A. AUFBAU DER ZAHLENSYSTEME

A1. Einleitung

Wir alle kennen die üblichen Zahlensysteme,

- die natürlichen Zahlen $\mathbb{N} \ni 17, 23, \ldots$,
- die ganzen Zahlen $\mathbb{Z} \ni -3, -8, \ldots$,
- die rationalen Zahlen $\mathbb{Q} \ni 1/2, -17/23, \ldots$,
- die reellen Zahlen $\mathbb{R} \ni \sqrt{3}, e, \pi, \ldots$,
- die komplexen Zahlen $\mathbb{C} \ni \sqrt{-1}, -1/2 + i\sqrt{3}/2, \ldots$.

Kann jemand aber erklären, was genau eine reelle Zahl sein sollte? Die reellen Zahlen sind ein Hauptthema dieses ersten Semesters; wir wollen damit anfangen, die genau zu untersuchen. Man kann die Eigenschaften der reellen Zahlen (dass sie einen vollständigen angeordneten Körper bilden—Definitionen später) als Axiome annehmen. Dann ist es aber nicht unbedingt intuitiv klar, ob es so einen Körper $\mathbb{R}$ gibt.

Stattdessen fangen wir hier mit den Peano'schen Axiomen für die natürlichen Zahlen an, wo es intuitiv klar sein sollte, dass die natürlichen Zahlen $\mathbb{N}$, die wir schon "kennen", genau diese Axiome erfüllen. Aus $\mathbb{N}$ konstruieren wir zunächst $\mathbb{Z}$, dann $\mathbb{Q}$, und schließlich $\mathbb{R}$. Wir können dabei zeigen, dass diese die jeweilig erwarteten Eigenschaften haben.

Wir setzen die folgenden Vorkenntnisse voraus: logische Grundbegriffen, Mengen, Abbildungen, Relationen, Äquivalenzklassen.

Ausser der Zahlensysteme selbst, betrachten wir in diesem Semester:

- Folgen und Reihen
  - Konvergenz, Häufungspunkte,
  - Cauchyfolgen, Vollständigkeit
  - absolute Konvergenz, Umordnungen
  - Dezimaldarstellung
- Funktionen und Stetigkeit
  - Grenzwerte, Monotonie, Zwischenwertsatz
  - Kompaktheit, gleichmäßige Stetigkeit
- Differentiation
  - Ableitungen, Differenzierbarkeit
  - Mittelwertsatz, Regel von l'Hôpital
  - monotone und konvexe Funktionen
  - Potenzreihen, elementare Funktionen (exp, log, sin, usw.)
  - Taylorapproximation, Taylorreihen

A2. Die natürlichen Zahlen

Man kann die natürlichen Zahlen aus reinen Mengen konstruieren, in dem man z.B. $0 := \emptyset$, $1 := \{\emptyset\} = \{0\}$, $2 := \{\emptyset, \{\emptyset\}\} = \{0, 1\}$, usw. setzt. Man braucht dazu aber Axiome für Mengentheorie, die eher unübersichtlich sind. Wir fangen lieber an mit der Annahme, dass es eine Menge $\mathbb{N}$ gibt, die den folgenden (intuitiven) Peano'schen Axiomen genügt. Diese Axiome sind eine Formalisierung der Idee des Zählens: man fängt mit 0 an; es kommt immer eine "nächste" Zahl; wir sehen eine Zahl nie wieder; wir erreichen alle natürliche Zahlen durch solches Zählen.

Der berühmte Mathematiker Leopold Kronecker sagte mal (in einem Vortrag bei der Berliner Naturforscherversammlung 1886) "Die ganzen Zahlen hat der liebe Gott gemacht, alles andere ist Menschenwerk."

a. *Die Peano'schen Axiome*

Die natürlichen Zahlen bilden eine Menge $\mathbb{N}$ mit einem ausgezeichneten Element $0 \in \mathbb{N}$ ("Null") und eine Abbildung $s : \mathbb{N} \to \mathbb{N}$ (die "Nachfolgerfunktion") mit den folgenden Eigenschaften:

1. $0 \notin s(\mathbb{N})$
2. s ist injektiv
3. Sei $A \subset \mathbb{N}$ eine Teilmenge mit $0 \in A$ und $s(A) \subset A$, dann gilt $A = \mathbb{N}$.

Bemerkung A2.1. Eine Menge X, die eine injektive aber nicht surjektive Abbildung $X \to X$ erlaubt, ist notwendigerweise unendlich groß. (Später erklären wir, was genau das heißt.)

Bemerkung A2.2. $s(0)$ nennen wir 1, $s(1)$ nennen wir 2, usw. Sobald wir Addition definieren, sehen wir, $s(n) = n+1$. Danach benutzen wir nie wieder die Notation $s(n)$.

Bemerkung A2.3. Wir definieren $\mathbb{N}^+ := \mathbb{N} \setminus \{0\} = \{1, 2, \ldots\}$, die positiven natürlichen Zahlen. (Manche Autoren nennen nur diese letzten die natürlichen Zahlen. Obwohl historisch Menschen erst spät 0 "entdeckten", natürlich ist "keine" eine gut mögliche Antwort zur Frage "Wie viele Äpfel hast du?")

b. *Verknüpfungen*

Eine *Verknüpfung* $\otimes$ auf einer Menge X ist eine Abbildung $\otimes\colon X \times X \to X$. (Wir schreiben $x \otimes y$ für $\otimes(x, y)$.) Die Verknüpfung heißt *kommutativ*, falls $x \otimes y = y \otimes x$ für alle $x, y \in X$. Sie heißt *assoziativ*, falls

$$x \otimes (y \otimes z) = (x \otimes y) \otimes z$$

für alle $x, y, z \in X$. Ein Element $e \in X$ mit $e \otimes x = x = x \otimes e$ für alle $x \in X$ heißt *Identität* (oder neutrales Element) für $\otimes$.

Bemerkung A2.4. Es gibt höchstens eine Identität. (Falls e und e' Identitäten sind, folgt $e = e \otimes e' = e'$.)

Beispiel A2.5. Addition und Multiplikation (auf $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ oder $\mathbb{R}$!) sind beide kommutativ und assoziativ. Subtraktion ist weder kommutativ ($2-1 \neq 1-2$) noch assoziativ

$(3-(2-1) = 2 \neq 0 = (3-2)-1)$. Für Addition ist 0 die Identität; für Multiplikation ist 1 die Identität.

c. *Induktionsbeweise*

Theorem A2.6. *Sei $E(n)$ eine Aussage für $n \in \mathbb{N}$. Falls*

- *$E(0)$ richtig ist (Induktionsbeginn), und*
- *$E(n) \implies E(s(n))$ für alle $n \in \mathbb{N}$ (Induktionsschritt),*

dann gilt $E(n)$ für alle $n \in \mathbb{N}$.

Beweis. Setzen wir $A := \{n \in \mathbb{N} : E(n) \text{ ist richtig}\}$. Der Induktionsbeginn sagt, $0 \in A$; der Induktionsschritt sagt, $n \in A \implies s(n) \in A$. Vom letzten Peanoaxiom folgt dann $A = \mathbb{N}$. □

Als Beispiel können wir beweisen, dass

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

(Wir definieren Addition usw. erst später, wollen aber hier schon ein Beispiel zeigen.) Für $n = 0$ laute die Aussage $0 = 0$, was richtig ist. Für den Induktionsschritt nehmen wir an, dass

$$1 + 2 + \cdots + n = \frac{(n+1)n}{2},$$

und wollen beweisen, dass

$$1 + 2 + \cdots + (n+1) = \frac{(n+2)(n+1)}{2}.$$

Wir haben

$$\begin{aligned} 1 + 2 + \cdots + (n+1) &= (1 + 2 + \cdots + n) + (n+1) \\ &= \frac{(n+1)n}{2} + (n+1) = \frac{(n+2)(n+1)}{2}. \end{aligned}$$

Eine *rekursive Definition* von f_n ist folgendes: man definiert zunächst f_0 und danach $f_{s(n)}$ mittels f_n. Dadurch ist f_n wohldefiniert für alle $n \in \mathbb{N}$.

`Ende der Vorlesung 2008 Apr 14`

A3. Strukturen auf den natürlichen Zahlen

a. *Arithmetik*

Wir wollen jetzt die Verknüpfungen $+$ und $\cdot$ auf $\mathbb{N}$ definieren.

Für $+$ definieren wir zunächst für jede $n \in \mathbb{N}$ eine Abbildung $\varphi_n : \mathbb{N} \to \mathbb{N}$ (die wir als $m \mapsto m+n$ verstehen wollen). Die rekursive Definition von φ_n lautet: $\varphi_0 := \mathrm{id}$, während $\varphi_{s(n)} := s \circ \varphi_n$ (d.h., $\varphi_{s(n)}(m) := s\big(\varphi_n(m)\big)$). Wir haben $\varphi_1 = s$, $\varphi_2 = s \circ s$, usw.

Lemma A3.1. *Für alle $n \in \mathbb{N}$ gilt $\varphi_n(0) = n$.*

Beweis. Wir benutzen Induktion. Für $n = 0$ gilt $\varphi_0(0) = 0$, weil $\varphi_0 = \mathrm{id}$. Für den Induktionsschritt nehmen wir an, $\varphi_n(0) = n$. Dann gilt $\varphi_{s(n)}(0) = s(\varphi_n(0)) = s(n)$, genau was wir zeigen mussten. □

Lemma A3.2. *Für alle $m, n \in \mathbb{N}$ gilt $\varphi_n(s(m)) = s(\varphi_n(m))$.*

Beweis. Sei $m \in \mathbb{N}$ beliebig; wir benutzen Induktion über n. Für $n = 0$ gilt $\varphi_0(s(m)) = s(m) = s(\varphi_0(m))$, weil $\varphi_0 = \mathrm{id}$. Für den Induktionsschritt haben wir, $\varphi_n(s(m)) = s(\varphi_n(m))$ und müssen zeigen, $\varphi_{s(n)}(s(m)) = s(\varphi_{s(n)}(m))$. Von der Definition haben wir aber

$$\begin{aligned} \varphi_{s(n)}(s(m)) &= s(\varphi_n(s(m))) \\ &= s(s(\varphi_n(m))) = s(\varphi_{s(n)}(m)). \end{aligned}$$ □

Korollar A3.3. *Für alle $m, n \in \mathbb{N}$ gilt $\varphi_n(m) = \varphi_m(n)$.*

Beweis. Sei $m \in \mathbb{N}$ beliebig; wir benutzen Induktion über n. Fürs Induktionsbeginn ($n = 0$) haben wir $\varphi_0(m) = m$, weil $\varphi_0 = \mathrm{id}$, und $\varphi_m(0) = m$ vom ersten Lemma.

Für den Induktionsschritt dürfen wir annehmen, dass $\varphi_n(m) = \varphi_m(n)$. Wie ist es für $s(n)$? Wir haben $\varphi_{s(n)}(m) = s(\varphi_n(m)) = s(\varphi_m(n)) = \varphi_m(s(n))$, wo wir am Ende das zweite Lemma brauchen. □

Jetzt definieren wir $m + n := \varphi_n(m)$. Das Vorangehende zeigt, diese ist eine kommutative Verknüpfung mit Identität 0. Wir haben $s(m) + n = s(m+n) = m + s(n)$ vom zweiten Lemma. (Weil $s(n) = n + 1$ ist das der Spezialfall $k = 1$ von Assoziativität.) Addition ist auch assoziativ:

Lemma A3.4. *Für alle $i, j, k \in \mathbb{N}$ gilt $(i+j)+k = i+(j+k)$.*

Beweis. Seien $i, j \in \mathbb{N}$ beliebig; wir benutzen Induktion über k. Für $k = 0$ gilt $(i+j)+0 = i+j = i+(j+0)$, weil 0 die Identität ist. Falls $(i+j)+k = i+(j+k)$ (Induktionsvoraussetzung) dann gilt

$$\begin{aligned} (i+j)+s(k) &= s((i+j)+k) = s(i+(j+k)) \\ &= i+s(j+k) = i+(j+s(k)). \end{aligned}$$ □

Wir definieren rekursiv auch die Multiplikation: $m \cdot n := \psi_n(m)$, wobei $\psi_0(m) := 0$ und $\psi_{s(n)}(m) := m + \psi_n(m)$. In ähnlicher Weise zeigt man, dass diese auch eine kommutative und assoziative Verknüpfung ist, mit 1 als Identität. Es gilt auch das Distributivgesetz:

$$(i+j) \cdot k = i \cdot k + j \cdot k.$$

Und wir haben

$$m \cdot n = 0 \iff m = 0 \text{ oder } n = 0.$$

b. *Ordnung*

Wichtig ist auch die Relation $\leq$ auf $\mathbb{N}$, definiert durch:

$$m \leq n \iff \exists d \in \mathbb{N} \text{ s.d. } m+d=n.$$

Wenn $m \leq n$, schreiben wir $n-m := d$. Diese Relation ist *reflexiv* ($m \leq m$ mit $d=0$) und *transitiv* ($i \leq j \leq k \implies i \leq k$ mit $k-i=(k-j)+(j-i)$). Weil wir auch

$$m \leq n \text{ und } n \leq m \implies m=n$$

haben, ist sie eine *partielle Ordnung*. Weil zusätzlich

$$m \leq n \text{ oder } n \leq m$$

für alle $m,n \in \mathbb{N}$, ist sie sogar eine *totale Ordnung*.

Wir schreiben auch $n \geq m$ für $m \leq n$ und schreiben $m<n$ falls $m \leq n$ und $m \neq n$. Wir haben $0 \leq n$ für alle $n \in \mathbb{N}$. Für $i,j \in \mathbb{N}$, $k \in \mathbb{N}^+$ gilt

$$i \cdot k \leq j \cdot k \iff i \leq j \iff i+k \leq j+k.$$

Zum Schluss bemerken wir, dass es keine $m,n \in \mathbb{N}$ mit $m<n<m+1$ gibt.

c. *Wohlordnung*

Eine Teilmenge $A \subset \mathbb{N}$ hat ein *Minimum* a_0, falls $a_0 \in A$ und $a_0 \leq a$ für alle $a \in A$.

Satz A3.5. *Die natürlichen Zahlen* $\mathbb{N}$ *sind* wohlgeordnet*, d.h., jede nichtleere Teilmenge besitzt ein Minimum.*

Beweis. Sei $A \subset \mathbb{N}$ eine Teilmenge ohne ein Minimum; wir zeigen $A=\emptyset$. Wir setzen

$$B := \{n \in \mathbb{N} : n \leq a \text{ für alle } a \in A\}.$$

Offensichtlich gilt $0 \in B$. Wir haben auch $A \cap B = \emptyset$, weil $n \in A \cap B$ ein Minimum für A wäre. Es sei nun $n \in B$. Wir haben $a \geq n$ für alle $a \in A$; weil $n \notin A$ gilt in der Tat $a>n$ und deshalb $a \geq n+1$. D.h., $n+1=s(n) \in B$. Vom Peano'schen Axiom folgt $B=\mathbb{N}$ und deshalb $A \subset \mathbb{N} \smallsetminus B = \emptyset$. □

Dieser Satz ist äquivalent zum 3. Axiom.

`Ende der Vorlesung 2008 Apr 15`

A4. Die ganzen Zahlen

Aus den natürlichen Zahlen konstruieren wir die ganzen und die rationalen Zahlen, um Subtraktion bzw. Division durchführen zu können. Das heißt, wir wollen inverse Elemente (bezüglich Addition bzw. Multiplikation) finden.

a. *Gruppen*

Definition A4.1. Sei $\otimes$ eine assoziative Verknüpfung auf einer Menge G mit Identität $e \in G$. Ein *Inverses* zu $g \in G$ bedeutet ein Element $\bar{g} \in G$ mit $g \otimes \bar{g} = e = \bar{g} \otimes g$. Dann heißt $(G,\otimes)$ eine *Gruppe*, falls zu jedem $g \in G$ es ein Inverses gibt. Eine Gruppe $(G,\otimes)$ heißt *kommutativ* (oder auch *Abel'sch*), falls $\otimes$ kommutativ ist.

Beispiel A4.2. Die folgenden sind alle Gruppen: $(\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{Q}^*,\cdot)$, $(\mathbb{Q}^+,\cdot)$, $(\mathbb{R},+)$, $(\mathbb{R}^*,\cdot)$, $(\mathbb{R}^+,\cdot)$, $(\mathbb{C},+)$, $(\mathbb{C}^*,\cdot)$, $(M^{2\times 2}\mathbb{R},+)$, $(GL_2\mathbb{R},\cdot)$. (Hier bedeutet * die Elemente, die nicht Null sind, und $^+$ die positiven. GL_2 ist die Menge aller 2×2 Matrizen, die ein Inverses haben.)

Die natürlichen Zahlen bilden keine Gruppe $(\mathbb{N},+)$, weil nur 0 ein Inverses hat. Wir können dieses Problem beheben, in dem wir zusätzliche Elemente -1, -2, usw. zufügen. Die genaue Definition erfolgt am einfachsten mithilfe sogenannter Äquivalenzklassen.

b. *Äquivalenzklassen*

Eine Relation $\sim$ auf einer Menge X heißt Äquivalenzrelation, falls sie

- reflexiv ($x \sim x$ für alle $x \in X$),
- transitiv ($x \sim y$ und $y \sim z \implies x \sim z$), und
- symmetrisch ($x \sim y \iff y \sim x$)

ist. Dann für jedes $x \in X$ heißt die Menge

$$[x] := \{y \in X : y \sim x\}$$

die Äquivalenzklasse von x. Die Menge $X/\!\sim$ aller Äquivalenzklassen bezüglich $\sim$ ist eine Zerlegung von X.

Beispiel A4.3 *(Ganze Zahlen modulo n).* Auf $\mathbb{N}$ definieren wir eine Äquivalenzrelation $\sim_n$ durch

$$i \sim_n j \iff \exists d \in N \text{ s.d. } i-j=dn \text{ oder } j-i=dn.$$

Es gibt genau n Äquivalenklassen:

$$[0]=[n], [1], [2], \dots, [n-1].$$

Man schreibt $\mathbb{Z}/n := \mathbb{N}/\!\sim_n$. Addition ist wohldefiniert und $(\mathbb{Z}/n,+)$ ist eine kommutative Gruppe mit n Elementen.

c. *Definition der ganzen Zahlen*

Aus den natürlichen Zahlen $\mathbb{N}$ wollen wir eine Gruppe $(\mathbb{Z},+)$ bauen, wo wir auch Subtraktion durchfürhen können. Die Idee ist, $m-n$ durch das Paar $(m,n) \in \mathbb{N}\times\mathbb{N}$ zu repräsentieren. Offensichtlich ist das nicht eindeutig: z.B. wollen wir, dass $17-23=1-7$.

Auf die Menge $\mathbb{N}\times\mathbb{N}$ aller geordneten Paare (m,n) natürlicher Zahlen definieren wir eine Relation:

$$(m,n) \sim (m',n') \iff m+n'=m'+n.$$

(Gern würden wir schreiben $\iff m-n=m'-n'$.) Diese Relation ist offensichtlich reflexiv und symmetrisch; sie ist auch transitiv und deshalb eine Äquivalenzrelation. Die *ganzen Zahlen* sind dann $\mathbb{Z} := (\mathbb{N}\times\mathbb{N})/\sim$; wir schreiben $[m,n] := [(m,n)]$. Als Beispiel,

$$[17,23] = \{(0,6),(1,7),(2,8),\dots\}.$$

Auf $\mathbb{N}\times\mathbb{N}$ definiert man $+$ komponentenweise:

$$(m,n)+(i,j) := (m+i,n+j).$$

Diese ist relationentreu im Sinne, dass

$$(m,n)\sim(m',n'),(i,j)\sim(i',j') \implies (m+i,n+j)\sim(m'+i',n'+j').$$

Das bedeutet, wir können $+$ auf $\mathbb{Z}$ definieren:

$$[m,n]+[i,j] := [m+i,n+j].$$

Das additive Inverse zu $[m,n]$ ist $[n,m]$. Dadurch wird $(\mathbb{Z},+)$ eine kommutative Gruppe.

Die bekannte Berechnung

$$(m-n)(i-j) = (mi+nj)-(ni+mj)$$

deutet schon die Definition

$$[m,n]\cdot[i,j] := [mi+nj, ni+mj]$$

an.

Auf $\mathbb{Z}$ definiert man $\le$ wie folgt:

$$[m,n]\le[m',n'] \iff m+n'\le m'+n.$$

Wir haben

$$\{x\in\mathbb{Z} : x\ge 0\} = \{[n,0] : n\in\mathbb{N}\}\cong\mathbb{N}.$$

Dadurch identifizieren wir die natürlichen Zahlen mit den nichtnegativen ganzen Zahlen. Ab jetzt schreiben wir die ganzen Zahlen auf normale Weise: $n=[n,0]$, $-n=[0,n]$.

A5. Die rationalen Zahlen

Um die rationalen Zahlen $\mathbb{Q}$ zu konstruieren, müssen wir Addition und Multiplikation zusammen betrachten.

a. *Kommutative Ringe*

Ein *Ring* $(R,\oplus,\otimes)$ besteht aus einer kommutativen Gruppe $(R,\oplus)$ (mit Identität $0\in R$) und einer assoziativen Verknüpfung $\otimes$ (mit Identität $1\in R$), wobei das Distributivgesetz

$$(a\oplus b)\otimes c = (a\otimes c)\oplus(b\otimes c),\quad c\otimes(a\oplus b) = (c\otimes a)\oplus(c\otimes b)$$

für alle $a,b,c\in R$ gilt.

Bemerkung A5.1. Manche Autoren erlauben, dass $\otimes$ keine Indentität 1 hat. (Dazu würden wir "Rng" für Ring ohne Identität schreiben.)

Bemerkung A5.2. Die "Addition" $\oplus$ muss kommutativ sein. Falls die "Multiplikation" $\otimes$ auch kommutativ ist, heißt $(R,\oplus,\otimes)$ ein *kommutativer Ring.*

Der Beweis, dass $(\mathbb{Z},+,\cdot)$ ein kommutativer Ring ist, ist zwar lang(weilig), beinhaltet aber keine Verwicklungen oder Schwierigkeiten.

Bemerkung A5.3. In einem Ring $(R,+,\cdot)$ schreiben wir $-a$ für die additive Inverse zu $a\in R$, $a-b := a+(-b)$ für die Differenz und $ab := a\cdot b$ für den Produkt zweier Elemente.

Bemerkung A5.4. Es gelten $0a=0$, $-0=0$, $(-1)a=-a$, $(-a)b=-(ab)$, usw.

Bemerkung A5.5. Für $n\in\mathbb{N}$ und $a\in R$ definieren wir $na, a^n\in R$ rekursiv:

$$0a=0,\quad 1a=1,\quad 2a=a+a,\quad\dots,\quad (n+1)a=na+a,$$
$$a^0=1,\quad a^1=a,\quad a^2=a\cdot a,\quad\dots,\qquad a^{n+1}=a^n\cdot a.$$

b. *Angeordnete Ringe*

Sei $(R,+,\cdot)$ ein Ring und $\le$ eine totale Ordnung auf R. Dann heißt R ein angeordneter Ring, wenn gelten:

- $x,y>0 \implies xy>0$ für alle $x,y\in R$.
- $x\le y \implies x+z\le y+z$ für alle $x,y,z\in R$.

Solch ein angeordneter Ring ist $\mathbb{Z}$.

Sei R ein Ring. Falls es $a,b\ne 0\in R$ gibt mit $ab=0$, dann heißen a und b *Nullteiler.* Ein angeordneter Ring hat keine Nullteiler.

`Ende der Vorlesung 2008 Apr 21`

c. *Körper*

Sei $(R,+,\cdot)$ ein kommutativer Ring. Wegen $0a=0$ hat 0 kein multiplikatives Inverses. Wir schreiben $R^* := R\smallsetminus\{0\}$. Falls jedes $a\in R^*$ ein multiplikatives Inverses hat, ist $(R^*,\cdot)$ eine kommutative Gruppe. In diesem Fall nennen wir R einen *Körper.*

Beispiel A5.6. Für eine Primzahl $p\in\mathbb{N}$ ist $(\mathbb{Z}/p,+,\cdot)$ ein Körper.

Im Ring $\mathbb{Z}$ haben nur ±1 Inverse; $\mathbb{Z}$ ist kein Körper. Wir können aber $\mathbb{Z}$ auf $\mathbb{Q}$ erweitern, den Körper der rationalen Zahlen. Die Konstruktion von $\mathbb{Q}$ aus $\mathbb{Z}$ und die von $\mathbb{Z}$ aus $\mathbb{N}$ sind sehr ähnlich.

Auf $\mathbb{Z}\times\mathbb{Z}^*$ definieren wir eine Äquivalenzrelation

$$(p,q)\sim(p',q') \iff pq'=p'q,$$

weil wir (p,q) als ${}^p/_q\in\mathbb{Q}$ verstehen wollen. Wir definieren $\mathbb{Q} := (\mathbb{Z}\times\mathbb{Z}^*)/\sim$ und schreiben ${}^p/_q := [(p,q)]$. Die Definitionen der Verknüpfungen sollten keine Überraschung sein:

$$p/_q + p'/_{q'} := (pq' + p'q)/_{qq'}, \quad p/_q \cdot p'/_{q'} := pp'/_{qq'}.$$

Jedes $p/_q \in \mathbb{Q}$ hat ein additives Inverses und zwar $-p/_q$. Jedes $p/_q \neq 0/_1 \in \mathbb{Q}$ hat als multiplikatives Inverses $q/_p$. Wir werden nicht im Detail checken, dass dadurch $\mathbb{Q}$ ein Körper wird.

Die (eindeutige) *minimale* (oder *teilerfremde*) *Darstellung* von $p/_q \in \mathbb{Q}$ ist $p_0/_{q_0}$ mit $q_0 \in \mathbb{N}$ so klein wie möglich gewählt. (Vom Wohlordnungsprinzip hat die Menge aller möglichen q ein Minimum.)

Auf $\mathbb{Q}$ haben wir eine totale Ordnung: für $q, q' > 0$ definieren wir

$$p/_q \leq p'/_{q'} \iff pq' \leq qp' \in \mathbb{Z}.$$

Dadurch wird $\mathbb{Q}$ ein angeordneter Körper. Das andere wichtigste Beispiel bilden die reellen Zahlen $\mathbb{R}$.

d. *Zusammenfassung*

Ein angeordneter Körper $(K, +, \cdot, \leq)$ ist durch folgenden Eigenschaften definiert:

1. $(K, +)$ ist eine kommutative Gruppe:
   - assoziativ: $a + (b + c) = (a + b) + c$,
   - kommutativ: $a + b = b + a$,
   - Identität und Inverse: $a + (-a) = 0$.
2. $(K^*, \cdot)$ ist eine kommutative Gruppe:
   - assoziativ: $a(bc) = (ab)c$,
   - kommutativ: $ab = ba$,
   - Identität und Inverse: $aa^{-1} = 1$.
3. $\leq$ ist ein totale Ordnung:
   - reflexiv: $a \leq a$,
   - transitiv: $a \leq b \leq c \implies a \leq c$,
   - Trichotomie: $a \leq b$ oder $b \leq a$, wobei beide nur im Falle $a = b$ glechzeitig gelten.
4. $\cdot$ ist über $+$ distributiv: $a(b + c) = ab + ac$.
5. $a > b \iff a - b > 0$.
6. $a, b > 0 \implies ab > 0$.

Insgesamt heißt das etwa, in jedem angeordneten Körper können wir Arithmetik genau so durchführen, wie wir sie aus der Schule kennen. Zum Beispiel, definieren wir $|a| \geq 0$ als

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

Um die reellen Zahlen zu konstruieren bzw. definieren, benutzen wir Konvergenz von Folgen, was mithilfe des Betrags $|a|$ definiert wird.

Weitere Eigenschaften eines angeordenten Körpers, die man leicht aus den Oberen beweisen kann, sind z.B.:

1. $a \leq 0, x \leq y \implies ay \leq ax$,
2. $x^2 \geq 0$, insbesondere $1 > 0$,
3. $x < 0 \implies x^{-1} < 0, -x > 0$,
4. $|xy| = |x||y|; \quad |x + y| \leq |x| + |y|$,
5. $|x - a| < \varepsilon \iff a - \varepsilon < x < a + \varepsilon$.

