Analysis

für Informatiker und Wirtschaftsinformatiker

Simon Fuhrmann
Christian M. Meyer

1. April 2005

Inhaltsverzeichnis

1 Zahlen, Induktion, Mengen

Vorlesung 27.10.2003

Wir stellen zuerst einige allgemeingültige *Rechenregeln* zusammen, die man **Axiome** nennt. Aus diesen Regeln leiten wir *neue* Sätze, Formeln und Ergebnisse her.

1.1 Die Körperaxiome

Definition §1.1

Wir gehen von einer Menge von Zahlen K aus (z.B. die Menge der reellen Zahlen), die zwei spezielle Elemente 0 und 1 enthalten soll. Weiter soll es zwei Verknüpfungen $+$ und $\cdot$ geben, die zwei beliebigen Elementen in K neue Elemente zuordnen. Zum Beispiel:

$$\begin{aligned} x + y &= u \\ x \cdot y &= v \end{aligned}$$

Dabei sollen folgende Axiome gelten, für alle x, y, z in K.

Kommutativgesetz

(K_+) $\quad x + y = y + x$
$(K_\cdot)$ $\quad x \cdot y = y \cdot x$

Assoziativgesetz

(A_+) $\quad (x + y) + z = x + (y + z)$
$(A_\cdot)$ $\quad (x \cdot y) \cdot z = x \cdot (y \cdot z)$

Distributivgesetz

(D) $\quad (x + y) \cdot z = x \cdot z + y \cdot z$
$\quad x \cdot (y + z) = x \cdot y + x \cdot z$

Existenz von Neutralelementen

(N_+) $\quad x + 0 = 0 + x = x$
$(N_\cdot)$ $\quad x \cdot 1 = 1 \cdot x = x$
$\quad x \cdot 0 = 0 \cdot x = 0$

Existenz von inversen Elementen

(I_+) Zu jedem x gibt es ein y mit $x + y = 0$
$(I_\cdot)$ Zu jedem $x \neq 0$ gibt es ein z mit $x \cdot z = 1$

Schreibe $y = -x$, $z = x^{-1} = \frac{1}{x}$

Sind alle diese Axiome erfüllt, so heißt $(K, 0, 1, +, \cdot)$ ein **Körper**. Beispiele dafür sind die reellen Zahlen $\mathbb{R}$ und die rationalen Zahlen $\mathbb{Q}$.

Aus den Körperaxiomen folgen schnell weitere einfache Regeln. Zum Beispiel:

$$(-1) \cdot x = -x$$

Herleitung:
$$\begin{aligned} 0 &= x + (-x) \\ &= x + (-1) \cdot x \\ &= 1 \cdot x + (-1) \cdot x \\ &= (1 + (-1)) \cdot x \\ &= 0 \cdot x = 0 \end{aligned}$$

Folgerung: $$(-x) \cdot (-y) = x \cdot y$$

denn:
$$\begin{aligned} (-x) \cdot (-y) &= (-1) \cdot x \cdot (-1) \cdot y \\ &= (-1) \cdot (-1) \cdot x \cdot y \\ &= -(-1) \cdot x \cdot y \\ &= 1 \cdot x \cdot y = x \cdot y \end{aligned}$$

Derartige Regeln gelten auch universell in *allen* Körpern. Es gibt neben $\mathbb{F}_2$, $\mathbb{R}$ und $\mathbb{Q}$ viele weitere wichtige Körper; Stichwort: *Kryptographie mit elliptischen Kurven.*

1.2 Beispielkörper: $\mathbb{F}_2$

Beispiel §1.2

Der Körper der Informatik $\mathbb{F}_2 = \{0, 1\}$. Wir definieren $+$ und $\cdot$ wie folgt:

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

Hier gelten alle Körperaxiome.

Kuriosum: In $\mathbb{F}_2$ gilt $0 = 2 = 1 + 1$, und $x + x = 0$ also $-x = x$ für alle x in $\mathbb{F}_2$.

1.3 Ordnung von Zahlen

Definition §1.3

Eine weitere wichtige Eigenschaft der reellen Zahlen ist ihre **Ordnung** oder **Anordnung**.

Definition: Eine Ordnung $<$ ist eine zweistellige Relation, bezüglich derer für alle x, y, z gelten soll:

(O_1) Entweder $x < y$ oder $x = y$ oder $y < x$, genau *einer* dieser Fälle tritt ein.
(O_2) Wenn gilt $x < y$ und $y < z$, dann soll auch gelten $x < z$

Ein Körper K ist angeordnet (kurz: ein *A.K.*), wenn $<$ eine Ordnung auf K ist, mit:

(O_{K+}) Wenn $x < y$, dann $x + z < y + z$
($O_{K\cdot}$) Wenn $x < y$ und wenn $z > 0$, dann $x \cdot z < y \cdot z$

Beispiele dafür sind $\mathbb{R}$ und $\mathbb{Q}$.

Beispiel: $\mathbb{F}_2$ ist *kein* *A.K.*! Denn wäre $\mathbb{F}_2$ bezüglich $<$ ein *A.K.*, so würde gelten:

$$\begin{aligned} x < y &\quad\rightsquigarrow\quad \underbrace{x+x}_{=0} < y + x \\ x < y &\quad\rightsquigarrow\quad x + y < \underbrace{y+y}_{=0} \end{aligned}$$

Dies wäre ein Widerspruch zu (O_1). □

Vereinbarung: $x \geq y$ falls $x > y$ oder $x = y$ Vorlesung 30.10.2003

x heißt **positiv** wenn $x > 0$, **negativ** wenn $x < 0$, **nicht positiv** wenn $x \leq 0$ und **nicht negativ** wenn $x \geq 0$.

Behauptung: Das Produkt von zwei positiven oder zwei negativen Elementen (Zahlen) ist positiv, das Produkt von einem positiven und einem negativen Element ist negativ.

Beweis: Falls $x, y > 0$, so ist $x \cdot y > 0 \cdot y = 0$. Ist $x < 0$, so ist $x + (-x) < -x$, also $-x > 0$. Ist $x, y < 0$ so ist $-x, -y > 0$ und $x \cdot y = -x \cdot (-y) > 0 \Rightarrow$ wenn $x, y < 0$, so ist ihr Produkt > 0. Ist $x > 0$, $y < 0$, so gilt $x \cdot y < 0 \cdot x = 0$.
Anmerkung: Insbesondere gilt für alle x, dass $x^2 \geq 0$, also $1 = 1^2 \geq 0$. Fazit: Die üblichen Vorzeichenregeln gelten also alle.

Bemerkung: Jeder angeordnete Körper ist unendlich. Denn:

$$1 > 0 \Rightarrow 1 + 1 > 1 \Rightarrow 1 + 1 + 1 > 1 + 1 \Rightarrow ...$$

1.4 Der Absolutbetrag $|x|$

Definition §1.4

Der **Absolutbetrag** (oder Betrag) eines Elementes x eines *A.K.* ist definiert wie folgt:

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

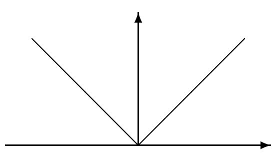


Abbildung 1: Die Betragsfunktion

Es gilt:

$$\begin{aligned} |x \cdot y| &= |x| \cdot |y| && (1) \\ |x| &= |-x| \geq 0 && (2) \\ |x + y| &\leq |x| + |y| \quad \text{(Dreiecksungleichung)} && (3) \\ |x - y| &\geq ||x| - |y|| && (4) \end{aligned}$$

Beweis: (1) und (2) sind klar. Zu (3):

$$\begin{array}{rrcccc} & |x| & \geq & \pm x & & \\ \text{also} & |x|+|y| & \geq & |x|+y & \geq & x+y \\ \text{und} & |x|+|y| & \geq & |x|-y & \geq & \underbrace{-x-y}_{=-(x+y)}, \\ \text{also} & |x|+|y| & \geq & |x+y| & & \end{array}$$

Zu (4):

$$\begin{array}{rrcccc} & |x| & = & |x+y-y| & \leq & |x-y|+|y| \\ & |y| & = & |y+x-x| & \leq & |y-x|+|x| \\ & |x|-|y| & \leq & |x-y| & & \\ & |y|-|x| & \leq & \underbrace{|y-x|}_{=|x-y|} & & \\ \Rightarrow & ||x|-|y|| & \leq & |x-y| & & \end{array}$$

1.5 Natürliche Zahlen und Induktion

Definition §1.5

1.5.1 Die Peano-Axiome

Schule: $\mathbb{N} = \{0, 1, 2, 3, ...\}$

Schreibe: $s(n) = n+1$, wobei $s(n) =$ Nachfolger von n, dann erfüllt $(\mathbb{N}, s, 0)$ die so genannten **Peano-Axiome**.

(P_1) Es gibt kein n mit $s(n) = 0$
(P_2) Ist $s(n) = s(m)$, dann ist $n = m$
(P_3) Axiom der vollständigen Induktion: Ist X eine Teilmenge von $\mathbb{N}$ mit $0 \in X$ und gilt: $x \in X \Rightarrow s(x) \in X$, so ist $X = \mathbb{N}$.

Satz: Erfüllt $(N, s', 0')$ die Peano-Axiome, so *ist* $N = \mathbb{N}$, genauer: Es gibt genau eine Abbildung $f\colon N \to \mathbb{N}$, die jedem $t \in N$ eine natürliche Zahl zuordnet, mit:

$$\begin{array}{rcl} f(0') & = & 0 \\ f(s'(t)) & = & f(t)+1 \end{array}$$

Beispiel: Wir behaupten $0+1+2+...+n = \frac{n\cdot(n+1)}{2}$ gilt für jede natürliche Zahl n.

Induktionsanfang: $n = 0 \Rightarrow$ Formel ist richtig. Auch aus $n = 1 \Rightarrow$ Formel ist richtig.

Induktionsschritt: Wir nehmen an das gilt für $0, 1, 2, ..., n$ und wir zeigen: es gilt auch für $n+1$.

$$\begin{array}{rcl} 0+1+\cdots+n+(n+1) & \overset{\text{IV}}{=} & \dfrac{n\cdot(n+1)}{2}+(n+1) \\ = \dfrac{n\cdot(n+1)}{2}+\dfrac{2\cdot(n+1)}{2} & = & \dfrac{(n+1)(n+2)}{2} \end{array}$$

Also gilt die Formel für $n+1$, wenn sie auch für n gilt.

Behauptung: Damit gilt die Formel für *alle natürlichen Zahlen*!

Setze: $X = \{n \in \mathbb{N} \mid \text{Formel gilt für } n\}$

Nach Induktionsanfang (I.A.) ist $0 \in X$, nach Induktionsschritt (I.S.) gilt, wenn $n \in X$, dann $s(n) = n + 1 \in X$.

Nach (P_3) ist $X = \mathbb{N}$.

Bemerkung: Wir können multiplizieren $(\cdot)$ und addieren $(+)$ mittels $s(\ldots)$ hinschreiben, durch:

$$\begin{aligned} 0 + n &= 0 \\ s(n) + m &= s(n+m) \end{aligned}$$

sowie:

$$\begin{aligned} 0 \cdot n &= 0 \\ s(n) \cdot m &= n \cdot m + m \end{aligned}$$

Es ist mühsam, Assoziativität, Kommutativität, etc. nachzuweisen.

1.5.2 Die ganzen Zahlen $\mathbb{Z}$

Die ganzen Zahlen $\mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \ldots\}$ stellen wir uns als vorzeichenbehaftete, natürliche Zahlen vor, mit den üblichen Rechenregeln. Die Menge $\mathbb{N}$ der natürlichen Zahlen und die Menge $\mathbb{Z}$ der ganzen Zahlen sind keine Körper.

Weil: $2 \cdot x = 1$ hat keine Lösung in $\mathbb{Z}$.

1.5.3 Die rationalen Zahlen $\mathbb{Q}$

Man erweitert $\mathbb{Z}$ zum Körper $\mathbb{Q} = \{\frac{a}{b} \mid a, b \in \mathbb{Z}, b \neq 0\}$ mit den üblichen Rechen- und Kürzungsregeln ($\frac{a}{b} = \frac{a \cdot c}{b \cdot c}, c \neq 0$).

$(\mathbb{Q}, 0, 1, +, \cdot, \geq)$ ist ein angeordneter Körper.

1.6 Mengen

Vorlesung 03.11.2003
Definition §1.6

Mengen sind die Bausteine der gesamten Mathematik. Unter einer Menge stellen wir uns eine Zusammenfassung gewisser **Elemente** (die ihrerseits Mengen sind) vor. Wenn ein Element x in einer Menge A ist, schreiben wir $x \in A$. Wenn x nicht in A ist, schreiben wir $x \notin A$. Sind A und B Mengen, und gilt für jedes $x \in A$, dass $x \in B$, so schreiben wir $A \subseteq B$ (A ist **Teilmenge** von B).

1.6.1 Existenz, Extensionalität und Aussonderung

Hier sind Axiome von *Zermelo Fraenkel*:

- (Ex) Existenz. Es gibt Mengen.
- (Ext) Extensionalität. $X = Y$ genau dann, wenn X und Y die gleichen Elemente haben, $X \subseteq Y$ und $Y \subseteq X$.
- (Aus) Aussonderung. Ist φ eine Eigenschaft und X eine Menge, so auch $\{x \in X \mid x \text{ hat Eigenschaft } \varphi\}$.

Beispiel: $\mathbb{R}$ ist eine Menge $\rightsquigarrow \{r \in \mathbb{R} \mid r > 0\}$ ist eine Menge.

Insbesondere sind **Durchschnitte** und **Komplemente** wieder Mengen.

Durchschnitt oder Schnittmenge
$A \cap B = \{a \in A \mid a \in B\}$

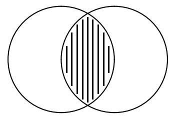


Abbildung 2: Schnittmenge

Komplement
$A \setminus B = \{a \in A \mid a \notin B\}$

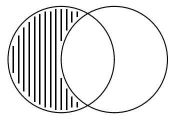


Abbildung 3: Komplement

1.6.2 Paarmengen-Axiom, Vereinigunsmengen-Axiom

($Paar$) Paarmengen-Axiom. Sind A und B Mengen, so auch $\{A, B\}$.
(Ver) Vereinigungsmengen-Axiom. Sind A und B Mengen, so auch $A \cup B = \{x \mid x \in A \text{ oder } x \in B\}$.

Vereinigungsmenge
$A \cup B = \{a \in A \text{ oder } a \in B\}$

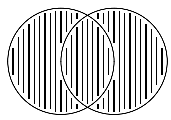


Abbildung 4: Vereinigungsmenge

Beispiel: $A = \{1, 2\}$ und $B = \{2, 3\}$

($Paar$) $\{A, B\} = \{\{1, 2\}, \{2, 3\}\}$
(Ver) $A \cup B = \{1, 2, 3\}$

1.6.3 Die leere Menge $\emptyset$

$\emptyset = \{\} = \{x \in X \mid x \neq x\}$
Für jedes x gilt also $x \notin \emptyset$

Beispiel: Mit $A = \{1, 5\}$ und $B = \{2, 511\}$ gilt $A \cap B = \emptyset$

Beachte auch: Die leere Menge ist Teilmenge jeder Menge: $\emptyset \subseteq X$. Z.B. $\emptyset \subseteq \{2, 3\}$, aber $\emptyset \notin \{2, 3\}$.

1.6.4 Potenzmengen-Axiom, Fundierungs-Axiom

(Pot) Potenzmengen-Axiom. Die Teilmengen einer Menge bilden wieder eine Menge, die Potenzmenge. $\mathcal{P}(X) = \{Y \mid Y \subseteq X\}$

($Fund$) Fundierungs-Axiom: Ist $X \neq \emptyset$ eine Menge, so gibt es ein Element $a \in X$ mit $a \cap X = \emptyset$. Folgerung: Es kommt nie vor, dass $X \in X$.

Beispiel: (Potenzmengen-Axiom) $X = \{2, 5, 11\}$
$\mathcal{P}(x) = \{\{2, 5\}, \{2, 11\}, \{5, 11\}, \{2\}, \{5\}, \{11\}, \{2, 5, 11\}, \emptyset\}$
Die Anzahl der Elemente beträgt 2^k.

Beweis durch Widerspruch (Fundierungs-Axiom):
Angenommen X ist eine Menge mit $X \in X$. Betrachte $\{X\}$, nach ($Fund$) gilt also $X \cap \{X\} = \emptyset$, d.h. $X \in X$. (Widerspruch)

1.6.5 Infinitude-Axiom

(Inf) Es gibt unendliche Mengen. Genauer gesagt: $\mathbb{N}$ existiert. Noch genauer: Es gibt eine Menge $X \neq \emptyset$ mit folgender Eigenschaft: Ist $x \in X$ so ist $x \cup \{x\} \in X$

1.6.6 Definition von $\mathbb{N}$ nach der Mengentheorie

Vorlesung 06.11.2003

Vorsicht: Natürlich stellen wir uns $\mathbb{N}$ nicht so vor; das ist die strenge mengentheoretische Definition.

$$\begin{array}{rcl} 0 & = & \emptyset \\ 1 & = & \emptyset \cup \{\emptyset\} = \{\emptyset\} = \{0\} \\ 2 & = & \{\emptyset\} \cup \{\{\emptyset\}\} = \{\emptyset, \{\emptyset\}\} = \{0, 1\} \\ 3 & = & \{\emptyset, \{\emptyset\}\} \cup \{\{\emptyset, \{\emptyset\}\}\} = \{0, 1, 2\} \\ & \vdots & \end{array}$$

$$0 \in 1 \in 2 \in 3 \in \ldots$$

1.6.7 Ersetzungs-Axiom, Auswahl-Axiom

(Ers) Ersetzungs-Axiom. Ist $\varphi(x, y)$ eine Vorschrift, die jeder Menge x eine Menge y zuordnet, und ist X eine Menge, so auch $\{y \mid x \in X \text{ und } \varphi(x, y)\}$.

($Ausw$) Auswahl-Axiom. Ist X eine Menge mit der Eigenschaft: Für alle $x, y \in X, x \neq y$ gilt $x \cap y = \emptyset$, so gibt es eine Menge $\mathbb{Z}$ mit: Für $x \in X$ ist $x \cap Z \neq \emptyset$.

Hinweis: Manchmal werden die Grundlagen über die sogenannten Kategorientheorie aufgebaut (Algebra-Topologie).

2 Die reellen Zahlen $\mathbb{R}$ und etwas Kombinatorik

2.1 Infimum und Supremum

Definition §2.1

Defnition: Ist A eine Teilmenge eines geordneten Körpers (z.B. $A \subseteq \mathbb{R}$), so heißt x die **obere Schranke** für A, falls $a \leq x$ für alle $a \in A$ gilt. Analog definiert man **untere Schranken**.

Beispiel: $A = \{x \in R \mid -1 \leq x < 3\}$
Obere Schranken: $7, 3, 4$
Untere Schranken: $-5, -2, -1$

Die kleinste obere Schranke heißt **Supremum**, die größte untere Schranke heißt **Infimum**. Schreibe:

$$\begin{aligned} \sup A = x &\iff x \text{ Supremum für } A \\ \inf B = z &\iff z \text{ Infimum für } B \end{aligned}$$

In unserem Beispiel $A = \{x \in \mathbb{R} \mid -1 \leq x < 3\}$ ist $\sup A = 3$ und $\inf A = -1$.

Falls	$\inf A \in A$	schreibe	$\inf A = \min A$
Falls	$\sup A \in A$	schreibe	$\sup A = \max A$

Die Menge A im Beispiel hat ein **Minimum** (-1) aber kein **Maximum**.

Beispiel: Setze $B = \{q \in \mathbb{Q} \mid q^2 \leq 2\}$
Diese Menge B hat obere Schranken in $\mathbb{Q}$ (z.B. 5). In dieser Menge $\mathbb{Q}$ gibt es kein Supremum für B. Dagegen gibt es in $\mathbb{R}$ natürlich ein Supremum, nämlich die irrationale Zahl $\sqrt{2}$. Das ist viel besser!

(Vollst) Ist A eine Menge der rellen Zahlen, die eine obere (untere) Schranke hat, dann hat A ein Supremum (Infimum) in $\mathbb{R}$. Man sagt, $\mathbb{R}$ ist vollständig.

Satz §2.2

Satz: Jeder angeordnete Körper, der in diesem Sinne vollständig ist, ist zu $\mathbb{R}$ **isomorph** (sieht genauso aus wie $\mathbb{R}$).

2.2 Die Fibonacci-Zahlen

Definition §2.3

Die Fibonacci-Zahlen sind rekursiv definiert wie folgt:

$f_0 = 0$, $f_1 = 1$, und $f_{n+2} = f_n + f_{n+1} \Longrightarrow 0, 1, 1, 2, 3, 5, 8, 13, \ldots$

2.3 Kombinatorische Formeln

2.3.1 Elemente der Potenzmenge

Satz §2.4

Ist A eine n-elementige Menge, so hat $\mathcal{P}(A)$ 2^n Elemente.

Beweis: Induktion nach n

I.A. $n = 0 \quad A = \emptyset \quad \mathcal{P}(\emptyset) = \{\emptyset\} \quad 2^0 = 1 \Longrightarrow$ stimmt
$n = 1 \quad A = \{a\} \quad \mathcal{P}(A) = \{\emptyset, \{a\}\} \quad 2^1 = 2 \Longrightarrow$ stimmt

I.S. $A = \{a\} \cap B$, wobei $a \notin B$ ist. A habe $n+1$ Elemente. Also hat B n Elemente. $\mathcal{P}(B)$ hat 2^n Elemente. Sei $X \subseteq A$. Ist $a \notin X$ so ist $X \subseteq B$, solche Mengen gibt es 2^n Stück.

Ist $a \in X$ betrachte $Y = X \setminus \{a\} \subseteq B$, davon gibt es auch 2^n Stück. (Beachte: $X = Y \cap \{a\}$)

Zusammen: $2^n + 2^n = 2^{n+1}$ Stück. □

Schreibe $\#A$ = Anzahl der Elemente von A

Wir haben gezeigt: $\#A = n \Longrightarrow \#\mathcal{P}(A) = 2^n$

2.3.2 Der Binomialkoeffizient

Satz §2.5

Für $k \leq n$ ist die Anzahl aller k-elementigen Teilmengen einer n-elementigen Menge der Binomialkoeffizient.

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

Beweis durch Induktion nach n:

I.A. $n = 0 \quad \emptyset \subseteq \emptyset \quad \binom{0}{0} = \frac{0!}{0!0!} = 1$

I.S. Sei $X \subseteq A$ k-elementig. Ist $a \notin X$, so ist $X \subseteq B$, solche X gibt es also $\binom{n}{k}$ Stück.

Ist $k \in X$, so ist $X \setminus \{a\} \subseteq B$ eine $k-1$-elementige Teilmenge von B. Solche Mengen gibt es $\binom{n}{k-1}$ Stück.

Insgesamt sind es also $\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}$ Stück. □

Schreibweise: Sind $a_0, a_1, \ldots, a_m$ Zahlen. Schreibe $a_0 + a_1 + \ldots + a_m = \sum\limits_{k=0}^{m} a_k$

2.3.3 Binomischer Lehrsatz

Satz §2.6

Es gilt für alle $n \in \mathbb{N}$ und alle Zahlen a, b

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

Beweis durch Induktion. Die Formel ist richtig für $n = 0$, $n = 1$ und $n = 2$.

$$
\begin{aligned}
\text{I.S.} \quad (a+b)^{n+1} &= (a+b)^n \cdot (a+b) \\
&= \left(\sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k\right)(a+b) \\
&= \left(\sum_{k=0}^{n} \binom{n}{k} a^{n+1-k} b^k\right) + \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^{k+1} \\
&= a^{n+1} + \sum_{k=0}^{n-1} \binom{n}{k+1} a^{n-k} b^{k+1} + b^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^k \qquad \square
\end{aligned}
$$

2.3.4 Geometrische Summe

Satz §2.7

Sei $x \neq 1$

$$\sum_{k=0}^{n} x^k = 1 + x + x^2 + x^3 + \ldots + x^n = \frac{1-x^{n+1}}{1-x}$$

Beweis:

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

2.3.5 Archimedisches Prinzip

Satz §2.8

(i) zu jedem $r \in \mathbb{R}$ gibt es ein $n \in \mathbb{N}$ mit $r < n$

(ii) zu allen $r, s \in \mathbb{R}$ mit $s > 0$ gibt es $n \in \mathbb{N}$ mit $sn > r$

(iii) zu allen $r, s \in \mathbb{R}$ mit $r < s$ gibt es $n \in \mathbb{N}$ mit $r + \frac{1}{n} < s$

Beweis:

(i) Widerspruchsbeweis:

Sonst gäbe es ein $r \in \mathbb{R}$ mit $r \geq n$ für *alle* $n \in \mathbb{N}$, das heißt r ist eine *obere Schranke* für $\mathbb{N}$. Setze $r_0 = \sup \mathbb{N}$, also gibt es ein $n \in \mathbb{N}$ mit $r_0 - 1 \leq n$, das heißt $r_0 \leq n + 1 < n + 2$. Ein Widerspruch. Also gilt (i).

(ii) Wähle $n > \frac{r}{s}$

(iii) Wähle $n < \frac{1}{s-r}$

2.3.6 Bernoulli'sche Ungleichung

Satz §2.9

Für $x > -1$ und $n \in \mathbb{N}$ gilt stets $(1+x)^n \geq 1 + nx$.

Beweis: Induktion nach n:

I.A. $n = 0$ $\quad (1+x)^0 \geq 1 + 0 = 1 \geq 1$

$n = 1$ $\quad (1+x)^1 \geq 1 + x \Leftrightarrow 1 + x \geq 1 + x$

I.S. $n \mapsto n+1$

$$
\begin{aligned}
(1+x)^{n+1} &= (1+x)^n(1+x) \geq (1+nx)(1+x) \\
&= 1 + nx + x + \underbrace{nx^2}_{\geq 0} \geq 1 + x(n+1)
\end{aligned}
$$

$\square$

3 Folgen und Reihen

3.1 Folgen

Definition §3.1

Definition: Eine **Folge** von (reellen) Zahlen ist eine Abbildung, die jeder natürlichen Zahl n ein Folgeglied a_n zuordnet.

Schreibe: $(a_n)_{n\in\mathbb{N}}$

3.1.1 Grenzwert oder Limes

Eine Zahl a heißt **Grenzwert** oder **Limes** der Folge $(a_n)_{n\in\mathbb{N}}$, wenn folgendes gilt:

Zu jedem $\varepsilon > 0$ gibt es ein $N = N_\varepsilon \in \mathbb{N}$, so dass für alle $k \geq N$ gilt:

$$\mid a_k - a \mid < \varepsilon$$

Wir schreiben dann $\lim\limits_{n\to\infty} a_n = a$ und sagen: Die Folge a_n konvergiert gegen a.

Beispiel:

(i) $a_n = n$

(ii) $b_n = \begin{cases} 1 & \text{falls } n \text{ gerade ist} \\ -1 & \text{falls } n \text{ ungerade ist} \end{cases}$

(iii) $c_n = c_0$ (konstante Folge)

(iv) $d_n = \frac{1}{n}$, $d_0 = 1$

Grenzwerte:

(i) $(a_n)_{n\in\mathbb{N}}$ konvergiert nicht.

(ii) $(b_n)_{n\in\mathbb{N}}$ konvergiert nicht.

(iii) $(c_n)_{n\in\mathbb{N}}$ konvergiert gegen $\lim\limits_{n\to\infty} c_n = c_0$

(iv) $(d_n)_{n\in\mathbb{N}}$ konvergiert gegen $\lim\limits_{n\to\infty} d_n = 0$

Warum? Sei $\varepsilon > 0$ gegeben. Wir wollen also $\mid d_k - 0 \mid < \varepsilon$, $\frac{1}{k} < \varepsilon$, also ist $\frac{1}{\varepsilon} < k$ gesucht.

Nach dem Archimedischen Prinzip finden wir $N \in \mathbb{N}$ mit $\frac{1}{\varepsilon} < N$. Für $k \geq N$ gilt damit $k > \frac{1}{\varepsilon}$, folglich $d_k = \frac{1}{k} < \varepsilon$.

Mit anderen Worten: $\mid d_k - 0 \mid < \varepsilon$ für $k \geq N$.

3.1.2 Häufungspunkte

Definition §3.2

Eine Zahl a heißt **Häufungspunkt** der Folge $(a_n)_{n\in\mathbb{N}}$, falls für jedes $\varepsilon > 0$ die Menge $\{k \in \mathbb{N} \big| \mid a_k - a \mid < \varepsilon\}$ unendlich ist.

Offensichtlich ist jeder Grenzwert ein Häufungspunkt. Vorlesung 13.11.2003

Beachte: Jeder Grenzwert ist ein Häufungspunkt, aber im Allgemeinen ist ein Häufungspunkt nicht notwendig ein Grenzwert.

Beispiel: $b_n = \begin{cases} 1 & \text{falls } n \text{ gerade ist} \\ -1 & \text{falls } n \text{ ungerade ist} \end{cases}$

Diese Folge hat zwei Häufungspunkte (1 und -1), aber keinen Grenzwert.

3.1.3 Konvergenz und Häufungspunkte

Satz §3.3

Satz: Eine konvergente Folge hat genau einen Häufungspunkt.

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ konvergent, $\lim\limits_{n\to\infty} a_n = a$. Sei $b \neq a$. Wir zeigen, dass b kein Häufungspunkt ist. Sei $\varepsilon = \left|\frac{b-a}{2}\right| > 0$. Wähle N_ε so, dass für alle $k \geq N_\varepsilon$ gilt: $|a_k - a| < \varepsilon$. Damit gilt

$$|b - a_k| = |b - a + a - a_k| \geq \Big| \underbrace{|b-a|}_{=2\varepsilon} - \underbrace{|a - a_k|}_{<\varepsilon} \Big| \geq \varepsilon$$

Also ist $\{k \in \mathbb{N} \,\big|\, |b - a_k| < \varepsilon\}$ endlich, das heißt b ist kein Häufungspunkt. □

Insbesondere hat eine konvergente Folge *genau einen* Grenzwert.

3.1.4 Beschränkte Folgen

Definition §3.4

Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt **beschränkt**, wenn es ein $r \in \mathbb{R}$ gibt, so dass $|a_k| \leq r$ für alle $k \in \mathbb{N}$ (d.h. $-r \leq a_k \leq r$).

Satz §3.5

Satz: Jede konvergente Folge ist beschränkt.

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ konvergent, $\lim\limits_{n\to\infty} a_n = a$ für $\varepsilon = 1$ gibt es also N_ε mit $k \geq N_\varepsilon \Rightarrow |a - a_k| < 1$.

Setze $s = \max\{|a_0|, |a_1|, \ldots, |a_{N_\varepsilon}|, |a|\}$. Für alle k gilt dann $|a_k| \leq s + 1$. □

Satz §3.6

Satz: Sind $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergent,

$$\lim_{n\to\infty} a_n = a, \qquad \lim_{n\to\infty} b_n = b$$

dann sind $(a_n + b_n)_{n\in\mathbb{N}}$, $(a_n \cdot b_n)_{n\in\mathbb{N}}$, $(c \cdot a_n)_{n\in\mathbb{N}}$ (mit $c \in \mathbb{R}$ beliebig) konvergent, und

$$\lim_{n\to\infty} (a_n + b_n) = a + b, \quad \lim_{n\to\infty} (a_n \cdot b_n) = ab, \quad \lim_{n\to\infty} (c \cdot a_n) = ca$$

Beweis für $(a_n + b_n)_{n\in\mathbb{N}}$: Sei $\varepsilon > 0$ gegeben. Es gibt N_1, $N_2 \in \mathbb{N}$ mit:

$k \geq N_1 \Rightarrow |a_k - a| < \frac{\varepsilon}{2}$
$k \geq N_2 \Rightarrow |b_k - b| < \frac{\varepsilon}{2}$

Setze $N = \max\{N_1, N_2\}$. Dann gilt für $k \geq N$:
$|(a_k + b_k) - (a + b)| \leq |a_k - a| + |b_k - b| < \varepsilon$

Beweis für $(a_n \cdot b_n)_{n\in\mathbb{N}}$: Beide Folgen sind beschränkt, also gibt es $k \in \mathbb{R}$ mit $|a_k|$, $|b_k| \leq k$ für alle $k \in \mathbb{N}$.

Sei $\varepsilon > 0$. Wähle N_1, N_2 so, dass gilt:

$k \geq N_1 \Rightarrow |a_k - a| < \frac{\varepsilon}{2k}$ (o.B.d.A. $k > 0$)
$k \geq N_2 \Rightarrow |b_k - b| < \frac{\varepsilon}{2k}$ (Setze $N = \max\{N_1, N_2\}$)

Für $k \geq N$ gilt:

$$\begin{aligned} |a_k \cdot b_k - a \cdot b| &= |a_k(b_k - b) + b(a_k - a)| \\ &\leq \underbrace{|a_k|}_{\leq k} \underbrace{|b_k - b|}_{<\frac{\varepsilon}{2k}} + \underbrace{|a_k - a|}_{<\frac{\varepsilon}{2k}} \underbrace{|b|}_{\leq k} < \varepsilon \end{aligned}$$

Beweis für $(c \cdot a_n)_{n\in\mathbb{N}}$: Diese Behauptung folgt mit $(b_n)_{n\in\mathbb{N}} = c$ (konstante Folge) aus der vorigen Behauptung. □

Satz §3.7

Satz: Ist $a_k \neq 0$ für alle $k \in \mathbb{N}$, und gilt $\lim\limits_{n\to\infty} a_n = a \neq 0$, so konvergiert die Folge $\left((\frac{1}{a})_n\right)_{n\in\mathbb{N}}$ gegen $\frac{1}{a}$.

Beweis: Sei $\varepsilon > 0$ gegeben. Wähle N so, dass gilt: $k \geq N \Rightarrow |a_k - a| < \frac{\varepsilon |a|^2}{2}, \frac{|a|}{2}$.
Dann gilt $|a_k - a| < \frac{|a|}{2}$, $|a| - |a_k| \leq |a_k - a| < \frac{|a|}{2} \Rightarrow \frac{|a|}{2} < |a_k|$.
Dann ist $\left|\frac{1}{a_k} - \frac{1}{a}\right| = \left|\frac{a-a_k}{a_k a}\right| = \frac{|a-a_k|}{|a_k|\cdot|a|} < \frac{|a|^2\varepsilon}{2\cdot|a_k|\cdot|a|} = \frac{|a|\varepsilon}{2\cdot|a_k|} < \varepsilon\frac{|a_k|}{|a_k|} = \varepsilon$. □

Beispiel:

Beispiel §3.8

$$\begin{aligned} a_n &= \frac{2n^3 + 2n^2 + n}{n^3 + 1} = \frac{n^3(2 + 2\frac{1}{n} + \frac{1}{n^2})}{n^3(1 + \frac{1}{n^3})} = \frac{2 + 2\frac{1}{n} + \frac{1}{n^2}}{1 + \frac{1}{n^3}} \\ \lim_{n\to\infty} a_n &= 2 \end{aligned}$$

Denn: $\left(2 + 2\frac{1}{n} + \frac{1}{n^2}\right)_{n\in\mathbb{N}}$ konvergiert, $\left(1 + \frac{1}{n^3}\right)_{n\in\mathbb{N}}$ konvergiert auch
$\Longrightarrow \left(\frac{1}{1+\frac{1}{n^3}}\right)_{n\in\mathbb{N}}$ konvergiert.

Damit konvergiert $\left(\frac{1}{1+\frac{1}{n^3}}\right)\left(2 + 2\frac{1}{n} + \frac{1}{n^2}\right)$.

3.1.5 Monotonie

Definition §3.9

Eine Folge $(a_n)_{n\in\mathbb{N}}$ heißt **(streng) monoton steigend** falls für alle n gilt:

$$a_n \leq a_{n+1} \qquad (a_n < a_{n+1})$$

Analog definiert man **(streng) monoton fallende** Folgen.

Satz §3.10

Satz: Eine monoton beschränkte Folge konvergiert.

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ monoton wachsend und beschränkt.
Sei $a = \sup\{a_k \mid k \in \mathbb{N}\}$. Für alle k gilt also $a_k \leq a$. Zu jedem $\varepsilon > 0$ gibt es ein N mit $a_N > a - \varepsilon$. Weiter gilt natürlich für alle $k \in \mathbb{N}$, dass $a_k < a + \varepsilon$. Für $k \geq N$ gilt $a_k > a - \varepsilon$, weil $a_k \geq a_N$ (denn die Folge ist monoton). Für $k \geq N$ gilt also $a - \varepsilon < a_k < a + \varepsilon$, das heißt:

$$|a - a_k| < \varepsilon$$

Für fallende Folgen nimmt man das Infimum. □

Beispiel §3.11

Beispiel: $a_0 = 0$, $a_{n+1} = \sqrt{6 + a_n}$ (rekursiv definiert).

Die Folge ist (a) beschränkt und (b) monoton, also konvergent.

(a) $|a_0| \leq 3$, ist $a_n \leq 3$, so $a_{n+1} \leq 3$, die Folge ist beschränkt.
(b) $a_0 = 0$, $a_1 = \sqrt{6} > a_0$
$a_{n+1} = \sqrt{6 + a_n} > \sqrt{6 + a_{n-1}} = a_n$
$\Rightarrow$ monoton, denn $0 < x < y \Leftrightarrow 0 < \sqrt{x} < \sqrt{y}$

3.1.6 Divergenz

Vorlesung 17.11.2003

Definition §3.12

Eine Folge, die nicht konvergent ist, heißt **divergent**. Wenn es zu jedem $r \in \mathbb{R}$ ein $N_r \in \mathbb{N}$ gibt mit $k \geq N_r \Rightarrow a_k \geq r$ schreiben wir:

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

Gibt es zu jedem $r \in \mathbb{R}$ ein N_r mit $k \geq N_r \Rightarrow a_k \leq r$ schreiben wir:

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

3.1.7 Teilfolgen

Definition §3.13

Ist $(a_n)_{n\in\mathbb{N}}$ eine Folge, und ist $(n_k)_{k\in\mathbb{N}}$ eine Folge natürlicher Zahlen ($\mathbb{N}$) mit $n_0 < n_1 < n_2 < \ldots$, dann heißt $(a_{nk})_{k\in\mathbb{N}}$ eine **Teilfolge** von $(a_n)_{n\in\mathbb{N}}$.

Beispiel:

$n_k = k$ $(a_n)_{n\in\mathbb{N}}$ ist Teilfolge von sich selbst.
$n_k = 2^k$ $(a_{2^k})_{k\in\mathbb{N}}$

Satz §3.14

Satz: Ist a ein Häufungspunkt der Folge $(a_n)_{n\in\mathbb{N}}$, so gibt es eine Teilfolge $(a_{nk})_{k\in\mathbb{N}}$, die gegen a konvergiert.

Beweis: Setze $n_0 = 0$ und definiere n_k rekursiv wie folgt: Zu $\varepsilon_k = \frac{1}{1+k}$ gibt es unendlich viele Folgenglieder a_l mit $l \geq n_k$ und $\mid a_l - a \mid < \varepsilon$. Setze $n_{k+1} = l$, $a_{n_{k+1}} = a_l$. Diese Teilfolge tut's. $\square$

Satz §3.15

Satz: (Bolzano-Weierstraß) Jede beschränkte Folge hat mindestens einen Häufungspunkt.

Beweis: Sei $(a_n)_{n\in\mathbb{N}}$ beschränkt. Sei $c = \inf\{a_n \mid n \in \mathbb{N}\}$, $d = \sup\{a_n \mid n \in \mathbb{N}\}$. Setze $M = \{x \in \mathbb{R} \mid$ es gibt $N \in \mathbb{N}$ mit $a_k \leq x$ für alle $k \geq N\}$. Offensichtlich ist $d \in M$, also $M \neq \emptyset$. Sei $a = \inf M$. Behauptung: a ist ein Häufungspunkt.

Für jedes $\varepsilon > 0$ gilt: Es gibt nur *endlich viele* k mit $a_k > a + \varepsilon$. Weiter gilt: $a - \varepsilon < a_k$ für unendlich viele $k \in \mathbb{N}$. Also gilt $\mid a - a_k \mid < \varepsilon$ für unendlich viele k, das heißt a ist ein Häufungspunkt. $\square$

Der größte Häufungspunkt solch einer Folge heißt **limes superior**, $\limsup\limits_{n\to\infty} a_n$, der kleinste Häufungspunkt heißt **limes inferior**, $\liminf\limits_{n\to\infty} a_n$.

3.1.8 Hilfsregeln zur Konvergenz

Bemerkung §3.16

Die folgenden Kriterien an eine Folge $(a_n)_{n\in\mathbb{N}}$ sind äquivalent:

(i) $(a_n)_{n\in\mathbb{N}}$ konvergiert, $\lim\limits_{n\to\infty} a_n = a$

(ii) Es gibt ein $k \in \mathbb{R}$ so, dass es zu jedem $\varepsilon > 0$ ein N_ε gibt mit: $l \geq N_\varepsilon \Rightarrow \mid a - a_l \mid < k \cdot \varepsilon$.

(iii) Es gibt $k \in \mathbb{R}$ so, dass es zu jedem $m \in \mathbb{N}$, $m > 0$ ein N_m gibt mit: $l \geq N_m \Rightarrow \mid a - a_l \mid < k \cdot \frac{1}{m}$.

Beweis:

(i) $\Rightarrow$ (ii) und (ii) $\Rightarrow$ (iii) sind klar. Wir zeigen (iii) $\Rightarrow$ (i):

Sei $\varepsilon > 0$, dann gibt es $m \in \mathbb{N}$ mit $\varepsilon > k \cdot \frac{1}{m}$ (Archimedisches Prinzip), also $\mid a - a_l \mid < k \cdot \frac{1}{m} < \varepsilon$.

3.1.9 Cauchy-Folge oder Fundamental-Folge

Definition §3.17

Eine Folge heißt **Cauchy-Folge** oder **Fundamental-Folge** wenn eine der beiden folgenden, gleichwertigen Bedingungen erfüllt ist:

(i) Zu jedem $\varepsilon > 0$ gibt es ein $N_\varepsilon \in \mathbb{N}$ so, dass für alle $k, l \geq N_\varepsilon$ gilt: $|a_k - a_l| < \varepsilon$.

(ii) Zu jedem $\varepsilon > 0$ gibt es ein $N_\varepsilon \in \mathbb{N}$ so, dass für alle $k \geq N_\varepsilon$ gilt: $|a_k - a_{N_\varepsilon}| < \varepsilon$.

Warum sind (i) und (ii) gleichwertig? Klar: (i) $\Rightarrow$ (ii) mit $l = N_\varepsilon$.
Wieso (ii) $\Rightarrow$ (i)? Für $k, l \geq N_\varepsilon$ gilt

$$|a_k - a_l| = |a_k - a_{N_\varepsilon} + a_{N_\varepsilon} - a_l| \leq |a_k - a_{N_\varepsilon}| + |a_l - a_{N_\varepsilon}| < 2\varepsilon$$

Satz §3.18

Satz: Für eine Folge $(a_n)_{n\in\mathbb{N}}$ reeller Zahlen sind äquivalent:

(i) $(a_n)_{n\in\mathbb{N}}$ ist konvergent.

(ii) $(a_n)_{n\in\mathbb{N}}$ ist eine Cauchy-Folge.

Beweis: (i) $\Rightarrow$ (ii) Sei $\lim\limits_{n\to\infty} a_n = a$. Zu $\varepsilon > 0$ wähle $N_\varepsilon \in \mathbb{N}$ mit: Für $k \geq N_\varepsilon$ gilt $\mid a - a_k \mid < \frac{\varepsilon}{2}$. Dann gilt für $k, l \geq N_\varepsilon$: $\mid a_k - a_k \mid = \mid a_k - a + a - a_l \mid \leq \mid a_k - a \mid + \mid a_l - a \mid < \varepsilon$.

Beweis: (ii) $\Rightarrow$ (i)
Wir nehmen an, $(a_n)_{n\in\mathbb{N}}$ ist eine Cauchy-Folge. Wir zeigen zuerst, die Folge ist **beschränkt**: Zu $\varepsilon = 1$ gibt es N_1 mit $\mid a_k - a_{N_1} \mid < 1$ für $k \geq N_1$. Setze $r = \max\{|a_0|, |a_1|, \ldots, |a_{N_1}|\} + 1$. Dann gilt für *alle* k $|a_k| \leq r$.

Nach Bolzano-Weierstrass gibt es einen Häufungspunkt a von $(a_n)_{n\in\mathbb{N}}$. Zeige jetzt $\lim\limits_{n\to\infty} a_n = a$. Sei $\varepsilon > 0$, wähle $N_\varepsilon \in \mathbb{N}$ so, dass $\mid a_k - a_l \mid < \frac{\varepsilon}{2}$ für alle k, $l \geq N_\varepsilon$. Weil a ein Häufungspunkt ist, gibt es $m \in \mathbb{N}$ mit $\mid a_m - a \mid < \varepsilon$, mit $m \geq N_\varepsilon$. Für $k \geq N_\varepsilon$ gilt dann:

$$\mid a - a_k \mid = \mid a - a_m + a_m - a_k \mid \leq \underbrace{\mid a - a_m \mid}_{<\frac{\varepsilon}{2}} + \underbrace{\mid a_m - a_k \mid}_{<\frac{\varepsilon}{2}} < \varepsilon$$

3.2 Reihen

Vorlesung 20.11.2003

Aus einer Folge $(a_n)_{n\in\mathbb{N}}$ konstruieren wir eine neue Folge $(s_n)_{n\in\mathbb{N}}$ durch $s_n = \sum\limits_{k=0}^{n} a_k$. Diese Folgenglieder heißen **Partialsummen**, die Folge $(s_n)_{n\in\mathbb{N}}$ heißt **Reihe**. Schreibe symbolisch: Definition §3.19

$$(s_n)_{n\in\mathbb{N}} = \sum_{n=0}^{\infty} a_n$$

Falls diese Reihe konvergiert, $\lim\limits_{n\to\infty} s_n = s$, schreibe

$$\sum_{n=0}^{\infty} a_n = s$$

Das Symbol $\sum\limits_{n=0}^{\infty} a_n$ hat also zwei verschiedene Bedeutungen!

3.2.1 Die geometrische Reihe

Beispiel §3.20

Die geometrische Reihe: $\sum\limits_{k=0}^{\infty} q^k$

Für	$q = 1$	$\Rightarrow$	$\sum\limits_{k=0}^{\infty} 1^k = \infty$	divergiert die Reihe.
Für	$q \neq 1$	$\Rightarrow$	wissen wir $\sum\limits_{k=0}^{l} q^k = \frac{1-q^{l+1}}{1-q} = s_l$	
Für	$\lvert q\rvert < 1$	$\Rightarrow$	konvergiert das gegen $\frac{1}{1-q}$ dann gilt $\lim\limits_{l\to\infty} \lvert q\rvert^l = 0$	

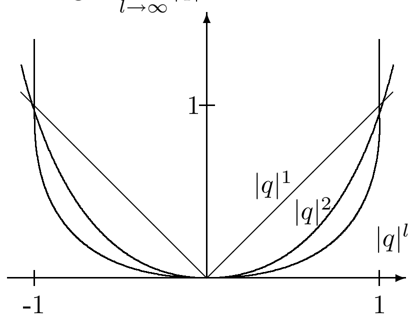


Abbildung 5: Die geometrische Reihe

Warum? Schreibe $\frac{1}{|q|} = 1 + x$ mit $x > 0$.
$(1+x)^l \geq 1 + lx$ (Bernoulli) also $|q|^l \leq \frac{1}{a+lx}$. Für jedes $\varepsilon > 0$ gibt es ein N_ε mit $\frac{1}{1+lx} < \varepsilon$ für $l \geq N_\varepsilon$ (Archimedisches Prinzip), also $|q| < \varepsilon$ für $l \geq N_\varepsilon \Rightarrow \lim\limits_{n\to\infty} |q|^n = 0$.
Für $|q| < 1$ ist also $\sum\limits_{n=0}^{\infty} q^n = \frac{1}{1-q}$.
Für $|q| \geq 1$ ist die geometrische Reihe divergent.

3.2.2 Die harmonische Reihe

Beispiel §3.21

Die harmonische Reihe: $\sum\limits_{k=1}^{\infty} \frac{1}{k}$ ist divergent.

Denn: $1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \ldots \geq 1 + \frac{1}{2} + 2\frac{1}{4} + 4\frac{1}{8} + \ldots$ Das wird beliebig groß!

Die Folge der Partialreihen ist also streng monoton und nicht beschränkt, also divergiert die Reihe.

Moral: Wenn $\sum\limits_{n=0}^{\infty} a_n$ konvergiert, dann *muss* $(a_n)_{n\in\mathbb{N}}$ sicher eine Nullfolge sein, $\lim\limits_{n\to\infty} a_n = 0$. Diese Bedingung alleine garantiert aber noch nicht die Konvergenz, wie wir an der harmonischen Reihe sehen.

4 Konvergenzkriterien für Reihen

4.1 Das Leibnizkriterium

Satz §4.1

Ist $(a_n)_{n\in\mathbb{N}}$ eine streng monoton fallende Nullfolge, so konvergiert die Reihe

$$\sum_{k=0}^{\infty}(-1)^k a_k$$

Beweis: Betrachte

$$\begin{aligned} s_{2l+1} &= \underbrace{(a_0 - a_1)}_{>0} + \underbrace{(a_2 - a_3)}_{>0} + \ldots + \underbrace{(a_{2l} - a_{2l+1})}_{>0} \\ s_{2l} &= a_0 - \underbrace{(a_1 - a_2)}_{>0} - \underbrace{(a_3 - a_4)}_{>0} - \ldots - \underbrace{(a_{2l-1} - a_{2l})}_{>0} \end{aligned}$$

Die Folge $(s_{2l-1})_{l\in\mathbb{N}}$ ist streng monoton wachsend. Die Folge $(s_{2l})_{l\in\mathbb{N}}$ ist streng monoton fallend.

Es gilt weiter: $0 < s_{2l+1} < s_{2l} < a_0$. Also konvergieren beide Folgen:

$$\lim_{l\to\infty} s_{2l+1} = s, \quad \lim_{l\to\infty} s_{2l} = t \quad \Rightarrow s \leq t$$

Da $\lim\limits_{l\to\infty} |s_{2l+1} - s_{2l}| = \lim\limits_{l\to\infty} |a_{2l+1}| = 0$ gilt $s = t = \lim\limits_{l\to\infty} \sum\limits_{k=0}^{\infty} (-1)^k a_k$.

$$|q|^l < \varepsilon \quad \text{für} \quad l \geq N_\varepsilon$$

Beispiel §4.2

Beispiel: Nach dem Leibnizkriterium konvergiert die **alternierende harmonische Reihe**:

$$\sum_{k=1}^{\infty}(-1)^k \frac{1}{k} \qquad (= \ln 2)$$

Anderes Beispiel:

$$\sum_{k=0}^{\infty}(-1)^k \frac{1}{2k+1} \qquad (= \tfrac{\pi}{4})$$

4.2 Absolute Konvergenz

Definition §4.3

Eine Reihe $\sum\limits_{n=0}^{\infty} a_n$ heißt **absolut konvergent** falls $\sum\limits_{n=0}^{\infty} |a_n|$ konvergiert.

