Grundlagen der Mathematik 1: Analysis

Andreas Gathmann

Vorlesungsskript RPTU Kaiserslautern 2023/24

Inhaltsverzeichnis (Grundlagen der Mathematik 1)

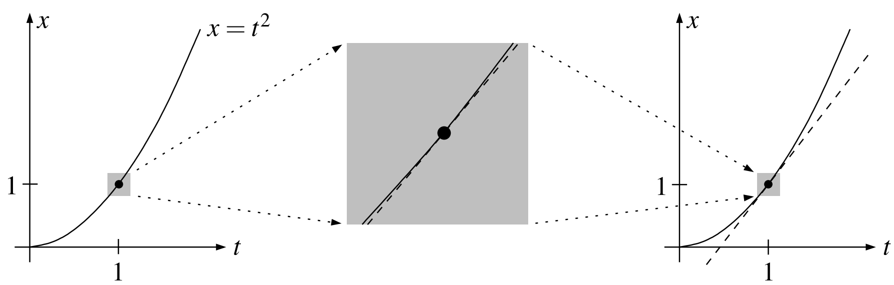
Grundlagen der Mathematik 1: Analysis

0. Einleitung und Motivation

In diesem Skript – so verspricht es der Titel – wollen wir uns die Grundlagen der Mathematik erarbeiten. Aber was ist das überhaupt, die „Grundlagen der Mathematik"? Es handelt sich hierbei um die Kombination zweier Themengebiete, die in der Tat das grundlegende Handwerkszeug für nahezu die gesamte Mathematik darstellen, nämlich

- der *Analysis*, d. h. der Untersuchung von Folgen und Grenzwerten, Stetigkeit, sowie der Differential- und Integralrechnung (zunächst in einer und im zweiten Semester dann auch in mehreren Variablen), und
- der *linearen Algebra*, d. h. der Theorie der Vektorräume, linearen Abbildungen und Gleichungssysteme.

Von beiden Gebieten habt ihr ja aus der Schule wahrscheinlich schon eine ungefähre Vorstellung. In der *Analysis* geht es grob gesagt darum, reelle Funktionen *lokal*, also in der Umgebung eines gewählten Punktes, zu untersuchen, und aus diesen Untersuchungen dann wieder Aussagen über die gesamte Funktion zurückzugewinnen. Betrachten wir z. B. ein Auto, das sich entlang einer geraden Strecke bewegt. Der Einfachheit halber nehmen wir an, dass die Position x des Autos nach der Zeit t (in geeigneten Einheiten) durch die Gleichung $x = t^2$ beschrieben werden kann, so dass die Bewegung durch die Kurve im folgenden Bild links dargestellt wird:



Wir wollen diese Bewegung nun nur in einer kleinen Umgebung eines fest gewählten Zeitpunkts, z. B. des Zeitpunkts $t = 1$ (und damit auch $x = 1$) betrachten. Im Bild oben haben wir diese Umgebung grau markiert und in der Mitte stark vergrößert dargestellt. Wir sehen, dass die Kurve in dieser Umgebung fast wie eine *Gerade* aussieht; wir haben diese Gerade gestrichelt eingezeichnet und im Bild rechts auch außerhalb der gewählten Umgebung fortgesetzt. Geometrisch ist diese Gerade natürlich einfach die *Tangente* an die Kurve an der Stelle $t = 1$. Physikalisch repräsentiert die Steigung dieser Geraden die *Geschwindigkeit* des Autos zum betrachteten Zeitpunkt, denn sie gibt ja gerade an, in welchem Verhältnis sich dort die Strecke x mit der Zeit t verändert. Aus der Schule wisst ihr auch schon, wie man diese Steigung ausrechnet: Man muss dazu die gegebene Funktion t^2 *differenzieren* – so dass man die Ableitung $2t$ erhält – und dort die betrachtete Stelle $t = 1$ einsetzen. Die Steigung ist in unserem Fall also gerade $2 \cdot 1 = 2$, und man rechnet sofort nach, dass $x = 2t - 1$ die Gleichung der oben eingezeichneten Tangente ist. Man sagt, dass die Gerade $x = 2t - 1$ eine *lineare Approximation* der ursprünglich gegebenen Funktion $x = t^2$ im Punkt $t = 1$ ist.

Wir können aus der Kenntnis der zurückgelegten Strecke zu jedem Zeitpunkt also die Geschwindigkeit des Autos durch Differenzieren bestimmen. Man kann sich natürlich auch die umgekehrte Frage stellen: Angenommen, der Kilometerzähler eures Autos ist kaputt, aber ihr beobachtet auf eurer Fahrt ständig eure Geschwindigkeit. Könnt ihr dann am Ende der Fahrt trotzdem ausrechnen, wie weit ihr gefahren seid? Dies ist offensichtlich die „Umkehrung" des Differenzierens – und auch hier wisst ihr aus der Schule natürlich schon, dass dies auf die *Integralrechnung* führen wird.

In der Praxis bewegt man sich mit dem Auto aber nicht immer nur auf einer geraden Strecke, und demzufolge braucht man für die Beschreibung solch einer Bewegung (und natürlich auch vieler anderer natürlich auftretender Prozesse) mehrere Variablen. Wir werden daher auch Abbildungen betrachten, die mehrere Variablen auf mehrere andere abbilden, wie z. B. die folgende Vorschrift, die zwei reelle Zahlen y_1 und y_2 in Abhängigkeit von zwei anderen x_1 und x_2 ausdrückt:

$$\begin{aligned} y_1 &= 2x_1 + x_2 \cos x_1 \\ y_2 &= x_1 \,\mathrm{e}^{x_2} - x_2 \end{aligned}$$

Genau wie oben werden wir uns auch hier wieder die Frage stellen, ob wir diese (in diesem Fall recht komplizierte) Funktion in der Nähe eines gegebenen Punktes nicht vielleicht durch eine einfache *lineare* Abhängigkeit annähern können. In der Tat ist dies möglich: Wir werden sehen, dass z. B. in einer kleinen Umgebung des Nullpunkts $(x_1,x_2) = (0,0)$ die obige Vorschrift näherungsweise die gleichen Ergebnisse liefert wie das lineare Gleichungssystem

$$\begin{aligned} y_1 &= 2x_1 + x_2 \\ y_2 &= x_1 - x_2. \end{aligned}$$

Auch wenn diese Gleichungen natürlich viel einfacher als die ursprünglichen sind, sollte offensichtlich sein, dass auch solche linearen Gleichungssysteme bei wachsender Zahl von Variablen (und in der Praxis sind Hunderte oder Tausende von Variablen keine Seltenheit) recht kompliziert werden können. Wir werden daher einen wesentlichen Teil dieser Vorlesung mit der *linearen Algebra*, also dem Studium derartiger linearer Gleichungssysteme, verbringen. Um dabei überhaupt erst einmal den Notationsaufwand in Grenzen zu halten, tut man dabei gut daran, die Start- und Zielvariablen nicht alle einzeln hinzuschreiben, sondern sie zu sogenannten *Vektoren* zusammenzufassen. In der Tat kann man in dieser Sichtweise die lineare Algebra als das Studium von linearen Abbildungen zwischen Vektorräumen beschreiben.

Wenn wir dann die linearen Abbildungen zwischen mehreren Variablen gut genug verstanden haben, können wir uns im zweiten Teil der Analysis schließlich daran machen, die Differential- und Integralrechnung auf den Fall von mehreren Variablen auszuweiten.

Wir haben damit jetzt relativ kurz umrissen, welcher mathematische Stoff uns in dieser Vorlesung erwartet. Es wird in diesem Skript aber nicht nur darum gehen, mathematische Resultate kennenzulernen. Mindestens ebenso wichtig ist es, das „mathematische Denken“ zu lernen, d. h. die Fähigkeit zu entwickeln, mit abstrakten Konzepten umzugehen, exakte logische Schlüsse zu ziehen und Beweise zu führen. Anders als in der Schule oder im Studium z. B. ingenieurswissenschaftlicher Fächer werden wir genau darauf achten, eine „wasserdichte“ Theorie aufzubauen: Jeder neue Begriff bzw. jeder neue Satz wird nur unter Verwendung des bisher Bekannten exakt definiert bzw. bewiesen. So sind z. B. Formulierungen in dem Stil „eine Funktion wird durch eine Gerade angenähert“ oder „eine Funktion heißt stetig, wenn man sie zeichnen kann, ohne den Stift abzusetzen“ als Veranschaulichung unserer Ideen zwar sehr sinnvoll, als exakte mathematische Formulierung jedoch schlichtweg unbrauchbar. Wir werden daher in dieser Vorlesung viel exakter arbeiten als ihr es wahrscheinlich aus der Schule gewohnt seid, und es ist wichtig, dass ihr diese exakte Denkweise verinnerlicht und anzuwenden lernt. Die Mathematik ist ein riesiges Gebäude – viel größer und komplexer als ihr es euch wahrscheinlich im Moment vorstellen könnt – das ständig höher gebaut wird, indem schon bewiesene Sätze auf neue Fälle angewendet oder wieder für neue Beweise verwendet werden. Wir starten gerade beim Fundament dieses Gebäudes und können es uns da wirklich nicht leisten herumzupfuschen.

Diese konsequent logische und exakte Herangehensweise ist zwar am Anfang wahrscheinlich ungewohnt, hat jedoch für euch auch einen Vorteil: Etwas überspitzt formuliert erzählen wir euch während des gesamten Studiums eigentlich nur Dinge, die logisch aus dem folgen, was ihr ohnehin schon wusstet. Dadurch ist die Mathematik wahrscheinlich das Studienfach, in dem man am wenigsten auswendig lernen muss – in dem es aber im Gegenzug auch am meisten auf das *Verständnis* des Stoffes ankommt. Je besser euer Verständnis für die Mathematik wird, um so mehr Dinge werden euch letztlich einfach „klar“ werden, so dass es euch dann auch viel leichter fällt, sie zu lernen.

Dieses Verständnis für die Mathematik bekommt man aber natürlich nur durch intensiven und *aktiven* Umgang mit dem Stoff, weswegen neben dem Studium der Vorlesung auch die Bearbeitung der Übungsaufgaben besonders wichtig ist.

Heißt das alles nun, dass wir nur mit logischen Argumenten die Mathematik sozusagen „aus dem Nichts“ aufbauen können? Nein, das geht natürlich nicht … von nichts kommt nichts. Man muss am Anfang immer gewisse Dinge als gegeben annehmen, also Aussagen als wahr voraussetzen, die man nicht mehr beweist bzw. beweisen kann, und auf denen dann die gesamte Theorie beruht. Derartige Annahmen bezeichnet man als **Axiome**. Natürlich versucht man in der Mathematik, mit möglichst wenigen und sehr elementaren Axiomen auszukommen, die hoffentlich niemand anzweifeln würde. In der modernen Mathematik ist es üblich, hierfür die grundlegenden Prinzipien der Logik und Mengenlehre zu verwenden (zu denen wir auch gleich in Kapitel 1 Genaueres sagen werden).

In dieser Vorlesung wollen wir uns das Leben allerdings etwas leichter machen und zusätzlich auch die *Existenz und elementaren Eigenschaften der reellen Zahlen* axiomatisch voraussetzen (um welche Eigenschaften es sich hierbei handelt, werden wir natürlich genau angeben). Man kann zwar nur aus den Axiomen der Logik und Mengenlehre beweisen, dass die reellen Zahlen existieren und dass sie die erwarteten Eigenschaften haben, der Beweis wäre zu diesem frühen Zeitpunkt im Studium aber sehr verwirrend und würde euch auch keine großartigen neuen Erkenntnisse bringen. In diesem Sinne starten wir also sozusagen doch nicht ganz beim Fundament unseres „Gebäudes Mathematik“, sondern bereits im ersten Stock.

Im weiteren Verlauf ist dieses Skript dann wie im Inhaltsverzeichnis angegeben in mehrere Teile gegliedert. Diese bauen der Reihe nach aufeinander auf, mit einer Ausnahme: Nach den in jedem Fall benötigten grundlegenden Anfangskapiteln 1 bis 3 sind die Teile „Grundlagen der Mathematik 1: Analysis“ (Kapitel 4 bis 12) und „Grundlagen der Mathematik 1: Lineare Algebra“ (Kapitel **??** bis **??**) *unabhängig voneinander* und können somit in beliebiger Reihenfolge oder auch parallel studiert werden.

Aber jetzt genug der Vorrede … beginnen wir nun also unser Studium der Mathematik mit den „Grundlagen der Grundlagen“, den für uns wesentlichen Prinzipien der Logik und Mengenlehre.

1. Etwas Logik und Mengenlehre

Bevor wir mit dem eigentlichen Inhalt der Vorlesung beginnen, müssen wir in diesem Kapitel kurz die exakte mathematische Sprache beschreiben, in der wir unsere Ergebnisse formulieren werden: die der Logik und Mengenlehre. Zentral hierbei sind die Begriffe der *Aussage* (in der Logik) und der *Menge* (in der Mengenlehre).

Da wir es hier mit den ersten beiden Begriffen überhaupt zu tun haben, die in der Mathematik vorkommen, können wir sie natürlich nicht durch bereits bekannte Dinge definieren oder mit bereits bekannten Resultaten ihre Eigenschaften herleiten. Wir müssen sie daher (wie schon in der Einleitung erwähnt) axiomatisch voraussetzen. Wir müssen *voraussetzen*, dass es sinnvoll ist, über logische Aussagen und deren Wahrheit zu reden, dass Mengen überhaupt existieren, dass man Mengen vereinigen und schneiden kann, aus ihnen Elemente auswählen kann, und noch einiges mehr. Wenn ihr euch zum Beispiel auf den Standpunkt stellt, dass ihr nicht an die Existenz von Mengen glaubt, wird euch niemand widerlegen können. Allerdings zweifelt ihr damit dann auch die Existenz der gesamten Mathematik an, wie sie heutzutage betrieben wird – und aus der Tatsache, dass ihr in dieser Vorlesung sitzt, schließe ich einmal, dass das nicht der Fall ist.

Glücklicherweise sind die Dinge, die wir benötigen, jedoch allesamt anschaulich sofort einleuchtend und euch natürlich aus der Schule auch schon hinlänglich bekannt. Ich möchte es euch (und mir) daher ersparen, an dieser Stelle eine vollständige und präzise axiomatische Formulierung der Logik und Mengenlehre hinzuschreiben, zumal das momentan sicher mehr verwirren als helfen würde und außerdem gerade im Bereich der Logik auch zu sehr in die Philosophie abdriften würde. Stattdessen wollen wir uns hier damit begnügen, die für uns wichtigsten Prinzipien und Notationen sowie beliebte Fehlerquellen in verständlicher Sprache zu erklären, auch wenn ein paar Dinge (insbesondere die Begriffsfestlegung – „Definition" möchte ich es eigentlich gar nicht nennen – einer Aussage und einer Menge) dadurch recht schwammig klingen werden. Außerdem werden wir in Beispielen zur besseren Verdeutlichung bereits hier die reellen Zahlen und ihre einfachsten Eigenschaften (die euch sicherlich bekannt sein werden) benutzen, auch wenn wir diese erst später formalisieren werden. Da es sicher niemanden von euch verwirren wird, werden wir auch die Schreibweise „$x \in \mathbb{R}$" für „x ist eine reelle Zahl" schon verwenden, bevor sie in den Notationen 1.12 und 1.14 offiziell eingeführt wird.

1.A Logik

Beginnen wir also mit der Logik. Unter einer **Aussage** verstehen wir (grob gesagt) ein sprachliches Gebilde, das entweder wahr oder falsch ist (wobei wir in der Mathematik natürlich letztlich daran interessiert sind, *wahre* Aussagen herzuleiten). Wichtig sind auch solche sprachlichen Gebilde, in denen freie **Variablen**, also Platzhalter, vorkommen, und die erst beim Einsetzen von Werten für diese Variablen Aussagen liefern. Man bezeichnet sie als **Aussageformen**.

Beispiel 1.1.

(a) $1+1=2$ ist eine wahre, $1+1=3$ eine falsche, und $1+1$ überhaupt keine Aussage.

(b) $x+1=2$ ist eine Aussageform, die beim Einsetzen von $x=1$ in eine wahre, beim Einsetzen jeder anderen reellen Zahl in eine falsche Aussage übergeht.

Bemerkung 1.2. Als Variablen in Aussageformen kann man beliebige Symbole benutzen. Üblich sind neben den normalen lateinischen Klein- und Großbuchstaben auch die griechischen Buchstaben, die wir zur Erinnerung hier auflisten:

$\mathrm{A}\ \alpha$ alpha	$\mathrm{B}\ \beta$ beta	$\Gamma\ \gamma$ gamma	$\Delta\ \delta$ delta	$\mathrm{E}\ \varepsilon$ epsilon	$\mathrm{Z}\ \zeta$ zeta	$\mathrm{H}\ \eta$ eta	$\Theta\ \vartheta$ theta
$\mathrm{I}\ \iota$ iota	$\mathrm{K}\ \kappa$ kappa	$\Lambda\ \lambda$ lambda	$\mathrm{M}\ \mu$ my	$\mathrm{N}\ \nu$ ny	$\Xi\ \xi$ xi	$\mathrm{O}\ o$ omikron	$\Pi\ \pi$ pi
$\mathrm{P}\ \rho$ rho	$\Sigma\ \sigma$ sigma	$\mathrm{T}\ \tau$ tau	$\Upsilon\ \upsilon$ ypsilon	$\Phi\ \varphi$ phi	$\mathrm{X}\ \chi$ chi	$\Psi\ \psi$ psi	$\Omega\ \omega$ omega

Oft verziert man Buchstaben auch noch mit einem Symbol oder versieht sie mit einem Index, um neue Variablen zu erhalten: So sind z. B. $x, x', \tilde{x}, \bar{x}, x_1, x_2, \dots$ alles Symbole für verschiedene Variablen, die zunächst einmal nichts miteinander zu tun haben (aber tunlichst für irgendwie miteinander zusammenhängende Objekte eingesetzt werden sollten, wenn man den Leser nicht vollends verwirren will).

Notation 1.3 (Zusammengesetzte Aussagen)**.** Sind A und B Aussagen, so lassen sich daraus wie folgt neue bilden:

Symbol	Wahrheitstafel				Bedeutung
A	w	f	w	f	
B	w	w	f	f	
$\neg A$	f	w			nicht A
$A \wedge B$	w	f	f	f	A und B
$A \vee B$	w	w	w	f	A oder B (oder beides): „nicht-ausschließendes Oder“
$A \Leftrightarrow B$	w	f	f	w	A und B sind gleichbedeutend / *äquivalent*, bzw. A genau dann, wenn B
$A \Rightarrow B$	w	w	f	w	aus A folgt B, bzw. wenn A dann B

Die sogenannte **Wahrheitstafel** in den mittleren vier Spalten ist dabei die eigentliche Definition der neuen zusammengesetzten Aussagen. Sie gibt in Abhängigkeit der Wahrheit von A und B (in den ersten beiden Zeilen) an, ob die zusammengesetzte Aussage wahr oder falsch ist.

Bemerkenswert ist hierbei wohl nur die Folgerungsaussage $A \Rightarrow B$, die keine Aussage über die Richtigkeit von A oder B separat macht, sondern nur sagt, dass B wahr ist, wenn auch A es ist. Ist hingegen A falsch, so ist die Folgerungsaussage $A \Rightarrow B$ stets wahr („aus einer falschen Voraussetzung kann man alles folgern“). So ist z. B. $0 = 1 \quad\Rightarrow\quad 2 = 3$ eine wahre Aussage. In der Regel wollen wir uns in der Mathematik aber natürlich mit wahren Aussagen beschäftigen, und neue wahre Aussagen aus alten herleiten. Gerade in Beweisen ist die übliche Verwendung der Notation $A \Rightarrow B$ daher, dass A eine bereits als wahr erkannte Aussage ist, und wir damit nun schließen wollen, dass auch B wahr ist.

Bemerkung 1.4 (Beweise mit Wahrheitstafeln)**.** Wollen wir kompliziertere zusammengesetzte Aussagen miteinander vergleichen, so können wir dies auch mit Hilfe von Wahrheitstafeln tun. So ist für zwei Aussagen A und B z. B.

$$A \Rightarrow B \quad \text{äquivalent zu} \quad (\neg A) \vee B,$$

denn wenn wir in der Wahrheitstafel

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

mit Hilfe der Definitionen von $\neg$ und $\vee$ aus Notation 1.3 zunächst $\neg A$ und dann $(\neg A) \vee B$ berechnen, sehen wir, dass das Ergebnis mit $A \Rightarrow B$ übereinstimmt. Nach der Bemerkung aus Notation 1.3 ist dies auch anschaulich klar: Die Folgerungsaussage $A \Rightarrow B$ ist ja genau dann wahr, wenn A falsch (also $\neg A$ wahr) ist, oder wenn B wahr ist (oder beides).

Genauso zeigt man die ebenfalls einleuchtende Aussage, dass

$$A \Leftrightarrow B \quad \text{äquivalent zu} \quad (A \Rightarrow B) \wedge (B \Rightarrow A)$$

ist – was auch die übliche Art ist, wie man eine Äquivalenz zeigt: Man zeigt separat die beiden Folgerungen $A \Rightarrow B$ und $B \Rightarrow A$.

Notation 1.5. Folgerungen („$\Rightarrow$") und Äquivalenzen („$\Leftrightarrow$") sind natürlich zwei verschiedene Dinge, die man nicht durcheinanderwerfen darf (auch wenn das in der Schule wahrscheinlich manchmal nicht so genau genommen wird). Es hat sich jedoch in der Mathematik eingebürgert, bei *Definitionen* von Begriffen durch eine äquivalente, definierende Eigenschaft die Sprechweise „wenn" anstatt des eigentlich korrekten „genau dann, wenn" zu verwenden: So würde man z. B. als Definition des Begriffs einer geraden Zahl hinschreiben

> „Eine ganze Zahl x heißt gerade, wenn $\frac{x}{2}$ eine ganze Zahl ist",

obwohl man genau genommen natürlich meint

> „Eine ganze Zahl x heißt *genau dann* gerade, wenn $\frac{x}{2}$ eine ganze Zahl ist".

Eine gewöhnliche Folgerungsaussage wie z. B. die wahre Aussage

> „Wenn eine ganze Zahl x positiv ist, dann ist auch $x+1$ positiv"

ist dagegen immer nur in einer Richtung zu verstehen; hier wird also nicht behauptet, dass mit $x+1$ auch x immer positiv sein muss (was ja auch falsch wäre).

Notation 1.6 (Quantoren)**.** Ist A eine Aussageform, in der eine freie Variable x vorkommt – wir schreiben dies dann auch als $A(x)$ – so setzen wir

Symbol	Bedeutung
$\forall x : A(x)$	für alle x gilt $A(x)$
$\exists x : A(x)$	es gibt ein x mit $A(x)$

Die beiden Symbole $\forall$ und $\exists$ bezeichnet man als **Quantoren**. Beachte, dass diese beiden Quantoren *nicht* miteinander vertauschbar sind: So besagt z. B. die Aussage

$$\forall x \in \mathbb{R}\ \exists y \in \mathbb{R} : y > x$$

„zu jeder reellen Zahl x gibt es eine Zahl y, die größer ist" (was offensichtlich wahr ist), während die Umkehrung der beiden Quantoren die Aussage

$$\exists y \in \mathbb{R}\ \forall x \in \mathbb{R} : y > x$$

„es gibt eine reelle Zahl y, die größer als jede reelle Zahl x ist" liefern würde (was ebenso offensichtlich falsch ist). Der Unterschied besteht einfach darin, dass im ersten Fall zuerst das x gewählt werden muss und dann ein y dazu existieren muss (das von x abhängen darf), während es im zweiten Fall *dasselbe* y für alle x sein müsste.

Bemerkung 1.7. Jede Aussage lässt sich natürlich auf viele Arten aufschreiben, sowohl als deutscher Satz als auch als mathematische Formel. Die gerade eben betrachtete Aussage könnte man z. B. auf die folgenden (absolut gleichwertigen) Arten aufschreiben:

(a) $\forall x \in \mathbb{R}\ \exists y \in \mathbb{R} : y > x$.

(b) Es sei $x \in \mathbb{R}$. Dann gibt es ein $y \in \mathbb{R}$ mit $y > x$.

(c) Zu jeder reellen Zahl gibt es noch eine größere.

Welche Variante man beim Aufschreiben wählt, ist weitestgehend Geschmackssache. Die Formulierung einer Aussage als deutscher Satz hat den Vorteil, dass wir sie oft leichter verstehen können, weil wir die deutsche Sprache schon länger kennen als die mathematische. Wenn wir uns jedoch erst einmal an die mathematische Sprache gewöhnt haben, wird auch sie ihre Vorzüge bekommen: Sie ist deutlich kürzer und besser logisch strukturiert. Wir werden im Folgenden beide Schreibweisen mischen und jeweils diejenige wählen, mit der unsere Aussagen (hoffentlich) am einfachsten verständlich werden.

Wenn wir mathematische Symbole verwenden, müssen wir diese aber auch stets in ihrer korrekten Notation und nicht als „Abkürzungen" für deutsche Wörter verwenden: Man würde die Aussage

„2 und 4 sind gerade Zahlen“ sicher niemals schreiben als „$2 \wedge 4$ sind gerade Zahlen“, und analog genauso wenig „Es gilt $n \geq 0$ für alle $n \in \mathbb{N}$“ als „Es gilt $n \geq 0 \, \forall n \in \mathbb{N}$“.

Bemerkung 1.8 (Negationen)**.** Es ist wichtig zu wissen, wie man von einer Aussage das Gegenteil, also die „Verneinung“ bildet. Da hierbei oft Fehler gemacht werden, wollen wir die allgemeinen Regeln hierfür kurz auflisten (die man wie in Bemerkung 1.4 auch wieder schnell mit Wahrheitstafeln zeigen könnte):

(a) $\neg(\neg A) \Leftrightarrow A$: Ist es falsch, dass A falsch ist, so bedeutet dies genau, dass A wahr ist.

(b) $\neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B)$: Das Gegenteil von „$A$ und B sind richtig“ ist „A oder B ist falsch“.

(c) $\neg(A \vee B) \Leftrightarrow (\neg A) \wedge (\neg B)$: Das Gegenteil von „$A$ oder B ist richtig“ ist „A und B sind falsch“.

(d) $\neg(\forall x : A(x)) \Leftrightarrow \exists x : \neg A(x)$: Das Gegenteil von „für alle x gilt $A(x)$“ ist „es gibt ein x, für das $A(x)$ falsch ist“.

(e) $\neg(\exists x : A(x)) \Leftrightarrow \forall x : \neg A(x)$: Das Gegenteil von „es gibt ein x, für das $A(x)$ gilt“ ist „für alle x ist $A(x)$ falsch“.

Man kann also sagen, dass eine Verneinung dazu führt, dass „und“ mit „oder“ sowie „für alle“ mit „es gibt“ vertauscht werden. So ist z. B. das Gegenteil der Aussage

„In Frankfurt haben *alle* Haushalte Strom *und* fließendes Wasser“

die Aussage

„In Frankfurt *gibt es* einen Haushalt, der keinen Strom *oder* kein fließendes Wasser hat“.

Beispiel 1.9.

(a) Wollen wir eine Folgerung $A \Rightarrow B$ verneinen, so können wir sie zunächst mit Bemerkung 1.4 zu $(\neg A) \vee B$ umformen, und erhalten nach Bemerkung 1.8 als Umkehrung dann $A \wedge \neg B$. Dies ist auch anschaulich einleuchtend: Die Folgerungsaussage „wenn A dann B“ ist genau dann falsch, wenn die Voraussetzung A zwar gilt, die Behauptung B aber nicht. Wir sehen also:

$$\boxed{\text{Die Verneinung einer Folgerung } A \Rightarrow B \text{ ist } A \wedge \neg B}$$

(und nicht etwa $A \Rightarrow \neg B$, wie man vielleicht denken könnte).

(b) Eine oft vorkommende Anwendung der Regeln für die Verneinung von Aussagen ist der sogenannte **Widerspruchsbeweis** bzw. Beweis durch **Kontraposition**. Nach Bemerkung 1.4 gesehen ist die Folgerung $A \Rightarrow B$ („aus A folgt B“) gleichbedeutend mit $(\neg A) \vee B$. Damit ist diese Aussage nach Bemerkung 1.8 (a) auch äquivalent zu $(\neg(\neg B)) \vee (\neg A)$, also zu $\neg B \Rightarrow \neg A$. Mit anderen Worten: Haben wir eine Schlussfolgerung $A \Rightarrow B$ zu beweisen, so können wir genauso gut $(\neg B) \Rightarrow (\neg A)$ zeigen, d. h. *wir können annehmen, dass die zu zeigende Aussage B falsch ist und dies dann zu einem Widerspruch führen bzw. zeigen, dass dann auch die Voraussetzung A falsch sein muss.*

Beispiel 1.10. Hier sind zwei Beispiele für die Anwendung der Prinzipien aus Bemerkung 1.8 und Beispiel 1.9 – und auch unsere ersten Beispiele dafür, wie man Beweise von Aussagen exakt aufschreiben kann.

(a) Einen Beweis durch Widerspruch könnte man z. B. so aufschreiben:

Behauptung: Für alle $x \in \mathbb{R}$ gilt $2x + 1 > 0$ oder $2x - 1 < 0$.

Beweis: Angenommen, die Behauptung wäre falsch, d. h. (nach Bemerkung 1.8 (c) und (d)) es gäbe ein $x \in \mathbb{R}$ mit

$$2x + 1 \leq 0 \quad (1) \qquad \text{und} \qquad 2x - 1 \geq 0 \quad (2).$$

Für dieses x würde dann folgen, dass

$$0 \overset{(1)}{\geq} 2x + 1 = 2x - 1 + 2 \overset{(2)}{\geq} 0 + 2 = 2.$$

Dies ist aber ein Widerspruch. Also war unsere Annahme falsch und somit die zu beweisende Aussage richtig. □

Das dabei verwendete Symbol „□" ist die übliche Art, das Ende eines Beweises zu kennzeichnen. Zur Verdeutlichung haben wir die beiden Ungleichungen mit (1) und (2) markiert, um später angeben zu können, wo sie verwendet werden.

(b) Manchmal weiß man von einer Aussage aufgrund der Aufgabenstellung zunächst einmal noch nicht, ob sie wahr oder falsch ist. In diesem Fall muss man sich dies natürlich zuerst überlegen – und, falls die Aussage falsch ist, ihre Negation beweisen. Als Beispiel dafür betrachten wir die Aufgabe

Man beweise oder widerlege: Für alle $x \in \mathbb{R}$ gilt $2x+1 < 0$ oder $2x-1 > 0$.

In diesem Fall merkt man schnell, dass die Aussage falsch sein muss, weil die Ungleichungen schon für den Fall $x = 0$ nicht stimmen. Man könnte als Lösung der Aufgabe unter Beachtung der Negationsregeln aus Bemerkung 1.8 also aufschreiben:

Behauptung: Die Aussage ist falsch, d. h. es gibt ein $x \in \mathbb{R}$ mit $2x+1 \geq 0$ und $2x-1 \leq 0$.

Beweis: Für $x = 0$ ist $2x+1 = 1 \geq 0$ und $2x-1 = -1 \leq 0$. □

Beachte, dass dies ein vollständiger Beweis ist: *Um eine allgemeine Aussage zu widerlegen, genügt es, ein Gegenbeispiel dafür anzugeben.*

Bemerkung 1.11. Bevor wir unsere kurze Auflistung der für uns wichtigen Prinzipien der Logik beenden, wollen wir noch kurz auf ein paar generelle Dinge eingehen, die man beim Aufschreiben mathematischer Beweise oder Rechnungen beachten muss.

Dass wir bei unseren logischen Argumenten sauber und exakt arbeiten – also z. B. nicht Folgerungen, die keine Äquivalenzen sind, in der falschen Richtung verwenden, „für alle" mit „es gibt" verwechseln oder ähnliches – sollte sich von selbst verstehen. Die folgende kleine Geschichte hilft vielleicht zu verstehen, was damit gemeint ist.

Ein Ingenieur, ein Physiker und ein Mathematiker fahren mit dem Zug nach Frankreich und sehen dort aus dem Fenster des Zuges ein schwarzes Schaf.

Da sagt der Ingenieur: „Oh, in Frankreich sind die Schafe schwarz!"

Darauf der Physiker: „Nein ... wir wissen jetzt nur, dass es in Frankreich mindestens ein schwarzes Schaf gibt."

Der Mathematiker: „Nein ... wir wissen nur, dass es in Frankreich mindestens ein Schaf gibt, das auf mindestens einer Seite schwarz ist."

Es gibt aber noch einen weiteren sehr wichtigen Punkt, der leider oft nicht beachtet wird: In der Regel werden wir beim Aufschreiben sowohl Aussagen notieren wollen, die wir erst noch zeigen wollen (um schon einmal zu sagen, worauf wir hinaus wollen), als auch solche, von denen wir bereits wissen, dass sie wahr sind (z. B. weil sie für die zu zeigende Behauptung als wahr vorausgesetzt werden oder weil sie sich logisch aus irgendetwas bereits Bekanntem ergeben haben). Es sollte offensichtlich sein, dass wir Aussagen mit derartig verschiedenen Bedeutungen für die Argumentationsstruktur nicht einfach kommentarlos nebeneinander schreiben dürfen, wenn noch jemand in der Lage sein soll, die Argumente nachzuvollziehen. Betrachten wir z. B. noch einmal unseren Beweis aus Beispiel 1.10 (a) oben, so wäre eine Art des Aufschreibens in folgendem Stil (wie man es leider oft sieht)

$$\begin{aligned}
&2x+1 > 0 \;\text{ oder }\; 2x-1 < 0 \\
&2x+1 \leq 0 \qquad\quad\; 2x-1 \geq 0 \\
&0 \geq 2x+1 = 2x-1+2 \geq 0+2 = 2
\end{aligned}$$

völlig inakzeptabel, obwohl hier natürlich letztlich die gleichen Aussagen stehen wie oben. Kurz gesagt:

> Von *jeder* aufgeschriebenen Aussage muss für den Leser *sofort* und *ohne eigenes Nachdenken* ersichtlich sein, welche Rolle sie in der Argumentationsstruktur spielt: Ist es z. B. eine noch zu zeigende Behauptung, eine Annahme oder eine Folgerung (und wenn ja, aus was)?

Dies bedeutet allerdings nicht, dass wir ganze Aufsätze schreiben müssen. Eine (schon recht platzoptimierte) Art, den Beweis aus Beispiel 1.10 (a) aufzuschreiben, wäre z. B.

Angenommen, es gäbe ein $x \in \mathbb{R}$ mit $2x+1 \leq 0$ und $2x-1 \geq 0$.

Dann wäre $0 \geq 2x+1 = 2x-1+2 \geq 0+2 = 2$, Widerspruch. □

01

1.B Mengenlehre

Nachdem wir die wichtigsten Regeln der Logik behandelt haben, wenden wir uns jetzt der Mengenlehre zu. Die gesamte moderne Mathematik basiert auf diesem Begriff der Menge, der ja auch schon aus der Schule hinlänglich bekannt ist. Zur Beschreibung, was eine Menge ist, zitiert man üblicherweise die folgende Charakterisierung von Georg Cantor (1845–1918):

„Eine **Menge** ist eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.“

Die in einer Menge zusammengefassten Objekte bezeichnet man als ihre **Elemente**.

Notation 1.12.

(a) Wir schreiben $x \in M$, falls x ein Element der Menge M ist, und $x \notin M$ andernfalls.

(b) Die einfachste Art, eine Menge konkret anzugeben, besteht darin, ihre Elemente in geschweiften Klammern aufzulisten, wobei es auf die Reihenfolge und Mehrfachnennungen nicht ankommt. So sind z. B. $\{1,2,3\}$ und $\{2,3,1,3\}$ zwei Schreibweisen für dieselbe Menge mit den drei Elementen 1, 2 und 3.

Beachte, dass die Elemente einer Menge nicht unbedingt Zahlen sein müssen – so ist z. B. $M = \{\{2,3\},\{1,3\}\}$ eine Menge mit zwei Elementen, die selbst wieder Mengen sind, nämlich $\{2,3\}$ und $\{1,3\}$. Mit der Notation aus (a) ist also z. B. $\{1,3\} \in M$. Insbesondere ist M nicht dasselbe wie die Menge $\{1,2,3\}$.

(c) Man kann die Elemente einer Menge auch durch eine beschreibende Eigenschaft angeben: $\{x : A(x)\}$ bezeichnet die Menge aller Objekte x, für die die Aussage $A(x)$ wahr ist, wie z. B. in $\{x \in \mathbb{R} : x^2 = 1\} = \{-1,1\}$.

(d) Die Menge $\{\}$ ohne Elemente, die sogenannte **leere Menge**, bezeichnen wir mit $\emptyset$.

(e) Eine Menge M heißt **Teilmenge** einer Menge N (geschrieben $M \subset N$), wenn jedes Element von M auch Element von N ist, bzw. in der Quantorenschreibweise von Notation 1.6 wenn

$$\forall x : x \in M \Rightarrow x \in N.$$

Man sagt in diesem Fall auch, dass N eine **Obermenge** von M ist (geschrieben $N \supset M$).

Beachte, dass M und N dabei auch gleich sein können; in der Tat ist offensichtlich

$$M = N \quad \text{genau dann, wenn} \quad M \subset N \text{ und } N \subset M.$$

Oft wird man eine Gleichheit $M = N$ von Mengen auch so beweisen, dass man separat $M \subset N$ und $N \subset M$ zeigt.

Wenn wir ausdrücken wollen, dass M eine Teilmenge von N und nicht gleich N ist, so schreiben wir dies als $M \subsetneq N$ und sagen, dass M eine **echte Teilmenge** von N ist. Es ist wichtig, dies von der Aussage $M \not\subset N$ zu unterscheiden, die bedeutet, dass M keine Teilmenge von N ist.

Achtung: Manchmal wird in der Literatur das Symbol „$\subset$“ für *echte* Teilmengen und „$\subseteq$“ für nicht notwendig echte Teilmengen verwendet.

(f) Hat eine Menge M nur endlich viele Elemente, so nennt man M eine **endliche Menge** und schreibt die Anzahl ihrer Elemente als $|M|$. Andernfalls setzt man formal $|M| = \infty$.

Bemerkung 1.13 (Russellsches Paradoxon). Die oben gegebene Charakterisierung von Mengen von Cantor ist aus mathematischer Sicht natürlich sehr schwammig. In der Tat hat Bertrand Russell kurz darauf bemerkt, dass sie sogar schnell zu Widersprüchen führt. Er betrachtet dazu

$$M = \{A : A \text{ ist eine Menge mit } A \notin A\}, \qquad (*)$$

also „die Menge aller Mengen, die sich nicht selbst als Element enthalten". Sicherlich ist es eine merkwürdige Vorstellung, dass eine Menge sich selbst als Element enthalten könnte – im Sinne von Cantors Charakterisierung wäre die Definition $(*)$ aber zulässig. Fragen wir uns nun allerdings, ob sich die so konstruierte Menge M selbst als Element enthält, so erhalten wir sofort einen Widerspruch: Wenn $M \in M$ gilt, so würde das nach der Definition $(*)$ ja gerade bedeuten, dass $M \notin M$ ist – und das wiederum, dass doch $M \in M$ ist. Man bezeichnet dies als das *Russellsche Paradoxon*.

Die Ursache für diesen Widerspruch ist, dass die Definition $(*)$ rückbezüglich ist: Wir wollen eine neue Menge M konstruieren, verwenden dabei aber auf der rechten Seite der Definition *alle Mengen*, also u. a. auch die Menge M, die wir gerade erst definieren wollen. Das ist in etwa so, als würdet ihr im Beweis eines Satzes die Aussage des Satzes selbst verwenden – und das ist natürlich nicht zulässig.

Man muss bei der Festlegung, was Mengen sind und wie man sie bilden kann, also eigentlich viel genauer vorgehen, als es Cantor getan hat. Heutzutage verwendet man hierzu in der Regel das im Jahre 1930 aufgestellte Axiomensystem von Zermelo und Fraenkel, das genau angibt, wie man aus bekannten Mengen neue konstruieren darf: z. B. indem man sie schneidet oder vereinigt, oder aus bereits bekannten Mengen Elemente mit einer bestimmten Eigenschaft auswählt. Wir wollen dies hier in dieser Vorlesung aber nicht weiter thematisieren und uns mit der naiven Mengencharakterisierung von Cantor begnügen (sowie der Versicherung meinerseits, dass schon alles in Ordnung ist, wenn wir neue Mengen immer nur aus alten konstruieren und keine rückbezüglichen Definitionen hinschreiben). Genaueres zum Zermelo-Fraenkel-Axiomensystem könnt ihr z. B. in [E, Kapitel 13] nachlesen.

Notation 1.14 (Reelle Zahlen). Unser wichtigstes Beispiel für eine Menge ist die Menge der **reellen Zahlen**, die wir mit $\mathbb{R}$ bezeichnen werden. Wir wollen die Existenz der reellen Zahlen in dieser Vorlesung axiomatisch voraussetzen und begnügen uns daher an dieser Stelle damit zu sagen, dass man sie sich als die Menge der Punkte auf einer Geraden (der „Zahlengeraden") vorstellen kann. Zusätzlich werden wir in den nächsten beiden Kapiteln die mathematischen Eigenschaften von $\mathbb{R}$ exakt angeben (und ebenfalls axiomatisch voraussetzen) – und zwar genügend viele Eigenschaften, um $\mathbb{R}$ dadurch eindeutig zu charakterisieren.

Ich möchte hier noch einmal betonen, dass man die Existenz und die Eigenschaften der reellen Zahlen eigentlich nicht voraussetzen müsste: Man kann das auch allein aus den Axiomen der Logik und Mengenlehre herleiten! Dies wäre jedoch relativ aufwendig und würde euch im Moment mehr verwirren als helfen, daher wollen wir hier darauf verzichten. Wer sich trotzdem dafür interessiert, kann die Einzelheiten hierzu in [E, Kapitel 1 und 2] nachlesen.

Außer den reellen Zahlen sind vor allem noch die folgenden Teilmengen von $\mathbb{R}$ wichtig:

(a) die Menge $\mathbb{N} = \{0, 1, 2, \dots\}$ der **natürlichen Zahlen** (Achtung: Es gibt Bücher, in denen die 0 nicht mit zu den natürlichen Zahlen gezählt wird!);

(b) die Menge $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ der **ganzen Zahlen**;

(c) die Menge $\mathbb{Q} = \{\frac{p}{q} : p, q \in \mathbb{Z}, q \neq 0\}$ der **rationalen Zahlen**.

Offensichtlich sind diese Mengen ineinander enthalten: Es gilt $\mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$. Teilmengen von $\mathbb{R}$, die durch Ungleichungen gegeben sind, schreiben wir in der Regel, indem wir die Ungleichungsbedingung als Index an das Symbol $\mathbb{R}$ schreiben, z. B. $\mathbb{R}_{\geq 0}$ für die Menge $\{x \in \mathbb{R} : x \geq 0\}$ aller nicht-negativen Zahlen.

Notation 1.15. Sind M und N Mengen, so bezeichnen wir mit ...

(a) $M \cap N := \{x : x \in M \text{ und } x \in N\}$ die **Schnittmenge** von M und N. Gilt $M \cap N = \emptyset$, so sagen wir, dass M und N **disjunkt** sind.

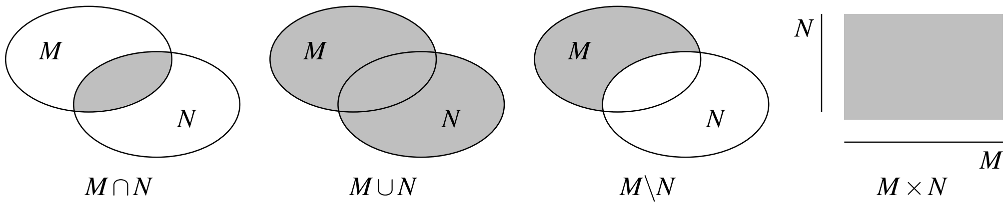
(b) $M \cup N := \{x : x \in M \text{ oder } x \in N\}$ die **Vereinigungsmenge** von M und N. Im Fall einer **disjunkten Vereinigung** mit $M \cap N = \emptyset$ schreiben wir statt $M \cup N$ auch $M \mathbin{\dot\cup} N$.

(c) $M \setminus N := \{x : x \in M \text{ und } x \notin N\}$ die **Differenzmenge** von M und N.

(d) $M \times N := \{(x,y) : x \in M, y \in N\}$ die **Produktmenge** bzw. das Produkt von M und N. Die Schreibweise (x,y) steht hierbei für ein **geordnetes Paar**, d. h. einfach für die Angabe eines Elements aus M und eines aus N (wobei es auch im Fall $M = N$ auf die Reihenfolge ankommt, d. h. (x,y) ist genau dann gleich (x',y') wenn $x = x'$ und $y = y'$). Im Fall $M = N$ schreibt man $M \times N = M \times M$ auch als M^2.

(e) $\mathscr{P}(M) := \{A : A \text{ ist Teilmenge von } M\}$ die **Potenzmenge** von M.

Das Symbol „:=“ bedeutet hierbei, dass der Ausdruck auf der linken Seite durch die rechte Seite definiert wird. Die Konstruktionen (a) bis (d) können durch die folgenden Bilder veranschaulicht werden. Natürlich sind sie auch für mehr als zwei Mengen möglich; aus der Schule kennt ihr zum Beispiel sicher den Fall $\mathbb{R}^3 = \mathbb{R} \times \mathbb{R} \times \mathbb{R}$.



Die Potenzmenge $\mathscr{P}(M)$ aller Teilmengen einer gegebenen Menge M lässt sich dagegen nicht so einfach durch ein Bild darstellen. Es ist z. B.

$$\mathscr{P}(\{0,1\}) = \big\{\emptyset, \{0\}, \{1\}, \{0,1\}\big\}.$$

Aufgabe 1.16. Wie lautet die Negation der folgenden Aussagen? Formuliere außerdem die Aussage (a) in Worten (also analog zu (b)) sowie die Aussage (b) mit Quantoren und anderen mathematischen Symbolen (also analog zu (a)).

(a) $\forall n \in \mathbb{N}\ \exists m \in \mathbb{N} : n = 2m$.

(b) Zwischen je zwei verschiedenen reellen Zahlen gibt es noch eine weitere reelle Zahl.

(c) Sind M, N, R Mengen mit $R \subset N \subset M$, so ist $M \setminus N \subset M \setminus R$.

Aufgabe 1.17. Es seien A, B, C Aussagen und M, N, R Mengen. Man zeige:

(a) $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$ und $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$.

(b) $M \cup (N \cap R) = (M \cup N) \cap (M \cup R)$ und $M \cap (N \cup R) = (M \cap N) \cup (M \cap R)$.

Aufgabe 1.18. Welche der folgenden Aussagen sind für beliebige gegebene x und M äquivalent zueinander? Zeige jeweils die Äquivalenz bzw. widerlege sie durch ein Gegenbeispiel.

(a) $x \in M$	(b) $\{x\} \subset M$	(c) $\{x\} \cap M \neq \emptyset$
(d) $\{x\} \in M$	(e) $\{x\} \setminus M = \emptyset$	(f) $M \setminus \{x\} = \emptyset$

Aufgabe 1.19. Man beweise oder widerlege: Für alle Mengen $A \subset M$ und $A' \subset M'$ gibt es Teilmengen B und C von M sowie B' und C' von M', so dass

$$(M \times M') \setminus (A \times A') = (B \times B') \mathbin{\dot\cup} (C \times C').$$

Könnt ihr die Aussage durch eine Skizze veranschaulichen?

2. Relationen und Funktionen

Nachdem wir Mengen eingeführt haben, wollen wir nun auch mehrere von ihnen miteinander in Beziehung setzen können. Das Grundkonzept hierfür ist das einer Relation.

Definition 2.1 (Relationen)**.** Es seien M und N zwei Mengen. Eine **Relation** zwischen M und N ist eine Teilmenge R des Produkts $M \times N$. Für $x \in M$ und $y \in N$ mit $(x,y) \in R$ sagen wir dann „*x steht (bezüglich R) in Relation zu y*". Ist $M = N$, so nennen wir R auch eine *Relation auf M*.

Bemerkung 2.2. Um eine Relation R anzugeben, also eine Teilmenge $R \subset M \times N$ zu definieren, müssen wir demzufolge einfach für alle Paare (x,y) mit $x \in M$ und $y \in N$ festlegen, ob $(x,y) \in R$ gelten, also ob x in Relation zu y stehen soll.

Wie wir in diesem Kapitel sehen werden, werden Relationen in der Mathematik für sehr unterschiedliche Konzepte verwendet – z. B. um Zahlen miteinander zu vergleichen wie in Beispiel 2.3, um eine Menge auf eine andere abzubilden wie in Abschnitt 2.A, oder um die Elemente einer Menge nach bestimmten Kriterien zu Klassen zusammenzufassen wie in Abschnitt 2.B. Dementsprechend sind für die Aussage „x steht bezüglich R in Relation zu y" auch je nach Anwendung ganz unterschiedliche Notationen üblich. Für allgemeine, nicht näher spezifizierte Relationen schreibt man hierfür oft xRy.

Beispiel 2.3 (Kleiner-Relation)**.** Für $M = N = \mathbb{R}$ betrachten wir die Relation

$$R = \{(x,y) : x,y \in \mathbb{R} \text{ mit } x < y\},$$

für die x also genau dann in Relation zu y steht, wenn $x < y$ gilt. Man nennt R deshalb auch die *Kleiner-Relation* auf $\mathbb{R}$. Die Notation „xRy" aus Bemerkung 2.2 stimmt in diesem Fall also mit der Schreibweise „$x < y$" überein, wenn man die Relation R direkt mit dem Symbol „$<$" bezeichnet. In der Tat ist es aus diesem Grund bei manchen Relationen üblich, sie gleich mit Symbolen statt mit Buchstaben zu benennen.

2.A Funktionen

Die mit Abstand wichtigsten Relationen sind ohne Zweifel die Funktionen, die ihr natürlich bereits hinlänglich aus der Schule kennt. Wir wollen sie hier nun exakt einführen und ihre ersten Eigenschaften untersuchen.

Definition 2.4 (Funktionen)**.** Es seien M und N zwei Mengen.

(a) Eine **Funktion** oder **Abbildung** f von M nach N, geschrieben $f : M \to N$, ist eine Relation zwischen M und N, bezüglich der jedes Element x von M zu *genau einem* Element y von N in Relation steht. Wir schreiben dies dann als $x \mapsto y$ oder $y = f(x)$ und sagen, y ist das **Bild** von x unter f bzw. der **Wert** von f in x.

(b) Für eine Funktion $f : M \to N$ bezeichnet man die Menge M als **Definitionsmenge**, **Startmenge** oder **Startraum** von f. Die Menge N heißt **Zielmenge** oder **Zielraum** von f.

Bemerkung 2.5.

(a) Um eine Funktion komplett festzulegen, müssen wir zuerst einmal den Start- und Zielraum angeben, und dann schließlich noch von jedem Element des Startraums sagen, auf welches Element des Zielraums es abgebildet wird. In welcher Form wir diese Zuordnung angeben – ob durch eine Formel, durch explizites Auflisten der Funktionswerte aller Elemente des Startraums, oder irgendwie anders – spielt dabei keine Rolle. So sind z. B.

$$f : \{0,1\} \to \mathbb{R},\ x \mapsto 2x^2, \quad g : \{0,1\} \to \mathbb{R},\ x \mapsto 2x^3, \quad h : \{0,1\} \to \mathbb{R},\ x \mapsto \begin{cases} 0 & \text{für } x = 0, \\ 2 & \text{für } x = 1 \end{cases}$$

trotz ihrer ganz verschieden aussehenden Vorschriften dieselbe Funktion, da alle drei den gleichen Start- und Zielraum haben und aus den gleichen Zuordnungen $0 \mapsto 0$ und $1 \mapsto 2$ bestehen. Mit anderen Worten sind zwei Funktionen $f, g\colon M \to N$ also genau dann gleich, wenn sie an jedem Punkt die gleichen Werte besitzen, also wenn gilt

$$\forall x \in M : f(x) = g(x).$$

(b) Man sieht leider oft, dass eine Funktion $f\colon M \to N$ als $f(x)$ geschrieben wird. Es ist wichtig zu verstehen, dass diese Notation gemäß Definition 2.4 falsch ist: Mit $f(x)$ wird *der Wert der Funktion f in einem Punkt $x \in M$* bezeichnet. Somit ist $f(x)$ (für gegebenes x) ein Element von N, und damit ein ganz anderes mathematisches Objekt als die Funktion selbst, die wir nur mit f bezeichnen und die eine Relation zwischen M und N ist. Dies mag auf den ersten Blick spitzfindig erscheinen – wir werden aber später noch oft Mengen sehen, deren Elemente Funktionen sind, und dann ist es natürlich wichtig, dies von der Menge ihrer Funktions*werte* zu unterscheiden.

Beispiel 2.6.

(a) Die Zuordnungen

$$f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \frac{1}{x} \qquad \text{und} \qquad g\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} x+1 & \text{für } x \leq 0 \\ x & \text{für } x \geq 0 \end{cases}$$

sind in dieser Form keine zulässigen Funktionsdefinitionen, weil im Fall von f der Zahl 0 kein gültiger Funktionswert zugeordnet wird und im Fall g für die Zahl 0 zwei (sich widersprechende) Festlegungen des Funktionswertes gemacht werden. Dies lässt sich jedoch in beiden Fällen leicht reparieren, z. B. indem man die Festlegungen abändert in

$$f\colon \mathbb{R}\backslash\{0\} \to \mathbb{R},\ x \mapsto \frac{1}{x} \qquad \text{und} \qquad g\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} x+1 & \text{für } x < 0 \\ x & \text{für } x \geq 0 \end{cases}.$$

(b) Zu jeder Menge M gibt es die **identische Abbildung**

$$\mathrm{id}_M\colon M \to M,\ x \mapsto x,$$

die jedes Element auf sich selbst abbildet.

(c) Ist $f\colon M \to N$ eine Abbildung und $A \subset M$ eine Teilmenge des Startraums, so erhält man durch die Einschränkung der Definitionsmenge von M auf A eine neue Abbildung, die wir mit

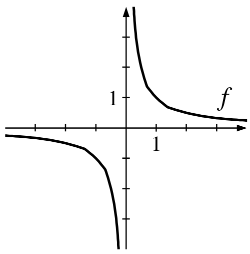
$$f|_A\colon A \to N,\ x \mapsto f(x)$$

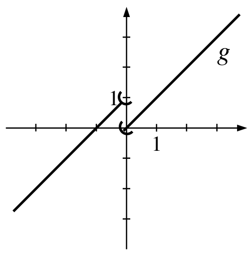
bezeichnen und die die **Einschränkung** von f auf A genannt wird. Genauso kann man natürlich auch die Zielmenge N auf eine Teilmenge B einschränken, wenn f nur Werte in B annimmt. Es ist üblich, bei einer derartigen Einschränkung der Zielmenge immer noch den gleichen Namen für die Abbildung zu verwenden, also dann $f\colon M \to B$ zu schreiben (auch wenn es sich dabei um eine andere Funktion als das ursprüngliche $f\colon M \to N$ handelt).

Bemerkung 2.7 (Graph einer Abbildung)**.** Zu einer Abbildung $f\colon M \to N$ heißt die Menge

$$\{(x, f(x)) : x \in M\} \quad \subset M \times N$$

der **Graph** von f. Sind M und N Teilmengen von $\mathbb{R}$, so ist dieser Graph also eine Teilmenge von $\mathbb{R}^2$, und man kann ihn leicht zeichnen und dadurch die Abbildung veranschaulichen. Für die Abbildungen aus Beispiel 2.6 (a) sieht dies z. B. so aus:





Beachte, dass dieser Graph nach den Definitionen 2.1 und 2.4 eigentlich sogar genau das gleiche ist wie die Funktion selbst, nämlich die Teilmenge des Produkts $M \times N$, die aus den Paaren (x,y) besteht, für die x bezüglich f in Relation zu y steht, also $y = f(x)$ gilt. Der Begriff des Graphen soll hier also nur noch einmal deutlich machen, dass man sich die Funktion gerade wirklich als ein derart „grafisches“ Objekt vorstellt und nicht als eine „Zuordnung“ von M nach N.

In der Definition 2.4 einer Abbildung $f\colon M \to N$ verlangen wir, dass jedem Element von M genau ein Element von N zugeordnet wird. Wir fordern jedoch nicht auch umgekehrt, dass jedes Element des Zielraums N das Bild von genau einem Element von M ist, oder dass es überhaupt als Bild eines Elements von M auftritt. Abbildungen, die diese Eigenschaften dennoch besitzen, haben spezielle Namen, die wir jetzt einführen wollen.

Definition 2.8 (Eigenschaften von Abbildungen)**.** Es sei $f\colon M \to N$ eine Abbildung.

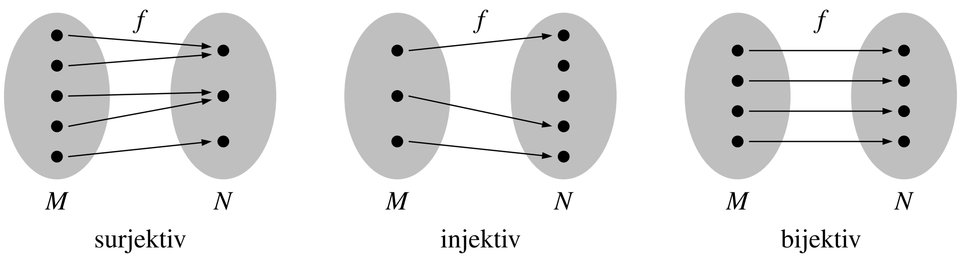
(a) Ist $y \in N$ und $x \in M$ mit $f(x) = y$, so heißt x ein **Urbild** von y unter f.

(b) Hat jedes $y \in N$ ...

  (i) *mindestens* ein Urbild, so heißt f **surjektiv**.
  In Quantoren bedeutet dies: $\forall y \in N \; \exists x \in M\colon \; f(x) = y$.

  (ii) *höchstens* ein Urbild, so heißt f **injektiv**.
  In Quantoren bedeutet dies: $\forall x_1, x_2 \in M\colon \; f(x_1) = f(x_2) \Rightarrow x_1 = x_2$. (Also: Haben zwei Elemente des Startraums das gleiche Bild, so müssen sie bereits dasselbe Element sein.)

  (iii) *genau* ein Urbild, ist f also surjektiv und injektiv, so heißt f **bijektiv**.



Beispiel 2.9. Betrachten wir noch einmal die Funktionen aus Beispiel 2.6 (a), so ist die Funktion f nicht surjektiv (und damit auch nicht bijektiv), da das Element 0 des Zielraums kein Urbild hat. Sie ist jedoch injektiv: Sind $x_1, x_2 \in \mathbb{R}\backslash\{0\}$ mit $f(x_1) = f(x_2)$, also $\frac{1}{x_1} = \frac{1}{x_2}$, so folgt durch Multiplikation mit $x_1 x_2$ sofort $x_1 = x_2$.

Die Funktion g dagegen ist surjektiv: Eine Zahl $y \in \mathbb{R}$ hat als Urbild $x = y$ für $y \geq 0$, und $x = y - 1$ für $y < 0$. Sie ist allerdings nicht injektiv, denn es ist $g(-1) = g(0) = 0$.

Beachte, dass diese Eigenschaften auch an den Graphen in Bemerkung 2.7 ablesbar sind: Surjektivität bzw. Injektivität bedeuten gerade, dass jede horizontale Gerade auf der Höhe eines Wertes im Zielraum den Funktionsgraphen in mindestens bzw. höchstens einem Punkt schneidet. Wichtig ist auch, dass diese Eigenschaften von der Wahl des Start- und Zielraums abhängen: So wird z. B. f bijektiv, wenn man den Zielraum $\mathbb{R}$ durch $\mathbb{R}\backslash\{0\}$ ersetzt, und g injektiv, wenn man den Startraum auf $\mathbb{R}_{\geq 0}$ einschränkt (in der Notation von Beispiel 2.6 (c) also $g|_{\mathbb{R}_{\geq 0}}$ betrachtet).

Aufgabe 2.10. Wie viele Abbildungen gibt es zwischen den Mengen $\{1,2,3,4\}$ und $\{1,2,3,4,5,6\}$? Wie viele von ihnen sind injektiv? 02

Bilder und Urbilder unter Abbildungen betrachtet man oft auch von ganzen Mengen statt nur von Punkten:

Definition 2.11 (Bild und Urbild von Mengen)**.** Es sei $f: M \to N$ eine Abbildung.

(a) Für $A \subset M$ heißt die Menge

$$f(A) := \{f(x) : x \in A\} \quad \subset N$$

(also die Menge aller Bilder von Punkten in A) das **Bild** von A unter f. Die Menge $f(M)$ nennt man auch das Bild von f.

(b) Ist $B \subset N$, so heißt die Menge

$$f^{-1}(B) := \{x \in M : f(x) \in B\} \quad \subset M$$

(also die Menge aller Urbilder von Punkten in B) das **Urbild** von B unter f.

Bemerkung 2.12. Die Grundidee der Notation in Definition 2.11 (a) ist: Schreiben wir als Argument einer Funktion $f: M \to N$ eine *Teilmenge* statt einem *Element* von M, so bedeutet dies, dass wir alle Werte $f(x)$ für $x \in M$ zusammen nehmen und diese wieder in einer Menge zusammenfassen. Diese Schreibweise verwendet man auch oft, wenn die Funktion aus einer Rechenverknüpfung besteht, wie z. B. in

$$\mathbb{N}+\tfrac{1}{2} := \left\{n+\tfrac{1}{2} : n \in \mathbb{N}\right\} = \left\{\tfrac{1}{2}, \tfrac{3}{2}, \tfrac{5}{2}, \dots\right\} \quad \text{oder} \quad 3\mathbb{Z} := \{3n : n \in \mathbb{Z}\} = \{\dots, -6, -3, 0, 3, 6, \dots\}.$$

Beispiel 2.13. Für die Funktion $f: \mathbb{R}\backslash\{0\} \to \mathbb{R},\ x \mapsto \frac{1}{x}$ aus Beispiel 2.6 (a) ist $f(\mathbb{R}_{>0}) = \mathbb{R}_{>0}$ und $f^{-1}(\{0\}) = \emptyset$.

Beispiel 2.14. Zwischen den Konstruktionen von Bild und Urbild aus Definition 2.11 und den Mengenoperationen aus Abschnitt 1.B gibt es sehr viele Beziehungen. Um einmal exemplarisch zu sehen, wie derartige Beziehungen aussehen und bewiesen werden können, wollen wir nun zeigen, dass für jede Abbildung $f: M \to N$ und zwei beliebige Teilmengen $A, B \subset M$ stets

$$f(A)\backslash f(B) \subset f(A\backslash B) \tag{$*$}$$

gilt.

Zum Beweis müssen wir zeigen, dass jedes Element der linken Menge auch in der rechten Menge liegt. Es sei also $y \in f(A)\backslash f(B)$ beliebig. Insbesondere ist damit $y \in f(A)$, nach Definition 2.11 (a) also $y = f(x)$ für ein $x \in A$. Würde nun auch $x \in B$ gelten, so hätten wir wegen $y = f(x)$ auch $y \in f(B)$, im Widerspruch zu $y \in f(A)\backslash f(B)$. Also ist $x \notin B$, und damit $x \in A\backslash B$. Damit besagt $y = f(x)$ aber gerade $y \in f(A\backslash B)$. Insgesamt haben wir somit die behauptete Teilmengenbeziehung $(*)$ gezeigt.

Beachte allerdings, dass in $(*)$ im Allgemeinen keine Gleichheit gilt: Für $f: \mathbb{R} \to \mathbb{R},\ x \mapsto x^2$ mit $A = \{-1, 1\}$ und $B = \{-1\}$ ist

$$f(A)\backslash f(B) = \{1\}\backslash\{1\} = \emptyset, \quad \text{aber} \quad f(A\backslash B) = f(\{1\}) = \{1\}.$$

Aufgabe 2.15. Beweise die folgenden Teilmengenbeziehungen und untersuche jeweils, ob auch die Gleichheit gilt.

(a) Für alle Mengen M, A, B gilt $M\backslash(A \cup B) \subset (M\backslash A) \cap (M\backslash B)$.

(b) Ist $f: M \to N$ eine Abbildung und $A \subset N$, so ist $f\left(f^{-1}(A)\right) \subset A$.

Aufgabe 2.16. Es sei $f: M \to N$ eine Abbildung. Finde für das Symbol $\square$ jeweils eine der Mengenbeziehungen $\subset, =, \supset$, so dass die folgenden Aussagen wahr werden, und beweise die so entstandenen Aussagen!

(a) $f(A) \cap f(B) \ \square \ f(A \cap B)$ für alle $A, B \subset M$.

(b) $f^{-1}(A) \cap f^{-1}(B) \ \square \ f^{-1}(A \cap B)$ für alle $A, B \subset N$.

Als Nächstes wollen wir nun die euch sicher bereits bekannte Verkettung, also die Hintereinanderausführung von Funktionen einführen.

Definition 2.17 (Verkettung von Funktionen)**.** Es seien $f\colon M \to N$ und $g\colon N \to R$ zwei Abbildungen (also so dass die Zielmenge von f gleich der Startmenge von g ist). Dann heißt die Abbildung

$$g \circ f\colon M \to R,\ x \mapsto g(f(x))$$

die **Verkettung** von f und g.

Bemerkung 2.18. Bei der Verkettung zweier Funktionen kommt es natürlich auf die Reihenfolge an, allein schon weil in der Situation von Definition 2.17 in der Regel der Zielraum von g ja nicht mit dem Startraum von f übereinstimmt und die „umgekehrte Verkettung“ $f \circ g$ damit gar nicht definierbar wäre. Beachte dabei, dass die Notation $g \circ f$ lautet, obwohl wir zuerst f (von M nach N) und dann g (von N nach R) anwenden – man liest $g \circ f$ daher manchmal auch als „g nach f“. Diese vielleicht etwas merkwürdig erscheinende Notation kommt einfach daher, dass die Buchstaben in der gleichen Reihenfolge stehen sollen wie bei der Abbildungsvorschrift $x \mapsto g(f(x))$.

Wir werden nun unser erstes *Lemma* beweisen – „Lemma“ ist griechisch und bedeutet eigentlich „Annahme“, aber in der Mathematik wird dieser Begriff für einen *Hilfssatz* verwendet, also für ein kleines Zwischenresultat, das vielleicht für sich genommen nicht übermäßig überraschend oder interessant ist, aber das in späteren Beweisen immer wieder nützlich sein wird. In unserem momentanen Fall geht es einfach darum, dass die Verkettung von Abbildungen *assoziativ* ist (siehe auch Definition 3.1):

Lemma 2.19 (Assoziativität der Verkettung)**.** *Sind $f\colon M \to N$, $g\colon N \to R$ und $h\colon R \to S$ drei Abbildungen, so gilt $h \circ (g \circ f) = (h \circ g) \circ f$. (Man schreibt für diese Abbildung daher oft auch einfach $h \circ g \circ f$.)*

Beweis. Nach Definition 2.4 können wir die Gleichheit zweier Funktionen zeigen, indem wir für jedes Element der Startmenge nachweisen, dass sein Bild unter beiden Funktionen übereinstimmt. Dies rechnen wir nun einfach durch wiederholtes Einsetzen von Definition 2.17 nach: Es gilt

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

und

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

Da diese beiden Ausdrücke übereinstimmen, ist das Lemma bewiesen. □

Am Ende dieses Abschnitts wollen wir schließlich noch das Konzept von Umkehrfunktionen bijektiver Funktionen einführen.

Definition 2.20 (Umkehrfunktionen)**.** Es sei $f\colon M \to N$ eine bijektive Funktion. Dann heißt

$$f^{-1}\colon N \to M,\ y \mapsto \text{das eindeutige Urbild von } y \text{ unter } f$$

die **Umkehrfunktion** bzw. **Umkehrabbildung** von f.

Bemerkung 2.21. Für die Umkehrfunktion f^{-1} einer bijektiven Funktion $f\colon M \to N$ gilt nach Konstruktion offensichtlich $f^{-1} \circ f = \mathrm{id}_M$ und $f \circ f^{-1} = \mathrm{id}_N$.

Gibt es umgekehrt zu einer Funktion $f\colon M \to N$ eine Abbildung $g\colon N \to M$ mit $g \circ f = \mathrm{id}_M$ und $f \circ g = \mathrm{id}_N$, so ist f bijektiv:

- f ist surjektiv: Ist $y \in N$ beliebig, so ist $x := g(y) \in M$ ein Urbild von y unter f, denn es ist $f(x) = f(g(y)) = y$.
- f ist injektiv: Sind $x_1, x_2 \in M$ mit $f(x_1) = f(x_2)$, so folgt durch Anwenden von g sofort auch $g(f(x_1)) = g(f(x_2))$, und damit $x_1 = x_2$.

Beispiel 2.22. Die Funktion $f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto x+1$ ist bijektiv, und ihre Umkehrabbildung ist $f^{-1}\colon \mathbb{R} \to \mathbb{R},\ x \mapsto x-1$. In der Tat gilt nämlich für alle $x \in \mathbb{R}$

$$(f^{-1} \circ f)(x) = (x+1)-1 = x \quad \text{und} \quad (f \circ f^{-1})(x) = (x-1)+1 = x.$$

Bemerkung 2.23 (Urbilder und Umkehrabbildungen)**.** Beachte, dass wir das Urbild einer Menge unter einer Abbildung $f: M \to N$ in Definition 2.11 (b) mit dem gleichen Symbol f^{-1} bezeichnet haben wie (im Fall einer bijektiven Abbildung) die Umkehrabbildung aus Definition 2.20. Das ist vielleicht etwas unglücklich gewählt, aber in der Literatur so fest verankert, dass wir hier nicht davon abweichen wollen. Bei genauem Hinschauen kann man aber auch immer feststellen, was gemeint ist:

> Ist $f: M \to N$ eine Abbildung, so bezeichnet ...
>
> ... $f^{-1}(B)$ für eine *Menge* $B \subset N$ das *Urbild* von B wie in Definition 2.11 (b); es existiert für jede Abbildung f.
>
> ... $f^{-1}(y)$ für ein *Element* $y \in B$ den *Wert der Umkehrabbildung* bei y wie in Definition 2.20; er existiert nur für bijektives f.

Letztlich hängen diese beiden Notationen aber auch eng miteinander zusammen: Ist f bijektiv und ist $x \in M$ mit $f(x) = y$, so ist $f^{-1}(y) = x$ (mit f^{-1} im Sinne der Umkehrabbildung) und $f^{-1}(\{y\}) = \{x\}$ (mit f^{-1} im Sinne des Urbildes).

Aufgabe 2.24.

(a) Untersuche die Abbildung $f: \mathbb{Z} \to \mathbb{Z}$, $x \mapsto 3x+2$ auf Injektivität und Surjektivität.

(b) Untersuche die Abbildung $f: \mathbb{R}^2 \to \mathbb{R}^2$, $(x,y) \mapsto (xy, x+1)$ auf Injektivität und Surjektivität.

(c) Man zeige: Sind $f: M \to N$ und $g: N \to R$ surjektiv, so ist auch $g \circ f: M \to R$ surjektiv.

Aufgabe 2.25. Es seien $f: M \to N$ und $g: N \to R$ bijektiv. Zeige, dass dann auch $f^{-1}: N \to M$ und $g \circ f: M \to R$ bijektiv sind.

Aufgabe 2.26. Man beweise oder widerlege:

(a) Sind $f: M \to N$ und $g: N \to R$ zwei Abbildungen und ist $g \circ f$ injektiv, so ist auch f injektiv.

(b) Sind $f: M \to N$ und $g: N \to R$ zwei Abbildungen und ist $g \circ f$ injektiv, so ist auch g injektiv.

Aufgabe 2.27. Es sei $f: M \to N$ eine Abbildung zwischen nicht-leeren Mengen. Man zeige:

(a) f ist genau dann surjektiv, wenn es eine Abbildung $g: N \to M$ gibt mit $f \circ g = \mathrm{id}_N$.

(b) f ist genau dann injektiv, wenn es eine Abbildung $g: N \to M$ gibt mit $g \circ f = \mathrm{id}_M$.

Aufgabe 2.28. Es sei $f: M \to N$ eine Abbildung zwischen zwei Mengen. Man beweise:

$$f \text{ ist surjektiv} \quad \Leftrightarrow \quad \text{für alle } A, B \subset N \text{ mit } f^{-1}(A) = f^{-1}(B) \text{ gilt } A = B.$$

2.B Äquivalenzrelationen

Am Anfang dieses Kapitels haben wir allgemeine Relationen eingeführt, als einzigen Spezialfall davon aber bisher nur die Funktionen ausführlicher betrachtet. Wir wollen daher nun noch einen ganz anderen wichtigen Typ von Relationen studieren, die sogenannten Äquivalenzrelationen.

Angenommen, wir möchten eine Menge M untersuchen, die uns zunächst einmal zu groß oder zu kompliziert erscheint. Es gibt dann zwei prinzipiell verschiedene Möglichkeiten, wie man daraus eine kleinere bzw. einfachere Menge machen kann:

- Wir können uns auf eine Teilmenge von M beschränken – dann schließen wir allerdings manche Elemente von M von unserer Untersuchung aus.
- Wir können zwar alle Elemente von M betrachten, aber manche von ihnen miteinander identifizieren bzw. zu Klassen zusammenfassen – d. h. sie als gleich bzw. „äquivalent" ansehen, wenn sie für das zu untersuchende Problem ähnliche Eigenschaften haben.

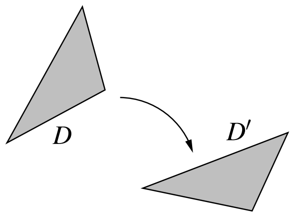
Diese zweite Idee der Identifizierung ähnlicher Elemente führt zum Begriff der Äquivalenzrelationen. Sie klingt vielleicht zunächst etwas abstrakt, ist euch aber sicher schon an vielen Stellen begegnet. Hier sind zwei einfache Beispiele dafür.

Beispiel 2.29.

(a) Eine analoge Uhr vereinfacht die recht große Menge aller (vergangenen und zukünftigen) Zeitpunkte, die wir uns als Zeitachse $M = \mathbb{R}$ vorstellen können, indem sie nach jeweils 12 Stunden wieder dasselbe anzeigt. Sie identifiziert also zwei Zeitpunkte $x, y \in \mathbb{R}$ (gemessen in Stunden) miteinander, wenn $x - y$ ein ganzzahliges Vielfaches von 12 ist. Dadurch „verkleinert" sie die ursprüngliche Zeitachse auf ein gut überschaubares Intervall von 12 Stunden – und wir alle wissen, dass uns ein Blick auf die Uhr in vielen Fällen ausreicht, wenn wir den aktuellen Zeitpunkt wissen wollen, auch wenn uns das nichts über das Datum oder die Tageszeit (vormittags oder nachmittags) sagt.

(b) Als „mathematischeres" Beispiel können wir die Menge M aller Dreiecke in der Ebene $\mathbb{R}^2$ betrachten.

Bekanntlich heißen zwei solche Dreiecke $D, D' \in M$ zueinander *kongruent*, wenn sie wie im Bild rechts durch eine Drehung und / oder Verschiebung auseinander hervorgehen – wir schreiben dies im Folgenden als $D \sim D'$. Zueinander kongruente Dreiecke werden oft miteinander identifiziert, nämlich immer dann, wenn es uns nur auf die Form bzw. Größe der Dreiecke, aber nicht auf ihre Lage in der Ebene ankommt.



Wenn wir z. B. sagen, dass die drei Seitenlängen ein Dreieck eindeutig bestimmen, dann meinen wir damit in Wirklichkeit, dass sie das Dreieck *bis auf Kongruenz* eindeutig bestimmen, also nur die Form und Größe festlegen, aber nicht die Lage des Dreiecks in $\mathbb{R}^2$. Formal kann man dies so ausdrücken: zu einem Dreieck D nennt man

$$\overline{D} := \{D' \in M : D' \sim D\},$$

also die Menge aller zu D kongruenten Dreiecke, die *Kongruenzklasse* von D. Die Menge aller dieser Kongruenzklassen bezeichnen wir mit

$$M/\sim \;:=\; \{\overline{D} : D \in M\}.$$

Man kann dann z. B. sagen, dass die Seitenlängen eines Dreiecks ein eindeutiges Element in $M/\sim$ bestimmen, also eine eindeutige Kongruenzklasse von Dreiecken festlegen – nicht aber ein eindeutiges Element von M.

Mit der Idee dieser Beispiele im Kopf wollen wir nun den Begriff der Äquivalenzrelation exakt definieren.

Definition 2.30 (Äquivalenzrelationen)**.** Es sei $\sim$ wie in Definition 2.1 eine Relation auf einer Menge M. Wie in Bemerkung 2.2 schreiben wir $x \sim y$, wenn x und y bezüglich $\sim$ in Relation stehen.

Man nennt $\sim$ eine **Äquivalenzrelation** auf M, wenn die folgenden Eigenschaften gelten:

(a) **Reflexivität**: Für alle $x \in M$ gilt $x \sim x$.

(b) **Symmetrie**: Sind $x, y \in M$ mit $x \sim y$, so gilt auch $y \sim x$.

(c) **Transitivität**: Sind $x, y, z \in M$ mit $x \sim y$ und $y \sim z$, so gilt auch $x \sim z$.

In diesem Fall sagt man statt $x \sim y$ auch, dass x (bezüglich dieser Relation) zu y **äquivalent** ist. Zu $x \in M$ heißt dann die Menge

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

aller Elemente, die zu x äquivalent sind, die **Äquivalenzklasse** bzw. einfach **Klasse** von x; jedes Element dieser Menge nennt man einen **Repräsentanten** dieser Klasse. Die Menge aller Äquivalenzklassen schreiben wir als

$$M/\sim \;:=\; \{\overline{x} : x \in M\}.$$

Beispiel 2.31.

(a) Das Beispiel 2.29 (a) einer analogen Uhr lässt sich mathematisch exakt wie folgt definieren: Auf $M = \mathbb{R}$ betrachten wir die Relation

$$x \sim y \;:\Leftrightarrow\; \text{es gibt ein } k \in \mathbb{Z} \text{ mit } x - y = 12k. \tag{$*$}$$

Dies ist in der Tat eine Äquivalenzrelation, denn für alle $x,y,z \in \mathbb{R}$ gilt:

- Reflexivität: Es gilt $x - x = 12 \cdot 0 = 12k$ mit $k = 0 \in \mathbb{Z}$, also $x \sim x$.
- Symmetrie: Es gelte $x \sim y$, also $x - y = 12k$ für ein $k \in \mathbb{Z}$. Durch Multiplikation mit -1 folgt dann auch $y - x = 12 \cdot (-k) = 12k'$ mit $k' := -k \in \mathbb{Z}$, und damit $y \sim x$.
- Transitivität: Es gelte $x \sim y$ und $y \sim z$, nach Definition der Relation also $x - y = 12k$ und $y - z = 12k'$ für gewisse $k, k' \in \mathbb{Z}$ (beachte, dass der Wert von k in $(*)$ von x und y abhängt und wir daher für die Differenz $y - z$ eine neue Variable k' brauchen). Durch Addition dieser beiden Gleichungen erhalten wir $x - z = 12(k + k') = 12k''$ mit $k'' := k + k' \in \mathbb{Z}$, und damit $x \sim z$.

Die Äquivalenzklasse z. B. von $2 \in \mathbb{R}$ ist

$$\overline{2} = \{x \in \mathbb{R} : x \sim 2\} = \{x \in \mathbb{R} : \text{es gibt ein } k \in \mathbb{Z} \text{ mit } x - 2 = 12k\} = \{2 + 12k : k \in \mathbb{Z}\},$$

also die Menge aller Zeitpunkte, zu denen die Uhr auf 2 steht. Jeder beliebige Zeitpunkt $x \in \mathbb{R}$, zu dem die Uhr auf 2 steht, ist ein Repräsentant dieser Klasse, und die Menge $M/\sim$ entspricht allen möglichen Ständen der Uhr.

(b) Die Kongruenz von Dreiecken aus Beispiel 2.29 (b) ist ebenfalls eine Äquivalenzrelation (es ist offensichtlich, dass sie die Eigenschaften aus Definition 2.30 erfüllt). Die Äquivalenzklassen sind in diesem Fall genau die Kongruenzklassen.

(c) Die Kleiner-Relation auf $\mathbb{R}$ aus Beispiel 2.3, also die Relation, für die für $x, y \in \mathbb{R}$ genau dann $x \sim y$ gilt, wenn $x < y$ ist, ist keine Äquivalenzrelation, da sie weder reflexiv noch symmetrisch ist.

Beachte, dass bei unseren Äquivalenzrelationen aus Beispiel 2.31 (a) und (b) jedes Element von M in genau einer Äquivalenzklasse liegt: Zu jedem Zeitpunkt hat eine analoge Uhr genau einen Stand, und jedes Dreieck in der Ebene liegt in genau einer Kongruenzklasse. Dies beschreibt genau unsere Idee, dass wir die Elemente von M auf eine bestimmte Art zu Klassen zusammenfassen wollen. Allgemein sind die Axiome einer Äquivalenzrelation aus Definition 2.30 anschaulich genau diejenigen, die man braucht, damit die Relation sinnvoll eine solche Identifizierung von Elementen zu Klassen beschreiben kann. Dies zeigt auch noch einmal der folgende zentrale Satz über Äquivalenzrelationen.

Satz 2.32 (Eigenschaften von Äquivalenzrelationen)**.** *Es sei $\sim$ eine Äquivalenzrelation auf einer Menge M.*

(a) *Für $x, y \in M$ gilt $x \sim y$ genau dann, wenn $\overline{x} = \overline{y}$. (Zwei Elemente sind also genau dann äquivalent zueinander, wenn sie die gleiche Äquivalenzklasse bestimmen.)*

(b) *Jedes Element $x \in M$ liegt in genau einer Äquivalenzklasse (nämlich in $\overline{x}$). Insbesondere ist M also die disjunkte Vereinigung aller Äquivalenzklassen. Man sagt dafür auch, dass die Äquivalenzklassen eine Partition von M bilden.*

Beweis.

(a) Es seien $x, y \in M$.

„$\Rightarrow$“: Es gelte $x \sim y$. Ist dann $z \in M$ mit $z \in \overline{x}$, also $z \sim x$, so ist nach der Transitivität wegen $x \sim y$ auch $z \sim y$, also $z \in \overline{y}$. Damit gilt $\overline{x} \subset \overline{y}$. Da mit $x \sim y$ wegen der Symmetrie aber auch $y \sim x$ gilt, folgt analog auch umgekehrt $\overline{y} \subset \overline{x}$, und somit insgesamt $\overline{x} = \overline{y}$.

„$\Leftarrow$“: Es sei nun $\overline{x} = \overline{y}$. Wegen der Reflexivität ist $x \sim x$, also $x \in \overline{x} = \overline{y}$, und damit $x \sim y$.

(b) Wegen der Reflexivität liegt natürlich jedes $x \in M$ in seiner eigenen Äquivalenzklasse $\overline{x}$. Ist nun auch $x \in \overline{y}$ für ein $y \in M$, also $x \sim y$, so gilt nach (a) bereits $\overline{y} = \overline{x}$. Also liegt x in genau einer Äquivalenzklasse von $\sim$, nämlich in $\overline{x}$. □

Aufgabe 2.33. Es sei $M = \{n \in \mathbb{Z} : |n| \leq 100\} = \{-100, -99, \dots, 0, \dots, 99, 100\}$. Welche der folgenden Relationen sind Äquivalenzrelationen auf M? Gib im Fall einer Äquivalenzrelation außerdem die Äquivalenzklasse $\overline{-34}$ explizit an.

(a) $x \sim y \;:\Leftrightarrow$ es gibt ein $n \in \mathbb{Z}$ mit $x = 2^n y$.

(b) $x \sim y \;:\Leftrightarrow\; xy \geq 0$.

Aufgabe 2.34. Welche der folgenden Relationen sind Äquivalenzrelationen auf $\mathbb{R}^2$? Im Fall einer Äquivalenzrelation berechne und skizziere man außerdem die Äquivalenzklassen von $(0,1) \in \mathbb{R}^2$ und $(1,1) \in \mathbb{R}^2$.

(a) $(x,y) \sim (x',y') \;:\Leftrightarrow$ es gibt ein $a \in \mathbb{R}\backslash\{0\}$ mit $x = a^2 x'$ und $y = ay'$;

03 (b) $(x,y) \sim (x',y') \;:\Leftrightarrow$ es gibt ein $a \in \mathbb{R}\backslash\{0\}$ mit $x = ay'$ und $y = ax'$.

3. Erste Eigenschaften der reellen Zahlen

In Notation 1.14 haben wir bereits die reellen Zahlen $\mathbb{R}$ als „Menge der Punkte auf einer Geraden" eingeführt. Man kann aber natürlich noch viel mehr Dinge mit den reellen Zahlen tun als sie als eine einfache Punktmenge zu betrachten: Man kann sie addieren, multiplizieren, die Größe von zwei Zahlen miteinander vergleichen, und noch einiges mehr. Wir wollen die Eigenschaften der reellen Zahlen in diesem und dem nächsten Kapitel exakt formalisieren, damit wir danach genau wissen, welche Eigenschaften von $\mathbb{R}$ wir in dieser Vorlesung axiomatisch voraussetzen. In der Tat werden diese Eigenschaften letztlich sogar ausreichen, um die reellen Zahlen eindeutig zu charakterisieren. Wir beginnen in diesem Kapitel aber zunächst einmal nur mit den „Grundrechenarten", also mit der Addition und der Multiplikation sowie ihren Umkehrungen, der Subtraktion und Division.