A6. Irrationale Zahlen

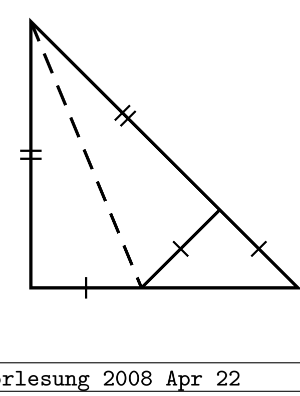
Betrachten wir die Gleichung $x^2 = a$ in einem angeordeten Körper K. Falls $a < 0$ gibt es keine Lösungen (weil $x^2 \geq 0$). Für $a = 0$ hat die Gleichung $x^2 = 0$ genau die eine Lösung $x = 0$ (weil es keine Nullteiler gibt). Für $a > 0$ gibt es entweder keine oder (mindestens) zwei Lösungen $\pm b$. (Offensichtlich ist $(-b)^2 = b^2$.)

Bemerkung A6.1. Im Allgemeinen in einem Körper K hat ein Polynom von Grad d höchstens d Nullstellen. Das heißt hier, $x^2 - a = 0$ hat höchstens zwei (und deshalb entweder keine oder genau zwei) Lösungen. Falls es die gibt, nennen wir die positive Lösung $\sqrt{a}$.

Satz A6.2. *Die Gleichung $x^2 = 2$ hat in $\mathbb{Q}$ keine Lösung.*

Algebraischer Beweis. Sei $p/_q$ die teilerfremde Darstellung einer Lösung. Dann gilt $p^2 = 2q^2$. Demzufolge ist p^2 (und deswegen auch p) gerade. Das heißt, wir können $p =: 2m$ schreiben. Dann haben wir aber $4m^2 = (2m)^2 = 2q^2$, also $2m^2 = q^2$. Demzufolge ist q^2 (und deswegen auch q) gerade. Das bedeutet aber, p und q sind nicht teilerfremd sondern haben 2 als gemeinsamen Teiler. □

Geometrischer Beweis. Sei T ein gleichschenkliges rechtwinkliges Dreieck mit ganzzahligen Seitenlängen. Dann können wir ein Kleineres konstruieren mit denselben Eigenschaften. Es kann aber (vom Wohlordnungsprinzip) nicht immer weiter so gehen! □



`Ende der Vorlesung 2008 Apr 22`

B. FOLGEN UND DIE REELLEN ZAHLEN

B1. Endliche Mengen

Wir sagen, zwei Mengen X und Y sind *gleich groß* (oder *gleichmächtig*), falls es eine Bijektion (eine injektive und surjektive Abbildung) $f\colon X \to Y$ gibt. Als Beispiel einer Menge mit n Elementen ($n \in \mathbb{N}$) nehmen wir

$$[n] := \{m \in \mathbb{N} : m < n\} = \{0, 1, \ldots, n-1\}.$$

Falls X und $[n]$ gleich groß sind, sagen wir, X ist *endlich* und $\#X = n$.

Definition B1.1. Sei $E(n)$ eine Aussage für $n \in \mathbb{N}$. Wir sagen, $E(n)$ gilt für *fast alle* $n \in \mathbb{N}$, falls es nur endlich viele Ausnahmen gibt, das heißt, falls die Menge $\{n \in \mathbb{N} : E(n) \text{ ist falsch}\}$ endlich ist.

Lemma B1.2. *Eine Teilmenge $X \subset \mathbb{N}$ hat eine obere Schranke genau dann, wenn sie endlich ist.*

Beweis. Falls n eine obere Schranke für X ist, dann ist X eine Teilmenge von $[n+1]$ und dadurch endlich. Umgekehrt benutzen wir Induktion über $\#X$. Im Fall $\#X = 0$ haben wir $X = \emptyset$. Dann gilt jedes $n \in \mathbb{N}$ als obere Schranke. Im Fall $\#X = n+1$ können wir $X = Y \cup \{a\}$ schreiben mit $\#Y = n$. Von der Induktionsvoraussetzung hat Y eine obere Schranke b. Dann ist $\max(a, b)$ eine obere Schranke für X. □

Satz B1.3. *Die Aussage $E(n)$ gilt für fast alle n genau dann, wenn es ein $n_0 \in \mathbb{N}$ so gibt, dass $E(n)$ für alle $n \geq n_0$ gilt.*

Beweis. Setzen wir $X := \{n : E(n) \text{ ist falsch}\}$. Dann folgt der Satz unmittelbar aus dem Lemma. □

Lemma B1.4. *Falls $A(n)$ und $B(n)$ Aussagen sind, die jeweils für fast alle n gelten, dann gilt auch die Aussage "$A(n)$ und $B(n)$" für fast alle n.*

Beweis. Es gibt n_A und n_B so, dass $A(n)$ für $n \geq n_A$ gilt und $B(n)$ für $n \geq n_B$ gilt. Dann gelten Beide für alle $n \geq \max(n_A, n_B)$. □

B2. Folgen und Monotonie

Sei X eine Menge. Eine *Folge* in X ist eine Abbildung $\mathbb{N} \to X$, wobei wir normalerweise den Wert in n nicht etwa $a(n)$ sondern a_n schreiben. Für die Folge selbst, schreiben wir dann

$$(a_n) = (a_n)_{n \in \mathbb{N}} = (a_0, a_1, a_2, \ldots).$$

Bemerkung B2.1. Es ist unwichtig, ob wir mit 0 oder 1 (oder einer beliebigen Zahl $n_0 \in \mathbb{N}$) anfangen. Zum Beispiel haben wir

$$(a_{n+1})_{n \in \mathbb{N}} = (a_1, a_2, \ldots) = (a_n)_{n \in \mathbb{N}^+} = (a_n)_{n \geq 1}.$$

Was uns interessiert, sind nicht die ersten Glieder einer Folge, sondern die unendlich Vielen die "später" kommen. (Der berühmte Mathematiker Hermann Weyl hat geschrieben, Mathematik "ist die Wissenschaft vom Unendlichen". In seinem Analysislehrbuch drückt es Erhard Behrends so aus: "Jugendsünden einer Folge sind ... unerheblich.")

Wird X durch $\leq$ total geordnet, können wir *monotone* und *beschränkte* Folgen betrachten.

Definition B2.2. Die Folge (a_n) heißt dann *beschränkt*, wenn es $x, y \in K$ gibt mit $x \leq a_n \leq y$ für (fast) alle n. Die Folge (a_n) heißt dann

- *monoton steigend*, wenn $a_{n+1} \geq a_n$;
- *monoton fallend*, wenn $a_{n+1} \leq a_n$;
- *streng monoton steigend*, wenn $a_{n+1} > a_n$;
- *streng monoton fallend*, wenn $a_{n+1} < a_n$,

jeweils für alle $n \in \mathbb{N}$. Die Folge heißt *monoton*, falls sie monoton steigend oder monoton fallend ist. (Oder natürlich auch beide gleichzeitig: für mathematiker ist "oder" immer *inklusiv*.) *Streng monoton* wird ähnlich definiert. Weil die Ordnung transitiv ist, folgt z.B. bei einer monoton steigenden Folge, dass $m \leq n \implies a_m \leq a_n$.

Betrachten wir jetzt Folgen in $\mathbb{N}$:

- $(0, 1, 1, 2, 3, 5, 8, \ldots)$ ist monoton steigend.
- $(1, 3, 5, 7, 9, 11, \ldots)$ ist streng monoton steigend.
- $(5, 2, 1, 0, 0, 0, 0, \ldots)$ ist monoton fallend.
- Es gibt keine streng monoton fallende Folge in $\mathbb{N}$.

Eine streng monoton steigende Folge in $\mathbb{N}$ ist nie beschränkt. (Wir haben sogar $a_n \geq n$.) Eine monoton fallende Folge in $\mathbb{N}$ ist immer beschränkt.

Seien $k = (k_n)$ eine streng monoton steigende Folge in $\mathbb{N}$ und $a = (a_k)$ eine beliebige Folge in einer beliebegen Menge X. Definieren wir eine Folge $b = (b_n)$ in X durch $b_n := a_{k_n}$, dann heißt b eine *Teilfolge* von a. Die Glieder von b sind Glieder von a und tauchen in der selben Reihenfolge auf. Falls a (streng) monoton ist, dann ist auch jede Teilfolge (streng) monoton.

B3. Konvergenz

Sei K jetzt ein angeordneter Körper. (Später im Semester haben wir fast immer $K = \mathbb{R}$. Um $\mathbb{R}$ konstruieren zu können, nehmen wir $K = \mathbb{Q}$.)

Definition B3.1. Die Folge (a_n) *konvergiert* gegen $x \in K$ (geschrieben $a_n \to x$), falls für jede beliebig kleine Toleranz $\varepsilon > 0 \in K$ fast alle a_n nah (innerhalb der Toleranz) an x sind. Das heißt, $x - \varepsilon < a_n < x + \varepsilon$ (oder äquivalent $|a_n - x| < \varepsilon$) für fast alle n.

Bemerkung B3.2. Wenn ε kleiner wird, gibt es natürlich mehrere Ausnahmen aber immer nur endlich viele.

Bemerkung B3.3. Äquivalent kann man die Definition wie folgt formulieren: $a_n \to x$ genau dann, wenn

$$\forall \varepsilon > 0 \quad \exists n_0 \quad \forall n \geq n_0 \quad |a_n - x| < \varepsilon.$$

Lemma B3.4. *Falls $a_n \to x$ und $a_n \to y$, dann folgt $x = y$.*

Beweis. Sonst setzen wir $\varepsilon := \frac{1}{2}|x-y| > 0$. Die Definition von Konvergenz besagt, $|x - a_n| < \varepsilon$ für fast alle n und $|y - a_n| < \varepsilon$ für fast alle n. Beide können aber nie gleichzeitig (für dasselbe n) gelten, weil $|x - a_n| + |y - a_n| \geq |x - y| = 2\varepsilon$. □

Definition B3.5. Falls (a_n) konvergiert, $a_n \to x$, dann heißt x der (vom Lemma eindeutig bestimmte) *Grenzwert* oder *Limes* von (a_n). Wir schreiben dann

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

Eine Folge, die gegen kein x konvergiert, heißt *divergent.*

Beispiel B3.6.

1. Falls $a_n = c$ für fast alle n, dann konvergiert a_n gegen c. (Fast alle glieder sind nicht nur nah an c sondern sind genau gleich c.)
2. $(1, 1, 0, 0, 1, 0, 0, 0, 0, 1, \ldots)$ ($a_n = 1$ falls $n = m^2$, sonst $a_n = 0$) ist nicht konvergent. (Nehmen wir $\varepsilon = 1/2$. Für jede n_0 gibt es $n \geq n_0$ mit $a_n = 1$ und $n' \geq n_0$ mit $a_{n'} = 0$. Dann für beliebiges x gilt $|a_n - x| + |a_{n'} - x| \geq |a_n - a_{n'}| = 1$.)
3. $a_n = (-1)^n$: $(+1, -1, +1, -1, \ldots)$ ist nicht konvergent. (Ähnlich.)
4. $a_n = n$: $(0, 1, 2, \ldots)$ ist nicht konvergent. (Nehmen wir an, $a_n \to x$, und setzten wir $\varepsilon = 1/2$. Dann existiert ein n_0 so, dass $|a_n - x| < \varepsilon$ für $n \geq n_0$. Dann gilt z.B.

$$\begin{aligned} 2\varepsilon &> |a_n - x| + |a_{n+1} - x| \\ &\geq |a_{n+1} - a_n| = |n + 1 - n| = 1, \end{aligned}$$

was die Wahl von ε widerspricht.

Gibt es ein "interessantes" Beispiel einer konvergenten Folge? Betrachten wir die Folge $a_n := 1/n$. Sicherlich konvergiert sie gegen 0, oder? Das stimmt in $\mathbb{Q}$ (oder in $\mathbb{R}$) aber nicht in jedem angeordneten Körper.

`Ende der Vorlesung 2008 Apr 28`

B4. Das Prinzip von Archimedes

Versuchen wir, einen Beweis zu finden, dass $1/n \to 0$. Für jedes $\varepsilon > 0$, nehmen wir einfach $n_0 > 1/\varepsilon$. Dann für $n \geq n_0$ gilt $1/n < \varepsilon$, d.h., $|a_n| < \varepsilon$. Woher aber kommt diese Zahl $n_0 > 1/\varepsilon$?

Wir errinern uns, jeder angeordneter Körper K hat eine Teilmenge isomorph zu $\mathbb{N}$ (nämlich: $\{n \cdot 1 : n \in \mathbb{N}\}$. Diese nennen wir einfach $\mathbb{N}$.

Definition B4.1. Ein angeordneter Körper K heißt *archimedisch* falls

$$\forall x \in K \quad \exists n \in \mathbb{N} \quad n > x.$$

Lemma B4.2. *Wir haben $1/n \to 0$ in K genau dann, wenn K archimedisch ist.*

Beweis.

$$\begin{aligned} 1/n \to 0 &\iff \forall \varepsilon > 0 \quad \exists n \quad 1/n < \varepsilon \\ &\iff \forall \varepsilon > 0 \quad \exists n > 1/\varepsilon \\ &\iff \forall x > 0 \quad \exists n > x. \end{aligned}$$ □

Lemma B4.3. *Der Körper $\mathbb{Q}$ ist archimedisch.*

Beweis. Sei $x \in \mathbb{Q}$. Falls $x < 0$, gibt es nichts zu tun. Sonst hat $x \geq 0$ eine Darstellung $x = n/m$, $n \in \mathbb{N}$, $m \in \mathbb{N}^+$. Weil $m \geq 1$, haben wir $n = mx \geq x$ und deshalb $n + 1 > x$. □

Bemerkung B4.4. Für einen angeordneten Körper K sind folgende äquivalent:

1. K ist archimedisch,
2. $\forall x > 0 \quad \exists n \in \mathbb{N} \quad x > 1/n$,
3. $\forall x, y > 0 \quad \exists n \in \mathbb{N} \quad nx > y$.

Diese letzte ist die Variante, die Archimedes benutzt hat um Längen vergleichen zu können.

B5. Rechenregeln für Grenzwerte

Bemerkung B5.1. Falls $a_n \to x$, dann konvergiert auch jede Teilfolge gegen x.

Satz B5.2. *Seien (a_n) und (b_n) konvergente Folgen mit $a_n \to x$ und $b_n \to y$. Dann gelten $a_n + b_n \to x + y$ und $a_n b_n \to xy$. Falls $y \neq 0$, dann gilt $a_n/b_n \to x/y$.*

Beweis. Wir beweisen nur $a_n/b_n \to x/y$, wobei wir die folgende Bemerkung machen: Möglicherweise ist $b_n = 0$ – aber nur für endlich viele n. In diesem Fall sind ein paar Glieder am Anfang der Folge $\left(a_n/b_n\right)$ gar nicht definiert. Das stört uns aber nicht, wenn wir konvergenz betrachten.

Zu jedem $\varepsilon > 0$ gibt es $n_0 \in \mathbb{N}$, so dass

$$|a_n - x| < \varepsilon, \quad |b_n - y| < \min(\varepsilon, |y|/2)$$

für alle $n \geq n_0$. Die Dreiecksungleichung ergibt $b_n \geq |y|/2$. Dann folgt:

$$\begin{aligned} \left|\frac{a_n}{b_n} - \frac{x}{y}\right| &= \left|\frac{a_n y - b_n x}{b_n y}\right| = \frac{\left|(a_n - x)y + x(y - b_n)\right|}{|b_n y|} \\ &\leq \frac{|a_n - x||y| + |x||y - b_n|}{|y|^2/2} \leq 2\frac{|x| + |y|}{|y|^2}\varepsilon. \end{aligned}$$ □

Lemma B5.3. *Seien (a_n) und (b_n) Folgen. Falls a_n beschränkt ist und $b_n \to 0$, dann gilt $a_n b_n \to 0$.* □

Lemma B5.4. *Seien (a_n) und (b_n) Folgen mit $a_n \to x$ und $b_n \to y$. Falls $a_n \le b_n$ für fast alle n, dann folgt $x \le y$.*

Beweis. Sei $\varepsilon > 0$ gegeben. Für fast alle n gilt

$$x - \varepsilon < a_n \le b_n < y + \varepsilon.$$

Deshalb haben wir $x < y + 2\varepsilon$ und zwar für jedes $\varepsilon > 0$. Daher ist $x \le y$. □

Bemerkung B5.5. Falls $a_n < b_n$ gilt natürlich $x \le y$ aber nicht unbedingt $x < y$.

B6. Cauchyfolgen

a. *Definition*

In einer konvergenten Folge $a_n \to x$ sind fast alle Glieder nah an x. Das bedeutet, dass sie auch nah bei einander sind. Der Mathematiker Augustin Cauchy hat diese Idee konkret gemacht.

Definition B6.1. Eine Folge (a_n) in einem angeordneten Körper heißt *Cauchyfolge*, wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt, so dass ab diesem Index alle Folgenglieder weniger als ε voneinander entfernt sind, d.h.:

$$\forall \varepsilon > 0 \quad \exists n_0 \in N \quad \forall n, m \ge n_0 \quad |a_m - a_n| < \varepsilon.$$

Lemma B6.2. *Eine Folge (a_n) ist genau dann eine Cauchyfolge, wenn es zu jedem $\varepsilon > 0$ ein $x \in K$ gibt, so dass $|x - a_n| < \varepsilon$ für fast alle n.*

Beweis. Falls $|x - a_n| < \varepsilon/2$ für fast alle n, dann gilt

$$|a_m - a_n| \le |a_m - x| + |a_n - x| < \varepsilon$$

für fast alle n und m. Umgekehrt, seien (a_n) eine Cauchyfolge und $\varepsilon > 0$. Es gibt $n_0 \in \mathbb{N}$, so dass $|a_m - a_n| < \varepsilon$ für alle $n, m \ge n_0$. Dann setzen wir einfach $x := a_{n_0}$. □

Korollar B6.3. *Jede Cauchyfolge ist beschränkt.* □

Korollar B6.4. *Jede konvergente Folge ist eine Cauchyfolge.*

Beweis. Falls $a_n \to x$, können wir im Lemma zu jedem $\varepsilon > 0$ dieses x benutzen. □

Satz B6.5. *Hat eine Cauchyfolge eine konvergente Teilfolge, so ist sie selbst konvergent.* □

Bisher, um zu zeigen, dass eine Folge divergent ist, haben wir in der Tat gezeigt, dass sie keine Cauchyfolge ist. Die Folge $a_n = n$ z.B. ist nicht beschränkt und deshalb keine Cauchyfolge. Die Glieder von $\big((-1)^n\big)$ z.B. bleiben zu weit von einander entfernt.

Konvergiert aber jede Cauchyfolge? Nicht unbedingt. Wir werden ein Beispiel geben, eine Folge in $\mathbb{Q}$, die gern gegen $\sqrt{2}$ konvergieren möchte, kann aber nicht, weil es keine $\sqrt{2}$ in $\mathbb{Q}$ gibt. Um zu zeigen, dass unser Beispiel doch eine Cauchyfolge ist, beweisen wir zunächst einen wichtigen Satz über Folgen, die monoton und beschränkt sind.

Lemma B6.6. *Sei (a_n) eine monoton steigende Folge in K. Seien $b < c$ Schranken mit Abstand $h := c - b$, so dass fast alle a_n dazwischen liegen, d.h., so dass $b \le a_n \le c$ für fast alle n. Dann gibt es auch Schranken $b' < c'$ mit Abstand $c' - b' = h/2$, so dass*

$$b' \le a_n \le c' \text{ für fast alle } n.$$

Beweis. Einer Umformulierung hilft sehr: wegen der Monotonie haben wir genau dann

$$x \le a_n \le y \text{ für fast alle } n,$$

wenn erstens $a_n \le y$ für alle n und zweitens es ein n gibt mit $a_n \ge x$.

Hier also existiert n mit $a_n \ge b$ und gilt $a_n \le c$ für alle n. Wir setzen $x := \frac{1}{2}(b + c)$ und betrachten zwei Fälle: Falls es n gibt mit $a_n \ge x$, setzen wir $b' := x$, $c' := c$. Sonst gilt $a_n \le x$ für alle n und wir setzen $b' := b$, $c' := x$. □

Satz B6.7. *Sei K ein Körper, der archimedisch angeordnet ist. Jede Folge in K, die beschränkt und monoton ist, ist eine Cauchyfolge.*

Beweis. Sei (a_n) (ohne Beschränkung der Allgemeinheit) eine monoton steigende Folge mit der Schranke $a_n \le C$. Dann liegen alle a_n zwischen a_0 und C. Sei $\varepsilon > 0$ gegeben. Weil K archimedisch ist, gibt es eine ganze Zahl m, so dass $(C - a_0)/\varepsilon < m < 2^m$. Wir wenden das Lemma m-mal an; am Ende liegen fast alle a_n zwischen Schranken mit Abstand kleiner als ε. Weil ε beliebig war, ist (a_n) eine Cauchyfolge. □

`Ende der Vorlesung 2008 Apr 29`

---

b. *Eine divergente Cauchyfolge*

Wir definieren rekursiv eine Folge $a_n \in \mathbb{Q}$:

$$a_0 := 2, \quad a_{n+1} := \tfrac{1}{2}\big(a_n + 2/a_n\big).$$

Wir haben dann z.B.

$$a_1 = 3/2 \approx 1.5,$$
$$a_2 = 17/12 \approx 1.4167,$$
$$a_3 = 577/408 \approx 1.4142.$$

Diese Folge ist offensichtlich positiv. Desweitern gilt

$$a_{n+1}^2 = \left(\frac{a_n}{2} + \frac{1}{a_n}\right)^2 = 2 + \left(\frac{a_n}{2} - \frac{1}{a_n}\right)^2 \ge 2.$$

Es folgt dann, dass a_n monoton fallend ist, weil $^{a_{n+1}}/_{a_n} = {}^1/_2 + {}^1/_{a_n^2}$. Deshalb ist (a_n) eine Cauchyfolge, weil (a_n) beschränkt und monoton ist und $\mathbb{Q}$ archimedisch ist.

Nehmen wir jetzt an, $a_n \to x$. Von den Rechnenregeln konvergiert auch $\left(a_n + {}^2/_{a_n}\right)$ und zwar gegen $x + {}^2/_x$. Diese Folge ist aber gleich $(2a_{n+1})$, die gegen $2x$ konvergiert. Das heißt, $2x = x + {}^2/_x$ und somit $x^2 = 2$. Weil diese Gleichung keine Lösung in $\mathbb{Q}$ hat, kann (a_n) keinen Grenzwert in $\mathbb{Q}$ haben. Das heißt, sie ist eine Cauchyfolge, die nicht konvergiert.

Bemerkung B6.8. Dieses iterative Verfahren zum Wurzelziehen ist nach Heron von Alexandria benannt. Für $k \in \mathbb{N}^+$ und $b > 0$ nennen wir $\sqrt[k]{b}$ (die) eine positive Lösung zu $x^k = b > 0$. Um $\sqrt[k]{b}$ zu finden, setzen wir $a_0 := b$ und

$$a_{n+1} := \frac{a_n}{k}\left(k - 1 + \frac{b}{a_n}\right).$$

Dieses Verfahren ist ein Spezialfall des Newtonverfahrens zur numerischen Lösung von Gleichungen $f(x) = b$.

c. *Häufungspunkte*

Definition B6.9. Wir sagen, $x \in K$ ist ein Häufungspunkt der Folge (a_n), falls (a_n) eine Teilfolge hat, die gegen x konvergiert.

Bemerkung B6.10. Eine Folge (a_n) hat Häufungspunkt x genau dann, wenn zu jedem $\varepsilon > 0$ es unendlich viele n gibt mit $x - \varepsilon < a_n < x + \varepsilon$.

Lemma B6.11. *Eine konvergente Folge hat einen einzigen Häufungspunkt, nämlich den Grenzwert. Eine divergente Cauchyfolge hat keinen Häufungspunkt.*

Beweis. Falls $a_n \to x$, dann konvergiert auch jede Teilfolge gegen x. Eine Cauchyfolge mit einer konvergente Teilfolge konvergiert auch selbst. □

B7. Die reellen Zahlen

Man könnte sagen, jede Cauchyfolge möchte gern konvergieren, wenn es nur einen passenden Grenzwert gäbe. Wir erweitern oder *vervollständigen* die rationalen Zahlen $\mathbb{Q}$ auf die reellen Zahlen $\mathbb{R}$, in dem wir Grenzwerte zu allen Cauchyfolgen hinzufügen. Der Grenzwert zu (a_n) wird einfach durch die Folge (a_n) selbst dargestellt. Natürlich gibt es aber mehrere Folgen, die denselben Grenzwert haben sollten, z.B. alle Teilfolgen von (a_n). Das heißt, wir müssen eine Äquivalenzrelation auf Cauchyfolgen definieren. Die reellen Zahlen werden als die Äquivalenzklassen definiert.

Definition B7.1. Sei F die Menge aller Cauchyfolgen in $\mathbb{Q}$. Mit den komponentenweisen Verknüpfungen

$$(a_n) + (b_n) := (a_n + b_n), \quad (a_n)(b_n) := (a_n b_n)$$

wird F ein Ring, in dem wir $\mathbb{Q}$ mit der Teilmenge aller konstanten Folgen identifizieren. Sei $N \subset F$ jetzt die Teilmenge aller *Nullfolgen*, d.h., aller Folgen, die gegen 0 konvergieren. Auf F definieren wir die Äquivalenzrelation

$$a \sim b \iff a - b \in N.$$

Die reellen Zahlen sind dann die Äquivalenzklassen dieser Relation: $\mathbb{R} := F/\sim$.

Bemerkung B7.2. Die folgende Eigenschaften sind die wichtigen, um zu zeigen, dass Addition und Multiplikation auf $\mathbb{R}$ wohldefiniert sind und dass $\mathbb{R}$ dadurch ein Ring wird:

- $a, b \in N \implies a + b \in N$,
- $a \in F,\ b \in N \implies ab \in N$.

In der abstrakten Algebra heißt das, N ist ein *Ideal* in F. Der Quotient $F/\sim$ wird dann F/N geschrieben und ist automatisch ein Ring.

Wir sagen, eine Cauchyfolge (a_n) ist *nichtpositiv*, falls für alle $\varepsilon > 0$ (in $\mathbb{Q}$) die Ungleichung $a_n < \varepsilon$ für fast alle n gilt. Man kann leicht zeigen, (a_n) ist genau dann nichtpositiv, wenn es eine äquivalente Folge $(b_n) \sim (a_n)$ mit $b_n \le 0$ für alle n gibt. Die Ordnung auf $\mathbb{R}$ wird so definiert, dass

$$[(a_n)] \le [(b_n)] \iff (a_n - b_n) \text{ nichtpositiv.}$$

Sei $0 < x \in \mathbb{R}$. Dann hat x eine Darstellung $x = [(a_n)]$, wobei $a_n > 0$ für alle n. Weil die Cauchyfolge a_n nicht gegen 0 konvergiert, ist 0 auch kein Häufungspunkt. Das heißt, es gibt $\varepsilon > 0$, so dass $a_n > \varepsilon$ für fast alle n. Wir können (a_n) mit einer äquivalente Folge ersetzen, wobei $a_n > \varepsilon$ für alle n. Es folgen zwei wichtigen Tatsachen: Erstens, dass $\left({}^1/_{a_n}\right)$ eine Cauchyfolge ist, die ein multiplikatives Inverses zu x liefert, wodurch $\mathbb{R}$ ein angeordneter Körper wird. Zweitens, dass $x \ge \varepsilon > 0$, was impliziert, dass die archimedische Eigenschaft der rationalen Zahlen auf die reellen Zahlen übertragen wird.

Lemma B7.3. *Zu jedem $x \in \mathbb{R}$ gibt es eine größte ganze Zahl $\lfloor x \rfloor \in \mathbb{Z}$, die kleiner gleich x ist.*

Beweis. Weil $\mathbb{R}$ archimedisch ist, gibt es $n \in \mathbb{N}$ mit $n > x$. Sei $T := \{m \in \mathbb{N} : n - m \le x\}$. Weil T nicht leer ist, hat T ein Minimum m_0. Dann gilt $\lfloor x \rfloor = n - m_0$. □

Bemerkung B7.4. Ähnlich definiert man $\lceil x \rceil := -\lfloor -x \rfloor$. Es gilt

$$x - 1 < \lfloor x \rfloor \le x \le \lceil x \rceil < x + 1.$$

Falls $x \in \mathbb{Z}$, gilt $\lfloor x \rfloor = x = \lceil x \rceil$; sonst gilt $\lfloor x \rfloor + 1 = \lceil x \rceil$.

