Skript zur Vorlesung “Analysis I” im SS2018 an der Bergischen Universität Wuppertal

Dr. Georg Biedermann

20. Juli 2018

Inhaltsverzeichnis

1 Einige Grundlagen

1.1 Mengen

Es ist hier weder nötig, noch haben wir die Zeit, eine komplette axiomatische Einführung in die Mengentheorie zu geben. Dies ist der Stoff einer eigenen Vorlesung. Beim intuitiven Gebrauch des Formalismus macht man keine Fehler, solange man sich einiger Probleme klar ist.

Das erst Problem ist, dass es nicht einfach ist, eine Mengen überhaupt zu definieren. Es gibt mehrere Möglichkeiten, Mengen einzuführen. Die beiden am meisten studierten Axiomatisierungen sind von Zermelo-Fraenkel und Gödel-Bernays.

Kurt Friedrich Gödel `https://de.wikipedia.org/wiki/Kurt_G%C3%B6del`
Paul Bernays `https://de.wikipedia.org/wiki/Paul_Bernays`
John von Neumann `https://de.wikipedia.org/wiki/John_von_Neumann`
Ernst Friedrich Ferdinand Zermelo `https://de.wikipedia.org/wiki/Ernst_Zermelo`
Adolf Abraham Halevi Fraenkel `https://de.wikipedia.org/wiki/Adolf_Abraham_Halevi_Fraenkel`

Die Theorie der Mengen wurde im 19. Jahrhundert durch den Mathematiker Georg Cantor (1845-1918) begründet.

Georg Cantor `https://de.wikipedia.org/wiki/Georg_Cantor`

Von ihm stammt die folgende Definition einer Menge: "Eine Menge ist eine Zusammenfassung bestimmter wohlunterschiedener Objekte unserer Anschauung oder unseres Denkens – welche die Elemente der Menge genannt werden – zu einem Ganzen."

Dies ist keine mathematische Definition im modernen Sinne und heißt nichts anderes als: Eine Menge ist ein Sack mit einigen Dingen darin, die Elemente der Menge genannt werden. Bespiele: $\{0\}, \{2, 5, 3/2\}, \mathbb{R}$, die Menge aller Studenten der Uni Wuppertal, die Menge aller Elementarteilchen im Universum, die Menge aller orientierungserhaltenden Rotationen von $\mathbb{R}^{19}$, die leere Menge $\varnothing$.

Das wichtigste Problem ist, dass nicht alles eine Menge sein kann. Die "Menge aller Mengen" ist keine Menge. Wenn man darauf besteht, die "Menge aller Mengen" als Menge zu behandeln, ergeben sich logische Widersprüche, mit denen die ganze Theorie zusammenbricht. Dies ist ein echtes Problem, nicht nur eine Schwierigkeit wie die Definition oben, die man mit viel Arbeit überwinden kann. Eine Lösung dieses Problems ist, eine weitere Stufe von "Mengen" einzuführen: die sogenannten Klassen. Alle Mengen zusammengefasst bilden eine Klasse. Die "Klasse aller Klassen" ist aber selber wieder keine Klasse, sondern ...

Bemerkung 1.1. Beachten Sie die folgenden Dinge beim Umgang mit Mengen, insbesondere auch bei der Unterscheidung zwischen Mengen und Folgen (die wir später einführen werden).

1. Eine Menge kann selbst wieder Element einer anderen Menge sein. Beispiele: $\{0, \{0\}\}$, die Menge aller Teilmengen einer Menge M. Die leere Menge ist Teilmenge *jeder* anderen Menge.

2. Eine Menge wird dadruch beschrieben, dass man ihre Elemente angibt. Dabei kommt es nicht auf die Reihenfolge an und es gibt keine Wiederholungen. Beispiel:

$$\{1, 2\} = \{2, 1\} \quad \{a, a\} = \{a\}$$

Notation 1.2. Wir sagen, A ist Teilmenge von C und schreiben $A \subset C$, wenn jedes Element von A auch Element von C ist.

Beachten Sie, dass $A \subset C$ auch die beiden Extremfälle $A = \varnothing$ und $A = C$ einschließt. Wir schreiben $A \subsetneq C$, wenn $A \subset C$ und $A \neq C$ gilt.

Bemerkung 1.3. Zwei Mengen A und B sind gleich, wenn sie dieselben Elemente haben. Insbesondere gilt:

$$A = B \iff (A \subset B \text{ und } B \subset A).$$

Wir beschreiben jetzt kurz die wichtigsten Operationen auf Mengen.

Notation 1.4. Seien A, B und C Mengen.

1. Die Vereinigung $A \cup B$ von A und B:

$$x \in A \cup B \Leftrightarrow x \in A \text{ oder } x \in B$$

   Beachten Sie, dass das mathematische "oder" immer ein einschließliches "oder" ist; d.h. $a \in A$ oder $x \in B$ oder beides.

2. Der Schnitt $A \cap B$ von A und B:

$$x \in A \cap B \Leftrightarrow x \in A \text{ und } x \in B$$

3. $A - B$ (lies "A ohne B"):

$$x \in A - B \Leftrightarrow x \in A \text{ und } x \notin B$$

Bemerkung 1.5. Die Tatsache, dass "oder" einschließlich gemeint ist, kann auch folgendermaßen ausgedrückt werden: es gilt für alle Mengen A und B immer

$$A \cap B \subset A \cup B.$$

Der Schnitt von A und B ist die Menge der Elemente, die sowohl in A, als auch in B sind. Für diese Elemente gilt insbesondere, dass sie entweder in A oder in B sind, also in $A \cup B$.

Wenn Sie das ausschließliche "oder" in Ihren mathematischen Betrachtungen benutzen wollen, dann müssen Sie es ausdrücklich sagen: "Das Element x ist in A oder in B, aber nicht in beiden gleichzeitig." Die entsprechende Menge ist

$$(A \cup B) - (A \cap B).$$

Notation 1.6. Sei $A \subset C$. Dann heißt $C - A$ das *Komplement* (nicht Kompliment!) von A in C.

Sei A eine Menge. Dann bilden alle Teilmengen von A wieder eine Menge, die sogenannte *Potenzmenge von A*. Wir schreiben dafür $P(A)$. Beispiel: $P(\{1,2\}) = \{\varnothing, \{1\}, \{2\}, \{1,2\}\}$ oder $P(\varnothing) = \{\varnothing\}$ (also $|P(\varnothing) \neq \varnothing$!).

Notation 1.7. Seien A und B Mengen. Dann heißt die Menge $A \times B$ der geordneten Paare (a, b) mit $a \in A$ und $b \in B$ das *kartesische Produkt* oder einfach nur *Produkt* von A und B.

$$A \times B = \{(a, b) \mid a \in A, b \in B\}$$

Man beachte, dass die Mengen $A \times B$ und $B \times A$ nicht gleich sind. Die Reihenfolge in den Paaren unterscheidet sich. Natürlich gibt es Bijektion (siehe Definition 1.14)

$$A \times B \to B \times A, \ (a, b) \mapsto (b, a),$$

trotzdem ist es wichtig, hier nichts zu vermischen.

Wenn wir Operationen wiederholen, so benutzen wir häufig die folgende sehr zu empfehlende Schreibweise. Seien $A_1, \ldots, A_n$ Mengen.

$$\bigcup_{i=1}^{n} A_i, \bigcap_{i=1}^{n} A_i, \prod_{i=1}^{n} A_i$$

1.2 Quantoren

Quantoren sind Operatoren der Prädikatenlogik. Wir werden sie einfach als abkürzende Schreibweise in formalen Aussagen benutzen. Es gibt zum einen den Allquantor $\forall$ ("für alle") und den Existenzquantor $\exists$ ("es gibt" oder "es existiert").

Beispiel 1.8. Die folgenden Beispiele sollen die Verwendung illustrieren.

1. Betrachten Sie folgende Aussage: "Alle Autos sind rot." Unabhängig davon, ob diese Aussage richtig ist oder nicht (sie ist es offensichtlich nicht), wollen wir sie einmal anders formulieren. Sei dazu A die Menge aller Autos und R die Menge aller roten Autos.

   $$\forall \ a \in A : a \in R$$

   Äquivalent dazu ist auch einfach: $A \subset R$.

2. Eine andere Aussage ist: “Es gibt ein grünes Auto.” Formal:

$$\exists a \in A : a \in G,$$

wobei G die Menge aller grünen Autos sei. Äquivalent dazu ist auch einfach: $A \cap G \neq \varnothing$. Beachten Sie, dass diese Aussage bedeutet, dass es mindestens ein grünes Auto gibt. Es kann auch mehrere geben und die Aussage ist immer noch richtig.

3. Quantoren werden oft zur Formulierung von Aussagen benutzt, die in unserer normalen Sprache sehr lange und manchmal grammatikalisch ungeschickt oder ungenau sind. Beispiel: Sei S die Menge der Studenten und Studentinnen an der Universität Wupertal.

$$\forall\ s \in S\ \exists\ n \in \mathbb{N},$$

also “für alle Studenten und Studentinnen gibt es eine natürliche Zahl” (dies könnte z.B. die Immatrikulationsnummer sein). So wie geschrieben, bedeutet es, dass es zu jedem Studenten/-in eine Nummer gibt; es kann aber sein (muss aber nicht), dass mehrere (oder alle) dieselbe Nummer haben.

4. Beachten Sie, dass wir immer von rechts nach links lesen und dass alles, was links steht von dem abhängt, was zuvor gekommen ist. Die Aussage

$$\exists\ n \in \mathbb{N}\ \forall\ s \in S$$

besagt, dass eine einzige Nummer für alle Studenten und Studentinnen gibt; jede(r) hat eine Nummer und diese Nummer ist für alle dieselbe! Das n existiert und ist für alle a dasselbe. Die erst Aussage oben bedeutete dagegen, dass es für jede(n) einzelne(n) eine Nummer gibt und diese Nummer wurde individuell vergeben, variiert also mit s.

Bemerkung 1.9. Es ist nützlich zu wissen, wie man Aussagen, die mit Quantoren geschrieben formuliert wurden, verneint: man vertauscht überall die Quantoren $\forall$ and $\exists$ und verneint am Schluss die Aussage. Bespiel:

$$\forall\ x \in X\ \exists\ y \in Y : f(x,y) \in Z$$

Die Verneinung lautet dann:

$$\exists\ x \in X\ \forall\ y \in Y : f(x,y) \notin Z$$

Anderes Beispiel: “Zu jedem Schloss im Haus habe ich einen Schlüssel.” Die Verneinung dieser Aussage lautet: “ Es gibt ein Schloss im Haus, zu dem ich keinen Schlüssel habe.”

In dieselbe Kerbe schlägt folgendes

Beispiel 1.10. Die Aussage

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

ist wahr. Zu jedem einzelnen $x \in \mathbb{R}$ gibt es sicherlich eine reelle Zahl $y \in \mathbb{R}$, die größer als x ist; setze z.B. $y = x + 1$.

Die Aussage

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

ist falsch. Es gibt keine reelle Zahl y, die größer als alle anderen reellen Zahlen ist.

1.3 Abbildungen

Es folgt die Definition des vielleicht wichtigsten Konzeptes in der ganzen Mathematik.

Definition 1.11. Seien X und Y zwei Mengen. Eine *Abbildung f von X nach Y* ist eine Teilmenge $R \subset X \times Y$, so dass gilt:

1. $\forall\ x \in X\ \exists\ y \in Y : (x, y) \in R$
2. $\forall\ x \in X\ \forall\ y_1, y_2 \in Y : \big((x, y_1) \in R \text{ und } (x, y_2) \in R\big) \implies y_1 = y_2$

Wir schreiben dann $f\colon X \to Y$. Wenn $(x, y) \in R$, so schreiben wir:

$$f(x) = y \quad \text{oder} \quad f\colon X \to Y,\ x \mapsto y.$$

Die Menge X heißt *Definitionsbereich* von f und Y heißt der *Wertebereich* (oder *Wertemenge*).

Die erste Bedingung für eine Abbildung $f\colon X \to Y$ bedeutet, dass jedes Element $x \in X$ abgebildet wird bzw. dass wir jedem x ein Bild zuordnen. So ist das Beispiel

$$2 \mapsto 5 \text{ von } \{1, 2\} \text{ nach } \{1, 5, 14\}$$

keine Abbildung, denn der 1 im Definitionsbereich wird nichts zugeordnet.

Die zweite Bedingung an eine Abbildung sagt, dass das Bild eines Elementes des Definitionsbereichs eindeutig sein muss. So ist z.B.

$$1 \mapsto 5, 2 \mapsto 1, 2 \mapsto 5$$

von $\{1, 2\}$ nach $\{1, 5, 14\}$ keine Abbildung, denn 2 hat kein eindeutiges Bild.

Bemerkung 1.12. Man beachte, dass die Angabe des Definitions- und Wertebereichs zwingend zur Definition einer Abbildung dazugehören! Siehe Beispiel 1.15.

Definition 1.13. Sei $f\colon X \to Y$ eine Abbildung. Die Menge

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

heißt *Bild von f*. Sie ist eine Teilmenge von Y.

Sei $B \subset Y$. Dann heißt die Menge

$$f^{-1}(B) = \{a \in A \mid f(a) \in B\}$$

das *Urbild von B unter f*. Sie ist eine Teilmenge von A. Wenn $B = \{b\}$, dann schreiben wir manchmal $f^{-1}(b)$, anstatt $f^{-1}(\{b\})$.

Definition 1.14. Sei $f\colon X \to Y$ eine Abbildung.

1. Die Abbildung f heißt *injektiv*, wenn gilt:

$$\forall\ x_1, x_2 \in X\ : f(x_1) = f(x_2) \implies x_1 = x_2$$

2. Die Abbildung f heißt *surjektiv*, wenn gilt:
$$\forall\ y \in Y\ \exists\ x \in X : f(x) = y$$
Also ist f surjektiv genau dann, wenn $f(X) = Y$.
3. Die Abbildung f heißt *bijektiv*, wenn sie injektiv und surjektiv ist.

Die Injektivität einer Abbildung f kann auch so formuliert werden:
$$\forall\ x_1, x_2 \in X : \big(x_1 \neq x_2 \implies f(x_1) \neq f(x_2)\big).$$
Das bedeutet, dass zwei voneinander verschiedene Elemente des Definitionsbereichs auch immer verschiedene Bilder haben. Eine Abbildung ist also genau dann nicht injektiv, wenn es zwei voneinander verschiedene Elemente mit demselben Bild gibt.

Die Surjektivität besagt, dass jedes Element des Wertebereichs getroffen wird, d.h. ein Urbild besitzt.

Beispiel 1.15. Die Abbildung $f\colon \mathbb{R}_{\geqslant 0} \to \mathbb{R}, f(x) = x^2$ ist nicht injektiv, denn z.B. $f(1) = f(-1)$. Allerdings ist die Abbildung $g\colon \mathbb{R}_{\geqslant 0} \to \mathbb{R}, g(x) = x^2$ injektiv.

Die Abbildung $f\colon \mathbb{R} \to \mathbb{R}_{\geqslant 0}, f(x) = x^2$ ist surjektiv. Allerdings ist die Abbildung $g\colon \mathbb{R} \to \mathbb{R}, g(x) = x^2$ nicht surjektiv.

Beachten Sie die verschiedenen Definitions -und Bildbereiche.

Lemma 1.16. *Sei $f\colon X \to Y$ eine Abbildung. Dann gilt:*

1. $\forall\ A \subset X : A \subset f^{-1}f(A)$
2. $\big(\forall\ A \subset X : A = f^{-1}f(A)\big) \iff f$ *injektiv*
3. $\forall\ B \subset Y : f(f^{-1}(B)) \subset B$
4. $\big(\forall\ B \subset Y : f(f^{-1}(B)) = B\big) \iff f$ *surjektiv*

Beweis. Eine gute Übung. □

Lemma 1.17. *Seien $A, B \subset C$, $X, Y \subset Z$ und $f\colon C \to Z$ eine Abbildung. Dann gilt:*

1. $f^{-1}(X \cap Y) = f^{-1}(X) \cap f^{-1}(Y)$
2. $f^{-1}(X \cup Y) = f^{-1}(X) \cup f^{-1}(Y)$
3. $f(A \cup B) = f(A) \cup f(B)$
4. $f(A \cap B) \subset f(A) \cap f(B)$

Beweis. Eine gute Übung. □

Definition 1.18. Die Abbildung $f\colon X \to Y$ ist *invertierbar* oder *umkehrbar*, wenn eine Abbildung $g\colon Y \to X$ existiert mit
$$f \circ g = \mathrm{id}_Y \quad \text{und} \quad g \circ f = \mathrm{id}_X\,.$$
Wir nennen g dann *die Umkehrabbildung von f*.

Wenn eine Abbildung f eine Umkehrabbildung besitzt, so ist diese Umkehrabbildung durch f eindeutig bestimmt.

Lemma 1.19. *Eine Abbildung von Mengen ist genau dann bijektiv, wenn sie eine Umkehrabbildung besitzt.*

Beweis. Eine gute Übung. □

Lemma 1.20. *Für jede Menge C gibt es eine Bijektion zwischen der Potenzmenge $P(C)$ und der Menge der Abbildungen von C auf die Menge $\{0,1\}$.*

Beweis. Sei $\mathrm{Abb}(C,\{0,1\})$ die Menge aller Abbildungen von C nach $\{0,1\}$. Wir beschreiben eine Abbildung

$$\alpha\colon\ P(C) \to \mathrm{Abb}(C,\{0,1\}),$$

indem wir eine Teilmenge $A \subset C$ auf diejenige Abbildung $\alpha(A)\colon\ C \to \{0,1\}$ schicken, die die Elemente von A auf 1 abbildet und den Rest auf 0. Umgekehrt definieren wir eine Abbildung folgendermaßen:

$$\begin{aligned} \beta\colon\ \mathrm{Abb}(C,\{0,1\}) &\to P(A) \\ (f\colon\ C \to \{0,1\}) &\mapsto \beta(f) = f^{-1}\{1\}. \end{aligned}$$

Wir zeigen jetzt $\alpha \circ \beta = \mathrm{id}_{\mathrm{Abb}(C,\{0,1\})}$ und $\beta \circ \alpha = \mathrm{id}_{P(C)}$.

Zuerst $\alpha \circ \beta = \mathrm{id}_{\mathrm{Abb}(C,\{0,1\})}$. Sei $f\colon\ C \to \{0,1\}$ eine Abbildung. Dann ist $\beta(f) = f^{-1}\{1\}$ das Urbild von 1. Diesem Urbild ordnet α genau diejenige Abbildung von C nach $\{0,1\}$ zu, die jedes Element des Urbildes auf 1 abbildet und alles andere auf 0. Das ist aber gerade, das was f macht. Also ist $\alpha \circ \beta(f) = f$.

Jetzt $\beta \circ \alpha = \mathrm{id}_{P(C)}$. Sei $A \subset C$. Dann ist $\alpha(A) =: g$ die Abbildung mit

$$g(c) = \begin{cases} 1 & \text{für } c \in A, \\ 0 & \text{für } c \notin A. \end{cases}$$

Dann ist aber $g^{-1}\{1\} = A$, also $\beta \circ \alpha(A) = A$ und damit $\beta \circ \alpha = \mathrm{id}_{P(C)}$. □

1.4 Kardinalität

Definition 1.21. Wir schreiben $|A|$ (oder vielleicht manchmal $\#A$) für die *Kardinalität* oder *Mächtigkeit* der Menge A. Dies ist die Anzahl der Elemente.

Beispiel: $|\{1,2\}| = 2$, $|P(\varnothing)| = 1$ siehe oben, $|\mathbb{N}| = \infty$. Das Letztere bedeutet einfach, dass die natürlichen Zahlen keine endliche Menge sind.

Satz 1.22. *Seien A und B Mengen.*

1. *Es gilt $|A| \leqslant |B|$ genau dann, wenn es eine injektive Abbildung von A nach B gibt.*
2. *Es gilt $|A| \leqslant |B|$ genau dann, wenn es eine surjektive Abbildung von B nach A gibt.*
3. *Zwei Mengen haben genau dann die gleiche Kardinalität, wenn es eine Bijektion zwischen ihnen gibt.*

Der Satz ist für endliche Mengen einfach zu beweisen. Der Beweis wird aber sehr schwierig, wenn man sich nicht auf endliche Mengen einschränkt. Das liegt daran, dass die "Anzahl der Elemente" dann keinen intrinsischen Sinn mehr hat. Wir wollen auch dieses Problem einfach übergehen. Das Augenmerk unserer Analysis-Vorlesung soll auf anderen Dingen liegen.

Lemma 1.23. *Für jede Menge A gilt:* $|P(A)| > |A|$.

Beweis. Wenn $A = \varnothing$, dann ist die Aussage wahr: $P(\varnothing) = \{\varnothing\}$. Also $|\varnothing| = 0 < 1 = P(\varnothing)$.

Sei von jetzt ab $A \neq \varnothing$. Man sieht leicht $|P(A)| \geqslant |A|$, denn es gibt eine Injektion $A \to P(A)$: jedem Element $a \in A$ ordnen wir die Teilmenge $\{a\}$ von A zu.

Wir zeigen jetzt: wenn $f\colon A \to P(A)$ eine Abbildung ist, so ist f nicht surjektiv. Insbesondere folgt dann wieder mit Satz 1.22: $|P(A)| > |A|$.

Also sei $f\colon A \to P(A)$ eine Abbildung. Für jedes $a \in A$ ist also $f(a)$ eine Teilmenge von A. Es kann also passieren, dass a selbst ein Element von $f(a)$ ist oder eben nicht. Wir betrachten die Menge M definiert durch

$$M = \{a \in A \mid a \notin f(a)\} \subset A.$$

Wir nehmen an, dass M ein Urbild $b \in A$ unter f besitzt: $\exists\ b \in A : f(b) = M$. Dann können wir folgenden Fälle für b unterscheiden:

$$b \in M \Longrightarrow b \notin f(b) = M$$

Aber eben auch:

$$b \notin M = f(b) \Longrightarrow b \in M$$

Beides führt zum Widerspruch. Also kann es b nicht gegeben haben. Damit ist f nicht surjektiv. □

Beachten Sie, dass diese Aussage von Lemma 1.23 für alle Mengen gilt ungeachtet der Tatsachte, dass sie endlich oder unendlich sind. Es folgt, dass es verschieden Formen von "unendlich" gibt. Die Menge $\mathbb{N}$ ist unendlich, aber die Menge $P(\mathbb{N})$ ist noch unendlicher!

Definition 1.24. Eine Menge M heißt *abzählbar*, wenn es eine Surjektion $\varphi\colon \mathbb{N} \to M$ gibt oder wenn M leer ist. Die Abbildung φ heißt dann *Abzählung von M* (oder "Aufzählung"?). Eine Menge, die nicht abzählbar ist, heißt *überabzählbar*.

Wenn eine Menge M abzählbar ist, so kann man ihre Elemente als Folge auflisten: $M = \{m_1, m_2, m_3, \ldots\}$, dabei ist $m_n = \varphi(n)$ für eine Abzählung φ.

Beispiel 1.25. Die Menge $\mathbb{N}$ ist per Definition abzählbar. Hier sind noch weitere Beispiele:

1. Endliche Mengen sind abzählbar.
2. Die Vereinigung zweier abzählbarer Mengen ist wieder abzählbar. Gegeben zwei Abzählungen $A = \{a_1, a_2, \ldots\}$ und $B = \{b_1, b_2, \ldots\}$, dann ist

   $$\{a_1, b_1, a_2, b_2, \ldots\} = A \cup B$$

   wieder eine Abzählung.

3. Die ganzen Zahlen $\mathbb{Z} = \mathbb{N} \cup \{0\} \cup -\mathbb{N}$ sind daher abzählbar.

4. Die rationalen Zahlen $\mathbb{Q}$ sind abzählbar. Als Beweis geben wir eine Abzählung aller positiven (nicht unbedingt gekürzten) Brüche an. Mit Aussage (1) (bzw. mit demselben Trick wie in (2)) folgt dann, dass ganz $\mathbb{Q}$ abzählbar ist. Man betrachte die folgende Tabelle, in der in Zeilen die verschiedenen Nenner und in den Spalten die verschiedenen Zähler aufgelistet werden.

	1	2	3	...
1	$\frac{1}{1}$	$\frac{2}{1}$	$\frac{3}{1}$	...
2	$\frac{1}{2}$	$\frac{2}{2}$	$\frac{3}{2}$	...
3	$\frac{1}{3}$	$\frac{2}{3}$	$\frac{3}{3}$	...
$\vdots$	$\vdots$	$\vdots$	$\vdots$	

Wir gehen jetzt der Reihe nach die Diagonalen von rechts oben nach links unten:

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

5. Die reellen Zahlen $\mathbb{R}$ sind überabzählbar. Wir zeigen dies, indem wir zeigen, dass schon das Intervall $]0,1[$ überabzählbar ist. Dazu benutzen wir, dass sich jede reelle Zahl eindeutig als Dezimalbruchentwicklung darstellen lässt: zu jedem $x \in\ ]0,1[$ existieren Koeffizienten $x_1, x_2, \ldots \in \{0,1,\ldots,9\}$, so dass

$$x = \sum_{i=1}^{\infty} x_i 10^{-i} = 0, x_1 x_2 x_3 \ldots$$

Für den Beweis dieser Tatsache und einer genaueren Beschreibung verweisen wir auf Parapgraph 5.4.

Wir nehmen jetzt eine beliebige Abbildung $\varphi\colon\ \mathbb{N} \to]0,1[$ und schreiben die Bilder und deren Dezimalbruchentwicklung in einer Liste wie folgt:

$$\begin{array}{r|l} \varphi(1) & 0, x_{1,1}\, x_{1,2}\, x_{1,3}\ \ldots \\ \varphi(2) & 0, x_{1,1}\, x_{1,2}\, x_{1,3}\ \ldots \\ \varphi(3) & 0, x_{1,1}\, x_{1,2}\, x_{1,3}\ \ldots \\ \vdots & \qquad \vdots \end{array}$$

Wir produzieren jetzt eine neue reelle Zahl, indem wir die Koeffizienten auf der Diagonal ändern: setze

$$y_n = \begin{cases} 5 & , \text{ wenn } x_{n,n} \neq 5 \\ 4 & , \text{ wenn } x_{n,n} = 5 \end{cases}$$

Dann kommt die Zahl $y = 0, y_1 y_2 y_3 \ldots$ nicht in der obigen List vor, denn sie unterscheidet sich an mindestens einer Stelle von jeder Zahl in der Liste. Damit liegt y nicht im Bild von φ. Es kann also keine surjektive Abbildung von $\mathbb{N}$ nach$]0,1[$ geben. Also ist $]0,1[$ überabzählbar.

2 Reelle und komplexe Zahlen

2.1 Die Lösbarkeit von Gleichungen

Wir wollen mehrere Gleichungen auf ihre Lösbarkeit hin untersuchen.

1. Gegeben $a, b, \in \mathbb{N}$, existiert eine Lösung für die Gleichung
$$x + a = b?$$
Diese Gleichung lässt sich immer in $\mathbb{Z}$, aber nicht unbedingt in $\mathbb{N}$ lösen.

2. Gegeben $a, b, c \in \mathbb{Z}$ mit $a \neq 0$, existiert eine Lösung für die Gleichung
$$ax + b = c?$$
Diese Gleichung lässt sich immer in $\mathbb{Q}$, aber nicht unbedingt in $\mathbb{Z}$ lösen.

3. Gegeben $a \in \mathbb{Q}, a > 0$, existiert eine Lösung für die Gleichung
$$x^2 = a?$$
Diese Gleichung lässt sich immer in $\mathbb{R}$, aber nicht in $\mathbb{Q}$ lösen.

Zu 3.: Setzen wir $a = 2$ so ist eine Lösing der Gleichung $x^2 = 2$ die Zahl $\sqrt{2}$. Die andere Lösung ist $-\sqrt{2}$. Es gilt aber:
$$\sqrt{2} \notin \mathbb{Q}.$$
Wir werden gleich zwei verschiedene Beweise geben, dass $\sqrt{2}$ keine rationale Zahl ist. Wir sollten uns aber zuerst folgende Frage stellen: wieso wollen wir eigentlich, dass eine Zahl wie $\sqrt{2}$ existiert. Sicherlich ist es schön und gut, quadratische Gleichungen zu lösen; aber wozu? Eine Antwort darauf kann geometrischer Natur sein. Ein Quadrat mit Seitenlänge 1 ist ein sehr natürliches, oft betrachtetes Objekt (auch außerhalb der reinen Mathematik!). Die Länge seiner Diagonal ist nach dem Satz des Pythagoras $\sqrt{1^2 + 1^2} = \sqrt{2}$. Wenn eine solche einfache geometrische Konstruktion zu einer nicht-rationalen Zahl führt, müssen wir einsehen, dass die rationalen Zahlen einfach nicht ausreichen. Also brauchen wir einen größeren Vorrat an Zahlen.

Jetzt zum Beweis, dass $\sqrt{2}$ nicht in $\mathbb{Q}$ liegt. Wir geben zuerst einen einfachen Beweis durch Widerspruch, der auf der Tatsache beruht, dass man eine ganze Zahl immer eindeutig (bis auf Vorzeichen und Reihenfolge) in Primfaktoren zerlegen kann.

Definition 2.1. Zwei ganzen Zahlen heißen *teilerfremd*, wenn ihr größter gemeinsamer Teiler 1 ist.

Bemerkung 2.2. Zwei ganze Zahlen sind genau dann teilerfremd, wenn in der (eindeutigen) Primfaktorzerlegung der ersten kein Primfaktor der zweiten auftaucht.

Definition 2.3. Für zwei ganze Zahlen a und b schreiben wir $a|b$, wenn ein $k \in \mathbb{Z}$ existiert, so dass
$$k \cdot a = b$$
gilt. Wir sagen dann: a *teilt* b.

Erster Beweis: Wir nehmen an, es existieren zwei teilerfremde $p, q \in \mathbb{Z}$, so dass $\sqrt{2} = \frac{p}{q}$ gilt. Die Annahme ist also $\sqrt{2} = \frac{p}{q} \in \mathbb{Q}$. Dies gilt es, zum Widerspruch zu führen. Es folgt dann:

$$p^2 = 2q^2 \qquad (*)$$

Dann ist p^2 gerade. Also ist p gerade. Das bedeutet: $4|p^2$. Aus der Gleichung $(*)$ folgt dann $2|q^2$, also ist q^2 gerade. Dann ist aber auch q gerade. Das ist ein Widerspruch zur Annahme, das p und q teilerfremd sind. $\square$

Wir geben jetzt einen zweiten direkten Beweis, d.h. ohne eine Annahme zum Widerspruch zu führen. Dazu stellen wir die Frage: Wie weit ist der Wert des Polynoms $P(x) := x^2 - 2$ an der rationalen Stelle $x = p/q$ von 0 verschieden? Wir werden zeigen, dass das Polynom P an der Stelle $x = p/q$ einen Wert vom Betrag mindestens $1/q^2$ hat. Insbesondere kann dann $\sqrt{2}$, das ja eine Nullstelle von P ist, nicht von der Form p/q, also rational sein.

Lemma 2.4. *Seien $p, q \in \mathbb{Z}$ und $P(x) = x^2 - 2$. Dann gilt $|P(p/q)| \geqslant 1/q^2$.*

Beweis. Wenn der Bruch p/q nicht gekürzt ist (also wenn p und q einen gemeinsamen Teiler haben), so ist q größer als der Nenner des gekürzten Bruches. Dann ist $1/q^2$ kleiner als der entsprechende Ausdruck für den gekürzten Bruch. Die behauptete Ungleichung für die gekürzte Darstellung der Zahl p/q ist also stärker. Wir können also ohne Beschränkung der Allgemeinheit annehmen, dass p und q teilerfremd sind (denn die Aussage, die wir beweisen werden, ist stärker als die Aussage für einen ungekürzten Bruch).

Seine also p und q teilerfremd; insbesondere sind also alle Faktoren 2 aus dem Bruch gekürzt. Das bedeutet, dass entweder p oder q ungerade ist, aber nicht beide.

Wir betrachten zuerst den Fall, dass p ungerade ist. Dann ist der Zähler von

$$P(p/q) = (p/q)^2 - 2 = (p^2 - 2q^2)/q^2$$

ungerade. Damit ist aber sein Absolutbetrag größer als 1, und es folgt:

$$|P(p/q)| = |(p/q)^2 - 2| = |p^2 - 2q^2|/q^2 \geqslant 1/q^2.$$

Also ist die Ungleichung in diesem Fall bewiesen.

Es bleibt, den anderen Fall, in dem q ungerade ist, zu betrachten. Dann ist p gerade und $2 \cdot 2 = 4$ teilt p^2. Der Term $2q^2$ hat aber genau einen Faktor 2, denn q ist ungerade. Also ist der Zähler $p^2 - 2q^2$ von $P(p/q)$ das Doppelte der ungeraden ganzen Zahl $p^2/2 - q^2$. Also ist auch hier der Betrag des Zähler mindestens 1 und es gilt wie oben: $|P(p/q)| \geqslant 1/q^2$. $\square$

Man kann zeigen, dass es in $\mathbb{R}$ zu jeder nicht-negativen reellen Zahl eine Wurzel gibt (tatsächlich gibt es für $a > 0$ immer genau zwei Wurzeln, eine positive und eine negative). Was passiert aber, wenn a negativ ist?

4. Gegeben $a \in \mathbb{R}$, dann existiert eine Lösung für die Gleichung

   $$x^2 = a?$$

   zwar nicht immer in $\mathbb{R}$, aber immer in den komplexen Zahlen $\mathbb{C}$.

Auch hier kann man also durch eine formale Konstruktion, den komplexen Zahlen $\mathbb{C}$, Abhilfe schaffen. Wieder stellt sich die Frage, ob es überhaupt sinnvoll ist, solche "komplexe" Zahlen zu betrachten, denn sowohl die abstrakte Konstruktion, als auch der Wunsch, Wurzeln aus negativen Zahlen zu ziehen, scheinen etwas an den Haaren herbeigezogen. Diese Frage läßt sich zum einen durch eine inzwischen jahrhunderte lange erfolgreiche Anwendung der komplexen Zahlen in Mathematik und Physik beantworten. Zum anderen gibt es auch hier wieder eine verblüffend einfache geometrische Interpretation der komplexen Zahlen, die jegliche Zweifel an ihrer "wirklichen Existenz" besitigen kann. Siehe Lemma 2.29.

2.2 Axiomatische Einführung der reelen Zahlen

Definition 2.5. Eine *Gruppe* ist eine Menge G und ein Element $e \in G$ zusammen mit einer *Verknüpfungsabbildung* $- \cdot -: G \times G \to G$, so dass folgende Bedingungen erfüllt sind:

1. Assoziativität: $\forall\ f, g, h \in G \quad (f \cdot g) \cdot h = f \cdot (g \cdot h)$
2. Neutrales Element: $\exists\ e \in G\ \forall\ g \in G \quad e \cdot g = g = g \cdot e$
   Man nennt e das *neutrale Element.*
3. Inverse Elemente: $\forall\ g \in G\ \exists\ g^{-1} \in G \quad g \cdot g^{-1} = e = g^{-1} \cdot g$.
   Man nennt dann g^{-1} das *zu g inverse Element.*

Man nennt eine Gruppe *abelsche* oder *kommutativ*, wenn zusa ̈tzlich das Kommutativitätsgesetz

$$\forall g, h \in G \quad gh = hg$$

gilt.

Man beachte, dass die Verknüpfung in einer Gruppe abstrakt oft als Multiplikation $- \cdot -$ geschrieben wird, dass sie aber im konkreten Fall nicht die Multiplikation von Zahlen sein muss. So wird z.B. in einer ableschen Gruppe die Verknüpfung oft mit einem Pluszeichen geschrieben. Auch hier muss nicht die Addition von Zahlen gemeint sein.

Beispiel 2.6. Wir geben einige Besipiele:

1. Die Paare $(\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{Q} - \{0\}, \cdot), (\mathbb{R}, +), (\mathbb{R} - \{0\}, \cdot), , (\mathbb{C}, +), (\mathbb{C} - \{0\}, \cdot)$ sind alles Beispiele für ablesche Gruppen.
2. Das Paar $(\mathbb{N}, +)$ ist *keine* Gruppe; es gibt keine inversen Elemente.
3. Die kleinste nicht-abelsche Gruppe ist die symmetrische Gruppe Σ_3; es ist die Gruppe aller Permutationen von drei Elementen. Die Verknüpfung hier ist die Hintereinanderausführung $- \circ -$ von Abbildungen.
4. Die Gruppe $\mathrm{G}l_n(\mathbb{R})$ der reellen invertiervaren $n \times n$-Matrizen mit Matrixmultiplikation bildet für $n \geqslant 2$ eine nicht-abelsche Gruppe.

Definition 2.7. Ein *Körper* ist eine Menge K mit zwei ausgezeichneten Elementen 0 und 1 in K mit $0 \neq 1$ und zwei Verknüpfungen

$$\begin{aligned} &\text{(Addition)} && - + -\colon\ K \times K \to K \quad \text{und} \\ &\text{(Multiplikation)} && - \cdot -\colon\ K \times K \to K, \end{aligned}$$

so dass folgende Bedingungen erfüllt sind:

1. Das Paar $(K, +)$ ist eine abelsche Gruppe mit neutralem Element 0.
2. Das Paar $(K - \{0\}, \cdot)$ ist eine abelsche Gruppe mit neutralem Element 1.
3. Für alle $a \in K$ gilt: $0 \cdot a = 0 = a \cdot 0$.
4. Für alle $a, b, c \in K$ gilt:

$$\begin{aligned} \text{(Distributivität)} \qquad & a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{und} \\ & (a + b) \cdot c = a \cdot c + b \cdot c \end{aligned}$$

Oft schreiben wir für $K - \{0\} = K^*$. Das *additive Inverse* eines Elements $a \in K$ wird $-a$ geschrieben; das *multiplikative Inverse* von a wird a^{-1} oder $1/a$ geschrieben.

Jeder Körper hat also mindestens zwei Elemente, nämlich eine 0 und 1. Tatsächlich gibt es einen Körper mit genau zwei Elementen.

Beispiel 2.8. Wir definieren auf der Menge $\mathbb{F}_2 = \{0, 1\}$ eine Addition mit 0 als neutralem Element und der zusätzlichen Gleichung:

$$1 + 1 = 0.$$

Auf der Menge $\mathbb{F}_2^* = \{1\}$ gibt es nur eine Gruppenverknüpfung: $1 \cdot 1 = 1$. Damit wird $(\mathbb{F}_2, +, \cdot)$ zu einem Körper!

Dieses Beispiel kommt Ihnen komisch vor? Betrachten Sie in $\mathbb{Z}$ Teilen mit Rest modulo 2. Dann sind die einzigen möglichen Reste 0 oder 1. Wenn Sie nun die übliche Addition und Multiplikation auf $\mathbb{Z}$ betrachten und dann Reste modulo 2 nehmen, kommen Sie genau zu obigen Körper: z.B. ungerade + gerade ergibt ungerade, also $1 + 0 = 1$; und ungerade + ungerade ergibt gerade, also $1 + 1 = 0$. Dieses Beispiel wird Ihnen ganz sicherlich in anderen Vorlesungen nochmal über den Weg laufen.

Lemma 2.9. *In jedem Körper gilt:* $(-1)^2 = 1$.

Beweis. Die 1 ist das neutrale Element bzgl. der Multiplikation und -1 ist das additive Inverse von 1. Also gilt in jedem Körper (tatsächlich in jedem Ring mit 1) mit dem Distributivitätsgesetz:

$$(-1) + (-1)^2 = (-1) \cdot 1 + (-1) \cdot (-1) = (-1)\big(1 + (-1)\big) = (-1) \cdot 0 = 0.$$

Wir addieren 1 auf jeder Seite und erhalten: $(-1) \cdot (-1) = 1$. □

Die reellen Zahlen erfüllen die "algebraischen Axiome" eines Körpers, aber es gibt weitere wichtige Eigenschaften von $\mathbb{R}$, die wir noch nicht axiomatisiert haben: eine davon ist die Anordnung der reellen Zahlen.

Definition 2.10. Ein Körper K heißt *angeordnet*, wenn es eine Teilmenge K_+ der *positiven Elemente* gibt, so dass gilt:

1. Für jedes Element $a \in K$ gilt genau eine der folgenden Aussagen:
$$a \in K_+, a = 0, a \in K - (K_+ \cup \{0\}) =: K_-$$
2. Aus $a, b \in K_+$ folgt $a + b \in K_+$ und $a \cdot b \in K_+$.

Wir schreiben $a > 0$, wenn $a \in K_+$ gilt. Wir nennen K_- die *negativen Elemente* und schreiben $a < 0$, wenn a negativ ist, also $a \in K_-$.

Man benützt weitere Abkürzungen, von denen wir nur einige aufzählen:

$$\begin{array}{lcl} a > b & \Leftrightarrow & a - b > 0 \\ a \geqslant b & \Leftrightarrow & (a - b > 0 \text{ oder } a = b) \\ a < b & \Leftrightarrow & b - a > 0 \end{array}$$

In angeordneten Körpern kann man folgende Aussage zeigen:

$$\begin{array}{lcl} (1)\ a > b \text{ und } b > c & \Rightarrow & a > c \ \ \text{(Transitivität)} \\ (2)\ a > b & \Rightarrow & a + c > b + c \\ (3)\ a > b & \Rightarrow & -a < -b \\ (4)\ a > b \text{ und } c > 0 & \Rightarrow & a \cdot c > b \cdot c \\ (5)\ a > b \text{ und } c < 0 & \Rightarrow & a \cdot c < b \cdot c \\ (6)\ 1 > 0 \text{ d.h. } 1 \in K_+ & & \\ (7)\ \forall a \in K - \{0\}\ a^2 > 0 & & \end{array}$$

Wir wollen Aussage (4) beweisen, um ein Beispiel solcher typischer Argumente zu geben: Sei $a > b$. Nach Definition dieser Schreibweise gilt also $a - b > 0$ (oder $a - b \in K_+$). Mit $c > 0$ folgt aus dem Distributivitätsgesetz und dem zweiten Anordnungsaxiom: $ac - bc = (a - b)c > 0$. Also $ac > bc$. □

Die Anordnung allein ist immer noch nicht die Eigenschaft, auf die wir bei der Axiomatisierung von $\mathbb{R}$ aus sind. Wir wollen eine ganz bestimmte Art der Anordnung definieren. Dazu brauchen wir folgende

Bemerkung 2.11. Für einen Körper K gibt es immer eine Abbildung $\varphi\colon\ \mathbb{Z} \to K$
$$n \mapsto \varphi(n) = n \cdot 1_K = \sum_{k=1}^{n} 1_K = 1_K + \ldots + 1_K.$$

Man beachte, dass die Abbildung nicht immer injektiv sein muss: für $\varphi\colon\ \mathbb{Z} \to \mathbb{F}_2$ werden alle geraden Zahlen auf 0 abgebildet. Für einen angeordneten Körper und alle $n \in \mathbb{N}$ gilt hingegen
$$\varphi(n + 1) = (n + 1) \cdot 1_K = n \cdot 1_K + 1_K > n \cdot 1_K = \varphi(n),$$
denn $1_K > 0_K$. Daraus folgt insbesondere, dass φ injektiv ist, wenn K ein angeordneter Körper ist.

Definition 2.12. Ein Körper K heißt *archimedisch angeordnet*, wenn K ein angeordneter Körper ist und das folgende *archimedische Axiom* erfüllt ist:
$$\forall a \in K\ \exists\ n \in \mathbb{N}\ a < n \cdot 1_K.$$

Archimedes von Syrakus `https://de.wikipedia.org/wiki/Archimedes`

Das Axiom geht aber wahrscheinlich zurück auf:
Eudoxos von Knidos `https://de.wikipedia.org/wiki/Eudoxos_von_Knidos`

Beispiel 2.13. Man beachte:

1. Der Körper der $\mathbb{Q}$ der rationalen Zahlen ist genauso wie $\mathbb{R}$ archimedisch angeordnet.
2. Es gibt angeordnete Körper, die nicht archimedisch angeordnet sind.
3. Nicht jeder Körper kann angeordnet werden. Als Beispiel dient der Körper $\mathbb{F}_2$ aus Beispiel 2.8, auf dem es keine Anordnung $<$ gibt, die mit Addition und Multiplikation kompatibel ist. Tatsächlich kann kein endlicher Körper angeordnet werden.
4. Auf den komplexen Zahlen $\mathbb{C}$ gibt es keine Anordnung, die mit den Körperaxiomen verträglich ist. Anschaulich kann man diese Aussage verstehen, wenn man sich die komplexen Zahlen als Ebene vorstellt. Im Gegensatz zur "Zahlengerade" $\mathbb{R}$ gibt es hier keine ausgewiesene Richtung, in der man anordnen könnte. (Dies ist natürlich kein mathematischer Beweis, sondern nur eine Veranschaulichung.)

Die letzte wichtige Eigenschaft der reellen Zahlen ist ihre Vollständigkeit. Diese Eigenschaft kann auf mehrere verschiedene Arten ausgedrückt werden. Wir werden den Begriff der Vollständigkeit erst im Abschnitt 4.2 definieren und dann studieren. Jetzt erwähnen wir nur, dass $\mathbb{Q}$ nicht vollständig ist, denn den rationalen Zahlen fehlen gewisse Punkte, z.B. $\sqrt{2}$.

Im folgenden Satz behaupten wir die Existenz und Eindeutigkeit der reellen Zahlen. Bevor wir die Eindeutigkeit besprechen, müssen wir den Begriff des Körperhomomorphismus und des Körperisomorphismus einführen.

Definition 2.14. Seien K und L Körper und sei $\varphi\colon K \to L$ eine Abbildung. Dann ist φ ein *Körperhomomorphismus*, wenn die Gleichungen

1. $\forall\ a, b \in K \quad \varphi(a+b) = \varphi(a) + \varphi(b)$
2. $\forall\ a, b \in K \quad \varphi(ab) = \varphi(a)\varphi(b)$

gilt.

Ein Körperhomomorphismus $\varphi\colon K \to L$ ist ein *Körperisomorphismus*, wenn es einen Körperhomomorphismus $\vartheta\colon L \to K$ gibt, so dass $\varphi \circ \vartheta = \mathrm{id}_L$ und $\vartheta \circ \varphi = \mathrm{id}_K$ gilt. Zwei Körper heißen *isomorph*, wenn es zwischen ihnen einen Körperisomorphismus gibt.

Bemerkung 2.15. Für den (einfachen) Beweis der folgenden beiden Aussagen verweisen wir auf die lineare Algebra.

1. Für einen Körperhomomorphismus $\varphi\colon K \to L$ gilt immer:
$$\varphi(0_K) = 0_L \quad \text{und} \quad \varphi(1_K) = 1_L.$$
2. Ein Körperhomomorphismus ist genau dann ein Körperisomorphismus, wenn er bijektiv ist.

Satz 2.16. *Die reellen Zahlen* $\mathbb{R}$ *sind ein archimedisch angeordneter vollständiger Körper. Jeder andere archimedisch angeordnete vollständige Körper ist isomorph zu* $\mathbb{R}$.