Eine absolut konvergente Reihe ist konvergent:

$$\begin{aligned} \left|\sum_{k=0}^{l} a_k - \sum_{k=0}^{m} a_k\right| &= \left|\sum_{k=l+1}^{m} a_k\right| \leq \sum_{k=l+1}^{m} |a_k| \\ &= \left|\sum_{k=0}^{l} |a_k| - \sum_{k=0}^{m} |a_k|\right| \qquad (\text{für } l \leq m) \end{aligned}$$

Das ist eine Cauchy-Folge, falls $\sum\limits_{n=0}^{\infty} |a_n|$ konvergiert.

Absolute Konvergenz $\Rightarrow$ Konvergenz. Die Umkehrung ist *falsch*, denn die alternierende harmonische Reihe konvergiert, die harmonische Reihe aber nicht.

Wir sagen, eine Eigenschaft trifft auf *fast alle* natürlichen Zahlen zu, falls sie nur auf endlich viele natürlichen Zahlen nicht zutrifft.

4.3 Das Majorantenkriterium

Satz §4.4

Ist $\sum\limits_{k=0}^{\infty} c_k$ konvergent, und gilt $|a_n| \leq c_n$ für fast alle $n \in \mathbb{N}$, dann ist $\sum\limits_{n=0}^{\infty} a_n$ absolut konvergent (insbesondere auch konvergent).

Beweis: Es gelte $|a_n| \leq c_n$ für fast alle $n \leq N$.
Damit $\sum\limits_{k=N}^{m} |a_k| \leq \sum\limits_{k=N}^{m} c_k$, die rechte Seite konvergiert, also ist $\sum\limits_{k=0}^{\infty} |a_k|$ beschränkt (und monoton), deswegen konvergent. $\square$

4.4 Das Quotientenkriterium

Satz §4.5

Ist $a_n \neq 0$ für (fast) alle n, und gibt es ein q mit $0 < q < 1$, so, dass $\left|\frac{a_{n+1}}{a_n}\right| \leq q$ für fast alle n, dann konvergiert die Reihe $\sum\limits_{n=0}^{\infty} a_n$ absolut.

Beweis: Wir nehmen an, alles gilt für *alle* $n \in \mathbb{N}$: $\left|\frac{a_n}{a_0}\right| \leq q$, $\left|\frac{a_2}{a_1}\right| \leq q$, usw.,
also $\sum\limits_{n=0}^{l} |a_n| \leq \sum\limits_{n=0}^{l} |a_0| q^n \leq |a_0| \sum\limits_{n=0}^{l} q^n$
Die rechte Seite ist eine geometrische Reihe, also konvergent. $\square$

4.5 Das Wurzelkriterium

Vorlesung 24.11.2003

Satz §4.6

Gibt es ein q mit $0 < q < 1$, so dass $\sqrt[n]{|a_n|} \leq q$ für (fast) alle n, so konvergiert die Reihe $\sum\limits_{n=0}^{\infty} a_n$ absolut.

Beweis: Wegen $|a_n| \leq q^n$ gilt $\sum\limits_{n=0}^{\infty} |a_n| \leq \sum\limits_{n=0}^{\infty} q^n = \frac{1}{1-q}$ wie beim Beweis des Quotientenkriterium.

4.6 Der Verdichtungssatz von Cauchy

Satz §4.7

Ist $a_n > 0$ für alle n, und ist $(a_n)_{n\in\mathbb{N}}$ monoton fallend, so konvergiert die Reihe $\sum\limits_{n=0}^{\infty} a_n$ genau dann, wenn die Reihe $\sum\limits_{n=0}^{\infty} 2^n a_{2^n}$ konvergiert.

Beweis: Setze $s_n = \sum\limits_{k=1}^{n} a_n$ und $t_n = \sum\limits_{k=0}^{n} 2^k a_{2^k}$.

Für $n < 2^{m+1}$ gilt dann

$$\begin{aligned} s_n &= a_1 + a_2 + \ldots + a_n \\ &= a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) \\ &\quad + \ldots + (a_{2^m} + a_{2^m+1} + \ldots + a_{2^{m+1}} - 1) \\ &\leq a_1 + 2a_2 + 4a_4 + \ldots + 2^m a_{2^m} = t_m \end{aligned}$$

Wenn $(t_m)_{m\in\mathbb{N}}$ konvergiert, so ist $(s_n)_{n\in\mathbb{N}}$ monoton wachsend und beschränkt, also konvergent.

Andersrum: Wenn $n \geq 2^{m+1}$, dann gilt entsprechend:

$$\begin{aligned} s_n &\geq a_1 + a_2 + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + \ldots \\ &\geq a_1 + a_2 + 2a_4 + 4a_8 + \ldots \\ &= \frac{1}{2}a_1 + \frac{1}{2}t_m \end{aligned}$$

Wenn also $(s_n)_{n\in\mathbb{N}}$ konvergiert, so auch die Folge $\left(\frac{1}{2}t_m\right)_{m\in\mathbb{N}}$, also auch $(t_m)_{m\in\mathbb{N}}$.
□

Beispiele

Beispiele §4.8

- Für $r > 1$, $r \in \mathbb{R}$ konvergiert die Reihe $\sum\limits_{n=1}^{\infty} \frac{1}{n^r}$.

  Beweis (Verdichtungssatz): Betrachte

  $\sum\limits_{n=0}^{\infty} 2^n \frac{1}{(2^n)^r} = \sum\limits_{n=0}^{\infty} q^n$ mit $q = 2^{1-r} < 1$, das konvergiert (geom. Reihe).

- Die Reihe $\sum\limits_{n=0}^{\infty} n^2 2^{-n}$ konvergiert.

  Beweis (Quotientenkriterium):

  $\left|\frac{(n+1)^2 2^{-(n+1)}}{n^2 2^{-n}}\right| = \frac{1}{2}\left(\frac{n+1}{n}\right)^2 = \frac{1}{2}\left(1+\frac{1}{n}\right)^2 \leq \frac{1}{2}\left(1+\frac{1}{3}\right)^2 = \frac{8}{9} < 1$

  für $n \geq 3$.

4.7 Die Exponentialreihe

Satz §4.9

Für jedes $x \in \mathbb{R}$ konvergiert die Reihe $\sum\limits_{k=0}^{\infty} \frac{1}{k!}x^k = \exp(x)$ absolut.

Beweis (Quotientenkriterium):

$\left|\frac{x^{k+1}}{x^k} \cdot \frac{k!}{(k+1)!}\right| = |x|\frac{1}{k+1} \leq \frac{1}{2} < 1$

falls $2|x| \leq k+1$. Für solche k können wir das Quotientenkriterium anwenden.

4.8 g-adische Entwicklung von Zahlen

Beispiel §4.10

Wir stellen und reelle Zahlen als *Dezimalbrüche* vor, z.B:

$$\sqrt{2} = 1,414213\ldots \quad \text{(10-adische Entwicklung)}$$

Andere g (z.B. $g = 2$ oder $g = 16$) sind als Basen in der Informatik relevant.

Sei $g \in \mathbb{N}$, $g \geq 2$, sei $l \in \mathbb{Z}$ und sei $(a_n)_{n\in\mathbb{N}}$ eine Folge mit $a_n \in \{0, 1, 2, \ldots, g-1\}$ für alle n. Betrachte die Reihe:

$$\sum_{n=0}^{\infty} a_n g^{l-n} = g^l \sum_{n=0}^{\infty} a_n \frac{1}{g^n}$$

Diese Reihe konvergiert absolut, denn

$$\sum_{n=0}^{\infty} a_n g^{l-n} \leq \sum_{n=0}^{\infty} (g-1) g^l g^{-n} = (g-1) g^l \frac{1}{1-\frac{1}{g}}$$

Für $a = \sum\limits_{n=0}^{\infty} a_n g^{l-n}$ schreibe: $a \overset{(g)}{=} a_0 \cdots a_l \,\mathbf{,}\, a_{l+1} \cdots$

bzw. $a \overset{(g)}{=} 0 \,\mathbf{,}\, \underbrace{0 \cdots 0}_{a-l} a_0\, a_1 \cdots$ (falls $l < 0$)

Beispiel ($g = 10$): $\frac{1}{6} = \frac{1}{10} + \frac{6}{100} + \frac{6}{1000} + \frac{6}{10000} + \cdots \overset{(10)}{=} 0,1666\ldots$

Denn: $\frac{1}{6} = \frac{1}{10} + \sum\limits_{k=0}^{\infty} \frac{6}{100} \cdot \frac{1}{10^k} = \frac{1}{10} + \frac{6}{100} \cdot \frac{1}{1-\frac{1}{10}} = \frac{1}{10} + \frac{6}{100} \cdot \frac{10}{9} = \frac{1}{6}$

Satz: Ist $a < 0$ eine reelle Zahl, so gibt es eine Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n \in \{0, \ldots, g-1\}$ und ein $l \in \mathbb{Z}$, mit $a_0 \neq 0$, so dass

$$a = \sum_{n=0}^{\infty} a_n g^{l-n} \qquad \text{Schreibe: } a \overset{(g)}{=} a_0 \cdots a_l \,\mathbf{,}\, a_{l+1} \cdots$$

Beweis: Ein Algorithmus. Für $r \in \mathbb{R}$ sei: Vorlesung 27.11.2003

$$\lfloor r \rfloor = \max\{n \in \mathbb{Z} \mid n \leq r\}$$

1. Setze $l = \min\{k \in \mathbb{Z} \mid g^{k+1} > a\}$
   Dann gilt also $g^{l+1} > a \geq g^l$
2. Setze
$$\begin{aligned} a_0 &= \lfloor a \cdot g^{-l} \rfloor & s_0 &= a_0 g^l \leq a \\ a_1 &= \lfloor (a - s_0) \cdot g^{-l+1} \rfloor & s_1 &= a_0 g^l + a_1 g^{l-1} \\ &\vdots & &\vdots \\ a_l &= \lfloor (a - s_{k-1}) g^{-l+k} \rfloor & s_k &= a_0 g^l + a_1 g^{l-1} + \ldots + a_k q^{l-k} \end{aligned}$$

Für $a = 0$ schreibe $a \overset{(g)}{=} 0$

Für $a < 0$ schreibe $a \overset{(g)}{=} -a_0\, a_1 \ldots$

Diese Darstellung zur Basis g ist nicht immer eindeutig, z.B. wissen wir aus der Schule, dass $1 = 0,\overline{9} = 0,999\ldots$, denn:

$$\begin{aligned} 0,999\ldots &= \frac{9}{10}+\frac{9}{100}+\frac{9}{1000}+\cdots = 9\left(\left(\sum_{k=0}^{\infty}10^{-k}\right)-1\right) \\ &= 9\left(\frac{1}{1-\frac{1}{10}}-1\right) = 9\left(\frac{10}{9}-1\right) = 1 \end{aligned}$$

Die Darstellung wird eindeutig, wenn man verlangt, dass nicht fast alle Ziffern $= g-1$ sein dürfen.

Beispiel: $g = 2$, $a = \frac{1}{3}$
Bestimme l: $\underbrace{2^{-2+1}}_{\frac{1}{2}} > \frac{1}{3} > \underbrace{2^{-2}}_{\frac{1}{4}} \Longrightarrow l = -2$

$$\begin{array}{ll} a_0 = \lfloor \frac{1}{3}\cdot 4\rfloor = 1 & s_0 = \frac{1}{4} \\ a_1 = \lfloor (\frac{1}{3}-\frac{1}{4})\cdot 8\rfloor = 0 & s_1 = \frac{1}{4} \\ a_2 = \lfloor (\frac{1}{3}-\frac{1}{4})\cdot 16\rfloor = 0 & s_2 = \frac{1}{4}+\frac{1}{16} = \frac{5}{16} \\ a_3 = \lfloor (\frac{1}{3}-\frac{5}{16})\cdot 32\rfloor = 0 & s_3 = \frac{5}{16} \\ \vdots & \vdots \end{array}$$

Also: $\frac{1}{3} \stackrel{(2)}{=} 0,01010101\ldots$

Denn: $\sum\limits_{k=0}^{\infty} 4^{-k} = \frac{1}{1-\frac{1}{4}} = \frac{4}{3} = 1 + \frac{1}{3} \Longrightarrow \sum\limits_{k=1}^{\infty} 4^{-k} = \frac{1}{3}$

5 Stetigkeit von Funktionen

5.1 Abbildungen

Definition §5.1

Sind X und Y Mengen, dann ist eine **Abbildung** f von X nach Y eine Vorschrift, die jedem $x \in X$ ein Element $y = f(x) \in Y$ zuordnet. Schreibweise:

$$f\colon X \longrightarrow Y \text{ oder } X \xrightarrow{f} Y \text{ oder } x \mapsto f(x) = y$$

Formal im Rahmen der Mengenlehre: f ist eine Menge von Paaren

$$f = \{(x, y) \mid f(x) = y\}$$

so dass für jedes $x \in X$ genau ein Paar (x, y) mit $y \in Y$ existiert.

Mit anderen Worten: Abbildungen sind **Schaubilder** oder **Graphen**.

Beispiele

- **Identische Abbildung** oder **Identität**:
  $id_x\colon X \longrightarrow Y$ für $X = Y, x \mapsto x$
- **Reelle Funktionen**:
  $f\colon X \longrightarrow Y$ für $X = \mathbb{R} = Y$
- **Absolutbetrag**:
  $f\colon \mathbb{R} \longrightarrow \mathbb{R}, x \mapsto |x|$
- **Polynomfunktion**:
  $p\colon \mathbb{R} \longrightarrow \mathbb{R}, x \mapsto a_0 + a_1x + a_2x^2 + \cdots + a_nx^x$ für $a_0, ..., a_n \in \mathbb{R}$
- **Konstante Abbildung**:
  $f\colon \mathbb{R} \longrightarrow \mathbb{R}, x \mapsto c$ für $c \in \mathbb{R}$ (fest)
- **Charakteristische Abbildung** der Menge $A \subseteq \mathbb{R}$
  $\chi_A\colon \mathbb{R} \longrightarrow \mathbb{R}$ für $x \mapsto \begin{cases} 1 & \text{falls } a \in A \\ 0 & \text{falls } a \notin A \end{cases}$

5.2 Stetigkeit

Definition §5.2

Sei $X \subseteq \mathbb{R}$ eine Menge der reellen Zahlen (z.B. $X = \mathbb{R}$) und $f\colon X \longrightarrow \mathbb{R}$ eine Abbildung. Dann heißt f **stetig**, wenn folgendes gilt: Ist $(x_n)_{n \in \mathbb{N}}$ eine Folge in X (d.h. $x_k \in X$ für alle k), die gegen $x \in X$ konvergiert, so gilt:

$$\lim_{n\to\infty} x_n = x \quad \Rightarrow \quad \lim_{n\to\infty} f(x_n) = f(x)$$

Beispiele

- $x \mapsto x$ ist stetig
  $\lim\limits_{n\to\infty} x_n = x \quad \Rightarrow \quad \lim\limits_{n\to\infty} f(x_n) = x = f(x)$
- $f(x) = \begin{cases} -1 & \text{falls } x < 0 \\ +1 & \text{falls } x \geq 0 \end{cases}$ ist nicht stetig.
  Setze $x_n = -\frac{1}{n}$, also $f(x_n) = -1$
  $\lim\limits_{n\to\infty} -\frac{1}{n} = 0, f(0) = 1$ Ein Widerspruch.

5.3 Lipschitz-stetige Funktionen

Eine Funktion heißt **lipschitz-stetig**, wenn es ein $L \in \mathbb{R}$ gibt mit:

$$|x-y| \cdot L \geq |f(x) - f(y)| \qquad \text{für alle } x, y \text{ im Definitionsbereich}$$

Lemma §5.3

Lemma: Lipschitz-stetige Funktionen sind stetig.

Beweis: Sei $(x_n)_{n\in\mathbb{N}}$ konvergent, $\lim\limits_{n\to\infty} x_n = x$.
Also gilt:

$$|f(x) - f(x_n)| \leq \underbrace{L \cdot |x - x_n|}_{\text{Nullfolge}}$$

also ist $(|f(x) - f(x_n)|)_{n\in\mathbb{N}}$ eine Nullfolge,
mit anderen Worten: $\lim\limits_{n\to\infty} f(x_n) = f(x)$ □

Beispiele

- Die Betragsfunktion $x \mapsto |x|$ ist lipschitz-stetig mit $L = 1$
- Die konstante Abbildung ist lipschitz-stetig mit $L = 0$
- Ebenso ist die Identität lipschitz-stetig mit $L = 1$

5.4 Rationale Operationen auf Funktionen

Vorlesung 01.12.2003

Satz §5.4

Sind f und g stetig, so auch die folgenden Funktionen:

$$\begin{aligned} f+g &= [x \mapsto f(x) + g(x)] \\ f \cdot g &= [x \mapsto f(x) \cdot g(x)] \\ c \cdot f &= [x \mapsto c \cdot f(x)] \text{ für ein konstantes } c \in \mathbb{R} \end{aligned}$$

Insbesondere sind alle Polynomfunktionen stetig:

$$p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

Beweis: Sei $x = \lim\limits_{n\to\infty} x_n$ und $f(x) = \lim\limits_{n\to\infty} f(x_n)$, $g(x) = \lim\limits_{n\to\infty} g(x_n)$.

Nach §3.6 gilt dann

$$f(x) + g(x) = \lim_{n\to\infty} \big(f(x_n) + g(x_n)\big)$$

$$f(x) \cdot g(x) = \lim_{n\to\infty} \big(f(x_n) \cdot g(x_n)\big)$$

5.5 Ring der stetigen Funktionen

Bemerkung §5.5

Für $X \subseteq \mathbb{R}$ sei

$$\mathrm{C}\,(X, \mathbb{R}) = \{f \colon X \longrightarrow \mathbb{R} \mid f \text{ ist stetig }\}$$

die Menge (der Ring) der stetigen Funktionen auf X.

5.6 Intervalle

Definition §5.6

Die folgenden Mengen bezeichnet man für alle $a < b$ als **Intervalle**:

$$
\begin{aligned}
[\,a,b\,] &= \{x \in \mathbb{R} \mid a \leq x \leq b\} \\
[\,a,b\,[&= \{x \in \mathbb{R} \mid a \leq x < b\} \\
]\,a,b\,] &= \{x \in \mathbb{R} \mid a < x \leq b\} \\
]\,a,b\,[&= \{x \in \mathbb{R} \mid a < x < b\}
\end{aligned}
$$

$$
\begin{aligned}
[\,a,\infty\,[&= \{x \in \mathbb{R} \mid a \leq x\} \\
]\,a,\infty\,[&= \{x \in \mathbb{R} \mid a < x\} \\
]\,-\infty,a\,] &= \{x \in \mathbb{R} \mid x \leq a\} \\
]\,-\infty,a\,[&= \{x \in \mathbb{R} \mid x < a\}
\end{aligned}
$$

5.7 Zwischenwertsatz

Satz §5.7

Sei J ein Intervall, sei $f\colon J \longrightarrow \mathbb{R}$ stetig, sei $a, b \in J$ mit $f(a) \leq f(b)$. Zu jedem y mit $f(a) \leq y \leq f(b)$ gibt es ein x zwischen a und b mit $f(x) = y$.

Beweis: Wir betrachten nur den Fall $f(a) < y < f(b)$ und nehmen an, dass $a \leq b$ ist (Der Fall $b < a$ geht analog). Sei $M = \{x \in [a,b] \mid f(x) \leq y\}$.

Da $a \in M$, ist $M \neq \emptyset$ und es existiert $r = \sup(M)$. Da r Supremum ist, gibt es eine Folge $(r_n)_{n\in\mathbb{N}}$ in M mit $\lim\limits_{n\to\infty} r_n = r$.

Dann gilt $\lim\limits_{n\to\infty} f(r_n) = f(r) \leq y$. Da $r < b$ gibt es eine Folge $(s_n)_{n\in\mathbb{N}}$ in $]r,b]$ mit $\lim\limits_{n\to\infty} s_n = r$. (z.B. $s_n = r + \frac{1}{n}$), dann ist $s_n \notin M$, d.h. $f(s_n) > y$, damit $\lim\limits_{n\to\infty} f(s_n) \geq y$. Damit folgt: $f(r) = y$. □

5.8 Umkehrfunktion

Folgerung §5.8

Sei $J = [a,b]$. Wenn $f \in \mathrm{C}(J, \mathbb{R})$ streng monoton wachsend ($s < t \Rightarrow f(s) < f(t)$ für alle $s, t \in J$) ist, dann hat f eine stetige **Umkehrfunktion** g, d.h.:

$$g(f(x)) = x \qquad \text{für alle } x \in J$$

Beweis: Für $s \neq t$ folgt $f(s) \neq f(t)$. Nach dem Zwischenwertsatz gibt es zu jedem y zwischen $f(a)$ und $f(b)$ *genau ein* $x \in J$ mit $f(x) = y$. Das liefert $g(y) = x$.

Warum ist g stetig? Angenommen $\lim\limits_{n\to\infty} y_n = y$. Zu zeigen: $\lim\limits_{n\to\infty} g(y_n) = g(y)$. Setze $g(y) = x$, $g(y_n) = x_n$, d.h. $y = f(x)$, $y_n = f(x_n)$. Zu zeigen: $\lim\limits_{n\to\infty} x_n = x$. Da $(x_n)_{n\in\mathbb{N}}$ beschränkt ist, können wir Bolzano-Weierstraß benutzen. Wenn $\lim\limits_{n\to\infty} x_n \neq x$ wäre, gäbe es eine Teilfolge $(x_{nk})_{k\in\mathbb{N}}$, die gegen $\hat{x} \neq x$ konvergiert. Dann gilt $\lim\limits_{k\to\infty} f(x_{nk}) = \lim\limits_{k\to\infty} g_{nk} = f(\hat{x})$.

Wegen $\lim\limits_{n\to\infty} y_n = y$ gilt aber sicher $\lim\limits_{k\to\infty} y_{nk} = y$, also ist $f(\hat{x}) = y = f(x)$ ein Widerspruch, denn $x \neq \hat{x}$, also $f(x) \neq f(\hat{x})$. □

Bemerkung: Ein entsprechender Satz gilt natürlich für streng monoton fallende stetige Funktionen.

Beispiel: $X = [0, \infty[, \quad n \geq 1, n \in \mathbb{N}. f(x) = x^n$. Das ist stetig (Polynom) und streng monoton. Für $r > 0$ gilt: $x^n + \underbrace{nx^{n-1}r + \cdots + r^n}_{>0} > x^n$

Nach dem Satz gibt es auf jedem Intervall $[a, b]$ eine Umkehrfunktion, also auf ganz $[0, \infty[$. Für $y = x^n$ ist diese Umkehrfunktion $g(y) = \sqrt[n]{y} = y^{1/n}$ stetig.

5.9 Satz von Weierstraß

Vorlesung 04.12.2003
Satz §5.9

Ist $J = [a, b]$ ein abgeschlossenes Intervall und ist $f: J \longrightarrow \mathbb{R}$ stetig, dann ist die Menge $f(J) = \{f(x) \mid x \in J\}$ wieder ein abgeschlossenes Intervall $f(J) = [c, d]$.

Insbesondere nimmt f ein Maximum und ein Minimum an, d.h. es gibt $u \in J$ mit $f(u) = c$ und es gibt $v \in J$ mit $f(v) = d$ aber $f(u) \leq f(x) \leq f(v)$ für alle $x \in J$.

Beweis: Angenommen $f(J)$ ist nicht nach oben beschränkt, d.h. es gibt Folgen $(x_n)_{n\in\mathbb{N}}$ in J mit $\lim\limits_{n\to\infty} f(x_n) = \infty$. Nach Bolzano-Weierstraß (§3.15) gibt es eine Teilfolge $(x_{nk})_{k\in\mathbb{N}}$ und $x = \lim\limits_{n\to\infty} x_{n_k} \in J$. Da f stetig ist, folgt:

$$f(x) = \lim_{k\to\infty} f(x_{n_k}) = \infty$$

Ein Widerspruch.

Sei $d = \sup f(J)$, dann gibt es eine Teilfolge $(x_{nk})_{k\in\mathbb{N}}$ mit

$$\lim_{k\to\infty} f(x_{n_k}) = f(x) = \lim_{n\to\infty} f(x_n) = d$$

Ebenso verfährt man mit $c = \inf f(J)$. Nach dem Zwischenwertsatz (§5.7) ist $f(J) = [c, d]$. Wir betrachten jetzt Folgen von Funktionen. In der Praxis sind das Folgen, die ein Aproximationsproblem besser und besser lösen.

5.10 Punktweise und gleichmäßige Konvergenz

Definition §5.10

Sei $X \subseteq \mathbb{R}$. Ist für jedes $n \in \mathbb{N}$ $f_n: X \longrightarrow \mathbb{R}$ eine Abbildung, so heißt $(f_n)_{n\in\mathbb{N}}$ **Funktionenfolge**.

Wir sagen $(f_n)_{n\in\mathbb{N}}$ **konvergiert punktweise** gegen eine Funktion $f: X \longrightarrow \mathbb{R}$ falls für alle $x \in X$ gilt:

$$\lim_{n\to\infty} f_n(x) = f(x)$$

Wir sagen $(f_n)_{n\in\mathbb{N}}$ **konvergiert gleichmäßig** gegen eine Funktion $f: X \longrightarrow \mathbb{R}$ falls zu jedem $\varepsilon > 0$ ein $N_\varepsilon \in \mathbb{N}$ existiert mit $|f_n(x) - f(x)| < \varepsilon$ für alle $x \in X$, für alle $n > N_\varepsilon$.

Gleichmäßige Konvergenz impliziert offensichtlich punktweise Konvegenz.

Beispiel

Beispiel §5.11

$X = [0,1]$ $\quad f_n(x) = x^n$ $\quad$ für $x = 1$ gilt $f_n(x) = 1^n = 1$, für $0 \leq x < 1$ gilt $\lim\limits_{n\to\infty} x^n = 0$ (vgl. §3.20), Die Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert punktweise gegen die unstetige Funktion:

$$f(x) = \begin{cases} 1 & \text{falls } x = 1 \\ 0 & \text{falls } 0 \leq x < 1 \end{cases}$$

5.11 Stetigkeit und gleichmäßige Konvergenz

Satz §5.12

Konvergiert die Folge $(f_n)_{n\in\mathbb{N}}$ gleichmäßig gegen die Funktion f und sind alle f_n stetig so ist auch f stetig.

Beweis: Sei $x \in X$, $(x_n)_{n\in\mathbb{N}}$ eine Folge in X mit $\lim\limits_{n\to\infty} x_n = x$. Sei $\varepsilon > 0$. Dann gibt es ein $m \in \mathbb{N}$ mit $|f_m(u) - f(u)| < \frac{\varepsilon}{3}$ für alle $u \in X$.

Wähle $N \in \mathbb{N}$ so, dass für $n > N$ gilt: $|f_m(x) - f_m(x_n)| < \frac{\varepsilon}{3}$ (da f_m stetig ist)
Für $n > N$ gilt also:

$$\begin{aligned} |f(x_n) - f(x)| &= |f(x_n) - f_m(x_n) + f_m(x_n) - f_m(x) + f_m(x) - f(x)| \\ &\leq |f(x_n) - f_m(x_n)| + |f_m(x_n) - f_m(x)| + |f_m(x) - f(x)| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} \leq \varepsilon \end{aligned}$$

also $\lim\limits_{n\to\infty} f(x_n) = f(x)$ $\square$

5.12 Grenzwert von rekursiv definierten Folgen

Anwendung §5.13

Ist f stetig, ist $(a_n)_{n\in\mathbb{N}}$ rekursiv definiert durch $a_{n+1} = f(a_n)$ und konvergiert $(a_n)_{n\in\mathbb{N}}$ gegen a, so ist $a = \lim\limits_{n\to\infty} a_{n+1} = \lim\limits_{n\to\infty} f(a_n) = f(\lim\limits_{n\to\infty} a_n) = f(a)$.

Beispiel

$a_{n+1} = \sqrt{6 + a_n}$, $\quad a_0 = 0$ (vgl. §3.11) $\quad (a_n)_{n\in\mathbb{N}}$ konvergiert, also gilt:

$$\begin{aligned} & a = \lim_{n\to\infty} a_n = \sqrt{6+a} > 0 \\ \Leftrightarrow\ & a^2 = 6 + a \\ \Leftrightarrow\ & a^2 - a - 6 = 0 \\ \Leftrightarrow\ & a = \frac{1}{2} \pm \sqrt{\frac{1}{4} + 6} = 3 \quad (\text{und } -2) \end{aligned}$$

Der Grenzwert der Folge a beträgt also 3, da die Folge streng monoton steigend ist mit $a_0 = 0$.

5.13 Die Potenzreihe

Definition §5.14

Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge. Setze $P_k(x) = \sum\limits_{n=0}^{k} a_n x^n$. Das ist eine Folge stetiger Funktionen. Diese Funktionenfolge heißt **Potenzreihe**. Schreibe symbolisch:

$$(P_k)_{k\in\mathbb{N}} = \sum_{n=0}^{\infty} a_n x^n$$

5.14 Konvergenzradius

Satz §5.15

Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge, $L = \limsup \sqrt[n]{|a_n|}$ (bzw. $L = \infty$ falls $\sqrt[n]{|a_n|}$ unbeschränkt).

Setze $R = \begin{cases} \frac{1}{L} & \text{falls } L \neq 0 \\ 0 & \text{falls } L = \infty \\ \infty & \text{falls } L = 0 \end{cases}$

Die Reihe $\sum\limits_{n=0}^{\infty} a_n x^n$ konvergiert absolut für $|x| < R$. Ist $r < R$, so konvergiert die Funktionenfolge $\sum\limits_{n=0}^{\infty} a_n x^n$ gleichmäßig auf $X = [-r, r]$ gegen eine stetige Funktion, die wir ebenfalls mit $\sum\limits_{n=0}^{\infty} a_n x^n$ bezeichnen.

R heißt **Konvergenzradius** der Reihe $\sum\limits_{n=0}^{\infty} a_n x^n$.

Für $|x| > R$ divergiert die Reihe $\sum\limits_{n=0}^{\infty} a_n x^n$. Für $|x| = R$ kann man keine allgemeine Aussage treffen.

Beweis: Wir betrachten den Fall $0 \leq L < \infty$. Dann ist $\sqrt[n]{|a_n x^n|} < |x| \sqrt[n]{|a_n|}$
Für $|x| < \frac{1}{L}$ gilt also: Es gibt ein $q < 1$ mit $\sqrt[n]{|a_n x^n|} < q$ für fast alle n.

Denn $\limsup\limits_{n\to\infty} \sqrt[n]{|a_n x^n|} \leq |x| \limsup\limits_{n\to\infty} \sqrt[n]{|a_n|} = |x| \cdot L < 1$
Nach dem Wurzelkriterium (§4.6) konvergiert die Reihe absolut.

Weiter ist $\sum\limits_{k=0}^{\infty} |a_n| \cdot |r|^n = \sum\limits_{k=0}^{\infty} (\sqrt[n]{|a_n|} r)^n$ eine Majorante für $|x| \leq r$, die gar nicht von x abhängt, daraus folgt die gleichmäßige Konvergenz auf $[-r, r]$. □

Beispiel §5.16

Beispiel: Die Exponentialfunktion $\exp(x) = \sum\limits_{n=0}^{\infty} \frac{1}{n!} x^n$ ist stetig. Die Exponentialreihe konvergiert auf jedem Intervall $[-r, r]$ gleichmäßig gegen $\exp(x)$.

$$L = \sqrt[n]{\frac{1}{n!}} = 0$$

5.15 Potenzreihen mit Entwicklungspunkt

Bemerkung §5.17

Manchmal betrachtet man Reihen der Form:

$$\sum_{n=0}^{\infty} a_n (x - x_o)^n \qquad \text{für festes } x_o \in \mathbb{R}$$

Man spricht von **Potenzreihen mit Entwicklungspunkt** x_0. Die Theorie ist die gleiche wie oben (ersetze $z = x - x_0$)

$$\sum_{k=0}^{\infty} a_n (x - x_0)^n \quad \text{konvergiert absolut auf }]-R + x_0,\ R + x_0[$$

5.16 Das Cauchy-Produkt

Hilfssatz §5.18

Wenn $\sum\limits_{n=0}^{\infty} a_n$ und $\sum\limits_{n=0}^{\infty} b_n$ absolut konvergieren, so konvergiert auch das **Cauchy-Produkt** $\sum\limits_{n=0}^{\infty} c_n$ (absolut) gegen $a \cdot b$ mit:

$$c_n = a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \cdots + a_n b_0 = \sum_{k=0}^{n} a_k b_{n-k}$$

Der Beweis ist einfach, aber länglich (siehe Forster Analysis I §8.3).

$$\begin{aligned} \left(\sum_{n=0}^{\infty} a_n\right)\left(\sum_{n=0}^{\infty} b_n\right) &= (a_0 + a_1 + a_2 + \cdots)(b_0 + b_1 + b_2 + \cdots) \\ &= a_0 b_0 + a_0 b_1 + a_1 b_0 + a_2 b_2 + a_1 b_1 + a_2 b_0 + \cdots \end{aligned}$$

5.17 Exponential- und Logarithmusfunktion

5.17.1 Die Exponentialfunktion

Vorlesung 08.12.2003

Satz §5.19

Es gilt:

(i) $\exp(0) = 1$

(ii) $\exp(-x) = \frac{1}{\exp(x)}$

(iii) $\exp(x + y) = \exp(x)\exp(y)$

Diese Eigenschaften sagen: exp ist ein **Homomorphismus** von der additiven Gruppe der reellen Zahlen in die multiplikative Gruppe der reellen Zahlen.

Beweis: (i) ist klar. Für (iii) betrachte:

$$\left(\sum_{n=0}^{\infty} \frac{1}{n!} x^n\right)\left(\sum_{n=0}^{\infty} \frac{1}{n!} y^n\right) = \sum_{n=0}^{\infty} c_n$$

mit $c_n = \sum\limits_{k=0}^{n} \frac{1}{k!} x^k \frac{1}{(n-k)!} y^{n-k} = \frac{1}{n!} \sum\limits_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = \frac{1}{n!}(x+y)^n$
also $c_n = \frac{1}{n!}(x+y)^n \Rightarrow \exp(x) \cdot \exp(y) = \exp(x+y)$

Damit folgt für (ii): $1 = \exp(0) = \exp(x + (-x)) = \exp(x) \cdot \exp(-x)$.
Also $\exp(-x) = \frac{1}{\exp x}$. □

5.17.2 Die Logarithmusfunktion

Satz §5.20

Die Exponentialfunktion ist streng monoton steigend, für jedes $y > 0$ gibt es genau ein $x \in \mathbb{R}$ mit $y = \exp(x)$. Setze $x = \ln(y)$. Dies ist die **Logarithmusfunktion** ln (natürlicher Logarithmus), die für alle positiven Zahlen definiert ist.

$$\ln :]0, \infty[\longrightarrow \mathbb{R}$$

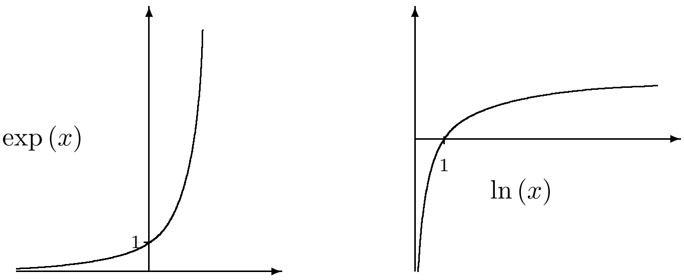


Abbildung 6: exp- und ln-Funktion

5.17.3 Die eulersche Zahl

Man setzt $\exp(1) = 1 + 1 + \frac{1}{2} + \frac{1}{6} + \cdots \approx 2,7$

Beweis: Für $r > 0$ gilt $\exp(r) = 1 + \underbrace{r + \frac{r^2}{2} + \frac{r^3}{6}}_{>0} + \cdots > 1$,

also $\exp(x + r) = \exp(x) \cdot \underbrace{\exp(r)}_{>1} > \exp(x)$,

also ist exp streng monoton steigend. Für $x \geq 0$ ist $\exp(x) > 0$, also gilt $\exp(x) > 0$ für alle $x \in \mathbb{R}$.

5.17.4 Die Exponentialfunktion zur Basis a

Für alle $n \in \mathbb{N}$ gilt: $\exp(n) = 1+n+\frac{n^2}{2}+\frac{n^3}{6}+\cdots \geq 1+n$, d.h. $\lim\limits_{n\to\infty} \exp(n) = \infty$.
Weiter gilt: $\exp(-n) = \exp(n)^{-1} \leq \frac{1}{1+n}$, also $\lim\limits_{n\to\infty} \exp(-n) = 0$.
Nach dem Zwischenwertsatz gibt es zu jedem $y > 0$ ein x mit $y = \exp(x)$.
Man setzt $\mathrm{e} = \exp(1) \approx 2,7$ und schreibt dann formal $\exp(x) = \mathrm{e}^x$.

Für $y, z > 0$ gilt: $\ln(yz) = \ln(y) + \ln(z)$. Für $a > 0$ schreiben wir für $x \in \mathbb{R}$.

$$a^x = \exp(x \cdot \ln a)$$
$$\mathrm{e}^x = \exp(x \cdot \ln \mathrm{e})$$

5.17.5 Der Logarithmus zur Basis a

Für alle $x \in \mathbb{N}$ gilt:

$$\begin{aligned}
x &= \underbrace{1 + 1 + 1 + \cdots + 1}_{x \text{ mal}} \\
a^x &= \exp\big((\underbrace{1 + 1 + 1 + \cdots + 1)}_{x \text{ mal}} \cdot \ln a\big) \\
&= \underbrace{\exp(\ln a) \cdot \exp(\ln a) \cdots \exp(\ln a)}_{x \text{ mal}} \\
&= \underbrace{a \cdot a \cdot a \cdots a}_{x \text{ mal}}
\end{aligned}$$

d.h. wir haben die übliche Potenzfunktion $a \longrightarrow a^n$ auf ganz $\mathbb{R}$ erweitert.

Offensichtlich gilt:

- $a^{(x+y)} = a^x \cdot a^y$
- $a^0 = 1$
- $a^{-x} = \frac{1}{a^x}$

Entsprechend definieren wir

$$\ln_a x = \frac{\ln x}{\ln a} \qquad \text{für } a, x > 0$$

als **Logarithmus zur Basis** a

Beispiel

$$\ln_a a^n = \frac{\ln a^n}{\ln a} = \frac{n \ln a}{\ln a} = n$$
$$\ln_2 1024 = 10$$

5.17.6 Wachstum der Exponentialfunktion

Satz §5.21

Für jedes Polynom $p(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$ gilt $\lim\limits_{k\to\infty} \frac{\exp(k)}{|p(k)|} = \infty$

"Die Exponentialfunktion wächst schneller als jedes Polynom."

Beweis: Für $x > 1$ gilt: $|p(x)| \leq |a_0 + a_1 + a_2 + \cdots + a_n| \cdot |x^n|$. Deshalb genügt es den Fall $p(x) = x^n$ zu betrachten. $\lim\limits_{k\to\infty} \frac{\exp(k)}{k} = \infty$ für n fest.
$\exp(k) \geq \frac{k^{n+1}}{(n+1)!}$ also $\frac{\exp(k)}{k^n} \geq \frac{k^{n+11!}}{(n+1)!}\frac{1}{k^n} = \frac{k}{(n+1)!}$
Ähnlich zeigt man $\lim\limits_{k\to\infty} \frac{\exp(-k)}{k^{-n}} = \lim\limits_{k\to\infty} k^n \cdot \exp(-k) = 0$ sowie $\lim\limits_{k\to\infty} \frac{\ln k}{\sqrt[n]{k}} = 0$

"Die Logarithmusfunktion wächst langsamer als jede Wurzelfunktion."

5.18 Hyperbolische Funktionen

Manchmal betrachtet man:

$$\cosh(x) = \frac{1}{2}(\mathrm{e}^x + \mathrm{e}^{-x}) \qquad \text{Cosinus hyperbolicus}$$

$$\sinh(x) = \frac{1}{2}(\mathrm{e}^x - \mathrm{e}^{-x}) \qquad \text{Sinus hyperbolicus}$$

5.19 Trigonometrische Funktionen

Definition §5.22

$$\cos(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$$
$$\sin(x) = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$$

Mit dem Quotientenkriterium sieht man schnell, dass beide Reihen für alle $x \in \mathbb{R}$ konvergieren.

Wie bei der Exponentialfunktion zeigt man:

- $\cos(0) = 1$
- $\sin(0) = 0$
- $\cos^2 x + \sin^2 x = 1$
- $\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y)$
- $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$

Vorlesung 11.12.2003

Man überlegt sich, dass für $0 < y \leq 3$ gilt:

$$c_2 = 1 - \frac{x^2}{2} < \cos x < 1 - \frac{x^2}{2} + \frac{x^4}{4} = c_4$$

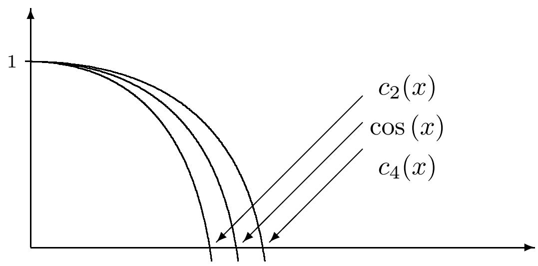


Abbildung 7: Annäherung der cos-Funktion

Weil c_2 und c_4 Nullstellen haben

$$c_2(\sqrt{2}) = 0, \qquad c_4(\sqrt{6-2\sqrt{3}}) = 0$$

hat auch cos eine Nullstelle (Zwischenwertsatz).
Diese Nullstelle liegt im Intervall $\left[\sqrt{2}, \sqrt{6-2\sqrt{3}}\right]$.

Wir definieren die Zahl π wie folgt: $\frac{\pi}{2}$ ist die kleinste positive Nullstelle von cos.
Es gilt $\pi = 3,14159\ldots$
Ähnlich zeigt man: Für $0 < x \leq 3$ ist $\sin x > 0$, insbesondere ist $\sin\frac{\pi}{2} > 0$, folglich gilt: $\sin\frac{\pi}{2} = 1$

Mit den Additionstheoremen folgt nun:

$\sin(x+\frac{\pi}{2}) = \cos x$	$\cos(x+\frac{\pi}{2}) = -\sin x$
$\sin(x+\pi) = -\sin x$	$\cos(x+\pi) = -\cos x$
$\sin(x+2\pi) = \sin x$	$\cos(x+2\pi) = \cos x$
$\cos\pi = -1$	$\sin\pi = 0$

6 Integration

Motivation: Wir wollen Flächeninhalte unter Kurven bestimmen.

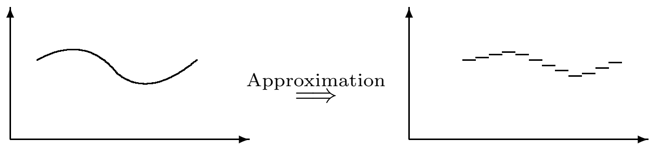


Abbildung 8: Integration

6.1 Beschränkte Funktionen

Definition §6.1

Sei $X \subseteq \mathbb{R}$. Eine Funktion $f\colon X \longrightarrow \mathbb{R}$ heißt **beschränkt**, wenn es eine Zahl $K \in \mathbb{R}$ gibt, so dass $|f(x)| \leq K$ für alle $x \in X$. Die Menge aller beschränkten Funktionen auf X

$$\{f\colon X \longrightarrow \mathbb{R} \mid f \text{ ist beschränkt}\}$$

heißt $\mathrm{B}(X, \mathbb{R})$ (Set of bounded functions).

Beispiel

- $X =]0, \infty[\,, \qquad f(x) = \frac{1}{x}$
  f ist nicht beschränkt.

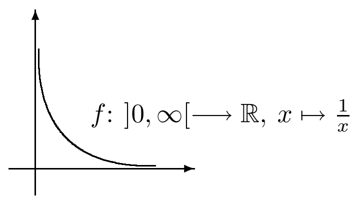


Abbildung 9: Hyperbel

- Dagegen ist z.B. $|sin(x)| \leq 1$ für alle x, also ist $\sin$ beschränkt.
- $X = [1, 5]\,, \qquad f(x) = \frac{1}{x}$
  Für $1 \leq x \leq 5$ gilt sicherlich $|\frac{1}{x}| \leq 1$, d.h. auf dem Intervall $[1, 5]$ ist die Funktion $f(x) = \frac{1}{x}$ beschränkt.

Rationale Operationen auf beschränkte Funktion

Sind $f, g \in \mathrm{B}(X, \mathbb{R})$ und $c \in \mathbb{R}$, so definieren wir:

$$\begin{aligned}(f+g)(x) &= f(x) + g(x) \in \mathrm{B}(X, \mathbb{R}) \\ (cf)(x) &= c \cdot f(x) \in \mathrm{B}(X, \mathbb{R}) \\ (f \cdot g)(x) &= f(x) \cdot g(x) \in \mathrm{B}(X, \mathbb{R})\end{aligned}$$

Beweis: Für alle $x \in X$ gelte: $|f(x)| \leq K$, $|g(x)| \leq L$ für gewisse $K, L \in \mathbb{R}$
Dann gilt also:

$$\begin{aligned} |f(x)+g(x)| &\leq |f(x)|+|g(x)| \leq K+L \\ |cf(x)| &= |c|\cdot|f(x)| \leq |c|\cdot|K| \\ |(fg)(x)| &= |f(x)\cdot g(x)| \leq K\cdot L \end{aligned}$$

Vorlesung 15.12.2003
Lemma §6.2

Lemma: Die Menge $\mathrm{B}(X,\mathbb{R})$ der beschränkten Funktionen auf X ist ein **reeller Vektorraum**, d.h. für alle $f, g, h \in \mathrm{B}(X,\mathbb{R})$ gilt:

- $(f+g)+h = f+g+h = f+(g+h)$
- $f+g = g+f$
- $f+(-1)f = 0$ (dabei steht 0 für die Nullfunktion $x \mapsto 0$)

sowie für alle $c, d \in \mathbb{R}$ und $f, g \in \mathrm{B}(X,\mathbb{R})$ gilt stets:

- $cf+dg \in \mathrm{B}(X,\mathbb{R})$
- $(c+d)f = cf+df$
- $c(f+g) = cf+cg$
- $(cd)f = cdf = c(df)$

$\mathrm{B}(X,\mathbb{R})$ ist ein Ring (zum Körper fehlt *nur* das multiplikative Inverse), denn $f(x)\cdot g(x) = 1$ hat keine Lösung falls f eine Nullstelle hat.

Bemerkung §6.3

Bemerkung: Ist $X = [a,b]$ ein abgeschlossenes Intervall und $f\colon X \longrightarrow \mathbb{R}$ stetig, so gilt durch den Satz von Weierstraß, dass f beschränkt ist, also:

$$\mathrm{C}([a,b],\mathbb{R}) \subseteq \mathrm{B}([a,b],\mathbb{R})$$

Eine beschränkte Funktion ist nicht notwendig stetig.
$f\colon [0,2] \longrightarrow \mathbb{R}$, $f(x) = 0$ für $0 \leq x \leq 1$, $f(x) = 1$ für $1 \leq x \leq 2$.

6.2 Supremumsnorm

Satz §6.4

Für $f \in \mathrm{B}(X,\mathbb{R})$ definieren wir die **Supremumsnorm**:

$$||f|| = \sup\{\,|f(x)| \;\big|\; x \in X\} \quad \Rightarrow \quad ||f|| \geq 0$$

Dann gilt:

(i) $||f|| = 0 \Leftrightarrow f(x) = 0$ für alle $x \in X$

(ii) $||c\cdot f|| = |c|\cdot||f||$ für alle $c \in \mathbb{R}$ und $f \in \mathrm{B}(X,\mathbb{R})$

(iii) $||f+g|| \leq ||f||+||g||$ für alle $f, g \in \mathrm{B}(X,\mathbb{R})$

Beweis: (i) ist klar.
Zu (ii) und (iii): Es gilt: $|c\cdot f(x)| = |c|\cdot|f(x)|$ also $||c\cdot f|| = |c|\cdot||f||$.
Wegen $|f(x)+g(x)| \leq |f(x)|+|g(x)| \leq ||f||+||g||$ für alle $x \in X$,
also $||f+g|| = \sup\{|f(x)+g(x)| \mid x \in X\} \leq ||f||+||g||$. □

6.3 Cauchy-Funktionenfolge

Definition §6.5

Eine Funktionenfolge $(f_n)_{n\in\mathbb{N}}$ in $\mathrm{B}(X,\mathbb{R})$ heißt **Cauchy-Folge**, falls zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert mit dem für alle $l, m \geq N$ gilt: $||f_l - f_m|| < \varepsilon$

Eine Folge $(f_n)_{n\in\mathbb{N}}$ konvergiert gegen f in $\mathrm{B}(X,\mathbb{R})$ falls $\lim\limits_{n\to\infty} ||f_n - f|| = 0$, die Bedingung sagt genau, dass $(f_n)_{n\in\mathbb{N}}$ gleichmäßig gegen f konvergiert (vgl. §5.10). f_n konvergiert also gegen f.

Für alle $\varepsilon > 0$ existiert $N \in \mathbb{N}$, so dass für alle $n \geq N$ und $x \in X$ gilt: $|f_n(x) - f(x)| < \varepsilon$, also gilt

$$\sup\{|f_n(x) - f(x)| \mid x \in X\} < \varepsilon \qquad \text{also} \qquad ||f_n - f|| < \varepsilon$$

Satz §6.6

Satz: Ist $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathrm{B}(X,\mathbb{R})$, dann gibt es $f \in \mathrm{B}(X,\mathbb{R})$ mit $\lim\limits_{n\to\infty} ||f_n - f|| = 0$. Schreibe:

$$f = \lim_{n\to\infty} f_n$$

Beweis: Da $|f_l(x) - f_m(x)| \leq ||f_l - f_m||$ für alle $x \in X$ gilt, ist die Folge $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$.