`Ende der Vorlesung 2008 Mai 5`

Satz B7.5. *Die rationalen Zahlen sind* dicht *in $\mathbb{R}$ im Sinne, dass zwischen jedem Paar $a < b \in \mathbb{R}$ es eine rationale Zahl $a < q < b$ gibt.*

Beweis. Weil $\mathbb{R}$ archimedisch ist, gibt es $n \in \mathbb{N}$, so dass ${}^1/_n < b - a$. Wir setzen $m := \lceil nb - 1 \rceil$. Dann gilt

$$a < b - {}^1/_n \le {}^m/_n < b.$$ □

B8. Vollständigkeit

Definition B8.1. Sei X durch $\leq$ total geordnet und sei $T \subset X$ eine Teilmenge. Ein $x \in X$ heißt eine *obere Schranke* für T, falls $t \leq x$ für alle $t \in T$. Falls $x \in X$ eine obere Schranke für T ist und es keine kleinere obere Schranke $x' < x$ gibt, heißt x die *kleinste obere Schranke* oder das *Supremum* von T. Wir schreiben $x = \sup T$. Falls $\sup T \in T$ nennen wir es auch das *Maximum* von T und schreiben $\max T$ dazu. Ähnlich definieren wir *untere Schranke*, *Infimum* (oder *größte untere Schranke*) und *Minimum*.

Wie wir schon erwähnt haben, ist $\mathbb{N}$ wohlgeordnet im Sinne, dass jede nicht leere Teilmenge $T \subset \mathbb{N}$ ein Minimum hat. Ein Maximum hat T genau dann, wenn sie eine obere Schranke hat. In $\mathbb{Q}$ hat $\{t : t^2 < 2\}$ zwar obere Schranken (jedes $x > 0$ mit $x^2 > 2$) aber kein Supremum (in $\mathbb{Q}$) weil $\sqrt{2} \notin \mathbb{Q}$. Die Teilmenge $T := \{x : x^2 < 1\}$ hat ein Supremeum $\sup T = 1$, was aber kein Maximum ist, weil $1 \notin T$.

Definition B8.2. Eine total geordnete Menge X heißt dann *ordnungsvollständig*, wenn jede nicht leere und nach oben beschränkte Teilmenge $T \subset X$ ein Supremum $\sup T \in X$ besitzt und jede nicht leere und nach unten beschränkte Teilmenge $T \subset X$ ein Infimum $\inf T \in X$ besitzt.

Das Beispiel $\{t : t^2 < 2\}$ zeigt, dass $\mathbb{Q}$ nicht ordnungsvollständig ist.

Satz B8.3. *Die reellen Zahlen sind ordnungsvollständig.*

Beweis. Sei $T \subset \mathbb{R}$ eine nicht leere und nach oben beschränkte Teilmenge. Weil es eine obere Schranke in $\mathbb{R}$ gibt, gibt es auch eine (größere) $b_0 \in \mathbb{Q}$. Weil T nicht leer ist, gibt es auch eine reelle Zahl – und deshalb auch ein $a_0 \in \mathbb{Q}$ – die keine obere Schranke für T ist.

Wir definieren rekursiv zwei monotone Folgen (a_n) und (b_n), wobei kein a_n aber jedes b_n eine obere Schranke für T ist. Dazu setzen wir zunächst $m_n := {}^1\!/_2(a_n + b_n)$. Falls m_n eine obere schranke für T ist, setzen wir $a_{n+1} := a_n$ und $b_{n+1} := m_n$. Sonst setzen wir $a_{n+1} := m_n$ und $b_{n+1} := b_n$. Die Folge (a_n) ist monoton steigend und nach oben (z.B. durch jedes b_n) beschränkt; (b_n) ist monoton fallend und nach unten beschränkt. Deshalb sind diese beiden Cauchyfolgen, die reellen Zahlen $x := [(a_n)]$ bzw. $y; = [(b_n)]$ darstellen.

Wir haben $b_n - a_n = {}^1\!/_{2^n}(b_0 - a_0) \to 0$, was heißt, $x = y$. Dann gilt $x = \sup T$: sicherlich ist x (wie jedes b_n) eine obere Schranke, für jedes $z < x$ gibt es aber n mit $z < a_n$ und deshalb ist z keine obere Schranke. □

Eine zweite Bedeutung von *Vollständigkeit* ist die Konvergenz aller Cauchyfolgen.

Lemma B8.4. *Jede Folge in $\mathbb{R}$ (oder in einer beliebigen total geordneten Menge) hat eine monotone Teilfolge.*

Beweis. Sei (a_n) eine Folge. Nennen wir (vorübergehend) ein Glied a_n *dominant*, falls $a_n \geq a_m$ für alle $m \geq n$. Falls es unendlich viele dominante Glieder gibt, nehmen wir genau diese als (monoton fallende) Teilfolge. Falls es nur endlich viele dominante Glieder gibt, nehmen wir zunächst alle Glieder nach der letzten Dominanten als Teilfolge. Diese hat kein dominantes Glied, d.h., nach jedem Glied kommt irgendwann ein noch Größeres. Also können wir eine streng monoton steigende Teilfolge finden. □

Lemma B8.5. *Jede monotone und beschränkte Folge in $\mathbb{R}$ ist konvergent.*

Beweis. Sei (a_n) eine beschränkte und monoton steigende Folge. Die nichtleere Menge $\{a_n : n \in \mathbb{N}\}$ ist nach oben beschränkt, und hat deshalb ein Supremum $x := \sup\{a_n\}$.

Wir zeigen, $a_n \to x$. Sei nun $\varepsilon > 0$. Es gibt ein n_0 mit $a_{n_0} > x - \varepsilon$, weil sonst $x - \varepsilon$ eine kleinere obere Schranke wäre. Für alle $n \geq n_0$ folgt $x - \varepsilon < a_n \leq x$. □

Satz B8.6. *Die reellen Zahlen sind* metrisch vollständig *im Sinne, dass jede Cauchyfolge konvergiert.*

Beweis. Sei (a_n) eine Cauchyfolge. Sie ist beschränkt und hat (1. Lemma) eine monotone Teilfolge, die (2. Lemma) konvergiert. Eine Cauchyfolge mit einer konvergenten Teilfolge konvergiert auch. □

Bemerkung B8.7. Sei K ein angeordneter Körper. Man kann folgendes beweisen: Falls K ordnungvollständig ist, folgt $K \cong \mathbb{R}$. Falls K archimedisch und metrisch vollständig ist, folgt auch $K \cong \mathbb{R}$.

Deshalb fangen viele Autoren mit Axiomen für $\mathbb{R}$ an: $\mathbb{R}$ sei ein angeordneter Körper, der (nach Geschmack) entweder ordnungsvollständig ist oder archimedisch und metrisch vollständig ist. Wenn man voraussetzt, dass es solch einen vollständigen angeordneten Körper gibt, dann folgen aus diesen Axiomen alle weitere Eigenschaften von $\mathbb{R}$.

Wir haben hingegen den Körper $\mathbb{R}$ explizit gebaut. Wir benutzen aber in Zukunft auch nicht mehr die Details unserer Konstruktion mittels Cauchyfolgen, sondern nur die eben erwähnten Eigenschaften, dass $\mathbb{R}$ ein vollständiger angeordneter Körper ist. Daraus folgt alles weitere.

`Ende der Vorlesung 2008 Mai 6`

C. REIHEN

C1. Definition

Definition C1.1. Sei $(a_n)_{n\geq 1}$ eine Folge in $\mathbb{R}$. Wir definieren dazu die Folge (s_n) der *Partialsummen*:

$$s_n := \sum_{k=1}^{n} a_k := a_1 + a_2 + \cdots + a_n,$$

welche eine Abkürzung für die rekursive Definition

$$s_0 := 0, \quad s_{n+1} := s_n + a_{n+1}$$

ist. Diese neue Folge nennen wir die zu (a_n) gehörgige *(unendliche) Reihe*. Falls sie konvergiert, nennen wir den Grenzwert die *Reihensumme*:

$$\lim_{n\to\infty} \sum_{k=1}^{n} a_k =: \sum_{k=1}^{\infty} a_k =: \sum a_k =: a_1 + a_2 + \cdots .$$

Wir benutzen $\sum a_k$ als Name sowohl für die Reihe selbst als auch für deren Grenzwert (falls dieser existiert).

Bemerkung C1.2. Natürlich tauchen auch Reihen der Form $\sum_{k=n_0}^{\infty} a_k$ auf, wo $n_0 \neq 1$.

Bemerkung C1.3. Sei (s_n) eine beliebige Folge (mit $s_0 = 0$). Wenn wir $a_n := s_n - s_{n-1}$ setzen, dann sind die s_n die Partialsummen der a_n. Eine Reihe $\sum a_k$ ist nichts anderes als die Folge (s_n) der Partialsummen. Wir entscheiden nur, dass wir uns für die Differenzen $a_k = s_k - s_{k-1}$ interessieren.

Satz C1.4 (Rechenregeln). *Seien $\sum b_n$ und $\sum c_n$ konvergente Reihen, und sei $a \in \mathbb{R}$. Die Folge $\sum(ab_n + c_n)$ konvergiert auch. Es gilt*

$$\sum_{n=1}^{\infty}(ab_n + c_n) = a\sum_{n=1}^{\infty} b_n + \sum_{n=1}^{\infty} c_n.$$

Beweis. Eine Reihe ist nichts mehr als eine Folge (der Partialsummen). Diese Rechenregeln folgen deshalb unmittelbar aus denen für Folgen. □

Bemerkung C1.5. Wir bekommen dann eine Teilfolge der Partialsummen, wenn wir Klammern setzen, z.B.

$$(a_1 + a_2) + (a_3) + (a_4 + a_5 + a_6) + (a_7 + a_8) + \cdots .$$

Genauer (aber wahrscheinlich nicht klarer) gesagt, sei (a_n) eine Folge mit Partialsummen (s_n), und sei (n_i) eine streng monoton steigende Folge in $\mathbb{N}$. Definieren wir die Teilfolge $t_i := s_{n_i}$, sind diese t_i die Partialsummen der Folge

$$b_i := s_{n_i} - s_{n_{i-1}} = \sum_{k=n_{i-1}+1}^{n_i} a_k.$$

Falls $\sum a_k$ existiert, haben wir

$$\sum b_n = \lim t_n = \lim s_k = \sum a_k,$$

weil jede Teilfolge einer konvergenten Folgen denselben Grenzwert hat. Natürlich kann es sein, dass $\sum b_n$ auch dann existiert, wenn $\sum a_k$ nicht existiert. (Eine divergente Folge s_k kann Häufungspunkte – und entsprechend konvergente Teilfolgen – haben.)

Beispiel C1.6. Sei $a_n := (-1)^n$. Dann existiert

$$\sum_{n=0}^{\infty} a_n = 1 - 1 + 1 - 1 + \cdots$$

nicht. Die Partialsummen haben zwei Häufungspunkte:

$$0 = (1-1)+(1-1)+\cdots , \quad 1 = 1+(-1+1)+(-1+1)+\cdots .$$

Natürlich dürfen wir nicht daraus $1 = 0$ schließen!

Wir wissen schon, eine Folge in $\mathbb{R}$ konvergiert genau dann, wenn sie eine Cauchyfolge ist. Dieser Satz können wir auf Reihen umsetzen, in dem wir merken, was die Differenz zweier Partialsummen ist. Für $n \geq m$ gilt nämlich

$$s_n - s_{m-1} = \sum_{k=m}^{n} a_k.$$

Satz C1.7 (Cauchykriterium). *Die Reihe $\sum a_k$ konvergiert genau dann, wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt, so dass für alle $n \geq m \geq n_0$ gilt*

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

Korollar C1.8. *Ist a_k keine Nullfolge, dann konvergiert $\sum a_k$ nicht.*

Beweis. Dies ist der Spezialfall $m = n$ im Satz. □

Umgekehrt kann man nichts sagen. Zum Beispiel ist $a_n = 1/n$ eine Nullfolge; die *harmonische Reihe* $\sum_{n=1}^{\infty} 1/n$ konvergiert aber nicht. Dagegen haben wir

$$\begin{aligned}
&1 + \tfrac12 + \left(\tfrac13 + \tfrac14\right) + \left(\tfrac15 + \tfrac16 + \tfrac17 + \tfrac18\right) + \left(\tfrac19 + \cdots\right) + \cdots \\
&> 1 + \tfrac12 + \left(\tfrac14 + \tfrac14\right) + \left(\tfrac18 + \tfrac18 + \tfrac18 + \tfrac18\right) + \left(\tfrac1{16} + \cdots\right) + \cdots \\
&= 1 + \tfrac12 + \quad \tfrac12 \quad + \quad \tfrac12 \quad + \quad \tfrac12 \quad + \cdots .
\end{aligned}$$

(Um dieses ausführlicher aufzuschreiben, würde man die Teilfolge $t_n := s_{2^n}$ der Partialsummen betrachten. Wir finden $t_{n+1} - t_n \geq 1/2$, was bedeutet, t_n divergiert.)

Das wichtigste Beispiel einer konvergenten Reihe ist die *geometrische Reihe* $\sum q^n$ für $|q| < 1$.

Lemma C1.9. *Sei $q \in \mathbb{R}$. Die Folge (q^n) ist genau dann eine Nullfolge, wenn $|q| < 1$.*

Beweis. Zu $|q| < 1$ setzen wir $x := 1/|q| - 1 > 0$. Die so gennante Bernoulli'sche Ungleichung (die man leicht mittels Induktion über n beweisen kann) sagt für $x > 0$ und $n \in \mathbb{N}$, dass $(1+x)^n \geq 1 + nx$. Das heißt hier, $1/|q|^n \geq 1 + nx > nx$. Dann gilt $|q^n| < \frac{1}{n}\frac{1}{x}$. Damit ist q^n der Produkt einer Nullfolge ($1/n$) mit einer (durch $1/x$) beschränkten Folge und deshalb selbst eine Nullfolge. □

Korollar C1.10. *Für $|q| < 1$ konvergiert die geometrisch Reihe*

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

Beweis. Mittels Induktion über n zeigt man

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

Dass diese Folge gegen $1/(1-q)$ konvergiert, folgt aus dem Lemma und unseren Rechenregeln. □

Bemerkung C1.11. Aus den Rechenregeln sehen wir (ohne Beschränkung auf $|q|$) folgendes. *Falls* $\sum q^n$ gegen x konvergiert, *dann* ist dieser Grenzwert $x = 1/(1-q)$:

$$\begin{aligned} x &= 1 + q + q^2 + q^3 + \cdots \\ &= 1 + q(1 + q + q^2 + \cdots) \\ &= 1 + qx. \end{aligned}$$

Natürlich konvergiert z.B. weder $1+2+4+\cdots$ gegen -1 noch $1-1+1-1+\cdots$ gegen $1/2$.

C2. Konvergenzkriterien und absolute Konvergenz

Weitere konvergente Beispiele finden wir unter den *alternierenden Reihen*, deren Glieder alternierend positiv und negativ sind.

Satz C2.1 (Leibnizkriterium). *Sei (a_k) eine monotone Nullfolge. Dann konvergiert $\sum(-1)^k a_k$.*

Beispiel C2.2. Die Reihe

$$1 - 1/2 + 1/3 - 1/4 + \cdots$$

ist konvergent, weil das Leibnizkriterium erfüllt ist. Später werden wir sehen, dass der Grenzwert $\log 2 \approx 0.693$ ist.

Beweis. Ohne Beschränkung der Allgemeinheit (im Folgenden mit „o.B.d.A." abgekürzt) ist a_k monoton fallend, und deshalb $a_k \geq 0$. Sei $s_n := \sum_{k=1}^{n}(-1)^k a_k$. Dann gilt

$$s_1 \leq s_3 \leq s_5 \leq \cdots \leq \cdots \leq s_4 \leq s_2 \leq s_0,$$

weil $a_n - a_{n-1} \leq 0$. Das heißt, die Teilfolgen (s_{2n+1}) bzw. (s_{2n}) sind beschränkt und monoton steigend bzw. fallend. Die haben deshalb Grenzwerte x bzw. y. Dann gilt $x - y = \lim\big(s_{2n+1} - s_{2n}\big) = \lim a_{2n} = 0$. □

Seien s_n und s_{n+1} zwei beliebige nacheinander folgende Partialsummen einer solchen Leibnizreihe. Dann liegt der Grenzwert immer zwischen diesen Beiden.

Definition C2.3. Eine Reihe $\sum a_k$ heißt *absolut konvergent*, wenn die Reihe $\sum |a_k|$ konvergent ist. Eine konvergente Reihe, die nicht absolut konvergent ist, heißt *bedingt konvergent*.

Bemerkung C2.4. Wenn wir uns für die Reihe $\sum a_k$ interessieren, scheint es auf erstem Blick vielleicht unwichtig ob $\sum |a_k|$ konvergiert. Absolut konvergente Reihe haben aber viele schöne Eigenschaften. Deshalb merken wir später bei jedem Konvergenzkriterium, ob es auch absolute Konvergenz liefert.

Beispiel C2.5. Die obige Reihe $1 - 1/2 + 1/3 - \cdots$ ist bedingt konvergent, weil die harmonische Reihe divergiert.

`Ende der Vorlesung 2008 Mai 13`

Bemerkung C2.6. Die Erweiterung der Dreiecksungleichung auf mehereren Größen,

$$\left|\sum_{k=1}^{n} a_k\right| \leq \sum_{k=1}^{n} |a_k|,$$

beweist man (wie sonst?) mittels Induktion über n.

Lemma C2.7. *Eine absolut konvergente Reihe ist konvergent und es gilt*

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

Beweis. Sei $\sum a_k$ eine absolut konvergente Reihe. Um Konvergenz zu betrachten, haben wir zur Verfügung bis jetzt nur das Cauchykriterium. Sei $\varepsilon > 0$. Weil $\sum |a_k|$ konvergiert, gibt es $n_0 \in \mathbb{N}$, so dass

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

für alle $n \geq m \geq n_0$. Dann gilt für solche m, n auch

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

Dadurch konvergiert $\sum a_k$. Weil für die Partialsummen $\left|\sum_1^n a_k\right| \leq \sum_1^n |a_k|$ gilt, folgt im Limes $\left|\sum_1^\infty a_k\right| \leq \sum_1^\infty |a_k|$. □

Satz C2.8 (Vergleichskriterium). *Seien $\sum a_k$ und $\sum b_k$ Reihen, für die $|a_k| \leq b_k$ für fast alle k gilt. Ist $\sum b_k$ konvergent, dann ist $\sum a_k$ absolut konvergent.*

Beweis. Sei $0 \leq |a_k| \leq b_k$ für alle $k > n_0$. Seien $s_n := \sum_1^n |a_k|$ die Partialsummen der Reihe $\sum |a_k|$. Die Folge (s_n) ist monoton steigend und durch

$$s_{n_0} + \sum_{k=n_0+1}^{\infty} b_k$$

beschränkt, also konvergent. (Wir hätten auch $n_0 = 0$ annehmen dürfen, dann wäre die Notation einfacher.) □

Beispiel C2.9. Um zu zeigen, dass $\sum 1/k^2$ konvergiert, setzen wir

$$b_k := \frac{1}{k-1} - \frac{1}{k} = \frac{1}{k(k-1)} > \left|\frac{1}{k^2}\right|$$

für $k \geq 2$. Die Reihe $\sum b_k$ konvergiert gegen 1, weil

$$\sum_{k=2}^{n} b_k = \left(\tfrac{1}{1} - \tfrac{1}{2}\right) + \left(\tfrac{1}{2} - \tfrac{1}{3}\right) + \cdots + \left(\tfrac{1}{n-1} - \tfrac{1}{n}\right) = 1 - \tfrac{1}{n}.$$

Nach dem Vergleichskreiterium konvergiert dann auch $\sum_{k=1}^{\infty} 1/k^2$. (Übrigens ist der Grenzwert $\pi^2/6$.)

Satz C2.10 (Wurzelkriterium). *Sei $\sum a_k$ eine Reihe. Falls es $q < 1$ gibt, so dass $\sqrt[k]{|a_k|} \leq q$ für fast alle $k \in \mathbb{N}$, dann ist die Reihe $\sum a_k$ absolut konvergent. Falls hingegen $\sqrt[k]{|a_k|} \geq 1$ für unendlich viele k, dann ist $\sum a_k$ divergent.*

Beweis. Falls $q < 1$ und $|a_k| \leq q^k$, dann liefert das Vergleichskriterium die absolute Konvergenz. Falls hingegen $|a_k| \geq 1$ für unendlich viele k, ist (a_k) keine Nullfolge. □

Bemerkung C2.11. Es ist sehr wichtig, dass $q < 1$ unabhängig von n ist. Dadurch ist $\sqrt[k]{|a_k|}$ weg von 1 beschränkt. Falls z.B. $\sqrt[k]{|a_k|} < 1$ und $\sqrt[k]{|a_k|} \to 1$, verspricht das Wurzelkriterium weder Konvergenz noch Divergenz. (Und gut so, weil sowohl die divergente $\sum 1/k$ als auch die konvergente $\sum 1/k^2$ diese beiden Eigenschaften haben.)

Korollar C2.12 (Quotientenkriterium). *Sei $\sum a_k$ eine Reihe, wobei $a_k \neq 0$ für fast alle k. Falls es $q < 1$ gibt, so dass $\left|a_{k+1}/a_k\right| \leq q$ für fast alle k, dann ist $\sum a_k$ absolut konvergent. Falls hingegen $\left|a_{k+1}/a_k\right| \geq 1$ für fast alle k, dann ist $\sum a_k$ divergent.*

Beweis. Falls $\left|a_{k+1}/a_k\right| \leq q$ für alle $k \geq n_0$, dann gilt

$$|a_k| \leq q^k \frac{|a_{n_0}|}{q^{n_0}}$$

und das Wurzelkriterium liefert absolute Konvergenz. Falls hingegen $\left|a_{k+1}/a_k\right| \geq 1$ für alle $k \geq n_0$, dann gilt $|a_k| \geq |a_{n_0}|$. Deshalb ist (a_k) keine Nullfolge. □

C3. Dezimaldarstellung

Satz C3.1 (Dezimaldarstellung). *Sei $n_0 \in \mathbb{N}$ und sei $(a_k)_{k \geq -n_0}$ eine Folge in $\{0, 1, \ldots, 9\}$. Dann konvergiert absolut*

$$\sum_{k=-n_0}^{\infty} \frac{a_k}{10^k} =: a_{-n_0} a_{-(n_0-1)} \cdots a_{-1} a_0 . a_1 a_2 a_3 \cdots .$$

Jede reelle Zahl $x \geq 0$ ist der Limes einer solchen Dezimaldarstellung.

Beweisskizze. Die Konvergenz folgt vom Vergleichskriterium:

$$\sum_{k=-n_0}^{\infty} \frac{|a_k|}{10^k} \leq \sum_{k=-n_0}^{\infty} \frac{9}{10^k} = 10^{n_0+1}.$$

Um eine Dezimaldarstellung von $x > 0$ zu finden, wählen wir zunächst $n_0 \in \mathbb{N}$ mit $10^{n_0} > x$. Wir definieren rekursiv $a_{-n_0} := 0$ und

$$a_{n+1} := \left\lfloor 10^{n+1} \left(x - \sum_{k=-n_0}^{n} \frac{a_k}{10^k} \right) \right\rfloor. \qquad \square$$

Bemerkung C3.2. Wenn wir führende Nullen weglassen, ist die Dezimaldarstellung fast immer eindeutig. Nur die rationalen Zahlen der Form $m/10^n$ haben jeweils zwei Darstellungen. Es gelten z.B. $0.999\cdots = 1 = 1.000\cdots$ und $0.24999\cdots = 1/4 = 0.25000\cdots$.

C4. Umordnungen

Die Formel $\sum a_k + \sum b_k = \sum(a_k + b_k)$ ist in einem Sinne eine unendliche Version des Kommutativgesetzes für Addition. Kann man aber eine konvergente Reihe beliebig umordnen ohne deren Summe zu ändern? Die Antwort hängt davon ab, ob die Reihe absolut oder bedingt konvergiert. Zunächst müssen wir sagen, was genau wir mit einer Umordnung meinen.

Definition C4.1. Sei $\varphi\colon \mathbb{N} \to \mathbb{N}$ eine Bijektion und sei $\sum_{k=0}^{\infty} a_k$ eine unendliche Reihe. Dann heißt die Reihe

$$\sum_{k=0}^{\infty} a_{\varphi}(k)$$

eine *Umordnung* von $\sum a_k$.

Satz C4.2. *Sei $\sum a_k$ eine absolut konvergente Reihe. Dann ist auch jede Umordnung absolut konvergent mit derselben Summe $\sum a_{\varphi(k)} = \sum a_k$.*

Beweisskizze. Wir dürfen annehmen, $a_k \geq 0$, d.h., $|a_k| = a_k$. Sei nun $T_n := \{\varphi(0), \varphi(1), \ldots, \varphi(n)\} \subset \mathbb{N}$. Wir setzen $i_n := \min(\mathbb{N} \setminus T_n)$ und $j_n := \max T_n$. Dann sind i_n und j_n Folgen in $\mathbb{N}$, die jeweils monoton steigend und unbeschränkt sind. Fur jedes $k \in \mathbb{N}$ merken wir,

$$\sum_{k=0}^{i_n} a_k \leq \sum_{k=0}^{n} a_{\varphi(k)} \leq \sum_{k=0}^{j_n} a_k.$$

Im Limes $n \to \infty$ gehen auch i_n und j_n nach unendlich. Deshalb gilt $\sum a_k \leq \sum a_{\varphi(k)} \leq \sum a_k$. □

Satz C4.3. *Sei $\sum a_k$ eine bedingt konvergente Reihe in $\mathbb{R}$ und sei $x \in \mathbb{R}$. Es gibt eine Umordnung, die gegen x konvergiert: $\sum a_{\varphi(k)} = x$.*

Beispiel C4.4. Wir wollen die Reihe $1 - {}^1\!/_2 + {}^1\!/_3 - \cdots$ umordnen, um dass die Summe $x = {}^1\!/_3$ ist. Wir nehmen zunächst $a_1 = 1$, danach genug negative Glieder um eine Partialsumme kleiner als x zu erreichen: $1 - {}^1\!/_2 - {}^1\!/_4 = {}^1\!/_4$. Danach genug positive Glieder um wieder eine Partialsumme größer als x zu erreichen: $1 - {}^1\!/_2 - {}^1\!/_4 + {}^1\!/_3 = {}^7\!/_{12}$. Dann nehmen wir wieder negative Glieder, usw.:

$$1 \quad - {}^1\!/_2 - {}^1\!/_4 \quad + {}^1\!/_3 \quad - {}^1\!/_6 - \cdots - {}^1\!/_{16} \quad + {}^1\!/_5 \quad - \cdots .$$

Beweisskizze. Seien (b_i) und (c_j) die Teilfolgen der positiven bzw. nichtpositiven Glieder. Beide Reihen $\sum b_i$ und $\sum c_j$ sind divergent. Wir können deshalb die obige Prozedur durchführen, in dem wir alternierend genug Glieder aus (b_i) bzw. (c_j) nehmen, bis die Partialsumme über bzw. unter x liegt. □

`Ende der Vorlesung 2008 Mai 19`

D. STETIGKEIT

D1. Offene und abgeschlossene Mengen

Definition D1.1. Für $\varepsilon > 0$ und $x \in \mathbb{R}$ heißt

$$\begin{aligned} B_\varepsilon(x) &:= \{p \in \mathbb{R} : |x-p| < \varepsilon\} \\ &= \{p \in \mathbb{R} : x-\varepsilon < p < x+\varepsilon\} \end{aligned}$$

der *ε-Ball um x* oder die *ε-Umgebung von x*.

Definition D1.2. Sei $D \subset \mathbb{R}$ eine Teilmenge. Ein Punkt $p \in D$ heißt ein *innerer Punkt von D*, falls es ein $\varepsilon > 0$ gibt, so dass $B_\varepsilon(p) \subset D$. Die Teilmenge D heißt *offen*, falls jeder Punkt $p \in D$ ein innerer Punkt ist.

Beispiele D1.3. Die Mengen $\mathbb{R}$ und $\emptyset$ sind auf triviale Weise offen. Die Menge $T := \{x \geq 0\}$ ist nicht offen, weil 0 kein innerer Punkt ist. Jeder Ball $B_\varepsilon(x)$ ist offen. (Sei $p \in B_\varepsilon(x)$ und sei $0 < \delta \leq \varepsilon - |p-x|$. Wir behaupten, dass $B_\delta(p) \subset B_\varepsilon(x)$. Für $q \in B_\delta(p)$ gilt nämlich

$$|q-x| \leq |q-p| + |p-x| < \delta + |p-x| \leq \varepsilon$$

von der Dreiecksungleichung.)