3.A Gruppen und Körper

Die Eigenschaften von Verknüpfungen wie der Addition oder Multiplikation reeller Zahlen werden mathematisch durch die Begriffe einer Gruppe bzw. eines Körpers beschrieben, die wir jetzt einführen wollen.

Definition 3.1 (Gruppen)**.** Eine **Gruppe** ist eine Menge G zusammen mit einer „Verknüpfung", d. h. einer Abbildung

$$*\colon G\times G\to G,\ (x,y)\mapsto x*y,$$

so dass die folgenden Eigenschaften (auch *Gruppenaxiome* genannt) gelten:

(a) (**Assoziativität**) Für alle $x,y,z\in G$ gilt $(x*y)*z=x*(y*z)$. Man schreibt diesen Ausdruck dann in der Regel auch einfach als $x*y*z$, weil die Reihenfolge der Klammerung ja egal ist.

(b) (Existenz eines neutralen Elements) Es gibt ein $e\in G$, für das $e*x=x*e=x$ für alle $x\in G$ gilt. Man nennt ein solches e ein **neutrales Element**, und verlangt davon zusätzlich:

(c) (Existenz von inversen Elementen) Für alle $x\in G$ gibt es ein $x'\in G$ mit $x'*x=x*x'=e$. Man nennt x' dann ein **inverses Element** zu x.

Wir bezeichnen eine solche Gruppe mit $(G,*)$. Wenn aus dem Zusammenhang klar ist, welche Verknüpfung gemeint ist, schreiben wir oft auch einfach nur G für die Gruppe.

Gilt zusätzlich zu den obigen Eigenschaften noch

(d) (**Kommutativität**) $x*y=y*x$ für alle $x,y\in G$,

so heißt $(G,*)$ eine **kommutative** oder **abelsche Gruppe**.

Bemerkung 3.2. Manchmal wird in der Definition einer Gruppe in Teil (b) lediglich $e*x=x$ und in Teil (c) lediglich $x'*x=e$ gefordert (man spricht dann auch von einem **linksneutralen** bzw. **linksinversen** Element). Man kann jedoch unter Verwendung der übrigen Gruppenaxiome zeigen, dass in diesem Fall automatisch auch $x*e=x$ und $x*x'=e$ gelten muss, also dass linksneutrale Elemente bereits immer neutrale und linksinverse Elemente immer inverse Elemente sind [G, Satz 1.7]. Die beiden Varianten der Definition einer Gruppe stimmen also letztlich überein.

Beispiel 3.3.

(a) $(\mathbb{R},+)$ ist eine abelsche Gruppe, denn die Addition ist (wie wir axiomatisch voraussetzen werden) eine Verknüpfung auf $\mathbb{R}$ mit den Eigenschaften:
   - $(x+y)+z=x+(y+z)$ für alle $x,y,z\in\mathbb{R}$;
   - $0\in\mathbb{R}$ ist ein neutrales Element, denn $0+x=x+0=x$ für alle $x\in\mathbb{R}$;
   - zu jedem $x\in\mathbb{R}$ ist $-x\in\mathbb{R}$ ein inverses Element, denn $(-x)+x=x+(-x)=0$;

- $x+y=y+x$ für alle $x,y\in\mathbb{R}$.

Auf die gleiche Art sind auch $(\mathbb{Q},+)$ und $(\mathbb{Z},+)$ abelsche Gruppen, jedoch nicht $(\mathbb{N},+)$: Hier existiert zwar noch ein neutrales Element 0, aber die Zahl $1\in\mathbb{N}$ hat kein Inverses mehr, denn es gibt kein $x\in\mathbb{N}$ mit $x+1=0$.

(b) $(\mathbb{R},\cdot)$ ist keine Gruppe: Die Multiplikation ist zwar assoziativ und kommutativ und hat das neutrale Element 1, aber die Zahl 0 hat kein Inverses – denn dies müsste ja eine Zahl $x\in\mathbb{R}$ sein mit $x\cdot 0=1$.

Nimmt man jedoch die 0 aus $\mathbb{R}$ heraus, so erhält man mit $(\mathbb{R}\backslash\{0\},\cdot)$ wieder eine abelsche Gruppe, bei der das neutrale Element 1 und das zu einem x inverse Element $\frac{1}{x}$ ist. Genauso funktioniert dies für $(\mathbb{Q}\backslash\{0\},\cdot)$, aber z. B. nicht für $(\mathbb{Z}\backslash\{0\},\cdot)$: Hier gibt es zwar noch ein neutrales Element 1, aber die Zahl $2\in\mathbb{Z}\backslash\{0\}$ hat kein Inverses mehr, denn es gibt kein $x\in\mathbb{Z}\backslash\{0\}$ mit $2\cdot x=1$.

(c) Hier ist noch ein Beispiel von einem ganz anderen Typ: Es sei M eine beliebige Menge und $G=\{f\colon M\to M \text{ bijektiv}\}$ die Menge aller bijektiven Abbildungen von M nach M. Da die Verkettung bijektiver Abbildungen nach Aufgabe 2.25 wieder bijektiv ist, definiert sie eine Verknüpfung auf G. In der Tat wird G damit zu einer Gruppe, denn die Verkettung ist assoziativ nach Lemma 2.19, die Identität id_M ist ein neutrales Element, und zu einem $f\in G$ ist die Umkehrabbildung f^{-1} aus Definition 2.20 ein inverses Element: Sie ist nach Aufgabe 2.25 selbst wieder bijektiv (also in G) und erfüllt $f^{-1}\circ f=f\circ f^{-1}=\mathrm{id}_M$ nach Bemerkung 2.21. Im Allgemeinen ist diese Gruppe jedoch nicht kommutativ.

Wir wollen nun ein paar einfache Eigenschaften von Gruppen beweisen, u. a. dass die in Definition 3.1 geforderten neutralen und inversen Elemente eindeutig sind und wir daher in Zukunft auch von *dem* neutralen und *dem* zu einem gegebenen Element inversen Element sprechen können.

Lemma 3.4 (Eigenschaften von Gruppen)**.** *Es seien $(G,*)$ eine Gruppe und $x,y\in G$.*

(a) *Es gibt genau ein neutrales Element (wie in Definition 3.1 (b)).*

(b) *Es gibt genau ein inverses Element zu x (wie in Definition 3.1 (c)).*

(c) *Sind x' und y' die inversen Elemente zu x bzw. y, so ist $y'*x'$ das inverse Element zu $x*y$.*

(d) *Ist x' das inverse Element zu x, so ist x das inverse Element zu x' („das Inverse des Inversen ist wieder das Ausgangselement“).*

Beweis.

(a) Sind e und $\tilde{e}$ neutrale Elemente, so folgt

$$\begin{aligned} e &= \tilde{e}*e && \text{(denn } \tilde{e} \text{ ist ein neutrales Element)}\\ &= \tilde{e} && \text{(denn } e \text{ ist ein neutrales Element).} \end{aligned}$$

(b) Sind x' und $\tilde{x}'$ inverse Elemente zu x, so gilt

$$\begin{aligned} x' &= e*x' && (e \text{ neutrales Element})\\ &= (\tilde{x}'*x)*x' && (\tilde{x}' \text{ ist ein inverses Element zu } x)\\ &= \tilde{x}'*(x*x') && \text{(Assoziativität)}\\ &= \tilde{x}'*e && (x' \text{ ist ein inverses Element zu } x)\\ &= \tilde{x}' && (e \text{ neutrales Element).} \end{aligned}$$

(c) Es gilt

$$(y'*x')*(x*y)=y'*(x'*x)*y=y'*e*y=y'*y=e$$

und analog auch $(x*y)*(y'*x')=e$. Damit ist $y'*x'$ das inverse Element zu $x*y$.

(d) Die Gleichung $x'*x=x*x'=e$ besagt direkt, dass x das inverse Element zu x' ist. □

Wie wir in Beispiel 3.3 (a) und (b) gesehen haben, erlauben die reellen Zahlen *zwei* grundlegende Gruppenstrukturen: die Addition und (nach Herausnahme der 0) die Multiplikation. Diese beiden Strukturen sind jedoch nicht unabhängig voneinander, da sie durch das Distributivgesetz $(x+y)\cdot z = xz+yz$ für alle $x,y,z \in \mathbb{R}$ miteinander verbunden sind. Eine derartige Kombination zweier Gruppenstrukturen bezeichnet man als einen Körper.

Definition 3.5 (Körper)**.** Ein **Körper** ist eine Menge K zusammen mit zwei Verknüpfungen

$$+ \colon K \times K \to K \quad (\text{genannt } \textit{Addition}) \qquad \text{und} \qquad \cdot \colon K \times K \to K \quad (\text{genannt } \textit{Multiplikation}),$$

so dass die folgenden Eigenschaften (auch *Körperaxiome* genannt) gelten:

(a) $(K,+)$ ist eine abelsche Gruppe. Wir bezeichnen ihr neutrales Element mit 0 und das zu einem $x \in K$ inverse Element mit $-x$.

(b) $(K\backslash\{0\},\cdot)$ ist ebenfalls eine abelsche Gruppe. Wir bezeichnen ihr neutrales Element mit 1 und das zu einem $x \in K\backslash\{0\}$ inverse Element mit x^{-1}.

(c) (**Distributivität**) Für alle $x,y,z \in K$ gilt $(x+y)\cdot z = (x \cdot z)+(y \cdot z)$.

Mit dieser Definition wollen wir nun also axiomatisch voraussetzen:

$$\boxed{\mathbb{R} \text{ ist ein Körper.}}$$

Um Verwirrungen zu vermeiden, werden wir die beiden Verknüpfungen in einem Körper immer mit den Symbolen „+“ und „ · “ bezeichnen. Ebenso werden wir (wie ihr es natürlich gewohnt seid) vereinbaren, dass man den Punkt bei der Multiplikation auch weglassen darf und bei ungeklammerten Ausdrücken zuerst die Multiplikationen und dann die Additionen ausgeführt werden, so dass man also z. B. die Distributivität aus Definition 3.5 (c) auch als $(x+y)z = xz+yz$ schreiben kann.

Es ist jedoch wichtig zu verstehen, dass wir ab jetzt *nicht* mehr voraussetzen werden, dass Addition und Multiplikation in einem Körper wie z. B. $\mathbb{R}$ genau die Verknüpfungen sind, „an die man als Erstes denken würde“ – was auch immer das heißen mag. Stattdessen sind es einfach irgendwelche zwei Verknüpfungen, die die Eigenschaften aus Definition 3.5 haben. Unsere zukünftigen Beweise über Körper wie z. B. $\mathbb{R}$ müssen wir also ausschließlich auf diesen Eigenschaften aufbauen.

Dieser axiomatische Zugang hat zwei Vorteile:

- Zum einen wissen wir dadurch genau, welche Eigenschaften der Grundrechenarten auf den reellen Zahlen wir eigentlich voraussetzen. Es sollte schließlich klar sein, dass wir eine *exakte* Mathematik nicht auf einer *anschaulichen* Vorstellung von $\mathbb{R}$ aufbauen können. Solltet ihr euch also z. B. später einmal dafür interessieren, wie man die Existenz der reellen Zahlen beweisen kann, so wüsstet ihr dann genau, was eigentlich zu beweisen ist: nämlich die Existenz einer Menge mit genau den Eigenschaften, die wir jetzt axiomatisch voraussetzen.
- Zum anderen werdet ihr im Laufe eures Studiums noch viele weitere Körper kennenlernen, z. B. in Kapitel 6 den sehr wichtigen Körper der komplexen Zahlen. Alle Resultate, die nur auf den Körperaxiomen aufbauen, übertragen sich dann also sofort auf diese neuen Fälle, ohne dass man sich darüber noch einmal neu Gedanken machen muss.

Beispiel 3.6.

(a) Neben $\mathbb{R}$ ist auch $\mathbb{Q}$ (mit den gleichen Verknüpfungen wie auf $\mathbb{R}$) ein Körper. Die ganzen Zahlen $\mathbb{Z}$ bilden mit diesen Verknüpfungen jedoch keinen Körper, da $(\mathbb{Z}\backslash\{0\},\cdot)$ nach Beispiel 3.3 (b) keine Gruppe ist. Ebenso ist $\mathbb{N}$ mit diesen Verknüpfungen kein Körper, da hier nach Beispiel 3.3 (a) bereits die Addition keine Gruppenstruktur liefert.

(b) Hier ist ein Beispiel für einen Körper, der sich ganz anders verhält als $\mathbb{R}$ und $\mathbb{Q}$. Wir definieren auf der Menge $K = \{g,u\}$ zwei Verknüpfungen durch die folgenden Tabellen.

+	g	u
g	g	u
u	u	g

und

·	g	u
g	g	g
u	g	u

Die Idee dieser Verknüpfungen ist, dass g für gerade und u für ungerade ganze Zahlen steht. So haben wir in der Tabelle z. B. $g+u$ als u definiert, weil die Addition einer geraden und einer ungeraden Zahl eine ungerade Zahl ergibt.

Man kann zeigen, dass K mit diesen beiden Verknüpfungen einen Körper bildet. Er wird in der Literatur mit $\mathbb{Z}_2$ bezeichnet, da seine Elemente die Reste ganzer Zahlen bei Division durch 2 beschreiben. Um zu beweisen, dass $\mathbb{Z}_2$ ein Körper ist, könnte man z. B. einfach die geforderten Eigenschaften für alle Elemente – es gibt ja nur zwei – explizit nachprüfen. In der Vorlesung „Algebraische Strukturen“ zeigt man allerdings, dass man die Körperaxiome hier auch viel eleganter direkt aus den Eigenschaften von $\mathbb{Z}$ folgern kann [G, Satz 7.10]. Wir wollen uns hier damit begnügen, die neutralen und inversen Elemente anzugeben:

- Das additive neutrale Element ist g, wie man leicht aus der Tabelle abliest. Im Sinne der Notationen von Definition 3.5 ist also $0=g$. Wegen $g+g=u+u=g=0$ sind die additiven inversen Elemente $-g=g$ und $-u=u$. Dies stimmt natürlich auch mit der Interpretation als gerade und ungerade Zahlen überein, da das Negative von einer geraden bzw. ungeraden Zahl ebenfalls wieder gerade bzw. ungerade ist.
- Das multiplikative neutrale Element in $\mathbb{Z}_2\backslash\{0\}$ ist u – in der Tat ist es ja auch das einzige Element in $\mathbb{Z}_2\backslash\{0\}$. Gemäß der Notation von Definition 3.5 ist also $1=u$.

Beachte, dass in diesem Körper $\mathbb{Z}_2$ die Gleichung $1+1=u+u=g=0$ gilt. Die Körperaxiome lassen es also zu, dass man bei fortgesetzter Addition der 1 irgendwann wieder zur 0 zurück kommt. Wir werden in dieser Vorlesung nicht viel mit dem Körper $\mathbb{Z}_2$ zu tun haben – wir haben ihn hier nur als Beispiel dafür angegeben, dass die Körperaxiome noch weit davon entfernt sind, die rationalen oder reellen Zahlen eindeutig zu charakterisieren.

Anschaulich kann man die Körperaxiome so interpretieren, dass ein Körper eine Menge ist, auf der „die vier Grundrechenarten existieren und die erwarteten Eigenschaften haben“. Wir wollen nun noch ein paar weitere dieser erwarteten Eigenschaften zeigen, die bereits aus den Körperaxiomen folgen und die wir dann beim Rechnen z. B. in $\mathbb{R}$ natürlich ständig benutzen werden.

Bemerkung 3.7. Es seien K ein Körper und $x,y\in K$.

(a) Wenden wir Lemma 3.4 (c) und (d) auf die (kommutative) Addition und Multiplikation an, so sehen wir sofort, dass

$$-(x+y)=(-x)+(-y) \quad \text{und} \quad -(-x)=x$$

sowie für $x,y\neq 0$

$$(xy)^{-1}=x^{-1}\cdot y^{-1} \quad \text{und} \quad (x^{-1})^{-1}=x.$$

(b) Etwas versteckt in Definition 3.5 steht in Teil (b) u. a. die Aussage, dass die Multiplikation überhaupt eine Verknüpfung auf $K\backslash\{0\}$ ist, also dass für $x,y\in K\backslash\{0\}$ auch $xy\in K\backslash\{0\}$ gilt. Äquivalent dazu bedeutet das:

$$\text{Ist } xy=0, \text{ so gilt } x=0 \text{ oder } y=0.$$

Lemma 3.8 (Eigenschaften von Körpern)**.** *In jedem Körper K gilt für alle $x,y\in K$:*

(a) $0\cdot x=0$.

(b) $x\cdot(-y)=-(xy)$.

(c) *Für $x\neq 0$ ist $-(x^{-1})=(-x)^{-1}$.*

Beweis.

(a) Es gilt

$$\begin{aligned} 0\cdot x &= (0+0)\cdot x && \text{(0 ist additives neutrales Element)} \\ &= 0\cdot x+0\cdot x, && \text{(Distributivität)} \end{aligned}$$

woraus durch Addition des additiven Inversen von $0\cdot x$ auf beiden Seiten die gewünschte Gleichung $0=0\cdot x$ folgt.

(b) Es ist

$$\begin{aligned} x\cdot(-y)+xy &= x\cdot(-y+y) && \text{(Distributivität)} \\ &= x\cdot 0 && (-y \text{ ist additives Inverses zu } y) \\ &= 0 && \text{(nach (a))}, \end{aligned}$$

daher ist $x\cdot(-y)$ das additive Inverse zu xy, d. h. es gilt $x\cdot(-y) = -(xy)$.

(c) Doppeltes Anwenden von (b), einmal für den linken und einmal für den rechten Faktor, ergibt

$$(-(x^{-1}))\cdot(-x) = -(x^{-1}\cdot(-x)) = -(-(x^{-1}\cdot x)) = -(-1) \overset{3.7(a)}{=} 1.$$

Also ist $-(x^{-1})$ das multiplikative Inverse zu $-x$, d. h. es ist $-(x^{-1}) = (-x)^{-1}$. □

Notation 3.9. In einem Körper K verwendet man üblicherweise die folgenden Notationen, von denen euch die meisten sicher bekannt sein werden:

(a) Für $x,y\in K$ setzt man $x-y := x+(-y)$. Ist $y\neq 0$, so setzt man $\frac{x}{y} := x\cdot y^{-1}$.

(b) Für $x\in K$ und $n\in\mathbb{N}$ definiert man die n-te **Potenz** von x als

$$x^n := \underbrace{x\cdot\;\cdots\;\cdot x}_{n\text{-mal}},$$

wobei dieser Ausdruck für $n=0$ als $x^0 := 1$ zu verstehen ist. Insbesondere legen wir also auch $0^0 := 1$ fest. Beachte, dass aus dieser Definition (und der Kommutativität der Multiplikation) unmittelbar die Potenzrechenregeln

$$x^m\cdot x^n = x^{m+n} \quad\text{und}\quad (xy)^n = x^n\cdot y^n$$

für alle $x,y\in K$ folgen. Ist $x\neq 0$, so definiert man zusätzlich Potenzen mit negativen ganzzahligen Exponenten durch $x^{-n} := (x^{-1})^n$.

Beachte, dass auch in einem beliebigen Körper K die Exponenten einer Potenz stets *ganze Zahlen* sind und keine Elemente aus K. Eine Potenz x^y für $x,y\in K$ lässt sich im Allgemeinen nicht definieren (auch wenn dies für $K=\mathbb{R}$ in vielen Fällen möglich ist, siehe Definition 9.7).

(c) Manchmal möchte man mehrere Elemente $x_m, x_{m+1}, x_{m+2},\dots,x_n$ in einem Körper (oder allgemeiner in einer additiv geschriebenen abelschen Gruppe) aufsummieren, die durch eine ganzzahlige Laufvariable indiziert werden, die von einem $m\in\mathbb{Z}$ bis zu einem $n\in\mathbb{Z}$ (mit $n\geq m$) läuft. Man schreibt dies dann als

$$\sum_{i=m}^{n} x_i := x_m + x_{m+1} + x_{m+2} + \cdots + x_n$$

(also mit einem großen griechischen Sigma, das an das Wort „Summe" erinnern soll). So steht z. B.

$$\sum_{i=1}^{n} i^2 = 1^2+2^2+3^2+\cdots+n^2 \qquad (*)$$

für die Summe aller Quadratzahlen bis n^2. Natürlich ist der Name der Laufvariablen dabei egal, und der Ausdruck $(*)$ hängt nicht von einem i ab (wie man auf der rechten Seite ja auch sieht). Außerdem kann man die Laufvariable verschieben, ohne den eigentlichen Ausdruck zu ändern: Setzt man z. B. $i=j+1$, also $j=i-1$, in der obigen Summe $(*)$, so läuft j dort von 0 bis $n-1$, wenn i von 1 bis n läuft, und wir können dieselbe Summe auch schreiben als

$$\sum_{j=0}^{n-1} (j+1)^2 = 1^2+2^2+3^2+\cdots+n^2.$$

Natürlich kann man diesen Ausdruck nun auch wieder genauso gut mit dem Buchstaben i statt j als $\sum_{i=0}^{n-1}(i+1)^2$ schreiben, oder den Index um mehr als 1 in die eine oder andere Richtung verschieben. Also:

Der Wert einer Summe ändert sich nicht, wenn man zur Laufvariablen im zu summierenden Ausdruck eine Konstante addiert, und dafür von der Ober- und Untergrenze der Summe diese Konstante abzieht.

Wir sagen in diesem Fall, dass die neue Darstellung der Summe durch eine **Indexverschiebung** (im Beispiel oben $i \mapsto i+1$) aus der alten hervorgeht.

Analog schreibt man

$$\prod_{i=m}^{n} x_i := x_m \cdot x_{m+1} \cdot x_{m+2} \cdot \;\cdots\; \cdot x_n$$

(mit einem großen griechischen Pi für das Produkt), wenn man die Körperelemente multiplizieren statt addieren möchte. Ist schließlich die Obergrenze einer Summe oder eines Produkts kleiner als die Untergrenze (man spricht dann von der **leeren Summe** bzw. dem **leeren Produkt**), so definiert man dies als

$$\sum_{i=m}^{n} x_i := 0 \quad \text{und} \quad \prod_{i=m}^{n} x_i := 1 \quad \text{für } n < m,$$

04 also als das additive bzw. multiplikative neutrale Element.

(d) Ist n eine natürliche Zahl, so fasst man diese oft auch als das Element

$$\sum_{i=1}^{n} 1 = \underbrace{1 + \cdots + 1}_{n\text{-mal}}$$

von K auf. Im Fall $K = \mathbb{R}$ ist dies dann einfach die natürliche Zahl $n \in \mathbb{N} \subset \mathbb{R}$ und liefert somit keine neue Notation, aber z. B. in $K = \mathbb{Z}_2$ aus Beispiel 3.6 (b) ist $2 = 1 + 1 = 0$.

Aufgabe 3.10. Zeige, dass in jedem Körper K die üblichen Rechenregeln

$$\frac{x}{y} + \frac{z}{w} = \frac{xw + yz}{yw} \quad \text{und} \quad \frac{x}{y} \cdot \frac{z}{w} = \frac{xz}{yw}$$

für Brüche gelten, wobei $x, y, z, w \in K$ mit $y, w \neq 0$.

Aufgabe 3.11. Es sei $a \in \mathbb{R}$ fest gegeben. Wir definieren auf $\mathbb{R}^2$ eine „Addition“ und „Multiplikation“ durch

$$(x_1, x_2) + (y_1, y_2) := (x_1 + y_1, x_2 + y_2) \quad \text{und} \quad (x_1, x_2) \cdot (y_1, y_2) := (x_1 y_1 + a x_2 y_2, x_1 y_2 + x_2 y_1).$$

Man prüft leicht durch explizite Rechnung nach, dass $\mathbb{R}^2$ mit dieser Addition eine kommutative Gruppe mit neutralem Element $(0,0)$ ist, dass auch die Multiplikation kommutativ ist, und dass diese beiden Operationen das Distributivgesetz erfüllen – ihr solltet euch kurz überlegen, warum das so ist, braucht das aber nicht aufzuschreiben. Man zeige stattdessen:

(a) Die Multiplikation ist assoziativ und besitzt ein neutrales Element.

(b) Im Fall $a = -1$ ist $(\mathbb{R}^2, +, \cdot)$ ein Körper, im Fall $a = 1$ jedoch nicht.

(Für $a = -1$ ist $(\mathbb{R}^2, +, \cdot)$ der sogenannte Körper der komplexen Zahlen, den wir in Kapitel 6 noch genau untersuchen werden.)

Aufgabe 3.12. Zu einem Körper K und einer Menge M mit $|M| \geq 2$ sei

$$V = \{f : f \text{ ist eine Abbildung von } M \text{ nach } K\}$$

die Menge aller reellwertigen Funktionen auf M. Für $f, g \in V$ definieren wir die Addition $f + g$ und Multiplikation $f \cdot g$ dieser Funktionen punktweise durch

$$f + g \colon M \to K,\ x \mapsto f(x) + g(x) \quad \text{und} \quad f \cdot g \colon M \to K,\ x \mapsto f(x) \cdot g(x).$$

(a) Zeige, dass V mit dieser Addition eine abelsche Gruppe ist.

(b) Ist V mit dieser Addition und Multiplikation ein Körper?

3.B Vollständige Induktion

Häufig möchte man in der Mathematik Aussagen beweisen, die von einer natürlichen Zahl abhängen – z. B. bei Formeln, die Summen oder Produkte wie in Notation 3.9 mit variablen Unter- oder Obergrenzen beinhalten. Die einfachste und bekannteste solcher Aussagen ist vermutlich die folgende Formel für die Summe aller natürlichen Zahlen bis zu einer gegebenen Obergrenze.

Satz 3.13 (Summenformel von Gauß). *Für alle $n \in \mathbb{N}$ gilt*

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

Beispiel 3.14. Für $n = 5$ ist z. B.

$$\sum_{k=1}^{5} k = 1+2+3+4+5 = 15 = \frac{5 \cdot 6}{2}.$$

Um derartige Aussagen zu beweisen, ist oft das Beweisverfahren der **(vollständigen) Induktion** nützlich, das wir jetzt einführen wollen.

Angenommen, wir wollen (wie z. B. in Satz 3.13) eine Aussage $A(n)$ für alle $n \in \mathbb{N}$ beweisen. Dann können wir dies tun, indem wir die folgenden beiden Dinge zeigen:

(a) (**Induktionsanfang**) Die Aussage $A(0)$ ist wahr.

(b) (**Induktionsschritt**) Für alle $n \in \mathbb{N}$ gilt $A(n) \Rightarrow A(n+1)$, d. h. *wenn* die Aussage $A(n)$ für ein gegebenes $n \in \mathbb{N}$ gilt (die „Induktionsannahme“ bzw. „Induktionsvoraussetzung“), *dann* gilt auch die Aussage $A(n+1)$ (der „Induktionsschluss“).

Haben wir diese beiden Dinge gezeigt, so folgt daraus nämlich die Gültigkeit von $A(n)$ für alle $n \in \mathbb{N}$: Die Aussage $A(0)$ haben wir mit dem Induktionsanfang gezeigt, und durch fortgesetztes Anwenden des Induktionsschritts $A(n) \Rightarrow A(n+1)$ für $n = 0, 1, 2, \dots$ erhalten wir dann auch

$$A(0) \Rightarrow A(1) \Rightarrow A(2) \Rightarrow A(3) \Rightarrow \cdots,$$

also die Gültigkeit von $A(n)$ für alle $n \in \mathbb{N} = \{0, 1, 2, 3, \dots\}$.

Derartige Induktionsbeweise sind immer dann sinnvoll, wenn die Aussagen $A(n)$ und $A(n+1)$ „ähnlich genug“ sind, so dass es beim Beweis von $A(n+1)$ hilft, die Gültigkeit von $A(n)$ voraussetzen zu dürfen.

Mit diesem Verfahren können wir nun die Summenformel aus Satz 3.13 beweisen:

Beweis von Satz 3.13. Wir zeigen die Formel mit Induktion über n.

Induktionsanfang ($n = 0$): Für $n = 0$ stimmen die beiden Seiten der zu zeigenden Gleichung überein, denn es ist

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

Induktionsschritt ($n \to n+1$): Als Induktionsvoraussetzung nehmen wir an, dass die zu beweisende Formel für ein gegebenes $n \in \mathbb{N}$ richtig ist, d. h. dass

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$$

gilt. (Beachte, dass wir diese Gleichung nicht für alle $n \in \mathbb{N}$ voraussetzen – dies wäre ja schon die gesamte zu zeigende Aussage!) Wir müssen zeigen, dass die entsprechende Gleichung dann auch für $n+1$ gilt, also dass

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

Dies ergibt sich nun leicht aus der folgenden Rechnung:

$$\begin{aligned}
\sum_{k=1}^{n+1} k &= \sum_{k=1}^{n} k + (n+1) && \text{(Abspalten des letzten Summanden für } k=n+1\text{)} \\
&= \frac{n(n+1)}{2} + (n+1) && \text{(nach Induktionsvoraussetzung)} \\
&= \frac{n^2+n+2n+2}{2} \\
&= \frac{(n+1)(n+2)}{2}.
\end{aligned}$$

Damit ist der Satz mit vollständiger Induktion bewiesen. □

Bemerkung 3.15. Offensichtlich erlaubt das Beweisverfahren der vollständigen Induktion die folgenden Abwandlungen:

(a) Im Induktionsschritt kann man, wenn es hilfreich ist, beim Beweis der Aussage $A(n+1)$ nicht nur die direkt vorangegangene Aussage $A(n)$, sondern *alle bereits gezeigten Aussagen* $A(0), A(1), \dots, A(n)$ voraussetzen.

(b) Möchte man die Aussage $A(n)$ nicht für alle $n \in \mathbb{N}$, sondern für alle $n \in \mathbb{Z}$ ab einem gewissen Startwert $n_0 \in \mathbb{Z}$ zeigen, so kann man als Induktionsanfang die Aussage $A(n_0)$ zeigen, und im Induktionsschritt dann die Folgerung $A(n) \Rightarrow A(n+1)$ für alle $n \geq n_0$.

Aufgabe 3.16. Zeige für alle $n \in \mathbb{N}$:

$$\text{(a)} \quad \sum_{k=1}^{n} \frac{1}{k(k+1)} = \frac{n}{n+1}, \qquad \text{(b)} \quad \prod_{k=1}^{n} \left(1 + \frac{1}{n+k}\right) = 2 - \frac{1}{n+1}.$$

Aufgabe 3.17. Zeige mit vollständiger Induktion: Ist $a \in \mathbb{R}\backslash\{0\}$ mit $a + \frac{1}{a} \in \mathbb{Z}$, so gilt für alle $n \in \mathbb{N}$ auch $a^n + \frac{1}{a^n} \in \mathbb{Z}$.

3.C Polynomfunktionen

Als erste Anwendung der Körpereigenschaften wollen wir zum Abschluss dieses Kapitels die euch sicher schon aus der Schule bekannten Polynomfunktionen behandeln – also die Funktionen, die sich aus den grundlegenden Körperoperationen Addition und Multiplikation bilden lassen.

Definition 3.18 (Polynomfunktionen und Nullstellen)**.** Es seien D eine Teilmenge eines Körpers K und $f\colon D \to K$ eine Funktion.

(a) Ist f von der Form

$$f(x) = \sum_{k=0}^{n} a_k x^k = a_n x^n + \cdots + a_1 x + a_0 \quad \text{für alle } x \in D$$

mit gewissen $a_0, \dots, a_n \in K$, so sagt man, dass f eine **Polynomfunktion** mit **Koeffizienten** $a_0, \dots, a_n$ ist. Ist n dabei so gewählt, dass der erste Koeffizient a_n ungleich Null ist, so heißt f eine Polynomfunktion vom **Grad** n und mit **Leitkoeffizient** a_n. Ist der Leitkoeffizient 1, so heißt f eine **normierte** Polynomfunktion.

Sind in der obigen Darstellung alle Koeffizienten $a_0, \dots, a_n$ gleich 0 (und ist f damit die Nullfunktion), so nennen wir f formal eine Polynomfunktion vom Grad $-\infty$. In diesem Fall hat f keinen Leitkoeffizienten.

(b) Ist $x_0 \in D$ mit $f(x_0) = 0$, so nennt man x_0 eine **Nullstelle** von f.

Das Besondere an Nullstellen von Polynomfunktionen ist, dass man sie wie im folgenden Satz als Linearfaktoren abspalten kann.

Satz 3.19 (Abspalten von Nullstellen in Polynomfunktionen)**.** *Es seien K ein Körper, $D \subset K$ und $f\colon D \to K$ eine Polynomfunktion vom Grad $n \in \mathbb{N}$.*

(a) *Ist $x_0 \in D$ eine Nullstelle von f, so gibt es eine Polynomfunktion $g\colon D \to K$ vom Grad $n-1$ mit $f(x) = (x-x_0)\,g(x)$ für alle $x \in D$ (d. h. man kann „den Linearfaktor $x - x_0$ abspalten").*

(b) *Die Funktion f hat höchstens n Nullstellen.*

Beweis. Wir zeigen die beiden Aussagen mit Induktion über n. Der Beweis von (a) ist dabei konstruktiv, d. h. er gibt auch ein Verfahren an, wie g berechnet werden kann (siehe Beispiel 3.20).

Der Induktionsanfang für $n = 0$ ist trivial, denn f ist dann eine Konstante ungleich 0 und hat somit keine Nullstellen. Für den Induktionsschluss nehmen wir an, dass die Aussagen des Satzes bis zu einem gegebenen n gelten, und betrachten $f\colon D \to K$, $x \mapsto a_{n+1}x^{n+1} + \cdots + a_1x + a_0$ vom Grad $n+1$, also mit $a_{n+1} \neq 0$.

(a) Wir definieren eine Polynomfunktion $\tilde{f}\colon D \to K$ durch

$$\begin{aligned}\tilde{f}(x) &:= f(x) - a_{n+1}x^n(x-x_0)\\ &= a_{n+1}x_0x^n + a_nx^n + a_{n-1}x^{n-1} + \cdots + a_1x + a_0\end{aligned}$$

für alle x. Ist $\tilde{f}$ die Nullfunktion, so sind wir fertig, da dann ja $f(x) = a_{n+1}x^n(x-x_0)$ für alle $x \in D$ gilt. Andernfalls ist $\tilde{f}$ nach Konstruktion eine Polynomfunktion von einem Grad kleiner als $n+1$ (der x^{n+1}-Term hebt sich ja gerade heraus), die immer noch die Nullstelle x_0 hat. Nach Induktionsvoraussetzung gibt es dann also eine Polynomfunktion $\tilde{g}\colon D \to K$ vom Grad kleiner als n mit $\tilde{f}(x) = (x-x_0)\,\tilde{g}(x)$ für alle $x \in D$, und somit ist

$$\begin{aligned}f(x) &= a_{n+1}x^n(x-x_0) + \tilde{f}(x)\\ &= (x-x_0)\cdot\underbrace{(a_{n+1}x^n + \tilde{g}(x))}_{=:g(x)}\end{aligned}$$

für alle $x \in D$, wobei g offensichtlich vom Grad n ist.

(b) Hat f keine Nullstelle, so sind wir fertig. Andernfalls wählen wir eine Nullstelle x_0 von f und schreiben $f(x) = (x-x_0)\,g(x)$ für alle $x \in D$ wie in (a) mit einer Polynomfunktion g vom Grad n. Nach Induktionsvoraussetzung hat g höchstens n Nullstellen, und nach Bemerkung 3.7 (b) sind die Nullstellen von f genau x_0 zusammen mit den Nullstellen von g. Also hat f höchstens $n+1$ Nullstellen. Damit ist die Behauptung mit Induktion bewiesen. □

Beispiel 3.20 (Polynomdivision)**.** Das Verfahren aus dem Beweis von Satz 3.19 (a) wird als *Polynomdivision* [G, Satz 10.19] bezeichnet: Man subtrahiert fortlaufend geeignete Vielfache von $x - x_0$ von f, so dass sich der jeweils höchste Term von f weghebt, und sammelt die dabei verwendeten Faktoren in g. Das folgende Schema, das genauso aussieht wie eine normale schriftliche Division, verdeutlicht dieses Verfahren am Beispiel der Funktion $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x^2 + 3x - 4$ mit Nullstelle $x_0 = 1$, die wir als $f(x) = (x-1)\,g(x)$ für alle $x \in \mathbb{R}$ schreiben wollen. Das Ergebnis ist in diesem Fall $g\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto x + 4$.

$$\begin{array}{lrcl}
f \longrightarrow & (x^2+3x-4) & : (x-1) = x+4 & \longleftarrow g \\
& -\ (x^2-\ \ x) & \cdot x & \\
\hline
\tilde{f} \longrightarrow & 4x-4 & & \\
& -\quad (4x-4) & \cdot 4 & \\
\hline
& 0 & &
\end{array}$$

Bemerkung 3.21. Satz 3.19 liefert uns zwar die neue Funktion nach dem Abspalten des Linearfaktors, er sagt uns hingegen nicht, wie wir überhaupt erst einmal eine Nullstelle von f finden können, oder ob es überhaupt Nullstellen gibt (die reelle Polynomfunktion $f(x) = x^2 + 1$ hat ja z. B. keine Nullstellen). In der Tat gibt es im Allgemeinen kein Verfahren, wie man Nullstellen von Polynomfunktionen exakt berechnen kann! Genauer gesagt gilt:

- Für Polynomfunktionen vom Grad höchstens 4 gibt es explizite Verfahren zur exakten Bestimmung der Nullstellen (für Grad 1 ist das klar, für Grad 2 gibt es die bekannte „*p*-*q*-Formel“ bzw. die quadratische Ergänzung, und für Grad 3 bzw. 4 sind die Formeln so lang, dass man im Allgemeinen nicht mehr mit ihnen arbeiten möchte).
- Für Polynomfunktionen vom Grad größer als 4 kann man beweisen(!), dass es keine derartigen Verfahren zur exakten Bestimmung der Nullstellen geben kann (das beweist man z. B. in der Vorlesung „Einführung in die Algebra“, die ihr im nächsten Studienjahr hören könnt). Aber:
- Für reelle Polynomfunktionen beliebigen Grades gibt es zumindest numerische Verfahren, die die Nullstellen (mit beliebiger Genauigkeit) näherungsweise bestimmen können.

Zum Schluss wollen wir nun noch zwei wichtige Konzepte für Polynomfunktionen untersuchen, die ihr beide im reellen Fall vielleicht schon aus der Schule kennt: den sogenannten Koeffizientenvergleich (also dass eine Polynomfunktion eindeutig ihre Koeffizienten bestimmt) und die Vielfachheit von Nullstellen. Es stellt sich jedoch heraus, dass man hierfür im allgemeinen Fall die Voraussetzung benötigt, dass die Definitionsmenge der betrachteten Funktionen unendlich viele Elemente besitzt.

Lemma 3.22 (**Koeffizientenvergleich**)**.** *Es seien K ein Körper, $D \subset K$ mit $|D| = \infty$, und $f\colon D \to K$ eine Polynomfunktion mit zwei Darstellungen*

$$f(x) = a_n x^n + \cdots + a_1 x + a_0 = b_n x^n + \cdots + b_1 x + b_0 \quad \textit{für alle } x \in D$$

für gewisse $a_0, \dots, a_n, b_0, \dots, b_n \in K$. (Beachte, dass wir dabei in beiden Darstellungen den gleichen höchsten Exponenten n wählen können, da wir nicht $a_n \neq 0$ und $b_n \neq 0$ vorausgesetzt haben.)

Dann gilt bereits $a_i = b_i$ für alle $i = 0, \dots, n$. Es ist also nicht möglich, „eine Polynomfunktion auf zwei verschiedene Arten hinzuschreiben“.

Beweis. Nach Voraussetzung ist die Polynomfunktion

$$D \to K,\ x \mapsto (a_n - b_n)x^n + \cdots + (a_1 - b_1)x + (a_0 - b_0) = f(x) - f(x) = 0$$

die Nullfunktion auf D. Da sie damit wegen $|D| = \infty$ unendlich viele Nullstellen besitzt, muss sie nach Satz 3.19 (b) vom Grad $-\infty$ sein. Also sind alle Koeffizienten dieser Polynomfunktion gleich 0, d. h. es ist $a_i = b_i$ für alle $i = 0, \dots, n$. □

Bemerkung und Notation 3.23 (Polynome)**.** Die Voraussetzung $|D| = \infty$ in Lemma 3.22 ist wirklich notwendig: So sind für $D = K = \mathbb{Z}_2$ wie in Beispiel 3.6 (b) z. B. $x \mapsto x$ und $x \mapsto x^2$ dieselbe Funktion, da sie beide 0 auf 0 und 1 auf 1 abbilden und in $\mathbb{Z}_2$ keine weiteren Elemente existieren.

In der Literatur bezeichnet man einen formalen Ausdruck der Form $a_n x^n + \cdots + a_1 x + a_0$ mit $a_0, \dots, a_n \in K$ als ein **Polynom** über K [G, Kapitel 9]. Jedes solche Polynom bestimmt natürlich eine Polynomfunktion von jeder Teilmenge D von K nach K, allerdings können verschiedene Polynome wie im eben angegebenen Beispiel durchaus dieselbe Polynomfunktion definieren: Über $\mathbb{Z}_2$ sind x und x^2 verschiedene Polynome, sie bestimmen aber dieselbe Polynomfunktion.

Mit dieser Notation ist die Aussage von Lemma 3.22 also, dass Polynome und Polynomfunktionen im Fall von unendlichen Definitionsmengen dasselbe sind. Da wir Polynomfunktionen im Folgenden in der Regel nur in diesem Fall unendlicher Definitionsmengen benötigen, werden wir die Begriffe Polynom und Polynomfunktion oft synonym verwenden. Wegen der Eindeutigkeit der Koeffizienten sind dann auch der Grad (und der Leitkoeffizient) einer Polynomfunktion f wie in Definition 3.18 (a) eindeutig bestimmt. Wir können daher eine Bezeichnung dafür einführen:

Definition 3.24 (Grad eines Polynoms)**.** Wir bezeichnen den **Grad** einer Polynomfunktion f (mit unendlicher Definitionsmenge) mit $\deg f \in \mathbb{N} \cup \{-\infty\}$ (vom englischen Wort „degree“).

Satz und Definition 3.25 (Vielfachheit von Nullstellen)**.** *Es seien K ein Körper, $D \subset K$ mit $|D| = \infty$ und $f\colon D \to K$ eine Polynomfunktion, die nicht die Nullfunktion ist. Dann lässt sich f (bis auf die Reihenfolge der Faktoren) eindeutig als*

$$f(x) = g(x) \cdot (x - x_1)^{a_1} \cdot \cdots \cdot (x - x_k)^{a_k} \quad \textit{für alle } x \in D$$

*schreiben, wobei $x_1, \dots, x_k \in D$ die verschiedenen Nullstellen von f sind, $a_1, \dots, a_k \in \mathbb{N}_{>0}$ gilt, und g eine Polynomfunktion ohne Nullstellen in D ist. In dieser Darstellung nennt man a_i für $i = 1, \dots, k$ die **Vielfachheit** der Nullstelle x_i (in der Literatur sind auch die Bezeichnungen **Ordnung** und **Multiplizität** der Nullstelle üblich).*

Beweis. Die Existenz einer solchen Darstellung ergibt sich sofort durch fortgesetztes Abspalten von Nullstellen gemäß Satz 3.19 (a). Wir zeigen nun die Eindeutigkeit mit Induktion über den Grad der Polynomfunktion. Dabei ist der Induktionsanfang für Grad 0 trivial, denn dann hat f keine Nullstellen, und es ist zwangsläufig $k = 0$ und $g = f$.

Für den Induktionsschritt bemerken wir zuerst, dass $x_1, \dots, x_k$ natürlich in jedem Fall als die Nullstellen von f eindeutig bestimmt sind. Wir nehmen also an, dass wir zwei Darstellungen

$$f(x) = g(x) \cdot (x - x_1)^{a_1} \cdot \dots \cdot (x - x_k)^{a_k} = h(x) \cdot (x - x_1)^{b_1} \cdot \dots \cdot (x - x_k)^{b_k}$$

wie in der Behauptung des Satzes haben. Im nullstellenfreien Fall $k = 0$ sind wir natürlich bereits fertig. Andernfalls liefert Division durch $x - x_1$ für alle $x \in D \backslash \{x_1\}$ (wir müssen x_1 hier herausnehmen, da wir sonst durch 0 teilen würden!)

$$\begin{aligned} & g(x) \cdot (x - x_1)^{a_1 - 1} \cdot (x - x_2)^{a_2} \cdot \dots \cdot (x - x_k)^{a_k} \\ = {} & h(x) \cdot (x - x_1)^{b_1 - 1} \cdot (x - x_2)^{b_2} \cdot \dots \cdot (x - x_k)^{b_k}. \end{aligned} \tag{$*$}$$

Wir haben also wieder zwei Darstellungen einer Polynomfunktion auf der immer noch unendlichen Menge $D \backslash \{x_1\}$. Da der Grad dieser Polynomfunktion nun um 1 kleiner ist als der von f, müssen diese Darstellungen aber nach der Induktionsvoraussetzung bereits übereinstimmen. Also gilt $g = h$, $a_1 - 1 = b_1 - 1$, $a_2 = b_2$, …, $a_k = b_k$, und damit stimmen auch die beiden ursprünglichen Darstellungen von f überein. □

Aufgabe 3.26. Bestimme die Nullstellen des reellen Polynoms $x^4 + 3x^3 - 4x$ und ihre Vielfachheiten.

Aufgabe 3.27. Es sei f ein reelles Polynom mit $f(x) = x^3$ für alle $x \in \{1, 2, 3, \dots, 8\}$. Welchen Grad kann f haben?

Aufgabe 3.28. Es sei f das reelle Polynom mit $f(x) = (x^2 - x + 1)^{2023}$ für alle $x \in \mathbb{R}$.

(a) Bestimme die Summe aller Koeffizienten von f.

(b) Bestimme die Summe aller Koeffizienten von geraden Potenzen von x in f.

Aufgabe 3.29.

(a) Bestimme alle reellen Polynome f mit $x f(x+1) = (x-1) f(x)$ für alle $x \in \mathbb{R}$.

(b) Bestimme alle reellen Polynome f mit $x f(x-1) = (x-1) f(x)$ für alle $x \in \mathbb{R}$. 05

Grundlagen der Mathematik 1: Analysis

4. Weitere Eigenschaften der reellen Zahlen

Wir haben uns nun das elementare Handwerkszeug für diese Vorlesung erarbeitet und beginnen jetzt mit dem Studium der eindimensionalen Analysis. Wer sich auch (bzw. zurzeit nur) für die lineare Algebra interessiert, kann ab diesem Zeitpunkt auch zusätzlich (bzw. ausschließlich) den Teil „Grundlagen der Mathematik 1: Lineare Algebra“ in den Kapiteln **??** bis **??** durcharbeiten.

Hier in diesem Kapitel wollen wir zunächst nach den gerade behandelten elementaren Eigenschaften der reellen Zahlen noch ein paar weitere untersuchen, die vor allem in der Analysis nützlich sind. Unter anderem wird sich daraus am Ende dieses Kapitels auch eine vollständige Charakterisierung der reellen Zahlen ergeben.

4.A Potenzen in Körpern

Wir beginnen mit zwei weiteren oft vorkommenden Formeln zu Potenzen, die sich allein aus den Eigenschaften aus Abschnitt 3.A herleiten lassen und somit nicht nur in den reellen Zahlen, sondern sogar in beliebigen Körpern gelten. Mit der ersten – der sogenannten (endlichen) geometrischen Reihe – können wir den Wert einer Summe fortlaufender Potenzen eines Körperelements explizit berechnen.

Satz 4.1 (**Endliche geometrische Reihe**). *Es seien K ein Körper, $q \in K\backslash\{1\}$ und $n \in \mathbb{N}$. Dann gilt*

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

Beweis. Die Aussage ließe sich leicht mit Induktion über n zeigen. Der folgende sehr einfache alternative Beweis hilft jedoch auch dabei, sich die Formel zu merken: Multiplizieren wir die gesuchte Summe $\sum_{k=0}^{n} q^k = 1+q+q^2+\cdots+q^n$ mit $1-q$, so heben sich fast alle Terme weg und wir erhalten sofort das gewünschte Ergebnis: Es ist

$$\begin{aligned}(1+q+q^2+\cdots+q^n)\cdot(1-q) &= 1+q+q^2+\cdots+q^n \\ &\qquad\; -q-q^2-\cdots-q^n-q^{n+1} \\ &= 1-q^{n+1},\end{aligned}$$

und damit für $q \neq 1$ wie behauptet

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

□

Beispiel 4.2. In $\mathbb{R}$ ist z. B.

$$1+2+4+8+16 = \sum_{k=0}^{4} 2^k \overset{4.1}{=} \frac{1-2^5}{1-2} = \frac{-31}{-1} = 31.$$

Die zweite Formel, die wir hier behandeln wollen, ist die sogenannte binomische Formel, die eine Verallgemeinerung der aus der Schule bekannten Formel $(x+y)^2 = x^2+2xy+y^2$ auf höhere Exponenten darstellt. Dazu benötigen wir zunächst die folgende Definition.

Definition 4.3 (Fakultät und Binomialkoeffizienten)**.**

(a) Für $n \in \mathbb{N}$ setzen wir

$$n! := \prod_{i=1}^{n} i = 1\cdot 2\cdot \cdots \cdot n \quad \in \mathbb{N} \qquad \text{(gesprochen „}n\text{-}\textbf{Fakultät}\text{“)},$$

wobei $0!$ gemäß Notation 3.9 (c) als 1 zu verstehen ist.

(b) Für $k, n \in \mathbb{N}$ mit $k \le n$ definiert man ferner die **Binomialkoeffizienten**

$$\binom{n}{k} := \frac{n!}{k!\,(n-k)!} \qquad \text{(gesprochen „}n\text{ über }k\text{“)},$$

die so genannt werden, weil sie in der binomischen Formel in Satz 4.7 auftreten. Sie sind aufgrund der Definition zunächst positive rationale Zahlen; wir werden aber in Bemerkung 4.6 sehen, dass sie sogar natürliche Zahlen sind.

Bemerkung 4.4.

(a) Offensichtlich ist $\binom{n}{0} = \binom{n}{n} = 1$ und $\binom{n}{k} = \binom{n}{n-k}$ für alle $k, n \in \mathbb{N}$ mit $k \le n$.

(b) Man kann im definierenden Ausdruck für $\binom{n}{k}$ die Faktoren von 1 bis $n-k$ kürzen und erhält damit die alternative Darstellung der Binomialkoeffizienten

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

d. h. $\binom{n}{k}$ ist ein Bruch mit k Zahlen im Zähler und k Zahlen im Nenner, wobei man „im Zähler von n nach unten und im Nenner von 1 nach oben zählt“. So ist z. B. $\binom{n}{1} = \frac{n}{1} = n$ und $\binom{n}{2} = \frac{n(n-1)}{1\cdot 2} = \frac{n(n-1)}{2}$.

Die wichtigste Identität zwischen den Binomialkoeffizienten ist die folgende:

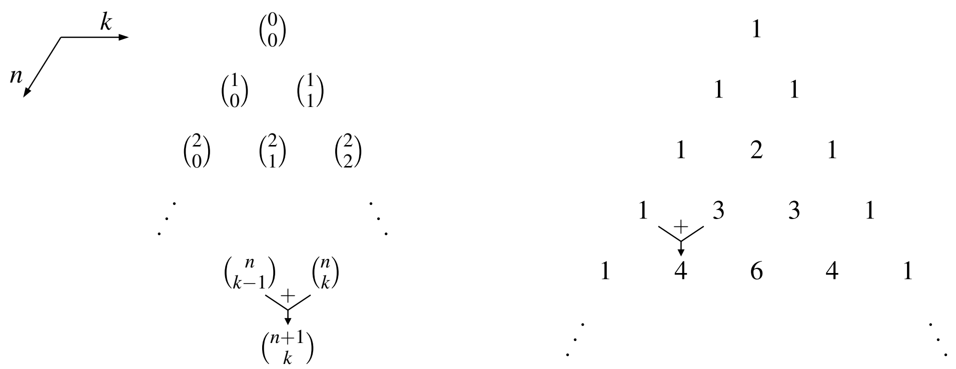
Lemma 4.5. *Für alle $n, k \in \mathbb{N}$ mit $1 \le k \le n$ gilt $\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$.*

Beweis. Dies ergibt sich durch einfaches Nachrechnen mit der Darstellung aus Bemerkung 4.4 (b):

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

□

Bemerkung 4.6 (Pascalsches Dreieck)**.** Man kann die Binomialkoeffizienten $\binom{n}{k}$ wie folgt in einem dreieckigen Schema, dem sogenannten **Pascalschen Dreieck**, anordnen.



Nach Bemerkung 4.4 (a) stehen auf den Schenkeln dieses Dreiecks nur Einsen, und nach Lemma 4.5 ergibt sich jede andere Zahl in diesem Diagramm als die Summe der beiden darüber stehenden. Insbesondere folgt daraus, dass alle Binomialkoeffizienten *natürliche* Zahlen sind – was aus der Definition aufgrund des Bruches ja nicht offensichtlich ist. Wir können sie damit für jeden Körper K gemäß Notation 3.9 (d) als Elemente von K auffassen (was wir gleich in Satz 4.7 auch tun werden).

Mit dieser Vorarbeit können wir nun die sehr wichtige binomische Formel beweisen. Ihr Name kommt übrigens von der lateinischen Vorsilbe „bi“ für „zwei“: Ein Binom ist eine Summe, die aus zwei Termen besteht, und die binomische Formel berechnet die Potenzen eines solchen Binoms. Beachte, dass die Binomialkoeffizienten in dieser Formel gemäß Notation 3.9 (d) als Elemente des Körpers K aufgefasst werden.

Satz 4.7 (Binomische Formel). *Es seien K ein Körper, $x,y \in K$ und $n \in \mathbb{N}$. Dann gilt*

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

Beweis. Wir beweisen die Formel mit Induktion über n. Für $n=0$ sind beide Seiten der Gleichung 1; die Aussage ist in diesem Fall also richtig. Für den Induktionsschritt nehmen wir nun an, dass die Gleichung für ein gegebenes $n \in \mathbb{N}$ richtig ist, und folgern daraus zunächst

$$\begin{aligned}
(x+y)^{n+1} &= (x+y)\cdot(x+y)^n \\
&= (x+y)\cdot \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} && \text{(nach Induktionsvoraussetzung)} \\
&= \sum_{k=0}^{n} \binom{n}{k} x^{k+1} y^{n-k} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k+1} && \text{(durch Ausmultiplizieren)} \\
&= \sum_{k=1}^{n+1} \binom{n}{k-1} x^k y^{n-k+1} + \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k+1} && \text{(Indexverschiebung } k \mapsto k-1 \text{ in der} \\
& && \text{ersten Summe, siehe Notation 3.9 (c)).}
\end{aligned}$$

Lösen wir hier nun aus der ersten Summe den Term für $k=n+1$ und aus der zweiten den für $k=0$ heraus, so können wir diesen Ausdruck auch schreiben als

$$\begin{aligned}
(x+y)^{n+1} &= \sum_{k=1}^{n} \left[\binom{n}{k-1} + \binom{n}{k}\right] \cdot x^k y^{n+1-k} + \binom{n}{n} x^{n+1} y^0 + \binom{n}{0} x^0 y^{n+1} \\
&= \sum_{k=1}^{n} \binom{n+1}{k} \cdot x^k y^{n+1-k} + x^{n+1} + y^{n+1} && \text{(nach Lemma 4.5).}
\end{aligned}$$

Die letzten beiden Summanden x^{n+1} und y^{n+1} sind hier aber genau diejenigen, die sich in der vorderen Summe ergeben, wenn man $k=n+1$ bzw. $k=0$ setzt. Also können wir die Summe über k gleich über alle Werte von 0 bis $n+1$ laufen lassen und erhalten

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

also genau die zu zeigende Gleichung für die Potenz $n+1$. Die binomische Formel ist damit durch Induktion bewiesen. □

Beispiel 4.8. Für $n=2$ ergibt Satz 4.7 wie erwartet die bekannte Formel

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

Bemerkung 4.9 (Kombinatorische Deutung der Binomialkoeffizienten)**.** Man kann sich die binomische Formel natürlich auch so entstanden denken, dass man den Ausdruck

$$(x+y)^n = \underbrace{(x+y)\cdot \cdots \cdot (x+y)}_{n\text{-mal}}$$

nach dem Distributivgesetz ausmultipliziert. Im Fall $n=3$ erhalten wir z. B. zunächst ohne Verwendung der Kommutativität der Multiplikation

$$\begin{aligned}
(x+y)^3 &= (x+y)\cdot(x+y)\cdot(x+y) \\
&= (x+y)\cdot(xx+xy+yx+yy) \\
&= xxx+xxy+xyx+xyy+yxx+yxy+yyx+yyy. && (*)
\end{aligned}$$

Insgesamt bekommen wir also eine Summe aus Produkten mit jeweils n Faktoren x oder y. Jede Möglichkeit, alle diese Faktoren separat als x oder y zu wählen, kommt dabei genau einmal vor. Fassen wir nun mit der Kommutativität der Multiplikation gleiche Terme zusammen, so erhalten wir den Term $x^k y^{n-k}$ also genau so oft, wie wir Möglichkeiten haben, aus den n Faktoren die k auszuwählen, die gleich x sein sollen. In $(*)$ oben bekommen wir z. B. den Term xy^2 dreimal, nämlich aus xyy, yxy und yyx. Nach der binomischen Formel ist der Vorfaktor von $x^k y^{n-k}$ in $(x+y)^n$ aber gerade $\binom{n}{k}$. Daher ist dieser Binomialkoeffizient genau die Anzahl der Möglichkeiten, aus n Objekten (hier: Faktoren) k auszuwählen (hier: diejenigen, bei denen wir x gewählt haben). Man kann sich die binomische Formel also als algebraische Formulierung dieser kombinatorischen Aussage vorstellen.

Aufgabe 4.10.

(a) Beweise für alle $k,n \in \mathbb{N}$ mit $k \le n$ die Gleichung

$$\sum_{m=k}^{n} \binom{m}{k} = \binom{n+1}{k+1}.$$

(b) Zeige mit Induktion über n, dass die Gleichung

$$x_1 + \cdots + x_n = d$$

für gegebenes $n \in \mathbb{N}_{>0}$ und $d \in \mathbb{N}$ genau $\binom{n+d-1}{n-1}$ Lösungen $(x_1,\dots,x_n)$ in *natürlichen* Zahlen $x_1,\dots,x_n$ besitzt.

Aufgabe 4.11. Für alle $n,p \in \mathbb{N}$ bezeichnen wir mit $s_p(n) := \sum_{k=1}^{n} k^p = 1^p + \cdots + n^p$ die Summe der p-ten Potenzen aller natürlichen Zahlen von 1 bis n.

(a) Beweise für alle $n,p \in \mathbb{N}$ die Formel

$$\binom{p+1}{0} s_0(n) + \binom{p+1}{1} s_1(n) + \cdots + \binom{p+1}{p} s_p(n) = (n+1)^{p+1} - 1.$$

(b) Zeige mit Hilfe von (a), dass s_2 ein Polynom in n ist, und berechne dieses Polynom explizit. Ist s_p für alle $p \in \mathbb{N}$ ein Polynom in n?

Aufgabe 4.12. Für ein gegebenes $n \in \mathbb{N}$ betrachten wir das Polynom

$$f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto x^n + x^{n-1} + \cdots + x + 1.$$

Man zeige:

(a) Ist n ungerade, so hat f genau eine Nullstelle. Was ist ihre Vielfachheit?

(b) Ist n gerade, so hat f keine Nullstelle.

4.B Geordnete Körper

Wir haben bisher von den reellen Zahlen nur die Körpereigenschaften, also die Eigenschaften der vier Grundrechenarten ausgenutzt – und dabei z. B. in Beispiel 3.6 (b) gesehen, dass es außer den reellen Zahlen auch noch ganz andere (und in der Tat sogar sehr viele) Körper gibt. Wir müssen also noch weitere Eigenschaften auflisten, um die reellen Zahlen eindeutig zu charakterisieren. Dies wollen wir im Rest dieses Kapitels tun.

Eine Eigenschaft der reellen Zahlen, die wir bisher völlig vernachlässigt haben, ist, dass man sie *ordnen* kann, also dass man zwei Zahlen der Größe nach vergleichen kann. Die Eigenschaften dieser Ordnungsrelation werden im Begriff des sogenannten *geordneten Körpers* formalisiert.

Definition 4.13 (Geordnete Körper)**.** Ein Körper K heißt **geordneter** oder **angeordneter Körper**, wenn in ihm eine Menge $P \subset K$ (die „Menge der positiven Zahlen“) gegeben ist, so dass die folgenden drei Eigenschaften gelten:

(a) Für alle $x \in K$ gilt *genau* eine der drei Eigenschaften $x = 0$, $x \in P$ oder $-x \in P$. (Im zweiten Fall nennt man x eine **positive** Zahl, im dritten eine **negative** Zahl.)

(b) Für alle $x, y \in P$ ist $x + y \in P$ („die Summe zweier positiver Zahlen ist positiv“).

(c) Für alle $x, y \in P$ ist $xy \in P$ („das Produkt zweier positiver Zahlen ist positiv“).

In diesem Fall schreibt man $x < y$ oder $y > x$ falls $y - x \in P$, und $x \le y$ oder $y \ge x$ falls $y - x \in P$ oder $y = x$. (Insbesondere ist also $x > 0$ genau dann, wenn $x \in P$, und $x < 0$ genau dann, wenn $-x \in P$; außerdem gilt nach (a) für $x, y \in K$ stets genau eine der Aussagen $x = y$, $x < y$ oder $y < x$.)

Beispiel 4.14.

(a) $\mathbb{R}$ ist ein geordneter Körper (was wir hier wiederum axiomatisch voraussetzen wollen). Nennt man in der Teilmenge $\mathbb{Q}$ von $\mathbb{R}$ genau die Zahlen positiv, die es auch in $\mathbb{R}$ sind, so ist damit auch $\mathbb{Q}$ ein geordneter Körper.

(b) Der Körper $\mathbb{Z}_2$ aus Beispiel 3.6 (b) kann nicht zu einem geordneten Körper gemacht werden: Das Element $1 = u$ ist nicht gleich 0, also müsste nach Definition 4.13 (a) genau eine der beiden Eigenschaften $1 \in P$ und $-1 \in P$ gelten. Dies ist aber unmöglich, da wegen $1 + 1 = 0$ in $\mathbb{Z}_2$ die Gleichung $-1 = 1$ gilt.

Bemerkung 4.15 (Partielle und totale Ordnungen)**.** Ist K ein geordneter Körper, so ist $\le$ wie in Definition 4.13 natürlich eine Relation auf K im Sinne von Definition 2.1. Sie besitzt die folgenden Eigenschaften für alle $x, y, z \in K$:

(a) **Reflexivität**: Es gilt $x \le x$.

(b) **Antisymmetrie**: Ist $x \le y$ und $y \le x$, so folgt $x = y$.

(c) **Transitivität**: Gilt $x \le y$ und $y \le z$, so folgt auch $x \le z$.

(d) **Totalität**: Es gilt (mindestens) eine der Aussagen $x \le y$ und $y \le x$. (Mit anderen Worten: „Zwei beliebige Elemente von K sind stets miteinander vergleichbar.“)

In der Tat folgt (a) unmittelbar aus Definition 4.13. Da $x \le y$ nach Definition äquivalent zu $y - x \in P$ oder $y - x = 0$ ist, und $y \le x$ zu $-(y - x) \in P$ oder $y - x = 0$, ergeben sich (b) und (d) außerdem aus Definition 4.13 (a). Die Aussage (c) schließlich ist trivial falls $x = y$ oder $y = z$; andernfalls gilt $y - x \in P$ oder $z - y \in P$, und damit $z - x = (y - x) + (z - y) \in P$, also $x < z$, nach Definition 4.13 (b).

Auf einer beliebigen Menge K (die also nicht notwendig ein Körper ist) nennt man eine Relation $\le$ mit den Eigenschaften (a), (b) und (c) eine **partielle Ordnung**. Gilt zusätzlich noch (d), so heißt $\le$ eine **(totale) Ordnung** auf K. Jeder geordnete Körper K liefert also eine totale Ordnung auf K.

Das Standardbeispiel für eine partielle Ordnung auf einer Menge ist die Teilmengenrelation auf der Potenzmenge $\mathscr{P}(M)$ einer beliebigen Menge M: Sind $A, B, C \in \mathscr{P}(M)$ Teilmengen von M, so gilt natürlich $A \subset A$; aus $A \subset B$ und $B \subset A$ folgt $A = B$; und aus $A \subset B$ und $B \subset C$ folgt $A \subset C$. Diese partielle Ordnung ist aber in der Regel nicht total: Für $M = \mathbb{N}$ sind die Teilmengen $A = \{0\}$ und $B = \{1\}$ nicht vergleichbar, denn es gilt weder $A \subset B$ noch $B \subset A$.

Wie schon bei den Körpern wollen wir nun auch hier für einen geordneten Körper kurz die wichtigsten Eigenschaften ableiten, die aus der Definition folgen (und die euch für die reellen Zahlen sicher bekannt sind). Wir werden sie im Folgenden verwenden, ohne jedes Mal darauf hinzuweisen.

Lemma 4.16 (Eigenschaften geordneter Körper)**.** *Für alle x, y, z in einem geordneten Körper K gilt:*

(a) *Ist $x < y$, so folgt $x + z < y + z$.*

(b) *Ist $x < y$ und $z > 0$, so gilt auch $xz < yz$. Ist dagegen $x < y$ und $z < 0$, so folgt $xz > yz$.*

(Ungleichungen drehen sich also bei der Multiplikation mit einer negativen Zahl um.)

(c) *Gilt $x \ne 0$, so ist $x^2 > 0$. Insbesondere ist also $1 > 0$.*

(d) *Wenn $0 < x < y$, dann folgt $0 < y^{-1} < x^{-1}$.*

Entsprechende Aussagen gelten natürlich auch für die nicht-strikten Ungleichungen $\le$ bzw. $\ge$.

Beweis.

(a) Ist $x < y$, also $y - x \in P$, so ist auch $(y+z) - (x+z) = y - x \in P$, also $x + z < y + z$.

(b) Gilt wieder $y - x \in P$ und $z \in P$, so folgt aus Definition 4.13 (c) auch $(y-x)z = yz - xz \in P$, also $xz < yz$. Ist hingegen $z < 0$, also $-z \in P$, so gilt diesmal $(y-x)(-z) = xz - yz \in P$, also $yz < xz$.

(c) Ist $x \in P$, so ist natürlich auch $x^2 \in P$ nach Definition 4.13 (c). Ist $-x \in P$, so folgt genauso $x^2 = (-x)^2 \in P$. Also ist für $x \neq 0$ in jedem Fall $x^2 > 0$. Insbesondere ist damit $1 = 1 \cdot 1 > 0$.

(d) Gilt $x \in P$, so folgt aus (c) und Definition 4.13 (c) zunächst $x^{-1} = x \cdot (x^{-1})^2 \in P$, also $x^{-1} > 0$. Genauso ergibt sich $y^{-1} > 0$. Ist nun $x < y$, so folgt aus (b) durch Multiplikation mit der positiven Zahl $x^{-1}y^{-1}$ die Ungleichung $xx^{-1}y^{-1} < yx^{-1}y^{-1}$, also $y^{-1} < x^{-1}$, was zu zeigen war. □

Notation 4.17 (Intervalle und Betrag)**.** Die folgenden Notationen verwendet man häufig in einem geordneten Körper K.

(a) Sind $a, b \in K$ mit $a \leq b$, so definiert man die folgenden Teilmengen von K:

- $[a,b] := \{x \in K : a \leq x \leq b\}$ (**abgeschlossene** bzw. **kompakte Intervalle**);
- $(a,b) := \{x \in K : a < x < b\}$ (**offene Intervalle**);
- $[a,b) := \{x \in K : a \leq x < b\}$ (**halboffene Intervalle**);
- $[a,\infty) := K_{\geq a} := \{x \in K : x \geq a\}$ (**uneigentliche Intervalle**);

und analog natürlich $(a,b]$, (a,∞), $(-\infty,b]$ und $(-\infty,b)$. Wenn wir derartige Intervalle im Fall $K = \mathbb{R}$ graphisch darstellen, deuten wir wie im Bild unten meistens durch Rundungen an den Intervallgrenzen an, ob die Randpunkte mit dazugehören sollen oder nicht. 06

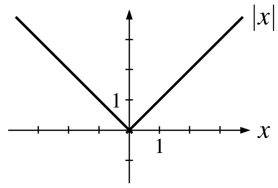
(b) Für $x \in K$ definieren wir den **Betrag** von x als

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

Insbesondere gilt also immer $|x| \geq 0$. Im Fall $K = \mathbb{R}$ sieht die Betragsfunktion natürlich wie im folgenden Bild rechts aus.

(a,b)

$[a,\infty)$



Die wichtigsten beiden Eigenschaften der Betragsfunktion sind ihre „Verträglichkeit" mit Addition und Multiplikation:

Lemma 4.18 (Eigenschaften der Betragsfunktion)**.** *Für alle x, y in einem geordneten Körper K gilt:*

(a) $|xy| = |x| \cdot |y|$. *Insbesondere ergibt sich daraus für $y = -1$, dass $|-x| = |x|$.*

(b) $x \leq |x|$.

(c) $|x + y| \leq |x| + |y|$. *Diese Ungleichung bezeichnet man als **Dreiecksungleichung** – wir werden in Bemerkung 6.10 (a) sehen, warum.*

Beweis.

(a) Wir machen eine Fallunterscheidung je nach Vorzeichen von x und y. Ist z. B. $x \geq 0$ und $y \leq 0$, so ist $xy \leq 0$ und damit nach Definition des Betrages $|x| = x$, $|y| = -y$ und $|xy| = -xy$. Zusammensetzen dieser Gleichungen ergibt die Behauptung $|xy| = -xy = x \cdot (-y) = |x| \cdot |y|$. Die anderen Fälle der möglichen Vorzeichenverteilungen beweist man genauso.

(b) Für $x \leq 0$ ist $x \leq 0 \leq |x|$; für $x \geq 0$ ist $x = |x|$.

(c) Nach (b), angewendet auf x und y, gilt (mit Lemma 4.16 (a))

$$x+y \leq |x|+|y|. \tag{1}$$

Wenden wir (b) hingegen auf $-x$ und $-y$ an, so folgt auch

$$-x-y \leq |-x|+|-y| = |x|+|y|. \tag{2}$$

Aber $|x+y|$ ist in jedem Fall eine der beiden Zahlen $x+y$ oder $-x-y$. Damit folgt die Behauptung $|x+y| \leq |x|+|y|$ aus (1) und (2). □

Bemerkung 4.19 (Dreiecksungleichung nach oben und unten)**.** Die Dreiecksungleichung aus Lemma 4.18 (c) schätzt den Betrag $|x+y|$ einer Summe nach oben ab. Offensichtlich gilt im Allgemeinen keine Gleichheit, wie das Beispiel $x=1$, $y=-1$ zeigt: Hier ist $|x+y| = 0 < 2 = |x|+|y|$.

Eine Abschätzung nach unten kann man erhalten, indem man Lemma 4.18 (c) auf die Zahlen $x+y$ und $-y$ anwendet: Man erhält dann nämlich $|x| = |(x+y)-y| \leq |x+y|+|-y| = |x+y|+|y|$ und damit $|x+y| \geq |x|-|y|$. Insgesamt haben wir also für alle $x,y \in K$

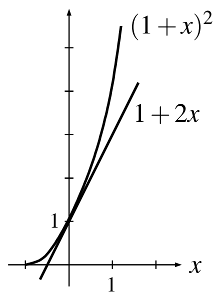
$$|x|-|y| \leq |x+y| \leq |x|+|y|.$$

Eine weitere Anwendung der Eigenschaften eines geordneten Körpers ist die folgende Ungleichung, die oftmals dann nützlich ist, wenn die Größe von Potenzen x^n mit der von Produkten $n \cdot x$ verglichen werden soll.

Satz 4.20 (Bernoullische Ungleichung). *Es seien K ein geordneter Körper, $x \in K$ mit $x \geq -1$, und $n \in \mathbb{N}$. Dann gilt $(1+x)^n \geq 1+nx$.*

Beweis. Wir zeigen die Aussage mit Induktion über n. Das Bild rechts unten veranschaulicht die Ungleichung im Fall $K=\mathbb{R}$ und $n=2$.

Der Induktionsanfang für $n=0$ ist klar: dann sind beide Seiten gleich 1, die Ungleichung ist also erfüllt.



Für den Induktionsschritt nehmen wir nun an, dass

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

für alle $x \geq -1$ und ein gegebenes $n \in \mathbb{N}$ gilt. Mit Lemma 4.16 (b) können wir diese Ungleichung mit der nach Voraussetzung nicht-negativen Zahl $1+x$ multiplizieren und erhalten

$$(1+x)^{n+1} \geq (1+nx)(1+x) = 1+(n+1)x+nx^2.$$

Nach Lemma 4.16 (c) ist nun $nx^2 \geq 0$ und damit $(1+x)^{n+1} \geq 1+(n+1)x$, was zu zeigen war. □

Aufgabe 4.21. Für welche $x,y \in \mathbb{R}$ bzw. $n \in \mathbb{N}$ gelten die folgenden Ungleichungen?

(a) $\dfrac{2xy}{x+y} \leq \dfrac{x+y}{2}$ (b) $\dfrac{4^n}{2n+1} < \dbinom{2n}{n} < 4^n$ (c) $n! < \left(\dfrac{n}{2}\right)^n$

4.C Supremum und Infimum

Wie bereits angekündigt wollen wir in diesem Abschnitt nun endlich eine eindeutige axiomatische Charakterisierung der reellen Zahlen angeben. Bisher haben wir nur gesehen, dass $\mathbb{R}$ ein geordneter Körper ist. Aber auch $\mathbb{Q}$ ist ein geordneter Körper, und daher müssen wir noch untersuchen, wie sich $\mathbb{R}$ von $\mathbb{Q}$ unterscheidet. Anschaulich würden wir dies wohl so formulieren, dass die Menge $\mathbb{Q}$ der rationalen Zahlen „Löcher“ auf der Zahlengeraden hat, also dass es Punkte wie z. B. $\sqrt{2}$ auf der Zahlengeraden gibt, die keiner rationalen Zahl entsprechen (siehe Lemma 4.36). Aber wie formuliert man so etwas exakt?

Um dies zu sehen, müssen wir genauer untersuchen, wann Teilmengen eines geordneten Körpers K größte bzw. kleinste Elemente haben. Dazu können wir natürlich zunächst auf die offensichtliche Art das Maximum und Minimum zweier Elemente von K definieren.

Definition 4.22 (Maximum und Minimum endlich vieler Zahlen)**.** Für zwei Elemente x und y eines geordneten Körpers definieren wir das **Maximum** und **Minimum** durch

$$\max(x,y) := \begin{cases} x & \text{falls } x \geq y, \\ y & \text{falls } y \geq x \end{cases} \quad \text{und} \quad \min(x,y) := \begin{cases} x & \text{falls } x \leq y, \\ y & \text{falls } y \leq x. \end{cases}$$

Analog kann man auch das Maximum und Minimum von mehr als zwei Zahlen definieren, *sofern es nur endlich viele sind.* Betrachten wir dagegen eine (unendliche) Teilmenge $M \subset K$, so können wir in der Regel nicht mehr erwarten, dass M ein kleinstes bzw. größtes Element hat, allein schon weil die Menge dann in folgendem Sinne unbeschränkt sein kann:

Definition 4.23 (Beschränkte Mengen)**.** Es sei M eine Teilmenge eines geordneten Körpers K.

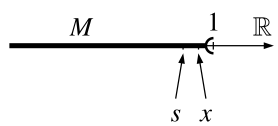
(a) Ein Element $s \in K$ heißt **obere Schranke** für M, wenn $x \leq s$ für alle $x \in M$. Existiert eine solche obere Schranke für M, so nennt man M **nach oben beschränkt**. Analog definiert man den Begriff einer **unteren Schranke** bzw. einer **nach unten beschränkten** Menge.

(b) Die Menge M heißt **beschränkt**, wenn sie nach oben und unten beschränkt ist, oder – äquivalent dazu – wenn die Menge ihrer Beträge beschränkt ist, also wenn es ein $s \in K$ gibt mit $|x| \leq s$ für alle $x \in M$.

Beispiel 4.24. Es sei $M = \mathbb{R}_{<1}$ im geordneten Körper $\mathbb{R}$.

(a) Die Menge M ist nach oben (aber nicht nach unten) beschränkt, denn $s = 1$ ist eine obere Schranke für M. Genauso ist auch $s = 2$ eine obere Schranke für M, auch wenn man anschaulich vielleicht sagen würde, dass diese Schranke „nicht so gut“ ist, weil $x \leq 2$ für alle $x \in M$ eine schwächere Aussage ist als $x \leq 1$ für alle $x \in M$.

(b) Ist $s \in M$, also $s < 1$, so gilt für den Mittelpunkt $x := \frac{s+1}{2}$ zwischen s und 1 wie im Bild rechts natürlich



$$x = \frac{s+1}{2} > \frac{s+s}{2} = s \quad \text{und} \quad x = \frac{s+1}{2} < \frac{1+1}{2} = 1.$$

Hieraus ergeben sich sofort zwei einfache Folgerungen:

- Es gibt keine größte Zahl in M (denn zu $s \in M$ liegt die größere Zahl x wegen $x < 1$ ebenfalls noch in M).
- Die Menge M ist aber nach (a) nach oben beschränkt, und die kleinste mögliche obere Schranke für M ist 1 (denn $s < 1$ kann keine obere Schranke mehr sein, da $x > s$ ebenfalls noch in M liegt).

Die Zahl 1 kann damit als „Obergrenze der Zahlen in M“ angesehen werden, auch wenn sie kein Element von M und daher keine größte Zahl in M ist. Dieses Konzept wollen wir jetzt exakt definieren.

Definition 4.25 (Supremum und Infimum)**.** Es sei M eine Teilmenge eines geordneten Körpers K.

(a) Eine Zahl $s \in K$ heißt ein **Supremum** von M, wenn s eine „kleinste obere Schranke für M“ ist, d. h. wenn gilt:

- s ist eine obere Schranke für M; und
- für jede obere Schranke s' für M gilt $s \leq s'$.

(b) Eine Zahl $s \in K$ heißt ein **Maximum** von M, wenn s eine „obere Schranke für M in M“ ist, d. h. wenn gilt:

- s ist eine obere Schranke für M; und
- $s \in M$.

Analog heißt $s \in K$ ein **Infimum** von M, wenn s eine größte untere Schranke für M ist, und ein **Minimum**, wenn s eine untere Schranke für M in M ist.

Bemerkung und Notation 4.26 ($\sup M$, $\max M$, $\inf M$, $\min M$)**.**

(a) Jedes Maximum s einer Menge M ist auch ein Supremum von M: Ist dann nämlich $s' \in K$ eine weitere obere Schranke, so folgt wegen $s \in M$ natürlich sofort $s \leq s'$.

(b) Wenn die Menge M ein Supremum besitzt, dann ist es eindeutig bestimmt: Sind nämlich s_1 und s_2 zwei kleinste obere Schranken, so folgt nach Definition 4.25 (a) sofort $s_1 \leq s_2$ (weil s_1 eine kleinste obere Schranke und s_2 auch eine obere Schranke ist) und $s_2 \leq s_1$ (weil s_2 Supremum und s_1 auch eine obere Schranke ist), also $s_1 = s_2$. Nach (a) ist damit auch ein Maximum von M eindeutig bestimmt, falls es existiert.

Wenn ein Supremum oder Maximum von M existiert, können wir also von *dem* Supremum bzw. *dem* Maximum von M reden. Wir bezeichnen es dann mit $\sup M$ bzw. $\max M$.

(c) Analog sind natürlich auch Infimum und Minimum von M eindeutig bestimmt, sofern sie existieren; wir bezeichnen sie dann mit $\inf M$ bzw. $\min M$.

Beispiel 4.27.

(a) Die Menge $M = \mathbb{R}_{\leq 1}$ hat offensichtlich das Maximum 1. Nach Bemerkung 4.26 (a) ist 1 damit auch das Supremum von M, d. h. es ist $\sup M = \max M = 1$.

(b) Das uneigentliche Intervall $M = \mathbb{R}_{<1}$ hat nach Beispiel 4.24 (b) kein Maximum, aber es gilt $\sup M = 1$.

(c) Das uneigentliche Intervall $M = \mathbb{R}_{>1}$ besitzt kein Supremum (und damit auch kein Maximum), da M nach oben unbeschränkt ist, also nicht einmal irgendeine obere Schranke für M existiert – insbesondere also keine kleinste. Auch die leere Menge hat kein Supremum, weil für sie jede reelle Zahl eine obere Schranke ist und damit keine kleinste obere Schranke existiert.

Analog ergeben sich natürlich Supremum, Maximum, Infimum und Minimum von allen Intervallen aus Beispiel 4.17 (a).

Aufgabe 4.28. Es seien A und B zwei Teilmengen eines geordneten Körpers K, die ein Supremum $\sup A$ bzw. $\sup B$ besitzen. Setzen wir $A+B := \{x+y : x \in A, y \in B\}$ und $-A := \{-x : x \in A\}$, so zeige man, dass auch die folgenden Suprema und Infima existieren und die behaupteten Werte haben:

(a) $\sup(A \cup B) = \max(\sup A, \sup B)$.

(b) $\inf(-A) = -\sup A$.

(c) $\sup(A+B) = \sup A + \sup B$.

Unsere bisher betrachteten Beispiele von Mengen ohne Supremum in Beispiel 4.27 (c) waren letztlich trivial – also die leere Menge sowie Mengen, die überhaupt keine obere Schranke besitzen. Ist es auch möglich, dass eine Menge zwar nicht leer und nach oben beschränkt ist, aber trotzdem keine *kleinste* obere Schranke hat? In der Tat ist dies in $\mathbb{R}$ im Gegensatz zu $\mathbb{Q}$ nicht möglich, und wie wir sehen werden ist genau dies der wesentliche Unterschied zwischen diesen beiden Körpern. Wir definieren daher zunächst:

Definition 4.29 (Supremumsaxiom)**.** Wir sagen, dass ein geordneter Körper das **Supremumsaxiom** erfüllt, wenn in ihm jede nicht leere, nach oben beschränkte Teilmenge ein Supremum besitzt. (Natürlich besitzt dann nach Aufgabe 4.28 (b) auch jede nicht leere, nach unten beschränkte Teilmenge ein Infimum.)

Die reellen Zahlen erfüllen also dieses Supremumsaxiom – das werden wir in dieser Vorlesung axiomatisch voraussetzen, und das ist nun endlich auch die letzte Eigenschaft der reellen Zahlen, die wir benötigen. Wenn wir dieses und das vorangegangene Kapitel zusammenfassen, setzen wir nun also insgesamt über die reellen Zahlen voraus:

$$\boxed{\mathbb{R} \text{ ist ein geordneter Körper, der das Supremumsaxiom erfüllt.}}$$

Wie schon in Notation 1.14 gesagt, kann man die Existenz der reellen Zahlen auch aus den Axiomen der Logik und Mengenlehre herleiten – und dann ist es natürlich ein beweisbarer Satz, dass $\mathbb{R}$ ein geordneter Körper ist, der das Supremumsaxiom erfüllt. Man kann sogar noch mehr zeigen, nämlich dass diese Eigenschaften die reellen Zahlen auch vollständig charakterisieren: $\mathbb{R}$ ist in der Tat der *einzige* geordnete Körper, der das Supremumsaxiom erfüllt. Der Beweis dieser Aussage ist jedoch sehr schwierig und soll hier nicht gegeben werden, zumal wir diese Aussage auch nicht benötigen werden. Wir werden lediglich in Bemerkung 4.37 noch sehen, dass $\mathbb{Q}$ das Supremumsaxiom in der Tat nicht erfüllt.

Für uns bedeutet diese Tatsache letztlich nur, dass wir ab jetzt alles, was wir mit den reellen Zahlen tun möchten, ausschließlich auf den Axiomen eines geordneten Körpers und dem Supremumsaxiom aufbauen können und werden.

Wir wollen nun ein paar erste elementare Folgerungen aus dem Supremumsaxiom ziehen. Seine wahre Stärke wird dieses Axiom jedoch erst im nächsten Kapitel bei der Untersuchung von Grenzwerten von Folgen zeigen.

Satz 4.30 ($\mathbb{R}$ ist **archimedisch geordnet**). *Die Teilmenge $\mathbb{N} = \{0,1,2,\dots\} \subset \mathbb{R}$ ist nach oben unbeschränkt.*

Beweis. Angenommen, $\mathbb{N}$ wäre nach oben beschränkt. Dann würde nach dem Supremumsaxiom $s := \sup \mathbb{N} \in \mathbb{R}$ existieren. Da s die *kleinste* obere Schranke ist, ist $s-1$ keine obere Schranke; es gibt also ein $n \in \mathbb{N}$ mit $n > s-1$. Dann ist aber $n+1 \in \mathbb{N}$ mit $n+1 > s$, im Widerspruch dazu, dass s eine obere Schranke für $\mathbb{N}$ ist. □

Bemerkung 4.31.

