogenannte imaginäre Einheit ist und als Lösung von $x^2 = -1$ definiert wird. Damit ergeben sich die oben angegebenen Definitionen für Addition und Multiplikation auf ganz natürliche Weise:

$$\begin{aligned}(a_1, b_1) + (a_2, b_2) &= a_1 + \mathrm{i}b_1 + a_2 + \mathrm{i}b_2 \\ &= (a_1 + a_2) + \mathrm{i}(b_1 + b_2) \\ &= (a_1 + a_2, b_1 + b_2)\end{aligned}$$

$$\begin{aligned}(a_1, b_1) \cdot (a_2, b_2) &= (a_1 + \mathrm{i}b_1)(a_2 + \mathrm{i}b_2) \\ &= a_1a_2 + \mathrm{i}a_1b_2 + \mathrm{i}a_2b_1 + \mathrm{i}^2b_1b_2 \\ &= a_1a_2 - b_1b_2 + \mathrm{i}(a_1b_2 + a_2b_1) \\ &= (a_1a_2 - b_1b_2, a_1b_2 + a_2b_1)\end{aligned}$$

Beide Schreibweisen haben ihre Vorteile: Mit der imaginären Einheit lässt sich leichter rechnen, da sich hier die Multiplikation aus der reellen Multiplikation ergibt. Die (a, b)-Schreibweise lässt uns die komplexen Zahlen als zweidimensionale Vektoren interpretieren, die sich besonders gut zur Visualisierung komplexer Zahlen eignet. Dabei nutzen wir die zweidimensionale Ebene: Auf der horizontalen Achse, die auch als *reelle Achse* bezeichnet wird, geben wir den Realteil der komplexen Zahl an und auf der vertikalen Achse, welche auch *imaginäre Achse* genannt wird, geben wir ihren Imaginärteil an. Damit können wir komplexe Zahlen z als Punkte in der sogenannten *komplexen Ebene* visualisieren und die Addition wird hierbei zu einer Vektoraddition. Probieren Sie dies gerne in den beiden folgenden Demos aus.

▶ Demo: Komplexe Zahlenebene

▶ Demo: Komplexe Addition

Wir haben bereits angegeben, dass $(\mathbb{C}, +, \cdot)$ ein Körper ist. Aber lässt sich dieser Körper auch ordnen und damit alle Folgerungen aus den Ordnungseigenschaften auf $\mathbb{C}$ übertragen?

Satz 2.49 (Die komplexen Zahlen lassen sich nicht ordnen)

$(\mathbb{C}, +, \cdot)$ ist **kein** geordneter Körper.

▼ Beweis

Wir führen einen Widerspruchsbeweis:

Angenommen, $\mathbb{C}$ wäre ein geordneter Körper, dann gibt es eine Menge positiver komplexer Zahlen P und es gilt entweder $(0, 1) = i \in P$ oder $i \notin P$.

Nehmen wir an, i ist positiv. Dann muss das Produkt zweier positiver Zahlen nach O3 auch positiv sein, also müssen auch $i^2 = -1$ und $(-1)^2 = 1$ positiv sein. Dies widerspricht aber der Eigenschaft O1, nach der nur entweder 1 oder -1 positiv sein kann (↯).

Wenn wir annehmen, dass i negativ ist, dann ist $-i$ positiv. Damit sind erneut $(-i)^2 = (-1)^2(i^2) = -1$ und $(-1)^2 = 1$ positiv, was auf den selben Widerspruch führt (↯).

$\mathbb{C}$ lässt sich also nicht ordnen.

□

Damit können wir zwei komplexe Zahlen also im Allgemeinen nicht mit $z_1 < z_2$ ordnen. Wie wir bereits bei der Definition 2.18 erwähnt haben, lässt sich aber auch auf manchen ungeordneten Körpern eine Metrik definieren, mit der wir den Abstand zwischen zwei komplexen Zahlen bestimmen können. Dazu führen wir die komplexe Konjugation ein, mit deren Hilfe wir den Betrag einer komplexen Zahl definieren:

Definition 2.50 (Komplexe Konjugation, komplexer Betrag)

Wir definieren die *komplex Konjugierte* $\overline{z}$ einer komplexen Zahl $z = (a, b) = a + \mathrm{i}b \in \mathbb{C}$ als

$$\overline{z} := (a, -b) = a - \mathrm{i}b.$$

Darüber hinaus definieren wir den (Absolut-)Betrag von $z \in \mathbb{C}$ als:

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

Satz 2.51 (Metrik der komplexen Zahlen, Konjugationseigenschaften)

Für beliebige $z_1, z_2 \in \mathbb{C}$ definiert $d(z_1, z_2) = |z_1 - z_2|$ eine Metrik auf $\mathbb{C}$.

Außerdem gelten für alle $z_1, z_2 \in \mathbb{C}$ die folgenden Eigenschaften für die komplexe Konjugation:

- $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$
- $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$

▼ Beweis zur Übung

Anders als wir es von reellen Zahlen gewöhnt sind, gilt hier bei einem Imaginärteil $b \neq 0$:

$$z \cdot z = z^2 = (a, b)^2 = a^2 + 2abi - b^2 \neq a^2 + b^2 = |z|^2,$$

also $z^2 \neq |z|^2$. Insbesondere ist z^2 im Allgemeinen komplex.

Der Betrag einer komplexen Zahl ist stets reell und somit lassen sich Beträge komplexer Zahle ordnen. Für den weiteren Verlauf der Vorlesung haben wir uns, anders als viele Lehrbücher, dagegen entschieden, alle Sätze und Definitionen direkt für komplexe Zahlen einzuführen und anschließend den reellen Spezialfall zu betrachten. Wir werden umgekehrt für den Großteil der Vorlesung mit reellen Zahlen arbeiten und nur ab und zu auf den allgemeineren komplexen Fall hinweisen. Hier hilft uns der Betrag, denn wenn alle Größenvergleiche in Sätzen und Definitionen nur in Verbindung mit Beträgen vorkommen, lassen sich die Schlussfolgerungen meist sehr einfach auf $\mathbb{C}$ erweitern.

Eine besondere Bedeutung haben komplexe Zahlen mit $|z| = 1$. Stellen wir diese auf der komplexen Ebene dar, liegen sie auf einem Einheitskreis um den Ursprung. Wenn wir zwei komplexe Zahlen z_1, z_2 mit $|z_1| = |z_2| = 1$ multiplizieren, ergibt sich für das Produkt:

$$|z_1 \cdot z_2|^2 = (z_1 \cdot z_2)\overline{(z_1 \cdot z_2)} = (z_1 \cdot z_2)(\overline{z_1} \cdot \overline{z_2}) = (z_1 \cdot \overline{z_1})(z_2 \cdot \overline{z_2}) = |z_1|^2 |z_2|^2 = 1.$$

Das Produkt liegt also ebenfalls auf dem Einheitskreis. In der folgenden Demo können Sie mit solchen Produkten experimentieren und werden dabei eventuell etwas feststellen:

▶ Demo: Komplexe Multiplikation

Das komplexe Produkt scheint einer Drehung zu entsprechen, wobei der Winkel zwischen den beiden Faktoren mit der reellen Achse addiert wird. Beispielsweise hat i einen Winkel von 90° zur reellen Achse, $i^2 = -1$ einen Winkel von $90^\circ + 90^\circ = 180^\circ$ und $(-1)i = -i$ einen Winkel von $180^\circ + 90^\circ = 270^\circ$. Auch für andere Prdukte und Winkel sieht es optisch so aus, als würde das komplexe Produkt einer Drehung entsprechen. Dies ist natürlich noch kein Beweis, wir werden aber im Laufe der Veranstaltung noch zeigen, warum dies immer gilt.

Betrachten wir eine allgemeine Komplexe Zahl $z = a + ib$ mit $|z| = \sqrt{a^2 + b^2} = r$, so ergibt sich

$$z = a + ib = r\left(\frac{a}{r} + i\frac{b}{r}\right) = r(\hat{a} + i\hat{b}) = r\hat{z}$$
$$|\hat{z}|^2 = \left|\frac{a}{r} + i\frac{b}{r}\right|^2 = \frac{a^2 + b^2}{r^2} = 1$$

Wir können also eine beliebige komplexe Zahl über ihren Betrag (Abstand zum Ursprung) und eine andere komplexe Zahl $\hat{z}$, die in der komplexen Ebene auf dem Einheitskreis liegt, darstellen. Für ein allgemeines Produkt komplexer Zahlen $z_1 = r_1\hat{z}_1, z_2 = r_2\hat{z}_2$ ergibt sich:

$$z_1 z_2 = r_1\hat{z}_1 r_2\hat{z}_2 = (r_1 r_2)(\hat{z}_1\hat{z}_2)$$
$$|z_1 z_2| = r_1 r_2 \underbrace{|\hat{z}_1\hat{z}_2|}_{=1} = r_1 r_2$$

Wir können also beliebige Produkte komplexer Zahlen auf Einheitskreisprodukte zurückführen. Der Betrag des Produkts entspricht dabei dem Produkt der Einzelbeträge. Außerdem addiert sich der Winkel der Faktoren zum Gesamtwinkel des Produkts. Letzteres können wir hier noch nicht beweisen, holen dies aber später in Kapitel 4 nach.

3 Folgen & Reihen

Viele Verfahren zur Lösung realer Probleme sind sogenannte *iterative Algorithmen*. Dabei startet man mit einem groben Schätzwert für die Lösung und berechnet darauf aufbauend eine zweite, (im Idealfall) bessere Schätzung. Dieses Vorgehen wiederholt (iteriert) man so lange, bis man schließlich bei einer zufriedenstellend genauen Lösung ankommt. Einfache Beispiele sind die Berechnungen beliebig genauer Lösungen für Wurzeln, Logarithmen, allgemeine Potenzen (Exponent ist keine ganze Zahl), oder von Zahlen wie π oder e. Wie können wir aber zeigen, dass ein solcher iterativer Ansatz auch wirklich zur gewünschten Lösung führt? Dem mathematischen Konzept hinter dieser Frage, dem sogenannten Grenzwert von Folgen und Reihen, werden wir uns in diesem Kapitel widmen. Folgen, Reihen und deren Grenzwerte sind eine weitere wichtige Grundlage für alle nachfolgenden Kapitel und werden Ihnen auch in anderen Vorlesungen begegnen.

3.1 Folgen und Grenzwerte

Definition 3.1 (Folge)

Unter einer Folge versteht man eine Abbildung, bei der jedem $n \in \mathbb{N}$ ein $a_n \in \mathbb{R}$ zugeordnet wird. Man schreibt für die Folge auch kurz (a_n) oder $(a_n)_{n\in\mathbb{N}}$.

Wir bezeichnen n als den *Folgenindex* und a_n als die *Folgenglieder* von (a_n).

Um eine bestimmte Folge zu definieren, geben wir eine Vorschrift zur Berechnung der Folgenglieder an.

Beispiel 3.2 (Folgen)

- Die Folge $(a, a, a, a, \ldots)$ kann definiert werden durch

$$a_n = a.$$

- Die Folge $\left(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots\right)$ kann definiert werden durch

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

- Die Folge $(-1, 1, -1, 1, \ldots)$ kann definiert werden durch

$$a_n = (-1)^n.$$

- Die Folge $\left(1, \frac{2^k}{\sqrt{2}}, \frac{3^k}{\sqrt{3}}, \frac{4^k}{\sqrt{4}}, \ldots\right)$ kann definiert werden durch

$$a_n = \frac{n^k}{\sqrt{n}}.$$

Eine besondere Art von Folgen sind die *rekursiven Folgen*, bei denen eine bestimmte Anzahl erster Folgenglieder $a_1, a_2, \ldots, a_h$ angegeben werden und sich jedes weitere Folgenglied a_n mit $n > h$ aus den Gliedern a_{n-1} bis a_{n-h} bestimmen lässt.

Beispiel 3.3 (rekursive Folgen)

- Die Fibonacci-Folge $(0, 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, ...)$ ist definiert durch

$$\begin{aligned} a_1 &= 0 \\ a_2 &= 1 \\ a_n &= a_{n-1} + a_{n-2}. \end{aligned}$$

- Die Mandelbrot-Folge ist für ein $c \in \mathbb{C}$ definiert durch

$$\begin{aligned} a_1 &= 0 \\ a_{n+1} &= a_n^2 + c. \end{aligned}$$

  Diese Folge definiert die sogenannte Mandelbrotmenge, welche in dieser Vorlesung besprochen wurde.

- Die ULAM-Folge ist für einen beliebigen Startwert $k \in \mathbb{N}$ definiert als

$$\begin{aligned} a_0 &= k \\ a_{n+1} &= \begin{cases} 1, & \text{wenn } a_n = 1, \\ 3a_n + 1, & \text{wenn } a_n \text{ ungerade ist,} \\ a_n/2, & \text{wenn } a_n \text{ gerade ist.} \end{cases} \end{aligned}$$

  Mit dem Startwert $a_0 = 3$ ergibt sich beispielsweise $(3, 10, 5, 16, 8, 4, 2, 1, 1, 1, 1, 1, ...)$.

  Es ist bis heute eine nicht bewiesene Vermutung, dass die ULAM-Folge unabhängig vom Startwert immer auf eine $(1, 1, 1, 1, ...)$-Folge hinauslaufen muss. Ein interessantes Video dazu finden Sie hier.

Im Zusammenhang mit Folgen werden wir häufiger von dem Begriff des Unendlichen (∞) Gebrauch machen. Dieser wird nun definiert.

Definition 3.4 (Unendlich)

Die Menge $\hat{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$ heißt *erweiterte reelle Zahlengerade*. Die zusätzlichen Elemente $-\infty$ und ∞ sind über die folgende Eigenschaft definiert:

$$\forall x \in \mathbb{R} : \ -\infty < x < \infty$$

Wichtig zu betonen ist, dass ∞ kein Element von $\mathbb{R}$ ist. Mit $\pm\infty$ kann also auch nicht wie mit gewöhnlichen reellen Zahlen gerechnet werden. Wenn wir also schreiben "$\ldots = \infty$", dann ist die linke Seite der Gleichung damit nicht reell. Wenn wir in Sätzen fordern, dass eine Lösung für eine bestimmte Gleichung existiert, so ist damit niemals die Lösung $\pm\infty$ gemeint. Wir wollen mit "

… $= \infty$" oder auch "… $\to \infty$" ausdrücken, dass der Term auf der linken Seite beliebig groß wird, wie in der folgenden Definition eines Grenzwerts für Folgen:

Definition 3.5 (Konvergente Folge, Divergenz, Grenzwert, Nullfolge)

Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt konvergent gegen $a \in \mathbb{R}$, wenn für jedes $0 < \varepsilon \in \mathbb{R}$ ein $n_0 \in \mathbb{N}$ existiert, sodass $|a_n - a| < \varepsilon$ für alle $n \geq n_0$ gilt, oder kurz:

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

Darüber hinaus vereinbaren wir:

1. Eine Folge, die nicht konvergiert, nennen wir *divergent*.
2. Falls (a_n) gegen a konvergiert, so nennen wir a den *Grenzwert* von (a_n) und schreiben

$$\lim_{n\to\infty} a_n = a \qquad \text{oder} \qquad a_n \to a \text{ für } n \to \infty.$$

3. Eine Folge, die gegen 0 konvergiert, nennen wir *Nullfolge*.

Dies ist unsere erste Quantorenaussage mit drei Quantoren hintereinander. Dabei geht man von links nach rechts vor: Für jedes $\varepsilon > 0$ müssen wir zeigen, dass ein $n_0 \in \mathbb{N}$ existiert, und dann muss ab diesem n_0 für jedes Folgenglied die Eigenschaft $|a_n - a| < \varepsilon$ gelten. Wir werden hier die Konvention nutzen, dass $\varepsilon > 0$ auch gleichzeitig $\varepsilon \in \mathbb{R}$ impliziert, sodass wir dies nicht mehr explizit mitschreiben werden. Die Bedingung für die Folgenkonvergenz nennt man auch das ε-Kriterium der Konvergenz. Solche ε-Kriterien werden ab sofort immer mal wieder auch in anderen Definitionen und Sätzen vorkommen.

Das Konvergenzkriterium können wir uns gut mit dem folgenden Bild veranschaulichen:

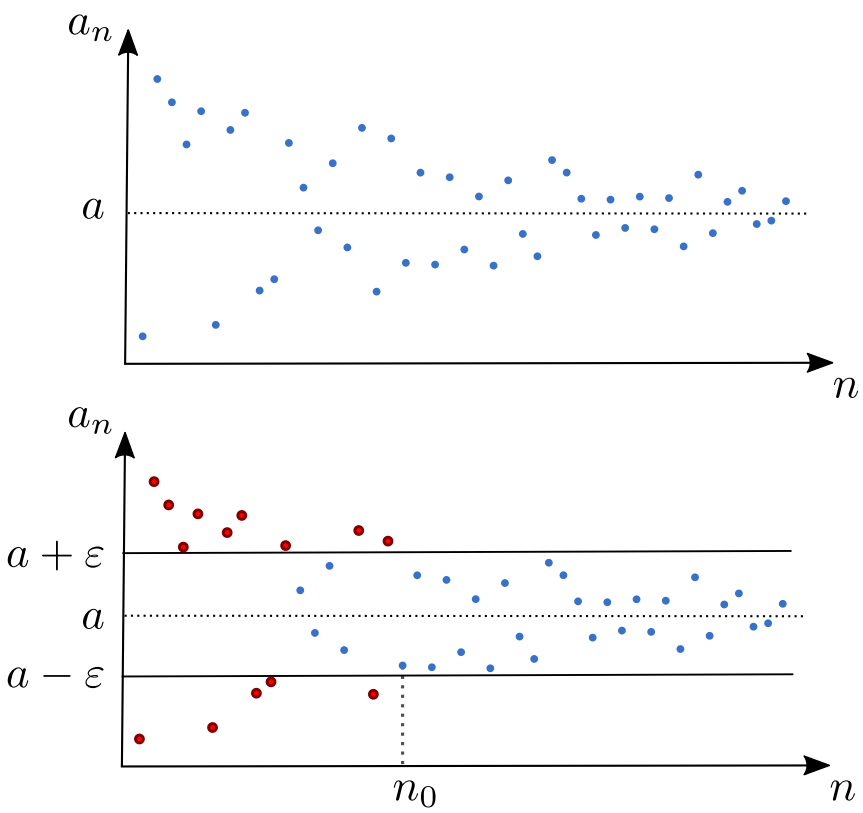


Bild von Ceranilo, CC BY-SA 4.0

Es muss für eine konvergente Folge stets ein n_0 geben, ab dem alle Folgenglieder in dem sogenannten ε-Schlauch um den Grenzwert a liegen. Sie können die Konvergenzaussage auch gerne mit der folgenden Demo testen:

▶ Demo: Folgenkonvergenz

Den n_0-Teil der Konvergenzaussage ($\exists n_0 \in \mathbb{N}\, \forall n \geq n_0$) kann man auch so übersetzen, dass die Aussage $|a_n - a| < \varepsilon$ nur für endlich viele Folgenglieder nicht gilt. Manchmal liest man auch die Formulierung, dass sie für “fast alle” Folgenglieder gilt.

Beispiel 3.6 (Grenzwerte und Konvergenz von Folgen)

- Die konstante Folge $a_n = a$ konvergiert gegen a. Dies folgt für beliebige n_0, denn

$$|a_n - a| = |a - a| = |0| = 0 < \varepsilon$$

  gilt immer.

- Für die Folge (b_n) mit $b_n = \frac{1}{n}$ gilt $\lim\limits_{n\to\infty} b_n = 0$.

  Herangehensweise:
  Für den Beweis rechnet man typischerweise zunächst ein paar Folgenglieder aus, damit man eine Idee bekommt, was der Grenzwert sein könnte. Dadurch vermutet man den Grenzwert $b = 0$.
  Dann betrachtet man $|b_n - b| = |1/n - 0| < \varepsilon$ und formt dies nach n um: $n > 1/\varepsilon$.
  $1/\varepsilon$ kann aber nicht n_0 sein, da $1/\varepsilon$ im Allgemeinen keine natürliche Zahl ist. Daher verwendet man die obere Gaußklammer (siehe Satz 2.27) und wählt $n_0 = \lceil 1/\varepsilon \rceil + 1$. Dabei stellt das $+1$ sicher, dass wir nicht (bei ungünstiger Wahl von ε) den Fall $n_0 = 1/\varepsilon$ erhalten, sonder stets $n_0 > 1/\varepsilon$.

  Damit gilt für $n \geq n_0 = \lceil 1/\varepsilon \rceil + 1$ die Konvergenzaussage. Diese Vorüberlegungen schreibt man normalerweise nicht in den Beweis, sondern geht für den formalen Beweis in umgekehrter Reihenfolge vor.

  Beweis:
  Sei $\varepsilon > 0$ und $n_0 = \left\lceil \frac{1}{\varepsilon} \right\rceil + 1$, dann folgt:

$$|b_n - 0| = |b_n| = \left|\frac{1}{n}\right| \overset{>0}{=} \frac{1}{n} \leq \frac{1}{n_0} = \frac{1}{\lceil 1/\varepsilon \rceil + 1} < \frac{1}{1/\varepsilon} = \varepsilon$$

  Damit ist die Folge (b_n) eine Nullfolge.

  ▶ Folge visualisieren

- Betrachten wir als letztes Beispiel die Folge $(c_n)_{n\in\mathbb{N}}$ mit $c_n = x^n$, deren Verhalten vom Wert $x \in \mathbb{R}$ abhängt:

  1. Falls $|x| < 1$ gilt $\lim\limits_{n\to\infty} x^n = 0$.
  2. Falls $x = 1$ gilt $\lim\limits_{n\to\infty} x^n = 1$.
  3. Falls $x = -1$, oder $|x| > 1$ divergiert die Folge.

  Führen Sie die Beweise in Ruhe zur Übung durch.

Wie bereits im letzten Kapitel im Bezug auf Wurzeln, Nullelemente und Suprema, interessiert uns neben der Existenz eines Grenzwerts auch dessen Eindeutigkeit. Mit anderen Worten: Ist es möglich, dass eine Folge zwei verschiedene Grenzwerte hat?

Satz 3.7 (Eindeutigkeit des Grenzwerts)

Der Grenzwert einer Folge ist, falls er existiert, eindeutig.

▼ Beweis

Sei $(a_n)_{n\in\mathbb{N}}$ eine konvergente Folge. Angenommen der Grenzwert ist nicht eindeutig und es existieren zwei Grenzwerte a und a' mit $a \neq a'$, sodass $\lim\limits_{n\to\infty} a_n = a$ und $\lim\limits_{n\to\infty} a_n = a'$.

Sind die Grenzwerte verschieden, dann gilt $|a - a'|/2 > 0$. Da das Grenzwertkriterium für ein beliebiges $\varepsilon > 0$ gilt, muss es auch für $\varepsilon = |a - a'|/2$ (d.h. $2\varepsilon = |a - a'|$) gelten.

Da sowohl a als auch a' Grenzwerte sind, muss ein n_0 existieren, sodass für alle $n \geq n_0$ gilt $|a_n - a| < \varepsilon$ und $|a_n - a'| < \varepsilon$. Damit gilt aber auch

$$|a - a'| = |a - a_n + a_n - a'| \leq |a_n - a| + |a_n - a'| < 2\varepsilon \text{ für } n \geq n_0.$$

Dies ist eine Widerspruch zu $2\varepsilon = |a - a'|$ (↯).

Damit folgt, dass $a = a'$ gelten muss.

□

Das ε-Kriterium kann man auch negieren und erhält damit das Kriterium für Divergenz: Eine Folge divergiert, wenn für alle $a \in \mathbb{R}$ gilt:

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

Um die Divergenz einer Folge zu beweisen, müssen wir also ein ε finden (die Wahl darf von a abhängen), sodass wir für jede natürliche Zahl n_0 eine größere natürliche Zahl n finden, für die $|a_n - a| \geq \varepsilon$ gilt. Alternativ kann man auch einen Widerspruchsbeweis führen und das ε-Kriterium für Konvergenz zu einem Widerspruch führen. Im Folgenden sehen wir uns diese Varianten an zwei einfachen Beispielen an:

Beispiel 3.8 (Divergenz von Folgen)

- Die Folge (a_n) mit $a_n = n^2$ divergiert.

  Divergenzbeweis mit Divergenzkriterium:
  Wir müssen für alle potentiellen Grenzwerte a das Divergenzkriterium zeigen. Für alle $a \leq 0$ gilt die Divergenzaussage trivial für $\varepsilon = 1$ und $n = n_0$.
  Für alle $a > 0$ kann man mit $\varepsilon = 1/4$ zeigen, dass für beliebige $n_0 \in \mathbb{N}$ ein $n \geq n_0$ existiert, nämlich $n = \max\{n_0, \lceil a + 1/2 \rceil\}$ sodass

$$\begin{aligned} |a_n - a| &\geq \left| \left\lceil a + \frac{1}{2} \right\rceil^2 - a \right| \geq \left| \left(a + \frac{1}{2} \right)^2 - a \right| \\ &= \left| a^2 + a + \frac{1}{4} - a \right| = \left| a^2 + \frac{1}{4} \right| > \frac{1}{4} = \varepsilon. \end{aligned}$$

  ▶ Folge visualisieren

- Die Folge (b_n) mit $b_n = (-1)^n$ divergiert.

  Divergenzbeweis durch Widerspruch:
  Angenommen, die Folge würde gegen ein $b \in \mathbb{R}$ konvergieren, dann gibt es nach Definition auch für $\varepsilon = 1$ ein $n_0 \in \mathbb{N}$, sodass $|b_n - b| < 1$ für alle $n \geq n_0$ gilt. Es gilt aber außerdem

$$\begin{aligned} 2 &= |b_{n+1} - b_n| \\ &= |(b_{n+1} - b) + (b - b_n)| \\ &\leq |b_{n+1} - b| + |b_n - b| \\ &< 1 + 1 = 2, \end{aligned}$$

  also der Widerspruch $2 < 2$ (↯). Somit muss die Annahme, dass (b_n) konvergiert, falsch sein.

  ▶ Folge visualisieren

Die Divergenz ist im ersten Fall relativ offensichtlich. Für solche Fälle können wir einen praktischen Zusammenhang zwischen Divergenz und Nullfolgen herstellen, der manche Divergenzbeweise vereinfacht:

Satz 3.9 (Divergenz inverser Nullfolgen)

Ist die Folge (a_n) eine Nullfolge mit $a_n \neq 0$, dann ist die Folge (b_n) mit $b_n = \frac{1}{a_n}$ divergent.

▼ Beweis

Da (a_n) eine Nullfolge ist, existiert für alle $\varepsilon_a > 0$ ein $n_0 \in \mathbb{N}$ und für alle $n \geq n_0$ gilt:

$$|a_n - 0| = |a_n| < \varepsilon_a.$$

Aus Satz 2.15(g) folgt

$$\left|\frac{1}{a_n}\right| > \frac{1}{\varepsilon_a}. \tag{*}$$

Die Beträge der Folgenglieder von (b_n) werden also beliebig groß.

Zu zeigen ist, dass (b_n) divergiert, also dass gilt

$$\forall b \in \mathbb{R}\ \exists \varepsilon > 0\ \forall n_0 \in \mathbb{N}\ \exists n \geq n_0 :\ |b_n - b| \geq \varepsilon. \tag{**}$$

$(*)$ gilt für alle ε_a, somit auch für $\varepsilon_a = \frac{1}{1+|b|}$ mit beliebigem b. Daher folgt aus $(*)$:

$$|b_n| = \left|\frac{1}{a_n}\right| > 1 + |b|$$

und somit

$$|b_n - b| \overset{S.2.17(e)}{\geq} ||b_n| - |b|| > |1 + |b| - |b|| = 1.$$

Damit haben wir gezeigt, dass die Aussage $(**)$ für $\varepsilon_b = 1$ gilt.

□

Mit diesem Satz können wir in vielen Fällen die Nullfolgeneigenschaft der inversen Folge prüfen, um die Divergenz zu beweisen. In diesen Fällen werden die Folgenglieder einer divergenten Folge beliebig groß (positiv oder negativ). Solche Folgen bezeichnet man als bestimmt divergent.

Definition 3.10 (Bestimmte Divergenz)

Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt *bestimmt divergent* gegen ∞, wenn $b_n = 1/a_n$ eine Nullfolge ist und ein $n_0 \in \mathbb{N}$ existiert, sodass für alle $n \geq n_0$ gilt $a_n > 0$. Wir schreiben

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

Analog definieren wir die bestimmte Divergenz gegen $-\infty$, wenn $b_n = 1/a_n$ eine Nullfolge ist und ein $n_0 \in \mathbb{N}$ existiert, sodass für alle $n \geq n_0$ gilt $a_n < 0$. Wir schreiben

$$\lim_{n\to\infty} a_n = -\infty.$$

Wir könnten zum Beispiel zeigen, dass $a_n = n^2$ bestimmt gegen ∞ divergiert. Betrachtet man die Menge der Folgenglieder ist diese Menge für (a_n) unbeschränkt. Da die Beschränktheit der Menge der Folgenglieder eine wichtige Rolle spielt, überträgt man die Definition auf die Folge selbst:

Definition 3.11 (Beschränkte Folgen)

Wir nennen eine Folge $(a_n)_{n\in\mathbb{N}}$ nach oben (bzw. nach unten) beschränkt, wenn die Menge $M = \{a_n \mid n \in \mathbb{N}\}$ nach oben (bzw. nach unten) beschränkt ist.

Eine sowohl nach oben als auch nach unten beschränkte Folge nennen wir beschränkt.

Für beschränkte Folgen gibt es immer ein $K \in \mathbb{R}$ mit $|a_n| \leq K$. Dies lässt sich einfach aus einer oberen Schranke K_1 und einer unteren Schranke K_2 bestimmen als $K = \max\{|K_1|, |_2|\}$. Machen Sie sich dies selbst an einem Beispiel klar, denn wir werden die Ungleichung $|a_n| \leq K$ nun häufiger in den nächsten Sätzen für beschränkte Folgen benutzen. Starten wir mit einem ersten einfachen Zusammenhang zwischen Beschränktheit und Konvergenz:

Satz 3.12 (Konvergenz ⇒ Beschränktheit)

Jede konvergente Folge ist beschränkt.

▼ Beweis

Sei $\lim\limits_{n\to\infty} a_n = a$ und $n_0 \in \mathbb{N}$ so gewählt, dass $|a_n - a| < 1$ für alle $n \geq n_0$.

Daraus folgt $|a_n| = |a + a_n - a| \leq |a| + |a_n - a| \leq |a| + 1$ für alle $n \geq n_0$.

Setze $K := \max\{|a_1|, |a_2|, \dots, |a_{n_0-1}|, |a| + 1\}$, dann ist $|a_n| \leq K$, bzw. $-K \leq a_n \leq K$ für alle $n \in \mathbb{N}$ und somit ist (a_n) beschränkt.

□

Die Kontraposition dieses Satzes lautet: Jede unbeschränkte Folge ist nicht konvergent (also divergent). Wir können also auch einen Divergenzbeweis führen, indem wir die Unbeschränktheit der Folge zeigen.

Achtung: Die Umkehrung des Satzes gilt nicht, da z. B. $a_n = (-1)^n$ beschränkt, aber nicht konvergent ist. Wir können aber trotzdem von Beschränktheit auf Konvergenz schließen, wenn wir noch eine weitere Eigenschaft der Folge zusätzlich fordern, die Monotonie:

Definition 3.13 (Monotone Folge)

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

- monoton wachsend, falls $a_n \leq a_{n+1}$ für alle $n \in \mathbb{N}$,
- streng monoton wachsend, falls $a_n < a_{n+1}$ für alle $n \in \mathbb{N}$,
- monoton fallend, falls $a_n \geq a_{n+1}$ für alle $n \in \mathbb{N}$,
- streng monoton fallend, falls $a_n > a_{n+1}$ für alle $n \in \mathbb{N}$.

Um die Monotonie für eine gegebene Folge zu zeigen, vergleicht man also die Größe von a_n mit der von a_{n+1}. Bei Folgen mit positiven Folgengliedern ist es auch oft hilfreich, hierzu a_{n+1}/a_n zu betrachten. Für eine monoton wachsende Folge muss der Quotient beispielsweise größer als 1 sein und für eine monoton fallende kleiner als 1.

Stellen wir uns eine beschränkte und gleichzeitig monotone, z.B. wachsende, Folge vor. Hier ist anschaulich klar, dass sich die Folgenglieder einem Grenzwert nähern müssen, da sie immer weiter wachsen, aber das Supremum nicht überschreiten können. Dieser Zusammenhang wird im nachfolgenden Satz bewiesen.

Satz 3.14 (Monotonie + Beschränktheit ⇒ Konvergenz)

Jede beschränkte monotone Folge ist konvergent. Genauer formuliert:

a. Ist $(a_n)_{n\in\mathbb{N}}$ monoton wachsend und nach oben beschränkt, so ist (a_n) konvergent und es gilt $\lim\limits_{n\to\infty} a_n = \sup\{a_n \mid n \in \mathbb{N}\}$.

b. Ist $(a_n)_{n\in\mathbb{N}}$ monoton fallend und nach unten beschränkt, so ist (a_n) konvergent und es gilt $\lim\limits_{n\to\infty} a_n = \inf\{a_n \mid n \in \mathbb{N}\}$.

▼ Beweis

a. Sei $(a_n)_{n\in\mathbb{N}}$ eine monoton wachsende und nach oben beschränkte Folge. Da (a_n) nach oben beschränkt ist, existiert ein Supremum $a \in \mathbb{R}$ (nach dem Vollständigkeitsaxiom). Damit gilt durch die Monotonie

$$a_1 \leq a_2 \leq \ldots \leq a_n \leq a_{n+1} \leq \ldots \leq a.$$

Es bleibt zu zeigen, dass gilt:

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

Sei $\varepsilon > 0$. Da a die kleinste obere Schranke ist, ist $a - \varepsilon$ keine obere Schranke. Somit gibt es ein n_0 mit $a - \varepsilon < a_{n_0}$. Da die Folge monoton ist gilt $a_{n_0} \leq a_n$ für $n \geq n_0$ und da a obere Schranke ist gilt $a_n < a$, sodass

$$|a - a_n| = a - a_n \leq a - a_{n_0} < \varepsilon.$$

Damit ist das Konvergenzkriterium erfüllt.

b. Der Beweis kann analog geführt werden.

□

Dies ist ein besonderes Konvergenzkriterium, da hier der Grenzwert nicht bekannt sein muss. Wir müssen lediglich eine obere (oder untere) Schranke finden (nicht zwangsweise das Supremum/Infimum) und die Monotonie nachweisen. Dies ist für manche Folgen einfacher, als den Grenzwert zu bestimmen.

Beispiel 3.15

Zur Anwendung des letzten Satzes betrachten wir Folge (e_n) mit

$$e_n = \left(1 + \frac{1}{n}\right)^n = \left(\frac{n+1}{n}\right)^n.$$

Diese Folge ist besonders interessant, weil einerseits die Basis $(1 + 1/n)$ offensichtlich gegen 1 konvergiert und $1^n = 1\ \forall n \in \mathbb{N}$ gilt. Allerdings gilt auch $\lim\limits_{n\to\infty} x^n = \infty$ für $x > 1$. Diese beiden Eigenschaften konkurrieren bei dieser Folge gegeneinander und es ist nicht offensichtlich, ob die Folge konvergiert, oder divergiert.

▼ Konvergenzbeweis

Mit Satz 3.14 können wir zeigen, dass (e_n) konvergiert (auch wenn wir dadurch nicht den Grenzwert bestimmen können). Wir zeigen dazu zunächst, dass die Folge monoton wächst, also $\frac{e_{n+1}}{e_n} > 1\ \forall n \in \mathbb{N}$. Damit die Terme etwas übersichtlicher werden, vergleichen wir $\frac{e_n}{e_{n-1}}$ für $n \geq 2$, was aber die identische Aussage liefert:

$$\begin{aligned}
\frac{e_n}{e_{n-1}} &= \frac{\left(\frac{n+1}{n}\right)^n}{\left(\frac{n}{n-1}\right)^{n-1}} \\
&= \left(\frac{n+1}{n}\right)^n \left(\frac{n-1}{n}\right)^{n-1} \\
&= \left(\frac{n+1}{n} \cdot \frac{n-1}{n}\right)^n \left(\frac{n}{n-1}\right) \\
&= \left(\frac{n^2-1}{n^2}\right)^n \left(\frac{n}{n-1}\right) \\
&= \left(1 - \frac{1}{n^2}\right)^n \left(\frac{n}{n-1}\right)
\end{aligned}$$

Für den linken Faktor können wir nun die Bernoullische Ungleichung anwenden (Satz 2.45): $(1+x)^n \geq 1 + nx$. Damit folgt:

$$\begin{aligned}
\frac{e_n}{e_{n-1}} &= \left(1 - \frac{1}{n^2}\right)^n \left(\frac{n}{n-1}\right) \\
&\geq \left(1 - \frac{n}{n^2}\right) \left(\frac{n}{n-1}\right) \\
&= \left(\frac{n^2-n}{n^2}\right) \left(\frac{n}{n-1}\right) = 1
\end{aligned}$$

Damit ist die Folge (e_n) monoton wachsend. Zur Anwendung von Satz 3.14 müssen wir noch die Beschränktheit von (e_n) zeigen.

Dazu betrachten wir die leicht modifizierte Folge $(\bar{e}_n)$ mit $\bar{e}_n = e_n\left(1 + \frac{1}{n}\right) = \left(1 + \frac{1}{n}\right)^{n+1}$. Offensichtlich gilt damit $\bar{e}_n > e_n\ \forall n \in \mathbb{N}$.

Wir werden nun zeigen, dass $\bar{e}_n$ monoton fällt. Somit ist $\bar{e}_1 = 2^2 = 4$ das größte Folgenglied und $e_n < \bar{e}_n < 4$ nach oben beschränkt.

Zeigen wir also noch, dass $\bar{e}_n$ monoton fällt, indem wir diesmal zeigen, dass $\frac{\bar{e}_{n-1}}{\bar{e}_n} > 1 \; \forall n \geq 2$. Die Umformungsschritte sind sehr ähnlich zu den letzten, daher hier nur die Kurzform:

$$\begin{aligned}
\frac{\bar{e}_{n-1}}{\bar{e}_n} &= \frac{\left(1+\frac{1}{n-1}\right)^n}{\left(1+\frac{1}{n}\right)^{n+1}} \\
&= \left(1+\frac{n}{n^2-1}\right)^n \left(1+\frac{1}{n+1}\right) \\
&\geq \left(1+\frac{n^2}{n^2-1}\right)\left(1+\frac{1}{n+1}\right) \\
&> \left(1+\frac{1}{n}\right)\left(1+\frac{1}{n+1}\right) = 1
\end{aligned}$$

Die letzte Abschätzung nutzt aus, dass $n^2 - 1 < n^2$ und damit $\frac{n}{n^2-1} > \frac{n}{n^2} = \frac{1}{n}$ ist.

So kann die Beschränktheit und Monotonie von (e_n) (und übrigens auch $(\bar{e}_n)$) gezeigt werden und damit folgt nach Satz 3.14, dass beide Folgen konvergieren. Die Grenzwerte kennen wir allerdings noch nicht. Aufgrund der Monotonie-Eigenschaften müssen diese aber auf jeden Fall zwischen $e_1 = 2$ und $\bar{e}_1 = 4$ liegen. Wir werden später noch zeigen, dass beide Folgen sogar gegen den gleichen Grenzwert konvergieren, der als die irrationale eulersche Zahl

$$e = 2.718281828459045235360287471352662497757247093699959574966967627 7 \ldots$$

bekannt ist.

▶ Folge visualisieren

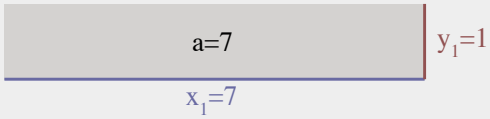
Der Satz eignet sich ebenfalls, um die Konvergenz von rekursiven Folgen zu beweisen. Dies war mit den bisherigen Methoden eher schwierig zu zeigen, da wir für rekursive Folgen im Allgemeinen keine direkte Formel zur Berechnung des n-ten Folgenglieds zur Verfügung haben.

Beispiel 3.16 (Konvergenz der Wurzelfolge)

Wir betrachten die rekursive Folge (x_n), die als *babylonisches Wurzelziehen* oder als *Heron-Formel* bekannt ist:

$$x_1 = a$$
$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right)$$

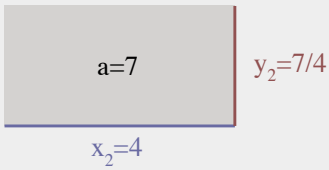
Diese Folge kann man geometrisch so interpretieren, dass wir ein Rechteck der Fläche a nach und nach in ein Quadrat gleichen Flächeninhalts umwandeln. Das Startrechteck hat dabei die Seitenlängen $x_1 = a$ und $y_1 = a/x_1 = 1$. Offensichtlich ist der Flächeninhalt a. Im folgenden Bild ist das Rechteck für $a = 7$ gezeigt:



Da es sich noch um kein Quadrat handelt, muss eine Seite des Rechtecks zu groß und die andere zu klein sein. Also bestimmen wir den Mittelwert aus den beiden Seitenlängen

$$x_2 = \frac{1}{2}\left(x_1 + \frac{a}{x_1}\right)$$

Im Beispiel von $a = 7$ ergibt sich somit $x_2 = 4$. Dies ist unsere zweite Schätzung für eine Rechteckseite. Damit der Flächeninhalt bei $a = 7$ bleibt, muss die zweite Seite die Länge $y_2 = a/x_2 = 7/4$ haben.



Wieder stellen wir fest, dass eine Seite größer als die andere ist. Daher bestimmen wir erneut den Mittelwert und erhalten damit $x_3 = 23/8 = 2.875$ und $y_3 = 56/23 \approx 2.435$.

Die Frage ist nun, ob diese Folge konvergiert. Intuitiv ist klar, dass der Grenzwert $\sqrt{a}$ sein muss, da wir uns mit jeden Schritt stärker einem Quadrat annähern. Der Konvergenzbeweis erfolgt wieder über Satz 3.14: Wir zeigen also, dass (x_n) nach unten beschränkt und monoton fallend ist.

Beschränktheit:
Wir zeigen, dass (x_n) nach unten beschränkt ist durch $\sqrt{a}$, also

$$x_n \geq \sqrt{a} \quad \Leftrightarrow \quad x_n^2 - a \geq 0,$$

denn $x_1 = a \geq \sqrt{a}$ und für $n > 1$

$$x_n^2 - a = \frac{1}{4}\left(x_{n-1} + \frac{a}{x_{n-1}}\right)^2 - a = \frac{1}{4}\left(x_{n-1} - \frac{a}{x_{n-1}}\right)^2 \geq 0$$

Monotonie:
Die Folge (x_n) ist monoton fallend, denn es gilt:

$$x_{n+1} - x_n = \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) - x_n = \frac{a}{2x_n} - \frac{x_n}{2} = \frac{a - x_n^2}{2x_n} \leq 0.$$

Damit muss nach Satz 3.14 die Folge (x_n) konvergieren. Wir werden nach dem nächsten Satz noch einmal auf diese Folge zurückkommen und beweisen, dass der Grenzwert $\sqrt{a}$ ist, wie die geometrische Interpretation schon vermuten lässt.

Viele Folgen lassen sich aus anderen Folgen zusammensetzen. Betrachten wir z.B. die Folge (a_n) mit $a_n = \frac{n+1}{n} = 1 + \frac{1}{n}$: Es liegt nahe, dass sich die Grenzwerte von $b_n = 1$ und $c_n = \frac{1}{n}$ addieren, und damit den Grenzwert $a = 1 + 0 = 1$ für (a_n) bilden. Solche Regeln für Kombinationen von Grenzwerten führen wir nun ein.

Definition 3.17 (Kombinierte Folgen)

Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ zwei konvergente Folgen und $c \in \mathbb{R}$. Dann definieren wir:

a. $(a_n)_{n\in\mathbb{N}} + (b_n)_{n\in\mathbb{N}} = (a_n + b_n)_{n\in\mathbb{N}}$

b. $c \cdot (a_n)_{n\in\mathbb{N}} = (c \cdot a_n)_{n\in\mathbb{N}}$

c. $(a_n)_{n\in\mathbb{N}} \cdot (b_n)_{n\in\mathbb{N}} = (a_n \cdot b_n)_{n\in\mathbb{N}}$

d. $\dfrac{(a_n)_{n\in\mathbb{N}}}{(b_n)_{n\in\mathbb{N}}} = \left(\dfrac{a_n}{b_n}\right)_{n\in\mathbb{N}}$ falls $b_n \neq 0$ für alle $n \in \mathbb{N}$

Satz 3.18 (Grenzwerte kombinierter Folgen)

Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ zwei konvergente Folgen und $c \in \mathbb{R}$. Dann gilt :

a. $\lim\limits_{n\to\infty} (a_n + b_n) = \lim\limits_{n\to\infty} (a_n) + \lim\limits_{n\to\infty} (b_n)$

b. $\lim\limits_{n\to\infty} c \cdot (a_n) = c \cdot \lim\limits_{n\to\infty} (a_n)$

c. $\lim\limits_{n\to\infty} (a_n) \cdot (b_n) = \lim\limits_{n\to\infty} (a_n) \cdot \lim\limits_{n\to\infty} (b_n)$

d. $\lim\limits_{n\to\infty} \frac{(a_n)}{(b_n)} = \frac{\lim\limits_{n\to\infty} (a_n)}{\lim\limits_{n\to\infty} (b_n)}$ falls $b_n \neq 0$ für alle $n \in \mathbb{N}$ und $\lim\limits_{n\to\infty} b_n \neq 0$.

▼ Beweis

Sei $a = \lim\limits_{n\to\infty} a_n$ und $b = \lim\limits_{n\to\infty} b_n$.

a. Sei $\varepsilon > 0$ vorgegeben, dann ist auch $\varepsilon/2 > 0$. Da beide Folgen konvergent sind, gibt es $n_a, n_b \in \mathbb{N}$ mit $|a_n - a| < \frac{\varepsilon}{2}$ für $n \geq n_a$ sowie $|b_n - b| < \frac{\varepsilon}{2}$ für $n \geq n_b$. Damit gilt für $n \geq \max(n_a, n_b)$

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

b. Ergibt sich aus Fall **(c)**, wenn wir $b_n = c$ als konstante Folge auffassen.

c. Nach Satz 3.12 sind die Folgen (a_n) und (b_n) beschränkt. Also gibt es $K_a, K_b \in \mathbb{R}$ mit $K_a \geq |a_n|$ und $K_b \geq |b_n|$ für alle $n \in \mathbb{N}$. Wir definieren $K := \max\{K_a, K_b, 1\}$. Sei nun $\varepsilon > 0$ vorgegeben. Da auch $\frac{\varepsilon}{2K} > 0$ und da beide Folgen konvergieren, gibt es $n_a, n_b \in \mathbb{N}$ sodass

$$|a_n - a| < \frac{\varepsilon}{2K} \quad \text{für} \quad n \geq n_a,$$
$$|b_n - b| < \frac{\varepsilon}{2K} \quad \text{für} \quad n \geq n_b.$$

Dann gilt für alle $n \geq \max(n_a, n_b)$

$$\begin{aligned} |a_n b_n - ab| &= |a_n b_n - a_n b + a_n b - ab| \\ &= |a_n \cdot (b_n - b) + (a_n - a) \cdot b| \\ &\leq |a_n| \cdot |b_n - b| + |a_n - a| \cdot |b| \\ &< K \cdot \frac{\varepsilon}{2K} + \frac{\varepsilon}{2K} \cdot K \\ &= \varepsilon \end{aligned}$$

d. Da $\frac{a_n}{b_n} = a_n \cdot \frac{1}{b_n}$, können wir den Beweis auf Fall **(c)** zurückführen, falls $\frac{1}{b_n} \to \frac{1}{b}$ für $n \to \infty$ und $b \neq 0$.

Wegen $b \neq 0$ ist $\frac{|b|}{2} > 0$ und es gibt ein $n_b \in \mathbb{N}$, sodass für alle $n \geq n_b$ gilt $|b_n - b| < \frac{|b|}{2}$. Daraus folgt $|b_n| > \frac{|b|}{2}$.

Zu einem vorgegebenen ε gibt es ein $n_0 \in \mathbb{N}$, sodass $|b_n - b| < \frac{\varepsilon |b|^2}{2}$ für alle $n \geq n_0$. Dann gilt auch für $n \geq \max(n_b, n_0)$

$$\left| \frac{1}{b_n} - \frac{1}{b} \right| = \frac{|b_n - b|}{|b_n| \cdot |b|} = \frac{1}{|b_n| \cdot |b|} \cdot |b_n - b| < \frac{2}{|b|^2} \cdot \frac{\varepsilon |b^2|}{2} = \varepsilon$$

Da $|b_n| \cdot |b| > |b| \frac{|b|}{2}$, ist $(|b_n| \cdot |b|)^{-1} < \frac{2}{|b|^2}$ und die im vorletzten Schritt durchgeführte Ersetzung ist korrekt.

Damit wurde $\lim_{n\to\infty} \frac{1}{b_n} = \frac{1}{b}$ gezeigt und die allgemeinere Aussage in **(d)** kann somit auf **(c)** zurückgeführt werden.

□

Beispiel 3.19

Sei $a_n = \dfrac{3n^2 + 13n}{n^2 - 2}$ mit $n \in \mathbb{N}$. Wegen

$$a_n = \frac{3n^2 + 13n}{n^2 - 2} = \frac{n^2 \left(3 + \frac{13}{n}\right)}{n^2 \left(1 - \frac{2}{n^2}\right)} = \frac{3 + \frac{13}{n}}{1 - \frac{2}{n^2}}$$

und da $\lim_{n\to\infty} \dfrac{1}{n} = 0$, folgt aus Satz 3.18(c), dass auch

$$\lim_{n\to\infty} \frac{1}{n^2} = \lim_{n\to\infty} \frac{1}{n} \cdot \frac{1}{n} = \left(\lim_{n\to\infty} \frac{1}{n}\right) \cdot \left(\lim_{n\to\infty} \frac{1}{n}\right) = 0 \cdot 0 = 0.$$

Damit gilt nach Satz 3.18(a)

$$\lim_{n\to\infty} \left(3 + \frac{13}{n}\right) = 3$$

sowie

$$\lim_{n\to\infty} \left(1 - \frac{2}{n^2}\right) = 1$$

und schließlich nach Satz 3.18(d)

$$\lim_{n\to\infty} \frac{3 + \frac{13}{n}}{1 - \frac{2}{n^2}} = \frac{\lim\limits_{n\to\infty} \left(3 + \frac{13}{n}\right)}{\lim\limits_{n\to\infty} \left(1 - \frac{2}{n^2}\right)} = \frac{3}{1} = 3.$$

▶ Folge visualisieren

Achtung: Die Rechenregeln aus Satz 3.18 sind zwar einfach auch auf drei oder mehr Einzelfolgen übertragbar, aber nicht, wenn es sich um n Summanden/Faktoren handelt. Sonst könnte man z.B. folgern, dass

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

gegen 0 konvergiert, was offensichtlich falsch ist. Genauso könnte man zu dem falschen Schluss kommen, dass die Folge

$$e_n = \left(1 + \frac{1}{n}\right)^n = \left(1 + \frac{1}{n}\right)\left(1 + \frac{1}{n}\right) ... \left(1 + \frac{1}{n}\right)$$

gegen 1 konvergiert. Die Anzahl an Summanden/Faktoren muss für jedes Folgenglied identisch sein, damit Satz 3.18 angewendet werden kann.

Mit den Rechenregeln aus Satz 3.18 kann man viele Folgengrenzwerte bestimmen, indem man sie in Teilfolgen zerlegt. Aber der Satz ist noch deutlich vielseitiger. Wir können damit zeigen, dass zwei Grenzwerte identisch sind, denn wenn

$$\lim_{n\to\infty} (a_n - b_n) = 0$$

gilt und sowohl (a_n) als auch (b_n) konvergieren, folgt aus Satzteil (a), dass

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n.$$

Beispiel 3.20 (identischer Grenzwert)

Wir haben in Beispiel 3.5 die zwei Folgen (e_n) und $(\bar{e}_n)$ betrachtet, mit

$$e_n = \left(1 + \frac{1}{n}\right)^n \qquad \bar{e}_n = \left(1 + \frac{1}{n}\right)^{n+1}.$$

Wir haben bereits bewiesen, dass beide Folgen konvergieren. Nun werden wir zeigen, dass beide Grenzwerte identisch sind, dazu betrachten wir

$$\begin{aligned} e_n - \bar{e}_n &= \left(1 + \frac{1}{n}\right)^n - \left(1 + \frac{1}{n}\right)^{n+1} \\ &= \left(1 + \frac{1}{n}\right)^n \left(1 - \left(1 + \frac{1}{n}\right)\right) \\ &= e_n \left(1 - \left(1 + \frac{1}{n}\right)\right) \\ &= e_n \left(-\frac{1}{n}\right). \end{aligned}$$

Wir können nun mit Satz 3.18 den Grenzwert des rechten Terms bestimmen, indem wir die Grenzwerte der beiden Faktoren bestimmen. Dafür müssen beide Faktoren konvergieren. (e_n) konvergiert nach Beispiel 3.5. Der rechte Faktor konvergiert gegen 0. Also gilt auch

$$\lim_{n\to\infty} (e_n - \bar{e}_n) = 0 \overset{S.3.11}{\Leftrightarrow} \lim_{n\to\infty} e_n = \lim_{n\to\infty} \bar{e}_n.$$

▶ Folgen visualisieren

Darüber hinaus lassen sich auch Grenzwerte rekursiver Folgen mit Satz 3.18 bestimmen, was mit unseren bisherigen Sätzen eher schwierig war. Für konvergente rekursive Folgen können wir ausnutzen, dass

$$\lim_{n\to\infty} a_{n+1} = \lim_{n\to\infty} a_n.$$

Der Beweis ist recht trivial und kann als Übung durchgeführt werden. Anschließend setzten wir links die Rekursionsformeln ein und stellen dann mithilfe von Satz 3.18 nach $\lim_{n\to\infty} a_n$ um. Das setzt natürlich voraus, dass $\lim_{n\to\infty} a_n$ existiert, was beispielsweise mit Satz 3.14 zuvor gezeigt werden kann.

Beispiel 3.21 (Grenzwert der Wurzelfolge)

Betrachten wir erneut die Wurzelfolge (x_n) aus Beispiel 3.16 mit

$$\begin{aligned} x_1 &= a \\ x_{n+1} &= \frac{1}{2}\left(x_n + \frac{a}{x_n}\right). \end{aligned}$$

Wir hatten in Beispiel 3.16 bereits die Konvergenz von (x_n) bewiesen, nun bestimmen wir den Grenzwert. Da wir bereits wissen, dass x_n konvergiert definieren wir

$$x := \lim_{n\to\infty} x_n$$

$$\begin{aligned} & \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} x_n \\ \Leftrightarrow \quad & \lim_{n\to\infty} \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) = \lim_{n\to\infty} x_n \\ \overset{S.3.11}{\Leftrightarrow} \quad & \frac{1}{2}\left(\lim_{n\to\infty}(x_n) + \frac{a}{\lim\limits_{n\to\infty}(x_n)}\right) = \lim_{n\to\infty} x_n \\ \Leftrightarrow \quad & \frac{1}{2}\left(x + \frac{a}{x}\right) = x \\ \Leftrightarrow \quad & x^2 + a = 2x^2 \\ \Leftrightarrow \quad & x^2 = a \\ \Leftrightarrow \quad & x = \pm\sqrt{a}. \end{aligned}$$

Es lässt sich leicht zeigen, dass die Folgenglieder immer positiv sind und für den Grenzwert daher $x = \sqrt{a}$ gilt.

Für das nächste nützliche Hilfsmittel benötigen wir eine Möglichkeit, Folgen und ihre Grenzwerte der Größe nach zu ordnen.

Satz 3.22 (Größenvergleich konvergenter Folgen)

Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ zwei konvergente Folgen mit $a_n \leq b_n$. Dann gilt auch $\lim\limits_{n\to\infty} a_n \leq \lim\limits_{n\to\infty} b_n$.

▼ Beweis

Sei $c_n = b_n - a_n$. Es genügt zu zeigen $\lim\limits_{n\to\infty} c_n \geq 0$, da c_n nach Satz 3.18 konvergent ist und $c_n \geq 0$ für alle $n \in \mathbb{N}$ nach Voraussetzung gilt. Angenommen $\lim\limits_{n\to\infty} c_n = -\varepsilon$ für $\varepsilon > 0$. Dann gäbe es ein n_0, sodass für alle $n \geq n_0$

$$|c_n - (-\varepsilon)| = |c_n + \varepsilon| < \varepsilon.$$

Dies würde bedeuten, dass $c_n < 0$, was nach Annahme nicht gelten kann (↯).

□

Vorsicht: Aus $a_n < b_n$ für alle $n \in \mathbb{N}$ folgt *nicht* $\lim\limits_{n\to\infty} a_n < \lim\limits_{n\to\infty} b_n$ wie zum Beispiel $a_n = 0$ und $b_n = \frac{1}{n}$ zeigt. Das $\leq$ im obigen Satz ist also sehr wichtig. Mit dem vorigen Satz kann das folgende Sandwich-Theorem bewiesen werden:

Satz 3.23 (Sandwich-Theorem)

Seien $(a_n)_{n\in\mathbb{N}}$, $(b_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ Folgen, für die ein n_0 existiert, sodass für alle $n \geq n_0$ gilt

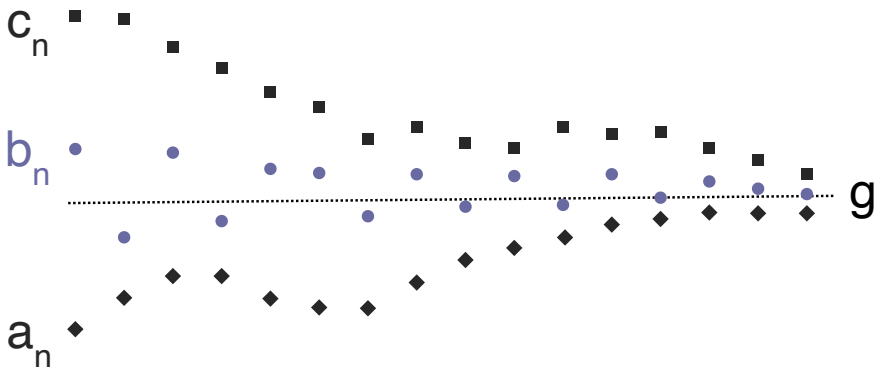
$$a_n \leq b_n \leq c_n.$$

Wenn $(a_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ konvergent sind und gegen den selben Grenzwert konvergieren, dann ist auch $(b_n)_{n\in\mathbb{N}}$ konvergent und es gilt:

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \lim_{n\to\infty} c_n.$$

▼ Beweis zur Übung

Anschaulich quetschen wir damit eine Folge zwischen zwei anderen Folgen ein, die sich beide auf den gleichen Grenzwert g zubewegen:



Quelle: Who2010, CC BY-SA 4.0, Wikimedia Commons

Beispiel 3.24 (n-te Wurzel aus n)

Wir wollen zeigen, dass für die Folge (a_n) mit $a_n = \sqrt[n]{n}$ gilt

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

Anmerkung: Für den Beweis nutzen wir aus, dass aus Satz 2.35 folgt: $a < b \Leftrightarrow \sqrt[n]{a} < \sqrt[n]{b} \quad (*)$

Der Grenzwert kann mit dem Sandwich-Theorem nachgewiesen werden. Wir benötigen also zwei Folgen, von denen die eine kleinere Folgenglieder als (a_n) und die anderen größere Folgenglieder als (a_n) besitzt.

Für die untere Folge wählen wir (u_n) mit $u_n = 1$, denn es gilt $1^n = 1 \leq n \overset{(*)}{\Leftrightarrow} 1 \leq \sqrt[n]{n}$.

Für die obere Folge wählen wir (o_n) mit

$$o_n = 1 + 2\frac{1}{\sqrt{n}}$$

Hier ist es nicht offensichtlich, warum $a_n \leq o_n$ gelten muss. Daher zeigen wir dies separat. Dazu nutzen wir eine Ungleichgung aus Kapitel 2.7 (Satz 2.46). Nach dieser gilt für $x \geq 0$ und $n \geq 2$ die Abschätzung

$$(1+x)^n > \frac{n^2x^2}{4}.$$

Mit $x = \frac{2}{\sqrt{n}}$ folgt daraus

$$\left(1 + \frac{2}{\sqrt{n}}\right)^n > \frac{n^2\left(\frac{2}{\sqrt{n}}\right)^2}{4} = n$$

$$\overset{(*)}{\Leftrightarrow} \quad \sqrt[n]{\left(1 + \frac{2}{\sqrt{n}}\right)^n} = 1 + \frac{2}{\sqrt{n}} = o_n > \sqrt[n]{n} = a_n.$$

Damit konnten wir zeigen, dass $o_n > a_n$ gilt für $n \geq 2$. Dass die Abschätzung erst ab $n = 2$ gilt, ist unerheblich, da endlich viele Folgenglieder den Grenzwert nicht ändern. Der Grenzwert von o_n ist ebenfalls 1 (Beweis zur Übung).

Damit gilt insgesamt

$$1 = \lim_{n\to\infty} u_n \leq \lim_{n\to\infty} a_n \leq \lim_{n\to\infty} o_n = 1.$$

Daraus folgt mit dem Sandwich-Theorem

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

Hiermit lässt sich ebenfalls leicht zeigen, dass für alle $k \geq 1$ ebenfalls gilt

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

In diesem Abschnitt haben wir den Großteil der Sätze vorgestellt, mit deren Hilfe wir die Konvergenz und Divergenz von Folgen nachweisen können. Zusammenfassend geben wir hier noch einmal eine kleine Übersicht der Sätze an. Wir geben jeweils mit an, ob man für den jeweiligen Konvergenzbeweis den Grenzwert kennen muss.

Satz/Def.	Konvergenz	Grenzwert?	Divergenz
Definition 3.5	✓	ja	✓
Satz 3.9	✗	-	✓
Satz 3.12	✗	-	✓
Satz 3.14	✓	nein	✗
Satz 3.18	✓	nein	✗
Satz 3.23	✓	ja	✗
Satz 3.27	✗	-	✓
Satz 3.34	✓	nein	✓

Die Tabelle enthält zur Vollständigkeit auch zwei Kriterien, die wir erst in den nächsten beiden Abschnitten formulieren werden.

3.2 Teilfolgen

Wir haben bereits ein paar Eigenschaften von beschränkten Folgen kennen gelernt. In diesem Abschnitt wollen wir den sogenannten Satz von Bolzano-Weierstraß beweisen. Dieser besagt, dass wir in jeder beschränkten Folge unendlich viele Folgenglieder so auswählen können, dass diese eine konvergente Folge ergeben. Also z.B. für $a_n = (-1)^n$ könnten wir alle geraden oder alle ungeraden Folgenglieder auswählen. Der Satz mag auf den ersten Blick nicht sonderlich nützlich erscheinen, aber wir werden ihn in späteren Beweisen immer mal wieder benötigen und die Aussage an sich ist interessant. Definieren wir uns zunächst den Begriff der Teilfolge.

Definition 3.25 (Teilfolge)

Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge und $n_1 < n_2 < n_3 < \ldots$ eine aufsteigende unendliche Folge natürlicher Zahlen, dann heißt $(a_{n_k})_{k\in\mathbb{N}} = a_{n_1}, a_{n_2}, \ldots$ eine Teilfolge der Folge $(a_n)_{n\in\mathbb{N}}$.

Für eine Teilfolge werden also nur einzelne Glieder der Folge berücksichtigt, ohne deren Reihenfolge zu ändern. Offensichtlich gilt, dass eine Teilfolge auch konvergent ist, wenn die ursprüngliche Folge schon konvergent ist, da das ε-Kriterium direkt übertragbar ist.

Satz 3.26

Jede Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ einer konvergenten Folge $(a_n)_{n\in\mathbb{N}}$ ist konvergent und es gilt

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

▼ Beweis

$$\forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall n > n_0 :\ |a_n - a_{n_0}| < \varepsilon.$$

Damit liegen nur endlich viele Glieder von (a_n) außerhalb von $(a-\varepsilon, a+\varepsilon)$ und damit auch nur endlich viele Glieder von a_{n_k}. Da jedes a_{n_k} auch ein a_n ist, gibt es ein $k_0 \leq n_0$, so dass $|a - a_{n_k}| < \varepsilon$ für $k \geq k_0$.

□

Interessanter wird es, wenn die ursprüngliche Folge nicht konvergent ist. Teilfolgen können hier dabei helfen, die Divergenz der ursprünglichen Folge zu beweisen.

Satz 3.27 (Divergenznachweis durch Teilfolgen)

Besitzt eine Folge $(a_n)_{n\in\mathbb{N}}$

a. eine divergente Teilfolge oder

b. zwei konvergente Teilfolgen $(a_{n_k})_{k\in\mathbb{N}}$ und $(a_{n_l})_{l\in\mathbb{N}}$ mit $\lim\limits_{k\to\infty}(a_{n_k}) \neq \lim\limits_{l\to\infty}(a_{n_l})$,

so ist die Folge divergent.

▼ Beweis

Wir schreiben Satz 3.26 als Implikation auf:

Folge konvergiert $\Rightarrow$ Alle Teilfolgen konvergieren gegen den selben Grenzwert

Wenn wir nun die Kontraposition formulieren (zur Erinnerung: $(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$ und ein negiertes "und" wird zu einem "oder"), dann ergibt sich:

Nicht alle Teilfolgen konvergieren oder haben nicht den selben Grenzwert $\Rightarrow$ Folge divergiert

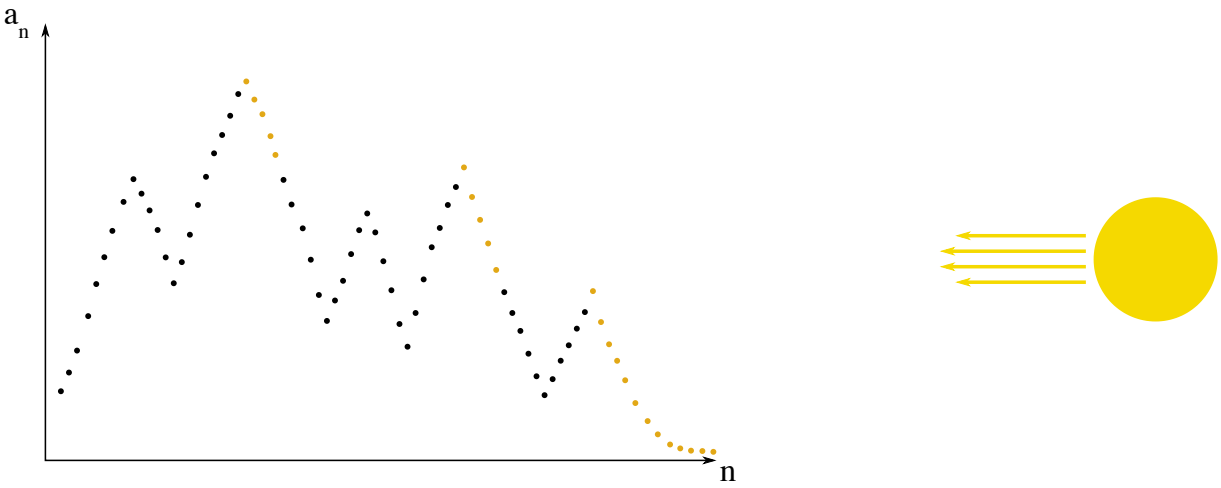
Die Aussage dieses Satzes ist also identisch mit der des letzten Satzes, es handelt sich lediglich dessen Kontraposition.

□

Beispiel 3.28

Sei $a_n = (-1)^n$ die untersuchte Folge. Es gibt zwei Teilfolgen $(-1)^{2n}$ und $(-1)^{2n+1}$ mit $\lim\limits_{n\to\infty}(-1)^{2n} = 1$ und $\lim\limits_{n\to\infty}(-1)^{2n+1} = -1$, daher ist Folge a_n divergent.

Der nächste Satz besagt, dass wir für jede beliebige Folge eine Teilfolge finden können, die entweder monoton wächst oder monoton fällt. Um den Beweis zu verstehen, nutzen wir sogenannte *Gipfelpunkte* einer Folge. Das sind Folgenglieder, die größer sind als alle nachfolgenden Folgenglieder. Betrachten wir dazu das folgende graphische Beispiel:



Es wird anschaulich klar, warum die gelben Punkte Gipfelpunkte genannt werden: Stellen wir uns die Folge wie die Silhouette eines Berges vor, auf welche die Sonne von rechts aus parallel zum Boden scheint, dann sind Gipfelpunkte solche Punkte der Bergkette, die im Sonnenlicht liegen (weil sie höher liegen, als alle Punkte rechts von ihnen). Mit diesen Punkten kann der folgende Satz sehr anschaulich bewiesen werden.

Satz 3.29

Jede Folge enthält eine monotone Teilfolge.

▼ Beweis

Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge. Wir betrachten die Menge der Gipfelpunkte in a_n $G = \{a_{n_1}, a_{n_2}, a_{n_3}, ...\}$ wobei $n_1 > n_2 > n_3 > ...$ und für alle $n > n_k$ gilt, dass $a_n \geq a_{n_k}$. Für die Menge G gibt es nun zwei Möglichkeiten:

1. G enthält unendlich viele Elemente.
   Dann bilden die Gilfelpunkte eine monoton fallende Teilfolge, da per Konstruktion gilt $a_{n_1} \geq a_{n_2} \geq a_{n_3} \geq ...$

2. G enthält nur endlich viele Elemente oder ist leer.
   Hier können wir eine rekursive Konstruktionsvorschrift für eine monoton steigende Teilfolge angeben:

   Sei l der Folgengliedindex des letzten Gipfelpunktes (oder 0, wenn G leer ist).
   Wir wählen $m_1 = l + 1$.
   Da a_{m_1} kein Gipfelpunkt ist, muss es einen Index m_2 geben, mit $a_{m_2} > a_{m_1}$.
   Da a_{m_2} auch kein Gipfelpunkt ist, muss es einen Index m_3 geben, mit $a_{m_3} > a_{m_2}$.
   Da a_{m_3} auch kein Gipfelpunkt ist, muss es einen Index m_4 geben, mit $a_{m_4} > a_{m_3}$.

   Dies lässt sich unendlich fortführen, und so entsteht eine (streng) monotone wachsende Teilfolge mit

   $$a_{m_1} < a_{m_2} < a_{m_3} < ...$$

□

Nun haben wir alle Voraussetzungen um den Satz von Bolzano-Weierstraß zu beweisen. Wie bereits angedeutet, hilft uns der Satz auf den ersten Blick nicht um Grenzwerte von Folgen zu bestimmen, aber wir werden ihn in späteren Kapiteln noch häufiger benötigen.

Satz 3.30 (Satz von Bolzano-Weierstraß)

Jede beschränkte Folge $(a_n)_{n\in\mathbb{N}}$ besitzt eine konvergente Teilfolge.

▼ Beweis

Nach Satz 3.29 enthält die Folge eine monotone Teilfolge (die natürlich auch beschränkt ist, wenn die ursprüngliche Folge beschränkt ist) und nach Satz 3.14 ist jede beschränkte monotone Folge konvergent.

□

Wir haben bereits festgestellt, dass Teilfolgen von divergenten Folgen gegen unterschiedliche Grenzwerte konvergieren können. Um hier die Benennung deutlicher zu machen, nennen wir Grenzwerte von Teilfolgen *Häufungspunkte*.

Definition 3.31 (Häufungspunkt)

Für eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt a *Häufungspunkt*, wenn es eine Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ von $(a_n)_{n\in\mathbb{N}}$ gibt und $\lim\limits_{k\to\infty} a_{n_k} = a$.

Der Satz von Bolzano-Weierstraß kann also auch so formuliert werden, dass jede beschränkte Folge mindestens einen Häufungspunkt besitzt. Genauso bedeuten die Sätze 3.26 und 3.27, dass eine konvergente Folge genau einen Häufungspunkt besitzt, wohingegen Folgen mit mehreren Häufungspunkten divergent sind. Die Umkehrung dieser Formulierung gilt nicht. Das heißt, auch wenn genau ein Häufungspunkt existiert, muss die Folge nicht zwingend konvergieren, wie wir im nachfolgenden Beispiel zeigen.

Beispiel 3.32

- Die Folge

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

  ist unbeschränkt, divergent und besitzt den Häufungspunkt 0.

- Die Folge $a_n = (-1)^n + \frac{1}{n}$ ist beschränkt und besitzt die Häufungspunkte -1 und 1.

3.3 Alternative Vollständigkeitsaxiome

Betrachten wir noch einmal eine der Mengen, mit denen wir Suprema und damit unser Vollständigkeitsaxiom motiviert hatten:

$$M_3 = \left\{ \frac{n}{n+1} \;\middle|\; \forall n \in \mathbb{N} \right\} = \left\{ \frac{1}{2}, \frac{2}{3}, \frac{4}{5}, \frac{5}{6}, \frac{6}{7}, \frac{7}{8}, \dots \right\}.$$

Mit dem Wissen aus diesem Kapitel würden Sie die Menge nun vermutlich mit einer Folge assoziieren, deren Grenzwert das Supremum von M_3 ist, und könnten diesen leicht bestimmen:

$$a_n = \frac{n}{n+1} = \frac{1}{1+\frac{1}{n}} \quad \Rightarrow \quad \lim_{n\to\infty} a_n = \frac{1}{1+0} = 1.$$

Die Existenz eindeutiger Suprema für beschränkte Teilmengen von $\mathbb{R}$ hatten wir in Kapitel 2.2 axiomatisch gefordert und daraus folgte, dass die reellen Zahlen (nach Definition 2.8) vollständig sind. Wir hatten aber auch bereits darauf hingewiesen, dass verschiedene Formulierungen für die Vollständigkeit eines Körpers existieren, welche als Basis für das Vollständigkeitsaxiom dienen können.

In diesem Kapitel werden wir uns mit zwei häufig genutzen Alternativen beschäftigen, die beide mit Folgen zusammenhängen. In Satz 2.21 hatten wir aus dem Vollständigkeitsaxiom die Existenz von Quadratwurzeln gefolgert, diesen Beweis werden wir hier noch einmal analog für beide Alternativen führen.

Cauchy-Folgen

Die erste Variante der Vollständigkeit bedient sich der sogenannten *Cauchy-Folgen*. Diese spielen darüber hinaus eine besondere Rolle für den Konvergenznachweis, wenn der explizite Grenzwert unbekannt ist.

Definition 3.33 (Cauchy-Folge)

Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt *Cauchy-Folge*, wenn gilt:

$$\forall \varepsilon > 0 \; \exists n_0 \in \mathbb{N} \; \forall n > n_0 : \; |a_n - a_{n_0}| < \varepsilon.$$

Die Definition ist ganz ähnlich zu dem ε-Kriterium für die Folgenkonvergenz, allerdings vergleichen wir hier nicht die Folgenglieder mit dem Grenzwert, sondern mit einem anderen Folgenglied a_{n_0}. Anders formuliert, gilt für Cauchy-Folgen, dass man immer ein Glied a_{n_0} findet, ab dem die Folgenglieder nicht mehr als ε voneinander abweichen, sich also beliebig nahe kommen. Damit liegt die intuitive Vermutung nahe, dass Cauchy-Folgen immer konvergieren, was wir im folgenden Satz beweisen:

Satz 3.34

Eine Folge ist genau dann konvergent, wenn sie eine Cauchy-Folge ist. Das bedeutet:

a. Jede konvergente Folge ist eine Cauchy-Folge.

b. Jede Cauchy-Folge ist konvergent.

▼ Beweis

a. Sei $(a_n)_{n\in\mathbb{N}}$ konvergent, sei $\varepsilon > 0$ und sei $a = \lim_{n\to\infty} a_n$. Definiere $\varepsilon' = \frac{\varepsilon}{2}$. Nach Definition existiert n_0, sodass $|a_n - a| < \varepsilon'$ für alle $n \geq n_0$ und damit

$$|a_n - a_{n_0}| = |a_n - a + a - a_{n_0}| \leq |a_n - a| + |a_{n_0} - a| < \varepsilon' + \varepsilon' = \varepsilon.$$

Damit ist jede konvergente Folge eine Cauchy-Folge.

b. Wir zeigen zunächst, dass jede Cauchy-Folge beschränkt ist: Sei $(a_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge. Dann gibt es zu $\varepsilon = 1$ ein n_0, sodass $|a_n - a_{n_0}| < 1$ für alle $n \geq n_0$. Daraus folgt für $n \geq n_0$

$$|a_n| = |a_{n_0} + (a_n - a_{n_0})| \leq |a_{n_0}| + |a_n - a_{n_0}| < |a_{n_0}| + 1.$$

Wähle $K = \max\{|a_1|, \dots, |a_{n_0-1}|, |a_{n_0}| + 1\}$, so gilt $|a_n| \leq K$ oder anders formuliert: $-K \leq a_n \leq K\ \forall n \in \mathbb{N}$. Damit ist jede Cauchy-Folge beschränkt und hat nach Satz 3.30 eine konvergente Teilfolge.