Definition D1.4. Sei $D \subset \mathbb{R}$. Ein Punkt $p \in \mathbb{R}$ heißt ein *Randpunkt von D*, falls p ein innerer Punkt weder von D noch vom Komplement $\mathbb{R} \smallsetminus D$ ist. Der *Rand von D* ist die Menge $\partial D \subset \mathbb{R}$ aller Randpunkte. Eine Teilmenge $D \subset \mathbb{R}$ heißt *abgeschlossen*, falls deren Komplement $\mathbb{R} \smallsetminus D$ offen ist.

Lemma D1.5. *Eine Teilmenge $D \subset \mathbb{R}$ ist genau dann abgeschlossen, wenn $\partial D \subset D$.*

Beweis. Sei $D \subset \mathbb{R}$ abgeschlossen und sei $p \in \partial D$ ein Randpunkt von D. Definitionsgemäß ist p kein innerer Punkt von $\mathbb{R} \smallsetminus D$. Weil diese Menge offen ist, bedeutet das, dass $p \notin \mathbb{R} \smallsetminus D$, d.h., $p \in D$.

Umgekehrt, sei $\partial D \subset D \subset \mathbb{R}$. Wir zeigen, dass $U := \mathbb{R} \smallsetminus D$ offen ist. Für $p \in U$ gilt $p \notin \partial D$, was heißt, p ist ein innerer Punkt entweder von D oder von U. Weil $p \notin D$, ist p in diesem Fall ein innerer Punkt von U. □

D2. Stetigkeit

Wir betrachten jetzt eine Funktion oder Abbildung $f : D \to \mathbb{R}$, deren Definitionsbereich eine beliebige Teilmenge $D \subset \mathbb{R}$ ist.

Definition D2.1. Sei $p \in D \subset \mathbb{R}$. Die Funktion $f : D \to \mathbb{R}$ heißt *stetig in p*, falls

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad f\big(B_\delta(p) \cap D\big) \subset B_\varepsilon\big(f(p)\big),$$

das heißt, falls zu jedem $\varepsilon > 0$ es ein $\delta > 0$ gibt, so dass

$$x \in D,\ |x-p| < \delta \implies |f(x) - f(p)| < \varepsilon.$$

Die Funktion f heißt *stetig* (auf D), falls sie in jedem Punkt $p \in D$ stetig ist.

Beispiel D2.2. Die Funktion $x \mapsto |x|$ ist stetig. (Wir wählen $\delta \leq \varepsilon$ und benutzen die Dreiecksungleichung in der Form $\big||x| - |p|\big| \leq |x-p|$.)

Beispiel D2.3. Die Signumfunktion

$$x \mapsto \operatorname{sgn} x := \begin{cases} -1, & x < 0, \\ 0, & x = 0, \\ +1, & x > 0 \end{cases}$$

ist stetig in jedem $p \neq 0$ (wir wählen $\delta \leq |p|$) aber unstetig in 0 (für $\varepsilon < 1$ gibt es kein passendes δ).

Beispiel D2.4. Die Funktion $f : \mathbb{R}^* \to \mathbb{R}$, $f(x) = \sin {}^1\!/_x$ ist stetig. Wenn wir aber eine Erweiterung definieren, in dem wir einen Wert $f(0)$ wählen (z.B. $f(0) := 0$), ist diese Erweiterung nie stetig in 0. (Auf jedem kleinen $B^*_\varepsilon(0)$ nimmt $\sin {}^1\!/_x$ alle Werte zwischen ± 1 an.)

Satz D2.5. *Sei $f : D \to \mathbb{R}$ eine Funktion, deren Definitionsbereich $D \subset \mathbb{R}$ offen ist. Die Funktion f ist stetig genau dann, wenn das Urbild $f^{-1}(U)$ jeder offenen Menge $U \subset \mathbb{R}$ offen ist.*

Beweis. Sei f stetig. Betrachten wir das Urbild $f^{-1}(U)$ einer offenen Menge $U \subset \mathbb{R}$. Sei $x \in f^{-1}(U)$, d.h., $y := f(x) \in U$. Weil U offen ist, gibt es $\varepsilon > 0$ mit $B_\varepsilon(y) \subset U$. Weil f stetig ist, gibt es zu diesem ε ein $\delta_1 > 0$, so dass

$$f\big(B_{\delta_1}(x) \cap D\big) \subset B_\varepsilon(y) \subset U.$$

Das heißt, $B_{\delta_1}(x) \cap D \subset f^{-1}(U)$. Weil D offen ist, gibt es $\delta_2 > 0$, so dass $B_{\delta_2}(x) \subset D$. Für $\delta := \min(\delta_1, \delta_2) > 0$ gilt deshalb $B_\delta(x) \subset f^{-1}(U)$.

Umgekehrt, seien die Urbilder aller offenen Mengen offen. Betrachten wir die Stetigkeit von f in $x \in D$. Wir setzen nochmal $y := f(x)$. Für jedes $\varepsilon > 0$ ist $f^{-1}\big(B_\varepsilon(y)\big)$ offen und enhält x. Deshalb gibt es $\delta > 0$ mit $B_\delta(x) \subset f^{-1}\big(B_\varepsilon(y)\big)$, d.h., $f\big(B_\delta(x)\big) \subset B_\varepsilon(y)$. □

Bemerkung D2.6. Topologie ist der Fachbereich, der sich mit stetigen Abbildungen im Allgemeinen beschäftigt. Diese Eigenschaft, dass die Urbilder von offenen Mengen offen sind, wird da als Definition für Stetigkeit benutzt.

Für $T \subset D$ heißt $f|_T : T \to \mathbb{R}$ die *Beschränkung* von f auf T, definiert durch $f|_T(t) := f(t)$ für alle $t \in T$.

Wie man von der Definition sieht, ist Stetigkeit eine lokale Eigenschaft. Falls f stetig ist, so ist auch die Beschränkung $f|_T$ auf jede Teilmenge $T \subset D$ stetig. Umgekehrt gilt folgendes.

Satz D2.7. *Sei $D = \bigcup_{\alpha \in A} U_\alpha \subset \mathbb{R}$ die Vereinigung der offenen Mengen U_α. Sei $f : D \to \mathbb{R}$ eine Funktion. Falls $f|_{U_\alpha}$ für jedes $\alpha \in A$ stetig ist, dann ist f stetig auf D.*

Beweis. Sei $V \subset \mathbb{R}$ offen. Wir wollen zeigen, dass $f^{-1}(V)$ offen ist. Sei $p \in f^{-1}(V) \subset D$. Dann gibt es $\alpha \in A$ mit $p \in U_\alpha$. Weil $f|_{U_\alpha}$ stetig ist, ist $f|_{U_\alpha}^{-1}(V) = f^{-1}(V) \cap U_\alpha$ offen. Deshalb gibt es $\varepsilon > 0$ mit $B_\varepsilon(p) \subset f^{-1}(V) \cap U_\alpha \subset f^{-1}(V)$. □

Bemerkung D2.8. Es ist hier sehr wichtig, dass die Mengen U_α offen sind. Zum Beispiel setzen wir $A := \{0, \pm 1\}$ mit $U_1 := \{x > 0\}$, $U_0 := \{0\}$, $U_{-1} := \{x < 0\}$. (Hier sind $U_{\pm 1}$ offen, U_0 aber nicht.) Dann gilt $\bigcup U_\alpha = \mathbb{R}$. Die Funktion $f(x) = \operatorname{sgn} x$ ist nicht stetig in 0. Jede der drei Beschränkungen $f|_{U_\alpha}$ ist doch stetig.

Nur dann, wenn T offen ist, impliziert die Stetigkeit von $f|_T$, dass f in jedem $t \in T$ stetig ist.

`Ende der Vorlesung 2008 Mai 20`

D3. Grenzwerte

Definition D3.1. Für $p \in \mathbb{R}$ und $\varepsilon > 0$ heißt

$$B^*_\varepsilon(p) := B_\varepsilon(p) \setminus \{p\} = \{x : 0 < |x - p| < \varepsilon\}$$

der *punktierte* ε*-Ball* um x. Ein Punkt $p \in \mathbb{R}$ heißt ein *Häufungspunkt* von der Teilmenge $D \subset \mathbb{R}$, falls in jedem punktierten Ball um p Punkte von D liegen, d.h., falls

$$\forall \varepsilon > 0 \quad B^*_\varepsilon(p) \cap D \neq \emptyset.$$

Lemma D3.2. *Sei $f\colon D \to \mathbb{R}$ eine Funktion auf D und seien $p, a \in \mathbb{R}$. Wir definieren $g\colon D \cup \{p\} \to \mathbb{R}$ wie folgt:*

$$g(x) := \begin{cases} a, & x = p, \\ f(x), & x \in D,\, x \neq p. \end{cases}$$

(Man könnte g eine Erweiterung bzw. eine Änderung von f nennen, falls $p \notin D$ bzw. $p \in D$.) Falls p kein Häufungspunkt von D ist, dann ist g stetig in p (unabhängig vom Wert $a = g(p)$).

Beweis. Weil p kein Häufungspunkt ist, gibt es $\delta > 0$, so dass $B^*_\delta(p) \cap D = \emptyset$. Das heißt,

$$g\big(B_\delta(p) \cap (D \cup \{p\})\big) = g\big(\{p\}\big) = \{a\},$$

welches natürlich für jedes ε eine Teilmenge von $B_\varepsilon(a)$ ist. □

Definition D3.3. Sei $f : D \to \mathbb{R}$ und sei p ein Häufungspunkt von D. Wir sagen, dass $a \in \mathbb{R}$ der *Grenzwert* von $f(x)$ für $x \to p$ ist, und schreiben

$$\lim_{x \to p} f(x) = a \in \mathbb{R} \quad \text{oder} \quad f(x) \to a \text{ für } x \to p,$$

falls folgendes gilt:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad f\big(B^*_\delta(p) \cap D\big) \subset B_\varepsilon(a).$$

Bemerkung D3.4. Der Punkt p muss nicht im Definitionsbereich D liegen. Auch im Falle $p \in D$ hat der Grenzwert $\lim_{x \to p} f(x)$ gar nichts mit dem Wert $f(p)$ zu tun. Der Grenzwert wird durch Werte in punktierten Bällen definiert, d.h., durch Werte in Punkten nah zu aber nicht gleich p.

Bemerkung D3.5. Sei g die Erweiterung bzw. Änderung von f durch $g(p) := a$ (wie im Lemma oben), wobei p jetzt ein Häufungspunkt von D ist. Die Funktion g ist genau dann in p stetig, wenn $f(x) \to a$ für $x \to p$.

Beispiel D3.6. Sei $D = \{x \in \mathbb{R} : x > 0 \text{ oder } x = -1\}$. Sei $f\colon D \to \mathbb{R}$ die Beschränkung $f = \operatorname{sgn}|_D$ der Signumfunktion auf D. Dann gilt $\lim_{x \to 0} f(x) = 1$. Weil -1 kein Häufungspunkt von D ist, bleibt der Ausdruck $\lim_{x \to -1} f(x)$ undefiniert.

Beispiel D3.7. Sei $f\colon \mathbb{R} \to \mathbb{R}$ stetig, und sei g die Änderung von f, wobei wir $g(0) := a \neq f(0)$ setzen. Dann ist g in 0 unstetig. Es gilt $\lim_{x \to 0} g(x) = f(0)$.

Lemma D3.8. *Sei $f\colon D \to \mathbb{R}$, sei $a \in \mathbb{R}$ und sei p ein Häufungspunkt von D. Es gilt $\lim_{x \to p} f(x) = a$ genau dann, wenn für jede Folge (x_n) in $D \setminus \{p\}$ mit Grenzwert $p \in \mathbb{R}$ gilt $\lim_{n \to \infty} f(x_n) = a$.*

Beweis. Wir nehmen zunächst an, $\lim_{x \to p} f(x) = a$. Sei (x_n) eine Folge in $D \setminus \{p\}$ mit Grenzwert p. Zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ mit

$$f\big(B^*_\delta(p) \cap D\big) = f\big(B_\delta(p) \cap (D \setminus \{p\})\big) \subset B_\varepsilon(a).$$

Weil $x_n \to p$, liegen fast alle x_n in $B_\delta(p)$ und deshalb in $B_\delta(p) \cap (D \setminus \{p\})$. Daher gilt $f(x_n) \in B_\varepsilon(a)$ für fast alle n. Weil ε beliebig war, gilt also $f(x_n) \to a$.

Umgekehrt, falls $\lim_{x \to p} f(x) = a$ *nicht* gilt, wollen wir eine Folge $x_n \to p$ mit $f(x_n) \not\to a$ finden. Dass a kein Grenzwert von f ist, heißt, es gibt $\varepsilon > 0$, so dass für kein $\delta > 0$ gilt

$$f\big(B^*_\delta(p) \cap D\big) \subset B_\varepsilon(a).$$

D.h., zu jedem $n \in \mathbb{N}$ können wir $\delta := {}^1\!/\!{}_n$ nehmen und ein $x_n \in B^*_{1/n}(p) \cap D$ mit $f(x_n) \notin B_\varepsilon(a)$ finden. Dadurch haben wir die Folge (x_n) definiert. Weil $|x_n - p| < {}^1\!/\!{}_n$, gilt $x_n \to p$. Weil $f(x_n) \notin B_\varepsilon(a)$, konvergiert $f(x_n)$ sicherlich nicht gegen a. □

Bemerkung D3.9. Hier am Ende haben wir die archimedische Eigenschaft der reellen Zahlen benutzt. Diese Charakterisierung von Stetigkeit mittels Konvergenz von Folgen gilt zwar in jedem metrischen Raum, nicht aber in jedem topologischen Raum.

Bemerkung D3.10. Eine Funktion $f : D \to \mathbb{R}$ ist genau dann stetig in $p \in D$, wenn für jede Folge (x_n) in D mit $x_n \to p$ gilt $f(x_n) \to f(p)$.

Satz D3.11. *Seien $f\colon D \to E \subset \mathbb{R}$ und $g\colon E \to \mathbb{R}$ Funktionen. Falls f stetig in $p \in D$ ist und g stetig in $f(p) \in E$ ist, dann ist die Komposition $g \circ f$ stetig in p.*

Beweis. $x_n \to p \implies y_n := f(x_n) \to f(p) \implies g(y_n) = g \circ f(x_n) \to g \circ f(p)$. □

Satz D3.12 (Rechenregeln). *Seien $f, g\colon D \to \mathbb{R}$ Funktionen auf $D \subset \mathbb{R}$ und sei p ein Häufungspunkt* *von D. Falls $\lim_{x\to p} f(x)$ und $\lim_{x\to p} g(x)$ existieren, dann gelten*

$$\lim_{x\to p} \big(f \pm g\big)(x) = \lim_{x\to p} f(x) \pm \lim_{x\to p} g(x),$$
$$\lim_{x\to p} \big(f \cdot g\big)(x) = \lim_{x\to p} f(x) \cdot \lim_{x\to p} g(x).$$

Die Funktion $f/_g$ ist auf der Menge $D' := \{x \in D : g(x) \neq 0\}$ definiert. Falls p ein Häufungspunkt von D' ist, gilt

$$\lim_{x\to p} \big(f/g\big)(x) = \lim_{x\to p} f(x) \Big/ \lim_{x\to p} g(x). \qquad \square$$

Korollar D3.13. *Seien $f, g\colon D \to \mathbb{R}$ stetig in $p \in D$. Dann sind auch $f \pm g$ und $f \cdot g$ stetig in p. Falls $g(p) \neq 0$, ist auch $f/_g$ stetig in p.* $\square$

Definition D3.14. Das Polynom $p\colon \mathbb{R} \to \mathbb{R}$ von Grad $n \in \mathbb{N}$ mit Koeffizienten $a_k \in \mathbb{R}$ (wobei $k = 0, \ldots, n$ und $a_n \neq 0$) ist die Funktion

$$p(x) := \sum_{k=0}^{n} a_k x^k.$$

Korollar D3.15. *Jedes Polynom $p\colon \mathbb{R} \to \mathbb{R}$ ist stetig.* $\square$

Definition D3.16. Sei $f : D \to \mathbb{R}$ eine Funktion und sei $p \in \mathbb{R}$ ein Häufungspunkt von D. Falls für jedes $\delta > 0$ gilt

$$\{x \in D : p - \delta < x < p\} \neq \emptyset,$$

sagen wir, p ist ein linksseitiger Häufungspunkt von D. Sei p ein linksseitiger Häufungspunkt. Wir sagen, $a \in \mathbb{R}$ ist der *linksseitige Grenzwert* von $f(x)$ für $x \to p$, falls

$$\forall\varepsilon > 0\ \exists\delta > 0 \quad x \in D,\, p-\delta < x < p \implies |f(x)-a| < \varepsilon.$$

Dazu schreiben wir

$$\lim_{x\to p^-} f(x) = \lim_{x\nearrow p} f(x) = a.$$

Entsprechend werden rechtsseitige Häufungspunkte und der rechtsseitige Grenzwert

$$\lim_{x\to p^+} f(x) = \lim_{x\searrow p} f(x)$$

definiert.

Beispiel D3.17. Die Signumfunktion hat in 0 die einseitigen Grenzwerte

$$\lim_{x\to 0^-} \operatorname{sgn} x = -1, \qquad \lim_{x\to 0^+} \operatorname{sgn} x = +1.$$

Bemerkung D3.18. Sei $f\colon D \to \mathbb{R}$ und sei p ein links- *und* rechtsseitiger Häufungspunkt von D (z.B. ein innerer Punkt von D). Dann existiert $\lim_{x\to p} f(x)$ genau dann, wenn die links- und rechtsseitige Grenzwerte beide existieren und gleich sind:

$$\lim_{x\to p^-} f(x) = \lim_{x\to p^+} f(x).$$

`Ende der Vorlesung 2008 Mai 26`

D4. Die erweiterten reellen Zahlen

a. Definition

Wir wissen, jede beschränkte Folge in $\mathbb{R}$ hat einen Häufungspunkt (d.h., hat eine konvergente Teilfolge). Ähnlich hat jede nichtleere und beschränkte Teilmenge in $\mathbb{R}$ sowohl Supremum als auch Infimum. Um ähnliche Eigenschaften auch für unbeschränkte Folgen und Teilmengen zu haben, ist es oft hilfreich, die reellen Zahlen zu erweitern auf

$$\overline{\mathbb{R}} := \mathbb{R} \cup \{+\infty, -\infty\}.$$

Hier sind $\pm\infty$ zwei zusäztliche Elemente, die keine reelle Zahlen sind. Auf $\overline{\mathbb{R}}$ können wir die totale Ordnung $\leq$ erweitern:

$$-\infty \leq x \leq +\infty \qquad \text{für alle } x \in \overline{\mathbb{R}}.$$

Die arithmetische Verknüpfungen können hingegen nur *teilweise* erweitert werden, insbesondere ist $\overline{\mathbb{R}}$ kein Ring oder Körper. Die Ausdrücke

$$\infty - \infty, \quad \pm\infty \cdot 0, \quad \frac{\pm\infty}{\pm\infty}, \quad \frac{x}{0}$$

bleiben *undefiniert.* Wir setzen

- $x + \infty := +\infty$ für $-\infty < x \in \overline{\mathbb{R}}$,
- $x - \infty := -\infty$ für $+\infty > x \in \overline{\mathbb{R}}$,
- $\pm\infty \cdot x := \pm\infty$ für $0 < x \in \overline{\mathbb{R}}$,
- $\pm\infty \cdot x := \mp\infty$ für $0 > x \in \overline{\mathbb{R}}$,
- $x/_{\pm\infty} := 0$ für $x \in \mathbb{R}$.

Für uns ist eher nur die Ordnung auf $\overline{\mathbb{R}}$ interessant. Mit der Anerkennung, dass manche Ausdrücke undefiniert sind, kann man aber auf $\overline{\mathbb{R}}$ Arithmetik etwa wie gewöhnt durchfürhen. Das Distributivgesetz sagt z.B., $a(b + c) = ab + ac$ im Sinne, dass entweder beide Seiten definiert und gleich sind oder Beide undefiniert sind.

Lemma D4.1. *In der total geordneten Menge $\overline{\mathbb{R}}$ hat jede Teilmenge $T \subset \overline{\mathbb{R}}$ ein Supremum und ein Infimum.*

Beweis. Wir zeigen nur, dass T ein Supremum hat. In jedem Fall ist $+\infty$ eine obere Schranke für T. Falls $+\infty \in T$ ist diese auch die kleinste und $\sup T = +\infty$. Sonst setzen wir $T' := T \cap \mathbb{R} = T \smallsetminus \{-\infty\}$ und behaupten, $\sup T = \sup T'$. Es gilt $\sup T' \in \mathbb{R}$, falls T' nichtleer und nach oben (durch $x \in \mathbb{R}$) beschränkt ist; es gilt $\sup T' = +\infty$, falls T' nach oben unbeschränkt ist; und es gilt $\sup T' = -\infty$, falls T' leer ist. In jedem Fall existiert $\sup T' \in \overline{\mathbb{R}}$. $\square$

Bemerkung D4.2. Es gilt $\sup T = +\infty$ genau dann, wenn T keine obere Schranke $x \in \mathbb{R}$ hat (z.B., wenn $+\infty \in T$). Es gilt $\sup T = -\infty$ genau dann, wenn $T \subset \{-\infty\}$.

Bemerkung D4.3. Wir sagen, $+\infty$ ist ein Häufungspunkt von T, falls $\sup\big(T \smallsetminus \{+\infty\}\big) = +\infty$. Entsprechendes gilt für $-\infty$.

b. Konvergenz

Definition D4.4. Sei (a_n) eine Folge in $\overline{\mathbb{R}}$. Falls für jede reelle Zahl $x \in \mathbb{R}$ gilt $a_n > x$ (bzw. $a_n < x$) für fast alle n, sagen wir $a_n \to +\infty$ (bzw. $a_n \to -\infty$).

Lemma D4.5. *Jede monotone Folge in $\overline{\mathbb{R}}$ hat einen Grenzwert in $\overline{\mathbb{R}}$.*

Beweis. Sei (a_n) monoton steigend. Falls (a_n) die konstante Folge $a_n = -\infty$ ist, konvergiert sie gegen $-\infty$. Sonst gilt wegen der Monotonie $a_n > -\infty$ für fast alle n. Falls (a_n) durch $x \in \mathbb{R}$ beschränkt ist, hat sie einen Grenzwert in $\mathbb{R}$. Sonst gibt es zu jedem $x \in \mathbb{R}$ ein $n_0 \in \mathbb{N}$ mit $a_{n_0} > x$. Weil (a_n) monoton ist, gilt dann $a_n > x$ für alle $n \geq n_0$. Das heißt (per Definition), dass $a_n \to +\infty$. □

Korollar D4.6. *Jede Folge in $\overline{\mathbb{R}}$ hat eine konvergente Teilfolge.*

Beweis. Wir haben schon gezeigt, jede Folge hat eine monotone Teilfolge. Vom Lemma konvergiert diese. □

Für eine reelle Funktion $f\colon D \to \mathbb{R}$ können wir nun unendliche Grenzwerte und Grenzwerte in Unendlich definieren. Sei $p \in \overline{\mathbb{R}}$ ein Häufungspunkt von D. Das heißt, es gibt (mindestens) eine Folge (x_n) in $D \setminus \{p\}$ mit Grenzwert p. Dann sagen wir $\lim_{x\to p} f(x) = a \in \overline{\mathbb{R}}$, falls für *jede* solche Folge (x_n) gilt $\lim f(x_n) = a$.

Falls $p, a \in \mathbb{R}$, haben wir schon gezeigt, diese Definition stimmt mit der Ursprunglichen üeberein.

Beispiel D4.7. Für $k \in \mathbb{N}^+$ betrachten wir die Funktion $x \mapsto x^k$. Es gilt $\lim_{x\to+\infty} x^k = +\infty$. (Falls $x_n \to +\infty$ gilt $x_n^k \to +\infty$, weil $x_n^k \geq x_n$ für $x_n \geq 1$.)

Beispiel D4.8. Es gilt $\lim_{x\to 0} 1/x^2 = +\infty$.

Wir können auch einseitigen Grenzwerten definieren. Zum Beispiel gilt

$$\lim_{x\to 0^+} 1/x^3 = +\infty, \qquad \lim_{x\to 0^-} 1/x^3 = -\infty,$$

was bedeutet, $\lim_{x\to 0} 1/x^3$ existiert nicht.

Bemerkung D4.9. Die Rechenregeln für

$$\lim_{x\to p} \big(f \pm g\big)(x), \quad \lim_{x\to p} \big(f \cdot g\big)(x)$$

gelten auch dann, wenn p bzw. die Grenzwerte unendlich sind. Dabei muss man aber die Existenz voraussetzen nicht nur von den beiden Grenzwerten $\lim f$ und $\lim g$ sondern auch von deren Summe bzw. Differenz bzw. Produkt. Es gilt auch

$$\lim_{x\to p} f(x) = 0 \iff \lim_{x\to p} \left|\frac{1}{f(x)}\right| = +\infty.$$

Beispiel D4.10. Seien $f(x) := x$ und $g(x) := x + c$. Wir betrachten Grenzwerte für $x \to +\infty$. Es gelten $f \to +\infty$ und $g \to +\infty$. Aus den Rechenregeln können wir $f + g \to +\infty$ und $fg \to +\infty$ schließen. Die Regeln geben aber keine Auskunft über die Grenzwerte von $f - g$ und f/g. (Natürlich können wir auf andere Weise zeigen, dass $f - g \to c$ und $f/g \to 1$.)

c. Limes superior und inferior

Sei (a_n) eine Folge in $\mathbb{R}$. Wir definieren die monoton fallende Folge $s_n := \sup\{a_k : k \geq n\}$. Dann existiert immer der *Limes superior* von (a_n),

$$\overline{\lim_{n\to\infty}}\, a_n := \limsup_{n\to\infty} a_n := \lim_{n\to\infty} s_n \in \overline{\mathbb{R}}.$$

Entsprechend definiert man den *Limes inferior*,

$$\underline{\lim}_{n\to\infty}\, a_n = \liminf_{n\to\infty} a_n \in \overline{\mathbb{R}}.$$

Bemerkung D4.11. Sei $H \subset \overline{\mathbb{R}}$ die Menge aller Häufungspunkte von (a_n). Dann gelten $\underline{\lim}\, a_n = \inf H$ und $\overline{\lim}\, a_n = \sup H$. (Weil diese zu H gehören, kann man auch $\min H$ und $\max H$ schreiben.) Die Folge (a_n) hat einen Limes genau dann, wenn

$$\underline{\lim}\, a_n = \overline{\lim}\, a_n.$$

Bemerkung D4.12. Ähnlich definiert man lim sup für reelle Funktionen. Sei $f\colon D \to \mathbb{R}$ und p ein Häufungspunkt von D. Dann gilt per Definition

$$\overline{\lim_{x\to p}}\, f(x) := \lim_{\delta\to 0^+} \sup f\big(B_\delta(p) \cap D\big).$$

Zum Beispiel gelten

$$\underline{\lim}_{x\to 0} \sin 1/x = -1, \qquad \overline{\lim_{x\to 0}} \sin 1/x = +1.$$

Es existiert $\lim_{x\to p} f(x)$ genau dann, wenn

$$\underline{\lim}_{x\to p}\, f(x) = \overline{\lim_{x\to p}}\, f(x).$$

`Ende der Vorlesung 2008 Mai 27`

D5. Intervalle und Monotonie

a. Intervalle

Definition D5.1. Sei X eine total geordnete Menge und seien $a, b \in X$. Wir definieren folgende Teilmengen in X:

$$\begin{aligned}
[a, b] &:= \{x \in X : a \leq x \leq b\},\\
(a, b) &:= \{x \in X : a < x < b\},\\
[a, b) &:= \{x \in X : a \leq x < b\},\\
(a, b] &:= \{x \in X : a < x \leq b\}.
\end{aligned}$$

Der interessanteste Fall ist $a < b$. Für $a \geq b$ gilt nämlich $(a, b) = [a, b) = (a, b] = \emptyset$ und für $a > b$ auch $[a, b] = \emptyset$. Für $a = b$ gilt $[a, a] = \{a\}$, eine einelementige Menge.

Bemerkung D5.2. Die französische Schreibweise – wo man $]a, b[$ für (a, b) usw. schreibt – hat sich unter Mathematikern nie so richtig durchgesetzt.

Definition D5.3. Sei X eine total geordnete Menge. Eine Teilmenge $I \subset X$ heißt ein *Intervall*, falls

$$x, y \in I \implies [x, y] \subset I.$$

(Das heißt, T enthält mit je zwei Punkten auch alle dazwischen liegenden.)