(a) Eine einfache, aber oft verwendete Folgerung aus Satz 4.30 ist, dass es zu jeder positiven Zahl $x \in \mathbb{R}_{>0}$ ein $n \in \mathbb{N}_{>0}$ gibt mit $\frac{1}{n} < x$: Da $\mathbb{N}$ nach oben unbeschränkt ist, ist insbesondere $\frac{1}{x}$ keine obere Schranke für $\mathbb{N}$. Also gibt es ein $n \in \mathbb{N}$ mit $n > \frac{1}{x}$, und damit nach Lemma 4.16 (d) mit $\frac{1}{n} < x$.

(b) Satz 4.30 mag zwar selbstverständlich erscheinen – man kann allerdings in der Tat geordnete Körper konstruieren, in denen diese Aussage falsch ist, in denen es also „unendlich große" Elemente gibt, die größer sind als jede Zahl, die man durch fortlaufende Addition der 1 erreichen kann [E, Kapitel 11].

Aufgabe 4.32. Bestimme Supremum, Infimum, Maximum und Minimum (sofern sie existieren) der Menge

$$M = \left\{ \frac{m+n}{mn} : m,n \in \mathbb{N}_{>0} \right\} \quad \subset \mathbb{R}.$$

07

Folgerung 4.33. *Jede nicht-leere, nach oben beschränkte Teilmenge von $\mathbb{Z}$ hat ein Maximum. Entsprechend hat jede nicht-leere, nach unten beschränkte Teilmenge von $\mathbb{Z}$ ein Minimum.*

*Insbesondere hat also jede nicht-leere Teilmenge von $\mathbb{N}$ ein Minimum. (Man sagt dafür auch: $\mathbb{N}$ ist **wohlgeordnet**.)*

Beweis. Es sei $M \subset \mathbb{Z}$ nicht-leer und nach oben beschränkt, mit oberer Schranke s. Nach Satz 4.30 gibt es weiterhin eine natürliche Zahl $b > s$, die dann natürlich auch eine obere Schranke für M ist.

Ist $a \in M$ nun ein beliebiges Element, so ist $M_{\geq a} = M \cap \{a, a+1, \dots, b\}$ eine nicht-leere endliche Menge, die demzufolge natürlich ein Maximum besitzt. Alle anderen Zahlen von M sind aber noch kleiner als a, so dass dieses Maximum also auch das Maximum von M ist. □

Bemerkung 4.34 (*Gaußklammer*). Nach Folgerung 4.33 kann man für alle $x \in \mathbb{R}$ die Zahl

$$\lfloor x \rfloor := \max\{k \in \mathbb{Z} : k \leq x\}$$

definieren, da die Menge aller $k \in \mathbb{Z}$ mit $k \leq x$ nach Satz 4.30 nicht leer ist (es gibt ja ein $n \in \mathbb{N}$ mit $n \geq -x$, wir können dann $k = -n$ setzen), und natürlich durch x nach oben beschränkt. Als die größte ganze Zahl kleiner oder gleich x kann man sie sich als „Abrundung“ von x vorstellen; es ist also z. B.

$$\left\lfloor \frac{7}{2} \right\rfloor = 3 \quad \text{und} \quad \left\lfloor -\frac{7}{2} \right\rfloor = -4.$$

Folgerung 4.35 ($\mathbb{Q}$ liegt **dicht** in $\mathbb{R}$)**.** *Jedes nicht-leere offene Intervall (a,b) in $\mathbb{R}$ enthält eine rationale Zahl.*

Beweis. Nach Bemerkung 4.31 (a) gibt es ein $n \in \mathbb{N}_{>0}$ mit $\frac{1}{n} < b-a$. Weiterhin ist die Menge

$$M := \left\{ k \in \mathbb{Z} : \frac{k}{n} > a \right\} = \{k \in \mathbb{Z} : k > na\}$$

nach Satz 4.30 nicht-leer und durch na nach unten beschränkt, und besitzt damit nach Folgerung 4.33 ein Minimum k. Für dieses Minimum gilt also:

(a) $k \in M$ und damit $\frac{k}{n} > a$;

(b) $k-1 \notin M$ und damit $\frac{k-1}{n} \le a$, d. h.

$$\frac{k}{n} = \frac{k-1}{n} + \frac{1}{n} < a + (b-a) = b.$$

Also ist $\frac{k}{n}$ eine rationale Zahl im offenen Intervall (a,b). $\square$

Als eine Konsequenz aus dieser Folgerung wollen wir wie bereits angekündigt nun sehen, dass die rationalen Zahlen das Supremumsaxiom nicht erfüllen, dass hier also der entscheidende Unterschied zwischen $\mathbb{Q}$ und $\mathbb{R}$ liegt. Wir untersuchen dazu zunächst, ob es eine Quadratwurzel aus 2 gibt, also eine positive Zahl x mit $x^2 = 2$ (die wir dann als $\sqrt{2}$ schreiben).

Lemma 4.36 (Irrationalität von $\sqrt{2}$)**.** *Es gibt keine rationale Zahl $x \in \mathbb{Q}$ mit $x^2 = 2$.*

Beweis. Angenommen, es gäbe eine rationale Zahl $x \in \mathbb{Q}$ mit $x^2 = 2$. Wir können x als gekürzten Bruch $x = \frac{p}{q}$ (mit $p,q \in \mathbb{Z}$ und $q \neq 0$) schreiben und erhalten

$$x^2 = \frac{p^2}{q^2} = 2, \quad \text{also} \quad p^2 = 2q^2. \tag{$*$}$$

Also muss p^2 und damit auch p selbst eine gerade Zahl, d. h. durch 2 teilbar sein. Wir können daher $p = 2r$ für eine ganze Zahl $r \in \mathbb{Z}$ setzen. Einsetzen in $(*)$ liefert

$$(2r)^2 = 2q^2, \quad \text{also} \quad q^2 = 2r^2.$$

Aber dann muss auch q^2 und damit q eine gerade Zahl sein – was ein Widerspruch dazu ist, dass die Darstellung von x als Bruch $\frac{p}{q}$ als gekürzt vorausgesetzt worden ist. $\square$

Wie ihr aus der Schule wisst, gibt es aber in den reellen Zahlen $\mathbb{R}$ eine Wurzel aus 2. Da unsere Axiome, dass $\mathbb{R}$ ein geordneter Körper ist, der das Supremumsaxiom erfüllt, die reellen Zahlen ja vollständig beschreiben, könnten wir die Existenz dieser Wurzel jetzt sogar aus unseren Axiomen beweisen: Die Menge $\{x \in \mathbb{R} : x^2 < 2\}$ ist offensichtlich eine nicht-leere, nach oben beschränkte Menge, die demzufolge ein Supremum s besitzt – und für dieses Supremum kann man zeigen, dass $s^2 = 2$ gilt, also dass s eine Wurzel aus 2 ist. Dieser Beweis ist jedoch recht technisch, und da wir in Folgerung 5.30 ohnehin die Existenz reeller Quadratwurzeln aus beliebigen nicht-negativen Zahlen aus unseren Axiomen beweisen werden, wollen wir hier darauf verzichten und für die folgende Bemerkung der Einfachheit halber annehmen, dass $\sqrt{2}$ in $\mathbb{R}$ existiert, zumal wir diese Bemerkung im Folgenden nicht verwenden werden.

Bemerkung 4.37 ($\mathbb{Q}$ erfüllt das Supremumsaxiom nicht)**.** Es sei $M := \{x \in \mathbb{Q} : x < \sqrt{2}\}$. Diese Menge ist offensichtlich nicht leer (denn $0 \in M$) und nach oben beschränkt (z. B. mit oberer Schranke 2). Würde $\mathbb{Q}$ das Supremumsaxiom erfüllen, müsste sie also ein Supremum $s := \sup M \in \mathbb{Q}$ besitzen. Nach Lemma 4.36 ist damit $s = \sqrt{2}$ ausgeschlossen, also kann nur $s < \sqrt{2}$ oder $s > \sqrt{2}$ gelten. Aber beides führt sofort zum Widerspruch:

- $s < \sqrt{2}$ kann keine obere Schranke für M sein, denn nach Folgerung 4.35 gäbe es dann eine rationale Zahl $x \in (s, \sqrt{2})$, also mit $x \in M$, aber $x > s$.
- $s > \sqrt{2}$ kann keine *kleinste* obere Schranke für M sein, denn wieder nach Folgerung 4.35 gäbe es nun eine rationale Zahl $s' \in (\sqrt{2}, s)$, die kleiner ist als s, aber immer noch eine obere Schranke für M (da aus $x \in M$, also $x < \sqrt{2}$, ja sofort auch $x < s'$ folgt).

Also erfüllt $\mathbb{Q}$ nicht das Supremumsaxiom.

Bemerkung 4.38 (Uneigentliche Suprema)**.** Nach dem Supremumsaxiom existiert das Supremum $\sup M$ für jede nicht leere, nach oben beschränkte Teilmenge $M \subset \mathbb{R}$. Oft ist es praktisch, diese Notation wie folgt auf beliebige Teilmengen von $\mathbb{R}$ zu erweitern:

(a) Ist $M = \emptyset$, so schreiben wir formal $\sup M = -\infty$. (Anschaulich: Jede reelle Zahl ist eine obere Schranke der leeren Menge, daher ist „die kleinste" davon $-\infty$.)

(b) Ist $M \neq \emptyset$ nach oben unbeschränkt, so schreiben wir formal $\sup M = \infty$. (Anschaulich: Ist M nach oben unbeschränkt, so ist keine reelle Zahl eine obere Schranke für M, die einzige und damit kleinste „obere Schranke" für M ist also ∞.)

Man spricht in diesem Fall von *uneigentlichen Suprema*. Analog setzt man natürlich $\inf \emptyset = \infty$ und $\inf M = -\infty$ für jede nach unten unbeschränkte Menge $M \subset \mathbb{R}$.

Mit „wir schreiben formal" ist dabei oben gemeint, dass $-\infty$ und ∞ natürlich keine Zahlen sind, mit denen man uneingeschränkt wie gewohnt rechnen kann. Manche Rechenoperationen wie z. B. $\infty + x := \infty$ für alle $x \in \mathbb{R}$ lassen sich zwar noch intuitiv sinnvoll definieren (siehe Bemerkung 5.42), aber andere wie z. B. $\infty - \infty$ nicht. Mit anderen Worten ist $\mathbb{R} \cup \{\pm\infty\}$ kein Körper. Trotzdem hat die Einführung uneigentlicher Suprema und Infima den Vorteil, dass $\sup M$ und $\inf M$ für *jede* Teilmenge M von $\mathbb{R}$ definiert sind, und Aussagen darüber oft auch in den neuen Fällen gültig bleiben. So gelten z. B. die Aussagen aus Aufgabe 4.28 sogar für beliebige Teilmengen von $\mathbb{R}$ (sofern in (c) rechts nicht der unbestimmte Ausdruck $\infty - \infty$ auftritt) – allerdings brauchen wir dafür natürlich einen neuen Beweis, da ja schon die Definition eines uneigentlichen Supremums eine andere als üblich ist.

5. Folgen und Grenzwerte

Nachdem wir die reellen Zahlen vollständig charakterisiert haben, wollen wir jetzt zur eigentlichen Analysis kommen. Der zentrale Begriff ist dabei der des Grenzwerts, den ihr ja sicher in der Schule schon in der einen oder anderen Form kennengelernt habt und den wir jetzt exakt einführen wollen. Wir beginnen dabei mit Grenzwerten von Folgen, da sie für den Anfang einfacher sind als die später auch noch wichtigen Grenzwerte von Funktionen.

5.A Grenzwerte von Folgen

Zur Untersuchung des Grenzwertbegriffs müssen wir als Erstes exakt definieren, was wir damit meinen, dass sich eine (unendlich lange) Folge reeller Zahlen einem Wert beliebig genau annähert.

Definition 5.1 (Folgen und Grenzwerte)**.**

(a) Eine **Folge** in einer Menge M ist eine Abbildung

$$\mathbb{N} \to M, \ \ n \mapsto a_n.$$

Man schreibt eine solche Folge als $(a_n)_{n\in\mathbb{N}}$, einfach nur als $(a_n)_n$, oder durch Aufzählen der Folgenglieder als $(a_0, a_1, a_2, \dots)$. Hin und wieder ist in der Literatur auch die noch weiter verkürzte Schreibweise (a_n) zu finden, die wir hier allerdings nicht verwenden wollen, um Verwechslungen der Folge $(a_n)_n$ mit einem zufällig eingeklammerten Folgenglied a_n zu vermeiden.

Manchmal ist es bequem, Folgen nicht beim Index 0, sondern bei einem anderen Startindex $n_0 \in \mathbb{Z}$ beginnen zu lassen – wenn man dies in der Notation deutlich machen möchte, schreibt man derartige Folgen als $(a_n)_{n\geq n_0}$.

In diesem Kapitel werden wir nur den Fall $M = \mathbb{R}$, also sogenannte *reelle Folgen* betrachten. Wir werden daher oft nur von einer Folge sprechen und damit dann immer eine reelle Folge meinen. Später werden wir auch noch andere Folgen kennenlernen, z. B. Folgen komplexer Zahlen in Abschnitt 6.C oder Folgen von Funktionen in Abschnitt 8.C.

(b) Eine Zahl $a \in \mathbb{R}$ heißt **Grenzwert** einer Folge $(a_n)_n$, wenn

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

Wir werden gleich in Lemma 5.5 sehen, dass eine Folge höchstens einen solchen Grenzwert besitzen kann. Wenn ein solches a existiert, können wir also sagen, dass a *der* Grenzwert der Folge $(a_n)_n$ ist. Man nennt die Folge in diesem Fall **konvergent** (gegen a) und schreibt dies als

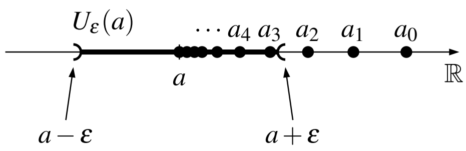
$$\lim_{n\to\infty} a_n = a$$

(die Bezeichnung kommt vom englischen Wort „limit“ bzw. dem lateinischen „limes“), oder manchmal auch als $a_n \to a$ (für $n \to \infty$). Existiert ein solcher Grenzwert nicht, so heißt die Folge **divergent**.

Bemerkung 5.2 (Anschauliche Deutung des Grenzwertbegriffs)**.** Um die Definition des Grenzwertes in leicht verständliche Worte zu fassen, führen wir ein paar intuitive Notationen ein. Für $a \in \mathbb{R}$ und $\varepsilon \in \mathbb{R}_{>0}$ heißt das offene Intervall

$$U_\varepsilon(a) := \{x \in \mathbb{R} : |x - a| < \varepsilon\} = (a - \varepsilon, a + \varepsilon)$$

die **ε-Umgebung** von a. Die Grenzwertbedingung besagt nun genau, dass in jeder solchen ε-Umgebung von a – egal wie klein das ε gewählt ist – alle Folgenglieder ab einem gewissen n_0 liegen, wobei dieses n_0 natürlich von dem gewählten ε abhängen darf. Im Beispielbild unten wäre das z. B. für $n_0 = 3$ der Fall, denn $a_3, a_4, a_5, \dots$ liegen alle in $U_\varepsilon(a)$.



Man kann diese Tatsache auch so ausdrücken, dass in jeder ε-Umgebung *alle bis auf endlich viele* Folgenglieder liegen müssen (nämlich alle bis auf evtl. $a_0,\dots,a_{n_0-1}$). In der Analysis verwendet man gerne den Buchstaben ε für eine kleine positive Zahl und die Sprechweise „**fast alle**" für „alle bis auf endlich viele", und kann damit die Grenzwertbedingung auch in Worten formulieren:

Eine Zahl a ist genau dann Grenzwert einer Folge, wenn in jeder ε-Umgebung von a fast alle Folgenglieder liegen.

Anschaulich bedeutet das natürlich einfach, dass sich die Folgenglieder immer mehr dem Grenzwert annähern. Beachte auch, dass dies insbesondere bedeutet, dass das Abändern oder Weglassen endlich vieler Folgenglieder nichts daran ändert, ob und gegen welchen Grenzwert eine Folge konvergiert. Der Startindex einer Folge wie in Definition 5.1 (a) ist für ihre Konvergenz also irrelevant.

Beispiel 5.3. Hier sind ein paar sehr wichtige Beispiele von Grenzwerten:

(a) Es ist offensichtlich, dass eine konstante Folge, in der alle Folgenglieder den gleichen Wert $a \in \mathbb{R}$ haben, gegen eben dieses a konvergiert, d. h. dass $\lim\limits_{n\to\infty} a = a$ gilt: Hier liegen ja sogar *alle* Folgenglieder in jeder beliebigen ε-Umgebung von a.

(b) Wir behaupten, dass $\lim\limits_{n\to\infty} \frac{1}{n} = 0$ gilt.

Um dies mit Hilfe der Definition 5.1 (b) zu beweisen, sei zunächst ein $\varepsilon \in \mathbb{R}_{>0}$ beliebig vorgegeben; wir müssen zeigen, dass fast alle Glieder der Folge $\left(\frac{1}{n}\right)_{n\geq 1}$ in der ε-Umgebung von 0 liegen. Dies ist aber sehr einfach: Nach der archimedischen Ordnung von $\mathbb{R}$ wie in Bemerkung 4.31 (a) gibt es ein $n_0 \in \mathbb{N}$ mit $\frac{1}{n_0} < \varepsilon$. Mit einem solchen n_0 gilt dann für alle $n \geq n_0$

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} \leq \frac{1}{n_0} < \varepsilon,$$

wobei wir die Rechenregeln für Ungleichungen aus Lemma 4.16 verwendet haben. Fast alle Folgenglieder, nämlich alle $\frac{1}{n}$ für $n \geq n_0$, liegen also in der ε-Umgebung von 0. Damit gilt nach Definition $\lim\limits_{n\to\infty} \frac{1}{n} = 0$.

Beachte, dass wir hierbei zu unserem ε gar kein *konkretes* n_0 angegeben haben, das die Grenzwertbedingung erfüllt. Wir hätten dies hier leicht tun können, z. B.

$$n_0 := \left\lfloor \frac{1}{\varepsilon} + 1 \right\rfloor$$

mit der Gaußklammer aus Bemerkung 4.34, denn dann ist n_0 eine natürliche Zahl größer als $\frac{1}{\varepsilon}$, und damit wie oben $\frac{1}{n_0} < \varepsilon$. Aber zur Überprüfung der Grenzwertbedingung ist es nicht nötig, ein konkretes n_0 anzugeben – erst recht nicht das *kleinste* (also „beste") mögliche n_0. In der Tat wäre eine solche Bestimmung des kleinstmöglichen n_0 für die allermeisten Folgen auch sehr aufwendig oder sogar gar nicht praktisch durchführbar.

(c) (**Geometrische Folge**) Es sei $q \in \mathbb{R}$ mit $|q| < 1$; wir behaupten, dass dann $\lim\limits_{n\to\infty} q^n = 0$ gilt.

Für $q = 0$ ist dies natürlich klar, da wir dann eine konstante Folge haben. Ansonsten sei wie in (b) wieder $\varepsilon \in \mathbb{R}_{>0}$ beliebig vorgegeben. Wir setzen $x := \frac{1}{|q|} - 1$, also $|q| = \frac{1}{1+x}$; wegen $|q| < 1$ ist natürlich $x > 0$. Nach Bemerkung 4.31 (a) gibt es nun ein $n_0 \in \mathbb{N}$ mit $\frac{1}{n_0} < \varepsilon x$. Es gilt dann für alle $n \geq n_0$

$$\begin{aligned}
|q^n - 0| = |q|^n &\overset{(1)}{=} \frac{1}{(1+x)^n} \\
&\overset{(2)}{\leq} \frac{1}{1+nx} && \text{(mit } x > 0 \text{ nach der Bernoulli-Ungleichung aus Satz 4.20)} \\
&\overset{(3)}{<} \frac{1}{nx} && \text{(wegen } 1 > 0\text{)} \\
&\overset{(4)}{\leq} \frac{1}{n_0 x} && \text{(wegen } n \geq n_0\text{)} \\
&\overset{(5)}{<} \varepsilon, && \left(\text{wegen } \frac{1}{n_0} < \varepsilon x\right)
\end{aligned}$$

woraus die Behauptung folgt.

Bemerkung 5.4 (Rückwärtsrechnen)**.** Wenn ihr euch die Rechnung in Beispiel 5.3 (c) angeschaut habt, werdet ihr vermutlich keine Probleme haben, sie nachzuvollziehen – aber euch sicher auch fragen, wie ihr darauf jemals selbst hättet kommen sollen. Insbesondere die Festlegungen von x und n_0 vor Beginn der Rechnung fallen ja doch sehr vom Himmel.

Die Antwort hierauf ist einfach, dass ich den Beweis zunächst „rückwärts“ durchgeführt habe, bevor ich angefangen habe, ihn aufzuschreiben. Ich habe also mit der Rechnung oben begonnen, bevor ich wusste, was z. B. n_0 später einmal sein würde, und mir etwa folgendes gedacht:

> *Okay, wir müssen sehen, dass $|q^n|$ für $n \to \infty$ kleiner als das gegebene ε wird. Nehmen wir der Einfachheit halber erst einmal $q > 0$ an, dann müssen wir also eine Ungleichungskette $q^n < \cdots < \varepsilon$ finden. Bisher wissen wir nichts darüber, wie sich Potenzen mit wachsendem Exponenten verhalten … aber wir hatten die Bernoulli-Ungleichung $(1+x)^n \geq 1+nx$ gezeigt, die Potenzen durch lineare Funktionen abschätzen kann. Um die anwenden zu können, könnten wir vielleicht $q = 1+x$ setzen? Nein, das hilft nicht, denn dann würde die Ungleichung $q^n = (1+x)^n \geq 1+nx$ ja in die falsche Richtung gehen. Also versuchen wir lieber $q = \frac{1}{1+x}$, das dreht „$\geq$“ zu „$\leq$“ um. Moment, gibt es so ein x überhaupt und erfüllt das die Voraussetzungen der Bernoulli-Ungleichung? Ja, die Gleichung ist ja äquivalent zu $x = \frac{1}{q} - 1$, und es ist $q < 1$, also $x > 0$, das passt. Damit haben wir die Schritte (1) und (2) oben.*
>
> *Jetzt müssen wir also $\frac{1}{1+nx}$ weiter abschätzen und sehen, warum dieser Term gegen 0 geht. Die 1 im Nenner stört. Also, wir könnten sicher auch mit ihr weiter rechnen, aber einfacher wäre der Ausdruck ohne sie. Es ist ja auch $\frac{1}{1+nx} < \frac{1}{nx}$, d. h. die Abschätzung geht in die richtige Richtung, und der neue Ausdruck $\frac{1}{nx}$ geht immer noch gegen 0. Also lassen wir die 1 in (3) einfach weg. Wie wir jetzt weiter machen können, wissen wir aus Beispiel 5.3 (b): Ist nun $n \geq n_0$ und $\frac{1}{n_0} < \varepsilon x$, so erhalten wir in (4) und (5) die gewünschte Abschätzung.*

Nachdem wir diese Überlegungen durchgeführt haben, können wir schließlich noch die Beträge wieder einarbeiten und den Beweis dann so aufschreiben wie oben.

Beachte, dass es natürlich viele verschiedene Arten gibt, derartige Abschätzungen durchzuführen. Aber nicht jede Abschätzung, die richtig ist, ist auch zielführend: So hätten wir z. B. in (3) oben auch versuchen können, den Term nx wegzulassen und die Abschätzung mit $\frac{1}{1+nx} < \frac{1}{1}$ fortzusetzen. Diese Ungleichung ist genauso richtig wie (3), aber der neue Ausdruck $\frac{1}{1} = 1$ geht offensichtlich mit $n \to \infty$ nicht mehr gegen 0, so dass wir die gewünschte Folgerung $\cdots < \varepsilon$ jetzt nicht mehr erreichen können. Man muss beim Abschätzen also stets einen geeigneten Mittelweg finden und aufpassen, dass man weder zu wenig noch zu viel abschätzt. Dadurch erfordern derartige Rechnungen oft eine geschickte und vielleicht nicht ganz offensichtliche Idee. Am Anfang ist das sicher ungewohnt, aber im Laufe der Zeit werdet ihr ein gewisses Gefühl dafür entwickeln, welche Art von Abschätzung in welchen Fällen sinnvoll sein könnte. Aber so oder so – für das reine *Nachvollziehen* einer Abschätzung, die jemand anders gefunden hat (wie z. B. wenn ihr den Beweis in Beispiel 5.3 (c) lest und verstehen wollt), sind solche Ideen natürlich nicht notwendig.

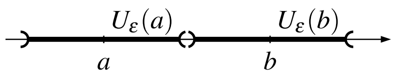
08

Wir wollen nun die bereits in Definition 5.1 versprochene Aussage beweisen, dass der Grenzwert einer Folge (sofern er existiert) immer eindeutig ist. Anschaulich ist diese Aussage natürlich sofort einleuchtend: Es können nicht fast alle Folgenglieder beliebig nahe an zwei verschiedenen Zahlen liegen. Denn wenn wir disjunkte ε-Umgebungen der beiden Grenzwerte wählen, kann jedes Folgenglied natürlich immer nur in einer der beiden Umgebungen liegen – und somit können nicht fast alle in *beiden* Umgebungen liegen. Formal aufgeschrieben sieht diese Beweisidee so aus:

Lemma 5.5 (Eindeutigkeit des Grenzwerts)**.** *Jede Folge hat höchstens einen Grenzwert.*

Beweis. Angenommen, die Aussage wäre falsch, d. h. es gäbe eine Folge $(a_n)_n$ mit $a_n \to a$ und $a_n \to b$ für gewisse $a, b \in \mathbb{R}$ mit $a \neq b$.

Wählen wir dann $\varepsilon := \frac{|a-b|}{2}$ als den halben Abstand zwischen a und b, so gilt wie im Bild rechts $U_\varepsilon(a) \cap U_\varepsilon(b) = \emptyset$, denn aus $x \in U_\varepsilon(a) \cap U_\varepsilon(b)$ würde mit der Dreiecksungleichung der Widerspruch



$$|a-b| = |a-x+x-b| \leq |a-x| + |b-x| < \varepsilon + \varepsilon = 2\varepsilon = |a-b|$$

folgen.

Aber wegen $a_n \to a$ gilt $a_n \in U_\varepsilon(a)$ für fast alle n (also für alle $n \geq n_1$ mit einem gewissen $n_1 \in \mathbb{N}$), und wegen $a_n \to b$ genauso $a_n \in U_\varepsilon(b)$ für fast alle n (also für alle $n \geq n_2$ mit einem gewissen $n_2 \in \mathbb{N}$). Damit folgt auch $a_n \in U_\varepsilon(a) \cap U_\varepsilon(b)$ für fast alle n (nämlich für alle n, bei denen beide Aussagen gelten, also für $n \geq \max(n_1, n_2)$), was ein Widerspruch zu $U_\varepsilon(a) \cap U_\varepsilon(b) = \emptyset$ ist und somit das Lemma beweist. □

Bemerkung 5.6. Wir sehen im Beweis von Lemma 5.5, dass die „fast alle"-Notation den Vorteil hat, dass wir uns oft das explizite Arbeiten mit dem n_0 aus Definition 5.1 (b) (von dem wir ja meistens ohnehin nicht wirklich wissen müssen, welchen Wert es genau hat) sparen können. Die einzige Eigenschaft, die wir hier wirklich gebraucht haben, ist die: Wenn eine Aussage $A(n)$ für fast alle n gilt, und eine weitere Aussage $B(n)$ ebenfalls für fast alle (aber nicht notwendig für die gleichen), dann gelten auch $A(n)$ und $B(n)$ zusammen für fast alle n – nämlich für alle bis auf die endlich vielen Ausnahmen für $A(n)$ und $B(n)$.

Natürlich gibt es auch Folgen ohne Grenzwert. Die einfachste Möglichkeit dafür ist, dass ihre Glieder unbeschränkt wachsen und sich somit keiner Zahl annähern können. Dies wollen wir jetzt formal untersuchen.

Definition 5.7 (Beschränkte Folgen)**.** Eine Folge $(a_n)_n$ heißt **beschränkt**, wenn die Menge ihrer Folgenglieder beschränkt ist, also wenn es ein $s \in \mathbb{R}$ gibt mit $|a_n| \leq s$ für alle $n \in \mathbb{N}$.

Lemma 5.8. *Jede konvergente Folge ist beschränkt.*

Beweis. Es sei $(a_n)_n$ eine konvergente Folge mit Grenzwert a. Dann gibt es zu $\varepsilon = 1$ ein n_0, so dass $|a_n - a| < \varepsilon = 1$ und damit nach der Dreiecksungleichung auch

$$|a_n| = |a_n - a + a| \leq |a_n - a| + |a| < 1 + |a|$$

für alle $n \geq n_0$ gilt. Damit ist dann aber $|a_n| \leq s$ für alle $n \in \mathbb{N}$, wenn wir

$$s := \max(|a_0|, |a_1|, \dots, |a_{n_0-1}|, 1 + |a|)$$

setzen. Also ist $(a_n)_n$ beschränkt. □

Beispiel 5.9.

(a) Die Folge

$$(a_n)_n = (1, 0, 2, 0, 3, 0, \dots)$$

ist unbeschränkt (da die Menge $\mathbb{N}$ ihrer Folgenglieder nach Satz 4.30 unbeschränkt ist) und damit nach Lemma 5.8 divergent.

(b) Auch die geometrische Folge $(q^n)_n$ aus Beispiel 5.3 (c) ist für $|q| > 1$ unbeschränkt und damit divergent: Ist nämlich $s \in \mathbb{R}_{>0}$ beliebig, so können wir nach der archimedischen Ordnung von $\mathbb{R}$ ein $n \in \mathbb{N}$ wählen mit $n > \frac{s}{|q|-1}$, und erhalten mit der Bernoulli-Ungleichung

$$|q|^n = (1+|q|-1)^n \overset{4.20}{\geq} 1+n(|q|-1) > n(|q|-1) > s.$$

Wie rechnet man nun aber Grenzwerte konkret aus, wenn man nicht jedes Mal wieder auf die Definition zurückgehen möchte? Glücklicherweise gibt es dafür die Grenzwertsätze, die besagen, dass man Grenzwerte mit Summen, Differenzen, Produkten und Quotienten vertauschen kann, und die damit oft eine konkrete Berechnung ermöglichen. Zum Beweis dieser Aussage benötigen wir zunächst ein Lemma.

Definition 5.10 (Nullfolgen)**.** Eine Folge heißt **Nullfolge**, wenn sie gegen 0 konvergiert. Offensichtlich konvergiert eine Folge $(a_n)_n$ damit nach Definition genau dann gegen $a \in \mathbb{R}$, wenn $(a_n - a)_n$ eine Nullfolge ist.

Lemma 5.11. *Ist $(a_n)_n$ eine beschränkte Folge und $(b_n)_n$ eine Nullfolge, so ist auch $(a_n b_n)_n$ eine Nullfolge.*

Beweis. Es sei $\varepsilon \in \mathbb{R}_{>0}$ beliebig. Da $(a_n)_n$ beschränkt ist, gilt $|a_n| \leq s$ für ein $s \in \mathbb{R}_{>0}$ und alle $n \in \mathbb{N}$. Da $(b_n)_n$ eine Nullfolge ist, ist weiterhin $|b_n| < \frac{\varepsilon}{s}$ für fast alle n. Also gilt für fast alle n auch die Abschätzung $|a_n b_n| = |a_n| \cdot |b_n| < s \cdot \frac{\varepsilon}{s} = \varepsilon$, d. h. $(a_n b_n)_n$ ist eine Nullfolge. □

Bemerkung 5.12 (Folgen mit Grenzwert ungleich 0)**.** Es sei $(a_n)_n$ eine Folge, die gegen einen Grenzwert $a \neq 0$ konvergiert. Eine unmittelbare, aber dennoch oft nützliche Folgerung aus der Grenzwertdefinition 5.1 ergibt sich, wenn wir dort $\varepsilon = \frac{|a|}{2} > 0$ setzen: Für fast alle n ist dann $a_n \in U_\varepsilon(a)$, mit der Dreiecksungleichung nach unten aus Bemerkung 4.19 also

$$|a_n| \geq |a| - |a - a_n| > |a| - \varepsilon = \frac{|a|}{2}.$$

Hat eine Folge also einen Grenzwert $a \neq 0$, so sind insbesondere auch fast alle Folgenglieder ungleich 0 (und betragsmäßig sogar größer als $\frac{|a|}{2}$).

Satz 5.13 (**Grenzwertsätze** für Folgen)**.** *Es seien $(a_n)_n$ und $(b_n)_n$ zwei konvergente Folgen mit $a_n \to a$ und $b_n \to b$. Dann gilt:*

(a) *$a_n + b_n \to a + b$ und $a_n - b_n \to a - b$.*

(b) *$a_n b_n \to ab$.*

(c) *Ist $b \neq 0$, so sind auch fast alle $b_n \neq 0$, und es gilt $\frac{a_n}{b_n} \to \frac{a}{b}$.*

Beweis.

(a) Es sei $\varepsilon \in \mathbb{R}_{>0}$ beliebig. Wegen $a_n \to a$ und $b_n \to b$ gilt $|a_n - a| < \frac{\varepsilon}{2}$ und $|b_n - b| < \frac{\varepsilon}{2}$ für fast alle n. Damit folgt für fast alle n (siehe Bemerkung 5.6) mit der Dreiecksungleichung

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

also wie behauptet $a_n + b_n \to a + b$. Die Aussage über die Differenz der Grenzwerte folgt natürlich genauso.

(b) Für alle $n \in \mathbb{N}$ gilt zunächst

$$a_n b_n - ab = a_n b_n - a_n b + a_n b - ab = a_n(b_n - b) + b(a_n - a). \tag{1}$$

Die Folge $(a_n)_n$ ist nach Voraussetzung konvergent und damit beschränkt nach Lemma 5.8. Weiterhin ist $b_n - b$ eine Nullfolge wegen $b_n \to b$. Also ist auch $(a_n(b_n - b))_n$, d. h. der erste Summand rechts in (1), nach Lemma 5.11 eine Nullfolge. Genauso ergibt sich, dass auch der zweite Summand $(b(a_n - a))_n$ eine Nullfolge ist. Damit ist (1) die Summe zweier Nullfolgen, nach (a) also ebenfalls eine Nullfolge. Dies zeigt $a_n b_n - ab \to 0$ und damit $a_n b_n \to ab$.

(c) Nach Bemerkung 5.12 sind mit $b \neq 0$ auch fast alle b_n ungleich 0, so dass wir (nach evtl. Weglassen endlich vieler Glieder) die Quotientenfolge $\left(\frac{a_n}{b_n}\right)_n$ betrachten können. Weil nach derselben Bemerkung dann sogar $|b_n| > \frac{|b|}{2}$ und damit $\left|\frac{1}{b_n}\right| < \frac{2}{|b|}$ gilt, ist die Folge $\left(\frac{1}{b_n}\right)_n$ außerdem beschränkt. Schreiben wir also

$$\frac{a_n}{b_n} - \frac{a}{b} = \frac{a_n b - a b_n}{b b_n} = \frac{a_n b - ab + ab - a b_n}{b b_n} = \frac{1}{b_n}(a_n - a) + \frac{a}{b b_n}(b - b_n), \tag{2}$$

so ergibt sich die Behauptung genauso wie in (b): $\left(\frac{1}{b_n}(a_n - a)\right)_n$ ist eine Nullfolge (nach Lemma 5.11 als Produkt der beschränkten Folge $\left(\frac{1}{b_n}\right)_n$ mit der Nullfolge $(a_n - a)_n$), analog ist auch $\left(\frac{a}{b b_n}(b - b_n)\right)_n$ eine Nullfolge. Damit ist (2) wieder die Summe zweier Nullfolgen, nach (a) also ebenfalls eine Nullfolge – woraus $\frac{a_n}{b_n} \to \frac{a}{b}$ folgt. □

Beispiel 5.14 (Grenzwerte von Quotienten von Polynomen)**.** Wollen wir den Grenzwert der Folge $\left(\frac{2n^2}{n^2+1}\right)_n$ bestimmen, so können wir nicht direkt die Grenzwertsätze anwenden, da Zähler und Nenner natürlich unbeschränkt sind und damit nach Lemma 5.8 divergieren. Durch Kürzen mit n^2 können wir die Folgenglieder aber umschreiben, so dass wir den Grenzwert dann mit Satz 5.13 in den Quotienten, die Summe und das Produkt hineinziehen können und (mit Beispiel 5.3)

$$\frac{2n^2}{n^2+1} = \frac{2}{1+\frac{1}{n^2}} = \frac{2}{1+\frac{1}{n}\cdot\frac{1}{n}} \to \frac{2}{1+0\cdot 0} = 2 \quad \text{für } n \to \infty$$

erhalten. Auf die gleiche Art kann man offensichtlich den Grenzwert jeder Folge berechnen, die ein Quotient von zwei Polynomfunktionen in n ist, indem man zuerst mit der höchsten auftretenden Potenz von n kürzt.

Bemerkung 5.15. Beachte, dass Satz 5.13 nur angewendet werden kann, wenn beide Grenzwerte $\lim\limits_{n\to\infty} a_n$ und $\lim\limits_{n\to\infty} b_n$ existieren – ansonsten macht der Satz keine Aussage. Eine Rechnung hinzuschreiben wie z. B.

$$\lim_{n\to\infty} \frac{n}{n+1} \cdot \frac{n+2}{n+3} \overset{(*)}{=} \lim_{n\to\infty} \frac{n}{n+1} \cdot \lim_{n\to\infty} \frac{n+2}{n+3} = \lim_{n\to\infty} \frac{1}{1+\frac{1}{n}} \cdot \lim_{n\to\infty} \frac{1+\frac{2}{n}}{1+\frac{3}{n}} = 1 \cdot 1 = 1$$

(mit Verweis an der Stelle $(*)$ auf den Grenzwertsatz 5.13 (b)) ist daher eigentlich *nicht korrekt*, da wir bei $(*)$ ja noch nicht überprüft haben, ob die Grenzwerte der beiden einzelnen Brüche auch wirklich existieren. Man müsste also theoretisch zuerst die Grenzwerte von $\frac{n}{n+1}$ und $\frac{n+2}{n+3}$ separat berechnen (bzw. ihre Existenz zeigen), und könnte dann erst die obige Rechnung $(*)$ hinschreiben. Da dies aber deutlich mehr Schreibaufwand wäre und die Darstellung auch unübersichtlicher machen würde, wollen wir vereinbaren, dass wir die Grenzwertsätze in einer Rechnung wie oben auch schon benutzen dürfen, wenn wir erst *nachträglich* überprüfen, dass die Einzelgrenzwerte existieren.

Aufgabe 5.16. Bestimme die Grenzwerte (sofern sie existieren)

$$\text{(a)} \ \lim_{n\to\infty} \frac{n^2+2}{n^2+n}, \qquad \text{(b)} \ \lim_{n\to\infty} \frac{n^3}{\binom{n}{3}}, \qquad \text{(c)} \ \lim_{n\to\infty} \frac{3^n}{n^3}.$$

Für (a) beweise man diesen Grenzwert zusätzlich direkt nach Definition, d. h. man gebe zu jedem $\varepsilon \in \mathbb{R}_{>0}$ ein $n_0 \in \mathbb{N}$ an, so dass die Grenzwertbedingung aus Definition 5.1 (b) gilt.

Aufgabe 5.17. Zu einer gegebenen Folge $(a_n)_{n\in\mathbb{N}_{>0}}$ definieren wir die Folge $(b_n)_{n\in\mathbb{N}_{>0}}$ ihrer Mittelwerte durch

$$b_n := \frac{a_1 + \cdots + a_n}{n} \quad \text{für alle } n \in \mathbb{N}_{>0}.$$

Man zeige: Ist $(a_n)_n$ konvergent mit Grenzwert $a \in \mathbb{R}$, so ist auch $(b_n)_n$ konvergent mit demselben Grenzwert a. (Hinweis: Zur Vereinfachung der Rechnung ist es nützlich, die Aussage zunächst für eine Nullfolge $(a_n)_n$ zu beweisen, und den allgemeinen Fall dann darauf zurückzuführen.)

Als Beispiele für divergente Folgen haben wir bisher nur unbeschränkte Folgen gesehen. Aber auch beschränkte Folgen können natürlich divergent sein, wie z. B. die Folge

$$\big((-1)^n\big)_{n\in\mathbb{N}} = (1,-1,1,-1,\dots),$$

in der alle geraden Folgenglieder gleich 1 und alle ungeraden gleich -1 sind, so dass für die gesamte Folge kein Grenzwert existieren kann. Formal können wir dies mit dem Begriff der Teilfolgen und Häufungspunkte ausdrücken.

Definition 5.18 (Umordnungen, Teilfolgen und Häufungspunkte)**.** Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge.

(a) Eine **Umordnung** von $(a_n)_{n\in\mathbb{N}}$ ist eine Folge der Form $(a_{\sigma(0)}, a_{\sigma(1)}, a_{\sigma(2)}, \dots) = (a_{\sigma(n)})_{n\in\mathbb{N}}$ für eine bijektive Abbildung $\sigma\colon \mathbb{N} \to \mathbb{N}$. Sie entsteht also einfach durch eine beliebige Permutation aller Folgenglieder.

(b) Eine **Teilfolge** von $(a_n)_{n\in\mathbb{N}}$ ist eine Folge der Form $(a_{n_0}, a_{n_1}, a_{n_2}, \dots) = (a_{n_k})_{k\in\mathbb{N}}$ für gewisse $n_0 < n_1 < n_2 < \cdots$, also eine Folge, die aus $(a_n)_{n\in\mathbb{N}}$ durch Auswählen bestimmter Folgenglieder unter Beibehaltung ihrer Reihenfolge entsteht.

(c) Eine Zahl $a \in \mathbb{R}$ heißt **Häufungspunkt** von $(a_n)_{n\in\mathbb{N}}$, wenn es eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$ gibt, die gegen a konvergiert.

Lemma 5.19 (Grenzwerte von Umordnungen und Teilfolgen)**.** *Konvergiert eine Folge $(a_n)_n$ gegen einen Grenzwert a, so konvergiert auch jede Umordnung und jede Teilfolge von $(a_n)_n$ gegen a.*

Insbesondere hat eine konvergente Folge also genau einen Häufungspunkt, nämlich ihren Grenzwert.

Beweis. Es sei $\varepsilon \in \mathbb{R}_{>0}$ beliebig. Da die Folge $(a_n)_n$ gegen a konvergiert, hat sie nur endlich viele Glieder, die außerhalb von $U_\varepsilon(a)$ liegen. Jede Umordnung oder Teilfolge von $(a_n)_n$ hat damit aber ebenfalls nur endlich viele Glieder außerhalb von $U_\varepsilon(a)$, und somit konvergiert eine solche Umordnung oder Teilfolge ebenfalls gegen a. □

Beispiel 5.20.

(a) Die oben betrachtete Folge $(a_n)_{n\in\mathbb{N}} = ((-1)^n)_{n\in\mathbb{N}} = (1,-1,1,-1,\dots)$ besitzt die beiden konstanten Teilfolgen

$$(a_{2n})_{n\in\mathbb{N}} = (1,1,1,\dots)$$
$$\text{und}\qquad (a_{2n+1})_{n\in\mathbb{N}} = (-1,-1,-1,\dots)$$

und damit die beiden Häufungspunkte 1 und -1. Sie ist also nach Lemma 5.19 divergent. In der Tat werden wir in Beispiel 5.23 (a) sehen, dass 1 und -1 auch die einzigen Häufungspunkte von $(a_n)_n$ sind.

(b) Für die Folge $(a_n)_n = (n+1)_n = (1,2,3,\dots)$ ist jede ihrer Teilfolgen unbeschränkt und damit nach Lemma 5.8 divergent. Also besitzt $(a_n)_n$ keine Häufungspunkte.

09

Zur konkreten Berechnung von Häufungspunkten sind oft die folgenden beiden Lemmata nützlich.

Lemma 5.21 (Äquivalente Charakterisierung von Häufungspunkten)**.** *Eine Zahl $a \in \mathbb{R}$ ist genau dann ein Häufungspunkt einer Folge $(a_n)_n$, wenn in jeder ε-Umgebung von a unendlich viele Folgenglieder a_n liegen.*

Beweis.

„$\Rightarrow$“: Konvergiert eine Teilfolge von $(a_n)_n$ gegen a, so liegen in jeder ε-Umgebung von a fast alle Glieder der Teilfolge und somit insbesondere auch unendlich viele Glieder von $(a_n)_n$.

„$\Leftarrow$“: Wir konstruieren eine Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ der gewünschten Art wie folgt: Als Startindex nehmen wir $n_0 = 0$. Ist nun für ein $k \in \mathbb{N}_{>0}$ der Index n_{k-1} bereits konstruiert, so wählen wir ein $n_k > n_{k-1}$ mit $|a_{n_k} - a| < \frac{1}{k}$ (dies ist möglich, da in der $\frac{1}{k}$-Umgebung von a nach Voraussetzung unendlich viele Folgenglieder liegen, also auch eines hinter $a_{n_{k-1}}$).

Die so konstruierte Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ konvergiert dann gegen a: Ist $\varepsilon \in \mathbb{R}_{>0}$ gegeben und $k_0 \in \mathbb{N}_{>0}$ mit $\frac{1}{k_0} < \varepsilon$, so ist $|a_{n_k} - a| < \frac{1}{k} \le \frac{1}{k_0} < \varepsilon$ für alle $k \ge k_0$. □

Lemma 5.22 (Mischfolgen)**.** *Es seien $(a_n)_{n\in\mathbb{N}}$ eine Folge sowie $(a_{n_k})_{k\in\mathbb{N}}$ und $(a_{m_k})_{k\in\mathbb{N}}$ zwei Teilfolgen, die zusammen die gesamte Folge ergeben, also so dass $\{n_k : k \in \mathbb{N}\} \cup \{m_k : k \in \mathbb{N}\} = \mathbb{N}$ gilt. Man sagt in diesem Fall auch, dass $(a_n)_{n\in\mathbb{N}}$ eine **Mischfolge** von $(a_{n_k})_{k\in\mathbb{N}}$ und $(a_{m_k})_{k\in\mathbb{N}}$ ist.*

Dann ist eine Zahl $a \in \mathbb{R}$ genau dann ein Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$, wenn a ein Häufungspunkt von $(a_{n_k})_{k\in\mathbb{N}}$ oder $(a_{m_k})_{k\in\mathbb{N}}$ ist.

Beweis. Eine Zahl a ist nach Lemma 5.21 genau dann ein Häufungspunkt von $(a_n)_n$, wenn in jeder ε-Umgebung von a unendlich viele Folgenglieder von $(a_n)_n$ liegen. Da $(a_n)_n$ eine Mischfolge von $(a_{n_k})_k$ und $(a_{m_k})_k$ ist, ist dies natürlich äquivalent dazu, dass in jeder solchen ε-Umgebung unendlich viele Folgenglieder von $(a_{n_k})_k$ oder $(a_{m_k})_k$ liegen, also dass a ein Häufungspunkt (mindestens) einer dieser beiden Teilfolgen ist. □

Beispiel 5.23.

(a) Die Folge $(a_n)_n = ((-1)^n)_{n\in\mathbb{N}}$ aus Beispiel 5.20 (a) ist eine Mischfolge ihrer positiven und negativen Glieder, die konstant gleich 1 bzw. -1 sind. Aus Lemma 5.22 folgt also, dass $(a_n)_n$ genau die beiden Häufungspunkte 1 und -1 hat.

(b) Die divergente Folge $(a_n)_n = (1,0,2,0,3,0,\dots)$ aus Beispiel 5.9 (a) ist eine Mischfolge der Folge $(1,2,3,\dots)$ (die nach Beispiel 5.20 (b) keine Häufungspunkte hat) und der konstanten Folge 0 (die natürlich 0 als einzigen Häufungspunkt hat). Damit hat $(a_n)_n$ nach Lemma 5.22 den einzigen Häufungspunkt 0. Wir sehen also (im Gegensatz zu Lemma 5.19), dass eine Folge mit genau einem Häufungspunkt nicht notwendig konvergiert.

5.B Konvergenzkriterien für Folgen

Nicht in allen Fällen lässt sich die Berechnung von Grenzwerten mit Hilfe der Grenzwertsätze auf bereits bekannte zurückführen. Wir benötigen daher noch weitere Techniken zur Grenzwertbestimmung und beginnen mit einem einfachen Vergleichskriterium.

Satz 5.24 (Verträglichkeit des Grenzwerts mit Ungleichungen)**.** *Es seien $(a_n)_n$ und $(b_n)_n$ konvergente Folgen mit $a_n \to a$ und $b_n \to b$. Dann gilt:*

(a) *Ist $a_n \le b_n$ für fast alle n, so auch $a \le b$.*

(b) *(**Einschachtelungssatz**) Ist $a = b$, konvergieren also beide Folgen gegen denselben Grenzwert, und ist $(c_n)_n$ eine weitere reelle Folge mit $a_n \le c_n \le b_n$ für fast alle n, so konvergiert auch $(c_n)_n$ gegen diesen Grenzwert.*

Beweis.

(a) Angenommen, es wäre $a > b$. Wir setzen $\varepsilon := \frac{a-b}{2}$. Wegen $a_n \to a$ und $b_n \to b$ wäre dann (nach Bemerkung 5.6) für fast alle n

$$a_n \in (a-\varepsilon, a+\varepsilon) \quad \text{und} \quad b_n \in (b-\varepsilon, b+\varepsilon).$$

Zusammensetzen liefert $a-\varepsilon < a_n \le b_n < b+\varepsilon$ für fast alle n, und damit $a-b < 2\varepsilon$ im Widerspruch zu $\varepsilon = \frac{a-b}{2}$.

(b) Es sei $\varepsilon \in \mathbb{R}_{>0}$ beliebig. Diesmal gilt wegen $a_n \to a$ und $b_n \to a$ für fast alle n

$$a_n \in (a-\varepsilon, a+\varepsilon) \quad \text{und} \quad b_n \in (a-\varepsilon, a+\varepsilon),$$

und damit $a-\varepsilon < a_n \le c_n \le b_n < a+\varepsilon$, also $c_n \in (a-\varepsilon, a+\varepsilon)$. Da $\varepsilon > 0$ beliebig war, folgt daraus wie behauptet $c_n \to a$. □

Bemerkung 5.25. Beachte, dass Satz 5.24 (a) *nicht* auch analog für die echte Ungleichung „<“ gilt: Ist z. B. $a_n = 0$ und $b_n = \frac{1}{n}$ für alle $n \ge 1$, so gilt zwar $a_n < b_n$ für alle n, aber die Grenzwerte beider Folgen sind natürlich gleich 0, d. h. es gilt nur $\lim\limits_{n\to\infty} a_n \le \lim\limits_{n\to\infty} b_n$ gemäß Satz 5.24 (a), aber nicht $\lim\limits_{n\to\infty} a_n < \lim\limits_{n\to\infty} b_n$.

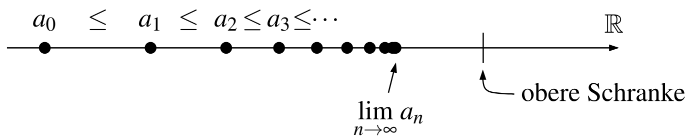
Aufgabe 5.26. Bestimme den Grenzwert

$$\lim_{n\to\infty} \sum_{k=1}^{n} \frac{n}{n^2+k}.$$

Alle unsere bisherigen Kriterien haben den entscheidenden Nachteil, dass sie Grenzwerte nur auf andere bereits bekannte zurückführen können. In der Praxis werden aber viele Größen wie z. B. Quadratwurzeln, π, e oder der Sinus und Kosinus einer gegebenen Zahl überhaupt erst als Grenzwerte geeigneter Folgen konstruiert. Die Konvergenz solcher Folgen werden wir also mit unseren bisherigen Methoden nie nachweisen können.

Wir benötigen daher auch Kriterien, mit denen man die Konvergenz einer Folge selbst dann nachweisen kann, wenn man ihren Grenzwert noch nicht vorher kennt oder gleichzeitig aus bereits bekannten anderen Grenzwerten berechnen kann. Im Gegensatz zu unseren Ergebnissen aus Abschnitt 5.A, die unverändert auch in $\mathbb{Q}$ gelten würden, handelt es sich hierbei nun um Resultate, die ganz zentral das Supremumsaxiom verwenden und daher nur in $\mathbb{R}$ gelten.

Das erste Kriterium dieser Art, das wir jetzt behandeln wollen, ist für Folgen anwendbar, deren Folgenglieder mit wachsendem n immer größer werden. Ist eine solche Folge nach oben beschränkt, so ist anschaulich klar, dass die Folgenglieder wie im Bild unten für wachsendes n „immer näher zusammen rücken“ müssen, was letztlich zur Konvergenz der Folge führen sollte. Dies ist der Inhalt des folgenden Satzes.



Definition 5.27 (Monotone und beschränkte Folgen)**.** Es sei $(a_n)_n$ eine Folge.

(a) Die Folge $(a_n)_n$ heißt **monoton wachsend** oder **steigend**, wenn $a_n \leq a_{n+1}$ für alle $n \in \mathbb{N}$, also $a_0 \leq a_1 \leq a_2 \leq \cdots$ und damit $a_m \leq a_n$ für alle $m \leq n$ gilt. Gilt sogar $a_n < a_{n+1}$ für alle $n \in \mathbb{N}$, so heißt $(a_n)_n$ **streng monoton wachsend** oder **steigend**.

Analog heißt $(a_n)_n$ **(streng) monoton fallend**, wenn $a_n \geq a_{n+1}$ (bzw. $a_n > a_{n+1}$) für alle $n \in \mathbb{N}$ gilt.

(b) Analog zu Definition 5.7 heißt $(a_n)_n$ **nach oben beschränkt**, wenn die Menge ihrer Folgenglieder nach oben beschränkt ist, also wenn es ein $s \in \mathbb{R}$ gibt mit $a_n \leq s$ für alle n.

Analog heißt $(a_n)_n$ **nach unten beschränkt**, wenn es ein $s \in \mathbb{R}$ gibt mit $a_n \geq s$ für alle n; die Folge ist also genau dann beschränkt, wenn sie nach oben und unten beschränkt ist.

Satz 5.28 (**Monotoniekriterium**)**.** *Jede monoton wachsende, nach oben beschränkte Folge $(a_n)_n$ in $\mathbb{R}$ ist konvergent. (Analog ist dann natürlich auch jede monoton fallende, nach unten beschränkte Folge konvergent.)*

Beweis. Da die Menge $M := \{a_n : n \in \mathbb{N}\} \subset \mathbb{R}$ aller Folgenglieder nicht leer und nach oben beschränkt ist, existiert $a := \sup M$ nach dem Supremumsaxiom. Wir behaupten, dass $a_n \to a$.

Es sei dazu $\varepsilon \in \mathbb{R}_{>0}$ beliebig. Da a die kleinste obere Schranke für M ist, ist $a - \varepsilon$ keine obere Schranke mehr. Es gibt also ein $n_0 \in \mathbb{N}$ mit $a_{n_0} > a - \varepsilon$. Für alle $n \geq n_0$ folgt dann

$$\begin{aligned} a - \varepsilon < a_{n_0} &\leq a_n && \text{(Monotonie)} \\ &\leq a && \text{(a ist obere Schranke der Folgenglieder)} \\ &< a + \varepsilon, \end{aligned}$$

also $|a_n - a| < \varepsilon$. Damit konvergiert $(a_n)_n$ gegen a. □

Als Beispiel für die Anwendung des Monotoniekriteriums wollen wir nun als Erstes zeigen, dass jede nicht-negative reelle Zahl eine Quadratwurzel besitzt. Auch wenn euch diese Tatsache aus der

Schule vielleicht „offensichtlich" erscheint (und wir sie in Bemerkung 4.37 auch schon ohne Beweis benutzt haben), folgt sie dennoch nicht unmittelbar aus den Axiomen für $\mathbb{R}$ und muss damit bewiesen werden. Eine Möglichkeit dafür ist, eine Folge $(a_n)_n$ zu konstruieren, deren Konvergenz wir beweisen können, und deren Grenzwert nur die gewünschte Wurzel sein kann. Die Konstruktion von $(a_n)_n$ ist dabei *rekursiv*, d. h. wir können (analog zur vollständigen Induktion) nicht direkt a_n für alle $n \in \mathbb{N}$ angeben, sondern legen nur das erste Folgenglied a_0 fest und geben dann eine Formel an, mit der für alle $n \in \mathbb{N}$ aus a_n das nächste Folgenglied a_{n+1} berechnet werden kann.

Lemma 5.29 (Existenz von Wurzeln)**.** *Es seien $c \in \mathbb{R}_{>0}$ und $a_0 \in \mathbb{R}_{>0}$ gegeben. Dann konvergiert die mit diesem Startwert a_0 rekursiv definierte Folge $(a_n)_n$ mit*

$$a_{n+1} := \frac{1}{2}\left(a_n + \frac{c}{a_n}\right) \quad \textit{für alle } n \in \mathbb{N} \tag{$*$}$$

gegen ein $a \in \mathbb{R}_{>0}$ mit $a^2 = c$.

Beweis. Wir zeigen die Konvergenz von $(a_n)_n$ mit dem Monotoniekriterium, indem wir nachweisen, dass die Folge nach unten beschränkt und monoton fallend ist.

Aus der Rekursionsvorschrift $(*)$ ist offensichtlich, dass mit c und a_0 auch alle Folgenglieder positiv sind. Die Folge ist also sicher durch 0 nach unten beschränkt. In der Tat gilt für alle $n \in \mathbb{N}$ sogar

$$a_{n+1}^2 = \frac{1}{4}\left(a_n + \frac{c}{a_n}\right)^2 = \frac{1}{4}a_n^2 + \frac{1}{2}c + \frac{1}{4}\frac{c^2}{a_n^2} = \frac{1}{4}a_n^2 - \frac{1}{2}c + \frac{1}{4}\frac{c^2}{a_n^2} + c = \frac{1}{4}\left(a_n - \frac{c}{a_n}\right)^2 + c \geq c$$

und somit $a_n^2 \geq c$ für alle $n \geq 1$. Daraus folgt für alle $n \geq 1$ aber auch

$$\frac{a_{n+1}}{a_n} = \frac{1}{2}\left(1 + \frac{c}{a_n^2}\right) \leq \frac{1}{2}\left(1 + \frac{c}{c}\right) = 1,$$

und damit $a_{n+1} \leq a_n$, die Folge ist also (mit Ausnahme evtl. des ersten Folgengliedes) monoton fallend. Damit konvergiert $(a_n)_n$ nach Satz 5.28, d. h. der Grenzwert $a := \lim_{n\to\infty} a_n$ existiert.

Um Informationen über den Grenzwert a zu bekommen, multiplizieren wir die Rekursionsgleichung $(*)$ zunächst mit a_n und erhalten $a_{n+1}a_n = \frac{1}{2}(a_n^2 + c)$ für alle $n \in \mathbb{N}$. Gehen wir in dieser Gleichung nun zum Grenzwert über, so ergibt sich

$$\lim_{n\to\infty} a_{n+1}a_n = \lim_{n\to\infty} \frac{1}{2}\left(a_n^2 + c\right), \quad \text{und damit} \quad a^2 = \frac{1}{2}\left(a^2 + c\right),$$

da die Folge $(a_{n+1})_n = (a_1, a_2, a_3, \dots)$ gegenüber $(a_n)_n = (a_0, a_1, a_2, \dots)$ ja nur um ein Folgenglied verschoben ist und somit als Teilfolge von $(a_n)_n$ ebenfalls gegen a konvergiert. Auflösen dieser Gleichung liefert nun sofort wie behauptet $a^2 = c$. Da alle a_n positiv sind, ergibt Satz 5.24 (a) außerdem auch $a \geq 0$; in der Tat ist wegen $a^2 = c > 0$ dann sogar $a > 0$. □

Folgerung und Definition 5.30. *Zu jedem $c \in \mathbb{R}_{\geq 0}$ gibt es eine eindeutig bestimmte Zahl $a \in \mathbb{R}_{\geq 0}$ mit $a^2 = c$. Wir nennen sie die **(Quadrat-)Wurzel** aus c und schreiben sie als $\sqrt{c}$.*

Beweis. Für $c = 0$ ist die Aussage mit $a = 0$ klar, daher können wir im Folgenden $c > 0$ annehmen. Die Existenz einer Wurzel a folgt dann direkt aus Lemma 5.29. Das Polynom $x \mapsto x^2 - c$ hat nun die positive Nullstelle a und die negative Nullstelle $-a$, und kann als Polynom vom Grad 2 nach Satz 3.19 (b) keine weiteren Nullstellen haben. Also ist die Wurzel a auch eindeutig bestimmt. □

Bemerkung 5.31 (Grenzwert rekursiver Folgen)**.** Das Monotoniekriterium bietet sich oft für rekursiv definierte Folgen $(a_n)_n$ wie in Lemma 5.29 an, da die Monotonie ja durch den Vergleich von a_{n+1} und a_n nachgewiesen werden kann. Sehr nützlich ist dabei auch der Trick, wie im Beweis des Lemmas in der Rekursionsgleichung zum Grenzwert überzugehen, um eine bestimmende Gleichung für den Grenzwert zu finden.

Beispiel 5.32. Die folgende Tabelle zeigt den Anfang der Folge aus Lemma 5.29 im Fall $c = 2$ und $a_0 = 1$. Beachte, dass die Folge „extrem schnell" konvergiert und daher sehr gut zur näherungsweisen Berechnung von Wurzeln geeignet ist – z. B. wenn man einem Computer, der bisher nur weiß, wie man die vier Grundrechenarten ausführt, das Wurzelziehen beibringen möchte. In der Tat kann man zeigen, dass jeder Schritt die Anzahl der korrekten Dezimalstellen (die in der Tabelle unten fett gedruckt sind) näherungsweise verdoppelt. Wir werden uns in dieser Vorlesung jedoch nicht weiter mit der „Geschwindigkeit" der Konvergenz von Folgen beschäftigen; derartige Fragestellungen werdet ihr später in den Vorlesungen zur Praktischen Mathematik untersuchen.

n	a_n
0	**1,**000000000000000000000000000000000000000000000000...
1	**1,5**00000000000000000000000000000000000000000000000...
2	**1,41**6666666666666666666666666666666666666666666666...
3	**1,41421**5686274509803921568627450980392156862745098...
4	**1,4142135623**746899106262955788901349101165596221 15...
5	**1,41421356237309504880168**9623502530243614981925776...
6	**1,414213562373095048801688724209698078569671875377**...

Aufgabe 5.33. Untersuche, ob die folgenden rekursiv definierten Folgen $(a_n)_n$ konvergieren, und bestimme im Fall der Konvergenz den Grenzwert.

(a) $a_0 = 1,\ a_n = \sqrt{a_{n-1} + 6}$ für alle $n \geq 1$;

(b) $a_0 = 1,\ a_1 = 2,\ a_n = \frac{1}{2}\,(a_{n-1} + a_{n-2})$ für alle $n \geq 2$.

(Hinweis: Leite zunächst eine Formel für die Differenz $a_n - a_{n-1}$ zweier aufeinander folgender Glieder her.)

Aufgabe 5.34. Es seien $a, b \in \mathbb{R}_{\geq 0}$. Zeige, dass die Quadratwurzel die folgenden Eigenschaften hat:

(a) Gilt $a \leq b$, so auch $\sqrt{a} \leq \sqrt{b}$ (Monotonie).

(b) $\sqrt{ab} = \sqrt{a} \cdot \sqrt{b}$.

(c) $\sqrt{ab} \leq \frac{a+b}{2}$.

(d) Konvergiert eine Folge $(a_n)_n$ in $\mathbb{R}_{\geq 0}$ gegen a, so gilt auch $\lim\limits_{n\to\infty} \sqrt{a_n} = \sqrt{a}$.

Aufgabe 5.35.

(a) Man zeige: Für alle $a, b \in \mathbb{Z}$ und $k, n \in \mathbb{N}$ ist $(a + b\sqrt{k})^n + (a - b\sqrt{k})^n \in \mathbb{Z}$.

(b) Bestimme die 100. Nachkommastelle (im Dezimalsystem) von $(2 + \sqrt{5})^{2023}$.

Aufgabe 5.36. Zeige, dass jedes (nicht-leere) offene Intervall in $\mathbb{R}$ unendlich viele rationale und unendlich viele irrationale Zahlen enthält.

Aufgabe 5.37 (Existenz höherer Wurzeln in $\mathbb{R}$)**.** In dieser Aufgabe wollen wir analog zu Lemma 5.29 und Folgerung 5.30 beweisen, dass jede nicht-negative reelle Zahl c für alle $k \in \mathbb{N}_{>0}$ eine eindeutige k-te Wurzel besitzt. Wir definieren dazu für ein gegebenes $c \in \mathbb{R}_{>0}$, $k \in \mathbb{N}_{>0}$ und einen beliebigen positiven Startwert a_0 die Folge $(a_n)_n$ rekursiv durch

$$a_{n+1} := \frac{1}{k}\left((k-1)\,a_n + \frac{c}{a_n^{k-1}}\right) \qquad \text{für alle } n \in \mathbb{N}.$$

Man beweise nun:

(a) Für alle $n \in \mathbb{N}$ ist $a_{n+1}^k \geq c$.

(b) Die Folge $(a_n)_n$ ist ab dem zweiten Folgenglied monoton fallend.

(c) Zu jeder Zahl $c \in \mathbb{R}_{\geq 0}$ gibt es ein eindeutiges $a \in \mathbb{R}_{\geq 0}$ mit $a^k = c$. Wir nennen dieses a die k-te **Wurzel** aus c und schreiben sie als $\sqrt[k]{c}$.

Aufgabe 5.38. Es sei $M \subset \mathbb{R}$ nicht leer und nach oben beschränkt. Zeige, dass für $s \in \mathbb{R}$ die folgenden Aussagen äquivalent sind:

(a) $s = \sup M$.

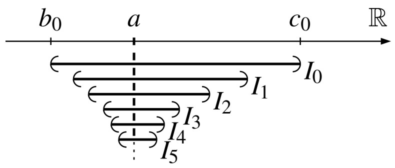
(b) s ist eine obere Schranke für M, und es gibt eine Folge $(a_n)_n$ von Elementen aus M mit $\lim\limits_{n\to\infty} a_n = s$.

10

Kombiniert man die Monotoniekriterien für wachsende und fallende Folgen miteinander, kann man einen Grenzwert wie folgt von beiden Seiten einschachteln.

Satz 5.39 (**Intervallschachtelung**)**.** *Gegeben sei für alle $n \in \mathbb{N}$ ein abgeschlossenes Intervall $I_n = [b_n, c_n]$ in $\mathbb{R}$, so dass $I_0 \supset I_1 \supset I_2 \supset \cdots$ (also die Intervalle ineinander liegen) und $\lim\limits_{n\to\infty}(c_n - b_n) = 0$ (also die Längen der Intervalle gegen 0 konvergieren).*

Dann gibt es genau ein $a \in \mathbb{R}$, das in allen diesen Intervallen liegt, und es gilt $b_n \to a$ und $c_n \to a$.



Beweis. Die Folge $(b_n)_n$ der unteren Intervallgrenzen ist monoton wachsend und nach oben beschränkt (z. B. durch c_0), nach dem Monotoniekriterium aus Satz 5.28 also konvergent. Genauso ist $(c_n)_n$ monoton fallend und nach unten beschränkt, und damit ebenfalls konvergent. Da die Längen der Intervalle nach Voraussetzung gegen 0 konvergieren, folgt nach Satz 5.13 also

$$\lim_{n\to\infty} c_n - \lim_{n\to\infty} b_n = \lim_{n\to\infty}(c_n - b_n) = 0 \quad \Rightarrow \quad \lim_{n\to\infty} b_n = \lim_{n\to\infty} c_n.$$

Es sei $a := \lim\limits_{n\to\infty} b_n = \lim\limits_{n\to\infty} c_n$ der gemeinsame Grenzwert dieser beiden Folgen.

Nach dem Beweis von Satz 5.28 ist a eine obere Schranke für alle b_n und eine untere Schranke für alle c_n. Es gilt also $a \in [b_n, c_n] = I_n$ für alle n. Ist umgekehrt $a' \in \mathbb{R}$ mit $a' \in I_n$ und damit $b_n \le a' \le c_n$ für alle n, so folgt daraus durch Grenzwertbildung mit Satz 5.24 auch $a \le a' \le a$, also $a' = a$. Somit gibt es genau eine Zahl in allen gegebenen Intervallen, nämlich a. □

Zum Abschluss dieses Abschnitts wollen wir analog zu den uneigentlichen Suprema in Bemerkung 4.38 auch *uneigentliche Grenzwerte* definieren, also festlegen, was es heißt, dass eine Folge „den Grenzwert ∞ besitzt". Dies hat den Vorteil, dass viele Aussagen über konvergente Folgen mit gewöhnlichen Grenzwerten in $\mathbb{R}$ auf diesen Fall verallgemeinert werden können.

Definition 5.40 (Uneigentliche Grenzwerte von Folgen)**.** Für eine Folge $(a_n)_n$ in $\mathbb{R}$ schreiben wir $a_n \to \infty$ bzw. $\lim\limits_{n\to\infty} a_n = \infty$, wenn

$$\forall s \in \mathbb{R}\ \exists n_0 \in \mathbb{N}\ \forall n \ge n_0 : a_n > s,$$

also wenn zu jeder vorgegebenen Schranke s fast alle Folgenglieder größer als s sind. Analog definiert man die Eigenschaft $\lim\limits_{n\to\infty} a_n = -\infty$.

Beachte, dass derartige Folgen natürlich insbesondere unbeschränkt und damit nach Lemma 5.8 divergent sind. Man bezeichnet sie als **bestimmt divergent** und nennt ∞ bzw. $-\infty$ einen **uneigentlichen Grenzwert**. Ist $(a_n)_n$ divergent und besitzt nicht in obigem Sinne den Grenzwert ∞ oder $-\infty$, so nennt man $(a_n)_n$ **unbestimmt divergent**.

Beispiel 5.41. Die Folge $(a_n)_n = (n)_n = (0,1,2,3,\dots)$ ist bestimmt divergent mit uneigentlichem Grenzwert $\lim\limits_{n\to\infty} a_n = \infty$. Die Folge $(b_n)_n = (1,0,2,0,3,0,\dots)$ aus Beispiel 5.9 (a) ist dagegen unbestimmt divergent (da z. B. nicht fast alle Folgenglieder größer als 1 sind).

Bemerkung 5.42 (Grenzwertsätze für uneigentliche Grenzwerte)**.** Die Grenzwertsätze aus Satz 5.13 gelten auch für uneigentliche Grenzwerte wie in Definition 5.40, wenn man die formalen Rechenregeln für ∞

$$\begin{aligned} a + \infty &= \infty && \text{für } a \in \mathbb{R}, \\ \infty + \infty &= \infty, \\ a \cdot \infty &= \infty && \text{für } a \in \mathbb{R}_{>0}, \\ \infty \cdot \infty &= \infty, \\ \frac{a}{\infty} &= 0 && \text{für } a \in \mathbb{R} \end{aligned}$$

und analog für $-\infty$ bzw. $a \in \mathbb{R}_{<0}$ definiert. Die Beweise dieser Aussagen sind letztlich analog zu denen von Satz 5.13, jedoch in den einzelnen Fällen immer etwas unterschiedlich, da die Bedingung für den Grenzwert ∞ aus Definition 5.40 ja formal anders aussieht als die eines endlichen Grenzwerts in Definition 5.1. Wir werden die Beweise hier nur exemplarisch in Aufgabe 5.43 betrachten.

Beachte aber, dass die Grenzwertsätze auch weiterhin keine Aussage liefern, wenn eine der betrachteten Folgen unbestimmt divergent ist oder sich Ausdrücke der Form $\infty - \infty$, $0 \cdot \infty$ oder $\frac{\infty}{\infty}$ ergeben, die sich nicht sinnvoll definieren lassen.

Aufgabe 5.43. Es seien $(a_n)_n$ und $(b_n)_n$ zwei reelle Zahlenfolgen.

(a) Man zeige: Gilt $a_n \to a \in \mathbb{R}_{>0}$ und $b_n \to \infty$, so ist auch $a_n b_n \to \infty$.

(b) Man zeige: Gilt $a_n \to \infty$ und $b_n \to \infty$, so ist auch $a_n + b_n \to \infty$.

(c) Kann man in (b) die Bedingung des Grenzwerts ∞ auch durch „nach oben unbeschränkt“ ersetzen, d. h. gilt für nach oben unbeschränkte Folgen $(a_n)_n$ und $(b_n)_n$ auch, dass $(a_n + b_n)_n$ nach oben unbeschränkt ist?

Bemerkung 5.44 (Monotoniekriterium mit uneigentlichem Grenzwert)**.** Auch das Monotoniekriterium lässt sich auf eine Variante mit uneigentlichen Grenzwerten erweitern: Ist eine reelle Folge $(a_n)_n$ zwar wie in Satz 5.28 monoton wachsend, aber nicht nach oben beschränkt, so gibt es zu jedem $s \in \mathbb{R}$ zunächst ein $n_0 \in \mathbb{N}$ mit $a_{n_0} > s$, und wegen der Monotonie dann auch mit $a_n > s$ für alle $n \geq n_0$. Nach Definition 5.40 ist damit dann also $\lim\limits_{n\to\infty} a_n = \infty$. Wir können Satz 5.28 also dahingehend verallgemeinern, dass jede monotone reelle Folge einen Grenzwert in $\mathbb{R} \cup \{\pm\infty\}$ hat, also konvergent oder bestimmt divergent ist.

5.C Limes superior und inferior

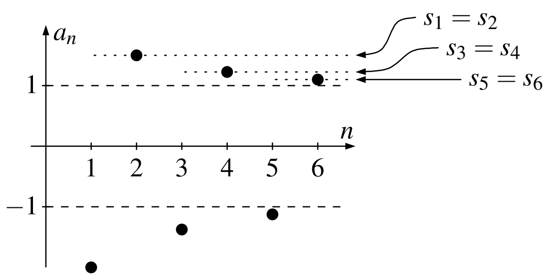
Bisher haben wir das Verhalten von Folgen „im Unendlichen“ in der Regel durch ihren Grenzwert beschrieben. Selbst wenn wir hierbei uneigentliche Grenzwerte wie in Definition 5.40 zulassen, funktioniert dies aber natürlich nicht bei unbestimmt divergenten Folgen, die keinen solchen Grenzwert besitzen.

Um solche Folgen zu untersuchen, können wir wie in Definition 5.18 (c) versuchen, auf Häufungspunkte auszuweichen. In der Tat wollen wir jetzt zeigen, dass *jede* Folge (zumindest im uneigentlichen Sinne) auch wirklich mindestens einen Häufungspunkt besitzt. Der Einfachheit halber betrachten wir hierfür zunächst nur beschränkte reelle Folgen, deren Häufungspunkte dann also in $\mathbb{R}$ liegen müssen. In diesem Fall werden wir in Folgerung 5.48 sehen, dass die folgende Konstruktion stets einen Häufungspunkt liefert – und zwar sogar einen ganz bestimmten, nämlich den größten.

Konstruktion 5.45. Es sei $(a_n)_n$ eine beschränkte reelle Folge. Wie im Bild unten für die Folge $(a_n)_n$ mit $a_n = (-1)^n\,(1+\frac{1}{n})$ dargestellt konstruieren wir nun die Hilfsfolge $(s_n)_n$ durch

$$s_n := \sup\{a_k : k \geq n\},$$

d. h. wir betrachten das Supremum aller Folgenglieder, wobei wir aber für s_n erst beim n-ten Folgenglied anfangen.



Beachte dabei, dass die Mengen $\{a_k : k \geq n\}$ nach Voraussetzung beschränkt sind und die Suprema s_n damit nach dem Supremumsaxiom in $\mathbb{R}$ existieren. Außerdem folgt aus der Beschränktheit der Folgenglieder, dass auch $(s_n)_n$ eine beschränkte Folge ist.

Darüber hinaus ist die Folge $(s_n)_n$ monoton fallend: Die obere Schranke s_n der Menge $\{a_k : k \geq n\}$ ja auch eine obere Schranke der Teilmenge $\{a_k : k \geq n+1\} \subset \{a_k : k \geq n\}$ und muss damit größer oder gleich der kleinsten oberen Schranke s_{n+1} von $\{a_k : k \geq n+1\}$ sein – d. h. es ist $s_{n+1} \leq s_n$.

Wir haben also gesehen, dass $(s_n)_n$ eine monoton fallende und (nach unten) beschränkte reelle Folge ist. Nach dem Monotoniekriterium aus Satz 5.28 besitzt sie also einen Grenzwert. Im oben dargestellten Beispiel ist dieser Grenzwert offensichtlich 1. Da die Häufungspunkte unserer Beispielfolge nach dem Mischfolgenlemma 5.22 genau ± 1 sind, ist der Grenzwert von $(s_n)_n$ hier also gerade der größte Häufungspunkt von $(a_n)_n$. Bevor wir zeigen, dass dies immer der Fall ist, geben wir der so konstruierten Zahl noch einen Namen.

Definition 5.46 (Limes superior und Limes inferior)**.** Für eine beschränkte reelle Folge $(a_n)_n$ definieren wir

$$\begin{aligned} &\text{den } \textbf{Limes superior} \quad && \limsup_{n\to\infty} a_n := \lim_{n\to\infty} \big(\sup\{a_k : k \geq n\}\big) \\ \text{und} \quad &\text{den } \textbf{Limes inferior} \quad && \liminf_{n\to\infty} a_n := \lim_{n\to\infty} \big(\inf\{a_k : k \geq n\}\big). \end{aligned}$$

In der Literatur sind hierfür auch die Schreibweisen $\overline{\lim}\, a_n$ bzw. $\underline{\lim}\, a_n$ üblich.

Das folgende Lemma zeigt, dass sich der Limes superior (und analog der Limes inferior) in gewissem Sinne wie eine „Mischung" aus einem Grenzwert und einem Häufungspunkt verhält: Während für einen Grenzwert a ja in jeder ε-Umgebung $U_\varepsilon(a)$ fast alle Folgenglieder, für einen Häufungspunkt nach Lemma 5.21 aber nur unendlich viele Glieder liegen müssen, ist der Limes superior die (eindeutig bestimmte) Zahl a, für die für alle ε fast alle Folgenglieder kleiner als $a+\varepsilon$, und unendlich viele größer als $a-\varepsilon$ sind.

Lemma 5.47. *Es seien $(a_n)_n$ eine beschränkte reelle Folge und $a \in \mathbb{R}$. Dann gilt $a = \limsup_{n\to\infty} a_n$ genau dann, wenn für alle $\varepsilon > 0$ die folgenden beiden Bedingungen erfüllt sind:*

(a) *Für fast alle n ist $a_n < a+\varepsilon$.*

(b) *Für unendlich viele n ist $a_n > a-\varepsilon$.*

(Eine analoge Aussage gilt natürlich auch für den Limes inferior.)

Beweis. Wie in Konstruktion 5.45 sei $s_n = \sup\{a_k : k \geq n\}$, so dass also $\limsup_{n\to\infty} a_n = \lim_{n\to\infty} s_n$ gilt. Weiterhin sei $\varepsilon > 0$ beliebig.

„$\Rightarrow$": Es gelte $s_n \to a$, also $s_n \in (a-\varepsilon, a+\varepsilon)$ für fast alle n. Wir müssen (a) und (b) zeigen.

Da s_n eine obere Schranke für $\{a_k : k \geq n\}$ (also insbesondere für a_n) ist, gilt $a_n \leq s_n < a+\varepsilon$ für fast alle n. Dies zeigt (a).

Um (b) zu zeigen, nehmen wir an, es gäbe nur endlich viele n mit $a_n > a-\varepsilon$. Dann gäbe es also ein $n_0 \in \mathbb{N}$ mit $a_n \leq a-\varepsilon$ für alle $n \geq n_0$, d. h. $a-\varepsilon$ wäre eine obere Schranke für alle diese Folgenglieder. Daraus folgt dann aber auch $s_n \leq a-\varepsilon$ für alle $n \geq n_0$, im Widerspruch zu $s_n \to a$.

„$\Leftarrow$": Wir setzen nun (a) und (b) voraus und müssen $s_n \in (a-\varepsilon, a+\varepsilon)$ für fast alle n zeigen.

Nach (a) gibt es ein $n_0 \in \mathbb{N}$ mit $a_n < a+\frac{\varepsilon}{2}$ für alle $n \geq n_0$. Damit ist für diese n auch $s_n = \sup\{a_k : k \geq n\} \leq a+\frac{\varepsilon}{2} < a+\varepsilon$.

Weiterhin gibt es nach (b) zu jedem $n \in \mathbb{N}$ ein $k \geq n$ mit $a_k > a-\varepsilon$, woraus natürlich auch $s_n = \sup\{a_k : k \geq n\} > a-\varepsilon$ folgt.

Insgesamt gilt also $s_n \in (a-\varepsilon, a+\varepsilon)$ für fast alle n. $\square$

Aus diesem Lemma ergibt sich nun die folgende wichtige Charakterisierung des Limes superior (und inferior), mit der man diese Zahl oft sehr einfach bestimmen kann.

Folgerung 5.48. *Für jede beschränkte reelle Folge* $(a_n)_n$ *ist* $\limsup\limits_{n\to\infty} a_n$ *der größte Häufungspunkt von* $(a_n)_n$*. (Analog ist dann natürlich* $\liminf\limits_{n\to\infty} a_n$ *der kleinste Häufungspunkt von* $(a_n)_n$*.)*

*Insbesondere besitzt also jede beschränkte reelle Folge einen Häufungspunkt (**Satz von Bolzano-Weierstraß** für* $\mathbb{R}$*).*

Beweis. Es seien $a = \limsup\limits_{n\to\infty} a_n$ und $\varepsilon > 0$ beliebig. Aus (a) und (b) von Lemma 5.47 folgt dann $a-\varepsilon < a_n < a+\varepsilon$ für unendlich viele n, d. h. nach Lemma 5.21 ist a ein Häufungspunkt von $(a_n)_n$.

Andererseits ist aber kein $b > a$ ein Häufungspunkt von $(a_n)_n$: Setzen wir nämlich $\varepsilon = \frac{b-a}{2} > 0$, so ist $a+\varepsilon = b-\varepsilon$, und somit gilt nach Lemma 5.47 (a) für fast alle n

$$a_n < a+\varepsilon = b-\varepsilon \quad \Rightarrow \quad a_n \notin (b-\varepsilon, b+\varepsilon).$$

11 Damit kann b kein Häufungspunkt von $(a_n)_n$ sein. $\square$

Beispiel 5.49.

(a) Die Folge $(a_n)_{n\in\mathbb{N}_{>0}}$ mit $a_n = (-1)^n\,(1+\frac{1}{n})$ ist eine Mischfolge aus den geraden Gliedern $a_{2n} = 1+\frac{1}{2n} \to 1$ und den ungeraden Gliedern $a_{2n+1} = -(1+\frac{1}{2n+1}) \to -1$, sie hat nach Lemma 5.22 also die einzigen Häufungspunkte 1 und -1. Aus Folgerung 5.48 ergibt sich damit sofort $\limsup\limits_{n\to\infty} a_n = 1$ und $\liminf\limits_{n\to\infty} a_n = -1$.

(b) Ist $(a_n)_n$ eine konvergente reelle Folge, so ist ihr Grenzwert a nach Lemma 5.19 der einzige Häufungspunkt. Also ist dann $\limsup\limits_{n\to\infty} a_n = \liminf\limits_{n\to\infty} a_n = a$ nach Folgerung 5.48.

Ist umgekehrt $(a_n)_n$ eine beschränkte reelle Folge mit $\limsup\limits_{n\to\infty} a_n = \liminf\limits_{n\to\infty} a_n =: a$, so folgt aus Lemma 5.47 (a) für alle $\varepsilon > 0$, dass $a-\varepsilon < a_n < a+\varepsilon$ für fast alle n ist – wobei sich die erste Ungleichung aus $\liminf\limits_{n\to\infty} a_n = a$ und die zweite aus $\limsup\limits_{n\to\infty} a_n = a$ ergibt. Also ist $(a_n)_n$ dann konvergent mit Grenzwert a.

Aufgabe 5.50. Berechne $\limsup\limits_{n\to\infty} a_n$ und $\liminf\limits_{n\to\infty} a_n$ für $a_n = \frac{1}{\sqrt{n^2+n+(-1)^n\cdot n}}$.

Aufgabe 5.51.

(a) Es seien $(a_n)_n$ und $(b_n)_n$ zwei beschränkte Folgen positiver Zahlen. Zeige, dass

$$\limsup_{n\to\infty}(a_n b_n) \le \limsup_{n\to\infty} a_n \cdot \limsup_{n\to\infty} b_n,$$

und gib ein Beispiel dafür an, dass hier im Allgemeinen keine Gleichheit gilt.

(b) Beweise, dass in (a) jedoch stets die Gleichheit gilt, wenn $(a_n)_n$ oder $(b_n)_n$ konvergent ist.

Bemerkung 5.52 (Uneigentliche Werte für Limes superior und Limes inferior)**.** Lässt man für Supremum, Infimum und Grenzwerte wie in Bemerkung 4.38 und Definition 5.40 formal auch $\pm\infty$ zu, so kann man den Limes superior und Limes inferior nach Bemerkung 5.44 genau wie in Definition 5.46 auch für beliebige reelle Folgen konstruieren und erhält dann Werte in $\mathbb{R}\cup\{\pm\infty\}$.

Aufgabe 5.53. Zeige, dass die Aussage von Folgerung 5.48 auch für beliebige reelle Folgen richtig ist, wenn man wie in Bemerkung 5.52 die uneigentlichen Werte $\pm\infty$ für Häufungspunkte sowie den Limes superior und inferior zulässt.

