V1G1 – Analysis I

Dozent:
Prof. Dr. Karl-Theodor Sturm

Mitschriften von:
Tien Nguyen Thanh

WS 2021/2022
Stand: 15. Januar 2022

Aktualisiert bis Vorlesung 22 am 12. Januar 2022.

Ich weiß, dass das ein wenig spät kommt, aber ich hoffe, dass ihr einen Nutzen hieraus ziehen könnt. Ursprünglich wollte ich das nur für mich machen, aber da unser Dozent kein Skript hat, gebe ich meines heraus. Grundsätzlich ist der Inhalt der Vorlesung wiederzufinden, aber oft umformuliert und teilweise auch von mir etwas kommentiert. In den Fußnoten stehen meist Randinformationen, die über den Tellerrand hinausführen. Der Anhang sollte auch interessant werden.

Feedback. Orthografische und inhaltliche Errata sowie konstruktive Kommentare bitte an `Gargantuar314@uni-bonn.de`. Auch Vorschläge zur Formulierung, fehlende Kapitel und Ausführungen, Konsistenz oder sogar zur Typografie sind willkommen. Ich werde mal schauen, ob ich sie dann einarbeite. Über das Layout und die mathematische Formelsetzung werde ich aber kaum mit mir diskutieren lassen, insbesondere die reine Gestaltung in schwarz/weiß. Aber vielleicht wird mich noch einer umstimmen?

Danksagung. Ein großer Dank geht an euch, die (hoffentlich) regelmäßig und aktiv dieses Skript nutzen. Des Weiteren bedanke ich mich bei allen, die mir Feedback und Korrekturhinweise gaben – hier bedanke ich mich insbesondere bei Julian Völlmecke und Solveig Tränkner.

Todo. Hinzufügen der fehlenden Beweise aus den Übungen. Vervollständigen des Anhangs. Vervollständigen der Schlagwortübersicht. Umsteigen auf `ntheorem`-Paket und Einrahmen von `theorem`-ähnlichen Umgebungen.

Grundlagen Aussagenlogik: logische Aussagen, Junktoren, Wahrheitstafeln, Tautologien. Beweisprinzip des indirekten Beweises. Prädikatenlogik: Quantoren, einige Regeln. Mengenlehre: Mengen, Element, Teilmenge, leere Menge, mengentheoretische Operatoren, Grenzen der naiven Mengenlehre. [1]

Vollständige Induktion Beweisprinzip der vollständigen Induktion. Summen- und Produktzeichen. Fakultät, Binomialkoeffizient. Pascalsches Dreieck, binomischer Lehrsatz. Kombinatorik: Permutation, Kombination. Potenzmenge. Multinomialkoeffizient, Multinomialsatz. [1,5]

Körper der reellen Zahlen Körperaxiome: Axiome eines Körpers, gefolgerte Rechenregeln, Verallgemeinerung der Assoziativität/Kommutativität/Distributivität. Anordnungsaxiome: Axiome der Anordnung, vergleichende Relationen, gefolgerte Rechenregeln. Archimedisches Axiom: Folgerungen, bernoullische Ungleichung. Betragsfunktion, Dreiecksungleichung. [1,5]

Konvergenz von Folgen I Begriffe: Folge, Konvergenz, Grenzwert/Limes, Beschränktheit. Grenzwertsätze. Ordnungserhaltung konvergenter Folgen. [1]

Vollständigkeitsaxiom Cauchy-Folgen. Vollständigkeitsaxiom. Intervallschachtelung: Intervallschachtelungsprinzip, Existenz der Wurzeln. Supremumseigenschaft. Äquivalenz aller vier Ansätze. Ausblick zu $\mathbb{R}$, Peano-Axiome von $\mathbb{N}$. Abzählbarkeit. [2,5]

Konvergenz von Folgen II Monotone Folgen. Uneigentliche Konvergenz, erweiterte Zahlengerade. Limes Superior/Inferior. Teilfolgen, Häufungspunkte, Satz von Bolzano-Weierstraß. [1,5]

Konvergenz von Reihen Reihen: harmonische/geometrische Reihe. Konvergenzkriterien: Cauchy-Kriterium, notwendige Bedingung einer Nullfolge, monoton und beschränkt, Leibniz-Kriterium, Majorantenkriterium, Quotientenkriterium, Wurzelkriterium. Riemannsche Zeta-Funktion. [1]

Absolut konvergente Reihen, summierbare Familien Absolut konvergente Reihen. Umordnung: Umordnungssatz. Summierbare Familien: großer Umordnungssatz, Doppelreihensatz. Produkt von Reihen: Faltung, Cauchy-Produkt. Binomische Reihe. [2]

Exponenzialreihe Exponenzialreihe: absolut konvergent, eulersche Zahl e. Restgliedabschätzung. Funktionalgleichung. Abschätzungen von exp. Grenzwertdarstellung. [1]

Komplexe Zahlen Körper der komplexen Zahlen: imaginäre Einheit i, Rechenregeln, komplexe Zahlenebene, Ähnlichkeit zu $\mathbb{R}^2$. Komplexe Konjugation, Betrag: Rechenregeln, Ungleichungen. Folgen in $\mathbb{C}$: Cauchy-Folgen, Grenzwertsätze, auch bzgl. Konjugation und Betrag. Reihen in $\mathbb{C}$: Konvergenzkriterien. Exponenzialreihe. Trigonometrische und hyperbolische Funktionen: Reihendarstellung, Identitäten (eulersche Formel, trigonometrischer Pythagoras), Additionstheoreme, Doppelwinkelfunktionen. Trigonometrische und hyperbolische Funktionen in $\mathbb{R}$, Abschätzungen. Kreiszahl π: Verhältnis zu exp/sin/cos, Periodizität, Phasenverschiebung, Nullstellen von sin/cos in $\mathbb{C}$. Polardarstellung: Multiplikation. Einheitswurzeln. [3]

Stetige Funktionen Funktionsbegriff: Beispiele, algebraische Operationen, Komposition, Umkehrung. Stetigkeit: ε-δ-Kriterium, Folgenkriterium, Äquivalenz beider. Folgerungen: Stetigkeit bekannter Funktionen, Stetigkeit unter algebraischen Operationen/Komposition. Zwischenwertsatz. Kompaktes Intervall. Extremwertsatz. Gleichmäßige Stetigkeit: gleichmäßige Stetigkeit auf kompakte Intervalle. Bijektive Funktionen. Stetigkeit in $\mathbb{C}$. Logarithmus: natürlicher Logarithmus, Logarithmus/Potenz zur beliebigen Basis, Potenzgesetze. Hyperbolische und trigonometrische Umkehrfunktion. [3]

Differenziation Differenzialquotient: Anschauung, Differenziation bekannter Funktionen. Charakterisierung der Differenzierbarkeit: Differenzenquotient, lineare Approximation. Ableitungsregeln: Summe, Produkt, Quotient, Kettenregel, Umkehrfunktionen. Höhere Ableitungen, stetige Differenzierbarkeit, Vektorräume der stetig differenzierbaren Funktionen. Extrema: lokales Extremum, Ableitung im Extremum, Satz von Rolle.

Inhaltsverzeichnis

1 Grundlagen

Wir werden uns zuerst mit einigen notwendigen logischen und mengentheoretischen Grundlagen beschäftigen. Dabei steigen wir nicht zu tief in die Materie ein, sondern behandeln alles (für unsere Zwecke ausreichend) etwas oberflächlich. Einen rigoroseren Zugang findet man in der Literatur dieser Gebiete.

1.1 Aussagenlogik

Definition 1.1 (Aussage)**.** Eine *Aussage* ist ein mathematischer oder sprachlicher Satz, der entweder *wahr* (1) oder *falsch* (0) ist.

Aussagen drücken damit keine Wahrscheinlichkeiten und auch kein „vielleicht“ aus. Ziel der Aussagenlogik ist es, aus gegebenen Aussagen neue (wahre) Aussagen abzuleiten.

Definition 1.2 (logische Operatoren, Junktoren)**.** Seien A und B zwei Aussagen. Es werden folgende Operatoren

Symbol	Name	Gesprochen
$\neg A$	*Negation*	„nicht A“
$A \wedge B$	*Konjunktion*	„A und B“, „sowohl A als auch B“
$A \vee B$	*Disjunktion*	„A oder B“ (dabei ist *nicht* das Exklusive Oder gemeint)
$A \Rightarrow B$	*Implikation*	„aus A folgt B“, „A impliziert B“, „wenn A, dann B“
$A \Leftrightarrow B$	*Äquivalenz*	„A ist äquivalent zu B“, „A genau dann, wenn B“, „A dann und nur dann, wenn B“ (engl. „iff – if and only if“)

durch die folgende *Wahrheitstafel* definiert:

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

Wir können uns die Wahrheitswerte wie folgt merken:

- $\neg A$ dreht den Wahrheitswert um.

  Das macht Sinn, denn wenn eine Aussage wahr war, wird sie danach falsch und andersherum.

- $A \vee B$ ist nur wahr, sobald eine der beiden oder beide Aussagen wahr ist. *Achtung*: Hiermit ist nicht das Exklusive Oder (auch „entweder ... oder“) gemeint, dass falsch ist, wenn beide Aussagen wahr sind.

  Auch das macht Sinn, denn wenn ich z. B. das eine *oder* das andere möchte, dann ist es egal, was von beiden eintritt – ich bin mit beiden zufrieden. Unzufrieden wäre ich hingegen, wenn weder das eine noch das andere eintritt.

- $A \wedge B$ ist nur wahr, wenn beide Aussagen wahr sind.

  Auch das macht Sinn: Wenn die Zitrone gelb und der Himmel blau ist, dann müssen schon beide Aussagen stimmen, damit diese Aussage insgesamt als richtig angesehen werden kann.

- $A \Rightarrow B$ ist nur falsch, wenn aus einem wahren A ein falsches B gefolgert wird. D. h., dass aus einer falschen Aussage alles gefolgert werden kann,[1] wohingegen aus einer wahren Aussage nur wahre Aussagen folgen.

  Das macht auch Sinn, denn wenn es regnet, ist die Straße zwar nass, aber wenn es nicht regnet, kann die Straße trotzdem nass sein (wenn z. B. jemand Wasser auf die Straße schüttete).

- $A \Leftrightarrow B$ ist genau dann wahr, wenn A und B denselben Wahrheitswert besitzen.

  Das macht Sinn, denn wenn zwei Dinge äquivalent („gleich“) sind, dann sollte man sie sich gegenseitig ersetzen können.

1.1.1 Gesetze der Aussagenlogik

Definition 1.3 (Tautologie)**.** Eine *Tautologie* ist eine Aussage, die immer wahr ist – unabhängig vom Wahrheitswert der darin enthaltenen Aussagen.

Tautologien, die eine Äquivalenz von Aussagen darstellen, können verwendet werden, um logische Aussagen äquivalent umzuformen und damit syntaktisch zu vereinfachen.

Satz 1.4 (Kontraposition)**.** *Für zwei Aussagen A und B gilt*

$$(A \Rightarrow B) \iff (\neg B \Rightarrow \neg A).$$

Satz 1.5. *Für zwei Aussagen A und B gilt*

$$(A \Rightarrow B) \iff (\neg A \vee B).$$

[1] Lat. *ex falso quodlibet* – „aus Falschem folgt Beliebiges“.

Beweis zu Satz 1.5. Wir beweisen den Satz mithilfe einer Wahrheitstafel.

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

Die Äquivalenz der beiden letzten Spalten ist stets wahr. □

Treten in einer zu zeigenden Tautologie noch mehr Aussagen auf, wächst die Wahrheitstafel dementsprechend um mehr Zeilen. Bei z. B. drei Aussagen A, B und C sind es schon $2^3 = 8$ Zeilen.

Satz 1.6. *Weitere Tautologien und Schlussregeln, die oft verwendet werden:*[2] *Seien A, B und C Aussagen. Dann gelten*

$$(A \Leftrightarrow B) \iff (\neg A \Leftrightarrow \neg B),$$
$$((A \Rightarrow B) \wedge A) \implies B \qquad \text{Modus ponens (Prinzip der vollständigen Induktion),}$$
$$((A \Rightarrow B) \wedge \neg B) \implies \neg A \qquad \text{Modus tollens (Prinzip des indirekten Beweises),}$$
$$((A \Rightarrow B) \wedge (B \Rightarrow C)) \implies (A \Rightarrow C) \qquad \text{Hypothetischer Syllogismus,}$$
$$A \implies (A \vee B),$$
$$\neg A \implies (A \Rightarrow B).$$

1.1.2 Indirekter Beweis

Aufbauend auf Satz 1.4 bzw. dem *Modus tollens* aus Satz 1.6 lässt sich das Beweisprinzip des *indirekten Beweises* (auch *Widerspruchsbeweis*, *reduction ad absurdum*) formulieren.

Wir wollen die Aussage $A \implies B$ beweisen. Dafür

- nehmen wir an, dass die Annahme $\neg B$ gilt und
- führen das zum Widerspruch mit A (wir erhalten $\neg B \implies \neg A$, was im Widerspruch zum wahren A steht).

Beispiel 1.7. *Satz von* Euklid*: Es gibt unendlich viele Primzahlen.*

Beweis: Angenommen es gibt nur endlich viele Primzahlen $p_1, p_2, \ldots, p_n$. Die Zahl $m = p_1 p_2 \ldots p_n + 1$ ist aber durch keine der obigen Primzahlen teilbar, weshalb m selbst eine Primzahl ist. Folglich gibt es mehr Primzahlen als die $p_1, p_2, \ldots, p_n$, ein Widerspruch, weshalb die Annahme falsch ist und somit unendlich viele Primzahlen existieren.

1.2 Prädikatenlogik

Definition 1.8 (Aussageform)**.** Aussagen, die von freien Variablen abhängen, heißen *Aussageformen.*

Eine Aussageform besitzt keinen Wahrheitswert. Nach Einsetzen in die Variablen wird die Aussageform zur Aussage und wir können dann einen Wahrheitswert zuordnen.

Beispiel 1.9.

[2] Noch weitere (nicht triviale) hilfreiche Tautologien, die der Dozent leider nicht nannte, sind

$$A \Rightarrow (B \Leftrightarrow C) \iff (A \Rightarrow B) \Rightarrow (A \Rightarrow C),$$
$$A \Rightarrow (B \wedge C) \iff (A \Rightarrow B) \wedge (A \Rightarrow C),$$
$$A \iff \neg(\neg A) \qquad \text{doppelte Negation,}$$
$$A \wedge (B \vee C) \iff (A \wedge B) \vee (A \wedge C) \qquad A \vee (B \wedge C) \iff (A \vee B) \wedge (A \vee C) \qquad \text{Distributivität,}$$
$$A \wedge (A \vee B) \iff A \qquad A \vee (A \wedge B) \iff A \qquad \text{Absorption,}$$
$$\neg(A \wedge B) \iff (\neg A \vee \neg B) \qquad \neg(A \vee B) \iff (\neg A \wedge \neg B) \qquad \text{de-morgansche Gesetze.}$$

1. Es gibt (mindestens) eine reelle Zahl x mit $x^2 + 1 = 0$.
2. Für alle natürlichen Zahlen x gilt $x^2 - 1 = (x-1)(x+1)$.

Die Gleichungen $x^2 + 1 = 0$ und $x^2 - 1 = (x-1)(x+1)$ sind Aussageformen. Erst nach Einsetzen gewisser x können wir bestimmen, ob sie wahr sind. Es stellt sich heraus, dass es kein x gibt, für das die erste Gleichung wahr ist, und alle x die zweite Gleichung erfüllen. Deshalb ist Punkt 1 falsch und Punkt 2 wahr.

Bezeichnung 1.10 (Quantoren)**.** Wir schreiben kürzer

- $\exists$ (*Existenzquantor*) für „es gibt ein" oder „es existiert ein" und
- $\forall$ (*Allquantor*) für „für alle".

Damit können wir die Aussagen aus dem obigen Beispiel 1.9 mit den Quantoren durch

- $\exists x \in \mathbb{R} \colon x^2 + 1 = 0$ und
- $\forall x \in \mathbb{N} \colon x^2 - 1 = (x-1)(x+1)$

ausdrücken.

Satz 1.11 (einige Regeln über Quantoren)**.**

1. *Quantoren lassen sich negieren. Sei hierfür $A(x)$ eine Aussage in x. Dann gilt*
$$\neg(\forall x \colon A(x)) \iff (\exists x \colon \neg A(x)), \quad \textit{daher auch} \quad \neg(\exists x \colon A(x)) \iff (\forall x \colon \neg A(x)).$$
Auch gilt für eine Aussage $A(x, y)$ in x und y
$$\neg(\exists x \colon \forall y \colon A(x,y)) \iff (\forall x \colon \exists y \colon \neg A(x,y)) \quad \textit{und} \quad \neg(\forall x \colon \exists y \colon A(x,y)) \iff (\exists x \colon \forall y \colon \neg A(x,y)).$$
2. *Ein Existenz- vor einem Allquantor lässt sich vertauschen. Sei $A(x, y)$ eine Aussage in x und y. Es gilt*
$$(\exists x \colon \forall y \colon A(x,y)) \implies (\forall y \colon \exists x \colon A(x,y)).$$
3. *Für zwei Aussagen $A(x)$ und $B(x)$ in x gilt*
$$(\exists x \colon A(x) \wedge B(x)) \implies ((\exists x \colon A(x)) \wedge (\exists x \colon B(x))).$$
4. *Im Allgemeinen gilt für eine Aussage $A(x)$ über eine Menge M* nicht
$$(\forall x \in M \colon A(x)) \implies (\exists x \in M \colon A(x)).$$

Bemerkung 1.12. Theoretisch lassen sich diese Aussagen korrekt beweisen, worauf wir hier aber verzichten werden. Trotzdem einige Bemerkungen:

1. Um eine Aussage mit beliebig vielen Quantoren zu verneinen, müssen wir uns von vorne nach hinten durcharbeiten und jeden All durch einen Existenzquantor ersetzen und andersherum. Zu allerletzt wird $A(x)$ bzw. $A(x, y)$ negiert.

   Das bedeutet auch insbesondere, dass wenn eine Allaussage falsch sein soll, nur ein *Gegenbeispiel* reicht. Ist hingegen ein Existenzaussage falsch, so müssen wir zeigen, dass *alle* Objekte die gegebene Eigenschaft nicht erfüllen.

2. Wenn es ein x gibt, sodass $A(x, y)$ für alle y gilt, so gibt es für jedes y auch ein x – nämlich erfüllt stets dasselbe x die Aussage $A(x, y)$. Die Umkehrung gilt hingegen im Allgemeinen *nicht*, da wenn es für jedes y zwar ein x gibt, muss es noch lange nicht für jedes y stets dasselbe x sein. *Beispiel*: Für $x = 0$ und allen $y \in \mathbb{N}$ gilt $A(x, y)$. Dann gibt es für jedes $y \in \mathbb{N}$ auch ein x, nämlich stets dasselbe $x = 0$. Gibt es hingegen für jedes $y \in \mathbb{N}$ ein x, so gilt die Umkehrung nicht, da für $y = 1$ z. B. $x = 1$, für $y = 2$ aber z. B. $x = 2$ die Aussage $A(x, y)$ erfüllt und wir somit nicht mit Sicherheit sagen können, dass ein x existiert, für das $A(x, y)$ für alle y wahr ist.

3. Der Existenzquantor kann distributiv auf eine Konjunktion verteilt werden. Wenn ein x sowohl $A(x)$ als auch $B(x)$ erfüllt, so erfüllt selbstverständlich dasselbe x auch $A(x)$ und $B(x)$ getrennt voneinander. Die Umkehrung gilt hingegen *nicht*: Denn ein x, dass $A(x)$ erfüllt, und ein x, dass $B(x)$ erfüllt, muss nicht dasselbe sein.

4. Ein Allquantor impliziert im Allgemeinen nicht den Existenzquantor. *Gegenbeispiel*: Die Menge M könnte ja leer sein.

1.3 Mengenlehre

Für unsere Zwecke reicht die naive Mengenlehre, begründet durch Georg Cantor.

Definition 1.13 (naive Menge und Elemente)**.** Eine *Menge* ist die Zusammenfassung bestimmter wohlunterscheidbarer Objekte unserer Anschauung oder unseres Denkens (welche die *Elemente* der Menge genannt werden) zu einem Ganzen.

Beispiel 1.14.

- Menge der natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, \dots\}$ (für diese Vorlesung beginnend bei 1), $\mathbb{N}_0 = \{0, 1, 2, 3, \dots\}$
- $M = \{x \in \mathbb{R} : A(x)\}$ für die Menge aller reellen Zahlen, für die $A(x)$ gilt
- $M = \{a, b, c, d\}$ als einfache Auflistung

Bezeichnung 1.15. Für zwei Mengen M und N bezeichnen

- $x \in M$ „*ist Element von*“,
- $M \subset N$ die *Teilmenge* M, jedes Element von M ist auch in N enthalten (nicht notwendigerweise echt, d. h. nicht notwendigerweise $M \neq N$),
- $M \subsetneq N$ die *echte Teilmenge* M, also $M \subset N$ und $M \neq N$, und
- $\varnothing = \{\}$ die *leere Menge*, die keine Elemente enthält ($\varnothing$ ist *nicht* $\{0\}$).

Definition 1.16 (mengentheoretische Operatoren)**.** Für zwei Mengen M und N werden folgende mengentheoretische Operatoren definiert:

Symbol	Name	Definition
$M \cap N$	*Durchschnitt*	$\{x : (x \in M) \text{ und } (x \in N)\}$
$M \cup N$	*Vereinigung*	$\{x : (x \in M) \text{ oder } (x \in N)\}$
$N^C = \complement N$	*Komplement*	$\{x \in X : x \notin N\}$ für eine übergeordnete Menge $X \supset N$
$M \setminus N = M \cap N^C$	*Differenz*	$\{x \in M : x \notin N\}$

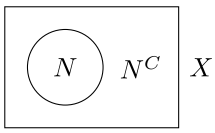


Abbildung 1.1: Komplementmenge $N^C = \complement N$.

Definition 1.17 (disjunkt)**.** Zwei Mengen M und N heißen *disjunkt*, falls $M \cap N = \varnothing$ gilt.

Die optische Ähnlichkeit der mengentheoretischen Operatoren $\wedge$ und $\vee$ mit den logischen Operatoren $\cap$ und $\cup$ ist kein Zufall. Bspw. gilt

$$\{x \in M : A(x) \wedge B(x)\} = \{x \in M : A(x)\} \cap \{x \in M : B(x)\},$$
$$\{x \in M : A(x) \vee B(x)\} = \{x \in M : A(x)\} \cup \{x \in M : B(x)\}.$$

1.3.1 Grenzen der naiven Mengenlehre

Im Laufe der Mathematikgeschichte stellte sich leider heraus, dass die naive Mengenlehre[3] sich zu unklar über erlaubte und nicht erlaubte Mengenbildungen äußert. Die Mengendefinition lässt also einen Interpretationsspielraum, was zu Widersprüchen führen kann.[4]

Nicht erlaubt ist z. B. die Konstruktion der Mengen M aller Mengen, die sich nicht selbst als Element enthalten (*Russellsche Antinomie*). Es stellt sich die Frage, ob M sich selbst enthält oder nicht: Alle Elemente

[3] Eigentlich mehrere naive Mengenlehren, wie z. B. von Georg Cantor, Richard Dedekind oder Gottlob Frege.
[4] Später wurde sie durch axiomatische Systeme ersetzt (*Zermelo-Fraenkel-Mengenlehre*).

von M haben per Definition die Eigenschaft, dass sie sich selbst nicht enthalten. Gilt $M \in M$, so kann M sich nach Definition nicht selbst enthalten, Widerspruch! Gilt $M \notin M$, so muss M nach Definition sich selbst enthalten, Widerspruch! Eine solche Konstruktion ist also unmöglich.

Analog dazu formulierte Bertrand Russel das *Barbierproblem*: Der Barbier ist derjenige, der ausschließlich all diejenigen rasiert, die sich nicht selbst rasieren. Rasiert er sich? Das führt unumgänglich zu einem Paradox.[5]

2 Vollständige Induktion

Die *vollständige Induktion* ist ein mächtiges Beweisverfahren, was sehr oft Anwendung findet. Ist $A(n)$ eine Aussage, die für alle $n \in \mathbb{N}$ gelten soll, können wir nicht nacheinander unendlich viele $A(1), A(2), \dots$ zeigen. Das Beweisverfahren schafft es aber, die Aussage *für alle* n zu zeigen.

2.1 Beweisprinzip der vollständigen Induktion

Wir setzen die Menge der natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, \dots\}$ vor. Die Aussage $A(n)$ gilt für alle $n \in \mathbb{N}$, falls:

1. Induktionsanfang: $A(1)$ gilt,
2. Induktionsschritt (Schritt von n auf $n+1$): falls $A(n)$, dann auch $A(n+1)$ für alle $n \in \mathbb{N}$ gilt.

Ende der Vorlesung 1 am 11. Oktober 2021

Als Beispiel betrachten wir den

Satz 2.1 (kleiner Gauss)**.** *Für alle* $n \in \mathbb{N}$ *gilt* $1 + 2 + 3 + \dots + n = \frac{1}{2}n(n+1)$.

Beweis durch vollständige Induktion. $A(n)$ bezeichne die behauptete Aussage.

1. Induktionsanfang ($n = 1$): $1 = \frac{1}{2}(1 \cdot 2)$, $A(1)$ ist wahr.
2. Induktionsschritt (angenommen $A(n)$ gilt, dann gilt auch $A(n+1)$):

$$1 + 2 + \dots + (n+1) = (1 + \dots + n) + (n+1) \overset{A(n)}{=} \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}. \qquad \square$$

Bemerkung 2.2. In solch einem Beweis muss immer spezifiziert werden, was die zu zeigende Aussage $A(n)$, was der Induktionsanfang (z. B. $n = 1$) und was der Induktionsschritt (z. B. von n auf $n+1$, von $n+1$ auf $n+2$) ist.

Bezeichnung 2.3 (Summenzeichen)**.** Für ganze Zahlen m, n und $a_k \in \mathbb{R}$ ist

$$\sum_{k=m}^{n} a_k = a_m + a_{m+1} + \dots + a_{n-1} + a_n, \quad \text{falls } m \le n,$$
$$\sum_{k=m}^{m} a_k = a_m \qquad \text{und} \qquad \sum_{k=m}^{m-1} a_k = 0 \quad \text{(leere Summe).}$$

Für die leere Summe ist die Indexmenge leer. Es ist Konvention, diese auf Null zu definieren.

Beispiel 2.4. Satz 2.1 lässt sich dann als $\sum_{k=1}^{n} k = \frac{1}{2}n(n+1)$ für alle $n \in \mathbb{N}$ schreiben.

Zur Übung besprechen wir noch weitere Induktionsbeweise als Beispiele.

Satz 2.5 (Summe der ersten n ungeraden Zahlen)**.** *Für alle* $n \in \mathbb{N}$ *gilt* $\sum_{k=1}^{n}(2k-1) = n^2$.

Das ist bestimmt wahr, denn $1 = 1^2$, $1 + 3 = 2^2$, $1 + 3 + 5 = 3^2$ und $1 + 3 + 5 + 7 = 4^2$.

Beweis. Sei $A(n)$ die zu zu zeigende Behauptung.

1. Induktionsanfang ($n = 1$): $1 = 1^2$, $A(1)$ ist wahr.

[5]Eine interessante Anekdote aus Paul R. Halmos' *Naive Mengenlehre* dazu befindet sich in Anhang B.

2. Induktionsschritt (von n auf $n+1$):

$$\sum_{k=1}^{n+1}(2k-1) = (2n+1) + \sum_{k=1}^{n}(2k-1) \overset{A(n)}{=} (2n+1) + n^2 = (n+1)^2. \qquad \square$$

Satz 2.6 (geometrische Reihe)**.** *Für alle reellen Zahlen $x \neq 1$ und für alle $n \in \mathbb{N}$ gilt*

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

(Hierbei ist $x^0 = 1$ auch für $x = 0$ definiert!)

Beweis. Sei $A(n)$ die zu zeigende Behauptung.

1. Induktionsanfang ($n = 1$): $1 + x = (1-x^2)/(1-x)$, $A(1)$ ist wahr.
2. Induktionsschritt (von n auf $n+1$):

$$\sum_{k=0}^{n+1} x^k = x^{n+1} + \sum_{k=0}^{n} \overset{A(n)}{=} \frac{1-x^{n+1}}{1-x} + x^{n+1} = \frac{1-x^{n+2}}{1-x}. \qquad \square$$

Bemerkung 2.7*.*

1. Statt dem Anfang bei $n = 1$ kann der Induktionsbeweis bei jeder Zahl $n_0 \in \mathbb{Z}$ starten und dann die Aussage für alle $n \in \mathbb{Z}$, $n \geq n_0$ liefern.
2. *Erweitertes Induktionsprinzip*: Die Aussagen $A(n)$ gelten für alle $n \in \mathbb{N}$, falls $A(1)$ gilt (Induktionsanfang) und falls $(A(1) \wedge A(2) \wedge \cdots \wedge A(n)) \implies A(n+1)$ gilt (erweiterter Induktionsschritt) – d. h., dass wir als Induktionsannahme nicht nur $A(n)$, sondern auch alle $A(1), \ldots, A(n)$ annehmen können.

   Dieses Beweisprinzip ist äquivalent zum bekannten Induktionsprinzip nach „Umetikettieren“: Setze $B(n) = A(1) \wedge \cdots \wedge A(n)$ und führe den Beweis von $B(n)$ durch. Insbesondere ist $B(1) = A(1)$.[1]

2.2 Binomischer Lehrsatz

Bevor wir zum binomischen Lehrsatz kommen, müssen wir noch einige Begriffe aus der Kombinatorik definieren.

Bezeichnung 2.8 (Produktzeichen)**.** Das *Produktzeichen* wird analog zum Summenzeichen definiert. Für ganze Zahlen m, n und $a_k \in \mathbb{R}$ bedeutet

$$\prod_{k=m}^{n} a_k = a_m \cdot a_{m+1} \cdots a_{n-1} \cdot a_n, \quad \text{falls } m \leq n,$$

$$\prod_{k=m}^{m} a_k = a_m \qquad \text{und} \qquad \prod_{k=m}^{m-1} a_k = 1 \quad \text{(leeres Produkt).}$$

Das leere Produkt wird konventionell auf Eins definiert.

Das Produktzeichen lässt sich äquivalent dazu auch rekursiv definieren:

$$\prod_{k=m}^{m-1} a_k = 1, \quad \prod_{k=m}^{n} = a_n \cdot \prod_{k=m}^{n-1} a_k \quad \text{für alle } n \geq m$$

verglichen mit dem Summenzeichen

$$\sum_{k=m}^{m-1} a_k = 0, \quad \sum_{k=m}^{n} = a_n + \sum_{k=m}^{n-1} a_k \quad \text{für allen} n \geq m$$

[1] Noch eine dritte Bemerkung: Induktion kann auch *rückwärts* von n auf $n-1$ erfolgen, sodass ggf. $A(n)$ nicht nur für $n \geq n_0$, sondern alle $n \in \mathbb{Z}$ gezeigt wurde. Ausgeklügeltere Induktionsbeweise können auch vorwärts und rückwärts in Sprüngen erfolgen (s. hierzu Cauchys Beweis für den *Satz vom arithmetischen und geometrischen Mittel*).

Definition 2.9 (Fakultät)**.** Für $n \in \mathbb{N}_0$ sei die *Fakultät* $n!$ (sprich „n Fakultät") definiert als

$$n! = \prod_{k=1}^{n} k = 1 \cdot 2 \cdots n.$$

Für $n = 0$ ist das Produkt ein leeres Produkt und es gilt $\prod_{k=1}^{0} k = 0! = 1$.

Äquivalent dazu ist die rekursive Definition

$$0! = 1, \quad (n+1)! = (n+1) \cdot n!\,.$$

Definition 2.10 (Binomialkoeffizient)**.** Für $k, n \in \mathbb{R}$ (meist aber $n, k \in \mathbb{N}_0$) sei der *Binomialkoeffizient* (sprich „n über k" oder „k aus n", engl. „n choose k") definiert als

$$\binom{n}{k} = \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} = \prod_{j=1}^{k} \frac{n-j+1}{j}. \tag{2.11}$$

Daraus folgt für $n, k \in \mathbb{N}_0$ unmittelbar

$$\binom{n}{k} = \begin{cases} \dfrac{n!}{k!\,(n-k)!} & \text{falls } 0, \leq k \leq n, \\ 0 & \text{falls } k > n. \end{cases}$$

Zusätzlich führen wir die Konvention $\binom{n}{k} = 0$ für negative k ein.

Bemerkung 2.12.

- In der Definition mit Fakultäten sehen wir, dass $k!$ und $(n-k)!$ als Faktoren zueinander symmetrisch sind, d. h., dass $\binom{n}{k} = \binom{n}{n-k}$ gilt. In (2.11) sehen wir, dass wenn $n-k$ statt k eingesetzt wird, mehr Faktoren sowohl im Zähler als auch im Nenner stehen, sie sich aber herauskürzen:
$$\binom{n}{n-k} = \frac{n(n-1)\cdots(n-k+1)\,\big[(n-k)\cdots(k+1)\big]}{\big[(n-k)\cdots(k+1)\big]\,k(k-1)\cdots 1}.$$
- Die Definition $\binom{n}{k} = 0$ für $k > n$ macht Sinn, denn für diesen Fall taucht im Zähler der Faktor Null auf.

Korollar 2.13.

$$\binom{n}{0} = \binom{n}{n} = 1, \qquad \binom{n}{1} = \binom{n}{n-1} = n.$$

Lemma 2.14. *Für alle $n \in \mathbb{N}_0$ und für alle $k \in \mathbb{Z}$ gilt*[2]

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

Beweis. Für $k \leq -2$ ergibt sich $0 = 0 + 0$, für $k = -1$ ergibt sich $1 = 0 + 1$ und für $k = 0$ ergibt sich $(n+1) = n + 1$. Damit ist das Lemma für $k \leq 0$ richtig.

Für $k > 0$ ergibt sich

$$\begin{aligned} \binom{n}{k} + \binom{n}{k+1} &= \frac{n(n-1)\cdots(n-k+1)}{k(k-1)\cdots 1} + \frac{n(n-1)\cdots(n-k)}{(k+1)k(k-1)\cdots 1} \\ &= \frac{n(n-1)\cdots(n-k+1)(k+1) + n(n-1)\cdots(n-k)}{(k+1)k\cdots 1} \\ &= \frac{n(n-1)\cdots(n-k+1)(k+1+n-k)}{(k+1)k\cdots 1} = \binom{n+1}{k+1}. \end{aligned}$$ □

Die Binomialkoeffizienten können wir übersichtlich im PASCALSCHEM *Dreieck* anordnen (s. Abb. 2.1). Jede Zeile n wird von oben nach unten beginnend bei null durchnummeriert, und jede Zeile hat von links nach rechts durchnummerierte Spalten k beginnend bei null (Spalten sind also eher „diagonal"). Jeder Eintrag in Zeile n und Spalte k entspricht dem Binomialkoeffizienten $\binom{n}{k}$. Bspw. gilt für den Eintrag $10 = \binom{5}{2}$ bei $n = 5$ und $k = 2$.

Eine Eigenschaft des PASCALSCHEN Dreiecks ist, dass ein Eintrag (bis auf den Einsen) genau die Summe der zwei darüberstehenden Einträge ist. Bspw. gilt $10 = 4 + 6$. Diese Eigenschaft folgt direkt aus Lemma 2.14.

[2] Auf Englisch heißt das Lemma PASCAL*'s rule*.

$$
\begin{array}{ccccccccccc}
 & & & & & 1 & & & & & \\
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1
\end{array}
$$

Abbildung 2.1: Pascalsches Dreieck bis Stufe 5.

Satz 2.15 (binomischer Lehrsatz)**.** *Für alle reellen x und y und $n \in \mathbb{N}_0$ gilt*

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

Einige Werte für n:

$$
\begin{aligned}
(x+y)^0 &= 1, \\
(x+y)^1 &= x+y, \\
(x+y)^2 &= x^2 + 2xy + y^2, \\
(x+y)^3 &= x^3 + 3x^2y + 3xy^2 + y^3, \\
(x+y)^5 &= x^5 + 5x^4y + 10x^3y^2 + 10x^2y^3 + 5xy^4 + y^5.
\end{aligned}
$$

Die Koeffizienten für ein n entsprechen den Einträgen des pascalschen Dreiecks der Zeile n.

Beweis. Wir beweisen den Satz durch vollständige Induktion nach n bei vorgegebenen x und y. Sei $A(n)$ die zu zeigende Aussage.

1. Induktionsanfang ($n = 0$): $(x+y)^0 = \binom{0}{0} = 1$, $A(0)$ ist wahr.

2. Induktionsschritt (von n auf $n+1$):

$$
\begin{aligned}
(x+y)^{n+1} = (x+y)(x+y)^n &\overset{A(n)}{=} \sum_{k=0}^{n} \binom{n}{k} x^{n-k+1} y^k + \sum_{k=0}^{n} \binom{n}{k} x^{n-k} y^{k+1} \\
&\overset{(*)}{=} \sum_{k=0}^{n} \binom{n}{k} x^{n-k+1} y^k + \sum_{k=1}^{n+1} \binom{n}{k-1} x^{n-k+1} y^k \overset{(\dagger)}{=} \sum_{k=0}^{n+1} \left(\binom{n}{k} + \binom{n}{k-1} \right) x^{n-k+1} y^k \\
&\overset{(2.14)}{=} \sum_{k=0}^{n+1} \binom{n+1}{k} x^{n-k+1} y^k. \qquad \square
\end{aligned}
$$

Zu $(*)$: Beim zweiten Summanden führten wir eine sog. *Indexverschiebung* durch. Die Anzahl der Summanden des Summenzeichens bleibt gleich (nämlich $n+1$). Da k nun stets um Eins größer ist, muss im Ausdruck jedes k um Eins kleiner werden, d. h. wir ersetzen alle k durch $k-1$.

Zu $(\dagger)$: Wir konnten die Summen zusammenfassen bzw. ausklammern, weil wir den Bereich des Laufindex anglichen. Dafür ergänzten wir zur ersten und zweiten Summe den Term $\binom{n}{k} x^{n-k+1} y^k$ für $k = n+1$ bzw. $k = 0$. Für beide k wird der Term Null.

Korollar 2.16. *Einsetzen von*

1. *$x = 1$ und $y = 1$ liefert $\sum_{k=0}^{n} \binom{n}{k} = 2^n$ und*
2. *$x = 1$ und $y = -1$ liefert $\sum_{k=0}^{n} (-1)^k \binom{n}{k} = 0$.*

Ende der Vorlesung 2 am 13. Oktober 2021

2.2.1 Kombinatorik

Der mathematische Bereich der *(klassischen) Kombinatorik* beschäftigt sich mit der Frage, wie viele Möglichkeiten es gibt, bestimmte Konfigurationen zu erhalten. Sie untersucht das Abzählen bzw. abzählbare Strukturen.

Die oben definierten Begriffe der Fakultät und des Binomialkoeffizienten haben dementsprechend auch eine kombinatorische Bedeutung.

Satz 2.17 (Permutation)**.** *Die Anzahl der Anordnungen von n verschiedenen Elementen ist $n!$.*

Beispiel 2.18.

- $n = 1$: a_1 $(1 = 1!)$
- $n = 2$: a_1a_2, a_2a_1 $(2 = 2!)$
- $n = 3$: $a_1a_2a_3$, $a_2a_3a_1$, $a_3a_1a_2$, $a_3a_2a_1$, $a_2a_3a_1$, $a_1a_3a_2$ $(6 = 3!)$

Beweis. Sei $N(n)$ die Anzahl der Anordnungen von n Elementen. Die zu zeigende Aussage ist $A(n) :\!= N(n) = n!$.

1. Induktionsanfang: $A(1)$ ist wahr (s. Beispiel 2.18).
2. Induktionsschritt: Wir nehmen an, dass $A(n)$ gilt. Gegeben seien $n + 1$ Elemente $a_1, \dots, a_{n+1}$, die angeordnet werden sollen.

   Für den ersten Platz gibt es $n + 1$ Möglichkeiten, ein Element von $a_1, \dots, a_{n+1}$ zu wählen. Für die letzten n Plätze ist jede Möglichkeit zur Anordnung der verbleibenden n Elemente erlaubt. Dafür gibt es $N(n) = n!$ Möglichkeiten zur Anordnung. Insgesamt gibt es also $(n + 1) \cdot n! = (n + 1)!$ Möglichkeiten. □

Bezeichnung 2.19. Wir nennen eine „Möglichkeit der Anordnung (der Elemente einer Menge)“ eine *Permutation (der Menge)*.

Satz 2.20 (Kombination)**.** *Die Anzahl der k-elementigen Teilmengen einer n-elementigen Menge ist $\binom{n}{k}$. (Insbesondere gilt daher, dass $\binom{n}{k}$ ganzzahlig ist.)*

Beweis. Wir beweisen den Satz durch vollständige Induktion über n (und nicht über k). Sei $A(n)$ die Aussage, dass der obige Satz für alle $0 \le k \le n$ gilt.

1. Induktionsanfang: Für $n = 0$ gibt es genau eine Anordnung der leeren Menge, nämlich sie selbst. $A(0)$ ist also wahr.
2. Induktionsschritt (von n nach $n + 1$): Gegeben sei die Menge $\{a_1, \dots, a_{n+1}\}$. Angenommen $A(n)$ ist wahr. Für jedes $0 \le k \le n + 1$ gilt: Alle k-elementigen Teilmengen zerfallen in zwei Klassen.

   (a) Teilmengen, die a_{n+1} enthalten.

   (b) Teilmengen, die a_{n+1} nicht enthalten.

   Für Punkt 2a gilt: Das sind die Mengen $\{a_{n+1}\} \cup M_{k-1}$ mit beliebigen $(k-1)$-elementigen Teilmengen M_{k-1} von $\{a_1, \dots, a_n\}$. Gemäß $A(n)$ gibt es davon $\binom{n}{k-1}$.

   Für Punkt 2b gilt: Das sind beliebige k-elementige Teilmengen von $\{a_1, \dots, a_n\}$. Gemäß $A(n)$ gibt es davon $\binom{n}{k}$.

   Insgesamt gibt es also $\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k}$. (Für $k > n$ oder $k < 0$ wird der Binomialkoeffizient im zu zeigenden Satz 0, was trivial ist.) Damit ist $A(n + 1)$ bewiesen. □

Bemerkung 2.21. Es ist tatsächlich nicht mehr notwendig, nochmal über k zu induzieren. Da wir für den Beweis ein beliebiges k betrachteten, sind die Schlussfolgerungen *für alle* k richtig – sogar für $k = n + 1$, denn dieser Fall gehört zu Punkt 2a.

Definition 2.22. Für eine Menge M ist die *Potenzmenge* $\mathcal{P}(M)$ die Menge aller Teilmengen von M.

Beispiel 2.23. Für $M = \{a_1, a_2, a_3\}$ ist die Potenzmenge

$$\mathcal{P}(M) = \{\emptyset, \{a_1\}, \{a_2\}, \{a_3\}, \{a_1, a_2\}, \{a_1, a_3\}, \{a_2, a_3\}, \{a_1, a_2, a_3\}\}.$$

Satz 2.24 (Mächtigkeit der Potenzmenge)**.** *Die Potenzmenge $\mathcal{P}(M)$ einer n-elementigen Menge M hat 2^n Elemente.*

Beweis. Es gibt $\binom{n}{k}$ Teilmengen mit k Elementen, also insgesamt $\sum_{k=0}^{n} \binom{n}{k} = 2^n$ Teilmengen (Korollar 2.16). □

Bemerkung 2.25*.* Eine Alternative zum Beweis: Jede Teilmenge T von M entspricht einer Abbildung

$$f\colon M \to \{0,1\} \quad \text{mit } \forall x \in M\colon f(x) = 1 \iff x \in T.$$

Es gibt 2^n solche Abbildungen.

2.2.2 Multinomialsatz

Satz 2.26 (Multinomialkoeffizient)**.** *Gegeben sei $k, n \in \mathbb{N}_0$ und $k_1 + \cdots + k_n = k$. Dann lassen sich k unterschiedliche Teilchen auf genau*

$$\frac{k!}{k_1!k_2!\cdots k_n!} =: \binom{k}{k_1, k_2, \ldots, k_n}$$

verschiedenen Weisen auf n Zellen so verteilen, dass in der i-ten Zelle genau k_i Teilchen sind ($i = 1, \ldots, n$).

Satz 2.27 (Multinomialsatz)**.** *Für alle $k, n \in \mathbb{N}$ und für alle $x_1, \ldots, x_n \in \mathbb{R}$ gilt*

$$(x_1 + x_2 + \cdots + x_n)^k = \sum_{k_1+\cdots+k_n=k} \frac{k!}{k_1!k_2!\cdots k_n!} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n},$$

wobei das Summenzeichen die Summation über alle Tupel $(k_1, \ldots, k_n) \in \mathbb{N}_0 \times \cdots \times \mathbb{N}_0$ mit $\sum_{i=1}^{n} k_i = k$ meint.

Die Beweise zu diesen beiden Sätzen erfolgen später in den Übungsaufgaben.

3 Der Körper der reellen Zahlen

Wir definieren die uns geläufigen reellen Zahlen über ihre Eigenschaften ($\mathbb{R}$ ist ein archimedisch angeordneter, vollständiger Körper). Ein Beweis, dass ein solche Menge wirklich existiert, werden wir nicht machen.