Lemma D5.4. *Sei $I \subset X$ ein Intervall, welches ein Infimum $a := \inf I$ und ein Supremum $b := \sup I$ besitzt. Falls $a > b$, gilt $I = \emptyset$, und falls $a = b$, gilt $I = [a, a] = \{a\}$. Falls $a < b$, gilt entweder*

$$I = [a, b], \quad I = [a, b), \quad I = (a, b], \quad \textit{oder} \quad I = (a, b),$$

abhängig davon, ob a bzw. b Elemente von I sind oder nicht.

Beweis. Auch ohne die Bedingung, dass I ein Intervall ist, gilt erstens $\sup I < \inf I$ nur dann, wenn $I = \emptyset$, und zweitens $\sup I = \inf I$ nur dann, wenn I eine einelementige Menge ist.

Sei jetzt $a := \inf I < \sup I =: b$. Weil a bzw. b Schranken für I sind, gilt $I \subset [a, b]$. Weil I ein Intervall ist, zeigen wir zunächst, dass $(a, b) \subset I$: Sei $x \in (a, b)$, d.h., $a < x < b$. Es gibt $p \in I$ mit $a \le p < x$, sonst wäre x eine untere Schranke grösser als a. Ähnlich gibt es $q \in I$ mit $x < q \le b$. Weil $p, q \in I$ und I ein Intervall ist, gilt $x \in [p, q] \subset I$.

Wir haben $(a, b) \subset I \subset [a, b]$. Es bleiben nur die vier aufgelisteten Möglichkeiten. □

Bemerkung D5.5. Jedes Intervall $I \subset \overline{\mathbb{R}}$ hat Supremum und Infimum, was heißt, das Lemma gilt für I.

Definition D5.6. Seien $a, b \in \overline{\mathbb{R}}$. Das Intervall $(a, b) \subset \mathbb{R}$ nennt man ein *offenes Intervall*, weil es eine offene Teilmenge ist. Dan Intervall $[a, b]$ nennt man ein *abgeschlossenes Intervall*. (Im Fall $a, b \in \mathbb{R}$ ist es eine abgeschlossene Teilmenge von $\mathbb{R}$. Die Halbgeraden $[a, +\infty)$ und $(-\infty, b]$ und die Gerade $\mathbb{R} = (-\infty, +\infty)$ sind zwar auch abgeschlossene Teilmengen aber keine abgeschlossene Intervalle.) Die Intervalle $(a, b]$ oder $[a, b)$ nennt man *halboffen*.

b. Monotonie

Definition D5.7. Seien X und Y total geordnete Mengen. Eine Abbildung $f : X \to Y$ heißt dann

- *monoton steigend*, wenn $f(x) \le f(x')$,
- *monoton fallend*, wenn $f(x) \ge f(x')$,
- *streng monoton steigend*, wenn $f(x) < f(x')$,
- *streng monoton fallend*, wenn $f(x) > f(x')$,

jeweils für alle $x < x' \in X$.

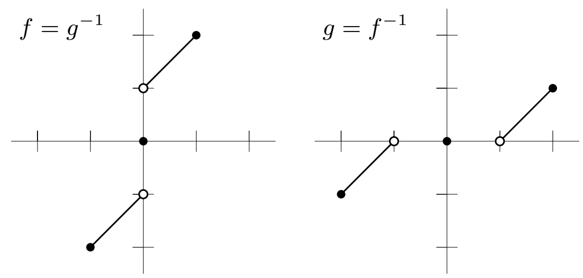
Bemerkung D5.8. Im Fall $X = \mathbb{N}$ ist f eine Folge in Y. Diese Definition stimmt mit der Bisherigen überein.

Lemma D5.9. *Eine streng monotone Abbildung Sei $f : X \to Y$ eine streng monotone Abbildung. Dann ist f injektiv. Auf $f(X) \subset Y$ gibt es also eine Umkehrabbildung*

$$f^{-1} : f(X) \to X.$$

Diese ist auch streng monoton. □

Beispiel D5.10. Sei $I := [-1, 1]$ das abgeschlossene Intervall. Sei die streng monoton steigende Funktion $f : I \to \mathbb{R}$ durch $f(x) := x + \operatorname{sgn} x$ definiert. In jedem Punkt $p \ne 0$ ist f stetig, in 0 jedoch nicht. Das Bild von f ist $J := f\big([-1, 1]\big) = [-2, -1) \cup \{0\} \cup (1, 2]$. Die Umkehrfunktion $g := f^{-1} : J \to \mathbb{R}$ ist stetig und streng monoton steigend. Wir haben $g(x) = x - \operatorname{sgn} x$ für alle $x \in J$. (Die Formel $x - \operatorname{sgn} x$ definiert auch eine Funktion $h : \mathbb{R} \to \mathbb{R}$, die weder stetig noch monoton ist.)



Der folgende Satz zeigt, dass die Stetigkeit von g zu erwarten ist, weil I ein Intervall ist. Umgekehrt können wir den Satz auf g nicht anwenden, weil T kein Intervall ist. Deshalb darf $f = g^{-1}$ unstetig sein.

Satz D5.11. *Sei $I \subset \mathbb{R}$ ein Intervall und sei $f : I \to \mathbb{R}$ streng monoton. Die Umkehrfunktion $f^{-1} : f(I) \to I \subset \mathbb{R}$ ist dann stetig.*

Bemerkung D5.12. Die Funktion f muss nicht stetig sein.

Beweis. Sei f (o.B.d.A.) streng monoton steigend. Sei $q = f(p) \in f(I)$. Um zu zeigen, dass f^{-1} in q stetig ist, sei $\varepsilon > 0$ gegeben. Wir müssen $\delta > 0$ finden, so dass für alle $y = f(x) \in f(I)$ gilt

$$|y - q| < \delta \implies |x - p| = \big|f^{-1}(y) - f^{-1}(q)\big| < \varepsilon.$$

Sei p zunächst ein innerer Punkt von I. Dann gibt es $x_\pm \in I$ mit $p - \varepsilon < x_- < p < x_+ < p + \varepsilon$. Weil f streng monoton steigend ist, gilt $f(x_-) < q < f(x_+)$. Wir setzen

$$\delta := \min\big(q - f(x_-), f(x_+) - q\big).$$

Dann gilt (wie gewünscht)

$$\begin{aligned} |y - q| < \delta &\implies y \in \big(f(x_-), f(x_+)\big) \\ &\implies x \in (x_-, x_+) \implies |x - p| < \varepsilon. \end{aligned}$$

Falls p ein Endpunkt von I ist, benutzen wir $[p, x_+)$ bzw. $(x_-, p]$ bzw. $[p, p]$ anstelle von (x_-, x_+). (Hier ist es wichtig, dass I ein Intervall ist.) □

Beispiel D5.13. Sei $k \in \mathbb{N}^+$. Die Funktion $x \mapsto x^k$ ist eine streng monoton steigende Abbildung $[0, +\infty) \to [0, +\infty)$. Die Umkehrfunktion schreibt man $x \mapsto \sqrt[k]{x}$ oder $x \mapsto x^{1/k}$. Sie ist auch stetig und streng monoton steigend.

`Ende der Vorlesung 2008 Juni 2`

D6. Zwischenwertsatz

Satz D6.1 (Zwischenwertsatz). *Sei $I \subset \mathbb{R}$ ein Intervall und sei $f\colon I \to \mathbb{R}$ stetig auf I. Dann ist das Bild $f(I)$ ein Intervall. Das heißt, f nimmt jeden Wert an, der zwischen zwei Funktionswerte $f(a)$ und $f(b)$ liegt.*

Beweis. Sei y ein dazwischen liegender Wert, $f(a) < y < f(b)$. Wir nehmen (o.B.d.A.) an, $a < b$. Sei $S := \big\{x \in [a, b] : f(x) \le y\big\}$. Die Teilmenge $S \subset \mathbb{R}$ ist nach oben (durch b) beschränkt und ist nicht leer (weil $a \in S$). Damit existiert $c := \sup S \in [a, b]$.

Wir behaupten, $f(c) = y$. Falls nicht, setzen wir $\varepsilon := |f(c) - y| > 0$. Weil f in c stetig ist, gibt es dazu ein $\delta > 0$, so dass für $|x - c| < \delta$ gilt $\big|f(x) - f(c)\big| < \varepsilon$. Im Fall $f(c) < y$ heißt das, $f(x) < y$ für $x \in B_\delta(c)$. Daher liegt $x \in S$ und c wäre kein obere Schranke. Im Fall $f(c) > y$ heißt es, $f(x) > y$ für $x \in B_\delta(c)$.. Damit wäre $c - \delta$ eine kleinere obere Schranke. □

Korollar D6.2. *Sei $f : [a, b] \to \mathbb{R}$ eine stetige Funktion mit $f(a)f(b) < 0$. Dann hat f eine Nullstelle $x \in [a, b]$, d.h., $f(x) = 0$.* □

Korollar D6.3 (Fixpunkt). *Seien $a \le b \in \mathbb{R}$ und sei $f\colon [a, b] \to [a, b]$ stetig. Dann hat f einen Fixpunkt $x = f(x)$.*

Beweis. (Aufgabe.) □

Bemerkung D6.4. Die Verallgemeinerung auf höhere Dimensionen heißt der Fixpunktsatz von Brouwer: Jede Selbstabbildung auf $[0, 1]^n$ hat einen Fixpunkt.

Korollar D6.5. *Sei $f\colon \mathbb{R} \to \mathbb{R}$ stetig und beschränkt. Dann hat f einen Fixpunkt.*

Beweis. Sei f durch $\pm C$ beschränkt. Dann ist mit f auch ihre Beschränkung auf $[-C, C]$ eine Abbildung in $[-C, C]$. □

Bemerkung D6.6. Sei K ein Kreis. Zu jedem punkt $x \in K$ gibt es einen *antipodischen* (genau gegenüber liegenden) Punkt x^*. Sei $f\colon K \to \mathbb{R}$ stetig. (Ein Beispiel einer stetigen Funktion auf einem Kreis wäre die jetzige Temperatur entlang dem Äquator.) Wenn wir $g(x) := f(x) - f(x^*)$ als periodische Funktion $\mathbb{R} \to \mathbb{R}$ betrachten, sehen wir, g hat eine Nullstelle, d.h., es gibt x mit $f(x) = f(x^*)$.

Der Satz von Borsuk–Ulam ist eine eine Verallgemeinerung auf höhere Dimensionen, die wichtig in der Topologie ist. Seien f und g zwei stetige Funktionen (z.B. Temperatur und Luftfeuchtigkeit) auf der Erdoberfläche. Dann gibt es antipodische Punkte x, x^* mit $f(x) = f(x^*)$ und $g(x) = g(x^*)$.

Satz D6.7. *Eine injektive und stetige Funktion auf einem Intervall ist streng monoton.*

Definition D6.8. Seien $a, b \in \mathbb{R}$. Die *lineare Interpolation* zwischen a und b ist die Funktion $\ell_a^b\colon [0, 1] \to \mathbb{R}$ mit $\ell_a^b(t) := (1 - t)a + tb$. Damit gelten $\ell_a^b(0) = a$ und $\ell_a^b(1) = b$. Das Bild $\ell_a^b\big([0, 1]\big)$ ist das abgeschlossene Intervall mit Endpunkten a und b.

Beweis. Sei $f\colon I \to \mathbb{R}$ stetig auf dem Intervall I. Wir nehmen an, f ist nicht streng monoton, und beweisen, f ist nicht injektiv. Dass f nicht streng mononton fallend ist, heißt, es gibt $a < a' \in I$ mit $f(a) \le f(a')$. Wir dürfen annehmen, $f(a) < f(a')$, weil sonst f offensichtlich nicht injektiv ist. Weil f nicht streng monoton fallend ist, gibt es ähnlich auch $b < b' \in I$ mit $f(b) \ge f(b')$. Wir interpolieren zwischen diesen beiden Paaren: sei $g\colon [0, 1] \to \mathbb{R}$ die Funktion

$$\begin{aligned} g(t) &:= f\big(\ell_a^b(t)\big) - f\big(\ell_{a'}^{b'}(t)\big) \\ &= f\big((1-t)a + tb\big) - f\big((1-t)a' + tb'\big). \end{aligned}$$

Es gelten

$$g(0) = f(a) - f(a') \le 0, \qquad g(1) = f(b) - f(b') \ge 0.$$

Also nach dem Zwischenwertsatz gibt es $t \in [0, 1]$ mit $g(t) = 0$, was heißt, $f\big(\ell_a^b(t)\big) = f\big(\ell_{a'}^{b'}(t)\big)$. Es gilt aber $\ell_a^b(t) < \ell_{a'}^{b'}(t)$, weil $a < a'$ und $b < b'$. Damit ist f nicht injektiv. □

D7. Kompaktheit

a. *Definition*

Lemma D7.1. *Seien $T \subset \mathbb{R}$ und $x \in \mathbb{R}$. Es gibt genau dann eine Folge (a_n) in T mit $a_n \to x$, wenn $x \in T \cup \partial T$, d.h., wenn x kein innerer Punkt von $\mathbb{R} \smallsetminus T$ ist.*

Beweis. Falls x ein innerer Punkt von $\mathbb{R} \smallsetminus T$ ist, dann gibt es $\varepsilon > 0$ mit $B_\varepsilon(x) \subset \mathbb{R} \smallsetminus T$. Eine Folge (a_n) in T bleibt ausserhalb $B_\varepsilon(x)$ und kann nicht gegen x konvergieren.

Umgekehrt, falls x kein innerer Punkt von $\mathbb{R} \smallsetminus T$ ist, gibt es zu jedem $n \in \mathbb{N}$ (mit $\varepsilon = 1/n$) ein $a_n \in T$, so dass $|a_n - x| < 1/n$. Dann konvergiert $a_n \to x$. □

Korollar D7.2. *Eine Teilmenge $T \subset \mathbb{R}$ ist genau dann abgeschlossen, wenn der Grenzwert zu jeder konvergenten Folge in T auch in T liegt.*

Beweis. Vom Lemma sind die möglichen Grenzwerte genau die Punkte in $T \cup \partial T$. Diese gleicht T genau dann, wenn T abgeschlossen ist. □

Definition D7.3. Eine Teilmenge $T \subset \mathbb{R}$ heißt *folgenkompakt*, falls jede Folge (a_n) in T eine konvergente Teilfolge hat, deren Grenzwert in T liegt.

Satz D7.4 (Bolzano–Weierstraß). *Eine Teilmenge $T \subset \mathbb{R}$ ist genau dann folgenkompakt, wenn sie abgeschlossen und beschränkt ist.*

Beispiel D7.5. Seien $a, b \in \mathbb{R}$. Das abgeschlossene Intervall $[a, b]$ ist folgenkompakt. Deshalb nennt man es auch ein *kompaktes Intervall*.

Beweis. Sei $T \subset \mathbb{R}$ abgeschlossen und beschränkt und sei (a_n) eine Folge in T. Als beschränkte reelle Folge hat sie eine Teilfolge (a_{n_k}), die gegen $x \in \mathbb{R}$ konvergiert. Weil T abgeschlossen ist liegt $x \in T$ (vom zweiten Lemma).

Umgekehrt betrachten wir zwei Fälle. Erstens, sei T (o.B.d.A. nach oben) unbeschränkt. Zu jedem $n \in \mathbb{N}$ wählen wir $a_n \in T$ mit $a_n > n$. Diese Folge hat keine konvergente Teilfolge. Zweitens, sei T nicht abgeschlossen. Dann gibt es $x \in \partial T \smallsetminus T$ und (vom ersten Lemma) eine Folge (a_n) in T, die gegen x konvergiert. Jede Teilfolge konvergiert auch gegen $x \notin T$. □

`Ende der Vorlesung 2008 Juni 3`

---

b. *Stetige Funktionen auf Kompakten Mengen*

Wir betrachten jetzt ein paar Sätze über eine stetige Funktion $f\colon K \to \mathbb{R}$, deren Definitionsbereich $K \subset \mathbb{R}$ folgenkompakt ist. Weil der wichtigste Fall $K = [a, b]$ ist, wird in manchen Büchern nur diesen Fall erwähnt.

Satz D7.6. *Sei $K \subset \mathbb{R}$ folgenkompakt und sei $f\colon K \to \mathbb{R}$ stetig. Dann ist $f(K) \subset \mathbb{R}$ auch folgenkompakt.*

Beweis. Sei (b_n) eine Folge in $f(K)$. Wir wählen $a_n \in K$ mit $f(a_n) = b_n$. Weil K folgenkompakt ist, hat (a_n) eine Konvergente Teilfolge $a_{n_k} \to x \in K$. Weil f stetig ist, gilt $f(a_{n_k}) \to f(x)$, d.h., die Teilfolge (b_{n_k}) konvergiert in $f(K)$. □

Der nächste Satz stammt 1861 von Karl Weierstraß.

Satz D7.7. *Sei $\emptyset \neq K \subset \mathbb{R}$ folgenkompakt und sei $f\colon K \to \mathbb{R}$ stetig. Dann nimmt f auf K sein Minimum und Maximum an, insbesondere ist f auf K beschränkt.*

Beweis. Vom ersten Satz ist die Menge $f(K)$ folgenkompakt, d.h., abgeschlossen und beschränkt. Weil sie nicht leer und beschränkt ist, gilt $\sup f(K) \in \partial f(K) \subset \mathbb{R}$. Weil sie abgeschlossen ist, ist dann $\sup f(K) \in f(K)$ ein Maximum. Entsprechendes gilt auch für $\inf f(K) \in f(K)$. □

Definition D7.8. Eine Funktion $f\colon D \to \mathbb{R}$ heißt dann *gleichmäßig stetig* auf D, falls

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in D \quad \forall y \in D$$
$$|x - y| < \delta \implies |f(x) - f(y)| < \varepsilon.$$

Bemerkung D7.9. Das f *stetig* auf D ist, heißt dasselbe – nur mit "$\exists \delta > 0$" und "$\forall x \in D$" getauscht. Da darf δ von x abhängen. Hier muss δ unabhängig von x (oder "gleichmäßig" für alle x) gewählt werden können. (Natürlich hängt δ immer noch von ε ab.)

Bemerkung D7.10. Wie wichtig der Begriff "gleichmäßig stetig" eigentlich ist, sehen wir erst später, z.B. bei der Integralrechnung.

Beispiel D7.11. Sei $f\colon (0, +\infty) \to \mathbb{R}$ die Funktion $x \mapsto {}^1\!/_x$. Dann ist f auf $(0, +\infty)$ zwar stetig aber nicht gleichmäßig stetig. Für jedes $a > 0$ ist die Beschränkung $f|_{[a,+\infty)}$ gleichmäßig stetig (auf $[a, +\infty)$).

Satz D7.12. *Sei $K \subset \mathbb{R}$ folgenkompakt und sei $f\colon K \to \mathbb{R}$ stetig. Dann ist f gleichmäßig stetig auf K.*

Beweis. Wir nehmen an, f ist nicht gleichmäßig stetig. Das heißt, es gibt $\varepsilon > 0$, so dass kein gleichmäßiges δ existiert. Insbesondere, für jedes $n \in \mathbb{N}^+$ geht es mit $\delta = {}^1\!/_n$ nicht, d.h., es gibt $x_n, y_n \in K$ mit

$$|x_n - y_n| < {}^1\!/_n \quad \text{aber} \quad \big|f(x_n) - f(y_n)\big| \geq \varepsilon.$$

Weil K folgenkompakt ist, gibt es eine konvergente Teilfolge $x_{n_k} \to p \in K$. Weil

$$|x_{n_k} - y_{n_k}| < {}^1\!/_{n_k} \to 0,$$

konvergiert auch $y_{n_k} \to p$. Weil f stetig ist, gilt

$$\lim_{k\to\infty} f(x_{n_k}) = f(p) = \lim_{k\to\infty} f(y_{n_k}).$$

Dieses widerspricht $|f(x_n) - f(y_n)| \geq \varepsilon$. □

Bemerkung D7.13. In der Topologie dient Kompaktheit oft als Ersatz für Endlichkeit. (Eine Funktion auf eine endliche Menge hat trivialerweise ein Maximum. Kompaktheit des Definitionsbereiches reicht aus, dasselbe zu schließen.) Man kann auch sagen, Kompaktheit erlaubt uns, aus lokaler Information (Stetigkeit) globale Eigenschaften (Existenz eines Maximums bzw. gleichmäßige Stetigkeit) herzuleiten.

E. ABLEITUNGEN

E1. Differenzierbare Funktionen

Wir glauben, die Welt ist deterministisch. (Oder: auch wenn wir dies philosophisch nicht glauben, machen wir sowieso in den Naturwissenschaften als ob ....) Deshalb haben die Gesetze der Naturwissenschaften oft folgende Form: die Änderungen eines Systems werden als Funktion des jetzigen Zustands beschrieben. Das Newton'sche Gesetz sagt, z.B., dass eine Kraft auf einem Objekt bewirkt eine Änderung dessen Geschwindigkeit.

Mathematisch werden solche Änderungen durch Ableitungen von Funktionen betrachtet. Geschwindigkeit ist die Ableitung der Position (bezüglich der Zeit); Beschleunigung ist die Ableitung der Geschwindigkeit. Das heißt, die allermeisten physikalischen Gesetze werden als Differentialgleichungen geschrieben.

Die Stetigkeit einer Funktion f in einem Punkt p kann man wie folgt verstehen: die konstante Funktion $x \mapsto f(p)$ ist eine gute Approximation für f. Genauer ausgedrückt, wie klein auch immer die Fehlertoleranz (ε) sein mag, gibt es immer eine hinreichend kleine (δ-) Umgebung von p, auf der die Funktionswerte $f(x)$ innerhalb der Toleranz um die Konstante $f(p)$ liegen.

Eine Konstante ist ein Polynom von Grad 0. Wir erwarten, dass viele Funktionen noch besser approximierbar sind, wenn wir Polynome von höherem Grad erlauben. Zunächst betrachten wir Polynome vom Grad (kleiner gleich) 1, d.h. (affin) lineare Funktionen der Form $g\colon x \mapsto ax + b$.

Natürlich legen wir jetzt einen höheren Standard für die Approximation an. Für $|h| < \delta$ verlangen wir, dass die Differenz $g(p+h) - f(p+h)$ nicht nur kleiner als ε sondern kleiner als $\varepsilon\,|h|$ ist. Der Graph von g – eine Gerade – ist die Tangente zum Graphen von f im Punkt $\big(p, f(p)\big)$. Dessen Steigung a heißt die Ableitung von f in p.

Um diese näher zu betrachten machen wir eine allgemeine Voraussetzung für die kommenden Wochen: sei $I \subset \mathbb{R}$ ein Intervall mit mehr als einem Punkt. Wir betrachten Funktionen $f\colon I \to \mathbb{R}$ und deren Eigenschaften um einen Punkt $p \in I$.

Definition E1.1. Die Funktion f heißt *differenzierbar* in p, falls

$$f'(p) := \lim_{x\to p} \frac{f(x) - f(p)}{x - p} = \lim_{h\to 0} \frac{f(p+h) - f(p)}{h} \in \mathbb{R}$$

existiert. Dann heißt $f'(p)$ die *Ableitung* von f in p.

Beispiel E1.2. Sei $f\colon x \mapsto ax+b$ eine affin lineare Funktion. In diesem Fall ist jeder *Differenzenquotient* $\frac{f(x)-f(p)}{x-p}$ gleich die Steigung a. Das heißt, der Limes existiert immer, und $f'(p) = a$ für alle p.

Bemerkung E1.3. Man schreibt auch

$$\frac{df(x)}{dx} := f'(x)$$

für die Ableitung. Diese Schreibweise ist besonders bequem, wenn eine Funktion (z.B. $x \mapsto x^3$) noch keinen Namen (etwa f) hat. Man schreibt auch manchmal ${}^{df}\!/_{dx}$, obwohl diese (streng gesehen) nicht konsequent ist.

Beispiel E1.4. Die Abbildung $x \mapsto x^3$ ist in jedem Punkt $x \in \mathbb{R}$ differenzierbar mit ${}^{dx^3}\!/_{dx} = 3x^2$. Es gilt nämlich

$$\begin{aligned}\lim_{h\to 0} \tfrac{1}{h}\big((x+h)^3 - x^3\big) &= \lim_{h\to 0} \tfrac{1}{h}\big(3x^2h + 3xh^2 + h^3\big)\\ &= \lim_{h\to 0}\big(3x^2 + 3xh + h^2\big) = 3x^2.\end{aligned}$$

Lemma E1.5. *Falls f differenzierbar in p ist, dann ist f stetig in p.*

Beweis. Für $x \to p$ gilt

$$f(x) - f(p) = \Big(\frac{f(x) - f(p)}{x - p}\Big)\cdot(x-p) \to f'(p)\cdot 0 = 0. \qquad \square$$

Beispiel E1.6. Die stetige Funktion $f\colon x \mapsto |x|$ ist in jedem Punkt $x \neq 0$ differenzierbar mit $f'(x) = \operatorname{sgn} x$. Sie ist aber in 0 nicht differenzierbar, weil der Differenzenquotient ${}^{(|h|-0)}\!/_h$ gleich $\operatorname{sgn} h$ ist. In diesem Fall gibt es links- und rechtsseitige Ableitungen

$$f'_\pm(0) := \lim_{h\to 0^\pm} {}^{|h|}\!/_h = \lim_{h\to 0^\pm} \operatorname{sgn} h = \pm 1.$$

Hier kann man sagen, $f'(0)$ existiert nicht, weil die einseitigen Ableitungen nicht gleich sind. (Die Ableitung $f'(x)$, die für $x \neq 0$ definiert ist, hat ein Sprung um $x = 0$.)

Beispiel E1.7. Die Funktion $x \mapsto \sqrt[3]{x}$ ist differenzierbar in jedem $x \neq 0$ aber nicht in $x = 0$. Hier existiert zwar

$$\lim_{x\to 0} \frac{\sqrt[3]{x} - 0}{x - 0} = +\infty \in \overline{\mathbb{R}},$$

die Definition erlaubt aber nicht, dass eine Ableitung unendlich ist.

Bemerkung E1.8. In der Physik gibt es viele Größen, die abhängig von der Zeit sind. Normalerweise schreibt man t für die Zeit (*time* auf Englisch). Die Ableitung einer Funktion $f(t)$ schreibt man meist mit einem Punkt anstelle eines Striches:

$$\dot f(t) := \frac{df(t)}{dt}.$$

Satz E1.9 (Rechenregeln). *Seien $f, g\colon I \to \mathbb{R}$ differenzierbar in $p \in I$. Dann sind auch $f \pm g$ und $f \cdot g$ in p differenzierbar und es gelten*

$$\begin{aligned}(f \pm g)'(p) &= f'(p) \pm g'(p),\\ (f\cdot g)'(p) &= f'(p)g(p) + f(p)g'(p).\end{aligned}$$

Falls $g(p) \neq 0$, ist auch ${}^{f}\!/_{g}$ in p differenzierbar mit Ableitung

$$\Big(\frac{f}{g}\Big)'(p) = \frac{f'(p)g(p) - f(p)g'(p)}{\big(g(p)\big)^2}. \qquad \square$$

Bemerkung E1.10. Vielleicht ist f/g nicht auf dem ganzen Intervall I definiert. Weil aber g in p stetig ist und $g(p) \neq 0$, gibt es ein Intervall J mit $p \in J \subset I$, auf dem g keine Nullstelle hat. Genauer gesagt sollten wir oben f/g als Funktion $J \to \mathbb{R}$ betrachten.

`Ende der Vorlesung 2008 Juni 9`

Beispiel E1.11. Die Ableitung eines Polynoms $P(x) := \sum a_k x^k$ ist das Polynom $P'(x) = \sum k a_k x^{k-1}$.

Bemerkung E1.12. Wenn man sehr vorsichtig ist und es zur Kenntnis nimmt, dass die „Gleichungen“ unten gar keine Bedeutung haben bis man etwa „durch dx teilt“, kann man als Kürzel oder Gedächtnishilfe die Rechenregeln so umformulieren:

$$\begin{aligned} d(u \pm v) &= du \pm dv, \\ d(uv) &= u\,dv + v\,du, \\ d(u/v) &= (v\,du - u\,dv)/v^2. \end{aligned}$$

Lemma E1.13. *Sei $0 \in I \subset \mathbb{R}$ und sei $f\colon I \to \mathbb{R}$ eine Funktion, für die $\bigl|f(x)\bigr| \leq x^2$ gilt. Dann ist f in 0 differenzierbar mit $f'(0) = 0$.*

Beweis. Es gilt $f(0) = 0$. Die Differenzenquotienten um 0 haben die Form $f(h)/h$. Es gilt $\bigl|f(h)/h\bigr| \leq |h| \to 0$. □

Beispiel E1.14. Die Abbildung

$$x \mapsto \begin{cases} x^2, & x \in \mathbb{Q}, \\ 0, & x \notin \mathbb{Q} \end{cases}$$

ist nur in dem einen Punkt $x = 0$ stetig. Vom Lemma ist sie dort auch differenzierbar (mit Ableitung 0).