Dieser Satz axiomatisiert den Körper der reellen Zahlen, indem er diejenigen Eigenschaften von $\mathbb{R}$ aufzählt, die ausreichen, um ihn bis auf Isomorphismus eindeutig zu bestimmen. Zu jedem anderen Körper K, der vollständig und archimedisch angeordnet ist, gibt es also eine bijektive Abbildung $\varphi\colon K \to \mathbb{R}$, die kompatibel mit Addition und Multiplikation im Sinne von Definition 2.14 ist. Das bedeutet, dass dann K in jedem mathematisch relevanten Kontext genauso gut ist wie $\mathbb{R}$; K ist $\mathbb{R}$ (via φ).

Wir werden im Abschnitt 4.2 erklären, was Vollständigkeit bedeutet. Damit eng verknüpft sind Konzepte wie Darstellung reeller Zahlen via Dezimalbrüche, Intervallschachtelung, die Existenz von Wurzeln positiver Zahlen. Wir werden den Satz 2.16 nicht beweisen. Ein solcher Beweis würde zwingend einen detailierten axiomatischen Aufbau der Zahlen angefangen mit den Peano-Axiomen der natürlichen Zahlen, Gruppenvervollständigung, Quotientenkörper und metrische Vervollständigung voraussetzen. In dieser Vorlesung über Analysis werden wir über diese wichtigen Grundlagen hinwegspringen und die reellen Zahlen als gegeben ansehen.

Giuseppe Peano `https://de.wikipedia.org/wiki/Giuseppe_Peano`

Definition 2.17. Ein *Intervall* ist eine Teilmenge von $\mathbb{R}$ aller Zahlen zwischen einer unteren Schranke a und einer oberen Schranke b. Dabei unterscheiden wir die folgenden Fälle:

1. Ein *offenes Intervall*:
$$]a,b[= \{x \in \mathbb{R} \mid a < x < b\}$$
2. Ein *(ab-)geschlossenes Intervall*:
$$[a,b] = \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$$
3. *Halboffene Intervalle*:
$$]a,b] = \{x \in \mathbb{R} \mid a < x \leqslant b\} \quad \text{und} \quad [a,b[= \{x \in \mathbb{R} \mid a \leqslant x < b\}$$

2.3 Einige grundlegende Dinge über $\mathbb{R}$

Definition 2.18. Der *Absolutbetrag* oder einfach *Betrag* einer reellen Zahl x ist wir folgt definiert:
$$|x| := \begin{cases} x & , \text{ für } x \geqslant 0 \\ -x & , \text{ für } x < 0 \end{cases}$$

Lemma 2.19. *Es gilt für alle* $x, y \in \mathbb{R}$:

(1) $|x| \geqslant 0$

(2) $|x| = 0 \iff x = 0$

(3) $|xy| = |x||y|$ *und* $|x^2| = |x|^2$

(4) $|x + y| \leqslant |x| + |y|$ (Dreiecksungleichung)

(5) $||x| - |y|| \leqslant |x - y|$

Der Name “Dreiecksungleichung” leuchtet eher ein, wenn man die entsprechende Aussage im zweidimensionalen Raum (und nicht nur in $\mathbb{R}$) betrachtet.

Beweis. Die Aussagen (1) und (2) folgen direkt aus der Definition des Absolutbetrages.

(3) Um (3) zu beweisen, machen wir eine Fallunterscheidung. Wenn $x, y \geqslant 0$, dann ist auch $xy > 0$ (weil $\mathbb{R}$ ein angeordneter Körper ist) und die Gleichung ist unmittelbar wahr. Falls $x \geqslant 0$ und $y < 0$ ist, dann folgt:

$$-y > 0 \Rightarrow -xy = x(-y) \geqslant 0$$

Also ist xy negativ und es gilt:

$$|xy| = -xy = x(-y) = |x||y|$$

Der Fall $x < 0$ und $y \geqslant 0$ läßt sich auf den gerade betrachteten fall zurückführen, indem man einfach die Variablen umbenennt. Der letzte Fall ist $x, y < 0$. Dann gilt aber $xy > 0$ und:

$$|x||y| = (-x)(-y) = xy = |xy|.$$

Damit ist der Beweis für (3) beendet.

(4) Wir beweisen jetzt die Dreiecksungleichung (4). Man erinnere sich, dass für alle $x \in \mathbb{R}$ immer

$$x \leqslant |x| \quad \text{und} \quad -x \leqslant |x|$$

gilt. Wir unterscheiden zwei Fälle:

Sei $a + b \geqslant 0$. Dann folgt:

$$|a + b| = a + b \leqslant |a| + |b|,$$

denn $a \leqslant |a|$ und $b \leqslant |b|$.

Im zweiten Fall sei $a + b < 0$. Dann folgt

$$|a + b| = -(a + b) = -a + (-b) \leqslant |a| + |b|,$$

denn $-a \leqslant |a|$ und $-b \leqslant |b|$. Damit sind alle Möglichkeiten für $a + b$ abgedeckt und die Aussage (4) bewiesen.

(5) Aussage (5) folgt leicht aus (4). Setze $z = x - y$. Dann gilt:

$$|x| = |z + y| \leqslant |z| + |y| = |x - y| + |y|$$

Also:

$$|x| - |y| \leqslant |x - y| \tag{2.20}$$

Setze jetzt $w = y - x$. Dann gilt:

$$|y| = |x + w| \leqslant |x| + |w| = |x| + |y - x| = |x| + |x - y|$$

Also:

$$|y| - |x| \leqslant |x - y| \tag{2.21}$$

Beide Ungleichungen 2.20 und 2.21 zusammen ergeben die Aussage (5). □

Lemma 2.22 (Bernoullische Ungleichung). *Für alle $x \in \mathbb{R}$ mit $x > -1$ und alle $n \in \mathbb{N}$ gilt:*

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

Die Ungleichung stammt von Jakob Bernoulli `https://de.wikipedia.org/wiki/Jakob_I._Bernoulli`.

Nicht zu verwechseln mit
Johann Bernoulli `https://de.wikipedia.org/wiki/Johann_Bernoulli`
oder
Daniel Bernoulli `https://de.wikipedia.org/wiki/Daniel_Bernoulli`

Beweis. Der Beweis erfolgt durch vollständige Induktion über n.

Für $n = 1$ ist die Ungleichung offensichtlich wahr. Das ist der Induktionsanfang.

Wir nehmen jetzt an, dass für ein $n \in \mathbb{N}$ die Aussage des Lemmas gilt. Wir müssen zeigen, dass dann die Aussage auch für $n+1$ richtig ist. Zunächst folgt

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

aus der Induktionsannahme, denn laut Voraussetzung ist $x > -1$, also $1+x > 0$. Dann gilt:

$$(1+x)^{n+1} = (1+x)^n(1+x) \geqslant (1+nx)(1+x) = 1+nx+x+nx^2 \geqslant 1+(n+1)x,$$

denn $nx^2 \geqslant 0$. Damit ist der Induktionsschritt beendet und das Lemma bewiesen. □

2.4 Die komplexen Zahlen

Ausgehend von $\mathbb{R}$ definieren wir jetzt die komplexen Zahlen $\mathbb{C}$.

Definition 2.23. Eine *komplexe Zahl* ist ein Paar $(a, b) \in \mathbb{R} \times \mathbb{R}$, wobei a der *Realteil* und b der *Imaginärteil* der komplexen Zahl genannt wird. Auf der Menge $\mathbb{C}$ der komplexen Zahlen definieren wir eine Addition durch

$$(a, b) + (c, d) = (a + c, b + d)$$

und eine Multiplikation durch

$$(a, b)(c, d) = (ac - bd, ad + bc).$$

Wir identifizieren die komplexe Zahle $(a, 0)$ mit der reellen Zahl a und betrachten auf diese Weise $\mathbb{R}$ als Teilmenge von $\mathcal{C}$. Wir kürzen $(0, 1) = i$ ab und nennen i die *imaginäre Einheit*. Wir schreiben

$$(a, b) = a + ib,$$

$\operatorname{Re}(a + ib) = a$ und $\operatorname{Im}(a + ib) = b$.

Satz 2.24. *Es gilt:*

1. *Die komplexen Zahlen $\mathbb{C}$ bilden einen $\mathbb{R}$-Vektorraum der Dimension* 2.
2. *Desweiteren ist $\mathbb{C}$ ein Körper.*

3. *Die Null des Körpers ist* $(0,0) = 0 + i \cdot 0 = 0$ *und die Eins ist* $(1,0) = 1 + i \cdot 0 = 1$.
4. *Das zu* $z = a + ib \neq 0$ *inverse Element ist*

$$z^{-1} = \frac{a - ib}{a^2 + b^2}.$$

Beweis. Übung. □

Definition 2.25. Wir definieren den *Betrag* einer komplexen Zahl $z = a + ib$ durch

$$|z| = |a + ib| = \sqrt{a^2 + b^2} \in \mathbb{R}.$$

Die zu $a + ib = z$ *konjugierte komplexe Zahl* ist $\overline{z} = a - ib$.

Proposition 2.26. *Es gilt für alle* $z, w \in \mathbb{C}$*:*

1. $z = \overline{z} \;\Leftrightarrow\; \operatorname{Im} z = 0 \;\Leftrightarrow\; z \in \mathbb{R}$
2. $\overline{z + w} = \overline{z} + \overline{w}$, $\overline{z \cdot w} = \overline{z} \cdot \overline{w}$, $\overline{\overline{z}} = z$
3. $|z| \geqslant 0$ *und* $(z = 0 \Longleftrightarrow |z| = 0)$
4. $|z| = |\overline{z}|$ *und* $z \cdot \overline{z} = |z|^2$
5. $z + \overline{z} = 2 \operatorname{Re} z$ *und* $z - \overline{z} = 2i \operatorname{Im} z$
6. *Das zu* $z \neq 0$ *inverse Element ist*

$$z^{-1} = \frac{\overline{z}}{|z|^2}.$$

7. $|z||w| = |zw|$
8. *Es gilt die Dreiecksungleichung:* $\forall\ w, z \in \mathbb{C}$

$$|w + z| \leqslant |w| + |z|.$$

Beweis. Sei $z = x + iy$ und $w = u + iv$. Die Aussagen (1) bis (6) überlassen wir dem Leser als Übung.

(7)

$$\begin{aligned} |z \cdot w|^2 &= (xu - yv)^2 + (xv + yu)^2 = x^2u^2 - 2xuyv + y^2v^2 + x^2v^2 + 2xvyu + y^2u^2 \\ &= (x^2 + y^2)(u^2 + v^2) = |z|^2|w|^2 \end{aligned}$$

Um (8) zu zeigen, bemerken wir zunächst für eine beliebige komplexe Zahhl $r = s + it$ folgende Ungleichung:

$$|\operatorname{Re} r| = |s| = \sqrt{s^2} \leqslant \sqrt{s^2 + t^2} = |r|.$$

Mit $r = z\overline{w}$ und (5) folgt:

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

Weil $|z + w|, |z|^2 + |w|^2 \geqslant 0$, dürfen wir die Wurzel ziehen und Gleichung (8) folgt. □

Die komplexen Zahlen $\mathbb{C} = \mathbb{R}^2$ bilden eine Ebene. Die Addition von komplexen Zahlen ist komponentenweise definiert. Sie ist also genau die Addition von Vektoren in $\mathbb{R}^2$. Dies liefert eine geometrische Interpretation der Addition komplexer Zahlen.

Die Multiplikation hat auch eine geometrische Interpretation. Um diese zu verstehen, führen wir Matrizen ein.

Definition 2.27. Für einen Körper K und natürliche Zahlen m, n sei

$$M_{m\times n}(K)$$

die Menge der $m \times n$-Matrizen mit Koeffizienten im Körper K, d.h. ein Element A dieser Menge ist ein Matrix

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix} = (a_{i,j})_{\substack{i=1,\dots,m \\ j=1,\dots,n}}$$

mit m Zeilen und n Reihen oder Spalten, deren Einträge $a_{1,1}\dots, a_{m,n}$ aus K stammen. Addition zweier Matrizen A und B in $M_{m\times n}(K)$ erfolgt komponentenweise:

$$A + B =: C = (c_{i,j})_{\substack{1\leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} \quad \text{mit} \quad c_{i,j} = a_{i,j} + b_{i,j}.$$

Matrizenmultiplikation funktioniert nach dem bekannten "Zeile mal Spalte"-Prinzip. Dabei ist zu beachten, dass die Multiplikation nur dann Sinn ergibt, wenn die erste Matrix dieselbe Anzahl an Spalten hat wie die zweite an Zeilen. Also:

$$- \cdot -\colon\ M_{\ell.m}(K) \times M_{m,n}(K) \to M_{\ell,n}(K)\ , (A, B) \mapsto A \cdot B$$

mit

$$A \cdot B =: C = (c_{i,j})_{\substack{1\leqslant i \leqslant m \\ 1 \leqslant j \leqslant n}} \quad \text{mit} \quad c_{i,j} = \sum_{k=1}^{m} a_{i,k} b_{k,j}.$$

In der linearen Algebra lernt man, dass für alle $m, n \in \mathbb{N}$ die Matrizen $M_{m\times n}(K)$ einen K-Vektorraum bilden und dass $M_{n\times n}(K)$ eine K-Algebra ist. Wir werden dies hier weder definieren, noch beweisen. Desweiteren lernt man dort, dass man jeder $m \times n$-Matrix eindeutig eine lineare Abbildung $K^n \to K^m$ zuordnen kann (und umgekehrt).

Bemerkung 2.28. Weiter betrachten wir folgende Menge von Matrizen:

$$C := \left\{ A \in M_{2\times 2}(\mathbb{R}) \mid A = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \right\}$$

Man überlegt sich leicht, dass die A zugeordnete lineare Abbildung zuerst Dehnung (oder Stauchung) um den Faktor $\sqrt{a^2 + b^2}$ und dann Drehung (in mathematisch positiven Sinn, also gegen der Uhrzeigersinn) um den Winkel α mit $\sin\alpha = \frac{a}{\sqrt{a^2+b^2}}$ (bzw. $\cos\alpha = \frac{b}{\sqrt{a^2+b^2}}$) ist.

Lemma 2.29. *Die Abbildung*

$$\varphi\colon\ \mathbb{C} \to C\ ,\ \varphi(a + bi) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}$$

ist ein Körperisomorphismus.

Beweis. In den Übungen. □

Damit haben wir es geschafft, auch der Multiplikation komplexer Zahlen eine geometrische Interpretation zu geben. So ist z.B. Multiplikation mit i dasselbe wie Matrixmultiplikation mit

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Dies ist die Drehung gegen den Uhrzeigersinn um 90°. Nun ist Drehen um 180° in $\mathbb{R}^2$ gerade Punktspiegelung am Ursprung und damit dasselbe wie Multiplikation mit -1. Es ist jetzt überhaupt nicht mehr verwunderlich, dass zwei mal Drehen um 90° die Punktspiegelung am Ursprung ergibt: $i^2 = -1$.

Wir kommen jetzt zum wichtigsten Grund, die komplexen Zahlen einzuführen. Erinnern Sie sich an den Pargraphen 2.1. Zuerst erinnern wir an die Definition von Polynomen.

Definition 2.30. Ein *Polynom* mit Variable x und Koeffizienten in $\mathbb{R}$ ist eine Funktion der Form

$$p\colon \mathbb{R} \to \mathbb{R},\ p(x) = \sum_{k=0}^{n} a_k x^k$$

mit Koeffizienten $a_i \in \mathbb{R}$ für $0 \leqslant i \leqslant n$. Analog definieren wir ein Polynom mit komplexen Koeffiezienten $a_i \in \mathbb{C}$. Das Polynom p hat den *Grad* n, wenn $a_n \neq 0$ gilt. Für die Menge dieser Polynome schreiben wir $\mathbb{R}[x]$ oder entsprechend $\mathbb{C}[x]$. Eine *Nullstelle* eines Polynoms p ist ein x_0 mit $p(x_0) = 0$.

Definition 2.31. Ein Körper K heißt *algebraisch abgeschlossen*, wenn jedes nicht-konstante Polynom mit Koeffizienten in K eine Nullstelle hat.

Satz 2.32 (Fundamentalsatz der Algebra)**.** *Der Körper $\mathbb{C}$ ist algebraisch abgeschlossen.*

Der erste Beweis dieses Satzes stammt von
Carl Friedrich Gauß `https://de.wikipedia.org/wiki/Carl_Friedrich_Gau%C3%9F`

Zum Beweis dieses Satzes werden üblicherweise Methoden der Funktionentheorie, der Galoistheorie oder der algebraischen Topologie benutzt. Man kann diesen Satz auch mit elementaren Methoden der reellen Analysis beweisen, aber wir verweisen auf spätere Vorlesungen. Aus dem Fundamentalsatz folgt direkt das

Korollar 2.33. *Zu einem beliebigen Polynom $p \in \mathbb{C}[x]$ vom Grad $n > 0$ gibt (nicht notwendigerweise verschiedene) $z_1, \dots, z_n \in \mathbb{C}$, so dass*

$$p(x) = \prod_{k}^{n} (x - z_k).$$

In Korollar 7.2 werden wir einen Teil der Aussage des Korollars beweisen.

Beispiel 2.34. Das Polynom $p(x) = x^2 + 2$ hat die Nullstellen $\pm i\sqrt{2}$. Es gilt:

$$x^2 + 2 = (x - i\sqrt{2})(x + i\sqrt{2}).$$

Das Polynom

$$x^2 - 2x + 1 = (x-1)^2$$

ist von Grad 2 hat aber nur eine Nullstelle 1. Diese Nullstelle kommt aber "doppelt" vor.

Es gehört zwar nicht zum Stoff der Vorlesung, aber es lohnt sich die Biographie von Évariste Galois `https://en.wikipedia.org/wiki/%C3%89variste_Galois` zu kennen. Er starb mit 22 Jahren in einem Duell, nachdem er noch in der Nacht zuvor im Wissen, dass er sterben wird, seine bahnbrechenden mathematischen Erkenntnisse aufgeschrieben hatte.

2.5 Binomialformeln

Definition 2.35. Wir definieren für ein $n \in \mathbb{N}$ den Ausdruck

$$n! = 1 \cdot 2 \cdot \ldots \cdot (n-1) \cdot n = \prod_{i=1}^{n} i.$$

Sprich: "n Fakultät". Wir definieren außerdem:

$$0! = 1.$$

Es gilt $n! \cdot (n+1) = (n+1)!$ für alle $n \in \mathbb{N}_0$.

Definition 2.36. Seien $n, k \in \mathbb{N}_0$ mit $n \geqslant k$. Wir definieren die *Binomialkoeffizienten*:

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

Wir definieren außerdem für $k \in \mathbb{Z}, k < 0$ und für $k \in \mathbb{N}, k > n+1$: $\binom{n}{k} = 0$.

Bemerkung 2.37. Es gilt für alle $n, k \in \mathbb{N}_0$:

$$\begin{aligned} \tbinom{n}{k} &= \tbinom{n}{n-k} & \tbinom{n}{0} &= \tbinom{n}{n} = 1 \\ \tbinom{n}{1} &= \tbinom{n}{n-1} = n & \tbinom{n}{2} &= \tfrac{n(n-1)}{2} \end{aligned}$$

Binomialkoeffizienten tauchen sehr häufig auf. Dies liegt an ihrer kombinatorischen Interpretation: $\binom{n}{k}$ ist die Anzahl der k-elementigen Teilmengen einer Menge mit n Elementen.

Lemma 2.38. *Für alle $n, k \in \mathbb{N}_0$ gilt:*

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

Dieses Lemma ist Grundlage für die Berechnung von Binomialkoeffizienten im Pascalschen Dreieck.

Blaise Pascal `https://de.wikipedia.org/wiki/Blaise_Pascal#Mathematik`

Beweis. Man rechnet einfach nach:

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

□

Lemma 2.39 (Binomialformel)**.** *Seien $a, b \in \mathbb{C}$ und $n \in \mathbb{N}_0$. Dann gilt:*

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

Beweis. Wir beweisen die Aussage durch Induktion über n. Die Aussage für $n = 0$ reduziert sich auf die wahre Gleichung

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

wegen unserer Definition $0! = 1$. Der Fall $n = 1$ ist auch leicht:

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

Für den Induktionsschritt nehmen wir an, dass für ein $n \in \mathbb{N}$ die Binomialformel schon bewiesen ist. Zu zeigen ist jetzt:

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

Wir beginnen auf der linken Seite:

$$
\begin{aligned}
(a+b)^{n+1} = (a+b)(a+b)^n &= (a+b)\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k} \\
&= a\left(\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}\right) + b\left(\sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k}\right) \\
&= \sum_{k=0}^{n}\binom{n}{k}a^{k+1}b^{n-k} + \sum_{k=0}^{n}\binom{n}{k}a^k b^{n-k+1} \\
&= a^{n+1} + \sum_{k=0}^{n-1}\binom{n}{k}a^{k+1}b^{n-k} + \sum_{k=1}^{n}\binom{n}{k}a^k b^{n-k+1} + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n}\binom{n}{k-1}a^k b^{n-k+1} + \sum_{k=1}^{n}\binom{n}{k}a^k b^{n-k+1} + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n}\left(\binom{n}{k-1} + \binom{n}{k}\right) a^{k+1}b^{n-k} + b^{n+1}
\end{aligned}
$$

Wir benutzen jetzt die Formel im Pascalschen Dreieck aus Lemma 2.38.

$$
\begin{aligned}
&= a^{n+1} + \sum_{k=1}^{n}\binom{n+1}{k}a^{k+1}b^{n-k} + b^{n+1} \\
&= \sum_{k=0}^{n+1}\binom{n+1}{k}a^k b^{n+1-k}
\end{aligned}
$$

□

Lemma 2.40. *Seien $a, b \in \mathbb{C}$ und $n \in \mathbb{N}_0$. Dann gilt:*

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

Beweis. Es gilt:

$$
\begin{aligned}
(a-b)\sum_{k=0}^{n} a^k b^{n-k} &= \sum_{k=0}^{n} a^{k+1} b^{n-k} - \sum_{k=0}^{n} a^k b^{n-k+1} \\
&= a^{n+1} + \sum_{k=0}^{n-1} a^{k+1} b^{n-k} - \sum_{k=1}^{n} a^k b^{n-k+1} - b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^{n} a^k b^{n-k+1} - \sum_{k=1}^{n} a^k b^{n-k+1} - b^{n+1} \\
&= a^{n+1} - b^{n+1}
\end{aligned}
$$

□

Lemma 2.41 ((endliche) geometrische Reihe)**.** *Für alle $x \in \mathbb{C}$ und $n \in \mathbb{N}$ gilt*

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

Für alle $x \in \mathbb{C} - \{-1\}$ gilt damit:

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

Beweis. Die erste Gleichung folgt aus Lemma 2.40, wenn man $a = 1$ setzt. Die zweite Gleichung ist dann offensichtlich. □

3 Folgen

3.1 Folgen und beschränkte Folgen

Definition 3.1. Sei B eine Menge. Eine *Folge in B* ist eine Abbildung $a\colon \mathbb{N} \to B$. Wir schreiben dann oft $a(n) = a_n$ für einzelne Folgeglieder und $(a_n)_{n\in\mathbb{N}}$ für die ganze Folge.

Im Gegensatz zu einer Menge kommt es bei einer Folge auf die Reihenfolge an und es kann Wiederholungen der Folgeglieder geben. So ist z.B.

$$\forall\, n \in \mathbb{N}\ a_n = 5$$

die konstante Folge mit Wert 5: $(5, 5, 5, \ldots)$.

Definition 3.2. Eine Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ heißt *nach oben beschränkt*, wenn es $S \in \mathbb{R}$ gibt, so dass für alle $n \in \mathbb{N}$ gilt: $x_n \leqslant S$. In diesem Fall nennen wir S eine *obere Schranke* für $(x_n)_{n\in\mathbb{N}}$.

Analog definieren wir *nach unten beschränkte Folgen* und *untere Schranken*:

$$\exists\, S \in \mathbb{R}\ \forall\, n \in \mathbb{N} : x_n \geqslant S.$$

Eine Folge in $\mathbb{R}$ heißt *beschränkt*, wenn sie nach unten und oben beschränkt ist. Eine Folge ist *unbeschränkt*, wenn sie nicht beschränkt ist.

Wir definieren denselben Begriff für Mengen. Eine Menge M heißt *nach oben beschränkt*, wenn es eine *obere Schranke* S gibt, also

$$\exists\, S \in \mathbb{R}\ \forall\, m \in M : m \leqslant S.$$

Analog definieren wir *nach unten beschränkte* und *beschränkte Mengen*.

Lemma 3.3. *Sei $(x_n)_{n\in\mathbb{N}}$ eine reelle beschränkte Folge. Dann existiert ein $T \in \mathbb{R}$, so dass für alle $n \in \mathbb{N}$ gilt: $|x_n| \leqslant T$.*

Beweis. Wenn die reelle Folge $(x_n)_{n\in\mathbb{N}}$ beschränkt ist, so existiert nach Definition eine untere Schranke S_1 und eine obere Schranke S_2. Setze $T = \max\{|S_1|, |S_2|\}$. Für alle $n \in \mathbb{N}$ folgt dann

$$S_1 \leqslant x_n \leqslant S_2 \implies |x_n| \leqslant \max\{|S_1|, |S_2|\} = T.$$

□

Die Formulierung des Lemmas eignet sich auch für Folgen in $\mathbb{C}$.

Definition 3.4. Eine Folge $(z_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ heißt *beschränkt*, wenn es ein $T \in \mathbb{R}$ gibt, so dass für alle $n \in \mathbb{N}$ gilt: $|z_n| \leqslant T$. Dieses T heißt dann eine *Schranke* für $(z_n)_{n\in\mathbb{N}}$.

3.2 Konvergenz und Grenzwerte

Folgen tauchen dann auf, wenn man einen Prozess immer wieder wiederholt. Dies dient meistens der Approximation einer unbekannten Größe. Um dies Idee zu formalisieren, führen wir den Begriff des Grenzwerts oder Limes ein.

Definition 3.5. Sei $(x_n)_{n\in\mathbb{N}}$ ein Folge in $\mathbb{R}$ (oder in $\mathbb{C}$). Wir sagen, dass diese Folge *konvergiert*, wenn gilt:

$$\exists\, x \in \mathbb{R}\ \forall\, \varepsilon > 0\ \exists\, N \in \mathbb{N}\ \forall\, n \geqslant N : |x - x_n| < \varepsilon.$$

In diesem Fall heißt x der *Grenzwert* oder *Limes* der Folge $(x_n)_{n\in\mathbb{N}}$. Eine Folge, die nicht konvergiert, heißt *divergent*.

Proposition 3.6. *Der Grenzwert einer konvergenten Folge ist eindeutig bestimmt. Sei also $(x_n)_{n\in\mathbb{N}}$ ein konvergente Folge mit Grenzwert a und b. Dann folgt $a = b$.*

Beweis. Sei $\varepsilon > 0$ gegeben.

$$|a - b| = |a - x_n + x_n - b| \leqslant |a - x_n| + |b - x_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Es ist $|a - b| \geqslant 0$, aber kleiner als jede positive Zahl. Also ist $a = b$. □

Obwohl der Beweis der vorigen Proposition sehr kurz und einfach ist, ist er doch bemerkenswert. Wir schließen aus der Aussage

$$\forall\ \varepsilon > 0 : |x| < \epsilon$$

die scheinbar viel stärkere Aussage $x = 0$ (im Beweis ist natürlich $x = a - b$). Dieser Schluß ist möglich, weil $|x| \in \mathbb{R}$ und $\mathbb{R}$ ein angeordneter Körper ist. Eine reelle Zahl ist entweder positiv oder negativ oder Null, hat aber nicht gleichzeitig zwei oder gar drei dieser Eigenschaften. Weil $|x| \geqslant 0$ für alle reellen (oder komplexen!) Zahlen gilt, $|x|$ aber eben auch kleiner als jede positive reelle Zahl ist, muss $|x|$, und damit x selber, Null sein.

Definition 3.7. Sei $(x_n)_{n\in\mathbb{N}}$ ein konvergente Folge. Dann schreiben wir auch $\lim_n x_n$ für ihren Grenzwert.

Bemerkung 3.8. Der Begriff der Konvergenz ändert sich nicht, wenn man in der Definition $<$ durch $\leqslant$ ersetzt.

Der folgende Definition ist eine nützliche Abkürzung.

Definition 3.9. Eine konvergent Folge, die gegen 0 konvergiert, heißt *Nullfolge.*

Wie läßt sich Konvergenz verstehen?

Erstens, um zu zeigen, dass eine Folge konvergiert, müssen Sie ihren Grenzwert schon kennen: $\exists x \in \mathbb{R}$ bedeutet, dass Sie x hinschreiben müssen, um den Beweis der Konvergenz überhaupt anfangen zu können.

Danach müssen Sie sich ein $\varepsilon \in \mathbb{R}, \varepsilon > 0$ hernehmen, und dann für dieses ε den Rest der Definition zeigen. Weil Sie an das ε keine weitere Bedingung gestellt haben, haben Sie dann die Aussage für alle $\varepsilon > 0$ bewiesen.

Das ε ist eine obere Schranke für die Abweichung der Folgeglieder von Grenzwert. Wenn ε sehr klein ist, dann sagt die Definition der Konvergenz, dass nicht unbedingt alle Folgeglieder im Intervall $]a - \varepsilon, a + \varepsilon[$ liegen müssen; aber es muss ein (vielleicht sehr großes) N geben, ab dem alle Folgeglieder $a_n, n \geqslant N$ in diesem kleinen Intervall liegen.

Lesen (und verstehen) Sie die folgende Aussage und ihren Beweis.

Lemma 3.10. *Die Folge* $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ *konvergiert und es gilt:* $\lim\limits_{n} \frac{1}{n} = 0.$

Beweis. Sei $\varepsilon > 0$ gegeben. Aufgrund der archimedischen Anordnung von $\mathbb{R}$ existiert $N \in \mathbb{N}$ mit $\frac{1}{\varepsilon} < N$. Dann gilt aber für alle $n \geqslant N$:

$$|0 - \frac{1}{n}| = \frac{1}{n} \leqslant \frac{1}{N} < \varepsilon.$$

Also konvergiert die Folge $(\frac{1}{n})_{n\in\mathbb{N}}$ gegen 0. □

Zu diesen Beweis: Wir kennen den Grenzwert der Folge (hier 0) noch bevor wir den Beweis anfangen. Dann geben wir uns ein beliebiges $\varepsilon > 0$ vor. Abhängend von diesem ε (man beachte die Reihenfolge der Quantoren!) bestimmen wir explizit ein N (hier mit Hilfe des archimedischen Axioms), so dass ab dem N-ten Folgeglied alle nachfolgenden Glieder in $]-\varepsilon, \varepsilon[$ liegen. Man sieht hier, dass für die Konvergenz nicht wichtig ist, was am Anfang der Folge passiert. Für $\varepsilon = \frac{1}{10}$ sind die ersten 9 Folgeglieder $1, \frac{1}{2}, \ldots, \frac{1}{9}$ nicht im Intervall $]-\varepsilon, \varepsilon[=]-\frac{1}{10}, \frac{1}{10}[$, danach aber alle. Für $\varepsilon = \frac{1}{100}$ sind die ersten 99 Folgeglieder nicht im Intervall $]-\varepsilon, \varepsilon[$, danach aber alle.

Beispiel 3.11. Die Folge

$$x_n = \begin{cases} 1 & , \text{ für } n \text{ gerade} \\ -1 & , \text{ sonst} \end{cases}$$

konvergiert nicht. Um das zu beweisen, zeigen wir die Verneinung der Definition der Konvergenz:

$$\forall\, x \in \mathbb{R}\ \exists\, \varepsilon > 0\ \forall\, N \in \mathbb{N}\ \exists\, n \geqslant N : |x - x_n| \geqslant \varepsilon.$$

Sei also ein $x \in \mathbb{R}$ gegeben.

1. Fall $x \geqslant 0$: Setze $\varepsilon = 1$. Für ein gegebenes N betrachte eine ungerade Zahl n größer als N. Dann gilt

$$|x - x_n| = x + 1 \geqslant 1.$$

2. Fall $x < 0$: Setze $\varepsilon = 1$. Für ein gegebenes N betrachte eine gerade Zahl n größer als N. Dann ist $x - 1 < 0$ und es gilt

$$|x - x_n| = |x - 1| = 1 - x \geqslant 1,$$

denn $-x > 0$.

Bemerkung 3.12. Man beachte auch, dass sich das Konvergenz- (oder Divergenz-)Verhalten einer Folge nicht ändert, wenn man endlich viele Folgeglieder abändert. So konvergiert z.B. die Folge

$$1, \frac{1}{2}, 15, 104, \frac{1}{5}, \frac{1}{6}, \frac{1}{7}, \ldots$$

genauso gegen Null wie die Folge $(1/n)_{n\in\mathbb{N}}$.

3.3 Bestimmte Divergenz reeller Folgen

Für reelle Folgen können wir den Begriff der Divergenz etwas genauer fassen.

Definition 3.13. Eine Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ heißt *bestimmt divergent gegen* ∞, wenn es zu jedem $M \in \mathbb{R}$ ein $N \in \mathbb{N}$ gibt, so dass für alle $n \geqslant N$ gilt: $M \leqslant x_n$.

$$\forall\, M \in \mathbb{R}\ \exists\, N \in \mathbb{N}\ \forall\, n \geqslant N : M \leqslant x_n.$$

Wir schreiben dann $\lim_n x_n = \infty$. (Wir sagen auch: "Die Folge $(x_n)_{n\in\mathbb{N}}$ wächst über alle Schranken.")

Eine Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ heißt *bestimmt divergent gegen* $-\infty$, wenn gilt:

$$\forall\, M \in \mathbb{R}\ \exists\, N \in \mathbb{N}\ \forall\, n \geqslant N : x_n \leqslant M.$$

Wir schreiben dann auch $\lim_n x_n = -\infty$.

Beispiel 3.14. (Nicht-)Beispiel für bestimmte Divergenz.

1. Die Folge $(n)_{n\in\mathbb{N}} = (1, 2, 3, \ldots)$ ist bestimmt divergent gegen $+\infty$.
2. Die Folge aus Beispiel 3.11 divergiert, ist aber nicht bestimmt divergent.
3. Die Folgen

$$(1, -1, 2, -1, 3, -1, 4, -1, \ldots) \quad \text{und} \quad (1, -1, 2, -2, 3, -3, \ldots)$$

divergieren, divergieren aber nicht bestimmt.

3.4 Einige allgemeine Sätze zur Konvergenz

Satz 3.15. *Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergente Folgen in $\mathbb{R}$ oder $\mathbb{C}$. Dann gilt:*

1. *Die Folgen $(a_n + b_n)_{n\in\mathbb{N}}$ konvergiert und*
$$\lim_n (a_n + b_n) = \lim_n a_n + \lim_n b_n$$

2. *Die Folgen $(a_n \cdot b_n)_{n\in\mathbb{N}}$ konvergiert und*
$$\lim_n (a_n \cdot b_n) = (\lim_n a_n) \cdot (\lim_n b_n).$$
Insbesondere folgt für alle $\lambda \in \mathbb{C}$:
$$\lim_n (\lambda \cdot b_n) = \lambda \cdot (\lim_n b_n).$$

3. *Wenn $(b_n)_{n\in\mathbb{N}}$ nicht gegen 0 konvergiert, so ist auch die Folge $(a_n/b_n)_{n\in\mathbb{N}}$ konvergent und*
$$\lim_n \frac{a_n}{b_n} = \frac{\lim_n a_n}{\lim_n b_n}.$$

Bevor wir den Satz beweisen, geben wir zunächste ein Beispiel, in dem man sieht, dass die Bedingung in (2), dass $(b_n)_{n\in\mathbb{B}}$ keine Nullfolge sein darf, notwendig ist.

Beispiel 3.16. Die Bedingung in (2), dass $(b_n)_{n\in\mathbb{N}}$ keine Nullfolge ist, ist essentiell, wie man an folgenden Beispielen sehen kann:

1. Für $(a_n)_{n\in\mathbb{N}} = (1)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}} = (\frac{1}{n})_{n\in\mathbb{N}}$ divergiert die Folge
$$n \mapsto a_n/b_n = \frac{1}{1/n} = n.$$

2. Oder etwas subtiler: Für $(a_n)_{n\in\mathbb{N}} = (\frac{1}{n})_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}} = (\frac{1}{n^2})_{n\in\mathbb{N}}$ divergiert die Folge
$$n \mapsto a_n/b_n = \frac{1}{n}/\frac{1}{n^2} = n.$$

Beweis. von Satz 3.15: Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergente Folgen in $\mathbb{R}$ oder $\mathbb{C}$ mit Grenzwerten a und b.

1. Zu zeigen ist, dass die Folge $(a_n + b_n)_{n\in\mathbb{N}}$ gegen $a + b$ konvergiert. Sei $\varepsilon > 0$. Aus $\lim_n a_n = a$ folgt, dass es $N_1 \in \mathbb{N}$ gibt, so dass für alle $n \in \mathbb{N}$ mit $n \geqslant N_1$ gilt:
$$|a - a_n| < \varepsilon/2.$$
Aus $\lim_n b_n = b$ folgt, dass es $N_2 \in \mathbb{N}$ gibt, so dass für alle $n \in \mathbb{N}$ mit $n \geqslant N_2$ gilt:
$$|b - b_n| < \varepsilon/2.$$
Dann folgt für alle $n \in \mathbb{N}$ mit $n \geqslant \max\{N_1, N_2\}$ mit der Dreiecksungleichung:
$$|(a + b) - (a_n + b_n)| = |(a - a_n) + (b - b_n)| \leqslant |a - a_n| + |b - b_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Damit haben wir $\lim_n(a_n + b_n) = a + b$ gezeigt.

2. Als nächstes ist zu zeigen, dass die Folge $(a_n b_n)_{n\in\mathbb{N}}$ gegen ab konvergiert. Sei $\varepsilon > 0$ gegeben.

Wir betrachten zunächst folgende Identität:

$$ab - a_n b_n = a(b - b_n) + b(a - a_n) - (a - a_n)(b - b_n)$$

Daraus erhalten wir durch die Dreiecksungleichung die folgende Abschätzung:

$$|ab - a_n b_n| \leqslant |a||b - b_n| + |b||a - a_n| + |a - a_n||b - b_n| \tag{3.17}$$

Wegen der Konvergenz der Folge $(a_n)_{n\in\mathbb{N}}$ gibt es ein $N_1 \in \mathbb{N}$, so dass für alle $n \geqslant N_1$

$$|a - a_n| < \min\{\frac{\varepsilon}{3|b|}, \sqrt{\frac{\varepsilon}{3}}\}$$

gilt. Man beachte, dass $\frac{\varepsilon}{3|b|}$ für $b = 0$ nicht definiert. Die Abschätzung gegen $\frac{\varepsilon}{3|b|}$ wird gebraucht, um den Summanden $|b||a - a_n|$ in (3.17) unter Kontrolle zu bekommen. Für $b = 0$ fällt dieser Term weg. Also ist in diesem Fall die Abschätzung gar nicht nötig, und wir können wie im Fall $b \neq 0$ weitermachen.

Weil die Folge $(b_n)_{n\in\mathbb{N}}$ gegen b konvergiert, gibt es ein $N_2 \in \mathbb{N}$, so dass für alle $n \geqslant N_2$

$$|b - b_n| < \min\{\frac{\varepsilon}{3|a|}, \sqrt{\frac{\varepsilon}{3}}\}$$

gilt. Wie oben muss man hier erwähnen, dass für $a = 0$ die Abschätzung nicht funktioniert, aber eben auch nicht gebraucht wird, denn der relevante Term $|a||b - b_n|$ in (3.17) fällt einfach weg.

Setze jetzt $n = \max\{N_1, N_2\}$. Dann gilt für alle $n \geqslant N$:

$$\begin{aligned} |ab - a_n b_n| &\leqslant |a||b - b_n| + |b||a - a_n| - |a - a_n||b - b_n| \\ &< |a|\frac{\varepsilon}{3|a|} + |b|\frac{\varepsilon}{3|b|} + \sqrt{\frac{\varepsilon}{3}} \cdot \sqrt{\frac{\varepsilon}{3}} = \varepsilon \end{aligned}$$

Damit haben wir $\lim_n(a_n b_n) = ab$ gezeigt.

Die Aussage, dass für alle $\lambda \in \mathbb{C}$

$$\lim_n(\lambda \cdot b_n) = \lambda \cdot (\lim_n b_n)$$

gilt, folgt aus dem vorhergenden Beweis, indem man $a_n = \lambda$ für alle $n \in \mathbb{N}$ setzt.

3. Sei jetzt zusätzlich $\lim_n b_n = b \neq 0$. Wir zeigen zunächst den Spezialfall, in dem $(a_n)_{n\in\mathbb{N}}$ die konstante Folge bei $a = 1$ ist. Zu zeigen ist also, dass $(1/b_n)_{n\in\mathbb{N}}$ gegen $1/b$ konvergiert.

Wegen $\lim_n b_n = b$ gibt es ein $N_3 \in \mathbb{N}$, so dass für alle $n \in \mathbb{N}, n \geqslant N_3$ gilt:

$$|b - b_n| < \frac{|b|}{2} =: C.$$

Es folgt

$$|b| = |b - b_n + b_n| = |b - b_n| + |b_n| < \frac{|b|}{2} + |b_n|,$$

und damit

$$C = \frac{|b|}{2} < |b_n|.$$

Wir haben also eine untere Schranke C für $|b_n|$ gefunden, die spätestens ab dem Folgeglied mit der Nummer N_3 gilt. Mit dieser Schranke können wir jetzt die Konvergenz von $n \mapsto a_n/b_n$ zeigen. Es gilt für alle $n \geqslant N_3$:

$$|1/b - 1/b_n| = \left|\frac{b_n - b}{bb_n}\right| < \frac{1}{C|b|}|b_n - b|$$

Man beachte jetzt, dass der Term $C_1 = \frac{1}{C|b|} = \frac{2}{b^2}$ konstant ist, d.h. nicht von n abhängt. Weil $(b_n)_{n\in\mathbb{N}}$ nach Voraussetzung gegen b konvergiert, gibt es ein $N_4 \in \mathbb{N}$, so dass für alle $n \geqslant N_4$ gilt:

$$|b - b_n| < \varepsilon/C_1.$$

Insgesamt folgt für alle $n \geqslant \max\{N_3, N_4\}$:

$$|1/b - 1/b_n| < C_1|b_n - b| < C_1 \cdot \frac{\varepsilon}{C_1} = \varepsilon$$

Also konvergiert die Folge $(1/b_n)_{n\in\mathbb{N}}$ gegen $1/b$ und der Spezialfall ist bewiesen.

Für eine beliebige konvergente Folge $(a_n)_{n\in\mathbb{N}}$ mit Grenzwert a folgt jetzt aus Teil 1 und dem Spezielfall, dass die Folge mit Folgegliedern $a_n/b_n = a_n \cdot \frac{1}{b_n}$ gegen

$$\frac{\lim_n a_n}{\lim_n b_n} = \frac{a}{b}$$

konvergiert, solange $b \neq 0$. □

Proposition 3.18. *Sei $(x_k)_{k\in\mathbb{N}}$ eine reelle Folge. Dann gilt:*

1. *Wenn $(x_k)_{k\in\mathbb{N}}$ bestimmt gegen $+\infty$ oder $-\infty$ konvergiert, dann ist $(1/x_k)_{k\in\mathbb{N}}$ eine Nullfolge.*
2. *Wenn $(x_k)_{k\in\mathbb{N}}$ eine Nullfolge und $x_k > 0$ für alle $k \in \mathbb{N}$ gilt, dann ist $(1/x_k)_{k\in\mathbb{N}}$ bestimmt divergent gegen $+\infty$.*
3. *Wenn $(x_k)_{k\in\mathbb{N}}$ eine Nullfolge und $x_k < 0$ für alle $k \in \mathbb{N}$ gilt, dann ist $(1/x_k)_{k\in\mathbb{N}}$ bestimmt divergent gegen $-\infty$.*

Beweis. Zu (1): Wir zeigen zuerst, dass $(1/x_k)_{k\in\mathbb{N}}$ eine Nullfolge ist, wenn $(x_k)_{k\in\mathbb{N}}$ bestimmt gegen $+\infty$ divergiert. Sei $\varepsilon > 0$. Nach Voraussetzung gibt es dann ein $N \in \mathbb{N}$, so dass für alle $n \geqslant N$

$$x_k > \frac{1}{\varepsilon}$$

gilt. Dann gilt aber für alle $n \geqslant N$

$$\frac{1}{x_k} < \varepsilon,$$

und damit ist $(1/x_k)_{k\in\mathbb{N}}$ eine Nullfolge.

Der Beweis für eine gegen $-\infty$ bestimmt divergente Folge geht genauso.

Zu (2): Sei $(x_k)_{k\in\mathbb{N}}$ eine Nullfolge mit $x_k > 0$ für alle $k \in \mathbb{N}$. Wir müssen zeigen, dass $(1/x_k)_{k\in\mathbb{N}}$ bestimmt gegen $+\infty$ divergiert. Sei dazu ein $M \in \mathbb{R}, M > 0$ gegeben. Weil $(x_k)_{k\in\mathbb{N}}$ eine Nullfolge ist, gibt es ein $N \in \mathbb{N}$, so dass für alle $k \geqslant N$ gilt:

$$\left|\frac{1}{x_k}\right| < \frac{1}{M}.$$

Weil für alle $k \in \mathbb{N}$ die Glieder x_k postitiv sind, folgt:

$$\frac{1}{x_k} = \left|\frac{1}{x_k}\right| < \frac{1}{M}. \tag{3.19}$$

Daraus ergibt sich aber wieder für alle $k \geqslant N$:

$$x_k > M.$$

Das war gerade zu zeigen.

Zu (3): Im Fall, in dem alle x_k negative sind, folgt anstelle von (3.19) die Abschätzung

$$-\frac{1}{x_k} = \left|\frac{1}{x_k}\right| < \frac{1}{M}.$$

Daraus ergibt sich für alle $k \geqslant N$:

$$\frac{1}{x_k} > -\frac{1}{M} \quad \Rightarrow \quad x_k < -M$$

Das zeigt, dass die Folge $(x_k)_{k\in\mathbb{N}}$ gegen $-\infty$ divergiert. □

Satz 3.20. *Sei $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge und $(b_n)_{n\in\mathbb{N}}$ eine beschränkte Folge. Dann ist die Folge $(a_n \cdot b_n)_{n\in\mathbb{N}}$ eine Nullfolge.*

Beweis. Zu zeigen ist, dass für beliebiges $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass für alle $n \geqslant N$ gilt: $|a_n b_n| < 0$.

Sei also ein $\varepsilon > 0$ gegeben. Weil $(b_n)_{n\in\mathbb{N}}$ beschränkt ist, existiert ein $C \in \mathbb{R}$, so dass für alle $n \in \mathbb{N}$ gilt: $|b_n| < C$. Weil $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge ist, gibt es ein $N \in \mathbb{N}$, so dass für alle $n \geqslant N$ gilt: $|a_n| < \frac{\varepsilon}{C}$. Insgesamt folgt:

$$|a_n b_n| < C|a_n| < C\frac{\varepsilon}{C} = \varepsilon.$$

□

Satz 3.21. *Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergente Folgen mit den Grenzwerten a und b. Wenn für alle $n \in \mathbb{N}$ die Ungleichung*

$$a_n \leqslant b_n$$

gilt, dann folgt $a \leqslant b$.

Beweis. Wir zeigen, dass aus $a > b$ folgt, dass es ein $n \in \mathbb{N}$ gibt mit $a_n > b_n$. Dies widerspricht der Voraussetzung im Satz.

