

Grundkurs Analysis
Bernd Kummer

Wer Arme auf den Arm nimmt,
anstatt ihnen unter die Arme zu greifen,
sollte besser die Beine unter die Arme nehmen,
denn es gibt mehr Arme als Beine auf der Welt.

K. Nissen, M. Reuter: Die neuen Leiden der jungen Wörter: Das aktuelle Wörterbuch zur Rächtschraiprehvorm, Knauer, München 1999

Dieses Script umfasst wichtige Definitionen, Sätze und Beweise des Analysis Grundkurses, Teil I und (den grössten Teil von) Teil II.

Es soll Anregung und Hilfe zum Nacharbeiten des Vorlesungsstoffes sein.

Es ist nicht in jedem Falle die chronologische Vorgehensweise der Vorlesung und auch nicht völlig identisch mit der tatsächlichen Stoffauswahl.

Ausserdem gibt es zu jedem Kapitel noch viel mehr zu sagen.

Hinweise auf Fehler werden dankbar entgegengenommen.

11. Juli 2011

Inhaltsverzeichnis

1	Elemente der Mengenlehre	7
1.1	Mengen und Funktionen	7
1.2	Mächtigkeit	9
1.3	Zahlbereiche	10
1.3.1	Natürlichen Zahlen und vollständige Induktion	10
1.3.2	Ganze Zahlen	12
1.3.3	Rationale Zahlen	13
1.3.4	Reelle Zahlen	13
1.3.5	Körperereigenschaften	16
1.3.6	Komplexe Zahlen	17
1.3.7	Polarkoordinaten und Wurzeln	17
2	Folgen im Reellen	19
2.1	Spezielle wichtige Grenzwerte	21
3	Metrische Räume	23
3.1	Grundlegende Begriffe	23
3.1.1	Konvergenz	23
3.1.2	Stetigkeit für verknüpfte Funktionen	26
3.2	Überdeckungssatz von Heine/Borel/Lebesgue	27
3.2.1	Beispiel einer Anwendung	29
3.2.2	Vollständige metrische Räume	30
3.3	Banachscher Fixpunktsatz	31
4	Reihen	33
4.1	Konvergenzbedingungen für Reihen	34
4.2	Quotientenkriterium, Wurzelkriterium	34
4.3	Absolute Konvergenz	35
4.4	Summen von Reihen	35
4.5	Alternierende Reihen	37
4.6	Produkte von absolut konvergenten Reihen	37
4.7	Potenzreihen	38
4.8	Eulersche Zahl e	39
5	Zusatz: Wichtige Funktionen	43
5.1	Die Potenz- und Logarithmus Funktion	43
5.2	Erweiterung von f auf \mathbb{R} und f^{-1} allgemein	45
6	Differentialrechnung	47
6.1	Ableitungsregeln	50
6.1.1	Winkelfunktionen und ihre Arcus-Funktionen	55
6.1.2	Hyperbel-Funktionen und ihre Area-Funktionen	56
6.1.3	Zusammenfassung Wichtige Ableitungen	57
6.2	Höhere und partielle Ableitungen	58

6.3	Taylor-Reihen	60
6.3.1	Taylor-Entwicklung für Funktionen $f : \mathbb{R}^m \rightarrow \mathbb{R}$	63
6.3.2	Anwendungen der Taylorentwicklung	65
6.4	Ableitungen für $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ und Satz über implizite Funktionen	68
6.4.1	Satz über implizite Funktionen	69
6.4.2	Lagrange Multiplikatoren für Extrema mit Nebenbedingungen	71
6.4.3	Newton-Methode zur Lösung von $F(x) = 0$:	72
7	Komplexe Zahlen und Funktionen	75
7.1	Potenzreihen und spezielle komplexe Funktionen	75
7.1.1	Potenzreihen, Exponential-, Sinus- und Kosinus-Funktion	75
7.1.2	Logarithmus	76
7.2	Fundamentalsatz der Algebra	77
7.2.1	Beweis des Fundamentalsatzes	77
7.2.2	Polynomdivision und konjugiert komplexe Nullstellen	79
7.3	Komplexe Differenzierbarkeit	79
8	Integralrechnung	83
8.1	Das bestimmte Integral für eine stetige Funktion	83
8.2	Berechnung von Stammfunktionen und Integrationsmethoden	86
8.2.1	Partielle Integration	87
8.2.2	Substitutionsmethode	88
8.2.3	Anwendungen der Substitutionsmethode	89
8.2.4	Partialbruchzerlegung	93
8.3	Uneigentliche Integrale	96
8.3.1	Laplace Transformierte	96
8.3.2	Gamma-Funktion	97
8.4	Näherungsformeln für Integrale	97
8.4.1	Trapezformel	97
8.4.2	Fassformel	98
8.4.3	Polynom-Approximation stetiger Funktionen: Bernstein-Polynome	98
8.5	Satz von Green (Gebiets- und Kurvenintegral; Dim. 2)	99
9	Differentialgleichungen (DGL)	103
9.1	Existenz und Eindeutigkeit von Lösungen	103
9.1.1	Der Raum $C[a, b]$	103
9.1.2	Eine reelle Funktion wird gesucht	104
9.1.3	Ein System von Funktionen wird gesucht	106
9.2	Existenzsatz für Lineare Systeme	106
9.3	Lineare DGL-Systeme mit konstanten Koeffizienten	107
9.3.1	Lösungsansätze	108
9.3.2	Beispiele	111
9.4	Lösungsansätze für spezielle Typen von DGL	113
9.4.1	Trennung der Variablen	113
9.4.2	Homogene DGL $y' = f(y/x)$	113
9.4.3	Lin. DGL 1. Ordnung (var. Koeff.)	113
9.4.4	Bernoulli-Gleichung	114
9.4.5	Riccati-Gleichung	115
9.4.6	Lineare DGL n-ter Ordnung (feste Koeff.)	115
10	Anhang	117
10.1	Konvexität und Norm-Äquivalenz im \mathbb{R}^n	117
10.2	Satz von Arzela-Ascoli	118
10.3	Daten einiger bekannter Mathematiker	119
	Index	120

Kapitel 1

Elemente der Mengenlehre

1.1 Mengen und Funktionen

Eine Menge ist eine „wohldefinierte Gesamtheit irgendwelcher Objekte“. Die Objekte heißen Elemente der Menge. Um zu sagen, dass a ein Element der Menge A ist, schreibt man $a \in A$. Die Menge ohne irgendein Element heißt leere Menge, Symbol \emptyset .

Eine Menge kann endlich oder unendlich viele Elemente enthalten. „Wohldefiniert“ meint stets, dass die Objekte, die zur Menge gehören sollen, zweifelsfrei erklärt sind.

Die Reihenfolge, in der sie eventuell aufgeschrieben sind, spielt dabei keine Rolle.

Als Mathematiker sprechen wir z.B. nicht von der Menge der guten Menschen, der Menge aller Fussballkenner oder der Menge aller Wahrsager, solange sie nicht „wohldefiniert“ sind.

Zwei Mengen sind gleich, wenn sie dieselben Elemente enthalten.

Beispiel der Beschreibung einer Menge

$$M = \left\{ x \mid x \text{ ganze Zahl, } x > 7, \frac{x}{2} \text{ ganze Zahl} \right\}$$

Lies:

$$\begin{aligned} & \text{Menge aller } x \text{ mit: } x \text{ ist ganzzahlig, } x \text{ ist größer als } 7 \text{ und } \frac{x}{2} \text{ ist ganzzahlig} \\ = & \text{ Menge der geraden Zahlen } 8, 10, 12, \dots \end{aligned}$$

Definition 1.1 (Vereinigung, Durchschnitt, Differenz, Teilmenge). Sind A und B Mengen, so kann man neue Mengen bilden, indem man definiert:

Vereinigung: $A \cup B = \{c \mid c \in A \text{ oder } c \in B\}$ („oder“ heißt hier nie „entweder oder“),

Durchschnitt: $A \cap B = \{c \mid c \in A \text{ und } c \in B\}$,

Differenz: $A \setminus B = \{c \mid c \in A \text{ und } c \notin B\}$.

Man sagt, A ist **Teilmenge** von B , wenn jedes Element aus A auch in B liegt; kurz $A \subset B$ (das erlaubt noch Gleichheit!) oder auch $B \supset A$.

(In manchen Büchern bezeichnet \subset auch nur echte Teilmengen, d.h. $A \neq B$; hier nicht).

Man beachte: \emptyset und B sind stets Teilmengen von B .

Beispiel 1.2. $A =$ Menge aller Großmütter und $B =$ Menge aller Mütter, also $A \subset B$.

Definition 1.3 (Potenzmenge). Die spezielle Menge C aller Teilmengen A einer gegebenen Menge B heißt **Potenzmenge** von B ; kurz $C = \mathcal{P}(B)$ oder $C = 2^B$.

Beispiel 1.4. $B = \{1, 2, 3\}$; $C = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}$ besteht aus $8 = 2^3$ Elementen, die selbst Mengen sind, nämlich Teilmengen von B .

Die Menge $\{1\}$ und das Element 1 der Menge B sind hier verschiedene Dinge.

Rechenregeln für Mengen. Man sieht leicht, dass für beliebige Mengen A, B, C gilt:

$$\begin{aligned}(A \cup B) \cap C &= (A \cap C) \cup (B \cap C), \\ (A \cap B) \cup C &= (A \cup C) \cap (B \cup C).\end{aligned}$$

Zum Beweis nimmt man an, dass ein Element in der linken Menge liegt und zeigt, dass es dann auch in der rechten liegen muss. Anschließend dasselbe mit „rechts-links“.

Analog folgt

$$A \cap B = (A \cup B) \setminus ((A \setminus B) \cup (B \setminus A)).$$

Aussagen: Eine Teilmenge $A \subset M$ kann man mit einer Aussage A' (über einer Gesamtheit M) „gleichsetzen“: Die Aussage A' (getroffen über ein Element $m \in M$) ist wahr, wenn m zu A gehört, sonst falsch.

Beispiel 1.5. $\mathbb{Z} =$ Menge der ganzen Zahlen, $A =$ Menge der positiven ganzen Zahlen; die Aussage A' bedeutet, dass $z \in \mathbb{Z}$ positiv ist.

Dann entspricht, für zwei Teilmengen A und B von M :

- $A \cup B$ dem logischen *oder*,
- $A \cap B$ dem logischen *und*,
- $A \subset B$ der Implikation „aus A' folgt B' “,
- und $M \setminus A$ der Negation von A' .

Relationen. Oft ist es sinnvoll zu definieren, wann zwei Elemente a, b einer Menge C in einer bestimmten Relation stehen.

Beispiel 1.6. $C = \mathbb{Z} =$ Menge der ganzen Zahlen; $a \leq b$ falls $b - a$ nichtnegativ ist.

Die Relation „ \leq “ besitzt (für beliebige a, b aus C) die Eigenschaften:

- (1) Es gilt stets $a \leq b$ oder $b \leq a$. (Relation ist vollständig)
- (2) Wenn $a \leq b$ und $b \leq c$, so folgt $a \leq c$. (Relation ist transitiv)
- (3) Wenn $a \leq b$ und $b \leq a$, so folgt $a = b$. (Relation ist antisymmetrisch)

Eine Relation „ \leq “ (definiert für Paare einer beliebigen Menge C) mit diesen 3 Eigenschaften nennt man auch **Ordnungsrelation** (auf C).

Oft schreibt man $a < b$ für $a \leq b$ und $a \neq b$. In dieser Form werden die Forderungen zu:
 - Für beliebige $a, b \in C$ gilt stets genau eine der Relationen $a < b, b < a$ bzw. $a = b$.
 - Wenn $a < b$ und $b < c$, so folgt $a < c$.

Eine Ordnungsrelation auf \mathbb{Z} muss nicht notwendig wie oben definiert sein (also dem entsprechen, was wir uns „natürlicherweise“, vorstellen), z.B. könnte man „ $b - a$ nichtnegativ“ durch „ $b - a$ nichtpositiv“ ersetzen, ohne dass (1), (2), (3) falsch werden.

Verzichtet man auf Eigenschaft (1), spricht man von Halbordnungen. Dann kann es passieren, dass gewisse a und b nicht vergleichbar sind.

Mitunter spricht man auch von Ordnungsrelation, wenn nur (1) und (2) richtig sind (wir nicht). Gilt dann sowohl $a \leq b$ als auch $b \leq a$, nennt man a und b äquivalent (im Sinne der Relation).

Hat eine Ordnungsrelation noch die Eigenschaft, dass es in jeder nichtleeren Teilmenge A von C ein kleinstes Element $x(A) \in A$ gibt, d.h., $x(A) \leq a \forall a \in A$, so heißt die Ordnung eine Wohlordnung.

Die „normale“ Ordnung der natürlichen Zahlen ist offenbar ein Wohlordnung, die der ganzen Zahlen nicht.

Eine andere Relation wird in derselben Menge $C = \mathbb{Z}$ definiert, wenn man schreibt $a \sim b$, falls a und b bei Division durch 3 denselben Rest lassen. Nun folgt:

- (1) Es gilt stets $a \sim a$. (Relation ist reflexiv)
- (2) Wenn $a \sim b$ und $b \sim c$, so folgt $a \sim c$. (Relation ist transitiv)
- (3) Wenn $a \sim b$, so folgt $b \sim a$. (Relation ist symmetrisch)

Eine Relation „ \sim “ mit diesen Eigenschaften (definiert für gewisse Elementenpaare einer beliebigen Menge C) nennt man auch **Äquivalenzrelation**.

Die Menge aller $b \in C$ mit $b \sim a$ heißt **Äquivalenzklasse** zum Element a und wird oft mit $[a]$ bezeichnet. Dann hat $a \sim b$ wegen (2) die Gleichheit $[a] = [b]$ zur Folge.

Die einfachste Äquivalenzrelation ist die Gleichheit. Allgemein kann man *äquivalent* lesen als *gleich unter einem bestimmten Gesichtspunkt*.

Für eine Relation \leq , die nur (1) und (2) (vollständig und transitiv) erfüllt, wird durch „ $a \sim b$ falls $a \leq b$ und $b \leq a$ “, eine Äquivalenzrelation erklärt.

Funktionen/Abbildungen

Definition 1.7. Sind A und B Mengen, so ist eine Funktion (auch Abbildung genannt) f von A in B (kurz $f : A \rightarrow B$) eine Vorschrift, die jedem Element $a \in A$ ein und nur ein (man sagt auch genau ein) Element $f(a) \in B$ zuordnet, das Bild von a .

Dabei heißt A Definitionsbereich (Urbildmenge) und B Wertebereich (auch Bildmenge) von f . Die Menge $f(A) = \{b \in B \mid b = f(a) \text{ für ein } a \in A\}$ heißt Bild von A . Die (eventuell leere) Menge $f^{-1}(b) = \{a \in A \mid f(a) = b\}$ heißt Urbildmenge von $b \in B$ (natürlich jeweils für f).

Beispiel 1.8. $A =$ Menge der Mieter eines Hauses, $B =$ Menge der natürlichen Zahlen, $f(a) =$ Alter von a .

Bemerkung: Man sieht am Beispiel, dass $f^{-1}(b)$ leer sein und damit das Bild $f(A)$ von A eine echte Teilmenge der Bildmenge B sein kann; $f(A) \subset B$, $f(A) \neq B$ (man schreibt dafür auch $f(A) \subsetneq B$). Definiert man $a' \sim a''$ falls $f(a') = f(a'')$, so entsteht eine **Äquivalenzrelation** \sim in A .

Gilt $B = f(A)$, so heißt $f : A \rightarrow B$ **surjektiv** (oder Abbildung auf B). Das ist also gleichbedeutend mit $f^{-1}(b) \neq \emptyset$ für alle (kurz \forall) $b \in B$.

Besteht $f^{-1}(b)$ nie aus mehreren Elementen, heißt f **injektiv**. Also bedeutet injektiv, dass $f(a') = f(a'')$ stets $a' = a''$ zur Folge hat; jedes $b \in B$ tritt höchstens einmal als Bild $f(a)$ auf.

Eine injektive und surjektive Funktion/Abbildung $f : A \rightarrow B$ heißt **bijektiv**. Mitunter schreibt man dann auch $f : A \leftrightarrow B$. Nun ordnet f zugleich jedem Element $b \in B$ genau ein Urbild a mit $f(a) = b$ zu; d.h., $\{a\} = f^{-1}(b)$. In diesem Fall schreibt man einfacher $a = f^{-1}(b)$ und nennt die so definierte *Funktion*

$$f^{-1} : B \rightarrow A$$

inverse Funktion zu f .

Man beachte, dass genau die bijektiven Funktionen eine inverse Funktion besitzen, obwohl die Urbildmengen $f^{-1}(b)$ allgemein definiert wurden.

Beispiel 1.9. $A =$ Menge der Teilnehmer am Berlin-Marathon, $f(a) =$ Startnummer von a ; $B =$ Menge der vergebenen Startnummern (zumindest sollte f jetzt bijektiv sein).

1.2 Mächtigkeit

Definition 1.10 (gleichmächtig). Zwei Mengen A und B heißen gleichmächtig, wenn es eine Bijektion $f : A \rightarrow B$ gibt.

Beispiel 1.11. 1. $A = \{1, 2\}$, $B = \{Max, Moritz\}$; $f(1) = Moritz$, $f(2) = Max$.

2. $A = \{1, 2, 3, 4, \dots\}$, $B = \{0, 2, 4, 6, \dots\}$; $f(a) = 2(a - 1)$.

Beispiele 1.12.

- Sind die Mengen $A = \{x, y, z\}$, $C = \{Max, Moritz\}$ gleichmächtig?
- Man zeige, dass $A = \{\dots, -6, -4, -2, 0, 2, 4, 6, \dots\}$ und $B = \{1, 4, 9, 16, 25, 36, \dots\}$ gleichmächtig sind.
- Sei $C = \mathcal{P}(B)$ die Potenzmenge einer Menge B . Man definiere für zwei Mengen A', A'' aus C :
 - $A' \sim A''$ falls A' und A'' gleichmächtig sind.
 - $A' \leq A''$ falls A' und eine Teilmenge A von A'' gleichmächtig sind.

Man zeige: \sim ist eine Äquivalenzrelation, \leq ist transitiv.

Welche Mengen sind jeweils gleichmächtig, welche nicht? (Beweise s. Vorlesung).

Die Mengen $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ der natürlichen Zahlen, die Menge $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$ der ganzen Zahlen und die Menge \mathbb{Q} der rationalen Zahlen sind gleichmächtig.

Die entsprechende Mächtigkeit wird auch *abzählbar* bzw. *abzählbar unendlich* genannt.

Die Menge der Punkte des Intervalls $(0, 1)$ und die einer Geraden sind gleichmächtig.

Die Menge der Punkte des Intervalls $[0, 1]$ und die eines Quadrates (mit Rand) sind gleichmächtig.

Die Menge der reellen Zahlen im Intervall $[0, 1]$ und \mathbb{N} sind nicht gleichmächtig.

Cantors Kontinuumshypothese: Es gibt keine Menge, die eine kleinere Mächtigkeit als die der Menge der reellen Zahlen und (zugleich) eine größere als die der natürlichen Zahlen besitzt. Die Kontinuumshypothese lässt sich mit den klassischen Mitteln der Mengenlehre weder beweisen noch widerlegen.

1.3 Zahlbereiche

1.3.1 Natürlichen Zahlen und vollständige Induktion

Die Menge der natürlichen Zahlen $\mathbb{N} = \{0, 1, 2, 3, \dots\}$.

Die natürlichen Zahlen repräsentieren die Mächtigkeit endlicher Mengen.

Für Interessenten: Sie lassen sich aber auch abstrakt definieren mittels der

Peano-Axiome der natürlichen Zahlen.

Die Menge \mathbb{N} ist (bis auf Umbenennung ihrer Elemente) durch die folgenden Eigenschaften bzw. Axiome definiert:

- (1) 0 ist eine natürliche Zahl (oft beginnt man mit 1).
- (2) Jede natürliche Zahl n hat genau einen Nachfolger $n' \in \mathbb{N}$; er erfüllt $n' \neq n$.
- (3) Jede natürliche Zahl n hat höchstens einen Vorgänger v , d.h. es gibt maximal ein v mit $v' = n$.
- (4) 0 ist nicht Nachfolger einer natürlichen Zahl (d.h., hat keinen Vorgänger).
- (5) Axiom zur vollständigen Induktion:
Trifft eine Aussage für die natürliche Zahl 0 zu und folgt, mit jedem $n \in \mathbb{N}$, aus ihrer Richtigkeit für n auch die Richtigkeit für den Nachfolger n' , so gilt sie für alle $n \in \mathbb{N}$.

Wie arbeitet man mit diesen Axiomen?

Durch den Begriff Nachfolger wird wegen (2) eine Funktion $f: \mathbb{N} \rightarrow \mathbb{N}$ definiert, sodass ($n' =$) $f(n) \neq n$. Dabei gilt wegen (1), (3): $0 \in \mathbb{N} \setminus f(\mathbb{N})$, und wegen (4) ist f injektiv (denn $u' = v'$ hat $u = v$ zur Folge).

Mittels Axiom (5) sieht man, dass die folgende Aussage richtig ist:

Wenn $n \neq 0$, so hat n genau einen Vorgänger.

Beweis. Die Aussage trifft auf $n = 0$ zu, da sie für dieses n nichts behauptet. Sie gilt auch für den (existierenden) Nachfolger von 0, weil 0 dessen einziger Vorgänger ist. Analog gilt sie für jeden Nachfolger $k = n'$ denn n ist sein Vorgänger (der einzige). Mit Axiom (5) folgt die Behauptung also für alle $n \in \mathbb{N}$. \square

Die Elemente von \mathbb{N} kann man nun nacheinander aufschreiben: $0, f(0), f(f(0)), f(f(f(0))), \dots$ und irgendwie umbenennen, z.B. $a_1 = f(0), a_2 = f(a_1), a_3 = f(a_2), \dots$

Dabei braucht man stets neue Symbole, denn die Elemente wiederholen sich nicht.

Beweis. Würde 0 nochmals auftreten, hätte 0 einen Vorgänger. Ansonsten betrachten wir das **erste** Element, das irgendwann doppelt auftritt; z.B. etwa a_k , und es sei a_m das **erste** Element mit $m \neq k$ und $a_k = a_m$. Dann hat jedoch $n = a_k$ die zwei verschiedenen Vorgänger a_{k-1} und a_{m-1} , was (4) widerspricht. \square

Es bleibt zu zeigen, dass \mathbb{N} nicht noch weitere Elemente enthalten kann, was wieder mit Axiom 5 geht (wie?).

Bemerkung: Die „normale“ $a \leq b$ Relation lässt sich nun dadurch definieren, dass a nicht später als b aufgeschrieben wird.

Axiom (5) kann man dann auch so lesen:

(6) In jeder nichtleeren Teilmenge $M \subset \mathbb{N}$ gibt es genau ein minimales Element $n = n(M)$.

Aus (5) folgt (6) per vollständiger Induktion, wenn man nacheinander Mengen M betrachtet, die $0, f(0), f(f(0)), \dots$ enthalten.

Umgekehrt folgt aus (6) auch (5): Wenn man nämlich definiert $M = \{n \mid \text{Aussage } A \text{ ist falsch für } n\}$, sichert die erste Annahme von Axiom (5): $0 \notin M$. Weiter ist $n(M) \in M$ falls $M \neq \emptyset$. Für den Vorgänger m von $n(M)$ (der nicht zu M gehört) ist dann A richtig, nach der zweiten Annahme des Axioms (5) ist aber damit A auch für dessen Nachfolger $n(M) = m'$ wahr, also $n(M) \notin M$. Daher führt die Annahme $M \neq \emptyset$ zu einem Widerspruch.

! Für viele Beweise (Beweise mittels vollständiger Induktion) ist das Axiom (5) besonders wichtig.

Beispiele 1.13.

1. Bernoullische Ungleichung

Es gilt stets

$$(1+x)^n \geq 1+nx \quad (\text{falls } x > -1 \text{ und } n \in \mathbb{N}). \quad (1.3.1)$$

Beweis: Für $n = 0$ und $n = 1$ trivial.

Gilt die Ungleichung schon für irgendein n , so folgt für (den Nachfolger) $n+1$ wegen $1+x > 0$:

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

Die Ungleichung ist dann also auch für $n+1$ richtig. Nach dem Induktionsaxiom gilt sie so für alle n . \square

Die Ungleichung wird noch viele nützliche Folgerungen zulassen.

2. Die Menge $A_n = \{1, 2, \dots, n\}$ besitzt genau 2^n Teilmengen.

Beweis: Für $n = 0$ ist $A_n = \emptyset$, und \emptyset ist die einzige Teilmenge; $2^0 = 1$.

Die Aussage gelte für irgendein n . Wir zeigen, dass sie dann auch für $n+1$ gilt.

Es gibt 2^n Teilmengen T_k ohne das Element $n+1$, sie sind die Teilmengen von A_n . Weiter gibt es noch die 2^n Teilmengen $T^k \cup \{n+1\}$. Das sind offenbar alle, insgesamt also $2^n + 2^n = 2^{n+1}$. \square

Man kann auch mit $n = 1$ (als Induktionsanfang) beginnen und die Aussage (per Induktion) dann für alle natürlichen Zahlen $n \geq 1$ beweisen.

3. Permutationen

Für die Menge $A_n = \{1, 2, \dots, n\}$ ($n > 0$) gibt es genau $n!$ mögliche Reihenfolgen (Permutationen) ihrer Elemente ($n!$ ist das Produkt der Zahlen $1, 2, \dots, n$).

Beweis: Für $n = 1$ ist das richtig, die Aussage gelte für irgendein $n \geq 1$. Es gibt also $n!$ mögliche Schlangen, wenn die Leute $1, 2, \dots, n$ an einem Postschalter anstehen. Zu irgendeiner Schlange komme Herr $n+1$ hinzu. Er kann sich ganz vorn einsortieren oder an beliebiger anderer Stelle. Das sind genau $n+1$ Möglichkeiten. Für die Gesamtzahl der Reihenfolgen erhalten wir damit wie verlangt $(n+1)n! = (n+1)!$. \square

Die Reihenfolgen heißen auch Permutationen. Jede stellt eine Bijektion f der Menge A_n auf sich dar, wenn man $f(k)$ als den Platz von k deutet.

Man definiert $0! = 1$, weil es für viele Aussagen zweckmäßig ist, speziell für die nächsten.

4. 6 aus 49

Es gibt genau $\binom{49}{6} = \frac{49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44}{1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot 6}$ verschiedene Tipps für dieses Spiel.

Beweis: Es gibt $n = 49 \cdot 48 \cdot 47 \cdot 46 \cdot 45 \cdot 44$ verschiedene Möglichkeiten, $k = 6$ Zahlen aus 49 auszuwählen, wenn man die Reihenfolge ihrer Auswahl berücksichtigt. Da aber uninteressant ist, in welcher Reihenfolge die Zahlen angekreuzt wurden, haben wir (wegen 3.) durch die Anzahl der Reihenfolgen $k!$ zu dividieren. \square

Für „ k aus n “ ($k \leq n$) erhält man also $\binom{n}{k}$ Möglichkeiten. Man definiert zweckmäßigerweise $\binom{n}{0} = 1$. Wir zeigen nun für $n, k \geq 1$:

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1}. \quad (1.3.2)$$

Beweis. Die Anzahl der Tipps $\binom{n+1}{k}$ für k aus $n+1$ kann man auch so bestimmen: Wir zählen erst alle Tipps, die die Zahl $n+1$ ignorieren. Das sind offenbar wieder $\binom{n}{k}$. Nun gibt es noch alle Tipps mit der Zahl $n+1$. Dann kann man nur noch $k-1$ übrige Zahlen aus $1, \dots, n$ auswählen; die Anzahl dieser Tipps ist somit $\binom{n}{k-1}$. Damit folgt (1.3.2) \square

5. Binomialkoeffizienten

Die n -te Potenz $(a+b)^n$ besteht (Ausmultiplizieren und Zusammenfassen!) immer aus Summen von Termen (Ausdrücken) der Form $c_{k,n} a^k b^{n-k}$, wobei $c_{k,n}$ eine Konstante ist und k die Zahlen $0, 1, \dots, n$ durchläuft.

Die $c_{k,n}$ heißen Binomialkoeffizienten (zur Potenz n).

Um einen Term mit den Potenzen $a^k b^{n-k}$ zu erhalten, muss man aus den n Produkten von $(a+b)$ genau k mal a zur Multiplikation auswählen (und $n-k$ mal b), wobei die Reihenfolge der Auswahl wieder gleichgültig ist. Wir „spielen also k aus n “ und haben, wie wir wissen, gerade $\binom{n}{k}$ Möglichkeiten.

Also gilt:

$$c_{k,n} = \binom{n}{k} \quad \text{und} \quad (a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}.$$

Vertauscht man a und b , sieht man nun $\binom{n}{k} = \binom{n}{n-k}$.

Wegen Formel (1.3.2) ergeben sich die Zahlen $\binom{n}{k}$ zeilenweise einfach aus dem Pascalschen Dreieck

$n = 0$	1					
$n = 1$		1	1			
$n = 2$		1	2	1		
$n = 3$		1	3	3	1	
$n = 4$	1	4	6	4	1	$k = 0, 1, \dots, 4$

Jede Zahl außer den Einsen ist die Summe der beiden darüber.

6. Zeilensumme des Pascalschen Dreiecks

Man beweise, dass die Zeilensumme des Pascalschen Dreiecks allgemein stets 2^n ist.

7. Summe der Quadratzahlen

Man beweise: $1^2 + 2^2 + \dots + n^2 = \frac{n(n+1)(2n+1)}{6}$.

1.3.2 Ganze Zahlen

Die Menge der ganzen Zahlen

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}.$$

wird definiert, damit die Gleichung

$$n + x = m \text{ für } n, m \in \mathbb{N} \text{ (und später } n, m \in \mathbb{Z} \text{) stets lösbar ist.}$$

Um sie exakt zu definieren, betrachte man alle geordneten Paare (m, n) , $m, n \in \mathbb{N}$ und definiere

$$(m, n) \sim (m', n') \text{ falls } m + n' = n + m'.$$

Dies ist eine Äquivalenzrelation. Wir zeigen Transitivität (Symmetrie und Reflexivität sind offensichtlich). Wenn noch $(m', n') \sim (m'', n'')$ gilt, folgt auch $m' + n'' = n' + m''$. Addition beider Gleichungen liefert $m + n' + m' + n'' = n + m' + n' + m''$, was nur richtig ist, wenn auch $m + n'' = n + m''$ gilt. Dies bedeutet aber $(m'', n'') \sim (m, n)$ wie verlangt.

Klar, es ist stets $(n, n) \sim (0, 0)$.

Wir definieren nun eine Addition der Paare durch:

$$(m, n) + (m', n') = (m + m', n + n'),$$

wobei rechts die Addition in \mathbb{N} benutzt wird. Ersetzt man (m', n') durch ein äquivalentes Paar (m'', n'') , folgt nach dieser Definition

$$(m, n) + (m'', n'') = (m + m'', n + n'').$$

Wegen $(m', n') \sim (m'', n'')$ folgt $m + m' + n + n'' = n + n' + m + m''$. Damit sieht man, dass beide Ergebnisse äquivalent sind. Ebenso kann man (m, n) durch ein äquivalentes Paar ersetzen; die Ergebnisse bleiben äquivalent. Man kann deshalb die Addition auf die Äquivalenzklassen $[m, n]$ zu den entsprechenden Paaren ausweiten, indem man sagt:

$$[m, n] + [m', n'] \text{ sei die Äquivalenzklasse von } (m + m', n + n').$$

Damit folgt

$$[m, n] + [n, m] = [m + n, n + m] = [0, 0].$$

Den natürlichen Zahlen n entsprechen die Elemente $[n, 0]$ und den negativen ganzen Zahlen $-n$ die Elemente $[0, n]$, $n > 0$.

1.3.3 Rationale Zahlen

Die Menge der rationalen Zahlen wird oft mit \mathbb{Q} bezeichnet. Die rationalen Zahlen sind alle „Ergebnisse“ von Brüchen der Form $\frac{p}{q}$ mit $p, q \in \mathbb{Z}$ und $q \neq 0$.

Motiv für die Definition: Jetzt ist die Gleichung $ax = b$ stets lösbar, wenn $a \neq 0$ und $a, b \in \mathbb{Z}$.

Zu beachten ist, dass $\frac{p}{q} = \frac{kp}{kq}$ für alle $k \in \mathbb{Z}$, $k \neq 0$ gilt. Also kann dieselbe rationale Zahl $a = \frac{p}{q}$ in verschiedener Weise aufgeschrieben werden. Um den (intuitiven und auf Rechenregeln basierenden) Begriff Ergebnis in der Definition rationaler Zahlen vermeiden, kann man sich exakter so helfen, dass man 2 Brüche $a = \frac{p}{q}$ und $b = \frac{p'}{q'}$ als äquivalent bezeichnet, wenn $p'q = pq'$, was mit der üblichen Division dasselbe wird wie $\left(\frac{p}{q}\right) / \left(\frac{p'}{q'}\right) = 1$.

Wie schon bei ganzen Zahlen ist diese Relation leicht als Äquivalenzrelation zu erkennen. Die rationalen Zahlen sind nun die Äquivalenzklassen aller Brüche obiger Form. Die Rechenregeln mit rationalen Zahlen definiert man wie üblich:

Addition:

$$\frac{p}{q} + \frac{p'}{q'} = \frac{pq' + p'q}{qq'} \text{ (bis auf Kürzen) bzw. in der Sprache der Äquivalenzklassen } \left[\frac{p}{q}\right] + \left[\frac{p'}{q'}\right] = \left[\frac{pq' + p'q}{qq'}\right].$$

Subtraktion:

$x - y$ ist definiert durch $x + (-y)$ wobei $-y$ das inverse Element zu y bzgl. Addition ist; also $y + (-y) = 0$ löst. Es ist eindeutig, denn wenn auch y' die Gleichung löst, folgt $y + y' = 0 = y + (-y)$. Addiert man $(-y)$ auf beiden Seiten, folgt so $y' = -y$.

Multiplikation:

$$\frac{p}{q} \cdot \frac{p'}{q'} = \frac{pp'}{qq'} \text{ (bis auf Kürzen) bzw. in der Sprache der Äquivalenzklassen } \left[\frac{p}{q}\right] \cdot \left[\frac{p'}{q'}\right] = \left[\frac{pp'}{qq'}\right].$$

Division:

$\frac{x}{y}$ ist definiert als xy^{-1} , wobei y^{-1} das inverse Element zu y bzgl. Multiplikation ist; also $yy^{-1} = 1$ löst, falls $y \neq 0$; d.h. $y^{-1} = \left[\frac{q}{p}\right]$ falls $y = \left[\frac{p}{q}\right]$; Eindeutigkeit wie bei Subtraktion.