Beispiel E1.15. Benutzen wir im Voraus den Cosinus, den wir erst später definieren, dann finden wir, die Funktion $f\colon x \mapsto x^2 \cos 1/x$, $x \neq 0$ ist mit $f(0) := 0$ stetig fortsetzbar. Die fortgesetzte f ist in jedem Punkt $x \in \mathbb{R}$ differenzierbar . Insbesondere gelten $f'(x) = \sin 1/x + 2x \cos 1/x$ für $x \neq 0$ (wie wir später erfahren) und $f'(0) = 0$ (vom Lemma). Die Abbildung $x \mapsto f'(x)$ ist damit in $x = 0$ unstetig.

Beispiel E1.16. Sei g die stetige Sägezahnfunktion

$$g\colon x \mapsto \Bigl|x - \lfloor x \rfloor - 1/2\Bigr|.$$

Sie ist differenzierbar in x genau dann, wenn $2x \notin \mathbb{Z}$. Sei

$$f(x) := \sum_{j=0}^{\infty} 2^{-j} g(2^j x).$$

Weil g beschränkt ist ($0 \leq g \leq 1/2$) konvergiert diese Reihe für jedes x. Damit ist f eine (periodische) Funktion $\mathbb{R} \to [0, 1]$. Man kann zeigen, f ist stetig auf $\mathbb{R}$ aber f ist in *keinem* Punkt $x \in \mathbb{R}$ differenzierbar.

Lemma E1.17. *Eine Funktion $f\colon I \to \mathbb{R}$ ist genau dann differenzierbar in $p \in I$, wenn es ein Funktion $f_1\colon I \to \mathbb{R}$ gibt, so dass $f(x) = f(p) + (x-p) f_1(x)$ und f_1 in p stetig ist. Es gilt dann $f'(p) = f_1(p)$.*

Beweis. Für $x \neq p$ gilt $f_1(x) = \frac{f(x)-f(p)}{x-p}$, der Differenzenquotient. Diese Funktion hat eine stetige Erweiterung auf $x = p$ genau dann, wenn die Ableitung $f'(p)$ existiert. □

Bemerkung E1.18. Dieses Lemma eignet sich für einen schönen Beweis der Rechenregeln.

Satz E1.19 (Kettenregel). *Sei $f\colon I \to J \subset \mathbb{R}$ differenzierbar in $p \in I$. Sei $g\colon J \to \mathbb{R}$ differenzierbar in $q := f(p) \in J$. Dann ist $g \circ f$ in p differenzierbar mit Ableitung*

$$(g \circ f)'(p) = g'(q) \cdot f'(p) = (g' \circ f)(p) \cdot f'(p).$$

Bemerkung E1.20. Mit $y = f(x)$ und $z = g(y) = g \circ f(x)$ kann man die Kettenregel auch wie folgt schreiben:

$$\frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}.$$

Beweis. Aus dem Lemma schreiben wir $f(x) = f(p) + (x-p) f_1(x)$ und $g(y) = g(q) + (y-q) g_1(y)$, wobei f_1 in p und g_1 in q stetig sind. Dann gilt

$$\begin{aligned} &(g \circ f)(x) \\ &\quad = g(q) + \bigl(f(x) - q\bigr) \cdot g_1\bigl(f(x)\bigr) \\ &\quad = (g \circ f)(p) + (x-p) \cdot f_1(x) \cdot g_1\bigl(q + (x-p) f_1(x)\bigr). \end{aligned}$$

Hier ist $f_1(x) \cdot g_1\bigl(q + (x-p) f_1(x)\bigr)$ eine Funktion, die stetig in p ist, und dessen Wert da $f'(p) \cdot g'(q)$ ist. Aus dem Lemma folgt die Kettenregel. □

Satz E1.21. *Sei $f\colon I \to \mathbb{R}$ stetig und injektiv (d.h., streng monoton). Sei J das Intervall $J = f(I)$ und sei $g = f^{-1}\colon J \to \mathbb{R}$ die (stetige) Umkehrfunktion. Falls f in $p \in I$ differenzierbar ist mit $f'(p) \neq 0$, dann ist g in $q := f(p)$ differenzierbar mit*

$$g'(q) = \frac{1}{f'(p)} = \frac{1}{(f' \circ g)(q)}.$$

Bemerkung E1.22. Sobald man weiß, dass g in q differenzierbar ist, ist die Formel für $g'(q)$ der Spezialfall der Kettenregel für $g \circ f = \mathrm{id}$.

Beweis. Mit $y := f(x)$ schreiben wir

$$f(x) - f(p) = (x-p) f_1(x)$$

als

$$y - q = \bigl(g(y) - g(q)\bigr) f_1(g(y))$$

um. Weil g stetig ist und f_1 in p stetig ist, ist $f_1 \circ g$ in q stetig. Weil $f_1 \circ g(q) = f_1(p) = f'(p) \neq 0$, ist auch $1/(f_1 \circ g)$ in q stetig. Weil $g(y) = g(q) + (y-q) \frac{1}{f_1(g(y))}$, ist g in q aus dem Lemma differenzierbar. □

Beispiel E1.23. Seien $f(x) := x^3$ und $g(y) := \sqrt[3]{y}$. Dann sind f und g stetige und streng monotone Umkehrfunktionen $\mathbb{R} \to \mathbb{R}$. Es gelten $f'(x) = 3x^2$ und

$$g'(y) = 1/f'(x) = \sqrt[3]{y}/3y, \qquad y \neq 0.$$

Weil $f'(0) = 0$, ist g in 0 nicht differenzierbar.

E2. Mittelwertsatz

Lemma E2.1. *Sei $f: I \to \mathbb{R}$ eine Funktion, die ihr Maximum (bzw. Minimum) in einem inneren Punkt $p \in I$ annimmt. Falls f in p differenzierbar ist, gilt $f'(p) = 0$.*

Beweis. Sei $f(p)$ ein Maximum. Wir betrachten den Differenzenquotient $Q := \frac{1}{h}\big(f(p+h) - f(p)\big)$. Für $h > 0$ gilt $Q \leq 0$ und für $h < 0$ gilt $Q \geq 0$. Das heißt, $f'_-(p) \geq 0$ und $f'_+(p) \leq 0$. Weil $f'(p) = f'_-(p) = f'_+(p)$ gilt $f'(p) = 0$. □

Satz E2.2 (Satz von Rolle). *Sei $f: [a,b] \to \mathbb{R}$ stetig mit $f(a) = f(b)$. Falls f in jedem inneren Punkt $x \in (a,b)$ differenzierbar ist, gibt es $c \in (a,b)$ mit $f'(c) = 0$.*

Beweis. Als stetige Funktion auf einem kompaktem Intervall, nimmt f ihre Maximum und Minimum an. Mindestens eins von denen muss in einem inneren Punkt $c \in (a,b)$ angenommen werden. (Der Wert $f(a) = f(b)$ ist nur dann Maximum *und* Minimum, wenn f konstant ist; in diesem Fall is $f(c)$ Maximum und Minimum für jedes $c \in (a,b)$.) Aus dem Lemma gilt $f'(c) = 0$. □

Bemerkung E2.3. Die Funktion $\sqrt{x(1-x)}$ ist stetig auf $[0,1]$. Sie ist nicht differenzierbar in 0 oder in 1. Wir dürfen den Satz trotzdem anwenden.

Der Mittelwertsatz stammt von Lagrange (1797).

Satz E2.4 (Mittelwertsatz). *Sei $f: [a,b] \to \mathbb{R}$ stetig und in jedem inneren Punkt differenzierbar. Dann gibt es $c \in (a,b)$ mit*

$$\frac{f(b) - f(a)}{b - a} = f'(c).$$

Beweis. Wir definieren

$$g(x) := f(a) + \frac{f(b) - f(a)}{b - a}(x - a).$$

(Der Graph von g ist eine Gerade – die Sekante durch die beiden Punkte $\big(a, f(a)\big)$ und $\big(b, f(b)\big)$ auf dem Graphen von f.) Wir wenden den Satz von Rolle auf $f - g$ an. Der Punkt c mit $f'(c) = g'(c)$ ist der Gesuchte. □

Bemerkung E2.5. Wie wir in den folgenden beiden Korollaren sehen, wendent man meistens den Mittelwertsatz „Umgekehrt" an: wir wissen etwas über die Ableitung f' und schließen daraus Information über die Funktionswerte.

`Ende der Vorlesung 2008 Juni 10`

Korollar E2.6. *Sei $f: I \to \mathbb{R}$ stetig und sei f in jedem inneren Punkt von I differenzierbar. Die Funktion f ist dann*

- *streng monoton steigend, falls $f'(x) > 0$,*
- *monoton steigend, falls $f'(x) \geq 0$,*
- *monoton fallend, falls $f'(x) \leq 0$,*
- *streng monoton fallend, falls $f'(x) < 0$,*

jeweils für jeden inneren Punkt $x \in I$.

Beweis. Wir betrachten den Fall $f'(x) > 0$; die Anderen sind ähnlich. Seien $a < b \in I$. Aus dem Mittelwertsatz gibt es $c \in [a,b]$, so dass $f(b) - f(a) = (b-a)f'(c)$. Es gelten $b - a > 0$ und $f'(c) > 0$, deshalb gilt

$$f(b) - f(a) = (b-a)f'(c) > 0. \qquad \square$$

Beispiel E2.7. Die Funktion $f(x) := x + x^2 \cos {}^1\!/\!_x$ ist differenzierbar in 0 mit Ableitung $f'(0) = 1$. Es gibt aber kein Intervall $I \ni 0$, auf dem f monoton ist. Deshalb haben wir den Satz über die Ableitung einer Umkehrfunktion sorgfältig formuliert.

Korollar E2.8. *Seien $f, g: I \to \mathbb{R}$ stetig und in jedem inneren Punkt differenzierbar mit $f' = g'$. Dann gilt $f(x) - g(x) \equiv c \in \mathbb{R}$.*

Beweis. Für die Differenz $f(x) - g(x)$ gilt $(f - g)' \equiv 0$. Aus dem letzten Korollar ist sie damit monoton steigend und monoton fallend, d.h., konstant. □

E3. Regel von l'Hôpital

Lemma E3.1 (Mittelwertsatz von Cauchy). *Seien die stetigen Funktionen $f, g: [a,b] \to \mathbb{R}$ in jedem inneren Punkt $p \in (a,b)$ differenzierbar. Es gibt dann $c \in (a,b)$ mit*

$$\big(f(b) - f(a)\big)\, g'(c) = \big(g(b) - g(a)\big)\, f'(c).$$

Beweis. Wir wenden den Satz von Rolle an auf

$$h(t) := \big(f(b) - f(a)\big)g(t) - \big(g(b) - g(a)\big)f(t).$$

Weil $h(a) = f(b)g(a) - f(a)g(b) = h(b)$ gibt es $c \in (a,b)$ mit $h'(c) = 0$. □

Bemerkung E3.2. Falls $g'(x) \neq 0$ für jedes $x \in (a,b)$, dann gilt (aus dem Satz von Rolle) auch $g(a) \neq g(b)$ und wir können den Satz in Bruchform schreiben:

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(c)}{g'(c)}.$$

Bemerkung E3.3. Falls g' stetig ist, können wir den Zwischenwertsatz anwenden: aus $g' \neq 0$ folgt, dass g' ein konstantes Vorzeichen hat. Damit ist g' streng monoton und deshalb hat g eine Umkehrfunktion, die differenzierbar ist. Betrachten wir die Funktion $f \circ g^{-1}$ auf $[g(a), g(b)]$, dann gibt uns der Mittelwertsatz ein $g(c) \in [g(a), g(b)]$ mit

$$\begin{aligned}\frac{f(c)}{g'(c)} &= \big(f \circ g^{-1}\big)(c) \\ &= \frac{\big(f \circ g^{-1}\big)g((b)) - \big(f \circ g^{-1}\big)(g(a))}{g(b) - g(a)} \\ &= \frac{f(b) - f(a)}{g(b) - g(a)}.\end{aligned}$$

Natürlich muss g' nicht stetig sein. Später werden wir aber sehen, dass ein Zwischenwertsatz auch für unstetige *Ableitungen* gilt.

Satz E3.4 (Regel von l'Hôpital). *Seien $I \subset \mathbb{R}$ ein Intervall, und $f, g\colon I \to \mathbb{R}$ stetige Funktionen. Weiter sei $p \in I$ mit $f(p) = 0 = g(p)$ und seien f und g differenzierbar in jedem Punkt $x \in I \smallsetminus \{p\}$ mit $g'(x) \neq 0$. Dann gilt auch $g(x) \neq 0$ für $x \neq p$. Falls existiert $\lim_{x\to p} {}^{f'(x)}/_{g'(x)} \in \overline{\mathbb{R}}$, dann gilt*

$$\lim_{x\to p} \frac{f(x)}{g(x)} = \lim_{x\to p} \frac{f'(x)}{g'(x)} \in \overline{\mathbb{R}},$$

insbesondere existiert dieser Grenzwert.

Beweis. Gäbe es ein $x \in I \smallsetminus \{p\}$ mit $g(x) = 0$, dann gäbe es nach dem Satz von Rolle ein c zwischen p und x mit $g'(c) = 0$, was unsere Voraussetzung widerspricht.

Für jedes x wenden wir den Mittelwertsatz von Cauchy an und zwar auf dem Intervall zwischen a und x. Wir finden $c = c(x)$ zwischen a und x mit $f(x)/g(x) = f'(c)/g'(c)$. Für $x \to p$ konvergiert $c(x) \to p$. Deshalb gilt

$$\lim_{x\to p} \frac{f(x)}{g(x)} = \lim_{x\to p} \frac{f'(c(x))}{g'(c(x))} = \lim_{y\to p} \frac{f'(y)}{g'(y)} \in \overline{\mathbb{R}},$$

falls letzteres existiert. □

Bemerkung E3.5. Natürlich reicht es aus, wenn $g' \neq 0$ auf einer Umgebung von p: wir ersetzen I mit diesem kleineren Intervall. Man darf aber nicht erlauben, dass das Vorzeichen von g' sich beliebig nah an p immer wieder ändert. (Sonst gibt es Gegenbeispiele, etwa $f(x) = x + \cos x \sin x$, $g(x) = e^{\sin x} f(x)$, $p = +\infty$. Hier existiert $\lim {}^{f'}/_{g'}$ aber nicht $\lim {}^{f}/_{g}$.)

Bemerkung E3.6. Es kann sein, dass $\lim {}^{f}/_{g}$ auch dann existiert, wenn $\lim {}^{f'}/_{g'}$ nicht existiert (z.B. mit $f(x) = x + \sin x$, $g(x) = x$, $p = \infty$).

Bemerkung E3.7. Es kann sein, dass $\lim_{x\to p} f'(x) = 0 = \lim_{x\to p} g'(x)$. Dann ist es zweckmäßig, die Existenz von $\lim {}^{f'}/_{g'}$ so zu untersuchen, dass man zunächst die Regel von l'Hôpital ein zweites mal anwendet. D.h., wir schauen zunächst, ob $\lim {}^{f''}/_{g''}$ existiert.

Korollar E3.8. *Die Regel gilt auch für $p = \pm\infty$. Falls die Grenzwerte*

$$\lim_{x\to+\infty} f(x) = 0 = \lim_{x\to+\infty} g(x), \qquad \lim_{x\to+\infty} {}^{f'(x)}/_{g'(x)} \in \overline{\mathbb{R}}$$

existieren (und falls $g'(x) \neq 0$), dann gilt

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)} = \lim_{x\to+\infty} \frac{f'(x)}{g'(x)} \in \overline{\mathbb{R}}.$$

(Und natürlich auch Entsprechendes für $x \to -\infty$.)

Beweisidee. Wir wenden die Regel von l'Hôpital auf die Funktionen $x \mapsto f({}^1/_x)$ und $x \mapsto g({}^1/_x)$ an. □

Korollar E3.9. *Die Regel gilt auch im Falle $\lim |f(x)| = \lim |g(x)| = +\infty$.*

Beweisidee. Wir wenden den Satz (bzw. das Korollar) auf die Funktionen $x \mapsto {}^1/_{g(x)}$ und $x \mapsto {}^1/_{f(x)}$ an. □

F. POTENZREIHEN

F1. Die Exponentialfunktion

a. *Eindeutigkeit*

Definition F1.1. Eine Funktion $f\colon \mathbb{R} \to \mathbb{R}$ heißt *Exponentialfunktion*, falls sie differenzierbar ist und

$$f'(x) = f(x), \qquad f(0) = 1.$$

Lemma F1.2. *Seien f und g Exponentialfunktionen. Dann gilt $f(-x) = {}^1\!/_{g(x)}$.*

Beweis. Für $h(x) := g(x)f(-x)$ gilt

$$h'(x) = g'(x)f(-x) - g(x)f'(-x) \equiv 0.$$

Deshalb ist $h(x) \equiv h(0) = 1$ konstant. □

Wenden wir das Lemma auf $g = f$ an, finden wir:

Korollar F1.3. *Für eine (beliebige) Exponentialfunktion f gilt $f(-x) = {}^1\!/_{f(x)}$.* □

Korollar F1.4. *Es gibt höchstens eine Exponentialfunktion.*

Beweis. Sind f und g Exponentialfunktionen, dann gilt $g(x) = {}^1\!/_{f(-x)} = f(x)$. □

Bemerkung F1.5. Die bekannten Eigenschaften der Exponentialfunktion folgen aus unserer abstrakten Definition. Wir warten aber, bis wir die Existenz bewiesen haben.

`Ende der Vorlesung 2008 Juni 16`

---

b. *Differentiation von Reihen*

Definition F1.6. Für $n \in \mathbb{N}$ definieren wir $n!$ rekursiv wie folgt: $0! := 1$ und $(n+1)! := (n+1)n!$. Damit gilt

$$n! := \prod_{i=1}^{n} i = 1 \cdot 2 \cdot 3 \cdot \cdots \cdot (n-1) \cdot n.$$

Lemma F1.7. *Für jedes $x \in \mathbb{R}$ konvergiert absolut die Reihe $\sum {}^{x^k}\!/_{k!}$.*

Beweis. Sei $x \in \mathbb{R}$ gegeben. Für jedes $k > 2|x|$ gilt

$$\left|\frac{x^k}{k!}\right| \Big/ \left|\frac{x^{k-1}}{(k-1)!}\right| = \frac{|x|}{k} < \frac{1}{2}.$$

Aus dem Quotientenkriterium konvergiert die Reihe. □

Korollar F1.8. $\lim \sqrt[k]{k!} = +\infty$.

Beweis. Falls es $a \in \mathbb{R}$ gibt, so dass $\sqrt[k]{k!} \le a$ für unendlich viele k gilt, betrachten wir die Reihe $\sum {}^{a^k}\!/_{k!}$. Aus dem Wurzelkriterium konvergiert sie nicht: für die Wurzel gilt ${}^a\!/_{\sqrt[k]{k!}} \ge 1$ für unendlich viele k. Dies widerspricht das Lemma. □

Sei nun (a_n) eine beschränkte Folge in $\mathbb{R}$. Sei p_n das Polynom

$$p_n(x) := \sum_{k=0}^{n} a_k \frac{x^k}{k!},$$

dessen Ableitung

$$p_n'(x) = \sum_{k=1}^{n} k a_k \frac{x^{k-1}}{k!} = \sum_{k=0}^{n-1} a_{k+1} \frac{x^k}{k!}$$

ist.

Aus dem Lemma oben und dem Vergleichskriterium wissen wir, dass

$$f(x) := \sum_{k=0}^{\infty} a_k \frac{x^k}{k!} = \lim_{n\to\infty} p_n(x)$$

für jedes $x \in \mathbb{R}$ konvergiert und deshalb eine Funktion $f\colon \mathbb{R} \to \mathbb{R}$ (die sogenannte Grenzfunktion) definiert. Dasselbe gilt für

$$g(x) := \sum_{k=0}^{\infty} a_{k+1} \frac{x^k}{k!} = \lim_{n\to\infty} p_n'(x).$$

Naheliegend ist die Vermutung, dass f differenzierbar ist mit Ableitung $f' = g$. Das Problem ist aber, dass Ableitungun als Limes definiert wurden.

Bemerkung F1.9. Zwei Limes darf man im Allgemeinen nicht vertauschen. Zum Beispiel gelten

$$\lim_{m\to\infty} \sqrt[m]{{}^1\!/_n} = 1, \qquad \text{für jedes } n \in \mathbb{N},$$
$$\lim_{n\to\infty} \sqrt[m]{{}^1\!/_n} = 0, \qquad \text{für jedes } m \in \mathbb{N}.$$

Deshalb gilt

$$\lim_{n\to\infty}\lim_{m\to\infty} \sqrt[m]{{}^1\!/_n} = 1 \ne 0 = \lim_{m\to\infty}\lim_{n\to\infty} \sqrt[m]{{}^1\!/_n}.$$

Bemerkung F1.10. Seien $f_n\colon I \to \mathbb{R}$ differenzierbare Funktionen. Sei $f\colon I \to \mathbb{R}$ einen Limes im Sinne, dass für jedes $x \in I$ konvergiert $f_n(x) \to f(x)$. Es muss nicht sein, dass f stetig ist, geschweige denn, dass f differenzierbar ist oder sogar $f' = \lim f_n'$. Später untersuchen wir allgemeinere Bedingungen für Konvergenz von Ableitungen und sehen noch erstaunlichere Beispiele. Vorerst erwähnen wir nur das Beispiel $I = [0,1]$, $f_n(x) = x^n$, wo der Limes $f(x) = 1 + \operatorname{sgn}(x-1)$ unstetig ist.

Satz F1.11 (Differenzierbare Potenzreihen). *Sei (a_n) eine beschränkte Folge und seien $f, g\colon \mathbb{R} \to \mathbb{R}$ wie oben durch Potenzreihen definiert. Dann ist f in jedem $p \in \mathbb{R}$ differenzierbar mit Ableitung $f'(p) = g(p)$.*

Beweis. Sei $p \in \mathbb{R}$ gegeben. Für $h \in \mathbb{R}$ sei $R(h)$ der Rest

$$F(h) := \left|\frac{f(p+h) - f(p)}{h} - g(p)\right|.$$

Sei $\varepsilon > 0$ gegeben. Wir müssen ein $\delta > 0$ finden, so dass $R(h) < \varepsilon$ für alle $|h| < \delta$. Aus der Definition von f und g haben wir

$$\begin{aligned} R(h) &\le \sum_{k=1}^{\infty} \left| \frac{a^k}{k!} \frac{(p+h)^k - p^k}{h} - \frac{a^k}{(k-1)!} p^{k-1} \right| \\ &= \sum_{1}^{\infty} \frac{|a_k|}{(k-1)!} \left| (p+h_k)^{k-1} - p^{k-1} \right| =: \sum b_{k-1}. \end{aligned}$$

Hier wenden wir für jedes k den Mittelwertsatz an auf die Funktion $x \mapsto x^k$ auf dem Intervall zwischen p und $p+h$; die Wert, die daraus kommt, nennen wir h_k.

Nun sei $r := |p| + 1$. Die absolute Konvergenz von $\sum |a_{k+1}| r^k / k!$ besagt, es gibt $n_0 \in \mathbb{N}$, so dass für $|x| \le r$ gilt

$$\sum_{k=n_0}^{\infty} \frac{|a_{k+1}|}{k!} |x|^k \le \sum_{k=n_0}^{\infty} \frac{|a_{k+1}|}{k!} r^k < \frac{\varepsilon}{4}.$$

Wir haben aber

$$b_k \le \frac{|a_{k+1}|}{k!} |p - h_{k+1}|^k + \frac{|a_{k+1}|}{k!} |p|^k.$$

Für $|h| < 1$ gilt $|h_k| < 1$ und deshalb $\sum_{k=n_0}^{\infty} b_k < \varepsilon/2$.

Für jedes $k = 0, 1, \dots, n_0 - 1$ ist die Funktion $x \mapsto x^k$ stetig. Deshalb gibt es $\delta_k > 0$, so dass für alle $|h| < \delta_k$ gilt

$$\frac{|a_{k+1}|}{k!} \left| (p+h)^k - p^k \right| < \frac{\varepsilon}{2n_0}.$$

Jetzt wählen wir $\delta \le \min(1, \delta_0, \dots, \delta_{n_0-1})$. Für $|h| < \delta$ gilt damit $|h_k| < |h| < \delta \le \delta_k$ und deshalb $b_k < \varepsilon/2n_0$ für $k < n_0$. Damit gilt $\sum_{k=0}^{n_0-1} b_k < \varepsilon/2$.

Deshalb gilt wie gewünscht $R(h) < \varepsilon$. □

Bemerkung F1.12. Weil g eine Funktion der selben Form ist, ist sie auch differenzierbar.

Definition F1.13. $\exp x := \displaystyle\sum_{k=0}^{\infty} \frac{x^k}{k!}$.

Korollar F1.14. *Die Funktion* exp *ist (die eindeutige) Exponentialfunktion:*

$$\exp' = \exp, \qquad \exp 0 = 1. \qquad \square$$

c. *Eigenschaften der Exponentialfunktion*

Lemma F1.15. *Die Exponentialfunktion* exp *ist positiv und streng monoton steigend, deshalb injektiv.*

Beweis. Weil $\exp(-x) = 1/\exp x$, gilt $\exp x \ne 0$. Aus dem Zwischenwertsatz und $\exp 0 = 1$ folgt $\exp x > 0$. Damit ist $\exp' x > 0$ und exp streng monoton steigend. □

Lemma F1.16. *Die Exponentialfunktion* exp *ist eine Bijektion* $\mathbb{R} \to (0, +\infty)$.

Beweis. Es gilt $\exp x > x + 1$. (Wir sehen das sofort von der Potenzreihe. Um es direkt aus der Definition einer Exponentialfunktion herzuleiten, wendet man den Mittelwertsatz auf $\exp|_{[0,x]}$ an.) Damit gilt

$$\lim_{x\to+\infty} \exp x = +\infty, \qquad \lim_{x\to-\infty} \exp x = \lim_{x\to+\infty} \frac{1}{\exp x} = 0.$$

Aus dem Zwischenwertsatz gilt $\exp(\mathbb{R}) = (0, +\infty)$. □

Satz F1.17. *Für* $a, b \in \mathbb{R}$ *gilt*

$$\exp(a+b) = \exp a \, \exp b.$$

Beweis. Sei $b \in \mathbb{R}$ gegeben. Setzen wir

$$g(x) := \exp(x+b)\exp(-x),$$

dann gilt

$$g'(x) = \exp'(x+b)\exp(-x) - \exp(x+b)\exp'(-x) = 0.$$

Deshalb ist $g(x)$ konstant, $g(x) \equiv g(0) = \exp b$. □

`Ende der Vorlesung 2008 Juni 17`

F2. Verwandte Funktionen

a. *Logarithmus*

Definition F2.1. Die Umkehrfunktion der Exponentialfunktion heißt *Logarithmus*,

$$\log = \exp^{-1} \colon (0, +\infty) \to \mathbb{R}.$$

Man schreibt auch ln = log um zu betonen, dass dieser der *natürliche* Logarithmus ist (und nicht etwa der Logarithmus zur Basis 10). Die Eigenschaften der Exponentialfunktion ergeben folgenden Eigenschaften des Logarithmus:

Satz F2.2. *Der Logarithmus ist streng monoton steigend. Es gelten*

$$\log'(x) = \frac{1}{\exp'(\log x)} = \frac{1}{x},$$
$$\lim_{x\to+\infty} \log x = +\infty, \qquad \lim_{x\to 0^+} \log x = -\infty,$$
$$\log(xy) = \log x + \log y. \qquad \square$$

b. *Allgemeine Potenzen*

Definition F2.3. Für $a \in (0, +\infty)$ und $x \in \mathbb{R}$ definieren wir

$$a^x := \exp(x \log a) > 0.$$

Die Eigenschaften von exp und log ergeben:

Satz F2.4. *Seien $a, b > 0$ und $x, y \in \mathbb{R}$. Es gilt $a^1 = a$. Ferner gilt*

$$a^{x+y} = a^x a^y \quad (\textit{insbes. } a^0 = 1 \textit{ und } a^{-x} = {}^1\!/_{a^x}).$$

Es gelten auch

$$(ab)^x = a^x b^x, \qquad \big(a^x\big)^y = a^{xy}. \qquad \square$$

Beweis. Wir beweisen eine Regel, die Anderen lassen wir als Aufgaben. Es gilt

$$\begin{aligned}(ab)^x &= \exp\big(x \log(ab)\big) = \exp\big(x(\log a + \log b)\big)\\ &= \exp\big(x \log a + x \log b\big)\\ &= \exp\big(x \log a\big)\ \exp\big(x \log b\big) = a^x b^x. \qquad \square\end{aligned}$$

Bemerkung F2.5. Schreibt man a^{b^x}, so heißt es $a^{(b^x)}$ und nicht $\big(a^b\big)^x$, welches man eher als a^{bx} schreibt.