Als nächstes zeigen wir, dass die Cauchy-Folge gegen den Grenzwert dieser Teilfolge konvergieren muss: Sei $(a_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge und $(a_{n_k})_{k\in\mathbb{N}}$ eine konvergente Teilfolge mit $\lim_{k\to\infty} a_{n_k} = a$. Sei $\varepsilon > 0$ und $\varepsilon' = \frac{\varepsilon}{3}$. Es gibt ein $k_0 \in \mathbb{N}$ mit $|a_{n_k} - a| < \varepsilon'$ für alle $k \geq k_0$. Ferner gibt es ein $n_0 \in \mathbb{N}$ mit $|a_n - a_{n_0}| < \varepsilon'$ für alle $n \geq n_0$. Sei $n \geq n_0$, dann gilt

$$a_n - a = (a_n - a_{n_0}) + (a_{n_0} - a_{n_k}) + (a_{n_k} - a),$$

wobei $k \geq k_0$ und $n_k \geq n_0$ gewählt wird. Damit gilt dann

$$|a_n - a| \leq \underbrace{|a_n - a_{n_0}|}_{\text{Cauchy}} + \underbrace{|a_{n_0} - a_{n_k}|}_{\text{Cauchy}} + \underbrace{|a_{n_k} - a|}_{\text{konv. Teilfolge}} < \varepsilon' + \varepsilon' + \varepsilon' = \varepsilon.$$

Somit haben wir insgesamt gezeigt, dass jede Cauchy-Folge konvergiert.

□

Ohne es explizit erwähnt zu haben, benutzt der Beweis für die Konvergenz von Cauchy-Folgen das Vollständigkeitsaxiom, denn die monotone, beschränkte Teilfolge bildet aus ihren Folgengliedern eine beschränkte Teilmenge von $\mathbb{R}$. Außerdem ist der Grenzwert einer monotonen Folge identisch mit dem

Supremum oder Infimum und diese existieren nur wegen der Vollständigkeit der reellen Zahlen (Axiom 3).

In vielen Lehrbüchern wird die Vollständigkeit der reellen Zahlen über die Konvergenz von Cauchy-Folgen definiert:

> **Alternative A zu Axiom 3**
>
> Jede Cauchy-Folge reeller Zahlen konvergiert gegen einen Grenzwert $g \in \mathbb{R}$.

Mit diesem Axiom kann (und muss) man wiederum beweisen, dass Suprema und Infima für beschränkte Teilmengen in $\mathbb{R}$ existieren. Die Entscheidung, welche Definition der Vollständigkeit man axiomatisch fordert ist größtenteils reine Geschmackssache. Man kann grundsätzlich aus jeder Variante die gleichen Schlüsse ziehen. Manche Varianten sind in bestimmten Fällen bei der Beweisführung aber einleuchtender. In Satz 2.21 hatten wir aus der Vollständigkeit der reellen Zahlen die Existenz von Quadratwurzeln in $\mathbb{R}$ abgeleitet.

Wenn man die Vollständigkeit über die Konvergenz von Cauchy-Folgen axiomatisch fordert, kann daraus ebenfalls die Existenz von Quadratwurzeln abgeleitet werden, indem Konvergenz und Grenzwert der Wurzelfolge aus Beispiel 3.16 und 3.21 gezeigt werden. Da jede konvergente Folge eine Cauchy-Folge ist, muss dies also auch für die Wurzelfolge gelten und der Grenzwert muss in $\mathbb{R}$ existieren. Häufiger wird jedoch zunächst aus den Cauchy-Folgen das sogenannte Intervallschachtelungsverfahren (siehe nächster Abschnitt) abgeleitet und damit die Wurzelexistenz bewiesen. Cauchy-Folgen werden uns aber im Kapitel 3.4 über Reihen wieder begegnen.

Intervallschachtelung

Wir definieren uns nun eine hilfreiche Schreibweise für bestimmte Mengen, die man Intervalle nennt:

Definition 3.35 (Intervalle)

Wir bezeichnen alle folgenden Mengen als Intervalle. Dabei werden diese jeweils einem bestimmten Intervalltyp zugeordnet:

- Abgeschlossene Intervalle:
  Seien $a, b \in \mathbb{R}, a \leq b$, dann ist $[a, b] := \{x \in \mathbb{R} \mid a \leq x \leq b\}$.
- Offene Intervalle:
  Seien $a, b \in \mathbb{R}, a < b$, dann ist $(a, b) := \{x \in \mathbb{R} \mid a < x < b\}$
  (Man schreibt manchmal auch $]a, b[$ statt (a, b))
- Halboffene Intervalle:
  Seien $a, b \in \mathbb{R}, a < b$:

  $[a, b) := \{x \in \mathbb{R} \mid a \leq x < b\}$
  $(a, b] := \{x \in \mathbb{R} \mid a < x \leq b\}$
- Uneigentliche Intervalle:
  Sei $a \in \mathbb{R}$:

  (uneigentlich und abgeschlossen)
  $[a, +\infty) := \{x \in \mathbb{R} \mid x \geq a\}$
  $(-\infty, a] := \{x \in \mathbb{R} \mid x \leq a\}$

  (uneigentlich und halboffen)
  $(-\infty, a) := \{x \in \mathbb{R} \mid x < a\}$
  $(a, +\infty) := \{x \in \mathbb{R} \mid x > a\}$

Außerdem führen wir die folgenden verkürzten Schreibweisen für bestimmte Intervalle ein:

- $\mathbb{R}_{>0} := (0, \infty)$
- $\mathbb{R}_{<0} := (-\infty, 0)$
- $\mathbb{R}_{\geq 0} := [0, \infty)$
- $\mathbb{R}_{\leq 0} := (-\infty, 0]$
- $\mathbb{R}_{\neq 0} := \mathbb{R}_{<0} \cup \mathbb{R}_{>0}$

Für ein Intervall $[a, b]$ bzw. (a, b) bezeichnet $|[a, b]| = b - a$ bzw. $|(a, b)| = b - a$ die *Länge des Intervalls*. Uneigentliche Intervalle haben die Länge ∞.

Aus der Definition folgt, dass das Intervall $[a, a]$, welches nur aus dem Punkt a besteht, die Länge 0 hat. Um die uneigentlichen abgeschlossenen Intervalle von den ‘normalen’ abgeschlossenen Intervallen zu unterscheiden, führen wir noch die Definition der Kompaktheit ein:

Definition 3.36 (Kompaktheit)

Wir nennen ein Intervall I *kompakt*, wenn es abgeschlossen und beschränkt ist.

Damit sind nur Intervalle der Form $[a, b]$ mit $a, b \in \mathbb{R}$ kompakt und Intervalle der Form $[a, \infty)$ nicht.

Der Unterschied zwischen offenen und geschlossenen Intervallen ist ähnlich zu dem Unterschied eines Maximums und eines Supremums, wie das folgende Beispiel zeigt.

Beispiel 3.37 (Sup/Inf/Min/Max von Intervallen)

- Für das abgeschlossene Intervall $I_1 = [a, b]$ gilt:
  $\inf(I_1) = \min(I_1) = a$
  $\sup(I_1) = \max(I_1) = b$
- Für das halboffene Intervall $I_2 = [a, b)$ gilt:
  $\inf(I_2) = \min(I_2) = a$
  $\sup(I_2) = a$, aber $\max(I_2)$ existiert nicht.
- Für das halboffene Intervall $I_3 = (a, b]$ gilt:
  $\inf(I_3) = a$, aber $\min(I_3)$ existiert nicht.
  $\sup(I_3) = \max(I_3) = b$

Mithilfe von Intervallen lässt sich eine sogenannte *Intervallschachtelung* definieren.

Definition 3.38 (Intervallschachtelung)

Eine Folge $(I_n)_{n\in\mathbb{N}}$ von abgeschlossenen Intervallen I_n heißt Intervallschachtelung, wenn die folgenden zwei Eigenschaften erfüllt sind:

- $\forall n \in \mathbb{N} \,:\, I_{n+1} \subset I_n$
- $\lim\limits_{n\to\infty} |I_n| = 0$

Die Intervallschachtelung begrenzt nach und nach immer enger eine bestimmte Zahl $x \in \mathbb{R}$. Daher eignet sie sich sehr gut, um die Existenz bestimmter reeller Zahlen zu beweisen. Wir weisen zunächst nach, dass so eine Zahl für jede Intervallschachtelung in $\mathbb{R}$ existieren muss.

Satz 3.39 (Konvergenz der Intervallschachtelung)

Für jede Intervallschachtelung (I_n) existiert genau ein eindeutiges $x \in \mathbb{R}$, für das gilt:

$$x \in I_n, \ \forall n \in \mathbb{N}$$

Wir sagen auch: die Intervallschachtelung konvergiert gegen x.

▼ Beweis

Betrachten wir die linken Intervallgrenzen der Schachtelung, so bilden diese eine monoton wachsende Folge (a_n). Analog bilden die rechten Intervallgrenzen eine monoton fallende Folge (b_n). Es gilt außerdem $a_n < b_n \ \forall n \in \mathbb{N}$. Also ist (a_n) nach oben und (b_n) nach unten beschränkt.

Damit gilt nach Satz 3.14 (bzw. nach Axiom 3), dass für beide Folgen ein Grenzwert in $\mathbb{R}$ existiert. Nennen wir diese Grenzwerte a und b, dann existiert ein Intervall $[a, b]$, das in jedem Intervall der Schachtelung liegt. Es gilt also noch zu zeigen, dass $a = b = x$ gilt, damit die Schachtelung genau gegen eine Zahl x konvergiert.

Da die Intervalllängen eine Nullfolge bilden folgt:

$$0 = \lim_{n\to\infty} |I_n| = \lim_{n\to\infty} (b_n - a_n) \overset{S.3.11}{=} \lim_{n\to\infty} b_n - \lim_{n\to\infty} a_n$$

$$\Leftrightarrow \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n.$$

Somit konvergiert die Intervallschachtelung gegen ein eindeutiges $x = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n$.

□

Auch die Konvergenz der Intervallschachtelung wird in manchen Lehrbüchern zur Definition der Vollständigkeit genutzt und für die reellen Zahlen $\mathbb{R}$ axiomatisch gefordert. Das entsprechende Alternativaxiom sähe dann so aus:

Alternative B zu Axiom 3

Jede Intervallschachtelung reeller Zahlen konvergiert gegen ein eindeutiges $x \in \mathbb{R}$.

Auch mit diesem Axiom könnte die Existenz von Suprema und Infima für beschränkte Mengen oder die Konvergenz von Cauchy-Folgen bewiesen werden. Es ist damit gleichwertig zu Axiom 3 und der

Alternative A zu Axiom 3. Wir können nun ein letztes Mal die Existenz von Quadratwurzeln beweisen, diesmal mithilfe einer Intervallschachtelung. Damit der Beweis auch ein paar neue Erkenntnisse bietet, verallgemeinern wir die Aussage direkt auf die Existenz k-ter Wurzeln.

Satz 3.40 (Existenz von Wurzeln)

Zu jedem $x \in \mathbb{R}_{>0}$ und jedem $k \in \mathbb{N}$ gibt es genau ein $y \in \mathbb{R}_{>0}$ mit $y^k = x$.

Dieses bezeichnen wir als die k-te Wurzel von x und nutzen die Schreibweise

$$y = x^{\frac{1}{k}} \text{ oder } y = \sqrt[k]{x}.$$

▼ Beweis

Es genügt $x > 1$ zu behandeln, denn den Fall $x < 1$ führt man durch den Übergang $x' = \frac{1}{x}$ auf den Fall $x > 1$ zurück. Für $x = 1$ ist $y = 1$ die Lösung.

Wir konstruieren per vollständiger Induktion eine Intervallschachtelung mit Intervallen $I_n = [a_n, b_n]$, für die gilt:

i. $a_n^k \leq x \leq b_n^k$
ii. $|I_n| = \left(\frac{1}{2}\right)^{n-1} |I_1|$.

Sei $I_1 = [1, x]$. Offensichtlich gelten **(i)** und **(ii)**, da $x > 1$.

I_{n+1} wird aus I_n gebildet, indem wir den Intervallmittelpunkt $m = \frac{1}{2}(a_n + b_n)$ bestimmen. Dann ist

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

Offensichtlich gilt **(i)** und $|I_{n+1}| = \frac{1}{2}|I_n|$. Damit folgt aus $|I_2| = \frac{1}{2}|I_1|$ auch **(ii)**. Die Folge der Intervalle bildet eine Intervallschachtelung, da $I_{n+1} \subset I_n$ und da es für jedes $\varepsilon > 0$ ein n gibt, sodass

$$\left(\frac{1}{2}\right)^{n-1} < \varepsilon |I_1|^{-1} \quad \Rightarrow \quad |I_n| < \varepsilon$$

Sei y die in allen Intervallen I_n liegende Zahl. Es bleibt zu zeigen, dass $y^k = x$ gilt. Zunächst zeigen wir, dass auch die Intervalle $I_n^k = [a_n^k, b_n^k]$ eine Intervallschachtelung bilden:

- $I_{n+1}^k \subset I_n^k$ gilt wegen $I_{n+1} \subset I_n$
- Für die Länge jedes Intervalls I_n^k gilt

$$\begin{aligned} \left|I_n^k\right| &= (b_n - a_n)(b_n^{k-1} + b_n^{k-2} a_n + \ldots + a_n^{k-1}) \\ &< |I_n| \cdot k \cdot b_1^{k-1}, \end{aligned}$$

  da $b_n > a_n \geq 1$ und $1 < b_n \leq b_1$.

Sei nun $\varepsilon > 0$ gegeben, so existiert ein Index υ, so dass $|I_\upsilon| < \varepsilon' = \frac{\varepsilon}{k \cdot b_1^{k-1}}$ und damit $\left|I_\upsilon^k\right| < \varepsilon$.

Da y in I_n liegt, liegt y^k in I_n^k für alle $n \in \mathbb{N}$. Außerdem liegt x in I_n^k für alle $n \in \mathbb{N}$, da **(i)** gilt. Da es nur genau eine Zahl gibt, die in allen Intervallen I_n^k liegt, gilt $y^k = x$.

Zu zeigen bleibt die Eindeutigkeit von y. Sei z eine weitere Zahl mit $z^k = x$ und $z \neq y$. Dann muss gelten $z > y$ oder $z < y$, woraus $z^k > y^k = x$ bzw. $z^k < y^k = x$ folgen würde, da $y > 1$ vorausgesetzt wurde. Dies steht im Widerspruch zur Annahme $z^k = x$ (↯). Also ist y eindeutig.

□

3.4 Reihen

Neben Folgen sind Reihen ein weiteres Hilfsmittel, um Probleme approximativ zu lösen. Im Prinzip könnte man Reihen als spezielle Folgen betrachten. Ihre besondere Form und große praktische Bedeutung führen dazu, dass wir sie separat vorstellen.

Definition 3.41 (Reihe)

Man nennt den Ausdruck $\sum_{k=1}^{\infty} a_k = a_1 + a_2 + \dots$ mit $a_k \in \mathbb{R}$ eine *(unendliche) Reihe* und $s_n = \sum_{k=1}^{n} a_k$ die *n-te Teilsumme* der Reihe.

Wenn die Folge der Teilsummen (s_n) konvergiert, dann heißt die Reihe *konvergent*. Eine nicht konvergente Reihe heißt *divergent*.

Im Falle der Konvergenz bezeichnet $\sum_{k=1}^{\infty} a_k$ nicht nur die Reihe bzw. Folge der Teilsummen, sondern auch den Grenzwert der Reihe.

Die Summation kann bei Reihen auch mit $k = 0$ (oder einem beliebigen anderen Index) beginnen, falls dies für die entsprechenden Beispiele zu einer einfacheren Darstellung führt. Da die Konvergenz einer Reihe über die Konvergenz der Teilsummen definiert ist, können wir das Cauchy-Konvergenzkriterium für Folgen auch auf Reihen übertragen:

Satz 3.42 (Cauchy-Konvergenzkriterium für Reihen)

Die Reihe $\sum_{k=1}^{\infty} a_k$ konvergiert genau dann, wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt, sodass $\left| \sum_{k=m}^{n} a_k \right| < \varepsilon$ für alle $n \geq m \geq n_0$.

▼ Beweis

Es gilt offensichtlich $s_n - s_m = \sum_{k=m+1}^{n} a_k$. Alles Weitere folgt direkt aus Satz 3.34.

□

Auch hier fällt wieder auf, dass die Konvergenzbedingung vollkommen unabhängig vom Startindex der Reihe ist. Ein paar endliche Summanden können lediglich den Grenzwert um einen endlichen Wert erhöhen/verringern, aber nichts an der Konvergenz oder Divergenz der Reihe ändern.

Beispiel 3.43 (harmonische Reihe)

Die Reihe $\sum_{k=1}^{\infty} \frac{1}{k}$ nennt man die *harmonische Reihe*. Diese Reihe ist divergent.

Beweis: Wäre die harmonische Reihe konvergent, müsste es für $\varepsilon = \frac{1}{2}$ ein n_0 geben, sodass

$$\forall n \geq m \geq n_0 : \left| \sum_{k=m}^{n} \frac{1}{k} \right| < \varepsilon.$$

Betrachten wir aber $m = n_0$ und $n = 2m - 1$, so folgt

$$\sum_{k=m}^{2m-1} \frac{1}{k} = \frac{1}{m} + \cdots + \frac{1}{2m-1} \geq \underbrace{\frac{1}{2m} + \cdots + \frac{1}{2m}}_{m \text{ Summanden}} = \frac{1}{2} = \varepsilon.$$

Damit erfüllt die harmonische Reihe nicht das Cauchy-Konvergenzkriterium und kann demnach nicht konvergent sein.

▶ Reihe visualisieren

Die harmonische Reihe ist eine sehr wichtige Vergleichsreihe, da sie “gerade so” nicht konvergiert. Sie können zur Übung einmal bestimmen, wie viele Folgenglieder sie benötigen, um auch nur eine Summe von 10 oder 20 zu erreichen. Die Divergenz bedeutet, dass man jede noch so hohe Summe mit der harmonischen Reihe erreichen kann. Diese Reihe zeigt auch, dass es Reihen geben kann, deren Reihenglieder zwar eine Nullfolge sind, aber die trotzdem nicht konvergieren.

Umgekehrt gilt aber für jede Reihe, dass sie divergiert, wenn die Reihenglieder *keine* Nullfolge bilden. Sind z.B. alle Folgenglieder größer als ein positives ε, so werden unendlich viele Summanden ε aufaddiert. Damit kann eine beliebig hohe Summe K durch $\lceil K/\varepsilon \rceil$ Summanden erreicht oder überschritten werden, wodurch die Folge der Partialsummen unbeschränkt, die Reihe also also divergent ist. Dies wird im folgenden Satz formal bewiesen:

Satz 3.44 (Notwendiges Konvergenzkriterium)

Sei $\sum_{k=1}^{\infty} a_k$ eine konvergente Reihe. Dann ist die Folge der Summanden (a_k) eine Nullfolge. Es gilt also $\lim_{k\to\infty} a_k = 0$.

▼ Beweis

Das ε-Kriterium für a_k folgt direkt aus dem Cauchy-Konvergenzkriterium der Reihe für $n = m$:

$$\forall \varepsilon > 0 \; \exists n_0 \in \mathbb{N} \; \forall m = n \geq n_0 : \; \left| \sum_{k=n}^{n} a_k \right| = |a_k| = |a_k - 0| < \varepsilon$$

□

Beispiel 3.45

Die Reihe $\sum_{k=1}^{\infty} (-1)^k$ ist divergent, da die Reihenglieder nicht gegen 0 konvergieren.

Häufig werden wir Reihen betrachten, bei denen alle Folgenglieder positiv sind. Damit ist die Folge der Partialsummen automatisch monoton wachsend und es folgt analog zu Satz 3.14 folgender Satz für Reihen.

Satz 3.46 (Reihenkonvergenz durch Teilsummenbeschränktheit)

Eine Reihe $\sum_{k=1}^{\infty} a_k$ mit $a_k \geq 0$ für alle $k \in \mathbb{N}$ konvergiert genau dann, wenn die Folge der Teilsummen beschränkt ist.

▼ Beweis

Aus $a_k \geq 0$ folgt, dass $s_n = \sum_{k=1}^{n} a_k$ monoton wachsend ist. Ist s_n zusätzlich beschränkt, so gilt nach Satz 3.14, dass s_n konvergiert und somit auch die Reihe konvergiert.

Für die zweite Richtung ("Die Reihe konvergiert $\Rightarrow$ s_n ist beschränkt") beweisen wir die Kontraposition, also "s_n ist unbeschränkt $\Rightarrow$ die Reihe divergiert". Dies haben wir aber bereits mit der Kontraposition von Satz 3.12 bewiesen.

□

Konkrete Grenzwerte von Reihen zu bestimmen, ist gar nicht so einfach. Eine Ausnahme bilden Fälle, in denen wir die n-te Partialsumme so weit vereinfachen können, dass wir die Reihe in eine normale Folge überführen können. Die Gleichheit des Folgenterms und der Partialsumme zeigt man in der Regel per vollständiger Induktion. Ein prominentes und sehr wichtiges Beispiel hierfür ist die *geometrische Reihe*.

Beispiel 3.47 (geometrische Reihe)

Die geometrische Reihe $\sum_{k=0}^{\infty} x^k$ konvergiert für $|x| < 1$ gegen den Grenzwert $\frac{1}{1-x}$.

Wir zeigen dazu zunächst per Induktion, dass

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

Induktionsanfang $n = 0$**:** $x^0 = 1 = \frac{1-x}{1-x}$

Induktionsvoraussetzung: Die Behauptung gelte für ein beliebiges $n \in \mathbb{N}_0$.

Induktionsschritt $n \to n + 1$**:** Es gilt nach Induktionsvoraussetzung $\sum_{k=0}^{n} x^k = \frac{1-x^{n+1}}{1-x}$, dann gilt auch

$$\begin{aligned}
\sum_{k=0}^{n+1} x^k &= \sum_{k=0}^{n} x^k + x^{n+1} \\
&\overset{IV}{=} \frac{1 - x^{n+1}}{1 - x} + x^{n+1} \\
&= \frac{1 - x^{n+1}}{1 - x} + x^{n+1}\frac{1 - x}{1 - x} \\
&= \frac{1 - x^{n+1}}{1 - x} + \frac{x^{n+1} - x^{n+2}}{1 - x} \\
&= \frac{1 - x^{n+2}}{1 - x}.
\end{aligned}$$

Damit gilt $s_n = \frac{1-x^{n+1}}{1-x}$ für alle $n \in \mathbb{N}_0$.

Da x^{n+1} für $|x| < 1$ gegen 0 konvergiert, folgt für den Grenzwert:

$$\lim_{n\to\infty} \sum_{k=0}^{n} x^k = \lim_{n\to\infty} \frac{1 - x^{n+1}}{1 - a} = \frac{1 - \lim\limits_{n\to\infty} (x^{n+1})}{1 - x} = \frac{1 - \lim\limits_{n\to\infty} (x^{n+1})}{1 - x} = \frac{1}{1 - x}.$$

▶ Geometrische Reihe für x=0.8 visualisieren

▶ Geometrische Reihe für x=1.2 visualisieren

In anderen Fällen lässt sich der Grenzwert durch die Ausnutzung sogenannter *Teleskopsummen* beweisen. Dabei versucht man die Reihenglieder a_k durch eine zweite Folge (b_k) so umzuschreiben, dass $a_k = b_k - b_{k+1}$. Damit ergibt sich für die n-te Partialsumme

$$\sum_{k=0}^{n} a_k = \sum_{k=0}^{n} (b_k - b_{k+1}) = b_0 - b_1 + b_1 - b_2 + b_2 - b_3 + \ldots + b_n - b_{n+1} = b_0 - b_{n+1}$$

Die Grenzwertbestimmung vereinfacht sich dadurch zu

$$\sum_{k=0}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=0}^{n} a_k = b_0 - \lim_{n\to\infty} b_{n+1}.$$

Beispiel 3.48 (Konvergenzbeweis durch Teleskopsummen)

Die Reihe $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$ konvergiert und es gilt $\lim_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k(k+1)} = 1$.

Um den Grenzwert herzuleiten nutzen wir folgende Identität $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$ für alle $k \in \mathbb{N}$. Damit lautet die n-te Teilsumme

$$s_n = 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \ldots - \frac{1}{n} + \frac{1}{n} - \frac{1}{n+1} = 1 - \frac{1}{n+1}$$

und es gilt $\lim_{n\to\infty} s_n = 1$.

▶ Reihe visualisieren

Ähnlich wie bei der Verknüpfung zweier konvergenter Folgen (Satz 3.18) können wir auch konvergente Reihen miteinander kombinieren.

Satz 3.49 (Rechnen mit konvergenten Reihen)

Seien $\sum_{k=1}^{\infty} a_k$ und $\sum_{k=1}^{\infty} b_k$ zwei konvergente Reihen, dann folgt:

a. Auch $\sum_{k=1}^{\infty} (a_k + b_k)$ und $\sum_{k=1}^{\infty} (a_k - b_k)$ sind konvergent und für die Grenzwerte gilt

$$\sum_{k=1}^{\infty} (a_k + b_k) = \sum_{k=1}^{\infty} a_k + \sum_{k=1}^{\infty} b_k,$$
$$\sum_{k=1}^{\infty} (a_k - b_k) = \sum_{k=1}^{\infty} a_k - \sum_{k=1}^{\infty} b_k.$$

b. Auch $\sum_{k=1}^{\infty} c \cdot a_k$ für ein beliebiges $c \in \mathbb{R}$ ist konvergent und für den Grenzwert gilt

$$\sum_{k=1}^{\infty} c \cdot a_k = c \cdot \sum_{k=1}^{\infty} a_k.$$

c. Für jedes $l \in \mathbb{N}$ mit $l > 1$ gilt:

$$\sum_{k=l}^{\infty} a_k \text{ konvergiert} \quad \Leftrightarrow \quad \sum_{k=1}^{\infty} a_k \text{ konvergiert.}$$

d. Gilt für die Folgenglieder $a_k \leq b_k \; \forall k \in \mathbb{N}$, so folgt

$$\sum_{k=1}^{\infty} a_k \leq \sum_{k=1}^{\infty} b_k.$$

▼ Beweis

Wir nutzen $s_n = \sum_{k=1}^{n} a_k$ und $t_n = \sum_{k=1}^{n} b_k$ als Kurzschreibweise für die n-ten Teilsummen der Reihen. Der Beweis nutzt Satz 3.18 zur Bestimmung der Grenzwerte kombinierter Folgen, wobei (s_n) und (t_n) die betrachteten Folgen sind.

a. Sei $u_n = \sum_{k=1}^{n} (a_k + b_k)$, dann gilt für den Grenzwert der Folge $(s_n + t_n)$ nach Satz 3.18(a): $\lim_{n\to\infty} s_n + \lim_{n\to\infty} t_n = \lim_{n\to\infty} (s_n + t_n) = \lim_{n\to\infty} u_n$, womit auch die kombinierte Reihe gegen den angegebenen Grenzwert konvergiert.

b. Da $s_n = \sum_{k=1}^{n} a_k$ gilt nach Satz 3.18(b): $\lim_{n\to\infty} c \cdot s_n = c \cdot \lim_{n\to\infty} s_n = c \cdot \sum_{k=1}^{\infty} a_k$, womit auch die skalierte Reihe gegen den angegebenen Grenzwert konvergiert.

c. Seien $r = \sum_{k=1}^{l-1} a_k$ und $s'_n = \sum_{k=l}^{n} a_k$ (beachten Sie, dass wir in Definition 2.33 Summen mit $n < l$ als 0 definiert haben). Damit entspricht der Grenzwert der Folge (s'_n) der Reihe $\sum_{k=l}^{\infty} a_k$. Es gilt $s'_n = s_n - \sum_{k=1}^{l-1} a_k = s_n - r$. Genauso gilt umgekehrt $s_n = s'_n + r$. Wir können r als konstante Folge interpretieren, womit nach Satz 3.18(a) gilt, dass s'_n genau dann konvergiert wenn s_n konvergiert.

d. Aus $a_k \leq b_k$ für alle $k \in \mathbb{N}$, $s_n \leq t_n$ für alle $n \in \mathbb{N}$, woraus mithilfe von Satz 3.22 die Behauptung folgt.

□

Um festzustellen, ob eine Reihe konvergiert, können wir im Prinzip die gleichen Kriterien anwenden, die wir bereits für Folgen kennengelernt haben. Allerdings ist dies durch die besondere Struktur der Reihen oft nicht leicht. In vielen Fällen können wir nicht, wie in den vorigen Beispielen, die Summenschreibweise vereinfachen. Daher gibt es besondere Konvergenzkriterien für Reihen, von denen wir im Folgenden einige kennenlernen werden. Wir beginnen mit dem Leibniz-Kriterium, welches wir für Reihen mit alternierenden Gliedern (d. h. Gliedern mit wechselnden Vorzeichen) anwenden können.

Satz 3.50 (Leibniz-Kriterium)

Sei $(a_k)_{k\in\mathbb{N}}$ eine monoton fallende Folge reeller nicht negativer Zahlen mit $\lim\limits_{k\to\infty} a_k = 0$. Dann konvergiert die alternierende Reihe $\sum\limits_{k=1}^{\infty} (-1)^k a_k$.

▼ Beweis

Für die Folge der Teilsummen gilt $s_n = \sum\limits_{k=1}^{n} (-1)^k a_k$.

Da (a_k) monoton fällt, gilt $a_{n+1} \leq a_n \; \forall n \in \mathbb{N}$.

Wir betrachten die geraden und ungeraden Teilsummen separat:
Da $s_{2n+2} - s_{2n} = -a_{2n+1} + a_{2n+2} \leq 0$, gilt: $s_2 \geq s_4 \geq s_6 \geq \ldots$
Da $s_{2n+3} - s_{2n+1} = a_{2n+2} - a_{2n+3} \geq 0$, gilt: $s_1 \leq s_3 \leq s_5 \leq \ldots$

Damit ist die Folge der geraden Teilsummen $(s_{2n})_{n\in\mathbb{N}}$ monoton fallend und die Folge der ungeraden Teilsummen $(s_{2n+1})_{n\in\mathbb{N}}$ monoton wachsend.

Da $s_{2n} \geq s_1$ und $s_{2n+1} \leq s_2$ sind beide Folgen beschränkt und es existieren die Grenzwerte $S = \lim\limits_{n\to\infty} s_{2n}$ und $S' = \lim\limits_{n\to\infty} s_{2n+1}$.

Es gilt $S = S'$, da $S - S' = \lim\limits_{n\to\infty} (s_{2n} - s_{2n+1}) = \lim\limits_{n\to\infty} a_{2n+1} = 0$.

An dieser Stelle ist die Konvergenz der Gesamtreihe (s_n) eigentlich schon klar, da alle Folgenglieder entweder gerade oder ungerade sind und beide Teilfolgen gegen S konvergieren. Damit muss auch jede andere Teilfolge und somit auch die Gesamtfolge gegen S konvergieren. Trotzdem führen wir noch den formalen Beweis über das ε-Kriterium:

Für alle $\varepsilon > 0$ existieren $n_1, n_2 \in \mathbb{N}$, sodass $|s_{2n} - S| < \varepsilon$ für $n \geq n_1$ und $|s_{2n+1} - S| < \varepsilon$ für $n \geq n_2$.
Wählen wir nun für die gesamte Folge (s_n) der Teilsummen $n_{12} = \max\{2n_1, 2n_2 + 1\}$ dann gilt $|s_n - S| < \varepsilon$ für alle $n \geq n_{12}$. Damit ist die Konvergenz bewiesen.

□

Beispiel 3.51 (alternierende harmonische Reihe)

Die alternierende harmonische Reihe $\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}$ konvergiert nach dem Leibniz-Kriterium, da $a_n = \frac{1}{n}$ eine Nullfolge ist.

▶ Reihe visualisieren

Absolut konvergente Reihen

Definition 3.52 (Absolute Konvergenz)

Wir nennen eine Reihe $\sum_{k=1}^{\infty} a_k$ absolut konvergent, wenn die Reihe der Absolutbeträge $\sum_{k=1}^{\infty} |a_k|$ konvergiert. Zur besseren Unterscheidung sprechen wir bei der Konvergenz von $\sum_{k=1}^{\infty} a_k$ (ohne Absolutbeträge) auch von der *gewöhnlichen Konvergenz*.

Durch die Beträge in $\sum |a_k|$ ist die Folge der Teilsummen stets monoton wachsend. Das bedeutet, wenn wir zusätzlich eine obere Schranke für $\sum |a_k|$ finden können, folgt die absolute Konvergenz von $\sum a_k$ über Satz 3.14. Für die absolute Konvergenz haben sich daher besondere Kriterien entwickelt, durch welche diese für viele Reihen leicht bestimmbar oder widerlegbar ist.

Das Kriterium der absoluten Konvergenz hat auch praktische Relevanz. Bei einer Reihe, die zwar gewöhnlich, aber nicht absolut konvergiert, kann sich beispielsweise der Grenzwert verändern, wenn wir die Reihensummanden vertauschen. Dies zeigen wir im folgenden Beispiel.

Beispiel 3.53 (Umordnung)

Für die alternierende harmonische Reihe

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \frac{1}{7} - \frac{1}{8} + \ldots + \frac{1}{2n-1} - \frac{1}{2n} \pm \ldots$$

haben wir bereits gezeigt, dass diese zwar gewöhnlich, aber nicht absolut konvergiert. Wir bezeichnen den Grenzwert der Reihe mit S.

Nun betrachten wir die folgende Umordnung:

$$\left(1 - \frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{3} - \frac{1}{6} - \frac{1}{8}\right) + \ldots + \left(\frac{1}{2n-1} - \frac{1}{4n-2} - \frac{1}{4n}\right) \pm \ldots$$

Dabei starten wir mit dem 1. ungeraden Summanden gefolgt von dem 1. und 2. geraden Summanden, dann folgt der 2. ungerade Summand und der 3. und 4. gerade, und so weiter. Da die Reihe unendlich viele Summanden hat, ist auch bei dieser Umordnung jeder Summand der ursprünglichen Reihenfolge irgendwann an der Reihe, also würde man vermuten, dass man den gleichen Grenzwert S erreicht. Allerdings lässt sich für den Grenzwert der Umordnung S' zeigen:

$$\begin{aligned}
S' &= \left(1 - \frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{3} - \frac{1}{6} - \frac{1}{8}\right) + \ldots + \left(\frac{1}{2n-1} - \frac{1}{4n-2} - \frac{1}{4n}\right) \pm \ldots \\
&= \left(\frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{6} - \frac{1}{8}\right) + \ldots + \left(\frac{1}{4n-2} - \frac{1}{4n}\right) \pm \ldots \\
&= \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \ldots + \frac{1}{2n-1} - \frac{1}{2n} \pm \ldots\right) \\
&= \frac{1}{2}\left(\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}\right) = \frac{1}{2}S
\end{aligned}$$

Diese Umordnung erreicht also einen anderen Grenzwert, nämlich $S' = \frac{1}{2}S$.

Für absolut konvergente Reihen können wir dagegen beweisen, dass jede Umordnung gegen denselben Grenzwert konvergiert und müssen uns daher über die Reihenfolge der Summation keine Gedanken machen.

Satz 3.54 (Reihenumordnung)

Sei $\sum_{k=1}^{\infty} a_k$ eine absolut konvergente Reihe. Dann konvergiert jede Umordnung der Glieder der Reihe gegen denselben Grenzwert.

▼ Beweis

Sei $\tau : \mathbb{N} \to \mathbb{N}$ eine Abbildung, die jedem $n_1 \in \mathbb{N}$ genau ein $n_2 \in \mathbb{N}$ zuordnet, sodass jedem $n_2 \in \mathbb{N}$ genau ein $n_1 \in \mathbb{N}$ zugeordnet wird.

Sei $\lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=1}^{n} a_k = A$. Wir zeigen, dass auch $\lim_{n\to\infty} \sum_{k=1}^{n} a_{\tau(k)} = A$ gilt.

Nach Voraussetzung gibt es für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, sodass $\sum_{k=n_0}^{\infty} |a_k| < \frac{\varepsilon}{2}$.

Daraus folgt

$$\left| A - \sum_{k=1}^{n_0-1} a_k \right| = \left| \sum_{k=n_0}^{\infty} a_k \right| \leq \sum_{k=n_0}^{\infty} |a_k| < \frac{\varepsilon}{2}.$$

Wähle n_1 so groß, dass $\{1, 2, \dots, n_0\} \subseteq \{\tau(1), \tau(2), \dots, \tau(n_1)\}$.
Dann gilt für alle $m \geq n_1$ $(\geq n_0)$:

$$\left| \sum_{k=1}^{m} a_{\tau(k)} - A \right| \leq \left| \sum_{k=1}^{m} a_{\tau(k)} - \sum_{k=1}^{n_0-1} a_k \right| + \left| \sum_{k=1}^{n_0-1} a_k - A \right| \leq \sum_{k=n_0}^{\infty} |a_k| + \frac{\varepsilon}{2} < \varepsilon.$$

Damit konvergiert die umgeordnete Reihe gegen denselben Grenzwert.

□

Die absolute Konvergenz erlaubt uns also, dass wir mit unendlichen Reihen genauso umgehen dürfen wie mit endlichen Summen. Außerdem werden wir nun einige nützliche Kriterien herleiten, welche uns erlauben, Reihen recht einfach auf absolute Konvergenz zu prüfen. Beginnen wir mit einem Satz über den Zusammenhang zwischen absoluter und gewöhnlicher Konvergenz.

Satz 3.55 (Absolute Konvergenz ⇒ gewöhnliche Konvergenz)

Wenn die Reihe $\sum_{k=1}^{\infty} a_k$ absolut konvergiert, so konvergiert sie auch im gewöhnlichen Sinne.

▼ Beweis

Aus der Dreiecksungleichung (Satz 2.17) folgt:

$$\left|\sum_{k=m}^{n} a_k\right| \leq \sum_{k=m}^{n} |a_k| = \left|\sum_{k=m}^{n} |a_k|\right|.$$

Damit gilt

$$\left|\sum_{k=m}^{n} |a_k|\right| < \varepsilon \quad \Rightarrow \quad \left|\sum_{k=m}^{n} a_k\right| < \varepsilon,$$

sodass die Cauchy-Konvergenzbedingung für die Reihe erfüllt ist, wenn die Bedingung für absolute Konvergenz erfüllt ist.

□

Wie bereits erwähnt, müssen wir für den Beweis der absoluten Konvergenz eine obere Schranke finden. Eine Möglichkeit dafür ist es, eine größere Reihe zu finden, von der wir wissen, dass sie konvergiert.

Satz 3.56 (Majorantenkriterium)

Sei $\sum_{k=1}^{\infty} c_k$ eine konvergente Reihe mit ausschließlich nicht-negativen Gliedern.

Wenn für die Reihe $\sum_{k=1}^{\infty} a_k$ ein $k_0 \in \mathbb{N}$ existiert, sodass für alle $k \geq k_0$ gilt $|a_k| \leq c_k$, dann konvergiert die Reihe $\sum_{k=1}^{\infty} a_k$ absolut.

▼ Beweis

Aus der Konvergenz von $\sum_{k=1}^{\infty} c_k$ folgt, dass für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ existiert, sodass für alle $n \geq m \geq n_0$ gilt:

$$\left| \sum_{k=m}^{n} c_k \right| < \varepsilon.$$

Da $|a_k| \leq c_k$ nach Voraussetzung für alle $k \geq k_0$ gilt, folgt mit $n_0' = \max\{n_0, k_0\}$

$$\sum_{k=m}^{n} |a_k| \leq \sum_{k=m}^{n} c_k = \left| \sum_{k=m}^{n} c_k \right| < \varepsilon$$

für alle $n \geq m \geq n_0'$.

Die Reihe $\sum_{k=1}^{\infty} |a_k|$ erfüllt also das Cauchy-Kriterium und konvergiert.

Somit konvergiert $\sum_{k=1}^{\infty} a_k$ absolut.

□

Beispiel 3.57 (allgemeine harmonische Reihe)

Die Reihe $\sum_{k=1}^{\infty} \frac{1}{k^m}$ ist eine Verallgemeinerung der harmonischen Reihe. Wir zeigen nun, dass diese für alle natürlichen Zahlen $m \geq 2$ (absolut) konvergiert.

Betrachten wir zunächst den Fall $m = 2$:
Nach Beispiel 3.4 ist die Reihe $\sum_{k=1}^{\infty} \frac{1}{k(k+1)}$ konvergent. Außerdem gilt $\frac{1}{k(k+1)} \geq \frac{1}{k(k+k)} = \frac{1}{2k^2}$. Damit gilt

$$\frac{1}{2}\sum_{k=1}^{\infty} \frac{1}{k^2} = \sum_{k=1}^{\infty} \frac{1}{2k^2} \leq \sum_{k=1}^{\infty} \frac{1}{k(k+1)}$$

Es gibt also eine größere Reihe, die konvergiert. Damit konvergiert nach dem Majorantenkriterium auch $\frac{1}{2}\sum_{k=1}^{\infty} \frac{1}{k^2}$ (absolut) und somit konvergiert nach Satz 3.49 ebenfalls die doppelte Reihe $\sum_{k=1}^{\infty} \frac{1}{k^2}$.

Für alle Exponenten $m > 2$ gilt $\frac{1}{k^m} \leq \frac{1}{k^2}$ und somit auch

$$\sum_{k=1}^{\infty} \frac{1}{k^m} \leq \sum_{k=1}^{\infty} \frac{1}{k^2},$$

womit nach dem Majorantenkriterium auch diese Reihen (absolut) konvergieren.

▶ Reihe visualisieren

Das Majorantenkriterium impliziert außerdem folgendes Divergenz-Kriterium.

Satz 3.58 (Minorantenkriterium)

Sei $\sum_{k=1}^{\infty} c_k$ eine divergente Reihe mit ausschließlich nicht-negativen Gliedern.

Wenn für die Reihe $\sum_{k=1}^{\infty} a_k$ ein $k_0 \in \mathbb{N}$ existiert, sodass für alle $k \geq k_0$ gilt $a_k \geq c_k$, dann divergiert auch $\sum_{k=1}^{\infty} a_k$.

▼ Beweis als Übung

Beispiel 3.59 (Minorantenkriterium)

Die Reihe $\sum_{k=1}^{\infty} \frac{1}{\sqrt{k}}$ divergiert, da $\frac{1}{\sqrt{k}} \geq \frac{1}{k}$ und die harmonische Reihe $\sum_{k=1}^{\infty} \frac{1}{k}$ divergiert.

▶ Reihe visualisieren

Die beiden folgenden Konvergenzkriterien basieren auf der Anwendung des Majorantenkriteriums auf die geometrische Reihe, wie man aus den Beweisen ablesen kann.

Satz 3.60 (Wurzelkriterium)

a. Wenn ein festes $q \in \mathbb{R}$ mit $0 < q < 1$ und $k_0 \in \mathbb{N}$ existieren, sodass

$$\forall k \geq k_0 : \sqrt[k]{|a_k|} \leq q,$$

dann konvergiert die Reihe $\sum_{k=1}^{\infty} a_k$ absolut.

b. Wenn $k_0 \in \mathbb{N}$ existiert, sodass

$$\forall k \geq k_0 : \sqrt[k]{|a_k|} \geq 1,$$

dann divergiert die Reihe $\sum_{k=1}^{\infty} a_k$.

Limesform:
Existiert der Grenzwert $a = \lim_{k\to\infty} \sqrt[k]{|a_k|}$, dann gilt:

- Die Reihe ist absolut konvergent, wenn $a < 1$.
- Die Reihe divergiert, wenn $a > 1$.
- Für $a = 1$ kann keine Aussage getroffen werden.

▼ Beweis

a. Mit Satz 2.36 folgt aus $\sqrt[k]{|a_k|} \leq q \Leftrightarrow |a_k| \leq q^k$. Damit konvergiert die Reihe $\sum_{k=1}^{\infty} a_k$ nach dem Majorantenkriterium absolut, da die geometrische Reihe $\sum_{k=0}^{\infty} q^k$ für $|q| < 1$ konvergiert.

b. Wenn aber $\sqrt[k]{|a_k|} \geq 1$ gilt, dann folgt daraus $|a_k| \geq 1$ und somit, dass a_k keine Nullfolge ist. Daher divergiert die Reihe in diesem Fall nach Satz 3.44.

Zur Limesform:
Wenn der Grenzwert $a = \lim_{k\to\infty} \sqrt[k]{|a_k|}$ existiert, dann gilt für alle $\varepsilon > 0$, also auch für $\varepsilon = |1 - a|$, dass ein $k_0 \in \mathbb{N}$ existiert, sodass für alle $n \geq k_0$ gilt:

$$\left| \sqrt[k]{|a_k|} - a \right| < |1 - a|$$

Ist $a < 1$ muss demnach $\sqrt[k]{|a_k|} < 1$ gelten, womit die Bedingungen für den ersten Fall des Wurzelkriteriums gegeben sind.

Ist $a > 1$, so kann a_k keine Nullfolge sein und damit divergiert die Reihe nach Satz 3.44.

Bemerkung: Für $a = 1$ ist die obige Argumentation nicht möglich, da sonst $\varepsilon = |1 - a| = 0$, was $\varepsilon > 0$ widersprechen würde.

□

Obwohl das Majorantenkriterium das Wurzelkriterium impliziert und damit eine stärkere Bedingung darstellt, ist gerade die Limesform des Wurzelkriteriums sehr hilfreich für Konvergenzbeweise, da sie den Beweis auf das Abarbeiten einer einfachen ‘Anleitung’ reduziert. Gerade für Reihen mit k im Exponenten ist das Wurzelkriterium ideal geeignet, wie das folgende Beispiel zeigt.

Beispiel 3.61 (Wurzelkriterium)

Wir betrachten die Reihe $\sum_{k=1}^{\infty} \left(x + \frac{1}{k}\right)^k$ mit $0 \leq x < 1$.

Es gilt $\sqrt[k]{\left(x + \frac{1}{k}\right)^k} = \left(x + \frac{1}{k}\right)$ und $\lim_{k\to\infty} \left(x + \frac{1}{k}\right) = x < 1$. Damit konvergiert die Reihe.

Satz 3.62 (Quotientenkriterium)

a. Wenn ein festes $q \in \mathbb{R}$ mit $0 < q < 1$ und $k_0 \in \mathbb{N}$ existieren, sodass

$$\forall k \geq k_0 : \; a_k \neq 0 \quad \wedge \quad \left| \frac{a_{k+1}}{a_k} \right| \leq q,$$

dann konvergiert die Reihe $\sum_{k=1}^{\infty} a_k$ absolut.

b. Wenn $k_0 \in \mathbb{N}$ existiert, sodass

$$\forall k \geq k_0 : \; a_k \neq 0 \quad \wedge \quad \left| \frac{a_{k+1}}{a_k} \right| \geq 1,$$

dann divergiert die Reihe $\sum_{k=1}^{\infty} a_k$.

Limesform:
Existiert der Grenzwert $a = \lim_{k \to \infty} \left| \frac{a_{k+1}}{a_k} \right|$, dann gilt:

- Die Reihe ist absolut konvergent, wenn $a < 1$.
- Die Reihe divergiert, wenn $a > 1$.
- Für $a = 1$ kann keine Aussage getroffen werden.

▼ Beweis

a. Wir betrachten für $k > k_0$ den Term

$$\left| \frac{a_k}{a_{k_0}} \right| = \left| \frac{a_k}{a_{k-1}} \right| \cdot \left| \frac{a_{k-1}}{a_{k-2}} \right| \cdots \left| \frac{a_{k_0+1}}{a_{k_0}} \right| \leq q \cdot q \cdot q \cdots q = q^{k-k_0}.$$

Damit ist

$$|a_k| \leq |a_{k_0}| \cdot q^{k-k_0} = \frac{|a_{k_0}|}{q^{k_0}} \cdot q^k = c \cdot q^k$$

mit $c = \frac{|a_{k_0}|}{q^{k_0}}$. Damit ist die Reihe $c \sum_{k=k_0}^{\infty} q^k$ eine konvergierende Majorante (geometrische Reihe) und $\sum_{k=1}^{\infty} a_k$ konvergiert absolut nach dem Majorantenkriterium.

b. Im Fall $\left| \frac{a_{k+1}}{a_k} \right| \geq 1$ folgt $|a_{k+1}| \geq |a_k|$, womit die Absolutbeträge monoton wachsen und keine Nullfolge bilden können.

Zur Limesform:
Beweisführung erfolgt analog zur Limesform des Wurzelkriteriums.

□

Man kann das Quotientenkriterium auch aus dem Wurzelkriterium herleiten, allerdings nicht umgekehrt. Daraus folgt, dass letzteres das stärkere Kriterium ist und man die absolute Konvergenz mancher Reihen nur mit dem Wurzelkriterium zeigen kann. In den meisten Fällen sind aber beide Kriterien einsetzbar und man entscheidet sich für jenes, bei dem sich die einfacheren Terme ergeben. Besonders für Reihen, in denen Fakultäten vorkommen, ist das Quotientenkriterium gut geeignet.

Beispiel 3.63 (Quotientenkriterium)

Die Reihe $\sum_{k=1}^{\infty} \frac{k^5}{k!}$ ist absolut konvergent nach Quotientenkriterium, denn es gilt:

$$\lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| = \lim_{k\to\infty} \left|\frac{(k+1)^5\cdot k!}{k^5\cdot(k+1)!}\right| = \lim_{k\to\infty} \left|\left(\frac{k+1}{k}\right)^5 \cdot \frac{1}{k+1}\right| = 1^5 \cdot 0 = 0 < 1.$$

Wir haben an dieser Stelle genug Vorarbeit geleistet, um Stellenwertsysteme zu verstehen. Also zum Beispiel das typische Dezimalsystem, in dem gilt $1/4 = 0.25$. Dieses Stellenwertsystem basiert auf der natürlichen Zahl 10. Wir können aber auch beliebige andere Basen wählen. Prominente Beispiele sind das Binärsystem (Basis 2) und das Hexadezimalsystem (Basis 16). Da die Stellenwertsysteme für die restliche Veranstaltung nicht relevant sind, haben wir diesen Teil in einen Exkurs Stellenwertsysteme ausgelagert.

3.5 Potenzreihen

Potenzreihen sind eine wichtige Klasse von Reihen, die einige angenehme Eigenschaften haben, sodass sie sich einfach handhaben lassen. Gleichzeitig kann man mit Hilfe von Potenzreihen viele Funktionen approximieren, was wir in späteren Kapiteln noch weiter ausführen werden. An dieser Stelle werden nur die ersten Grundlagen der Potenzreihen eingeführt.

Definition 3.64 (Potenzreihe)

Sei $(a_k)_{k\in\mathbb{N}_0}$ eine Folge und $x, x_0 \in \mathbb{R}$, dann ist eine *Potenzreihe* $P(x, x_0)$ mit *Entwicklungspunkt* x_0 definiert als

$$P(x, x_0) \;=\; \sum_{k=0}^{\infty} a_k \cdot (x - x_0)^k \;=\; a_0 + a_1\,(x - x_0) + a_2(x - x_0)^2 + \ldots$$

Mit dem häufig gewählten Entwicklungspunkt $x_0 = 0$ ergibt sich die Potenzreihenform:

$$P(x) := P(x, 0) \;=\; \sum_{k=0}^{\infty} a_k \cdot x^k \;=\; a_0 + a_1 x + a_2 x^2 + \ldots$$

Beispiel 3.65 (Potenzreihen)

- Geometrische Reihe, $a_k = 1$:

$$P_{\text{geo}}(x) = \sum_{k=0}^{\infty} x^k$$

- Exponentialreihe, $a_k = \frac{1}{k!}$:

$$P_{\exp}(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

- Logarithmusreihe mit Entwicklungspunkt $x_0 = 1$ und $a_k = \frac{(-1)^{k+1}}{k}$, $a_0 = 0$:

$$P_{\log}(x, 1) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} (x - 1)^k$$

Die Exponentialreihe werden wir in diesem Abschnitt noch gesondert betrachten, die Logarithmusreihe wird uns erst in einem späteren Kapitel wieder begegnen.

Für Potenzreihen fragen wir uns nicht nur, ob die Reihe konvergiert, sondern für welche $x \in \mathbb{R}$ sie konvergiert. Durch die spezielle Struktur der Potenzreihe ergeben sich außerdem Vereinfachungen für die Konvergenzuntersuchung. Betrachten wir dazu folgendes Beispiel:

Beispiel 3.66 (Konvergenz von Potenzreihen)

Wir wollen eine beliebige Potenzreihe auf Konvergenz prüfen:

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k$$

Mit dem Wurzelkriterium und dem Quotientenkriterium ergibt sich, dass

$$\lim_{k\to\infty} \sqrt[k]{|a_k (x - x_0)^k|} = |x - x_0| \lim_{k\to\infty} \sqrt[k]{|a_k|}$$

und

$$\lim_{k\to\infty} \left| \frac{a_{k+1}(x - x_0)^{k+1}}{a_k (x - x_0)^k} \right| = |x - x_0| \lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right|$$

jeweils kleiner als 1 sein müssen, damit die Reihe (absolut) konvergiert. Die Limesbildung erfolgt in beiden Fällen nur über die Folge (a_k).

Nehmen wir an, wir erhalten zum Beispiel

$$\lim_{k\to\infty} \sqrt[k]{|a_k|} = 20.$$

Dann vereinfacht sich das Wurzelkriterium zu $20\,|x - x_0| < 1$, also konvergiert die Reihe in diesem Fall für $x_0 - \frac{1}{20} < x < x_0 + \frac{1}{20}$.

Wenn $\lim\limits_{k\to\infty} \sqrt[k]{|a_k|}$ bestimmt gegen $\pm\infty$ divergiert, dann divergiert auch die Potenzreihe für $x \neq x_0$. Am Entwicklungspunkt $x = x_0$ konvergiert offensichtlich jede Potenzreihe.

Die Menge der x-Werte, für die eine Potenzreihe konvergiert, ist stets ein zusammenhängendes Intervall, da der folgende Satz gilt:

Satz 3.67 (Konvergenz von Potenzreihen)

a. Konvergiert eine Potenzreihe $P(x, x_0)$ in einem Punkt c, so konvergiert sie in jedem Punkt x mit $|x - x_0| < |c - x_0|$ absolut.

b. Wenn $P(x, x_0)$ in einem Punkt c *nicht* absolut konvergiert (also nur gewöhnlich konvergiert oder divergiert), dann divergiert $P(x, x_0)$ für alle $|x - x_0| > |c - x_0|$.

▼ Beweis

Wir führen den Beweis für $x_0 = 0$, da sich andere Entwicklungspunkte mit der Substitution $x \leftarrow x - x_0$ und $c \leftarrow c - x_0$ auf diesen Fall zurückführen lassen.

a. Da $P(c)$ konvergiert, müssen die Reihenglieder a_k eine Nullfolge bilden. Damit gibt es eine obere Schranke $K \in \mathbb{R}$ mit $\left|a_k c^k\right| \leq K$ für alle $k \in \mathbb{N}$. Dann ist aber auch

$$\left|a_k c^k\right| \;=\; \left|a_k c^k \cdot \frac{x^k}{x^k}\right| \;=\; \left|a_k x^k\right| \cdot \left|\left(\frac{c}{x}\right)^k\right| \;\leq\; K.$$

Daraus folgt mit $|x| < |c|$

$$\left|a_k x^k\right| \leq q^k K \quad \text{mit} \quad q = \left|\frac{x}{c}\right| < 1.$$

Die Reihe $K \sum_{k=0}^{\infty} q^k$ konvergiert, da sie ein Vielfaches der geometrischen Reihe ist. Also konvergiert nach Majorantenkriterium die betragsmäßig kleinere Reihe $P(x) = \sum_{k=0}^{\infty} a_k x^k$ absolut.

b. Für den zweiten Teil des Satzes muss, wenn $P(c)$ nicht absolut konvergiert, nach Wurzelkriterium für fast alle k gelten:

$$|c| \sqrt[k]{|a_k|} \geq 1.$$

Daraus folgt, dass für $|x| > |c|$ auch gilt

$$1 \;\leq\; |c| \sqrt[k]{|a_k|} \;<\; |x| \sqrt[k]{|a_k|}.$$

Somit kann $P(x)$ für $|x| > |c|$ nach Wurzelkriterium nicht absolut konvergieren.

□

Den sich so ergebenden zusammenhängenden Konvergenzbereich charakterisiert man über den sogenannten *Konvergenzradius*.

Definition 3.68 (Konvergenzradius)

Sei $P(x, x_0)$ eine Potenzreihe. Wenn für ein $r \in \mathbb{R}_{\geq 0}$ gilt, dass $P(x, x_0)$ für alle $|x - x_0| < r$ konvergiert und für alle $|x - x_0| > r$ divergiert, dann nennt man r den *Konvergenzradius* von $P(x, x_0)$.

Allgemein kann man den Konvergenzradius als Supremum definieren. Es gilt $r = \sup \{y = |x - x_0| | \text{Reihe konvergiert für } x\}$.

Beispiel 3.69 (Konvergenzradius)

Die Potenzreihe $P(x) = \sum_{k=1}^{\infty} \frac{x^k}{k}$ mit $a_k = \frac{1}{k}$ und $x_0 = 0$) ist für $x = -1$ identisch zur alternierenden harmonischen Reihe, welche konvergiert (siehe Beispiel 3.51). Daraus folgt nach dem letzten Satz, dass $P(x)$ für alle $|x| < |-1| = 1$ absolut konvergiert.

Für $x = 1$ ist $P(x)$ identisch zur harmonischen Reihe, welche divergiert (siehe (Beispiel 3.43). Daher divergiert $P(x)$ für alle $|x| > 1$.

Damit ist $r = 1$ der Konvergenzradius von $P(x)$.

Im Folgenden werden wir noch die Exponentialreihe als sehr prominente Potenzreihe genauer betrachten, welche uns für den Rest der Veranstaltung begleiten wird.

Die Exponentialreihe

Definition 3.70 (Exponentialreihe)

Wir bezeichnen die Reihe

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$$

als *Exponentialreihe*. Wir definieren darüber hinaus die *Eulersche Zahl* als

$$e := \exp(1) = \sum_{k=0}^{\infty} \frac{1}{k!} = 1 + \frac{1}{1} + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \ldots$$

Satz 3.71 (Konvergenz der Exponentialreihe)

Für jedes $x \in \mathbb{R}$ ist $\exp(x)$ absolut konvergent.

▼ Beweis

Die Exponentialreihe ist eine Potenzreihe mit $a_k = \frac{1}{k!}$ und $x_0 = 0$. Damit folgt nach Quotientenkriterium analog zu Beispiel 3.66:

$$|x| \lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| = |x| \lim_{k\to\infty} \frac{k!}{(k+1)!} = |x| \lim_{k\to\infty} \frac{1}{k+1} = 0 < 1.$$

Damit ist die Exponentialreihe absolut konvergent für alle $x \in \mathbb{R}$.

□

Die Exponentialreihe definiert die Exponentialfunktion $\exp(x)$, für die es zahlreiche Anwendungen in den Naturwissenschaften und auch in Gesellschafts- und Wirtschaftwissenschaften gibt. Die Exponentialreihe ermöglicht die Approximation der Werte der Exponentialfunktion durch endliche Summation. Mit ihrer Hilfe werden z.B. Wachstums- und Zerfallsprozesse beschrieben.

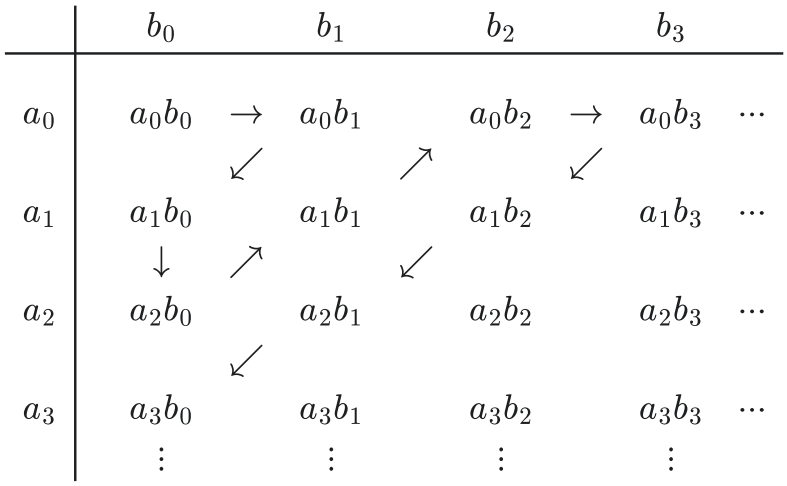
Eine fundamentale Eigenschaft der Exponentialreihe ist

$$\exp(x+y) = \exp(x)\exp(y).$$

In manchen Lehrbüchern wird sie sogar hauptsächlich darüber definiert. Aus dieser lassen sich nahezu alle weiteren Eigenschaften der Reihe ableiten. Für den Beweis benötigen wir eine Verallgemeinerung des Distributivgesetzes (“jeder mit jedem”), damit wir unendliche Summen miteinander multiplizieren können. Dabei taucht dasselbe Problem auf, welches uns bereits bei der Abzählbarkeit von $\mathbb{Q}$ begegnet ist (Kapitel 2.6): Wenn wir zwei unendliche Summen multiplizieren, ergibt sich:

$$\begin{aligned}(a_0 + a_1 + a_2 + \ldots)\,(b_0 + b_1 + b_2 + \ldots) = \\ a_0b_0 + a_0b_1 + a_0b_2 + \ldots + a_1b_0 + a_1b_1 + a_1b_2 + \ldots + a_2b_0 + a_2b_1 + a_2b_2 + \ldots + \ldots\end{aligned}$$

Hierbei verbirgt sich hinter jedem … eine unendliche Anzahl von Summanden. Bei einer Zuordnung der natürlichen Zahlen in dieser Reihenfolge würden bereits vor dem Term a_1b_0 alle natürlichen Zahlen zugewiesen sein. Die Lösungsidee ist ebenfalls dieselbe, welche wir schon bei den rationalen Zahlen angewendet haben: wir stellen uns die Summanden der ersten Reihe als Zeilen und die der zweiten als Spalten einer Tabelle vor. Die sich ergebenden Produkte innerhalb der Tabelle nummerieren wir dann diagonal durch:



Dabei ist die Summe der Indizes in jeder Diagonale konstant. Es ergibt sich also die Summationsreihenfolge:

$$(a_0b_0) + (a_0b_1 + a_1b_0) + (a_0b_2 + a_1b_1 + a_2b_0) + ... + (a_0b_n + a_1b_{n-1} + a_2b_{n-2} + ... + a_nb_0) + ...$$

Mit dieser Abzählung können wir jedem Summanden des Produkts der zwei Reihen genau eine natürliche Zahl zuweisen. Die Frage ist nun, wann diese Produktreihe konvergiert und ob das Produkt auch dem Produkt der Grenzwerte der beiden Einzelreihen entspricht? Der französische Mathematiker Augustin-Louis Cauchy hat für dieses Reihenprodukt erneut seinen Namen hergegeben.

Satz 3.72 (Cauchy-Produkt)

Seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{k=0}^{\infty} b_k$ absolut konvergente Reihen. Für $n \in \mathbb{N}$ sei

$$c_n := \sum_{k=0}^{n} a_k \cdot b_{n-k} = a_0 b_n + a_1 b_{n-1} + \ldots + a_n b_0.$$

Dann ist die Reihe

$$\sum_{k=0}^{\infty} c_k = \left(\sum_{k=0}^{\infty} a_k \right) \cdot \left(\sum_{k=0}^{\infty} b_k \right)$$

absolut konvergent.

▼ Beweis

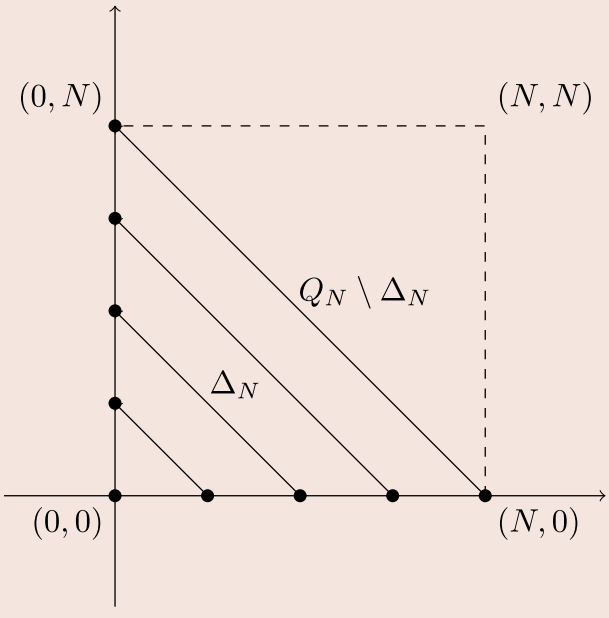
Wir definieren

$$c_n := \sum_{\substack{k+l=n \\ k,l \in \mathbb{N}_0}} a_k \cdot b_l \qquad \text{und} \qquad C_N := \sum_{n=0}^{N} c_n = \sum_{n=0}^{N} \sum_{\substack{k+l=n \\ k,l \in \mathbb{N}_0}} a_k \cdot b_l,$$

sowie die beiden Mengen

$$Q_N := \{(k,l) \in \mathbb{N}_0 \times \mathbb{N}_0 \mid k \leq N, l \leq N\},$$
$$\Delta_N := \{(k,l) \in \mathbb{N}_0 \times \mathbb{N}_0 \mid k + l \leq N\}.$$

Die Mengen werden in der unten stehenden Abbildung dargestellt. Offensichtlich gilt $Q_{\lfloor N/2 \rfloor} \subset \Delta_N \subset Q_N$ für $N \geq 2$.



Die Multiplikation der Teilsummen $A_N := \sum_{n=0}^{N} a_n$ und $B_N := \sum_{n=0}^{N} b_n$ liefert

$$A_N \cdot B_N = \sum_{k,l \in Q_N} a_k \cdot b_l.$$

Da $\Delta_N \subset Q_N$ gilt $A_N B_N - C_N = \sum_{k,l \in Q_N \setminus \Delta_N} a_k \cdot b_l.$

Für die Teilsummen $A_N^* := \sum_{n=0}^{N} |a_n|$ und $B_N^* := \sum_{n=0}^{N} |b_n|$ erhält man

$$A_N^* B_N^* = \sum_{k,l \in Q_N} |a_k| \cdot |b_l|.$$

Ferner folgt aus $Q_{\lfloor N/2 \rfloor} \subset \Delta_N \;\Rightarrow\; Q_N \setminus \Delta_N \subset Q_N \setminus Q_{\lfloor N/2 \rfloor}$, womit gilt

$$|A_N B_N - C_N| \leq \sum_{k,l \in Q_N \setminus Q_{\lfloor N/2 \rfloor}} |a_k|\,|b_l| = A_N^* B_N^* - A_{\lfloor N/2 \rfloor}^* B_{\lfloor N/2 \rfloor}^*.$$

Da beide Reihen absolut konvergent sind, konvergiert $A_N^* B_N^*$ und ist also eine Cauchy-Folge, sodass für $N \to \infty$ die obige Differenz gegen 0 konvergiert und damit auch

$$\lim_{N\to\infty} C_N = \lim_{N\to\infty} A_N B_N = \lim_{N\to\infty} A_N \cdot \lim_{N\to\infty} B_N.$$

Damit wurde gezeigt, dass $\sum_{n=0}^{\infty} c_n$ konvergiert. Die absolute Konvergenz folgt aus

$$\sum_{n=0}^{\infty} |c_n| \leq \sum_{n=0}^{\infty} \sum_{k=0}^{n} |a_k| \cdot |b_{n-k}|\,.$$

□

Die Konvergenz des Cauchy-Produkts zweier absolut konvergenter Reihen gilt übrigens erneut nur für absolut konvergente Reihen, aber nicht zwangsweise für konvergente Reihen.

Wenden wir uns nun den wichtigsten Eigenschaften der Exponentialreihe zu:

Satz 3.73 (Eigenschaften der Exponentialreihe)

Für die Exponentialreihe $\exp(x)$ gelten folgende Eigenschaften:

a. $\forall x, y \in \mathbb{R}: \exp(x+y) = \exp(x) \cdot \exp(y)$

b. $\forall x \in \mathbb{R}: \exp(-x) = \frac{1}{\exp(x)}$

c. $\forall x \in \mathbb{R}: \exp(x) > 0$

d. $\forall n \in \mathbb{Z}: \exp(n) = e^n$

▼ Beweis

a. Wir bilden das Cauchy-Produkt der beiden absolut konvergenten Reihen $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ und $\sum_{k=0}^{\infty} \frac{y^k}{k!}$. Für c_n gilt dann

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

Die einzelnen Umformungen folgen aus dem binomischen Lehrsatz. Wir erhalten nach obiger Umformung

$$\exp(x) \cdot \exp(y) = \sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} = \exp(x+y).$$

b. Aufgrund von **(a)** gilt

$$\exp(x) \cdot \exp(-x) = \exp(x-x) = \exp(0) = 1.$$

Daraus folgt $\exp(x) \neq 0$ und $\exp(-x) = \frac{1}{\exp(x)}$.

c. Für $x \geq 0$ gilt

$$\exp(x) = 1 + x + \frac{x^2}{2} + \ldots \geq 1 > 0,$$

da $\frac{x^k}{k!} > 0$. Für $x < 0$ folgt $-x > 0$ und damit $\exp(x) = \frac{1}{\exp(-x)} > 0$, wie aus **(a)** und **(b)** folgt.

d. Wir zeigen per Induktion über $n \in \mathbb{N}_0$, dass $\exp(n) = e^n$ gilt.
Induktionsanfang $n = 0 : \exp(0) = 1 = e^0$.
Induktionsvoraussetzung: Wir nehmen an, die Behauptung gilt für ein $n \in \mathbb{N}_0$
Induktionsschritt $n \to n+1$:
Sei $\exp(n) = e^n$ und $\exp(1) = e$ (laut Definition 3.70). Nach **(a)** gilt dann

$$\exp(n+1) = \exp(n) \cdot \exp(1) = e^n \cdot e^1 = e^{n+1}.$$

Damit ist der Beweis für $n \geq 0$ komplett. Mittels **(b)** gilt für $n \in \mathbb{N}$

$$\exp(-n) = \frac{1}{\exp(n)} = \frac{1}{e^n} = e^{-n}.$$

Damit ist der Satz für alle $n \in \mathbb{Z}$ bewiesen.

□

Teil (d) lässt sich auf $x \in \mathbb{R}$ ausdehnen. Es ist dann $e^x = \exp(x) = \exp(n + \delta) = e^n \cdot \exp(\delta)$ mit $\delta = x - n$, $n = \lfloor x \rfloor \in \mathbb{Z}$. Diese Darstellung kann man nutzen, um reellwertige Exponenten für eine allgemeine Basis, nicht nur der Basis e, zu definieren. Wir werden dies später weiter vertiefen.

Wir sind bisher noch nicht auf den zweiten Teil der Definition 3.70 eingegangen, der Definition der Eulerschen Zahl e. Wir hatten e bereits im Zusammenhang mit den Folgen (e_n) und $(\bar{e}_n)$ (Beispiel 3.15 und 3.20) im Kapitel zu Folgen erwähnt. Die Folgen waren definiert als

$$e_n = \left(1 + \frac{1}{n}\right)^n \qquad \bar{e}_n = \left(1 + \frac{1}{n}\right)^{n+1}$$

Es ist alles andere als offensichtlich, dass die Grenzwerte dieser beiden Folgen und der Reihenwert von $\exp(1)$ die gleiche Zahl e ergeben sollen. Dies werden wir im folgenden Satz beweisen.

Satz 3.74 (Exponentialfunktion als Folgengrenzwert)

Es gilt für alle $x \in \mathbb{R}$

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n.$$

Insbesondere gilt mit $x = 1$ für die Eulersche Zahl:

$$e = \sum_{k=0}^{\infty} \frac{1}{k!} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n.$$

▼ Beweis

Zur Übersichtlichkeit beweisen wir nur den Fall $x = 1$ (also den zweiten Teil der Aussage) ausführlich und klären anschließend, wie wir den allgemeinen Fall beweisen würden. Betrachten wir also den zweiten Teil der Aussage: Dafür nutzen wir das Sandwich-Theorem und zeigen, dass die Folge der Partialsummen

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

zwischen den zwei Folgen (e_n) und $(\bar{e}_n)$ liegt, mit

$$e_n = \left(1 + \frac{1}{n}\right)^n, \qquad \bar{e}_n = \left(1 + \frac{1}{n}\right)^{n+1}$$

Der Konvergenzbeweis, sowie der Nachweis, dass der Grenzwert beider Folgen identisch ist, wurde in Beispiel 3.15 und Beispiel 3.20 geführt.

1. Teil ($e_n \leq s_n$):

Für diesen müssen wir lediglich den binomischen Lehrsatz (Satz 2.44) auf die Folge (e_n) anwenden:

$$
\begin{aligned}
e_n &= \left(1+\frac{1}{n}\right)^n \\
&= \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} \\
&= \sum_{k=0}^{n} \frac{n(n-1)(n-2)\cdots(n-(k-1))}{k!} \frac{1}{n^k} \\
&= \sum_{k=0}^{n} \frac{n(n-1)(n-2)\cdots(n-(k-1))}{n^k} \frac{1}{k!} \\
&= \sum_{k=0}^{n} \underbrace{\frac{n}{n}\frac{n-1}{n}\frac{n-2}{n}\cdots\frac{n-(k-1)}{n}}_{\leq 1} \frac{1}{k!} \\
&\leq \sum_{k=0}^{n} \frac{1}{k!} = s_n
\end{aligned}
$$

2. Teil ($\bar{e}_n \geq s_n$):

Der zweite Teil ist schon deutlich anspruchsvoller zu zeigen. Wir nutzen hier neben dem binomischen Lehrsatz noch verschiedene Tricks, unter anderem die Bernoullische Ungleichung (Satz 2.45) und Indexverschiebungen. Letztere nutzen aus, dass die folgenden Summen identisch sind:

$$
\sum_{k=0}^{n} a_k = a_0 + a_1 + a_2 + \ldots + a_n = \sum_{k=1}^{n+1} a_{k-1} = a_0 + a_1 + a_2 + \ldots + a_n
$$

In der Summation wird der Index um Eins erhöht und im Reihenterm um Eins verringert, wodurch sich der Summenwert nicht ändert.

Fangen wir an:

$$\bar{e}_n = \left(1+\frac{1}{n}\right)^{n+1}$$

↓ Binomischer Lehrsatz

$$= \sum_{k=0}^{n+1} \binom{n+1}{k} \frac{1}{n^k}$$

↓ Ersten Summanden aus der Summe holen

$$= 1 + \sum_{k=1}^{n+1} \binom{n+1}{k} \frac{1}{n^k}$$

↓ Indexverschiebung

$$= 1 + \sum_{k=0}^{n} \binom{n+1}{k+1} \frac{1}{n^{k+1}}$$

$$= 1 + \sum_{k=0}^{n} \frac{(n+1)n(n-1)(n-2)\cdots(n-(k-1))}{(k+1)!} \frac{1}{n^{k+1}}$$

↓ Nenner der Brüche vertauschen

$$= 1 + \sum_{k=0}^{n} \frac{(n+1)n(n-1)(n-2)\cdots(n-(k-1))}{n^{k+1}} \frac{1}{(k+1)!}$$

↓ Alle Faktoren im ersten Zähler sind $\geq n-(k-1)$

$$\geq 1 + \sum_{k=0}^{n} \frac{(n-(k-1))^{k+1}}{n^{k+1}} \frac{1}{(k+1)!}$$

$$= 1 + \sum_{k=0}^{n} \left(1-\frac{k-1}{n}\right)^{k+1} \frac{1}{(k+1)!}$$

↓ Bernoullische Ungleichung $(1-x)^n \geq 1-nx$

$$\geq 1 + \sum_{k=0}^{n} \left(1-\frac{(k-1)(k+1)}{n}\right) \frac{1}{(k+1)!}$$

$$= 1 + \sum_{k=0}^{n} \frac{1}{(k+1)!} - \frac{1}{n} \sum_{k=0}^{n} \frac{(k-1)(k+1)}{(k+1)!}$$

↓ Indexverschiebung in 1. Summe

$$= \left(1 + \sum_{k=1}^{n+1} \frac{1}{k!}\right) + \left(-\frac{1}{n} \sum_{k=0}^{n} \frac{(k-1)}{k!}\right)$$

↓ Aufnahme von $1 = 1/0!$ in 1. Summe

$$= \left(\sum_{k=0}^{n+1} \frac{1}{k!}\right) + \left(-\frac{1}{n} \sum_{k=0}^{n} \frac{(k-1)}{k!}\right)$$

↓ letzten Summanden aus 1. Summe weglassen

$$\geq \left(\sum_{k=0}^{n} \frac{1}{k!}\right) + \underbrace{\left(-\frac{1}{n} \sum_{k=0}^{n} \frac{(k-1)}{k!}\right)}_{\geq 0}$$

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

Dass der 2. geklammerte Term nicht negativ ist, ist natürlich nicht offensichtlich. Das zeigen wir nun separat:

$$
\begin{aligned}
-\frac{1}{n}\sum_{k=0}^{n}\frac{(k-1)}{k!} \quad &\downarrow \text{Herausziehen des } k=0 \text{ Summanden} \\
&= -\frac{1}{n}\left(-1+\sum_{k=1}^{n}\frac{(k-1)}{k!}\right) \\
&= -\frac{1}{n}\left(-1+\sum_{k=1}^{n}\left(\frac{k}{k!}-\frac{1}{k!}\right)\right) \\
&= -\frac{1}{n}\left(-1+\sum_{k=1}^{n}\left(\frac{1}{(k-1)!}-\frac{1}{k!}\right)\right) \\
&= -\frac{1}{n}\left(-1+\left(\frac{1}{0!}-\frac{1}{1!}+\frac{1}{1!}-\frac{1}{2!}+\frac{1}{2!}\mp\ldots-\frac{1}{(n-1)!}+\frac{1}{(n-1)!}-\frac{1}{n!}\right)\right) \\
&\quad \downarrow \text{Ausnutzung der Teleskopsumme} \\
&= -\frac{1}{n}\left(-1+\left(\frac{1}{0!}-\frac{1}{n!}\right)\right) \\
&= -\frac{1}{n}\left(-1+1-\frac{1}{n!}\right) \\
&= -\frac{1}{n}\left(-\frac{1}{n!}\right) \\
&= \frac{1}{n\cdot n!} > 0
\end{aligned}
$$

Damit haben wir mit (e_n) und $(\bar{e}_n)$ zwei Folgen, die gegen den gleichen Grenzwert konvergieren (Beispiel 3.15 und 3.20) und die Teilsummenfolge von oben und unten begrenzen. Damit gilt nach dem Sandwich-Theorem:

$$\lim_{N\to\infty} e_n = \lim_{N\to\infty} s_n = \lim_{N\to\infty} \bar{e}_n.$$

Da wir definiert haben, dass der Reihenwert $\lim\limits_{N\to\infty} s_n$ die Eulersche Zahl ist, konvergieren auch die beiden Folgen gegen e.