Schließlich vereinbart man für $q, q' > 0$ (was man notfalls durch Multiplikation von Zähler und Nenner mit -1 erreicht) $\left[\frac{p}{q}\right] < \left[\frac{p'}{q'}\right]$ falls $pq' < p'q$ (also mittels Multiplikation mit qq').

Bemerkung: Man beachte, dass damit das Rechnen mit rationalen Zahlen vollständig auf das mit ganzen Zahlen zurückgeführt wird.

1.3.4 Reelle Zahlen

Die Menge der reellen Zahlen wird zumeist mit \mathbb{R} bezeichnet.

Motiv für die Definition: Jetzt gibt es (unter anderem) Wurzeln für positive Zahlen.

Die Definition reeller Zahlen ist (vielleicht) mittels Dedekindscher Schnitte am einfachsten.

Definition 1.14 (Schnitt). *Zwei nichtleere Teilmengen A und B von \mathbb{Q} heißen Schnitt (A, B) , wenn gilt:*

- (1) $A \cup B = \mathbb{Q}$,
- (2) $a < b$ für alle $a \in A$ und $b \in B$
- (3) in B existiert kein minimales Element.

Definition 1.15 (reelle Zahl). *Eine reelle Zahl r ist ein Schnitt (A, B) .*

Einer rationalen Zahl r kann man durch

$$A_r = \{a \in \mathbb{Q} \mid a \leq r\}, \quad B_r = \{b \in \mathbb{Q} \mid b > r\} \quad (1.3.3)$$

eindeutig einen Schnitt zuordnen, sodass A_r ein maximales Element besitzt, naemlich r . Umgekehrt kann man alle solche Schnitte als rationale Zahl r auffassen (es gibt also eine Bijektion zwischen den rationalen Zahlen und den Schnitten mit maximalem Element in A).

Die Schnitte ohne ein maximales Element in A bilden nun unsere neuen, *irrationalen* Zahlen. Solche gibt es!

Beispiel: $A = \{a \in \mathbb{Q} \mid a \leq 0 \text{ oder } (a > 0 \text{ und } a^2 < 2)\}$, $B = \{b \in \mathbb{Q} \mid b > 0 \text{ und } b^2 > 2\}$.

Da keine rationale Zahl $\frac{p}{q}$ die Gleichung $\left(\frac{p}{q}\right)^2 = 2$ erfüllt, ist hier $A \cup B = \mathbb{Q}$. In der Tat gilt

Satz 1.16. *Die Wurzel aus 2 ist nicht rational.*

Beweis indirekt. Angenommen $\left(\frac{p}{q}\right)^2 = 2$. Dann findet man auch p und q derart, dass der Bruch nicht mehr gekürzt werden kann. Es folgt nun $p^2 = 2q^2$, also ist p^2 gerade. Damit muss auch p gerade sein (denn das Quadrat ungerader Zahlen ist immer ungerade). Sei also $p = 2k$. Wir erhalten dann $p^2 = 4k^2$ und $q^2 = \frac{1}{2}(4k^2)$. Das zeigt aber, dass auch q^2 und somit q gerade sein muss. Es kann also dennoch immer gekürzt werden. Dieser Widerspruch beweist den Satz. \square

Es bleiben also tatsaechlich Schnitte übrig ohne ein maximales Element in A .

Jetzt ist die Gleichung $x^2 = c$ stets lösbar für $c \geq 0$ (mit Schnitten analog zu $\sqrt{2}$, s. Beispiel).

Wir leiten nun die wichtigsten Eigenschaften reeller Zahlen her. Seien $r = (A, B), r' = (A', B')$ reelle Zahlen. Wir definieren

$$r \leq r' \text{ falls } A \subset A' \text{ bzw. } r < r' \text{ falls } A \subset A' \text{ und } A \neq A'.$$

Diese Ordnung stimmt also mit der üblichen Ordnung rationaler (und aus der Schule bekannter reeller) Zahlen überein.

Man sieht zugleich, dass $r = (A, B)$ in ein beliebig kleines Intervall $[a, b]$ mit rationalen $a < b$ ($a \in A$ und $b \in B$) gepackt werden kann.

Beweis: Intervallschachtelung durch Halbierung: Für den Beweis reicht eine spezielle Intervallschachtelung. Dazu startet man mit beliebigen $a' \in A$ und $b' \in B$ und bildet den Mittelpunkt $c = \frac{1}{2}(a' + b')$, der wieder rational ist. Ist $c \in A$, betrachtet man weiter das rechte halbierte Intervall, d.h., man setzt $a'' = c$ und $b'' = b'$; ist $c \in B$, setzt man $a'' = a'$ und $b'' = c$.

Nun sind $a'' \in A$ und $b'' \in B$ sowie $b'' - a'' = \frac{1}{2}(b' - a')$. Erneut betrachtet man den Mittelpunkt c zu a'', b'' , usw. Nach endlich vielen Schritten erhält man die gesuchten Elemente. \square

Folgerung:

Sind zwei reelle Zahlen $r < r'$ verschieden, unterscheiden sich A bzw. A' . Damit gibt es ein rationales x mit $r < x < r'$. Also gibt es auch (kleine) rationale Intervalle $[a, b], [a', b']$ (zu r, r' wie oben) mit $b < x < a'$.

Intervallschachtelung lässt sich verallgemeinern.

Satz 1.17 (Satz über Intervallschachtelung). *Seien, für $k = 1, 2, 3, \dots$ jeweils r_k und s_k reelle Zahlen mit den Eigenschaften*

$$r_k \leq r_{k+1} < s_{k+1} \leq s_k \quad \text{und} \quad s_{k+1} - r_{k+1} = \frac{1}{2}(s_k - r_k).$$

(Dann sind $I_k = [r_k, s_k]$ Intervalle mit $I_{k+1} \subset I_k$ und jeweils halber Länge)

Behauptung: Dann gibt es genau eine reelle Zahl r , sodass $r \in I_k$ für alle k gilt.

Beweis. Man definiere den Schnitt

$$B = \{b \in \mathbb{Q} \mid b > s_k \text{ für wenigstens ein } k \geq 1\} \text{ und } A = \mathbb{Q} \setminus B. \quad (1.3.4)$$

Zunächst definiert (A, B) per Definition eine reelle Zahl $r = (A, B)$. Weiter folgt $r \geq r_n \forall n$: Wäre $r < r_n$ für ein n , folgte auch $b < r_n$ für ein $b \in B$ wegen *Intervallschachtelung 1*. Mit diesem $b \in B$ folgt über (1.3.4) die Existenz eines k , sodass $b > s_k$. Wegen $s_k > r_n$, ist dann aber auch die umgekehrte Ungleichung $b > r_n$ richtig, was offenbar unmöglich ist. Also muss $r \geq r_n \forall n$ richtig sein.

Analog zeigt man $r \leq s_n \forall n$. Also ist $r \in I_k$ für alle k richtig.

Die Einzigkeit von r ergibt sich aus der Tatsache, dass die Länge der Intervalle I_k beliebig klein wird. \square

Tatsächlich gelten beide Aussagen auch dann, wenn die Intervalle beliebig klein werden. Sie zu halbieren, ist nur eine von vielen Möglichkeiten. Weiter sieht man leicht (mittels $\frac{p}{q}$ Darstellung): Ist x rational und r irrational, so ist $x + r$ irrational und (wenn $x \neq 0$) auch xr irrational.

Man beweise damit (indem man Vielfache xr von $r = \sqrt{2}$ betrachtet)

Lemma 1.18. *Zwischen zwei verschiedenen rationalen Zahlen liegt stets eine irrationale.*

Existenz des Supremums und Infimums

Eine Menge $M \subset \mathbb{R}$ heißt nach oben beschränkt, wenn es eine Konstante $K \in \mathbb{R}$ gibt, sodass $m \leq K \forall m \in M$.

Wir fragen nun nach einer scharfen Abschätzung, d.h., nach der kleinsten all solcher Zahlen K . Diese heißt Supremum von M ; $r = \sup M$. Analog heißt die größte untere Schranke Infimum, $\inf M$.

Satz 1.19. *Jede nach oben beschränkte, nichtleere Menge $M \subset \mathbb{R}$ besitzt ein Supremum $\sup M$.*

Beweis. Wenn M ein größtes Element besitzt, sind wir fertig; es ist das gesuchte Supremum. Wir betrachten den anderen Fall und benutzen Intervallschachtelung als Beweismittel. Mit irgendeiner oberen Schranke K und irgendeinem Element $m \in M$ bilden wir $r_1 = m$ und $s_1 = K$.

Dann ist $r_1 < s_1$. Weiter sei t_1 der Mittelpunkt des Intervalls $[r_1, s_1]$, $t_1 = \frac{1}{2}(r_1 + s_1)$. Nun gibt es zwei Fälle.

- 1) t_1 ist ebenfalls obere Schranke für M .
Dann ersetzen wir s_1 durch t_1 , d.h. wir betrachten weiter die Zahlen $r_2 = r_1$, $s_2 = t_1$.
- 2) t_1 ist keine obere Schranke für M . Dann gibt es ein $m' > t_1$, $m' \in M$.
Dann ersetzen wir r_1 durch t_1 , d.h. wir betrachten weiter die Zahlen $r_2 = t_1$, $s_2 = s_1$.

In beiden Fällen gibt es nun ein $m' \in M$ mit $m' > r_2$, und $s_2 > r_2$ ist wieder obere Schranke für M . Mit dem Intervall $[r_2, s_2]$ und dem Mittelpunkt $t_2 = \frac{1}{2}(r_2 + s_2)$ wiederholen wir nun diesen Schluss. So erhalten wir eine Folge von Intervallen $I_k = [r_k, s_k]$, wobei gilt für $k = 2, 3, \dots$:

$$r_{k-1} \leq r_k < s_k \leq s_{k-1},$$

r_k ist keine obere Schranke für M , s_k ist obere Schranke für M ,

$$s_k - r_k = \frac{1}{2}(s_{k-1} - r_{k-1}).$$

Die durch diese Intervallschachtelung definierte reelle Zahl $r \in I_k$ erfüllt nun $r = \sup M$. \square

Bemerkung: Man kann auch einfacher definieren: $B = \{b \in \mathbb{Q} \mid b > m \forall m \in M\}$, $A = \mathbb{Q} \setminus B$, $r = (A, B)$. Damit beweist man:

1. r ist obere Schranke: Sonst gilt $r < m$ für ein festes $m \in M$.
Dann gibt es nach Definition von „ $<$ “ wegen $B(r) \setminus B(m) \neq \emptyset$ (das sind die „rechten Mengen“ zum entsprechenden Schnitt) auch ein $b \in B(r) \setminus B(m)$. Somit ist $b \in B = B(r)$ und $b \in A(m)$, also gilt auch $b > m \geq b$, was offenbar unmöglich ist.
2. Keine obere Schranke r' ist kleiner als r : Sonst ist $r' < r$. Wähle $b' \in B' \setminus B$.
Nun ist $r' < b'$ (nach Definition von „ $<$ “) und $b' \leq m'$ für wenigstens ein $m' \in M$ (weil $b' \notin B$).
Es folgt so $r' < b' \leq m'$; also kann r' keine obere Schranke für M sein.

Rechenregeln

Die üblichen Rechenregeln mit rationalen Zahlen sind nun auf beliebige reelle Zahlen zu erweitern unter Anwendung der Definition eines Schnittes. Das ist etwas mühselig, aber möglich.

Addition: $r'' = r + r'$: Man bilde

$$A'' = A + A' := \{a'' \mid a'' = a + a' \text{ mit } a \in A, a' \in A'\}$$

und

$$B'' = B + B' := \{b'' \mid b'' = b + b' \text{ mit } b \in B, b' \in B'\}.$$

Nun kann man $r'' = r + r'$ als $r'' = (A'', B'')$ definieren (Dazu muss man $A'' \cup B'' = \mathbb{Q}$ zeigen).

Multiplikation positiver Zahlen $r'' = rr'$: Man bilde

$$A'' = AA' := \{a'' \mid a'' = aa' \text{ mit } a \in A, a' \in A' \text{ beide positiv}\} \cup \{a'' \mid a'' \in \mathbb{Q}, a'' \leq 0\}$$

und

$$B'' = BB' := \{b'' \mid b'' = bb' \text{ mit } b \in B, b' \in B'\}.$$

Nun kann man wieder $r'' = rr'$ als $r'' = (A'', B'')$ definieren. (Dazu ist ebenfalls $A'' \cup B'' = \mathbb{Q}$ zu zeigen).

Multiplikation mit -1: $r'' = -r, r \neq 0$. Man bilde

$$A'' = -B = \{a'' \mid a'' = -b \text{ mit } b \in B\}$$

und

$$B'' = -A := \{b'' \mid b'' = -a \text{ mit } a \in A\}.$$

Im Falle eines rationalen r hat jetzt B'' ein minimales Element β , aber A'' besitzt kein maximales.

Um der Forderung (3) eines Schnittes zu genügen, schieben wir dann einfach β aus B'' zu A'' .

Nun kann man wieder r'' als $r'' = (A'', B'')$ definieren.

In üblicher Weise lassen sich die restlichen Operationen erklären; die Multiplikation zweier negativer Zahlen mittels $r'' = (-r)(-r')$, die Multiplikation von $r < 0$ mit $r' > 0$ mittels $r'' = (-1)(-rr')$, die Multiplikation mit 0 als Null.

Natürlich rechnet praktisch kein Mensch explizit in Form von Schnitten. Statt dessen benutzt man eine zweckmäßige Darstellung der Zahl, z.B. im Dezimalsystem. Man beachte jedoch, dass dann $r = 1.111111\dots$ schon mit einer unendlichen Reihe identifiziert wird: $r = \sum_{k=0}^{\infty} 10^{-k}$.

1.3.5 Körpereigenschaften

Die rationalen Zahlen \mathbb{Q} und die reellen Zahlen \mathbb{R} bilden jeweils eine Menge (mit Elementen $0, 1, a, b, c, \dots$) in welcher eine **Addition und Multiplikation** erklärt sind, die den folgenden Forderungen genügen:

$$\begin{array}{llll} a + b & = & b + a & \text{und } ab & = & ba & \text{(Kommutativität)} \\ (a + b) + c & = & a + (b + c) & \text{und } (ab)c & = & a(bc) & \text{(Assoziativität)} \\ a + 0 & = & a & \text{und } a \cdot 1 & = & a & \text{(es gibt neutrale Elemente, genannt 0 bzw. 1)} \\ a(b + c) & = & ab + ac & & & & \text{(Distributivgesetz)} \end{array}$$

Desweiteren gilt: $a + x = b$ ist stets lösbar und $ax = b$ ist lösbar falls $a \neq 0$.

Die Lösbarkeitsforderungen können auch durch die formal schwächere Bedingung $a + x = 0$ ist lösbar und $ax = 1$ ist lösbar falls $a \neq 0$ ersetzt werden.

(Wieso ist dann auch $a + y = b$ lösbar? Wenn $a + x = 0$, folgt nämlich $a + x + b = b$. Damit erfüllt $y = x + b$ offenbar $a + y = b$. Analog für $ax = b$.)

Eine Menge mit einer solchen Addition und Multiplikation nennt man (Zahlen-) Körper. Oft verlangt man noch $0 \neq 1$, um den trivialen Fall $M = \{0\}$ auszuschließen.

\mathbb{Q} und \mathbb{R} sind nicht die einzigen Zahlkörper, z.B. wird auch die Menge $M = \{0, 1\}$ mit den Rechenregeln

$$\begin{aligned} 1 + 0 &= 0 + 1 = 1, & 0 + 0 &= 0, & 1 + 1 &= 0 \\ 1 \cdot 0 &= 0 \cdot 1 = 0, & 0 \cdot 0 &= 0, & 1 \cdot 1 &= 1 \end{aligned}$$

ein Zahlkörper (man prüfe, ob die Axiome/Forderungen erfüllt sind).

Die Menge der komplexen Zahlen bildet einen weiteren Körper. Das Motiv für deren Definition besteht darin, dass man auch Wurzeln für negative Zahlen erhält, was erstaunlich viele Konsequenzen für die Lösbarkeit von Gleichungen besitzt.

1.3.6 Komplexe Zahlen

Wir betrachten die Punkte (x, y) des \mathbb{R}^2 jetzt als eine Menge von „Zahlen“, die dadurch entstehen, dass wir die Elemente der Form $(x, 0)$ als reelle Zahl x interpretieren, das Element $(0, 1)$ dagegen als eine Zahl i neuen Typs, die der Gleichung $i^2 = -1$ genügt. Wir wollen die Grundrechenarten auf alle Elemente (x, y) ausdehnen, sodass erneut ein Zahlkörper entsteht.

Für diese Paare schreiben wir auch $x + iy$ und nennen x Realteil, y Imaginärteil der komplexen Zahl $z = x + iy$ (x, y reell). Das Zeichen „+“ hat hierbei zunächst nur den Charakter eines Symbols.

Die Definition von Addition und Multiplikation erfolgt allerdings genauso, wie wir mit diesem „+“ rechnen würden:

Addition:

$$(x + iy) + (x' + iy') := (x + x') + i(y + y')$$

(reelle Addition von Realteil und Imaginärteil auf rechter Seite)

Multiplikation:

Die Multiplikation ist unter Berücksichtigung von $i^2 = -1$ ebenso definiert wie die gewöhnliche Multiplikation der Binome $(x + iy)$ und $(x' + iy')$, wenn „+“ als Addition interpretiert wird, naemlich

$$(x + iy) \cdot (x' + iy') = xx' + ix'y' + iyx' + i^2yy' = (x \cdot x' - y \cdot y') + i(x \cdot y' + x' \cdot y)$$

(reelle Multiplikation „ \cdot “ und Addition „+“ bzw. Subtraktion „-“ auf rechter Seite).

Man beachte, dass die Addition komplexer Zahlen damit der Addition (von Vektoren) im \mathbb{R}^2 entspricht. Die Multiplikation ist neu und hat auch nichts mit dem Skalarprodukt zu tun.

Mit diesen Operationen erhalten wir tatsächlich einen Zahlkörper, wie man mit einiger Mühe anhand der Axiome nachrechnet.

Die neutralen Elemente sind $0 + 0i$, kurz einfach als 0 bezeichnet, in Bezug auf die Addition, sowie $1 + 0i$, kurz 1 , für die Multiplikation.

Die zu $x + iy$ inversen Elemente sind $-x + i(-y)$, kurz $-(x + iy)$, bzgl. Addition und

$$\frac{1}{x + iy} = \frac{x - iy}{(x + iy) \cdot (x - iy)} = \frac{x - iy}{x \cdot x + y \cdot y} = \frac{x - iy}{x^2 + y^2}$$

bzgl. Multiplikation.

Die Zahl $\bar{z} = x - iy$ heißt *konjugiert komplexe Zahl* zu $z = x + iy$.

Die nichtnegative Zahl $\sqrt{x^2 + y^2}$ (Länge der Strecke zwischen 0 und z) heißt *Betrag* von z , kurz $|z|$.

Damit ist

$$\bar{z} \cdot z = x^2 + y^2 = |z|^2 \quad \text{und} \quad z^{-1} = \frac{\bar{z}}{|z|^2}.$$

Die Gleichung $z^2 = -1$ hat nun (genau) zwei Lösungen: $-i$ und i .

Die Menge der komplexen Zahlen bezeichnen wir mit dem Symbol \mathbb{C} . Für kompliziertere Aussagen ist es nützlich, auf die Eulersche Darstellung komplexer Zahlen (s. Kap. 7.1) zurückzugreifen, die auf dem Zusammenhang zwischen e -Funktion und \sin bzw. \cos (hergestellt mittels ihrer Taylor-Reihen) basiert.

In \mathbb{Q} und \mathbb{R} hat man noch zusätzliche Eigenschaften der Ordnungsrelation: Es ist genau eine der Relationen $a > 0$, $a = 0$, $a < 0$ richtig.

Aus $a > 0$ und $b > 0$ folgt $a + b > 0$ und $ab > 0$.

Zu jedem a existiert eine natürliche Zahl n mit $n > a$ (Archimedisches Axiom).

Zahlkörper mit diesen Eigenschaften heißen auch Archimedisch. Für komplexe Zahlen gibt es keine ähnlich schöne Ordnung.

1.3.7 Polarkoordinaten und Wurzeln

Die Gleichung $z^2 = -1$ hat nun (genau) zwei Lösungen: $\pm i$. Jede komplexe Zahl $z = x + iy$ lässt sich in der Form

$$z = r(\cos \varphi + i \sin \varphi) \tag{1.3.5}$$

schreiben. Die komplexe Zahl in der Klammer hat wegen $\sin^2(\varphi) + \cos^2(\varphi) = 1$ den Betrag 1, also gilt $r = |z|$.

Für $r > 0$ ist φ im Intervall $[0, 2\pi)$ eindeutig durch die Gleichungen

$$\cos \varphi = \frac{x}{r} \quad \text{und} \quad \sin \varphi = \frac{y}{r}$$

festgelegt:

$$\varphi = \frac{\pi}{2} \text{ if } x = 0, y > 0, \quad \varphi = \frac{3\pi}{2} \text{ if } x = 0, y < 0, \quad \varphi = \arctan\left(\frac{y}{x}\right) \text{ if } x \neq 0$$

Achtung: Auf Periode π bei \arctan achten und immer überlegen, in welchem Orthanten (x, y) liegt! Oft legt man die Werte von \arctan auch ins Intervall $(-\pi/2, \pi/2)$.

Die Darstellung (1.3.5) von z heißt Darstellung in Polarkoordinaten (r und φ).

Ihr Sinn ist bei der Multiplikation erkennbar: Hat nämlich eine zweite komplexe Zahl ζ die Form

$$\zeta = s(\cos \omega + i \sin \omega),$$

so erhält man für das Produkt aus den Additionstheoremen für Sinus und Cosinus:

$$\begin{aligned} \zeta \cdot z &= rs(\cos \varphi + i \sin \varphi)(\cos \omega + i \sin \omega) \\ &= rs[(\cos \varphi \cos \omega - \sin \varphi \sin \omega) + i(\sin \varphi \cos \omega + \sin \omega \cos \varphi)] \\ &= rs(\cos(\varphi + \omega) + i \sin(\varphi + \omega)). \end{aligned}$$

Die Beträge werden multipliziert, die Winkel addiert.

Durch vollständige Induktion ergibt sich hieraus (Satz von Moivre) insbesondere für alle $n \in \mathbb{N}$:

$$\begin{aligned} z^n &= r^n (\cos(n\varphi) + i \sin(n\varphi)) \\ z^{-n} &= \frac{1}{z^n} = \frac{\bar{z}^n}{|z^n|^2} = \frac{r^n (\cos(n\varphi) - i \sin(n\varphi))}{r^{2n}} = r^{-n} (\cos(-n\varphi) + i \sin(-n\varphi)). \end{aligned}$$

Wir bestimmen nun die Lösungen der Gleichung $\zeta = z^{\frac{1}{n}} = \sqrt[n]{z}$, $n \in \mathbb{N} \setminus \{0\}$, d.h., die n -te Wurzel. Die Gleichung bedeutet

$$z = \zeta^n = s^n (\cos(n\omega) + i \sin(n\omega)), \text{ d.h. } r = s^n \text{ sowie } \cos \varphi + i \sin \varphi = \cos(n\omega) + i \sin(n\omega).$$

Die zweite Gleichung gilt genau dann, wenn

$$n\omega = \varphi + 2k\pi$$

mit einer ganzen Zahl k richtig ist. Man beachte, dass dies für genau n verschiedene Werte $w \in [0, 2\pi)$ gilt, nämlich

$$\omega = \frac{\varphi + 2k\pi}{n} \text{ mit } k = 0, 1, \dots, n-1.$$

Damit besitzt $z = r(\cos \varphi + i \sin \varphi)$ die n verschiedenen Wurzeln

$$z = r^{\frac{1}{n}} \left[\cos\left(\frac{\varphi + 2k\pi}{n}\right) + i \sin\left(\frac{\varphi + 2k\pi}{n}\right) \right]; \quad k = 0, 1, \dots, n-1.$$

Den Wert zu $k = 0$ nennt man auch *Hauptwert*. Analog bestimme man die n Lösungen z von $\zeta^n = z^m$. Zusammenfassend erhält man so für rationale Exponenten q die Darstellung des Hauptwertes als

$$z^q = r^q (\cos(q\varphi) + i \sin(q\varphi)),$$

Kapitel 2

Folgen im Reellen

Wir betrachten in diesem Abschnitt (unendliche) Folgen $\{a_n\}$ ($n = 1, 2, 3, \dots$) reeller Zahlen.

Eine solche Folge kann als Funktion aufgefasst werden, die jeder natürlichen Zahl $n \geq 1$ eine reelle Zahl a_n zuschreibt. Dabei wird es gleichgültig sein, mit welcher Zahl n man beginnt; weil sich alle folgenden (wesentlichen) Aussagen und Definitionen *nur auf große n* beziehen werden.

Gilt eine Aussage für alle n ab einem bestimmten n_0 (also für alle n bis auf endlich viele Ausnahmen) so sagt man auch, sie gelte für *fast alle n* .

Definition 2.1. Eine Folge $\{a_n\}$ heißt konvergent, falls es eine (reelle) Zahl a gibt, sodass mit jedem positiven ε die Ungleichung

$$(*) \quad |a_n - a| < \varepsilon$$

richtig ist, sofern nur n hinreichend groß (größer als ein gewisses $n(\varepsilon)$) ist.

Kurz: Wie klein $\varepsilon > 0$ auch gewählt wird, $(*)$ gilt stets für fast alle n .

Die Zahl a heißt in diesem Falle Grenzwert (oder Limes) der Folge.
Für diesen Sachverhalt werden folgende Bezeichnungen benutzt:

$$a_n \rightarrow a, \quad \lim_{n \rightarrow \infty} a_n = a, \quad \{a_n\} \rightarrow a.$$

Oft schreiben wir auch nur $\lim a_n = a$, wenn (wie hier) klar ist, dass $n \rightarrow \infty$ streben soll.

Eine nicht-konvergente Folge heißt divergent.

Ferner vereinbaren wir:

$\lim a_n = \infty$, falls mit jedem $\varepsilon > 0$ gilt: $a_n > \frac{1}{\varepsilon}$ für fast alle n . Diese Folge ist dann divergent.

Beispiele 2.2 (für konvergente Folgen).

- $a_n = \frac{1}{n}$, $n(\varepsilon) > \frac{1}{\varepsilon}$, $\lim a_n = 0$
- $a_n = \frac{\sqrt{n}}{n}$, $\lim a_n = 0$
- $a_n = \frac{n}{1+n}$, $\lim a_n = 1$

Beispiele 2.3 (für divergente Folgen).

- $a_n = (-1)^n$
- $a_n = n$

Definition 2.4 (beschränkt). Eine Folge $\{a_n\}$ heißt beschränkt, wenn es eine obere Schranke für die Beträge aller ihrer Glieder a_n gibt:

$$\exists S : |a_n| < S \quad \forall n \in \mathbb{N} \quad (S \text{ unabhängig von } n).$$

Offenbar ist jede konvergente Folge beschränkt, aber nicht jede beschränkt Folge auch konvergent (s.o.).

Definition 2.5 (monoton, streng monoton). Eine Folge $\{a_n\}$ heißt *monoton wachsend* (auch *uneigentlich monoton wachsend*), wenn

$$a_n \leq a_{n+1} \quad \forall n \in \mathbb{N}.$$

Ist stets $a_n < a_{n+1}$, nennt man die Folge *oft streng monoton* (auch *strikt*) *wachsend*.

Analog definiert man *monoton fallende Folgen*. Schließlich heißt eine Folge *monoton*, wenn sie *monoton wächst* oder *monoton fällt*.

Definition 2.6 (Teilfolge). Eine (unendliche) Teilfolge der Folge $\{a_n\}$ entsteht durch die Auswahl unendlich vieler ihrer Glieder und Beibehalten ihrer Reihenfolge.

Das bedeutet, dass wir jeder natürlichen Zahl k ein $n(k)$ mit $n(k) < n(k+1)$ zuordnen und eine Folge $\{b_k\}$ durch $b_k = a_{n(k)}$ definieren.

Man sieht leicht:
Jede Folge kann nur einen Grenzwert besitzen, und im Falle $\lim a_n = a$ hat auch jede Teilfolge der Folge $\{a_n\}$ den Grenzwert a . (Übungsaufgabe)

Die nachstehenden Aussagen sind für das Rechnen mit Folgen grundlegend.

Satz 2.7. Jede beschränkte und monotone (reelle) Folge ist konvergent.

Beweis. Angenommen $a_n \leq a_{n+1}$ für alle n . Wir zeigen, dass die kleinste obere Schranke $S := \sup\{a_n \mid n \geq 1\}$ Grenzwert der betrachteten Folge ist. Dazu sei $\varepsilon > 0$ beliebig fixiert. Weil $S - \varepsilon$ keine obere Schranke aller a_n ist, findet sich ein spezielles n , etwa $n(\varepsilon)$, mit welchem $a_{n(\varepsilon)} > S - \varepsilon$ gilt. Nutzt man nun die Monotonie aus, folgt erst recht

$$a_n \geq a_{n(\varepsilon)} > S - \varepsilon \quad \forall n > n(\varepsilon).$$

Da andererseits $S \geq a_n$ gilt, folgt schon $|a_n - S| < \varepsilon$ für alle $n > n(\varepsilon)$. Im Falle $a_n \geq a_{n+1}$ betrachtet man analog die größte untere Schranke $S := \inf\{a_n \mid n \geq 1\}$. \square

Lemma 2.8. (Einschachteln) Sind $\{a_n\}$ und $\{b_n\}$ zwei konvergente Folgen mit demselben Grenzwert a und gilt für die Elemente c_n einer weiteren Folge $a_n \leq c_n \leq b_n$ für fast alle n , so gilt auch $\lim c_n = a$.

Beweis. Zu jedem $\varepsilon > 0$ existieren nach Voraussetzung $n_1(\varepsilon)$ und $n_2(\varepsilon)$, sodass

$$|a_n - a| < \varepsilon \quad \forall n > n_1(\varepsilon) \quad \text{und} \quad |b_n - a| < \varepsilon \quad \forall n > n_2(\varepsilon)$$

richtig sind. Damit folgt aber auch $|c_n - a| < \varepsilon$ für fast alle n . \square

Lemma 2.9 (Rechnen mit Folgen). Es seien zwei konvergente Folgen $\{a_n\}$ und $\{b_n\}$ gegeben mit den Grenzwerten a bzw. b . Dann konvergiert

- die Folge $\{c_n\}$ mit $c_n = a_n + b_n$ gegen $a + b$
- die Folge $\{p_n\}$ mit $p_n = a_n \cdot b_n$ gegen $a \cdot b$
- die Folge $\{q_n\}$ mit $q_n = \frac{a_n}{b_n}$ gegen $\frac{a}{b}$, falls $b \neq 0$

Beweis. Die letzte Folge ist nur für $b_n \neq 0$ definiert ist, was jedoch für große n aus $b \neq 0$ folgt. Für die endlich vielen übrigen n denke man sich irgendwelche Zahlen q_n geschrieben. Wir betrachten nur die Folge $\{p_n\}$.

Schreibt man $a_n = a + x_n$, $b_n = b + y_n$, so konvergieren offenbar x_n und y_n gegen Null.

Wegen $p_n = a \cdot b + (a \cdot y_n + b \cdot x_n + x_n y_n)$ sieht man, dass die Klammer im Betrag kleiner als jedes positive ε wird, wenn n hinreichend groß ist. Man setze etwa

$$\delta = \frac{\varepsilon}{3M} \quad \text{mit} \quad M = |a| + |b| + 1$$

und wähle $n(\varepsilon)$ so groß, dass $|x_n|$ und $|y_n|$ beide $< \delta$ sind, falls $n > n(\varepsilon)$. \square

Der folgende Satz liefert eine der am häufigsten angewandten Aussagen der Analysis.

Satz 2.10 (von Bolzano und Weierstraß). Jede beschränkte (reelle) Folge besitzt eine konvergente Teilfolge.

Beweis. Man definiere die Suprema

$$S_n = \sup\{a_m \mid m > n\}.$$

Sie sind monoton, d.h. $S_{n+1} \leq S_n$, weil die Menge der fraglichen m kleiner wird. Sie sind beschränkt, weil alle $|a_m|$ beschränkt sind. Also existiert nach Satz 2.7 der Limes $S = \lim S_n$. Sei $n(1) = 1$. Beginnend mit $k = 1$ wähle man eine natürliche Zahl $m > n(k)$, sodass $a_m > S_{n(k)} - \frac{1}{k}$. Sie existiert, weil $S_{n(k)} - \frac{1}{k}$ keine obere Schranke ist. Andererseits ist auch $S_{n(k)} \geq a_m$. Anschliessend setzen wir $n(k+1) = m$, bilden k aus $k+1$ und wiederholen die procedure. Wegen

$$S_{n(k)} - \frac{1}{k} < a_{n(k+1)} < S_{n(k)}, \quad (k = 1, 2, \dots)$$

konvergiert die eingeschachtelte Teilfolge der $a_{n(k+1)}$ nun ebenfalls gegen S . \square

Die folgende Definition ist grundlegend für die gesamte Analysis.

Definition 2.11 (Cauchy-Folge). *Eine Folge $\{a_n\}$ heißt Cauchy-Folge oder auch Fundamentalfolge, wenn zu jedem $\varepsilon > 0$ ein solches $n(\varepsilon)$ existiert, dass gilt*

$$(1) \quad |a_m - a_n| < \varepsilon \quad \forall n, m > n(\varepsilon).$$

Man sieht leicht, dass jede konvergente Folge eine Cauchy-Folge ist (wieso?). Umgekehrt gilt:

Satz 2.12. *Jede (reelle) Cauchy-Folge ist konvergent.*

Beweis. Wählt man $\varepsilon = 1$, so gibt es nach Voraussetzung ein $n(1)$ mit $|a_m - a_{n(1)+1}| < 1 \quad \forall m > n(1)$. Die Folge ist somit beschränkt, besitzt also eine konvergente Teilfolge. Wir bezeichnen deren Elemente mit $a_{m(k)}$ ($k = 1, 2, \dots$) und den Grenzwert mit g . Sei nun $\varepsilon > 0$ beliebig. Dann gibt es wegen Konvergenz der Teilfolge ein $k(\varepsilon)$, sodass

$$|a_{m(k)} - g| < \varepsilon \quad \forall k > k(\varepsilon).$$

Mit dem laut (1) existierenden $n(\varepsilon)$ folgt weiter für $n > n(\varepsilon)$: $|a_n - a_{m(k)}| < \varepsilon$, wenn nur k hinreichend groß ist, nämlich so groß, dass auch $m(k) > n(\varepsilon)$. Beides zusammen liefert

$$|a_n - g| \leq |a_n - a_{m(k)}| + |a_{m(k)} - g| < 2\varepsilon \quad \forall n > n(\varepsilon).$$

Da $\varepsilon > 0$ beliebig war, bedeutet das aber $g = \lim a_n$. \square

In den angegebenen Sätzen ist es wichtig, dass wir reelle Folgen betrachten. Mit Elementen aus \mathbb{Q} kann man eine wachsende, beschränkte Folge $\{a_n\}$ (z.B. aus den ersten n Dezimalstellen von $\sqrt{2}$) bilden, die nicht gegen ein $a \in \mathbb{Q}$ konvergiert.