3.1 Körperaxiome

Axiom 3.1 (Körperaxiome)**.** $\mathbb{R}$ ist eine Menge, auf der zwei Verknüpfungen definiert sind (Addition und Multiplikation):

$$+\colon \mathbb{R} \times \mathbb{R} \to \mathbb{R}, \quad (x,y) \mapsto x + y, \qquad\qquad \cdot\colon \mathbb{R} \times \mathbb{R}, \quad (x,y) \mapsto xy,$$

d. h. jedem Paar (x, y) mit $x, y \in \mathbb{R}$ sind genau ein Element $x + y \in \mathbb{R}$ (Summe) und genau ein Element $xy \in \mathbb{R}$ (Produkt) zugeordnet (aus Konvention lassen wir den Malpunkt weg, falls es den Sinn nicht verfälscht). Hierfür gelten die *Körperaxiome.*

(A1) $x + (y + z) = (x + y) + z$ für alle $x, y, z \in \mathbb{R}$ *(Assoziativgesetz).*

(A2) $x + y = y + x$ für alle $x, y \in \mathbb{R}$ *(Kommutativgesetz).*

(A3) Es gibt eine Zahl $0 \in \mathbb{R}$, sodass $x + 0 = x$ für alle $x \in \mathbb{R}$ *(Existenz der Null).*

(A4) Zu jedem $x \in \mathbb{R}$ existiert eine Zahl $-x \in \mathbb{R}$, sodass $x + (-x) = 0$ *(Existenz des Negativen).*

(A5) $x(yz) = (xy)z$ für alle $x, y, z \in \mathbb{R}$ *(Assoziativgesetz).*

(A6) $xy = yx$ für alle $x, y \in \mathbb{R}$ *(Kommutativgesetz).*

(A7) Es gibt eine Zahl $1 \in \mathbb{R} \setminus \{0\}$, sodass $x1 = x$ für alle $x \in \mathbb{R}$ *(Existenz der Eins).*

(A8) Zu jedem $x \in \mathbb{R} \setminus \{0\}$ existiert eine Zahl $x^{-1} \in \mathbb{R}$, sodass $xx^{-1} = 1$ *(Existenz des Inversen).*

(A9) $x(y + z) = xy + yz$ für alle $x, y, z \in \mathbb{R}$ *(Distributivgesetz).*

Weiterhin gelten folgende Regeln, um das Schreiben zu vereinfachen: Punkt- vor Strichrechnung (Klammern können wir weggelassen, falls wir mit dieser Regel den Sinn nicht verfälschen)

$$x - y := x + (-y), \qquad \frac{1}{x} := x^{-1}, \qquad \text{und} \qquad \frac{x}{y} = x/y = xy^{-1}.$$

3.1.1 Folgerungen aus den Körperaxiomen

Die Axiome (A1) bis (A4) besagen, dass $(\mathbb{R}, +)$ eine *Gruppe*[1] ist. Hieraus folgen insbesondere

(a) Die 0 ist eindeutig.

Beweis. Wäre 0^* ein weiteres neutrales Element, dann $0^* \overset{(A3)}{=} 0^* + 0 \overset{(A2)}{=} 0 + 0^* \overset{(A3)}{=} 0.$ □

(b) Die Negation ist eindeutig.

Beweis. Wären x' und x'' Zahlen, die $x + x' = 0 = x + x''$ erfüllen, dann

$$x' \overset{(A3)}{=} x' + 0 \overset{\text{Vor.}}{=} x' + (x + x'') \overset{(A1)}{=} (x' + x) + x'' \overset{(A2)}{=} (x + x') + x'' \overset{(A4)}{=} 0 + x'' \overset{(A2)}{=} x'' + 0 \overset{(A3)}{=} x''.$$ □

(c) $-0 = 0$.

Beweis. $0 \overset{(A4)}{=} 0 + (-0) \overset{(A2)}{=} (-0) + 0 \overset{(A3)}{=} -0.$ □

(d) $\forall x\colon -(-x) = x$.

Beweis. Aus $0 = x + (-x) = (-x) + x$ und der Eindeutigkeit des Negativen folgt $x = -(-x)$. □

(e) $\forall x, y, z \in \mathbb{R}\colon x + y = z \iff x = z - y$.

(f) $\forall x, y \in \mathbb{R}\colon -(x + y) = (-x) + (-y)$.

Die Axiome (A5) bis (A8) besagen, dass $(\mathbb{R} \setminus \{0\}, \cdot)$ eine *Gruppe* ist.[2] Daher gelten insbesondere

(a') Die 1 ist eindeutig.

(b') Das Inverse x^{-1} ist eindeutig.

(c') $1^{-1} = 1$.

(d') $(x^{-1})^{-1} = x$.

(e') $xy = z \iff x = zy^{-1}$ $(y \neq 0)$.

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

Mit dem Distributivgesetz folgt

(g) $\forall x, y, z \in \mathbb{R}\colon (x + y)z = z(x + y) = zx + zy = xz + yz$.

(h) $\forall x\colon x0 = 0$.

Beweis. $x0 = x(0 + 0) = x0 + x0$. Aus Punkt (e) folgt $x0 = 0$. □

(i) $\forall x, y\colon xy = 0 \iff (x = 0) \vee (y = 0)$ (*Nullteilerfreiheit*).

Beweis. Die Rückrichtung folgt aus Axiom (A6) und Punkt (h).
Für die Hinrichtung nehmen wir an, dass $xy = 0$ und $\neg(x = 0)$ ist. Wir zeigen $y = 0$. Mit den Punkten (e') und (h) gilt $y = 0$. Analoges gilt für $\neg(y = 0)$, wenn wir x und y vertauschen. □

Bemerkung 3.2.

[1] Die Gruppenaxiome sind: Abgeschlossenheit, Assoziativität, Existenz des neutralen Elements, Existenz eines inversen Elements zu jedem Element. Das neutrale Element sowie ein Element und sein Inverses sind kommutativ. In $\mathbb{R}$ gilt zusätzlich Kommutativität für alle Elemente, sodass $\mathbb{R}$ eine *abelsche Gruppe* ist.

[2] Damit $\mathbb{R}$ wirklich eine Gruppe ist, müssten wir noch die Abgeschlossenheit zeigen, d. h., dass das multiplikative Inverse eines jeden Elements nicht 0 ist, was nicht zu $\mathbb{R} \setminus \{0\}$ gehört. Das ist aber trivial mit einem Widerspruchsbeweis und Punkt (h).

- Jede Menge $\mathbb{K}$ mit Verknüpfungen $+$ und $\cdot$, für die Axiome (A1) bis (A9) gelten, heißen *Körper*.
- *Beispiele für Körper*: $\mathbb{R}, \mathbb{Q}, \mathbb{C}$ sowie $\mathbb{F}_p = \{0, 1, \ldots, p-1\}$ für Primzahlen p mit Addition und Multiplikation modulo p.

  Der Körper $\mathbb{F}_2 = \{0, 1\}$ für $p = 2$ ist der kleinstmögliche Körper. Die beiden Verknüpfungen werden definiert als

$y\backslash x$	0	1
0	0	1
1	1	0

$y\backslash x$	0	1
0	0	0
1	0	1

  wobei links die Addition und rechts die Multiplikation definiert ist. Eine Interpretation dafür ist die Verknüpfung von geraden und ungeraden Zahlen unter Addition und Multiplikation.

Ende der Vorlesung 3 am 18. Oktober 2021

Weitere Folgerungen sind

(j) $\forall x\colon -x = (-1)x$.

Beweis. $x + (-1)x = 1x + (-1)x = (1 + (-1))x = 0x = 0$. □

(k) $\forall x, y\colon (-x)(-y) = xy$.

Beweis. $(-x)(-y) = (-x)(-1)y = (-1)(-x)y = ((-1)(-x))y = xy$. □

3.1.2 Verallgemeinerte Assoziativität, Kommutativität und Distributivität

Bemerkung 3.3. Für alle $x, y, z \in \mathbb{R}$ gibt es 12 Möglichkeiten, die Summe der Zahlen zu bilden: $(x + y) + z$, $x + (y + z)$, $(z + x) + y$, … (3! = 6 Möglichkeiten mit jeweils zwei Möglichkeiten, Klammern zu setzen). Alle Möglichkeiten stellen dieselbe Zahl dar, die wir mit $x + y + z$ bezeichnen.

Das können wir verallgemeinern: Für alle $x \in \mathbb{N}$ und für alle $x_1, \ldots, x_n \in \mathbb{R}$ gilt das *allgemeine Assoziativgesetz*

$$x_1 + x_2 + \cdots + x_n = x_1 + (x_2 + (x_3 + \cdots + (x_{n-1} + x_n) \ldots)) = (\ldots((x_1 + x_2) + x - 3) + \cdots + x_{n-1}) + x_n.$$

und das *allgemeines Kommutativgesetz*

$$x_1 + x_2 + \cdots + x_n = x_n + x_{n-1} + \cdots + x_1. \tag{3.4}$$

Für (3.4) schreiben wir auch $\sum_{k=1}^{n} x_k = x_1 + \cdots + x_n$.

Aus dem allgemeinen Kommutativgesetz folgt

Satz 3.5 (Doppelsumme)**.** *Für alle $m, n \in \mathbb{N}$ und für alle $x_{ij} \in \mathbb{R}$ gilt*

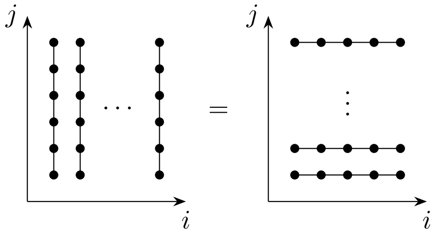
$$\sum_{i=1}^{m} \left(\sum_{j=1}^{n} x_{ij} \right) = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} x_{ij} \right),$$

d. h., dass voneinander unabhängige Summenzeichen vertauscht werden können.

Beweis (nicht sehr formal). Es gilt

$$\begin{aligned} \sum_{i=1}^{m} \sum_{j=1}^{n} x_{ij} &= \sum_{j=1}^{n} x_{1j} + \sum_{j=1}^{n} x_{2j} + \cdots + \sum_{j=1}^{n} x_{mj} \\ &= \sum_{i=1}^{m} x_{i1} + \sum_{i=1}^{m} x_{i2} + \cdots + \sum_{i=1}^{m} x_{in} = \sum_{j=1}^{n} \sum_{i=1}^{m} x_{ij} \end{aligned}$$

anhand der Anschauung



□

Bemerkung 3.6. Das *allgemeine Distributivgesetz* lautet

$$\left(\sum_{i=1}^{m} x_i\right)\left(\sum_{j=1}^{n} y_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} x_i y_j.$$

Definition 3.7 (Potenz)**.** Für alle $n \in \mathbb{N}_0$ und für alle $x \in \mathbb{R}$ sei die *Potenz* $x^n \in \mathbb{R}$ (gesprochen „x hoch n“) rekursiv definiert durch $x^0 := 1$ und $x^{n+1} := x \cdot x^n$.

Für $x \neq 0$ ist ferner $x^{-n} := \left(x^{-1}\right)^n$, wobei x^{-1} das inverse Element der Multiplikation ist.

Damit folgt

(l) $\forall m, n \in \mathbb{Z}$, $\forall x, y \in \mathbb{R}$ (x, y jeweils nicht 0, falls der zugehörige Exponent negativ ist) gilt $x^m x^n = x^{m+n}$, $(x^m)^n = x^{mn}$ und $x^n y^n = (xy)^n$.

3.2 Anordnungsaxiome

Axiom 3.8 (Anordnungsaxiome)**.** Es gibt eine Teilmenge $\{x \in \mathbb{R} : x > 0\}$ von positiven Zahlen mit

(A10) Für alle $x \in \mathbb{R}$ gilt genau eine der folgenden Aussagen: $x > 0$ (positiv), $x = 0$ (neutral), $x < 0$ (negativ) (*Trichotomie*).

(A11) Für alle $x, y \in \mathbb{R}$ gilt (*Abgeschlossenheit bzgl. der Addition und Multiplikation*)

$$x > 0 \wedge y > 0 \implies x + y > 0 \wedge xy > 0.$$

Ein Beispiel für einen Körper, der sich nicht anordnen lässt, ist der Körper der komplexen Zahlen $\mathbb{C}$. Für $x = y \neq 0$ gilt aufgrund der Abgeschlossenheit und den Körperaxiomen stets $x^2 > 0$, aber in $\mathbb{C}$ lassen sich immer Wurzeln ziehen, also insbesondere auch von negativen Zahlen $x^2 < 0$.

Bezeichnung 3.9. Wir definieren die Relationen $>$, $<$, $\geq$ und $\leq$ durch[3]

$$\begin{aligned}
x > y &\iff x - y > 0\\
x < y &\iff y > x\\
x \geq y &\iff (x > y \wedge x = y) \iff \neg(x < y)\\
x \leq y &\iff (x < y \wedge x = y) \iff \neg(x > y).
\end{aligned}$$

[3]Vielleicht verwirrt es, dass wir noch einmal die „>“-Relation definieren, obwohl es in den Axiomen schon definiert wurde. Müsste man das nicht beweisen?

Betrachten wir die Anordnungsaxiome genauer: Wir nehmen die Teilmenge aller Elemente aus $\mathbb{R}$ (Idee: Menge der positiven Zahlen), die einen *unären Operator* „> 0“ (gesprochen: „ist positiv“) definieren. Was dieser Operator ist, wird durch die beiden Axiome ausgedrückt: Addition und Multiplikation ist unter dieser Teilmenge abgeschlossen, und Elemente können entweder drin oder nicht drin liegen (exakter wäre zu definieren, dass nur $x > 0$, $x = 0$ oder $-x > 0$ gilt, da wir den „<“-Operator noch nicht definiert haben).

Anschließend definieren wir aus dem unären Operator die anderen *binären Operatoren*, die das Schreiben vereinfachen, also anstatt jedes mal $x - y > 0$ einfach kurz $x > y$ schreiben zu können etc.

3.2.1 Folgerungen aus den Anordnungsaxiomen

Offensichtlich liegt allen folgenden Variablen die Menge $\mathbb{R}$ zugrunde.

(m) $\forall x, y$ gilt genau eines von $x < y$, $x = y$, $x > y$.

(n) $x < y \wedge y < z \implies x < z$ (*Transitivität*).

Beweis. $x < y \wedge y < z \implies x - y > 0 \wedge z - y > 0 \implies z - x = (z - y) + (y - x) > 0$. □

(o) $x < y \implies \forall z\colon x + z < y + z$ (*Translation*).

Beweis. $(y + z) - (x + z) = y - x > 0$. □

(p) $x < y \implies (-x) > (-y)$ (*Spiegelung*).

Beweis. $(-x) - (-y) = y - x > 0$. □

(q) $x < y \wedge u < v \implies x + u < y + v$.

(r) $x < y \wedge t > 0 \implies tx < ty$.

(s) $x < y \wedge t < 0 \implies tx > ty$ (*Achtung*: Relationszeichen dreht sich um).

(t) $0 \leq x < y \wedge 0 \leq u < v \implies xu < yv$.

(u) $\forall x \neq 0\colon x > 0 \iff x^{-1} > 0,\ x < 0 \iff x^{-1} < 0$.

Beweis. Angenommen es gilt $x > 0 \wedge x^{-1} < 0$, dann $-1 = x\big(-x^{-1}\big) > 0$. Widerspruch. □

(v) $0 < x < y \implies x^{-1} > y^{-1}$.

(w) $\forall x \neq 0\colon x^2 > 0$, insbesondere $1 > 0$.

Beweis. Falls $x > 0$, dann gilt die Aussage wegen Axiom (A11). Falls $x < 0$, dann gilt die Aussage wieder wegen Axiom (A11), denn $x^2 = (-x)(-x) > 0 \cdot 0 = 0$. □

Bezeichnung 3.10 (nichtnegativ)**.** $x \in \mathbb{R}$ heißt *nichtnegativ*, falls $x \geq 0$ ist.

3.2.2 Archimedisches Axiom

Axiom 3.11 (ARCHIMEDISCHES Axiom)**.** Sei $\mathbb{N}$ in $\mathbb{R}$ eingebettet, d. h. $\mathbb{N} \subset \mathbb{R}$.

(A12) Für alle $x \in \mathbb{R}$ gibt es ein $n \in \mathbb{N}$, sodass $n > x$ gilt.[4]

Wir können daraus folgern:

(x) $\forall \varepsilon > 0\colon \exists n \in \mathbb{N}\colon \varepsilon > 1/n$. Diese Aussage wird für den Konvergenzbegriff wichtig.

(y) $\forall x \in \mathbb{R}, x \geq 0\colon \exists! n \in \mathbb{N}_0\colon n \leq x < n + 1$.[5] Dabei bedeutet $\exists!$ „existiert genau ein“.

Definition 3.12 (GAUSS-Klammer)**.** Für $n \in \mathbb{N}$ mit $n \leq x < n + 1$ schreiben wir $\lfloor x \rfloor \coloneqq n$ für alle $x \in \mathbb{R}$.

Satz 3.13 (BERNOULLISCHE Ungleichung)**.** *Für alle reellen $x \geq -1$ und für alle $n \in \mathbb{N}$ gilt*

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

[4] Traditionell (wie auch in vielen Lehrbüchern) wird das archimedische Axiom wie folgt äquivalent formuliert: Für alle positiven $x, y \in \mathbb{R}$ gibt es ein $n \in \mathbb{N}$, sodass $nx > y$ gilt.

[5] Ein Beweis auch für $x < 0$ und $n \in \mathbb{Z}$: Sei $A = \{m \in \mathbb{Z} \mid m \leq x\}$. A ist nicht leer, denn es gibt ein $m > -x$, weshalb $-m \in A$ gilt. Sei nun $n = \sup A$, woraus $n \leq x$ folgt. Aufgrund der Supremumseigenschaft ist $(n + 1) \notin A$, also $n + 1 > x$. Wichtig ist, dass das Supremum von unendlichen Mengen ganzer Zahlen definiert ist. Der Satz, so wie er in der Vorlesung genannt wurde, braucht kein Supremum unendlicher Mengen.

Der Beweis erfolgt in der Übung.
Einige Schlussfolgerungen aus Satz 3.13:

Satz 3.14.

1. *Für alle reellen $a > 1$ gibt es zu jedem $K \in \mathbb{R}$ ein $n \in \mathbb{N}$, sodass $a^n > K$ gilt.*

2. *Für alle reellen $0 < a < 1$ gibt es zu jedem $\varepsilon > 0$ ein $n \in \mathbb{N}$, sodass $a^n < \varepsilon$ gilt.*

Beweis.

1. Wir wenden die bernoullische Ungleichung auf $x = a - 1 > 0$ an: $a^n = (1 + x)^n \geq 1 + nx$. Mit dem archimedischen Axiom gibt es ein natürliches $n > (K - 1)/x$, woraus $a^n > 1 + (K - 1)/x \cdot x = K$ folgt.

2. Wir wenden Punkt 1 auf $1/a$ statt a und $K = 1/\varepsilon$ an. □

3.2.3 Absolutbetrag

Definition 3.15 (Absolutbetrag)**.** Für $x \in \mathbb{R}$ definieren wir

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

und sagen dafür „Betrag von x“ oder „x absolut“.

Es ist besser, wenn die Fälle der abschnittsweise definierten Funktion sich nicht überschneiden, wie es hier zu sehen ist (wir könnten auch $|x| = -x$ für $x \leq 0$ definieren).

Definition 3.16 (Maximum und Minimum)**.** Ferner definieren wir für $x, y \in \mathbb{R}$

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

Gelegentlich tauchen auch folgende Schreibweisen auf:

$$x \vee y := \max(x, y), \qquad x \wedge y := \min(x, y).$$

Offenbar gilt $|x| = \max(x, -x)$, d. h., dass der Betrag immer größer als die Zahl und ihr Negatives ist. Das ist hilfreich für Beweise mit Beträgen.

Satz 3.17 (Betragsfunktion, Dreiecksungleichung)**.** *Für die Betragsfunktion $|\cdot|$ gilt:*

1. *Ist $x \in \mathbb{R}$, so gilt $|x| \geq 0$ und $|x| = 0$ genau dann, wenn $x = 0$ ist.*

2. *Für alle $x, y \in \mathbb{R}$ gilt $|xy| = |x| \cdot |y|$.*

3. *Für alle $x, y \in \mathbb{R}$ gilt $|x + y| \leq |x| + |y|$, was als* Dreiecksungleichung *bekannt ist.*

Jeder Körper mit einer Abbildung $|\cdot|\colon K \to \mathbb{R}_+$ und den drei obigen Aussagen heißt *bewerteter Körper*.

Beweis.

1. Die Aussage ist trivial.

2. Die Aussage ist für $x, y \geq 0$ trivial. Für allgemeine x, y wählen wir $\bar{x}, \bar{y} \geq 0$ mit $x = \pm\bar{x}$ und $y = \pm\bar{y}$. Dann gilt $|xy| = |\pm\bar{x}\bar{y}| = |\bar{x}\bar{y}| = |\bar{x}| \cdot |\bar{y}| = |x| \cdot |y|$.

3. Wegen $x \leq |x|$ und $y \leq |y|$ gilt $x + y \leq |x| + |y|$. Ferner ist $-x \leq |x|$ und $-y \leq |y|$, daher haben wir auch $-(x + y) = (-x) + (-y) \leq |x| + |y|$. Aus beiden Ungleichungen folgt dann $|x + y| \leq |x| + |y|$. □

4 Konvergenz von Folgen I

Die Konvergenz reeller Zahlenfolgen wird ein zentrales Thema in Analysis I sein.

4.1 Begriffe Folge, Konvergenz, Grenzwert

Definition 4.1 (Folge)**.** Eine *Folge* reeller Zahlen ist eine Abbildung $\mathbb{N} \to \mathbb{R}$, $n \mapsto a_n$ und schreiben dafür $(a_n)_{n\in\mathbb{N}} = (a_1, a_2, a_3, \dots)$ oder kurz (a_n).

Im Allgemeinen muss der Index nicht bei $n = 1$ beginnen.

Definition 4.2 (erweiterter Folgenbegriff)**.** Eine Folge ist allgemein eine Abbildung $\{k, k+1, k+2, \dots\} \to \mathbb{R}$ für $k \in \mathbb{Z}$. Wir schreiben kurz $(a_n)_{n\geq k} = (a_k, a_{k+1}, a_{k+2}, \dots)$.

Meistens beginnen Zahlenfolgen aber bei $k = 0$ oder $k = 1$.

Beispiel 4.3.

- $(1/n)_{n\in\mathbb{N}} = (1, \frac{1}{2}, \frac{1}{3}, \dots)$.
- $\left((-1)^n\right)_{n\in\mathbb{N}} = (1, -1, 1, -1, \dots)$.
- $(2^n)_{n\in\mathbb{N}_0} = (1, 2, 4, 8, \dots)$.
- Die Folge der Fibonacci-Zahlen $(f_n)_{n\in\mathbb{N}_0}$ mit

$$f_0 = 0, \quad f_1 = 1, \quad f_n = f_{n-1} + f_{n-2},$$

also $(0, 1, 1, 2, 3, 5, 8, 13, 21, \dots)$.

Ende der Vorlesung 4 am 20. Oktober 2021

Definition 4.4 (Konvergenz)**.** Eine Folge $(a_n)_{n\in\mathbb{N}}$ *konvergiert gegen eine Zahl* $a \in \mathbb{R}$, falls es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, für das $|a_n - a| < \varepsilon$ für alle $n \geq N$ gibt. Oder in Zeichen:

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

Eine Folge heißt *konvergent* falls sie gegen ein $a \in \mathbb{R}$ konvergiert. Falls das nicht der Fall ist, heißt (a_n) *divergent*.

Bemerkung 4.5. Das N ist hierbei als $N = N(\varepsilon)$ zu verstehen, da offensichtlich die Wahl von N vom vorgegebenen ε abhängen wird. Allgemein bedeutet das auch, dass N größer wenn ε kleiner wird. D. h. auch, dass in der Definition ε beliebig durch z. B. $\frac{1}{4}\varepsilon$ ersetzt werden könnte, da wir dann einfach das neue $N(\varepsilon)$ als das alte $N\left(\frac{1}{4}\varepsilon\right)$ definieren. Genauso äquivalent wären $\frac{1}{10}\varepsilon$, ε^2, etc. statt ε. In der Praxis kommt es beim Probieren häufig vor, dass nicht direkt ε rauskommt, sondern eben $\frac{1}{10}\varepsilon$, ε^2, etc.

Definition 4.6 (Grenzwert, Limes)**.** Das a in Definition 4.4 heißt *Grenzwert* oder *Limes* der Folge (a_n), in Zeichen

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

Diese Definition zu verstehen ist sehr wichtig. Wir sagen, dass wir für *beliebig kleine* ε stets ein Folgeglied finden können, wo ab diesem Glied alle folgenden Glieder näher als ε am Grenzwert a sind. Oft spricht man auch von einer *ε-Umgebung* um a: Eine Folge (a_n) ist genau dann konvergent, wenn für alle $\varepsilon > 0$ *schließlich alle* Folgenglieder in der Umgebung $\{x \in \mathbb{R} \colon a - \varepsilon < x < a + \varepsilon\}$ liegen. Das können wir auch geometrisch in Abb. 4.1 veranschaulichen. Da ε beliebig klein werden kann, müssen alle Folgenglieder (für hinreichend große n) beliebig nah an a kommen.

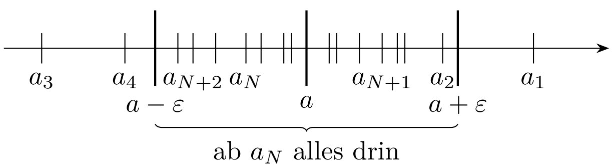


Abbildung 4.1: Visualisierung des Konvergenzbegriffs.

Definition 4.7 (Nullfolge)**.** Eine Folge heißt *Nullfolge*, falls sie gegen 0 konvergiert.

Satz 4.8. *Der Grenzwert einer konvergenten Folge ist eindeutig.*

Beweis durch Widerspruch. Die Annahme ist, dass es zwei Grenzwerte a und $b \neq a$ gibt. Um die Annahme zu widerlegen, reicht es, ein ε als Gegenbeispiel zu finden. Dafür setzen wir $\varepsilon = \frac{1}{2}|a-b| > 0$. Also existieren N_a und N_b mit $|a_n - a| < \varepsilon$ für alle $n \geq N_a$, und $|a_n - b| < \varepsilon$ für alle $n \geq N_b$. Wir setzen nun $N \coloneqq \max(N_a, N_b)$, d. h., dass mit der Dreiecksungleichung für alle $n \geq N$ dann

$$|a-b| = |(a-a_n) + (a_n - b)| \leq |a - a_n| + |a_n - b| < \varepsilon + \varepsilon = 2\varepsilon = |a-b|$$

gilt, was aber $|a-b| < |a-b|$ bedeuten würde. Ein Widerspruch. □

Alternativ hätten wir auch $\varepsilon = |a-b|/4$ wählen können, was dann zum Widerspruch $|a-b| < \frac{1}{2}|a-b|$ führt. Ähnlich sind weitere Variationen des Beweises möglich.

Beispiel 4.9.

1. Die konstante Folge mit $a_n \coloneqq b$ für alle n konvergiert gegen b.
2. Die alternierende Folge $\left((-1)^n\right)_{n\in\mathbb{N}_0} = (1, -1, 1, -1, \dots)$ divergiert.

   Beweis. Wenn die Folge doch gegen $a \in \mathbb{R}$ konvergiert, können wir $\varepsilon = 1$ wählen, wozu es ein N gibt mit $|a_n - a| < 1$ und natürlich auch $|a_{n+1} - a| < 1$ für alle $n \geq N$. Daraus folgt $2 > |a_n - a| + |a - a_{n+1}| \geq |(a_n - a) + (a - a_{n+1})| = |a_n - a_{n+1}| = 2$, Widerspruch. □

3. $(1/n)_{n\in\mathbb{N}}$ ist eine Nullfolge.

   Beweis. Für jedes $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$ mit $N > 1/\varepsilon$ (archimedisches Axiom, Punkt (x)). Daher gilt für alle $n \geq N$: $|1/n - 0| \leq 1/N < \varepsilon$, also ist 0 der Grenzwert. □

4. $\left(2^n\right)_n$ und die Fibonacci-Folge sind divergent, da beide unbeschränkt sind.
5. Wir betrachten die Folge $\left(x^n\right)_{n\in\mathbb{N}}$ mit $x \in \mathbb{R}$. Dann gibt es vier Fälle:
   (a) $x = 1$ liefert eine konstante Folge, die gegen 1 konvergiert.
   (b) $x = -1$ liefert die alternierende Folge, die divergiert.
   (c) $|x| > 1$ liefert eine unbeschränkte Folge, die divergiert.
   (d) $|x| < 1$ ist eine Nullfolge.

   Beweis zu (d). Aufgrund Satz 3.14 gibt es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$, für das $|x|^N < \varepsilon$ gilt. Daher ist $|x^n - 0| = |x|^n \leq |x|^N < \varepsilon$. □

Definition 4.10 (beschränkte Folge)**.** $(a_n)_{n\in\mathbb{N}}$ heißt *beschränkt*, falls es ein $K \in \mathbb{R}$ gibt, sodass $|a_n| \leq K$ für alle $n \in \mathbb{N}$ gilt. Analog ist (a_n) *nach oben beschränkt*, falls es ein K gibt, sodass $a_n \leq K$ für alle n ist.

Beispiel 4.11. $\left(x^n\right)_n$ mit $|x| > 1$ ist unbeschränkt, weil es nach Satz 3.14 ein $K \in \mathbb{R}$ gibt, sodass $|x|^n > K$ für alle n ist.

Satz 4.12. *Jede konvergente Folge ist beschränkt.*

Beweis. Wir nehmen an, dass (a_n) gegen a konvergiert. Wählen wir $\varepsilon = 1$, so gibt es ein N mit $|a_n - a| < 1$ für alle $n \geq N$, also gilt nach der Dreiecksungleichung

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

Setzen wir $K = \max\{|a_1|, |a_2|, \dots, |a_{n-1}|, 1 + |a|\}$, gilt dann $|a_n| \leq K$ für alle $n \in \mathbb{N}$. □

4.2 Grenzwertsätze

Satz 4.13 (Summe/Differenz und Produkt konvergenter Folgen)**.** *Sind (a_n) und (b_n) konvergent, so konvergiert auch ihre Summe $(a_n + b_n)$ und ihr Produkt $(a_n b_n)$ und es gilt*

$$\lim_{n\to\infty}(a_n + b_n) = \Big(\lim_{n\to\infty} a_n\Big) + \Big(\lim_{n\to\infty} b_n\Big), \qquad \lim_{n\to\infty}(a_n b_n) = \Big(\lim_{n\to\infty} a_n\Big)\Big(\lim_{n\to\infty} b_n\Big).$$

Die Subtraktion folgt aus $(\lim_{n\to\infty} a_n) + (\lim_{n\to\infty}(-1))(\lim_{n\to\infty} b_n)$

Beweis. Zuerst zur Summenfolge: Es konvergieren $a_n \to a$ und $b_n \to b$ für $n \to \infty$. Nach Voraussetzung gibt es zu jedem $\varepsilon > 0$ Zahlen $N_a, N_b \in \mathbb{N}$, sodass $|a_n - a| < \frac{1}{2}\varepsilon$ für alle $n \geq N_a$, und $|a_n - b| < \frac{1}{2}\varepsilon$ für alle $n \geq N_b$ gilt. Setzen wir nun $N := \max(N_a, N_b)$, erhalten wir mit der Dreiecksungleichung für alle $n \geq N$

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

(Alternativ können wir auch $|a_n - a| < \varepsilon$ und $|b_n - b| < \varepsilon$ nehmen, woraus $|(a_n + b_n) - (a + b)| < 2\varepsilon$ folgt.)

Jetzt zur Produktfolge: Da (a_n) konvergiert, existiert eine obere Schranke $K > 0$, sodass $|a_n| < K$ für alle n (Satz 4.12). Nach Voraussetzung gibt es für jedes $\varepsilon > 0$ Zahlen $N_a, N_b \in \mathbb{N}$, sodass $|a_n - a| < \varepsilon$ für alle $n \geq N_a$ und $|b_n - b| < \varepsilon$ für alle $n \geq N_b$ gilt. Mit der Dreiecksungleichung folgt daraus

$$|a_n b_n - ab| = |a_n b_n - a_n b + a_n b - ab| = |a_n(b_n - b) + (a_n - a)b| \leq |a_n| \cdot |b_n - b| + |a_n - a| \cdot |b| < K\varepsilon + \varepsilon|b| = \varepsilon(K + |b|).$$

Da $K + |b|$ eine Konstante ist und ε beliebig klein werden kann, kann auch $\varepsilon(K + |b|)$ beliebig klein werden.

Nun nochmals zurück, um den Beweis besser aufzuschreiben: Wir wählen nun N so, dass für alle $n \geq N$ gilt:

$$|a_n - a| < \frac{\varepsilon}{K + |b|}, \quad |b_n - b| < \frac{\varepsilon}{K + |b|} \implies \cdots \implies |a_n b_n - ab| < \frac{\varepsilon}{K + |b|}(K + |b|) = \varepsilon.$$ □

Beispiel 4.14. Die geometrische Reihe konvergiert für $x \in \mathbb{R}$ mit $|x| < 1$ gegen

$$\lim_{n\to\infty} \frac{1 - x^{n+1}}{1 - x} = \frac{1}{1 - x} - \frac{1}{1 - x} \lim_{n\to\infty} x^{n+1} = \frac{1}{1 - x}.$$

Bevor wir die Division zweier konvergenter Folgen betrachten, brauchen wir noch eine alternative Form der Dreiecksungleichung.

Lemma 4.15 (umgekehrte Dreiecksungleichung)**.** *Für alle $x, y \in \mathbb{R}$ gilt*

$$|x - y| \geq |x| - |y|.$$

Beweis. Für alle reellen Zahlen a, b gilt gemäß der Dreiecksungleichung $|a| \geq |a + b| - |b|$. Nach einsetzen von $a = x + z$ und $b = -z$ haben wir $|x + z| \geq |x| - |z|$ (auch eine nützliche Ungleichung). Setzen wir nun $z = -y$, erhalten wir das Lemma.[1] □

Satz 4.16 (Quotient konvergenter Folgen)**.** *Seien (a_n) und (b_n) konvergente Folgen mit $\lim_{n\to\infty} b_n = b \neq 0$. Dann existiert ein $k \in \mathbb{N}$ mit $b_n \neq 0$ für alle $n \geq k$, und die Folge $(a_n/b_n)_{n\geq k}$ konvergiert mit dem Limes*

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

Bemerkung 4.17. Die Existenz des oben genannten k ist gesichert, weil wir ein $0 < \varepsilon < |b|$ wählen können, und es dafür ein $k \in \mathbb{N}$ geben muss mit $|b_n - b| < \varepsilon < |b|$ für alle $n \geq k$. Diese b_n sind dann entweder alle positiv oder negativ, je nachdem ob b positiv oder negativ ist. Mit anderen Worten: Wir warten so lange, bis k hinreichend groß ist und die b_n in der Nähe von b liegen, also ungleich Null sind.

Beweis. Wir versuchen den Sachverhalt auf bekanntes zurückzuführen: Konvergiert $(1/b_n)_{n\geq k}$ gegen $1/b$, dann folgt der Rest aus dem vorherigen Satz über Produktfolgen: $a_n/b_n = a \cdot (1/b_n)$.

[1]Die „Mutter“ aller Dreiecksungleichungen ist auch sehr nützlich: $||x| - |y|| \leq |x \pm y| \leq |x| + |y|$. Das haben wir in der Übung bewiesen.

Nun zur Konvergenz von $(1/b_n)$: Nach Voraussetzung ist $b \neq 0$. Mit $\varepsilon = \frac{1}{2}|b|$ folgt, dass ein $N \geq k$ existiert, sodass

$$|b_n| \geq |b| - |b_n - b| > |b| - \varepsilon = \frac{1}{2}|b|$$

für alle $n \geq N$. Dabei folgte die erste Abschätzung aus Lemma 4.15, die zweite aus der Konvergenz von $b_n \to b$. Ferner gibt es zu jedem $\varepsilon > 0$ ein $N' \in \mathbb{N}$, sodass $|b_n - b| < \frac{1}{2}\varepsilon|b|^2$ für alle $n \geq N'$ gilt. Daher haben wir für alle $n \geq \max(N, N')$

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

Alternativ hätten wir für die letzte Abschätzung auch $|1/b_n - 1/b| < \varepsilon(|b| - \varepsilon)/|b|$ erhalten können.[2]

Beispiel 4.18. Um den Grenzwert rationaler Funktionen in n können wir bestimmen, indem wir den größten Exponenten aus Zähler und Nenner ausklammern und dann kürzen. Viele Summanden reduzieren sich dann zu Nullfolgen.

$$\lim_{n\to\infty} \frac{7n^2 + 13n + 24}{n^2 - 2n - 8} = \lim_{n\to\infty} \frac{7 + 13/n + 24/n^2}{1 - 2/n - 8/n^2} = \frac{\lim_{n\to\infty}\left(7 + 13/n + 24/n^2\right)}{\lim_{n\to\infty}(1 - 2/n - 8/n^2)} = \frac{7+0+0}{1-0-0} = 7.$$

Beachte, dass der Nenner $1 - 2/n - 8/n^2 \neq 0$ für natürliche n ist.

Satz 4.19 (Ordnungserhaltung konvergenter Folgen)**.** *Falls (a_n) und (b_n) konvergent sind, gilt dann*

$$\forall n\colon a_n \leq b_n \implies \lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n.$$

Beweis. Wir betrachten die Folge $c_n := b_n - a_n \geq 0$. Wäre $c := \lim_{n\to\infty} c_n < 0$, so folgt aus der Konvergenz, dass es ein $\varepsilon = -c > 0$ gibt, sodass $|c_n - c| < \varepsilon = -c$ gilt. Weil $c_n \geq 0 > c$, können wir die Betragsstriche weglassen und erhalten $c_n - c < -c \iff c_n < 0$, ein Widerspruch.[3] □

Korollar 4.20. *Falls $K \leq a_n \leq L$ für alle $n \in \mathbb{N}$ ist, gilt dann $K \leq \lim_{n\to\infty} a_n \leq L$.*

Das Korollar erhalten wir durch Abschätzung einer Folge durch konstante Folgen $(K)_n$ und $(L)_n$ von oben und unten.

Achtung: Satz 4.19 und Korollar 4.20 gelten nicht mit $<$ an Stelle von $\leq$. *Gegenbeispiel*: Es gilt $a_n = 1/n > 0$ für alle $n \in \mathbb{N}$, aber $\lim_{n\to\infty} a_n = 0$.

Ende der Vorlesung 5 am 25. Oktober 2021

5 Das Vollständigkeitsaxiom

5.1 Motivation

Mit den bisherigen Axiomen können wir aus den natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, \ldots\}$ nacheinander die uns bekannten Zahlenbereiche konstruieren.

Die Menge der ganzen Zahlen $\mathbb{Z} = \mathbb{N} \cup \{0\} \cup -\mathbb{N}$ lässt sich dann als Menge der Äquivalenzklassen $\mathbb{Z} = \{m - n : m, n \in \mathbb{N}\}/\sim$ mit der Äquivalenzrelation $m - n \sim m' - n'$, falls $m + n' = m' + n$, definieren.[1] Die

[2] Noch ein hilfreicher Grenzwertsatz ist die *k-te Wurzel einer konvergenten Folge*: Ist $a_n \to a$ nichtnegativ, so gilt $\sqrt[k]{a_n} \to \sqrt[k]{a}$ für alle $k \in \mathbb{N}$. *Beweis*: Generell gilt für $x \geq y \geq 0$ mit dem binomischen Lehrsatz und anschließendem Weglassen von positiven Summanden: $x^k = (y+(x-y))^k \geq y^k + (x-y)^k \iff x^k - y^k \geq (x-y)^k$. Um den Fall $x \leq y$ zu beachten, können wir Betragsstriche setzen: $|x^k - y^k| \geq |x - y|^k$. Ersetzen wir $x = \sqrt[k]{a_n}$ und $y = \sqrt[k]{a}$, erhalten wir nach umstellen $|\sqrt[k]{a_n} - \sqrt[k]{a}| \leq \sqrt[k]{|a_n - a|}$. Mit $|a_n - a| \leq \varepsilon^k$ folgt die Behauptung.

[3] Der Dozent gab einen (eher intuitiven) Beweis, dass c_n für große n in eine $\frac{1}{2}|c|$-Umgebung kommen muss, also $c_n \leq \frac{1}{2}c < 0$.

[1] Eine kurze (unsaubere) Definition von Äquivalenzrelationen: Eine *Äquivalenzrelation* $\sim$ über eine Menge M ist eine Relation, für die noch zusätzlich $a \sim a$ (*Reflexivität*), $a \sim b \iff b \sim a$ (*Symmetrie*) und $(a \sim b \land b \sim c) \implies a \sim c$ (*Transitivität*) für alle $a, b, c \in M$ erfüllt. Intuitiv sind das auch Eigenschaften des Begriffs „Äquivalenz“. Eine *Äquivalenzklasse* ist nun die Menge aller Elemente, die zueinander äquivalent sind, also $[a]_\sim := \{b \in M : a \sim b\}$. Die Menge aller *Äquivalenzklassen* $M/\sim$ ist $M/\sim = \{[a]_\sim : a \in M\}$.

Äquivalenzrelation sagt, wann zwei Darstellungen $m - n$ und $m' - n'$ dieselbe Zahl repräsentieren, nämlich wenn dieselbe Differenz rauskommt. $\mathbb{Z}$ ist nun die Menge aller dieser Differenzen.

Ähnlich lässt sich die Menge der rationalen Zahlen $\mathbb{Q}$ definieren als die Menge der Äquivalenzklassen $\mathbb{Q} = \{m/n : m \in \mathbb{Z}, n \in \mathbb{Z}\}/\sim$ mit der Äquivalenzrelation $m/n \sim m'/n'$, falls $mn' = m'n$ gilt (Brüche bleiben nach Kürzen bzw. Erweitern dieselben). $\mathbb{Q}$ ist nun die Menge aller dieser Quotienten.

$\mathbb{Q}$ und $\mathbb{R}$ sind nun beides archimedisch angeordnete Körper. Aber wir wissen, dass $\mathbb{Q}$ nicht „vollständig“ ist, es weist „Löcher“[2] auf. Es gibt nämlich „Zahlen“, die nicht in $\mathbb{Q}$ enthalten sind, wie z. B Wurzeln wie $\sqrt{2} \notin \mathbb{Q}$. Denn wäre $\sqrt{2} = m/n \in \mathbb{Q}$, dann ist $2m^2 = n^2$. In der Primfaktorzerlegung der linken Seite steht eine Zweierpotenz mit ungeradem Exponenten, während auf der Zerlegung der rechten Seite ein gerader Exponent auftritt, was unmöglich ist.

Das Ziel ist, die Löcher von $\mathbb{Q}$ zu stopfen. Dazu gibt es verschiedene Ansätze:

1. *Cauchy-Folgen*,
2. *Intervallschachtelungen* und
3. *Dedekindsche Schnitte*, was in der Übung erfolgt.

5.2 Cauchy-Folgen

Definition 5.1 (Cauchy-Folge)**.** Eine Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen heißt *Cauchy-Folge*, wenn es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, für das $|a_m - a_n| < \varepsilon$ für alle $m, n \geq N$ gilt.

Bemerkung 5.2.

1. Die Idee hier ist, den Begriff der Konvergenz zu definieren, ohne den Limes zu verwenden. Anstatt zu sagen, dass Folgenglieder sich einem Grenzwert annähern, können Folgenglieder sich gegenseitig annähern. Diese Definition ist hilfreich, denn nicht für alle Folgen ist der Grenzwert gemäß den Grenzwertsätzen, Null- und konstanten Folgen definiert, aber die Cauchy-Folge, wie z. B. für die rekursive Folge $a_{n+1} = 1 + 1/a_n$.
2. Es reicht <u>nicht</u>, hier nur $|a_n - a_{n+1}| < \varepsilon$ für zwei aufeinanderfolgende Glieder zu verifizieren. *Beispiel*: Für die *harmonische Reihe* $a_n := \sum_{k=1}^{n} 1/k$ gilt $|a_n - a_{n+1}| = 1/(n+1) < \varepsilon$ für hinreichend große n aufgrund des archimedischen Axioms. Trotzdem divergiert die Reihe.[3]

Satz 5.3. *Jede konvergente Folge reeller Zahlen ist eine Cauchy-Folge.*

Beweis. Angenommen (a_n) konvergiert gegen $a \in \mathbb{R}$. Dann gibt es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$, sodass $|a_n - a| < \frac{1}{2}\varepsilon$ für alle $n \geq N$ gilt. Daraus folgt wegen der Dreiecksungleichung $|a_n - a_m| \leq |a_n - a| + |a - a_m| < \varepsilon$ für alle $n, m \geq N$. □

5.3 Vollständigkeitsaxiom (Vollständigkeit in $\mathbb{R}$)

Mit Satz 5.3 haben wir gezeigt, dass jede konvergente Folge eine Cauchy-Folge ist, also Cauchy-Folgen unserer Idee von Konvergenz entsprechen. Die Umkehrung müssen wir aber als Axiom voraussetzen.

Axiom 5.4 (Vollständigkeitsaxiom)**.** Jede Cauchy-Folge reeller Zahlen konvergiert in $\mathbb{R}$.

Beispiel 5.5. Wir betrachten die Dezimalbruchentwicklung einer reellen Zahl $x \in \mathbb{R}$ mit $x \geq 0$. Setze dafür $x_n := 10^{-n}\lfloor 10^n x\rfloor$ als die Zahl, die dadurch entsteht, indem wir nach der n-ten Nachkommastelle alle Ziffern abschneiden. (x_n) liefert eine Cauchy-Folge in $\mathbb{Q}$, denn $|x_n - x_m| \leq 10^{-N}$ für alle $n, m \geq N$. Die Folge (x_n) konvergiert in $\mathbb{Q}$ genau dann, wenn $x \in \mathbb{Q}$ ist.

Mit dem obigen Beispiel sehen wir, wie wir aus $\mathbb{Q}$ mithilfe von rationalen Cauchy-Folgen jede reelle Zahl konstruieren können: Sei $\hat{\mathbb{Q}} := \{(a_n) : \text{Cauchy-Folgen in } \mathbb{Q}\}$ mit der Äquivalenzrelation $((a_n) \sim (b_n)) :\Longleftrightarrow ((a_n - b_n) \text{ ist Nullfolge})$. Dann ist $\mathbb{R} := \hat{\mathbb{Q}}/\sim$, wobei $\mathbb{Q}$ in $\mathbb{R}$ eingebettet wird, indem wir jedes $q \in \mathbb{Q}$ mit der konstanten Folge $(q, q, q, \dots)$ identifizieren.

[2] „Löcher“ ist ein irreführendes Wort. Genau wie $\mathbb{R}$ ist auch $\mathbb{Q}$ „unendlich dicht“ (zwischen zwei „Zahlen“ gibt es stets eine „dazwischen“). Der Unterschied zwischen den beiden liegt vielmehr in der *Abzählbarkeit*.

[3] Das kann man zeigen, indem man geschickt abschätzt und somit zeigt, dass die Folge nach oben unbeschränkt ist:

$$1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) + \dots > 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \dots = 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} + \cdots$$

5.4 Intervallschachtelung

Definition 5.6 (Intervall)**.** Für $a, b \in \mathbb{R}$ mit $a \leq b$ sind

- $[a, b] \coloneqq \{x \in \mathbb{R} : a \leq x \leq b\}$ das *abgeschlossen Intervall*,
- $(a, b) \coloneqq \{x \in \mathbb{R} : a < x < b\}$ das *offene Intervall* und
- $[a, b)$ sowie $(a, b]$ *halboffene Intervalle.*

a und b heißen *Randpunkte* der jeweiligen Intervalle. $|b - a|$ heißt *Länge* des Intervalls, kurz $\operatorname{diam}([a, b])$.