Bemerkung F2.6. Weil $a^{x+1} = a^x a^1 = a^x a$, stimmt diese Definition im Falle $x \in \mathbb{N}$ mit der alten rekursiven Definition überein.

Bemerkung F2.7. Für $x = {}^m\!/_n \in \mathbb{Q}$ gilt

$$a^{m/n} = \sqrt[n]{a^m} = \big(\sqrt[n]{a}\big)^m.$$

Satz F2.8. *Die Funktionen $x \mapsto a^x$ und $a \mapsto a^x$ sind differenzierbar. Es gelten*

$$\frac{da^x}{dx} = a^x \log a, \qquad \frac{da^x}{da} = a^x \frac{x}{a} = x a^{x-1}.$$

Bemerkung F2.9. Die Umkehrfunktion zu $x \mapsto a^x$ heißt Logarithmus zur Basis a:

$$\log_a \colon (0, +\infty) \to \mathbb{R}, \quad \log_a(a^x) = x, \quad \log_a y = \frac{\log y}{\log a}.$$

Definition F2.10. Die Euler'sche Zahl ist

$$e := \exp 1 = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} + \cdots.$$

Damit gelten $\log e = 1$, $\log_e x = \log x$ und $e^x = \exp x$.

Lemma F2.11. *Es gilt $2\,{}^2\!/_3 < e < 2\,{}^3\!/_4$.*

Beweis. Die untere Schranke ist einfach die Summe $1 + 1 + {}^1\!/_2 + {}^1\!/_6$ der ersten vier Summanden. Für die obere Schranke gilt

$$\begin{aligned}e &= 1 + 1 + \frac{1}{2}\left(1 + \frac{1}{3} + \frac{1}{3\cdot 4} + \frac{1}{3\cdot 4\cdot 5} + \cdots\right)\\ &< 1 + 1 + \frac{1}{2}\left(1 + \frac{1}{3} + \frac{1}{3\cdot 3} + \frac{1}{3\cdot 3\cdot 3} + \cdots\right)\\ &= 2 + \frac{1}{2}\cdot\frac{3}{2}. \qquad \square\end{aligned}$$

Bemerkung F2.12. Mithilfe von Taylorreihen können wir später solche Abschätzungen schneller machen. Es gilt

$$e \approx 2.71828182845904523536028747135266 2\cdots.$$

Später zeigen wir auch, dass die Zahl e irrational ist. Sie ist sogar *transzendent*, d.h., es gibt kein Polynom mit ganzzahligen Koeffizienten, das e als Nullstelle hat.

c. *Hyperbelfunktionen*

Definition F2.13. Eine Funktion $f \colon \mathbb{R} \to \mathbb{R}$ heißt *gerade* (bzw. *ungerade*), falls für jedes $x \in \mathbb{R}$ gilt $f(x) = f(-x)$ (bzw. $f(x) = -f(-x)$).

Lemma F2.14. *Jede Funktion $f \colon \mathbb{R} \to \mathbb{R}$ hat eine Zerlegung $f = f_+ + f_-$, wobei f_+ gerade und f_- ungerade ist.*

Beweis. Wir setzen $f_\pm(x) := \frac{1}{2}\big(f(x) \pm f(-x)\big)$. $\square$

Definition F2.15. Der gerade Teil von exp heist *Cosinus hyperbolicus*, der ungerade Teil *Sinus hyperbolicus*:

$$\cosh x := \frac{e^x + e^{-x}}{2}, \qquad \sinh x := \frac{e^x - e^{-x}}{2}.$$

Deren Quotient heißt *Tangens hyperbolicus*

$$\tanh x := \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}}.$$

Satz F2.16. *Es gelten*

$$\begin{gathered}(\cosh x)^2 - (\sinh x)^2 \equiv 1,\\ \sinh' = \cosh, \qquad \cosh' = \sinh,\\ \tanh' x = \frac{1}{(\cosh x)^2} = 1 - (\tanh x)^2. \qquad \square\end{gathered}$$

Bemerkung F2.17. Für $n = 2, 3, \ldots$ schreibt man normalerweise $\sinh^n x := (\sinh x)^n$ (und ähnlich bei den anderen Hyperbel- und Winkelfunktionen). Das ist sicherlich inkonsequent aber trotzdem bequem. Unberührt davon bedeutet $\sinh^{-1}$ die Umkehrfunktion und nicht ${}^1\!/_{\sinh}$.

Bemerkung F2.18. Die Funktionen sinh und tanh sind streng monoton steigend und haben damit Umkehrfunktionen

$$\sinh^{-1} \colon \mathbb{R} \to \mathbb{R}, \qquad \tanh^{-1} \colon (-1, +1) \to \mathbb{R}.$$

Der Cosinus hyperbolicus ist auf $[0, +\infty)$ streng monoton steigend, und damit gibt es $\cosh^{-1} \colon [1, +\infty) \to [0, +\infty)$.

Lemma F2.19. *Es gelten*

$$\begin{aligned}\sinh^{-1} x &= \log\Big(x + \sqrt{1 + x^2}\Big),\\ \tanh^{-1} x &= \log\sqrt{\frac{1+x}{1-x}},\\ \cosh^{-1} x &= \log\Big(x + \sqrt{x^2 - 1}\Big).\end{aligned}$$

Beweis. Wir beweisen nur die erste Formel, die anderen sind ähnlich. Es gilt

$$\begin{aligned}\log\Big(\sinh x + \sqrt{1 + \sinh^2 x}\Big) &= \log(\sinh x + \cosh x)\\ &= \log\exp x = x. \qquad \square\end{aligned}$$

Bemerkung F2.20. Umgekehrt kann man es auch beweisen:

$$\begin{aligned}\sinh\Big(\log\big(x+\sqrt{1+x^2}\big)\Big) &= \frac{1}{2}\left(x+\sqrt{1+x^2}-\frac{1}{x+\sqrt{1+x^2}}\right)\\ &= \frac{1}{2}\left(\frac{x^2+2x\sqrt{1+x^2}+1+x^2\ -\ 1}{x+\sqrt{1+x^2}}\right)=x.\end{aligned}$$

F3. Trigonometrische Funktionen

Falls die Ableitung f' einer differenzierbaren Funktion $f\colon I\to\mathbb{R}$ auch differenzierbar ist, heißt $f''=(f')'$ die zweite Ableitung von f. Zum Beispiel, sei $f\colon x\mapsto e^{\lambda x}$. Dann gilt $f'=\lambda f$ und $f''=\lambda^2 f$. Suchen wir nach Lösungen von $f''=\alpha f$ mit $\alpha>0$, kennen wir deshalb schon zwei: $x\mapsto\exp(\pm\sqrt{\alpha}x)$. (Oder deren linearen Kombinationen $\cosh\sqrt{\alpha}x$ und $\sinh\sqrt{\alpha}x$.)

Für $\alpha<0$ hingegen kennen wir noch keine Lösungen. Die Gleichung $f''=-f$ heißt Schwingungsgleichung und beschreibt z.B. die Bewegung einer Feder.

`Ende der Vorlesung 2008 Juni 23`

Definition F3.1. Sei (a_k) die Folge

$$0,1,0,-1,0,1,0,-1,\ldots,$$

d.h., für $j\in\mathbb{N}$ gilt

$$a_{4j}=0,\quad a_{4j+1}=1,\quad a_{4j+2}=0,\quad a_{4j+3}=-1.$$

Dann ist *Sinus* die differenzierbare Funktion

$$\sin x := \sum_{k=0}^{\infty} a_k\frac{x^k}{k!} = x-\frac{x^3}{3!}+\frac{x^5}{5!}-\cdots.$$

Die Ableitung heißt *Cosinus*:

$$\cos x := \sin' x = \sum_{k=0}^{\infty} a_{k+1}\frac{x^k}{k!} = 1-\frac{x^2}{2!}+\frac{x^4}{4!}-\cdots.$$

Bemerkung F3.2. Weil $a_{k+2}=-a_k$ gilt $\sin''=\cos'=-\sin$. Es gelten auch $\sin 0=0$ und $\cos 0=1$.

Lemma F3.3. *Es gilt* $(\sin x)^2+(\cos x)^2\equiv 1$.

Beweis. Dazu sei $g(x):=(\sin x)^2+(\cos x)^2$. Wir rechnen $g'(x)=2\sin x\cos x-2\cos x\sin x=0$. Damit ist g konstant, $g(x)\equiv g(0)=1$. □

Bemerkung F3.4. Natürlich kann man dies direkt aus den Potenzreihen schließen, es ist aber viel schöner und schneller die Differentialgleichung zu benutzen.

Satz F3.5. *Sei $f\colon\mathbb{R}\to\mathbb{R}$ eine Lösung der Differentialgleichung $f''=-f$. Seien $a:=f(0)$ und $b:=f'(0)$. Dann gilt $f(x)=a\cos x+b\sin x$.*

Beweis. Offensichtlich ist $a\cos x+b\sin x$ eine Lösung der Schwingsungsgleichung mit den gennanten Anfangsbedingungen. Wir müssen Eindeutigkeit beweisen. Dazu seien

$$\begin{aligned}g(x)&:=\cos x\,f(x)-\sin x\,f'(x),\\ h(x)&:=\sin x\,f(x)+\cos x\,f'(x).\end{aligned}$$

Aus der Differentialgleichung finden wir $g'\equiv 0$ und $h'\equiv 0$. Deshalb sind g und h konstant, $g(x)\equiv g(0)=a$ und $h(x)\equiv h(0)=b$. Jetzt betrachten wir folgende Summe, $\cos x$ mal die erste Gleichung plus $\sin x$ mal die Zweite:

$$a\cos x+b\sin x=\big(\cos^2 x+\sin^2 x\big)f(x)=f(x).\qquad\square$$

Korollar F3.6. *Sinus ist ungerade und Cosinus gerade.*

Beweis. Die Funktion $x\mapsto\cos(-x)$ löst die Schwingungsgleichung mit denselben Anfangsbedingungen wie bei cos selbst. Der Satz liefert $\cos(-x)=\cos x$, die Ableitung davon ist $\sin(-x)=-\sin x$. □

Bemerkung F3.7. Hier geht es schneller, die Potenzreihen zu benutzen; wir wollten aber üben, den Satz anzuwenden.

Satz F3.8 (Additionstheorem). *Für alle $x,y\in\mathbb{R}$ gelten*

$$\begin{aligned}\sin(x+y)&=\sin x\cos y+\sin y\cos x,\\ \cos(x+y)&=\cos x\cos y-\sin y\sin x.\end{aligned}$$

Beweis. Sei $y\in\mathbb{R}$ gegeben. Die Funktionen $x\mapsto\sin(x+y)$ und $x\mapsto\cos(x+y)$ lösen die Schwingungsgleichung. Aus dem Satz folgen die gewünschten Formeln. □

Lemma F3.9. *Es gilt* $\cos 2<0$.

Beweis. Weil die Potenzreihe für cos alternierend ist, aus dem Leibnizkreiterium wissen wir, dass die Partialsummen alternierend oberhalb und unterhalb der Reihensumme liegen. Deshalb gilt

$$\begin{aligned}\cos 2 &= 1-\frac{2^2}{2!}+\frac{2^4}{4!}-\frac{2^6}{6!}+\cdots\\ &< 1-\frac{2^2}{2!}+\frac{2^4}{4!}=1-2+\frac{2}{3}=-\frac{1}{3}<0.\qquad\square\end{aligned}$$

Definition F3.10. $\pi:=2\inf\{x>0:\cos x=0\}$.

Lemma F3.11. *Es gilt $0<\pi<4$. In ${}^{\pi}\!/_2$ gelten $\cos{}^{\pi}\!/_2=0$ und $\sin{}^{\pi}\!/_2=1$.*

Beweis. Wir wissen $\cos 0=1$ und $\cos 2<0$. Aus dem Zwischenwertsatz hat cos eine Nullstelle in $(0,2)$. Es folgt, dass ${}^{\pi}\!/_2\in[0,2)$. Weil cos stetig ist, ist die Menge aller Nullstellen abgeschlossen. Das heißt, ${}^{\pi}\!/_2\neq 0$ ist eine Nullstelle, die kleinste Positive. Damit gilt auch $\sin^2{}^{\pi}\!/_2=1$. Weil cos auf $(0,{}^{\pi}\!/_2)$ positiv ist, ist sin streng monoton steigend und damit auch positiv. Deshalb gilt $\sin{}^{\pi}\!/_2=+1$. □

Bemerkung F3.12. Es gilt

$$\pi \approx 3.14159265358979323846264338327950288841971\cdots.$$

Korollar F3.13. *Sinus und Cosinus sind* 2π*-periodische Funktionen. Es gelten*

$$\begin{aligned} \sin(x+{}^{\pi}/_{2}) &= \cos x, & \cos(x+{}^{\pi}/_{2}) &= -\sin x,\\ \sin(x+\pi) &= -\sin x, & \cos(x+\pi) &= -\cos x,\\ \sin(x+2\pi) &= \sin x, & \cos(x+2\pi) &= \cos x. \end{aligned}$$

Beweis. Die erste Zeile folgt aus dem Additionstheorem, die Zweite aus der Ersten, die Dritte aus der Zweiten. □

Bemerkung F3.14. Der mathematisch sauberste Weg, Winkel auch in Grad messen zu können, ist folgendes. Für $x \in \mathbb{R}$ definieren wir $x^\circ := x\frac{\pi}{180}$. Damit gelten z.B. $30^\circ = {}^{\pi}/_{6}$ und $90^\circ = {}^{\pi}/_{2}$.

Satz F3.15. *Es gelten*

$$\begin{aligned} \sin 0^\circ = \sin 0 = 0 = {}^{\sqrt{0}}/_{2} = \cos {}^{\pi}/_{2} = \cos 90^\circ,\\ \sin 30^\circ = \sin {}^{\pi}/_{6} = {}^{1}/_{2} = {}^{\sqrt{1}}/_{2} = \cos {}^{\pi}/_{3} = \cos 60^\circ,\\ \sin 45^\circ = \sin {}^{\pi}/_{4} = {}^{1}/_{\sqrt{2}} = {}^{\sqrt{2}}/_{2} = \cos {}^{\pi}/_{4} = \cos 45^\circ,\\ \sin 60^\circ = \sin {}^{\pi}/_{3} = {}^{\sqrt{3}}/_{2} = {}^{\sqrt{3}}/_{2} = \cos {}^{\pi}/_{6} = \cos 30^\circ,\\ \sin 90^\circ = \sin {}^{\pi}/_{2} = 1 = {}^{\sqrt{4}}/_{2} = \cos 0 = \cos 0^\circ. \end{aligned}$$
□

Definition F3.16. Wir definieren $\tan x := {}^{\sin x}/_{\cos x}$ auf der Teilmenge

$$\{x \in \mathbb{R} : \cos x \neq 0\} = \{x : {}^{x}/_{\pi} - {}^{1}/_{2} \notin \mathbb{Z}\}.$$

Lemma F3.17. *Damit ist* tan *eine* π*-periodische Funktion. Es gelten*

$$\begin{aligned} \tan' x &= 1 + \tan^2 x = {}^{1}/_{\cos^2 x},\\ \tan(x+y) &= \frac{\tan x + \tan y}{1 - \tan x \tan y}. \end{aligned}$$
□

Bemerkung F3.18. Besonders ausserhalb des deuschtsprachigen Raumes benutzt man auch den Cotangens, den Sekans und den Cosekans:

$$\cot x := \frac{1}{\tan x}, \quad \sec x := \frac{1}{\cos x}, \quad \csc x := \frac{1}{\sin x}.$$

Damit schreibt man z.B. $\tan' x = 1 + \tan^2 x = \sec^2 x$.

Lemma F3.19. *Sinus ist auf* $[-{}^{\pi}/_{2}, {}^{\pi}/_{2}]$ *streng monoton steigend und damit eine Bijektion*

$$\sin\colon [-{}^{\pi}/_{2}, {}^{\pi}/_{2}] \to [-1, 1].$$

Cosinus ist auf $[0, \pi]$ *streng monoton fallend und damit eine Bijektion*

$$\cos\colon [0, \pi] \to [-1, 1].$$

Tangens ist auf jedem Intervall in seinem Definitionsbereich – insbes. auf $(-{}^{\pi}/_{2}, {}^{\pi}/_{2})$ *– streng monoton steigend und damit eine Bijektion*

$$\tan\colon (-{}^{\pi}/_{2}, {}^{\pi}/_{2}) \to \mathbb{R}.$$

Beweis. Die Ableitung $\sin' = \cos$ ist (nach der Definition von π) positiv auf $[0, {}^{\pi}/_{2})$, und – weil sie gerade ist – auch auf $({}^{\pi}/_{2}, 0]$.

Weil $\sin x = \cos(x - {}^{\pi}/_{2})$, ist sin auf $(0, \pi)$ positiv, d.h., die Ableitung $\cos' = -\sin$ ist negativ.

Die Ableitung $\sec^2$ ist positiv in jedem Punkt im Definitionsbereich von tan. □

Definition F3.20. Die Umkehrfunktionen heißen Arcus Sinus, Arcus Cosinus bzw. Arcus Tangens:

$$\begin{aligned} \arcsin &:= \sin^{-1}\colon [-1,1] \to [-{}^{\pi}/_{2}, {}^{\pi}/_{2}],\\ \arccos &:= \cos^{-1}\colon [-1,1] \to [0,\pi],\\ \arctan &:= \tan^{-1}\colon \mathbb{R} \to (-{}^{\pi}/_{2}, {}^{\pi}/_{2}). \end{aligned}$$

Lemma F3.21. *Es gelten*

$$\begin{aligned} \arccos x &= \frac{\pi}{2} - \arcsin x,\\ \arcsin' x &= \frac{1}{\sqrt{1-x^2}} = -\arccos' x,\\ \arctan' x &= \frac{1}{1+x^2}. \end{aligned}$$
□

F4. Stetigkeit der Ableitung

Die einfachsten Beispiele unstetiger Funktionen haben einen Sprung wie bei sgn x. Wenn wir versuchen eine differenzierbare Funktion f zu finden, deren Ableitung einen Sprung hat, geht es nicht. Die Funktion $x \mapsto |x|$ ist in 0 (wo die Ableitung springt) nicht differenzierbar. Unser einfachstes Beispiel, wo f' existiert aber unstetig ist, war $f(x) = x^2 \cos {}^{1}/_{x}$. Warum es nicht einfacher geht (z.B. mit Sprung) erklärt uns ein Satz von Darboux:

Satz F4.1 (Zwischenwertsatz für Ableitungen). *Sei* f *eine differenzierbare Funktion* $f\colon I \to \mathbb{R}$*. Dann ist* $f'(I)$ *ein Intervall.*

Beweis. Seien $a < b \in I$. Für jeden Wert s zwischen $f'(a)$ und $f'(b)$ müssen wir $c \in (a,b)$ finden mit $f'(c) = s$. Wir dürfen (o.B.d.A.) annehmen, $f'(a) < s < f'(b)$. Sei $m := \frac{f(b)-f(a)}{b-a}$. Im Falle $s = m$ erhalten wir aus dem Mittelwertsatz das gewünschte c.

Im Falle $s < m$ definieren wir $g(x) := \frac{f(x)-f(a)}{x-a}$, stetig fortgesetzt durch $g(a) := f'(a)$. Wir haben

$$g(a) = f'(a) < s < m = g(b).$$

Aus dem Zwischenwertsatz für g gibt es $d \in (a,b)$ mit $g(d) = s$. Aus dem Mittelwertsatz für $f|_{[a,d]}$ gibt es dann $c \in (a,d) \subset (a,b)$ mit $f'(c) = s$.

Im Falle $s > m$ geht es entsprechend mit $g(x) := \frac{f(b)-f(x)}{b-x}$. □

F5. Höhere Ableitungen

Definition F5.1. Sei $f\colon I \to \mathbb{R}$ eine differenzierbare Funktion. Falls die Ableitung $f'\colon I \to \mathbb{R}$ auch differenzierbar ist, sagen wir, f ist *zweimal differenzierbar*, und wir nennen $f'' := \big(f'\big)'$ die *zweite Ableitung* von f. Ähnliches gilt für die *dritte Ableitung* f'''. Weil die Notation sonst unhandlich wird, schreiben wir auch:

$$f^{(0)} := f, \quad f^{(1)} := f', \quad f^{(2)} := f'', \quad f^{(3)} = f'''.$$

Rekursiv, sei $f\colon I \to \mathbb{R}$ eine k-mal differenzierbare Funktion. Falls die k^{te} Ableitung $f^{(k)}\colon I \to \mathbb{R}$ auch differenzierbar ist, sagen wir, f ist $(k+1)$*-mal differenzierbar*, und wir nennen $f^{(k+1)} := \big(f^{(k)}\big)'$ die $(k+1)^{te}$ *Ableitung* von f.

Falls f k-mal differenzierbar ist und $f^{(k)}$ stetig ist, sagen wir, f ist k-mal *stetig differenzierbar*, und schreiben $f \in C^k(I)$.

Falls f k-mal differenzierbar ist für jedes $k \in \mathbb{N}$, sagen wir, f ist *beliebig oft* oder *unendlich oft* differenzierbar, und schreiben $f \in C^\infty(I)$.

Bemerkung F5.2. Natürlich kann man nicht unendlich oft differenzieren – es gibt keine „∞^{te} Ableitung“.

Bemerkung F5.3. Eine k-mal differenzierbare Funktion ist immer $(k-1)$-mal stetig differenzierbar.

Bemerkung F5.4. Sei (a_n) beschränkt und sei

$$f(x) := \sum a_k \frac{x^k}{k!}.$$

Wir haben gezeigt, $f' = \sum a_{k+1} x^k/_{k!}$. Weil diese Ableitung dieselbe Form hat, ist sie auch differenzierbar, usw. Das heißt, $f \in C^\infty(\mathbb{R})$. Es gilt

$$f^{(n)} = \sum_{k=0}^{\infty} a_{k+n} \frac{x^k}{k!}, \qquad \text{insbes. } f^{(n)}(0) = a_n.$$

a. *Konvexität*

Definition F5.5. Eine Funktion $f\colon I \to \mathbb{R}$ heißt *konvex*, falls sie immer unter ihre lineare Interpolationen liegt, d.h., für alle $x, y \in I$ und alle $t \in (0,1)$ gilt $f\big(\ell_x^y(t)\big) \le \ell_{f(x)}^{f(y)}(t)$, d.h.,

$$f\big((1-t)x + ty\big) \le (1-t)f(x) + tf(y).$$

Bemerkung F5.6. Falls wir $\le$ mit $<$ bzw. $\ge$ bzw. $>$ ersetzen, haben wir die Definition von *streng konvex* bzw. *konkav* bzw. *streng konkav.*

Bemerkung F5.7. Für $a < b < c \in I$ gilt $b = \ell_a^c(t)$ mit $t = \frac{b-a}{c-a}$ und $1 - t = \frac{c-b}{c-a}$. Die folgenden vier Ungleichungen sind algebraisch äquivalente Umformulierungen der Konvexitätsungleichung:

$$f(b) \le \frac{c-b}{c-a} f(a) + \frac{b-a}{c-a} f(c),$$
$$\frac{f(b)-f(a)}{b-a} \le \frac{f(c)-f(a)}{c-a},$$
$$\frac{f(c)-f(a)}{c-a} \le \frac{f(c)-f(b)}{c-b},$$
$$\frac{f(b)-f(a)}{b-a} \le \frac{f(c)-f(b)}{c-b}.$$

Geometrisch heißen alle vier Formulierungen, dass $f(b)$ unterhalb der Sekante durch $f(a)$ und $f(c)$ liegt. Für eine konvexe Funktion f ist der Differenzenquotient $\frac{f(y)-f(x)}{y-x}$ damit monoton steigend sowohl in x als auch in y.

Satz F5.8. *Eine differenzierbare Funktion $f\colon I \to \mathbb{R}$ ist genau dann konvex, wenn f' monoton steigend ist.*

Beweis. Sei f konvex und seien $x < y \in I$. Weil für kleines h gilt

$$\frac{f(x+h)-f(x)}{h} \le \frac{f(y)-f(x)}{y-x} \le \frac{f(y+h)-f(y)}{h},$$

folgt, dass $f'(x) \le \frac{f(y)-f(x)}{x-y} \le f'(y)$.

Umgekehrt seien $a < b < c \in I$. Nach dem Mittelwertsatz gibt es $x \in (a,b)$ und $y \in (b,c)$ mit

$$\frac{f(b)-f(a)}{b-a} = f'(x) \le f'(y) = \frac{f(c)-f(b)}{c-b}. \qquad \square$$

Beispiel F5.9. Die Exponentialfunktion ist konvex (weil die Ableitung steigend ist). Es folgt z.B., dass

$$e^{(x+y)/2} \le \frac{e^x + e^y}{2}.$$

Für $a = e^x$ und $b = e^y$ ist dies die Ungleichung zwischen aritmetischem und geometrischem Mittel:

$$\sqrt{ab} \le \frac{a+b}{2}.$$

Korollar F5.10. *Eine zweimal differenzierbare Funktion f ist genau dann konvex, wenn $f'' \ge 0$.*

Beweis. Falls f' differenzierbar ist, gilt $f'' \ge 0$ genau dann, wenn f' monoton steigend ist. $\square$

F6. Taylorapproximation

a. *Taylorpolynome*

Definition F6.1. Für $n \in \mathbb{N}$ sei $f : I \to \mathbb{R}$ eine n-mal differenzierbare Funktion auf dem Intervall I. (Für $n = 0$ verstehen wir damit, f ist stetig.) Für $p \in I$ definieren wir $T_n(x) = T_{n,p}^f(x)$, das n^{te} *Taylorpolynom* von f an der Entwicklungsstelle p, wie folgt:

$$T_{n,p}^f(x) := \sum_{k=0}^{n} f^{(k)}(p) \frac{(x-p)^k}{k!}.$$

Die Bedingungen oben setzen wir in diesem Abschnitt voraus.

Bemerkung F6.2. Für jedes $0 \le k \le n$ gilt $T_n^{(k)}(p) = f^{(k)}(p)$. (Das Taylorpolynom ist das einzige Polynom vom Grad höchtens n mit dieser Eigenschaft).

Bemerkung F6.3. Die Ableitung des Taylorpolynoms ist das Taylorpolynom der Ableitung:

$$\left(T^f_{n,p}\right)'(x) = \sum_{k=0}^{n-1} f^{(k+1)}(p)\frac{(x-p)^k}{k!} = T^{(f')}_{n-1,p}(x).$$

Bemerkung F6.4. Wir erwarten, dass T_n eine sehr gute Approximation für f in der Nähe von p ist (sogar, die bestmögliche Approximation durch ein Polynom von Grad höchstens n). Das heißt, wir erwarten, dass das n^{te} *Restglied* $f(x) - T_n(x)$ klein ist.

Satz F6.5 (Taylorapproximation). *Es gilt*

$$\lim_{x\to p} \frac{f(x) - T_n(x)}{(x-p)^n} = 0.$$

Beweis. Wir beweisen den Satz mittels Induktion über n. Für $n = 0$ ist diese Formel einfach eine Formulierung der Stetigkeit: $\lim f(x) - f(p) = 0$.

Für $n = 1$ ist sie eine Formulierung der Differenzierbarkeit, dass T_1 eine gute lineare Approximation zu f ist. (Hier muss man die Versuchung widerstehen, die Regel von l'Hôpital anzuwenden. Es gilt zwar

$$\lim_{x\to p} \frac{f(x) - T_1(x)}{x-p} = \lim_{x\to p} f'(x) - T_1'(x) = \lim_{x\to p} f'(x) - f'(p)$$

falls letzteres existiert. Wir haben aber nicht vorausgesetzt, dass f' stetig ist, sondern nur, dass f' existiert.)