Mit ganzzahligen a_n wird Konvergenz trivial, weil sie $a_{n+1} = a_n$ für alle großen n bedeutet (Man nehme $\varepsilon < 1$).

2.1 Spezielle wichtige Grenzwerte

(A1) Sei $a_n = \sqrt[n]{n}$. Dann ist $\lim a_n = 1$.

Hierzu lässt sich die Bernoullische Ungleichung (s. Kap. 1.3.1, S. 11)

$$(1+x)^n \geq 1+nx \quad (x > -1, n \in \mathbb{N})$$

sehr effektiv benutzen. Beweistrick: Man setze $1+x = \sqrt[n]{n}$ ($= \sqrt[n]{\sqrt[n]{n}}$).

Dann ist $x = \sqrt[n]{\sqrt[n]{n}} - 1$, und es folgt $(1+x)^n = \sqrt[n]{n} \geq 1+n(\sqrt[n]{\sqrt[n]{n}} - 1)$, und damit auch $\sqrt[n]{n} \geq 1+n\sqrt[n]{\sqrt[n]{n}} - n$. Division durch n liefert:

$$1/\sqrt[n]{n} = \sqrt[n]{n}/n \geq 1/n + \sqrt[n]{\sqrt[n]{n}} - 1.$$

Also müssen - wegen $\lim 1/\sqrt[n]{n} = 0 = \lim 1/n$ - die positiven Zahlen $\sqrt[n]{\sqrt[n]{n}} - 1$ gegen Null streben.

Das liefert uns $\sqrt[n]{\sqrt[n]{n}} \rightarrow 1$ und schliesslich (mit den bekannten Eigenschaften der Wurzelfunktion) auch $\sqrt[n]{n} \rightarrow 1$.

Der Trick war wichtig. Setzt man einfacher $1+x = \sqrt[n]{n}$, so folgt analog nur $2 \geq 1/n + \sqrt[n]{n}$.

(A2) n -te Wurzel aus $\frac{1}{n}$: $a_n = \left(\frac{1}{n}\right)^{1/n} = \frac{1}{\sqrt[n]{n}}$, also $\lim a_n = 1$.

(A3) $a_n = \sqrt[n]{x}$, $\lim a_n = 1$ (für beliebige reelle $x > 0$).

1) Falls $x > 1$, gilt fuer große n : $1 < \sqrt[n]{x} < \sqrt[n]{n}$ und rechte Seite konvergiert gegen 1.

2) Falls $x < 1$, betrachte man $y = \frac{1}{x} > 1$: $\sqrt[n]{x} = \sqrt[n]{\frac{1}{y}} = \frac{1}{\sqrt[n]{y}}$

(A4) $a_n = \frac{n}{x^n}$, $\lim a_n = 0$, für beliebiges $x > 1$.

Es ist $\frac{a_{n+1}}{a_n} = \frac{n}{n+1} \cdot \frac{1}{x} < \frac{1}{x} < 1$. Also ist $a_{n+1} < q a_n$ für große n und $q = \frac{1}{x} < 1$, und deshalb $a_{n+m} < q^m a_n$ mit $q^m \rightarrow 0$ für $m \rightarrow \infty$.

(A5)

$$a_n = \sqrt{n+1} - \sqrt{n}, \quad \lim a_n = 0.$$

Als Bruch auffassen und mit der Summe erweitern:

$$a_n = \frac{(\sqrt{n+1} - \sqrt{n})(\sqrt{n+1} + \sqrt{n})}{\sqrt{n+1} + \sqrt{n}} = \frac{n+1-n}{\sqrt{n+1} + \sqrt{n}} \rightarrow 0$$

(A6) $a_n = |x|^n/n!$ Dann ist $\lim a_n = 0$.

Es gibt ein n_0 mit $|x| < n_0$. Dann ist der Quotient $q = \frac{|x|}{n_0+1}$ sicher kleiner als 1. Wir schreiben

$$a_{n_0+m} = \left(\frac{|x|}{1} \cdot \frac{|x|}{2} \cdots \frac{|x|}{n_0} \right) \left(\frac{|x|}{n_0+1} \cdot \frac{|x|}{n_0+2} \cdots \frac{|x|}{n_0+m} \right).$$

Das erste Produkt nennen wir C ; es hängt nicht von m ab. Der Quotient $\frac{|x|}{n_0+1}$ und alle folgenden sind $\leq q < 1$. Deshalb erhalten wir mit $m \rightarrow \infty$

$$a_{n_0+m} \leq C q^m \quad \text{wobei } q^m \rightarrow 0.$$

(A7) $a_n = \frac{2n+1}{n+\sqrt{n}-5}$; $\lim a_n = 2$

wegen $\frac{2n+1}{n+\sqrt{n}-5} = \frac{2+1/n}{1-\sqrt{n}/n+5/n}$ und $1/n \rightarrow 0$, $\frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}} \rightarrow 0$, $5/n \rightarrow 0$.

(A8) $a_n = \frac{2^n}{n^2}$, $\lim a_n = \infty$. Es gilt

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

Für große n (z.B. $n > 10$), ist also $\frac{a_{n+1}}{a_n} \geq p := 3/2 > 1$. Es folgt so mittels $a_{n+1} \geq p a_n$ die Behauptung. Durch Betrachten von $\frac{a_{n+1}}{a_n}$ kann man auch (A6) oder z.B. $\lim \frac{2^n}{n^8} = \infty$ zeigen.

Kapitel 3

Metrische Räume

3.1 Grundlegende Begriffe

Der im Reellen behandelte Begriff der Konvergenz lässt sich sehr allgemein für Folgen definieren, deren Glieder in Mengen liegen, für die ein „Abstand ihrer Elemente“ erklärt ist. Dies erlaubt, auch von der Konvergenz anderer Objekte, die keine reellen Zahlen sind, in einem wohldefinierten Sinne zu sprechen.

Definition 3.1 (Metrik, metrischer Raum). *Ein Paar (X, d) heißt metrischer Raum, wenn X eine Menge ist und d eine Funktion, die je zwei Elementen $x, y \in X$ eine reelle Zahl $d(x, y)$ derart zuordnet, dass gilt:*

- (1) $d(x, y) = d(y, x) \geq 0$
- (2) $d(x, y) = 0$ genau dann, wenn $x = y$
- (3) $d(x, z) \leq d(x, y) + d(y, z)$ sofern $x, y, z \in X$ (Dreiecksungleichung).

Eine solche Funktion d heißt Metrik oder Abstand. Ist eine Metrik bereits erklärt, spricht man häufig einfach vom metrischen Raum X . Seine Elemente nennen wir auch Punkte.

„Metrik“ ist eine Verallgemeinerung des im Reellen durch den Betrag gegebenen Abstandes $d(x, y) = |x - y|$, der obige Forderungen offensichtlich erfüllt.

Wenn wir von \mathbb{R} sprechen, werden wir stets annehmen, dass die Metrik in dieser Weise gegeben ist. Das bedeutet allerdings nicht, dass es hier nur diese eine Metrik gibt! Z.B. sind $d(x, y) = 8|x - y|$ und komplizierter $d(x, y) = \frac{|x-y|}{1+|x-y|}$ ebenfalls Metriken in \mathbb{R} . Man kann sogar definieren: $d(x, y) = 0$ falls $x = y$; $d(x, y) = 1$ falls $x \neq y$ (diskrete Metrik).

Es kann also dieselbe Menge mit verschiedenen Abständen versehen werden, was zu unterschiedlichen metrischen Räumen mit denselben Elementen führt.

Bezeichnung:

Es sei $B(x, \varepsilon) = \{y \in X \mid d(x, y) < \varepsilon\}$ die „offene Kugel“ um x mit Radius $\varepsilon > 0$.

Oft wird sie auch mit $B^0(x, \varepsilon)$ bezeichnet, wenn man $B(x, \varepsilon)$ als die abgeschlossene Kugel $B(x, \varepsilon) = \{y \in X \mid d(x, y) \leq \varepsilon\}$ definiert. Ausserdem nimmt man im Deutschen oft auch K statt B (ball). In der Folge ist also $B = B^0$.

Bemerkung: „Kugeln“ im metrischen Raum können äußerst „unregelmäßige Gebilde“ sein.

Wenn $X = \mathbb{R}$, so ist $B(x, \varepsilon)$ das Intervall $(x - \varepsilon, x + \varepsilon)$ (ohne die Randpunkte).

Die folgenden Begriffe sind für die Analysis, Geometrie und Topologie von fundamentaler Bedeutung und basieren (zumindest im metrischen Raum) allein auf dem Abstandsbegriff und der Konvergenz von Folgen. Deshalb ist dieses Kapitel schon hier eingeordnet.

3.1.1 Konvergenz

Eine Folge $\{x_n\}$ von Elementen aus X konvergiert gegen $x \in X$, sofern jede Kugel $B(x, \varepsilon)$ fast alle x_n enthält, d.h. alle x_n mit $n > n(\varepsilon)$. Dann schreiben wir wieder $x = \lim x_n$ und sagen, x sei Grenzelement

oder auch Limes der Folge.

Offenbar ist dies im Falle $X = \mathbb{R}$ die schon bekannte Konvergenz.

Definition 3.2 (innerer Punkt, offen, abgeschlossen). Sei M eine Teilmenge von X .

- Ein Punkt $x \in M$ heißt innerer Punkt von M , wenn $B(x, \varepsilon) \subset M$ für wenigstens ein $\varepsilon > 0$.
- Die Menge aller inneren Punkte von M bezeichnet man mit $\text{int}M$ (von interior = Inneres).
- M heißt offen, wenn M nur innere Punkte enthält (also $\text{int}M = M$ ist).
- Die Komplemente offener Mengen, d.h. alle $A = X \setminus M$ (M offen), heißen abgeschlossen.

Damit gibt es wenigstens zwei zugleich offene und abgeschlossene Mengen: Den ganzen Raum X und die leere Menge.

Außerdem sind alle „Kugeln“ $B(x, \varepsilon)$ offen, denn wenn $x' \in B(x, \varepsilon)$, so findet sich wegen $d(x', x) < \varepsilon$ ein $\delta > 0$ mit $\delta + d(x', x) < \varepsilon$, und jeder Punkt $x'' \in B(x', \delta)$ liegt deshalb wegen $d(x'', x) \leq d(x'', x') + d(x', x) < \delta + d(x', x) < \varepsilon$ (Dreiecksungleichung!) ebenfalls in $B(x, \varepsilon)$.

Alle einelementigen Teilmengen von X sind abgeschlossen. Darüberhinaus gilt

Lemma 3.3. Es seien I eine beliebige Menge und M_i für jedes $i \in I$ eine Teilmenge von X . Dann ist

- die Vereinigung aller Mengen M_i offen, sofern alle M_i offen sind
- der Durchschnitt aller M_i abgeschlossen, sofern alle M_i abgeschlossen sind

Ist I eine endliche Menge, so gilt außerdem:

- Der Durchschnitt aller M_i ist offen, sofern alle M_i offen sind.
- Die Vereinigung aller M_i ist abgeschlossen, sofern alle M_i abgeschlossen sind.

Beweis. Ist I leer, wird die Aussage trivial. Wir betrachten den interessanteren Fall.

Seien alle M_i offen, und sei $x \in M := \bigcup_{i \in I} M_i$.

Fixiert man ein $i_0 \in I$ mit $x \in M_{i_0}$, so gibt es eine Kugel $B(x, \varepsilon)$, die ganz in M_{i_0} liegt, folglich auch in M .

Sei nun x aus dem Durchschnitt aller M_i . Zu jedem $i \in I$ gibt es, da M_i offen ist, ein $\varepsilon(i)$, sodass $B(x, \varepsilon(i)) \subset M_i$. Für endliches I bleibt das kleinste der $\varepsilon(i)$, $\varepsilon = \min_{i \in I} \{\varepsilon(i)\}$ positiv, und es ist $B(x, \varepsilon)$ in allen M_i enthalten, was zu zeigen war.

Die Aussagen zur Abgeschlossenheit folgen durch Komplementbildung (Übungsaufgabe). □

Lemma 3.3 sagt insbesondere, dass der Durchschnitt D aller abgeschlossenen Mengen A , die eine festgewählte Teilmenge M von X enthalten, wieder abgeschlossen ist. Offenbar gibt es keine kleinere abgeschlossene Menge (im Sinne der Mengeninklusion), die M ebenfalls enthält.

- D heißt Abschließung der Menge M , kurz $D = cM$ (lies closure = Abschließung)
- Für die Differenz $cM \setminus \text{int}M$ ist die Bezeichnung bdM (boundary = Rand, Grenze) üblich
- Ein Häufungspunkt x^* von M ist dadurch charakterisiert, dass in jeder Kugel $B(x^*, \varepsilon)$ wenigstens zwei Punkte aus M liegen (und damit unendlich viele, wieso?). Man überlege sich, dass dann x^* immer zu cM gehören muss.
- Die Punkte aus M , die keine Häufungspunkte (von M) sind, heißen isolierte Punkte von M .
- Eine Teilmenge A von M heißt dicht in M , wenn es zu jedem $x \in M$ und $\varepsilon > 0$ ein $a \in A$ mit $x \in B(a, \varepsilon)$ gibt.
(Die letzte Bedingung fordert mit anderen Worten, dass M in cA enthalten ist (vgl. Lemma 3.4).)
- Enthält M eine abzählbare dichte Teilmenge A , so heißt M separabel.

Beispiel:

In \mathbb{R} sei $M = \{x \in \mathbb{R} \mid 0 < x \leq 1\} \cup \{\pi\}$. Dann gilt:

$$\text{int}M = (0, 1) := \{x \in \mathbb{R} \mid 0 < x < 1\},$$

$$cM = M \cup \{0\},$$

π ist einziger isolierter Punkt von M .

Weiter ist \mathbb{R} separabel, weil die Menge \mathbb{Q} aller rationalen Zahlen abzählbar und dicht in \mathbb{R} ist. Oft ist es nützlich, die Abgeschlossenheit mit Hilfe konvergenter Folgen zu charakterisieren.

Lemma 3.4. *Eine Menge $M \subset X$ ist abgeschlossen genau dann, wenn*

(*) *der Limes jeder in M enthaltenen konvergenten Folge $\{x_n\}$ ebenfalls zu M gehört.*

(äquivalent dazu: wenn jeder Häufungspunkt von M in M liegt)

Beweis. Bezeichne W die Komplementärmenge $W = X \setminus M$.

1. Sei M abgeschlossen, also W offen. Für $x \in W$ betrachten wir eine in W enthaltene Kugel $B(x, \varepsilon)$. Sind alle x_n aus M (also insbesondere nicht in B) kann offenbar x nicht Grenzwert der Folge sein. Damit gilt (*).
2. Sei M nicht abgeschlossen, also W nicht offen. Dann gibt es ein $x \in W \setminus \text{int}W$. Die Kugeln $B(x, \frac{1}{n})$ schneiden daher stets die Menge M . Wählt man jeweils x_n aus dem Durchschnitt $B(x, \frac{1}{n}) \cap M$, erhält man eine Folge in M , die gegen x konvergiert. Weil x nicht in M liegt, ist dann (*) falsch.

□

Definition 3.5 (Kompaktheit). *Eine Teilmenge $M \subset X$ heißt **kompakt**, wenn jede (unendliche) Folge $\{x_n\}$ in M eine (unendliche) konvergente Teilfolge $\{x_{n(k)}\}$ besitzt, deren Limes x^* ebenfalls zu M gehört.*

Betrachtet man insbesondere konvergente Folgen $\{x_n\}$ in M , sieht man nach Lemma 3.4 mit (*): *Kompakte Mengen sind abgeschlossen.* (Denn unendliche Teilfolgen konvergenter Folgen haben denselben Limes wie die Ausgangsfolge).

In $X = \mathbb{R}$ ist die Menge X offen und abgeschlossen, aber nicht kompakt. Z.B. hat die Folge $x_n = n$ keine konvergente Teilfolge. Dagegen sind hier alle „abgeschlossenen“, beschränkten Intervalle $[a, b]$ kompakt (und unter allen Intervallen nur diese). Alle „offenen“ Intervalle (a, b) sind offen im Sinne obiger Definition. Das „halboffene“ Intervall $[a, b)$ ist dagegen (für $a < b$) weder offen noch abgeschlossen. Also sind Mengen nicht automatisch offen, wenn sie nicht abgeschlossen sind!

Sind A, B zwei kompakte Teilmengen eines metrischen Raumes, so auch ihre Vereinigung und ihr Durchschnitt. Abgeschlossene Teilmengen kompakter Mengen sind ebenfalls kompakt (Übungsaufgabe).

Definition 3.6 (Stetigkeit). *Es seien $(X, d_X), (Y, d_Y)$ metrische Räume, $f : X \rightarrow Y$ und $x^* \in X$. f heißt **stetig** in x^* , falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ derart existiert, dass*

(*) $f(x) \in B_Y(f(x^*), \varepsilon) \quad \forall x \in B_X(x^*, \delta)$.

Oben soll „ B_X “ darauf hinweisen, dass die Kugel in X liegt und mit der entsprechenden Metrik d_X definiert ist. Man beachte, dass Stetigkeit und Konvergenz immer von den gewählten Abständen abhängen.

Wenn $X = Y = \mathbb{R}$, so arbeiten wir mit dem „normalen“ Abstand $d(x', x) = |x' - x|$. Dann bedeutet die Bedingung (*) offenbar

$$|f(x) - f(x^*)| < \varepsilon \quad \text{für alle } x \text{ mit } |x - x^*| < \delta.$$

Interpretation:

Der Funktionswert $f(x)$ unterscheidet sich beliebig wenig von $f(x^*)$, wenn sich nur x hinreichend wenig von x^* unterscheidet; der „Unterschied“ wird mit den Abständen gemessen.

Die Bedingung (*) kann als Fehlerabschätzung interpretiert werden: Wird x^* mit Genauigkeit δ bestimmt, d.h., ist $d_X(x, x^*) < \delta$, so ist der Fehler bei der Funktionswertberechnung $d_Y(f(x), f(x^*))$ kleiner als ε .

Man überlege sich:

$f : X \rightarrow Y$ ist in $x^* \in X$ genau dann stetig, wenn aus der Konvergenz $x_n \rightarrow x^*$ im Raume X stets die Konvergenz der Funktionswerte $f(x_n) \rightarrow f(x^*)$ im Bildraum Y folgt (in der Vorlesung bewiesen).

Definition 3.7. *Ist $M \subset X$ und f stetig in allen $x^* \in M$, so heißt f stetig auf M .*

Man beachte, dass ein jeweils zu ε gehörendes δ auch von $x^* \in M$ abhängen kann: $\delta = \delta(\varepsilon, x^*)$.

Definition 3.8 (Gleichmäßig stetig). *Lässt sich $\delta = \delta(\varepsilon)$ unabhängig von $x^* \in M$ festlegen, so heißt f gleichmäßig (glm.) stetig auf M .*

Mit anderen Worten: f heißt **gleichmäßig stetig auf M** , falls zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, so dass (*) für alle $x^* \in M$ gilt (mit demselben δ).

Die reelle Funktion $f(x) = 2x$ ist gleichmäßig stetig, weil man stets $\delta = \frac{1}{2}\varepsilon$ setzen kann.

Dagegen ist $f(x) = x^2$ nicht gleichmäßig stetig auf $M = \mathbb{R}$ (große x^* verlangen „besonders kleine“ δ).

Ohne Stetigkeit verstanden zu haben besteht man nirgends eine Analysis Prüfung.

3.1.2 Stetigkeit für verknüpfte Funktionen

Es möge nun ein weiterer metrischer Raum (Z, d_Z) gegeben sein und f mittels zweier Funktionen $g : X \rightarrow Z$ und $h : Z \rightarrow Y$ durch $f(x) = h(g(x))$ erklärt sein. Dann gilt

Lemma 3.9. *f ist stetig in x^* , falls gilt: g ist stetig in x^* und h ist stetig in $g(x^*)$.*

Beweis. Wir benutzen Folgen. Gelte $x^* = \lim x_n$.

Für die Bilder $z_n = g(x_n)$ gilt dann $g(x^*) = \lim z_n$, da g in x^* stetig ist. Analog ist wegen der Stetigkeit von h in $g(x^*)$: $h(g(x^*)) = \lim h(z_n)$. Das bedeutet nach Definition von f gerade $f(x^*) = \lim f(x_n)$, was zu zeigen war. \square

Beispiele 3.10.

1. Siehe vor allem die Übungen.
2. Ist x_0 ein fixiertes Element in X , so bildet die durch $f(x) = d(x, x_0)$ definierte Abstandsfunktion X in \mathbb{R} ab. Sie ist stetig. Wegen der Dreiecksungleichung gilt nämlich mit beliebigen x, x^* :

$$f(x) - f(x^*) = d(x, x_0) - d(x^*, x_0) \leq (d(x, x^*) + d(x^*, x_0)) - d(x^*, x_0) = d(x, x^*).$$

Analog folgt durch Vertauschen von x und x^* : $f(x^*) - f(x) \leq d(x, x^*)$.

Also gilt stets $|f(x) - f(x^*)| \leq d(x, x^*)$. Die Stetigkeitsbedingung (*) ist also mit $\delta = \varepsilon$ erfüllt.

3. Als Funktion von \mathbb{R} in sich sind $f(x) = e^x$ und $s(x) = x^2$ auf \mathbb{R} stetig, aber nicht gleichmäßig stetig. Man berechne δ für $\varepsilon = 1$ und verschiedene x^* !

Dagegen ist e^x gleichmäßig stetig auf der Menge $M = \{x \in \mathbb{R} \mid x \leq 0\}$.

Man weise dies z.B. mittels $\delta = \frac{\varepsilon}{3}$ (für $\varepsilon \leq 1$) nach.

4. Sei jetzt $f(x) = 0$ für $x \leq 0$, $f(x) = 1$ für $x > 0$.
 - (a) Offenbar ist f in $x = 0$ nicht stetig, also auch nicht stetig auf $M = \{x \in \mathbb{R} \mid x \leq 0\}$.
 - (b) Betrachtet man jedoch f als nur auf M definiert, also als Funktion von M in \mathbb{R} , dann wird f auf M sogar gleichmäßig stetig (da Punkte $x \in \mathbb{R} \setminus M$ in (*) nicht mehr berücksichtigt werden). Dem zweiten Fall entspricht das Vorgehen, M selbst als metrischen Raum (mit dem Abstand d) aufzufassen und f als Funktion von M in \mathbb{R} anzusehen (die auf M eingeschränkte Funktion f , kurz f_M oder $f|_M$). In dieser Form wird der folgende wichtige Satz zumeist angewandt.

Satz 3.11 (-ein- Satz von Weierstraß/ Existenz des Maximums). *Es seien X ein kompakter metrischer Raum und $f : X \rightarrow \mathbb{R}$ stetig. Dann gibt es ein $x^* \in X$, sodass $f(x) \leq f(x^*) \forall x \in X$. (Existenz des Maximums von f über X).*

Beweis. Es sei $s = \sup\{f(x) \mid x \in X\}$.

Fall 1: Ist s unbeschränkt, so gibt es eine Folge $\{x_n\}$ in X mit $f(x_n) > n$ für $n = 1, 2, \dots$

Fall 2: Andernfalls lässt sich eine Folge $\{x_n\}$ so wählen, dass $f(x_n) > s - \frac{1}{n}$ gilt.

Da X kompakt ist, besitzt die Folge in jedem Fall eine konvergente Teilfolge $\{x_{n(k)}\}$; sei x^* das entsprechende Grenzelement. Da f in x^* stetig ist, konvergieren die Werte $f(x_{n(k)})$ gegen $f(x^*)$. Damit kann Fall 1 nicht zutreffen, da auch $n(k)$ unbeschränkt wächst. Im Fall 2 sichern nun $f(x_{n(k)}) > s - \frac{1}{n(k)}$ und die Stetigkeit die Ungleichung $f(x^*) \geq s$, was direkt die Behauptung ergibt. \square

Analog beweist man die Existenz eines Minimalpunktes x^* ; $f(x^*) \leq f(x) \forall x \in X$.

Kompakte Mengen spielen auch in Zusammenhang mit **gleichmäßiger Stetigkeit** eine wichtige Rolle.

Satz 3.12. [gleichmäßige Stetigkeit] *Jede auf einer kompakten Teilmenge $M \subset X$ stetige Funktion ist dort gleichmäßig stetig.*

Beweis. Indirekt: Angenommen f ist auf M stetig, aber nicht gleichmäßig stetig. Dann gibt es ein festes $\varepsilon > 0$, sodass zu jedem $\delta > 0$ wenigstens ein $x^* \in M$ existiert, mit dem (*) falsch ist; also ein weiteres $x \in B_X(x^*, \delta)$ existiert mit

$$(**) \quad d_Y(f(x), f(x^*)) \geq \varepsilon.$$

Dieses x kann zu $X \setminus M$ gehören, muss also nicht selbst in M liegen.

Wir lassen δ die Folge $\delta_n = \frac{1}{n}$ durchlaufen und nennen die zugeordneten Elemente nun x_n^* und x_n . Die Folge der $x_n^* \in M$ besitzt (da M kompakt) eine konvergente Teilfolge, die gegen ein $x^{**} \in M$ konvergiert. Um nicht doppelt indizieren zu müssen, nehmen wir ohne Beschränkung der Allgemeinheit an, dass schon $x^{**} = \lim x_n^*$ gilt. Wegen $d_X(x_n, x_n^*) < \delta_n$ muss dann aber auch $x^{**} = \lim x_n$ gelten.

Da f in x^{**} stetig ist, folgt somit $\lim f(x_n) = f(x^{**}) = \lim f(x_n^*) < \varepsilon$ für große n bewirkt. Dies widerspricht aber unserer Konstruktion nach (**). Also muss der Satz richtig sein. \square

3.2 Überdeckungssatz von Heine/Borel/Lebesgue

Definition 3.13 (Überdeckung). *Es seien M eine Teilmenge eines metrischen Raumes (X, d) , I eine beliebige (nichtleere) Menge, und M_i (für alle $i \in I$) seien Teilmengen von X , deren Vereinigung $\bigcup M_i$ die Menge M enthält.*

Man sagt dann, dass die Mengenfamilie (formal eine Menge von Mengen) $\{M_i \mid i \in I\}$ eine Überdeckung von M bildet. Sind alle M_i offen, nennt man die Überdeckung offen.

Ein für viele Existenzbeweise und Konstruktionen in metrischen Räumen zentraler Satz ist der folgende

Satz 3.14 (Überdeckungssatz von Heine/Borel/Lebesgue). *Für jede offenen Überdeckung $\{M_i \mid i \in I\}$ einer kompakten Teilmenge M von X gibt es eine endliche Teilmenge E von I , sodass auch schon $\{M_i \mid i \in E\}$ eine Überdeckung von M ist.*

Beweis. Zunächst ein Beweis für den Spezialfall $X = \mathbb{R}$: Die kompakte Menge M muss in einem Intervall $[a, b]$ liegen. Angenommen der Satz sei falsch. Wir teilen das Intervall (Mittelpunkt c) und betrachten

$$A = M \cap [a, c], \quad B = M \cap [c, b].$$

Lassen sich beide Mengen von je endlich vielen M_i überdecken, so wird $M = A \cup B$ von diesen insgesamt endlich vielen Mengen M_i überdeckt. Also kann man eine der beiden Mengen nur mit unendlich vielen M_i überdecken, o.B.d.A. sei das A .

Mit dem Intervall $[a', b'] = [a, c]$, dessen Mittelpunkt c' und den neuen Mengen

$$A' = M \cap [a', c'], \quad B' = M \cap [c', b'].$$

wiederholen wir diese Überlegung. In dieser Weise erhalten wir eine Intervallschachtelung $[\alpha_k, \beta_k]$ mit der jeweils gilt:

$M \cap [\alpha_k, \beta_k]$ lässt sich nur durch unendlich viele der Mengen M_i überdecken.

Der durch die Intervallschachtelung definierte Punkt x erfüllt nun

$$x \in [\alpha_k, \beta_k] \quad \forall k = 1, 2, 3, \dots$$

und $x \in M$ (da x Limes von Punkten $x_k \in M \cap [\alpha_k, \beta_k]$ ist).

Damit gibt es aber eine Menge $M_{i(x)}$ mit $x \in M_{i(x)}$. Da sie offen ist und $\lim \alpha_k = x = \lim \beta_k$ gilt, muss nun auch $[\alpha_k, \beta_k] \in M_{i(x)}$ für große $k > k_0$ gelten. Dann kann man $M \cap [\alpha_k, \beta_k]$ aber bereits mit der einen Menge $M_{i(x)}$ überdecken, im Widerspruch zur Konstruktion der Intervalle. Dieser Widerspruch beweist den Satz im reellen Fall. \square

Beweis. (Der allgemeine Fall)

Nun müssen die Intervalle durch geeignete andere Mengen ersetzt werden. Das werden die Kugeln eines δ -Netzes sein. Der Beweis wird uns noch zwei interessante „Nebenresultate“ liefern.

1. Wir zeigen im ersten und entscheidenden Teil, dass M schon in der Vereinigung abzählbar vieler M_i liegt.

Wir betrachten dazu alle Kugeln $B(x, \delta)$ mit $x \in M$ und beliebig fixiertem Radius $\delta > 0$.

- 1.1. Zuerst überlegen wir uns, dass M in der Vereinigung endlich vieler solcher Kugeln liegen muss.

Sei $x_1 \in M$ beliebig gewählt. Beginnend mit $n = 1$ bilden wir die Vereinigung

$$W(n) = \bigcup_{1 \leq k \leq n} B(x_k, \delta).$$

Ist die Restmenge $M \setminus W(n)$ leer, sind wir fertig, denn M liegt in $W(n)$.

Andernfalls wählen wir ein $x_{n+1} \in M \setminus W(n)$. Es erfüllt offenbar

$$\delta \leq d(x_{n+1}, x_k) \quad \forall k = 1, 2, \dots, n.$$

Damit kann die entstehende Folge $\{x_n\}$ keine konvergente Teilfolge besitzen. Weil das mit $n \rightarrow \infty$ der Kompaktheit von M widerspricht, muss obige Konstruktion für ein gewisses n abbrechen.

Bemerkung: Ist $M \subset W(n)$, so sagt man auch, dass die Menge $\{x_1, \dots, x_n\}$ ein endliches δ -Netz für M bildet.

Wir haben also unter 1.1 gezeigt, dass für eine kompakte Menge M stets ein endliches δ -Netz existiert, wie klein $\delta > 0$ auch gewählt wird.

- 1.2. Wir lassen nun δ die Folge $\delta_m = \frac{1}{m}$ ($m = 1, 2, \dots$) durchlaufen und vereinigen alle oben konstruierten endlichen δ_m -Netze, bestehend etwa aus $x_1^m, \dots, x_{n(m)}^m$. Die entstehende Menge aller x_k^m ist abzählbar, (d.h. hat höchstens die Mächtigkeit von \mathbb{N}). Sie heie A , ihre Elemente seien umbenannt in a_1, \dots, a_k, \dots . Ist $x \in M$ und m beliebig, so findet sich stets ein $a_k \in A$ mit $d(x, a_k) < \frac{1}{m}$. Man nehme ein entsprechendes Element des δ_m -Netzes. (Eine kompakte Menge M ist also separabel, und A ist eine dichte Teilmenge von M).
- 1.3. An dieser Stelle betrachten wir erstmals die Mengen M_i . Sei $x \in M$ irgendwie fixiert. Nach Voraussetzung gibt es ein $i(x) \in I$, sodass $x \in M_{i(x)}$. Diese Menge ist offen. Also gibt es ein $\delta(x) > 0$, sodass $B(x, \delta(x))$ ganz in $M_{i(x)}$ liegt. Sei nun m so gro, dass $\frac{1}{m} < \frac{\delta(x)}{2}$. Dann gibt es, wie gerade festgestellt, ein k mit $d(x, a_k) < \frac{1}{m}$ bzw. $x \in B(a_k, \frac{1}{m})$. Fr jedes Element b dieser Kugel gilt

$$d(b, x) \leq d(b, a_k) + d(a_k, x) < \frac{2}{m} < \delta(x); \quad \text{also } b \in B(x, \delta(x)).$$

Damit ist $B(a_k, \frac{1}{m})$ enthalten in $B(x, \delta(x))$ und erst recht in $M_{i(x)}$. Wir haben so gezeigt: Es gibt zu jedem $x \in M$ eine Kugel $B(a_k, \frac{1}{m})$, die x enthlt und ganz in einer der Mengen M_i liegt.

Sei schlielich $J(k, m) = \{i \in I \mid B(a_k, \frac{1}{m}) \subset M_i\}$, und sei Q die (abzhlbare) Menge aller (k, m) mit $J(k, m) \neq \emptyset$. Ordnen wir allen $(k, m) \in Q$ genau ein $j(k, m) \in J(k, m)$ zu, so liegt jedes $x \in M$ in einer der Mengen $M_{j(k, m)}$. Das sind nur abzhlbar viele. Die Bilder der Abbildung j definieren so eine abzhlbare Teilmenge I' , dass die Vereinigung $\bigcup M_i$ mit $i \in I'$ die Menge M ebenfalls berdeckt.

2. Wir haben nur noch zu zeigen, wie man von einer abzhlbaren offenen berdeckung von M zu einer endlichen gelangt. Seien dazu die Elemente von I' durchnummeriert; i_1, i_2, \dots und sei zur Abkrzung $G_k = M_{i_k}$, sowie $V_n = G_1 \cup G_2 \cup \dots \cup G_n$.

Wir zeigen, dass M in V_n enthalten ist, wenn nur n hinreichend gro ist.

Andernfalls whle man (wie blich!) $x_n \in M \setminus V_n$. Wieder existiert, da M kompakt ist, eine konvergente Teilfolge $x_{n(k)}$ mit Grenzelement $x^* \in M$.

Nach Konstruktion von I' gehrt x^* zu wenigstens einer der Mengen G_n , etwa zu G_m .

Diese Menge ist nach Voraussetzung offen; wir whlen ein $\delta > 0$, sodass $B(x^*, \delta)$ in G_m enthalten ist. Dann liegt $B(x^*, \delta)$ aber auch vollstndig in allen (greren) Mengen V_n mit $n \geq m$.

Auerdem gilt $x_{n(k)} \in B(x^*, \delta)$ fr fast alle k wegen $x^* = \lim x_{n(k)}$.

Beides zusammen zeigt, dass $x_{n(k)} \in V_{n(k)}$ fr alle hinreichend groen k richtig ist.

Dies widerspricht allerdings der Wahl von $x_n \in M \setminus V_n$. Der berdeckungssatz ist damit bewiesen. □

Der Beweis zeigte zugleich:

Kompakte Mengen sind separabel und besitzen mit beliebigen $\varepsilon > 0$ ein endliches ε -Netz.