Sei $a > b$. Wähle $\varepsilon = \frac{a-b}{2}$. Dann existiert $N \in \mathbb{N}$, so dass für alle $n \geqslant N$ gilt:

$$|a - a_n| < \varepsilon \quad \text{und} \quad |b - b_n| < \varepsilon.$$

(Aus der Konvergenz der beiden Folgen bekommt man a priori zwei verschiedene Werte für N, dann kann man aber einfach das größere der beiden wählen.) Insbesondere gilt dann aber

$$a_N \in]a - \varepsilon, a + \epsilon[\quad \text{und} \quad b_N \in]b - \varepsilon, b + \varepsilon[$$

und

$$a - \varepsilon = a - \frac{a-b}{2} = \frac{a+b}{2} = b + \frac{a-b}{2} = b + \varepsilon.$$

Die beiden Intervalle überschneiden sich nicht und es gilt:

$$a_N > \frac{a+b}{2} > b_N.$$

□

Bemerkung 3.22. Seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergente Folgen mit den Grenzwerten a und b. Man beachte, dass man aus der Ungleichung $a_n < b_n$ für alle $n \in \mathbb{N}$ nicht $a < b$ folgern kann.

Ein einfaches Gegenbeispiel ist $b_n = \frac{1}{n}$ und $a_n = 0$ für alle $n \in \mathbb{N}$. Es gilt aber: $\lim_n b_n = 0 = \lim_n a_n$.

Satz 3.23 (Sandwichsatz)**.** *Seien $(a_n)_{n\in\mathbb{N}}, (b_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ Folgen in $\mathbb{R}$, so dass für alle $n \in \mathbb{N}$ gilt:*

$$a_n \leqslant b_n \leqslant c_n$$

Wenn beide Folgen $(a_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ konvergieren und $\lim_n a_n = \lim_n c_n = d$ gilt, dann konvergiert auch die Folge $(b_n)_{n\in\mathbb{N}}$ gegen d.

Beweis. Sei $\varepsilon > 0$. Zu zeigen ist, dass es ein N gibt, so dass für alle $n \geqslant N$

$$|d - b_n| < \varepsilon$$

gilt. Wir wissen aber, dass ein N existiert, so dass für alle $n \geqslant N$

$$|d - a_n| < \varepsilon \quad \text{und} \quad |d - c_n| < \varepsilon$$

gelten. Damit sind beide a_n und b_n im Intervall $]d - \varepsilon, d + \varepsilon[$ enthalten. Es gilt damit für alle $n \geqslant N$:

$$b_n \in]a_n, c_n[\subset]d - \varepsilon, d + \varepsilon[.$$

Dies war zu zeigen. □

Lemma 3.24. *Eine Folge in $\mathbb{C}$ konvergiert genau dann, wenn die Folgen der Real- und Imaginärteile in $\mathbb{R}$ konvergieren.*

Beweis. Sei $(z_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{C}$, die gegen $z = a + ib$ konvergiert. Wir zeigen, dass dann die Folge $(a_n)_{n\in\mathbb{N}}$ gegeben durch

$$a_n = \mathrm{Re}\,(z_n)$$

und $(b_n)_{n\in\mathbb{N}}$ gegeben durch

$$b_n = \mathrm{Im}\,(z_n)$$

jeweils gegen $a = \mathrm{Re}\,(z)$ und $b = \mathrm{Im}\,(z)$ konvergieren. Sei $\varepsilon > 0$. Dann existiert ein $N \in \mathbb{N}$, so dass für alle $n \geqslant N$

$$|a - a_n| = |\mathrm{Re}\,(z) - \mathrm{Re}\,(z_n)| = |\mathrm{Re}\,(z - z_n)| \leqslant |z - z_n| < \varepsilon$$

und

$$|b - b_n| = |\mathrm{Im}\,(z) - \mathrm{Im}\,(z_n)| = |\mathrm{Im}\,(z - z_n)| \leqslant |z - z_n| < \varepsilon$$

gilt. Dies zeigt die Behauptung und damit eine Richtung des Lemmas.

Seien jetzt $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ Folgen in $\mathbb{R}$, die gegen a und b konvergieren. Wir müssen zeigen, dass dann

$$\lim_{n\to\infty}(a_n + ib_n) = a + ib = z$$

gilt. Diese Aussage folgt aber direkt aus Satz 3.15. □

Wir geben zum Schluß dieses Abschnitts noch eine nützliche Umformulierung der Konvergenz.

Lemma 3.25. *Sei $(\varepsilon_k)_{k\in\mathbb{N}}$ eine Nullfolge in $\mathbb{R}$ mit $\varepsilon_k > 0$ für alle $k \in \mathbb{N}$. Eine Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ oder $\mathbb{C}$ konvergiert genau dann gegen x, wenn es für jedes $k \in \mathbb{N}$ ein $N_k \in \mathbb{N}$ gibt, so dass für alle $n \geqslant N_k$*

$$|x - x_n| < \varepsilon_k.$$

gilt.

Beweis. Wenn $(x_n)_{n\in\mathbb{N}}$ gegen x konvergiert, dann existiert zu jedem $\varepsilon_k > 0$ ein $N_k \in \mathbb{N}$, so dass für alle $n \geqslant N_k$ gilt:

$$|x - x_n| < \varepsilon_k.$$

Damit impliziert die Konvergenz die alternative Formulierung.

Um die Rückrichtung zu zeigen, nehmen wir die alternative Formulierung der Konvergenz im Lemma. Sei $(\varepsilon_k)_{k\in\mathbb{N}}$ eine Nullfolge in $\mathbb{R}$ mit $\varepsilon_k > 0$ für alle $k \in \mathbb{N}$. Sei $(x_n)_{n\in\mathbb{N}}$ ein Folge, so dass für jedes $k \in \mathbb{N}$ ein $N_k \in \mathbb{N}$ gibt, so dass für alle $n \geqslant N_k$

$$|x - x_n| < \varepsilon_k.$$

Zu einem beliebigen $\varepsilon > 0$ gibt es also ein $k \in \mathbb{N}$, so dass $0 < \varepsilon_k < \varepsilon$ gilt, denn $(\varepsilon_k)_{k\in\mathbb{N}}$ ist eine Nullfolge. Dann gibt es aber nach Voraussetzung ein $N_k \in \mathbb{N}$, so dass für alle $n \geqslant N_k$

$$|x - x_n| < \varepsilon_k < \varepsilon$$

gilt. Das bedeutet gerade $\lim_n x_n = x$. □

Bemerkung 3.26. Sei $(x_n)_{n\in\mathbb{N}}$ ein Folge, von der wir zeigen wollen, dass sie gegen x konvergiert. Das vorige Lemma besagt, dass es z.B. ausreicht zu zeigen, dass es für alle $k \in \mathbb{N}$ ein $N_k \in \mathbb{N}$ gibt, so dass für alle $n \geqslant N_k$

$$|x - x_n| < \frac{1}{k}$$

oder auch

$$|x - x_n| < \frac{1}{2^k}$$

gilt. Es reicht z.B. auch zu zeigen, dass für alle $n \in \mathbb{N}$

$$|x - x_n| < \frac{1}{n}$$

gilt, obwohl das nicht immer zu erwarten ist. Mit Hilfe des Lemmas kann man die Menge der $\varepsilon > 0$ diskretisieren und dann dem Sachverhalt angepasst wählen. Die wichtige Bedingung an die ε_k ist natürlich, dass $\varepsilon_k \to 0$ für $k \to \infty$.

3.5 Beispiele zur Konvergenz von Folgen

Beispiel 3.27. Für alle $k \in \mathbb{N}$ konvergiert die Folge $(\frac{1}{n^k})_{n\in\mathbb{N}}$ gegen Null, denn es gilt für alle $k, n \in \mathbb{N}$:

$$\left|0 - \frac{1}{n^k}\right| = \frac{1}{n^k} \leqslant \frac{1}{n}$$

Wir wissen schon nach Lemma 3.10, dass die rechte Seite gegen Null konvergiert. Die Behauptung folgt dann aus Satz 3.23, weil

$$0 < \frac{1}{n^k} < \frac{1}{n}.$$

Beispiel 3.28. Die Folge $\frac{3n^4+n^2-1}{2n^4-n^3+2}$ konvergiert gegen $\frac{3}{2}$. Dazu formen wir zuerst um

$$\frac{3n^4 + n^2 - 1}{2n^4 - n^3 + 2} = \frac{3 + \frac{1}{n^2} - \frac{1}{n^4}}{2 - \frac{1}{n} + \frac{2}{n^4}},$$

indem wir den Bruch mit n^4 kürzen. Der Nenner $2 - \frac{1}{n} + \frac{2}{n^4}$ konvergiert nach Satz 3.15(1) gegen 2, weil sowohl $\frac{1}{n}$, als auch $\frac{1}{n^4}$ nach Beispiel 3.27 Nullfolgen sind. Wir können als Satz 3.15(3) anwenden und es folgt:

$$\lim_n \frac{3n^4 + n^2 - 1}{2n^4 - n^3 + 2} = \frac{3 + \lim_n \frac{1}{n^2} - \lim_n \frac{1}{n^4}}{2 - \lim_n \frac{1}{n} + \lim_n \frac{2}{n^4}} = \frac{3}{2}.$$

Proposition 3.29. *Seien*

$$p(x) = \sum_{k=0}^{s} a_k x^k \quad \textit{und} \quad q(x) = \sum_{k=0}^{t} b_k x^k$$

in $\mathbb{C}[x]$ *mit* $a_s, b_t \neq 0$. *Betrachte die Folge* $\left(\frac{p(n)}{q(n)}\right)_{n\in\mathbb{N}}$. *Dann gilt:*

1. *Fall $s = t$: Die Folge* $\left(\frac{p(n)}{q(n)}\right)_{n\in\mathbb{N}}$ *konvergiert gegen* $\frac{a_s}{b_s}$.

2. *Fall $s < t$: Die Folge* $\left(\frac{p(n)}{q(n)}\right)_{n\in\mathbb{N}}$ *konvergiert gegen Null.*

3. *Fall $s > t$ und $a_s b_t > 0$: Die Folge* $\left(\frac{p(n)}{q(n)}\right)_{n\in\mathbb{N}}$ *divergiert bestimmt gegen* $+\infty$.

4. *Fall $s > t$ und $a_s b_t < 0$: Die Folge* $\left(\frac{p(n)}{q(n)}\right)_{n\in\mathbb{N}}$ *divergiert bestimmt gegen* $-\infty$.

Beweis. Zu 1.: Wir kürzen den Bruch $\frac{p(n)}{q(n)}$ um den Faktor n^s. Der Zähler hat dann die Form

$$\sum_{k=0}^{s} a_k n^{k-s} = a_s + a_{s-1}\frac{1}{n} + a_{s-2}\frac{1}{n^2} + \ldots + a_1\frac{1}{n^{s-1}} + a_0\frac{1}{n^s}.$$

Der Nenner sieht folgendermaßen aus:

$$\sum_{k=0}^{s} b_k n^{k-s} = b_s + b_{s-1}\frac{1}{n} + b_{s-2}\frac{1}{n^2} + \ldots + b_1\frac{1}{n^{s-1}} + b_0\frac{1}{n^s}.$$

Dieser Nenner konvergiert gegen $b_s \neq 0$ nach Satz 3.15(1). Damit lassen sich auch alle anderen Teile des Satzes 3.15 anwenden. Es folgt:

$$\lim_{n\to\infty}\frac{p(n)}{q(n)} = \frac{a_s}{b_s}.$$

Zu 2.: Wir kürzen den Bruch $\frac{p(n)}{q(n)}$ um den Faktor n^t. Der Zähler hat dann die Form

$$\sum_{k=0}^{s} a_k n^{k-t} = a_s\frac{1}{n^{t-s}} + a_{s-1}\frac{1}{n^{t-s+1}} + a_{s-2}\frac{1}{n^{t-s+2}} + \ldots + a_1\frac{1}{n^{t-1}} + a_0\frac{1}{n^t}.$$

Der Zähler konvergiert nach Satz 3.15(1) gegen Null. Der Nenner sieht folgendermaßen aus:

$$\sum_{k=0}^{t} b_k n^{k-t} = b_t + b_{t-1}\frac{1}{n} + b_{t-2}\frac{1}{n^2} + \ldots + b_1\frac{1}{n^{t-1}} + b_0\frac{1}{n^t}.$$

Dieser Nenner konvergiert gegen $b_t \neq 0$. Damit lassen sich alle Teile des Satzes 3.15 anwenden. Es folgt:

$$\lim_{n\to\infty}\frac{p(n)}{q(n)} = 0.$$

Zu 3. und 4.: Indem man wieder mit n^t kürzt, sieht man dass der Nenner gegen $b_t \neq 0$ konvergiert, der Zähler aber bestimmt divergiert. Dabei entscheidet das Vorzeichen von a_s, ob der Zähler gegen $+\infty$ oder $-\infty$ divergiert. Das Vorzeichen von b_t entscheidet dann, ob sich das Divergenzverhalten verändert oder nicht. Die Bedingung $a_s b_t > 0$ sagt, dass die Vorzeichen von a_s und b_t gleich sind und in diesem Fall divergiert der Bruch bestimmt gegen $+\infty$. Im Fall $a_s b_t < 0$ sind die Vorzeichen entgegengesetzt und der Bruch divergiert bestimmt gegen $-\infty$. □

Lemma 3.30. *Sei $a \in \mathbb{R}, a > 0$.*

1. *Wenn $a > 1$, dann gibt es zu jedem $S \in \mathbb{R}$ ein $n \in \mathbb{N}$ mit $a^n > S$.*
2. *Wenn $0 < a < 1$, dann gibt es zu jedem $\varepsilon > 0$ ein $n \in \mathbb{N}$ mit $a^n < \varepsilon$.*

Beweis. Zu 1.: Sei $S \in \mathbb{R}, S > 0$ gegeben. Setze $x = a - 1 \geqslant 0$. Dann ist die Bernoullische Ungleichung 2.22 anwendbar und es gilt:

$$a^n = (1 + x)^n \geqslant 1 + nx.$$

Nach dem archimedischen Axiom gibt es ein $n \in \mathbb{N}$, so dass $nx > S$. Für dieses n gilt dann:

$$a^n = (1 + x)^n \geqslant 1 + nx > 1 + S > S.$$

Zu 2.: Wir setzen $b = \frac{1}{a} > 1$. Sei $\varepsilon > 0$ gegeben. Dann gibt es nach Teil 1 zu $S = \frac{1}{\varepsilon}$ ein $n \in \mathbb{N}$ mit $b^n > S = \frac{1}{\varepsilon}$. Dann folgt:

$$a^n = \frac{1}{b^n} < \varepsilon.$$

□

Beispiel 3.31. Sei $a \in \mathbb{R}$.

1. Für alle $-1 < a < 1$ gilt: $\lim_n a^n = 0$.
2. Für alle $a > 1$ divergiert die Folge $(a^n)_{n\in\mathbb{N}}$ bestimmt gegen $+\infty$.
3. Für alle $a < -1$ divergiert die Folge $(a^n)_{n\in\mathbb{N}}$.

Beweis. Zu 1.: Der Fall $a = 0$ ist offensichtlich, denn für alle $n \in \mathbb{N}$ gilt dann $a^n = 0$. Sei von jetzt ab also $0 < |a| < 1$. Sei $\varepsilon > 0$. Nach Lemma 3.30(2) gibt es ein $N \in \mathbb{N}$ mit

$$|a^N| = |a|^N < \varepsilon.$$

Dann gilt aber für alle $n \geqslant N$:

$$|a^n| < |a^N| < \varepsilon$$

denn $0 < |a| < 1$. Dies zeigt die Behauptung.

Zu 2.: Wenn $a > 1$, dann ist $0 < \frac{1}{a} < 1$. Die Aussage folgt dann aus Teil 1 zusammen mit Proposition 3.18(2). Die Aussage folgt natürlich auch direkt aus Teil 1 von Lemma 3.30.

Zu 3. : Aus Teil 2 folgt, dass die Folge $(|a^n|)_{n\in\mathbb{N}}$ bestimmt gegen $+\infty$ divergiert. Weil $a < -1$, alterniert das Vorzeichen der Folgeglieder a^n. Damit divergiert die Folge immer noch, aber sie divergiert nicht mehr bestimmt gegen $\pm\infty$. □

Beispiel 3.32. Sei $z \in \mathbb{C}$ mit $|z| < 1$. Dann gilt $\lim_n z^n = 0$. Der Beweis dieser Aussage ist Wort für Wort derselbe wie der von Beispiel 3.31(1).

4 Cauchy-Folgen und Vollständigkeit

4.1 Cauchy-Folgen

Augustin Louis Cauchy `https://de.wikipedia.org/wiki/Augustin-Louis_Cauchy`

Definition 4.1. Eine Folge $(x_n)_{n\in\mathbb{N}}$ heißt *Cauchy-Folge*, wenn es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für alle $m, n \geqslant \mathbb{N}$

$$|x_m - x_n| < \varepsilon$$

gilt.

Man beachte, dass die Definition einer Cauchy-Folge keinen weiteren Referenzpunkt (wie z.B. bei konvergenten Folge den Grenzwert) voraussetzt.

Satz 4.2. *Jede konvergente Folge ist eine Cauchy-Folge.*

Beweis. Sei $(x_n)_{n\in\mathbb{N}}$ eine konvergente Folge mit Grenzwert x. Sei $\varepsilon > 0$ gegeben. Zu zeigen ist die Definition einer Cauchy-Folge für $(x_n)_{n\in\mathbb{N}}$. Weil die Folge gegen x konvergiert, gibt es ein $N \in \mathbb{N}$, so dass für alle $n \geqslant N$ gilt:

$$|x - x_n| < \frac{\varepsilon}{2}.$$

Dann folgt für alle $m, n \geqslant N$:

$$|x_m - x_n| = |x_m - x| + |x - x_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

□

Der folgende Satz gilt unabhängig von jeder Vollständigkeitsbedingung. So ist er auch für Cauchy-Folgen in $\mathbb{Q}$ (oder in jedem beliebigen metrischen Raum) wahr.

Satz 4.3. *Jede Cauchy-Folge ist beschränkt.*

Beweis. Sei $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge. Wir müssen ein $C \in \mathbb{R}$ finden, so dass für alle $n \in \mathbb{N}$ gilt: $|x_n| < C$.

Wir wählen $\varepsilon = 1$. Weil $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge ist, gibt es $N \in \mathbb{N}$, so dass für alle $m, n \geqslant N$ gilt: $|x_m - x_n| < 1$. Insbesondere gilt dies für $m = N$, d.h. für alle $n \geqslant N$ gilt:

$$|x_N - x_n| < 1.$$

Also sind alle bis vielleicht auf die ersten $N - 1$ Folgeglieder im Intervall $]x_N - 1, x_n + 1[$ enthalten. Beachte:

$$x_N - 1, x_N + 1 \leqslant |x_N| + 1.$$

Setze jetzt $C = \max\{|x_N| + 1, |x_1|, \ldots, |x_{N-1}|\}$. Dann gilt für alle $n \in \mathbb{N}$: $|x_n| < C$. □

Korollar 4.4. *Jede konvergente Folge (in $\mathbb{R}$ oder $\mathbb{C}$) ist beschränkt.*

Beweis. Dies folgt direkt aus den vorigen Sätzen 4.2 und 4.3. □

4.2 Vollständigkeit

Definition 4.5. Eine Teilmenge M von $\mathbb{R}$ oder $\mathbb{C}$ heißt vollständig, wenn jede Cauchy-Folge in M konvergiert und ihr Grenzwert auch in M liegt.

Axiom 4.6. Die reellen Zahlen sind vollständig.

Die Vollständigkeit von $\mathbb{R}$ hat wichtige Konsequenzen: das Intervallschachtelungsprinzip 4.13, die Dezimalbruchentwicklung 5.4 und die Existenz von Supremum und Infimum 6.19 einer nicht-leeren beschränkten Teilmenge von $\mathbb{R}$.

Definition 4.7. Sei I ein Intervall. Dabei spielt es keine Rolle, ob es von der Form $[a, b]$, $]a, b[$, $[a, b[$ oder $]a, b]$ für $a < b$ ist. Sein *Durchmesser* ist definiert durch

$$\operatorname{diam}(I) = b - a.$$

Definition 4.8. Eine *Intervallschachtelung* ist eine Folge $(I_k)_{k\in\mathbb{N}}$ von abgeschlossenen Intervallen, so dass

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

Definition 4.9. Wir sagen, dass für eine Teilmenge M von $\mathbb{R}$ das *Intervallschachtelungprinzip* gilt, wenn es zu jeder Intervallschachtelung $(I_k)_{k\in\mathbb{N}}$ mit $I_1 \subset M$ eine Zahl $x \in M$ gibt, die in allen Intervallen $I_k, k \in \mathbb{N}$ enthalten ist.

Die folgenden Proposition sagt aus, dass das x, dessen Existenz im Intervallschachtelungsprinzip behauptet wird, immer eindeutig bestimmt ist.

Proposition 4.10. *Sei $M \subset \mathbb{R}$ eine Menge, für die das Intervallschachtelungsprinzip gilt. Sei $(I_k)_{k\in\mathbb{N}}$ eine Intervallschachtelung in M und seien $x, x' \in M$ Elemente, die in allen Intervallen I_k liegen. Dann ist $x = x'$.*

Beweis. Seien x und x' in M zwei solche Elemente. Weil x und x' aber beide in allen Intervallen liegen, gilt für alle $k \in \mathbb{N}$

$$|x - x'| < \operatorname{diam}(I_k).$$

Weil $\lim_k \operatorname{diam}(I_k) = 0$ nach Voraussetzung, gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $k \geqslant N$ gilt:

$$|x - x'| < \operatorname{diam}(I_k) < \varepsilon.$$

Damit folgt $x = x'$. □

Definition 4.11. Eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{R}$ heißt *(monoton) wachsend*, wenn für alle $k \in \mathbb{N}$ gilt: $a_k \leqslant a_{k+1}$.

In diesem Sinne ist z.B. auch eine konstante Folge wachsend.

Im Beweis des Satzes 4.13 benötigen wir das folgende

Lemma 4.12. *Sei $(a_n)_{n\in\mathbb{N}}$ eine wachsende konvergente Folge mit Grenzwert a. Dann gilt für alle $n \in \mathbb{N}$ auch $a_n \leqslant a$.*

Beweis. Wir nehmen an, es gäbe ein $N \in \mathbb{N}$ mit $a_N > a$. Weil aber

$$a < a_N \leqslant a_{N+1} \leqslant a_{N+2} \leqslant \dots$$

folgt dann sofort für alle $n \geqslant N$

$$|a - a_n| \geqslant |a - a_N|,$$

und die Folge $(a_n)_{n\in\mathbb{N}}$ kann sicherlich nicht gegen a konvergieren. □

Satz 4.13. *In* $\mathbb{R}$ *gilt das Intervallschachtelungsprinzip.*

Beweis. Sei $(I_k)_{k\in\mathbb{N}}$ eine Intervallschachtelung mit $I_k = [a_k, b_k]$. Zu zeigen ist, dass es ein $x \in \mathbb{R}$ gibt mit $x \in I_k$ für alle $k \in \mathbb{N}$.

Wir zeigen zuerst, dass $(a_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge ist. Sei also $\varepsilon > 0$ gegeben. Nach Definition einer Intervallschachtelung gilt

$$\operatorname{diam}(I_k) \to 0 \quad \text{für} \quad n \to \infty.$$

Also gibt es ein $N \in \mathbb{N}$, so dass für alle $k \geqslant N$

$$\operatorname{diam}(I_k) < \varepsilon.$$

Dann gilt für alle $m, n \geqslant N$, dass $a_m, a_n \in I_N$ und damit

$$|a_m - a_n| \leqslant \operatorname{diam}(I_N) < \varepsilon.$$

Damit erfüllt $(a_k)_{k\in\mathbb{N}}$ die Bedingung für Cauchy-Folgen.

Nach dem Vollständigkeitsaxiom 4.6 konvergiert die Folge $(a_k)_{k\in\mathbb{N}}$ gegen einen Grenzwert $x \in \mathbb{R}$. Wir zeigen jetzt im nächsten Schritt, dass x das gesuchte Element.

Weil $(a_k)_{k\in\mathbb{N}}$ eine monoton wachsende Folge mit Grenzwert x ist, gilt nach Lemma 4.12 für alle $k \in \mathbb{N}$ die Ungleichung $a_k \leqslant x$. Es gilt aber auch für alle $k \in \mathbb{N}$ eben $x \leqslant b_k$. Gäbe es ein $K \in \mathbb{N}$ mit $x > b_K$, so gälte

$$|x - a_n| \geqslant |x - b_K|,$$

denn $a_n \leqslant b_n \leqslant b_K < x$. Damit könnte aber die Folge $(a_n)_{n\in\mathbb{N}}$ wieder nicht gegen x konvergieren.

Insgesamt folgt, dass für alle $k \in \mathbb{N}$ gilt: $x \in [a_k, b_k] = I_k$. □

Im vorigen Beweis haben wir in essentieller Weise die Vollständigkeit von $\mathbb{R}$ ausgenutzt. Tatsächlich ist das Intervallschachtelungsprinzip äquivalent zur Vollständigkeit. Dazu beweisen wir die Umkehrung von Satz 4.13.

Satz 4.14. *Das Intervallschachtelungsprinzip für* $\mathbb{R}$ *impliziert die Vollständigkeit von* $\mathbb{R}$.

Beweis. Wir müssen zeigen, dass jede Cauchy-Folge in $\mathbb{R}$ einen Grenzwert hat. Dazu sei $(a_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$. Zu jedem $k \in \mathbb{N}$ gibt es dann (setze $\varepsilon = 2^{-k}$) eine Zahl $N_k \in \mathbb{N}$, so dass für alle $n, m \geqslant N_k$ die Abschätzung

$$|a_m - a_n| < 2^{-k}$$

gilt. Mit anderen Worten gibt es eine Folge solcher N_k und wir können ohne Einschränkung annehmen, dass $N_0 < N_1 < N_2 < \ldots$. Wir definieren jetzt zu jedem $k \in \mathbb{N}$ ein Intervall

$$I_k = \{x \in \mathbb{R} \mid |x - a_{N_k}| \leqslant 2^{-k+1}\}.$$

Alle diese Intervalle sind abgeschlossen.

Wir behaupten, dass für alle $k \in \mathbb{N}$ auch $I_{k+1} \subset I_k$ gilt. Sei $x \in I_{k+1}$, d.h.

$$|x - a_{N_{k+1}}| \leqslant 2^{-k}.$$

Desweiteren wissen wir nach Wahl von N_k, dass

$$|a_{N_{k+1}} - a_{N_k}| \leqslant 2^{-k}.$$

Mit der Dreiecksungleichung folgt also:

$$|x - a_{N_k}| \leqslant |x - a_{N_{k+1}}| + |a_{N_{k+1}} - a_{N_k}| \leqslant 2^{-k} + 2^{-k} = 2^{-k+1}.$$

Das beudeutet aber nichts anderes als $x \in I_k$.

Die Folge der Durchmesser der Intervalle $\left(\frac{1}{2^k}\right)_{n\in\mathbb{N}}$ ist eine Nullfole. Also ist $(I_k)_{k\in\mathbb{N}}$ tatsächlich eine Intervallschachtelung und aus dem Intervallschachtelungsprinzip folgt, dass es ein $x_0 \in \mathbb{R}$ gibt mit $x_0 \in I_k$ für alle $k \in \mathbb{N}$. Nach Definition der I_k gilt also für alle $k \in \mathbb{N}$:

$$|x_0 - a_{N_k}| \leqslant 2^{-k+1}.$$

Wir behaupten jetzt, dass $\lim_{n\to\infty} a_n = x_0$. Wieder nach Wahl der N_k gilt für alle $n \geqslant N_k$:

$$|x_0 - a_n| \leqslant |x - a_{N_k}| + |a_{N_k} - a_n| \leqslant 2^{-k+1} + 2^{-k} = \frac{3}{2^{-k}} < 2^{-k+2}.$$

Die Konvergenz folgt jetzt aus Lemma 3.25 bzw. Bemerkung 3.26. □

Satz 4.15. *Der Körper* $\mathbb{C}$ *ist vollständig.*

Beweis. Sei $(z_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge von komplexen Zahlen. Wir behaupten, dass dann die Folge der Realteile $(a_n = \mathrm{Re}\,(z_n))_{n\in\mathbb{N}}$ und Imaginärteile $(b_n = \mathrm{Im}\,(z_n))_{n\in\mathbb{N}}$ auch Cauchy-Folgen sind. Dies folgt sehr einfach aus den Ungleichungen

$$|a_m - a_n| = |\mathrm{Re}\,(z_m) - \mathrm{Re}\,(z_n)| = |\mathrm{Re}\,(z_m - z_n)| \leqslant |z_m - z_n|$$

und

$$|b_m - b_n| = |\mathrm{Im}\,(z_m) - \mathrm{Im}\,(z_n)| = |\mathrm{Im}\,(z_m - z_n)| \leqslant |z_m - z_n|.$$

Die Vollständigkeit von $\mathbb{R}$ impliziert jetzt, dass beide Folgen $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergieren. Die Aussage des Satzes folgt dann aus Lemma 3.24. □

5 Reihen

5.1 Reihen und konvergente Reihen

Definition 5.1. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ oder $\mathbb{C}$. Die Folge der *Partialsummen*

$$s_n = \sum_{k=0}^{n} a_k$$

heißt *die der Folge $(a_n)_{n\in\mathbb{N}}$ zugeordnete unendliche Reihe* oder einfach *Reihe*. Wir sagen, dass *die Reihe konvergiert*, wenn die Folge $(s_n)_{n\in\mathbb{N}}$ der Partielsummen konvergiert. Wir schreiben den Grenzwert dann als

$$\sum_{k=0}^{\infty} a_k.$$

Wir sagen, dass *die Reihe divergiert* (*bestimmt gegen $\pm\infty$ divergiert*, *beschränkt ist*), wenn $(s_n)_{n\in\mathbb{N}}$ divergiert (bestimmt gegen $\pm\infty$ divergiert, beschränkt ist).

Notation 5.2. Man beachte:

1. Um die Notation zu vereinfachen, bezeichnet Der Term $\sum_{k=0}^{\infty} a_k$ oft auch einfach die Reihe selber, d.h. die Folge $(s_n = \sum_{k=0}^{n} a_k)_{n\in\mathbb{N}}$. Damit macht man einen begrifflichen Fehler, aber

2. Wir betrachten manchmal auch Reihen, die bei einem anderen Index als 0 starten: z.B. $\sum_{k=1}^{\infty} a_k$, $\sum_{k=-5}^{\infty} a_k$, $\sum_{k=m}^{\infty} a_k$ für ein $m \in \mathbb{Z}$, ... Die Bezeichnung ist selbsterklärend. Konvergenz/Divergenz bedeutet immer Konvergenz/Divergenz der Folge der Partialsummen (auch wenn diese Folgen vielleicht bei $0, 1, -5$ oder $m \in \mathbb{Z}$ anfangen).

Jeder Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ ist wie oben beschrieben eine Reihe zugeordnet. Umgekehrt kann man jeder Reihe eine Folge zuordnen, indem man

$$a_0 = s_0 \quad \text{und} \quad a_n = s_n - s_{n-1} \quad \text{für} \quad n > 0$$

setzt. Das bedeutet, dass jede Definition oder Aussage über Folgen zu einer Definition oder Aussage über Reihen uminterpretiert werden kann und umgekehrt. Genau das haben wir z.B. für den Begriff der Konvergenz schon gemacht. Die Tatsache, dass in $\mathbb{R}$ jede Cauchy-folge konvergiert, schlägt sich in der Theorie der Reihe auf folgende Art nieder.

Satz 5.3 (Cauchysches Konvergenzkriterium)**.** *Die Reihe $\sum_{k=0}^{\infty} a_k$ konvergiert genau dann, wenn für alle $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass für alle $n \geqslant m \geqslant N$*

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

gilt.

Beweis. Sei $\varepsilon > 0$ und es existiere ein $N \in \mathbb{N}$, so dass für alle $n \geqslant m \geqslant N$

$$\left|\sum_{k=m}^{n} a_k\right| < \varepsilon.$$

Diese Bedingung sagt gerade, dass die Folge der Partialsummen eine Cauchy-Folge bilden:

$$|s_n - s_{m-1}| = \left|\sum_{k=0}^{n} a_k - \sum_{k=0}^{m-1} a_k\right| = \left|\sum_{k=m}^{n} a_k\right| < \varepsilon$$

Weil $\mathbb{R}$ und $\mathbb{C}$ vollständig sind und deshalb alle Cauchy-Folgen konvergieren, konvergieren die Partialsummen und damit die Reihe $\sum_{k=0}^{\infty} a_k$.

Die Umkehrung ist einfach. Eine Reihe konvergiert genau dann, wenn ihre Partialsummen konvergieren. Nach Satz 4.2 ist dann die Folge der Partialsummen eine Cauchy-Folge. Das ist aber gerade das Cauchysche Konvergenzkriterium. □

Proposition 5.4. *Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ oder $\mathbb{C}$. Wenn die Reihe $\sum_{k=0}^{\infty} a_k$ konvergiert, so ist $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge.*

Beweis. Nach dem Cauchy-Konvergenzkriterium 5.3 gibt es also zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $m, n \geqslant N$

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

gilt. Man kann jetzt einfach $m = n \geqslant N$ wählen und es folgt $\lim_n a_n = 0$. □

Dass die vorhergehende Proposition keine Umkehrung zuläßt, zeigt folgendes wichtiges Beispiel. Beachte hierzu auch Satz 8.5.

5.2 Die harmonische Reihe

Definition 5.5. Die Reihe

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \dots$$

heißt *harmonische Reihe.*

Satz 5.6. *Die harmonische Reihe divergiert.*

Beweis. Dazu betrachte man folgende Abschätzungen:

$$1 \geqslant \frac{1}{2} \qquad \frac{1}{2} \geqslant \frac{1}{2} \qquad \frac{1}{3} + \frac{1}{4} = \frac{7}{12} \geqslant \frac{1}{2} \qquad \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} = \frac{533}{840} \geqslant \frac{1}{2}$$

Wir addieren jetzt immer die Terme $\frac{1}{2^k+1}$ bis $\frac{1}{2^{k+1}}$ für $k \geqslant 1$ auf. Dies sind 2^k Terme und jeder einzelne dieser Term ist größer oder gleich $\frac{1}{2^{k+1}}$. Zum Bespiel:

$$\frac{1}{3} + \frac{1}{4} \geqslant \frac{1}{4} + \frac{1}{4} = \frac{1}{2} \qquad \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} \geqslant \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{1}{2}$$

Es folgt also:

$$\sum_{n=2^k+1}^{2^{k+1}} \frac{1}{n} \geqslant 2^k \cdot \frac{1}{2^{k+1}} = \frac{1}{2}$$

Damit bekommt man für die Folge der Partialsummen

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

Diese Folge ist also nicht nach oben beschränkt und kann wegen Korollar 4.4 nicht konvergieren. Genauer folgt natürlich, dass die harmonische Reihe bestimmt gegen $+\infty$ divergiert. □

5.3 Die geometrische Reihe

Die bei weitem wichtigste Reihe in der Mathematik ist die geometrische Reihe.

Definition 5.7. Für ein $x \in \mathbb{C}$ heißt

$$\sum_{n=0}^{\infty} x^n = 1 + x + x^2 + x^3 + \dots$$

die *geometrische Reihe für x*.

Satz 5.8. *Für jedes $x \in \mathbb{C}$ mit $|x| < 1$ konvergiert die zu x gehörende geometrische Reihe und es gilt:*

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

Beweis. Nach Lemma 2.41 gilt für alle $x \in \mathbb{C}$ mit $|x| < 1$

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

Wegen Lemma 3.31 (bzw. 3.32 für den komplexen Fall) ist $(x^k)_{k\in\mathbb{N}}$ eine Nullfolge. Nach Satz 3.15 folgt dann

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

□

Beispiel 5.9. Für $x = \frac{1}{2}$ gilt

$$1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \dots = \sum_{n=0}^{\infty} 2^{-n} = \frac{1}{1-\frac{1}{2}} = 2.$$

Lemma 5.10. *Für alle $x \in \mathbb{R}$ mit $|x| \geqslant 1$ divergiert die geometrische Reihe $\sum_{n=0}^{\infty} x^n$.*

Beweis. Wenn $|x| > 1$, dann kann die Folge $(x^n)_{n\in\mathbb{N}}$ nicht konvergieren, siehe Beispiel 3.31. Für $x = 1$ ist $(x^n)_{n\in\mathbb{N}}$ die konstante Folge bei 1. Für $x = -1$ springt diese Folge immer zwischen 1 und -1. Insbesondere ist $(x^n)_{n\in\mathbb{N}}$ in keinem der Fälle eine Nullfolge. Wegen Proposition 5.4 kann die entsprechende Reihe also nicht konvergieren. □

5.4 b-adische Brüche und Dezimalbruchentwicklung

Was meinen wir üblicherweise, wenn wir $12,782$ schreiben? Antwort:

$$1 \cdot 10^1 + 2 \cdot 10^0 + 7 \cdot 10^{-1} + 8 \cdot 10^{-2} + 2 \cdot 10^{-3}.$$

Im Dezimalsystem bedeuten die Ziffern einer Zahl, wie oft die entsprechende Potenz von Zehn auftaucht, Um den Wert der Zahl (eigentlich Ziffernfolge) zu bekommen, muss man dann diese Vielfachen von Zehnerpotenzen aufaddieren. Dabei muss die Darstellung in Zehnerpotenzen (also die Ziffernfolge) nicht unbedingt abbrechen. So gilt z.B.

$$\frac{1}{9} = 0,111\ldots = \sum_{k=1}^{\infty} 10^{-k}.$$

Man beachte, dass dies (bis auf den Anfang bei $k = 1$) die geometrische Reihe bzgl. $\frac{1}{10}$ ist. Tatsächlich gilt laut Satz 5.8:

$$\sum_{k=1}^{\infty} 10^{-k} = -1 + \sum_{k=0}^{\infty} 10^{-k} = \frac{1}{1-\frac{1}{10}} - 1 = \frac{10}{9} - 1 = \frac{1}{9}.$$

Manche Zahlen haben eine Dezimalbruchentwicklung, die in einer sich wiederholenden, also periodischen Zahlenfolge endet. Wir notieren dies mit Hilfe eines Querstriches über den Ziffern (was natürlich nichts mit komplexer Konjugation zu tun hat). Z.B.:

$$\begin{aligned} 0,31414141\ldots = 0,3\overline{14} = 3\cdot 10^{-1} + 14\cdot 10^{-3} + 14\cdot 10^{-5} + \ldots &= \frac{3}{10} + \frac{14}{1000}\sum_{k=0}^{\infty}\frac{1}{10^{2k}} \\ &= \frac{3}{10} + \frac{14}{1000}\cdot\frac{1}{1-\frac{1}{100}} = \frac{3}{10} + \frac{14}{1000}\cdot\frac{100}{99} = \frac{3}{10} + \frac{14}{990} \\ &= \frac{297+14}{990} = \frac{311}{990} \end{aligned}$$

Die Dezimalbruchdarstellung reller Zahlen beruht auf dem 10er-System. Man kann allgemeiner reelle Zahlen durch b-adische Brüche mit $b \in \mathbb{N}, b \geqslant 2$ darstellen. Das Dezimalsystem ist einfach der Spezialfall $b = 10$.

Definition 5.11. Sei $b \in \mathbb{N}, b \geqslant 2$. Ein *(unendlicher) b-adischer Bruch* ist eine Reihe der Form

$$\pm \sum_{n=-k}^{\infty} a_k b^{-n}$$

mit $a_k \in \{0, 1, \ldots, b-1\}$ und $k \geqslant 0$.

Satz 5.12. *Die Partialsummen eines b-adischen Bruches sind eine Cauchy-Folge.*

Bemerkung 5.13. Insbesondere konvergiert jeder b-adische Bruch gegen eine reelle Zahl und wir sagen dann, dass der b-adische Bruch diese reelle Zahl darstellt.

Beweis. Es reicht zu zeigen, dass $(s_n = \sum_{\nu=-k}^{n} a_k b^{-\nu})_{n\geqslant -k}$ eine Cauchy-Folge ist. Die Aussage folgt dann auch für negative b-adische Brüche.

Sei $\varepsilon > 0$. Weil $b \geqslant 2$, folgt $0 < b^{-1} < 1$. Nach Lemma 3.30(2) existiert ein $N \in \mathbb{N}$ mit $b^{-N} < \varepsilon$. Es gilt dann für alle $n \geqslant m \geqslant N$:

$$\begin{aligned} |s_n - s_m| = \sum_{\nu=m+1}^{n} a_\nu b^{-\nu} &\leqslant (b-1)\sum_{\nu=m+1}^{n} b^{-\nu} = (b-1)b^{-m-1}\sum_{\nu=0}^{\nu=n-m-1} b^{-\nu} \\ &\leqslant (b-1)b^{-m-1}\frac{1}{1-b^{-1}} = b^{-m} \leqslant b^{-N} < \varepsilon \end{aligned}$$

□

Satz 5.14. *Sei $b \in \mathbb{N}, b \geqslant 2$. Dann lässt sich jede reelle Zahl als b-adischer Bruch darstellen.*

Beweis. Es reicht, eine reelle Zahl $x \geqslant 0$ als b-adischen Bruch zu entwickeln. Für negative reelle Zahlen folgt die Aussage dann durch Multiplikation mit -1.

Sei $x \geqslant 0$. Nach Lemma 3.30(1) gibt es ein $N \in \mathbb{N}$ mit $x < b^{N+1}$. Sei k die kleinste Zahl mit der Eigenschaft $0 \leqslant x < b^{k+1}$.

Wir konstruieren jetzt mit vollständiger Induktion eine Folge $(a_\nu)_{\nu \geqslant -k}$ in $\mathbb{N}$ mit $0 \leqslant a_\nu \leqslant b-1$, so dass für alle $n \geqslant -k$

$$x = \xi_n + \sum_{\nu=-k}^{n} a_\nu b^{-\nu} \quad \text{mit} \quad 0 \leqslant \xi_n < b^{-n}$$

gilt. Weil $\lim_n b^{-n} = 0$ nach Beispiel 3.31(1), folgt mit dem Sandwichsatz 3.23 auch $\lim_n \xi_n = 0$ und damit

$$x = \sum_{\nu=-k}^{\infty} a_\nu b^{-\nu}.$$

Induktionanfang $n = -k$: Es gilt $0 \leqslant x < b^{k+1}$ bzw. $0 \leqslant xb^{-k} < b$. Also gibt es ein $a_{-k} \in \{0, 1, \ldots, b-1\}$ und ein $\delta \in \mathbb{R}, 0 \leqslant \delta < 1$ mit

$$xb^{-k} = a_{-k} + \delta.$$

Setze jetzt $\xi_{-k} = \delta b^k$. Dann erhält man

$$x = a_{-k} b^k + \xi_{-k} \quad \text{mit} \quad 0 \leqslant \xi_{-k} < b^k$$

Das ist die Behauptung für $n = -k$.

Induktionannahme: Es gibt $a_{-k}, a_{-k+1}, \ldots, a_n \in \mathbb{N}$ mit $0 \leqslant a_\nu \leqslant b-1$ und $\xi_n \in \mathbb{R}$ mit $0 \leqslant \xi_n < b^{-n}$, so dass für alle $-k \leqslant \nu \leqslant n$ gilt:

$$x = \xi_n + \sum_{\nu=-k}^{n} a_\nu b^{-\nu}.$$

Induktionsschritt: Wir müssen ein $a_{n+1} \in \{0, 1, \ldots, b-1\}$ und ein $0 \leqslant \xi_{n+1} < b^{-n-1}$ finden, so dass

$$x = \xi_{n+1} + \sum_{\nu=-k}^{n+1} a_\nu b^{-\nu}$$

gilt.

Wir haben ξ_n so konstruiert, dass $0 \leqslant \xi_n < b^{-n}$ bzw. $0 \leqslant \xi_n b^{n+1} < b$ gilt. Daher gibt es ein $a_{n+1} \in \{0, 1, \ldots, b-1\}$ und ein $\delta \in \mathbb{R}, 0 \leqslant \delta < 1$ mit

$$\xi_n b^{n+1} = a_{n+1} + \delta.$$

Setze $\xi_{n+1} = \delta b^{-n-1}$. Dann erhält man

$$x = (a_{n+1} + \delta) b^{-n-1} + \sum_{\nu=-k}^{n} a_\nu b^{-\nu} = \xi_{n+1} + \sum_{\nu=-k}^{n+1} a_\nu b^{-\nu}$$

mit $0 \leqslant \xi_{n+1} < b^{-n-1}$. □

Bemerkung 5.15. Im Falle $b = 10$ gilt:

$$0,\overline{9} = \sum_{k=1}^{\infty} 9 \cdot 10^{-k} = -9 + 9\sum_{k=0}^{\infty} 10^{-k} = -9 + \frac{9}{1-\frac{1}{10}} = 1.$$

Die Dezimalbruchdarstellung einer reellen Zahl ist eindeutig bis auf die gerade erwähnte Identität. Allgemeiner gilt für alle $b \in \mathbb{N}, b \geqslant 2$: wenn zwei b-adische Brüche dieselbe reelle Zahl darstellen, so sind sie entweder gleich oder sie unterscheiden sich in der folgenden Form:

$$a_n a_{n-1} \ldots a_1 a_0 a_{-1} \ldots a_{-m+1} a_{-m} \overline{b-1} = a_n a_{n-1} \ldots a_1 a_0 a_{-1} \ldots a_{-m+1}(a_{-m} + 1)$$

mit $n \in \mathbb{N}_0, m \in \mathbb{Z}, m \leqslant n$.

Ein endlicher b-adischer Bruch ist offensichtlich eine rationale Zahl. Mit Hilfe des Prozesses in der Einleitung zu diesem Abschnitt sieht man, dass jeder periodische b-adische Bruch genauso in $\mathbb{Q}$ liegt. Dies folgt im wesentlichen aus der Konvergenz der geometrischen Reihe.

Proposition 5.16. *Sei $b \in \mathbb{N}, b \geqslant 2$. Ein nicht endlicher und nicht periodischer b-adischer Bruch ist irrational.*

Wir werden den Beweis nicht komplett ausführen. Er folgt aber aus dem vorher Gesagten und dem folgenden

Lemma 5.17. *Sei $p \in \mathbb{N}$ und $b \in \mathbb{N}, b \geqslant 2$. Dann besitzt $\frac{1}{p}$ eine entweder endliche oder periodische b-adische Darstellung.*

Beweis. Vielleicht später. □

6 Der Satz von Bolzano-Weierstraß, monotone Folgen, Infima und Suprema

6.1 Teilfolgen und der Satz von Bolzano-Weierstraß

Bernard Bolzano `https://en.wikipedia.org/wiki/Bernard_Bolzano`
Karl Weierstraß `https://de.wikipedia.org/wiki/Karl_Weierstra%C3%9F`

Definition 6.1. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in einer Menge. Ferner sei $n_0 < n_1 < n_2 < \ldots$ eine Folge natürlicher Zahlen. Dann heißt die Folge $(a_{n_k})_{k\in\mathbb{N}}$ eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$.