Fall $x \in \mathbb{N}$:

Wir haben bisher nur den zweiten Teil der Aussage bewiesen. Die allgemeine Aussage ist sehr anschaulich für eine natürliche Zahl $x = m \in \mathbb{N}$:
Für eine konvergente Folge konvergiert jede Teilfolge gegen den gleichen Grenzwert. Betrachten wir also die Teilfolge (e_{mn}), also

$$e_{mn} = \left(1+\frac{1}{mn}\right)^{mn},$$

dann muss diese gegen denselben Grenzwert wie e_n, also gegen e konvergieren. Wenn wir nun für die allgemeinere Folge der ersten Aussage

$$\left(1+\frac{m}{n}\right)^{n}$$

ebenfalls die mn-te Teilfolge betrachten, ergibt sich:

$$\lim_{n\to\infty}\left(1+\frac{m}{mn}\right)^{mn}=\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{mn}=\left(\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{n}\right)^{m}=e^{m}.$$

Fall $x\in\mathbb{R}$:

Für $x=0$ ist die Aussage trivial. Für $x<0$ lässt sich die Argumentation auf einen Fall $x>0$ zurückführen, denn es gilt (Beweis zur Übung):

$$\lim_{n\to\infty}\left(1-\frac{x}{n}\right)^{n}=\frac{1}{\lim\limits_{n\to\infty}\left(1+\frac{x}{n}\right)^{n}}.$$

Wir müssen also den Satz nur für $x>0$ beweisen. Wenn Sie aber erneut den Beweis für $x=1$ durchgehen, werden Sie feststellen, dass Sie lediglich in den Termen x^k hinzufügen müssen. Alle Umformungen bleiben wahr (da $x^k>0$), die Terme werden nur etwas länger.

□

▶ Exponentialreihe und Exponentialfolge(n) visualisieren

Es gibt noch weitere Darstellungen der Eulerschen Zahl z.B. $e=\lim\limits_{n\to\infty}\frac{n}{\sqrt[n]{n!}}$, die wir hier aber nicht beweisen werden. Manche Beweise werden mit der Folgendarstellung der Exponentialreihe deutlich einfacher. Zum Beispiel könnten wir erneut zeigen, dass gilt

$$\begin{aligned}\exp(x)\exp(y)&=\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^{n}\lim_{n\to\infty}\left(1+\frac{y}{n}\right)^{n}\\&=\lim_{n\to\infty}\left(\left(1+\frac{x}{n}\right)\left(1+\frac{y}{n}\right)\right)^{n}\\&=\lim_{n\to\infty}\left(1+\frac{x+y}{n}+\frac{xy}{n^2}\right)^{n}\\&=\lim_{n\to\infty}\left(1+\frac{x+y}{n}\right)^{n}\\&=\exp(x+y)\end{aligned}\quad.$$

Den vorletzten Umformungsschritt müsste man noch mit dem binomischen Lehrsatz begründen, aber ansonsten würden Sie vermutlich zustimmen, dass dieser Beweis einfacher von der Hand geht als unsere erste Variante mit Cauchy-Produkten.

3.6 Komplexe Folgen und Reihen

Zum Abschluss des Kapitels wollen wir noch ein paar Worte zu Folgen und Reihen komplexer Zahlen verlieren. Besonders die komplexe Exponentialreihe wird uns noch häufiger begegnen, da diese in der Praxis häufig genutzt wird. Welche Eigenschaften aus Kapitel 3 können wir auch auf komplexe Folgen

und Reihen anwenden und welche nicht? Der Unterschied zwischen $\mathbb{R}$ und $\mathbb{C}$ ist, dass wir zwei reelle Zahlen vergleichen (ordnen) können, zwei komplexe Zahlen jedoch nicht. Da es bei Folgen und Reihen immer wieder um Größenvergleiche ging, liegt der Schluss nahe, dass die meisten Sätze im Komplexen nicht anwendbar sind. Allerdings fällt bei genauerer Betrachtung auf, dass die meisten Aussagen lediglich eine Menge erfordern, auf der eine Metrik definiert ist (sogenannte *metrische Räume*). Somit können wir fast alle Sätze auch auf komplexe Folgen und Reihen anwenden. Sätze, in denen betragsfreie Größenvergleiche vorkommen (wie im Sandwich-Theorem) können wir ebenfalls auf komplexe Reihen übetragen. Dazu betrachten wir den Betrag der zu vergleichenden Größen.

Als das prominenteste Beispiel betrachten wir die komplexe Exponentialreihe

Beispiel 3.75 (komplexe Exponentialreihe)

Die komplexe Exponentialreihe ist für $z \in \mathbb{C}$ definiert als

$$\exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

Die Reihe konvergiert für alle $z \in \mathbb{C}$ absolut nach dem Quotientenkriterium, da

$$\lim_{k\to\infty} \left| \frac{|z|^{k+1} k!}{|z|^k (k+1)!} \right| = |z| \lim_{k\to\infty} \frac{k!}{(k+1)!} = |z| \lim_{k\to\infty} \frac{1}{k+1} = 0 < 1.$$

Sie sehen, dass der Beweis analog zum Konvergenzbeweis der reellwertigen Exponentialreihe funktioniert und genauso lassen sich alle Eigenschaften aus Satz 3.73 bis auf $\exp(x) > 0$ auch auf komplexwertige Argumente übertragen.

▶ Demo: Komplexe Exponentialreihe

Für komplexe Potenzreihen ist der Konvergenzradius tatsächlich der Radius eines Kreises in der komplexen Ebene, dies können Sie in der folgenden Demo für die komplexe geometrische Reihe ausprobieren:

▶ Demo: Konvergenzradius der komplexen geometrischen Reihe mit Entwicklungspunkt

4 Funktionen

Funktionen sind ein zentrales Hilfsmittel, um den Zusammengang zwischen abhängigen Größen formal zu beschreiben. Sie werden damit in praktisch allen Wissenschaftsbereichen eingesetzt. In der Informatik existiert das klassische Konzept der Funktion, die ein Programmstück beschreibt, das aus den Werten der Eingabeparameter einen Resultatwert berechnet. Die theoretische Informatik beschäftigt sich unter anderem damit, welche Funktionen überhaupt auf einem Rechner, wie er heute genutzt wird, berechenbar sind und wie hoch der Berechnungsaufwand ist. Zur Darstellung des Berechnungsaufwandes wird dieser oft in Abhängigkeit von einer Eingabegröße mit einigen typischen Funktionen verglichen. Neben dem Berechnungsaufwand spielt natürlich auch die Genauigkeit der Berechnungen eine Rolle. Im Rechner wird immer nur mit endlicher Genauigkeit gearbeitet. Deshalb müssen viele Elemente aus $\mathbb{R}$ durch rationale Zahlen approximiert werden, so dass fast alle Berechnungen streng genommen nur Approximationen sind. Der Approximationsgüte widmet sich die Numerik, die an der Grenze zwischen angewandter Mathematik und Informatik angesiedelt ist.

Definition 4.1 (Funktion, Definitionsbereich, Bildbereich, Bild)

Seien A und B zwei nichtleere Mengen. Eine *Funktion* f ist eine Vorschrift, die jedem Element $x \in A$ ein eindeutiges Element $y \in B$ zuordnet. Wir schreiben auch $f : A \to B$. Das zugeordnete Element $y \in B$ schreiben wir auch als $f(x)$.

Wir nennen A den *Definitionsbereich* der Funktion und B den *Bild-* oder *Zielbereich* der Funktion.

Wir nennen die Menge $f(A) \subseteq B$ mit $f(A) = \{f(x) \mid x \in A\}$ *Bildmenge* oder *Bild* von f.

Wir haben uns bisher bereits mit speziellen Funktionen beschäftigt, die von jeder natürlichen Zahl $n \in \mathbb{N}$ auf eine reelle Zahl $a_n \in \mathbb{R}$ abbilden. Diese speziellen Funktionen haben wir Folgen (und Reihen) genannt. In der Analysis beschäftigt man sich aber in den meisten Fällen mit Funktionen, deren Definitionsbereich die gesamten reellen Zahlen oder ein bestimmtes Teilintervall der reellen Zahlen ist.

Verinnerlichen Sie die Unterschiede von Definitionsbereich, Bildbereich und dem Bild von f. Wenn wir für eine Funktion f schreiben $f : \mathbb{R} \to \mathbb{R}$, dann ist für den Definitionsbereich entscheidend, dass die Funktion auch für alle $x \in \mathbb{R}$ definiert ist. Wir könnten also zum Beispiel nicht $f(x) = \frac{1}{x}$ auf ganz $\mathbb{R}$ definieren, da die Funktion für 0 nicht definiert wäre. Der Bildbereich dagegen darf mehr Elemente enthalten, als wir mit f tatsächlich erreichen. Hier können wir also für reellwertige Funktionen stets alle reellen Zahlen angeben, denn nicht jedes Element des Bildbereichs muss getroffen werden (das Bild von f ist eine Teilmenge der Bildmenge von f). Für $f(x) = x^2$ wäre es also vollkommen legitim zu schreiben $f : \mathbb{R} \to \mathbb{R}$, obwohl das *Bild* von f nur $f(\mathbb{R}) = \mathbb{R}_{\geq 0}$ ist. Achten Sie darauf, das Bild und den Bildbereich von f nicht gleichzusetzen.

Beispiel 4.2 (Funktionen)

- $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x$ ist die Identitätsfunktion. Die Bildmenge ist $f(\mathbb{R}) = \mathbb{R}$.
- $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = c$, $c \in \mathbb{R}$ ist eine konstante Funktion. Die Bildmenge ist $f(\mathbb{R}) = \{c\}$.
- $f : \mathbb{N} \to \mathbb{R}$ mit $f(n) = a_n$ ist eine Folge. Die Bildmenge ist $f(\mathbb{N}) = \{a_1, a_2, a_3, ...\}$.
- $f : \mathbb{R}_{\geq 0} \to \mathbb{R}$ mit $f(x) = \sqrt{x}$ ist die Quadratwurzelfunktion. Die Bildmenge ist $f(\mathbb{R}_{\geq 0}) = \mathbb{R}_{\geq 0}$.

Wir können Funktionen auch aus Teilbereichen zusammensetzen:

- $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = \begin{cases} 0 & \text{wenn } x < 0 \\ 1 & \text{wenn } x \geq 0 \end{cases}$ ist die *Heaviside-Funktion*.
- $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = \begin{cases} 1 & \text{wenn } x = 0 \\ \frac{1}{x} & \text{sonst.} \end{cases}$
- $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = \begin{cases} 2x & \text{für } x \in (2, \infty) \\ -x^2 & \text{für } x \in [-2, 2] \\ 0 & \text{für } x \in (-\infty, -2) \end{cases}$

Wir können auch sehr merkwürdige Funktionen definieren, die man sich gar nicht mehr richtig vorstellen kann:

- $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = \begin{cases} 1 & \text{wenn } x \in \mathbb{Q} \\ 0 & \text{sonst.} \end{cases}$ ist die *Dirichlet-Funktion*.

Auch wenn wir in dieser Veranstaltung stets Funktionen betrachten werden, die von Zahlen auf Zahlen abbilden, ist der Funktionsbegriff für allgemeine Mengen definiert und nichts hält uns im Allgemeinen davon ab, Funktionen zwischen beliebigen Mengen zu definieren:

- $f : \{\text{Montag, Dienstag, Mittwoch, Donnerstag, Freitag, Samstag, Sonntag}\} \to \{\text{mag ich, geht so, mag ich nicht}\}$ mit

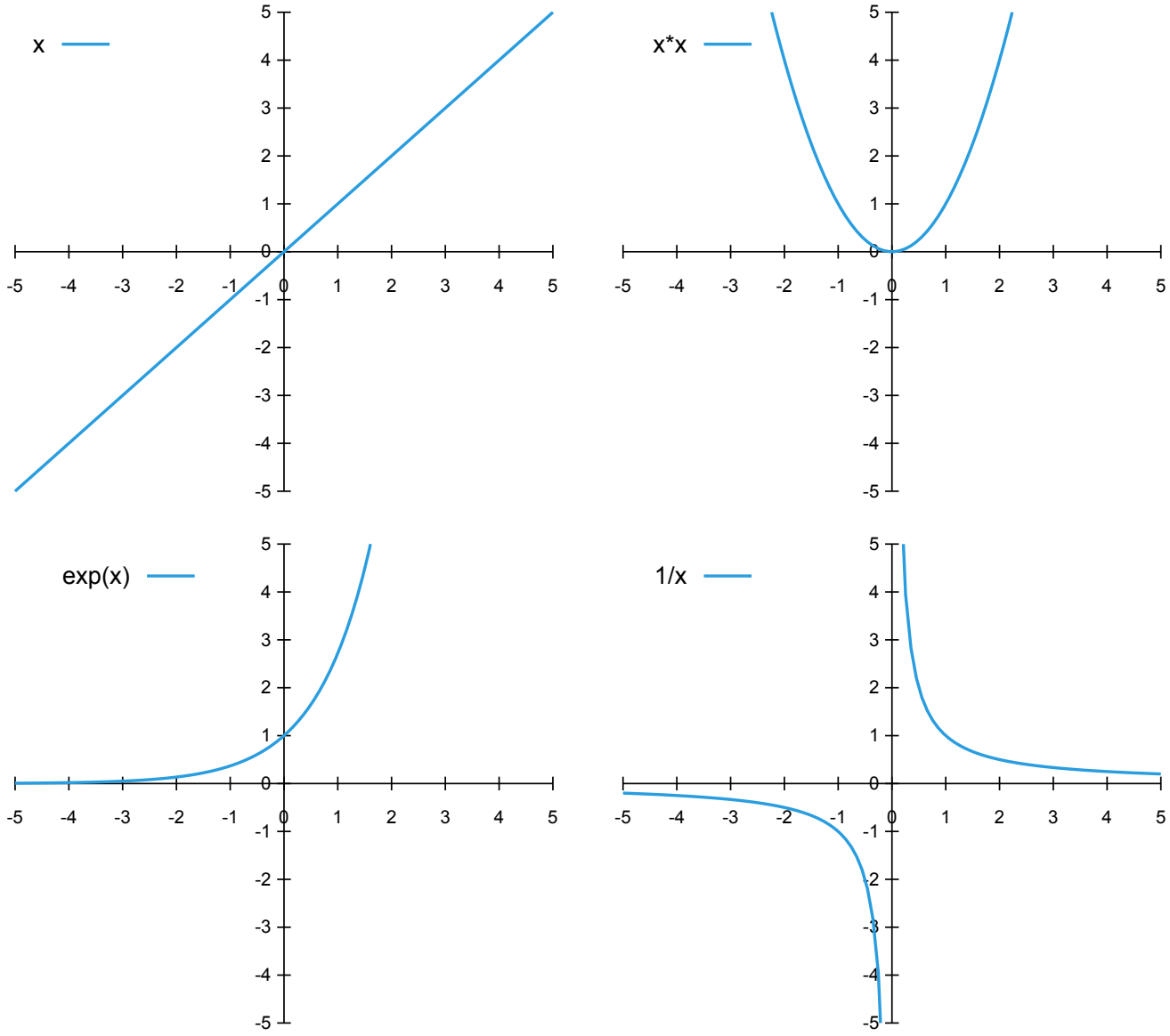
$$f(x) = \begin{cases} \text{mag ich} & \text{wenn } x \in \{\text{Freitag, Samstag, Sonntag}\} \\ \text{geht so} & \text{wenn } x \in \{\text{Montag, Donnerstag}\} \\ \text{mag ich nicht} & \text{sonst.} \end{cases}$$

Bei der Definition zusammengesetzter Funktionen muss darauf geachtet werden, dass man jedem x ein *eindeutiges* Ziel zuordnet. Es darf also niemals zwei verschiedene Bildwerte für dasselbe x geben.

Wenn Definitions- und Bildbereich ein Teil der reellen Zahlen sind, dann kann man eine Funktion auch mit einem sogenannten Graphen assoziieren. Dazu definiert man für eine Funktion $f : A \to B$ den Graphen als

$$\Gamma_f := \{(x, y) \in \mathbb{R} \times \mathbb{R} \mid x \in A \wedge y = f(x)\},$$

also die Menge aller Tupel $(x, f(x))$, welche wir zur Visualisierung in einem x-$f(x)$-Diagramm (wird auch häufig x-y-Diagramm genannt) darstellen können. Es gibt Funktionen, wie die oben definierte Dirichlet-Funktion, die sich nicht sinnvoll zeichnen lassen. Im Folgenden zeigen wir ein paar Beispiele für Funktionsgraphen.



Bevor wir zum Hauptthema dieses Kapitels (der Stetigkeit) kommen können, starten wir mit ein paar grundlegenden Definitionen für Funktionen, damit wir diese besser in ihren Eigenschaften beschreiben können.

4.1 Grundlegende Funktionseigenschaften

Definition 4.3 (Injektivität, Surjektivität, Bijektivität)

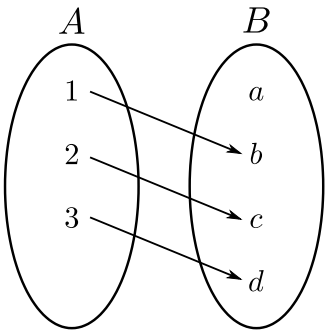
1. Eine Funktion $f : A \to B$ heißt *injektiv*, wenn zu jedem $y \in B$ höchstens ein $x \in A$ mit $f(x) = y$ gehört. In Kurzschreibweise:

$$\forall x_1, x_2 \in A : \Big(x_1 \neq x_2 \;\Rightarrow\; f(x_1) \neq f(x_2) \Big)$$

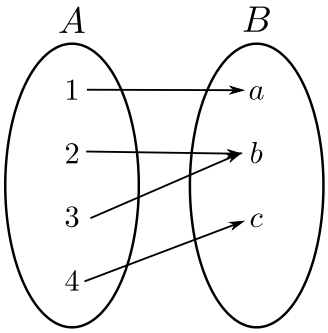
2. Eine Funktion $f : A \to B$ heißt *surjektiv*, wenn jedes $y \in B$ als Abbild mindestens eines $x \in A$ auftaucht. In Kurzschreibweise:

$$\forall y \in B \; \exists x \in A : f(x) = y$$

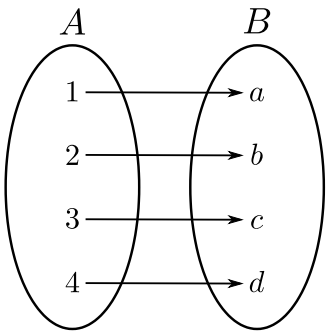
3. Eine Funktion $f : A \to B$ heißt *bijektiv*, wenn sie injektiv und surjektiv ist.



injektiv
nicht surjektiv



surjektiv
nicht injektiv



injektiv & surjektiv
$\Leftrightarrow$ bijektiv

Beispiel 4.4 (Bijektivität)

$f(x) = x^2$ ist bijektiv auf $A = \mathbb{R}_{\geq 0}$, $B = \mathbb{R}_{\geq 0}$

Um dies zu zeigen, müssen wir die Injektivität und die Surjektivität zeigen:

Injektivität:
Wenn $x \neq y$, dann muss entweder $x > y$ oder $x < y$ gelten. Daraus folgt aber sofort nach Satz 2.36, dass auch $x^2 > y^2$ bzw. $x^2 < y^2$ gelten muss, also $x^2 \neq y^2$. Damit ist die Injektivität bewiesen. Alternativ könnte man auch die Kontraposition der Injektivitätsbedingung betrachten:

$$f(x_1) = f(x_2) \Rightarrow x_1 = x_2$$

Diese Aussage könnte man mit der Eindeutigkeit der Quadratwurzeln beweisen.

Surjektivität:
Die Surjektivität besagt, dass für jedes $y \in \mathbb{R}_{\geq 0}$ ein $x \in \mathbb{R}_{\geq 0}$ existieren muss, sodass $x^2 = y$. Dies haben wir bereits mehrfach mit der Existenz von Quadratwurzeln gezeigt.

Damit ist die Funktion $f(x) = x^2$ bijektiv, wenn sie auf die positiven reellen Zahlen eingeschränkt wird.

Betrachten wir dagegen $f(x)$ als Funktion $f : \mathbb{R} \to \mathbb{R}$, dann ist sie weder injektiv noch bijektiv. Sie ist nicht injektiv, da $f(-x) = x^2 = f(x)$. Sie ist nicht surjektiv, da für $f(x) < 0$ kein x existiert mit $x^2 = f(x) < 0$ (Nach Satz 2.15(f)).

Für surjektive und bijektive Funktionen ist $B = f(A)$. Wir können den Definitionsbereich auch auf ein Teilintervall einschränken und sagen, dass f auf diesem Teilintervall bijektiv ist. Die Eigenschaften haben unter anderem eine Bedeutung bei der Lösung von Gleichungen der Form $f(x) = y$. Falls f injektiv ist, gibt es höchstens eine Lösung, falls f surjektiv ist, gibt es mindestens eine Lösung.

Die beiden folgenden Definitionen beschreiben die Kombination von Funktionen zur Erzeugung neuer, zusammengesetzter Funktionen. Wir werden später untersuchen, welche Eigenschaften der Funktionen sich auf die zusammengesetzten Funktionen übertragen.

Definition 4.5 (Rationale Operationen auf Funktionen)

Seien $f, g : A \to \mathbb{R}$ Funktionen und $c \in \mathbb{R}$. Dann sind die Funktionen $f + g : A \to \mathbb{R}$, $cf : A \to \mathbb{R}$, $fg : A \to \mathbb{R}$ definiert durch

$$\begin{aligned} (f+g)(x) &:= f(x) + g(x), \\ (cf)(x) &:= cf(x), \\ (f \cdot g)(x) &:= f(x)g(x). \end{aligned}$$

Sei $A' := \{x \in A \mid g(x) \neq 0\}$, dann ist die Funktion $\frac{f}{g} : A' \to \mathbb{R}$ definiert durch

$$\left(\frac{f}{g}\right)(x) := \frac{f(x)}{g(x)}.$$

Definition 4.6 (Komposition/Verkettung von Funktionen)

Seien $f : A \to \mathbb{R}$ und $g : B \to \mathbb{R}$ Funktionen und $f(A) \subseteq B$. Dann ist die Funktion $g \circ f : A \to \mathbb{R}$ definiert durch $(g \circ f)(x) = g(f(x))$.

Beispiel 4.7

Sei $f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ durch $f(x) = \exp(x)$ definiert und $g : \mathbb{R}_{\geq 0} \to \mathbb{R}$ durch $g(x) = \sqrt{x}$.
Dann ergibt sich die Komposition $(g \circ f)(x) = g(f(x)) = \sqrt{\exp(x)}$.

Eine besondere Komposition ergibt sich, wenn die beiden Funktionswirkungen sich gegenseitig aufheben, also $f(g(x)) = x$ gilt. In diesem Fall spricht man davon, dass g die Umkehrfunktion von f ist.

Definition 4.8 (Umkehrfunktion)

Für eine Funktion $f : A \to B$ nennen wir $f^{-1} : B \to A$ die *Umkehrfunktion* von f, wenn die folgenden beiden Eigenschaften gelten:

- $(f^{-1} \circ f)(x) = f^{-1}(f(x)) = x, \quad \forall x \in A$
- $(f \circ f^{-1})(x) = f(f^{-1}(x)) = x, \quad \forall x \in B$

Beispiel 4.9 (Umkehrfunktionen 1)

- Für

$$f : \mathbb{R} \to \mathbb{R}, f(x) = 2x + 4$$

ist die Umkehrfunktion

$$f^{-1} : \mathbb{R} \to \mathbb{R}, f^{-1}(x) = \frac{1}{2}x - 2,$$

da f^{-1} für den gesamten Bildbereich von f definiert ist und

$$f^{-1}(f(x)) = \frac{1}{2}(2x + 4) - 2 = f(f^{-1}(x)) = 2(\frac{1}{2}x - 2) + 4 = x.$$

- Für $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$ ist $g : \mathbb{R}_{>0} \to \mathbb{R}, g(x) = \sqrt{x}$ **keine** Umkehrfunktion, da zum Einen g nur auf $\mathbb{R}_{>0}$, also nicht auf dem gesamten Bildbereich $\mathbb{R}$ von f, definiert ist und zum Anderen $g(f(-2)) = 2 \neq -2$.

- Für $f : \mathbb{R}_{>0} \to \mathbb{R}, f(x) = x^2$ ist $g : \mathbb{R} \to \mathbb{R}_{>0}, g(x) = \sqrt{|x|}$ **keine** Umkehrfunktion, da zwar

$$f(g(x)) = \left(\sqrt{|x|}\right)^2 = |x| = x, \quad \forall x \in \mathbb{R}_{>0}$$

gilt, aber

$$g(f(x)) = \sqrt{|x^2|} = \sqrt{x^2} = |x|$$

nicht für den gesamten Definitionsbereich von g gleich x ist.

Wir können Umkehrfunktionen auf dem Bildbereich B nur dann und genau dann definieren, wenn f bijektiv ist.

Satz 4.10 (Bijektivität und Umkehrfunktionen)

Für eine Funktion $f : A \to B$ existiert die Umkehrfunktion genau dann, wenn f bijektiv ist.

▼ Beweis

Richtung 1: f ist bijektiv $\Rightarrow$ Umkehrfunktion existiert.

Die Bijektivität bedeutet, dass die Zuordnung von f in beide Richtungen eindeutig ist, zu jedem x existiert also ein eindeutiger Funktionswert $f(x)$ und zu jedem $f(x)$ ein eindeutiges x. Damit können wir die Zuordnung umkehren und somit die Umkehrfunktion definieren.

Richtung 2: Umkehrfunktion existiert $\Rightarrow$ f ist bijektiv.

Die Umkehrfunktion muss als Funktion eine eindeutige Zuordnung von $f(x)$ zu x sein. Daraus folgt, dass f injektiv ist. Außerdem kann die Umkehrfunktion nur dann auf B definiert sein, wenn es ein $x \in A$ gibt für alle $y \in B$ mit $f(x) = y$. Damit muss f surjektiv sein.

□

Häufig werden nicht-bijektive Funktionen f auf ein Teilintervall des Definitionsbereichs eingeschränkt, auf dem sie bijektiv sind. Für dieses gibt man anschließend eine Umkehrfunktion an.

Beispiel 4.11 (Umkehrfunktionen 2)

- Für $f : \mathbb{R} \to \mathbb{R}, f(x) = x^2$ existiert keine Umkehrfunktion, da f nicht injektiv ist.
- Für $f : \mathbb{R}_{>0} \to \mathbb{R}, f(x) = x^2$ existiert ebenfalls keine Umkehrfunktion, da als Bildbereich alle reellen Zahlen angegeben sind, das Bild der Funktion aber nur die positiven Zahlen abdeckt. f ist also nicht surjektiv. $g(x) = \sqrt{x}$ kann also keine Umkehrfunktion sein, da z.B. $g(-2)$ nicht definiert ist.
- Für $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}, f(x) = x^2$ existiert die Umkehrfunktion $g(x) = \sqrt{x}$, denn g ist für alle $x \in \mathbb{R}_{>0}$ definiert und es gilt $g(f(x)) = \sqrt{x^2} = |x| = x \; \forall x \in \mathbb{R}_{>0}$.

Der Unterschied zwischen der zweiten und dritten Definition im letzten Beispiel kann leicht übersehen werden. Man sollte sich daher angewöhnen, für die Definition von Umkehrfunktionen ganz genau auf Definitions- und Bildbereich zu achten. Übrigens: Wenn wir wissen, dass f bijektiv ist, genügt es eine der beiden Bedingungen aus Definition 4.8 für eine mögliche Umkehrfunktion zu prüfen, da die zweite dann aus der ersten folgt.

Als Nächstes übertragen wir ein paar Definitionen der Folgen auf allgemeine Funktionen.

Definition 4.12 (Beschränktheit)

Eine Funktion $f : A \to \mathbb{R}$ heißt *nach oben beschränkt*, *nach unten beschränkt* bzw *beschränkt*, wenn die Bildmenge $f(A)$ die jeweilige Eigenschaft besitzt.

Eine Funktion ist also nach oben beschränkt, wenn es für alle $x \in A$ ein $K_1 \in \mathbb{R}$ gibt mit $K_1 \geq f(x)$ bzw. nach unten beschränkt, wenn es ein $K_2 \in \mathbb{R}$ gibt mit $K_2 \leq f(x)$. Wir werden für beschränkte Funktionen auch häufiger $K = \max\{|K_1|, |K_2|\}$ nutzen, da wir damit die beiden Beschränktheitsbedingungen kompakt als

$$|f(x)| < K, \ \forall x \in A$$

schreiben können.

Definition 4.13 (Monotonie)

Sei $A \subseteq \mathbb{R}$ und $f : A \to \mathbb{R}$ eine Funktion,

$$\text{dann heißt } f \begin{cases} \text{monoton wachsend,} & \text{falls } f(x) \leq f(x') \\ \text{streng monoton wachsend,} & \text{falls } f(x) < f(x') \\ \text{monoton fallend,} & \text{falls } f(x) \geq f(x') \\ \text{streng monoton fallend,} & \text{falls } f(x) > f(x') \end{cases}$$

für alle $x, x' \in A$ mit $x < x'$.

Beispiel 4.14

1. Die Identitätsfunktion ist (streng) monoton wachsend, da aus $x < x'$ offensichtlich $f(x) = x < f(x') = x'$ folgt.
2. Die konstante Funktion $f(x) = c$ ist monoton wachsend und fallend, da $c \leq c$ gilt.
3. Die Funktion $f(x) = x^2$ ist streng monoton wachsend auf dem Intervall $[0, \infty)$ und streng monoton fallend auf dem Intervall $(-\infty, 0]$

   Beweis (siehe auch Satz 2.36):
   $x, x' \in [0, +\infty)$: Im Fall von $0 = x < x'$ folgt $0 = x^2 < (x')^2$. Für die restlichen Fälle ist

   $$0 < x < x' \Rightarrow ((xx' < x'x') \wedge (xx < xx')) \Rightarrow x^2 < xx' < (x')^2.$$

   $x, x' \in (-\infty, 0]$: Im Fall von $x < x' = 0$ folgt $x^2 > 0 = (x')^2$. Für die restlichen Fälle ist

   $$x < x' \Rightarrow ((xx' > x'x') \wedge (xx > xx')) \Rightarrow x^2 > xx' > (x')^2.$$

Aus strenger Monotonie können wir unter anderem die Injektivität der Funktion folgern und damit die Existenz einer Umkehrfunktion (auf dem Bild von f). Dies ist eine einfache Möglichkeit zu zeigen, dass eine Umkehrfunktion existiert.

Satz 4.15 (Monotonie und Umkehrfunktionen)

Sei $A \subseteq \mathbb{R}$ und $f : A \to B$ eine Funktion mit $B := f(A) \subseteq \mathbb{R}$.

Ist f streng monoton, dann existiert die Umkehrfunktion $f^{-1} : B \to A$ und ist ebenfalls streng monoton (im gleichen Sinne).

▼ Beweis

Aus der strengen Monotonie folgt $x \neq x' \Rightarrow f(x) \neq f(x')$ und damit ist f injektiv. Da $B = f(A)$ ist f surjektiv. Damit ist f bijektiv und die Umkehrfunktion $g(x) = f^{-1}(x)$ existiert. Es bleibt also noch die Monotonie der Umkehrfunktion zu zeigen.

Wir betrachten den Fall, dass f streng monoton wachsend ist. Seien $y_1, y_2 \in f(A)$ mit $y_1 < y_2$ gegeben. Zu zeigen ist $g(y_1) < g(y_2)$.

Wäre $g(y_1) = g(y_2)$, dann wäre

$$y_1 = f(g(y_1)) = f(g(y_2)) = y_2,$$

da

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

Analog folgt aus $g(y_1) > g(y_2)$, dass

$$y_1 = f(g(y_1)) > f(g(y_2)) = y_2,$$

was im Widerspruch zu $y_1 < y_2$ steht (↯). Damit bleibt nur $g(y_1) < g(y_2)$.

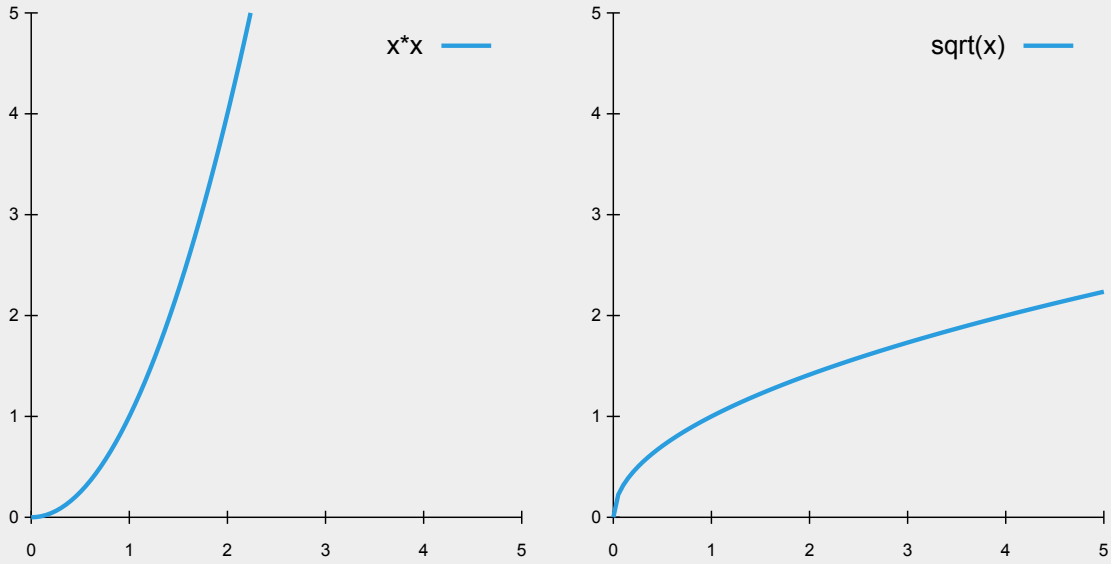
Den Fall der streng fallenden Monotonie beweist man analog.

□

Beispiel 4.16

Sei $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ mit $f(x) = x^k$, $k \in \mathbb{N}$ und $k \geq 2$.

f ist (nach Satz 2.36) streng monoton wachsend und bildet von $\mathbb{R}_{>0}$ bijektiv auf $\mathbb{R}_{>0}$ ab. Die Umkehrfunktion $f^{-1} : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$ mit $f^{-1}(x) = \sqrt[k]{x}$ existiert nach dem vorigen Satz und ist ebenfalls streng monoton wachsend.



4.2 Grenzwerte von Funktionen

Wir haben uns bereits in Kapitel 3 eingehend mit Grenzwerten von Folgen und Reihen (also einer speziellen Art von Funktionen) beschäftigt. Wir wollen nun diesen Grenzwertbegriff auf beliebige Funktionen verallgemeinern, damit klar ist, was wir meinen, wenn wir z.B.

$$\lim_{x \to 5} f(x)$$

schreiben. Diese Funktionsgrenzwerte werden uns für den Rest der Veranstaltung begleiten, da sie für die Differential- und Integralrechnung essenziell sind. Damit wir im obigen Beispiel den Grenzwert bei $x = 5$ bestimmen können, muss entweder $x = 5$ selbst zum Definitionsbereich gehören, oder es muss Elemente im Definitionsbereich geben, die $x = 5$ beliebig nahe kommen. Die Funktion könnte also zum Beispiel auch auf $[1, 5)$ definiert sein. Diese beiden Fälle fasst man unter dem Begriff des *Berührpunktes* zusammen.

Definition 4.17 (Berührpunkt)

Sei $A \subseteq \mathbb{R}$ und $a \in \mathbb{R}$. Dann ist a ein *Berührpunkt* von A, falls in jeder ε-Umgebung von a, d.h. im Intervall $(a - \varepsilon,\ a + \varepsilon)$ mit $\varepsilon > 0$, mindestens ein Element aus A liegt.

Dies ist äquivalent zu der Aussage, dass es eine Folge (x_n) gibt, deren Folgenglieder im Definitionsbereichs von f liegen und die gegen a konvergiert, also $\lim\limits_{n\to\infty} x_n = a$. Für $A = \mathbb{R}_{>0}$ sind also alle $x \geq 0$ (auch 0 selbst) Berührpunkte, $x = -1$ aber nicht. In den meisten Fällen ist dies sofort durch den Definitionsbereich klar und man muss sich meist keine großen Gedanken um Berührpunkte machen. Kommen wir also zur allgemeinen Grenzwertdefinition für Funktionen:

Definition 4.18 (Grenzwerte von Funktionen)

Sei $f : A \to \mathbb{R}$ mit $A \subseteq \mathbb{R}$ und $a \in \mathbb{R}$ ein Berührpunkt von A.

Man definiert dann $\lim\limits_{x\to a} f(x) = c$ für ein $c \in \mathbb{R}$, falls für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in A$, die gegen a konvergiert, gilt, dass die Folge der Funktionswerte $f(x_n)$ gegen c konvergiert, also

$$\lim_{n\to\infty} x_n = a \quad \Rightarrow \quad \lim_{n\to\infty} f(x_n) = c.$$

Analog definieren wir:

- $\lim\limits_{x\to\infty} f(x) = c$, wenn A nach oben unbeschränkt ist und für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $\lim\limits_{n\to\infty} x_n = \infty$ gilt $\lim\limits_{n\to\infty} f(x_n) = c$.
- $\lim\limits_{x\to-\infty} f(x) = c$, wenn A nach unten unbeschränkt ist und für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $\lim\limits_{n\to\infty} x_n = -\infty$ gilt $\lim\limits_{n\to\infty} f(x_n) = c$.

Achtung: Wir sprechen nur von einem Grenzwert, wenn $c \in \mathbb{R}$, also sind $c = \pm\infty$ keine gültigen Grenzwerte. Wir werden später häufig Sätze sehen in denen 'wenn der Grenzwert existiert' ein Teil der Aussage ist. Dafür muss der Grenzwert eine reelle Zahl sein. Wir werden später noch sogenannte *uneigentliche Grenzwerte* einführen, die den Fall $c = \pm\infty$ behandeln.

Wir haben also Grenzwerte von Funktionen über Folgengrenzwerte definiert. Unser Wissen über Folgengrenzwerte aus dem letzten Kapitel können wir also auch hier wieder nutzen. Oft ist es hilfreich den Definitionsbereich in zwei Teilbereiche zu trennen:

Definition 4.19 (linksseitiger und rechtsseitiger Grenzwert)

Sei $f\colon A \subseteq \mathbb{R} \to \mathbb{R}$ eine Funktion und $c \in \mathbb{R}$. Dann definieren wir

1. Den rechtsseitigen Grenzwert:

$$\lim_{x \searrow a} f(x) = c,$$

wenn a ein Berührpunkt von $A \cap (a, \infty)$ ist und für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in A, x_n > a$ und $\lim\limits_{n\to\infty} x_n = a$ gilt: $\lim\limits_{n\to\infty} f(x_n) = c$.

2. Den linksseitigen Grenzwert:

$$\lim_{x \nearrow a} f(x) = c,$$

wenn a ein Berührpunkt von $A \cap (-\infty, a)$ ist und für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $x_n \in A, x_n < a$ und $\lim\limits_{n\to\infty} x_n = a$ gilt: $\lim\limits_{n\to\infty} f(x_n) = c$.

Man spricht auch vom Grenzwert “von oben” (rechtsseitiger) und “von unten” (linksseitig). Manchmal sieht man auch die Schreibweise $\lim\limits_{x\to a^-} f(x)$ für den linksseitigen und $\lim\limits_{x\to a^+} f(x)$ für den rechtsseitigen Grenzwert.

Satz 4.20

Sei $f\colon A \subseteq \mathbb{R} \to \mathbb{R}$ eine Funktion und $a \in A$. Es gilt folgende Äquivalenz:

$$\lim_{x\to a} f(x) = f(a) \quad \Leftrightarrow \quad \lim_{x \nearrow a} f(x) = \lim_{x \searrow a} f(x) = f(a)$$

▼ Beweis als Übung

Beispiel 4.21

1. Für $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$ gilt $\lim\limits_{x \to a} f(x) = a^2$, denn für jede Folge (x_n) mit $\lim\limits_{n \to \infty} x_n = a$ gilt nach den Rechenregeln für konvergente Folgen (Satz 3.18)

$$\lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} x_n^2 = \left(\lim_{n \to \infty} x_n \right) \cdot \left(\lim_{n \to \infty} x_n \right) = a^2.$$

2. Die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = \begin{cases} 1 & \text{für } x \neq 1 \\ 2 & \text{für } x = 1 \end{cases}$ hat keinen Grenzwert bei $x = 1$, denn für die Folge $a_n = 1$ konvergiert $f(a_n)$ gegen $f(1) = 2$, aber für die Folge $b_n = 1 + 1/n > 1$ konvergiert $f(b_n)$ gegen 1. Somit gibt es zwei Folgen, die beide gegen $x = 1$ konvergieren und deren Funktionswerte sogar denselben Grenzwert haben, aber dieser Grenzwert ist ungleich $f(1)$:

$$\lim_{x \nearrow 1} (x) = \lim_{x \searrow 1} f(x) = 1 \neq f(1).$$

Die Rückrichtung dieses Satzes eignet sich besonders gut, um Grenzwerte zusammengesetzter Funktionen zu bestimmen, oder zu zeigen, dass diese nicht existieren. Mit dem Werkzeugt der Grenzwertbestimmung gerüstet, können wir uns nun dem Hauptthema dieses Kapitels widmen: der Stetigkeit.

4.3 Stetige Funktionen

Für die weiteren Untersuchungen von Funktionen ist es oft notwendig, dass die betrachtete Funktion ein “gutmütiges Verhalten” besitzt. Damit wollen wir die Sprünge einer Funktion, wie wir sie im letzten Beispiel gesehen haben, ausschließen. Anders gesagt: Wenn wir eine beliebige Folge betrachten, die gegen einen Grenzwert a konvergiert, dann ist unser gewünschtes “gutartiges Verhalten”, dass die Funktionswerte der Folge gegen $f(a)$ konvergieren. Dies fasst man unter dem Begriff der *Stetigkeit* zusammen.

Definition 4.22 (Stetigkeit)

Sei $f : A \to \mathbb{R}$ eine Funktion und $a \in A$. Die Funktion f heißt *stetig im Punkt* a, falls

$$\lim_{x \to a} f(x) = f(a).$$

f heißt *stetig* (in A), falls f in jedem Punkt aus A stetig ist.

Beispiel 4.23

1. Die konstante Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = c$ ist überall stetig, da für jede Folge (x_n) die (konstante) Folge der Funktionswerte $(f(x_n))$ gegen c konvergiert.
2. Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$ ist in jedem Punkt stetig nach Beispiel 4.21.
3. Die Heaviside-Funktion $f : \mathbb{R} \to \mathbb{R}$

$$f(x) = \begin{cases} 1 & \text{für } x \geq 0 \\ 0 & \text{für } x < 0 \end{cases}$$

ist in jedem $x \in \mathbb{R} \setminus \{0\}$ stetig (dort ist der Beweis analog zu 1.) und ist nicht stetig in 0, da $\lim\limits_{x \searrow 0} f(x) = 1 \neq \lim\limits_{x \nearrow 0} f(x) = 0$.

4. Die Dirichlet-Funktion

$$f(x) = \begin{cases} 1 & \text{falls } x \in \mathbb{Q}, \\ 0 & \text{falls } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

ist in keinem Punkt stetig. Zum Beweis betrachten wir für ein $a \in \mathbb{Q}$ die Folge $a_n = a + \frac{\sqrt{2}}{n}$, die offensichtlich gegen a konvergiert. Allerdings ist $\sqrt{2} \notin \mathbb{Q}$ und somit auch jedes Folgenglied von a_n: Also ist $f(a_n) = 0 \; \forall n \in \mathbb{N}$ und damit

$$\lim_{n \to \infty} f(a_n) = 0 \neq f(a) = 1.$$

Analog können wir uns für ein $b \in \mathbb{R} \setminus \mathbb{Q}$ eine Folge (b_n) konstruieren (zum Beispiel über eine Intervallschachtelung) die gegen b konvergiert, aber deren Folgenglieder alle rational sind, wodurch mit

$$\lim_{n \to \infty} f(b_n) = 1 \neq f(b) = 0$$

folgt, dass die Dirichlet-Funktion auch in keinem irrationalen Punkt stetig ist.

Definition 4.22 kann man auch so ausdrücken, dass wir für stetige Funktionen (bzw. an stetigen Stellen der Funktion) den Limes in die Funktion ziehen dürfen, also:

$$\lim_{x \to a} f(x) = f\left(\lim_{x \to a} x\right) = f(a).$$

Dies nutzen wir, um zu zeigen, dass die Exponentialfunktion $\exp(x)$ stetig ist.

Beispiel 4.24 (Stetigkeit der Exponentialfunktion)

Die Exponentialfunktion $\exp : \mathbb{R} \to \mathbb{R}, \exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ ist für alle $x \in \mathbb{R}$ stetig.

Für den Beweis zeigen wir zunächst, dass die Funktion in $x = 0$ stetig ist. Sei dazu (x_n) eine beliebige Nullfolge, dann gibt es ein n_0, sodass für alle $n \geq n_0$ gilt $|x_n - 0| = |x_n| < 1$. Wir zeigen damit nun, dass $\exp(x_n)$ gegen $\exp(0)$ konvergiert:

$$\begin{aligned} 0 \leq |\exp(x_n) - \exp(0)| = |\exp(x_n) - 1| &= \left| \left(\sum_{k=0}^{\infty} \frac{x_n^k}{k!} \right) - 1 \right| = \left| \sum_{k=1}^{\infty} \frac{x_n^k}{k!} \right| \\ &\leq \sum_{k=1}^{\infty} \frac{|x_n|^k}{k!} \leq \sum_{k=1}^{\infty} |x_n|^k \\ &= \left(\sum_{k=0}^{\infty} |x_n|^k \right) - |x_n|^0 = \frac{1}{1 - |x_n|} - 1 = \frac{|x_n|}{1 - |x_n|} \end{aligned}$$

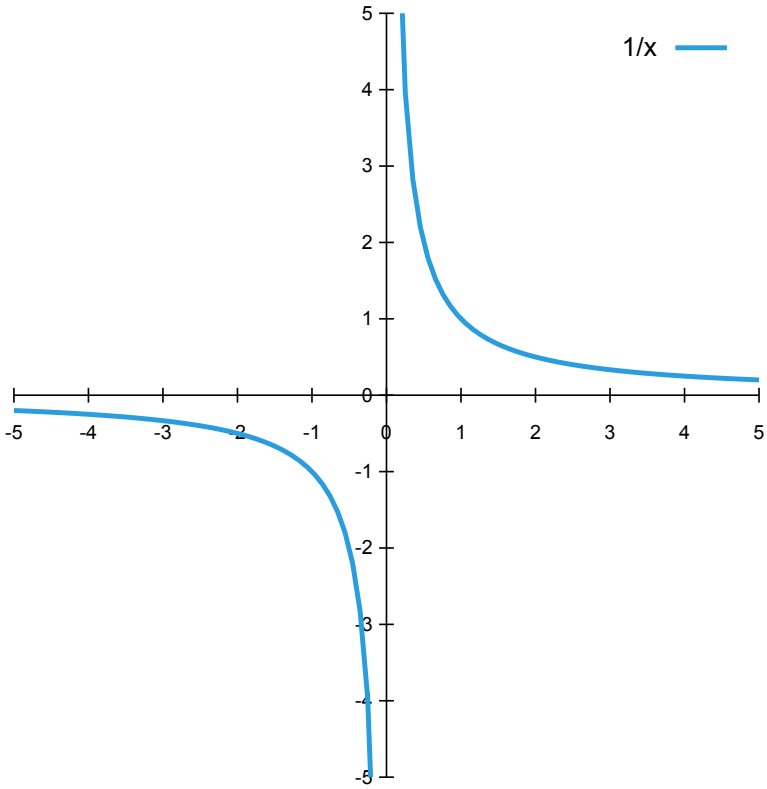
Da x_n eine Nullfolge ist konvergiert $\frac{|x_n|}{1-|x_n|}$ gegen 0. Es folgt nach dem Sandwich-Theorem, dass $|\exp(x_n) - 1|$ auch gegen 0 konvergieren muss. Somit gilt $\lim_{x \to 0} \exp(x) = \exp(0) = 1$ und $\exp(x)$ ist in $x = 0$ stetig.

Wir dürfen also den Limes in die Funktion ziehen für Grenzwerte gegen 0. Dies nutzen wir nun, um die Stetigkeit in einem beliebigen Punkt a zu zeigen. Sei dazu (y_n) eine beliebige Folge mit Grenzwert a, dann ist $x_n = y_n - a$ eine Nullfolge. Damit ergibt sich:

$$\begin{aligned} \lim_{n \to \infty} \exp(y_n) = \lim_{n \to \infty} \exp(x_n + a) &= \lim_{n \to \infty} \left(\exp(x_n) \cdot \exp(a) \right) \\ &= \exp\left(\lim_{n \to \infty} x_n \right) \cdot \exp(a) = \exp(0) \exp(a) = \exp(a) \end{aligned}$$

Damit ist die Stetigkeit von $\exp(x)$ für jedes $a \in \mathbb{R}$ bewiesen.

Bisher müssen wir für einen Stetigkeitsbeweis immer den Umweg über Folgen gehen. Da der Grenzwert für beliebige Folgen gelten muss, kann dies bei komplizierteren Funktionen schnell sehr aufwendig werden. Wir hätten also gerne eine alternative Methode für den Stetigkeitsbeweis. Oft wird Stetigkeit damit erklärt, dass wir den Graphen einer Funktion zeichnen können, ohne den Stift abzusetzen. Allerdings eignet sich diese Anschauung nicht für einen Beweis, da wir zum Beispiel immer nur einen begrenzten Abschnitt einer Funktion zeichnen können und auch nur mit begrenzt hoher Auflösung. Außerdem gibt es Funktionen, bei denen wir zwar den Stift absetzen müssen, die aber trotzdem stetig sind, wie beispielsweise $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$ mit $f(x) = x^{-1}$.



Der Graph erfordert definitiv das Absetzen des Stiftes, allerdings liegt die “Absetzstelle” nicht im Definitionsbereich, da $x = 0$ ausgeschlossen werden muss. Für alle anderen Punkte $x \neq 0$ lässt sich die Stetigkeit leicht zeigen (zur Übung). Somit wäre diese Funktion stetig, obwohl der Graph nicht danach aussieht. Wir benötigen also ein handlicheres Kriterium für die Stetigkeit, das sich besser für einen Beweis eignet. Das nun folgende ε-δ-Kriterium sieht auf den ersten Blick zwar komplizierter aus, ist aber sehr hilfreich für diesen Zweck. Es beschreibt die “Gutmütigkeit” der Funktion in etwa wie folgt: wenn sich x und a immer näher kommen, dann sollen sich $f(x)$ und $f(a)$ auch immer näher kommen.

Satz 4.25 (Epsilon-Delta-Kriterium der Stetigkeit)

Sei $A \subseteq \mathbb{R}$ und $f : A \to \mathbb{R}$ eine Funktion. f ist genau dann im Punkt $a \in A$ stetig, wenn gilt:

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in A : \; |x - a| < \delta \;\Rightarrow\; |f(x) - f(a)| < \varepsilon$$

In Worten: Zu jedem positiven ε finden wir ein positives δ, sodass für alle x-Werte, deren Abstand zur untersuchten Stelle a kleiner als δ ist, folgt, dass der Abstand ihres Funktionswerts zu $f(a)$ kleiner als ε ist.

▼ Beweis

Der Beweis wird in zwei Schritten geführt, indem wir in jedem Schritt eine Richtung der Äquivalenzbeziehung beweisen:

1. ε, δ existieren $\Rightarrow$ Stetigkeit:

   Es gebe zu jedem $\varepsilon > 0$ ein $\delta > 0$, sodass $|f(x) - f(a)| < \varepsilon$ für alle $x \in A$ mit $|x - a| < \delta$. Es ist zu zeigen, dass für jede Folge $(x_n)_{n \in \mathbb{N}}$ mit $x_n \in A$ und $\lim\limits_{n \to \infty} x_n = a$ gilt, dass $\lim\limits_{n \to \infty} f(x_n) = f(a)$.

   Sei $\varepsilon > 0$ vorgegeben und $\delta > 0$ gemäß der Voraussetzung gewählt. Da $\lim\limits_{n \to \infty} x_n = a$ existiert ein $n_0 \in \mathbb{N}$, sodass $|x_n - a| < \delta$ für alle $n \geq n_0$. Nach Voraussetzung ist damit $|f(x_n) - f(a)| < \varepsilon$ für alle $n \geq n_0$. Also gilt $\lim\limits_{n \to \infty} f(x_n) = f(a)$ und f ist stetig.

2. Stetigkeit $\Rightarrow \varepsilon, \delta$ existieren:

   Für jede Folge $x_n \in A$ mit $\lim\limits_{n \to \infty} x_n = a$ gelte $\lim\limits_{n \to \infty} f(x_n) = f(a)$. Es ist zu zeigen, dass zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, sodass $|f(x) - f(a)| < \varepsilon$ für alle $x \in A$ mit $|x - a| < \delta$. Wir führen einen Widerspruchsbeweis:

   Angenommen es existiert ein $\varepsilon > 0$, sodass für jedes $\delta > 0$ mindestens ein $x \in A$ existiert mit

   $$|x - a| < \delta \quad \text{und} \quad |f(x) - f(a)| \geq \varepsilon. \tag{*}$$

   Wenn die obige Aussage für jedes $\delta > 0$ gilt, dann gilt sie auch für $\delta = 1/n$, $n \in \mathbb{N}$. Es gibt also für beliebige $n \in \mathbb{N}$ ein $x_n \in A$ mit $|x_n - a| < \frac{1}{n}$. Diese x_n bilden eine Folge (x_n), die gegen a konvergiert, denn es gibt für beliebige $\varepsilon' > 0$ ein $n_0' = \lceil 1/\varepsilon \rceil$, sodass für alle $n \geq n_0'$ gilt $|x_n - a| < \frac{1}{n} \leq \varepsilon'$.

   Wegen der Stetigkeit von f muss für diese Folge (x_n) gelten

   $$\lim_{n \to \infty} f(x_n) = f(a),$$

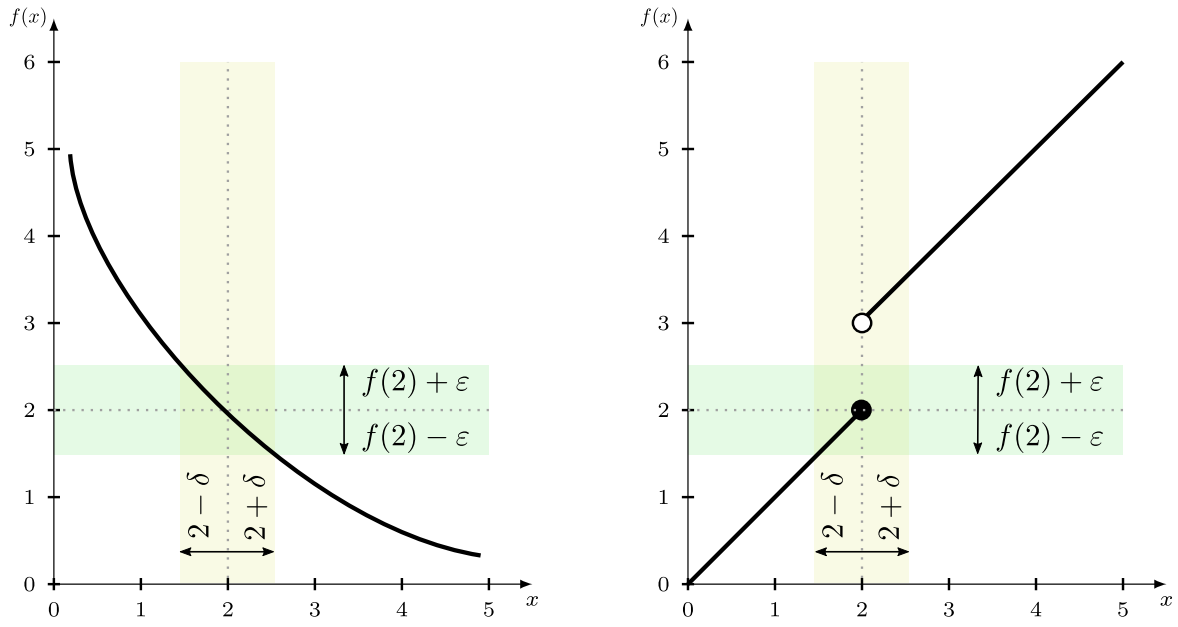
oder anders gesagt: für jedes $\varepsilon'' > 0$ muss es ein n_0'' geben, sodass für alle $n \geq n_0$ gilt:

$$|f(x_n) - f(a)| < \varepsilon''.$$

Dies steht im Widerspruch zu (*) (↯).

□

Betrachten wir zunächst eine graphische Interpretation des ε-δ-Kriteriums:



Links ist der Fall einer stetigen Funktion gezeigt. Für das grüne Funktionswertintervall $(f(a) - \varepsilon, f(a) + \varepsilon)$ finden wir stets ein gelbes Definitionsbereichsintervall $(a - \delta, a + \delta)$, sodass alle Funktionswerte des gelben Bereichs auch im grünen liegen. Wir könnten links das ε, also den grünen Bereich, beliebig klein machen und würden trotzdem immer einen δ (gelben) Bereich finden, für den dies gilt. Bei der unstetigen Funktion rechts gibt es eine 1 Längeneinheiten große Lücke bei $x = 2$. Wenn wir also hier z.B. $\varepsilon = 0.5$ wählen, dann können wir den gelben Bereich um die unstetige Stelle noch so groß oder klein machen, es würde immer Funktionswerte im gelben Bereich geben, die nicht im grünen Bereich liegen.

Beispiel 4.26 (Epsilon-Delta-Stetigkeitsbeweise)

- Wir starten mit einem einfachen Beispiel, der Identitätsfunktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x$. Da hier $|f(x) - f(a)| = |x - a|$ gilt, können wir $\delta = \varepsilon$ wählen, denn dann gilt:

$$|x - a| < \delta \quad \Rightarrow \quad |f(x) - f(a)| = |x - a| < \delta = \epsilon,$$

  womit das ε-δ-Kriteriums erfüllt ist. Somit ist $f(x) = x$ stetig.

- $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^2$ ist stetig. Auch wenn wir das bereits gezeigt haben, führen wir den Beweis hier erneut mit dem ε-δ-Kriterium, denn hier ist der Beweis nicht trivial, aber trotzdem noch so einfach, dass man die Herangehensweise verstehen kann.

▼ Beweis

Man startet in der Regel mit einer Nebenrechnung als Vorüberlegung, bevor man anschließend den eigentlichen Beweis formuliert.

Wir beginnen mit der Annahme, dass $|x - a| < \delta$ gilt. Dann betrachten wir $|f(x) - f(a)|$ und versuchen für diesen Term einen größeren Ausdruck zu finden, der nicht von x abhängt. Dabei dürfen wir $|x - a| < \delta$ benutzen und alle Ungleichungen, die wir bereits kennen oder vorher beweisen. Besonders die Dreiecksungleichungen (Satz 2.17) sind hier immer wieder hilfreich. Fangen wir an:

$$\begin{aligned}
|f(x) - f(a)| &= \left|x^2 - a^2\right| \\
&= |(x - a)(x + a)| && \text{(3. binomische Formel)} \\
&= |x + a|\,|x - a| \\
&< |x + a|\,\delta \\
&= |x - a + a + a|\,\delta \\
&\leq (|x - a| + |2a|)\,\delta && \text{(Dreiecksungleichung)} \\
&< (\delta + |2a|)\,\delta.
\end{aligned}$$

Wenn wir also $\varepsilon = (\delta + |2a|)\,\delta$ wählen, dann würde das ε-δ-Kriterium erfüllt sein. Da wir aber eine Vorschrift für δ angeben müssen und nicht für ε, müssen wir diese Bedingung noch nach δ umstellen: $\delta = \sqrt{\varepsilon + |a|^2} - |a|$. Hier müssen wir uns noch kurz versichern, dass $\delta > 0$ gilt (daher entfällt die 2. Lösung $-\sqrt{\varepsilon + |a|^2} - |a|$).

Die bisherigen Schritte waren wie bereits gesagt nur eine Vorberechnung für den eigentlichen Beweis, der nun aber einfach (in umgekehrter Reihenfolge) zu führen ist: Sei $\varepsilon > 0$ und $a \in \mathbb{R}$ beliebig, wähle $\delta = \sqrt{\varepsilon + |a|^2} - |a|$, dann folgt für alle $x \in \mathbb{R}$ mit $|x - a| < \delta$:

$$
\begin{aligned}
|f(x) - f(a)| &= \left|x^2 - a^2\right| \\
&= |(x-a)(x+a)| && \text{(3. binomische Formel)} \\
&= |x+a|\,|x-a| \\
&< |x+a|\,\delta \\
&= |x-a+a+a|\,\delta \\
&\leq (|x-a| + |2a|)\,\delta && \text{(Dreiecksungleichung)} \\
&< (\delta + |2a|)\,\delta = \varepsilon.
\end{aligned}
$$

Somit ist das ε-δ-Kriterium erfüllt und f ist stetig.

□

Wir können jetzt mit dem Folgenkriterium (Definition 4.22 bzw. Satz 4.20) oder dem ε-δ-Kriterium (Satz 4.25) die Stetigkeit von Funktionen nachweisen. Die folgenden zwei Sätze vereinfachen den Nachweis der Stetigkeit komplizierter Funktionen, indem sie diese auf Kombinationen einfacherer Funktionen zurückführen.

Satz 4.27 (Operationen auf stetigen Funktionen)

Seien $f, g : A \to \mathbb{R}$ Funktionen, die in $a \in A$ stetig sind, und sei $c \in \mathbb{R}$. Dann sind auch die Funktionen

a. $f + g : A \to \mathbb{R}$
b. $c \cdot f : A \to \mathbb{R}$
c. $f \cdot g : A \to \mathbb{R}$

im Punkt a stetig. Ist $g(a) \neq 0$, so ist auch die Funktion

d. $\frac{f}{g} : A' \to \mathbb{R}$

in a stetig. Dabei ist $A' = \{x \in A \mid g(x) \neq 0\}$.

▼ Beweis

Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in A (bzw. A') und $\lim\limits_{n\to\infty} x_n = a$. Es ist zu zeigen:

$$\lim_{n\to\infty} (f+g)(x_n) = (f+g)(a)$$
$$\lim_{n\to\infty} (c\cdot f)(x_n) = (c\cdot f)(a)$$
$$\lim_{n\to\infty} (f\cdot g)(x_n) = (f\cdot g)(a)$$
$$\lim_{n\to\infty} \left(\frac{f}{g}\right)(x_n) = \left(\frac{f}{g}\right)(a)$$

Nach Voraussetzung ist $\lim\limits_{n\to\infty} f(x_n) = f(a)$ und $\lim\limits_{n\to\infty} g(x_n) = g(a)$. Die Behauptung folgt aus den Rechenregeln für Folgen (siehe Satz 3.18).

□

Beispiel 4.28

- Die Identitätsfunktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x$, ist nach Beispiel 4.26 stetig. Damit ist nach Satz 4.27 auch jede Funktion $g : \mathbb{R} \to \mathbb{R}$ mit $g(x) = cx^n$ für alle $c \in \mathbb{R}$ und $n \in \mathbb{N}$ stetig (Beweis durch (b) und mehrmalige Anwendung von (c)).
- Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \frac{\exp(x)}{x^2+1}$, ist stetig, da
  - die Exponentialfunktion $\exp(x)$ stetig ist (Beispiel 4.24),
  - konstante Funktionen $f(x) = c$ und $f(x) = x^2$ stetig sind (Beispiel 4.21),
  - nach Satz 4.27(a) somit auch $1 + x^2$ stetig ist und
  - wegen $1 + x^2 \neq 0$ nach Satz 4.27(d) $\frac{\exp(x)}{x^2+1}$ in jedem Punkt $a \in \mathbb{R}$ stetig ist.

Satz 4.29 (Komposition stetiger Funktionen)

Seien $f : A \to \mathbb{R}$ und $g : B \to \mathbb{R}$ Funktionen mit $f(A) \subseteq B$. Die Funktion f sei in $a \in A$ und g in $b = f(a) \in B$ stetig.

Dann ist die Funktion $g \circ f : A \to \mathbb{R}$ in a stetig.

▼ Beweis

Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in A mit $\lim\limits_{n\to\infty} x_n = a$.

Da f stetig in a ist, gilt $\lim\limits_{n\to\infty} f(x_n) = f(a)$. Nach Voraussetzung ist $y_n = f(x_n) \in B$ und $\lim\limits_{n\to\infty} y_n = b$. Da g in b stetig ist, gilt auch $\lim\limits_{n\to\infty} g(y_n) = g(b)$. Deshalb folgt

$$\lim_{n\to\infty} (g \circ f)(x_n) = \lim_{n\to\infty} g(f(x_n)) = \lim_{n\to\infty} g(y_n) = g(b) = g(f(a)) = (g \circ f)(a).$$

□

Beispiel 4.30

Seien $f : \mathbb{R} \to \mathbb{R},\ f(x) = \exp(x)$ und $g : \mathbb{R} \to \mathbb{R},\ g(x) = x^2$. f und g sind stetig in $\mathbb{R}$, daher ist auch $h(x) = \exp(x^2)$ stetig in $\mathbb{R}$.

Nachdem wir nun einige Techniken kennen, um die Stetigkeit einer Funktion nachzuweisen, wollen wir nun nützliche Folgerungen aus der Stetigkeit ableiten. Als Anschauung können Sie immer das “Zeichnen des Graphen ohne den Stift abzusetzen” im Hinterkopf behalten.

Satz 4.31 (Zwischenwertsatz)

Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) < 0 < f(b)$ oder $f(a) > 0 > f(b)$. Dann existiert ein $c \in (a, b)$ mit $f(c) = 0$.

Allgemeiner gilt für alle $y \in \mathbb{R}$: Wenn $f(a) < y < f(b)$ oder $f(a) > y > f(b)$, dann existiert ein $d \in (a, b)$ mit $f(d) = y$.

▼ Beweis

Erster Teil:
Der Beweis erfolgt über eine Intervallschachtelung (Satz 3.39). Wir betrachten den Fall $f(a) < 0 < f(b)$. Der andere Fall $f(a) > 0 > f(b)$ ist analog zu beweisen.

Wir definieren zunächst eine Intervallschachtelung mit den Eigenschaften

1. $[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$,
2. $|[a_{n+1}, b_{n+1}]| = (b_{n+1} - a_{n+1}) = \frac{1}{2}(b_n - a_n) = \frac{1}{2}|[a_n, b_n]|$,
3. $f(a_n) < 0 < f(b_n)$.

Dass so eine Schachtelung existiert, beweisen wir durch vollständige Induktion:

- **Induktionsanfang:** $[a_1, b_1] = [a, b]$.
- **Induktionsvoraussetzung:** Für $n \in \mathbb{N}$ existiert ein Intervall $I_n = [a_n, b_n]$ mit den oben genannten Eigenschaften.
- **Induktionsschritt:** Wir bestimmen den Mittelpunkt $m = \frac{a_n + b_n}{2}$ des Intervalls $[a_n, b_n]$, welches nach Induktionsvoraussetzung existiert, und konstruieren daraus das nächste Intervall:

$$[a_{n+1}, b_{n+1}] = \begin{cases} [a_n, m] & \text{falls } f(m) \geq 0 \\ [m, b_n] & \text{falls } f(m) < 0 \end{cases}$$

Es ist leicht nachzuprüfen, dass die obigen Bedingungen 1.–3. für die Intervalle gelten. Nach Satz 3.39 existiert ein eindeutiges $c \in \mathbb{R}$ mit $c \in [a_n, b_n]\ \forall n \in \mathbb{N}$.

Da f stetig ist, gilt außerdem:

$$\lim_{x \to c} f(x) = \lim_{n \to \infty} f(a_n) = \lim_{n \to \infty} f(b_n) = f(c)$$

und gleichzeitig auch (per Konstruktion)

$$f(a_n) < 0 < f(b_n).$$

Dies ist nur für den Grenzwert 0 möglich. Es gilt also $f(c) = 0$.

Zweiter Teil:
Man könnte die Intervallschachtelung aus dem ersten Teil des Beweises analog mit y statt mit 0 als Vergleichspunkt konstruieren. Wir führen hier aber einen etwas

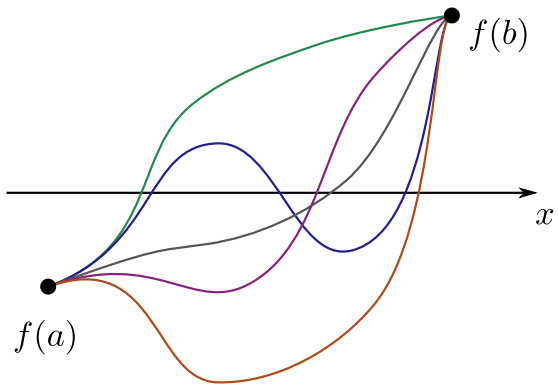
anderen Beweis, der auf den ersten Teil des Satzes aufbaut. Wir betrachten den Fall $f(a) < y < f(b)$, der andere Fall lässt sich analog beweisen.

Wir definieren eine Funktion $g : [a, b] \to \mathbb{R}$ mit $g(x) = f(x) - y$. Wegen der Stetigkeit von f und von konstanten Funktionen ist nach Satz 4.27 auch g stetig. Es gilt $g(a) < 0 < g(b)$, sodass nach dem ersten Teil des Satzes ein $d \in [a, b]$ mit $g(d) = 0$ existiert, für das nach Konstruktion dann auch $f(d) = y$ gilt.

□

Vorsicht: Der Satz gilt nicht für $f : [a, b] \subset \mathbb{Q} \to \mathbb{R}$, da in $\mathbb{Q}$ die Intervallschachtelung, welche wir zum Beweis benutzt haben, nicht konvergieren muss. Zum Beispiel ist $f(x) = x^2 - 2$ stetig, aber die Nullstelle $\sqrt{2}$ liegt nicht in $\mathbb{Q}$.

Graphisch kann man sich den Zwischenwertsatz mit dem im folgenden Bild veranschaulichen: Es ist nicht möglich, eine durchgängige Linie von $f(a)$ zu $f(b)$ zu zeichnen (die stetige Funktion), welche die x-Achse nicht mindestens einmal schneidet.



Beispiel 4.32 (Bisektion und Regula-Falsi)

Die Erkenntnis des Zwischenwertsatzes wird in der Praxis genutzt um Nullstellen zu bestimmen. Kennt man nämlich zwei Funktionswerte $f(a)$ und $f(b)$ mit unterschiedlichem Vorzeichen, dann wissen wir (für stetige Funktionen) dass es zwischen a und b eine Nullstelle x_0 geben muss. Eine einfache Möglichkeit, sich x_0 zu nähern, ist das sogenannte Bisektionsverfahren. Hierbei wird genau so verfahren, wie bereits im Beweis des Zwischenwertsatzes: Wir bestimmen zunächst den Mittelpunkt $c = \frac{1}{2}(a+b)$. Es ist entweder $f(c) = 0$, dann ist $x_0 = c$, oder $f(c) \neq 0$, dann ersetzt c eine der beiden Grenzen a oder b (ja nach Vorzeichen von $f(c)$). Die Schritte können nun für dieses kleinere Intervall wiederholt werden. Somit grenzt man die Nullstelle in jedem Iterationsschritt weiter ein.

Die Bisektionsmethode kann verbessert werden durch das sogenannte Regula-Falsi-Verfahren. Hierbei teilt man das Intervall $[a, b]$ nicht am Mittelpunkt sondern bestimmt die Nullstelle der Verbindungslinie durch die Punkte $(a, f(a))$ und $(b, f(b))$ (Sekante). Diese ergibt sich über:

$$c = b - f(b)\frac{b-a}{f(b)-f(a)}.$$

Die restlichen Schritte erfolgen analog zur Bisektion. Da es bei beiden Verfahren unwahrscheinlich ist, dass genau der Punkt c mit $f(c) = 0$ gefunden wird, muss das Verfahren irgendwann abgebrochen werden. Dazu muss eine Abbruchbedingung definiert werden. Übliche Bedingungen sind $|f(c)| < \varepsilon$ oder $b - a < \varepsilon$ für ein kleines $\varepsilon > 0$.

Die Verfahren konvergieren für stetige Funktionen immer gegen eine Nullstelle, die Konvergenz kann aber relativ langsam sein, d.h. es werden viele Schritte benötigt. Das Regula-Falsi-Verfahren können Sie in der nachfolgenden Demo testen.

▶ Demo: Regula-Falsi

Satz 4.33 (Intervallabbildung stetiger Funktionen)

Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine stetige Funktion. Dann ist auch $J = f(I) \subseteq \mathbb{R}$ ein Intervall.

▼ Beweis

Wir setzten $q = \sup(f(I)) \in \mathbb{R} \cup \{\infty\}$ und $p = \inf(f(I)) \in \mathbb{R} \cup \{-\infty\}$. Zunächst zeigen wir $(p, q) \subseteq f(I)$.

Sei dazu $y \in \mathbb{R}$ beliebig gewählt, sodass $p < y < q$. Nach Definition von p und q gibt es dann $a, b \in I$ mit $f(a) < y < f(b)$. Nach dem Zwischenwertsatz existiert $x \in I$ mit $f(x) = y$. Also ist $y \in f(I)$. Damit ist $(p, q) \subseteq f(I)$ bewiesen und $f(I)$ muss eines der folgenden vier Intervalle sein: $[p, q], (p, q], [p, q), (p, q)$.

□

Satz 4.34 (Umkehrfunktionen stetiger Funktionen)

Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ stetig und streng monoton (wachsend oder fallend). Sei $J = f(I)$, dann bildet f das Intervall I bijektiv auf J ab und die Umkehrfunktion $f^{-1} : J \to \mathbb{R}$ ist stetig.

▼ Beweis

Die Monotonie und Existenz von f^{-1} wurde bereits in Satz 4.15 gezeigt. f ist als streng monotone Funktion injektiv. Da für den Definitionsbereich von f^{-1} gilt $J = f(I)$, ist die Abbildung auch surjektiv und damit bijektiv. J ist nach Satz 4.33 ebenfalls ein Intervall. Damit bleibt nur noch die Stetigkeit von f^{-1} zu zeigen.

Zum Beweis der Stetigkeit benutzen wir das ε-δ-Kriterium (Satz 4.25) und nehmen an, dass f streng monoton wachsend ist. Für ein $b \in J$ existiert ein $a = f^{-1}(b) \in I$, d.h. $b = f(a)$. Ist $b \in J$ kein Randpunkt von J, dann existiert $\varepsilon' > 0$, sodass $[b - \varepsilon', b + \varepsilon'] \subseteq J$.

Wir zeigen nun die Stetigkeit von f^{-1} in b. Da f bijektiv und stetig ist, ist auch a kein Randpunkt von I. Damit können wir ein $\varepsilon > 0$ wählen, sodass $[a - \varepsilon, a + \varepsilon] \subseteq I$ und für die Elemente x des Intervalls gilt $|x - a| = |f^{-1}(f(x)) - f^{-1}(f(a))| < \varepsilon$.

Sei $b_1 = f(a - \varepsilon)$ und $b_2 = f(a + \varepsilon)$. Wegen der strengen Monotonie von f gilt $b_1 < b < b_2$. Ferner bildet f das Intervall $[a - \varepsilon, a + \varepsilon]$ bijektiv auf das Intervall $[b_1, b_2]$ ab.
Wir wählen $\delta = \min(b - b_1, b_2 - b)$. Dann gilt $f^{-1}((b - \delta, b + \delta)) \subseteq (a - \varepsilon, a + \varepsilon)$.
Für $y \in (b - \delta, b + \delta)$ gilt offensichtlich $|y - b| < \delta$ und gleichzeitig für die Funktionswerte $|f^{-1}(y) - f^{-1}(b)| < \varepsilon$. Damit ist das ε-δ-Kriterium erfüllt und f^{-1} ist stetig nach Satz 4.25.

Damit haben wir die Stetigkeit von f^{-1} in allen Punkten gezeigt, die keine Randpunkte von J sind. Für Randpunkte $b \in J$ erfolgt der Beweis durch Betrachtung von $[a, a - \varepsilon]$ bzw. $[a, a + \varepsilon]$ und ein ansonsten analoges Vorgehen. Für monoton fallende Funktionen erfolgt der Beweis ebenfalls analog.

□

Beispiel 4.35

$f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$, $f(x) = x^2$ ist stetig (Beispiel 4.21) und streng monoton wachsend (Beispiel 4.14). Damit ist die Umkehrfunktion $f^{-1}(x) = \sqrt{x}$ ebenfalls stetig auf $f(\mathbb{R}_{>0}) = \mathbb{R}_{>0}$.

Der folgende Satz betrachtet stetige Funktionen auf kompakten Intervallen (zur Erinnerung: kompakt bedeutet abgeschlossen und beschränkt, siehe Definition 3.35).