Insbesondere gibt es also auch vom Satz von Bolzano-Weierstraß die erweiterte Form, dass jede (nicht notwendig beschränkte) reelle Folge einen (evtl. uneigentlichen) Häufungspunkt hat.

5.D Mächtigkeiten von Mengen

Zum Abschluss dieses Kapitels wollen wir noch ein Thema behandeln, das wir auch schon deutlich früher hätten ansprechen können, aber das durch unsere Ergebnisse zu Folgen nun einfacher zu untersuchen ist: die Frage nach der „Größe“ von Mengen. Gibt es zu zwei Mengen M und N eine bijektive Abbildung $f: M \to N$ und damit wie im rechten Bild von Definition 2.8 eine 1:1-Beziehung zwischen ihren Elementen, so können wir uns diese beiden Mengen anschaulich als „gleich groß“ vorstellen. Ist z. B. $M = \{x_1, \dots, x_n\}$ eine endliche Menge mit n Elementen, so ist $N = \{f(x_1), \dots, f(x_n)\}$ (da f surjektiv ist und somit alle Elemente von N trifft) – und in dieser Aufzählung der Elemente von N steht auch kein Element doppelt, da f injektiv ist. Also hat N dann ebenfalls n Elemente, d. h. genauso viele Elemente wie M.

Wir wollen dieses Konzept nun für unendliche Mengen untersuchen. Ist es auch in diesem Fall noch sinnvoll, sich zwei Mengen M und N als „gleich groß“ vorzustellen, wenn eine bijektive Abbildung $f: M \to N$ zwischen ihnen existiert? Gibt es überhaupt „verschieden große unendliche Mengen“? Da diese Frage intuitiv nicht mehr besonders gut zugänglich ist, sollten wir natürlich zunächst erst einmal exakt definieren, worüber wir reden wollen.

Definition 5.54 (Gleichmächtige und abzählbare Mengen)**.**

(a) Zwei Mengen M und N heißen **gleichmächtig**, wenn es zwischen ihnen eine bijektive Abbildung $f: M \to N$ gibt.

(b) Eine Menge M heißt …

- **abzählbar unendlich**, wenn sie gleichmächtig zu $\mathbb{N}$ ist, also wenn es eine bijektive Abbildung $f: \mathbb{N} \to M$ gibt.
- **abzählbar**, wenn sie endlich oder abzählbar unendlich ist.
- **überabzählbar**, wenn sie nicht abzählbar ist.

Bemerkung 5.55. Die Gleichmächtigkeit erfüllt formal die Eigenschaften einer Äquivalenzrelation:

(a) Jede Menge M ist gleichmächtig zu sich selbst (mit der Identität $\mathrm{id}_M: M \to M$).

(b) Ist M gleichmächtig zu N, so ist auch N gleichmächtig zu M (denn mit einer bijektiven Abbildung $f: M \to N$ ist nach Aufgabe 2.25 auch ihre Umkehrung $f^{-1}: N \to M$ bijektiv).

(c) Ist M gleichmächtig zu N und N gleichmächtig zu R, so ist auch M gleichmächtig zu R (denn die Verkettung bijektiver Abbildungen ist nach Aufgabe 2.25 wieder bijektiv).

Man könnte daher versucht sein zu sagen, dass die Gleichmächtigkeit eine Äquivalenzrelation auf der Menge aller Mengen ist. Dies ist jedoch nicht ganz korrekt, da wir wie in Bemerkung 1.13 erläutert keine „Menge aller Mengen“ bilden können.

Beispiel 5.56.

(a) Wie wir am Anfang dieses Abschnitts gesehen haben, sind zwei endliche Mengen M und N genau dann gleichmächtig, wenn sie gleich viele Elemente haben, also wenn $|M| = |N|$ gilt. Insbesondere ist eine endliche Menge also nie gleichmächtig zu einer echten Teilmenge von ihr: Wenn wir von einer endlichen Menge Elemente entfernen, wird sie in diesem Sinne „kleiner“ – was natürlich nicht allzu überraschend sein sollte.

(b) Für unendliche Mengen ist dies jedoch falsch: Die Menge $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ ist gleichmächtig zu ihrer echten Teilmenge $\mathbb{N}\backslash\{0\} = \{1, 2, 3, \dots\}$, z. B. durch die bijektive Abbildung

$$f: \mathbb{N} \to \mathbb{N}\backslash\{0\},\ n \mapsto n+1,$$

die jede Zahl um 1 erhöht.

Bemerkung 5.57.

(a) Die abzählbar unendlichen Mengen sind genau diejenigen, die sich als Aufzählung in der Form $M = \{x_0, x_1, x_2, \dots\}$ mit $x_m \neq x_n$ für alle $m, n \in \mathbb{N}$ mit $m \neq n$ schreiben lassen: Die Funktion $f \colon \mathbb{N} \to M,\ n \mapsto x_n$ ist dann die geforderte bijektive Abbildung. Dies erklärt auch den Begriff „*abzählbar* unendlich". Wir sehen so auch z. B. schon, dass auch die Menge $\mathbb{Z}$ der ganzen Zahlen abzählbar ist, da wir sie z. B. als $\mathbb{Z} = \{0, -1, 1, -2, 2, -3, 3, \dots\}$ schreiben können.

(b) Aus (a) ergibt sich direkt, dass jede Teilmenge M einer abzählbaren Menge N wieder abzählbar ist: Ist $N = \{x_0, x_1, x_2, \dots\}$ abzählbar (mit einer evtl. abbrechenden Aufzählung), so lässt sich jede Teilmenge $M \subset N$ durch Weglassen gewisser Elemente als $M = \{x_{n_0}, x_{n_1}, x_{n_2}, \dots\}$ für geeignete (evtl. endlich viele) $n_0 < n_1 < n_2 < \cdots$ schreiben, ist damit also ebenfalls abzählbar.

Abzählbare Mengen bleiben aber nicht nur abzählbar, wenn man Elemente von ihnen entfernt. Man kann sie umgekehrt auch noch um „sehr viele" Elemente vergrößern, ohne dass sie dadurch überabzählbar werden. Konkret wollen wir jetzt zeigen, dass wir sogar abzählbar viele abzählbare Mengen vereinigen können und dabei immer noch eine abzählbare Menge erhalten. Im folgenden Satz sind diese abzählbaren Mengen dazu mit M_i bezeichnet, wobei i die Mengen durchnummeriert und damit selbst abzählbar viele Werte (in der sogenannten Indexmenge) annehmen kann.

Satz 5.58 (Abzählbare Vereinigungen abzählbarer Mengen sind abzählbar)**.** *Es seien I eine abzählbare Indexmenge sowie M_i für alle $i \in I$ eine abzählbare Menge. Dann ist die Vereinigung aller dieser Mengen M_i, geschrieben $\bigcup_{i \in I} M_i$, ebenfalls abzählbar.*

Beweis. Nach Bemerkung 5.57 (a) können wir die Elemente von I sowie allen M_i mit $i \in I$ in der Form

$$I = \{i_0, i_1, i_2, \dots\} \qquad \text{und} \qquad M_{i_k} = \{x_{k,0}, x_{k,1}, x_{k,2}, \dots\} \quad \text{für } k \in \mathbb{N}$$

aufzählen (wobei einige dieser Mengen auch endlich sein können, so dass die Aufzählungen dann also irgendwo abbrechen). Wir können die Elemente aller M_i damit in der folgenden Form auflisten und abzählen:

$$\begin{array}{llllll}
M_{i_0}: & x_{0,0} & x_{0,1} & x_{0,2} & x_{0,3} & x_{0,4} \ \cdots \\
M_{i_1}: & x_{1,0} & x_{1,1} & x_{1,2} & x_{1,3} & \\
M_{i_2}: & x_{2,0} & x_{2,1} & x_{2,2} & & \\
M_{i_3}: & x_{3,0} & & & &
\end{array}$$

Wir haben also

$$\bigcup_{i \in I} M_i = \{x_{0,0},\ x_{0,1}, x_{1,0},\ x_{0,2}, x_{1,1}, x_{2,0},\ x_{0,3}, x_{1,2}, x_{2,1}, x_{3,0},\ x_{0,4}, x_{1,3}, x_{2,2}, \dots\}.$$

Dabei müssen wir in dieser Aufzählung alle nicht vorhandenen Positionen (wenn einige der Mengen I oder M_i endlich sind) und bereits vorher vorgekommene Elemente (wenn die M_i nicht disjunkt sind) weglassen. Auf diese Art sehen wir also, dass $\bigcup_{i \in I} M_i$ endlich oder abzählbar unendlich sein muss. Man bezeichnet die obige Abzählart auch als das **Cantorsche Diagonalverfahren**. □

Beispiel 5.59.

(a) Sind M und N abzählbare Mengen, so ist nach Satz 5.58 auch ihr Produkt $M \times N$ abzählbar, da man es als abzählbare Vereinigung $\bigcup_{m \in M} (\{m\} \times N)$ abzählbarer Mengen schreiben kann.

(b) Für ein festes $q \in \mathbb{N}_{>0}$ ist die Menge $M_q = \left\{ \frac{p}{q} : p \in \mathbb{Z} \right\}$ aller rationalen Zahlen, die sich als Bruch mit Nenner q schreiben lassen, bijektiv zu $\mathbb{Z}$ und damit nach Bemerkung 5.57 (a) abzählbar. Damit ist nach Satz 5.58 auch die Menge $\mathbb{Q} = \bigcup_{q \in \mathbb{N}_{>0}} M_q$ aller rationalen Zahlen abzählbar. Auch wenn es der ersten Intuition vermutlich widerspricht, gibt es in diesem Sinne also „genauso viele" rationale wie natürliche Zahlen.

Auch wenn wir mit Satz 5.58 jetzt von sehr vielen Mengen sehen können, dass sie abzählbar sind, gibt es dennoch unendliche Mengen, die „zu groß“ sind, um eine bijektive Abbildung nach $\mathbb{N}$ zuzulassen. Das einfachste Beispiel hierfür ist die Menge der reellen Zahlen.

Satz 5.60. *Die Menge $\mathbb{R}$ ist überabzählbar.*

Beweis. Dies ist eine unmittelbare Folge der Intervallschachtelung aus Satz 5.39: Angenommen, wir hätten eine Aufzählung $\mathbb{R} = \{x_0, x_1, x_2, \dots\}$ der reellen Zahlen. Wir konstruieren dann rekursiv eine Intervallschachtelung $I_0 \supset I_1 \supset I_2 \supset \cdots$ aus Intervallen I_n der Länge $\frac{1}{3^n}$, von der wir lediglich verlangen, dass für alle n die Zahl x_n nicht in I_n liegt. Dies ist natürlich sehr einfach: Wir wählen für I_{n+1} immer das linke Drittel von I_n – es sei denn, die Zahl x_{n+1} liegt in diesem Drittel; dann wählen wir für I_{n+1} das rechte Drittel von I_n.

Nach Satz 5.39 gibt es nun aber ein $a \in \mathbb{R}$, das in allen I_n liegt und damit keine der Zahlen $x_n \in \mathbb{R}\backslash I_n$ sein kann, also im Widerspruch zur Annahme in unserer Aufzählung von $\mathbb{R}$ nicht enthalten ist. $\square$

Zusammenfassend können wir also sagen, dass es im Sinne der Gleichmächtigkeit zwar „genauso viele“ natürliche wie ganze oder rationale, aber „deutlich mehr“ reelle Zahlen gibt.

Beispiel 5.61. Nach Beispiel 5.59 (b) gibt es eine Aufzählung $\mathbb{Q} = \{x_0, x_1, x_2, \dots\}$ der rationalen Zahlen. Fassen wir diese als Folge $(x_n)_{n\in\mathbb{N}}$ auf, so liegen in jeder ε-Umgebung jeder reellen Zahl nach Aufgabe 5.36 unendlich viele Folgenglieder. Also ist jede reelle Zahl ein Häufungspunkt dieser Folge: Eine (abzählbare) Folge kann durchaus überabzählbar viele Häufungspunkte haben!

Aufgabe 5.62. Untersuche die folgenden Mengen auf Abzählbarkeit:

(a) die Menge aller zweielementigen Teilmengen von $\mathbb{N}$;

(b) die Menge aller endlichen Teilmengen von $\mathbb{N}$;

(c) die Menge aller Teilmengen von $\mathbb{N}$ (also die Potenzmenge $\mathscr{P}(\mathbb{N})$ von $\mathbb{N}$).

6. Komplexe Zahlen

Bisher haben wir uns nahezu ausschließlich mit dem euch aus der Schule bekannten Körper $\mathbb{R}$ der reellen Zahlen befasst. In der Mathematik – und zwar sowohl in der Analysis als auch in der Algebra – ist jedoch noch ein weiterer damit eng zusammenhängender Körper sehr wichtig: der Körper der komplexen Zahlen. Die Idee dabei ist, einen Körper zu konstruieren, der die reellen Zahlen als Teilmenge enthält (man nennt so etwas auch eine *Körpererweiterung* von $\mathbb{R}$), und in dem jede nicht konstante Polynomfunktion (wie z. B. $x \mapsto x^2 + 1$) eine Nullstelle besitzt. Außerdem werden die komplexen Zahlen zur Einführung und Untersuchung der Winkelfunktionen in Kapitel 9.B sehr nützlich sein.

6.A Die Konstruktion der komplexen Zahlen

Im Gegensatz zu den reellen Zahlen brauchen wir die komplexen nicht mehr axiomatisch vorauszusetzen – wir können sie explizit aus den reellen konstruieren.

Definition 6.1 (Komplexe Zahlen)**.** Die Menge der **komplexen Zahlen** ist definiert als $\mathbb{C} := \mathbb{R}^2$. Wir betrachten auf dieser Menge die beiden Verknüpfungen

$$(x_1,y_1)+(x_2,y_2) := (x_1+x_2, y_1+y_2) \qquad \text{(Addition)}$$
$$\text{und} \qquad (x_1,y_1)\cdot(x_2,y_2) := (x_1x_2 - y_1y_2, x_1y_2 + y_1x_2) \qquad \text{(Multiplikation).}$$

Notation 6.2 (Komplexe Zahlen in der Form $x+\mathrm{i}y$)**.** Beachte, dass für komplexe Zahlen, deren zweite Komponente 0 ist, die Addition und Multiplikation

$$(x_1,0)+(x_2,0) = (x_1+x_2,0) \qquad \text{bzw.} \qquad (x_1,0)\cdot(x_2,0) = (x_1x_2,0)$$

genauso definiert ist wie für reelle Zahlen. Wir schreiben die komplexe Zahl $(x,0) \in \mathbb{C}$ daher in der Regel auch einfach als x und fassen auf diese Art die reellen Zahlen als Teilmenge der komplexen auf. Setzen wir weiterhin $\mathrm{i} := (0,1)$, so gilt

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

sowie für alle $x,y \in \mathbb{R}$

$$x+\mathrm{i}y = (x,0)+(0,1)(y,0) = (x,0)+(0,y) = (x,y).$$

Diese Darstellung als $x+\mathrm{i}y$ ist in der Tat auch die übliche Schreibweise für eine komplexe Zahl – wir werden Elemente von $\mathbb{C}$ ab jetzt immer in dieser Form schreiben. Diese Notation hat den Vorteil, dass sich die Rechenregeln für die Addition und Multiplikation aus Definition 6.1 ganz von selbst ergeben, wenn man i als Variable auffasst und die Gleichung $\mathrm{i}^2 = -1$ berücksichtigt: Es ist dann nämlich wie erwartet

$$(x_1+\mathrm{i}y_1)+(x_2+\mathrm{i}y_2) = (x_1+x_2)+\mathrm{i}\,(y_1+y_2)$$
$$\text{und} \qquad (x_1+\mathrm{i}y_1)\cdot(x_2+\mathrm{i}y_2) = (x_1x_2-y_1y_2)+\mathrm{i}\,(x_1y_2+x_2y_1) = x_1x_2+\mathrm{i}^2y_1y_2+\mathrm{i}x_1y_2+\mathrm{i}x_2y_1.$$

Wenn man in ingenieurswissenschaftliche Bücher schaut, werden die komplexen Zahlen dort in der Tat sogar oft so eingeführt: Man nehme einfach an, dass es eine Zahl i mit $\mathrm{i}^2 = -1$ gibt, und rechne damit dann ganz normal weiter, als wäre nichts Besonderes passiert. Es sollte aber hoffentlich klar sein, dass eine solche „Definition“ aus mathematischer Sicht unsinnig ist: Wenn wir bisher nur die reellen Zahlen kennen, gibt es nach Lemma 4.16 (c) einfach keine Zahl, deren Quadrat gleich -1 ist – und diese Situation wird natürlich auch nicht dadurch besser, dass wir diesem nicht existierenden Objekt einen Namen i geben. Stattdessen müssen wir den Umweg über die korrekte Konstruktion aus Definition 6.1 gehen, die uns garantiert, dass $\mathbb{C}$ erst einmal widerspruchsfrei definiert ist, und können dann erst im Nachhinein untersuchen, welche Eigenschaften der reellen Zahlen sich tatsächlich auf die komplexen übertragen. Dies sind nämlich auch nicht alle – so werden wir z. B. in Lemma 6.6 und Bemerkung 6.8 sehen, dass $\mathbb{C}$ zwar ein Körper, aber kein geordneter Körper ist. 12

Definition 6.3 (Real- und Imaginärteil, Konjugation und Betrag)**.** Es sei $z = x + \mathrm{i}y \in \mathbb{C}$ mit $x, y \in \mathbb{R}$ wie in Notation 6.2.

(a) Man nennt x den **Realteil** und y den **Imaginärteil** von z; die Notation hierfür ist $x = \operatorname{Re} z$ und $y = \operatorname{Im} z$.

(b) Man nennt

$$\bar{z} := x - \mathrm{i}y \qquad \text{die zu } z \textbf{ komplex konjugierte Zahl}$$

$$\text{und} \qquad |z| := \sqrt{x^2 + y^2} \in \mathbb{R}_{\geq 0} \qquad \text{den } \textbf{Betrag} \text{ von } z$$

(mit der reellen Wurzel aus Definition 5.30).

Bemerkung 6.4. Offensichtlich lassen sich der Real- und Imaginärteil von $z = x + \mathrm{i}y \in \mathbb{C}$ ausdrücken als

$$\operatorname{Re} z = \frac{1}{2}(z + \bar{z}) \qquad \text{und} \qquad \operatorname{Im} z = -\frac{\mathrm{i}}{2}(z - \bar{z}),$$

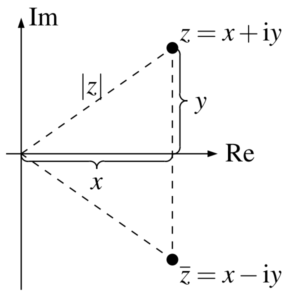
während der Betrag wegen $z\bar{z} = (x + \mathrm{i}y)(x - \mathrm{i}y) = x^2 - \mathrm{i}^2 y^2 = x^2 + y^2$ auch als

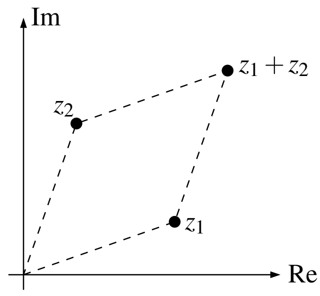
$$|z| = \sqrt{z\bar{z}}$$

geschrieben werden kann.

Bemerkung 6.5 (Geometrische Interpretation von $\mathbb{C}$)**.** Geometrisch können wir die Elemente von $\mathbb{C} = \mathbb{R}^2$ natürlich als Punkte der Ebene, der sogenannten **komplexen Zahlenebene**, zeichnen. Wir wollen jetzt sehen, wie man die oben eingeführten Operationen für komplexe Zahlen in dieser Zahlenebene grafisch veranschaulichen kann. Da diese Interpretation zwar für die Vorstellung sehr wichtig ist, aber nicht für unsere späteren exakten Rechnungen benötigt wird, wollen wir dabei ein paar einfache und sicherlich bekannte Prinzipien der Schulgeometrie ohne Beweis verwenden.

Zunächst einmal ist klar, dass die *reellen* Zahlen in $\mathbb{C}$, also diejenigen der Form $x + \mathrm{i} \cdot 0$, genau die auf der horizontalen Achse sind. Der Betrag $|z|$ einer komplexen Zahl ist nach Definition genau der Abstand des Punktes z vom Ursprung, und die komplexe Konjugation entspricht einer Spiegelung an der reellen Achse (wie im Bild unten links). Ebenso offensichtlich ist, dass zwei komplexe Zahlen genau so addiert werden, wie ihr in der Schule Vektoren im $\mathbb{R}^2$ addiert habt, also indem man die Verbindungsstrecken vom Ursprung zu z_1 und z_2 wie im folgenden Bild rechts zu einem Parallelogramm zusammensetzt.



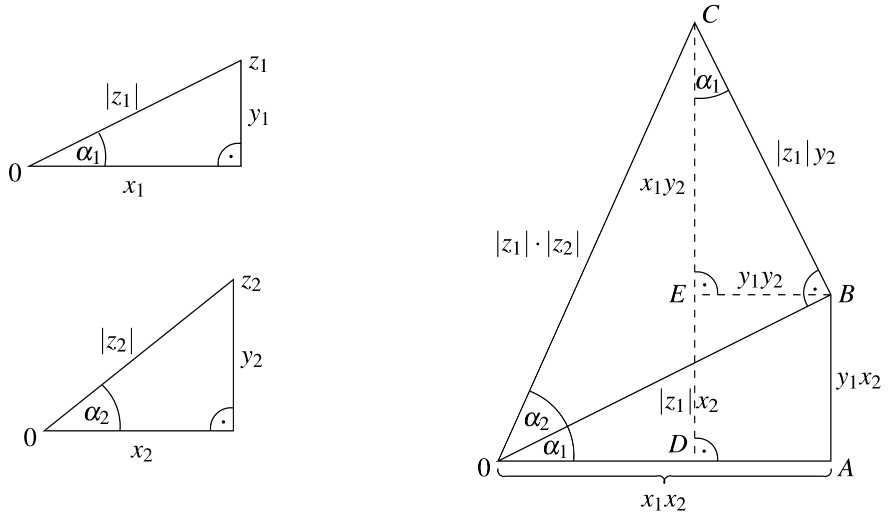


Die Multiplikation dagegen ist schon interessanter. Der Einfachheit halber beschränken wir uns im Bild unten auf den Fall, in dem Real- und Imaginärteil beider Zahlen positiv sind – die anderen Fälle lassen sich analog behandeln. Wir haben dort links zwei komplexe Zahlen z_1 und z_2 wie oben dargestellt (und zusätzlich die Winkel eingezeichnet, die die Verbindungsstrecken zum Ursprung mit der positiven reellen Achse einschließen), und die zugehörigen rechtwinkligen Dreiecke rechts wie folgt zusammengesetzt:

(a) Das Dreieck für z_1 haben wir um den Faktor x_2 zum Dreieck $0AB$ gestreckt.

(b) Das Dreieck für z_2 haben wir um den Faktor $|z_1|$ gestreckt und um den Winkel α_1 gedreht, so dass das Dreieck $0BC$ mit Seitenlängen $|z_1|x_2$, $|z_1|y_2$ und $|z_1|\cdot|z_2|$ entstanden ist (insbesondere hat dieses Dreieck mit dem aus (a) also eine gemeinsame Kante $0B$ mit der Seitenlänge $|z_1|x_2$).

(c) CD ist die Senkrechte auf $0A$, und BE die Senkrechte auf CD. Damit ist das Dreieck CEB ähnlich zu $0AB$, es ist daher die Streckung des Dreiecks für z_1 um den Faktor y_2 und hat Seitenlängen x_1y_2, y_1y_2 und $|z_1|y_2$.



Aus diesem Bild lesen wir nun direkt ab, dass C die Koordinaten $(x_1x_2 - y_1y_2, x_1y_2 + y_1x_2)$ hat, also genau der Punkt $z_1 \cdot z_2$ ist. Da dieser Punkt den Betrag $|z_1|\cdot|z_2|$ hat und den Winkel $\alpha_1 + \alpha_2$ mit der positiven reellen Achse einschließt, sehen wir anschaulich:

Komplexe Zahlen werden multipliziert, indem man ihre Beträge *multipliziert* und ihre Winkel *addiert*.

Wir wollen nun sehen, dass die Addition und Multiplikation auf $\mathbb{C}$ die erwarteten Eigenschaften haben, also die Struktur eines Körpers bilden. Unsere Ergebnisse aus Kapitel 3 und Abschnitt 4.A gelten somit unverändert auch für die komplexen Zahlen.

Lemma 6.6. *$\mathbb{C}$ ist ein Körper.*

Beweis (siehe auch Aufgabe 3.11). Die Kommutativität der Addition und Multiplikation ist aus der Definition offensichtlich. Die Assoziativität der Addition und Multiplikation sowie die Distributivität rechnet man einfach nach; wir zeigen hier exemplarisch die Distributivität: Für drei komplexe Zahlen $z_1 = x_1 + \mathrm{i}y_1$, $z_2 = x_2 + \mathrm{i}y_2$, $z_3 = x_3 + \mathrm{i}y_3$ folgt (letztlich wegen der Distributivität in $\mathbb{R}$)

$$\begin{aligned}(z_1+z_2)z_3 &= ((x_1+x_2)+\mathrm{i}(y_1+y_2))\cdot(x_3+\mathrm{i}y_3)\\ &= (x_1+x_2)x_3-(y_1+y_2)y_3+\mathrm{i}((x_1+x_2)y_3+(y_1+y_2)x_3)\\ &= \big(x_1x_3-y_1y_3+\mathrm{i}(x_1y_3+y_1x_3)\big)+\big(x_2x_3-y_2y_3+\mathrm{i}(x_2y_3+y_2x_3)\big)\\ &= z_1z_3+z_2z_3.\end{aligned}$$

Das additive neutrale Element ist 0, das additive Inverse zu $z = x + \mathrm{i}y$ natürlich $-z = -x - \mathrm{i}y$. Das multiplikative neutrale Element ist 1, das multiplikative Inverse zu $z = x + \mathrm{i}y \neq 0$ ist

$$\frac{x}{x^2+y^2}+\mathrm{i}\,\frac{-y}{x^2+y^2}, \quad \text{denn} \quad \left(\frac{x}{x^2+y^2}+\mathrm{i}\,\frac{-y}{x^2+y^2}\right)\cdot(x+\mathrm{i}y) = \frac{(x-\mathrm{i}y)(x+\mathrm{i}y)}{x^2+y^2} = \frac{x^2+y^2}{x^2+y^2} = 1. \qquad \square$$

Beispiel 6.7 (Division komplexer Zahlen)**.** Erwähnenswert ist an Lemma 6.6 wohl vor allem die Existenz einer Division, da ja zunächst einmal nicht offensichtlich ist, wie man für eine komplexe Zahl $z = x + \mathrm{i}y$ das multiplikative Inverse $\frac{1}{z} = \frac{1}{x+\mathrm{i}y}$ wieder in der Form $x' + \mathrm{i}y'$ schreiben kann. Die Merkregel hierfür ist, dass man diesen Bruch mit $\overline{z}$ zu $\frac{1}{z} = \frac{\overline{z}}{z\overline{z}}$ erweitert, so dass der Nenner zu der nach Bemerkung 6.4 *reellen* Zahl $z\overline{z} = |z|^2$ wird und somit das i aus dem Nenner verschwindet. So ist z. B.

$$\frac{1}{1+2\mathrm{i}} = \frac{1-2\mathrm{i}}{(1+2\mathrm{i})(1-2\mathrm{i})} = \frac{1-2\mathrm{i}}{1^2+2^2} = \frac{1}{5} - \frac{2}{5}\mathrm{i}.$$

Bemerkung 6.8 ($\mathbb{C}$ ist kein geordneter Körper)**.** $\mathbb{C}$ ist zwar ein Körper, kann aber nicht zu einem *geordneten* Körper gemacht werden. Andernfalls müsste nämlich i^2 als Quadrat einer Zahl ungleich 0 nach Lemma 4.16 (c) positiv sein – was aber natürlich ein Widerspruch ist, da andererseits $\mathrm{i}^2 = -1$ nach demselben Lemma auch eine negative Zahl sein müsste.

Es ergibt also keinen Sinn zu fragen, welche von zwei gegebenen komplexen Zahlen größer ist als die andere. Damit sind unsere Ergebnisse aus den Abschnitten 4.B und 4.C auf die komplexen Zahlen nicht anwendbar; z. B. sind die Begriffe von Supremum und Infimum sowie Maximum und Minimum für Teilmengen von komplexen Zahlen nicht definiert.

6.B Eigenschaften der komplexen Zahlen

Auch wenn $\mathbb{C}$ kein geordneter Körper ist, haben wir in Definition 6.3 (b) bereits wie für $\mathbb{R}$ auch für $\mathbb{C}$ eine Betragsfunktion eingeführt, die immer reelle Werte annimmt und es uns somit erlaubt, komplexe Zahlen *betragsmäßig* miteinander zu vergleichen. Wir wollen nun sehen, dass diese komplexe Betragsfunktion in der Tat sogar die gleichen Eigenschaften wie die reelle Betragsfunktion in Lemma 4.18 (a) und (c) hat, auch wenn der Beweis dafür in $\mathbb{C}$ ganz anders ist als in $\mathbb{R}$.

Lemma 6.9 (Eigenschaften der komplexen Konjugation und Betragsfunktion)**.** *Für alle $z_1, z_2 \in \mathbb{C}$ gilt*

(a) $\overline{\overline{z_1}} = z_1$, $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$ *und* $\overline{z_1\, z_2} = \overline{z_1} \cdot \overline{z_2}$;

(b) $|\operatorname{Re} z_1| \le |z_1|$ *und* $|\operatorname{Im} z_1| \le |z_1|$;

(c) $|z_1\, z_2| = |z_1| \cdot |z_2|$;

(d) $|z_1 + z_2| \le |z_1| + |z_2|$ ***(Dreiecksungleichung)****.*

Beweis. Wie üblich sei $z_1 = x_1 + \mathrm{i}y_1$ und $z_2 = x_2 + \mathrm{i}y_2$.

(a) Dies rechnet man einfach nach: Es ist $\overline{\overline{z_1}} = \overline{x_1 - \mathrm{i}y_1} = x_1 + \mathrm{i}y_1 = z_1$, sowie

$$\overline{z_1 + z_2} = \overline{x_1 + x_2 + \mathrm{i}\,(y_1 + y_2)} = x_1 + x_2 - \mathrm{i}\,(y_1 + y_2) = \overline{z_1} + \overline{z_2}$$

und

$$\begin{aligned}\overline{z_1\, z_2} &= \overline{x_1x_2 - y_1y_2 + \mathrm{i}\,(x_1y_2 + y_1x_2)} = x_1x_2 - y_1y_2 - \mathrm{i}\,(x_1y_2 + y_1x_2) = (x_1 - \mathrm{i}y_1)(x_2 - \mathrm{i}y_2) \\ &= \overline{z_1} \cdot \overline{z_2}.\end{aligned}$$

(b) Es gilt

$$|\operatorname{Re} z_1| = |x_1| = \sqrt{x_1^2} \overset{5.34\,(a)}{\le} \sqrt{x_1^2 + y_1^2} = |z_1|;$$

analog folgt dies auch für den Imaginärteil.

(c) Bei der geometrischen Deutung der komplexen Multiplikation in Bemerkung 6.5 haben wir dies bereits anschaulich gesehen; man rechnet es aber auch mit (a) sofort nach: Nach Bemerkung 6.4 ist

$$|z_1\, z_2| = \sqrt{z_1\, z_2 \cdot \overline{z_1\, z_2}} = \sqrt{z_1\, \overline{z_1}\, z_2 \overline{z_2}} \overset{5.34\,(b)}{=} \sqrt{z_1\, \overline{z_1}} \cdot \sqrt{z_2\, \overline{z_2}} = |z_1| \cdot |z_2|.$$

(d) Zunächst ist nach (a) und Bemerkung 6.4

$$|z_1+z_2|^2 = (z_1+z_2)\,\overline{(z_1+z_2)} = z_1\,\overline{z_1}+z_2\,\overline{z_2}+z_1\,\overline{z_2}+z_2\,\overline{z_1} = |z_1|^2+|z_2|^2+z_1\,\overline{z_2}+\overline{z_1\,\overline{z_2}}$$
$$= |z_1|^2+|z_2|^2+2\,\mathrm{Re}(z_1\,\overline{z_2}).$$

Mit (b) können wir nun den dabei auftretenden Realteil abschätzen durch

$$\mathrm{Re}(z_1\,\overline{z_2}) \overset{4.18\,(b)}{\leq} |\mathrm{Re}(z_1\,\overline{z_2})| \leq |z_1\,\overline{z_2}| \overset{(c)}{=} |z_1|\cdot|\overline{z_2}| = |z_1|\cdot|z_2|,$$

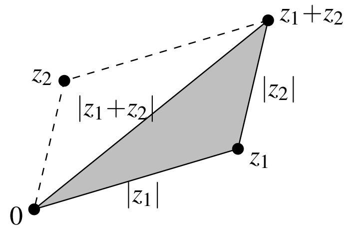
und erhalten so

$$|z_1+z_2|^2 \leq |z_1|^2+|z_2|^2+2\,|z_1|\cdot|z_2| = (|z_1|+|z_2|)^2.$$

Wurzelziehen liefert nun die Behauptung. □

Bemerkung 6.10.

(a) Die Dreiecksungleichung hat eine sehr anschauliche Bedeutung, die auch ihren Namen erklärt: Nach der geometrischen Interpretation der Addition komplexer Zahlen aus Bemerkung 6.5 besagt sie einfach, dass eine Seite in einem Dreieck (wie $|z_1+z_2|$ im Bild rechts) höchstens so lang ist wie die Summe der beiden anderen (hier $|z_1|$ und $|z_2|$).



(b) Wenn ihr gleichzeitig die Parallelvorlesung „Algebraische Strukturen“ hört, werdet ihr sicher sehen, dass Lemma 6.9 (a) gerade besagt, dass die komplexe Konjugation $z \mapsto \overline{z}$ ein Gruppenhomomorphismus von $(\mathbb{C},+)$ nach $(\mathbb{C},+)$ und von $(\mathbb{C}\backslash\{0\},\cdot)$ nach $(\mathbb{C}\backslash\{0\},\cdot)$ ist. Zusammen macht dies die komplexe Konjugation zu einem Körperhomomorphismus von $\mathbb{C}$ nach $\mathbb{C}$ (in der Tat sogar zu einem Körperisomorphismus, da die Abbildung $\mathbb{C}\to\mathbb{C}$, $z\mapsto\overline{z}$ natürlich bijektiv ist).

Wie schon am Anfang dieses Kapitels erwähnt, besteht aber die wesentliche Eigenschaft der komplexen Zahlen darin, dass in $\mathbb{C}$ jede Polynomfunktion eine Nullstelle besitzt. Beachte, dass dies ganz und gar nicht offensichtlich ist – da man jede komplexe Zahl als $x+\mathrm{i}y$ mit $x,y\in\mathbb{R}$ und $\mathrm{i}^2=-1$ schreiben kann, sieht es ja eher so aus, als ob wir durch den Übergang von $\mathbb{R}$ nach $\mathbb{C}$ nur eine „Quadratwurzel aus -1“ hinzugefügt haben, also nur der Polynomfunktion $z^2+1=0$ (oder bestenfalls noch anderen quadratischen Polynomfunktionen) eine Nullstelle gegeben haben. Dass dies in der Tat auch für Polynomfunktionen beliebigen Grades gilt, und zwar sogar noch, wenn sie auch komplexe Koeffizienten haben dürfen, ist der Inhalt des sogenannten Fundamentalsatzes der Algebra:

Satz 6.11 (**Fundamentalsatz der Algebra**). *Jede nicht konstante komplexe Polynomfunktion hat eine Nullstelle in $\mathbb{C}$.*

Beweisidee. Es gibt mehrere (völlig) verschiedene Möglichkeiten, den Fundamentalsatz der Algebra zu beweisen. Leider sind alle diese Beweise für uns aber momentan noch zu schwierig, und so muss ich euch für einen exakten Beweis dieses Satzes auf weiterführende Vorlesungen vertrösten – in den Vorlesungen „Einführung in die Funktionentheorie“, „Einführung in die Algebra“ und „Einführung in die Topologie“ könnt ihr z. B. drei ganz verschiedene Beweise dieses Satzes sehen. Wir können aber auch jetzt zumindest schon eine *Beweisidee* angeben, die hoffentlich dafür ausreicht, dass ihr den Satz glaubt und ein Gefühl dafür bekommt, warum er richtig ist.

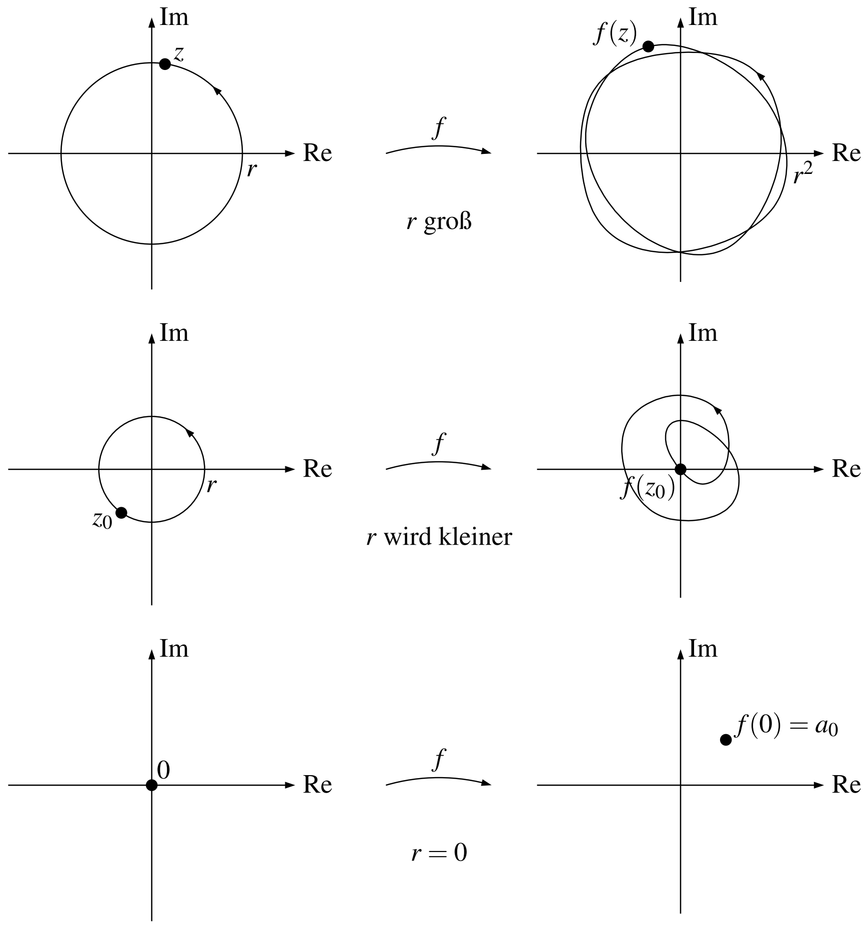
Es sei dazu f eine komplexe, nicht konstante Polynomfunktion, die wir der Einfachheit halber natürlich als normiert annehmen können. Es ist also

$$f(z) = z^n + a_{n-1}z^{n-1}+\cdots+a_1z+a_0$$

für gewisse $n\in\mathbb{N}_{>0}$ und $a_0,\dots,a_{n-1}\in\mathbb{C}$. Wie können wir uns eine solche Funktion grafisch vorstellen? Da ihre Start- und Zielmenge $\mathbb{C}$ ist, können wir ihren Graphen, der ja dann in $\mathbb{C}\times\mathbb{C}=\mathbb{R}^2\times\mathbb{R}^2=\mathbb{R}^4$ liegt, nicht mehr wirklich zeichnen. In den Bildern unten haben wir daher den Startraum $\mathbb{C}$ links und den Zielraum $\mathbb{C}$ rechts dargestellt, und für einige Punkte im Startraum die zugehörigen Bildpunkte im Zielraum eingezeichnet.

Als Erstes wählen wir uns einmal eine feste, sehr große Zahl $r \in \mathbb{R}_{>0}$ und schauen, was passiert, wenn wir mit z den Kreis um 0 mit Radius r durchlaufen. Wenn unsere Funktion einfach $z \mapsto z^n$ wäre, dann wüssten wir genau, wie $f(z)$ auf dieser Kreislinie aussehen würde: Da bei der komplexen Multiplikation nach Bemerkung 6.5 ja gerade Beträge multipliziert und Winkel addiert werden, ist die n-te Potenz einer komplexen Zahl mit Betrag r und Winkel α genau die Zahl mit Betrag r^n und Winkel $n\,\alpha$. Läuft also z einmal beim Radius r im Kreis herum, d. h. α von 0 bis 2π, so läuft z^n beim Radius r^n genau n-mal im Kreis herum, nämlich mit Winkel $n\alpha$ von 0 bis $2n\pi$.

Nun ist unsere Polynomfunktion zwar nicht wirklich genau $z \mapsto z^n$, aber für sehr große Beträge von z ist der Term z^n in $f(z)$ mit der höchsten z-Potenz natürlich betragsmäßig viel größer als die anderen Terme $a_{n-1}z^{n-1} + \cdots + a_1 z + a_0$. Anschaulich bedeutet das, dass $f(z)$ immer „in der Nähe" von z^n ist. Wenn also z^n beim Radius r^n insgesamt n-mal auf einer exakten Kreislinie herumläuft, wird $f(z)$ ein klein wenig von diesem Weg abweichen, aber letztlich immer noch n-mal um den Ursprung herumlaufen. Das Bild unten zeigt in der ersten Zeile einen solchen möglichen Weg für $n = 2$, bei dem also $f(z)$ in einem ungefähren Abstand von r^2 zweimal um den Ursprung läuft, während z einmal auf dem Kreis mit Radius r entlang läuft.



Was passiert nun, wenn wir den Radius r des Kreises für z langsam kleiner machen und zu schließlich 0 werden lassen, so wie im Bild von oben nach unten dargestellt? Natürlich wird sich dann auch der von $f(z)$ durchlaufene Weg in irgendeiner Form langsam ändern. Wir können nicht viel darüber aussagen, wie diese Änderung genau aussieht – klar ist die Situation aber natürlich, wenn der Radius wie in der unteren Zeile des Bildes gleich 0 geworden ist: Dann ist der Kreis für z zu einem Punkt zusammengeschrumpft, und folglich muss natürlich auch der Weg von $f(z)$ von der ursprünglichen Schleife zu einem Punkt (nämlich zum Punkt $f(0) = a_0$) zusammenschrumpfen. Aber es ist anschaulich klar, dass man einen geschlossenen Weg, der ursprünglich n-mal um den Ursprung herumgelaufen ist, nicht auf einen Punkt zusammenziehen kann, ohne ihn dabei mindestens einmal über den Nullpunkt zu ziehen. Und genau an so einer Stelle, wo der Weg für $f(z)$ den Nullpunkt trifft, haben wir natürlich, was wir wollen: eine Nullstelle z_0 von f, so wie in der mittleren Zeile oben im Bild. □

Auch wenn diese Beweisidee jetzt hoffentlich sehr anschaulich war, wäre es doch noch ein sehr weiter Weg für uns, diese Argumente zu einem exakten Beweis zu machen. Ein wichtiger fehlender Punkt ist z. B., dass wir irgendwie formalisieren müssten, was es genau heißt, dass „sich $f(z)$ langsam ändert, wenn sich z langsam ändert". Denn nur wenn sich der Weg für $f(z)$ oben langsam und kontinuierlich ändert, können wir schließen, dass wir ihn irgendwann einmal über den Nullpunkt ziehen müssen.

Bemerkung 6.12.

(a) Nach Satz 3.19 (a) folgt durch wiederholte Anwendung des Fundamentalsatzes der Algebra sofort, dass jede komplexe Polynomfunktion komplett in Linearfaktoren zerfällt, dass sich also jede solche Polynomfunktion f mit $\deg f = n \in \mathbb{N}_{>0}$ als

$$f(z) = c\,(z - z_1) \cdot \cdots \cdot (z - z_n)$$

für gewisse $c, z_1, \dots, z_n \in \mathbb{C}$ mit $c \neq 0$ schreiben lässt. Manchmal wird in der Literatur auch diese Aussage als Fundamentalsatz der Algebra bezeichnet.

(b) Der Fundamentalsatz der Algebra garantiert uns zwar die Existenz einer Nullstelle einer nicht konstanten komplexen Polynomfunktion, er sagt uns aber nicht, wie wir eine solche Nullstelle konkret finden können. In der Tat haben wir ja schon in Bemerkung 3.21 erwähnt, dass es zur exakten Bestimmung von Nullstellen von Polynomfunktionen im Allgemeinen nur für kleine Grade explizite Formeln gibt. Einen sehr einfachen und oft vorkommenden Fall, in dem sich die Nullstellen jedoch schnell finden lassen, wollen wir hier kurz erwähnen:

Beispiel 6.13. Es sei $f\colon \mathbb{C} \to \mathbb{C}$ eine Polynomfunktion mit $\deg f = 2$ und *reellen* Koeffizienten, das der Einfachheit halber wieder normiert sei, d. h. es sei $f(z) = z^2 + pz + q$ für gewisse $p, q \in \mathbb{R}$. In diesem Fall lassen sich die (komplexen) Nullstellen von f schnell berechnen: Aus $z^2 + pz + q = 0$ folgt durch quadratische Ergänzung

$$\left(z + \frac{p}{2}\right)^2 = \frac{p^2}{4} - q =: D.$$

Für $D \geq 0$ ergeben sich durch Wurzelziehen natürlich die (reellen) Nullstellen $-\frac{p}{2} \pm \sqrt{D}$. Für $D < 0$ gibt es keine reellen Lösungen, aber wegen $\mathrm{i}^2 = -1$ erhalten wir stattdessen die beiden komplexen Lösungen $-\frac{p}{2} \pm \mathrm{i}\sqrt{-D}$.

Aufgabe 6.14. Für $n = 1, 2, 3$ bestimme und skizziere man die Menge aller $z \in \mathbb{C}$, für die die Gleichung $2 \operatorname{Im} z \cdot \operatorname{Im} \frac{1}{z} = n$ gilt.

Aufgabe 6.15.

(a) Zeige (ohne Verwendung des Fundamentalsatzes der Algebra), dass auch in $\mathbb{C}$ Quadratwurzeln existieren, also dass es zu jedem $w \in \mathbb{C}$ ein $z \in \mathbb{C}$ gibt mit $z^2 = w$.

(b) Beweise den Fundamentalsatz der Algebra für Polynome vom Grad 2.

Aufgabe 6.16. Stelle die folgenden Zahlen in der Form $x + \mathrm{i}y$ mit $x, y \in \mathbb{R}$ dar:

(a) $z = \frac{2+\mathrm{i}}{1-\mathrm{i}}$;

(b) $z = \left(\frac{1+\mathrm{i}}{\sqrt{2}}\right)^{-2023}$;

(c) alle Lösungen der Gleichung $z^4 + z^2 + 1 = 0$;

(d) alle $z \in \mathbb{C}$ mit $\left|\frac{z-1}{z-\mathrm{i}}\right| < 1$.

6.C Reelle und komplexe Folgen

Auch wenn $\mathbb{C}$ nach Bemerkung 6.8 kein geordneter Körper ist, können wir mit Hilfe der Betragsfunktion aus Definition 6.3 sagen, was es bedeutet, dass sich eine Folge komplexer Zahlen einem Grenzwert annähert. In der Tat können wir die reelle Grenzwertdefinition 5.1 (b) wörtlich auf den reellen Fall übertragen:

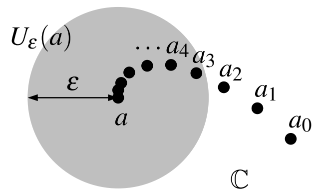
Definition 6.17 (Grenzwerte komplexer Folgen)**.** Eine komplexe Zahl a heißt **Grenzwert** einer Folge $(a_n)_n$ in $\mathbb{C}$, wenn

$$\forall \varepsilon \in \mathbb{R}_{>0}\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0 : |a_n - a| < \varepsilon.$$ 13

Bemerkung 6.18 (Anschauliche Deutung des Grenzwertbegriffs in $\mathbb{C}$)**.** Definieren wir für ein *komplexes* $a \in \mathbb{C}$ und *reelles* $\varepsilon \in \mathbb{R}_{>0}$ analog zu Bemerkung 5.2 wieder die **ε-Umgebung** von a als

$$U_\varepsilon(a) := \{x \in \mathbb{C} : |x - a| < \varepsilon\},$$

so ist dies wie im Bild rechts nun ein Kreis in der komplexen Ebene mit Mittelpunkt a und Radius ε. Die Grenzwertbedingung besagt weiterhin, dass in jeder ε-Umgebung fast alle Folgenglieder liegen, und kann damit wieder so interpretiert werden, dass sich die Folgenglieder immer mehr dem Grenzwert nähern.



Bemerkung 6.19 (Übertragung der Grenzwerteigenschaften von $\mathbb{R}$ auf $\mathbb{C}$)**.** Aufgrund der gleichen Grenzwertdefinition 5.1 bzw. 6.17 sowie der gleichen Eigenschaften der Betragsfunktion aus Lemma 4.18 bzw. 6.9 (insbesondere der Dreiecksungleichung) gelten sehr viele Resultate über reelle Grenzwerte genauso auch für komplexe. *In der Tat übertragen sich alle Definitionen und Sätze aus Abschnitt 5.A mit wörtlich den gleichen Beweisen unmittelbar auf komplexe Folgen $(a_n)_n$:*

(a) die Definitionen von Nullfolgen, Häufungspunkten und beschränkten Folgen (wobei die Schranke s mit $|a_n| \leq s$ für alle n natürlich weiterhin reell bleibt);

(b) die Eindeutigkeit des Grenzwerts (und damit die Notation $\lim\limits_{n\to\infty} a_n$), der Grenzwert $\lim\limits_{n\to\infty} q^n = 0$ für $|q| < 1$, die Beschränktheit konvergenter Folgen, die Grenzwertsätze und die äquivalenten Charakterisierungen von Häufungspunkten.

Ihr könnt euch gerne selbst davon überzeugen und Abschnitt 5.A noch einmal unter der Voraussetzung durchlesen, dass alle Folgen nun komplex sind – es werden keinerlei Änderungen erforderlich sein. Wir werden die Ergebnisse dieses Abschnitts daher im Folgenden auch im Komplexen verwenden, ohne jedes Mal wieder darauf hinzuweisen. Um solche Aussagen in Zukunft für den reellen und komplexen Fall gleichzeitig aufschreiben zu können, vereinbaren wir:

Im Folgenden steht $\mathbb{K}$ immer für einen der Körper $\mathbb{R}$ oder $\mathbb{C}$.

Die Inhalte der Abschnitte 5.B und 5.C benötigen jedoch wirklich einen geordneten Körper und nicht nur das Konzept des Abstandes zweier Zahlen. Ergebnisse wie das Monotoniekriterium, die Intervallschachtelung oder die Existenz eines Limes superior haben daher keine Entsprechung im Komplexen.

Aufgabe 6.20. Es sei $(a_n)_n$ eine Folge in $\mathbb{C}$. Beweise, dass $(a_n)_n$ genau dann gegen die komplexe Zahl a konvergiert, wenn die Folgen $(\operatorname{Re} a_n)_n$ und $(\operatorname{Im} a_n)_n$ ihrer Real- und Imaginärteile gegen $\operatorname{Re} a$ bzw. $\operatorname{Im} a$ konvergieren.

Da wir die Konvergenzkriterien aus Abschnitt 5.B, mit denen wir die Konvergenz einer Folge auch ohne Kenntnis oder gleichzeitige Berechnung des Grenzwerts beweisen konnten, in $\mathbb{C}$ nicht mehr zur Verfügung haben, wollen wir nun noch zwei sehr wichtige Konvergenzkriterien behandeln, die sowohl in $\mathbb{R}$ als auch in $\mathbb{C}$ gelten.

Satz 6.21 (Satz von Bolzano-Weierstraß). *Jede beschränkte Folge $(a_n)_n$ in $\mathbb{K}$ besitzt einen Häufungspunkt.*

Beweis. Für den Fall $\mathbb{K} = \mathbb{R}$ haben wir dies bereits in Folgerung 5.48 gesehen: Der Limes superior von $(a_n)_n$ ist ein Häufungspunkt.

Im Fall $\mathbb{K} = \mathbb{C}$ stellen wir zunächst fest, dass nach Lemma 6.9 (b) mit $(a_n)_n$ auch die reellen Folgen $(\operatorname{Re} a_n)_n$ und $(\operatorname{Im} a_n)_n$ beschränkt sind. Nach dem Satz von Bolzano-Weierstraß für $\mathbb{R}$ (den wir ja schon bewiesen haben) gibt es also zunächst eine Teilfolge von $(a_n)_n$, in der die Realteile gegen ein $a \in \mathbb{R}$ konvergieren, und dann *innerhalb dieser Teilfolge* eine weitere Teilfolge, in der auch die Imaginärteile gegen ein $b \in \mathbb{R}$ konvergieren. Nach Aufgabe 6.20 konvergiert diese Teilfolge dann gegen $a + ib$, d. h. $a + ib$ ist ein Häufungspunkt von $(a_n)_n$. □

Das letzte wichtige Konvergenzkriterium, das wir hier beweisen wollen – das sogenannte Cauchy-Kriterium – sieht fast so aus wie die Definition der Konvergenz. Der Unterschied besteht lediglich darin, dass wir nicht verlangen, dass sich die Folgenglieder *einem gegebenen Grenzwert* immer weiter annähern, sondern nur, dass sie sich *untereinander* beliebig nahe kommen. Auf diese Art müssen wir den Grenzwert der Folge also wiederum nicht vorher kennen, um das Kriterium anwenden zu können. Im Gegensatz zu unseren bisherigen Kriterien hat das Cauchy-Kriterium aber auch noch den weiteren entscheidenden Vorteil, dass es *äquivalent* zur Konvergenz ist und somit auch zum Beweis der Divergenz einer Folge verwendet werden kann.

Die Eigenschaft, dass sich die Folgenglieder untereinander immer näher kommen, sieht formal wie folgt aus.

Definition 6.22 (Cauchyfolgen)**.** Eine Folge $(a_n)_n$ in $\mathbb{K}$ heißt **Cauchyfolge**, wenn

$$\forall \varepsilon \in \mathbb{R}_{>0} \ \exists n_0 \in \mathbb{N} \ \forall m,n \geq n_0 : |a_m - a_n| < \varepsilon.$$

Bemerkung 6.23. Jede konvergente Folge ist eine Cauchyfolge: Ist $(a_n)_n$ konvergent mit Grenzwert $a \in \mathbb{K}$, so gibt es zu jedem $\varepsilon \in \mathbb{R}_{>0}$ ein $n_0 \in \mathbb{N}$ mit $|a_n - a| < \frac{\varepsilon}{2}$ für alle $n \geq n_0$. Dann gilt nach der Dreiecksungleichung aber auch für alle $m,n \geq n_0$

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

d. h. $(a_n)_n$ ist eine Cauchyfolge.

Diese Tatsache, dass eine konvergente Folge immer eine Cauchyfolge ist, ist also sehr einfach zu zeigen und wäre z. B. auch in $\mathbb{Q}$ richtig: Wenn die Folgenglieder immer mehr gegen einen Grenzwert streben, müssen sie sich natürlich auch untereinander immer näher kommen. Die Umkehrung dagegen ist weit weniger klar: Da $\mathbb{Q}$ ja „Löcher" auf der Zahlengeraden hat, könnte es ja sein, dass sich die Glieder einer rationalen Folge zwar immer näher kommen, aber sich an einem solchen Loch häufen und daher kein Grenzwert der Folge in $\mathbb{Q}$ existiert. Dass so etwas in $\mathbb{R}$ oder $\mathbb{C}$ nicht passieren kann, weil es dort keine solchen Löcher gibt, wird als *Vollständigkeit* dieser Körper bezeichnet (siehe auch Definition **??**). Um dies zu zeigen, benötigen wir zunächst ein kleines Lemma analog zu Lemma 5.8:

Lemma 6.24. *Jede Cauchyfolge in $\mathbb{K}$ ist beschränkt.*

Beweis. Nicht nur die Aussage, sondern auch ihr Beweis ist völlig analog zu Lemma 5.8: Es sei $(a_n)_n$ eine Cauchyfolge in $\mathbb{K}$. Dann gibt es zu $\varepsilon = 1$ ein n_0, so dass $|a_m - a_n| < \varepsilon = 1$ für alle $m,n \geq n_0$ ist. Insbesondere gilt dies also für $m = n_0$, und damit erhalten wir nach der Dreiecksungleichung für alle $n \geq n_0$

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

Damit folgt nun aber $|a_n| \leq s$ für alle $n \in \mathbb{N}$, wenn wir

$$s := \max(|a_0|, |a_1|, \dots, |a_{n_0-1}|, 1 + |a_{n_0}|)$$

setzen. Also ist $(a_n)_n$ beschränkt. □

Satz 6.25 (**Cauchy-Kriterium** für Folgen, **Vollständigkeit** von $\mathbb{K}$)**.** *Jede Cauchyfolge in $\mathbb{K}$ konvergiert.*

Nach Bemerkung 6.23 konvergiert eine Folge in $\mathbb{K}$ also genau dann, wenn sie eine Cauchyfolge ist.

Beweis. Es sei $(a_n)_n$ eine Cauchyfolge in $\mathbb{K}$. Dann ist $(a_n)_n$ nach Lemma 6.24 beschränkt und besitzt damit nach dem Satz 6.21 von Bolzano-Weierstraß einen Häufungspunkt a. Wir behaupten, dass $(a_n)_n$ sogar schon gegen a konvergiert.

Um dies zu zeigen, sei $\varepsilon \in \mathbb{R}_{>0}$ beliebig. Da $(a_n)_n$ eine Cauchyfolge ist, gibt es ein $n_0 \in \mathbb{N}$, so dass

$$|a_m - a_n| < \frac{\varepsilon}{2} \quad \text{für alle } m, n \geq n_0.$$

Weil a ein Häufungspunkt von $(a_n)_n$ ist, gilt nach Lemma 5.21 weiterhin

$$|a_m - a| < \frac{\varepsilon}{2} \quad \text{für unendlich viele } m,$$

und damit insbesondere für ein $m \geq n_0$. Wir können diese beiden Ungleichungen also miteinander kombinieren und erhalten für ein solches m nach der Dreiecksungleichung

$$|a_n - a| = |a_n - a_m + a_m - a| \leq |a_n - a_m| + |a_m - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

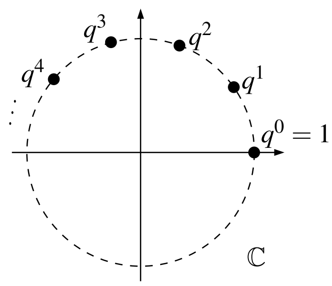
für alle $n \geq n_0$. Damit ist $(a_n)_n$ konvergent gegen a. □

Beispiel 6.26 (Noch einmal die geometrische Folge)**.** Wir betrachten noch einmal die geometrische Folge $(q^n)_n$ für ein $q \in \mathbb{K}$. Aus Beispiel 5.3 (c) und 5.9 (b) wissen wir bereits, dass $(q^n)_n$ für $|q| < 1$ gegen 0 konvergiert und für $|q| > 1$ divergiert. Außerdem ist klar, dass die Folge für $q = 1$ konstant ist und damit konvergiert. Wir zeigen nun mit dem Cauchy-Kriterium in den übrigen Fällen, also wenn $|q| = 1$ und $q \neq 1$, dass die Folge divergiert. Dazu müssen wir also beweisen, dass $(q^n)_n$ keine Cauchyfolge ist, d. h. (nach den Regeln der Negation aus Bemerkung 1.8)

$$\exists \varepsilon > 0 \; \forall n_0 \in \mathbb{N} \; \exists n, m \geq n_0 : |q^n - q^m| \geq \varepsilon.$$

Um dies zu zeigen, setzen wir $\varepsilon := |q-1| > 0$. Nun sei $n_0 \in \mathbb{N}$ beliebig; wir setzen dann $n = n_0 + 1$ und $m = n_0$. Mit diesen Werten folgt

$$|q^n - q^m| = |q^{n_0+1} - q^{n_0}| = |q^{n_0}(q-1)| = \underbrace{|q|^{n_0}}_{=1} \cdot \underbrace{|q-1|}_{=\varepsilon} = \varepsilon.$$



Also ist $(q^n)_n$ keine Cauchyfolge und damit nach Satz 6.25 nicht konvergent. Das Bild rechts illustriert dies: Nach der geometrischen Interpretation der komplexen Multiplikation aus Bemerkung 6.5 läuft die Folge für $|q| = 1$ und $q \neq 1$ „mit konstanter Geschwindigkeit“ auf dem Einheitskreis herum und nähert sich somit keinem Grenzwert beliebig an.

Aufgabe 6.27. Es sei $(a_n)_n$ eine Folge in $\mathbb{K}$. Man zeige: Gibt es ein $q \in \mathbb{R}_{\geq 0}$ mit $q < 1$, so dass

$$|a_{n+1} - a_n| \leq q^n \quad \text{für alle } n \in \mathbb{N},$$

so ist $(a_n)_n$ eine Cauchyfolge.

Aufgabe 6.28. Für ein fest gegebenes $c \in \mathbb{C}$ mit $|c| < \frac{1}{4}$ definieren wir eine komplexe Folge $(a_n)_n$ rekursiv durch

$$a_0 = 0 \quad \text{und} \quad a_{n+1} = a_n^2 + c \quad \text{für alle } n \in \mathbb{N}.$$

Zeige, dass $(a_n)_n$ konvergiert.

(Hinweis: Zeige zunächst, dass $\frac{1}{4} + |c|$ eine obere Schranke für die Beträge aller Folgenglieder ist.)

7. Reihen

Wir wollen uns nun mit einem speziellen Typ von Folgen beschäftigen, der in der Praxis sehr häufig vorkommt: nämlich Folgen, die in der Form

$$(a_0,\ a_0+a_1,\ a_0+a_1+a_2,\ \dots)$$

gegeben sind, deren Grenzwert wir also anschaulich als die „unendliche Summe" $a_0+a_1+a_2+\cdots$ auffassen können. Derartige Folgen bezeichnet man als Reihen. Wir werden solche Reihen im reellen und komplexen Fall gleichzeitig betrachten und arbeiten daher im Folgenden in der Regel über dem Körper $\mathbb{K}$ wie in Bemerkung 6.19.

7.A Grenzwerte von Reihen

Da Reihen letztlich nichts anderes als spezielle Folgen sind, können wir die Definition und die ersten Eigenschaften von Folgen und Grenzwerten natürlich unmittelbar auf unsere neue Situation übertragen. Dies wollen wir nun im ersten Abschnitt dieses Kapitels tun.

Definition 7.1 (Reihen)**.** Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{K}$. Dann heißt die Folge $(s_N)_{N\in\mathbb{N}}$ mit

$$s_N = \sum_{n=0}^{N} a_n = a_0+a_1+\cdots+a_N$$

die Folge der **Partialsummen** von $(a_n)_n$ bzw. die zu $(a_n)_n$ gehörige **Reihe**. Wir bezeichnen sowohl diese Reihe als auch ihren Grenzwert $\lim\limits_{N\to\infty} s_N$ (sofern er existiert) mit

$$\sum_{n=0}^{\infty} a_n \quad \text{bzw.} \quad a_0+a_1+a_2+\cdots.$$

Genau wie bei Folgen kann auch eine Reihe bei einem anderen Startindex $n_0\in\mathbb{Z}$ als bei 0 anfangen; in diesem Fall schreiben wir sie natürlich als

$$\sum_{n=n_0}^{\infty} a_n \quad \text{bzw.} \quad a_{n_0}+a_{n_0+1}+a_{n_0+2}+\cdots.$$

Bemerkung 7.2.

(a) Da jede Reihe nach Definition eine Folge ist, übertragen sich die Begriffe Konvergenz und Divergenz, Beschränktheit usw. aus Kapitel 5 direkt auf Reihen.

(b) Die Doppelbelegung des Symbols $\sum_{n=0}^{\infty} a_n$ sowohl für die Reihe (also die Folge ihrer Partialsummen) als auch für ihren Grenzwert ist zwar mathematisch unschön, aber in der Literatur so fest verankert, dass wir hier nicht davon abweichen wollen. Es sollte dadurch keine Verwirrung entstehen: Wenn wir von Eigenschaften einer Folge reden, also z. B. sagen, dass $\sum_{n=0}^{\infty} a_n$ konvergiert oder divergiert, so meinen wir natürlich die Partialsummenfolge – während z. B. in Gleichungen der Form $\sum_{n=0}^{\infty} a_n = a$ der Grenzwert der Reihe gemeint ist. Wenn Verwechslungen zu befürchten sind, können wir natürlich auch immer die eindeutige Schreibweise $\left(\sum_{n=0}^{N} a_n\right)_{N\in\mathbb{N}}$ für die Reihe und $\lim\limits_{N\to\infty} \sum_{n=0}^{N} a_n$ für ihren Grenzwert benutzen.

Beispiel 7.3.

(a) (**Unendliche geometrische Reihe**) Wir betrachten die Reihe $\sum_{n=0}^{\infty} q^n$ für ein $q\in\mathbb{K}$. Für $q=1$ ist diese Reihe $1+1+1+\dots$ natürlich unbeschränkt und damit divergent. Ansonsten haben wir in Satz 4.1 gesehen, dass

$$\sum_{n=0}^{N} q^n = \frac{1-q^{N+1}}{1-q}.$$

Der Grenzwert für $N \to \infty$ ergibt sich nun sofort aus Beispiel 6.26: Da $\lim\limits_{N\to\infty} q^{N+1}$ nur für $|q| < 1$ existiert und dann gleich 0 ist, erhalten wir also

$$\sum_{n=0}^{\infty} q^n = \lim_{N\to\infty} \sum_{n=0}^{N} q^n = \frac{1}{1-q} \qquad \text{für } |q| < 1, \tag{$*$}$$

während die Reihe in allen anderen Fällen divergiert.

Ein interessanter konkreter Fall dieser Reihe ist die Frage, ob die Dezimalzahl $0,9999\ldots$ gleich 1 oder „etwas kleiner" als 1 ist. Dies können wir nun beantworten, denn die einzig mögliche mathematisch korrekte Definition dieser Zahl ist natürlich die geometrische Reihe

$$0,9999\ldots = \sum_{n=1}^{\infty} 9 \cdot 10^{-n} = \frac{9}{10} \cdot \sum_{n=0}^{\infty} \left(\frac{1}{10}\right)^n \overset{(*)}{=} \frac{9}{10} \cdot \frac{1}{1-\frac{1}{10}} = 1.$$

Die Zahl $0,9999\ldots$ ist daher wirklich *gleich* 1 – in diesem Fall ist die Dezimaldarstellung einer reellen Zahl also nicht eindeutig.

(b) (Teleskopreihen) Wir wollen den Grenzwert der Reihe $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ bestimmen. Normalerweise lassen sich derartige Reihen nicht ohne weiteres berechnen, aber in diesem ganz speziellen Fall können wir einen Trick anwenden: Wegen $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$ können wir die Partialsummen der Reihe schreiben als

$$\begin{aligned}
\sum_{n=1}^{N} \frac{1}{n(n+1)} &= \sum_{n=1}^{N} \left(\frac{1}{n} - \frac{1}{n+1}\right) \\
&= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \cdots + \left(\frac{1}{N} - \frac{1}{N+1}\right) \\
&= \frac{1}{1} - \frac{1}{N+1}.
\end{aligned}$$

Derartige Reihen, bei denen sich in den Partialsummen durch geeignete Differenzen alle Terme bis auf einen Start- und Endterm wegheben, bezeichnet man als **Teleskopreihen** (weil die Summe sozusagen wie ein Teleskop „zusammengeschoben" werden kann). Der Grenzwert der Reihe lässt sich dann natürlich einfach berechnen; in diesem Fall ist er

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = \lim_{N\to\infty} \sum_{n=1}^{N} \frac{1}{n(n+1)} = \lim_{N\to\infty} \left(1 - \frac{1}{N+1}\right) = 1.$$

(c) (**Harmonische Reihe**) Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ divergiert: Für die Partialsummen mit Index $N = 2^k$ gilt

$$\begin{aligned}
\sum_{n=1}^{2^k} \frac{1}{n} &= 1 + \tfrac{1}{2} + \left(\tfrac{1}{3} + \tfrac{1}{4}\right) + \left(\tfrac{1}{5} + \tfrac{1}{6} + \tfrac{1}{7} + \tfrac{1}{8}\right) + \cdots + \left(\tfrac{1}{2^{k-1}+1} + \cdots + \tfrac{1}{2^k}\right) \\
&\geq 1 + \tfrac{1}{2} + \left(\tfrac{1}{4} + \tfrac{1}{4}\right) + \left(\tfrac{1}{8} + \tfrac{1}{8} + \tfrac{1}{8} + \tfrac{1}{8}\right) + \cdots + \left(\tfrac{1}{2^k} + \cdots + \tfrac{1}{2^k}\right) \\
&= 1 + \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} + \cdots + \tfrac{1}{2} \\
&= 1 + \tfrac{k}{2}.
\end{aligned}$$

Da $1 + \frac{k}{2}$ mit k unbeschränkt wächst, ist die gegebene Reihe also unbeschränkt und damit nach Lemma 5.8 divergent. 14

Die folgenden einfachen Rechenregeln für Reihen – die Verträglichkeit mit Summen, Differenzen, Multiplikation mit Konstanten sowie im Fall des Körpers $\mathbb{R}$ mit Ungleichungen – ergeben sich sofort aus denen für Folgen in Kapitel 5.

Lemma 7.4 (Rechenregeln für Reihen). *Es seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ konvergente Reihen in $\mathbb{K}$. Dann gilt:*

(a) $\displaystyle\sum_{n=0}^{\infty} (a_n + b_n) = \sum_{n=0}^{\infty} a_n + \sum_{n=0}^{\infty} b_n$ *und* $\displaystyle\sum_{n=0}^{\infty} (a_n - b_n) = \sum_{n=0}^{\infty} a_n - \sum_{n=0}^{\infty} b_n$.

(b) *Für* $c \in \mathbb{K}$ *ist* $\sum_{n=0}^{\infty} c a_n = c \cdot \sum_{n=0}^{\infty} a_n$.

(c) *Ist* $\mathbb{K} = \mathbb{R}$ *und* $a_n \le b_n$ *für alle* n, *so ist* $\sum_{n=0}^{\infty} a_n \le \sum_{n=0}^{\infty} b_n$.

Beweis. Alle behaupteten Aussagen gelten trivialerweise für die Partialsummen der Reihen (also wenn die Summen bis zu einem festen $N \in \mathbb{N}$ laufen). Übergang zum Grenzwert liefert dann mit den Sätzen 5.13 und 5.24 die Behauptungen. □

Eine analoge direkte Verträglichkeit mit der Multiplikation ist natürlich nicht zu erwarten, weil ja schon für die Partialsummen $\left(\sum_{n=0}^{N} a_n\right) \cdot \left(\sum_{n=0}^{N} b_n\right)$ nicht dasselbe ist wie $\sum_{n=0}^{N} a_n b_n$. Wir werden aber später in Satz 7.34 noch eine Formel für das Produkt von Reihen finden.

Bevor wir nun mit der Herleitung allgemeiner Konvergenzkriterien für Reihen beginnen, wollen wir noch zwei sehr einfache Hilfsaussagen festhalten, die aber dennoch oft nützlich sind. Die erste von ihnen ist so einfach, dass sie üblicherweise als *Trivialkriterium* bezeichnet wird: Eine Reihe kann höchstens dann konvergieren, wenn die aufsummierten Zahlen zumindest gegen 0 konvergieren.

Lemma 7.5 (Trivialkriterium). *Ist die Reihe* $\sum_{n=0}^{\infty} a_n$ *konvergent, so ist* $(a_n)_n$ *eine Nullfolge.*

Beweis. Existiert der Grenzwert $\sum_{n=0}^{\infty} a_n$, so folgt aus den Grenzwertsätzen

$$a_N = \sum_{n=0}^{N} a_n - \sum_{n=0}^{N-1} a_n \to \sum_{n=0}^{\infty} a_n - \sum_{n=0}^{\infty} a_n = 0 \quad \text{für } N \to \infty.$$ □

Beispiel 7.6.

(a) Die Reihe $\sum_{n=0}^{\infty} \frac{2n^2}{n^2+1}$ ist divergent, denn nach Beispiel 5.14 ist $\lim\limits_{n\to\infty} \frac{2n^2}{n^2+1} = 2 \ne 0$.

(b) Das Trivialkriterium ist nicht umkehrbar: So ist z. B. zwar $\left(\frac{1}{n}\right)_n$ eine Nullfolge, aber die harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ nach Beispiel 7.3 (c) trotzdem divergent. Man kann mit diesem Kriterium also immer nur die Divergenz einer Reihe nachweisen, aber nie die Konvergenz.

Dieses Beispiel zeigt auch noch etwas anderes: Bezeichnen wir mit $a_n = \sum_{k=1}^{n} \frac{1}{k}$ die Partialsummen der harmonischen Reihe, so ist die Folge $(a_n)_n$ zwar divergent, aber die Folge $(a_{n+1} - a_n)_n = \left(\frac{1}{n+1}\right)_n$ konvergiert trotzdem gegen 0, d. h. es gilt

$$\forall \varepsilon \in \mathbb{R}_{>0} \ \exists n_0 \in \mathbb{N} \ \forall n \ge n_0 : |a_{n+1} - a_n| < \varepsilon.$$

Um die Äquivalenz zwischen konvergenten Folgen und Cauchyfolgen zu erhalten, genügt es in der Definition 6.22 einer Cauchyfolge also nicht, zwei *benachbarte* Folgenglieder a_n und a_{n+1} miteinander zu vergleichen, sondern wir müssen zwei *beliebige* Folgenglieder a_m und a_n (mit $m, n \ge n_0$) nehmen!

Lemma 7.7. *Eine Reihe* $\sum_{n=0}^{\infty} a_n$ *mit* $a_n \in \mathbb{R}_{\ge 0}$ *für alle* $n \in \mathbb{N}$ *ist genau dann konvergent, wenn sie beschränkt ist.*

Beweis. Da alle aufsummierten Zahlen reell und nicht-negativ sind, ist die Folge ihrer Partialsummen monoton wachsend. Für eine reelle, monoton wachsende Folge ist die Konvergenz nach Lemma 5.8 und dem Monotoniekriterium aus Satz 5.28 aber äquivalent zur Beschränktheit. □

7.B Konvergenzkriterien für Reihen

Wie im Fall von Folgen im letzten Kapitel wollen wir nun einige Kriterien herleiten, mit denen man die Konvergenz einer Reihe beweisen kann, ohne ihren Grenzwert zu kennen. Dabei bleiben natürlich alle Ergebnisse aus Abschnitt 5.B unverändert anwendbar, da Reihen ja letztlich auch nur Folgen sind. Es gibt aber einige zusätzliche Kriterien, die speziell auf den Fall von Reihen zugeschnitten und meistens einfacher zu überprüfen sind. Wir beginnen dabei mit einem Kriterium für reelle Reihen, in denen abwechselnd positive und negative Glieder aufsummiert werden.

Satz 7.8 (Leibniz-Kriterium). *Ist $(a_n)_n$ eine monoton fallende Nullfolge in $\mathbb{R}_{\geq 0}$, so ist die Reihe*

$$\sum_{n=0}^{\infty} (-1)^n a_n = a_0 - a_1 + a_2 - a_3 \pm \cdots$$

konvergent, und ihre Partialsummen sind abwechselnd obere und untere Schranken für ihren Grenzwert. (Derartige reelle Reihen, bei denen sich das Vorzeichen in der Summe immer abwechselt, nennt man ***alternierend****.)*

Beweis. Es sei $s_N = \sum_{n=0}^{N} (-1)^n a_n$, also $(s_N)_N$ die Folge der Partialsummen der betrachteten Reihe. Da $(a_n)_n$ monoton fallend und die Differenz zweier aufeinander folgender Glieder von $(a_n)_n$ damit nicht negativ ist, ist die Folge $(s_{2N})_N$ der *geraden* Partialsummen monoton fallend: Es gilt

$$s_{2N+2} = s_{2N} \underbrace{- a_{2N+1} + a_{2N+2}}_{\geq 0} \leq s_{2N}.$$

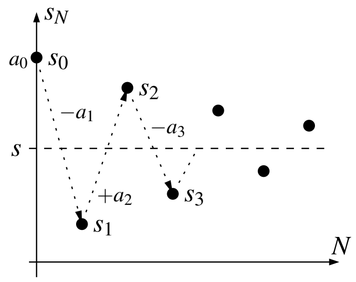
Analog ist die Folge $(s_{2N+1})_N$ der ungeraden Partialsummen monoton wachsend, wie auch das Bild unten rechts zeigt. Damit haben wir ineinander liegende Intervalle

$$[s_1, s_2] \supset [s_3, s_4] \supset [s_5, s_6] \supset \cdots,$$

die eine Intervallschachtelung definieren, da die Länge

$$s_{2N} - s_{2N-1} = a_{2N}$$

dieser Intervalle mit $N \to \infty$ gegen 0 konvergiert. Nach Satz 5.39 konvergieren also die geraden und ungeraden Partialsummen monoton fallend bzw. wachsend gegen den gleichen Grenzwert s. Insbesondere sind die geraden und ungeraden Partialsummen also obere bzw. untere Schranken für s.



Außerdem liegen damit in jeder ε-Umgebung von s fast alle geraden und fast alle ungeraden Partialsummen, und somit konvergiert auch die gesamte Folge der Partialsummen gegen s. □

Beispiel 7.9 (Alternierende harmonische Reihe). Die Reihe

$$\sum_{n=1}^{\infty} (-1)^n \frac{1}{n} = -\frac{1}{1} + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} \mp \cdots$$

ist nach dem Leibniz-Kriterium konvergent, denn $\frac{1}{n}$ ist eine monoton fallende Nullfolge. Ihren Grenzwert können wir momentan noch nicht berechnen (in der Tat ist er gleich $-\log 2$, wie wir in Beispiel 11.15 (a) sehen werden), aber nach Satz 7.8 liegt er sicher zwischen den ersten beiden Partialsummen $-\frac{1}{1} = -1$ und $-\frac{1}{1} + \frac{1}{2} = -\frac{1}{2}$.

Übrigens ist diese Reihe (ganz im Gegensatz z. B. zur Folge aus Beispiel 5.37) eine, die „extrem langsam" konvergiert: Um hier den Grenzwert auf k Nachkommastellen genau zu berechnen, müssen wir natürlich mindestens die ersten 10^k Summanden mitnehmen, denn der 10^k-te Summand ist ja 10^{-k} und ändert somit in jedem Fall noch die k-te Nachkommastelle.

Wir können an dieser alternierenden harmonischen Reihe aber noch eine weitere überraschende Eigenschaft sehen. Dazu sortieren wir die aufzusummierenden Zahlen mal etwas um und schreiben unsere Reihe als

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

Das Prinzip hierbei ist, dass die Terme $(-1)^n \frac{1}{n} \ldots$

- für ungerade n der Reihe nach als erste Summanden in den Klammern stehen,
- für gerade, aber nicht durch 4 teilbare n der Reihe nach als zweite Summanden in den Klammern stehen,
- für durch 4 teilbare n der Reihe nach außerhalb der Klammern stehen.

Es ist klar, dass wir hier wirklich nur die Summanden umsortiert, also keinen vergessen oder doppelt hingeschrieben haben. Rechnen wir jetzt aber mal die Klammern aus, so erhalten wir

$$-\frac{1}{2}+\frac{1}{4}-\frac{1}{6}+\frac{1}{8}-\frac{1}{10}+\frac{1}{12}-\frac{1}{14}+\frac{1}{16}\mp\cdots$$

und damit genau die Hälfte der ursprünglichen Reihe! Da die Reihe nicht den Wert 0 hat (wie wir oben schon gesehen haben, liegt ihr Wert ja zwischen -1 und $-\frac{1}{2}$), haben wir ihren Wert durch das Umsortieren also tatsächlich geändert und müssen damit wohl oder übel feststellen:

Das Umordnen der Summanden in einer konvergenten Reihe kann ihren Grenzwert ändern.

Das ist natürlich extrem lästig, weil uns das sozusagen die Kommutativität der Addition im Fall von unendlichen Summen kaputt macht – was völlig der Intuition widerspricht und natürlich auch beim Rechnen mit solchen Reihen große Probleme bereitet. Glücklicherweise gibt es einen relativ eleganten Ausweg aus dieser Situation: Es gibt eine Eigenschaft von Reihen, die etwas stärker als die normale Konvergenz ist, in vielen Fällen aber dennoch erfüllt ist und die Umsortierbarkeit ohne Änderung des Grenzwerts garantiert. Diese wollen wir jetzt einführen.

Definition 7.10 (Absolute Konvergenz)**.** Eine Reihe $\sum_{n=0}^{\infty} a_n$ in $\mathbb{K}$ heißt **absolut konvergent**, wenn die Reihe $\sum_{n=0}^{\infty} |a_n|$ ihrer Beträge konvergiert, also nach Lemma 7.7 wenn diese Reihe $\sum_{n=0}^{\infty} |a_n|$ beschränkt ist.

(Der Name kommt einfach daher, dass man den Betrag einer Zahl oft auch als *Absolutbetrag* bezeichnet.)

Für Reihen, in denen nur nicht-negative reelle Zahlen aufsummiert werden, stimmen die Begriffe „konvergent“ und „absolut konvergent“ offensichtlich überein. Wir wollen nun sehen, dass der Begriff der absoluten Konvergenz für allgemeine Reihen wirklich „stärker“ als die gewöhnliche Konvergenz ist, also dass aus der absoluten Konvergenz einer Reihe auch die Konvergenz folgt. Dazu müssen wir zunächst das Cauchy-Kriterium aus Satz 6.25 auf Reihen übertragen. Auch hier ist dieses Kriterium wieder besonders deswegen wichtig, weil es zum einen zur Konvergenz *äquivalent* ist (man mit ihm also Konvergenz genauso wie Divergenz nachweisen kann) und es außerdem in $\mathbb{R}$ und $\mathbb{C}$ gleichermaßen funktioniert.

Folgerung 7.11 (**Cauchy-Kriterium** für Reihen)**.** *Eine Reihe $\sum_{n=0}^{\infty} a_n$ in $\mathbb{K}$ ist genau dann konvergent, wenn*

$$\forall \varepsilon > 0 \; \exists n_0 \in \mathbb{N} \; \forall m \geq n \geq n_0 : \left| \sum_{k=n+1}^{m} a_k \right| < \varepsilon.$$

Beweis. Nach Definition ist die Reihe $\sum_{n=0}^{\infty} a_n$ genau dann konvergent, wenn die Folge $(s_N)_N$ der Partialsummen mit $s_N = \sum_{n=0}^{N} a_n$ konvergiert. Wenden wir das Cauchy-Kriterium für Folgen aus Satz 6.25 auf $(s_N)_N$ an, sehen wir, dass dies genau dann der Fall ist, wenn

$$\forall \varepsilon > 0 \; \exists n_0 \in \mathbb{N} \; \forall m,n \geq n_0 : |s_n - s_m| < \varepsilon.$$

Natürlich können wir hier aus Symmetriegründen $m \geq n$ annehmen, und aus $s_m - s_n = \sum_{k=n+1}^{m} a_k$ folgt dann sofort die Behauptung. □

Lemma 7.12. *Jede absolut konvergente Reihe in $\mathbb{K}$ ist konvergent.*

Beweis. Es sei $\sum_{n=0}^{\infty} a_n$ eine absolut konvergente Reihe, d. h. die Reihe $\sum_{n=0}^{\infty} |a_n|$ sei konvergent. Nach dem Cauchy-Kriterium aus Folgerung 7.11 gibt es also zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ mit

$$\left| \sum_{k=n+1}^{m} |a_k| \right| = \sum_{k=n+1}^{m} |a_k| < \varepsilon$$

für alle $m \geq n \geq n_0$. Dann ist nach der Dreiecksungleichung aber erst recht

$$\left| \sum_{k=n+1}^{m} a_k \right| \leq \sum_{k=n+1}^{m} |a_k| < \varepsilon,$$

und damit ist wiederum nach dem Cauchy-Kriterium auch die Reihe $\sum_{n=0}^{\infty} a_n$ konvergent. □

Beispiel 7.13. Die alternierende harmonische Reihe $\sum_{n=1}^{\infty}(-1)^n \frac{1}{n}$ ist nach Beispiel 7.9 konvergent. Sie ist aber nicht absolut konvergent, da die harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ nach Beispiel 7.3 (c) divergiert. Die Umkehrung von Lemma 7.12 gilt also nicht.

Als Nächstes hatten wir behauptet, dass die absolute Konvergenz einer Reihe sicher stellt, dass man die Summanden ohne Änderung des Grenzwerts umordnen kann. Dies wollen wir jetzt zeigen.

Definition 7.14 (Umordnungen einer Reihe)**.** Es sei $\sum_{n=0}^{\infty} a_n$ eine Reihe in $\mathbb{K}$ und $\sigma : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Dann heißt die Reihe $\sum_{n=0}^{\infty} a_{\sigma(n)}$ (die offensichtlich aus den gleichen Summanden besteht, nur evtl. in anderer Reihenfolge) eine **Umordnung** von $\sum_{n=0}^{\infty} a_n$.