Definition 5.7 (Intervallschatelung)**.** Eine *Intervallschachtelung* ist eine Folge $(I_n)_{n\in\mathbb{N}}$ von *abgeschlossenen* Intervallen $I_n = |a_n, b_n|$ mit

1. $I_{n+1} \subset I_n$ für alle $n \in \mathbb{N}$, und
2. $\lim_{n\to\infty} \operatorname{diam}(I_n) = 0$.

Die Idee ist, dass wir einen Grenzwert von oben und unten eingrenzen, und diese Intervalle immer kleiner werden lassen, bis das Intervall dem Grenzwert „entspricht“.

Satz 5.8 (Intervallschachtelungsprinzip)**.** *Für jede Intervallschachtelung $(I_n)_{n\to\infty}$ in $\mathbb{R}$ gibt es eine reelle Zahl $x \in \bigcap_{n\in\mathbb{N}} I_n$. Oder mit anderen Worten: Es gibt ein $x \in \mathbb{R}$, sodass $x \in I_n$ für alle $n \in \mathbb{N}$ gilt.*[4]

Die Beweisidee ist, die linken und rechten Randpunkte der Intervalle als zwei monotone Folgen aufzufassen, wobei die linken Randpunkte eine Cauchy-Folge $a_n \to a$ implizieren. Nun können wir aber a zwischen beiden Folgen „schachteln“.

Beweis. Gegeben sei die Intervallschachtelung $(I_n)_n$ mit $I_n = [a_n, b_n]$, $(a_n)_n$ eine nicht fallende Folge, $(b_n)_n$ eine nicht wachsende Folge und $\operatorname{diam}(I_n) = (b_n - a_n)_n$ eine Nullfolge.

Dann ist (a_n) eine Cauchy-Folge, denn aufgrund der Nullfolge gibt es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit $b_N - a_N = |(b_N - a_N) - 0| < \varepsilon$. Wegen der Monotonie gilt $a_N \leq a_n$ und $a_m \leq b_n \leq b_N$ für alle $m, n \geq N$. Daher ist $|a_n - a_m| \leq b_N - a_N < \varepsilon$, weshalb (a_n) eine Cauchy-Folge ist. Das bedeutet aber auch, dass nach dem Vollständigkeitsaxiom ein Grenzwert $a_n \to a \in \mathbb{R}$ für $n \to \infty$ existiert.

Da die Ordnungsrelation $\leq$ unter Konvergenz erhalten bleibt (Korollar 4.20), gilt $a_N \leq a \leq b_N$ für alle $N \in \mathbb{N}$, woraus auch $a \in \bigcap_{N\in\mathbb{N}} I_N$ folgt. □

Bemerkung 5.9.

1. Das Intervallschachtelungsprinzip gilt <u>nicht</u> für offene Intervalle. Z. B. gilt $\bigcap_{n\in\mathbb{N}}\big(0, n^{-1}\big) = \varnothing$.
2. Die Zahl $x \in \bigcap_{n\in\mathbb{N}} I_n$ bei Schachtelung abgeschlossener Intervalle ist eindeutig. Für jedes weitere $x' \in \bigcap I_n$ würde nämlich $|x - x'| \leq \operatorname{diam}(I_n) \to 0$ für $n \to \infty$ gelten, also $|x - x'| = 0$ und damit $x = x'$.[5]

Tatsächlich gilt für alle archimedisch angeordneten Körper

Satz 5.10. *Das Vollständigkeitsaxiom ist äquivalent zum Intervallschachtelungsprinzip.*

Beweis. Die Hinrichtung haben wir mit dem Intervallschachtelungsprinzip schon bewiesen. Es bleibt noch die Rückrichtung zu zeigen.

Sei (a_n) eine Cauchy-Folge, weshalb für alle $k \in \mathbb{N}$ ein $N_k \in \mathbb{N}$ existiert, sodass $|a_n - a_m| < \varepsilon = 2^{-k}$ für alle $m, n \geq N_k$ gilt.

Wir definieren die Intervalle

$$I_k \coloneqq \left[a_{N_k} - 2^{-k+1}, a_{N_k} + 2^{-k+1}\right]$$

für alle $k \in \mathbb{N}$. Nun ist die Folge $(I_k)_{k\in\mathbb{N}}$ eine Intervallschachtelung aus den folgenden Gründen: 1. Die I_k sind abgeschlossen. 2. Des Weiteren ist $I_{k+1} \subset I_k$, denn für alle $x \in I_{k+1}$ gelten $\left|x - a_{N_{k+1}}\right| \leq 2^{-k}$, weil $a_{N_{k+1}}$ der Mittelpunkt von I_{k+1} ist, und $\left|a_{N_{k+1}} - a_{N_k}\right| < 2^{-k}$ wegen des Cauchy-Kriteriums. Deshalb folgt

$$|x - a_{N_k}| \leq \left|x - a_{N_{k+1}}\right| + \left|a_{N_{k+1}} - a_{N_k}\right| < 2^{-k} + 2^{-k} = 2^{-k+1},$$

[4] Dieser Satz ist äquivalent zum *Einschnürungssatz*, bewiesen durch Gauss: Ist $a_n \leq c_n \leq b_n$ für alle $n \in \mathbb{N}$ und $\lim a_n = \lim b_n$, so folgt daraus $\lim c_n = \lim a_n = \lim b_n$. Die Aussage ist auch ein Korollar von Satz 4.19.

[5] Eine dritte Bemerkung: Als Nebenprodukt folgt aus dem Beweis, dass $\bigcap_n I_n$ auch Grenzwert von (a_n) und (b_n) ist.

also $x \in I_k$. 3. Ferner ist offensichtlich $\operatorname{diam}(I_k) \to 0$ für $k \to \infty$. Damit ist $(I_k)_{k\in\mathbb{N}}$ eine Intervallschachtelung.

Nach dem Intervallschachtelungsprinzip gibt es ein $x \in \bigcap_{k\in\mathbb{N}} I_k$ (also ein x in allen Intervallen), wofür stets $|x - a_{N_k}| \le 2^{-k+1}$ gilt (a_{N_k} ist der Mittelpunkt von I_k). Damit und mit $|a_{N_k} - a_n| < 2^{-k}$ (Cauchy-Kriterium) erhalten wir

$$|x - a_n| \le |x - a_{N_k}| + |a_{N_k} - a_n| < 2^{-k+1} + 2^{-k} = 3 \cdot 2^{-k} = \varepsilon$$

für alle $k \in \mathbb{N}$, weshalb $a_n \to x$ für $n \to \infty$. □

In diesem Beweis nahmen wir eine Cauchy-Folge (a_n) und konstruierten darum eine Intervallschachtelung symmetrisch um a_n, wobei sich die Länge der Intervalle sukzessive halbiert. Diese Intervallschachtelung umschließt ein x, sodass auch die Cauchy-Folge gegen x konvergiert, was das Vollständigkeitsaxiom ist.

Beispiel 5.11 (für Intervallschachtelungen).

1. Für alle $x \in \mathbb{R}$ sei $I_k = \left[x - k^{-1}, x + k^{-1}\right]$. Dann gilt $\{x\} = \bigcap_{k\in\mathbb{N}} I_k$. Ähnliche Beispiele wären $I_k = \left[x, x + 2^{-k}\right]$, etc.

2. Seien f_n und F_n die Flächen der den Einheitskreis einbeschriebenen bzw. umschriebenen n-Ecke. Seien $I_k := [f_{3\cdot 2^{-k}}, F_{3\cdot 2^{-k}}]$, wobei offensichtlich $I_{k+1} \subset I_k$ gilt. Dann approximieren die I_k die Kreiszahl π: $\bigcap_{k\in\mathbb{N}} I_k = \{\pi\}$ durch die Folge von Dreiecken, Sechsecken, Zwölfecken, etc.

   Hierbei ist $f_6 = \frac{3}{2}\sqrt{3}$ und $F_6 = 2\sqrt{3}$. Schon Archimedes wusste $f_{192} = 3 + 10/71$ und $F_{192} = 3 + 10/70$.

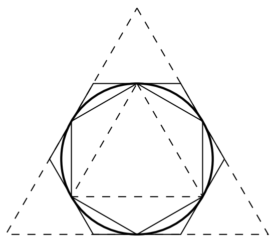


Abbildung 5.1: Erste Annäherung der Kreisfläche durch Dreiecke f_3, F_3 und Sechsecke f_6, F_6.

5.4.1 Wurzeln

Satz 5.12 (Existenz der Wurzeln). *Für alle $x > 0$ und für alle $k \in \mathbb{N}$ gibt es genau ein $y > 0$ mit $y^k = x$.*

Bezeichnung 5.13. $y = \sqrt[k]{x} = x^{1/k}$ bezeichnet die *k-te Wurzel* von x.

Beweis. Wir betrachten zunächst den Fall $x > 1$. Hierfür definieren wir eine Intervallschachtelung $I_n = [a_n, b_n]$ rekursiv mit $a_1 = 1$, $b_1 = x$, und $c_n = \frac{1}{2}(a_n + b_n)$. Dann ist $I_{n+1} = [a_{n+1}, b_{n+1}]$ mit

$$\begin{cases} a_{n+1} := a_n,\ b_{n+1} := c_n & \text{falls } x \le c_n^k, \\ a_{n+1} := c_n,\ b_{n+1} := b_n & \text{falls } x > c_n^k. \end{cases}$$

(I_n) erfüllt alle Bedingungen einer Intervallschachtelung: Die I_n sind abgeschlossen, es gilt $I_{n+1} \subset I_n$ und $\operatorname{diam}(I_{n+1}) = \frac{1}{2}\operatorname{diam}(I_n) = 2^{-n}(x-1) \to 0$ für alle $n \in \mathbb{N}$. Deshalb folgt aus dem Intervallschachtelungsprinzip, dass es ein $y \in \bigcap_{n\in\mathbb{N}} I_n$ gibt.

Nun betrachten wir die Intervalle $J_n = \left[a_n^k, b_n^k\right]$. Auch (J_n) ist eine Intervallschachtelung: Die J_n sind abgeschlossen, es gilt $J_{n+1} \subset J_n$ und wir haben

$$\operatorname{diam}(J_n) = b_n^k - a_n^k = (b_n - a_n)\left(b_n^{k-1} + a_n b_n^{k-2} + \cdots + a_n^{k-2} b_n + a_n^{k-1}\right) \le \operatorname{diam}(I_n) \cdot k x^{k-1} \to 0$$

für $n \to \infty$. Die letzte Abschätzung gilt wegen $x \ge b_n \ge a_n \iff x^{k-1} \ge b_n^{k-1} \ge a_n^{k-1}$.

Durch die rekursive Konstruktion von I_n bleibt stets $a_n^k \le x \le b_n^k$ für alle $n \in \mathbb{N}$ erhalten, woraus $x \in J_n$ für alle $n \in \mathbb{N}$ folgt. Aus $y \in \bigcap_{n\in\mathbb{N}} I_n$ folgt auch $y^k \in J_n$ für alle $n \in \mathbb{N}$, weshalb aufgrund der Eindeutigkeit des Intervallschachtelungsprinzips $y^k = x$ gelten muss. Das beweist den Fall $x > 1$.

Der Fall $x = 1$ ist trivial. Für $0 < x < 1$ ist $1/x > 1$, sodass die eindeutige Wurzel $\sqrt[k]{1/x}$ existiert. Damit ist $\left(\sqrt[k]{1/x}\right)^{-k} = x$. □

c_n war unsere momentane Approximation von y, und durch die Fallunterscheidung schauten wir, ob unsere Schätzung über oder unter y liegt. Dementsprechend schränkten wir das Suchintervall I_n in die richtige Richtung ein, sodass stets $y \in I_n$ gilt. Dann „potenzierten wir das Intervall I_n" und erhielten J_n, welches nun sowohl x als auch y^k enthält, und aufgrund der Eindeutigkeit des Intervallschachtelungsprinzips sind sie identisch.[6]

Im obigen Beweis haben wir $b_n - a_n$ aus $b_n^k - a_n^k$ ausgeklammert. Das wurde in der Übung gezeigt.

Ende der Vorlesung 6 am 27. Oktober 2021

5.5 Supremumseigenschaft

(Auch *lineare Vollständigkeit* oder *Ordnungsvollständigkeit*)

Definition 5.14 (beschränkt, Maximum/Minimum, Supremum/Infimum)**.**

1. Eine Menge $M \subset \mathbb{R}$ heißt *nach oben beschränkt* (oder *von oben beschränkt*), falls es ein $K \in \mathbb{R}$ gibt mit $x \leq K$ für alle $x \in M$. In diesem Fall heißt K *obere Schranke* von M.
2. Ein $K \in \mathbb{R}$ heißt *Maximum* von M, falls es eine obere Schranke von M ist und zu M gehört, d. h. $K \in M$ und $\forall x \in M: x \leq K$. Wir schreiben $K = \max M$.
3. Ein $K \in \mathbb{R}$ heißt *Supremum* von M, fall es die kleinste obere Schranke von M ist, d. h. K ist eine obere Schranke von M und für alle oberen Schranken K' von M gilt $K' \geq K$. Hierfür schreiben wir $K = \sup M$.
4. Entsprechend werden definiert:
   (a) *untere Schranke* von M,
   (b) *Minimum* von M, $L = \min M$, und
   (c) *Infimum* von M, $L = \inf M$.

Bemerkung 5.15. Offenbar gilt

- Falls ein Maximum (bzw. Minimum) existiert, dann ist es auch ein Supremum (bzw. Infimum).
- Obere und untere Schranken sind nicht eindeutig.
- Supremum, Infimum, Maximum und Minimum von M sind (falls sie existieren) eindeutig.

Beispiel 5.16.

1. Für $M = [a, b]$ ist $\min M = a$ und $\max M = b$.
2. Für $M = (a, b)$ ist $\inf M = a$ und $\sup M = b$. Es existieren weder Minimum noch Maximum.
3. Für $M = \{n/(n+1) : n \in \mathbb{N}\}$ ist $\sup M = 1$. Es existiert kein Maximum.
4. Für $M = \{n^2/2^n : n \in \mathbb{N}\}$ ist $\max M = \frac{9}{8}$.
5. $\mathbb{N}$ ist wegen des archimedischen Axioms nicht nach oben beschränkt. Das archimedische Axiom lautet in Symbolen: $\forall x \in \mathbb{R}: \exists n \in \mathbb{N}: n \geq x$. Das ist äquivalent zu $\neg(\exists x \in \mathbb{R}: \forall n \in \mathbb{N}: n \leq x)$.

Satz 5.17 (Supremumseigenschaft von $\mathbb{R}$)**.** *Jede nach oben beschränkte Menge $M \subset \mathbb{R}$ mit $M \neq \varnothing$ besitzt ein Supremum. Also $\forall M \subset \mathbb{R}$: ($M$ nach oben beschränkt $\iff \exists \sup M \in \mathbb{R}$).*

Beweis mittels Intervallschachtelung. Wir definieren die Intervalle $I_n = [a_n, b_n]$ rekursiv mit den drei Eigenschaften

1. $a_n \in M$,
2. b_n ist eine obere Schranke von M und

[6] Der zugrundeliegende Algorithmus ist ähnlich zum HERON-*Verfahren* zur Bestimmung von Quadratwurzeln $\sqrt{x}$: Man wähle eine erste Approximation c_1, bilde das Intervall $[c_1, x/c_1] \ni x$ und schränke das Intervall mit $c_{n+1} = \frac{1}{2}(c_1 + x/c_1)$, also dem Durchschnitt der Randpunkte.

3. $b_n - a_n \leq 2^{1-n}(b_1 - a_1)$ (also $\operatorname{diam}(I_n) \to 0$).

Am Anfang können wir ein beliebiges $a_1 \in M$ und eine beliebige obere Schranke b_1 von M wählen. Nun der Schritt von n auf $n+1$: Wir setzen $c_n = \frac{1}{2}(b_n + a_n)$.

- Falls $M \cap (c_n, b_n] = \varnothing$, dann ist c_n eine obere Schranke von M. Dann setzen wir $b_{n+1} \coloneqq c_n$ und $a_{n+1} \coloneqq a_n$.
- Falls $M \cap (c_n, b_n] \neq \varnothing$, dann gibt es ein $a_{n+1} \in M$ mit $a_{n+1} > c_n$. In dem Fall setzen wir $b_{n+1} = b_n$.

Damit erfüllen auch a_{n+1}, b_{n+1} die obigen Bedingungen. Da auch $I_{n+1} \subset I_n$ gilt, gibt es gemäß dem Intervallschachtelungsprinzip ein $c \in \bigcap_{n\in\mathbb{N}} I_n$ mit $a_n \leq c \leq b_n$ für alle n.

Nun zeigen wir, dass c das Supremum von M ist: c ist eine obere Schranke von M, denn für alle $x \in M$ und für alle n gilt

$$x \leq b_n = a_n + 2^{1-n}(b_1 - a_1) \leq c + 2^{1-n}(b_1 - a_1),$$

sodass wegen der Ordnungserhaltung von Folgen (Satz 4.19) $x \leq c$ für $n \to \infty$ ist. c ist aber auch die kleinste obere Schranke von M, da für alle oberen Schranken K von M gilt

$$K \geq a_n = b_n - 2^{1-n}(b_1 - a_1) \geq c - 2^{1-n}(b_1 - a_1),$$

sodass $K \geq c$ für $n \to \infty$ wegen Satz 4.19 ist. □

Tatsächlich gilt in jedem angeordneten Körper folgender

Satz 5.18. *Die Supremumseigenschaft ist äquivalent zu Vollständigkeits- und archimedischem Axiom.*

Beweis. Die Rückrichtung haben wir gerade bewiesen. Die Hinrichtung erfolgt in der Übung. □

5.6 Zusammenfassung

Hiermit haben wir die reellen Zahlen $\mathbb{R}$ eingeführt. $\mathbb{R}$ ist ein vollständiger, archimedisch angeordneter Körper.

Ferner kann man zeigen (das werden wir nicht tun):

- $\mathbb{R}$ existiert. (Man zeigen, dass der Raum aller Cauchy-Folgen $\hat{\mathbb{Q}}$ unter Äquivalenzrelation alle Eigenschaften, die wir an $\mathbb{R}$ stellen, erfüllt; s. Unterkap. 5.3.)
- $\mathbb{R}$ ist hierdurch im wesentlich eindeutig charakterisiert, d. h. alle anderen Strukturen mit den Axiomen sind isomorph zu $\mathbb{R}$.

  Beispiel 5.19. $\mathbb{R} \to \hat{\mathbb{R}} \coloneqq \{x \in \mathbb{R} : x > 0\}$, $(u, v) \mapsto (e^u, e^v)$, $(0, 1) \mapsto (\hat{0} = 1, \hat{1} = e)$, $u + v \mapsto x \mathbin{\hat{+}} y = xy$, $u \cdot v \mapsto x \mathbin{\hat{\cdot}} y = x^{\ln y}$.

5.6.1 Einige Bemerkungen zu $\mathbb{N}$

- Wenn man $\mathbb{R}$ bereits kennt, erhält man $\mathbb{N} = \{1, 2, 3, \dots\} \subset \mathbb{R}$ als *kleinste Induktive Teilmenge von* $\mathbb{R}$. Hierbei heißt eine Menge $M \subset \mathbb{R}$ *induktiv*, falls
  - $1 \in M$ und
  - $(n + 1) \in M$ für alle $n \in M$ gilt (vgl. Beweisprinzip der vollständigen Induktion, Unterkap. 2.1).

  Offenbar ist der Durchschnitt (beliebig vieler) induktiver Mengen wieder induktiv, also auch der Durchschnitt aller induktiver Mengen. Dieser Durchschnitt sei die *kleinste* induktive Menge.

  Dieser Ansatz ist sehr fragwürdig, da wir $\mathbb{R}$ *ausgehend von* $\mathbb{N}$ definiert und konstruiert haben.
- Ohne einen Rückgriff auf $\mathbb{R}$ wird $\mathbb{N}$ durch die *Peano-Axiome* von Giuseppe Peano (1858–1939) definiert:

  $\mathbb{N}$ ist eine Menge mit einem ausgezeichnetem Element $1 \in \mathbb{N}$ (oder 0, oder 2, etc.) und
  - $\forall n \in \mathbb{N} \colon \exists v(n) \in \mathbb{N} \setminus \{1\}$ (die Idee eines „Nachfolgers“);
  - $\forall m, n \in \mathbb{N} \colon m = n \iff v(n) = v(m)$ (zwei Zahlen sind genau dann gleich, wenn sie denselben Nachfolger haben – die Nachfolgerfunktion ist also injektiv);
  - $\forall M \subset \mathbb{N} \colon (1 \in M \wedge v(M) \subset M) \implies (M = \mathbb{N})$ (*Induktionsaxiom*, die Menge der natürlichen Zahlen ist eindeutig).[7]

[7] Oft findet man in der Literatur eine „ziemlich äquivalente“ Formulierung $\forall M \colon (1 \in M \wedge v(M) \subset M) \implies \mathbb{N} \subset M$. Dieses Axiom ist wichtig, denn die vorherigen Axiome garantieren nicht, dass mit der Folge $1, v(1), v(v(1)), v(v(v(1))), \dots$ alle natürlichen Zahlen durchlaufen werden. Ein Ring mit $v^k(n) = n$ würde nämlich auch alle Axiome erfüllen.

5.7 Abzählbarkeit

Definition 5.20 (Abbildung)**.** Gegeben seien Mengen A und B. Eine *Abbildung* (oder *Funktion*) $f: A \to B$ ist eine Zuordnung, die jedem $x \in A$ genau ein $f(A) \in B$ zuordnet (das „*Bild* von x unter f"), kurz $x \mapsto f(x)$. Wir bezeichnen $f(A) = \{f(x) : x \in A\} = \{y \in B : (\exists x \in A \colon y = f(x))\} \subset B$.

Definition 5.21 (surjetiv, injektiv, bijektiv)**.** Sei f eine Abbildung. f heißt

- *surjektiv* (selten auch „Abbildung *auf* B"), falls $f(A) = B$,
- *injektiv* (auch „eindeutig"), falls $x = x' \iff f(x) = f(x')$ für alle $x, x' \in A$ gilt und
- *bijektiv*, falls f surjektiv und injektiv ist.

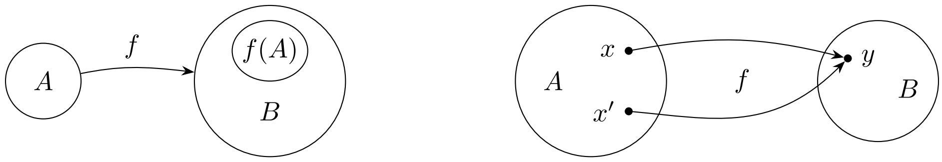


Abbildung 5.2: Links: f ist *nicht surjektiv*. Rechts: f ist *nicht injektiv*.

Die folgenden Definitionen gehen auf Georg Cantor zurück.

Definition 5.22 (Gleichmächtigkeit, Abzählbarkeit)**.**

1. Mengen A und B heißen *gleichmächtig*, falls es eine bijektive Abbildung $f: A \to B$ gibt.
2. Eine Menge A heißt *abzählbar unendlich*, falls sie zu $\mathbb{N}$ gleichmächtig ist.
3. Eine Menge A heißt *abzählbar*, falls sie leer, endlich oder abzählbar unendlich ist. Andernfalls heißt A *überabzählbar*.

Die Idee der Gleichmächtigkeit ist, dass eine Reihenfolge gefunden werden kann, in der alle Elemente der Menge vorkommen, sprich wir können die Elemente der Menge M durchnummerieren mit $M = \{a_1, a_2, a_3, \ldots\}$. Damit erhalten wir eine Bijektion zwischen den Indizes (den natürlichen Zahlen) und den Elementen.

Ende der Vorlesung 7 am 03. November 2021

Beispiel 5.23.

1. $\mathbb{N}$, $\mathbb{Z}$ und $\mathbb{Q}$ sind abzählbar.

   $\mathbb{N}$ ist offensichtlich abzählbar. Für $\mathbb{Z}$ können wir die Reihenfolge $\mathbb{Z} = \{0, 1, -1, 2, -2, 3, -3, \ldots\}$ definieren, sodass wir jedem Element seine Stelle in der Auflistung (eine natürliche Zahl) zuordnen:

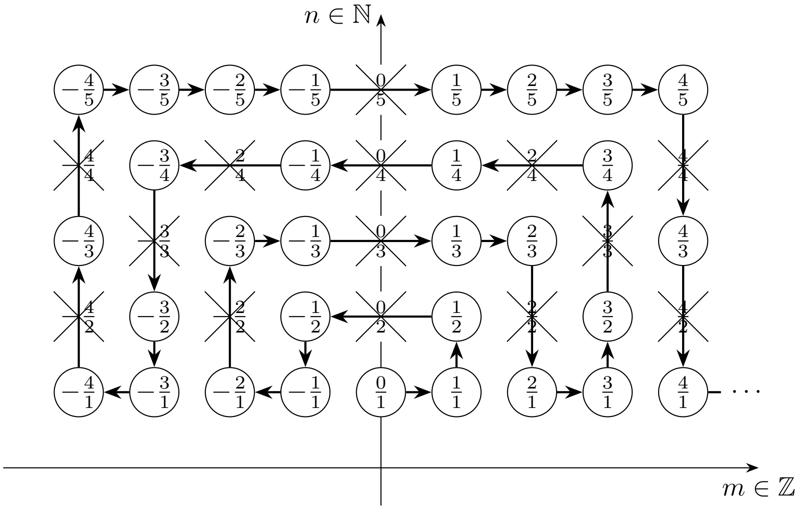
   $$\begin{matrix} 0 & 1 & -1 & 2 & -2 & 3 & -3 & \ldots \\ \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & \ldots \end{matrix} \quad \text{bzw.} \quad f(n) = \begin{cases} \frac{1}{2}n & \text{falls } n \text{ gerade,} \\ \frac{1}{2}(1-n) & \text{falls } n \text{ ungerade.} \end{cases}$$

   Für $\mathbb{Q}$ erfolgt ein Beweis weiter unten.

2. $\mathbb{R}$ ist überabzählbar. Auch das beweisen wir unten.

Satz 5.24. *$\mathbb{Q}$ ist abzählbar.*

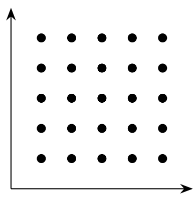
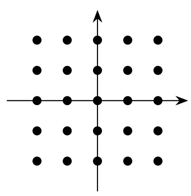
Beweis. Jeden Bruch $q \in \mathbb{Q}$ können wir als Quotienten $q = m/n$ mit teilerfremden $m \in \mathbb{Z}$ und $n \in \mathbb{N}$ schreiben. Wir betrachten alle (m, n) als Punkte im Gitter $\mathbb{Z}^2 \subset \mathbb{R}^2$:



Die teilerfremden (m, n) werden behalten, während die nicht teilerfremden (m, n) rausgestrichen werden. Wir durchlaufen alle Gitterpunkte in „konzentrischen Kreisen“ um 0 und betrachten dabei nur die teilerfremden Paare (m, n). Damit können wir alle q durchnummerieren.[8] □

Bemerkung 5.25.

1. Mit analogen Verfahren können wir auch $\mathbb{N} \times \mathbb{N}$ oder $\mathbb{Z} \times \mathbb{Z}$ durchnummerieren.

2. Die naive Idee, zuerst alle q mit $n = 1$, dann $n = 2$ etc. durchzugehen funktioniert *nicht*, da es unendlich viele q in der Reihe $n = 1$ gibt und wir nie zu $n > 1$ kommen.

Satz 5.26. $\mathbb{R}$ *ist nicht abzählbar.*

Beweis durch Intervallschachtelung. Wir nehmen an, dass $\mathbb{R} = \{x_1, x_2, \ldots\}$ abzählbar ist. Nun definieren wir eine Intervallschachtelungen $(I_n)_n$ rekursiv: $I_1 = [x_1 + 1, x_1 + 2]$; zerlege I_n in drei gleichlange Teilintervalle und wähle eines dieser als I_{n+1} so aus, dass x_{n+1} nicht in I_{n+1} enthalten ist. Nach Intervallschachtelungsprinzip gibt es ein $s \in \mathbb{R}$ mit $s \in \bigcap_n I_n$. Das bedeutet aber auch, dass es ein $k \in \mathbb{N}$ geben muss mit $s = x_k \in \mathbb{R}$. Nach Konstruktion der I_k ist $x_k \notin I_k$, aber $s \in I_k$, ein Widerspruch. □

Im Beweis reicht es nicht, I_n zu halbieren, da x_{n+1} die Mitte von I_n sein kann und damit in beiden Teilintervallen liegt.

Alternativer Beweis. Wir verwenden die dyadische (binäre) Darstellung der reellen Zahlen. Angenommen $[0, 1) = \{x_1, x_2, \ldots\}$ ist abzählbar, d. h. wir können eine Liste aller Zahlen

$$\begin{aligned} x_1 &= 0, \underline{a_{11}} a_{12} a_{13} \ldots \\ x_1 &= 0, a_{11} \underline{a_{12}} a_{13} \ldots \\ x_1 &= 0, a_{11} a_{12} \underline{a_{13}} \ldots \\ &\vdots \end{aligned}$$

[8] Das ist eine alternative Version des *ersten* CANTORSCHEN *Diagonalargument*. Für die originale Idee nehmen wir nur den ersten Quadraten der Abbildung, also $\mathbb{Q}^+$. Nun zeichnen wir Diagonalen mit Anstieg -1 und verbinden sie so parallel zu den Achsen, dass eine „Schlangenlinie“ über die Diagonalen entsteht. Streichen wir noch die nicht teilerfremden Paare (m, n), haben wir eine Reihenfolge und $\mathbb{Q}^+$ und $\mathbb{N}$ sind gleichmächtig. Fügen wir vor der Schlangenlinie die Null, und hinter jedem (m, n) noch $(-m, n)$ hinzu, sind auch $\mathbb{Q}$ und $\mathbb{N}$ gleichmächtig. Mehr dazu s. `https://de.wikipedia.org/wiki/Cantors_erstes_Diagonalargument`.

mit $a_{ij} \in \{0,1\}$ für alle $i, j \in \mathbb{N}$.

Nun konstruieren wir eine neue Zahl $y = 0, b_{11} b_{12} b_{13} \dots$ mit $b_i = 1 - a_{ii}$ für alle i. Nun unterscheiden sich y und jedes x_i an der i-ten Stelle, weshalb y nicht in der Liste liegt, also $y \notin \{x_i : i \in \mathbb{N}\} = [0,1)$, Widerspruch.[9] □

Bemerkung 5.27. Streng genommen müssen wir für den Beweis noch zeigen, dass jede reelle Zahl in $[0,1)$ sich im Binärsystem darstellen lässt, also $z = \sum_{i=0}^{\infty} z_i 2^{-i}$ in $\mathbb{R}$ konvergiert und jede binäre Repräsentation einer reellen Zahl auch eindeutig ist.

Korollar 5.28. $\mathbb{R} \setminus \mathbb{Q}$ *ist nicht abzählbar.*

Bemerkung 5.29. Folgende Aussagen werden noch in den Übungen gezeigt.

1. Die Menge *aller Teilmengen* aus $\mathbb{N}$ ist überabzählbar. Die Menge aller *endlichen* Teilmengen aus $\mathbb{N}$ ist abzählbar.

2. Die Menge *aller Folgen* mit Werten in $\{0,1\}$ ist überabzählbar. Die Menge aller *endlichen* Folgen mit Werten in $\mathbb{Q}$ ist abzählbar.

5.7.1 Kontinuumshypothese

(Cantor 1878). *Es existiert keine Menge, deren Mächtigkeit zwischen der von* $\mathbb{N}$ *und von* $\mathbb{R}$ *liegt.*

Diese Aussage ist mit den üblichen Axiomen der Mengenlehre (*Zermelo-Frankel-Mengenlehre*) weder beweisbar noch widerlegbar (Nichtwiderlegbarkeit durch Kurt Gödel 1940, Nichtbeweisbarkeit durch Paul Cohen 1963).[10]

6 Konvergenz von Folgen II

6.1 Monotone Folgen

Definition 6.1 (schwach waschend/fallend, monoton)**.** Eine Folge $(a_n)_{n\in\mathbb{N}}$ von reellen Zahlen heißt

- *schwach wachsend* (nicht fallend), falls $a_n \leq a_{n+1}$ für alle $n \in \mathbb{N}$ gilt,
- *schwach fallend* (nicht wachsend), falls $a_n \geq a_{n+1}$ für alle $n \in \mathbb{N}$ gilt, und
- *monoton*, falls sie schwach wachsend oder schwach fallend ist.

Satz 6.2 (Konvergenz beschränkter monotoner Folgen)**.** *Jede beschränkte monotone Folge* $(a_n)_{n\in\mathbb{N}}$ *reeller Zahlen konvergiert, und zwar*

- *gegen* $\sup A = \{a_n\}$, *falls* (a_n) *schwach wachsend ist, und*
- *gegen* $\inf A = \{a_n\}$, *falls* (a_n) *schwach fallend ist.*

Beweis. Wir beweisen den Satz für schwach wachsende Folgen. Für schwach fallende Folgen erfolgt der Beweis analog.

Sei $a = \sup A$. Dann gibt es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$, sodass $a - \varepsilon \leq a_N$ gilt (ansonsten wäre $a - \varepsilon$ eine obere Schranke der Folge). Mit der Monotonie gilt nun für alle $n \geq N$: $a - \varepsilon \leq a_n \leq a$, also $a_n \to a$ für $n \to \infty$. □

Vgl. den Satz mit Satz 4.12, der besagt, dass jede konvergente Folge beschränkt ist.

[9] Der Beweis (aber im Dezimalsystem) ist bekannt als das *zweite cantorsche Diagonalargument*.

[10] Unser Dozent rechnete fälschlicherweise die Unentscheidbarkeit der Kontinuumshypothese Gödel im Jahr 1938 zu.

6.2 Uneigentliche Konvergenz

Unser Ziel ist es, sagen zu können, dass *jede* (auch unbeschränkte) monotone Folge „konvergiert“. Unsere bisherige Definition 4.4 im (strikten) Sinne erlaubt nur Konvergenz gegen reelle Zahlen. Wir erweitern also den Begriff, sodass unbeschränkte monotone Folgen auch gegen $+\infty$ oder $-\infty$ konvergieren können.

Definition 6.3 (erweiterte Zahlengerade)**.** Sei die *erweiterte Zahlengerade* (auch die *erweiterten reellen Zahlen*) $\overline{\mathbb{R}} := \mathbb{R} \cup \{-\infty, +\infty\}$ die um zwei Elemente $\pm\infty$ („plus/minus unendlich“) erweiterte Menge der reellen Zahlen $\mathbb{R}$.

Für $\pm\infty$ werden folgende Rechenregeln für alle $x \in \mathbb{R}$ definiert:

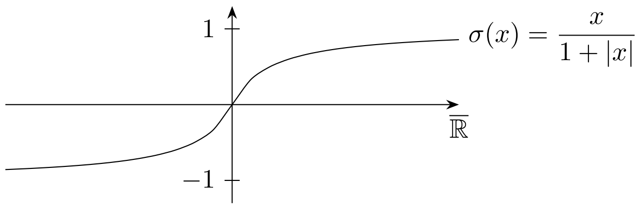
- $-\infty < x < +\infty$;
- $(\pm\infty) + x = \pm\infty$, genauso gelten hierfür Kommutativität, Assoziativität und Distributivität;
- für $x > 0$: $(\pm\infty)x = \pm\infty$, hierfür gelten auch Kommutativität, Assoziativität, Distributivität;
- für $x < 0$: $(\pm\infty)x = \mp\infty$;
- $(-\infty) + (+\infty)$ ist nicht definiert;
- $(\pm\infty) \cdot 0$ ist nicht definiert.[1]

Ferner sind $\overline{\mathbb{R}} = [-\infty, +\infty]$, $\mathbb{R} = (-\infty, +\infty)$ und für alle $a, b \in \overline{\mathbb{R}}$ mit $a < b$ die Intervalle $[a, b]$, $[a, b)$, $(a, b]$ und (a, b) wie gewohnt (d. h. analog zu Definition 5.6) definiert.

Definition 6.4 (uneigentliche Konvergenz)**.** Eine Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen *konvergiert uneigentlich gegen* $+\infty$ (geschrieben $a_n \to +\infty$), falls für jedes $k \in \mathbb{R}$ ein $N \in \mathbb{N}$ existiert, sodass $a_n \geq k$ für alle $n \geq N$ gilt. Analog ist die *uneigentliche Konvergenz gegen* $-\infty$ definiert ($a_n \to -\infty$).

Achtung: Die Folge $(a_n)_n$ *konvergiert*, falls sie gegen ein $a \in \mathbb{R}$ konvergiert. Sie ist hingegen *uneigentlich konvergent*, falls sie gegen $+\infty$ oder $-\infty$ konvergiert. Wenn wir also nicht spezifizieren, ist Konvergenz in $\mathbb{R}$ gemeint, und Divergenz kann uneigentliche Konvergenz oder Divergenz in $\overline{\mathbb{R}}$ bedeuten.

Bemerkung 6.5. Durch die bijektive Abbildung $\sigma\colon \overline{\mathbb{R}} \to [-1, 1]$, $x \mapsto x/(1 + |x|)$ mit $\sigma(+\infty) := 1$ und $\sigma(-\infty) := -1$ werden die Ordnungsstrukturen von $\bar{R}$ und $[-1, 1]$ ineinander überführt.[2] Ferner gilt deshalb: Eine Folge $(x_n)_{n\in\mathbb{N}}$ konvergiert (uneigentlich) gegen ein $x \in \bar{R}$ genau dann, wenn die Folge $(\sigma(x_n))_{n\in\mathbb{N}}$ gegen ein $\sigma(x) \in [-1, 1]$ konvergiert.



Achtung: Beim Rechnen mit $\pm\infty$ gelten nicht alle Regeln aus $\mathbb{R}$ wie gewohnt (ansonsten kommen Widersprüche auf). Dasselbe gilt auch für das Rechnen mit uneigentlich konvergenten Folgen.

6.3 Limes Superior und Limes Inferior

Definition 6.6 (Limes Superior/Inferior)**.** Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge reeller Zahlen und seien $b_n = \sup\{a_k : k \geq n\}$ und $c_n = \inf\{a_k : k \leq n\}$. Dann sind $(b_n)_{n\in\mathbb{N}}$ eine schwach fallende und $(c_n)_{n\in\mathbb{N}}$ eine schwach wachsende Folge[3] in $\overline{\mathbb{R}}$, die auch (uneigentlich in $\overline{\mathbb{R}}$) konvergieren. Dann sind

$$\limsup_{n\to\infty} a_n := \overline{\lim_{n\to\infty}}\, a_n := \lim_{n\to\infty} b_n \qquad \text{und} \qquad \liminf_{n\to\infty} a_n := \underline{\lim_{n\to\infty}}\, a_n := \lim_{n\to\infty} c_n$$

[1] Weitere nicht definierte Ausdrücke sind $x/0$, $(\pm\infty)/(\pm\infty)$ und $(\pm\infty)/(\mp\infty)$. Hingegen gilt zusätzlich für die Division $x/(\pm\infty) = 0$, $(\pm\infty)/x = \pm\infty$ für $x > 0$ und $(\pm\infty)/x = \mp\infty$ für $x < 0$.

[2] σ ist ein sog. *Homöomorphismus*, da die Abbildung bijektiv und stetig sowie die Umkehrabbildung auch stetig ist. Damit werden die Strukturen zwischen beiden Räumen erhalten (sind also *homöomorph*).

[3] Eine kurze Begründung dafür: Nach Definition ist $b_n \geq a_k$ für alle $k \geq n$, also auch $b_n \geq a_k$ für alle $k \geq n+1$, d. h. b_n ist eine obere Schranke von $\{a_k : k \geq n+1\}$. Aufgrund der Minimalität des Supremums muss dann $b_n \geq b_{n+1} = \sup\{a_k : k \geq n+1\}$ sein. Analog lässt sich für c_n argumentieren.

der *Limes Superior* bzw. der *Limes Inferior*.[4]

Wegen der Monotonie von (b_n) bzw. (c_n) gilt

$$\limsup_{n\to\infty} a_n = \inf_n b_n \qquad \text{und} \qquad \liminf_{n\to\infty} a_n = \sup_n c_n,$$

oder mit anderen Worten

Lemma 6.7. *Es gelten*

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \sup_{k\geq n} a_k = \inf_{n\in\mathbb{N}} \sup_{k\geq n} a_k \qquad \textit{und} \qquad \liminf_{n\to\infty} a_n = \lim_{n\to\infty} \inf_{k\geq n} a_k = \sup_{n\in\mathbb{N}} \inf_{k\geq n} a_k.$$

Beispiel 6.8.

1. Für $(a_n)_{n\in\mathbb{N}}$ mit $a_n = n$ sind $b_n = \sup\{a_k : k \geq n\} = +\infty$ und $c_n = \inf\{a_k : k \geq n\} = n$ für alle n. Damit sind $\limsup_{n\to\infty} a_n = \lim_{n\to\infty} b_n = +\infty$ und $\liminf_{n\to\infty} a_n = \lim_{n\to\infty} c_n = +\infty$.
2. Für $a_n = (-1)^n(1 - 1/n)$, also $\left(0, +\frac{1}{2}, -\frac{2}{3}, +\frac{3}{4}, -\frac{4}{5}, \ldots\right)$, sind $b_n = \sup\{a_k : k \geq n\} = +1$ und $c_n = \inf\{a_k : k \geq n\} = -1$. Damit sind $\limsup_{n\to\infty} a_n = +1$ und $\liminf_{n\to\infty} a_n = -1$.
3. Für $a_n = (-1)^n(1 + 1/n)$ sind

$$b_n = \begin{cases} 1 + \frac{1}{n} & \text{für gerade } n, \\ 1 + \frac{1}{n+1} & \text{für ungerade } n, \end{cases} \qquad \text{und} \qquad c_n = \begin{cases} 1 + \frac{1}{n} & \text{für ungerade } n, \\ 1 + \frac{1}{n+1} & \text{für gerade } n. \end{cases}$$

   Daraus folgen $\limsup_{n\to\infty} a_n = \lim_{n\to\infty} b_n = +1$ und $\liminf_{n\to\infty} = \lim_{n\to\infty} c_n = -1$.

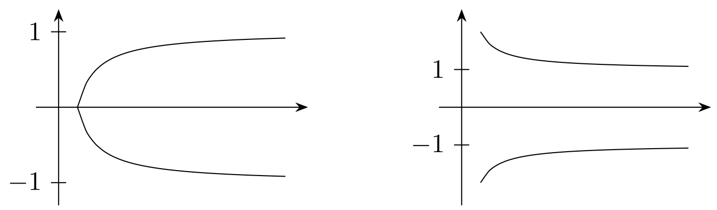


Abbildung 6.1: Veranschaulichung der Beispiele. Links: $a_n = (-1)^n(1 - 1/n)$. Rechts: $a_n = (-1)^n(1 + 1/n)$.

Bemerkung 6.9. Es gilt stets $\liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n$, was aus $\inf_{n\geq k} a_k \leq \sup_{n\geq k} a_k$ und der Ordnungserhaltung von konvergierenden Folgen folgt (Satz 4.19).

Ende der Vorlesung 8 am 08. November 2021

Satz 6.10. *Jede monotone Folge in $\mathbb{R}$ konvergiert in $\overline{\mathbb{R}}$.*

Beweis. O. B. d. A. sei $(a_n)_n$ schwach wachsend. Dann unterscheiden wir zwei Fälle:

1. Wenn $\sup_n a_n$ endlich ist, dann ist $(a_n)_n$ nach oben beschränkt, und aufgrund Satz 6.2 konvergiert die Folge gegen $\sup_n a_n$.
2. Wenn $\sup_n a_n = +\infty$, dann gibt es für jedes $k \in \mathbb{R}$ ein $N \in \mathbb{N}$ mit $a_N \geq k$ (andernfalls ist $\sup_n a_n$ nicht die kleinste obere Schranke mehr). Aufgrund Monotonie gilt $a_n \geq k$ für alle $n \geq N$. Das ist die Definition der uneigentlichen Konvergenz, also $a_n \to +\infty$ für $n \to \infty$. □

Satz 6.11. *Eine reelle Zahlenfolge $(a_n)_{n\in\mathbb{N}}$ konvergiert (bzw. konvergiert in $\overline{\mathbb{R}}$) genau dann, wenn $\limsup a_n = \liminf a_n \in \mathbb{R}$ (bzw. $\limsup a_n = \liminf a_n \in \overline{\mathbb{R}}$). In diesem Fall ist $\lim a_n = \liminf a_n = \limsup a_n$.*

[4] Die Idee des Limes Superior bzw. Inferior ist es, eine neue Weise einzuführen, um Folgen zu beschreiben. Bisher kennen wir nur die Konvergenz gegen eine Zahl, aber was ist, wenn die Folge divergiert? Dann kann es trotzdem sein, dass sie beschränkt ist (wie z. B. $((-1)^n)_n$, aber in $\overline{\mathbb{R}}$ eigentlich immer), und wir können versuchen Aussagen über die Grenzwerte der Schranken zu machen. Mit dem Limes Superior bzw. Inferior können wir also den „Grenzwert" zu einem „Grenzbereich" erweitern, worin ab einem bestimmten Folgenglied alle Folgenglieder im Intervall $[\liminf a_n - \varepsilon,\ \limsup a_n + \varepsilon]$ für beliebig kleine ε liegen.

Beweis. Wir zeigen die Behauptung für Konvergenz in $\mathbb{R}$. Für die uneigentliche Konvergenz kann man ähnlich verfahren.

Zur Hinrichtung: Angenommen $(a_n)_n$ konvergiert gegen $a \in \mathbb{R}$. Dann gibt es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$, sodass $|a_n - a| < \varepsilon \iff a - \varepsilon < a_n < a + \varepsilon$ für alle $n \geq N$ gilt.

Aus $a_n < a + \varepsilon$ folgt $b_N := \sup_{n \geq N} a_n \leq a + \varepsilon$. Da (b_n) eine schwach fallende folgende Folge ist, gilt für alle $n \geq N$ dann $b_n \leq b_N \leq a + \varepsilon$, also $\lim_{n\to\infty} b_n \leq a + \varepsilon$.