Setze $f(x) = \lim\limits_{n\to\infty} f_n(x)$ für alle $x \in X$. $(f_n)_{n\in\mathbb{N}}$ konvergiert gleichmäßig gegen f. Sei $\varepsilon > 0$, dann existiert $N \in \mathbb{N}$, so dass für alle $x \in X$ und $l, m \geq N$ gilt: $|f_l(x) - f_m(x)| < \varepsilon$, $f(x) = \lim\limits_{l\to\infty} f_l(x)$ also gilt: $|f(x) - f_n(x)| \leq \varepsilon$ für alle $x \in X$ also gilt: $||f - f_n|| < \varepsilon$ für alle $n \geq N$, also gilt:

$$\lim_{n\to\infty} ||f - f_n|| = 0$$

Warum gilt $f \in \mathrm{B}(X,\mathbb{R})$?
Zu $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$, so dass für alle $m \geq N$ und $x \in X$ gilt:

$$|f(x) - f_m(x)| < \varepsilon$$

Wähle $m > N$ fest. f_m ist beschränkt, d.h. es gibt ein $K \in \mathbb{R}$ mit $|f_m(x)| \leq K$ für alle $x \in X$.
$|f(x)| = |f(x) - f_m(x) + f_m(x)| \leq |f(x) - f_m(x)| + |f_m(x)| \leq \varepsilon + K$ $\square$

6.4 Banachraum

Bemerkung §6.7

Ein reeller Vektorraum heißt **normiert**, wenn es eine Abbildung $||\cdot||\colon V \longrightarrow \mathbb{R}$ gibt mit:

- $||x|| \geq 0$ und $||x|| = 0 \Leftrightarrow x = 0$ für alle $x \in V$
- $||cx|| = |c| \cdot ||x||$ für alle $c \in \mathbb{R}$, $x \in V$
- $||x + y|| \leq ||x|| + ||y||$ für alle $x, y \in V$

Ein normierter Raum heißt **Banachraum**, wenn jede Cauchy-Folge konvergiert. Nach §6.6 ist $\mathrm{B}(X,\mathbb{R})$ ein *Banachraum.*

Bemerkung: Wenn $X = [a, b]$ ein abgeschlossenes Intervall ist, so gilt: Bemerkung §6.8

$$\mathrm{C}(X, \mathbb{R}) \subseteq \mathrm{B}(X, \mathbb{R}) \quad \text{(vgl. §6.3)}$$

Nach §6.6 gilt: Wenn $(f_n)_{n \in \mathbb{N}}$ eine Cauchy-Folge in $\mathrm{C}(X, \mathbb{R})$ ist, so ist $f = \lim\limits_{n \to \infty} f_n$ stetig nach §5.12.

Mit anderen Worten: Ist $X = [a, b]$ ein abgeschlossenes Intervall, dann ist $\mathrm{C}(X, \mathbb{R})$ mit der Supremumsnorm $|| - ||$ ein Banachraum.

Für den Rest dieses Kapitels nehmen wir an, dass $X = [a, b]$ ein abgeschlossenes Intervall ist.

6.5 Zerlegungen

Definition §6.9

Eine Menge $\{a = s_0 < s_1 < \cdots < s_k = b\}$ heißt **Zerlegung** von $X = [a, b]$ der Länge k.

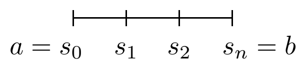


Abbildung 10: Zerlegung

Eine Zerlegung $Z_1 = \{a = s_0 < s_1 < \cdots < s_k = b\}$ wird von einer Zerlegung $Z_2 = \{a = t_0 < t_1 < \cdots < t_l = b\}$ verfeinert, wenn gilt: $Z_1 \subseteq Z_2$

Sind Z_1, Z_2 Zerlegungen, so offensichtlich auch $Z_3 = Z_1 \cup Z_2$ und es gilt:

$$Z_1 \subseteq Z_3 \text{ und } Z_2 \subseteq Z_3$$

6.6 Stufenfunktionen

Ist $Z = \{s_0 < s_1 < \cdots < s_k\}$ eine Zerlegung von X und $c_1 \ldots c_k, d_1 \ldots d_k \in \mathbb{R}$, so heißt die Abbildung

$$f\colon X \longrightarrow \mathbb{R} \text{ mit } f(x) = \begin{cases} c_k & \text{falls } x = s_k \\ d_k & \text{falls } s_k < x < s_{k+1} \end{cases}$$

Stufenfunktion/Treppenfunktion bezüglich Z. Stufenfunktionen sind offensichtlich beschränkt.

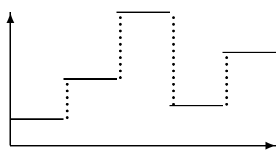


Abbildung 11: Stufenfunktion

Ist f eine Stufenfunktion bzgl. einer Zerlegung Z_1 und ist Z_2 eine Zerlegung von X und $Z_1 \subseteq Z_2$, so ist f auch eine Stufenfunktion bzgl. Z_2.

Die Menge Step $(X, \mathbb{R})$ aller Stufenfunktionen ist aber ein Untervektorraum vom Raum der beschränkten Funktionen $\mathrm{B}(X, \mathbb{R})$, damit ist f Stufenfunktion bzgl. Z_1, g Stufenfunktion bzgl. Z_2, also sind f und g Stufenfunktionen bzgl. $Z = Z_1 \cup Z_2$, also ist auch $f+g$ eine Stufenfunktion bzgl. Z und $c \cdot f$ Stufenfunktion bzgl. Z_1.

6.7 Das Integral

Das Integral einer Stufenfunktion $f \in \text{Step}(X, \mathbb{R})$ bzgl. einer Zerlegung $Z = \{s_0 < s_1 < s_2 < \cdots < s_k\}$ ist definiert als: Definition §6.10

$$\int_a^b f(x)\,\mathrm{dx} = \sum_{l=1}^{k} (s_l - s_{l-1})\, d_l = \sum_{l=1}^{k} (s_l - s_{l-1}) \cdot f\left(\frac{s_l + s_{l-1}}{2}\right)$$

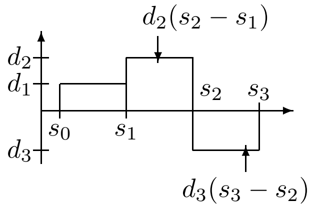


Abbildung 12: Stufenfunktion

Achtung: Flächen unterhalb der x-Achse werden negativ gezählt.

Lemma: Ist f Stufenfunktion bzgl. Z_1 und Z_2 eine Verfeinerung von Z_1 mit $Z_1 = \{s_0 < s_1 < \cdots < s_l\}$, $Z_2 = \{t_0 < t_1 < \cdots < t_m\}$, $Z_1 \subseteq Z_2$ so gilt: Vorlesung 18.12.2003 Lemma §6.11

$$\sum_{k=1}^{l} (s_k - s_{k-1}) \cdot f\left(\frac{s_k + s_{k-1}}{2}\right) = \sum_{k=1}^{m} (t_k - t_{k-1}) \cdot f\left(\frac{t_k + t_{k-1}}{2}\right)$$

Das heißt der Ausdruck $\int_a^b f(x)\,\mathrm{dx}$ hängt *nicht* von der gewählten Zerlegung ab.

Beweis: Induktion nach $m - l$
Für $m = l$ ist $Z_1 = Z_2$.

Für $m = l + 1$ gibt es ein $n \in \mathbb{N}$ mit:
$s_0 = t_0,\ s_1 = t_1,\ \ldots,\ s_n = t_n,\ t_{n+1} < s_{n+1} = t_{n+2}, s_{n+2} = t_{n+3},\ \ldots,\ s_l = t_{l+1}$
$s_0 = t_0,\ s_1 = t_1,\ \ldots,\ s_n = t_n,\ s_{n+1} = t_{n+1}$
und $s_{n+1} - s_n = (t_{n+1} - t_n) + (t_{n+2} - t_{n+1})$
$f\left(\frac{s_{n+1}+s_n}{2}\right) = f\left(\frac{t_{n+1}+t_n}{2}\right) = f\left(\frac{t_{n+2}+t_{n+1}}{2}\right)$

$$\begin{aligned} & \sum_{k=1}^{l} (s_k - s_{k-1}) \cdot f\left(\frac{s_k + s_{k-1}}{2}\right) \\ = \ & \sum_{k=1}^{n} (s_k - s_{k-1}) \cdot f\left(\frac{s_k + s_{k-1}}{2}\right) + (s_{n+1} - s_n) \cdot f\left(\frac{s_{n+1} + s_n}{2}\right) \end{aligned}$$

$$
\begin{aligned}
& + \sum_{k=n+2}^{l} (s_k - s_{k-1}) f\left(\frac{s_k + s_{k-1}}{2}\right) \\
= \quad & \sum_{k=1}^{n} (t_k - t_{k-1}) \cdot f\left(\frac{t_k + t_{k-1}}{2}\right) \\
& + (t_{n+1} - t_n) \cdot f\left(\frac{t_{n+1} + t_n}{2}\right) + (t_{n+2} - t_{n+1}) \cdot f\left(\frac{t_{n+2} + t_{n+1}}{2}\right)
\end{aligned}
$$

Da Z_2 und Z_1 durch sukzessives Hinzufügen von Punkten entsteht, folgt die Behauptung. □

6.8 Der Integrationsoperator

Bemerkung §6.12

Für $f, g \in \mathrm{Step}\,(X, \mathbb{R})$ und $c \in \mathbb{R}$ gilt:

$$\int_a^b (f+g)(x)\ \mathrm{dx} = \int_a^b f(x)\ \mathrm{dx} + \int_a^b g(x)\ \mathrm{dx}$$

$$\int_a^b (cf)(x)\ \mathrm{dx} = c \int_a^b f(x)\ \mathrm{dx}$$

Mit anderen Worten: Der **Integrationsoperator**

$$\mathrm{Step}\,(X, \mathbb{R}) \longrightarrow \mathbb{R}, \quad f \mapsto \int_a^b f(x) dx$$

ist eine lineare Abbildung (man sagt auch ein **lineares Funktional**).

6.9 Regelfunktionen

Definition §6.13

Eine beschränkte Funktion $f \in \mathrm{B}\,(X, \mathbb{R})$ heißt **Regelfunktion**, falls es eine Folge $(f_n)_{n \in \mathbb{N}}$ von Stufenfunktionen gibt, die gleichmäßig gegen f konvergiert. Die Menge aller Regelfunktionen ist ein Untervektorraum von $\mathrm{B}\,(X, \mathbb{R})$.

$$\mathrm{Step}\,(X, \mathbb{R}) \subseteq \mathrm{Reg}\,(X, \mathbb{R}) \subseteq \mathrm{B}\,(X, \mathbb{R})$$

6.10 Gleichmäßige Stetigkeit

Satz §6.14

Wenn $X = [a, b]$ ein abgeschlossenes Intervall ist und $f \colon X \longrightarrow \mathbb{R}$ stetig ist, dann gibt es ein $\varepsilon > 0$ und $\delta > 0$, so dass für alle $u, v \in X$ gilt:

$$|u - v| < \delta \Rightarrow |f(u) - f(v)| < \varepsilon$$

Man sagt auch f ist **gleichmäßig stetig**.

Beweis: Wir führen einen Widerspruchsbeweis und nehmen an, die Behauptung sei falsch. Es gibt ein $\varepsilon > 0$, so dass für alle $\delta > 0$ ein $u, v \in X$ existiert mit:

$$|u - v| < \delta \text{ und } |f(u) - f(v)| \geq \varepsilon$$

Setze $\delta = \frac{1}{n} > 0$. Es gibt $u_n, v_n \in X$ mit $|u_n - v_n| < \frac{1}{n}$ und $|f(u_n) - f(v_n)| \geq \varepsilon$.

Weil X ein abgeschlossenes Intervall ist, finden wir eine konvergente Teilfolge $(u_{nk})_{k\in\mathbb{N}}$ und eine konvergente Teilfolge $(u_{nk})_{k\in\mathbb{N}}$ und $\left(v_{n_k l}\right)_{l\in\mathbb{N}}$ (Bolzano-Weierstraß §3.15).
Schreibe: $u'_l = u_{n_{k_l}}$, $v'_n = v_{n_{k_l}}$, $u = \lim\limits_{l\to\infty} u'_l$, $v = \lim\limits_{l\to\infty} v'_l$, $u = v$ da $\lim\limits_{l\to\infty} u'_l - v'_l = 0$

$$\begin{aligned}
\varepsilon &\leq |f(u'_l) - f(v'_l)| \\
&= |f(u'_l) - f(u) + f(u) - f(v'_l)| \\
&= |f(u'_l) - f(u) + f(v) - f(v'_l)| \\
&\leq |f(u'_l) - f(u)| + |f(v) - f(v'_l)|
\end{aligned}$$

Da f stetig ist und der rechte Ausdruck eine Nullfolge, ergibt sich ein Widerspruch. □

6.11 Stetige Regelfunktionen

Satz §6.15

Auf einem abgeschlossenen Intervall $X = [a, b]$ ist jede stetige Funktion eine Regelfunktion:

$$\mathrm{C}\,(X, \mathbb{R}) \subseteq \mathrm{Reg}\,(X, \mathbb{R})$$

Beweis: Sei $f\colon X \longrightarrow \mathbb{R}$ stetig. Zu jedem $\varepsilon > 0$ finden wir $\delta > 0$, so dass gilt:

$$|f(u) - f(v)| < \varepsilon \;\text{ falls }\; |u - v| < \delta$$

Wähle $m \in \mathbb{N}$ mit $\frac{b-a}{m} < \delta$. Setze $t_0 = a, t_1 = t_0 + \frac{b-a}{m}, t_2 = t_1 + \frac{b-a}{m}, \ldots, t_n = b$, also $t_k = a + k\frac{b-a}{m}$ für $k \in 0, \ldots, n$.
Dann ist $a = t_0 < t_1 < t_2 < \cdots < t_n = b$ eine Zerlegung von X. Setze $c_k = f(t_k)$. Definiere $g \in \mathrm{Step}\,(X, \mathbb{R})$ durch:

$$g(x) = \begin{cases}
c_0 & \text{falls } t_0 \leq x < t_1 \\
c_1 & \text{falls } t_1 \leq x < t_2 \\
\vdots & \vdots \\
c_{n-1} & \text{falls } t_{n-1} \leq x < t_m \\
c_n & \text{falls } t_m = x
\end{cases}$$

Für $t_{k-1} \leq t_k$ gilt dann:

$$0 = t_{k-1} - t_{k-1} \leq x - t_{k-1} \leq t_k - t_{k-1} < \delta$$

also $|f(x) - f(t_{k-1})| < \varepsilon$. $f(t_{k-1}) = g(x)$, also auch $|f(x) - g(x)| < \varepsilon$ für alle $x \in X$, also gilt $||f - g|| \leq \varepsilon$.

Wähle also $g_n \in \mathrm{Step}\,(X, \mathbb{R})$ mit $||f - g_n|| \leq \frac{1}{n}$, so ist $(g_n)_{n\in\mathbb{N}}$ eine Folge in $\mathrm{Step}\,(X, \mathbb{R})$, die gleichmäßig gegen f konvergiert, also $f \in \mathrm{Reg}\,(X, \mathbb{R})$.

6.12 Das Riemann-Integral

Definition §6.16

Für eine Regelfunktion $f \in \mathrm{Reg}\,(X, \mathbb{R})$ definieren wir das **Riemann-Integral** wie folgt:

$$\int_a^b f(x)\;\mathrm{dx} = \lim_{n\to\infty} \int_a^b g_n(x)\;\mathrm{dx}$$

Wobei $(g_n)_{n\in\mathbb{N}}$ eine beliebige Folge von Stufenfunktionen ist, die gleichmäßig gegen f konvergiert.
Natürlich müssen wir überlegen, ob diese Definition überhaupt Sinn macht. Konvergiert die rechte Seite, ist es egal, welche Folge man wählt. Für eine Stufenfunktion g gilt offensichtlich:

$$\left|\int_a^b g(x)\ \mathrm{dx}\right| \leq \int_a^b |g(x)|\ \mathrm{dx} = ||g|| \cdot (b-a)$$

Da die Folge $(g_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge bildet, ist auch die Folge $\int_a^b g_n(x)\ \mathrm{dx}$ eine Cauchy-Folge.

$$\left|\int_a^b (g_n(x) - g_m(x))\ \mathrm{dx}\right| \leq |g_n - g_m| \cdot (b-a)$$

Das heißt die rechte Seite konvergiert jedenfalls.

Angenommen $(h_n)_{n\in\mathbb{N}}$ ist eine weitere Folge von Stufenfunktionen, die gleichmäßig gegen f konvergiert, dann ist

$$||h_n - g_n|| = ||h_n - f + f - g_n|| \leq ||h_n - f|| + ||f - g_n||$$

$$\left|\int_a^b (h_n - g_n)(x)\ \mathrm{dx}\right| \leq ||h_n - g_n||(b-a) \leq (||h_n - f|| + ||f - g_n||) \cdot (b-a)$$

Also gilt: $\lim\limits_{n\to\infty} \int_a^b (h_n - g_n)(x)\ \mathrm{dx} = 0 = \lim\limits_{n\to\infty} \int_a^b h_n(x)\ \mathrm{dx} - \int_a^b g_n(x)\ \mathrm{dx}$
also folgt: $\lim\limits_{n\to\infty} \int_a^b g_n(x)\ \mathrm{dx} = \lim\limits_{n\to\infty} \int_a^b h_n(x)\ \mathrm{dx}$

6.13 Eigenschaften des Integrationsoperators

Vorlesung 05.01.2004
Satz §6.17

Sind f und g Regelfunktionen und $c \in \mathbb{R}$, so gilt:

- $\int_a^b (f+g)(x)\ \mathrm{dx} = \int_a^b f(x)\ \mathrm{dx} + \int_a^b g(x)\ \mathrm{dx}$
- $\int_a^b (cf)(x)\ \mathrm{dx} = c\int_a^b f(x)\ \mathrm{dx}$
- $\left|\int_a^b f(x)\ \mathrm{dx}\right| \leq \int_a^b |f(x)|\ \mathrm{dx} \leq ||f||(b-a)$

Denn diese Aussagen stimmen für Stufenfunktionen, also nach den Rechenregeln für konvergente Zahlenfolgen auch für Regelfunktionen. Mit anderen Worten: Der Integrationsoperator

$$\mathrm{Reg}\,(X, \mathbb{R}) \longrightarrow \mathbb{R}, \quad f \mapsto \int_a^b f(x)\ \mathrm{dx}$$

ist eine **lineare Abbildung**.

$$\left|\int_a^b f(x)\ \mathrm{dx}\right| \leq \int_a^b |f(x)|\ \mathrm{dx} \leq ||f||\ (b-a)$$

Beispiel

Beispiel §6.18

- Für die konstante Funktion $f(x) = C$ gilt: $\int_a^b f(x)\,\mathrm{dx} = C(b-a)$, denn die konstanten Funktionen sind Stufenfunktionen.
- Für die Identität $f(x) = x$ gilt: Setze $t_k = a + k\frac{b-a}{n}$
  $\sum_{k=0}^{n-1} f(t_k)\frac{b-a}{n} = \sum_{k=0}^{n-1}(a + k\frac{b-a}{n})\frac{b-a}{n} = n \cdot a\frac{b-a}{n} + \left(\frac{b-a}{n}\right)^2 \sum_{k=0}^{n-1} k$
  $= ab - a^2 + \left(\frac{b-a}{n}\right)^2 \frac{n(n-1)}{2} = ab - a^2 + (b^2 + 2ab + a^2)\frac{1}{2}(1 - \frac{1}{n})$
  Im Grenzwert $n \to \infty$ erhalten wir also:
  $ab - a^2 + \frac{1}{2}(b^2 - 2ab + a^2) = \frac{1}{2}(b^2 - a^2)$
  Das war mühsam, wir brauchen bessere Methoden! Aber: Die Methode ist gut für *numerische Zwecke*.

6.14 Numerische Berechnung von Integralen

Bemerkung §6.19

Sei $f\colon [a,b] \longrightarrow \mathbb{R}$ stetig. Für $N \in \mathbb{N}$ setze $t_{k,N} = a + k\frac{b-a}{N}$ und betrachte:

$$s_N = \frac{b-a}{N}\sum_{k=0}^{N-1} f(t_k)$$

Für $N \to \infty$ konvergiert s_N gegen $\int_a^b f(x)\,\mathrm{dx}$, also ist s_N für große N eine gute Approximation für das Integral.

6.15 Funktionsklassen

Bemerkung §6.20

Die Regelfunktionen sind eine echte Teilmenge der beschränkten Funktionen:

$$\mathrm{Reg}\,(X,\mathbb{R}) \subset \mathrm{B}\,(X,\mathbb{R})$$

Beispiel

Sei $f \in \mathrm{B}\,(X,\mathbb{R})$ mit $X = [a,b]$ und

$$f(x) = \begin{cases} 1 & \text{falls } x \in \mathbb{Q} \\ 0 & \text{falls } x \in \mathbb{R}\backslash\mathbb{Q} \end{cases}$$

Das ist die Dirichlet'sche Sprungfunktion. Sie ist *keine* Regelfunktion und *nicht Riemann-integrierbar.*

Bemerkung §6.21

In der Klasse der beschränkten Funktionen haben wir folgende Hierarchie:
Diese sechs Klassen von Funktionen bilden jeweils einen **reellen Vektorraum**. Das Integral ist eine **lineare Abbildung** in die reellen Zahlen auf die unteren fünf Funktionsklassen.

$$
\begin{tikzcd}[column sep=small]
 & \mathrm{B}\,(X,\mathbb{R}) \arrow[d, dash] & \\
 & \mathrm{Reg}\,(X,\mathbb{R}) \arrow[dl, dash] \arrow[dr, dash] & \\
\mathrm{Step}\,(X,\mathbb{R}) \arrow[dr, dash] & & \mathrm{C}\,(X,\mathbb{R}) \arrow[dl, dash] \\
 & \{\text{Konstante Funktionen}\} \arrow[d, dash] & \\
 & \{\text{Nullfunktion}\} &
\end{tikzcd}
$$

Abbildung 13: Funktionsklassen

6.16 Eigenschaften von Regelfunktionen

Bemerkung §6.22

Sind f, g Regelfunktionen, so ist auch $f \cdot g\colon x \mapsto f(x) \cdot g(x)$ wieder eine Regelfunktion.
Es gilt aber im Allgemeinen *nicht*, dass:

$$\int_a^b f(x)g(x)\ \mathrm{dx} = \int_a^b f(x)\ \mathrm{dx} \int_a^b g(x)\ \mathrm{dx}$$

6.17 Monotonie des Integrals

Satz §6.23

Sind $f, g \in \mathrm{Reg}\,(X,\mathbb{R})$ und gilt $f(x) \leq g(x)$ für alle $x \in X$, so schreiben wir:

$$f \leq g$$

Satz: Es gilt $\int_a^b f(x)\ \mathrm{dx} \leq \int_a^b g(x)\ \mathrm{dx}$ falls $f \leq g$.

Beweis: Falls f, g Stufenfunktionen sind, ist das klar.
Seien jetzt $(f_n)_{n\in\mathbb{N}}$ und $(g_n)_{n\in\mathbb{N}}$ Folgen von Stufenfunktionen, die gegen f bzw. g konvergieren und gelte $f \leq g$.

Setze: $\tilde{f}_n(x) = f_n(x) - ||f_n - f||$
$\tilde{g}_n(x) = g_n(x) + ||g_n - g||$

Dann sind $\left(\tilde{f}_n\right)_{n\in\mathbb{N}}$ und $(\tilde{g}_n)_{n\in\mathbb{N}}$ Stufenfunktionen, die immer noch gegen f bzw. g konvergieren. Es gilt $\tilde{f}_n(x) \leq f(x) \leq g(x) \leq \tilde{g}_n(x)$ für alle x und alle n, also $\int_a^b \tilde{f}_n(x)\ \mathrm{dx} \leq \int_a^b \tilde{g}_n(x)\ \mathrm{dx}$, also $\int_a^b f(x)\ \mathrm{dx} \leq \int_a^b g(x)\ \mathrm{dx}$ □

6.18 Mittelwertsatz der Integralrechnung

Vorlesung 08.01.2004
Satz §6.24

Seien $p \in \mathrm{Reg}\,(X,\mathbb{R})\,,\ f \in \mathrm{C}\,(X,\mathbb{R})$ mit $p(x) \geq 0$ für alle $x \in X$, dann gibt es ein $t \in [a, b]$ mit:

$$\int_a^b f(x)p(x)\ \mathrm{dx} = f(t) \int_a^b p(x)\ \mathrm{dx}$$

Wichtiger Spezialfall: $p(x) = 1$ für alle $x \in X$

$$\int_a^b f(x)\ \mathrm{dx} = (b-a) \cdot f(t)$$

Beweis: Sei $M = \max\{f(x) \mid x \in X\}$, $m = \min\{f(x) \mid x \in X\}$

Diese Extrema existieren nach dem Satz von Weierstraß (§5.9).

Es gibt also $m \cdot p(x) \leq f(x)\, p(x) \leq M \cdot p(x)$ für alle $x \in X$. Wegen der Monotonie des Integrales folgt:

$$m \int_a^b p(x)\, \mathrm{dx} \leq \int_a^b f(x)\, p(x)\, \mathrm{dx} \leq M \int_a^b p(x)\, \mathrm{dx}$$

Also gibt es eine Zahl q mit

$$q \int_a^b f(x)\, \mathrm{dx} = \int_a^b f(x)\, p(x)\, \mathrm{dx} \text{ mit } m \leq q \leq M$$

Nach dem Zwischenwertsatz (§5.7) gibt es ein $t \in [a, b]$ mit $f(t) = q$. $\square$

Integrale berechnen können wir im Moment noch nicht, dazu brauchen wir Ableitungen. Mit einem verfeinertem Integralbegriff lassen sich *mehr* Funktionen integrieren $\longrightarrow$ *Lebesgue-Integral, Maßtheorie.*

7 Differentiation

Problemstellung: Sei $f\colon X \longrightarrow \mathbb{R}$ eine stetige Funktion. Sei $t_0 \in X$. Wir wollen f nahe bei t_0 durch eine Gerade, das heißt eine **affin-lineare** Funktion approximieren.

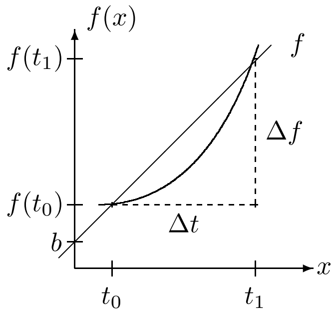


Abbildung 14: Differentiation

Geradengleichung:	$l(t) = m \cdot t + b$
Steigung der Geraden:	$m = \frac{\Delta f}{\Delta t} = \frac{f(t_1)-f(t_o)}{t_1-t_0}$

Weiter wollen wir: $l(t_0) = f(t_0)$. Nach Umformen erhalten wir:

$$l(t) = \frac{f(t_1) - f(t_0)}{t_1 - t_0} \cdot (t - t_0) + f(t_0)$$

dann gilt: $l(t_0) = f(t_0)$.
Um das Problem am Anfang zu lösen, betrachten wir den Fall $\Delta t \to 0$. Im Folgenden wird diese Idee präzisiert.

7.1 Einschränkungen und Fortsetzungen

Definition §7.1

Sei $Y \subseteq X \subseteq \mathbb{R}$ und sei $f\colon X \longrightarrow \mathbb{R}$ eine Funktion. Dann liefert f eine Funktion auf Y,

$$f\mid_Y\colon Y \longrightarrow \mathbb{R}, \quad y \mapsto f(y)$$

die **Einschränkung** von f auf die Teilmenge $Y \subseteq X$.
Sei umgekehrt $g\colon Y \longrightarrow \mathbb{R}$ eine Funktion. Wir sagen $f\colon X \longrightarrow \mathbb{R}$ setzt g fort (**Fortsetzung von** g), falls $f\mid_Y = g$.

Im Allgemeinen hat solch ein g (unendlich) viele Fortsetzungen.

7.2 Stetige Fortsetzungen

Definition §7.2

Jetzt nehmen wir an, dass g *stetig* ist und suchen **stetige Fortsetzungen**.

Ist $(y_n)_{n\in\mathbb{N}}$ eine Folge in Y, die gegen $x \in X$ konvergiert, und ist f eine stetige Fortsetzung von g, so gilt natürlich $\lim\limits_{n\to\infty} g(y_n) = f(x)$.

Die linke Seite hängt nur von y auf g ab, sie legt $f(x)$ eindeutig fest! Wir schreiben in dieser Situation:

$$\lim_{y\to x} g(y) = f(x)$$

Beispiel

Beispiel §7.3

$X = [-1, 1]$, $Y = X \setminus \{0\}$

- $g(y) = y^2$ für $y \in Y$, stetige Fortsetzung $f(x) = x^2$

$$\lim_{y \to 0} y^2 = 0$$

- $g(y) = \begin{cases} 1 & \text{falls } \; 0 < y \leq 1 \\ -1 & \text{falls } -1 \leq y < 0 \end{cases}$

  für $y \in Y$. Auf Y ist g eine stetige Abbildung!

  Aber g hat *keine* stetige Fortsetzung auf X denn:

$$\lim_{n \to \infty} g\left(\frac{1}{n}\right) = \lim_{n \to \infty} 1 = 1$$

  aber

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

  das heißt der Grenzwert $\lim_{y \to 0} g(y)$ existiert *nicht*.

7.3 Differenzierbarkeit

Definition §7.4

Es sei $X =]a, b[$ ein offenes Intervall, $f\colon X \longrightarrow \mathbb{R}$ stetig, $t_0 \in X$.
Wir sagen, f ist in t_0 **differenzierbar** mit Ableitung $f'(t_0) = c$, falls die Funktion $p(h) = \frac{f(t_0+h)-f(t_0)}{h}$ (Differenzenquotient) eine stetige Fortsetzung in 0 hat:

$$\lim_{h \to 0} p(h) = c = f'(t_0)$$

Andere Schreibweisen für die Ableitung:

$$f'(t_0) = \dot{f}(t_0) = \left.\frac{df}{dt}\right|_{t=t_0} = \left.\frac{d}{dt} f(t)\right|_{t=t_0}$$

Beispiel

Vorlesung 12.01.2004

- $f(t) = c$ für alle $t \in X$

  $p(h) = \frac{f(t_0+h)-f(t_0)}{h} = \frac{c-c}{h} = 0$

  Es gibt eine stetige Fortsetzung $p(0) = 0 = f'(t_0)$ für alle $t_0 \in X$

- $f(t) = mt + b$ für alle $m, b \in \mathbb{R}$

  $p(h) = \frac{f(t_0+h)-f(t_0)}{h} = \frac{m(t_0+h)+b-(mt_0+b)}{h} = \frac{mh}{h} = m$

  Es gibt eine stetige Fortsetzung $p(0) = m = f'(t_0)$ für alle $t_0 \in X$

- $f(t) = t^2$ (Parabelgleichung) für alle $t \in X$

  $p(h) = \frac{f(t_0+h)-f(t_0)}{h} = \frac{(t_0+h)^2-t_0^2}{h} = \frac{t_0^2+2t_0h+h^2-t_0{}^2}{h} = 2t_0 + h$

  Es gibt eine stetige Fortsetzung $p(0) = 2t_0 = f'(t_0)$ für alle $t_0 \in X$

- $f(t) = \frac{1}{t}$ für $t \neq 0$

  $p(h) = \frac{f(t_0+h)-f(t_0)}{h} = \frac{\frac{1}{t_0+h} - \frac{1}{t_0}}{h} = \frac{t_0-(t_0+h)}{t_0(t_0+h)h} = \frac{-1}{t_0(t_0-h)}$

  Es gibt eine stetige Fortsetzung $p(0) = \frac{-1}{{t_0}^2} = f'(t_0)$ für alle $t_0 \in X$

- $f(t) = |t|$

  $p(h) = \frac{f(t_0+h)-f(t_0)}{h} = \frac{|t_0+h|-|t_0|}{h}$

  1. Fall: $t_0 > 0$ für $h > -t_0$ ist $t_0 + h > 0$
     also $p(h) = \frac{t_0+h-t_0}{h} = 1$, also $p(0) = 1 = f'(t_0)$
  2. Fall: $t_0 < 0$ für $h < t_0$ ist $t_0 + h < 0$
     also $p(h) = \frac{-(t_0+h)+t_0}{h} = -1$, also $p(0) = -1 = f'(t_0)$
  3. Fall: $t_0 = 0$
     also $p(h) = \frac{|h|}{h} = \begin{cases} 1 & \text{falls } h > 0 \\ -1 & \text{falls } h < 0 \end{cases}$

  Es gibt *keine* stetige Fortsetzung in $h = 0$, also ist die Betragsfunktion in der Stelle $t_0 = 0$ nicht differenzierbar.

Bemerkung: Es gibt Funktionen, die in *keiner* Stelle differenzierbar sind. Bemerkung §7.6
Setze $\phi(h) = p(h) - p(0)$, dann gilt:

$$f(t_0 + h) = f(t_0) + h \cdot p(0) + h \cdot \phi(0)$$

Dann sieht man also: f ist in t_0 differenzierbar genau dann, wenn es eine Zahl $c \in \mathbb{R}$ und eine stetige Funktion ϕ mit $\phi(0) = 0$ gibt, so dass

$$f(t_0 + h) = f(t_0) + c \cdot h + h \cdot \phi(h)$$

Dann gilt $f'(t_0) = c$

7.4 Rechenregeln für das Differenzieren

Seien f, g in t_0 differenzierbar, dann gilt:

(i) Die Summe $f + g\colon t \mapsto f(t) + g(t)$ ist in t_0 differenzierbar:
$(f + g)(t_0) = f'(t_0) + g'(t_0)$

(ii) Für $c \in \mathbb{R}$ ist $cf\colon t \mapsto cf(t)$ in t_0 differenzierbar:
$(cf)'(t_0) = cf'(t_0)$

(iii) **Leibnizregel:** Das Produkt $f \cdot g\colon t \mapsto f(t) \cdot g(t)$ ist in t_0 differenzierbar:
$(fg)(t_0) = f'(t_0)g(t_0) + f(t_0)g'(t_0)$

(iv) Ist $g(t) \neq 0$ für alle $t \in X$, so ist $\left(\frac{f}{g}\right)\colon t \mapsto \frac{f(t)}{g(t)}$ in t_0 differenzierbar:
$\left(\frac{f}{g}\right)'(t_0) = \frac{f'(t_0)g(t_0)-f(t_0)g'(t_0)}{g(t)^2}$

Spezialfall: $\left(\frac{1}{g}\right)'(t_0) = -\frac{g'(t_0)}{g(t_0)^2}$

Beweis

Setze: $p(h) = \frac{f(t_0+h)-f(t_0)}{h} \qquad p(t_0) = f'(t_0)$

$q(h) = \frac{g(t_0+h)-g(t_0)}{h} \qquad q(t_0) = g'(t_0)$

(i) $\frac{(f+g)(t_0+h)-(f+g)(t_0)}{h} = \frac{f(t_0+h)-f(t_0)+g(t_0+h)-g(t_0)}{h} = p(h) + q(h)$

$(f+g)'(t_0) = p(0) + q(0)$

(ii) $\frac{cf(t_0+h)-cf(t_0)}{h} = cp(h)$

(iii) $\frac{f(t_0+h)g(t_0+h)-f(t_0)g(t_0)}{h} = \frac{f(t_0+h)-f(t_0)}{h}g(t_0) + f(t_0)\frac{g(t_0+h)+g(t_0)}{h}$

$= p(h) \cdot g(t_0) + f(t_0+h) \cdot q(h)$

Für $n \to 0$ erhält man $p(0) \cdot g(t_0) + f(t_0) \cdot q(0)$

(iv) Für (iv) genügt es den Spezialfall $\left(\frac{1}{g}\right)'(t_0)$ zu betrachten, weil $\frac{f}{g} = f \cdot \frac{1}{g}$

$\frac{\frac{1}{g(t_0+h)} - \frac{1}{g(t_0)}}{h} = \frac{g(t_0)-g(t_0+h)}{g(t_0)g(t_0+h)h} = -q(h)\frac{1}{g(t_0)g(t_0+h)}$

Für $n \to 0$ erhält man $q(0)\frac{-1}{g(t_0)^2}$

7.5 Differenzierbare Funktionen

Definition §7.7

Ist eine Funktion f an *jeder* Stelle $t_0 \in X$ differenzierbar, so heißt f **differenzierbar**, die Funktion $f' \colon t \mapsto f'(t)$ heißt **Ableitung** von f.
Falls f zusätzlich stetig ist, heißt f **stetig differenzierbar** oder C^1-Funktion.

$$\mathrm{C}^1(X, \mathbb{R}) = \{f \colon X \longrightarrow \mathbb{R} \mid f \text{ ist stetig differenzierbar}\}$$

Nach den Rechenregeln ist $\mathrm{C}^1(X, \mathbb{R})$ ein *reeller Vektorraum*.

Beispiel

Beispiel §7.8

- $f(t) = t^n \qquad n \in \mathbb{N}$

  Vollständige Induktion über n. Behauptung: $f(t) = t^n \Rightarrow f'(t) = nt^{n-1}$

  I.A.: $n = 0 \quad f(t) = 1 \quad f'(t) = 0$
  $n = 1 \quad f(t) = t \quad f'(t) = 1$

  I.S.: $t^{n+1} = t^n \cdot t \overset{\text{Ableitung}}{\longrightarrow} n \cdot t^{n-1}t + t^n \cdot 1 = n \cdot t^n + t^n = (n+1)t^n$

  □

- $f(t) = \frac{1}{t^n}$ Rechenregel (iv) mit $g(t) = t^n$

  $-\frac{g'(t)}{g(t)^2} = \frac{-n \cdot t^{n-1}}{t^{2n}} = -n\frac{1}{t^{n+1}}$

  Fazit: Für $n \in \mathbb{Z}$ und $f(t) = t^n$ gilt also $f'(t) = n \cdot t^{n-1}$

- $f(t) = \cos(t) \qquad f'(t) = -\sin(t)$
- $f(t) = \sin(t) \qquad f'(t) = \cos(t)$
- $f(t) = \exp(t) \qquad f'(t) = \exp(t)$

7.6 Die Kettenregel

Vorlesung 15.01.2004 Satz §7.9

Satz: Seien $g : X \longrightarrow \mathbb{R}$ und $f : Y \longrightarrow \mathbb{R}$ Funktionen mit $g(x) \subseteq Y$ und sei $s_0 = g(t_0)$ für $t_0 \in X$. Wenn g in t_0 differenzierbar ist, und f in s_0 differenzierbar ist, dann ist die Komposition (Hintereinanderausführung) $f \circ g : X \longrightarrow Y \longrightarrow \mathbb{R}$ in t_0 differenzierbar mit Ableitung:

$$(f \circ g)'(t_0) = f'(s_0)g'(t_0) = f'\left(g(t_0)\right)g'(t_0)$$

Achtung: Unterscheide *Komposition* $f \circ g$ vom *Produkt* $f \cdot g$:

$$(f \circ g)(t) = f\left(g(t)\right) \quad \not\equiv \quad (f \cdot g)(t) = f(t) \cdot g(t)$$

Beweis: Betrachte $p(h) \cdot h = f(s_0 + h) - f(s_0)$:

$\frac{f(g(t_0+h))-f(g(t_0))}{h} = \frac{f(g(t_0+h))-f(s_0)}{h} = \frac{g(t_0+h)-g(t_0)}{h} \cdot p\left(g(t_0+h) - g(t_0)\right)$

Im $\lim h \to 0$ erhält man $g'(t_0) \cdot f'(s_0)$ □

7.7 Extrema

Erinnerung §7.10

Definition: Eine Funktion $f : X \longrightarrow \mathbb{R}$ hat ein **Maximum** (ein **Minimum**) in $t_0 \in X$, falls $f(t) \leq f(t_0)$ (falls $f(t) \geq f(t_0)$) für alle $t \in X$ gilt.

Liegt ein Minimum oder ein Maximum vor, spricht man von eimem **Extremum**.

Satz: Ist f differenzierbar in t_0 und hat ein Extremum in t_0, dann ist

$$f'(t_0) = 0$$

Beweis: Wir betrachten den Fall, dass f in t_0 ein Maximum hat. Dann ist $f(t_0 + h) \leq f(t_0)$. Also gilt:

$$p(h) = \frac{f(t_0 + h) - f(t_0)}{h} \quad \begin{cases} \leq 0 & \text{falls } h > 0 \\ \geq 0 & \text{falls } h < 0 \end{cases}$$

In $h = 0$ gilt (weil p stetig ist) $p(0) = 0$.

Anwendung §7.11

Anwendung: Sei $X = [a, b]$, $f : X \longrightarrow \mathbb{R}$ sei stetig und differenzierbar in allen $t \in]a, b[$. Um Extrema von f zu finden, betrachtet man die Randpunkte a und b, sowie alle $t \in]a, b[$ mit $f'(t) = 0$. An einer dieser Stellen liegen die Extrema.

Beispiel: Wie kann man die Oberfläche eine Blechdose (mit Volumen = 1) minimieren?

$$\begin{array}{rrcl} \text{Radius} & r & = & r \\ \text{Höhe} & h & = & \dfrac{1}{\pi r^2} \quad \text{wegen } V = 1 \\ \text{Volumen} & V & = & \pi r^2 \cdot h \\ \text{Oberfläche} & A & = & 2\pi r^2 + h \cdot 2\pi r \\ & & = & 2\pi r^2 + \dfrac{2\pi r}{\pi r^2} = 2\pi r^2 + \dfrac{2}{r} \end{array}$$

Ein Intervall $X = [a, b]$ ist physikalisch sinnvoll. Suche nun Extremum $A' = 0$:

$$A' = 4\pi r - \frac{2}{r^2} \quad \Leftrightarrow \quad 4\pi r = \frac{2}{r^2} \quad \Leftrightarrow \quad r^3 = \frac{1}{2\pi} \quad \Leftrightarrow \quad r = \sqrt[3]{\frac{1}{2\pi}} \approx 0,572\ldots$$

Für a sehr klein und b sehr groß sind $A(a)$ und $A(b)$ ebenfalls sehr groß, in $r = \frac{1}{\sqrt[3]{2\pi}}$ liegt ein Minimum vor.

Beispiel §7.12

Beispiel: Jede Polynomfunktion $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ ist stetig differenzierbar mit Ableitung $p'(x) = na_n x^{n-1} + \cdots + 2a_2 x + a_1$, das bleibt auch für Potenzreihen richtig (§5.14).

$$\begin{aligned} f(x) &= \sum_{k=0}^{\infty} a_k x^k \\ f'(x) &= \sum_{k=0}^{\infty} (k+1) a_{k+1} x^k \quad \text{für alle } |x| < R \end{aligned}$$

Wenn $\sum\limits_{k=0}^{\infty} a_k x^k$ für diese x konvergiert (Beweis vielleicht später). Damit kann man zum Beispiel die Ableitung der exp-Funktion berechnen:

$$\begin{aligned} \exp\, x &= \sum_{k=0}^{\infty} \frac{1}{k!} x^k \\ \exp' x &= \sum_{k=0}^{\infty} (k+1) \frac{1}{(k+1)!} x^k = \exp x \end{aligned}$$

7.8 Der Satz von Rolle

Lemma §7.13

Sei $f\colon [a, b] \longrightarrow \mathbb{R}$ stetig, $f(a) = f(b)$, sei f in allen $t \in]a, b[$ differenzierbar. Dann gibt es ein $t_0 \in]a, b[$ mit $f'(t_0) = 0$.

Beweis:

1. Fall: f ist *konstant*, klar.
2. Fall: f ist nicht konstant, es gibt also ein $t \in]a, b[$ mit $f(t) \neq f(a)$.

   Wir nehmen an, $f(t) > f(a)$ (Der andere Fall $f(t) < f(a)$ geht analog). Da f stetig ist, hat f ein Maximum in Punkt $t_0 \in [a, b]$, also $f(t_0) \geq f(t) > f(a) = f(b)$ also $t_0 \in]a, b[$.

   Also gilt $f'(t_0) = 0$. □

7.9 Mittelwertsatz der Differentialrechnung

Satz §7.14

Ist $f\colon [a, b] \longrightarrow \mathbb{R}$ stetig und differenzierbar in allen $t \in]a, b[$, so gibt es ein $t_0 \in]a, b[$ mit:

$$f'(t_0) = \frac{f(b) - f(a)}{b - a}$$

Beweis: Betrachte $g(t) = f(t) - m(t-a)$, mit $m = \frac{f(b)-f(a)}{b-a}$.
Es gilt $g(a) = f(a)$, $g(b) = f(b) - m(b-a) = f(a)$.

Nach dem Satz von Rolle gilt: Es gibt ein $t_0 \in]a,b[$ mit $g'(t_0) = 0$.
Aber $g'(t) = f'(t) - m$, also $f'(t_0) = m$. □

Folgerungen: Sei $f\colon [a,b] \longrightarrow \mathbb{R}$ differenzierbar

(a) gilt $f'(t) \geq 0$ für alle $t \in]a,b[$, so ist f monoton steigend.
(b) gilt $f'(t) \leq 0$ für alle $t \in]a,b[$, so ist f monoton fallend.
(c) ist $f'(t) = 0$ für alle $t \in]a,b[$, so ist f konstant.
(d) ist $|f'(t)| \leq k$ für alle $t \in]a,b[$, so ist f lipschitz-stetig:
$|f(u) - f(v)| \leq k|u-v|$

Beweise:

(a) Sei $a < u < v < b$, damit $\frac{f(v)-f(u)}{v-u} = f'(t_0) \geq 0$ also $f(v) \geq f(u)$.
(b) Geht analog.
(c) Folgt aus (a) und (b).
(d) $\frac{|f(v)-f(u)|}{|v-u|} = |f'(t_0)| \leq k$ □

7.10 Mehrfach stetig differenzierbare Funktionen

Vorlesung 19.01.2004
Definition §7.15

Sei $X =]a,b[$ und $f\colon X \longrightarrow \mathbb{R}$. Falls f stetig differenzierbar ist und falls $f'\colon X \longrightarrow \mathbb{R}$ stetig differenzierbar ist, so heißt f **zweimal stetig differenzierbar**. f'' ist die Ableitung von f'. Analog definiert man k-mal stetig differenzierbare Funktionen. Setze:

$$\mathrm{C}^k(X,\mathbb{R}) = \{f\colon X \longrightarrow \mathbb{R} \mid f \text{ ist } k\text{-mal stetig differenzierbar}\}$$

Offensichtlich ist $\mathrm{C}^k(X,\mathbb{R})$ ein *reeller Vektorraum*.
Die Funktionen in $\mathrm{C}^\infty(X,\mathbb{R}) = \bigcap\limits_{k\geq 1} \mathrm{C}^k(X,\mathbb{R})$ heißen **glatte Funktionen**.

Wir haben also eine Hierarchie von reellen Vektorräumen:

$$\begin{array}{c} \mathrm{C}(X,\mathbb{R}) \\ | \\ \mathrm{C}^1(X,\mathbb{R}) \\ | \\ \mathrm{C}^2(X,\mathbb{R}) \\ \vdots \\ \mathrm{C}^k(X,\mathbb{R}) \\ | \\ \mathrm{C}^\infty(X,\mathbb{R}) \end{array}$$

Abbildung 15: Funktionsklassen: Stetig differenzierbare Funktionen

Die Rechenregeln für Ableitungen sagen, dass der Differentialoperator $\frac{d}{dx}\colon f \mapsto f'$, der jeder Funktion f die Ableitung f' zuordnet, eine *lineare Abbildung* ist:

$$\mathrm{C}^k(X,\mathbb{R}) \xrightarrow{\frac{d}{dx}} \mathrm{C}^{k-1}(X,\mathbb{R})$$

8 Die Hauptsätze der Differential- und Integralrechnung

8.1 Schreibweisen in der Integralrechnung

Definition §8.1

Sei $X = [a, b]$ und $f\colon X \longrightarrow \mathbb{R}$ eine Regelfunktion. Für $a \leq u \leq v \leq b$ setze:

$$f(x)\Big|_u^v = f(v) - f(u)$$

Die Einschränkung von f auf $[u, v]$ ist eine Regelfunktion, deren Integral wir schreiben als $\int_u^v f(x)\,\mathrm{dx}$. Setze außerdem $f(x)\big|_v^u = -f(x)\big|_u^v$.

$$\int_v^u f(x)\,\mathrm{dx} = -\int_u^v f(x)\,\mathrm{dx}$$

Für festes $t_0 \in [a, b]$ betrachte die reelle Funktion $F\colon X \longrightarrow \mathbb{R}$ mit

$$F(t) = \int_{t_o}^t f(x)\,\mathrm{dx}$$

8.2 I. Hauptsatz der Differential- und Integralrechnung

Satz §8.2

Sei $f\colon [a, b] \longrightarrow \mathbb{R}$ stetig und $t_0 \in [a, b]$ sowie $F(t) = \int_{t_0}^t f(x)\,\mathrm{dx}$. Dann ist F in jedem Punkt $t \in [a, b]$ stetig differenzierbar und

$$F' = f$$

Beweis

$\frac{1}{h}\big(F(t+h) - F(t)\big) = \frac{1}{h}\left(\int_{t_0}^{t+h} f(x)\,\mathrm{dx} - \int_{t_0}^{t} f(x)\,\mathrm{dx}\right) = \frac{1}{h}\int_t^{t+h} f(x)\,\mathrm{dx}$
$= \frac{1}{h} h \cdot f(r)$ für ein $r \in [t, t+h]$ und für $h > 0$ (Der Fall $h < 0$ geht analog).

Der letzte Schritt folgt mit dem Mittelwertsatz der Integralrechnung (§6.24). Im Limes $h \to 0$ konvergiert das gegen $f(t)$, weil f stetig ist. Also ist $F'(t) = f(t)$ □

8.3 II. Hauptsatz der Differential- und Integralrechnung

Satz §8.3

Sei $f\colon]a, b[\longrightarrow \mathbb{R}$ stetig differenzierbar. Für $a < u \leq v < b$ gilt:

$$\int_u^v F'(x)\,\mathrm{dx} = F(x)\Big|_u^v = F(v) - F(u)$$

Beweis

Setze $g(t) = \int_a^b F'(x)\,\mathrm{dx}$. Nach dem I. Hauptsatz gilt jedenfalls $g'(t) = F'(t)$. Also gilt:

$$g'(t) - F'(t) = 0 = (g - F)'(t)$$

Nach dem Mittelwertsatz der Differentialrechnung (§7.14c) ist $g - F$ konstant. Also gilt: $g(t) = F(t) - F(u)$, denn $g(u) = 0$. □

8.4 Stammfunktionen

Bemerkung §8.4

Ist F stetig differenzierbar mit Ableitung $F' = f$, so heißt F **Stammfunktion** zu f. Integrale lösen bedeutet also eigentlich das Finden von solchen Stammfunktionen.