Satz 4.36 (Minimum und Maximum auf kompakten Intervallen)

Auf einem kompakten Intervall $[a, b]$ ist jede stetige Funktion $f : [a, b] \to \mathbb{R}$ beschränkt und nimmt ihr Minimum und Maximum an.

D.h. es existieren ein $c \in [a, b]$, sodass $f(c) = \sup\{f(x) \mid x \in [a, b]\}$, und ein $d \in [a, b]$, sodass $f(d) = \inf\{f(x) \mid x \in [a, b]\}$.

▼ Beweis

Wir führen den Beweis für das Maximum: Sei $y = \sup\{f(x) \mid x \in [a, b]\} \in \mathbb{R} \cup \{\infty\}$. Dann existiert eine Folge (x_n) mit $x_n \in [a, b]$ für alle $n \in \mathbb{N}$, sodass $\lim_{n\to\infty} f(x_n) = y$, da die Funktion f stetig ist. Da (x_n) beschränkt ist (das Intervall aus dem die Folge gewählt wird ist beschränkt), besitzt sie nach den Satz von Bolzano-Weierstraß (Satz 3.30) eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ mit $\lim_{k\to\infty} x_{n_k} = c \in [a, b]$.

Aus der Stetigkeit folgt nun $f(c) = \lim_{k\to\infty} f(x_{n_k}) = y$ und damit insbesondere auch $y = f(c) \in \mathbb{R}$. Damit ist f nach oben beschränkt und nimmt in c das Maximum an.

Der Beweis für das Minimum ist analog zu führen.

□

Der vorherige Satz gilt nicht für offene oder halboffene Intervalle. So ist zum Beispiel die Funktion $f(x) = \frac{1}{x}$ im Intervall $(0, 1)$ stetig, aber unbeschränkt. Die Funktion $f(x) = x$ ist im Intervall $(0, 1)$ stetig und beschränkt, nimmt aber weder das Supremum 1 noch das Infimum 0 an.

Die Stetigkeit und ihre Folgerungen wirken an dieser Stelle recht trivial, da wir uns die Funktionen noch gut als Graphen vorstellen können. Wir können die Regeln aber später auch teilweise auf mehrdimensionale oder komplexwertige Funktionen erweitern, die in sehr vielen praktischen Anwendungen eine Rolle spielen. Für viele dieser Funktionen versagt die Vorstellungskraft. Mit den hier eingeführten Regeln und Folgerungen können wir für diese Funktionen aber zum Beispiel trotzdem zeigen, dass auf einem kompakten Definitionsgebiet irgendwo ein Minimum und Maximum existieren muss. Diese Extrema sind in der Praxis von großem Interesse, da sie eine optimale Lösung des von der Funktion beschriebenen Problems darstellen.

Zuletzt betrachten wir noch die sogenannte *gleichmäßige Stetigkeit*, bei der eine stärkere Bedingung an die Funktion gestellt wird, als bei der herkömmlichen Stetigkeit. Ähnlich wie beim Satz von Bolzano-Weierstraß wird an dieser Stelle noch nicht gleich klar, wofür wir diese zweite Art der Stetigkeit

benötigen. Aber wir werden im Kapitel über Integrale einen sehr wichtigen Satz beweisen, der die gleichmäßige Stetigkeit einer Funktion erfordert.

Definition 4.37 (Gleichmäßige Stetigkeit)

Eine Funktion $f : A \to \mathbb{R}$ heißt *gleichmäßig stetig*, wenn gilt:

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x, y \in A : \; |x - y| < \delta \Rightarrow \; |f(x) - f(y)| < \varepsilon.$$

Machen Sie sich den Unterschied zu Satz 4.25 bewusst. Dort haben wir gefordert, dass für alle a im Definitionsbereich gilt

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall x \in A : \; |x - a| < \delta \Rightarrow \; |f(x) - f(a)| < \varepsilon.$$

Hierbei konnte das δ sowohl von a, als auch von ε abhängen. Bei der gleichmäßigen Stetigkeit übernimmt das y die Rolle von a, steht aber nun nach dem δ in der Bedingung. Daher darf für die gleichmäßige Stetigkeit das δ nur noch von ε abhängen. Man kann sich die gleichmäßige Stetigkeit also damit vorstellen, dass eine Grenze für die Schwankung von Funktionswerten für jede Intervalllänge von x-Werten existiert. Die Funktion kann also auf einer vorgegebenen Intervalllänge nicht beliebig stark ansteigen.

Beispiel 4.38 (Gleichmäßige Stetigkeit)

Die Funktion $f : \mathbb{R}_{>0} \to \mathbb{R}_{>0}, f(x) = \sqrt{x}$ ist gleichmäßig stetig. Vor dem Beweis machen wir uns dies erst anschaulich klar: die Wurzelfunktion steigt für kleine x stärksten an und flacht für größere Argumente immer weiter ab, damit können wir bereits vermuten, dass die Funktion gleichmäßig stetig ist. Für den formalen Beweis geht man ähnlich vor wie in Beispiel 4.26:

Ein kleiner Hilfsbeweis vorweg:
Es gilt $(a+b)^2 = a^2 + 2ab + b^2 \geq a^2 + b^2$, also nach Satz 2.36 auch $a + b \geq \sqrt{a^2 + b^2}$. Für $a = \sqrt{x}$ und $b = \sqrt{y}$ und $x, y > 0$ folgt weiter

$$\sqrt{x} + \sqrt{y} \geq \sqrt{x+y} = \sqrt{x - (-y)} \geq \sqrt{||x| - |-y||} = \sqrt{|x-y|},$$

wobei der vorletzte Schritt Satz 2.17(e) ausnutzt. Es gilt also insgesamt

$$\sqrt{x} + \sqrt{y} \geq \sqrt{|x-y|} \; (*).$$

Nun zum Beweis der gleichmäßigen Stetigkeit:
Sei $\varepsilon > 0$ und $\delta = \varepsilon^2 > 0$. Dann folgt für alle $x, y > 0$ aus $|x - y| < \delta$:

$$\begin{aligned} |f(x) - f(y)| = |\sqrt{x} - \sqrt{y}| &= \left| (\sqrt{x} - \sqrt{y}) \frac{\sqrt{x} + \sqrt{y}}{\sqrt{x} + \sqrt{y}} \right| \\ &= \frac{|x-y|}{\sqrt{x} + \sqrt{y}} \\ &\overset{(*)}{\leq} \frac{|x-y|}{\sqrt{|x-y|}} = \sqrt{|x-y|} < \sqrt{\delta} = \varepsilon. \end{aligned}$$

Damit ist die gleichmäßige Stetigkeit von $f(x) = \sqrt{x}$ bewiesen.

Den nun folgenden Satz werden wir, wie bereits erwähnt, später noch benötigen.

Satz 4.39 (Gleichmäßige Stetigkeit auf kompakten Intervallen)

Ist eine Funktion $f : A \to \mathbb{R}$ auf einem kompakten Intervall $[a, b] \subseteq A$ stetig, dann ist sie dort auch gleichmäßig stetig.

▼ Beweis

Wir führen einen Beweis durch Widerspruch:
Angenommen f sei nicht gleichmäßig stetig. Die Negation der Bedingung der gleichmäßigen Stetigkeit sieht wie folgt aus:

$$\exists \varepsilon > 0 \; \forall \delta > 0 \; \exists x, y \in A : \; |x - y| < \delta \wedge \; |f(x) - f(y)| \geq \varepsilon.$$

Da dies für alle $\delta > 0$ gilt, können wir $\delta = \frac{1}{n}$ für alle $n \in \mathbb{N}$ wählen und die entsprechenden x, y mit obiger Eigenschaft x_n und y_n nennen. Dann gibt es ein $\varepsilon > 0$, sodass für alle $n \in \mathbb{N}$ zwei Punkte $x_n, y_n \in [a, b]$ existieren mit $|x_n - y_n| < \frac{1}{n}$ und $|f(x_n) - f(y_n)| \geq \varepsilon$.

Die Folge (x_n) ist beschränkt (da $x_n \in [a, b]$) und besitzt nach dem Satz von Bolzano-Weierstraß (Satz 3.30) eine konvergente Teilfolge (x_{n_k}) mit $\lim\limits_{k\to\infty} x_{n_k} = c \in [a, b]$. Aus $|x_n - y_n| < \frac{1}{n}$ folgt $\lim\limits_{k\to\infty} |x_{n_k} - y_{n_k}| = 0$ und damit $\lim\limits_{k\to\infty} y_{n_k} = c$.

Da f stetig ist, gilt auch

$$\lim_{k\to\infty} (f(x_{n_k}) - f(y_{n_k})) = f(c) - f(c) = 0,$$

was im Widerspruch zur Annahme $|f(x_n) - f(y_n)| \geq \varepsilon$ für alle $n \in \mathbb{N}$ steht (↯), sodass f gleichmäßig stetig auf $[a, b]$ sein muss.

□

4.4 Elementare Funktionen

Wir wollen in diesem Abschnitt einmal die wichtigsten Funkionenklassen gebündelt vorstellen. Diese werden auch häufig als *elementare Funktionen* bezeichnet, auch wenn dies keine offizielle mathematische Definition ist. Im letzten Abschnitt haben wir gesehen, dass Kombinationen stetiger Funktionen wieder stetig sind. Daher ist es hilfreich, ein gutes Repertoire an solchen stetigen Basisfunktionen zu kennen, aus denen man sich viele (aber nicht alle) in der Praxis relevanten Funktionen zusammenbauen kann, die dann wiederum stetig sind.

Polynome und rationale Funktionen

In diesem Abschnitt betrachten wir zwei wichtige Klassen von Funktionen, die durch ihre einfache Analysierbarkeit und die Möglichkeiten, mit ihnen andere, komplexere Funktionen zu approximieren, in der Analysis eine große Bedeutung haben.

Definition 4.40 (Polynom)

Seien $n \in \mathbb{N}_0$ und $a_0, a_1, \dots, a_n \in \mathbb{R}$. Wir nennen $p : \mathbb{R} \to \mathbb{R}$ mit

$$p(x) = a_n x^n + \cdots + a_2 x^2 + a_1 x + a_0$$

eine *Polynomfunktion* oder einfach *Polynom*.

Das größte $n \in \mathbb{N}_0$ mit $a_n \neq 0$ nennen wir *Grad des Polynoms*.

Wir können Polynome auch kombinieren. Wenn $p(x) = \sum_{i=0}^{n} a_i x^i$ und $q(x) = \sum_{i=0}^{m} b_i x^i$ zwei Polynome sind, und $c \in \mathbb{R}$ ein Skalar, dann ergeben folgende Operationen wieder ein Polynom, wie man leicht nachrechnen kann:

$$\begin{aligned}
(p+q)(x) &= \sum_{i=0}^{\max\{n,m\}} (a_i + b_i)\, x^i, \\
(c \cdot p)(x) &= \sum_{i=0}^{n} (c \cdot a_i)\, x^i, \\
(p \cdot q)(x) &= \sum_{i=0}^{n+m} c_i x^i \quad \text{mit} \quad c_k = \sum_{j=0}^{k} a_j b_{k-j}.
\end{aligned}$$

Die Division von zwei Polynomen ist allerdings kein Polynom mehr, sondern eine sogenannte *rationale Funktion*:

Definition 4.41 (rationale Funktionen)

Seien $p, q : \mathbb{R} \to \mathbb{R}$ Polynome und $A = \{x \in \mathbb{R} \mid q(x) \neq 0\}$. Dann ist $r : A \to \mathbb{R}$ mit

$$r(x) = \left(\frac{p}{q}\right)(x) = \frac{p(x)}{q(x)}$$

eine *rationale Funktion*.

Mit der Anwendung von Satz Satz 4.27 ist jedes Polynom und auch jede rationale Funktion stetig. Wichtig ist vielleicht noch zu erwähnen, dass wir Satz 4.27 nur endlich oft anwenden dürfen. Unendliche Summen von Potenzen (also Potenzreihen) sind nicht automatisch stetig auch wenn ihre

Summanden und Teilsummen jeweils stetig sind. Wenn wir in den beiden letzten Definitionen $x \in \mathbb{C}$ und $a_n \in \mathbb{C}$ wählen, erhalten wir *komplexe* Polynome und *komplexe* rationale Funktionen.

Beispiel 4.42

Ein Anwendungsgebiet von Polynomen ist die Approximation von Funktionen. Ziel ist es, mithilfe eines Polynoms eines vorgegebenen Grades eine Funktion möglichst gut anzugleichen. Dazu werden oft *Stützstellen* verwendet. Es gibt eine Menge von Punkten (x_i, y_i), die von der zu approximierenden Funktion angenommen werden oder aus einer Messung stammen und nun durch eine Funktion beschrieben werden sollen. Es soll nun ein Polynom $p(x)$ gefunden werden, sodass $p(x_i) = y_i$ für alle vorgegebenen Wertepaare (x_i, y_i).

Wir betrachten im Folgenden einen einfachen Algorithmus, der zu $n+1$ Stützstellen (x_i, y_i) $(0 \leq i \leq n)$ ein Polynom $p(x)$ vom Grad n erzeugt, sodass $p(x_i) = y_i$. Das Polynom wird schrittweise generiert, indem das folgende Vorgehen gewählt wird. Wir nehmen dazu an, dass $x_i \neq x_j$ für alle $i \neq j$ gilt, mit $i, j \in \{0, \dots, n\}$:

$$\begin{aligned} p_0(x) &= y_0 \\ p_{k+1}(x) &= p_k(x) + (y_{k+1} - p_k(x_{k+1})) \prod_{j=0}^{k} \frac{x - x_j}{x_{k+1} - x_j} \quad \text{für } 0 \leq k < N. \end{aligned}$$

$p_n(x)$ ist dann die gesuchte Polynomfunktion vom Grad n und es gilt $p_n(x_i) = y_i$ für alle $i \in \{0, \dots, n\}$. Man kann sogar zeigen, dass $p_n(x)$ eindeutig ist, d.h., es gibt keine Polynomfunktion vom gleichen oder niedrigeren Grad, die alle Punkte exakt rekonstruiert.

Wir betrachten als einfaches Beispiel die Punktemenge $\{(0,1), (2,3), (3,0)\}$, die zur Konstruktion der folgenden Polynomfunktion führt:

$$\begin{aligned} p_0(x) &= 1 \\ p_1(x) &= p_0(x) + (y_1 - p_0(x_1)) \frac{x - x_0}{x_1 - x_0} \\ &= 1 + (3-1)\frac{x-0}{2-0} \;=\; x + 1 \\ p_2(x) &= p_1(x) + (y_2 - p_1(x_2)) \frac{x - x_0}{x_2 - x_0} \frac{x - x_1}{x_2 - x_1} \\ &= 1 + x + (0-4)\frac{x-0}{3-0}\frac{x-2}{3-2} \\ &= -\frac{4}{3}x^2 + \frac{11}{3}x + 1. \end{aligned}$$

Die Beispielfunktion gibt die vorgegebenen Punkte exakt wieder und hat einen glatten Verlauf. Es sollte allerdings erwähnt werden, dass dies nicht für alle Funktionen der Fall ist und in vielen Fällen die Hinzunahme neuer Punkte zu einer Polynomfunktion führt, die sehr abrupte Änderungen aufweist und die vorgegeben Funktion nur unzureichend approximiert. Es werden deshalb in der Regel komplexere Verfahren zur Funktionsapproximation mit Hilfe von Polynomfunktionen verwendet, auf die wir nicht weiter eingehen.

▶ Demo: Polynomrekonstruktion durch Stützstellen

Aus der Schule wissen Sie vermutlich, dass jedes Polynom vom Grad n maximal n Nullstellen besitzen kann. Aber warum ist das eigentlich so? Um dies zu zeigen, müssen wir erst eine Art Division mit Rest von Polynomen betrachten, die sogenannte Polynomdivision. Wenn wir ein Polynom $p(x)$ vom Grad n durch ein Polynom $q(x)$ vom Grad m mit $m \leq n$ teilen, ergibt sich ein neues Polynom $s(x)$ vom Grad $n - m$ und ein Restpolynom $r(x)$, dessen Grad kleiner ist als m. Vergleichen Sie dies mit der Division mit Rest Satz 2.29: Auch hier ist das Restglied stets kleiner als die Zahl, durch die wir teilen, und das Ergebnis der Division ist kleiner als die Ausgangszahl. Da der Beweis recht technisch ist, betrachten wir zuerst ein Beispiel für die Polynomdivision:

Beispiel 4.43

Wir berechnen $(3x^4 + x^3 - 2x) : (x^2 + 1)$ nach folgendem Schema:

$$\begin{array}{rl}
(3x^4 + x^3 \quad - 2x \quad\;) : (x^2+1) = 3x^2 + x - 3 \\
\underline{-(3x^4 \quad + 3x^2)} & \\
\quad x^3 - 3x^2 - 2x & \\
\underline{-(x^3 \quad + \;\; x)} & \\
\quad - 3x^2 - 3x & \\
\underline{-(-3x^2 \quad - 3)} & \\
\quad - 3x \;\; + 3 &
\end{array}$$

Die Division wird abgebrochen, da $\mathrm{Grad}(-3x + 3) < \mathrm{Grad}(x^2 + 1)$. Damit ist $s(x) = 3x^2 + x - 3$ und $r(x) = -3x + 3$, also

$$3x^4 + x^3 - 2x = \left(x^2 + 1\right)\left(3x^2 + x - 3\right) - 3x + 3.$$

Wenn das Verfahren der Polynomdivision an Ihrer Schule nicht eingeführt wurde, oder es schon in Vergessenheit geraten ist, und Sie keine Angst vor etwas Fremdscham haben, dann können Sie sich ein berühmtes Youtube-Video dazu ansehen, welches das Verfahren musikalisch ins Gehirn einbrennt.

Nun also noch der eigentliche Satz und Beweis:

Satz 4.44 (Polynomdivision)

Seien $p(x)$ und $q(x)$ Polynome vom Grad n und m mit $m \leq n$. Dann gibt es Polynome $s(x)$ und $r(x)$, so dass

$$p(x) = s(x)q(x) + r(x).$$

Es gilt

$$\mathrm{Grad}(s) = \mathrm{Grad}(p) - \mathrm{Grad}(q) \quad \text{und} \quad \mathrm{Grad}(r) < \mathrm{Grad}(q).$$

▼ Beweis

Sei $p(x) = a_n x^n + \ldots + a_1 x + a_0$ und $q(x) = b_m x^m + \ldots + b_1 x + b_0$. Subtrahiert man von $p(x)$ das Polynom

$$\frac{a_n}{b_m} x^{n-m} \cdot q(x),$$

so erhält man ein Polynom $p_1(x)$ mit $n_1 = \mathrm{Grad}(p_1) < \mathrm{Grad}(p) = n$.

Falls $n_1 < m$, dann sei $r(x) = p_1(x)$ und $s(x) = \frac{a_n}{b_m} x^{m \cdot n}$ und die obige Darstellung wurde erreicht. Ansonsten wird mit $p_1(x)$ und $q(x)$ wie für $p(x)$ und $q(x)$ beschrieben fortgefahren. Diese Schritte werden so lange iteriert, bis ein Polynom $p_i(x)$ mit $\mathrm{Grad}(p_i) < \mathrm{Grad}(q)$ erzeugt wurde. Ein solches Polynom entsteht, da sich der Grad beim Übergang von $p_{i+1}(x)$ auf $p_i(x)$ jeweils um 1 reduziert.

Nun muss noch die Eindeutigkeit der Polynome $s(x)$ und $r(x)$ gezeigt werden. Sei $p(x) = s'(x)q(x) + r'(x)$ mit $s'(x) \neq s(x)$ und $r'(x) \neq r(x)$ eine weitere Darstellung. Dann folgt

$$s(x)q(x) + r(x) = s'(x)q(x) + r'(x) \quad \Leftrightarrow \quad (s'(x) - s(x))\, q(x) = r(x) - r'(x)$$

und damit auch

$$\mathrm{Grad}(s(x) - s'(x))q(x) = \mathrm{Grad}(r(x) - r'(x)) < \mathrm{Grad}q(x).$$

Der Grad eines Produktes zweier Polynome ($q(x)$ und $s(x) - s'(x)$) kann allerdings nicht kleiner sein als der Grad einer der Polynomfaktoren ($q(x)$), es sei denn, es gilt $s(x) - s'(x) = 0$, wonach auch $r(x) - r'(x) = 0$ gilt. Die Polynome $r(x)$ und $s(x)$ sind also eindeutig.

□

Als nächstes betrachten wir bestimmte Polynome für $q(x)$, sogenannte Linearfaktoren $q(x) = x - x_1$ für ein $x_1 \in \mathbb{R}$. Für diese können wir Folgendes zeigen:

Satz 4.45 (Linearfaktoren)

Ein Polynom $p(x)$ lässt sich genau dann ohne Rest durch $q(x) = x - x_1$ teilen ($x_1 \in \mathbb{R}$), wenn x_1 eine Nullstelle von $p(x)$ ist.

▼ Beweis

Wir führen den Beweis in zwei Richtungen. Wir hatten im letzten Satz bereits gezeigt, dass der Grad des Restglieds r einer Polynomdivision immer kleiner ist als des Polynoms q, durch das wir teilen. Der Grad von $q(x) = x - x_1$ ist 1, also ist $r(x) = c$ eine Konstante $c \in \mathbb{R}$.

Richtung 1:

Wenn $r(x) = 0$ ist, dann folgt $p(x_1) = (x_1 - x_1)s(x_1) + 0 = 0$, somit ist x_1 eine Nullstelle von $p(x)$.

Richtung 2:

Wenn x_1 eine Nullstelle ist, dann folgt $p(x_1) = 0 = (x_1 - x_1)s(x_1) + c = c$, somit ist $r(x) = c = 0$.

□

Kombinieren wir Satz 4.44 und Satz 4.45, dann folgt, dass wir $p(x)$ als $(x - x_1)q(x)$ darstellen können, wenn x_1 eine Nullstelle von $p(x)$ ist. Außerdem muss dann $q(x)$ um einen Grad kleiner sein als $p(x)$. Dies können wir maximal n (Grad von $p(x)$) mal durchführen. Damit hat ein Polynom vom Grad n höchstens n Nullstellen.

Für komplexe Polynome lässt sich sogar zeigen, dass diese genau n Nullstellen haben. Dies beweist man, indem man zeigt, dass jedes komplexe Polynom mindestens eine Nullstelle hat. Der formale Beweis hierfür ist leider ein wenig ausufernd, weswegen wir ihn nur graphisch motivieren, bzw. dafür auf das Vorlesungsvideo von Prof. Weitz verweisen. Anschließend kann man wie im letzten Absatz folgern, dass jede Nullstelle als Linearfaktor abgespalten werden kann, wodurch sich der Grad um Eins verringert. Da auch das übrige Polynom wieder mindestens eine komplexe Nullstelle haben muss, ergibt sich nach n-maliger Anwendung dieser Argumentation, dass ein komplexes Polynom vom Grad n genau n Nullstellen besitzt.

Die reelle Exponentialfunktion

Wir haben im Kapitel zu Potenzreihen bereits die Exponentialreihe eingeführt. Die Exponentialfunktion definieren wir nun über eben diese Reihe:

Definition 4.46 (Exponentialfunktion)

Wir definieren die Funktion $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ als

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

Wie wir bereits in Beispiel 4.24 gesehen haben, ist die Exponentialfunktion stetig. Die meisten weiteren Eigenschaften der Exponentialfunktion leiten sich aus den Eigenschaften der Exponentialreihe ab und werden hier nur noch einmal wiederholt:

Satz 4.47 (Eigenschaften der Exponentialfunktion)

Für die Exponentialfunktion $\exp(x)$ gelten folgende Eigenschaften:

a. $\forall x, y \in \mathbb{R}: \ \exp(x+y) = \exp(x) \cdot \exp(y)$

b. $\forall x \in \mathbb{R}: \ \exp(-x) = \frac{1}{\exp(x)}$

c. $\forall x \in \mathbb{R}: \ \exp(x) > 0$

d. $\forall n \in \mathbb{Z}: \ \exp(n) = e^n$

e. $\forall x \in \mathbb{R}: \ \exp(x) = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n$

f. $\exp\colon \mathbb{R} \to \mathbb{R}_{>0}$ ist streng monoton wachsend und bijektiv.

g. Es gilt $\lim_{x\to 0} \frac{\exp(x)-1}{x} = 1.$

▼ Beweis

Die Beweise für (a)-(d) folgen aus Satz 3.73 und (e) aus Satz 3.74. Die Beweise für (f) und (g) überlassen wir Ihnen zur Übung.

Sie werden in ihrem Studium noch häufig das Wachstum zweier Funktionen miteinander vergleichen, um zum Beispiel Aussagen darüber zu treffen, wie viel langsamer ein Algorithmus wird, wenn man die Elemente einer darin vorkommenden Menge verdoppelt (z.B. einen Sortierungsalgorithmus auf ein doppelt so großes Array anwendet). Dabei ist es hilfreich, für die elementaren Funktionen zu wissen, wie schnell diese im Verhältnis zueinander wachsen. Vergleichen wir also das Wachstum von x^k mit dem von $\exp(x)$:

Satz 4.48 (Erster Satz vom Wachstum)

Für beliebige $n \in \mathbb{N}_0$ gilt:

$$\lim_{x\to\infty} \frac{\exp(x)}{x^n} = \infty$$

$$\lim_{x\to-\infty} \exp(x)x^n = 0$$

▼ Beweis

Für $x > 0$ sind alle Terme der Exponentialreihe positiv. Damit gilt durch Weglassen aller Terme bis auf einen

$$\exp(x) > \frac{x^{n+1}}{(n+1)!}$$

und für den Kehrwert

$$0 < \frac{x^n}{\exp(x)} < \frac{(n+1)!}{x}.$$

Da die rechte und linke Seite der Abschätzung für $x \to \infty$ gegen 0 konvergieren, gilt nach dem Sandwich-Theorem

$$\lim_{x\to\infty} \frac{x^n}{\exp(x)} = 0.$$

Nach Satz 3.9 muss der Kehrwert divergieren:

$$\lim_{x\to\infty} \frac{\exp(x)}{x^n} = \infty.$$

Der zweite Teil des Satzes folgt mit der Substitution $x \to -x$ aus dem ersten:

$$\lim_{-x\to-\infty} \exp(-x)(-x)^n = (-1)^n \lim_{x\to\infty} \frac{x^n}{\exp(x)} = 0.$$

□

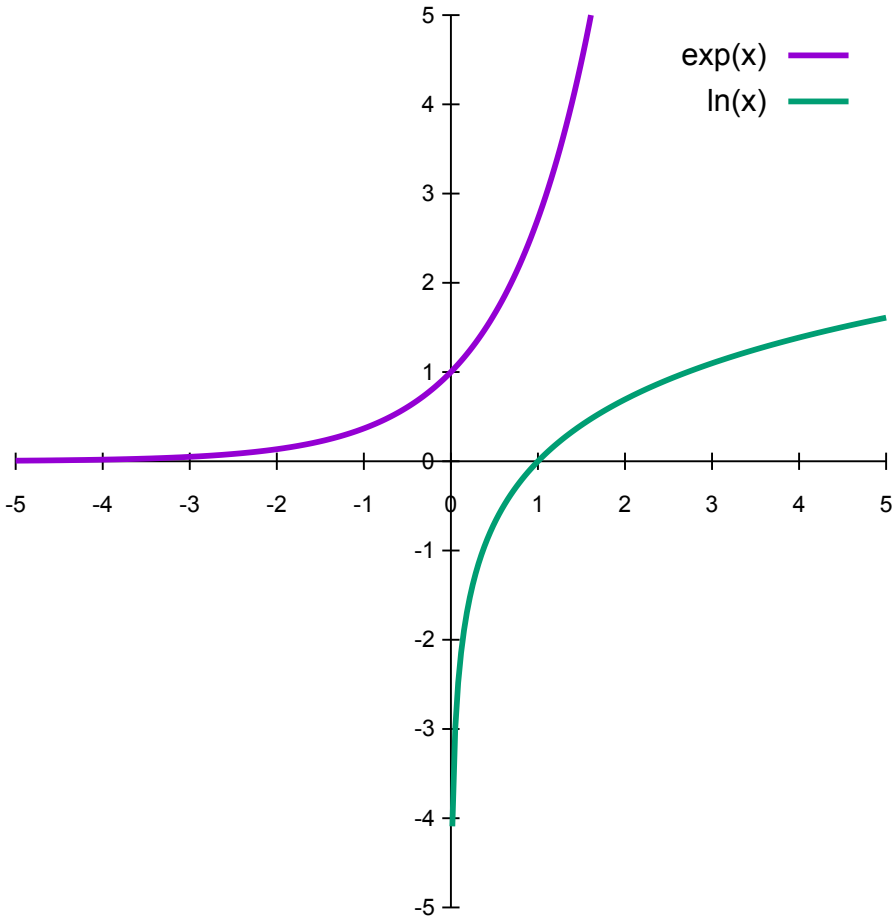
Mit anderen Worten: Die Exponentialfunktion wächst für $x \to \infty$ schneller gegen ∞ als jedes x^n und fällt für $x \to -\infty$ schneller gegen 0 als jedes x^{-n}.

Logarithmen und allgemeine Potenzen

Da die Exponentialfunktion eine bijektive Abbildung $\exp : \mathbb{R} \to \mathbb{R}_{>0}$ ist, existiert eine Umkehrfunktion.

Definition 4.49 (natürlicher Logarithmus)

Wir bezeichnen die Umkehrfunktion der Exponentialfunktion $\ln : \mathbb{R}_{>0} \to \mathbb{R}$ als den *natürlichen Logarithmus*.



Aus der Definition ergeben sich direkt die folgenden Eigenschaften:

Satz 4.50 (Eigenschaften des natürlichen Logarithmus)

Für den natürlichen Logarithmus gilt:

a. $\ln(\exp(x)) = \exp(\ln(x)) = x$

b. $\ln(1) = 0$ und $\ln(e) = 1$

c. $\ln(x) \begin{cases} < 0 & \text{für } x \in (0,1) \\ = 0 & \text{für } x = 1 \\ > 0 & \text{für } x > 1 \end{cases}$

d. $\forall x, y \in \mathbb{R}_{>0}: \ln(xy) = \ln(x) + \ln(y)$

e. $\forall x \in \mathbb{R}_{>0}, n \in \mathbb{Z}: \ln(x^n) = n \ln(x)$

f. Der natürlich Logarithmus ist stetig.

▼ Beweise als Übung

Wir werden später noch sehen, wie wir Werte des natürlichen Logarithmus, wie z.B. $\ln(2)$ beliebig genau approximieren können. Analog zum Satz vom Wachstum der Exponentialfunktion gibt es auch einen nützlichen Wachstumsvergleich für Logarithmen:

Satz 4.51 (Zweiter Satz vom Wachstum)

Für beliebige $n \in \mathbb{N}$ gilt

$$\lim_{x \to \infty} \frac{\ln(x)}{\sqrt[n]{x}} = 0.$$

Der Logarithmus wächst also schwächer als jede Wurzelfunktion.

▼ Beweis

Folgt aus dem ersten Satz vom Wachstum (Satz 4.48) mit der Substitution $x = e^{nx}$.

□

Wie Sie sich vielleicht erinnern, hatten wir für die harmonische Reihe damals gesagt, dass diese “gerade so” nicht konvergiert: Sie wächst zwar auf beliebig hohe Werte an, braucht dafür aber extrem

viele Summanden. Ähnlich verhält es sich mit Logarithmen, und man kann sogar zeigen, dass gilt:

$$\sum_{k=1}^{n} \frac{1}{k} \approx \ln(n) + \gamma \quad \text{mit} \quad \gamma = 0.577215\ldots$$

γ ist die Euler-Mascheroni-Konstante und die Näherung wird für große n immer genauer. Logarithmisches Wachstum wird in der Informatik oft als Vergleichswert verwendet, um die Komplexität eines Algorithmus zu beschreiben (also wie viel länger muss dieser rechnen, wenn man die Größe eines Eingabeelements, z.B. eines Arrays, verdoppelt).

Mithilfe der Logarithmen können wir unsere Potenzdefinition aus Definition 2.34 auch auf allgemeine positive Basen und beliebige Exponenten erweitern:

Definition 4.52 (allgemeine Exponentialfunktion)

Für $a \in \mathbb{R}_{>0}$ sei die *allgemeine Exponentialfunktion zur Basis* a definiert als $\exp_a : \mathbb{R} \to \mathbb{R}$ mit

$$\exp_a(x) := \exp(x \ln(a)).$$

Wir schreiben auch a^x anstelle von $\exp_a(x)$.

Damit der Ausdruck a^x wohldefiniert ist, muss sich für $x = n \in \mathbb{Z}$ wieder unsere alte Definition mit ganzzahligem Exponent ergeben. Außerdem zeigen wir im folgenden Satz, dass sich auch alle anderen Eigenschaften, die wir schon von ganzzahligen Potenzen kennen, übertragen lassen:

Satz 4.53 (Eigenschaften allgemeiner Potenzen)

Sei $a, b \in \mathbb{R}_{>0}$. Es gilt für die allgemeine Exponentialfunktion

a. $\exp_a(x)$ ist stetig für alle $x \in \mathbb{R}$.

b. $\forall n \in \mathbb{Z}: \exp_a(n) = a^n$

c. $\forall x, y \in \mathbb{R}: a^{x+y} = a^x a^y$

d. $\forall x, y \in \mathbb{R}: (a^x)^y = a^{xy}$

e. $\forall x \in \mathbb{R}: a^x b^x = (ab)^x$

f. $\forall p \in \mathbb{Z}, q \in \mathbb{N} \setminus \{1\}: a^{\frac{p}{q}} = \sqrt[q]{a^p}$

▼ Beweis

a. Die Stetigkeit folgt, da $\exp_a$ die Komposition der stetigen Funktionen $\exp$ und $\ln$ ist, die nach Satz 4.29 stetig ist.

b. Beweis per Induktion:

Induktionsannahme:
$n = 1$: $\exp_a 1 = \exp(1 \ln(a)) = a = a^1$

Induktionsvoraussetzung:
Wir nehmen an, die Behauptung gilt für ein $n \in \mathbb{N}$.

Induktionsschritt:
Es folgt für $n + 1$:

$$a^{n+1} = \exp((n+1)\ln(a)) = \exp(n\ln(a))\exp(1\ln(a)) \overset{IV, IA}{=} a^n \cdot a = a^{n+1}.$$

Damit gilt die Behauptung für alle $n \in \mathbb{N}$.

Sie gilt auch offensichtlich für $n = 0$. Mit

$$\exp_a(-n) = \exp(-n\ln(a)) = \frac{1}{\exp(n\ln(a))} = \frac{1}{\exp_a(n)} = \frac{1}{a^n} = a^{-n}$$

gilt sie auch für alle negativen ganzen Zahlen.

c. $a^{x+y} = \exp((x+y)\ln(a)) = \exp(x\ln(a))\exp(y\ln(a)) = a^x a^y.$

d. $(a^x)^y = \exp(y\ln(a^x)) = \exp(yx\ln(a)) = a^{xy}.$

e. $a^x b^x = \exp(x\ln(a))\exp(x\ln(b)) = \exp(x(\ln(a)+\ln(b))) = \exp(x\ln(ab)) = (ab)^x.$

f. Aus (d) folgt: $(a^{\frac{p}{q}})^q = a^p$. Damit ist $x = a^{\frac{p}{q}}$ die Zahl, die für $y = a^p$ die Gleichung $x^q = y$ löst. Die Lösung x dieser Gleichung haben wir als die k-te Wurzel definiert. Also gilt: $x = \sqrt[q]{y}$ bzw. $a^{\frac{p}{q}} = \sqrt[q]{a^p}$.

□

Ähnlich zur Exponentialfunktion kann auch der Logarithmus für allgemeine Basen $a \in \mathbb{R}_{>0} \setminus \{1\}$ definiert werden.

Definition 4.54 (Logarithmen zu allgemeinen Basen)

Sei $a \in \mathbb{R}_{>0} \setminus \{1\}$, dann ist der *Logarithmus zur Basis* a definiert als $\log_a : \mathbb{R}_{>0} \to \mathbb{R}$ mit

$$\log_a(x) := \frac{\ln(x)}{\ln(a)}.$$

Der allgemeine Logarithmus löst die Gleichung $a^x = b$ mit $x = \log_a(b)$. Die Schreibweise zu Logarithmen ist in der Literatur nicht einheitlich. Manchmal bezeichnet $\log(x)$ den natürlichen Logarithmus, den wir mit $\ln(x)$ bezeichnen. Manchmal sieht man auch die Schreibweise $\lg(x)$ für den dekadischen Logarithmus ($a = 10$) und $\mathrm{lb}(x)$ für den binären Logarithmus ($a = 2$).

Beispiel 4.55 (irrationale Logarithmen)

Wir haben vor einiger Zeit in Beispiel 2.31 gezeigt, dass $\sqrt{2}$ nicht rational ist. Auch Logarithmen sind in viele Fällen nicht rational und hier ist der Widerspruchsbeweis noch deutlich offensichtlicher. Damals konnten wir Logarithmen leider noch nicht nutzen, weswegen $\sqrt{2}$ das Standardbeispiel für eine irrationale Zahl ist.

Wir zeigen hier exemplarisch, dass $\log_2 5$ nicht rational ist.

▼ Beweis

Es gilt

$$a^x = b \quad \Leftrightarrow \quad x = \log_a b$$

also in diesem Fall:

$$2^x = 5 \quad \Leftrightarrow \quad x = \log_2 5.$$

Wir zeigen nun, dass das x nicht rational sein kann per Widerspruchsbeweis. Angenommen, es gäbe eine Lösung von $2^x = 5$ mit $x \in \mathbb{Q}$, also $x = p/q$ für ein $p \in \mathbb{Z}$ und $q \in \mathbb{N}$ (für dieses Beispiel muss x übrigens kein gekürzter Bruch sein).

Dann folgt:

$$2^{\frac{p}{q}} = 5.$$

Wenn wir beide Seiten der Gleichung 'hoch q' nehmen, ergibt sich:

$$(2^{\frac{p}{q}})^q = 2^p = 5^q.$$

Die linke Seite der Gleichung ist ein Produkt aus Zweien und damit eine gerade Zahl. Dagegen ist die rechte Seite ein Produkt aus Fünfen, also ungerade. Dies ist ein Widerspruch, woraus folge, dass keine rationale Lösung von $2^x = 5$ existiert.

Die Existenz einer reellen Lösung ist über die Stetigkeit der Exponentialfunktion garantiert, welche auf der Konvergenz von Folgen und damit auf der Vollständigkeit der reellen Zahlen beruht. $\log_2 5$ ist also irrational.

□

Beispiel 4.56

Die Exponentialfunktion kann genutzt werden, um Wachstums- oder Zerfallsprozesse zu beschreiben. So kann radioaktiver Zerfall durch die Formel

$$n(t) = n(0) \cdot e^{-\lambda t}$$

beschrieben werden. Dabei ist $t \geq 0$ der Zeitparameter und $n(t)$ die Anzahl der Kerne, die noch nicht zerfallen sind. Entsprechend ist $n(0)$ die Anzahl der Kerne zu Beginn des Experiments. λ ist eine materialspezifische Zerfallskonstante, die zum Beispiel für Jod-131 $10^{-6}s^{-1}$ beträgt. Wenn wir diesen Wert auf Tage umrechnen erhalten wir $\lambda = 0.086d^{-1}$.

Die Halbwertszeit eines Stoffes ist die Zeit τ, nach der die Hälfte der Kerne zerfallen ist. Es muss also gelten

$$n(\tau) = n(0) \cdot e^{-\lambda \tau} = \frac{n(0)}{2}$$

und somit

$$\ln\left(e^{-\lambda\tau}\right) = \ln\left(\frac{1}{2}\right) \quad \Leftrightarrow \quad -\lambda\tau = \ln\left(\frac{1}{2}\right) \quad \Leftrightarrow \quad \tau = \frac{\ln(2)}{\lambda},$$

da $\ln\left(\frac{1}{2}\right) = -\ln(2)$. Für Jod 131 beträgt die Halbwertszeit ca. 8 Tage.

Hyperbolische Funktionen

Wir betrachten in diesem Abschnitt Funktionen, die normalerweise nicht zum Schulstoff gehören, aber trotzdem an vielen Stellen in der Praxis vorkommen. Außerdem bereiten sie die Definition der Sinus- und Cosinusfunktion vor. Beginnen wir mit der Definition einer wichtigen Funktionseigenschaft, der *Symmetrie*.

Definition 4.57 (Funktionssymmetrie)

Sei $A \subseteq \mathbb{R}$ eine Menge für die gilt: $x \in A \Rightarrow -x \in A$.

Wir nennen eine Funktion $f : A \to \mathbb{R}$

- *achsensymmetrisch* oder *gerade*, wenn

$$\forall x \in A : f(-x) = f(x),$$

- *punktsymmetrisch* oder *ungerade*, wenn

$$\forall x \in A : f(-x) = -f(x).$$

Beispiel 4.58

- $f(x) = c$ ist gerade.
- $f(x) = \frac{1}{x}$ ist ungerade.
- $f(x) = \frac{1}{x^2+1}$ ist gerade.
- $f(x) = x^3 + 4x^2$ ist weder gerade noch ungerade.

Wir können jede Funktion mit symmetrischem Definitionsbereich in eine gerade Funktion $g(x)$ und eine ungerade Funktion $u(x)$ aufteilen, sodass

$$\begin{aligned} f(x) &= g(x) + u(x) \\ f(-x) &= g(-x) + u(-x) = g(x) + (-u(x)). \end{aligned}$$

Daraus folgt die Aufteilung

$$\begin{aligned} g(x) &= \frac{f(x) + f(-x)}{2} \\ u(x) &= \frac{f(x) - f(-x)}{2}. \end{aligned}$$

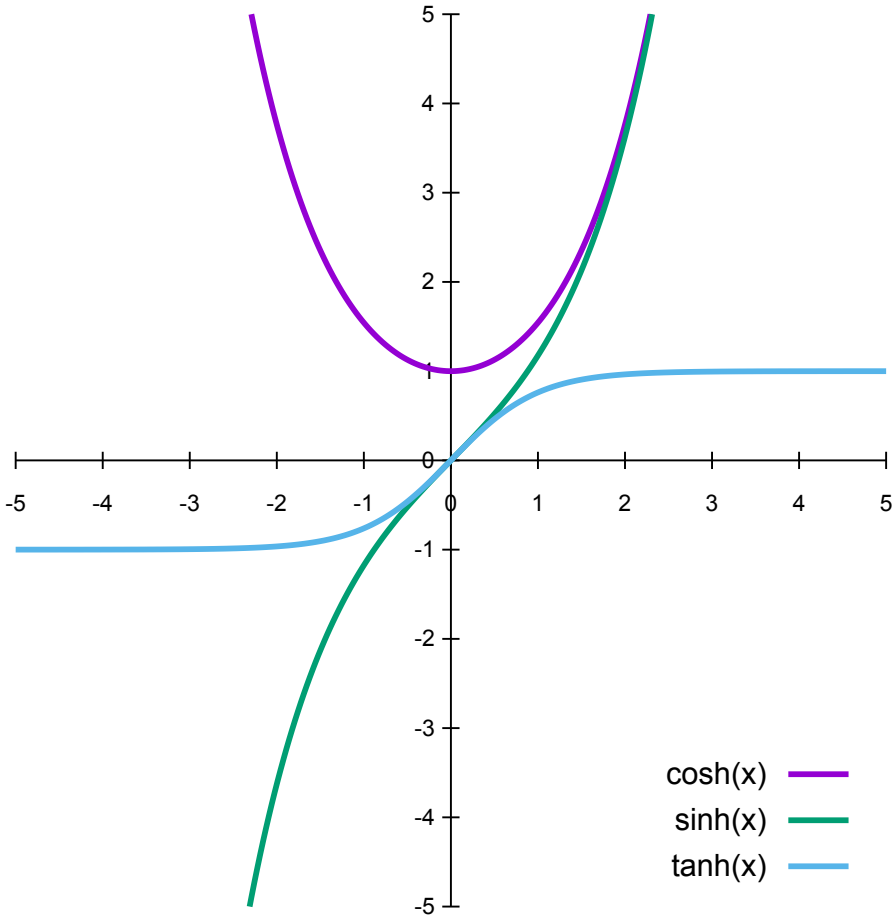
Wir können leicht erkennen, dass $g(x)$ auf diese Art immer gerade und $u(x)$ immer ungerade ist. Wenden wir dies auf die Exponentialfunktion an, lassen sich die *hyperbolischen* Funktionen definieren:

Definition 4.59 (hyperbolische Funktionen)

Wir definieren die hyperbolischen Funktionen (jeweils $\mathbb{R} \to \mathbb{R}$):

$$\begin{aligned} \cosh(x) &:= \frac{e^x + e^{-x}}{2} && \text{(Cosinus hyperbolicus)} \\ \sinh(x) &:= \frac{e^x - e^{-x}}{2} && \text{(Sinus hyperbolicus)} \\ \tanh(x) &:= \frac{\sinh(x)}{\cosh(x)} && \text{(Tangens hyperbolicus)} \end{aligned}$$

Die hyperbolischen Funktionen sind stetig, da die Exponentialfunktion stetig ist.



Wir werden an dieser Stelle noch nichts zu den hyperbolischen Funktionen beweisen, da die Beweise alle recht einfach aus der Definition folgen und sich gut als Übung eignen. An dieser Stelle sollten Sie die Funktionen nur einmal gesehen haben und mit dem geraden und ungeraden Anteil der Exponentialfunktion in Verbindung bringen. Hier eine kleine Auswahl an Zusammenhängen:

Satz 4.60 (Eigenschaften der hyperbolischen Funktionen)

Für die hyperbolischen Funktionen gelten für alle $x, y \in \mathbb{R}$ die folgenden Eigenschaften:

a. $\exp(x) = \cosh(x) + \sinh(x)$

b. $\cosh^2(x) - \sinh^2(x) = 1$

c. $\cosh(x) = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}$

d. $\sinh(x) = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}$

e. $\cosh(x+y) = \cosh(x)\cosh(y) + \sinh(x)\sinh(y)$

f. $\sinh(x+y) = \sinh(x)\cosh(y) + \cosh(x)\sinh(y)$

▼ Beweis als Übung

Die komplexe Exponentialfunktion

Analog zur Definition der reellen Exponentialfunktion können wir die komplexe Exponentialfunktion über die komplexe Exponentialreihe definieren:

Definition 4.61 (komplexe Exponentialfunktion)

Wir definieren die *komplexe Exponentialfunktion* als

$$\exp : \mathbb{C} \to \mathbb{C} \quad \text{mit} \quad \exp(z) = e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

Offensichtlich sind die reelle und komplexe Exponentialfunktion für reelle Argumente identisch, weswegen wir dieselbe Abkürzung ($\exp$) verwenden. Da die Reihe für beliebige $z \in \mathbb{C}$ konvergiert, ist die komplexe Exponentialfunktion wohldefiniert. Wir können fast alle Eigenschaften der Exponentialfunktion auch im Komplexen anwenden. Aber $\exp(z) > 0$ sowie die Monotonie lassen sich im Komplexen nicht folgern, da die komplexen Zahlen kein geordneter Körper sind.

Die komplexe Exponentialfunktion ist die wohl wichtigste komplexe Funktion, welche an zahlreichen Stellen in Naturwissenschaften und Technik auftaucht. Sie beschreibt, wie wir noch zeigen werden,

allgemeine Schwingungen. Auch Funktionen, die Sie eventuell mit Schwingungen assoziieren (Sinus und Cosinus) lassen sich auf die komplexe Exponentialfunktion zurückführen.

Wir haben gezeigt, dass wir die komplexe Konjugation einer Summe oder eines Produktes in die einzelnen Summanden/Faktoren ziehen können. Damit ergibt sich mit $z = a + \mathrm{i}b$ für $a, b \in \mathbb{R}$:

$$\overline{e^z} = e^{\overline{z}} = e^{a-\mathrm{i}b}.$$

Für den Betrag der komplexen Exponentialfunktion ergibt sich demnach

$$|e^z| = \sqrt{e^z \overline{e^z}} = \sqrt{e^{a+\mathrm{i}b} e^{a-\mathrm{i}b}} = \sqrt{e^{2a}} = e^a.$$

Der Betrag eines Funktionswerts e^z entspricht also der reellen Exponentialfunktion des Realteils von z. Im nächsten Abschnitt beschäftigen wir uns mit rein imaginären Argumenten der Exponentialfunktion, also $z = \mathrm{i}x, x \in \mathbb{R}$. Hierfür ergibt sich nach der vorigen Betrachtung

$$\left|e^{\mathrm{i}x}\right| = e^0 = 1.$$

Funktionswerte für rein imaginäre Argumente liegen also auf dem Einheitskreis (mit Radius 1) in der komplexen Ebene. Wir werden im folgenden Abschnitt den Real- und Imaginärteil dieser Funktionswerte genauer betrachten.

Trigonometrische Funktionen

Wir kombinieren nun unser Wissen aus den zwei vorangegangenen Kapiteln. Aus der Schule kennen Sie vielleicht die Darstellung der trigonometrischen Funktionen am Einheitskreis: Der x-Wert zu einem Punkt auf dem Einheitskreis ist der Cosinus des entsprechenden Winkels, und der dazugehörige y-Wert ist der Sinus. Etwas Ähnliches machen wir nun auch in der komplexen Ebene: Hier ist der x-Wert der Realteil und der y-Wert der Imaginärteil eines Funktionswertes der komplexen Exponentialfunktion.

▶ Demo: Sinus und Cosinus in der komplexen Ebene

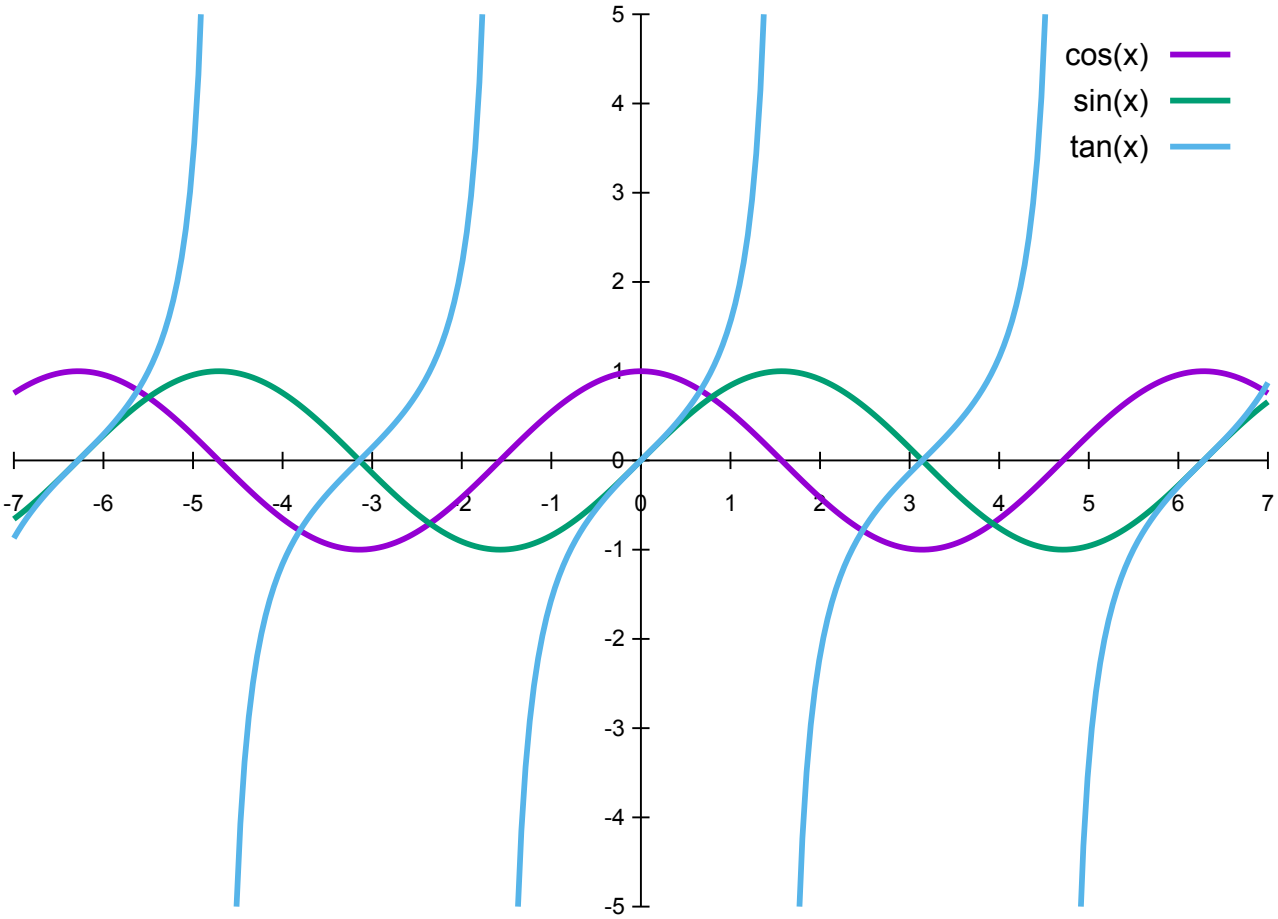
Definition 4.62 (Trigonometrische Funktionen)

Wir definieren die trigonometrischen Funktionen $\sin\colon \mathbb{R} \to \mathbb{R}$, $\cos\colon \mathbb{R} \to \mathbb{R}$ und $\tan\colon \{x \in \mathbb{R} \mid \cos(x) \neq 0\} \to \mathbb{R}$ als

$$\cos(x) := \mathrm{Re}(e^{\mathrm{i}x}) = \frac{e^{\mathrm{i}x} + e^{-\mathrm{i}x}}{2} \qquad \text{(Cosinus)}$$

$$\sin(x) := \mathrm{Im}(e^{\mathrm{i}x}) = \frac{e^{\mathrm{i}x} - e^{-\mathrm{i}x}}{2\mathrm{i}} \qquad \text{(Sinus)}$$

$$\tan(x) := \frac{\sin(x)}{\cos(x)} \qquad \text{(Tangens).}$$



Hierbei wurde ausgenutzt, dass für den Realteil $a = \mathrm{Re}(z)$ einer komplexen Zahl $z = a + \mathrm{i}b$ gilt $a = (z + \overline{z})/2$, sowie für den Imaginärteil: $b = \mathrm{Im}(z) = (z - \overline{z})/(2\mathrm{i})$. Sinus und Cosinus sind stetig, da sie sich aus der stetigen Exponentialfunktion ergeben. Vermutlich fällt Ihnen die große Ähnlichkeit zu den hyperbolischen Funktionen auf, tatsächlich gilt $\cosh(\mathrm{i}x) = \cos(x)$ und $\sinh(\mathrm{i}x) = \mathrm{i}\sin(x)$. Wir können für die trigonometrischen Funktionen ähnliche Eigenschaften nachweisen, wie für die hyperbolischen Funktionen:

Satz 4.63 (Eigenschaften der trigonometrischen Funktionen)

Für die trigonometrischen Funktionen gelten für alle $x, y \in \mathbb{R}$ die folgenden Eigenschaften:

a. $\exp(\mathrm{i}x) = \cos(x) + \mathrm{i}\sin(x)$ (*Eulersche Formel*)

b. $\cos^2(x) + \sin^2(x) = 1$

c. $|\sin(x)| \leq 1$ und $|\cos(x)| \leq 1$

d. $\cos(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}$

e. $\sin(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}$

f. $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$

g. $\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y)$

▼ Beweis

a. Folgt direkt aus der Definition.

b. $\left(\frac{e^{\mathrm{i}x}+e^{-\mathrm{i}x}}{2}\right)^2 + \left(\frac{e^{\mathrm{i}x}-e^{-\mathrm{i}x}}{2\mathrm{i}}\right)^2 = (e^{2\mathrm{i}x} + 2 + e^{-2\mathrm{i}x})/4 + (e^{2\mathrm{i}x} - 2 + e^{-2\mathrm{i}x})/(-4) = 1.$

c. Wäre einer der Beträge größer als 1, dann folgt $\sin^2(x) + \cos^2(x) = |\sin(x)|^2 + |\cos(x)|^2 > 1$, was im Widerspruch zu (b) steht.

d. Hierfür muss die Exponentialreihe in ihre geraden und ungeraden Anteile zerlegt werden

$$\begin{aligned}
\frac{e^{\mathrm{i}x} + e^{-\mathrm{i}x}}{2} &= \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{(\mathrm{i}x)^k}{k!} + \frac{(-\mathrm{i}x)^k}{k!}\right) \\
&= \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{(\mathrm{i}x)^{2k}}{(2k)!} + \frac{(-\mathrm{i}x)^{2k}}{(2k)!}\right) + \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{(\mathrm{i}x)^{2k+1}}{(2k+1)!} + \frac{(-\mathrm{i}x)^{2k+1}}{(2k+1)!}\right) \\
&= \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{(\mathrm{i}x)^{2k}}{(2k)!} + \frac{(\mathrm{i}x)^{2k}}{(2k)!}\right) + \frac{1}{2}\sum_{k=0}^{\infty}\left(\frac{(\mathrm{i}x)^{2k+1}}{(2k+1)!} - \frac{(\mathrm{i}x)^{2k+1}}{(2k+1)!}\right) \\
&= \sum_{k=0}^{\infty}\frac{\mathrm{i}^{2k}x^{2k}}{(2k)!} \\
&= \sum_{k=0}^{\infty}(-1)^k\frac{x^{2k}}{(2k)!}
\end{aligned}$$

e. Analog zu (d).

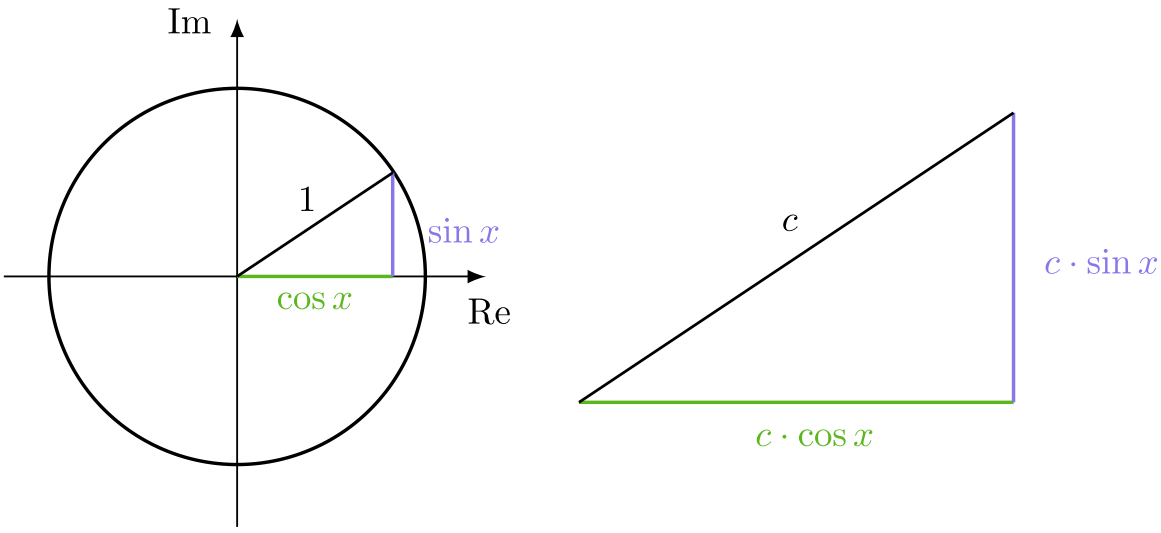
f. Wir beginnen von rechts:

$$
\begin{aligned}
&\cos(x)\cos(y) - \sin(x)\sin(y)\\
&= \frac{(e^{\mathrm{i}x} + e^{-\mathrm{i}x})(e^{\mathrm{i}y} + e^{-\mathrm{i}y})}{4} - \frac{(e^{\mathrm{i}x} - e^{-\mathrm{i}x})(e^{\mathrm{i}y} - e^{-\mathrm{i}y})}{4\mathrm{i}^2}\\
&= \frac{e^{\mathrm{i}(x+y)} + e^{\mathrm{i}(x-y)} + e^{\mathrm{i}(y-x)} + e^{-\mathrm{i}(x+y)}}{4} + \frac{e^{\mathrm{i}(x+y)} - e^{\mathrm{i}(x-y)} - e^{\mathrm{i}(y-x)} + e^{-\mathrm{i}(x+y)}}{4}\\
&= \frac{e^{\mathrm{i}(x+y)} + e^{-\mathrm{i}(x+y)}}{2} = \cos(x+y).
\end{aligned}
$$

g. Analog zu (f).

□

Die letzten beiden Erkenntnisse des Satzes gehören zu einer Reihe sogenannter *Additionstheoreme*. Diese sind in der Schule nur sehr schwer nachvollziehbar, lassen sich aber hier recht einfach aus der Darstellung über die komplexe Exponentialfunktion herleiten. Die Eigenschaft (b) ist im Grunde der *Satz des Pythagoras*. Die bekanntere Form $a^2 + b^2 = c^2$ für ein rechtwinkliges Dreieck mit Katheten a, b und Hypotenuse c ergibt sich, wenn man das Dreieck aus der letzten Demo um den Faktor c vergrößert (siehe nachfolgende Abbildung).



Wir können über die Reihendarstellung nun die ungefähre Lage bestimmter Funktionswerte auf dem Einheitskreis berechnen. So gilt beispielsweise

$$
\cos(1) = 1 - \frac{1}{2!} + \frac{1}{4!} - \frac{1}{6!} + \frac{1}{8!} - \frac{1}{10!} + \ldots = \underbrace{1 - \frac{1}{2!}}_{>0} + \underbrace{\frac{1}{4!} - \frac{1}{6!}}_{>0} + \underbrace{\frac{1}{8!} - \frac{1}{10!}}_{>0} + \ldots > 0
$$

und

$$
\sin(1) = \frac{1}{1!} - \frac{1}{3!} + \frac{1}{5!} - \frac{1}{7!} + \frac{1}{9!} - \frac{1}{11!} + \ldots = \underbrace{\frac{1}{1!} - \frac{1}{3!}}_{>0} + \underbrace{\frac{1}{5!} - \frac{1}{7!}}_{>0} + \underbrace{\frac{1}{9!} - \frac{1}{11!}}_{>0} + \ldots > 0.
$$

Damit sind sowohl Sinus als auch Cosinus für $x = 1$ positiv und e^{i} liegt im ersten Quadranten auf dem Einheitskreis. Wir haben bereits in Kapitel 2.7 die Vermutung formuliert, dass das Produkt zweier komplexer Zahlen einer Drehung in der komplexen Ebene entspricht. Demnach müsste der Wert von $e^{2\mathrm{i}} = e^{\mathrm{i}} \cdot e^{\mathrm{i}}$ doppelt so weit von der reellen Achse aus gedreht sein wie e^{i}. Trotzdem hilft uns das nicht wirklich weiter, ohne die Lage eines Punktes (außer $x = 0$) genau zu kennen. Ähnlich wie auch schon bei der reellen Exponentialfunktion, bei der wir die Eulersche Zahl einfach als $\exp(1)$ definiert haben (ohne den Wert genau zu kennen), werden wir nun hier einen Punkt der komplexen Exponentialfunktion (genauer gesagt ihres Realteils, also der Cosinusfunktion) festlegen. Damit definieren wir eine zweite sehr wichtige mathematische Konstante, die Kreiszahl π. Wie bei der Eulerschen Zahl müssen wir dafür den genauen Wert von π gar nicht kennen und werden trotzdem viele wichtige Schlüsse für die trigonometrischen Funktionen daraus ableiten. Für die Definition von π beweisen wir zunächst zwei nützliche Abschätzungen.

Satz 4.64 (Abschätzungen für Sinus und Cosinus)

Für $x \in (0, 2]$ gilt:

$$1 - \frac{x^2}{2} < \cos(x) < 1 - \frac{x^2}{2} + \frac{x^4}{4!}$$
$$x - \frac{x^3}{3!} < \sin(x) < x.$$

▼ Beweis

Wir führen den Beweis für den Cosinus, die Sinusabschätzung erfolgt analog.

Betrachten wir zwei aufeinanderfolgende Summanden der Cosinusreihe (siehe Satz 4.63d):

$$\frac{|a_{k+1}|}{|a_k|} = \frac{\frac{x^{2k+2}}{(2k+2)!}}{\frac{x^{2k}}{(2k)!}} = \frac{x^2}{(2k+1)(2k+2)}.$$

Für $x \leq 2$ folgt $\frac{|a_{k+1}|}{|a_k|} < 1$. Jeder Summand ist demnach betragsmäßig kleiner als sein Vorgänger. Daraus folgt:

$$\cos(x) = 1 - \tfrac{x^2}{2} + \underbrace{\left(\frac{x^4}{4!} - \frac{x^6}{6!}\right)}_{>0} + \underbrace{\left(\frac{x^8}{8!} - \frac{x^10}{10!}\right)}_{>0} + \ldots + \underbrace{\left(\frac{x^{2k}}{(2k)!} - \frac{x^{2k+2}}{(2k+2)!}\right)}_{>0} + \ldots > 1 - \tfrac{x^2}{2}$$

und analog

$$\cos(x) = 1 - \tfrac{x^2}{2} + \tfrac{x^4}{4!} + \underbrace{\left(-\frac{x^6}{6!} + \frac{x^8}{8!}\right)}_{<0} - \ldots + \underbrace{\left(-\frac{x^{2k}}{(2k)!} + \frac{x^{2k+2}}{(2k+2)!}\right)}_{<0} - \ldots < 1 - \tfrac{x^2}{2} + \tfrac{x^4}{4!}.$$

Also insgesamt $1 - \frac{x^2}{2} < \cos(x) < 1 - \frac{x^2}{2} + \frac{x^4}{4!}$.

□

Solche Abschätzungen ergeben sich häufig für alternierende Reihen und sind sehr nützlich, um den Reihengrenzwert abzuschätzen. Dies nutzen wir im folgenden Satz:

Satz 4.65 (Nullstelle des Cosinus)

$\cos(x)$ besitzt im Intervall $(0, 2)$ genau eine Nullstelle.

▼ Beweis

Aus der Abschätzung des Cosinus folgt $\cos(2) < 1 - 2 + 2/3 = -1/3$. Es gilt außerdem $\cos(0) = 1$.

Nach dem Zwischenwertsatz existiert also mindestens ein $x \in (0, 2)$ mit $\cos(x) = 0$. Es fehlt noch zu zeigen, dass dies die einzige Nullstelle in diesem Intervall ist. Dafür zeigen wir, dass $\cos(x)$ in diesem Intervall streng monoton fallend ist. Wir nutzen dazu ein weiteres Additionstheorem (Beweis zur Übung):

$$\cos(x) - \cos(y) = -2 \sin\left(\frac{x+y}{2}\right) \sin\left(\frac{x-y}{2}\right).$$

Für $x, y \in [0, 2)$ und $x > y$ gilt die Abschätzung für den Sinus aus Satz 4.64, also:

$$\sin\left(\frac{x+y}{2}\right) > \frac{x+y}{2} > 0$$
$$\sin\left(\frac{x-y}{2}\right) > \frac{x-y}{2} > 0$$

Damit folgt insgesamt $\cos(x) - \cos(y) < 0$, und somit ist $\cos(x)$ im Intervall $(0, 2)$ streng monoton fallend. Demnach kann nur eine Nullstelle existieren.

□

Da wir nun wissen, dass genau eine Nullstelle in diesem Intervall existiert, können wir dieser einen Namen geben: $\pi/2$. Darüber definiert sich die berühmte Kreiszahl. Wir werden später noch zeigen, dass π auch das Verhältnis zwischen Umfang und Durchmesser eines Kreises beschreibt, womit die Konstante normalerweise assoziiert wird. An dieser Stelle würde uns die Kreisumfangdefinition jedoch nicht weiterhelfen, weswegen wir diese (für Sie vielleicht etwas ungewöhnlich wirkende) Definition wählen.

Definition 4.66 (Pi)

Sei s die Nullstelle der Cosinusfunktion im Intervall $(0, 2)$. Dann definieren wir

$$\pi := 2s$$

Wir bezeichnen Vielfache und Teile von π, wie 2π oder 0.2π auch als *Winkel im Bogenmaß*. Wir können diese außerdem in das im Alltag gebräuchlichere *Gradmaß* umrechnen. Wenn x ein Winkel im Bogenmaß ist, dann ist der entsprechende Gradmaßwinkel $180x/\pi$. Zur Verdeutlichung schreiben wir einen kleinen Kreis an den Winkel im Gradmaß, also z.B. 360°.

Obwohl wir den Wert von π noch nicht genau kennen (wir wissen bisher, dass $1 < \pi/2 < 2$, denn wir haben gezeigt, dass $\cos(1) > 0$ und $\cos(2) < 0$ gilt), können wir damit jetzt für Vielfache und Teile von π die Sinus und Cosinusfunktionswerte bestimmen:

Satz 4.67 (Folgerungen aus der Definition von Pi)

Es gilt für die Funktionswerte der trigonometrischen Funktionen

x	0	$\pi/6$	$\pi/4$	$\pi/3$	$\pi/2$	π	$3\pi/2$	2π
$\cos(x)$	1	$\frac{\sqrt{3}}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{1}{2}$	0	-1	0	1
$\sin(x)$	0	$\frac{1}{2}$	$\frac{1}{\sqrt{2}}$	$\frac{\sqrt{3}}{2}$	1	0	-1	0
$\tan(x)$	0	$\frac{1}{\sqrt{3}}$	1	$\sqrt{3}$	–	0	–	0

Allgemein gilt für alle $x \in \mathbb{R}$:

$$\begin{aligned} \cos(x+\pi/2) &= -\sin(x) & \qquad \sin(x+\pi/2) &= \cos(x) \\ \cos(x+\pi) &= -\cos(x) & \qquad \sin(x+\pi) &= -\sin(x) \\ \cos(x+2\pi) &= \cos(x) & \qquad \sin(x+2\pi) &= \sin(x) \end{aligned}$$

▼ Beweis

Wir betrachten hier exemplarisch ein paar ausgewählte Werte der Tabelle, die übrigen können Sie analog als Übung bestimmen:

Aus $\sin^2(\pi/2) + \cos^2(\pi/2) = 1$ und $\cos(\pi/2) = 0$ folgt $|\sin(\pi/2)| = 1$. Da wir bereits mit der Abschätzung des Sinus argumentiert haben, dass der Sinus bei $x = \pi/2$ positiv ist, folgt $\sin(\pi/2) = 1$.

Es gilt also insgesamt $e^{\mathrm{i}\pi/2} = 0 + \mathrm{i} = \mathrm{i}$ und damit folgt

$$e^{\mathrm{i}\pi} = e^{\mathrm{i}\pi/2} e^{\mathrm{i}\pi/2} = \mathrm{i}^2 = -1 + \mathrm{i} \cdot 0,$$

demnach ist $\cos(\pi) = -1$ und $\sin(\pi) = 0$. Analog folgt für $e^{\mathrm{i}\pi/4} = a + \mathrm{i}b$, dass

$$e^{\mathrm{i}\pi/4} e^{\mathrm{i}\pi/4} = (a + \mathrm{i}b)^2 = (a^2 - b^2) + \mathrm{i}(2ab) = e^{\mathrm{i}\pi/2} = \mathrm{i}$$

gelten muss, also $a^2 - b^2 = 0$ und $2ab = 1$. Daraus folgt $a = b = 1/\sqrt{2}$ und insgesamt $\cos(\pi/4) = \sin(\pi/4) = 1/\sqrt{2}$. Aus ähnlichen Überlegungen ergeben sich die übrigen Tabellenwerte.

Für den zweiten Teil des Satzes benutzt man die Additionstheoreme aus Satz 4.63 und die entsprechenden Tabelleneinträge des ersten Satzteils:

$$\begin{aligned} \cos(x+\pi/2) &= \cos(x)\cos(\pi/2) - \sin(x)\sin(\pi/2) & &= \cos(x)\cdot 0 - \sin(x)\cdot 1 = -\sin(x) \\ \cos(x+\pi) &= \cos(x)\cos(\pi) - \sin(x)\sin(\pi) & &= \cos(x)\cdot(-1) - \sin(x)\cdot 0 = -\cos(x) \\ \cos(x+2\pi) &= \cos(x)\cos(2\pi) - \sin(x)\sin(2\pi) & &= \cos(x)\cdot 1 - \sin(x)\cdot 0 = \cos(x) \end{aligned}$$

Die Äquivalenzen für den Sinus beweist man analog.

□

Ein Nebenprodukt des letztes Satzes ist die Formel, die von Mathematiker*innen zur “schönsten Formel der Welt” gekürt wurde:

$$e^{\mathrm{i}\pi} + 1 = 0.$$

Diese Formel enthält immerhin fünf der wichtigsten mathematischen Konstanten: das Nullelement, das Einselement, die Eulersche Zahl, die imaginäre Einheit und die Kreiszahl. In der folgenden Demo wird diese Formel benutzt, um einen anschaulichen Zusammenhang zwischen unserer Definition von π und der herkömmlichen Variante über den Kreisumfang herzustellen.

▶ Demo: Pi und der Kreis

Wegen der letzten beiden Eigenschaften des vorigen Satzes sind Sinus und Cosinus (und Tangens und die komplexe Exponentialfunktion) sogenannte *periodische Funktionen*:

Definition 4.68 (Periodische Funktion)

Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt *periodische Funktion*, wenn es ein $p > 0$ gibt, so dass

$$f(x) = f(x + p)$$

für alle $x \in \mathbb{R}$.

Das kleinste $p \in \mathbb{R}_{>0}$ mit der obigen Eigenschaft heißt *Periode der Funktion* f.

$\sin(x)$ und $\cos(x)$ haben also die Periode 2π. Zuletzt geben wir noch alle Nullstellen der Sinus und Cosinusfunktion an:

Satz 4.69 (Nullstellen von Sinus und Cosinus)

Die Sinusfunktion besitzt genau die Nullstellen $x_k = k\pi$ mit $k \in \mathbb{Z}$.

Die Cosinusfunktion besitzt genau die Nullstellen $x_k = k\pi + \pi/2$ mit $k \in \mathbb{Z}$.

▼ Beweis

Wir haben bereits gezeigt, dass $x_0 = \pi/2$ die einzige Nullstelle der Cosinusfunktion für $x \in [0, 2]$ ist, also ist x_0 auch die einzige Nullstelle im kleineren Intervall $[0, \pi/2]$. Da der Cosinus achsensymmetrisch ist, ist x_0 auch die einzige Nullstelle im Intervall $(-\pi/2, \pi/2]$.

Da $\cos(x + \pi) = -\cos(x)$ gilt, sind demnach $x_0 = \pi/2$ und $x_1 = \pi/2 + \pi$ die einzigen Nullstellen in $(-\pi/2, \pi/2 + \pi]$.

Dieses Intervall hat eine Länge von 2π. Da der Cosinus 2π-periodisch ist, liegen die einzigen Nullstellen somit bei $x_k = k\pi + \pi/2 \; \forall k \in \mathbb{Z}$.

Die Sinusnullstellen folgen durch $\sin(x) = -\cos(x + \pi/2)$.

□

Wenn wir Sinus, Cosinus und Tangens jeweils auf ein Gebiet einschränken, auf dem sie streng monoton sind, also zum Beispiel $\cos : [0, \pi] \to [-1, 1]$, können wir jeweils Umkehrfunktionen definieren. Diese nennt man Arcussinus $\arcsin(x)$, Arcuscosinus $\arccos(x)$ und Arcustangens $\arctan(x)$. Oft werden sie auch mit $\operatorname{asin}(x)$, $\operatorname{acos}(x)$ und $\operatorname{atan}(x)$ bezeichnet. Mit den Eigenschaften dieser Funktionen werden wir uns in den Übungen beschäftigen.

Die Polarkoordinaten komplexer Zahlen

Als Nebenprodukt aus den Erkenntnissen des letzten Kapitels können wir noch eine alternative Schreibweise für komplexe Zahlen einführen, mit welcher sich manche Rechnungen deutlich vereinfachen lassen. Wir hatten bei der Einführung komplexer Zahlen bereits gezeigt, dass wir jede komplexe Zahl z als $z = |z|\,\hat{z}$ für ein $\hat{z} \in \{z \in \mathbb{C} \mid |\hat{z}| = 1\}$. Damit liegt $\hat{z}$ in der komplexen Eben auf dem Einheitskreis und es gibt ein $\phi \in \mathbb{R}$ mit

$$z = |z|\, e^{\mathrm{i}\phi} = re^{\mathrm{i}\phi}.$$

Damit können wir komplexe Zahlen über einen Radius (Abstand zum Ursprung $(0, 0)$) und einen Winkel zur reellen Achse ϕ dargestellen. Dies zeigen wir einmal formal mit folgendem Satz.

Satz 4.70 (Polarkoordinaten komplexer Zahlen)

Für jede komplexe Zahl $z \in \mathbb{C}$ existiert ein $\phi \in \mathbb{R}$, sodass

$$z = |z|\, e^{\mathrm{i}\phi} = |z| \cos\phi + \mathrm{i}\,|z| \sin\phi.$$

Für $z \neq 0$ ist ϕ bis auf eine Addition mit Vielfachen von 2π eindeutig.

Das Paar $(|z|, \phi)$ bezeichnen wir als *Polarkoordinaten* von z und ϕ als *Argument* von z.

▼ Beweis

Für $z = 0$ gilt für jede Wahl von ϕ offensichtlich $z = 0 = |0|\, e^{\mathrm{i}\phi}$.

Für $z \neq 0$ betrachten wir zunächst den Fall $\mathrm{Im}(z) \geq 0$. Wir setzen $\hat{z} := \frac{z}{|z|} = \hat{a} + \mathrm{i}\hat{b}$ mit $\hat{a}, \hat{b} \in \mathbb{R}$. Dann ist $\left|\hat{z}\right|^2 = \hat{a}^2 + \hat{b}^2 = 1$ und $\hat{b} \geq 0$.