Wenn wir zurückgehen auf die ursprüngliche Definition einer Folge, dann ist eine Folge $(a_n)_{n\in\mathbb{N}}$ eine Abbildung $\varphi\colon \mathbb{N} \to M$. Eine Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ ist dann nichts anderes als die Verknüpfung von φ mit einer streng monoton steigenden Abbildung $\vartheta\colon \mathbb{N} \to \mathbb{N}$, also

$$(a_n)_{n\in\mathbb{N}} = (\varphi\colon \mathbb{N} \to M) \quad \text{und} \quad (a_{n_k})_{k\in\mathbb{N}} = (\mathbb{N} \xrightarrow{\vartheta} \mathbb{N} \xrightarrow{\varphi} M)$$

mit

$$a_{n_k} = \varphi(\vartheta(k)).$$

Beispiel 6.2. Die Folge $(\frac{1}{2n})_{n\in\mathbb{N}}$ ist eine Teilfolge von $(\frac{1}{n})_{n\in\mathbb{N}}$:

$$a_n = \frac{1}{n} \quad , \; a_{n_k} = \frac{1}{2k}$$

oder anders ausgedrückt:

$$\varphi\colon \; \mathbb{N} \to \mathbb{N}, n \mapsto \frac{1}{n} \quad , \vartheta\colon \; \mathbb{N} \to \mathbb{N}, n \mapsto 2n \quad , \; a_{n_k} = \varphi(\vartheta(k)) = \frac{1}{2k}.$$

Es gilt offensichtlich:

Lemma 6.3. *Jede Teilfolge einer konvergenten Folge mit Grenzwert x konvergiert selbst gegen x.*

Der Beweis dieses Lemmas ist dem sorgfältigen Leser überlassen.

Satz 6.4 (Bolzano-Weierstraß)**.** *Jede beschränkte Folge in $\mathbb{R}$ besitzt eine konvergente Teilfolge.*

Beweis. Sei $(x_n)_{n\in\mathbb{N}}$ eine beschränkte Folge mit unterer Schranke A und oberer Schranke B. Wir konstruieren rekursiv eine Folge von abgeschlossenen Intervallen $(I_k)_{k\in\mathbb{N}}$, so dass für alle $k \in \mathbb{N}$ gilt:

1. Es liegen unendlich viele Folgeglieder von $(a_n)_{n\in\mathbb{N}}$ in I_k.
2. Es gilt $I_k \subset I_{k-1}$.
3. Es gilt $\operatorname{diam}(I_k) \leqslant 2^{-k}\operatorname{diam}(I_0)$.

Wir setzen $I_0 = [A, B]$. Jetzt nehmen wir an, dass die ersten n Intervalle schon konstruiert sind. Wir müssen jetzt erklären, wie wir das $(n+1)$-te Intervall definieren wollen.

Betrachte den Mittelpunkt $C_{n+1} = \frac{A_n+B_n}{2}$ des Intervalls $I_n = [A_n, B_n]$. Wir wissen, dass nach Konstruktion von I_n unendlich viele Folgeglieder $(a_n)_{n\in\mathbb{N}}$ in I_n liegen. Das bedeutet, dass in mindestens einem der Intervalle $[A_n, C_{n+1}]$ oder $[C_{n+1}, B_n]$ wieder unendlich viele Folgeglieder liegen müssen. Wir wählen dasjenige Intervall, in dem unendlich viele Folgeglieder liegen als neues Intervall. Setze also

$$I_{n+1} = [A_{n+1}, B_{n+1}] \quad \text{mit } A_{n+1} = A_n, B_{n+1} = C_{n+1},$$

wenn unendlich viele der a_n in diesem Intervall liegen, und ansonsten setze

$$I_{n+1} = [A_{n+1}, B_{n+1}] \quad \text{mit } A_{n+1} = C_{n+1}, B_{n+1} = B_n.$$

Es bleibt klar, dass die abgeschlossenen Intervalle $(I_k)_{k\in\mathbb{N}}$ eine Intervallschachtelung bilden, denn die obigen Bedingungen 2 und 3 sind erfüllt. Insbesondere gibt es genau ein $x \in \mathbb{R}$ mit $x \in I_k$ für alle $k \in \mathbb{N}$.

Induktiv wählen wir jetzt $b_j \in I_j$. Setze $b_0 = a_0 \in I_0$. Seien die ersten j Folgeglieder schon gewählt. Wir wähle jetzt $b_{j+1} = a_{n_j} \in I_{j+1}$, so dass $b_{j+1} = a_{n_j} \neq a_\ell$ für alle $\ell < n_j$. Eine solches b_{j+1} existiert, denn es gibt in I_{j+1} unendlich viele Folgeglieder von $(a_n)_{n\in\mathbb{N}}$. Dann ist $(b_j)_{j\in\mathbb{N}}$ eine Teilfolge von $(a_n)_{n\in\mathbb{N}}$.

Es reicht jetzt zu zeigen, dass $\lim_j b_j = x$. Es gilt aber

$$|x - b_j| \leqslant \operatorname{diam}(I_j) = 2^{-j}\operatorname{diam}(I_0) \to 0 \quad \text{für} \quad j \to \infty.$$

Dies zeigt mit Hilfe von Lemma 3.25 bzw. Bemerkung 3.26, dass $(b_j)_{j\in\mathbb{N}}$ konvergiert und beendet den Beweis. □

Korollar 6.5. *Jede beschränkte Folge in $\mathbb{C}$ besitzt eine konvergente Teilfolge.*

Beweis. Wenn eine komplexe Folge $(z_n)_{n\in\mathbb{N}}$ beschränkt ist, dann sind die reellen Folgen der jeweiligen Real- und Imaginärteile $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ auch beschränkt. Nach Satz 6.4 besitzt dann die Folge $(a_n)_{n\in\mathbb{N}}$ der Realteile eine konvergente Teilfolge $(a_{n_j})_{j\in\mathbb{N}}$. Die entsprechende Folge $(b_{n_j})_{j\in\mathbb{N}}$ ist dann beschränkt und besitzt nach 6.4 eine konvergente Teilfolge $(b_{n_{j_\ell}})_{\ell\in\mathbb{N}}$. Nach Lemma 3.24 konvergiert dann die Teilfolge $(z_{n_{j_\ell}} = a_{n_{j_\ell}} + ib_{n_{j_\ell}})_{\ell\in\mathbb{N}}$. □

Beispiel 6.6. Die komplexe Folge $(i^n)_{n\in\mathbb{N}}$ ist beschränkt. Eine konvergente (in der Tat konstante) Teilfolge ist $(i^{4n} = 1)_{n\in\mathbb{N}}$.

6.2 Monotone Folgen

Der erste Teil der nächsten Definition wurde schon in 4.11 gegeben.

Definition 6.7. Sei $(x_n)_{n\in\mathbb{N}}$ eine reelle Folge. Die Folge

1. heißt *(monoton) steigend/wachsend*, wenn für alle $n \in \mathbb{N}$ gilt: $x_n \leqslant x_{n+1}$.
2. heißt *streng monoton steigend/wachsend*, wenn für alle $n \in \mathbb{N}$ gilt: $x_n < x_{n+1}$.
3. heißt *(monoton) fallend*, wenn für alle $n \in \mathbb{N}$ gilt: $x_n \geqslant x_{n+1}$.
4. heißt *streng (monoton) steigend*, wenn für alle $n \in \mathbb{N}$ gilt: $x_n > x_{n+1}$.

Eine Folge heißt *monoton*, wenn sie monoton fallend oder monoton steigend ist.

Beachte, dass der Begriff einer fallenden oder steigenden Folge nur in $\mathbb{R}$ Sinn ergibt.

Satz 6.8. *Jede monotone und beschränkte Folge konvergiert.*

Beweis. Sei $(x_n)_{n\in\mathbb{N}}$ eine monoton steigende und beschränkte Folge. Weil die Folge beschränkt ist, existiert nach dem Satz 6.4 von Bolzano-Weierstraß bzw. seinem komplexen Analogon 6.5 eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ mit Grenzwert x.

Wir zeigen, dass x tatsächlich der Grenzwert der gesamten Folge $(x_n)_{n\in\mathbb{N}}$ ist. Dazu machen wir eine Vorüberlegung: Weil die Folge $(x_n)_{n\in\mathbb{N}}$ und damit auch ihre Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ steigend sind, folgt nach Lemma 4.12, dass für alle $k \in \mathbb{N}$ gilt: $x_{n_k} \leqslant x$. Teilfolgen haben aber die folgende Eigenschft: Zu jedem $n \in \mathbb{N}$ gibt es aber ei $j \in \mathbb{N}$ mit

$$n_j \leqslant n \leqslant n_{j+1}.$$

Insbesondere folgt dann aus der Monotonie:

$$x_{n_j} \leqslant x_n \leqslant x_{n_{j+1}} \leqslant x. \tag{6.9}$$

Jetzt zum Beweis der Konvergenz. Sei $\varepsilon > 0$ gegeben. Dann gibt es ein $K \in \mathbb{N}$, so dass für alle $k \geqslant K$ gilt:

$$|x - x_{n_k}| < \varepsilon.$$

gilt. Setze $N = n_k$. Es folgt dann aus (6.9) für alle $n \geqslant N$

$$|x - x_n| \leqslant |x - x_{n_k}| < \varepsilon.$$

Das war zu zeigen. Der Beweis für eine fallende beschränkte Folge ist analog. □

6.3 Suprema, Infima und Häufungspunkte

Definition 6.10. Sei $(z_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{C}$. Ein $z \in \mathbb{C}$ heißt *Häufungspunkt* oder *Häufungswert* der Folge $(z_n)_{n\in\mathbb{N}}$, wenn es zu jedem $\varepsilon > 0$ unendlich viele Folgeglieder z_n gibt, so dass

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

gilt.

Bemerkung 6.11. Man beachte folgende einfache Tatsachen.

1. Eine Folge besitzt einen Häufungspunkt genau dann, wenn sie eine konvergente Teilfolge besitzt. Der Grenzwert dieser Teilfolge ist dann ein Häufungspunkt der ursprünglichen Folge.
2. Nach Bolzano-Weierstraß 6.4 bzw. Korollar 6.5 besitzt jede beschränkte Folge (in $\mathbb{R}$ oder $\mathbb{C}$) mindestens einen Häufungspunkt.
3. Eine konvergente Folge besitzt genau einen Häufungspunkt, nämlich ihren Grenzwert. Genauer gilt: Eine Folge konvergiert genau dann, wenn sie beschränkt ist und genau einen Häufungspunkt besitzt.

Beispiel 6.12. Betrachte folgende einfache Beispiele:

1. Die Folge $\big((-1)^n\big)_{n\in\mathbb{N}}$ hat die Häufungspunkte 1 und -1.
2. Die Folge $\big((-1)^n + \frac{1}{n}\big)_{n\in\mathbb{N}}$ hat auch die Häufungspunkte 1 und -1.
3. Die Folge $\big(n + (-1)^n n\big)_{n\in\mathbb{N}}$ hat einen Häufungspunkt 0.

Definition 6.13. Sei $\varnothing \neq M \subset \mathbb{R}$. Das *Supremum von M*

$$\sup M = \sup\{x \in M\} = \sup_{x\in M} x$$

ist die kleinste obere Schranke für M. Mit anderen Worten, es gilt

1. $\forall\, x \in M : x \leqslant \sup M$ und
2. $\forall\, S \in \mathbb{R} : (\forall\, x \in M : x \leqslant S) \implies \sup M \leqslant S$.

Wir definieren außerdem $\sup \varnothing = -\infty$.

Definition 6.14. Sei $\varnothing \neq M \subset \mathbb{R}$. Das *Infimum von M*

$$\inf M = \inf\{x \in M\} = \inf_{x \in M} x$$

ist die größte untere Schranke für M. Mit anderen Worten, es gilt

1. $\forall\, x \in M : x \geqslant \inf M$ und
2. $\forall\, S \in \mathbb{R} : (\forall\, x \in M : x \geqslant S) \implies \inf M \geqslant S$.

Wir definieren außerdem $\inf \varnothing = +\infty$.

Der Beweis des folgenden Lemmas ist dem Leser überlassen.

Lemma 6.15. *Wenn Infima oder Suprema einer Menge existieren, dann sind sie eindeutig bestimmmt.*

Beispiel 6.16. Es gilt

$$\sup\{\frac{1}{n} \mid n \in \mathbb{N}\} = 1 \quad \text{und} \quad \inf\{\frac{1}{n} \mid n \in \mathbb{N}\} = 0.$$

Definition 6.17. Sei A eine Menge, die ein Supremum besitzt. Falls das Supremum selbst ein Element von A ist, so nennen wir es *das Maximum von A*. Genauso sagen wir, dass A ein *Minimum besitzt*, wenn A ein Infimum besitzt und dieses Infimum eine Element von A ist.

Bemerkung 6.18. Im vorigen Beispiel sieht man den Unterschied zwischen Minimum und Infimum. Während das Minimum der Menge $\{\frac{1}{n} \mid n \in \mathbb{N}\}$ nicht existiert, gibt es das Infimum.

Satz 6.19. *Jede nicht-leere beschränkte Teilmenge von $\mathbb{R}$ besitzt ein Supremum und ein Infimum.*

Beweis. Sei $\varnothing \neq A \subset \mathbb{R}$ und beschränkt. Wir zeigen, dass A ein Supremum besitzt. Der Beweis der Existenz eines Infimums ist dann analog.

Wegen $A \neq \varnothing$ können wir ein $x_0 \in A$ auswählen. Weil A beschränkt ist, gibt es eine obere Schranke. Wir konstruieren jetzt induktiv eine Folge $[x_n, K_n]$ von Intervallen, so dass für alle $n \in \mathbb{N}$ folgende Eigenschaften erfüllt sind:

(1) $[x_{n+1}, K_{n+1}] \subset [x_n, K_{n+1}]$,

(2) $x_n \in A$,

(3) K_n ist eine obere Schranke von A,

(4) $K_n - x_n \leqslant 2^{-n}(K_0 - x_0)$.

Für $n = 0$ haben wir x_0 und K_0 schon konstruiert. Seien $x_0, \dots, x_n$ und $K_0, \dots, K_n$ schon konstruiert. Setze $M_n = \frac{K_n + x_n}{2}$. Dann gibt es zwei Fälle:

1. $A \cap]M, K_n] = \varnothing$. Dann ist M eine obere Schranke von A und wir definieren $x_{n+1} = x_n$ und $K_{n+1} = M$.

2. $A \cap]M, K_n] \neq \varnothing$. Dann gibt es ein $x_{n+1} \in A$ mit $x_{n+1} > M$. Wir setzen weiter $K_{n+1} = K_n$.

Es ist klar, dass auch für das neue Intervall $[x_{n+1}, K_{n+1}]$ die Eigenschaften (1) bis (4) gelten.

Nun ist $(K_n)_{n \in \mathbb{N}}$ eine monoton fallende und nach unten beschränkte Folge, denn es gilt z.B. $x_0 < K_n$ für alle $n \in \mathbb{N}$. Nach Satz 6.8 konvergiert die Folge gegen einen Grenzwert K. Wir behaupten jetzt, dass K das Supremum der Menge A ist.

Erstens gilt für alle $n \in \mathbb{N}$ und alle $a \in A$, dass $a \leqslant K_n$. Daraus folgt $a \leqslant K$. Also ist K eine obere Schranke.

Sei jetzt K' eine weitere obere Schranke für A. Wir nehmen an, es gälte $K' < K$, d.h. $K - K' > 0$. Dann gäbe es ein $n \in \mathbb{N}$, so dass

$$K_n - x_n \leqslant 2^{-n}(K_0 - x_0) < K - K' \leqslant K_n - K'.$$

Daraus folgt aber $K' < x_n$. Dies ist ein Widerspruch dazu, dass K' eine obere Schranke ist. Also ist die Annahme falsch und es muss $K \leqslant K'$ gelten. Insgesamt folgt $K = \sup A$. $\square$

Definition 6.20. Für eine Folge $(a_n)_{n \in \mathbb{N}}$ reeller Zahlen definieren wir den *Limes superior*

$$\limsup_{n \to \infty} a_n = \lim_{n \to \infty} \sup\{a_k \mid k \geqslant n\} \in \mathbb{R} \cup \{\pm\infty\}$$

und den *Limes inferior*

$$\liminf_{n \to \infty} a_n = \lim_{n \to \infty} \inf\{a_k \mid k \geqslant n\} \in \mathbb{R} \cup \{\pm\infty\}.$$

Beispiel 6.21. Man sieht:

1. Es gilt $\limsup_n (-1)^n = 1$ und $\liminf_n (-1)^n = -1$.
2. Es gilt $\limsup_n (n + (-1)^n n) = +\infty$ und $\liminf_n (n + (-1)^n n) = 0$.

Lemma 6.22. *Sei $(a_n)_{n \in \mathbb{N}}$ ein Folge in $\mathbb{R}$. Ist die Folge nicht nach unten beschränkt, dann gilt $\liminf_n a_n = -\infty$. Ist die Folge nach unten beschränkt, dann ist $\liminf_n a_n \in \mathbb{R}$ der kleinste Häufungspunkt der Folge. Analoge Aussagen gelten für $\limsup$.*

Der Beweis des Lemmas ist eine Übung.

7 Wurzeln in $\mathbb{R}$

Wir erinnern daran, dass $\mathbb{C}[x]$ die Menge aller Polynome mit Koeffizienten in $\mathbb{C}$ bezeichnet, entsprechend bezeichnet $\mathbb{R}[x]$ diejenigen Polynome mit reellen Koeffizienten.

Lemma 7.1. *Für jedes $p \in \mathbb{C}[x]$ von Grad $n \geqslant 1$ existiert ein $q \in \mathbb{C}[x]$, so dass für alle $a \in \mathbb{C}$*

$$p(x) - p(a) = (x - a)q(x)$$

gilt. Dabei hat q den Grad $n - 1$.

Beweis. In Lemma 2.40 haben wir für $n \geqslant 1$ die folgende Identität bewiesen:

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

Sei jetzt $p(x) = \sum_{j=0}^{n} b_j x^j \in \mathbb{C}[x]$ mit $b_n \neq 0$. Betrachte:

$$p(x) - p(a) = \sum_{j=0}^{n} b_j x^j - \sum_{j=0}^{n} b_j a^j = \sum_{j=1}^{n} b_j (x^j - a^j)$$

Man beachte, dass der konstante Term wegfällt und die Summe rechts tatsächlich erst bei $j = 1$ startet. Dann gilt mit Lemma 2.40:

$$p(x) - p(a) = \sum_{j=1}^{n} b_j (x^j - a^j) = \sum_{j=1}^{n} b_j \left((x-a) \sum_{k=0}^{j-1} x^{j-k-1} a^j \right) = (x-a)q(x)$$

wobei

$$q(x) = \sum_{j=1}^{n} a^j b_j \left(\sum_{k=0}^{j-1} x^{j-k-1} \right) \in \mathbb{C}[x]$$

mit Grad $n-1$. □

Korollar 7.2. *Ein Polynom vom Grad n hat höchstens n Nullstellen.*

Beweis. Wir beweisen diese Aussage durch vollständige Induktion über n. Ein Polynom $p(x) = a_1 X + a_0$ mit $a_1 \neq 0$ vom Grad 1 hat genau eine Nullstelle, nämlich $-\frac{a_0}{a_1}$. Sei die Aussage jetzt für alle Polynome von Grad $\leqslant n-1$ schon bewiesen, d.h. Polynome von Grad $\leqslant n-1$ haben höchstens $n-1$ Nullstellen. Wenn ein Polynom p vom Grad n eine Nullstelle a hat, so gilt nach Lemma 7.1

$$p(x) = p(x) - p(a) = (x-a)q(x).$$

Hier ist q von Grad $n-1$ und hat nach Induktionsannahme höchstens $n-1$ Nullstellen. Damit hat p höchstens n Nullstellen. □

Man beachte, dass dieses Korollar nichts über die Existenz von Nullstellen aussagt. Wir zeigen jetzt die Existenz von Nullstellen in $\mathbb{R}$ für bestimmte Polynome.

Definition 7.3. Sei $a \in \mathbb{C}$. Wir sagen $x_0 \in \mathbb{C}$ *ist eine Wurzel von* a, wenn x eine Nullstelle des Polynoms $x^2 - a$ ist. Allgemeiner, für eine $k \in \mathbb{N}, k \geqslant 2$ sagen wir $x_0 \in \mathbb{C}$ *ist eine k-te Wurzel von* a, wenn x eine Nullstelle des Polynoms $x^k - a$ ist.

Satz 7.4. *Sei $k \in \mathbb{N}, k \geqslant 2$ und $a \in \mathbb{R}, a > 0$. Dann konvergiert die rekursiv definierte Folge*

$$x_{n+1} = \frac{1}{k}\left((k-1)x_n + \frac{a}{x_n^{k-1}} \right)$$

für jeden Anfangswert $x_0 > 0$ gegen die einzige positive Lösung der Gleichung $x^k = a$.

Beweis. Weil aus $0 \leqslant y_1 < y_2$ die Ungleichung $y_1^k < y_2^k$ folgt, kann es höchstens eine positive Lösung geben.

Falls der Grenzwert x der Folge $(x_n)_{n\in\mathbb{N}}$ existiert, so ist er tatsächlich eine Lösung der Gleichung $x^k = a$. Dies folgt, indem wir zunächst die Rekursionsvorschrift mit kx_n^{k-1} multiplizieren:

$$kx_n^{k-1}x_{n+1} = (k-1)x_n^k + a.$$

Durch Übergang zum Grenzwert folgt:

$$kx^k = (k-1)x^k + a \implies x^k = a.$$

Also ist x eine k-te Wurzel.

Es bleibt noch zu zeigen, dass der Grenzwert x wirklich existiert bzw. dass die Folge $(x_n)_{n\in\mathbb{N}}$ konvergiert. Wir werden dazu zeigen, dass die Folge ab x_1 monoton fallend ist.

Wir bemerken zuerst, dass für alle $n \in \mathbb{N}$ die Folgeglieder x_n positiv sind. Das folgt leicht aus der Rekursionsformel durch Induktion, denn $a, x_0 > 0$. Man betrachte jetzt:

$$0 < x_{n+1} = \frac{1}{k}\left((k-1)x_n + \frac{a}{x_n^{k-1}}\right) = x_n\left(1 + \frac{1}{k}\big(\frac{a}{x_n^k} - 1\big)\right). \tag{7.5}$$

Daraus folgt für alle $n \in \mathbb{N}$

$$\frac{1}{k}\big(\frac{a}{x_n^k} - 1\big) > -1.$$

Damit ist die Bernoullische Ungleichung 2.22 anwendbar und es folgt für alle $n \in \mathbb{N}$:

$$\begin{aligned} & \Big(1 + \tfrac{1}{k}\big(\tfrac{a}{x_n^k} - 1\big)\Big)^k && \geqslant 1 + \big(\tfrac{a}{x_n^k} - 1\big) = \tfrac{a}{x_n^k} \\ \implies & x_{n+1}^k && = x_n^k\Big(1 + \tfrac{1}{k}\big(\tfrac{a}{x_n^k} - 1\big)\Big)^k \geqslant a \end{aligned}$$

Für alle $n \in \mathbb{N}$ mit $n \geqslant 1$ folgt also

$$\begin{aligned} & 0 < \tfrac{a}{x_n^k} \leqslant 1 \\ \implies & \tfrac{1}{k}\big(\tfrac{a}{x_n^k} - 1\big) \leqslant 0 \\ \implies & 1 + \tfrac{1}{k}\big(\tfrac{a}{x_n^k} - 1\big) \leqslant 1 \\ \implies & x_{n+1} = x_n\Big(1 + \tfrac{1}{k}\big(\tfrac{a}{x_n^k} - 1\big)\Big) \leqslant x_n \end{aligned}$$

Damit ist $(x_n)_{n\in\mathbb{N}}$ eine fallende Folge, die, wie wir oben schon gesehen haben, durch 0 nach unten beschränkt ist. Aus Satz 6.8 folgt, dass $(x_n)_{n\in\mathbb{N}}$ einen Grenzwert besitzt. (Man beachte, dass durchaus $x_0 < x_1$ gelten kann. Aber ab $n = 1$ ist $(x_n)_{n\in\mathbb{N}}$ monoton fallend!) □

Definition 7.6. Sei $a \in \mathbb{R}, a \geqslant 0$. Wir schreiben $\sqrt{a} = a^{\frac{1}{2}}$ für die nicht-negative Nullstelle des Polynoms $x^2 - a$, die nach Satz 7.4 existiert und eindeutig bestimmt ist. Für $k \in \mathbb{N}, k \geqslant 2$ schreiben wir allgemeiner $\sqrt[k]{a} = a^{\frac{1}{k}}$ für die nicht-negative Lösung der Gleichung $x^k = a$, die wieder nach Satz 7.4 existiert und eindeutig bestimmt ist.

Bemerkung 7.7. Für alle $a, b \in \mathbb{R}$ gilt offensichtlich:

$$a = b \implies a^2 = b^2.$$

Die Umkehrung gilt nicht! Tatsächlich gilt nur:

$$a^2 = b^2 \iff a = \pm b.$$

Das Quadrieren einer Gleichung ist keine Äquivalenzumformung. Es können neue Lösungen dazu kommen, wie man hier sieht. So gilt z.B.

$$\sqrt{a^2} = a = (\sqrt{a})^2$$

dann und nur dann, wenn $a \geqslant 0$.

Bemerkung 7.8. Sei $a > 0$. Es gilt:

$$0 = x^2 - a = (x - \sqrt{a})(x + \sqrt{a}).$$

Daher hat die Gleichung $x^2 = a$ genau eine positive Lösung $x = \sqrt{a}$ und eine negative Lösung $x = -\sqrt{a}$.

Korollar 7.9. *Die rationalen Zahlen sind nicht vollständig.*

Beweis. Die Folge $(x_n)_{n\in\mathbb{N}}$ aus dem vorigen Beweis besteht aus rationalen Zahlen, solange $a \in \mathbb{Q}$. Weil die Folge konvergiert, ist sie insbesondere eine Cauchy-Folge nach Satz 4.2. Wir haben also z.B. für $a = 2$ eine Cauchy-Folge in $\mathbb{Q}$, die gegen $\sqrt{2} \notin \mathbb{Q}$ konvergiert, siehe Lemma 2.4. □

Proposition 7.10. *Es gilt:*

1. *Für alle* $a > 0$ *gilt:* $\lim_n \sqrt[n]{a} = 1$.
2. $\lim_{n\to\infty} \sqrt[n]{n} = 1$.

Beweis. Zu 1.: 1. Fall $a > 1$: Setze $x_n = \sqrt[n]{a} - 1 > 0$. Aus der Bernoulli-Ungleichung 2.22 folgt dann:

$$\begin{aligned} & a = (1 + x_n)^n \geqslant 1 + n x_n \\ \Longrightarrow\quad & 0 < x_n \leqslant \frac{a-1}{n} \\ \Longrightarrow\quad & \lim_n(\sqrt[n]{a} - 1) = \lim_n x_n = 0 \\ \Longrightarrow\quad & \lim_n \sqrt[n]{a} = 1 \end{aligned}$$

(Hier gehen natürlich auch die Sätze 3.15 und 3.23 ein.)
2. Fall $a = 1$: Dann ist die Folge $(\sqrt[n]{1})_{n\in\mathbb{N}}$ konstant 1.
3. Fall $0 < a < 1$: Die Aussage folgt aus Teil 1, wenn man zu a^{-1} übergeht.

Zu 2.: Setze $x_n = \sqrt[n]{n} - 1 > 0$. Aus der binomischen Formel 2.39 folgt dann:

$$\begin{aligned} & n = (1 + x_n)^n \geqslant \frac{n(n-1)}{2} x_n^2 \\ \Longrightarrow\quad & 0 < x_n \leqslant \sqrt{\frac{2}{n-1}} \\ \Longrightarrow\quad & \lim_n x_n = 0 \\ \Longrightarrow\quad & \lim_n \sqrt[n]{n} = 1 \end{aligned}$$

□

Definition 7.11. Für ein $q = \frac{n}{m} \in \mathbb{Q}$ und $x \in \mathbb{R}, x > 0$ definieren wir

$$x^q = x^{\frac{n}{m}} = \sqrt[m]{x^n}.$$

Wir werden später in Abschnitt 9.2 x^α für alle $\alpha \in \mathbb{R}$ definieren.

8 Konvergenzkriterien für Reihen

8.1 Das Monotoniekriterium

Satz 8.1 (Monotoniekriterium)**.** *Es sei $\sum_{k=1}^{\infty} a_k$ eine Reihe mit $a_k \in \mathbb{R}, a_k \geqslant 0$ für alle $k \in \mathbb{N}$. Die Reihe konvergiert genau dann, wenn die Folge ihrer Partialsummen beschränkt ist.*

Beweis. Sei $s_n = \sum_{k=1}^{n} a_k$. Da $a_k \geqslant 0$ für alle $k \in \mathbb{N}$, folgt für alle $n \in \mathbb{N}$

$$s_{n+1} = s_n + a_{n+1} \geqslant s_n.$$

Das bedeutet, dass die Folge $(s_n)_{n\in\mathbb{N}}$ der Partialsummen monoton steigt. Nach Voraussetzung ist diese Folge auch beschränkt. Also konvergiert die Folge nach Satz 6.8. □

Proposition 8.2. *Für alle $s \in \mathbb{Q}$ gilt:*

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \begin{cases} \textit{konvergiert, falls} & s > 1 \\ \textit{divergiert, falls} & s \leqslant 1 \end{cases}$$

Beweis. Sei $s > 1$. Für $n \in \mathbb{N}$ wähle $p \in \mathbb{N}$, so dass $2^p - 1 \geqslant n$. Es folgt

$$s_{2^p-1} = s_n + \sum_{k=n+1}^{2^p-1} \frac{1}{k^s} \geqslant s_n,$$

denn $k^{-s} > 0$. Weiter gilt:

$$\begin{aligned} s_n \leqslant s_{2^p-1} &= \sum_{k=1}^{2^p-1} \frac{1}{k^s} = \sum_{\ell=0}^{p-1} \sum_{k=2^\ell}^{2^{\ell+1}-1} \frac{1}{k^s} \leqslant \sum_{\ell=0}^{p-1} \frac{2^\ell}{2^{\ell s}} = \sum_{\ell=0}^{p-1} \left(\frac{1}{2^{s-1}}\right)^\ell \\ &< \sum_{\ell=0}^{\infty} \left(\frac{1}{2^{s-1}}\right)^\ell = \frac{1}{1-2^{1-s}}, \end{aligned}$$

denn die geometrische Reihe für $\frac{1}{2^{s-1}}$ konvergiert, weil gilt:

$$s > 1 \implies 2^{s-1} > 1 \implies 0 < \frac{1}{2^{s-1}} < 1.$$

Wir haben damit eine obere Schranke für alle s_n gefunden. Die Konvergenz der Reihe $\sum_{n=1}^{\infty} \frac{1}{n^s}$ folgt jetzt aus Satz 8.1, denn $n^s \geqslant 0$ für alle $n \in \mathbb{N}$.

Die Divergenz der Reihe für $s \leqslant 1$ folgt aus der Tatsache, dass de harmonische Reihe $(s = 1)$ divergiert und dem weiter unten bewisenen Minorantenkriterium **ref**. □

Bemerkung 8.3. Wir haben gerade gezeigt, dass für alle $s \in \mathbb{R}, s > 1$ die *Riemannsche Zeta-Funktion* gegeben durch

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

wohldefiniert ist. Es ist aber einfach zu zeigen, dass ζ durch dieselbe Formel für alle $s \in \mathbb{C}$ mit $\mathrm{Re}\,(s) > 1$ definiert werden kann. Mit Mitteln der Funktionentheorie kann man dann die Funktion (meromorph) auf die ganze komplexe Zahlenebene fortsetzen, wobei $s = 1$ der einzige Pol ist. Die Frage nach Nullstellen dieser Funktion ist die wichtigste offene Frage der Mathematik. Man kann zeigen, dass für alle negativen ganzen Zahlen $s = -2n, n \in \mathbb{N}$ die Funktion Nullstellen hat. Diese heißen die "trivialen Nullstellen" der Riemannschen Zeta-Funktion. Die *Riemannsche Vermutung* besagt, dass sich alle anderen sogenannten nicht trivialen Nullstellen der Riemannschen Zeta-Funktion auf der Gerade $\mathrm{Re}\,(s) = \frac{1}{2}$ befinden. Wenn Sie diese Vermutung beweisen, ist Ihnen nicht nur eine Million USD, sondern auch ewiger Ruhm sicher. Ich bin mir ziemlich sicher, dass das Bestehen der Analysis I Klausur eine notwendige Bedingung dafür ist. Schauen Sie doch mal folgende Seiten an:
`https://de.wikipedia.org/wiki/Bernhard_Riemann`
`https://de.wikipedia.org/wiki/Riemannsche_%CE%B6-Funktion`
`https://de.wikipedia.org/wiki/Riemannsche_Vermutung`
`https://www.youtube.com/watch?v=W4GQSGSOOpg` (speziell ab 5:00 bis 8:15, sie werden es wiedererkennen!)

Und weil es gerade passt:
Leonhard Euler `https://de.wikipedia.org/wiki/Leonhard_Euler`

Die Formel, die Euler schon als jungen Mann berühmt machte, lautet:

$$\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} = 1 + \frac{1}{4} + \frac{1}{9} + \ldots = \frac{\pi^2}{6}.$$

Wir werden diese Formel nicht beweisen. Sie wird aber im obigen Video in den ersten 15 Minuten bewiesen.

8.2 Leibniz-Konvergenzkriterium für alternierende Reihen

Gottfried Wilhelm Leibniz `https://de.wikipedia.org/wiki/Gottfried_Wilhelm_Leibniz`

Definition 8.4. Wir sagen, dass $\sum_{n=0}^{\infty} a_n$ eine *alternierende Reihe* ist, wenn für alle $n \in \mathbb{N}$ reelle Zahlen $b_n \in \mathbb{R}$ mit $b_n \geqslant 0$ gibt, so dass $a_n = (-1)^n b_n$.

Satz 8.5 (Leibniz-Konvergenzkriterium für alternierende Reihen)**.** *Sei $(b_n)_{n \in \mathbb{N}}$ eine reelle fallende Nullfolge.*

1. *Dann konvergiert die alternierende Reihe $\sum_{n=0}^{\infty}(-1)^n b_n$.*
2. *Sei $s_n = \sum_{k=0}^{n}(-1)^k b_k$ die k-te Partialsumme und $S = \sum_{n=0}^{\infty}(-1)^n b_n$. Dann gilt für alle $n \in \mathbb{N}$:*

$$|S - s_n| = |\sum_{k=n+1}^{\infty}(-1)^n b_n| \leqslant b_{n+1}.$$

Beweis. Wenn $(b_n)_{n\in\mathbb{N}}$ eine reelle fallende Folge mit $\lim_n b_n = 0$ ist, dann folgt aus einer Version von Lemma 4.12 für fallende Folgen, dass $b_n \geqslant 0$ für alle $n \in \mathbb{N}$.

Sei jetzt $s_n = \sum_{k=0}^{n}(-1)^n b_n$.

Zu 1.: Weil $(b_n)_{n\in\mathbb{N}}$ eine fallende Folge und $b_n \geqslant 0$ für alle $n \in \mathbb{N}$, folgt

$$s_{2n+2} - s_{2n} = b_{2n+2} - b_{2n+1} \leqslant 0.$$

Also $s_{2n} \geqslant s_{2n+2}$. Die Folge $(s_{2n})_{n\in\mathbb{N}}$ ist eine monoton fallende Folge, die durch 0 nach unten beschränkt ist. Aus Satz 6.8 folgt, dass $(s_{2n})_{n\in\mathbb{N}}$ einen Grenzwert S besitzt. Wieder mit der Version von Lemma 4.12 für fallende Folgen, sieht man für alle $n \in \mathbb{N}$: $s_{2n} \geqslant S$.

Die Folge $(s_{2n+1})_{n\in\mathbb{N}}$ hingegen ist monoton steigend: für alle $n \in \mathbb{N}$ folgt

$$s_{2n+3} - s_{2n+1} = b_{2n+2} - b_{2n+3} \geqslant 0.$$

Die Folge ist nach oben durch b_0 beschränkt und besitzt daher nach Satz 6.8 einen Grenzwert S'. Mit Lemma 4.12 folgt für alle $n \in \mathbb{N}$: $s_{2n+1} \leqslant S'$.

Es gilt mit Satz 3.15:

$$S - S' = \lim_{n\to\infty} s_{2n} - \lim_{n\to\infty} s_{2n-1} = \lim_{n\to\infty}(s_{2n} - s_{2n-1}) = \lim_{n\to\infty} b_{2n} = 0.$$

Also $S = S'$.

Wir behaupten jetzt:

$$\sum_{n=0}^{\infty}(-1)^n b_n = S$$

Sei $\varepsilon > 0$ gegeben. Dann gibt es $N_1, N_2 \in \mathbb{N}$, so dass für alle $n \geqslant N_1$

$$|S - s_{2n}| < \varepsilon$$

und für alle $n \geqslant N_2$

$$|S - s_{2n+1}| < \varepsilon$$

gilt. Setze jetzt $N = \max\{2N_1, 2N_2 + 1\}$. Dann gilt für alle $n \geqslant N$:

$$|S - s_n| < \varepsilon.$$

Dies zeigt die Konvergenz und damit Aussage 1.

Zu 2.: Wir haben oben schon gezeigt, dass für alle $n \in \mathbb{N}$ gilt:

$$s_{2n-1} \leqslant S \leqslant s_{2n} \quad \text{und natürlich genauso} \quad s_{2n+1} \leqslant S \leqslant s_{2n}.$$

Damit folgt aber:

$$|S - s_{2n}| \leqslant |s_{2n-1} - s_{2n}| = b_{2n} \quad \text{und} \quad |S - s_{2n+1}| \leqslant |s_{2n} - s_{2n+1}| = b_{2n+1}$$

Somit haben wir den geraden und ungeraden Fall abgearbeitet. Dies beweist Teil 2. □

Beispiel 8.6. Die Reihe

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

konvergiert. Dies folgt direkt aus dem Leibniz-Kriterium 8.5, denn $\left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ ist eine monoton fallende Nullfolge.

Tatsächlich gilt:

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = \ln 2.$$

Wir werden das später **ref** zeigen.

Beispiel 8.7. Die Reihe

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots$$

konvergiert nach Satz 8.5, denn $\left(\frac{1}{2n+1}\right)_{n\in\mathbb{N}}$ ist eine monoton fallende Nullfolge. Leibniz konnte sogar den Grenzwert bestimmen.

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4}.$$

Wir werden das später **ref** zeigen.

8.3 Absolute Konvergenz

Definition 8.8. Eine Reihe $\sum_{n=0}^{\infty} a_n$ von komplexen Zahlen $a_0, a_1, \dots$ heißt *absolut konvergent*, wenn die Reihe

$$\sum_{n=0}^{\infty} |a_n|$$

konvergiert.

Lemma 8.9. *Eine absolut konvergent Reihe konvergiert.*

Beweis. Weil $\sum_{n=0}^{\infty} |a_n|$ konvergiert, gibt es nach dem Cauchy-Kriterium 5.3 zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so das für alle $n \geqslant m \geqslant N$ gilt:

$$\left| \sum_{k=m}^{n} a_k \right| \leqslant \sum_{k=m}^{n} |a_k| < \varepsilon.$$

Wieder mit Cauchy-Kriterium folgt daraus die Konvergenz von $\sum_{n=0}^{\infty} a_n$. $\square$

Bemerkung 8.10. Man beachte:

1. Für Reihen von reellen positiven Zahlen stimmt Konvergenz mit absoluter Konvergenz überein.

2. Konvergenz impliziert nicht ohne weitere Bedingung absolute Konvergenz. Als Beispiel betrachte man die Reihe

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}.$$

Wir wissen wegen des Leibniz-Kriteriums 8.5, dass diese Reihe konvergiert. Aber wir wissen auch 5.6, dass die Reihe

$$\sum_{n=1}^{\infty} \left| \frac{(-1)^{n-1}}{n} \right| = \sum_{n=1}^{\infty} \frac{1}{n}$$

divergiert.

3. Im Abschnitt 8.8 werden wir sehen, dass der Begriff der Konvergenz einer Reihe nicht unserer Intuition entspricht. Es sind die absolut konvergenten Reihen, die sich gut verhalten, und zwar nicht nur im Zusammenhang mit Umordnungen, sondern in vielen anderen Hinsichten.

8.4 Das Majorantenkriterium

Definition 8.11. Seien $(a_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ Folge in $\mathbb{R}$ mit $0 \leqslant a_n \leqslant c_n$ für alle $n \in \mathbb{N}$. Dann heißt $\sum_{n=0}^{\infty} c_n$ eine *Majorante* von $\sum_{n=0}^{\infty} a_n$.

Satz 8.12 (Majorantenkriterium). *Seien $(a_n)_{n\in\mathbb{N}}$ und $(c_n)_{n\in\mathbb{N}}$ Folge in $\mathbb{C}$ und es gelte:*

1. $|a_n| \leqslant |c_n|$ *für alle* $n \in \mathbb{N}$. $\left(\sum_{n=0}^{\infty} |c_n|\right.$ *sei eine Majorante von* $\left.\sum_{n=0}^{\infty} |a_n|.\right)$
2. *Die Reihe* $\sum_{n=0}^{\infty} |c_n|$ *konvergiere.*

Dann ist die Reihe $\sum_{n=0}^{\infty} a_n$ *absolut konvergent und es gilt:*

$$\left| \sum_{n=0}^{\infty} a_n \right| \leqslant \sum_{n=0}^{\infty} |c_n|.$$

Beweis. Weil $\sum_{n=0}^{\infty} |c_n|$ konvergiert, gibt es nach dem Cauchy-Kriterium 5.3 zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so das für alle $n \geqslant m \geqslant N$ gilt:

$$\left| \sum_{k=m}^{n} a_k \right| \leqslant \sum_{k=m}^{n} |a_k| \leqslant \sum_{k=m}^{n} |c_k| < \varepsilon.$$

Wieder mit Cauchy-Kriterium folgt daraus die Konvergenz von $\sum_{n=0}^{\infty} |a_n|$. Die behauptete Abschätzung des Grenzwertes in Teil 2 folgt jetzt:

$$\begin{aligned} \left| \sum_{k=0}^{\infty} a_k \right| &= \left| \lim_{n\to\infty} \sum_{k=0}^{n} a_k \right| = \lim_{n\to\infty} \left| \sum_{k=0}^{n} a_k \right| \\ &\leqslant \lim_{n\to\infty} \sum_{k=0}^{n} |a_k| \leqslant \lim_{n\to\infty} \sum_{k=0}^{n} |c_k| = \sum_{k=0}^{\infty} |c_k| \end{aligned}$$

Dabei benutzen wir Satz 3.21 und das folgende Lemma. □

Lemma 8.13. *Sei $(z_n)_{n\in\mathbb{N}}$ eine Folge komplexer Zahlen, die gegen $z \in \mathbb{C}$ konvergiere. Dann gilt:*

$$\lim_{n\to\infty} |z_n| = |\lim_{n\to\infty} z_n| = z.$$

Beweis. Für alle $z, w \in \mathbb{C}$ gilt:

$$\big||z| - |w|\big| \leqslant |z - w|.$$

Diese Ungleichung folgt aus der Dreiecksungleichung 2.26(8) in $\mathbb{C}$, wobei der Beweis wörtlich derselbe ist wie der Beweis von Lemma 2.19(5).

Wenn $\lim_{n\to\infty} z_n = z$, dann gibt es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $n \geqslant N$ gilt:

$$\big||z| - |z_n|\big| \leqslant |z - z_n| < \varepsilon.$$

□

Korollar 8.14. *Sei $\sum_{n=0}^{\infty} c_n$ eine divergente Reihe reeller Zahlen mit $c_n \geqslant 0$ für alle $n \in \mathbb{N}$. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ mit $a_n \geqslant c_n$. Dann divergiert auch die Reihe $\sum_{k=0}^{\infty} a_n$.*

Beweis. Wenn die Reihe $\sum_{k=0}^{\infty} a_n$ konvergieren würde, so wäre sie eine Majorante für die Reihe $\sum_{n=0}^{\infty} c_n$, die nach dem Majorantenkriterium auch konvergieren müsste. □

8.5 Das Quotientenkriterium

Definition 8.15. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge. Wir sagen, dass eine Bedingung für *fast alle* a_n gilt, wenn die Bedingung für alle außer endlich vielen a_n gilt.

Satz 8.16 (Quotientenkriterium)**.** *Sei $\sum_{n=0}^{\infty} a_n$ eine Reihe von komplexen Zahlen a_n und es gelte $a_n \neq 0$ für fast alle $n \in \mathbb{N}$. Weiter existiere der Grenzwert*

$$q = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|.$$

Dann gilt:

1. *Wenn $q < 1$, so ist $\sum_{n=0}^{\infty} a_n$ absolut konvergent.*
2. *Wenn $q > 1$, so divergiert $\sum_{n=0}^{\infty} a_n$.*

Beweis. Wir betrachten zuerst den Fall

$$q = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1.$$

Wir wählen $q_1 \in \mathbb{R}$, so dass $q < q_1 < 1$ und setzen $\varepsilon = q_1 - q > 0$. Dann existiert ein $N \in \mathbb{N}$, so dass für alle $n \geqslant N$ gilt:

$$\left| q - \left|\frac{a_{n+1}}{a_n}\right| \right| < \varepsilon = q_1 - q.$$

Daraus folgt:

$$\left|\frac{a_{n+1}}{a_n}\right| < q + \varepsilon = q_1.$$

Durch wiederholte Anwendung dieser Ungleichung erhalten wir folgende Ungleichungen:

$$|a_n| \leqslant q_1|a_{n-1}| \leqslant \ldots \leqslant |a_N| q_1^{n-N}.$$

Also gilt für alle $n \geqslant N$:

$$\begin{aligned}
\sum_{k=N}^{n} |a_k| &\leqslant |a_N| \sum_{k=N}^{n} q_1^{k-N} = |a_N| \sum_{k=0}^{n-N} q_1^k < |a_N| \sum_{k=0}^{\infty} q_1^k \\
&= |a_N| \tfrac{1}{1-q_1} \\
\Longrightarrow \quad \sum_{k=0}^{n} |a_k| &\leqslant |a_N| \tfrac{1}{1-q_1} + \sum_{k=0}^{N-1} |a_K|
\end{aligned}$$

Aus dem Monotoniekriterium 8.1 folgt jetzt die Konvergenz der Reihe $\sum_{n=0}^{\infty} a_n$.

Sei jetzt $q > 1$. Wir wählen $q_1 \in \mathbb{R}$ mit $q > q_1 > 1$. Dann folgt wie oben, dass es ein $N \in \mathbb{N}$ gibt, so dass für alle $n \geqslant N$ gilt:

$$\left| \frac{a_{n+1}}{a_n} \right| \geqslant q_1.$$

Es folgt für alle $n \geqslant N$:

$$|a_n| \geqslant |a_N| q_1^{n-N}.$$

Weil $q_1 > 1$ ist, divergiert die Folge $(|a_n|)_{n \in \mathbb{N}}$. Damit ist die Folge $(a_n)_{n \in \mathbb{N}}$ keine Nullfolge und die Reihe $\sum_{n=0}^{\infty} a_n$ kann wegen Proposition 5.4 nicht konvergieren. □

Bemerkung 8.17. Im Fall $q = \lim_n \left| \frac{a_{n+1}}{a_n} \right| = 1$ ist keine Aussage möglich. Als Beispiel hierfür dient die Aussage von Proposition 8.2:

$$\sum_{n=1}^{\infty} \frac{1}{n^s} \quad \begin{cases} \text{konvergiert, falls} & s > 1 \\ \text{divergiert, falls} & s \leqslant 1 \end{cases}$$

Dabei gilt

$$\left| \frac{a_{n+1}}{a_n} \right| = \left(\frac{n}{n+1} \right)^s \to 1$$

für $n \to \infty$.