Satz 7.15 (Umordnungssatz). *Jede Umordnung einer absolut konvergenten Reihe ist ebenfalls absolut konvergent und konvergiert gegen denselben Grenzwert.*

Beweis. Es seien $\sum_{n=0}^{\infty} a_n$ eine absolut konvergente Reihe und $\sigma : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Ferner sei $\varepsilon > 0$ beliebig. Da $\sum_{n=0}^{\infty} |a_n|$ nach Voraussetzung konvergiert, gibt es nach dem Cauchy-Kriterium aus Folgerung 7.11 ein $n_0 \in \mathbb{N}$ mit $\sum_{k=n+1}^{m} |a_k| < \frac{\varepsilon}{2}$ für alle $m \geq n \geq n_0$. Insbesondere haben wir für $n = n_0$ und $m \to \infty$ also

$$\sum_{k=n_0+1}^{\infty} |a_k| \leq \frac{\varepsilon}{2} < \varepsilon. \tag{1}$$

Da σ surjektiv ist, können wir ein $n_0' \geq n_0$ wählen, so dass alle Summanden $a_0, \dots, a_{n_0}$ bis zum n_0'-ten Term der Umordnung aufgetreten sind, also so dass

$$\{0, 1, \dots, n_0\} \subset \{\sigma(0), \sigma(1), \dots, \sigma(n_0')\} \tag{2}$$

gilt. Wir betrachten nun für beliebiges $n \geq n_0'$ die Summe

$$\sum_{k=0}^{n} (a_{\sigma(k)} - a_k) = (a_{\sigma(0)} - a_0) + (a_{\sigma(1)} - a_1) + \cdots + (a_{\sigma(n)} - a_n).$$

Wegen (2) und $n \geq n_0' \geq n_0$ treten in dieser Summe alle Glieder $a_0, \dots, a_{n_0}$ sowohl einmal mit positivem als auch einmal mit negativem Vorzeichen auf, heben sich also heraus. Die übrigen a_n mit $n > n_0$ können sich ebenfalls herausheben, oder mit einem positiven oder negativen Vorzeichen auftreten. Wir können dies symbolisch schreiben als

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

wobei die Summe hier über gewisse (endlich viele) $k > n_0$ läuft und für jedes solche k das Vorzeichen von a_k positiv oder negativ sein kann. Damit können wir diesen Ausdruck mit der Dreiecksungleichung betragsmäßig abschätzen durch

$$\left| \sum_{k=0}^{n} (a_{\sigma(k)} - a_k) \right| = \left| \sum_{k} \pm a_k \right| \leq \sum_{k} |a_k| \leq \sum_{k=n_0+1}^{\infty} |a_k| \overset{(1)}{<} \varepsilon.$$

Daraus ergibt sich $\sum_{n=0}^{\infty}(a_{\sigma(n)} - a_n) = 0$, und damit nach den üblichen Rechenregeln aus Lemma 7.4

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

Die Umordnung $\sum_{n=0}^{\infty} a_{\sigma(n)}$ konvergiert also gegen den gleichen Grenzwert wie die ursprüngliche Reihe. Wenden wir dieses Ergebnis nun auch noch auf die Reihe $\sum_{n=0}^{\infty} |a_n|$ an, so erhalten wir genauso $\sum_{n=0}^{\infty} |a_{\sigma(n)}| = \sum_{n=0}^{\infty} |a_n|$, woraus die absolute Konvergenz der Umordnung folgt. □

15

Bemerkung 7.16 (Summen mit abzählbar unendlicher Indexmenge)**.** Da es bei „unendlichen Summen“ im Fall der absoluten Konvergenz also nicht auf die Reihenfolge der Summanden ankommt, können wir damit auch derartige Summen definieren, bei denen die Summanden zunächst einmal überhaupt keine vorgegebene Reihenfolge haben, sondern durch eine beliebige abzählbar unendliche Menge I indiziert werden: Ist $a_i \in \mathbb{K}$ für alle $i \in I$, so wählen wir eine bijektive Abbildung $\sigma : \mathbb{N} \to I$. Ist dann die Reihe $\sum_{n=0}^{\infty} a_{\sigma(n)}$ absolut konvergent, so schreiben wir den Wert dieser Reihe als

$$\sum_{i\in I} a_i := \sum_{n=0}^{\infty} a_{\sigma(n)} \quad \in \mathbb{K}.$$

Dies hängt dann nach dem Umordnungssatz 7.15 nicht von der Wahl von σ ab, da sich die durch eine andere Bijektion entstehende Reihe nur durch eine Umordnung unterscheidet und somit nichts an der absoluten Konvergenz bzw. dem Grenzwert der Reihe ändert.

Ist die Reihe $\sum_{n=0}^{\infty} a_{\sigma(n)}$ hingegen nicht absolut konvergent, so können wir $\sum_{i\in I} a_i$ nicht sinnvoll definieren.

Aufgabe 7.17. Zeige die folgende Verallgemeinerung des Leibniz-Kriteriums ins Komplexe: Ist $(a_n)_n$ eine reelle, monoton fallende Nullfolge, so konvergiert die Reihe $\sum_{n=0}^{\infty} a_n x^n$ für alle $x \in \mathbb{K}$ mit $|x| = 1$ und $x \neq 1$.

(Hinweis: Untersuche die Reihe $\displaystyle\sum_{n=0}^{\infty} a_n (x-1) x^n$.)

Aufgrund der schönen Eigenschaften absolut konvergenter Reihen werden wir uns im Folgenden oftmals eher für die absolute als für die „gewöhnliche“ Konvergenz von Reihen interessieren. Wir wollen nun ein paar Kriterien zusammentragen, mit denen man die absolute Konvergenz von Reihen in vielen Fällen einfach nachprüfen kann. Das erste von ihnen ist eigentlich sehr offensichtlich:

Satz 7.18 (**Majoranten-/Minorantenkriterium**)**.** *Es seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ zwei Reihen in $\mathbb{K}$ mit $|a_n| \leq |b_n|$ für fast alle n.*

(a) *Ist $\sum_{n=0}^{\infty} b_n$ absolut konvergent, so auch $\sum_{n=0}^{\infty} a_n$.*

*(Man nennt $\sum_{n=0}^{\infty} b_n$ in diesem Fall eine konvergente **Majorante** von $\sum_{n=0}^{\infty} a_n$.)*

(b) *Ist $\sum_{n=0}^{\infty} a_n$ divergent, so auch $\sum_{n=0}^{\infty} |b_n|$.*

*(Man nennt $\sum_{n=0}^{\infty} a_n$ in diesem Fall eine divergente **Minorante** von $\sum_{n=0}^{\infty} |b_n|$.)*

Beweis.

(a) Ist $\sum_{n=0}^{\infty} b_n$ absolut konvergent, also $\sum_{n=0}^{\infty} |b_n|$ beschränkt, so ist wegen $|a_n| \leq |b_n|$ für fast alle n auch $\sum_{n=0}^{\infty} |a_n|$ beschränkt, und damit $\sum_{n=0}^{\infty} a_n$ nach Lemma 7.7 absolut konvergent.

(b) Ist die Reihe $\sum_{n=0}^{\infty} a_n$ nicht konvergent, so ist sie nach Lemma 7.12 insbesondere auch nicht absolut konvergent, und daher kann nach (a) auch $\sum_{n=0}^{\infty} b_n$ nicht absolut konvergent sein, d. h. die Reihe $\sum_{n=0}^{\infty} |b_n|$ divergiert. □

Beispiel 7.19.

(a) Die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$ ist konvergent: Wegen $\frac{1}{(n+1)^2} \leq \frac{1}{n(n+1)}$ für $n \geq 1$ ist $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ nach Beispiel 7.3 (b) eine (absolut) konvergente Majorante von $\sum_{n=1}^{\infty} \frac{1}{(n+1)^2}$. Damit konvergiert die Reihe

$$\frac{1}{1^2} + \sum_{n=1}^{\infty} \frac{1}{(n+1)^2} = \frac{1}{1^2} + \sum_{n=2}^{\infty} \frac{1}{n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2}.$$

Beachte, dass man auf diese Art mit Hilfe des Majorantenkriteriums zwar die Konvergenz der Reihe $\sum_{n=1}^{\infty} \frac{1}{n^2}$ beweisen, aber nicht ihren Grenzwert bestimmen kann (in der Tat kann man zeigen, dass der Wert dieser Reihe gleich $\frac{\pi^2}{6}$ ist).

(b) Für $k \geq 2$ ist die Reihe $\sum_{n=1}^{\infty} \frac{1}{n^k}$ konvergent, denn wegen $\frac{1}{n^k} \leq \frac{1}{n^2}$ für alle n ist $\sum_{n=1}^{\infty} \frac{1}{n^2}$ nach (a) eine konvergente Majorante.

(c) Die Reihe $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$ dagegen ist divergent, denn wegen $\frac{1}{\sqrt{n}} \geq \frac{1}{n}$ ist die harmonische Reihe $\sum_{n=1}^{\infty} \frac{1}{n}$ aus Beispiel 7.3 (c) eine divergente Minorante.

Aufgabe 7.20. Es sei $(a_n)_n$ eine Folge in $\mathbb{K}$ mit $a_n \neq -1$ für alle $n \in \mathbb{N}$. Man zeige:

$$\sum_{n=0}^{\infty} a_n \text{ ist absolut konvergent} \quad \Leftrightarrow \quad \sum_{n=0}^{\infty} \frac{a_n}{1+a_n} \text{ ist absolut konvergent.}$$

Wenn man mit dem Majorantenkriterium die (absolute) Konvergenz einer Reihe nachweisen möchte, stellt sich natürlich die Frage, wo man eine konvergente Majorante herbekommt. Sehr oft kann man hierfür einfach eine geometrische Reihe $\sum_{n=0}^{\infty} q^n$ für ein $q \in \mathbb{R}_{>0}$ mit $q < 1$ wie in Beispiel 7.3 (a) verwenden. Aus diesem Ansatz ergeben sich in der Tat die folgenden beiden allgemeinen Kriterien, die sehr oft anwendbar sind:

Satz 7.21 (**Quotientenkriterium**). *Es sei* $\sum\limits_{n=0}^{\infty} a_n$ *eine Reihe in* $\mathbb{K}$ *mit* $a_n \neq 0$ *für fast alle* n. *Dann gilt:*

(a) *Ist* $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$, *so ist die Reihe* $\sum\limits_{n=0}^{\infty} a_n$ *absolut konvergent.*

(b) *Ist* $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| > 1$, *so ist die Reihe* $\sum\limits_{n=0}^{\infty} a_n$ *divergent.*

Der Fall $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = \infty$ *ist dabei in (b) zugelassen. Ist die Folge* $\left(\left|\frac{a_{n+1}}{a_n}\right|\right)_n$ *jedoch unbestimmt divergent oder konvergiert sie gegen* 1, *so macht das Quotientenkriterium keine Aussage.*

Beweis. Es sei $a := \lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| \in \mathbb{R}_{\geq 0} \cup \{\infty\}$.

(a) Ist $a < 1$, so können wir ein $\varepsilon > 0$ wählen, so dass auch noch $q := a + \varepsilon < 1$ gilt. Wegen $\lim\limits_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = a$ gibt es dann ein $n_0 \in \mathbb{N}$, so dass

$$\left|\frac{a_{n+1}}{a_n}\right| < a + \varepsilon = q \quad \text{für alle } n \geq n_0,$$

und damit $|a_{n+1}| < q\,|a_n|$. Daraus ergibt sich für alle $n \geq n_0$

$$|a_n| < q\,|a_{n-1}| < q^2\,|a_{n-2}| < \cdots < q^{n-n_0}\,|a_{n_0}|.$$

Also ist die Reihe $\sum_{n=0}^{\infty} q^{n-n_0}\,|a_{n_0}|$ eine Majorante der gegebenen Reihe $\sum_{n=0}^{\infty} a_n$. Wegen $q < 1$ konvergiert sie nach Beispiel 7.3 (a) absolut, denn es ist

$$\sum_{n=0}^{\infty} q^{n-n_0}\,|a_{n_0}| = q^{-n_0}\,|a_{n_0}| \sum_{n=0}^{\infty} q^n = q^{-n_0}\,|a_{n_0}| \cdot \frac{1}{1-q}.$$

Die zu beweisende Aussage folgt damit aus dem Majorantenkriterium von Satz 7.18.

(b) Ist $a \in \mathbb{R}$ mit $a > 1$, so können wir ein $\varepsilon > 0$ finden mit $a - \varepsilon > 1$. In diesem Fall gibt es nach der Grenzwertbedingung ein $n_0 \in \mathbb{N}$, so dass

$$\left|\frac{a_{n+1}}{a_n}\right| > a - \varepsilon > 1 \quad \text{für alle } n \geq n_0.$$

Beachte, dass dies auch im Fall $a = \infty$ gilt, denn auch dann sind ja insbesondere fast alle $\left|\frac{a_{n+1}}{a_n}\right|$ größer als 1.

Also gilt $|a_{n+1}| > |a_n|$ für alle $n \geq n_0$. Damit ist $(|a_n|)_n$ ab n_0 aber eine monoton wachsende Folge positiver Zahlen, und kann somit keine Nullfolge sein. Die gegebene Reihe divergiert also nach dem Trivialkriterium aus Lemma 7.5. □

Das zweite, recht ähnliche Kriterium, das auf dem Vergleich mit der geometrischen Reihe beruht, benutzt die höheren Wurzeln aus Aufgabe 5.37.

Satz 7.22 (**Wurzelkriterium**). *Für jede Reihe* $\sum_{n=0}^{\infty} a_n$ *in* $\mathbb{K}$ *gilt:*

(a) *Ist* $\limsup_{n\to\infty} \sqrt[n]{|a_n|} < 1$, *so ist die Reihe* $\sum_{n=0}^{\infty} a_n$ *absolut konvergent.*

(b) *Ist* $\limsup_{n\to\infty} \sqrt[n]{|a_n|} > 1$, *so ist die Reihe* $\sum_{n=0}^{\infty} a_n$ *divergent.*

Der Fall $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = \infty$ *ist dabei in (b) wieder zugelassen. Ist jedoch* $\limsup_{n\to\infty} \sqrt[n]{|a_n|} = 1$, *so macht das Wurzelkriterium keine Aussage.*

Beweis. Es sei $a := \limsup_{n\to\infty} \sqrt[n]{|a_n|} \in \mathbb{R}_{\geq 0} \cup \{\infty\}$.

(a) Für $a < 1$ sei wieder $\varepsilon > 0$ mit $q := a + \varepsilon < 1$. Nach Lemma 5.47 (a) gilt dann

$$\sqrt[n]{|a_n|} < a + \varepsilon = q, \qquad \text{also} \quad |a_n| < q^n$$

für fast alle n. Also ist $\sum_{n=0}^{\infty} q^n$ eine Majorante der gegebenen Reihe. Da diese wegen $q < 1$ nach Beispiel 7.3 (a) (absolut) konvergiert, konvergiert auch $\sum_{n=0}^{\infty} a_n$ nach dem Majorantenkriterium aus Satz 7.18 absolut.

(b) Ist $a \in \mathbb{R}$ mit $a > 1$, so wählen wir ein $\varepsilon > 0$ mit $a - \varepsilon > 1$. Diesmal folgt dann aus Lemma 5.47 (b), dass

$$\sqrt[n]{|a_n|} > a - \varepsilon > 1$$

für unendlich viele n. Beachte, dass dies auch im Fall $a = \infty$ gilt, weil die Folge $(\sqrt[n]{|a_n|})_n$ dann eine Teilfolge mit uneigentlichem Grenzwert ∞ hat.

Damit ist aber auch $|a_n| > 1$ für unendlich viele n. Also ist $(a_n)_n$ keine Nullfolge, und die gegebene Reihe divergiert nach dem Trivialkriterium aus Lemma 7.5. □

Bemerkung 7.23 (Vergleich von Quotienten- und Wurzelkriterium). Das Quotientenkriterium hat gegenüber dem Wurzelkriterium den Vorteil, dass sich der Quotient $\left|\frac{a_{n+1}}{a_n}\right|$ oft einfacher berechnen lässt als die Wurzel $\sqrt[n]{|a_n|}$. Allerdings benötigen wir im Quotientenkriterium einen Grenzwert der Quotientenfolge, während im Wurzelkriterium der Limes superior der Wurzelfolge genügt, der ja nach Bemerkung 5.52 zumindest im uneigentlichen Sinne stets existiert.

Dies liegt daran, dass wir für die Induktion im Beweis von Satz 7.21 brauchten, dass fast alle Quotienten in (a) kleiner als $a + \varepsilon$ und in (b) größer als $a - \varepsilon$ sind, so dass a dort der Grenzwert der Quotientenfolge sein musste. Im Beweis von Satz 7.22 brauchten wir dagegen in (a) zwar auch, dass fast alle Wurzeln kleiner als $a + \varepsilon$ sind, aber in (b) reichten unendlich viele Wurzeln größer als $a - \varepsilon$.

Mit dieser Beobachtung sieht man allerdings mit Hilfe von Lemma 5.47 auch, dass wir den Grenzwert im Quotientenkriterium von Satz 7.21 in (a) durch den Limes superior und in (b) durch den Limes inferior ersetzen könnten, um so noch allgemeinere Aussagen zu erhalten. Hat die Quotientenfolge jedoch mehrere Häufungspunkte, von denen einer größer als 1 und einer kleiner als 1 ist, so lässt sich aus der Idee des Quotientenkriteriums aber endgültig keine Aussage über die Konvergenz der Reihe mehr herleiten.

Beispiel 7.24.

(a) Betrachten wir für ein $q \in \mathbb{K}$ mit $q \neq 0$ die geometrische Reihe $\sum_{n=0}^{\infty} q^n$ selbst, ist also $a_n = q^n$ in der Notation von Satz 7.21 und 7.22, so ist offensichtlich $\left|\frac{a_{n+1}}{a_n}\right| = \sqrt[n]{|a_n|} = |q|$ unabhängig von n, und sowohl Quotienten- als auch Wurzelkriterium reproduzieren einfach das Ergebnis aus Beispiel 7.3 (a) in den Fällen mit $|q| \neq 1$.

(b) Für die alternierende harmonische Reihe $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$ aus Beispiel 7.9 macht das Quotientenkriterium keine Aussage, denn dort gilt

$$\left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{(-1)^{n+1}/(n+1)}{(-1)^n/n}\right| = \frac{n}{n+1} = \frac{1}{1+\frac{1}{n}} \to 1$$

für $n \to \infty$. Dies war natürlich zu erwarten, da diese Reihe ja auch weder absolut konvergent noch divergent ist.

(c) Die Reihe $\sum_{n=0}^{\infty} \left(\frac{n}{2n+1}\right)^n$ ist nach dem Wurzelkriterium (absolut) konvergent, denn es ist

$$\sqrt[n]{\left(\frac{n}{2n+1}\right)^n} = \frac{n}{2n+1} = \frac{1}{2+\frac{1}{n}} \to \frac{1}{2}$$

für $n \to \infty$.

(d) Für alle $x \in \mathbb{K}$ ist die Reihe

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots$$

nach dem Quotientenkriterium absolut konvergent, denn es gilt

$$\left| \frac{x^{n+1}/(n+1)!}{x^n/n!} \right| = \frac{|x|}{n+1} \to 0$$

für $n \to \infty$. Auf diese Art haben wir also letztlich eine Funktion von $\mathbb{K}$ nach $\mathbb{K}$ definiert, die jedem x den Wert der Reihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ zuordnet.

Es handelt sich bei diesem letzten Beispiel genau um die Exponentialfunktion, die ihr zumindest im reellen Fall bereits aus der Schule kennt. Sie ist aber letztlich nur ein spezielles Beispiel für eine sehr große Klasse von Funktionen, die sich in der Form $x \mapsto \sum_{n=0}^{\infty} a_n x^n$ für gewisse $a_n \in \mathbb{K}$ schreiben lassen. Wir wollen derartige Funktionen, die in dieser Vorlesung immer wieder vorkommen werden, daher jetzt einführen und etwas genauer untersuchen.

7.C Potenzreihen

Potenzreihen kann man sich in gewissem Sinne als Verallgemeinerung der Polynome aus Abschnitt 3.C vorstellen: statt endlicher Summen $a_0 + a_1 x + \cdots + a_n x^n$ in einer Variablen x betrachten wir nun unendliche Reihen der Form

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + a_2 x^2 + \cdots .$$

Wir beginnen mit der formalen Definition solcher Reihen, zusammen mit dem wohl wichtigsten Beispiel: der Exponentialfunktion.

Definition 7.25 (Potenzreihen und die Exponentialfunktion)**.**

(a) Ist $(a_n)_n$ eine Folge in $\mathbb{K}$ und $x \in \mathbb{K}$, so heißt die Reihe $\sum_{n=0}^{\infty} a_n x^n$ die **Potenzreihe** in x mit Koeffizienten $(a_n)_n$. Ist $D \subset \mathbb{K}$ die Menge aller x, für die diese Reihe konvergiert, so können wir die Potenzreihe offensichtlich als Funktion von D nach $\mathbb{K}$ auffassen.

Der Startindex einer Potenzreihe darf auch größer als 0 sein (dann kann man die ersten Koeffizienten ja gleich 0 setzen), aber nie kleiner als 0 – eine Potenzreihe in x enthält nach Definition keine negativen Potenzen von x.

(b) Die **Exponentialfunktion** ist die Potenzreihenfunktion

$$\exp\colon \mathbb{K} \to \mathbb{K} \qquad \text{mit} \qquad \exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!}$$

(die nach Beispiel 7.24 (d) für alle $x \in \mathbb{K}$ absolut konvergiert).

Aus dem Wurzelkriterium können wir sofort eine allgemeine Aussage ableiten, auf welchen Gebieten derartige Potenzreihen konvergieren: nämlich auf um 0 zentrierten Intervallen (im Fall $\mathbb{K} = \mathbb{R}$) bzw. auf Kreisen um 0 (im Fall $\mathbb{K} = \mathbb{C}$).

Satz und Definition 7.26 (Konvergenzgebiete von Potenzreihen, Formel von **Cauchy-Hadamard**). *Es sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe über $\mathbb{K}$ und*

$$r = \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{|a_n|}} \quad \in \mathbb{R}_{\geq 0} \cup \{\infty\}$$

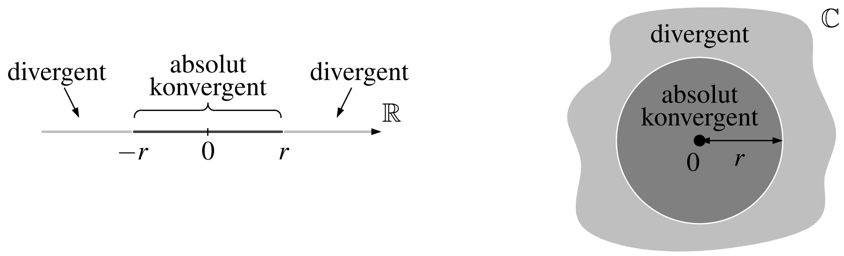
(beachte, dass dieser Wert nach Bemerkung 5.52 in jedem Fall existiert). Dann gilt:

(a) *für alle $x \in \mathbb{K}$ mit $|x| < r$ ist die Potenzreihe absolut konvergent;*

(b) *für alle $x \in \mathbb{K}$ mit $|x| > r$ ist die Potenzreihe divergent.*

Im Fall $|x| = r$ kann keine allgemeine Aussage über die Konvergenz der Reihe getroffen werden.

*Die geometrische Deutung dieser Konvergenzaussagen im reellen bzw. komplexen Fall zeigt das folgende Bild. Man nennt r den **Konvergenzradius** und $\{x \in \mathbb{K} : |x| < r\}$ das **Konvergenzgebiet** der Potenzreihe.*



16

Beweis. Wir wenden das Wurzelkriterium aus Satz 7.22 auf die Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$ an: Nach Aufgabe 5.51 (b) ist für alle $x \in \mathbb{K}$

$$\limsup_{n\to\infty} \sqrt[n]{|a_n x^n|} = \limsup_{n\to\infty} \left(|x| \cdot \sqrt[n]{|a_n|}\right) = |x| \cdot \limsup_{n\to\infty} \sqrt[n]{|a_n|} = \frac{|x|}{r},$$

also konvergiert die Reihe für $|x| < r$ absolut und divergiert für $|x| > r$. □

Bemerkung 7.27. Beachte, dass die Eigenschaften (a) und (b) aus Satz 7.26 den Konvergenzradius eindeutig charakterisieren als

$$r = \sup \left\{ |x| : x \in \mathbb{K} \text{ mit } \sum_{n=0}^{\infty} a_n x^n \text{ konvergent} \right\}. \qquad (*)$$

So ist z. B. auch ohne Berechnung des Ausdrucks für r in Satz 7.26 klar, dass die Exponentialreihe aus Definition 7.25 (b) den Konvergenzradius ∞ hat, da sie ja auf ganz $\mathbb{K}$ konvergiert. In der Tat wird die Gleichung $(*)$ in der Literatur auch oft als Definition des Konvergenzradius einer Potenzreihe benutzt.

Es sollte nicht überraschen, dass man nicht nur mit dem Wurzelkriterium, sondern auch mit dem Quotientenkriterium eine Aussage über den Konvergenzradius einer Potenzreihe treffen kann. Allerdings ist diese nicht ganz so universell, da sie wie in Satz 7.21 die Existenz des Grenzwerts der Quotientenfolge der Koeffizienten voraussetzt.

Satz 7.28 (Alternative Formel für den Konvergenzradius einer Potenzreihe). *Es sei $\sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe in $\mathbb{K}$ mit $a_n \neq 0$ für fast alle n. Existiert dann der Grenzwert*

$$r = \lim_{n\to\infty} \left| \frac{a_n}{a_{n+1}} \right| \quad \in \mathbb{R}_{\geq 0} \cup \{\infty\},$$

so ist dies der Konvergenzradius der Potenzreihe.

Beweis. Wegen

$$\lim_{n\to\infty} \left| \frac{a_{n+1} x^{n+1}}{a_n x^n} \right| = \lim_{n\to\infty} \left(|x| \cdot \left| \frac{a_{n+1}}{a_n} \right| \right) = |x| \cdot \lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = \frac{|x|}{r}$$

konvergiert die Potenzreihe nach Satz 7.21 absolut für $|x| < r$ und divergiert für $|x| > r$. Nach Bemerkung 7.27 ist r damit der Konvergenzradius der Reihe. □

Beispiel 7.29.

(a) Die Potenzreihe $\sum_{n=1}^{\infty} \frac{x^n}{n}$ hat nach Satz 7.28 den Konvergenzradius

$$r = \lim_{n\to\infty} \left| \frac{1/n}{1/(n+1)} \right| = \lim_{n\to\infty} \frac{n+1}{n} = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right) = 1,$$

konvergiert also für alle $x \in \mathbb{K}$ mit $|x| < 1$ absolut und divergiert für alle x mit $|x| > 1$. Für $|x| = 1$ treten in der Tat verschiedene Fälle auf: Im Fall $x = 1$ erhalten wir die harmonische Reihe, die nach Beispiel 7.3 (c) divergiert, während wir für $x = -1$ die alternierende harmonische Reihe haben, die nach Beispiel 7.9 konvergiert. Dies zeigt noch einmal, dass unsere obigen nur von $|x|$ abhängigen Kriterien auf dem Rand des Konvergenzgebiets wirklich keine allgemeine Aussage machen können.

(b) Für den Konvergenzradius der Potenzreihe $\sum_{n=0}^{\infty} n\,x^n$ gilt nach Satz 7.26 und 7.28

$$r = \frac{1}{\limsup\limits_{n\to\infty} \sqrt[n]{n}} \qquad \text{sowie} \qquad r = \lim_{n\to\infty} \left| \frac{n}{n+1} \right| = 1.$$

Vergleich dieser beiden Ergebnisse liefert also $\limsup\limits_{n\to\infty} \sqrt[n]{n} = 1$. Für alle $\varepsilon > 0$ gilt damit nach Lemma 5.47 (a), dass $\sqrt[n]{n} < 1 + \varepsilon$ für fast alle n. Natürlich ist aber auch $\sqrt[n]{n} \geq 1$ für alle n, und damit ergibt sich zusammen

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

(was gar nicht so offensichtlich ist, da der Ausdruck $\sqrt[n]{n}$ für wachsendes n ja durch das n unter der Wurzel größer, durch das Ziehen der n-ten Wurzel aber kleiner wird).

Aufgabe 7.30. Untersuche die folgenden Reihen auf Konvergenz (im Fall (c) in Abhängigkeit von $x \in \mathbb{R}$):

$$\text{(a)} \quad \sum_{n=0}^{\infty} \frac{n}{(3 + (-1)^n)^n}, \qquad \text{(b)} \quad \sum_{n=0}^{\infty} \frac{n!}{n^n}, \qquad \text{(c)} \quad \sum_{n=0}^{\infty} \frac{n}{n^2+1} x^n.$$

Eine Berechnung des Grenzwerts im Fall der Konvergenz ist nicht erforderlich.

Aufgabe 7.31. Zeige, dass jede Potenzreihe $\sum_{n=0}^{\infty} a_n x^n$ in $\mathbb{K}$ denselben Konvergenzradius wie ihre „formale Ableitung“ $\sum_{n=1}^{\infty} n\,a_n x^{n-1}$ hat, aber nicht notwendig für die gleichen x konvergiert.

(Wir werden in Folgerung 10.27 sehen, dass dies dann auch genau die „gewöhnliche“ Ableitung der ursprünglichen Reihe ist.)

Aufgabe 7.32. Es sei $(a_n)_n$ eine Folge in $\mathbb{K}$. Man zeige:

(a) Ist a_n der Quotient von zwei Polynomen in n, so machen weder das Quotienten- noch das Wurzelkriterium eine Aussage über die Konvergenz der Reihe $\sum_{n=0}^{\infty} a_n$.

(b) Macht das Wurzelkriterium keine Aussage über die Konvergenz der Reihe $\sum_{n=0}^{\infty} a_n$, so macht auch das Quotientenkriterium keine Aussage darüber.

Aufgabe 7.33 (Alternative Darstellung der Exponentialfunktion)**.** Zeige, dass für alle $x \in \mathbb{K}$

$$\exp(x) = \lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n$$

gilt. (Hinweis: Eine Möglichkeit besteht darin, durch eine geeignete Abschätzung zu zeigen, dass

$$\left(\sum_{k=0}^{n} \frac{x^k}{k!}\right) - \left(1 + \frac{x}{n}\right)^n$$

mit $n \to \infty$ gegen 0 konvergiert.)

Der Ausdruck $\lim\limits_{n\to\infty} \left(1 + \frac{x}{n}\right)^n$ in dieser Darstellung der Exponentialfunktion hat übrigens eine sehr anschauliche Interpretation: Wenn ihr 1 Euro zu einem Zinssatz x ein Jahr lang anlegt und sich die

Bank bereit erklärt, nicht einmal am Ende des Jahres den Betrag x an Zinsen auszuzahlen, sondern stattdessen n-mal einen Zinssatz von $\frac{x}{n}$ bezahlt, so habt ihr dadurch am Ende des Jahres aufgrund des Zinseszinses natürlich mehr Geld als bei einer einmaligen Zinszahlung, nämlich genau $\left(1+\frac{x}{n}\right)^n$. Die gerade gezeigte Formel besagt, dass ihr selbst für $n \to \infty$, also wenn ihr die Bank zu einer unendlichen Aufteilung der Zinsen auf diese Art überreden könntet, dadurch kein unendliches Vermögen aufbauen könntet, sondern am Ende des Jahres lediglich den Betrag von $\exp(x)$ hättet.

Zum Schluss dieses Kapitels wollen wir nun noch Produkte von Reihen untersuchen. Haben wir z. B. zwei Potenzreihen $\sum_{k=0}^{\infty} a_k x^k$ und $\sum_{l=0}^{\infty} b_l x^l$, so würden wir erwarten, dass wir ihr Produkt für alle x im Durchschnitt der Konvergenzgebiete der beiden Reihen wie folgt durch „unendliches Ausmultiplizieren" berechnen und wieder zu einer neuen Potenzreihe zusammenfassen können:

$$\begin{aligned}\left(\sum_{k=0}^{\infty} a_k x^k\right) \cdot \left(\sum_{l=0}^{\infty} b_l x^l\right) &= (a_0 + a_1 x + a_2 x^2 + \cdots) \cdot (b_0 + b_1 x + b_2 x^2 + \cdots) \\ &\overset{?}{=} a_0 b_0 + (a_0 b_1 + a_1 b_0) x + (a_0 b_2 + a_1 b_1 + a_2 b_0) x^2 + \cdots \\ &= \sum_{n=0}^{\infty} c_n x^n\end{aligned}$$

mit $c_n = \sum_{k=0}^{n} a_k b_{n-k}$. Wir wollen nun zeigen, dass dies in der Tat erlaubt ist – und zwar auch für allgemeine Reihen, nicht nur für Potenzreihen (also wenn wir uns das x in der obigen Rechnung wegdenken). Da die einzelnen ausmultiplizierten Summanden in der entstehenden Reihe dabei aber irgendwie sortiert werden müssen, sollte es in Anbetracht des Umordnungssatzes 7.15 nicht überraschen, dass wir für die Gültigkeit dieser Rechnung die *absolute* Konvergenz der Reihen benötigen.

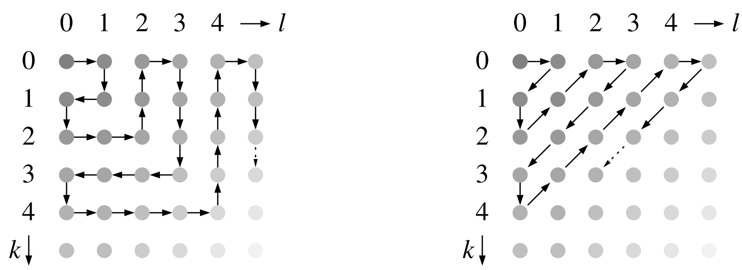
Satz 7.34 (**Cauchy-Produkt** von Reihen). *Es seien $\sum_{k=0}^{\infty} a_k$ und $\sum_{l=0}^{\infty} b_l$ zwei absolut konvergente Reihen in $\mathbb{K}$. Setzen wir dann*

$$c_n = \sum_{k=0}^{n} a_k b_{n-k}$$

für alle $n \in \mathbb{N}$, so ist auch die Reihe $\sum_{n=0}^{\infty} c_n$ absolut konvergent, und es gilt

$$\left(\sum_{k=0}^{\infty} a_k\right) \cdot \left(\sum_{l=0}^{\infty} b_l\right) = \sum_{n=0}^{\infty} c_n. \qquad (*)$$

Beweis. Nach Voraussetzung existieren die Grenzwerte $A := \sum_{k=0}^{\infty} |a_k|$ und $B := \sum_{l=0}^{\infty} |b_l|$ in $\mathbb{R}$. Wir zeigen, dass die Summe $\sum_{(k,l)\in\mathbb{N}^2} a_k b_l$ über die nach Beispiel 5.59 (a) abzählbare Indexmenge $\mathbb{N} \times \mathbb{N}$ im Sinne von Bemerkung 7.16 existiert und mit dem Wert beider Seiten der Gleichung $(*)$ übereinstimmt. Dazu betrachten wir die beiden im folgenden Bild dargestellten Aufzählungen von $\mathbb{N}^2$: eine „quadratische", und eine „schräge" wie im Cantorschen Diagonalverfahren im Beweis von Satz 5.58.



(a) (b)

(a) Summieren wir zunächst die Beträge $|a_k b_l|$ in der Reihenfolge dieser „quadratischen" Aufzählung, so erhalten wir nach der Summation von höchstens n^2 Termen, also denen im Quadrat links oben mit n Zeilen und Spalten, maximal den Wert

$$\sum_{k=0}^{n-1} \sum_{l=0}^{n-1} |a_k b_l| = \left(\sum_{k=0}^{n-1} |a_k|\right) \cdot \left(\sum_{l=0}^{n-1} |b_l|\right) \le A \cdot B.$$

Die Reihe über alle $|a_k b_l|$ ist (in dieser Aufzählungsreihenfolge) also beschränkt. Damit ist die Reihe über $a_k b_l$ nach Definition 7.10 absolut konvergent, und nach Bemerkung 7.16 existiert somit die Reihe $\sum_{(k,l)\in\mathbb{N}^2} a_k b_l$.

Um den Wert dieser Reihe zu bestimmen, können wir nach Lemma 5.19 auch nur eine Teilfolge der Partialsummenfolge betrachten. Nehmen wir hierzu die Partialsummen, bei denen wir die ersten n^2 Terme aufaddieren, und lassen dort n gegen ∞ gehen, so erhalten wir

$$\sum_{(k,l)\in\mathbb{N}^2} a_k b_l = \lim_{n\to\infty} \sum_{k=0}^{n-1}\sum_{l=0}^{n-1} a_k b_l = \lim_{n\to\infty} \left(\sum_{k=0}^{n-1} a_k\right)\cdot\left(\sum_{l=0}^{n-1} b_l\right) \overset{5.13}{=} \left(\sum_{k=0}^{\infty} a_k\right)\cdot\left(\sum_{l=0}^{\infty} b_l\right),$$

also die linke Seite von $(*)$.

(b) In dieser „schrägen“ Aufzählungsreihenfolge können wir den Wert von $\sum_{(k,l)\in\mathbb{N}^2} a_k b_l$ analog als Grenzwert für $N\to\infty$ der Partialsummen bestimmen, bei denen wir die ersten N Diagonalen aufsummieren. Da die Summe der $a_k b_l$ entlang der Diagonale mit $k+l=n$ gerade c_n ist, erhalten wir so

$$\sum_{(k,l)\in\mathbb{N}^2} a_k b_l = \lim_{N\to\infty} \sum_{n=0}^{N-1} c_n = \sum_{n=0}^{\infty} c_n$$

und damit durch Vergleich mit (a) die behauptete Gleichheit $(*)$. Weil die ersten N Diagonalen außerdem im Quadrat links oben mit N Zeilen und Spalten enthalten sind, haben wir weiterhin mit der Dreiecksungleichung

$$\sum_{n=0}^{N-1} |c_n| = \sum_{n=0}^{N-1} \left|\sum_{k=0}^{n} a_k b_{n-k}\right| \le \sum_{n=0}^{N-1}\sum_{k=0}^{n} |a_k|\cdot|b_{n-k}| \le \left(\sum_{k=0}^{N-1} |a_k|\right)\cdot\left(\sum_{l=0}^{N-1} |b_l|\right) \le A\cdot B$$

und damit auch die absolute Konvergenz von $\sum_{n=0}^{\infty} c_n$ gezeigt. □

Die wohl wichtigste Anwendung des Cauchy-Produkts erhalten wir im Fall der Exponentialfunktion.

Folgerung 7.35 (Funktionalgleichung der Exponentialfunktion)**.** *Für alle $x,y\in\mathbb{K}$ gilt*

$$\exp(x)\cdot\exp(y) = \exp(x+y).$$

Beweis. Nach Beispiel 7.24 (d) und Definition 7.25 (b) konvergiert die Exponentialreihe auf ganz $\mathbb{K}$ absolut. Also gilt für alle $x,y\in\mathbb{K}$

$$\begin{aligned}
\exp(x)\cdot\exp(y) &= \left(\sum_{k=0}^{\infty} \frac{x^k}{k!}\right)\cdot\left(\sum_{l=0}^{\infty} \frac{y^l}{l!}\right) \\
&= \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} \frac{x^k}{k!}\cdot\frac{y^{n-k}}{(n-k)!}\right) && \text{(nach Satz 7.34)} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!}\cdot\left(\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}\right) \\
&= \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} && \text{(nach Satz 4.7)} \\
&= \exp(x+y),
\end{aligned}$$

was zu zeigen war. □

Bemerkung 7.36 (Produkte von Potenzreihen)**.** Sind $\sum_{k=0}^{\infty} a_k x^k$ und $\sum_{l=0}^{\infty} b_l x^l$ zwei Potenzreihen in x mit Konvergenzradien r_1 bzw. r_2, so gilt nach Satz 7.34 für alle $x\in\mathbb{K}$ mit $|x| < \min(r_1,r_2)$ (wo also nach Satz 7.26 beide Potenzreihen absolut konvergieren)

$$\left(\sum_{k=0}^{\infty} a_k x^k\right)\cdot\left(\sum_{l=0}^{\infty} b_l x^l\right) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_k b_{n-k}\right) x^n,$$

d. h. das Produkt zweier Potenzreihen mit Konvergenzradien r_1 und r_2 ist wieder eine Potenzreihe, deren Konvergenzradius mindestens $\min(r_1,r_2)$ beträgt und die durch „unendliches Ausmultiplizieren“ berechnet werden kann.

Aufgabe 7.37. Es sei $q \in \mathbb{C}$ mit $|q| < 1$. Berechne das Cauchy-Produkt $\left(\sum_{n=0}^{\infty} q^n\right)^2$ und damit den Wert der Reihe $\sum_{n=0}^{\infty} n\, q^n$.

Aufgabe 7.38. Für alle $n \in \mathbb{N}$ sei $a_n = \frac{(-1)^n}{\sqrt{n+1}}$. Zeige in diesem Fall, dass die Reihe $\sum\limits_{n=0}^{\infty} a_n$ zwar konvergiert, aber dass ihr Cauchy-Produkt mit sich selbst wie in Satz 7.34

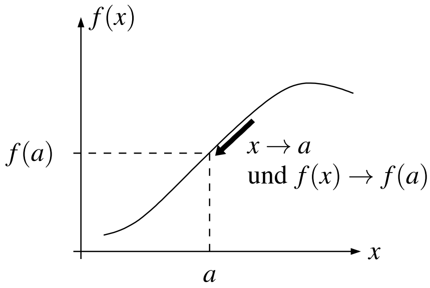
$$\sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k a_{n-k}$$

divergiert.

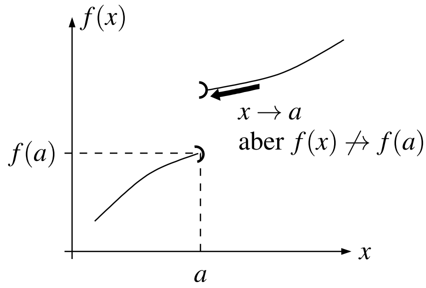
8. Stetigkeit

Nachdem wir uns gerade ausführlich mit Grenzwerten von Folgen und Reihen befasst haben, wollen wir den Grenzwertbegriff nun auf Funktionen einer reellen (oder evtl. komplexen) Variablen ausdehnen, also auf Funktionen $f: D \to \mathbb{K}$ mit einer Definitionsmenge $D \subset \mathbb{K}$. Dies führt zum zentralen Begriff der Stetigkeit solcher Funktionen.

Anschaulich ist die Frage dabei: Wenn eine Stelle $a \in D$ mit Funktionswert $f(a)$ gegeben ist, und wir nun andere Punkte $x \in D$ in der Nähe von a betrachten, liegt dann auch $f(x)$ in der Nähe von $f(a)$? Im Bild unten links ist dies der Fall: Laufen wir entlang des dick eingezeichneten Pfeils mit x auf den Punkt a zu, so nähern sich auch die Funktionswerte $f(x)$ dem Punkt $f(a)$. In diesem Fall werden wir sagen, dass f im Punkt a stetig ist.



f ist stetig in a



f ist nicht stetig in a

Im Bild oben rechts dagegen führt der Sprung im Funktionsgraphen zu einem anderen Verhalten: Nähern wir uns hier entlang des dick eingezeichneten Pfeils mit x dem Punkt a, so nähert sich $f(x)$ nicht dem Wert $f(a)$, sondern dem oberen Punkt der Sprungstelle. In diesem Fall ist f im Punkt a unstetig.

Um dies mathematisch exakt zu formulieren, wollen wir jetzt den Begriff von Funktionsgrenzwerten einführen. Im linken Fall können wir dann sagen, dass $f(x)$ mit $x \to a$ gegen $f(a)$ konvergiert, während dies im Fall rechts nicht so ist.

8.A Grenzwerte von Funktionen

Wie eben erläutert wollen wir das Verhalten von Funktionen $f: D \to \mathbb{K}$ mit $D \subset \mathbb{K}$ untersuchen, wenn wir uns einem vorgegebenen Punkt a nähern. Dieser Wert a kann dabei, muss aber nicht unbedingt selbst Element von D sein. Wir sollten aber natürlich sicherstellen, dass wir uns zumindest innerhalb von D dem Punkt a beliebig annähern können – also anschaulich gesprochen, dass a entweder in D oder am Rand von D liegt. Formal bedeutet dies, dass a im Sinne der folgenden Definition ein Berührpunkt von D sein muss.

Definition 8.1 (Berührpunkte)**.** Es sei $D \subset \mathbb{K}$ eine Menge. Eine Zahl $a \in \mathbb{K}$ heißt **Berührpunkt** von D, wenn jede ε-Umgebung von a (siehe Bemerkung 5.2) mindestens einen Punkt aus D enthält, also wenn es zu jedem $\varepsilon > 0$ ein $x \in D$ gibt mit $|x - a| < \varepsilon$. Die Menge aller Berührpunkte von D wird mit $\overline{D}$ bezeichnet und heißt der **Abschluss** von D.

Beispiel 8.2.

(a) Für ein $D \subset \mathbb{K}$ ist jedes $a \in D$ Berührpunkt von D: Wir können in diesem Fall in der Definition 8.1 einfach $x = a$ für jedes ε wählen. Es gilt also stets $D \subset \overline{D}$.

(b) Für ein offenes reelles Intervall $D = (a,b) \subset \mathbb{R}$ ist $\overline{D} = [a,b]$ das zugehörige abgeschlossene Intervall. Die Randpunkte a und b sind also Berührpunkte von D, die nicht selbst zu D gehören.

17

Für derartige Berührpunkte können wir nun Grenzwerte von Funktionen definieren. Die Konstruktion ist völlig analog zur Definition 5.1 (b) des Grenzwerts von Folgen:

Definition 8.3 (Grenzwerte von Funktionen)**.** Es seien $D \subset \mathbb{K}$ eine Menge und $f\colon D \to \mathbb{K}$ eine Funktion. Ferner sei $a \in \overline{D}$ ein Berührpunkt von D.

Dann heißt eine Zahl $c \in \mathbb{K}$ **Grenzwert** von f in a, wenn

$$\forall \varepsilon \in \mathbb{R}_{>0}\ \exists \delta \in \mathbb{R}_{>0}\ \forall x \in D : |x-a| < \delta \Rightarrow |f(x)-c| < \varepsilon.$$

Wie schon im Fall von Folgen werden wir sehen (siehe Bemerkung 8.12), dass ein solcher Grenzwert im Fall der Existenz eindeutig ist, so dass wir also von *dem* Grenzwert von f in a sprechen können. Wir schreiben dies dann als

$$\lim_{x\to a} f(x) = c \qquad \text{bzw.} \qquad \lim_{\substack{x\to a\\ x\in D}} f(x) = c$$

oder auch als „$f(x) \to c$ für $x \to a$“, und sagen, dass f in a **konvergent** ist gegen c. Existiert ein solcher Grenzwert nicht, so heißt f **divergent** in a.

Bemerkung 8.4. Gilt in Definition 8.3 sogar $a \in D$, so kommt als Grenzwert c nur $f(a)$ in Frage: Wir können dann nämlich $x = a$ in der Grenzwertbedingung von Definition 8.3 setzen (so dass $|x-a| < \delta$ in jedem Fall erfüllt ist) und erhalten damit $|f(a)-c| < \varepsilon$ für alle ε – was nur möglich ist, wenn $c = f(a)$ ist.

Definition 8.5 (Stetigkeit)**.** Es seien $D \subset \mathbb{K}$ eine Menge und $f\colon D \to \mathbb{K}$ eine Funktion.

(a) Ist $a \in D$, so heißt f **stetig** in a, wenn der Grenzwert $\lim\limits_{x\to a} f(x)$ existiert (und nach Bemerkung 8.4 damit zwangsläufig gleich $f(a)$ ist), d. h. wenn

$$\forall \varepsilon \in \mathbb{R}_{>0}\ \exists \delta \in \mathbb{R}_{>0}\ \forall x \in D : |x-a| < \delta \Rightarrow |f(x)-f(a)| < \varepsilon.$$

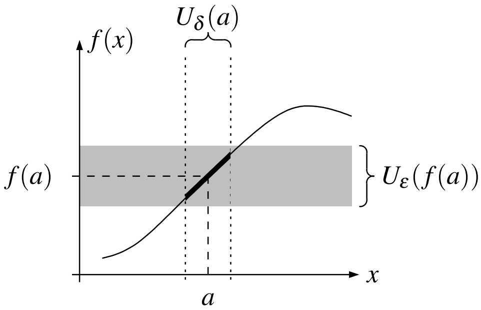
Die Funktion f heißt **stetig** (auf D), wenn sie in jedem Punkt $a \in D$ stetig ist.

(b) Ist $a \in \overline{D} \backslash D$, so heißt f **stetig fortsetzbar** nach a, wenn der Grenzwert $c = \lim\limits_{x\to a} f(x)$ existiert. (In diesem Fall erhält man nämlich eine in a stetige Funktion

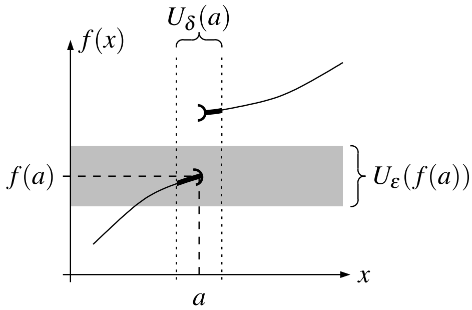
$$\tilde{f}\colon D \cup \{a\} \to \mathbb{K},\ x \mapsto \begin{cases} f(x) & \text{für } x \in D, \\ c & \text{für } x = a, \end{cases}$$

die man als stetige Fortsetzung von f nach a bezeichnet.)

Bemerkung 8.6 (Anschauliche Deutung des Grenzwertbegriffs)**.** Das Bild unten zeigt noch einmal das Beispiel vom Anfang dieses Kapitels mit den eben eingeführten Notationen: Nach Definition ist eine Funktion $f\colon D \to \mathbb{K}$ genau dann stetig in einem Punkt $a \in D$, wenn es zu jeder (beliebig kleinen) ε-Umgebung $U_\varepsilon(f(a))$ von $f(a)$ eine δ-Umgebung $U_\delta(a)$ von a gibt, in der alle Punkte von D nach $U_\varepsilon(f(a))$ abgebildet werden, also so dass $f(D \cap U_\delta(a)) \subset U_\varepsilon(f(a))$ gilt. Im Bild unten bedeutet dies, dass zu jedem auch noch so schmal gewählten grauen horizontalen Streifen um $f(a)$ eine Einschränkung von f auf eine hinreichend kleine Umgebung von a dazu führt, dass alle Funktionswerte dort (im Bild unten dick eingezeichnet) in dem gewählten Streifen liegen. Dies entspricht genau der ursprünglichen Motivation, dass eine kleine Änderung von x um a herum auch nur zu einer kleinen Änderung der Funktionswerte $f(x)$ um $f(a)$ führen darf. Bei der linken Funktion ist dies also der Fall, bei der rechten aufgrund der Sprungstelle jedoch nicht.



f ist stetig in a



f ist nicht stetig in a

Die ebenfalls oft gehörte geometrische Interpretation, dass eine reelle Funktion stetig ist, wenn man „ihren Graphen zeichnen kann, ohne den Stift abzusetzen", ist übrigens etwas mit Vorsicht zu genießen, wie die Beispiele 8.7 (e) und (f) unten zeigen.

Beispiel 8.7.

(a) Die Identität $f\colon \mathbb{K} \to \mathbb{K},\ x \mapsto x$ ist stetig: Sind $a \in \mathbb{K}$ und $\varepsilon > 0$ gegeben, so setze man $\delta := \varepsilon$. Dann gilt natürlich für alle $x \in \mathbb{K}$ mit $|x-a| < \delta$, dass $|f(x) - f(a)| = |x-a| < \delta = \varepsilon$. Analog zeigt man, dass konstante Funktionen stetig sind.

(b) Wir zeigen, dass die Betragsfunktion $f\colon \mathbb{K} \to \mathbb{K},\ x \mapsto |x|$ stetig ist. Es seien dazu $a \in \mathbb{K}$ und $\varepsilon > 0$ gegeben. Wir setzen wieder $\delta := \varepsilon$. Dann folgt für alle $x \in \mathbb{K}$ mit $|x-a| < \delta$ mit Hilfe der Dreiecksungleichung nach unten

$$|x-a| \geq |x| - |a| \quad \text{und} \quad |x-a| \geq |a| - |x|.$$

Da $|f(x) - f(a)| = \big||x| - |a|\big|$ aber eine der beiden Zahlen $|x| - |a|$ und $|a| - |x|$ sein muss, ergibt sich in jedem Fall

$$|f(x) - f(a)| \leq |x-a| < \delta = \varepsilon.$$

Damit ist f stetig.

(c) Analog ist die komplexe Konjugation $f\colon \mathbb{C} \to \mathbb{C},\ z \mapsto \overline{z}$ (siehe Notation 6.2) stetig: Sind $a \in \mathbb{C}$ und $\varepsilon > 0$ gegeben, so setzen wir $\delta := \varepsilon$ und erhalten für alle $z \in \mathbb{C}$ mit $|z-a| < \delta$

$$|f(z) - f(a)| = |\overline{z} - \overline{a}| = |\overline{z-a}| = |z-a| < \delta = \varepsilon.$$

(d) Die Funktion

$$f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} 0 & \text{für } x \neq 0, \\ 1 & \text{für } x = 0 \end{cases}$$

(siehe Bild unten) ist in $a = 0$ nicht stetig. Wollen wir dies formal zeigen, müssen wir die Negation der Bedingung aus Definition 8.5 (a) beweisen, d. h.

$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x \in \mathbb{R} : |x-a| < \delta \text{ und } |f(x) - f(a)| \geq \varepsilon.$$

(Beachte dabei, dass die Negation der Aussage „$|x-a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon$" nach Beispiel 1.9 (a) die angegebene Bedingung „$|x-a| < \delta$ *und* $|f(x) - f(a)| \geq \varepsilon$" ist, und nicht etwa eine Folgerung „$|x-a| < \delta \Rightarrow |f(x) - f(a)| \geq \varepsilon$"!)

Dies zu zeigen ist hier aber sehr einfach: Setzen wir $\varepsilon = \frac{1}{2}$ und ist $\delta > 0$ beliebig, so können wir $x = \frac{\delta}{2}$ setzen und erhalten $|x-a| = \frac{\delta}{2} < \delta$ und $|f(x) - f(a)| = |0-1| = 1 \geq \varepsilon$.

Anders ausgedrückt existiert in diesem Fall der Grenzwert $\lim\limits_{x \to 0} f(x)$ nicht. Falls ihr jetzt gedacht hättet, dass dieser Grenzwert doch existiert und gleich 0 ist, so habt ihr damit sicher gemeint, dass sich $f(x)$ dem Wert 0 nähert, wenn x in der Nähe von 0, *aber nicht gleich* 0 ist. Der Fall $x = 0$ (bzw. $x = a$) ist in Definition 8.3 aber nicht ausgeschlossen! Wenn wir dies ausschließen wollten, so müssten wir die Definitionsmenge von f auf $\mathbb{R}\backslash\{0\}$ einschränken und würden dann in der Tat

$$\lim_{\substack{x\to 0\\ x\neq 0}} f(x) = 0$$

erhalten.

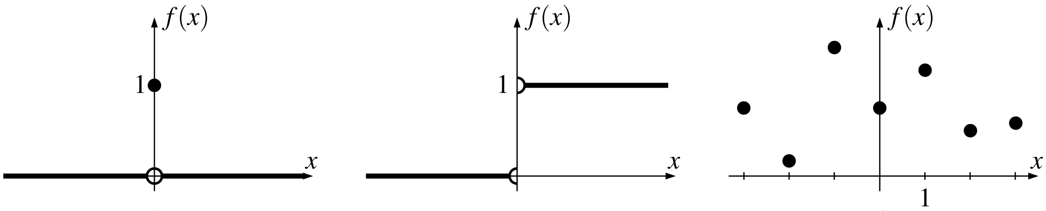
Beachte jedoch, dass es hier in der Literatur zwei verschiedene Konventionen gibt: In manchen Büchern werden Funktionsgrenzwerte so definiert, dass $\lim\limits_{x\to a}$ immer für $\lim\limits_{\substack{x\to a\\ x\neq a}}$ steht.

(e) Die unten im Bild dargestellte Funktion

$$f\colon \mathbb{R}\backslash\{0\} \to \mathbb{R},\ x \mapsto \begin{cases} 0 & \text{für } x<0,\\ 1 & \text{für } x>0 \end{cases}$$

ist stetig – ja, wirklich! Sie ist nämlich an jedem Punkt *der Definitionsmenge*, also an jedem $a \neq 0$ stetig, weil sie in der Nähe eines jeden solchen Punktes (genauer: in der $|a|$-Umgebung von a) konstant ist. Die Funktion f ist aber natürlich *nicht stetig fortsetzbar* nach 0: Der Grenzwert $\lim\limits_{x\to 0} f(x)$ existiert nicht.

(f) *Jede* Funktion $f\colon \mathbb{Z}\to\mathbb{R}$ ist stetig (siehe unten). Das liegt anschaulich einfach daran, dass wir hier gar keine Möglichkeit haben, ein gegebenes $a\in\mathbb{Z}$ ein wenig so zu verändern, dass es immer noch in der Definitionsmenge liegt. Formal können wir in der Bedingung aus Definition 8.5 (a) für jedes gegebene $\varepsilon>0$ immer $\delta=\frac{1}{2}$ setzen und haben damit sicher gestellt, dass $|x-a|<\delta$ mit $x\in\mathbb{Z}$ nur für $x=a$ erfüllt ist, womit dann natürlich auch $|f(x)-f(a)|=0<\varepsilon$ ist.



(d) (e) (f)

Bemerkung 8.8 (Funktionen mit Grenzwert ungleich 0)**.** Es seien $D\subset\mathbb{K}$, $f\colon D\to\mathbb{K}$ eine Funktion und $a\in\overline{D}$ mit $c:=\lim\limits_{x\to a} f(x)\neq 0$. Aus Definition 8.3 für $\varepsilon=\frac{|c|}{2}$ erhalten wir dann wie in Bemerkung 5.12 ein $\delta>0$, so dass $|f(x)-c|<\varepsilon$ und damit

$$|f(x)| = |f(x)-c+c| \geq |c|-|f(x)-c| > |c|-\varepsilon = \frac{|c|}{2} > 0$$

für alle $x\in D$ mit $|x-a|<\delta$ gilt.

Insbesondere ergibt sich im Fall $a\in D$ also, dass eine in a stetige Funktion f mit $f(a)\neq 0$ auch in einer δ-Umgebung von a ungleich 0 ist. Beispiel 8.7 (d) zeigt (bei $a=0$), dass dies für unstetige Funktionen im Allgemeinen natürlich falsch ist.

Eine analoge Aussage gilt im Fall $\mathbb{K}=\mathbb{R}$ auch ohne Beträge: Eine in a stetige reelle Funktion f mit $f(a)>0$ ist auch in einer δ-Umgebung von a positiv.

Aufgabe 8.9. Es seien $m,n\in\mathbb{N}_{>0}$. Berechne den Grenzwert $\lim\limits_{x\to 1}\frac{x^m-1}{x^n-1}$.

Aufgabe 8.10. Zeige durch Rückgang auf die ε-δ-Definition der Stetigkeit, dass die Funktion $f\colon[-1,1]\to\mathbb{R}$, $x\mapsto\sqrt{1-x^2}$ stetig ist.

Wie im Fall von Folgengrenzwerten wollen wir nun natürlich auch für Grenzwerte von Funktionen ein paar einfache Rechenregeln zeigen, z. B. dass solche Grenzwerte wie in Satz 5.13 mit Summen und Produkten vertauschen. Glücklicherweise lassen sich Grenzwerte von Funktionen mit Hilfe des folgenden Satzes immer auf Grenzwerte von Folgen zurückführen, so dass wir viele unserer Ergebnisse dann sofort von Folgen auf Funktionen übertragen können:

Satz 8.11 (Folgenkriterium). *Es seien $D \subset \mathbb{K}$ und $f: D \to \mathbb{K}$ eine Funktion.*

(a) ***(Folgenkriterium für Funktionsgrenzwerte)*** *Für $a \in \overline{D}$ und $c \in \mathbb{K}$ gilt*

$$\lim_{x\to a} f(x) = c \quad \Leftrightarrow \quad \text{Für jede Folge } (x_n)_n \text{ in } D \text{ mit } x_n \to a \text{ gilt } f(x_n) \to c.$$

(b) ***(Folgenkriterium für Stetigkeit)*** *Für $a \in D$ gilt*

$$f \text{ ist stetig in } a \quad \Leftrightarrow \quad \text{Für jede Folge } (x_n)_n \text{ in } D \text{ mit } x_n \to a \text{ gilt } f(x_n) \to f(a).$$

Es ist dann also

$$\lim_{n\to\infty} f(x_n) = f\left(\lim_{n\to\infty} x_n\right),$$

d. h. „eine stetige Funktion f vertauscht mit der Grenzwertbildung von Folgen“.

Beweis. Wir beweisen zunächst Teil (a).

„$\Rightarrow$“ Es seien $\lim_{x\to a} f(x) = c$ und $(x_n)_n$ eine Folge in D mit $x_n \to a$; wir müssen $f(x_n) \to c$ zeigen. Dazu sei $\varepsilon > 0$ beliebig. Wegen $\lim_{x\to a} f(x) = c$ gibt es ein $\delta > 0$, so dass $|f(x) - c| < \varepsilon$ für alle $x \in D$ mit $|x - a| < \delta$ gilt. Wegen $x_n \to a$ ist aber $|x_n - a| < \delta$ für fast alle n, und damit dann auch $|f(x_n) - c| < \varepsilon$ für diese n. Damit gilt $f(x_n) \to c$.

„$\Leftarrow$“ Wir zeigen diese Richtung durch einen Widerspruchsbeweis und nehmen also an, dass c kein Grenzwert von $f(x)$ in a ist, d. h. (durch Negation der Definition 8.3)

$$\exists \varepsilon > 0 \ \forall \delta > 0 \ \exists x \in D : |x - a| < \delta \text{ und } |f(x) - c| \geq \varepsilon.$$

Wir wählen nun ein solches ε. Indem wir $\delta = \frac{1}{n}$ für $n \in \mathbb{N}_{>0}$ setzen, erhalten wir für alle n ein $x_n \in D$ mit $|x_n - a| < \frac{1}{n}$ und $|f(x_n) - c| \geq \varepsilon$. Für diese Folge gilt dann aber $x_n \to a$ und $f(x_n) \not\to c$ im Widerspruch zur Annahme.

Teil (b) folgt nun mit Definition 8.5 (a) sofort aus (a). □

Bemerkung 8.12. Mit Hilfe des Folgenkriteriums können wir nun sehr schnell viele Resultate über Grenzwerte von Folgen auf Funktionen übertragen. So folgt z. B. sofort, dass Grenzwerte von Funktionen immer eindeutig sind, sofern sie existieren: Sind $D \subset \mathbb{K}$, $f: D \to \mathbb{K}$, $a \in \overline{D}$ und $f(x) \to c$ für $x \to a$, so können wir wegen $a \in \overline{D}$ eine Folge $(x_n)_n$ in D mit $x_n \to a$ wählen, und erhalten mit Satz 8.11 (a) dann auch $f(x_n) \to c$. Da Folgengrenzwerte nach Lemma 5.5 aber eindeutig sind, ist dies für höchstens ein c möglich.

Die folgenden Rechenregeln ergeben sich ebenfalls sofort aus dem Folgenkriterium und sorgen auch dafür, dass wir für sehr viele Funktionen ohne weitere Rechnung direkt die Stetigkeit nachweisen können:

Satz 8.13 (Grenzwertsätze für Funktionen**).** *Es seien $D \subset \mathbb{K}$ und $f, g: D \to \mathbb{K}$ zwei Funktionen. Weiterhin sei $a \in \overline{D}$, so dass beide Grenzwerte $\lim_{x\to a} f(x)$ und $\lim_{x\to a} g(x)$ existieren. Dann gilt*

$$\lim_{x\to a}(f(x) + g(x)) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x).$$

Eine analoge Aussage gilt auch für $f(x) - g(x)$, $f(x) \cdot g(x)$ und $\frac{f(x)}{g(x)}$ (letzteres natürlich nur, falls $\lim_{x\to a} g(x) \neq 0$).

Insbesondere sind für $a \in D$ also mit f und g auch $f + g$, $f - g$, $f \cdot g$ und $\frac{f}{g}$ in a stetig (letzteres wiederum nur, falls $g(a) \neq 0$).

Beweis. Beachte im Fall $\frac{f}{g}$ zunächst, dass die Definitionsmenge dieses Quotienten nicht ganz D, sondern die evtl. kleinere Menge $D' = \{x \in D : g(x) \neq 0\}$ ist. Um überhaupt über den Grenzwert von $\frac{f(x)}{g(x)}$ für $x \to a$ sprechen zu können, müssen wir also zuerst überprüfen, dass a ein Berührpunkt von D' ist. Dies folgt aber aus Bemerkung 8.8, die besagt, dass g wegen $\lim\limits_{x\to a} g(x) \neq 0$ in einer ε-Umgebung von a nirgends 0 wird, so dass D und D' dort also übereinstimmen.

Die eigentliche Behauptung des Lemmas ist nun eine direkte Übertragung der Grenzwertsätze für Folgen aus Satz 5.13. Wir betrachten hier nur den Fall der Addition, da die anderen drei Fälle wörtlich genauso bewiesen werden. Dazu berechnen wir den Grenzwert von $f(x)+g(x)$ mit dem Folgenkriterium aus Satz 8.11 (a): Es sei $(x_n)_n$ eine beliebige Folge in D mit $x_n \to a$. Dann gilt nach dem Folgenkriterium für f und g

$$\lim_{n\to\infty} f(x_n) = \lim_{x\to a} f(x) \quad \text{und} \quad \lim_{n\to\infty} g(x_n) = \lim_{x\to a} g(x)$$

und damit nach Satz 5.13 (a)

$$\lim_{n\to\infty} (f(x_n)+g(x_n)) = \lim_{n\to\infty} f(x_n) + \lim_{n\to\infty} g(x_n) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x),$$

d. h. $f(x_n)+g(x_n)$ konvergiert für jede solche Folge gegen $\lim\limits_{x\to a} f(x) + \lim\limits_{x\to a} g(x)$, also immer gegen dieselbe Zahl. Wiederum nach dem Folgenkriterium – diesmal für $f+g$ – ergibt sich damit also wie gewünscht auch

$$\lim_{x\to a}(f(x)+g(x)) = \lim_{x\to a} f(x) + \lim_{x\to a} g(x). \qquad \square$$

Beispiel 8.14. Jede *rationale Funktion*, also jede Funktion der Form $f(x) = \frac{p(x)}{q(x)}$ mit Polynomfunktionen $p(x)$ und $q(x)$, lässt sich natürlich mit den vier Grundrechenarten aus der Identität und den konstanten Funktionen zusammensetzen. Damit folgt aus Satz 8.13, dass jede solche Funktion auf jeder Definitionsmenge $D \subset \{x \in \mathbb{K} : q(x) \neq 0\}$ – also überall dort, wo f überhaupt definiert werden kann – stetig ist.

18

Als Nächstes wollen wir zeigen, dass auch Verkettungen stetiger Funktionen wieder stetig sind. Dazu beweisen wir einen etwas allgemeineren Satz, der analog zur Vertauschbarkeit stetiger Funktionen mit Folgengrenzwerten in Satz 8.11 (b) ist, und der auch oft zur Berechnung von Grenzwerten nützlich ist.

Satz 8.15 (Grenzwert einer Verkettung)**.** *Es seien $f: D \to \mathbb{K}$ mit $D \subset \mathbb{K}$ sowie $g: D' \to \mathbb{K}$ mit $D' \subset \mathbb{K}$ zwei Funktionen mit $f(D) \subset D'$. Ferner sei $a \in \overline{D}$, so dass $\lim\limits_{x\to a} f(x)$ existiert, in D' liegt, und g in diesem Punkt stetig ist. Dann gilt*

$$\lim_{x\to a}(g\circ f)(x) = g\big(\lim_{x\to a} f(x)\big),$$

d. h. „für stetige g kann man die Anwendung von g mit der Grenzwertbildung vertauschen".

Insbesondere folgt für $a \in D$ also aus der Stetigkeit von f in a und der von g in $f(a)$ auch die Stetigkeit von $g \circ f$ in a, d. h. die Verkettung stetiger Funktionen ist stetig.

Beweis. Wir zeigen das Lemma wieder mit dem Folgenkriterium aus Satz 8.11 (a). Es sei also $(x_n)_n$ eine beliebige Folge in D mit $x_n \to a$. Weil der Grenzwert $c := \lim\limits_{x\to a} f(x)$ nach Voraussetzung existiert, gilt $f(x_n) \to c$ nach dem Folgenkriterium für f. Da weiterhin g in c stetig ist, gilt nach dem Folgenkriterium für g auch $(g\circ f)(x_n) = g(f(x_n)) \to g(c)$. Die Aussage des Lemmas ergibt sich damit aus dem Folgenkriterium für $g \circ f$. $\square$

Aufgabe 8.16. Zeige, dass die Funktion

$$f: [0,1] \to \mathbb{R},\ x \mapsto \begin{cases} x & \text{für } x \in \mathbb{Q}, \\ 1-x & \text{für } x \notin \mathbb{Q} \end{cases}$$

genau im Punkt $a = \frac{1}{2}$ stetig ist.

Aufgabe 8.17. Man beweise: Ist $f\colon \mathbb{R} \to \mathbb{R}$ eine stetige Funktion, für die die Funktionalgleichung

$$f(x+y) = f(x) + f(y) \qquad \text{für alle } x, y \in \mathbb{R}$$

gilt, so gibt es ein $a \in \mathbb{R}$ mit $f(x) = ax$ für alle $x \in \mathbb{R}$, d. h. f ist eine lineare Funktion.

Bleibt die Aussage richtig, wenn man überall $\mathbb{R}$ durch $\mathbb{C}$ ersetzt?

Genau wie bei Folgen wollen wir nun auch für Funktionen im reellen Fall uneigentliche Grenzwerte einführen, und zwar sowohl in der Start- als auch in der Zielmenge: Wir wollen sowohl Grenzwerte der Form $\lim_{x\to\infty} f(x)$ definieren als auch sagen, was es bedeutet, dass der Grenzwert einer Funktion gleich ∞ ist. Die folgende Definition ist völlig analog zu Definition 5.40.

Definition 8.18 (Uneigentliche Grenzwerte von Funktionen)**.** Es seien $D \subset \mathbb{R}$ und $f\colon D \to \mathbb{R}$ eine Funktion.

(a) Für $a \in \overline{D}$ schreiben wir $\lim_{x\to a} f(x) = \infty$, wenn

$$\forall s \in \mathbb{R}\ \exists \delta > 0\ \forall x \in D : |x-a| < \delta \Rightarrow f(x) > s.$$

Wie im Fall von Folgen in Definition 5.40 spricht man in diesem Fall von einem **uneigentlichen Grenzwert** bzw. sagt, dass f für $x \to a$ **bestimmt divergiert**.

(b) Ist D nach oben unbeschränkt (so dass man x in $f(x)$ überhaupt beliebig groß werden lassen kann), so schreibt man $\lim_{x\to\infty} f(x) = c$ für ein $c \in \mathbb{R}$, wenn

$$\forall \varepsilon > 0\ \exists r \in \mathbb{R}\ \forall x \in D : x \geq r \Rightarrow |f(x) - c| < \varepsilon.$$

Kombiniert man dies nun noch mit (a), so erhält man die Schreibweise $\lim_{x\to\infty} f(x) = \infty$ für

$$\forall s \in \mathbb{R}\ \exists r \in \mathbb{R}\ \forall x \in D : x \geq r \Rightarrow f(x) > s.$$

Beachte, dass diese letzten beiden Notationen sogar exakt mit der Definition von Folgengrenzwerten aus Definition 5.1 (b) und 5.40 übereinstimmen, wenn man sie auf eine Folge (a_n) als Funktion mit Definitionsmenge $\mathbb{N}$ anwendet.

Analog definiert man derartige Grenzwerte mit $-\infty$ statt ∞.

Bemerkung 8.19. Man prüft leicht nach, dass mit Definition 8.18 sowohl das Folgenkriterium für Funktionsgrenzwerte aus Satz 8.11 (a) als auch die Grenzwertsätze aus Satz 8.13 auch für diese uneigentlichen Grenzwerte gelten, wenn man die üblichen Rechenregeln für $\pm\infty$ verwendet.

8.B Eigenschaften stetiger Funktionen

Nachdem wir nun von vielen Funktionen gesehen haben, wie man ihre Stetigkeit nachweisen kann, wollen wir jetzt untersuchen, was wir davon haben, wenn wir wissen, dass eine gegebene Funktion stetig ist. Dazu wollen wir einige sehr anschauliche Aussagen formal beweisen, die für *reelle* stetige Funktionen auf einem abgeschlossenen Intervall $[a,b]$ mit $a < b$ gelten. Die erste von ihnen besagt, dass eine solche Funktion mit je zwei Funktionswerten auch jeden Wert dazwischen annehmen muss – was bei einer kontinuierlichen Änderung der Funktionswerte natürlich zu erwarten ist.

Satz 8.20 (Zwischenwertsatz). *Es sei $f\colon [a,b] \to \mathbb{R}$ eine stetige Funktion. Dann gibt es zu jedem c zwischen $f(a)$ und $f(b)$ ein $x \in [a,b]$ mit $f(x) = c$.*

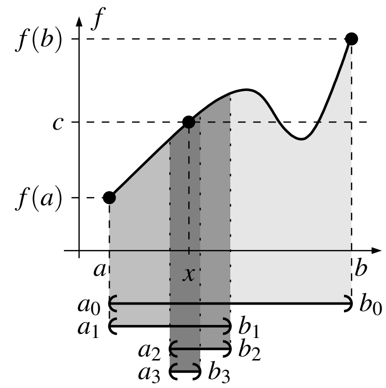
Beweis. Wir können ohne Einschränkung annehmen, dass wie im Bild unten rechts $f(a) \leq f(b)$ und damit $f(a) \leq c \leq f(b)$ gilt. Ausgehend von $[a_0, b_0] := [a,b]$ konstruieren wir nun rekursiv eine Intervallschachtelung

$$[a,b] = [a_0,b_0] \supset [a_1,b_1] \supset [a_2,b_2] \supset \cdots$$

mit in jedem Schritt halbierter Länge der Intervalle, so dass $f(a_n) \leq c \leq f(b_n)$ für alle $n \in \mathbb{N}$ gilt:

Ist $[a_n, b_n]$ für ein $n \in \mathbb{N}$ bereits konstruiert, so betrachten wir den Funktionswert $f(\frac{a_n+b_n}{2})$ in der Intervallmitte.

- Ist $f(\frac{a_n+b_n}{2}) \geq c$ (wie im Bild im Fall $n = 0$), so ersetzen wir die rechte Intervallgrenze durch den Mittelpunkt, setzen also $a_{n+1} := a_n$ und $b_{n+1} := \frac{a_n+b_n}{2}$.
- Ist dagegen $f(\frac{a_n+b_n}{2}) < c$ (wie im Fall $n = 1$ im Bild rechts), so ersetzen wir die linke Intervallgrenze durch den Mittelpunkt, setzen also $a_{n+1} := \frac{a_n+b_n}{2}$ und $b_{n+1} := b_n$.



Für den nach Satz 5.39 durch diese Intervallschachtelung definierten Punkt $x \in [a, b]$ gilt dann $a_n \to x$ und $b_n \to x$, nach dem Folgenkriterium aus Satz 8.11 also

$$f(x) = \lim_{n\to\infty} f(a_n) \leq c \leq \lim_{n\to\infty} f(b_n) = f(x)$$

und damit $f(x) = c$. □

Als Nächstes wollen wir zeigen, dass eine stetige Funktion auf einem abgeschlossenen Intervall wie in Satz 8.20 immer beschränkt ist.

Definition 8.21 (Beschränkte und monotone Funktionen)**.** Es seien $D \subset \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion. Dann heißt f auf D ...

(a) (nach oben bzw. unten) **beschränkt**, wenn die Menge $f(D) \subset \mathbb{R}$ aller Bildpunkte von f (nach oben bzw. unten) beschränkt ist.

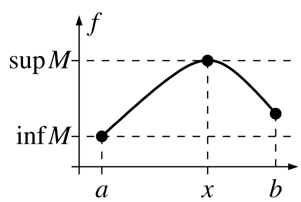
(b) **monoton wachsend** oder **steigend** (bzw. **streng monoton wachsend** oder **steigend**), wenn für alle $x, y \in D$ mit $x < y$ gilt, dass $f(x) \leq f(y)$ (bzw. $f(x) < f(y)$). Analog definiert man **(streng) monoton fallend**.

Satz 8.22. *Jede stetige Funktion $f : [a, b] \to \mathbb{R}$ auf einem abgeschlossenen Intervall ist beschränkt.*

Beweis. Angenommen, f wäre unbeschränkt. Dann gäbe es zu jedem $n \in \mathbb{N}$ ein $x_n \in [a, b]$ mit $|f(x_n)| > n$. Beachte, dass die Folge $(f(x_n))_n$ dann natürlich unbeschränkt, die Folge $(x_n)_n$ aber beschränkt ist (weil ja stets $x_n \in [a, b]$ gilt). Nach dem Satz 6.21 von Bolzano-Weierstraß besitzt $(x_n)_n$ also eine konvergente Teilfolge $(x_{n_k})_k$. Der Grenzwert x dieser Teilfolge liegt nach Satz 5.24 (a) ebenfalls in $[a, b]$ und damit in der Definitionsmenge von f. Nach dem Folgenkriterium aus Satz 8.11 (b) müsste dann aber die Folge $(f(x_{n_k}))_k$ gegen $f(x)$ konvergieren – was ein Widerspruch dazu ist, dass diese Folge nach Konstruktion unbeschränkt und damit divergent ist. □

Bemerkung 8.23. Für nicht abgeschlossene Intervalle ist Satz 8.22 natürlich im Allgemeinen falsch, wie das Beispiel $f(x) = \frac{1}{x}$ auf dem offenen Intervall $(0, 1)$ zeigt.

Wir haben gerade gesehen, dass das Bild $M = f([a, b])$ einer stetigen Funktion auf einem abgeschlossenen Intervall $[a, b]$ immer beschränkt ist und damit also stets zwischen $\inf M$ und $\sup M$ liegt. Wir wollen nun zeigen, dass Infimum und Supremum dieser Menge in der Tat sogar Minimum und Maximum sind, also dass diese beiden Zahlen auch als Werte von f angenommen werden – so wie z. B. im Bild rechts $f(x) = \sup M$ ist.



Satz 8.24 (Satz vom Maximum und Minimum). *Jede stetige Funktion $f : [a, b] \to \mathbb{R}$ nimmt ein Maximum und Minimum an, d. h. die Menge $M = f([a, b]) = \{f(x) : x \in [a, b]\}$ hat ein Maximum und Minimum.*

Beweis. Wir zeigen die Aussage für das Maximum; das Minimum ergibt sich analog. Die Menge M ist natürlich nicht leer und nach Satz 8.22 beschränkt, also existiert $s := \sup M$. Da s die kleinste obere Schranke für M ist, ist $s - \frac{1}{n}$ für alle $n \in \mathbb{N}_{>0}$ dann keine obere Schranke mehr. Wir finden also ein $x_n \in [a,b]$ mit

$$s - \frac{1}{n} < f(x_n) \leq s. \tag{$*$}$$

Nach dem Einschachtelungssatz 5.24 (b) konvergiert $(f(x_n))_n$ damit gegen s. Nun können wir aber wieder nach dem Satz 6.21 von Bolzano-Weierstraß aus $(x_n)_n$ eine konvergente Teilfolge $(x_{n_k})_k$ auswählen, die gegen ein $x \in [a,b]$ konvergiert. Weil f in x stetig ist, gilt nach dem Folgenkriterium aus Satz 8.11 (b) damit

$$f(x) = f\left(\lim_{k\to\infty} x_{n_k}\right) = \lim_{k\to\infty} f(x_{n_k}) = s. \qquad \square$$

Die Ergebnisse aus den Sätzen Satz 8.20, 8.22 und 8.24 lassen sich übrigens einfach in einer einzigen Aussage zusammenfassen:

Folgerung 8.25. *Ist $f\colon [a,b] \to \mathbb{R}$ eine stetige Funktion auf einem abgeschlossenen Intervall $[a,b]$, so ist das Bild von f ebenfalls ein abgeschlossenes Intervall $[c,d]$.*

Beweis. Nach Satz 8.24 existieren $c := \min f([a,b])$ und $d := \max f([a,b])$. Insbesondere gilt damit $f([a,b]) \subset [c,d]$, wobei die Werte c und d von f angenommen werden. Nach dem Zwischenwertsatz werden damit von f aber auch alle Werte zwischen c und d angenommen, d. h. es ist in der Tat $f([a,b]) = [c,d]$. $\square$

Aufgabe 8.26. Man zeige:

(a) Es seien $f\colon \mathbb{R} \to \mathbb{R}$ eine stetige Funktion mit $f(0) = 0$ sowie $g\colon \mathbb{R}\backslash\{0\} \to \mathbb{R}$ eine beschränkte stetige Funktion.

Dann ist die Funktion $f \cdot g\colon \mathbb{R}\backslash\{0\} \to \mathbb{R}$, $x \mapsto f(x) \cdot g(x)$ stetig in den Nullpunkt fortsetzbar.

(b) Jede bijektive, monoton wachsende Funktion $f\colon [a,b] \to [c,d]$ zwischen abgeschlossenen reellen Intervallen ist stetig.

Eine der wichtigsten Anwendungen dieser Aussage ist die Konstruktion von (stetigen) Umkehrfunktionen für streng monotone Funktionen:

Satz 8.27 (Existenz und Stetigkeit von Umkehrfunktionen)**.** *Es sei $f\colon [a,b] \to [c,d]$ eine streng monoton wachsende, stetige Funktion mit $c = f(a)$ und $d = f(b)$. Dann ist f bijektiv, und ihre Umkehrfunktion $f^{-1}\colon [c,d] \to [a,b]$ ist ebenfalls streng monoton wachsend und stetig.*

Analog gilt dies mit „streng monoton fallend“ statt „streng monoton wachsend“.

Beweis. Die Abbildung f ist surjektiv nach Folgerung 8.25. Sie ist auch injektiv, da sie streng monoton wachsend ist. Also ist f bijektiv, und die Umkehrfunktion $f^{-1}\colon [c,d] \to [a,b]$ existiert. Sie ist notwendigerweise streng monoton wachsend, denn wenn es $x,y \in [c,d]$ mit $x < y$ und $f^{-1}(x) \geq f^{-1}(y)$ gäbe, würde sich daraus durch Anwenden der streng monotonen Funktion f der Widerspruch $f(f^{-1}(x)) \geq f(f^{-1}(y))$, also $x \geq y$ ergeben. Nach Aufgabe 8.26 (b) ist f^{-1} damit auch stetig. $\square$

Beispiel 8.28 (Wurzelfunktionen)**.** Es seien $n \in \mathbb{N}_{>0}$ und $R \in \mathbb{R}_{>0}$ gegeben. Dann ist die Funktion $f\colon [0,R] \to [0,R^n]$, $x \mapsto x^n$ nach Lemma 4.16 streng monoton wachsend und nach Beispiel 8.14 stetig. Also ist die Umkehrfunktion $f^{-1}\colon [0,R^n] \to [0,R]$, $x \mapsto \sqrt[n]{x}$, die wir bereits aus Aufgabe 5.37 kennen, ebenfalls streng monoton wachsend und stetig. Betrachtet man diese Aussage für alle R zusammen, ist damit auch die Wurzelfunktion $f^{-1}\colon \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$, $x \mapsto \sqrt[n]{x}$ streng monoton wachsend und stetig. Ihr Graph ensteht wie im Bild unten durch Spiegelung des Graphen von f an der gestrichelt eingezeichneten Diagonalen.

Aufgabe 8.29. Man beweise:

(a) Jede stetige Funktion $f\colon [a,b] \to [a,b]$ hat einen Fixpunkt, d. h. ein $x \in [a,b]$ mit $f(x) = x$.
Ist f darüber hinaus monoton wachsend, so konvergiert die rekursiv definierte Folge $(x_n)_n$ mit $x_{n+1} = f(x_n)$ für alle $n \in \mathbb{N}$ für ein beliebiges $x_0 \in [a,b]$ gegen einen Fixpunkt von f.

(b) Ist $f\colon \mathbb{R} \to \mathbb{R}$ eine stetige Funktion mit $f(x) = f(x+1)$ für alle $x \in \mathbb{R}$ (d. h. „f ist periodisch mit Periodenlänge 1“), dann gibt es ein $a \in \mathbb{R}$ mit $f(a) = f(a+\frac{1}{2})$. (Anschaulich bedeutet dies z. B., dass es auf dem Äquator der Erde (mit Umfang 1 und Koordinate x) stets zwei gegenüberliegende Punkte gibt, an denen die gleiche Temperatur $f(x)$ herrscht.)

Aufgabe 8.30. Es sei $f\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ eine stetige Funktion mit $\lim\limits_{x\to\infty} f(x) = \infty$. Zeige, dass f ein Minimum annimmt.

Aufgabe 8.31. Man zeige:

(a) Es gibt keine stetige Funktion $f\colon \mathbb{R} \to \mathbb{R}$, unter der jede reelle Zahl genau zwei Urbilder hat.