Wegen $\cos 0 = 1$ und $\cos\pi = -1$ und $a \in [-1, 1]$ gibt es nach dem Zwischenwertsatz ein $\pi \in [0, \pi]$ mit $\cos\phi = a$. Wir wählen $\phi = \mathrm{acos}(\hat{a})$.

Aus $\hat{a}^2 + \hat{b}^2 = 1$ folgt damit

$$\hat{b}^2 = 1 - \hat{a}^2 = 1 - \cos^2\phi = \sin^2\phi$$

und wegen $\hat{b} \geq 0$ gilt $\sin\phi = \hat{b}$.

Den Fall $\mathrm{Im}(z) < 0$ führen wir auf den ersten Fall zurück, indem wir $\bar{z}$ betrachten. Dort gilt wieder $\mathrm{Im}(z) > 0$ und wir finden wie oben beschrieben ein $\phi \in [0, \pi]$ mit

$$\bar{z} = |\bar{z}|\, e^{\mathrm{i}\phi} = |z|\, e^{\mathrm{i}\phi} \Leftrightarrow z = |\bar{z}|\, e^{\bar{\mathrm{i}}\phi} = |z|\, e^{-\mathrm{i}\phi}.$$

Es gibt also ein $-\phi \in [-\pi, 0]$, das die Voraussetzungen erfüllt und somit für beliebige z ein $\phi \in [-\pi, \pi]$.

Wir zeigen nun die Eindeutigkeit für $z \neq 0$ bis auf additive Vielfache von 2π: Sei $z = |z|\, e^{\mathrm{i}\psi}$ eine weitere Darstellung von z. Dann ist

$$1 = \frac{z}{z} = \frac{|z|\, e^{\mathrm{i}\phi}}{|z|\, e^{\mathrm{i}\psi}} = e^{\mathrm{i}(\phi-\psi)} = \cos(\phi-\psi) + \mathrm{i}\sin(\phi-\psi).$$

Daraus folgt $\sin(\phi-\psi) = 0$ und $\cos(\phi-\psi) = 1$. Aus Satz 4.69 folgt schließlich

$$\phi - \psi = 2\pi k, \quad k \in \mathbb{Z}.$$

□

Hieraus ergibt sich endlich ein simpler Beweis für unsere Vermutung, dass die Multiplikation komplexer Zahlen einer Drehung entspricht, wobei sich die Beträge multiplizieren und die Winkel addieren:

$$z_1 \cdot z_2 = |z_1|\, e^{\mathrm{i}\phi_1} \cdot |z_2|\, e^{\mathrm{i}\phi_2} = |z_1 \cdot z_2|\, e^{\mathrm{i}(\phi_1+\phi_2)}.$$

Ein weiteres Anwendungsbeispiel ist die deutliche Vereinfachung der Bestimmung komplexer Wurzeln und Potenzen.

Beispiel 4.71 (komplexe Wurzeln und Potenzen)

- Wir wollen $(\sqrt{2}-\sqrt{2}i)^6$ bestimmen, also z^6 für $z=\sqrt{2}-\sqrt{2}i$. Es ist

  $$|z| = |\sqrt{2}-\sqrt{2}i| = \sqrt{4} = 2$$

  und

  $$\arg(z) = -\arg(\bar{z}) = \phi = -\operatorname{acos}\frac{\sqrt{2}}{2} = -\frac{\pi}{4}.$$

  (Hierbei verwendet man häufig die Schreibweise $\arg(z)$ für die Funktion, die z auf den Winkel ϕ abbildet.)

  Also ist $z = 2e^{-i\pi/4}$ und

  $$z^6 = 2^6\left(e^{-i\pi/4}\right)^6 = 64e^{-i3\pi/2} = 64(\cos(-3\pi/2)+i\sin(-3\pi/2)) = 64i.$$

- Wir wollen eine Lösung u für $u^2 = w$ mit $w = 3+3\sqrt{3}i$ finden.

  Wir bestimmen erst die Polarkoordinaten von w:

  $$\begin{aligned} |w| &= |3+3\sqrt{3}i| = 6 \\ \arg(w) = \phi &= \operatorname{acos}\frac{1}{2} = \frac{\pi}{3} \\ w &= 6e^{i\pi/3}. \end{aligned}$$

  Nun können wir realtiv einfach eine Lösung von $u^2 = w$ bestimmen, indem wir aus $|w|$ die Wurzel ziehen und das Argument ϕ von w halbieren:

  $$u = \sqrt{6}e^{i\pi/6} = \sqrt{6}(\cos(\pi/6)+i\sin(\pi/6)) = \frac{\sqrt{18}}{2} + i\frac{\sqrt{6}}{2}.$$

  Auch beliebige n-te Wurzeln lassen sich auf diese Weise bestimmen.

5 Differentialrechnung

Nachdem wir bis zu diesem Punkt viel Vorarbeit geleistet haben, um Funktionen mathematisch genau zu beschreiben, können wir uns in diesem und dem folgenden Kapitel dem Herzen der Analysis zuwenden: der Differential- und Integralrechnung. Hierbei benötigen wir erneut Grenzwerte, denn in beiden Fällen spielen unendlich kleine Abstände eine entscheidende Rolle.

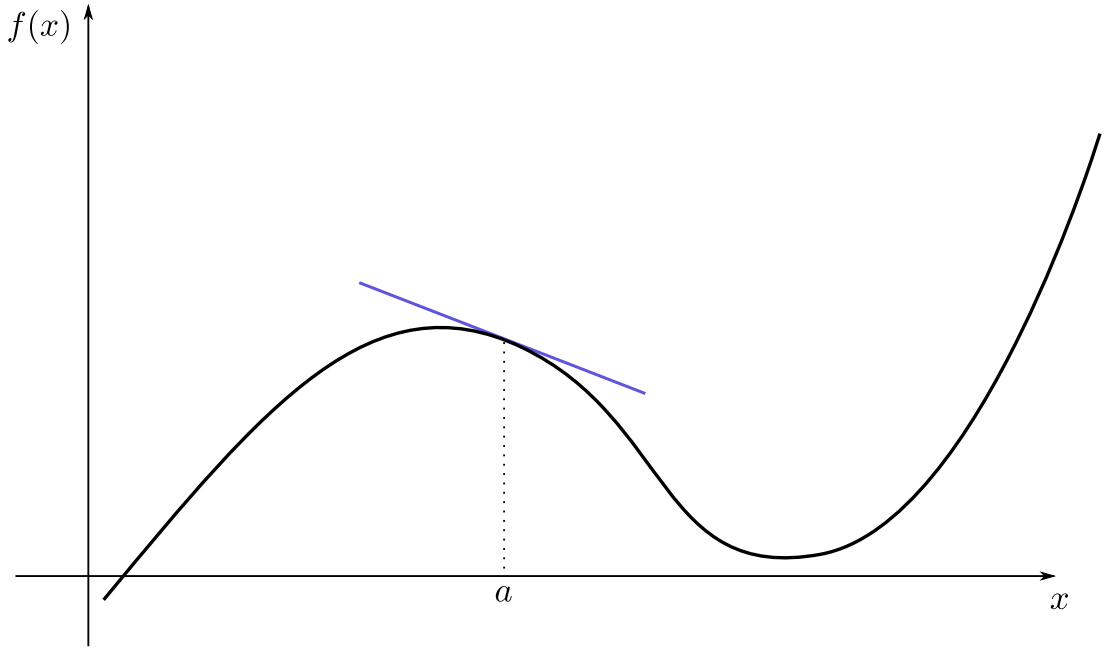
Die Differentialrechnung geht auf Leibniz und Newton zurück und ist die Basis für Differentialgleichungen, mit denen sich viele Vorgänge in der Natur und in technischen Systemen modellieren lassen. Darüber hinaus wird sie benutzt, um Funktionen zu optimieren, also einen Punkt x des Definitionsbereichs zu finden, an dem der Funktionswert $f(x)$ maximal groß oder klein wird. Solche Optimierungen kommen in allen Naturwissenschaften und auch in der Informatik sehr häufig vor, da es oft einfacher ist, eine Funktion zu definieren, deren Optimum eine bestimmte Eigenschaft hat, als das Optimum direkt anzugeben.

Wir betrachten in diesem Kapitel die Differentialrechnung mit einer Veränderlichen, wie sie auch in der Schule behandelt wird. Nach einer generellen Einführung leiten wir die Ableitungsregeln her, führen höhere Ableitungen ein und nutzen schließlich die Ableitungen, um diverse Eigenschaften von Funktionen zu analysieren.

5.1 Differenzierbarkeit

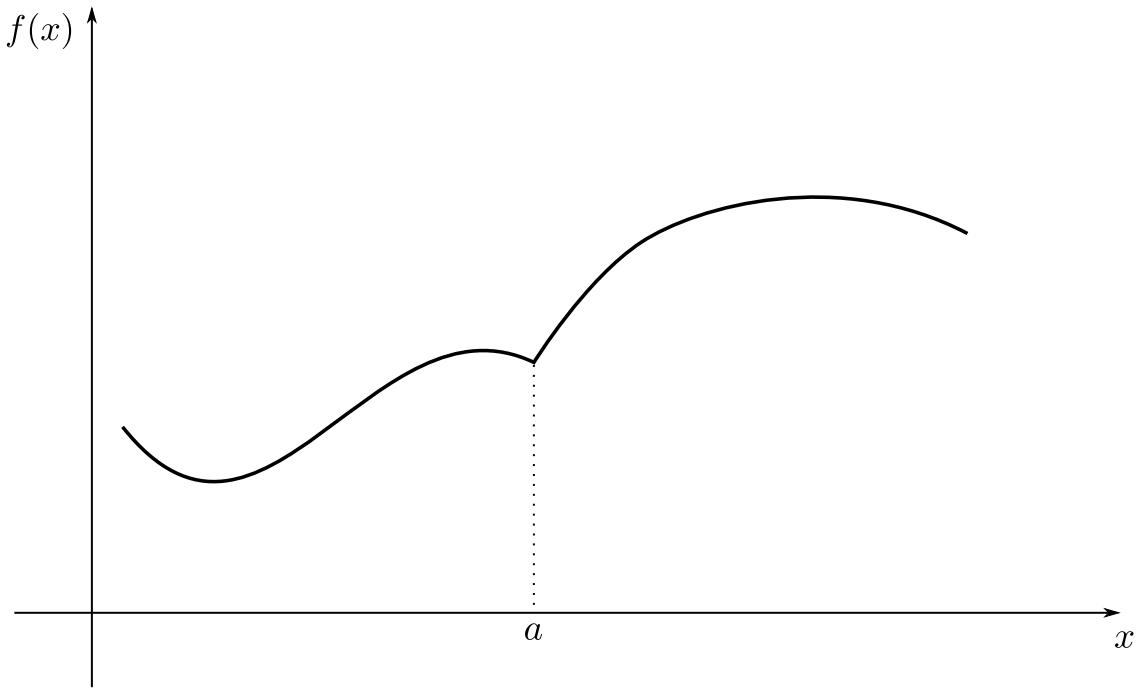
In der Differentialrechnung arbeiten wir mit sogenannten *Ableitungen* von Funktionen. Wir werden feststellen, dass die Stetigkeit einer Funktion nicht ausreicht, um eine Ableitung bestimmen zu können. Man könnte sagen, stetige Funktionen sind noch nicht “gutmütig” genug. Deswegen wurde der Begriff der *Differenzierbarkeit* eingeführt, der die Existenz einer Ableitung sicherstellt.

Wenn wir uns zurückerinnern, hatte die Stetigkeit einer Funktion salopp bedeutet, dass ähnliche x-Werte auch ähnliche Funktionswerte besitzen, die Funktion also nicht plötzlich springt. Dies erlaubt es uns, den Graphen einer Funktion durchgängig zu zeichnen. Kennen wir also einen Punkt einer stetigen Funktion, wissen wir auch, dass es weitere Punkte der Funktion in jeder noch so kleinen Umgebung des Punktes gibt. Die Differenzierbarkeit erweitert dieses Wissen, indem sie uns eine Aussage darüber erlaubt, in welcher “Richtung” die benachbarten Punkte auf der Funktion liegen. Diese Richtung kann man sich mithilfe einer Tangente an einem Funktionspunkt visualisieren:



Für $x = a$ können wir die Lage benachbarter Funktionswerte näherungsweise durch die blaue Tangente beschreiben.

Es gibt Funktionen, die zwar stetig sind, aber deren Graph einen "Knick" besitzt, hier ist es nicht möglich eine klare Richtung (in Form einer Tangente) für die Lage umliegender Punkte anzugeben, da es zwei verschiedene Möglichkeiten für die Richtung gibt. Genau diese "Knickfreiheit" ist mit der Differenzierbarkeit einer Funktion gemeint.



Für $x = a$ gibt es keine Tangente, welche die Lage benachbarter Funktionswerte (in beide Richtungen) gut beschreibt.

▶ Demo: Tangenten

Als Newton und Leibniz den Ableitungsbegriff entwickelten, hat man sich die Bestimmung einer Ableitung in etwa so vorgestellt: Man möchte wissen, wie sich die Funktionswerte $f(x)$ verändern, wenn wir x um ein winzig kleines Stück dx verändern. Für eine lineare Funktion, wie zum Beispiel $f(x) = 5x$ ergibt sich

$$\begin{aligned} f(x+dx) &= 5(x+dx) \\ &= 5x + 5 \cdot dx \\ &= f(x) + 5 \cdot dx. \end{aligned}$$

Das heißt, wenn wir den x-Wert um ein kleines Stück vergrößern, dann nimmt der Funktionswert um das Fünffache dieses Wertes zu, die *Änderungsrate* der Funktion ist also 5.

Für eine nichtlineare Funktion, wie zum Beispiel $f(x) = x^2$, ergibt sich mit der gleichen Überlegung

$$\begin{aligned} f(x+dx) &= x^2 + 2x \cdot dx + (dx)^2 \\ &= f(x) + 2x \cdot dx + (dx)^2 \\ &\approx f(x) + 2x \cdot dx. \end{aligned}$$

Die Idee der Begründer der "Infinitesimalrechnung" (wie die Differential- und Integralrechnung zusammenfassend genannt wird) war, dass dx so klein gewählt wird, dass höhere Potenzen von dx, wie zum Beispiel $(dx)^2$, vernachlässigbar klein sind und wir somit sagen können: Wenn wir x um ein kleines Stück dx vergrößern, dann vergrößert sich der Funktionswert um das $2x$-fache von dx. Die *Änderungsrate* der Funktion am Punkt x ist also $2x$. Das Vernachlässigen des $(dx)^2$-Terms löste in der Mathematik damals große Debatten aus, denn es wirkte so, als würde man nicht exakt arbeiten, sondern eher runden.

Sehen wir uns als letztes Beispiel noch $f(x) = x^n$ an:

$$\begin{aligned} f(x+dx) &= (x+dx)^n \\ &= \sum_{k=0}^{n} \binom{n}{k} x^{n-k} (dx)^k \\ &= x^n + nx^{n-1}dx + \frac{n(n+1)}{2} x^{n-2}(dx)^2 + \ldots + (dx)^n \\ &\approx f(x) + nx^{n-1}dx \end{aligned}$$

Hier ergibt sich also eine Veränderung, die ein (nx^{n-1})-faches von dx ist.

Was den Mathematiker*innen damals fehlte und weswegen diese Rechnungen in der damaligen Zeit recht "unmathematisch/ungenau" wirkten, war ein sauber definierter Grenzwertbegriff. Mit dem Vorwissen aus den vorigen Kapiteln fällt Ihnen vermutlich auf, dass sich die gewohnten Ableitungen ergeben, wenn wir für

$$\frac{f(x+dx) - f(x)}{dx}$$

in den Rechnungen zuvor der Grenzwert ($dx \to 0$) betrachten. Man verwendet den Ausdruck dx mittlerweile nicht mehr, sondern benutzt stattdessen normalerweise eine kleine Änderung h. Allerdings sieht man das dx noch in der Schreibweise der Ableitung $\frac{df(x)}{dx}$, auch wenn in der Schule häufiger f' geschrieben wird. Das d steht hierbei immer für eine unendlich kleine Differenz der Funktionswerte $(df(x))$ bzw. der Argumente (dx).

Definition 5.1 (Differenzierbarkeit, Ableitung)

Sei $a \in A \subseteq \mathbb{R}$ und sei $f : A \to \mathbb{R}$ eine Funktion. f heißt *in a differenzierbar*, falls der Grenzwert

$$\lim_{\substack{x \to a \\ x \in A \setminus \{a\}}} \frac{f(x) - f(a)}{x - a}$$

existiert. Wir schreiben für den Grenzwert $f'(a)$ oder $\frac{df}{dx}(a)$ oder $\frac{df(x)}{dx}\Big|_{x=a}$ und nennen ihn die *Ableitung* an der Stelle a.

Ersetzt man $x = a + h$, ergibt sich die alternative Schreibweise

$$f'(a) \;=\; \frac{df(x)}{dx}\Big|_{x=a} \;=\; \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

Eine Funktion heißt *differenzierbar*, wenn der Grenzwert für jedes $a \in A$ existiert. In diesem Fall bezeichnen wir $f'(x) = \frac{df}{dx}(x)$ als die *Ableitung* von f.

In der obigen Definition bedeutet $\lim\limits_{\substack{x \to a \\ x \in A \setminus \{a\}}}$, dass x gegen a geht, dabei aber nur Werte aus $A \setminus \{a\}$ annimmt, damit wir nicht durch 0 teilen.

Zur Erinnerung: Ein Grenzwert existiert nur dann, wenn er eine reelle Zahl ist. Ein Grenzwert von $f'(a) = \pm\infty$ bedeutet also, dass der Grenzwert nicht existiert.

Man kann die Grenzwertbildung in der Definition der Ableitung auch geometrisch interpretieren. Dabei entspricht $\frac{f(x)-f(a)}{x-a}$ der Steigung der Sekante des Graphen von f im Punkt $(a, f(a))$. Beim Grenzübergang geht die Sekante in die Tangente im Punkt a über, welche die Tangentengleichung $f(a) + f'(a)(x - a)$ hat. In den folgenden beiden Demos können Sie mit der graphischen Interpretation der Ableitung experimentieren.

▶ Demo: Ableitungen 1

▶ Demo: Ableitungen 2

Die Ableitung einer Funktion hat in Naturwissenschaften und Technik eine wichtige Bedeutung, da sie die momentane Änderungsrate einer Funktion angibt. Zum Beispiel gibt die Ableitung einer Funktion, welche den Ort eines Objektes in Abhängigkeit zur Zeit beschreibt, die Geschwindigkeit des Objektes an.

Beispiel 5.2 (Differenzierbarkeit)

Um die Differenzierbarkeit einer Funktion zu zeigen, nutzen wir unser Wissen über Stetigkeit: Für eine stetige Funktion können wir die Grenzwertbildung durch "Einsetzen" durchführen. Allerdings stört hierbei das h im Nenner, welches gegen 0 laufen soll. Wir versuchen also in der Regel, den Grenzwertquotienten so umzuformen, dass das h im Nenner gekürzt werden kann. Anschließend können wir für eine stetige Funktion $h = 0$ einfach einsetzen. Für die folgenden Beispiele sei $c \in \mathbb{R}$ eine beliebige Konstante.

1. Untersuche die Funktion $f : \mathbb{R} \to \mathbb{R},\ f(x) = c$

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lim_{h \to 0} \frac{0}{h} \\ &= \lim_{h \to 0} 0 \\ &= 0 \end{aligned}$$

2. Untersuche die Funktion $f : \mathbb{R} \to \mathbb{R},\ f(x) = cx$

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lim_{h \to 0} \frac{c(x+h) - cx}{h} \\ &= \lim_{h \to 0} \frac{ch}{h} \\ &= \lim_{h \to 0} c \\ &= c \end{aligned}$$

3. Untersuche die Funktion $f : \mathbb{R} \to \mathbb{R},\ f(x) = x^2$

$$\begin{aligned} f'(x) &= \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} \\ &= \lim_{h \to 0} \frac{(x+h)^2 - x^2}{h} \\ &= \lim_{h \to 0} \frac{2xh + h^2}{h} \\ &= \lim_{h \to 0} (2x + h) \\ &= 2x \end{aligned}$$

4. Untersuche die Funktion $f : \mathbb{R} \setminus \{0\} \to \mathbb{R},\ f(x) = \frac{1}{x}$

$$
\begin{aligned}
f'(x) &= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \\
&= \lim_{h\to 0} \left(\frac{1}{h}\cdot\left(\frac{1}{x+h}-\frac{1}{x}\right)\right) \\
&= \lim_{h\to 0} \frac{x-(x+h)}{hx(x+h)} \\
&= \lim_{h\to 0} \frac{-1}{x(x+h)} \\
&= -\frac{1}{x^2}
\end{aligned}
$$

5. Untersuche die Funktion $f : \mathbb{R} \to \mathbb{R},\ f(x) = \exp(x)$

$$
\begin{aligned}
f'(x) &= \lim_{h\to 0} \frac{\exp(x+h)-\exp(x)}{h} \\
&= \lim_{h\to 0} \frac{\exp(x)\cdot(\exp(h)-1)}{h} \\
&= \exp(x)\cdot \lim_{h\to 0} \frac{\exp(h)-1}{h} \\
&= \exp(x)\cdot 1 \\
&= \exp(x)
\end{aligned}
$$

Die vorletzte Umformung folgt aus Satz 4.47(g). Die Ableitung kann analog für die komplexe Exponentialfunktion bestimmt werden, so dass auch gilt $\exp'(z) = \exp(z)\ \forall z \in \mathbb{C}$.

Differenzierbarkeit stellt eine stärkere Forderung an eine Funktion dar als die Stetigkeit, denn aus Differenzierbarkeit folgt Stetigkeit, wie der folgende Satz zeigt.

Satz 5.3 (Aus Differenzierbarkeit folgt Stetigkeit)

Wenn eine Funktion $f : A \to \mathbb{R}$ in $a \in A$ differenzierbar ist, so ist sie in a auch stetig.

▼ Beweis

Sei (x_n) eine beliebige Folge mit $\lim\limits_{n\to\infty} x_n = a$ und $x_n \neq a \ \forall n \in \mathbb{N}$. Es gilt

$$\lim_{n\to\infty} (x_n - a) = 0.$$

Da f in a differenzierbar ist, existiert außerdem der Grenzwert

$$\lim_{n\to\infty} \frac{f(x_n) - f(a)}{x_n - a} = f'(a).$$

Nach Satz 3.18 dürfen wir beide Grenzwerte multiplizieren und den Limes vor das Produkt ziehen:

$$\begin{aligned} 0 &= 0 \cdot f'(a) \\ &= \left(\lim_{n\to\infty} (x_n - a)\right) \left(\lim_{n\to\infty} \frac{f(x_n) - f(a)}{x_n - a}\right) \\ &= \lim_{n\to\infty} \left((x_n - a)\frac{f(x_n) - f(a)}{x_n - a}\right) \\ &= \lim_{n\to\infty} f(x_n) - f(a) \end{aligned}$$

Daraus folgt $\lim\limits_{n\to\infty} f(x_n) = f(a)$, und damit ist $f(x)$ in a stetig.

□

Die Umkehrung von Satz 5.3 gilt *nicht*, wie das folgende Beispiel zeigt.

Beispiel 5.4

Die Funktion $f(x) = |x|$ ist stetig in jedem $a \in \mathbb{R}$, da $\lim\limits_{x \to a} |x| = |a|$ (wurde auf Übungsblatt 6 gezeigt).

Betrachten wir allerdings die Folge $x_n = (-1)^n/n$, so gilt $\lim\limits_{n \to \infty} x_n = 0$, aber der Grenzwert

$$\lim_{x \to 0} \frac{|x| - |0|}{x - 0} = \lim_{n \to \infty} \frac{|x_n|}{x} = \lim_{n \to \infty} \frac{\frac{1}{n}}{(-1)^n \frac{1}{n}} = (-1)^n$$

existiert nicht. Damit ist $|x|$ im Punkt 0 nicht differenzierbar.

Für Funktionen, die aus zwei Teilstücken bestehen, eignet sich analog zu Satz 4.20 der folgende Satz, bei dem wir die Differenzierbarkeit auf rechts- und linksseitige Grenzwerte zurückführen.

Satz 5.5

Eine Funktion $f : [a, b] \to \mathbb{R}$ ist genau dann für ein $c \in (a, b)$ differenzierbar, wenn die links- und rechtsseitigen Grenzwerte

$$f'_-(c) = \lim_{x \nearrow c} \frac{f(x) - f(c)}{x - c} \quad \text{und} \quad f'_+(c) = \lim_{x \searrow c} \frac{f(x) - f(c)}{x - c}$$

existieren und übereinstimmen. In diesem Fall gilt

$$f'(c) \;=\; f'_-(c) \;=\; f'_+(c).$$

▼ Beweis

Beweis in zwei Richtungen.

$\Rightarrow$-Richtung:

Wenn f differenzierbar in c ist, dann gilt für jede Folge (x_n), die gegen c konvergiert

$$\lim_{n \to \infty} \frac{f(x_n) - f(c)}{x_n - c} = f'(c).$$

Insbesondere gilt dies auch für Folgen mit $x_n > c \; \forall n \in \mathbb{N}$ und für Folgen mit $x_n < c \; \forall n \in \mathbb{N}$, also

$$\lim_{x \nearrow c} \frac{f(x) - f(c)}{x - c} = f'(c) = \lim_{x \searrow c} \frac{f(x) - f(c)}{x - c}.$$

$\Leftarrow$-Richtung:

Die rechts- und linksseitigen Grenzwerte existieren und stimmen überein. Sei (x_n) eine beliebige Folge mit $x_n \neq c \; \forall n \in \mathbb{N}$, die gegen c konvergiert. Wir müssen nun zeigen, dass der Differenzenquotient für die Folge (x_n) auch gegen den links- und rechtsseitigen Grenzwert konvergiert.

Wir teilen dazu die Folgenglieder von (x_n) auf die zwei Mengen $A = \{x_n \mid x_n > c\}$ und $B = \{x_n \mid x_n < c\}$ auf. Wenn A bzw. B nur endlich viele Folgenglieder enthält, setzen wir n_a bzw. n_b auf den höchsten Folgenindex, der in A bzw. B enthalten ist. Anderenfalls bilden die Elemente von A bzw. B eine Teilfolge (a_n) bzw. (b_n) von (x_n), die nach Satz 3.26 ebenfalls gegen c konvergiert, und es gilt $a_n > c$ bzw. $b_n < c$.

Da der rechtssseitige Grenzwert existiert, gibt es für jedes $\varepsilon > 0$ ein $n_a \in \mathbb{N}$, sodass für alle $n \geq n_a$ gilt

$$\left| \frac{f(a_n) - f(c)}{a_n - c} - f'_+(c) \right| < \varepsilon.$$

Da der linksseitige Grenzwert existiert, gibt es für jedes $\varepsilon > 0$ ein $n_b \in \mathbb{N}$, sodass für alle $n \geq n_b$ gilt

$$\left| \frac{f(b_n) - f(c)}{b_n - c} - f'_-(c) \right| < \varepsilon.$$

Da die beiden Grenzwerte übereinstimmen, können wir für ein beliebiges $\varepsilon > 0$ ein $n_0 = \max\{n_a, n_b\}$ wählen, und es gilt für alle $n \geq n_0$

$$\left| \frac{f(x_n) - f(c)}{x_n - c} - f'_-(c) \right| = \left| \frac{f(x_n) - f(c)}{x_n - c} - f'_+(c) \right| < \varepsilon.$$

Da dies für beliebige Folgen (x_n) gilt, folgt

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} = f'_+(c) = f'_-(c).$$

□

▶ Demo: Links- und rechtsseitiger Grenzwert des Differenzenquotienten

Früher dachte man, dass die nicht differenzierbaren Stellen einer Funktion immer nur einzelne Punkte betreffen können, eine stetige Funktion aber dazwischen immer stückweise differenzierbar ist. Einen Gegenbeweis liefern die sogenannten Weierstraß-Funktionen, die in jedem Punkt stetig, aber in keinem Punkt differenzierbar sind. Ein Beipiel wäre die Funktion

$$f(x) = \sum_{k=0}^{\infty} a^k \cos(b^k \pi x)$$

mit $0 < a < 1$ und $ab \geq 1$. Graphisch haben solche Funktionen eine fraktale Struktur. Wenn wir also näher an den Graphen heranzoomen, zeigt sich wieder der gleiche Verlauf der Funktion, wie auf einer größeren Skala, wie das folgende Bild zeigt:

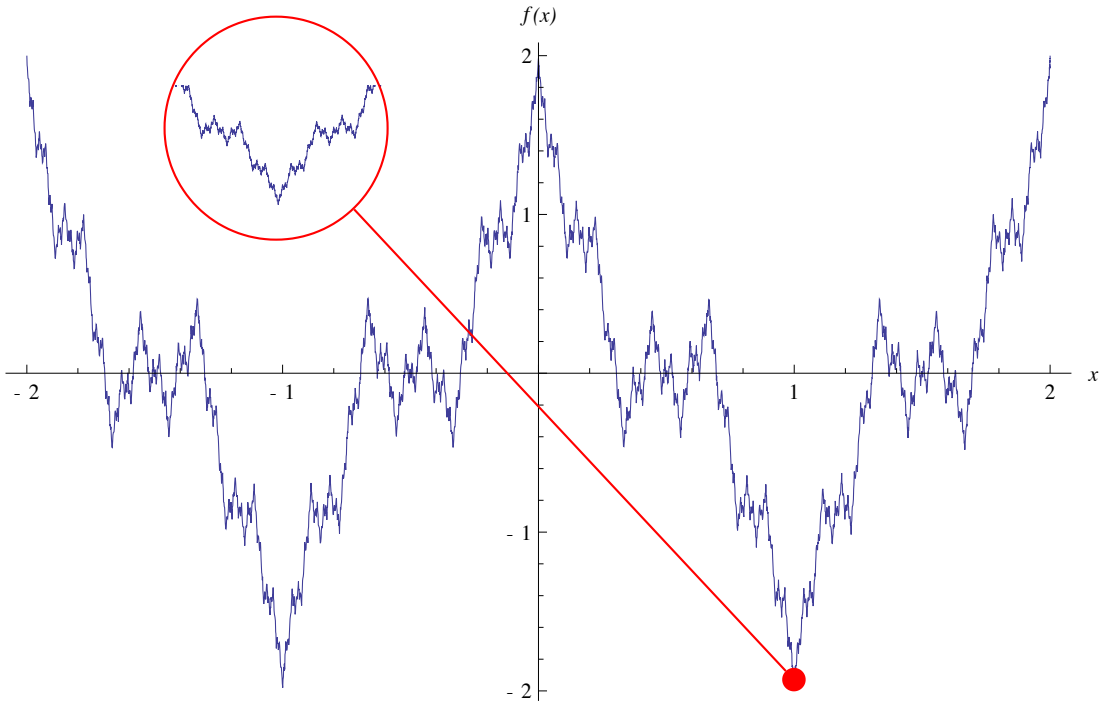


Bild von Eeyore22

Zur Berechnung der Ableitung einer Funktion ist die Definition 5.1 nur in einfachen Fällen praktikabel. Für komplexere Funktionen kann die damit verbundene Grenzwertberechnung schnell sehr aufwendig werden. Daher leiten wir nun ein paar Ableitungsregeln her, mit denen wir viele Ableitungen auf schon bekannte elementare Fälle zurückführen können.

Satz 5.6 (Ableitungsregeln)

Seien $f, g : A \to \mathbb{R}$ in $a \in A$ differenzierbar und $c \in \mathbb{R}$. Dann sind auch die Funktionen $f + g$, $c \cdot f$ und $f \cdot g$ in a differenzierbar und es gelten folgende Rechenregeln:

a. Linearität:

$$(f+g)'(a) = f'(a) + g'(a)$$
$$(c \cdot f)'(a) = c \cdot f'(a)$$

b. Produktregel:

$$(f \cdot g)'(a) = f'(a) \cdot g(a) + f(a) \cdot g'(a)$$

Ist außerdem $g(x) \neq 0$ für alle $x \in A$, so ist auch die Funktion $\frac{f}{g}$ differenzierbar, und es gilt:

c. Quotientenregel:

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a) \cdot g(a) - f(a) \cdot g'(a)}{g(a)^2}$$

▼ Beweis

a. Folgt direkt aus der Definition des Differenzenquotienten und den Rechenregeln für Grenzwerte.

b. Es gilt

$$\begin{aligned}
(f \cdot g)'(a) &= \lim_{h \to 0} \frac{f(a+h)g(a+h) - f(a)g(a)}{h} \\
&= \lim_{h \to 0} \left(\frac{1}{h} \cdot \Big[f(a+h)\left(g(a+h) - g(a)\right) + \left(f(a+h) - f(a)\right) \cdot g(a) \Big] \right) \\
&= \lim_{h \to 0} \left(f(a+h) \frac{g(a+h) - g(a)}{h} \right) + \lim_{h \to 0} \left(\frac{f(a+h) - f(a)}{h} g(a) \right) \\
&= f(a) \cdot g'(a) + f'(a) \cdot g(a)
\end{aligned}$$

c. Wir beweisen zuerst den Spezialfall $f(x) = 1$.

$$\begin{aligned}
\left(\frac{1}{g}\right)'(a) &= \lim_{h \to 0} \left(\frac{1}{h} \left(\frac{1}{g(a+h)} - \frac{1}{g(a)} \right) \right) \\
&= \lim_{h \to 0} \left(\frac{1}{g(a+h)g(a)} \left(\frac{g(a) - g(a+h)}{h} \right) \right) \\
&= -\frac{g'(a)}{g(a)^2}.
\end{aligned}$$

Allgemein gilt dann

$$\left(\frac{f}{g}\right)'(a) = \left(f\frac{1}{g}\right)'(a)$$
$$\overset{\text{nach (b)}}{=} f'(a)\frac{1}{g(a)} + f(a)\frac{-g'(a)}{g(a)^2}$$
$$= \frac{f'(a)\cdot g(a) - f(a)\cdot g'(a)}{g(a)^2}$$

□

Beispiel 5.7

1. $f_n(x) = x^n$ für $n \in \mathbb{N}$ ist differenzierbar und hat die Ableitung $f_n'(x) = nx^{n-1}$.

   Beweis per vollständiger Induktion über n:

   Induktionsanfang: Für $n = 1$ und $n = 2$ wurde die Behauptung im Beispiel 5.2 gezeigt. Damit ist der Induktionsanfang bewiesen.

   Induktionsvoraussetzung: Für ein beliebiges $n \geq 2$ gelte $f_{n-1}'(x) = (n-1)x^{n-2}$.

   Induktionsschritt $(n-1) \to n$**:**

   $$f_n(x) = f_1(x)f_{n-1}(x)$$
   $$f_n'(x) = f_1'(x)f_{n-1}(x) + f_1(x)f_{n-1}'(x)$$
   $$= x^{n-1} + x(n-1)x^{n-2}$$
   $$= nx^{n-1}$$

   Damit gilt die Behauptung für alle $n \in \mathbb{N}$.

2. Ein Polynom $f(x) = \sum_{k=0}^{n} a_k x^k$ mit $n \in \mathbb{N}_0$ und $a_k \in \mathbb{R}$ ist differenzierbar in $x \in \mathbb{R}$ und es gilt:

   $$f'(x) = \begin{cases} \sum_{k=1}^{n} k\, a_k\, x^{k-1} & \text{für } n \geq 1, \\ 0 & \text{für } n = 0. \end{cases}$$

   Der Beweis folgt aus dem ersten Beispiel und der Linearität.

3. Die Funktion $f(x) = \frac{1}{x^n}$ mit $n \in \mathbb{N}$ ist differenzierbar in $x \in \mathbb{R} \setminus \{0\}$ und es gilt nach Quotientenregel:

   $$f'(x) = \frac{-(nx^{n-1})}{(x^n)^2} = -nx^{-n-1}.$$

Die folgenden beiden Sätze führen Regeln zur Bestimmung der Ableitung von Umkehrfunktionen und verketteten Funktionen ein.

Satz 5.8 (Ableitung der Umkehrfunktion)

Sei $I \subset \mathbb{R}$ ein Intervall, das aus mehr als einem Punkt besteht, $f : I \to \mathbb{R}$ eine stetige streng monotone Funktion, und $g = f^{-1} : J \to \mathbb{R}$ mit $J = f(I)$ die Umkehrfunktion von f.

Wenn f in $a \in I$ differenzierbar ist und es gilt $f'(a) \neq 0$, dann ist g in $b = f(a)$ differenzierbar und es gilt:

$$g'(b) = \frac{1}{f'(a)} = \frac{1}{f'(g(b))}.$$

▼ Beweis

Sei $b_n \in J \setminus \{b\}$ eine Folge mit $\lim_{n\to\infty} b_n = b$. Wir setzen $a_n = g(b_n)$. a_n existiert, da $b_n \in f(I)$.

Da g stetig ist (siehe Satz 4.34), ist $\lim_{n\to\infty} a_n = a$ und da $g : J \to I$ bijektiv ist, ist $a_n \neq a$ für $b_n \neq b$. Es gilt

$$\lim_{n\to\infty} \frac{g(b_n) - g(b)}{b_n - b} = \lim_{n\to\infty} \frac{a_n - a}{f(a_n) - f(a)} = \frac{1}{f'(a)}.$$

Der Grenzwert existiert, da die Grenzwerte im Zähler und Nenner existieren.

□

Beispiel 5.9

1. Ableitung des Logarithmus $\ln : \mathbb{R}_{>0} \to \mathbb{R}$:

   $\ln(x)$ ist die Umkehrfunktion von $\exp(x)$. Damit gilt nach Satz 5.8 mit $f(x) = \exp(x)$ und $g(x) = \ln(x)$:

   $$\ln'(x) = \frac{1}{\exp'(\ln(x))} = \frac{1}{\exp(\ln(x))} = \frac{1}{x}.$$

2. Ableitung der Wurzelfunktion $\sqrt{x} : \mathbb{R}_{>0} \to \mathbb{R}_{>0}$:

   $g(x) := \sqrt{x}$ ist die Umkehrfunktion von $f(x) := x^2$. Es gilt $f'(x) = 2x$ und für $x > 0$ gilt $f'(x) > 0$. Damit folgt nach Satz 5.8:

   $$g'(x) = \frac{1}{f'(g(x))} = \frac{1}{2g(x)} = \frac{1}{2\sqrt{x}}.$$

   Achtung: $g(x) = \sqrt{x}$ wäre für $x = 0$ nicht differenzierbar.

Satz 5.10 (Kettenregel)

Seien $f : A \subseteq \mathbb{R} \to \mathbb{R}$ und $g : B \subseteq \mathbb{R} \to \mathbb{R}$ Funktionen. Sei f in $a \in A$ differenzierbar und g in $b = f(a)$ differenzierbar. Dann ist die Komposition (Verkettung) der beiden Funktionen $g \circ f : A \to \mathbb{R}$ in a differenzierbar und es gilt

$$(g \circ f)'(a) \;=\; g'\left(f(a)\right) \cdot f'(a).$$

▼ Beweis

Definiere die Funktion $g^* : B \to \mathbb{R}$ durch

$$g^*(y) := \begin{cases} \frac{g(y)-g(b)}{y-b}, & \text{für } y \neq b \\ g'(b), & \text{für } y = b \end{cases}.$$

Da g in b differenzierbar ist, gilt

$$\lim_{y \to b} g^*(y) = g^*(b) = g'(b)$$

sowie

$$g(y) - g(b) = g^*(y) \cdot (y - b).$$

Daraus folgt

$$\begin{aligned} (g \circ f)'(a) &= \lim_{x \to a} \frac{g(f(x)) - g(f(a))}{x - a} \\ &= \lim_{x \to a} \frac{g^*(f(x)) \cdot (f(x) - f(a))}{x - a} \\ &= \lim_{x \to a} g^*(f(x)) \cdot \lim_{x \to a} \frac{f(x) - f(a)}{x - a} \\ &= g'(f(a)) \cdot f'(a) \end{aligned}$$

□

Beispiel 5.11 (Kettenregel)

1. Ableitung von $\sqrt{x^2+2}$:

$$\begin{aligned} f(x) &= x^2+2, & f'(x) &= 2x, \\ g(y) &= \sqrt{y}, & g'(y) &= \frac{1}{2\sqrt{y}}, \\ h(x) &= g(f(x)), & h'(x) &= \frac{1}{2\sqrt{x^2+2}} \cdot 2x = \frac{x}{\sqrt{x^2+2}}. \end{aligned}$$

2. Ableitung von e^{cx} für ein $c \in \mathbb{R}$:

$$\begin{aligned} f(x) &= cx, & f'(x) &= c \\ g(y) &= e^y, & g'(y) &= e^y \\ h(x) &= g(f(x)), & h'(x) &= e^{cx} \cdot c = ce^{cx}. \end{aligned}$$

   Der Beweis funktioniert analog für die komplexe Exponentialfunktion, also mit $x, c \in \mathbb{C}$.

Mit dem letzten Beispiel ergeben sich insbesondere die Ableitungen der trigonometrischen Funktionen.

Beispiel 5.12 (Ableitungen trigonometrischer Funktionen)

Nach Beispiel 5.11 gilt $(\exp(cx))' = c\exp(cx)$. Damit folgt:

$$\begin{aligned} (\cos(x))' &= \frac{(\exp(\mathrm{i}x))' + (\exp(-\mathrm{i}x))'}{2} \\ &= \frac{\mathrm{i}\exp(\mathrm{i}x) - \mathrm{i}\exp(-\mathrm{i}x)}{2} \\ &= \frac{\mathrm{i}\exp(\mathrm{i}x) - \mathrm{i}\exp(-\mathrm{i}x)}{2} \cdot \frac{\mathrm{i}}{\mathrm{i}} \\ &= \frac{-\exp(\mathrm{i}x) + \exp(-\mathrm{i}x)}{2\mathrm{i}} \\ &= -\sin(x) \end{aligned}$$

$$\begin{aligned} (\sin(x))' &= \frac{(\exp(\mathrm{i}x))' - (\exp(-\mathrm{i}x))'}{2\mathrm{i}} \\ &= \frac{\mathrm{i}\exp(\mathrm{i}x) + \mathrm{i}\exp(-\mathrm{i}x)}{2\mathrm{i}} \\ &= \frac{\exp(\mathrm{i}x) + \exp(-\mathrm{i}x)}{2} \\ &= \cos(x) \end{aligned}$$

Die Idee der Ableitung einer Funktion kann rekursiv angewendet werden, indem man die Ableitung der Ableitung bildet, sofern diese existiert. Dies führt zu folgender Definition höherer Ableitungen.

Definition 5.13 (Ableitungen höherer Ordnung)

Für eine Funktion $f : A \subseteq \mathbb{R} \to \mathbb{R}$ und $k \in \mathbb{N}_0$ ist die *k-te Ableitung* (oder auch die *Ableitung k-ter Ordnung*) $f^{(k)}(a)$ von f in $a \in A$ rekursiv definiert als

$$\begin{aligned} f^{(0)}(a) &:= f(a), \\ f^{(k+1)}(a) &:= \left(f^{(k)}(a)\right)', \text{ falls die Ableitung von } f^{(k)}(a) \text{ in } a \in A \text{ existiert.} \end{aligned}$$

Wenn für eine Funktion f die k-te Ableitung für alle $a \in A$ existiert, ist f k-mal differenzierbar.

Wenn die k-te Ableitung zusätzlich stetig ist, ist f k-mal stetig differenzierbar. Dann schreibt man auch: f ist C^k-stetig.

Man schreibt auch $\dfrac{d^k f(a)}{dx^k}$ für die k-te Ableitung von f im Punkt a.

Beispiel 5.14

1. Die Funktion $f(x) = e^{cx}$ ist unendlich oft stetig differenzierbar. Also ist sie C^∞-stetig und es gilt:
$$f^{(k)}(x) = c^k e^{cx}.$$

2. Die Funktion $f(x) = x^n$ für $n \in \mathbb{N}$ ist unendlich oft stetig differenzierbar. Also ist sie C^∞-stetig und es gilt:
$$f^{(k)}(x) = \begin{cases} \frac{n!}{(n-k)!} x^{(n-k)}, & \text{für } 0 \leq k \leq n, \\ 0, & \text{für } k > n. \end{cases}$$

3. Die Funktion $f(x) = |x|$ ist stetig, aber nicht differenzierbar, also C^0-stetig.

4. Die Funktion $f(x) = |x|^3$ ist zweimal stetig-differenzierbar, also C^2-stetig (Beweis und Ableitungen zur Übung).

5. Die Funktion
$$f(x) = \begin{cases} 0 & \text{für } x = 0, \\ x^2 \sin\left(\frac{1}{x}\right) & \text{sonst.} \end{cases}$$
ist stetig und differenzierbar, aber die erste Ableitung ist nicht stetig in $x = 0$. Also ist f nur C^0-stetig (Beweis und Ableitung zur Übung).

5.2 Lokale Extrema

Wir wollen nun mit Hilfe der Ableitungen von Funktionen deren Eigenschaften analysieren. Dazu wird eine Reihe von Sätzen hergeleitet, die das Verhalten von Funktionen auf Intervallen beschreiben. Zuerst definieren wir den Begriff des *lokalen Extremums*.

Definition 5.15 (Lokale Extrema)

Eine Funktion $f : (a, b) \to \mathbb{R}$ nimmt in $x \in (a, b)$ ein *lokales Maximum* (bzw. *lokales Minimum*) $f(x)$ an, wenn ein $\varepsilon > 0$ existiert, sodass $f(x) \geq f(y)$ (bzw. $f(x) \leq f(y)$) für alle y mit $|x - y| < \varepsilon$ gilt.

Extremum ist der Oberbegriff für Minimum und Maximum. Wir nennen diese Extrema *lokal*, weil z.B. ein lokales Maximum nur in einer gewissen Umgebung das Maximum aller in dieser Umgebung liegenden Funktionswerte ist. An einer anderen Stelle könnte die Funktion noch ein weiteres, größeres Maximum besitzen, oder sogar nach oben unbeschränkt sein. Wir werden Bedingungen für lokale Extrema herleiten, da diese auf Basis der Ableitung, die das Verhalten einer Funktion in der Umgebung eines Punktes beschreibt, bestimmt werden können. Das Maximum/Minimum *aller* Funktionswerte einer Funktion nennt man *globales Minimum/Maximum*. Der x-Wert, in dem das Extremum angenommen wird, bezeichnet man auch als *Extremstelle*. Die Bestimmung globaler Extrema ist nur in bestimmten Funktionsklassen möglich und geht über den Inhalt dieser Vorlesung hinaus.

Genau genommen ist für konstante Funktionen $f(x) = c$, $c \in \mathbb{R}$, jeder Punkt ein lokales Minimum *und* ein lokales Maximum. In manchen Lehrbüchern wird daher zusätzlich der Begriff des *strengen* lokalen Extremums eingeführt, bei dem man in der obigen Definition $\leq$ durch $<$ bzw. $\geq$ durch $>$ ersetzt. Wir werden in den Beweisen darauf hinweisen, aber ansonsten immer nur von lokalen Extrema sprechen.

Satz 5.16 (Notwendige Bedingung für lokale Extrema)

Wenn die Funktion $f : (a, b) \to \mathbb{R}$ im Punkt $x \in (a, b)$ ein lokales Extremum annimmt und in x differenzierbar ist, dann ist $f'(x) = 0$.

▼ Beweis

Wenn die Funktion f in x ein lokales Maximum annimmt, dann existiert $\varepsilon > 0$, sodass $f(y) \leq f(x)$ für alle $y \in (x - \varepsilon, x + \varepsilon)$.

Für $x < y$ gilt dann

$$\frac{f(y) - f(x)}{y - x} \leq 0$$

und damit

$$f'(x) = \lim_{y \searrow x} \frac{f(y) - f(x)}{y - x} \leq 0.$$

Für $x > y$ gilt analog

$$\frac{f(y) - f(x)}{y - x} \geq 0$$

und damit

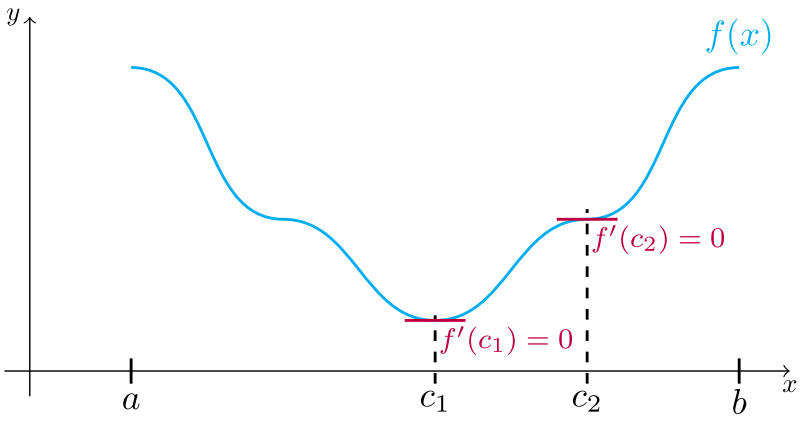
$$f'(x) = \lim_{y \nearrow x} \frac{f(y) - f(x)}{y - x} \geq 0.$$

Aus $0 \leq f'(x) \leq 0$ folgt insgesamt $f'(x) = 0$.

Der Beweis für lokale Minima ist analog zu führen.

□

Die Umkehrung des letzten Satzes gilt nicht. An einer Stelle c mit $f'(c) = 0$ muss also nicht zwangsweise ein lokales Extremum liegen (Mathematiker*innen sagen, die Bedingung ist nicht hinreichend). Die folgende Abbildung zeigt die graphische Interpretation des Satzes: An einem lokalen Extremum (in der Abbildung bei c_1) hat die Tangente eine Steigung von 0, verläuft also parallel zur x-Achse. Allerdings erfüllt die gezeigte Funktion auch an einer anderen Stelle c_2 die Bedingung $f'(c_2) = 0$, an der jedoch kein Extremum vorliegt.



Im letzten Satz ist zu beachten, dass wir die Funktion f über dem offenen Intervall (a, b) untersucht haben. Wenn wir $f : [a, b] \to \mathbb{R}$ untersuchen und ein Minimum oder Maximum in einem der beiden Randpunkte angenommen wird, so ist dort die erste Ableitung nicht notwendigerweise gleich 0, wie man sich leicht überlegen kann.

Die nun folgenden Sätze beschäftigen sich mit dem Verhalten von Funktionen in Intervallen, wenn Informationen über die Randpunkte vorliegen. Wir betrachten dabei oft Funktionen, die auf dem abgeschlossenen Intervall $[a, b]$ definiert sind und im offenen Intervall (a, b) differenzierbar sind. Damit ist $f'(x)$ nur für $x \in (a, b)$ definiert.

Satz 5.17 (Satz von Rolle)

Sei $a < b$ und $f : [a, b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) = f(b)$, die in (a, b) differenzierbar ist. Dann gilt:

$$\exists c \in (a, b) : f'(c) = 0$$

▼ Beweis

Nach Satz 4.36 nimmt f als stetige Funktion auf dem kompakten Intervall $[a, b]$ ihr Minimum und Maximum an.

Falls f konstant ist, gilt $f'(c) = 0$ für alle $c \in (a, b)$.

Ist f nicht konstant, so gibt es ein $x \in (a, b)$ mit $f(x) < f(a)$ oder $f(x) > f(a)$. Damit kann entweder das Minimum oder das Maximum nicht an den Intervallgrenzen liegen. Es gibt also ein Extremum in einem Punkt $c \in (a, b)$, und nach der notwendigen Bedingung für Extrema gilt dort $f'(c) = 0$.

□

Satz 5.18 (Erster Mittelwertsatz der Differentialrechnung)

Sei $a < b$ und $f : [a, b] \to \mathbb{R}$ eine stetige Funktion, die in (a, b) differenzierbar ist. Dann gilt:

$$\exists c \in (a, b) : \frac{f(b) - f(a)}{b - a} = f'(c)$$

▼ Beweis

Wir definieren eine Funktion $g : [a, b] \to \mathbb{R}$, mit

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Die Funktion g ist stetig in $[a, b]$, differenzierbar in (a, b) und $g(b) = f(a) = g(a)$.

Dann existiert nach dem Satz von Rolle ein $c \in (a, b)$ mit $g'(c) = 0$. Aus

$$g'(c) = f'(c) - \frac{f(b) - f(a)}{b - a} = 0$$

folgt die Behauptung.

□

Beide Sätze kann man geometrisch interpretieren. Die Steigung der Sekante durch die Punkte $(a, f(a))$ und $(b, f(b))$ entspricht der Steigung der Tangente im Punkt $(c, f(c))$, wie die folgende Demo zeigt.

▶ Demo: Erster Mittelwertsatz der Differentialrechnung

Mit Hilfe der Ableitungen lassen sich auch Aussagen über die Monotonie von Funktionen treffen.

Satz 5.19 (Zusammenhang zwischen Monotonie und Ableitung)

Sei $f : [a, b] \to \mathbb{R}$ stetig und in (a, b) differenzierbar. Dann gilt:

a. $\forall x \in (a, b) : f'(x) \geq 0 \Leftrightarrow f$ monoton wachsend in $[a, b]$

b. $\forall x \in (a, b) : f'(x) > 0 \Rightarrow f$ streng monoton wachsend in $[a, b]$

c. $\forall x \in (a, b) : f'(x) \leq 0 \Leftrightarrow f$ monoton fallend in $[a, b]$

d. $\forall x \in (a, b) : f'(x) < 0 \Rightarrow f$ streng monoton fallend in $[a, b]$

▼ Beweis

a. **⇒-Richtung:**
Wähle beliebige $c, d \in [a, b]$ mit $c < d$. Der Mittelwertsatz (Satz 5.18) angewendet auf das Intervall $[c, d]$ liefert ein $y \in (c, d)$ mit $f'(y) = \frac{f(d)-f(c)}{d-c}$. Da $f'(x) \geq 0$ für alle $x \in (a, b)$, ist auch f'(y) $, daher ist auch $\frac{f(d)-f(c)}{d-c} \geq 0$. Da $c < d$ muss auch $f(c) < f(d)$ sein. Also ist die Funktion monoton wachsend.

⇐-Richtung:
Nehmen wir nun an, dass f monoton wachsend in $[a, b]$ ist. Dann ist für alle $x, y \in (a, b)$ mit $x \neq y$ der Differenzenquotient $\frac{f(x)-f(y)}{x-y} \geq 0$. Daraus folgt durch den Grenzübergang die Behauptung:

$$\lim_{x \to y} \frac{f(x) - f(y)}{x - y} = f'(y) \geq 0$$

Die Beweise für **(b)**–**(d)** sind analog zu führen, wobei zu beachten ist, dass in den Fällen **(b)** und **(d)** nur eine Richtung bewiesen werden muss und kann.

□

Es sollte beachtet werden, dass in dem vorherigen Satz nur die beiden Aussagen **(a)** und **(c)** genau-dann-wenn-Beziehungen sind, d.h. dass die Folgerungen in beide Richtungen gelten. Für die beiden Punkte **(b)** und **(d)** gilt nur die Folgerung von links nach rechts. Dies kann man sich leicht anhand der Funktion $f(x) = x^3$ überlegen: Trotz $f'(0) = 0$ ist die Funktion streng monoton wachsend.

Aus den vorherigen Sätzen können wir nun eine hinreichende Bedingung für strenge lokale Extrema herleiten.

Satz 5.20 (Hinreichende Bedingung für lokale Extrema)

Sei $f : (a, b) \to \mathbb{R}$ eine differenzierbare Funktion, die im Punkt $x \in (a, b)$ zweimal differenzierbar ist.

Falls $f'(x) = 0$ und $f''(x) > 0$ (bzw. $f''(x) < 0$), dann nimmt f in x ein lokales Minimum (bzw. Maximum) an.

▼ Beweis

Wir beweisen den Satz für lokale Minima. Sei also $f''(x) > 0$.

Da $f''(x) = \lim\limits_{y \to x} \frac{f'(y) - f'(x)}{y - x} > 0$ ist, existiert ein $\varepsilon > 0$, sodass $\frac{f'(y) - f'(x)}{y - x} > 0$ für alle $y \in (x - \varepsilon, x + \varepsilon)$.

Da außerdem $f'(x) = 0$ folgt $f'(y) < 0$ für $y \in (x - \varepsilon, x)$ und $f'(y) > 0$ für $y \in (x, x + \varepsilon)$.

Nach Satz 5.19 ist f damit streng monoton fallend in $(x - \varepsilon, x)$ und streng monoton wachsend in $(x, x + \varepsilon)$. Damit muss f in x ein (sogar strenges) lokales Minimum besitzen.

Der Beweis für lokale Maxima ist analog zu führen.

□

Die Bedingungen sind nur hinreichend, aber nicht notwendig: Wenn die Bedingung gilt, gibt es ein lokales Extremum, aber nicht für jedes lokale Extremum gilt diese Bedingung. Dies kann man an der Funktion $f(x) = x^4$ erkennen: Diese besitzt in $x = 0$ ein lokales Minimum, trotzdem ist $f''(0) = 0$. Für solche Fälle, in denen $f'(x) = f''(x) = 0$, aber $f^{(n)}(x) \neq 0$ für $n > 2$ gilt, kann der folgende Satz genutzt werden, den wir hier ohne Beweis angeben, da der Beweis über den Vorlesungsinhalt hinaus geht.

Satz 5.21

Sei $f : (a, b) \to \mathbb{R}$ eine differenzierbare Funktion, die im Punkt $x \in (a, b)$ $n+1$-mal differenzierbar ist. Falls

$$f'(x) = f^{(2)}(x) = \ldots = f^{(n)}(x) = 0 \quad \text{und} \quad f^{(n+1)}(x) \neq 0,$$

dann besitzt f in x

- ein strenges lokales Minimum, falls n ungerade ist und $f^{(n+1)}(x) > 0$,
- ein strenges lokales Maximum, falls n ungerade ist und $f^{(n+1)}(x) < 0$,
- kein Extremum, falls n gerade ist.

Da die erste Ableitung, wenn sie existiert, auch wieder eine Funktion ist, können wir fragen, wo die erste Ableitung (lokal) maximal oder minimal wird, also an welchem Punkt wir die (lokal) größte oder kleinste Steigung der ursprünglichen Funktion haben. Solche Punkte nennt man *Wendepunkte*. Wendepunkte kann man allerdings auch für nicht differenzierbare Funktionen definieren. Dafür führen wir zunächst den Begriff der *Konvexität* einer Funktion ein.

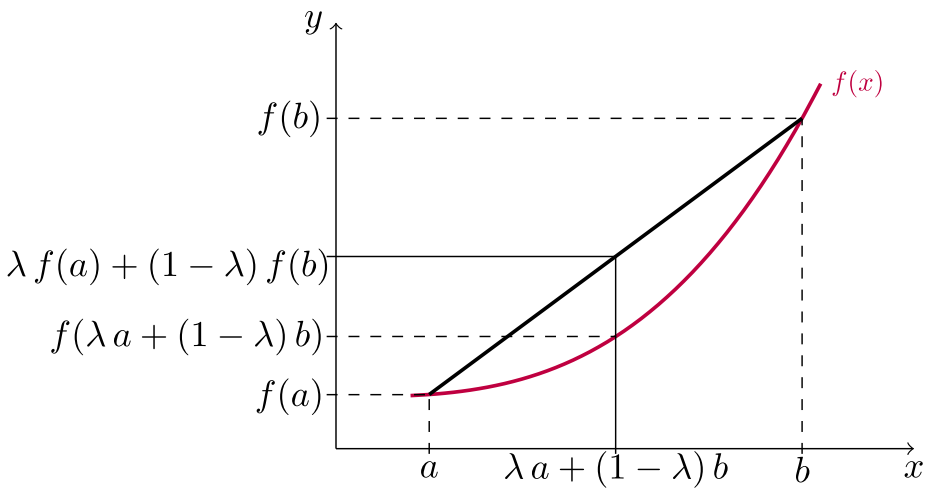
Definition 5.22 (Konvexität)

Eine Funktion $f : (a, b) \to \mathbb{R}$ heißt *konvex*, wenn für alle $x_1, x_2 \in (a, b)$ und alle $\lambda \in (0, 1)$ gilt

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

Die Funktion heißt *konkav*, wenn $-f$ konvex ist.

Gilt in der obigen Bedingung sogar $<$ anstelle von $\leq$, so nennen wir f *streng konvex* (bzw $-f$ *streng konkav*).



Die Abbildung zeigt den Graphen einer konvexen Funktion. Die Funktion hängt nach unten durch, d.h., die Funktionswerte liegen unterhalb der Verbindungsgeraden zwischen den Funktionswerten an den Intervallgrenzen. Man kann sich die Konvexität auch so vorstellen, dass jede Verbindungslinie zwischen zwei Punkten oberhalb der Funktion verläuft und den Graphen der Funktion nicht schneidet. Konvexe Funktionen haben einige sehr angenehme Eigenschaften: Zum Beispiel ist jedes lokale Minimum einer konvexen Funktion gleichzeitig das globale Minimum. Dies kann bei der Optimierung, d.h. dem Finden von globalen Extrema, ausgenutzt werden.

Der folgende Satz stellt den Zusammenhang zwischen Konvexität und zweiter Ableitung her.

Satz 5.23

Eine zweimal differenzierbare Funktion $f : (a, b) \to \mathbb{R}$ ist genau dann konvex, wenn $f''(x) \geq 0$ für alle $x \in (a, b)$.

▼ Beweis

⇐-Richtung:

Sei $f''(x) \geq 0$ für alle $x \in (a, b)$. Dann ist f' nach Satz 5.19 monoton wachsend. Sei $x = \lambda x_1 + (1 - \lambda)x_2$ für $a < x_1 < x_2 < b$ und $0 < \lambda < 1$, dann gilt offensichtlich $x_1 < x < x_2$.

Nach Mittelwertsatz 1 existieren dann $y_1 \in (x_1, x)$ und $y_2 \in (x, x_2)$ mit

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(y_1) \leq f'(y_2) = \frac{f(x_2) - f(x)}{x_2 - x}.$$

Da $x - x_1 = (1 - \lambda)(x_2 - x_1)$ und $x_2 - x = \lambda(x_2 - x_1)$ folgt

$$\frac{f(x) - f(x_1)}{1 - \lambda} \leq \frac{f(x_2) - f(x)}{\lambda} \quad \Rightarrow \quad f(x) \leq \lambda f(x_1) + (1 - \lambda)f(x_2).$$

Damit ist die Funktion konvex.

⇒-Richtung:

Sei nun f konvex. Angenommen es gäbe ein $x_0 \in (a, b)$ mit $f''(x_0) < 0$. Sei $c = f'(x_0)$ und $g(x) = f(x) - c(x - x_0)$. g ist zweimal differenzierbar und $g'(x_0) = 0$ sowie $g''(x_0) = f''(x_0) < 0$. Damit besitzt g nach Satz 5.20 an der Stelle x_0 ein (strenges) lokales Maximum.

Es gibt damit ein $\varepsilon > 0$, so dass $(x_0 - \varepsilon, x_0 + \varepsilon) \subset (a, b)$ und $g(x_0 - \varepsilon) < g(x_0)$, $g(x_0 + \varepsilon) < g(x_0)$. Aufgrund der Definition von g folgt damit auch

$$f(x_0) = g(x_0) > \frac{1}{2}(g(x_0 - \varepsilon) + g(x_0 + \varepsilon)) = \frac{1}{2}(f(x_0 - \varepsilon) + f(x_0 + \varepsilon)).$$

Wähle nun $x_1 = x_0 - \varepsilon$, $x_2 = x_0 + \varepsilon$ und $\lambda = \frac{1}{2}$. Dann ist $x_0 = \lambda x_1 + (1 - \lambda)x_2$ und

$$f(\lambda x_1 + (1 - \lambda)x_2) > \lambda f(x_1) + (1 - \lambda)f(x_2).$$

Dies steht im Widerspruch zur Konvexität von f.

□

Konvexe Bereiche einer Funktion haben also eine positive zweite Ableitung. Man sagt auch, dass der Graph der Funktion hier linksgekrümmt ist, bzw. für konkave Bereiche rechtsgekrümmt ist. Ein Wendepunkt ist nun ein Wechselpunkt dieser Krümmungsrichtung.

Definition 5.24 (Wendepunkt)

Eine stetige Funktion $f : (a, b) \to \mathbb{R}$ hat in $x \in (a, b)$ einen Wendepunkt, wenn Intervalle (α, x) und (x, β) existieren, für die eine der zwei folgenden Bedingungen gilt:

1. f ist in (α, x) streng konvex und in (x, β) streng konkav.
2. f ist in (α, x) streng konkav und in (x, β) streng konvex.

Zu beachten ist, dass die obige Definition nicht voraussetzt, dass f in x differenzierbar ist, oder überhaupt irgendwo differenzierbar ist.

Beispiel 5.25

Die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit

$$f = \begin{cases} \sqrt{x} & \text{für } x \geq 0 \\ -\sqrt{|x|} & \text{für } x < 0 \end{cases}$$

ist in $x = 0$ nicht differenzierbar, aber es gilt $f''(x) > 0$ für $x < 0$ und $f''(x) < 0$ für $x > 0$. Damit sind diese Bereiche nach Satz 5.23 streng konvex bzw. streng konkav (links- und rechtsgekrümmt), und bei $x = 0$ hat die Funktion einen Wendepunkt.

In den meisten Fällen haben wir es aber mit beliebig oft stetig differenzierbaren Funktionen zu tun, und dann gilt, dass Wendepunkte der Funktion gleichzeitig Extrema der ersten Ableitung sind. Die Bedingungen aus Satz 5.16 und Satz 5.20 lassen sich also auf Wendepunkte übertragen. Diese geben wir hier nur zur Vollständigkeit ohne Beweis an:

Satz 5.26

Eine dreimal differenzierbare Funktion $f : (a, b) \to \mathbb{R}$ besitzt einen Wendepunkt in x genau dann, wenn f' ein lokales Extremum in x besitzt. Damit folgt:

- Notwendige Bedingung für Wendepunkt:
  Wenn f in x einen Wendepunkt hat, dann ist $f''(x) = 0$.
- Hinreichende Bedingung für Wendepunkt:
  Wenn $f''(x) = 0$ und $f'''(x) \neq 0$, dann hat f in x einen Wendepunkt.

Auch hier kann man Funktionen finden, bei denen das hinreichende Kriterium verletzt ist, aber trotzdem ein Wendepunkt vorliegt, wie zum Beispiel $f(x) = x^5$ bei $x = 0$. In solchen Fällen kann man Satz 5.21 auf Wendepunkte übertragen.

5.3 Grenzwerte an Randbereichen

In diesem Kapitel verwenden wir erste Ableitungen, um Grenzwerte von Funktionen an Stellen zu berechnen, an denen die Funktion eigentlich nicht mehr definiert ist, nämlich an den Randbereichen des Definitionsbereiches, also an Parameterwerten, die nicht im Definitionsbereich, aber beliebig nahe am Definitionsbereich der Funktion liegen. Ein Beispiel wäre $x = 0$ für die Funktion $f(x) = 1/x$.

Oft streben Funktionen in Randbereichen immer größere (positive oder negative) Werte an. Analog zur *bestimmten Divergenz* bei Folgen definieren wir sogenannte *uneigentliche Grenzwerte* für Funktionen.

Definition 5.27 (Uneigentlicher Grenzwert)

Sei $f : A \to \mathbb{R}$ und a ein Häufungspunkt von A. Falls für alle $K \in \mathbb{R}$ ein $\delta > 0$ existiert, sodass $f(x) > K$ für $|x - a| < \delta$, so schreibt man $\lim\limits_{x \to a} f(x) = \infty$.

Anstelle von $\lim\limits_{x \to a} - f(x) = \infty$ schreibt man auch $\lim\limits_{x \to a} f(x) = -\infty$.

Speziell interessieren wir uns für Funktionen, die auf offenen oder halboffenen Intervallen definiert sind und deren Grenzwerte wir an der halboffenen Grenze bestimmen wollen. Hier gibt es drei Möglichkeiten: Entweder die Funktion konvergiert gegen eine reelle Zahl c, oder die Funktionswerte divergieren und werden betragsmäßig immer größer, streben also gegen ∞ oder $-\infty$, oder es gibt am Rand eine unbestimmte Divergenz (also analog zu Folgen wie $a_n = (-1)^n$).

Beispiel 5.28 (Grenzwerte an Randpunkten)

- $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = \frac{1}{x^2}$ hat den uneigentlichen Grenzwert $\lim_{x\to 0} f(x) = \infty$.
  Beweis: Für $K \in \mathbb{R}$ gibt es $\delta = 1/\sqrt{|K|}$, sodass für $|x - 0| < \delta$ folgt $f(x) = 1/x^2 = 1/|x|^2 > 1/\delta^2 = |K| > K$.
- $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = \frac{e^x - 1}{x}$ hat nach Satz 4.47 den Grenzwert $\lim_{x\to 0} f(x) = 1$.
- $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = \frac{1}{x}$ hat keinen Grenzwert und auch keinen uneigentlichen Grenzwert für $\lim_{x\to 0} f(x)$, denn es gilt $\lim_{x\nearrow 0} f(x) = -\infty \neq \lim_{x\searrow 0} f(x) = \infty$. Den Beweis für die beiden einseitigen Grenzwerte führt man wie im ersten Beispiel.
- $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = \sin\left(\frac{1}{x}\right)$ hat keinen Grenzwert und auch keinen uneigentlichen Grenzwert für $\lim_{x\to 0} f(x)$. (Beweis als Übung).

Wir wissen bereits, dass wir (mit Hilfe von Satz 3.18) den Grenzwert eines Quotienten $f(x)/g(x)$ als

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \frac{\lim_{x\to a} f(x)}{\lim_{x\to a} g(x)}$$

bestimmen können. Das ganze funktioniert allerdings **nur dann**, wenn die beiden Grenzwerte $\lim_{x\to a} f(x)$ und $\lim_{x\to a} g(x)$ existieren (und damit insbesondere reelle Zahlen $c \neq \infty$ sind) und wenn zusätzlich $\lim_{x\to a} g(x) \neq 0$ ist. In den folgenden beiden Fällen ist das aber nicht mehr der Fall:

$$\lim_{x\to 0} \frac{\exp(x) - 1}{\sin(x)} \quad \text{oder} \quad \lim_{x\to\infty} \frac{x^2}{\exp(x)}$$

Der erste Fall führt zu $\frac{0}{0}$, der zweite zu $\frac{\infty}{\infty}$, was beides nicht definiert ist. Beide Grenzwerte existieren aber dennoch, wir haben nur (noch) nicht die Mittel, um sie zu berechnen.

Meist ergeben sich solche Situationen an den Rändern des Definitionsbereiches, wo dann durch Null geteilt werden würde (z.B. $x = 0$ bei $1/x$) oder Funktionen gegen Unendlich gehen (z.B. $x \to \infty$ bei $\exp(x)$). Dies muss aber nicht zwangsweise bedeuten, dass eine Funktion $f(x)/g(x)$ am Rand des Definitionsbereiches divergiert. Falls der Nenner $g(x)$ "gleich schnell" oder "schneller" als der Zähler $f(x)$ gegen Null strebt, kann sich ein reeller Grenzwert ergeben (wie z.B. bei $f(x) = x^2/x$). Den Grenzwert können wir in solchen Fällen mithilfe der sogenannten *Regel von L'Hospital* bestimmen. Für deren Beweis benötigen wir zunächst eine Erweiterung des Mittelwertsatzes.

Satz 5.29 (Zweiter Mittelwertsatz der Differentialrechnung)

Seien $f, g : [a, b] \to \mathbb{R}$ zwei Funktionen, die in $[a, b]$ stetig und in (a, b) differenzierbar sind. Sei außerdem $g'(x) \neq 0$ für alle $x \in (a, b)$. Dann ist $g(a) \neq g(b)$ und es existiert ein $c \in (a, b)$ mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

▼ Beweis

Wir beweisen zuerst $g(a) \neq g(b)$: Wäre $g(a) = g(b)$, so gäbe es nach dem Satz von Rolle ein $c \in (a, b)$ mit $g'(c) = 0$, was aber laut Voraussetzung ausgeschlossen ist.

Für die Hilfsfunktion $h(x) = f(x) - \frac{f(b)-f(a)}{g(b)-g(a)} (g(x) - g(a))$ gilt $h(a) = h(b) = f(a)$. Damit existiert nach dem Satz von Rolle $c \in (a, b)$ mit $h'(c) = 0$ und somit gilt $h'(c) = f'(c) - \frac{f(b)-f(a)}{g(b)-g(a)} g'(c) = 0$.

Wegen $g'(c) \neq 0$ können wir daraus die Behauptung folgern:

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

□

Daraus folgt nun die Regel von L'Hospital (gesprochen "lopital").

Satz 5.30 (Regel von L'Hospital)

Seien $f, g : (a, b) \to \mathbb{R}$ zwei differenzierbare Funktionen, und sei $g(x) \neq 0$ und $g'(x) \neq 0$ für alle $x \in (a, b)$. Gilt einer der beiden Fälle

a. $\lim\limits_{x \searrow a} f(x) = \lim\limits_{x \searrow a} g(x) = 0$

b. $\lim\limits_{x \searrow a} |f(x)| = \lim\limits_{x \searrow a} |g(x)| = \infty$,

und existiert der Grenzwert $\lim\limits_{x \searrow a} \frac{f'(x)}{g'(x)}$ (eigentlich oder uneigentlich), dann existiert auch der Grenzwert $\lim\limits_{x \searrow a} \frac{f(x)}{g(x)}$ (eigentlich oder uneigentlich) und es gilt

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = \lim_{x \searrow a} \frac{f'(x)}{g'(x)}.$$

Dies gilt analog für die rechte Intervallgrenze $\lim\limits_{x \nearrow b} \frac{f(x)}{g(x)}$ sowie für $a, b = \pm\infty$.

▼ Beweis

Wir beweisen nur **Fall (a)**.

f und g sind auf (a, b) differenzierbar und damit auch stetig (Satz 5.3). Da der rechtsseitige Grenzwert für $x \searrow a$ existiert und der Definitionsbereich nur aus Werten $x > a$ besteht, können wir die Funktionen f und g durch das Setzen von $f(a) := 0$ und $g(a) := 0$ stetig auf $[a, b)$ erweitern.

Sei (x_n) eine Folge, die von oben gegen a konvergiert, also $x_n \in (a, b)$ und $x_n \searrow a$. Für jedes x_n gibt es dann nach Satz 5.29 ein $c_n \in (a, x_n)$ mit

$$\frac{f'(c_n)}{g'(c_n)} = \frac{f(x_n) - f(a)}{g(x_n) - g(a)} = \frac{f(x_n) - 0}{g(x) - 0} = \frac{f(x_n)}{g(x_n)}.$$

Sei (c_n) die Folge der Werte, für die die obige Gleichung gilt. Da $(x_n) \searrow a$ und $c_n \in (a, x_n)$, konvergiert auch $(c_n) \searrow a$. Somit folgt insgesamt

$$\lim_{n \to \infty} \frac{f(x_n)}{g(x_n)} = \lim_{n \to \infty} \frac{f'(c_n)}{g'(c_n)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

Die letzte Umformung gilt, da der Grenzwert $\lim\limits_{x \searrow a} \frac{f'(x)}{g'(x)}$ existiert und daher für jede Folge $x_n \to a$ (oder $c_n \to a$) angenommen wird.

Die weiteren Fälle, also $\lim\limits_{x \nearrow b}$ und $a, b = \pm\infty$, beweist man analog.

□

Die Forderung nach der Existenz von $\lim\limits_{x \searrow a} \frac{f'(x)}{g'(x)}$ im eigentlichen oder uneigentlichen Sinne bedeutet, dass der Grenzwert entweder eine reelle Zahl (*eigentlicher Grenzwert*) oder $\pm\infty$ (*uneigentlicher Grenzwert*) sein muss. Beispiel 5.32 diskutiert einen Fall, wo der Grenzwert weder eigentlich noch uneigentlich existiert und der Satz daher nicht angewendet werden kann.

Da die Regel sowohl für rechts- als auch für linksseitige Grenzwerte identisch definiert ist, kann man sie auch allgemein für Grenzwerte $x \to c$ anwenden, wenn die Funktionen auf beiden Seiten der betreffenden Stelle c definiert sind, wie z.B. $\lim\limits_{x \to 0} 1/x$.

Die Gültigkeit für die Regel von L'Hospital kann man sich für den Fall **(a)** auch vereinfacht folgendermaßen plausibel machen: Wenn die Funktionen f und g differenzierbar sind, können wir beide in einem sehr kleinen Bereich um a durch ihre jeweilige Tangente approximieren, also

$$f(x) \approx f'(a)x + n_f \qquad g(x) \approx g'(a)x + n_g.$$

Da beide Funktionen in $x = a$ eine Nullstelle haben (da wir Fall (a) des Satzes betrachten), müssen auch die Tangenten bei $x = a$ eine Nullstelle haben. Aus $f'(x)a + n_f = 0$ folgt $n_f = -f'(x)a$ und analog $n_g = -g'(a)a$. Insgesamt können wir die Funktionen also in der Nähe von a approximieren durch

$$f(x) \approx f'(a)(x - a) \qquad g(x) \approx g'(a)(x - a).$$

Für den Quotienten ergibt sich

$$\frac{f(x)}{g(x)} \approx \frac{f'(a)(x-a)}{g'(a)(x-a)} = \frac{f'(a)}{g'(a)}.$$

Je näher wir $x = a$ kommen, desto besser wird die Näherung, sodass beide Seiten im Grenzfall identisch sind.