Bemerkung 8.18. Die Bedingung des Quotientenkriteriums für Konvergenz ist nicht äquivalent zu der Bedingung:

$$\exists\ N \in \mathbb{N}\ \forall\ n \geqslant N \quad \left| \frac{a_{n+1}}{a_n} \right| < 1. \tag{8.19}$$

Die Bedingung lautet:

$$\exists\ 0 < q_1 < 1\ \exists\ N \in \mathbb{N}\ \forall\ n \geqslant N \quad \left| \frac{a_{n+1}}{a_n} \right| < q_1,$$

und dabei ist q_1 unabhängig von n! Dass die Bedingung (8.19) nicht hinreichend für die Konvergenz einer Reihe ist, sieht man am einfachsten bei der harmonischen Reihe. Es gilt zwar

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{n}{n+1} < 1$$

für alle $n \in \mathbb{N}$, aber die Reihe divergiert. Es gibt kein $q_1 < 1$ mit

$$\left|\frac{a_{n+1}}{a_n}\right| = \frac{n}{n+1} < q_1 < 1,$$

wie im Quotientenkriterium gefordert, denn $\lim_n \frac{n}{n+1} = 1$.

Bemerkung 8.20. Das Quotientenkriterium gibt eine hinreichende Bedingung für Konvergenz. Diese Bedingung ist aber nicht notwendig. Als Beispiel betrachte man

$$\sum_{n=1}^{\infty} \frac{1}{n^2}.$$

Wir wissen schon, dass diese Reihe konvergiert, aber das Quotientenkriterium ist nicht erfüllt, denn es gilt:

$$\lim_{n\to\infty} \frac{n^2}{(n+1)^2} = 1.$$

Es kann also kein $q_1 < 1$ wie im Quotientenkriterium geben.

8.6 Das Wurzelkriterium

Satz 8.21 (Wurzelkriterium)**.** *Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{C}$ und*

$$q = \limsup_{n\to\infty} \sqrt[n]{|a_n|}.$$

Dann gilt:

1. *Wenn $q < 1$, so ist $\sum_{n=0}^{\infty} a_n$ absolut konvergent.*
2. *Wenn $q > 1$, so divergiert $\sum_{n=0}^{\infty} a_n$.*

Beweis. Wir betrachten zuerst den Fall $q < 1$. Wir wählen $q_1 \in \mathbb{R}$, so dass $q < q_1 < 1$ und setzen $\varepsilon = q_1 - q > 0$. Der Limes superior q ist der größte Häufungspunkt der Folge $\left(\sqrt[n]{|a_n|}\right)_{n\in\mathbb{N}}$, d.h. zu beliebigem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass für alle $n \geqslant N$

$$\sqrt[n]{|a_n|} \leqslant q + \varepsilon = q_1.$$

gilt. Weil $0 < q < 1$, folgt:

$$\forall n \geqslant N : |a_n| \leqslant q^n$$

$$\implies \quad \sum_{k=N}^{\infty} |a_k| \leqslant \sum_{k=N}^{\infty} q^k \leqslant \sum_{k=0}^{\infty} q^k = \frac{1}{1-q}$$

Damit haben wir gezeigt, dass die Reihe $\sum_{n=N}^{\infty} |a_n|$ eine konvergente Majorante $\sum_{k=N}^{\infty} q^k$. Nach dem Majorantenkriterium 8.12 konvergiert die Reihe $\sum_{n=N}^{\infty} |a_n|$ und dann natürlich auch die Reihe $\sum_{n=0}^{\infty} |a_n|$.

Sei jetzt $1 < q = \limsup_n \sqrt[n]{|a_n|}$. Weil der Limes superior der größte Häufungspunkt ist, existieren unendlich viele $n \in \mathbb{N}$ mit

$$\sqrt[n]{|a_n|} < 1.$$

Insbesondere gilt für diese n auch $|a_n| > 1$, und damit ist $(a_n)_{n\in\mathbb{N}}$ keine Nullfolge. Nach Proposition 5.4 kann die Reihe $\sum_n a_n$ dann nicht konvergieren. □

Bemerkung 8.22. Es gelten dieselben Vorsichtshinweise wie für das Quotientenkriterium.

1. Im Fall $q = \limsup_n \sqrt[n]{|a_n|} = 1$ kann keine Aussage getroffen werden.
2. Es reicht für das Wurzelkriterium **nicht** aus, dass

$$\sqrt[n]{|a_n|} < 1$$

für alle $n \in \mathbb{N}$.

Bemerkung 8.23. Für eine beliebige Folge $(a_n)_{n\in\mathbb{N}}$ in $\mathbb{C}$ kann man

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} \leqslant \limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$$

zeigen. Das bedeutet, dass das Wurzelkriterium stärker ist als das Quotientenkriterium; es liefert öfter eine hinreichende Bedingung für absolute Konvergenz. Tatsächlich gibt es Beispiele für absolut konvergente Reihen, die das Wurzelkriterium, aber nicht das Quotientenkriterium erfüllen.

8.7 Summen und Cauchy-Produkte von Reihen

Satz 8.24. *Seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ konvergente Reihen. Dann konvergiert auch für alle $\alpha, \beta \in \mathbb{C}$ die Reihe*

$$\sum_{n=0}^{\infty} (\alpha a_n + \beta b_n)$$

und es gilt

$$\sum_{n=0}^{\infty} (\alpha a_n + \beta b_n) = \alpha \cdot \sum_{n=0}^{\infty} a_n + \beta \cdot \sum_{n=0}^{\infty} b_n.$$

Bemerkung 8.25. Warnung: Aus der Konvergenz von $\sum_{n=0}^{\infty}(\alpha a_n + \beta b_n)$ kann man nicht auf die Konvergenz von $\sum_{n=0}^{\infty} a_n$ oder $\sum_{n=0}^{\infty} b_n$ schließen.

Beweis. Man betrachte die Partialsummen

$$\sum_{k=0}^{n} (\alpha a_k + \beta b_k) = \alpha \cdot \sum_{k=0}^{n} a_k + \beta \cdot \sum_{k=0}^{\infty} b_k.$$

Die Aussage folgt dann aus Satz 3.15. □

Bemerkung 8.26. Für zwei endliche Summen $\sum_{i=0}^{n} a_i$ and $\sum_{j=0}^{n} b_j$:

$$\left(\sum_{i=0}^{n} a_i\right)\left(\sum_{j=0}^{n} b_j\right) = \sum_{k=0}^{n} \sum_{i+j=k} a_i b_j.$$

Definition 8.27. Seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ komplexe Reihen. Setze

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{i+j=n} a_i b_j.$$

Dann heißt die Reihe $\sum_{k=0}^{\infty} c_n$ das *Cauchy-Produkt* der Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$.

Satz 8.28. *Seien* $\sum_{n=0}^{\infty} a_n$ *und* $\sum_{n=0}^{\infty} b_n$ *absolut konvergente komplexe Reihen. Dann konvergiert das Cauchy-Produkt*

$$\sum_{k=0}^{\infty} c_n$$

der beiden Reihen mit

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{i+j=n} a_i b_j,$$

und es gilt:

$$\sum_{k=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right).$$

Beweis. Für ein $N \in \mathbb{N}$ betrachten wir die Teilmengen von $\mathbb{N} \times \mathbb{N}$

$$\begin{aligned} Q_N &= \{(m,n) \in \mathbb{N} \times \mathbb{N} \mid 0 \leqslant m, n \leqslant N\} \\ \Delta_N &= \{(m,n) \in \mathbb{N} \times \mathbb{N} \mid 0 \leqslant m + n \leqslant N\}. \end{aligned}$$

Setze

$$A_N = \sum_{n=0}^{N} a_n, \ A_N^{|-|} = \sum_{n=0}^{N} |a_n|, \quad B_N = \sum_{n=0}^{N} b_n \quad \text{und} \quad B_N^{|-|} = \sum_{n=0}^{N} |b_n|.$$

Dann gilt für alle $N \in \mathbb{N}$:

$$\begin{aligned} C_N &= \sum_{n=0}^{N} c_n = \sum_{(m,n)\in\Delta_N} a_m b_n \\ A_N B_N &= \left(\sum_{n=0}^{N} a_n\right)\left(\sum_{n=0}^{N} b_n\right) = \sum_{(m,n)\in Q_N} a_m b_n \\ A_N^{|-|} B_N^{|-|} &= \left(\sum_{n=0}^{N} |a_n|\right)\left(\sum_{n=0}^{N} |b_n|\right) = \sum_{(m,n)\in Q_N} |a_m b_n| \end{aligned}$$

und

$$\Delta_N \subset Q_N \subset \Delta_{2N} \subset Q_{2N}.$$

Jetzt folgt, dass $\sum_{n=0}^{\infty} c_n$ absolut konvergiert (und insbesondere konvergiert), denn

$$\begin{aligned} \sum_{n=0}^{\infty} |c_n| &= \lim_{N\to\infty} \left| \sum_{(m,n)\in\Delta_N} a_m b_n \right| \leqslant \lim_{N\to\infty} \sum_{(m,n)\in\Delta_N} |a_m b_n| \leqslant \lim_{N\to\infty} \sum_{(m,n)\in Q_N} |a_m b_n| \\ &= \lim_{N\to\infty} A_N^{|-|} B_N^{|-|} = \left(\lim_{N\to\infty} A_N^{|-|}\right)\left(\lim_{N\to\infty} B_N^{|-|}\right) = \left(\sum_{n=0}^{\infty} |a_n|\right)\left(\sum_{n=0}^{\infty} |b_n|\right), \end{aligned}$$

wobei der letzte Term ein Produkt endlicher Zahlen ist, denn $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ konvergieren absolut. (Aus den Teilmengenbeziehungen oben oder aus dem, was wir jetzt noch zeigen, folgt, dass diese Ungleichungen tatsächlich Gleichungen sind.)

Es bleibt die Formel für den Grenzwert $\sum_{n=0}^{\infty} c_n$ zu zeigen. Wir betrachten:

$$\begin{aligned}|A_{2N}B_{2N} - C_{2N}| = \left| \sum_{(m,n)\in Q_{2N}-\Delta_{2N}} a_m b_n \right| &\leqslant \sum_{(m,n)\in Q_{2N}-\Delta_{2N}} |a_m b_n| \\ &\leqslant \sum_{(m,n)\in Q_{2N}-Q_N} |a_m b_n| = A_{2N}^{|-|} B_{2N}^{|-|} - A_N^{|-|} B_N^{|-|}\end{aligned}$$

Nun wissen wir, dass die Folge $(A_N^{|-|} B_N^{|-|})_{N\in\mathbb{N}}$ konvergiert, ihr Grenzwert ist $\sum_{n=0}^{\infty} |c_n|$. Deshalb ist diese Folge eine Cauchy-Folge. Daher finden wir zu jedem $\varepsilon > 0$ ein $M \in \mathbb{N}$, so dass für alle $N \geqslant M$ gilt:

$$|A_{2N}B_{2N} - C_{2N}| \leqslant A_{2N}^{|-|} B_{2N}^{|-|} - A_N^{|-|} B_N^{|-|} < \varepsilon.$$

Es folgt:

$$\sum_{k=0}^{\infty} c_n = \lim_{N\to\infty} C_{2N} = \lim_{N\to\infty} A_{2N}B_{2N} = \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right).$$

□

8.8 Umordnung von Reihen

Satz 8.29. *Sei $\sum_{n=1}^{\infty} a_n$ eine konvergente, aber nicht absolut konvergente Reihe. Dann existiert zu jedem $a \in \mathbb{R} \cup \{\pm\infty\}$ eine Bijektion $\sigma\colon \mathbb{N} \to \mathbb{N}$, so dass*

$$\sum_{n=1}^{\infty} a_{\sigma(n)} = a.$$

Satz 8.30. *Sei $\sum_{n=1}^{\infty} a_n$ eine absolut konvergente Reihe. Sei $\sigma\colon \mathbb{N} \to \mathbb{N}$ eine Bijektion und setze $b_n = a_{\sigma(n)}$. Dann ist auch $\sum_{n=1}^{\infty} b_n$ absolut konvergent und es gilt:*

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} b_n.$$

9 Potenzreihen

9.1 Potenzreihen und ihr Konvergenzradius

Definition 9.1. Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{C}$ und $z \in \mathbb{C}$. Dann heißt die Funktion

$$P(z) = \sum_{n=0}^{\infty} a_n z^n$$

Potenzreihe in z mit Koeffizienten in $\mathbb{C}$. Wir schreiben $\mathbb{C}[[z]]$ ($\mathbb{R}[[z]]$) für die Menge der Potenzreihe in z mit Koeffizienten in $\mathbb{C}$ ($\mathbb{R}$).

Potenzreihen sind Verallgemeinerungen von Polynomen. Warum diese Verallgemeinerung? Wie schon bei der Einführung immer grßerer Zahlenbereiche werden wir hier vom Wunsch getrieben, Lösungen von Gleichungen zu bekommen. Um das zu erklären, greifen wir voraus und betrachten die Ableitung f' einer differenzierbaren Funktion f. Wir möchten gerne Lösungen folgender Differentialgleichungen hinschreiben:

$$f' = f$$

Diese Gleichung taucht bei natürlichen Wachstums- oder Zerfallsprozessen auf. Die nächste Gleichung

$$f'' = -f$$

beschreibt einen harmonischen Oszillator. Sie können sich leicht davon überzeugen, dass beide Gleichung keine Polynome als Lösungen zulassen: Ableiten verringert den Grad eines Polynoms, daher ist der Grad auf der linken Seite strikt kleiner als der Grad auf der rechten Seite. Wir brauchen also einen größeren Vorrat an Funktionen. Dazu dienen die Potenzreihen. (Eine Lösung der ersten Gleichung ist die Exponentialfunktion, Lösungen der zweiten Gleichung sind Sinus und Cosinus, siehe Abschnitt 9.2.)

Satz 9.2. *Sei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{C}$ und*

$$P(z) = \sum_{n=0}^{\infty} a_n z^n$$

eine Potenzreihe. Sei $z_0 \neq 0$ und $P(z_0)$ konvergent. Dann konvergiert $P(z)$ für alle $z \in \mathbb{C}$ mit $|z| < |z_0|$.

Beweis. Weil $P(z_0) = \sum_{n=0}^{\infty} a_n z_0^n$ konvergiert, ist nach Proposition 5.4 die Folge $(a_n z_0^n)_{n\in\mathbb{N}}$ eine Nullfolge. Sie ist insbesondere beschränkt (Kor. 4.4), d.h.

$$\exists\ C \in \mathbb{R}, C > 0\ \forall\ n \in \mathbb{N} : |a_n z_0^n| < C.$$

Setze $q = |z/z_0|$. Dann gilt für beliebiges $z \in \mathbb{C}$:

$$|a_n z^n| = |a_n z_0^n| \left|\frac{z}{z_0}\right|^n \leqslant Cq^n.$$

Solange $|z| < |z_0|$ ist, gilt für den Quotienten $q < 1$. In diesem Falle ist die geometrische Reihe $C \cdot \sum_{n=0}^{\infty} q^n$ eine konvergente Majorante für $P(z)$. Aus dem Majorantenkriterium 8.12 folgt jetzt der Satz. □

Bemerkung 9.3. Der Satz 9.2 sagt, dass, wenn die Potenzreihe $P(z) = \sum_{n=0}^{\infty} a_n z^n$ an einem bestimmten Punkt $z_0 \in \mathbb{C}$ konvergiert, dann konvergiert sie schon absolut im Inneren des Kreises um den Ursprung mit Radius $|z_0|$. Man sei gewarnt, dass auf dem Rand des Kreises ohne weitere Bedingung keine Aussage möglich ist.

Definition 9.4. Sei $P(z) = \sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe. Dann heißt

$$r(P) = \sup\{|z| \in \mathbb{R}_{\geqslant 0} \mid z \in \mathbb{C}, P(z) \text{ konvergiert}\} \in \mathbb{R}_{\geqslant 0} \cup \{+\infty\}$$

der *Konvergenzradius der Potenzreihe P.*

Wenn eine Potenzreihe den Konvergenzradius $+\infty$ hat, dann bedeutet das, dass sie auf ganz $\mathbb{C}$ absolut konvergiert.

Lemma 9.5. *Jede Potenzreihe besitzt einen Konvergenzradius.*

Beweis. Für eine beliebige Potenzreihe ist die Menge $\{|z| \in \mathbb{R}_{\geqslant 0} \mid z \in \mathbb{C}, P(z) \text{ konvergiert}\}$ nicht-leer, denn sie enthält immer $z = 0$. Wenn sie nach oben beschränkt ist, folgt die Existenz aus Satz 6.19. Wenn sie nicht nach oben beschränkt ist, so ist $r(P) = +\infty$. □

Proposition 9.6. *Sei $P(z) = \sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe. Die Zahl $r \in \mathbb{R}_{\geqslant 0} \cup \{+\infty\}$ ist genau dann der Konvergenzradius von $P(z)$, wenn*

1. *$P(z)$ für alle $z \in \mathbb{C}$ mit $|z| < r$ absolut konvergiert und*
2. *$P(z)$ für alle $z \in \mathbb{C}$ mit $|z| > r$ divergiert.*

Beweis. Aus der Definition des Konvergenzradius als Supremum und Satz 9.2 folgt, dass der Konvergenzradius $r(P)$ die beiden Eigenschaften erfüllt.

Sei jetzt r eine positive reelle Zahl oder $+\infty$, die die beiden Eigenschaften erfüllt.

Annahme: $r < r(P)$. Weil nach Definition gilt

$$r(P) = \{r \in \mathbb{R}_{\geqslant 0} \mid P(r) \text{ konvergiert}\},$$

würde P auch an der Stelle $z = r + \delta \in \mathbb{R}$ mit $\delta = \frac{r+r(P)}{2}$ konvergieren im Widerspruch zu Eigenschaft (2).

Annahme: $r > r(P)$. Dann würde P nach Eigenschaft (1) auch an der Stelle $z = r - \delta \in \mathbb{R}$ mit $\delta = \frac{r+r(P)}{2}$ konvergieren im Widerspruch zur Definition des Konvergenzradius als Supremum. □

Im folgenden Satz (und nur hier!) verwenden wir die folgende Konvention:

$$\frac{1}{0} = +\infty \quad \text{und} \quad \frac{1}{\infty} = 0.$$

Satz 9.7. *Sei $P(z) = \sum_{n=0}^{\infty} a_n z^n$ eine Potenzreihe mit Koeffizienten in $\mathbb{C}$. Dann gilt:*

1. *Für $L = \limsup_{n\to\infty} \sqrt[n]{|a_n|}$ gilt $r(P) = L^{-1}$.*
2. *Wenn $q = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|, q \in \mathbb{R}_{\geqslant 0} \cup \{+\infty\}$ existiert, dann gilt $r(P) = q^{-1}$.*

Beweis. Aussage 1 folgt direkt aus dem Wurzelkriterium 8.21: es gilt

$$1 > \limsup_{n\to\infty} \sqrt[n]{|a_n z^n|} = |z| \limsup_{n\to\infty} \sqrt[n]{|a_n|} = |z| \cdot L$$

genau dann, wenn $|z| < L^{-1}$, und die Potenzreihe P konvergiert dann. Sie divergiert für $|z| > L^{-1}$ wieder nach Satz 8.21. Also gilt $r(P) = L^{-1}$ nach Proposition 9.6.

Aussage (2) folgt aus dem Quotientenkriterium 8.16.

$$1 > \lim_{n\to\infty} \left|\frac{a_{n+1} z^{n+1}}{a_n z^n}\right| = |z| \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| = |z| \cdot q$$

Wie oben folgt, dass $r(P) = q^{-1}$. □

9.2 Die Exponentialfunktion und trigonometrische Funktionen

Definition 9.8. Die *Exponentialfunktion* ist definiert durch

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

Lemma 9.9. *Der Konvergenzradius der Exponentialfunktion ist* $+\infty$.

Beweis. Wir wenden das Quotientenkriterium 8.16 und Satz 9.7 an:

$$\lim_{n\to\infty} \left| \frac{n!}{(n+1)!} \right| = \lim_{n\to\infty} \frac{1}{n+1} = 0,$$

also $r(\exp) = +\infty$. □

Proposition 9.10 (Funktionalgleichung der Exponentialfunktion)**.** *Für alle* $w, z \in \mathbb{C}$ *gilt:*

$$\exp(w) \cdot \exp(z) = \exp(w+z).$$

Beweis. Wir benutzen Satz 8.28 und bilden das Cauchy-Produkt der Reihen $\exp(w)$ und $\exp(z)$:

$$\begin{aligned}
\exp(w) \cdot \exp(z) &= \left(\sum_{n=0}^{\infty} \frac{w^n}{n!} \right) \left(\sum_{n=0}^{\infty} \frac{z^n}{n!} \right) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{w^k}{k!} \frac{z^{n-k}}{(n-k)!} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} w^k z^{n-k} = \sum_{n=0}^{\infty} \frac{1}{n!} (w+z)^n \\
&= \exp(w+z)
\end{aligned}$$

□

Definition 9.11. Wir nennen

$$e = \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!}$$

die *Eulersche Zahl.*

Lemma 9.12. *Es gilt:*

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

Beweis. Übung □

Wir haben bisher in Definition 7.11 den Ausdruck e^x nur für $x \in \mathbb{Q}$ definiert. Die Exponentialfunktion ist aber für alle $z \in \mathbb{C}$ definiert. Es bietet sich an, diese Tatsache zu benutzen, um Potenzen mit beliebigen komplexen Zahlen (nicht nur rationale Zahlen) zu definieren. Dazu müssen wir aber zuerst sicherstellen, dass für alle $x \in \mathbb{Q}$

$$e^x = \exp(x)$$

gilt. Für alle $x \in \mathbb{N}$ folgt das durch Induktion aus der Funktionalgleichung: z.B.

$$e^2 = \exp(1) \cdot \exp(1) = \exp(1+1) = \exp(2)$$

usw. Sei jetzt $p \in \mathbb{Q}, p = \frac{m}{n}$ mit $m, n \in \mathbb{Z}$. In Definition 7.11 haben wir

$$e^p = \sqrt[n]{e^m}$$

gesetzt. Mit der Funktionalgleichung folgt

$$(e^p)^n = (\exp(p))^n = \exp(np) = \exp(m) = e^m.$$

Also ist e^m eine positive (denn $e = \exp(1) > 0$) Lösung der Gleichung $x^n = e^p$, und es folgt mit Satz 7.4, dass

$$\exp(p) = e^p$$

für alle $p \in \mathbb{Q}$ gilt. Das war unser Ziel. Von jetzt ab, können wir die Terme e^x und $\exp(x)$ für alle $x \in \mathbb{C}$ synonym verwenden. Tatsächlich wird dadurch e^x erst für alle $x \in \mathbb{C} - \mathbb{Q}$ definiert. Wir werden in Definition 12.3 sogar noch weitergehen und den Ausdruck z^α für alle $\alpha \in \mathbb{C}$ und $z \in \mathbb{C}$ definieren.

Proposition 9.13. *Für alle $x, y \in \mathbb{R}$ und $z \in \mathbb{C}$ gilt:*

1. $\exp(0) = 1$ *und* $\exp(x) = e^x \in \mathbb{R}$
2. $x < y \implies \exp(x) < \exp(y)$
3. $\exp(x) = e^x > 0$ *und* $\exp(z) = e^z \neq 0$
4. $\exp(\overline{z}) = \overline{\exp(z)}$
5. $|\exp(ix)| = |e^{ix}| = 1$

Beweis. 1. Folgt direkt aus der Definition und dem vor der Proposition ausgeführten Argument.

2. Für $0 \leqslant x < y$ folgt $\exp(x) < \exp(y)$ direkt aus der Definition. Für $x < y \leqslant 0$ gilt nach dem gerade Gezeigten: $\exp(-y) < \exp(-x)$. Damit und mit der Funktionalgleichung folgt dann:

$$\exp(-y) = \exp(y)^{-1} > \exp(x)^{-1} = \exp(-x).$$

3. Aus 1. und 2. folgt $e^x > 1$ für alle $x \in \mathbb{R}, x \geqslant 0$. Weil $e^{-x} = (e^x)^{-1}$ folgt $0 < e^x < 1$ für alle $x \in \mathbb{R}, x < 0$. Desweiteren gilt für alle $z \in \mathbb{C}$: $1 = e^z \cdot e^{-z}$ nach der Funktionalgleichung. Damit kann e^z keine Nullstellen haben.

4. Folgt wieder direkt aus der Definition und der Aussage, dass für jede konvergente Folge $(z_n)_{n \in \mathbb{N}}$ in $\mathbb{C}$ gilt: $\lim_n \overline{z_n} = \overline{\lim_n z_n}$.

5. Folgt aus 4.: für alle $x \in \mathbb{R}$

$$|e^{ix}|^2 = e^{ix} e^{\overline{ix}} = e^{ix} e^{-ix} = 1.$$

Es folgt $|e^{ix}| = \pm 1$. Aber $|e^{ix}| \geqslant 0$, also $|e^{ix}| = 1$. □

Definition 9.14. Die 1-*dimensionale Einheitssphäre* oder einfach 1-Sphäre ist

$$\mathbb{S}^1 = \{z \in \mathbb{C} \mid |z| = 1\} = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}.$$

Korollar 9.15. *Die Exponentialfunktion ist ein Gruppenhomomorphismus*

$$\exp\colon (\mathbb{C}, +) \to (\mathbb{C} - \{0\}, \cdot)$$

von der additiven Gruppe und die multiplikative Gruppe der komplexen Zahlen. Genauso ist die Funktion

$$\exp(i-)\colon \mathbb{R} \to \mathbb{S}^1, x \mapsto \exp(ix) = e^{ix}$$

ein Gruppenhomomorphismus der Gruppe $(\mathbb{R}, +)$ *in die Gruppe* $(\mathbb{S}^1, \cdot)$.

Sobald wir den Logarithmus einführen, können wir zeigen, dass die Exponentialfunktion ein Gruppenisomorphisus ist. Natürlich sind aber $\mathbb{R}$ und $\mathbb{S}^1$ nicht isomorph als Gruppen.

Beweis. Dies folgt direkt aus der Funktionalgleichung 9.10 und Proposition 9.13(1). □

Definition 9.16. Für beliebiges $x \in \mathbb{R}$ setzen wir

$$\operatorname{Re}(e^{ix}) = \cos(x) \quad \text{und} \quad \operatorname{Im}(e^{ix}) = \sin(x).$$

Lemma 9.17. *Für alle* $x \in \mathbb{R}$ *gilt:*

$$e^{ix} = \cos(x) + i\sin(x).$$

Mit anderen Worten:

$$\cos(x) = \frac{1}{2}\left(e^{ix} + e^{-ix}\right) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n} = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} \pm \ldots \quad \textit{und}$$
$$\sin(x) = \frac{1}{2i}\left(e^{ix} - e^{-ix}\right) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} x^{2n+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \ldots.$$

Diese Reihen konvergieren auf ganz $\mathbb{R}$. *Weiter gilt:*

$$\cos(x)^2 + \sin^2(x) = 1.$$

Beweis. Nach Definition gilt:

$$e^{ix} = \operatorname{Re}(e^{ix}) + i \cdot \operatorname{Im}(e^{ix}) = \cos(x) + i\sin(x).$$

Nach Proposition 2.26(5) gilt:

$$\cos(x) = \operatorname{Re}(e^{ix}) = \frac{1}{2}\left(e^{ix} + e^{-ix}\right).$$

Wenn man jetzt die Potenzreihe der Exponentialfunktion einsetzt, erhält man

$$\frac{1}{2}\left(e^{ix} + e^{-ix}\right) = \frac{1}{2}\left(\sum_{n=0}^{\infty} \frac{(ix)^n + (-ix)^n}{n!}\right) = \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} x^{2n},$$

denn ungerade Potenzen heben sich gegenseitig weg. Eine analoge Rechnung gilt für den Sinus. Beide Reihen konvergieren überall, denn die Potenzreihen für e^{ix} und e^{-ix} konvergieren überall. Ferner gilt:

$$1 = |e^{ix}|^2 = \cos(x)^2 + \sin^2(x).$$

$\square$

Die Potenzreihen im letzten Lemma eignen sich dazu, Sinus und Cosinus für komplexe Zahlen zu definieren, wenn das einmal benötigt wird. Um weitere Eigenschaften der Funktionen $\exp, \cos$ und $\sin$ zu beweisen, werden wir etwas später Methoden der Differentialrechnung anwenden.

10 Stetige Funktionen

Nach unendlich langer Zeit, die wir mit Folgen und Reihen, verbracht haben, können wir uns nun dem eigentlichen Thema der reellen Analysis I zuwenden: Funktionen und ihre Eigenschaften.

10.1 Definition und Beispiele

Es folgt jetzt eine weitere zentrale Definition der Analysis.

Definition 10.1. Es sei $I \subset \mathbb{R}$. Sei $f\colon I \to \mathbb{R}$ eine Funktion. Schließlich sei $x_0 \in I$. Die Funktion f ist *stetig in* x_0, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x \in I$ folgt:

$$|x - x_0| < \delta \implies |f(x) - f(x_0)| < \varepsilon.$$

Die Funktion f heißt *stetig*, wenn sie für alle $x_0 \in I$ stetig ist. Wir schreiben $\mathcal{C}(I, \mathbb{R})$ für die Menge aller stetigen Abbildungen von I nach $\mathbb{R}$.

Bemerkung 10.2. Mit Quantoren ausgedrückt lautet die Definition:

$$\forall\ \varepsilon > 0\ \exists\ \delta > 0\ \forall\ x \in I, |x - x_0| < \delta :\ |f(x) - f(x_0)| < \varepsilon.$$

oder äquivalent

$$\forall\ \varepsilon > 0\ \exists\ \delta > 0\ \forall\ x \in\ ]x_0 - \delta, x_0 + \delta[\ :\ f(x) \in\ ]f(x_0) - \varepsilon, f(x_0) + \varepsilon[$$

oder äquivalent

$$\forall\ \varepsilon > 0\ \exists\ \delta > 0\ f\big(]x_0 - \delta, x_0 + \delta[\ \cap I\big) \subset\ ]f(x_0) - \varepsilon, f(x_0) + \varepsilon[.$$

Der Absolutbetrag ergibt auch komplexe Zahlen Sinn, deshalb können wir die erste Formulierung der Stetigkeit auch für Funktionen von $\mathbb{C}$ nach $\mathbb{C}$ verwenden.

Definition 10.3. Sei $U \subset \mathbb{C}$ und $f\colon U \to \mathbb{C}$ ein Funktion. Sei $z_0 \in U$. Die Funktion f ist *stetig in* z_0, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $z \in U$ gilt:

$$|z - z_0| < \delta \implies |f(z) - f(z_0)| < \varepsilon.$$

Mit Quantoren ausgedrückt lautet die Definition genauso wie oben:

$$\forall\, \varepsilon > 0\ \exists\, \delta > 0\ \forall\, z \in U, |z - z_0| < \delta : \ |f(z) - f(z_0)| < \varepsilon.$$

Die Funktion f heißt *stetig*, wenn sie für alle $z_0 \in U$ stetig ist. Wir schreiben $\mathcal{C}(U, \mathbb{C})$ für die Menge aller stetigen Abbildungen von U nach $\mathbb{C}$.

Im Gegensatz zu $\mathbb{R}$ haben wir in $\mathbb{C}$ keine Intervalle. Eine gute Verallgemeinerung von Intervallen ist die folgende

Definition 10.4. Für $z_0 \in \mathbb{C}$ und $\delta > 0$ heiße die Menge

$$U_\delta(z_0) = \{z \in \mathbb{C} \mid |z - z_0| < \delta\}$$

die *offene Scheibe um* z_0 *mit Radius* δ oder einfach *offene* δ*-Scheibe um* z_0.

Bemerkung 10.5. Für ein $x_0 \in \mathbb{R}$ kann man natürlich auch die Menge $U_\delta(x_0)$ betrachten. Wenn man aber exklusiv im Reellen arbeitet, ergibt es oft mehr Sinn, die Menge

$$U_\delta(x_0) \cap \mathbb{R} = \{x \in \mathbb{R} \mid |x - x_0| < \delta\} =]x_0 - \delta, x_0 + \delta[.$$

zu betrachten. Das ist dann gerade das offene δ-Intervall um x_0. Daher kann man, wie schon erwähnt, $U_\delta(z_0)$ als eine zwei-dimensionale Verallgemeinerung von offenen Intervallen verstehen.

Bemerkung 10.6. In der Situation von Definition 10.3 können wir jetzt eine weitere äquivalente Formulierung geben. Die Funktion f ist *stetig in* z_0, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für alle $x \in \mathbb{R}$ gilt:

$$f\big(U_\delta(z_0)\big) \subset U_\varepsilon\big(f(z_0)\big).$$

Bemerkung 10.7. Der Leser soll sich folgendes vor Augen führen.

1. Man beachte bei der Definition der Stetigkeit in einem Punkt x_0 (oder z_0), dass δ von x_0 und ε abhängen darf, aber nicht von x.

2. Mit den Formulierungen in Definition 10.1 und 10.3, sowie Bemerkung 10.6 haben wir die Stetigkeit nicht nur für Funktionen von Teilmengen von $I \subset \mathbb{R}$ nach $\mathbb{R}$ oder Teilmengen von $U \subset \mathbb{C}$ nach $\mathbb{C}$, sondern auch von I nach $\mathbb{C}$ oder U nach $\mathbb{R}$ erklärt.

3. Die Definition der Stetigkeit an einem Punkt x_0 besagt, dass es zu jeder (noch so kleinen) ϵ-Umgebung des Bildpunktes $f(x_0)$ eine (eventuell noch viel kleinere) δ-Umgebung des Urbildpunktes x_0 gibt, die ganz nach $U_\varepsilon(f(x_0))$ abgebildet wird:

$$f\big(U_\delta(x_0)\big) \subset U_\varepsilon\big(f(x_0)\big).$$

Beispiel 10.8. Die Funktion $\mathrm{id}\colon \mathbb{R} \to \mathbb{R}, x \mapsto x$ ist stetig. Um das zu beweisen, seien $x_0 \in \mathbb{R}$ und $\varepsilon > 0$ beliebig. Wir müssen jetzt ein $\delta > 0$ finden, so dass

$$f\big(U_\delta(x_0)\big) \subset U_\varepsilon\big(f(x_0)\big)$$

mit $f = \mathrm{id}$ gilt. Setze $\delta = \varepsilon$. Dann gilt:

$$\mathrm{id}\big(U_\delta(x_0)\big) = U_\delta(x_0) = U_\varepsilon(x_0) = U_\varepsilon\big(\mathrm{id}(x_0)\big).$$

Damit ist $f = \mathrm{id}$ überall stetig.

Beispiel 10.9. Wir definieren die Funktion $g\colon \mathbb{R} \to \mathbb{R}$ durch

$$g(x) = \begin{cases} 0 & , \text{ für } x < 0 \\ 1 & , \text{ für } x \geqslant 0 \end{cases}$$

Dann ist die Funktion in $x_0 = 0$ nicht stetig. Dazu müssen wir zeigen:

$$\exists\, \varepsilon > 0\ \forall\, \delta > 0\ \exists\, x \in \mathbb{R}, |x - x_0| < \delta \quad \text{und} \quad |f(x) - f(x_0)| \geqslant \varepsilon.$$

Setze jetzt $\varepsilon = \frac{1}{2}$ (jedes $0 < \varepsilon < 1$ läßt dasselbe Argument zu). Man betrachte das Urbild von $U_{\frac{1}{2}}(f(0)) = U_{\frac{1}{2}}(1) =]1-\varepsilon, 1+\varepsilon[$:

$$f^{-1}\left(]1-\varepsilon, 1+\varepsilon[\right) = f^{-1}\left(\left]\frac{1}{2}, \frac{3}{2}\right[\right) = [0, +\infty[.$$

Man sieht, dass 0 im Urbild enthalten ist, aber keine negative Zahl, auch wenn sie noch so nahe an Null ist. Dies widerspricht der Stetigkeit in $x_0 = 0$: für alle $\delta > 0$ gibt es $x = -\delta/2$, so dass $|x - x_0| = \delta/2 < \delta$ und $|f(x) - f(x_0)| = |0 - 1| = 1 \geqslant \frac{1}{2} = \varepsilon$.

Beispiel 10.10. Jede konstante Funktion ist stetig. Beweis: Übung!

Beispiel 10.11. Für jedes $n \in \mathbb{N}$ ist die Funktion $f\colon \mathbb{C} \to \mathbb{C}, x \mapsto x^n$ stetig. Für den Beweis sei $\varepsilon > 0$ und $z_0 \in \mathbb{C}$. Es gilt für alle $n \in \mathbb{N}$:

$$|z^n - z_0^n| = |z - z_0| \cdot \left| \sum_{k=0}^{n-1} z^k z_0^{n-k} \right|.$$

Daraus erhält man die Idee: Wähle $0 < \delta < \min\{1, \frac{\varepsilon}{n(|z_0|+1)}\}$. Dann gilt für alle $z \in U_\delta(z_0)$:

$$|z - z_0| < \delta < 1 \implies |z| \leqslant |z_0| + 1$$

und damit:

$$\left| \sum_{k=0}^{n-1} z^k z_0^{n-k} \right| \leqslant \sum_{k=0}^{n-1} |z|^k |z_0|^{n-k} \leqslant n(|z| + 1)$$

Also folgt für alle $z \in U_\delta(z_0)$:

$$|z^n - z_0^n| = |z - z_0| \cdot \left| \sum_{k=0}^{n-1} z^k z_0^{n-k} \right| < \delta \cdot n(|z| + 1) < \varepsilon.$$

Also ist f bei z_0 stetig. Weil z_0 beliebig war, ist f überall stetig.

Definition 10.12. Seien $f, g\colon U \to \mathbb{C}$ Funktionen, wobei U irgendeine Menge ist. Sei $\alpha \in \mathbb{C}$. Dann definieren wir die Funktionen

$$f + g\colon U \to \mathbb{C}, (f+g)(z) = f(z) + g(z),$$

sowie

$$f \cdot g\colon U \to \mathbb{C}, (f \cdot g)(z) = f(z)g(z)$$

und

$$\alpha \cdot f\colon U \to \mathbb{C}, (\alpha \cdot f)(z) = \alpha \cdot f(z).$$

Satz 10.13. *Seien $U \subset \mathbb{C}$, $z_0 \in U$ und $\alpha \in \mathbb{C}$. Seien $f, g\colon U \to \mathbb{C}$ Funktionen, die in z_0 stetig sind. Dann sind auch*

$$f + g, f \cdot g \quad \text{und} \quad \alpha f\colon U \to \mathbb{C}$$

in z_0 stetig.

Beweis. Stetigkeit von $f+g$: Sei $\varepsilon > 0$ gegeben. Aus der Stetigkeit von f wissen wir, dass es ein $\delta_1 > 0$ gibt, so dass für alle $z \in U$ gilt:

$$|z - z_0| < \delta_1 \implies |f(z) - f(z_0)| < \frac{\varepsilon}{2}.$$

Aus der Stetigkeit von g wissen wir, dass es ein $\delta_2 > 0$ gibt, so dass für alle $z \in U$ gilt:

$$|z - z_0| < \delta_2 \implies |g(z) - g(z_0)| < \frac{\varepsilon}{2}.$$

Dann folgt für alle $z \in U$ mit $|z - z_0| < \min\{\delta_1, \delta_2\}$ aus der Dreicksungleichung:

$$|(f+g)(z) - (f+g)(z_0)| \leqslant |f(z) - f(z_0)| + |g(z) - g(z_0)| \leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

Als ist $f + g$ an der Stelle z_0 stetig.

Stetigkeit von $f \cdot g$: Sei $\varepsilon > 0$ gegeben.

1. Fall $f(z_0) = 0$ oder $g(z_0) = 0$: Wir müssen einfach zeigen, dass es ein $\delta > 0$ gibt, so dass aus $z \in U \cap U_\delta(z_0)$ folgt

$$|(f \cdot g)(z) - (f \cdot g)(z_0)| = |f(z)| \cdot |g(z)| < \varepsilon.$$

Aus der Stetigkeit von f wissen wir, dass es ein $\delta_1 > 0$ gibt, so dass für alle $z \in U$ gilt:

$$|z - z_0| < \delta_1 \implies |f(z) - f(z_0)| < \sqrt{\varepsilon}.$$

Aus der Stetigkeit von g wissen wir, dass es ein $\delta_2 > 0$ gibt, so dass für alle $z \in U$ gilt:

$$|z - z_0| < \delta_2 \implies |g(z) - g(z_0)| < \sqrt{\varepsilon}.$$

Dann folgt für alle $z \in U$ mit $|z - z_0| < \min\{\delta_1, \delta_2\}$:

$$|f(z)| \cdot |g(z)| \leqslant \sqrt{\varepsilon} \cdot \sqrt{\varepsilon} = \varepsilon.$$

2. Fall $f(z_0) \neq 0, g(z_0) \neq 0$: Dann betrachte man folgende Gleichung:

$$\begin{aligned}(f \cdot g)(z) - (f \cdot g)(z_0) = f(z_0)\big(g(z) - g(z_0)\big) &+ g(z_0)\big(f(z) - f(z_0)\big) \\ &+ \big(f(z) - f(z_0)\big)\big(g(z) - g(z_0)\big)\end{aligned} \tag{10.14}$$

Aus der Stetigkeit von f wissen wir, dass es ein $\delta_1 > 0$ gibt, so dass für alle $z \in U$ gilt:

$$|z - z_0| < \delta_1 \implies |f(z) - f(z_0)| < \min\left\{\sqrt{\frac{\varepsilon}{3}}, \frac{\varepsilon}{3|g(z_0)|}\right\}.$$

Aus der Stetigkeit von g wissen wir, dass es ein $\delta_2 > 0$ gibt, so dass für alle $z \in U$ gilt:

$$|z - z_0| < \delta_2 \implies |g(z) - g(z_0)| < \min\left\{\sqrt{\frac{\varepsilon}{3}}, \frac{\varepsilon}{3|f(z_0)|}\right\}.$$

Dann folgt aus (10.14):

$$\begin{aligned}|(f\cdot g)(z)-(f\cdot g)(z_0)| &\leqslant |f(z_0)|\big|g(z)-g(z_0)\big|+|g(z_0)|\big|f(z)-f(z_0)\big| \\ &\quad +\big|f(z)-f(z_0)\big|\cdot\big|g(z)-g(z_0)\big| \\ &\leqslant |f(z_0)|\tfrac{\varepsilon}{3|f(z_0)|}+|g(z_0)|\tfrac{\varepsilon}{3|g(z_0)|}+\sqrt{\tfrac{\varepsilon}{3}}\cdot\sqrt{\tfrac{\varepsilon}{3}} \\ &= \tfrac{\varepsilon}{3}+\tfrac{\varepsilon}{3}+\tfrac{\varepsilon}{3}=\varepsilon\end{aligned}$$

Wir haben jetzt in allen Fällen gezeigt, dass $f\cdot g$ an der Stelle z_0 stetig ist.

Zu $\alpha\cdot f$: Dies folgt einfach aus 2., indem man für g die konstante Funktion bei $g(z)=\alpha$ nimmt. □

Definition 10.15. Sei $U\subset\mathbb{C}$ und $f\colon U\to\mathbb{C}$ eine Funktion. Wir schreiben $\mathrm{Null}(f)$ für die *Nullstellenmenge von* f, d.h.

$$\mathrm{Null}(f)=\{z\in\mathbb{C}\mid f(z)=0\}.$$

Proposition 10.16. *Sei $z_0\in U\subset\mathbb{C}$ und $f\colon U\to\mathbb{C}$ stetig bei z_0. Sei $z_0\in U$ keine Nullstelle von f, also $f(z_0)\neq 0$. Dann existiert ein $\delta>0$, so dass auch für alle $z\in U\cap U_\delta(z_0)$*

$$f(z)\neq 0 \qquad \left(\textit{genauer } 0<\frac{1}{2}|f(z_0)|\leqslant|f(z)|\right)$$

gilt.

Beweis. Wähle $\varepsilon=|f(z_0)/2|$. Weil f stetig in z_0 ist, gibt es ein $\delta>0$, so dass für alle $z\in U\cap U_\delta(z_0)$ gilt:

$$\big||f(z)|-|f(z_0)|\big|\leqslant|f(z)-f(z_0)|<\varepsilon=\left|\frac{f(z_0)}{2}\right|.$$

Dann folgt aber:

$$0<\frac{1}{2}|f(z_0)|\leqslant|f(z)|.$$

Insbesondere folgt $f(z)\neq 0$ für alle $z\in U_\delta(z_0)\cap U$. □

Definition 10.17. Eine *rationale Funktion* ist eine Funktion der Form

$$f\colon\ \mathbb{C}-\mathrm{Null}(q)\to\mathbb{C}, f(z)=\frac{p(z)}{q(z)},$$

wobei p und q Polynome mit Koeffizienten in $\mathbb{C}$.

Satz 10.18. *Sei $z_0\in U\subset\mathbb{C}$ und $f\colon U\to\mathbb{C}$ eine Funktion, so dass $f(z_0)\neq 0$ und f stetig in z_0 ist. Dann ist die Funktion*

$$g=\frac{1}{f}\colon\ U-\mathrm{Null}(f)\to\mathbb{C}, g(z)=\frac{1}{f(z)}$$

stetig bei z_0.

Beweis. Sei $\varepsilon > 0$ gegeben. Wegen der Stetigkeit von f bei z_0 gibt es nach Proposition 10.16 ein $\delta_1 > 0$ und ein $C > 0$, so dass für alle $z \in U \cap U_{\delta_1}(z_0)$

$$|f(z)| > C$$

gilt. Auch wegen der Stetigkeit von f bei z_0 gibt es ein $\delta_2 > 0$, so dass für alle $z \in U \cap U_{\delta_2}(z_0)$

$$|f(z) - f(z_0)| < \varepsilon \cdot C \cdot |f(z_0)|$$

gilt. Setze $\delta = \min\{\delta_1, \delta_2\}$. Dann gilt für alle $z \in U \cap U_\delta(z_0)$:

$$|g(z) - g(z_0)| = |f(z)^{-1} - f(z_0)^{-1}| = \left|\frac{f(z_0) - f(z)}{f(z) \cdot f(z_0)}\right| \leqslant \frac{\varepsilon \cdot C|f(z_0)|}{C|f(z_0)|} = \varepsilon.$$

Also ist g bei z_0 stetig. □

Beispiel 10.19. Wir haben folgende Beispiele für stetige Funktionen.

1. Polynome sind nach Satz 10.13 auf ganz $\mathbb{C}$ stetig.
2. Rationale Funktionen sind auf ihrem Definitionsbereich (d.h. außerhalb der Nullstellen des Nenners) stetig. Dies folgt aus den Sätzen 10.13 und 10.18.
3. Die Betragsfunktion $|-|\colon \mathbb{C} \to \mathbb{C}$ ist stetig. Dies folgt aus der Ungleichung

   $$\big||z| - |w|\big| \leqslant |z - w|$$

   für alle $z, w \in \mathbb{C}$.