Auerdem ist er ein typisches Beispiel fr einen „nichtkonstruktiven Beweis“: Wie man die Menge E letztlich findet, ist algorithmisch nicht nachvollziehbar. Satz 3.14 ist ferner umkehrbar, d.h. es gilt

Satz 3.15. *Wenn sich aus jeder offenen berdeckung einer Teilmenge M von X endlich viele Mengen auswhlen lassen, die M ebenfalls berdecken, so ist M kompakt.*

Beweis. Angenommen M ist nicht kompakt. Gibt es eine Folge $\{x_n\}$ in M , die gegen $x^* \in X \setminus M$ konvergiert, knnte man $M_n = X \setminus \{x_n\}$ setzen, um im Widerspruch zur Voraussetzung eine offene berdeckung von M zu finden, die keine endliche Teilberdeckung zulsst.

Andernfalls gibt es, da M nicht kompakt ist, eine Folge $\{x_n\}$ in M , die berhaupt keine konvergente Teilfolge besitzt. Ist δ_n hinreichend klein, knnen dann in der Kugel $B(x_n, \delta_n)$ hchstens endlich viele der Folgeelemente liegen.

Bildet man M_n aus dieser Kugel und fgt zu den so erhaltenen Mengen noch die Menge W hinzu, die aus X durch Entfernen aller x_n entsteht, findet man wiederum eine offene berdeckung, die keine endliche Teilberdeckung zulsst. □

3.2.1 Beispiel einer Anwendung

Gleichmäßige Stetigkeit mittels Überdeckungssatz

Wir wissen schon: von Satz 3.12:

Eine auf einer kompakten Menge $M \subset X$ definierte stetige Funktion $f : M \rightarrow Y$ ist dort gleichmäßig stetig.

Wir beweisen dies hier mit dem Überdeckungssatz. Zu gegebenem $\varepsilon > 0$ existiert wegen der Stetigkeit zu jedem $x \in M$ ein $\delta(x)$, sodass für alle Punkte der (offenen) Kugel $B(x, \delta(x)) = \{x' \mid d_X(x', x) < \delta(x)\}$ gilt:

$$d_Y(f(x'), f(x)) < \varepsilon.$$

Die Kugeln $B(x, \frac{1}{2}\delta(x))$, $x \in M$ überdecken offenbar M . Also lässt sich M auch durch endlich viele solcher Kugeln $B_1 = B(x_1, \frac{1}{2}\delta(x_1)), \dots, B_n = B(x_n, \frac{1}{2}\delta(x_n))$ überdecken.

Sei nun $x^* \in M$ beliebig gewählt und $0 < \delta < \min_{1 \leq k \leq n} \{\frac{1}{2}\delta(x_k)\}$.

Für jedes $x \in B(x^*, \delta)$ folgt dann: Es gibt ein k mit $x^* \in B_k$. Der Abstand $d(x, x_k)$ erfüllt

$$d(x, x_k) < d(x, x^*) + d(x^*, x_k) < \delta + \frac{1}{2}\delta(x_k) < \delta(x_k).$$

Damit folgt aber nach Wahl der Kugeln

$$d_Y(f(x), f(x_k)) < \varepsilon \quad \text{und} \quad d_Y(f(x^*), f(x_k)) < \varepsilon,$$

was mit der Dreiecksungleichung

$$d_Y(f(x), f(x^*)) \leq d_Y(f(x), f(x_k)) + d_Y(f(x_k), f(x^*)) < 2\varepsilon$$

liefert. Das heißt, zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$, sodass aus $x^* \in M$ und $d(x, x^*) < \delta$ stets folgt:

$$d_Y(f(x), f(x^*)) < 2\varepsilon.$$

Dies bedeutet gleichmäßige Stetigkeit von f auf M .

Maximum

Um etwa das Maximum $f(x^*)$ einer stetigen Funktion f von X in \mathbb{R} auf einem kompakten metrischen Raum zu bestimmen (vgl. Satz 3.12), wähle man ein endliches $\frac{1}{m}$ -Netz $A(m) = \{a_1, \dots, a_{n(m)}\}$. (Ist m groß, werden das recht viele Punkte). Durch Berechnen aller zugehörigen Funktionswerte findet man einen Punkt $a^*(m) \in A(m)$ mit $f(a^*(m)) \geq f(a)$ für alle $a \in A(m)$.

Um zu sichern, dass der Fehler des berechneten Maximalwertes kleiner als ein vorgegebenes α wird, d.h. dass

$$(***) \quad f(x^*) - f(a^*(m)) < \alpha$$

gilt, muss „nur“ m hinreichend groß sein.

Da nämlich f auf X gleichmäßig stetig ist, findet sich ein $\delta(\alpha)$ entsprechend der Definition der gleichmäßigen Stetigkeit, wonach für beliebige x_0 und x_1 mit $d(x_0, x_1) < \delta(\alpha)$ stets $|f(x_0) - f(x_1)| < \alpha$ gilt.

Ist $\frac{1}{m} < \delta(\alpha)$, so gibt es ein $a \in A(m)$ mit $d(x^*, a) < \delta(\alpha)$.

Setzt man $x_0 = x^*$ und $x_1 = a$, folgt also $f(x^*) - f(a) < \alpha$, und wegen $f(a^*(m)) \geq f(a)$ erst recht (***) .

Wächst m unbeschränkt, erhält man nun eine Folge von Punkten $a^*(m)$. Diese besitzt eine unendliche, konvergente Teilfolge $\{a^*(m(k))\}$ mit einem gewissen Häufungspunkt a^* .

Mit jedem $\alpha > 0$ gilt (***) dann für fast alle $m(k)$ (nämlich wenn $\frac{1}{m(k)} < \delta(\alpha)$). Folglich muss auch $f(x^*) - f(a^*) < \alpha$ richtig sein. Da dies für alle positiven α gilt, kann nur $f(a^*) = f(x^*)$ sein, d.h. wir haben in a^* sogar einen Punkt aus X mit maximalem Funktionswert gefunden, nicht nur diesen Wert selbst.

Diese Diskussion liefert ein theoretisch denkbares, aber praktisch oft problematisches Verfahren zur Lösung der Aufgabe, das Maximum von f auf X zu bestimmen. Schwierigkeiten resultieren aus der oft hohen Anzahl der benötigten Punkte in A , der Frage ihrer günstigsten Anordnung und aus erforderlichen Abschätzungen bezüglich der Größe $\delta(\alpha)$, die ja m bestimmt. Bei „kleinen Aufgaben“ kann man jedoch häufig so vorgehen.

3.2.2 Vollständige metrische Räume

Zu den Existenzaussagen der Sätze 3.12 und 3.14 wollen wir eine weitere, wesentliche hinzufügen. Sie benutzt eine Abschätzung fuer $\sum_{k=1}^n q^k$ (siehe 1. Übung) zusammen mit der (für den reellen Fall bewiesenen) Eigenschaft, dass Cauchy-Folgen konvergent sind.

Letzteres ist allerdings in metrischen Räumen i.A. nicht erfüllt, was die folgenden Definitionen motiviert.

Definition 3.16 (Cauchy-Folge). Sei (X, d) ein metrischer Raum. Eine Folge $\{x_n\}$ in X heißt Cauchy-Folge (auch Fundamentalfolge), wenn zu jedem $\varepsilon > 0$ ein $n(\varepsilon)$ derart existiert, dass $d(x_n, x_m) < \varepsilon$ für alle $n > n(\varepsilon)$ und $m > n(\varepsilon)$ gilt.

Definition 3.17. Ein metrischer Raum (X, d) heißt vollständig, wenn jede seiner Cauchy-Folgen in X konvergiert.

Damit ist \mathbb{R} (wieder mit $d(x, y) = |x - y|$) ein vollständiger metrischer Raum. Die Menge \mathbb{Q} der rationalen Zahlen, versehen mit demselben Abstand, ist dagegen nicht vollständig. Zum Beispiel bildet eine in \mathbb{R} gegen die Zahl $\sqrt{2}$ konvergierende Folge rationaler Zahlen als konvergente Folge eine Cauchy-Folge. Diese Folge ist auch in \mathbb{Q} eine Cauchy-Folge, aber keine rationale Zahl ist Limes dieser Folge. Aus dem gleichen Grunde wäre $X = \mathbb{R} \setminus \{0\}$ nicht vollständig.

Die Menge \mathbb{N} der natürlichen Zahlen (mit d) ist dagegen wieder vollständig, da hier nur solche Folgen die Cauchy-Eigenschaft besitzen, die von einem gewissen Index an konstant sind (man nehme z.B. $\varepsilon = \frac{1}{2}$). Um zu sehen, dass es auch weniger triviale vollständige Räume und kompakte Mengen gibt als die bisher notierten, betrachten wir den folgenden Raum.

Vollständigkeit und Kompaktheit im Raum \mathbb{R}^p

In der Menge aller „ p -tupel“ reeller Zahlen

$$X := \mathbb{R}^p := \{x \mid x = (x_1, x_2, \dots, x_p), x_k \in \mathbb{R} \quad \forall k = 1, 2, \dots, p\}$$

lässt sich (z.B.) durch

$$d(x, y) := d_{\max}(x, y) := \max\{|x_1 - y_1|, |x_2 - y_2|, \dots, |x_p - y_p|\}$$

ein Abstand definieren. (Abstandseigenschaften nachrechnen!)

Um die Elemente einer Folge in X zu kennzeichnen, schreiben wir jetzt x^n (Index oben), um sie nicht mit den Komponenten x_k von x zu verwechseln (Potenzen brauchen wir hier nicht). Es ist also $x^n = (x_1^n, \dots, x_p^n)$. Die Bedingung $d(x^n, x^m) < \varepsilon$ bedeutet

$$|x_k^n - x_k^m| < \varepsilon \quad \forall k = 1, 2, \dots, p.$$

Konvergenz einer Folge $\{x^n\}$ in (X, d) heißt daher, dass für jedes feste k die Folge der reellen Zahlen $\{x_k^n\}$ (für $n \rightarrow \infty$) einen Grenzwert, etwa x_k^* , besitzt.

Der Limes der Folge $\{x^n\}$ in X ist dann $x^* = (x_1^*, \dots, x_p^*)$.

Analog ist $\{x^n\}$ eine Cauchy-Folge in (X, d) , genau dann, wenn alle reellen Folgen $\{x_k^n\}$ (k fest, $n \rightarrow \infty$) Cauchy-Folgen in \mathbb{R} sind.

In diesem Falle ist aber jede einzelne konvergent (Jede Cauchy-Folge in \mathbb{R} ist konvergent!). Also existiert dann $x^* = \lim x^n$ in X .

Der Raum \mathbb{R}^p , versehen mit dem obigen Maximum-Abstand d_{\max} , ist demnach vollständig.

Kompaktheit in \mathbb{R}^p

Eine Teilmenge M von \mathbb{R}^p heie beschrnkt, wenn es eine Zahl C gibt, sodass $d(x, 0) < C$ für alle $x \in M$ richtig ist, wobei 0 das aus p (reellen) Nullen bestehende p -tupel bezeichnet.

Das heißt, es gilt mit jedem festen k : $|x_k^n| < C$ für alle $x^n \in M$.

Sei $\{x^n\}$ eine unendliche Folge in M .

Da die Folge $\{x_1^n\}$ reeller Zahlen beschrnkt ist, besitzt sie eine (in \mathbb{R}) konvergente Teilfolge.

Deren Indizes mögen die Teilmenge N_1 von \mathbb{N} bilden.

Die Folge der Zahlen x_2^n mit $n \in N_1$ ist gleichfalls beschrnkt. Aus ihr lässt sich also erneut eine konvergente Teilfolge aussondern, die durch eine unendliche Teilmenge N_2 von N_1 bestimmt ist.

Nach p Schritten erhalten wir somit eine Menge N_p ($\subset N_{p-1} \subset \dots \subset N_1$), die eine unendliche Teilfolge der Ausgangsfolge charakterisiert, sodass jede der Komponentenfolgen $\{x_k^n\}$ ($n \in N_p$) konvergiert.

Damit konvergiert die durch N_p gegebene Teilfolge der Ausgangsfolge $\{x^n\}$ in X :
Es existiert $x^* = \lim x^n$ ($n \in N_p$).

Ist M abgeschlossen, so gilt auch noch $x^* \in M$; die Menge M ist dann also kompakt. Diese Überlegungen liefern:

Lemma 3.18. *In $X = \mathbb{R}^p$ (versehen mit obigem Abstand d_{\max}) ist jede beschränkte und abgeschlossene Menge kompakt; der Raum selbst ist vollständig und separabel.*

Beweis. Um die Separabilität zu sehen, bilde man A aus allen x mit sämtlich rationalen Koordinaten x_1, \dots, x_p . Der Rest wurde gerade hergeleitet. \square

Wir wissen schon dass alle kompakten Mengen abgeschlossen sein müssen. Ist andererseits M nicht beschränkt, so wähle man eine Folge mit unbeschränkt wachsendem Abstand $d(x^n, 0)$ und zeige, dass sie keine konvergente Teilfolge besitzen kann (Übungsaufgabe).

Damit sieht man, dass *nur* die beschränkten und abgeschlossenen Teilmengen kompakt in \mathbb{R}^p sind.

3.3 Banachscher Fixpunktsatz

Satz 3.19. [*Banachscher Fixpunktsatz*] *Es sei f eine Funktion, die einen (nichtleeren!) vollständigen metrischen Raum (X, d) kontraktiv in sich abbildet, d.h., $f : X \rightarrow X$ erfüllt mit einem gewissen festen $q \in (0, 1)$:*

$$(*) \quad d(f(x), f(y)) \leq q \cdot d(x, y) \quad \forall x, y \in X.$$

Dann gibt es einen und nur einen Punkt x^* in X mit $f(x^*) = x^*$.

Bemerkung: x^* heißt Fixpunkt der Funktion f , q heißt Kontraktionskonstante.

Beweis. Existenz: Beginnend mit einem beliebig fixierten $x_0 \in X$ und $n = 0$ bilde man nach der Vorschrift $x_{n+1} = f(x_n)$ eine Folge $\{x_n\}$ in X . Wegen (*) gilt dann

$$\begin{aligned} d(x_2, x_1) &= d(f(x_1), f(x_0)) \leq q \cdot d(x_1, x_0), \\ d(x_3, x_2) &= d(f(x_2), f(x_1)) \leq q \cdot d(x_2, x_1) \leq q^2 \cdot d(x_1, x_0), \end{aligned}$$

und allgemein über vollständige Induktion für $n > 0$:

$$d(x_{n+1}, x_n) = d(f(x_n), f(x_{n-1})) \leq q \cdot d(x_n, x_{n-1}) \leq q \cdot q^{n-1} \cdot d(x_1, x_0) = q^n \cdot d(x_1, x_0).$$

Sind n und m größer als ein gewisser Index p und gilt $m = n + t$, ($t > 0$ ganzzahlig), so ergibt sich mittels mehrmaliger Anwendung der Dreiecksungleichung

$$d(x_n, x_m) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \dots + d(x_{n+t-1}, x_{n+t}) \leq (q^n + q^{n+1} + \dots + q^{n+t-1}) d(x_1, x_0).$$

Wegen $n > p$ ist die Summe in der Klammer kleiner als $q^p (q + q^2 + q^3 + \dots) = q^p \frac{q}{1-q}$.

Ist p groß genug ($> p(\varepsilon)$), wird daher für $n, m > p$ der Abstand $d(x_n, x_m)$ kleiner als jedes vorgegebene $\varepsilon > 0$. Die konstruierte Folge ist also eine Cauchy-Folge; somit existiert wegen der Vollständigkeit von X der Limes $x^* = \lim x_n \in X$. Offenbar gilt damit auch $x^* = \lim x_{n+1} = \lim f(x_n)$.

Da f nach (*) insbesondere stetig ist, muss also $f(x^*) = x^*$ sein.

Eindeutigkeit: Gäbe es einen weiteren Fixpunkt x^{**} , so folgte wieder über (*):

$$d(x^{**}, x^*) = d(f(x^{**}), f(x^*)) \leq q \cdot d(x^{**}, x^*),$$

was wegen $q < 1$ nur mit $d(x^{**}, x^*) = 0$, also $x^{**} = x^*$ möglich ist. Damit ist der Satz richtig. \square

Man beachte, dass dieser Standardbeweis konstruktiv ist und zu einer simplen Methode für die näherungsweise Berechnung des Fixpunktes führt, man bilde einfach

$$x_{n+1} = f(x_n).$$

Diese Methode heißt auch *Methode der sukzessiven Approximation*.

Der Fehler nach n Schritten $d(x_n, x^*)$ lässt sich mittels q^n abschätzen. Tun Sie das!

Beispiel 3.20. Wir betrachten die Iterationssformel $x_{n+1} = \frac{1}{2}(x_n + \frac{a}{x_n})$ zur Berechnung von $s = \sqrt{a}$ für $a > 0$.

Um zu zeigen, dass mit jedem Anfangspunkt $x_0 > 0$ die entstehende Folge gegen s konvergiert, „erfinden“ wir den metrischen Raum $X = \{x \in \mathbb{R} \mid x \geq s\}$ und die Funktion $f(x) = \frac{1}{2}(x + \frac{a}{x})$.

Die Gleichung $f(x) = x$ bedeutet dann tatsächlich $x^2 = a$. Was ist zu zeigen?

1. X ist ein vollständiger metrischer Raum (Ist mit Betragsabstand klar).
2. f bildet X in sich ab, also $f(x) \geq s$ wenn $x \geq s$. Dazu überlegt man sich (für $x > 0$)

$$s \leq f(x) \quad \Leftrightarrow \quad s \leq \frac{1}{2} \left(x + \frac{a}{x} \right) \quad \Leftrightarrow \quad 0 \leq \frac{1}{2} (x^2 - 2sx + a) \quad \Leftrightarrow \quad 0 \leq (x - s)^2.$$

D.h., wenn nur $x_0 > 0$ ist, liegen $x_1 = f(x_0)$ und alle anderen $x_{k+1} = f(x_k)$ in X .

3. f ist kontraktiv auf X . Man rechnet aus

$$f(x) - f(y) = \frac{1}{2} \left(x - y + \frac{a}{x} - \frac{a}{y} \right) = \frac{1}{2} \left(x - y + a \frac{y - x}{xy} \right)$$

Wegen $\frac{a}{xy} \leq 1$ für $x, y \in X$, liefert dies $|f(x) - f(y)| \leq \frac{1}{2}|x - y|$.

Also konvergiert die mit $x_{n+1} = \frac{1}{2}(x_n + \frac{a}{x_n})$ erzeugte Folge gegen \sqrt{a} , falls $x_0 > 0$ (Nachprüfen!).

Kapitel 4

Reihen

Definition: Einer Folge $\{a_n\}$ (reeller Zahlen) lässt sich in natürlicher Weise einer Folge $\{s_n\}$ zuordnen, die Folge der Partialsummen $s_n = a_1 + a_2 + \dots + a_n$, kurz $s_n = \sum_{k=1}^n a_k$ für $1 \leq k \leq n$. Konvergiert die Folge $\{s_n\}$ gegen s , so sagt man, dass die (unendliche) Reihe

$$a_1 + a_2 + \dots$$

gegen s konvergiert und schreibt dafür $s = \sum_{k=1}^{\infty} a_k$. Andernfalls spricht man von Divergenz dieser Reihe.

Wir betrachten also spezielle Folgen, nämlich solche, die aus Partialsummen bestehen.

Ein wichtiges Beispiel ist die geometrische Reihe mit $a_n = q^n$, $n = 1, 2, 3, \dots$

Wegen $q^1 + \dots + q^n = \frac{q - q^{n+1}}{1 - q}$ (man multipliziert beide Seiten mit $1 - q$) ist offenbar die Folge $\{s_n\}$ für $0 < q < 1$ konvergent mit dem Limes $s = \frac{q}{1 - q}$.

Ist eine Reihe mit $a_n \geq 0 \forall n$ gegeben, so sind die Partialsummen monoton ($s_n \leq s_{n+1}$). Sie sind also nach Satz 2.7 konvergent, wenn sie (nach oben) beschränkt sind: $s_n \leq S \forall n$; S unabhängig von n . Also gilt

Satz 4.1. *Eine Reihe mit nichtnegativen Gliedern a_n konvergiert, wenn die Folge ihrer Partialsummen (nach oben) beschränkt ist.*

Es ist trivial, dass die Beschränktheit der Partialsummen stets notwendig für die Konvergenz einer Reihe ist. Ebenso leicht sieht man: Eine Reihe konvergiert nur dann, wenn $\lim a_n = 0$ (wieso?).

Die Bedingung $\lim a_n = 0$ ist allerdings nicht hinreichend für die Konvergenz der Reihe, wie das Beispiel der *harmonischen Reihe* $a_n = \frac{1}{n}$ zeigt: Die Reihe lässt sich schreiben als

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

Indem die folgenden Summanden jeweils zu Termen von 2^k Summanden zusammengefasst werden ($k = 3, 4, 5, \dots$) mit dem letzten Summanden $\frac{1}{2^{k+1}}$, sieht man, dass die Summe jeder Klammer mindestens $\frac{1}{2}$ wird. Damit sind die zugehörigen Partialsummen nicht beschränkt. Die harmonische Reihe divergiert.

Andererseits ist die Summe jeder Klammer nicht größer als 1. Um also eine vorgegebene Summe s zu erreichen, braucht man mindestens s Klammern und somit mindestens 2^{s-1} Summanden. Dies ist eine grobe Abschätzung, die allerdings zeigt, dass die Reihe „recht langsam“ divergiert.

Wichtigstes Mittel, um die Konvergenz einer Reihe $\sum_n b_n$ zu zeigen, ist der Vergleich mit einer schon bekannten konvergenten Reihe $\sum_n a_n$, konkret oft mit der geometrischen.

Lemma 4.2 (Vergleichskriterium). *Die Reihe $\sum_n a_n$ konvergiere, und es gelte $0 \leq b_n \leq a_n$ für hinreichend große n , etwa für $n > n_0$. Dann konvergiert auch die Reihe $\sum_n b_n$.*

Beweis. Die monotonen Partialsummen der Reihe $\sum_{n > n_0} b_n$ sind durch die Summe $s = \sum_{n > n_0} a_n$ beschränkt.

Satz 4.1 liefert so Konvergenz von $\sum_{n > n_0} b_n$, was offenbar Konvergenz der gesamten Reihe bedeutet. \square

Kann man ähnlich auf Divergenz schließen?
Um entscheiden zu können, ob Reihen konvergieren, brauchen wir anwendbare Kriterien für die Konvergenz einer Reihe.

4.1 Konvergenzbedingungen für Reihen

Wir betrachten zunächst eine beliebige Reihe $\sum_n a_n$. Sie sei konvergent mit der Summe s .

Durch willkürliches Zusammenfassen jeweils endlich vieler ihrer Element und Beibehalten der Reihenfolge bilden wir eine neue Reihe

$$(a_1 + \dots + a_{m_1}) + (a_{m_1+1} + \dots + a_{m_2}) + (a_{m_2+1} + \dots + a_{m_3}) + \dots$$

mit Elementen $b_1 + b_2 + b_3 + \dots$.

Lemma 4.3 (Assoziativität). *Die oben konstruierte Reihe $\sum b_n$ konvergiert ebenfalls gegen den nach Voraussetzung existierenden Grenzwert s der Originalreihe.*

Beweis. Die erste Reihe $\sum_n a_n$ hat nach Voraussetzung konvergente Partialsummen $s_n \rightarrow s$. Den Partialsummen t_n der zweiten Reihe entsprechen Partialsummen s_{m_n} der ersten, wobei m_n mit n unbeschränkt wächst. Die Teilfolge $\{s_{m_n}\}$ der Folge aller Partialsummen s_n konvergiert aber gegen denselben Wert wie die Ausgangsfolge, also gegen s . \square

Beispiel 4.4. Sei $b_n = \frac{1}{n(n+1)}$, $n \geq 1$. Um die Konvergenz von $\sum_n b_n$ zu zeigen, stellen wir fest, dass jeweils $b_n = \frac{1}{n} - \frac{1}{n+1}$ ist. Die Reihe ist damit eine Zusammenfassung der folgenden Reihe $\sum_n a_n$:

$$1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \frac{1}{4} + \dots$$

Deren Partialsummen sind abwechselnd 1 bzw. $1 - \frac{1}{n} = \frac{n-1}{n}$ und konvergieren offenbar gegen 1. Die Partialsummen t_n der Reihe $\sum b_n$ bilden die Teilfolge $s_1 = \frac{1}{2}$, $s_2 = \frac{2}{3}$, $s_3 = \frac{3}{4}$, \dots . Sie konvergiert (als Teilfolge einer konvergenten Folge) offenbar ebenfalls gegen 1. Also ist auch $\sum_n b_n = 1$.

Mit Hilfe dieser Reihe $\sum_n b_n$ lässt sich Konvergenz von $\sum_{k \geq 1} \frac{1}{k^2}$ zeigen: Wegen $\frac{1}{k^2} < \frac{1}{k(k-1)} = b_{k-1}$ für $k > 1$ und Konvergenz von $\sum_n b_n$ folgt die Konvergenz durch Einschachteln (Vergleichskriterium). Es gilt $\sum_{k \geq 1} \frac{1}{k^2} = \frac{\pi^2}{6}$ (ohne Beweis).

ABER: Lemma 4.3 ist nicht umkehrbar: Wenn $\sum_n b_n$ konvergiert, folgt nichts für die Ausgangsreihe $\sum_n a_n$. Man betrachte einfach $(1-1) + (1-1) + \dots$.

4.2 Quotientenkriterium, Wurzelkriterium

Nun sei $\sum_n a_n$ eine Reihe mit positiven Elementen a_n . Die beiden nächsten Konvergenzbedingungen basieren auf dem Vergleich mit der geometrischen Reihe entsprechend Lemma 4.2.

Satz 4.5 (Quotientenkriterium von d'Alembert). *Die Reihe $\sum_n a_n$ konvergiert, wenn es eine Zahl q mit $0 < q < 1$ derart gibt, dass $\frac{a_{n+1}}{a_n} \leq q$ für fast alle n gilt.*

Satz 4.6 (Wurzelkriterium von Cauchy). *Die Reihe $\sum_n a_n$ konvergiert, wenn es eine Zahl q mit $0 < q < 1$ derart gibt, dass $\sqrt[n]{a_n} = (a_n)^{\frac{1}{n}} \leq q$ für fast alle n gilt.*

Beweis. In beiden Fällen gilt für $n > n_0$, $a_{n+1} \leq qa_n$ bzw. $a_n \leq q^n$.

Die zweite Ungleichung zeigt direkt, dass die Partialsummen s_n der fraglichen Reihe durch $a_1 + \dots + a_n \leq s_{n_0} + \frac{q}{1-q}$ nach oben abgeschätzt werden können, also Konvergenz vorliegt. Die erste Ungleichung zeigt, dass für große $n_1 > n_0$ sowohl $a_{n_1} < q$ als auch $a_{n_1+k} \leq q^{k+1}$ gilt. Die monotonen Partialsummen s_n können daher durch $a_1 + \dots + a_n \leq s_{n_1} + \frac{q}{1-q}$ nach oben abgeschätzt werden. \square

Analoge Aussagen lassen sich über die Divergenz (Partialsommen werden beliebig groß) treffen.

Satz 4.7 (Satz 4.5'). Die Reihe $\sum_n a_n$ divergiert, wenn es eine Zahl $q > 1$ derart gibt, dass $\frac{a_{n+1}}{a_n} \geq q$ für fast alle n gilt.

Satz 4.8 (Satz 4.6'). Die Reihe $\sum_n a_n$ divergiert, wenn es eine Zahl $q > 1$ derart gibt, dass $(a_n)^{\frac{1}{n}} \geq q$ für fast alle n gilt (es reichen schon unendlich viele n mit dieser Eigenschaft).

Beweis. Übungsaufgabe. □

Beispiele 4.9.

1. $a_n = \frac{x^n}{n!}$ ($x > 0$).

Es ist $\frac{a_{n+1}}{a_n} = \frac{x}{n+1} < \frac{1}{2}$ für große n . Also liegt Konvergenz für $\sum_n a_n$ vor.

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

Die n -te Wurzel liefert $\frac{n+2}{2n+1} = \frac{1+\frac{2}{n}}{2+\frac{1}{n}} < \frac{3}{4}$ für große n . Also: $\sum_n a_n$ konvergiert.

3. $a_n = nx^n$.

$\frac{a_{n+1}}{a_n} = x \frac{n+1}{n}$. Für $0 \leq x < 1$ Konvergenz.

4.3 Absolute Konvergenz

Wir betrachten jetzt auch Reihen $\sum_n a_n$, deren Summanden a_n beliebiges Vorzeichen haben dürfen.

Definition 4.10 (absolut konvergent). Eine Reihe $\sum_n a_n$ mit $a_n \in \mathbb{R}$ heißt absolut konvergent, wenn die Reihe $\sum_n |a_n|$ konvergiert.

Lemma 4.11 (absolut konvergent \Rightarrow konvergent.). Jede absolut konvergente Reihe ist auch konvergent.

Beweis. Es sei $\varepsilon > 0$ beliebig fixiert. Die Summe $S = \sum_{n=1}^{\infty} |a_n|$ existiert nach Voraussetzung. Wir betrachten für $m > n > n_0$ die Partialsommen der Originalreihe $s_n = a_1 + \dots + a_n$ und $s_m = a_1 + \dots + a_n + \dots + a_m$. Offenbar gilt $|s_m - s_n| = |a_{n+1} + \dots + a_m| \leq |a_{n+1}| + \dots + |a_m| \leq S - S_{n_0}$, wobei S_{n_0} die entsprechende Partialsomme der Absolutreihe kennzeichnet. Ist $n_0 > n(\varepsilon)$, so gilt nach Voraussetzung $S - S_{n_0} < \varepsilon$. Die Partialsommen $\{s_n\}$ bilden somit eine Cauchy-Folge. Nach Satz 2.12 konvergiert sie deshalb, was zu zeigen war. □

Bemerkung:

Diese hinreichende Konvergenzbedingung (Lemma 4.11) ist keine notwendige! Z.B. konvergiert die Reihe

$$(A1) \quad \sum_n \frac{(-1)^n}{n} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} \dots$$

(wie wir später sehen werden) ohne dass sie absolut konvergiert.

4.4 Summen von Reihen

Zwei konvergente Reihen $\sum_n a_n$ und $\sum_n b_n$ mit den Summen a bzw. b kann man formal addieren, indem man die Reihe mit den Summanden

$$c_n = a_n + b_n$$

bildet. Die Summe dieser Reihe ist dann gerade $a + b$, was sich unmittelbar aus den Partialsommen S_k dieser Reihe ergibt.

Interessant bleibt nur die Frage, ob dieselbe Summe $a + b$ bei beliebiger Klammersetzung entsteht.

Wir betrachten deshalb die Reihe mit den Summanden $a_1 + b_1 + a_2 + b_2 + \dots$

Für ihre Partialsommen s_k gilt dann offenbar $s_{2k} = S_k$ und $s_{2k-1} = S_k - b_k$.

Wegen $\lim S_k = a + b$ und $\lim b_k = 0$ (da die Reihe $\sum_n b_n$ konvergiert) müssen also auch die Partialsommen s_k gegen $a + b$ konvergieren. Im folgenden Satz erlauben wir, die Summanden einer Reihe umzusortieren und auf mehrere Reihen zu verteilen.

Satz 4.12 (Umordnungssatz). Sei $\sum_n a_n$ eine absolut konvergente Reihe, s ihre Summe ($n \geq 1$). Die Glieder der Reihe mögen in irgendeiner Art auf endlich oder unendlich viele Reihen

$$\sum_n b_{p,n} \quad (\text{Reihe } p \text{ für } p = 1, 2, 3, \dots)$$

aufgeteilt sein. Dann konvergieren auch diese Reihen absolut, und für ihre Summen $s(p) = \sum_n b_{p,n}$ gilt

$$\sum_p s(p) = s.$$

Beweis.

(I) Sei stets $a_k \geq 0$.

0. Wir betrachten zuerst den Fall, dass es nur eine Reihe $s(1) = \sum_n b_{1,n}$ gibt. Wir schreiben b_n statt $b_{1,n}$. Sei $\varepsilon > 0$ und $s = \lim a_n$, $s' = \lim b_n$ (wir können sogar $s, s' = \infty$ zulassen). Dann gibt es ein n , sodass $s_n = a_1 + \dots + a_n > s - \varepsilon$ (wir setzen $s - \varepsilon = \frac{1}{\varepsilon}$ falls $s = \infty$).

Die ersten n Summanden der Ausgangsreihe mögen sich in der zweiten Reihe unter den Elementen b_1, \dots, b_m befinden. Das ist für große m erfüllt. Dann gilt natürlich auch

$$s'_m := b_1 + \dots + b_m \geq s_n > s - \varepsilon.$$

Damit folgt $s' \geq s$. Andererseits kann man die Reihen vertauschen und den Schluss wiederholen. Das liefert $s \geq s'$, also $s = s'$.

1. Der Fall unendlich vieler Reihen. Sei dazu $\varepsilon > 0$.
Wegen der vorausgesetzten absoluten Konvergenz gibt es ein n_0 , sodass

$$(*) \quad a_{n_0+1} + a_{n_0+2} + \dots < \varepsilon.$$

Wie schon gezeigt, wird die Summe durch Umsortieren nicht größer, durch Weglassen von Summanden offenbar auch nicht.

Für großes q_0 treten die endlich vielen (wesentlichen) Elemente a_1, \dots, a_{n_0} schon alle in den Reihen mit Index $p \leq q_0$ auf. Es findet sich weiter ein n_1 , sodass sie bereits in den Partialsummen bis Index n_1

$$b_{p,1} + \dots + b_{p,n_1} = s_{n_1}(p) \text{ dieser } q_0 \text{ Reihen vorkommen.}$$

Jede Partialsumme von Reihe „ p “ ist deshalb durch $s_{n_1}(p) + \varepsilon$ beschränkt. Letzteres gilt erst recht für $p > q_0$.

Also konvergieren alle diese Reihen, ihre Summen $s(p)$ existieren. Weiter ist wegen (*)

$$s(1) + s(2) + \dots + s(q_0) > s - \varepsilon,$$

da dies schon für die n_1 -Partialsummen gilt. Andererseits gilt

$$s_n(1) + s_n(2) + \dots + s_n(q_0) \leq s \quad \forall n.$$

Diese Ungleichungen zeigen $|s(1) + \dots + s(q_0) - s| < \varepsilon$, d.h. $\sum_p s(p) = s$.

(II) Die Zahlen a_k besitzen beliebiges Vorzeichen.