Analog verfahren wir mit $a_n > a - \varepsilon$, sodass wir $c_N := \inf_{n \geq N} a_n \geq a - \varepsilon$, also $\lim_{n\to\infty} c_n \geq a - \varepsilon$ erhalten. Mit $\liminf a_n \leq \limsup a_n$ erhalten wir

$$a - \varepsilon \leq \lim_{n\to\infty} c_n = \liminf_{n\to\infty} a_n \leq \limsup_{n\to\infty} a_n = \lim_{n\to\infty} b_n \leq a + \varepsilon,$$

wobei für beliebig kleine $\varepsilon > 0$ Gleichheit gilt.

Zur Rückrichtung: Angenommen es gilt $\liminf a_n = a = \limsup a_n \in \mathbb{R}$. Dann gibt aufgrund des Konvergenzkriteriums für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$, sodass $b_N \leq a + \varepsilon$ und $c_N \geq a - \varepsilon$ gilt. Wegen der Definition der Suprema/Infima haben wir $a_n \leq a + \varepsilon$ und $a_n \geq a - \varepsilon$ für alle $n \geq N$, also $|a_n - a| < \varepsilon$ und daher $a_n \to a$. □

6.4 Teilfolgen und Häufungspunkte

Definition 6.12 (Teilfolge, Häufungspunkt)**.**

1. Sei $(a_n)_{n\in\mathbb{N}}$ eine reelle Zahlenfolge und $(n_k)_{k\in\mathbb{N}}$ eine strikt wachsende Folge in $\mathbb{N}$. Dann heißt die Folge $(a_{n_k})_{k\in\mathbb{N}} = (a_{n_1}, a_{n_2}, a_{n_3}, \dots)$ *Teilfolge* der Folge $(a_n)_{n\in\mathbb{N}}$.

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

Beispiel 6.13.

1. $a_n = (-1)^n$, dann sind 1 und -1 Häufungspunkte, denn $1 = \lim_{k\to\infty} a_{2k}$ und $-1 = \lim_{k\to\infty} a_{2k+1}$ (also $n_k = 2k$ bzw. $n_k = 2k+1$).

2. Teilfolgen der Folge $(1/n)_{n\in\mathbb{N}}$ sind z. B. $\left(2^{-k}\right)_{k\in\mathbb{N}}$, $(1/k!)_{k\in\mathbb{N}}$, $(1/p)_{p \text{ prim}}$, etc.

Bemerkung 6.14. Jede Teilfolge einer *konvergenten Folge* konvergiert und besitzt denselben Grenzwert (das folgt sofort aus der Konvergenzdefinition). D. h. jede konvergente Folge besitzt genau einen Häufungspunkt, nämlich den Grenzwert selbst.[5]

Lemma 6.15. *$a \in \mathbb{R}$ ist genau dann ein Häufungspunkt von $(a_n)_{n\in\mathbb{N}}$, wenn es für jedes $\varepsilon > 0$ unendlich viele $n \in \mathbb{N}$ gibt mit $|a_n - a| < \varepsilon$.*

Beweis. Zur Hinrichtung: Ist a ein Häufungspunkt von (a_n), so gibt es eine Folge $(n_k)_k$ natürlicher Zahlen und ein $K \in \mathbb{N}$, sodass $|a_{n_k} - a| \leq \varepsilon$. Also gibt es unendlich viele $n \in \mathbb{N}$, die das erfüllen, nämlich $n \in \{n_K, n_{K+1}, n_{K+2}, \dots\}$.

Zur Rückrichtung: Wir werden eine streng wachsende Folge $(n_k)_k$ in $\mathbb{N}$ rekursiv konstruieren, sodass die Teilfolge $(a_{n_k})_k$ gegen a konvergiert. Dazu wählen wir $\varepsilon = 1/k$ im Konvergenzkriterium. Sei n_1 so gewählt mit $|a_{n_1} - a| < 1/1$. Für folgende $k > 1$ wählen wir n_k so von den unendlich vielen n, dass $|a_{n_k} - a| < 1/k$ und n_k größer als alle vorher gewählten $n_1, n_2, \dots, n_{k-1}$ ist. Damit gilt $a_{n_k} \to a$. □

Nun zu einem wichtigen Satz der Analysis:

Satz 6.16 (Bolzano-Weierstrass)**.** *Für jede beschränkte reelle Zahlenfolge $(a_n)_{n\in\mathbb{N}}$ gilt:*

1. *$\liminf_{n\to\infty} a_n$ und $\limsup_{n\to\infty} a_n$ sind Häufungspunkte.*

2. *Für jeden Häufungspunkt $a \in \mathbb{R}$ gilt*

$$\liminf_{n\to\infty} a_n \leq a \leq \limsup_{n\to\infty} a_n.$$

Mit anderen Worten: $\liminf_{n\to\infty} a_n$ und $\limsup_{n\to\infty} a_n$ sind der kleinste bzw. größte Häufungspunkt.

[5] Die Umkehrung gilt auch: Konvergiert jede Teilfolge, so konvergiert auch die Folge selbst, denn die Folge selbst ist ja auch eine Teilfolge von sich selbst.

Beweis.

1. Wir konstruieren rekursiv $(n_k)_{k\in\mathbb{N}}$ in $\mathbb{N}$ mit $a_{n_k} \to \limsup_{n\to\infty}$ für $k \to \infty$. Die Konstruktion für $a_{n_k} \to \liminf_{n\to\infty}$ ist ähnlich.

   Für den Rekursionsanfang setzen wir $n_1 = 1$. Für den Rekursionsschritt sei n_k gegeben. Wir betrachten

   $$b_{1+n_k} := \sup\{a_{1+n_k}, a_{2+n_k}, a_{3+n_k}, \ldots\}.$$

   Diese Zahl ist reell, weil $(a_n)_n$ beschränkt ist. Da sie auch die kleinste obere Schranke der obigen Menge ist, gibt es ein $n_{k+1} \in \{1+n_k, 2+n_k, 3+n_k, \ldots\}$ mit $a_{n_{k+1}} \geq b_{1+n_k} - 1/k$ (andernfalls wäre auch $b_{1+n_k} - 1/k$ eine obere Schranke). Daher gilt für alle $k \in \mathbb{N}$:

   $$b_{1+n_k} \geq a_{n_{k+1}} \geq b_{1+n_k} - \frac{1}{k}.$$

   Nun wissen wir, dass $(b_{1+k_n})_k$ monoton schwach fällt und deshalb konvergiert (Satz 6.2). Aus der eben gezeigten Ungleichung folgt also für beliebig kleine $1/k$, dass $(a_{n_k})_{k\in\mathbb{N}}$ konvergiert, nämlich gegen den Grenzwert von $(b_{1+n_k})_k$. Da $(b_{1+n_k})_k$ eine Teilfolge von $(b_n)_n$ (es sei $b_n := \sup_{k\geq n} a_k$) ist, konvergieren $(a_{n_k})_k$ und $(b_{n_k})_k$ gegen den Grenzwert von $(b_n)_n$, also $a_{n_k} \to \limsup_n a_n$.

2. Sei $a = \lim_{k\to\infty} a_{n_k}$. Dann gibt es für jedes $\varepsilon > 0$ ein $K \in \mathbb{N}$, sodass $|a - a_{n_k}| < \varepsilon$ für alle $k \geq K$ gilt. Daraus folgt $a \leq a_{n_k} + \varepsilon \leq \sup\{a_n : n \geq n_k\} + \varepsilon$. Für $k \to \infty$ ergibt sich $a \leq \limsup_{n\to\infty} a_n + \varepsilon$, und für beliebig kleine $\varepsilon > 0$ schließlich $a \leq \limsup_{n\to\infty} a_n$. Analog würden wir $a \geq \liminf_{n\to\infty} a_n$ erhalten. □

Korollar 6.17. *Jede beschränkte Zahlenfolge besitzt (mindestens) eine konvergente Teilfolge.*[6]

Korollar 6.18. *Jede reelle Zahlenfolge besitzt (mindestens) eine in $\overline{\mathbb{R}}$ konvergente Teilfolge.*

7 Konvergenz von Reihen

7.1 Unendliche Reihen

Definition 7.1 (Partialsumme, Reihe)**.** Gegeben sei eine reelle Zahlenfolge $(a_n)_{n\in\mathbb{N}}$ (nicht in $\overline{\mathbb{R}}$!). Dann ist $(s_n)_{n\in\mathbb{N}}$ die *Folge der Partialsummen* $s_n = \sum_{k=1}^n a_k$ (zum Teil auch $(a_n)_{n\in\mathbb{N}_0}$ und $s_n = \sum_{k=0}^n a_k$).
Als *unendliche Reihe* $\sum_k a_k$ bezeichnen wir

- einerseits die Folge $(s_n)_{n\in\mathbb{N}}$ und
- andererseits deren Grenzwert $\lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=1}^n a_k = \sum_{k=1}^\infty a_k$.

Beispiel 7.2.

1. Die *geometrische Reihe* $\sum_{k=0}^\infty q^k = 1/(1-q)$ für alle $q \in \mathbb{R}$ mit $|q| < 1$.

2. Die *harmonische Reihe* $\sum_{n=1}^\infty 1/n = +\infty$. Die geometrische und harmonische Reihe sind sehr wichtig („Mütter aller Reihen").

   Beweis (unformal). Es gilt

   $$1 + \underbrace{\frac{1}{2}}_{\geq 1/2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\geq 1/2} + \underbrace{\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}}_{\geq 1/2} + \underbrace{\frac{1}{9} + \cdots + \frac{1}{16}}_{\geq 1/2} + \underbrace{\frac{1}{17} + \cdots + \frac{1}{32}}_{\geq 1/2} + \cdots$$

   Wir können das auch alternativ durch die Cauchy-Folge der Partialsummen beweisen: Für alle $N \in \mathbb{N}$ gibt es $m > n \geq N$, sodass $|s_n - s_m| \geq \frac{1}{2}$ gilt, nämlich $m = N$ und $n = 2N$. Das ergibt $|s_n - s_m| = |1/(N+1) + \cdots + 1/(2N)| \geq \frac{1}{2}$. □

3. Die b-adischen Brüche $\sum_{k=0}^\infty a_k b^{-k}$ mit Basis $b \in \mathbb{N} \setminus \{1\}$ und Ziffern $a_k \in \{0, 1, \ldots, b-1\}$. Die Partialsummen konvergieren in $\mathbb{R}$ (wird in den Übungen bewiesen).

[6] In den meisten Lehrbüchern wird das als Satz von Bolzano-Weierstraß bezeichnet.

7.2 Konvergenzkriterien für Reihen

Satz 7.3 (Cauchy-Kriterium für Reihen). *$\sum_k a_k$ konvergiert genau dann, wenn für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, sodass $|\sum_{k=n}^m a_k| < \varepsilon$ für alle $m, n \geq N$ gilt.*

Beweis. Das folgt direkt aus dem Cauchy-Kriterium für die Folge $(s_n)_n$. Die Folge $(s_n)_n$ konvergiert genau dann, wenn für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, sodass $|s_n - s_m| < \varepsilon$ für alle $m, n \geq N$ gilt. □

Ende der Vorlesung 9 am 10. November 2021

Satz 7.4. *Eine notwendige (aber nicht hinreichende) Bedingung für die Konvergenz der Reihe $\sum_k a_k$ ist $\lim_{n\to\infty} a_n = 0$.*

Beispiel 7.5. Ein Beispiel dafür, dass die Bedingung nicht hinreichende ist, ist die harmonische Reihe. Sie divergiert, obwohl $1/n \to 0$ für $n \to \infty$.

Beweis. Aus $s_n = \sum_{k=1}^n a_k \to s$ folgt $a_n = s_n - s_{n-1}$, also $a_n \to s - s = 0$ für $n \to \infty$. (Alternativ funktioniert auch das Cauchy-Kriterium für $m = n$.) □

Satz 7.6 (monotone Reihen). *Eine Reihe $\sum a_k$ mit $a_k \geq 0$ konvergiert genau dann, wenn die Folge ihrer Partialsummen beschränkt ist.*

Bemerkung 7.7. Ohne Beschränktheit der Partialsummen würde die Reihe trotzdem in $\overline{\mathbb{R}}$ konvergieren.

Beweis. Offensichtlich ist $s_n = \sum_{k=1}^n a_k$ eine schwach wachsende Folge. Daher konvergiert die Reihe in $\mathbb{R}$ genau dann, wenn sie beschränkt ist (Sätze 4.12 und 6.2). Auch aufgrund der Monotonie konvergiert die Reihe stets in $\overline{\mathbb{R}}$ (Satz 6.10). □

Der folgende Satz ist nützlich, da bei alternierenden Reihen $\sum_k (-1)^k a_k$ die Reihe der Beträge $\sum_k |a_k|$ oft nicht konvergiert und deshalb die Dreiecksungleichung nicht hilft.

Satz 7.8 (alternierende Reihen, Leibniz-Kriterium). *Sei $(a_n)_{n\in\mathbb{N}}$ eine schwach fallende Nullfolge. Dann konvergiert die alternierende Reihe $\sum_{n=1}^\infty (-1)^n a_n$.*

Beweis. Sei $(s_k)_k$ die Folge der Partialsummen von $(a_n)_n$. Wir stellen

$$s_k - s_{k-2} = (-1)^k a_k + (-1)^{k-1} a_{k-1} = (-1)^k (a_k - a_{k-1})$$

für alle k fest. Da $(a_n)_n$ schwach fällt, gilt $a_k - a_{k-1} \leq 0$ und damit $s_1 \leq s_3 \leq s_5 \leq \cdots$ für ungerade k sowie $\cdots \leq s_6 \leq s_4 \leq s_2$ für gerade k. Ferner ist $s_{2k} - s_{2k-1} = (-1)^{2k} a_{2k}$ nichtnegativ und es konvergiert gegen 0 wegen $a_{2k} \to 0$. Also bilden die Intervalle $[s_{2k-1}, s_{2k}]$ eine Intervallschachtelung. Nach dem Intervallschachtelungsprinzip gibt es ein $s \in \mathbb{R}$ in allen Intervallen mit $s = \lim_{k\to\infty} s_{2k} = \lim_{k\to\infty} s_{2k-1}$ und demnach $s = \lim_{k\to\infty} s_k$. □

Beispiel 7.9. $\sum_{n=1}^\infty (-1)^{n+1}/n = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$ konvergiert.

Der folgende Satz ist auch wichtig, da wir somit Reihen auf andere konvergente Reihen zurückführen können.[1]

Satz 7.10 (Majorantenkriterium). *Sei $(a_n)_n$ eine Folge reeller Zahlen und $(b_n)_n$ eine Folge nichtnegativer Zahlen mit der Eigenschaft, dass es ein $N \in \mathbb{N}$ existiert, sodass $|a_n| \leq b_n$ für alle $n \geq N$ gibt. Dann gilt: Konvergiert $\sum_n b_n$, s konvergiert auch $\sum_n a_n$ und es gilt $|\sum_{n=N}^\infty a_n| \leq \sum_{n=N}^\infty b_n$.*

Die Reihe $\sum_{n=N}^\infty b_n$ heißt Majorante *von $\sum_{n=N}^\infty a_n$.*

Beweis. Mit dem Cauchy-Kriterium für $(b_n)_n$ gibt es für jedes $\varepsilon > 0$ ein $N' \geq N$ mit $\sum_{k=m}^n b_k \leq \varepsilon$ für alle $n \geq m \geq N'$. Unter Zunahme der Dreiecksungleichung und der Majoranteneigenschaft können wir

$$\left| \sum_{k=m}^n a_k \right| \leq \sum_{k=m}^n |a_k| \leq \sum_{k=m}^n b_k \leq \varepsilon$$

für alle $n \geq m \geq N'$ schließen, sodass $\sum a_k$ gemäß dem Cauchy-Kriterium konvergiert. Wegen der Monotonie der Limiten (Satz 4.19) gilt $|\sum_{k=N}^\infty| \leq \sum_{k=N}^\infty |a_k| \leq \sum_{k=n}^\infty b_k$. □

[1] Damit verwandt ist das sog. *Minorantenkriterium*: Seien $(a_n)_n$ und $(b_n)_n$ Reihen. Ist $a_n \geq b_n \geq 0$ für alle $n \geq N$ für ein gewisses N, dann gilt: Divergiert die *Minorante* $\sum_n b_n$, dann auch $\sum_n a_n$.

Der folgende Satz versucht eine Folge zu charakterisieren, die sich ab einem Folgenglied wie die geometrische Reihe verhält.

Satz 7.11 (Quotientenkriterium). *Sei $(a_n)_n$ eine reelle Zahlenfolge mit der Eigenschaft, dass es ein $0 < q < 1$ und $N \in \mathbb{N}$ gibt, sodass $a_n \neq 0$ und $|a_{n+1}/a_n| \leq q$ für alle $n \geq N$ gilt. Dann konvergiert auch $\sum a_k$.*

Beweis. Nach rekursiver Anwendung der Voraussetzung folgern wir $a_n \leq q^{n-N}|a_N|$ für alle $n \geq N$. Weil $q < 1$, ist $\sum_{n=N}^k b_n$ mit $b_n = q^{n-N}|a_N|$ eine Majorante von $(a_n)_n$, denn $a_n \leq q^{n-N}|a_N| \leq \sum_{n=N}^k q^{n-N}|a_N|$. Damit gilt $\sum_{n=N}^\infty b_n = |a_N| \sum_{n=0}^\infty q^n = |a_N|/(1-q)$ für $k \to \infty$. ☐

Eine Modifikation des vorherigen Satzes:

Korollar 7.12 (Modifikation des Quotientenkriteriums). *Sei $(a_n)_n$ eine reelle Zahlenfolge mit $a_n \neq 0$ für alle $n \geq N \in \mathbb{N}$. Dann*

1. *konvergiert $(a_n)_n$, wenn $\limsup_{n\to\infty}|a_{n+1}/a_n| < 1$ ist, und*
2. *divergiert $(a_n)_n$, wenn $\liminf_{n\to\infty}|a_{n+1}/a_n| > 1$ ist.*

Beweis.

1. Sei $p = \limsup|a_{n+1}/a_n| < 1$. Nach Lemma 6.7 und def Definition des Limes Superior gibt es für jedes $\varepsilon > 0$ ein N, sodass für alle $n \geq N$ gilt
$$\left|\frac{a_{n+1}}{a_n}\right| - p \leq \sup_{n\geq N}\left|\frac{a_{n+1}}{a_n}\right| - p \leq \varepsilon \implies \left|\frac{a_{n+1}}{a_n}\right| \leq p + \varepsilon.$$
Nun ist aber $p < 1$ und wir können ein ε so klein wählen, sodass $q := p + \varepsilon < 1$ ist. Folglich konvergiert die Reihe nach dem Quotientenkriterium.
2. Sei $p' = \liminf|a_{n+1}/a_n| > 1$. Analog zu oben gilt
$$\varepsilon \geq p' - \inf_{n\geq N}\left|\frac{a_{n+1}}{a_n}\right| \geq p' - \left|\frac{a_{n+1}}{a_n}\right| \implies \left|\frac{a_{n+1}}{a_n}\right| \geq p' - \varepsilon.$$
Da $p' > 1$, können wir ein ε so klein wählen, dass $p' - \varepsilon \geq 1$ ist. Folglich ist (a_n) keine Nullfolge und $\sum a_n$ divergiert nach Satz 7.4. ☐

Bemerkung 7.13.

1. Zum Nachweis der Konvergenz der Reihe $\sum a_n$ reicht nicht, dass $|a_{n+1}/a_n| < 1$ für alle n ist. Der Bruch muss *uniform kleiner* 1 sein. *Beispiel*: Die harmonische Reihe divergiert, aber $|a_{n+1}/a_n| = n/(n+1) < 1$.
2. Die Bedingung $\limsup_{n\to\infty}|a_{n+1}/a_n| < 1$ ist nicht notwendig für die Konvergenz. *Beispiel*: $\sum 1/n^2$ konvergiert, aber
$$\limsup_{n\to\infty}|\frac{a_{n+1}}{a_n}| = \limsup_{n\to\infty}\left(\frac{n}{n+1}\right)^2 = 1.$$

Satz 7.14 (Wurzelkriterium). *Sei $(a_n)_n$ eine Folge reeller Zahlen. Setze $r := \limsup_{n\to\infty}|a_n|^{1/n}$. Dann gilt*

1. *Ist $r < 1$, so konvergiert die Reihe $\sum_n a_n$.*
2. *Ist $r > 1$, so divergiert die Reihe $\sum_n a_n$.*

Beweis.

1. Sei $q \in (r, 1)$. Aus der Konvergenz folgt (ähnlich zum Beweis des Quotientenkriteriums), dass es für alle $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, sodass
$$|a_n|^{1/n} \leq \sup_{n\geq N}|a_n|^{1/n} < r + \varepsilon$$
für alle $n \geq N$. Wählen wir ε klein genug, gilt dann $r + \varepsilon \leq q$, also gibt es ein N sodass für alle $n \geq N$: $|a_n|^{1/n} \leq q$. Aus dem Majorantenkriterium folgt $\sum_{n=N}^\infty|a_n| \leq \sum_{n=N}^\infty q^n < \infty$, also konvergiert die Reihe.

2. Aus $r > 1$ folgt, dass es unendlich viele $n \in \mathbb{N}$ gibt mit $|a_n|^{1/n} > 1$ (gäbe es nur endlich viele so wäre $\sup|a_n|^{1/n} \leq 1$ nach den endlich vielen n). $(a_n)_n$ ist also keine Nullfolge und $\sum_n a_n$ divergiert nach Satz 7.4. $\square$

Bemerkung 7.15. Gilt das Quotientenkriterium, dann folgt auch das Wurzelkriterium: Für alle $n \in \mathbb{N}$ gilt mit $|a_{n+1}/a_n| \leq q < 1$

$$\begin{aligned} r = \limsup_{n\to\infty}|a_{N+n}|^{1/(N+n)} &= \limsup_{n\to\infty}\left(\left(\left|\frac{a_{N+n}}{a_{N+n-1}}\right| \cdot \left|\frac{a_{N+n-1}}{a_{N+n-2}}\right| \cdots \left|\frac{a_{N+1}}{a_N}\right|\right)^{1/(N+n)} \cdot |a_N|^{1/(N+n)}\right) \\ &\leq \limsup_{n\to\infty}\left(q^{n/(N+n)}|a_N|^{1/(N+n)}\right) = q. \end{aligned}$$

Aus dem Wurzelkriterium folgt aber im Allgemeinen nicht das Quotientenkriterium.

Beispiel 7.16. Sei

$$a_n = \begin{cases} 2^{-n} & \text{falls } n \text{ gerade,} \\ 3^{-n} & \text{falls } n \text{ ungerade.} \end{cases}$$

Für ungerade n folgt daraus

$$\frac{a_{n+1}}{a_n} = \frac{2^{-(n+1)}}{3^{-n}} = \frac{1}{2}\left(\frac{3}{2}\right)^n \to \infty,$$

also funktioniert das Quotientenkriterium nicht. Aber das Wurzelkriterium liefert $\lim_{n\to\infty}|a_{2n}|^{1/(2n)} = \frac{1}{2} < 1$.

Obwohl wir irrationale Exponenten noch nicht definiert haben, behaupten wir aber trotzdem folgende allgemeinere Aussage.

Satz 7.17. *Für $s \in \mathbb{R}$ ist die Reihe $\sum_{n=1}^\infty 1/n^s$*

1. *konvergent genau dann, wenn $s > 1$;*
2. *divergent genau dann, wenn $s \leq 1$.*

Bezeichnung 7.18. Im Fall $s > 1$ heißt

$$\zeta(s) := \sum_{n=1}^\infty \frac{1}{n^s}$$

die RIEMANNSCHE *Zeta-Funktion.*

EULER (1734) zeigte einige Werte der Reihe, wie z. B. $\zeta(2) = \frac{1}{6}\pi^2$, $\zeta(4) = \pi^4/90$, $\zeta(6) = \pi^6/954$, usw.

Beweis.

1. Für $s > 1$ folgt mit der geometrischen Reihe (und der Darstellung $n = 2^k + l$)

$$\begin{aligned} \sum_{n=1}^\infty \frac{1}{n^s} = \sum_{k=0}^\infty \sum_{l=0}^{2^k-1} \frac{1}{(2^k+l)^s} &= 1 + \left(\frac{1}{2^s} + \frac{1}{3^s}\right) + \left(\frac{1}{4^s} + \cdots + \frac{1}{7^s}\right) + \left(\frac{1}{8^s} + \cdots + \frac{1}{15^s}\right) + \cdots \\ &\leq 1 + 2\cdot\frac{1}{2^s} + 4\cdot\frac{1}{4^s} + 8\cdot\frac{1}{8^s} + \cdots = \sum_{k=0}^\infty \frac{2^k}{2^{ks}} = \sum_{k=0}^\infty \frac{1}{(2^{s-1})^k} = \frac{1}{1-2^{1-s}} < \infty. \end{aligned}$$

   Also konvergiert die Reihe.

2. Für $s = 1$ ist die Reihe die harmonische Reihe, die divergiert. Für $s < 1$ folgt die Divergenz aus dem Majorantenkriterium, denn $1/n < 1/n^s$.[2] $\square$

Ende der Vorlesung 10 am 15. November 2021

[2] Eigentlich folgt die Aussage aus dem Minorantenkriterium.

Korollar 7.19. *Für $b, c \in \mathbb{R}$, $b \neq 0$ konvergiert*

$$\sum_{n=1} \frac{n^c}{b^n}$$

genau dann, wenn eine der folgenden Bedingungen erfüllt ist:

1. $|b| > 1$,

2. $b = -1$ *und* $c < 0$,

3. $b = 1$ *und* $c < -1$.

Beweis.

1. Für aufeinanderfolgende Glieder gilt

$$\left|\frac{a_{n+1}}{a_n}\right| = \left(\frac{n+1}{n}\right)^c \cdot \frac{1}{|b|} \xrightarrow{n\to\infty} \frac{1}{|b|}.$$

Aus dem Quotientenkriterium folgt, dass die Reihe konvergiert, wenn $|b| > 1$, und divergiert, wenn $|b| \leq 1$. (Genauer: Da $((n+1)/n)^c$ gegen 1 konvergiert, kommt der Term beliebig nah an 1 ran, also $((n+1)/n)^c < 1 + \varepsilon$ für alle $n \geq N$. Damit können wir ε klein genug wählen, sodass $1 + \varepsilon \leq |b|$ für $|b| > 1$ ist. Daraus folgt

$$\left(\frac{n+1}{n}\right)^c \cdot \frac{1}{|b|} < \frac{1+\varepsilon}{|b|} \leq 1,$$

weshalb die Reihe nach dem Quotientenkriterium konvergiert. Für $|b| \leq 1$ ist das Verhältnis aufeinanderfolgender Glieder größergleich 1, also divergiert die Reihe.)

2. Für $b = -1$ ist $(n^c/b^n)_n = (-1)^n n^c$ eine alternierende Folge. Dabei ist die Folge genau dann eine Nullfolge, wenn $c < 0$. Ist also $c < 0$, dann konvergiert die Reihe nach dem Leibniz-Kriterium. Ist $c \geq 0$, dann divergiert sie (Satz 7.4).

3. Für $b = 1$ und $c < -1$ ist das die riemannsche Zeta-Funktion für $s = -c$ (Satz 7.17). □

8 Absolut konvergente Reihen und summierbare Familien

8.1 Motivation

Welche Manipulationen dürfen wir bei Reihen vornehmen, ohne die Konvergenz oder den Wert der Reihe zu verändern?

- *Sind Reihen kommutativ, d. h. können wir Summanden beliebig umordnen?* Nein, im Allgemeinen sind sie es nicht. In der alternierenden harmonischen Reihe könnten wir statt der normalen Reihenfolge immer ein positives, und dann zwei negative Glieder addieren. Die Aussage

$$\underbrace{1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots}_{\ln 2} \stackrel{?}{=} \underbrace{1 - \frac{1}{2} - \frac{1}{4} + \frac{1}{3} - \frac{1}{6} - \frac{1}{8} + \frac{1}{5} - \frac{1}{10} - \frac{1}{12} + \frac{1}{7} - \cdots}_{(\ln 2)/2}$$

ist falsch.[1]

[1]Der Grenzwert der Umordnung lässt sich wie folgt zeigen: Es gilt für die $3n$-ten Partialsummen der Umordnung

$$t_{3n} = \sum_{i=1}^{n}\left(\frac{1}{2i-1} - \frac{1}{4i-2} - \frac{1}{4i}\right) = \sum_{i=1}^{n}\left(\frac{1}{4i-2} - \frac{1}{4i}\right) = \frac{1}{2}\sum_{i=1}^{n}\left(\frac{1}{2i-1} - \frac{1}{2i}\right) = \frac{1}{2}s_{2n}.$$

Damit ist $\lim_{n\to\infty} t_{3n} = \frac{1}{2}\lim_{n\to\infty} s_{2n} = \frac{1}{2}\ln 2$.

- Ferner sei $(a_n)_n$ eine alternierende Folge. Dann gibt es für jedes $s \in \mathbb{R}$ eine Umordnung $f\colon \mathbb{N} \to \mathbb{N}$, sodass $\sum_{n=1}^{\infty} a_{f(n)}$ gegen s konvergiert.[2]

  Beweis: Wir wählen zuerst der Reihe nach die positiven Glieder und summieren sie, bis erstmals s überschritten wird. Dann wählen wir die größten der unverbrauchten negativen Glieder und addieren sie, bis s wieder unterschritten wird. Das wiederholen wir.

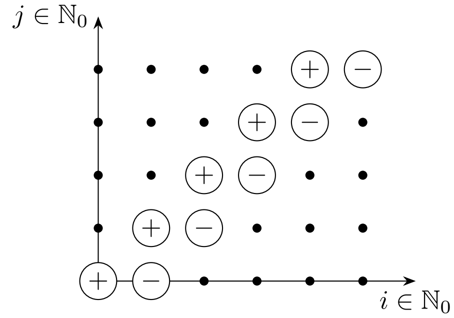
- *Können wir Summenzeichen wie bei endlichen Summen vertauschen?* Nein, im Allgemeinen können wir das nicht. Sei bspw.

$$a_{ij} = \begin{cases} 1 & \text{falls } i = j, \\ -1 & \text{falls } j = i+1, \\ 0 & \text{sonst.} \end{cases}$$

  Dann ist

$$\sum_{i=0}^{\infty} \left(\sum_{j=0}^{\infty} a_{ij} \right) \stackrel{?}{=} \sum_{j=0}^{\infty} \left(\sum_{i=0}^{\infty} a_{ij} \right)$$

  falsch, denn $\sum_{i=0}^{\infty} a_{ij} = 0$ für alle $i \in \mathbb{N}_0$, aber $\sum_{j=0}^{\infty} a_{ij} = 0$ nur für alle $j \in \mathbb{N}$, also nicht für $i = 0$.



- *Sind Reihen assoziativ, d. h. können wir beliebig Klammern setzen?* Nein, im Allgemeinen nicht. Bspw. ist folgende Aussage falsch:

$$(1-1) + (1-1) + (1-1) + \cdots \stackrel{?}{=} 1 + (-1+1) + (-1+1) + (-1+1) + \cdots$$

Definition 8.1 (absolut konvergente Reihen)**.** Eine Reihe $\sum_{n=1}^{\infty} a_n$ heißt *absolut konvergent*, falls die Reihe $\sum_{n=1}^{\infty} |a_n|$ konvergiert.

Bemerkung 8.2.

1. $\sum_n |a_n|$ konvergiert genau dann, wenn $\sup_k \sum_{n=1}^{k} |a_n| < \infty$. Weil die Folge der Partialsummen $\sum_{n=1}^{k} |a_n|$ monoton wächst, folgt das aus den Sätzen 4.12 und 6.2.

2. Wenn $\sum_n |a_n|$ konvergiert, konvergiert auch $\sum_n a_n$. Das folgt aus der Dreiecksungleichung und dem Cauchy-Kriterium für Reihen.

   Achtung: Die Umkehrung gilt nicht, wie die alternierende harmonische Reihe zeigt.

8.2 Umordnung

Definition 8.3 (Umordnung)**.** Eine Reihe $\sum_n b_n$ heißt *Umordnung* der gegebenen Reihe $\sum_n a_n$, falls eine Bijektion $f\colon \mathbb{N} \to \mathbb{N}$ existiert mit $b_n = a_{f(n)}$.

Die neue Reihe summiert also dieselben Folgenglieder, aber in anderer Reihenfolge. Für *endliche* Summen bleibt der Wert der Summe unter Umordnung unverändert.

Satz 8.4 (Umordnungssatz)**.** *Jede Umordnung einer absolut konvergenten Reihe ist wieder absolut konvergent und besitzt denselben Wert.*

[2] Bekannt als *Riemannscher Umordnungssatz.*

Beweis. Sei $A \coloneqq \sum_{n=1}^{\infty} a_n$ eine absolut konvergente Reihe und $\sum_{n=1}^{\infty} a_{f(n)}$ eine Umordnung der Reihe. Wir wollen zeigen, dass $\sum_{n=1}^{\infty} a_{f(n)} = A$.

Wegen der absoluten Konvergenz von $\sum_n a_n$ folgt aus dem Cauchy-Kriterium, dass es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ mit $\sum_{n=N}^{\infty} |a_n| \leq \varepsilon$ gibt. Daraus folgt mit der Dreiecksungleichung

$$\left| A - \sum_{n=1}^{N-1} a_n \right| = \left| \sum_{n=N}^{\infty} a_n \right| \leq \sum_{n=N}^{\infty} |a_n| \leq \varepsilon.$$

Wir wählen nun ein $N' \in \mathbb{N}$ so groß, dass $\{f(1), f(2), \dots, f(N')\}$ mindestens $\{1, 2, \dots, N-1\}$ enthält. Dann gilt für alle $k \geq N'$ mit der Dreiecksungleichung

$$\left| A - \sum_{n=1}^{k} a_{f(n)} \right| \leq \underbrace{\left| A - \sum_{n=1}^{N-1} a_n \right|}_{\leq \varepsilon} + \underbrace{\left| \sum_{n=1}^{N-1} a_n - \sum_{n=1}^{k} a_{f(n)} \right|}_{\substack{\sum a_{f(n)} \text{ enthält mind.} \\ \text{alle } a_n (n \in \{1,\dots,N-1\})}} \leq \varepsilon + \left| \sum_{n=N}^{\infty} a_n \right| \leq 2\varepsilon.$$

Also gilt $\sum_{n=1}^{\infty} a_{f(n)} = A$. □

Wenden wir das auf eine Reihe $\sum |a_n|$ an, heißt das, dass $\sum |a_n|$ genau dann konvergiert, wenn eine Umordnung $\sum |a_{f(n)}|$ konvergiert.

8.3 Summierbare Familien

Definition 8.5 (Familie)**.** Sei I eine beliebige *abzählbare* Indexmenge und $a \colon I \to \mathbb{R}$ eine Abbildung. Anstatt von der Abbildung $a \colon I \to \mathbb{R}$ sprechen wir auch von einer *Familie* $(a_i)_{i \in I}$. Im Fall von $I = \mathbb{N}$ ist es unsere bekannte *Folge*.

Definition 8.6 (absolute Summierbarkeit)**.** Sei $f \colon \mathbb{N} \to I$ eine Abzählung von I und $\sum_{n=1}^{\infty} |a_{f(n)}|$ eine absolut konvergente Reihe. Dann heißt die Familie $(a_i)_{i \in I}$ *(absolut) summierbar* und wir definieren ihre *Summe* als $\sum_{i \in I} a_i \coloneqq \sum_{n=1}^{\infty} a_{f(n)}$.

Die Indizes der Indexmenge können alles mögliche sein, solange es abzählbar viele sind. Mit f definieren wir eine Reihenfolge der Indizes, sodass wir über sie summieren können.

Bemerkung 8.7.

1. Ist $g \colon \mathbb{N} \to I$ eine andere Abzählung von I, so existiert eine Bijektion $\varphi \colon \mathbb{N} \to \mathbb{N}$ mit $g = f \circ \varphi$, die die Argumente von f zu g passend umordnet. Mit $c_n \coloneqq a_{f(n)}$ und $b_n \coloneqq a_{g(n)} = c_{\varphi(n)}$ gilt daher

$$\sum_{n=1}^{\infty} a_{g(n)} = \sum_{n=1}^{\infty} b_n = \sum_{n=1}^{\infty} c_{\varphi(n)} \overset{(8.4)}{=} \sum_{n=1}^{\infty} c_n = \sum_{n=1}^{\infty} a_{f(n)}.$$

   (Dabei können wir Satz 8.4 verwenden, weil $\sum_n |a_{f(n)}|$ absolut konvergiert.) Das bedeutet, dass die Summierbarkeit und der Wert der Summe unabhängig von jeglicher Umordnung ist.

2. $(a_i)_{i \in I}$ ist genau dann summierbar, wenn ein $K \in \mathbb{R}$ existiert, sodass $\sum_{i \in J} |a_i| \leq K$ für alle endlichen Teilmengen J von I gilt.

Beweis zu Punkt 2. Ist (a_i) summierbar, so ist $K = \sum_{n=1}^{\infty} |a_{f(n)}| < \infty$. Für jede endliche Teilmenge $J \subset I$ können wir ein $N \in \mathbb{N}$ finden, sodass $\{f(1), \dots, f(N)\}$ mindestens alle Elemente aus J enthält. Sodann haben wir

$$\sum_{i \in J} |a_i| \leq \sum_{n=1}^{N} |a_{f(n)}| \leq K.$$

Gibt es hingegen ein K, sodass $\sum_{i \in J} |a_i| \leq K$ für alle endlichen $J \subset I$ erfüllt ist, so wird die Ungleichung auch für alle endliche Mengen $\{f(1), \dots, f(N)\}$ erfüllt. Folglich ist die Reihe der Partialsummen $\sum_{n=1}^{N} a_{f(n)} \leq K$ beschränkt und damit $\sum_{n=1}^{\infty} |a_{f(n)}| \leq K$. □

Für I nehmen wir typischerweise $\mathbb{N} \times \mathbb{N}$, $\mathbb{Z} \times \mathbb{Z}$ oder $\mathbb{Z}^d = \{z = \{z_1, \dots, z_d\} : z_i \in \mathbb{Z}\}$ für ein $d \in \mathbb{N}$.
Der folgende Satz ist eine diskrete Version und damit ein Spezialfall des *Satzes von* FUBINI für Integrale.[3]

Satz 8.8 (großer Umordnungssatz)**.** *Seien I und L abzählbare Mengen und I_ℓ für alle $\ell \in L$ paarweise disjunkte Teilmengen von I mit $I = \bigcup_{\ell \in L} I_\ell$ (d. h. $\{I_\ell : \ell \in L\}$ ist eine Partition von I, wir schreiben kurz $\dot{\bigcup}_{\ell \in L} I_\ell$). Ferner sei $(a_i)_{i \in I}$ eine summierbare Familie reeller Zahlen. Dann ist jede Teilfamilie $(a_i)_{i \in I_\ell}$ mit $\ell \in L$ summierbar, also auch die Familie $(s_\ell)_{\ell \in L}$ der Summen $s_\ell \coloneqq \sum_{i \in I_\ell} a_i$ und es gilt*

$$\sum_{i \in I} a_i = \sum_{\ell \in L} s_\ell = \sum_{\ell \in L} \left(\sum_{i \in I_\ell} a_i \right).$$

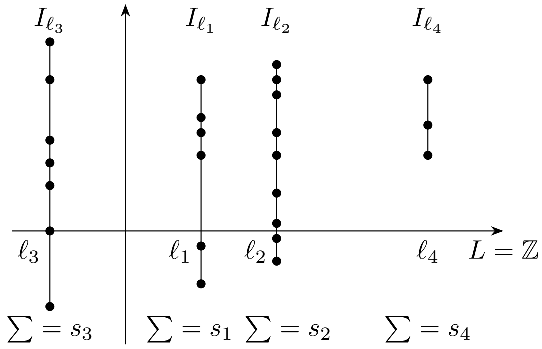


Abbildung 8.1: Beispiel und Visualisierung von Satz 8.8.
Die Menge aller Punkte ist $I = \bigcup_{\ell \in L} I_\ell$ und es gilt $\sum_{\ell \in I} s_\ell = \sum_{i \in I} a_i = \sum_{\ell \in L} \big(\sum_{i \in I_\ell} a_i \big)$.

Ende der Vorlesung 11 am 17. November 2021

Beweis. Wir zeigen die behaupteten Aussagen aus dem Satz.

1. Summierbarkeit der Teilfamilien $(a_i)_{i \in I_\ell}$: Es gilt $\sum_{i \in I_\ell} |a_i| \leq \sum_{i \in I} |a_i| < \infty$ für alle $\ell \in L$.

2. Summierbarkeit der Familie $(s_\ell)_{\ell \in L}$: Weil die Teilfamilien summierbar sind, also $\sum_{i \in I_\ell} |a_i|$ konvergiert, konvergiert auch $\big|\sum_{i \in I_\ell} a_i\big|$. Daher können wir eine geeignete endliche Teilmenge $J_\ell \subset I_\ell$ finden, sodass die endliche Summe $\big|\sum_{i \in J_\ell} a_i\big|$ beliebig nah (ε-nah) an den „Grenzwert" $\big|\sum_{i \in I_\ell} a_i\big|$ kommt. Also gibt es für jedes $\varepsilon > 0$ ein $J_\ell \subset I_\ell$, sodass gilt:

$$\left| \sum_{i \in I_\ell} a_i \right| \leq \left| \sum_{i \in J_\ell} a_i \right| + \varepsilon.$$

Sei $\varphi \colon \mathbb{N} \to L$ eine (bijektive) Abzählung der Indexmenge L (bzw. für endliche Mengen $\varphi \colon \{1, \dots, N\} \to L$, was trivial ist). Dann gilt für alle $N_0 \in \mathbb{N}$

$$\sum_{n=1}^{N_0} \big|s_{\varphi(n)}\big| = \sum_{n=1}^{N_0} \left| \sum_{i \in I_{\varphi(n)}} a_i \right| \leq \sum_{n=1}^{N_0} \left(\left| \sum_{i \in J_{\varphi(n)}} a_i \right| + \frac{1}{2^n} \right) \leq \sum_{i \in \bigcup_{n=1}^{N_0} J_{\varphi(n)}} |a_i| + 1 \leq \sup_{\text{endliche } J \subset I} \left(\sum_{i \in J} |a_i| + 1 \right).$$

Dabei haben wir für die vorletzte Ungleichung die Dreiecksungleichung verwendet und $1/2^n$ durch die geometrische Reihe abgeschätzt. Für die letzte Ungleichung haben wir die Supremumseigenschaft angewendet (da $(a_i)_{i \in I}$ summierbar und demnach nach oben beschränkt ist). Damit ist $(s_\ell)_{\ell \in L}$ nach oben beschränkt und folglich summierbar.

[3] Der *Satz von* FUBINI beschreibt, dass ein mehrdimensionales Integral durch Komposition eindimensionaler Integrale ausgedrückt werden kann, also $\int_{[a,b] \times [c,d]} f(x, y) \, \mathrm{d}(x, y) = \int_c^d \int_a^b f(x, y) \, \mathrm{d}x \, \mathrm{d}y$.

3. Sei a die Summe von $(a_i)_{i\in I}$. Seien f und φ bijektive Abzählungen von I bzw. L (der Einfachheit halber nehmen wir an, dass I und L abzählbar unendlich sind).

   Weil die Familie $(a_i)_{i\in I}$ summierbar ist, folgt aus dem Cauchy-Kriterium für Reihen: Für jedes $\varepsilon > 0$ gibt es ein $K_0 \in \mathbb{N}$, sodass für alle $K \geq K_0$

$$\left|a - \sum_{k=1}^{K} a_{f(k)}\right| \leq \sum_{k=K+1}^{\infty} |a_{f(k)}| \leq \varepsilon$$

   gilt. Dabei verwendeten wir für die erste Ungleichung die Dreiecksungleichung.

   Weil auch jede Teilfamilie $(a_i)_{i\in I_{\varphi(n)}}$ summierbar ist, gilt (analog zu oben) wegen dem Cauchy-Kriterium und der Dreiecksungleichung für jede Teilfamilie: Für jedes $\varepsilon > 0$ gibt es eine endliche Teilmenge $J_{\varphi(n)} \subset I_{\varphi(n)}$, sodass für alle endlichen Teilmengen $J^*_{\varphi(n)}$ mit $J_{\varphi(n)} \subset J^*_{\varphi(n)} \subset I_{\varphi(n)}$ gilt:

$$\left|s_{\varphi(n)} - \sum_{i\in J^*_{\varphi(n)}} a_i\right| \leq \frac{\varepsilon}{2^n}.$$

   Wir betrachten $\{f(1),\dots,f(K_0)\}$ mit K_0 wie oben festgelegt. Da es endlich viele sind, können wir ein hinreichend großes $N_0 \in \mathbb{N}$ wählen, sodass $\{f(1),\dots,f(K_0)\} \subset \bigcup_{n=1}^{N_0} I_{\varphi(n)}$ gilt. Genauso können wir aber hinreichend große, endliche Teilmengen $J^*_{\varphi(n)}$ mit $J_{\varphi(n)} \subset J^*_{\varphi(n)} \subset I_{\varphi(n)}$ wählen ($J_{\varphi(n)}$ und $J^*_{\varphi(n)}$ wie oben definiert), sodass

$$\{f(1),\dots,f(K_0)\} \subset \bigcup_{n=1}^{N_0} J^*_{\varphi(n)} \tag{8.9}$$

   erfüllt ist.