Beispiel 5.31

- $f(x) = x^2,\ g(x) = x$. Es ist $f(0) = g(0) = 0$ und daher gilt

$$\lim_{x\to 0} \frac{x^2}{x} \overset{\text{L'H } \frac{0}{0}}{=} \lim_{x\to 0} \frac{(x^2)'}{(x)'} = \lim_{x\to 0} \frac{2x}{1} = 0.$$

- $f(x) = \exp(x) - 1,\ g(x) = \sin x$. Es ist $f(0) = g(0) = 0$ und daher gilt

$$\lim_{x\to 0} \frac{\exp(x) - 1}{\sin(x)} \overset{\text{L'H } \frac{0}{0}}{=} \lim_{x\to 0} \frac{(\exp(x) - 1)'}{(\sin(x))'} = \lim_{x\to 0} \frac{\exp(x)}{\cos(x)} = \frac{1}{1} = 1.$$

- $f(x) = x^2,\ g(x) = \exp(x)$. Es ist $\lim\limits_{x\to\infty} f(x) = \lim\limits_{x\to\infty} g(x) = \infty$. Daher gilt

$$\lim_{x\to\infty} \frac{x^2}{\exp(x)} \overset{\text{L'H } \frac{\infty}{\infty}}{=} \lim_{x\to\infty} \frac{2x}{\exp(x)} \overset{\text{L'H } \frac{\infty}{\infty}}{=} \lim_{x\to\infty} \frac{2}{\exp(x)} = 0.$$

- Für $x \ln(x)$ soll der Grenzwert für $x \searrow 0$ bestimmt werden. Hier hätten wir den Fall $0 \cdot (-\infty)$, der aber nicht vom Satz von L'Hospital abgedeckt wird. Wir können aber geschickt umformen:

$$\lim_{x\searrow 0} x \ln(x) = \lim_{x\searrow 0} \frac{\ln(x)}{\frac{1}{x}} \overset{\text{L'H } \frac{\infty}{\infty}}{=} \lim_{x\searrow 0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = -\lim_{x\searrow 0} \frac{x^2}{x} = \lim_{x\searrow 0} x = 0.$$

- Mit dem vorigen Grenzwert können wir eine alte Frage beantworten, die Sie sich vielleicht schon in Kapitel 2 gestellt haben, als Potenzen definiert wurden. Unsere bisherigen Definitionen haben den Fall 0^0 nie mit eingeschlossen. Dazu bestimmen wir nun den folgenden Grenzwert:

$$\lim_{x\searrow 0} x^x = \lim_{x\searrow 0} \exp(x \ln(x)) = \exp\left(\lim_{x\searrow 0} x \ln(x)\right) = \exp(0) = 1.$$

  Hierbei nutzen wir die Stetigkeit von $\exp$ aus, wodurch wir den Limes in die Exponentialfunktion ziehen dürfen.

Das dritte und vierte Beispiel zeigt, dass man die Regel auch in Fällen anwenden kann, in denen es zunächst nicht danach aussieht, da wir keinen Quotienten von zwei Funktionen betrachten. Die nützlichsten Umformungen für solche Situationen sind diese:

- "$0 \cdot \infty$" führt auf "$\frac{0}{0}$":

$$f(x) \cdot g(x) = \frac{f(x)}{\frac{1}{g(x)}}$$

- "0^0" bzw. "∞^0" führt auf "$0 \cdot \infty$":

$$f(x)^{g(x)} = \exp\left(g(x) \cdot \ln(f(x))\right)$$

Da $\exp$ stetig ist, darf man hier den Grenzwert in die Funktion ziehen.

- "$\infty - \infty$" führt auf "$\frac{0}{0}$":

$$f(x) - g(x) = \frac{1}{\frac{1}{f(x)}} - \frac{1}{\frac{1}{g(x)}} = \frac{\frac{1}{g(x)} - \frac{1}{f(x)}}{\frac{1}{f(x)g(x)}}$$

In diesen Fällen müssen allerdings die Ableitungen aus Zähler und Nennerfunktion der jeweiligen Umformung gebildet werden. Die Regel von L'Hospital darf auch mehrfach angewendet werden, wenn die Grenzwerte der ersten Ableitungen wiederum die Bedingungen der Regel von L'Hospital erfüllen. Es ist allerdings immer darauf zu achten, dass die Voraussetzungen erfüllt sind. Wenn die Grenzwerte der Funktionen $\neq 0$ (bzw. $\neq \infty$) sind, oder wenn der Grenzwert $\lim \frac{f'(x)}{g'(x)}$ nicht existiert, so gelten die Regeln nicht, wie in diesen Beispielen gezeigt wird:

Beispiel 5.32

- Der Grenzwert $\lim\limits_{x\to 1} \frac{x^2}{x}$ ist offensichtlich 1. Da Zähler und Nenner an 1 nicht 0 und nicht ∞ sind, dürfen wir L'Hospital nicht anwenden. Das würde nämlich ergeben

$$\lim_{x\to 1} \frac{x^2}{x} = 1 \neq \lim_{x\to 1} \frac{2x}{1} = 2.$$

- Bestimme den Grenzwert $\lim\limits_{x\to\infty} \frac{x+\sin(x)}{x}$. Mit $f(x) = x + \sin(x)$ und $g(x) = x$ gilt zwar $\lim\limits_{x\to\infty} f(x) = \lim\limits_{x\to\infty} g(x) = \infty$, aber der Grenzwert

$$\lim_{x\to\infty} \frac{f'(x)}{g'(x)} = \lim_{x\to\infty} \frac{1+\cos(x)}{1}$$

existiert nicht (weil Cosinus oszilliert). Daher können wir der Satz von L'Hospital nicht anwenden. Der gesuchte Grenzwert existiert aber dennoch

$$\lim_{x\to\infty} \frac{x+\sin(x)}{x} = \lim_{x\to\infty} \left(1 + \frac{\sin(x)}{x}\right) = 1 + \lim_{x\to\infty} \frac{\sin(x)}{x} = 1 + 0 = 1.$$

5.4 Kurvendiskussion

An diesem Punkt haben wir alle Regel für Ableitungen und Extremwertbestimmung gesehen. Hier fassen wir einige der bisherigen Resultate noch einmal zusammen und nutzen sie zur Analyse von Funktionen in Form einer Kurvendiskussion. Sei $f : A \to \mathbb{R}$ eine in A differenzierbare Funktion. Wir analysieren für f die folgenden Eigenschaften.

Symmetrie

Nach Definition 4.57 prüfen wir für Achsensymmetrie, ob $f(-x) = f(x)$, und für Punktsymmetrie, ob $f(-x) = -f(x)$ für alle $x \in A$ gilt.

Verhalten am Rand des Definitionsbereichs

Interessant sind Häufungspunkte a von A, die nicht zu A gehören, sowie $-\infty$ und ∞, falls A nach unten oder oben unbeschränkt ist. Die folgenden Grenzwerte werden für Häufungspunkte untersucht:

$$\lim_{x \nearrow a} f(x) \quad \text{oder} \quad \lim_{x \searrow a} f(x).$$

Dabei ist jeweils festzustellen, ob ein eigentlicher oder uneigentlicher Grenzwert existiert. Für unbeschränkte Definitionsbereiche A werden die Grenzwerte

$$\lim_{x \to \infty} f(x) \quad \text{oder} \quad \lim_{x \to -\infty} f(x)$$

untersucht. Auch hier wird wieder die Frage nach der Konvergenz oder der Existenz der uneigentlichen Grenzwerte gestellt.

Für die Fälle $\lim\limits_{x \to \infty} f(x)$ und $\lim\limits_{x \to -\infty} f(x)$ kann das asymptotische Verhalten analysiert werden. Dazu wird eine Gerade $g(x) = \alpha x + \beta$ mit $\alpha, \beta \in \mathbb{R}$ gesucht, sodass

$$\lim_{x \to \infty} (f(x) - g(x)) = 0 \quad \text{bzw.} \quad \lim_{x \to -\infty} (f(x) - g(x)) = 0.$$

Die Gerade heißt Asymptote von f für $x \to \infty$ (bzw. für $x \to -\infty$). Die folgenden Bedingungen charakterisieren eine Asymptote und erlauben eine direkte Bestimmung der Werte von α und β:

$$\begin{aligned} \lim_{x \to \infty} \frac{f(x)}{x} &= \alpha \quad \text{bzw.} \quad \lim_{x \to -\infty} \frac{f(x)}{x} = \alpha, \\ \lim_{x \to \infty} (f(x) - \alpha x) &= \beta \quad \text{bzw.} \quad \lim_{x \to -\infty} (f(x) - \alpha x) = \beta. \end{aligned}$$

Die zweite Bedingung folgt direkt aus der Bedingung an die Asymptote, die erste Bedingung folgt aus

$$\lim_{x \to \infty} \left(\frac{f(x) - g(x)}{x} \right) = \lim_{x \to \infty} \left(\frac{f(x) - \alpha x - \beta}{x} \right) = \lim_{x \to \infty} \left(\frac{f(x)}{x} - \alpha \right),$$

oder aus dem entsprechenden Grenzwert gegen $-\infty$.

Nullstellen

Eine Nullstelle liegt vor, wenn $f(x) = 0$.

Extrempunkte und Monotonie

Hier nutzen wir entweder direkt die Definition 5.15 oder die notwendigen und hinreichenden Kriterien für Extrema (Satz 5.16, 5.20), also

$$f'(x) = 0 \quad \wedge \quad f''(x) \neq 0.$$

Über das Verhalten der Funktion am Rand des Definitionsbereichs kann festgestellt werden, ob es sich eventuell sogar um globale Extrema handelt.

Da sich an Extrempunkten das Vorzeichen der ersten Ableitung ändert und wir über das Vorzeichen der ersten Ableitung die Monotonie bestimmen können (Satz 5.19), teilen die Extrempunkte die Funktion in monotone Teilintervalle: Von einem Hoch- zu einem Tiefpunkt fällt die Funktion, von einem Tief- zu einem Hochpunkt steigt die Funktion. Ist der Extrempunkt mit dem höchsten (niedrigsten) x-Wert ein Hochpunkt, muss die Funktion danach (davor) fallen und bei einem Tiefpunkt steigen.

Wendepunkt und Konvexität

Ein Wendepunkt liegt dann vor, wenn sich das Vorzeichen der zweiten Ableitung ändert. Wir nutzen entweder direkt diese Bedingung aus Definition 5.24 oder das Resultat aus Satz 5.26, mit dem wir die notwendigen und hinreichenden Bedingungen der Extrempunkte auf Wendepunkte übertragen konnten, also:

$$f''(x) = 0 \quad \wedge \quad f'''(x) \neq 0.$$

Analog zum Zusammenhang zwischen Extrema und Monotonie, teilen Wendepunkte die Funktion in konvexe und konkave Teilintervalle.

Funktionsgraph

Aus den bisherigen Analysen lassen sich erste Aussagen über den Verlauf der Funktion machen. Diese können durch die Bestimmung des Verlaufs der Funktion, durch punktweises Abtasten, und die graphische Darstellung noch ergänzt werden.

Beispiel 5.33

Als Beispiel untersuchen wir die Funktion

$$f(x) = \frac{x^3 + 2x^2}{x^2 + 2x + 1} = \frac{x^2(x+2)}{(x+1)^2}$$

mit Definitionsbereich $A = \mathbb{R} \setminus \{-1\}$.

Symmetrie

Da $1 \in A$ aber $-1 \notin A$, kann keine Symmetrie vorliegen.

Verhalten am Rand des Definitionsbereichs

Interessant ist das Verhalten der Funktion für x gegen einen der Werte $-\infty,\ \infty,\ -1$.

Im Punkt -1 gilt: Der Nenner $(x+1)^2$ ist > 0 für alle $x \in \mathbb{R} \setminus \{-1\}$. Der Zähler $x^2(x+2)$ ist > 0 für $x > -2$. Damit ist $f(x)$ positiv für $x > -2$. Es gilt

$$\lim_{x \nearrow -1} = \lim_{x \searrow -1} f(x) = \infty,$$

da $\lim\limits_{x \to -1} x^2(x+2) = 1$ und $\lim\limits_{x \to -1} (x+1)^2 = 0$.

Für die Grenzwerte $x \to \infty$ und $x \to -\infty$ gilt

$$\lim_{x \to \infty} f(x) = \lim_{x \to \infty} \frac{x^3 + 2x^2}{x^2 + 2x + 1} = \lim_{x \to \infty} \frac{x+2}{1 + \frac{2}{x} + \frac{1}{x^2}} = \infty$$

$$\lim_{x \to -\infty} f(x) = \lim_{x \to -\infty} \frac{x^3 + 2x^2}{x^2 + 2x + 1} = \lim_{x \to -\infty} \frac{x+2}{1 + \frac{2}{x} + \frac{1}{x^2}} = -\infty$$

Hier hätte auch die Regel von L'Hospital (mehrfach) angewandt werden können, da der Grenzwert im uneigentlichen Sinne existiert:

$$\lim_{x \to \infty} \frac{x^3 + 2x^2}{x^2 + 2x + 1} \overset{\text{L'H } \frac{\infty}{\infty}}{=} \lim_{x \to \infty} \frac{3x^2 + 4x}{2x + 2} \overset{\text{L'H } \frac{\infty}{\infty}}{=} \lim_{x \to \infty} \frac{6x + 4}{2} = \infty$$

$$\lim_{x \to -\infty} \frac{x^3 + 2x^2}{x^2 + 2x + 1} \overset{\text{L'H } \frac{\infty}{\infty}}{=} \lim_{x \to -\infty} \frac{3x^2 + 4x}{2x + 2} \overset{\text{L'H } \frac{\infty}{\infty}}{=} \lim_{x \to -\infty} \frac{6x + 4}{2} = -\infty$$

Die Asymptote $g(x) = \alpha x + \beta$ für $x \to \infty$ wird wie folgt berechnet:

$$\alpha = \lim_{x \to \infty} \frac{x^2(x+2)}{x(x+1)^2} = \lim_{x \to \infty} \frac{x^2 + 2x}{x^2 + 2x + 1} = 1$$

$$\beta = \lim_{x \to \infty} \left(\frac{x^2(x+2)}{(x+1)^2} - x \right) = \ldots = \lim_{x \to \infty} \frac{-x}{x^2 + 2x - 1} = 0$$

Nullstellen

Es gilt $f(x) = \frac{x^2(x+2)}{(x+1)^2} = 0$ für $x = 0$ oder $x = -2$.

Der Schnittpunkt mit der y-Achse liegt an der Stelle 0, da $f(0) = 0$.

Extrempunkte und Monotonie

Zur Bestimmung der Extremstellen berechnen wir die erste Ableitung mithilfe der Quotientenregel:

$$\begin{aligned} f'(x) &= \frac{(3x^2+4x)(x+1)^2 - 2(x+1)(x^3+2x^2)}{(x+1)^4} \\ &= \frac{3x^3+4x^2+3x^2+4x-2x^3-4x^2}{(x+1)^3} \\ &= \frac{x^3+3x^2+4x}{(x+1)^3} \\ &= \frac{x(x^2+3x+4)}{(x+1)^3}. \end{aligned}$$

$f'(x)$ besitzt eine Nullstelle im Punkt 0. Für den Term in der Klammer des Zählers gilt $x^2+3x+4 = \left(x+\frac{3}{2}\right)^2 + \frac{7}{4} \geq \frac{7}{4} > 0$ für alle $x \in \mathbb{R}$. Damit ist der Zähler > 0 für $x > 0$, $= 0$ für $x = 0$ und < 0 für $x < 0$. Der Nenner ist > 0 für $x > -1$ und < 0 für $x < -1$, sodass

$$f'(x) \begin{cases} > 0 & \text{für } x \in (-\infty, -1) \cup (0, \infty), \\ < 0 & \text{für } x \in (-1, 0), \\ = 0 & \text{für } x = 0. \end{cases}$$

Daraus können wir die folgenden Aussagen über den Verlauf von f ableiten:

- f ist streng monoton wachsend auf $(-\infty, -1)$.
- f streng monoton fallend auf $(-1, 0)$.
- f ist streng monoton wachsend auf $(0, \infty)$.
- f hat in $x = 0$ ein lokales Minimum.

Wie wir gleich sehen werden, gilt $f''(0) = 2 > 0$, sodass auch die hinreichende Bedingung für das lokale Extremum erfüllt ist. Ein globales Extremum existiert nicht, da die Funktionswerte uneigentlich gegen $-\infty$ und ∞ konvergieren.

Wendepunkte und Konvexität

Zur Bestimmung der Wendepunkte bestimmen wir die zweite Ableitung:

$$
\begin{aligned}
f''(x) &= \left(\frac{x(x^2+3x+4)}{(x+1)^3} \right)' \\
&= \frac{(3x^2+6x+4)(x+1)^3 - (x^3+3x^2+4x)3(x+1)^2}{(x+1)^6} \\
&= \frac{3x^3+6x^2+4x+3x^2+6x+4-3x^3-9x^2-12x}{(x+1)^4} \\
&= \frac{-2x+4}{(x+1)^4} \\
&= \frac{2(2-x)}{(x+1)^4}.
\end{aligned}
$$

Der Nenner ist im Definitionsbereich der Funktion positiv, während der Zähler für $x < 2$ positiv und für $x > 2$ negativ ist. Im Punkt $x = 2$ hat $f''(x)$ eine Nullstelle. Damit gilt

$$
f''(x) \begin{cases} > 0 & \text{für } (-\infty, -1) \cup (-1, 2) \\ = 0 & \text{für } x = 2 \\ < 0 & \text{für } (2, \infty). \end{cases}
$$

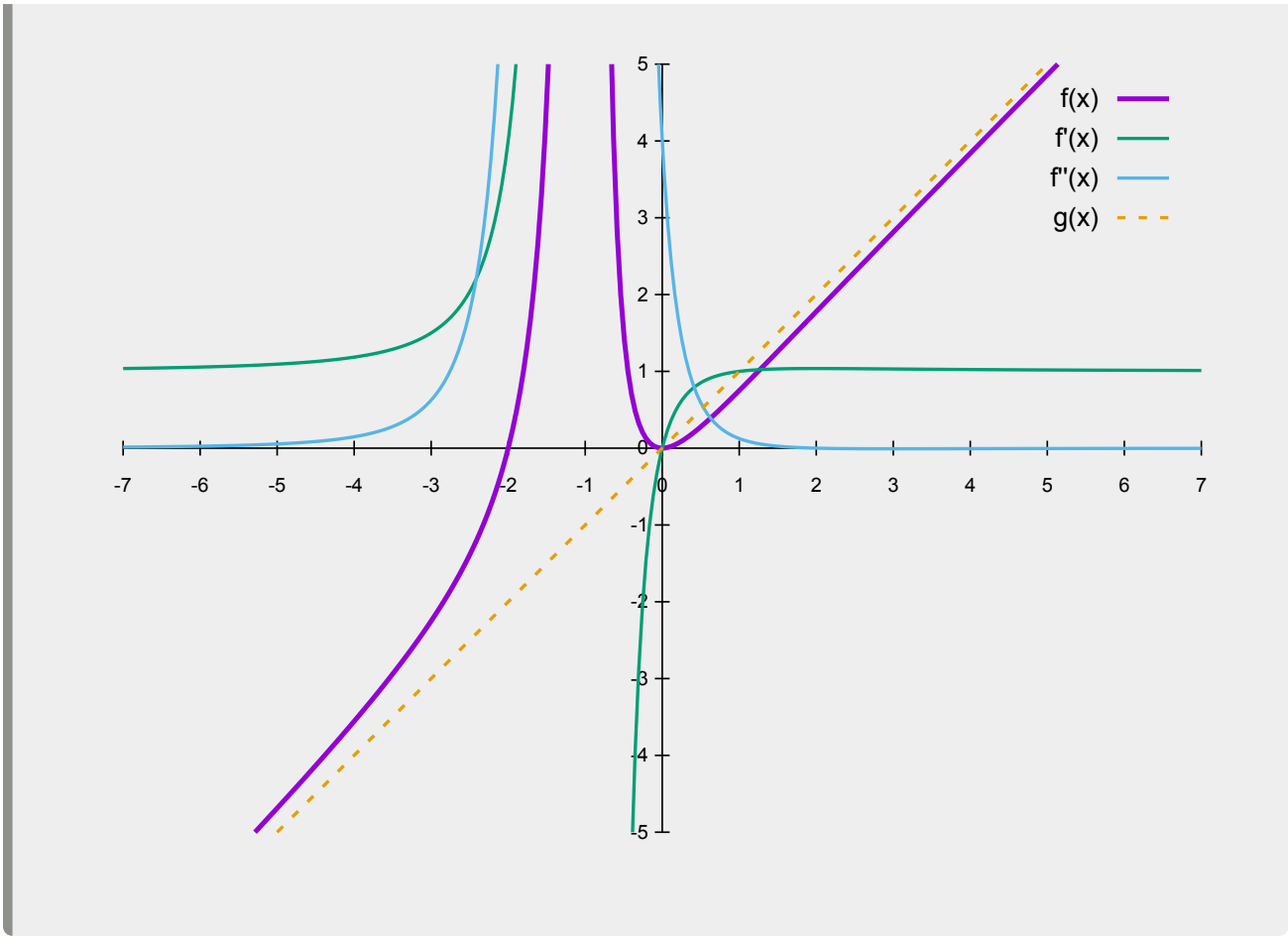
Die Funktion hat an der Stelle $x = 2$ einen Wendepunkt. Für die dritte Ableitung gilt

$$
\begin{aligned}
f'''(x) &= \left(\frac{4-2x}{(x+1)^4} \right)' \\
&= \frac{-2(x+1)^4 - (4-2x) \cdot 4 \cdot (x+1)^3}{(x+1)^8} \\
&= \frac{-2(x+1) - (4-2x)4}{(x+1)^5} \\
&= \frac{6x-18}{(x+1)^5} \\
&= \frac{6(x-3)}{(x+1)^5}
\end{aligned}
$$

Da $f''(2) = 0$ und $f'''(2) = \frac{-6}{3^5} = -\frac{1}{27} \neq 0$, ist die hinreichende Bedingung für einen Wendepunkt erfüllt.

Funktionsgraph

Die folgende Graphik zeigt den Verlauf der Funktion $f(x)$ im Intervall $[-7, 7]$, sowie auch die ersten beiden Ableitungen $f'(x)$ und $f''(x)$ und die Asymptote $g(x)$:



5.5 Der Satz von Taylor

Wir haben bereits die Darstellung verschiedener Funktionen, wie der Exponentialfunktion, der Cosinus- oder Sinus-Funktion, durch unendliche Reihen kennen gelernt. In diesem Kapitel betrachten wir nun die Approximation von Funktionen durch Potenzreihen in allgemeiner Form. Die Taylorsche Formel versucht, den Verlauf einer Funktion in der Umgebung einer Stelle x_0 durch Kenntnis ihrer Ableitungen an dieser Stelle zu approximieren. Die sich daraus ergebende *Taylorentwicklung* der Funktion hat hohe Relevanz in der Praxis, da hiermit Funktionen wie Sinus, Cosinus, Logarithmen, allgemeine Potenzen, Wurzeln, etc. durch einfache Polynome (unter gewissen Bedingungen) beliebig genau approximiert werden können.

Satz 5.34 (Taylorsche Formel)

Sei $f : A \to \mathbb{R}$ eine $(n+1)$-mal stetig differenzierbare Funktion, und seien die Funktionen $T_n, R_n : A \to \mathbb{R}$ für ein beliebiges $n \in \mathbb{N}$ und $x_0 \in A$ definiert als

$$T_n(x) := \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \qquad \text{und} \qquad R_n(x) := f(x) - T_n(x).$$

Dann existiert zu jedem $x \in A \setminus \{x_0\}$ ein y zwischen x und x_0, sodass

$$R_n(x) = \frac{f^{(n+1)}(y)}{(n+1)!}(x - x_0)^{n+1}.$$

▼ Beweis

Wir betrachten hier nur den Fall $x < x_0$. Für $x > x_0$ funktioniert der Beweis genauso.

Die Funktion R_n ist $(n+1)$-mal differenzierbar, da f nach Voraussetzung $(n+1)$-mal differenzierbar ist und $T_n(x)$ als Polynom unendlich oft differenzierbar ist.

Wir definieren eine Hilfsfunktion $g_n : A \to \mathbb{R}$ mit

$$g_n(x) := (x - x_0)^{n+1}.$$

Es gilt $R_n(x_0) = g_n(x_0) = 0$, woraus folgt

$$\frac{R_n(x)}{g_n(x)} = \frac{R_n(x) - R_n(x_0)}{g_n(x) - g_n(x_0)}.$$

Hier sind die Bedingungen für den zweiten Mittelwertsatz (Satz 5.29) gegeben. Es existiert also ein $y_1 \in (x, x_0)$, sodass

$$\frac{R_n(x)}{g_n(x)} = \frac{R_n'(y_1)}{g_n'(y_1)} = \frac{R_n'(y_1)}{(n+1)(y_1 - x_0)^n}.$$

Auch für die ersten Ableitungen gilt $R_n'(x_0) = g_n'(x_0) = 0$, woraus wieder folgt

$$\frac{R_n(x)}{g_n(x)} = \frac{R_n'(y_1)}{g_n'(y_1)} = \frac{R_n'(y_1) - R_n'(x_0)}{g_n'(y_1) - g_n'(x_0)}.$$

Hierauf können wir erneut den zweiten Mittelwertsatz (Satz 5.29) anwenden. Es existiert also ein $y_2 \in (y_1, x_0)$, sodass

$$\frac{R_n(x)}{g_n(x)} = \frac{R_n''(y_2)}{g_n''(y_2)} = \frac{R_n''(y_2)}{(n+1) \cdot n \cdot (y_2 - x_0)^{n-1}}.$$

Dies können wir insgesamt $(n+1)$-mal durchführen, und erhalten schließlich

$$\frac{R_n(x)}{g_n(x)} = \frac{R_n^{(n+1)}(y_{n+1})}{g_n^{(n+1)}(y_{n+1})} = \frac{R_n^{(n+1)}(y_{n+1})}{(n+1)!}$$

für ein $y_{n+1} \in (y_n, x_0) \subset (y_{n-1}, x_0) \subset \cdots \subset (y_1, x_0) \subset (x, x_0)$.

Für $R_n(x)$ gilt

$$R_n^{(n+1)}(x) = f^{(n+1)}(x),$$

da

$$T_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

ein Polynom n-ten Grades ist, weswegen die $(n+1)$-te Ableitung $T_n^{(n+1)}(x)$ Null ist.

Insgesamt folgt also die Behauptung:

$$\frac{R_n(x)}{g_n(x)} = \frac{R_n^{(n+1)}(y_{n+1})}{(n+1)!} = \frac{f^{(n+1)}(y_{n+1})}{(n+1)!} \Leftrightarrow R_n(x) = \frac{f^{(n+1)}(y)}{(n+1)!}(x - x_0)^{n+1}.$$

□

Machen Sie sich die Bedeutung des Satzes bewusst: Der Unterschied zwischen der Funktion f und dem Polynom T_n entspricht dem Restglied $R_n(x)$. In vielen Fällen (aber nicht immer) ist dieses Restglied besonders für geringe Differenzen $|x - x_0|$ klein. Wir hatten sogar schon gezeigt, dass die im Restglied enthaltene Folge $(x - x_0)^{n+1}/(n+1)!$ eine Nullfolge ist (da die Exponentialreihe $\sum x^k/k!$ für jedes $x \in \mathbb{R}$ konvergiert). Daher können wir das Restglied in vielen praxisrelevanten Fällen für ein ausreichend hohes n vernachlässigen und anstelle der Funktion, die beliebig komplex sein kann, ein einfaches Polynom auswerten.

Wir hatten auch schon gesehen, dass wir eine Funktion $f(x)$ an einer Stelle x_0 durch ihre Tangente $T(x) = f(x) + f'(x) \cdot (x - x_0)$ approximieren können. Dies ist sogar die beste lineare Approximation von f an dieser Stelle, mit den Eigenschaften $f(x_0) = T(x_0)$ und $f'(x_0) = T'(x_0)$. Die Tangente $T(x)$ ist nur ein Spezialfall der Taylorschen Formel für $n = 1$. Die Polynome $T_n(x)$ gehen noch einen Schritt weiter, da sie (wie man recht leicht sieht) im Punkt x_0 bis zur n-ten Ableitung mit f übereinstimmen:

$$\forall k \in \{0, \ldots, n\} : f^{(k)}(x_0) = T_n^{(k)}(x_0).$$

In diesem Sinne definieren wir für beliebig oft differenzierbare Funktionen die sogenannte *Taylorreihe*.

Definition 5.35 (Taylorreihe, Taylorpolynom)

Sei $f : A \to \mathbb{R}$ eine beliebig oft differenzierbare Funktion in $x_0 \in A$. Dann heißt

$$T[f, x_0](x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

die *Taylorreihe* von f im (Entwicklungs-)Punkt x_0.

Die n-te Teilsumme der Taylorreihe $T_n[f, x_0](x)$ bezeichnen wir als *Taylorpolynom vom Grad* n mit *Entwicklungspunkt* x_0:

$$T_n[f, x_0](x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k.$$

Wenn Sie sich an Definition 3.64 zurückerinnern, dann wird klar, dass jede Taylorreihe eine Potenzreihe ist. Potenzreihen hatten einen bestimmten Konvergenzradius (vgl. Definition 3.68), in dem sie konvergieren. Machen Sie sich bei der Funktionsapproximation durch Taylorreihen Folgendes unbedingt bewusst:

- Die Taylorreihe kann unter Umständen für alle $x \in A \setminus \{x_0\}$ divergieren. D.h. sie kann einen Konvergenzradius von 0 haben.
- Auch wenn die Taylorreihe von f im Punkt x konvergiert, konvergiert sie nicht notwendigerweise gegen $f(x)$.
- Die Taylorreihe konvergiert genau für die $x \in A$ gegen $f(x)$, für die das Restglied $R_n(x)$ für $n \to \infty$ gegen 0 konvergiert.

Beispiel 5.36

Sei $f(x) = \sin(x)$ und $x_0 = 0$. Die Ableitungen lauten

$$f'(x) = \cos x,\ f''(x) = -\sin x,\ f'''(x) = -\cos x,\ f^{(4)}(x) = \sin x,\ \text{usw.}$$

Außerdem ist $\sin(0) = 0$ und $\cos(0) = 1$. Damit ergibt sich für die Ableitungen in x_0

$$\left(f^{(n)}(x_0)\right)_{n\in\mathbb{N}} = (0, 1, 0, -1, 0, 1, 0, -1, ...).$$

Wir erhalten für die Taylorreihe

$$\begin{aligned} T[\sin, 0](x) &= \frac{\sin(0)}{0!}x^0 + \frac{\cos(0)}{1!}x^1 + \frac{-\sin(0)}{2!}x^2 + \frac{-\cos(0)}{3!}x^3 + \frac{\sin(0)}{4!}x^4 + ... \\ &= 0 + \frac{1}{1!}x + 0 \cdot x^2 - \frac{1}{3!}x^3 + 0 \cdot x^4 + ... \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}x^{2k+1}. \end{aligned}$$

Wir zeigen nun, dass die Potenzreihe auch gegen die Funktion konvergiert. Dafür betrachten wir das Restglied $R_n(x)$ aus Satz 5.34:

Falls $n = 2m$, also n gerade ist, dann ist

$$|R_{2m}(x)| = \left|(-1)^{m+1}\frac{\cos(y)}{(2m+1)!}x^{2m+1}\right| \leq \frac{|x|^{2m+1}}{(2m+1)!}.$$

Falls $n = 2m-1$, also n ungerade ist, dann ist

$$|R_{2m-1}(x)| = \left|(-1)^{m}\frac{\sin(y)}{(2m)!}x^{2m}\right| \leq \frac{|x|^{2m}}{(2m)!}.$$

Die Schranken gelten für alle y, da $\cos(x)$ und $\sin(x)$ nur Werte im Intervall $[-1, 1]$ annehmen können. Damit gilt $\lim\limits_{n\to\infty} R_n(x) = 0$ für alle x und wir können für beliebige x die Sinusfunktion und ihre Taylorreihe gleichsetzen.

Diese Reihendarstellung des Sinus kennen wir natürlich schon aus Satz 4.63. Auch die Taylorreihen der Exponentialfunktion und des Cosinus sowie die der hyperbolischen Funktionen entsprechen den bekannten Reihendarstellungen. Es lässt sich sogar allgemein zeigen, dass eine Potenzreihendarstellung eindeutig ist, die Taylorreihe also immer mit der Potenzreihendarstellung einer Funktion übereinstimmt.

Der Verlauf der Funktion $\sin(x)$ und der Taylorpolynome $T_n[\sin, 0](x)$ aus dem letzten Beispiel wird in der nachfolgenden Demo gezeigt. Es wird deutlich, dass z.B. mit $T_6[\sin, 0](0)$ eine relativ gute Approximation im Intervall $(-\pi, \pi)$ erreicht wird. Außerhalb des Intervalls wird der Verlauf der Funktion aber nicht mehr gut approximiert. Dies zeigt, dass zur Approximation des periodischen Verlaufs der Sinus-Funktion über mehrere Perioden Taylorpolynome eines deutlich höheren Grades verwendet werden müssen. Bei der praktischen Berechnung kann man sich aber die Periodizität der

Sinus-Funktion zunutze machen, so dass nur Werte aus dem Intervall $(-\pi, \pi)$ berechnet werden müssen.

▶ Demo: Taylor-Approximation des Sinus

Als Nächstes betrachten wir ein etwas spannenderes Beispiel, die Taylorreihe des Logarithmus.

Beispiel 5.37

$f(x) = \ln(x)$ ist für $x = 0$ nicht definiert, daher können wir nicht $x_0 = 0$ als Entwicklungspunkt wählen. Wir könnten nun z.B. $x_0 = 1$ wählen, aber dann haben wir in der Taylorreihe immer die etwas umständlichen Terme $(x-1)^k$. Um dies zu vermeiden, entwickelt man stattdessen die Funktion $f(x) = \ln(x+1)$, da diese für $x = 0$ definiert ist.

Wir entwickeln also die Taylorreihe für $f(x) = \ln(x+1)$ im Entwicklungspunkt 0.

Es gilt

$$f'(x) = \frac{1}{x+1},\ f''(x) = \frac{-1}{(x+1)^2},\ f'''(x) = \frac{2}{(x+1)^3},\ f^{(4)}(x) = \frac{-6}{(x+1)^4}$$

und allgemein

$$f^{(k)}(x) = (-1)^{k+1}\frac{(k-1)!}{(x+1)^k}.$$

Die allgemeine Formel sollte man streng genommen noch z.B. per vollständiger Induktion nachweisen, aber das überlassen wir Ihnen als Übung. In diesem Beispiel ist die allgemeine Formel auch wenig überraschend.

Für $x = 0$ ergibt sich

$$f^{(k)}(0) = (-1)^{k+1}\frac{(k-1)!}{(1)^k} = (-1)^{k+1}(k-1)!.$$

Damit ist die Taylorreihe für $f(x) = \ln(x+1)$

$$T[f,0](x) = \ln(1) + \sum_{k=1}^{\infty}\frac{(-1)^{k+1}(k-1)!}{k!}x^k = \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}x^k.$$

Die Reihe konvergiert nach dem Quotientenkriterium (Satz 3.62) für $|x| < 1$, da

$$\left|\frac{a_{k+1}}{a_k}\right| = \left|\frac{(-1)^{k+2}\cdot x^{k+1}\cdot k}{(-1)^{k+1}\cdot(k+1)\cdot x^k}\right| = \frac{k}{k+1}|x| \overset{k\to\infty}{\longrightarrow} |x| < 1.$$

Für $x = 1$ entsteht die Reihe $\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}$, die nach dem Leibniz-Kriterium (Satz 3.50) konvergiert.

Für das Restglied gilt

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

Falls $x > 0$ gilt $y \in (0,x)$ und damit gilt $\left|\frac{x}{1+y}\right| < 1$ für $x \in (0,1)$ und damit auch $\lim_{n\to\infty} R_n(x) = 0$. Um zu zeigen, dass das Restglied auch für $x \in (-1,0)$ und $x = 1$ gegen 0 konvergiert, benötigen wir eine andere Art der Darstellung des Restgliedes, die

wir hier aber nicht einführen werden. Insofern sei für diese Beweise auf die einschlägigen Lehrbücher verwiesen.

Insgesamt kann man zeigen, dass die Taylorreihe für $x \in (-1, 1]$ gegen die Funktion $\ln(x+1)$ konvergiert. Damit ergibt sich unter Anderem für den Grenzwert der alternierenden harmonischen Reihe

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = \ln(1+1) = \ln(2).$$

▶ Demo: Taylor-Approximation des Logarithmus

Taylorreihen und der Satz von Taylor mit seiner endlichen Summation sind wichtige Hilfsmittel, die insbesondere bei der numerischen Lösung von Gleichungen und Differentialgleichungen eine große praktische Bedeutung haben. Ein wichtiges Beispiel ist das folgende *Newton-Verfahren*.

Beispiel 5.38 (Newton-Verfahren)

Wir hatten in Beispiel 4.32 bereits zwei numerische Verfahren kennengelernt, um Nullstellen einer Funktion zu finden (Regula-Falsi und Bisektion). Diese können mithilfe der Taylor-Approximationen verbessert werden, wodurch sich das sogenannten Newton-Verfahren ergibt. Dabei benutzen wir zunächst eine beliebige (auch beliebig schlechte) Schätzung der Nullstelle x_0 und nutzen diese als Entwicklungspunkt der Taylorreihe. Anschließend nähern wir $f(x)$ über die lineare Approximation $T_1[f, x_0](x)$ an:

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0).$$

Nun bestimmen wir die Nullstelle dieser Näherung:

$$0 = f(x_0) + f'(x_0)(x - x_0) \quad \Leftrightarrow \quad x = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Die Nullstelle der linearen Approximation nehmen wir dann als nächsten Schätzwert x_1 der Nullstelle von f und beginnen mit dem Verfahren von vorne.

Im Gegensatz zum Bisektionsverfahren oder der Regula Falsi benötigen wir hier nur einen Startwert. Das Newton-Verfahren konvergiert für viele Startwerte deutlich schneller als die anderen beiden Verfahren. Die Konvergenz gegen eine Nullstelle ist aber nicht für jeden Startwert garantiert. Das Newton-Verfahren lässt sich auch für mehrdimensionale Funktionen verallgemeinern und wird sehr häufig für Optimierungen eingesetzt.

Übrigens: Für $f(x) = x^2 - a$ ergibt sich mit $f'(x) = 2x$ und $x_0 = a$ die Folge, die wir aus Beispiel 3.21 kennen, mit der die Nullstelle von f, also $\sqrt{a}$, angenähert werden kann:

$$\begin{aligned} x_0 &= a, \\ x_{n+1} &= x_n - \frac{f(x_n)}{f'(x_n)} \\ &= x_n - \frac{x_n^2 - a}{2x_n} \\ &= \frac{1}{2}\left(x_n + \frac{a}{x_n}\right). \end{aligned}$$

Die damals eingeführte Wurzelfolge ist also ein Spezialfall des Newton-Verfahrens.

▶ Demo: Newton-Verfahren

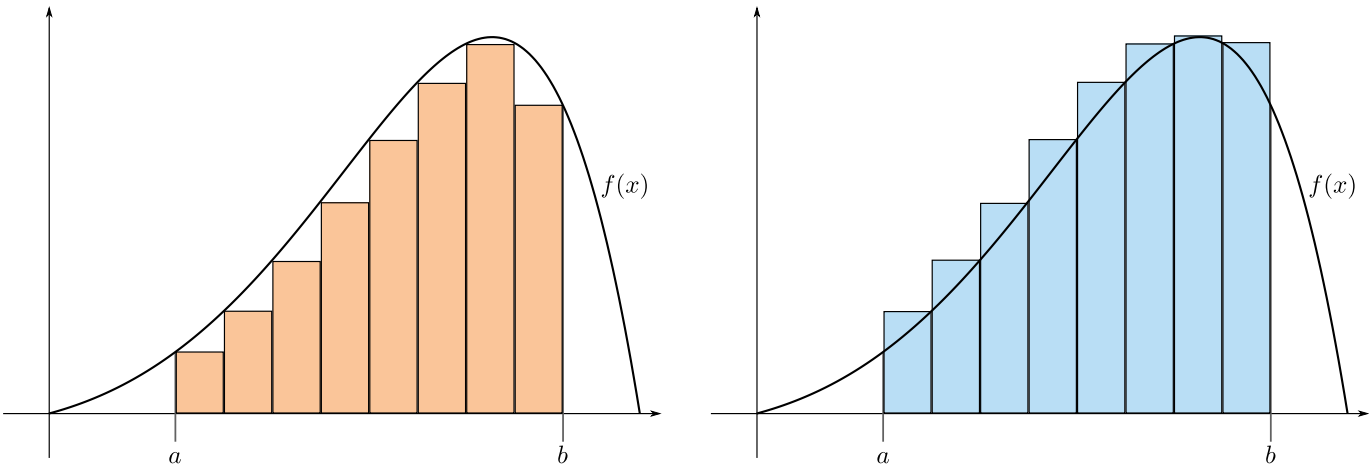
6 Integralrechnung

Neben der Differentiation ist die Integration ein zweites zentrales Thema der Analysis. Historisch entwickelten sich erste Versionen der Integralrechnung bereits im alten Babylonien, in Ägypten, sowie im antiken Griechenland und China. Sie wurde unter anderem zur Landvermessung benötigt, wobei häufig Flächen bestimmt werden mussten, die nicht durch gerade Linien umschlossen wurden. Inhalte geradlinig begrenzter Flächen wie Dreiecke, Rechtecke oder Trapeze waren bekannt. Also versuchte man krummlinig begrenzte Flächen durch Unterteilung der Fläche in geradlinig begrenzte Flächen zu approximieren. Auch hier taucht wieder der Begriff der Approximation auf, der durch eine immer weitergehende Verfeinerung (Grenzwertbetrachtung) schließlich zum *exakten* Ergebnis führt. Sie können dies in der folgenden Demo für die Bestimmung einer Kreisfläche ausprobieren.

▶ Demo: Approximation von Pi

Dies war tatsächlich lange Zeit die gängigste Methode, um immer genauere Approximationen von π zu bestimmen: Man spannte ein regelmäßiges n-Eck um einen Kreis vom Radius 1 und ein weiteres innerhalb des Kreises. Der Wert des Kreisflächeninhalts musste zwischen den Flächeninhalten der beiden n-Ecke liegen. Damals rechneten Mathematiker dies per Hand aus. Für ein Sechseck können Sie dies selbst einmal recht einfach ausprobieren. Der griechische Mathematiker Archimedes bestimmte bereits ca. 250 v. Chr. mit dieser Methode über ein 96-Eck den Wert von π bis auf zwei Nachkommastellen genau. In Europa machte die Bestimmung von π lange Zeit keine weiteren großen Fortschritte, allerdings waren parallel in China bereits um 500 n. Chr. sieben Nachkommastellen von π bekannt. 1430 gelang es dem persischen Astronomen Al-Khashi, den Flächeninhalt eines $3 \cdot 2^{28}$-Ecks zu berechnen und damit 16 Nachkommastellen zu bestimmen. Ludolph van Ceulen erhöhte 1615 auf 35 Stellen mithilfe eines 2^{62}-Ecks, alles per Hand berechnet (zum Vergleich: Ein typischer Computer rechnet mit 8 (single precision) oder 16 (double precision) Nachkommastellen von π). Angeblich soll Ceulen für die Berechnung 30 Jahre seines Lebens investiert haben (dann doch lieber ein Semester in MafI2 etwas über Integrale lernen).

Ganz ähnlich gehen wir nun auch für andere krummlinig begrenzte Flächen vor: Wir approximieren die Fläche, welche von einer Funktion $f(x)$ und dem Intervall $[a, b]$ eingeschlossen wird, durch Rechtecke, die unterhalb der Funktion liegen. Dadurch ergibt sich eine untere Grenze für den Flächeninhalt der gesuchten Fläche. Analog können wir größere Rechtecke um die Funktionsfläche legen und erhalten eine Abschätzung nach oben. Wenn wir immer mehr und schmalere Rechtecke verwenden, sollte sich als Grenzwert die gesuchte Fläche ergeben.



Basiert auf Bild von Svenlx (Sven Laux), CC BY-SA 3.0

Ähnlich wie schon beim Ableitungsbegriff müssen wir uns wieder fragen, ob wir dies überhaupt für jede Funktion so durchführen können. Nachdem Sie bereits ein paar verrückte Funktionen in den letzten Kapiteln kennengelernt haben, wird es Sie nicht verwundern, dass wir auch hier wieder Funktionen definieren können, die nicht integrierbar sind. Allerdings stößt man in der Praxis eher seltener auf solche Funktionen (anders als nicht stetige oder nicht differenzierbare Funktionen). Daher werden wir den Begriff der sogenannten Riemann-Integrierbarkeit zwar formal einführen, beschränken uns dabei aber nur auf die nötigsten Sätze und Definitionen und werden uns dann verstärkt den Techniken des Integrierens zuwenden.

In diesem Kapitel haben wir fünf Hauptziele:

1. Den Integralbegriff sowie die Integrierbarkeit einer Funktion über die bereits erwähnten Rechtecksummen definieren.
2. Beweisen, dass alle stetigen Funktionen integrierbar sind (tatsächlich sind sogar viele unstetige Funktionen integrierbar).
3. Eine Methode herleiten, mit der wir Integrale über Grenzwerte bestimmen können.
4. Die Integration und das Differenzieren (Ableiten) in Beziehung setzen mit dem sogenannten Hauptsatz der Differential- und Integralrechnung (aus der Schule kennen Sie vermutlich bereits den Begriff der Stammfunktion, deren Ableitung die zu integrierende Funktion ist).
5. Mithilfe des Hauptsatzes werden wir Integrationstechniken entwickeln, welche das Integrieren einiger Funktionen vereinfachen.

6.1 Riemann-Integrierbarkeit

Wenden wir uns also unserem ersten Ziel zu. Für eine saubere Definition der oben motivierten Rechtecksummen benötigen wir sogenannte *Zerlegungen* und *Treppenfunktionen*.

Definition 6.1 (Zerlegung)

Gegeben seien ein Intervall $[a, b] \subset \mathbb{R}$ und eine endliche Anzahl von Punkten $x_0, x_1, \dots, x_n$ mit $a = x_0 < x_1 < \dots < x_n = b$.

Dann heißt $Z = (x_0, \dots, x_n)$ eine *Zerlegung* von $[a, b]$ und

$$|Z| := \max \{x_i - x_{i-1} \mid i = 1, \dots, n\}$$

ist das *Feinheitsmaß* der Zerlegung Z. Eine Zerlegung heißt *äquidistant*, wenn die Intervalle $[x_{i-1}, x_i]$ für $i = 1, \dots, n$ alle gleich groß sind.

Definition 6.2 (Treppenfunktion)

Sei $Z = (x_0, \dots, x_n)$ eine Zerlegung des Intervalls $[a, b]$, dann heißt eine Funktion $\varphi : [a, b] \to \mathbb{R}$ Treppenfunktion, wenn sie auf jedem Teilintervall (x_{k-1}, x_k) konstant ist. Wir bezeichnen mit $\mathcal{T}[a, b]$ die Menge aller Treppenfunktionen auf $[a, b]$.

Beispiel 6.3 (Treppenfunktion)

Die Funktion $\varphi : [-1, 5] \to \mathbb{R}$ mit

$$\varphi(x) = \begin{cases} 7, & \text{für } x \in [-1, 0) \\ 2, & \text{für } x \in [0, 0.5] \\ 1, & \text{für } x \in (0.5, 3) \\ 6, & \text{für } x = 3 \\ 5, & \text{für } x \in (3, 5] \end{cases}$$

ist eine Treppenfunktion.

Machen Sie sich bewusst, dass die Treppenstufen nicht gleich lang sein müssen (d.h., die Zerlegung muss nicht äquidistant sein) und dass die Treppenstufen nicht nur steigen oder nur fallen müssen. Außerdem sind die Werte der Treppenfunktion an den Zerlegungspunkten beliebig wählbar. Die im letzten Bild dargestellten Rechtecke können wir uns nun als Flächen unter einer Treppenfunktion vorstellen und das Integral einer Treppenfunktion als die Summe der Rechteckflächen definieren.

Definition 6.4 (Integral für Treppenfunktionen)

Sei $\varphi \in \mathcal{T}[a,b]$ eine Treppenfunktion bezüglich einer Zerlegung $Z = (a = x_0, x_1, \dots, x_n = b)$ und seien c_k die konstanten Funktionsabschnitte von φ, also $\varphi(x) = c_k$ für $x \in (x_{k-1}, x_k)$.

Dann definieren wir das *Integral* von φ auf dem Intervall $[a,b]$ als

$$\int_a^b \varphi(x)\,\mathrm{d}x := \sum_{k=1}^{n} c_k\,(x_k - x_{k-1})\,.$$

Auf der rechten Seite der Integraldefinition ist jeder Summand eine der Rechteckflächen, welche die Treppenfunktion mit der x-Achse einschließt. Dabei ist c_k die Höhe des k-ten Rechtecks und $(x_k - x_{k-1})$ dessen Breite. Negative c_k führen zu einer Verringerung des Integralwerts. Beachten Sie, dass die beliebig wählbaren Werte der Treppenfunktion an den Zerlegungspunkten keinen Einfluss auf den Wert des Integrals haben. Genauso hat das Hinzufügen weiterer Zwischenpunkte in der Zerlegung keine Veränderung des Integralwerts zur Folge (ein Rechteck wird lediglich in mehrere Teilrechtecke gleicher Höhe aufgeteilt).

Beispiel 6.5 (Integral für Treppenfunktionen)

Für die Treppenfunktion $\varphi : [-1,5] \to \mathbb{R}$ mit

$$\varphi(x) = \begin{cases} 7, & \text{für } x \in [-1,0) \\ 2, & \text{für } x \in [0,0.5] \\ 1, & \text{für } x \in (0.5,3) \\ 6, & \text{für } x = 3 \\ 5, & \text{für } x \in (3,5] \end{cases}$$

ist

$$\int_{-1}^{5} \varphi(x)\,\mathrm{d}x = 7\cdot(0-(-1)) + 2\cdot(0.5-0) + 1\cdot(3-0.5) + 5\cdot(5-3) = 20.5.$$

Wie bereits angedeutet, werden wir Treppenfunktionen nutzen, um eine Funktion (und deren Integral) von oben und unten zu begrenzen. Dafür benötigen wir folgenden einfachen Satz.

Satz 6.6 (Monotonie des Treppenfunktionsintegrals)

Seien $\varphi, \psi \in \mathcal{T}[a,b]$ zwei Treppenfunktionen. Dann gilt

$$\forall x \in [a,b] : \varphi(x) \leq \psi(x) \quad \Rightarrow \quad \int_a^b \varphi(x)\,\mathrm{d}x \leq \int_a^b \psi(x)\,\mathrm{d}x.$$

▼ Beweis als Übung

Nun können wir die Treppenfunktionsintegrale nutzen, um damit eine untere und obere Grenze für das eigentlich gesuchte Funktionsintegral zu definieren, indem wir die Funktion zwischen einer größeren und einer kleineren Treppenfunktion einschließen. Die Integrale solcher Treppenfunktionen definieren wir als sogenannte Ober- bzw. Untersummen.

Definition 6.7 (Obersumme, Untersumme)

Sei $f : [a,b] \to \mathbb{R}$ eine beliebige beschränkte Funktion und $Z = (x_0, x_1, \ldots, x_n)$ eine Zerlegung von $[a,b]$. Seien außerdem $\overline{\varphi}, \underline{\varphi} \in \mathcal{T}[a,b]$ definiert als

$$\overline{\varphi}(x) := \overline{c}_k \quad \text{für} \quad x \in [x_{k-1}, x_k) \quad \text{mit} \quad \overline{c}_k := \sup\{f(x) \mid x_{k-1} \leq x < x_k\}$$

bzw.

$$\underline{\varphi}(x) := \underline{c}_k \quad \text{für} \quad x \in [x_{k-1}, x_k) \quad \text{mit} \quad \underline{c}_k := \inf\{f(x) \mid x_{k-1} \leq x < x_k\}$$

und $\overline{\varphi}(b) := \underline{\varphi}(b) := f(b)$.

Dann bezeichnen wir

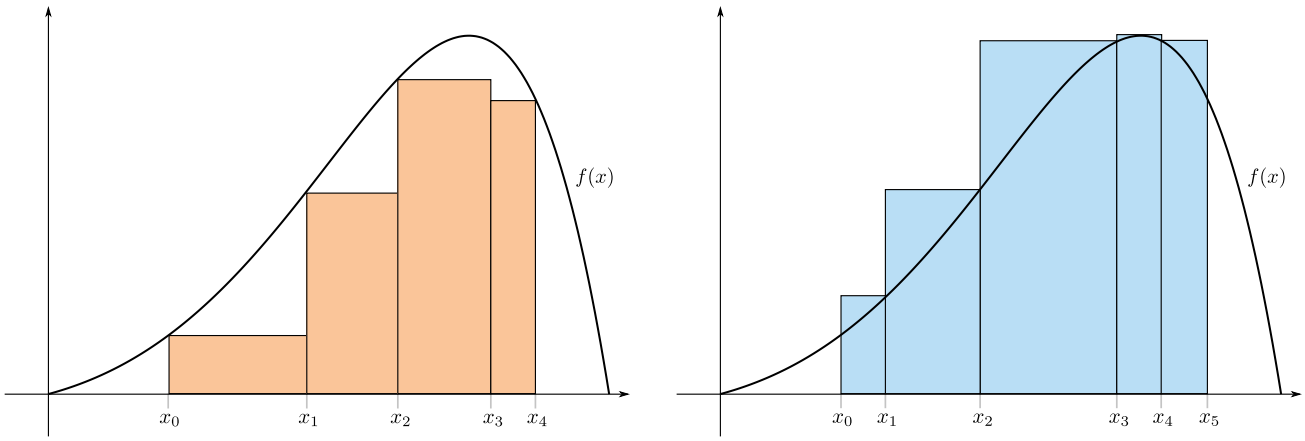
$$\overline{S}(Z,f) := \int_a^b \overline{\varphi}(x)\,\mathrm{d}x = \sum_{k=1}^{n} \overline{c}_k (x_k - x_{k-1})$$

als eine *Obersumme* von f und

$$\underline{S}(Z,f) := \int_a^b \underline{\varphi}(x)\,\mathrm{d}x = \sum_{k=1}^{n} \underline{c}_k (x_k - x_{k-1})$$

als eine *Untersumme* von f.

Im folgenden Bild zeigen die linke/rechte Abbildung jeweils eine Ober- und eine Untersumme der Funktion.



Basiert auf Bild von Svenlx (Sven Laux), CC BY-SA 3.0

Da per Konstruktion $\underline{\varphi}(x) \leq f(x) \leq \overline{\varphi}(x)$ gilt, ist nach Satz 6.6 auch jede Untersumme kleiner oder gleich jeder Obersumme. Wenn wir die Menge aller Obersummen betrachten, ist diese also nach unten beschränkt und besitzt daher ein Infimum. Genauso ist die Menge der Untersummen nach oben beschränkt und besitzt ein Supremum. Dieses Infimum bzw. Supremum bezeichnen wir als das sogenannte Ober- bzw. Unterintegral von f:

Definition 6.8 (Oberintegral, Unterintegral)

Sei $f : [a, b] \to \mathbb{R}$ eine beliebige beschränkte Funktion. Dann bezeichnen wir

$$\overline{\int_a^b} f(x)\,\mathrm{d}x := \inf\left\{\overline{S}(Z, f) \mid Z \text{ ist Zerlegung von } [a, b]\right\}$$

als das *Oberintegral* von f und

$$\underline{\int_a^b} f(x)\,\mathrm{d}x := \sup\left\{\underline{S}(Z, f) \mid Z \text{ ist Zerlegung von } [a, b]\right\}$$

als das *Unterintegral* von f.

Anschaulich wird klar, dass die gesuchte Fläche unter einer Funktion genau ihrem Ober- bzw. Unterintegral entsprechen sollte. Dies ist der Gedanke hinter der Definition der Riemann-Integrierbarkeit einer Funktion, womit wir unser erstes Etappenziel erreicht haben.

Definition 6.9 (Riemann-Integrierbarkeit, Integral)

Eine beschränkte Funktion $f : [a, b] \to \mathbb{R}$ heißt *(Riemann-)integrierbar*, wenn

$$\underline{\int_a^b} f(x)\,\mathrm{d}x = \overline{\int_a^b} f(x)\,\mathrm{d}x.$$

In diesem Fall definieren wir das *Integral* von f über $[a, b]$ als

$$\int_a^b f(x)\,\mathrm{d}x := \underline{\int_a^b} f(x)\,\mathrm{d}x.$$

Beachten Sie, dass wir die Integrierbarkeit damit nur für beschränkte Funktionen (zumindest beschränkt auf $[a, b]$) definiert haben, denn anderenfalls wäre die Obersumme nicht definiert. Integrale unbeschränkter Funktionen betrachten wir erst in Kapitel 6.5. Wenn wir bis dahin von einer integrierbaren Funktion sprechen, ist also immer gleichzeitig eine beschränkte Funktion gemeint.

Es sollte noch erwähnt werden, dass das Vorwort "Riemann-" nicht nur zu Ehren des deutschen Mathematikers Georg Friedrich Bernhard Riemann (1826–1866) verwendet wird. Es gibt in der Mathematik verschiedene Definitionen der Integrierbarkeit und es existieren Funktionen, welche bezüglich einer Definition integrierbar sind, aber bezüglich einer anderen nicht. Für unsere Zwecke ist die Riemann-Integrierbarkeit vollkommen ausreichend und wir werden in Zukunft daher meistens vereinfacht von der *Integrierbarkeit* einer Funktion sprechen, obwohl wir eigentlich streng genommen die *Riemann-Integrierbarkeit* meinen.

Übrigens: Das Integralzeichen entstand aus dem Buchstaben "S", da man sich zu dessen Entstehungszeit ein Integral als eine **S**umme vorgestellt hat (ähnlich wie wir jetzt auch). Wenn wir uns $\mathrm{d}x$ erneut als infinitesimale Länge vorstellen (siehe Kapitel 5.1), dann sind die Produkte $f(x)\,\mathrm{d}x$ Flächeninhalte sehr schmaler Rechtecke, aus denen sich die Integralfläche zusammensetzt. Jedes Rechteck hat an der Stelle x die Höhe $f(x)$ und die Breite $\mathrm{d}x$.

Die Integrationsvariable x ist übrigens beliebig wählbar (wie jede Funktionsvariable), wir müssen nur aufpassen, dass sie nicht mit einer der Grenzen übereinstimmt. Es gilt also

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b f(y)\,\mathrm{d}y = \int_a^b f(t)\,\mathrm{d}t.$$

Obwohl wir den Integralbegriff nun definiert haben, können wir damit noch nicht viel anfangen, da wir unendlich viele Treppenfunktionen testen müssten, um die Definition für eine bestimmte Funktion zu prüfen. Daher können wir bisher nur für sehr einfache Beispiele die Integrierbarkeit nachweisen oder widerlegen.

Beispiel 6.10

- Jede Treppenfunktion $\varphi : [a, b] \to \mathbb{R}$ ist integrierbar, denn $\int_a^b \varphi(x)\,\mathrm{d}x$ ist sowohl eine Ober- als auch eine Untersumme von φ. Diese Ober- bzw. Untersumme ist gleich dem Ober- bzw Unterintegral, denn keine Untersumme kann kleiner als eine Obersumme werden und umgekehrt (Satz 6.6). Damit ist für Treppenfunktionen das allgemeine Integral identisch zum Integral von Treppenfunktionen aus Definition 6.4.

- Jede konstante Funktion ist auch eine Treppenfunktion und somit integrierbar und es gilt

$$\int_a^b c\,\mathrm{d}x = c\,(b-a).$$

- Die Dirichlet-Funktion $f : [0, 1] \to \mathbb{R}$

$$f(x) = \begin{cases} 1 & \text{falls } x \in \mathbb{Q}, \\ 0 & \text{falls } x \in \mathbb{R} \setminus \mathbb{Q}, \end{cases}$$

  ist nicht integrierbar, da jedes Teilintervall $[x_{k-1}, x_k]$ jeder Zerlegung sowohl rationale als auch irrationale Zahlen enthält. Damit gilt für jede Treppenfunktion einer Obersumme $\overline{\varphi}(x) = 1$ und für jede Treppenfunktion einer Untersumme $\underline{\varphi}(x) = 0$. Damit folgt insgesamt

$$\overline{\int_a^b} f(x)\,\mathrm{d}x = 1(b-a) \neq 0 = \underline{\int_a^b} f(x)\,\mathrm{d}x.$$

Ähnlich wie bereits im Fall der Stetigkeit (vgl. Satz 4.25) können wir eine alternative Integrierbarkeitsbedingung durch ein ε-Kriterium herleiten, mit dessen Hilfe wir die Integrierbarkeit von Funktionen leichter beweisen können.

Satz 6.11 (Einschließung zwischen Treppenfunktionen)

Eine Funktion $f : [a, b] \to \mathbb{R}$ ist genau dann integrierbar, wenn zu jedem $\varepsilon > 0$ eine Obersumme $\overline{S}(Z, f)$ und eine Untersumme $\underline{S}(Z', f)$ existieren mit

$$\overline{S}(Z, f) - \underline{S}(Z', f) \leq \varepsilon.$$

▼ Beweis

⇒-Richtung:

Wenn f integrierbar ist, dann sind Ober- und Unterintegral identisch. Sei I der Wert dieses Ober- und Unterintegrals. Damit existiert nach der Definition des Unterintegrals eine Untersumme mit $I - \underline{S}(Z', f) \leq \varepsilon/2$, da I die kleinste obere Schranke aller Untersummen ist (machen Sie sich dies noch einmal in Ruhe klar). Genauso existiert nach der Definition des Oberintegrals eine Obersumme mit $\overline{S}(Z, f) - I \leq \varepsilon/2$.

Daraus folgt

$$\overline{S}(Z, f) - \underline{S}(Z', f) = \overline{S}(Z, f) - I + I - \underline{S}(Z', f) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$$

und damit folgt die Behauptung.

⇐-Richtung:

Wenn umgekehrt das ε-Kriterium gilt, dann können wir durch Widerspruch folgern, dass f integrierbar sein muss: Angenommen, f ist nicht integrierbar, dann ist der Oberintegralwert O ungleich dem Unterintegralwert U. Sei $\varepsilon = (O - U)/2$, dann existieren auch für dieses ε Ober- und Untersummen mit

$$\overline{S}(Z, f) - \underline{S}(Z', f) \leq \varepsilon.$$

Da O als das Infimum aller Obersummen definiert ist, muss gelten $O \leq \overline{S}(Z, f)$. Es folgt damit

$$\overline{S}(Z, f) - \underline{S}(Z', f) \leq \varepsilon < 2\varepsilon = O - U \leq \overline{S}(Z, f) - U,$$

also insgesamt $U < \underline{S}(Z', f)$. Dies steht aber im Widerspruch dazu, dass U das Supremum aller Untersummen ist (↯).

□

Ausgestattet mit dem handlicheren ε-Kriterium können wir uns nun unserem zweiten Ziel zuwenden und die Integrierbarkeit aller stetigen Funktionen beweisen.

Satz 6.12 (Stetige Funktionen sind integrierbar)

Sei $f : A \to \mathbb{R}$ eine auf $[a, b] \subseteq A$ stetige Funktion. Dann ist f auf $[a, b]$ integrierbar.

▼ Beweis

Wir hatten bereits in Satz 4.39 bewiesen, dass eine stetige Funktion auf einem kompakten Intervall sogar gleichmäßig stetig ist. Für alle $\varepsilon > 0$ existiert also wegen der gleichmäßigen Stetigkeit von f ein $\delta > 0$, sodass

$$\forall x, y \in [a, b] \text{ mit } |x - y| < \delta : |f(x) - f(y)| < \varepsilon.$$

Wir wählen $n \in \mathbb{N}$ so groß, dass gilt $\frac{b-a}{n} < \delta$. Nun definieren wir eine Zerlegung $Z = (x_0, x_1, ..., x_n)$ von $[a, b]$ mit

$$x_k := a + k\,\frac{b-a}{n}.$$

Auf jedem Teilintervall der Zerlegung existiert nach Satz 4.36 ein minimaler und ein maximaler Funktionswert von f. Seien s_k die Stellen, an denen f die Minima auf $[x_{k-1}, x_k]$ annimmt, und t_k die Stellen der Maxima. Die Zerlegung definiert eine Ober- bzw. Untersumme, wobei die konstanten Funktionswerte der Treppenfunktionen jeweils den Minima und Maxima auf den Teilintervallen entsprechen:

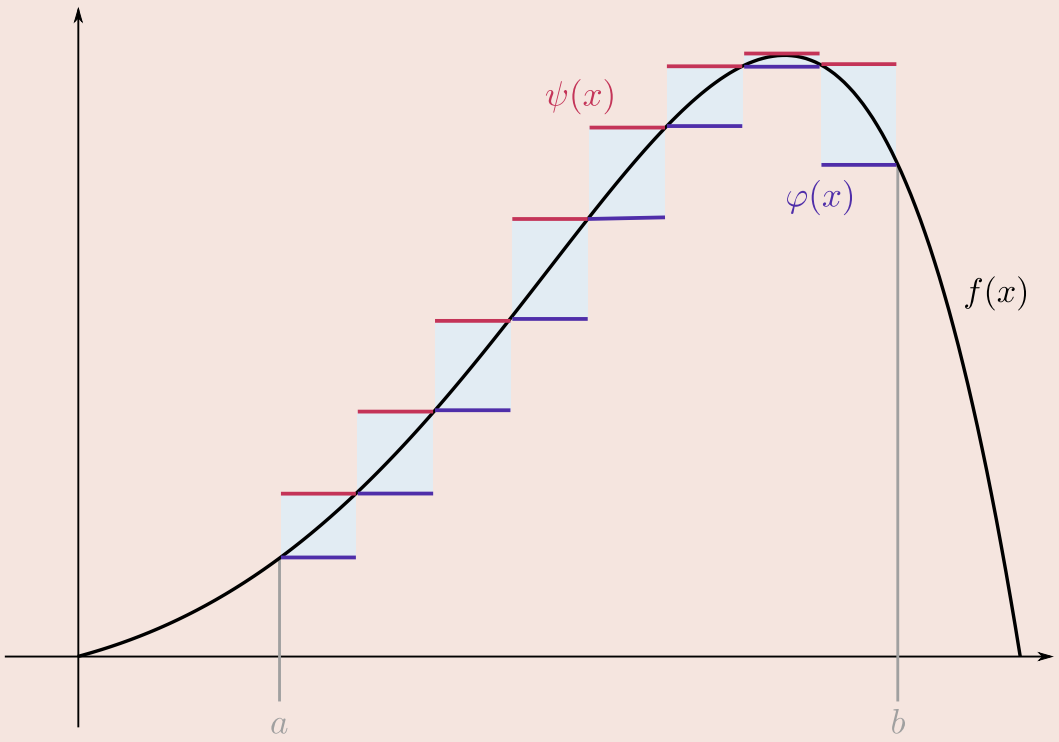
$$\underline{S}(Z, f) = \sum_{k=1}^{n} f(s_k)(x_k - x_{k-1}) = \sum_{k=1}^{n} f(s_k)\frac{b-a}{n},$$
$$\overline{S}(Z, f) = \sum_{k=1}^{n} f(t_k)(x_k - x_{k-1}) = \sum_{k=1}^{n} f(t_k)\frac{b-a}{n}.$$

Wegen $x_k - x_{k-1} = \frac{b-a}{n} < \delta$ gilt auch $t_k - s_k < \delta$. Wegen der gleichmäßigen Stetigkeit von f gilt dann $f(t_k) - f(s_k) < \varepsilon$. Es folgt insgesamt

$$\begin{aligned}\overline{S}(Z, f) - \underline{S}(Z, f) &= \frac{b-a}{n} \sum_{k=1}^{n} (f(t_k) - f(s_k)) \\ &< \frac{b-a}{n}\, n\, \varepsilon \\ &= \varepsilon(b-a)\end{aligned}$$

und somit nach Satz 6.11 die Integrierbarkeit von f.

Eine graphische Darstellung der Konstruktion sehen Sie in der untenstehenden Abbildung.



Für den Beweis haben wir gezeigt, dass die blaue Fläche beliebig klein werden kann. Die blauen Rechtecke haben nach unserer Konstruktion jeweils eine Breite von $\frac{b-a}{n}$ und eine Höhe von maximal ε (aufgrund der gleichmäßigen Stetigkeit). Es gibt n Rechtecke, also ergibt sich eine Gesamtfläche von $(b-a)\varepsilon$. Da wir ε beliebig klein wählen können und $(b-a)$ konstant ist, wird auch die Fläche beliebig klein.

□

Wie bereits erwähnt, muss eine Funktion aber nicht zwangsweise stetig sein, um integrierbar zu sein. Wir zeigen zum Beispiel im folgenden Satz, dass monotone Funktionen unabhängig von ihrer Stetigkeit immer integrierbar sind.

Satz 6.13 (Monotone Funktionen sind integrierbar)

Sei $f : A \to \mathbb{R}$ eine auf $[a, b] \subseteq A$ monotone Funktion. Dann ist f auf $[a, b]$ integrierbar.

▼ Beweis

Für den Beweis nehmen wir an, dass f auf $[a, b]$ monoton wächst (der monoton fallende Fall lässt sich analog beweisen).

Wir nutzen erneut die Zerlegung $Z = (x_0, x_1, ..., x_n)$ von $[a, b]$ mit

$$x_k := a + k \cdot \frac{b-a}{n}$$

und die zugehörigen Ober- und Untersummen. Da f monoton wächst, entsprechen die Treppenfunktionswerte der Untersumme jeweils den Funktionswerten von $f(x)$ an den linken Intervallgrenzen und die der Obersumme den rechten Teilintervallgrenzen.

Für die Differenz gilt dann:

$$\begin{aligned}
\overline{S}(Z, f) - \underline{S}(Z, f) &= \sum_{k=1}^{n} f(x_k)\,(x_k - x_{k-1}) - \sum_{k=1}^{n} f(x_{k-1})\,(x_k - x_{k-1}) \\
&= \sum_{k=1}^{n} \left(f(x_k) - f(x_{k-1})\right)(x_k - x_{k-1}) \\
&= \frac{b-a}{n} \sum_{k=1}^{n} \left(f(x_k) - f(x_{k-1})\right) \\
&= \frac{b-a}{n} \left(f(b) - f(a)\right).
\end{aligned}$$

Für ein genügend großes n wird die Integraldifferenz der beiden Treppenfunktionen also beliebig klein. Damit folgt die Integrierbarkeit von f auf $[a, b]$ nach Satz 6.11.

□

Wenn wir also eine unstetige Funktion in monotone Teilintervalle zerlegen können, dann ist f auch auf jedem Teilintervall integrierbar (wir werden später noch zeigen, dass daraus auch die Integrierbarkeit für das gesamte Intervall folgt).

Beispiel 6.14

Die Funktion $f : [0, 2] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} x, & \text{für } x \in [0, 1) \\ -x + 4, & \text{für } x \in [1, 2] \end{cases}$$

ist nicht stetig in $x = 1$. Sie ist aber monoton wachsend für $x \in [0, 1]$ und monoton fallend für $x \in [1, 2]$. Daher ist sie auf beiden Intervallen integrierbar.

Im Rahmen dieser Vorlesung reichen uns die angeführten Sätze für die Integrierbarkeit aus. Wir haben versucht, diesen Abschnitt auf ein Minimum zu begrenzen, da die Integrierbarkeit wie bereits erwähnt für die allermeisten praxisrelevanten Funktionen auf ihrem Definitionsbereich gegeben ist und daher kaum überprüft werden muss. Kritisch wird die Integrierbarkeit immer in Fällen, in denen die Funktion auf dem Integrationsintervall unbeschränkt ist, oder am Rand des Intervalls nicht definiert ist. Solche Fälle betrachten wir in Kapitel 6.5.

6.2 Integralbestimmung durch Grenzwerte

Dieser Abschnitt ist besonders für uns als Informatiker*innen interessant, da wir hier die Grundlage für die numerische Integralbestimmung legen. Wir werden noch feststellen, dass manche sehr bedeutsame Integraltypen gar nicht analytisch lösbar sind, weswegen numerische Lösungsverfahren in solchen Fällen die einzige Möglichkeit zur Integralbestimmung darstellen. In solchen Fällen wird die zu integrierende Funktion an verschiedenen Stellen ausgewertet (der Funktionswert bestimmt) und der Integralwert über einfache Flächen (wie z.B. Rechtecke oder Trapeze) approximiert. Für eine Approximationslösung will man natürlich wissen, unter welchen Kriterien diese gegen die zu approximierende Größe, also in diesem Fall den tatsächlichen Integralwert, konvergiert. Diesem Thema widmen wir uns nun.

In der folgenden Demo können Sie verschiedene Unter- und Obersummen für eine Funktion erzeugen.

▶ Demo: Unter- und Obersummen

Wir vermuten, dass wir den Wert des Integrals bestimmen können, wenn wir die Zerlegung immer feiner wählen, das Feinheitsmaß der Zerlegung also gegen Null konvergiert. Um dies zu beweisen, definieren wir uns zunächst, was wir mit einer *Verfeinerung* einer Zerlegung meinen.

Definition 6.15 (Verfeinerung, Überlagerung)

Seien Z, Z' Zerlegungen beliebiger Intervalle.

Eine Zerlegung $\tilde{Z}$ wird *Verfeinerung* von Z genannt, wenn $\tilde{Z}$ dasselbe Intervall zerlegt und alle Punkte von Z enthält.

Eine Zerlegung $\hat{Z}$, die genau die Punkte der Zerlegungen Z und Z' enthält, nennen wir *Überlagerung* von Z und Z' und schreiben dafür $\hat{Z} = Z + Z'$.

Nun betrachten wir, was mit Ober-/Untersummen passiert, wenn wir von einer Zerlegung auf eine andere wechseln.

Satz 6.16 (Zerlegungswechsel)

Sei $f(x)$ auf $[a,b]$ beschränkt mit $|f(x)| \leq K$ und sei Z eine Zerlegung von $[a,b]$ mit Feinheitsmaß $|Z|$. Die Zerlegung $\tilde{Z}$ von $[a,b]$ entstehe aus Z durch Hinzunahme eines zusätzlichen Punktes. Dann gilt:

a. $\underline{S}(Z,f) \leq \underline{S}(\tilde{Z},f) \leq \underline{S}(Z,f) + 2K\,|Z|$
b. $\overline{S}(Z,f) \geq \overline{S}(\tilde{Z},f) \geq \overline{S}(Z,f) - 2K\,|Z|$

▼ Beweis

Wir zeigen den Fall **(a)**. Den Fall **(b)** beweist man analog.

Sei $Z = (x_0, \dots, x_n)$ und sei $\tilde{x} \in (x_{k-1}, x_k)$ der zusätzlich eingefügte Punkt für ein beliebiges $k \in \{1, \dots, n\}$, sodass gilt

$$\tilde{Z} = (x_0, \dots, x_{k-1}, \tilde{x}, x_k, \dots, x_n)\,.$$

Seien außerdem

$$\begin{aligned} m_k &= \inf\{f(x) \mid x \in [x_{k-1}, x_k]\}, \\ \tilde{m}_1 &= \inf\{f(x) \mid x \in [x_{k-1}, \tilde{x}]\} \geq m_k, \\ \tilde{m}_2 &= \inf\{f(x) \mid x \in [\tilde{x}, x_k]\} \geq m_k \end{aligned}$$

die von der Verfeinerung betroffenen Funktionswerte der Treppenfunktionen der zu Z und $\tilde{Z}$ gehörigen Untersummen. Dann folgt:

$$\underline{S}(\tilde{Z},f) \;=\; \underline{S}(Z,f) - (x_k - x_{k-1})\,m_k + (\tilde{x} - x_{k-1})\,\tilde{m}_1 + (x_k - \tilde{x})\,\tilde{m}_2$$

Dies ist äquivalent zu

$$\begin{aligned} \underline{S}(\tilde{Z},f) - \underline{S}(Z,f) &= (\tilde{x} - x_{k-1})\,\tilde{m}_1 + (x_k - \tilde{x})\,\tilde{m}_2 - (x_k - x_{k-1})\,m_k \\ &= \tilde{x}\tilde{m}_1 - x_{k-1}\tilde{m}_1 + x_k\tilde{m}_2 - \tilde{x}\tilde{m}_2 - x_k m_k + x_{k-1}m_k \\ &= \tilde{x}\tilde{m}_1 - x_{k-1}\tilde{m}_1 + x_k\tilde{m}_2 - \tilde{x}\tilde{m}_2 - x_k m_k + x_{k-1}m_k \\ &\quad \underbrace{+\tilde{x}m_k - \tilde{x}m_k}_{\text{0 addiert}} \\ &= \underbrace{(\tilde{x} - x_{k-1})}_{\geq 0}\,\overbrace{(\tilde{m}_1 - m_k)}^{\leq 2K}_{\geq 0} + \underbrace{(x_k - \tilde{x})}_{\geq 0}\,\overbrace{(\tilde{m}_2 - m_k)}^{\leq 2K}_{\geq 0} \end{aligned}$$

Es folgt also zum Einen

$$\underline{S}(\tilde{Z},f) - \underline{S}(Z,f) \geq 0,$$

womit der linke Teil der Behauptung gilt. Zum Anderen folgt

$$\begin{aligned}\underline{S}(\tilde{Z}, f) - \underline{S}(Z, f) &\leq 2K\,(\tilde{x} - x_{k-1}) + 2K\,(x_k - \tilde{x}) \\ &= 2K\,\underbrace{(x_k - x_{k-1})}_{\leq |Z|} \\ &\leq 2K\,|Z|\,,\end{aligned}$$

womit der rechte Teil der Behauptung gilt.