Beispiel: $f(x) = x^n$ hat die Stammfunktion $F(x) = \frac{x^{n+1}}{n+1} + c$. Die Konstante c ist egal, wir setzen $c = 0$. Mit dem II. Hauptsatz ist also:

$$\int_u^v x^n \, \mathrm{dx} = \left(\frac{x^{n+1}}{n+1} + c \right) \Bigg|_u^v = \frac{1}{n+1}(v^{n+1} - u^{n+1})$$

Satz §8.5

Satz: Es sei $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen auf $X =]a, b[$. Wenn $(f_n)_{n\in\mathbb{N}}$ punktweise gegen $f: X \longrightarrow \mathbb{R}$ konvergiert, jedes f_n stetig differenzierbar ist und wenn die Folge $({f'}_n)_{n\in\mathbb{N}}$ gleichmäßig gegen $g: X \longrightarrow \mathbb{R}$ konvergiert, so ist f stetig differenzierbar und $f' = g$.

Erinnerung: Eine Folge von Funktionen $(h_n)_{n\in\mathbb{N}}$ konvergiert *punktweise* gegen eine Funktion h, falls $\lim\limits_{h\to\infty} = h(x)$ für alle $x \in X$.
Die Konvergenz ist gleichmäßig, falls $\lim\limits_{n\to\infty} ||h_n - h|| = 0$, dabei ist
$||h_n - h|| = \sup \{ |h_n(x) - h(x)| \bigm| x \in X \}$.
Gleichmäßige Konvergenz impliziert punktweise Konvergenz.

Beweis (von Satz 5): Für jedes $t \in X$ gilt:

$$\int_{t_0}^t g(x) \, \mathrm{dx} = \lim_{h\to\infty} \int_{t_0}^t f_n'(x) \, \mathrm{dx} = \lim_{h\to\infty} \big(f_n(t) - f_n(t_0)\big) = f(t) - f(t_0)$$

Also gilt: $f' = g$. □
Warum gilt das erste '=' im Beweis?

Satz §8.6

Satz: Ist $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen, die gleichmäßig gegen f konvergiert, so ist

$$\lim_{n\to\infty} \int_u^v f_n(x) \, \mathrm{dx} = \int_u^v f(x) \, \mathrm{dx}$$

Beweis: $\left| \int_u^v f(x) - f_n(x) \, \mathrm{dx} \right| \leq \int_u^v |f(x) - f_n(x)| \, \mathrm{dx} \leq |u - v| \cdot ||f - f_n||$ □

Folgerung

Folgerung §8.7

Sei $f(t) = \sum\limits_{h=0}^{\infty} a_n t^n$ eine Potenzreihe, die für alle $|t| < R$ konvergiert. Dann ist $f'(t) = \sum\limits_{h=0}^{\infty} a_{n+1}(n+1)t^n$ und $\int_0^t f(t) \, \mathrm{dt} = \sum\limits_{h=1}^{\infty} \frac{1}{n} a_{n-1} t^n$ für $|t| < R$.
Denn das ist genau die Situation von Satz 5 mit $f_n(t) = \sum\limits_{k=0}^{n} a_k t^k$

8.5 Tabelle mit Ableitungen

Vorlesung 22.01.2004
Paragraph §8.8

$f(x)$	$\sin x$	$\cos x$	x^n	$\exp x$	$\sinh x$	$\cosh x$	$\ln x$
$f'(x)$	$\cos x$	$-\sin x$	$n \cdot x^{n-1}$	$\exp x$	$\cosh x$	$\sinh x$	$\frac{1}{x}$

Liest man die Tabelle anders herum, erhält man Stammfunktionen, damit lassen sich Integrale berechnen, z.B.:

$$\int_u^v \cos x \text{ dx} \;=\; \sin x\Big|_u^v \;=\; \sin v - \sin u$$

Es folgen weitere Methoden zum Lösen von Integralen:

8.6 Partielle Integration

Paragraph §8.9

Es gilt:

$$\int_a^b f'(x)g(x) \text{ dx} = f(x)g(x)\Big|_a^b - \int_a^b f(x)g'(x) \text{ dx}$$

falls f und g stetig differenzierbar sind, denn $(f \cdot g)' = f' \cdot g + f \cdot g'$ (Leibniz-Regel).

Beispiel: $\int_a^b \ln x \text{ dx} = \int_a^b \ln x \cdot 1 \text{ dx}$ mit $g(x) = \ln x$, $f'(x) = 1$ und $f(x) = x$:

$$\int_a^b \ln x \text{ dx} = \int_a^b 1 \cdot \ln x \text{ dx} = x \cdot \ln x\Big|_a^b - \int_a^b x \cdot \frac{1}{x} \text{ dx} = b \cdot \ln b - a \cdot \ln a + (b-a)$$

8.7 Integration durch Substitution

Paragraph §8.10

Das ist das Integral der Kettenregel:

$$\int_u^v f(g(x)) \cdot g'(x) \text{ dx} = \int_a^b f(s) \text{ ds} \qquad \begin{aligned} g(v) &= b \\ g(u) &= a \end{aligned}$$

Beispiel: Mit Substitution $x = \ln t$:

$$\int_a^b \frac{2 \cdot \exp x + 1}{\exp x + 1} \text{ dx} = \int_c^d \frac{2t+1}{t+1} \cdot \frac{1}{t} \text{ dt} \qquad \begin{aligned} g\,(t) &= \ln t \\ g'(t) &= \tfrac{1}{t} \end{aligned}$$

Es gibt viele Tafeln (z.B. Bronstein, Handbuch der Mathematik) mit Integralen und viele weitere Tricks zum Berechnen von Integralen. Moderne Computer-Algebra wie MAPLE oder MATHEMATICA lösen viele Integrale, allerdings manchmal mit Fehlern!

8.8 Uneigentliche Integrale

Paragraph §8.11

Bisher können wir nur Regelfunktionen auf Intervallen $X = [a, b]$ integrieren. Machmal will man einen allgemeinen Definitionsbereich.

Fall 1: Sei $f\colon [a, \infty[\longrightarrow \mathbb{R}$ eine Funktion, die auf jedem Intervall $[a, r]$ eine Regelfunktion ist, z.B.: f ist eine stetige Funktion.

Falls der Grenzwert $\lim\limits_{r \to \infty} \int_a^r f(x)\,\mathrm{dx}$ existiert, schreibt man:

$$\int_a^\infty f(x)\,\mathrm{dx} = \lim_{r \to \infty} \int_a^r f(x)\,\mathrm{dx}$$

Analog definiert man:

$$\int_{-\infty}^a f(x)\,\mathrm{dx} = \lim_{r \to -\infty} \int_r^a f(x)\,\mathrm{dx}$$

Fall 2: Schließlich sieht man, warum das Sinn macht:

$$\int_{-\infty}^\infty f(x)\,\mathrm{dx} = \int_{-\infty}^a f(x)\,\mathrm{dx} + \int_a^\infty f(x)\,\mathrm{dx}$$

Beispiel:

(a) $\int_0^r \sin x\,\mathrm{dx} = -\cos x\Big|_0^r = \cos 0 - \cos r = 1 - \cos r$

Der Grenzwert für $r \to \infty$ existiert nicht.

(b) $\int_r^0 \exp x\,\mathrm{dx} = \exp x\Big|_r^0 = \exp 0 - \exp r = 1 - \exp r$

$\lim\limits_{r \to -\infty} \exp r = 0$ also $\int_{-\infty}^0 \exp x = 1$

Fall 3: Ist $X = [a, b[$ und ist f auf dem Interfall $[0, r]$ eine Regelfunktion für $a < r < b$, und existiert $\lim\limits_{r \to b} \int_a^b f(x)\,\mathrm{dx}$, setzt man

$$\int_a^b f(x)\,\mathrm{dx} = \lim_{r \to b} \int_a^r f(x)\,\mathrm{dx}$$

Analoge Definition für $X =]a, b]$ oder $X =]a, b[$.

8.9 Das Integralvergleichs-Kriterium

Satz §8.13

Sei $f\colon [1, \infty[\longrightarrow \mathbb{R}$ monoton fallend. Dann gilt für alle n:

$$0 \leq \left(\sum_{k=1}^n f(k) - \int_1^{n+1} f(x)\,\mathrm{dx} \right) \leq f(1)$$

und der Term in der Mitte konvergiert für $n \to \infty$.

$$0 \leq \lim_{n \to \infty} \left(\sum_{k=1}^n f(k) - \int_1^{n+1} f(x)\,\mathrm{dx} \right) \leq f(1)$$

Die Reihe konvergiert genau dann, wenn das Integral konvergiert.

Beweis: Für $x \in [k, k+1]$ gilt

$$f(k) \geq f(x) \geq f(k+1) \qquad \text{also} \qquad f(k) \geq \int_k^{k+1} f(x)\,\mathrm{dx} \geq f(k+1)$$

Wir addieren diese Ungleichungen:

$$\begin{array}{rccccc} & \sum_{k=1}^{n} f(k) & \geq & \int_1^{n+1} f(x)\,\mathrm{dx} & \geq & \sum_{k=2}^{n+1} f(k) \\ \Leftrightarrow & 0 & \geq & \int_1^{n+1} f(x)\,\mathrm{dx} - \sum_{k=1}^{n} f(k) & \geq & f(n+1) - f(1) \\ \Leftrightarrow & 0 & \leq & \sum_{k=1}^{n} f(k) - \int_1^{n+1} f(x)\,\mathrm{dx} & \leq & f(1) - f(n+1) \end{array}$$

Also ist $\sum_{k=1}^{n} f(x) - \int_1^{n+1} f(x)\,\mathrm{dx}$ eine beschränkte monotone Folge und damit konvergent. □

Vorlesung 26.01.2004

Beispiel §8.14

Beispiel

- $f(x) = \frac{1}{x}$

  $\lim\limits_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k} - \int\limits_1^{n+1} \frac{1}{x}\,\mathrm{dx} = \lim\limits_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k} - \ln x\Big|_1^{n+1} = \lim\limits_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k} - \ln(n+1)$

  Die harmonische Reihe wächst also ungefähr wie der Logarithmus.

- $f(x) = \frac{1}{x^s} \qquad s \geq 2 \qquad \int_1^{n+1} \frac{1}{x^s}\,\mathrm{dx} = \frac{-1}{s-1} \cdot \frac{1}{x^{s-1}}\Big|_1^{n+1}$

  $\lim\limits_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k^s} - \int\limits_1^{n+1} \frac{1}{x^s}\,\mathrm{dx} = \lim\limits_{n\to\infty} \sum_{k=1}^{n} \frac{1}{k^s} - \frac{1}{s-1} + \frac{1}{s-1} \cdot \frac{1}{(n+1)^{s-1}} \leq 1$

  Also konvergiert die Reihe $\sum_{k=1}^{\infty} \frac{1}{k^s}$ für $s \geq 2$.

8.10 Die Taylor-Entwicklung

Satz §8.15

Es gilt $f(t) = f(t_0) + f(t) - f(t_0) = f(t_0) + \int\limits_{t_0}^{t} f'(s)\,\mathrm{ds}$.

Satz: Sei $f \in \mathrm{C}^{n+1}(X, \mathbb{R})$, also $n+1$-mal stetig differenzierbar. Sei $t_0 \in X$. Dann gilt die **Taylor-Formel**:

$$f(t) = \underbrace{\sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(t_0) \cdot (t-t_0)^k}_{\text{Taylor}-\text{Polynom}} + \underbrace{R_n(t, t_0)}_{\text{Restglied}}$$

$f^{(k)}(t_0)$ ist die k-te Ableitung an der Stelle t_0.

$$R_n(t, t_0) = \frac{1}{n!} \int_{t_0}^{t} (t-s)^n \cdot f^{(n+1)}(s)\,\mathrm{ds} = \frac{1}{(n+1)!} \cdot f^{(n+1)}(s_0) \cdot (t-t_0)^{n+1}$$

mit $|s_0 - t_0| \leq |t - t_0|$.

Beweis: Induktion nach n. Für $n = 0$ steht folgendes da:

$$\begin{aligned} f(t) &= \tfrac{1}{0!} f^{(0)} f(t_0) \cdot 1 + \tfrac{1}{0!} \int_{t_0}^{t} 1 \cdot f^{(1)}(s) \text{ ds} \\ &= f(t_0) + \int_{t_0}^{t} f'(s) \text{ ds} \\ &= f(t_0) + f(t) - f(t_0) = f(t) \end{aligned}$$

Induktionsschritt: Wir nehmen an, die Formel stimmt für n, $f \in \mathrm{C}^{n+1}(X, \mathbb{R})$, wir wollen den Satz für $n+1$ beweisen.

$$\begin{aligned} f(t) &= \sum_{k=0}^{n} \tfrac{1}{k!} f^{(k)}(t_0) \cdot (t - t_0)^k + R_n(t, t_0) \\ R_n(t, t_0) &= \tfrac{1}{n!} \int_{t_0}^{t} f^{(n+1)}(s) \cdot (t-s)^n \text{ ds} \\ &= \tfrac{(-1)^n}{n!} \cdot \tfrac{(s-t)^{n+1}}{n+1} f^{(n+1)}(s) \Big|_{s=t_0}^{s=t} - \tfrac{(-1)^n}{n!} \int_{t_0}^{t} f^{(n+2)}(s) \tfrac{(s-t)^{n+1}}{n+1} \text{ ds} \\ &= \tfrac{1}{(n+1)!} (t - t_0) f^{(n+1)}(t_0) + R_{n+1}(t, t_0) \qquad \square \end{aligned}$$

Die 2. Formel ist der Mittelwertsatz der Integralrechnung.

Bemerkung: Ist $f \in \mathrm{C}^{\infty}(X, \mathbb{R})$ glatt, kann man die **Taylor-Reihe** — Bemerkung §8.16

$$\sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(t_0) \cdot (t - t_0)^k$$

konstruieren. Trotzdem gilt i.A. *nicht*, dass $f(t) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(t_0) \cdot (t - t_0)^k$.

Wichtiger als die Taylor-Reihe ist die Taylor-Formel.

8.11 Lokale Extrema

Definition §8.17

Wir sagen f hat in t_0 ein **striktes lokales Maximum (striktes lokales Minimum)**, falls es ein $r > 0$ gibt, so dass für alle t mit $|t - t_0| < r$ gilt:

$$f(t) < f(t_0) \quad (\text{bzw. } f(t) > f(t_0))$$

Satz: Ist $f \in \mathrm{C}^2(X, \mathbb{R})$ und $t_0 \in X$ mit $f'(t_0) = 0$ und $f''(t_0) < 0$ (bzw. $f''(t_0) > 0$), dann hat f in t_0 ein striktes lokales Maximum (striktes lokales Minimum).

Beweis:

$$\begin{aligned} f(t) &= f(t_0) + f'(t_0)(t - t_0) + \tfrac{1}{2} \int_{t_0}^{t} f''(s)(t-s)^2 \text{ ds} \\ &= f(t_0) + \tfrac{1}{2} \int_{t_0}^{t} f''(s)(t-s)^2 \text{ ds} \end{aligned}$$

Weil f' stetig ist, gibt es ein $r > 0$ mit $f'(s) < 0$ für alle s mit $|s - t_0| < r$.

Für $|t - t_0| < r$ und $t \neq t_0$ ist also $\int_{t_0}^{t} f'(s) \cdot (t-s)^2 \text{ ds} < 0$, also $f(t) < f(t_0)$

Für das strikte lokale Minimum argumentiert man entsprechend.

Beispiele:

Beispiel §8.18

- $f(t) = \cos(t) \qquad f'(x) = -\sin(x) \qquad f''(x) = -\cos(x)$

  $f'(0) = 0 \qquad f''(0) = -\cos 0 = -1$

  Lokales Maximum in 0.

- $f(t) = t^3 \qquad f'(t) = 3t^2 \qquad f''(t) = 6t$

  $f'(0) = 0 \qquad f''(0) = 0$

  Hier liefert der Satz keine Information.

- $f(t) = t^4 \qquad f'(t) = 4t^3 \qquad f''(t) = 12t^2$

  $f'(0) = 0 = f''(0)$

  Hier liefert der Satz auch keine Information, obwohl 0 ein Minimum ist.

9 Algorithmen

Vorlesung 29.01.2004

Als erstes Problem betrachten wir die **Polynominterpolation**.

9.1 Polynominterpolation

Gegeben sind *Stützstellen* $s_0, s_1, \ldots, s_n$ und *Wertvorgaben* $w_0, w_1, \ldots, w_n$. Gesucht ist eine Polynomfunktion f mit $f(s_0) = w_0$, $f(s_1) = w_1, \ldots, f(s_n) = w_n$.

Die Stützstellen, Wertvorgaben und Lösungen sind im Folgenden *reell* oder *komplex.*

9.1.1 Überlegungen über Polynome

Paragraph §9.1

Sei $f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$ eine Polynomfunktion. Die Variable x und die Koeffizienten $a_0, \ldots, a_n$ dürfen *reell* oder *komplex* sein.

Ist $a_n \neq 0$ so heißt n der Grad des Polynoms f, $n = \deg(f)$.
Für das Nullpolynom setzt man $\deg(0) = -\infty$.

Beispiele

- $\deg(f) = 0$ falls $f(x) = 3$ für alle x
- $\deg(f) = 1$ falls $f(x) = 3x - 7$ für alle x
- $\deg(f) = 2$ falls $f(x) = (3x - 7)(x + 5)$ für alle x

Der Grad ist **wohldefiniert**, d.h. die Koeffizienten $a_n, \ldots, a_0$ lassen sich von der Polynomfunktion eindeutig ablesen:

$$f(0) = a_0, \quad f'(0) = a_1, \quad f''(0) = 2a_2, \quad \ldots, \quad f^{(k)}(x) = k! a_k$$

9.1.2 Rechenregeln für den Grad

Für Polynome gilt stets:

- $\deg(f \cdot g) = \deg(f) + \deg(g)$
- $\deg(f + g) \leq \max\{\deg(f), \deg(g)\}$

Beispiel: $\deg\big((x^7 + 3) + (4x - x^7)\big) = \deg(4x + 3) = 1$

Ist f ein Polynom mit $\deg(f) = n$, so ist $g(x) = f(x + x_0)$ ebenfalls ein Polynom mit $\deg(f) = \deg(g)$.

9.1.3 Polynome und Nullstellen

Satz §9.2

Ist f ein Polynom mit $\deg(f) \geq 0$, und ist x_0 eine **Nullstelle** (d.h. $f(x_0) = 0$), dann gibt es ein Polynom g mit $f(x) = g(x)(x - x_0)$ mit $\deg(g) = \deg(f) - 1$.

Beweis: Falls $x_0 = 0$ eine Nullstelle ist, so ist $a_0 = 0$, also:

$$f(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x = (a_n x^n + \ldots + a_1)(x - 0)$$

Für x_0 beliebig, betrachte $h(x) = f(x + x_0)$, dann ist $h(0) = f(x_0) = 0$ also $h(x) = h_1(x - x_0)(x - x_0) = g(x)$. □

Folgerung §9.3

Folgerung: Ist f ein Polynom mit $\deg(f) = n \geq 0$ (d.h. $f \neq 0$), dann hat f höchstens n verschiedene Nullstellen.

Beweis: Seien x_1, x_2, …, x_r Nullstellen von f. Dann gilt nach dem vorigen Satz:

$$\underbrace{f(x)}_{\deg=n} = \underbrace{h(x)}_{\deg(h)=n-r\geq 0} (x - x_1)(x - x_2)\cdots(x - x_r)$$

9.1.4 Das Lagrange-Polynom

Definition §9.4

Seien $s_0, s_1, \ldots, s_n$ gegeben mit $s_i \neq s_j$ für $i \neq j$. Das k-te **Lagrange-Polynom** ist:

$$L_k(x) = \frac{x - s_0}{s_k - s_0} \cdot \frac{x - s_1}{s_k - s_1} \cdot \ldots \cdot \frac{x - s_{k-1}}{s_k - s_{k-1}} \cdot \frac{x - s_{k+1}}{s_k - s_{k+1}} \cdot \ldots \cdot \frac{x - s_n}{s_k - s_n}$$

Dann ist $L_k(x)$ ein Polynom vom Grad n.

$$L_k(s_j) = \begin{cases} 0 & \text{falls } k \neq j \\ 1 & \text{falls } k = j \end{cases}$$

Setzt man $f(x) = w_0 L_0(x) + w_1 L_1(x) + \cdots + w_n L_n(x)$ dann gilt $\deg(f) \leq n$.

Weiter gilt: $f(s_j) = w_0 L_1(s_j) + \cdots + w_n L_n(s_j) = u_j L_j(s_j) = w_j$

Satz §9.5

Satz: Seien $s_0, \ldots, s_n$ paarweise verschiedene Stützstellen. Für Wertvorgaben $w_0, \ldots, w_n$ gibt es genau ein Polynom f mit $f(s_j) = w_j$ für alle $j = 0 \ldots n$ mit $\deg(f) \leq n$.

Beweis: Eine explizite Lösung für f haben wir mit den Lagrange-Polynomen schon gefunden. Eine andere Lösung gibt es nicht, denn angenommen f_1 wäre eine weitere Lösung mit $\deg(f_1) \leq n$, $f_1(s_j) = w_j$ mit $j = 0 \ldots n$, $\deg(f) \leq n$, $f(s_j) = w_j$ mit $j = 0 \ldots n$. Also gilt $\deg(f - f_1) \leq n$ und $(f - f_1)(s_j) = w_j - w_j = 0$. Also hat $f - f_1$ $n + 1$ Nullstellen. Also ist $f = f_1$. □

Bemerkung §9.6

Bemerkung: Betrachte $h(x) = q(x)(x - s_0)(x - s_1)\cdots(x - s_n)$ für q beliebig. Dann gilt $h(s_j) = 0$ für alle j, d.h. $h + f$ löst auch das Interpolationsproblem, aber $\deg(h + f) \geq n$ im Allgemeinen.

9.1.5 Anzahl der Rechenoperationen

Vorlesung 02.02.2004

Gesucht: Eindeutige Lösung für das Polynom f mit $\deg(f) \leq n$ und den Stützstellen $s_1, \ldots, s_n$ mit $f(s_j) = q_j$.

Definition §9.7

Insbesondere bestimmt das $n+1$-Tupel $(w_0, w_1, \ldots, w_n)$ das Polynom eindeutig. Wir betrachten jetzt wieviel Rechenoperationen benötigt werden.

Schreibweise: Ist $(c_k)_{k\in\mathbb{N}}$ eine Folge nicht negativer Zahlen und $(a_k)_{k\in\mathbb{N}}$ eine weitere solche Folge, dann sagt man, a_k ist von der Ordnung $O(c_k)$ falls es eine Konstante $K \geq 0$ gibt mit:

$$c_k \cdot K \geq a_k \quad \text{für alle } k$$

Das Ordnungssymbol heißt auch Landau'sches O-Symbol.

Beispiel: Sind die Lagrange-Polynome berechnet, so ist die Anzahl der Rechenoperationen, die man braucht um

$$f(x) = \sum_{k=0}^{n} a_k L_k(x)$$

zu berechnen: $O\big((n+1)^2\big)$

9.1.6 Optimierung der Rechenoperationen

Schnelle Auswertung von Polynomen:

Naiv:

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n$$

$$\begin{array}{rl} n & \text{Additionen} \\ \frac{n(n+1)}{2} & \text{Multiplikationen} \end{array}$$

Horner-Schema:

$$f(x) = ((a_n x + a_{n-1})x + a_{n-2})x + \cdots a_0$$

$$\left.\begin{array}{ll} n & \text{Additionen} \\ n & \text{Multiplikationen} \end{array}\right\} O(n) \text{ Rechenoperationen}$$

Problem: Gegeben sind die Polynome Problem §9.8

$$p(x) = a_n x^n + \cdots + a_0 \qquad \text{mit } \deg(p) \leq n$$

$$q(x) = b_n x^n + \cdots + b_0 \qquad \text{mit } \deg(q) \leq n$$

$$p(x) \cdot q(x) = c_{2n} x^{2n} + \cdots + c_0$$

Frage: Wie berechnet man die Koeffizienten $c_{2n}, \ldots, c_0$? Es gilt:

$$c_k = \sum_{j=0}^{2n} a_j b_{k-j}$$

Die direkte Methode braucht $O(n^2)$ Rechenoperationen.

Idee zur Lösung: Wähle Stützstellen $s_0, s_1, \ldots, s_{2n}$ dann gilt: Idee §9.9

$$f(s_j) = p(s_j) \cdot q(s_j)$$

Rekonstruiere f aus den Daten $f(s_j), j = 0, \ldots, 2n$ durch Interpolation.

9.1.7 Die n-ten Einheitswurzeln in $\mathbb{C}$

Definition §9.10

Für $n \in \mathbb{N}, n \geq 2$ betrachte Lösungen der Gleichung $x^n = 1$, das heißt die Nullstellen von

$$x^n - 1$$

Dieses Polynom hat *höchstens* n verschiedene Nullstellen (3. Folgerung). Natürlich gibt es höchstens 2 reelle Nullstellen. Es gibt in $\mathbb{C}$ genau n Nullstellen. Setze:

$$\vartheta = \cos\left(\frac{2\pi}{n}\right) + i \cdot \sin\left(\frac{2\pi}{n}\right) \in \mathbb{C}$$

Es gilt: Für $0 \leq k < l \leq n-1$ ist $\vartheta^k \neq \vartheta^l$ und $\vartheta^{k \cdot n} = 1$.

Die komplexen Nullstellen des Polynoms $x^n - 1$ sind also genau:

$$1 = \vartheta^0, \vartheta, \vartheta^2, \vartheta^3, \ldots, \vartheta^{n-1}$$

Beachte: $\vartheta = \cos\left(\frac{2\pi}{n}\right) + i \cdot \sin\left(\frac{2\pi}{n}\right)$ hängt vom gegebenen n ab. Es gilt immer: $\vartheta^{k+n} = \vartheta^k$. Nach der 3. Folgerung gilt also auch

$$(x^n - 1)^n = (x-1)(x-\vartheta)(x-\vartheta^2)\cdots \quad \text{(Abspaltung der Nullstellen)}$$

9.1.8 Gerade und ungerade Teilpolynome

Überlegung §9.11

Jedes Polynom $f(x) = a_n x^n + \cdots + a_0$ lässt sich zerlegen in einen **geraden Teil** und einen **ungeraden Teil**. Es gilt:

$$\begin{aligned}
f(x) &= f_g(x^2) + x \cdot f_u(x^2) \\
f_g(t) &= a_0 + a_2 t + a_4 t^2 + \cdots \\
f_u(t) &= a_1 + a_3 t + a_5 t^2 + \cdots \\
f_g(x^2) + x \cdot f_u(x^2) &= a_0 + a_2 x^2 + a_4 x^4 + \cdots \\
&+ x(a_1 + a_3 x^2 + a_5 x^4 + \cdots)
\end{aligned}$$

Algorithmus §9.12

Algorithmus: Gegeben sind Polynome p und q, wir suchen $p \cdot q$.

1. Schritt:	Auswerten von p und q auf geeigneten Einheitswurzeln. Sei $m = 2^r > \deg(p), \deg(q)$.
Setze:	$\vartheta = \cos\left(\frac{2\pi}{m}\right) + i \cdot \sin\left(\frac{2\pi}{m}\right)$ (m-te Einheitswurzel)
Gesucht:	$p(1), p(\vartheta), p(\vartheta^2), \ldots, p(\vartheta^{n-1})$ und $p(\vartheta^k) = p_g(\vartheta^{2k}) + \vartheta p_u(\vartheta^{2k})$
Daten:	$p(x) = a_{n-1}x^{n-1} + \cdots + a_0$ $A = (a_0, a_1, a_2, \ldots, a_{m-1})$
Ausgabe:	$P\big(p(1), p(\vartheta), p(\vartheta^2), \ldots, p(\vartheta^{n-1})\big)$

```
procedure Ausw(m, A, alpha, P)
  if m = 1 then
    P := (A(0))
  else
    m2 := m / 2

  AG := (A(0), A(2), ..., A(m - 2))
  AU := (A(1), A(3), ..., A(m - 1))

  Ausw(m2, AG, alpha^2, PG)
  Ausw(m2, AU, alpha^2, PU)

  for j := 0 to m2 - 1 do
    P(j) := PG(j) + theta * PU(j)

  P(j + m2) := PG(j) - theta * PU(j)
end.
```

Wir haben benutzt, dass $\vartheta^{k+m_2} = -\vartheta^k$ und $\vartheta^0, \vartheta^2, \vartheta^4, \ldots$ die $\frac{m}{2}$-ten Einheitswurzeln sind.

Anzahl der benötigten Multiplikationen
$M(m) = \#$Multiplikationen für $m = 2^r$

$m = 1$:	$M(1) = 0$
$m = 2^r > 1$:	Wir erhalten wir rekursiv $M(2^r) = 2M(2^{r-1})$

Also gilt $M(m) = \frac{m}{2}\log_2(m)$. Die Anzahl der Rechenoperationen zum Berechnen

$$P = (p(1), p(\vartheta), \ldots, p(\vartheta^{m-1})$$

mit $O(m \log_2 m)$

9.1.9 Das Interpolationsproblem

Vorlesung 05.02.2004
Paragraph §9.14

Ist $z = (z_0, z_1, \ldots, z_{m-1}) \in \mathbb{C}^m$, setze

$$V(z) = \begin{pmatrix} 1 & z_0 & z_0^2 & \cdots & z_0^{m-1} \\ 1 & z_1 & z_1^2 & \cdots & z_1^{m-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & z_{m-1} & z_{m-1}^2 & \cdots & z_{m-1}^{m-1} \end{pmatrix}$$

$V(z)$ ist eine $m \times m$-Matrix, die **Vandermondsche Matrix** zu z.

Ist $a = (a_0, a_1, \ldots, a_{m-1})$, so ist

$$V(z) \cdot a^T = \begin{pmatrix} a_0 + a_1 z_0 + a_2 z_0^2 + \ldots + a_{m-1} z_0^{m-1} \\ \vdots \end{pmatrix}$$

das heißt, für $p(x) = a_0 + a_1 x + \ldots + a_{m-1}x^{m-1}$ gilt: $V(z)a^T = \begin{pmatrix} p(z_0) \\ p(z_1) \\ \vdots \\ p(z_{m-1}) \end{pmatrix}$.

Idee: Man multipliziere diese Gleichung von links mit $V(z)^{-1}$, dann erhält man:

$$a^T = V(z)^{-1} = \begin{pmatrix} p(z_0) \\ p(z_1) \\ \vdots \\ p(z_{m-1}) \end{pmatrix}$$

Satz §9.15

Satz: Sei $z = (\vartheta^0, \vartheta^1, \ldots, \vartheta^{m-1})$ und $w = (\vartheta^0, \vartheta^{-1}, \ldots, \vartheta^{1-m})$. Dann gilt:

$$V(z) \cdot V(w) = \begin{pmatrix} m & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & m \end{pmatrix} \quad \Leftrightarrow \quad V(z)^{-1} = \frac{1}{m} V(w)$$

Beweis: $V(z)_{ij} = z_i^j = (\vartheta^i)^j = \vartheta^{ij}$, $V(w)_{jk} = z_j^{-k} = (\vartheta^{-j})^k = \vartheta^{-jk}$, also

$$(V(z) \cdot V(w))_{i,k} = \sum_{j=0}^{m-1} \vartheta^{ij} \cdot \vartheta^{-jk} = \sum_{j=0}^{m-1} \vartheta^{j(i-k)} = \begin{cases} m & \text{falls } i = k \\ \frac{\vartheta^{(i-k)m}-1}{\vartheta^{i-k}-1} & \text{falls } i \neq k \end{cases}$$

Damit haben wir das Interpolationsproblem gelöst!

$$(a_0, a_1, \ldots, a_{m-1}) = V(\vartheta^0, \ldots, \vartheta^{1-m}) \cdot \begin{pmatrix} p(\vartheta^0) \\ \vdots \\ p(\vartheta^{m-1}) \end{pmatrix}$$

Daten: $(p(1), p(\vartheta), p(\vartheta^2), \ldots, p(\vartheta^{m-1}))$
Gesucht: $(a_0, a_1, \ldots, a_{m-1})$ Koeffizienten

```
procedure Koeff (m, B, ϑ, P)
  Ausw(m, B, ϑ^-1, C)
  P = 1/m C
end
```

Wenn man das aufruft mit $B = (p(1), p(\vartheta), \ldots, p(\vartheta^{m-1}))$ erhält man $P = (a_0, a_1, \ldots, a_{m-1})$

9.1.10 Zusammenfassung

Paragraph §9.16

Gegeben sind Polynome p und q, die multipliziert werden sollen.

- Wähle l minimal mit $z^l > 2\deg(p), 2\deg(q)$.
- Rufe `Ausw` für p und q auf, um $(p(\vartheta^0), \ldots, p(\vartheta^{m-1}))$ und $(q(\vartheta^0), \ldots, q(\vartheta^{m-1}))$ zu berechnen.
- Bestimme $B = (p(\vartheta^0) \cdot q(\vartheta^0), \ldots, p(\vartheta^{m-1}) \cdot q(\vartheta^{m-1}))$.
- Berechne Koeffizienten des zugehörigen Interpolationspolynom mit `Koeff`.

Rechenaufwand: $O(m \log_2 m)$ mit $m = 2^l$, $\deg(p), \deg(q) \leq k$.
Aufwand: $O(k \log_2 k)$.

Das Verfahren ist gut, wenn p und q ähnlichen Grad haben. Wegen

$$\vartheta = \cos\left(\frac{2\pi}{m}\right) + i \sin\left(\frac{2\pi}{m}\right)$$

braucht man eine gute Gleitzahlarithmetik, selbst für ganzzahlige Polynome!

9.1.11 Anwendung

Paragraph §9.17

Schnelle Multiplikation großer Zahlen. M, N natürliche Zahlen in g-adischer Entwicklung,

$$\begin{aligned} M &= a_0 + ga_1 + g^2 a_2 + \ldots + g^k a_k \\ N &= b_0 + gb_1 + g^2 b_2 + \ldots + g^k b_k \end{aligned}$$

Unser Algorithmus bestimmt $p \cdot q$ für $p(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_k x^k$ und $q(x) = b_0 + b_1 x + b_2 x^2 + \ldots + b_k x^k$.

$$N \cdot M = p(g) \cdot q(g) = c_0 + c_1 g + c_2 g^2 + \ldots + c_{2k} g^{2k}$$

Rechenaufwand: $O(k \log_2 k)$

Literatur:

- Schönhage-Strassen: *Schnelle Multiplikation großer Zahlen*, Computing 7 (1971), Seiten 281-292
- Knuth, *Arithmetik*, Springer, Berlin, 2001.

9.2 Das Newton-Verfahren

Nun betrachten wir das **Newton-Verfahren** zur Bestimmtung von **Nullstellen**.

9.2.1 Problemstellung

Funktion $f\colon X \longrightarrow \mathbb{R}$, gesucht sind die Nullstellen von f, also $f(x_0) = 0$

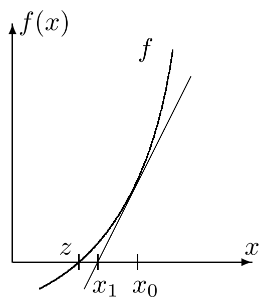


Abbildung 16: Newton-Verfahren

Idee: $f(z) = 0$ Nullstelle, x_0 liegt *nahe* bei z. Konstruiere bessere Approximation $x_1, x_2, x_3, \ldots$ bis eine gute Näherungslösung für z gefunden ist.
x_1 ist Nullstelle der Tangente. Die Steigung ist $f'(x_0) = \frac{f(x_0)}{x_0 - x_1}$.

$$x_0 - x_1 = \frac{f(x_0)}{f'(x_0)} \qquad x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}$$

Setze $\varphi(x) = x - \frac{f(x)}{f'(x)}$. $x_1 = \varphi(x_0)$, $x_2 = \varphi(x_1)$, …, $x_{n+1} = \varphi(x_n)$.

Konvergiert das gegen z?

9.2.2 Konvergenz des Newton-Verfahrens

Satz §9.18

Sei f eine C^2-Funktion, $f(z) = 0$ und $f'(z) \neq 0$. Dann konvergiert das Newton-Verfahren gegen z, sogar gut, falls x_0 nahe genug bei z liegt.

Beweis: Wegen $f'(z) \neq 0$ gibt es ein $r > 0$, so dass $f'(x) \neq 0$ falls $|x - z| < r$. Dann ist auch $\varphi(x) = x - \frac{f(x)}{f'(x)}$ ebenfalls eine C^2-Funktion.

$$\varphi'(x) = 1 - \frac{1}{f'(x)^2} \cdot \big(f'(x)^2 + f(x) f''(x)\big)$$

also $\varphi'(x) = 0$. Nach der Taylor-Formel gilt:
$\varphi(x) - \varphi(z) = (x - z)^2 \cdot \underbrace{\varphi''(s)\frac{1}{2}}_{\leq k}$ für ein s mit $|z - s| \leq |x - z| \leq r$.

$$|\varphi(x) - \varphi(z)| = |x - z|^2 \cdot k$$

Die Iteration liefert also eine Nullfolge, im Grenzwert erhalten wir $\varphi(z) = z$. $\square$

Beispiele

- **Wurzelziehen:** $f(x) = a - x^2$
  $\varphi(x) = \frac{1}{2}x + \frac{9}{2x}$
  $x_{n+1} = \frac{1}{2}x_n + \frac{9}{2x_n}$
- **Division:** $f(x) = \frac{1}{x} - a$
  $\varphi(x) = 2x - ax^2$
  $x_{n+1} = 2x_n - ax_n^2$

10 Metrische Räume und Normen

Vorlesung 13.04.2004

Gegeben sei eine (nicht leere) Menge X. Wir wollen auf X einen **Abstandsbegriff**. Der Abstand zweier Elemente in X soll eine nicht negative reelle Zahl sein. Eine Einheit wie zum Beispiel *Kilometer* oder *Stunden* lassen wir weg, weil wir sie nicht brauchen.

Erinnerung: $X \times X$ ist die Menge aller **Paare** (u, v) mit $u, v \in X$.

10.1 Die Axiome der Metrik

Definition §10.1

Eine Abbildung $d\colon X \times X \longrightarrow \mathbb{R}$ heißt **Metrik** auf X, falls folgendes gilt:

(M_1) $d(u, v) = d(v, u) \geq 0$ gilt für alle $u, v \in X$ (**symmetrisch**, **positiv**)
(M_2) $d(u, v) = 0$ genau dann, wenn $u = v$ gilt.
(M_3) $d(u, w) \leq d(u, v) + d(v, w)$ für alle $u, v, w \in X$ (**Dreiecksungleichung**)

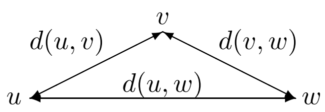


Abbildung 17: Die Dreiecksungleichung

Ist d eine Metrik auf X, so heißt (X, d) **metrischer Raum**. Die Axiome (M_1), (M_2) und (M_3) formalisieren Alltagserfahrungen mit *physikalischen* Abständen.

10.2 Beispiele

Beispiel §10.2

Beispiel (a) X mit reeller Metrik:

$$X = \mathbb{R},\ d(u, v) = |u - v|$$

Behauptung: Das ist eine Metrik. (M_1) gilt offensichtlich, ebenso (M_2). Nach Analysis I §1.4 gilt auch (M_3).

Beispiel (b) X mit diskreter Metrik:

$$d(u, v) = \begin{cases} 0 & \text{falls} \quad u = v \\ 1 & \text{falls} \quad u \neq v \end{cases}$$

(M_1) und (M_2) gelten. Zu (M_3): Für $d(u, w) = 0$ ist nichts zu zeigen. Falls $d(u, w) = 1 \Rightarrow u \neq w$. Für jedes $v \in X$ gilt dann: $v \neq w$ oder $v \neq u$, also $d(v, w) = 1$ oder $d(v, u) = 1$, also $d(u, w) \leq d(v, w) + d(v, u)$.

Diese Metrik heißt **diskrete Metrik** (weil der Wertebereich nur bestimmte Werte annehmen kann).

Beispiel (c) Computerbildschirm:

$$I = \{1, 2, 3, \dots, 1024\} \quad J = \{1, 2, 3, \dots, 768\} \quad X = I \times J$$
$$d: X \times X \longrightarrow \mathbb{R} \qquad d((u_1, v_1), (u_2, v_2)) = |u_1 - u_2| + |v_1 - v_2|$$

(M_1) und (M_2) gelten. (M_3) gilt auch:

$$\begin{array}{rcl} |u_1 - u_3| & \leq & |u_1 - u_2| + |u_2 - u_3| \\ |v_1 - v_3| & \leq & |v_1 - v_2| + |v_2 - v_3| \end{array}$$

also $d((u_1, v_1), (u_3, v_3)) \leq d((u_1, v_1), (u_2, v_2)) + d((u_2, v_2), (u_3, v_3)) \Rightarrow (M_3)$

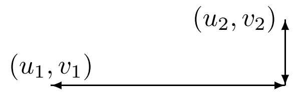


Abbildung 18: Pixelabstand am Bildschirm

Beobachtung: Ist (X, d) ein metrischer Raum, und ist $A \subseteq X$ eine (nicht leere) Teilmenge, so ist die Einschränkung der Metrik auf $A \times A$ wieder eine Metrik, man sagt (A, d) ist ein **metrischer Teilraum** (oder **Unterraum**) von (X, d). Beobachtung §10.3

10.3 Die offene r-Kugel

Definition §10.4

Es sei (X, d) ein metrischer Raum, $r \in \mathbb{R}$, $r > 0$.
Für $u \in X$ ist $B_r(u) = \{v \in X \mid d(u, v) < r\}$ **die offene r-Kugel um** u.

In den vorherigen Beispielen sehen die Kugeln wie folgt aus:

Zu Beispiel (a) X mit reeller Metrik:

$$B_r(u) =]u - r, u + r[\subseteq \mathbb{R}$$

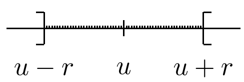


Abbildung 19: Offene r-Kugel in $\mathbb{R}$

Zu Beispiel (b) X mit diskreter Metrik:

$$\begin{array}{rcl} B_1(u) & = & \{v \in X | d(u, v) < 1\} = \{u\} \\ B_2(u) & = & \{v \in X | d(u, v) < 2\} = X \end{array}$$

Zu Beispiel (c) Computerbildschirm:

$$\begin{aligned} B_1((u_1,v_1)) &= \{(u_1,v_1)\} \\ B_{\frac{3}{2}}((u_1,v_1)) &= \{(u_1,v_1),(u_1\pm 1,v_1),(u_1,v_1\pm 1)\} \end{aligned}$$

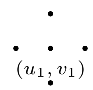


Abbildung 20: Offene r-Kugel von Bildschirmpixeln

10.4 Folgen und metrische Räume

Definition: Eine **Folge** in einem metrischen Raum (X,d) ist eine Abbildung $\mathbb{N} \longrightarrow X$, die jeder natürlichen Zahl n ein Folgeglied $x_n \in X$ zuordnet. Schreibe wie gehabt $(x_n)_{n\in\mathbb{N}}$.

Sei (X,d) ein metrischer Raum, $(x_n)_{n\in\mathbb{N}}$ eine Folge in X. Wir sagen $(x_n)_{n\in\mathbb{N}}$ Definition §10.5
konvergiert gegen $x \in X$, $\lim\limits_{n\to\infty} x_n = x$, falls gilt: Für jedes $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$ (abhängig von ε), so dass für alle $k \geq N$ gilt $d(x_k,x) < \varepsilon$.

$$d(x,y) < \varepsilon \quad \text{ist äquivalent zu} \quad y \in B_\varepsilon(x)$$

Eine Folge hat höchstens *einen* Grenzwert: Falls $\lim\limits_{n\to\infty} x_n = x$ und $\lim\limits_{n\to\infty} x_n = y$, so gilt $x = y$, denn sonst wäre $d(x,y) > 0$. Setze $\varepsilon = \frac{1}{2}d(x,y) > 0$. Für $k \geq N$ gilt dann $d(x_k,x) < \varepsilon$, $d(x_k,y) < \varepsilon$. Mit der Dreiecksungleichung folgt:

$$d(x,y) \leq d(x,x_k) + d(x_k,y) < 2\varepsilon = d(x,y) \quad \text{(Widerspruch)}$$

Zu Beispiel (a) X mit reeller Metrik:

$X = \mathbb{R},\ d(u,v) = |u-v| \rightsquigarrow$ Konvergenzbegriff aus Analysis I §3.1

Zu Beispiel (b) X mit diskreter Metrik:

Setze $\varepsilon = 1$. Wenn $\lim\limits_{n\to\infty} x_n = x$, so gibt es also $N \in \mathbb{N}$ mit: $d(x_k,x) < 1$ für alle $k \geq N$, d.h. $x_k = x$ für alle $k \geq N$.

Zu Beispiel (c) Computerbildschrim:

Es gilt wie in Beipiel (b), dass konvergente Folgen ab einem bestimmten Folgenglied konstant sind.

Vorlesung 16.04.2004

Beispiel: $X = \mathbb{R}^n = \{(v_1,\ldots,v_n) \mid v_1,\ldots,v_n \in \mathbb{R}\}$
Metrik: $d(u,v) = |u_1-v_1| + |u_2-v_2| + \cdots + |u_n-v_n|$

Das ist eine Metrik, die **Manhattan-Taxi-Metrik**. Eine Folge von Vektoren $(v_k)_{k\in\mathbb{N}}$ in X konvergiert gegen einen Vektor u genau dann, wenn für jedes $i = 1,\ldots,n$ die Folge der i-ten Einträge $(v_{i,k})_{k\in\mathbb{N}}$ gegen u_i konvergiert.

Warum? $d(u,v) \geq |u_i - v_i|$ für jedes $i = 1,\ldots,n$

10.5 Cauchy-Folge

Definition §10.6

Eine Folge $(v_n)_{n\in\mathbb{N}}$ im metrischen Raum heißt **Cauchy-Folge**, wenn sie folgende Eigenschaft hat:

(CF) Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$ so, dass für alle $k, \ell \geq N$ gilt: $d(v_k, v_\ell) < \varepsilon$

Für $X = \mathbb{R}$, $d(u, v) = |u - v|$ ist das genau die Definition aus Analysis I (§3.17).

Satz: Ist $(x_n)_{n\in\mathbb{N}}$ eine konvergente Folge in einem metrischen Raum X, dann ist $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge. Satz §10.7

Beweis: Sei $\varepsilon > 0$. Es gibt ein $N \in \mathbb{N}$ so, dass für alle $k \geq N$ gilt $d(x_k, x) < \frac{\varepsilon}{2}$, wobei $x = \lim\limits_{n\to\infty} x_n$. Für $k, \ell \geq N$ gilt:

$$d(x_k, x_\ell) \leq \underbrace{d(x_k, x)}_{<\frac{\varepsilon}{2}} + \underbrace{d(x_\ell, x)}_{<\frac{\varepsilon}{2}} < \varepsilon$$

□

10.6 Vollständigkeit

Definition §10.8

Ein metrischer Raum (X, d) heißt **vollständig**, wenn *jede* Cauchy-Folge in X einen Grenzwert in X hat.

Zu Beispiel (a) X mit reeller Metrik:

(X, d) ist vollständig (nach Satz Analysis I §3.18).

Zu Beispiel (b) X mit diskreter Metrik:

(X, d) ist immer vollständig:

Setze $\varepsilon = \frac{1}{2}$. Sei $(x_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge. Für $k, \ell \geq N$ gilt also $d(x_k, x_\ell) < \frac{1}{2}$, also $x_k = x_\ell$. Die Folge wird konstant und konvergiert also.

Zu Beispiel (c) Computerbildschirm:

Der Bildschirm ist ebenfalls vollständig.

Beispiel $X = \mathbb{Q}$, $d(u, v) = |u - v|$:

(X, d) ist *nicht* vollständig, es gibt eine Folge $(x_n)_{n\in\mathbb{N}}$ in $\mathbb{Q}$, die gegen $\sqrt{2} \notin \mathbb{Q}$ konvergiert (in $\mathbb{R}$).

Beispiel $A =]0, 1] = \{x \in \mathbb{R} \mid 0 < x \leq 1\}$, $d(u, v) = |u - v|$:

(A, d) ist *nicht* vollständig. $x_0 = 1,\ x_n = \frac{1}{n} \in A$ konvergiert gegen $0 \notin A$, bildet also eine Cauchy-Folge die *in* A nicht konvergiert.

Bemerkung:

(1) Unterräume metrischer Räume müssen nicht unbedingt vollständig sein ($\rightarrow$ später genaueres Kriterium).

(2) Die Begriffe *Konvergenz, Cauchy-Folge* und *Vollständigkeit* hängen nicht nur von der Menge X, sondern auch und vor allem von der Metrik ab.

10.7 Abgeschlossenheit

Definition §10.9

Eine Teilmenge A in einem metrischen Raum (X, d) heißt **abgeschlossen**, wenn für jede Folge $(a_n)_{n\in\mathbb{N}}$ in A gilt:

Falls $\lim\limits_{n\to\infty} a_n = x$ existiert (in X), so gilt $x \in A$.

Beispiel: $A =]0,1]$, $X = \mathbb{R}$, $d(u,v) = |u-v|$ ist *nicht* abgeschlossen, denn $a_n = \frac{1}{n} \in A$ hat keinen Grenzwert in A: $\lim\limits_{n\to\infty} a_n = 0 \notin A$

Erinnerung: Abgeschlossene Intervalle im Sinne der Analysis I sind abgeschlossen in $\mathbb{R}$ mit $|u-v| = d(u,v)$.

$$[r,s], \qquad [r,\infty[, \qquad]-\infty,s] \qquad \Rightarrow \text{ abgeschlossen.}$$

Bemerkung: X und $\emptyset$ sind immer abgeschlossen im Sinne der Definition.

10.8 Eigenschaften abgeschlossener Mengen

Satz §10.10

(a) **Vereinigungen** von *endlich vielen* abgeschlossenen Mengen sind abgeschlossen.

(b) **Durchschnitte** von *beliebig vielen* abgeschlossenen Mengen sind abgeschlossen.

Beweis von (a) mit Induktion nach der Anzahl der abgeschlossenen Mengen. Für *eine* Menge ist das banal. Seien $A_1, A_2, \ldots, A_k, A_{k+1}$ abgeschlossene Mengen. Setze $B = A_1 \cup A_2 \cup \cdots \cup A_k, C = A_{k+1}$. Nach Induktionsannahme dürfen wir annehmen, dass B und C abgeschlossen sind, zu zeigen ist die Abgeschlossenheit von $B \cup C$.
Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in $B \cup C$, mit dem Grenzwert $x = \lim\limits_{n\to\infty} x_n$. Dann gibt es eine Teilfolge $(x_{nk})_{k\in\mathbb{N}}$, für die alle Folgenglieder in B oder alle Folgenglieder in C liegen. Dann gilt:

$$\lim_{k\to\infty} x_{nk} = x = \lim_{n\to\infty} x_n$$

Da B und C abgeschlossen sind, gilt: $x \in B$ oder $x \in C$, jedenfalls $x \in B \cup C$.