   Ferner können wir aber für jedes $N \geq N_0$ ein hinreichend großes $K \geq K_0$ finden, sodass

$$\bigcup_{n=1}^{N} J^*_{\varphi(n)} \subset \{f(1),\dots,f(K)\}.$$

   Damit können wir für alle $N \geq N_0$ abschätzen:

$$\begin{aligned}
\left|a - \sum_{n=1}^{N} s_{\varphi(n)}\right| &= \left|\left(a - \sum_{k=1}^{K} a_{f(k)}\right) - \left(\sum_{n=1}^{N} s_{\varphi(n)} - \sum_{i\in\bigcup_{n=1}^{N} J^*_{\varphi(n)}} a_i\right) + \left(\sum_{k=1}^{K} a_{f(k)} - \sum_{i\in\bigcup_{n=1}^{N} J^*_{\varphi(n)}} a_i\right)\right| \\
&\leq \underbrace{\left|a - \sum_{k=1}^{K} a_{f(k)}\right|}_{\leq\varepsilon} + \underbrace{\left|\sum_{n=1}^{N} s_{\varphi(n)} - \sum_{i\in\bigcup_{n=1}^{N} J^*_{\varphi(n)}} a_i\right|}_{\leq\sum_{n=1}^{N}\varepsilon/2^n\leq\varepsilon} + \left|\sum_{k=1}^{K} a_{f(k)} - \sum_{i\in\bigcup_{n=1}^{N} J^*_{\varphi(n)}} a_i\right| \\
&\leq \left|\sum_{k=1}^{K} a_{f(k)} - \sum_{i\in\bigcup_{n=1}^{N} J^*_{\varphi(n)}} a_i\right| + 2\varepsilon \overset{(*)}{\leq} \sum_{k=K_0+1}^{K} |a_{f(k)}| + 2\varepsilon \leq 3\varepsilon,
\end{aligned}$$

   wobei wir für $(*)$ die Beziehung (8.9) verwendeten und noch fehlende positive Terme $|a_{f(k)}|$ bis K_0+1 zur Summe hinzufügten (um die Notation zu vereinfachen). Folglich finden wir zu jedem $\varepsilon > 0$ ein N_0, sodass für alle $N \geq N_0$ das Cauchy-Kriterium gilt.[4] □

Bemerkung 8.10. Beide Umordnungssätze (Sätze 8.4 und 8.8) gelten auch ohne Voraussetzung der absoluten Konvergenz bzw. Summierbarkeit, falls alle $a_i \geq 0$.

Konvergiert die Reihe, so ist sie auch absolut konvergent. Divergiert die Reihe, so muss sie uneigentlich gegen ∞ konvergieren. Das bedeutet aber, dass auch jede Umordnung uneigentlich gegen ∞ konvergiert, denn andernfalls würde das dem Umordnungssatz widersprechen.

Genauso lässt sich für die Summierbarkeit von Familien und deren Zerlegung in summierbare Teilfamilien argumentieren.

[4]Zur Intuition des Beweises: Wir betrachten folgendes Bild:

Korollar 8.11 (Doppelreihensatz)**.** *Sei* $(a_{j\ell})_{(j,\ell)\in J\times L}$ *eine summierbare Familie. Dann gilt*

$$\sum_{(j,\ell)\in J\times L} a_{j\ell} = \sum_{j\in J}\left(\sum_{\ell\in L} a_{j\ell}\right) = \sum_{\ell\in L}\left(\sum_{j\in J} a_{j\ell}\right),$$

d. h. die Summenzeichen lassen sich vertauschen.

8.3.1 Beispiele für Umordnung

Beweis. Wir verwenden den obigen Satz mit $I = J \times L$ und $I_\ell = J \times \{\ell\}$. Dann ist $I = \dot{\bigcup}_{\ell\in L} I_\ell$. Daraus folgt $\sum a_{j\ell} = \sum_{\ell\in L}(\sum_{j\in J} a_{j\ell})$. Vertauschen der Rollen von J und L liefert die zweite Gleichheit. □

Beispiel 8.12. Für $J, L = \mathbb{N}_0$ definieren wir

$$a_{j\ell} = \begin{cases} 1 & \text{falls } j = 1, \\ -1 & \text{falls } j = \ell + 1, \\ 0 & \text{sonst.} \end{cases}$$

Dann gilt $\sum_\ell \sum_j a_{j\ell} = 0$, aber $\sum_j \sum_\ell a_{j\ell} = +1$. Das ist auch offensichtlich, da $(a_{j\ell})_{(j,\ell)}$ nicht absolut summierbar ist!

Beispiel 8.13. Mit der geometrischen Reihe gilt

$$\sum_{k=2}^{\infty}(\zeta(k) - 1) = \sum_{k=2}^{\infty}\sum_{n=2}^{\infty}\frac{1}{n^k} = \sum_{n=2}^{\infty}\sum_{k=2}^{\infty}\frac{1}{n^k} = \sum_{n=2}^{\infty}\frac{1}{n(n-1)} = \sum_{n=2}^{\infty}\left(\frac{1}{n-1} - \frac{1}{n}\right).$$

Bezeichnung 8.14 (Wahrscheinlichkeitsverteilung, Erwartungswert)**.** Eine Folge $(p_k)_{k\in\mathbb{N}_0}$ heißt *Wahrscheinlichkeitsverteilung* auf N_0, falls $p_k > 0$ für alle k und $\sum_{k=0}^{\infty} p_k = 1$.

Als *Erwartungswert* dieser Wahrscheinlichkeitsverteilung definieren wir $\mu = \sum_{k=0}^{\infty} kp_k$, falls dieser endlich ist bzw. konvergiert.

Satz 8.15. *Für den Erwartungswert gilt* $\mu = \sum_{j=0}^{\infty}\left(\sum_{k>j} p_k\right)$.

Beweis. Wir setzen

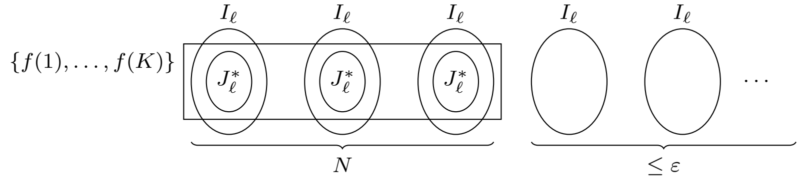
$$p_{jk} := 1_{\{k>j\}} p_k = \begin{cases} 0 & \text{falls } k \le j, \\ 1 & \text{falls } k > j. \end{cases} \quad \text{für } j, k \in \mathbb{N}_0.$$

Hierbei bezeichnet $1_{\{k>j\}}$ die sog. *Indikatorfunktion* auf $\{k > j\}$. Dann gilt mit dem Doppelreihensatz mit $p_k \ge 0$

$$\mu = \sum_{k=0}^{\infty} kp_k = \sum_{k=0}^{\infty}\left(\sum_{j=0}^{\infty} p_{jk}\right) = \sum_{j=0}^{\infty}\left(\sum_{k=0}^{\infty} p_{jk}\right) = \sum_{j=0}^{\infty}\left(\sum_{k>j} p_k\right).$$ □

Beispiel 8.16. Für $p \in (0,1)$ definieren wir $p_k = p(1-p)^{k-1}$ für alle $k \in \mathbb{N}$ und $p_0 = 0$. Diese Verteilung gibt die Wahrscheinlichkeit, dass der k-te Versuch zum ersten Mal zum Erfolg führt, an. Nun ist

$$\sum_{k=0}^{\infty} p_k = p\sum_{k=1}^{\infty}(1-p)^{k-1} = p\sum_{k=0}^{\infty}(1-p)^k = p\frac{1}{1-(1-p)} = 1,$$



Ziel ist es, zu jedem ε hinreichend große N zu finden, sodass die Summe der ersten N Teilfamilien so groß wird, sodann der Rest kleiner ε ist. Das können wir erreichen, indem wir hinreichend große endliche Teilmengen $J_\ell^* \subset I_\ell$ betrachten, und darum eine größere Menge $\{f(1), \dots, f(K)\}$ konstruieren, sodass $J_\ell^* \subset \{f(1), \dots, f(K)\}$ für die ersten N Mengen J_ℓ^*. Da wir wissen, dass die Summe der $(a_i)_{i\in I}$ mit $I \supset \{f(1), \dots, f(K)\}$ gegen a konvergiert, können wir dementsprechend N groß genug wählen.

also ist p eine Wahrscheinlichkeitsverteilung. Des Weiteren ist

$$\sum_{k=0}^{\infty} = kp_k = \sum_{j=0}^{\infty} \sum_{k=j+1}^{\infty} p(1-p)^{k-1} = \sum_{j=0}^{\infty} p\left(\frac{1}{1-(1-p)} - \frac{1-(1-p)^j}{1-(1-p)}\right) = \sum_{j=0}^{\infty} (1-p)^j = \frac{1}{p}.$$

Interpretation: Bei einer Erfolgswahrscheinlichkeit p müssen wir im Mittel $1/p$ Versuche durchführen, um einmal zu gewinnen.[5]

8.4 Produkte von Reihen

Definition 8.17 (diskrete Faltung)**.** Für zwei reelle Folgen $(a_n)_{n\in\mathbb{N}_0}$ und $(b_n)_{n\in\mathbb{N}_0}$ definieren wir deren *Faltung* als Folge $(c_n)_{n\in\mathbb{N}_0}$ mit Glieder

$$c_n := (a*b)_n := \sum_{k=0}^{n} a_k b_{n-k}.$$

Allgemein für Familien $(a_n)_{n\in\mathbb{Z}}$ und $(b_n)_{n\in\mathbb{Z}}$ definieren wir

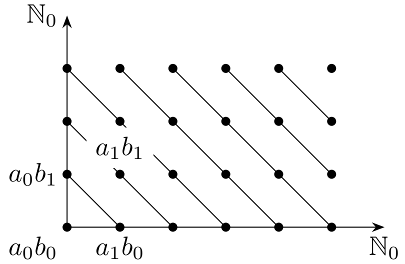
$$(a*b)_n := \sum_{k\in\mathbb{Z}} a_k b_{n-k}$$

(vorausgesetzt die Reihe konvergiert).

Satz 8.18 (Cauchy-Produkt)**.** *Sind $\sum a_n$ und $\sum b_n$ absolut konvergente Reihen, so definiert die Faltung der Folgenglieder eine absolut konvergente Reihe $\sum_n (a*b)_n$, genannt* Cauchy-Produkt *der Reihen*

$$\left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) = \sum_{j,l=0}^{\infty} a_j b_l = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} a_n b_{n-k}\right) = \sum_{n=0}^{\infty} (a*b)_n.$$

Ende der Vorlesung 12 am 22. November 2021



Beweis. Seien $A := \sum_{j=0}^{\infty}|a_j|$, $B := \sum_{k=0}^{\infty}|b_k|$, $a := \sum_{j=1}^{\infty} a_j$ und $b := \sum_{k=1}^{\infty} b_k$.

Zuerst wollen wir zeigen, dass die Familie $(a_j b_k)_{(j,k)\in\mathbb{N}^2}$ summierbar ist. Für alle $N \in \mathbb{N}$ gilt nämlich

$$\sum_{j,k=0}^{N} |a_j b_k| = \sum_{j=0}^{N}\left(\sum_{k=0}^{N} |a_j|\cdot|b_k|\right) \le \sum_{j=0}^{N} |a_j| B \le AB < \infty.$$

Deshalb können wir den Doppelreihensatz anwenden und erhalten

$$\sum_{(j,k)\in\mathbb{N}_0^2} a_j b_k = \sum_{j=0}^{\infty}\left(\sum_{k=0}^{\infty} a_j b_k\right) = \sum_{j=0}^{\infty} a_j b = ab.$$

Wir betrachten nun eine Zerlegung von $\mathbb{N}_0 \times \mathbb{N}_0 = I = \dot{\bigcup}_{\ell=0}^{\infty} I_\ell$ mit

$$I_\ell := \{(j,k) \in \mathbb{N}_0 \times \mathbb{N}_0 : j + k = \ell\} = \{(j, \ell - j) : j \in \{0, 1, \dots, \ell\}\},$$

[5] Diese Verteilung ist als *geometrische Verteilung* bekannt.

also eine Zerlegung in geordneten Paaren mit vorgegebener Summe (den Geraden im Bild). Aus dem großen Umordnungssatz folgt

$$\sum_{(j,k)\in\mathbb{N}_0^2} a_j b_k = \sum_{\ell=0}^{\infty} \sum_{(j,k)\in I_\ell} a_j b_k = \sum_{\ell=0}^{\infty}\sum_{j=0}^{\ell} a_j b_{\ell-j} = \sum_{\ell=0}^{\infty}(a*b)_\ell.$$

Damit sind der obige Ausdruck und ab gleich der Summe der Familie. □

8.4.1 Binomische Reihe

Satz 8.19 (binomische Reihe)**.** *Für alle $x, \alpha \in \mathbb{R}$ mit $|x| < 1$ ist die Reihe*

$$B(x,\alpha) := \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n = 1 + \alpha x + \frac{\alpha(\alpha-1)}{1\cdot 2}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{1\cdot 2\cdot 3}x^3 + \cdots$$

absolut konvergent.

Die Summe ist endlich, falls $\alpha \in \mathbb{N}$.

Beweis. Mit dem Quotientenkriterium folgt die absolute Konvergenz der Reihe:

$$\left|\frac{|a_{n+1}|}{|a_n|}\right| = \left|\frac{|\alpha - n|\cdot|x|}{n+1}\right| \xrightarrow{n\to\infty} |x| < 1.$$ □

Lemma 8.20. *Für alle $\alpha, \beta, x \in \mathbb{R}$ mit $|x| < 1$ gilt*

$$B(x,\alpha)\cdot B(x,\beta) = B(x,\alpha+\beta).$$

Beweis. Es gilt für das Cauchy-Produkt (Satz 8.18)[6]

$$B(x,\alpha)\cdot B(x,\beta) = \sum_{n=0}^{\infty}\left(\sum_{k=0}^{n}\binom{\alpha}{k}x^k\binom{\beta}{n-k}x^{n-k}\right) = \sum_{n=0}^{\infty}\binom{\alpha+\beta}{n}x^n = B(x,\alpha+\beta).$$ □

Satz 8.21. *Für $\alpha \in \mathbb{Q}^+$ und für $x \in \mathbb{R}$ mit $|x| < 1$ gilt*

$$B(x,\alpha) = (1+x)^\alpha.$$

Beweis. Sei $\alpha = p/q$. Mit dem binomischen Lehrsatz sehen wir $(1+x)^p = B(x,p)$. Mit dem gerade bewiesenen Lemma erhalten wir $B(x,p) = B(x,q\alpha) = (B(x,\alpha))^q$, also $(1+x)^\alpha = (1+x)^{p/q} = B(x,\alpha)$. □

9 Die Exponenzialreihe

Definition 9.1 (Exponenzialreihe, eulersche Zahl)**.** Für alle $x \in \mathbb{R}$ sei die *Exponenzialreihe*

$$\exp(x) := \sum_{n=0}^{\infty}\frac{x^n}{n!}.$$

Insbesondere ist die *EULERSCHE Zahl*

$$e := \exp(1) = \sum_{n=0}^{\infty}\frac{1}{n!} = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \cdots \approx 2{,}7182818\ldots$$

[6] Der Dozent verwendete hier die *CHU-VANDERMONDESCHE Identität*, die wir noch nicht bewiesen haben. Führen wir die *fallende Faktorielle* $\alpha^{\underline{n}} := \prod_{i=0}^{n-1}(\alpha - i)$ ein, können wir den Binomialkoeffizienten als $\binom{\alpha}{n} = \alpha^{\underline{n}}/n!$ schreiben. Damit wird die Identität

$$\frac{(\alpha+\beta)^{\underline{n}}}{n!} = \binom{\alpha+\beta}{n} \stackrel{!}{=} \sum_{k=0}^{n}\binom{\alpha}{k}\binom{\beta}{n-k} = \sum_{k=0}^{n}\frac{\alpha^{\underline{k}}\beta^{\underline{n-k}}}{k!(n-k)!} \iff (\alpha+\beta)^{\underline{n}} = \sum_{k=0}^{n}\binom{n}{k}\alpha^{\underline{k}}\beta^{\underline{n-k}}.$$

Das können wir analog zum normalen binomischen Lehrsatz via Induktion über n beweisen.

9.1 Eigenschaften

Satz 9.2.

1. *Für alle $x \in \mathbb{R}$ ist die Exponenzialreihe absolut konvergent.*
2. *Restgliedabschätzung: Für alle $x \in \mathbb{R}$ und $N \in \mathbb{N}$ mit $|x| \leq 1 + \frac{1}{2}N$ gilt*

$$\left| \exp(x) - \sum_{n=0}^{N} \frac{x^n}{n!} \right| \leq 2 \frac{|x|^{N+1}}{(N+1)!}.$$

Die Restgliedabschätzung ist – öfter als man denkt – nützlich, und lässt sich auch einfach merken: Der Rest ist kleiner als das doppelte des ersten Restglieds.

Beweis.

1. Nach dem Quotientenkriterium gilt

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{|x|}{n+1} \xrightarrow{n\to\infty} 0.$$

2. Aus der Dreiecksungleichung und der geometrischen Reihe folgt

$$\begin{aligned} \left| \exp(x) - \sum_{n=0}^{N} \frac{x^N}{n!} \right| &\leq \sum_{n=N+1}^{\infty} \frac{|x|^n}{n!} = \frac{|x|^{N+1}}{(N+1)!} \left(1 + \frac{|x|}{N+2} + \frac{|x|^2}{(N+2)(N+3)} + \cdots \right) \\ &\leq \frac{|x|^{N+1}}{(N+1)!} \left(1 + \frac{|x|}{N+2} + \frac{|x|^2}{(N+2)^2} + \frac{|x|^3}{(N+2)^3} + \cdots \right) = \frac{|x|^{N+1}}{(N+1)!} \sum_{l=0}^{\infty} \left(\frac{|x|}{N+2} \right)^l \\ &\leq 2 \frac{|x|^{N+1}}{(N+1)!}. \end{aligned}$$

□

Satz 9.3 (Funktionalgleichung)**.** *Für alle $x, y \in \mathbb{R}$ gilt*

$$\exp(x) \cdot \exp(y) = \exp(x+y).$$

Beweis. Mit dem Cauchy-Produkt für absolut konvergente Reihen und dem binomischen Lehrsatz gilt

$$\exp(x) \cdot \exp(y) = \left(\sum_{n=0}^{\infty} \frac{x^n}{n!} \right) \left(\sum_{n=0}^{\infty} \frac{y^n}{n!} \right) = \sum_{n=0}^{\infty} \left(\sum_{k=0}^{n} \frac{x^k}{k!} \cdot \frac{y^{n-k}}{(n-k)!} \right) = \sum_{n=0}^{\infty} \frac{(x+y)^n}{n!} = \exp(x+y).$$

□

Bemerkung 9.4. Die Exponenzialfunktion exp definiert einen Gruppenhomöomorphismus von der additiven Gruppe $(\mathbb{R}, +)$ auf die multiplikative Gruppe $((0, \infty), \cdot)$.

Korollar 9.5.

1. *Für alle $x \in \mathbb{R}$ gilt*

$$\exp(x) > 0 \qquad \textit{und} \qquad \exp(-x) = \frac{1}{\exp(x)}.$$

2. *Für alle $x \in \mathbb{R}$ und $q \in \mathbb{Q}$ gilt*

$$\exp(qx) = (\exp(x))^q.$$

Insbesondere folgt daraus auch $\exp(q) = e^q$.

Beweis.

1. Aus der Funktionalgleichung folgt

$$\exp(x) = \exp\left(\frac{x}{2}\right) \exp\left(\frac{x}{2}\right) = \left(\exp\left(\frac{x}{2}\right) \right)^2 \geq 0.$$

Ferner gilt $\exp(x) \neq 0$, denn für alle $x \in \mathbb{R}$ gilt $1 = \exp(0) = \exp(x) \cdot \exp(-x)$, und aufgrund Nullteilerfreiheit ist keiner der Faktoren Null.

Ferner folgt daraus $1/\exp(x) = \exp(-x)$.

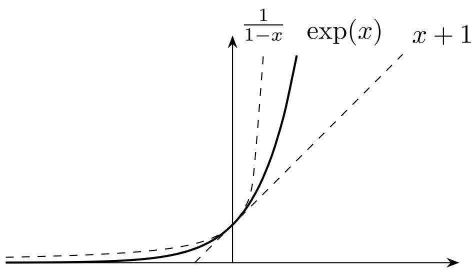
2. Wir beweisen zuerst $\exp(nx) = (\exp(x))^n$ für natürliche n. Der Anfang $n = 1$ ist klar. Der Schritt:
$$\exp((n+1)x) = \exp(x) \cdot \exp(nx) = \exp(x) \cdot \exp(x)^n = (\exp(x))^{n+1}.$$
Ferner gilt
$$\exp(-nx) = \frac{1}{\exp(nx)} = \frac{1}{(\exp(x))^n} = \exp(x)^{-n},$$
also gilt die Gleichung auch für alle ganzen Zahlen. Schließlich erhalten wir für alle $q = k/m$, $k \in \mathbb{Z}$ und $m \in \mathbb{N}$
$$(\exp(qx))^m = \exp(mqx) = \exp(kx) = (\exp(x))^k \implies \exp(qx) = (\exp(x))^{k/m} = (\exp(x))^q. \qquad \square$$

Proposition 9.6.

1. *Es gilt* $\exp(x) \geq 1 + x$ *für alle* $x \in \mathbb{R}$.
2. *Es gilt* $\exp(x) \leq 1/(1-x)$ *für alle* $x < 1$.

Beweis.

1. Für $x \geq 0$ gilt natürlich $\sum_{n=0}^{\infty} x^n/n! \geq 1 + x$.
   Für $x < 0$ betrachten wir $\exp(-x)$ mit $x > 0$. Für $|x| < 1$ ist $\exp(-x) = \sum_{n=0}^{\infty}(-x)^n/n!$ nach dem Leibniz-Kriterium konvergent, denn die Reihe ist alternierend, der Betrag $x^n/n!$ ist monoton fallend und konvergiert gegen Null. Aus dem *Beweis* des Leibniz-Kriteriums (s. Satz 7.8) geht hervor, dass für die Partialsummen $s_{2n-1} \geq s_1$ für alle $n \geq 1$ gilt, also auch $\exp(-x) \geq s_1 = 1 - x$.
   Diese Abschätzung gilt auch für $x \geq 1$, denn es gilt stets $\exp(-x) > 0$, aber $1 - x \leq 0$. Schließlich erhalten wir $\exp(-x) \geq 1 - x$ für alle $x > 0$, also $\exp(x) \geq 1 + x$ für alle $x \in \mathbb{R}$.
2. Das folgt aus $\exp(-x) \geq 1 - x$ durch Kehrwertbildung:
$$\exp(x) = \frac{1}{\exp(-x)} \leq \frac{1}{1-x},$$
   gültig für alle $x < 1$. $\square$



9.2 Grenzwertdarstellung

Satz 9.7 (Grenzwertdarstellung der Exponenzialfunktion)**.** *Für alle* $x \in \mathbb{R}$ *gilt*
$$\lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n = \exp(x).$$
Insbesondere gilt $e = \lim_{n\to\infty}(1 + 1/n)^n$.

Der obige Satz impliziert auch, dass $\lim_{n\to\infty}(1 + x/n)^n$ existiert.

Beweis. Zuerst betrachten wir den Fall $x \geq 0$. Nach dem binomischen Lehrsatz gilt für alle $n \in \mathbb{N}$ mit $x \geq 0$
$$\left(1 + \frac{x}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \frac{x^k}{n^k} = 1 + \sum_{k=1}^{n} \frac{x^k}{k!} \cdot \underbrace{\frac{n(n-1)\cdots(n-k+1)}{n \cdot n \cdots n}}_{k \text{ Faktoren}, \leq 1} \leq 1 + \sum_{k=1}^{\infty} \frac{x^k}{k!} = \exp(x).$$

Umgekehrt gilt mit dem binomischen Lehrsatz für alle $n \geq m$ mit $x \geq 0$

$$\Big(1 + \frac{x}{n}\Big)^n = \sum_{k=0}^{n} \binom{n}{k} \frac{x^k}{n^k} \geq \sum_{k=0}^{m} \frac{x^k}{k!} \Big(\frac{n}{n} \cdot \frac{n-1}{n} \cdots \frac{n-k+1}{n}\Big) \xrightarrow{n\to\infty} \sum_{k=0}^{m} \frac{x^k}{k!} \cdot 1.$$

(*Achtung*: Wir haben in der obigen Abschätzung nur bis m summiert.) Das bedeutet nach dem Satz von Bolzano-Weierstraß

$$\liminf_{n\to\infty} \Big(1 + \frac{x}{n}\Big)^n \geq \sum_{k=0}^{m} \frac{x^k}{k!} \quad \text{für jedes } m \in \mathbb{N}.$$

Lassen wir nun $m \to \infty$, erhalten wir

$$\liminf_{n\to\infty} \Big(1 + \frac{x}{n}\Big)^n \geq \sum_{k=0}^{\infty} \frac{x^k}{k!} = \exp(x).$$

Zusammen erhalten wir für alle $x \geq 0$

$$\lim_{n\to\infty} \Big(1 + \frac{x}{n}\Big)^n = \exp(x).$$

Nun zum Fall $x < 0$ bzw. $\exp(-x)$ für $x > 0$. Für $x > 0$ und hinreichend große n gilt

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

Mit $1 + x/n > 0$ und der Ordnungserhaltung unter $\limsup$ folgt daraus[1]

$$\limsup_{n\to\infty} \Big(1 - \frac{x}{n}\Big)^n \leq \limsup_{n\to\infty} \frac{1}{(1 + x/n)^n} \leq \frac{1}{\liminf_{n\to\infty} (1 - x/n)^n} = \frac{1}{\exp(x)} = \exp(-x).$$

Ferner gilt für alle $x > 0$: Für jedes $\varepsilon > 0$ gibt es hinreichend große n, sodass

$$\Big(1 + \frac{x+\varepsilon}{n}\Big)\Big(1 - \frac{x}{n}\Big) = 1 + \frac{\varepsilon}{n} - \frac{x(x+\varepsilon)}{n^2} \geq 1$$

(z. B. für alle $n \geq n_0$ mit $\varepsilon n_0 \geq x(x+\varepsilon)$). Daraus folgt wie oben mit der Ordnungserhaltung von $\liminf$

$$\liminf_{n\to\infty} \Big(1 - \frac{x}{n}\Big)^n \geq \liminf_{n\to\infty} \frac{1}{(1 + (x+\varepsilon)/n)^n} = \lim_{n\to\infty} \frac{1}{(1 + (x+\varepsilon)/n)^n} = \frac{1}{\exp(x+\varepsilon)} = \exp(-x-\varepsilon),$$

wobei wir beim Übergang von $\liminf$ auf $\lim$ ausgenutzt haben, dass der Bruch (für $x + \varepsilon \geq 0$) konvergiert. Nun folgt aus der Funktionalgleichung und Proposition 9.6

$$\exp(-x-\varepsilon) = \exp(-x) \cdot \exp(-\varepsilon) \geq \exp(-x)(1-\varepsilon).$$

Schließlich[2] erhalten wir für jedes $\varepsilon > 0$

$$(1-\varepsilon)\exp(-x) \leq \liminf \Big(1 - \frac{x}{n}\Big)^n \leq \limsup \Big(1 - \frac{x}{n}\Big)^n \leq \exp(-x) \implies \lim_{n\to\infty} \Big(1 - \frac{x}{n}\Big)^n = \exp(-x). \qquad \square$$

Ende der Vorlesung 13 am 24. November 2021

Beispiel 9.8. Sei x ein Zinssatz bei jährlicher Anlage, $x/2$ bei halbjähriger Anlage, $x/12$ bei monatlicher, $x/365$ bei täglicher Anlage, usw. Das Kapital K_0 nach einem Jahr wäre dann

$$K_0(1+x), \quad K_0\Big(1 + \frac{x}{2}\Big)^2, \quad K_0\Big(1 + \frac{x}{12}\Big)^{12}, \quad K_0\Big(1 + \frac{x}{365}\Big)^{365}, \quad \ldots, \quad K_0 \exp(x).$$

Für bspw. $x = 1$ erhalten wir

$$2K_0; \quad 2{,}25K_0; \quad \ldots; \quad (2{,}718\ldots)K_0.$$

[1] Hier wurde ein Lemma verwendet, dass wir noch nicht bewiesen haben: Sei $A \subset \mathbb{R}$ eine nach unten beschränkte Menge und $f\colon \{\inf A\} \cup A \to \mathbb{R}$ eine monoton (schwach) fallende Abbildung. Dann gilt $\sup f(A) \leq f(\inf A)$. *Beweis*: Angenommen $\sup f(A) > f(\inf A)$. Weil das Supremum die kleinste obere Schranke ist, gibt es ein $a \in A$ mit $f(a) > f(\inf A)$. Aus der Monotonie folgt aber $a \leq \inf A$, also $a = \inf A = \min A$. Aus der Monotonie folgt aber auch $\sup f(A) = f(\min A) = f(\inf A)$, Widerspruch.

[2] Vor diesem Schritt schätzte der Dozent $\exp(-x) \cdot \exp(-\varepsilon) \geq \exp(-x)(1 - 2\varepsilon)$ ab und verwendete die Restgliedabschätzung nach dem $N = 0$-ten Glied, was $|\exp(x) - 1| \leq 2\varepsilon$ ergibt. Meiner Meinung nach ist das aber nicht nötig.

10 Komplexe Zahlen

10.1 Motivation und Definition

Eine Möglichkeit, um die *komplexen Zahlen* $\mathbb{C}$ zu konstruieren, ist über folgende *Körpererweiterung* von $\mathbb{R}$ nach $\mathbb{C}$.

Angenommen es gibt einen Körper K, der $\mathbb{R}$ enthält (aber nicht notwendigerweise alle Eigenschaften von $\mathbb{R}$ hat) und ein Element i mit $i^2 = -1$ enthält.

Dann sind aufgrund Abgeschlossenheit von K für $x, y, u, v \in \mathbb{R} \subset \mathbb{C}$ auch $(x+iy) \in K$ und $(u+iv) \in K$ mit

$$(x+iy)+(u+iv) = (x+u)+i(y+v) \qquad \text{und} \qquad (x+iy)\cdot(u+iv) = (xu-yv)+i(yu+xv).$$

Damit ist die Menge $K_0 := \{x+iy : x, y \in \mathbb{R}\}$ abgeschlossen bzgl. Addition $+$ und Multiplikation $\cdot$, es gibt $0, 1 \in K_0$ und K_0 ist abgeschlossen unter additive und multiplikative Inversenbildung. Diesen Körper K_0 bezeichnen wir mit $\mathbb{C}$.

Wir wollen es aber „von Hinten“ aufziehen und nicht die Existenz eines solchen Körpers zuerst voraussetzen (da wir es mit unserem Kenntnisstand noch nicht wissen).

Satz 10.1 (Körper der komplexen Zahlen)**.** *Die Menge* $\mathbb{R}\times\mathbb{R} = \{(x,y) : x, y \in \mathbb{R}\}$ *der geordneten Paare reeller Zahlen definiert einen* Körper *mit* Addition *und* Multiplikation*, definiert durch*

$$(x,y)+(u,v) := (x+u, y+v) \qquad \textit{und} \qquad (x,y)\cdot(u,v) := (xu-yv, yu+xv),$$

neutrale Elemente $0 := (0,0)$*,* $1 := (1,0)$ *sowie* Inverse*, definiert durch*

$$-(x,y) := (-x,-y) \qquad \textit{und} \qquad (x,y)^{-1} := \left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right) \quad \textit{falls } (x,y) \neq (0,0)$$

mit $x, y, u, v \in \mathbb{R}$.

Beweis. Der Nachweis ist wirklich nur stures nachrechnen, weshalb wir uns hier sehr kurz fassen.

- Gruppenstruktur bzgl. Addition: trivial.
- Gruppenstruktur bzgl. Multiplikation:
  - Kommutativität: $(x,y)\cdot(u,v) = (xu-yv, yu+xv) = (ux-vy, uy+vx) = (u,v)\cdot(x,y)$, wobei wir für die mittlere Gleichheit die Kommutativität in $\mathbb{R}$ nutzten.
  - Neutrales Element: $(x,y)\cdot(1,0) = \cdots = (x,y)$.
  - Inverses:
  $$(x,y)\cdot\left(\frac{x}{x^2+y^2}, \frac{-y}{x^2+y^2}\right) = \cdots = \left(\frac{x^2+y^2}{x^2+y^2}, \frac{-xy+yx}{x^2+y^2}\right) = (1,0).$$
  - Assoziativität: $(x,y)\cdot((u,v)\cdot(s,t)) = \cdots = ((x,y)\cdot(u,v))\cdot(s,t)$.
  - Distributivität: $(x,y)((u,v)+(s,t)) = \cdots = (x,y)\cdot(u,v)+(x,y)\cdot(s,t)$. $\square$

Wir beobachten: Für $(x,0)$ und $(u,0)$ mit $x, u \in \mathbb{R}$ gilt

$$(x,0)+(u,0) = (x+u,0) \qquad \text{sowie} \qquad (x,0)\cdot(u,0) = (xu,0).$$

Deshalb bildet $\{(x,y) \in \mathbb{C} : y = 0\}$ einen *Teilkörper* von $\mathbb{C}$, der *isomorph* zu $\mathbb{R}$ ist. Daher identifizieren wir $\mathbb{R}$ mit $\{(x,0) : x \in \mathbb{R}\} \subset \mathbb{C}$ und schreiben $x = (x,0) \in \mathbb{C}$ für alle $x \in \mathbb{R}$. Damit ist $\mathbb{R} \subset \mathbb{C}$.

Ferner setzen wir $i := (0,1)$, woraus $i^2 = (0,1)\cdot(0,1) = (-1,0) = -1 \in \mathbb{R}$ sowie

$$(x,y) = (x,0)+(0,y) = x(1,0)+y(0,1) = x+iy$$

für alle $x, y \in \mathbb{R}$ folgt. Damit können wir den Körper der komplexen Zahlen definieren.

Definition 10.2 (Körper der komplexen Zahlen, imaginäre Einheit, Real-/Imaginärteil)**.** Wir setzen den Körper aus Satz 10.1 vor. Sei $i := (0,1)$ mit $i^2 = -1$ die *imaginäre Einheit*. Dann lässt sich jede *komplexe Zahl* des Körpers schreiben als $z = (x,y) = x+iy$ mit *Realteil* $\operatorname{Re}(z) = x$ und *Imaginärteil* $\operatorname{Im}(z) = y$ von z.

$$\mathbb{C} := \{x+iy : x, y \in \mathbb{R}\}$$

heißt *Körper der komplexen Zahlen*.

Achtung: Der Imaginärteil $\mathrm{Im}(z)$ einer komplexen Zahl z sollte nicht mit dem Bild $\mathrm{im}(z)$ einer Abbildung z verwechselt werden.

Die komplexen Zahlen lassen sich gut in einer $\mathbb{R} \times \mathbb{R}$-Ebene, einem kartesischen Koordinatensystem visualisieren, der sog. *komplexen* bzw. GAUSSSCHEN *Zahlenebene*.

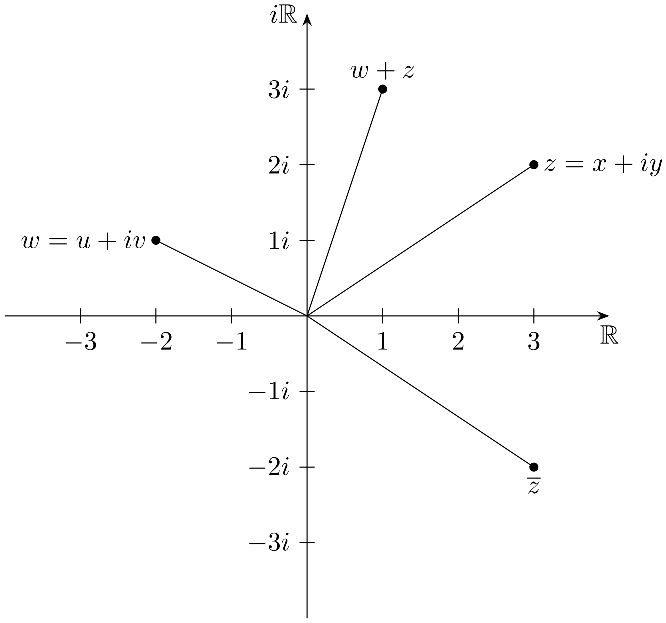


Abbildung 10.1: Komplexe Zahlenebene.

Dabei entspricht

- $\mathbb{C}$ der Ebene,
- $\mathbb{R}$ der x-Achse (also dem reellen „Zahlenstrahl“) und
- $i\mathbb{R} = \{ix : x \in \mathbb{R}\}$, die rein *imaginären Zahlen* der y-Achse.

Addition zweier komplexer Zahlen entspricht der Addition zweier Vektoren in $\mathbb{R}^2$ (indem wir einen der Pfeile an die Spitze des anderen Pfeils verschieben – die Spitze des ersten Pfeils ist die Summe). Die Multiplikation funktioniert aber doch etwas anders . . .

10.2 Komplexe Konjugation und Betrag

Definition 10.3 (komplexe Konjugation)**.** Die Abbildung $\bar{\cdot}\colon \mathbb{C} \to \mathbb{C}$ mit

$$z = (x + iy) \mapsto \overline{z} = (x - iy)$$

ist die *komplexe Konjugation.* $\overline{z}$ ist die *komplex Konjugierte* von z.

In der komplexen Zahlenebene entspricht die Konjugation der Spiegelung an der x-Achse.

Lemma 10.4 (Rechenregeln mit Konjugierten)**.** *Für die Konjugation gilt mit $z = x + iy \in \mathbb{C}$ und $w \in \mathbb{C}$:*

1. $\overline{\overline{z}} = z$.
2. $\overline{z + w} = \overline{z} + \overline{w}$.
3. $\mathrm{Re}(z) = \frac{1}{2}(z + \overline{z})$ *und* $\mathrm{Im}(z) = \frac{1}{2}(z - \overline{z})$.
4. $z\overline{z} = x^2 + y^2$, *also ist* $z\overline{z}$ *reell und* ≥ 0. *Ferner ist* $z\overline{z} = 0$ *genau dann, wenn* $z = 0$.[1]

Beweis. Das erfolgt durch einfaches nachrechnen. □

[1] Eine weitere (wie ich finde) nützliche Regel: Für $a \in \mathbb{R}$ und $z \in \mathbb{C}$ gilt $\overline{az} = a\overline{z}$.

Definition 10.5 (Betrag)**.** Der Betrag einer komplexen Zahl $z = x + iy \in \mathbb{C}$ ist

$$|z| := \sqrt{z\overline{z}} = \sqrt{x^2 + y^2}.$$

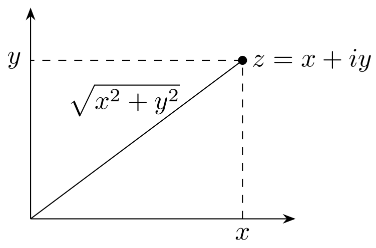


Abbildung 10.2: Geometrische Visualisierung des Betrags.

Lemma 10.6 (Rechenregeln mit Beträgen)**.** *Für Beträge gilt mit $z = x + iy \in \mathbb{C}$ und $w \in \mathbb{C}$:*

1. $|z| \geq 0$, *und* $|z| = 0$ *genau dann, wenn* $z = 0$.
2. $|z| = |\overline{z}|$.
3. $|\mathrm{Re}(z)| \leq |z|$ *und* $|\mathrm{Im}(z)| \leq |z|$.
4. $|z + w| \leq |z| + |w|$ (Dreiecksungleichung).
5. $|zw| = |z| \cdot |w|$.
6. $1/z = \overline{z}/|z|^2$ *für* $z \neq 0$.

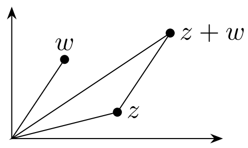


Abbildung 10.3: Geometrische Visualisierung der Dreiecksungleichung.

Beweis. Die Beweise folgen direkt aus den Definitionen. Deshalb zeigen wir hier nur eine Auswahl an Aussagen.

2. Es gilt $|z| = \sqrt{z\overline{z}} = \sqrt{\overline{z}z} = |\overline{z}|$.
3. Für $z = x + iy$ gilt $|z|^2 = \mathrm{Re}(z)^2 + \mathrm{Im}(z)^2 \geq \mathrm{Re}(z)^2$. Analog für $\mathrm{Im}(z)^2$.
4. Es gilt

$$\begin{aligned} |z+w|^2 = (z+w)(\overline{z}+\overline{w}) = z\overline{z} + (z\overline{w} + \overline{z}w) + w\overline{w} &= |z|^2 + 2\,\mathrm{Re}(z\overline{w}) + |w|^2 \\ &\leq |z|^2 + 2|z\overline{w}| + |w|^2 = |z|^2 + 2|zw| + |w|^2 = (|z| + |w|)^2. \end{aligned}$$

Anschließendes Wurzelziehen ergibt die Behauptung, wobei Beträge stets nichtnegativ sind. □

Ende der Vorlesung 14 am 29. November 2021

Bemerkung 10.7.

1. Auf $\mathbb{C}$ können wir keine Ordnung einführen, denn aus den Ordnungsaxiomen würde $z^2 \geq 0$ für alle $z \in \mathbb{C}$ folgen. Das steht aber im Widerspruch zu $i^2 + 1 = 0$ denn eine nichtnegative Zahl plus eine positive Zahl kann nicht Null ergeben.
2. Falls wir $z \geq 0$ schreiben, so meinen wir stets: „$z \in \mathbb{R}$ und $z \geq 0$“.
3. Im Allgemeinen gilt <u>nicht</u> $z^2 = |z|^2$ (z. B. $i^2 = -1 \neq 1 = |i|^2$).

10.3 Folgen in $\mathbb{C}$

Definition 10.8 (komplexe Folge)**.** Sei $(c_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{C}$.

1. Die Folge *konvergiert gegen* $c \in \mathbb{C}$, falls es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, sodass $|c_n - c| \leq \varepsilon$ für alle $n \geq N$ gilt. In Zeichen:
$$\forall \varepsilon \geq 0\colon \exists N \in \mathbb{N}\colon \forall n \geq N\colon |c_n - c| \leq \varepsilon.$$
2. Die Folge ist eine *Cauchy-Folge*, falls es für jedes $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, sodass $|c_n - c_m| \leq \varepsilon$ für alle $n, m \geq N$ gilt. In Zeichen:
$$\forall \varepsilon \geq 0\colon \exists N \in \mathbb{N}\colon \forall n, m \geq N\colon |c_n - c_m| \leq \varepsilon.$$

Die Definition für komplexe Folgen sind *identisch* zu den Definition für reelle Folgen, wobei wir hier beachten sollten, das eine etwas andere Betragsdefinition gilt (vgl. Definitionen 4.4 und 5.1).

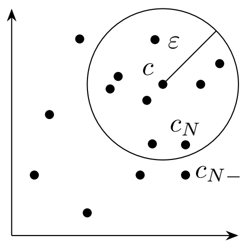


Abbildung 10.4: Visualisierung des Konvergenzbegriffs in der komplexen Zahlenebene.

10.3.1 Cauchy-Folgen

Satz 10.9 (Konvergenz in $\mathbb{C}$, komplexe Cauchy-Folgen)**.** *Sei $(c_n)_n$ eine komplexe Zahlenfolge mit $c_n = a_n + ib_n$, d. h. $a_n = \mathrm{Re}(c_n)$, $b_n = \mathrm{Im}(c_n)$.*

1. *$(c_n)_n$ konvergiert genau dann gegen $c = a + ib$, wenn $(a_n)_n$ gegen $a = \mathrm{Re}(c) \in \mathbb{R}$ und $(b_n)_n$ gegen $b = \mathrm{Im}(c) \in \mathbb{R}$ konvergiert. (Eine komplexe Folge konvergiert genau dann, wenn die Folgen der Real- und Imaginärteile konvergieren.)*
2. *$(c_n)_n$ ist genau dann eine Cauchy-Folge in $\mathbb{C}$, wenn $(a_n)_n$ und $(b_n)_n$ Cauchy-Folgen sind.*

D. h. wir können sehr leicht mit komplexen Zahlenfolgen umgehen, wenn wir die Real- und Imaginärteile getrennt betrachten.

Beweis.

1. Sei $c_n \to c$ in $\mathbb{C}$, d. h. dass es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, sodass $|c_n - c| \leq \varepsilon$ für alle $n \geq N$ gilt. Aus Lemma 10.6 folgt für alle $n \geq N$ dann
$$|a_n - a| = |\mathrm{Re}(c_n - c)| \leq |c_n - c| \leq \varepsilon.$$
Damit konvergiert $(a_n)_n$ gegen $a \in \mathbb{R}$. Analog gilt für alle $n \geq N$
$$|b_n - b| = |\mathrm{Im}(c_n - c)| \leq |c_n - c| \leq \varepsilon,$$
also konvergiert auch $(b_n)_n$ gegen $b \in \mathbb{R}$.
Für die Hinrichtung seien $a_n \to a$ und $b_n \to b$. Für jedes $\varepsilon > 0$ gibt es also ein $N_a \in \mathbb{N}$, sodass $|a_n - a| \leq \varepsilon$ für alle $n \geq N_a$ gilt, und ein $N_b \in \mathbb{N}$, sodas $|b_n - b| \leq \varepsilon$ für alle $n \geq N_b$ gilt. Setzen wir nun $N := \max\{N_a, N_b\}$, dann gilt mit der Betragsdefinition für alle $n \geq N$
$$|c_n - c|^2 = |a_n - a|^2 + |b_n - b|^2 \leq 2\varepsilon^2 \implies |c_n - c| \leq \varepsilon\sqrt{2}.$$
Damit konvergiert $(c_n)_n$ gegen $c \in \mathbb{C}$.
2. Der Beweis erfolgt identisch, indem wir c durch c_m an jeder Stelle des Beweises ersetzen. $\square$

Korollar 10.10 (Vollständigkeit von $\mathbb{C}$)**.** *$\mathbb{C}$ ist vollständig, denn jede Cauchy-Folge in $\mathbb{C}$ konvergiert.*

Beweis. Da die Cauchy-Folgen $(\mathrm{Re}(c_n))_n$ und $(\mathrm{Im}(c_n))_n$ in $\mathbb{R}$ konvergieren, konvergiert auch die Cauchy-Folge $(c_n)_n$ mit $c_n = \mathrm{Re}(c_n) + i\,\mathrm{Im}(c_n)$ in $\mathbb{C}$. $\square$

10.3.2 Grenzwertsätze

Korollar 10.11 (Grenzwertsätze in $\mathbb{C}$)**.** *Für zwei komplexe Folgen $c_n \to c$ und $d_n \to d$ gilt*

$$\lim_{n\to\infty}(c_n + d_n) = \lim_{n\to\infty} c_n + \lim_{n\to\infty} d_n = c + d \qquad \text{und} \qquad \lim_{n\to\infty}(c_n d_n) = \Big(\lim_{n\to\infty} c_n\Big)\Big(\lim_{n\to\infty} d_n\Big) = cd,$$

und, falls $d_n \neq 0$ und $d \neq 0$ für alle $n \geq N$ mit hinreichend großem N,

$$\lim_{n\to\infty}\frac{c_n}{d_n} = \frac{\lim_{n\to\infty} c_n}{\lim_{n\to\infty} d_n} = \frac{c}{d}.$$

Beweis. Die Summe und das Produkt der komplexen Folgen lassen sich umschreiben zu

$$c_n + d_n = (\operatorname{Re}(c_n) + \operatorname{Re}(d_n)) + i(\operatorname{Im}(c_n) + \operatorname{Im}(d_n))$$

und

$$c_n d_n = (\operatorname{Re}(c_n)\operatorname{Re}(d_n) - \operatorname{Im}(c_n)\operatorname{Im}(d_n)) + i(\operatorname{Re}(c_n)\operatorname{Im}(d_n) + \operatorname{Im}(c_n)\operatorname{Re}(d_n)).$$

Damit konnten wir die Konvergenz der neuen Folgen auf die Konvergenz reeller Folgen zurückführen.