Trick: Wir wenden den schon bewiesenen ersten Teil des Satzes auf die Reihen mit nichtnegativen Gliedern $A_k = a_k + |a_k|$ sowie $B_{p,k} = b_{p,k} + |b_{p,k}|$ an. Nun gilt

$$a_k = A_k - |a_k|, \quad b_{p,k} = B_{p,k} - |b_{p,k}|,$$

was sich offenbar auf entsprechende Partialsummen überträgt. □

4.5 Alternierende Reihen

Eine Reihe $\sum_n a_n$ heißt alternierend, wenn zwei aufeinanderfolgende Summanden stets verschiedenes Vorzeichen haben. Ein Beispiel ist die schon erwähnte Reihe

$$(A1) \quad \sum_n \frac{(-1)^n}{n} = -1 + \frac{1}{2} - \frac{1}{3} + \frac{1}{4} \dots$$

Sie besitzt die zusätzlichen Eigenschaften

$$(E1) \quad |a_{n+1}| \leq |a_n| \quad \forall n \quad \text{und}$$

$$(E2) \quad \lim |a_n| = 0.$$

Satz 4.13 (Leibnitz, alternierende Reihen). *Jede alternierende Reihe mit den Eigenschaften (E1) und (E2) konvergiert.*

Beweis. Sei o.B.d.A. (ohne Beschränkung der Allgemeinheit) das erste Element a_1 wie oben negativ. Für die Partialsummen mit geradem Index s_{2k} ist dann stets $s_{2(k+1)} \leq s_{2k}$ (Man setze Klammern um jeweils 2 aufeinanderfolgende Elemente).

Für Partialsummen mit ungeradem Index gilt $s_{2k+1} = a_1 + (a_2 + a_3) + \dots + (a_{2k} + a_{2k+1})$ und, da alle Klammern nichtnegativ sind, $s_{2(k+1)+1} \geq s_{2k+1}$

Beide Partialsummenfolgen sind dann in umgekehrtem Sinne monoton, wonach für alle $n > 2k + 1$ folgt:

$$s_{2k+1} \leq s_n \quad (n \text{ ungerade}) \quad \text{bzw.} \quad s_n \leq s_{2k} \quad (n \text{ gerade}).$$

Weiter konvergiert die Differenz $|s_{2k+1} - s_{2k}| = |a_{2k+1}|$ gegen Null. Damit müssen

$$c = \inf\{s_{2k} \mid k \in \mathbb{N}\} \quad \text{und} \quad d = \sup\{s_{2k+1} \mid k \in \mathbb{N}\}$$

existieren und gleich sein. Wählt man $n(\varepsilon)$ hinreichend groß, gilt also tatsächlich

$$|s_n - c| < \varepsilon \quad \forall n > n(\varepsilon).$$

□

4.6 Produkte von absolut konvergenten Reihen

Es seien $\sum_n a_n$ und $\sum_n b_n$ zwei absolut konvergente Reihen mit den Summen a bzw. b .

Setzt man $p(k) = s(k)t(k)$, wobei $s(k)$ und $t(k)$ die k -te Partialsumme der ersten bzw. zweiten Reihe kennzeichnet, so gilt offenbar $\lim p(k) = \lim s(k)t(k) = ab$.

Andererseits kann man sich alle Produkte der Form $a_n b_m$ irgendwie hintereinander aufgeschrieben denken, was zu einer weiteren Reihe führt, und nach der Summe dieser neuen Reihe fragen. Z.B. bilden wir die Reihe

$$\begin{aligned} c_1 &= a_1 b_1, \quad c_2 = a_1 b_2, \quad c_3 = a_2 b_1, \\ c_4 &= a_2 b_2, \quad c_5 = a_1 b_3, \quad c_6 = a_2 b_3, \quad c_7 = a_3 b_1, \quad c_8 = a_3 b_2, \\ c_9 &= a_3 b_3, \quad c_{10} = a_1 b_4, \quad c_{11} = a_2 b_4, \quad \dots \end{aligned}$$

Hier ist gerade $c_{k \cdot k} = p(k)$ und damit auch $ab = \lim c_{k \cdot k}$.

Wir überlegen uns, dass auch die Reihe $\sum_n c_n$ gegen ab konvergiert.

Dazu sei m beliebig fixiert und $k \cdot k$ die größte Quadratzahl $\leq m$; $k = k(m)$.

Weiter sei $C(m)$ die m -te Partialsumme der Reihe $\sum_n c_n$. Die Differenz $d = C(m) - p(k)$ erfüllt dann

$$|d| \leq |b_{k+1}|(|a_1| + \dots + |a_k|) + |a_{k+1}|(|b_1| + \dots + |b_k|) \leq |b_{k+1}|A + |a_{k+1}|B =: h(k),$$

wobei $A = \sum_n |a_n|$ und $B = \sum_n |b_n|$ gesetzt wurde. Ist nun $\varepsilon > 0$ beliebig gegeben, so findet sich zunächst ein k_0 derart, dass $h(k) < \frac{\varepsilon}{2}$ und gleichzeitig auch $|p(k) - ab| < \frac{\varepsilon}{2}$ für alle $k > k_0$ gilt.

Weiter gibt es ein m_0 , sodass $k(m_0) > k_0$ und damit auch $k(m) > k_0 \quad \forall m > m_0$ gilt. Dies liefert aber $|C(m) - ab| < \varepsilon \quad \forall m > m_0$; also gilt tatsächlich $\sum_n c_n = ab$.

Dieselbe Überlegung zeigt, dass auch die Reihe $\sum_n |c_n|$ gegen AB konvergiert. Unter Berücksichtigung des Umordnungssatzes erhalten wir so:

Satz 4.14 (Produktreihen). *Jede Reihe, die aus allen Produkten $a_n b_m$ der Glieder zweier absolut konvergenter Reihen $\sum_k a_k$ und $\sum_k b_k$ gebildet wird (und nur aus diesen), konvergiert gegen das Produkt der entsprechenden Summen (und ist selbst absolut konvergent).*

4.7 Potenzreihen

Definition 4.15 (Potenzreihe). *Reihen der Form $R(x) = \sum_n a_n x^n$, wobei $n \in \mathbb{N}$ die natürlichen Zahlen durchläuft und a_n sowie x reell sind, heißen (reelle) Potenzreihen.*

Bemerkung: Mit gegebenen a_n fragt man nun nach der Menge X aller $x \in \mathbb{R}$, für welche die Reihe (absolut) konvergiert. Dieselbe Frage ist später im Zusammenhang mit komplexwertigen a_n und x interessant (s. Kap. 7.1). Der folgende Satz zeigt, dass die Menge X eine sehr einfache Gestalt hat.

Satz 4.16 (reelle Potenzreihe). *Auf die Reihe $R(x)$ trifft genau eine der folgenden drei Aussagen zu:*

- (A1) *Sie konvergiert nur für $x = 0$,*
- (A2) *Sie konvergiert für alle x absolut,*
- (A3) *Es gibt ein $r \in \mathbb{R}$, sodass $R(x)$ für alle x mit $|x| < r$ absolut konvergiert und für kein x mit $|x| > r$ konvergiert.*

Bemerkungen:

1. Man setzt r in den beiden ersten Fällen gleich 0 bzw. ∞ und nennt r auch Konvergenzradius. Die Bezeichnung hat mit dem Fall komplexer Reihen zu tun, wo r der Radius eines Kreises wird. Hier kennzeichnet r natürlich ein Intervall.
2. Man beachte, dass im dritten Fall nichts für $|x| = r$ ausgesagt wird.
3. Mittels $L = \limsup \sqrt[n]{|a_n|}$ kann man die Fälle durch $L = \infty$, $L = 0$ bzw. $r = \frac{1}{L}$ charakterisieren. Dabei ist $\limsup \sqrt[n]{|a_n|}$ das Supremum aller $p \in \mathbb{R}$, sodass gilt $\sqrt[n]{|a_n|} > p$ für unendliche viele n . Äquivalent dazu: $L = \inf \{ q \in \mathbb{R} \mid \sqrt[n]{|a_n|} \leq q \text{ für fast alle } n \}$. Gilt $\sqrt[n]{|a_n|} > p$ für jedes p und unendliche viele n , setzt man $L = \infty$. Existiert $L' = \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}$, so ist (nach obiger Definition) einfach $L = L'$.

Die Bezeichnung $L = \limsup \sqrt[n]{|a_n|}$ kommt daher: Falls $0 \leq p < L$, so gibt es unendl. viele n mit $\sqrt[n]{|a_n|} > p$. Setzt man $p = L - 1/k$, findet man mit entsprechenden $n = n(k)$ eine Teilfolge aller n , sodass $L = \lim_{k \rightarrow \infty} \sqrt[n(k)]{|a_{n(k)}|}$. Da es nur endl. viele n mit $\sqrt[n]{|a_n|} > L + \varepsilon$ gibt, ist L zugleich das Supremum aller Limes dieser Form.

Beweis. Sei $L \in (0, \infty)$ und $r = \frac{1}{L}$. Man benutze das Wurzel-Kriterium. Ist $|x| \leq s < r$, so folgt $sL < 1$, und mit jedem $\varepsilon > 0$ und großen n :

$$q_n := \sqrt[n]{|a_n|} \cdot |x|^n = |x| \sqrt[n]{|a_n|} < s(L + \varepsilon).$$

Für kleine ε wird so $q_n \leq q < 1$. Deshalb folgt absolute Konvergenz für $R(x)$. Derselbe Schluss ist für $L = 0$ und $r = \infty$ möglich.

Divergenz: Wieder sei $L \in (0, \infty)$ und $r = \frac{1}{L}$. Ist nun $|x| > r$, nehme man $\varepsilon > 0$, mit $|x|(L - \varepsilon) > 1$. Nun gibt es unendlich viele n mit $\sqrt[n]{|a_n|} > L - \varepsilon$, also auch $|x| \sqrt[n]{|a_n|} > 1$. Wegen $\sqrt[n]{|a_n|} \cdot |x|^n = |x| \sqrt[n]{|a_n|} > 1$, ist auch $|a_n x^n| > 1$. Die Reihe $R(x)$ kann daher nicht konvergieren (Summanden gehen nicht gegen Null). Hier ist dieser Schluss mit $L = \infty$ und $r = 0$ ebenso anwendbar. \square

Ergänzung: Man kann auch direkt zeigen, dass gilt $r = \sup\{|y| \mid R(y) \text{ konvergiert}\}$, hat dann aber noch nicht die Beschreibung von r als Kehrwert von L :

Beweis. Sei $r \in (0, \infty)$, andernfalls gilt (A1) oder (A2). Nach Definition von r existiert ein $y \neq 0$, sodass die Reihe für $x = y$ konvergiert. Dann muss gelten $|a_n y^n| < 1$ für hinreichend große n , etwa $n > n_0$.

Für jedes feste x mit $|x| < |y|$ folgt so $|x| = q \cdot |y|$ (mit $q = \frac{|x|}{|y|} < 1$) und $|a_n \cdot |x^n| = q^n \cdot |a_n| \cdot |y^n| < q^n$. Durch Vergleich mit der geometrischen Reihe folgt so die absolute Konvergenz der Reihe $R(x)$ für alle x mit $|x| < |y|$. Da y beliebig gewählt war (mit konvergentem $R(y)$) folgt absolute Konvergenz für jedes feste x mit $|x| < r$.

Divergenz für $|x| = t > r$: Nach Definition von r gibt es zu jedem $t > r$ ein y mit $|y| < t$, sodass $R(y)$ nicht konvergiert. Wenn $|a_n t^n| < 1$ für alle großen n gilt, konvergiert $R(y)$ sogar absolut, wie oben gesehen. Also muss $|a_n t^n| \geq 1$ für unendlich viele n gelten. Dann folgt aber, dass $R(t)$ nicht konvergieren kann, und auch $R(-t)$ erfüllt die notwendige Konvergenzbedingung $\lim_{n \rightarrow \infty} |a_n t^n| = 0$ nicht. Also ist auch $R(x)$ nicht konvergent. \square

Die Bedeutung von Potenzreihen wird später im Zusammenhang mit Taylor-Reihen (s. Kap. 6.3) klar. Sie erlauben oft die einfache Darstellung der Funktionswerte komplizierterer Funktionen.

Beispiele 4.17. *Jeder der drei Fälle ist möglich.*

$$(A1) \quad R(x) = \sum (n^n) x^n$$

$$(A2) \quad R(x) = \sum \frac{x^n}{n!}$$

$$(A3) \quad R(x) = \sum n x^n.$$

4.8 Eulersche Zahl e

Definition 4.18 (Eulersche Zahl). *Die Eulersche Zahl e , die in vielfältigen Bereichen der Mathematik eine sehr wichtige Rolle spielt, ist definiert als Summe der Reihe*

$$e = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \dots + \frac{1}{n!} + \dots,$$

also $e = \sum_{n \geq 0} a_n$ mit $a_n = \frac{1}{n!}$ ($n \geq 0$).

- Wir überlegen uns, dass die angegebene Reihe tatsächlich konvergiert, die Definition also sinnvoll ist. Dazu vergleichen wir sie mit der Summe

$$s = 1 + 1 + \frac{1}{2^1} + \frac{1}{2^2} + \frac{1}{2^3} + \dots + \frac{1}{2^{n-1}} + \dots$$

Da $2^{n-1} \leq n!$ für $n \geq 2$ gilt, lassen sich die Summanden $\frac{1}{n!} \leq \frac{1}{2^{n-1}}$ einzeln abschätzen.

Die Summe s ist aber gerade 3, wie eine Vergleich mit der geometrischen Reihe für $q = \frac{1}{2}$ ergibt. Nach Lemma 4.2 konvergiert also die Reihe $\sum \frac{1}{n!}$ ($n \geq 0$) wirklich.

- Als nächstes zeigen wir (mit einigen Tricks), dass für e gleichzeitig gilt

$$e = \lim_{n \rightarrow \infty} b_n \quad \text{mit} \quad b_n = \left(1 + \frac{1}{n}\right)^n.$$

Daher kann man auch definieren $e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$.

Beweis. (der Gleichheit): Nach dem binomischen Satz ist

$$b_n = 1 + \binom{n}{1} \left(\frac{1}{n}\right)^1 + \dots + \binom{n}{k} \left(\frac{1}{n}\right)^k + \dots + \binom{n}{n} \left(\frac{1}{n}\right)^n,$$

was sich schreiben lässt als

$$(1) \quad b_n = 1 + f_n(1) \frac{1}{1!} + \dots + f_n(k) \frac{1}{k!} + \dots + f_n(n) \frac{1}{n!},$$

mit den Faktoren

$$(2) \quad f_n(k) = \binom{n}{n} \cdot \binom{n-1}{n} \cdot \dots \cdot \binom{n-k+1}{n} \quad (n \text{ Faktoren} \leq 1)$$

(Zähler von $\binom{n}{k}$ durch n^k dividiert). Wegen $f_n(k) \leq 1$ ($\forall k$), gilt somit $b_n \leq e \forall n$. Ersetzt man in (2) n durch $n+1$ und lässt $k \leq n$ fest (dann erhöhen sich jeweils Zähler und Nenner um 1), werden die einzelnen Faktoren nicht kleiner:

$$f_{n+1}(k) = \binom{n+1}{k} \cdot \binom{n}{n+1} \cdot \binom{n-1}{n} \cdot \dots \cdot \binom{n-k+2}{n+1} \geq f_n(k)$$

(denn mit positiven α, β gilt stets $\frac{\alpha+1}{\beta+1} \geq \frac{\alpha}{\beta} \Leftrightarrow \alpha \leq \beta$), und im Vergleich mit b_n kommt ein weiterer positiver Summand $f_{n+1}(n+1) \frac{1}{(n+1)!}$ zur Summe (1) für b_{n+1} hinzu. Damit gilt

$$b_n < b_{n+1} \leq e.$$

Die monotone und beschränkte Folge $\{b_n\}$ konvergiert somit gegen eine Zahl $z = \lim b_n \leq e$. Wir zeigen indirekt $z = e$. Angenommen, es ist $z < e$. Dann muss für hinreichend große k auch

$$(3) \quad z < 1 + \frac{1}{1!} + \dots + \frac{1}{k!}$$

wegen der Konvergenz dieser Partialsummen gegen e richtig sein. Wir halten ein solches $k = k_0$ fest und betrachten b_n für große $n > k_0$. Sicher gilt (weil wir rechts Summanden aus der Binomialsumme (1) für b_n weglassen)

$$(4) \quad b_n \geq 1 + f_n(1) \frac{1}{1!} + \dots + f_n(k_0) \frac{1}{k_0!}.$$

In der rechten Seite dieser Ungleichung konvergieren alle k_0 Faktoren $f_n(p)$ ($1 \leq p \leq k_0$) gegen 1, wenn $n \rightarrow \infty$, denn jeder einzelne Faktor in den k_0 Produkten $f_n(p)$, siehe (2), konvergiert gegen 1, wenn $n \rightarrow \infty$. Mit anderen Worten,

$$\lim_{n \rightarrow \infty} \left(1 + f_n(1) \frac{1}{1!} + \dots + f_n(k_0) \frac{1}{k_0!} \right) = 1 + \frac{1}{1!} + \dots + \frac{1}{k_0!} > z.$$

Für große n muss deshalb die rechte Seite von (4) und erst recht b_n ebenfalls größer als z werden. Wegen Monotonie kann andererseits b_n nie größer als $z = \lim b_n$ sein. Dieser Widerspruch zeigt, dass nur $z = e$ sein kann. \square

Bemerkung: Wir haben so zwei Darstellungen der Zahl e :

1. Als Summe der Reihe $e = \sum_{n=0}^{\infty} \frac{1}{n!}$ (mit $\frac{1}{0!} = 1$)
2. Als Grenzwert der Folge $e = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} \right)^n$ ($n \geq 1, n \in \mathbb{N}$)

Die Konvergenz der Partialsummen der Reihe gegen e erfolgt viel schneller als die der Folge.

Wir können nun weitere Grenzwerte $\lim a_n$ berechnen.

$$(G1) \quad a_n = \left(1 - \frac{1}{n} \right)^n$$

Man betrachtet den Kehrwert

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

Der erste Faktor konvergiert gegen e , der zweite gegen 1. Also ist $\lim a_n = \frac{1}{e} = e^{-1}$.

$$(G2) \quad a_n = \left(1 + \frac{x}{n} \right)^n, \quad x > 0 \text{ fixiert.}$$

Dann ist

$$a_n = \left(1 + \frac{1}{\frac{n}{x}} \right)^n = \left[\left(1 + \frac{1}{\frac{n}{x}} \right)^{\frac{n}{x}} \right]^x =: (h_n)^x.$$

Die Zahl $\frac{n}{x}$ lässt sich durch Einschachteln abschätzen: $n_1 \leq \frac{n}{x} < n_1 + 1$, wobei n_1 ganzzahlig. Für den Term h_n ergibt sich dann

$$\left(1 + \frac{1}{n_1 + 1} \right)^{n_1} \leq h_n \leq \left(1 + \frac{1}{n_1} \right)^{n_1 + 1}. \quad (4.8.1)$$

Schreibt man die linke Seite als

$$\left(1 + \frac{1}{n_1 + 1}\right)^{n_1 + 1} \left(1 + \frac{1}{n_1 + 1}\right)^{-1},$$

sieht man, dass sie mit wachsendem n_1 (und n) gegen e konvergiert. Analog behandelt man die rechte Seite, sie strebt ebenfalls gegen e . Also ist auch $\lim h_n = e$.

Aus der Stetigkeit (sie wird später allgemein im metrischen Raum definiert) der Funktion $f(y) = y^x$ für $y > 0$, folgt damit schließlich $\lim a_n = \lim (h_n)^x = e^x$.

$$(G3) \quad a_n = \left(1 + \frac{x}{n}\right)^n, \quad x < 0 \text{ fixiert.}$$

Nutzt man jetzt (G1), lässt sich analog zu (G2) zeigen, dass wiederum $\lim a_n = e^x$ gilt.

Bemerkung: Ein Vergleich von $(1 + \frac{x}{n})^n$ und $\sum \frac{x^n}{n!}$ ($n \geq 0$, mit $0! = 1$), der sich ähnlich wie im schon studierten Fall $x = 1$ nachvollziehen lässt, zeigt, dass auch diese Summe gegen e^x konvergiert, also

$$e^x = \sum_{n \geq 0} \frac{x^n}{n!} \quad \forall x \in \mathbb{R}. \quad (4.8.2)$$

Die entsprechenden Partialsummen sind schon nach wenigen Schritten gute Approximationen des tatsächlichen Wertes und werden für numerische Approximation auch genutzt. Ähnliche Reihendarstellungen für andere Funktionen lernen wir im Zusammenhang mit Taylor-Reihen (Kap. 6.3, S. 60) kennen. Die Reihe (4.8.2) ist die Taylor-Reihe für $f(x) = e^x$.

Für Spezialisten:

Für welche $\alpha > 0$ (speziell $\alpha > 1$) konvergiert die Folge $(x(n))_{n \in \mathbb{N}}$ mit $x(1) = \alpha, \dots, x(n+1) = \alpha^{x(n)}$? Welchen „Konvergenzradius“ r hat die Reihe $\sum x^n \tan(n)$?

Kapitel 5

Zusatz: Definition und Eigenschaften einiger wichtiger Funktionen

5.1 Die Potenz- und Logarithmus Funktion

Wir betrachten in diesem Abschnitt die Potenzfunktion $z = x^y$ mit $x, y \in \mathbb{R}$ und $x > 0$.

Wir haben bisher (streng genommen) nur die Grundrechenarten für reelle Zahlen definiert. Deshalb soll hier wenigstens einmal ausführlich gezeigt werden, welche Schritte nötig sind, um reelle Potenzen und deren Umkehrfunktionen korrekt zu definieren.

1. Zunächst kann man x^p für natürliche Zahlen p wie üblich als p -faches Produkt, und
2. $x^0 = 1$ sowie $x^{-p} = \frac{1}{x^p}$ definieren.
3. $x^{\frac{1}{p}}$ ($p > 1$ natürliche Zahl) wird als Umkehrfunktion (Inverse) zu $f(x) = x^p$ definiert. Hierzu kann man die Stetigkeit und Monotonie von $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ ausnutzen (wobei $\mathbb{R}_+ = \{x \in \mathbb{R} \mid x \geq 0\}$).
Stetigkeit folgt per vollständige Induktion bzgl. p , da $f = x^1$ stetig ist und $x^{p+1} = x \cdot x^p$ dann das Produkt zweier stetiger Funktionen ist. Monotonie in x folgt analog.
Damit ist $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ bijektiv, und die Inverse f^{-1} existiert und ist erneut stetig (das war Übungsaufgabe) und monoton:

$$x = f^{-1}(y) \Leftrightarrow y = f(x) := x^p.$$

Wir schreiben dann $x = y^{\frac{1}{p}}$.

Man beachte, dass wegen $f^{-1} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ nur $x \geq 0$ möglich ist, wir kümmern uns also nur um nichtnegative Wurzeln.

4. $f(x) = x^{\frac{p}{q}}$ für positive natürliche Zahlen p, q lässt sich nun durch $x^{\frac{p}{q}} = (x^p)^{\frac{1}{q}}$ definieren. Dies führt (nach den Regeln der Multiplikation) zu der Identität

$$(x^p)^{\frac{1}{q}} = x^{\frac{p}{q}} = \left(x^{\frac{1}{q}}\right)^p \quad (\text{Beweis durch Nachrechnen})$$

5. Das Gestatten negativer natürlicher Zahlen p und q ändert hieran nichts, wenn man nun 2. auf rationale Exponenten erweitert.

Damit ist die Definition von x^y für rationale Exponenten y abgeschlossen. Man beachte, dass jetzt für rationale u und v auch gilt:

$$(*) \quad x^{u+v} = x^u x^v \quad \text{und} \quad \lim_{u \rightarrow 0} x^u = 1.$$

Tatsächlich, mit $u = \frac{p}{q}$ und $v = \frac{p'}{q'}$ folgt $u + v = \frac{pq' + p'q}{qq'}$ und

$$x^u x^v = x^{\frac{p}{q}} x^{\frac{p'}{q'}} = (x^{pq'})^{\frac{1}{qq'}} (x^{p'q})^{\frac{1}{qq'}} = (x^{pq' + p'q})^{\frac{1}{qq'}} = x^{\frac{pq' + p'q}{qq'}} = x^{u+v}.$$

Ausgangspunkt für den Limes in (*) ist der uns schon bekannte Grenzwert $\lim_{n \rightarrow \infty} x^{\frac{1}{n}} = 1$ und das Einschachtelungsprinzip, wir lassen hier die Details weg und verweisen auf (4.8.1).

Den eigentlich interessanten Teil der Definitionen bilden reelle Exponenten und die Inverse Funktion. Sie können eleganter mit den beiden kommenden Hilfsätzen (nächster Abschnitt) behandelt werden. Wer alles explizit mag, kann so vorgehen:

6. Reelle Exponenten: $f(x) = x^r$; wir untersuchen zuerst den Fall $r > 1$, $x > 1$.

Man kann jetzt $r = \lim y(k)$ mit rationalen $y(k)$ ($k = 1, 2, 3, \dots$) schreiben. Dann sind die reellen Zahlen $z_k = x^{y(k)}$ stets definiert.

Existenz des Grenzwertes der z_k :

Wegen Konvergenz der $y(k)$ gibt es rationale, monoton wachsende α_k und monoton fallende β_k mit

$$\alpha_k \rightarrow r, \quad \beta_k \rightarrow r \quad \text{und} \quad y(k) \in (\alpha_k, \beta_k).$$

Wir schreiben $\beta_k = \alpha_k + h_k$; dann ist h_k wieder rational und wegen (*) auch $\lim x^{h_k} = 1$.

Zu jedem $\varepsilon > 0$ gibt es deshalb ein $k(\varepsilon)$, sodass

$$r - \varepsilon < \alpha_k < y(k) < \beta_k < r + \varepsilon \quad \text{und} \quad |x^{h_k} - 1| < \varepsilon \quad \forall k > k(\varepsilon).$$

Also ist wegen Monotonie der Funktion $g(u) = x^u$ für rationale u und (*):

$$x^{\alpha_k} < x^{y(k)} = z_k < x^{\beta_k} = x^{\alpha_k} x^{h_k} < x^{\alpha_k} (1 + \varepsilon) \quad \forall k > k(\varepsilon).$$

Wegen Beschränktheit $x^{\alpha_1} < x^{\alpha_k} < x^{\beta_k} < x^{\beta_1}$ und Monotonie, ergibt sich die Konvergenz der reellen Folge x^{α_k} . Sei $u \in \mathbb{R}$ ihr Limes. Ist $k > k(\varepsilon)$ so groß, dass auch $|x^{\alpha_k} - u| < \varepsilon$ gilt, etwa für $k > K(\varepsilon)$, folgt schließlich $u - \varepsilon < z_k < (u + \varepsilon)(1 + \varepsilon) \quad \forall k > K(\varepsilon)$.

Die Folge aller z_k ist deshalb eine Cauchy-Folge, denn zu beliebigem $\varepsilon' > 0$ finden wir zuerst ein $\varepsilon > 0$ mit $u - \frac{1}{2}\varepsilon' < u - \varepsilon < (u + \varepsilon)(1 + \varepsilon) < u + \frac{1}{2}\varepsilon'$, wählen dann $K' = K(\varepsilon)$ und sehen, dass nun

$$|z_m - z_n| < \varepsilon' \quad \forall m, n > K'$$

gilt. Damit hat die Folge $z_k = x^{y(k)}$ stets einen Limes $z \in \mathbb{R}$, wenn $r = \lim y(k)$ mit rationalen $y(k)$ gilt.

Eindeutigkeit des Grenzwertes der z_k :

Wir zeigen, dass der Limes $z = \lim x^{y(k)}$ nicht von der speziellen Wahl der Folge $\{y(k)\}$ abhängt (dann folgt insbesondere $z = u$). Ist nämlich $\{y'(k)\}$ eine weitere solche Folge mit entsprechendem Limes $z' = \lim x^{y'(k)}$, so können wir die neue Folge $\{Y(k)\}$ in der Form $y(1), y'(1), y(2), y'(2), \dots$ bilden. Auch sie ist rational und konvergiert gegen r . Mit zugeordneten Limes $Z = \lim x^{Y(k)}$ folgt dann $z = Z = z'$, weil die beiden ersten Folgen unendlich Teilfolgen der letzteren $\{Y(k)\}$ sind. Also kann man definieren

$$(**) \quad x^r = z, \quad \text{wobei } z \text{ der eindeutig bestimmte Limes } z = \lim x^{y(k)} \text{ für rationale } y(k) \rightarrow r \text{ ist.}$$

Monotonie und Stetigkeit bleiben nach dieser Limes-Definition erhalten.

Schließlich betrachten wir beliebige $r > 0$, $x > 0$.

Analog zu oben könnte man mit den entsprechenden Monotonie-Eigenschaften in diesen Fällen verfahren. Man kann die fehlenden Fälle aber auch auf die schon bekannten zurückführen.

Ist $r > 1$ und $0 < x < 1$, setze man $y = \frac{1}{x} > 1$ und definiere $x^r = \frac{1}{y^r}$.

Ist $0 < r < 1$, bilde man $s = \frac{1}{r}$ und definiere x^r als Umkehrfunktion zu x^s , d.h.

$$y = x^r \quad \Leftrightarrow \quad x = y^s = y^{\frac{1}{r}}.$$

Für negative Exponenten r setzt man schließlich wieder $x^r = \frac{1}{x^{-r}}$. Damit gilt nun (**) auch für beliebige $r \in \mathbb{R}$ und $x > 0$.

Die Logarithmus-Funktion

Nun ist $f(x) = a^x$ für festes $a > 0$ und $x \in \mathbb{R}$ eindeutig definiert, positiv und stetig.

Für $a > 1$ ist f monoton wachsend (warum?) und erfüllt $\lim_{x \rightarrow -\infty} a^x = 0$ und $\lim_{x \rightarrow \infty} a^x = \infty$.

Für $a < 1$ ist f fallend (setze $a = \frac{1}{b}$, $b > 1$ und beachte $a^x = \frac{1}{b^x} = b^{-x}$) und erfüllt $\lim_{x \rightarrow -\infty} a^x = \infty$ und

$$\lim_{x \rightarrow \infty} a^x = 0.$$

Vom Trivialfall $a = 1$ abgesehen, besitzt $f : (-\infty, \infty) \rightarrow (0, \infty)$ also eine (stetige) Inverse $f^{-1} : (0, \infty) \rightarrow (-\infty, \infty)$, definiert durch $x = f^{-1}(y) \Leftrightarrow y = a^x$ (mit $0 < y < \infty$).

Sie hängt von der Konstanten a ab und heißt Logarithmus von y zur Basis a , kurz

$$x = \log_a y \quad (= f^{-1}(y)).$$

Also gilt

$$(L0) \quad x = \log_a y \Leftrightarrow y = a^x$$

Diese Definition hat wegen $a^{x+x'} = a^x a^{x'}$ die Konsequenz, dass mit $y = a^x$ und $y' = a^{x'}$ (bzw. mit $x = \log_a y$ und $x' = \log_a y'$) sofort folgt: $yy' = a^{x+x'}$. Also ist nach Definition $x + x' = \log_a(yy')$ und

$$(L1) \quad x + x' = \log_a y + \log_a y' = \log_a(yy').$$

Der Logarithmus eines Produkts ist die Summe der einzelnen Logarithmen. Weiter folgt für $p \in \mathbb{R}$

$$(L2) \quad \log_a(y^p) = p \log_a y \quad (\text{Potenzen im Argument erscheinen außen als Faktor}),$$

denn wegen $a^x = y$ und $a^{px} = (a^x)^p = y^p$, gilt auch $px = \log_a y^p$.

Schließlich kann man Logarithmen zu unterschiedlichen Basen a und a' miteinander verbinden. Wir betrachten den natürlichen Logarithmus zur Basis $a' = e$, der als $\ln y$ geschrieben wird, und zeigen

$$(L3) \quad \log_a y = \frac{\ln y}{\ln a}.$$

Wieder sei $x = \log_a y$. Nun nutzt man (L2) und die Definition von \ln :

$$e^z = a^x \Leftrightarrow z = \ln(a^x) = x \ln a.$$

So folgt $\frac{z}{x} = \ln a$, also wegen $a^x = y$ auch

$$\ln a = \frac{\ln(a^x)}{\log_a y} = \frac{\ln y}{\log_a y}.$$

Nach Ausmultiplizieren ist das (L3).

Der Logarithmus $\log_a y$ zur Basis a und $\ln y$ unterscheiden sich also nur durch den Faktor $\frac{1}{\ln a}$. Ferner ist $e^{\ln y} = y$, denn $x = \ln y$ bedeutet ja $e^x = y$. Analog: $y^p = e^{p \ln y}$.

5.2 Erweiterung von f auf \mathbb{R} und f^{-1} allgemein

Dieses Vorgehen (zur Definition von x^y und der Inversen) kann durch die beiden folgenden Hilfssätze verallgemeinert und verkürzt werden:

Eine reelle Funktion heißt streng monoton wachsend, wenn gilt $f(y) > f(x)$ falls $y > x$. Analog streng monoton fallend. Es folgt damit sofort, dass f injektiv ist, denn $x \neq y$ hat $f(x) \neq f(y)$ zur Folge.

Ist f auch noch stetig, kann man die inverse Funktion charakterisieren (die offenbar im selben Sinne monoton sein muss).

Hilfssatz 5.1 (monotone, stetige Inverse). *Sei f eine streng monotone (wachsend oder fallend) und stetige reelle Funktion mit $f(\mathbb{R}) = Y$. Dann existiert die Inverse $f^{-1} : Y \rightarrow \mathbb{R}$ und ist ebenfalls stetig. ($x = f^{-1}(y) \Leftrightarrow f(x) = y$)*

Beweis. Existenz von f^{-1} :

Sei f streng monoton wachsend. Klar, zu jedem $y \in Y$ gibt es wegen strenger Monotonie genau ein x mit $f(x) = y$.

Stetigkeit von f^{-1} : Gelte $y_n \rightarrow y$ in Y und sei $x_n = f^{-1}(y_n)$, $x = f^{-1}(y)$.

Wegen Monotonie liegen dann alle x_n (für grosse n) im Intervall $[x - 1, x + 1]$, denn ausserhalb ist f entweder $< f(x - 1) < y$ oder $> f(x + 1) > y$. Die beschränkte Folge der x_n hat eine konvergente Teilfolge (Bolzano-Weierstraß, Kap.2.10), etwa fuer $n \in N' \subset N$. Also existiert $x^* = \lim_{n \in N', n \rightarrow \infty} x_n$. Damit folgt für $n \in N'$ und $n \rightarrow \infty$ mit Stetigkeit von f : $y = \lim y_n = \lim f(x_n) = f(x^*)$. Falls $x^* \neq x$ hätten wir 2 Urbilder $x^* \neq x$ von y . Dies ist ausgeschlossen wegen Monotonie. Also muss $x_n = f^{-1}(y_n)$ stets für eine unendliche Teilfolge aller y_n gegen $f^{-1}(y) = x$ konvergieren. Daraus folgt aber sofort die Stetigkeit (warum?). Analog kann man schließen bei umgekehrter Monotonie.