Beweis von (b): Sei $A_i, i \in I$ eine Menge abgeschlossener Teilmengen von X. Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge im Durchschnitt mit $\lim\limits_{n\to\infty} x_n = x$. Zu zeigen ist nun, dass auch x im Durchschnitt liegt.
Für jedes i gilt: $(x_n)_{n\in\mathbb{N}}$ liegt in A_i, also gilt $x \in A_i$ (mit A_i abgeschlossen). Also liegt x im Durchschnitt. $\square$

Beispiel: (a) ist im Allgemeinen falsch, wenn man unendlich viele Mengen zulässt. $X = \mathbb{R},\ d(u,v) = |u - v|$

$A_1 = \{1\},\ A_2 = [\frac{1}{2}, 1],\ A_n = [\frac{1}{n}, 1]$

$\bigcup\limits_{i=1}^{\infty} A_i =]0,1] \quad \Rightarrow$ nicht abgeschlossen.

Satz: Ist (X, d) ein vollständiger metrischer Raum und $A \subseteq X$, so ist A vollständig als Unterraum genau dann, wenn A abgeschlossen ist. Satz §10.11

Beweis: Angenommen A ist abgeschlossen, so ist zu zeigen, dass A vollständig ist. Sei $(a_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in A, also existiert $x = \lim\limits_{n\to\infty} a_n$ (weil X vollständig ist). Weil A abgeschlossen in X ist, gilt $x \in A$.

Angenommen A ist vollständig. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in A, die gegen $x \in X$ konvergiert. Zu zeigen ist $x \in A$. Dann ist $(a_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge, also gilt $x = \lim\limits_{n\to\infty} a_n \in A$, weil A vollständig ist. $\square$

Vorlesung 20.04.2004

Vollständigkeit und Abgeschlossenheit: Für folgende Aussagen gilt Äquivalenz:

- X ist vollständig $\Leftrightarrow$ jede Cauchy-Folge konvergiert.
- $A \subseteq X$ ist abgeschlossen $\Leftrightarrow$ jede konvergente Folge $(a_n)_{n\in\mathbb{N}}$ in A hat ihren Grenzwert in A.

Beispiele:

- $]0,1] \subseteq \mathbb{R}$ ist nicht abgeschlossen
- $[0,1] \subseteq \mathbb{R}$ ist abgeschlossen
- $[0,5] \cup [201, 395]$ ist abgeschlossen
- Abgeschlossene Intervalle in $\mathbb{R}$ (oder endliche Vereinigungen davon) sind abgeschlossen.
- In einem diskreten metrischen Raum oder *Bildschirm-Raum* ist *jede* Teilmenge abgeschlossen.

Bemerkung: Ist A ein Unterraum von (X, d) und ist A vollständig, dann ist A abgeschlossen in X (folgt aus dem Beweis von Satz §10.11).

10.9 Vektorräume und Normen

Eine wichtige Klasse metrischer Räume stammt aus der linearen Algebra, die **Vektorräume**.

Definition §10.12

Definition: Sei V ein **reeller Vektorraum** (über den Körper $\mathbb{R}$). V darf endlich oder unendlich viele Dimensionen haben. Ein Längenmaß für Vektoren heißt **Norm**, wenn gilt:

$$\underbrace{||\ \cdot\ ||}_{\text{Norm}}: V \longrightarrow \mathbb{R}, \quad u \mapsto ||u||$$

(N_1)	$\|\|u\|\| \geq 0$	
	$\|\|u\|\| = 0 \Leftrightarrow u = 0$	für alle $u \in V$
(N_2)	$\|\|ru\|\| = \|r\| \cdot \|\|u\|\|$	für alle $u \in V, r \in \mathbb{R}$
(N_3)	$\|\|u+v\|\| \leq \|\|u\|\| + \|\|v\|\|$	für alle $u, v \in V$
	(Dreiecksungleichung)	

Das Paar $(V, ||\cdot||)$ heißt dann normierter Raum. Jeder normierte Raum ist auch ein metrischer Raum.

10.9.1 Die 1-Norm

$V = \mathbb{R}^n = \{\text{reelle } n\text{-Tupel}\}$

$$||u||_1 = |u_1| + |u_2| + \ldots + |u_n|$$

Dies ist eine Norm, die **1-Norm**. Beweis:

Zu (N_1) Das gilt, klar.
Zu (N_2) $||ru||_1 = |ru_1| + \ldots + |ru_n| = |r| \cdot |u_1| + \ldots + |r| \cdot |u_n| = |r| \cdot ||u||$
Zu (N_3) $||u+v||_1 = |u_1+v_1| + \ldots + |u_n+v_n| \leq |u_1| + |v_1| + \ldots + |u_n| + |v_n| = ||u||_1 + ||v||_1$

Lemma §10.13

Lemma: Ist $(V, ||\cdot||)$ ein normierter Raum, dann ist $d(u,v) = ||u-v||$ eine Metrik.

(M_1) $d(u,v) = ||u-v|| \geq 0$ **(positiv)**
$d(u,v) = ||u-v|| = |-1| \cdot ||u-v|| = ||(-1)\cdot(u-v)||$
$= ||v-u|| = d(v,u)$ **(symmetrisch)**
(M_2) $d(u,v) = ||u-v|| = 0 \Leftrightarrow u-v = 0 \Leftrightarrow u = v$
(M_3) $||u-w|| = ||(u-v)+(v-w)|| \leq ||u-v|| + ||v-w||$

10.9.2 Das euklidische Skalarprodukt

Definition §10.14

Sei $V = \mathbb{R}^n$. Für $u, v \in \mathbb{R}^n$ setze $\langle u \mid v \rangle = u_1v_1 + \ldots + u_nv_n$.

$$\langle \cdot \mid \cdot \rangle: V \times V \longrightarrow \mathbb{R}$$

Es gelten folgende Rechenregeln:

(i)	$\langle u \mid u \rangle \geq 0$	
	$\langle u \mid u \rangle = 0 \Leftrightarrow u = 0$	für alle $u \in V$
(ii)	$\langle u \mid v+w \rangle = \langle u \mid v \rangle + \langle u \mid w \rangle$	
	$\langle u+v \mid w \rangle = \langle u \mid w \rangle + \langle v \mid w \rangle$	für alle $u, v, w \in V$
	(bilinear)	
(iii)	$\langle ru \mid v \rangle = r\langle u \mid v \rangle = \langle u \mid rv \rangle$	für alle $u, v \in V$, $r \in \mathbb{R}$
(iv)	$\langle u \mid v \rangle = \langle v \mid u \rangle$	für alle $u, v \in V$

10.9.3 Die euklidische Norm oder 2-Norm

$$\text{Setze } ||u||_2 = \sqrt{\langle\, u \mid u \,\rangle} = \sqrt{u_1^2 + u_2^2 + \ldots + u_n^2}$$

10.9.4 Die Cauchy-Schwarz-Ungleichung

Satz §10.15

$$\text{Es gilt:} \quad |\langle\, u \mid v \,\rangle| \leq ||u||_2 \cdot ||v||_2 \quad \text{für alle } u, v \in V = \mathbb{R}^n$$

Beweis: Setze $a = \langle\, v \mid v \,\rangle$ und $b = -\langle\, u \mid v \,\rangle$. Dann gilt:

$$\begin{aligned} 0 &\leq \langle\, au + bv \mid au + bv \,\rangle = a^2\langle\, u \mid u \,\rangle + b^2\langle\, v \mid v \,\rangle + ab\langle\, u \mid v \,\rangle + ab\langle\, v \mid u \,\rangle \\ &= a^2\langle\, u \mid u \,\rangle + b^2\langle\, v \mid v \,\rangle + 2ab\langle\, u \mid v \,\rangle \\ & \text{also} \\ 0 &\leq ||v||_2^4 \cdot ||u||_2^2 + \langle\, u \mid v \,\rangle^2 \cdot ||v||_2^2 - 2||v||_2^2 \cdot \langle\, u \mid v \,\rangle^2 \end{aligned}$$

Falls $||v||_2^2 = 0 \Leftrightarrow v = 0$ nach (i), also gilt die Cauchy-Schwarz-Ungleichung. Falls $||v||_2^2 \neq 0$ teilen wir durch $||v||_2^2$.

$$0 \leq ||v||_2^2 \cdot ||u||_2^2 + \langle\, u \mid v \,\rangle^2 - 2\langle\, u \mid v \,\rangle^2$$
$$\langle\, u \mid v \,\rangle^2 \leq (||u||_2 \cdot ||v||_2)^2 \quad \text{also} \quad |\langle\, u \mid v \,\rangle| \leq ||u||_2 \cdot ||v||_2$$

Bemerkung: Es gilt für alle $u, v \in V$:

$$||u + v||_2^2 = ||u||_2^2 + ||v||_2^2 + 2\langle\, u \mid v \,\rangle$$

Satz §10.16

Satz: Die 2-Norm ist eine Norm. Beweis:

Zu (N_1) Das ist klar.

Zu (N_2) $||ru||_2 = \sqrt{\langle\, ru \mid ru \,\rangle} = \sqrt{r^2\langle\, u \mid u \,\rangle} = r||u||_2$

Zu (N_3)
$$\begin{aligned} ||u + v||_2^2 &= ||u||_2^2 + ||v||_2^2 + 2\langle\, u \mid v \,\rangle \\ &\leq^* ||u||_2^2 + ||v||_2^2 + 2||u||_2 \cdot ||v||_2 \\ &= (||u||_2 + ||v||_2)^2 \text{ also } ||u + v||_2 \leq ||u||_2 + ||v||_2 \end{aligned}$$ □

* An dieser Stelle wird die Cauchy-Schwarz-Ungleichung verwendet.

Es gibt auf $\mathbb{R}^n$ viele weitere Normen.

10.9.5 Die Maximums- oder Supremumsnorm

$$||u||_\infty = \max\{|u_1|, |u_2|, \ldots, |u_n|\}$$

Das ist die **Maximums-** oder **Supremumsnorm**. Man rechnet schnell nach, dass das eine Norm ist.

10.9.6 Verhältnis der Normen zueinander

Satz §10.17

Für alle drei Normen auf $\mathbb{R}^n$ gilt:

$$||u||_1 \geq ||u||_2 \geq ||u||_\infty \geq \frac{1}{n}||u||_1 \quad \text{für alle } u \in \mathbb{R}^n$$

Beweis:

$$\begin{aligned} (|u_1| + \ldots + |u_n|)^2 &\geq |u_1|^2 + \ldots + |u_n|^2 \\ |u_1|^2 + \ldots + |u_n|^2 &\geq \max\{|u_1|, \ldots, |u_n|\} \\ n \cdot \max\{|u_1|, \ldots, |u_n|\} &\geq |u_1| + \cdots + |u_n| \end{aligned}$$

□

Konsequenz: Alle drei Normen liefern die selben konvergenten Folgen beziehungsweise die selben Cauchy-Folgen!

Satz §10.18

Satz: $\mathbb{R}^n$ ist bezüglich jeder der drei Normen $|| \cdot ||_1$, $|| \cdot ||_2$, $|| \cdot ||_\infty$ vollständig.
Beweis: Es genügt, die $|| \cdot ||_1$-Norm zu betrachten, denn alle drei Normen liefern die gleichen Cauchy-Folgen bzw. konvergenten Folgen nach Satz §10.17.

Sei $(u_k)_{k\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}^n$ (bezüglich der 1-Norm). Wegen $||u||_1 = |u_1|+|u_2|+\ldots+|u_n|$ gilt für jeden Eintrag u_j: $||u||_1 \geq |u_j|$. Für jedes $j = 1 \ldots n$ ist deshalb die Folge $(u_{j,k})_{k\in\mathbb{N}}$ eine Cauchy-Folge in $\mathbb{R}$, also konvergent nach Analysis I.
Setze $v_j = \lim\limits_{k\to\infty} u_{j,k}$ und $v = (v_1, \ldots, v_n)$. Behauptung: $\lim\limits_{k\to\infty} u_k = v$.

Wähle zu gegebenen $\varepsilon > 0$ eine Zahl $N \in \mathbb{N}$ so, dass $|u_{j,k} - v_j| < \varepsilon$ für alle $k \geq N$ und für alle $j = 1 \ldots n$. Dann gilt $||u_k - v||_1 < n \cdot \varepsilon$, damit folgt die Konvergenz. □

Beispiel §10.19

Beispiel: $X = C([a, b], \mathbb{R}) = \{f\colon [a, b] \longrightarrow \mathbb{R} \mid f \text{ stetig}\}$
Mit der Norm $||f||_\infty = \sup\{|f(x)|, x \in [a, b]\}$ ist dieser Vektorraum vollständig.

Definition §10.20

Definition: Ein vollständiger, normierter Raum heißt **Banachraum**.
Beispiele für Banachräume:

Endlich dimensionale Banachräume:

- $\mathbb{R}$ mit dem Absolutbetrag
- $\mathbb{R}^n$ mit der 1-Norm
- $\mathbb{R}^n$ mit der 2-Norm
- $\mathbb{R}^n$ mit der Supremumsnorm

Unendlich dimensionale Banachräume:

- $\mathrm{C}\,([a, b], \mathbb{R})$ mit der Supremumsnorm
- $\mathrm{Reg}\,([a, b], \mathbb{R})$ mit der Supremumsnorm

Beispiele, die *keine* Banachräume sind:

- Polynomfunktionen mit der Supremumsnorm
- $\mathrm{Step}\,([a, b], \mathbb{R})$ mit der Supremumsnorm

11 Stetigkeit von Funktionen

Wir betrachten Abbildungen zwischen metrischen Räumen.

11.1 Stetigkeit

Definition §11.1

Seien (X, d_X) und (Y, d_Y) metrische Räume. Eine Abbildung $f\colon X \longrightarrow Y$ heißt **stetig**, wenn für jede konvergente Folge $(x_n)_{n\in\mathbb{N}}$ in X gilt:

$$f(\lim_{n\to\infty} x_n) = \lim_{n\to\infty} f(x_n)$$

Beispiel: $X \subseteq \mathbb{R}$, $f\colon X \longrightarrow \mathbb{R}$ stetig $\Leftrightarrow$ stetig im Sinne der Analysis I.

Verfeinerung: $f\colon Y \longrightarrow Y$ ist **stetig im Punkt** $x \in X$, falls für jede Folge $(x_n)_{n\in\mathbb{N}}$ mit $\lim\limits_{n\to\infty} x_n = x$ gilt:

$$\lim_{n\to\infty} f(x_n) = f(x)$$

Also ist f stetig genau dann, wenn f in jedem Punkt $x \in X$ stetig ist.

11.2 Lipschitz-stetige Funktionen

Definition §11.2

Eine Funktion $f\colon X \longrightarrow Y$ heißt **L-Lipschitz-stetig**, wenn es ein $L \in \mathbb{R}$ gibt mit:

$$d_Y\big(f(u), f(v)\big) \leq L \cdot d_X(u, v) \qquad \text{für alle } u, v \in X$$

Lipschitz-stetige Funktionen sind stetig.

Satz §11.3

Satz: Ist $X \xrightarrow{f} Y$ stetig und $Y \xrightarrow{g} Z$ stetig, so ist $g \circ f\colon X \longrightarrow Z$ ebenfalls stetig.

Beweis: Angenommen $(x_n)_{n\in\mathbb{N}}$ konvergiert gegen $x \in X$.
Also $\lim\limits_{n\to\infty} f(x_n) = f(x)$. Nun ist $(f(x_n))_{n\in\mathbb{N}}$ eine konvergente Folge in Y, also $\lim\limits_{n\to\infty} g(f(x_n)) = g(f(x))$. □

11.3 Das ε-δ-Kriterium für Stetigkeit

Satz §11.4

Eine Abbildung $f\colon X \longrightarrow Y$ zwischen metrischen Räumen ist stetig in $u \in X$ genau dann, wenn die Abbildung das ε-δ-Kriterium erfüllt:

Zu jedem $\varepsilon > 0$ und jedem $x \in X$ gibt es ein $\delta > 0$, so dass aus $d_X(u, x) < \delta$ folgt:

$$d_Y\big(f(u), f(x)\big) < \varepsilon$$

Beweis: Angenommen, f ist stetig und das ε-δ-Kriterium gilt *nicht*: Dann gibt es ein $x \in X$ und ein $\varepsilon > 0$, so dass für jedes $n \in \mathbb{N}$ ein $u_n \in X$ existiert mit $d_X(u_n, x) < \frac{1}{n}$ und $d_Y(f(u_n), f(x)) \geq \varepsilon$. Es folgt also $\lim\limits_{n\to\infty} f(u_n) = f(x)$. Widerspruch.

Angenommen das ε-δ-Kriterium gilt. Sei $\lim\limits_{n\to\infty} x_n = x$. Zu jedem $\varepsilon > 0$ finden wir $\delta > 0$ mit: $d_X(x_n, x) < \delta \Rightarrow d_Y(f(x_n), f(x)) < \varepsilon$. Es gibt $N \in \mathbb{N}$ so, dass aus $k \geq N$ folgt $d_X(x_n, x) < \delta$. Dann folgt also $d_Y(f(x_n), f(x)) < \varepsilon$ für alle $k \geq N$, also $\lim\limits_{n\to\infty} f(x_n) = f(x)$. □

Beispiel: Sei (X, d) ein metrischer Raum, $x_0 \in X$. Setze $f(x) = d(x, x_0)$. Das ist 1-**Lipschitz-stetig**: Beispiel §11.5

$$|f(x) - f(y)| = |d(x, x_0) - d(y, x_0)| \leq d(x, y)$$

Beispiel: $C([a, b], \mathbb{R}) \longrightarrow \mathbb{R}, f \mapsto \int_a^b f(x)\ \mathrm{dx}$ Beispiel §11.6
Das ist stetig, wenn man auf $C([a, b], \mathbb{R})$ die $|| \cdot ||_\infty$-Norm nimmt, denn:

$$\begin{aligned} \Big| \int_a^b f(x)\ \mathrm{dx} - \int_a^b g(x)\ \mathrm{dx} \Big| &= \Big| \int_a^b f(x) - g(x)\ \mathrm{dx} \Big| \leq \int_a^b |f(x) - g(x)|\ \mathrm{dx} \\ &\leq \int_a^b ||f - g||_\infty\ \mathrm{dx} = ||f - g||_\infty (b - a) \end{aligned}$$

also ist die Funktion $f \mapsto \int_a^b f(x)\ \mathrm{dx}$ sogar L-Lipschitz-stetig mit $L = b - a$.

Vorlesung 27.04.2004

Beispiel: $\mathrm{C}([a, b], \mathbb{R}) \longrightarrow \mathrm{C}([a, b], \mathbb{R}), \quad f \mapsto F, \quad F(x) = \int_a^x f(s)\ \mathrm{ds}$
Beide Seiten versehen wir mit der $|| \cdot ||_\infty$-Norm. Dann ist das ebenfalls eine lineare Lipschitz-stetige Abbildung.

11.4 Lineare Abbildungen in normierten Räumen

Satz §11.7

Für eine lineare Abbildung $f \colon V \longrightarrow W$ zwischen normierten Vektorräumen sind die folgenden Aussagen äquivalent:

(i) f ist stetig

(ii) f ist L-Lipschitz-stetig für eine Zahl $L \in \mathbb{R}$

(iii) Es gilt $||f(v)||_W \leq L \cdot ||v||_V$ für ein $L \in \mathbb{R}$ und alle $v \in V$

Beweis: (ii) $\Rightarrow$ (i) ist klar.

(iii) $\Rightarrow$ (ii): Angenommen (iii) gilt. Für $u, v \in V$ gilt dann:

$$L \cdot ||u - v||_V \geq ||f(u) - f(v)||_W = ||f(u - v)||_W$$

Das ist die Lipschitz-Bedingung, d.h. (iii) $\Rightarrow$ (ii)

(i) $\Rightarrow$ (iii): Wir nehmen an (i) gilt. Benutze ε-δ-Kriterium.
Zu $\varepsilon = 1$ und $v = 0 \in V$ gibt es also ein $\delta > 0$ mit:

$$||v - 0||_V < \delta \quad \Rightarrow \quad ||f(v) - f(0)||_W < \varepsilon = 1$$

Ist $v = 0$, so ist $f(v) = 0$, ist $v \neq 0$, so setze $r = \frac{\delta}{2||v||_V}$. Dann ist
$||r - v||_V = |r| \cdot ||v||_V = \frac{\delta}{2||v||_V} ||v||_V = \frac{\delta}{2} < \delta$, also $||f(rv)||_W = |r| \cdot ||f(v)||_W < 1$,
$r \neq 0$ da $v \neq 0$, also: $||f(v)||_W \leq \frac{2}{\delta} ||v||_V = L \cdot ||v||_V$ $\square$

Bemerkung: Dieser Beweis zeigt auch: Eine lineare Abbildung, die in einem Punkt stetig ist, ist *überall* stetig.

Beispiel §11.8

Beispiel: Sei $V = \mathrm{C}^1([a,b],\mathbb{R})$ der Vektorraum der 1-mal stetig differenzierbaren Funktionen und $W = \mathrm{C}([a,b],\mathbb{R})$ der Vektorraum der stetigen Funktionen. Sei f eine lineare Abbildung mit:

$$f \mapsto f' = \frac{\mathrm{d}}{\mathrm{dx}} f$$

Ist das stetig?

1. Versuch: Beide Vektoräume, V und W, mit der $||\cdot||_\infty$-Norm.
$f_n(x) = x^n, f_n'(x) = nx^{n-1}$. Es gilt also: $||f_n'||_\infty = n||f_{n-1}||_\infty$.
Angenommen die Abbildung wäre stetig. Also gibt es $L \in \mathbb{R}$ (nach vorigem Satz) mit $||f'||_\infty \leq L \cdot ||f||_\infty$. Setze $[a,b] = [0,1]$. $||f_n||_\infty = 1, ||f_n'||_\infty = n$. Die Gleichung $L \cdot ||f_n||_\infty = L \geq ||f_n'||_\infty = n$ ergibt einen Widerspruch, da sie nicht für alle $n \in \mathbb{N}$ gilt.

2. Versuch: Gleiche Vektorräume, gleiche Abbildungen, *andere* Normen.

Auf V:	$\|\|f\|\|_V = \|\|f\|\|_\infty + \|\|f'\|\|_\infty$	
Auf W:	$\|\|f\|\|_W = \|\|f\|\|_\infty$	(wie gehabt)

Dann gilt: $||f'||_W = ||f'||_\infty \leq ||f'||_\infty + ||f||_\infty = ||f||_V$. Das Ableiten ist also bezüglich dieser neuen Norm 1-Lipschitz-stetig!

Es ist also ganz wichtig, welche Norm betrachtet wird, wenn man über Stetigkeit spricht.

11.5 Stetigkeit von linearen Abbildungen

Lemma §11.9

Lemma: Ist $(V, ||\cdot||_1)$ ein normierter Vektorraum und $f\colon \mathbb{R}^n \longrightarrow V$ eine lineare Abbildung, dann ist f stetig bezüglich der $||\cdot||_1$-Norm auf $\mathbb{R}^n$.

Beweis: Standardbasis für den $\mathbb{R}^n$:

$$\begin{aligned} e_1 &= (1,0,\ldots,0)\\ e_2 &= (0,1,\ldots,0)\\ &\vdots\\ e_n &= (0,\ldots,0,1) \end{aligned}$$

Für $v \in \mathbb{R}^n, v = (v_1, v_2, \ldots, v_n) = \sum\limits_{k=1}^{n} v_k e_k,\ f(v) = \sum\limits_{k=1}^{n} v_k f(e_k)$ gilt:

$$||f(v)||_V \leq |v_1| \cdot ||f(e_1)||_1 + |v_2| \cdot ||f(e_2)||_1 + \cdots \leq (|v_1| + \cdots + |v_n|) \cdot L$$

mit $L = \max\{||f(e_1)||_1, \ldots, ||f(e_n)||_1\}$, also $||f(v)||_V \leq L \cdot ||v||_1$ □

Lemma §11.10

Lemma: Sei $||\cdot||$ eine beliebige Norm auf $\mathbb{R}^n$. Dann gibt es ein $R > 0$ so, dass $||u|| \geq R$ gilt für alle $u \in \mathbb{R}^n$ mit $||u||_1 = 1$.

Beweis (Widerspruchsbeweis): Wenn das falsch wäre, gäbe es also zu jedem $k \in \mathbb{N}$ einen Vektor $u_k \in \mathbb{R}^n$ mit $||u_k||_1 = 1$ und $||u_k|| < \frac{1}{k}$. Für jede Komponente $u_{i,k}$ gilt $|u_{i,k}| \leq 1$ für alle i, k und $u_k = (u_{1,k}, \ldots, u_{n,k})$.

Nach dem Satz von Bolzano-Weierstraß (Analysis I §3.15) gibt es eine konvergente Teilfolge $(u_{i,k})_{k\in\mathbb{N}}$. Ersetze die ursprüngliche Folge durch diese Teilfolge (umnummerieren der Indizes). Mache das selbe für $i = 2$ (Konvergenz für $i = 1$ bleibt erhalten), und so weiter.

Wir erhalten so eine neue Folge (Teilfolge) $(v_k)_{k\in\mathbb{N}}$ so, dass *jede* Komponentenfolge $(v_{i,k})_{k\in\mathbb{N}}$ gegen v_j konvergiert. Dann konvergiert also die Folge $(v_k)_{k\in\mathbb{N}}$ gegen $(v_1, \ldots, v_n)$ bezüglich der 1-Norm. Wegen $||v_k||_1 = 1$ und $||v_k|| < \frac{1}{k}$ folgt ein Widerspruch, denn $||v||_1 = 1$ und $||v|| = 0$.
Die Abbildung $(\mathbb{R}^n, ||\cdot||_1) \longrightarrow (\mathbb{R}^n, ||\cdot||)$ ist ja stetig. □

Zusammenfassung

Vorlesung 30.04.2004

Lemma 9:	$(\mathbb{R}^n, \|\|\cdot\|\|_1) \xrightarrow{f} (V, \|\|\cdot\|\|), f$ ist linear $\Rightarrow f$ ist stetig
Lemma 10:	$(\mathbb{R}^n, \|\|\cdot\|\|) \Rightarrow \exists\, R > 0.\ \forall\, u \in \mathbb{R}^n.\ \|\|u\|\|_1 = 1 \Rightarrow \|\|u\|\| \geq R$

11.6 Fundamentalsatz über endlichdimensionale normierte Vektorräume

Theorem §11.11

Es seien $(V, ||\cdot||_V)$ und $(W, ||\cdot||_W)$ normierte Vektorräume und $f\colon V \longrightarrow W$ eine lineare Abbildung. Wenn V endliche Dimension hat (z.B. $V = \mathbb{R}^n$), dann ist f (automatisch) stetig.

Beweis: Sei $b_1, \ldots, b_n$ eine Basis für V. Für $v \in V$ haben wir:

$$v = v_1 b_1 + v_2 b_2 + \cdots + v_n b_n, \qquad (v_1, \ldots, v_n) \in \mathbb{R}^n$$

Setze $||v||_1 = |v_1| + \cdots + |v_n|$. Nach Lemma 9 ist $(V, ||\cdot||_1) \xrightarrow{f} (W, ||\cdot||_W)$ stetig. Nach Lemma 10 gibt es ein $R > 0$ so, dass $||u||_V \geq R$ für alle $u \in V$ mit $||u||_1 = 1$ gilt. Für $u \neq 0$ gilt: $||\frac{1}{||u||_1} u||_V = \frac{1}{||u||_1} ||u||_V = 1 \geq R$, also

$$||u||_V \geq R||u||_1 \quad \Leftrightarrow \quad \frac{1}{R}||u||_V \geq ||u||_1$$

□

$$\begin{tikzcd}
(V, ||\cdot||_V) \arrow[dr, "f\ \text{stetig}"] \arrow[dd, "id"', "\text{stetig}"] & \\
& (W, ||\cdot||_W) \\
(V, ||\cdot||_1) \arrow[ur, "f"', "\text{Lipschitz-stetig}"] &
\end{tikzcd}$$

Abbildung 21: Stetigkeit im endlichdimensionalen Vektorraum

Folgerung §11.12

1. Folgerung: Ist V ein endlichdimensionaler Vektorraum, z. B. $\mathbb{R}^n = V$ und sind $||\cdot||_a$ und $||\cdot||_b$ zwei Normen auf V, dann gibt es Zahlen $A, B \in \mathbb{R}$ mit $||u||_a \leq B \cdot ||u||_b$ sowie $||u||_b \leq A \cdot ||u||_a$ für alle $u \in V$. Alle Normen auf $\mathbb{R}^n$ sind also äquivalent.

Folgerung §11.13

2. Folgerung: Alle Normen auf $\mathbb{R}^n$ liefern den gleichen Konvergenzbegriff.

3. Folgerung: *Jeder* endlichdimensionale normierte Vektorraum (z.B. $\mathbb{R}^n$) ist ein **Banachraum**. Folgerung §11.14

Im Unendlichen ist das im Allgemeinen falsch.

11.7 Banach'scher Fixpunktsatz

Theorem §11.15

Es sei (X, d) ein vollständiger metrischer Raum (z.B. $X = \mathbb{R}^n$ oder eine abgeschlossene Teilmenge von $\mathbb{R}^n$) und sei $f\colon X \longrightarrow X$ eine **kontrahierende** Abbildung, d.h. f ist L-Lipschitz-stetig für ein $L < 1$, also $d\big(f(u), f(v)\big) \leq L \cdot d(u, v)$. Dann hat f genau einen **Fixpunkt** $x \in X$, d.h. $f(x) = x$.

Beweis: (a) Es gibt höchstens *einen* Fixpunkt, denn sonst würde gelten: $f(u) = u,\ f(v) = v, u \neq v \Rightarrow d(u, v) = r > 0$, d.h. $d\big(f(u), f(v)\big) = d(u, v) \leq L \cdot d(u, v)$. Ein Widerspruch. Also kann es höchstens einen Fixpunkt geben.

(b) Es gibt (wenigstens) einen Fixpunkt. Wir finden ihn algorithmisch. Wähle $x_0 \in X$ beliebig. Setze $x_1 = f(x_0)$ und rekursiv $x_{n+1} = f(x_n)$.
Behauptung: Die Folge $(x_n)_{n \in \mathbb{N}}$ ist eine Cauchy-Folge.

$$
\begin{aligned}
d(x_2, x_1) &= d\big(f(x_1), f(x_0)\big) \leq L \cdot d(x_1, x_0) \\
d(x_3, x_2) &= d\big(f(x_2), f(x_1)\big) \leq L \cdot d(x_2, x_1) \leq L^2 \cdot d(x_1, x_0) \\
&\vdots \\
d(x_{n+1}, x_n) &= d\big(f(x_n), f(x_{n-1})\big) \leq L^n \cdot d(x_1, x_0)
\end{aligned}
$$

Es gilt also:

$$
\begin{aligned}
d(x_{k+m}, x_k) &\leq d(x_{k+m}, x_{k+m-1}) + d(x_{k+m-1}, x_{k+m-2}) + \cdots \\
&\leq L^k (L^{m-1} + L^{m-2} + \cdots + L + 1) \cdot d(x_1, x_0) \\
&\leq L^k \cdot \frac{1 - L^m}{1 - L} \cdot d(x_1, x_0) \\
&\leq L^k \cdot \underbrace{\frac{1}{1 - L} \cdot d(x_1, x_0)}_{\text{konstant}}
\end{aligned}
$$

Wählt man k groß genug, so gilt: $L^k \cdot \frac{1}{1-L} \cdot d(x_1, x_0) < \frac{\varepsilon}{2}$ für beliebiges $\varepsilon > 0$, also $d(x_{k+m}, x_{k+\ell}) \leq d(x_{k+m}, x_k) + d(x_{k+\ell}, x_k) < \varepsilon$. Also ist die Folge $(x_k)_{k \in \mathbb{N}}$ eine Cauchy-Folge. Setze $x = \lim\limits_{k \to \infty} x_k$. Behauptung: x ist ein Fixpunkt.

$$
f(x) = f(\lim_{k \to \infty} x_k = \lim_{k \to \infty} f(x_k) = \lim_{k \to \infty} x_{k+1} = x
$$

$\square$

Der Beweis liefert ein gutes Näherungsverfahren, um x zu finden. Wähle $x_0 \in X$, setze $x_{k+1} = f(x_k)$. Wir haben die Fehlerabschätzung:

$$
d(x, x_k) \leq L^k \cdot \frac{1}{1 - L} \cdot d(x_0, x_1)
$$

Vorlesung 07.05.2004
Beispiel §11.16

Beispiele: Sei $X = [1, \infty[\subseteq \mathbb{R}$ mit $|u - v| = d(u, v)$ ein metrischer Raum. X ist abgeschlossen in $\mathbb{R}$, also vollständig. Betrachte $f(x) = \frac{x}{2} + \frac{1}{x}$:

$$f(x) = x \quad \Leftrightarrow \quad x = \frac{x}{2} + \frac{1}{x} \quad \Leftrightarrow \quad x^2 = 2 \quad \Leftrightarrow \quad x = \sqrt{2}$$

- Für $x \in X$ gilt $f(x) = x$: es gilt jedenfalls $f(x) > 0$

  $$\frac{x}{2} + \frac{1}{x} \geq 1 \Leftrightarrow \frac{x^2}{2} - x + 1 \geq 0 \Leftrightarrow x^2 - 2x + 2 \geq 0 \Leftrightarrow (x-1)^2 + 1 \geq 0$$

  also $x \in X \implies f(x) \in X$

- Finde L: $|f(u) - f(v)| = |\frac{u-v}{2} + \frac{v-u}{uv}| = |u - v| \cdot \left|\frac{1}{2} - \frac{1}{uv}\right|$
  Behauptung: $\left|\frac{1}{2} - \frac{1}{uv}\right| \leq \frac{1}{2}$
  $\frac{1}{2} - \frac{1}{uv} \leq \frac{1}{2} \Leftrightarrow -\frac{1}{uv} \leq 0 \quad \surd$
  $\frac{1}{uv} - \frac{1}{2} \leq \frac{1}{2} \Leftrightarrow \frac{1}{uv} \leq 1 \quad$ weil $\quad u, v \geq 1 \quad \surd$
  Also gilt $|f(u) - f(v)| \leq \frac{1}{2} \cdot |u - v|$

  Setze $x_0 = 1{,}5 = \frac{3}{2}$, $x_1 = f(\frac{3}{2}) = \frac{3}{4} + \frac{2}{3} = \frac{17}{12}$, $|x_1 - x_0| = \left|\frac{3}{2} - \frac{17}{12}\right| = \frac{1}{12}$.
  Nach n Iterationen haben wir also $d(\sqrt{2}, x_n) \leq \frac{1}{2^n} \cdot 2 \cdot \frac{1}{12} = \frac{1}{2^{n+1}} \cdot \frac{1}{3}$.
  Für $n = 10$ haben wir einen Fehler von $\leq \frac{1}{2^{11}} \cdot \frac{1}{3} = 1{,}63 \cdot 10^{-4}$.

12 Offene Mengen und differenzierbare Abbildungen

Erinnerung: Ein offenes Intervall $J \subseteq \mathbb{R}$ ist von der Form $]a,b[,]-\infty,a[,]a,\infty[$.

Ist $x \in J$ so gibt es ein $r > 0$ mit $]x-r, x+r[\subseteq J$. Wir benutzen diese Eigenschaft als Definition offener Mengen.

Definition §12.1

Definition: Sei (X,d) ein metrischer Raum, $U \subseteq X$. Wir sagen U ist **offen** wenn es für jedes $u \in U$ ein $r > 0$ gibt, so dass $B_r(u) \subseteq U$ (Erinnerung: $B_r(u) = \{v \in X \mid d(u,v) < r\}$). Offene Intervalle in $\mathbb{R}$ haben offensichtlich diese Eigenschaft.
Beachte: Die **leere Menge** $\emptyset$ ist offen, ebenso ist X offen.

12.1 Offene und abgeschlossene Mengen

Satz §12.2

Satz: Sei (X,d) ein metrischer Raum. Eine Menge $U \subseteq X$ ist offen genau dann, wenn ihr Komplement $X \setminus U = \{x \in X \mid x \notin U\}$ abgeschlossen ist. Eine Menge $A \subseteq X$ ist abgeschlossen genau dann, wenn ihr Komplement $X \setminus A$ offen ist.

Beweis: Vorüberlegung: Wenn gilt $\lim\limits_{n\to\infty} x_n = x$, dann liegt in jeder $B_r(x)$ mindestens ein x_n.
Angenommen, $U \subseteq X$ ist offen, setze $A = X \setminus U$. Falls $A = \emptyset$ ist A abgeschlossen. Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in A, die gegen $x \in X$ konvergiert. Zu zeigen: $x \in A$. Wenn $x \notin A$, dann $x \in U$, also gibt es ein $r > 0$ mit $B_r(x) \subseteq U$, das heißt $B_r(x) \cap A = \emptyset$. Es gibt aber ein n mit $a_n \in B_r(x)$, also $B_r(x) \cap A \neq \emptyset$. Das ist ein Widerspruch. Also ist A abgeschlossen.

$$(U \text{ offen} \implies X \setminus U \text{ abgeschlossen})$$

Sei nun $A \subseteq X$ abgeschlossen, sei $U = X \setminus A$. Falls $U = \emptyset$, ok. Sei $u \in U$. Wenn es zu jedem $k \geq 1$ ein $a_k \in A$ gäbe mit $a_k \in B_{\frac{1}{k}}(u)$, so $\lim\limits_{k\to\infty} a_k = u \in A$. Das ist ein Widerspruch. Also gibt es ein k mit $B_{\frac{1}{k}}(u) \cap A = \emptyset$, das heißt $B_{\frac{1}{k}}(u) \subseteq U$, U offen.

$$(A \text{ abgeschlossen} \implies U = X \setminus A \text{ offen})$$

$$\begin{array}{lcccc} X \setminus (X \setminus U) = U & \rightsquigarrow & X \setminus U \text{ abges.} & \implies & U \text{ offen} \\ X \setminus (X \setminus A) = A & \rightsquigarrow & X \setminus A \text{ offen} & \implies & A \text{ abges.} \end{array}$$

Achtung: Es gibt viele Mengen, die weder offen noch abgeschlossen sind. Beispiel: $Y = [0,1[\subseteq \mathbb{R}$ ist nicht abgeschlossen. Y ist auch nicht offen. Es gibt Mengen, die sowohl offen als auch abgeschlossen sind (*abgeschloffen*), zum Beispiel sind $\emptyset$ und X immer abgeschlossen und offen.

12.2 Durchschnitte und Vereinigungen offener Mengen

Satz §12.3

Satz: Endliche Durchschnitte und beliebige Vereinigung offener Mengen sind offen.

Beweis: Benutze §10.10. Sei $U_1, \ldots, U_n$ offen.

$$X \setminus (U_1 \cap \ldots \cap U_n) = \underbrace{X \setminus U_1}_{\text{abges.}} \cup \ldots \cup \underbrace{X \setminus U_n}_{\text{abges.}}$$

Das ist also abgeschlossen nach §10.10, also ist $U_1 \cap \ldots \cap U_n$ offen. Sei $\{U_i | i \in I\}$ eine Menge von offenen Mengen. Behauptung: $\bigcup\limits_{i \in I} U_i$ ist offen.

$X \setminus (\bigcup\limits_{i \in I} U_i) = \bigcap\limits_{i \in I} (X \setminus U_i)$ ist abgeschlossen, also ist $\bigcup\limits_{i \in I} U_i$ offen. □

12.3 Kurven

Wir wollen differenzierbare Abbildungen zwischen Banach-Räumen (zum Beispiel $\mathbb{R}^n$) studieren und fangen an mit **Kurven**.

Definition §12.4

Definition: Sei $J \subseteq \mathbb{R}$ ein Intervall (offen, halboffen, oder abgeschlossen) und $(V, ||\cdot||)$ ein Banachraum. Eine stetige Abbildung $c\colon J \longrightarrow V$ heißt **stetige Kurve** in V. $t \mapsto c(t) \in V$.

Beispiele:

- $V = \mathbb{R}^2,\ J = \mathbb{R},\ c_1(t) = (\cos t, \sin t) \in \mathbb{R}^2$
- $V = \mathbb{R}^3,\ J = \mathbb{R},\ c_2(t) = (\cos t, \sin t, t) \in \mathbb{R}^3$

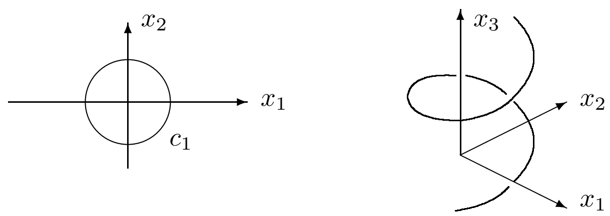


Abbildung 22: Funktionsgraphen von c_1 und c_2

Satz §12.5

Satz: Ist $V = \mathbb{R}^n$, $c\colon J \longrightarrow \mathbb{R}^n$ eine Abbildung, dann wird c beschrieben durch die Koordinaten/Komponenten $c(t) = (c_1(t), c_2(t), \ldots, c_n(t))$. c ist genau dann stetig, wenn alle einzelnen c_k stetig sind.

Beweis: Sei $e_k = (0, \ldots, 1, \ldots, 0) \in \mathbb{R}^n$ an der k-ten Stelle 1.
Daraus folgt $c(t) = c_1(t) \cdot e_1 + c_2(t) \cdot e_2 + \ldots + c_n(t) \cdot e_n$. Sei $(t_m)_{m \in \mathbb{N}}$ eine Folge in J, $\lim\limits_{m \to \infty} t_m = t$, alle c_k seien stetige Funktionen. Wähle $\varepsilon > 0$ so, dass $|c_k(t_m) - c_k(t)| < \varepsilon$ für alle k, alle $m \geq N \Longrightarrow ||c(t_k) - c(t)||_1 \leq n \cdot \varepsilon$, also $\lim\limits_{m \to \infty} c(t_m) = c(t)$, weil die $||\cdot||_1$-Norm die gleiche Konvergenz liefert, wie $||\cdot||$.
Also: Alle c_k stetig, also ist c stetig. $c_k = P_k \cdot c$, $P_k(v_1, \ldots, v_k) = v_k$ ist stetig.
Also c stetig, alle c_k stetig. □

12.4 Differenzierbare Kurven

Vorlesung 11.05.2004

Definition §12.6

Eine auf einem offenen Intervall $I \subseteq \mathbb{R}$ definierte stetige Kurve c mit Werten in einem normierten Vektorraum V heißt **differenzierbar** in $t_0 \in I$, falls die Abbildung $p\colon h \mapsto \frac{1}{h}\big(c(t_0+h) - c(t_0)\big)$ eine stetige Fortsetzung in $h = 0$ hat, falls also der Grenzwert

$$\dot{c}(t_0) = p(0) = \lim_{h\to 0} \frac{c(t_0+h) - c(t_0)}{h}$$

existiert. Der Vektor $\dot{c}(t_0) \in V$ heißt **Geschwindigkeitsvektor** oder **Tangentialvektor** an c im Punkt t_0 und $||\dot{c}(t_0)||$ ist die **Geschwindigkeit** von c in t_0.

Falls c in jedem $t \in I$ differenzierbar ist, heißt c **differenzierbar** und falls die Abbildung $\dot{c}\colon I \longrightarrow V,\ t \mapsto \dot{c}(t)$ stetig ist, heißt c **stetig differenzierbar**.

Satz §12.7

Satz: Ist $V = \mathbb{R}^n$ (mit beliebiger Norm), dann ist $c\colon I \longrightarrow V$ genau dann differenzierbar in $t_0 \in I$, wenn es die Komponenten $c_1, c_2, \ldots, c_n$ sind. Entsprechendes gilt für (stetige) Differenzierbarkeit.

Beweis: $p(h) = \frac{1}{h}\big(c(t_0+h) - c(t_0)\big) = \big(p_1(h), p_2(h), \ldots, p_n(h)\big)$ Die Behauptung folgt aus Satz 5. □

Bemerkung: Dann gilt:

$$c'(t) = p(0) = \big(c_1'(t), c_2'(t), \ldots, c_n'(t)\big)$$

Beispiel: Die Kurven aus den Beispielen in §12.4 sind stetig differenzierbar:

$$\begin{aligned} c(t) &= (\cos t, \sin t, t) \\ c'(t) &= (-\sin t, \cos t, 1) \\ ||c'(t)||_2 &= \sqrt{(-\sin t)^2 + (\cos t)^2 + 1^2} = \sqrt{1+1} = \sqrt{2} \end{aligned}$$

Bemerkung §12.8

Bemerkung: Entsprechend definiert man **k-mal (stetig) differenzierbare** Kurven. Wenn c zweimal stetig differenzierbar ist, heißt die 2. Ableitung $\ddot{c}(t)$ die **Beschleunigung** im Punkt t.

Vereinbarung §12.9

Vereinbarung: Ist $I = [a,b]$ ein abgeschlossenes Intervall, $c\colon I \longrightarrow V$ eine Kurve, so heißt c (stetig) differenzierbar, wenn es ein $r > 0$ gibt und eine (stetig) differenzierbare Kurve $e\colon]a-r, b+r[\longrightarrow V$ mit $e(t) = c(t)$ für alle $t \in I$.

12.5 Verknüpfungen von differenzierbaren Kurven

Satz §12.10

Satz: Seien $I, J \subseteq \mathbb{R}$ Intervalle, $f\colon I \longrightarrow J$ eine (stetig) differenzierbare Abbildung, $c\colon J \longrightarrow V$ eine (stetig) differenzierbare Kurve. Dann ist die Hintereinanderausführung $e = c \circ f\colon I \longrightarrow V,\ t \mapsto c\big(f(t)\big)$ eine (stetig) differenzierbare Kurve und $\dot{e}(t) = f'(t) \cdot \dot{c}\big(f(t)\big)$

Beweis: Schreibe $s_0 = f(t_0)$.

$$p(h) = \tfrac{1}{h}\big(c(s_0 + h) - c(s_0)\big) \quad p(0) = \dot{c}(s_0)$$
$$q(k) = \tfrac{1}{k}\big(f(t_0 + k) - f(t_0)\big) \quad q(0) = f'(t_0)$$

Dann ist:

$$\begin{aligned} \frac{e(t_0+k) - e(t_0)}{k} &= \frac{c\big(f(t_0+k)\big) - c\big(f(t_0)\big)}{k} \\ &= \frac{p\big(f(t_0+k) - f(t_0)\big) \cdot \big(f(t_0+k) - f(t_0)\big)}{k} \\ &= p\big(f(t_0+k) - f(t_0)\big) \cdot q(k) \end{aligned}$$

Für $k \to 0$ ergibt sich: $p(0) \cdot q(0) = f'(t_0) \cdot \dot{c}(t_0)$ □

Satz: Ist $f\colon I \longrightarrow \mathbb{R}$ (stetig) differenzierbar und $c\colon J \longrightarrow V$ eine (stetig) differenzierbare Kurve, so ist $e\colon I \longrightarrow V,\ t \mapsto f(t) \cdot c(t)$ eine (stetig) differenzierbare Kurve mit der Ableitung: Satz §12.11

$$\dot{e}(t) = f'(t) \cdot c(t) + f(t) \cdot \dot{c}(t)$$

Beweis: Gleiche Idee wie bei der Leibniz-Regel (siehe Analysis I §7.6) □

12.6 Kurvenlänge

Definition §12.12

Die **Länge** einer stetig differenzierbaren Kurve $c\colon [a, b] \longrightarrow V$ ist:

$$L(c) = \int_a^b ||\dot{c}(t)||\ \mathrm{dt}$$

Beispiel: Seien $u, v \in V$ und $c\colon [0, 1] \longrightarrow V,\ t \mapsto (1-t)u + tv = u + t(v-u)$ eine Kurve, so gilt:

$$\begin{aligned} \dot{c}(t) &= \lim_{h\to 0} \frac{c(t+h) - c(t)}{h} \\ &= \lim_{h\to 0} \frac{u + (t+h)(v-u) - u - t(v-u)}{h} = v - u \\ L(c) &= \int_0^1 ||\dot{c}(t)||\ \mathrm{dt} = \int_0^1 ||v-u||\ \mathrm{dt} = ||v-u||(1-0) = d(u, v) \end{aligned}$$

Für gerade Strecken erhält man also die gewöhnliche Streckenlänge.

12.7 Umparameterisierung

Bemerkung §12.13

Ist $f\colon I \longrightarrow J$ eine streng monotone bijektive stetig differenzierbare Abbildung und $c\colon J \longrightarrow V$ eine Kurve, so heißt $e = c \circ f\colon I \longrightarrow V,\ t \mapsto c\big(f(t)\big)$ die zu f gehörige **Umparameterisierung** von c.

Man kann zeigen:

$$L(e) = L(c)$$

Beweisidee:

$$\begin{aligned} L(e) &= \int_a^b ||\dot{c}(t)||\,\mathrm{dt} = \int_a^b ||f'(t)\cdot\dot{c}(f(t))||\,\mathrm{dt} = \int_a^b |f'(t)|\cdot||\dot{c}(f(t))||\,\mathrm{dt} \\ &= \int_{f(a)}^{f(b)} ||\dot{c}(s)||\,\mathrm{ds} = L(c) \quad \text{falls } f' \geq 0 \end{aligned}$$

Falls $f' \leq 0$ dreht sich das Vorzeichen um. □

12.8 Reelle Funktionen

Kurven sind Abbildungen von $\mathbb{R}$ in Vektorräume, jetzt betrachten wir Abbildungen von Vektorräumen nach $\mathbb{R}$.

Es sei V ein Banachraum und $U \subseteq V$ eine offene Menge. Zu jedem $u \in U$ gibt es also ein $r > 0$ so, dass U die offene r-Kugel $B_r(u) = \{v \in V \mid ||u-v|| < r\}$ enthält.

Wir interessieren uns für stetige Abbildungen $f\colon U \longrightarrow \mathbb{R}$. Ist $V = \mathbb{R}^n$, dann ist $f(v) = f(v_1, v_2, \ldots, v_n)$. Um die Stetigkeit von f zu zeigen, genügt es *nicht* die Stetigkeit der Funktionen $\mathbb{R} \ni t \mapsto f(v_1, v_2, \ldots, v_{j-1}, t, v_{j+1}, \ldots, v_n)$ zu fordern.