Lemma 10.20. *Sei $U \subset \mathbb{C}$, $z_0 \in U$ und sei $f\colon U \to \mathbb{C}$ eine Funktion. Die Funktion f ist in z_0 stetig genau dann, wenn die Funktionen*

$$\operatorname{Re} f\colon U \to \mathbb{R}, z \mapsto \operatorname{Re}(f(z))$$

und

$$\operatorname{Im} f\colon U \to \mathbb{R}, z \mapsto \operatorname{Im}(f(z))$$

in z_0 stetig sind.

Beweis. "⇒" Wenn f stetig bei z_0 ist, dann folgt die Aussage, dass $\operatorname{Re} f$ und $\operatorname{Im} f$ auch bei z_0 stetig sind, direkt aus den Ungleichungen

$$|\operatorname{Re} f(z) - \operatorname{Re} f(z_0)| = |\operatorname{Re}\big(f(z) - f(z_0)\big)| \leqslant |f(z) - f(z_0)|$$

und

$$|\operatorname{Im} f(z) - \operatorname{Im} f(z_0)| = |\operatorname{Im}\big(f(z) - f(z_0)\big)| \leqslant |f(z) - f(z_0)|.$$

"⇐" Seien $\operatorname{Re} f$ und $\operatorname{Im} f$ bei z_0 stetig. Es gilt:

$$f = \operatorname{Re} f + i \cdot \operatorname{Im} f.$$

Dann folgt die Aussage aus Satz 10.13. □

Beispiel 10.21. Insbesondere sind die Funktionen $\operatorname{Re}, \operatorname{Im} : \mathbb{C} \to \mathbb{R}$ und komplexe Konjugation stetig.

Der folgende Satz besagt, dass die Verknüpfung von stetigen Funktionen wieder stetig ist.

Satz 10.22. *Seien $U, V \subset \mathbb{C}$ und $z_0 \in U$. Seien $f\colon U \to V$, $g\colon V \to \mathbb{C}$ Funktionen. Außerdem sei f stetig in z_0 und g stetig in $f(z_0)$. Dann ist $g \circ f\colon U \to \mathbb{C}$ stetig in z_0.*

Beweis. Sei ε gegeben. Es ist zu zeigen, dass es ein $\delta > 0$ gibt, so dass für alle $z \in U$ gilt:

$$|z - z_0| < \delta \implies |(g \circ f)(z) - (g \circ f)(z_0)| < \varepsilon.$$

Weil g stetig an der Stelle $f(z_0)$ ist, gibt es ein $\delta_1 > 0$, so dass für alle $w \in V$ gilt:

$$|w - f(z_0)| < \delta_1 \implies |g(w) - g(f(z_0))| < \varepsilon. \tag{10.23}$$

Weil f stetig an der Stelle z_0 ist, gibt es ein $\delta_2 > 0$, so dass für alle $z \in U$ gilt:

$$|z - z_0| < \delta_2 \implies |f(z) - f(z_0)| < \delta_1. \tag{10.24}$$

Wegen $f(U) \subset V$ können wir also aus beiden Bedingungen (10.23) und (10.24) schließen, dass für alle $z \in U$ gilt:

$$|z - z_0| < \delta_2 \implies |(g \circ f)(z) - (g \circ f)(z_0)| < \varepsilon.$$

Dies beweist die Stetigkeit der Verknüpfung $g \circ f$ bei z_0. □

Proposition 10.25. *Für alle $n \in \mathbb{N}$ ist die Funktion $\sqrt[n]{-}\colon \mathbb{R}_{\geqslant 0} \to \mathbb{R}_{\geqslant 0}$ stetig.*

Beweis. Übungsblatt 9.

□

Beispiel 10.26. Jede Funktion $\mathbb{R}_{\geqslant 0} \to \mathbb{R}$, $x \mapsto x^q$ für $q = \frac{m}{n} \in \mathbb{Q}$ ist stetig. Diese Funktion ist die Hintereinanderausführung der Abbildung

$$x \mapsto x^{\frac{1}{n}} = \sqrt[n]{x},$$

die nach Proposition 10.25 stetig ist, und der Funktion

$$y \mapsto y^m,$$

die nach Satz 10.13 stetig ist. Also ist sie nach Satz 10.22 stetig.

Satz 10.27. *Sei $U \subset \mathbb{C}$ und $f\colon U \to \mathbb{C}$ eine Funktion, die in $z \in U$ stetig ist. Sei $(z_n)_{n\in\mathbb{N}}$ ein Folge in U, die gegen z konvergiert. Dann konvergiert die Folge $\big(f(z_n)\big)_{n\in\mathbb{N}}$ gegen $f(z)$.*

Beweis. Wir müssen zeigen, dass es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so dass für alle $n \geqslant N$ gilt: $|f(z) - f(z_n)| < \varepsilon$.

Weil f stetig in z ist, gibt es ein $\delta > 0$, so dass für alle $w \in U \cap U_\delta(z)$ gilt:

$$|f(z) - f(z_n)| < \varepsilon.$$

Weil die Folge $(z_n)_{n\in\mathbb{N}}$ gegen z konvergiert, gibt es ein $N \in \mathbb{N}$, so dass für alle $n \geqslant N$

$$|z - z_n| < \delta$$

gilt. Alle Folgeglieder z_n liegen in U. Also können wir beide Aussagen zusammensetzen. Für alle $n \geqslant N$ gilt $|f(z) - f(z_n)| < \varepsilon$, und damit konvergiert $\big(f(z_n)\big)_{n\in\mathbb{N}}$ gegen $f(z)$. □

10.2 Berührpunkte und Häufungspunkte von Mengen, Grenzwerte von Funktionen

Definition 10.28. Sei $U \subset \mathbb{C}$. Ein $b \in \mathbb{C}$ heißt *Berührpunkt von* U, wenn für jedes $\delta > 0$

$$U \cap U_\delta(b) \neq \varnothing.$$

Anders ausgedrückt bedeutet das, dass es in jeder offenen δ-Scheibe um b mindestens einen Punkt aus U gibt.

Wir haben in Definition 6.10 den Begriff des Häufungspunktes für eine Folge definiert. Hier definieren wir ihn für eine Menge.

Definition 10.29. Sei $U \subset \mathbb{C}$. Ein $b \in \mathbb{C}$ heißt *Häufungspunkt von* U, wenn für jedes $\delta > 0$ die Menge $U \cap U_\delta(b)$ unendlich viele Elemente enthält.

Bemerkung 10.30. Man beachte:

1. Jedes $u \in U$ ist ein Berührpunkt von U, denn $u \in U$ ist in jeder offenen δ-Scheibe um u.
2. Der Punkt b ist genau dann ein Berührpunkt von U, wenn es eine Folge in U gibt, die gegen b konvergiert.
3. Jeder Häufungspunkt von U ist auch Berührpunkt von U; bei einem Häufungspunkt muss es unendlich viele Punkte aus U (und nicht nur einen) in jeder Umgebung geben. In diesem Sinne ist ein Punkt b genau dann Häufungspunkt einer Menge U, wenn es eine nicht-konstante Folge in U gibt, die gegen b konvergiert.
4. Der Punkt b ist genau dann Häufungspunkt von U, wenn er Berührpunkt der Menge $U - \{b\}$ ist.

Beispiel 10.31. Es lohnt sich, sich folgende Beispiele klar zu machen.

1. Sei $\delta > 0$ und $z \in \mathbb{C}$ Die Menge der Berührpunkte von $U_\delta(z)$ ist
$$\{w \in \mathbb{C} \,|\, |z - w| \leqslant \delta\},$$
also die "abgeschlossene δ-Scheibe" um z. Hier stimmen Berühr- und Häufungspunkte überein.
2. Sei $U = \{0\}$. Dann ist 0 ein Berührpunkt von U, aber kein Häufungspunkt.
3. Sei $U = \{\frac{1}{n} \,|\, n \in \mathbb{N}\}$. Dann ist die Menge der Berührpunkt von U gerade $U \cup \{0\}$, aber 0 ist der einzige Häufungspunkt von U.
4. Die Menge der Berührpunkte von $\mathbb{Q}$ ist $\mathbb{R}$.
5. Die Menge der Berührpunkte von $\{e^{iq} \,|\, q \in \mathbb{Q}\}$ ist $\mathbb{S}^1$.

Die nächste Definition führt eine sehr suggestive und oft hilfreiche Limes-Schreibweise ein, die aber oft von Studenten falsch verstanden und benutzt wird. Das entscheidende Wort ist fett gedruckt.

Definition 10.32. Sei $U \subset \mathbb{C}$. Sei $b \in \mathbb{C}$ ein Berührpunkt von U. Sei $f\colon U \to \mathbb{C}$ eine Funktion. Wir schreiben

$$\lim_{\substack{x \in U \\ x \to b}} f(x) = c,$$

wenn für **jede** Folge $(x_n)_{n\in\mathbb{N}}$ in U, die gegen b konvergiert, die Folge $\big(f(x_n)\big)_{n\in\mathbb{N}}$ gegen c konvergiert. Oft wird U in der Notation unterdrückt, wenn es aus dem Kontext klar ist, und man schreibt

$$\lim_{x\to b} f(x) = c$$

oder auch

$$f(x) \to c \quad \text{für} \quad x \to b.$$

Lemma 10.33. *Sei $U \subset \mathbb{C}$. Seien $b \in \mathbb{C}$ ein Berührpunkt von U und $f\colon U \to \mathbb{C}$ eine Funktion. Dann sind die folgenden Aussagen äquivlent:*

1. $\lim_{z\to b} f(z) = c$
2. *Für alle $\varepsilon > 0$ existiert ein $\delta > 0$, so dass für alle $z \in U_\delta(b) \cap U$ gilt:*

$$|f(z) - c| < \varepsilon.$$

Beweis. “(2) $\Rightarrow$ (1)” Weil b ein Berührpunkt von U ist, gibt es mindestens eine Folge in U, die gegen b konvergiert. Sei $(z_n)_{n\in\mathbb{N}}$ so eine Folge. Wir müssen zeigen, dass $\big(f(z_n)\big)_{n\in\mathbb{N}}$ gegen c konvergiert.

Sei $\varepsilon > 0$ beliebig. Nach (2) existiert dann ein $\delta > 0$, so dass für alle $z \in U_\delta(b) \cap U$ gilt:

$$|f(z) - c| < \varepsilon.$$

Weil $(z_n)_{n\in\mathbb{N}}$ gegen b konvergiert, gibt es ein $N \in \mathbb{N}$, so dass für alle $n \geqslant N$

$$|z_n - b| < \delta \quad \text{bzw.} \quad z_n \in U_\delta(b) \cap U$$

gilt. Aus (2) folgt dann:

$$|f(z_n) - c| < \varepsilon.$$

Also konvergiert $\big(f(z_n)\big)_{n\in\mathbb{N}}$ gegen c.

“(1) $\Rightarrow$ (2)” Um dies zu zeigen, können wir auch zeigen, dass aus der Verneinung von (2) die Verneinung von (1) folgt. Die Verneinung von (2) lautet:

$$\exists\, \varepsilon > 0\ \forall\, \delta > 0\ \exists\, z \in U : |z - b| < \delta \ \text{ und } \ |f(z) - c| \geqslant \varepsilon.$$

Wir nehmen dieses $\varepsilon > 0$ und setzen $\delta = \frac{1}{n}$ mit $n \in \mathbb{N}$. Also existiert nach der Verneinung von (2) ein $z_n \in U$ mit

$$|z_n - b| < \frac{1}{n} \quad \text{und} \quad |f(z_n) - c| \geqslant \varepsilon.$$

Weil $n \in \mathbb{N}$ beliebig ist, erhalten wir so eine Folge $(z_n)_{n\in\mathbb{N}}$ in U, die gegen b konvergiert, deren Bildfolge $\big(f(z_n)\big)_{n\in\mathbb{N}}$ aber nicht gegen c konvergiert. Das ist die Verneinung von (1). □

Satz 10.34. *Sei $U \subset \mathbb{C}$ und $z \in U$. Sei $f\colon U \to \mathbb{C}$ eine stetige Funktion. Dann gilt:*

$$\lim_{\substack{x \in U \\ x \to z}} f(x) = f(z).$$

Der vorige Satz ist nur eine Umformulierung des Satzes 10.27 mit unserer neuen Schreibweise. Ein Beweis besteht einfach darin, den Ausdruck

$$\lim_{\substack{x \in U \\ x \to z}} f(x)$$

korrekt nach Definition 10.32 zu interpretieren.

Bemerkung 10.35. Sei $U \subset \mathbb{C}$ und $z \in U$. Sei $f\colon U \to \mathbb{C}$ eine stetige Funktion. Sei b ein Berührpunkt von U. Man beachte, dass der Berührpunkt b nicht unbedingt eine Element von U sein muss. Falls es ein $c \in \mathbb{C}$ gibt, so dass

$$\lim_{\substack{x \in U \\ x \to b}} f(x) = c$$

gilt, kann man die Funktion f auf $U \cup \{b\}$ fortsetzen, indem man einfach $f(b) = c$ setzt. Der Satz 10.27 bzw. eine neue Form 10.34 sagt dann, dass die neue Funktion auf $U \cup \{b\}$ stetig ist. Dies gibt einem also eine Methode eine stetige Funktion von ihrem Definitionsgebiet auf die Menge der Berührpunkte des Definitionsgebietes stetig fortzusetzen.

Die folgenden Definitionen sind Variationen der Definition 10.32. Wir erlauben uns daher, dieses Mal etwas kürzer und laxer zu sein. Dieselbe Warnung wie für Definition 10.32 gilt auch hier: die Bedingung muss für jede Folge gelten.

Definition 10.36. Sei $I \subset \mathbb{R}$. Sei $b \in \mathbb{R}$ ein Berührpunkt von I. Sei $f\colon I \to \mathbb{R}$ eine Funktion.

1. Wir schreiben
$$\lim_{\substack{x \in I \\ x \to b}} f(x) = +\infty,$$
wenn für jede Folge $(x_n)_{n\in\mathbb{N}}$ in I, die gegen b konvergiert, die Folge $\big(f(x_n)\big)_{n\in\mathbb{N}}$ bestimmt gegen $+\infty$ divergiert. Man schreibt auch
$$f(x) \to +\infty \quad \text{für} \quad x \to b.$$
Eine analoge Schreibweise definieren wir für $-\infty$.

2. Wir schreiben
$$\lim_{x\to+\infty} f(x) = c \quad (\text{bzw.} \quad \lim_{x\to-\infty} f(x) = c)$$
wenn für jede Folge $(x_n)_{n\in\mathbb{N}}$, die gegen $+\infty$ $(-\infty)$ bestimmt divergiert, die Folge $\big(f(x_n)\big)_{n\in\mathbb{N}}$ bestimmt gegen c konvergiert. Man schreibt auch
$$f(x) \to c \quad \text{für} \quad x \to +\infty \ (-\infty).$$

3. Man definiert analoge Schreibweisen
$$\lim_{x\to\pm\infty} f(x) = \pm\infty$$
wenn für jede Folge $(x_n)_{n\in\mathbb{N}}$, die gegen $+\infty$ $(-\infty)$ bestimmt divergiert, die Folge $\big(f(x_n)\big)_{n\in\mathbb{N}}$ die Folge $\big(f(x_n)\big)_{n\in\mathbb{N}}$ bestimmt gegen $+\infty$ $(-\infty)$ divergiert. Man schreibt auch
$$f(x) \to +\infty\ (-\infty) \quad \text{für} \quad x \to +\infty\ (-\infty).$$

4. Wir schreiben
$$\lim_{x\to b-} f(x) = c, \pm\infty$$
wenn für jede Folge $(x_n)_{n\in\mathbb{N}}$, die erstens gegen b konvergiert und zweitens
$$\forall\ n \in \mathbb{N}\ x_n \leqslant b$$
erfüllt, die Folge $\big(f(x_n)\big)_{n\in\mathbb{N}}$ bestimmt gegen c (oder $+\infty$, $-\infty$) konvergiert.

5. Wir schreiben
$$\lim_{x\to b+} f(x) = c, \pm\infty$$
wenn für jede Folge $(x_n)_{n\in\mathbb{N}}$, die erstens gegen b konvergiert und zweitens
$$\forall\ n \in \mathbb{N}\ x_n \geqslant b$$
erfüllt, die Folge $\big(f(x_n)\big)_{n\in\mathbb{N}}$ bestimmt gegen c (oder $+\infty$, $-\infty$) konvergiert.

Diese Schreibweise soll in den folgenden Beispielen illustriert werden:

Lemma 10.37. *Sei $p(x) = \sum_{k=0}^n a_k x^k$ ein Polynom mit $a_0, \ldots, a_n \in \mathbb{R}$ und $a_n > 0$. Dann gilt*
$$\lim_{x\to+\infty} p(x) = +\infty$$
und
$$\lim_{x\to-\infty} p(x) = \begin{cases} +\infty & , \text{ für } n \text{ gerade} \\ -\infty & , \text{ für } n \text{ ungerade} \end{cases}$$
Für $a_n < 0$ dreht sich die Aussage entsprechend um.

Das Lemma läßt sich verkürzt so ausdrücken: für x gegen $+\infty$ oder $-\infty$ wird das Verhalten eines Polynoms vom Term mit der höchsten Potenz kontrolliert.

Beweis. Wir wissen, dass für jedes $n \in \mathbb{N}$ gilt: $\frac{1}{x^n} \to 0$ für $x \to +\infty$. Dies folgt z.B. aus Proposition 3.18(1). Betrachte dann
$$\frac{\sum_{k=0}^n a_k x^k}{x^n} = \sum_{k=0}^n a_k x^{k-n} = a_n + \frac{a_{n-1}}{x} + \ldots + \frac{a_0}{x^n}.$$
Dieser Ausdruck konvergiert gegen die Konstante $a_n > 0$ für $x \to \infty$. Weil aber für $x \to +\infty$ der Nenner x^n bestimmt gegen $+\infty$ divergiert, muss dann auch der Zähler $\sum_{k=0}^n a_k x^k$ gegen $+\infty$ divergieren. Die anderen Aussagen zeigt man genauso. □

Beispiel 10.38. Die Beweise folgenden Beispiele sind dem Leser überlassen.

1. Sei $a \in \mathbb{R}$. Für dei Funktion $\mathbb{R} - \{0\} \to \mathbb{R}, x \mapsto \frac{1}{x-a}$ gilt:

$$\lim_{x \to a-} \frac{1}{x-a} = -\infty \ , \quad \lim_{x \to a+} \frac{1}{x-a} = +\infty$$

2. Es gilt:

$$\mathbb{R} - \{0\} \to \mathbb{R} \ , \quad \frac{1}{x} \cdot \sin(x) \to 0 \ \text{ für } \ x \to 0.$$

   Beweis auf dem Übungsblatt 9. Also läßt sich diese Funktion stetig auf ganz $\mathbb{R}$ fortsetzen, indem man als Funktionswert bei $x = 0$ einfach 0 definiert.

Solange man sich im Klaren ist, wie diese Aussage zu interpretieren ist (nämlich via Def. 10.36), kann man das nächste Lemma auch folgendermaßen zusammenfassen: die Exponentialfunktion wächst schneller als jede Potenz.

Lemma 10.39. *Es gilt für alle $n \in \mathbb{N}$:*

1. $e^x \to +\infty$ *für* $x \to +\infty$
2. $e^x \to 0$ *für* $x \to -\infty$
3. $\frac{x^n}{e^x} = x^n e^{-x} \to 0$ *für* $x \to +\infty$
4. $\frac{e^x}{x^n} \to +\infty$ *für* $x \to +\infty$
5. $x^n e^x \to 0$ *für* $x \to -\infty$

Beweis. Zu 1. In der Exponentialreihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ strebt jeder einzelne Term gegen $+\infty$ für $x \to \infty$, also auch die ganze Summe.

Zu 2. Dies folgt jetzt aus der Funktionalgleichung

$$e^x \cdot e^{-x} = 1$$

und der Aussage von Proposition 3.18(1).

Die Aussagen 3. und 4. sind äquivalent. Wie im Beweis von Lemma 10.37 kann man den folgenden Term links betrachten und abschätzen:

$$\frac{e^x}{x^n} = \frac{1}{x^n} \sum_{k=0}^{\infty} \frac{x^k}{k!} = \sum_{k=0}^{\infty} \frac{x^{k-n}}{k!} \geqslant \frac{x}{(n+1)!}$$

Dabei läßt man alle Terme, die ja alle positiv sind, außer dem für $k = n + 1$ weg. Aber diese rechte Seite divergiert bestimmt gegen $+\infty$ für $x \to +\infty$.

Genauso wie Aussage 2 aus Aussage 1 folgt, beweist man 5 mit Hilfe von 3. Man benutzt die Funktionalgleichung und Proposition 3.18(1). $\square$

10.3 Der Zwischenwertsatz

Nachdem wir die Stetigkeit von Funktionen eingeführt haben und dann dafür gesorgt haben, dass wir viele Beispiele von stetigen Funktionen haben, müssen wir jetzt erklären, warum dieses neue (und am Anfang schwierige) Konzept interessant ist. Wiese betrachtet man stetige Funktionen?

Die erste Antwort liegt in ihrer Definition selbst: der Wert $f(x)$ einer stetigen Funktion f an einem Punkt x wird durch die Werte der Funktion in einer kleinen Umgebung Umgebung von x kontrolliert. Anders formuliert: der Wert $f(y)$ an einem Punkt y, der "nahe bei x ist", liegt selbst "nahe bei $f(x)$". Proposition 10.16 ist ein typische Aussage, die den vorigen Slogan illustriert.

Eine weitere Antwort auf die Frage, warum man stetige Funktionen betrachtet, geben die folgenden Sätze. Wir weisen darauf hin, dass diese Aussagen wesentlich auf der Vollständigkeit von $\mathbb{R}$ beruht.

Proposition 10.40 (Zwischwertsatz, Version 1)**.** *Sei $f\colon [a,b] \to \mathbb{R}$ eine stetig Funktion, so dass entweder*

1. *$f(a) < 0$ und $f(b) > 0$ oder*
2. *$f(a) > 0$ und $f(b) < 0$*

gilt. Dann existiert ein $c \in [a,b]$, so dass $f(c) = 0$.

Beweis. Wir betrachten den ersten Fall $f(a) < 0$ und $f(b) > 0$. Der andere Fall folgt aus dem ersten, indem man zu $-f$ übergeht. Wir werden induktiv eine Folge von Intervallen $(I_n = [A_n, B_n])_{n\in\mathbb{N}}$ definieren mit folgenden Eigenschaften: für alle $n \geqslant 0$

1. $I_{n+1} \subset I_n$
2. $B_n - A_n = \mathrm{diam}(I_n) \leqslant 2^{-n}\mathrm{diam}(I_0) = \frac{B_0 - A_0}{2^n}$
3. $f(A_n) \leqslant 0$ und $f(B_n) \geqslant 0$.

Wir setzen $A_0 = a$ und $B_0 = b$. Seien jetzt $I_0, \ldots, I_n$ schon definiert. Betrachte $M = \frac{B_n - A_n}{2}$. Es gibt jetzt zwei Fälle:

1. $f(M) \geqslant 0$. Dann setze $A_{n+1} = A_n$ und $B_{n+1} = M$.
2. $f(M) < 0$. Dann setze $A_{n+1} = M$ und $B_{n+1} = B_n$.

Die so konstruierte Folge $(I_n = [A_n, B_n])_{n\in\mathbb{N}}$ erfüllt dann offensichtlich die drei oben aufgeführten Eigenschaften. Die Folge $(A_n)_{n\in\mathbb{N}}$ ist monoton steigend und nach oben beschränkt, also besitzt die einen Grenzwert c. Die Folge $(B_n)_{n\in\mathbb{N}}$ ist monoton fallend und nach unten beschränkt, also besitzt die einen Grenzwert c'. Außerdem gilt für alle $n \in \mathbb{N}$: $A_n < B_n$. Aus Eigenschaft (3) oben folgt

$$0 \leqslant c' - c = \lim_{n\to\infty} B_n - \lim_{n\to\infty} A_n = \lim_{n\to\infty} (B_n - A_n) \leqslant (B_0 - A_0) \lim_{n\to\infty} 2^{-n} = 0,$$

also $c = c'$. Weil f stetig ist, folgt aus Satz 10.27:

$$0 \leqslant \lim_{n\to\infty} f(B_n) = f(c) = \lim_{n\to\infty} f(A_n) \leqslant 0.$$

Es folgt $f(c) = 0$ und der Satz ist bewiesen. □

Beispiel 10.41. Aus Lemma 10.37 folgt, dass ein Polynom von ungeradem Grad immmer eine Stelle mit negativen Funktionswert und eine andere mit positiven Funktionswert besitzt. Nach Proposition 10.40 besitzt also jedes Polynom von ungeradem Grad in $\mathbb{R}$ mindestens eine Nullstelle.

Man bezeichnet nicht nur den nächsten Satz, sondern auch Propositiion 10.40 als Zwischenwertsatz. Die beiden Aussagen sind äquivalent.

Proposition 10.42 (Zwischwertsatz, Version 2)**.** *Sei $f\colon [a,b] \to \mathbb{R}$ eine stetig Funktion und sei $f(a) \leqslant \gamma \leqslant f(b)$. Dann existiert ein $c \in [a,b]$ mit $f(c) = \gamma$.*

Beweis. Weil f stetig ist, ist auch die Funktion

$$g\colon [a,b] \to \mathbb{R}, g(x) = f(x) - \gamma$$

stetig. Nun gilt aber

$$g(a) = f(a) - \gamma \leqslant 0 \quad \text{und} \quad g(b) = f(b) - \gamma \geqslant 0.$$

Nach Proposition 10.40 gibt es dann ein $c \in [a,b]$ mit

$$0 = g(c) = f(c) - \gamma,$$

also $f(c) = \gamma$. □

Der folgende Satz kann verkürzt so ausgedrückt werden: auf einer abgeschlossenen Kreisscheibe nimmt eine stetige Funktion ihr Supremum und Infimum an.

Satz 10.43. *Eine stetige Funktion auf einem beschräkten abgeschlossenen Intervall nimmt ihr Maximum und Minimum an. D.h. mit den Bezeichnungen des vorigen Satzes existiere $c, d \in [a,b]$ mit $f(c) = \alpha$ und $f(d) = \beta$.*

Beweis. Wir zeigen die Existenz von c; die Existenz von d ist analog. Nach Voraussetzung gilt:

$$\alpha = \sup\{f(x) \mid a \leqslant x \leqslant b\}.$$

Also gibt es eine Folge $(x_n)_{n\in\mathbb{N}}$ in $[a,b]$, so dass $f(x_n)_{n\in\mathbb{N}}$ gegen α konvergiert. Nach dem Satz 6.4 von Bolzano-Weierstraß besitzt $(x_n)_{n\in\mathbb{N}}$ eine konvergente Teilfolge $(x_{n_k})_{k\in\mathbb{N}}$ mit Grenzwert c. Weil f stetig ist, folgt dann aus Satz 10.27

$$f(c) = f(\lim_{k\to\infty} x_{n_k}) = \lim_{k\to\infty} f(x_{n_k}) = \alpha.$$

Weil die Folge $(x_{n_k})_{k\in\mathbb{N}}$ in $[a,b]$ liegt, folgt dann aus dem Sandwichsatz 3.23, dass $c = \lim_{k\to\infty} x_{n_k} \in [a,b]$. □

Definition 10.44. Für $z \in \mathbb{C}$ und $\delta > 0$ heißt die Menge

$$B_\delta(z) = \overline{U_\delta(z)} = \{w \in \mathbb{C} \mid |z - w| \leqslant \delta\}$$

die *abgeschlossene Kreisscheibe um z mit Radius δ* oder einfach *die abgeschlosse δ-Scheibe um z.*

Der folgende Satz ist eine komplexe Version des Satzes 10.43. Der Beweis ist derselbe.

Satz 10.45. *Sei $z \in \mathbb{C}$ und $\delta > 0$. Sei $f\colon B_\delta(z) \to \mathbb{C}$ stetig. Dann existieren $w_1, w_2 \in B_\delta(z)$, so dass*

$$f(w_1) = \sup_{w \in B_\delta(z)} |f(w)|$$

und

$$f(w_2) = \inf_{w \in B_\delta(z)} |f(w)|.$$

Beweis. Weil f stetig, ist nach Satz 10.22 auch die Funktion

$$|-| \circ f\colon B_\delta(z) \to \mathbb{R}, w \mapsto |f(w)|$$

stetig. Für ein $\gamma \in \mathbb{R}$ mit

$$\inf_{w \in B_\delta(z)} |f(w)| \leqslant \gamma \leqslant \sup_{w \in B_\delta(z)} |f(w)|$$

existiert eine Folge $(w_n)_{n \in \mathbb{N}}$ in $B_\delta(z)$ mit

$$\lim_{n \to \infty} f(w_n) = \gamma.$$

Nach der komplexen Version des Satzes von Bolzano-Weierstraß, Korollar 6.5, gibt es eine Teilfolge $(w_{n_k})_{k \in \mathbb{N}}$ in $B_\delta(z)$ mit einem Grenzwert c.
Beh: $f(c) = \gamma$. Nach Satz 10.27 folgt

$$f(c) = f(\lim_{k \to \infty} w_{n_k}) = \lim_{k \to \infty} f(w_{n_k}) = \gamma.$$

Beh. $c \in B_\delta(z)$. Es gilt für alle $n \in \mathbb{N}$, dass $w_n \in B_\delta(z)$, also $|z - w_n||le\delta$. Dann folgt aber für alle $k \in \mathbb{N}$

$$|z - c| = |z - \lim_{k \to \infty} w_{n_k}| = \lim_{k \to \infty} |z - w_{n_k}| \leqslant \delta,$$

weil die Betragsfunktion stetig ist. □

Die Sätze 10.43 und 10.45 können verallgemeinert werden. In deren Beweis war das Hauptargument der Satz von Bolzano-Weierstraß, d.h. die Tatsache, dass jede Folge im Definitionsgebiet der Funktion eine Teilfolge besitzt, die im Definitionsgebiet konvergiert. Abgeschlossene beschränkte Intervalle in $\mathbb{R}$ und abgeschlossene beschränkte Kreisscheiben in $\mathbb{C}$ haben diese Eigenschaft. Aber es gibt noch viele anderen Teilmengen von $\mathbb{R}$ oder $\mathbb{C}$, die diese Eigenschaft haben. Diese Mengen heißen *kompakt*. Wir überlassen diesen Begriff der Vorlesung "Analysis 2".

Definition 10.46. Seien $a, b \in \mathbb{R}$. Wir schreiben oft

$$\begin{aligned}]-\infty, a] &= \{x \in \mathbb{R} \mid x \leqslant a\} & \quad]-\infty, a[&= \{x \in \mathbb{R} \mid x < a\} \\ ]a, +\infty[&= \{x \in \mathbb{R} \mid a < x\} & \quad [a, +\infty[&= \{x \in \mathbb{R} \mid a \leqslant x\}. \end{aligned}$$

Diese Mengen heißen *uneigentliche Intervalle*. Die bisher üblichen Intervalle von Definition 2.17 heißen auch *eigentliche Intervalle*.

Korollar 10.47. *Sei $I \subset \mathbb{R}$ ein eigentlichen oder uneigentlichen Intervall und sei $f\colon I \to \mathbb{R}$ eine stetige Funktion. Dann ist das Bild $f(I)$ wieder ein Intervall.*

Beweis. Setze

$$A = \inf f(I) \in \mathbb{R} \cup \{-\infty\} \quad \text{und} \quad B = \sup f(I) \in \mathbb{R} \cup \{+\infty\}.$$

Sei $A < \gamma < B$. Dann gibt es nach der Definition von A und B Zahlen $a, b \in I$ mit

$$f(a) \leqslant \gamma \leqslant f(b).$$

Nach dem Zwischenwertsatz 10.42 gibt es also ein $c \in [a, b]$ mit $f(c) = \gamma$. (Man kann nicht Satz 10.43 anwenden, denn I ist nicht unbedingt ein beschränktes Intervall.) Es folgt $\gamma \in f(I)$. Weil γ beliebig war, folgt $]A, B[\subset f(I)$. Es folgt, dass $f(I)$ selbst $[A, B],]A, B[, [A, B[$ oder $]A, B]$ ist. □

10.4 Gleichmäßige Konvergenz und die Supremumsnorm

Wir möchten gerne zeigen, dass nicht nur Polynome, sondern auch Potenzreihen (innerhalb ihres Konvergenzradius) stetig sind. Weil Potenzreihen Grenzwerte sind (für jedes z ist der Wert $P(z)$ der Grenzwert $\sum_n a_n z^n$), müssen wir uns über Folgen von Funktionen und der Grenzwerte, selber wieder Funktionen, Gedanken machen.

Definition 10.48. Sei $U \subset \mathbb{C}$ und sei $(f_n)_{n\in\mathbb{N}}$ ein Folge von Funktionen $f_n\colon U \to \mathbb{C}$. Wir nehmen ferner an, dass für jedes $z \in U$ die Folge $(f_n(z))_{n\in\mathbb{N}}$ in $\mathbb{C}$ konvergiert. Wir definieren dann die *Grenzfunktion f der Folge $(f_n)_{n\in\mathbb{N}}$* durch

$$f(z) = \lim_{n\to\infty} f_n(z).$$

Die so definierte Funktion ist eine Abbildung $f\colon U \to \mathbb{C}$. Wir sagen dann, dass die Folge $(f_n)_{n\in\mathbb{N}}$ *punktweise* gegen f konvergiert. Dabei heißt f der *punktweise Limes*, *punktweise Grenzwert* oder die *punktweise Grenzfunktion* der Fogle $(f_n)_{n\in\mathbb{N}}$.

Punktweise Konvergenz ist der erste und sehr wichtige Konvergenzbegriff für Funktionenfolgen. Leider ist er auch sehr naiv und hat schlechte Eigenschaften. So interagiert punktweise Konvergenz nicht mit Stetigkeit: *sei $(f_n)_{n\in\mathbb{N}}$ eine Folge stetiger Funktionen, die punktweise gegen f konvergiere, dann muss die punktweise Grenzfunktion f nicht unbedingt stetig sein.* Wir brauchen einen stärkeren Konvergenzbegriff, der garantiert, dass der Limes einer Folge stetiger Funktionen wieder stetig ist. Schließlich wollen wir am Schluss unter anderem die Stetigkeit von Potenzreihen allgemein beweisen.

Definition 10.49. Sei $U \subset \mathbb{C}$ und $f\colon U \to \mathbb{C}$ eine Abbildung. Die *Supremumsnorm* von f ist gegeben durch

$$\|f\|_\infty^U = \|f\|_\infty = \sup\{|f(z)| \mid z \in U\} \in \mathbb{R}_{\geqslant 0} \cup \{+\infty\}.$$

Dabei unterdrückt man fast immer das U von der Notation. Für die Menge aller Funktionen $f\colon U \to \mathbb{C}$ mit $\|f\|_\infty^U < +\infty$ (d.h. $\|f\|_\infty^U \in \mathbb{R}$) schreiben wir $L^\infty(U)$.

Lemma 10.50. *Die Supremumsnorm erfüllt auf $L^\infty(U)$ die Eigenschaften einer Norm, d.h. für alle $f, g \in L^\infty(U)$ und $\lambda \in \mathbb{C}$ gilt*

1. $\|f\|_\infty \geqslant 0$
2. $\|f\|_\infty = 0 \iff f = 0$
3. $\|\lambda \cdot f\|_\infty = |\lambda| \cdot \|f\|_\infty$
4. $\|f + g\|_\infty \leqslant \|f\|_\infty + \|g\|_\infty$

Beweis. Einzig Aussage (4) bedarf eines Beweises. Für jedes $z \in U$ gilt

$$|f(z) + g(z)| \leqslant |f(z)| + |g(z)|.$$

nach der üblichen Dreiecksungleichung in $\mathbb{C}$. Wir bilden jetzt das Supremum dieser Terme über $z \in U$:

$$\begin{aligned}\|f + g\|_\infty &= \sup\{|f(z) + g(z)| \mid z \in U\} \leqslant \sup\{|f(z)| + |g(z)| \mid z \in U\} \\ &\leqslant \sup\{|f(z)| \mid z \in U\} + \sup\{|g(z)| \mid z \in U\} = \|f\|_\infty + \|g\|_\infty.\end{aligned}$$

□

Definition 10.51. Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n\colon U \to \mathbb{C}$ für $U \subset \mathbb{C}$. Sei $f\colon U \to \mathbb{C}$ eine weitere Funktion. Wir sagen, dass die Folge $(f_n)_{n\in\mathbb{N}}$ *gleichmäßig gegen f* konvergiert, wenn gilt:

$$\lim_{n\to\infty} \|f - f_n\|_\infty = 0.$$

Mit anderen Worten: wenn für alle $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so dass für alle $n \geqslant N$ gilt:

$$\|f - f_n\|_\infty < \varepsilon.$$

Lemma 10.52. *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n\colon U \to \mathbb{C}$ für $U \subset \mathbb{C}$. Sei $f\colon U \to \mathbb{C}$ eine Funktion. Die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert genau dann gleichmäßig gegen f, wenn es für alle $\varepsilon > 0$ und alle $z \in U$ ein $N \in \mathbb{N}$ gibt, so dass*

$$|f(z) - f_n(z)| < \varepsilon.$$

Insbesondere konvergiert $(f_n)_{n\in\mathbb{N}}$ auch punktweise gegen f.

Beweis. "$\Rightarrow$" Die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiere gleichmäßig gegen f, d.h.

$$\forall\ \varepsilon > 0\ \exists\ N \in \mathbb{N}\ \forall\ n \geqslant N : \|f - f_n\|_\infty < \varepsilon.$$

Nach Definition der Supremumsnorm gilt für alle $z \in U$:

$$\varepsilon > \|f - f_n\|_\infty = \sup_{z\in U} |f(z) - f_n(z)| \geqslant |f(z) - f_n(z)|.$$

Erstens konvergiert damit $(f_n)_{n\in\mathbb{N}}$ punktweise und zweitens können wir schließen:

$$\forall\ \varepsilon > 0\ \exists\ N \in \mathbb{N}\ \forall\ n \geqslant N\ \forall\ z \in U : |f(z) - f_n(z)| < \varepsilon.$$

(Man beachte, dass man $\forall$-Quantoren untereinander vertauschen kann. Dies gilt natürlich auch für $\exists$-Quantoren. Man darf nur nicht ein $\exists$ mit einem $\forall$ vertauschen!) Dies ist äquivalent zu:

$$\forall\ \varepsilon > 0\ \exists\ N \in \mathbb{N}\ \forall\ z \in U\ \forall\ n \geqslant N : |f(z) - f_n(z)| < \varepsilon.$$

"$\Leftarrow$" Dieses Mal starten wir mit der Voraussetzung

$$\forall\ \varepsilon > 0\ \exists\ N \in \mathbb{N}\ \forall\ z \in U\ \forall\ n \geqslant N : |f(z) - f_n(z)| < \varepsilon$$

bzw. mit der äquivalenten Form

$$\forall\ \varepsilon > 0\ \exists\ N \in \mathbb{N}\ \forall\ n \geqslant N\ \forall\ z \in U : |f(z) - f_n(z)| < \varepsilon.$$

Aber die Aussage $\forall\ z \in U : |f(z) - f_n(z)| < \varepsilon$ impliziert sofort

$$\|f - f_n\|_\infty = \sup_{z \in U} |f(z) - f_n(z)| < \varepsilon.$$

Damit sind wir wieder bei Aussage 1 angekommen. □

Bemerkung 10.53. Die punktweise Konvergenz einer Funktionenfolge $(f_n\colon U \to \mathbb{C})_{n\in\mathbb{N}}$ gegen eine Funktion f kann folgendermaßen ausgedrückt werden:

$$\forall\ \varepsilon > 0\ \forall\ z \in U\ \exists\ N \in \mathbb{N}\ \forall\ n \geqslant N : |f(z) - f_n(z)| < \varepsilon.$$

Gleichmäßige Konvergenz der Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ mit $f_n\colon U \to \mathbb{C}$ gegen eine Funktion f ist äquivalent zu:

$$\forall\ \varepsilon > 0\ \exists\ N \in \mathbb{N}\ \forall\ z \in U\ \forall\ n \geqslant N : |f(z) - f_n(z)| < \varepsilon.$$

Man beachte die Reihenfolge der Quantoren: bei punktweiser Konvergenz kann N von z abhängen ("die Konvergenz ist punktweise"), während bei gleichmäßiger Konvergenz das N nicht vom Punkt z abhängen darf (in diesem Sinne ist die Konvergenz gleichmäßig). Man sieht auch auf diese Weise, dass gleichmäßige Konvergenz eine stärkere Bedingung an die Funktionenfolge ist als punktweise Konvergenz.

Bemerkung 10.54. Aus Lemma 10.52 folgt insbesondere, dass der gleichmäßige Limes einer Funktionenfolge $(f_n)_{n\in\mathbb{N}}$, wenn er existiert, immer gleich dem punktweisen Limes ist.

Satz 10.55. *Sei $(f_n)_{n\in\mathbb{N}}$ eine Folge stetiger Funktionen $f_n\colon U \to \mathbb{C}$ für $U \subset \mathbb{C}$, die gleichmäßig gegen $f\colon U \to \mathbb{C}$ konvergiert. Dann ist f stetig.*

Beweis. Sei $z_0 \in U$ beliebig. Wir müssen zeigen, dass f in z_0 stetig ist. Sei $\varepsilon > 0$ gegeben. Weil $(f_n)_{n\in\mathbb{N}}$ gleichmäßig gegen f konvergiert, gibt es ein $N \in \mathbb{N}$, so dass für alle $n \geqslant N$ und alle $z \in U$ gilt:

$$|f(z) - f_n(z)| < \frac{\varepsilon}{3}.$$

Weil f_N stetig ist, gibt es ein $\delta > 0$, so dass für alle $z \in U \cap U_\delta(z_0)$ gilt:

$$|f_N(z) - f_N(z_0)| < \frac{\varepsilon}{3}.$$

Insgesamt gilt also für alle $z \in U \cap U_\delta(z_0)$:

$$|f(z) - f(z_0)| \leqslant |f(z) - f_N(z)| + |f_N(z) - f_N(z_0)| + |f_N(z_0) - f(z_0)| \leqslant \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Also ist f stetig bei z_0. Weil z_0 beliebig war, ist f stetig. □

Proposition 10.56. *Sei $U \subset \mathbb{C}$ und $(f_n\colon U \to \mathbb{C})_{n\in\mathbb{N}}$ eine Folge von Funktionen. Die folgenden Aussagen sind äquivalent:*

1. *Die Folge $(f_n\colon U \to \mathbb{C})_{n\in\mathbb{N}}$ konvergiert gleichmäßig.*
2. *Für alle $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass für alle $m, n \geqslant N$ gilt:*
$$\|f_m - f_n\|_\infty < \varepsilon.$$
3. *Für alle $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass für alle $m, n \geqslant N$ und alle $z \in U$ gilt:*
$$|f_m(z) - f_n(z)| < \varepsilon.$$

Diese Proposition sagt, dass es reicht, eine Cauchy-Folge bezüglich der Supremumsnorm $\| - \|_\infty$ zu sein, um gleichmäßige Konvergenz zu garantieren. Die entsprechende Aussage für Folgen ist das Vollständigkeitsaxiom 4.6, die entsprechende Aussage für Reihen ist das Cauchy-Kriterium 5.3.

Beweis. "(1) $\Rightarrow$ (2)" Sei f die Grenzfunktion der gleichmäßig konvergenten Folge $(f_n)_{n\in\mathbb{N}}$. Sei $\varepsilon > 0$. Dann gibt es ein $N \in \mathbb{N}$, so dass für alle $n \geqslant N$ gilt:
$$\|f - f_n\|_\infty < \varepsilon/2.$$
Dann folgt für alle $m, n \geqslant N$ nach Lemma 10.50(4):
$$\|f_m - f_n\|_\infty \leqslant \|f_m - f\|_\infty + \|f - f_n\|_\infty < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$
"(2) $\Rightarrow$ (3)" Es gilt für alle $z \in U$:
$$|f_m(z) - f_n(z)| \leqslant \sup_{z\in U} |f_m(z) - f_n(z)| = \|f_m - f_n\|_\infty < \varepsilon$$
Das zeigt Aussage (3).

"(3) $\Rightarrow$ (1)" Die Bedingung (3) sagt insbesondere, dass für jedes $z \in U$ die Folge $(f_n(z))_{n\in\mathbb{N}}$ eine Cauchy-Folge ist. Nach Satz 4.15 ist $\mathbb{C}$ vollständig. Deshalb hat die Folge $(f_n(z))_{n\in\mathbb{N}}$ für jedes $z \in U$ einen Grenzwert. Sei $f\colon U \to \mathbb{C}$ die dadurch definierte punktweise Grenzfunktion.

Es bleibt zu zeigen, dass die Konvergenz gleichmäßig ist. Sei $\varepsilon > 0$ und benutze (3), um ein N zu wählen, so dass für alle $z \in U$ und $m, n \geqslant N$ gilt:
$$|f_m(z) - f_n(z)| < \varepsilon/2.$$
Wir halten jetzt n fest und bilden den Grenzwert für $m \to \infty$. Dann gilt für alle $z \in U$ und $n \geqslant N$:
$$|f(z) - f_n(z)| \leqslant \varepsilon/2 < \varepsilon.$$
□

Proposition 10.57. *Sei $(g_n\colon U \to \mathbb{C})_{n\in\mathbb{N}}$ eine Folge. Es gebe für jedes $n \in \mathbb{N}_0$ ein $M_n \in \mathbb{R}$, so dass für alle $z \in U$*
$$|g_n(z)| \leqslant M_n$$
gelte. Wenn $\sum_{n=0}^\infty M_n$ konvergiert, dann konvergiert $\sum_{n=0}^\infty g_n$ gleichmäßig auf U.

Beweis. Setze $f_n = \sum_{i=0}^n g_i \colon U \to \mathbb{C}$. Wenn $\sum_{n=0}^\infty M_n$ konvergiert, dann gibt es nach dem Cauchykriterium für Reihen 5.3 zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$, so dass für alle $n \geqslant m \geqslant N$ gilt:

$$|f_n(z) - f_{m-1}(z)| = \left|\sum_{i=m}^n g_i(z)\right| \leqslant \sum_{i=m}^n M_n < \varepsilon.$$

Die gleichmäßige Konvergenz die Folge $(f_n)_{n\in\mathbb{N}}$ folgt jetzt aus Proposition 10.56. Indem wir $\sum_{i=0}^n g_i = f_n$ zurücksubstituieren, folgt die Proposition. □

Korollar 10.58. *Sei $P(z) = \sum_{k=0}^\infty c_k z^k$ eine Potenzreihe in $\mathbb{C}[[z]]$ mit Konvergenzradius $r(P) > 0$ ($r(P) = +\infty$ eingeschlossen). Für alle $n \in \mathbb{N}$ sei $f_n \colon \mathbb{C} \to \mathbb{C}$ das Polynom gegeben durch*

$$p_n(z) = \sum_{k=0}^n c_k z^k.$$

Schließlich sei $0 < \varepsilon < r(P)$ beliebig. Dann gilt:

1. *Die Folge $(p_n)_{n\in\mathbb{N}}$ konvergiert für $|z| \leqslant r(P) - \varepsilon$ gleichmäßig gegen P.*

2. *Die Potenzreihe P ist auf $U_{r(P)}(0)$ eine stetige Funktion.*

3. *Der Konvergenzradius der Potenzreihe*

$$Q(z) = \sum_{k=1}^\infty k c_{k-1} z^{k-1}$$

ist $r(P)$. Diese Funktion ist auch stetig auf $U_{r(P)}(0)$.