Sei nun $n > 1$. Als Induktionsvoraussetzung dürfen wir den Fall $n-1$ benutzen, was wir für die Funktion f' tun:

$$\lim_{x\to p} \frac{f'(x) - \left(T^f_{n,p}\right)'(x)}{(x-p)^{n-1}} = \lim_{x\to p} \frac{f'(x) - T^{f'}_{n-1,p}(x)}{(x-p)^{n-1}} = 0.$$

Jetzt benutzen wir doch die Regel von l'Hôpital:

$$\lim_{x\to p} \frac{f(x) - T^f_{n,p}(x)}{(x-p)^n} = \lim_{x\to p} \frac{f'(x) - \left(T^f_{n,p}\right)'(x)}{n(x-p)^{n-1}} = 0. \qquad \square$$

Bemerkung F6.6. Durch die Induktion im obigen Beweis benutzen wir die Regel von l'Hôpital insgesamt $(n-1)$-mal, um zu schließen, dass:

$$\lim_{x\to p} \frac{f(x) - T_n(x)}{(x-p)^n} = \cdots = \lim_{x\to p} \frac{f^{(n-1)}(x) - T_n^{(n-1)}(x)}{n!(x-p)}.$$

Danach benutzen wir die Definition der Differenzierbarkeit von $f^{(n-1)}$. Falls f n-mal *stetig* differenzierbar wäre, könnten wir l'Hôpital ein n^{tes} Mal benutzen und danach die Definition der Stetigkeit.

Bemerkung F6.7. Der Satz sagt, für festes n ist $T_n(x)$ eine sehr gute Approximation für $f(x)$ im Limes $x \to p$. In dem Sinne ist er eine Version höherer Ordnung von unserer Überlegungen bei der Definition von Differenzierbarkeit.

b. Lagrange'sche Form

Satz F6.8 (Lagrange'sche Form des Restglieds). *Zu jedem $x \in I \setminus \{p\}$ gibt es ein y zwischen p und x, so dass*

$$f(x) = T^f_{n-1,p}(x) + f^{(n)}(y)\frac{(x-p)^n}{n!}.$$

Bemerkung F6.9. Wenn man auf der rechten Seite y durch p ersetzt, bekommt man $T_n(x)$.

Beweis. Wir schreiben $R(t) = f(t) - T_{n-1}(t)$ für das Restglied. Wir definieren

$$h(t) := R(t)\frac{(x-p)^n}{n!} - R(x)\frac{(t-p)^n}{n!}.$$

Es gilt $T_{n-1}^{(n)} \equiv 0$ und deshalb

$$h^{(n)}(t) = f^{(n)}(t)\frac{(x-p)^n}{n!} - R(x).$$

Das heißt, wir suchen nach einer Nullstelle y von $h^{(n)}$. Für $0 \le k \le n-1$ gilt $R^{(k)}(p) = 0$ und deshalb auch $h^{(k)}(p) = 0$. Wir definieren $y_0 := x$; es ist klar, dass $h(y_0) = 0$. Rekursiv wissen wir, dass $h^{(k)}(p) = 0 = h^{(k)}(y_k)$; der Mittelwertsatz gibt uns ein y_{k+1} zwischen p und y_k mit $h^{(k+1)}(y_{k+1}) = 0$. Dann ist y_n die gesuchte Nullstelle y. $\square$

Bemerkung F6.10. Ein alternativer Beweis benutzt die zwei Funktionen $g(t) := T^f_{n-1,t}(x)$ und $h(t) := (x-t)^n$ auf dem Intervall zwischen x und p. Wendet man den Mittelwertsatz von Cauchy auf g und h an, bekommt man das gewünschte y.

Bemerkung F6.11. Die Lagrange'sche Form ist eine Version höherer Ordnung des Mittelwertsatzes.

`Ende der Vorlesung 2008 Juni 30`

Beispiel F6.12. Sei $f = \exp$ die Exponentialfunktion und $p = 0$. Dann gilt $T_n(x) = \sum_0^n {x^k}/{k!}$. Für $x > 0$ wissen wir aus der Lagrange'schen Form, dass es $y \in (0, x)$ gibt mit

$$0 < \sum_{k=n}^{\infty} \frac{x^k}{k!} = \exp x - T_{n-1}(x) = (\exp y)\frac{x^n}{n!} < (\exp x)\frac{x^n}{n!}.$$

Zum Beispiel für $x = 0$ gilt $0 < e - T_{n-1}(1) < {e}/{n!}$.

Satz F6.13. *Die Euler'sche Zahl e ist irrational.*

Beweis. Nehmen wir an, $e = {p}/{q}$ für $p, q \in \mathbb{N}$. Weil wir schon wissen, ${8}/{3} < e < {11}/{4}$, gilt $q > 4$. Für $n := q + 1$ lautet die obige Restgliedformel $q!e - q!T_q(1) < {e}/{n}$. Die linke Seite ist aber eine positive ganze Zahl, und die rechte Seite ist kleiner als 1. Widerspruch! $\square$

c. Taylorreihen

Falls $f \in C^\infty(I)$, gibt es zu jedem $n \in \mathbb{N}$ ein Taylorpolynom T_n. Dann ist es eine natürliche Frage, ob für festes x Konvergenz $T_n(x) \to f(x)$ für $n \to \infty$ zu erwarten ist.

Definition F6.14. Seien $f \in C^\infty(I)$ und $p \in I$. Zu jedem $n \in \mathbb{N}$ gibt es dann ein Taylorpolynom $T_n = T^f_{n,p}$. Für $x \in \mathbb{R}$ heißt dann die Reihe

$$\lim_{n\to\infty} T_n(x) = \sum_{k=0}^{\infty} f^{(k)}(p)\frac{(x-p)^k}{k!}$$

die *Taylorreihe* von f (an der Entwicklungsstelle p, ausgewertet im Punkt x).

Beispiel F6.15. Die Funktion $\exp\colon \mathbb{R} \to \mathbb{R}$ hat in $p \in \mathbb{R}$ die Taylorreihe $\sum \exp p\,{}^{(x-p)^k}/_{k!}$. Diese ist $\exp p$ mal die Reihe, die $\exp(x-p)$ definiert. D.h., sie konvergiert für jedes $x \in \mathbb{R}$ und zwar gegen $\exp x = \exp p \exp(x-p)$.

Definition F6.16. Seien $f \in C^\infty(I)$ und $p \in I$. Falls es $\varepsilon > 0$ gibt, so dass für jedes $x \in B_\varepsilon(p) \cap I$ die Taylorreihe gegen $f(x)$ konvergiert (d.h., $T_n(x) \to f(x)$), dann heißt f *analytisch* in p. Falls f in jedem $p \in I$ analytisch ist, sagen wir einfach f ist *analytisch*.

Bemerkung F6.17. Aus dem Wurzelkriterium kann man festlegen, für welches $\varepsilon > 0$ die Taylorreihe auf $B_\varepsilon(p)$ konvergent ist. Ob sie dann gegen $f(x)$ konvergiert ist eine schwierigere Frage. Dazu muss man zeigen, das Restglied $f - T_n$ konvergiert gegen Null. Ausser der Langrange'schen Form helfen dabei auch die vielen anderen bekannten Abschätzungen des Restglieds, die wir hier nicht erwähnen.

Beispiel F6.18. Man kann zeigen, dass die Funktion $f\colon x \mapsto -\log(1-x)$, die nur für $x < 1$ definiert ist, in jedem $p < 1$ analytisch ist. Für $k > 0$ gilt nämlich

$$f^{(k)}(x) = \frac{(k-1)!}{(1-x)^k}.$$

Die Taylorreihe an der Entwicklungsstelle $p = 0$ (z.B.) ist deshalb

$$\sum_{k=1}^{\infty} \frac{x^k}{k}.$$

Aus dem Quotienten- oder Wurzelkriterium wissen wir, diese konvergiert für $|x| < 1$ und divergiert für $|x| > 1$.

Tatsächlich konvergiert sie genau für $-1 \le x < 1$ und zwar immer gegen $f(x)$. Zum Beispiel gilt

$$1 - {}^1/_2 + {}^1/_3 - \cdots = \log 2.$$

Aus der Lagrange'schen Form $\frac{1}{n}\big(\frac{x}{1-y}\big)^n$ des Restglieds (mit y zwischen 0 und x) kann man dieses nur für $|x| < {}^1/_2$ zeigen. Ratsam ist es, die „Cauchy'sche Form“ anzuwenden (und für $x = -1$ auch den „Abel'schen Grenzwertsatz“).

d. Ein nichtanalytisches Beispiel

Die Exponentialfunktion wächst schneller als jedes Polynom. Genauer gesagt gilt folgendes:

Lemma F6.19. *Für jedes Polynom $P\colon \mathbb{R} \to \mathbb{R}$ gilt*

$$\lim_{x\to+\infty} \frac{e^x}{P(x)} = +\infty.$$

Beweis. Wir benutzen Induktion über den Grad n von P. Für $n = 0$ haben wir schon gezeigt, $\lim e^x = +\infty$. Für $n > 0$ hat P' Grad $n-1$. Aus der Induktionsvoraussetzung gilt damit $\lim {}^{e^x}/_{P'(x)} = +\infty$. Dann sagt die Regel von l'Hôpital,

$$\lim_{x\to+\infty} \frac{e^x}{P(x)} = \lim_{x\to+\infty} \frac{e^x}{P'(x)} = +\infty. \qquad \square$$

Korollar F6.20. $\lim_{y\to 0^+} P\big({}^1/_y\big)e^{-1/y} = 0.$ $\square$

Satz F6.21. *Sei $f\colon \mathbb{R} \to \mathbb{R}$ die Funktion*

$$f(x) := \begin{cases} e^{-1/x}, & x > 0, \\ 0, & x \le 0. \end{cases}$$

Dann ist $f \in C^\infty(\mathbb{R})$, aber f ist in 0 nicht analytisch.

Beweis. Wir definieren Polynome P_k rekursiv: $P_0(x) = 1$,

$$P_{k+1}(x) := x^2\big(P_k(x) - P_k'(x)\big).$$

Dann gilt

$$f^{(k)}(x) = P_k\big({}^1/_x\big)e^{-1/x}$$

für jedes $k \in \mathbb{N}$ und jedes $x > 0$. Für $x < 0$ gilt offensichtlich $f^{(k)}(x) = 0$. Es gilt auch $f_-^{(k)}(0) = 0$.

Um zu zeigen, dass $f \in C^\infty$, müssen wir deshalb zeigen, dass die rechtsseitigen Ableitungen $f_+^{(k)}(0)$ auch 0 sind. Es gilt aber

$$f_+^{(k+1)}(0) := \lim_{h\to 0^+} \frac{f^{(k)}(h) - 0}{h} = \lim_{h\to 0^+} {}^1/_h P_k\big({}^1/_h\big)e^{-1/h}$$

und letzteres ist (aus dem Korollar zum Lemma) Null, weil $xP(x)$ wieder ein Polynom ist.

Weil $f^{(k)}(0) = 0$ für alle k gilt, ist die Taylorreihe $\sum 0x^k$. Für jedes x konvergiert diese Reihe offensichtlich gegen 0, welches nur für $x \le 0$ gleich $f(x)$ ist. Das heißt, es gibt keinen Ball um 0, wo die Konvergenz immer gegen $f(x)$ ist; die Funktion f ist in 0 nicht analytisch. $\square$

`Ende der Vorlesung 2008 Juli 1`

G. KLEINE TEILMENGEN IN $\mathbb{R}$

G1. Abzählbare Mengen

Wir errinnern uns, zwei Mengen X und Y sind *gleich groß* (oder *gleichmächtig*), falls es eine Bijektion (eine injektive und surjektive Abbildung) $f\colon X \to Y$ gibt. Falls X und Y gleich groß sind und Y und Z gleich groß sind, folgt es, dass X und Z gleich groß sind.

Falls X und $\mathbb{N}$ gleich groß sind, sagen wir, X ist *abzählbar unendlich*. (Wenn $f\colon \mathbb{N} \to X$ bijektiv ist, können wir die Elemente von X wie folgt „abzählen“: $f(0), f(1), f(2), \ldots$.) Wir schreiben $\#X = \aleph_0$. (Der erste hebräische Buchstabe $\aleph$ heißt Aleph.) Falls X entweder endlich oder abzählbar unendlich ist, sagen wir, X ist *abzählbar*; falls nicht, sagen wir, X ist *überabzählbbar*.

Lemma G1.1. *Jede Teilmenge $T \subset \mathbb{N}$ ist abzählbar.*

Beweis. Wir versuchen eine streng monoton steigende Abbildung $f\colon \mathbb{N} \to T$ und eine Folge von Teilmengen $T_n \subset T \subset \mathbb{N}$ rekursiv zu definieren. Zunächst setzen wir $T_0 := T$. Der rekursive Schritt fängt mit der Frage an, ob T_n leer ist. Falls nein, setzen wir $f(n) := \min T_n$ und $T_{n+1} := T_n \smallsetminus \{f(n)\}$; wir merken, $f(n) < t$ für alle $t \in T_{n+1}$, und wir machen weiter. Falls aber $T_n = \emptyset$, brechen wir die Rekursion ab.

Falls die Rekursion in n abbricht (weil T_n leer ist), heißt das, $f\colon \{0, \ldots, n-1\} \to X$ ist nicht nur injektiv (weil streng monoton) sonder auch surjektiv. Damit gilt $\#X = n$.

Falls die Rekursion nie abbricht, haben wir eine Injektion $f\colon \mathbb{N} \to X$ definiert. Weil f streng monoton ist, gilt $f(n) \geq n$. Damit gehört jedes $x \in X$ zum Bild $f\big(\{0, \ldots, x\}\big)$. Insbesondere ist f surjektiv. $\square$

Korollar G1.2. *Falls $f\colon X \to \mathbb{N}$ injektiv ist, dann ist X abzählbar.*

Beweis. Die Teilmenge $f(X) \subset \mathbb{N}$ ist aus dem Lemma abzählbar und $f\colon X \to f(X)$ ist eine Bijektion. $\square$

Korollar G1.3. *Seien X abzählbar und $Y \subset X$. Dann ist Y abzählbar.*

Beweis. Es gibt eine Bijektion $f\colon X \to T \subset \mathbb{N}$. (Es ist uns unwichtig, aber hier ist die Teilmenge T entweder $\{0, \ldots, n-1\}$ oder $\mathbb{N}$ selbst.) Damit ist $f|_Y$ eine Injektion $Y \to \mathbb{N}$. Aus dem ersten Korollar ist Y abzählbar. $\square$

Satz G1.4. *Eine nichtleere Menge X ist genau dann abzählbar, wenn es eine surjektion $f\colon \mathbb{N} \to X$ gibt.*

Beweis. Eine Richtung ist ganz einfach: Falls X abzählbar unendlich ist, gibt es sogar eine Bijektion $\mathbb{N} \to X$. Sonst gibt es $n \in \mathbb{N}$ und eine Bijektion $f\colon \{m < n\} \to X$. Weil X nicht leer ist, gilt $n > 0$. Wir erweitern f auf eine Surjektion $f\colon \mathbb{N} \to X$, in dem wir $f(m) := f(0)$ für $m \geq n$ setzen.

Umgekehrt, sei $f\colon \mathbb{N} \to X$ surjektiv. Wir definieren $g\colon X \to \mathbb{N}$ mit

$$g(x) := \min\{n \in \mathbb{N} : f(n) = x\}.$$

Damit ist g injektiv, weil

$$g(x) = g(y) \implies x = f\big(g(x)\big) = f\big(g(y)\big) = y.$$

Aus dem Korollar oben ist X abzählbar. $\square$

Beispiel G1.5. Die ganzen Zahlen $\mathbb{Z}$ sind abzählbar unendlich, weil wir diese wir folgt abzählen können:.

$$0, -1, 1, -2, 2, \ldots.$$

(Das entspricht die Bijektion $f\colon \mathbb{N} \to \mathbb{Z}$ mit $f(2n) := n$ und $f(2n-1) := -n$.)

Beispiel G1.6. Die positiven rationalen Zahlen $\mathbb{Q}^+$ sind abzählbar unendlich, weil wir diese auch auflisten können (jetzt mit Duplikaten):

$$1/1,\ 1/2, 2/1,\ 1/3, 2/2, 3/1,\ 1/4, 2/3, 3/2, 4/2,\ \ldots.$$

Dieser Trick heißt das erste Cantor'sche Diagonalverfahren. Hier können wir auch eine Formel für die Surjektion $f\colon \mathbb{N} \to \mathbb{Q}^+$ geben: für $m < n$ gilt

$$f\left(\binom{n}{2} + m\right) = \frac{m+1}{n-m}.$$

Aus einer Kombination dieser beiden Tricks folgt, dass $\mathbb{Q}$ auch abzählbar unendlich ist. Wir zeigen allgemeiner

Satz G1.7. *Sei A eine abzählbare Indexmenge. Für jedes $\alpha \in A$ sei X_α eine abzählbare Menge. Dann ist die Vereinigung*

$$X := \bigcup_{\alpha \in A} X_\alpha$$

auch abzählbar.

Beweis. Wir dürfen annehmen, jedes X_α ist nicht leer, weil wir sonst A mit $\{\alpha \in A : X_\alpha \neq\}$ ersetzen. Falls $A = \emptyset$, gilt auch $X = \emptyset$, eine endliche Menge. Sonst gibt es sowohl eine Surjektion $g\colon \mathbb{N} \to A$ als auch für jedes $\alpha \in A$ eine Surjektion $f_\alpha\colon \mathbb{N} \to X_\alpha$. Wir definieren $f\colon \mathbb{N} \to X$ mit dem Diagonalverfahren:

$$f\left(\binom{n}{2} + m\right) = f_{g(m)}(n-m).$$

Das heißt, wir listen die Elemente von X wir folgt auf:

$$f_{g(0)}(0),\ f_{g(0)}(1), f_{g(1)}(0),\ f_{g(0)}(2), f_{g(1)}(1), f_{g(2)}(0),\ \ldots.$$

$\square$

Sind die reellen Zahlen auch abzählbar? Gibt es überhaupt überabzählbare Mengen? Das zweite Cantor'sche Diagonalverfahren antwortet diese Fragen.

Satz G1.8. *Sei P die Menge $P := \{T \subset \mathbb{N}\}$ aller Teilmengen von $\mathbb{N}$. Dann ist P überabzählbar.*

Beweis. Nehmen wir an, P ist abzählbar, dann gibt es eine Surjektion $F\colon \mathbb{N} \to P$. Sei $T \subset \mathbb{N}$ wie folgt definiert:

$$T := \{n \in \mathbb{N} : n \notin F(n)\}.$$

Für alle $n \in \mathbb{N}$ gilt $T \neq F(n)$ (weil $n \in T \iff n \notin F(n)$). Das heißt, $T \in P$ ist eine Teilmenge, die nicht zum Bild $F(\mathbb{N})$ gehört. Widerspruch! □

Eine Binärdarstellung $b^0.b^1b^2b^3\cdots$ einer reellen Zahl $x \in [0,2]$ kann man mit der Teilmenge $\{n : b_n = 1\}$ identifizieren. Das zeigt noch nicht, dass $[0,2]$ überabzählbar ist, weil die Binärdarstellung nicht immer eindeutig ist. Mit einer kleinen Änderung kommen wir diesem Problem vorbei.

Satz G1.9. *Das Intervall $(0,1)$ ist überabzählbar.*

Beweis. Seien $x_1, x_2, \ldots \in (0,1)$. Es genügt, ein $x \in (0,1)$ zu konstruieren, das ungleich jedes x_i ist. Dazu sei (für jedes i)

$$0.d_i^1 d_i^2 d_i^3 \cdots = x_i$$

eine Dezimaldarstellung von x_i. Wir setzen

$$d^i := \begin{cases} 3, & \text{falls } d_i^i \geq 5, \\ 6, & \text{falls } d_i^i \leq 4. \end{cases}$$

Dann hat $x := 0.d^1d^2d^3\cdots$ diese als eindeutige Dezimaldarstellungen (weil 0 bzw. 9 nie auftauchen). Deshalb gilt $x \neq x_i$, weil die Dezimaldarstellungen ungleich sind. □

Bemerkung G1.10. Damit ist auch jede offene Menge in $\mathbb{R}$ überabzählbar.

```
Ende der Vorlesung 2008 Juli 7
```

G2. Nullmengen

Definition G2.1. Sei $I \subset \mathbb{R}$ ein Intervall. Seine *Länge* ist

$$\operatorname{len}(I) := \sup I - \inf I \in [0, +\infty].$$

Definition G2.2. Eine Teilmenge $N \subset \mathbb{R}$ heißt dann eine *Nullmenge*, wenn zu jedem $\varepsilon > 0$ es offene Intervalle I_n, $n \in \mathbb{N}$ gibt, so dass

$$N \subset \bigcup_{n=0}^{\infty} I_n \qquad \text{und} \qquad \sum_{n=0}^{\infty} \operatorname{len}(I_n) < \varepsilon.$$

Kurz sagt man: eine Nullmenge hat – für jedes $\varepsilon > 0$ – eine Überdeckung durch Intervalle mit totaler Länge kleiner als ε.

Bemerkung G2.3. Natürlich ist die leere Menge $\emptyset$ ein Beispiel einer Nullmenge. Man sollte aber den Begriff „Nullmenge“ nicht mit der leeren Menge verwechseln.

Lemma G2.4. *Jede abzählbare Teilmenge $X \subset \mathbb{R}$ ist eine Nullmenge.*

Beweis. Sei $\varepsilon > 0$ gegeben und sei $f\colon \mathbb{N} \to X$ eine Surjektion. Wir setzen

$$I_n := B_{\varepsilon/2^{n+3}}\big(f(n)\big).$$

Damit gilt $\operatorname{len}(I_n) = \varepsilon/2^{n+2}$ und

$$\sum \operatorname{len} I_n = \varepsilon\big(1/4 + 1/8 + \cdots\big) = \varepsilon/2 < \varepsilon. \qquad \square$$

Lemma G2.5. *Sei A eine abzählbare Indexmenge und für jedes $\alpha \in A$ sei X_α eine Nullmenge. Dann ist die Vereinigung $X := \bigcup X_\alpha$ wieder eine Nullmenge.*

Beweisskizze. Sei $f\colon \mathbb{N} \to A$ eine Surjektion. Für jedes $m \in \mathbb{N}$ mit $\alpha := f(n) \in A$ sei $E_m = \{I_m^n : n \in \mathbb{N}\}$ eine Überdeckung von X_α durch Intervalle I_m^n mit totaler Länge kleiner als $\varepsilon/2^{m+1}$. Dann ist die Vereinigung $E := \bigcup_m E_m = \{I_m^n : m, n \in \mathbb{N}\}$ eine Überdeckung von X durch (abzählbar viele) Intervalle mit totaler Länge kleiner als $\sum \varepsilon/2^{m+1} = \varepsilon$. (Die Reihen, die hier auftauchen, konvergieren absolut.) □

Definition G2.6. Sei $f\colon D \to \mathbb{R}$ differenzierbar. Ein *kritischer Punkt* von f ist ein $x \in D$ mit $f'(x) = 0$. Falls x ein kritischer Punkt von f ist, heißt $f(x)$ ein *kritischer Wert* von f.

Die höher-dimensionale Erweiterung des nächsten Satzes heißt der Satz von Sard und ist u.a. in der Differentialtopologie sehr wichtig.

Satz G2.7. *Sei $f\colon \mathbb{R} \to \mathbb{R}$ stetig differenzierbar (d.h., $f \in C^1$). Dann ist die Menge aller kritischen Werte von f eine Nullmenge.*

Beweis. Wir zeigen dasselbe für $f\colon [0,1] \to \mathbb{R}$. Dann folgt der Satz aus dem letzten Lemma, in dem wir $\mathbb{R}$ durch abzählbar viele Intervalle $[n, n+1]$ Überdecken.

Sei $f\colon [0,1] \to \mathbb{R}$ stetig differenzierbar und sei $\varepsilon > 0$. Weil $[0,1]$ folgenkompakt ist, ist die stetige Ableitung $f'\colon [0,1] \to \mathbb{R}$ sogar gleichmäßig stetig. Deshalb gibt es $\delta > 0$, so dass für alle $x, y \in [0,1]$ gilt

$$|x - y| < \delta \implies \big|f'(x) - f'(y)\big| < \varepsilon.$$

Wir wählen $n > 1/\delta$ und betrachten die Teilintervalle $I_j := \big[j/n, (j+1)1/n\big]$ für $0 \leq j < n$.

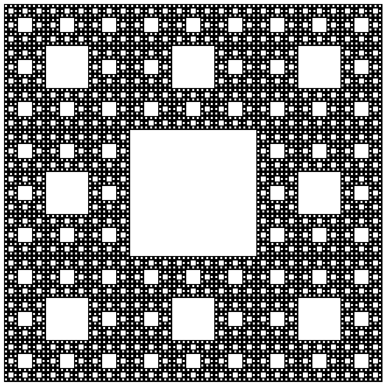
Sei j jetzt fest. Falls es einen kritischen Punkt $x \in I_j$ gibt, dann gilt $|f'(y)| < \varepsilon$ für alle $y \in I_j$ (weil $|y - x| < \delta$ und $f'(x) = 0$). Nach dem Zwischenwertsatz ist $f(I_j)$ ein Intervall $[m, M]$, wobei $m = f(y)$ das Minimum von f auf I_j ist und $M = f(z)$ das Maximum. Nach dem Mittelwertsatz gilt dann

$$\operatorname{len}\big(f(I_j)\big) = f(z) - f(y) < \varepsilon|z - y| \leq \varepsilon/n.$$

(Hier folgt alles natürlich aus unserer Voraussetzung, dass I_j einen kritischen Punkt enthält.)

Jeder kritische Wert von f gehört natürlich zu einem solchen $f(I_j)$ (wo I_j einen kritischen Punkt enthält). Die Vereinigung dieser Intervalle $f(I_j)$ überdeckt die kritischen Werte mit totaler Länge kleiner als ε. □

Definition G2.8. Die *Cantormenge* $C \subset [0,1]$ ist die Menge aller Zahlen, die eine Ternärdarstellung $0.d^1d^2d^3\cdots_3$ ohne 1 besitzen, d.h. jedes $d^i \in \{0,2\}$.

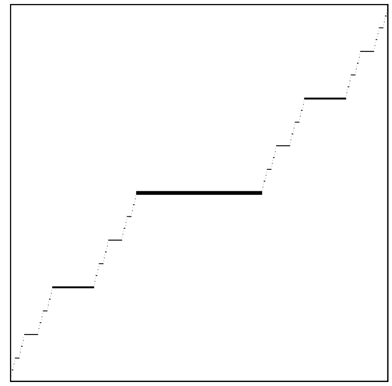


Bemerkung G2.9. Man kann C auch mit einem sogenannten iterierten Funktionensystem konstruieren. Dazu seien $g(x) := x/3$ und $h(x) := (x+2)/3$ zwei stetige Abbildungen $[0,1] \to [0,1]$. Wir definieren rekursiv $C_0 := [0,1]$ und $C_{n+1} := g(C_n) \cup h(C_n) \subset C_n$. Dann gilt $C = \bigcap C_n$.

Bemerkung G2.10. Es folgt, dass C (wie alle C_n) abgeschlossen und deshalb folgenkompakt ist. Jedes C_n ist die Vereinigung von 2^n Intervallen mit je Länge 3^{-n}. Es folgt, dass C eine Nullmenge ist.

Bemerkung G2.11. Jeder Punkt $x \in C$ ist ein Häufungspunkt von C und auch ein Randpunkt von C. (D.h., $\partial C = C$.) Damit ist C auch *nirgendwo dicht*.

Definition G2.12. Die Cantorfunktion ist eine stetige und monoton steigende Surjektion $f\colon [0,1] \to [0,1]$. Für $x = 0.d^1d^2d^3\cdots_3 \in C$ gilt $f(x) := 0.b^1b^2b^3\cdots_2$, wobei $b^i := d^i/2$. Auf jedem Intervall in $[0,1] \smallsetminus C$ ist f konstant.



Bemerkung G2.13. Diese Funktion ist eine Surjektion $C \to [0,1]$. Damit ist C überabzählbar.

Bemerkung G2.14. Für $x \in C$ existiert $f'(x)$ nicht. Für $x \in [0,1] \smallsetminus C$ gilt $f'(x) = 0$.

Bemerkung G2.15. In dem man skalierte Kopien von dieser Funktion aufsummiert, kann man auch eine Funktion g konstruieren, die stetig und *streng* monoton ist, so dass ausserhalb einer Nullmenge (wo g' nicht existiert) gilt $g' = 0$.

`Ende der Vorlesung 2008 Juli 8`