Vorlesung 14.05.2004

Erinnerung an Analysis I: Wir hatten $p(h) = \begin{cases} \frac{f(t_0+h)-f(t_0)}{h} & \text{falls } h \neq 0 \\ f'(t_0) & \text{falls } h = 0 \end{cases}$

also $f(t_0+h) - f(t_0) = h \cdot p(h) = p(0)\cdot h + h\cdot(p(h)-p(0)) = f'(t_0)\cdot h + |h|\cdot\lambda(h)$

mit $\lambda(h) = \begin{cases} p(h)-p(0) & \text{für } h > 0 \\ 0 & \text{für } h = 0 \\ p(0)-p(h) & \text{für } h < 0 \end{cases} \quad \Longrightarrow \quad \lambda$ ist stetig!

12.9 Differenzierbare reelle Funktionen

Definition §12.14

Definition: Es sei U eine offene Menge in einem normierten Vektorraum V, $f\colon U \longrightarrow \mathbb{R}$ stetig und $u_0 \in U$. Wähle $r > 0$ so, dass $B_r(u_0) \subseteq U$. Wir sagen f ist **differenzierbar im Punkt** u_0, falls es eine stetige lineare Abbildung $A\colon V \longrightarrow \mathbb{R}$ gibt, und eine stetige Funktion $\lambda\colon B_r(0) \longrightarrow \mathbb{R}$ mit $\lambda(0) = 0$ so, dass für alle Vektoren $h \in V$ mit $||h|| < r$ gilt:

$$f(u_0+h) - f(u_0) = A(h) + ||h||\cdot\lambda(h)$$

Wenn U ein offenes Intervall in $\mathbb{R}$ ist, ist das genau die Definition aus Ana I.

Bemerkung: Wegen $\lambda(h) = \frac{f(u_0+h)-f(u_0)-A(h)}{||h||}$ für $h \neq 0$ ist λ eindeutig durch A bestimmt.

Die lineare Abbildung $A\colon V \longrightarrow \mathbb{R}$ ist in dieser Situation ebenfalls eindeutig bestimmt, und wir nennen $df(u_0) := A\colon V \longrightarrow \mathbb{R}$ das **Differential** (oder die Ableitung) von f im Punkt u_0.

Paragraph §12.15

Warum ist A eindeutig bestimmt?

$$\begin{aligned}
A(h) &= f(u_0+h) - f(u_0) - ||h|| \cdot \lambda(h) \\
\tilde{A}(h) &= f(u_0+h) - f(u_0) - ||h|| \cdot \tilde{\lambda}(h) \\
A(h) - \tilde{A}(h) &= ||h||(\tilde{\lambda}(h) - \lambda(h)) \quad \text{für } 0 < t \leq 1 \\
A(th) - \tilde{A}(th) &= ||th||(\tilde{\lambda}(th) - \lambda(th)) \\
tA(h) - t\tilde{A}(h) &= t \cdot ||h||(\tilde{\lambda}(th) - \lambda(th)) \\
\underbrace{A(h) - \tilde{A}(h)}_{\text{unabh. von } t} &= ||h|| \underbrace{(\tilde{\lambda}(th) - \lambda(th))}_{\text{gegen 0 für } t \to 0}
\end{aligned}$$

also $A(h) = \tilde{A}(h)$ für alle h mit $||h|| < r$, damit folgt $A = \tilde{A}$ □

Die Ableitung ordnet also einem Punkt $u \in U$ eine lineare Abbildung

$$df(u)\colon V \longrightarrow \mathbb{R}$$

zu. Eine Funktion, die in jedem Punkt u differenzierbar ist, heißt eine **differenzierbare Abbildung**. Um zu sagen, was es heißt, dass eine Funktion stetig differenzierbar ist, brauchen wir eine Norm auf der Menge aller stetigen linearen Abbildungen von V nach $\mathbb{R}$.

12.10 Dualraum

Definition: Es sei $(V, ||\cdot||_V)$ ein normierter Vektorraum und Definition §12.16

$$V^* := \{\varphi\colon V \longrightarrow \mathbb{R} \mid \varphi \text{ ist eine stetige lineare Abbildung}\}$$

die Menge aller stetigen, linearen Abbildungen von V nach $\mathbb{R}$ (manchmal sagt man auch **Linearformen**). V^* heißt **Dualraum** oder **Dual** von V und ist ein Vektorraum: Für $\varphi, \psi \in V^*$, $c \in \mathbb{R}$ setze $\varphi + \psi = [u \mapsto \varphi(u) + \psi(u)]$, $c\varphi = [u \mapsto c\varphi(u)]$. Dann sind $\varphi + \psi$, $c\varphi$ wieder stetige lineare Abbildungen. Für $\varphi \in V^*$ setzen wir $||\varphi|| = \sup\{||\varphi(u)|| \mid u \in V, ||u|| < 1\}$.

Satz: $(V^*, ||\cdot||)$ ist ein normierter Vektorraum. Satz §12.17

Beweis: Ist $\varphi \in V^*$ und $\varphi \neq 0$ (das heißt φ ist nicht die Null-Abbildung), so gibt es $v \in V$ mit $\varphi(v) \neq 0$ und es folgt:

$$||\varphi|| \geq \left|\varphi\left(\frac{1}{||v||_V} \cdot v\right)\right| = \frac{1}{||v||_V} \cdot |\varphi(v)| > 0 \quad \text{mit} \quad |\varphi(v)| \neq 0$$

Aus $|\varphi(v)+\psi(v)| \leq |\varphi(v)|+|\psi(v)|$ folgt $||\varphi+\psi|| \leq ||\varphi||+||\psi||$ und aus $|c\varphi(v)| = |c| \cdot |\varphi(v)|$ folgt $||c\varphi|| = |c| \cdot ||\varphi||$, also ist $||\cdot||\colon V^* \longrightarrow \mathbb{R}$ eine Norm. □

Bemerkungen

Bemerkung §12.18

(a) Es gilt: $|\varphi(v)| \leq ||\varphi|| \cdot ||v||_V$, $||\varphi||$ ist die kleinstmögliche Lipschitz-Konstante für φ

(b) V^* ist sogar ein Banachraum: Für eine Cauchy-Folge $(\varphi_n)_{n\in\mathbb{N}}$ in V^*, setze $\varphi(v) = \lim\limits_{n\to\infty} \varphi_n(v)$; das macht Sinn, weil $\varphi_n(v)$ eine Cauchy-Folge in $\mathbb{R}$ ist.

Wenn f differenzierbar ist, und die Abbildung $df\colon U \longrightarrow V^*, u \mapsto df(u)$ stetig ist (bezüglich der gerade definierten Norm), dann heißt f **stetig differenzierbar**. Paragraph §12.19

12.11 Erweiterung der Kettenregel

Satz §12.20

Satz (Kettenregel von $\mathbb{R} \longrightarrow V \longrightarrow \mathbb{R}$): Ist $c\colon I \longrightarrow V$ eine stetig differenzierbare Kurve und ist $c(I) \subseteq U$, und ist weiter $f\colon U \longrightarrow \mathbb{R}$ stetig differenzierbar, dann ist $g := f \circ c\colon I \longrightarrow \mathbb{R}, t \mapsto f(c(t))$ stetig differenzierbar und es gilt $g'(t) = df(c(t))(\dot{c}(t))$.

Beweis: $p(h) = \frac{1}{h}(c(t_0 + h) - c(t_0))$

$$\begin{aligned}
& \frac{1}{h}\Big(f\big(c(t_0+h)\big) - f\big(c(t_0)\big)\Big) \\
= \; & \frac{1}{h}\Big(f\big(c(t_0) + \big(c(t_0+h) - c(t_0)\big)\big) - f\big(c(t_0)\big)\Big) \\
= \; & \frac{1}{h}\big(df\big(c(t_0)\big)\big(c(t_0+h) - c(t_0)\big) + ||c(t_0+h) - c(t_0)||\lambda\big(c(t_0+h) - c(t_0)\big)\big) \\
= \; & df\big(c(t_0)\big)\big(p(h)\big) \pm ||p(h)|| \cdot \lambda\big(c(t_0+h) - c(t_0)\big)
\end{aligned}$$

Das ist stetig in $h = 0$ und ergibt dort $g'(t_0) = df\big(c(t_0)\big)\big(\dot{c}(t_0)\big) \pm \big[||\dot{c}(t_0)|| \cdot \lambda(0)\big]$

12.12 Richtungsableitung

Definition §12.21

Definition: Speziell sei $0 \in I$, $u_0 \in U \subseteq V$, $v \in V$ und $c\colon I \longrightarrow U, t \mapsto u_0 + tv$. Dann $(f \circ c)'(0) = df\big(c(0)\big)\big(\dot{c}(0)\big) = df(u_0)(v)$.

Für $v \in V$ heißt die Zahl $df(u_0)(v) =: D_v f(u_0)$ auch **Richtungsableitung** von f in Richtung v. Speziell für $V = \mathbb{R}^n$ und $v = e_j = (0, \ldots, 1, \ldots, 0)$ an der j-ten Stelle 1 schreibt man

$$\frac{\partial f}{\partial x_j}(u_0) = df(u_0)(e_j) = D_{e_j} f(u_0)$$

Die Richtungsableitung einer Funktion $f\colon U \longrightarrow \mathbb{R}$, $U \subseteq R^n$ in Koordinatenrichtungen erhält man also, indem man die Variablen $x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_n$ fest läßt und nur die j-te Koordinate variiert, man leitet also nach der j-ten Variable ab und betrachtet die anderen Variablen als Konstanten:

$g(x) = f(x_1, \ldots, x_{j-1}, x, x_{j+1}, \ldots, x_n)$
$g'(x_j) = \frac{\partial f}{\partial x_j}(x)$

Vorlesung 18.05.2004

Beispiel: Sei $f\colon \mathbb{R}^3 \longrightarrow \mathbb{R}$ eine Funktion mit $f(x_1, x_2, x_3) = x_1^2 + x_2^2 - x_2 x_3$. Es ergeben sich die folgenden Richtungsableitungen:

$\frac{\partial f}{\partial x_1}(v_1, v_2, v_3) = 2v_1$
$\frac{\partial f}{\partial x_2}(v_1, v_2, v_3) = 2v_2 - v_3$
$\frac{\partial f}{\partial x_3}(v_1, v_2, v_3) = -v_2$

Bemerkung §12.22

Bemerkung: Mit den Richtungsableitungen kann man das Differential ganz konkret angeben:

$h_k(t) = u_0 + t(0, \ldots, 0, 1, 0, \ldots, 0)$ an der k-ten Stelle 1
$\dot{h}_k(t) = (0, \ldots, 0, 1, 0, \ldots, 0)$ an der k-ten Stelle 1
$df(u_0)\big(\dot{h}_k(0)\big) = \frac{\partial f}{\partial x_k}(u_0)$

Also gilt für $h = (h_1, h_2, \ldots, h_n)$:

$$df(u_0)(h) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(u_0)h_k$$

Weil h Linearkombination der Vektoren $(0, \ldots, 0, 1, 0, \ldots, 0)$ ist, setzt man den **Gradienten** ein:

$$\nabla f(u_0) = \Big(\frac{\partial f}{\partial x_1}(u_o), \frac{\partial f}{\partial x_2}(u_o), \ldots, \frac{\partial f}{\partial x_n}(u_o)\Big) \qquad \textit{(sprich: Nabla } f \textit{ von } u_0\textit{)}$$

Mit dieser Bezeichnung und Notation gilt für $h \in \mathbb{R}^n$, $h = (h_1, h_2, \ldots, h_n)$:

$$df(u_0)(h) = \sum_{k=1}^{n} \frac{\partial f}{\partial x_k}(u_0)h_k = \langle\, \nabla f(u_0) \mid h \,\rangle$$

Beispiel: Sei $V = U = \mathbb{R}^2$ und $f(x, y) = x^2 - 2xy$.

$\frac{\partial f}{\partial x} = 2x - 2y \qquad \frac{\partial f}{\partial y} = -2x$

$\nabla f(u_1, u_2) = (2u_1 - 2u_2, -2u_1)$

$df(u_1, u_2)(h_1, h_2) = (2u_1, -2u_2)h_1 - 2u_1h_2 = \langle\, \nabla f(u_1, u_2) \mid (h_1, h_2) \,\rangle$

Bemerkung: Ist $f\colon V \longrightarrow \mathbb{R}$ linear und stetig ($f \in V^*$), so ist f stetig differenzierbar und $df(u_0)(h) = f(h)$, denn $f(u_0 + h) - f(u_0) = f(h) = df(u_0)(h)$. Bemerkung §12.23

Satz: Sei $U \subseteq \mathbb{R}^n$ offen, $f\colon U \longrightarrow \mathbb{R}$ stetig, dann ist f stetig differenzierbar genau dann, wenn alle Ableitungen $\frac{\partial f}{\partial x_k}$ existieren und stetige Funktionen auf U sind. Satz §12.24

Beweis: Wenn f stetig differenzierbar ist, so existieren die partiellen Ableitungen $\frac{\partial f}{\partial x_k}$ und sind stetig (vgl: §12.22). Für die umgekehrte Implikation betrachte: $g(u) = f(u) - \langle\, \nabla f(u) \mid u - u_0 \,\rangle$. Dann hat g die gleichen Differenzierbarkeitseigenschaften wie f und $\nabla g(u_0) = \nabla f(u_0)$. Es gilt nun für $u = (u_1, u_2, \ldots, u_n)$, $u_0 = (u_{1,0}, u_{2,0}, \ldots, u_{n,0})$:

$$\begin{aligned} g(u) - g(u_0) = \quad & \textstyle\sum_{k=1}^{n} \quad g(u_1, u_2, \ldots, u_{k-1}, u_k \quad, u_{k+1,0}, \ldots, u_{n,0}) \\ & \quad - \quad\;\; g(u_1, u_2, \ldots, u_{k-1}, u_{k,0}, u_{k+1,0}, \ldots, u_{n,0}) \end{aligned}$$

Da die Funktionen $\frac{\partial g}{\partial x_k}$ stetig sind und $\frac{\partial g}{\partial x_k}(u_0) = 0$ gilt, gibt es zu jedem $\varepsilon > 0$ ein $\delta > 0$ mit $|\frac{\partial g}{\partial x_k}(u)| < \varepsilon$ falls $||u - u_0|| < \delta$. Mit dem Mittelwertsatz der Differentialrechnung (Analysis I, §7.16) folgt:

$$|g(u) - g(u_0)| \leq \sum_{k=1}^{n} ||u - u_0||\varepsilon = n||u - u_o||\varepsilon$$

Das heißt $dg(u_0) = 0$, die Funktion g ist in u_0 differenzierbar. Man kann leicht die Stetigkeit nachweisen. $\square$

Beispiele

Beispiel §12.25

- Sei $V = \mathbb{R}^n$ und $f(x) = ||x||_2^2 = x_1^2 + x_2^2 + \cdots + x_n^2$.
  $\frac{\partial f}{\partial x_k} = 2x_k \qquad \nabla f(x) = 2x$
  (Also ist f stetig differenzierbar.)
- Sei $V = \mathbb{R}^3$ und $f(x, y, z) = 2xy - y\cos(z)$.
  $\nabla f(x, y, z) = \big(2y, 2x - \cos(z), y\sin(z)\big)$

12.13 Vektorraum der stetig differenzierbaren Funktionen

Bemerkung §12.26

Ist V ein Banachraum und $U \subseteq V$ eine offene Menge, setze:

$$\mathrm{C}^1(U, \mathbb{R}) = \{f\colon U \longrightarrow \mathbb{R} \mid \ f \text{ ist stetig differenzierbar}\}$$

Für beliebige $f, g \in \mathrm{C}^1(U, \mathbb{R})$ und $c \in \mathbb{R}$ setzt man $(f + g)(u) = f(u) + g(u)$ und $(cf)(u) = cf(u)$, damit gilt: $d(f + g) = df + dg$ sowie $d(cf) = c \cdot df$

Mit $(f \cdot g)(u) = f(u) \cdot g(u)$ ist $(fg) \in \mathrm{C}^1(U, \mathbb{R})$ und $d(fg) = g \cdot df + f \cdot dg$ (Leibnizregel).

13 Ableitungen vektorwertiger Funktionen

Wir hatten betrachtet:

Kurven	$J \xrightarrow{c} V$	Geschwindigkeit
Reelle Funktionen	$V \xrightarrow{f} \mathbb{R}$	Differential

13.1 Die Operatornorm

Definition §13.1

Seien V, W Banachräume, $L(V,W) = \{f\colon V \longrightarrow W \mid f \text{ ist stetig und linear}\}$, zum Beispiel $V = \mathbb{R}^n$, $W = \mathbb{R}^m \rightsquigarrow L(\mathbb{R}^n, \mathbb{R}^m)$ entspricht eindeutig $\mathbb{R}^{n\times m}$, also der Menge der $\{n \times m$-Matrizen$\}$. Offensichtlich ist $L(V,W)$ ein Vektorraum, wenn man setzt:

$$\begin{array}{lll} (f+g)(v) = f(v) + g(v) & \text{für} & f, g \in L(V,W) \\ (cf)(v) = cf(v) & \text{für} & f \in L(V,W), r \in \mathbb{R} \end{array}$$

Für $f \in L(V,W)$ setze

$$\sup\{||f(v)||_W \bigm| ||v||_V \leq 1\} = ||f||$$

Diese Norm heißt **Operatornorm** auf $L(V,W)$.

Satz §13.2

Satz: Sind (V,W) Banachräume (zum Beispiel $V = \mathbb{R}^n$, $W = \mathbb{R}^m$), dann ist $L(V,W)$ mit der Operatornorm ein Banachraum.

Beweis: Warum ist das eine Norm? Ist $f \neq 0$, so gibt es $v \in V$ mit $f(v) = w \neq 0$, also

$$\frac{1}{||v||_V} \cdot f(v) = f\underbrace{\left(\frac{1}{||v||_V} \cdot v\right)}_{\text{Einheitsvektor}} \neq 0 \quad \Longrightarrow \quad ||f|| \neq 0$$

Dreiecksungleichung für die Operatornorm:

$$||f(v) + g(v)||_W \leq ||f(v)||_W + ||g(v)||_W$$

Das bleibt im Supremum richtig, also $||f + g|| \leq ||f|| + ||g||$.

Zur Vollständigkeit: Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $L(V,W)$. Für $v \in V$ ist dann $(f_n(v))_{n\in\mathbb{N}}$ eine Cauchy-Folge in W, definiere $f(v) = \lim\limits_{n\to\infty} f_n(v)$.

Bemerkung §13.3

Bemerkung: Sind X, Y, Z Banachräume, $X \xrightarrow{f} Y \xrightarrow{g} Z$ stetige lineare Abbildungen, so gilt $||g \circ f|| \leq ||g|| \cdot ||f||$ (submultiplikativ).

13.2 Differenzierbarkeit

Definition §13.4

Definition: Seien V, W Banachräume (zum Beispiel $V = \mathbb{R}^n$, $W = \mathbb{R}^m$), sei $U \subseteq V$ offen, sei $f\colon U \longrightarrow W$ stetig. Sei $u \in U$, sei $r > 0$ mit $B_r(u) \subseteq U$. Wir sagen f ist **differenzierbar in** u, falls es eine stetige Abbildung $\lambda\colon B_r(0) \longrightarrow W$ gibt mit $\lambda(0) = 0$ *und* $A \in L(V,W)$ mit

$$f(u+h) - f(u) = A(h) + ||h|| \cdot \lambda(h) \quad \text{für alle } h \in B_r(0)$$

Falls f in jedem $u \in U$ differenzierbar ist, heißt f **differenzierbar**. Die definierte Gleichung zeigt, dass die lineare Abbildung $A \in L(V, W)$ durch f eindeutig bestimmt ist, wir schreiben $Df(u) = A$, also $Df(u) \in L(V, W)$:

$$f(u + h) - f(u) = Df(u)(h) + ||h|| \cdot \lambda(h)$$

Falls die Abbildung $U \longrightarrow L(V, W), u \mapsto Df(u)$ stetig ist, heißt f **stetig differenzierbar**.

Beispiel §13.5

Bemerkung: Diese neue Definition passt zu den alten, bekannten Differenzierbarkeitsbegriffen:

- Analysis I: $f\colon U \longrightarrow \mathbb{R}$, $U \subseteq \mathbb{R}$ offen, $f'(u) \cdot h = Df(u)(h)$ mit $u \in U \subseteq \mathbb{R}$, $h \in \mathbb{R}$
- Kurven: $c\colon J \longrightarrow V$, $J \subseteq \mathbb{R}$ offenes Intervall, $\dot{c}(t) \in V$, $Dc(t)(h) = \dot{c}(t) \cdot h$, $h \in \mathbb{R}$, $t \in J$
- Reelle Funktionen: $U \subseteq V$ offen, $f\colon U \longrightarrow \mathbb{R}$, $df(u)\colon V \longrightarrow \mathbb{R}$ linear, $Df(u)(h) = df(u)(h)$

13.3 Beispiele

Beispiel: Ist $f\colon V \longrightarrow W$ stetig und linear, dann ist f stetig differenzierbar:

$$f(u + h) - f(u) = f(h) \quad \text{also} \quad Df(u)(h) = f(h)$$

Paragraph §13.6

Beispiel: Ist $V = \mathbb{R}^n$, $W = \mathbb{R}^m$, $f\colon V \longrightarrow W$.
$f(v) = f(v_1, v_2, \ldots, v_n) = (f_1(v_1, \ldots, v_n), f_2(v_1, \ldots, v_n), \ldots, f_m(v_1, \ldots, v_n))$
Dann ist f (stetig) differenzierbar genau dann, wenn es die Komponenten

$$f_1,\ f_2,\ \ldots,\ f_m$$

sind. Setze $f_k\colon V \longrightarrow \mathbb{R}$, $k = 1, 2, \ldots, m$. Für die Ableitung gilt dann:

$$Df(u)(h) = (w_1, w_2, \ldots, w_n) \quad \text{mit} \quad h = (h_1, h_2, \ldots, h_n)$$

$$w_\ell = \sum_{k=1}^{n} \frac{\partial f_\ell}{\partial x_k}(u) h_k$$

In Matrixschreibweise erhalten wir:

$$\left(\frac{\partial f_\ell}{\partial x_k}(u)\right)_{\substack{\ell = 1 \ldots m \\ k = 1 \ldots n}} \cdot \begin{pmatrix} h_1 \\ \vdots \\ h_n \end{pmatrix} = \begin{pmatrix} w_1 \\ \vdots \\ w_m \end{pmatrix}$$

Beispiel §13.7

Beispiel: $f\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^2$, $(x, y, z) \mapsto (2x^2 z, -xy)$

$$\begin{pmatrix} \frac{\partial f_1}{\partial x} & \frac{\partial f_1}{\partial y} & \frac{\partial f_1}{\partial z} \\ \frac{\partial f_2}{\partial x} & \frac{\partial f_2}{\partial y} & \frac{\partial f_2}{\partial z} \end{pmatrix} = \begin{pmatrix} 4xz & 0 & 2x^2 \\ -y & -x & 0 \end{pmatrix} \rightsquigarrow Df(x, y, z)$$

13.4 Zweimal stetig differenzierbare Funktionen

Vorlesung 25.05.2004

Definition §13.8

Definition: Sei $U \subseteq V$ eine offene Teilmenge. Eine Funktion $U \longrightarrow W$ heißt **zweimal stetig differenzierbar**, falls f einmal stetig differenzierbar ist und falls ihre Ableitung Df ebenfalls stetig differenzierbar ist.

$$\begin{array}{ll} f\colon U \longrightarrow W & u \mapsto f(u) \\ Df\colon U \longrightarrow L(V,W) & u \mapsto Df(u) \\ D^2f\colon U \longrightarrow L(V,L(V,W)) & u \mapsto D^2f(u) \end{array}$$

$Df(u)$ ist eine stetige lineare Abbildung von V nach W. $DDf = D^2f$ ist eine Abbildung mit $D^2f\colon U \longrightarrow L(V,L(V,W))$, das heißt $D^2f(u)(v)$ ist eine lineare Abbildung von V nach W, also $D^2f(u)(v_1)(v_2) \in W$ mit $u \in U$ und $v_1, v_2 \in V$.

Beispiel §13.9

Beispiel: Sei $U = \mathbb{R}^2 = V, W = \mathbb{R}$ und $f\colon V \longrightarrow W$ eine Abbildung mit $f(x,y) = x^2 - 2xy$

1. Ableitung: Gradient $\nabla f(x,y) = (2x - 2y, -2x)$
2. Ableitung: Bilde partielle Ableitungen von ∇f:

$$\begin{array}{ll} \frac{\partial^2 f}{\partial x \partial x}(x,y) = 2 & \frac{\partial^2 f}{\partial x \partial y}(x,y) = -2 \\ \frac{\partial^2 f}{\partial y \partial x}(x,y) = -2 & \frac{\partial^2 f}{\partial y \partial y}(x,y) = 0 \end{array}$$

13.5 Die Hesse-Matrix

Definition §13.10

Definition: Sei $U \subseteq \mathbb{R}^n$ eine offene Menge und sei $f\colon U \longrightarrow \mathbb{R}$ zweimal stetig differenzierbar. Die Matrix

$$H(f)(u) = \left(\frac{\partial^2 f}{\partial x_j \partial x_i}(u) \right)_{i,j=1}^{n} \in \mathbb{R}^{n \times n}$$

heißt **Hesse-Matrix** von f im Punkt $u \in U$. Gradient und Hesse-Matrix braucht man zur Berechnung lokaler Extrema von reellen Funktionen.

13.6 Lokale Extrema

Erinnerung §13.11

Erinnerung: Ein $u \in U$ heißt **lokales Maximum** einer Funktion $f\colon U \longrightarrow \mathbb{R}$, falls es ein $r > 0$ gibt, mit $f(u) \geq f(u')$ für alle $u' \in B_r(u)$. Entsprechend definiert man ein **lokales Minimum**. Gilt $f(u) > f(u')$ für alle $u' \neq u$ in $B_r(u)$ spricht man von einem **strikten lokalen Maximum** (Entsprechend: **striktes lokales Minimum**).

Satz §13.12

Satz: Sei $f\colon U \longrightarrow \mathbb{R}$ stetig differenzierbar. Falls f in $u \in U$ ein lokales Extremum hat, gilt $df(u) = 0$. Falls $U \subseteq \mathbb{R}^n$ bedeutet das:

$$\frac{\partial f}{\partial x_i}(u) = 0, \qquad \text{für } i = 1, \ldots, n$$

Beweis: Sei $v \in V$. Setze $c(t) = u + vt$ und $g(t) = f\big(c(t)\big)$. Dann ist die Abbildung $g\colon]-r, r[\longrightarrow \mathbb{R}$ stetig differenzierbar und hat ein lokales Extremum in $0 \in \mathbb{R}$, weil $g(0) = f(u+tv) \leq f(u)$ beziehungsweise $g(0) = f(u+tv) \geq f(u))$ ist. Also gilt $g'(0) = 0$ (siehe Analysis I). Mit der Kettenregel gilt nun:
$g'(0) = df\big(c(0)\big)\big(\dot{c}(0)\big) = df(u)(v) = 0$, also $df(u) = 0$. Da df in $\mathbb{R}^n$ durch ∇f gegeben ist, gilt $\frac{\partial f}{\partial x_i}(u) = 0$ für $i = 1, \ldots, n$ $\square$

Beispiel: Beispiel §13.13

- $f(x,y) = x^2 + y^2, \qquad \nabla f(x,y) = (2x, 2y)$
  $\nabla f(x,y) = 0 \quad \Leftrightarrow \quad (x,y) = 0$
- $f(x,y) = x^2 - y^2, \qquad \nabla f(x,y) = (2x, -2y)$
  $\nabla f(x,y) = 0 \quad \Leftrightarrow \quad (x,y) = 0$
  $(0,0)$ ist aber *kein* lokales Extremum dieser Funktion.

13.7 Linearität der zweiten Ableitung

Bemerkung §13.14

Ist $f\colon U \longrightarrow \mathbb{R}$ stetig differenzierbar, dann ist $df(u) = Df(u)$ eine lineare Abbildung $V \longrightarrow \mathbb{R}$. Die 2. Ableitung $D^2 f(u)\colon V \longrightarrow L(V,\mathbb{R}) = V^*$ (Dualraum von V) ist ebenfalls eine lineare Abbildung. Für $v, w \in V$ ist $D^2 f(u)(v)(w) \in \mathbb{R}$. Wir schreiben kurz:

$$D^2 f(u)(v,w) = D^2 f(u)(v)(w)$$

So geschrieben ist $D^2 f(u)$ eine Abbildung $V \times V \longrightarrow \mathbb{R},\ (v,w) \mapsto D^2 f(u)(v,w)$. Diese Abbildung $D^2 f(u)(v,w)\colon V \times V \longrightarrow \mathbb{R}$ ist offensichtlich in jedem der beiden Argumente **linear**, das heißt $D^2 f(u)$ ist **bilinear**. Für alle $v, v_1, v_2, w, w_1, w_2 \in V, r \in \mathbb{R}$ gilt:

(i) $D^2 f(u)(v_1 + v_2, w) = D^2 f(u)(v_1, w) + D^2 f(u)(v_2, w)$

(ii) $D^2 f(u)(v, w_1 + w_2) = D^2 f(u)(v, w_1) + D^2 f(u)(v, w_2)$

(iii) $D^2 f(u)(rv, w) = r D^2 f(u)(v,w) = D^2 f(u)(v, rw)$

Die zweite Ableitung ist eine bilineare Abbildung. Ist $U \subseteq \mathbb{R}^n, v = (v_1, \ldots, v_n)$ und $w = (w_1, \ldots, w_n)$, so gilt:

$$D^2 f(u)(v,w) = \sum_{i,j=1}^{n} v_j \frac{\partial^2 f}{\partial x_i \partial x_j}(u) \cdot w_j = (v_1, \ldots, v_n) \cdot H(f)(u) \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

Die Bilinearform $D^2 f(u)$ wird also durch die Hesse-Matrix gegeben.

13.8 Die Kettenregel

Satz §13.15

Satz: Sind V_1, V_2, V_3 Banachräume, $U_1 \subseteq V_1, U_2 \subseteq V_2$ offen und $f\colon U_1 \longrightarrow V_2$, $g\colon U_2 \longrightarrow V_3$ Funktionen mit $f(U_1) \subseteq U_2$, so gilt: Wenn f, g stetig differenzierbar sind, so auch $g \circ f\colon U_1 \longrightarrow V_3$ mit $D(g \circ f)(u) = Dg\big(f(u)\big) \circ Df(u)$

Beweis: Der Beweis funktioniert analog der Kettenregel für Kurven und reelle Funktionen. □

13.9 Die Symmetrie der zweiten Ableitung

Lemma §13.16

Lemma: Sei $r > 0$, $U = \{(x,y) \in \mathbb{R}^2 \mid |x|, |y| < r\}$ und $f\colon U \longrightarrow \mathbb{R}$ zweimal stetig differenzierbar, dann gilt:

$$\frac{\partial^2 f}{\partial x \partial y}(0,0) = \frac{\partial^2 f}{\partial y \partial x}(0,0)$$

Beweis: Setze: $F(x) = f(x,y) - f(x,0)$. Dann ist F stetig differenzierbar mit $F'(x) = \frac{\partial f}{\partial x}(x,y) - \frac{\partial f}{\partial x}(x,0)$. Nach dem Mittelwertsatz (Analysis I, §7.14) ist $F(x) - F(0) = F'(\hat{x})x$ für ein $\hat{x}$ mit $|\hat{x}| \leq |x|$, damit gilt:

$$f(x,y) - f(x,0) - f(0,y) + f(0,0) = \frac{\partial f}{\partial x}(\hat{x},y)x + \frac{\partial f}{\partial x}(\hat{x},0)x$$

Setze $G(y) = \frac{\partial f}{\partial x}(\hat{x},y)$, dann gibt es $\hat{y}$ mit $|\hat{y}| \leq |y|$ so, dass:

$$G(y) - G(0) = G'(\hat{y})y \qquad G'(y) = \frac{\partial^2 f}{\partial y \partial x}(\hat{x},y)$$

Insgesamt: $f(x,y) - f(x,0) - f(0,y) + f(0,0) = \frac{\partial^2 f}{\partial y \partial x}(\hat{x},\hat{y})xy$. Fängt man mit y an, benutzt dann x, erhält man: $f(x,y) - f(x,0) - f(0,y) + f(0,0) = \frac{\partial^2 f}{\partial x \partial y}(\tilde{x},\tilde{y})xy$ mit $|\hat{x}|, |\tilde{x}| \leq |x|; |\hat{y}|, |\tilde{y}| \leq |y|$ Für $xy \neq 0$ hat man also: $\frac{\partial^2 f}{\partial x \partial y}(\tilde{x},\tilde{y}) = \frac{\partial^2 f}{\partial y \partial x}(\hat{x},\hat{y})$. Lässt man x, y gegen 0 gehen, streben auch $\hat{x}, \tilde{x}, \hat{y}, \tilde{y}$ gegen 0. Da f zweimal *stetig* differenzierbar ist, gibt im Grenzwert: $\frac{\partial^2 f}{\partial x \partial y}(0,0) = \frac{\partial^2 f}{\partial y \partial x}(0,0)$ □

Satz: Sei V ein Banachraum, $U \subseteq V$ offen, $f\colon U \longrightarrow \mathbb{R}$ zweimal stetig differenzierbar. Dann gilt: Vorlesung 28.05.2004 Satz §13.17

$$D^2 f(u)(v_1, v_2) = D^2 f(u)(v_2, v_1) \quad \text{für alle } u \in U,\ v_1, v_2 \in V$$

Die Bilinearform $D^2 f(u)$ ist **symmetrisch**. Für $V = \mathbb{R}^n$ gilt speziell, dass die Hesse-Matrix symmetrisch ist:

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(u) = \frac{\partial^2 f}{\partial x_j \partial x_i}(u) \quad \text{für alle } i, j$$

Beweis: Setze $g(x,y) = f(u + xv_1 + yv_2)$ für $x, y \in \mathbb{R}$, dann ist

$\frac{\partial g}{\partial x}(x,y) = Df(u + xv_1 + yv_2)(v_1)$
$\frac{\partial g}{\partial y}(x,y) = Df(u + xv_1 + yv_2)(v_2)$
$\frac{\partial}{\partial y}\frac{\partial g}{\partial x}(x,y) = D^2 f(u + xv_1 + yv_2)(v_1)(v_2)$
$\frac{\partial}{\partial x}\frac{\partial g}{\partial y}(x,y) = D^2 f(u + xv_1 + yv_2)(v_2)(v_1)$

Folgerung: Ist $U \subseteq \mathbb{R}^n$ offen, $v\colon U \longrightarrow \mathbb{R}^n$ stetig differenzierbar, so kann man fragen, ob es eine stetig differenzierbare Funktion $f\colon U \longrightarrow \mathbb{R}$ gibt mit $v(u) = \nabla f(u)$ für alle $u \in U$. Folgerung §13.18

Beispiel: (a) $v\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2, (x,y) \mapsto (y,x)$ $\quad f\colon \mathbb{R}^2 \longrightarrow \mathbb{R}, (x,y) \mapsto xy$
(b) $v\colon \mathbb{R}^2 \longrightarrow \mathbb{R}^2, (x,y) \mapsto (y,-x)$ $\quad f$ gibt es hier nicht!

Eine notwendige Bedingung dafür ist, dass $\frac{\partial v_i}{\partial x_j} = \frac{\partial v_j}{\partial x_i}$ für alle i, j gilt. Ist noch spezieller $n = 3$, setzt man

$$\operatorname{rot}(v) = \begin{pmatrix} \frac{\partial v_3}{\partial x_2} - \frac{\partial v_2}{\partial x_3} \\ \frac{\partial v_1}{\partial x_3} - \frac{\partial v_3}{\partial x_1} \\ \frac{\partial v_2}{\partial x_1} - \frac{\partial v_1}{\partial x_2} \end{pmatrix}$$

dann lautet die notwendige Bedingung $\operatorname{rot}(v) = 0$. $\operatorname{rot}(v)\colon \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ heißt **Rotation von** v.

14 Lokale Extrema reeller Funktionen

Wir betrachten eine zweimal stetig differenzierbare Funktion $f\colon U \longrightarrow \mathbb{R}$, $U \subseteq V$ eine offene Teilmenge eines Banachraumes. Gesucht sind (lokale) Extrema (Maxima oder Minima) von f.

14.1 Schritt 1: Finden von potentiellen Extrema

Paragraph §14.1

Falls f ein Extremum in $u \in U$ hat, so gilt $df(u) = 0$, vergleiche §13.12. Zuerst sucht man also die Punkte $u \in U$ mit $df(u) = 0$; falls $U \subseteq \mathbb{R}^n$ bedeutet das, man sucht Punkte $u \in U \subseteq \mathbb{R}^n$ mit $\frac{\partial f}{\partial x_i}(u) = 0$ für $i \in \{1, \ldots, n\}$.

14.2 Schritt 2: Validieren der gefundenen Stellen

Paragraph §14.2

Angenommen, $df(u) = 0$. Setze $g(s) = f(u + sv)$ für einen festen Vektor $v \in V$, $v \neq 0$ und genügend kleine $s \in \mathbb{R}$. Wenn f ein Extremum in u hat, dann hat g ein Extremum in 0. Nun ist $g\colon]-r, r[\longrightarrow \mathbb{R}$ eine Funktion, die wir mit den Mitteln aus Analysis I behandeln können. Es gilt $g'(s) = df(u + sv)(v)$. $g''(f) = D^2 f(u + sv)(v, v)$.

Schreibweise: Wir setzen $h(v_1, v_2) = D^2 f(u)(v_1, v_2)$.
Wenn f in u ein lokales Maximum hat, muss $g''(0) \leq 0$ gelten, für ein Minimum muss gelten $g''(0) \geq 0$.

Also: f hat ein lokales Minimum in $u \implies h(0,0) \geq 0$ für alle $v \in V$
f hat ein lokales Maximum in $u \implies h(0,0) \leq 0$ für alle $v \in V$

Eine symmetrische Bilinearform $h : V \times V \longrightarrow \mathbb{R}$ heißt

positiv definit	falls $v \neq 0$	$\implies$	$h(v, v) > 0$
positiv semi-definit	falls $v \neq 0$	$\implies$	$h(v, v) \geq 0$
negativ definit	falls $v \neq 0$	$\implies$	$h(v, v) < 0$
negativ semi-definit	falls $v \neq 0$	$\implies$	$h(v, v) \leq 0$
indefinit	ansonsten		

Beispiele

$f(x, y) = x^2 - y^2 \quad \rightsquigarrow \quad \nabla f(x, y) = (2x, -2y)$
$\nabla f(x, y) = (0, 0) \Leftrightarrow (x, y) = (0, 0)$ ist einziger Kandidat für lokales Extremum.

Hesse-Matrix: $H(f)(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \rightsquigarrow Df(0,0)(v, v) = v^T \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} v$

Mit konkreten Werten:

$$Df(0,0)\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right) = (1, 0) \cdot \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 2$$
$$Df(0,0)\left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = (0, 1) \cdot \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -2$$

Das heißt, $D^2 f(0,0)$ ist weder positiv semi-definit noch negativ semi-definit, also kann in $(0,0)$ kein lokales Extremum liegen.

14.3 Hinreichendes Kriterium für lokale Extrema

Wir suchen jetzt ein hinreichendes Kriterium für die Existenz lokaler Extrema. Sei $df(u) = 0$, $g(s) = f(u + sv)$. Dann wissen wir aus Analysis I §8.16:

$$\begin{aligned} g(s) &= g(0) + g'(0) \cdot s + \int_0^s (s-t) \cdot g''(t) \text{ dt} \\ f(u+sv) &= f(u) + \int_0^s D^2 f(u+tv)(v,v) \cdot (s-t) \text{ dt} \end{aligned}$$

Angenommen, es gibt ein $\delta > 0$ so, dass für alle $v \in V$ mit $||v|| = 1$ gilt: $D^2 f(u)(v,v) > \delta$, dann gilt für alle w nahe bei u, dass $D^2 f(w)(v,v) \geq \frac{\delta}{2}$, der Ausdruck im Integral wird also positiv, $g(s) > g(0)$ für genügend kleine $s \neq 0$.

Satz: Sei $U \subseteq V$ offen, $f\colon U \longrightarrow \mathbb{R}$ zweimal stetig differenzierbar, sei $u \in U$ mit $df(u) = 0$. Falls es ein $\delta > 0$ gibt so, dass für alle $v \in V$ mit $||v|| = 1$ gilt: $D^2 f(u)(v,v) > \delta$, so hat f in u ein striktes Minimum, das heißt es gibt $r > 0$ so, dass für alle $w \in U$ mit $||u - w|| \leq r$ und $w \neq u$ gilt: $f(u) < f(w)$. Satz §14.4

Gilt entsprechend $D^2 f(u)(v,v) < -\delta$ mit $||v|| = 1$, so hat f in u ein striktes lokales Maximum.

Vorlesung 01.06.2004

Beweis: Es gibt ein $\varepsilon > 0$ so, dass aus $||w - v|| < \varepsilon$ folgt: $D^2 f(u)(v,v) > \frac{\delta}{2}$ für alle $v \in V$ mit $||v|| = 1$, denn: Setze $A(v,v) = D^2 f(w)(v,v) - D^2 f(u)(v,v)$. $||A(v,v)|| \leq ||A|| \cdot ||v||^2$ (gilt für Operatornorm).

$$D^2 f(w)(v,v) = D^2 f(u)(v,v) - A(v,v)$$

$$|D^2 f(w)(v,v)| \geq \underbrace{|D^2 f(u)(v,v)|}_{\geq \delta} - \underbrace{|A(v,v)|}_{> \frac{\delta}{2}}$$

($A(v,v) > \frac{\delta}{2}$ gilt falls w nahe bei u liegt.)
Es gilt: $g(s) = f(u + sv), g(s) = f(u) + \int_0^s D^2 f(u + ts)(v,v) \cdot (s - t) \text{ dt}$ □

Bemerkung: Wenn V *endliche Dimension* hat und $D^2 f(u)$ positiv definit ist, dann gibt es automatisch so ein $\delta > 0$, denn $||v|| = \sqrt{D^2 f(u)(v,v)}$ ist eine Norm auf V. Weil auf endlich dimensionalen Vektorräumen alle Normen äquivalent sind (nach §11.12) gibt es ein $\delta > 0$ mit $||v|| = 1, ||v||' > \sqrt{\delta}$, also $D^2 f(u)(v,v) > \delta$. Bemerkung §14.5

Satz: Sei $U \subseteq \mathbb{R}^n$ eine offene Menge und $f\colon U \longrightarrow \mathbb{R}$ zweimal stetig differenzierbar. Sei $u \in U$ mit $df(u) = 0$ (das heißt $\nabla f(u) = 0$). Dann hat f ein striktes lokales Minimum (Maximum), falls $D^2 f(u)$ positiv definit (negativ definit) ist. Satz §14.6

Beispiel: Beispiel §14.7

- $f(x,y) = x^2 + y^2, \quad \nabla f(x,y) = (2x, 2y), \quad \nabla f(x,y) = 0 \Leftrightarrow (x,y) = (0,0)$

  $H(f)(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, $H(f)(x,y)$ ist positiv definit, denn:

  $$\begin{pmatrix} v_1 & v_2 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = 2v_1^2 + 2v_2^2 > 0 \quad \text{für } (v_1, v_2) \neq (0,0)$$

- $f(x,y) = 4xy, \quad \nabla f(x,y) = (4y, 4x), \quad \nabla f(x,y) = 0 \Leftrightarrow (x,y) = (0,0)$

  $H(f)(x,y) = \begin{pmatrix} 0 & 4 \\ 4 & 0 \end{pmatrix}$

  $\begin{pmatrix} 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 4 \\ 4 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 4 + 4 = 8 > 0$

  $\begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 4 \\ 4 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = -4 - 4 = -8 < 0$

  Die Hesse-Matrix ist indefinit, also gibt es *kein* lokales Extremum.

- $f(x,y) = xy^2, \quad \nabla f(x,y) = (y^2, 2xy), \quad \nabla f(x,y) = 0 \Leftrightarrow y = 0$

  $H(f)(x,y) = \begin{pmatrix} 0 & 2y \\ 2y & 2x \end{pmatrix}$

  $H(f)(x,y)$ ist positiv semi-definit für $x \geq 0$

  $H(f)(x,y)$ ist negativ semi-definit für $x \leq 0$

  $\Rightarrow$ Hier kann man Satz §14.6 nicht anwenden.

14.4 Definitheitskriterien

Paragraph §14.8

Im Folgenden soll kurz ein Kriterium betrachtet werden, mit dem sich die Definitheit einer Bilinearform leicht ermitteln lässt. Weitere Verfahren und Kriterien werden in der linearen Algebra vorgestellt. Ist $T\colon V \longrightarrow V$ eine lineare Abbildung und $v \in V$, $v \neq 0$ ein Vektor mit $T(v) = c \cdot v$, so heißt v **Eigenvektor** zum **Eigenwert** c. Es gilt der folgende Satz:

Ist $T\colon \mathbb{R}^n \longrightarrow \mathbb{R}^n$ linear und symmetrisch (selbstadjungierend, das heißt es gilt $\langle Tv \,|\, w \rangle = \langle v \,|\, Tw \rangle$ für alle $v, w \in \mathbb{R}^n$, oder auch: T wird durch eine symmetrische Matrix beschrieben), dann gibt es eine Basis $v_1, v_2, \ldots, v_n$ des $\mathbb{R}^n$ aus paarweise orthogonalen Einheitsvektoren, die Eigenvektoren von T sind, das heißt es gilt: $\langle v_i \,|\, v_j \rangle = 0$, $\langle v_i \,|\, v_i \rangle = 1$ sowie $Tv_i = c_i v_i$ für $i, j = 1, \ldots, n$ $i \neq j$ und $c_1, c_2, \ldots, c_n$ Eigenwerte von T mit $c_1 \leq c_2 \leq \cdots \leq c_n$.

Die Bilinearform $(v, w) \mapsto \langle Tv \,|\, w \rangle$ ist:

positiv definit	$\Leftrightarrow$	$0 < c_1 \leq c_2 \leq \cdots \leq c_n$
positiv semi-definit	$\Leftrightarrow$	$0 \leq c_1 \leq c_2 \leq \cdots \leq c_n$
negativ definit	$\Leftrightarrow$	$c_n \leq \cdots \leq c_2 \leq c_1 < 0$
negativ semi-definit	$\Leftrightarrow$	$c_n \leq \cdots \leq c_2 \leq c_1 \leq 0$
indefinit	$\Leftrightarrow$	$c_1 < 0 < c_n$

Bezüglich der neuen Basis $\{v_1, v_2, \ldots, v_n\}$ hat T Diagonalgestalt:

$$T = \begin{pmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & c_n \end{pmatrix}, \text{ also: } \quad \begin{aligned} \det(T) &= c_1 \cdot c_2 \cdot \ldots \cdot c_n \\ \operatorname{spur}(T) &= c_1 + c_2 + \ldots + c_n \end{aligned}$$

Falls T positiv definit ist, gilt also:	$\det(T) > 0$ und $\operatorname{spur}(T) > 0$
Falls T negativ definit ist, gilt also:	$\det(-T) > 0$ und $\operatorname{spur}(T) < 0$

Satz §14.9

Satz: Im $\mathbb{R}^2$ gilt: $\det(T) > 0$ und $\mathrm{spur}(T) > 0 \Rightarrow T$ ist positiv definit.

Beweis: $\det(T) = c_1 \cdot c_2 > 0$ und $\mathrm{spur}(T) > 0 \Rightarrow c_1 > 0$ und $c_2 > 0$ $\square$

Bemerkung §14.10

Bemerkung: Betrachtet man den $\mathbb{R}^3$, so reicht das Kriterium §14.9 nicht, zum Beispiel ist die Matrix

$$M = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 3 \end{pmatrix}$$

indefinit, obwohl gilt: $\det(M) = 3 > 0$ und $\mathrm{spur}(M) = 1 > 0$.

Das Hurwitz-Kriterium

Satz §14.11

Satz: Ist $M = \left(m_{ij}\right)_{i,j=1}^n$ symmetrisch, so ist M posititv definit genau dann, wenn für alle $k = 1, 2, \ldots, n$ gilt:

$$\det \begin{pmatrix} m_{11} & \cdots & m_{1k} \\ \vdots & \ddots & \vdots \\ m_{k1} & \cdots & m_{kk} \end{pmatrix} > 0$$

Beweis: Siehe Lineare Algebra, zum Beispiel Fischer: „Lineare Algebra“. $\square$

Bemerkung: M ist negativ definit genau dann, wenn $-M$ positiv definit ist.

Beispiel §14.12

Beispiel: Die Matrix

$$M = \begin{pmatrix} 2 & \frac{1}{2} & 0 \\ \frac{1}{2} & 2 & \frac{1}{3} \\ 0 & \frac{1}{3} & 2 \end{pmatrix}$$

ist nach dem Hurwitz-Kriterium positiv definit, denn:

$$\det(2) = 2 > 0 \qquad \det \begin{pmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & 2 \end{pmatrix} = \frac{15}{4} > 0 \qquad \det(M) = 8 - \frac{2}{9} - \frac{1}{2} > 0$$

15 Mittelwertsatz und das lokale Inverse

Vorlesung 04.06.2004

15.1 Beschränkte Funktionen

Es sei $J = [a,b] \subseteq \mathbb{R}$ ein abgeschlossenes Intervall, V ein Banachraum. Eine Abbildung $f\colon J \longrightarrow V$ heißt **beschränkt** falls es ein $R > 0$ gibt, so $||f(t)|| \leq R$ für alle $t \in J$. Die beschränkten Funktionen auf J bilden einen reellen Vektorraum (das heißt $||f(t)|| \leq R$, $||g(t)|| \leq S \;\Rightarrow ||(f+g)(t) \leq R+S$). Setze $||f||_\infty = \sup\{||f(t)|| \bigm| t \in J\}$, damit ist $||\cdot||_\infty$ eine Norm auf diesem Vektorraum, die **Supremumsnorm**. Setze $\mathrm{B}\,(J,V) = \{f\colon J \longrightarrow V \mid f \text{ ist beschränkt}\}$ mit dieser Norm $||\cdot||_\infty$. In Analysis I hatten wir den Spezialfall $V = \mathbb{R}$.

Lemma §15.1

Lemma: $\mathrm{B}\,(J,V)$ ist ein Banachraum.

Beweis: Sei $(f_n)_{n\in\mathbb{N}}$ eine Cauchy-Folge in $\mathrm{B}\,(J,V)$. Wegen $||f_k(t) - f_\ell(t)|| \leq ||f_k - f_\ell||$ ist $\big(f_k(t)\big)_{k\in\mathbb{N}}$ eine Cauchy-Folge in V. Setze $f(t) = \lim\limits_{k\to\infty} f_k(t)$. Wegen $||f_k(t)|| \leq ||f_k||_\infty$ ist $(||f_k(t)||)_{k\in\mathbb{N}}$ beschränkt, die Schranke hängt nicht von t ab, also $f \in \mathrm{B}\,(J,V)$.