□

Aus den Übungen wissen wir schon, dass mit $a, b \in Y$ auch jedes c zwischen a und b zu Y gehört (Beweis durch Intervallschachtelung).

Hilfssatz 5.2 (stetige Erweiterung einer reellen Funktion von \mathbb{Q} auf \mathbb{R}). *Sei eine reellwertige Funktion f auf einem rationalem Intervall $(a, b) \cap \mathbb{Q}$ definiert und dort gleichmässig stetig. Dann kann sie für alle reellen (irrationalen) $x \in (a, b)$ so definiert werden, dass sie insgesamt stetig auf (a, b) wird.*

Beweis. Sei $x \in (a, b)$ reell, irrational. Wir wollen setzen $f(x) = \lim f(p_n)$, $p_n \rightarrow x$, p_n rational. Dazu brauchen wir

- (1) die Existenz des Limes,
- (2) dass der Limes derselbe bleibt, wenn wir eine andere rationale Folge $q_n \rightarrow x$ nehmen,
- (3) Stetigkeitsnachweis für f .

(1) Existenz des Limes

Sei $\varepsilon > 0$ und $\delta = \delta(\varepsilon)$ entsprechend glm. Stetigkeit gewählt. Für große n, m (wenn $|p_n - x| < \delta/2$ und $|p_m - x| < \delta/2$) liegen alle p_n in einer δ -Kugel (Intervall) um p_m . Dann folgt also $|f(p_n) - f(p_m)| < \varepsilon$. Damit bilden die Funktionswerte $f(p_n)$ eine Cauchy-Folge in \mathbb{R} . Diese ist konvergent; also existiert $r = \lim f(p_n)$.

(2) Wohldefiniertheit

Nimmt man eine andere Folge $q_n \rightarrow x$, gilt $|p_n - q_n| < \delta$ für große n , also ebenfalls $|f(p_n) - f(q_n)| < \varepsilon$. Der Limes $r' = \lim f(q_n)$ (dessen Existenz wir unter (1) gerade gezeigt haben) muss also (weil $\varepsilon > 0$ beliebig klein sein darf) mit r zusammenfallen. Damit ist die Definition $f(x) = \lim f(p_n)$ eindeutig, d.h. nicht abhängig von der speziellen Wahl der rationalen Folge $\{p_n\}$ mit $x = \lim p_n$.

(3) Stetigkeit von f

Wenn $x_n \rightarrow x$ (reell oder nicht), findet sich (nach Definition von $f(x_n)$ als Limes) zu jedem n ein rationales p_n , sodass $|f(p_n) - f(x_n)| < \varepsilon$ und $|p_n - x_n| < \frac{1}{n}$. Weil nun $p_n \rightarrow x$ gilt, folgt $f(x) = \lim f(p_n)$. Damit gilt auch $|f(x) - f(x_n)| < |f(x) - f(p_n)| + |f(p_n) - f(x_n)| < 2\varepsilon$ für grosse n (sodass $|f(p_n) - f(x)| < \varepsilon$). Daher muss (weil $\varepsilon > 0$ beliebig klein sein darf) auch $f(x) = \lim f(x_n)$ gelten. \square

Bemerkung: Der Hilfssatz wird falsch, wenn man nur voraussetzt (statt glm. stetig), dass f in jedem rationalen $p \in (a, b)$ stetig ist. Beispiel: $f(p) = \sin(\frac{1}{p-\sqrt{2}})$. Die Stetigkeit des Sinus folgt aus dem Additionstheorem $\sin(x + h_n) = \sin x \cos h_n + \cos x \sin h_n$ und den nach Def. am Kreis fuer $h_n \rightarrow 0$ offensichtlichen Limes $\lim \sin h_n = 0$ und $\lim \cos h_n = 1$. Wenn $p \in \mathbb{Q}$, wird der Nenner nicht Null. Dann ist f (Verknüpfung zweier stetiger Funktionen) auf \mathbb{Q} stetig. Wir können nun rationale $p_n \rightarrow \sqrt{2}$ finden, sodass $\frac{1}{p_n - \sqrt{2}}$ beliebig nahe an einer Nullstelle $\mu_n = n\pi$ des Sinus liegt. Man nehme z.B. y_n sodass $\frac{1}{y_n - \sqrt{2}} = n\pi$ und wähle p_n rational, hinreichend nahe bei y_n . In diesem Fall strebt $f(p_n)$ gegen $r = 0$. Analog gibt es rationale $q_n \rightarrow \sqrt{2}$, sodass $\frac{1}{q_n - \sqrt{2}}$ beliebig nahe an einer Maximumstelle $\mu_n = \frac{1}{2}\pi + 2n\pi$ des Sinus liegt. In diesem Fall strebt $f(q_n)$ gegen $r' = 1$. Man kann also $f(\sqrt{2})$ trotz Stetigkeit auf \mathbb{Q} nicht unabhängig von der speziellen Folge $p_n \rightarrow \sqrt{2}$ definieren.

Als Beispiel einer Anwendung betrachten wir erneut den Logarithmus

Sei $f(x) = c^x$ für rationale $x = \frac{m}{n}$ in (a, b) mit $0 < a$ und $c > 1$ (fest). f ist definiert durch $f(x) = c^{\frac{m}{n}} = (\sqrt[n]{c})^m = \sqrt[n]{c^m} > 1$. Damit gilt für rationale p, h : $f(p + h) = c^{p+h} = c^p c^h$ und $f(p + h) - f(p) = c^p(c^h - 1)$. Nutzt man $c^h \rightarrow 1$ wenn $h \rightarrow 0$ (rational) sowie Beschränktheit $c^p < K$ für alle $p \in (a, b)$, so folgt mit $|f(p + h) - f(p)| = c^p|c^h - 1| \leq K|c^h - 1|$ die glm. Stetigkeit von f auf $(a, b) \cap \mathbb{Q}$: Zu gegebenen ε mache man $\delta > 0$ klein genug, sodass $|c^h - 1| < \frac{\varepsilon}{K}$ für $|h| < \delta$ gilt.

Also kann Hilfssatz 5.2 angewandt werden, um c^x über $c^x = \lim c^{p_n}$, $p_n \rightarrow x$ zu definieren, wonach die resultierende reelle Funktion stetig wird; zunächst für $x > 0$, aber analog auch für $x < 0$, wenn man (wie üblich) definiert, dass dann $c^x = \frac{1}{c^{-x}}$ ist.

Schließlich setzt man wieder $c^0 = 1$, womit die nun auf ganz \mathbb{R} definierte Funktion $f(x) = c^x$ überall stetig (und streng monoton) wird.

Nach Hilfssatz 5.1 existiert auf $f(\mathbb{R}) = Y = \{y \mid y > 0\}$ die inverse Funktion f^{-1} .

$f^{-1}(y)$ wird *Logarithmus zur Basis c von y* genannt und mit $\log_c y$ bezeichnet. Für $0 < c < 1$, wonach c^x streng monoton fällt, ist dieselbe Konstruktion möglich. Man kann aber einfacher direkt definieren: $c^x = \frac{1}{d^x}$, wobei $d = \frac{1}{c} > 1$.

Kapitel 6

Differentialrechnung

Definition 6.1 (differenzierbar, Ableitung). *Eine reelle Funktion f heißt in x^* differenzierbar, wenn gilt:*

Für alle Folgen $u_k \rightarrow 0$ mit $u_k \neq 0$ ($k = 1, 2, 3, \dots$) haben die Differenzenquotienten

$$a_k = \frac{f(x^* + u_k) - f(x^*)}{u_k}$$

denselben Limes $a = \lim_{k \rightarrow \infty} a_k$.

In diesem Fall heißt a Ableitung von f in x^ und wird oft mit $a = f'(x^*)$ oder $a = Df(x^*)$ bezeichnet.*

Die Differenzierbarkeitsbedingung wird oft kürzer als

$$a = \lim_{u \rightarrow 0, u \neq 0} \frac{f(x^* + u) - f(x^*)}{u}$$

geschrieben.

Man beachte jedoch, dass dies wie oben bedeuten soll: Der Limes existiert für alle solche u -Folgen und hängt nicht von ihrer speziellen Wahl ab.

Wir reden also nicht, wie oft in der Schule, allein von rechts- und linksseitigen Limes.

Beispiele 6.2.

1) $f(x) = x^2$, $x^* = 4$.

Für $u \neq 0$ gilt $f(x^* + u) - f(x^*) = [(x^*)^2 + 2x^*u + u^2] - (x^*)^2 = 2x^*u + u^2$, also

$$\frac{f(x^* + u) - f(x^*)}{u} = 2x^* + u \quad \text{und} \quad a = 2x^* = 8.$$

2) $f(x) = |x|$, $x^* = 0$.

Für $u_k = u \rightarrow 0$ mit $u > 0$ gilt: $\lim_{u \rightarrow 0} \frac{f(x^* + u) - f(x^*)}{u} = 1$;

Für $u_k = u \rightarrow 0$ mit $u < 0$ gilt: $\lim_{u \rightarrow 0} \frac{f(x^* + u) - f(x^*)}{u} = -1$.

Also ist f in $x^* = 0$ nicht differenzierbar.

Andere Formulierung

Man schreibe $f(x^* + u_k) - f(x^*)$ mit einem Rest $r(u_k)$ als

$$(*) \quad f(x^* + u_k) - f(x^*) = au_k + r(u_k).$$

Nach Division durch u_k bedeutet $Df(x^*) = a$ dann die Konvergenz

$$a_k - a = \frac{f(x^* + u_k) - f(x^*)}{u_k} - a = \frac{r(u_k)}{u_k} \rightarrow 0$$

für alle Folgen $u_k \rightarrow 0$ mit $u_k \neq 0$.

Das heißt:

2. $Df(x^*) = a \Leftrightarrow$ Es gilt $f(x^* + u) = f(x^*) + au + r(u)$ mit einer Funktion $r = r(u)$, die für $u \neq 0$ die Bedingung

$$\lim_{u \rightarrow 0} \frac{r(u)}{|u|} = 0$$

erfüllt. Auch hier muss die Limes-Beziehung für alle Folgen $u_k \rightarrow u$ gelten.

Um zu kennzeichnen, dass r eine solche Funktion ist, schreibt man oft auch $r(u) = o(u)$ und sagt, r sei eine o -Funktion.

Ebenfalls äquivalent dazu ist:
 3. $Df(x^*) = a \Leftrightarrow$ Die Differenz $f(x^* + u) - f(x^*)$ und die lineare Funktion $L(u) = au$ unterscheiden sich nur durch eine o -Funktion.

Man kann weiter die reelle Zahl a mit der linearen Funktion L , die u in $Lu = au$ überführt, identifizieren (alle linearen Funktionen von \mathbb{R} in \mathbb{R} haben offenbar diese Gestalt) und ebenso $Df(x^*)$ als lineare Funktion auffassen. Dann wird 3. zu

4. $Df(x^*) = L \Leftrightarrow$ Die Differenz $f(x^* + u) - f(x^*)$ und die lineare Funktion $L : \mathbb{R} \rightarrow \mathbb{R}$ unterscheiden sich nur durch eine Funktion $r = r(u)$ vom Typ o .

Versteht man Ableitungen wie oben als lineare Funktionen, kann man den Begriff der Differenzierbarkeit wesentlich allgemeiner fassen:

(in der Bachelor- Analysis I-Vorlesung 2010/2011 lassen wir dies zunächst aus, man kann also zu Abschnitt 6.1 springen. In Analysis II allerdings nicht.)

Dazu sei X ein linearer Vektorraum über \mathbb{R} , der mit einer Norm $\|x\|$ versehen wurden. Hier ersetzt $\|x\|$ den Betrag von x im Reellen (oder den Euklidischen Abstand vom Nullpunkt im \mathbb{R}^n) und muss erfüllen ($\forall r \in \mathbb{R}, \forall x, x_1, x_2 \in X$):

$$\begin{aligned} \|x\| \geq 0, \quad \|x\| = 0 &\Leftrightarrow x = 0, \\ \|rx\| &= |r| \cdot \|x\|, \\ \|x_1 + x_2\| &\leq \|x_1\| + \|x_2\|. \end{aligned} \tag{6.0.1}$$

In diesem Falle nennt man X einen normierten Raum (über \mathbb{R} ; analog über dem Körper der komplexen Zahlen). Mit der Vereinbarung $d(x_1, x_2) = \|x_1 - x_2\|$ wird X zu einem speziellen metrischen Raum.

Zwei Normen $\|\cdot\|_1, \|\cdot\|_2$ für X heißen *äquivalent*, wenn es positive Konstanten α, β gibt mit

$$\alpha\|x\|_1 \leq \|x\|_2 \leq \beta\|x\|_1 \quad \forall x \in X.$$

Man stelle sich im folgenden X und Y als \mathbb{R} oder als \mathbb{R}^n vor. Die betrachteten Normen sind z.B. die Euklidische $\|x\| = \sqrt{\sum_{i=1}^n x_i^2}$ für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ (oder irgendwelche anderen; s. Normäquivalenz im \mathbb{R}^n ; Anhang). In \mathbb{R} steht einfach $\|x\|$ für den Betrag $|x|$. Die linearen und stetigen Abbildungen $L : X \rightarrow Y$ sind dann die additiven und homogenen.

(Fréchet-) Differenzierbarkeit:

Sei nun $f : X \rightarrow Y$ eine Funktion zwischen normierten Räumen X und Y und $L : X \rightarrow Y$ eine lineare (und stetige) Abbildung. Dann nennt man L (Fréchet-) Ableitung von f in x^* , falls gilt

$$\begin{aligned} f(x^* + u) &= f(x^*) + L(u) + r(u) \text{ mit einer Funktion } r : X \rightarrow Y, \\ \text{die für } u \neq 0, u \in X \text{ die Bedingung } \lim_{u \rightarrow 0} \frac{\|r(u)\|_Y}{\|u\|_X} &= 0 \text{ erfüllt.} \end{aligned}$$

Bemerkungen:

- Man schreibt oft Lu statt $L(u)$.
- Der Index in der Limes-Bedingung kennzeichnet die jeweilige Norm.
- Die Normen können durch äquivalente ersetzt werden.
- der Begriff der o -Funktion wird wieder durch die Limes-Bedingung definiert.
- Die Konvergenz $u \rightarrow 0$ in X bedeutet $\lim \|u\|_X = 0$ (im Reellen).
- Die so definierte Ableitung L heißt auch Fréchet-Ableitung (hier einfach Ableitung).

Funktionen vom Typ o , äquivalente Definition: $\forall \varepsilon > 0 \exists \delta > 0$, sodass gilt:

$$\|o(u)\|_Y \leq \varepsilon \|u\|_X \quad \forall u \in X \text{ mit } 0 < \|u\|_X < \delta.$$

Der Wert $o(0)$ spielt keine Rolle in den Definitionen. Man setzt i.a. aber

$$o(0) = 0, \quad \text{damit } o \text{ im Nullpunkt stetig wird.}$$

Funktionen vom Typ o sind dann einfach diejenigen Funktionen von X in Y , deren Ableitung L im Nullpunkt existiert verschwindet; $L(u) = 0 \forall u$.

Beispiele im Reellen: $o(x) = x^2$, $o(x) = x^3$, $o(x) = |x|^{\frac{3}{2}}$, aber $p(x) = x$ ist keine o -Typ Funktion. Die Funktion o muss nicht stetig sein in $x \neq 0$. Z.B. kann $o(x) = x^2$ für rationale x und $o(x) = 0$ für irrationale sein.

Bemerkung 1:

Die Funktionen o sind also diejenigen, welche die Nullabbildung $Lu = 0_Y$ ($\forall u$) als Ableitung im Nullpunkt $x^* = 0_X$ besitzen.

Mittels der Ableitung werden die Funktionswerte von f in $x = x^* + u$ (wobei x nahe x^* ist) durch eine affin-lineare Funktion angenähert und nur Fehler von der Größenordnung $\|o(u)\|_Y$ zulassen.

Ohne $o(u)$ zu benutzen, heißt das wieder äquivalent: $\forall \varepsilon > 0 \exists \delta > 0$, sodass

$$(*) \quad \|f(x^* + u) - f(x^*) - Lu\|_Y \leq \varepsilon \|u\|_X \quad \forall u \in X \text{ mit } \|u\|_X < \delta.$$

Wie klein man $\varepsilon > 0$ auch wählt, die Differenz $f(x^* + u) - f(x^*)$ soll (in der Norm des Bildraumes) bis auf einen Fehler $\leq \varepsilon \|u\|_X$ mit Lu übereinstimmen, wenn nur $\|u\|_X$ hinreichend klein ist ($< \delta$).

In allgemeinen normierten Räumen X und Y (komplizierter als \mathbb{R} oder als \mathbb{R}^n) sind additive und homogene Abbildungen nicht notwendig stetig, dann muss man Stetigkeit voraussetzen. Zur Erinnerung an die Algebra Vorlesung:

$$\begin{aligned} \text{additiv heißt:} \quad & L(x' + x'') = L(x') + L(x'') \quad \forall x', x'' \in X \\ \text{homogen heißt:} \quad & L(rx) = rL(x) \quad \forall x \in X, r \in \mathbb{R}. \end{aligned}$$

Man beachte, dass $L(0_X) = 0_Y$ mit den Nullelementen der beiden Räume gilt.

Bemerkung 2:

Die Frage, ob eine lineare Abbildung L Ableitung von f in x^* ist, hängt nur davon ab, wie f auf einer (beliebig kleinen) Kugel um x^* definiert ist. Es ist daher auch möglich, in derselben Weise wie oben von der Differenzierbarkeit einer Funktion f in x^* zu sprechen, wenn f nur auf einer offenen Menge $\Omega \subset X$ mit $x^* \in \Omega$ erklärt ist (einer Umgebung von x^*). Die Funktion L selbst bildet stets ganz X in Y ab.

Bemerkung 3:

Das Rechnen mit Ableitungen lässt sich häufig auf die folgenden Eigenschaften von Funktionen vom o -Typ zurückführen: Sind o_1 und o_2 zwei derartige Funktionen (in entsprechenden Räumen), so auch ihre Summe bzw. die durch $o(u) = o_1(o_2(u))$ definierte Funktion. Letztere bleibt auch dann eine Funktion vom o -Typ, wenn eine der beiden gegebenen Funktionen vom o -Typ ist und die andere -etwa o_2 - nur einer (Lipschitz-) Abschätzung der Form

$$\|o_2(u)\| \leq C \|u\| \quad (\text{für normkleine } u)$$

genügt. Dies ist anhand der Beziehung (man erweitert nur einen Bruch)

$$\frac{\|o(u)\|}{\|u\|} = \left(\frac{\|o_1(o_2(u))\|}{\|o_2(u)\|} \right) \cdot \left(\frac{\|o_2(u)\|}{\|u\|} \right) \leq \frac{\|o_1(o_2(u))\|}{\|o_2(u)\|} C$$

zu sehen, wobei die Nenner jeweils verschieden von Null seien.

Andere Schreibweisen für Ableitungen: $\frac{df(x)}{dx}$, $\frac{dy}{dx}$ für $y = f(x)$ sowie $f'(x)$.

Ableitungen für $f: \mathbb{R}^n \rightarrow \mathbb{R}^m$:

Aus der Algebra ist klar, wie lineare Funktionen $L: \mathbb{R}^n \rightarrow \mathbb{R}^m$ für $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ aussehen.

Ist $m = 1$, wird L durch ein n -tupel $a = (a_1, \dots, a_n) \in \mathbb{R}^n$ definiert als $L(x) = a_1 x_1 + \dots + a_n x_n$ (also durch das natürliche Skalarprodukt von a und x).

Für $m > 1$ besteht $L(x) = (L_1(x), L_2(x), \dots, L_m(x))$ aus m Funktionen dieser Form:

$$\begin{aligned} L_1(x) &= a_{11}x_1 + \dots + a_{1n}x_n \\ L_2(x) &= a_{21}x_1 + \dots + a_{2n}x_n \\ &\vdots \\ L_m(x) &= a_{m1}x_1 + \dots + a_{mn}x_n \end{aligned}$$

Die Koeffizienten a_{ik} bilden eine (m, n) -Matrix A , und $L(x)$ ist das Produkt von A und dem Vektor x , aufgefasst als Spaltenvektor.

Für $m = 1$ wird also $Df(x^*)$ als lineare Abbildung von \mathbb{R}^n in \mathbb{R} durch n reelle Zahlen a_1, \dots, a_n , kurz durch ein Element $a \in \mathbb{R}^n$ charakterisiert, nämlich in der Weise, dass

$$Df(x^*)u = Lu = a_1u_1 + \dots + a_nu_n$$

mit den Komponenten von u gilt. Analog wird $Df(x^*)$ für $m > 1$ durch eine (m, n) -Matrix beschrieben. Die Zahlen a_i bzw. a_{ik} ergeben sich aus den später zu definierenden partiellen Ableitungen. Die affine Funktion g mit $g(x) = f(x^*) + Df(x^*)(x - x^*)$ heißt *Linearisierung* von f in x^* .

6.1 Ableitungsregeln

Grundlegende Eigenschaften von Ableitungen (s. Vorlesung/Übung)

- 1. Eindeutigkeit:** $Df(x^*)$ ist eindeutig bestimmt (muss allerdings nicht immer existieren !)
- 2. Additivität:** $D(f + g)(x^*) = Df(x^*) + Dg(x^*)$ (falls f und g beide von X in Y abbilden)
- 3. Homogenität:** $D(rf)(x^*) = rDf(x^*)$ (für reelle r)
- 4. Kettenregel:** $Df(x^*)(u) = [Dg(h(x^*)) \cdot Dh(x^*)](u) = Dg(h(x^*))(Dh(x^*)u)$ für $f(x) = g(h(x))$.

Die letzte formulieren wir wegen ihrer Wichtigkeit als Satz, der zugleich eine Aussage über die Existenz der Ableitung von h trifft.

Satz 6.3 ((Erweiterte) Kettenregel). *Es möge f mittels zweier Abbildungen $h : X \rightarrow Z$ und $g : Z \rightarrow Y$ durch $f(x) = g(h(x))$ für alle $x \in X$ erklärt sein [kurz $f = g \circ h$]. $x^* \in X$ sei fixiert. Dann gilt*

(1) *Kettenregel:*

Sind g und h in $z^ := h(x^*)$ bzw. in x^* differenzierbar mit Ableitungen L_g bzw. L_h , so ist f in x^* differenzierbar und $Df(x^*)(u) = L_g(L_h(u)) \quad \forall u \in X$.*

(2) *Inverse Abbildung:*

Seien f in x^ und g in $h(x^*)$ differenzierbar und h in x^* stetig. Ferner sei die Ableitung $L_g = Dg(h(x^*)) : Z \rightarrow Y$ bijektiv (eindeutig auf Y).*

Dann ist auch h in x^ differenzierbar, und $Dh(x^*)(u) = L_g^{-1}(L_f(u)) \quad \forall u \in X$.*

Bemerkung:

Für reelle Funktionen f, g, h bedeutet Bijektivität von L_g einfach $Dg(h(x^*)) \neq 0$. Der Abbildung $L_g(L_h)$ entspricht dem Produkt der Zahlen $Dg(h(x^*))$ und $Dh(x^*)$. Für diesen wichtigen Fall geben wir zunächst unter Satz 6.4 einen einfacheren Beweis.

Satz 6.4 ((Erweiterte) Kettenregel für reelle Funktionen $f(x) = g(h(x))$).

(1) *Sei $f(x) = g(h(x))$: Dann gilt für differenzierbare g und h : $f'(x) = g'(h(x))h'(x)$.*

(2) *Sind f und g differenzierbar (in x bzw. $h(x)$), ist weiter $g'(h(x)) \neq 0$ und h stetig, so ist auch h differenzierbar in x , und es gilt $h'(x) = \frac{f'(x)}{g'(h(x))}$.*

Beweis.

(1) Die erste Aussage ist die „normale“ Kettenregel und folgt aus dem Studium der Differenzenquotienten für $u \neq 0$, $u \rightarrow 0$ im nichttrivialen Fall $h(x+u) \neq h(x)$:

$$(1.1) \quad \frac{f(x+u) - f(x)}{u} = \frac{g(h(x+u)) - g(h(x))}{u} = \frac{g(h(x+u)) - g(h(x))}{h(x+u) - h(x)} \cdot \frac{h(x+u) - h(x)}{u}.$$

Für $u \rightarrow 0$ wissen wir wegen $h(x+u) \rightarrow h(x)$, dass gilt:

$$\lim_{u \rightarrow 0, u \neq 0} \frac{g(h(x+u)) - g(h(x))}{h(x+u) - h(x)} = g'(h(x))$$

(man setze $y^* = h(x)$, $y = h(x+u) = y^* + v$, und beachte $v \rightarrow 0$). Außerdem ist

$$\lim_{u \rightarrow 0, u \neq 0} \frac{h(x+u) - h(x)}{u} = h'(x).$$

Zusammen zeigt das die Behauptung

$$\lim_{u \rightarrow 0, u \neq 0} \frac{f(x+u) - f(x)}{u} = g'(h(x))h'(x).$$

(2) Für die zweite Aussage wissen wir, dass

$$\lim_{u \rightarrow 0, u \neq 0} \frac{f(x+u) - f(x)}{u} = f'(x)$$

und

$$\lim_{u \rightarrow 0, u \neq 0} \frac{g(h(x+u)) - g(h(x))}{h(x+u) - h(x)} = g'(h(x)) \neq 0$$

gilt, denn $h(x+u) \rightarrow h(x)$ wird durch Stetigkeit von h gesichert. Benutzen wir wieder (1.1), folgt

$$(1.2) \quad \frac{h(x+u) - h(x)}{u} = \frac{f(x+u) - f(x)}{u} \left(\frac{g(h(x+u)) - g(h(x))}{h(x+u) - h(x)} \right)^{-1}.$$

Der Ausdruck

$$\left(\frac{g(h(x+u)) - g(h(x))}{h(x+u) - h(x)} \right)^{-1} = \frac{1}{\frac{g(h(x+u)) - g(h(x))}{h(x+u) - h(x)}}$$

konvergiert nun gegen $\frac{1}{g'(h(x))}$, die gesamte rechte Seite von (1.2) also gegen

$$\frac{f'(x)}{g'(h(x))}.$$

Also konvergiert $\frac{h(x+u) - h(x)}{u}$ gegen denselben Wert, was gerade unsere Behauptung ist. □

Spezialfall der inversen Funktion h zu g (im Reellen):

Jetzt ist $h(x) = g^{-1}(x)$, also $g(h(x)) = x (= f(x))$. Damit folgt im Falle $g'(h(x)) \neq 0$:

$$h'(x) = \frac{1}{g'(h(x))}.$$

Beweis von Satz 6.3; der allgemeinen Fall:

(1) Nach Voraussetzung gilt (mit Elementen aus den entsprechenden Räumen)

$$g(z^* + w) = g(z^*) + L_g(w) + o_g(w) \quad \text{und} \quad h(x^* + u) = h(x^*) + L_h(u) + o_h(u).$$

Wir können daher mit $w := h(x^* + u) - h(x^*)$ und $L = L_g \cdot L_h$ schreiben

$$\begin{aligned} f(x^* + u) - f(x^*) &= g(h(x^* + u)) - g(h(x^*)) \\ &= L_g(h(x^* + u) - h(x^*)) + o_g(w) \\ &= L_g(L_h(u) + o_h(u)) + o_g(w) \\ &= L(u) + L_g(o_h(u)) + o_g(w) = L(u) + o(u). \end{aligned}$$

Man beachte Bemerkung 3 zu o -Type Funktionen und schreibe $o_g(w)/\|u\| = o_g(w)/\|w\| \cdot \|w\|/\|u\|$. Damit ist (1) richtig.

(2) Die inverse Funktion -wir bezeichnen sie kürzer mit I - ist definiert durch

$$I(y) = z \Leftrightarrow L_g(z) = y.$$

Da L_g linear ist, folgt dasselbe für I .

Sind nämlich $z_1 = I(y_1)$ und $z_2 = I(y_2)$, so heißt das $y_1 = L_g(z_1)$, $y_2 = L_g(z_2)$ und somit $y_1 + y_2 = L_g(z_1 + z_2)$ bzw. $z_1 + z_2 = I(y_1 + y_2)$, wonach I additiv ist. Die Homogenität bleibt als Übungsaufgabe.

Ferner ist I stetig (das gilt allgemein in Banach-Räumen wegen Banachs Inversensatz (eine nicht-triviale Aussage für Räume mit unendlicher Dimension) und ist bei endlicher Dimension einfach elementar zu zeigen).

Damit wird die durch $L(u) = I(L_f(u))$ erklärte Funktion von X in Z erneut linear.

Sie kommt also als Kandidat für $Dh(x^*)$ in Frage. Nach Voraussetzung gilt nun mit $z^* = h(x^*)$:

$$f(x^* + u) - f(x^*) = L_f(u) + o_f(u) \quad \text{und} \quad g(z^* + w) - g(z^*) = L_g(w) + o_g(w).$$

Setzen wir erneut

$$(2.1) \quad w = h(x^* + u) - h(x^*),$$

so sind die linken Seiten beider Gleichungen gleich. Also ist dann auch

$$L_f(u) + o_f(u) = L_g(w) + o_g(w).$$

Wendet man auf beide Seiten dieser Gleichung die Abbildung I an, erhält man $L(u) + I(o_f(u)) = w + I(o_g(w))$ bzw.

$$(2.2) \quad w - L(u) = I(o_f(u)) - I(o_g(w)).$$

Der Rest sind o -Abschätzungen.

Für $I(o_f(u))$ dürfen wir nach Bemerkung 3 bereits $o(u)$ schreiben. Um zu sehen, dass auch der zweite Term vom Typ $o = o(u)$ ist (womit der Satz bewiesen wäre nach Definition von w), schreiben wir die Gleichung (2.2) in der Form

$$w + I(o_g(w)) = L(u) + I(o_f(u))$$

und beachten die Stetigkeit von h in x^* . Sie sichert mittels (2.1), dass $\|w\|$ beliebig klein wird, wenn nur $\|u\|$ hinreichend klein ist. Damit wird aber auch $\|o_g(w)\|$ klein gegenüber $\|w\|$, insbesondere

$$\|o_g(w)\| \leq \frac{\|w\|}{2}.$$

Das liefert für normkleine u , etwa für $\|u\| < \delta$:

$$\begin{aligned} \|w\| &\leq \|w + I(o_g(w))\| + \|-I(o_g(w))\| \leq \|L(u) + o(u)\| + \frac{\|w\|}{2} \\ \text{und } \|w\| &\leq 2\|L(u) + o(u)\| \leq C\|u\| \text{ mit einer gewissen Konstanten } C. \end{aligned}$$

Wegen $\frac{\|o_g(w)\|}{\|u\|} = \left(\frac{\|o_g(w)\|}{\|w\|}\right) \left(\frac{\|w\|}{\|u\|}\right) \leq C \frac{\|o_g(w)\|}{\|w\|}$ wird somit $\frac{\|o_g(w)\|}{\|u\|}$ tatsächlich beliebig klein, ist also vom o -Typ, was dann auch für $I(o_g(w))$ zutreffen muss.

Zusammen mit (2.1) und (2.2) erhalten wir so die angegebenen Eigenschaften für h . □

Spezialfall der inversen Funktion h zu $g : X \rightarrow X$ (allgemein):

Wieder ist $h(x) = g^{-1}(x)$, also $g(h(x)) = x (= f(x))$. Sei g in $h(x^*)$ differenzierbar und h in x^* stetig. Ferner sei die Ableitung $L_g = Dg(h(x^*)) : X \rightarrow X$ bijektiv und L_g^{-1} ihre Inverse. Dann ist h in x^* differenzierbar, und $Dh(x^*) = L_g^{-1}$.

Weitere Regeln der Differentiation:

5. Produktregel

Sind g und h reelle und in x^* differenzierbare Funktionen, so gilt für $f(x) = g(x)h(x)$:

$$Df(x^*) = g(x^*)Dh(x^*) + h(x^*)Dg(x^*).$$

Man multipliziere aus

$$\begin{aligned} f(x^* + u) &= g(x^* + u) h(x^* + u) \\ &= (g(x^*) + L_g(u) + o_g(u)) (h(x^*) + L_h(u) + o_h(u)) \\ &= g(x^*)h(x^*) + g(x^*)L_h(u) + h(x^*)L_g(u) + \dots \end{aligned}$$

und rechne nach, dass alle restlichen Produkte \dots vom Typ $o(u)$ sind.

Dies gilt analog, wenn f und g differenzierbare Funktionen von \mathbb{R}^n in \mathbb{R} sind oder wenn beide in \mathbb{R}^m abbilden und das Produkt als Skalarprodukt $\langle f(x), g(x) \rangle$ gelesen wird.

6. Quotientenregel

Sind g und h reelle und in x^* differenzierbare Funktionen und ist $h(x^*) \neq 0$, so gilt für die (mit x nahe x^* definierte) Funktion $f(x) = \frac{g(x)}{h(x)}$:

$$Df(x^*) = \frac{h(x^*)Dg(x^*) - g(x^*)Dh(x^*)}{h(x^*)^2}.$$

Man nutze, dass nach Kettenregel die Funktion $F(x) = \frac{1}{h(x)}$ die Ableitung $F'(x) = -1 \cdot h(x)^{-2} Dh(x^*)$ besitzt und wende Produkt- und Kettenregel auf $f(x) = g(x)(h(x))^{-1}$ an:

$$Df(x^*) = (h(x^*))^{-1} Dg(x^*) + g(x^*)(-1 \cdot (h(x^*))^{-2} Dh(x^*)) = \frac{h(x^*)Dg(x^*) - g(x^*)Dh(x^*)}{h(x^*)^2}.$$

Mittels Kettenregel lassen sich weitere Ableitungen leicht berechnen:

7. Inverse Matrizen

$$A = A(x) : DA^{-1} = -A^{-1}DAA^{-1}.$$

Beweis. Es ist $A^{-1}A = AA^{-1} = E$ in jedem x . Daher gilt nach Produktregel nach Differenzieren von $E = (A(x))^{-1}A(x)$:

$$0 = D(A(x))^{-1}A(x) + (A(x))^{-1}DA(x) \quad \text{und somit} \quad D(A(x))^{-1} = -(A(x))^{-1}DA(x)(A(x))^{-1}$$

□

8. Ableitungen konkreter Funktionen

Ableitung des sinus:

$$\begin{aligned} \frac{\sin(x+h) - \sin x}{h} &= \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} = \frac{(\sin x)(\cos h - 1) + \cos x \sin h}{h} \\ &= \frac{(\sin x)(\cos h - 1)}{h} + \cos x \frac{\sin h}{h}. \end{aligned}$$

Für $h \rightarrow 0$ sind uns die Einzel-Limes aus der Übung bekannt:

$$\frac{\cos h - 1}{h} \rightarrow 0, \quad \frac{\sin h}{h} \rightarrow 1 \quad (\text{s.u.}).$$

Also ist die Ableitung von $f(x) = \sin x$ gerade $\cos x$, kurz: $(\sin x)' = \cos x$.