Wir werden später in Korollar 16.3 zeigen, dass die Reihe Q innerhalb des Konvergenzradius $r(Q) = r(P)$ die Ableitung von P ist, d.h. Potenzreihe dürfen gliedweise abgeleitet werden.

Beweis. Teil 1: Zunächst ist klar, dass die Folge $(p_n)_{n\in\mathbb{N}}$ innerhalb des Konvergenzradius von P punktweise gegen P konvergiert. Wir müssen zeigen, dass die Konvergenz gleichmäßig ist. Nach Voraussetzung ist $0 < r(P) - \varepsilon < r(P)$. Dann ist P an der Stelle $z = r(P) - \varepsilon \in \mathbb{R}$ absolut konvergent und es gilt für alle $|z| < r(P) - \varepsilon$:

$$|c_n z^n| \leqslant |c_n| \big(r(P) - \varepsilon\big)^n = M_n.$$

Nach Definition von $r(P)$ konvergiert die Reihe

$$P\big(r(P) - \varepsilon\big) = \sum_{n=0}^\infty c_n \big(r(P) - \varepsilon\big)^n = \sum_{n=0}^\infty M_n$$

absolut. Propositon 10.57 liefert dann die gleichmäßige Konvergenz der Folge $(p_n)_{n\in\mathbb{N}}$.

Für alle $n \in \mathbb{N}$ sind die Funktionen p_n stetig, denn es sind ja Polynome. Weil die Folge $(p_n)_{n\in\mathbb{N}}$ auf $U_{r(P)-\varepsilon}(0)$ gleichmäßig konvergiert, folgt aus Satz 10.55, dass dann auch die Grenzfunktion P auf $U_{r(P)-\varepsilon}(0)$ stetig ist. Aber $\varepsilon > 0$ ist beliebig. Also ist P auf ganz $U_{r(P)}(0)$ stetig. Das ist Teil 2.

Für Teil 3 zeigen wir zuerst, dass $r(P) = r(Q)$. Wir wissen schon:

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

Nach dem Wurzelkriterium folgt dann:

$$r(Q)^{-1} = \limsup_{n\to\infty} \sqrt[n]{nc_{n-1}} = \left(\lim_{n\to\infty} \sqrt[n]{n}\right)\left(\limsup_{n\to\infty} \sqrt[n]{c_n}\right) = r(P)^{-1},$$

also $r(P) = r(Q)$. Jetzt können wir die Teile 1 und 2 statt auf P auf Q anwenden und es folgt die Stetigkeit von Q auf $U_{r(P)}(0)$. □

Beispiel 10.59. Aus Korollar 10.61 folgt, dass die Exponentialfunktion, Sinus und Cosinus auf ganz $\mathbb{C}$ stetig sind, denn sie haben Konvegenzradius $+\infty$.

Nachtrag: Potenzreihen bei $a \in \mathbb{C}, a \neq 0$

Bisher haben wir Potenzreihen der Form

$$P(z) = \sum_{n=0}^{\infty} c_n z^n$$

mit $c_n \in \mathbb{C}$ betrachtet. Dies ist eine Potenzreihe im Entwicklungspunkt $a = 0$. Allgemeiner betrachet man folgende Funktionen.

Definition 10.60. Sei $a \in \mathbb{C}$. Die Reihe

$$P(z) = \sum_{n=0}^{\infty} c_n (z-a)^n$$

mit $c_0, c_1, \ldots \in \mathbb{C}$ heißt *Potenzreihe in z mit Entwicklungspunkt* a oder einfach *Potenzreihe um a*.

Wieder betrachten wir eine solche Potenzreihe als Funktion in z. Für Potenzreihe um $a = 0$ haben wir in 9.2 und 9.6 gezeigt, dass sie einen Konvergenzradius r zwischen 0 und $+\infty$ besitzen und im Inneren des Kreises um 0 mit Radius r absolut konvergieren. Außerdem haben wir gerade gezeigt 10.61, dass innerhalb der Konvergenzscheibe die Partialsummen $\sum_{k=0}^{n} c_k(z-a)^k$ gleichmäßig konvergieren und dass deshalb die Potenzreihe (als Funktion in z) dort stetig ist. Wenn man in sämtlichen Beweisen den Term $|z|$, der den Abstand zum Entwicklunspunkt $a = 0$ mißt, durch den Term $|z-a|$, also wieder den Abstand zum Entwicklungspunkt, ersetzt, so bleiben alle Argumente und Aussagen gültig.

Korollar 10.61. *Sei $a \in \mathbb{C}$ und $P(z) = \sum_{k=0}^{\infty} c_k(z-a)^k$ eine Potenzreihe um a mit Konvergenzradius $r(P) > 0$ ($r(P) = +\infty$ eingeschlossen).*

1. *Die Potenzreihe P ist auf $U_{r(P)}(a)$ eine stetige Funktion.*
2. *Der Konvergenzradius der Potenzreihe*

$$Q(z) = \sum_{k=1}^{\infty} kc_{k-1}(z-a)^{k-1}$$

ist $r(P)$. Diese Funktion ist auch stetig auf $U_{r(P)}(a)$.

Der Konvergenzradius berechnet sich wie vorher durch Satz 9.7. Man beachte, dass die Konvergenzscheiben jetzt den Mittelpunkt a haben.

10.5 Monotone Funktionen und Stetigkeit von Umkehrfunktionen

Definition 10.62. Sei $I \subset \mathbb{R}$ und $f\colon I \to \mathbb{R}$ eine Funktion. Wir nennen f

1. *steigend* oder *monoton steigend*, wenn aus $x, y \in I, x < y$ die Ungleichung $f(x) \leqslant f(y)$ folgt.
2. *streng steigend* oder *streng monoton steigend*, wenn aus $x, y \in I, x < y$ die Ungleichung $f(x) < f(y)$ folgt.
3. *fallend* oder *monoton fallend*, wenn aus $x, y \in I, x < y$ die Ungleichung $f(x) \geqslant f(y)$ folgt.
4. *streng fallend* oder *streng monoton fallend*, wenn aus $x, y \in I, x < y$ die Ungleichung $f(x) > f(y)$ folgt.

Beispiel 10.63. Die Funktionen $x \mapsto x^n$ für alle $n \in \mathbb{N}$ ist auf $\mathbb{R}_{\geqslant 0}$ streng monoton steigend. Die Exponentialfunktion ist auf ganz $\mathbb{R}$ streng monoton steigend.

Bemerkung 10.64. Streng montone Funktionen sind injektiv.

Sei $U \subset \mathbb{C}$ und $f\colon U \to \mathbb{C}$ injektiv. Sei $V = f(U) \subset \mathbb{C}$ das Bild von f. Wir begehen jetzt einen kleinen Fehler und bezeichnen die (neue) Abbildung $U \to V, z \mapsto f(z)$ wieder mit f. Auf diese Weise haben wir f surjektiv gemacht; f ist jetzt also bijektiv und besitzt eine Umkehrfunktion $g\colon V \to U$, siehe Definition 1.18.

Warnung: die Umkehrfunktion von f wird oft mit f^{-1} bezeichnet. Verwechseln Sie dies nicht mit der Funktion

$$\frac{1}{f}\colon U - \mathrm{Null}(f) \to \mathbb{C}, z \mapsto f(z)^{-1} = \frac{1}{f(z)}.$$

Satz 10.65. *Sei $I \subset \mathbb{R}$ ein (un-)eigentliches Intervall und $f\colon I \to \mathbb{R}$ eine streng monotone und stetige Funktion. Sei $D = f(I)$ das Bild von f und sei $g\colon D \to I$ die Umkehrfunktion von f. Dann ist g selbst streng monoton und stetig.*

Beweis. Weil f streng monoton ist, ist f injektiv. Also bildet f das Intervall I bijektiv auf D ab. Insbesondere existiert die Umkehrfunktion $g\colon D \to I$ und g ist bijektiv.

Behauptung: g ist streng monoton. Genauer ist g streng monoton wachsend, wenn f streng monoton wachsend ist, und g ist streng monoton fallend, wenn f streng monoton fallend ist.

Wir nehmen an f ist streng monoton wachsend. Seien $y_1, y_2 \in D, y_1 < y_2$. Weil f bijektiv ist, gibt es $x_1, x_2 \in I$ mit

$$f(x_1) = y_1 \quad \text{und} \quad f(x_2) = y_2.$$

Das bedeutet

$$g(y_1) = x_1 \quad \text{und} \quad g(y_2) = x_2.$$

Nach Voraussetzung ist

$$f(x_1) = y_1 < y_2 = f(x_2).$$

Weil f strong monoton wachsend ist, muss $x_1 < x_2$ gelten. Also folgt

$$g(y_1) = x_1 < x_2 = g(x_2),$$

und g ist streng monoton wachsend.

Der Beweis, dass g streng monoton fallend ist, wenn f streng monoton fallend, folgt aus dem gerade Bewiesenen durch Übergang zu $-f$.

Behauptung: g ist stetig.

Wir nehmen wieder zuerst an, dass f und g streng monoton wachsend sind. Der andere Fall folgt durch Übergang von f zu $-f$. Sei $\varepsilon > 0$ und sei $b \in D$. Sei $g(b) = a \in I$ bzw. $f(a) = b$. Wir werden die Stetigkeit von g in b zeigen.

1. Fall: a ist kein Randpunkt des Intervalls I. Es reicht die Definition der Stetigkeit für genügend kleine $\varepsilon > 0$ nachzuweisen. Also können wir voraussetzen, dass ε so klein ist, dass das ganze Intervall $[a - \varepsilon, a + \varepsilon]$ in I liegt, denn a ist kein Randpunkt. Setze

$$f(a - \varepsilon) = b_1 \quad \text{und} \quad f(a + \varepsilon) = b_2.$$

Korollar 10.47 sagt, dass D wieder ein Intervall ist. Dann gilt also : $b_1 < b < b_2$ und f bildet das Intervall $[a-\varepsilon, a+\varepsilon]$ bijektiv auf $[b_1, b_2] \subset D$ ab. Wähle $\delta = \min\{b-b_1, b_2-b\}$. Dann gilt:

$$g\big(\,]b - \delta, b + \delta[\,\big) \subset\,]a - \varepsilon, a + \varepsilon[.$$

Also ist g in b stetig.

2. Fall: a ist unterer Randpunkt des Intervall I. (Wenn I kein Minimum besitzt, so tritt dieser Fall nicht ein.) Dann kann man wie oben $\varepsilon > 0$ so klein machen, dass

$$[a, a + \varepsilon] \subset I.$$

Setzte dann $b_1 = b$ und $f(a + \varepsilon) = b_2$ und fahre fort wie oben.

2. Fall: b ist oberer Randpunkt des Intervall I. (Wenn I kein Maximum besitzt, so tritt dieser Fall nicht ein.) Dann kann man wie oben $\varepsilon > 0$ so klein machen, dass

$$[b - \varepsilon, b] \subset I.$$

Setzte dann $b_2 = b$ und $f(a - \varepsilon) = b_1$ und fahre fort wie oben. □

11 Differenzierbare Funktionen und ihre Ableitung

11.1 Differenzierbare Funktionen

Zu jeweils zwei verschiedenen Punkte $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$ gibt es genau eine Gerade G, die durch diese Punkte geht. Wir nehmen an, dass G nicht parallel zur y-Achse ist, d.h. $x_1 \neq x_2$. Sei $x_1 < x_2$. Dann ist die Steigung dieser Geraden $m = \frac{y_2 - y_1}{x_2 - x_1}$ ("Anstieg durch Fortschritt"). Dieder Term heißt auch oft *Differenzenquotient*. Wir können dann für Punkte $(x, y) \in G$ die Gleichung

$$y = m(x - x_1) + y_1 = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1) + y_1$$

schreiben.

Sei $f\colon\ ]a,b[\to \mathbb{R}$ eine Funktion. Wir wollen an einem Punkt $x_0 \in]a,b[$ versuchen, die Funktion durch ihre Tangente im Punkt x_0 zu approximieren. Dazu müssen wir uns überlegen, wie man dise Tangente zu definieren hat und unter welchen Bedingungen diese Definition Sinn ergibt.

Unsere Idee ist, die Tangente ("die Berührende") durch Sekanten ("die Schneidende") zu approxiemieren. Eine Sekante, die den Graphen von f an den Stellen x_0 und x schneidet, hat die Gleichung

$$S(z) = \frac{f(x) - f(x_0)}{x - x_0}(z - x_0) + f(x_0).$$

Wenn wir die Stelle x immer näher bei x_0 wählen, so wird sich die Sekante der Tangente, sofern sie existiert, annähern. Daher:

Definition 11.1. Sei $a < x_0 < b$ und $f\colon\ ]a,b[\to \mathbb{R}$ eine Funktion. Die Funktion f ist *differenzierbar in* x_0, wenn der Grenzwert

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) = \frac{df}{dx}(x_0)$$

existiert. Dieser Grenzwert heißt dann der *Differentialquotient* oder *Ableitung von* f *an der Stelle* x_0. Wenn f für alle $x_0 \in]a,b[$ in x_0 differenzierbar ist, so heißt f *differenzierbar* und die so definierte Funktion

$$f' = \frac{df}{dx}\colon\ ]a,b[\to \mathbb{R}$$

heißt die *Ableitung von* f. Die *Tangente von* f *an der Stelle* x_0 ist dann gegeben durch

$$T(x) = f'(x_0)(x - x_0) + f(x_0).$$

Für die Menge der differenzierbaren Funktionen schreiben wir $\mathcal{D}(]a,b[, \mathbb{R})$.

Es gibt viele äquivalente Umformulierungen der Differenzierbarkeit. Man kann den Abstand $x - x_0$ durch h ersetzen, also $x = x_0 + h$. Der Differntialquotient schreibt sich dann als

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = f'(x_0).$$

Weitere äquivalente Formulierung findet man, wenn man den Standpunkt, dass die Tangente die Funktion approximieren soll, Ernst nimmt.

Lemma 11.2. *Sei $a < x_0 < b$ und $f\colon\ ]a,b[\to \mathbb{R}$ eine Funktion. Dann sind die folgenden Aussagen äquivalent:*

1. *Die Funktion f ist in x_0 differenzierbar.*
2. *Es gibt genau eine Zahl $f'(x_0) \in \mathbb{R}$, so dass für alle $\varepsilon > 0$ ein $\delta > 0$ existiert mit*

$$\forall\, x \in]a,b[\ : |x - x_0| < \delta \ \Rightarrow\ \left| \frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right| < \varepsilon.$$

3. *Es gibt genau eine Zahl $f'(x_0) \in \mathbb{R}$, so dass für alle $\varepsilon > 0$ ein $\delta > 0$ existiert mit*

$$\forall\, x \in]a,b[\ : |x - x_0| < \delta \ \Rightarrow\ \left| f(x) - f'(x_0)(x - x_0) - f(x_0) \right| < \varepsilon \cdot |x - x_0|.$$

Beweis. "(1) ⇔ (2)" Man wendet Lemma 10.33 auf $x_0 = b$, $f'(x_0) = c$ und die Funktion

$$g\colon\ ]a,b[\,-\{x_0\} \to \mathbb{R},\ g(x) = \frac{f(x) - f(x_0)}{x - x_0}$$

an. Dann folgt die Äquivalenz von (1) und (2).

"(2) ⇔ (3)" Man multipliziert oder dividiert mit $x - x_0 \neq 0$. □

Bemerkung 11.3. Wir setzen voraus, dass f an der Stelle x_0 differenzierbar ist. Teil 3 des vorigen Lemmas bedeutet dann, dass die Tangente von f unter allen Geraden in $\mathbb{R}^2$, die durch den Punkt $(x_0, f(x_0)$ gehen, diejenige ist, die f im Intervall $]x_0 - \delta, x_0 + \delta[$ am besten approximiert.

Definition 11.4. Sei $f\colon [a,b] \to \mathbb{R}$ eine Funktion.

1. Die Funktion f ist *rechtsseitig in a differenzierbar*, wenn der Grenzwert

   $$\lim_{x\to a+} \frac{f(x) - f(a)}{x - a} = f'(a)$$

   existiert. Dieser Grenzwert heißt die *rechsseitige Ableitung von f an der Stelle a.*

2. Die Funktion f ist *linksseitig in b differenzierbar*, wenn der Grenzwert

   $$\lim_{x\to b-} \frac{f(x) - f(b)}{x - b} = f'(b)$$

   existiert. Dieser Grenzwert heißt die *rechsseitige Ableitung von f an der Stelle b.*

Definition 11.5. Sei $z_0 \in U \subset \mathbb{R}$ und $f\colon U \to \mathbb{C}$ eine Funktion. Die Funktion f ist *komplex differenzierbar*, wenn der Grenzwert

$$\lim_{z\to\ z_0} \frac{f(z) - f(z_0)}{z - z_0} = f'(z_0) \in \mathbb{C}$$

existiert. Dieser Grenzwert heißt dann die *komplexe Ableitung von f an der Stelle z_0*. Wenn f für alle $z_0 \in U$ in z_0 komplex differenzierbar ist, so heißt f *komplex differenzierbar.*

Bemerkung 11.6. Wenn eine Funktion an einer Stelle komplex differenzierbar ist, dann ist sie insbesondere dort auch reell differenzierbar.

Definition 11.7. Sei $z_0 \in U \subset \mathbb{R}$ und $f\colon U \to \mathbb{C}$ eine Funktion. Die Funktion f ist *holomorph in z_0*, wenn es ein $\delta > 0$ gibt, so dass $U_\delta(z_0) \subset U$ und für alle $w \in U_\delta(z_0)$ die Funktion in w komplex differenzierbar ist. Eine Funktion heißt *holomorph*, wenn sie auf ihrem ganzen Definitionsgebiet holomorph ist.

Die Theorie holomorpher Funktion nennt man im Deutschen Funktionentheorie. Sie ist ein wichtiges Gebiet sowohl der klassischen, als auch modernen Analysis.

Beispiel 11.8. Betrachte $f(x) = x^n$ und $a \in \mathbb{C}$. Es gilt für alle $x \in \mathbb{C} - \{a\}$:

$$x^n - a^n = (x - a)\sum_{k=0}^{n-1} a^k x^{n-k} = x^{n-1} + ax^{n-2} + a^2x^{n-3} + \ldots + a^{n-2}x + a^{n-1}.$$

Also ist $x \mapsto x^n$ komplex differenzierbar (holomorph) mit Ableitung

$$f'(a) = \lim_{x\to a} \frac{x^n - a^n}{x - a} = \sum_{k=0}^{n-1} a^k (\lim_{x\to a} x^{n-k}) = na^{n-1}.$$

Beispiel 11.9. Sei $c \in \mathbb{C}$. Wir berechnen die Ableitung von $x \mapsto \exp(cx) = e^{cx}$ an der Stelle $x \in \mathbb{C}$. Dafür betrachte man:

$$\begin{aligned}\frac{e^{c(x+h)} - e^{cx}}{h} &= e^{cx}\frac{e^{ch}-1}{h} = e^{cx}h^{-1}\left(-1 + \sum_{k=0}^{\infty}\frac{(ch)^k}{k!}\right) \\ &= e^{cx}\left(-1 + \sum_{k=0}^{\infty}\frac{c^k h^{k-1}}{k!}\right) = e^{cx}\left(c + \frac{c^2}{2}h + \ldots\right)\end{aligned}$$

Es folgt:

$$\frac{d}{dx}e^{cx} = \lim_{h\to 0}\frac{e^{c(x+h)} - e^{cx}}{h} = e^{cx}\lim_{h\to 0}\big(c + \frac{c^2}{2}h + \ldots\big) = ce^{cx}$$

Deshalb ist die Exponentialfunktion exp überall komplex differenzierbar (holomorph) und es gilt

$$\exp'(cx) = c\exp(x).$$

Definition 11.10. Sei $f\colon U \to \mathbb{C}$ in x differenzierbar und $f'\colon U \to \mathbb{C}$ die Ableitung.

1. Wenn f' in x stetig ist, dann heißt f *in x stetig differenzierbar*.
2. Wenn f' in x differenzierbar ist, dann heißt f *in x zweimal differenzierbar*. Für die zweite Ableitung schreiben wir $f''(x)$.
3. Induktiv definiert man die Begriffe n-mal (stetig) differenzierbar und unendlich oft differenzierbar.

11.2 Grundlegende Sätze

Satz 11.11. *Sei f in x_0 (reell oder komplex) differenzierbar. Dann ist f in x_0 stetig.*

Beweis. Nach Voraussetzung existiert der Grenzwert nach dem ersten Gleichheitszeichen:

$$\begin{aligned}\lim_{x\to x_0} f(x) &= \lim_{x\to x_0}\left(f(x_0) + \frac{f(x)-f(x_0)}{x-x_0}(x-x_0)\right) = f(x_0) + f'(x_0)\left(\lim_{x\to x_0} x - x_0\right) \\ &= f(x_0).\end{aligned}$$

Aus Lemma 10.33 folgt dann die Stetigkeit von f. □

Satz 11.12. *Seien $f, g\colon U \to \mathbb{C}$ in $a \in U$ differenzierbar. Sei $\lambda \in \mathbb{C}$. Dann sind $f+g, f\cdot g$ und $\lambda \cdot f$ in a differenzierbar und es gilt:*

$$\begin{aligned}(f+g)'(a) &= f'(a) + g'(a), \\ (\lambda\cdot f)'(a) &= \lambda \cdot f'(a) \\ (f\cdot g)'(a) &= f'(a)g(a) + f(a)g'(a),\end{aligned}$$

Wenn $g(x_0) \neq 0$, dann ist $\frac{f}{g}$ in a differenzierbar und es gilt:

$$\left(\frac{f}{g}\right)'(a) = \frac{f'(a)g(a) - f(a)g'(a)}{g(a)^2}.$$

Beweis. Diese Rechenreglen folgen direkt aus den Rechenregeln für Grenzwerte.

$$\begin{aligned}(f+g)'(a) &= \lim_{h\to 0}\frac{(f+g)(a+h)-(f+g)(a)}{h}\\ &= \lim_{h\to 0}\frac{f(a+h)-f(a)}{h}+\lim_{h\to 0}\frac{g(a+h)-g(a)}{h}\\ &= f'(a)+g'(a)\end{aligned}$$

Die Aussage über die Ableitung von λf folgt aus der nächsten Rechnung, indem man für g die konstante Funktion bei λ einsetzt. Jetzt also zur Produktregel:

$$\begin{aligned}(f\cdot g)'(a) &= \lim_{h\to 0}\frac{1}{h}\big[f\cdot g)(a+h)-(f\cdot g)(a)\big]\\ &= \lim_{h\to 0}\frac{1}{h}\big[f(a+h)g(a+h)-f(a)g(a+h)+f(a)g(a+h)-f(a)g(a)\big]\\ &= \lim_{h\to 0}\frac{f(a+h)-f(a)}{h}g(a+h)-f(a)\lim_{h\to 0}\frac{g(a+h)-g(a)}{h}\\ &= f'(a)g(a)+f(a)g'(a)\end{aligned}$$

Dabei benutzen wir, dass g nach Satz 11.11 in a stetig und deswegen $\lim_{h\to 0} g(a+h)=g(a)$ gilt.

Schließlich zur Quotientenregel: Weil g stetig ist und $g(a)\neq 0$, wissen wir nach Proposition 10.16, dass g auch in einer kleinen Umgebung von a nicht verschwindet; also ist $g(a+h)\neq 0$ für kleines h.

$$\begin{aligned}\left(\frac{f}{g}\right)'(a) &= \lim_{h\to 0}\frac{1}{h}\left(\frac{f(a+h)}{g(a+h)}-\frac{f(a)}{g(a)}\right)\\ &= \lim_{h\to 0}\frac{1}{g(a+h)g(a)}\left(\frac{f(a+h)g(a)-f(a)g(a+h)}{h}\right)\\ &= \frac{1}{g(a)^2}\lim_{h\to 0}\frac{1}{h}\Big(f(a+h)g(a)-f(a)g(a+h)+f(a)g(a+h)-f(a)g(a+h)\Big)\\ &= \frac{1}{g(a)^2}\left(g(a)\lim_{h\to 0}\frac{f(a+h)-f(a)}{h}-f(a)\lim_{h\to 0}\frac{g(a+h)-g(a)}{h}\right)\\ &= \frac{f'(a)g(a)-f(a)g'(a)}{g(a)^2}\end{aligned}$$

□

Bemerkung 11.13. Die Gleichungen $(f+g)'=f'+g'$ und $(\lambda f)'=\lambda\cdot f'$ bedeuten, dass der Operator

$$\frac{d}{dx}=(-)'\colon\ \mathcal{D}(U,\mathbb{C})\to\mathcal{C}(U,\mathbb{C})$$

eine lineare Abbildung ist.

Beispiel 11.14. Es gilt nach Definition für alle $x\in\mathbb{R}$:

$$e^{ix}=\cos(x)+i\sin(x).$$

Nach Beispiel 11.9 gilt:

$$\left(e^{ix}\right)'=ie^{ix}.$$

Beides zusammen ergibt:

$$\begin{aligned}\cos'(x) + i\sin'(x) &= \frac{d}{dx}\big(\cos(x) + i\sin(x)\big) = \big(e^{ix}\big)' = ie^{ix} = i\big(\cos(x) + i\sin(x)\big)\\ &= i\cos(x) - \sin(x)\end{aligned}$$

Durch Vergleich der Real- und Imaginärteile folgt:

$$\cos'(x) = -\sin(x) \quad \text{und} \quad \sin'(x) = \cos(x).$$

Beispiel 11.15. Es gilt:

$$\mathrm{id}'(x) = \frac{d}{dx}x = \lim_{h\to 0}\frac{(x+h)-x}{h} = 1$$

Aus der Produktregel folgt dann

$$\frac{d}{dx}x^2 = \left(\frac{d}{dx}x\right)\cdot x + x \cdot \left(\frac{d}{dx}x\right) = 2x$$

Per vollständiger Induktion folgt:

$$\frac{d}{dx}x^n = \frac{d}{dx}(x^{n-1}\cdot x) = \left(\frac{d}{dx}x^{n-1}\right)x + x^{n-1}\left(\frac{d}{dx}x\right) = (n-1)x^{n-2}x + x^{n-1} = nx^{n-1}.$$

Beispiel 11.16. Die Betragsfunktion $|-|\colon \mathbb{R} \to \mathbb{R}$ ist in 0 nicht differenzierbar. Betrachte dazu aus der einen Seite

$$\lim_{h\to 0-}\frac{|h|}{h} = -1$$

und auf der anderen Seite

$$\lim_{h\to 0+}\frac{|h|}{h} = 1.$$

Also existiert der Grenzwert von $\frac{|h|}{h}$ für $h \to 0$ nicht.

11.3 Die Kettenregel

Satz 11.17 (Kettenregel)**.** *Seien $U, V \subset \mathbb{C}$ und seien $f\colon U \to V$ und $g\colon V \to \mathbb{C}$ Funktionen. Sei $x \in U$, f in x differenzierbar und g in $f(x)$ differenzierbar. Dann ist*

$$g \circ f\colon U \to \mathbb{C},\ x \mapsto g(f(x))$$

in x differenzierbar und es gilt:

$$(g \circ f)'(x) = f'(x)\cdot g'(f(x)).$$

Beweis. Setze $z = f(x) \in V$. Für diesen Beweis definieren wir eine Hilfsfunktion

$$\widetilde{g}\colon V \to \mathbb{C},\ \widetilde{g}(w) = \begin{cases} \frac{g(w)-g(z)}{w-z} & , \text{ für } w \neq z \\ g'(z) & , \text{ für } w = z \end{cases}.$$

Weil g in $z = f(x)$ differenzierbar ist, ist $\widetilde{g}$ stetig in z und es gilt erstens

$$\lim_{w \to z} \widetilde{g}(w) = g'(z) = g'(f(x)) \tag{11.18}$$

und zweitens für alle $w \in V$

$$g(w) - g(z) = \widetilde{g}(w)(w - z).$$

Damit folgt:

$$\begin{aligned}
(g \circ f)'(x) &= \lim_{y \to x} \frac{g(f(y)) - g(f(x))}{y - x} = \lim_{y \to x} \frac{\widetilde{g}(f(y))(f(y) - f(x))}{y - x} \\
&= \left(\lim_{y \to x} \widetilde{g}(f(y)) \right) \left(\lim_{y \to x} \frac{f(y) - f(x)}{y - x} \right)
\end{aligned}$$

Die Funktion f ist differenzierbar in x und deshalb dort insbesondere stetig nach Satz 11.11. Also folgt

$$\lim_{y \to x} f(y) = f(x) = z.$$

Mit Gleichung (11.18) folgt dann:

$$\begin{aligned}
(g \circ f)'(x) &= \left(\lim_{y \to x} \widetilde{g}(f(y)) \right) \left(\lim_{y \to x} \frac{f(y) - f(x)}{y - x} \right) \\
&= g'(f(x)) \cdot f'(x).
\end{aligned}$$

□

Die folgende Aussage gilt auch für Funktionen im Komplexen. Da wir den Satz 10.65 über die Stetigkeit der Umkehrfunktion benutzen und diesen nur für Funktionen im Reellen bewiesen haben, wird der folgende Satz auch nur für solche Funktionen formuliert.

Satz 11.19 (Ableitung der Umkehrfunktion)**.** *Sei I ein (un-) eigentliches Intervall in $\mathbb{R}$ und sei $f\colon I \to \mathbb{R}$ streng monoton mit Bild $J = f(I)$. Sei $g\colon J \to I$ die Umkehrfunktion von f. Wenn f an der Stelle $x \in I$ differenzierbar ist und $f'(x) \neq 0$, dann ist g an der Stelle $f(x) = y \in J$ differenzierbar und es gilt:*

$$g'(y) = \frac{1}{f'(x)} = \frac{1}{f'(g(y))}.$$

Beweis. Wir werden zeigen, dass der folgende Limes existiert und dass gilt

$$g'(y) = \lim_{z \to y} \frac{g(z) - g(y)}{z - y} = \frac{1}{f'(g(y))}.$$

Sei $(z_n)_{n \in \mathbb{N}}$ eine beliebige Folge in $J - \{y\}$ mit Grenzwert y. Setze $x_n = g(z_n)$; dann folgt $f(x_n) = z_n$. Weil f in x differenzierbar ist, erhält man jetzt:

$$\begin{aligned}
\lim_{n \to \infty} \frac{g(z_n) - g(y)}{z_n - y} &= \lim_{n \to \infty} \frac{x_n - x}{f(x_n) - f(x)} = \frac{1}{f'(x)} \\
&= \frac{1}{f'(g(y))}.
\end{aligned}$$

Dabei haben wir benutzt, dass nach Satz 11.11 f stetig in x ist und deshalb gilt:

$$\lim_{y \to x} f(y) = f(x).$$

Weil die Folge $(z_n)_{n \in \mathbb{N}}$ beliebig war, folgt damit die Aussage des Satzes. □

12 Der Logarithmus und allgemeine Potenzen

Die Exponentialfunktion $\exp\colon \mathbb{R} \to \mathbb{R}$ ist differenzierbar und streng monoton steigend mit Bild $\mathbb{R}_{>0}$. Also besitzt exp eine Umkehrfunktion.

Definition 12.1. Die Umkehrfunktion der reellen Exponentialfunktion heißt *natürlicher Logarithmus* und wir schreiben

$$\ln\colon \mathbb{R}_{>0} \to \mathbb{R}.$$

Proposition 12.2. *Es gilt:*

1. $\forall\, x \in \mathbb{R}_{>0} : e^{\ln x} = x$ *und* $\forall\, x \in \mathbb{R} : \ln(e^x) = x$
2. *Der Funktion* ln *ist streng monoton steigend.*
3. *Der natürliche Logarithmus ist differenzierbar und es gilt:*

$$\ln'(x) = \frac{d}{dx}\ln(x) = \frac{1}{x}.$$

4. *Es gilt die Funktionalgleichung des Logarithmus: für alle* $a, b \in \mathbb{R}_{>0}$*:*

$$\ln(ab) = \ln(a) + \ln(b)$$

Beweis. Teile 1 und 2 sind offensichtlich. Die Differenzierbarkeit in Teil 3 folgt aus Satz 11.19, genauso wie die Formel für die Ableitung:

$$\ln'(x) = \frac{1}{\exp'(\ln(x))} = \frac{1}{\exp(\ln(x))} = \frac{1}{x}$$

für alle $x > 0$.

Um die Funktionalgeichung zu beweisen, seien $a, b > 0$ und $\ln(a) = x$ und $\ln(b) = y$. Aus der Funktionalgleichung der Exponentialfunktion folgt:

$$e^{x+y} = e^x \cdot e^y.$$

Darauf wenden wir den Logarithmus an:

$$\ln(ab) = \ln(e^x e^y) = \ln(e^{x+y}) = x + y = \ln(a) + \ln(b).$$

□

Definition 12.3. Sei $\alpha \in \mathbb{C}$ und $x \in \mathbb{R}_{>0}$. Wir definieren

$$x^\alpha = \exp(\alpha \ln(x)) = e^{\alpha \ln(x)}.$$

Satz 12.4. *Seien* $\alpha, \beta \in \mathbb{C}$ *und* $x, y \in \mathbb{R}_{>0}$*. Dann gilt:*

1. $x^\alpha x^\beta = x^{\alpha+\beta}$
2. $(x^\alpha)^\beta = x^{\alpha\beta}$
3. $x^\alpha y^\alpha = (xy)^\alpha$

4. $\left(\frac{1}{x}\right)^{-1} = x^{-\alpha}$

5. $\frac{d}{dx}x^\alpha = \alpha x^{\alpha-1}$

6. $\frac{d}{d\alpha}x^\alpha = \ln(x)x^\alpha$

Beweis. Die ersten vier Aussagen folgen auf einfache Weise aus der Definition 12.3. Zu Teil 5:

$$\frac{d}{dx}x^\alpha = \frac{d}{dx}e^{\alpha\ln(x)} = \frac{\alpha}{x}e^{\alpha\ln(x)} = \frac{\alpha}{x}x^\alpha = \alpha x^{\alpha-1}$$

mit Hilfe der Kettenregel 11.17. Zu Teil 6:

$$\frac{d}{d\alpha}x^\alpha = \frac{d}{d\alpha}e^{\alpha\ln(x)} = \ln(x)e^{\alpha\ln(x)} = \ln(x)x^\alpha$$

wieder mit der Kettenregel. □

Teil 1 des folgenden Lemmas sagt umgangssprachlich, dass der Logarithmus langsamer als jede Potenz wächst.

Lemma 12.5. *Es gilt:*

1. $\lim_{x\to\infty} \frac{\ln(x)}{x^\alpha} = 0$

2. $\lim_{x\to 0} x^\alpha \ln(x) = 0$

Beweis. Folgt aus Lemma 10.39. □

13 Lokale und globale Extrema, Mittelwertsatz und Konsequenzen

13.1 Lokale Extrema

Definition 13.1. Sei $U \subset \mathbb{C}$ und $f\colon U \to \mathbb{C}$. Wir sagen, dass f in $z \in U$

1. ein *lokales Maximum* (*lokales Minimum*) besitzt, wenn ein $\delta > 0$ existiert, so dass für alle $w \in U_\delta(z) \cap U$ gilt:

$$f(w) \leqslant f(z) \quad (f(w) \geqslant f(z)).$$

Ein *lokales Extremum* ist ein lokales Maximum oder lokales Minimum.

2. ein *globales* oder *absolutes Maximum* (*globales/absolutes Minimum*) besitzt, wenn für alle $w \in U$ gilt:

$$f(w) \leqslant f(z) \quad (f(w) \geqslant f(z)).$$

Ein *globales/absolutes Extremum* ist ein globales Maximum oder globales Minimum.

Satz 13.2. *Sei $f\colon\ ]a,b[\to \mathbb{R}$ ein Funktion, die in $x \in]a,b[$ ein lokales Extremum besitzt und dort differenzierbar ist. Dann gilt $f'(x) = 0$.*

Beweis. Betrachte den Fall, in dem f ein x ein lokles Maximum hat. Der Fall für ein lokales Minimum folgt aus diesem ersten Fall durch Übergang zu $-f$.

Es exisiert ein $\delta > 0$, so dass $f(w) \leqslant f(z)$ für alle $w \in U_\delta(z) \cap U$ gilt. Weil f in x differenzierbar ist, existieren die folgenden Grenzwert und sind gleich:

$$0 \leqslant \lim_{y \to x-} \frac{f(y) - f(x)}{y - x} = f'(x) = \lim_{y \to x+} \frac{f(y) - f(x)}{y - x} \leqslant 0.$$

Es bleibt nichts anderes übrig als $f'(x) = 0$. □

Bemerkung 13.3. Es gilt folgendes zu beachten:

1. Die Bedingung $f'(x) = 0$ ist eine notwendige Bedingung dafür, dass x ein lokales Extremum ist. Sie ist nicht hinreichend. Als Beispiel vetrachte man $f(x) = x^3$. Hier gilt $f'(0) = 0$, aber 0 ist kein lokales Extremum.

2. Nach Satz 10.43 nimmt eine stetige Funktion $f\colon [a, b] \to \mathbb{R}$ ihr globales Minimum und globales Maximum an. Hier ist die Funktion auf einem abgeschlossenen Intervall definiert. Diese globalen Extrema können aber am Rand liegen und es muss dort nicht unbedingt $f'(x) = 0$ gelten. Als Beispiel betrachte man die Identität

$$\mathrm{id}\colon [0, 1] \to [0, 1], x \mapsto x.$$

   Das globale Maximum liegt bei 1 und die (linksseitige) Ableitung dort ist auch 1.

Satz 13.4 (Satz von Rolle)**.** *Sei $f\colon [a, b] \to \mathbb{R}$ eine stetige Funktion, die auf $]a, b[$ differenzierbar ist. Wenn $f(a) = f(b)$, dann existiert ein $x \in]a, b[$ mit $f'(x) = 0$.*

Michel Rolle `https://de.wikipedia.org/wiki/Michel_Rolle`

Beweis. Falls f konstant ist, müssen wir nichts zeigen. Jedes $x \in]a, b[$ erfüllt die Aussage.

Sei f nicht konstant. Dann gibt es ein $y \in]a, b[$ mit $f(y) > f(a)$ oder $f(y) < f(a)$. Das bedeutet, dass entweder das globale Maximum oder das globale Minimum im Inneren $]a, b[$ angenommen wird. Dort ist f differenzierbar. Also folgt nach Satz 13.2, dass es ein $x \in]a, b[$ gibt mit $f'(x) = 0$. □

Der folgende Satz ist ein einfaches Korollar aus dem Satz von Rolle.

Satz 13.5 (Der Mittelwertsatz der Differentialrechnung)**.** *Sei $f\colon [a, b] \to \mathbb{R}$ eine stetige Funktion, die auf $]a, b[$ differenzierbar ist. Dann existiert ein $x \in]a, b[$ mit*

$$f'(x) = \frac{f(b) - f(a)}{b - a}.$$

Beweis. Wir definieren die Hilfsfunktion $F\colon [a, b] \to \mathbb{R}$ durch

$$F(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a).$$

Sie hat dieselben Eigenschaften wir f: sie ist auf $[a, b]$ stetig und auf $]a, b[$ differenzierbar. Außerdem gilt:

$$F(a) = f(a) = F(b).$$

Nach dem Satz 13.4 von Rolle existiert ein $x \in]a,b[$ mit

$$0 = F'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}.$$

□

Korollar 13.6 (Schrankensatz, Version 1)**.** *Sei $f\colon [a,b] \to \mathbb{R}$ eine stetige Funktion, die auf $]a,b[$ differenzierbar ist. Es gebe $m, M \in \mathbb{R}$, so dass für alle $x \in]a,b[$ gilt:*

$$m \leqslant f'(x) \leqslant M.$$

Dann gilt für alle $x_1, x_2 \in [a,b]$ mit $x_1 \leqslant x_2$:

$$m(x_2 - x_1) \leqslant f(x_2) - f(x_1) \leqslant M(x_2 - x_1).$$

Beweis. Seien $x_1, x_2 \in [a,b]$ mit $x_1 \leqslant x_2$ beliebig. Aus dem Mittelwertsatz 13.5 folgt, dass es ein $x \in [x_1, x_2]$ gibt mit

$$f(x_2) - f(x_1) = f'(x)(x_2 - x_1).$$

Wenn aber für $f'(x)$ Schranken m und M existieren, folgt daraus die Aussage

$$m(x_2 - x_1) \leqslant f(x_2) - f(x_1) \leqslant M(x_2 - x_1).$$

□

Korollar 13.7. *Sei $f\colon [a,b] \to \mathbb{R}$ eine stetige Funktion, die auf $]a,b[$ differenzierbar ist. Wenn $f'(x) = 0$ für alle $x \in]a.b[$ gilt, dann ist f konstant.*

Beweis. Dies ist ein Spezialfall des Korollars 13.6 für $m = M = 0$. □

Korollar 13.8. *Seien $f, g\colon [a,b] \to \mathbb{R}$ stetige Funktion, auf $]a,b[$ differenzierbar, so dass für alle $x \in]a.b[$*

$$f'(x) = g'(x)$$

gilt. Dann gibt es ein $c \in \mathbb{R}$, so dass $g = f + c$ gilt.

Beweis. Wende das Korollar 13.7 auf die Funktion $F = f - g$ an. □

Korollar 13.9 (Schrankensatz, Version 2)**.** *Sei $I \subset \mathbb{R}$ und $f\colon I \to \mathbb{C}$ differenzierbar. Es gebe ein $L \in \mathbb{R}, L \geqslant 0$, so dass*

$$\|f'\|_\infty \leqslant L.$$

Dann folgt für alle $x_1, x_2 \in I$:

$$|f(x_2) - f(x_1)| \leqslant L|x_2 - x_1|.$$

Beweis. Dies folgt aus Korollar 13.6 mit $m = -L$ und $M = L$. □

Satz 13.10 (Monotoniekriterium)**.** *Sei $f\colon\ ]a,b[\to \mathbb{R}$ differenzierbar. Dann gilt:*

1. $\forall\, a < x < b : f'(x) \geqslant 0 \;\Leftrightarrow\; f$ *ist auf $]a,b[$ monoton wachsend*

2. $\forall\, a < x < b : f'(x) > 0 \;\Rightarrow\; f$ *ist auf* $]a, b[$ *streng monoton wachsend*

3. $\forall\, a < x < b : f'(x) \leqslant 0 \;\Leftrightarrow\; f$ *ist auf* $]a, b[$ *monoton fallend*

4. $\forall\, a < x < b : f'(x) < 0 \;\Rightarrow\; f$ *ist auf* $]a, b[$ *streng monoton fallend*

Wenn f sogar auf $[a, b]$ definiert und dort stetig, so gelten die Aussagen 1 bis 4, wenn man auf der rechten Seite $]a, b[$ durch $[a, b]$ ersetzt.

Beweis. Es sei $a < x_1 < x_2 < b$. Aus dem Mittelwertsatz 13.5 folgt, dass es ein $x \in [x_1, x_2]$ gibt mit

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(x).$$

Äquivalent dazu ist

$$f(x_2) - f(x_1) = f'(x)(x_2 - x_1),$$

wobei $x_2 - x_1 > 0$. Also gilt $f'(x) \geqslant 0$ genau dann, wenn $f(x_2) \geqslant f(x_2)$. In diesem Fall ist f monoton wachsend.

Wenn $f'(x) > 0$, dann folgt $f(x_2) > f(x_1)$ und f ist streng monoton wachsend.

Die anderen Fälle sind analog. □

Bemerkung 13.11. Man beachte, dass aus der Tatsache, dass f streng monoton wachsend ist, nicht ohne weiteres folgt, dass die Ableitung f' immer positiv ist. Ein Gegenbeispiel ist die Funktion $x \mapsto x^3$ an der Stelle 0. Also sind im vorigen Satz die Implikationen (2) und (4) nicht umkehrbar.

Bemerkung 13.12. Das Monotoniekriterium 13.10 liefert ein Möglichkeit, Funktionen auf lokale Extrema hin zu testen. Als Beispiel betrachten wir

$$f\colon \mathbb{R} \to \mathbb{R},\; f(x) = x^4.$$

Es ist offensichtlich, dass f in $x = 0$ ein lokales Minimum hat. Wir können versuchen mittels der Ableitung auf ein lokales Minimum zu testen. Es gilt

$$f'(0) = 4x^3|_{x=0} = 0.$$

Das ist aber nicht(!) hinreichend als Bedingung für ein lokales Minimum. Man sieht aber, dass f' in 0 einen Vorzeichenwechsel hat: $f'(x) < 0$ für alle $x < 0$ und $f'(x) > 0$ für alle $x > 0$. Das beudeutet, dass f auf $\mathbb{R}_{<0}$ streng monoton fallend und auf $\mathbb{R}_{>0}$ streng monoton wachsend ist. Weil f insbesondere stetig ist, folgt, dass f in 0 ein lokales Minimum hat.

Natürlich ist 0 ein globales Miminum und die einfachste Weise, das zu sehen, ist zu bemerken, dass

$$x^4 > 0 \implies x \neq 0.$$

Ein besseres Beispiel für den Test mittels Vorzeichenwechsel ist die Funktion

$$g\colon \mathbb{R} \to \mathbb{R}\;,\quad g(x) = x(1 + x^4) = x^5 + x^4.$$

Diese Funktion hat ein lokales, aber kein globales Minimum in $x = 0$. Die Ableituug ist

$$g'(x) = 5x^4 + 4x^3 = x^3(5x + 4).$$

Man sieht, dass $g'(0) = 0$ und dass g' einen Vorzeichenwechsel durchläuft, von Minus nach Plus. Es folgt, dass 0 ein lokales Minimum von g ist.

Definition 13.13. Sei $I \subset \mathbb{R}$ und $f\colon I \to \mathbb{R}$. Wir sagen, dass $x \in I$

- ein *strenges lokales Maximum* ist, wenn ein $\delta > 0$ existiert, so dass für alle $y \in]x-\delta, x+\delta[, y \neq x$ gilt: $f(y) < f(x)$.
- ein *strenges lokales Minimum* ist, wenn ein $\delta > 0$ existiert, so dass für alle $y \in]x-\delta, x+\delta[, y \neq x$ gilt: $f(y) > f(x)$.

Satz 13.14. *Sei $f\colon]a,b[\to \mathbb{R}$ eine differenzierbare Funktion. Sei $a < x < b$ ein Punkt, in dem f zweimal differenzierbar ist, und es gelte*

$$f'(x) = 0 \quad \textit{und} \quad f''(x) > 0 \quad (\textit{bzw. } f''(x) < 0).$$

Dann besitzt f in x ein strenges lokales Minimum (Maximum).