(b) Jede stetige Funktion $f\colon \mathbb{R} \to \mathbb{R}$, die offene Intervalle auf offene Intervalle abbildet, ist streng monoton.

(c) Ist $f\colon \mathbb{R} \to \mathbb{R}$ eine beschränkte stetige Funktion, so gibt es eine Gerade in $\mathbb{R}^2$, die mit dem Graphen von f mindestens drei Punkte gemeinsam hat.

19

8.C Gleichmäßige Konvergenz und Stetigkeit

Wir haben nun einige schöne Eigenschaften stetiger Funktionen gesehen und auch Methoden kennengelernt, mit denen wir von vielen Funktionen ihre Stetigkeit nachweisen können. Allerdings haben wir dabei bisher eine wichtige Klasse von Funktionen ausgelassen – nämlich solche, die durch den Grenzwert einer konvergenten Folge oder Reihe definiert sind, wie z. B. die Exponentialfunktion oder ganz generell allgemeine Potenzreihen wie in Definition 7.25. Zur Untersuchung der Stetigkeit derartiger Funktionen starten wir mit einem einfachen Beispiel.

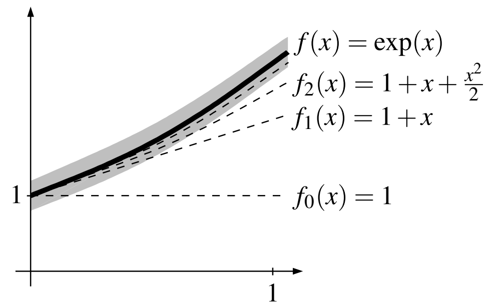
Beispiel 8.32. In Definition 7.25 (b) hatten wir die Exponentialfunktion als $\exp x = \sum_{k=0}^{\infty} \frac{x^k}{k!}$ für alle $x \in \mathbb{C}$ definiert, also als den Grenzwert

$$\exp x = \lim_{n\to\infty} f_n(x) \quad \text{mit} \quad f_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}.$$

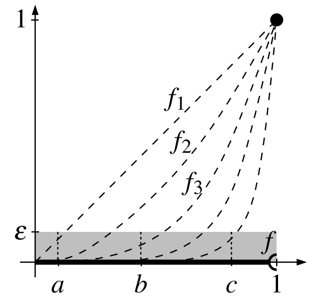
Natürlich ist jede einzelne Partialsumme f_n nach Beispiel 8.14 eine stetige Funktion in x. Da sich diese Partialsummen für $n \to \infty$ immer mehr der Exponentialfunktion annähern, würden wir nun hoffen, dass aus der Stetigkeit aller f_n auch die Stetigkeit der Grenzfunktion, also der Exponentialfunktion folgt. Allgemein fragen wir uns also: Sind $f_n\colon D \to \mathbb{K}$ für $n \in \mathbb{N}$ stetige Funktionen auf einer Menge $D \subset \mathbb{K}$, so dass für alle $x \in D$ der Grenzwert

$$f(x) := \lim_{n\to\infty} f_n(x)$$

existiert (wir sagen in diesem Fall auch, dass die Funktionen f_n *punktweise gegen f konvergieren* – siehe Definition 8.33), ist dann auch diese Grenzfunktion $f\colon D \to \mathbb{K}$ stetig? Der Fall der reellen Exponentialfunktion auf dem Intervall $[0,1]$ ist im folgenden Bild links dargestellt, wobei die einzelnen f_n gestrichelt und die Grenzfunktion f dick eingezeichnet ist.



$$f_n(x) = \sum_{k=0}^{n} \frac{x^k}{k!}, \quad f(x) = \exp(x)$$



$$f_n(x) = x^n, \quad f(x) = \begin{cases} 0 & \text{für } x < 1 \\ 1 & \text{für } x = 1 \end{cases}$$

Es sieht hier also bereits so aus, als ob die Grenzfunktion wie gewünscht stetig ist, und in der Tat werden wir auch sehen, dass dies bei der Exponentialfunktion wirklich der Fall ist. Allerdings ist die Situation im Allgemeinen leider nicht ganz so schön, wie man es sich wünschen würde. Betrachten wir z. B. einmal die Funktionen

$$f_n \colon D = [0,1] \to \mathbb{R},\ x \mapsto x^n$$

wie im Bild oben rechts, so existiert nach Beispiel 5.3 (c) zwar der Grenzwert

$$f(x) := \lim_{n\to\infty} f_n(x) = \lim_{n\to\infty} x^n = \begin{cases} 0 & \text{für } x < 1, \\ 1 & \text{für } x = 1 \end{cases}$$

für alle $x \in D$, aber die Grenzfunktion ist hier offensichtlich *nicht* stetig! Wir halten also fest:

> Konvergiert eine Folge stetiger Funktionen $f_n \colon D \to \mathbb{K}$ punktweise gegen eine Grenzfunktion $f \colon D \to \mathbb{K}$, so muss f nicht notwendig stetig sein!

Analog zum Fall der Umordnungen von Reihen in Beispiel 7.9 wird auch hier der Ausweg aus diesem Problem darin bestehen, eine stärkere Form der Konvergenz einer Folge stetiger Funktionen einzuführen, die letztlich die Stetigkeit der Grenzfunktion sicherstellt.

In der Tat können wir an unserem obigen Beispiel $f_n(x) = x^n$ auch schon motivieren, wie dieses stärkere Kriterium aussehen wird, denn man sieht an diesem Bild bereits recht deutlich, wo das Problem liegt: Es ist zwar richtig, dass für jedes $x \in [0,1)$ der Grenzwert $\lim\limits_{n\to\infty} x^n$ gleich 0 ist, d. h. dass $x^n < \varepsilon$ für alle n ab einem gewissen n_0 gilt – aber dieses n_0 hängt extrem vom betrachteten Punkt x ab und wird insbesondere für $x \to 1$ immer größer. So kann man z. B. für den Wert $x = a$ im Bild oben rechts noch $n_0 = 1$ wählen, beim Wert $x = b$ braucht man mindestens $n_0 = 3$, beim Wert $x = c$ schon mindestens $n_0 = 5$. Je weiter sich x dem Wert 1 nähert, um so größer muss man dieses n_0 wählen – bis es im Grenzfall $x = 1$ schließlich gar kein solches n_0 mehr gibt, so dass 0 nicht mehr der Grenzwert der Folge $\lim\limits_{n\to\infty} x^n$ ist. Im Bild oben links hingegen kann man für die dargestellte ε-Umgebung um f z. B. $n_0 = 3$ *für alle* x (in dem dort betrachteten Intervall $[0,1]$) wählen, denn $f_3, f_4, f_5, \dots$ liegen komplett in dem grau eingezeichneten Streifen.

Es kommt bei der Grenzwertdefinition also anscheinend darauf an, ob man das verlangte n_0 unabhängig vom betrachteten Punkt x wählen kann. Dies führt auf die folgende Definition:

Definition 8.33 (Gleichmäßige Konvergenz)**.** Es seien $D \subset \mathbb{K}$ und $f \colon D \to \mathbb{K}$ eine Funktion. Weiterhin sei für alle $n \in \mathbb{N}$ eine Funktion $f_n \colon D \to \mathbb{K}$ gegeben – man nennt $(f_n)_n$ dann auch eine **Funktionenfolge** auf D.

(a) Ist $\lim\limits_{n\to\infty} f_n(x) = f(x)$ für alle $x \in D$, d. h. gilt

$$\forall x \in D\ \forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0 : |f_n(x) - f(x)| < \varepsilon,$$

so nennt man $(f_n)_n$ **punktweise konvergent** gegen f. Beachte, dass das n_0 hierbei nicht nur von ε, sondern auch vom betrachteten Punkt x abhängen darf.

(b) Kann man n_0 in (a) auch unabhängig von x wählen, d. h. gilt sogar

$$\forall \varepsilon > 0 \; \exists n_0 \in \mathbb{N} \; \forall x \in D \; \forall n \geq n_0 : |f_n(x) - f(x)| < \varepsilon,$$

so heißt die Funktionenfolge $(f_n)_n$ auf D **gleichmäßig konvergent** gegen f.

Bemerkung 8.34.

(a) Beachte, dass die gleichmäßige Konvergenz nach Definition kein „punktweises Konzept“ ist, also nicht an jedem Punkt der Definitionsmenge D separat überprüft werden kann. Es ergibt also z. B. keinen Sinn, zu sagen, eine Funktionenfolge auf D sei „in jedem Punkt von D gleichmäßig konvergent“. Stattdessen muss man bei der Bestimmung der gleichmäßigen Konvergenz immer alle Punkte von D gleichzeitig betrachten.

(b) Natürlich ist jede gleichmäßig konvergente Funktionenfolge $(f_n)_n$ auf D auch punktweise konvergent mit der gleichen Grenzfunktion f.

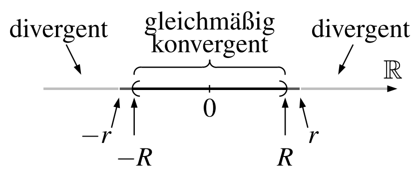
Wollen wir also die gleichmäßige Konvergenz von $(f_n)_n$ untersuchen, so werden wir in der Regel zunächst mit der punktweisen Konvergenz beginnen und für alle $x \in D$ die Grenzwerte $f(x) := \lim_{n\to\infty} f_n(x)$ bestimmen. Existiert dann einer dieser Grenzwerte nicht, so ist die Funktionenfolge damit nicht punktweise, also auch nicht gleichmäßig konvergent. Ansonsten ist die so bestimmte Funktion f die Grenzfunktion, mit der wir für die gleichmäßige Konvergenz die Bedingung aus Definition 8.33 (b) überprüfen müssen.

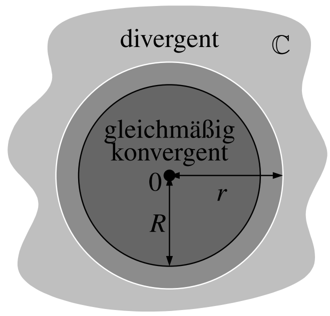
Unser wichtigstes Beispiel von Funktionenfolgen sind Potenzreihen wie z. B. die Exponentialreihe in Beispiel 8.32, und glücklicherweise sind diese in folgendem Sinne immer gleichmäßig konvergent.

Satz 8.35 (Gleichmäßige Konvergenz von Potenzreihen)**.** *Jede Potenzreihe $\sum_{k=0}^{\infty} a_k x^k$ in $\mathbb{K}$ mit Konvergenzradius r ist gleichmäßig konvergent auf jeder Menge der Form*

$$K_R := \{x \in \mathbb{K} : |x| \leq R\} \qquad \text{für } 0 \leq R < r$$

(d. h. die Folge $(f_n)_n$ der Partialsummen $f_n(x) = \sum_{k=0}^{n} a_k x^k$ konvergiert gleichmäßig auf jedem K_R gegen die Grenzfunktion f mit $f(x) = \sum_{k=0}^{\infty} a_k x^k$).





Mit anderen Worten konvergieren Potenzreihen also gleichmäßig auf jedem abgeschlossenen Intervall (für $\mathbb{K} = \mathbb{R}$) bzw. Kreis (für $\mathbb{K} = \mathbb{C}$) innerhalb des Konvergenzgebiets.

Beweis. Wir weisen das Kriterium aus Definition 8.33 (b) direkt nach. Es sei dazu $\varepsilon > 0$ beliebig. Wegen $R < r$ konvergiert die Reihe $\sum_{k=0}^{\infty} a_k R^k$ nach Satz 7.26 absolut. Es gibt also ein $n_0 \in \mathbb{N}$, so dass

$$\left| \sum_{k=0}^{\infty} |a_k| \cdot R^k - \sum_{k=0}^{n} |a_k| \cdot R^k \right| = \sum_{k=n+1}^{\infty} |a_k| \cdot R^k < \varepsilon$$

für alle $n \geq n_0$ gilt. Dann folgt für alle $n \geq n_0$ und $x \in \mathbb{K}$ mit $|x| \leq R$ aber auch

$$|f(x) - f_n(x)| = \left| \sum_{k=0}^{\infty} a_k x^k - \sum_{k=0}^{n} a_k x^k \right| = \left| \sum_{k=n+1}^{\infty} a_k x^k \right| \leq \sum_{k=n+1}^{\infty} |a_k| \cdot |x|^k \leq \sum_{k=n+1}^{\infty} |a_k| \cdot R^k < \varepsilon.$$

Da wir unser n_0 hierbei unabhängig von $x \in K_R$ wählen konnten, ist die Potenzreihe auf K_R also gleichmäßig konvergent. □

Beachte, dass man den Wert von R in Satz 8.35 beliebig nahe an r wählen darf. Insbesondere findet man also zu jedem $x \in \mathbb{K}$ im Konvergenzgebiet $D = \{x \in \mathbb{K} : |x| < r\}$ ein R, so dass x in K_R enthalten ist. Da die gleichmäßige Konvergenz gemäß Bemerkung 8.34 (a) nicht punktweise überprüft werden kann, bedeutet dies jedoch *nicht*, dass die Potenzreihe auch auf ganz D gleichmäßig konvergiert! Hier ist ein einfaches Gegenbeispiel dafür:

Beispiel 8.36. Die geometrische Reihe $f(x) = \sum_{k=0}^{\infty} x^k$ hat nach Beispiel 7.3 (a) das Konvergenzgebiet $D = \{x \in \mathbb{K} : |x| < 1\}$, also den Konvergenzradius 1. Wir behaupten, dass f auf D nicht gleichmäßig konvergent ist, d. h. dass die Umkehrung der Bedingung aus Definition 8.33 (b)

$$\exists \varepsilon > 0 \; \forall n_0 \in \mathbb{N} \; \exists x \in D \; \exists n \geq n_0 : \left| \sum_{k=0}^{\infty} x^k - \sum_{k=0}^{n} x^k \right| \geq \varepsilon$$

gilt. Dazu wählen wir $\varepsilon := 1$; es sei $n_0 \in \mathbb{N}$ gegeben, und wir setzen $n := n_0$. Für alle $x \in D$ mit $x > 0$ ist nun nach der Formel für die geometrische Reihe aus Beispiel 7.3 (a)

$$\left| \sum_{k=0}^{\infty} x^k - \sum_{k=0}^{n} x^k \right| = \sum_{k=n+1}^{\infty} x^k = x^{n+1} \cdot \sum_{k=0}^{\infty} x^k = \frac{x^{n+1}}{1-x}.$$

Nähert sich x nun innerhalb von D dem Wert 1, so wächst dieser Ausdruck offensichtlich unbeschränkt an. Also gibt es insbesondere ein $x \in D$, für das dieser Ausdruck mindestens gleich $1 = \varepsilon$ ist, was zu zeigen war.

Wir kommen nun zu dem zentralen Satz, der in Beispiel 8.32 die Motivation für die Einführung der gleichmäßigen Konvergenz war:

Satz 8.37 (Der gleichmäßige Grenzwert stetiger Funktionen ist stetig)**.** *Es seien $D \subset \mathbb{K}$ und $(f_n)_n$ eine Folge stetiger Funktionen $f_n \colon D \to \mathbb{K}$, die gleichmäßig gegen eine Grenzfunktion $f \colon D \to \mathbb{K}$ konvergiert. Dann ist auch f stetig.*

Beweis. Wir weisen das ε-δ-Kriterium aus Definition 8.5 (a) für f nach. Es seien dazu $a \in D$ und $\varepsilon > 0$ beliebig. Da $(f_n)_n$ gleichmäßig gegen f konvergiert, gibt es ein $n_0 \in \mathbb{N}$, so dass

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3} \quad \text{für alle } x \in D \text{ und alle } n \geq n_0 \tag{1}$$

gilt (insbesondere also auch für $x = a$). Wir können hier also $n := n_0$ setzen. Wegen der Stetigkeit von f_n gibt es dann ein $\delta > 0$ mit

$$|f_n(x) - f_n(a)| < \frac{\varepsilon}{3} \quad \text{für alle } x \in D \text{ mit } |x - a| < \delta. \tag{2}$$

Insgesamt folgt damit für alle $x \in D$ mit $|x - a| < \delta$ nach der Dreiecksungleichung

$$\begin{aligned} |f(x) - f(a)| &= |f(x) - f_n(x) + f_n(x) - f_n(a) + f_n(a) - f(a)| \\ &\leq \underbrace{|f(x) - f_n(x)|}_{< \frac{\varepsilon}{3} \text{ nach (1)}} + \underbrace{|f_n(x) - f_n(a)|}_{< \frac{\varepsilon}{3} \text{ nach (2)}} + \underbrace{|f_n(a) - f(a)|}_{< \frac{\varepsilon}{3} \text{ nach (1)}} \\ &< \varepsilon. \end{aligned}$$

□

Folgerung 8.38 (Stetigkeit von Potenzreihen)**.** *Jede Potenzreihe in $\mathbb{K}$ ist in ihrem Konvergenzgebiet stetig.*

Beweis. Wir betrachten eine Potenzreihenfunktion f in $\mathbb{K}$ mit Konvergenzradius r, also eine Funktion der Form $f(x) = \lim_{n \to \infty} f_n(x)$ mit $f_n(x) = \sum_{k=0}^{n} a_k x^k$ für alle $x \in \mathbb{K}$ mit $|x| < r$.

Es sei nun ein $c \in \mathbb{K}$ mit $|c| < r$ gegeben; wir müssen zeigen, dass f in c stetig ist. Wähle dazu ein $R > 0$ mit $|c| < R < r$. Dann ist f nach Satz 8.35 auf $K_R = \{x \in \mathbb{K} : |x| \leq R\}$ der gleichmäßige Grenzwert der stetigen Partialsummen f_n. Also ist diese Grenzfunktion f nach Satz 8.37 auf K_R und damit insbesondere in c stetig. □

Beispiel 8.39.

(a) Die Exponentialfunktion $\exp\colon \mathbb{K} \to \mathbb{K}$ ist als Potenzreihe mit Konvergenzradius ∞ nach Folgerung 8.38 auf ganz $\mathbb{K}$ stetig.

(b) Die reelle Funktionenfolge $(x^n)_n$ auf $[0,1]$ aus Beispiel 8.32 ist nach Satz 8.37 nicht gleichmäßig konvergent, da ihre Grenzfunktion nicht stetig ist. (Natürlich könnte man dies auch analog zu Beispiel 8.36 direkt nachrechnen.)

Aufgabe 8.40 (Supremumsnorm)**.**

(a) Zu einer Funktion $f\colon D \to \mathbb{K}$ mit $D \subset \mathbb{K}$ definieren wir die **Supremumsnorm**

$$\|f\|_{\sup} := \sup\{|f(x)| : x \in D\} \quad \in \mathbb{R}_{\geq 0} \cup \{\infty\}.$$

Zeige, dass eine Funktionenfolge $(f_n)_n$ auf D genau dann gleichmäßig gegen f konvergiert, wenn $\|f_n - f\|_{\sup} \to 0$ für $n \to \infty$ gilt.

(b) Zeige, dass die reelle Funktionenfolge $(f_n)_n$ mit $f_n(x) = \frac{nx}{1+nx}$ zwar nicht auf $\mathbb{R}_{>0}$, aber auf jedem Intervall $[a, \infty)$ mit $a > 0$ gleichmäßig konvergiert.

Aufgabe 8.41.

(a) Für $n \in \mathbb{N}_{>0}$ sei $f_n\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto \sqrt{x^2 + \frac{1}{n}}$. Zeige, dass die Funktionenfolge $(f_n)_n$ auf $\mathbb{R}$ gleichmäßig konvergiert.

(b) Zeige, dass die Exponentialreihe auf $\mathbb{R}$ nicht gleichmäßig konvergiert.

Aufgabe 8.42 (Koeffizientenvergleich für Potenzreihen)**.** Es sei $f(x) = \sum_{n=0}^{\infty} a_n x^n$ eine Potenzreihe über $\mathbb{K}$ mit Konvergenzradius mindestens $r > 0$.

(a) Zeige mit vollständiger Induktion: Ist $f(x) = 0$ für alle $x \in \mathbb{K}$ mit $|x| < r$, so gilt bereits $a_n = 0$ für alle $n \in \mathbb{N}$ (d. h. ist der Wert der Reihe gleich 0 für alle diese x, so sind bereits alle Koeffizienten der Reihe gleich 0).

(b) Man zeige: Ist $g(x) = \sum_{n=0}^{\infty} b_n x^n$ eine weitere Potenzreihe mit Konvergenzradius mindestens r und gilt $f(x) = g(x)$ für alle $x \in \mathbb{K}$ mit $|x| < r$, so ist bereits $a_n = b_n$ für alle $n \in \mathbb{N}$.

Zum Abschluss dieses Kapitels wollen wir nun noch das Konzept der *gleichmäßigen Stetigkeit* einführen, das wir später (z. B. in Satz 12.12) noch benötigen werden und das eine sehr ähnliche Idee wie die gleichmäßige Konvergenz hat:

Definition 8.43 (Gleichmäßige Stetigkeit)**.** Es seien $D \subset \mathbb{K}$ und $f\colon D \to \mathbb{K}$ eine Funktion. Bekanntlich heißt die Funktion f nach Definition 8.5 stetig, wenn sie in jedem Punkt $a \in D$ stetig ist, also wenn gilt

$$\forall a \in D\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in D : |x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon.$$

Beachte, dass δ hierbei (analog zu n_0 in Definition 8.33 (a)) nicht nur von ε, sondern auch vom betrachteten Punkt a abhängen darf. Kann man δ jedoch auch unabhängig von a wählen, gilt also

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall a \in D\ \forall x \in D : |x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon,$$

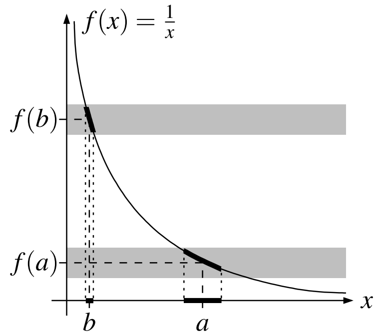
so nennt man f auf D **gleichmäßig stetig**.

Bemerkung 8.44. Genau wie die gleichmäßige Konvergenz (siehe Bemerkung 8.34 (a)) ist auch die gleichmäßige Stetigkeit einer Funktion $f\colon D \to \mathbb{K}$ kein punktweises Konzept, sondern kann nur nachgewiesen werden, indem man alle Punkte von D gleichzeitig betrachtet.

Beispiel 8.45. Wir behaupten, dass die nach Beispiel 8.14 stetige Funktion

$$f\colon \mathbb{R}_{>0} \to \mathbb{R},\ x \mapsto \frac{1}{x}$$

nicht gleichmäßig stetig ist. Anschaulich ist diese Aussage im Bild rechts dargestellt: Die Stetigkeit von f bedeutet ja gerade, dass wir zu jeder ε-Umgebung $U_\varepsilon(f(x))$ eines Bildpunktes $f(x)$ eine δ-Umgebung von x finden, die ganz nach $U_\varepsilon(f(x))$ abgebildet wird. Zum Punkt $x = a$ haben wir in der Skizze eine ε-Umgebung von $f(a)$ grau eingezeichnet, und auf der x-Achse eine dazu passende δ-Umgebung wie in Bemerkung 8.6 dick markiert. Wenn wir nun auf einen (viel) kleineren Wert $x = b$ gehen und das gleiche ε wie oben behalten, so sehen wir, dass wir das zugehörige δ viel kleiner wählen müssen. Wenn sich x der 0 nähert, müssen wir bei gleich bleibendem ε das δ in der Tat sogar gegen 0 gehen lassen. Dies bedeutet, dass wir für festgehaltenes ε kein δ finden können, das in *jedem* Punkt $x > 0$ funktioniert – und das wiederum bedeutet genau, dass f nicht gleichmäßig stetig ist.



Wollen wir diese Aussage formal beweisen, so müssen wir die Negation der Bedingung aus Definition 8.43 nachweisen, d. h.

$$\exists \varepsilon > 0 \; \forall \delta > 0 \; \exists a \in D \; \exists x \in D : |x-a| < \delta \text{ und } |f(x) - f(a)| \geq \varepsilon.$$

Dies ist schnell gezeigt: Wir setzen $\varepsilon := 1$; es sei $\delta > 0$ beliebig. Dann wählen wir $x = \delta$ und $a = \frac{\delta}{1+\delta}$, und es gilt wegen $x > a$

$$|x-a| = \delta - \frac{\delta}{1+\delta} < \delta \qquad \text{und} \qquad |f(x)-f(a)| = \left| \frac{1}{\delta} - \frac{1+\delta}{\delta} \right| = 1 \geq \varepsilon.$$

Wir wollen nun aber zeigen, dass eine stetige Funktion auf einem abgeschlossenen reellen Intervall glücklicherweise immer gleichmäßig stetig ist, so dass wir in diesem Fall die eigentlich stärkere Bedingung der gleichmäßigen Stetigkeit immer umsonst mitgeliefert bekommen.

Satz 8.46. *Jede stetige Funktion $f\colon [a,b] \to \mathbb{R}$ auf einem abgeschlossenen reellen Intervall ist dort auch gleichmäßig stetig.* 20

Beweis. Angenommen, f wäre nicht gleichmäßig stetig. Dann würde wie in Beispiel 8.45 das Gegenteil der Bedingung aus Definition 8.43 gelten, d. h.

$$\exists \varepsilon > 0 \; \forall \delta > 0 \; \exists x, y \in [a,b] : |y-x| < \delta \text{ und } |f(y) - f(x)| \geq \varepsilon.$$

Wählen wir ein solches ε, so finden wir also mit $\delta = \frac{1}{n}$ zu jedem $n \in \mathbb{N}_{>0}$ Zahlen $x_n, y_n \in [a,b]$ mit $|y_n - x_n| < \frac{1}{n}$ und $|f(y_n) - f(x_n)| \geq \varepsilon$. Insbesondere ist dann $\lim\limits_{n\to\infty}(y_n - x_n) = 0$. Wählen wir nun nach dem Satz 6.21 von Bolzano-Weierstraß eine gegen ein $x \in [a,b]$ konvergente Teilfolge $(x_{n_k})_k$ von $(x_n)_n$, so folgt

$$\lim_{k\to\infty} x_{n_k} = x \qquad \text{und} \qquad \lim_{k\to\infty} y_{n_k} = \lim_{k\to\infty} x_{n_k} + \lim_{k\to\infty}(y_{n_k} - x_{n_k}) = x + 0 = x,$$

d. h. die entsprechende Teilfolge von $(y_n)_n$ konvergiert ebenfalls gegen x. Wegen der Stetigkeit von f in x ergibt sich dann aber nach dem Folgenkriterium aus Satz 8.11 (b)

$$\lim_{k\to\infty}(f(y_{n_k}) - f(x_{n_k})) = \lim_{k\to\infty} f(y_{n_k}) - \lim_{k\to\infty} f(x_{n_k}) = f(x) - f(x) = 0,$$

im Widerspruch zu $|f(y_{n_k}) - f(x_{n_k})| \geq \varepsilon$ für alle k. □

Aufgabe 8.47 (Lipschitz-Stetigkeit)**.** Es sei $D \subset \mathbb{K}$. Eine Funktion $f\colon D \to \mathbb{K}$ heißt **Lipschitz-stetig**, wenn es ein $L \in \mathbb{R}_{\geq 0}$ gibt, so dass

$$|f(x) - f(y)| \leq L \cdot |x-y|$$

für alle $x, y \in D$. Man nennt L in diesem Fall eine **Lipschitz-Konstante** für f. Man zeige:

(a) Ist f Lipschitz-stetig, so ist f auch gleichmäßig stetig.

(b) Die Wurzelfunktion $f\colon [0,1] \to \mathbb{R}, x \mapsto \sqrt{x}$ ist gleichmäßig stetig, aber nicht Lipschitz-stetig.

9. Spezielle Funktionen

Nachdem wir jetzt schon recht viel allgemeine Theorie kennengelernt haben, wollen wir diese nun anwenden, um einige bekannte spezielle Funktionen zu studieren (oder überhaupt erst einmal exakt zu definieren), die ihr bereits aus der Schule kennt: die Exponential- und Logarithmusfunktion, die allgemeine Potenz x^a für $a \in \mathbb{R}$, die Winkelfunktionen und ihre Umkehrfunktionen. Ausgangspunkt aller dieser Funktionen ist letztlich die in Definition 7.25 (b) bereits eingeführte Exponentialfunktion

$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots \quad \text{für } x \in \mathbb{K}.$$

Aus Folgerung 7.35 wissen wir schon, dass diese Funktion die Funktionalgleichung

$$\exp(x+y) = \exp(x) \cdot \exp(y) \quad \text{für alle } x, y \in \mathbb{K}$$

erfüllt. Außerdem ist sie nach Beispiel 8.39 (a) stetig, und aus der Reihendarstellung sieht man sofort, dass $\exp(0) = 1$.

Die weiteren Eigenschaften der Exponentialfunktion sind im reellen und komplexen Fall trotz der gleich lautenden Definition sehr unterschiedlich. Wir werden diese beiden Fälle im Folgenden daher separat untersuchen.

9.A Logarithmen und allgemeine Potenzen

Wir beginnen mit der reellen Exponentialfunktion und zeigen zunächst einige ihrer wichtigen Eigenschaften.

Satz 9.1 (Eigenschaften der reellen Exponentialfunktion)**.**

(a) *Es gilt* $\exp(x) > 0$ *für alle* $x \in \mathbb{R}$.

(b) *Die Funktion* $\exp\colon \mathbb{R} \to \mathbb{R}$ *ist streng monoton wachsend.*

(c) *Für alle* $n \in \mathbb{N}$ *gilt*

$$\lim_{x\to\infty} \frac{\exp(x)}{x^n} = \infty \qquad \textit{und} \qquad \lim_{x\to-\infty} x^n \exp(x) = 0.$$

Insbesondere ist also $\lim_{x\to\infty} \exp(x) = \infty$ *und* $\lim_{x\to-\infty} \exp(x) = 0$.

(d) *Für die Zahl* $\mathrm{e} := \exp(1)$ *gilt* $2 < \mathrm{e} < 3$. *(Man nennt* e *die* ***Eulersche Zahl****; eine explizite näherungsweise Berechnung der Exponentialreihe zeigt, dass* $\mathrm{e} = 2{,}71828\ldots$*.)*

Beweis.

(a) Für $x \geq 0$ ist dies aus der Reihendarstellung offensichtlich. Für $x \leq 0$ folgt aus der Funktionalgleichung

$$\exp(x) \cdot \exp(-x) = \exp(0) = 1 \quad \text{und damit} \quad \exp(x) = \frac{1}{\exp(-x)},$$

was nun wegen $-x \geq 0$ ebenfalls positiv ist.

(b) Es seien $x, y \in \mathbb{R}$ mit $x < y$. Wegen $y - x > 0$ folgt aus der Reihendarstellung der Exponentialfunktion dann $\exp(y-x) > 1$. Da nach (a) außerdem $\exp(x) > 0$ gilt, erhalten wir mit der Funktionalgleichung wie gewünscht

$$\exp(y) = \exp(x) \cdot \exp(y-x) > \exp(x) \cdot 1 = \exp(x).$$

(c) Für $x > 0$ ergibt sich aus der Reihendarstellung der Exponentialfunktion natürlich

$$\exp(x) > \frac{x^{n+1}}{(n+1)!} \quad \text{und damit} \quad \frac{\exp(x)}{x^n} > \frac{x}{(n+1)!}.$$

für alle $n \in \mathbb{N}$. Wegen $\lim\limits_{x\to\infty} \frac{x}{(n+1)!} = \infty$ folgt damit auch $\lim\limits_{x\to\infty} \frac{\exp(x)}{x^n} = \infty$.

Die Aussage für $x \to -\infty$ zeigt man analog: Für $x < 0$ ist $-x > 0$, und da wir in (a) schon gesehen haben, dass $\exp(x) = \frac{1}{\exp(-x)}$ gilt, erhalten wir

$$\exp(x) = \frac{1}{\exp(-x)} < \frac{1}{(-x)^{n+1}/(n+1)!} \quad \text{und damit} \quad |x^n \exp(x)| < \frac{(n+1)!}{|x|}.$$

Wegen $\lim\limits_{x\to-\infty} \frac{(n+1)!}{|x|} = 0$ ergibt sich damit auch $\lim\limits_{x\to-\infty} x^n \exp(x) = 0$.

(d) Aus der Exponentialreihe erhalten wir sofort

$$\mathrm{e} > \frac{1}{0!} + \frac{1}{1!} = 2,$$

und wegen $n! = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot n \geq 1 \cdot 2 \cdot 2 \cdot \cdots \cdot 2 = 2^{n-1}$ mit Hilfe der geometrischen Reihe aus Beispiel 7.3 (a) auch

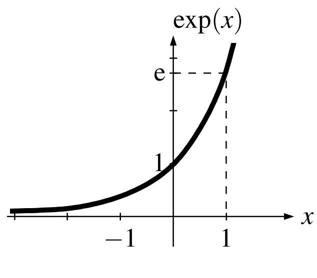
$$\mathrm{e} = \frac{1}{0!} + \sum_{n=1}^{\infty} \frac{1}{n!} < 1 + \sum_{n=1}^{\infty} \frac{1}{2^{n-1}} = 1 + \sum_{n=0}^{\infty} \frac{1}{2^n} = 1 + \frac{1}{1-\frac{1}{2}} = 3. \qquad \square$$

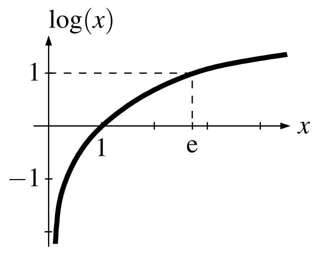
Aufgabe 9.2 (Irrationalität von e)**.** Finde für alle $n \in \mathbb{N}_{>0}$ explizit eine natürliche Zahl $a \in \mathbb{N}$ mit $\frac{a}{n!} < \mathrm{e} < \frac{a+1}{n!}$, und zeige so, dass e irrational ist.

Bemerkung 9.3.

(a) Da die (uneigentlichen) Grenzwerte von $\exp(x)$ für $x \to \pm\infty$ nach Satz 9.1 (b) gleich ∞ bzw. 0 sind, bedeutet die Aussage desselben Satzes für $n > 0$ gerade, dass sich in diesen Grenzwerten, die ja von der Form $\frac{\infty}{\infty}$ bzw. $\pm\infty \cdot 0$ sind, die Exponentialfunktion gegenüber der Potenz x^n durchsetzt. Man sagt auch, „die Exponentialfunktion ist für $x \to \pm\infty$ stärker als jede Potenz“.

(b) Im Bild unten links haben wir den Graphen der reellen Exponentialfunktion gemäß Satz 9.1 skizziert. Da exp nach Beispiel 8.39 (a) stetig und nach Satz 9.1 (b) streng monoton wachsend ist, existiert nach Satz 8.27 eine Umkehrfunktion (wie in Beispiel 8.28 zunächst für Start- und Zielmenge $[-R,R] \to [\exp(-R), \exp(R)]$ für alle $R > 0$, dann durch den Übergang $R \to \infty$ aber auch für $\mathbb{R} \to \mathbb{R}_{>0}$):





Definition 9.4 (Logarithmus)**.** Die Umkehrfunktion der Exponentialfunktion $\exp\colon \mathbb{R} \to \mathbb{R}_{>0}$ wird mit

$$\log\colon \mathbb{R}_{>0} \to \mathbb{R},\ x \mapsto \log(x)$$

bezeichnet und heißt die (natürliche) **Logarithmusfunktion**. Sie ist im Bild oben rechts dargestellt. Statt $\log(x)$ ist oft auch die Bezeichnung $\ln(x)$ üblich.

Notation 9.5 (Schreibweise spezieller Funktionen)**.** Bei den speziellen Funktionen, die wir in diesem Kapitel kennenlernen werden, ist es zur Vereinfachung der Notation oft üblich, die Klammern beim Funktionsargument wegzulassen, wenn es sich nur um eine einfache Zahl oder Variable handelt, also z. B. $\log x$ statt $\log(x)$ zu schreiben. Ist das Funktionsargument jedoch ein zusammengesetzter Ausdruck, sind die Klammern zwingend erforderlich: $\log(x+y)$ kann man nicht als $\log x + y$ schreiben, da $\log x + y$ immer als $(\log x) + y$ zu verstehen ist.

Bemerkung 9.6 (Eigenschaften der Logarithmusfunktion)**.** Unsere bisher gezeigten Eigenschaften der Exponentialfunktion übertragen sich natürlich sofort auf die Logarithmusfunktion:

(a) log ist stetig und streng monoton wachsend nach Satz 8.27.

(b) $\log 1 = 0$ und $\log \mathrm{e} = 1$.

(c) Die Grenzwerte aus Satz 9.1 (b) übertragen sich durch Vertauschen von Start- und Zielraum auf den Logarithmus als $\lim\limits_{x\to\infty} \log x = \infty$ und $\lim\limits_{x\to 0} \log x = -\infty$.

(d) Wenden wir die Funktionalgleichung der Exponentialfunktion auf $\log x$ und $\log y$ für $x, y > 0$ an, so erhalten wir

$$\exp(\log x + \log y) = \exp(\log x) \cdot \exp(\log y) = x \cdot y$$

und damit durch Logarithmieren

$$\log x + \log y = \log(x \cdot y) \qquad \text{für } x, y \in \mathbb{R}_{>0},$$

was die **Funktionalgleichung der Logarithmusfunktion** genannt wird.

Eine der wichtigsten Anwendungen der Logarithmusfunktion ist, dass man mit ihr allgemeine Potenzen definieren kann – also Potenzen der Form x^a, wobei a nun nicht mehr wie bisher in $\mathbb{Z}$ liegen muss, sondern eine beliebige reelle Zahl sein kann:

Definition 9.7 (Allgemeine Potenzen)**.** Für $x \in \mathbb{R}_{>0}$ und $a \in \mathbb{R}$ definieren wir die **Potenz**

$$x^a := \exp(a \log x)$$

(wir werden in Bemerkung 9.9 (a) noch sehen, dass dies für $a \in \mathbb{Z}$ mit unserer alten Definition aus Notation 3.9 (b) übereinstimmt – was dann auch diese neue, allgemeinere Definition motiviert).

Lemma 9.8 (Rechenregeln für allgemeine Potenzen)**.** *Für alle $x, y \in \mathbb{R}_{>0}$ und $a, b \in \mathbb{R}$ gilt*

(a) $x^0 = 1$ *und* $x^1 = x$;

(b) $x^{a+b} = x^a \cdot x^b$ *und* $x^{-a} = \frac{1}{x^a}$;

(c) $x^{ab} = (x^a)^b$;

(d) $(xy)^a = x^a \cdot y^a$.

Beweis. Alle Beweise sind einfaches Nachrechnen mit Hilfe der Funktionalgleichung:

(a) $x^0 = \exp(0 \cdot \log x) = \exp(0) = 1$ und $x^1 = \exp(\log x) = x$.

(b) Es ist

$$x^{a+b} = \exp((a+b)\log x) = \exp(a \log x + b \log x) = \exp(a \log x) \cdot \exp(b \log x) = x^a \cdot x^b.$$

Setzen wir in dieser Gleichung $b = -a$, so ergibt sich ferner $1 = x^a \cdot x^{-a}$ und damit $x^{-a} = \frac{1}{x^a}$.

(c) Es gilt

$$(x^a)^b = \exp(b \log(x^a)) = \exp(b \log(\exp(a \log x))) = \exp(ab \log x) = x^{ab}.$$

(d) Es ist

$$(xy)^a = \exp(a \log(xy)) = \exp(a \log x + a \log y) = \exp(a \log x) \cdot \exp(a \log y) = x^a \cdot y^a. \qquad \square$$

Bemerkung 9.9.

(a) Aus Lemma 9.8 (a) und (b) folgt insbesondere, dass im Fall $a \in \mathbb{N}$ für unsere allgemeine Potenz aus Definition 9.7

$$x^a = x^{1+\cdots+1} = \underbrace{x \cdot \cdots \cdot x}_{a\text{-mal}} \qquad \text{und genauso} \qquad x^{-a} = x^{-1-\cdots-1} = \underbrace{\frac{1}{x} \cdot \cdots \cdot \frac{1}{x}}_{a\text{-mal}}$$

gilt, dass sie dann also mit der alten Definition der Potenz aus Notation 3.9 (b) übereinstimmt.

(b) Nach Lemma 9.8 (c) ist $x \mapsto x^{\frac{1}{a}}$ für $a \neq 0$ eine Umkehrfunktion zu $x \mapsto x^a$, denn es ist

$$(x^a)^{\frac{1}{a}} = x^1 = x \qquad \text{und} \qquad \left(x^{\frac{1}{a}}\right)^a = x^1 = x$$

für alle $x \in \mathbb{R}_{>0}$ und $a \in \mathbb{R}\backslash\{0\}$.

Da Umkehrfunktionen eindeutig sind und wir im Fall $a \in \mathbb{N}_{>0}$ aus Aufgabe 5.37 und Beispiel 8.28 bereits wissen, dass die Wurzelfunktion $x \mapsto \sqrt[a]{x}$ ebenfalls eine solche Umkehrfunktion ist, sehen wir also, dass $x^{\frac{1}{a}} = \sqrt[a]{x}$ für alle $x \in \mathbb{R}_{>0}$ und $a \in \mathbb{N}_{>0}$ gilt.

(c) Mit der Eulerschen Zahl e aus Satz 9.1 (d) ist offensichtlich $\mathrm{e}^a = \exp(a \log \mathrm{e}) = \exp a$ für alle $a \in \mathbb{R}$. Man verwendet daher in der Regel die einfachere Potenzschreibweise e^a für $\exp a$.

(d) Beachte, dass wir die allgemeine Potenz x^a mit $a \in \mathbb{R}$ nur für positive x definieren konnten, weil für negative Zahlen kein Logarithmus existiert. In der Tat ist es auch einleuchtend, dass ein Ausdruck wie z. B. $(-1)^{\sqrt{2}}$ nicht sinnvoll definiert werden kann, da nicht einmal klar ist, ob diese Zahl positiv oder negativ sein sollte.

9.B Winkelfunktionen

Nach der reellen wollen wir nun die komplexe Exponentialfunktion studieren, die uns schließlich zu den Winkelfunktionen führen wird. Wie in Bemerkung 9.9 (c) werden wir dabei die Exponentialfunktion auch im Komplexen in der Regel mit $z \mapsto \mathrm{e}^z$ bezeichnen (obwohl es keine allgemeine Potenz w^z für $w, z \in \mathbb{C}$ gibt). Ihre wesentlichen Eigenschaften, die wir benötigen, um den Zusammenhang mit Winkelfunktionen herstellen zu können, sind die folgenden:

Satz 9.10 (Eigenschaften der komplexen Exponentialfunktion)**.** *Es gilt:*

(a) $\mathrm{e}^{\overline{z}} = \overline{\mathrm{e}^z}$ *für alle* $z \in \mathbb{C}$*;*

(b) $|\mathrm{e}^{\mathrm{i}x}| = 1$ *für alle* $x \in \mathbb{R}$.

Beweis.

(a) Für die Partialsummen $f_n(z) := \sum_{k=0}^n \frac{z^k}{k!}$ der Exponentialfunktion folgt natürlich $f_n(\overline{z}) = \overline{f_n(z)}$ durch fortgesetztes Anwenden von Lemma 6.9 (a). Da die komplexe Konjugation $z \mapsto \overline{z}$ nach Beispiel 8.7 (c) stetig ist, ergibt sich also nach dem Folgenkriterium für Stetigkeit aus Satz 8.11 (b)

$$\mathrm{e}^{\overline{z}} = \lim_{n\to\infty} f_n(\overline{z}) = \lim_{n\to\infty} \overline{f_n(z)} = \overline{\lim_{n\to\infty} f_n(z)} = \overline{\mathrm{e}^z}.$$

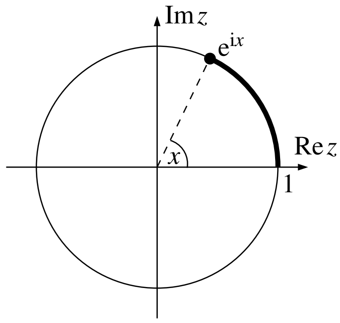
(b) Wegen $|z| = \sqrt{z\overline{z}}$ (siehe Bemerkung 6.4) erhalten wir nun mit (a)

$$|\mathrm{e}^{\mathrm{i}x}| = \sqrt{\mathrm{e}^{\mathrm{i}x} \cdot \overline{\mathrm{e}^{\mathrm{i}x}}} = \sqrt{\mathrm{e}^{\mathrm{i}x} \cdot \mathrm{e}^{\overline{\mathrm{i}x}}} = \sqrt{\mathrm{e}^{\mathrm{i}x} \cdot \mathrm{e}^{-\mathrm{i}x}} = \sqrt{\mathrm{e}^{\mathrm{i}x-\mathrm{i}x}} = \sqrt{\mathrm{e}^0} = 1. \qquad \square$$

Bemerkung 9.11. Satz 9.10 (b) besagt gerade, dass die komplexe Zahl e^{ix} für *reelle* x immer auf dem Rand des Einheitskreises liegt. Multiplizieren wir zwei solche Zahlen e^{ix} und e^{iy} für $x,y \in \mathbb{R}$ miteinander, so erhalten wir einerseits nach der Funktionalgleichung der Exponentialfunktion die Zahl

$$e^{ix} \cdot e^{iy} = e^{i(x+y)}$$

(d. h. die Exponenten addieren sich), andererseits haben wir aber auch schon in Bemerkung 6.5 gesehen, dass sich bei der komplexen Multiplikation die Winkel, die die Zahlen mit der positiven reellen Achse einschließen, ebenfalls addieren. Wir können den Exponenten x der Zahl e^{ix} daher wie im Bild rechts als ein Maß für diesen Winkel auffassen.



In der Tat werden wir in Aufgabe 9.16 sehen, dass dieses x genau die (im Bild oben rechts dick eingezeichnete) Länge des Kreisbogens ist, der von 1 zu der Zahl e^{ix} führt – man sagt auch, dass x der im **Bogenmaß** gemessene Winkel ist. Wir werden diese Aussage im Folgenden nicht benötigen, sondern verwenden sie hier nur als Motivation dafür, dass Real- und Imaginärteil von e^{ix} (also die beiden Koordinaten dieses Punktes in der Ebene) dann wie aus der Schule bekannt der Kosinus bzw. Sinus des Winkels x sein sollten. Diese Idee machen wir nun zu unserer *Definition* von Kosinus und Sinus.

Definition 9.12 (Kosinus und Sinus)**.** Für $x \in \mathbb{R}$ definieren wir **Kosinus** und **Sinus** als die reellen Zahlen

$$\cos x := \operatorname{Re}(e^{ix}) \qquad \text{und} \qquad \sin x := \operatorname{Im}(e^{ix}),$$

so dass also $e^{ix} = \cos x + i \sin x$ die Zerlegung der komplexen Exponentialfunktion in Real- und Imaginärteil ist.

Bevor wir die Eigenschaften dieser beiden Funktionen studieren, wollen wir erst einmal zwei einfache alternative Darstellungsweisen notieren:

Lemma 9.13 (Alternative Darstellung von Kosinus und Sinus)**.**

(a) *Für alle $x \in \mathbb{R}$ ist*

$$\cos x = \frac{1}{2}\left(e^{ix} + e^{-ix}\right) \qquad \textit{und} \qquad \sin x = -\frac{i}{2}\left(e^{ix} - e^{-ix}\right).$$

(b) *Kosinus und Sinus lassen sich darstellen als reelle Potenzreihen mit Konvergenzradius* ∞

$$\cos x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} \qquad = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \pm \cdots,$$

$$\sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \cdots.$$

Insbesondere sind Kosinus und Sinus nach Folgerung 8.38 also stetige Funktionen auf $\mathbb{R}$.

Beweis.

(a) Dies folgt aus den allgemeinen Formeln $\operatorname{Re} z = \frac{1}{2}(z + \overline{z})$ und $\operatorname{Im} z = -\frac{i}{2}(z - \overline{z})$ (siehe Bemerkung 6.4) zusammen mit Satz 9.10 (a).

(b) Die Potenzreihendarstellungen ergeben sich mit (a) sofort aus

$$e^{ix} = \sum_{n=0}^{\infty} \frac{(ix)^n}{n!} = 1 + i\frac{x^1}{1!} - \frac{x^2}{2!} - i\frac{x^3}{3!} + \frac{x^4}{4!} + i\frac{x^5}{5!} - \frac{x^6}{6!} - i\frac{x^7}{7!} + \cdots$$

$$\text{und} \qquad e^{-ix} = \sum_{n=0}^{\infty} \frac{(-ix)^n}{n!} = 1 - i\frac{x^1}{1!} - \frac{x^2}{2!} + i\frac{x^3}{3!} + \frac{x^4}{4!} - i\frac{x^5}{5!} - \frac{x^6}{6!} + i\frac{x^7}{7!} + \cdots,$$

da man konvergente Reihen nach Lemma 7.4 (a) gliedweise addieren kann. Weil diese Reihendarstellung für alle $x \in \mathbb{R}$ gilt, ist der Konvergenzradius dieser Potenzreihen gleich ∞. □ 21

Wir listen nun zunächst die einfachsten Eigenschaften von Kosinus und Sinus auf, die sich direkt aus der Definition ergeben.

Satz 9.14 (Eigenschaften von Kosinus und Sinus). *Für alle $x, y \in \mathbb{R}$ gilt*

(a) $\cos(-x) = \cos x$ *und* $\sin(-x) = -\sin x$. *(Der Graph von* $\cos$ *ist also achsensymmetrisch zur vertikalen Achse, der von* $\sin$ *punktsymmetrisch zum Ursprung).*

(b) $(\cos x)^2 + (\sin x)^2 = 1$*; insbesondere ist also* $|\cos x| \leq 1$ *und* $|\sin x| \leq 1$.

(c) ***(Additionstheoreme)***

$$\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y$$
$$\text{und} \qquad \sin(x \pm y) = \sin x \cos y \pm \cos x \sin y$$

(wobei die Gleichungen so zu verstehen sind, dass an beiden Stellen das obere oder an beiden Stellen das untere Vorzeichen zu nehmen ist).

Beweis.

(a) Dies folgt z. B. unmittelbar aus Lemma 9.13 (a).

(b) Nach Satz 9.10 (b) ist $(\cos x)^2 + (\sin x)^2 = (\mathrm{Re}(\mathrm{e}^{\mathrm{i}x}))^2 + (\mathrm{Im}(\mathrm{e}^{\mathrm{i}x}))^2 = |\mathrm{e}^{\mathrm{i}x}|^2 = 1$.

(c) Einerseits gilt nach der Funktionalgleichung der komplexen Exponentialfunktion

$$\begin{aligned} \mathrm{e}^{\mathrm{i}(x+y)} &= \mathrm{e}^{\mathrm{i}x} \cdot \mathrm{e}^{\mathrm{i}y} = (\cos x + \mathrm{i}\,\sin x)\,(\cos y + \mathrm{i}\,\sin y) \\ &= \cos x \cos y - \sin x \sin y + \mathrm{i}\left(\sin x \cos y + \cos x \sin y\right), \end{aligned}$$

andererseits nach Definition aber auch

$$\mathrm{e}^{\mathrm{i}(x+y)} = \cos(x+y) + \mathrm{i}\,\sin(x+y).$$

Vergleich von Real- und Imaginärteil liefert nun die behaupteten Formeln für $\cos(x+y)$ und $\sin(x+y)$. Die Formeln für $\cos(x-y)$ und $\sin(x-y)$ ergeben sich daraus durch den Übergang $y \to -y$ mit (a). □

Notation 9.15 ($\cos^n x$ und $\sin^n x$). Für $n \in \mathbb{N}$ schreibt man zur Abkürzung oft auch $\cos^n x$ und $\sin^n x$ anstatt $(\cos x)^n$ und $(\sin x)^n$. Die Formel aus Satz 9.14 (b) schreibt sich dann z. B. kürzer als $\cos^2 x + \sin^2 x = 1$. Beachte aber, dass dies leicht zu Verwechslungen führen kann, weil wir die Umkehrfunktion einer bijektiven Funktion f in Definition 2.20 ja mit $x \mapsto f^{-1}(x)$ bezeichnet haben, dies aber nach dieser neuen Notation auch als $x \mapsto (f(x))^{-1} = \frac{1}{f(x)}$ interpretiert werden könnte – was natürlich etwas völlig anderes ist. Wir werden daher für Kosinus und Sinus die Schreibweise $\cos^{-1}(x)$ bzw. $\sin^{-1}(x)$ überhaupt nicht verwenden, und den Umkehrfunktionen dieser beiden Funktionen andere Namen geben (siehe Definition 9.25).

Aufgabe 9.16 (Bogenmaß).

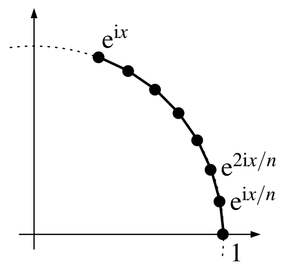
(a) Berechne den Grenzwert $\lim\limits_{x \to 0} \frac{1-\cos x}{x^2}$.

(b) Es sei $x \in \mathbb{R}_{\geq 0}$. Wir wollen zeigen, dass die „Bogenlänge" entlang des Einheitskreises von 1 nach $\mathrm{e}^{\mathrm{i}x} \in \mathbb{C}$ gleich x ist und $\mathrm{e}^{\mathrm{i}x}$ damit als der Punkt auf dem Einheitskreis aufgefasst werden kann, der mit der positiven reellen Achse den Winkel x „im Bogenmaß" einschließt.

Für alle $n \in \mathbb{N}_{>0}$ unterteilen wir dazu den Kreisbogen wie im Bild durch die Zwischenpunkte $\mathrm{e}^{\mathrm{i}kx/n}$ mit $k = 0, \ldots, n$. Die Länge des geraden Streckenzuges, der diese Punkte der Reihe nach miteinander verbindet, ist dann

$$L_n = \sum_{k=0}^{n-1} \left|\mathrm{e}^{\mathrm{i}(k+1)x/n} - \mathrm{e}^{\mathrm{i}kx/n}\right|.$$

Zeige, dass $\lim\limits_{n \to \infty} L_n = x$.



Aufgabe 9.17. Es sei $x \in \mathbb{R}$. Finde und beweise eine explizite Formel für die Summen

$$\sum_{k=0}^{n} \cos(kx) \quad \text{und} \quad \sum_{k=0}^{n} \sin(kx).$$

Als Nächstes wollen wir die Nullstellen und die Periodizität von Kosinus und Sinus untersuchen. Aus Bemerkung 9.11 (und dem, was ihr aus der Schule wisst) ist z. B. klar, dass diese beiden Funktionen die Periode 2π besitzen sollten. Aber bisher wissen wir überhaupt noch nicht, was π eigentlich genau ist! Wie ihr euch vielleicht schon denken könnt, wird auch hier der Ausweg wieder darin bestehen, die Sache rückwärts anzugehen und die Zahl π über die Eigenschaften der Kosinus- und Sinusfunktion zu *definieren*. Dazu benötigen wir das folgende Lemma.

Lemma 9.18.

(a) *Für alle* $x \in (0,2)$ *ist* $\sin x > 0$.

(b) *Die Kosinusfunktion ist im Intervall* $[0,2]$ *streng monoton fallend, und es gilt* $\cos 0 > 0$ *sowie* $\cos 2 < 0$.

Beweis. Wir bemerken zunächst, dass die Summanden der Exponentialreihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ für $0 < x \le 2$ ab dem x^1-Term betragsmäßig monoton fallend sind, denn für $n \ge 1$ ist

$$\left| \frac{x^{n+1}/(n+1)!}{x^n/n!} \right| = \frac{x}{n+1} \le \frac{2}{1+1} = 1.$$

Dasselbe gilt dann natürlich auch für die Glieder der Kosinus- und Sinusreihe, die nach Lemma 9.13 (b) ja bis auf das Vorzeichen genau die geraden bzw. ungeraden Terme der Exponentialreihe sind. Da die Kosinus- und Sinusreihe zudem alternierend sind, sind ihre Partialsummen nach Satz 7.8 damit abwechselnd obere und untere Schranken für den Grenzwert (sofern wir mindestens bis zum x^1-Term aufsummieren, ab dem die Summanden betragsmäßig monoton fallen).

(a) Nach unserer Vorbemerkung folgt nun sofort für alle $x \in (0,2)$

$$\sin x \ge x - \frac{x^3}{3!} = x\left(1 - \frac{x^2}{6}\right) > x\left(1 - \frac{2^2}{6}\right) = \frac{x}{3} > 0.$$

(b) Natürlich ist $\cos 0 = 1 > 0$. Für $\cos 2$ gilt wieder nach der Vorbemerkung

$$\cos 2 \le 1 - \frac{2^2}{2!} + \frac{2^4}{4!} = 1 - \frac{4}{2} + \frac{16}{24} = -\frac{1}{3} < 0.$$

Es bleibt also nur noch die strenge Monotonie zu zeigen. Es seien dazu $x, y \in [0,2]$ mit $x < y$ gegeben. Mit den Additionstheoremen aus Satz 9.14 (c) ergibt sich

$$\cos\left(\frac{y+x}{2} \pm \frac{y-x}{2}\right) = \cos\frac{y+x}{2} \cdot \cos\frac{y-x}{2} \mp \sin\frac{y+x}{2} \cdot \sin\frac{y-x}{2}.$$

Subtraktion dieser beiden Gleichungen voneinander liefert nun

$$\cos\Bigg(\underbrace{\frac{y+x}{2} + \frac{y-x}{2}}_{=y}\Bigg) - \cos\Bigg(\underbrace{\frac{y+x}{2} - \frac{y-x}{2}}_{=x}\Bigg) = -2\sin\underbrace{\frac{y+x}{2}}_{\in(0,2)} \cdot \sin\underbrace{\frac{y-x}{2}}_{\in(0,2)},$$

mit (a) also $\cos y - \cos x < 0$ und damit wie behauptet $\cos x > \cos y$. □

Da die Kosinusfunktion nach Lemma 9.13 (b) stetig ist, ergibt sich mit dem Zwischenwertsatz 8.20 aus Lemma 9.18 (b) also, dass es *genau ein* $x \in (0,2)$ gibt mit $\cos x = 0$. Aus der Interpretation von Bemerkung 9.11 sehen wir, dass diese Nullstelle des Kosinus genau bei $x = \frac{\pi}{2}$ auftreten sollte, also dort wo e^{ix} auf der imaginären Achse liegt. Wir benutzen dies nun, um die Zahl π zu *definieren*:

Definition 9.19 (Die Zahl π)**.** Wir definieren die Zahl $\pi \in \mathbb{R}$ als das Doppelte der (nach obigen Überlegungen eindeutig bestimmten) Nullstelle der Kosinusfunktion im Intervall $[0,2]$. (Eine näherungsweise Berechnung dieser Nullstelle zeigt, dass $\pi = 3,14159\ldots$)

Bemerkung 9.20.

(a) Unsere Definition 9.19 ist natürlich nicht die einzig mögliche Art, wie man die Zahl π definieren kann. Man könnte genauso gut auch irgendeine andere charakteristische Eigenschaft dieser Zahl als Definition verwenden, wie z. B. (was ja häufig gesagt wird) den Flächeninhalt des Einheitskreises. Allerdings wissen wir bisher noch gar nicht, wie man derartige Flächeninhalte überhaupt definieren und berechnen kann. Für uns hat die etwas merkwürdig scheinende Definition 9.19 daher einfach den Vorteil, dass sie am schnellsten zu den gewünschten Resultaten führt (insbesondere auch ohne Flächeninhalte berechnen zu können).

(b) Da die Kosinusfunktion nach Lemma 9.18 (b) auf $[0,2]$ streng monoton fallend ist, ist ihre Einschränkung auf das Intervall $[0,\frac{\pi}{2}]$ damit eine bijektive, streng monoton fallende Funktion, die von $\cos 0 = 1$ nach $\cos\frac{\pi}{2} = 0$ läuft. Wir wollen nun sehen, wie wir aus diesem Abschnitt der Kosinusfunktion die Kosinus- und Sinusfunktion auf ganz $\mathbb{R}$ rekonstruieren können.

Satz 9.21 (Periodizität von Kosinus und Sinus)**.**

(a) *An den Stellen* $x \in \{0,\frac{\pi}{2},\pi,\frac{3\pi}{2},2\pi\}$ *nehmen* $\mathrm{e}^{\mathrm{i}x}$, $\cos x$ *und* $\sin x$ *die folgenden Werte an:*

x	0	$\frac{\pi}{2}$	π	$\frac{3\pi}{2}$	2π
$\mathrm{e}^{\mathrm{i}x}$	1	i	-1	$-\mathrm{i}$	1
$\cos x$	1	0	-1	0	1
$\sin x$	0	1	0	-1	0

(b) *Kosinus und Sinus sind* 2π*-periodisch, d. h. es gilt* $\cos(x+2\pi) = \cos x$ *und* $\sin(x+2\pi) = \sin x$ *für alle* $x \in \mathbb{R}$.

(c) *Für alle* $x \in \mathbb{R}$ *ist* $\cos(\pi - x) = -\cos x$ *und* $\sin(\frac{\pi}{2} \pm x) = \cos x$.

Beweis.

(a) Nach Definition 9.19 ist $\cos\frac{\pi}{2} = 0$. Für den Sinus gilt damit $\sin^2\frac{\pi}{2} = 1 - \cos^2\frac{\pi}{2} = 1 - 0 = 1$ nach Satz 9.14 (b); da außerdem $\sin\frac{\pi}{2} > 0$ nach Lemma 9.18 (a) gilt, muss also $\sin\frac{\pi}{2} = 1$ sein. Damit ist $\mathrm{e}^{\mathrm{i}\frac{\pi}{2}} = \cos\frac{\pi}{2} + \mathrm{i}\sin\frac{\pi}{2} = \mathrm{i}$.

Die übrigen behaupteten Werte für $\mathrm{e}^{\mathrm{i}n\frac{\pi}{2}}$ mit $n \in \{0,1,2,3,4\}$ folgen damit sofort nach den Rechenregeln für Potenzen aus $\mathrm{e}^{\mathrm{i}n\frac{\pi}{2}} = (\mathrm{e}^{\mathrm{i}\frac{\pi}{2}})^n = \mathrm{i}^n$. Aufteilen dieser Zahlen in Real- und Imaginärteil liefert dann die Werte für Kosinus und Sinus in der Tabelle.

(b) Nach (a) ist $\mathrm{e}^{\mathrm{i}\cdot 2\pi} = 1$ und damit $\mathrm{e}^{\mathrm{i}\cdot(x+2\pi)} = \mathrm{e}^{\mathrm{i}x}\cdot\mathrm{e}^{\mathrm{i}\cdot 2\pi} = \mathrm{e}^{\mathrm{i}x}$. Aufteilen in Real- und Imaginärteil liefert wieder die Behauptung.

(c) Wiederum mit (a) ist

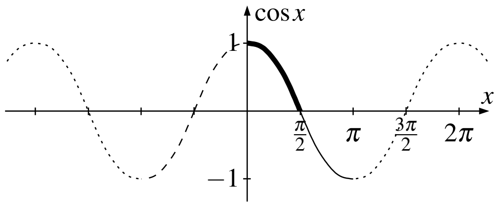
$$\cos(\pi - x) = \cos\pi\,\cos x + \sin\pi\,\sin x = (-1)\cdot\cos x + 0\cdot\sin x = -\cos x$$

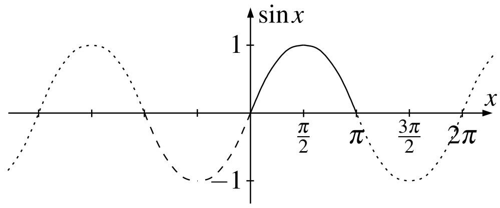
$$\text{und}\quad \sin\left(\frac{\pi}{2} \pm x\right) = \sin\frac{\pi}{2}\,\cos x \pm \cos\frac{\pi}{2}\,\sin x = 1\cdot\cos x \pm 0\cdot\sin x = \cos x$$

nach den Additionstheoremen aus Satz 9.14 (c). □

Bemerkung 9.22. Mit Satz 9.21 können wir nun aus dem Verlauf der Kosinusfunktion im Intervall $[0,\frac{\pi}{2}]$ (nach Bemerkung 9.20 (b) streng monoton fallend von 1 nach 0 verlaufend; im Bild unten dick eingezeichnet) die gesamte Kosinus- und Sinusfunktion rekonstruieren:

- Satz 9.21 (c) für $x \in [0,\frac{\pi}{2}]$ legt zunächst Kosinus und Sinus im Intervall $[0,\pi]$ fest: Die Graphen verlaufen hier „genauso" wie beim Kosinus von 0 bis $\frac{\pi}{2}$, nur gedreht bzw. gespiegelt (im Bild unten als durchgezogene Linie markiert);
- Satz 9.14 (a) bestimmt Kosinus und Sinus damit dann auch im Intervall $[-\pi,\pi]$ (im Bild gestrichelt eingezeichnet);
- Satz 9.21 (b) schließlich besagt dann, dass dieser Verlauf von Kosinus und Sinus in beide Richtungen 2π-periodisch fortgesetzt wird (wie im Bild gepunktet eingezeichnet).





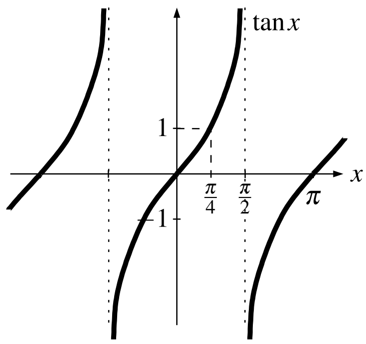
Wie in der Schule definiert man schließlich noch den Tangens:

Definition 9.23 (Tangens)**.** Für alle $x \in \mathbb{R}\backslash\{\frac{\pi}{2} + n\pi : n \in \mathbb{Z}\}$, also nach Bemerkung 9.22 für alle x mit $\cos x \neq 0$, heißt

$$\tan x := \frac{\sin x}{\cos x}$$

der **Tangens** von x.

Bemerkung 9.24 (Eigenschaften des Tangens)**.** Die wesentlichen Eigenschaften des Tangens ergeben sich unmittelbar aus denen des Kosinus und Sinus:



(a) Der Tangens ist auf seiner Definitionsmenge als Quotient stetiger Funktionen stetig.

(b) Es ist

$$\tan(-x) = \frac{\sin(-x)}{\cos(-x)} = \frac{-\sin x}{\cos x} = -\tan x$$

nach Satz 9.14 (a) (d. h. der Graph von tan ist punktsymmetrisch zum Ursprung), und

$$\tan(x+\pi) = \frac{\sin(x+\pi)}{\cos(x+\pi)} = \frac{-\sin x}{-\cos x} = \tan x$$

nach Bemerkung 9.22 (d. h. tan ist periodisch mit Periode π). Der Graph der Tangensfunktion ist wegen dieser Symmetrien also bereits durch den Graphen im Intervall $[0, \frac{\pi}{2})$ bestimmt.

(c) Es ist $\tan 0 = \frac{\sin 0}{\cos 0} = \frac{0}{1} = 0$ sowie $\tan\frac{\pi}{4} = 1$, denn nach Satz 9.21 (c) ist $\sin\frac{\pi}{4} = \cos\frac{\pi}{4}$. Weiterhin ist

$$\lim_{\substack{x\to\pi/2\\ x<\pi/2}} \tan x = \infty,$$

da in diesem Grenzfall $\sin x \to 1$ gilt und $\cos x$ von der positiven Seite gegen 0 konvergiert.

(d) Die Tangensfunktion ist auf $[0, \frac{\pi}{2})$ (und damit nach der Symmetrie aus (b) auch auf $(-\frac{\pi}{2}, \frac{\pi}{2})$) streng monoton wachsend: Für $x, y \in [0, \frac{\pi}{2})$ mit $x < y$ ist $\sin x < \sin y$ und $\cos x > \cos y > 0$ nach Bemerkung 9.22, also

$$\tan x = \frac{\sin x}{\cos x} < \frac{\sin y}{\cos y} = \tan y.$$

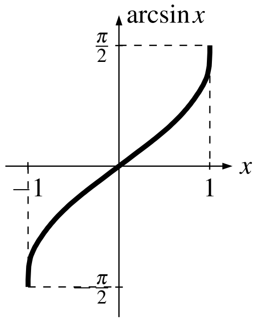
Wir können nun natürlich auf den Intervallen, auf denen die Winkelfunktionen stetig und streng monoton sind, nach Satz 8.27 die Umkehrfunktionen definieren. Aufgrund der Periodizität haben wir dabei in allen drei Fällen eine Wahl, *welches* solche Intervall wir betrachten. Üblicherweise verwendet man die folgenden:

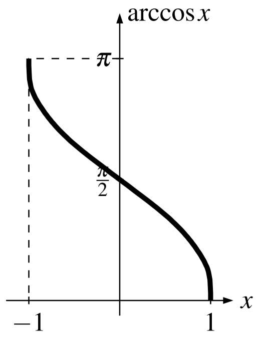
Definition 9.25 (Umkehrfunktionen der Winkelfunktionen)**.** Die Umkehrfunktion von ...

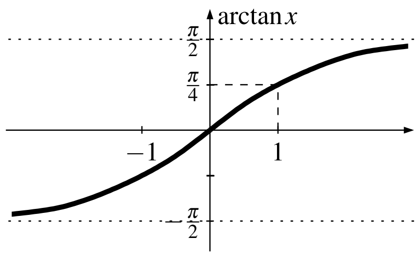
(a) $\sin\colon [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1]$ heißt **Arkussinus** $\arcsin\colon [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$.

(b) $\cos\colon [0, \pi] \to [-1, 1]$ heißt **Arkuskosinus** $\arccos\colon [-1, 1] \to [0, \pi]$.

(c) $\tan\colon \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \to \mathbb{R}$ heißt **Arkustangens** $\arctan\colon \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.





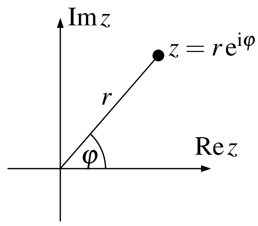


Bemerkung 9.26.

(a) Beachte, dass die Arkusfunktionen *nur in den betrachteten Intervallen* die Umkehrfunktionen der Winkelfunktionen sind; es ist also z. B. $\arcsin(\sin x) = x$ nur für $x \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, wohingegen z. B. $\arcsin(\sin \pi) = \arcsin 0 = 0$ ist.

(b) Wie bei den anderen bisher betrachteten Umkehrfunktionen ergeben sich die Eigenschaften der Arkusfunktionen natürlich wieder direkt aus denen der Winkelfunktionen. So sind z. B. alle Arkusfunktionen stetig und streng monoton nach Satz 8.27 (wachsend für arcsin und arctan, fallend für arccos), und die wichtigsten speziellen Werte und Symmetrien sind wie im Bild oben dargestellt.

Als Anwendung der Arkusfunktionen wollen wir zum Abschluss dieses Kapitels nun noch die sogenannte Polarkoordinatendarstellung komplexer Zahlen behandeln, mit der sich Rechnungen in $\mathbb{C}$ oft wesentlich vereinfachen lassen. Die Idee dabei ist einfach, dass man durch Multiplikation einer Zahl $e^{i\varphi}$ auf dem Einheitskreis (siehe Bemerkung 9.11) mit einer positiven reellen Zahl r jeden Punkt der komplexen Zahlenebene (mit Ausnahme des Nullpunkts) erreichen können sollte, wobei wie im Bild rechts r gerade der Betrag und φ der Winkel des betrachteten Punktes ist. Man kann einen solchen Punkt also auch durch Angabe der Werte von r und φ (anstatt durch Real- und Imaginärteil) charakterisieren:



Satz 9.27 (Polarkoordinatendarstellung)**.**

(a) *Jede komplexe Zahl $z \in \mathbb{C}\backslash\{0\}$ lässt sich in der Form $z = re^{i\varphi}$ mit $r \in \mathbb{R}_{>0}$ und $\varphi \in \mathbb{R}$ schreiben.*

(b) *Die Darstellung aus (a) ist eindeutig bis auf ganzzahlige Vielfache von 2π in φ, d. h. für $r, r' \in \mathbb{R}_{>0}$ und $\varphi, \varphi' \in \mathbb{R}$ gilt $re^{i\varphi} = r'e^{i\varphi'}$ genau dann wenn $r' = r$ und $\varphi' - \varphi = 2\pi n$ für ein $n \in \mathbb{Z}$.*

*Man nennt r und φ die **Polarkoordinaten** von z.*

Beweis.

(a) Wir setzen $r := |z|$. Es bleibt also noch zu zeigen, dass es ein $\varphi \in \mathbb{R}$ gibt mit $z = |z|e^{i\varphi}$, d. h. dass sich die komplexe Zahl $\frac{z}{|z|}$, die ja jetzt Betrag 1 hat, in der Form $e^{i\varphi}$ schreiben lässt. Aufgeteilt in Real- und Imaginärteil bedeutet dies genau, dass wir $\frac{z}{|z|} =: x + iy$ als $\cos\varphi + i\sin\varphi$ schreiben wollen, d. h. wir suchen zu $x, y \in \mathbb{R}$ mit $x^2 + y^2 = 1$ eine Zahl $\varphi \in \mathbb{R}$ mit

$$\cos\varphi = x \qquad \text{und} \qquad \sin\varphi = y. \qquad (*)$$

Wegen $x^2 + y^2 = 1$ ist insbesondere $|x| \leq 1$, wir können also $\alpha := \arccos x$ setzen, so dass schon einmal $\cos\alpha = x$ gilt. Nun ist mit Satz 9.14 (b)

$$y^2 = 1 - x^2 = 1 - \cos^2\alpha = \sin^2\alpha$$

und damit $y = \pm \sin \alpha$. Im Fall des Vorzeichens „+“ ergibt nun $\varphi := \alpha$, im Fall des Vorzeichens „−“ hingegen $\varphi := -\alpha$ die gewünschten Relationen (∗).

(b) Zunächst ist $r\,\mathrm{e}^{\mathrm{i}\varphi} = r'\,\mathrm{e}^{\mathrm{i}\varphi'}$ genau dann, wenn $r = r'$ und $\mathrm{e}^{\mathrm{i}\varphi} = \mathrm{e}^{\mathrm{i}\varphi'}$ (für die Richtung „⇐“ nehmen wir auf beiden Seiten den Betrag). Weiterhin ist $\mathrm{e}^{\mathrm{i}\varphi} = \mathrm{e}^{\mathrm{i}\varphi'}$, also $\mathrm{e}^{\mathrm{i}(\varphi'-\varphi)} = 1$, äquivalent zu

$$\cos(\varphi' - \varphi) = 1 \qquad \text{und} \qquad \sin(\varphi' - \varphi) = 0.$$

Aus Bemerkung 9.22 ergibt sich nun, dass dies genau dann der Fall ist, wenn $\varphi' - \varphi$ ein ganzzahliges Vielfaches von 2π ist. □

Beispiel 9.28.

(a) In Polarkoordinaten ist insbesondere die Multiplikation zweier komplexer Zahlen sehr einfach: Es ist

$$(r_1\,\mathrm{e}^{\mathrm{i}\varphi_1}) \cdot (r_2\,\mathrm{e}^{\mathrm{i}\varphi_2}) = (r_1 r_2)\,\mathrm{e}^{\mathrm{i}(\varphi_1+\varphi_2)},$$

was die algebraische Version der geometrischen Aussage aus Bemerkung 6.5 ist, dass sich bei der Multiplikation komplexer Zahlen die Beträge multiplizieren und die Winkel addieren.

(b) (Einheitswurzeln) In Polarkoordinaten können wir in $\mathbb{C}$ besonders einfach die Gleichung $z^n = 1$ für $n \in \mathbb{N}_{>0}$ lösen. Schreiben wir nämlich $z = r\,\mathrm{e}^{\mathrm{i}\varphi}$, so wollen wir also

$$z^n = r^n\,\mathrm{e}^{\mathrm{i}n\varphi} = 1 = 1\,\mathrm{e}^{\mathrm{i}\cdot 0},$$

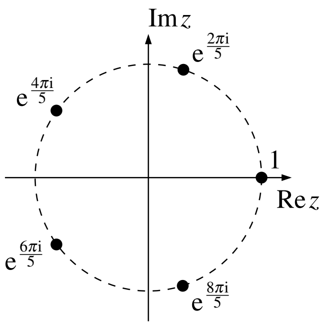
was nach der Eindeutigkeitsaussage aus Satz 9.27 (b) bedeutet, dass

$$r^n = 1 \qquad \text{und} \qquad n\varphi = 2\pi k \text{ für ein } k \in \mathbb{Z},$$

also $r = 1$ und $\varphi = \frac{2\pi k}{n}$ für ein $k \in \mathbb{Z}$. Da man die Polarkoordinaten stets so wählen kann, dass $\varphi \in [0, 2\pi)$ ist, genügt es dabei, sich auf die Werte $k \in \{0, 1, \dots, n-1\}$ zu beschränken. Die komplexen Lösungen der Gleichung $z^n = 1$ sind also genau die n Zahlen

$$z_k = \mathrm{e}^{\frac{2\pi \mathrm{i} k}{n}} \qquad \text{für } k = 0, \dots, n-1.$$

Man bezeichnet diese Zahlen als die n-ten **Einheitswurzeln**; sie bilden die Ecken eines regelmäßigen n-Ecks (wie im Bild rechts für das Beispiel $n = 5$ dargestellt).



Aufgabe 9.29.

(a) Bestimme $\cos \frac{\pi}{6}$ und $\sin \frac{\pi}{6}$.

(b) Stelle alle komplexen Lösungen der Gleichung $z^6 = -8$ sowohl in Polarkoordinaten $z = r\mathrm{e}^{\mathrm{i}\varphi}$ als auch ohne Verwendung von Winkelfunktionen in der Form $z = x + \mathrm{i}y$ dar. Wo liegen sie in der komplexen Zahlenebene?

Aufgabe 9.30.

(a) Zeige, dass

$$\tan(x+y) = \frac{\tan x + \tan y}{1 - \tan x \cdot \tan y}$$

für alle $x, y \in \mathbb{R}$, für die diese Ausdrücke definiert sind.

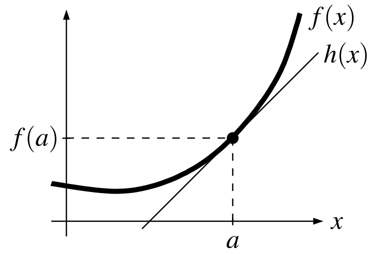
(b) Skizziere die Funktion

$$f \colon \mathbb{R}\backslash\{-1\} \to \mathbb{R},\ x \mapsto \arctan x + \arctan\left(\frac{1-x}{1+x}\right)$$

und begründe dabei alle wesentlichen qualitativen Merkmale des Funktionsgraphen.

10. Differentialrechnung

Wir kommen nun zum wohl wichtigsten Teil der Analysis (in einer Veränderlichen), der sogenannten Differentialrechnung. Ziel der Differentialrechnung ist es, wie im Bild rechts eine Funktion $f: D \to \mathbb{K}$ in einem gegebenen Punkt $a \in D \subset \mathbb{K}$ *linear zu approximieren*, d. h. eine Gerade h zu finden, die f in einer kleinen Umgebung von a möglichst gut annähert. Mit anderen Worten können wir h als *Tangente* an den Graphen von f im Punkt a auffassen.



In der Praxis ist dies natürlich oft wünschenswert, denn immer wenn wir aus irgendwelchen Gründen wissen, dass wir die Funktion f nur in der Nähe von a benötigen werden, dann können wir die womöglich sehr komplizierte Funktion f näherungsweise durch eine Gerade ersetzen, also durch eine viel einfacher zu behandelnde Funktion.

Wie kann man nun diese Tangente h bestimmen? Als Gerade durch den Punkt $(a, f(a))$ muss sie natürlich von der Form $h(x) = f(a) + c\,(x-a)$ für ein $c \in \mathbb{K}$ sein, wobei dieses c die Steigung der Geraden angibt. Wir möchten also erreichen, dass

$$f(x) \approx f(a) + c\,(x-a),$$

wobei das Symbol „$\approx$" hier nicht exakt definiert ist, sondern nur den anschaulichen Sachverhalt „ist für x in der Nähe von a in etwa gleich" beschreiben soll. Es müsste dann also

$$c \approx \frac{f(x)-f(a)}{x-a}$$

sein. Aufgrund unserer Vorarbeiten wissen wir aber natürlich nun, wie man dies mathematisch exakt formulieren muss: Die beste Näherung erhalten wir für den Grenzwert

$$c = \lim_{x\to a} \frac{f(x)-f(a)}{x-a}$$

(sofern er existiert). Derartige Grenzwerte wollen wir nun also in der Differentialrechnung studieren. 22

10.A Ableitungen von Funktionen

Bevor wir den obigen Grenzwert von $\frac{f(x)-f(a)}{x-a}$ für $x \to a$ exakt definieren können, müssen wir noch kurz eine (recht schwache) Bedingung an die Definitionsmenge D der betrachteten Funktion stellen: Da dieser Quotient nur für $x \in D\backslash\{a\}$ definiert ist, muss a nach Definition 8.3 ein Berührpunkt von $D\backslash\{a\}$ sein, damit der Grenzwert dieses Ausdrucks für $x \to a$ überhaupt definierbar ist, also damit man sich innerhalb von $D\backslash\{a\}$ dem Punkt a beliebig nähern kann. Wir wollen diese Bedingung nun formalisieren.

Definition 10.1 (Isolierte Punkte)**.** Es sei $D \subset \mathbb{K}$. Ein Punkt $a \in D$ heißt **isolierter Punkt** von D, wenn es eine ε-Umgebung von a gibt, die außer a keinen Punkt von D enthält (also „wenn man sich innerhalb von $D\backslash\{a\}$ dem Punkt a nicht beliebig nähern kann").

Beispiel 10.2.

(a) Die Menge $\mathbb{Z}$ besteht nur aus isolierten Punkten.

(b) Intervalle in $\mathbb{R}$ – egal ob offene, halboffene, abgeschlossene oder uneigentliche – haben keine isolierten Punkte (solange sie nicht ein einpunktiges Intervall $[a,a]$ sind). Vereinigungen derartiger Intervalle haben ebenfalls keine isolierten Punkte.

Durch Negation der Bedingung aus Definition 10.1 sehen wir, dass ein Punkt $a \in D$ *kein* isolierter Punkt von D ist, wenn in jeder ε-Umgebung von a ein Punkt von $D \backslash \{a\}$ liegt, also wenn a ein Berührpunkt von $D \backslash \{a\}$ ist. Um Grenzwerte für $x \to a$ mit $x \in D$ und $x \neq a$ in jedem Punkt $a \in D$ bilden zu können, machen wir wie oben erläutert jetzt also die

Grundvoraussetzung für dieses Kapitel: Die Definitionsmengen aller betrachteten Funktionen haben keine isolierten Punkte.

Damit können wir nun wie oben motiviert die Steigungen der lokalen linearen Approximationen einer gegebenen Funktion als Grenzwerte berechnen:

Definition 10.3 (Differenzierbarkeit)**.** Es seien $D \subset \mathbb{K}$ und $f\colon D \to \mathbb{K}$ eine Funktion.

(a) Die Funktion f heißt **differenzierbar** in $a \in D$, wenn der Grenzwert

$$f'(a) := \lim_{\substack{x \to a \\ x \neq a}} \frac{f(x) - f(a)}{x - a}$$

in $\mathbb{K}$ existiert (ein uneigentlicher Grenzwert $\pm\infty$ im reellen Fall wie in Definition 8.18 ist hier also nicht zugelassen). Die Zahl $\frac{f(x)-f(a)}{x-a}$ bezeichnet man oft als **Differenzenquotient**, ihren Grenzwert $f'(a)$ – sofern er existiert – als **Differentialquotient** oder **Ableitung** von f in a. Da es offensichtlich ist, dass wir hier bei der Grenzwertbildung $x \neq a$ beachten müssen, werden wir diese Bedingung in der Regel nicht jedes Mal wieder explizit hinschreiben.

(b) Man nennt f differenzierbar (auf D), wenn f in jedem Punkt von D differenzierbar ist. Offensichtlich erhält man dann eine Funktion $f'\colon D \to \mathbb{K}$, $x \mapsto f'(x)$, die die Ableitungsfunktion (oder ebenfalls kurz Ableitung) von f genannt wird.

Wie oben erläutert ist $f'(a)$ also (zumindest im reellen Fall) die Steigung der Tangenten an den Graphen von f im Punkt $(a, f(a))$.

Beispiel 10.4.

(a) Jede Gerade $f\colon \mathbb{K} \to \mathbb{K}$, $x \mapsto mx + b$ mit $m, b \in \mathbb{K}$ ist differenzierbar mit Ableitung m, denn

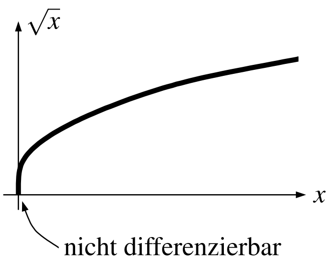
$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} \frac{(mx+b) - (ma+b)}{x - a} = m.$$

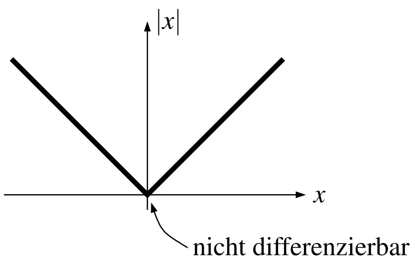
Dies ist geometrisch natürlich klar, denn eine solche Gerade ist ihre eigene Tangente in jedem Punkt und hat damit überall die Steigung m.