Die kritischen Grenzwerte für $h > 0$:

Der Kreissektor zum Winkel h hat im Einheitskreis die Fläche $S(h) = \pi \frac{h}{2\pi}$.

Die Fläche $d(h)$ des (kleineren) Dreiecks zum Kreissektor mit Winkel h ist im Einheitskreis nach der Flächenformel $F = \frac{1}{2}ab \sin \gamma$ gerade $d(h) = \frac{1}{2} \sin h < S(h)$.

Andererseits hat das mit den Katheten der Länge 1 und $\tan h$ konstruierte rechtwinklige Dreieck die Fläche $D(h) = \frac{1}{2} \tan h > S(h)$.

Damit gilt:

$$\frac{1}{2} \sin h < S(h) = \frac{1}{2} h < \frac{1}{2} \tan h$$

und Division durch $h > 0$ ergibt

$$\frac{\sin h}{h} < 1 < \frac{\tan h}{h} = \frac{\sin}{h} \cdot \frac{1}{\cos h}.$$

Wegen $\cos h \rightarrow 1$ für $h \rightarrow 0$ folgt damit die Behauptung $\frac{\sin h}{h} \rightarrow 1$.

Weiter ist

$$\frac{\cos h - 1}{h} = \frac{\cos h - \sin^2 h - \cos^2 h}{h} = \frac{\cos h - \cos^2 h}{h} - \frac{\sin^2 h}{h} = \frac{\cos h(1 - \cos h)}{h} - \frac{\sin h \sin h}{h}.$$

Hier strebt $r(h) = \frac{\sin h \sin h}{h}$ gegen Null. So folgt aus obiger Gleichung über $\frac{\cos h - 1}{h} = -\cos h \frac{\cos h - 1}{h} + r(h)$ und $\cos h \rightarrow 1$, dass $\frac{1 - \cos h}{h}$ gegen Null streben muss.

Ableitung des cosinus:

Es ist $\cos x = \sin(\frac{\pi}{2} + x)$ (das sieht man am Einheitskreis). Damit folgt nach Kettenregel:

$$(\cos x)' = \left(\sin \left(\frac{\pi}{2} + x \right) \right)' = \cos \left(\frac{\pi}{2} + x \right).$$

Aber es ist auch $\cos(\frac{\pi}{2} + x) = \sin(\pi + x) = -\sin x$. Also folgt $(\cos x)' = -\sin x$.

Ableitung des tangens:

Es ist $\tan x = \frac{\sin x}{\cos x}$ und nach Quotientenregel (für $\cos x \neq 0$):

$$(\tan x)' = \frac{\cos x \cos x - (-\sin x)(\sin x)}{(\cos x)^2} = \frac{(\cos x)^2 + (\sin x)^2}{(\cos x)^2} = 1 + (\tan x)^2 = \frac{1}{(\cos x)^2}.$$

Ableitung von $\ln x$ ($x > 0$):

Für betragskleine $h \rightarrow 0$ ist auch $x + h > 0$, sodass man wie folgt rechnen kann:

$$\frac{\ln(x+h) - \ln x}{h} = \frac{\ln\left(\frac{x+h}{x}\right)}{h} = \frac{\ln\left(1 + \frac{h}{x}\right)}{h} = \frac{1}{x} \cdot \frac{x}{h} \cdot \ln\left(1 + \frac{1}{\frac{x}{h}}\right) = \frac{1}{x} \ln\left[\left(1 + \frac{1}{\frac{x}{h}}\right)^{\frac{x}{h}}\right].$$

Wegen $\frac{x}{h} \rightarrow \infty$ für $h \rightarrow 0$, strebt das Argument $\left(1 + \frac{1}{\frac{x}{h}}\right)^{\frac{x}{h}}$ von \ln gegen e . Also wird (wegen $\ln e = 1$ und der Stetigkeit der Logarithmusfunktion) der Limes gerade $\frac{1}{x}$; kurz: $(\ln x)' = \frac{1}{x}$.

Ableitung von $\log_a x$ ($x > 0, a > 0$ fest):

Wegen $\log_a x = \frac{\ln x}{\ln a}$ ist $(\log_a x)' = \frac{(\ln x)'}{\ln a} = \frac{1}{x \ln a}$.

Die Ableitung der \ln -Funktion kann man nutzen, um weitere Ableitungen mittels der (erweiterten) Kettenregel zu bestimmen. Jetzt fungiert stets

$$\ln y = g(y) \quad \text{mit der Ableitung} \quad g'(y) = \frac{1}{y}$$

als die Funktion g in Satz 6.4. Diese Methode der Ableitungs-Berechnung wird oft auch kurz als „Logarithmische Ableitung“ bezeichnet. Es handelt sich aber um keine neue Ableitung sondern nur um die Technik, sie zu finden (also keine schöne Bezeichnung).

Kurzregel: Wir logarithmieren erst die fragliche Funktion h (die im aktuellen x positiv sein muss) und differenzieren dann.

Ableitung von $h = x^r$ ($x > 0, r$ reell):

Wir haben $\ln h(x) = \ln x^r = r \ln x$. Damit folgt nach erweiterter Kettenregel mit $f(x) = r \ln x = g(h(x))$ und $g(y) = \ln y$:

$$f'(x) = \frac{r}{x}, \quad g'(y) = \frac{1}{y} \quad \text{sowie wegen} \quad y = h(x): \quad h'(x) = \frac{\frac{r}{x}}{g'(h(x))} = \frac{\frac{r}{x}}{\frac{1}{x^r}} = rx^{r-1}.$$

Speziell $\left(\frac{1}{x}\right)' = (x^{-1})' = -x^{-2}$ und $(\sqrt{x})' = \left(x^{\frac{1}{2}}\right)' = \frac{1}{2}x^{-\frac{1}{2}} = \frac{1}{2\sqrt{x}}$.

x negativ, $r = n$ nat. Zahl:

Wenn n gerade, so $x^n = (-x)^n$. Also wird $f'(x) = -f'(|x|) = -n|x|^{n-1} = nx^{n-1}$.

Wenn n ungerade, so $x^n = -|x|^n$. Also wird $f'(x) = f'(|x|) = n|x|^{n-1} = nx^{n-1}$.

Die Formel gilt also auch für negative Argumente x , wenn n eine nat. Zahl ist. Analog argumentiert man mit dem Kehrwert bei negativen ganzen n .

Ableitung von $h = e^x$:

Wir haben $\ln h(x) = \ln e^x = x (= f(x))$. Damit folgt jetzt wie oben nach erweiterter Kettenregel

$$h'(x) = \frac{f'(x)}{g'(h(x))} = \frac{1}{\frac{1}{h(x)}} = h(x) = e^x.$$

Ableitung von $h = a^x, a > 0$:

Wir haben $\ln h(x) = \ln a^x = x \ln a (= f(x))$.

Damit folgt analog wegen $h'(x) = \frac{f'(x)}{g'(h(x))}$ und $g'(y) = \frac{1}{y}$ sowie $f'(x) = \ln a$:

$$h'(x) = \frac{\ln a}{\frac{1}{h(x)}} = h(x) \ln a = a^x \ln a$$

Anderer Weg: Wegen $\ln h(x) = x \ln a$ ist $h(x) = e^{x \ln a}$ und somit $h'(x) = (\ln a)e^{x \ln a}$.

Ableitung von $h(x) = x^x$ ($x > 0$):

Wieder betrachte man $\ln h(x) = x \ln x (= f(x))$.

Die Ableitung von f ist nach Produktregel $f'(x) = \ln x + x \frac{1}{x} = 1 + \ln x$. Es folgt so

$$h'(x) = \frac{1 + \ln x}{\frac{1}{h(x)}} = (1 + \ln x)x^x.$$

Ableitung von $h(x) = \sqrt[x]{x} = x^{\frac{1}{x}}$ ($x > 0$):

Nun ist $\ln h(x) = \frac{1}{x} \ln x (= f(x))$.

Die Ableitung von f ist nach Produktregel $\frac{-1}{x^2} \ln x + \frac{1}{x} \cdot \frac{1}{x} = \frac{1}{x^2}(1 - \ln x)$. Es folgt so

$$h'(x) = \frac{1}{x^2}(1 - \ln x) h(x) = \frac{1}{x^2}(1 - \ln x)x^{\frac{1}{x}}.$$

Ableitung von $h(x) = \exp(e^x)$: Setze $\exp(e^x) = e^u$ mit $u = e^x$.

Betrachte $f(x) = e^x = \ln h(x) = g(h(x))$, $g(y) = \ln y$, $y = h(x)$.

g' existiert und ist stetig, f' ebenfalls, $g' \neq 0$ in $y = h(x) > 0$. Also ist h differenzierbar und

$$h'(x) = \frac{f'(x)}{g'(h(x))} = \frac{e^x}{\frac{1}{h(x)}}, \text{ also } h'(x) = e^x \exp(e^x).$$

Produktregel mit log. Ableitung Sei $h(x) = u(x)v(x)$, mit positive Faktoren.

Betrachte $f(x) = \ln h(x) = \ln(u(x)v(x)) = \ln u(x) + \ln v(x)$. $f' = \frac{1}{u(x)}u'(x) + \frac{1}{v(x)}v'(x)$. Also ist f differenzierbar und

$$h'(\bar{x}) = \frac{\frac{1}{u(\bar{x})}u'(\bar{x}) + \frac{1}{v(\bar{x})}v'(\bar{x})}{1/h(\bar{x})} = u(\bar{x})v'(\bar{x}) + v(\bar{x})u'(\bar{x}).$$

Ist $u < 0, v > 0$ (nahe \bar{x}), nutze man $h = -\tilde{h} := u * (-v)$ und wende obige Regel an. Analog für $u < 0, v < 0$. Wenn $v(\bar{x}) = 0$, sagt die Produktregel dass $h'(\bar{x}) = u(\bar{x})v'(\bar{x})$. Weil die Differenzenquotienten nun die Form

$$a_n = \frac{1}{h_n} [u(\bar{x} + h_n)v(\bar{x} + h_n) - 0] = u(\bar{x} + h_n) \frac{v(\bar{x} + h_n) - v(\bar{x})}{h_n}$$

besitzen, folgt $\lim a_n = u(\bar{x})v'(\bar{x})$ tatsächlich .

6.1.1 Winkelfunktionen und ihre Arcus-Funktionen

$h(x) = \arcsin x$ (*arcus sinus*):

arcus sinus ordnet per Definition jedem $x \in [-1, 1]$ denjenigen Winkel (Bogen) $h(x) \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ zu, für welchen gilt

$$x = \sin h(x).$$

Natürlich erfüllt auch $h(x) + 2k\pi$ (k ganzzahlig) die Gleichung; die Auswahl des Intervalls $[-\frac{\pi}{2}, \frac{\pi}{2}]$ ist daher -wie auch bei den nächsten Funktionen- nicht zwingend, aber üblich.

Damit ist (für $-1 < x < 1$ und wieder nach Kettenregel mit $g = \sin$), $h'(x) = \frac{1}{\cos(h(x))}$.

Nutzt man weiter $\cos h(x) = \sqrt{1 - \sin^2 h(x)}$ sowie $\sin h(x) = x$, folgt

$$h'(x) = \frac{1}{\sqrt{1 - x^2}}, \quad -1 < x < 1.$$

$h(x) = \arccos x$ (*arcus cosinus*):

arcus cosinus ordnet per Definition jedem $x \in [-1, 1]$ denjenigen Winkel (Bogen) $h(x) \in [0, \pi]$ zu, für welchen gilt $x = \cos h(x)$. Es folgt nun analog zu oben

$$h'(x) = -\frac{1}{\sqrt{1 - x^2}} \quad -1 < x < 1.$$

$\arctan x$ (*arcus tangens*), $\operatorname{arccot} x$ (*arcus cotangens*): Ebenso ordnet $\arctan x$ reellen x einen Winkel $h(x)$ mit $\tan h(x) = x$ zu; arccot analog. $h(x)$ wird oft aus $[0, \pi]$ genommen; aus Symmetriegründen ist es günstig $\arctan x \in (-\frac{\pi}{2}, \frac{\pi}{2})$ zu wählen. Variiert φ in $(-\frac{\pi}{2}, \frac{\pi}{2})$, nimmt $\tan \varphi$ jeden reellen Wert genau einmal an (und $\cos \varphi > 0$), dasselbe gilt für $\cot \varphi$, wenn φ in $(0, \pi)$ variiert ($\sin \varphi > 0$).

Die Ableitungen berechnet man in bewährter Weise:

$$\tan h(x) = x \quad \Rightarrow \quad (1 + \tan^2 h(x))h'(x) = 1 \quad \Rightarrow \quad (1 + x^2)h'(x) = 1.$$

Also: $(\arctan x)' = \frac{1}{1+x^2}$ und ebenso $(\operatorname{arccot} x)' = -\frac{1}{1+x^2}$.

Vorsicht!

Wegen des periodischen Verhaltens der Originalfunktionen, kann man die Werte der Umkehrfunktionen auch in andere Intervalle verlegen. Benutzt man diese Funktionen etwa in Computerprogrammen, sollte man sich deshalb vergewissern, wo die zugeordneten Werte (für den Compiler) tatsächlich liegen.

6.1.2 Hyperbel-Funktionen und ihre Area-Funktionen

$\cosh x = \frac{1}{2}(e^x + e^{-x})$ (*cosinus hyperbolicus*), $\sinh x = \frac{1}{2}(e^x - e^{-x})$ (*sinus hyperbolicus*) lassen sich direkt differenzieren und besitzen die Ableitungen

$$(\cosh x)' = \sinh x \quad \text{bzw.} \quad (\sinh x)' = \cosh x$$

(ohne eine Minus, vgl. Ableitungen von \sin und \cos).

Der *tangens hyperbolicus* ist wieder durch $\tanh x = \frac{\sinh x}{\cosh x}$ erklärt, ebenso $\coth x = \frac{1}{\tanh x}$. Man rechnet direkt und leicht nach, dass

$$(\cosh x)^2 - (\sinh x)^2 = 1$$

ist (hyperbolischer Pythagoras), denn

$$(\cosh x)^2 - (\sinh x)^2 = \frac{1}{4}(e^{2x} + 2 + e^{-2x}) - \frac{1}{4}(e^{2x} - 2 + e^{-2x}) = 1.$$

So folgt

$$\begin{aligned} (\tanh x)' &= \frac{\cosh x \cosh x - \sinh x \sinh x}{(\cosh x)^2} = \frac{1}{(\cosh x)^2}, \\ (\coth x)' &= \frac{\sinh x \sinh x - \cosh x \cosh x}{(\sinh x)^2} = -\frac{1}{(\sinh x)^2}. \end{aligned}$$

Damit erhält man wie für \arcsin usw. die Ableitungen ihrer Umkehrfunktionen (Inversen) *area cosinus hyperbolicus*, *area sinus hyperbolicus* usw.

Area Funktionen:

$f(x) = \operatorname{arcosh} x$ (*area cosinus hyperbolicus*):

Sei $f(x)$ erklärt durch $y = f(x) \Leftrightarrow \cosh y = x$. f ist die Umkehrfunktion (inverse Funktion) zu \cosh . Sie ist definiert für solche x , die Lösungen y von $\cosh y = x$ zulassen, $x = \frac{1}{2}(e^y + e^{-y})$, und heißt $f(x) = \operatorname{arcosh} x$.

Für $x = 1$ gibt es genau eine Lösung, nämlich $y = 0$, für $x > 1$ genau zwei (Symmetrie).

Sei $x > 1$ und $y = f(x) > 0$. Dann gilt $\cosh f(x) = x$.
Mit den Funktionen $h = x$ und $g(y) = \cosh y$ erhalten wir:

$$g'(f(x)) = \sinh f(x) \neq 0, \quad h'(x) = 1, \quad f'(x) = \frac{1}{\sinh f(x)} = \frac{1}{\sqrt{\cosh^2 f(x) - 1}}.$$

Wegen der Definition $\cosh f(x) = x$ folgt so

$$f'(x) = \frac{1}{\sqrt{x^2 - 1}} \quad (x > 1).$$

Für später: $\int \frac{dx}{\sqrt{x^2 - 1}} = \operatorname{arcosh} x + C, (x > 1)$.

$f(x) = \operatorname{arsinh} x$ (*area sinus hyperbolicus*):

Sei $f(x)$ erklärt durch $y = f(x) \Leftrightarrow \sinh y = x$. f ist die Umkehrfunktion zu \sinh . Sie ist definiert durch

$$\sinh y = x, \quad x = \frac{1}{2}(e^y - e^{-y}),$$

und heißt $\operatorname{arsinh} x$.

Da die Ableitung $\cosh y$ von $\sinh y$ positiv ist, ist $\sinh y$ streng wachsend (Mittelwertsatz, Satz 6.11 S. 59).

Beachtet man noch $\lim_{y \rightarrow \infty} \sinh y = \infty$ und $\lim_{y \rightarrow -\infty} \sinh y = -\infty$, sieht man, dass die Gleichung genau eine (reelle) Lösung $y = f(x)$ hat.

Nun gilt $\sinh f(x) = x$.
Mit $h = x$ und $g(y) = \sinh y$ erhalten wir:

$$g'(f(x)) = \cosh f(x) > 0, \quad h'(x) = 1, \quad f'(x) = \frac{1}{\cosh f(x)} = \frac{1}{\sqrt{\sinh^2 f(x) + 1}}.$$

Wegen der Definition $\sinh f(x) = x$ folgt so

$$f'(x) = \frac{1}{\sqrt{x^2 + 1}}.$$

Für später: $\int \frac{dx}{\sqrt{x^2+1}} = \operatorname{arsinh} x + C$

Übungsaufgabe: Man bestimme die Definitionsbereiche und Ableitungen von $\tanh x = \frac{\sinh x}{\cosh x}$, $\coth x = \frac{\cosh x}{\sinh x}$ und deren Umkehrfunktionen $\operatorname{artanh} x$, $\operatorname{arcoth} x$.

6.1.3 Zusammenfassung Wichtige Ableitungen

Funktion	Ableitung	
$\operatorname{arcsin} x$	$\frac{1}{\sqrt{1-x^2}}$	$(-1 < x < 1)$
$\operatorname{arccos} x$	$-\frac{1}{\sqrt{1-x^2}}$	$(-1 < x < 1)$
$\operatorname{arctan} x$	$\frac{1}{1+x^2}$	(alle x)
$\operatorname{arccot} x$	$-\frac{1}{1+x^2}$	(alle x)
$\operatorname{arsinh} x$	$\frac{1}{\sqrt{x^2+1}}$	(alle x)
$\operatorname{arcosh} x$	$\pm \frac{1}{\sqrt{x^2-1}}$	$(x > 1)$

Da \cosh symmetrisch (bzw. gerade) ist, besteht die Umkehrfunktion wie bei \sqrt{x} aus zwei „Ästen“.

$\operatorname{artanh} x$	$\frac{1}{1-x^2}$	$(x < 1)$
$\operatorname{arcoth} x$	$\frac{1}{1-x^2}$	$(x > 1)$

Logarithmus-Darstellung der area-Funktionen:

$$\begin{aligned} \operatorname{arsinh} x &= \ln(x + \sqrt{x^2 + 1}) \quad (\text{alle } x) \\ \operatorname{arcosh} x &= \ln(x \pm \sqrt{x^2 - 1}) \quad (x > 1) \\ \operatorname{artanh} x &= \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right) \quad (|x| < 1) \\ \operatorname{arcoth} x &= \frac{1}{2} \ln\left(\frac{1+x}{x-1}\right) \quad (|x| > 1). \end{aligned}$$

Beweis für Logarithmus-Darstellung von arsinh . (Rest analog zu zeigen):

$$y = \operatorname{arsinh} x \Leftrightarrow x = \sinh y \Leftrightarrow x = \frac{1}{2}(e^y - e^{-y}).$$

Setze $u = e^y$ und berechne $u > 0$:

$$x = \frac{1}{2} \left(u - \frac{1}{u} \right), \quad 2ux = u^2 - 1, \quad u = x \pm \sqrt{x^2 + 1}, \quad y = \ln u.$$

Negative Lösung unmöglich bei $u = e^y$. Also

$$\operatorname{arsinh} x = y = \ln(x + \sqrt{x^2 + 1}).$$

□

Später brauchen wir in diesem Zusammenhang oft auch

- 1) $\int \sin^2 x dx = \frac{1}{2}(x - \sin x \cos x)$
- 2) $\int \cos^2 x dx = \frac{1}{2}(x + \sin x \cos x)$
- 3) $\int \sinh^2 x dx = \frac{1}{2}(\sinh x \cosh x - x)$
- 4) $\int \cosh^2 x dx = \frac{1}{2}(\sinh x \cosh x + x)$

Die Integrale 1) - 4) werden später mit partieller Integration berechnet, sie sind hier der besseren Übersicht halber schon aufgeführt.

6.2 Höhere und partielle Ableitungen

Höhere Ableitungen. Im Reellen ist die Ableitung von f zum Argument x als Funktion deutbar, die x jeweils die reelle Zahl $g(x) := Df(x)$ zuordnet. Ist g erneut differenzierbar (dazu muss $Df(x)$ für alle x aus einem offenen Intervall existieren), heißt ihre Ableitung $Dg(x)$ auch zweite Ableitung von f in x : Kurz $Dg(x) = D^2f(x)$ bzw. $\frac{dg}{dx} = \frac{d^2f}{dx^2}$ bzw. $Dg(x) = f''(x) = f^{(2)}(x)$. Analog bildet man durch weiteres Differenzieren n -te Ableitungen $f^{(n)}(x)$ von f .

Weiter kann $Df(x)$ für eine Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ als diejenige Matrix A vom Typ (m, n) aufgefasst werden, die die lineare Abbildung $Df(x)$ eindeutig beschreibt.

Ist f auf einer offenen Teilmenge Ω des \mathbb{R}^n differenzierbar, können wir $g = Df$ als eine Abbildung von Ω in die Menge der (m, n) -Matrizen auffassen, also $g : \Omega \rightarrow \mathbb{R}^k$, $k = mn$.

Mittels Dg definiert man dann in gleicher Weise die zweite Ableitung von f .

Man beachte, dass die zugeordneten „Matrizen“ dann eine höhere Dimension bekommen:

Falls $f : \mathbb{R}^n \rightarrow \mathbb{R}^1$, so ist $Df(x) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ durch n Zahlen, also einen Punkt des \mathbb{R}^n definiert. Damit bildet g die Teilmenge $\Omega \subset \mathbb{R}^n$ in den Raum \mathbb{R}^n ab. Die Ableitung Dg von g wird folglich durch eine (n, n) -Matrix charakterisiert.

Partielle Ableitungen. Es sei erneut $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Wir stellen nun die Frage, wie für ein fixiertes x die Differenzierbarkeit von f in x festgestellt werden kann und wie die n Komponenten der Ableitung $Df(x)$ dann zu bestimmen sind.

Dazu halten wir einen Index $i \in \{1, \dots, n\}$ fest und betrachten Punkte der Form

$$(x_1, \dots, x_{i-1}, x_i + h, x_{i+1}, \dots, x_n),$$

die wir abkürzend mit $x(h)$ bezeichnen ($h \in \mathbb{R}$).

Definition 6.5 (partiell differenzierbar, partielle Ableitung, Gradient). *Gilt $\lim_{h \rightarrow 0} \frac{f(x(h)) - f(x)}{h} = L_i$ ($\in \mathbb{R}$) für jede Nullfolge h , so heißt f partiell differenzierbar bzgl. x_i ; der Grenzwert L_i heißt partielle Ableitung von f bzgl. x_i . Er wird gewöhnlich mit $\frac{\partial f(x)}{\partial x_i}$ bezeichnet. Existieren alle n partiellen Ableitungen, wird f partiell differenzierbar genannt. Der Vektor der partiellen Ableitungen heißt auch Gradient von f (im aktuellen Punkt).*

Beispiel 6.6. $f(x_1, x_2) = x_1 \cos x_2$: $\frac{\partial f(x)}{\partial x_1} = \cos x_2$; $\frac{\partial f(x)}{\partial x_2} = -x_1 \sin x_2$.

Aus der Definition der Differenzierbarkeit von f folgt unmittelbar: Existenz der Ableitung $Df(x) \Rightarrow$ Existenz der partiellen Ableitungen $\frac{\partial f(x)}{\partial x_i}$, bzw. differenzierbar \Rightarrow partiell differenzierbar, und die lineare Abbildung $Df(x)$ wird durch den Gradienten definiert:

$$Df(x)(u) = \sum_{i=1}^n \frac{\partial f(x)}{\partial x_i} u_i \quad (\text{bei Differenzierbarkeit})$$

Um dies zu sehen, betrachte man die Linearisierung $g(x+u) = f(x) + Df(x)(u)$ in Punkten der Form $x+u = x(h)$.

Aber: Aus der Existenz des Gradienten folgt noch nicht die Differenzierbarkeit!

Beispiel 6.7 (Gegenbeispiel). *Seien $x, y \in \mathbb{R}$: $f(x, y) = \min\{|x|, |y|\}$ für $x \geq 0$ und $y \geq 0$, $f(x, y) = 0$ sonst.*

Die Funktion ist auf den Achsen Null. Für $(x^, y^*) = (0, 0)$ ist der Gradient der Nullvektor des \mathbb{R}^2 , doch f ist in (x^*, y^*) nicht differenzierbar (Man betrachte f in Punkten (t, t)).*

Der Zusammenhang wird einfacher, wenn partielle Ableitungen überall existieren und stetig sind.

Satz 6.8 (partiell differenzierbar + Stetigkeit der partiellen Ableitungen \Rightarrow differenzierbar). *Es sei $f : \mathbb{R}^n \rightarrow \mathbb{R}$ auf einer offenen Teilmenge Ω des \mathbb{R}^n partiell differenzierbar, und die partiellen Ableitungen $\frac{\partial f(x)}{\partial x_i}$ seien (als Funktionen von Ω in \mathbb{R}) dort stetig. Dann ist f auf Ω differenzierbar.*

Beweis. Wir betrachten den Fall $n = 2$, der Beweis für größere n läuft analog. Seien $(x^*, y^*) \in \Omega$ und $(u, v) \in \mathbb{R}^2$. Wir setzen

$$Df(x^*, y^*)(u, v) = \frac{\partial f(x^*, y^*)}{\partial x} u + \frac{\partial f(x^*, y^*)}{\partial y} v.$$

(Trick: Zerlegen der Differenzen)

Dann gilt nach Anwenden des (reellen) Mittelwertsatzes mit gewissen q_1 und q_2 in $(0, 1)$:

$$\begin{aligned} f(x^* + u, y^* + v) - f(x^*, y^*) &= [f(x^* + u, y^* + v) - f(x^* + u, y^*)] + [f(x^* + u, y^*) - f(x^*, y^*)] \\ &= \frac{\partial f(x^* + u, y^* + q_1 v)}{\partial y} v + \frac{\partial f(x^* + q_2 u, y^*)}{\partial x} u \\ &= Df(x^*, y^*)(u, v) \\ &\quad + \left[\left(\frac{\partial f(x^* + q_2 u, y^*)}{\partial x} + \frac{\partial f(x^* + u, y^* + q_1 v)}{\partial y} \right) - Df(x^*, y^*) \right] (u, v). \end{aligned}$$

Der Ausdruck $r(u, v)$ in der letzten Zeile ist das Skalarprodukt des Vektors (u, v) mit einem Vektor $W(u, v, q_1, q_2)$. Da $q_1, q_2 \in (0, 1)$ und die partiellen Ableitungen nach Voraussetzung stetig sind, strebt W gegen Null für $(u, v) \rightarrow 0$. Damit ist dieser Ausdruck vom Typ $o(u, v)$, was den Satz beweist. \square

Bemerkung: Die aus der Gestalt der Ableitung

$$Df(x^*, y^*)(u, v) = \frac{\partial f(x^*, y^*)}{\partial x} u + \frac{\partial f(x^*, y^*)}{\partial y} v$$

und aus

$$f(x^* + u, y^* + v) = f(x^*, y^*) + \frac{\partial f(x^*, y^*)}{\partial x} u + \frac{\partial f(x^*, y^*)}{\partial y} v + o(u, v)$$

resultierende Veränderung von f nahe (x^*, y^*) wird formal auch als

$$df = \frac{\partial f(x^*, y^*)}{\partial x} dx + \frac{\partial f(x^*, y^*)}{\partial y} dy$$

geschrieben.

Die rechte Seite heißt dann vollständiges Differential von f . (Analog bei mehr Variablen).

Beispiel 6.9. $f(x, y) = x^y$ ($x > 0$).

$\frac{\partial f}{\partial x} = yx^{y-1}$ (man betrachtet y als Konstante, differenziert bzgl. x),

$\frac{\partial f}{\partial y} = x^y \ln x$ (man betrachtet x als Konstante, differenziert bzgl. y).

Mit Ableitungen in $(2, 3)$ weiß man so wegen $\frac{\partial f(2,3)}{\partial x} = 3 \cdot 2^2$ und $\frac{\partial f(2,3)}{\partial y} = 2^3 \ln 2$, dass

$$f(2 + u, 3 + v) = f(2, 3) + 12u + (8 \ln 2)v + r = 8 + 12u + (8 \ln 2)v + r$$

ist, mit einem Rest r , der mit jedem $\varepsilon > 0$ für hinreichend kleine $|u|$ und $|v|$ im Betrag nicht größer als $\varepsilon \sqrt{u^2 + v^2}$ ist, denn $\frac{|r|}{\sqrt{u^2 + v^2}}$ strebt gegen Null für $(u, v) \rightarrow (0, 0)$.

Grundlegend für die Anwendung von Ableitungen sind die folgenden Sätze.

Satz 6.10 (Extremalpunkte). *Es seien $f : X \rightarrow \mathbb{R}$ eine in x^* differenzierbare Funktion, X ein normierter Raum, und es gelte $f(x) \geq f(x^*)$ für alle x in einer gewissen Kugel $B(x^*, \varepsilon)$ (d.h. x^* ist ein lokaler Minimalpunkt von f). Dann ist $Df(x^*) = 0$ (die Nullabbildung).*

Beweis. Es gilt $f(x^* + u) = f(x^*) + Df(x^*)(u) + o(u) \geq f(x^*)$ für normkleine u .

Damit ist $Df(x^*)(u) + o(u) \geq 0$.

Setzt man insbesondere $u = rv$ ($v \in X$ beliebig fixiert, $r \downarrow 0$) und dividiert durch r , so folgt $Df(x^*)(v) + \frac{o(rv)}{r} \geq 0$, und wegen $\frac{o(rv)}{r} \rightarrow 0$ muss gelten $Df(x^*)(v) \geq 0$. Da v beliebig aus X war, die Ungleichung also auch für $-v$ gilt, ist $Df(x^*)(v) = 0$. \square

Satz 6.11 (Mittelwertsatz der Diff.-Rechnung, reell). *Es seien $f : \mathbb{R} \rightarrow \mathbb{R}$ eine differenzierbare Funktion und $a < b$. Dann existiert ein $\theta \in (a, b)$, sodass*

$$f(b) - f(a) = Df(\theta)(b - a).$$

Beweis. Wir bilden die Funktion $g(x) = f(x) - f(a) - \frac{(f(b)-f(a))(x-a)}{b-a}$.

Sie erfüllt $g(a) = g(b) = 0$ und $Dg(x) = Df(x) - \frac{f(b)-f(a)}{b-a}$.

Es ist zu zeigen, dass $Dg(\theta) = 0$ für ein $\theta \in (a, b)$ gilt.

Ist g konstant auf $[a, b]$, so verschwindet in $\theta = \frac{a+b}{2}$ die Ableitung.

Ist das Minimum von g auf $[a, b]$ negativ, so ist jeder Minimalpunkt θ innerer Punkt des Intervalls und erfüllt $Dg(\theta) = 0$, (Satz 6.10). Andernfalls ist das Maximum von g auf $[a, b]$ positiv, und wir erhalten $Dg(\theta) = 0$ für jeden Maximalpunkt. \square

Bemerkung: Der Satz kann auch direkt mittels Satz über die Taylor-Entwicklung bewiesen werden, s. Bemerkung und entsprechenden Beweis dazu.

Satz 6.12 (Mittelwertsatz für \mathbb{R}^n). *Es seien $f : \mathbb{R}^n \rightarrow \mathbb{R}$ differenzierbar und $x, y \in \mathbb{R}^n$.*

Dann gibt es eine Zahl $\theta \in (0, 1)$, sodass $f(y) - f(x) = Df(x + \theta(y-x))(y-x)$.

Beweis. Wir wenden Satz 6.11 auf die reelle Funktion $g(t) = f(x + t(y-x))$ in den Punkten $a = 0$ und $b = 1$ an. Offenbar ist $g(a) = f(x)$ und $g(b) = f(y)$. Über die Kettenregel folgt

$$Dg(t) = Df((x + t(y-x))(y-x),$$

was die Behauptung liefert. \square

Bemerkung: Wie im Reellen liegt der fragliche Punkt $z = x + \theta(y-x)$ im offenen Intervall zwischen x und y (das jetzt eine Strecke im \mathbb{R}^n wird). Oft kennt man z nicht explizit, kann aber die Kenntnis einer Abschätzung der Form $\|Df(z)\| \leq C$ für diese z ausnutzen, um

$$|f(y) - f(x)| = |Df(x + \theta(y-x))(y-x)| \leq C\|y-x\|$$

abzuschätzen.

6.3 Taylor-Reihen (Brook Taylor *18.8.1685)

Sei $f : \mathbb{R} \rightarrow \mathbb{R}$ eine hinreichend oft differenzierbare Funktion und $a \in \mathbb{R}$ fixiert. Dann kann die Differenz $f(x) - f(a)$ genauer als durch den Mittelwertsatz abgeschätzt werden, wenn man höhere Ableitungen benutzt.

Satz 6.13 (Taylor). *Existiert überall die k -te Ableitung $f^{(k)}$ von f für $k \leq n+1$, so gilt*

$$(1) \quad f(x) = f(a) + f'(a) \frac{(x-a)^1}{1!} + \dots + f^{(n)}(a) \frac{(x-a)^n}{n!} + R_n(x)$$

wobei das Restglied $R_n(x)$ mit einem geeigneten Θ zwischen a und x die Form

$$(2) \quad R_n(x) = f^{(n+1)}(\Theta) \frac{(x-a)^{n+1}}{(n+1)!}$$

hat.

Bemerkung: Die obige Darstellung von f heißt Taylorentwicklung n -ter Ordnung von f . Die angegebene Darstellung von $R_n(x)$ ist die Restglied-Darstellung von J.-L. Lagrange (*1736). Ist f beliebig oft differenzierbar, so heißt die unendliche Reihe (Potenzreihe)

$$f(a) + f'(a) \frac{(x-a)^1}{1!} + \dots + f^{(n)}(a) \frac{(x-a)^n}{n!} + \dots$$

Taylorreihe für f in a .

Der Mittelwertsatz der Differentialrechnung ist nun der Taylor-Satz mit $n = 0$.