Beweis. Sei $f''(x) > 0$. Der Fall $f''(x) < 0$ geht durch Übergang zu $-f$ aus diesem ersten Fall hervor. Nach Voraussetzung gilt also

$$\lim_{\xi \to x} \frac{f'(\xi) - f'(x)}{\xi - x} = f''(x) > 0.$$

Es gibt demnach ein $\varepsilon > 0$, so dass für alle $a < \xi < b$ mit $0 < |\xi - x| < \varepsilon$

$$\frac{f'(\xi) - f'(x)}{\xi - x} > 0$$

gilt. Hier wurde benutzt, dass f' in x stetig ist; das ist in Ordnung, weil f' in x sogar differenzierbar ist.. Da $f'(x) = 0$ folgt daraus:

- $\forall\ x - \varepsilon < \xi < x : f'(\xi) < 0$
- $\forall\ x < \xi < x + \varepsilon : f'(\xi) > 0$

Nach dem Monotoniekriterium 13.10 ist f deshalb auf $[x - \varepsilon, x]$ streng monoton fallend und auf $[x, x+\varepsilon]$ streng monoton wachsend. Das bedeutet, dass f in x ein strenges lokales Minimum besitzt. □

Bemerkung 13.15. Satz 13.14 gibt ein hinreichende Bedingung für die Existenz eines lokalen strengen Extremums im Inneren eines Intervalls, aber keine notwendige Bedingung. Die Funktion $x \mapsto x^4$ hat in 0 ein strenges lokales Minimum. Es gilt jedoch $f''(0) = 0$.

13.2 Konvexe und konkave Funktionen

Definition 13.16. Sei $I \subset \mathbb{R}$ ein (un-)eigentliches Intervall. Eine Funktion $f\colon I \to \mathbb{R}$ heißt *konvex*, wenn für alle $x_1, x_2 \in I$ und alle $\lambda \in \mathbb{R}, 0 < \lambda < 1$ gilt:

$$f\big(\lambda x_1 + (1-\lambda)x_2\big) \leqslant \lambda f(x_1) + (1-\lambda) f(x_2).$$

Eine Funktion $f\colon I \to \mathbb{R}$ heißt *konkav*, wenn $-f$ konvex ist.

Bemerkung 13.17. Beachte:

1. Sei $x_1 < x_2$. Die Konvexitätsbedingung sagt dann, dass die Gerade (Sekante) durch die Punkte $(x_1, f(x_1)$ und $(x_2, f(x_2)$ im Intervall $]x_1, x_2[$ oberhalb des Graphen von f verläuft.

2. Die Funktion f ist genau dann konkav, wenn für alle alle $x_1, x_2 \in I$ und alle $\lambda \in \mathbb{R}, 0 < \lambda < 1$ gilt:

$$f\big(\lambda x_1 + (1-\lambda)x_2\big) \geqslant \lambda f(x_1) + (1-\lambda)f(x_2).$$

Diese Bedingung besagt, dass die Sekante durch $(x_1, f(x_1)$ und $(x_2, f(x_2)$ untrhalb des Graphen von f verläuft.

Beispiel 13.18. Die Exponentialfunktion ist konvex. Der (natürliche) Logarithmus ist konkav. Die Funktion $x \mapsto x^2$ ist konvex. Die Funktion $x \mapsto x^3$ ist auf $\mathbb{R}_{\geqslant 0}$ konvex und auf $\mathbb{R}_{\leqslant 0}$ konkav.

Satz 13.19. *Sei $I \subset \mathbb{R}$ ein offenes (un-)eigentliches Intervall und sei $f\colon I \to \mathbb{R}$ eine zweimal differenzierbare Funktion. Die Funktion f ist genau dann konvex, wenn für alle $x \in I$ gilt: $f''(x) \geqslant 0$.*

Beweis. "$\Leftarrow$" Sei $f''(x) \geqslant 0$ für alle $x \in I$. Dann ist f' auf I nach Satz 13.10 monoton steigend. Seien $x_1, x_2 \in I, x_1 < x_2, 0 < \lambda < 1$ und setze $x = \lambda x_1 + (1-\lambda)x_2$. Dann gilt: $x_1 < x < x_2$. Nach dem Mittelwertsatz existiert ein $\xi_1 \in]x_1, x[$ und ein $\xi_2 \in]x, x_2[$ mit

$$\frac{f(x) - f(x_1)}{x - x_1} = f'(\xi_1) \leqslant f'(\xi_2) = \frac{f(x_2) - f(x)}{x_2 - x}.$$

Es gilt:

$$x - x_1 = (1-\lambda)(x_2 - x_1) \quad \text{und} \quad x_2 - x = \lambda(x_2 - x_1).$$

Daraus folgt:

$$\frac{f(x) - f(x_1)}{1 - \lambda} \leqslant \frac{f(x_2) - f(x)}{\lambda}.$$

Nach Auflösen der Brüche folgt:

$$f(x) \leqslant \lambda f(x_1) + (1-\lambda)f(x_2).$$

Also ist f konvex.

"$\Rightarrow$" Sei f konvex. Wir nehmen an, es gäbe ein $x_0 \in I$ mit $f''(x_0) < 0$. Setze $c = f'(x_0)$ und

$$\varphi\colon I \to \mathbb{R},\ \varphi(x) = f(x) - c(x - x_0).$$

Dann ist φ zweimal differenzierbar mit

$$\varphi(x_0) = f(x_0),\ \varphi'(x_0) = 0 \quad \text{und} \quad \varphi''(x_0) = f''(x_0) < 0.$$

Nach Satz 13.14 ist x_0 ein strenges lokales Maximum von φ. Es gibt also ein $h > 0$, so dass $[x_0 - h, x_0 + h] \subset I$ und

$$\varphi(x_0 - h) < \varphi(x_0) \quad \text{und} \quad \varphi(x_0 + h) < \varphi(x_0).$$

Daraus folgt:

$$\begin{aligned}
f(x_0) = \varphi(x_0) &> \frac{1}{2}\big(\varphi(x_0 - h) + \varphi(x_0 + h)\big) \\
&= \frac{1}{2}\big(f(x_0 - h) - c((x_0 - h) - x_0) + f(x_0 + h) - c((x_0 + h) - x_0)\big) \\
&= \frac{1}{2}\big(f(x_0 - h) + f(x_0 + h)\big)
\end{aligned}$$

Das steht aber im Widerspruch dazu, dass f konvex ist. (Setze $x_1 = x_0 - h$, $x_2 = x_0 + h$ und $\lambda = \frac{1}{2}$. Dann ist $x_0 = \lambda x_1 + (1 - \lambda)x_2$). □

13.3 Die Regeln von de l'Hôpital

Guillaume François Antoine, marquis de L'Hôpital `https://de.wikipedia.org/wiki/Guillaume_Fran%C3%A7ois_Antoine,_Marquis_de_L%E2%80%99Hospital`

Lemma 13.20. *Sei $f\colon\,]0, a[\to \mathbb{R}$ eine differenzierbare Funktion mit*

$$\lim_{x\to 0+} f(x) = 0 \quad \textit{und} \quad \lim_{x\to 0+} f'(x) = c.$$

Dann gilt: $\lim_{x\to 0+} f(x)/x = c$.

Beweis. Wir betrachten den Spezialfall $c = 0$. Die Voraussetzung $\lim_{x\to 0+} f'(x) = 0$ bedeutet, dass ein $\delta > 0$ existiert, so dass für alle $0 < x < \delta$ gilt:

$$\lim_{y\to x} \frac{|f(x) - f(y)|}{|x - y|} = |f'(x)| \leqslant \varepsilon.$$

Daraus folgt, dass für alle $x, y \in\,]0, \delta[$ gilt:

$$|f(x) - f(y)| \leqslant \varepsilon|x - y|.$$

Wir bilden jetzt auf beiden Seiten den Grenzwert für $y \to 0$. Wegen der Voraussetzung $f(y) = 0$ geht die Ungleichung über in folgende Ungleichung:

$$|f(x)| = |f(x) - \lim_{y\to 0+} f(y)| \leqslant \varepsilon|x - \lim_{y\to 0+} y| = \varepsilon|x|.$$

Daraus folgt für alle $0 < x < \delta$

$$\frac{|f(x)|}{x} \leqslant \varepsilon.$$

Das ist die Behauptung im Fall $c = 0$.

Der allgemeine Fall folgt aus dem Spezialfall, indem man die Funktion $g(x) = f(x) - cx$ betrachtet. □

Lemma 13.21. *Sei $f\colon\,]a, +\infty[\, \to \mathbb{R}$ eine differenzierbare Funktion mit*

$$\lim_{x\to +\infty} f'(x) = c.$$

Dann gilt $\lim_{x\to +\infty} f(x)/x = c$.

Beweis. Wir betrachten zuerst den Spezialfall $c = 0$. Wegen $\lim_{x\to\infty} f'(x) = 0$ gibt es zu jedem $\varepsilon > 0$ ein $x_0 > a$, so dass für alle $x \geqslant x_0$ gilt:

$$|f'(x)| \leqslant \frac{\varepsilon}{2}.$$

Aus dem Schrankensatz folgt für alle $x \geqslant x_0$:

$$|f(x) - f(x_0)| \leqslant \frac{\varepsilon}{2}(x - x_0).$$

Bisher galt nur $x_0 > 0$, aber wir können ohne Beschränkung der Allgemeinheit annehmen, dass $x_0 > \max\{a, 0\}$. Für alle $x \geqslant \max\{x_0, \frac{2|f(x_0)|}{\varepsilon}\}$ folgt dann:

$$\begin{aligned}\frac{|f(x)|}{x} &\leqslant \frac{|f(x) - f(x_0)|}{x} + \frac{|f(x_0)|}{x} \leqslant \frac{|f(x) - f(x_0)|}{x - x_0} + \frac{|f(x_0)|}{x} \\ &\leqslant \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.\end{aligned}$$

Das ist die Behauptung im Spezialfall $c = 0$.

Der allgemeine Fall folgt aus dem Spezialfall, indem man die Funktion $g(x) = f(x) - cx$ betrachtet. □

Satz 13.22 (de L'Hôpital)**.** *Seie $f, g\colon\ ]a, b[\to \mathbb{R}$ differenzierbare Funktionen, wobei $-\infty \leqslant a < b \leqslant +\infty$. Es gelte $g'(x) \neq 0$ für alle $x \in I$. Ferner existiere der Limes*

$$\lim_{x\to b-} \frac{f'(x)}{g'(x)}.$$

Dann folgt:

1. *Wenn $\lim\limits_{x\to b-} g(x) = \lim\limits_{x\to b-} f(x) = 0$, dann ist $g(x) \neq 0$ für alle $x \in I$ und*
$$\lim_{x\to b-} \frac{f(x)}{g(x)} = c.$$
2. *Wenn $\lim\limits_{x\to b-} g(x) = \pm\infty$, so existiert ein $a < x_0 < b$ mit $g(x) \neq 0$ für alle $x \geqslant x_0$ und es gilt*
$$\lim_{x\to b-} \frac{f(x)}{g(x)} = c.$$

Analoge Aussagen gelten für $\lim_{x\to a+}$.

Beweis. Wir beweisen Teil 2, indem wir die Aussage auf Lemma 13.21 zurückfhren. Teil 1 wird analog mit Hilfe von Lemma 13.20 bewiesen.

Behauptung: g ist streng monoton und g' wechselt auf $]a, b[$ nicht das Vorzeichen.

Wenn es $x_1, x_2 \in]a, b[$ gäbe, so dass $g'(x_1)$ und $g'(x_2)$ entgegen gesetze Vorzeichen hätten, so wäre g nach dem Monotoniekriterium 13.10 in einen gewissen Bereich monoton steigend und in einem anderen Bereich monoton fallens. Dann besäße g ein lokales Extremum und nach Satz 13.2 gäbe es ein $x \in [x_1, x_2] \subset]a, b[$ mit $g'(x) = 0$. Diese Möglichkeit ist explizit ausgeschlossen worden. Es folgt, dass g' nicht das Vorzeichen wechselt. Wieder nach dem Monotonitekriterium folgt, dass g streng monoton (fallend oder steigend) ist.

Fall 1: g fällt streng monoton. Dieser Fall geht aus dem Fall 2 hervor, indem man zu $-g$ übergeht.

Fall 2: g wächst streng monoton. Es gilt $\lim\limits_{x\to b-} g(x) = +\infty$. Sei $g(]a,b[) =]A,+\infty[$, wobei $A = \lim_{x\to a+} g(x)$. Sei $h\colon\ ]A,+\infty[\to]a,b[$ die Umkehrfunktion von g. Betrachte die Hilfsfunktion

$$F = f\circ h\colon\ ]A,+\infty[\to \mathbb{R}.$$

Nach der Kettenregel 11.17 und dem Satz 11.19 über die Ableitung der Umkehrfunktion folgt für alle $y > A$:

$$F'(y) = f'(h(y))h'(y) = \frac{f'(h(y))}{g'(h(y))}.$$

Die Voraussetzung sagt:

$$\lim_{y\to+\infty} F'(y) = \lim_{x\to b-} \frac{f'(x)}{g'(x)} = c.$$

Wir können jetzt das Lemma 13.21 auf F anwenden. Es besagt dann:

$$\lim_{y\to\infty} \frac{F(y)}{y} = c.$$

Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $I =]a,b[$ mit Grenzwert b. Setze $y_n = g(x_n)$. Dann folgt $\lim_{n\to\infty} g(x_n) = \infty$ und es gilt:

$$\lim_{n\to\infty} \frac{f(x_n)}{g(x_n)} = \lim_{n\to\infty} \frac{f(h(y_n))}{y_n} = \lim_{n\to\infty} \frac{F(y_n)}{y_n} = c.$$

Weil die Folge $(x_n)_{n\in\mathbb{N}}$ beliebig war, folgt die Aussage von Teil 2. $\square$

Beispiel 13.23. Es gilt nach de L'Hôpital:

$$\lim_{\substack{x\to 0\\ x\neq 0}} \frac{\sin(x)}{x} = \lim_{x\to 0} \frac{\cos(x)}{1} = 1.$$

14 Die Zahl π, weitere trigonometrische Funktionen und Polarkoordinaten

14.1 Definition von π

Zur Erinnerung aus Lemma 9.17:

$$\cos(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k} = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} \pm \dots$$

$$\sin(x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1} = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \pm \dots.$$

Satz 14.1 (Abschätzung der Restglieder)**.** *Es gilt*

$$\cos(x) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k)!} x^{2k} + r_{2n+2}(x)$$
$$\sin(x) = \sum_{k=0}^{n} \frac{(-1)^k}{(2k+1)!} x^{2k+1} + r_{2n+3}(x),$$

mit

$$|r_{2n+2}(x)| \leqslant \frac{|x|^{2n+2}}{(2n+2)!} \quad \textit{für} \quad |x| \leqslant 2n+3,$$
$$|r_{2n+3}(x)| \leqslant \frac{|x|^{2n+3}}{(2n+3)!} \quad \textit{für} \quad |x| \leqslant 2n+4.$$

Beweis. Es ist

$$r_{2n+2}(x)) = \pm \frac{x^{2n+2}}{(2n+2)!} \left(1 - \frac{x^2}{(2n+3)(2n+4)} \pm \dots \right).$$

Für $k \geqslant 1$ setze

$$a_k = \frac{x^{2k}}{(2n+3)(2n+4) \cdot \ldots \cdot (2n+2k+2)}.$$

Damit gilt erstens

$$r_{2n+2}(x)) = \pm \frac{x^{2n+2}}{(2n+2)!} (1 - a_1 + a_2 - a_3 \pm \dots) \tag{14.2}$$

und zweitens für alle $|x| \leqslant 2n+3$

$$a_k = a_{k-1} \frac{x^2}{(2n+2k+1)(2n+2k+2)}.$$

Solange $|x| \leqslant 2n+3$, ist die Folge $(a_k)_{k \in \mathbb{N}}$ eine streng monoton fallende Nullfolge mit

$$1 > a_1 > a_2 > a_3 > \ldots > 0.$$

Wie im Beweis des Leibnizkriteriums 8.5 für alternierende Reihen folgt daraus

$$0 \leqslant 1 - a_1 + a_2 - a_3 \pm \ldots \leqslant 1.$$

Aus (14.2) folgt dann die Behauptung

$$|r_{2n+2}(x)| \leqslant \frac{|x|^{2n+2}}{(2n+2)!}.$$

Der Beweis der Abschätzung des Restglieds des Sinus geht analog. □

Lemma 14.3. *Es gilt:*

1. $\cos(0) = 1$

2. $\cos(2) \leqslant -\frac{1}{3}$

Beweis. Weil $e^0 = 1$ folgt:

$$\cos(0) = \mathrm{Re}\left(e^{i\cdot 0}\right) = 1.$$

Um die Abschätzung zu beweisen, nutzen wir den Satz 14.1:

$$\cos(x) = 1 - \frac{x^2}{2} + r_4(x),$$

wobei für $|x| \leqslant 5$

$$|r_4(x)| \leqslant \frac{|x|^4}{24}.$$

Für $x = 2$ ergibt sich

$$\cos(2) = 1 - 2 + r_4(2) \quad \text{und} \quad |r_4(2)| \leqslant \frac{16}{24} = \frac{2}{3}.$$

Es folgt:

$$\cos(2) \leqslant -1 + \frac{2}{3} = -\frac{2}{3}.$$

□

Lemma 14.4. *Für alle* $x \in \,]0, 2]$ *ist* $\sin(x) > 0$.

Beweis. Für $x \neq 0$ können wir die folgende Umformung machen:

$$\sin(x) = x + r_3(x) = x\left(1 + \frac{r_3(x)}{x}\right).$$

Nach der Abschätzung des Restglied 14.1 gilt

$$|r_3(x)| \leqslant \frac{|x|^3}{3!}.$$

für alle $|x| \leqslant 4$. Für alle $0 < x \leqslant 2$ gilt also

$$\left|\frac{r_3(x)}{x}\right| \leqslant \frac{|x|^3}{3!} \cdot \frac{1}{|x|} = \frac{|x|^2}{3!} \leqslant \frac{4}{6} = \frac{2}{3}$$

und damit

$$\sin(x) = x\left(1 + \frac{r_3(x)}{x}\right) \geqslant x\left(1 - \frac{2}{3}\right) = \frac{x}{3} > 0$$

für alle $0 < x \leqslant 2$. □

Proposition 14.5. *Der Cosinus hat im Intervall* $]0, 2]$ *genau eine Nullstelle.*

Beweis. Der Cosinus ist stetig. Es gilt $\cos(0) = 1$ und $\cos(2) < 0$ nach Lemma 14.3. Daher hat der Cosinus nach dem Zwischenwertsatz 10.40 in $]0, 2]$ mindestens eine Nullstelle.

Die Ableitung des Cosinus ist $-\sin$ und diese ist auf $]0, 2]$ kleiner als Null nach Lemma 14.4. Aus dem Monotoniekriterium 13.10 folgt dann, dass der Cosinus in $[0, 2]$ streng monoton fallend ist. Also hat er höchstens eine Nullstelle in $[0, 2]$. □

Definition 14.6. Nach Proposition 14.5 hat der Cosinus in $]0, 2]$ genau eine Nullstelle x_0. Wir definieren:

$$\pi = 2x_0.$$

14.2 Grundlegende Aussagen über Sinus und Cosinus

Proposition 14.7 (Additionstheoreme von Sinus und Cosinus)**.** *Für alle $x, y \in \mathbb{R}$ gilt:*

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$$
$$\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y)$$

Beweis. Mit Hilfe der Funktionalgleichung der Exponentialfunktion folgt:

$$\begin{aligned}\cos(x+y) + i\sin(x+y) = e^{i(x+y)} = e^{ix}\cdot e^{iy} &= \Big(\cos(x) + i\sin(x)\Big)\Big(\cos(y) + i\sin(y)\Big)\\ &= \cos(x)\cos(y) + i\cos(x)\sin(y) + i\sin(x)\cos(y) - \sin(x)\sin(y)\\ &= \cos(x)\cos(y) - \sin(x)\sin(y) + i\big(\sin(x)\cos(y) + \cos(x)\sin(y)\big)\end{aligned}$$

Jetzt vergleichen wir Real- und Imaginärteile und die behaupteten Gleichungen folgen. □

Lemma 14.8. *Es gilt:*

$$e^{i\frac{\pi}{2}} = i,\ e^{i\pi} = -1,\ e^{i\frac{3\pi}{2}} = -i,\ e^{2\pi i} = 1.$$

Beweis. Nach Definition gilt: $\cos(\frac{\pi}{2}) = 0$. Es folgt:

$$\sin(\frac{\pi}{2})^2 = 1 - \cos(\frac{\pi}{2})^2 = 1$$

Nach Lemma 14.4 ist $\sin(\frac{\pi}{2}) > 0$. Es folgt also:

$$\sin(\frac{\pi}{2}) = 1.$$

Insbesondere folgt dann auch:

$$e^{i\frac{\pi}{2}} = \cos(\frac{\pi}{2}) + i\sin(\frac{\pi}{2}) = i.$$

Die anderen Gleichung folgen aus $e^{i\frac{n\pi}{2}} = i^n$. □

Lemma 14.9. *Es gilt:*

$$\cos(x+2\pi) = \cos(x) \qquad \sin(x+2\pi) = \sin(x)$$
$$\cos(x+\pi) = -\cos(x) \qquad \sin(x+\pi) = -\sin(x)$$

Beweis. Diese Aussagen folgen direkt aus dem Additionstheorem 14.7 und den speziellen Werten aus Lemma 14.8. □

Lemma 14.10. *Es gilt:*

1. *Ist $x \in \mathbb{R}$ eine Nullstelle des Sinus, dann existiert ein $n \in \mathbb{Z}$ mit $x = n\pi$.*
2. *Ist $x \in \mathbb{R}$ eine Nullstelle des Cosinus, dann existiert ein $n \in \mathbb{Z}$ mit $x = n\pi + \frac{\pi}{2}$.*

Beweis. Wir beweisen nur Teil 2. Der andere Teil geht ähnlich. Wir wissen schon nach Proposition 14.5 bzw. Defintion 14.6, dass die Zahlen der Form $n\pi + \frac{\pi}{2}$ für $n \in \mathbb{Z}$ Nullstellen des Cosinus sind, denn es gilt $\cos(\frac{\pi}{2}) = 0 = \cos(\frac{3\pi}{2})$ nach Lemma 14.9 und der Cosinus ist auf Grund desselben Lemmas 2π-periodisch.

Es bleibt zu zeigen, dass es keine anderen Nullstellen gibt. Wir nehmen an, dass x eine Nullstelle ist, die nicht von der Form $n\pi + \frac{\pi}{2}$ ist. Wegen den Gleichung in 14.9 ist dann $x + k\pi$ für alle $k \in \mathbb{Z}$ auch eine Nullstelle des Cosinus. Dann können wir mittels dieser Periodizität ein $\widetilde{x} \in]0, \pi[$ finden mit $\cos(\widetilde{x}) = 0$. Man addiert einfach πk für ein $k \in \mathbb{Z}$, bis $\widetilde{x} = x + \pi k$ in $]0, \pi[$ liegt.

Wir wissen aus Proposition 14.5, dass $\frac{\pi}{2}$ die einzige Nullstelle in $]0, 2]$ ist. Wieder mittels der Gleichungen 14.9 kann es dann aber keine weitere Nullstelle in $]0, 2] \cup [\frac{\pi}{2}, \frac{3\pi}{2}[=]0, \frac{3\pi}{2}[$ geben. Das Intervall deckt also den Bereich $]0, \pi[$ ab, in dem $\widetilde{x}$ liegen müsste; $\widetilde{x}$ und daher auch x können nicht existieren. □

Satz 14.11. *Wir fassen zusammen:*

1. *Die Funktionen* $\sin, \cos\colon \mathbb{R} \to \mathbb{R}$ *sind periodisch mit Periode* 2π*, d.h. für alle* $x \in \mathbb{R}$ *gilt:*
$$\sin(z) = \sin(x + 2\pi) \quad \textit{und} \quad \cos(x) = \cos(x + 2\pi).$$

2. *Die Exponentialfunktion ist periodisch mit Periode* $2\pi i$*, d.h. für alle* $z \in \mathbb{C}$ *gilt:*
$$\exp(z) = \exp(z + 2\pi i).$$

Beweis. Wir haben Teil 1 schon in Lemma 14.9 bewiesen. Teil 2 folgt jetzt mittels der Formel
$$e^{i\varphi} = \cos(\varphi) + i\sin(\varphi)$$
für alle $\varphi \in \mathbb{R}$. □

a

14.3 Weitere trigonometrische Funktionen

Wir wissen jetzt also, dass der Kosinus im Intervall $]-\frac{\pi}{2}, \frac{\pi}{2}[$ positiv ist. Wegen $\sin' = \cos$ folgt aus dem Monotoniekriterium, dass der Sinus auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$ streng monoton steigend ist. Deshalb existiert dort seine Umkehrfunktion. Ähnlich argumentiert man für den Cosinus; er ist auf $[0, \pi]$ streng monoton fallend.

Definition 14.12. Wir definieren:

1. Die Umkehrfunktion des Sinus $\sin\colon [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1]$ ist
$$\arcsin\colon [-1, 1] \to [-\frac{\pi}{2}, \frac{\pi}{2}]$$
der *Arcus sinus.*

2. Die Umkehrfunktion des Cosinus $\cos\colon [0, \pi] \to [-1, 1]$ ist
$$\arccos\colon [-1, 1] \to [0, \pi]$$
der *Arcus cosinus.*

Proposition 14.13. *Die Funktionen* arcsin *und* arccos *sind stetig und im Intervall* $]-1,1[$ *unendlich oft differenzierbar und es gilt:*

$$\arcsin'(x) = \frac{1}{\sqrt{1-x^2}} \quad \textit{und} \quad \arccos'(x) = \frac{1}{\sqrt{x^2-1}}$$

Beweis. Der (Co-)Sinus ist auf ganz $\mathbb{R}$ unendlich oft differenzierbar. Also sind nach Satz 10.65 der Arcus Sinus und der Arcus Cosinus auf ihrem Definitionsgebiet $[-1,1]$ stetig.

Für die Differenzierbarkeit müssen wir die Nullstellen der Ableitungen von sin und cos vermeiden. Diese liegen gerade dort, wo sin und cos die Werte ± 1 annehmen. Nach dem Satz über die Ableitung der Umkehrfunktion sind die Funktion arcsin und arccos auf $]-1,1[$ unendlich oft differenzierbar. Es gilt weiter:

$$\sin'(x) = \cos(x)$$

Die Formel für die Ableitung der Umkehrfunktion ist dann:

$$\arcsin'(x) = \big(\sin'(\arcsin(x)\big)^{-1} = \frac{1}{\cos(\arcsin(x))}$$

Um weiter zu rechnen, benutzen wir

$$\sin^2 + \cos^2 = 1 \quad \Leftrightarrow \quad \cos(x) = \sqrt{1-\sin(x)^2}.$$

Also:

$$\arcsin'(x) = \frac{1}{\cos(\arcsin(x))} = \frac{1}{\sqrt{1-\sin(\arcsin(x))^2}} = \frac{1}{\sqrt{1-x^2}}.$$

Genauso rechnet man nach:

$$\arccos'(x) = \frac{1}{-\sin(\arccos(x))} = \frac{1}{-\sqrt{1-\cos(\arccos(x))^2}} = \frac{1}{\sqrt{x^2-1}}.$$

□

Definition 14.14. Eine weitere fundamentale Funktion ist der *Tangens* definiert durch

$$\tan\colon\ ]-\frac{\pi}{2},\frac{\pi}{2}[\ \to \mathbb{R},\ \tan(x) = \frac{\sin(x)}{\cos(x)}.$$

Natürlich kann man diese Funktion auf $\mathbb{R} - \mathrm{Null}(\cos) = \mathbb{R} - \{\frac{\pi}{2} + k\pi \mid k \in \mathbb{Z}\}$ definieren.

Lemma 14.15. *Die Funktion* $\tan\colon\ ]-\frac{\pi}{2},\frac{\pi}{2}[\ \to \mathbb{R}$ *ist differenzierbar und es gilt:*

$$\tan'(x) = \frac{1}{\cos(x)^2}.$$

Beweis. Nach der Quotientenregel 11.12 ist der Tangens differenzierbar und es gilt:

$$\tan'(x) = \frac{d}{dx}\left(\frac{\sin(x)}{\cos(x)}\right) = \frac{\cos(x)^2 + \sin(x)^2}{\cos(x)^2} = \frac{1}{\cos(x)^2}.$$

□

Lemma 14.16. *Der Tangens ist im Intervall* $]-\frac{\pi}{2},\frac{\pi}{2}[$ *streng monoton wachsend und bildet das Intervall bijektiv auf* $\mathbb{R}$ *ab.*

Beweis. Wir haben gerade gesehen, dass $\tan'(x) = \cos(x)^{-2} > 0$ für alle $-\frac{\pi}{2} < x < \frac{\pi}{2}$. Aus dem Monotoniekriterium 13.10 folgt, dass tan auf $]-\frac{\pi}{2},\frac{\pi}{2}[$ streng monoton wachsend ist. Insbesondere bildet tan das Intervall $]-\frac{\pi}{2},\frac{\pi}{2}[$ bijektiv auf sein Bild ab.

Wir zeigen jetzt:

$$\tan(x) \to +\infty \quad \text{für} \quad x \to \frac{\pi}{2}.$$

Wegen $\tan(-x) = -\tan(x)$ folgt dann automatisch

$$\tan(x) \to -\infty \quad \text{für} \quad x \to -\frac{\pi}{2}.$$

Damit ist das Bild des Tangens

$$\tan\Big(]-\frac{\pi}{2},\frac{\pi}{2}[\Big) = \mathbb{R}$$

und das Lemma ist bewiesen.

Sei jetzt $(x_n)_{n\in\mathbb{N}}$ ein Folge in $]-\frac{\pi}{2},\frac{\pi}{2}[$ mit

$$\lim_{n\to\infty} x_n = \frac{\pi}{2}.$$

Wir können auch noch voraussetzen, dass $x_n > 0$ für alle $n \in \mathbb{N}$. Dann ist für alle $n \in \mathbb{N}$

$$y_n = \frac{\cos(x_n)}{\sin(x_n)} > 0,$$

und es gilt:

$$\lim_{n\to\infty} y_n = \frac{\lim_{n\to\infty}\sin(x_n)}{\lim_{n\to\infty}\cos(x_n)} = \frac{\cos(\pi/2)}{\sin(\pi/2)} = 0.$$

Aus Proposition 3.18 folgt dann

$$+\infty = \lim_{n\to\infty} \frac{1}{y_n} = \lim_{n\to\infty} \tan(x_n).$$

□

Definition 14.17. Die Umkehrfunktion des Tangens $\tan\colon\ ]-\frac{\pi}{2},\frac{\pi}{2}[\ \to \mathbb{R}$ ist der *Arcus tangens*

$$\arctan\colon \mathbb{R} \to\ ]-\frac{\pi}{2},\frac{\pi}{2}[.$$

Lemma 14.18. *Der Arcus tangens ist differenzierbar und es gilt für alle* $x \in \mathbb{R}$

$$\arctan'(x) = \frac{1}{1+x^2}.$$

Beweis. Weil der Tanges auf $]-\frac{\pi}{2},\frac{\pi}{2}[$ differenzierbar ist und ohne Nullstelle der Ableitung, ist der Arcus tangens differenzierbar, und es gilt:

$$\arctan'(x) = \big(\tan'(\arctan(x))\big)^{-1} = \cos(\arctan(x))^2$$

Setze $y = \arctan x$. Es folgt:

$$x^2 = \tan(y)^2 = \frac{\sin(y)^2}{\cos(y)^2} = \frac{1-\cos(y)^2}{\cos(y)^2} = \frac{1}{\cos(y)^2} - 1$$

Das ist äquivalent zu:

$$\cos(y)^2 = \frac{1}{1+x^2}.$$

Wir machen jetzt unsere Substitution rückgängig und es folgt:

$$\arctan'(x) = \cos(\arctan(x))^2 = \frac{1}{1+x^2}.$$

□

14.4 Polarkoordinaten

Wir können eine Ebene parametrisieren, indem wir ein rechtwinkliges Koordinatensystem wählen. Das bedeutet, dass man eine Ursprung, die "Null", auswählt und dazu noch zwei Richtungen, die einen rechten Winkel zueinander haben. Im $\mathbb{R}^2$ ist uns Null, aka. der Nullvektor, schon vorgegeben. Wir können noch eine $\mathbb{R}$-Basis $\{v_1, v_2\} \subset \mathbb{R}^2$ wählen, so dass erstens $|v_1| = |v_2| = 1$ und $< v_1, v_2 >= 0$ ist. Man nennt dies eine "Orthonormal-Basis". Im $\mathbb{R}^2$ ist eine (kanonische) Wahl $v_1 = (1, 0)$ und $v_2 = (0, 2)$. Natürlich kann man andere (orthonormale) Basen wählen. Ein Punkt (Vektor) $v \in \mathbb{R}$ läßt sich dann eindeutig als Linearkombination der Basis schreiben:

$$v = \lambda_1 v_1 + \lambda_2 v_2 \ , \ \lambda_1, \lambda_2 \in \mathbb{R}.$$

Im Spezialfall $v_1 = (1, 0)$ und $v_2 = (0, 1)$ ließt sich das als

$$v = \lambda_1(1, 0) + \lambda_2(0, 1) = (\lambda_1, \lambda_2) \in \mathbb{R}^2. \tag{14.19}$$

Wenn wir $\mathbb{R}^2$ mit $\mathbb{C}$ identifizieren, so identifiziert man $(1, 0)$ mit 1 und $(0, 1)$ mit i. Also bilden 1 und i eine Orthonormal-Basis von $\mathbb{C}$ über $\mathbb{R}$. In der Gleichung (14.19) wird dann λ_1 zum Realteil und λ_2 zum Imaginärteil von v.

Die Wahl eines Koordinatensystems ist oft sehr angenehm für Berechnungen, aber immer vom Betrachter abhängig und völlig willkürlich. Es gibt viele andere Möglichkeiten, die Ebene zu parametrisieren. Eine weitere wichtige Möglichkeit sind Polarkoordinaten. Man kann einen Punkt in $\mathbb{C} - \{0\}$ auch dadurch beschreiben, dass man den Winkel angibt, den die Gerade durch 0 und den Punkt mit der positive reellen Achse bildet, und den Abstand der Punktes von 0.

Satz 14.20 (Polarkoordinaten von $\mathbb{C}$). *Jede komplexe Zahl z läßt sich schreiben als*

$$z = r \cdot e^{i\varphi},$$

wobei $\varphi \in \mathbb{R}$ und $r = |z| \in \mathbb{R}_{\geqslant 0}$. Für $z \neq 0$ ist φ bis auf Addition von $2\pi k, k \in \mathbb{Z}$ eindeutig bestimmt.

Die Zahl φ ist der Winkel im Bogenmaß zwischen dem Ortsvektor von z (bzw. der Geraden durch z und 0) und der positiven reellen Achse, wenn man bei der positiven x-Achse anfängt und im mathematisch positive Sinn (also gegen den Uhrzeigersinn) bis zum Ortsvektor von z geht. Man nennt φ auch *das Argument* und

$$z = r \cdot e^{i\varphi},$$

die *Polarkoordinatendarstellung* von z.

Beweis. Gegeben eine komplexe Zahl $z = re^{i\varphi}$ in Polarkoordinatendarstellung, dann ist

$$x = r\cos(\varphi) = \mathrm{Re}\,(z) \quad \text{und} \quad y = r\sin(\varphi) = \mathrm{Im}\,(z),$$

also $z = x + iy$ in der bisher üblichen Schreibweise.
Für $z = 0$ ist $z = 0 \cdot e^{i\varphi}$ mit beliebigem φ eine Polarkoordinatendarstellung. Sei $z \neq 0$. Setze

$$r = |z|\ ,\ w = \frac{z}{|z|} = \frac{z}{r}\ ,\ a = \mathrm{Re}\,(w) \ \text{ und } \ b = \mathrm{Im}\,(w),$$

so dass gilt:

$$w = a + ib\ ,\ |w| = 1\ ,\ a^2 + b^2 = |w|^2 = 1, \quad \text{und} \quad |a|, |b| \leqslant 1 \tag{14.21}$$

Setze jetzt

$$\varphi = \begin{cases} \arccos(a) & , \text{ wenn } b \geqslant 0 \\ -\arccos(a) & , \text{ wenn } b < 0 \end{cases}$$

Dann gilt $w = e^{i\varphi}$ denn für $\varphi \in [0, \pi]$

$$e^{i\varphi} = \cos(\varphi) + i\sin(\varphi) = a + ib = w,$$

wie man mit einer Fallunterscheidung nach $b \geqslant 0$ und $b < 0$ und den Gleichungen (14.21) leicht sehen kann. Es folgt:

$$z = |z|\frac{z}{|z|} = rw = re^{i\varphi}.$$

□

Wir haben insbesondere gezeigt:

Proposition 14.22. *Sei $z \in \mathbb{Z}$ eine Zahl mit $|z| = 1$. Dann existiert ein eindeutig bestimmtes $0 \leqslant \varphi < 2\pi$ mit*

$$z = e^{i\varphi}.$$

Bemerkung 14.23. Man beachte, dass sich die Multiplikation komplexer Zahlen in Polarkoordinaten besonders einfach schreiben läßt. Seien $r, s \geqslant 0$ und $\varphi, \vartheta \in \mathbb{R}$. Für

$$z = re^{i\varphi} \quad \text{und} \quad w = se^{i\vartheta}$$

gilt:

$$zw = \left(re^{i\varphi}\right)\left(w = se^{i\vartheta}\right) = rse^{\varphi + \vartheta}.$$

Damit erhält man wieder die geometrische Interpretation der komplexen Multiplikation aus Bemerkung 2.28 bzw. Lemma 2.29.

14.5 Sinus hyperbolicus, Cosinus hyperbolicus und die Area-Funktionen

Definition 14.24. Wir definieren die Funktionen *Sinus hyperbolicus*

$$\sinh\colon \mathbb{R} \to \mathbb{R} \;\; \sinh(x) = \frac{1}{2}(e^x - e^{-x})$$

und *Cosinus hyperbolicus*

$$\cosh\colon \mathbb{R} \to \mathbb{R} \;\; \cosh(x) = \frac{1}{2}(e^x + e^{-x}).$$

Lemma 14.25. *Es gilt für alle $x \in \mathbb{R}$:*

$$\cosh(x)^2 - \sinh(x)^2 = 1$$

Beweis. Aus den Definitionen folgt:

$$\begin{aligned}\cosh(x)^2 - \sinh(x)^2 &= \frac{1}{4}\left((e^x + e^{-x})^2 - (e^x - e^{-x})^2\right)\\ &= \frac{1}{4}\left(e^{2x} + 2e^x e^{-x} + e^{-2x} - e^{2x} + 2e^x e^{-x} - e^{-2x}\right)\\ &= e^x e^{-x} = 1.\end{aligned}$$

□

Lemma 14.26. *Die Funktionen sind überall unendlich oft differenzierbar und es gilt:*

$$\sinh'(x) = \cosh(x) \quad \textit{und} \quad \cosh'(x) = \sinh(x).$$

Der Sinus hyperbolicus ist streng monoton steigend.

Beweis. Man rechnet einfach nach:

$$\begin{aligned}\sinh'(x) &= \frac{1}{2}\frac{d}{dx}(e^x - e^{-x}) = \frac{1}{2}(e^x + e^{-x}) = \cosh(x)\\ \cosh'(x) &= \frac{1}{2}\frac{d}{dx}(e^x + e^{-x}) = \frac{1}{2}(e^x - e^{-x}) = \sinh(x)\end{aligned}$$

Die Tatsache, dass sinh streng monoton steigend ist, folgt aus dem Monotoniekriterium und der Tatsache, dass $e^x, e^{-x} > 0$ für alle $x \in \mathbb{R}$. □

15 Die komplexen Zahlen sind algebraisch abgeschlossen

Satz 15.1 (Fundamentalsatz der Algebra)**.** *Der Körper $\mathbb{C}$ ist algebraisch abgeschlossen. Genauer: für $n \in \mathbb{N}$ seien $a_0, \ldots, a_n \in \mathbb{C}, a_n \neq 0$ und*

$$p(z) = \sum_{k=0}^{n} a_k z^k.$$

Dann existiert ein $z \in \mathbb{C}$ mit $p(z) = 0$.

Hier ist es wichtig, dass $n = 0$ explizit ausgeschlossen ist. Ein Polynom vom Grad $n = 0$ ist konstant. Wenn es nicht gerade konstant Null ist, hat es keine Nullstelle.

Beweis. Ohne Beschränkung der Allgemeinheit können wir annhemen, dass $a_n = 1$ ist. Wenn nicht, so können wir die Gleichung $p(z) = 0$ durch $a_n \neq 0$ teilen.

Setze $\mu = \inf_{z \in \mathbb{C}} |p(z)|$. Diese Infimum existiert nach dem Satz 6.19, denn die Menge $\{|p(z)| \mid z \in \mathbb{C}\}$ ist offensichtlich nicht leer und nach unten beschränkt.

Wenn $|z| = R$, folgt

$$|p(z)| = \left| z^n + \sum_{k=0}^{n-1} a_k z^k \right| = R^n \left| 1 + \sum_{k=0}^{n-1} a_k z^{k-n} \right|,$$

und daher

$$|p(z)| \to +\infty \quad \text{für} \quad R \to \infty.$$

(Das ist auch Gegenstand von Lemma 10.37.) Also existiert ein $R_0 > 0$, so dass für alle $z \in \mathbb{C}$ mit $|z| > R_0$ gilt: $|p(z)| > \mu$. Es folgt:

$$\mu = \inf_{z \in B_{R_0}(0)} |p(z)|,$$

wobei

$$B_{R_0}(0) = \{w \in \mathbb{C} \mid |w| \leqslant R_0\}$$

der abgeschlossene Ball mit Radius R_0 um 0 ist. Weil p stetig ist, nimmt p nach Satz 10.45 auf $B_{R_0}(0)$ sein Infimum an. Das heißt, es existiert ein $z_0 \in B_{R_0}(0)$ mit $|p(z_0)| = \mu$.

Behauptung: $\mu = 0$.

Wenn das nicht der Fall ist, setze

$$q(z) = \frac{p(z + z_0)}{p(z_0)}.$$

Dann ist q ein nicht-konstantes Polynom mit $q(0) = 1$ und $|q(z)| \geqslant 1$ für alle $z \in \mathbb{C}$. (Die letzte Aussage folgt, weil p bei z_0 sein globales Minimum hat.)

Es gibt ein kleinstes $k \in \mathbb{N}, 1 \leqslant k \leqslant n$, so dass

$$q(z) = 1 + b_k z^k + \ldots + b_n z^n$$

mit $b_k \neq 0$. Nach Proposition 14.22 gibt es ein $\varphi \in \mathbb{R}$ mit

$$e^{ik\varphi} b_k = -|b_k|.$$

Dann folgt für alle $r > 0$:

$$\begin{aligned} |q(re^{i\varphi})| &\leqslant 1 + |b_k (re^{i\varphi})^k| + b_{k+1} r^{k+1} \ldots + b_n r^n \\ &= 1 - r^k |b_k| + b_{k+1} r^{k+1} \ldots + b_n r^n \\ &= 1 - r^k \big(|b_k| - b_{k+1} r - \ldots - b_n r^{n-k} \big) \end{aligned}$$

Aber für genügend kleines $r > 0$ ist der Ausdruck in Klammern positiv, also

$$|q(re^{i\varphi})| < 1.$$

Das ist ein Widerpsruch. □

16 Gleichmäßige Konvergenz und Differentiation, Ableitung von Potenzreihen

Das Ziel dieses Abschnitts ist zu zeigen, dass Potenzreihen innerhalb ihres Konvergenzradius unendlich oft differenzierbare Funktionen sind. Dazu benötigen wir zunächst, wie sich gleichmäßige Konvergenz und Differentiation zueinander verhalten. Man beachte, dass Differentiation und eine wie auch immer konvergierende Funktionenfolge zwei verschiedene Grenzprozesse sind, und am Ende wollen wir diese Prozesse miteinander vertauschen: die Ableitung der Grenzfunktion ist die Grenzfunktion der Ableitungen der Funktionen in der Folge. Dazu eine Warnung: **Zwei Grenzprozesse lassen sich nicht ohne weiteres vertauschen!**

Beispiel 16.1. Definiere für $m, n \in \mathbb{N}$ die Terme

$$s_{m,n} = \frac{n}{m+n}.$$

Dann gilt für alle $n \in \mathbb{N}$

$$\lim_{m\to\infty} s_{m,n} = \lim_{m\to\infty} \frac{n}{m+n} = 0,$$

also

$$\lim_{n\to\infty} \lim_{m\to\infty} s_{m,n} = \lim_{n\to\infty} 0 = 0.$$

Andererseits gilt für alle $m \in \mathbb{N}$:

$$\lim_{n\to\infty} s_{m,n} = \lim_{n\to\infty} \frac{n}{m+n} = \lim_{n\to\infty} \frac{1}{\frac{m}{n}+1} = 1,$$

also

$$\lim_{m\to\infty} \lim_{n\to\infty} s_{m,n} = \lim_{m\to\infty} 1 = 1.$$

Man muss große Sorgfalt walten lassen, wenn man zwei Grenzwerte vertauschen will. Die nächsten beiden Sätze werden das für uns bewerkstelligt.

Satz 16.2. *Sei $\left(f_n\colon [a,b] \to \mathbb{R}\right)_{n\in\mathbb{N}}$ eine Folge differenzierbarer Funktionen, so dass*

1. *ein Punkt $x_0 \in [a,b]$ existiert, an dem die Folge $\left(f_n(x_0)\right)_{n\in\mathbb{N}}$ konvergiert und*
2. *die Folge $(f'_n)_{n\in\mathbb{N}}$ gleichmäßig konvergiert.*

Dann konvergiert die Folge $(f_n)_{n\in\mathbb{N}}$ gleichmäßig auf $[a,b]$ gegen eine differenzierbare Funktion f und es gilt für alle $x \in [a,b]$

$$f'(x) = \lim_{n\to\infty} f'_n(x).$$

Korollar 16.3. *Sei $a \in \mathbb{C}$ und sei $P(z) = \sum_{k=0}^{\infty} c_k(z-a)^k$ eine Potenzreihe um a mit Konvergenzradius $r(P) > 0$ ($r(P) = +\infty$ eingeschlossen). Die Potenzreihe P ist auf $U_{r(P)}(a)$ eine unendlich oft differenzierbare Funktion und es gilt*

$$P'(z) = \sum_{k=1}^{\infty} kc_{k-1}z^{k-1}.$$

Als Slogan ausgedrückt: Potenzreihen lassen sich im Inneren der Konvergenzscheibe gliedweise ableiten.

Beweis. Die Potenzreihe P ist nach Korollar 10.61 auf $U_{r(P)}(a)$ der gleichmäßige Limes der Polynome

$$p_k(z) = \sum_{j=0}^{k} c_k(z-a)^k.$$

Das Korollar sagt auch aus, dass

$$p_k'(z) = \sum_{j=1}^{k} kc_{k-1}(z-a)^{k-1}$$

auf $U_{r(P)}(a)$ der gleichmäßige gegen die Potenzreihe

$$Q(z) = \sum_{j=1}^{\infty} kc_{k-1}(z-a)^{k-1}$$

konvergiert. Der Satz

besagt dann, dass

$$p_k'(z) = \sum_{j=1}^{k} kc_{k-1}(z-a)^{k-1}$$

gleichmäßig gegen die P' konvergiert. Also ist P auf $U_{r(P)}(a)$ differenzierbar und es gilt $P' = Q$. Wir können dasselbe Argument wieder auf Q anstelle von P anwenden und schließen □