Für den Quotienten gilt

$$\frac{c_n}{d_n} = \frac{c_n\overline{d_n}}{|d_n|^2},$$

woraus auch wieder die Konvergenz der komplexen Folge folgt, da die Real- und Imaginärteile wieder aus Summe, Produkte und Quotienten konvergenter reeller Folgen bestehen. □

Korollar 10.12 (Konjugation einer konvergenten Folge)**.** *Mit $(c_n)_n$ ist auch $(\overline{c_n})_n$ eine komplexe Cauchy-Folge und aus $c_n \to c$ folgt $\overline{c_n} \to \overline{c}$. Oder mit anderen Worten*

$$\lim_{n\to\infty}\overline{c_n} = \overline{\lim_{n\to\infty} c_n}.$$

Beweis. Es gilt $\overline{c_n} = \operatorname{Re}(c_n) - i\operatorname{Im}(c_n)$ und $\overline{c} = \operatorname{Re}(c) - i\operatorname{Im}(c)$. □

Korollar 10.13 (Betrag einer konvergenten Folge)**.** *Gilt für eine komplexe Folge $c_n \to c$, so gilt auch $|c_n| \to |c|$.*

Die Umkehrung gilt hingegen (wie in den reellen Zahlen) nicht.

Beweis. Es gilt $|c_n| = \sqrt{\operatorname{Re}(c_n)^2 + \operatorname{Im}(c_n)^2}$ und $|c| = \sqrt{\operatorname{Re}(c)^2 + \operatorname{Im}(c)^2}$. Damit ist $|c_n|$ eine reelle Folge und wir können die Grenzwertsätze in $\mathbb{R}$ anwenden. □

10.4 Reihen in $\mathbb{C}$

Definition 10.14 (komplexe Reihe, Konvergenz, absolute Konvergenz)**.** Die *komplexe Reihe* $\sum_{n=1}^{\infty} c_n$ mit $c_n \in \mathbb{C}$ heißt *konvergent*, falls die Folge $(s_k)_{k\in\mathbb{N}}$ der Partialsummen $s_k = \sum_{n=1}^{k} c_n$ in $\mathbb{C}$ konvergiert.

Die Reihe $\sum_{n=1}^{\infty} c_n$ heißt *absolut konvergent*, falls die Reihe $\sum_{n=1}^{\infty}|c_n|$ in $\mathbb{R}$ konvergiert (bzw. die Folge $t_k = \sum_{n=1}^{k}|c_n|$ in $\mathbb{R}$ konvergiert.)

Proposition 10.15. *Sei $\sum_n c_n$ eine komplexe Reihe.*

1. *Eine absolut konvergente Reihe in $\mathbb{C}$ ist auch konvergent in $\mathbb{C}$.*
2. Majorantenkriterium*: Gilt stets $|c_n| \leq a_n \in \mathbb{R}_+$ für hinreichend große n, und konvergiert $\sum_n a_n$ in $\mathbb{R}_+$, so ist auch $\sum_n c_n$ absolut konvergent.*
3. Quotientenkriterium*: Gilt*
$$\limsup_{n\to\infty}\left|\frac{c_{n+1}}{c_n}\right| < 1,$$
dann konvergiert die Reihe absolut.
4. Wurzelkriterium*: Gilt*
$$\limsup_{n\to\infty}|c_n|^{1/n} < 1,$$
dann konvergiert die Reihe absolut.

Beweis. Die Beweise sind analog zu den Sätzen in $\mathbb{R}$ mit der Eigenschaft, dass $|\mathrm{Re}(z)|, |\mathrm{Im}(z)| \leq |z|$ ist. □

Bemerkung 10.16.

1. Alle obigen Sätze gelten, da wir die Aussagen über komplexe Reihen durch Beträge auf Aussagen in den reellen Zahlen reduzierten.

2. Es gibt kein *Leibniz-Kriterium*, da der Begriff einer alternierenden Reihe in $\mathbb{C}$ nicht sinnvoll ist.

10.5 Exponenzialreihe in $\mathbb{C}$

Satz 10.17. *Für alle $z \in \mathbb{C}$ ist die Exponenzialreihe*

$$\exp(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

absolut konvergent.

Beweis (analog zu $\mathbb{R}$). Es gilt nach dem Quotientenkriterium

$$\limsup_{n\to\infty} \left| \frac{c_{n+1}}{c_n} \right| = \limsup_{n\to\infty} \frac{|z|}{n+1} = 0 < 1.$$ □

Proposition 10.18.

1. Restgliedabschätzung*: Für alle $N \in \mathbb{N}$ und für $z \in \mathbb{C}$ mit $|z| \leq 1 + \frac{1}{2}|N|$ gilt*

$$\left| \exp(z) - \sum_{n=0}^{N} \frac{z^n}{n!} \right| \leq 2 \frac{|z|^{N+1}}{(N+1)!}.$$

2. Funktionalgleichung*: Für alle $z, w \in \mathbb{C}$ gilt*

$$\exp(z) \cdot \exp(w) = \exp(z + w).$$

3. Für alle $z \in \mathbb{C}$ gelten

$$\exp(z) \neq 0 \qquad \textit{und} \qquad \exp(-z) = \frac{1}{\exp(z)}.$$

4. Für alle $z \in \mathbb{C}$ gilt

$$\exp(\overline{z}) = \overline{\exp(z)}.$$

Beweis. Aussagen Punkte 1 bis 3 sind sehr ähnlich zum reellen Pendant und werden analog bewiesen.
Zu Punkt 4: Für alle $n \in \mathbb{N}$ gilt für die n-te Partialsumme

$$\sum_{k=0}^{n} \frac{(\overline{z})^k}{k!} = \sum_{k=0}^{n} \overline{\left(\frac{z^k}{k!} \right)} = \overline{\sum_{k=0}^{n} \frac{z^k}{k!}}.$$

Für $n \to \infty$ folgt aus Korollar 10.12 dann $\exp(\overline{z}) = \overline{\exp(z)}$. □

10.6 Trigonometrische und hyperbolische Funktionen

Definition 10.19 (Trigonometrische und hyperbolische Funktionen)**.** Für alle $z \in \mathbb{C}$ definieren wir

- $\cosh(z) = \frac{1}{2}(\exp(z) + \exp(-z))$ (*Cosinus hyperbolicus*),
- $\sinh(z) = \frac{1}{2}(\exp(z) - \exp(-z))$ (*Sinus hyperbolicus*),
- $\cos(z) = \cosh(iz) = \frac{1}{2}(\exp(iz) + \exp(-iz))$ (*Cosinus*),
- $\sin(z) = \frac{1}{i}\sinh(iz) = \frac{1}{2i}(\exp(iz) - \exp(-iz))$ (*Sinus*),

sowie, falls der Nenner jeweils nicht Null ist,

$$\tanh(z) = \frac{\sinh(z)}{\cosh(z)}, \qquad \coth(z) = \frac{\cosh(z)}{\sinh(z)}, \qquad \tan(z) = \frac{\sin(z)}{\cos(z)}, \qquad \cot(z) = \frac{\cos(z)}{\sin(z)}$$

(in der Reihenfolge *Tangens hyperbolicus*, *Cotangens hyperbolicus*, *Tangens*, *Cotangens*).[2]

Auch tanh eignet sich wie $x \mapsto x/(1+|x|)$ als Abbildung, die $\overline{\mathbb{R}}$ isomorph auf $[-1,1]$ abbildet.

10.6.1 Eigenschaften und Identitäten

Satz 10.20 (Reihendarstellung)**.**

- *Es gelten*
$$\cosh(z) = \sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!} \qquad \textit{und} \qquad \sinh(z) = \sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!}.$$
- *Es gelten*
$$\cos(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!} = 1 - \frac{z^2}{2} + \frac{z^4}{24} - \frac{z^6}{6!} + \frac{z^8}{8!} - \cdots$$
und
$$\sin(z) = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} = z - \frac{z^3}{6} + \frac{z^5}{5!} - \frac{z^7}{7!} + \cdots$$

Beweis.

1. Aus den Definitionen folgen
$$\cosh(z) = \frac{1}{2}\left(\sum_{n=0}^{\infty} \frac{z^n}{n!} + \sum_{n=0}^{\infty} \frac{(-z)^n}{n!}\right) = \sum_{\substack{n=0 \\ n \text{ gerade}}}^{\infty} \frac{z^n}{n!} = \sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!}$$
und
$$\sinh(z) = \frac{1}{2}\left(\sum_{n=0}^{\infty} \frac{z^n}{n!} - \sum_{n=0}^{\infty} \frac{(-z)^n}{n!}\right) = \sum_{\substack{n=0 \\ n \text{ ungerade}}}^{\infty} \frac{z^n}{n!} = \sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!}.$$
2. Aus dem eben Gezeigtem folgen
$$\cos(z) = \sum_{n=0}^{\infty} \frac{(iz)^{2n}}{(2n)!} = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n}}{(2n)!}$$
und
$$\sin(z) = \sum_{n=0}^{\infty} \frac{i^{2n+1}}{i} \frac{z^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!}.$$
□

[2] cosh, sinh, cos und sin lassen sich wie folgt geometrisch visualisieren: Zeichnen wir den Einheitskreis mit $x^2 + y^2 = 1$, so liegt jeder Punkt $(\cos\theta, \sin\theta)$ auf dem Kreis, wobei der Radius mit der x-Achse den Winkel θ einschließt. Zeichnen wir die Einheitshyperbel mit $x^2 - y^2 = 1$, so liegt jeder Punkt $(\cosh A, \sinh A)$ auf der Hyperbel, wobei die Strecke vom Ursprung zu dem Punkt mit der x-Achse und der Hyperbel die Fläche $\frac{1}{2}A$ einschließt.

Proposition 10.21.

1. cos *und* cosh *sind gerade Funktionen, wohingegen* sin *und* sinh *ungerade Funktionen sind. D. h. es gelten*

$$\cos(-z) = \cos(z), \qquad \cosh(-z) = \cosh(z),$$
$$\sin(-z) = -\sin(z), \qquad \sinh(-z) = -\sinh(z).$$

2. *Für alle* $z \in \mathbb{C}$ *gelten*

$$\exp(z) = \cosh(z) + \sinh(z) \qquad \textit{und} \qquad \exp(iz) = \cos(z) + i\sin(z).$$

Die letzte Gleichung heißt auch EULERSCHE Formel.

3. *Für alle* $z \in \mathbb{C}$ *gelten*

$$\cosh^2(z) - \sinh^2(z) = 1 \qquad \textit{und} \qquad \cos^2(z) + \sin^2(z) = 1.$$

Die letzte Gleichung heißt trigonometrischer PYTHAGORAS *(in Anlehnung an den Satz des* PYTHAGORAS*).*

Bezeichnung 10.22. Es ist Konvention, für trigonometrische und hyperbolische Funktionen die Potenzen an die Funktion zuschreiben, z. B. $\sin^2(z)$ anstatt von $(\sin(z))^2$. Genauso lässt man oft Klammern weg, falls der Sinn nicht verfälscht wird, also z. B. $\sin^2 z = \sin^2(z)$.

Beweis.

1. Einsetzen in die Definition liefert die Behauptung.
2. Anwenden der Definition liefert die Behauptung.
3. Aus $a \coloneqq \exp(z)$ und $b \coloneqq \exp(-z)$ folgt

$$1 = \exp(z) \cdot \exp(-z) = ab = \left(\frac{a+b}{2}\right)^2 - \left(\frac{a-b}{2}\right)^2 = \cosh^2 z - \sinh^2 z.$$

Die zweite Gleichung folgt durch Einsetzen von iz statt z. □

Ende der Vorlesung 15 am 06. Dezember 2021

Satz 10.23 (Additionstheoreme[3])**.**

1. *Für alle* $z, w \in \mathbb{C}$ *gelten*

$$\cosh(z \pm w) = \cosh z \cosh w \pm \sinh z \sinh w,$$
$$\sinh(z \pm w) = \sinh z \cosh w \pm \cosh z \sinh w.$$

2. *Für alle* $z, w \in \mathbb{C}$ *gelten*

$$\cos(z \pm w) = \cos z \cos w \mp \sin z \sin w,$$
$$\sin(z \pm w) = \sin z \cos w \pm \cos z \sin w.$$

Beweis.

[3] Ich habe mir die Freiheit genommen, hier auch Subtrationstheoreme hinzuzufügen, die sich aus den Additionstheorem unter Beachtung der Parität, d. h. ob die Funktionen (un-)gerade sind, ergeben.

1. Aus den Definitionen und der Funktionalgleichung von exp folgt

$$\begin{aligned}
2\cosh(z+w) &= \exp(z+w) + \exp(-z-w) = \exp(z)\exp(w) + \exp(-z)\exp(-w) \\
&= \frac{\big(\exp(z)+\exp(-z)\big)\big(\exp(w)+\exp(-w)\big)}{2} + \frac{\big(\exp(z)-\exp(-z)\big)\big(\exp(w)-\exp(-w)\big)}{2} \\
&= 2\cosh z\cosh w + 2\sinh z\sinh w.
\end{aligned}$$

Analog gilt

$$\begin{aligned}
2\sinh(z+w) &= \exp(z+w) - \exp(-z-w) = \exp(z)\exp(w) - \exp(-z)\exp(-w) \\
&= \frac{\big(\exp(z)-\exp(-z)\big)\big(\exp(w)+\exp(-w)\big)}{2} + \frac{\big(\exp(z)+\exp(-z)\big)\big(\exp(w)-\exp(-w)\big)}{2} \\
&= 2\sinh z\cosh w + 2\cosh z\sinh w.
\end{aligned}$$

2. Die Aussagen folgen aus dem gerade Gezeigtem, indem wir iz für z einsetzen. □

Korollar 10.24 (Doppelwinkelfunktionen)**.**

1. Für alle $z \in \mathbb{C}$ gelten

$$\begin{aligned}
\cosh(2z) &= \cosh^2 z + \sinh^2 z = 2\cosh^2 z - 1 = 2\sinh^2 z + 1, \\
\sinh(2z) &= 2\cosh z\sinh z.
\end{aligned}$$

2. Für alle $z \in \mathbb{C}$ gelten

$$\begin{aligned}
\cos(2z) &= \cos^2 z - \sin^2 z = 1 - 2\sin^2 z = 2\cos^2 z - 1, \\
\sin(2z) &= 2\sin z\cos z.
\end{aligned}$$

3. Für alle $z, w \in \mathbb{C}$ gelten

$$\begin{aligned}
\sin z - \sin w &= 2\cos\frac{z+w}{2}\sin\frac{z-w}{2}, \\
\cos z - \cos w &= -2\sin\frac{z+w}{2}\sin\frac{z-w}{2}.
\end{aligned}$$

Beweis.

1. Das folgt, wenn wir $w = z$ setzen.
2. Das folgt, wenn wir $w = z$ setzen.
3. Das folgt, wenn wir Punkt 2 der Additionstheoreme auf $\frac{1}{2}(z+w)$ und $\frac{1}{2}(z-w)$ anwenden:

$$\begin{aligned}
\sin z &= \sin\left(\frac{z+w}{2} + \frac{z-w}{2}\right) = \sin\frac{z+w}{2}\cos\frac{z-w}{2} + \cos\frac{z+w}{2}\sin\frac{z-w}{2}, \\
\sin w &= \sin\left(\frac{z+w}{2} - \frac{z-w}{2}\right) = \sin\frac{z+w}{2}\cos\frac{z-w}{2} - \cos\frac{z+w}{2}\sin\frac{z-w}{2}, \\
\sin z - \sin w &= 2\sin\frac{z-w}{2}\cos\frac{z+w}{2}.
\end{aligned}$$ □

10.7 Trigonometrische und hyperbolische Funktionen für reelle Argumente

Proposition 10.25. *Sei $x \in \mathbb{R}$.*

1. $\cosh x$, $\sinh x$, $\cos x$ und $\sin x$ sind reellwertig. Es gelten

$$\cos x = \mathrm{Re}(e^{ix}) \qquad \textit{und} \qquad \sin x = \mathrm{Im}(e^{ix}).$$

2. *Es gilt*
$$\cos^2 x + \sin^2 x = 1.$$
Damit liegen die Punkte $(\cos x, \sin x)$ *auf dem Einheitskreis in* $\mathbb{C} \cong \mathbb{R}^2$.[4] *Dabei ist* x *die Bogenlänge auf der Einheitskreislinie von* $1 = (1,0)$ *nach* $e^{ix} = (\cos x, \sin x)$ *(Übungsaufgabe).*

3. *Für alle* $x \in \mathbb{R}$ *gelten*
$$|\cos x| \leq 1, \qquad |\sin x| \leq 1 \qquad \textit{und} \qquad \cosh x \geq 1.$$
Für alle $x \in [-2,2]$ *gelten die Abschätzungen*
$$\cos x \leq 1, \qquad \cos x \geq 1 - \frac{x^2}{2}, \qquad \cos x \leq 1 - \frac{x^2}{2} + \frac{x^4}{24}.$$
Für alle $x \in [0,2]$ *gelten die Abschätzungen*
$$\sin x \leq x, \qquad \sin x \geq x - \frac{x^3}{6}, \qquad \sin x \leq x - \frac{x^3}{6} + \frac{x^5}{120}.$$

Beweis. Punkte 1 und 2 folgen als Spezialfall aus $\mathbb{C}$.

Zu Punkt 3: Die Reihendarstellung von cos und sin sind alternierend. In jeder Reihe ist die Folge der Beträge ihrer Summanden monoton fallend, falls
$$\frac{x^{n+2}}{(n+2)!} \leq \frac{x^n}{n!} \iff \frac{x^2}{(n+1)(n+2)} \text{für } x \in [0,2],$$
was für alle $n \geq 1$ der Fall ist. Das bedeutet, dass im Restglied immer kleinere Glieder abwechselnd addiert und subtrahiert werden, d. h. das weggelassene Restglied ist entweder nichtnegativ oder nichtpositiv. ☐

10.8 Definition von π

Definition 10.26 (Kreiszahl)**.** $\frac{1}{2}\pi$ ist die eindeutige Nullstelle von $x \mapsto \cos x$ im Intervall $[0,2]$. Dabei ist $\pi \approx 3{,}1415\ldots$

Proposition 10.27. *Die Nullstelle* $\frac{1}{2}\pi$ *existiert und ist eindeutig.*

Beweis. Nächste Woche mittels Zwischenwertsatz. ☐

Proposition 10.28. *Es gelten*

1. $\cos(\frac{1}{2}\pi) = 0$ *und* $\sin(\frac{1}{2}\pi) = 1$*; sowie*
2. $\exp(\frac{1}{2}\pi i) = i$, $\exp(\pi i) = -1$, $\exp(\frac{3}{2}\pi i) = -i$ *und* $\exp(2\pi i) = 1$.

Beweis.

1. Aus dem trigonometrischen Pythagoras folgt $\sin(\frac{1}{2}\pi) = \pm 1$. Ferner ist sin nichtnegativ auf dem Intervall $[0,2]$, also ist $\sin(\frac{1}{2}\pi) = +1$.
2. Aus der eulerschen Formel folgt $\exp(\frac{1}{2}\pi i) = \cos(\frac{1}{2}\pi) + i\sin(\frac{1}{2}\pi)$. Die restlichen Werte können wir durch Potenzieren und der Funktionalgleichung von exp ermitteln. ☐

x	$\exp(ix)$	$\cos x$	$\sin x$
$\frac{1}{2}\pi$	i	0	1
π	-1	-1	0
$\frac{3}{2}\pi$	$-i$	0	-1
2π	1	1	0

Satz 10.29. *Für alle* $z \in \mathbb{C}$ *gelten*

[4] Hier muss spezifiziert werden, was mit dem Isomorphismus gemeint ist. Die beiden Räume sind maximal in der Kategorie der metrischen Räume isomorph zueinander. Aber $\mathbb{R}^2$ ist im Gegensatz zu $\mathbb{C}$ kein Körper.

1. $\exp(z + i\frac{1}{2}\pi) = i\exp(z)$,
2. $\exp(z + i\pi) = -\exp(z)$ *und*
3. $\exp(z + 2\pi i) = \exp(z)$.

Damit ist die Exponenzialfunktion periodisch *in* $\mathbb{C}$ *mit Periode* $2\pi i$.

Beweis. Die Gleichungen folgen aus der Funktionalgleichung und der eben gezeigten Proposition 10.28. □

Korollar 10.30 (Phasenverschiebung)**.** *Für alle* $z \in \mathbb{C}$ *gelten*

1. $\cos(z + \frac{1}{2}\pi) = -\sin z$ *und* $\sin(z + \frac{1}{2}\pi) = -\cos z$,
2. $\cos(z + \pi) = -\cos z$ *und* $\sin(z + \pi) = -\sin z$ *sowie*
3. $\cos(z + 2\pi) = \cos z$ *und* $\sin(z + 2\pi) = \sin z$.

Damit sind die trigonometrischen Funktionen periodisch *in* $\mathbb{C}$ *mit Periode* 2π.

Achtung: Für exp ist die Periode $2\pi i$ imaginär, während für sin und cos die Periode 2π reell.

Beweis.

1. Einsetzen in die Definition der Funktionen und Anwenden der Funktionalgleichung liefern die Identitäten.
2. Das folgt aus zweifachem Anwenden von Punkt 1, d. h. $\cos(z + \pi) = \cos((z + \frac{1}{2}\pi) + \frac{1}{2}\pi)$ etc.
3. Das folgt aus zweifachem Anwenden von Punkt 2, d. h. $\cos(z + 2\pi) = \cos((z + \pi) + \pi)$ etc. □

Satz 10.31 (Nullstellen in $\mathbb{R}$)**.** *Die Menge aller Nullstellen von* cos *in* $\mathbb{R}$ *ist*

$$\{x \in \mathbb{R} : \cos x = 0\} = \left\{\left(k + \frac{1}{2}\right)\pi : k \in \mathbb{Z}\right\}$$

und von sin *ist*

$$\{x \in \mathbb{R} : \sin x = 0\} = \{k\pi : k \in \mathbb{Z}\}.$$

Beweis. Aus der Definition von $\frac{1}{2}\pi$ und dem Fakt, dass cos ungerade ist, folgt, dass $\frac{1}{2}\pi$ die einzige Nullstelle von cos im Intervall $(-\frac{1}{2}\pi, \frac{1}{2}\pi]$ ist. Wegen $\cos(x + \pi) = -\cos x$ (Phasenverschiebung) sind $\frac{1}{2}\pi$ und $\frac{1}{2}\pi + \pi$ die einzigen Nullstellen im Intervall $(-\frac{1}{2}\pi, \frac{3}{2}\pi]$. Da die Länge dieses Intervalls genau die Periode von cos entspricht, ist die Menge aller Nullstellen $\{\frac{1}{2}\pi, \frac{3}{2}\pi\} \cdot 2\pi\mathbb{Z} = \{(k + \frac{1}{2})\pi : k \in \mathbb{Z}\}$.

Für die Nullstellen von sin verwenden wir $\sin z = -\cos(z + \frac{1}{2}\pi)$ (Phasenverschiebung). Damit erhalten wir $\{k\pi : k \in \mathbb{Z}\}$ für die Menge aller Nullstellen. □

Korollar 10.32. *Es ist genau dann* $\exp(z) = 1$, *wenn* $z = 2k\pi i$ *für ein* $k \in \mathbb{Z}$.

Beweis. Zur Rückrichtung: Ist $z = 2k\pi i$ für ein $k \in \mathbb{Z}$, so folgt mit Proposition 10.28

$$\exp(z) = \exp(2k\pi i) = (\exp(2\pi i))^k = 1^k = 1.$$

Zur Hinrichtung: Sei $z = x + iy \in \mathbb{C}$ mit $x, y \in \mathbb{R}$ und $\exp(z) = 1$. Dann folgt mit der eulerschen Formel

$$1 = \exp(z) = \exp(x + iy) = \exp(x)\exp(iy) = \exp(x)(\cos y + i\sin y).$$

Nehmen wir nun den Betrag, erhalten wir

$$1 = |\exp(x)| \cdot |\cos^2 y + \sin^2 y| = \exp(x).$$

Dabei ist $x = 0$ die einzige Stelle mit $\exp(x) = 1$, denn $x \mapsto \exp(x)$ ist streng monoton. Betrachten wir ferner Imaginär- und Realteil, erhalten wir

$$0 = \mathrm{Im}(1) = \mathrm{Im}(\cos y + i\sin y) = \sin y,$$

und

$$1 = \mathrm{Re}(1) = \mathrm{Re}(\cos y + i\sin y) = \cos y.$$

Die Nullstellen von sin sind $y = k\pi$ mit $k \in \mathbb{Z}$. Aus $\cos y = 1$ und der Periodizität von cos folgt, dass k gerade sein muss. □

Korollar 10.33 (Nullstellen in $\mathbb{C}$)**.** *Die Menge aller Nullstellen von* cos *in* $\mathbb{C}$ *ist*

$$\{x \in \mathbb{C} : \cos x = 0\} = \left\{ \left(k + \frac{1}{2}\right)\pi : k \in \mathbb{Z} \right\}$$

und von sin *ist*

$$\{x \in \mathbb{C} : \sin x = 0\} = \{k\pi : k \in \mathbb{Z}\}.$$

sin und cos haben also dieselben Nullstellen wie in $\mathbb{R}$.

Beweis. Sei $z \in \mathbb{C}$ mit $\sin z = 0$. Aus der Definition des Sinus folgt

$$0 = \sin z = \frac{\exp(iz) - \exp(-iz)}{2i} \iff \exp(-iz) = \exp(iz).$$

Nun können wir durch $\exp(-iz)$ dividieren, da die Exponenzialfunktion nie den Wert Null annimmt (Proposition 10.18):

$$1 = \frac{\exp(iz)}{\exp(-iz)} = \exp(2iz).$$

Nun folgt aus Korollar 10.32, dass das genau dann so ist, wenn für ein $k \in \mathbb{Z}$

$$2iz = 2k\pi i \iff z = k\pi.$$

Das sind die Nullstellen von sin. Durch Phasenverschiebung ergeben sich daraus die Nullstellen von cos.[5] □

10.9 Polardarstellung

Proposition 10.34 (Eindeutigkeit der Polardarstellung)**.** *Für jedes* $z \in \mathbb{C} \setminus \{0\}$ *gibt es eindeutig bestimmte* $r \in (0, \infty)$ *und* $\varphi \in [0, 2\pi)$*, sodass gilt:*

$$z = r \exp(i\varphi) = r(\cos\varphi + i \sin\varphi) = r(\cos\varphi, \sin\varphi).$$

Bezeichnung 10.35 (Betrag, Argument)**.** $r = |z|$ ist der *Betrag* und $\varphi =: \arg(z)$ ist das *Argument* von $z = r\exp(i\varphi)$.

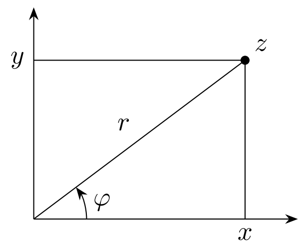


Abbildung 10.5: Kartesische Darstellung und Polardarstellung $z = x + iy = (r\cos\varphi, r\sin\varphi)$.

Diese Darstellung existiert natürlich auch für $z = 0$, wobei aber ϕ nicht mehr eindeutig ist.

Beweis. Nächste Woche mit Zwischenwertsatz. □

Bezeichnung 10.36 (Polardarstellung)**.** Das Paar

$$(r, \varphi) \in (0, \infty) \times [0, 2\pi)$$

heißt *Polardarstellung* von $z = r\exp(i\varphi)$. Im erweiterten Sinne ist jedes

$$(r, \varphi) \in [0, \infty) \times \mathbb{R}$$

auch eine Darstellung von z, wobei r eindeutig ist, sich die φ aber um Vielfache von 2π unterscheiden.

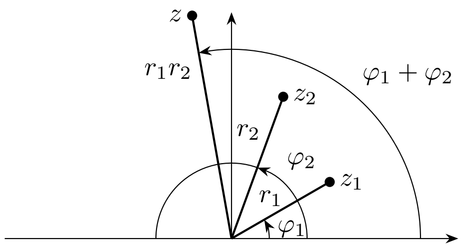
[5] Der Beweis stammt von mir. In der Vorlesung gab der Dozent keinen Beweis.

Korollar 10.37 (Multiplikation in Polardarstellung). *Für alle $z_1 = r_1 \exp(i\varphi_1) \in \mathbb{C}$ und $z_2 = r_2 \exp(i\varphi_2) \in \mathbb{C}$ gilt*

$$z_1 z_2 = r \exp(i\varphi) \quad \textit{mit } r = r_1 r_2 \textit{ und } \varphi = \varphi_1 + \varphi_2$$

(Produkt der Beträge und Summe der Argumente).[6]

Beweis. Das folgt aus der Tatsache, dass $\exp(w)\exp(w') = \exp(w + w')$ für alle $w, w' \in \mathbb{C}$. ☐



Bemerkung 10.38. Was wir uns leicht merken können: Für $x, y \in \mathbb{R}$ gilt

$$\exp(x + iy) = \exp(x)(\cos y + i \sin y).$$

Dabei ist $\exp(x)$ der Betrag und y das Argument der komplexen Zahl.

10.10 Einheitswurzeln

Satz 10.39 (Einheitswurzeln). *Sei $n \in \mathbb{N}$. Die Gleichung $z^n = 1$ hat genau n verschiedene Lösungen in $\mathbb{C}$, nämlich*

$$\xi_k = \exp\left(\frac{k}{n} 2\pi i\right) \quad \textit{für } k = 0, 1, \dots, n-1$$

und werden als n-te Einheitswurzeln *bezeichnet.*

Beweis. Jedes der ξ_k sind tatsächlich Lösungen, denn für alle k gilt

$$(\xi_k)^n = \exp(k2\pi i) = (\exp(2\pi i))^k = 1^k = 1.$$

Um zu zeigen, dass es auch die einzigen Lösungen sind, betrachten wir allgemein $z^n = 1$ mit $z \in \mathbb{C}$. Daraus folgt $|z| = 1$ und folglich $z = \exp(i\varphi)$. Weil $z^n = 1 = \exp(in\varphi)$ gelten muss, folgt aus Korollar 10.32, dass $n\varphi = 2k\pi i \iff \varphi = (n/k)2\pi i$ für alle $k \in \mathbb{Z}$ gelten muss. Daraus folgen die Einheitswurzeln $\xi_k = z = \exp(i(k/n)2\pi)$. Wegen der Periodizität von exp können wir uns auf $0 \le i(k/n)2\pi < 2\pi i \iff 0 \le k < n$ beschränken. ☐

Verbinden wir „benachbarte" Einheitswurzeln, erhalten wir ein regelmäßiges n-Eck, wobei eine Ecke bei $(1, 0) = 1$ liegt.

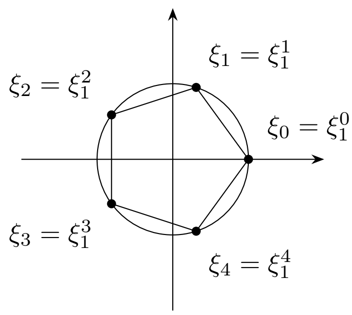


Abbildung 10.6: Einheitswurzeln für $n = 5$.

Ende der Vorlesung 16 am 08. Dezember 2021

[6] Das gilt natürlich auch für die Division (Quotient der Beträge und Differenz der Argumente).

11 Stetige Funktionen

11.1 Funktionen

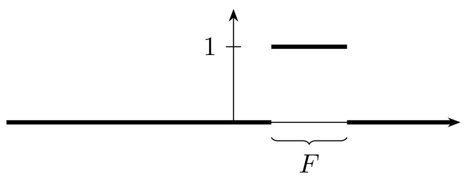
Definition 11.1 (Funktion in $\mathbb{R}$)**.** Eine *reellwertige Funktion* auf einer Teilmenge $D \subset \mathbb{R}$ ist eine Abbildung $f\colon D \to \mathbb{R}$, also die jedem $x \in D$ ein $f(x) \in \mathbb{R}$ zuordnet. D heißt *Definitionsbereich* von f, und $f(D) \coloneqq \{f(x)\colon x \in D\} \subset \mathbb{R}$ heißt *Wertebereich* von f.

Später werden wir auch komplexwertige Funktionen betrachten.

Beispiel 11.2. Sei $D \subset \mathbb{R}$ ein Definitionsbereich sowie $f, g\colon D \to \mathbb{R}$ Funktionen. Für alle $x \in D$ können z. B. wir folgende Funktionen definiert:

1. *Konstante Funktion*: $f(x) = c$ für ein $c \in \mathbb{R}$.
2. *Identische Abbildung*: $f(x) = x$. Oft schreiben wir $\mathrm{id}_D\colon D \to D$, $x \mapsto x$.
3. *Betragsfunktion*: $f(x) = |x|$.
4. *Exponenzialfunktion*: $f(x) = \exp(x)$.
5. *Trigonometrische Funktionen*: $f(x) = \sin x$ und $g(x) = \cos x$.
6. *Hyperbolische Funktionen*: $f(x) = \sinh x$ und $g(x) = \cosh x$.
7. Ferner können wir für $D \subset \mathbb{R} \setminus \{(k + \frac{1}{2}\pi) : k \in \mathbb{Z}\}$ definieren: $f(x) = \tan x, \cot x, \tanh x, \coth x$.
8. *Indikatorfunktion* (gelegentlich auch *charakteristische Funktion*): Für $F \subset D$ definieren wir

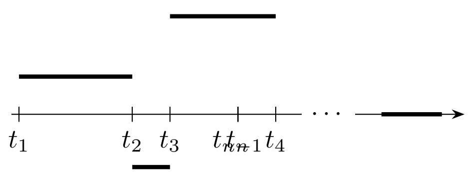
$$f = 1_F \qquad \text{durch} \qquad f(x) = \begin{cases} 1 & \text{falls } x \in F, \\ 0 & \text{sonst.} \end{cases}$$



Z. B. $1_{\mathbb{Q}}$ für $F = \mathbb{Q}$.

9. *Treppenfunktion*:

$$f = \sum_{k=1}^{n} c_k 1_{(t_k, t_{k+1}]} \qquad \text{für Intervalle } (t_k, t_{k+1}].$$[1]



10. *Gauss-Klammer* oder *Abrundungsfunktion*:

$$\lfloor x \rfloor = \sup\{k \in \mathbb{Z} : k \leq x\} = \left(\sum_{k \in \mathbb{Z}} k 1_{[k,k+1)}\right)(x).$$

Achtung: f ist eine *Funktion* $x \mapsto f(x)$ (z. B. $f = \sin$), während $f(x)$ eine *Zahl* ist (z. B. 0 für $x = \pi$).

[1] Der Dozent gab hier abgeschlossene Intervalle an, was aber nicht funktioniert.

Definition 11.3 (algebraische Operationen)**.** Gegeben seien zwei Funktionen $f, g\colon D \to \mathbb{R}$ und $\lambda \in \mathbb{R}$. Dann definieren wir *punktweise*[2]

$$(f+g)(x) \coloneqq f(x)+g(x), \qquad (\lambda f)(x) \coloneqq \lambda \cdot f(x), \qquad (fg)(x) \coloneqq f(x)\cdot g(x).$$

Ferner definieren wir auf $D' = \{x \in D : f(x) \neq 0\}$

$$\left(\frac{f}{g}\right)(x) = \frac{f(x)}{g(x)}.$$

Beispiel 11.4.

1. *Polynomfunktion*: $f\colon D \to \mathbb{R}$ mit beliebigen $D \subset \mathbb{R}$ hat die Form

   $$f(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

   mit $n \in \mathbb{N}$ und $a_0, \ldots, a_n \in \mathbb{R}$ mit $a_n \neq 0$. Sie entsteht mit den algebraischen Operationen aus der Identität und der konstanten Funktion.

2. *Rationale Funktion*: $f = g/h$ mit Polynomfunktionen f, g.

Definition 11.5 (Komposition)**.** Seien $f\colon D \to \mathbb{R}$ und $g\colon E \to \mathbb{R}$ Funktionen mit $f(D) \subset E$. Dann ist deren *Komposition* (auch *Verknüpfung*, *Hintereinanderausführung*)

$$g \circ f\colon D \xrightarrow{f} E \xrightarrow{g} \mathbb{R}$$

definiert durch

$$(g \circ f)(x) \coloneqq g(f(x)).$$

Achtung: Die Komposition ist <u>keine</u> Multiplikation der Funktionen.

Definition 11.6 (Umkehrfunktion)**.** Die Funktion $g\colon E \to \mathbb{R}$ heißt *Inverse* oder *Umkehrfunktion* zur Funktion $f\colon D \to \mathbb{R}$, falls

$$g \circ f\colon \mathrm{id}_D \qquad \text{und} \qquad f \circ g\colon \mathrm{id}_E$$

(insbesondere $g(E) = D$ und $f(D) = E$) gilt.

11.2 Stetigkeit

Hierfür gibt es zwei (tatsächlich äquivalente) Definitionen.

Definition 11.7 (ε-δ-Kriterium für Stetigkeit)**.**

1. Eine Funktion $f\colon D \to \mathbb{R}$ heißt *stetig im Punkt* $x_0 \in D$, falls für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, sodass

   $$|f(x) - f(x_0)| \leq \varepsilon \qquad \text{für alle } x \in D \text{ mit } |x - x_0| \leq \delta$$

   gilt. Oder in Zeichen:

   $$\forall \varepsilon > 0\colon \exists \delta > 0\colon \forall x \in D\colon \quad |x - x_0| \leq \delta \implies |f(x) - f(x_0)| \leq \varepsilon.$$

2. Eine Funktion $f\colon D \to \mathbb{R}$ heißt *stetig*, falls sie in jedem Punkt $x_0 \in D$ stetig ist.

Definition 11.8 (Folgenkriterium für Stetigkeit)**.** Eine Funktion $f\colon D \to \mathbb{R}$ heißt *stetig in* $x_0 \in D$, falls für jede Folge $(x_n)_{n\in\mathbb{N}}$ in D und $x_n \to x_0$ gilt $f(x_n) \to f(x_0)$. Oder anders formuliert: Für jede Folge $x_n \to x_0$ gilt

$$\lim_{n\to\infty} f(x_n) = f\Big(\lim_{n\to\infty} x_n\Big).$$

[2] *Punktweise* bedeutet, dass wir sie für jede Stelle $x \in D$ einzeln definieren.

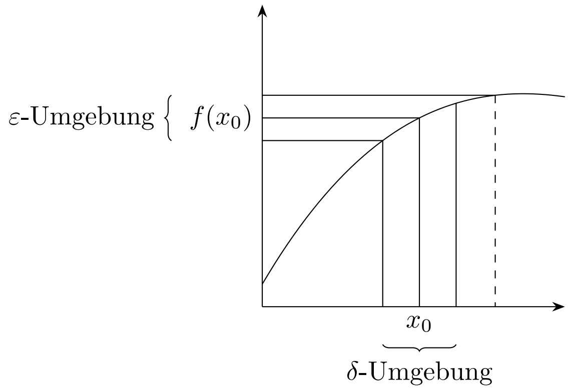


Abbildung 11.1: Darstellung des ε-δ-Kriteriums für Stetigkeit. Zu jeder ε-Umgebung finden wir eine δ-Umgebung, die innerhalb der Urbilder $f^{-1}(x_0 - \varepsilon)$ und $f^{-1}(x_0 + \varepsilon)$ liegen.

Die Idee der Stetigkeit ist, dass wir über eine Funktion f an einer Stelle x_0 sagen können, dass sie an dem Punkt „zusammenhängend" ist und keine „Sprünge" macht. D. h., dass eine noch so kleine Änderung ε um $f(x_0)$ nur auf eine kleine Änderung δ um x_0 zurückzuführen ist bzw. eine δ-Änderung um x_0 nur eine ε-Änderung um $f(x_0)$ bewirkt (und keine großen wie Sprünge).

Eine ähnliche Idee beschreibt das Folgenkriterium. Laufen wir von $-\infty$ (von links) und $+\infty$ (von rechts) Richtung x_0 und schauen uns dabei die Funktionswerte an, so können zwei Szenarien vorkommen: Treffen sich beide Grenzwerte der Funktionswerte, so gibt es keinen Sprung; treffen sie sich nicht, so gibt es einen Sprung. Viele Prozesse in der Natur sind stetig.[3]

Satz 11.9 (Äquivalenz der Stetigkeitsdefinitionen)**.** *Definitionen 11.7 und 11.8 sind äquivalent.*

Beweis. Wir wollen die Funktion $f\colon D \to \mathbb{R}$ an $x_0 \in D$ betrachten. Für die Äquivalenz zeigen wir die Hinrichtung und deren Kontraposition.

Nehmen wir zuerst das ε-δ-Kriterium an, d. h. es gibt für jedes $\varepsilon > 0$ ein $\delta > 0$, sodass für alle $x \in \mathbb{R}$ mit $|x - x_0| \leq \delta$ gilt: $|f(x) - f(x_0)| \leq \varepsilon$. Sei nun $(x_n)_n$ eine beliebige Folge in D, die für $n \to \infty$ gegen x_0 konvergiert. Aufgrund der Konvergenz finden wir für jedes gegebene δ ein $N \in \mathbb{N}$, sodass $|x_n - x| \leq \delta$ für alle $n \geq N$ ist. Insbesondere finden wir also für jedes $\varepsilon > 0$ ein N, sodass $|f(x_n) - f(x)| \leq \varepsilon$ für alle $n \geq N$ gilt. Damit erfüllt die Folge $(f(x_n))_n$ die Konvergenzdefinition von Folgen, also $f(x_n) \to f(x_0)$ für $n \to \infty$, was das Folgenkriterium ist.

Für die Rückrichtung nehmen wir an, dass das ε-δ-Kriterium nicht gelte, d. h. es gibt ein $\varepsilon > 0$, sodass für alle $\delta > 0$ ein x mit $|x_n - x| \leq \delta$ existiert, wofür aber $|f(x) - f(x_0)| > \varepsilon$ gilt. Wählen wir nun $\delta = 1/n$, gibt es für jedes $n \in \mathbb{N}$ ein x_n mit $|x_n - x| \leq 1/n$, aber $|f(x_n) - f(x)| > \varepsilon$. Damit haben wir eine Folge $(x_n)_n$ konstruiert, die zwar gegen x_0 konvergiert, wofür aber $f(x_n) \not\to f(x_0)$ gilt. Folglich ist das Folgenkriterium nicht erfüllt. □

Proposition 11.10. *Die Exponenzialfunktion* $x \mapsto \exp(x)$ *ist stetig auf* $\mathbb{R}$.

Beweis. Mit der Restgliedabschätzung für exp gilt für alle x mit $|x - x_0| \leq 1$

$$|\exp(x) - \exp(x_0)| = \left|\frac{\exp(x)}{\exp(x_0)} - 1\right| \exp(x_0) = |\exp(x - x_0) - 1| \cdot \exp(x_0) \leq 2|x - x_0| \cdot \exp(x_0).$$

Für jedes $\varepsilon > 0$ können wir ein

$$\delta = \min\left\{1, \frac{\varepsilon}{2\exp(x_0)}\right\} > 0$$

wählen, sodass für alle $|x - x_0| \leq \delta$ auch die obige Ungleichung weiter durch ε abgeschätzt werden kann. □

Proposition 11.11. *Für alle* $k \in \mathbb{N}$ *ist* $x \mapsto x^{1/k}$ *stetig auf* $\mathbb{R}_+$.

[3]Letztendlich beschreibt die Stetigkeit einen Begriff der Nähe: Sind im Definitionsbereich Werte „nah" beieinander, so sind sie auch nach Anwendung von f „nah" beieinander. Dieser erweiterte Begriff spielt auch in der Topologie eine wichtige Rolle.

Beweis. Wir zeigen zunächst
$$\left|x^{1/k} - x_0^{1/k}\right| \leq |x - x_0|^{1/k}.$$

Setzen wir $y = x^{1/k}$ und $y_0 = y_0^{1/k}$, und nehmen wir o. B. d. A. $y \geq y_0$ an, gilt nach dem binomischen Lehrsatz
$$y^k = (y_0 + (y - y_0))^k = y_0^k + \underbrace{\binom{k}{1} y_0^{k-1}(y - y_0)}_{\geq 0} + \cdots + \underbrace{\binom{k}{k-1} y_0 (y - y_0)^{k-1}}_{\geq 0} + (y - y_0)^k \geq y_0^k + (y - y_0)^k.$$

Um auch den Fall $y < y_0$ zu beachten, schreiben wir allgemeiner
$$|x - x_0| = \left|y^k - y_0^k\right| \geq |y - y_0|^k = \left|x^{1/k} - x_0^{1/k}\right|^k \implies |x - x_0|^{1/k} \geq \left|x^{1/k} - x_0^{1/k}\right|.$$

Wählen wir nun $\delta = \varepsilon^k > 0$, erhalten wir für alle x mit $|x - x_0| \leq \delta$
$$\left|x^{1/k} - x_0^{1/k}\right|^k \leq |x - x_0|^{1/k} \leq \left(\varepsilon^k\right)^{1/k} = \varepsilon.$$ □

Proposition 11.12. *Die konstante Funktion, die Identität und die Betragsfunktion sind stetig auf $\mathbb{R}$.*

Proposition 11.13 (Stetigkeit unter algebraischen Operationen)**.** *Mit Funktionen f und g sind auch die Funktionen $f + g$, $f \cdot g$ und f/g (falls $g(x_0) \neq 0$) stetig in x_0.*

Beweis. Wir reduzieren alles auf die Grenzwertsätze von Folgen. Sei $(x_n)_n$ eine beliebige Folge mit $x_n \to x_0$. Nach Annahme gilt $f(x_n) \to f(x_0)$ und $g(x_n) \to g(x_0)$, und daher $(f+g)(x_n) \to (f+g)(x_0)$ und $(fg)(x_n) \to (fg)(x_0)$.