□

Damit haben wir gezeigt, dass wir durch eine Verfeinerung einer Ober-/Untersumme stets dichter an das Ober-/Unterintegral herankommen. Wir können den letzten Satz auch mehrmals hintereinander ausführen, um alle Punkte einer zweiten Zerlegung Z' in die erste Zerlegung Z einzufügen, um so eine Überlagerung $Z + Z'$ der beiden Zerlegungen zu erzeugen. Dabei ergibt sich dann beispielsweise für die Untersumme

$$\underline{S}(Z, f) \;\leq\; \underline{S}(Z + Z', f) \;\leq\; \underline{S}(Z, f) + 2pK\,|Z|\,,$$

wobei p die Anzahl der eingefügten Punkte angibt. Damit lässt sich nun zeigen, dass wir den Integralwert als Grenzwert immer feinerer Ober- und Untersummen bestimmen können.

Satz 6.17 (Integralwertbestimmung)

Sei $f : [a, b] \to \mathbb{R}$ eine beschränkte Funktion und (Z_n) eine Folge von Zerlegungen mit

$$\lim_{n\to\infty} |Z_n| = 0.$$

Dann gilt

a. $\lim_{n\to\infty} \underline{S}(Z_n, f) = \underline{\int_a^b} f(x)\,\mathrm{d}x,$

b. $\lim_{n\to\infty} \overline{S}(Z_n, f) = \overline{\int_a^b} f(x)\,\mathrm{d}x.$

▼ Beweis

Wir führen den Beweis wieder nur für Fall **(a)**. Fall **(b)** wird analog bewiesen.

Seien $I := \underline{\int_a^b} f(x)\,\mathrm{d}x$ und $S := \lim_{n\to\infty} \underline{S}(Z_n, f)$. Wir wollen zeigen, dass $I = S$ gilt.

Da I das Supremum aller Untersummen ist, gibt es für alle $\varepsilon > 0$ eine Zerlegung Z_ε, sodass für die Untersumme gilt

$$0 \le I - \underline{S}(Z_\varepsilon, f) < \varepsilon. \qquad (*)$$

(Ansonsten wäre $I - \varepsilon$ größer als alle Untersummen, also eine obere Schranke, was im Widerspruch zur Supremumseigenschaft von I steht.)

Nach Satz 6.16 gilt

$$\underline{S}(Z_n, f) \;\le\; \underline{S}(Z_n + Z_\varepsilon, f) \;\le\; \underline{S}(Z_n, f) + 2pK\,|Z_n|\,,$$

wobei p die Anzahl der in Z_n einzufügenden Zwischenpunkte für die Überlagerung mit Z_ε ist.

Es gilt

$$\lim_{n\to\infty} \left(\underline{S}(Z_n, f) + 2pK\,|Z_n|\right) = S + 2pK \cdot 0 = S.$$

Somit folgt nach dem Sandwich-Theorem (Satz 3.23), dass auch gilt

$$\lim_{n\to\infty} \underline{S}(Z_n + Z_\varepsilon, f) = S.$$

Außerdem ist

$$\underline{S}(Z_\varepsilon, f) \;\le\; \underline{S}(Z_n + Z_\varepsilon, f) \;\le\; I,$$

weshalb zusammen mit $(*)$ folgt:

$$0 \leq I - \underline{S}(Z_n + Z_\varepsilon, f) \; < \; \varepsilon.$$

Da dies für ein beliebiges $\varepsilon > 0$ gilt, gilt es auch für $\varepsilon = \frac{1}{n}$. Damit folgt nach dem Sandwich-Theorem

$$0 \; = \; I - \lim_{n \to \infty} \underline{S}(Z_n + Z_\varepsilon, f) \; = \; I - S$$

und damit die Behauptung $I = S$.

□

Mit dem letzten Satz haben wir nun unser nächstes Etappenziel erreicht. Jetzt können wir über Satz 6.12 und Satz 6.13 die Integrierbarkeit einer Funktion sicherstellen, womit laut Definition der Integrierbarkeit das Ober- und Unterintegral identisch sind. Anschließend können wir mit Satz 6.17 den Wert des Integrals bestimmen.

Beispiel 6.18 (Integralbestimmung durch Grenzwerte)

Wir nutzen hier stets die Zerlegungsfolge Z_n mit Teilpunkten $x_{n,k} = a + k\frac{b-a}{n}$, womit für das Feinheitsmaß $|Z_n| = x_{n,k} - x_{n,k-1} = \frac{b-a}{n}$ gilt.

1. Wir wollen $f(x) = x$ auf $[0, b]$ integrieren (also $a = 0$). f ist stetig und monoton, also auf jeden Fall integrierbar (Satz 6.12 oder Satz 6.13). Wir bestimmen die Obersumme. Da f monoton wächst, wird das Supremum im Intervall $[x_{k-1}, x_k]$ stets an der rechten Intervallgrenze x_k angenommen. Die Obersumme ergibt:

$$\begin{aligned}
\overline{S}(Z_n, f) &= \sum_{k=1}^{n} f(x_{n,k})\,(x_{n,k} - x_{n,k-1}) \\
&= \sum_{k=1}^{n} \left(0 + k\frac{b-0}{n}\right) \frac{b-0}{n} \\
&= \frac{b^2}{n^2} \sum_{k=1}^{n} k \\
&= \frac{b^2}{n^2} \frac{n(n+1)}{2} \\
&= \frac{n^2+n}{n^2} \frac{b^2}{2} \\
&\overset{n\to\infty}{\to} \frac{b^2}{2}.
\end{aligned}$$

   Also gilt $\int_0^b x \,\mathrm{d}x = \frac{b^2}{2}$.

2. Wir wollen $f(x) = e^x$ auf $[0, b]$ integrieren. f ist stetig und monoton, also auf jeden Fall integrierbar (Satz 6.12 oder Satz 6.13). Wir bestimmen die Obersumme. Da f monoton wächst, befindet sich das Supremum stets an der rechten Intervallgrenze.

$$
\begin{aligned}
\overline{S}(Z_n, f) &= \sum_{k=1}^{n} f(x_{n,k})\,(x_{n,k} - x_{n,k-1}) \\
&= \sum_{k=1}^{n} e^{kb/n}\,\frac{b-0}{n} \\
&= \frac{b}{n} \sum_{k=1}^{n} \left(e^{b/n}\right)^k \\
&= \frac{b}{n} \left(\sum_{k=0}^{n} \left(e^{b/n}\right)^k - 1 \right) \\
&= \frac{b}{n} \left(\frac{1 - \left(e^{b/n}\right)^n}{1 - e^{b/n}} - 1 \right) \\
&= \frac{b}{n} \left(\frac{1 - e^b}{1 - e^{b/n}} - 1 \right) \\
&\overset{h:=b/n}{=} h \left(\frac{1 - e^b}{1 - e^h} - 1 \right) \\
&= \left(1 - e^b\right) \frac{h}{1 - e^h} - h.
\end{aligned}
$$

Durch die Substitution $h = b/n$ wird aus $n \to \infty$ der Grenzübergang $h \to 0$ und es folgt

$$
\lim_{h \to 0} \frac{h}{1 - e^h} \overset{\text{l'Hospital}}{=} \lim_{h \to 0} \frac{1}{-e^h} = -1
$$

und damit insgesamt

$$
\int_0^b e^x \,\mathrm{d}x = e^b - 1.
$$

▶ Demo: Integrale als Grenzwert von Ober- und Untersummen

Wenn wir wissen, dass die Funktion integrierbar ist, können wir den letzten Satz noch weiter vereinfachen, da wir dann einen beliebigen Zwischenpunkt auf jedem Teilintervall der Zerlegung nutzen dürfen.

Satz 6.19 (Riemannsche Zwischensummen)

Sei $f : [a, b] \to \mathbb{R}$ eine integrierbare Funktion und (Z_n) eine Zerlegungsfolge mit $\lim\limits_{n\to\infty} |Z_n| = 0$. Für jede Zerlegung $Z_n = (x_0, x_1, ..., x_m)$ sei $\varphi_n \in \mathcal{T}[a, b]$ eine Treppenfunktion mit

$$\varphi_n(x) := f(\xi_k) \text{ für } x \in [x_{k-1}, x_k) \text{ mit beliebigem } \xi_k \in [x_{k-1}, x_k]$$

und $\varphi_n(b) = f(b)$. Dann konvergiert die Folge (S_n) der *Riemannschen Zwischensummen*

$$S_n := \int_a^b \varphi_n(x)\,\mathrm{d}x \;=\; \sum_{k=1}^{n} f(\xi_k)\,(x_k - x_{k-1})$$

gegen das Integral:

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} S_n.$$

▼ Beweis

Hier nur als Beweisskizze, formulieren Sie den Beweis gerne zur Übung in Ruhe aus:

$\varphi(x)$ liegt stets zwischen den zwei Treppenfunktionen, welche die Unter- bzw. Obersumme zur Zerlegung Z_n definieren (machen Sie sich klar, warum). Die Folge der Obersummen und Untersummen konvergiert nach Satz 6.17 gegen das Ober- und Unterintegral, deren Werte aufgrund der Integrierbarkeit von f nach Definition 6.9 übereinstimmen. Aus dem Sandwich-Theorem (Satz 3.23) folgt dann die Satzaussage.

□

Wie bereits in der Einleitung zu diesem Kapitel erwähnt, werden tatsächlich in der Praxis viele Integrale auf ähnliche Art berechnet, indem diese in Rechtecksummen aufgeteilt werden. Dadurch muss jetzt nur noch die Funktion an viele Stellen ausgewertet werden, und wenn wir dies an immer mehr Stellen der Funktion durchführen, wird das Integral immer genauer approximiert (Feinheit nimmt immer stärker zu). In manchen Fällen werden die Punkte, an denen die Funktion ausgewertet wird, sogar zufällig ausgewählt, dann spricht man von einer sogenannten *Monte-Carlo-Integration*. Dies wird besonders häufig für mehrdimensionale Integrale angewendet, bei denen nicht nur über ein Intervall, sondern über eine ganze Fläche, oder ein Volumen, oder noch höherdimensionale Gebiete integriert werden muss. Ein typisches Beispiel aus der Computergrafik sind Raytracing-Verfahren, bei denen die Beleuchtung an einem Punkt im Raum berechnet werden muss. Hier muss theoretisch über jede mögliche Richtung, aus der Licht auf den betrachteten Punkt treffen könnte, der entsprechende Lichtanteil integriert werden. Dies wäre analytisch gar nicht möglich, allerdings kann das Integral über eine endliche Auswahl von Lichtstrahlen (Funktionsauswertungen) approximiert werden.

Mithilfe von Satz 6.17 können wir nun außerdem recht einfach die wichtigsten Eigenschaften von Integralen beweisen.

Satz 6.20 (Integraleigenschaften 1)

Seien f und g zwei auf $[a,b]$ integrierbare Funktionen und $\lambda, \mu \in \mathbb{R}$. Dann gilt:

a. **Linearität:** Auch $\lambda f + \mu g$ ist integrierbar und es gilt

$$\int_a^b \lambda f(x) + \mu g(x)\,\mathrm{d}x = \lambda \int_a^b f(x)\,\mathrm{d}x + \mu \int_a^b g(x)\,\mathrm{d}x.$$

b. **Monotonie:** Wenn $f(x) \leq g(x)\ \forall x \in [a,b]$, dann gilt

$$\int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x.$$

c. **Beschränktheit:** Auch $|f(x)|$ ist integrierbar und es gilt

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leq \int_a^b |f(x)|\ \mathrm{d}x.$$

▼ Beweis

Die Eigenschaften folgen im Wesentlichen aus den Eigenschaften von Folgen, da wir nun Ober- und Unterintegral über Folgen von Ober- und Untersummen darstellen können.

a. Sei Z_n eine Zerlegungsfolge von $[a,b]$ mit $|Z_n| \to 0$. Machen Sie sich zur Übung klar, dass mit jeder Zerlegung Z für die Ober- bzw. Untersumme gilt

$$\lambda \underline{S}(Z,f) + \mu \underline{S}(Z,g) = \underline{S}(Z, \lambda f + \mu g).$$

Da f und g integrierbar sind, gilt

$$\lim_{n\to\infty} \underline{S}(Z_n,f) = \lim_{n\to\infty} \overline{S}(Z_n,f) = I_f,$$
$$\lim_{n\to\infty} \underline{S}(Z_n,g) = \lim_{n\to\infty} \overline{S}(Z_n,g) = I_g.$$

Es folgt dann aus den Kombinationsregeln für Folgengrenzwerte Satz 3.18

$$\lim_{n\to\infty} \left(\underline{S}(Z_n,\lambda f) + \underline{S}(Z_n,\mu g)\right) = \lambda \lim_{n\to\infty} \underline{S}(Z_n,f) + \mu \lim_{n\to\infty} \underline{S}(Z_n,g) = \lambda I_f + \mu I_g$$
$$\lim_{n\to\infty} \left(\overline{S}(Z_n,\lambda f) + \overline{S}(Z_n,\mu g)\right) = \lambda \lim_{n\to\infty} \overline{S}(Z_n,f) + \mu \lim_{n\to\infty} \overline{S}(Z_n,g) = \lambda I_f + \mu I_g.$$

Daraus folgt die Integrierbarkeit und die angegebene Gleichung.

b. Im Fall der Gleichheit ist die Aussage trivial. Für den $<$-Fall sei Z_n eine Zerlegungsfolge von $[a,b]$ mit $|Z_n| \to 0$. Dann ist das Infimum/Supremum von f auf jedem Teilintervall kleiner als das Infimum/Supremum von g. Es gilt also für Ober/Untersummen:

$$\underline{S}(Z_n,f) \ \leq\ \underline{S}(Z_n,g) \quad \text{und} \quad \overline{S}(Z_n,f) \ \leq\ \overline{S}(Z_n,g).$$

Damit gilt die Ungleichung nach Satz 3.22 auch für die Grenzwerte der Folgen:

$$\int_a^b f(x)\,\mathrm{d}x \;\leq\; \int_a^b g(x)\,\mathrm{d}x.$$

c. Den Beweis der Integrierbarkeit von $|f(x)|$ lassen wir Ihnen zu Übungszwecken. Es gilt $f(x) \leq |f(x)|$ sowie $-f(x) \leq |f(x)|$. Aus **(b)** folgt dann sofort die Behauptung.

□

Satz 6.21 (Integraleigenschaften 2)

Eine Funktion $f : [a, b] \to \mathbb{R}$ ist genau dann integrierbar auf $[a, b]$, wenn sie auf $[a, c]$ und $[c, b]$ integrierbar ist für ein $c \in (a, b)$. Es gilt außerdem

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x.$$

▼ Beweis

Sei $\tilde{Z}_n$ eine Folge von Zerlegungen von $[a, c]$ und $\hat{Z}_n$ eine Folge von Zerlegungen von $[c, b]$ mit $\left|\tilde{Z}_n\right| \to 0$ und $\left|\hat{Z}_n\right| \to 0$ für $n \to \infty$. Dann ist $Z_n := \tilde{Z}_n + \hat{Z}_n$ für jedes n eine Zerlegung von $[a, b]$ und es gilt $|Z_n| \leq \max\left\{\left|\tilde{Z}_n\right|, \left|\hat{Z}_n\right|\right\} \to 0$ für $n \to \infty$. Für alle $n \geq 0$ gilt

$$\begin{aligned}\overline{S}(\tilde{Z}_n, f) + \overline{S}(\hat{Z}_n, f) &= \overline{S}(Z_n, f),\\ \underline{S}(\tilde{Z}_n, f) + \underline{S}(\hat{Z}_n, f) &= \underline{S}(Z_n, f).\end{aligned}$$

Mit einem Grenzübergang und Satz 3.18 folgt Teil 2 der Behauptung. Es fehlt also noch der genau-dann-wenn-Zusammenhang für die Integrierbarkeit.

⇐-Richtung:
Wenn f auf beiden Abschnitten integrierbar ist, dann konvergieren Ober- und Untersummen auf beiden Teilintervallen für die Zerlegungen $\tilde{Z}_n$ und $\hat{Z}_n$ jeweils gegen denselben Grenzwert. Das heißt, die kombinierte Folge aus Ober- und Untersummen ist eine Cauchy-Folge und es gibt zu $\varepsilon/2 > 0$ ein n_0, sodass für $m, n \geq n_0$ gilt

$$\overline{S}(\tilde{Z}_n, f) - \underline{S}(\tilde{Z}_m, f) \leq \frac{\varepsilon}{2} \quad \text{bzw.} \quad \overline{S}(\hat{Z}_n, f) - \underline{S}(\hat{Z}_m, f) \leq \frac{\varepsilon}{2}.$$

Nach Addition folgt für die kombinierte Zerlegung Z_n dann

$$\overline{S}(Z_n, f) - \underline{S}(Z_n, f) \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Damit ist die Bedingung für die Integrierbarkeit auf $[a, c]$ nach Satz 6.11 erfüllt.

⇒-Richtung:
Sei f auf $[a, b]$ integrierbar, dann gibt es nach Satz 6.11 für alle $\varepsilon > 0$ Unter- und Obersummen mit

$$\overline{S}(Z, f) - \underline{S}(Z', f) \leq \varepsilon.$$

Nach Satz 6.16 wird eine Untersumme größer und eine Obersumme kleiner, wenn wir neue Punkte hinzufügen, daher gilt nach hinzufügen von c zur Zerlegung weiterhin

$$\overline{S}(Z + c, f) - \underline{S}(Z' + c, f) \leq \varepsilon.$$

Die Treppenfunktion der Obersumme ist überall größer als die Treppenfunktion der Untersumme. Das heißt, auf jedem Teilintervall gilt ebenfalls die Abschätzung und somit auch auf $[a, c]$ und $[c, b]$. Damit folgt nach Satz 6.11 die Integrierbarkeit auf diesen Intervallen.

□

Beispiel 6.22

In Beispiel 6.14 konnten wir bisher nur die Integrierbarkeit der zusammengesetzten Funktion $f : [0, 2] \to \mathbb{R}$ mit

$$f(x) = \begin{cases} x, & \text{für } x \in [0, 1) \\ -x + 4, & \text{für } x \in [1, 2] \end{cases}$$

auf den Teilintervallen $[0, 1]$ und $[1, 2]$ zeigen. Aus Satz 6.21 folgt nun, dass f auch auf dem gesamten Definitionsbereich $[0, 2]$ integrierbar ist und mit Satz 6.20 und Beispiel 6.18 folgt dann

$$\int_0^2 f(x)\,\mathrm{d}x = \int_0^1 f(x)\,\mathrm{d}x + \int_1^2 f(x)\,\mathrm{d}x.$$

Da $f(1) = -1 + 4 = 3 \neq 1$ gilt, nutzen wir die Treppenfunktion $\varphi : [0, 1] \to \mathbb{R}$ mit $\varphi(x) = 0$ für $x \in [0, 1)$ und $\varphi(1) = 2$. Somit lässt sich $f(x)$ auf $[0, 1]$ schreiben als $f(x) = x + \varphi(x)$. Zusammen mit Satz 6.20, Definition 6.4 und Beispiel 6.18 können wir nun den Integralwert bestimmen:

$$\begin{aligned}
\int_0^2 f(x)\,\mathrm{d}x &= \int_0^1 f(x)\,\mathrm{d}x + \int_1^2 f(x)\,\mathrm{d}x \\
&= \int_0^1 x + \varphi(x)\,\mathrm{d}x + \int_1^2 -x + 4\,\mathrm{d}x \\
&= \int_0^1 x\,\mathrm{d}x + \int_0^1 \varphi(x)\,\mathrm{d}x - \int_1^2 x\,\mathrm{d}x + \int_1^2 4\,\mathrm{d}x \\
&= \int_0^1 x\,\mathrm{d}x + \int_0^1 \varphi(x)\,\mathrm{d}x - \left(\int_0^2 x\,\mathrm{d}x - \int_0^1 x\,\mathrm{d}x\right) + \int_1^2 4\,\mathrm{d}x \\
&= \frac{1^2}{2} + 0 \cdot (1 - 0) - \left(\frac{2^2}{2} - \frac{1^2}{2}\right) + 4 \cdot (2 - 1) \\
&= \frac{1}{2} + 0 - \frac{3}{2} + 4 \\
&= 3.
\end{aligned}$$

Der Wert eines Integrals wurde bisher nur für die Grenzen $a < b$ ermittelt. Die folgende Definition erweitert dies auf beliebige Grenzen.

Definition 6.23 (beliebige Integralgrenzen)

Sei f integrierbar auf $[a, b]$ und $c \in [a, b]$, dann definieren wir

$$\int_b^a f(x)\,\mathrm{d}x := -\int_a^b f(x)\,\mathrm{d}x$$
$$\int_c^c f(x)\,\mathrm{d}x := 0$$

Mit diesen Definitionen lässt sich Satz 6.21 auch für beliebige Integrationsgrenzen anwenden. Wenn wir die untere Integralgrenze konstant wählen und die obere als Funktionsvariable verwenden, können wir über das Integral eine Funktion F definieren:

$$F(x) = \int_a^x f(t)\,\mathrm{d}t.$$

Beachten Sie, dass wir für die Integrationsgrenze und für die Funktionsvariable von f unterschiedliche Variablen verwenden müssen. Den Zusammenhang von der Integralfunktion F, welche wir als *Stammfunktion* definieren werden und der integrierten Funktion f werden wir im nun folgenden Kapitel herleiten.

6.3 Der Hauptsatz der Differential- und Integralrechnung

Als Sie in der Schule Integralwerte berechnet haben, wurden diese höchstwahrscheinlich nicht durch eine Verfeinerung von Ober- und Untersummen bestimmt. Aus der Schule wissen Sie vermutlich schon/noch, dass die Integration so etwas wie die Umkehrung der Ableitung ist: Wenn wir eine Funktion F finden, deren Ableitung die zu integrierende Funktion f ist, also $F'(x) = f(x)$, dann können wir den Integralwert mittels

$$\int_a^b f(x)\,\mathrm{d}x = F(b) - F(a)$$

bestimmen. Wenn Sie dies mit Ihrer mittlerweile hoffentlich gut ausgeprägten mathematischen Skepsis betrachten, wird Ihnen vielleicht auffallen, wie wenig offensichtlich dieser Zusammenhang ist. Was haben Ableitungen mit Flächen unter Funktionsgraphen zu tun? Warum können wir das Integral einer beliebig komplizierten Funktion f über einem beliebig großen Intervall $[a, b]$ exakt bestimmen, indem wir nur zwei Werte einer Funktion F an den Intervallgrenzen voneinander abziehen?

Der Satz, der diesen Zusammenhang beschreibt, wird als *Hauptsatz der Differential- und Integralrechnung* oder auch als *Fundamentalsatz der Analysis* bezeichnet. Er bekam einen so bedeutsam klingenden Namen, weil er zwei fundamentale Themengebiete der Analysis, welche auf den ersten Blick keinen Zusammenhang zu haben scheinen, miteinander in Beziehung setzt. In diesem Kapitel werden wir diesen Satz nun beweisen. Dafür benötigen wir zunächst einen dritten Mittelwertsatz (siehe auch Satz 5.18 und Satz 5.29).

Satz 6.24 (Mittelwertsatz der Integralrechnung)

Sei $f : [a, b] \to \mathbb{R}$ integrierbar und seien m und M eine untere bzw. obere Schranke von f, also $\forall x \in [a, b] : m \leq f(x) \leq M$. Dann gilt

$$m(b-a) \;\leq\; \int_a^b f(x)\,\mathrm{d}x \;\leq\; M(b-a).$$

Ist f außerdem stetig, dann existiert ein $c \in (a, b)$ mit

$$\int_a^b f(x)\,\mathrm{d}x = (b-a)f(c).$$

▼ Beweis

Integration der Ungleichungen $m \leq f(x) \leq M$ ergibt mit Satz 6.20 den ersten Teil der Behauptung:

$$\int_a^b m\,\mathrm{d}x \;=\; m(b-a) \;\leq\; \int_a^b f(x)\,\mathrm{d}x \;\leq\; \int_a^b M\,\mathrm{d}x \;=\; M(b-a).$$

Seien nun

$$\begin{aligned} m &= \min\{f(x) \mid x \in [a, b]\}, \\ M &= \max\{f(x) \mid x \in [a, b]\}, \\ I &= \int_a^b f(x)\,\mathrm{d}x. \end{aligned}$$

Die Werte m und M werden nach Satz 4.36 an zwei Punkten $x_1 < x_2 \in [a, b]$ angenommen. An welchem der beiden jeweils min/max angenommen werden, ist für den weiteren Beweis irrelevant.

Aufgrund des ersten Teil des Satzes gilt

$$m(b-a) \leq I \leq M(b-a) \quad \Leftrightarrow \quad m \leq \frac{1}{b-a} I \leq M,$$

also $\frac{1}{b-a} I \in [m, M]$.

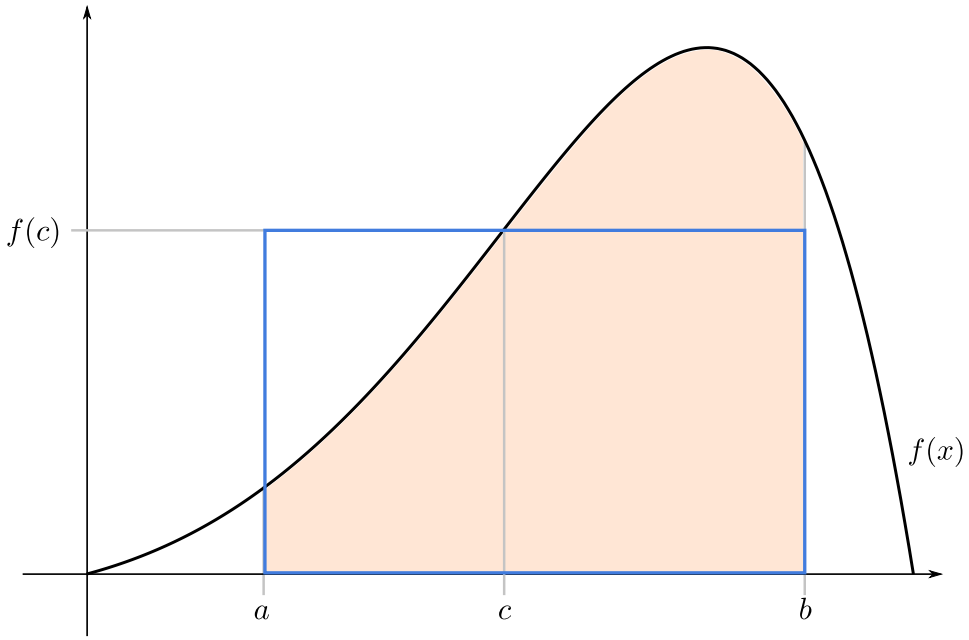
Nach dem Zwischenwertsatz (Satz 4.31) nimmt $f(x)$ auf (x_1, x_2) jeden Wert zwischen m und M an. Also existiert ein $c \in (x_1, x_2) \subseteq [a, b]$ mit

$$f(c) = \frac{1}{b-a} I = \frac{1}{b-a} \int_a^b f(x)\,\mathrm{d}x,$$

woraus die Behauptung folgt.

□

Der Mittelwertsatz der Integralrechnung hat auch eine geometrische Bedeutung: $f(c)$ ist der Wert, den eine konstante Funktion haben müsste, um auf $[a, b]$ den gleichen Integralwert zu erreichen. Dies ist in folgender Abbildung visualisiert, wo die rötliche Integralfläche genauso groß ist wie die blau umrandete Rechteckfläche.



In der folgenden Geogebra-App können Sie den Mittelwertsatz interaktiv ausprobieren:

▶ Demo: Mittelwertsatz der Integralrechnung

$f(c)$ bezeichnet man auch als den Mittelwert der Funktion auf dem Intervall $[a, b]$. Bei einer normalen Mittelwertberechnung endlicher Mengen summieren wir alle Werte auf und teilen durch die Anzahl der Werte. Hier integrieren wir alle Funktionswerte auf (was, wie Sie mittlerweile wissen, auch eine Art unendlich feine Summe darstellt) und teilen anschließend durch die Länge des Integrationsintervalls. Man kann auch gewichtete Funktionsmittelwerte definieren, indem man zusätzlich eine Gewichtsfunktion $w : [a, b] \to \mathbb{R}_{>0}$ definiert. Auch hier lässt sich ganz analog eine Verallgemeinerung des Mittelwertsatzes zeigen (siehe z.B. Satz 18.7 in Forsters “Analysis 1”). Darüber definiert sich der gewichtete Funktionsmittelwert als

$$f(c) = \frac{\int_a^b w(x) f(x)\,\mathrm{d}x}{\int_a^b w(x)\,\mathrm{d}x},$$

was erneut analog zum endlichen Fall interpretiert werden kann (gewichtete Summe geteilt durch die Summe der Gewichte).

Nun haben wir alles vorbereitet, was wir zum Beweis des Hauptsatzes benötigen. Für den Rest des Kapitels sei $I \subset \mathbb{R}$ ein aus mindestens zwei Punkten bestehendes offenes, halboffenes oder abgeschlossenes endliches oder unendliches Intervall.

Satz 6.25 (Hauptsatz der Differential- und Integralrechnung)

Sei $f : I \to \mathbb{R}$ eine stetige Funktion und $c \in I$. Sei außerdem $F : I \to \mathbb{R}$ für $x \in I$ definiert als

$$F(x) = \int_c^x f(t)\,\mathrm{d}t.$$

Dann folgt:

a. F ist stetig differenzierbar und es gilt $F'(x) = f(x)$.

b. Für beliebige $a, b \in I$ gilt $\int_a^b f(t)\,\mathrm{d}t = F(b) - F(a)$.

▼ Beweis

a. Für $h \neq 0$ gilt wegen Satz 6.21 und Satz 6.24

$$\begin{aligned}
\frac{F(x+h) - F(x)}{h} &= \frac{1}{h}\left(\int_c^{x+h} f(t)\,\mathrm{d}t - \int_c^x f(t)\,\mathrm{d}t\right) \\
&= \frac{1}{h}\left(\int_x^c f(t)\,\mathrm{d}t + \int_c^{x+h} f(t)\,\mathrm{d}t\right) \\
&\overset{S.6.21}{=} \frac{1}{h}\int_x^{x+h} f(t)\,\mathrm{d}t \\
&\overset{S.6.24}{=} \frac{1}{h}(hf(d_h)) \\
&= f(d_h),
\end{aligned}$$

für ein $d_h \in [x, x+h]$. Im Limes $h \to 0$ gilt $d_h \to x$, und da f stetig ist, gilt auch $f(d_h) \to f(x)$, und es folgt

$$\lim_{h\to 0} \frac{F(x+h) - F(x)}{h} = F'(h) = \lim_{h\to 0} f(d_h) = f(x).$$

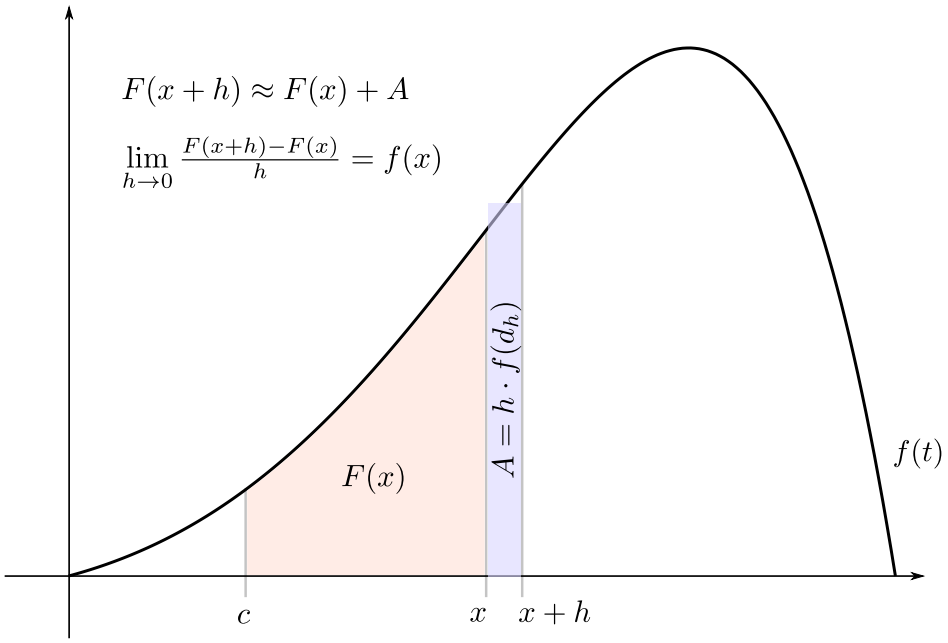
b. Die Behauptung folgt direkt aus der Additivität des Integrals (Satz 6.21):

$$F(b) - F(a) = \int_c^b f(t)\,\mathrm{d}t - \int_c^a f(t)\,\mathrm{d}t = \int_c^b f(t)\,\mathrm{d}t + \int_a^c f(t)\,\mathrm{d}t = \int_a^b f(t)\,\mathrm{d}t.$$

□

Man kann sich den Hauptsatz auch wieder anhand einer geometrischen Interpretation plausibel machen. Die Fläche unter dem Graphen von $f(t)$ im Intervall $[a, x]$ ist $F(x)$. Wenn wir das x nun um ein kleines h nach rechts verschieben, erhalten wir die Fläche $F(x+h)$. Wenn wir die hinzukommende Fläche als Rechteck approximieren, dann hat dieses Rechteck eine Fläche von

$f(d_h) \cdot h$. Dabei ist d_h ein beliebiger Wert zwischen x und $x+h$ (da wir in Satz 6.19 gezeigt haben, dass beliebige Zwischenwerte für den Grenzwertübergang zum Integral geeignet sind). Es ist also $F(x+h) \approx F(x) + f(d_h) \cdot h$, wobei die Näherung immer besser wird, je kleiner wir das h wählen. Wenn wir dies nach $f(d_h)$ umstellen und den Grenzwert $h \to 0$ bilden, ergeben sich die restlichen Schritte analog zum Beweis. Dies ist in folgendem Bild veranschaulicht:



Auch die zweite Aussage ist geometrisch gut verständlich, wenn wir ein $c < a$ nutzen: Die Fläche über $[c, b]$ ergibt sich als Summe der Flächen über $[c, a]$ und $[a, b]$ und umgekehrt erhalten wir die Fläche über $[a, b]$, wenn wir die anderen beiden voneinander subtrahieren.

Zuletzt definieren wir uns die Funktion F noch als *Stammfunktion* und zeigen ein Beispiel.

Definition 6.26 (Stammfunktion, unbestimmtes Integral)

Sei $f : [a, b] \to \mathbb{R}$. Eine Funktion $F(x)$ mit der Eigenschaft $F'(x) = f(x)$ auf $[a, b]$ nennen wir *Stammfunktion* oder *unbestimmtes Integral* von $f(x)$. Wir schreiben dafür auch abkürzend

$$F(x) = \int f(x)\,\mathrm{d}x.$$

Wir führen außerdem folgende Kurzschreibweise ein:

$$[F(x)]_a^b := F(b) - F(a).$$

Manchmal sieht man auch die Schreibweise

$$F(x)\Big|_a^b = F(b) - F(a),$$

wobei hier bei längeren Funktionsvorschriften manchmal nicht ganz klar ist, wo der Ausdruck der Stammfunktion beginnt und wo er endet, was durch die oben eingeführte Klammerschreibweise klarer ist.

Wenn $F(x)$ eine Stammfunktion von f ist, dann ist offensichtlich auch $F(x) + c$ für eine beliebige Konstante $c \in \mathbb{R}$ eine Stammfunktion von f. In der Regel bestimmen wir aber einfach *eine* Stammfunktion von f, da für die Integralbestimmung nach dem zweiten Teil des Hauptsatzes die Wahl der Konstante irrelevant ist.

Beispiel 6.27

Wir haben in Beispiel 5.9 bereits gesehen, dass für $F(x) = \ln(x)$ mit $x > 0$ gilt $F'(x) = f(x) = \frac{1}{x}$. Für $a, b > 0$ folgt damit

$$\int_a^b \frac{1}{x}\,\mathrm{d}x = F(b) - F(a) = [\ln(x)]_a^b = \ln(b) - \ln(a) = \ln\left(\frac{b}{a}\right).$$

Da wir nun das Integrieren und das Differenzieren in Beziehung gesetzt haben, können wir im nächsten Kapitel unsere bekannten Ableitungsregeln nutzen, um damit zwei wichtige Integrationstechniken herzuleiten.

6.4 Die Kunst des Integrierens

Im Gegensatz zum Differenzieren, bei dem wir auch noch so kompliziert verschachtelte Funktionen (mit genügend Zeit und Lust) problemlos ableiten können, ist das Integrieren in den meisten Fällen nicht nur ein einfaches Abarbeiten von Regeln. Man hört auch oft Sätze wie "Differenzieren ist Handwerk, Integrieren ist Kunst!". Trotzdem haben natürlich alle Künstler*innen gewisse Techniken, welche ihnen bei der Erschaffung eines Meisterwerks helfen, und diese wollen wir Ihnen hier natürlich auch für das Kunstwerk der Integration beibringen.

Dabei nutzen wir im Wesentlichen den Hauptsatz (Satz 6.25) aus dem letzten Kapitel zusammen mit den Ableitungsregeln aus Kapitel 5. Wenn wir die Stammfunktion einer Funktion kennen, so können wir das bestimmte Integral durch Auswertung der Stammfunktion einfach bestimmen. Daher ist die erste "Technik" das simple Erstellen einer Ableitungsliste für die häufigsten Funktionen, die uns durch den Hauptsatz umgekehrt als Stammfunktionsliste dient. Wir zeigen hier eine kleine Auswahl, aber zögern Sie nicht, diese um eigene Einträge zu ergänzen. Die Richtigkeit der Angaben können Sie selbst durch einfaches Ableiten überprüfen.

$f(x) = F'(x)$	$F(x) = \int f(x)\,\mathrm{d}x$	**Definitionsbereich von f**
c	cx	$\mathbb{R}$ für $c \in \mathbb{R}$
x^n	$\frac{1}{n+1}x^{n+1}$	$\mathbb{R}$ für $n \in \mathbb{N}$
$\frac{1}{x^n}$	$-\frac{1}{n-1}\,\frac{1}{x^{n-1}}$	$\mathbb{R} \setminus \{0\}$ für $n \in \mathbb{N} \setminus \{1\}$,

		Definitionsbereich von
x^α	$\frac{1}{\alpha+1}x^{\alpha+1}$	$\mathbb{R}_{>0}$ für $\alpha \in \mathbb{R} \setminus \{-1\}$
$\frac{1}{x}$	$\ln(\lvert x\rvert)$	$\mathbb{R} \setminus \{0\}$
e^x	e^x	$\mathbb{R}$
a^x	$\frac{a^x}{\ln(a)}$	$\mathbb{R}$ für $a > 0, a \neq 1$
$\ln(\lvert x\rvert)$	$x \cdot \ln(\lvert x\rvert) - x$	$\mathbb{R} \setminus \{0\}$
$\frac{1}{1+x^2}$	$\arctan(x)$	$\mathbb{R}$
$\frac{1}{1-x^2}$	$\frac{1}{2}\ln\left(\left\lvert\frac{x+1}{x-1}\right\rvert\right)$	$\mathbb{R} \setminus \{-1, 1\}$
$\frac{1}{\sqrt{1+x^2}}$	$\operatorname{arsinh}(x)$	$\mathbb{R}$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin(x)$	$(-1, 1)$
$\sin(x)$	$-\cos(x)$	$\mathbb{R}$
$\cos(x)$	$\sin(x)$	$\mathbb{R}$
$\tan(x)$	$-\ln(\lvert\cos(x)\rvert)$	$\mathbb{R} \setminus \left\{x = \frac{\pi}{2} + k \cdot \pi \mid k \in \mathbb{Z}\right\}$

Hierbei sind die Funktionen *Arcussinus* $\arcsin(x)$, *Arcuscosinus* $\arccos(x)$ und *Arcustangens* $\arctan(x)$ die Umkehrfunktionen von Sinus, Cosinus und Tagens, welche oft auch mit $\operatorname{asin}(x)$, $\operatorname{acos}(x)$ und $\operatorname{atan}(x)$ bezeichnet werden. Wir hatten diese Funktionen in Kapitel 4.4 schon einmal kurz erwähnt. Die Funktion *Areasinus hyperbolicus* $\operatorname{arsinh}(x) := \ln\left(x + \sqrt{x^2+1}\right)$ ist die Umkehrfunktion des Sinus hyperbolicus $\sinh(x)$.

Als Nächstes können wir die Produktregel (Satz 5.6) für Integrale von Produkten zweier Funktionen anwenden. Wenn wir eine Funktion $(f \cdot g)(x)$ ableiten, ergibt sich mit der Produktregel

$$(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x).$$

Dies können wir nun genauso zur Integralbestimmung nutzen, was den folgenden Satz ergibt:

Satz 6.28 (Partielle Integration)

Seien $f, g : [a, b] \to \mathbb{R}$ zwei auf $[a, b]$ stetig differenzierbare Funktionen. Dann gilt

$$\int_a^b f(x)g'(x)\,\mathrm{d}x = [f(x)g(x)]_a^b - \int_a^b f'(x)g(x)\,\mathrm{d}x.$$

bzw. in unbestimmter Integralform

$$\int f(x)g'(x)\,\mathrm{d}x = f(x)g(x) - \int f'(x)g(x)\,\mathrm{d}x.$$

▼ Beweis

Wenn f und g stetig differenzierbar sind, dann ist nach Satz 4.27 und Satz 5.6 auch $f \cdot g$ stetig differenzierbar. Seien $H, h : [a, b] \to \mathbb{R}$ zwei Funktionen mit

$$H(x) := f(x)g(x) \quad \text{und} \quad h(x) := H'(x).$$

Dann ist H die Stammfunktion von h und es gilt nach der Produktregel (Satz 5.6)

$$h(x) = H'(x) = f'(x)g(x) + f(x)g'(x).$$

Daher gilt

$$\begin{aligned}\int_a^b f'(x)g(x)\,\mathrm{d}x + \int_a^b f(x)g'(x)\,\mathrm{d}x &= \int_a^b \left(f'(x)g(x) + f(x)g'(x)\right)\,\mathrm{d}x \\ &= \int_a^b h(x)\,\mathrm{d}x \\ &= [H(x)]_a^b \\ &= [f(x)g(x)]_a^b.\end{aligned}$$

Durch Umstellen nach $\int_a^b f(x)g'(x)\,\mathrm{d}x$ ergibt sich die Behauptung.

□

Partielle Integration ist dann gut einsetzbar, wenn wir ein Produkt aus zwei Funktionen integrieren, von einem der beiden Faktoren die Stammfunktion kennen (g), und der andere durch Ableiten einfacher wird (f). In manchen Fällen führt man ein Integral durch partielle Integration auch auf sich selbst zurück (besonders bei den trigonometrischen Funktionen). Beide Varianten werden in den folgenden Beispielen gezeigt.

Beispiel 6.29

1. Berechne $\int xe^x \,\mathrm{d}x$

$$\begin{aligned}\int \underbrace{x}_{f}\underbrace{e^x}_{g'} \,\mathrm{d}x &= \underbrace{x}_{f}\underbrace{e^x}_{g} - \int \underbrace{1}_{f'}\underbrace{e^x}_{g} \,\mathrm{d}x \\ &= xe^x - e^x \\ &= (x-1)e^x\end{aligned}$$

2. Berechne $\int x^2e^x \,\mathrm{d}x$

$$\begin{aligned}\int \underbrace{x^2}_{f}\underbrace{e^x}_{g'} \,\mathrm{d}x &= \underbrace{x^2}_{f}\underbrace{e^x}_{g} - \int \underbrace{2x}_{f'}\underbrace{e^x}_{g} \,\mathrm{d}x \\ &= x^2e^x - 2(x-1)e^x \\ &= (x^2-2x+2)e^x\end{aligned}$$

3. Berechne $\int \ln(x) \,\mathrm{d}x$ (für ein Intervall aus $\mathbb{R}_{>0}$)

$$\begin{aligned}\int \ln(x) \,\mathrm{d}x &= \int \underbrace{\ln(x)}_{f}\underbrace{1}_{g'} \,\mathrm{d}x \\ &= \underbrace{\ln(x)}_{f}\underbrace{x}_{g} - \int \underbrace{\frac{1}{x}}_{f'}\underbrace{x}_{g} \,\mathrm{d}x \\ &= x\ln(x) - \int 1 \,\mathrm{d}x \\ &= x\ln(x) - x\end{aligned}$$

4. Berechne $\int \sin^2(x) \,\mathrm{d}x$

$$\begin{aligned}\int \sin^2(x) \,\mathrm{d}x &= \int \underbrace{\sin(x)}_{f}\underbrace{\sin(x)}_{g'} \,\mathrm{d}x \\ &= \underbrace{\sin(x)}_{f}\underbrace{(-\cos(x))}_{g} - \int \underbrace{\cos(x)}_{f'}\underbrace{(-\cos(x))}_{g} \,\mathrm{d}x \\ &= -\sin(x)\cos(x) + \int \underbrace{\cos^2(x)}_{1-\sin^2(x)} \,\mathrm{d}x \\ &= -\sin(x)\cos(x) + \int 1 \,\mathrm{d}x - \int \sin^2(x) \,\mathrm{d}x \\ \Rightarrow \int \sin^2(x) \,\mathrm{d}x &= \frac{1}{2}(x - \sin(x)\cos(x))\end{aligned}$$

Genauso kann auch aus der Kettenregel eine Integralregel abgeleitet werden, welche wir als *Substitutionsregel* bezeichnen.

Satz 6.30 (Integration durch Substitution)

Sei $f : I \to \mathbb{R}$ eine stetige Funktion und $g : [a, b] \to \mathbb{R}$ stetig differenzierbar mit $g([a, b]) \subseteq I$. Dann gilt

$$\int_a^b f(g(t))\, g'(t)\, \mathrm{d}t = \int_{g(a)}^{g(b)} f(x)\, \mathrm{d}x,$$

bzw. in unbestimmter Integralform mit einer Stammfunktion F von f

$$\int f(g(t))\, g'(t)\, \mathrm{d}t = F(g(t)).$$

▼ Beweis

Die Integrale existieren, weil f, g, g' und nach Satz 4.29 auch die Komposition $f(g(t))$ stetig sind.

Sei $F : I \to \mathbb{R}$ eine Stammfunktion von f, also $F'(x) = f(x)$ bzw. $F(x) = \int f(x)\, \mathrm{d}x$. Nach der Kettenregel (Satz 5.6) gilt

$$(F(g(t)))' = F'(g(t))\, g'(t) = f(g(t))\, g'(t),$$

bzw. in Integralform

$$F(g(t)) = \int f(g(t))\, g'(t)\, \mathrm{d}t,$$

woraus mit den Integrationsgrenzen $a, b \in I$ und einem beliebigen $c \in I$ folgt

$$\begin{aligned}\int_a^b f(g(t))\, g'(t)\, \mathrm{d}t &= F(g(b)) - F(g(a)) \\ &= \int_c^{g(b)} f(t)\, \mathrm{d}t - \int_c^{g(a)} f(t)\, \mathrm{d}t \\ &= \int_{g(a)}^{g(b)} f(x)\, \mathrm{d}x.\end{aligned}$$

□

Beispiel 6.31 (Substitutionsregel)

Wir wollen $\int_1^2 (2t-2)^9 \,\mathrm{d}t$ bestimmen. Mit $f(x) = \frac{1}{2}x^9$ und $g(t) = 2t-2$ gilt $g'(t) = 2$, und wir können das Integral wie folgt umformen:

$$\begin{aligned}
\int_1^2 (2t-2)^9 \,\mathrm{d}t &= \int_1^2 f(g(t))\, g'(t) \,\mathrm{d}t \\
&= \int_{g(1)}^{g(2)} f(x) \,\mathrm{d}x \\
&= \int_0^2 \frac{1}{2} x^9 \,\mathrm{d}x \\
&= \left[\frac{1}{2}\frac{1}{10}x^{10}\right]_0^2 \\
&= \frac{1}{20}\left(2^{10} - 0^{10}\right) \\
&= \frac{512}{10}
\end{aligned}$$

Mit der verkürzten Schreibweise $\mathrm{d}g(x) := g'(x)\,\mathrm{d}x$ lässt sich die Substitutionsregel auch schreiben als

$$\int f(g(x))\,\mathrm{d}g(x) = F(g(x)).$$

Dies ist besonders leicht zu merken, da man hier lediglich das x einer herkömmlichen Integration durch $g(x)$ ersetzt. Außerdem ergibt sich daraus die Schreibweise für die Ableitung (siehe Definition 5.1), wenn wir $\mathrm{d}x$ etwas unsauber wie eine Konstante dividieren:

$$\mathrm{d}g(x) = g'(x)\,\mathrm{d}x \quad \Leftrightarrow \quad \frac{\mathrm{d}g(x)}{\mathrm{d}x} = g'(x).$$

Die Abhängigkeit von x vernachlässigen wir dabei oft und schreiben $\mathrm{d}g$ anstellen von $\mathrm{d}g(x)$.

Die Wahl einer geeigneten Substitution kann den Weg zu einer Lösung stark vereinfachen, wobei eine falsche Wahl das Integral auch noch weiter verkomplizieren kann. Daher ist es nie verkehrt, mehrere Substitutionen für dasselbe Integral auszuprobieren, um ein besseres Gefühl dafür zu bekommen. Letztendlich gibt es hier keine klaren Regeln, wann welche Substitution zum Erfolg führt. Stattdessen ist es eher eine Mischung aus Intuition, Glück und Übung.

Beispiel 6.32 (Substitution)

1. Wir wollen $\int_1^2 (2t-2)^9 \,\mathrm{d}t$ bestimmen.

   Wir substituieren $g := 2t-2$, nutzen also die innere Funktion $g(t) = 2t-2$. Damit ergibt sich $\mathrm{d}g(t) = 2\,\mathrm{d}t$, also $\frac{1}{2}\,\mathrm{d}g = \mathrm{d}t$. Für die Grenzen ergibt sich $g(1) = 0$ sowie $g(2) = 2$. Insgesamt folgt:

$$\begin{aligned}\int_1^2 (2t-2)^9 \,\mathrm{d}t &= \int_0^2 g^9 \frac{1}{2}\,\mathrm{d}g \\ &= \left[\frac{g^{10}}{2\cdot 10}\right]_0^2 \\ &= \frac{1}{20}\left(2^{10} - 0^{10}\right) \\ &= \frac{512}{10}\end{aligned}$$

2. Wir wollen $\int \tan(t)\,\mathrm{d}t = \int \frac{\sin(t)}{\cos(t)}\,\mathrm{d}t$ bestimmen.

   Wir substituieren $\cos(t) = g$, also $\mathrm{d}g = -\sin t\,\mathrm{d}t$. Damit ergibt sich:

$$\begin{aligned}\int \tan(t)\,\mathrm{d}t &= \int \frac{\sin(t)}{\cos(t)}\,\mathrm{d}t \\ &= \int \frac{1}{\cos(t)} \underbrace{\sin(t)\,\mathrm{d}t}_{=-\,\mathrm{d}g} \\ &= -\int \frac{1}{g}\,\mathrm{d}g \\ &= -\ln\left(|g|\right) \\ &= -\ln\left(|\cos(t)|\right).\end{aligned}$$

Unter bestimmten Voraussetzungen bieten sich bestimmte Substitutionen besonders an, woraus wir die folgenden Regeln ableiten können.

Satz 6.33 (Integralvereinfachungen)

a. $\int f(ax+b)\,\mathrm{d}x = \frac{1}{a}F(ax+b) \quad (\text{mit } F'(x) = f(x))$

b. $\int f(x)f'(x)\,\mathrm{d}x = \frac{1}{2}f^2(x)$

c. $\int \frac{f'(x)}{f(x)}\,\mathrm{d}x = \ln\left(|f(x)|\right)$

▼ Beweis als Übung

Beispiel 6.34 (Integralvereinfachungen)

1. $\int \ln(4x)\,\mathrm{d}x = \frac{1}{4}(4x\,(\ln(4x)-1)) = x\,(\ln(4x)-1)$

   Hierbei wurde Vereinfachung (a) mit $F(x) = x\ln(x) - x = x(\ln(x)-1)$ gewählt, da $F'(x) = \ln(x)$.

2. $\int \underbrace{\sin(x)}_{f(x)}\,\underbrace{\cos(x)}_{f'(x)}\,\mathrm{d}x = \frac{1}{2}\sin^2(x)$

   Hierbei wurde Vereinfachung (b) verwendet.

3. $\int \frac{1}{x\ln(x)}\,\mathrm{d}x = \int \frac{\frac{1}{x}}{\ln(x)}\,\mathrm{d}x = \ln(|\ln(x)|)$

   Hierbei wurde Vereinfachung (c) verwendet. Beachten Sie, dass das Integral nur für $x > 0$ definiert ist, weswegen der Betrag im inneren Logarithmus nicht genutzt werden muss.

In manchen Fällen kann es auch sinnvoll sein, einen Ausdruck im Integranden nicht durch eine Variable (wie z.B. $g = 3t+1$) zu ersetzen, sondern umgekehrt die Integrationsvariable als Funktion zu definieren. Dafür betrachten wir das folgende Beispiel. Dort zeigen wir außerdem, dass sich der Wert von π auch mit unserer Definition (als erste Nullstelle des Cosinus, siehe Definition 4.66) als die Fläche des Einheitskreises herausstellt.

Beispiel 6.35 (Fläche des Einheitskreises)

Der Graph der Funktion $f(x) = \sqrt{1-x^2}$ beschreibt für $x \in [-1, 1]$ einen Halbkreis mit Radius $r = 1$, da

$$x^2 + f(x)^2 = x^2 + 1 - x^2 = 1.$$

Daher sollte gelten

$$\int_{-1}^{1} f(x)\,\mathrm{d}x = \frac{\pi}{2}.$$

Wir nutzen die Substitution $x = \sin(u)$, woraus folgt $\mathrm{d}x = \mathrm{d}\sin(u) = \cos(u)\,\mathrm{d}u$, bzw. $u(x) = \arcsin(x)$. Es gilt für die Grenzen $u(-1) = -\pi/2$ und $u(1) = \pi/2$. Damit ergibt sich also

$$\begin{aligned}\int_{-1}^{1} \sqrt{1-x^2}\,\mathrm{d}x &= \int_{-\pi/2}^{\pi/2} \sqrt{1-\sin^2(u)}\,\mathrm{d}\sin(u) \\ &= \int_{-\pi/2}^{\pi/2} \sqrt{\cos^2(u)}\,\cos(u)\,\mathrm{d}u \\ &= \int_{-\pi/2}^{\pi/2} \cos^2(u)\,\mathrm{d}u.\end{aligned}$$

Es gilt außerdem

$$\cos^2(u) = \left(\frac{e^{\mathrm{i}u} + e^{-\mathrm{i}u}}{2}\right)^2 = \frac{1}{4}(e^{2\mathrm{i}u} + 2 + e^{-2\mathrm{i}u}) = \frac{1}{2}(\cos(2u) + 1),$$

womit weiter folgt

$$\begin{aligned}\int_{-\pi/2}^{\pi/2} \cos^2(u)\,\mathrm{d}u &= \int_{-\pi/2}^{\pi/2} \frac{1}{2}(\cos(2u) + 1)\,\mathrm{d}u \\ &= \left[\frac{1}{2}\left(\frac{1}{2}\sin(2u) + u\right)\right]_{-\pi/2}^{\pi/2} \\ &= \frac{1}{2}(0 + \pi/2 - (0 - \pi/2)) \\ &= \frac{\pi}{2}.\end{aligned}$$

Bei der Integration rationaler Funktionen zerlegt man diese durch Polynomdivision und Partialbruchzerlegung in Polynome und Restbrüche. Diese kann man dann einzeln deutlich einfacher integrieren. Durch Substitutionen können auch viele andere Funktionen auf rationale Funktionen zurückgeführt werden, wie das folgende Beispiel zeigt.

Beispiel 6.36 (Integration einer rationalen Funktion)

Das Integral $\int \frac{e^{3x}+3}{e^x+1}\,\mathrm{d}x$ soll bestimmt werden. Mit der Substitution $t = e^x$ und $\mathrm{d}t = e^x\,\mathrm{d}x$ ergibt sich daraus das Integral einer rationalen Funktion:

$$\int \frac{e^{3x}+3}{e^x+1}\,\mathrm{d}x = \int \frac{t^3+3}{t(t+1)}\,\mathrm{d}t = \int \frac{t^3+3}{t^2+t}\,\mathrm{d}t$$

Mithilfe einer Polynomdivision zerlegen wir die rationale Funktion in ihren polynomiellen Anteil und den echt gebrochenen Rest (siehe Satz 4.44):

$$\begin{aligned}(t^3+3) : (t^2+t) &= t-1+\frac{t+3}{t^2+t}\\ \underline{-(t^3+t^2)}\quad&\\ -t^2+3\quad&\\ \underline{-(-t^2-t)}\quad&\\ t+3\quad&\end{aligned}$$

Damit ist $\frac{t+3}{t(t+1)}$ echt gebrochen rational (Polynomgrad im Zähler $<$ Polynomgrad im Nenner).

Mittels einer Partialbruchzerlegung zerlegen wir diesen Rest in Brüche mit linearem Nenner:

$$\frac{t+3}{t(t+1)} = \frac{A}{t} + \frac{B}{t+1} = \frac{A(t+1)+Bt}{t(t+1)} = \frac{(A+B)t+A}{t(t+1)}.$$

Damit $(A+B)t + A = t+3$ gilt, muss $A = 3$ und $B = -2$ gelten, es folgt also

$$\frac{t+3}{t(t+1)} = \frac{3}{t} - \frac{2}{t+1}.$$

Die Summanden der so zerlegten rationalen Funktion können wir nun integrieren.

$$\begin{aligned}\int \frac{t^3+3}{t(t+1)}\,\mathrm{d}t &= \int (t-1)\,\mathrm{d}t + \int \frac{3}{t}\,\mathrm{d}t - \int \frac{2}{t+1}\,\mathrm{d}t\\ &= \frac{1}{2}t^2 - t + 3\ln(|t|) - 2\ln(|t+1|)\end{aligned}$$

Mit der Rücksubstitution $t = e^x$ ergibt sich das gesuchte unbestimmte Integral

$$\int \frac{e^{3x}+3}{e^x+1}\,\mathrm{d}x = \frac{1}{2}e^{2x} - e^x + 3x - 2\ln(e^x+1).$$

Auch wenn wir für viele Funktionen eine Stammfunktion finden können, die aus elementaren Funktionen (siehe Kapitel 4.4) zusammengesetzt ist, funktioniert das nicht für jeden Integranden.

Typische Beispiele sind die Fehlerfunktion erf und der Integralsinus Si, die definiert sind als

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} \, \mathrm{d}t,$$
$$\operatorname{Si}(x) := \int_0^x \frac{\sin(t)}{t} \, \mathrm{d}t.$$

Beide Funktionen lassen sich nicht als Kombination elementarer Funktionen schreiben, obwohl der Integrand eine solche Kombination darstellt. Die Integralwerte müssen hier also numerisch angenähert werden, wie im vorangegangenen Kapitel gezeigt.

6.5 Uneigentliche Integrale

Bisher haben wir das Integral in Definition 6.9 nur für kompakte Intervalle $[a, b]$ definiert. Außerdem umfasst die Definition nur Funktionen, die auf diesem Intervall beschränkt sind. Dies wollen wir nun erweitern, um auch unbeschränkte Funktion sowie auch über uneigentliche Intervalle (z.B. $[0, \infty)$) integrieren zu können. Auch hier hilft uns wieder eine Grenzwertbestimmung: Wir bestimmen erst das Integral für ein kompaktes Intervall und bilden dann den Limes gegen die uneigentliche Grenze (oder die Stelle der Unbeschränktheit).

Definition 6.37 (Integrale über uneigentliche Intervalle)

Sei $f : [a, \infty) \to \mathbb{R}$ eine über jedes Intervall $[a, c]$ für $a < c < \infty$ integrierbare Funktion. Dann definieren wir das *uneigentliche Integral* als

$$\int_a^\infty f(x) \, \mathrm{d}x := \lim_{c \to \infty} \int_a^c f(x) \, \mathrm{d}x,$$

falls der Grenzwert existiert. In diesem Fall sagen wir auch, dass das Integral *konvergiert* (oder anderenfalls *divergiert*).

Analog definieren wir das uneigentliche Integral für eine Funktion $f : (-\infty, a] \to \mathbb{R}$, welche auf $[c, a]$ für $-\infty < c < a$ integrierbar ist, als

$$\int_{-\infty}^a f(x) \, \mathrm{d}x := \lim_{c \to -\infty} \int_c^a f(x) \, \mathrm{d}x.$$

Für das Integrationsintervall $(-\infty, \infty)$ definieren wir das uneigentliche Integral für ein beliebiges $c \in \mathbb{R}$ als

$$\int_{-\infty}^\infty f(x) \, \mathrm{d}x := \int_{-\infty}^c f(x) \, \mathrm{d}x + \int_c^\infty f(x) \, \mathrm{d}x,$$

wobei beide Integrale auf der rechten Seite existieren müssen, damit das uneigentliche Integral konvergiert.

Beispiel 6.38

1. Berechne $\int_0^\infty e^{-x}\,\mathrm{d}x$

$$\begin{aligned}\int_0^\infty e^{-x}\,\mathrm{d}x &= \lim_{c\to\infty}\int_0^c e^{-x}\,\mathrm{d}x\\ &= \lim_{c\to\infty}\left[-e^{-x}\right]_0^c\\ &= \lim_{c\to\infty}\left(1-e^{-c}\right) = 1\end{aligned}$$

   Das uneigentliche Integral existiert damit und ist konvergent.

2. Berechne $\int_1^\infty \frac{1}{x^\alpha}\,\mathrm{d}x$

$$\begin{aligned}\int_1^\infty \frac{1}{x^\alpha}\,\mathrm{d}x &= \lim_{c\to\infty}\int_1^c \frac{1}{x^\alpha}\,\mathrm{d}x\\ &= \lim_{c\to\infty}\left[\frac{1}{1-\alpha}x^{1-\alpha}\right]_1^c\\ &= \lim_{c\to\infty}\frac{c^{1-\alpha}-1}{1-\alpha}\\ &= \begin{cases}\frac{1}{\alpha-1} & \text{für } \alpha>1 \text{ (konvergent)}\\ \infty & \text{für } \alpha<1 \text{ (divergent)}\end{cases}\end{aligned}$$

   Für $\alpha = 1$ siehe nächstes Beispiel.

3. Berechne $\int_1^\infty \frac{1}{x}\,\mathrm{d}x$

$$\begin{aligned}\int_1^\infty \frac{1}{x}\,\mathrm{d}x &= \lim_{c\to\infty}\int_1^c \frac{1}{x}\,\mathrm{d}x\\ &= \lim_{c\to\infty}\left[\ln(x)\right]_1^c\\ &= \lim_{c\to\infty}\ln(c)\\ &= \infty\end{aligned}$$

   Das Integral divergiert, d.h., es existiert nicht.

4. Berechne $\int_{-\infty}^\infty \frac{1}{1+x^2}\,\mathrm{d}x$

$$\begin{aligned}\int_{-\infty}^\infty \frac{1}{1+x^2}\,\mathrm{d}x &= \lim_{c\to-\infty}\int_c^0 \frac{1}{1+x^2}\,\mathrm{d}x + \lim_{c\to\infty}\int_0^c \frac{1}{1+x^2}\,\mathrm{d}x\\ &= \lim_{c\to-\infty}\left[\arctan(x)\right]_c^0 + \lim_{c\to\infty}\left[\arctan(x)\right]_0^c\\ &= -\lim_{c\to-\infty}\arctan(c) + \lim_{c\to\infty}\arctan(c)\\ &= -\left(-\frac{\pi}{2}\right)+\frac{\pi}{2}\\ &= \pi\end{aligned}$$

   Das uneigentliche Integral existiert damit und ist konvergent.

5. Berechne $\int_0^\infty \sin(x)\,\mathrm{d}x$

$$\begin{aligned}\int_0^\infty \sin(x)\,\mathrm{d}x &= \lim_{c\to\infty}\int_0^c \sin(c)\,\mathrm{d}x \\ &= \lim_{c\to\infty}\left[-\cos(x)\right]_0^c \\ &= \lim_{c\to\infty}\left(-\cos(c)+1\right)\end{aligned}$$

Der Grenzwert existiert nicht, damit ist das Integral divergent.

Achtung: Im Allgemeinen gilt

$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x \;\neq\; \lim_{c\to\infty}\int_{-c}^{c} f(x)\,\mathrm{d}x,$$

da in der Definition gefordert wird, dass beide Grenzwerte $\int_{-\infty}^{c} f(x)\,\mathrm{d}x$ und $\int_{c}^{\infty} f(x)\,\mathrm{d}x$ existieren müssen. Ein Gegenbeispiel ist $\int_{-\infty}^{\infty} x\,\mathrm{d}x$: Für den ersten Grenzwert gilt

$$\lim_{c\to\infty}\int_{-c}^{c} x\,\mathrm{d}x \;=\; \lim_{c\to\infty}\left[\frac{1}{2}x^2\right]_{-c}^{c} \;=\; 0,$$

aber das uneigentliche Integral

$$\int_{-\infty}^{\infty} x\,\mathrm{d}x \;=\; \int_{-\infty}^{0} x\,\mathrm{d}x + \int_{0}^{\infty} x\,\mathrm{d}x$$

existiert nicht, da jedes Teilintegral divergiert.

Als Nächstes weiten wir die Integraldefinition auch auf beliebige offene oder halboffene Integrale aus, wobei wir nun nicht mehr fordern, dass die Funktion auf dem jeweiligen Intervall beschränkt ist.

Definition 6.39 (Integration über offene/halboffene Invervalle)

Sei $f:(a,b)\to\mathbb{R}$, mit $-\infty \le a < b \le \infty$, auf jedem kompakten Intervall $[\alpha,\beta]\subset(a,b)$ integrierbar. Sei $c\in(a,b)$ beliebig gewählt.

Dann definieren wir

$$\int_c^b f(x)\,\mathrm{d}x := \lim_{\beta\nearrow b}\int_c^\beta f(x)\,\mathrm{d}x,$$

$$\int_a^c f(x)\,\mathrm{d}x := \lim_{\alpha\searrow a}\int_\alpha^c f(x)\,\mathrm{d}x$$

$$\int_a^b f(x)\,\mathrm{d}x := \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x$$

falls die jeweiligen Grenzwerte existieren.

Interessant sind in der vorherigen Definition Funktionen, für die $\lim\limits_{x \searrow a} f(x) = \pm\infty$ und $\lim\limits_{x \nearrow b} f(x) = \pm\infty$.

Beispiel 6.40

1. Berechne $\int_0^1 \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x$

   $$\int_0^1 \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \lim_{c \nearrow 1} \int_0^c \frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \lim_{c \nearrow 1} \left[\arcsin(x)\right]_0^c = \lim_{c \nearrow 1} \arcsin(c) = \frac{\pi}{2}$$

2. Berechne $\int_0^1 \ln(x)\,\mathrm{d}x$

   $$\int_0^1 \ln(x)\,\mathrm{d}x = \lim_{c \searrow 0} \int_c^1 \ln(x)\,\mathrm{d}x = \lim_{c \searrow 0} \left[x\ln(x) - x\right]_c^1 = -1 - \lim_{c \searrow 0} c\ln(c) = -1$$

   Hierfür haben die folgende Nebenrechnung verwendet:

   $$\lim_{c \searrow 0} c\ln(c) = \lim_{c \searrow 0} \frac{\ln(c)}{c^{-1}} \underset{\text{l'Hospital}}{=} \lim_{c \searrow 0} \frac{c^{-1}}{-c^{-2}} = -\lim_{c \searrow 0} c = 0$$

Falls sowohl Integrationsbereich als auch zu integrierende Funktion unbeschränkt sind, dann spaltet man in mehrere uneigentliche Integrale auf und fordert, dass *alle* uneigentlichen Teilintegrale existieren müssen.

Beispiel 6.41

Sei $f : (0, \infty) \to \mathbb{R}$ mit $f(x) = \min\left\{\ln(x), \dfrac{\ln(16)}{x^2}\right\} = \begin{cases} \ln(x), & \text{für } x \leq 2, \\ \frac{\ln(16)}{x^2}, & \text{für } x > 2. \end{cases}$

Die Funktion ist stetig, daher integrierbar.

$$\begin{aligned}
\int_0^\infty f(x)\,\mathrm{d}x &= \int_0^2 f(x)\,\mathrm{d}x + \int_2^\infty f(x)\,\mathrm{d}x \\
&= \lim_{a \searrow 0} \int_a^2 \ln(x)\,\mathrm{d}x + \lim_{b \nearrow \infty} \int_2^b \frac{\ln(16)}{x^2}\,\mathrm{d}x \\
&= \lim_{a \searrow 0} \left[x\ln(x) - x\right]_a^2 + \lim_{b \nearrow \infty} \left[-\frac{\ln(16)}{x}\right]_2^b \\
&= 2\ln(2) - 2 - \underbrace{\lim_{a \searrow 0} a\ln(a)}_{=0} - \underbrace{\lim_{b \nearrow \infty} \frac{\ln(16)}{b}}_{=0} + \frac{\ln(16)}{2} \\
&= 4\ln(2) - 2
\end{aligned}$$

Die letzte Umformung gilt, da $\ln(16) = \ln(2^4) = 4\ln(2)$.

7 Funktionen im $\mathbb{R}^n$

Bisher haben wir uns mit Funktionen beschäftigt, die eine einzelne Variable (*univariat*) auf einen einzelnen Wert (*skalarwertig*) abgebildet haben. In der Praxis reichen diese *univariaten skalarwertigen* Funktionen in der Regel nicht aus, um damit reale Problem zu modellieren. Meistens hängt das Ergebnis von mehreren relevanten Faktoren ab, so dass wir mehrere Variablen (*multivariat*) brauchen. Oft ist das Ergebnis auch nicht nur eine reelle Zahl, sondern ein Vektor oder ein Tupel von Zahlen (*vektorwertig*).

In diesem Kapitel lernen wir multivariate und vektorwertige Funktionen und Differentialrechnung auf diesen Funktionen kennen. Wir führen zunächst Grundlagen des Vektorraum $\mathbb{R}^n$ der n-dimensionalen reellen Zahlen ein und verallgemeinern Begriffe wie Folgen, Konvergenz und Stetigkeit. Danach schauen wir uns parametrische Kurven (univariate vektorwertige Funktionen) und parametrische Flächen (bivariate vektorwertige Funktionen) an, welche interessante Anwendungen z.B. in der Physik oder der Computergraphik haben. Ein wichtiger Anwendungsfall in der Optimierung oder im Machine Learning ist die Minimierung einer skalarwertigen Kostenfunktion (*Loss Function*), die von mehreren (oft sehr vielen) Variablen abhängt. Wir werden dafür Extremwertbestimmung von multivariaten skalarwertigen Funktionen definieren und ein paar Algorithmen für dieses Problem sehen.

Für den Schritt zu multivariaten und/oder vektorwertigen Funktionen (wir nennen sie einfach *mehrdimensionale* Funktionen) lassen sich viele Eigenschaften, die wir für “normale” eindimensionale Funktionen kennengelernt haben, in ähnlicher Form definieren oder direkt auf den eindimensionalen Fall zurückführen. Einige Aspekte sind im Mehrdimensionalen allerdings etwas komplexer und auch weniger intuitiv. Im Mathematikstudium würden wir mit der mehrdimensionalen Analysis problemlos ein bis zwei Semester füllen. Im Informatikstudium beschränken wir uns aber auf die für die Informatik wichtigsten Aspekte und werden aus Zeitgründen auch die meisten Beweise nur grob umreißen oder ganz weglassen. Die gezeigten Beispiele beschränken sich oft auf Funktionen mit zwei oder drei Variablen, deren Verhalten man sich geometrisch noch gut vorstellen kann. Die mathematischen Konzepte sind aber für beliebige Dimensionen gültig.

Wir beginnen mit grundlegenden Definitionen des sogenannten *Euklidischen Raums*, übertragen den Begriff der Stetigkeit auf mehrdimensionale Funktionen und beschäftigen uns anschließend für den Großteil des Kapitels mit der mehrdimensionalen Differenzialrechnung. Die mehrdimensionale Integration würde leider den zeitlichen Rahmen der Veranstaltung sprengen. Hier verweisen wir Sie bei Interesse auf die einschlägigen Lehrbücher.

7.1 Grundlagen des $\mathbb{R}^n$

Definition 7.1 (n-dimensionaler Euklidischer Raum)

Wir bezeichnen das n-fache kartesische Produkt ($n \in \mathbb{N}$) der reellen Zahlen als *n-dimensionalen euklidischen Raum*, oder kurz $\mathbb{R}^n$, also

$$\mathbb{R}^n := \underbrace{\mathbb{R}\times\mathbb{R}\times\cdots\times\mathbb{R}}_{n\text{-mal}} = \{(x_1, \dots, x_n) \mid x_i \in \mathbb{R}, \text{ für } i = 1, \dots, n\}.$$

Die n-Tupel $\mathbf{x} \in \mathbb{R}^n$ bezeichnen wir als *Vektoren*.

Zur Unterscheidung bezeichnen wir die Elemente aus $\mathbb{R}^1 = \mathbb{R}$ als *Skalare*.

Vektoren $\mathbf{x} \in \mathbb{R}^n$ werden wir zur leichteren Unterscheidung von Skalaren $x \in \mathbb{R}$ immer in fettgedruckter Schrift schreiben. n-dimensionale Vektoren kann man als

$$\text{Spaltenvektoren } \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{oder} \quad \text{Zeilenvektoren } (x_1, \dots, x_n)$$

auffassen. An vielen Stellen wird es wichtig sein, dass wir diese zwei strikt unterscheiden. Es gilt also

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \neq (x_1, \dots, x_n)\,.$$

Mit $\mathbf{x} \in \mathbb{R}^n$ meinen wir stets einen Spaltenvektor. Um im Fließtext Platz zu sparen, nutzen wir die Notation

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_1, \dots, x_n)^\mathsf{T},$$

wobei das hochgestellte $^\mathsf{T}$ der Operator für Transponieren ist, der einen Zeilenvektor in einen Spaltenvektor umwandelt und umgekehrt. Dieser Operator wird in unserem Exkurs Matrizen definiert. Wenn Sie Definition 2.47 und Definition 7.1 vergleichen, werden Sie feststellen, dass die Menge der komplexen Zahlen sehr ähnlich zum $\mathbb{R}^2$ ist, allerdings mit besonderen Operatoren, die die komplexen Zahlen zu einem Körper machen. Im Allgemeinen definieren wir im euklidischen Raum $\mathbb{R}^n$ nur die folgenden Operationen.

Definition 7.2 (Operationen auf dem $\mathbb{R}^n$)

Seien $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$. Wir definieren die Operationen

1. *Vektoraddition*

$$+ : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n \quad \text{mit} \quad \mathbf{x} + \mathbf{y} := (x_1 + y_1, \dots, x_n + y_n)^\mathsf{T}$$

2. *Multiplikation mit Skalar*

$$\cdot : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n \quad \text{mit} \quad \lambda \cdot \mathbf{x} := (\lambda x_1, \dots, \lambda x_n)^\mathsf{T}$$

3. *Euklidisches Skalarprodukt*

$$\cdot : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \quad \text{mit} \quad \mathbf{x} \cdot \mathbf{y} := \mathbf{x}^\mathsf{T}\mathbf{y} = \sum_{k=1}^{n} x_k y_k = x_1 y_1 + \cdots + x_n y_n$$

4. *Euklidische Norm*

$$\| \, \| : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R} \quad \text{mit} \quad \|\mathbf{x}\| := \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{x_1^2 + \cdots + x_n^2}$$

Die Schreibweise für eine Funktionsdefinition mit mehreren Argumenten, zum Beispiel $f : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$, bedeutet, dass die Funktion f ein Element aus $\mathbb{R}$ und ein Element aus $\mathbb{R}^n$ auf ein Element aus $\mathbb{R}^n$ abbildet.

Mit dem Ausdruck $\mathbf{x}^\mathsf{T}\mathbf{y}$ kann man das Skalarprodukt $\mathbf{x} \cdot \mathbf{y}$ auch als *Matrixprodukt* schreiben, welches in unserem Exkurs Matrizen definiert wird. Wir empfehlen Ihnen, sich diesen durchzulesen, wenn Sie Ihre Matrixkenntnisse aus der Schule bzw. MaI1 ein wenig auffrischen möchten.

Das Multiplikationszeichen ist durch die obige Definition mehrfach definiert. Man muss daher aus den beiden Faktoren schließen, ob es sich um ein Skalarprodukt zweier Vektoren handelt ($\mathbf{x} \cdot \mathbf{y}$), um eine Multiplikation mit einem Skalar ($x \cdot \mathbf{y}$), oder einfach nur um eine "normale" Multiplikation reeller Zahlen ($x \cdot y$). Beachten Sie, dass man im Allgemeinen keine Multiplikation zwischen zwei Vektoren definiert, die wieder auf einen Vektor gleicher Dimension abbildet. Wichtige Ausnahmen sind das Kreuzprodukt im $\mathbb{R}^3$ (siehe unten) und die Multiplikation komplexer Zahlen.