(b) Es sei $f\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$, $x \mapsto \sqrt{x}$ die Wurzelfunktion. Dann ist für alle $a \in \mathbb{R}_{\geq 0}$

$$\lim_{x \to a} \frac{\sqrt{x} - \sqrt{a}}{x - a} = \lim_{x \to a} \frac{\sqrt{x} - \sqrt{a}}{(\sqrt{x} + \sqrt{a})(\sqrt{x} - \sqrt{a})} = \lim_{x \to a} \frac{1}{\sqrt{x} + \sqrt{a}} = \begin{cases} \frac{1}{2\sqrt{a}} & \text{für } a > 0, \\ \infty & \text{für } a = 0. \end{cases}$$

Also ist f differenzierbar auf $\mathbb{R}_{>0}$ mit Ableitung $f'(x) = \frac{1}{2\sqrt{x}}$, aber nicht differenzierbar auf $\mathbb{R}_{\geq 0}$. Anschaulich ist f nicht differenzierbar in 0, weil f dort „unendliche Steigung“ hat und die lineare Approximation daher eine senkrechte Gerade sein müsste (siehe Bild unten links).





(c) Die Betragsfunktion $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto |x|$ ist nicht differenzierbar in 0 nach dem Folgenkriterium, denn die Folge $(x_n)_n$ mit $x_n = \frac{(-1)^n}{n}$ konvergiert gegen 0, aber der Grenzwert
$$\lim_{n\to\infty} \frac{f(x_n)-f(0)}{x_n - 0} = \lim_{n\to\infty} \frac{|x_n|}{x_n} = \lim_{n\to\infty} \frac{1/n}{(-1)^n/n} = \lim_{n\to\infty} (-1)^n$$
existiert nicht. Anschaulich ist f deswegen nicht differenzierbar in 0, weil der Funktionsgraph dort einen „Knick" hat und sich die Funktion daher dort nicht durch eine Gerade approximieren lässt (siehe Bild oben rechts).

Wir haben gerade mit Beispiel 10.4 (b) und (c) zwei Beispiele von Funktionen gesehen, die (in einem Punkt) stetig, aber nicht differenzierbar sind. Wir wollen nun zeigen, dass umgekehrt aber jede differenzierbare Funktion stetig ist. Hierfür benötigen wir das folgende Lemma, das wir auch später noch einmal verwenden werden.

Lemma 10.5 (Äquivalentes Kriterium für Differenzierbarkeit)**.** *Es seien $D \subset \mathbb{K}$, $f\colon D \to \mathbb{K}$ eine Funktion und $a \in D$. Dann sind die folgenden beiden Bedingungen äquivalent:*

(a) *f ist differenzierbar in a.*

(b) *Es gibt eine in a stetige Funktion $\varphi\colon D \to \mathbb{K}$ mit $f(x) - f(a) = \varphi(x)\cdot(x-a)$ für alle $x \in D$.*

In diesem Fall ist dann $\varphi(a) = f'(a)$.

Beweis. Die gegebene Bedingung an φ legt diese Funktion für alle $x \neq a$ offensichtlich fest als den Differenzenquotienten
$$\varphi(x) = \frac{f(x)-f(a)}{x-a}.$$
Damit besagt (b) also genau, dass diese Funktion stetig nach a fortsetzbar ist. Dies ist gemäß Definition 8.5 (b) exakt dasselbe wie die Aussage, dass der Grenzwert
$$\lim_{\substack{x\to a\\ x\neq a}} \varphi(x) = \lim_{\substack{x\to a\\ x\neq a}} \frac{f(x)-f(a)}{x-a} \tag{$*$}$$
existiert, also dass f in a differenzierbar ist. Ist dies der Fall, so ist der Ausdruck $(*)$ dann aber sowohl gleich der stetigen Fortsetzung $\varphi(a)$ von φ in a als auch gleich der Ableitung $f'(a)$. □

Bemerkung 10.6. Anschaulich gibt die Funktion φ aus Lemma 10.5 im Punkt x genau die Steigung der Geraden durch die Punkte $(a, f(a))$ und $(x, f(x))$ an – die im Grenzfall $x \to a$ dann zur Tangentensteigung wird.

Folgerung 10.7. *Es seien $D \subset \mathbb{K}$, $f\colon D \to \mathbb{K}$ eine Funktion und $a \in D$. Ist f differenzierbar in a, so ist f auch stetig in a.*

Beweis. Ist f differenzierbar in a, so gibt es nach Lemma 10.5 eine in a stetige Funktion $\varphi\colon D \to \mathbb{K}$ mit $f(x) = f(a) + \varphi(x)\,(x-a)$. Insbesondere existiert also der Grenzwert $\lim\limits_{x\to a} \varphi(x) = \varphi(a)$, und damit nach den Grenzwertsätzen aus Satz 8.13 auch
$$\lim_{x\to a} f(x) = \lim_{x\to a} \big(f(a) + \varphi(x)\,(x-a)\big) = f(a) + \varphi(a)\,(a-a) = f(a),$$
d. h. f ist stetig in a. □

Genau wie bei unserer Untersuchung der Stetigkeit wollen wir nun zeigen, dass sich die Differenzierbarkeit von Funktionen auf Summen, Differenzen, Produkte, Quotienten, Verkettungen, Umkehrfunktionen und schließlich auch auf Potenzreihen überträgt – und auch wie man dann die Ableitungen dieser neuen Funktionen berechnet. Wir beginnen mit den vier Grundrechenarten.

Satz 10.8 (Rechenregeln für Ableitungen)**.** *Die Funktionen $f, g\colon D \to \mathbb{K}$ seien differenzierbar in $a \in D$. Dann gilt:*

(a) *$f \pm g$ ist differenzierbar in a mit Ableitung $(f \pm g)'(a) = (f' \pm g')(a)$.*

(b) ***(Produktregel)** fg ist differenzierbar in a mit $(fg)'(a) = (f'g + fg')(a)$.*

(c) ***(Quotientenregel)*** *Ist* $g(a) \neq 0$, *so ist* $\frac{f}{g}$ *differenzierbar in* a *mit* $\left(\frac{f}{g}\right)'(a) = \left(\frac{f'g-fg'}{g^2}\right)(a)$.

Beweis.

(a) Wir führen den Beweis hier nur für die Addition, der für die Subtraktion ist analog:

$$(f+g)'(a) = \lim_{x\to a} \frac{f(x)+g(x)-(f(a)+g(a))}{x-a} = \lim_{x\to a} \Big(\underbrace{\frac{f(x)-f(a)}{x-a}}_{\to f'(a)} + \underbrace{\frac{g(x)-g(a)}{x-a}}_{\to g'(a)} \Big)$$
$$= (f'+g')(a).$$

(b) Da g in a differenzierbar, nach Folgerung 10.7 also auch stetig ist, gilt $\lim\limits_{x\to a} g(x) = g(a)$. Damit ergibt sich nach den Rechenregeln für Grenzwerte aus Satz 8.13

$$\begin{aligned}(fg)'(a) &= \lim_{x\to a} \frac{f(x)g(x)-f(a)g(a)}{x-a} \\ &= \lim_{x\to a} \frac{f(x)g(x)-f(a)g(x)+f(a)g(x)-f(a)g(a)}{x-a} \\ &= \lim_{x\to a} \Big(\underbrace{\frac{f(x)-f(a)}{x-a}}_{\to f'(a)} \cdot \underbrace{g(x)}_{\to g(a)} + f(a) \cdot \underbrace{\frac{g(x)-g(a)}{x-a}}_{\to g'(a)} \Big) \\ &= (f'g+fg')(a).\end{aligned}$$

(c) Da g als differenzierbare Funktion nach Folgerung 10.7 auch stetig ist, ist g wegen $g(a) \neq 0$ nach Bemerkung 8.8 in einer ε-Umgebung von a nirgends 0. Damit stimmt die Definitionsmenge von $\frac{f}{g}$ dort mit D überein. Also ist a kein isolierter Punkt dieser Definitionsmenge, und wir können sinnvoll über die Ableitung von $\frac{f}{g}$ in a sprechen.

Die eigentliche Berechnung dieser Ableitung ist nun analog zu (b):

$$\begin{aligned}\left(\frac{f}{g}\right)'(a) &= \lim_{x\to a} \frac{\frac{f(x)}{g(x)} - \frac{f(a)}{g(a)}}{x-a} \\ &= \lim_{x\to a} \frac{1}{g(x)g(a)} \cdot \frac{f(x)g(a)-f(a)g(a)+f(a)g(a)-f(a)g(x)}{x-a} \\ &= \lim_{x\to a} \underbrace{\frac{1}{g(x)g(a)}}_{\to 1/(g(a))^2} \cdot \Big(\underbrace{\frac{f(x)-f(a)}{x-a}}_{\to f'(a)} \cdot g(a) - f(a) \cdot \underbrace{\frac{g(x)-g(a)}{x-a}}_{\to g'(a)} \Big) \\ &= \left(\frac{f'g-fg'}{g^2}\right)(a).\end{aligned}$$

□

Beispiel 10.9.

(a) Wir zeigen mit (auf- und absteigender) Induktion über n, dass die Ableitung der Potenzfunktion $f(x) = x^n$ gleich $f'(x) = nx^{n-1}$ für alle $n \in \mathbb{Z}$ ist. Der Induktionsanfang für $n = 0$ ergibt sich aus Beispiel 10.4 (a). Wissen wir nun für ein festes $n \in \mathbb{Z}$, dass die Ableitung von $x \mapsto x^n$ gleich $x \mapsto nx^{n-1}$ ist, so folgt mit der Produktregel für die Ableitung von $f(x) = x^{n+1} = x^n \cdot x$

$$f'(x) = nx^{n-1} \cdot x + x^n \cdot 1 = (n+1)x^n$$

(da die Ableitung der Funktion $x \mapsto x$ nach Beispiel 10.4 (a) die konstante Funktion 1 ist), und mit der Quotientenregel für die Ableitung von $f(x) = x^{n-1} = \frac{x^n}{x}$ analog

$$f'(x) = \frac{nx^{n-1} \cdot x - x^n \cdot 1}{x^2} = (n-1)x^{n-2}.$$

(b) Aus der Produktregel und Beispiel 10.4 (a) folgt insbesondere für alle $c \in \mathbb{K}$ und jede differenzierbare Funktion $f\colon D \to \mathbb{K}$, dass $(cf)' = c \cdot f'$.

(c) Mit den Regeln aus Satz 10.8 (und Beispiel 10.4 (a)) können wir nun offensichtlich die Ableitung jeder rationalen Funktion, also jeder Funktion der Form $f(x) = \frac{p(x)}{q(x)}$ mit Polynomfunktionen p und q berechnen. Ist z. B. $f(x) = \sum_{k=0}^{n} a_k x^k$ selbst eine Polynomfunktion, so ist $f'(x) = \sum_{k=1}^{n} k a_k x^{k-1}$ nach (a), (b) und Satz 10.8 (a).

Wir kommen jetzt zu Verkettungen und Umkehrfunktionen.

Satz 10.10 (Kettenregel). *Es seien $f: D \to \mathbb{K}$ und $g: D' \to \mathbb{K}$ zwei Funktionen, so dass $f(D) \subset D'$. Ist dann $a \in D$, so dass f differenzierbar in a und g differenzierbar in $f(a)$ ist, so ist auch die Verkettung $g \circ f$ differenzierbar in a, und es gilt*

$$(g \circ f)'(a) = g'(f(a)) \cdot f'(a),$$

d. h. „die Ableitung einer Verkettung ist das Produkt der beiden Ableitungen“.

Beweis. Da f und g in a bzw. $f(a)$ differenzierbar sind, gibt es nach Lemma 10.5 Funktionen $\varphi: D \to \mathbb{K}$ und $\psi: D' \to \mathbb{K}$, die in a bzw. $f(a)$ stetig sind, und für die

$$\begin{aligned} & f(x) - f(a) = \varphi(x)\,(x-a) && \text{für alle } x \in D \\ \text{bzw.} \quad & g(y) - g(f(a)) = \psi(y)\,(y - f(a)) && \text{für alle } y \in D' \end{aligned}$$

sowie $\varphi(a) = f'(a)$ und $\psi(f(a)) = g'(f(a))$ gelten. Setzen wir nun $y = f(x)$, so erhalten wir durch Einsetzen der ersten Gleichung in die zweite

$$g(f(x)) - g(f(a)) = \psi(f(x))\,(f(x) - f(a)) = \psi(f(x))\,\varphi(x)\,(x-a)$$

für alle $x \in D$. Da f und φ in a sowie ψ in $f(a)$ stetig sind, ist nun aber auch $x \mapsto \psi(f(x))\,\varphi(x)$ in a stetig, und somit ergibt sich aus der Richtung „(b) $\Rightarrow$ (a)“ von Lemma 10.5 angewendet auf $g \circ f$, dass diese Funktion in a differenzierbar ist mit $(g \circ f)'(a) = \psi(f(a))\,\varphi(a) = g'(f(a)) \cdot f'(a)$. □

Satz 10.11 (Ableitung der Umkehrfunktion)**.** *Es seien $D, D' \subset \mathbb{K}$ und $f: D \to D'$ eine bijektive Funktion mit Umkehrfunktion $f^{-1}: D' \to D$. Ist dann $a \in D$ ein Punkt, so dass f differenzierbar in a ist mit $f'(a) \neq 0$, und so dass f^{-1} stetig in $f(a)$ ist, so ist auch f^{-1} differenzierbar in $f(a)$ mit*

$$(f^{-1})'(f(a)) = \frac{1}{f'(a)}.$$

Beweis. Wir berechnen die Ableitung von f^{-1} in $b := f(a)$ mit dem Folgenkriterium aus Satz 8.11. Es sei dazu $(y_n)_n$ eine beliebige Folge in $D' \backslash \{b\}$ mit $y_n \to b$. Da f^{-1} nach Voraussetzung in b stetig ist, gilt dann $x_n := f^{-1}(y_n) \to f^{-1}(b) = a$. Also ist $\lim\limits_{n\to\infty} \frac{f(x_n)-f(a)}{x_n - a} = f'(a)$, und somit nach dem Grenzwertsatz 5.13 (c) für Quotienten

$$\lim_{n\to\infty} \frac{f^{-1}(y_n) - f^{-1}(b)}{y_n - b} = \lim_{n\to\infty} \frac{x_n - a}{f(x_n) - f(a)} = \frac{1}{f'(a)},$$

woraus die Behauptung folgt. □

Bemerkung 10.12. Im Fall einer reellen, streng monotonen Funktion f benötigen wir die Voraussetzung der Stetigkeit von f^{-1} in $f(a)$ in Satz 10.11 nicht, da dies nach Satz 8.27 automatisch erfüllt ist. Die Bedingung $f'(a) \neq 0$ ist hingegen auch in diesem Fall nicht überflüssig: Das Beispiel der reellen Funktion $f: \mathbb{R} \to \mathbb{R}$, $x \mapsto x^3$ mit $f'(0) = 0$ zeigt, dass eine differenzierbare, streng monotone Funktion in einem Punkt auch Ableitung Null haben kann.

Beispiel 10.13. In den Sätzen 10.10 und 10.11 werden nicht alle Ableitungen an derselben Stelle a, sondern manche auch an $f(a)$ ausgewertet. Man macht dies „automatisch“ richtig, wenn man wie in den folgenden beiden Beispielen für die Definitions- und Wertemengen der beteiligten Funktionen bestimmte Variablennamen festlegt und darauf achtet, dass Funktionen und ihre Ableitungen immer an der entsprechenden Variablen ausgewertet werden.

(a) Die Funktion $h\colon \mathbb{R}\to\mathbb{R}$, $x\mapsto y=\sqrt{x^2+1}$ ist von der Form $h=g\circ f$ mit $f\colon x\mapsto u=x^2+1$ und $g\colon u\mapsto y=\sqrt{u}$. Ihre Ableitung ergibt sich daher nach der Kettenregel aus Satz 10.10 zu

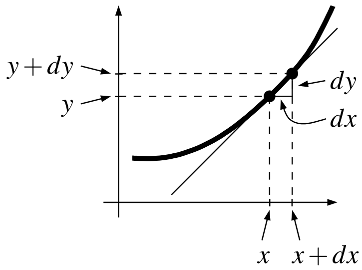
$$h'(x)=g'(u)\cdot f'(x)=\frac{1}{2\sqrt{u}}\cdot 2x=\frac{x}{\sqrt{x^2+1}},$$

denn $f'(x)=2x$ und $g'(u)=\frac{1}{2\sqrt{u}}$ nach Beispiel 10.9 (c) und 10.4 (b).

(b) Nach Beispiel 10.9 (a) ist die Ableitung der Potenzfunktion $f\colon \mathbb{R}_{>0}\to\mathbb{R}_{>0}$, $x\mapsto y=x^n$ für $n\in\mathbb{N}_{>0}$ gleich $f'(x)=nx^{n-1}$. Damit ist die Ableitung ihrer Umkehrfunktion, also der n-ten Wurzelfunktion $f^{-1}\colon y\mapsto x=\sqrt[n]{y}=y^{1/n}$, nach Satz 10.11 gleich

$$(f^{-1})'(y)=\frac{1}{f'(x)}=\frac{1}{nx^{n-1}}=\frac{1}{n\,(\sqrt[n]{y})^{n-1}}=\frac{1}{n}\cdot y^{\frac{1}{n}-1}.$$

Notation 10.14 (Differentialschreibweise)**.** Die Regeln aus den Sätzen 10.10 und 10.11 lassen sich leicht mit Hilfe der sogenannten Differentialschreibweise merken: Man legt hierzu wie in Beispiel 10.13 für eine Funktion $f\colon D\to\mathbb{K}$ bestimmte Variablennamen für Definitions- und Wertemenge fest, etwa $y=f(x)$, und schreibt die Ableitung $f'(x)$ dann als formalen Quotienten $\frac{dy}{dx}$, wobei die „Differentiale" dx und dy wie im Bild rechts für eine (unendlich kleine) Differenz in den x- und y-Werten stehen sollen.



Wichtig dabei ist, dass dies nur eine formale Schreibweise ist – es gibt nicht wirklich Objekte dx und dy, die hier durcheinander geteilt werden. Dennoch nehmen die Sätze 10.10 und 10.11 in dieser Schreibweise eine sehr natürliche Form an, die so aussieht, als könnte man mit diesen „Brüchen" wirklich rechnen:

(a) (Verkettung) Ist $h=g\circ f$ eine Verkettung und setzen wir $u=f(x)$, $y=g(u)$ und damit $y=h(x)$, so besagt Satz 10.10 einfach $\frac{dy}{dx}=\frac{dy}{du}\cdot\frac{du}{dx}$, so als ob man hier mit du erweitern würde.

(b) (Umkehrfunktion) Ist $y=f(x)$, also $x=f^{-1}(y)$, so würden wir die Ableitung von f^{-1} ja als $\frac{dx}{dy}$ schreiben, und damit sagt Satz 10.11 gerade $\frac{dx}{dy}=\left(\frac{dy}{dx}\right)^{-1}$, so als ob man hier einfach den Kehrwert des Bruches $\frac{dy}{dx}$ bilden würde.

Aufgabe 10.15. Es seien $f,g\colon\mathbb{R}\to\mathbb{R}$ stetige Funktionen. Man beweise oder widerlege:

(a) Ist f differenzierbar in 0 und $g(0)=0$, dann ist $f\cdot g$ differenzierbar in 0.

(b) Ist f differenzierbar in 0 und $f(0)=0$, dann ist $f\cdot g$ differenzierbar in 0.

(c) Sind f und g differenzierbar in 0 mit $f(0)=0$ und $f(x)g(x)=x$ für alle $x\in\mathbb{R}$, dann gilt $g(0)\neq 0$.

Aufgabe 10.16. Es seien $D\subset\mathbb{R}$ und $f\colon D\to\mathbb{R}$ eine Funktion. Zeige, dass für jedes $a\in D$ und $c\in\mathbb{R}$ die folgenden beiden Bedingungen äquivalent sind:

(a) Die Funktion f ist differenzierbar in a mit $f'(a)=c$.

(b) Für zwei beliebige gegen a konvergente Folgen $(x_n)_n$ und $(y_n)_n$ in D mit $x_n\le a\le y_n$ und $x_n\neq y_n$ für alle $n\in\mathbb{N}$ gilt

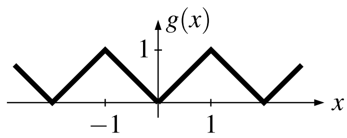
$$\lim_{n\to\infty}\frac{f(y_n)-f(x_n)}{y_n-x_n}=c.$$

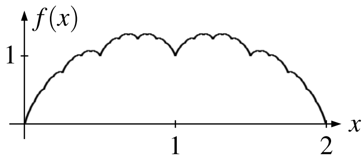
Aufgabe 10.17. Wir betrachten wie im Bild rechts dargestellt die „Zickzackfunktion“ $g\colon \mathbb{R} \to \mathbb{R}$, die periodisch mit Periodenlänge 2 ist und auf $[-1,1]$ mit der Betragsfunktion übereinstimmt. Zeige, dass dann die Funktion

$$f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \sum_{n=0}^{\infty} \frac{g(2^n x)}{2^n}$$

in jedem Punkt stetig und in keinem Punkt differenzierbar ist.

(Hinweis: Für die Differenzierbarkeit ist Aufgabe 10.16 nützlich.)





10.B Extremwerte und der Mittelwertsatz

Nachdem wir nun schon einige Ableitungen berechnen können, wollen wir uns als Nächstes anschauen, welche Informationen man über eine differenzierbare Funktion aus ihrer Ableitung erhalten kann. Am wichtigsten ist dabei, dass man mit Hilfe der Ableitung sehr leicht die Stellen finden kann, an denen eine Funktion ihre größten bzw. kleinsten Werte annimmt. Dies zu untersuchen ist natürlich nur für reelle Funktionen sinnvoll, und daher beschränken wir uns im Folgenden auf solche.

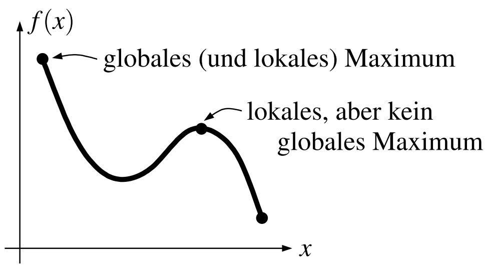
Definition 10.18 (Extrema)**.** Es seien $D \subset \mathbb{R}$, $f\colon D \to \mathbb{R}$ eine Funktion und $a \in D$. Man sagt, …

(a) f habe in a ein **(globales) Maximum**, wenn $f(a) \geq f(x)$ für alle $x \in D$.

(b) f habe in a ein **lokales Maximum**, wenn es ein $\varepsilon > 0$ gibt, so dass $f(a) \geq f(x)$ für alle $x \in D$ mit $|x-a| < \varepsilon$. Gilt sogar $f(a) > f(x)$ für alle $x \in D$ mit $|x-a| < \varepsilon$ und $x \neq a$, so nennt man das lokale Maximum **isoliert**.

Analog definiert man globale und lokale (isolierte) **Minima**. Hat f in a ein (globales, lokales, isoliertes) Maximum oder Minimum, so sagt man auch, dass f dort ein (globales, lokales, isoliertes) **Extremum** hat.

Bemerkung 10.19. Ein globales Maximum (analog Minimum) bedeutet also gerade, dass f dort den größten aller möglichen Funktionswerte annimmt; ein lokales Maximum dagegen nur, dass f in einer kleinen Umgebung des betrachteten Punktes den größten Wert hat. Offensichtlich ist also jedes globale Maximum auch ein lokales. Das Bild rechts zeigt, dass die Umkehrung nicht notwendig richtig ist. 23



Wie wir im Bild schon sehen, zeichnet sich ein lokales Extremum, *das nicht am Rand des Definitionsbereichs liegt*, dadurch aus, dass die Ableitung, also die Steigung der Funktion, dort gleich 0 ist:

Lemma 10.20 (Notwendige Bedingung für lokale Extrema)**.** *Hat eine Funktion $f\colon (a,b) \to \mathbb{R}$ in einem Punkt $x \in (a,b)$ ein lokales Extremum und ist f dort differenzierbar, so gilt $f'(x) = 0$.*

Beweis. Wir beweisen das Lemma für ein Maximum; der Beweis für ein Minimum ist natürlich analog. Nach eventuellem Verkleinern des Definitionsintervalls (a,b) können wir weiterhin ohne Beschränkung der Allgemeinheit annehmen, dass f in x sogar ein globales Maximum hat. Wähle nun eine Folge $(x_n)_n$ in (a,b) mit $x_n \to x$ und $x_n > x$ für alle $n \in \mathbb{N}$ – also eine Folge in D, die sich dem Punkt x von rechts nähert. Die Ableitung $f'(x)$, die nach Voraussetzung existiert, können wir dann nach dem Folgenkriterium aus Satz 8.11 als

$$f'(x) = \lim_{n\to\infty} \frac{f(x_n) - f(x)}{x_n - x}$$

berechnen. Nun ist der Zähler dieses Bruches immer kleiner oder gleich 0 (weil f in x ein Maximum hat), und der Nenner immer größer als Null – und damit folgt $f'(x) \leq 0$ nach Satz 5.24 (a). Durch eine Folge, die sich von links dem Punkt x nähert, erhält man genauso $f'(x) \geq 0$, und damit letztendlich $f'(x) = 0$. □

Bemerkung 10.21. Hat eine reelle Funktion $f\colon [a,b] \to \mathbb{R}$ auf einem abgeschlossenen Intervall in einem Punkt $x \in [a,b]$ ein lokales Extremum, so gibt es also zwei Möglichkeiten:

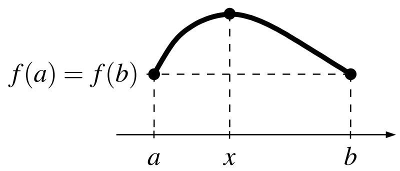
(a) $x \in (a,b)$: Nach Lemma 10.20 muss dann $f'(x) = 0$ sein, falls f dort differenzierbar ist. Beachte aber, dass die Bedingung $f'(x) = 0$ nicht hinreichend dafür ist, dass in x ein lokales Extremum vorliegt – dies zeigt das Beispiel der Funktion $f(x) = x^3$, für die zwar $f'(0) = 0$ gilt, die bei $x = 0$ aber kein Extremum hat. Punkte $x \in (a,b)$, für die $f'(x) = 0$ gilt, die also als lokales Extremum im Inneren des Definitionsintervalls in Frage kommen, werden oft **kritische Punkte** genannt. Wir werden später noch sehen, wie man feststellen kann, ob ein kritischer Punkt wirklich ein lokales Extremum ist oder nicht (siehe Bemerkung 10.25 und Satz 11.18).

(b) $x = a$ oder $x = b$: In diesem Fall muss die Ableitung von f in x nicht notwendig 0 sein (wie z. B. beim globalen Maximum der Funktion in Bemerkung 10.19). Solche Extremwerte am Rand des Definitionsintervalls nennt man **Randextrema**.

Als Nächstes wollen wir mit Hilfe der Ableitung einer reellen Funktion ihre Monotonieeigenschaften untersuchen. Wir beweisen dazu zunächst zwei einfache Resultate, die wir auch noch für spätere Anwendungen benötigen werden.

Satz 10.22 (Satz von Rolle). *Es sei $f\colon [a,b] \to \mathbb{R}$ eine stetige Funktion, die auf (a,b) differenzierbar ist. Gilt dann $f(a) = f(b)$, so gibt es ein $x \in (a,b)$ mit $f'(x) = 0$.*

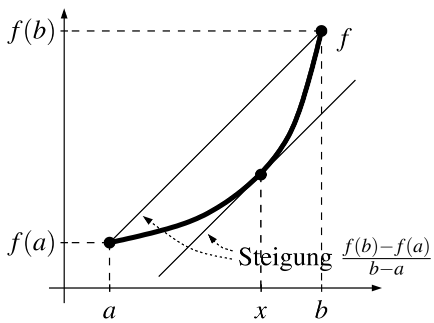
Beweis. Da f stetig ist, nimmt f nach Satz 8.24 auf dem Intervall $[a,b]$ Maximum und Minimum an. Sind diese beide gleich $f(a) = f(b)$, so ist f offensichtlich konstant und wir können ein beliebiges $x \in (a,b)$ wählen. Andernfalls können wir ohne Beschränkung der Allgemeinheit annehmen, dass das Maximum von f größer als $f(a) = f(b)$ ist, also im Inneren des Definitionsintervalls angenommen wird. Dort gilt dann aber $f'(x) = 0$ nach Lemma 10.20. □



Satz 10.23 (Mittelwertsatz).

(a) *(1. Version) Es sei $f\colon [a,b] \to \mathbb{R}$ eine stetige Funktion, die auf (a,b) differenzierbar ist. Dann gibt es ein $x \in (a,b)$ mit $f'(x) = \frac{f(b)-f(a)}{b-a}$. Mit anderen Worten wird die Steigung der Geraden zwischen $(a, f(a))$ und $(b, f(b))$ also wie im Bild rechts an einer Stelle $x \in (a,b)$ als Tangentensteigung angenommen.*

(b) *(2. Version) Es seien $f,g\colon [a,b] \to \mathbb{R}$ stetige, auf (a,b) differenzierbare Funktionen. Dann gibt es ein $x \in (a,b)$ mit*
$$f'(x) \cdot (g(b) - g(a)) = g'(x) \cdot (f(b) - f(a)).$$



Beweis. Es genügt, die allgemeinere Aussage (b) zu zeigen, da sich Teil (a) sofort daraus ergibt, wenn man $g(x) = x$ setzt. Wir betrachten dazu die Funktion
$$h\colon [a,b] \to \mathbb{R},\ x \mapsto (f(x) - f(a))\,(g(b) - g(a)) - (g(x) - g(a))\,(f(b) - f(a)).$$
Mit f und g ist auch h auf $[a,b]$ stetig und auf (a,b) differenzierbar; außerdem ist $h(a) = h(b) = 0$. Nach dem Satz 10.22 von Rolle gibt es also ein $x \in (a,b)$ mit $h'(x) = 0$, und wegen
$$h'(x) = f'(x)\,(g(b) - g(a)) - g'(x)\,(f(b) - f(a))$$

ergibt sich daraus genau die Behauptung. □

Der Mittelwertsatz wirkt auf den ersten Blick etwas unscheinbar, ist in der Tat aber sehr wichtig, da er es erlaubt, einen Differenzenquotienten (und damit letztlich die Differenz zweier Funktionswerte) durch einen Differentialquotienten (also eine Ableitung) auszudrücken. Wie bereits angekündigt ist ein erstes Beispiel hierfür, dass man die Monotonie von Funktionen mit Hilfe von Ableitungen untersuchen kann.

Folgerung 10.24 (Monotonie differenzierbarer Funktionen)**.** *Es sei $f\colon [a,b] \to \mathbb{R}$ eine stetige, auf (a,b) differenzierbare Funktion. Gilt dann für alle $x \in (a,b)$ . . .*

(a) *$f'(x) \geq 0$ (bzw. $f'(x) > 0$), so ist f monoton (bzw. streng monoton) wachsend.*

(b) *$f'(x) \leq 0$ (bzw. $f'(x) < 0$), so ist f monoton (bzw. streng monoton) fallend.*

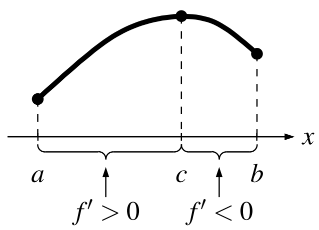
(c) *$f'(x) = 0$, so ist f konstant.*

Beweis. Es seien $x,y \in [a,b]$ mit $x < y$. Nach dem Mittelwertsatz 10.23 (a) angewendet auf das Intervall $[x,y]$ gibt es dann ein $c \in (x,y) \subset (a,b)$ mit $f(y) - f(x) = f'(c)\,(y-x)$. Im Fall (a) ist nun $f'(c) \geq 0$ bzw. $f'(c) > 0$, und damit $f(y) - f(x) \geq 0$ bzw. $f(y) - f(x) > 0$, d. h. f ist monoton (bzw. streng monoton) wachsend. Teil (b) ergibt sich natürlich genauso, und (c) folgt aus der Kombination der beiden Teile. □

Bemerkung 10.25 (Hinreichendes Kriterium für lokale Extrema)**.** Mit Folgerung 10.24 ergibt sich ein einfaches *hinreichendes* Kriterium für ein (lokales) Extremum: Ist $f\colon [a,b] \to \mathbb{R}$ stetig und auf (a,b) differenzierbar, und gibt es ein $c \in (a,b)$ mit

$$f'(x) > 0 \text{ für alle } x < c \text{ und } f'(x) < 0 \text{ für alle } x > c$$

(d. h. hat f' einen Vorzeichenwechsel von $+$ nach $-$ in c), so ist f nach Folgerung 10.24 streng monoton wachsend auf $[a,c]$ und streng monoton fallend auf $[c,b]$, d. h. f hat ein isoliertes lokales Maximum in c. Eine analoge Aussage gilt natürlich auch für ein Minimum.



Wir sehen also, dass man mit Hilfe der Ableitung gut die Extrema von differenzierbaren Funktionen finden kann. Um dies in der Praxis auch anwenden zu können, müssen wir aber auch noch in der Lage sein, von komplizierteren Funktionen – z. B. den „speziellen Funktionen“ aus Kapitel 9 – die Ableitung zu berechnen oder überhaupt erst einmal ihre Differenzierbarkeit nachzuweisen. Da diese Funktionen oftmals über Potenzreihen definiert sind, müssen wir uns also mit der Differenzierbarkeit solcher Potenzreihen (oder allgemeiner von Funktionenfolgen) beschäftigen. Entscheidend hierfür ist die folgende Aussage, die ebenfalls zentral den Mittelwertsatz verwendet.

Satz 10.26 (Vertauschbarkeit von Differentiation und Grenzwertbildung)**.** *Es seien $D \subset \mathbb{R}$ ein Intervall und $(f_n)_n$ mit $f_n\colon D \to \mathbb{R}$ eine Folge differenzierbarer Funktionen. Wir setzen voraus, dass*

- *$(f_n)_n$ punktweise gegen eine Grenzfunktion $f\colon D \to \mathbb{R}$ konvergiert, und*
- *die Ableitungen f_n' stetig sind und gleichmäßig gegen eine Grenzfunktion $g\colon D \to \mathbb{R}$ konvergieren.*

Dann ist f differenzierbar mit $f' = g$ (d. h. „Differentiation und Grenzwertbildung können vertauscht werden“; es ist $\left(\lim_{n\to\infty} f_n\right)' = \lim_{n\to\infty} f_n'$).

Beweis. Für alle $a \in D$ zeigen wir direkt mit der Grenzwertdefinition, dass $\lim_{x\to a} \frac{f(x)-f(a)}{x-a} = g(a)$. Es sei also $\varepsilon > 0$ beliebig. Da g nach Satz 8.37 als gleichmäßiger Grenzwert stetiger Funktionen stetig ist, gibt es zunächst ein $\delta > 0$, so dass

$$|g(x) - g(a)| < \frac{\varepsilon}{3} \qquad \text{für alle } x \in D \text{ mit } |x-a| < \delta. \tag{1}$$

Außerdem konvergiert $(f'_n)_n$ nach Voraussetzung gleichmäßig gegen g, d. h. es gibt auch ein (von x unabhängiges) $n_0 \in \mathbb{N}$ mit

$$|f'_n(x) - g(x)| < \frac{\varepsilon}{3} \qquad \text{für alle } x \in D \text{ und } n \geq n_0. \tag{2}$$

Es seien nun $x \in D$ mit $x \neq a$ und $|x-a| < \delta$ sowie $n \geq n_0$ beliebig. Nach dem Mittelwertsatz 10.23 (a) gibt es dann ein c zwischen a und x (für das also insbesondere auch $|c-a| < |x-a| < \delta$ gilt) mit

$$\frac{f_n(x) - f_n(a)}{x-a} = f'_n(c). \tag{3}$$

Setzen wir dies nun alles zusammen, so erhalten wir mit der Dreiecksungleichung

$$\begin{aligned}\left| \frac{f_n(x) - f_n(a)}{x-a} - g(a) \right| &\stackrel{(3)}{=} |f'_n(c) - g(a)| = |f'_n(c) - g(c) + g(c) - g(a)| \\ &\leq |f'_n(c) - g(c)| + |g(c) - g(a)| \\ &\stackrel{(1),(2)}{<} \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \frac{2\varepsilon}{3},\end{aligned}$$

wobei wir (1) und (2) für den Punkt c angewendet haben. Nehmen wir hier nun den Grenzwert für $n \to \infty$, so erhalten wir daraus mit Satz 5.24 (a) für alle x mit $|x-a| < \delta$

$$\left| \frac{f(x) - f(a)}{x-a} - g(a) \right| \leq \frac{2\varepsilon}{3} < \varepsilon.$$

Da $\varepsilon > 0$ beliebig war, bedeutet dies aber genau, dass $f'(a) = \lim\limits_{x \to a} \frac{f(x)-f(a)}{x-a} = g(a)$. □

Auch wenn der Beweis dieses Satzes recht kompliziert war, ist die Aussage doch sehr einfach anzuwenden. So ergibt sich z. B. in dem für uns wichtigsten Fall von Potenzreihen:

Folgerung 10.27 (Differenzierbarkeit von Potenzreihen)**.** *Es sei $f(x) = \sum_{k=0}^{\infty} a_k x^k$ eine reelle Potenzreihe mit Konvergenzradius r. Dann ist f im Konvergenzgebiet $(-r,r)$ differenzierbar mit Ableitung $f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$, d. h. „Potenzreihen können gliedweise differenziert werden".*

Beweis. Es sei $c \in (-r,r)$; wir wollen zeigen, dass f in c differenzierbar ist mit Ableitung $f'(c) = \sum_{k=1}^{\infty} k a_k c^{k-1}$. Wähle dazu ein R mit $|c| < R < r$. Dann ist die Folge der Partialsummen $f_n(x) = \sum_{k=0}^{n} a_k x^k$ nach Satz 8.35 auf $(-R,R)$ gleichmäßig konvergent gegen f. Die Ableitung dieser Partialsummenfunktionen sind nach Beispiel 10.9 (c) die stetigen Funktionen $f'_n(x) = \sum_{k=1}^{n} k a_k x^{k-1}$. Da die Reihe $g(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$ nach Aufgabe 7.31 den gleichen Konvergenzradius r wie f hat, konvergieren genauso auch die f'_n auf $(-R,R)$ gleichmäßig gegen g.

Damit haben wir alle Voraussetzungen überprüft, um Satz 10.26 auf die Folge $(f_n)_n$ auf $(-R,R)$ anwenden zu können. Der Satz liefert uns also $f' = g$ auf $(-R,R)$, und damit insbesondere auch im Punkt c. □

Mit Folgerung 10.27 (und unseren vorherigen Resultaten) können wir jetzt endlich von „praktisch allen" Funktionen die Ableitungen berechnen:

Beispiel 10.28 (Ableitungen spezieller Funktionen)**.**

(a) Die Ableitung der (reellen) Exponentialfunktion $f(x) = \mathrm{e}^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$ ergibt sich durch gliedweises Differenzieren zu

$$f'(x) = \sum_{n=1}^{\infty} \frac{n x^{n-1}}{n!} = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!} = \mathrm{e}^x;$$

die Exponentialfunktion ist also gleich ihrer eigenen Ableitung.

(b) Die Ableitung der Sinusfunktion $f(x) = \sin x = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1}$ ist analog

$$f'(x) = \sum_{n=0}^{\infty} (2n+1) \cdot \frac{(-1)^n}{(2n+1)!} x^{2n} = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = \cos x.$$

Genauso berechnet man $\cos'(x) = -\sin x$. Aus der Quotientenregel von Satz 10.8 (c) ergibt sich damit

$$\tan'(x) = \left(\frac{\sin}{\cos}\right)'(x) = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = 1 + \tan^2 x.$$

24

(c) Die Ableitungen der Umkehrfunktionen zu (a) und (b) folgen nun sofort aus Satz 10.11: Mit $f\colon x \mapsto y = \mathrm{e}^x$, also $f'(x) = \mathrm{e}^x$ und $f^{-1}(y) = \log y$ ist z. B.

$$\log'(y) = \frac{1}{f'(x)} = \frac{1}{\mathrm{e}^x} = \frac{1}{y}.$$

Für $f\colon x \mapsto y = \sin x$, also $f'(x) = \cos x$ und $f^{-1}(y) = \arcsin y$ ist analog

$$\arcsin'(y) = \frac{1}{f'(x)} = \frac{1}{\cos x} \overset{(*)}{=} \frac{1}{\sqrt{1-\sin^2 x}} = \frac{1}{\sqrt{1-y^2}},$$

wobei wir in $(*)$ die Gleichung aus Satz 9.14 (b) benutzen (sowie dass der Arkussinus streng monoton wachsend ist, so dass wir hier das positive Vorzeichen der Wurzel nehmen müssen). Genauso zeigt man

$$\arccos'(y) = -\frac{1}{\sqrt{1-y^2}} \qquad \text{und} \qquad \arctan'(y) = \frac{1}{1+y^2}.$$

(d) Die Ableitung der Potenzfunktion $f(x) = x^a = \mathrm{e}^{a \log x}$ (mit festem Exponenten a und variabler Basis x) ist nach der Kettenregel aus Satz 10.10

$$f'(x) = \mathrm{e}^{a \log x} \cdot \frac{a}{x} = x^a \cdot \frac{a}{x} = a x^{a-1},$$

das Ergebnis ist also für alle $a \in \mathbb{R}$ das gleiche wie schon für die speziellen Exponenten in Beispiel 10.9 (a) und 10.13 (b). Die Ableitung der Potenzfunktion $f(x) = a^x = \mathrm{e}^{x \log a}$ mit fester Basis und variablem Exponenten ist hingegen ebenfalls nach der Kettenregel

$$f'(x) = \mathrm{e}^{x \log a} \cdot \log a = \log a \cdot a^x.$$

Bemerkung 10.29.

(a) Man kann zeigen, dass die Aussage von Satz 10.26 (und damit auch von Folgerung 10.27) genauso auch im komplexen Fall gilt. Da wir in unserem Beweis dieser Aussagen den Mittelwertsatz verwendet haben (der nur in $\mathbb{R}$ gilt), benötigt man hierfür jedoch andere Argumente. Wir werden den komplexen Fall im Folgenden in dieser Vorlesung aber nicht benötigen – die Untersuchung komplex differenzierbarer Funktionen bzw. Potenzreihen ist der wesentliche Inhalt der Vorlesung „Einführung in die Funktionentheorie", die ihr im zweiten Studienjahr hören könnt.

(b) Ohne die Voraussetzung der (gleichmäßigen) Konvergenz der Folge der Ableitungen in Satz 10.26 wäre die Aussage im Allgemeinen falsch: Betrachten wir z. B. die Folge $f_n\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \frac{1}{n}\sin(nx)$, so gilt zwar $\lim\limits_{n\to\infty} f_n(x) = 0$ für alle x, d. h. $(f_n)_n$ konvergiert punktweise (und in der Tat auch gleichmäßig) gegen die Nullfunktion, die ja Ableitung 0 hat – aber die Folge der Ableitungen $f_n'(x) = \cos(nx)$ konvergiert für $n \to \infty$ noch nicht einmal punktweise! Hier können Differentiation und Grenzwertbildung also nicht vertauscht werden.

Aufgabe 10.30. Bestimme alle lokalen und globalen Extrema der Funktion $f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \frac{\mathrm{e}^x}{1+2|x|}$.

Aufgabe 10.31. Es sei $f\colon [a,b] \to \mathbb{R}$ eine differenzierbare Funktion. Man zeige:

(a) Ist $f'(a) > 0$ und $f'(b) < 0$, so gibt es ein $x \in (a,b)$ mit $f'(x) = 0$.

(b) Für alle c zwischen $f'(a)$ und $f'(b)$ gibt es ein $x \in [a,b]$ mit $f'(x) = c$. (Ableitungen erfüllen also den Zwischenwertsatz, obwohl sie nach Aufgabe 10.33 (d) nicht stetig sein müssen.)

Aufgabe 10.32. Untersuche die Funktionenfolge $(f_n)_n$ mit $f_n\colon \mathbb{R}_{\geq 0} \to \mathbb{R},\ x \mapsto x^{n+1}\mathrm{e}^{-nx}$ auf gleichmäßige Konvergenz.

Aufgabe 10.33. Finde $m \in \mathbb{N}$ und $n \in \mathbb{N}_{>0}$, so dass $f\colon [-1,1] \to \mathbb{R},\ x \mapsto \begin{cases} x^m \sin \frac{1}{x^n} & \text{für } x \neq 0 \\ 0 & \text{für } x = 0 \end{cases}$ ...

(a) unstetig ist.

(b) stetig, aber nicht differenzierbar ist.

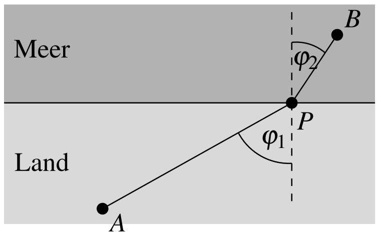
(c) differenzierbar mit unbeschränkter Ableitung ist.

(d) differenzierbar mit beschränkter, aber unstetiger Ableitung ist.

(e) differenzierbar mit stetiger Ableitung ist.

Skizziere für kleine Werte von m und n auch die Graphen dieser Funktionen!

Aufgabe 10.34. Ein Rettungsschwimmer, der sich an Land am Punkt A befindet, möchte eine im Meer ertrinkende Person am Punkt B retten. Er läuft dazu zunächst entlang einer geraden Linie zu einem Punkt P am Ufer, und schwimmt von dort wieder entlang einer geraden Linie nach B. Wenn er mit der Geschwindigkeit v_1 laufen und mit der Geschwindigkeit v_2 schwimmen kann, wo muss er dann den Punkt P wählen, damit er möglichst schnell bei B ist? Zeige, dass diese minimale Zeit genau dort erreicht wird, wo



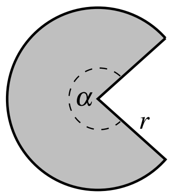
$$\frac{\sin \varphi_1}{\sin \varphi_2} = \frac{v_1}{v_2}.$$

Für die Physiker und physikalisch Interessierten unter euch: Dies ist übrigens genau das Brechungsgesetz für Licht – auch Licht bewegt sich so, dass es schnellstmöglich ans Ziel kommt!

Aufgabe 10.35. Wir betrachten wie im Bild unten rechts Kreissektoren mit variablem Öffnungswinkel $\alpha \in (0, 2\pi)$ und Radius $r \in \mathbb{R}_{>0}$.

Welcher solche Kreissektor hat bei vorgegebenem Flächeninhalt F den kleinstmöglichen Umfang U? Bestimme für diesen Fall r, α und U in Abhängigkeit von F.

(Der Umfang beinhaltet dabei auch die beiden Geradenstücke zum Mittelpunkt. Die Formeln für den Flächeninhalt und Umfang eines Kreissektors können als bekannt vorausgesetzt werden.)



Aufgabe 10.36. Es sei $f\colon D \to \mathbb{R}$ eine auf einem Intervall D definierte Funktion. Wir setzen voraus, dass es $b, c \in \mathbb{R}_{>0}$ gibt mit $|f(x) - f(y)| \leq c\,|x-y|^b$ für alle $x, y \in D$ mit $x \neq y$. Man zeige:

(a) f ist gleichmäßig stetig.

(b) Ist $b > 1$, so ist f konstant.

Aufgabe 10.37. Zeige mit Hilfe des Mittelwertsatzes für alle $x, y \in \mathbb{R}$:

(a) $|\sin x - \sin y| \leq |x-y|$;

(b) $\left|\mathrm{e}^{-x^2} - \mathrm{e}^{-y^2}\right| \leq \sqrt{\frac{2}{\mathrm{e}}} \cdot |x-y|$.

Aufgabe 10.38. Es sei $f\colon \mathbb{R}_{\geq 0} \to \mathbb{R}$ differenzierbar und beschränkt. Zeige mit Hilfe des Mittelwertsatzes, dass es eine Folge $(x_n)_n$ in $\mathbb{R}_{\geq 0}$ gibt mit $\lim\limits_{n\to\infty} x_n = \infty$ und $\lim\limits_{n\to\infty} f'(x_n) = 0$.

11. Anwendungen der Differentialrechnung

Im letzten Kapitel haben wir gesehen, wie wir von „nahezu allen“ Funktionen ihre Ableitung berechnen können und welche elementaren Eigenschaften der Funktion man daran ablesen kann. In diesem Kapitel wollen wir nun zwei weitere Anwendungen vorstellen, die sich aus der Differentialrechnung ergeben. Der Einfachheit halber beschränken wir uns dabei auf reelle Funktionen.

11.A Die Regel von de l’Hôpital

Als Erstes wollen wir eine einfache Regel vorstellen, mit der man oft Grenzwerte berechnen kann, die anders nur schwer zu bestimmen wären: nämlich Grenzwerte der Form $\lim\limits_{x\to a} \frac{f(x)}{g(x)}$, bei denen die normalen Grenzwertsätze aus Satz 8.13 bzw. Bemerkung 8.19 nicht anwendbar sind, weil sich die unbestimmten Quotienten „$\frac{0}{0}$“ oder „$\frac{\pm\infty}{\pm\infty}$“ ergeben würden.

Es gibt viele Varianten dieser Regel, je nachdem, bei welchen der im folgenden Satz vorkommenden Grenzwerten auch uneigentliche Grenzwerte $\pm\infty$ zugelassen sind. In der Praxis treten alle diese Varianten auch oft auf. Um die Beweisidee des Satzes klar herauszustellen, beschränken wir uns hier aber zunächst (wie auch bei unseren bisherigen Beweisen von Rechenregeln für Grenzwerte) auf den Fall, in dem keine uneigentlichen Grenzwerte vorkommen, und geben die möglichen Verallgemeinerungen dann in der anschließenden Bemerkung 11.2 an.

Satz 11.1 (**Regel von de l’Hôpital**, Grundversion)**.** *Es seien $f,g\colon (a,b)\to\mathbb{R}$ differenzierbare Funktionen mit $g'(x)\neq 0$ für alle $x\in(a,b)$. Wir nehmen ferner an, dass*

$$\lim_{x\to a} f(x) = \lim_{x\to a} g(x) = 0$$

(so dass der Grenzwert $\lim\limits_{x\to a}\frac{f(x)}{g(x)}$ also formal von der unbestimmten Form „$\frac{0}{0}$“ ist). Existiert dann der Grenzwert $\lim\limits_{x\to a}\frac{f'(x)}{g'(x)}$ in $\mathbb{R}$, so auch $\lim\limits_{x\to a}\frac{f(x)}{g(x)}$, und es gilt

$$\lim_{x\to a}\frac{f(x)}{g(x)} = \lim_{x\to a}\frac{f'(x)}{g'(x)}.$$

Beweis. Wegen $\lim\limits_{x\to a} f(x) = \lim\limits_{x\to a} g(x) = 0$ können wir f und g durch $f(a) := g(a) := 0$ stetig nach $[a,b)$ fortsetzen. Beachte außerdem, dass g auf (a,b) nirgends gleich 0 sein kann, denn sonst gäbe es im Widerspruch zur Voraussetzung nach dem Satz 10.22 von Rolle zwischen a und dieser Nullstelle von g eine Nullstelle von g'.

Wir zeigen nun den Grenzwert $\lim\limits_{x\to a}\frac{f(x)}{g(x)}$ mit dem Folgenkriterium. Es sei also $(x_n)_n$ eine Folge in (a,b) mit $x_n\to a$. Nach dem Mittelwertsatz 10.23 (b) gibt es dann für alle $n\in\mathbb{N}$ ein $c_n\in(a,x_n)$ mit

$$f'(c_n)\cdot(g(x_n)-g(a)) = g'(c_n)\cdot(f(x_n)-f(a)), \quad \text{also} \quad \frac{f(x_n)}{g(x_n)} = \frac{f'(c_n)}{g'(c_n)}$$

wegen $f(a)=g(a)=0$ und der Nullstellenfreiheit von g und g'. Nun konvergiert wegen $a<c_n<x_n$ mit $(x_n)_n$ aber auch $(c_n)_n$ gegen a, und damit ergibt sich

$$\lim_{n\to\infty}\frac{f(x_n)}{g(x_n)} = \lim_{n\to\infty}\frac{f'(c_n)}{g'(c_n)} = \lim_{x\to a}\frac{f'(x)}{g'(x)},$$

was die Behauptung mit dem Folgenkriterium zeigt. □

Bemerkung 11.2 (**Regel von de l’Hôpital**, Varianten)**.** Der Satz 11.1 von de l’Hôpital hat die folgenden Varianten, die wir im Folgenden ebenfalls verwenden werden. Die Beweise sollen hier nicht gegeben werden – sie lassen sich mit analogen, allerdings oft technisch etwas aufwendigeren Methoden führen.

(a) Die Regel gilt auch, wenn $\lim\limits_{x\to a} f(x)$ und $\lim\limits_{x\to a} g(x)$ beide im uneigentlichen Sinne gleich $\pm\infty$ sind, wir also beim Grenzwert $\lim\limits_{x\to a} \frac{f(x)}{g(x)}$ den unbestimmten Ausdruck „$\frac{\pm\infty}{\pm\infty}$“ haben.

(b) Statt einer Annäherung von rechts an die Grenze der Definitionsmenge (in unserem Fall also an den Punkt a am Rand des Intervalls (a,b)) ist natürlich auch eine Annäherung von links oder eine beidseitige Annäherung möglich. Dabei sind die Fälle $a = -\infty$ und $b = \infty$ zugelassen.

(c) Die Regel gilt auch, wenn der Grenzwert von $\frac{f'(x)}{g'(x)}$ nur im uneigentlichen Sinne existiert, also gleich $\pm\infty$ ist.

Existiert der Grenzwert von $\frac{f'(x)}{g'(x)}$ dagegen auch im uneigentlichen Sinne nicht, so macht die Regel von de l'Hôpital keine Aussage – wir können daraus dann also nicht schließen, dass auch der ursprüngliche Grenzwert von $\frac{f(x)}{g(x)}$ nicht existiert!

Beispiel 11.3.

(a) Für jedes $n \in \mathbb{N}_{>0}$ ist der Grenzwert $\lim\limits_{x\to\infty} \frac{\log x}{x^n}$ von der Form „$\frac{\infty}{\infty}$“. Um ihn mit der Regel von de l'Hôpital (mit $f(x) = \log x$ und $g(x) = x^n$) zu bestimmen, differenzieren wir also Zähler und Nenner separat und erhalten den Bruch $\frac{f'(x)}{g'(x)} = \frac{1/x}{nx^{n-1}}$. Da der Nenner dieses Bruchs für $x > 0$ nirgends gleich 0 ist und der Grenzwert

$$\lim_{x\to\infty} \frac{1/x}{nx^{n-1}} = \lim_{x\to\infty} \frac{1}{nx^n} = 0$$

existiert, folgt mit Satz 11.1 (bzw. der Verallgemeinerung aus Bemerkung 11.2) also auch

$$\lim_{x\to\infty} \frac{\log x}{x^n} \overset{\text{„}\frac{\infty}{\infty}\text{“}}{\underset{\downarrow}{=}} \lim_{x\to\infty} \frac{1/x}{nx^{n-1}} = 0. \qquad (*)$$

Analog zu Bemerkung 5.15 zur Anwendung von Grenzwertsätzen schreibt man dabei die Anwendung der Regel von de l'Hôpital oftmals gleich wie in der oben in $(*)$ mit „$\frac{\infty}{\infty}$“ bezeichneten Gleichung, und überprüft erst nachträglich, dass der neu entstandene Bruch einen (evtl. uneigentlichen) Grenzwert hat und sein Nenner stets ungleich 0 ist.

(b) Analog erhält man

$$\lim_{x\to 0} x^n \log x = \lim_{x\to 0} \frac{\log x}{x^{-n}} \overset{\text{„}\frac{-\infty}{\infty}\text{“}}{\underset{\downarrow}{=}} \lim_{x\to 0} \frac{1/x}{-nx^{-n-1}} = \lim_{x\to 0} -\frac{x^n}{n} = 0,$$

da der nach dem Differenzieren entstandene Nenner $-nx^{-n-1}$ für $x > 0$ ungleich 0 ist. Zusammen mit (a) sehen wir in diesem Sinne also, dass „der Logarithmus für $x \to 0$ oder $x \to \infty$ schwächer ist als jede Potenz“ – in den beiden betrachteten Grenzwerten setzt sich jeweils die Funktion x^n durch. Dies ist natürlich ganz analog zu der Aussage von Bemerkung 9.3 (a), dass die Exponentialfunktion schneller als jede Potenz wächst. Beachte auch, dass wir in der zweiten Rechnung oben gesehen haben, dass es sich auch bei einem ursprünglichen Ausdruck der Form „$0 \cdot (\pm\infty)$“ lohnen kann, ihn künstlich als Bruch umzuschreiben, um dann die Regel von de l'Hôpital anwenden zu können.

(c) Falls sich nach einmaliger Anwendung von Satz 11.1 immer noch ein Bruch der Form „$\frac{0}{0}$“ oder „$\frac{\pm\infty}{\pm\infty}$“ ergibt, kann man den Satz natürlich auch mehrfach hintereinander anwenden, wie z. B. in dem Grenzwert

$$\lim_{x\to\infty} \frac{x^2}{e^x} \overset{\text{„}\frac{\infty}{\infty}\text{“}}{\underset{\downarrow}{=}} \lim_{x\to\infty} \frac{2x}{e^x} \overset{\text{„}\frac{\infty}{\infty}\text{“}}{\underset{\downarrow}{=}} \lim_{x\to\infty} \frac{2}{e^x} = 0$$

(wegen $e^x \neq 0$ für alle x), den wir aber natürlich auch schon aus Satz 9.1 (c) kannten.

Bemerkung 11.4 (Die Regel von de l'Hôpital für Folgengrenzwerte)**.** Manchmal kann man auch den Grenzwert einer reellen Folge $(a_n)_n$ mit der Regel von de l'Hôpital berechnen. Kann man die Folge nämlich – betrachtet als Abbildung $\mathbb{N} \to \mathbb{R}$ – zu einer Abbildung $\mathbb{R} \to \mathbb{R}$ fortsetzen, also auch für reelles n betrachten, und existiert dann der Grenzwert für *reelle* $n \to \infty$, so existiert er dann natürlich auch für *natürliche* $n \to \infty$, und hat denselben Wert. Für die Berechnung des Grenzwerts für reelle n haben wir dann aber wieder die Regel von de l'Hôpital zur Verfügung.

Betrachten wir als Beispiel hierfür einmal für gegebenes $x \in \mathbb{R}$ den Grenzwert $\lim\limits_{n\to\infty} \left(1+\frac{x}{n}\right)^n$, den wir in Aufgabe 7.33 mit viel Aufwand zu e^x berechnet haben. Da wir Potenzen inzwischen auch für reelle n definiert haben, können wir den Ausdruck $\left(1+\frac{x}{n}\right)^n$ nun aber auch als Funktion einer reellen Variablen n auffassen und seinen Grenzwert für $n \to \infty$ mit der Regel von de l'Hôpital berechnen: Es ist

$$\begin{aligned}
\lim_{n\to\infty} \left(1+\frac{x}{n}\right)^n &= \lim_{n\to\infty} \exp\left(n \log\left(1+\frac{x}{n}\right)\right) && \text{(Definition 9.7)} \\
&= \exp\left(\lim_{n\to\infty} n \log\left(1+\frac{x}{n}\right)\right) && \text{(Stetigkeit von exp, Satz 8.15)} \\
&= \exp\left(\lim_{n\to\infty} \frac{\log(1+\frac{x}{n})}{n^{-1}}\right) \\
&= \exp\left(\lim_{n\to\infty} \frac{-x/n^2}{1+\frac{x}{n}} \cdot \frac{1}{-1/n^2}\right) && \left(\text{„}\tfrac{0}{0}\text{“, Differenzieren nach } n \text{, beachte } -\tfrac{1}{n^2} \neq 0\right) \\
&= \exp x.
\end{aligned}$$

Aufgabe 11.5. Berechne die folgenden Grenzwerte:

(a) $\lim\limits_{\substack{x\to 0\\ x>0}} \dfrac{\log(\tan(2x))}{\log(\tan(3x))}$ (b) $\lim\limits_{x\to 1} \left(\dfrac{1}{x-1} - \dfrac{1}{\log x}\right)$ (c) $\lim\limits_{\substack{x\to 0\\ x>0}} x^x$

Aufgabe 11.6. Es sei $f\colon [a,b] \to \mathbb{R}$ stetig und auf (a,b) differenzierbar, so dass $\lim\limits_{x\to a} f'(x)$ existiert. Zeige, dass f dann auch in a differenzierbar ist und $f'(a) = \lim\limits_{x\to a} f'(x)$ gilt.

11.B Taylor-Entwicklung

Als weitere Anwendung der Differentialrechnung wollen wir nun unsere ursprüngliche Idee der linearen Approximation einer (reellen) Funktion $f\colon D \to \mathbb{R}$ in einem Punkt $a \in D$ erweitern und uns fragen, ob wir vielleicht noch bessere Näherungen bekommen können, wenn wir als Näherungsfunktion statt einer *linearen* Funktion eine Polynomfunktion von höherem Grad verwenden. Betrachten wir z. B. statt unserer bisherigen Näherung vom Anfang von Kapitel 10

$$f(x) \approx f(a) + c_1\,(x-a)$$

den Ansatz

$$f(x) \approx f(a) + c_1\,(x-a) + c_2\,(x-a)^2,$$

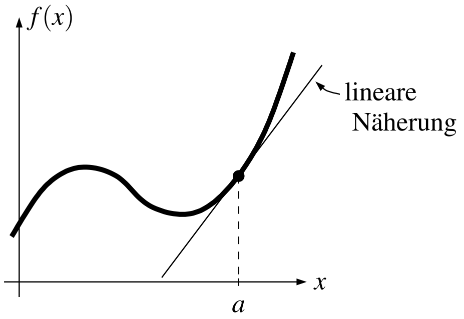
bei dem wir auch einen quadratischen Term zulassen (den wir proportional zu $(x-a)^2$ statt zu x^2 wählen, damit er am Näherungspunkt a selbst verschwindet), so können wir wie im Bild unten erwarten, dass wir eine viel bessere Näherung erhalten, da die Näherungsfunktion ja jetzt eine quadratische Parabel ist und damit auch ein wenig die Krümmung von f an der Stelle a nachbilden kann. Natürlich können wir dies dann auch noch weiter treiben und Polynomfunktionen höheren Grades zulassen: Wenn wir für ein $n \in \mathbb{N}$ einen Ansatz

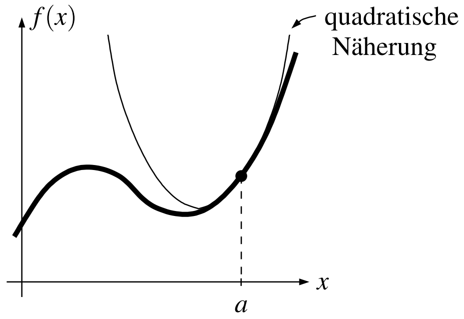
$$f(x) \approx \sum_{k=0}^{n} c_k\,(x-a)^k$$

machen und dabei die Koeffizienten c_k geschickt wählen, sollte die Näherung mit wachsendem n immer besser werden. Wir können sogar versuchen, den Grenzübergang $n \to \infty$ zu machen und uns fragen, ob wir mit einer Potenzreihe

$$\sum_{k=0}^{\infty} c_k\,(x-a)^k$$

im Grenzfall vielleicht nicht nur eine ganz besonders gute Näherung, sondern sogar *genau* die Funktion f zurück erhalten, also ob wir f letztlich als Potenzreihe in $x-a$ schreiben können – schließlich haben wir ja auch wie z. B. die Exponentialfunktion schon einige Funktionen gesehen, die wir von vornherein bereits als Potenzreihe geschrieben haben.





Um diese Idee zu verfolgen, wollen wir nun als Erstes untersuchen, welche Koeffizienten c_k wir in den obigen Polynomen bzw. Reihen wählen sollten. Da wir bereits wissen, dass der lineare Koeffizient c_1 gerade die Ableitung $f'(a)$ ist, sollte es nicht überraschen, dass wir für die höheren Koeffizienten c_k mit $k > 1$ höhere Ableitungen benötigen. Diese wollen wir daher jetzt einführen.

Definition 11.7 (Höhere Ableitungen)**.** Es seien $f\colon D \to \mathbb{R}$ eine Funktion und $n \in \mathbb{N}$.

(a) Die Funktion f heißt ***n*-mal differenzierbar** auf D, wenn alle fortgesetzten Ableitungen $f^{(0)} := f$, $f^{(1)} := f'$, $f^{(2)} := f'' := (f')'$, $\dots$, $f^{(n)} := (f^{(n-1)})'$ existieren.

(b) Die Funktion f heißt ***n*-mal stetig differenzierbar** auf D, wenn zusätzlich $f^{(n)}$ stetig ist.

(c) Existieren die höheren Ableitungen $f^{(n)}$ für alle n, so heißt f **unendlich oft differenzierbar** auf D.

Die Menge aller n-mal stetig differenzierbaren Funktionen auf D wird mit $C^n(D)$ bezeichnet (der Buchstabe C kommt vom englischen Wort „continuous" für „stetig"). Inbesondere ist also $C^0(D)$ die Menge aller stetigen und $C^\infty(D)$ die Menge aller unendlich oft differenzierbaren Funktionen auf D.

Wie diese höheren Ableitungen die oben betrachteten Koeffizienten c_k bestimmen, sieht man am besten wie im folgenden Satz am Beispiel von Potenzreihen, die ja bereits in einer derartigen Form geschrieben sind. (Beachte, dass dies auch noch einmal die Aussage aus Aufgabe 8.42 zeigt, dass die Koeffizienten einer Potenzreihe mit Konvergenzradius ungleich 0 durch die durch sie definierte Funktion bereits eindeutig bestimmt sind.)

Satz 11.8 (Taylor-Formel für Potenzreihen). *Es seien $a \in \mathbb{R}$ und $f(x) = \sum_{k=0}^{\infty} c_k (x-a)^k$ eine reelle Potenzreihe in $x-a$ mit Konvergenzradius $r > 0$, so dass wir f also als Funktion $f\colon (a-r, a+r) \to \mathbb{R}$ auffassen können.*

Dann ist f auf $(a-r, a+r)$ unendlich oft differenzierbar, und für die Koeffizienten c_k der Potenzreihe gilt

$$c_k = \frac{f^{(k)}(a)}{k!} \qquad \text{für alle } k \in \mathbb{N}.$$

Mit anderen Worten ist also

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k = f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots$$

für alle $x \in (a-r, a+r)$.

Beweis. Nach Folgerung 10.27 ist jede Potenzreihe in ihrem Konvergenzgebiet differenzierbar, und ihre Ableitung ist wieder eine Potenzreihe (mit demselben Konvergenzradius), die sich durch gliedweises Differenzieren berechnen lässt. Insbesondere ist f damit also unendlich oft differenzierbar, und die höheren Ableitungen sind

$$f^{(n)}(x) = \sum_{k=n}^{\infty} k(k-1)\cdots(k-n+1)\,c_k\,(x-a)^{k-n}$$

für alle $n \in \mathbb{N}$. Setzen wir hier nun $x = a$ ein, so ist $(x-a)^{k-n}$ gleich 0 für $k > n$ und 1 für $k = n$. In der obigen Summe bleibt dann also nur der Term für $k = n$ übrig, und wir erhalten wie behauptet

$$f^{(n)}(a) = n(n-1)\cdots(n-n+1)\,c_n = n!\,c_n. \qquad \square$$

Beispiel 11.9. Wir betrachten die Funktion $f\colon \mathbb{R} \to \mathbb{R}$, $f(x) = \frac{1}{1+x^2}$ und möchten die 10. Ableitung $f^{(10)}(0)$ im Nullpunkt berechnen. Natürlich könnte man jetzt mit Hilfe der Regeln von Satz 10.8 alle fortgesetzten Ableitungen von f berechnen und schließlich in dem so gefundenen Ausdruck für $f^{(10)}$ den Wert $x = 0$ einsetzen – dies wäre aber sehr zeitaufwendig. Viel schneller geht es mit der Taylor-Formel: Nach der geometrischen Reihe können wir f ja für $|x| < 1$ gemäß

$$f(x) = \frac{1}{1+x^2} = \frac{1}{1-(-x^2)} = \sum_{n=0}^{\infty} (-x^2)^n = 1 - x^2 + x^4 - x^6 + x^8 - x^{10} \pm \cdots$$

als Potenzreihe in x schreiben. Satz 11.8 mit $a = 0$ und $k = 10$ sagt uns also für den Koeffizienten von x^{10} in dieser Reihe, der ja offensichtlich gleich -1 ist, dass

$$-1 = \frac{f^{(10)}(0)}{10!}, \qquad \text{und damit} \qquad f^{(10)}(0) = -10!.$$

Wir sehen an diesem Beispiel schon, dass die Taylor-Formel für Potenzreihen auch dann nützlich ist, wenn die Funktion f ursprünglich gar nicht als Potenzreihe gegeben ist, sondern wir nur wissen, dass es eine solche Darstellung als Potenzreihe gibt. In der Tat benutzt die Taylor-Formel in der Form

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}\,(x-a)^k \qquad (*)$$

ja auch gar nicht mehr die Koeffizienten der ursprünglichen Reihe, sondern nur noch die Tatsache, dass sich f überhaupt irgendwie als Potenzreihe schreiben lässt. Gilt die Formel $(*)$ also vielleicht sogar für jede unendlich oft differenzierbare Funktion f?

Leider ist (wie wir gleich sehen werden) die Antwort auf diese Frage nein. Für viele in der Praxis vorkommende Funktionen ist die Antwort allerdings auch ja – und daher lohnt es sich, die Sache doch noch weiter zu verfolgen. Wir geben der rechten Seite von $(*)$, bzw. den Partialsummen dieser Reihe, daher zunächst einen Namen.

Definition 11.10 (Taylor-Polynom und Taylor-Reihe)**.** Es seien $D \subset \mathbb{R}$ ein Intervall, $f\colon D \to \mathbb{R}$ eine Funktion, sowie $a \in D$ ein fest gewählter Punkt.

(a) Die Funktion f sei n-mal differenzierbar für ein $n \in \mathbb{N}$. Dann heißt die Polynomfunktion

$$T^n_{f,a}\colon D \to \mathbb{R},\ x \mapsto \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}\,(x-a)^k$$

das n**-te Taylor-Polynom** von f mit Entwicklungspunkt a; offensichtlich ist $\deg T^n_{f,a} \le n$.

(b) Ist f unendlich oft differenzierbar, so heißt die Potenzreihe in $x-a$

$$T_{f,a}(x) := \lim_{n\to\infty} T^n_{f,a}(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}\,(x-a)^k$$

die **Taylor-Reihe** von f mit Entwicklungspunkt a. Beachte, dass zunächst nicht klar ist, ob diese Potenzreihe einen Konvergenzradius größer als 0 hat, also ob sie überhaupt in irgendeinem Punkt x (außer a) konvergiert – und dass, selbst wenn sie konvergiert, nicht klar ist, ob sie als Funktion im Konvergenzgebiet mit f übereinstimmt.

Beispiel 11.11.

(a) Lässt sich eine Funktion f auf einem Intervall $(a-r,a+r)$ mit $a \in \mathbb{R}$ und $r>0$ als Potenzreihe in $x-a$ (mit Konvergenzradius mindestens r) schreiben, so besagt Satz 11.8 gerade, dass die Taylor-Reihe $T_{f,a}(x)$ genau diese Reihe ist, also dass $T_{f,a}(x) = f(x)$ für alle
25 $x \in (a-r,a+r)$ gilt.

(b) Wir betrachten die Funktion

$$f\colon \mathbb{R}_{>0} \to \mathbb{R},\ x \mapsto \log x$$

und bestimmen ihre Taylor-Reihe mit Entwicklungspunkt $a=1$. Die Ableitungen von f sind einfach zu berechnen: Wegen $f'(x) = x^{-1}$ ist

$$f^{(k)}(x) = (-1)\cdot(-2)\cdots(-(k-1))\cdot x^{-k} = (-1)^{k-1}\cdot(k-1)!\cdot x^{-k}$$

und damit $f^{(k)}(1) = (-1)^{k-1}\cdot(k-1)!$ für alle $k>0$. Die Taylor-Reihe von f mit Entwicklungspunkt 1 ist damit

$$\begin{aligned} T_{f,1}(x) = \log 1 + \sum_{k=1}^{\infty} \frac{(-1)^{k-1}\cdot(k-1)!}{k!}\,(x-1)^k &= \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}\,(x-1)^k \\ &= (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} \mp \cdots . \end{aligned}$$

Wie in Beispiel 7.29 (a) hat diese Potenzreihe in $x-1$ den Konvergenzradius 1, sie konvergiert also für $|x-1|<1$, d. h. für $x \in (0,2)$, und divergiert für $|x-1|>1$, also für $x<0$ oder $x>2$. Damit ist schon einmal klar, dass die Taylor-Reihe $T_{f,1}(x)$ für $x>2$ sicher *nicht* die ursprüngliche Funktion $f(x) = \log x$ darstellt, da sie dort ja nicht einmal konvergiert. Aber auch für $x \in (0,2)$ ist noch nicht klar, dass wirklich $T_{f,1}(x) = f(x)$ gilt: Das folgende Beispiel zeigt, dass eine Taylor-Reihe auch im Fall der Konvergenz nicht mit der ursprünglichen Funktion übereinstimmen muss.

Aufgabe 11.12. Es sei

$$f\colon \mathbb{R} \to \mathbb{R},\ x \mapsto \begin{cases} \exp\left(-\frac{1}{x^2}\right) & \text{für } x \neq 0, \\ 0 & \text{für } x = 0. \end{cases}$$

Zeige, dass f unendlich oft differenzierbar ist, und dass die Taylor-Reihe $T_{f,0}$ die Nullfunktion ist (also insbesondere zwar überall konvergiert, aber außer im Nullpunkt nirgends mit f übereinstimmt). Skizziere auch den Graphen von f.

(Hinweis: Man zeige mit vollständiger Induktion, dass alle Ableitungen von f in 0 gleich 0 und für $x \neq 0$ von der Form $\frac{p(x)}{q(x)}\exp(-\frac{1}{x^2})$ für gewisse Polynomfunktionen p und q sind.)

Wir benötigen also ein Kriterium, mit dem wir eine Funktion f mit ihren Taylor-Polynomen $T^n_{f,a}$ bzw. ihrer Taylor-Reihe $T_{f,a}$ vergleichen können, so dass wir letztlich nachprüfen können, ob eine (konvergierende) Taylor-Reihe auch wirklich gleich der ursprünglichen Funktion ist. Dies liefert der folgende Satz:

Satz 11.13 (Taylor-Formel). *Es seien $n \in \mathbb{N}$, $D \subset \mathbb{R}$ ein Intervall, und $f\colon D \to \mathbb{R}$ eine $(n+1)$-mal differenzierbare Funktion. Ferner seien $a,x \in D$. Dann gibt es ein c zwischen a und x mit*

$$f(x) - T^n_{f,a}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}\,(x-a)^{n+1}.$$

*Man bezeichnet $f(x) - T^n_{f,a}(x)$ auch als das **Restglied** des n-ten Taylor-Polynoms und schreibt es als $R^n_{f,a}(x)$.*

Beweis. Für den Fall $x=a$ ist die Aussage trivial, da dann beide Seiten der zu zeigenden Formel gleich 0 sind. Für $x \neq a$ behaupten wir, dass die Aussage unmittelbar aus dem Mittelwertsatz 10.23

(b) angewendet auf die beiden Funktionen

$$F\colon D \to \mathbb{R},\ t \mapsto T^n_{f,t}(x) = f(t) + \frac{f'(t)}{1!}(x-t) + \frac{f''(t)}{2!}(x-t)^2 + \cdots + \frac{f^{(n)}(t)}{n!}(x-t)^n$$
$$\text{und} \quad G\colon D \to \mathbb{R},\ t \mapsto (x-t)^{n+1}$$

folgt. In der Tat sind F und G dann differenzierbar mit

$$\begin{aligned} F'(t) &= f'(t) + \Big(-\frac{f'(t)}{1!} + \frac{f''(t)}{1!}(x-t)\Big) + \Big(-\frac{f''(t)}{1!}(x-t) + \frac{f'''(t)}{2!}(x-t)^2\Big) + \cdots \\ &\qquad + \Big(-\frac{f^{(n)}(t)}{(n-1)!}(x-t)^{n-1} + \frac{f^{(n+1)}(t)}{n!}(x-t)^n\Big) \\ &= \frac{f^{(n+1)}(t)}{n!}(x-t)^n \\ \text{und} \quad G'(t) &= -(n+1)(x-t)^n \end{aligned}$$

sowie $F(x) - F(a) = f(x) - T^n_{f,a}(x) = R^n_{f,a}(x)$ und $G(x) - G(a) = -(x-a)^{n+1}$. Der Mittelwertsatz 10.23 (b) liefert also ein c zwischen a und x mit

$$-\frac{f^{(n+1)}(c)}{n!}(x-c)^n \cdot (x-a)^{n+1} = -(n+1)(x-c)^n \cdot R^n_{f,a}(x),$$

d. h. wie behauptet $R^n_{f,a}(x) = \frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}$. □

Bemerkung 11.14.

(a) Für $n = 0$ ist Satz 11.13 exakt der Mittelwertsatz 10.23 (a). Wir können die Taylor-Formel also auch als eine Verallgemeinerung des Mittelwertsatzes auffassen.

(b) Setzen wir in der Formel aus Satz 11.13 noch den Ausdruck aus Definition 11.10 (a) ein, so erhalten wir für jede $(n+1)$-mal differenzierbare Funktion f

$$f(x) = \underbrace{f(a) + \frac{f'(a)}{1!}(x-a) + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n}_{=T^n_{f,a}(x)} + \underbrace{\frac{f^{(n+1)}(c)}{(n+1)!}(x-a)^{n+1}}_{=R^n_{f,a}(x)}$$

für ein c zwischen a und x. Das Restglied des n-ten Taylor-Polynoms hat also genau die Form des $(n+1)$-ten Gliedes der Taylor-Reihe – bis auf den Unterschied, dass man die Ableitung dort an einer Zwischenstelle c anstatt am Entwicklungspunkt a nehmen muss.

(c) Offensichtlich gilt für eine unendlich oft differenzierbare Funktion f nach Satz 11.13 genau dann $T_{f,a}(x) = f(x)$, wenn $\lim\limits_{n\to\infty} R^n_{f,a}(x) = 0$. Wenn man die Taylor-Reihe oder die Taylor-Polynome mit der ursprünglichen Funktion vergleichen möchte, muss man also in irgendeiner Form das Restglied abschätzen. Hier sind zwei Beispiele dafür.

Beispiel 11.15 (Restgliedabschätzung)**.**

(a) Wenden wir Satz 11.13 auf die Taylor-Reihe der Funktion $f\colon \mathbb{R}_{>0} \to \mathbb{R},\ x \mapsto \log x$ mit Entwicklungspunkt $a = 1$ aus Beispiel 11.11 (b) an, so erhalten wir mit den dort berechneten Ableitungen $f^{(n+1)}(x) = (-1)^n \frac{n!}{x^{n+1}}$ für jedes $x \in \mathbb{R}_{>0}$

$$R^n_{f,1}(x) = f(x) - T^n_{f,1}(x) = \frac{(-1)^n}{n+1} \cdot \frac{(x-1)^{n+1}}{c_n^{n+1}}$$

für ein c_n zwischen 1 und x (wir haben den Zwischenwert hier mit c_n statt c bezeichnet, da es natürlich für jedes n ein anderer sein wird). Ist nun $x \in [1,2]$, so ist aber stets $c_n \geq 1$ und $|x-1| \leq 1$, und wir erhalten die Abschätzung

$$|R^n_{f,1}(x)| \leq \frac{1}{n+1},$$

was mit $n \to \infty$ gegen 0 konvergiert. Also konvergieren die Taylor-Polynome für $n \to \infty$ zumindest auf $[1,2]$ wirklich gegen die Funktion f: Es gilt

$$\log x = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} (x-1)^k \qquad \text{für alle } x \in [1,2]. \tag{$*$}$$

Insbesondere ergibt sich damit für $x = 2$ der Wert der alternierenden harmonischen Reihe zu

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

Wir werden später in Beispiel 12.39 (a) übrigens noch sehen, dass die Gleichung $(*)$ sogar für alle $x \in (0,2]$ gilt (also für alle x, für die die Taylor-Reihe überhaupt konvergiert), aber mit unserer bisherigen Formel für das Restglied aus Satz 11.13 können wir das noch nicht beweisen.

(b) Wenn wir als Näherung einer Funktion nur an einem bestimmten Taylor-Polynom (und nicht an der kompletten Reihe) interessiert sind, kann uns Satz 11.13 sagen, wie groß der Fehler ist, den wir dabei machen. Betrachten wir z. B. die Sinusfunktion $f\colon \mathbb{R} \to \mathbb{R}$, $x \mapsto \sin x$, so ist am Entwicklungspunkt 0

$$T_{f,0}^4(x) = x - \frac{x^3}{6}$$

(wie man aus Lemma 9.13 (b) sofort abliest, denn nach Satz 11.8 ist ja jede Potenzreihe ihre eigene Taylor-Reihe). Wegen $f^{(5)}(x) = \cos x$ besagt Satz 11.13 für $n = 4$ nun für alle $x \in \mathbb{R}$

$$R_{f,0}^4(x) = \sin x - \left(x - \frac{x^3}{6}\right) = \frac{\cos c}{5!} x^5$$

für ein c zwischen 0 und x. Wenn wir nun z. B. nur an Werten $x \in \mathbb{R}$ mit $|x| \le \frac{1}{2}$ interessiert sind, so können wir diesen Ausdruck wegen $|\cos c| \le 1$ abschätzen zu

$$|R_{f,0}^4(x)| \le \frac{(\frac{1}{2})^5}{5!} = \frac{1}{3840},$$

d. h. wenn wir für $|x| \le \frac{1}{2}$ den Sinus durch sein viertes Taylor-Polynom $x - \frac{x^3}{6}$ ersetzen, machen wir dabei einen Fehler von höchstens $\frac{1}{3840} \approx 0{,}0003$.

Aufgabe 11.16.

(a) Berechne das Taylor-Polynom $T_{f,1}^2$ für die Funktion $f(x) = \sqrt{x}$ und zeige die Restgliedabschätzung $|f(x) - T_{f,1}^2(x)| \le \frac{1}{20}$ für alle $x \in \left[\frac{1}{2}, \frac{3}{2}\right]$.

(b) Berechne $f^{(10)}(0)$ sowie das Taylor-Polynom $T_{f,0}^{10}$ für die Funktion $f(x) = \frac{\cos(x^5)}{1-2x^6}$.

Aufgabe 11.17. Es sei $f\colon \mathbb{R} \to \mathbb{R}$ eine zweimal differenzierbare Funktion mit $f'' = f$ sowie $f(0) = f'(0) = 1$. Berechne die Taylor-Reihe von f mit Entwicklungspunkt 0 und zeige durch eine Restgliedabschätzung, dass $f = \exp$ die Exponentialfunktion ist.

Als weitere Anwendung der Taylor-Formel wollen wir nun noch ein einfaches hinreichendes Kriterium für lokale Extrema geben. Wir hatten bisher ja nur in Lemma 10.20 gesehen, dass an einem lokalen Extremum, das nicht am Rand der Definitionsmenge liegt, ein kritischer Punkt vorliegen, also die erste Ableitung verschwinden muss – dass diese Bedingung aber nicht für ein lokales Extremum ausreicht. Mit Hilfe höherer Ableitungen und der Taylor-Formel können wir nun ein Kriterium angeben, das nahezu immer und ohne allzu großen Aufwand entscheiden kann, ob wirklich ein lokales Extremum vorliegt oder nicht:

Satz 11.18 (Extremwertkriterium). *Es seien $n \in \mathbb{N}_{>0}$ und $f\colon (a,b) \to \mathbb{R}$ eine n-mal stetig differenzierbare Funktion. Für ein $c \in (a,b)$ gelte $f'(c) = f''(c) = \cdots = f^{(n-1)}(c) = 0$ und $f^{(n)}(c) \neq 0$.*

(a) *Ist n gerade und $f^{(n)}(c) > 0$, so hat f in c ein isoliertes lokales Minimum.*

(b) *Ist n gerade und $f^{(n)}(c) < 0$, so hat f in c ein isoliertes lokales Maximum.*

(c) *Ist n ungerade, so hat f in c kein lokales Extremum.*

Beweis. Es sei ohne Beschränkung der Allgemeinheit $f^{(n)}(c) > 0$ (der Fall $f^{(n)}(c) < 0$ ist analog). Da $f^{(n)}$ nach Voraussetzung stetig ist, gilt nach Bemerkung 8.8 dann sogar $f^{(n)}(x) > 0$ in einer δ-Umgebung von c. Die Taylor-Formel aus Satz 11.13 besagt nun, dass es für alle $x \in (c-\delta, c+\delta)$ ein x' zwischen c und x gibt mit

$$f(x) - T_{f,c}^{n-1}(x) = \frac{f^{(n)}(x')}{n!}\,(x-c)^n.$$

Wegen $f'(c) = f''(c) = \cdots = f^{(n-1)}(c) = 0$ ist aber $T_{f,c}^{n-1}(x) = f(c)$, und damit also

$$f(x) - f(c) = \frac{f^{(n)}(x')}{n!}\,(x-c)^n. \qquad (*)$$

Da mit x auch x' in $(c-\delta, c+\delta)$ liegt, ist $f^{(n)}(x')$ in jedem Fall positiv. Also ist der Term $(*)$ für $x \in (c-\delta, c+\delta)\backslash\{c\}$

- immer größer als 0 falls n gerade ist; in diesem Fall hat f dann also ein isoliertes lokales Minimum in c;
- größer als 0 für $x > c$ und kleiner als Null für $x < c$ wenn n ungerade ist; in diesem Fall hat f also kein lokales Extremum in c. □

Bemerkung 11.19. Anschaulich kann man die Idee von Satz 11.18 kurz so zusammenfassen: Es sei f eine Funktion, von der wir an einer Stelle c wissen wollen, ob ein lokales Extremum vorliegt. Ist nun die n-te Ableitung von f die erste, die am Punkt c nicht verschwindet, so enthält das Taylor-Polynom $T_{f,c}^n$ nur den konstanten Term und den vom Grad n – und damit sagt uns die Idee der Taylor-Näherung, dass in der Nähe von c

$$f(x) \approx T_{f,c}^n(x) = f(c) + \frac{f^{(n)}(c)}{n!}\,(x-c)^n$$

gelten sollte. Da der Ausdruck auf der rechten Seite eine einfache Potenzfunktion ist, sieht man ihm aber natürlich sofort sein Verhalten um den Punkt c herum an: Für gerades n gibt es je nach Vorzeichen von $f^{(n)}(c)$ ein isoliertes lokales Minimum oder Maximum, und für ungerades n kein lokales Extremum. Mit dieser Idee lässt sich übrigens auch die Aussage des Satzes sehr leicht merken!