Falls $g(x_0) \neq 0$, so ist auch $g(x_n) \neq 0$ für schließlich alle n und damit
$$\left(\frac{f}{g}\right)(x_n) = \frac{f(x_n)}{g(x_n)} \to \frac{f(x_0)}{g(x_0)} = \left(\frac{f}{g}\right)(x_0).$$ □

Korollar 11.14. *Jede Polynomfunktion ist stetig auf $\mathbb{R}$. Jede rationale Funktion ist stetig auf ihrem Definitionsbereich.*

Proposition 11.15 (Stetigkeit unter Komposition)**.** *Ist eine Funktion $f\colon D \to \mathbb{R}$ stetig in x_0 und eine Funktion $g\colon E \to \mathbb{R}$ stetig in $f(x_0)$, so ist auch $g \circ f$ stetig in x_0.*

Beweis. Wir benutzen das Folgenkriterium: Sei $(x_n)_n$ eine beliebige Folge in D mit $x_n \to x_0$. Aufgrund der Stetigkeit von f gilt für die Folge der $f(x_n)$ in E auch $f(x_n) \to f(x_0)$. Wiederum wegen der Stetigkeit von g gilt für die Folge $g(f(x_n))$ auch $g(f(x_n)) \to g(f(x_0))$. Damit gilt für $x_n \to x_0$ auch $g(f(x_n)) \to g(f(x_0))$ und das Folgenkriterium in x_0 ist für $g \circ f$ erfüllt. □

Korollar 11.16. *Ist f eine stetige Funktion, so ist auch $f^{1/k}$ stetig (in x_0 bzw. auf D).*

Insbesondere ist auch $x \mapsto x^s$ für jedes $s \in \mathbb{Q}$ stetig auf $(0, \infty)$ (bzw. für jedes $s \in \mathbb{Q}^+$ auf $[0, \infty)$) und $x \mapsto |x|^s$ für jedes $s \in \mathbb{Q}$ stetig auf $\mathbb{R} \setminus \{0\}$ (bzw. für jedes $s \in \mathbb{Q}^+$ auf $\mathbb{R}$).

Ende der Vorlesung 17 am 13. Dezember 2021

Satz 11.17 (Zwischenwertsatz)**.** *Eine stetige Funktion $f\colon [a, b] \to \mathbb{R}$ nimmt für jeden Wert γ zwischen $f(a)$ und $f(b)$ an, d. h. es gibt für jedes $\gamma \in [f(a), f(b)]$ ein $c \in [a, b]$ mit $f(c) = \gamma$.*

Ist insbesondere $f(a) \leq 0 \leq f(b)$, so hat f eine Nullstelle.

Bemerkung 11.18. Der Definitionsbereich muss ein Intervall sein. Der Satz ist bspw. für $D = \mathbb{Q} \cap [a, b]$ oder für $D = \{a, b\}$ falsch.

Beweis. O. B. d. A. können wir annehmen, dass $f(a) \leq \gamma \leq f(b)$ gilt (für $f(a) \geq f(b)$ funktioniert der Beweis analog).

Wir definieren eine Intervallschachtelung durch Halbierung: Seien am Anfang $a_1 \coloneqq a$ und $b_1 \coloneqq b$. Danach definieren wir

$$a_{n+1} = \begin{cases} a_n & \text{falls } f\left(\dfrac{a_n + b_n}{2}\right) \geq \gamma, \\ \dfrac{a_n + b_n}{2} & \text{sonst;} \end{cases} \qquad b_{n+1} = \begin{cases} \dfrac{a_n + b_n}{2} & \text{falls } f\left(\dfrac{a_n + b_n}{2}\right) \geq \gamma, \\ b_n & \text{sonst.} \end{cases}$$

Per Induktion können wir $f(a_n) \leq \gamma \leq f(b_n)$ und $|b_n - a_n| = 2^{1-n}|b - a|$ für alle $n \in \mathbb{N}$ zeigen (was relativ einfach ist). Aus dem Intervallschachtelungsprinzip folgt nun, dass es ein c gibt mit

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

Aus der Stetigkeit via Folgenkriterium und dem Sandwichlemma folgt

$$f(a_n) \leq \gamma \leq f(b_n) \implies f(c) = \lim_{n\to\infty} f(a_n) \leq \gamma \leq \lim_{n\to\infty} = f(c) \implies f(c) = \gamma. \qquad \square$$

Der Zwischenwertsatz ist sehr mächtig, da wir mit ihm Aussagen über die Bildmenge einer stetigen Funktion machen und daraus auch einige Eigenschaften ermitteln können.

Korollar 11.19. *Jedes Polynom ungeraden Grades, also*

$$f\colon \mathbb{R} \to \mathbb{R} \quad \textit{mit} \quad f(x) = a_{2n+1}x^{2n+1} + a_{2n}x^{2n} + \cdots + a_0 \quad \textit{mit } a_{2n+1} \neq 0,$$

hat mindestens eine reelle *Nullstelle*

Beweis. O. B. d. A. sei $a_{2n+1} > 0$. Dann konvergieren die Funktionswerte uneigentlich: $\lim_{x\to\infty} f(x) = \infty$ und $\lim_{x\to-\infty} f(x) = -\infty$. Folglich gibt es reelle Zahlen $a, b \in \mathbb{R}$ mit $a < b$ und $f(a) < 0 < f(b)$. Damit existiert eine Nullstelle in $[a, b]$. $\square$

Korollar 11.20. *Sei $I \in \mathbb{R}$ ein beliebiges Intervall (offen, abgeschlossen, halboffen, beschränkt, unbeschränkt). Ist eine Funktion $f\colon I \to \mathbb{R}$ stetig auf I, so ist auch $f(I)$ ein Intervall.*

Beweis. Für $I = \varnothing$ ist offensichtlich $f(I) = \varnothing \subset \mathbb{R}$ auch ein „Intervall“. Deshalb nehmen wir o. B. d. A. $I \neq \varnothing$ an. Für $I = [x, x]$ ist auch offensichtlich $f(I) = [f(x), f(x)]$ ein Intervall. Deshalb nehmen wir auch $\operatorname{diam}(I) > 0$ an.

Seien $A = \inf f(I) \in [-\infty, \infty)$ und $B = \sup f(I) \in (-\infty, \infty]$ (in $\overline{\mathbb{R}}$). Für jedes $\gamma \in (A, B)$ gibt es $a, b \in I$ mit $f(a) < \gamma < f(b)$. Weil f insbesondere auf (a, b) stetig ist, gibt es nach Zwischenwertsatz ein $c \in I$ mit $f(c) = \gamma$. Daraus folgt $\gamma \in f(I)$ für alle $\gamma \in (A, B)$, oder $(A, B) \subset f(I) \subset ([A, B] \cap \mathbb{R})$ (geschnitten mit $\mathbb{R}$, da wir nicht erlauben wollen, dass $\pm\infty \in f(I)$). Somit haben wir $f(I)$ geschachtelt, und $f(I)$ kann nur eines von (A, B), $[A, B]$, $(A, B]$ oder $[A, B)$ (geschnitten mit $\mathbb{R}$) sein. $\square$

Der Grund, wieso wir nochmals a und b einführen mussten ist, dass A und B nicht notwendigerweise Funktionswerte sein müssen.

Definition 11.21 (kompaktes Intervall)**.** Ein Intervall $I \subset \mathbb{R}$ heißt *kompakt*, falls es abgeschlossen und beschränkt ist.

Satz 11.22 (Extremwertsatz)**.** *Das Bild eines kompakten Intervalls unter einer stetigen Abbildung ist ein kompaktes Intervall. Insbesondere nimmt jede stetige Funktion auf einem kompakten Intervall ihre Extrema an, d. h. für ein stetiges $f\colon I \to \mathbb{R}$ mit $I \subset \mathbb{R}$ kompakt gibt es $p, q \in I$ mit*

$$f(p) = \sup f(I) \qquad \textit{und} \qquad f(q) = \inf f(I).$$

Oder mit anderen Worten: Es existieren $p, q \in I$, sodass $f(q) \leq f(x) \leq f(p)$ für alle $x \in I$ gilt und damit $f(I) = [f(q), f(p)]$.

Bemerkung 11.23. Der Satz stimmt für offene/halboffene Intervalle nicht. Bspw. ist $f\colon x \mapsto 1/x$ auf $I = (0, 1]$ definiert, wobei I beschränkt ist. Aber $f(I) = [1, \infty)$ ist ein *unbeschränktes* Intervall, d. h. es gibt kein $p \in I$ mit $f(p) = \infty \notin \mathbb{R}$.

Beweis des Extremwertsatzes. Sei I ein kompaktes Intervall und $f\colon I \to \mathbb{R}$ eine stetige Funktion. Seien $A = \inf f(I)$ und $B = \sup f(I)$ (wobei $A, B \in \overline{\mathbb{R}}$). Dann existieren Folgen $(a_n)_n$ und $(b_n)_n$ in I mit $f(a_n) \to A$ und $f(b_n) \to B$.[4]

Dabei sind die Folgen $(a_n)_n$ und $(b_n)_n$ beschränkt, da I beschränkt ist. Folglich können wir den Satz von Bolzano-Weierstraß anwenden, sodass $(a_n)_n$ und $(b_n)_n$ einen Häufungspunkt q bzw. p besitzen, d. h. es gibt Teilfolgen $(a_{n_k})_k$ und $(b_{n_k})_k$ mit $a_{n_k} \to q$ bzw. $b_{n_k} \to p$.

Wegen der Abgeschlossenheit von I ist $p, q \in I$ und f ist in p und q definiert. Weil in einer konvergenten Folge jede Teilfolge gegen denselben Grenzwert konvergiert, folgt aus der Stetigkeit von f

$$A = \lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} f(a_{n_k}) = f\Big(\lim_{n\to\infty} a_{n_k}\Big) = f(q)$$

und analog

$$B = \lim_{n\to\infty} f(b_{n_k}) = f(p).$$

Wegen dem Zwischenwertsatz gilt $[f(p), f(q)] \subset f(I) \subset [A, B]$ und damit $f(I) = [f(q), f(p)]$. □

Definition 11.24 (gleichmäßige Stetigkeit)**.** Eine Funktion $f\colon D \to \mathbb{R}$ heißt *gleichmäßig stetig*, falls für jedes $\varepsilon > 0$ ein $\delta > 0$ existiert, sodass

$$|f(x) - f(x_0)| \le \varepsilon \qquad \text{für alle } x, x_0 \in D \text{ mit } |x - x_0| \le \delta$$

gilt. Oder in Zeichen:

$$\forall \varepsilon > 0\colon \exists \delta > 0\colon \forall x, x_0 \in D\colon \quad |x - x_0| \le \delta \implies |f(x) - f(x_0)| \le \varepsilon.$$

Bemerkung 11.25*.* Wenn wir uns die Definition der Stetigkeit

$$\forall x_0 \in D\colon \forall \varepsilon > 0\colon \exists \delta > 0\colon \forall x \in D\colon \quad |x - x_0| \le \delta \implies |f(x) - f(x_0)| \le \varepsilon$$

anschauen, so ist das $\forall x_0 \in D$ nach Innen gerutscht. Das bedeutet, dass δ bei der gleichmäßigen Stetigkeit nur noch von ε abhängen darf, wohingegen δ bei der (punktweisen) Stetigkeit von ε und x_0 abhängen durfte.

Damit ist die gleichmäßige Stetigkeit eine stärkere Bedingung, und sie impliziert auch die einfache Stetigkeit.[5]

Satz 11.26. *Sei I ein kompaktes Intervall. Dann ist die Funktion $f\colon I \to \mathbb{R}$ genau dann stetig, wenn sie gleichmäßig stetig ist.*[6]

Beweis. Die Rückrichtung ist trivial (s. Bemerkung 11.25).

Die Hinrichtung zeigen wir indirekt: Angenommen f ist nicht gleichmäßig stetig, d. h. es gibt ein $\varepsilon > 0$, sodass für jedes $\delta = 1/n > 0$ Stellen $x_n, y_n \in I$ mit $|x_n - y_n| \le \delta = 1/n$ und $|f(x_n) - f(y_n)| > \varepsilon$ gibt.

$(x_n)_n$ und $(y_n)_n$ sind Folgen in I und damit beschränkt. Mit dem Satz von Bolzano-Weierstraß gibt es eine Teilfolge $(x_{n_k})_k$ mit $x_{n_k} \to p \in I$ (ein Häufungspunkt). Aus der Bedingung $|x_n - y_n| \le 1/n$ folgt $y_{n_k} - x_{n_k} \to 0$, also $y_{n_k} \to p$. Aufgrund Stetigkeit von f erhalten wir

$$f(y_{n_k}) - f(x_{n_k}) \to f(p) - f(p) = 0,$$

ein Widerspruch zur Annahme $|f(y_{n_k}) - f(x_{n_k})| > \varepsilon$ für alle $k \in \mathbb{N}$. □

Beispiel 11.27*.*

1. Die Funktion $x \mapsto \sin(1/x)$ auf $(0, 1] \subset \mathbb{R}$ ist *nicht* gleichmäßig stetig.

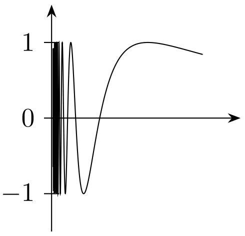
[4]Leider hat auch hier der Dozent eine Aussage verwendet, die wir noch nicht bewiesen haben: Ist $M \subset \mathbb{R}$, so gibt es eine Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n \in M$ für alle n und $a_n \to \sup M$. *Beweis*: Für jedes $n \in \mathbb{N}$ muss es ein $a_n \in M$ mit $\sup M - 1/n \le a_n$ geben; andernfalls wäre $\sup M - 1/n$ eine kleinere obere Schranke als $\sup M$. Nun gilt auch $a_n \le \sup M$, also $a_n \to \sup M$ nach dem Sandwichlemma.

[5]Wir können den Unterschied auch im Sinne von Approximation verstehen: Bei der Stetigkeit können wir für jede Stelle x_0 und jeden Maximalfehler ε in den Funktionswerten eine δ-Umgebung finden, sodass der Fehler in dieser Umgebung kleiner ε ist, d. h. die $f(x)$ mit $x \in [x_0 - \delta, x_0 + \delta]$ approximieren $f(x_0)$ gut. Dabei kann aber die δ-Umgebung von x_0 abhängen.

Bei der gleichmäßigen Stetigkeit haben wir hingegen eine gleichmäßige Approximation: Für jeden Maximalfehler ε gibt es ein δ, sodass wir uns sicher sein können, dass die δ-Umgebung um *jede Stelle* x_0 ausreichend gute Approximation ist.

Eine weitere Eigenschaft: Jede Cauchy-Folge bleibt unter einer gleichmäßig stetigen Funktion wieder eine Cauchy-Folge, was wir bei punktweiser Stetigkeit nicht behaupten können.

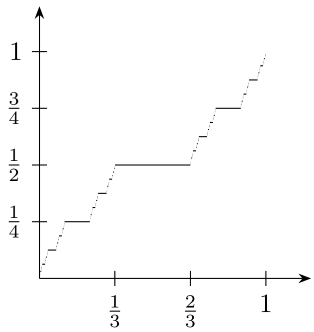
[6]Der Satz ist auch als *Satz von* H*EINE* bekannt.



2. Die Funktion $x \mapsto 1/x$ auf $I = (0,1]$ ist zwar stetig, aber nicht gleichmäßig stetig, denn für $\varepsilon = 1$ gibt es für jedes $\delta > 0$ Stellen $x = \min\{\frac{1}{2}, \delta\} \in I$ und $x' = 2x \in I$ mit $|x - x'| = x \leq \delta$, aber $|f(x) - f(x')| = 1/(2x) \geq 1 = \varepsilon$.

3. Die *Verteilungsfunktion auf der* CANTOR*-Menge* bzw. CANTOR*-Verteilung* (auch engl. *devil's staircase*) auf $[0,1]$ ist gleichmäßig stetig.

   Die Funktionen wird über die rekursiv definierten Cantor-Mengen konstruiert. Das Intervall $[0,1]$ wird gedrittelt, wobei für das mittlere Drittel $(\frac{1}{3}, \frac{2}{3})$ die Funktion den Wert $\frac{1}{2}$ annimmt. Dann bleiben noch das linke Drittel $L = [0, \frac{1}{3}]$ und das rechte Drittel $R = [\frac{2}{3}, 1]$ übrig. Auf diesen beiden Machen wir dieselbe Konstruktion, wobei das mittlere Drittel von L den Wert $\frac{1}{4}$, das mittlere Drittel von R den Wert $\frac{3}{4}$ annimmt. Das können wir so fortführen.



   Nach dem k-ten Schritt ist kein „Sprung“ größer als 2^{-k}.

Satz 11.28. *Sei $D \subset \mathbb{R}$ ein Intervall und $f\colon D \to \mathbb{R}$ stetig und strikt monoton. Dann bildet f das Intervall D bijektiv auf das Intervall $D' = f(D)$ ab und die Umkehrfunktion $f^{-1}\colon D' \to \mathbb{R}$ ist stetig und strikt monoton.*

Beweis. Wegen Korollar 11.20 ist D' ein Intervall. Weil f strikt monoton, ist es auch injektiv. Zusammen ist also $f\colon D \to D'$ bijektiv. Folglich existiert $f^{-1}\colon D' \to D$ und ist strikt monoton, denn für z. B. monoton wachsende Funktion gilt $x < y \iff f(x) < f(y) \iff f^{-1}(f(x)) < f^{-1}(f(y))$ und analog für monoton fallende.

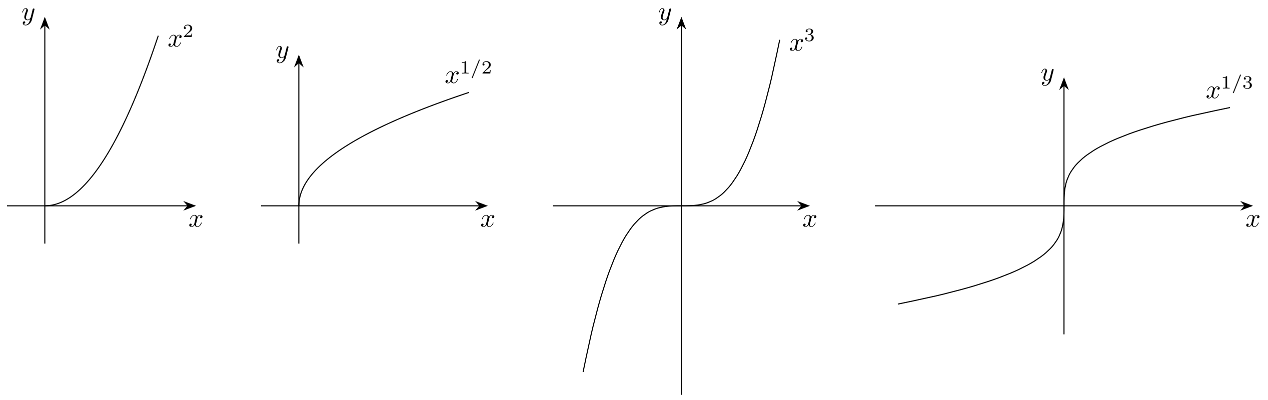
Nun müssen wir nur noch die Stetigkeit zeigen, die wir indirekt mittels dem Folgenkriterium zeigen wollen. Angenommen es gäbe eine Stelle $y \in D'$, in der f^{-1} nicht stetig wäre, d. h. es gibt eine Folge $(y_n)_n$ in D' mit $y_n \to y$, wofür $x_n \coloneqq f^{-1}(y_n) \not\to f^{-1}(y) =: x$ gilt. Weil $(y_n)_n$ konvergiert, ist die Folge beschränkt (Satz 4.12). Damit gibt es ein kompaktes Intervall $f(D_0)$ mit $D_0 \subset I$, das alle y_n enthält. Aufgrund der Monotonie muss auch D_0 selbst kompakt sein, d. h. $(x_n)_n$ ist beschränkt. Nach Bolzano-Weierstraß gibt es eine konvergente Teilfolge $x_{n_k} \to \bar{x} \in D_0$ mit $\bar{x} \neq x$, denn $(x_n)_n$ konvergiert nicht, sodass es mindestens zwei verschiedene Häufungspunkte geben muss. Daraus folgt aufgrund der Stetigkeit von f nun

$$f(\bar{x}) = \lim_{n\to\infty} f(x_{n_k}) = \lim_{n\to\infty} y_{n_k} = y = f(x),$$

und aufgrund Bijektivität $\bar{x} = x$, ein Widerspruch. □

Beispiel 11.29. Für $k \in \mathbb{N}$ ist die Potenzfunktion $f\colon \mathbb{R}_+ \to \mathbb{R}_+$, $x \mapsto x^k$ strikt monoton wachsend und stetig. Sie bildet $\mathbb{R}_+$ also bijektiv auf $\mathbb{R}_+$ ab. Die Umkehrfunktion ist die k-te Wurzel $x \mapsto x^{1/k} = \sqrt[k]{x}$.

Für ungerade k ist $x \mapsto x^k$ auf ganz $\mathbb{R}$ strikt monoton. Somit können wir $x \mapsto x^{1/k}$ auf ganz $\mathbb{R}$ definieren.



Ende der Vorlesung 18 am 15. Dezember 2021

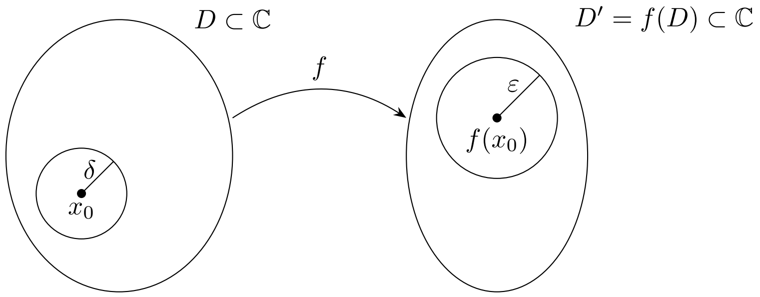
11.3 Stetigkeit für Funktionen in $\mathbb{C}$

Viele Begriffe, die über Stetigkeit in $\mathbb{R}$ einführten, können wir nach $\mathbb{C}$ übertragen, wie z.Ḃ. die Stetigkeit selbst.

Definition 11.30 (Stetigkeit in $\mathbb{C}$)**.** Für eine komplexwertige Funktion $f\colon D \to \mathbb{C}$ mit $D \subset \mathbb{C}$ ist *stetig in* $z_0 \in \mathbb{C}$, falls

1. es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, sodass $|f(z) - f(z_0)| \leq \varepsilon$ für alle $z \in D$ mit $|z - z_0| \leq \delta$ gilt; oder falls
2. es für alle Folgen $(z_n)_n$ in D mit $z_n \to z_0$ für $n \to \infty$ gilt: $f(z_n) \to f(z_0)$.

Beide sind auch in $\mathbb{C}$ äquivalent.



Jedoch ist nicht alles nach $\mathbb{C}$ übertragbar, da die Ordnung fehlt. Bspw. gibt es keinen Zwischenwertsatz oder Extremwertsatz.

Proposition 11.31. $z \mapsto \exp(z)$ *ist stetig auf* $\mathbb{C}$.

Beweis. Der Beweis erfolgt exakt wie zuvor. □

Proposition 11.32. *Folgende Aussagen sind für eine Funktion $f\colon D \to \mathbb{C}$ äquivalent.*

1. f ist stetig.

2. $\operatorname{Re}(f)$ *und* $\operatorname{Im}(f)$ *sind stetig.*

3. $\overline{f}$ ist stetig.

Beweis. Aus Punkt 1 folgt Punkt 2, denn wenn $f(z_n) \to f(z_0)$ konvergiert, dann auch Real- und Imaginärteil konvergieren (Satz 10.9). Umgekehrt folgt Punkt 2 aus Punkt 1, denn die Summe und das Produkt stetiger Funktionen ist wieder stetig (nämlich $f = \operatorname{Re}(f) + \mathrm{i}\operatorname{Im}(f)$).

Genauso ergibt sich die Äquivalenz der Punkte 2 und 3. □

Korollar 11.33. *Die Funktionen $x \mapsto \sin x$ und $x \mapsto \cos x$ sind stetig auf $\mathbb{R}$.*

Beweis. Weil die Funktion $y \mapsto \exp(\mathrm{i}y) = \cos y + \mathrm{i} \sin y$ stetig ist, sind auch $y \mapsto \cos y = \mathrm{Re}(\exp(\mathrm{i}y))$ und $y \mapsto \sin y = \mathrm{Im}(\exp(\mathrm{i}y))$ stetig auf $\mathbb{R}$. $\square$

Alternativer Beweis über ε-δ-Kriterium. Mit den Additionstheorem gilt für jedes $x \in \mathbb{R}$ und jedes $\delta > 0$

$$\cos(x + \delta) - \cos x = \cos x(\cos \delta - 1) - \sin x \sin \delta.$$

Nun können wir $0 \leq |\sin \delta| \leq \delta$ und $0 \leq 1 - \cos \delta \leq \delta^2/2$ abschätzen, d. h. nach dem Sandwichlemma werden die abgeschätzten Terme 0 für $\delta \to 0$. Damit können wir für jedes gegebene $\varepsilon > 0$ hinreichend kleine $\delta > 0$ finden, sodass $|\cos(x + \delta) - \cos x| \leq \varepsilon$ ist.

Analog können wir für $\delta \to 0$

$$\sin(x + \delta) - \sin x = \sin x \underbrace{(\cos \delta - 1)}_{\to 0} - \cos x \underbrace{\sin \delta}_{\to 0} \to 0$$

schließen. Eine andere Möglichkeit wäre durch Phasenverschiebung $\sin x = \cos(x + \frac{1}{2}\pi)$ und dem Fakt, das Komposition stetiger Funktionen wieder stetig sind. $\square$

Korollar 11.34. *Die Funktion $x \mapsto \cos x$ auf $[0, 2]$ besitzt eine eindeutige Nullstelle (nämlich $\frac{1}{2}\pi$, vgl. Proposition 10.27).*

Beweis. Zuerst zeigen wir, dass eine Nullstelle existiert. Aus der Restgliedabschätzung des Kosinus erhalten wir

$$\cos 2 \leq 1 - \frac{2^2}{2} + \frac{2^4}{24} < 0 \qquad \text{und} \qquad \cos 0 \geq 1 - \frac{0^2}{2} > 0.$$

Nach dem Zwischenwertsatz gibt es eine Nullstelle.

Nun zeigen wir, dass cos auf $[0, 2]$ streng monoton fällt und damit nach Satz 11.28 bijektiv ist. Mit der Restgliedabschätzung des Sinus gilt für $x \in (0, 2]$

$$x^2 < 6 \iff \frac{x^3}{6} < x \iff 0 < x - \frac{x^3}{6} \leq \sin x.$$

Für alle $0 \leq x < x' \leq 2$ gilt nun mit den Doppelwinkelfunktionen

$$\cos x' - \cos x = -2 \sin \frac{x' + x}{2} \sin \frac{x' - x}{2} < 0,$$

denn die Argumente der Sinusfaktoren liegen in $(0, 2]$, weshalb beide Sinusfaktoren positiv sind.

Damit ist cos auf $[0, 2]$ bijektiv und die Nullstelle ist eindeutig.[7] $\square$

11.4 Logarithmus

Proposition 11.35 (natürlicher Logarithmus)**.**

1. *Die Exponenzialfunktion* $\exp\colon \mathbb{R} \to \mathbb{R}$ *ist stetig und strikt monoton. Deshalb bildet sie* $\mathbb{R}$ *bijektiv auf* $\mathbb{R}_+^* = (0, \infty)$ *ab.*

2. *Die Umkehrfunktion ist der* natürliche Logarithmus *(oder* Logarithmus zur Basis e*)*

$$\log\colon \mathbb{R}_+^* \to \mathbb{R}$$

(auch mit ln *abgekürzt) Er genügt der Funktionalgleichung*

$$\log(xy) = \log(x) + \log(y) \qquad \textit{für alle } x, y \in \mathbb{R}_+^*.$$

Beweis.

[7] Der Dozent zeigte in der Vorlesung nur, dass das Infimum aller Nullstellen auch eine Nullstelle ist via Stetigkeit: Es gibt eine Folge $(y_n)_n$ von Nullstellen, die gegen das Infimum y konvergiert. Aufgrund der Stetigkeit haben wir $0 = \cos y_n \to \cos y$, also $\cos y = 0$.

Dann fuhr er fort zu behaupten, dass cos auf $(\frac{1}{2}\pi, \frac{3}{2}\pi)$ keine weitere Nullstellen besitzt, was auch korrekt ist. Das haben wir aber nur gezeigt, als wir schon die Eindeutigkeit von $\frac{1}{2}\pi$ angenommen haben, also ein Zirkelschluss.

1. Um die strikte Monotonie nachzuweisen, betrachten wir $x < y = x + h$ mit $h > 0$. Dann gilt

$$\exp(y) = \exp(x)\exp(h) \geq \exp(x)(1 + h) > \exp(x),$$

also ist exp streng monoton wachsend.

Nun zur Bildmenge: Weiterhin gilt $n < \exp(n)$ für alle $n \in \mathbb{N}$, also divergiert auch $\exp(n)$ bzw. $\lim_{x\to\infty} \exp(x) = \infty$. Genauso ist $\lim_{x\to-\infty} \exp(x) = \lim_{x\to\infty} 1/\exp(x) = 0$. Nach dem Zwischenwertsatz ist also $f(\mathbb{R}) = (0, \infty)$.

2. Sei $a = \log(x)$ und $b = \log(y)$. Dann gilt

$$\exp(a + b) = \exp(a)\exp(b) = xy \implies \log(xy) = a + b = \log x + \log y. \qquad \square$$

Bemerkung 11.36. log bildet die multiplikative Gruppe $(\mathbb{R}^*_+, \cdot)$ isomorph auf die additive Gruppe $(\mathbb{R}, +)$ ab; sie macht also das Umgekehrte von exp.

Definition 11.37 (Logarithmus, Potenz zur Basis a)**.** Für $a \in \mathbb{R}^*_+ \setminus \{1\}$ definieren wir den *Logarithmus zur Basis a* als

$$\log_a(x) = \frac{\log(x)}{\log(a)} \qquad \text{für alle } x \in \mathbb{R}^*_+$$

und die *Potenz bzw. Exponenzialfunktion zur Basis a*

$$a^x = \exp(x\log(a)) \qquad \text{für alle } x \in \mathbb{R}$$

Nun haben wir Potenzen für alle reellen Exponenten definiert.

Satz 11.38 (Potenzgesetze)**.** *Für alle $a, b \in \mathbb{R}^*_+$, für alle $x, y \in \mathbb{R}$ und für alle $m, n \in \mathbb{N}$ gelten*

1. $a^x a^y = a^{x+y}$,
2. $(a^x)^y = a^{xy}$,
3. $a^x b^x = (ab)^x$,
4. $a^{m/n} = \sqrt[n]{a^m}$ *mit* $a^m = a \cdots a$ *(m Faktoren),*
5. $a^{\log_a(x)} = x$ *und* $\log_a(a^x) = x$*, d. h. die Funktionen* $x \mapsto \log_a(x)$ *und* $x \mapsto a^x$ *sind invers zueinander.*

Beweis.

1. Es gilt per Funktionalgleichung von exp

$$a^x a^y = \exp(x\log(a))\exp(y\log(a)) = \exp((x + y)\log(a)) = a^{x+y}.$$

2. Weil log die Umkehrfunktion von exp ist, gilt

$$(a^x)^y = \exp(y\log(a^x)) = \exp(y\log(\exp(x\log(a)))) = \exp(yx\log(a)) = a^{xy}.$$

3. Aus den Funktionalgleichungen folgt

$$a^x b^x = \exp(x\log(a))\exp(x\log(b)) = \exp(x(\log(a) + \log(b))) = \exp(x\log(ab)) = (ab)^x.$$

4. Aus Punkt 1 folgt für alle $m \in \mathbb{N}$

$$a^m = \exp(m\log(a)) = \exp(\underbrace{\log(a) + \cdots + \log(a)}_{m \text{ mal}}) = \underbrace{\exp(\log(a)) + \cdots + \exp(\log(a))}_{m \text{ mal}} = \underbrace{a \cdots a}_{m \text{ mal}}.$$

Mit Punkt 2 folgt $(a^m)^{1/n} = a^{m/n}$.

5. Es gilt

$$a^{\log_a(x)} = \exp(\log_a(x)\log(a)) = \exp\left(\frac{\log(x)}{\log(a)}\log(a)\right) = \exp(\log(x)) = x \qquad \text{für alle } x \in \mathbb{R}_+^*$$

und analog

$$\log_a(a^x) = \frac{\log(a^x)}{\log(a)} = \frac{\log(\exp(x\log(a)))}{\log(a)} = \frac{x\log(a)}{\log(a)} = x \qquad \text{für alle } x \in \mathbb{R}. \qquad \square$$

Bemerkung 11.39.

1. *Achtung*: Wir können keinen Logarithmus zur Basis $a = 1$ definieren, da zum einen $x \mapsto 1^x$ nicht bijektiv ist, und zum anderen wir in der Definition durch $\log(a) = \log(1) = 0$ teilen würden.
2. Wichtige Spezialfälle für $\log_a$ sind $a = 2$, $a = 10$ und $a = \mathrm{e}$. Für letzteres schreiben wir auch $\ln(x) = \log(x) = \log_{\mathrm{e}}(x)$.
3. Für $a, b \in \mathbb{R}_+ \setminus \{0, 1\}$ gilt
$$\log_b(x) = c\log_a(x) \qquad \text{mit } c = \frac{\log(a)}{\log(b)}.$$
Den Graphen skalieren wir hier also bei Basiswechsel in y-Richtung.
4. Für alle $x \in \mathbb{R}$ gilt $\exp(x) = \mathrm{e}^x$.
5. Für rationale x stimmt die neue Definition a^x mit der früheren überein.

Bemerkung 11.40. Alternativ hätten wir die allgemeine Potenz für rationale x definieren können: Für alle Basen $a \in \mathbb{R}_+$ definieren wir a^x mit $x \in \mathbb{Q}$ als

$$a^x \coloneqq (\exp(\log(a)))^x = \exp(x\log(a)).$$

Anschließend können wir durch *stetige Fortsetzung*[8] für alle irrationalen x

$$a^x \coloneqq \lim_{\substack{y\in\mathbb{Q}\\ y\to x}} a^y$$

definieren (per Definition ist a^x stetig). Dabei existiert der Grenzwert, denn exp ist stetig auf $\mathbb{R}$, und damit gilt $\exp(y\log(a)) \to \exp(x\log(a))$ für $y \in \mathbb{Q}$.

11.5 Umkehrfunktionen der hyperbolischen Funktionen

Definition 11.41 (Areafunktionen)**.** Die hyperbolischen Funktionen

$$\sinh\colon \mathbb{R} \to \mathbb{R}, \qquad \cosh\colon [0,\infty) \to [1,\infty), \qquad \tanh\colon \mathbb{R} \to (-1{,}1)$$

sind alle stetig, strikt monoton wachsend und damit bijektiv. Ihre Umkehrfunktionen sind

- $\operatorname{Ar}\sinh\colon \mathbb{R} \to \mathbb{R}$ (*Area sinus hyperbolici*),
- $\operatorname{Ar}\cosh\colon [1,\infty) \to [0,\infty)$ (*Area cosinus hyperbolici*) und
- $\operatorname{Ar}\tanh\colon (-1,1) \to \mathbb{R}$ (*Area tangens hyperbolici*).[9]

[8] *Stetige Fortsetzung* heißt, dass die neu definierte Funktion einen größeren Definitionsbereich hat, aber trotzdem in allen Punkten der originalen Funktion übereinstimmt, plus dass sie auch noch stetig ist.

[9] Es gilt genauer: $\operatorname{Ar}\sinh x = \log(x + \sqrt{x^2+1})$ und $\operatorname{Ar}\cosh x = \log(x + \sqrt{x^2-1})$.

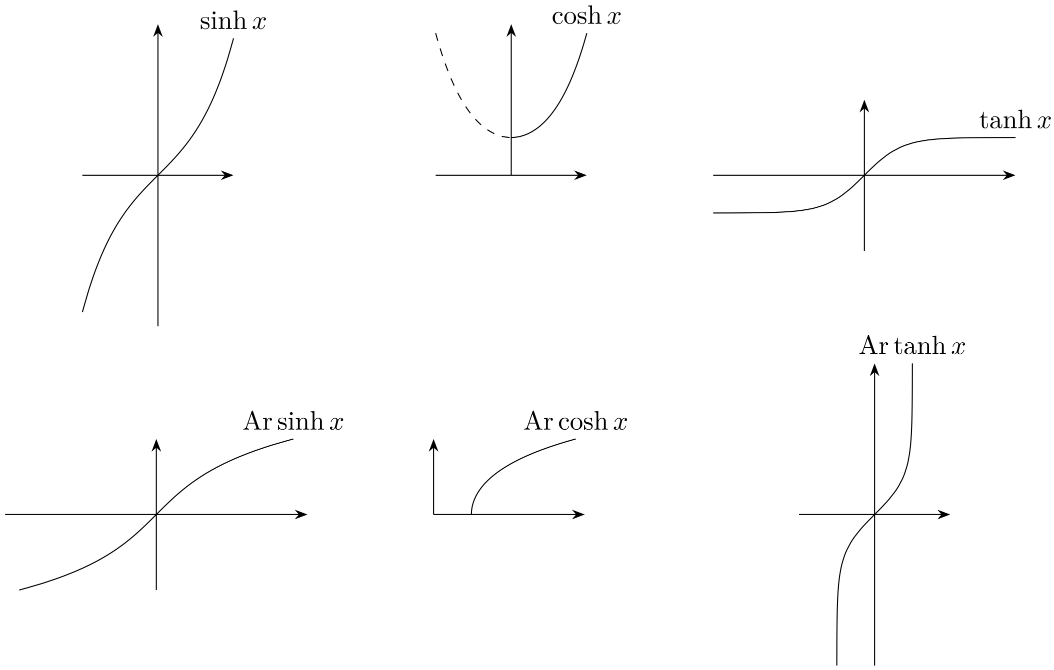


Abbildung 11.2: Hyperbolische Funktionen und deren Umkehrfunktionen.

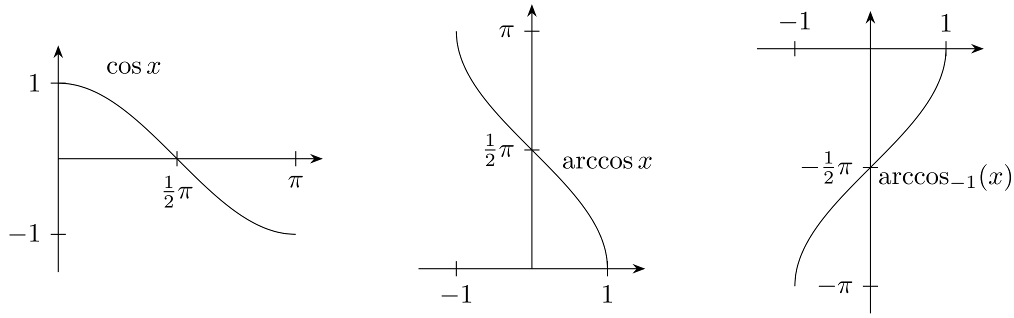
11.6 Umkehrfunktionen der trigonometrischen Funktionen

Satz 11.42 (Arkuskosinus)**.**

1. $\cos\colon [0,\pi] \to [-1,1]$ *ist stetig und strikt monoton fallend, also bijektiv. Die Umkehrfunktion*

$$\arccos\colon [-1,1] \to [0,\pi]$$

heißt Arcus cosinus *(oder genauer,* „Hauptzweig des Arcus cosinus“*).*



2. Für jedes $k \in \mathbb{Z}$ ist $\cos\colon [k\pi,(k+1)\pi] \to [-1,1]$ *bijektiv und die Umkehrfunktion*

$$\arccos_k\colon [-1,1] \to [k\pi,(k+1)\pi]$$

heißt der k-ter Nebenzweig des Arcus cosinus*.*

Beweis.

1. Der Kosinus ist bekanntlich stetig (Korollar 11.33). Zur Monotonie: Wir wissen bereits aus dem Beweis zu Korollar 11.34, dass cos auf $[0,2]$ streng monoton fällt, also insbesondere auch auf $[0,\frac{1}{2}\pi]$. Mittels Phasenverschiebung gilt $\cos(\pi - x) = -\cos(-x) = -\cos x$, also gilt für $\pi - x < \pi - y$ mit $x, y \in [0,\frac{1}{2}\pi]$

$$x > y \implies \cos x < \cos y \implies -\cos x > -\cos y \implies \cos(\pi - x) > \cos(\pi - y),$$

weshalb cos auch auf $[\frac{1}{2}\pi,\pi]$ monoton fallend ist.[10]

[10]Der Dozent gab hier einen anderen Beweis: Für $x, y \in [0,\pi]$ mit $x < y$ gilt $\cos y - \cos x = -2\sin(\frac{1}{2}(y+x))\sin(\frac{1}{2}(y-x))$. Nun sind die beiden Sinuswerte positiv, also der gesamte Ausdruck negativ. Damit fällt cos strikt monoton auf $[0,\pi]$.

2. Mittels Phasenverschiebung gilt $\cos(x + k\pi) = \pm \cos x$ für alle $x \in [0, \pi]$. Da unter Multiplikation mit -1 die Funktion nach wie vor stetig und strikt monoton ist (dann aber strickt monoton wachsend), ist auch cos auf $[k\pi, (k+1)\pi]$ bijektiv. $\square$

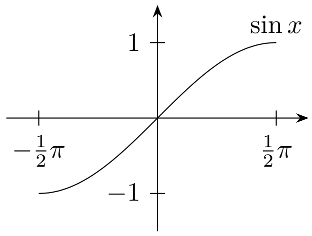
Ende der Vorlesung 19 am 20. Dezember 2021

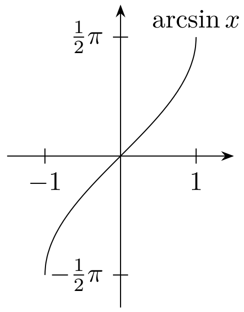
Korollar 11.43 (Arkussinus)**.** $\sin\colon [-\frac{1}{2}\pi, \frac{1}{2}\pi]$ *ist stetig und strikt monoton, also bijektiv. Die Umkehrfunktion heißt* Arcus sinus*:*

$$\arcsin\colon [-1,1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Für jedes $k \in \mathbb{Z}$ gibt es den k-ten Nebenzweig

$$\arcsin_k\colon [-1,1] \to \left[\left(k - \frac{1}{2}\right)\pi, \left(k + \frac{1}{2}\right)\pi\right].$$



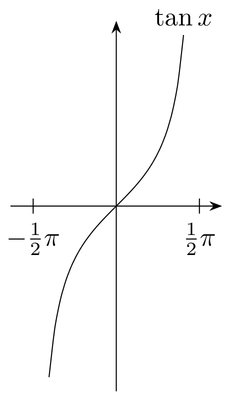


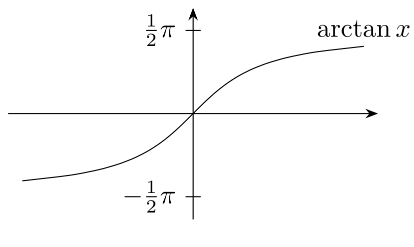
Korollar 11.44. $\tan\colon (-\frac{1}{2}\pi, \frac{1}{2}\pi) \to \mathbb{R}$ *ist stetig und strikt monoton, also bijektiv. Die Umkehrfunktion heißt* Arcus tangens*:*

$$\arctan\colon \mathbb{R} \to \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Für jedes $k \in \mathbb{Z}$ gibt es den k-ten Nebenzweig

$$\arctan_k\colon \mathbb{R} \to \left(\left(k - \frac{1}{2}\right)\pi, \left(k + \frac{1}{2}\right)\pi\right).$$





Korollar 11.45 (Eindeutigkeit der Polardarstellung)**.** *Für jedes $z \in \mathbb{C} \setminus \{0\}$ gibt es eindeutig bestimmte $r \in (0, \infty)$ und $\varphi \in [0, 2\pi)$, sodass $z = r\exp(\mathrm{i}\varphi)$ gilt.*

Beweis. Sei $z = x + \mathrm{i}y$. Da z eindeutig durch x und y bestimmt ist und die Wurzelfunktion bijektiv ist, ist $r := |z| = \sqrt{x^2 + y^2}$ eindeutig bestimmt.

Das Problem an dem Beweis ist, dass wir noch nicht gezeigt haben, dass sin auf $[0, \pi]$ positiv ist. Das über z. B. Restgliedabschätzung zu zeigen, ist meiner Meinung nach schwer.

Zudem gilt $z = r\cos\varphi + \mathrm{i}r\sin\varphi$ mit $r \neq 0$. Vergleichen wir Real- und Imaginärteil, so gilt $x/r = \cos\varphi$ und $y/r = \sin\varphi$. Da cos nicht auf dem gesamten Intervall $[0, 2\pi)$ bijektiv ist, müssen wir eine Fallunterscheidung bzgl. y bzw. $y/r = \sin\varphi$ durchführen:

$$\varphi := \begin{cases} \arccos(x/r) \in [0,\pi] & \text{falls } y \geq 0, \\ \arccos_1(x/r) \in (\pi, 2\pi) & \text{falls } y < 0. \end{cases}$$

Weil arccos und $\arccos_1$ bijektiv sind, ist auch φ eindeutig bestimmt. □

Bemerkung 11.46. Insbesondere folgt für $x, y > 0$: $\varphi := \arctan(y/x)$.

12 Differenziation

12.1 Differenzialquotient

Definition 12.1 (Differenzierbarkeit, Ableitung)**.** Sei $D \subset \mathbb{R}$ ein nichtleeres Intervall und $f\colon D \to \mathbb{R}$ eine Funktion. Dann heißt f *differenzierbar im Punkt* $x_0 \in D$, falls der Grenzwert

$$\lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{f(x) - f(x_0)}{x - x_0}$$

in $\mathbb{R}$ existiert. Dieser Grenzwert heißt dann *Ableitung* oder *Differenzialquotient von* f *im Punkt* x_0.

Bezeichnung 12.2. Wir schreiben

$$f'(x_0) := \left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=x_0} := \lim_{\substack{h \to 0 \\ h \neq 0 \\ x_0 + h \in D}} \frac{f(x_0 + h) - f(x_0)}{h} := \lim_{\substack{x \to x_0 \\ x \neq x_0 \\ x \in D}} \frac{f(x) - f(x_0)}{x - x_0}.$$

Da uns klar ist, dass die Folgenglieder aus D kommen und der Nenner nicht 0 werden kann, können wir einige Bedingungen unter lim weglassen:

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

12.1.1 Geometrische Interpretation

Der *Differenzenquotient*

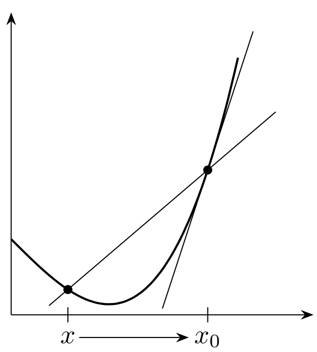
$$\frac{f(x) - f(x_0)}{x - x_0}$$

gibt die Steigung der *Sekante* durch die Punkte $(x, f(x))$ und $(x_0, f(x_0))$ an.