Schließlich gilt: Zu jedem $\varepsilon > 0$ gibt es ein $N \in \mathbb{N}$ mit $||f_k - f_t|| < \frac{\varepsilon}{2}$ für $k, \ell \in \mathbb{N}$. Dann gilt $||f(t) - f_k(t)|| = \lim\limits_{n\to\infty} ||f_n(t) - f_k(t)|| \leq \frac{\varepsilon}{2}$ für $k \geq N$, also $\lim\limits_{n\to\infty} f_n = f$.

15.2 Stufenfunktionen

Definition §15.2

Eine beschränkte Funktion f heißt **Stufenfunktion** falls es eine Zerlegung

$$s_o = a < s_1 < \ldots < s_N = b$$

gibt, so dass f auf jedem Intervall $]s_{k-1}, s_k[$ konstant ist. Die Menge $\mathrm{Step}\,(J,V)$ aller Stufenfunktionen ist ein Untervektorraum von $\mathrm{B}\,(J,V)$.

15.3 Regelfunktionen

Eine beschränkte Funktion $f \in \mathrm{B}\,(J,V)$ die sich als Grenzfunktion einer Folge von Stufenfunktionen schreiben lässt, heißt **Regelfunktion**.

$$f = \lim_{n\to\infty} f_n \quad \text{für} \quad f_n \in \mathrm{Step}\,(J,V)$$

Die Menge der Regelfunktionen ist ebenfalls ein Untervektorraum von $\mathrm{B}\,(J,V)$, die Menge der beschränkten Funktionen. Die stetigen Funktionen sind Teilmenge der Regelfunktionen.

$$\begin{array}{c} \mathrm{B}\,(J,V) \\ | \\ \mathrm{Reg}\,(J,V) \\ | \\ \mathrm{Step}\,(J,V) \end{array}$$

Abbildung 23: Funktionsklassen

Lemma §15.3

Lemma: Reg (J, V) ist ein Banachraum (bezüglich der $||\cdot||_\infty$-Norm).

Beweis: Wir zeigen: Reg (J, V) ist *abgeschlossen* in B (J, V), dann ist es nach §10.11 ein Banachraum.
Sei also $(f_n)_{n\in\mathbb{N}}$ eine konvergente Folge mit $\lim\limits_{n\to\infty} f_n = f \in \mathrm{B}\,(J,V)$ und sei $f_n \in \mathrm{Reg}\,(J,V)$ für alle n. Zu zeigen: $f \in \mathrm{Reg}\,(J,V)$. Zu jedem n gibt es eine Stufenfunktion g_n mit $||f_n - g_n||_\infty < \frac{1}{n}$. Idee: Zeige $\lim\limits_{n\to\infty} g_n = f$. Wähle $N \in \mathbb{N}$ so, dass aus $k \geq N$ folgt $||f_k - f|| \leq \frac{\varepsilon}{2}$ (für gegebenes $\varepsilon > 0$). Dann gilt

$$||g_k - f||_\infty = ||g_k - f_k + f_k - f||_\infty \leq \underbrace{||g_k - f_k||_\infty}_{<\frac{1}{k}} + \underbrace{||f_k - f||_\infty}_{<\frac{\varepsilon}{2}} < \varepsilon$$

15.4 Das Riemann-Integral

Definition §15.4

Definition: Sei $f\colon J \longrightarrow V$ eine Stufenfunktion zur Zerlegung

$$s_0 = a < s_1 < s_2 < \ldots < s_N = b$$

sei $f(t) = c_k$ für alle $t \in]s_{k-1}, s_k[$, setze

$$\int_a^b f(t)\ \mathrm{dt} = c_1(s_1 - s_0) + c_2(s_2 - s_1) + \ldots + c_N(s_N - s_{N-1})$$

Die Abbildung Step $(J, V) \longrightarrow V, f \mapsto \int_a^b f(t)\ \mathrm{dt}$ ist linear, denn

$$\begin{aligned}\int_a^b (f_1 + f_2)(t)\ \mathrm{dt} &= \int_a^b f_1(t)\ \mathrm{dt} + \int_a^b f_2(t)\ \mathrm{dt} \\ \int_a^b cf(t)\ \mathrm{dt} &= c\int_a^b f(t)\ \mathrm{dt}\end{aligned}$$

und L-Lipschitz-stetig mit $L = b - a$, denn

$$||\int_a^b f(t)\ \mathrm{dt}|| \leq \int_a^b ||f(t)||\ \mathrm{dt} \leq \int_a^b ||f||_\infty\ \mathrm{dt} = (b-a)\cdot||f||_\infty$$

Ist $(f_n)_{n\in\mathbb{N}}$ eine Folge von Stufenfunktionen, die gegen die Regelfunktion f konvergiert, so konvergiert auch die Folge $\int_a^b f_n(t)\ \mathrm{dt}$ (weil das Integrieren Lipschitz-stetig ist), setze $\int_a^b f(t)\ \mathrm{dt} = \lim\limits_{n\to\infty} \int_a^b f_n(t)\ \mathrm{dt}$.
Das ist **wohldefiniert** (das heißt es hängt *nicht* von der Wahl der Folge ab), denn angenommen

$$\lim_{n\to\infty} f_n = f = \lim_{n\to\infty} g_n \quad \text{für} \quad f_n, g_n \in \mathrm{Step}\,(J,V)$$

dann gilt $\lim\limits_{n\to\infty} (f_n - g_n) = 0$, damit $\lim\limits_{n\to\infty} \int_a^b \big(f_n(t) - g_n(t)\big)\ \mathrm{dt} = 0$, also $\lim\limits_{n\to\infty} \int_a^b f_n(t)\ \mathrm{dt} = \lim\limits_{n\to\infty} \int_a^b g_n(t)\ \mathrm{dt}$. □

Satz §15.5

Satz: Das Riemann-Integral ist eine Lipschitz-stetige lineare Abbildung, $\mathrm{Reg}\,(J,V) \longrightarrow V, f \mapsto \int_a^b f(t)\ \mathrm{dt}$ und es gilt:

$$\begin{aligned} \|\int_a^b f(t)\ \mathrm{dt}\| &\leq \int_a^b \|f(t)\|\ \mathrm{dt} \ \leq\ \|f\|_\infty \cdot (b-a) \\ \int_a^b (f+g)(t)\ \mathrm{dt} &= \int_a^b f(t)\ \mathrm{dt}\ +\ \int_a^b g(t)\ \mathrm{dt} \end{aligned}$$

Satz §15.6

Satz: Ist $f\colon J \longrightarrow V$ stetig, $J=[a,b]$, so ist f **gleichmäßig stetig**, das heißt zu jedem $\varepsilon > 0$ gibt es $\delta > 0$ mit:

$$|s-t| < \delta \quad \Longrightarrow \quad \|f(s)-f(t)\| < \varepsilon$$

Beweis (vergleiche Analysis I §6.14): Angenommen das wäre nicht so. Dann gibt es ein $\varepsilon > 0$ und $s_n, t_n \in J$ mit $|s_n - t_n| < \frac{1}{n}$, $\|f(s_n) - f(t_n)\| \geq \varepsilon$. Nach Bolzano-Weierstraß dürfen wir annehmen, dass $(s_n)_{n\in\mathbb{N}}$, $(t_n)_{n\in\mathbb{N}}$ konvergieren, $\lim\limits_{n\to\infty} s_n = \lim\limits_{n\to\infty} t_n = t$. Aber $\|f(t_n) - f(s_n)\| \geq \varepsilon$. Im Grenzwert $n \to \infty$ ergibt dies aber $\|f(t) - f(t)\| = 0$. (Widerspruch) □

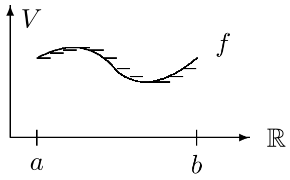
15.5 Regelfunktionen und Integration

Vorlesung 08.06.2004

Satz §15.7

Satz: Auf einem abgeschlossenen Intervall $J=[a,b]$ ist jede stetige Funktion $f\colon J \longrightarrow V$ eine Regelfunktion, also $\mathrm{C}\,(J,V) \subseteq \mathrm{Reg}\,(J,V)$. Insbesondere ist für jede stetige Funktion das Riemann-Integral definiert.

Beweis: Sei $f\colon J \longrightarrow V$ stetig, sei $\varepsilon > 0$. Wähle $\delta > 0$ so, dass aus $|s-t| < \delta$ folgt: $\|f(s) - f(t)\| < \varepsilon$ (Das geht nach Satz §15.6). Wähle $m \in \mathbb{N}$ so, dass $\frac{b-a}{m} < \delta$, setze $s_k = a + k\frac{b-a}{m}$, $k = 0,\dots,m$, setze $c_k = f(s_k)$.



$$g(s) = \begin{cases} c_0 & \text{für } s_0 \leq s < s_1 \\ c_1 & \text{für } s_1 \leq s < s_2 \\ \vdots & \\ c_m & \text{für } s_m = b \end{cases}$$

Abbildung 24: Stetige Regelfunktionen

Nach Konstruktion gilt: $\|g(s) - f(s)\| < \varepsilon$ also $\|g - f\|_\infty < \varepsilon$. Damit gilt: $f \in \mathrm{Reg}\,(J,V)$ □

Dieses Verfahren liefert einen Algorithmus zur numerischen Berechnung von Integralen (Wähle m groß).

Bemerkung §15.8

Bemerkung: Ist $V = \mathbb{R}^n$ und $f(s) = \big(f_1(s), f_2(s), \dots, f_n(s)\big)$ eine Regelfunktion (zum Beispiel eine stetige Funktion), dann gilt:

$$\int_a^b f(s)\ \mathrm{ds} = \left(\int_a^b f_1(s)\ \mathrm{ds}, \int_a^b f_2(s)\ \mathrm{ds}, \dots, \int_a^b f_n(s)\ \mathrm{ds}\right)$$

denn das stimmt für Stufenfunktionen, also auch für Regelfunktionen.

Beispiel:

- $V = \mathbb{R}^2, \quad J = [0,3], \quad f(s) = (s^2, s)$
  $\int_0^3 f(s)\ \mathrm{ds} = \left(\frac{1}{3}s^3, \frac{1}{2}s^2\right) \Big|_0^3 = \left(9, \frac{9}{2}\right)$
- $V = \mathbb{R}^2, \quad J = [a,b], \quad g(s) = \big(\cos(s), \sin(s)\big)$
  $\int_a^b g(s)\ \mathrm{ds} = \big(\sin(s), -\cos(s)\big) \Big|_a^b = \big(\sin(b) - \sin(a), \cos(a) - \cos(b)\big)$

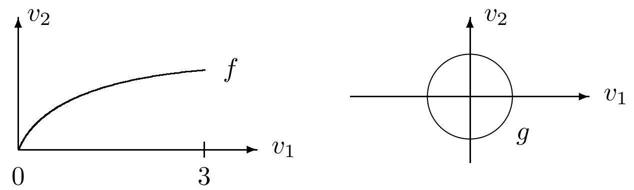


Abbildung 25: Funktionsgraphen von f und g

Satz §15.9

Satz: Ist $c\colon J \longrightarrow V$ stetig, $a \le t_0 \le b$ und $g(s) = \int_{t_0}^s c(t)\ \mathrm{dt}$, dann ist g stetig differenzierbar und $\dot{g} = c$ (dabei setzen wir $\int_{t_0}^s c(t)\ \mathrm{dt} = -\int_s^{t_0} c(t)\ \mathrm{dt}$, falls $s < t_0$).

Beweis: $\int_{t_0}^{t+h} c(s)\ \mathrm{ds} - \int_{t_0}^{t} c(s)\ \mathrm{ds} = g(t+h) - g(t) = \int_t^{t+h} c(s)\ \mathrm{ds}$, also:

$$\begin{aligned} g(t+h) - g(t) - h \cdot c(t) &= \int_t^{t+h} (c(s) - c(t))\ \mathrm{ds} \\ ||g(t+h) - g(t) - h \cdot c(t)|| &\le \int_t^{t+h} ||c(s) - c(t)||\ \mathrm{ds} \\ &\le |h| \sup\big\{ ||c(s) - c(t)|| \ \big|\ |s-t| \le h \big\} \end{aligned}$$

Also folgt $g(t+h) - g(t) - h \cdot c(t) = |h|\lambda(h)$ mit λ stetig und $\lambda(0) = 0$, also $\dot{g}(t) = c(t)$ □

Insbesondere gilt: $g(b) - g(a) = \int_a^b c(t)\ \mathrm{dt} = \int_a^b \dot{g}(t)\ \mathrm{dt}$

Folgerung: Ist $g\colon [a,b] \longrightarrow V$ eine stetig differenzierbare Kurve, so gilt:

$$\int_a^b \dot{g}(s)\ \mathrm{ds} = g(b) - g(a)$$

(Wende den Satz auf $c = \dot{g}$ an).

15.6 Mittelwertsatz der Integralrechnung

Lemma §15.10

Lemma: Ist $f\colon J \longrightarrow L(V,W)$ eine Regelfunktion, so gilt: $t \mapsto f(t)(v) \in W$ ist ebenfalls eine Regelfunktion (für $v \in V$ fest gewählt) und

$$\int_a^b f(t)(v)\ \mathrm{dt} = \left(\int_a^b f(t)\ \mathrm{dt}\right)(v)$$

Beweis: Das ist für Stufenfunktionen sicher richtig und bleibt im Grenzwert auch für Regelfunktionen richtig. □

Satz §15.11

Satz: Seien V, W Banachräume, $U \subseteq V$ offen und $f\colon U \longrightarrow W$ stetig differenzierbar. Sind $u \in U$ und $v \in V$ so, dass $u + sv \in U$ für $0 \leq s \leq 1$, dann gilt:

$$f(u+v) - f(u) = \int_0^1 Df(u+sv)(v)\ \mathrm{ds} = \left(\int_0^1 Df(u+sv)\ \mathrm{ds}\right)(v)$$

Beweis: Setze $g(s) = f(u+sv), g\colon [0,1] \longrightarrow V$, $\dot{g}(s) = Df(u+sv)(v)$, also $\int_0^1 \dot{g}(s)\ \mathrm{ds} = g(1) - g(0) = f(u+v) - f(u) = \int_0^1 Df(u+sv)(v)\ \mathrm{ds}$ □

Folgerung §15.12

Folgerung: Seien V, W Banachräume, $U \subseteq V$ offen und $f\colon U \longrightarrow W$ stetig differenzierbar. Sind $u \in U$ und $v \in V$ so, dass $u + sv \in U$ für $0 \leq s \leq 1$, dann gilt:

$$||f(u+v) - f(u)|| \leq ||v|| \cdot R$$
$$R = \sup \left\{ ||Df(u+sv)|| \;\middle|\; 0 \leq s \leq 1 \right\}$$

Beweis: Der Beweis folgt aus dem Mittelwertsatz der Integralrechnung:

$$\begin{aligned} ||f(u+v) - f(u)|| &\leq \int_0^1 ||Df(u+sv)(v)||\ \mathrm{ds} \\ &\leq \int_0^1 ||Df(u+sv|| \cdot ||v||\ \mathrm{ds} \\ &\leq \int_0^1 R \cdot ||v||\ \mathrm{ds} = R \cdot ||v|| \end{aligned}$$

□

Bemerkung: Die Abbildung $t \mapsto ||Df(u+tv)||$ ist stetig, also existiert solch ein R (siehe Analysis I §5.9).

Satz §15.13

Satz: Ist V ein Banachraum, $U = \left\{ f \in L(V,W) \;\middle|\; ||f|| < 1 \right\}$ und ist $f \in U$, dann hat die Abbildung $v \mapsto v - f(v) = (\mathrm{id}_V - f)(v)$ ein stetiges Inverses $g \in L(V,V)$.

Beweis: Setze $g_n = \sum_{k=0}^n f^k = \mathrm{id}_V + f + f^2 + \cdots + f^n$ (geometrische Summe). $(\mathrm{id}_V - f)(\mathrm{id}_V + f + f^2 + \cdots + f^n) = \mathrm{id}_V - f^{n+1}$ sowie $||f^n|| \leq ||f||^n$. Es gilt: $\lim\limits_{n\to\infty} ||f||^n = 0$ (weil $||f|| < 1$), folglich ist $\lim\limits_{n\to\infty} f^n = 0$. Aus der Gleichung erhält man im Limes $\lim\limits_{n\to\infty} (\mathrm{id}_V - f) \sum_{k=0}^n f^k = \mathrm{id}_V$, denn $(g_n)_{n\in\mathbb{N}}$ ist eine Cauchy-Folge in $L(V,V)$, also $\lim\limits_{n\to\infty} g_n = \sum_{k=0}^\infty f^k = g$, also: $(\mathrm{id}_V - f) \circ g = \mathrm{id}_V$. Man rechnet genauso nach: $g \circ (\mathrm{id}_V - f) = \mathrm{id}_V$, also $g = (\mathrm{id}_V - f)^{-1}$. □

Man nennt $\sum\limits_{k=0}^{\infty} f^k$ die **Neumann'sche Reihe** (= geometrische Reihe) zu f.

$$(\mathrm{id}_V - f)^{-1} = \sum_{k=0}^{\infty} f^k$$

15.7 Der Satz vom lokalen Inversen

Vorlesung 11.06.2004

Satz §15.14

Satz: Seien V, W Banachräume (zum Beispiel V, W isomorph zu $\mathbb{R}^n$), $U \subseteq V$ offen, $f: U \longrightarrow W$ stetig differenzierbar.
Sei $u_0 \in U$, setze $T = Df(u_0): V \longrightarrow W$. Falls T ein Vektorraumisomorphismus ist, gibt es ein $r > 0$ und ein Inverses $g: B_r\big(f(u_0)\big) \longrightarrow U$ zu f. g ist stetig differenzierbar und $Dg(f(u_0)) = T^{-1}$.

Mit anderen Worten: Die Gleichung $f(x) = y$ lässt sich für y nahe $f(u_0)$ nach x auflösen.

Den Beweis zerlegen wir in Einzelschritte:

Satz §15.15

Satz: Satz 14 ist richtig, falls f eine stetige lineare Abbildung ist.

Beweis: Wenn f linear ist, so gilt $Df(u_0)(v) = f(v)$, das heißt $T = f$ ist ein Vektorraumisomorphismus. Dann ist $T^{-1}: W \longrightarrow V$ ein Vektorraumisomorphismus (lineare Algebra) *und* T^{-1} ist stetig, denn falls $\dim V$ endlich ist (§11.11) ist das sicher richtig. Es ist auch im Unendlichdimensionalen richtig (Banachs Satz von der offenen Abbildung, vergleiche Heuser → Funktionalanalysis). Jedenfalls ist $T^{-1} = f^{-1}$ dann stetig, setze $g = T^{-1}$. □

Satz §15.16

Satz: Satz 14 ist in der folgenden Situation richtig:
Falls $V = W$, $u_0 = 0 = f(u_0)$ und $T = \mathrm{id}_V$.

Beweis: **(a)** Behauptung: Es gibt ein $r > 0$ so, dass für $||u|| \leq r$ folgt: $Df(u)$ hat ein stetiges Inverses. Denn: Die Abbildung $u \mapsto Df(u)$ ist stetig, also gibt es ein $r > 0$ so, dass folgendes gilt:

$$||u|| \leq r \quad \Longrightarrow \quad ||Df(u) - \underbrace{Df(0)}_{=\mathrm{id}_V}|| < 1$$

Wegen Satz 13 ist $Df(u)$ invertierbar (Neumann'sche Reihe). □

(b) Setze $p(x) = x - f(x)$, dann gilt $p(0) = 0$ und $Dp(0) = D\mathrm{id}_V - Df(0) = 0$. Es gibt also ein $r > 0$ so, dass $||Dp(u)|| \leq \frac{1}{2}$ für $||u|| < r$, und $Df(u)$ hat ein stetiges Inverses. Nach Folgerung 12 ist $||p(u)|| \leq \frac{1}{2} \cdot r$ für $||u|| \leq r$.

(c) Behauptung: Ist $||u|| \leq \frac{r}{2}$ so gibt es genau ein u mit $||u|| \leq r$ und $f(u) = v$. Denn: Betrachte $q(x) = v + p(x) = v + x - f(x)$. Für $||x|| \leq r$ gilt

$$||q(x)|| \leq \underbrace{||v||}_{\leq \frac{r}{2}} + \underbrace{||p(x)||}_{\leq \frac{r}{2}} \leq r$$

Setze $X = \{x \in V \mid ||x|| \leq r\}$, das ist abgeschlossen in V, also vollständig, $q: X \longrightarrow X$ ist stetig. Es gilt $||q(x) - q(y)|| = ||p(x) - p(y)|| \leq \frac{1}{2} \cdot ||x - y||$ (nach Teil (b)). Wir können Banachs Fixpunktsatz anwenden! $\Longrightarrow$ Es gibt genau ein Element $u \in X$ mit $q(u) = u$, das heißt $v + u - f(u) = u \;\Leftrightarrow\; v = f(u)$.

(d) Setze $g(v) = u$, das ist eine Umkehrfunktion. Warum ist g stetig? Denn $v = p(v) + f(v)$ (folgt aus dem Anfang von (b)) also

$$||v_1 - v_2|| \leq \underbrace{||p(v_1) - p(v_2)||}_{\frac{1}{2}||v_1 - v_2||} + ||f(v_1) - f(v_2)||$$

Mit $g(u_1) = v_1$ und $g(u_2) = v_2$ ergibt sich:

$$\begin{aligned} ||g(u_1) - g(u_2)|| &\leq \tfrac{1}{2} \cdot ||g(u_1) - g(u_2)|| + ||u_1 + u_2|| \\ ||g(u_1) - g(u_2)|| &\leq 2 \cdot ||u_1 - u_2|| \quad \Longrightarrow \quad g \text{ ist stetig} \end{aligned}$$

(e) g ist differenzierbar in jedem Punkt v_1 mit $||v_1|| < \frac{r}{2}$. Denn:
Setze $g(v_1+h) = u$, $v_1+h = f(u)$, $g(v_1) = u_1$, $v_1 = f(u_1)$ und $h = f(u)-f(u_1)$.

$$\begin{aligned} & ||g(v_1 + h) - g(v_1) - Df(u_1)^{-1}(h)|| \\ = \;& ||u - u_1 - Df(u_1)^{-1}(f(u) - f(u_1))|| \\ & \big[f(u) - f(u_1) = Df(u_1)(u - u_1) + ||u - u_1||\lambda(u - u_1)\big] \\ = \;& ||u - u_1 - Df(u_1)^{-1}(Df(u_1)(u - u_1) + ||u - u_1||\lambda(u - u_1))|| \\ \leq \;& ||Df(u_1)\underbrace{\lambda(u - u_1)}_{\hookrightarrow 0}|| \cdot \underbrace{||u - u_1||}_{\hookrightarrow 0} \end{aligned}$$

also ist g differenzierbar in v_1 und $Dg(v_1) = Df(u_1)$.

(f) g ist stetig differenzierbar, weil $Df(g(v_1))^{-1}$ stetig von v_1 abhängt.

Damit ist Satz 16 bewiesen.

Beweis von Satz 14: Setze $\tilde{f}(x) = T^{-1}(f(x + u_0) - f(u_0))$, damit ist $\tilde{f}$ stetig differenzierbar, $\tilde{f}\colon V \longrightarrow V$, $\tilde{f}(0) = 0$ und $D\tilde{f}(0) = T^{-1}Df(u_0) = \mathrm{id}_V$.

Nach Satz 16 hat $\tilde{f}$ ein lokales Inverses $\tilde{g}$. Löst man $\tilde{f}$ nach f auf, liefert $\tilde{g}$ ein lokales Inverses g zu f. $\square$

Beispiel: $f(x) = x^2$, $Df(1)(v) = 2v$ mit $V = \mathbb{R} = W$.
Satz 14 sagt: f hat ein lokales Inverses, nämlich $g(y) = \sqrt{y}$.

15.8 Satz über implizite Funktionen

Sind X, Y, Z Banachräume (z.B. $\mathbb{R}^k, \mathbb{R}^\ell, \mathbb{R}^m$), $U \subseteq X$ und $V \subseteq Y$ offen, so ist $X \times Y$ ein Banachraum und $U \times V \subseteq X \times Y$ offen. Ist $f\colon U \times V \longrightarrow Z$ stetig differenzierbar und ist $(u, v) \in U \times V$, dann ist Vorlesung 15.06.2004

$$Df(u, v)(x, y) = D_1f(u, v)(x) + D_2f(u, v)(y)$$

mit $D_1f(u, v)\colon X \longrightarrow Z$ und $D_2f(u, v)\colon Y \longrightarrow Z$ oder ganz kompakt:

$$Df(u, v) = \big(D_1f(u, v), D_2f(u, v)\big)$$

$$\big(D_1f(u, v), D_2f(u, v)\big) \begin{pmatrix} x \\ y \end{pmatrix} = D_1f(u, v)(x) + D_2f(u, v)(y)$$

Satz §15.17

Satz: Seien X, Y, Z Banachräume, $U \subseteq X$ und $V \subseteq Y$ offen. $f\colon U \times V \longrightarrow Z$ sei stetig differenzierbar. Falls $D_2f(u, v)\colon Y \longrightarrow Z$ ein Isomorphismus ist, gibt es $U_0 \subseteq U$ offen mit $u \in U_0$ und eine stetig differenzierbare Funktion $g\colon U_0 \longrightarrow V$ mit $g(u) = v$ und $f\big(x, g(x)\big) = f(u, v) = \mathit{const}$ für alle $x \in U_0$. (Idee: Man will die Gleichung $f(x, y) = \mathit{const}$ lokal nach y auflösen.

Beweis: Definiere $z_0 = f(u,v)$, $\tilde{f}\colon U \times V \longrightarrow Z$, $\tilde{f}(u,v) = \big(u, f(u,v)\big)$. Dann gilt:
$D\tilde{f}(u,v)\begin{pmatrix} x \\ y \end{pmatrix} = \left(x, Df(u,v)\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} \mathrm{id}_X & 0 \\ D_1f(u,v) & D_2f(u,v) \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$
Diese Matrix ist invertierbar mit Inverser:

$$\begin{pmatrix} \mathrm{id}_X & 0 \\ A & D_2f(u,v)^{-1} \end{pmatrix}, \quad A = -D_2f(u,v)^{-1}D_1f(u,v)$$

Nach Satz §15.14 hat $\tilde{f}$ ein lokales Inverses $\tilde{p}$ definiert auf $U_0 \times W_0 \subseteq X \times Z$ mit $f(u,v) \in U_0 \times W_0$
$\tilde{p}(x,z) = \big(x, p(x,z)\big)$. Setze $g(x) = p(x, z_0)$, dann ist g stetig differenzierbar und $\big(x, f(x, g(x))\big) = \tilde{f}\big(x, g(x)\big) = \tilde{f}\big(x, p(x, z_0)\big) = \tilde{f}\big(\tilde{p}(x, z_0)\big) = (x, z_0)$, also $f\big(x, g(x)\big) = z_0 = \mathit{const}$ $\square$

Beispiel §15.18

Beispiel: $X = Y = Z = U = V = \mathbb{R}$, $\quad f(x,y) = x^2 - y^2$
$Df(x,y) = \big(D_1f(x,y), D_2f(x,y)\big) = (2x, -2y)$. Für $y \neq 0$ ist $D_2f(x,y)$ invertierbar, z.B. in $(x,y) = (0,1)$ ist $D_2f(0,1) = -2$.
Es gibt also g mit $f\big(x, g(x)\big) = f(0,1) = -1$, $f\big(x, g(x)\big) = x^2 - g(x)^2 = -1$, also $g(x) = \sqrt{1+x^2}$, $g(0) = 1$

Anwendung §15.19

Erste Anwendung: *Niveaumengen („Höhenlinien“) reeller Funktionen.*
Sei $f\colon U \longrightarrow \mathbb{R}$ eine reelle Funktion, $U \subseteq V$ offen in einem Banachraum V (z.B. $V = \mathbb{R}^2$). Sei $u \in U$ mit $df(u) \neq 0$. Weil $df(u)\colon V \longrightarrow \mathbb{R}$ eine lineare Abbildung ist, ist $H = \ker\big(df(u)\big) = \{v \in V \mid df(u)(v) = 0\}$ ein Untervektorraum von V und als Teilmenge in V abgeschlossen, also ebenfalls ein Banachraum. Wähle $w \in V$ mit $df(u)(w) \neq 0$ (das heißt $w \notin H$), dann lässt sich jeder Vektor $v \in V$ eindeutig zerlegen in $v = h + sw$ mit $h \in H$ und $s \in \mathbb{R}$, also $V = H \oplus (\mathbb{R}w)$.

Schreibe $\tilde{f}(h, sw) = f(u + h + sw)$, $\tilde{f}$ ist eine reellwertige Funktion, es gilt:

$$D_1\tilde{f}(0,0)(h) = df(u)(h) = 0$$
$$D_2\tilde{f}(0,0)(s) = df(u)(w) \cdot s \neq 0$$

also ist $D_2\tilde{f}(0,0)$ invertierbar und es gibt g mit $\tilde{f}\big(h, g(h)\big) = \tilde{f}(0,0)$ für $h \in H$ nahe bei 0.

$$\tilde{f}\big(h, g(h)\big) = f\big(u + h + g(h)\cdot w\big) = f\big(p(h)\big) \quad \text{mit } p\colon U_0 \longrightarrow V,\ U_0 \subseteq H$$

Also $f \circ p = \mathit{const}$, p parameterisiert diese Niveaumenge (Höhenlinie) nahe u.

15.9 Extrema mit Nebenbedingungen

Vorlesung 18.06.2004
Definition §15.20

Definition: Sei $U \subseteq V$ offen in einem Banachraum V,

$$q\colon U \longrightarrow \mathbb{R}, \quad f\colon U \longrightarrow \mathbb{R}$$

q und f stetig differenzierbar. Setze $q^{-1}(0) = N = \{u \in U \mid q(u) = 0\}$ als Nebenbedingung, das sind die Nullstellen von q. Ein Punkt $u \in N$ heißt **Maximum/Minimum** oder **Extremum** von f **mit** der **Nebenbedingung** $q = 0$, falls für alle $v \in N$ gilt:

$f(u) \geq f(v) \iff$ Maximum von f mit Nebenbedingung
$f(u) \leq f(v) \iff$ Minimum von f mit Nebenbedingung

15.10 Der Lagrange-Multiplikator

Satz §15.21

Satz: Sei $U \subseteq V$ offen, $f, q\colon U \longrightarrow \mathbb{R}$ stetig differenzierbar, sei $u \in U$ mit $q(u) = 0$ und $dq(u) \neq 0$. Wenn u ein Extremum von f mit Nebenbedingung $q = 0$ ist, dann gibt es $r \in \mathbb{R}$ mit $df(u) = r \cdot dq(u)$. Die Zahl r heißt **Lagrange-Multiplikator**. Falls $U \subseteq \mathbb{R}^n$, heißt das

$$\nabla f(u) = r \cdot \nabla q(u) \quad \Longleftrightarrow \quad \frac{\partial f}{\partial x_i}(u) = r \cdot \frac{\partial q}{\partial x_i}(u)$$

mit $i = 1, 2, \ldots, n$ für letzteres, welches ein Gleichungssystem darstellt.

Beweis: Weil $dq(u) \neq 0$ und $dq(u)(w) \neq 0$, nach der vorigen Diskussion finden wir $g\colon W \longrightarrow \mathbb{R}$, $W \subseteq H = \ker\big(dq(u)\big)$, mit $q(u + h + g(h) \cdot w) = 0 = \mathit{const}$ für $h \in W$ und $dg(0) = 0$. Ist u ein Extremum von f in u, so verschwindet die Ableitung der Funktion $h \mapsto f(u + h + g(h) \cdot w)$ an der Stelle $h = 0$, also

$$\begin{aligned} 0 &= df(u)\underbrace{D\,[h \mapsto u + h + g(h) \cdot w](0)(v)}_{v + dg(u)(v) \cdot w} \quad \text{für } v \in H \text{ beliebig} \\ &= df(u)\big(dg(0)(v) \cdot w\big) + df(u)(v) \quad \text{das heißt } H \subseteq \ker\big(df(u)\big) \end{aligned}$$

Setze $df(u)(w) = r \cdot \underbrace{dq(u)(w)}_{\neq 0}$ mit $r = \frac{df(u)(w)}{dq(u)(w)}$.

Jedes $v \in V$ lässt sich schreiben als $v = h + sw$ mit $h \in H$,

$$\begin{aligned} & df(u)(v) = df(u)(h + sw) = df(u)(w) \cdot s \\ & dq(u)(v) = dq(u)(h + sq) = df(u)(w) \cdot s \\ \Rightarrow \quad & df(u)(v) = dq(u)(v) \cdot r \quad \Rightarrow \quad df(u) = dq(u) \cdot r \quad \text{für alle } v \in V \end{aligned}$$

□

Beispiel §15.22

Beispiel: $U = \{(x, y, z) \mid x, y, z > 0\}$, $f(x, y, z) = x + y + z$.
Nebenbedingung: $xyz = 1$, was ist das Minimum unter dieser Nebenbedingung?
$q(x, y, z) = xyz - 1$, $\nabla q(yz, xz, xy) \neq 0$ für $x, y, z \in U \;\rightarrow\;$ gut!
$\nabla f(x, y, z) = (1, 1, 1)$. Ist u Extremum, gilt $r(yz, xz, xy) = (1, 1, 1)$ also $x = y = z = 1$. Man sieht schnell ein: Das ist auch ein Minimum von f unter der Nebenbedingung $xyz = 1$.

15.11 Extrema mit mehreren Nebenbedingungen

Satz §15.23

Satz: Sei $U \subseteq V$ offen, $f\colon U \longrightarrow \mathbb{R}$ stetig differenzierbar und $q_1, q_1, \ldots, q_m$ stetig differenzierbare Funktionen mit $q_1, q_1, \ldots, q_m : U \longrightarrow \mathbb{R}$. Die Menge N bezeichnet die Nebenbedingungen:

$$N = \{u \in U \mid q_1(u) = q_2(u) = \ldots = q_m(u) = 0\}$$

Auf dieser Menge suchen wir Extrema von f, das heißt $u \in N$ mit:

$$v \in N \;\Rightarrow\; f(u) \geq f(v) \rightsquigarrow \text{Maximum}, \quad f(u) \leq f(v) \rightsquigarrow \text{Minimum}$$

jeweils mit Nebenbedingung $q_1, q_2, \ldots, q_m = 0$.

Falls u ein Extremum von f mit dieser Nebenbedingung $q_1 = q_2 = \ldots = q_m$ ist, und falls $dq_1(u) \neq 0, dq_2(u) \neq 0, \ldots, dq_m(u) \neq 0$, so gibt es Zahlen $r_1, r_2, \ldots, r_m$ mit $df(u) = r_1 dq_1(u) + r_2 dq_2(u) + \ldots + r_m dq_m(u)$. Diese Zahlen heißen **Lagrange-Multiplikatoren**.

Beweis: Wie vorher mit $m = 1$.

In Koordinaten heißt das: $U \subseteq \mathbb{R}^n$, $n \leq m$. $\nabla f(u) = r_1 \nabla q_1(u) + \ldots + r_m \nabla q_m(u)$

$$\begin{pmatrix} \frac{\partial f}{\partial x_1}(u) \\ \vdots \\ \frac{\partial f}{\partial x_n}(u) \end{pmatrix} = \begin{pmatrix} \frac{\partial q_1}{\partial x_1}(u) & \cdots & \frac{\partial q_m}{\partial x_1}(u) \\ \vdots & & \vdots \\ \frac{\partial q_1}{\partial x_n}(u) & \cdots & \frac{\partial q_m}{\partial x_n}(u) \end{pmatrix} \cdot \begin{pmatrix} r_1 \\ \vdots \\ r_m \end{pmatrix}$$

Beispiel: Sei $A \in \mathbb{R}^n$ eine symmetrische $(n \times n)$-Matrix (das heißt $A = A^T$ oder $a_{ij} = a_{ji}$ für alle i, j). Betrachte Beispiel §15.24

$$f\colon \mathbb{R}^n \longrightarrow \mathbb{R}, x \mapsto x^T A x = \langle x | Ax \rangle = \sum_{i,j=1}^{n} a_{ij} x_i x_j$$

Nebenbedingung: $||x||_2 = 1$, das heißt $x^T \cdot x = \sum\limits_{i=1}^{n} x_j^2 = 1$, also $a(x) = x^T \cdot x - 1$, gesucht: Extrema. Wir müssen df und dq ausrechnen.

$$\begin{aligned} f(u+h) - f(u) &= (u+h)^T A(u+h) - u^T Au = h^T Au + u^T Ah + h^T Ah \\ &= 2u^T Ah + h^T Ah \quad \text{also} \quad df(u)(h) = 2u^T Ah \\ \text{also } df(u)(h) &= 2u^T Ah \quad \text{und} \quad dq(u)(h) = 2u^T h \end{aligned}$$

(vorige Rechnung mit $A = \mathbb{1}$). In einem Extremum mit Nebenbedingung $q = 0$ gilt also: $2ru^T = 2u^T A$ also $ru = Au$, das ist ein **Eigenwertproblem**. Also: Die Extrema von f unter der Nebenbedingung $q = 0$ sind Einheits-Eigenvektoren.

Vorlesung 22.06.2004

Satz: Sei $A \subseteq \mathbb{R}^n$ abgeschlossen und beschränkt (das heißt es gibt ein $R > 0$ so, dass $||a|| \leq R$ für alle $a \in A$). Falls $f\colon A \longrightarrow \mathbb{R}$ stetig ist, gibt es $m, M \in A$ so, dass $f(m) \leq f(a) \leq f(M)$ für alle $a \in A$. Insbesondere hat f ein Maximum und ein Minimum. Satz §15.25

Beweis: Wir zeigen, dass $M \in A$ so existiert. Angenommen, f ist nicht nach oben beschränkt, dann gibt es zu jedem $k \in \mathbb{N}$ ein $a_k \in A$ mit $f(a_k) \geq k$. Weil A beschränkt ist, ist $x_j = \{a_j \mid a = (a_1, a_2, \ldots, a_n) \in A\}$ beschränkt für $j = 1, 2, \ldots, n$. Wir finden eine Teilfolge a'_k so, dass die Folge $a_{k,1}$ konvergiert: $a'_k = (a'_{k,1}, a'_{k,2}, \ldots, a'_{k,n})$. Das machen wir nacheinander für alle Komponenten, dann erhalten wir eine Teilfolge a''_k für die alle Komponentenfolgen konvergieren, also existiert $\lim\limits_{k \to \infty} a''_k = \tilde{a} \in A$. Aber: $f(a''_k) \geq k$ und $\lim\limits_{k \to \infty} f(a''_k) = f(\tilde{a})$. Ein Widerspruch. Also existiert $L = \sup\{f(a) \mid a \in A\}$. Zu jedem $k \in \mathbb{N}$ gibt es $a_k \in A$ mit $f(a_k) \geq L - \frac{1}{k}$. Wir können wie vorher zu einer konvergenten Teilfolge a'_k übergehen. $f(a'_k) \geq L - \frac{1}{k}$, $\lim\limits_{k \to \infty} a_k = \tilde{a}$, $\lim\limits_{k \to \infty} f(a'_k) = f(\tilde{a}) = L$. Setze $M = \tilde{a}$, damit sind wir fertig. Die Existenz von $m \in A$ folgt, indem man die Funktion $-f$ betrachtet. $\square$

Insbesondere folgt also: Die Funktion $f(x) = \langle\, x \,|\, Ax \,\rangle$ hat auf der beschränkten Menge $\{x \in \mathbb{R}^n \mid ||x||_2 = 1\}$ ein Minimum *und* ein Maximum.

Folgerung: Jede reelle symmetrische Matrix $A \in \mathbb{R}^{n \times n}$ hat wenigstens einen Eigenvektor. (Genauer: $\mathbb{R}^n$ hat sogar eine Basis aus Eigenvektoren von A). Folgerung §15.26

16 Gewöhnliche Differentialgleichungen

16.1 Überblick

In einer gewöhnlichen Differentialgleichung tauchen nur Ableitungen nach *einer* reellen Variablen auf, in partiellen Differentialgleichungen tauchen dagegen *partielle* Ableitungen auf. Beispiele:

- Gewöhnliche Differentialgleichung: $y' = y^2$
- Partielle Differentialrechnung: $\frac{\partial^2 f}{\partial x \partial x}(x, y) + \frac{\partial^2 f}{\partial y \partial y}(x, y) = 0$

Paragraph §16.1

Viele Naturgesetze sind Differentialgleichungen, zum Beispiel **Newtons Gesetz**: $\vec{F} = m \cdot \vec{a}$. Das ist die Bewegung eines Teilchens im $\mathbb{R}^3$ in einem Kraftfeld ($\vec{F}$ gibt das Kraftfeld an, $\vec{a}$ die Beschleunigung).

Man kann das Gesetz als gewöhnliche Differentialgleichung 2. Ordnung (es tauchen 2. Ableitungen auf) schreiben:

$$F = m \cdot \ddot{c} \qquad \text{mit} \qquad c \colon \mathbb{R} \longrightarrow \mathbb{R}^3,\ t \mapsto c(t)$$

Ebenso kann man $F = m \cdot a$ eine gewöhnliche Differentialgleichung 1. Ordnung im $\mathbb{R}^6$ betrachten, indem man eine neue Funktion v mit $v = \dot{c}$ definiert. Es ergibt sich: $F = m \cdot \dot{v}$

16.2 Das Vektorfeld

Definition §16.2

Definition: Sei V ein Banachraum (z.B. $V = \mathbb{R}^n$) und $U \subseteq V$ offen. Eine stetige Abbildung $\xi \colon U \longrightarrow V$ heißt **Vektorfeld** auf U. Jedem $u \in U$ wird also stetig ein Vektor $\xi(u) \in V$ zugeordnet.

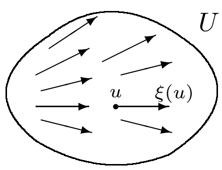


Abbildung 26: Das Vektorfeld

Beispiel:

- $\xi(u) = w_0 \in V$ **(konstantes Vektorfeld)**
- $\xi(u) = u$

16.3 Die Integralkurve

Definition §16.3

Definition: Eine stetig differenzierbare Kurve $c\colon J \longrightarrow V$ heißt **Integralkurve** (bezüglich ξ), falls für jedes $t \in J$ gilt: $\dot{c}(t) = \xi\big(c(t)\big)$

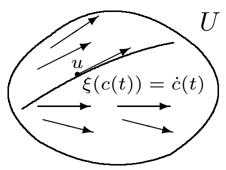


Abbildung 27: Die Integralkurve

Diese Bedingung ist eine gewöhnliche Differentialgleichung für c, die wir lösen wollen. Wir sehen später: *Jede* gewöhnliche Differentialgleichung lässt sich so formulieren.

Beispiel: Es werden wieder die Beispiele aus §16.2 betrachtet:

- $c(t) = u + t \cdot w_0 \;\Rightarrow\; \dot{c}(t) = w_0 = \xi\big(c(t)\big)$
- $\xi(u) = u \;\Rightarrow\; c(t) = ?$ (Hier ist die Lösung noch unklar!)

16.4 Lokaler Fluss

Wir fassen viele Integralkurven in einer Funktion zusammen.

Definition §16.4

Definition: Sei $\xi\colon U \longrightarrow V$ ein Vektorfeld, sei $u \in U$, $r > 0$ und sei $U_0 \subseteq U$ eine offene Teilmenge mit $u \in U_0$. Eine Funktion $\varphi\colon]-r, r[\times U_0 \longrightarrow U$ mit $(t, v) \mapsto \varphi(t, v)$ heißt **lokaler Fluss**, falls gilt:

- $\varphi(0, v) = v$ für alle $v \in U_0$
- Für jedes $v \in U_0$ ist die Kurve $c_v(t) = \varphi(t, v)$ eine Integralkurve für ξ.

Betrachte $c_v(0) = v$ für $v \in U_0$, $\dot{c}_v(t) = \xi\big(c_v(t)\big)$

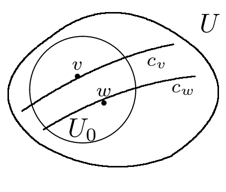


Abbildung 28: Lokaler Fluss

Beispiel: $\xi(u) = w_0$, $\varphi(t, v) = v + t \cdot w_0$ ist ein Fluss, denn:
$\varphi(0, v) = v$, $c_v(t) = v + t \cdot w_0$, $\dot{c}_v(t) = w_0 = \xi\big(c_v(t)\big)$

Für $r > 0$ setze $\overline{B_r}(x) = \{y \in V \mid ||y - x|| \leq r\}$. Das ist die **abgeschlossene r-Kugel** um v.

Satz (Lokale Existenz und Eindeutigkeit des Flusses): Sei $\xi\colon U \longrightarrow V$ ein L-Lipschitz-stetiges Vektorfeld und sei $u \in U$, dann gibt es $a, b > 0$ und einen lokalen Fluss $\varphi\colon]-b, b[\times B_a(u) \longrightarrow U$. Ist $\tilde{\varphi}\colon]-\tilde{b}, \tilde{b}[\times B_{\tilde{a}}(u) \longrightarrow U$ ein anderer Fluss, so gilt $\varphi(t,v) = \tilde{\varphi}(t,v)$ wann immer: Satz §16.5

$$(t,v) \in \big(]-b,b[\times B_a(u)\big) \cap \big(]-\tilde{b},\tilde{b}[\times B_{\tilde{a}}(u)\big)$$

Also für alle (u,v) im gemeinsamen Definitionsbereich.

Vorlesung 25.06.2004

Beweis: Wähle $1 > a > 0$ so, dass $\overline{B_{2a}}(u) \subseteq U$ (zum Beispiel $B_{3a}(u) \subseteq U$). Für $x \in \overline{B_{2a}}$ gilt dann $||\xi(x) - \xi(u)|| \le 2aL$, also $||\xi(x)|| \le ||\xi(u)|| + 2aL = k$. Wähle $b > 0$ so, dass $b < \frac{a}{k}$ und $b < \frac{1}{L}$. Setze $I = [-b, b]$ und

$$X_v = \{c\colon I \longrightarrow U \mid c \text{ stetig}, \ c(0) = v\} \quad \text{für } v \in \overline{B_{2a}}(u)$$

Die Menge X_v ist vollständig in der Supremumsnorm.
Setze $S_v(c)(t) = v + \int_0^t \xi(c(s))\,\mathrm{ds}$, $S_v(c)(0) = v$,

$$||S_v(c)(t) - v|| = ||\int_0^t \xi(c(s))\,\mathrm{ds}|| \le \int_0^t ||\xi(c(s))||\,\mathrm{ds} \le |t| \cdot k \le b \cdot k < a$$

Das heisst $S_v(c)(t) \in \overline{B_a}(v)$ und $S_v\colon X_v \longrightarrow X_v$ ist eine stetige Abbildung. Für $c_1, c_2 \in X_v$:

$$||S_v(c_1)(t) - S_v(c_2)(t)|| = ||\int_0^t \big(\xi(c_1(s)) - \xi(c_2(s))\big)\,\mathrm{ds}||$$

$$\le \int_0^t \underbrace{||\xi(c_1(s)) - \xi(c_2(s))||}_{L\cdot||c_1(s)-c_2(s)||}\,\mathrm{ds} \le L \cdot ||c_1 - c_2||_\infty \cdot b$$

also S_v ist Lb-Lipschitz-stetig und wegen $Lb < 1$ kontrahierend. Nach Banachs Fixpunktsatz gibt es $c_v \in X_v$ mit $S_v(c_v) = c_v$. $c_v(t) = v + \int_0^t \xi(c_v(s))\,\mathrm{ds}$, das ist stetig differenzierbar.

$$\dot{c}_v(t) = \xi\big(c_v(t)\big), \ c_v(0) = v$$

Setze $\varphi(t,v) = c_v(t)$, das ist der gewünschte Fluss; weil c_v ein Fixpunkt ist, ist c_v eindeutig bestimmt nach Banachs Fixpunktsatz. $\square$

16.5 Picard-Iteration

Bemerkung: Der Satz liefert einen *Algorithmus* zur näherungsweisen Lösung der Differentialgleichung: Setze $c_0(t) = v$, $c_1(t) = S_v(c_o)$, $c_2(t) = S_v(c_1)$, …, $c_{n+1}(t) = S_v(c_n)$. Das Verfahren heißt **Picard-Iteration**. Bemerkung §16.6

Beispiel: $U = V = \mathbb{R}$, $\xi(u) = u$, DGL $\dot{c} = c = \xi(c)$
$c_0(t) = v$
$c_1(t) = v + \int_0^t c_0(s)\,\mathrm{ds} = v + \int_0^t v\,\mathrm{ds} = v + vt$
$c_2(t) = v + \int_0^t (v + vs)\,\mathrm{ds} = v + vt + \frac{1}{2}vt^2$
$c_n(t) = v(1 + t + \frac{1}{2}t^2 + \frac{1}{6}t^3 + \ldots)$
Das konvergiert gegen $c_v(t) = v\exp(t)$, $\dot{c}_v(t) = v\exp(t) = c_v(t)$, $c_v(0) = v$

Satz §16.7

Satz (gleiche Voraussetzungen und Bezeichnungen wie in Satz §16.5): Die Abbildung $\overline{B_a}(u) \longrightarrow C(I,V), v \mapsto c_v$ ist Lipschitz-stetig, und der Fluss φ ist stetig.