Aufgabe 11.20. Zeige, dass für jede zweimal stetig differenzierbare Funktion $f\colon \mathbb{R} \to \mathbb{R}$ die folgenden drei Bedingungen äquivalent sind:

(a) Für alle $x \in \mathbb{R}$ gilt $f''(x) \geq 0$.

(b) Für alle $x, y \in \mathbb{R}$ gilt $f(y) \geq f(x) + f'(x)\cdot(y-x)$.

(c) Für alle $x, y \in \mathbb{R}$ und $\lambda \in [0,1]$ gilt $(1-\lambda)f(x) + \lambda f(y) \geq f((1-\lambda)x + \lambda y)$.

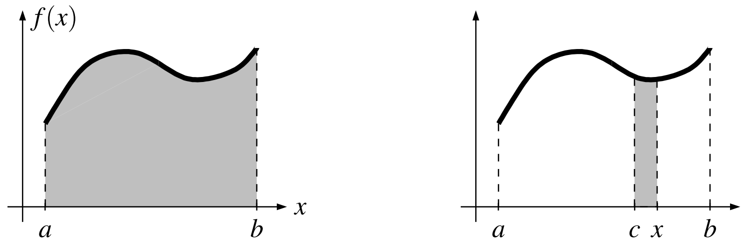
Eine Funktion, die diese Bedingungen erfüllt, heißt *konvex*. Was bedeuten die drei Bedingungen anschaulich?

12. Integralrechnung

Als Abschluss der Analysis in einer Veränderlichen wollen wir nach der Differentiation nun noch die Integration betrachten. Wie auch schon im letzten Kapitel wollen wir uns dabei auf den reellen Fall beschränken, da sich die Integralrechnung über $\mathbb{C}$ ganz anders verhält. In der Tat sind komplexe Integrale (oder allgemein die komplexe Analysis) der wesentliche Inhalt der Vorlesung „Einführung in die Funktionentheorie“, die ihr im zweiten Studienjahr hören könnt.

Die Integralrechnung kann man auf zweierlei Arten motivieren. Ist $f\colon [a,b] \to \mathbb{R}$ eine reelle Funktion, so können wir die folgenden beiden Fragestellungen betrachten:

- (Flächenberechnung) Wie groß ist die Fläche, die unter dem Graphen von f liegt (im Bild unten links grau eingezeichnet) – oder allgemeiner, wie kann man den Flächeninhalt gekrümmter Flächen berechnen?



- (Umkehrung der Differentiation) Gibt es eine differenzierbare Funktion $F\colon [a,b] \to \mathbb{R}$, deren Ableitung F' gleich f ist – und wenn ja, wie können wir ein solches F bestimmen? Diese Frage hat oft auch eine anschauliche Bedeutung: Beschreibt eine Funktion z. B. die Position eines Gegenstandes in Abhängigkeit von der Zeit, so ist die Ableitung dieser Funktion, also die lokale Positionsänderung pro Zeiteinheit, natürlich einfach die Geschwindigkeit des Gegenstandes. Wenn wir von der Ableitung auf die ursprüngliche Funktion zurück schließen wollen, möchten wir anschaulich also aus der Kenntnis der Geschwindigkeit zu jedem Zeitpunkt die von dem Gegenstand zurückgelegte Wegstrecke berechnen können.

Es ist leicht einzusehen, dass diese beiden Probleme sehr eng miteinander zusammenhängen: Bezeichnen wir für $c \in [a,b]$ mit $F(c)$ die Fläche, die über dem Intervall $[a,c]$ unter dem Graphen von f liegt, so ist $F(x) - F(c)$ für $x \in [a,b]$ natürlich gerade die Fläche unter f zwischen c und x (im Bild oben rechts grau eingezeichnet). Für x nahe bei c ist dies näherungsweise eine Rechteckfläche der Breite $x - c$ und Höhe $f(c)$, d. h. es ist

$$F(x) - F(c) \approx (x-c) \cdot f(c), \qquad \text{und damit} \qquad \frac{F(x)-F(c)}{x-c} \approx f(c).$$

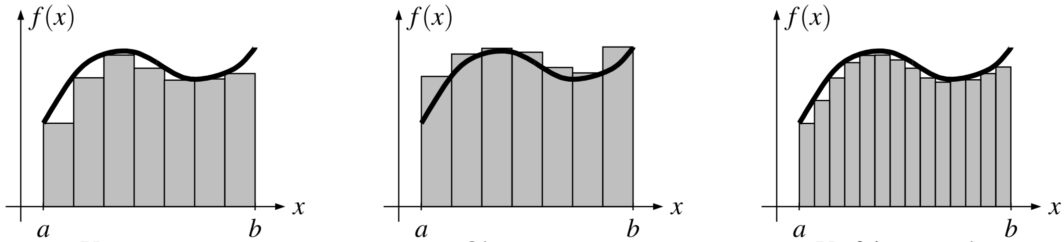
Im Grenzfall $x \to c$ sollte also $F' = f$ gelten, d. h. das Problem der Flächenberechnung unter dem Graphen einer Funktion sollte automatisch auch zur Umkehrung der Differentiation führen.

Wir werden uns im Folgenden zunächst in Abschnitt 12.A mit dem ersten Problem der Flächenberechnung beschäftigen, und daraufhin dann in Abschnitt 12.B den Zusammenhang zur Umkehrung der Differentiation herstellen.

12.A Das Riemann-Integral

Um den Flächeninhalt unter dem Graphen einer Funktion $f\colon [a,b] \to \mathbb{R}$ untersuchen zu können, müssen wir natürlich zunächst erst einmal mit einer exakten Definition dieses Konzepts beginnen.

Die Idee hierfür ist einfach: Wir zerlegen das Intervall $[a,b]$ in viele kleine Teilintervalle, und approximieren die Fläche unter dem Graphen von f durch Rechteckflächen über diesen Teilintervallen, indem wir wie im Bild unten als Höhe der Rechtecke einmal das Minimum und einmal das Maximum von f auf den betrachteten Teilintervallen wählen. Auf diese Art erhalten wir leicht zu berechnende Flächen, die im Fall des Minimums etwas kleiner und im Fall des Maximums etwas größer als die gesuchte Fläche sind. Wenn wir die Zerlegung in die Teilintervalle immer feiner machen (wie z. B. im Bild unten rechts), sollten diese Flächen dann von unten bzw. oben gegen den gesuchten Flächeninhalt unter dem Graphen von f konvergieren.



Untersumme — Obersumme — Verfeinerung der Untersumme

Wir wollen diese Idee nun mathematisch exakt definieren. Um die Theorie möglichst allgemein zu halten, wollen wir uns dabei nicht auf stetige Funktionen beschränken. Dies heißt natürlich, dass f auf den betrachteten Teilintervallen nicht mehr notwendig ein Minimum und Maximum hat (siehe Satz 8.24), sondern dass wir im Allgemeinen nur ein Infimum und Supremum erhalten – und das auch nur dann, wenn wir voraussetzen, dass f beschränkt ist.

Definition 12.1 (Zerlegungen, Unter- und Obersummen)**.** Es sei $f\colon [a,b] \to \mathbb{R}$ eine beschränkte Funktion.

(a) Eine endliche Teilmenge $Z = \{x_0, x_1, \dots, x_n\}$ von Punkten in $[a,b]$ mit $a,b \in Z$ bezeichnen wir als eine **Zerlegung** des Intervalls $[a,b]$. Wir vereinbaren, dass wir in dieser Schreibweise die $x_0, \dots, x_n$ immer so anordnen wollen, dass $a = x_0 < x_1 < \dots < x_n = b$ gilt. Sind Z, Z' zwei Zerlegungen von $[a,b]$ mit $Z \subset Z'$, so nennen wir Z' eine **Verfeinerung** von Z.

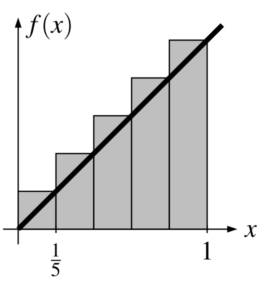
(b) Ist $Z = \{x_0, \dots, x_n\}$ eine Zerlegung von $[a,b]$, so heißt

$$\mathrm{US}(f,Z) := \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \inf f([x_{i-1}, x_i]) \qquad \text{die } \textbf{Untersumme}\text{, und analog}$$

$$\mathrm{OS}(f,Z) := \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \sup f([x_{i-1}, x_i]) \qquad \text{die } \textbf{Obersumme}$$

von f bezüglich Z. 26

Beispiel 12.2. Wir betrachten die Funktion $f\colon [0,1] \to \mathbb{R}$, $x \mapsto x$, und für gegebenes $n \in \mathbb{N}_{>0}$ die Zerlegung $Z_n = \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$. Natürlich ist das Supremum von f auf einem Teilintervall $\left[\frac{i-1}{n}, \frac{i}{n}\right]$ genau der Funktionswert $\frac{i}{n}$ an der rechten Intervallgrenze, und damit ist die Obersumme von f bezüglich Z_n (also die für den Fall $n = 5$ im Bild rechts eingezeichnete graue Fläche) gleich



$$\mathrm{OS}(f,Z_n) = \sum_{i=1}^{n} \left(\frac{i}{n} - \frac{i-1}{n}\right) \cdot \frac{i}{n} = \frac{1}{n^2} \sum_{i=1}^{n} i \overset{3.13}{=} \frac{1}{n^2} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2n}.$$

Analog müssen wir für die Untersumme jeweils den Funktionswert $\frac{i-1}{n}$ an der linken Intervallgrenze nehmen, und erhalten

$$\mathrm{US}(f,Z_n) = \sum_{i=1}^{n} \left(\frac{i}{n} - \frac{i-1}{n}\right) \cdot \frac{i-1}{n} = \frac{1}{n^2} \sum_{i=1}^{n} (i-1) = \frac{1}{n^2} \sum_{i=0}^{n-1} i = \frac{1}{n^2} \cdot \frac{(n-1)n}{2} = \frac{n-1}{2n}.$$

Lemma 12.3 (Eigenschaften von Unter- und Obersummen)**.** *Es seien $f\colon [a,b] \to \mathbb{R}$ eine beschränkte Funktion und Z, Z' zwei Zerlegungen von $[a,b]$. Dann gilt:*

(a) *Ist Z' eine Verfeinerung von Z, so ist $\operatorname{US}(f,Z') \geq \operatorname{US}(f,Z)$ und $\operatorname{OS}(f,Z') \leq \operatorname{OS}(f,Z)$.*

(b) $\operatorname{US}(f,Z) \leq \operatorname{OS}(f,Z')$.

Beweis.

(a) Da jede Verfeinerung von Z durch endliches Hinzufügen von weiteren Unterteilungspunkten entsteht, genügt es, den Fall zu betrachten, dass Z' durch Hinzufügen eines weiteren Punktes aus Z entsteht, also dass $Z = \{x_0, \dots, x_n\}$ und $Z' = \{x_0, x_1, \dots, x_{k-1}, x', x_k, \dots, x_n\}$ ist. Nach Definition ist dann

$$\begin{aligned}\operatorname{US}(f,Z') = &\sum_{i \neq k} (x_i - x_{i-1}) \cdot \inf f([x_{i-1}, x_i]) \\ &+ (x' - x_{k-1}) \cdot \inf f([x_{k-1}, x']) + (x_k - x') \cdot \inf f([x', x_k]).\end{aligned}$$

In dieser Summe sind nun die beiden Infima in der zweiten Zeile größer oder gleich dem Infimum der größeren Menge $f([x_{k-1}, x_k])$. Also erhalten wir wie gewünscht

$$\operatorname{US}(f,Z') \geq \sum_{i \neq k} (x_i - x_{i-1}) \cdot \inf f([x_{i-1}, x_i]) + (x_k - x_{k-1}) \cdot \inf f([x_{k-1}, x_k]) = \operatorname{US}(f,Z).$$

Die Aussage über die Obersumme beweist man natürlich analog.

(b) Da $Z \cup Z'$ eine gemeinsame Verfeinerung von Z und Z' ist, erhalten wir mit (a)

$$\operatorname{US}(f,Z) \leq \operatorname{US}(f, Z \cup Z') \leq \operatorname{OS}(f, Z \cup Z') \leq \operatorname{OS}(f,Z'),$$

wobei die mittlere Ungleichung gilt, weil das Infimum einer Menge immer kleiner oder gleich dem Supremum ist. □

Aufgabe 12.4. Es seien $f, g\colon [a,b] \to \mathbb{R}$ zwei beschränkte Funktionen, Z eine Zerlegung von $[a,b]$ und $c \in \mathbb{R}_{\geq 0}$. Man zeige:

(a) $\operatorname{OS}(f+g, Z) \leq \operatorname{OS}(f,Z) + \operatorname{OS}(g,Z)$;

(b) $\operatorname{OS}(cf, Z) = c \cdot \operatorname{OS}(f,Z)$;

(c) $\operatorname{OS}(|f|, Z) - \operatorname{US}(|f|, Z) \leq \operatorname{OS}(f,Z) - \operatorname{US}(f,Z)$.

Lemma 12.3 (b) besagt insbesondere, dass *jede* Obersumme eine obere Schranke für *alle* Untersummen ist. Die Menge aller Untersummen ist also nach oben beschränkt. Wir können damit das Supremum aller Untersummen (und genauso das Infimum aller Obersummen) bilden:

Definition 12.5 (Unter- und Oberintegral)**.** Es sei $f\colon [a,b] \to \mathbb{R}$ beschränkt. Dann heißt

$\operatorname{UI}(f) := \sup\{\operatorname{US}(f,Z) : Z \text{ Zerlegung von } [a,b]\}$ das **Unterintegral**, und analog

$\operatorname{OI}(f) := \inf\{\operatorname{OS}(f,Z) : Z \text{ Zerlegung von } [a,b]\}$ das **Oberintegral**

von f.

Anschaulich bedeutet dies im Fall des Unterintegrals einfach, dass wir – wie in der Einleitung zu diesem Abschnitt erklärt – versuchen, die Untersummen (durch fortgesetztes Verfeinern der Zerlegungen) möglichst groß zu machen, so dass wir uns letztlich immer mehr dem eigentlich gesuchten Flächeninhalt unter dem Graphen von f nähern. Das Supremum dieser Untersummen, also das Unterintegral, sollte demnach bereits der gesuchte Flächeninhalt unter dem Graphen von f sein. Das gleiche gilt natürlich auch für das Oberintegral, so dass wir insgesamt erwarten würden, dass Unter- und Oberintegral übereinstimmen und gleich dem gesuchten Flächeninhalt sind.

Leider ist dies unter den schwachen Voraussetzungen, die wir bisher an f gestellt haben, im Allgemeinen nicht der Fall, wie wir gleich in Beispiel 12.9 (d) sehen werden. Für beliebiges f erhalten wir zunächst nur die folgende Ungleichung.

Lemma 12.6. *Für jede beschränkte Funktion $f\colon [a,b] \to \mathbb{R}$ gilt $\operatorname{UI}(f) \leq \operatorname{OI}(f)$.*

Beweis. Angenommen, es wäre $\mathrm{UI}(f) > \mathrm{OI}(f)$. Wir wählen dann eine beliebige Zahl $c \in \mathbb{R}$ mit $\mathrm{UI}(f) > c > \mathrm{OI}(f)$. Da $\mathrm{UI}(f)$ die kleinste obere Schranke für die Untersummen ist, ist c keine obere Schranke mehr, d. h. es gibt eine Zerlegung Z von $[a,b]$ mit $\mathrm{US}(f,Z) > c$. Analog finden wir auch eine Zerlegung Z' von $[a,b]$ mit $\mathrm{OS}(f,Z') < c$. Dann ist im Widerspruch zu Lemma 12.3 (b) aber $\mathrm{US}(f,Z) > c > \mathrm{OS}(f,Z')$. □

Definition 12.7 (Integrierbarkeit)**.** Es sei $f\colon [a,b] \to \mathbb{R}$ beschränkt. Gilt dann $\mathrm{UI}(f) = \mathrm{OI}(f)$, so nennen wir f **(Riemann-)integrierbar**, und definieren das **Integral** von f als diesen Wert

$$\int_a^b f(x)\,dx := \mathrm{UI}(f) = \mathrm{OI}(f).$$

Bemerkung 12.8.

(a) Die Schreibweise $\int_a^b f(x)\,dx$ ist an die Differentialschreibweise aus Notation 10.14 angelehnt und soll andeuten, dass man sich das Integral entsprechend unserer Konstruktion anschaulich als eine „unendliche Summe kleiner Rechteckflächen“ vorstellen kann. Dabei steht das Integralzeichen $\int$ als stilisiertes S weiterhin für eine Summe, und die aufsummierten Rechtecke haben die Höhe $f(x)$ und Breite dx (siehe Notation 10.14), also die Fläche $f(x)\,dx$. Die Integrationsvariable x ist damit analog zur Laufvariablen in einer Summe und kann daher auch durch einen anderen Buchstaben ersetzt werden, darf aber natürlich nicht gleichzeitig noch für etwas anderes (z. B. die Ober- oder Untergrenze) verwendet werden: Ein Ausdruck z. B. der Form $\int_a^x f(x)\,dx$ ergibt keinen Sinn, genauso wenig wie eine Summe $\sum_{n=1}^n a_n$.

(b) Es gibt mehrere Arten, den Flächeninhalt unter dem Graphen einer Funktion zu definieren. Neben der hier behandelten Riemannschen Integrationstheorie über Unter- und Obersummen, die wohl die einfachste Herangehensweise ist, ist die „zweitwichtigste“ Möglichkeit das sogenannte *Lebesgue-Integral*, das zwar komplizierter zu definieren ist, dafür aber allgemeiner ist in dem Sinne, dass eine größere Klasse von Funktionen integrierbar wird. Wir werden in dieser Vorlesung jedoch nur die Riemannsche Integrationstheorie behandeln und daher statt von Riemann-Integrierbarkeit einfach immer nur von Integrierbarkeit reden. Die Lebesguesche Integrationstheorie könnt ihr im zweiten Studienjahr in der Vorlesung „Maß- und Integrationstheorie“ kennenlernen.

Beispiel 12.9.

(a) Ist $f(x) = c$ (mit $c \in \mathbb{R}$) eine konstante Funktion, so sind die Infima und Suprema von f auf allen Teilintervallen gleich c. Damit ist dann $\mathrm{US}(f,Z) = \mathrm{OS}(f,Z) = c\,(b-a)$ für alle Unterteilungen Z und somit auch $\mathrm{UI}(f) = \mathrm{OI}(f) = c\,(b-a)$. Also ist f integrierbar mit $\int_a^b f(x)\,dx = c\,(b-a)$ (was natürlich auch genau der Flächeninhalt für $x \in [a,b]$ unter dem Graphen von f ist).

(b) Wie in Beispiel 12.2 betrachten wir noch einmal die Funktion $f\colon [0,1] \to \mathbb{R}$, $x \mapsto x$ mit den Zerlegungen $Z_n = \{0, \frac{1}{n}, \dots, 1\}$. Da das Unterintegral nach Definition eine obere Schranke für alle Untersummen (und analog das Oberintegral eine untere Schranke für alle Obersummen) ist, folgt aus der Rechnung von Beispiel 12.2 sowie Lemma 12.6

$$\frac{n-1}{2n} = \mathrm{US}(f,Z_n) \le \mathrm{UI}(f) \le \mathrm{OI}(f) \le \mathrm{OS}(f,Z_n) = \frac{n+1}{2n},$$

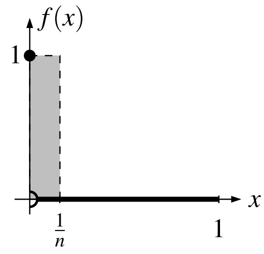
und damit durch Grenzwertbildung $n \to \infty$ nach Satz 5.24 (a)

$$\frac{1}{2} \le \mathrm{UI}(f) \le \mathrm{OI}(f) \le \frac{1}{2},$$

d. h. $\mathrm{UI}(f) = \mathrm{OI}(f) = \frac{1}{2}$. Also ist f integrierbar mit $\int_0^1 f(x)\,dx = \frac{1}{2}$ – was anschaulich ja auch die Dreiecksfläche unter dem Graphen von f ist.

(c) Es sei f die (unstetige) Funktion

$$f\colon [0,1] \to \mathbb{R},\ x \mapsto \begin{cases} 1 & \text{für } x = 0, \\ 0 & \text{für } x > 0. \end{cases}$$



Für die gleiche Zerlegung $Z_n = \{0, \frac{1}{n}, \frac{2}{n}, \dots, 1\}$ wie in (b) ist diesmal $\mathrm{US}(f, Z_n) = 0$ und $\mathrm{OS}(f, Z_n) = \frac{1}{n}$ (im Bild rechts ist die Obersumme eingezeichnet). Also folgt wieder

$$0 = \mathrm{US}(f, Z_n) \le \mathrm{UI}(f) \le \mathrm{OI}(f) \le \mathrm{OS}(f, Z_n) = \frac{1}{n},$$

und damit wie in (b) durch Grenzwertbildung für $n \to \infty$

$$0 \le \mathrm{UI}(f) \le \mathrm{OI}(f) \le 0 \quad \Rightarrow \quad \mathrm{UI}(f) = \mathrm{OI}(f) = 0.$$

Damit ist f integrierbar mit $\int_0^1 f(x)\,dx = 0$ – was auch anschaulich einleuchtend ist, denn unter dem einen Punkt, an dem der Funktionswert gleich 1 ist, liegt ja kein Flächeninhalt größer als Null.

(d) Wir betrachten die Funktion

$$f\colon [0,1] \to \mathbb{R},\ x \mapsto \begin{cases} 1 & \text{für } x \in \mathbb{Q}, \\ 0 & \text{für } x \notin \mathbb{Q}. \end{cases}.$$

Da in jedem Teilintervall von $[0,1]$ nach Aufgabe 5.36 sowohl rationale als auch irrationale Zahlen liegen, ist auf jedem solchen Teilintervall das Infimum von f gleich 0 und das Supremum gleich 1. Damit folgt $\mathrm{US}(f, Z) = 0$ und $\mathrm{OS}(f, Z) = 1$ für jede Zerlegung Z, d. h. es ist auch $\mathrm{UI}(f) = 0$ und $\mathrm{OI}(f) = 1$. Also ist f nicht integrierbar – mit unseren Definitionen können wir den Flächeninhalt unter dem Graphen von f nicht sinnvoll definieren.

Aufgabe 12.10. Zeige durch eine explizite Berechnung von Ober- und Untersummen, dass

$$\text{(a)} \quad \int_0^a e^x\,dx = e^a - 1 \qquad\qquad \text{(b)} \quad \int_0^a x^n\,dx = \frac{1}{n+1}\,a^{n+1}$$

für alle $a \in \mathbb{R}_{>0}$ und $n \in \mathbb{N}$. (Hinweis: Aufgabe 4.11 ist für (b) nützlich.)

Bevor wir die wichtigsten Eigenschaften integrierbarer Funktionen untersuchen, wollen wir zunächst noch ein einfaches Kriterium für die Integrierbarkeit beweisen, das implizit auch bereits in unseren Rechnungen von Beispiel 12.9 versteckt ist.

Lemma 12.11 (Riemannsches Integrabilitätskriterium). *Es sei $f\colon [a,b] \to \mathbb{R}$ eine beschränkte Funktion.*

(a) *f ist genau dann integrierbar, wenn es zu jedem $\varepsilon > 0$ eine Zerlegung Z von $[a,b]$ gibt mit $\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon$.*

(b) *f ist genau dann integrierbar mit Integral $\int_a^b f(x)\,dx = c$, wenn es zu jedem $\varepsilon > 0$ Zerlegungen Z und Z' von $[a,b]$ gibt mit $\mathrm{OS}(f,Z) < c + \varepsilon$ und $\mathrm{US}(f,Z') > c - \varepsilon$.*

Beweis.

„$\Rightarrow$“ Es sei f integrierbar mit $\int_a^b f(x)\,dx = \mathrm{UI}(f) = \mathrm{OI}(f) = c$. Da $\mathrm{OI}(f)$ nach Definition die größte untere Schranke für die Obersummen von f ist, ist $c + \frac{\varepsilon}{2}$ keine untere Schranke mehr, d. h. es gibt eine Zerlegung Z von $[a,b]$ mit $\mathrm{OS}(f,Z) < c + \frac{\varepsilon}{2}$. Analog gibt es eine Zerlegung Z' von $[a,b]$ mit $\mathrm{US}(f,Z') > c - \frac{\varepsilon}{2}$, was bereits (b) zeigt. Außerdem erfüllt die Zerlegung $Z \cup Z'$ dann auch die Eigenschaft von (a), denn nach Lemma 12.3 (a) ist

$$\mathrm{OS}(f, Z \cup Z') - \mathrm{US}(f, Z \cup Z') \le \mathrm{OS}(f,Z) - \mathrm{US}(f,Z') < \left(c + \frac{\varepsilon}{2}\right) - \left(c - \frac{\varepsilon}{2}\right) = \varepsilon.$$

„$\Leftarrow$“ Für Teil (a) haben wir zu jedem $\varepsilon > 0$ eine Zerlegung Z wie in der Voraussetzung, und damit

$$\mathrm{OI}(f) - \mathrm{UI}(f) \le \mathrm{OS}(f,Z) - \mathrm{US}(f,Z) < \varepsilon,$$

da das Ober- bzw. Unterintegral eine untere bzw. obere Schranke für die Ober- bzw. Untersummen sind. Nimmt man hier den Grenzwert für $\varepsilon \to 0$, so ergibt sich $\mathrm{OI}(f) - \mathrm{UI}(f) \le 0$, mit Lemma 12.6 also $\mathrm{OI}(f) = \mathrm{UI}(f)$. Damit ist f dann integrierbar.

Für Teil (b) gibt es stattdessen für jedes $\varepsilon > 0$ Zerlegungen Z und Z' von $[a,b]$ mit

$$c - \varepsilon < \mathrm{US}(f,Z') \le \mathrm{UI}(f) \le \mathrm{OI}(f) \le \mathrm{OS}(f,Z) < c + \varepsilon,$$

woraus im Grenzfall $\varepsilon \to 0$ die Ungleichungskette $c \le \mathrm{UI}(f) \le \mathrm{OI}(f) \le c$ folgt, d. h. f ist integrierbar mit Integral c. □

Als erste Anwendung dieses Kriteriums wollen wir nun untersuchen, wie die Integrierbarkeit mit der Stetigkeit einer Funktion zusammenhängt. Dazu haben wir in Beispiel 12.9 (c) schon gesehen, dass integrierbare Funktionen nicht notwendig stetig sein müssen. Die Umkehrung ist jedoch immer richtig:

Satz 12.12. *Ist $f\colon [a,b] \to \mathbb{R}$ stetig, so ist f auch integrierbar auf $[a,b]$.*

Beweis. Nach Satz 8.22 ist f beschränkt, so dass wir also die Begriffe dieses Kapitels anwenden können. Wir zeigen die Integrierbarkeit von f mit dem Kriterium aus Lemma 12.11 (a).

Es sei also $\varepsilon > 0$ gegeben. Da f auf dem abgeschlossenen Intervall $[a,b]$ stetig ist, ist f dort nach Satz 8.46 sogar gleichmäßig stetig. Es gibt also ein $\delta > 0$, so dass $|f(y) - f(z)| < \frac{\varepsilon}{b-a}$ für alle $y,z \in [a,b]$ mit $|y-z| < \delta$. Wir wählen nun eine Zerlegung $Z = \{x_0, \dots, x_n\}$ von $[a,b]$ mit $x_i - x_{i-1} < \delta$ für alle i, d. h. alle Teilintervalle sollen kürzer als δ sein. Dann gilt

$$\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \big(\sup f([x_{i-1},x_i]) - \inf f([x_{i-1},x_i])\big).$$

Als stetige Funktion nimmt f auf jedem Teilintervall $[x_{i-1},x_i]$ nach Satz 8.24 an einer Stelle y_i ein Maximum und an einer Stelle z_i ein Minimum an. Da y_i und z_i beide im Intervall $[x_{i-1},x_i]$ liegen, dessen Länge ja kleiner als δ ist, ist natürlich auch $|y_i - z_i| < \delta$ und damit $|f(y_i) - f(z_i)| < \frac{\varepsilon}{b-a}$ nach Wahl von δ. Wir können oben also weiterrechnen und erhalten

$$\mathrm{OS}(f,Z) - \mathrm{US}(f,Z) = \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot (f(y_i) - f(z_i)) < \sum_{i=1}^{n} (x_i - x_{i-1}) \cdot \frac{\varepsilon}{b-a} = \varepsilon,$$

woraus nun mit Lemma 12.11 (a) die Behauptung folgt. □ 27

Als Nächstes wollen wir die wichtigsten elementaren Eigenschaften von integrierbaren Funktionen herleiten.

Satz 12.13 (Eigenschaften des Integrals)**.** *Es seien $f,g\colon [a,b] \to \mathbb{R}$ integrierbare Funktionen. Dann gilt:*

(a) *$f+g$ ist ebenfalls integrierbar auf $[a,b]$, und es gilt*

$$\int_a^b (f(x)+g(x))\,dx = \int_a^b f(x)\,dx + \int_a^b g(x)\,dx.$$

(b) *Für alle $c \in \mathbb{R}$ ist $c\,f$ ebenfalls integrierbar auf $[a,b]$, und es gilt*

$$\int_a^b c\,f(x)\,dx = c \cdot \int_a^b f(x)\,dx.$$

(c) *Ist $f \le g$, d. h. $f(x) \le g(x)$ für alle $x \in [a,b]$, so ist $\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx$.*

(d) *$|f|$ ist ebenfalls integrierbar auf $[a,b]$, und es gilt die **Dreiecksungleichung***

$$\left|\int_a^b f(x)\,dx\right| \le \int_a^b |f(x)|\,dx.$$

Beweis. Wir verwenden das Riemannsche Integrabilitätskriterium aus Lemma 12.11.

(a) Da f und g integrierbar sind, gibt es für alle $\varepsilon > 0$ nach Lemma 12.11 (b) Zerlegungen Z und Z' von $[a,b]$ mit

$$\mathrm{OS}(f,Z) < \int_a^b f(x)\,dx + \frac{\varepsilon}{2} \qquad \text{und} \qquad \mathrm{OS}(g,Z') < \int_a^b g(x)\,dx + \frac{\varepsilon}{2}.$$

Nach Aufgabe 12.4 (a) und Lemma 12.3 (a) folgt daraus

$$\begin{aligned} \mathrm{OS}(f+g,Z\cup Z') &\le \mathrm{OS}(f,Z\cup Z') + \mathrm{OS}(g,Z\cup Z') \le \mathrm{OS}(f,Z) + \mathrm{OS}(g,Z') \\ &< \int_a^b f(x)\,dx + \int_a^b g(x)\,dx + \varepsilon. \end{aligned}$$

Analog finden wir für die Untersummen Zerlegungen $\tilde{Z}$ und $\tilde{Z}'$ mit

$$\mathrm{US}(f+g,\tilde{Z}\cup\tilde{Z}') > \int_a^b f(x)\,dx + \int_a^b g(x)\,dx - \varepsilon.$$

Die Behauptung folgt nun aus Lemma 12.11 (b) angewendet auf $f+g$.

(b) Für $c=0$ ist die Aussage trivial. Es sei nun $c>0$. Zu gegebenem $\varepsilon>0$ gibt es dann wieder eine Zerlegung Z von $[a,b]$ mit $\mathrm{OS}(f,Z) < \int_a^b f(x)\,dx + \frac{\varepsilon}{c}$. Damit folgt aus Aufgabe 12.4 (b) dann

$$\mathrm{OS}(cf,Z) = c\cdot\mathrm{OS}(f,Z) < c\cdot\int_a^b f(x)\,dx + \varepsilon.$$

Eine analoge Abschätzung bekommen wir natürlich auch wieder für die Untersummen. Damit folgt die Behauptung für $c>0$ aus Lemma 12.11 (b).

Für $c<0$ ergibt sich die Behauptung genauso aus der analog zu zeigenden Aussage $\mathrm{OS}(cf,Z) = c\cdot\mathrm{US}(f,Z)$.

(c) Aus $f\le g$ folgt sofort $\mathrm{OS}(f,Z)\le\mathrm{OS}(g,Z)$ für jede Zerlegung Z, und damit durch Übergang zum Infimum über alle Z auch $\int_a^b f(x)\,dx = \mathrm{OI}(f) \le \mathrm{OI}(g) = \int_a^b g(x)\,dx$.

(d) Wir zeigen zunächst die Integrierbarkeit von $|f|$. Dazu sei wieder $\varepsilon>0$ gegeben; nach Lemma 12.11 (a) können wir eine Zerlegung Z von $[a,b]$ wählen mit $\mathrm{OS}(f,Z)-\mathrm{US}(f,Z)<\varepsilon$. Mit Aufgabe 12.4 (c) folgt dann aber auch

$$\mathrm{OS}(|f|,Z)-\mathrm{US}(|f|,Z) \le \mathrm{OS}(f,Z)-\mathrm{US}(f,Z) < \varepsilon,$$

und damit ist $|f|$ nach Lemma 12.11 (a) integrierbar. Die Abschätzung des Integrals erhalten wir nun aus (c): Wegen $f\le|f|$ und $-f\le|f|$ ist sowohl

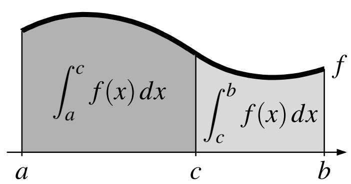
$$\int_a^b f(x)\,dx \le \int_a^b |f(x)|\,dx \qquad \text{als auch} \qquad -\int_a^b f(x)\,dx \overset{\text{(b)}}{=} \int_a^b -f(x)\,dx \le \int_a^b |f(x)|\,dx,$$

woraus sich die Behauptung ergibt, da $\left|\int_a^b f(x)\,dx\right|$ in jedem Fall eine dieser beiden linken Seiten ist. □

Eine weitere sehr anschauliche Eigenschaft von Integralen ist die sogenannte Additivität: für jede Zwischenstelle $c\in(a,b)$ ist die Fläche unter dem gesamten Graphen von $f\colon [a,b]\to\mathbb{R}$ gleich der Summe der Flächen von a bis c und von c bis b.

Satz 12.14 (Additivität des Integrals)**.** *Es seien $f\colon [a,b]\to\mathbb{R}$ eine Funktion und $c\in(a,b)$. Ist f dann sowohl auf $[a,c]$ als auch auf $[c,b]$ integrierbar, so auch auf $[a,b]$, und es gilt*

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$



Beweis. Der Beweis ist sehr ähnlich zu dem von Satz 12.13 (a). Es sei $\varepsilon > 0$ gegeben. Da f auf $[a,c]$ und $[c,b]$ integrierbar ist, gibt es Zerlegungen Z bzw. Z' dieser beiden Intervalle, so dass

$$\mathrm{OS}(f|_{[a,c]},Z) < \int_a^c f(x)\,dx + \frac{\varepsilon}{2} \qquad \text{und} \qquad \mathrm{OS}(f|_{[c,b]},Z') < \int_c^b f(x)\,dx + \frac{\varepsilon}{2}.$$

Die Obersumme von f bezüglich der Zerlegung $Z \cup Z'$ ist dann offensichtlich gerade die Summe dieser beiden Teilobersummen, d. h. wir haben

$$\mathrm{OS}(f,Z\cup Z') = \mathrm{OS}(f|_{[a,c]},Z) + \mathrm{OS}(f|_{[c,b]},Z') < \int_a^c f(x)\,dx + \int_c^b f(x)\,dx + \varepsilon,$$

und eine analoge Aussage auch genauso für die Untersummen. Damit folgt die Behauptung aus Lemma 12.11 (b). □

Notation 12.15 (Integrale mit vertauschten Grenzen)**.** Bisher haben wir Integrale $\int_a^b f(x)\,dx$ nur für $a \le b$ definiert. Ist hingegen $a > b$, so vereinbaren wir die Notation

$$\int_a^b f(x)\,dx := -\int_b^a f(x)\,dx, \qquad (*)$$

wenn f auf $[b,a]$ integrierbar ist. Dies hat den Vorteil, dass die Formel aus Satz 12.14 (im Fall der Integrierbarkeit) dann nicht nur für $a \le c \le b$, sondern für beliebige a,b,c gilt: Ist z. B. $a < b < c$, so ist nach Satz 12.14

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx,$$

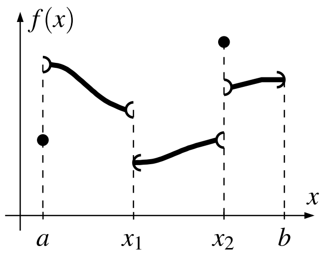
was (durch Subtraktion von $\int_b^c f(x)\,dx$ auf beiden Seiten) mit der Konvention $(*)$ wieder die gleiche Form

$$\int_a^b f(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx$$

wie in Satz 12.14 hat.

Beispiel 12.16 (Stückweise stetige Funktionen)**.** Es sei $f\colon [a,b] \to \mathbb{R}$ eine Funktion. Wir nennen einen Punkt $c \in [a,b]$ eine **Sprungstelle** von f, wenn die drei Zahlen

$$\lim_{\substack{x\to c\\ x<c}} f(x), \quad \lim_{\substack{x\to c\\ x>c}} f(x) \quad \text{und} \quad f(c)$$



existieren, aber nicht alle gleich sind (falls c einer der Randpunkte des Intervalls ist, gibt es den Grenzwert natürlich nur von einer der beiden Seiten). Man nennt f **stückweise stetig**, wenn f wie im Bild rechts stetig bis auf endlich viele Sprungstellen ist.

Eine solche stückweise stetige Funktion ist stets integrierbar:

(a) Es seien $a = x_0 < x_1 < \cdots < x_n = b$ die Sprungstellen und Randpunkte des Definitionsintervalls. Auf jedem Teilintervall $[x_{i-1},x_i]$ für $i = 1,\dots,n$ ist f dann eine stetige Funktion mit evtl. abgeänderten Funktionswerten an den Rändern, also die Summe aus einer stetigen Funktion und geeigneten Vielfachen der „Sprungfunktionen"

$$[x_{i-1},x_i] \to \mathbb{R},\ x \mapsto \begin{cases} 1 & \text{für } x = x_{i-1}, \\ 0 & \text{für } x > x_{i-1} \end{cases} \qquad \text{und} \qquad [x_{i-1},x_i] \to \mathbb{R},\ x \mapsto \begin{cases} 1 & \text{für } x = x_i, \\ 0 & \text{für } x < x_i. \end{cases}$$

Da eine stetige Funktion und diese Sprungfunktionen nach Satz 12.12 und Beispiel 12.9 (c) integrierbar sind, ist nach Satz 12.13 (a) und (b) auch $f|_{[x_{i-1},x_i]}$ integrierbar.

(b) Nach der Additivität aus Satz 12.14 ist f damit auch auf $[a,b]$ integrierbar.

Aufgabe 12.17. Zeige, dass die folgenden Funktionen integrierbar sind:

(a) eine beliebige monotone Funktion $f\colon [a,b] \to \mathbb{R}$;

(b) $f\colon [-1,1] \to \mathbb{R},\ x \mapsto \begin{cases} \sin\frac{1}{x} & \text{für } x \ne 0, \\ 0 & \text{für } x = 0; \end{cases}$

(c) $f\colon [0,1] \to \mathbb{R},\ x \mapsto \begin{cases} \frac{1}{q} & \text{für } x \in \mathbb{Q} \text{ mit gekürzter Darstellung } x = \frac{p}{q} \text{ für } p \in \mathbb{N} \text{ und } q \in \mathbb{N}_{>0}, \\ 0 & \text{für } x \in \mathbb{R}\backslash\mathbb{Q}. \end{cases}$

Aufgabe 12.18. Für eine Zerlegung $Z = \{x_0, \dots, x_k\}$ eines Intervalls $[a,b]$ definieren wir die *Feinheit* $l(Z)$ als den größten Abstand $\max\{x_i - x_{i-1} : i = 1, \dots, k\}$ zwischen zwei benachbarten Punkten von Z.

Es seien nun eine Folge (Z_n) von Zerlegungen $Z_n = \{x_{n,0}, \dots, x_{n,k_n}\}$ von $[a,b]$ mit $\lim\limits_{n\to\infty} l(Z_n) = 0$ sowie Zwischenpunkte $\xi_{n,i} \in [x_{n,i-1}, x_{n,i}]$ für $n \in \mathbb{N}$ und $i = 1, \dots, k_n$ gegeben. Zeige, dass dann für jede stetige Funktion $f\colon [a,b] \to \mathbb{R}$

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \sum_{i=1}^{k_n} (x_{n,i} - x_{n,i-1})\, f(\xi_{n,i})$$

gilt, also dass man das Integral statt mit Ober- und Untersummen auch mit einer beliebigen „Zwischensumme“ berechnen kann.

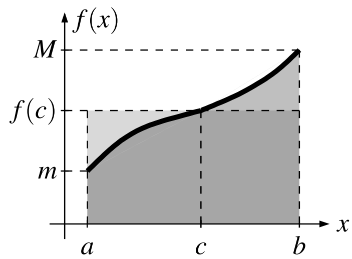
12.B Stammfunktionen

Unsere bisherigen Ergebnisse erlauben es uns zwar, von vielen Funktionen die Integrierbarkeit nachzuweisen, aber noch nicht, den Wert des Integrals dann auch explizit zu berechnen. Wie wir in der Einleitung dieses Kapitels schon motiviert haben, ist das zentrale Resultat für solche Berechnungen die Aussage, dass die Integration die Umkehrung der Differentiation ist. Um dies zu zeigen, benötigen wir zur Vorbereitung noch einen kleinen und sehr anschaulichen Hilfssatz.

Satz 12.19 (Mittelwertsatz der Integralrechnung). *Es sei $f\colon [a,b] \to \mathbb{R}$ eine stetige Funktion. Dann gibt es einen Punkt $c \in [a,b]$ mit*

$$\int_a^b f(x)\,dx = f(c) \cdot (b-a)$$

(d. h. die Fläche unter dem Graphen ist wie im Bild rechts gleich der Fläche eines Rechtecks, dessen Höhe ein Funktionswert $f(c)$ auf dem betrachteten Intervall ist).



Beweis. Nach Satz 8.24 nimmt f als stetige Funktion auf einem abgeschlossenen Intervall ein Maximum M und Minimum m an. Also folgt aus Beispiel 12.9 (a) und Satz 12.13 (c)

$$m\,(b-a) = \int_a^b m\,dx \le \int_a^b f(x)\,dx \le \int_a^b M\,dx = M\,(b-a),$$

und damit

$$m \le \frac{1}{b-a} \cdot \int_a^b f(x)\,dx \le M.$$

Da f stetig ist, gibt es nun nach dem Zwischenwertsatz 8.20 ein $c \in [a,b]$ mit $f(c) = \frac{1}{b-a} \cdot \int_a^b f(x)\,dx$ – was genau die Behauptung war. □

Bemerkung 12.20. Mit Notation 12.15 gilt die Gleichung $\int_a^b f(x)\,dx = f(c)\,(b-a)$ für ein c zwischen a und b auch für den Fall $a > b$: Anwenden des Mittelwertsatzes 12.19 auf das Intervall $[b,a]$ liefert dann zunächst $\int_b^a f(x)\,dx = f(c)\,(a-b)$, woraus wir aber durch Multiplikation mit -1 wieder die Form $\int_a^b f(x)\,dx = f(c)\,(b-a)$ erhalten können.

Satz 12.21 (Hauptsatz der Differential- und Integralrechnung). *Es sei $f\colon [a,b] \to \mathbb{R}$ eine stetige Funktion. Dann ist die Funktion*

$$F\colon [a,b] \to \mathbb{R},\ x \mapsto \int_a^x f(t)\,dt$$

(bei der wir also das Integral von f berechnen und dabei die Obergrenze als Variable nehmen) differenzierbar mit $F' = f$.

Beweis. Wir zeigen die Differenzierbarkeit von F in einem Punkt $c \in [a,b]$ mit dem Folgenkriterium. Es sei also $(x_n)_n$ eine beliebige Folge in $[a,b]\backslash\{c\}$ mit $x_n \to c$. Dann gilt für alle n

$$\begin{aligned} F(x_n) - F(c) &= \int_a^{x_n} f(t)\,dt - \int_a^c f(t)\,dt \\ &= \int_c^{x_n} f(t)\,dt && \text{(Satz 12.14 bzw. Notation 12.15)} \\ &= f(z_n)\,(x_n - c) && \text{(Satz 12.19 bzw. Bemerkung 12.20)} \end{aligned}$$

für ein z_n zwischen c und x_n. Beachte, dass wegen $x_n \to c$ auch $z_n \to c$ gilt, da z_n ja immer zwischen c und x_n liegt. Da f stetig ist, haben wir damit

$$\lim_{n\to\infty} \frac{F(x_n) - F(c)}{x_n - c} = \lim_{n\to\infty} f(z_n) = f(c)$$

nach dem Folgenkriterium für Stetigkeit aus Satz 8.11 (b). Wiederum nach dem Folgenkriterium gemäß Satz 8.11 (a) bedeutet dies nun aber gerade wie gewünscht

$$F'(c) = \lim_{x\to c} \frac{F(x) - F(c)}{x - c} = f(c). \qquad \square$$

Beispiel 12.22. Die Voraussetzung der Stetigkeit im Hauptsatz 12.21 ist wirklich notwendig: Betrachten wir z. B. noch einmal die unstetige Funktion

$$f\colon [0,1] \to \mathbb{R},\ x \mapsto \begin{cases} 1 & \text{für } x = 0, \\ 0 & \text{für } x > 0 \end{cases}$$

aus Beispiel 12.9 (c), so ist hier

$$F(x) = \int_0^x f(t)\,dt = 0 \quad \text{für alle } x \in [0,1].$$

Diese Funktion ist zwar differenzierbar, hat als Ableitung jedoch die Nullfunktion und nicht f.

Für die Integralberechnung benötigen wir also Funktionen, deren Ableitung die ursprünglich gegebene Funktion ist. Wir geben solchen Funktionen daher einen besonderen Namen. Wegen Beispiel 12.22 beschränken wir uns dabei auf stetige Funktionen.

Definition 12.23 (Stammfunktionen)**.** Es seien $D \subset \mathbb{R}$ ein Intervall und $f\colon D \to \mathbb{R}$ eine stetige Funktion. Dann heißt eine differenzierbare Funktion $F\colon D \to \mathbb{R}$ mit $F' = f$ eine **Stammfunktion** von f.

Folgerung 12.24 (Integralberechnung mit Stammfunktionen)**.** *Es sei $f\colon [a,b] \to \mathbb{R}$ eine stetige Funktion. Dann gilt:*

(a) *f besitzt eine Stammfunktion.*

(b) *Sind F und G zwei Stammfunktionen von f, so unterscheiden sich diese nur um eine additive Konstante, d. h. es gibt ein $c \in \mathbb{R}$ mit $F - G = c$.*

(c) *Ist F eine Stammfunktion von f, so gilt*

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

Beweis.

(a) folgt unmittelbar aus Satz 12.21: Die dort angegebene Funktion $x \mapsto \int_a^x f(t)\,dt$ ist eine Stammfunktion von f.

(b) Nach Voraussetzung ist $(F-G)' = F' - G' = f - f = 0$. Damit ist $F - G$ nach Folgerung 10.24 (c) konstant.

(c) Nach dem Hauptsatz 12.21 sind sowohl F als auch $x \mapsto \int_a^x f(t)\,dt$ Stammfunktionen von f. Also gibt es nach (b) eine Konstante $c \in \mathbb{R}$ mit

$$F(x) - \int_a^x f(t)\,dt = c$$

für alle $x \in [a,b]$. Einsetzen von $x = a$ liefert $F(a) = c$, und damit

$$F(x) - F(a) = \int_a^x f(t)\,dt.$$

Für $x = b$ ergibt sich nun die Behauptung. □

Notation 12.25 (Unbestimmte Integrale)**.** Man schreibt die Differenz $F(b) - F(a)$ in Folgerung 12.24 (c) oft auch als $\big[F(x)\big]_{x=a}^{b}$ oder $F(x)\big|_{x=a}^{b}$ (wenn die Integrationsvariable aus dem Zusammenhang klar ist, auch als $\big[F(x)\big]_a^b$ oder $F(x)\big|_a^b$), und die dortige Gleichung damit als

$$\int_a^b f(x)\,dx = \big[F(x)\big]_a^b.$$

Da dies für beliebige Integrationsgrenzen a und b gilt, vereinfacht man diese Notation oft noch weiter und schreibt gemäß Folgerung 12.24 einfach

$$\int f(x)\,dx = F(x), \tag{$*$}$$

für die Aussage, dass F eine Stammfunktion von f ist. Man bezeichnet dies dann auch als ein **unbestimmtes Integral**. Beachte aber, dass $(*)$ nur eine symbolische Schreibweise ist, die erst nach dem Einsetzen von Grenzen zu einer echten Gleichheit in $\mathbb{R}$ wird! Dies kann man alleine schon daran sehen, dass x von der Notation her auf der linken Seite eine Integrationsvariable ist, auf der rechten Seite aber wie eine freie Variable aussieht (siehe Bemerkung 12.8 (a)). Außerdem ist mit F z. B. auch $F + 1$ eine Stammfunktion von f, d. h. wir können sowohl

$$\int f(x)\,dx = F(x) \qquad \text{als auch} \qquad \int f(x)\,dx = F(x) + 1$$

schreiben – was natürlich sofort zum Widerspruch $F(x) = F(x) + 1$ führen würde, wenn man dies als echte Gleichungen von Funktionen betrachten dürfte.

Beispiel 12.26. Wir können nun viele Integrale konkret berechnen, indem wir vom Integranden eine Stammfunktion suchen:

(a) Ist $f(x) = x^a$ für ein $a \in \mathbb{R}\backslash\{-1\}$, so ist $F(x) = \frac{1}{a+1}x^{a+1}$ nach Beispiel 10.28 (d) eine Stammfunktion von f, d. h. mit Notation 12.25 gilt

$$\int x^a\,dx = \frac{1}{a+1}x^{a+1} \qquad \text{für } a \neq -1.$$

Konkret können wir damit z. B. das Integral aus Beispiel 12.9 (b) auch ohne komplizierte Berechnung von Ober- und Untersummen bestimmen: Es ist einfach

$$\int_0^1 x\,dx = \left[\frac{1}{2}x^2\right]_{x=0}^{1} = \frac{1}{2}\cdot 1^2 - \frac{1}{2}\cdot 0^2 = \frac{1}{2}.$$

(b) Für $a = -1$ ist die Formel aus (a) natürlich nicht anwendbar. Wir haben aber glücklicherweise mit dem Logarithmus schon eine Funktion kennengelernt, deren Ableitung nach Beispiel 10.28 (c) gleich $\frac{1}{x}$ ist: Es ist also

$$\int \frac{1}{x}\,dx = \log x$$

für Integrationsintervalle in $\mathbb{R}_{>0}$ (so dass der Logarithmus dort definiert ist). Falls das Integrationsintervall in $\mathbb{R}_{<0}$ liegt, können wir als Stammfunktion $\log(-x)$ nehmen, denn auch die Ableitung dieser Funktion ist ja gleich $\frac{1}{x}$. Insgesamt ist damit

$$\int \frac{1}{x}\,dx = \log|x|.$$

(c) Durch die Ableitungen der speziellen Funktionen, die wir in Beispiel 10.28 berechnet haben, sehen wir genauso z. B.

$$\int \mathrm{e}^x\,dx = \mathrm{e}^x, \qquad \int \cos x\,dx = \sin x \qquad \text{und} \qquad \int \sin x\,dx = -\cos x.$$

Nach der Einführung von Stammfunktionen wollen wir unseren Integralbegriff nun noch etwas erweitern. Bisher konnten wir Integrale nur auf abgeschlossenen Intervallen $[a,b]$ im Definitionsbereich des Integranden berechnen. Oft treten allerdings Fälle auf, in denen eine oder beide Integrationsgrenzen entweder nicht mehr im Definitionsbereich liegen oder aber gleich $\pm\infty$ sind. Auch in diesen Fällen kann man durch eine einfache Grenzwertbildung das Integral definieren.

Definition 12.27 (Uneigentliche Integrale)**.**

(a) Es sei $f\colon [a,b) \to \mathbb{R}$ eine reelle Funktion (wobei der Fall $b=\infty$ zugelassen ist). Wir nehmen weiterhin an, dass f auf jedem abgeschlossenen Intervall $[a,c]$ mit $a \le c < b$ integrierbar ist. Existiert dann der Grenzwert

$$\int_a^b f(x)\,dx := \lim_{\substack{c\to b\\ c<b}} \int_a^c f(x)\,dx \qquad \in \mathbb{R}\cup\{\pm\infty\},$$

so nennen wir ihn das **uneigentliche Integral** von f auf $[a,b)$. Liegt dieser Grenzwert zusätzlich in $\mathbb{R}$, so heißt das uneigentliche Integral $\int_a^b f(x)\,dx$ **konvergent**, andernfalls **divergent**. Analog definiert man uneigentliche Integrale im Fall $f\colon (a,b] \to \mathbb{R}$.

(b) Es sei nun $f\colon (a,b) \to \mathbb{R}$, wobei die Fälle $a=-\infty$ bzw. $b=\infty$ wieder zugelassen sind. Für ein $c \in (a,b)$ setzen wir dann

$$\int_a^b f(x)\,dx := \int_a^c f(x)\,dx + \int_c^b f(x)\,dx,$$

sofern die rechte Summe (von zwei uneigentlichen Integralen gemäß (a)) in $\mathbb{R}\cup\{\pm\infty\}$ existiert. Beachte, dass diese Summe dann wegen der Additivität des Integrals aus Satz 12.14 nicht von der Wahl des Zwischenpunktes c abhängt. Wie im Fall (a) spricht man auch hier von einem (beidseitig) uneigentlichen Integral bzw. von der Konvergenz oder Divergenz dieses Integrals.

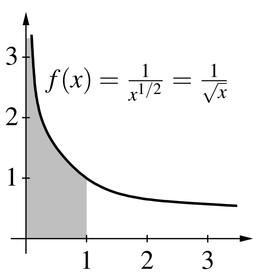
Beispiel 12.28.

(a) Für $a \in \mathbb{R}_{>0}\backslash\{1\}$ ist

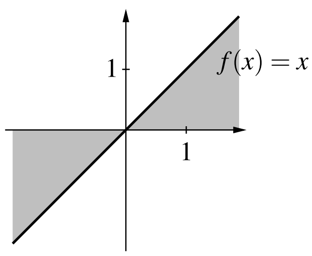
$$\int_0^1 \frac{1}{x^a}\,dx = \lim_{\substack{c\to 0\\ c>0}} \int_c^1 \frac{1}{x^a}\,dx = \lim_{\substack{c\to 0\\ c>0}} \left[\frac{1}{1-a}x^{1-a}\right]_c^1 = \frac{1}{1-a} \lim_{\substack{c\to 0\\ c>0}} (1-c^{1-a})$$

$$= \begin{cases} \frac{1}{1-a} & \text{für } a<1, \\ \infty & \text{für } a>1. \end{cases}$$

Das uneigentliche Integral konvergiert also genau für $a<1$. Anschaulich bedeutet dies, dass die Fläche von 0 bis 1 unter dem Graphen von $x \mapsto \frac{1}{x^a}$ in diesem Fall (wie im Bild unten links für $a=\frac{1}{2}$ dargestellt) endlich ist, obwohl sie nach oben eine unendliche Ausdehnung hat.



(a)



(b)

(b) Das uneigentliche Integral der Identität $f(x) = x$ auf $(-\infty, \infty)$ existiert nicht, denn bei der Wahl des Zwischenpunktes 0 erhalten wir den unbestimmten Ausdruck

$$\begin{aligned}\int_{-\infty}^{\infty} x\,dx &= \int_{-\infty}^{0} x\,dx + \int_{0}^{\infty} x\,dx = \lim_{c\to-\infty}\int_{c}^{0} x\,dx + \lim_{c\to\infty}\int_{0}^{c} x\,dx \\ &= \lim_{c\to-\infty}\left[\frac{1}{2}x^2\right]_c^0 + \lim_{c\to\infty}\left[\frac{1}{2}x^2\right]_0^c = \frac{1}{2}\lim_{c\to-\infty} -c^2 + \frac{1}{2}\lim_{c\to\infty} c^2 \\ &= -\infty + \infty.\end{aligned}$$

Auch anschaulich ist im Bild oben rechts ersichtlich, dass sich dieser Flächeninhalt aus einer unendlich großen negativen und positiven Fläche zusammensetzt. Beachte, dass wir nicht das gleiche Ergebnis erhalten hätten, wenn wir das uneigentliche Integral symmetrisch um die vertikale Achse als

$$\lim_{c\to\infty}\int_{-c}^{c} x\,dx = \lim_{c\to\infty}\left[\frac{1}{2}x^2\right]_{-c}^{c} = \frac{1}{2}\lim_{c\to\infty}(c^2 - c^2) = 0$$

28 definiert hätten!

12.C Integrationsregeln

In Abschnitt 12.B haben wir alle Stammfunktionen zur Berechnung von Integralen letztlich „durch Zufall“ gefunden – also weil wir uns einfach an eine Funktion erinnern konnten, deren Ableitung wir schon einmal berechnet haben und bei der für diese Ableitung dann die gegebene Funktion herauskam. Daher müssen wir uns jetzt natürlich fragen, wie man Stammfunktionen berechnen kann, wenn man nicht gerade zufällig eine solche sieht. Gibt es analog zur Berechnung von Ableitungen auch Regeln, mit denen man, wenn man die Stammfunktionen einiger spezieller Funktionen kennt, auch die Stammfunktionen z. B. ihrer Produkte, Quotienten oder Verkettungen berechnen kann?

Leider gibt es keine solchen universellen Regeln. Dies ist auch der Grund dafür, dass in mathematischen Formelsammlungen oft seitenweise Tabellen von Stammfunktionen stehen, während man für das Differenzieren aufgrund der Produkt-, Quotienten- und Kettenregel keine derartigen Tabellen benötigt. Es gibt jedoch auch für die Integration ein paar Regeln, mit denen man Integrale oft berechnen kann – nur ist es je nach der betrachteten Funktion mehr oder weniger schwierig (oder evtl. sogar unmöglich), einen Weg zu finden, um mit diesen Regeln ans Ziel zu kommen.

Wir wollen nun die wichtigsten derartigen Regeln behandeln. Die erste ist im wesentlichen nur die „umgekehrte Richtung“ der Produktregel der Differentiation:

Satz 12.29 (**Partielle Integration** bzw. **Produktintegration**). *Es seien $u, v\colon [a,b] \to \mathbb{R}$ stetig differenzierbare Funktionen. Dann gilt*

$$\int_a^b u'(x)v(x)\,dx = \big[u(x)v(x)\big]_a^b - \int_a^b u(x)v'(x)\,dx$$

(bzw. als unbestimmtes Integral $\int u'(x)v(x)\,dx = u(x)v(x) - \int u(x)v'(x)\,dx$).

Beweis. Nach der Produktregel aus Satz 10.8 (b) ist uv eine Stammfunktion von $(uv)' = u'v + uv'$. Also ist $\int_a^b (u'(x)v(x) + u(x)v'(x))\,dx = \big[u(x)v(x)\big]_a^b$. Die Behauptung folgt dann durch Subtraktion von $\int_a^b u(x)v'(x)\,dx$ nach Satz 12.13. □

Beispiel 12.30. Die Regel aus Satz 12.29 nennt sich „partielle Integration“, weil bei der Berechnung des Integrals auf der linken Seite mit Hilfe der rechten neben einem „ausintegrierten Anteil“ noch ein anderes Integral übrig bleibt – nämlich eines, bei dem wir von einem Faktor des ursprünglichen Integrals die Ableitung und vom anderen eine Stammfunktion gebildet haben. Die Anwendung dieser Regel ist also vor allem dann sinnvoll, wenn dieses neue Integral bereits bekannt oder zumindest einfacher als das ursprüngliche ist. Hier sind zwei Beispiele dafür.

(a) Zur Berechnung von $\int x \cos x \, dx$ setzen wir (mit den Notationen von Satz 12.29)

$$\begin{aligned} u(x) &= \sin x \qquad & u'(x) &= \cos x \\ v(x) &= x & v'(x) &= 1 \end{aligned}$$

und erhalten

$$\int x \cos x \, dx = x \sin x - \int \sin x \, dx = x \sin x + \cos x.$$

Wir haben bei der Anwendung der partiellen Integration also den Faktor x differenziert und den Faktor $\cos x$ integriert. Möchte man dies in der Rechnung deutlich machen (und die Funktionen u, u', v, v' nicht explizit hinschreiben), so notiert man dies auch oft als

$$\int \underset{\downarrow}{x} \underset{\uparrow}{\cos x} \, dx = x \sin x - \int \sin x \, dx = x \sin x + \cos x.$$

Beachte, dass die umgekehrte Wahl hier nicht zum Ziel geführt hätte: Die Rechnung

$$\int \underset{\uparrow}{x} \underset{\downarrow}{\cos x} \, dx = \frac{x^2}{2} \cos x + \int \frac{x^2}{2} \sin x \, dx$$

ist zwar korrekt, aber das neue Integral ist hier komplizierter als das ursprüngliche.

(b) Das Integral $\int \log x \, dx$ lässt sich mit einem Trick ebenfalls durch partielle Integration berechnen:

$$\int \log x \, dx = \int \underset{\uparrow}{1} \cdot \underset{\downarrow}{\log x} \, dx = x \log x - \int x \cdot \frac{1}{x} \, dx = x \log x - x.$$

Dieser Trick funktioniert hier, weil aus der (komplizierten) Logarithmusfunktion beim Ableiten die sehr viel einfachere Funktion $\frac{1}{x}$ entsteht. Auf die gleiche Art kann man übrigens auch die Integrale der Arkusfunktionen aus Definition 9.25 berechnen, da auch diese bei der Differentiation sehr viel einfacher werden (siehe Beispiel 10.28 (c)).

Die zweite wichtige Integrationsregel ergibt sich analog aus der Kettenregel der Differentiation.

Satz 12.31 (Substitutionsregel). *Es seien $f\colon [a,b] \to \mathbb{R}$ eine stetig differenzierbare Funktion und $g\colon D \to \mathbb{R}$ eine stetige Funktion auf einer Teilmenge $D \subset \mathbb{R}$ mit $f([a,b]) \subset D$. Dann gilt*

$$\int_a^b g(f(x)) \, f'(x) \, dx = \int_{f(a)}^{f(b)} g(y) \, dy.$$

Beweis. Nach Folgerung 8.25 ist $f([a,b])$ ein abgeschlossenes Intervall, und damit hat g nach Folgerung 12.24 (a) dort eine Stammfunktion G. Die Kettenregel aus Satz 10.10 liefert dann $(G \circ f)'(x) = g(f(x)) \, f'(x)$. Damit ist die linke Seite der zu beweisenden Gleichung

$$\int_a^b (G \circ f)'(x) \, dx = \big[G(f(x))\big]_{x=a}^{b} = G(f(b)) - G(f(a)),$$

und die rechte Seite ebenfalls

$$\int_{f(a)}^{f(b)} g(y) \, dy = \big[G(y)\big]_{y=f(a)}^{f(b)} = G(f(b)) - G(f(a)). \qquad \square$$

Bemerkung 12.32. Die Substitutionsregel nimmt in der Differentialschreibweise aus Notation 10.14 eine besonders einfache Form an: Setzen wir $y = f(x)$ und damit $\frac{dy}{dx} = f'(x)$, und bezeichnen wir die Integrationsgrenzen mit $x_1 = a$ und $x_2 = b$ bzw. $y_1 = f(a)$ und $y_2 = f(b)$, so schreibt sich die Substitutionsregel als

$$\int_{x_1}^{x_2} g(y) \, \frac{dy}{dx} \, dx = \int_{y_1}^{y_2} g(y) \, dy,$$

oder analog zu Notation 12.25 einfach als

$$\int g(y) \, \frac{dy}{dx} \, dx = \int g(y) \, dy,$$

wenn man darauf achtet, dass die Grenzen passend zur Integrationsvariablen gewählt werden. Die Regel sieht dann also einfach wie ein „formales Erweitern mit dx“ aus.

Beispiel 12.33. Die Substitutionsregel bietet sich natürlich immer dann an, wenn die zu integrierende Funktion eine Verkettung von zwei anderen Funktionen ist oder enthält – und insbesondere dann, wenn die Ableitung der inneren Funktion zusätzlich auch noch als Faktor im Integranden steht.

(a) Beim Integral $\int x\mathrm{e}^{x^2}\,dx$ stellen wir fest, dass sich im Integranden eine verkettete Funktion e^{x^2} befindet, und dass die Ableitung $2x$ der inneren Funktion x^2 auch (bis auf die Konstante 2) zusätzlich noch als Faktor im Integranden steht. Wir substituieren also $y = x^2$, so dass $\frac{dy}{dx} = 2x$ in der Notation von Bemerkung 12.32 gilt. Damit folgt also

$$\int x\mathrm{e}^{x^2}\,dx = \frac{1}{2}\int \frac{dy}{dx}\,\mathrm{e}^y\,dx \overset{12.31}{=} \frac{1}{2}\int \mathrm{e}^y\,dy = \frac{1}{2}\mathrm{e}^y = \frac{1}{2}\mathrm{e}^{x^2}.$$

Im Fall eines bestimmten Integrals hätten wir bei der Anwendung von Satz 12.31 die Grenzen mitsubstituieren müssen:

$$\int_a^b x\mathrm{e}^{x^2}\,dx = \frac{1}{2}\int_a^b \frac{dy}{dx}\,\mathrm{e}^y\,dx \overset{12.31}{=} \frac{1}{2}\int_{a^2}^{b^2} \mathrm{e}^y\,dy = \frac{1}{2}\left[\mathrm{e}^y\right]_{y=a^2}^{b^2} = \frac{1}{2}\left(\mathrm{e}^{b^2} - \mathrm{e}^{a^2}\right).$$

Beachte, dass der Faktor x im Integranden bei diesem Beispiel ganz wesentlich dafür war, dass die Substitutionsregel zum Ziel geführt hat: Ohne diesen Faktor hätten wir mit derselben Substitution

$$\int \mathrm{e}^{x^2}\,dx = \int \frac{1}{2x}\cdot\frac{dy}{dx}\,\mathrm{e}^y\,dx = \frac{1}{2}\int \frac{\mathrm{e}^y}{\sqrt{y}}\,dy$$

erhalten – was zwar auch richtig ist, aber nicht weiter hilft, weil das neue Integral auch nicht einfacher als das ursprüngliche zu berechnen ist. In der Tat kann man zeigen, dass sich die Stammfunktion von e^{x^2} nicht durch die uns bisher bekannten „speziellen Funktionen“ ausdrücken lässt.

(b) Besonders einfach wird die Substitutionsregel im Fall der sogenannten *linearen Substitution*: Ist f eine beliebige stetige Funktion, deren Stammfunktion F wir kennen, so können wir damit immer auch die Stammfunktion von $f(ax+b)$ mit $a,b \in \mathbb{R}$ und $a \neq 0$ bestimmen, da die innere Ableitung hier eine Konstante ist: Substituieren wir $y = ax+b$, so ergibt sich wegen $\frac{dy}{dx} = a$

$$\int f(ax+b)\,dx = \frac{1}{a}\int f(y)\frac{dy}{dx}\,dx \overset{12.31}{=} \frac{1}{a}\int f(y)\,dy = \frac{1}{a}F(y) = \frac{1}{a}F(ax+b).$$

Konkret ist also z. B.

$$\int \frac{1}{2x+3}\,dx = \frac{1}{2}\log|2x+3|,$$

da $x \mapsto \log|x|$ nach Beispiel 12.26 (b) eine Stammfunktion von $x \mapsto \frac{1}{x}$ ist.

Wir hatten in Beispiel 12.33 (a) bereits erwähnt, dass sich die Stammfunktionen von Funktionen, die aus unseren speziellen Funktionen zusammengesetzt sind, manchmal nicht wieder auf diese Art schreiben lassen. Für viele Klassen von Funktionen ist dies aber doch der Fall – z. B. für rationale Funktionen der Form $\frac{p}{q}$ für zwei Polynome p und q. Wir wollen dies hier nun zeigen, der Einfachheit halber allerdings nur für den Fall, dass das Nennerpolynom q in verschiedene Linearfaktoren zerfällt und größeren Grad als das Zählerpolynom p hat. Der Trick besteht in diesem Fall darin, den Ausdruck $\frac{p}{q}$ als Summe von Brüchen zu schreiben, die nur eine Konstante im Zähler und einen einzigen Linearfaktor im Nenner haben. Derartige Funktionen der Form $\frac{c}{x-a}$ lassen sich mit einer linearen Substitution wie in Beispiel 12.33 (b) dann einfach zu $c\log|x-a|$ integrieren.

Lemma 12.34 (Partialbruchzerlegung). *Es seien $n \in \mathbb{N}_{>0}$ und $a_1,\dots,a_n \in \mathbb{R}$ verschieden. Ferner sei p ein reelles Polynom mit $\deg p < n$. Dann gilt*

$$\frac{p(x)}{(x-a_1)\cdots(x-a_n)} = \sum_{i=1}^n \frac{c_i}{x-a_i} \quad \textit{für} \quad c_i := \frac{p(a_i)}{(a_i-a_1)\cdots(a_i-a_{i-1})(a_i-a_{i+1})\cdots(a_i-a_n)}$$

für alle $x \in \mathbb{R}\backslash\{a_1,\dots,a_n\}$.

Beweis. Offensichtlich ist

$$f(x) := p(x) - \sum_{i=1}^{n} p(a_i) \cdot \frac{(x-a_1)\cdots(x-a_{i-1})(x-a_{i+1})\cdots(x-a_n)}{(a_i-a_1)\cdots(a_i-a_{i-1})(a_i-a_{i+1})\cdots(a_i-a_n)} \qquad (*)$$

ein Polynom mit $\deg f < n$. Es hat aber jedes a_k für $k = 1,\dots,n$ als Nullstelle, denn es gilt

$$f(a_k) = p(a_k) - p(a_k) \cdot \frac{(a_k-a_1)\cdots(a_k-a_{k-1})(a_k-a_{k+1})\cdots(a_k-a_n)}{(a_k-a_1)\cdots(a_k-a_{k-1})(a_k-a_{k+1})\cdots(a_k-a_n)} = 0,$$

da nach Einsetzen von $x = a_k$ in der Summe über i in $(*)$ alle Terme mit $i \neq k$ einen Faktor 0 im Zähler haben und damit verschwinden. Nach Satz 3.19 (b) ist f also das Nullpolynom. Division von $(*)$ durch $(x-a_1)\cdots(x-a_n)$ liefert damit wie behauptet

$$0 = \frac{p(x)}{(x-a_1)\cdots(x-a_n)} - \sum_{i=1}^{n} \frac{c_i}{x-a_i}. \qquad \square$$

Bemerkung 12.35. Die Formel für die Koeffizienten c_i in Lemma 12.34 lässt sich leicht merken: Man erhält c_i, indem man $x = a_i$ im ursprünglichen Ausdruck $\frac{p(x)}{(x-a_1)\cdots(x-a_i)\cdots(x-a_n)}$ einsetzt – bis auf den Linearfaktor $x - a_i$ im Nenner, den man dabei weglassen muss, da er ansonsten ja auch zu einem Faktor 0 im Nenner führen würde.

Beispiel 12.36. Um das Integral $\int \frac{x}{x^2+3x+2}\,dx$ zu berechnen, führen wir eine Partialbruchzerlegung des Integranden durch: Mit $x^2+3x+2 = (x+1)(x+2)$ ist

$$\frac{x}{(x+1)(x+2)} = \frac{c_1}{x+1} + \frac{c_2}{x+2},$$

wobei sich $c_1 = \frac{-1}{1} = -1$ durch Einsetzen von $x = -1$ in $\frac{x}{x+2}$ und $c_2 = \frac{-2}{-1} = 2$ durch Einsetzen von $x = -2$ in $\frac{x}{x+1}$ ergibt. Also ist

$$\int \frac{x}{x^2+3x+2}\,dx = \int \left(-\frac{1}{x+1} + 2\,\frac{1}{x+2}\right) dx = -\log|x+1| + 2\log|x+2|$$

nach einer linearen Substitution wie in Beispiel 12.33 (b).

Als letzte Rechenregel zur Bestimmung von Integralen wollen wir nun noch untersuchen, wie man Integrale von Funktionen berechnen kann, die als Grenzwerte von Funktionenfolgen entstehen – also z. B. von Potenzreihen. Die Situation ist hier sehr viel einfacher als sie es bei der Differentiation in Satz 10.26 war.

Satz 12.37 (Vertauschbarkeit von Integration und Grenzwertbildung)**.** *Es seien* $f_n\colon [a,b] \to \mathbb{R}$ *stetige Funktionen, die gleichmäßig gegen eine (nach Satz 8.37 dann automatisch ebenfalls stetige) Grenzfunktion* $f\colon [a,b] \to \mathbb{R}$ *konvergieren. Dann gilt*

$$\int_a^b \left(\lim_{n\to\infty} f_n(x)\right) dx = \int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx,$$

d. h. „Grenzwertbildung und Integration können vertauscht werden".

Beweis. Die erste behauptete Gleichheit ist natürlich nichts weiter als die Definition von f. Für die zweite sei $\varepsilon > 0$ gegeben. Wegen der gleichmäßigen Konvergenz von $(f_n)_n$ gibt es ein $n_0 \in \mathbb{N}$, so dass $|f_n(x) - f(x)| < \frac{\varepsilon}{2(b-a)}$ für alle $x \in [a,b]$ und $n \geq n_0$. Damit ergibt sich nach Satz 12.13

$$\begin{aligned}\left|\int_a^b f_n(x)\,dx - \int_a^b f(x)\,dx\right| &= \left|\int_a^b (f_n(x)-f(x))\,dx\right| \leq \int_a^b |f_n(x)-f(x)|\,dx \leq \int_a^b \frac{\varepsilon}{2(b-a)}\,dx \\ &= \frac{\varepsilon}{2} < \varepsilon.\end{aligned}$$

Nach Definition des Grenzwerts bedeutet dies genau $\int_a^b f_n(x)\,dx \to \int_a^b f(x)\,dx$ für $n \to \infty$, was zu zeigen war. $\square$

Bemerkung 12.38 (Integration von Potenzreihen)**.** Insbesondere bedeutet Satz 12.37, dass Potenzreihen (die in jedem abgeschlossenen Intervall innerhalb des Konvergenzgebiets nach Satz 8.35 ja gleichmäßig konvergieren) gliedweise integriert werden können: Sind $f(x) = \sum_{k=0}^{\infty} c_k x^k$ eine Potenzreihe und $f_n(x) = \sum_{k=0}^{n} c_k x^k$ ihre Partialsummen, so folgt

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} \int_a^b f_n(x)\,dx = \lim_{n\to\infty} \left[\sum_{k=0}^{n} \frac{c_k}{k+1} x^{k+1} \right]_a^b = \left[\sum_{k=0}^{\infty} \frac{c_k}{k+1} x^{k+1} \right]_a^b$$

(sofern $[a,b]$ komplett im Konvergenzintervall der Potenzreihe liegt), d. h. als unbestimmtes Integral geschrieben ist

$$\int f(x)\,dx = \sum_{k=0}^{\infty} \frac{c_k}{k+1} x^{k+1}.$$

Dies zeigt noch einmal deutlich die besonders schönen Eigenschaften von Potenzreihen: Innerhalb ihres Konvergenzgebiets kann man mit ihnen praktisch „wie mit Polynomen rechnen“, d. h.

- sie können wie erwartet addiert und multipliziert werden (Lemma 7.4 und Bemerkung 7.36);
- sie sind beliebig oft differenzierbar und ihre Ableitungen können gliedweise berechnet werden (Folgerung 10.27 und Satz 11.8);
- Integrale können gliedweise berechnet werden;
- „viele Funktionen“ können als Potenzreihe (nämlich als ihre Taylor-Reihe, siehe Kapitel 11) geschrieben werden.

Beispiel 12.39.

(a) Wir betrachten die Funktion $f\colon \mathbb{R}_{>0} \to \mathbb{R}$, $x \mapsto \log x$. Die Ableitung dieser Funktion ist natürlich $f'(x) = \frac{1}{x}$. Nun können wir dies für $|x-1| < 1$ mit Hilfe der geometrischen Reihe (siehe Beispiel 7.3 (a)) als

$$f'(x) = \frac{1}{1+(x-1)} = 1 - (x-1) + (x-1)^2 - (x-1)^3 \pm \cdots$$

schreiben. Diese Potenzreihe kann jetzt aber nach Bemerkung 12.38 gliedweise integriert werden, und darum ist

$$(x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} \pm \cdots$$

für $|x-1| < 1$, also auf $(0,2)$, eine Stammfunktion von f'. Nach Folgerung 12.24 (b) kann sich diese von der ursprünglichen Funktion f nur um eine additive Konstante unterscheiden – Einsetzen von $x = 1$ liefert aber auch sofort, dass diese Konstante gleich 0 ist. Also erhalten wir die Darstellung

$$\log x = (x-1) - \frac{(x-1)^2}{2} + \frac{(x-1)^3}{3} - \frac{(x-1)^4}{4} \pm \cdots$$

für alle $x \in (0,2)$ (die wir in Beispiel 11.15 (a) bereits für $x \in [1,2]$ bewiesen hatten).

(b) Eine analoge Rechnung können wir auch mit der Funktion $f\colon (-1,1) \to \mathbb{R}$, $x \mapsto \arctan x$ durchführen: Hier ist die Ableitung

$$f'(x) = \frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 \pm \cdots,$$

und damit ist

$$x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \pm \cdots$$

eine Stammfunktion von f', die sich von f wiederum nur um eine additive Konstante unterscheiden kann. Auch hier ist diese Konstante wegen $\arctan 0 = 0$ wieder gleich 0, und wir erhalten auf $(-1,1)$ ohne irgendwelche komplizierte Rechnungen die Potenzreihendarstellung der Arkustangens-Funktion

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} \pm \cdots .$$

Aufgabe 12.40. Berechne die folgenden (z. T. unbestimmten bzw. uneigentlichen) Integrale:

$$\text{(a)} \int_{-1}^{1} \frac{1}{\sqrt{2x+3}}\,dx \qquad \text{(b)} \int_0^\infty x^2\,\mathrm{e}^{-2x}\,dx \qquad \text{(c)} \int \frac{x^3+x^2+1}{x^3-x}\,dx$$

$$\text{(d)} \int_{\mathrm{e}}^{\mathrm{e}^2} \frac{\log(\log x)}{x\log x}\,dx \qquad \text{(e)} \int \frac{1}{1+\mathrm{e}^x}\,dx \qquad \text{(f)} \int_0^{\frac{\pi}{2}} \sin^8 x\cos^3 x\,dx$$

Aufgabe 12.41. Zeige mit Induktion über $n \in \mathbb{N}$, dass

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^n x\,dx = \begin{cases} \pi \cdot \frac{1}{2} \cdot \frac{3}{4} \cdot \frac{5}{6} \cdot \cdots \cdot \frac{n-1}{n} & \text{falls } n \text{ gerade,} \\ 2 \cdot \frac{2}{3} \cdot \frac{4}{5} \cdot \frac{6}{7} \cdot \cdots \cdot \frac{n-1}{n} & \text{falls } n \text{ ungerade.} \end{cases}$$

Aufgabe 12.42 (Integralkriterium für Reihen). Es sei $f\colon \mathbb{R}_{\geq 1} \to \mathbb{R}_{\geq 0}$ eine stetige und monoton fallende Funktion. Man zeige:

(a) Das uneigentliche Integral $\int_1^\infty f(x)\,dx$ hat das gleiche Konvergenzverhalten wie die Reihe $\sum_{n=1}^\infty f(n)$, d. h. es sind entweder beide konvergent oder beide divergent.

(b) Für $a \in \mathbb{R}$ konvergiert die Reihe $\sum_{n=1}^\infty \frac{1}{n^a}$ genau dann, wenn $a > 1$. (Dies ist eine Verallgemeinerung von Beispiel 7.3 (c) und 7.19 auf reelle Exponenten.)

Gilt die Aussage (a) auch ohne die Voraussetzung, dass f monoton fallend ist?

Aufgabe 12.43. Es sei $f\colon [0,1] \to \mathbb{R}$ eine stetige Funktion. Man zeige:

(a) $\lim\limits_{n\to\infty} \int_0^1 f(x)\,x^n\,dx = 0.$

(b) Gilt $f(x) = f(1-x)$ für alle $x \in [0,1]$, so ist $\int_0^1 x\,f(x)\,dx = \frac{1}{2}\int_0^1 f(x)\,dx$.

Aufgabe 12.44. Es seien $a,b \in \mathbb{R}_{>0}$ und $f\colon [0,a] \to [0,b]$ eine bijektive, stetig differenzierbare Funktion. Man zeige:

(a) Ist f monoton wachsend mit $f(0)=0$ und $f(a)=b$, dann gilt $\int_0^a f(x)\,dx + \int_0^b f^{-1}(x)\,dx = ab$.

(b) Ist f monoton fallend mit $f(0)=b$ und $f(a)=0$, dann gilt $\int_0^a f(x)\,dx = \int_0^b f^{-1}(x)\,dx$.

Was bedeuten diese Aussagen geometrisch?

Aufgabe 12.45 (Binomische Reihe)**.** Für $a \in \mathbb{R}$ definieren wir die **verallgemeinerten Binomialkoeffizienten** durch

$$\binom{a}{n} := \frac{a\cdot(a-1)\cdot \cdots \cdot (a-n+1)}{n!}$$

für alle $n \in \mathbb{N}$. Wir betrachten nun auf $D=(-1,1)$ die Funktion $f\colon D \to \mathbb{R},\ f(x) = (1+x)^a$. Man zeige:

(a) Die Taylor-Reihe von f mit Entwicklungspunkt 0 ist gegeben durch $T_{f,0}(x) = \sum_{n=0}^\infty \binom{a}{n} x^n$ und konvergiert auf D.

(b) Es gilt sogar $(1+x)^a = \sum_{n=0}^\infty \binom{a}{n} x^n$ für alle $x \in D$, d. h. die Taylor-Reihe stellt wirklich die ursprüngliche Funktion dar. (Hinweis: Zeige zunächst, dass die Ableitung von $\frac{T_{f,0}}{f}$ gleich 0 ist.)

(c) Die Funktion arcsin lässt sich auf D als Potenzreihe schreiben. Berechne diese Potenzreihe explizit!

Was ergibt sich aus der binomischen Reihe in den Spezialfällen $a \in \mathbb{N}$ bzw. $a=-1$?

29

Literatur

[B] A. Beutelspacher, *Lineare Algebra*, Vieweg-Verlag (2003)
[BF] M. Barner, F. Flohr, *Analysis 1*, de Gruyter Lehrbuch (2000)
[E] H.-D. Ebbinghaus et al., *Zahlen*, Springer-Verlag (1988)
[Fi] G. Fischer, *Lineare Algebra*, Vieweg-Verlag (2002)
[Fo1] O. Forster, *Analysis 1*, Vieweg-Verlag (2011)
[Fo2] O. Forster, *Analysis 2*, Vieweg-Verlag (2010)
[G] A. Gathmann, *Algebraische Strukturen*, Vorlesungsskript RPTU Kaiserslautern (2023), https://agag-gathmann.math.rptu.de/ags
[GK] G.-M. Greuel, T. Keilen, *Lineare Algebra I und II*, Vorlesungsskript TU Kaiserslautern (1999), www.math.uni-tuebingen.de/~keilen/download/LectureNotes/linearealgebra.pdf
[J] K. Jänich, *Lineare Algebra*, Springer-Verlag (2010)
[K1] K. Königsberger, *Analysis 1*, Springer-Verlag (2003)
[K2] K. Königsberger, *Analysis 2*, Springer-Verlag (2003)
[M] T. Markwig, *Grundlagen der Mathematik*, Vorlesungsskript TU Kaiserslautern (2011), www.math.uni-tuebingen.de/~keilen/download/LectureNotes/grundlagen11.pdf

Index