Analog zu reellen Zahlen definieren wir die *Subtraktion* zweier Vektoren durch Addition mit dem negierten Vektor

$$\mathbf{x} - \mathbf{y} \;:=\; \mathbf{x} + (-\mathbf{y}) \;=\; (x_1 - y_1, \dots, x_n - y_n)^\mathsf{T}$$

und die Division eines Vektors durch einen Skalar als Multiplikation mit dessen Kehrwert

$$\frac{\mathbf{x}}{\lambda} \;:=\; \frac{1}{\lambda} \cdot \mathbf{x} \;=\; \left(\frac{x_1}{\lambda}, \dots, \frac{x_n}{\lambda}\right)^\mathsf{T}.$$

Mit der Vektoraddition **(1)** und der skalaren Multiplikation **(2)** wird der $\mathbb{R}^n$ zu einem *Vektorraum* — ein Begriff, der Ihnen in MaI-1 bestimmt schon begegnet ist. Da die Vektorraumeigenschaften direkt aus

den Körpereigenschaften der reellen Zahlen folgen und wenig überraschend sind, verzichten wir hier auf eine explizite Nennung. Mit dem euklidischen Skalarprodukt **(3)** und der dadurch induzierten euklidischen Norm **(4)** bekommen wir den *euklidischen Vektorraum* $\mathbb{R}^n$ und können (in beliebigen Dimensionen!) Winkel und Längen messen: Für zwei Vektoren $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ sind $\|\mathbf{x}\|$ und $\|\mathbf{y}\|$ ihre Längen, und es gilt

$$\mathbf{x} \cdot \mathbf{y} \;=\; \|\mathbf{x}\|\, \|\mathbf{y}\| \cos\left(\angle(\mathbf{x}, \mathbf{y})\right),$$

wobei $\angle(\mathbf{x}, \mathbf{y})$ den Winkeln zwischen den beiden Vektoren $\mathbf{x}$ und $\mathbf{y}$ bezeichnet. Das Skalarprodukt berechnet also den (skalierten) Cosinus des Winkels, so dass wir über den Arcuscosinus den Winkel bestimmen können:

$$\angle(\mathbf{x}, \mathbf{y}) \;=\; \arccos\left(\frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \cdot \|\mathbf{y}\|}\right).$$

Der Vollständigkeit halber erwähnen wir noch das *Kreuzprodukt*, welches wir aber **nur für den Spezialfall** $n = 3$ wie folgt definieren können

$$\times : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3 \quad \text{mit} \quad \mathbf{x} \times \mathbf{y} := \begin{pmatrix} x_2 y_3 - x_3 z_2 \\ x_3 y_1 - x_1 z_3 \\ x_1 y_2 - x_2 z_1 \end{pmatrix}.$$

Das Kreuzprodukt $\mathbf{z} = \mathbf{x} \times \mathbf{y}$ produziert einen Vektor, der sowohl auf $\mathbf{x}$ als auch auf $\mathbf{y}$ senkrecht steht, für den also gilt $\mathbf{z} \cdot \mathbf{x} = \mathbf{z} \cdot \mathbf{y} = 0$ (warum wohl?). Ebenso wie das Skalarprodukt hängt auch das Kreuzprodukt mit dem Winkel zwischen den beiden Vektoren zusammen:

$$\|\mathbf{x} \times \mathbf{y}\| \;=\; \|\mathbf{x}\|\, \|\mathbf{y}\| \sin\left(\angle(\mathbf{x}, \mathbf{y})\right).$$

Die euklidische Norm ist zwar die Standard-Norm im $\mathbb{R}^n$, aber es gibt noch die allgemeineren *p-Normen*, die je nach Anwedung besser geeignet sein können:

Definition 7.3 (p-Norm)

Wir definieren für $p \in [1, \infty)$ und $\mathbf{x} \in \mathbb{R}^n$ die *p-Norm* von $\mathbf{x}$ durch eine Funktion

$$\|\cdot\|_p : \mathbb{R}^n \to \mathbb{R} \quad \text{mit} \quad \|\mathbf{x}\|_p := \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}.$$

Für $p = 2$ erhalten wir die *euklidische Norm*

$$\|\mathbf{x}\|_2 = \|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2}$$

und für $p = \infty$ die *Maximumsnorm* oder *Tschebyschew-Norm*

$$\|\mathbf{x}\|_\infty := \max\left\{|x_1|, |x_2|, \ldots, |x_n|\right\}.$$

Für den Namen Tschebyschew gibt es in der Literatur mehrere Schreibweisen, zum Beispiel Tschebyscheff, Chebyshev oder Čebyšev. Sie können zu Übungszwecken beweisen, dass die Maximumsnorm ihren Namen verdient, denn es gilt

$$\lim_{p\to\infty} \|\mathbf{x}\|_p = \|\mathbf{x}\|_\infty.$$

Beispiel 7.4 (p-Norm)

Für $\mathbf{x} = (-6, 3, 2)^\mathsf{T}$ gilt

$$\begin{aligned} \|\mathbf{x}\|_1 &= |-6| + |3| + |2| = 11 \\ \|\mathbf{x}\|_2 &= \sqrt{(-6)^2 + 3^2 + 2^2} = 7 \\ \|\mathbf{x}\|_\infty &= \max\{|-6|, |3|, |2|\} = 6 \end{aligned}$$

Besonders die p-Normen für $p = 1, 2$ und die Maximumsnorm kommen in der Praxis häufig zum Einsatz. In der folgenden Demo können Sie mit der p-Norm für verschiedene Werte von p experimentieren.

▶ Demo: Einheitskreis bzgl. p-Normen

Mithilfe der p-Normen können wir eine Metrik (siehe Definition 2.18), also eine Abstandsfunktion für zwei Vektoren, definieren:

Satz 7.5 (Metrik im $\mathbb{R}^n$)

Für $p \geq 1$ definiert die Funktion $d_p : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ mit

$$d_p(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|_p$$

eine Metrik auf dem $\mathbb{R}^n$.

Da wir nun eine Möglichkeit haben, Abstände zwischen Vektoren zu bestimmen, können wir analog zum reellen Fall Folgen und Grenzwerte definieren, und damit schließlich die Stetigkeit von Funktionen.

Definition 7.6 (Folge im $\mathbb{R}^n$)

Unter einer Folge im $\mathbb{R}^n$ versteht man eine Abbildung, bei der jedem $k \in \mathbb{N}$ ein $\mathbf{x}^{(k)} \in \mathbb{R}^n$ zugeordnet wird. Wir schreiben für die Folge auch kurz $\left(\mathbf{x}^{(k)}\right)_{k\in\mathbb{N}}$ oder $\left(\mathbf{x}^{(k)}\right)$.

Vergleichen Sie die Definition mit Definition 3.1. Wir haben den Folgenindex k hier nach oben gestellt, damit dieser nicht mit dem Index für die k-te Komponente x_k des Vektors $\mathbf{x}$ verwechselt werden kann. So bezeichnet z.B. $x_2^{(5)}$ die zweite Komponente des fünften Folgengliedes. Wir können nun auch für Folgen im $\mathbb{R}^n$ einen Grenzwert definieren (vgl. Definition 3.5).

Definition 7.7 (Grenzwerte von Folgen im $\mathbb{R}^n$)

Eine Folge $\left(\mathbf{x}^{(k)}\right)$ heißt konvergent gegen den Grenzwert $\mathbf{x} \in \mathbb{R}^n$, wenn gilt

$$\lim_{k\to\infty} \left\|\mathbf{x}^{(k)} - \mathbf{x}\right\| = 0.$$

In dem Fall schreiben wir auch $\lim\limits_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x}$.

Da die Norm wieder auf eine reelle Zahl abbildet, konnten wir hier die vektoriellen Konvergenz auf eine reelle Nullfolge zurückführen. Zu beachten ist, dass wir anstelle des Betrages für reelle Folgen im n-dimensionalen die euklidische Norm verwenden. Man könnte sich fragen, warum wir keine andere p-Norm verwendet haben. Es lässt sich aber zeigen, dass eine Folge, welche bezüglich einer Norm gegen einen Grenzwert konvergiert, auch bezüglich jeder anderen Norm gegen denselben Grenzwert konvergiert. Daher werden wir im Folgenden zur Vereinfachung immer die euklidische Norm in Sätzen und Definitionen verwenden. Der Grenzwert ist, falls er existiert, eindeutig. Zur Grenzwertbestimmung können wir die Grenzwerte der einzelnen Vektorkomponenten bestimmen, wie der folgende Satz zeigt.

Satz 7.8

Eine Folge $\left(\mathbf{x}^{(k)}\right)$ im $\mathbb{R}^n$ konvergiert genau dann gegen einen Grenzwert $\mathbf{x}$, wenn jede Komponente $x_i^{(k)}$ gegen x_i konvergiert, also

$$\lim_{k\to\infty} \mathbf{x}^{(k)} = \mathbf{x} \quad \Leftrightarrow \quad \forall i \in \{1, \dots, n\} : \lim_{k\to\infty} x_i^{(k)} = x_i.$$

▼ Beweis

Wir zeigen zuerst, dass folgende Ungleichungskette gilt:

$$0 \;\leq\; \|\mathbf{x}\|_\infty \;\leq\; \|\mathbf{x}\| \;\leq\; \sqrt{n}\|\mathbf{x}\|_\infty. \qquad (*)$$

Einsetzen der Definitionen der Norm und Quadrieren der Ungleichungen liefert

$$0 \leq \max\left\{x_1^2, \dots, x_n^2\right\} \leq \sum_{i=1}^{n} x_i^2 \leq n \max\left\{x_1^2, \dots, x_n^2\right\}.$$

Sei $x_m = \max\left\{x_1^2, \dots, x_n^2\right\}$, dann ergibt sich

$$0 \leq x_m^2 \leq \sum_{i=1}^{n} x_i^2 \leq n x_m^2,$$

womit die Gültigkeit der Aussage offensichtlich ist. Damit lässt sich die Satzaussage jetzt sehr einfach beweisen.

⇐-Richtung:
Wenn die Folge komponentenweise konvergiert, dann auch bzgl. der Maximumsnorm, d.h.

$$\lim_{k\to\infty} \left\|\mathbf{x}^{(k)} - \mathbf{x}\right\|_\infty = 0.$$

Nach Sandwich-Theorem angewendet auf den ersten, dritten und vierten Term der Ungleichung $(*)$ gilt dann auch

$$\lim_{k\to\infty} \left\|\mathbf{x}^{(k)} - \mathbf{x}\right\| \to 0,$$

so dass die Folge bzgl. der euklidischen Norm (und daher bzgl. allen Normen) konvergiert.

⇒-Richtung:
Wenn umgekehrt die Folge konvergiert, also

$$\left\|\mathbf{x}^{(k)} - \mathbf{x}\right\| \to 0$$

gilt, dann folgt nach Sandwich-Theorem angewendet auf den ersten, zweiten und dritten Term der Ungleichung $(*)$ auch

$$\left\|\mathbf{x}^{(k)} - \mathbf{x}\right\|_\infty \to 0.$$

Wenn das Betragsmaximum aller Komponenten gegen Null konvergiert, dann konvergiert jede Komponente gegen Null, so dass jede Komponente der Folge konvergiert.

Beispiel 7.9

Wir betrachten die Folge $\mathbf{x}^{(k)} = \left(\frac{1}{k}, \frac{2k}{3k+4}\right)^\mathsf{T}$. Der Grenzwert dieser Folge ist $\left(0, \frac{2}{3}\right)^\mathsf{T}$, da

$$\lim_{k\to\infty} \frac{1}{k} = 0 \quad \text{und} \quad \lim_{k\to\infty} \frac{2k}{3k+4} = \frac{2}{3}.$$

▶ Visualisierung

Wie Sie sehen, hat es sich gelohnt, dass wir so viel Zeit mit der eindimensionalen Analysis verbracht haben, denn häufig lassen sich mehrdimensionale Probleme auf eindimensionale Fälle zurückführen. So lassen sich beispielsweise auch ganz analog Cauchy-Folgen im mehrdimensionalen definieren, jede konvergente Folge ist auch hier wieder eine Cauchy-Folge und jede Cauchy-Folge konvergiert im $\mathbb{R}^n$. Damit ist der $\mathbb{R}^n$ ein vollständiger metrischer Raum. Bei Interesse können Sie die entsprechenden Definitionen und Beweise in den einschlägigen Lehrbüchern zur Analysis 2 nachlesen.

Als nächstes definieren wir die Stetigkeit für multivariate und vektorwertige Funktionen, die vom $\mathbb{R}^n$ in den $\mathbb{R}^m$ abbilden. Vergleichen Sie dies mit Definition 4.18 und Definition 4.22, um zu sehen, dass es konsistent mit der eindimensionalen Definition ist (für $n = m = 1$).

Definition 7.10 (Stetigkeit im $\mathbb{R}^n$)

Seien $X \subseteq \mathbb{R}^n$ und $Y \subseteq \mathbb{R}^m$ zwei Mengen und $\mathbf{f} : X \to Y$ eine Abbildung von X nach Y.

Wir nennen $\mathbf{f}$ stetig in $\mathbf{a} \in X$, wenn für jede gegen $\mathbf{a}$ konvergente Folge $\left(\mathbf{x}^{(k)}\right)$ mit $\mathbf{x}^{(k)} \in X$ gilt

$$\lim_{k\to\infty} \mathbf{f}\left(\mathbf{x}^{(k)}\right) = \mathbf{f}(\mathbf{a}).$$

Wir nennen $\mathbf{f}$ stetig, wenn $\mathbf{f}$ stetig in jedem $\mathbf{a} \in X$ ist.

Wir nutzen übrigens auch für vektor- und skalarwertige Funktionen die optische Unterscheidung, dass wir erstere fett drucken ($\mathbf{f}$) und letztere nicht (f). Auch das ε-δ-Kriterium der Stetigkeit Satz 4.25

können wir auf mehrdimensionale Funktionen übertragen, wenn wir die darin vorkommenden Beträge durch Normen ersetzen:

Satz 7.11 (Epsilon-Delta-Kriterium für mehrdimensionale Funktionen)

Seien $X \subseteq \mathbb{R}^n$ und $Y \subseteq \mathbb{R}^m$ zwei Mengen und $\mathbf{f} : X \to Y$ eine Abbildung von X nach Y.

Die Funktion $\mathbf{f}$ ist genau dann im Punkt $\mathbf{a} \in X$ stetig, wenn

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; \forall \mathbf{x} \in X : \; \|\mathbf{x} - \mathbf{a}\| < \delta \;\Rightarrow\; \|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{a})\| < \varepsilon.$$

Da wir Grenzwerte komponentenweise betrachten können und für die reellwertigen Grenzwerte der Komponenten die bekannten Rechenregeln für Grenzwerte und stetige Funktionen (Satz 3.18 bzw. Satz 4.27) gelten, können wir auch im Mehrdimensionalen die Stetigkeit der Komponenten auf die Stetigkeit der gesamten Funktion übertragen, wie das folgende Beispiel zeigt.

Beispiel 7.12 (Stetigkeit mehrdimensionaler Funktionen 1)

Sei $\mathbf{f} : \mathbb{R}^2 \to \mathbb{R}^2$ mit

$$\mathbf{f}(\mathbf{x}) = \mathbf{f}(x, y) = \begin{pmatrix} \frac{xy}{1+x^2+y^2} \\ \exp(x)\sin(y) \end{pmatrix}.$$

Wir wollen zeigen, dass $\mathbf{f}$ stetig ist. Dazu sei $\left(\mathbf{x}^{(k)}\right)$ eine beliebige Folge, die gegen $\mathbf{a} = (x_a, y_a)^\mathsf{T}$ konvergiert. Wir bezeichnen mit (x_k) und (y_k) die reellen Folgen der beiden Komponenten, d.h., $\mathbf{x}^{(k)} = (x_k, y_k)^\mathsf{T}$. Nach Satz Satz 7.8 können wir die Grenzwerte der beiden Komponenten f_1 und f_2 von $\mathbf{f}$ einzeln betrachten, um den Grenzwert zu bestimmen:

$$\begin{aligned}
\lim_{k\to\infty} f_1\left(\mathbf{x}^{(k)}\right) &= \lim_{k\to\infty} \frac{x_k y_k}{1 + x_k^2 + y_k^2} \\
&= \frac{\left(\lim\limits_{k\to\infty} x_k\right)\left(\lim\limits_{k\to\infty} y_k\right)}{1 + \left(\lim\limits_{k\to\infty} x_k^2\right) + \left(\lim\limits_{k\to\infty} y_k^2\right)} \\
&= \frac{a_x a_y}{1 + a_x^2 + a_y^2} \\
&= f_1(\mathbf{a})
\end{aligned}$$

und

$$\begin{aligned}
\lim_{k\to\infty} f_2\left(\mathbf{x}^{(k)}\right) &= \lim_{k\to\infty} \exp(x_k)\sin(y_k) \\
&= \left(\lim_{k\to\infty} \exp(x_k)\right)\left(\lim_{k\to\infty} \sin(y_k)\right) \\
&= \exp(a_x)\sin(a_y) \\
&= f_2(\mathbf{a}),
\end{aligned}$$

wobei wir Satz 3.18 ausgenutzt haben, um die Grenzwerte in Summen, Produkte und Quotienten zu ziehen, und die Stetigkeit von $\exp(x)$, $\sin(x)$ und x^2 ausgenutzt haben, um die Grenzwertbestimmung durch Einsetzen vorzunehmen.

Damit folgt insgesamt

$$\lim_{k\to\infty} \mathbf{f}\left(\mathbf{x}^{(k)}\right) = \mathbf{f}(\mathbf{a})$$

und somit die Stetigkeit von $\mathbf{f}$.

Wie das letzte Beispiel zeigt, ist der Stetigkeitsnachweis kombinierter stetiger Funktionen sehr ähnlich zum eindimensionalen Fall. Anders sieht es bei stückweise definierten Funktionen aus, hier ist der Stetigkeitsnachweis im mehrdimensionalen deutlich schwieriger, da wir uns aus unendlich vielen Richtungen dem Grenzwert nähern können (und nicht nur aus zwei Richtungen, vgl. Satz 4.20).

Beispiel 7.13 (Stetigkeit mehrdimensionaler Funktionen 2)

Sei $f : \mathbb{R}^2 \to \mathbb{R}$ mit

$$f(\mathbf{x}) = f(x, y) = \begin{cases} \frac{xy}{x^2+y^2} & \text{für } \mathbf{x} \neq \mathbf{0}, \\ 0 & \text{sonst.} \end{cases}$$

Für $\mathbf{x} \neq \mathbf{0}$ weist man die Stetigkeit leicht wie im letzten Beispiel nach, da die Funktion hier nur aus stetigen Anteilen besteht.

Für $\mathbf{x} = \mathbf{0}$ ist die Funktion nicht stetig, denn für Funktionswerte der Nullfolge $\mathbf{x}^{(k)} = (1/k, 1/k)^\mathsf{T}$ ergibt sich

$$\lim_{k\to\infty} f\left(\mathbf{x}^{(k)}\right) = \lim_{k\to\infty} \frac{\frac{1}{k} \cdot \frac{1}{k}}{\frac{1}{k^2} + \frac{1}{k^2}} = \lim_{k\to\infty} \frac{\frac{1}{k^2}}{2\frac{1}{k^2}} = \frac{1}{2} \neq f(0, 0) = 0.$$

Dies ist nicht das einzige Gegenbeispiel: Wählt man z.B. $\mathbf{x}^{(k)} = (1/k, 2/k)^\mathsf{T}$, dann ist $\lim_{k\to\infty} f(\mathbf{x}^{(k)}) = 2/5$. Je nachdem, aus welcher "Richtung" man sich hier $\mathbf{0}$ nähert, ergibt sich ein anderer Grenzwert. Somit ist die Funktion immer unstetig, egal wie der Wert bei $\mathbf{0}$ definiert wird.

▶ Visualisierung

7.2 Parametrische Kurven und Flächen

In diesem Abschnitt nehmen wir uns die folgenden mehrdimensionalen Funktionstypen vor:

1. Abbildungen von $\mathbb{R}$ nach $\mathbb{R}^n$, welche wir als *parametrische Kurven* bezeichnen
2. Abbildungen von $\mathbb{R}^2 \to \mathbb{R}^n$, welche wir als *parametrische Flächen* bezeichnen

Bei beiden Funktionstypen interessiert uns besonders der Begriff der Ableitung(en), da diese eine interessante geometrische Bedeutung haben.

Kurven

Beginnen wir also mit der Definition einer Kurve.

Definition 7.14 (Parametrische Kurve)

Eine stetige Abbildung $\mathbf{f} : I \to \mathbb{R}^n$, wobei $I \subseteq \mathbb{R}$ ein eigentliches oder uneigentliches Intervall ist, bezeichnen wir als *Kurve*.

Beispiel 7.15 (Parametrische Kurven)

- Eine Gerade im durch einen Punkt $\mathbf{o} \in \mathbb{R}^n$ in Richtung $\mathbf{d} \in \mathbb{R}^n$ wird beschrieben durch die Kurve $\mathbf{f} : \mathbb{R} \to \mathbb{R}^n$ mit

$$\mathbf{f}(t) = \mathbf{o} + t\mathbf{d}.$$

▶ Visualisierung

- Ein Kreis vom Radius $r > 0$ wird im $\mathbb{R}^2$ beschrieben durch die Kurve $\mathbf{f} : [0, 2\pi] \to \mathbb{R}^2$ mit

$$\mathbf{f}(t) = \begin{pmatrix} r\cos(t) \\ r\sin(t) \end{pmatrix}.$$

▶ Visualisierung

- Die Funktion $\mathbf{f}(t) : [0, \infty) \to \mathbb{R}^2$ mit

$$\mathbf{f}(t) = \begin{pmatrix} t\cos(t) \\ t\sin(t) \end{pmatrix}.$$

beschreibt eine archimedische Spirale (Warum waren die nochmal cool?).

▶ Visualisierung

- Seien $r > 0$ und $c \neq 0$ reelle Zahlen. Die Kurve $\mathbf{f} : \mathbb{R} \to \mathbb{R}^3$ mit

$$\mathbf{f}(t) = \begin{pmatrix} r\cos(t) \\ r\sin(t) \\ ct \end{pmatrix}$$

beschreibt eine Schraubenlinie.

▶ Visualisierung

- Wir haben eigentlich schon sehr lange mit Kurven gearbeitet, ohne diese so zu nennen. Denn für eine "normale" Funktion $f : \mathbb{R} \to \mathbb{R}$ ist der Graph der Funktion $\mathbf{\Gamma}_f : \mathbb{R} \to \mathbb{R}^2$ mit

$$\mathbf{\Gamma}_f(x) = \begin{pmatrix} x \\ f(x) \end{pmatrix}$$

ebenfalls eine Kurve.

▶ Visualisierung

Wir verwenden hier häufig t als Funktionsargument, da man sich dieses oft als zeitliche Variable vorstellen kann: Zum Zeitpunkt t befindet sich ein Punkt auf der Kurve am Ort $\mathbf{f}(t)$. Als Nächstes definieren wir die Ableitung einer Kurve, die gut zu dieser geometrischen Anschauung passt.

Definition 7.16 (Differenzierbarkeit einer Kurve, Tangentialvektor)

Sei $\mathbf{f} : I \to \mathbb{R}^n$ eine Kurve. Wir nennen $\mathbf{f}$ differenzierbar, wenn jede Komponente $f_1(t), \dots, f_n(t)$ differenzierbar ist.

In diesem Fall nennen wir für ein $t \in I$ den Vektor

$$\mathbf{f}'(t) = (f_1'(t), f_2'(t), \dots, f_n'(t))^\mathsf{T}$$

den *Tangentialvektor* der Kurve $\mathbf{f}$ zum Parameterwert t.

Geometrische Interpretation: Den Tangentialvektor können wir als Limes von Sekanten auffassen:

$$\mathbf{f}'(t) \;=\; \lim_{h \to 0} \frac{\mathbf{f}(t+h) - \mathbf{f}(t)}{h}.$$

Physikalische Interpretation: Wenn $\mathbf{f}(t)$ einen Punkt auf der Kurve zum Zeitpunkt t beschreibt, dann ist der Tangentialvektor $\mathbf{f}'(t)$ der Geschwindigkeitsvektor dieses Punktes zum Zeitpunkt t und $\|\mathbf{f}'(t)\|$ ist der Betrag der Geschwindigkeit.

Beispiel 7.17 (Tangentialvektoren)

- Für einen Kreis vom Radius $r > 0$ sind Funktion und Tangentialvektor

$$\mathbf{f}(t) = \begin{pmatrix} r\cos(t) \\ r\sin(t) \end{pmatrix} \quad \text{und} \quad \mathbf{f}'(t) = \begin{pmatrix} -r\sin(t) \\ r\cos(t) \end{pmatrix}.$$

▶ Visualisierung

- Für die archimedische Spirale sind Funktion und Tangentialvektor

$$\mathbf{f}(t) = \begin{pmatrix} t\cos(t) \\ t\sin(t) \end{pmatrix} \quad \text{und} \quad \mathbf{f}'(t) = \begin{pmatrix} \cos(t) - t\sin(t) \\ \sin(t) + t\cos(t) \end{pmatrix}.$$

▶ Visualisierung

- Für eine Schraubenlinie sind Funktion und Tangentialvektor

$$\mathbf{f}(t) = \begin{pmatrix} r\cos(t) \\ r\sin(t) \\ ct \end{pmatrix} \quad \text{und} \quad \mathbf{f}'(t) = \begin{pmatrix} -r\sin(t) \\ r\cos(t) \\ c \end{pmatrix}$$

▶ Visualisierung

Wir wollen nun die Länge einer Kurve bestimmen. Dazu approximieren wir die Kurve zunächst mit mehreren Geradenabschnitten. Die aufsummierte Länge dieser Geradenabschnitte ist eine untere Schranke für die Kurvenlänge. Wenn wir die Abschnitte immer weiter unterteilen ergibt sich im Grenzwert für eine unendlich feine Unterteilung die Kurvenlänge. Sollte Sie dies an unsere Definition von Integralen erinnern, dann können Sie sich gedanklich auf die Schulter klopfen, denn tatsächlich ergibt sich als Endresultat eine Integralformel.

Satz 7.18 (Kurvenlänge durch Integration)

Jede stetig differenzierbare Kurve $\mathbf{f} : [a, b] \to \mathbb{R}^n$ ist die Länge der Kurve

$$L = \int_a^b \|\mathbf{f}'(t)\| \, \mathrm{d}t.$$

▶ Anschauliche Begründung

Beispiel 7.19 (Kreisbogenlänge)

Die Kurve $\mathbf{f} : [0, 2\pi] \to \mathbb{R}^2$ mit

$$\mathbf{f}(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

beschreibt einen Kreis vom Radius $r = 1$.

Für den Tangentialvektor gilt

$$\mathbf{f}'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}.$$

Die Bogenlänge (bzw. der Kreisumfang) beträgt damit

$$\begin{aligned} L &= \int_0^{2\pi} \|\mathbf{f}'(t)\| \, \mathrm{d}t \\ &= \int_0^{2\pi} \sqrt{(-\sin(t))^2 + (\cos(t))^2} \, \mathrm{d}t \\ &= \int_0^{2\pi} 1 \, \mathrm{d}t \\ &= 2\pi. \end{aligned}$$

Parametrische Flächen

Eine detaillierte Betrachtung allgemeiner vektorwertiger Funktionen, die vom $\mathbb{R}^n$ in den $\mathbb{R}^m$ abbilden, geht über den Vorlesungsstoff hinaus. Wir möchten hier nur die Grundidee für die Ableitung einer solchen Funktion möglichst anschaulich einführen und in der Vorlesung ein paar Anwendungsbeispiele zeigen, um eventuell Ihr Interesse für das Thema zu wecken. In der Praxis hat man es beispielsweise mit solchen Funktionen zu tun, wenn man die Eigenschaften von Oberflächen technischer Bauteile untersuchen will. In vielen Bereichen werden an diese Oberflächen sehr hohe Anforderungen gestellt (z.B. in der Luft- und Raumfahrt oder der Automobilbranche). Diese Anforderungen lassen sich in Eigenschaften der Ableitungen von Flächenfunktionen übersetzen. Ein anderes Anwendungsgebiet ist die Texturierung (oder genauer gesagt Parametrisierung) von 3D-Modellen. Hierbei werden Farbwerte einer zweidimensionalen Fläche (der Textur) auf ein dreidimensionales Objekt (das Modell) abgebildet. Wünschenswert sind Texturen, bei denen es zu möglichst wenig Verzerrung bei dieser Abbildung kommt. Diese Verzerrungen lassen sich ebenfalls mit Funktionseigenschaften assoziieren.

Wir Betrachten im Folgenden Beispiele für *parameterische Flächen*, also Abbildungen vom $\mathbb{R}^2$ in den $\mathbb{R}^3$. Für die beiden Koordinaten von $\mathbf{x} \in \mathbb{R}^2$ verwendet man häufig die Variablen $u := x_1$ und $v := x_2$, und für die Funktion schreiben wir dann entweder $\mathbf{f}(\mathbf{x})$ oder $\mathbf{f}(u, v)$.

Beispiel 7.20 (parameterische Flächen)

- Eine Kugeloberfläche vom Radius $r > 0$ wird beschrieben durch die Funktion $\mathbf{f} : [0, 2\pi] \times [0, \pi] \to \mathbb{R}^3$ mit

$$\mathbf{f}(u, v) = \begin{pmatrix} r\cos(u)\sin(v) \\ r\sin(u)\sin(v) \\ r\cos(v) \end{pmatrix}.$$

  Hierbei geben die beiden Intervalle in der Funktionsdefinition die Parameterbereiche für u und v an, also $u \in [0, 2\pi]$ und $v \in [0, \pi]$.

  Wenn wir $u = 0$ konstant wählen, dann ergibt sich eine Halbkreiskurve in der x-z-Ebene

$$\mathbf{f}(0, v) = \begin{pmatrix} r\sin(v) \\ 0 \\ r\cos(v) \end{pmatrix}.$$

  Umgekehrt ergibt sich für ein konstantes $v = \pi/2$ eine Kreiskurve in der x-y-Ebene

$$\mathbf{f}(u, \pi/2) = \begin{pmatrix} r\cos(u) \\ r\sin(u) \\ 0 \end{pmatrix}.$$

  Man kann sich die Entstehung der Fläche so vorstellen, als dass die Halbkreiskurve sich einmal um 360° dreht und damit die Kreisoberfläche überstreicht.

  ▶ Visualisierung

- Eine Zylindermanteloberfläche vom Radius $r > 0$ und Höhe $h > 0$ wird beschrieben durch die Funktion $\mathbf{f} : [0, 2\pi] \times [0, h] \to \mathbb{R}^3$ mit

$$\mathbf{f}(u, v) = \begin{pmatrix} r\cos(u) \\ r\sin(u) \\ v \end{pmatrix}.$$

  Auch hier lässt sich eine Kreiskurve in der x-y-Ebene erkennen, die in z-Richtung durch Variation von v die Zylindermantelfläche überstreicht.

  ▶ Visualisierung

- Die Oberfläche eines Torus mit Innenradius $r > 0$ und Außenradius $R > 0$ wird beschrieben durch die Funktion $\mathbf{f} : [0, 2\pi] \times [0, 2\pi] \to \mathbb{R}^3$ mit

$$\mathbf{f}(u, v) = R \cdot \begin{pmatrix} \cos(u) \\ \sin(u) \\ 0 \end{pmatrix} + r \cdot \begin{pmatrix} \cos(u)\cos(v) \\ \sin(u)\cos(v) \\ \sin(v) \end{pmatrix}.$$

Auch hier lässt sich die Funktion zerlegen in eine Kreiskurve vom Radius r, deren Mittelpunkt sich durch Variation von u auf einer zur ersten Kreisebene senkrechten Kreisbahn vom Radius R bewegt, sodass die Kreiskurve somit die Torusoberfläche überstreicht.

▶ Visualisierung

Die letzten Beispiele zeigen, dass man sich eine parametrische Fläche auch als Kurve vorstellen kann, welche durch den zweiten Parameter variiert wird und dabei die sich ergebende Fläche überstreicht. Also zum Beispiel eine Kurve entlang des Paramters u, bei der sich jeder Punkt entlang einer zweiten Kurve mit Parameter v bewegt. In diesem Kontext bedeutet die Stetigkeit der Funktion, dass das Parametergebiet bei der Abbildung nicht auseinandergerissen wird, sondern eine zusammenhängende Fläche bleibt.

Wir würden auch gerne die "Knickfreiheit", also Differenzierbarkeit einer solchen Fläche mathematisch überprüfen können. Durch die Interpretation, dass eine Fläche nichts anderes ist, als die Variation einer Kurve, können wir auch Ableitungen parametrischer Flächen auf Ableitungen parametrischer Kurven zurückführen: Wir halten erst den Parameter v konstant und bestimmen die Ableitung der sich ergebenden u-Kurve, was die u-Tangente ergibt. Anschließend halten wir den Parameter u konstant und betrachten die Ableitung der sich ergebenden v-Kurve, was die v-Tangente ergibt. Die u- und v-Tangenten spannen am aktuellen Punkt $\mathbf{f}(u,v)$ die *Tangentialebene* $T_{\mathbf{f}}$ von $\mathbf{f}$ auf, die eine Verallgemeinerung der Tangente einer univariaten Funktion darstellt:

$$T_{\mathbf{f}(u,v)}(s,t) \;=\; \mathbf{f}(u,v) + s\,\mathbf{t}_u + t\,\mathbf{t}_v.$$

Beispiel 7.21 (Tangentialebene)

- Wir untersuchen die Mantelfläche eines Zylinders vom Radius $r > 0$ und Höhe $h > 0$, also $\mathbf{f} : [0, 2\pi] \times [0, h] \to \mathbb{R}^3$ mit

$$\mathbf{f}(u, v) = \begin{pmatrix} r\cos(u) \\ r\sin(u) \\ v \end{pmatrix}.$$

  Wir erhalten die u-Tangente, indem wir v als Konstante auffassen und die resultierende u-Kurve nach u ableiten:

$$\mathbf{t}_u = \frac{\partial \mathbf{f}}{\partial u}(u, v) = \begin{pmatrix} -r\sin(u) \\ r\cos(u) \\ 0 \end{pmatrix}.$$

  Die merkwürdige Notation $\frac{\partial \mathbf{f}}{\partial u}$ bezeichnet die *partielle Ableitung*, welche wir aber erst im nächsten Kapitel einführen werden. Analog erhalten wir die v-Tangente durch Festhalten von u und Ableiten nach v:

$$\mathbf{t}_v = \frac{\partial \mathbf{f}}{\partial v}(u, v) = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

  Die Tangentialebene wird von den beiden Tangentialvektoren $\mathbf{t}_u$ und $\mathbf{t}_v$ aufgespannt und verläuft damit stets parallel zur z-Achse, und abhängig von u auch parallel zum Vektor $(-r\sin(u), r\cos(u), 0)^\mathsf{T}$. Für $u = v = 0$ ergibt sich als Tangentialebene beispielsweise die y-z-Ebene.

  ▶ Visualisierung

Die Existenz der einzelnen Tangentialvektoren reicht leider nicht aus, um die "Knickfreiheit" (Differenzierbarkeit) einer solchen Fläche zu garantieren. Allerdings kann man zeigen, dass aus der Stetigkeit aller Einträge der Tangentialvektoren auch die Differenzierbarkeit der Funktion folgt.

7.3 Extrema im $\mathbb{R}^n$

In diesem Kapitel beschäftigen wir uns mit multivariaten skalarwertigen Funktionen $f\colon \mathbb{R}^n \to \mathbb{R}$, die n Eingabevariablen $\mathbf{x} = (x_1, \dots, x_n)^\mathsf{T}$ auf einen Ausgabewert $f(\mathbf{x}) = f(x_1, \dots, x_n)$ abbilden. Solche Funktionen werden Ihnen im Studium häufig begegnen, zum Beispiel beim maschinellen Lernen, wo ein Modell an Daten gefittet wird, indem eine Kostenfunktion (oder Loss-Function) f durch geschickte Wahl der Modellparameter $x_1, \dots, x_n$ minimiert werden soll. Weitere Beispiele gibt es im Kontext der Optimierung, wo auch entweder (unerwünschte) Kosten minimiert oder ein (gewünschter) Ertrag maximiert werden soll.

Wir sind also an der Bestimmung der Extrema der Funktion $f(\mathbf{x})$ interessiert und werden hieraus erneut Bedingungen für erste und zweite Ableitungen definieren. Auch hier kann eine Visualisierung der Funktionen helfen, die Zusammenhänge besser zu verstehen. Aus diesem Grund betrachten wir in diesem Kapitel häufig Funktionen $f : \mathbb{R}^2 \to \mathbb{R}$, da wir diese noch einfach als Graph (in diesem Fall eine parametrische Fläche) visualisieren können:

$$\mathbf{\Gamma}_f(x, y) = \begin{pmatrix} x \\ y \\ f(x, y) \end{pmatrix}.$$

Dabei werden die Funktionswerte $f(x, y)$ auf der z-Achse aufgetragen, so dass sich ein Höhenfeld über der x-y-Ebene ergibt. Wie in der folgenden Demo gezeigt, können wir diese Höhenfelder als Fläche, als Höhenlinien (ähnlich wie bei Geländekarten) und durch Farbkodierungen der z-Werte darstellen.

▶ Demo: Funktion als Höhenfeld plotten

Damit wir die Extrema, also Minima und Maxima, von multivariaten Funktionen bestimmen können, müssen wir diese zunächst definieren. Wir verallgemeiner dafür die Definition von lokalen Extrema univariater Funktionen:

Definition 7.22 (Lokale Extrema im $\mathbb{R}^n$)

Eine Funktion $f: U \subseteq \mathbb{R}^n \to \mathbb{R}$ nimmt in $\mathbf{x} \in U$ ein *lokales Maximum* (bzw. *lokales Minimum*) $f(\mathbf{x})$ an, wenn ein $\varepsilon > 0$ existiert, sodass $f(\mathbf{x}) \geq f(\mathbf{y})$ (bzw. $f(\mathbf{x}) \leq f(\mathbf{y})$) für alle $\mathbf{y}$ mit $\|\mathbf{x} - \mathbf{y}\| < \varepsilon$.

Wenn Sie diese Definition mit Definition 5.15 vergleichen, wird Ihnen wieder auffallen, dass hier lediglich die Beträge durch Normen und die skalaren Variablen durch Vektoren ersetzt wurden. Für $n = 1$ ergibt sich aus Definition 7.22 der eindimensionale Fall, also Definition 5.15.

Im univariaten Fall haben wir Extrema bestimmt, indem wir die erste und zweite Ableitung untersucht haben: Die erste Ableitung muss verschwinden (Satz 5.16), und wenn zusätzlich die zweite Ableitung größer bzw. kleiner als Null ist, dann liegt ein Minimum bzw. Maximum vor (Satz 5.20). Bei multivariaten Funktionen ist das ganz ähnlich, nur dass wir hier mehrere Variablen haben, nach denen wir die Funktion ableiten können. Wir brauchen also zunächst eine Verallgemeinerung der ersten Ableitung:

Definition 7.23 (Partielle Ableitung)

Eine Funktion $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ ist im Punkt $\mathbf{x} = (x_1, x_2, \dots, x_n)^\mathsf{T} \in U$ partiell differenzierbar bezüglich der i-ten Koordinate, falls der Grenzwert

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) := \lim_{h \to 0} \frac{f(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n) - f(\mathbf{x})}{h}$$

existiert. In diesem Fall heißt der Grenzwert die i-te *partielle Ableitung* von f.

Wenn alle partiellen Ableitungen $\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n}$ für alle $\mathbf{x} \in U$ existieren, heißt die Funktion *partiell differenzierbar*. Sind die partiellen Ableitungen zusätzlich stetig, heißt f *stetig partiell differenzierbar* oder C^1-stetig.

Das Bestimmen von partiellen Ableitungen ist eigentlich ganz einfach. Wenn wir eine univariate Funktion $f_i : \mathbb{R} \to \mathbb{R}$ definieren mit

$$f_i(\xi) := f(x_1, \dots, x_{i-1}, \xi, x_{i+1}, \dots, x_n),$$

die von der Funktion f alle Parameter außer dem i-ten Parameter $x_i = \xi$ festhält, dann entspricht die i-te partielle Ableitung von $f(\mathbf{x})$ gerade der herkömmlichen Ableitung von $f_i(\xi)$:

$$\frac{\partial f}{\partial x_i}(x_1, \dots, x_n) \;=\; \lim_{h \to 0} \frac{f_i(x_i + h) - f_i(x_i)}{h} \;=\; f_i'(x_i).$$

Eine Funktion partiell abzuleiten ist also nichts anderes, als alle Parameter bis auf den, nach dem gerade abgeleitet wird, als Konstanten aufzufassen und ansonsten wie bei einer univariaten Funktion zu differenzieren.

Der Wert der i-ten partiellen Ableitung gibt die Steigung der Tangente an den Funktionsgraphen in Richtung der i-ten Koordinatenachse an. Das haben wir bei parametrischen Flächen schon in Form der u- und v-Tangenten gesehen, die als partielle Ableitungen der Fläche nach u und v definiert waren.

▶ Demo: Partielle Ableitungen

Im Gegensatz zu univariaten Funktion folgt bei multivariaten Funktionen aus der partiellen Differenzierbarkeit nicht die Stetigkeit der Funktion. Erst wenn die Funktion *stetig* partiell Differenzierbar ist, also alle partiellen Ableitungen stetig sind, ist die Funktion auch stetig, hat also keine Sprünge.

Beispiel 7.24 (partielle Ableitungen)

Sei $f : \mathbb{R}^3 \to \mathbb{R}$ mit $f(x, y, z) = 2x^2 + 3xy + z$. Die partiellen Ableitungen lauten

$$\frac{\partial f}{\partial x}(x, y, z) = 4x + 3y$$
$$\frac{\partial f}{\partial y}(x, y, z) = 3x$$
$$\frac{\partial f}{\partial z}(x, y, z) = 1$$

Wenn wir alle partiellen Ableitungen in einen Vektor schreiben, erhalten wir den *Gradienten* der Funktion f, dessen geometrische Bedeutung wir im nächsten Kapitel kennenlernen werden.

Definition 7.25 (Gradient)

Für eine partiell differenzierbare Funktion $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ heißt der Vektor

$$\nabla f(\mathbf{x}) := \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix} \in \mathbb{R}^n$$

aus allen n partiellen Ableitungen der *Gradient* von f am Punkt $\mathbf{x}$.

Auch wenn wir es in diesem Kapitel nicht brauchen, definieren wir der Vollständigkeit halber noch die Verallgemeinerung des Gradienten für multivariate und *vektorwertige* Funktionen, die sogenannte *Jacobi-Matrix*:

Definition 7.26 (Jacobi-Matrix)

Für eine partiell differenzierbare Funktion $\mathbf{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ mit Komponenten $f_1, \dots, f_m$, also

$$\mathbf{f}(\mathbf{x}) \;=\; (f_1(\mathbf{x}), \dots, f_m(\mathbf{x}))^\mathsf{T},$$

heißt die $(m \times n)$-Matrix

$$\mathbf{J}_\mathbf{f}(\mathbf{x}) \;:=\; \begin{pmatrix} \frac{\partial f_1(\mathbf{x})}{\partial x_1} & \frac{\partial f_1(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_1(\mathbf{x})}{\partial x_n} \\ \frac{\partial f_2(\mathbf{x})}{\partial x_1} & \frac{\partial f_2(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_2(\mathbf{x})}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m(\mathbf{x})}{\partial x_1} & \frac{\partial f_m(\mathbf{x})}{\partial x_2} & \cdots & \frac{\partial f_m(\mathbf{x})}{\partial x_n} \end{pmatrix}$$

die *Jacobi-Matrix* von $\mathbf{f}$ am Punkt $\mathbf{x}$.

Wenn wir die beiden letzten Definitionen vergleichen, fällt auf, dass in der i-ten Zeile der Jacobi-Matrix gerade der (transponierte) Gradient ∇f_i der i-ten Komponente von $\mathbf{f}$ steht.

Im Folgenden beschränken wir uns aber wieder auf multivariate skalarwertige Funktionen.

Beispiel 7.27 (Gradient)

1. Sei $f : \mathbb{R}^3 \to \mathbb{R}$ mit $f(x, y, z) = 2x^2 + 3xy + z$ die Funktion aus dem vorherigen Beispiel. Der Gradient dieser Funktion ist

$$\nabla f(x, y) = \begin{pmatrix} 4x + 3y \\ 3x \\ 1 \end{pmatrix}.$$

2. Sei $f : \mathbb{R}^n \to \mathbb{R}$ mit

$$f(\mathbf{x}) = \|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

f berechnet also den Abstand des Punktes $\mathbf{x}$ vom Ursprung bzw. die Länge des Vektors $\mathbf{x}$, je nachdem, ob wir $\mathbf{x}$ als Punkt oder Vektor auffassen. Am Punkt $\mathbf{x} = \mathbf{0}$ ist die Funktion nicht partiell differenzierbar. Für $\mathbf{x} \neq \mathbf{0}$ ergeben sich (mit Kettenregel) die partiellen Ableitungen

$$\frac{\partial f}{\partial x_i} = \frac{1}{2\|\mathbf{x}\|} 2x_i = \frac{x_i}{\|\mathbf{x}\|}.$$

und somit der Gradient

$$\nabla f(\mathbf{x}) = \frac{\mathbf{x}}{\|\mathbf{x}\|}.$$

Kommen wir nun zum Zusammenhang zwischen lokalen Extrema und partiellen Ableitungen. Zunächst liefert der folgende Satz eine notwendige Bedingung für ein Extremum (analog zu Satz 5.16):

Satz 7.28 (Notwendige Bedingung für lokales Extremum im $\mathbb{R}^n$)

Sei $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ eine partiell differenzierbare Funktion. Besitzt f im Punkt $\mathbf{x} \in U$ ein lokales Extremum, dann ist

$$\nabla f(\mathbf{x}) = \mathbf{0} = (0, \dots, 0)^\mathsf{T}.$$

▼ Beweis

Definiere die univariaten Funktionen

$$f_i(\xi) := f(x_1, \dots, x_{i-1}, \xi, x_{i+1}, \dots, x_n)$$

für $i = 1, \dots, n$. Wenn f in $\mathbf{x} = (x_1, x_2, ..., x_i, ..., x_n)^\mathsf{T}$ ein lokales Extremum hat, dann haben auch die Teilfunktionen f_i in x_i ein lokales Extremum. Nach Satz 5.16 gilt dann $f_i'(x_i) = 0$. Diese Ableitungen sind aber genau die partiellen Ableitung von f. Es gilt also

$$\forall i \in \{1, \dots, n\} : \frac{\partial f}{\partial x_i}(\mathbf{x}) = f_i'(x_i) = 0 \quad \Leftrightarrow \quad \nabla f(\mathbf{x}) = \mathbf{0}.$$

□

Genau genommen darf $\mathbf{x} \in U$ in der obigen Definition kein Randpunkt von U sein. Es muss ein $\varepsilon > 0$ existieren, sodass alle $\mathbf{y} \in \mathbb{R}^n$ mit $\|\mathbf{x} - \mathbf{y}\| < \varepsilon$ auch in U liegen. An Randpunkten kann ansonsten ein Extremum vorliegen, ohne dass die notwendige Bedingung gilt (daher hatten wir auch in Definition 5.15 ein offenes Intervall verwendet).

Beispiel 7.29 (Notwendige Bedingung)

1. Die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = x^2 + y^2$ nimmt in $(0, 0)$ ein lokales (und auch globales) Minimum an, da die Funktion an jeder anderen Stelle strikt positiv ist. Daher muss hier auch die notwendige Bedingung gelten. Wir bestimmen den Gradienten

$$\nabla f(x, y) = (2x, 2y)^\mathsf{T}$$

und es gilt wie erwartet

$$\nabla f(0, 0) = (0, 0)^\mathsf{T}.$$

▶ Visualisierung

2. Wir untersuchen die Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = x^2 - y^2$. Der Gradient ist

$$\nabla f(x, y) = (2x, -2y)^\mathsf{T}$$

und verschwindet daher an $(0, 0)$. Allerdings nimmt die Funktion in $(0, 0)$ kein lokales Extremum, sondern eine Sattelstelle an, wie die Visualisierung zeigt. Dies verdeutlicht, dass die *notwendige* Bedingung von Satz 7.28 eben nicht hinreichend ist.

▶ Visualisierung

Man kann sich die notwendige Bedingung auch am Graphen einer Funktion mit zwei Parametern verdeutlichen. Dieser ist, wie schon weiter oben geschrieben, die parametrische Fläche $\mathbf{\Gamma}_f : \mathbb{R}^2 \to \mathbb{R}^3$ mit

$$\mathbf{\Gamma}_f(u, v) = \begin{pmatrix} u \\ v \\ f(u, v) \end{pmatrix}.$$

Die Tangentialebene dieser Fläche wird durch die beiden Tangenten in u- und v-Richtung aufgespannt. Diese Tangenten werden durch die partiellen Ableitungen nach u und v bestimmt:

$$\frac{\partial \mathbf{\Gamma}_f}{\partial u}(u, v) = \begin{pmatrix} 1 \\ 0 \\ \frac{\partial f}{\partial u}(u, v) \end{pmatrix} \quad \text{und} \quad \frac{\partial \mathbf{\Gamma}_f}{\partial v}(u, v) = \begin{pmatrix} 0 \\ 1 \\ \frac{\partial f}{\partial v}(u, v) \end{pmatrix}$$

Die z-Komponenten der beiden Tangenten sind die partiellen Ableitungen von f. An einem Extremum verschwinden diese, sodass die Tangentialebene parallel zur xy-Ebene ist. Das kann man in der folgenden Demo-App gut ausprobieren.

▶ Demo: Partielle Ableitungen an Extremstellen

Wir haben im obigen Beispiel gesehen, dass die notwendige Bedingung für ein Extremum nicht hinreichend ist. Analog zum univariaten Fall benötigen wir hier zweite Ableitungen für eine hinreichende Bedingung.

Definition 7.30 (Zweite Ableitungen im $\mathbb{R}^n$)

Sei $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ eine C^1-stetige Funktion. Die *partielle Ableitung zweiter Ordnung* in die Raumrichtungen i und j an der Stelle $\mathbf{x}$ ist

$$\frac{\partial^2 f}{\partial x_j \, \partial x_i}(\mathbf{x}) \;:=\; \frac{\partial \left(\frac{\partial f}{\partial x_i} \right)}{\partial x_j}(\mathbf{x}).$$

Wenn wir zweimal nach derselben Variable x_i ableiten, schreiben wir auch

$$\frac{\partial^2 f}{\partial x_i^2}(\mathbf{x}) \;:=\; \frac{\partial^2 f}{\partial x_i \, \partial x_i}(\mathbf{x}).$$

Existieren alle partiellen Ableitungen zweiter Ordnung an allen $\mathbf{x} \in U$, so ist f zweimal partiell differenzierbar. Sind alle partiellen Ableitungen zweiter Ordnung zusätzlich stetig, so ist f zweimal stetig partiell differenzierbar oder C^2-stetig.

Da wir bei n Variablen jede der n partiellen Ableitungen noch einmal nach allen n Variablen partiell ableiten können, erhalten wir insgesamt n^2 partielle Ableitungen zweiter Ordnung erhalten. Diese können wir in der sogenannten *Hesse-Matrix* anordnen:

Definition 7.31 (Hesse-Matrix)

Sei $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ eine C^2-stetige Funktion. Die aus den partiellen Ableitungen zweiter Ordnung gebildete quadratische Matrix

$$\nabla^2 f(\mathbf{x}) \;:=\; \begin{pmatrix} \frac{\partial^2 f(x)}{\partial x_1^2} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f(x)}{\partial x_n^2} \end{pmatrix}$$

nennen wir die *Hesse-Matrix* der Funktion f an der Stelle $\mathbf{x}$.

Übrigens: Die Hesse-Matrix ist die Jacobi-Matrix des Gradienten von von f.

Beispiel 7.32 (Hesse-Matrix)

Sei $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = 2xy + 3x + 4x^2y$. Die ersten partiellen Ableitungen lauten

$$\frac{\partial f(x,y)}{\partial x} = 2y + 3 + 8xy \quad \text{und} \quad \frac{\partial f(x,y)}{\partial y} = 2x + 4x^2.$$

Erneutes partielles Ableiten nach x bzw. y liefert die partiellen Ableitungen zweiter Ordnung

$$\frac{\partial^2 f(x,y)}{\partial x^2} = 8y, \quad \frac{\partial^2 f(x,y)}{\partial x \partial y} = 2 + 8x, \quad \frac{\partial^2 f(x,y)}{\partial y \partial x} = 2 + 8x, \quad \frac{\partial^2 f(x,y)}{\partial y^2} = 0,$$

und damit die Hessematrix

$$\nabla^2 f(x,y) = \begin{pmatrix} 8y & 2+8x \\ 2+8x & 0 \end{pmatrix}.$$

Beim letzten Beispiel ist Ihnen vielleicht aufgefallen, dass die gemischten zweiten Ableitungen identisch waren, wodurch die Hesse-Matrix symmetrisch wurde. Dies ist nicht zufällig so, sondern nach dem *Satz von Schwarz* für C^2-stetig Funktionen immer der Fall.

Satz 7.33 (Satz von Schwarz)

Sei $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ eine C^2-stetige Funktion. Dann gilt

$$\frac{\partial^2 f}{\partial x_i \, \partial x_j}(\mathbf{x}) = \frac{\partial^2 f}{\partial x_j \, \partial x_i}(\mathbf{x}) \quad \forall i, j \in \{1, \dots, n\}.$$

Die hinreichende Bedingung für ein lokales Extremum ist bei multivariaten Funktionen leider weniger handlich als bei univariaten Funktionen, wo wir nach Satz 5.20 aus dem Vorzeichen der zweiten Ableitung (zusammen mit der notwendigen Bedingung $f'(x) = 0$) die Existenz und Art eines lokalen Extremums bestimmen konnten: Minimum falls $f''(x) > 0$, Maximum falls $f''(x) < 0$, und bei $f''(x) = 0$ keine Aussage möglich.

Für multivariate Funktionen übernimmt die Hesse-Matrix die Rolle der zweiten Ableitung, welche für ein Minimum bzw. Maximum positiv definit bzw. negativ definit sein muss.

Definition 7.34 (Definitheit einer Matrix)

Eine Matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$ ist

- *positiv definit*, wenn $\forall \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\} : \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} > 0$
- *positiv semidefinit*, wenn $\forall \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\} : \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} \geq 0$
- *negativ definit*, wenn $\forall \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\} : \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} < 0$
- *negativ semidefinit*, wenn $\forall \mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\} : \mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} \leq 0$
- *indefinit*, wenn $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ existieren mit $\mathbf{x}^\mathsf{T}\mathbf{A}\mathbf{x} < 0$ und $\mathbf{y}^\mathsf{T}\mathbf{A}\mathbf{y} > 0$.

Die Definitheit einer Matrix hängt mit den Eigenwerten der Matrix zusammen (die Sie in MafI1 gelernt haben sollten): Bei einer positiv definiten Matrix sind alle Eigenwerte positiv, bei einer negativ definiten Matrix sind sie negativ, und bei einer indefiniten Matrix gibt es sowohl positive als auch negative Eigenwerte. Bei einer positiv (bzw. negativ) semidefiniten Matrix sind die Eigenwerte ≥ 0 (bzw. ≤ 0).

Wenn man die Definitheit einer Matrix als mehrdimensionales Analogon zu einem Vorzeichen akzeptiert (für $n = 1$ ergibt sich tatsächlich daraus eine einfache Vorzeichenbetrachtung von $\mathbf{A}$), dann ist auch die hinreichende Bedingung nicht weiter überraschend. Vergleichen Sie diese mit Satz 5.20.

Satz 7.35 (Hinreichende Bedingung für lokale Extrema im $\mathbb{R}^n$)

Sei $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ eine C^2-stetige Funktion.

Falls $\nabla f(\mathbf{x}) = \mathbf{0}$ und die Hesse-Matrix $\nabla^2 f(\mathbf{x})$ in $\mathbf{x}$ positiv (bzw. negativ) definit ist, dann nimmt f in $\mathbf{x}$ ein lokales Minimum (bzw. Maximum) an.

Ist die Hesse-Matrix in $\mathbf{x}$ indefinit, so liegt kein Extremum vor.

Wenn die Hesse-Matrix semidefinit ist, kann keine Aussage getroffen werden. Übrigens hängt die Definitheit der Hesse-Matrix auch mit der Konvexität der Funktion zusammen (vgl. Satz 5.23). Ist $\nabla^2 f(\mathbf{x})$ für alle $\mathbf{x} \in U \subseteq \mathbb{R}^n$ positiv (bzw. negativ) semidefinit, so ist f auf U konvex (bzw. konkav) und lokale Minima (bzw. Maxima) sind auch globale Minima (bzw. Maxima).

Da die Bestimmung der Definitheit einer Matrix im Allgemeinen mehrere Wochen einer Veranstaltung zur linearen Algebra füllen kann, beschränken wir uns hier auf den zweidimensionalen Fall, der bei Funktionen $f : U \subseteq \mathbb{R}^2 \to \mathbb{R}$ anwendbar ist. Hier gibt es eine handliche Regel für die Definitheit der Hesse-Matrix. Der Fall von $(n \times n)$-Matrizen wird im Exkurs Matrizen erklärt.

Satz 7.36 (Definitheit für 2×2-Matrizen)

Eine symmetrische (2×2)-Matrix $\mathbf{A} = \begin{pmatrix} a & c \\ c & b \end{pmatrix} \in \mathbb{R}^{2\times 2}$ ist

- genau dann positiv definit, wenn $a > 0 \wedge |\mathbf{A}| = ab - c^2 > 0$
- genau dann negativ definit, wenn $a < 0 \wedge |\mathbf{A}| = ab - c^2 > 0.$
- genau dann indefinit, wenn $|\mathbf{A}| = ab - c^2 < 0.$

Wenn Sie nicht mehr wissen, was die Determinante $|\mathbf{A}|$ einer Matrix $\mathbf{A}$ ist, schauen Sie im Exkurs Matrizen (oder Ihren MafI1-Unterlagen) nach.

Beispiel 7.37 (Bestimmung von Extrema)

1. Für $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = x^2 + y^2 + 1$ ergibt sich

$$\nabla f(x, y) = \begin{pmatrix} 2x \\ 2y \end{pmatrix} \quad \text{und} \quad \nabla^2 f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

   Weil $\nabla f(0, 0) = 0$ und $\nabla^2 f$ für alle (x, y) positiv definit ist, hat f in $(0, 0)^\mathsf{T}$ ein lokales (und globales) Minimum. Der Graph von f ist ein nach oben geöffneter Paraboloid.

   ▶ Visualisierung

2. Für $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = 42 - x^2 - y^2$ ergibt sich

$$\nabla f(x, y) = \begin{pmatrix} -2x \\ -2y \end{pmatrix} \quad \text{und} \quad \nabla^2 f(x, y) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}.$$

   Weil $\nabla f(0, 0) = 0$ und $\nabla^2 f$ für alle (x, y) negativ definit ist, hat f in $(0, 0)^\mathsf{T}$ ein lokales (und globales) Maximum. Der Graph von f ist ein nach unten geöffneter Paraboloid.

   ▶ Visualisierung

3. Für $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = 3 + x^2 - y^2$ ergibt sich

$$\nabla f(x, y) = \begin{pmatrix} 2x \\ -2y \end{pmatrix} \quad \text{und} \quad \nabla^2 f(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}.$$

   Weil $\nabla f(0, 0) = 0$ und $\nabla^2 f$ indefinit ist, hat f in $(0, 0)^\mathsf{T}$ kein lokales Extremum. Der Graph von f ist eine sogenannte Sattelfläche.

   ▶ Visualisierung

4. Für $f : \mathbb{R}^2 \to \mathbb{R}$ mit $f(x, y) = x^3 + y^3 - 3xy$ ergibt sich

$$\nabla f(x, y) = \begin{pmatrix} 3x^2 - 3y \\ 3y^2 - 3x \end{pmatrix} \quad \text{und} \quad \nabla^2 f(x, y) = \begin{pmatrix} 6x & -3 \\ -3 & 6y \end{pmatrix}.$$

   Der Gradient verschwindet, wenn $x^2 = y$ und $y^2 = x$, also wenn $x^4 = x$, woraus sich die zwei Lösungen $(0, 0)^\mathsf{T}$ und $(1, 1)^\mathsf{T}$ ergeben. An diesen beiden Stellen können Extrema angenommen werden (müssen aber nicht). Wir setzen die beiden Punkte in die Hesse-Matrix ein. Für $(0, 0)$ ergibt sich

$$\nabla^2 f(0, 0) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}.$$

   Diese Matrix ist indefinit, da $\left|\nabla^2 f(0, 0)\right| = -9 < 0$. Somit ist $(0, 0)$ keine Extremstelle. Für den zweiten Punkt ergibt sich

$$\nabla^2 f(1,1) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}.$$

Diese Matrix ist positiv definit, da der obere linke Eintrag positiv ist und $\left|\nabla^2 f(1,1)\right| = 27 > 0$. Damit nimmt f in $(1,1)^\mathsf{T}$ ein lokales Minimum an.

▶ Visualisierung

7.4 Optimierung multivariater Funktionen

Im letzten Beispiel haben wir mit einfachen Funktionen gearbeitet, für die wir die lokalen Extrema dann auch recht einfach ausrechnen konnten. In der Praxis sind die Probleme aber meistens deutlich komplexer: Die zu minimierenden Funktionen sind viel komplizierter als simple quadratischen Polynome und hängen auch nicht nur von zwei oder drei Variablen ab, sondern von sehr vielen Parametern. Hier kommen wir oft mit Kopfrechnen oder Papier und Bleistift an unsere Grenzen und müssen die Lösung numerisch bestimmen.

Abstiegsmethoden mit Liniensuche

In diesem Kapitel werden wir zwei numerische Optimierungsmethoden sehen, die in der Praxis häufig und erfolgreich eingesetzt werden: Die Methode des steilsten Abstiegs und das Newton-Verfahren. Beide Methoden minimieren eine multivariate skalarwertige Funktion $f : \mathbb{R}^n \to \mathbb{R}$, von der gefordert wird, dass sie C^1-stetig (für steilsten Abstieg) oder C^2-stetig (für Newton-Verfahren) ist. Beide Algorithmen arbeiten iterativ: Sie beginnen bei einem Startwert $\mathbf{x}^{(0)}$ und verbessern diesen sukzessive, sodass die Folge $\mathbf{x}^{(0)}, \mathbf{x}^{(1)}, \mathbf{x}^{(2)}, \ldots$ gegen den Parametervektor $\mathbf{x}^*$ konvergiert, für den f ein lokales Minimum $f(\mathbf{x}^*)$ annimmt. Hier gilt nach Satz 7.28 also $\nabla f(\mathbf{x}^*) = 0$.

In jeder Iteration wird der Punkt $\mathbf{x}^{(k)} \in \mathbb{R}^n$ dabei um eine Schrittweite $\alpha^{(k)} \in \mathbb{R}_{>0}$ in eine Richtung $\mathbf{d}^{(k)} \in \mathbb{R}^n$ verschoben, um so einen neuen Punkt $\mathbf{x}^{(k+1)}$ mit kleinerem Funktionswert $f(\mathbf{x}^{(k+1)}) < f(\mathbf{x}^{(k)})$ zu erhalten:

$$\mathbf{x}^{(k+1)} := \mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{d}^{(k)}.$$

Hierfür muss der Vektor $\mathbf{d}^{(k)}$ eine Richtung sein, in der sich der Funktionswert auch tatsächlich verringert. Das Bestimmen dieser sogenannten *Abstiegsrichtungen* $\mathbf{d}^{(k)}$ und der entsprechenden *Schrittweiten* $\alpha^{(k)}$ ist die Hauptaufgabe der Optimierungsalgorithmen, die wie folgt aufgebaut sind:

Input: Startpunkt $\mathbf{x}^{(0)}$, Toleranz ε
Output: Approximiertes Minimum $\mathbf{x}^{(k)} \approx \mathbf{x}^*$
$k = 0$
repeat
  Bestimme Abstiegsrichtung $\mathbf{d}^{(k)}$
  Bestimme Schrittweite $\alpha^{(k)}$
  Neuer Punkt $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{d}^{(k)}$
  $k = k + 1$
until $\left\|\nabla f(\mathbf{x}^{(k)})\right\| \leq \varepsilon$

Die Wahl der Schrittweite $\alpha^{(k)}$ (im Machine Learning auch *Lernrate* genannt) hat starken Einfluss auf die Konvergenz der Methode: Ist sie zu groß, divergiert die Methode; ist sie zu klein, braucht der Algorithmus sehr viele Iterationen bis zur Konvergenz. Es gibt es beliebig komplizierte Methoden, um die Schrittweiten zu bestimmen, bis hin zu einer eigenen numerischen Optimierung für die optimale Schrittweite entlang der Richtung $\mathbf{d}^{(k)}$. Eine einfache Heuristik ist, mit einer bestimmten Schrittweite zu starten (z.B. $\alpha^{(k)} = 1$), und diese sukzessive zu halbieren, bis der Funktionswert kleiner wird, also $f\big(\mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{d}^{(k)}\big) < f\big(\mathbf{x}^{(k)}\big)$.

Bei der Wahl der Abstiegsrichtungen unterscheiden sich nun die Methoden. Wir beginnen mit der Methode des steilsten Abstiegs.

Methode des steilsten Abstiegs

Stellen wir uns die zu minimierende Funktion $f(\mathbf{x})$ als Höhenfeld bzw. als Graphen vor. Wenn wir an einer Position $\big(\mathbf{x}^{(k)}, f\big(\mathbf{x}^{(k)}\big)\big)$ auf dem Höhenfeld stehen und möglichst schnell in ein Minimum ("ins Tal") kommen möchten, ist es naheliegend, möglichst steil bergab zu laufen. Die Frage ist jetzt, wie wir die Suchrichtung $\mathbf{d}^{(k)}$ finden, die am steilsten bergab zeigt.

Hier helfen uns wieder (partielle) Ableitungen, die uns am Punkt $\mathbf{x}$ die Steigung der Funktion geben. Wenn

$$\mathbf{e}_i := (0, \dots, 0, \underbrace{1}_{i\text{-te Stelle}}, 0, \dots, 0)^\mathsf{T}$$

der i-te Basisvektor ist, dann ist die partielle Ableitung nach der i-ten Koordinate gerade die Ableitung in Richtung $\mathbf{e}_i$, also die Steigung der Funktion in Richtung $\mathbf{e}_i$

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{e}_i) - f(\mathbf{x})}{h}.$$

Das sieht man leicht durch Vergleich der letzten Gleichung mit Definition 7.23 sieht. Da wir aber aus allen möglichen Richtungen die Richtung des steilsten Abstiegs finden wollen, müssen wir die Funktion "in alle Richtungen" ableiten. Dies führt auf den Begriff der *Richtungsableitung*.

Definition 7.38 (Richtungsableitung)

Sei $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ partiell differenzierbar, $\mathbf{x} \in U$ ein Punkt und $\mathbf{v} \in \mathbb{R}^n$ ein Richtungsvektor mit $\|\mathbf{v}\| = 1$. Wenn der Grenzwert

$$\mathrm{D}_{\mathbf{v}} f(\mathbf{x}) \;:=\; \lim_{h \to 0} \frac{f(\mathbf{x} + h\mathbf{v}) - f(\mathbf{x})}{h}$$

existiert, dann nennen wir diesen die *Richtungsableitung* von f im Punkt $\mathbf{x}$ in Richtung $\mathbf{v}$.

Hierbei entspricht die Richtungsableitung in Richtung $\mathbf{e}_i$ der partiellen Ableitung nach x_i. Jetzt können wir die Funktion f in jede Richtung ableiten und so die Steigung in alle Richtungen ermitteln. Der folgende Satz gibt uns einen interessanten Zusammenhang zwischen der Richtungsableitung und dem Gradienten:

Satz 7.39 (Gradient und Richtungsableitung)

Für eine C^1-stetige Funktion $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ gilt für jeden Punkt $\mathbf{x} \in U$ und jede Richtung $\mathbf{v} \in \mathbb{R}^n$ mit $\|\mathbf{v}\| = 1$

$$\mathrm{D}_{\mathbf{v}} f(\mathbf{x}) \;=\; \nabla f(\mathbf{x}) \cdot \mathbf{v}.$$

Die geometrische Bedeutung des Skalarproduktes kennen wir bereits: Wenn θ der Winkel zwischen $\mathbf{v}$ und $\nabla f(\mathbf{x})$ ist, dann gilt

$$\mathrm{D}_{\mathbf{v}} f(\mathbf{x}) \;=\; \nabla f(\mathbf{x}) \cdot \mathbf{v} \;=\; \|\nabla f(\mathbf{x})\| \cos(\theta).$$

Da der Cosinus sein Maximum an $\theta = 0$ annimmt, ist die Richtungsableitung maximal (d.h., maximale Steigung bergauf), wenn $\mathbf{v}$ und ∇f in die gleiche Richtung zeigen. Probieren Sie das in der folgenden Demo einmal aus.

▶ Demo: Richtungsableitung und Gradient

Mit der gleichen Argumentation wird die Steigung minimal (d.h., maximales Gefälle bergab), wenn $\theta = 180°$, also $\mathbf{v} = -\nabla f(\mathbf{x})$. Der Gradient ist also die Richtung des steilsten Anstiegs und der negative Gradient die Richtung des steilsten Abstiegs. Genau diese Richtung haben wir gesucht, so dass unser Algorithmus damit komplett ist:

Algorithmus: Methode des steilsten Abstiegs (Gradient Descent)

Input: Startpunkt $\mathbf{x}^{(0)}$, Toleranz ε
Output: Approximiertes Minimum $\mathbf{x}^{(k)} \approx \mathbf{x}^*$
$k = 0$
repeat
  Bestimme Abstiegsrichtung: $\mathbf{d}^{(k)} = -\nabla f(\mathbf{x}^{(k)})$
  Bestimme Schrittweite $\alpha^{(k)}$
  Neuer Punkt $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{d}^{(k)}$
  $k = k + 1$
until $\left\|\nabla f(\mathbf{x}^{(k)})\right\| \leq \varepsilon$

In der folgenden Demo können Sie selber in die Rolle dieses Algorithmus schlüpfen, indem Sie den Punkt $\mathbf{x}$ in Richtung des negativen Gradienten verschieben.

▶ Demo: Verschieben in Gradientenrichtung

Die Methode des steilsten Abstiegs, auch *Gradientenabstieg* oder *Gradient Descent* genannt, hat den Vorteil, dass man von der zu minimierenden Funktion nur Funktionswerte $f(\mathbf{x})$ und Gradienten $\nabla f(\mathbf{x})$ auswerten muss, und nicht (wie bei vielen anderen Methoden) ein Gleichungssystem lösen muss. Daher kann man diese Methode auch für eine sehr große Anzahl n von Parametern einsetzen. Aus diesem Grund ist Gradient Descent (bzw. dessen Variante *Stochastic Gradient Descent*) *die* Standardmethode im Deep Learning, um eine Loss Function zu minimieren. Der Nachteil des Gradientenabstiegs ist, dass das Bestimmen geeigneter Schrittweiten nicht trivial ist, und dass der Algorithmus teilweise sehr viele Iterationen benötigt.

Newton-Verfahren

Der Gradientenabstieg hat die Funktion f in jedem Punkt $\mathbf{x}^{(k)}$ lokal durch ihre ersten partiellen Ableitungen angenähert, bzw. das Verhalten der Funktion durch den Gradienten abgeschätzt. Diesen Trick haben wir bei der Taylor-Approximation in Kapitel 5.5 schon einmal gesehen, und er lässt sich auf multivariate Funktionen erweitern. Die multivariate Taylor-Approximation erster Ordnung verwendet den Funktionswert $f(\mathbf{x}_0)$ und den Gradienten $\nabla f(\mathbf{x}_0)$ am Entwicklungspunkt $\mathbf{x}_0$:

$$T_1[f, \mathbf{x}_0](\mathbf{x}) \;=\; f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^\mathsf{T} (\mathbf{x} - \mathbf{x}_0)\,.$$

Wenn die Funktion f nicht nur C^1-stetig, sondern sogar C^2-stetig ist, können wir die Taylor-Approximation zweiter Ordnung bestimmen, welche zusätzlich zum Gradienten ∇f noch die Hesse-Matrix $\nabla^2 f$ verwendet:

$$T_2[f, \mathbf{x}_0](\mathbf{x}) \;=\; f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^\mathsf{T} (\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0)^\mathsf{T} \nabla^2 f(\mathbf{x}_0)\,(\mathbf{x} - \mathbf{x}_0)\,.$$

Dass die Taylor-Approximation zweiter Ordnung die Funktion f deutlich besser approximiert als die Taylor-Approximation erster Ordnung, kann man in der folgenden Demo gut sehen.

▶ Demo: Multivariate Taylor-Approximation

Die Idee des Newton-Verfahrens ist nun, die Funktion f in Iteration k am (Entwicklungs-)Punkt $\mathbf{x}^{(k)}$ durch ihr Taylorpolynom zweiter Ordnung anzunähern. Ein Taylorpolynom zweiter Ordnung ist nur eine quadratische Funktion, die sich deutlich einfacher minimieren lässt. Durch die Minimierung des Taylorpolynoms erhalten wir somit eine neue Näherung $\mathbf{x}^{(k+1)}$. Wir sparen uns der Übersichtlichkeit halber die Iterationsindizes $^{(k)}$ und schreiben die Taylorapproximation am Punkt $\mathbf{x}$ als

$$f(\mathbf{x}+\mathbf{d}) \;\approx\; f(\mathbf{x}) + \nabla f(\mathbf{x})^\mathsf{T}\mathbf{d} + \frac{1}{2}\mathbf{d}^\mathsf{T}\nabla^2 f(\mathbf{x})\mathbf{d}.$$

Um das Minimum zu finden, fassen wir das Taylorpolynom als Funktion $p(\mathbf{d})$ auf und berechnen den Gradienten:

$$p(\mathbf{d}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^\mathsf{T}\mathbf{d} + \frac{1}{2}\mathbf{d}^\mathsf{T}\nabla^2 f(\mathbf{x})\mathbf{d},$$

$$\nabla p(\mathbf{d}) = \nabla f(\mathbf{x}) + \nabla^2 f(\mathbf{x})\mathbf{d}.$$

Für das Berechnen des Gradienten haben wir die folgenden Ableitungsregeln für Funktionen von Vektoren verwendet, welche sich recht einfach nachrechnen lassen:

Satz 7.40 (Vektor-Ableitungen)

a. Für $\mathbf{y} \in \mathbb{R}^n$ und $f : \mathbb{R}^n \to \mathbb{R}$ mit $f(\mathbf{x}) = \mathbf{y}^\mathsf{T}\mathbf{x} = \mathbf{x}^\mathsf{T}\mathbf{y}$ ist $\nabla f(\mathbf{x}) = \mathbf{y}$.

b. Für $f : \mathbb{R}^n \to \mathbb{R}$ mit $f(\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{x}$ ist $\nabla f(\mathbf{x}) = 2\mathbf{x}$.

c. Für $\mathbf{M} \in \mathbb{R}^{n\times n}$ und $f : \mathbb{R}^n \to \mathbb{R}$ mit $f(\mathbf{x}) = \mathbf{x}^\mathsf{T}\mathbf{M}\mathbf{x}$ ist $\nabla f(\mathbf{x}) = \mathbf{M}\mathbf{x} + \mathbf{M}^\mathsf{T}\mathbf{x}$.

Ist $\mathbf{M}$ symmetrisch, ergibt sich $\nabla f(\mathbf{x}) = 2\mathbf{M}\mathbf{x}$.

Der Gradient muss am Minimum dem Nullvektor entsprechen, daher gilt

$$\nabla f(\mathbf{x}) + \nabla^2 f(\mathbf{x})\mathbf{d} = \mathbf{0} \quad\Leftrightarrow\quad \nabla^2 f(\mathbf{x})\mathbf{d} = -\nabla f(\mathbf{x}) \quad\Leftrightarrow\quad \mathbf{d} = -\big(\nabla^2 f(\mathbf{x})\big)^{-1}\nabla f(\mathbf{x}).$$

Wir bekommen also den Update-Vektor $\mathbf{d}$, indem wir das lineare $(n \times n)$-Gleichungssystem $\nabla^2 f(\mathbf{x}) \cdot \mathbf{d} = -\nabla f(\mathbf{x})$ lösen. Wenn Sie gut aufgepasst haben, ist Ihnen aufgefallen, dass der resultierende Vektor $\mathbf{d}$ zunächst nur die *notwendige* Bedingung (Satz 7.28) erfüllt. Damit wirklich ein Minimum vorliegt, muss noch die *hinreichende* Bedingung (Satz 7.35) erfüllt sein, d.h., die Hesse-Matrix von p (welche identisch mit $\nabla^2 f(\mathbf{x})$ ist) muss positiv definit sein. In der Praxis werden daher meist Kostenfunktionen verwendet, welche entweder konvex sind (und damit ist die Hesse-Matrix positiv definit), oder bei denen durch zusätzlich Bedingungen an die Funktion (sogenannte *Regularisierungen*) lokale Konvexität erzwungen wird.

Damit haben wir jetzt alle Zutaten für den Newton-Algorithmus parat:

Algorithmus: Newton-Verfahren

Input: Startpunkt $\mathbf{x}^{(0)}$, Toleranz ε
Output: Approximiertes Minimum $\mathbf{x}^{(k)} \approx \mathbf{x}^*$
$k = 0$
repeat
  Bestimme Abstiegsrichtung: $\mathbf{d}^{(k)} = -\left(\nabla^2 f\left(\mathbf{x}^{(k)}\right)\right)^{-1} \nabla f\left(\mathbf{x}^{(k)}\right)$
  Bestimme Schrittweite $\alpha^{(k)}$
  Neuer Punkt $\mathbf{x}^{(k+1)} = \mathbf{x}^{(k)} + \alpha^{(k)}\mathbf{d}^{(k)}$
  $k = k + 1$
until $\left\|\nabla f(\mathbf{x}^{(k)})\right\| \leq \varepsilon$

Der Newton-Algorithmus ist zwar aufgrund des zu lösenden linearen Gleichungssystems pro Iteration viel aufwendiger als der Gradientenabstieg, konvergiert dafür aber in deutlich weniger Iterationen. Wenn die Problemgröße es zulässt, ist die Newton-Methode daher meist die bessere Wahl. Das robuste und effiziente Lösen linearer Gleichungssysteme ist ein Thema der Master-Vorlesung "Effizientes und paralleles wissenschaftliches Rechnen".