Beweis: Sei $x, y \in \overline{B_a}(u)$. Dann gilt

$$\begin{aligned} & ||S_x(c_y)(t) - c_y(t)|| = ||S_x(c_y)(t) - S_y(c_y)(t)|| \\ = \quad & ||x - y + \int_0^t \big(\xi(c_y(s)) - \xi(c_y(s))\big) \,\mathrm{ds}|| = ||x-y|| \end{aligned}$$

Setze $R = b \cdot L < 1$, dann gilt:

$$\begin{aligned} ||S_x^n(c_y) - c_y|| &\leq ||S_x^n(c_y) - S_x^{n-1}(c_y)|| + ||S_x^{n-1}(c_y) - S_x^{n-2}(c_y)|| \\ &+ \ldots + ||S_x(c_y) - c_y|| \leq (R^{n-1} + R^{n-2} + \ldots + R + 1)||x-y|| \\ &\Longrightarrow ||c_x - c_y||_\infty \leq \tfrac{1}{1-R}||x-y|| \end{aligned}$$

Damit ist φ stetig:

$$\begin{aligned} & ||\varphi(t,v) - \varphi(t+t_1, v+v_1)|| \leq ||\varphi(t,v) - \varphi(t+t_1, v)|| \\ + \quad & ||\varphi(t+t_1, v) - \varphi(t+t_1, v+v_1)|| \\ = \quad & \underbrace{||c_v(t) - c_v(t+t_1)||}_{\substack{\text{klein für } t_1 \text{ klein,} \\ \text{weil } c_v \text{ stetig ist}}} + \underbrace{||c_v(t+t_1) - c_{v+v_1}(t+t_1)||}_{\substack{\text{klein für } t_1 \text{ klein,} \\ \text{wegen oben}}} \end{aligned}$$

16.6 Lokale Lipschitz-Stetigkeit

Definition §16.8

Definition: Eine Abbildung $f\colon U \longrightarrow W$ heißt **lokal Lipschitz-stetig** falls es zu jedem $u \in U$ ein $r > 0$ und ein L gibt, so dass für alle $x, y \in B_r(u) \cap U$ gilt:

$$||f(x) - f(y)|| \leq L||x-y||$$

Solch eine Abbildung ist stetig, und eine Lipschitz-stetige Abbildung ist lokal Lipschitz-stetig.

Satz §16.9

Satz: Ist $f\colon U \longrightarrow W$ stetig differenzierbar, dann ist f lokal Lipschitz-stetig.

Beweis: Sei $u \in U$, wähle $\overline{r} > 0$ so, dass $B_{\overline{r}}(u) \subseteq U$. Weil f stetig differenzierbar ist, gibt es $r \leq \overline{r}$ so, dass $||Df(x) - Df(u)|| \leq 1$ für alle $x \in B_r(u)$. Es folgt $||Df(x)|| \leq ||Df(u)|| + 1 = L$. Dann gilt für $x, y \in B_r(u)$:

$$\begin{aligned} & ||f(x) - f(y)|| = ||\int_0^1 Df(sx + (1-s)y)(x-y) \,\mathrm{ds}|| \\ \leq \quad & \int_0^1 \underbrace{||Df(sx + (1-s)y)||}_{\leq L} \cdot ||x-y|| \,\mathrm{ds} \leq L \cdot ||x-y|| \end{aligned}$$

Beispiel: $f(x) = x^2$ ist nicht Lipschitz-stetig, aber lokal Lipschitz-stetig.

$$\begin{array}{ccccc} \text{L.S.} & \Rightarrow & \text{lokal L.S.} & \Rightarrow & \text{stetig} \\ & \not\Leftarrow & \Uparrow \ \not\Downarrow & \not\Leftarrow & \\ & & \text{stetig differenzierbar} & & \end{array}$$

Folgerung: Ist ξ lokal Lipschitz-stetig (zum Beispiel ξ stetig differenzierbar) und ist $u \in U$, dann gibt es einen Fluss $\varphi\colon]-b, b[\times B_a(u) \longrightarrow U$, der lokal eindeutig bestimmt ist. Folgerung §16.10

Satz (Existenz von globalen Lösungen): Sei $\xi\colon U \longrightarrow V$ ein lokal Lipschitz-stetiges Vektorfeld, sei $u \in U$, seien c_1, c_2 Integralkurven: Satz §16.11
$c_i\colon J_i \longrightarrow U$, $c_1(0) = u = c_2(0)$. Dann gilt:

$$c_1(t) = c_2(t) \quad \text{für alle } t \in J_1 \cap J_2$$

Beweis: Sei $Q = \{b > 0 \mid c_1(t) = c_2(t) \text{ gilt für alle } 0 \leq t \leq b\}$. Dann ist $Q \neq \emptyset$ nach Satz §16.5. Beachte: Ist $0 \leq \tilde{b} \leq b \in Q \Rightarrow \tilde{b} \in Q$.

1. Fall: $Q =]0, \infty[$ dann $c_1(t) = c_2(t)$ für *alle* $t \geq 0$
2. Fall: $b = \sup(Q) < \infty$. Wenn die Behauptung falsch wäre, gäbe es $b_1 \in J_1 \cap J_2$ mit $b_1 > b$. Betrachte $\tilde{c}_i(t) = c_i(t + b)$. Das sind zwei Integralkurven, $\tilde{c}_1(0) = \tilde{c}_2(0)$. Nach Satz §16.5 gilt $\tilde{c}_1(t) = \tilde{c}_2(t)$ für ein kleines $t > 0$. (Widerspruch)

   Gleiches Argument für $t < 0$.

Folgerung: Ist $\xi\colon U \longrightarrow V$ ein lokal Lipschitz-stetiges Vektorfeld und $u \in U$. Setze $J(u) = \cup\{J \mid \text{ es gibt eine Integralkurve } c\colon J \longrightarrow V \text{ mit } c(0) = u\}$ Das ist ein offenes Intervall, nach Satz §16.11 gibt es eine mögliche Lösung der Differentialgleichung $\dot{c} = \xi(c)$ zum Anfangswert $c(0) = u$. Folgerung §16.12

Beispiel: Beispiel §16.13

(a) $U = \mathbb{R}$, $\xi(u) = u$, Lösung $c_v(t) = v\exp(t)$, $J(v) = \mathbb{R}$ für *jedes* $v \in R$.

(b) $\xi(u) = -u^2$, $U = \mathbb{R}$, $\dot{c}(t) = -c(t)^2$.
Raten den Lösung: $c(t) = \frac{1}{t+t_0} = (t + t_0)^{-1}$, $c(0) = \frac{1}{t_0}$,
$\dot{c}(t) = -(t + t_0)^{-2} = \frac{-1}{(t+t_0)^2} = -c(t)^2$

Anfangswert: $u = (c_0) = \frac{1}{t_0}$, $u \neq 0$. $c_n(t) = \frac{1}{t+\frac{1}{u}}$.

für $u > 0$ $\quad J(u) =]-\frac{1}{u}\infty[$
für $u < 0$ $\quad J(u) =]-\infty, -\frac{1}{u}[$
für $u = 0$ $\quad c(t) = 0 = const$

16.7 Die klassische Differentialgleichung

Vorlesung 29.06.2004

Sei $F\colon U \longrightarrow \mathbb{R}^n$ eine lokal Lipschitz-stetige (reell- oder vektorwertige) Funktion in mehreren Argumenten mit $U \subseteq \mathbb{R}^m$. Eine gewöhnliche Differentialgleichung im *klassischen* Sinne sieht so aus: Paragraph §16.14

$$c^{(k)}(t) = F\big(c(t), \dot{c}(t), \ddot{c}(t), \ldots, c^{(k-1)}(t), t, p\big) \quad \text{mit konstantem Parameter } p$$

Wie passt das zu Flüssen, Integralkurven und Vektorfeldern?

Betrachte zuerst den Fall $k = 1$, $\dot{c}(t) = F\big(c(t), t, p\big)$. Paragraph §16.15

Definiere: $\xi(u_1, s, u_2) = \big(F(u_1, s, u_2), 1, 0\big)$.

Klassisch: $F\colon U_1 \times I \times U_2 \longrightarrow V$ mit $U_1 \subseteq V$ offen, $I \subseteq \mathbb{R}$ ein offenes Intervall mit $0 \in I$ und $U_2 \subseteq W$ offen.
$\xi\colon U_1 \times I \times U_2 \longrightarrow V \times \mathbb{R} \times W$, also ist ξ ein lokal Lipschitz-stetiges Vektorfeld, definiert auf der offenen Menge $U_1 \times I \times U_2 \subseteq V \times \mathbb{R} \times W$.

Sei $\tilde{c}$ eine Lösung der Differentialgleichung $\dot{\tilde{c}}(t) = \xi\big(\tilde{c}(t)\big)$.
Schreibe $\tilde{c} = (c_1, c_2, c_3)$. Lösung zum Startwert $\tilde{c} = (u, 0, p)$:

$\dot{c}_2(t) = 1, \quad c_2(0) = 0 \quad \Rightarrow \quad c_2(t) = t$
$\dot{c}_3(t) = 0, \quad c_3(o) = p \quad \Rightarrow \quad c_3(t) = p = const$
$\dot{c}_1(t) = F\big(c_1(t), c_2(t), c_3(t)\big) = F\big(c_1(t), t, p\big)$

Das heißt die erste Komponente c_1 der Integralkurve $\tilde{c} = (c_1, c_2, c_3)$ ist genau die Lösung der Differentialgleichung: $\dot{c}_1(t) = F\big(c_1(t), t, p\big)$, schreibe $c(t) = c_1(t)$. Die Differentialgleichung $\dot{c} = F\big(c, t, p\big)$ hat also eine eindeutige Lösung zum Anfangswert $c(0) = u$.

Bemerkung: Wir haben immer den Fall betrachtet, dass $c(0)$ als Anfangswert vorgegeben wurde. Das reicht auch vollkommen aus: Falls man $\dot{c}(t) = \xi\big(c(t)\big)$ mit $c(t_0) = u$ lösen will, setze $\tilde{c} = c(t + t_0)$, dann gilt $\dot{\tilde{c}}(t) = \dot{c}(t + t_0)$ (Kettenregel). Löst also $\tilde{c}$ die Differentialgleichung Bemerkung §16.16

$$\dot{\tilde{c}}(t) = \xi\big(\tilde{c}(t)\big), \quad \tilde{c}(0) = u$$

dann ist $c(t) = \tilde{c}(t - t_0)$ eine Integralkurve mit $c(t_0) = \tilde{c}(0) = u$.

Eine *klassische* gewöhnliche Differentialgleichung k-ter Ordnung Paragraph §16.17

$$c^{(k)}(t) = F\big(c(t), \dot{c}(t), \dots, c^{(k-1)}(t), t, p\big)$$

mit den Anfangswerten $c(0) = v_0$, $\dot{c}(0) = v_1$, …, $c^{(k-1)}(0) = v_{k-1}$ und

$$F\colon U_0 \times U_1 \times U_2 \times \cdots \times U_{k-1} \times I \times \tilde{U} \longrightarrow V \quad \text{lokal Lipschitz-stetig}$$

Setze $\xi(u_0, u_1, \dots, u_{k-1}, s, w) = (u_1, u_2, \dots, u_{k-2}, F(u_1, \dots, w), 1, 0)$
dann ist ξ ein lokal Lipschitz-stetiges Vektorfeld:

$$\xi\colon U_0 \times U_1 \times \cdots \times U_{k-1} \times I \times \tilde{U} \longrightarrow \underbrace{V \times \cdots \times V}_{k\text{-Stück}} \times \mathbb{R} \times W$$

Sei $\tilde{c}$ die zugehörige Integralkurve, $\dot{\tilde{c}}(t) = \xi(\tilde{c}(t))$, $\tilde{c}(t) = \big(c_0(t), \dots, c_{k-1}(t), t, p\big)$ und $\dot{c}_0 = c_1 = c_0^{(1)}$, $\dot{c}_1 = c_2 = c_0^{(2)}$, …, $\dot{c}_{k-2} = c_{k-1} = c_0^{(k-1)}$, $\dot{c}_{k-1} = F(c_0, c_1, \dots, c_{k-1}, t, p) \Longrightarrow c^{(k)}(t) = F\big(c(t), \dot{c}(t), \dots, c^{(k-1)}(t), t, p\big)$, $c = c_0$.

Schlagwort: Eine Differentialgleichung k-ter Ordnung wird auf ein System von k Differentialgleichungen 1. Ordnung zurückgeführt.

Beispiel (a): $\ddot{c} + c = 0$: $c_1 = c$, $c_2 = \dot{c}_1$, $\begin{pmatrix} \dot{c}_1 \\ \dot{c}_2 \end{pmatrix} = \begin{pmatrix} c_2 \\ -c_1 \end{pmatrix}$ also Beispiel §16.18

$$\xi(u_1, u_2) = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}, \ (\dot{c}_1, \dot{c}_2) = \xi(c_1, c_2)$$

Das ist eine Differentialgleichung der Form $\xi(u) = A(u)$ mit $A\colon V \longrightarrow V$ linear. Man spricht von einer **linearen Differentialgleichung**. Hier *raten* wir: sin und cos haben die Eigenschaft $\ddot{c} + c = 0 \Leftrightarrow \ddot{c} = -c$. Ansatz mit $a, b \in \mathbb{R}$:

$$\begin{aligned} c(t) &= a \cdot \sin(t) + b \cdot \cos(t) \\ \dot{c}(t) &= a \cdot \cos(t) - b \cdot \sin(t) \\ \ddot{c}(t) &= -a \cdot \sin(t) - b \cdot \cos(t) = -c(t) \end{aligned}$$

Anfangswerte: $c(0) = b$, $\dot{c}(0) = a$.
Bei einer Differentialgleichung k-ter Ordnung gibt man vor: $c(0)$, $\dot{c}(0)$, $\ldots$, $c^{(k-1)}(0)$, also gibt man k verschiedene Anfangswerte vor.
Zu den Anfangswerten $c(0) = b$, $\dot{c}(0) = a$ ist also $c(t) = a \cdot \sin(t) + b \cdot \cos(t)$ die eindeutig bestimmte Lösung. Es gilt: $J(b, a) = \mathbb{R}$.

Beispiel (b) für eine (nicht lineare) Differentialgleichung: $F(u, t) = g(u) \cdot f(t)$ Beispiel §16.19
(Differentialgleichung mit getrennten Variablen).
Angenommen, g hat keine Nullstellen, $u \in U \subseteq \mathbb{R}$:

$$\dot{c}(t) = g\big(c(t)\big) \cdot f(t) \quad \Longrightarrow \quad \frac{\dot{c}(t)}{g\big(c(t)\big)} = f(t)$$

Setze $F(t) = \int_0^t f(s)\,\mathrm{ds}$, $H(v) = \int_u^v \frac{1}{g(x)}\,\mathrm{dx}$, Anfangswert $c(0) = u$.
Behauptung: Für die Lösung c gilt $F(t) = H(c(t))$, denn

$$f(t) = \frac{d}{dt} \int_u^{c(t)} \frac{1}{g(x)}\,\mathrm{dx} = \frac{d}{dt} \left(\int_0^t \frac{1}{g(c(t))} \cdot \dot{c}(s)\,\mathrm{ds} \right) = \frac{1}{g(c(t))} \cdot \dot{c}(t)$$

Für $t = 0$ haben wir $F(0) = 0 = H(c(0)) = \int_u^v \frac{1}{g(x)}\,\mathrm{dx} = 0$

Beispiel (c): $F(u, t) = u^2$ also $g(u) = u^2$, $f(t) = 1$. Das ist eine Differentialgleichung mit getrennten Variablen. $F(t) = t$.

$$\begin{aligned} H(v) = \int_u^v \frac{1}{x^2}\,\mathrm{dx} = \left[-\frac{1}{x} \right]_{x=u}^{x=v} &= \frac{1}{u} - \frac{1}{v} &\Longrightarrow \quad t = \frac{1}{u} - \frac{1}{c(t)} \\ \frac{1}{c(t)} = \frac{1}{u} - t = \frac{1}{u} - \frac{ut}{u} &= \frac{1 - ut}{u} &\Longrightarrow \quad c(t) = \frac{u}{1 - ut} \end{aligned}$$

Das ist die Lösung der Differentialgleichung.

Vorlesung 02.07.2004

Beispiel (d): $\ddot{c} = F(c)$ (zum Beispiel Newtons Gesetz für ein zeitunabhängiges Kraftfeld), $U \subseteq \mathbb{R}$ offen, $F\colon U \longrightarrow \mathbb{R}, u \in U$. Ansatz: $V(x) = -\int_u^x F(r)\,\mathrm{d}r$

$$
\begin{aligned}
\rightsquigarrow \quad & \ddot{c}(t) + \frac{\partial V}{\partial x}\big(c(t)\big) = 0 \\
& \ddot{c}(t)\dot{c}(t) + \dot{c}(t)\frac{\partial V}{\partial x}\big(c(t)\big) = 0 \\
\Rightarrow \quad & \frac{1}{2}\dot{c}(t)^2 + V\big(c(t)\big) = E = \mathit{const} \\
\Rightarrow \quad & \dot{c}(t) = \sqrt{2E - 2V(c(t))}
\end{aligned}
$$

Das ist der Energieerhaltungssatz (kinetische Energie + potentielle Energie = *const*), eine Differentialgleichung in getrennten Variablen, die wir wie in (c) lösen können: $H(x) = \int_v^x \frac{1}{\sqrt{2E-2V(r)}}\,\mathrm{d}r$

Beispiel (e): $\ddot{c} = -kc, \quad k > 0$ reelle Konstante (Federgleichung)
$u = 0 = c(0), \quad \dot{c}(0) = c_0 > 0$
Potential: $V(x) = \int_u^x kr\,\mathrm{d}r = \frac{1}{2}x^2k$
$\dot{c}(t) = \sqrt{2E - c(t)^2k}, \quad E = \frac{1}{2}c_0^2$
$\dot{c}(t) = \sqrt{c_0^2 - c(t)^2k}, F(t) = t$
$H(x) = \int_0^x \frac{1}{c_0^2 - kr^2}\,\mathrm{d}r = \frac{1}{\omega}\arcsin(x\frac{\omega}{c_0})$ mit $\omega = \sqrt{k}, \quad t = \frac{1}{\omega}\arcsin\big(c(t)\frac{\omega}{c_0}\big)$
$\Rightarrow \sin(\omega t) = c(t)\frac{\omega}{c_0} \Rightarrow c(t) = \frac{c_0}{\omega}\sin(\omega t)$
Diese Funktion löst die Differentialgleichung für alle t.

Faustregel: Der Existenz- und Eindeutigkeitssatz garantiert die Existenz und Eindeutigkeit der Lösung. Mit solchen Tricks und Methoden, wie in den Beispielen (d) und (e) sucht man eine Lösung ohne auf Einzelheiten im Definitionsbereich zu achten, sondern prüft ganz zum Schluss, ob die gefundene Funktion eine Lösung ist.

Weitere Literatur:

- Walter, Gewöhnliche Differentialgleichungen
- Braun, Gewöhnliche Differentialgleichungen

17 Lineare DGL und globaler Fluss

17.1 Der globale Fluss

Definition §17.1

Definition: Wir betrachten eine lokal Lipschitz-stetige Funktion $\xi\colon U \longrightarrow V$ mit $U \subseteq V$ offen und V Banachraum. Zu $u \in U$ hatten wir das *maximale* Lösungsintervall $J(u) \subseteq \mathbb{R}$ zur maximalen Integralkurve $c_u\colon J(u) \longrightarrow U$ mit $c_u(0) = u$. Setze $\Omega = \bigcup J(u) \times \{u\} \subseteq \mathbb{R} \times U$. Definiere $\varphi\colon \Omega \longrightarrow U,\ (t,u) \mapsto \varphi(t,u) = c_u(t)$. Wir nennen φ den **globalen Fluss** zu ξ und Ω seinen **Definitionsbereich**. Für $(t,u) \in \Omega$ schreibe $c_u(t) = \varphi(t,u) = \varphi_t(u)$.

Beispiel §17.2

Beispiel:

- $U = \mathbb{R}, \quad \xi(u) = u, \quad c_u(t) = ue^t = \varphi_t(u),$
  $J(u) = \mathbb{R} \Rightarrow \Omega = \mathbb{R} \times \mathbb{R}$
- $U = \mathbb{R}, \quad \xi(u) = -u^2,$

$$J(u) = \begin{cases}]-\frac{1}{u}, \infty[& \text{falls } u > 0 \\ ]-\infty, -\frac{1}{u}[& \text{falls } u < 0 \\ \mathbb{R} & \text{falls } u = 0 \end{cases}$$

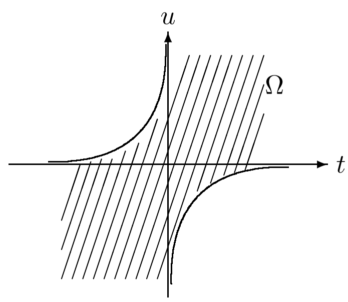


Abbildung 29: Globaler Fluss

Satz §17.3

Satz: Sei $\xi\colon U \longrightarrow V$ lokal Lipschitz-stetig. Sei $(u,t) \in \Omega$. Dann gilt

$$J\big(\varphi_t(u)\big) = \{s - t \mid s \in J(u)\} = J(u) - t$$

und $\varphi_{s+t}(u) = (\varphi_s \circ \varphi_t)(u) = (\varphi_t \circ \varphi_s)(u)$.

Beweis: Betrachte $c_1(t) = \varphi_t\big(\varphi_s(u)\big)$ und $c_2(t) = \varphi_{s+t}(u)$
$c_1(0) = \varphi_s(u) = c_u(s) = \varphi_s(u) = c_2(0)$
$\dot{c}_1(t) = \xi\big(c_1(t)\big)$ und $\dot{c}_2(t) = \xi\big(c_2(t)\big)$
Mit dem Satz über die Eindeutigkeit der Lösung folgt: $c_1(t) = c_2(t)$ $\square$

Satz §17.4

Satz: $\Omega \subseteq \mathbb{R} \times V$ ist offen und $\varphi\colon \Omega \longrightarrow U$ ist stetig.

Beweis: Sei $(t,u) \in \Omega$, sei $Q = \{r > 0 \mid$ für jedes $t \in]0,r[$ gibt es $a,b > 0$ mit $]\,t-b, t+b\,[\times B_a(u) \subseteq \Omega$ und φ ist stetig in $(t,v)\}$. Nach §16.7 ist $Q \neq \emptyset$.
1. Fall: $Q =]0,\infty[\quad \to$ OK.
2. Fall: $Q \neq]0,\infty[$, also existiert $q = \sup(Q) < \infty$

Behauptung: $q \notin J(u) \quad \to$ OK.
Angenommen $q \in J(u)$. Setze $\overline{u} = \varphi_q(u) \in U$, es gibt einen lokalen Fluss $\tilde{\varphi}$ nahe $\overline{u}$, $\tilde{\varphi}\colon]-r,r[\times B_r(\overline{u}) \longrightarrow U$. Weiter gibt es $\delta > 0$ so, dass $t \in]q-\delta, q[\Rightarrow \varphi_t(u) \in B_{\frac{r}{4}}(\overline{u})$ (weil Integralkurven stetig sind). Wähle $t_1 < q$ mit $q - \delta < t_1$ und $q - t_1 < \frac{r}{4}$. Es gibt dann $s > 0$ mit $\varphi\big(]t_1 - s, t_1 + s[\times B_{\frac{r}{2}}(\overline{u})\big) \subseteq B_{\frac{r}{2}}(\overline{u})$. Für $t \in]\,t_1 - r, t_1 + r[$ setze $\psi(t,v) = \tilde{\varphi}\big(t - t_1, \varphi(t_1, v)\big) \Rightarrow \psi(t_1, v) = \varphi(t_1, v)$. Damit ist φ doch in (q,u) stetig. Ein Widerspruch. $\square$

17.2 Die Exponentialfunktion

Vorlesung 06.07.2004
Definition §17.5

Definition: Es sei V ein Banachraum (z.B. $V = \mathbb{R}^n$), $T\colon V \longrightarrow V$ stetig und linear, also $T \in L(V,V)$. Setze $g_n(T) = \mathrm{id}_V + T + \frac{1}{2}T^2 + \cdots + \frac{1}{n!}T^n = \sum_{k=0}^n \frac{1}{k!}T^k$, das ist eine stetige lineare Abbildung, $g_n \in L(V,V)$, denn g_n ist linear und Summe stetiger Abbildungen.
Wähle $R > 0$, setze $X_R = \big\{T \in L(V,V) \;\big|\; ||T|| \le R\big\}$, $g_n\colon X_R \longrightarrow L(V,V)$ ist stetig für jedes $n \in \mathbb{N}$. Es gilt:

$$\Big|\Big|g_N(T) - g_{N-\ell}(T)\Big|\Big| = \Big|\Big|\sum_{k=N+1}^{N+\ell} \frac{1}{k!}T^k\Big|\Big| \le \sum_{k=N+1}^{N+\ell} \frac{1}{k!}||T||^k \le \sum_{k=N+1}^{N+\ell} \frac{1}{k!}R^k$$

$$\Rightarrow ||g_N - g_{N-\ell}||_\infty \le \sum_{N+1}^{N+\ell} \frac{1}{k!}R^k$$

Das wird beliebig klein, falls N groß genug ist. Die Folge $(g_n)_{n\in\mathbb{N}}$ ist also eine Cauchy-Folge, der Grenzwert $\sum\limits_{k=0}^{\infty} \frac{1}{k!}T^k = \exp(T) = e^T$ ist eine stetige lineare Abbildung $X_R \longrightarrow L(V,V)$. Das gilt für jedes $R > 0$, also ist die Abbildung $\exp\colon L(V,V) \longrightarrow L(V,V)$ *stetig*.

Sei $A \in L(V,V)$.

$$A\int_0^t g_n(sA)\,\mathrm{ds} = A\int_0^t (A + As + A\frac{s^2}{2} + \cdots)\,\mathrm{ds} = At + \frac{1}{2}At^2 + \cdots = g_{n+1}(tA) - 1$$

Im Grenzwert $n \to \infty$ ergibt das: $A\int_0^t e^{As}\,\mathrm{ds} = e^{At} - 1$
Die linke Seite ist stetig differenzierbar, also gilt:

$$\frac{\mathrm{d}}{\mathrm{dt}}(e^{At} - 1) = Ae^{At} \quad \Rightarrow \quad \frac{\mathrm{d}}{\mathrm{dt}}e^{At} = Ae^{At} = e^{At}A$$

Betrachte das Vektorfeld ξ auf $L(V,V) \times L(V,V)$ mit $\xi(X,Y) = (XY, 0)$. Es ergibt sich die Differentialgleichung $(\dot{X}, \dot{Y}) = (XY, 0) \Rightarrow Y = Y_0 = const$
Lösung: $X(t) = X_0 e^{Y_0 t}$, $\dot{X}(t) = X_0 e^{Y_0 t} Y_0 = X(t)Y(t)$ mit $X(0) = X_0 e^{Y_0 \cdot 0} = X_0 \cdot 1 = X_0$
Der zugehörige globale Fluss ist $\varphi_t(X,Y) = (Xe^{tY}, Y)$. Nach Satz §17.3 gilt $\varphi_{s+t} = \varphi_s \circ \varphi_t$, $(Xe^{(t+s)Y}, Y) = \varphi_s(Xe^{tY}, Y) = (Xe^{tY}e^{sY}, Y)$. Insbesondere gilt: $e^{(s+t)Y} = e^{sY} \cdot e^{tY}$

Satz §17.6

Satz: Sei V ein Banachraum, $A \in L(V,V)$ und $s,t \in \mathbb{R}$ beliebig. Es gilt stets: $e^{(s+t)A} = e^{sA} \cdot e^{tA}$. Insbesondere gilt $e^{sA} \cdot e^{-sA} = \mathrm{id}_V$. Damit ist e^{sA} immer *invertierbar.*

Beweis: Siehe oben. □

Achtung: Es gilt im Allgemeinen *nicht*: $e^{A+B} = e^A \cdot e^B$.

17.3 Homogene lineare Differentialgleichungen

Definition §17.7

Definition: Sei $I \subseteq \mathbb{R}$ ein offenes Intervall (mit $0 \in I$), sei $A\colon I \longrightarrow L(V,V)$ lokal Lipschitz-stetig. Die Differentialgleichung $\dot{c}(t) = A(t) \cdot c(t)$ heißt **homogene lineare Differentialgleichung**.

Beispiel §17.8

Beispiel: Ist $A \in L(V,V)$ fest, $\alpha\colon I \longrightarrow \mathbb{R}$ lokal Lipschitz-stetig, $A(t) = A\alpha(t)$, so gilt für $c(t) = e^{\int_0^t \alpha(s)\,\mathrm{d}s A} c_0$ mit einem festen Vektor $c_0 \in V$. Es gilt $c(0) = c_0$, $\dot{c}(t) = \alpha(t) \cdot A \cdot e^{\int_0^t \alpha(s)\,\mathrm{d}s A} c_0 = A(t) \cdot c(t)$

Also ist c die Integralkurve dieser Differentialgleichung zum Anfangswert c_0. Wichtiger Spezialfall: Ist $\alpha = const$, $\dot{c}(t) = Ac(t)$, so spricht man von einer **homogenen linearen Differentialgleichung** mit **konstanten Koeffizienten**.

Beispiel §17.9

Beispiel (a): $f\colon I \longrightarrow \mathbb{R} = V, \quad \dot{c}(t) = f(t) \cdot c(t) \quad \Rightarrow \quad c(t) = e^{\int_0^t f(s)\,\mathrm{d}s}$
Konkret: $f(t) = t^n, \quad \dot{c}(t) = t^n \cdot c(t), \quad c(t) = e^{\int_0^t s^n\,\mathrm{d}s} = e^{\frac{1}{n+1}t^{n+1}}$

Beispiel (b): $\ddot{c}(t) = -c(t), \quad c(t) = c_1(t), \quad \dot{c}(t) = c_2(t)$
$\Rightarrow \dot{c}_2(t) = -c_1(t), \quad \dot{c}_1(t) = c_2(t)$

$$\dot{\tilde{c}}(t) = \begin{pmatrix} \dot{c}_1(t) \\ \dot{c}_2(t) \end{pmatrix} = \begin{pmatrix} c_2(t) \\ -c_1(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix} = A\tilde{c}(t)$$

Lösung: $\tilde{c}(t) = e^{At}\tilde{c}_0$. Was also ist e^{At}?

$$A^0 = E,\ A^1 = A,\ A^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = -E,\ A^3 = A^2 A = -A,$$

$$A^4 = (A^2)^2 = E,\ A^{4k} = A^k$$

$$e^{tA} = E + tA + t^2 A^2 \tfrac{1}{2} + t^3 A^3 \tfrac{1}{6} + \cdots = \begin{pmatrix} \sum\limits_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!} & \sum\limits_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!} \\ -\sum\limits_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!} & \sum\limits_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!} \end{pmatrix}$$

$$= \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$

$$\tilde{c}(t) = e^{At}\tilde{c}(0)$$

$$\begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix} = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} c_1(0) \\ c_2(0) \end{pmatrix} \text{ und } \begin{pmatrix} c_1(0) \\ c_2(0) \end{pmatrix} = \begin{pmatrix} c(0) \\ \dot{c}(0) \end{pmatrix}$$

Also: $c(t) = c(0)\cos(t) + \dot{c}(0)\sin(t)$

Beispiel (c): Sei $A = \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix}$ eine obere Dreiecksmatrix. Was ist e^{At}?
$A^0 = E,\ A^1 = A,\ A^2 = \begin{pmatrix} 1 & 2r \\ 0 & 1 \end{pmatrix},\ \ldots,\ A^k = \begin{pmatrix} 1 & kr \\ 0 & 1 \end{pmatrix}$
$e^{tA} = E + At + \frac{1}{2}A^2t^2 + \cdots = \begin{pmatrix} e^t & e^t tr \\ 0 & e^t \end{pmatrix}$, denn $\sum\limits_{k=0}^{\infty} \frac{1}{k!}t^k kr = tr \sum\limits_{k=0}^{\infty} \frac{1}{k!}t^k$.

17.4 Die Exponentialfunktion und Matritzen

Vorlesung 09.07.2004

$$\exp\left(\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} t\right) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}$$
$$\exp\left(\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} t\right) = \begin{pmatrix} e^t & e^t tr \\ 0 & e^t \end{pmatrix}$$

Paragraph §17.10

Im Allgemeinen ist es schwierig e^{At} für eine Matrix A auszurechnen. Es gibt ein paar hilfreiche Regeln:

(i) Ist $A = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{pmatrix}$ und A_ℓ quadratische Matrix, dann

$$e^{At} = \begin{pmatrix} e^{A_1t} & 0 & \cdots & 0 \\ 0 & e^{A_2t} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{A_rt} \end{pmatrix}$$

Insbesondere ist $A = \begin{pmatrix} a_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & a_{nn} \end{pmatrix}$ Diagonalmatrix, dann

$$e^{At} = \begin{pmatrix} e^{a_{11}t} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{a_{nn}t} \end{pmatrix}$$

(ii) Ist v ein **Eigenvektor** von A, das heißt $Av = av$ und $a \in \mathbb{R}$, $v \neq 0$, dann $e^{At}v = e^{at}v$, insbesondere ist v ein Eigenvektor von e^{At} zum **Eigenwert** e^{at}.

(iii) Ist B invertierbar, so gilt $(B^{-1}AB)^k = B^{-1}A^kB$, also $B^{-1}e^{At}B = e^{B^{-1}AtB}$. Das heißt zum Beispiel: Ist $D = B^{-1}AB$ eine Diagonalmatrix, $D = \begin{pmatrix} d_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & d_n \end{pmatrix}$

$$\Rightarrow \quad e^{tD} = e^{tB^{-1}AB} = \begin{pmatrix} e^{td_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{td_n} \end{pmatrix} = B^{-1}e^{tA}B$$

$$\Rightarrow \quad e^{tA} = B \begin{pmatrix} e^{td_1} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & e^{td_n} \end{pmatrix} B^{-1}$$

Beispiel: DGL $y'' + 4y' - 2y = 0$. Umschreiben auf DGL 1. Ordnung: $c_2 = \dot{c}_1$, $c_1 = y$, $\dot{c}_2 + 4c_2 - 2c_1 = 0$ Beispiel §17.11

$$\dot{c}_1 = c_2, \quad \dot{c}_2 = 2c_1 - 4c_2$$

$$\underbrace{\begin{pmatrix} \dot{c}_1(t) \\ \dot{c}_2(t) \end{pmatrix}}_{=\dot{c}} = \underbrace{\begin{pmatrix} 0 & 1 \\ 2 & -4 \end{pmatrix}}_{=A} \underbrace{\begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix}}_{=c} \qquad \text{Lösung: } c(t) = e^{At}c(0)$$

Suche **Eigenwerte** von A: Über das charakteristische Polynom,

$$\det \begin{pmatrix} 0 - x & 1 \\ 2 & -4 - x \end{pmatrix} = x^2 + 4x - 2$$

Nullstellen $\lambda_1 = -1 + \sqrt{6}$, $\lambda_2 = -1 - \sqrt{6}$, das sind die Eigenwerte von A.
Eigenvektoren: Löse $(A - \lambda_i)v_i = 0$, zum Beispiel
$v_1 = \begin{pmatrix} 1 \\ -2 + \sqrt{6} \end{pmatrix}$, $v_2 = \begin{pmatrix} 1 \\ -2 - \sqrt{6} \end{pmatrix}$.
Also: $e^{At}v_1 = e^{(-2+\sqrt{6})t}v_1$, $e^{At}v_2 = e^{(-2-\sqrt{6})t}v_2$
$e^{At}(r_1v_1 + r_2v_2) = r_1e^{(-2+\sqrt{6})t}v_1 + r_2e^{(-2-\sqrt{6})t}v_2 = \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix}$
Anfangsbedingung: $c_1(0)$, $c_2(0)$ sind vorgegeben. $r_1v_1 + r_2v_2 = \begin{pmatrix} c_1(0) \\ c_2(0) \end{pmatrix}$,

$$\begin{pmatrix} 1 & 1 \\ -2 + \sqrt{6} & -2 - \sqrt{6} \end{pmatrix} \begin{pmatrix} r_1 \\ t_2 \end{pmatrix} = \begin{pmatrix} c_1(0) \\ c_2(0) \end{pmatrix}$$

$$\begin{pmatrix} r_1 \\ r_2 \end{pmatrix} = \tfrac{1}{-2\sqrt{6}} \begin{pmatrix} -2 - \sqrt{6} & -1 \\ 2 - \sqrt{6} & 1 \end{pmatrix} \begin{pmatrix} c_1(0) \\ c_2(0) \end{pmatrix}$$

$$\begin{aligned} \Rightarrow \quad & c_1(t) = r_1e^{(-2+\sqrt{6})t} + t_2e^{(-2-\sqrt{6})t} \\ = \quad & \frac{1}{-2\sqrt{6}} \left(c_1(0)(2 + \sqrt{6}) + c_2(0) \right) e^{(-2+\sqrt{6})t} \\ - \quad & \frac{1}{2\sqrt{6}} \left(c_1(0)(2 - \sqrt{6}) + c_2(0) \right) e^{(-2-\sqrt{6})t} \end{aligned}$$

mit $c_1 = y, c_2 = y'$, Startwerte $y(0) = c_1(0), y'(0) = c_2(0)$. Es folgt:
$\begin{pmatrix} 0 & 1 \\ 2 & -4 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -2+\sqrt{6} & -2-\sqrt{6} \end{pmatrix} \begin{pmatrix} -2+\sqrt{6} & 0 \\ 0 & -2-\sqrt{6} \end{pmatrix} \begin{pmatrix} -2-\sqrt{6} & -1 \\ 2-\sqrt{6} & 1 \end{pmatrix} \frac{1}{-2\sqrt{6}}$
$\Rightarrow e^{At} = -\frac{1}{2\sqrt{6}} \begin{pmatrix} 1 & 1 \\ -2+\sqrt{6} & -2-\sqrt{6} \end{pmatrix} \begin{pmatrix} e^{(-2+\sqrt{6})t} & 0 \\ 0 & e^{(-2-\sqrt{6})t} \end{pmatrix} \begin{pmatrix} -2-\sqrt{6} & -1 \\ 2-\sqrt{6} & 1 \end{pmatrix}$
$\dot{c}(t) = A(t)c(t)$

17.5 Inhomogene lineare Differentialgleichungen

Definition §17.12

Definition: Sei $J \subseteq \mathbb{R}$ ein offenes Intervall mit $0 \in J$, $A\colon J \longrightarrow L(V,V)$ und $b\colon J \longrightarrow V$, beide lokal Lipschitz-stetig. Die Differentialgleichung

$$\dot{c}(t) = A(t)c(t) + b(t)$$

heißt **inhomogene lineare Differentialgleichung**.

Satz §17.13

Satz: Der Definitionsbereich des globalen Flusses ist $\Omega = J \times V$, das heißt die Lösungen sind für *alle* $t \in J$ definiert.

Beweis: Sei $I \subseteq J$ abgeschlossen und beschränkt mit $0 \in I$.
Sei $R = \sup\{||A(t)|| \mid t \in I\}$ (das existiert, weil I abg. und beschränkt ist).
$F(x,t) = A(t)x + b(t), \quad ||F(x,t) - F(y,t)|| = ||A(t)(x-y)|| \leq R||x-y||$
Setze $c_0 = v$, $c_{n+1}(t) = \int_0^t F(c_n(s), s)\ \mathrm{ds} + v$ (Picard-Iteration)
$||c_{n+1}(t) - c_n(t)|| = ||\int_0^t F(c_n(s),s)\ \mathrm{ds} - \int_0^t F(c_{n-1}(s),s)\ \mathrm{ds}|| \leq |\int_0^t R||c_s(s) - c_{n-1}(s)||\ \mathrm{ds}| \leq R^n \underbrace{|t|^n ||c_1 - c_0||_\infty}_{=n!}$
Die $(c_n)_{n\in\mathbb{N}}$ bilden als stetige Kurven $I \longrightarrow V$ eine Cauchy-Folge mit Grenzwert $c\colon I \longrightarrow V$ mit $c(t) = \int_0^t F(c(s),s)\ \mathrm{ds} + v$.

$$\begin{aligned} \Longrightarrow \quad & \dot{c}(t) = A(t)c(t) + b(t) \quad \text{für alle } t \in I \\ & c(0) = v \end{aligned}$$

Das Argument geht für jedes abgeschlossene und beschränkte Intervall $I \subseteq J$, also für alle $t \in J$. □

Bemerkung §17.14

Bemerkung: Einige elementare Beobachtungen: Sei $F(x,t) = A(t)x + b(t)$, die Differentialgleichung $\dot{c}(t) = F(c(t),t)$ hat folgende Eigenschaften:

(i) Ist c_1 eine Lösung der homogenen linearen DGL, also $\dot{c}_1(t) = A(t)c_1(t)$ und ist $c_2(t)$ eine Lösung der inhomogenen linearen Differentialgleichung, also $\dot{c}_2(t) = A(t)c_2(t) + b(t)$, so ist $c = c_1 + c_2$ ebenfalls eine Lösung der inhomogenen linearen Differentialgleichung $\dot{c}(t) = A(t)c(t) + b(t)$ zum Anfangswert $c(0) = c_1(0) + c_2(0)$.

(ii) Kann die homogene Differentialgleichung *eindeutig* lösen und findet man *eine* Lösung der inhomogenen linearen Differentialgleichung, so hat man damit *alle* Lösungen der inhomogenen linearen Differentialgleichung zu beliebigen Anfangswerten.

Vorlesung 13.07.2004

Satz §17.15

Satz: Sei $\Phi_t(u)$ die Lösung des homogenen Problems, das heißt $c(t) = \Phi_t(u)$ hat die Eigenschaft, dass

$$\dot{c}(t) = A(t)c(t) \quad \text{mit} \quad c(0) = u$$

Dann gilt $\Phi_t(u+v) = \Phi_t(u) + \Phi_t(v)$, weil die Differentialgleichung linear ist. Das heißt Φ_t ist eine lineare Abbildung mit $\Phi_t \colon V \longrightarrow V$ und Φ_t ist invertierbar.

Beweis: Ist $\xi \colon V \times J \longrightarrow V \times J$ ein Vektorfeld mit $(u,v) \mapsto \big(A(r)u, 1\big)$ und Fluss $\varphi_t(u,r) = \big(\tilde{\varphi}_t(u,r), r+t\big)$ dann gilt $\Phi_t(u) = \tilde{\varphi}_t(u,0)$ also ist Φ_t invertierbar.□

Ist $A(t) = \alpha(t)A$ (siehe Beispiel §17.8), so ist $c(t) = e^{\int_0^t \alpha(s)\,\mathrm{d}s A}u$, das heißt in dieser Situation ist $\Phi_t = e^{\int_0^t \alpha(s)\,\mathrm{d}s A}$

Bemerkung §17.16

Bemerkung: Wir finden *eine* Lösung des inhomogenen Systems mit dem folgenden Ansatz: $\dot{u}(t) = \Phi_t^{-1}\big(b(t)\big)$, das heißt $u(t) = \int_0^t \Phi_t^{-1}\big(b(s)\big)\,\mathrm{d}s$. Setze

$$\begin{aligned}
c(t) &= \Phi_t\big(u(t)\big) \\
\dot{c}(t) &= \frac{\partial \Phi_t}{\partial t}\big(u(t)\big) + \Phi_t \dot{u}(t) = A(t)\Phi_t\big(u(t)\big) + \Phi_t\Phi_t^{-1}\big(b(t)\big) \\
&= A(t)c(t) + b(t)
\end{aligned}$$

Damit haben wir eine Lösung des inhomogenen Systems gefunden. Man nennt diesen Ansatz **Variation der Konstanten**.

Beispiel §17.17

Beispiel: Sei $V = \mathbb{R}^2$, $J = \mathbb{R}$ und $\dot{c}(t) = A(t)c(t) + b(t)$ mit

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},\ b(t) = \begin{pmatrix} 0 \\ t \end{pmatrix} \quad \Rightarrow \quad \begin{pmatrix} \dot{c}_1(t) \\ \dot{c}_2(t) \end{pmatrix} = \begin{pmatrix} c_2(t) \\ -c_1(t) \end{pmatrix} + \begin{pmatrix} 0 \\ t \end{pmatrix}$$

1. Schritt: Löse die homogene Differentialgleichung $\dot{c}(t) = A(t)c(t)$ vollständig.
$\Phi_t(u) = e^{\int_0^t A\,\mathrm{d}s}u = e^{tA}u = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} u \quad$ (e^{At} aus Beispiel §17.9)

2. Schritt: Finde *eine* Lösung der inhomogenen linearen Differentialgleichung. In dieser Situation ergibt sich: $\Phi_s^{-1} = \Phi_{-s}$

$$\begin{aligned}
u(t) &= \int_0^t \begin{pmatrix} \cos(s) & -\sin(s) \\ \sin(s) & \cos(s) \end{pmatrix} \begin{pmatrix} 0 \\ s \end{pmatrix} \mathrm{d}s = \int_0^t \begin{pmatrix} -s\sin(s) \\ s\cos(s) \end{pmatrix} \mathrm{d}s \\
&= \begin{pmatrix} t\cos(t) - \sin(t) \\ t\sin(t) + \cos(t) - 1 \end{pmatrix} \\
c(t) &= \Phi_t u(t) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} t\cos(t) - \sin(t) \\ t\sin(t) + \cos(t) - 1 \end{pmatrix} \\
&= \begin{pmatrix} t\cos^2(t) - \cos(t)\sin(t) + 2\sin^2(t) + \sin(t)\cos(t) - \sin(t) \\ -t\sin(t)\cos(t) + \sin^2(t) + t\sin(t)\cos(t) + \cos^2(t) - \cos(t) \end{pmatrix} \\
&= \begin{pmatrix} t - \sin(t) \\ 1 - \cos(t) \end{pmatrix}
\end{aligned}$$

Das ist eine Lösung der inhomogenen linearen Differentialgleichung zum Anfangswert $(0,0)$

3. Schritt: Allgemeine Lösung des inhomogenen Problems zum Anfangswert $c(0) = v = (v_1, v_2)$.

$$\begin{aligned} c(t) &= \Phi_t(v) + \begin{pmatrix} t - \sin(t) \\ 1 - \cos(t) \end{pmatrix} \\ &= \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} + \begin{pmatrix} t - \sin(t) \\ 1 - \cos(t) \end{pmatrix} \\ &= \begin{pmatrix} \cos(t) \\ -\sin(t) \end{pmatrix} v_1 + \begin{pmatrix} \sin(t) \\ \cos(t) \end{pmatrix} v_2 + \begin{pmatrix} t - \sin(t) \\ 1 - \cos(t) \end{pmatrix} \end{aligned}$$

17.6 Fundamentalsystem

Bemerkung §17.18

Bemerkung: Ist $V = \mathbb{R}^n$ und sind $v_{(1)}, \ldots, v_{(n)} \in \mathbb{R}^n$ linear unabhängige Vektoren, so gilt für alle $v \in \mathbb{R}^n$ mit $v = v_{(1)} r_1 + \cdots + v_{(n)} r_n$:

$$\Phi_t(v) = \Phi_t(v_{(1)} r_1) + \cdots + \Phi_t(v_{(n)}) r_n = c_{(1)}(t) r_1 + \cdots + c_{(n)}(t) r_n$$

Damit ist $c(t) = r_1 c_{(1)}(t) + \cdots + r_n c_{(n)}(t)$ Lösung der homogenen Differentialgleichung zum Anfangswert $c(0) = v$. Man nennt $(c_{(1)}, \ldots, c_{(n)})$ ein **Fundamentalsystem** von Lösungen der linearen homogenen Differentialgleichung.

Beispiel §17.19

Beispiel: Sei $V = \mathbb{R}$, $J =]-1, \infty[$ und $2(t+1)^2 y'' - (t+1)y' + y = 0$. Setze $c_1(t) = y(t),\ c_2(t) = y'(t)$

$$\left\{ \begin{aligned} 2(t+1)^2 \dot{c}_2(t) - (t+1)c_2(t) + c_1(t) &= 0 \\ \dot{c}_1(t) &= c_2(t) \end{aligned} \right.$$

$$\Rightarrow \left\{ \begin{aligned} \dot{c}_2(t) &= -\tfrac{1}{2} \tfrac{1}{(t+1)^2} c_1(t) + \tfrac{1}{2} \tfrac{1}{1+t} c_2(t) \\ \dot{c}_1(t) &= c_2(t) \end{aligned} \right.$$

$$\Rightarrow \begin{pmatrix} \dot{c}_1(t) \\ \dot{c}_2(t) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{1}{2(t+1)^2} & \frac{1}{2(t+1)} \end{pmatrix} \begin{pmatrix} c_1(t) \\ c_2(t) \end{pmatrix}$$

Ansatz: Setze $c_1(t) = (t+1)^a$ mit $a \in \mathbb{R}, a > 0$, also $c_2(t) = a(t+1)^{a-1} = \dot{c}_1(t)$ und $\dot{c}_2(t) = a(a-1)(t+1)^{a-2}$

$$\begin{aligned} a(a-1)(t+1)^{a-2} &= \frac{-1}{2(t+1)^2}(t+1)^a + \frac{a}{2(t+1)}(t+1)^{a-1} \\ &= \frac{1}{2}(t+1)^{a-2} + \frac{1}{2} a (t+1)^{a-2} \\ \Leftrightarrow \qquad a(a-1) &= \frac{1}{2}(a-1) \\ \Rightarrow \qquad a_1 = 1, &\qquad a_2 = \frac{1}{2} \end{aligned}$$

Lösung: $c_{(1)} = \begin{pmatrix} t+1 \\ 1 \end{pmatrix}, \quad c_{(2)} = \begin{pmatrix} \sqrt{t+1} \\ \frac{1}{2\sqrt{t+1}} \end{pmatrix}$

Die allgemeine Lösung der Differentialgleichung ist also von der Form

$$c = \begin{pmatrix} t+1 \\ 1 \end{pmatrix} r_1 + \begin{pmatrix} \sqrt{t+1} \\ \frac{1}{2\sqrt{t+1}} \end{pmatrix} r_2$$

Die bisherigen Methoden funktionieren hier nicht, da $A(t)$ nicht durch $\alpha(t) \cdot A$ ausdrückt werden kann.

Literaturverzeichnis

Analysis Bücher

- Bröcker, *Analysis 1/2*, Spektrum Verlag, Heidelberg, 1995.
- Forster, *Analysis 1/2*, Vieweg: Mathematik Grundkurs, Braunschweig, 1999.
- Heuser, *Lehrbuch der Analysis. Teil 1/Teil 2*, Teubner, Stuttgart, 1991.
- Königsberger, *Analysis 1/2*, Springer, Berlin, 1997.
- Walter, *Analysis 1/2*, Springer, Berlin, 1997.
- Walter, *Gewöhnliche Differentialgleichungen*, Springer, Berlin, 1972.

Allgemeine Bücher

- Beutelspacher, *Das ist o.B.d.A. trivial!*
  Eine Gebrauchsanleitung zur Formulierung mathematischer Gedanken mit vielen praktischen Tips für Studierende der Mathematik und Informatik. Friedr. Vieweg & Sohn, Braunschweig, 1997.
- Courant and Robbins, *Was ist Mathematik?*, Fourth edition, Springer, Berlin, 1992.
- Ebbinghaus et al., *Zahlen*, Springer, Berlin, 1983.
- Graham, Knuth and Patashnik, *Concrete mathematics*, Addison-Wesley, Reading, MA, 1994.
- Schönhage-Strassen: *Schnelle Multiplikation großer Zahlen*, Computing 7 (1971), Seiten 281-292
- Knuth, *Arithmetik*, Springer, Berlin, 2001.

Abbildungsverzeichnis

Index