Der *Differenzialquotient*

$$\left.\frac{\mathrm{d}f(x)}{\mathrm{d}x}\right|_{x=x_0} = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

(falls er existiert) gibt die Steigung der *Tangente* im Punkt $(x_0, f(x_0))$ an.



12.1.2 Physikalische Interpretation

Falls $f(x)$ den Ort beschreibt, der zum Zeitpunkt x eingenommen wird, so ist $f'(x)$ die *Momentangeschwindigkeit* zum Zeitpunkt x, während $(f(x+h)-f(x))/h$ die *mittlere Geschwindigkeit* im Zeitintervall $(x, x+h)$ angibt.

12.2 Beispiele

Proposition 12.3. *Für $f(x) = c \in \mathbb{R}$ auf $\mathbb{R}$ ist $f'(x) = 0$.*

Proposition 12.4. *Für $f(x) = x^n$ auf $\mathbb{R}$ mit einem $n \in \mathbb{N}$ ist $f'(x) = nx^{n-1}$.*

Beweis. Es gilt mit dem binomischen Lehrsatz

$$\begin{aligned}\frac{(x+h)^n - x^n}{h} &= \frac{1}{h}\left(\binom{n}{1}x^{n-1}h + \binom{n}{2}x^{n-2}h^2 + \cdots + \binom{n}{n}h^n\right)\\ &= nx^{n-1} + \binom{n}{2}x^{n-2}h + \cdots + \binom{n}{n}h^{n-1} \xrightarrow{h\to 0} nx^{n-1}.\end{aligned}$$ □

Proposition 12.5. *Für $f(x) = 1/x$ auf $D = \mathbb{R}\setminus\{0\}$ ist $f'(x) = -1/x^2$.*

Beweis. Es gilt

$$\frac{1}{h}\left(\frac{1}{x+h} - \frac{1}{x}\right) = \frac{1}{h}\cdot\frac{x-(x+h)}{(x+h)x} = -\frac{1}{(x+h)x} \xrightarrow{h\to 0} -\frac{1}{x^2}.$$ □

Proposition 12.6. *Für $f(x) = \mathrm{e}^{cx}$ auf $\mathbb{R}$ mit $c \in \mathbb{R}$ ist $f'(x) = c\mathrm{e}^{cx}$.*

Beweis. Aus der Restgliedabschätzung der Exponenzialfunktion folgt

$$|\exp(y) - (1+y)| \le 2\frac{|y|^2}{2!} = |y|^2 \implies \left|\frac{\exp(y)-1}{y} - 1\right| \le |y|.$$

Für $y \to 0$ konvergiert also $(\mathrm{e}^y - 1)/y \to 1$ nach dem Sandwichlemma. Damit gilt

$$\frac{\mathrm{e}^{c(x+h)} - \mathrm{e}^{cx}}{h} = \mathrm{e}^{cx}\frac{\mathrm{e}^{ch}-1}{h} = c\mathrm{e}^{cx}\frac{\mathrm{e}^{ch}-1}{ch} \xrightarrow{h\to 0} c\mathrm{e}^{cx}.$$ □

Proposition 12.7. *Für $f(x) = a^x$ auf $\mathbb{R}$ mit $a \in \mathbb{R}_+^*$ ist $f'(x) = \log(a)a^x$.*

Beweis. Das folgt aus $a^x = \mathrm{e}^{\log(a)x}$ und der Ableitung der Exponenzialfunktion. □

Proposition 12.8. *Es gilt $\sin'(x) = \cos x$ und $\cos'(x) = -\sin x$.*

Beweis. Aus der Restgliedabschätzung für den Sinus erhalten wir für $0 \le h \le 2$

$$h - \frac{h^3}{6} \le \sin h \le h \iff 1 - \frac{h^2}{6} \le \frac{\sin h}{h} \le 1.$$

Da $h \to 0$, muss irgendwann $|h| \le 2$ sein (für negative h drehen wir die Relationszeichen einfach um, was aber die Konvergenz nicht ändert). Damit folgt aus dem Sandwichlemma $(\sin h)/h \to 1$ für $h \to \infty$.

Ähnlich folgern wir für den Kosinus

$$1 - \frac{h^2}{2} \le \cos h \le 1 \iff -\frac{h}{2} \le \frac{\cos h - 1}{h} \le 0 \implies \frac{\cos h - 1}{h} \xrightarrow{h\to 0} 0.$$

Damit erhalten wir aus den Additionstheoremen

$$\begin{aligned}\sin'(x) &= \lim_{h\to 0}\frac{\sin(x+h) - \sin x}{h} = \lim_{h\to 0}\frac{\sin x\cos h + \cos x \sin h - \sin x}{h}\\ &= \sin x \underbrace{\lim_{h\to 0}\frac{\cos h - 1}{h}}_{\to 0} + \cos x \underbrace{\lim_{h\to 0}\frac{\sin h}{h}}_{\to 1} = \cos x.\end{aligned}$$

Analog erhalten wir

$$\cos'(x) = \lim_{h\to 0}\frac{\cos x(\cos h - 1) - \sin x \sin h}{h} = -\sin x.$$ □

Proposition 12.9. *Es gilt* $\sinh'(x) = \cosh x$ *und* $\cosh'(x) = \sinh x$.

Beweis. Aus der Definition und der Ableitung der Exponenzialfunktion folgt

$$\begin{aligned}\sinh'(x) &= \lim_{h\to 0} \frac{1}{h}\left(\frac{\mathrm{e}^{x+h} + \mathrm{e}^{-(x+h)}}{2} - \frac{\mathrm{e}^{x} + \mathrm{e}^{-x}}{2}\right) = \frac{1}{2}\lim_{h\to 0} \frac{\mathrm{e}^{x+h} - \mathrm{e}^{x}}{h} + \frac{1}{2}\lim_{h\to 0}\frac{\mathrm{e}^{-(x+h)} - \mathrm{e}^{-x}}{h} \\ &= \frac{\mathrm{e}^{x}}{2} + \frac{-\mathrm{e}^{-x}}{2} = \cosh x.\end{aligned}$$

Analog erhalten wir[1]

$$\cosh'(x) = \lim_{h\to 0} \frac{\mathrm{e}^{x+h} - \mathrm{e}^{-(x+h)} - \mathrm{e}^{x} + \mathrm{e}^{-x}}{2h} = \lim_{h\to 0}\frac{\mathrm{e}^{x+h} - \mathrm{e}^{x}}{2h} - \lim_{h\to 0}\frac{\mathrm{e}^{-(x+h)} - \mathrm{e}^{-x}}{2h} = \sinh x. \qquad \square$$

Ende der Vorlesung 20 am 22. Dezember 2021

12.3 Sätze zur Differenzierbarkeit und Ableitungsregeln

Satz 12.10. *Sei* $f\colon D \to \mathbb{R}$ *eine Funktion auf* $D \subset \mathbb{R}$ *und sei* $x_0 \in D$ *eine Stelle. Dann sind folgende Aussagen äquivalent:*

1. f *ist in* x_0 *differenzierbar.*
2. *Es existiert eine Funktion* $\varphi\colon D \to \mathbb{R}$*, die in* x_0 *stetig ist, mit*
$$f(x) = f(x_0) + (x - x_0)\varphi(x) \qquad \text{für alle } x \in D. \tag{12.11}$$
3. *Es existiert eine affine lineare Abbildung* $L\colon \mathbb{R} \to \mathbb{R}$ *(also eine Funktion der Form* $L(x) = ax + b$ *bzw.* $L(x) = a_0(x - x_0) + b_0$*) mit*
$$L(x_0) = f(x_0) \qquad \text{und} \qquad \lim_{x\to x_0} \frac{f(x) - L(x)}{x - x_0} = 0. \tag{12.12}$$

Dabei sind

$$\varphi(x) = \frac{f(x) - f(x_0)}{x - x_0} \qquad \text{und} \qquad L(x) = f(x_0) + f'(x_0)(x - x_0)$$

(wobei φ *in* x_0 *stetig fortgesetzt wird).*

Beweis.

- Aus Punkt 1 folgt Punkt 2: Aus der Differenzierbarkeit von f in x_0 können wir schließen, dass für den Differenzenquotient die Funktion
$$x \mapsto \frac{f(x) - f(x_0)}{x - x_0}$$
eine *stetige Fortsetzung* in x_0 besitzt (für $x = x_0$ ist nämlich der Bruch nicht definiert). Wir setzen
$$\varphi\colon D \to \mathbb{R}, \quad \varphi(x) \coloneqq \begin{cases} \dfrac{f(x) - f(x_0)}{x - x_0} & \text{falls } x \neq x_0, \\ \displaystyle\lim_{x'\to x_0} \frac{f(x') - f(x_0)}{x' - x_0} & \text{falls } x = x_0. \end{cases}$$
Per Festlegung ist φ stetig in x_0 und erfüllt (12.11) sowohl für alle $x \neq x_0$ als auch für $x = x_0$ (in dem Fall ist $(x - x_0)\varphi(x) = 0$).

[1] Der Dozent meinte, dass man den Beweis für die hyperbolischen Funktionen wie die trigonometrischen durchführen kann mit den zugehörigen Additionstheoremen. Ich sehe leider nicht, wie das gehen soll, da wir keine einfache Möglichkeit haben, $(\sinh x)/x \to 1$ und $(\cosh x - 1)/x \to 0$ zu zeigen.

- Aus Punkt 2 folgt Punkt 3: Sei $L\colon \mathbb{R} \to \mathbb{R}$, $L(x) = f(x_0) + f'(x_0)(x - x_0)$. Tatsächlich gilt $f'(x_0) = \varphi(x_0)$. Nun überprüfen wir (12.12). Einsetzen liefert

$$\frac{f(x) - L(x)}{x - x_0} = \frac{\varphi(x)(x - x_0) - \varphi(x_0)(x - x_0)}{x - x_0} = \varphi(x) - \varphi(x_0) \xrightarrow{x \to x_0} 0.$$

  Die Konvergenz gilt, weil φ in x_0 stetig ist und daher $\varphi(x) \to \varphi(x_0)$ für $x \to x_0$.

- Aus Punkt 3 folgt Punkt 1: Sei $L\colon \mathbb{R} \to \mathbb{R}$ mit $L(x) = ax + b = a(x - x_0) + b_0$. Aus $L(x_0) = f(x_0)$ folgt $b_0 = L(x_0) = f(x_0)$, also

$$0 = \lim_{x \to x_0} \frac{f(x) - L(x)}{x - x_0} = \lim_{x \to x_0} \frac{f(x) - a(x - x_0) - f(x_0)}{x - x_0} = -a + \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} \iff f'(x_0) = a.$$

  Folglich ist f in x_0 differenzierbar. □

Im Allgemeinen hängen φ und L vom gegebenen x_0 ab.

Bezeichnung 12.13 (lineare Approximation)**.** L heißt *lineare Approximation* von f in x_0.

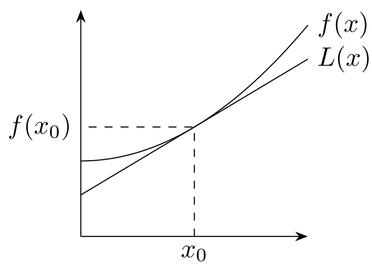


Abbildung 12.1: Lineare Approximation $L(x) = f(x_0) + (x - x_0)f'(x_0)$ von f in x_0.
Die Gerade stimmt mit f in x_0 überein und hat besitzt Anstieg $f'(x_0)$.

Korollar 12.14. *Ist f in x_0 differenzierbar, dann ist f auch stetig.*

Beweis. Das folgt direkt aus Satz 12.10, denn f ist eine Addition und Multiplikation stetiger Funktionen, insbesondere φ. □

Satz 12.15 (Ableitungsregeln Summe, Produkt, Quotient)**.** *Seien $f, g\colon D \to \mathbb{R}$ in x_0 differenzierbar. Dann sind $f + g$, fg und, falls $g(x_0) \neq 0$, auch f/g in x_0 differenzierbar. Dabei gelten*

1. $(f + g)'(x_0) = f'(x_0) + g'(x_0)$ (Linearität),
2. $(fg)'(x_0) = f(x_0)g'(x_0) + f'(x_0)g(x_0)$ (LEIBNIZSCHE Produktregel) *und*
3. $(f/g)'(x_0) = (f'(x_0)g(x_0) - f(x_0)g'(x_0))/g^2(x_0)$ (Quotientenregel).

Beweis. Die Existenz der Ableitungen ist ein Nebenprodukt, wenn wir zeigen, wie sich die Ableitungen berechnen.

1. Aus der Linearität des Limes folgt

$$\begin{aligned}(f + g)'(x_0) &= \lim_{h \to 0} \frac{(f(x_0 + h) + g(x_0 + h)) - (f(x_0) + g(x_0))}{h} \\ &= \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} + \lim_{h \to 0} \frac{g(x_0 + h) - g(x_0)}{h} \\ &= f'(x_0) + g'(x_0).\end{aligned}$$

2. Weil aufgrund der Differenzierbarkeit f auch in x_0 stetig ist, erhalten wir durch geschicktes Umformen

$$\begin{aligned}
(fg)'(x_0) &= \lim_{h\to 0} \frac{f(x_0+h)g(x_0+h) - f(x_0)g(x_0)}{h} \\
&= \lim_{h\to 0} \frac{f(x_0+h)g(x_0+h) - f(x_0+h)g(x_0) + f(x_0+h)g(x_0) - f(x_0)g(x_0)}{h} \\
&= \left(\lim_{h\to 0} f(x_0+h)\right)\left(\lim_{h\to 0} \frac{g(x_0+h)-g(x_0)}{h}\right) + \left(\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}\right) g(x_0) \\
&= f(x_0)g'(x_0) + f'(x_0)g(x_0).
\end{aligned}$$

3. Aus der Differenzierbarkeit in x_0 folgt auch die Stetigkeit von g, also $g(x_0+h) \to g(x_0)$ für $h \to 0$. Außerdem folgt aus der Stetigkeit und $g(x_0) \neq 0$, dass $g(x) \neq 0$ für alle x in einer hinreichend kleinen Umgebung um x_0 gilt. Darum sind die Brüche in der folgenden Berechnung für kleine h definiert:

$$\begin{aligned}
\left(\frac{f}{g}\right)'(x_0) &= \lim_{h\to 0} \frac{f(x_0+h)g(x_0) - f(x_0)g(x_0+h)}{g(x_0+h)g(x_0)h} \\
&= \lim_{h\to 0} \frac{1}{g(x_0+h)g(x_0)} \left(\frac{f(x_0+h)-f(x_0)}{h} g(x_0) - \frac{g(x_0+h)-g(x_0)}{h} f(x_0) \right) \\
&= \frac{1}{g^2(x_0)} (f'(x_0)g(x_0) - g'(x_0)f(x_0)).
\end{aligned}$$ □

Korollar 12.16. *Es gelten für alle $x \in \mathbb{R}$, sodass die Nenner nicht Null werden,*

$$\tan' x = \frac{1}{\cos^2 x} \qquad \textit{und} \qquad \cot' x = -\frac{1}{\sin^2 x}.$$

Beweis. Aus der Quotientenregel und $\tan x = (\sin x)/(\cos x)$ folgen

$$\tan' x = \frac{\cos x \sin' x - \sin x \cos' x}{\cos^2 x} = \frac{\cos^2 x + \sin^2 x}{\cos^2 x} = \frac{1}{\cos^2 x}.$$

$\cot' x$ ergibt sich analog. □

Satz 12.17 (Kettenregel)**.** *Seien $f\colon D \to E$ und $g\colon E \to \mathbb{R}$ Funktionen, und sei f differenzierbar in $x_0 \in D$ und g differenzierbar in $y_0 = f(x_0) \in E$. Dann folgt, dass auch $g \circ f\colon D \to \mathbb{R}$ in x_0 differenzierbar ist und es gilt:*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Beweis. Aus Satz 12.10 folgt, dass es zu gegebenen x_0 und $y_0 = f(x_0)$ zwei Funktionen $F\colon D \to E$ und $G\colon E \to \mathbb{R}$ existieren, die in x_0 bzw. y_9 stetig sind und

$$f(x) - f(x_0) = (x - x_0)F(x_0) \qquad \text{für alle } x \in D \tag{12.18}$$

sowie

$$g(y) - g(y_0) = (y - y_0)G(y_0) \qquad \text{für alle } y \in E \tag{12.19}$$

erfüllen. Dabei sind $F(x_0) = f'(x_0)$ und $G(y_0) = g'(y_0)$. Setzen wir $y = f(x)$ und $y_0 = f(x_0)$ in (12.19) ein, erhalten wir

$$(g \circ f)(x) - (g \circ f)(x_0) \overset{(12.19)}{=} (f(x) - f(x_0))G(f(x)) \overset{(12.18)}{=} (x - x_0)F(x)G(f(x)).$$

Nun ist die Abbildung $\varphi(x) \coloneqq F(x)G(f(x))$ stetig in x_0. Daraus und mit der eben gezeigten Gleichung folgt, dass die Abbildung $g \circ f$ differenzierbar in x_0 ist und es gilt

$$(g \circ f)'(x_0) = \varphi(x_0) = F(x_0)G(f(x_0)) = f'(x_0)g'(f(x_0)).$$ □

Beispiel 12.20. Es gilt $(e^f)'(x) = e^{f(x)}f'(x)$.

Satz 12.21 (Differenzierbarkeit der Umkehrfunktion)**.** *Sei $D \subset \mathbb{R}$ ein nichttriviales Interval. Sei $f\colon D \to E$ eine Funktion, die surjektiv auf $E \subset \mathbb{R}$, streng monoton und differenzierbar in $x_0 \in D$ ist mit $f'(x_0) \neq 0$. Dann ist die Umkehrfunktion $g := f^{-1}\colon E \to D$ in $y_0 = f(x_0)$ differenzierbar mit*

$$g'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(g(y_0))}.$$

Bemerkung 12.22. Wir können uns den Satz wie folgt merken: Wenn wir wissen, dass g differenzierbar in y_0 ist, so folgt aus $(f \circ g)(y) = y$ und der Kettenregel $f'(g(y)) \cdot g'(y) = 1$.

Das ist jedoch kein Beweis, da wir die Differenzierbarkeit voraussetzen, die wir für den Satz noch zeigen müssen.

Beweis. Weil f differenzierbar in x_0 ist, gibt es gemäß Satz 12.10 eine Funktion $F\colon D \to E$, die stetig in x_0 ist mit

$$f(x) - f(x_0) = (x - x_0)F(x) \qquad \text{für alle } x \in D. \tag{12.23}$$

Nun ist nach 12.10

$$F(x) = \begin{cases} \dfrac{f(x) - f(x_0)}{x - x_0} & \text{falls } x \neq x_0, \\ f'(x_0) & \text{falls } x = x_0. \end{cases}$$

Es gilt $F(x_0) = f'(x_0) \neq 0$ nach Voraussetzung. Für alle anderen $x \neq x_0$ gilt auch $F(x) \neq 0$, da f streng monoton ist und der Zähler $f(x) - f(x_0) \neq 0$ im Differenzenquotienten ist.

Setzen wir $y = f(x)$ und $g(y) = x$ in (12.23) ein, ergibt sich

$$y - y_0 = (g(y) - g(y_0))F(g(y)) \iff g(y) - g(y_0) = (y - y_0)G(y),$$

wobei $G(y) := 1/F(g(y))$ wohldefiniert und stetig in x_0 ist. Mit Satz 12.10 folgt daher, dass g differenzierbar in y_0 ist mit

$$g'(y_0) = G(y_0) = \frac{1}{F(g(y_0))} = \frac{1}{f'(g(y_0))}.$$ □

Korollar 12.24. *Es gilt* $\log'(x) = 1/x$ *auf* $D = (0, \infty)$.

Beweis. Sei $f = \log$. Für $g = \exp$ gilt $f = g^{-1}$ und daher

$$f'(x) = \frac{1}{g'(f(x))} = \frac{1}{\exp(\log(x))} = \frac{1}{x}.$$ □

Proposition 12.25. *Für $x \in (-1, 1)$ gilt* $\arcsin' x = 1/\sqrt{1 - x^2}$.[2]

Beweis. Aus dem trigonometrischen Pythagoras folgt

$$\cos x = \sqrt{1 - \sin^2 x}$$

und daher

$$\arcsin' x = \frac{1}{\sin'(\arcsin x)} = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1 - x^2}}.$$ □

Proposition 12.26. *Es gilt* $\arctan' x = 1/(1 + x^2)$ *für alle* $x \in \mathbb{R}$.

Beweis. Es gilt

$$\tan' y = \frac{1}{\cos^2 y} = \frac{\sin^2 y + \cos^2 y}{\cos^2 y} = 1 + \tan^2 y.$$

Daraus folgt

$$\arctan' x = \frac{1}{\tan'(\arctan x)} = \frac{1}{1 + x^2}.$$ □

Ende der Vorlesung 21 am 10. Januar 2022

[2] Es gilt $\arccos' x = -1/\sqrt{1 - x^2} = -\arcsin' x$.

12.4 Höhere Ableitungen

Definition 12.27 (höhere Ableitung)**.** Gegeben sei eine Funktion $f\colon D \to \mathbb{R}$, die in D differenzierbar ist. Wir betrachten $f'\colon D \to \mathbb{R}$. Falls f' im Punkt $x_0 \in D$ differenzierbar ist, so heißt die Ableitung von f' *zweite Ableitung* von f:

$$f''(x_0) = (f')'(x_0) = \left.\frac{\mathrm{d}^2}{\mathrm{d}x^2} f(x)\right|_{x=x_0}.$$

Allgemein ist die *k-te Ableitung* ($k \in \mathbb{N}_0$) rekursiv definiert als

$$f^{(0)}(x) = f(x), \qquad f^{(k)}(x) = \left(f^{(k-1)}\right)'(x) = \frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{\mathrm{d}^{k-1}}{\mathrm{d}x^{k-1}} f\right)(x);$$

vorausgesetzt f ist bereits $(k-1)$-mal differenzierbar gewesen und die Ableitung von $f^{(k-1)}$ existiert.

Achtung: Die Ableitung ist selbst wieder eine Funktion. Daher muss in der Notation

$$\frac{\mathrm{d}^k f(x)}{\mathrm{d}x^k}$$

das x im Zähler und im Nenner *dasselbe* sein. Wünschen wir eine Auswertung an $x = x_0$, schreiben wir

$$\left.\frac{\mathrm{d}^k f(x)}{\mathrm{d}x^k}\right|_{x=x_0}.$$

In dem Fall ist der Ausdruck eine Zahl.

Definition 12.28 (beliebig oft differenzierbar)**.** Eine Funktion heißt *unendlich* oder *beliebig oft differenzierbar*, falls die Funktion k-mal differenzierbar ist für alle $k \in \mathbb{N}$.

Beispiel 12.29. Sei f ein Polynom in x mit Grad n. Dann ist $f^{(k)} \equiv 0$ die Nullabbildung für alle $k > n$.

Definition 12.30 (stetige Differnzierbarkeit)**.** Sei $k \in \mathbb{N}$. Eine Funktion f heißt *k-mal stetig differenzierbar*, falls die Funktion k-mal differenzierbar und $f^{(k)}$ stetig ist.

Das folgende Beispiel zeigt, warum Differenzierbarkeit nicht stetige Differenzierbarkeit impliziert.

Beispiel 12.31. Wir definieren

$$f\colon \mathbb{R} \to \mathbb{R}, \quad f(x) = \begin{cases} x^2 \sin\left(\dfrac{1}{x}\right) & \text{falls } x \neq 0, \\ 0 & \text{falls } x = 0. \end{cases}$$

f ist auf ganz $\mathbb{R}$ differenzierbar. Für $x \neq 0$ folgt direkt aus den Ableitungsregeln

$$f'(x) = 2x\sin\left(\frac{1}{x}\right) - x^2\cos\left(\frac{1}{x}\right)\frac{1}{x^2} = 2x\sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right). \tag{12.32}$$

Für $x = 0$ wenden wir den Differenzenquotienten an:

$$\begin{aligned} 0 = \liminf_{h\to 0} \frac{-h^2}{h} \leq \liminf_{h\to 0} \frac{f(h)-f(0)}{h} \leq \limsup_{h\to 0} \frac{f(h)-f(0)}{h} \leq \liminf_{h\to 0} \frac{h^2}{h} = 0 \\ \implies \lim_{h\to 0} \frac{f(h)-f(0)}{h} = 0, \end{aligned}$$

wobei wir $-1 \leq \sin(1/x) \leq 1$ verwendeten.

Jedoch ist f' nicht in 0 stetig. Es gilt

$$-2x \leq 2x\sin\left(\frac{1}{x}\right) \leq 2x \xrightarrow{x\to 0} 0,$$

aber $1/x \to \infty$ für $x \to 0$. Darum oszilliert $\cos(1/x)$ in $[-1, 1]$, weshalb $f'(x)$ gemäß (12.32) nicht konvergiert.

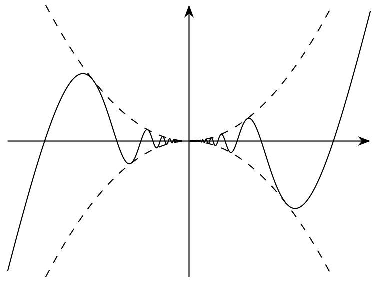


Abbildung 12.2: $f(x) = x^2 \sin(1/x)$. Die y-Achse wurde um den Faktor 3 skaliert.

Wir hätten auch andere Gegenbeispiele konstruieren können, wie z. B. $x^2 \sin(1/x^2)$. Wichtig ist, dass für $x \to 0$ der Kosinus oszilliert, es also keinen Faktor x oder so gibt, sondern eher $1/x$ und dergleichen.

Definition 12.33 (Vektorraum der stetig differenzierbaren Funktionen)**.** Wir betrachten alle Funktionen $f\colon D \to \mathbb{R}$ und definieren:

- $\mathcal{C}(D) := \mathcal{C}^0(D)$ der Vektorraum aller stetigen f.
- $\mathcal{C}^k(D)$ der Vektorraum aller k-mal stetig differenzierbaren f.
- $\mathcal{C}^\infty(D)$ der Vektorraum aller beliebig oft stetig differenzierbaren f.

Aufgrund der Linearität der Differenzierbarkeit, die aus der Linearität der Grenzwertbildung folgt, bilden die Mengen tatsächlich Vektorräume.

Beispiel 12.34. Die Klasse $\mathcal{C}^\infty$ enthält mehr als nur die Menge aller Polynome. Wir definieren

$$f\colon \mathbb{R} \to \mathbb{R}, \quad f(x) = \begin{cases} \exp\left(-\dfrac{1}{x}\right) & \text{falls } x > 0, \\ 0 & \text{falls } x \leq 0. \end{cases}$$

Wir wollen zeigen, dass f in $\mathcal{C}^\infty$ liegt und $f^{(k)}(0) = 0$ für alle $k \in \mathbb{N}$. Dann ist nämlich f kein Polynom, denn für ein Polynom p gilt $p^{(k)}(0) = k!a_k$ für alle $k \in \mathbb{N}$. Dafür müsste dann alle Koeffizienten Null sein.

Das erreichen wir, indem wir zeigen, dass es für jedes $k \in \mathbb{N}$ ein Polynom $p_k\colon \mathbb{R} \to \mathbb{R}$ gibt mit

$$f^{(k)}(x) = \begin{cases} p_k\left(\dfrac{1}{x}\right) \exp\left(-\dfrac{1}{x}\right) & \text{falls } x > 0, \\ 0 & \text{falls } x \leq 0. \end{cases}$$

Wir machen es induktiv.

Zum Induktionsanfang $k = 0$: Für $k = 0$ mit $p_0\colon x \mapsto 1$ ist die Aussage offensichtlich richtig.

Zum Induktionsschritt $k \to k+1$: Für $x > 0$ werden die Folgenglieder zur Bestimmung des Differenzialquotienten irgendwann alle positiv. Deshalb können wir den positiven Zweig verwenden und erhalten per Produktregel

$$f^{(k+1)}(x) = \underbrace{\left(p_k'\left(\frac{1}{x}\right)\frac{-1}{x^2} + p_k\left(\frac{1}{x}\right)\frac{1}{x^2}\right)}_{:= p_{k+1}(1/x)} \exp\left(-\frac{1}{x}\right) = p_{k+1}\left(\frac{1}{x}\right) \exp\left(-\frac{1}{x}\right).$$

Für $x < 0$ werden die Folgenglieder zur Bestimmung des Differenzialquotienten irgendwann alle negativ. Damit gilt $f^{(k+1)}(x) = 0$.

Für $x = 0$ gelten

$$f^{(k+1)}(0) = \lim_{\substack{h \to 0 \\ h < 0}} \frac{f^{(k)}(h) - f^{(k)}(0)}{h} = 0$$

oder

$$f^{(k+1)}(0) = \lim_{\substack{h \to 0 \\ h > 0}} \frac{f^{(k)}(h) - f^{(k)}(0)}{h} = \lim_{\substack{h \to 0 \\ h > 0}} \frac{1}{h} p_k\left(\frac{1}{h}\right) \exp\left(-\frac{1}{h}\right) = \lim_{t \to \infty} \frac{t p_k(t)}{\exp(t)} = 0$$

mit $t = 1/h$, da die Exponenzialfunktion schneller als jedes Polynom wächst.

12.5 Lokale Extrema

Definition 12.35 (lokales Minimum/Maximum/Extremum)**.** Eine Funktion $f\colon D \to \mathbb{R}$ hat in $x_0 \in D$ ein *lokales Minimum*, falls es ein $\varepsilon > 0$ gibt, sodass x_0 das Minimum in $f|_{D\cap[x_0-\varepsilon,x_0+\varepsilon]}$ ist.[3,4]

Analog definieren wir ein *lokales Maximum*. Ein *lokales Extremum* x_0 von f ist entweder ein lokales Maximum oder ein lokales Minimum.

Satz 12.36. *Besitzt eine Funktion $f\colon (a,b) \to \mathbb{R}$ ein lokales Extremum $x_0 \in (a,b)$ und ist dort differenzierbar, gilt dann $f'(x_0) = 0$.*

Bemerkung 12.37.

1. Die Umkehrung gilt <u>nicht</u>. Bspw. ist für $f(x) = x^3$ die Ableitung $f'(0) = 0$, aber f besitzt keine Extrema.
2. Die Aussage ist auf abgeschlossenen Intervallen auch falsch, also für $f\colon [a,b] \to \mathbb{R}$. Bspw. sei $f(x) = x$ auf [0,1]. Nun besitzt f lokale Extrema in $x \in \{0,1\}$, aber es gilt $f'(x) = 1$ für alle $x \in [0,1]$.

Beweis. Angenommen f hat ein lokales Minimum in x_0, d. h. es gibt ein $\varepsilon > 0$ mit $[x_0-\varepsilon, x_0+\varepsilon] \subset (a,b)$, sodass

$$f(x_0) = \min_{|y-x_0|\le\varepsilon} f(y).$$

Daraus folgt für den Differenzenquotienten mit $h \neq 0$

$$\frac{1}{h}\underbrace{(f(x_0+h) - f(x_0))}_{>0} \begin{cases} > 0 & \text{falls } h > 0, \\ < 0 & \text{falls } h < 0, \end{cases}$$

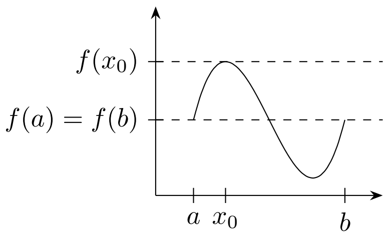
$$\implies \liminf_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h} \le 0 \le \limsup_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h}.$$

Weil f in x_0 differenzierbar ist, existiert der Grenzwert des Differenzenquotienten für $h \to 0$, also

$$f'(x_0) = \liminf_{h\to 0} \frac{f(x_0+h) - f(x_0)}{h} = \limsup h \to 0 \frac{f(x_0+h) - f(x_0)}{h} = 0. \qquad \square$$

Satz 12.38 (Rolle)**.** *Sei $f\colon [a,b] \to \mathbb{R}$ eine stetige Funktion mit $f(a) = f(b)$, und sei f differenzierbar auf (a,b). Dann gibt es ein $x_0 \in (a,b)$ mit $f'(x_0) = 0$.*

Insbesondere liegt zischen je zwei Nullstellen von f eine Nullstelle von f'.



Beweis. Für konstante f ist der Satz trivial (es gilt dann $f'(x) = 0$ für alle $x \in (a,b)$).

Wir nehmen nun an, dass f nicht konstant ist. Aufgrund der Stetigkeit auf dem kompakten Intervall $[a,b]$ gibt es nach dem Extremwertsatz ein $x_0 \in [a,b]$, sodass $f(x_0)$ ein Extremum von f ist. Zudem gilt $x_0 \notin \{a,b\}$, denn andernfalls wäre $f(x_0) = f(a) = f(b)$, also wären alle Extremwerte gleich $f(a)$ und f damit konstant, was der Annahme widerspricht. Folglich ist $x_0 \in (a,b)$. Aus Satz 12.36 folgt $f'(x_0) = 0$. $\square$

Ende der Vorlesung 22 am 12. Januar 2022

[3] $f|_M$ ist die Einschränkung von f auf die Menge M.

[4] *Alternativ*: x_0 ist ein *lokales Minimum* von $f\colon D \to \mathbb{R}$, wenn es ein $\varepsilon > 0$ gibt, sodass $f(x) \ge f(x_0)$ für alle $x \in D$ mit $|x - x_0| < \varepsilon$ gilt.

A Trickkiste

Wie löse ich eine Aufgabe? Das sollte für Mathematikstudenten eine sehr wichtige Frage sein. Es gibt eigentlich zwei Antworten:

- Es gibt generelle Problemlösestrategien, die auf jeder Aufgabe anwendbar sind. Im Englischen werden Leute, die gut darin sind, als *problem solver* bezeichnet.

  Ich fasse mal ein paar dieser Strategien in Fragen zusammen:

  - Wo habe ich das schon einmal gesehen? Woran erinnert mich die Aufgabe?
  - Kann ich eine Analogie machen oder eine Veranschaulichung zeichnen?
  - Was ist gegeben? Was kann ich aus dem Gegebenen schlussfolgern?
  - Was ist gesucht? Welche Aussagen muss ich davor zeigen, um zum gesuchten Ergebnis zu kommen? (Diese und die vorherige Frage komplementieren sich, sodass man durch wiederholtes Vorwärts- und Rückwärtsarbeiten womöglich beide Enden in der Mitte zu einer Lösung zusammenführen.)
  - Kann ich mir ein konkretes Beispiel/einen konkreten Fall anschauen und die Aufgabe dafür lösen (*toy problems*)? Was beobachte ich? Kann ich den Beweis verallgemeinern?
  - Kann ich eine Lösung/ein Beispiel raten? Welche Heuristiken könnte ich verwenden, um die Lösungssuche zu vereinfachen (das ist die Definition einer Heuristik)? Wie müsste eine Lösung wahrscheinlich aussehen?
  - Und noch weitere Strategien. An dieser Stelle ist George Pólyas *Schule des Denkens: Vom lösen mathematischer Probleme („How to solve it“).* zu empfehlen.

- Es gibt gewisse mathematische Tricks (sie wirken wie Zauberei, sind aber trotzdem mathematisch korrekt), die in spezifischen Aufgaben oder Aufgabentypen Anwendung finden. Oft sind sie die Idee, die man haben muss, um auf die Lösung zu kommen, und nur manchmal gibt es alternative *brute force*-Lösungswege.

  Was man sich beim anschauen der Musterlösung oder Lösung anderer fragen sollte ist: Was war der Trick? Wie hätte ich darauf kommen können? Auf welche Arten von Aufgaben kann ich den Trick anwenden? Worauf müsste ich achten, um mich an diesen Trick zu erinnern?

  Diese Sammlung von Tricks habe ich immer meine *mathematische Trickkiste* genannt – andere nennen sie vielleicht *Werkzeugkasten* oder *dreckige Tricks*, aber ich bleibe bei der Trickkiste. Diese Tricks sind oftmals nicht spezifisch für einen mathematischen Fachbereich, sondern können auch in anderen Bereichen Anwendung finden.

Die folgende Zusammenstellung ist (wie es sich für einen aktiv genutzten Werkzeugkasten gehört) nicht sehr strukturiert. Und leider ist es so, dass man erst ein Werkzeug richtig versteht, wenn man es im Kontext nutzt und sieht, also *üben, üben, üben!*

A.1 Logik und Mengenlehre

A.2 Sonstiges

- Es ist extrem wichtig, sicher in Termumformungen zu sein. Dazu gehören Ausmultiplizieren, Gleichungen und Ungleichungen äquivalent umformen, die binomischen Formeln und der binomische Lehrsatz, Potenz- und Wurzelgesetze, Rechnen mit Brüchen, etc. pp.
- Ausklammern und Faktorisieren ist auch sehr nützlich, aber ungemein schwerer.
  - Man kennt vielleicht einfach die Faktorisierung. Dazu sollten gehören:
    * die binomischen Formeln (auch $x^2 \pm 2x + 1 = (x \pm 1)^2$ und $x^2 - 1 = (x+1)(x-1)$),
    * der binomische Lehrsatz, insbesondere auch $(x+y)^3 = x^3 + 3x^2y + 3xy^2 + y^3$,
    * $x^3 - y^3 = (x^2 + xy + y^2)$,
    * $x^n - y^n = (x-y)(x^{n-1} + x^{n-2}y + x^{n-3}y^2 + \cdots + y^{n-1})$,
    * $x^n + y^n = (x+y)(x^{n-1} - x^{n-2}y + x^{n-3}y^2 - x^{n-4}y^3 + \cdots + y^{n-1})$ für ungerade n.
  - Man kann und sollte manchmal auch raten.

    - Z. B. indem man sich den Grad des zu faktorisierenden Terms anschaut, und dann überlegt, welchen Grad die Faktoren haben müssen. Bspw. $x^3 - y^3 = (x - y)(x^2 + xy + y^2)$; links steht Grad 3, rechts zwei Faktoren mit Grad 1 und Grad 2.
    - Ein nützlicher Trick ist das Raten von Nullstellen x_0 von Polynomen in x, sodass dieser als Faktor $x - x_0$ ausgeklammert werden kann. Bspw. hat $p = x^3 - 3x - 2$ die Nullstelle $x_0 = -1$, sodass p einen Faktor $x - x_0 = x + 1$ hat, d. h. $p = (x + 1)(x^2 - x + -2)$. Der zweite Faktor hat wiederum einen Faktor $x_0 = 2$ und daher $p = (x + 1)(x - 2)(x + 1)$. Wir hätten also auch noch einmal $x_0 = 1$ als Nullstelle raten können.
    - Natürlich sollte man beim Raten auch sein Gehirn ein wenig einschalten um zu schauen, welche Terme in den Faktoren da sein müssen, um den ursprünglichen Term zu erhalten.
- Substitution sollte man auch sehen können, wie z. B. $(x^2 + y^2)^2 = x^4 + 2x^2y^2 + y^4$, was $(a + b)^2$ für $a = x^2$ und $b = y^2$ ist.

B Halmos' *Naive Mengenlehre* über die russellsche Antinomie

Halmos' *Naive Mengenlehre*[1] behandelt die russellsche Antinomie. Hier ein Auszug:

> **Aussonderungsaxiom**: *Zu jeder Menge A und jeder Bedingung $S(x)$ gibt es eine Menge B, deren Elemente genau jene x aus A sind, für die $S(x)$ gilt.* [...]
>
> Man erhält eine amüsante und zugleich recht lehrreiche Anwendung des Aussonderungsaxioms, wenn man für $S(x)$ die Aussage
>
> $$\text{nicht } (x \in x)$$
>
> $[x \notin x]$ setzt. [...] Unabhängig von der vorgegebenen Menge A folgt: wenn $B = \{x \in A : x \notin x\}$, so gilt für alle y
>
> $$y \in B \textit{ dann und nur dann, wenn } (y \in A \textit{ und } y \notin y). \qquad (*)$$
>
> Kann $B \in A$ gelten? Die Antwort heißt nein! Selbstverständlich gilt entweder $B \in B$ (ungewöhnlich, aber nicht offenkundig unmöglich), oder $B \notin B$; im Falle $B \in A$ würde aber aus $B \in B$ auf Grund von $(*)$ $B \notin B$ folgen, also ein Widerspruch, andererseits aus $B \notin B$ ebenfalls nach $(*)$ $B \in B$, also abermals ein Widerspruch. Demnach kann auf keinen Fall $B \in A$ gelten, also haben wir $B \notin A$. Das Interessante an dieser Folgerung ist die Feststellung, daß es etwas gibt (nämlich B), das nicht zu A gehört; die Menge A war bei dieser Betrachtung völlig beliebig. Damit haben wir mit anderen Worten bewiesen:
>
> *keine Menge enthält alles,*
>
> oder eindrucksvoller:
>
> *es gibt keine Allmenge.*
>
> Noch anders gewendet: eine Gesamtheit, die alle Objekte (Mengen) unserer Theorie enthalten würde, könnte nicht selbst Objekt unserer Theorie sein.
>
> In älteren (präaxiomatischen) Zugängen zur Mengenlehre hatte man die Existenz einer Allmenge als selbstverständlich angenommen und damit auch naiv geglaubt, daß jede Aussage $S(x)$ im absoluten Sinne, nämlich in der vermeintlichen Allmenge, eine „Erfüllungs"-Menge bestimme. In obiger Argumentation fällt dann aber bei der Definition von B, also auch in $(*)$, der Zusatz $x \in A$ (bzw. $y \in A$) fort, und es offenbart sich an der Frage, ob $B \in B$ oder $B \notin B$, ein eklatanter Widerspruch in der Theorie, bekannt als *russellsche Paradoxie*. Die Nutzanwendung liegt darin, daß es – insbesondere in der Mathematik – unmöglich ist, aus dem Nichts etwas zu schaffen. Zur Beschreibung einer Menge genügt es also nicht, ein paar magische Worte auszusprechen (selbst wenn sie eine formal richtig gebildete Aussage wie „$x \notin x$" darstellen), sondern man muß schon eine Menge zur Verfügung haben (in unserer Argumentation hieß sie A) auf deren Elemente sich die magischen Worte anwenden lassen.

[1] Halmos, Paul R., Ostermann, Fritz & Armbrust, Manfred (Übers.): *Naive Mengenlehre.* Göttingen 1968: Vandenhoeck & Ruprecht, 2. Auflg. Kap. 2, S. 19f.

C Cauchys exponenzielle Funktionalgleichung

Die Frage nach der Exponenzialfunktion für allgemeine Basen $a \in \mathbb{R}$ können wir uns wie folgt stellen:

Aufgabe C.A (Cauchys exponenzielle Funktionalgleichung)**.** Finde alle Funktionen $f\colon \mathbb{Q} \to \mathbb{R}$, die die Funktionalgleichung

$$f(x+y) = f(x) + f(y) \qquad \text{für alle } x, y \in \mathbb{Q}$$

erfüllen.

Lösung. Allgemein ist $f(x) = f(x/2 + x/2) = f(x/2)^2 \geq 0$. Nun können wir zwei Fälle für $a \coloneqq f(1)$ unterscheiden.

- Sei $a > 1$. Dann können wir folgende Schritte betrachten.
  - $f(0)$: Es gilt $f(1) = f(1+0) = f(1)f(0) > 0$, also $f(0) \neq 0$ und damit $f(0) = f(0)^2 \iff f(0) = 1$.
  - f auf $\mathbb{N}$: Für alle $n \in \mathbb{N}$ gilt per Induktion $f(nx) = f(x + \cdots + x) = f(x) \cdots f(x) = f(x)^n$ mit $x \in \mathbb{Q}$, also insbesondere $f(n) = a^n$.
  - f auf $\mathbb{Q}_+$: Für $x = m/n \in \mathbb{Q}_+$ erhalten wir aus dem gerade Gezeigtem $f(m/n)^n = f(n(m/n)) = f(m) = a^m \iff f(m/n) = \sqrt[n]{a^m} = a^{m/n}$.
  - f auf $\mathbb{Q}_-$: Allgemein erhalten wir für $x < 0$: $1 = f(0) = f(x-x) = f(x)f(-x) \iff f(-x) = 1/f(x)$, also $f(-m/n) = 1/f(m/n) = 1/a^{m/n} = a^{-m/n}$ für $x = -m/n \in \mathbb{Q}_-$.

  Insgesamt erhalten wir also $f(x) = a^x$ für alle $x \in \mathbb{Q}$ und $a > 0$.
- Sei $a = 0$. Dann ist $f(0) = f(1-1) = f(1)f(-1) = 0$ und damit $f(x) = f(x+0) = f(x)f(0) = 0$ für alle $x \in \mathbb{Q}$.

Schließen wir die triviale Lösung $f = 0$ aus, so erfüllt *nur* die verallgemeinerte Exponenzialfunktion a^x die Funktionalgleichung in $\mathbb{Q}$. Aber wie können wir das auf $\mathbb{R}$ erweitern? Tatsächlich geht das nicht so einfach und wir müssen Annahmen machen.

Eine Möglichkeit ist anzunehmen, dass f in einem Punkt $x_0 \in \mathbb{R} \setminus \mathbb{Q}$ stetig ist, d. h. $\lim_{x \to x_0} f(x) = f(x_0)$ mit $x \in \mathbb{Q}$. Dann ist f auch auf gesamt $\mathbb{R}$ stetig. Wir können nämlich für jedes irrationale x_1 via Vollständigkeitsaxiom eine rationale Cauchy-Folge $x \to x_1$ finden, und es gilt

$$\lim_{x \to x_1} f(x) = \lim_{x \to x_0} f(x_1 + x - x_0) = f(x_1 - x_0) \lim_{x \to x_0} f(x) = f(x_1 - x_0) f(x_0) = f(x_1).$$

Damit ist f auf $\mathbb{R}$ stetig. Damit können wir $f\colon \mathbb{Q} \to \mathbb{R}$ eindeutig auf $\mathbb{R}$ stetig fortsetzen: $f(x) = a^x$ oder $f(x) = 0$ für alle $x \in \mathbb{R}$.