Beweis des Taylorschen Satzes. Seien a und x beliebig fixiert und sei die Konstante $R_n = R_n(x)$ durch Formel (1) definiert als Differenz zwischen $f(x)$ und der Partialsumme. Wir zeigen, dass R_n sich tatsächlich in der Form (2) schreiben lässt.

Trick: Wir definieren hierzu eine Funktion g durch

$$g(y) = f(y) + f'(y) \frac{(x-y)^1}{1!} + \dots + f^{(k)}(y) \frac{(x-y)^k}{k!} + \dots + f^{(n)}(y) \frac{(x-y)^n}{n!} + \frac{(x-y)^{n+1}}{(x-a)^{n+1}} R_n$$

y ersetzt a in der n -ten Partialsumme. Dazu noch $\frac{(x-y)^{n+1}}{(x-a)^{n+1}} R_n$ addieren.

Offenbar ist g differenzierbar.

Für $y = x$ gilt $g(y) = f(x)$. Für $y = a$ folgt aus (1) ebenfalls $g(y) = f(x)$. Also ist $g(x) = g(a)$.

Die stetige Funktion g hat im offenen Intervall (a, x) bzw. (x, a) einen globalen Maximalpunkt oder Minimalpunkt Θ . Daher verschwindet dort ihre (existierende) Ableitung: $g'(\Theta) = 0$ (Das ist dasselbe Argument wie beim Beweis des Mittelwertsatzes, den man auch direkt benutzen könnte).

Wir rechnen die Ableitung von g aus.

Mittels Produktregel für die (y -)Ableitung der Terme $f^{(k)}(y) \frac{(x-y)^k}{k!}$ folgt

$$\begin{aligned} \left(f^{(k)}(y) \frac{(x-y)^k}{k!} \right)' &= f^{(k+1)}(y) \frac{(x-y)^k}{k!} - k f^{(k)}(y) \frac{(x-y)^{k-1}}{k!} \\ &= f^{(k+1)}(y) \frac{(x-y)^k}{k!} - f^{(k)}(y) \frac{(x-y)^{k-1}}{(k-1)!} \quad (k \leq n) \end{aligned}$$

und

$$\left(f^{(k+1)}(y) \frac{(x-y)^{k+1}}{(k+1)!} \right)' = f^{(k+2)}(y) \frac{(x-y)^{k+1}}{(k+1)!} - f^{(k+1)}(y) \frac{(x-y)^k}{k!} \quad (k < n)$$

Bei der Summenbildung über alle k heben sich hier alle Ausdrücke mit positivem und negativem Vorzeichen weg bis auf den letzten Ausdruck $f^{(n+1)}(y) \frac{(x-y)^n}{n!}$.

Mit der Ableitung von $\frac{(x-y)^{n+1}}{(x-a)^{n+1}} R_n$ bzgl. y ist daher

$$g'(y) = f^{(n+1)}(y) \frac{(x-y)^n}{n!} - \frac{(n+1)(x-y)^n}{(x-a)^{n+1}} R_n$$

Für $y = \Theta$ folgt so wegen $g'(\Theta) = 0$,

$$f^{(n+1)}(y) \frac{(x-y)^n}{n!} = \frac{(n+1)(x-y)^n}{(x-a)^{n+1}} R_n.$$

Beachtet man $\Theta \neq x$ und löst nach R_n auf, erhält man nun direkt die Behauptung (2). \square

Die Taylorreihe (als unendliche Reihe betrachtet) ist eine spezielle Potenzreihe.

Sie kann (muss aber nicht) konvergieren. Ihr Limes -falls er für ein x existiert- muss auch nicht mit $f(x)$ zusammenfallen. Das alles hängt vom Verhalten des Restgliedes $R_n(x)$ für $n \rightarrow \infty$ ab (bzw. von den höheren Ableitungen $f^{(n+1)}(\delta)$).

Für das Restglied gibt es verschiedene Darstellungen, u.a. auch in Integralform (haben wir noch nicht). Die aktuelle ist die Darstellung nach Lagrange.

Es sei nun x auf ein Intervall der Form $X = (a-c, a+c)$ (mit $c > 0$) beschränkt. Wir erhalten dann direkt aus dem bewiesenen Satz.

Folgerung (Gleichmäßige Konvergenz der Taylorreihe)

Gilt eine Abschätzung der Form $|R_n(x)| < \varepsilon(n)$ für alle $x \in X$, mit $\varepsilon(n) \rightarrow 0$ für wachsendes n , so konvergiert die Taylorreihe mit jedem x aus X , und ihr Wert fällt mit $f(x)$ zusammen. Für die n -ten Partialsummen $F_n(x)$ der Reihe gilt dann

$$|f(x) - F_n(x)| < \varepsilon(n) \quad (\forall x \in X).$$

In diesem Falle sagt man, die Reihe (die Funktionenfolge F_n) konvergiere gleichmäßig bzgl. $x \in X$ gegen die Funktion $f = f(x)$ bzw. F_n konvergiert glm. gegen f .

Gleichmäßige Konvergenz einer Funktionenfolge F_n gegen f (Funktionen von X in \mathbb{R}) bedeutet also die Konvergenz der Zahlenfolge $s_n = \sup_{x \in X} |f(x) - F_n(x)|$ gegen Null.

Für Funktionen mit Bildern in einem metrischen Raum Y ersetzt man einfach $|f(x) - F_n(x)|$ durch $d_Y(f(x), F_n(x))$.

Achtung! Auch wenn die Taylorreihe einer Funktion f stets konvergieren sollte, muss ihr Wert nicht mit $f(x)$ zusammenfallen.

Beispiel 6.14. Man betrachte die Funktion $f(x) = \exp(-x^{-2})$ für $x \neq 0$ und $f(0) = 0$.

Dann ist f beliebig oft differenzierbar, und es sind alle Ableitungen (in $x = 0$) Null.

Der Wert der mit $a = 0$ gebildeten Taylorreihe ist damit für jedes x gleich Null, im Gegensatz zu $f(x) > 0$.

Beweis. Wenn $x \neq 0$ so: $f' = 2x^{-3} \exp(-x^{-2})$, $f'' = -6x^{-4} \exp(-x^{-2}) + 2x^{-3}(2x^{-3} \exp(-x^{-2})) = (-6x^{-4} + 4x^{-6}) \exp(-x^{-2})$.

Allgemein: wenn eine Ableitung schon die Form $f^{(n)} = p(x) \exp(-x^{-2})$ mit p = "Polynom" mit negativen Exponenten, besitzt, folgt dasselbe für

$$f^{(n+1)} = p'(x) \exp(-x^{-2}) + 2x^{-3} p(x) \exp(-x^{-2}) = (p'(x) + 2x^{-3} p(x)) \exp(-x^{-2}) = q(x) \exp(-x^{-2}).$$

Jeder Ausdruck der Form $a(x) = x^{-k} \exp(-x^{-2}) = \left(\frac{1}{x}\right)^k \cdot \frac{1}{\exp\left(\left(\frac{1}{x}\right)^2\right)}$ strebt bei $x \rightarrow 0$ gegen Null. Mit $y = 1/x \rightarrow \infty$ folgt naemlich $|a(x)| = |y^k|/e^{y^2} < |y^k|/e^y \rightarrow 0$.

Also gilt auch $f^{(n)}(x)$ für $x \rightarrow 0$ und jedem festen n . Außerdem ist analog auch

$$\lim_{x \rightarrow 0} \frac{f^{(n)}(x)}{x} = 0 \quad (x \neq 0).$$

Nutzt man das mit $n = 0, 1, 2, \dots$ folgt so die Existenz der Ableitungen und $f^{(n)}(0) = 0$. □

Differenzieren (und glm. Konvergenz) von Potenzreihen innerhalb des Konvergenzradius
Beschreibt die Taylorreihe (als wichtigste Potenzreihe) die Funktion f für $|x - a| < r$, kann man nach der Reihendarstellung von f' fragen. Solange x im Innern des Konvergenzintervalls der f -Reihe liegt, gibt es keine Probleme, denn die durch gliedweises Differenzieren gewonnene Reihe erweist sich als stetig im Argument x und fällt mit $f'(x)$ zusammen. Dies sichert der folgende

Satz 6.15 (Ableitung einer Potenzreihe). *Eine Reihe*

$$(1) \quad \sum_{k \geq 0} a_k x^k$$

konvergiere für alle x mit $|x| < r$ jeweils gegen ein $f(x)$.

1. Dann ist für $|x| \leq s < r$ die Konvergenz gleichmäßig, d.h. die Partialsummen $S_n(x)$ der Reihe erfüllen

$$(*) \quad |f(x) - S_n(x)| \leq \varepsilon(n) \quad \forall x : |x| \leq s \text{ mit gewissen } \varepsilon(n) \rightarrow 0 \text{ für } n \rightarrow \infty.$$

Außerdem ist f auf $[-s, s]$ stetig.

2. Die durch Differenzieren der Summanden gewonnene Reihe

$$(2) \quad \sum_{k \geq 1} k a_k x^{k-1},$$

konvergiert ebenfalls gleichmäßig in Bezug auf $|x| \leq s$, und zwar gegen $g(x) = f'(x)$.

3. Die Konvergenzintervalle beider Reihen sind gleich.

Folgerung 6.16 (aus 3.). *Man kann auch erneut differenzieren. Man kann außerdem beide Reihen unter 1. und 2. vertauschen. So erhält man dieselbe Aussage für das (später definierte) Integral einer Potenzreihe.*

Beweis. 1. Für alle $|x| \leq s < r$ und große n gilt eine Abschätzung der Form $|a_k x^k| \leq q^k$ mit festem $q \in (0, 1)$ (siehe Potenzreihen), die (*) und zugleich absolute Konvergenz der Reihe sichert.

Stetigkeit von f : Sei $x \in [-s, s]$ fixiert, $\varepsilon > 0$ und $y \in [-s, s]$. Dann gilt

$$\begin{aligned} |f(y) - f(x)| &\leq \left| S_n(y) - S_n(x) + \sum_{k > n} a_k y^k - \sum_{k > n} a_k x^k \right| \\ &\leq |S_n(y) - S_n(x)| + \left| \sum_{k > n} a_k y^k \right| + \left| \sum_{k > n} a_k x^k \right|. \end{aligned}$$

Für große n sind die beiden letzten Summen jeweils kleiner als $\frac{\varepsilon}{3}$. Wir halten ein solches n fest und nutzen Stetigkeit der endlichen Summen S_n . Dann folgt auch $|S_n(y) - S_n(x)| < \frac{\varepsilon}{3}$, wenn $|y - x|$ hinreichend klein ist.

Zusammen ergibt dies $|f(y) - f(x)| \leq \varepsilon$. Also ist f stetig als Funktion von $[-s, s]$ in \mathbb{R} . Wenn man nicht $y \in [-s, s]$ fordert, folgt bei kleinem $|y - x|$ immer noch $|y| < s' < r$ mit einem $s' > s$. Der Schluss lässt sich also mit s' statt s wiederholen, wonach die Stetigkeit von f auf $[-s, s]$ folgt.

2. Mittels Wurzelkriterium sieht man, dass das Konvergenzintervall beider Reihen gleich bleibt. $|ka_k|^{1/k-1} = k^{1/k-1}|a_k|^{1/k-1}$ wobei $\lim_{k \rightarrow \infty} k^{1/k-1} = 1$ und $|a_k|^{1/k} = (|a_k|^{1/k-1})^{k-1}$.

In beiden Fällen ist dieselbe Größe $l_s := \limsup_{k \rightarrow \infty} |a_k|^{1/k-1} = \limsup_{k \rightarrow \infty} |a_k|^{1/k}$ für das Konvergenzintervall entscheidend ($r = \frac{1}{l_s}$ wissen wir schon). Also konvergiert die zweite Reihe gegen eine auf $[-s, s]$ stetige Funktion $g(x)$.

Es bleibt zu zeigen, dass $g(x) = f'(x)$ gilt.

Die konvergenten Partialsummen σ_n von (2) erfüllen $\sigma_n(x) = S'_n(x)$.

Sei $\varepsilon > 0$ beliebig fixiert und $|h| > 0$ klein genug, sodass $|g(\theta) - g(x)| < \varepsilon$ für alle θ zwischen x und $x+h$.

Wir approximieren die Funktionswerte $f(x+h) - f(x)$ durch die Partialsummen S_n .

$$f(x+h) - f(x) = S_n(x+h) - S_n(x) + r_n.$$

Dabei können wir wegen der glm. Konvergenz abschätzen:

$|S_n(x+h) - f(x+h)| < \delta(n)$ und $|S_n(x) - f(x)| < \delta(n)$, wobei $\delta(n) \rightarrow 0$ für $n \rightarrow \infty$ also $|r_n| < 2\delta(n)$.

Es folgt

$$\frac{f(x+h) - f(x)}{h} = \frac{S_n(x+h) - S_n(x)}{h} + \frac{r_n}{h}.$$

Für grosse $n > n(h, \varepsilon)$ gilt weiter, dass $w_n = \frac{r_n}{h}$ die Ungleichung $|w_n| < \varepsilon$ erfüllt, und nach Mittelwertsatz 6.11 folgt $S_n(x+h) - S_n(x) = S'_n(\theta)h = \sigma_n(\theta)h$, also auch $\frac{f(x+h) - f(x)}{h} = \sigma_n(\theta) + w_n$ mit θ zwischen x und $x+h$.

Schließlich können wir nutzen, dass $|\sigma_n(\theta) - g(\theta)| < \varepsilon$ und $|g(\theta) - g(x)| < \varepsilon$ sind. Wir erhalten so: $\frac{f(x+h) - f(x)}{h} = g(x) + W$ mit $|W| < 3\varepsilon$ für betragskleine h .

Da $\varepsilon > 0$ beliebig war, folgt so die Existenz und Eindeutigkeit des Limes

$$\lim_{h \rightarrow 0, h \neq 0} \frac{f(x+h) - f(x)}{h} = g(x).$$

Damit existiert $f'(x)$ und fällt mit $g(x)$ zusammen. □

Folgerung 6.17. Die Taylorreihe

$$(1) \quad f(a) + f'(a) \frac{(x-a)^1}{1!} + \dots + f^{(n)}(a) \frac{(x-a)^n}{n!} + \dots$$

konvergiere gegen $f(x)$ für jedes x mit $|x-a| < r$. Sei $0 < s < r$.

1. Dann erfolgt für $|x-a| \leq s$ die Konvergenz gleichmäßig, d.h. die Partialsummen $S_n(x)$ der Reihe erfüllen $|f(x) - S_n(x)| \leq \varepsilon(n)$ für alle x mit $|x-a| \leq s$ und mit demselben $\varepsilon(n) \rightarrow 0$ für $n \rightarrow \infty$.
2. Weiter konvergiert die durch gliedweises Differenzieren gewonnene Reihe

$$(2) \quad 0 + f'(a) \frac{(x-a)^0}{1!} + 2f''(a) \frac{(x-a)^1}{2!} + \dots + nf^{(n)}(a) \frac{(x-a)^{n-1}}{n!} + \dots$$

ebenfalls gleichmäßig in Bezug auf $|x-a| \leq s$ gegen $g(x) = f'(x)$.

Beweis. Man muss nur $a_k = \frac{f^{(k)}(a)}{k!}$ setzen und $x-a$ durch eine neue Variable ξ ersetzen. □

6.3.1 Taylor-Entwicklung für Funktionen $f: \mathbb{R}^m \rightarrow \mathbb{R}$

Für $f: \mathbb{R}^m \rightarrow \mathbb{R}$ kann man $f(b) - f(a)$ abschätzen, indem man die reelle Funktion $g(t) = f(a + t(b-a))$ mit $g(0) = f(a)$ und $g(1) = f(b)$ betrachtet ($t \in \mathbb{R}$).

Zur Abkürzung setzen wir $b-a = v$.

Die Funktion g hat die Form $g(t) = f(u(t))$ wobei $u: \mathbb{R} \rightarrow \mathbb{R}^m$ affin linear ist; $u(t) = a + tv$ mit der konstanten Ableitung $Du(t) = v$.

Die Kettenregel liefert deshalb bei Differenzierbarkeit von f für jedes feste t :

$$(k1) \quad g'(t) = Df(u(t))Du(t),$$

das ist das Skalarprodukt

$$(s1) \quad g'(t) = \sum_i \frac{\partial f(u(t))}{\partial x_i} v_i \text{ von } Df(u(t)) \rightarrow \mathbb{R}^m \text{ und } v.$$

Es wird oft geschrieben als

$$g'(t) = \langle Df(a + tv), v \rangle \quad (\text{kanonisches Skalarprodukt im } \mathbb{R}^m.)$$

Für jedes feste i ist $\frac{\partial f}{\partial x_i}$ eine Abbildung von \mathbb{R}^m in \mathbb{R} .

Bei Stetigkeit der zweiten partiellen Ableitungen (oder äquivalent bei stetiger Differenzierbarkeit von $Df : \mathbb{R}^m \rightarrow \mathbb{R}^m$) ist auch die von t abhängige Funktion

$$Df(a + tv) = Df(u(t))$$

mit Werten im \mathbb{R}^m stetig differenzierbar. Sie ist entscheidend für $g''(t)$ und hat explizit die Form

$$Df(u(t)) = (F_1(t), F_2(t), \dots, F_m(t))$$

mit $F_i(t) = \frac{\partial f(u(t))}{\partial x_i} \in \mathbb{R}$.

Die Kettenregel angewandt auf jede Komponente F_i liefert nun, ebenso wie in (k1),

$$(k2) \quad F'_i(t) = D \left(\frac{\partial f(u(t))}{\partial x_i} \right) Du(t) = D \left(\frac{\partial f(u(t))}{\partial x_i} \right) v \quad \text{für jedes feste } i.$$

Das ist nun das Skalarprodukt (wir müssen den Lauf-Index ändern, weil i jetzt fest ist)

$$(s2) \quad \sum_k \frac{\partial}{\partial x_k} \frac{\partial f(u(t))}{\partial x_i} v_k,$$

was mittels der zweiten partiellen Ableitungen von f zu

$$(s3) \quad F'_i(t) = \sum_k \frac{\partial^2 f(u(t))}{\partial x_i \partial x_k} v_k \quad \text{für jedes feste } i \text{ wird.}$$

Somit wird

$$g''(t) = \sum_i \left[\sum_k \frac{\partial^2 f(u(t))}{\partial x_i \partial x_k} v_k \right] v_i.$$

In Matrixschreibweise (mit v als Spaltenvektor) bedeutet das gerade

$$(s4) \quad g''(t) = \langle D^2 f(a + tv)v, v \rangle = v^T D^2 f(a + tv)v.$$

Dabei ist die Matrix $D^2 f = \frac{\partial^2 f}{\partial x_i \partial x_k}$ in $a + tv$ gegeben.

Diese Matrix ist (nach einem Satz von Schwarz 8.11, Beweis mittels Integralrechnung später) symmetrisch, d. h. $\frac{\partial^2 f(u(t))}{\partial x_i \partial x_k} = \frac{\partial^2 f(u(t))}{\partial x_k \partial x_i}$, sofern die 2. Ableitungen stetig vom Argument abhängen. Taylorentwicklung für g liefert unser Ergebnis (mit $x = b$):

$$\begin{aligned} f(x) &= f(a) + \langle Df(a), x - a \rangle + R_1(x), \quad \text{wobei} \\ R_1(x) &= \frac{1}{2} \langle D^2 f(a + \theta(x - a))(x - a), x - a \rangle \end{aligned}$$

mit einem gewissen $\theta \in (0, 1)$ gilt.

Analog -nur mit größerem formalen Aufwand- kann man Entwicklungen höherer Ordnung angeben.

Die Idee besteht wieder darin, f auf der Strecke zwischen x und a als reelle Funktion zu betrachten.

Beispiele:

$$f(x) = \sin x.$$

Wir betrachten die Taylorentwicklung für $a = 0$.

Mit $f'(0) = \cos 0 = 1$, $f''(0) = -\sin 0 = 0$, $f'''(0) = -\cos 0 = -1$ und $f^{(4)}(0) = \sin 0 = 0$ usw. folgt

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} \dots$$

Da das Restglied R_n einheitlich ($\forall x$) nicht größer als $\frac{x^{n+1}}{(n+1)!}$ ist, konvergiert die Reihe tatsächlich auf jedem Intervall gleichmäßig gegen $\sin x$.

Analog erhält man

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} \dots$$

$$f(x) = e^x.$$

Jetzt sind alle Ableitungen in $a = 0$ identisch Eins und das n -te Restglied ist nicht größer als

$$\frac{x^{n+1}e^x}{(n+1)!} \quad (\text{für } x > a = 0).$$

Deshalb konvergiert die Taylor-Reihe von e^x im selben Sinne wie oben; wir erhalten die schon bekannte Formel

$$e^x = \sum_{n \geq 0} \frac{x^n}{n!}.$$

Bleibt x auf ein Intervall $[a, b]$ beschränkt, konvergiert die Reihe gleichmäßig.

Fixiert man irgendein n , so wird andererseits die Differenz zwischen e^x und der Summe

$$\frac{x^0}{0!} + \dots + \frac{x^n}{n!}$$

beliebig groß, wenn nur x hinreichend groß ist. Es liegt also keine gleichmäßige Konvergenz bezüglich aller x aus \mathbb{R} vor.

Der Konvergenzradius der e^x -Reihe ist dennoch unbeschränkt, da sie für alle x konvergiert.

6.3.2 Anwendungen der Taylorentwicklung

1. Fehlerabschätzungen $f(x) - f(a)$ wird für x nahe a beliebig genau approximiert.

2. Monotonieverhalten Wenn f' nahe $x = a$ existiert (d.h. für alle x aus einer Umgebung von a) und f' in a stetig ist, so hat $f'(a) > 0$ zur Folge, dass für alle $x < y$ (x, y nahe a) gilt: $f(y) > f(x)$.

Beweis. Man benutze den Mittelwertsatz mit $a = x$ und $b = y$ und beachte, dass aufgrund der Stetigkeit von f' für θ in der Nähe von a noch $f'(\theta) > 0$ richtig ist. \square

Die Bedingung $f'(a) > 0$ ist nicht notwendig für Monotonie: $f(x) = x^3$.

Denn: Ist f differenzierbar, aber die Ableitung in a nicht stetig, kann die Monotonie-Aussage " $f(y) > f(x)$ (falls $f'(a) > 0$ und $x < y$ nahe a)" falsch werden. Man studiere mit $a = 0$ das Beispiel

$$f(x) = \begin{cases} x + x^2 \sin(1/x) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0 \end{cases}$$

3. Extrema Wenn f'' nahe a existiert (d.h. für alle x aus einer Umgebung von a) und f'' in a stetig ist, so haben $f''(a) > 0$ und $f'(a) = 0$ zusammen zur Folge, dass a ein (strenger) lokaler Minimalpunkt von f ist, d.h. wenn $|x - a|$ hinreichend klein und $x \neq a$, so $f(x) > f(a)$.

Beweis. Man benutze den Taylor-Satz mit $n = 1$ und beachte, dass wegen Stetigkeit von f'' noch $f''(\theta) > 0$ richtig ist. \square

Ist $f''(a) > 0$ nennt man f konvex (nahe a), ist $f''(a) < 0$, nennt man f konkav (nahe a).

Dies ist nicht die allgemeinste Definition von konvex und konkav (siehe Anhang). Punkte a , in welchen f'' das Vorzeichen wechselt, heißen auch Wendepunkte. Punkte a mit $f'(a) = 0$ heißen oft stationär.

4. Grenzwerte: l' Hospital-sche Regel Wenn $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = 0$ für $x \rightarrow a$ und $f'(x), g'(x)$ für x nahe a existieren, in a stetig sind und $g'(a) \neq 0$ gilt, so ist

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{f'(a)}{g'(a)}.$$

(Beispiel: $\lim_{x \rightarrow 0} \frac{x}{\sin x}$)

Man schreibe: $f(x) = f(a) + f'(\theta_1)(x - a)$, $g(x) = g(a) + g'(\theta_2)(x - a)$, beachte $f(a) = g(a) = 0$, dividiere und nutze $f'(\theta_1) \rightarrow f'(a)$, $g'(\theta_2) \rightarrow g'(a)$.

Falls $f'(a) = g'(a) = 0$, nehme man auch höhere Ableitungen.

5. Differentialgleichungen einfachsten Typs

5.1 Angenommen f ist für x nahe $a = 0$ differenzierbar und erfüllt dort $f'(x) = f(x)$. Dann folgt auch

$$f''(x) = f'(x), \quad f^{(3)}(x) = f''(x), \quad f^{(4)}(x) = f^{(3)}(x), \dots$$

Daher ergibt sich für jedes n und alle x nahe a :

$$\begin{aligned} f(x) &= f(a) + f'(a) \frac{(x-a)^1}{1!} + \dots + f^{(n)}(a) \frac{(x-a)^n}{n!} + R_n(x) \\ &= f(a) \left[1 + \frac{x^1}{1!} + \dots + \frac{x^n}{n!} \right] + R_n(x), \quad \text{wobei } R_n(x) = f(\theta) \frac{x^{n+1}}{(n+1)!}. \end{aligned}$$

Da f stetig ist, bleibt $f(\theta)$ beschränkt. Wegen $R_n(x) \rightarrow 0$ ist damit $f(x) = f(a)e^x$.

Mit anderen Worten:

Die (reelle) Differentialgleichung $f'(x) = f(x)$ ($\forall x$ nahe 0) hat nur Lösungen der Form $f(x) = c e^x$. Diese f sind offenbar auch wirklich Lösungen, c bestimmt den Wert von f in a .

5.2 Wir studieren analog die Lösungen von $f''(x) = f(x)$. Nun folgt -da f differenzierbar ist und nahe 0 mit f'' zusammenfällt-

$$f^{(3)}(x) = f'(x), \quad f^{(4)}(x) = f^{(2)}(x) = f(x), \quad f^{(5)}(x) = f^{(3)}(x) = f'(x), \dots$$

Daher ergibt sich für jedes n und alle x nahe $a = 0$:

$$\begin{aligned} f(x) &= f(a) + f'(a) \frac{(x-a)^1}{1!} + \dots + f^{(2n+1)}(a) \frac{(x-a)^{2n+1}}{(2n+1)!} + R_{2n+1}(x) \\ &= f(a) \left[1 + \frac{x^2}{2!} + \dots + \frac{x^{2n}}{(2n)!} \right] + f'(a) \left[x + \frac{x^3}{3!} + \dots + \frac{x^{2n+1}}{(2n+1)!} \right] + R_{2n+1}(x). \end{aligned}$$

Dabei ist $R_{2n+1}(x) = f(\theta) \frac{x^{2n+2}}{(2n+2)!}$. Da f stetig ist, bleibt $f(\theta)$ beschränkt (für x nahe 0). Wegen $R_{2n+1}(x) \rightarrow 0$ und der absoluten Konvergenz beider Reihen ist damit $f(x) = f(a) \cosh x + f'(a) \sinh x$. Wieso $\cosh x$ und $\sinh x$?

Man beachte dass $\cosh x = \frac{1}{2}(e^x + e^{-x})$, $\sinh x = \frac{1}{2}(e^x - e^{-x})$ per Definition gilt (auch für komplexe x). Dann gilt: $(\cosh x)' = \sinh x$, $(\sinh x)' = \cosh x$ und $(\cosh x)^2 - (\sinh x)^2 = 1$. Dadurch hat $\cosh x$ die (absolut konvergente) Taylorreihe ($a = 0$):

$$\begin{aligned} \cosh x &= 1 + \frac{x^2}{2!} + \dots + \frac{x^{2n}}{(2n)!} + \dots; \quad \text{analog} \\ \sinh x &= x + \frac{x^3}{3!} + \dots + \frac{x^{2n+1}}{(2n+1)!} \dots \quad (\text{vergleiche } \sin x \text{ und } \cos x) \end{aligned}$$

Wir erhalten: Die (reelle) Differentialgleichung

$$f''(x) = f(x) \quad (\forall x \text{ nahe } 0)$$

hat nur Lösungen der Form $f(x) = f(a) \cosh x + f'(a) \sinh x$, d.h.

$$f(x) = \frac{1}{2}\alpha(e^x + e^{-x}) + \frac{1}{2}\beta(e^x - e^{-x}).$$

Variation von α und β liefert alle Funktionen f der Form $\hat{\alpha}e^x + \hat{\beta}e^{-x}$.

Diese f sind offenbar auch Lösungen von $f'' = f$ (einsetzen und beide Seiten ausrechnen!).

Übungsaufgabe: Sei f für x nahe 0 zweimal differenzierbar und gelte dort $f(x) + f''(x) = 0$. Zeigen Sie, dass dann $f(x) = \alpha \sin x + \beta \cos x$ gilt (α und β sind Konstanten).

6. Krümmungskreis (etwas Geometrie) Wir wollen den Graphen einer (hinreichend glatten) Funktion $f = f(x)$ nahe $(x, y) = (a, f(a))$ durch die Punkte eines Kreises in der (x, y) -Ebene mit Mittelpunkt (u, v) und Radius r approximieren:

$$(y - v)^2 + (x - u)^2 = r^2.$$

Dann ist

$$y(x) = v + \sqrt{r^2 - (x - u)^2} = v + \sqrt{w(x)} \quad \text{mit } w(x) = r^2 - (x - u)^2, w' = -2(x - u),$$

und (Taylor) $f(x) = f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2 + R_2^f(x)$.

Die Approximation sei so genau, dass

$$(*) \quad \frac{|f(x) - y(x)|}{|x - a|^2} \rightarrow 0 \quad \text{für } x \rightarrow a, x \neq a.$$

Wir schreiben die Taylor-Entwicklung von y auf:

$$y(x) = y(a) + y'(a)(x - a) + \frac{1}{2}y''(a)(x - a)^2 + R_2^y(x)$$

wobei gilt:

$$\begin{aligned} y'(x) &= (\sqrt{w(x)})' = -\frac{x - u}{\sqrt{w}}. \\ y''(x) &= -\left(\frac{x - u}{\sqrt{w}}\right)' = -\left[\frac{1}{\sqrt{w}} - \frac{(x - u)w'}{2\sqrt{w}^3}\right] \\ &= -\frac{1}{\sqrt{w}^3} \left[w - \frac{1}{2}(x - u)w'\right] = -\frac{1}{\sqrt{w}^3} [w + (x - u)^2] = -\frac{r^2}{\sqrt{w}^3}. \end{aligned}$$

Ist $f^{(3)}$ stetig, folgt $\lim_{x \rightarrow a} \frac{R_2(x)}{|x - a|^2} = 0$ für beide Restglieder.

Wegen (*) muss nun gelten $y(a) = f(a)$, $y'(a) = f'(a)$ und $y''(a) = f''(a)$, also mit $w = w(a)$:

$$\begin{aligned} y(a) &= v + \sqrt{w} = f(a), \\ y'(a) &= \frac{(u - a)}{\sqrt{w}} = f'(a), \\ y''(a) &= -\frac{r^2}{\sqrt{w}^3} = f''(a). \end{aligned}$$

Dann folgt auch

$$(**) \quad r^2 = -f''(a)\sqrt{w}^3$$

$$(***) \quad u = a + f'(a)\sqrt{w}, \quad v = f(a) - \sqrt{w}.$$

Setzt man r^2 und u in $w = r^2 - (a - u)^2$ ein, erhält man

$$w = r^2 - (a - u)^2 = -f''(a)\sqrt{w}^3 - (f'(a))^2 w.$$

Also muss auch gelten

$$w(1 + (f'(a))^2) = -f''(a)\sqrt{w}^3 \quad \text{bzw.} \quad -\frac{1 + (f'(a))^2}{f''(a)} = \sqrt{w}.$$

Mit der Gleichung (**) ergibt sich so

$$\begin{aligned} r^2 &= -f''(a)\sqrt{w}^3 = \frac{(1 + (f'(a))^2)^3}{(f''(a))^2} \\ \Rightarrow r &= \frac{\sqrt{(1 + (f'(a))^2)^3}}{|f''(a)|} \\ \Rightarrow \frac{1}{r} &= \frac{|f''(a)|}{\sqrt{(1 + (f'(a))^2)^3}}. \end{aligned}$$

Die Größe $\frac{1}{r}$ (bei Krümmungsradius $r > 0$) heißt auch Krümmung von f im Punkt a , der berechnete Kreis heißt Krümmungs- bzw. Schmiegunskreis.

Sein Mittelpunkt hat nach (***) die (von a abhängigen) Koordinaten

$$\begin{aligned} u &= a + f'(a)\sqrt{w} = a - f'(a)\frac{1 + (f'(a))^2}{f''(a)} \\ v &= f(a) - \sqrt{w} = f(a) + \frac{1 + (f'(a))^2}{f''(a)}. \end{aligned}$$

Der geometrische Ort aller Krümmungsmittelpunkte (bei variablem a) heißt Evolute der Kurve. Oft ist man an der Orientierung der Kurve interessiert. Deshalb setzt man dann als Krümmung

$$k = \frac{f''(a)}{\sqrt{(1 + f'(a)^2)^3}}.$$

Positivem k entspricht dann eine (nahe a) konvexe Kurve, negativem k eine konkave. Ende Anwendungen der Taylor-Reihe, obwohl das lange noch nicht alle sind.

6.4 Ableitungen für $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ und Satz über implizite Funktionen

Für $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ bilden die Komponenten f_i ($i = 1, \dots, m$) jeweils \mathbb{R}^n in \mathbb{R} ab. Dabei ist f genau dann differenzierbar, wenn jede der Komponenten f_i differenzierbar ist (in einem Punkt oder auf einer gegebenen offenen Menge im \mathbb{R}^n).

Die lineare Funktion $Df(x^*) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ wird damit (falls sie existiert) durch die Ableitungen $Df_i(x^*) : \mathbb{R}^n \rightarrow \mathbb{R}$ definiert:

$$Df(x^*)(u_1, \dots, u_n) = \begin{pmatrix} Df_1(x^*)(u_1, \dots, u_n) \\ \dots \\ Df_m(x^*)(u_1, \dots, u_n) \end{pmatrix} \in \mathbb{R}^m.$$

Die Ableitungen $Df_i(x^*)$ sind mit dem Gradienten von f_i

$$\nabla f_i(x^*) = \text{grad } f_i(x^*) = \left(\frac{\partial f_i(x^*)}{\partial x_1}, \dots, \frac{\partial f_i(x^*)}{\partial x_n} \right) \in \mathbb{R}^n$$

durch das Skalarprodukt

$$Df_i(x^*)(u_1, \dots, u_n) = \frac{\partial f_i(x^*)}{\partial x_1} u_1 + \dots + \frac{\partial f_i(x^*)}{\partial x_n} u_n = \langle \text{grad } f_i(x^*), u \rangle$$

bestimmt. Um dies formal zu kennzeichnen, schreibt dann meist

$$df_i = \frac{\partial f_i(x^*)}{\partial x_1} dx_1 + \dots + \frac{\partial f_i(x^*)}{\partial x_n} dx_n$$

und nennt df_i auch vollständiges (bzw. totales) Differential von f_i in x^* .

Bildet man eine Matrix $J(x^*)$ aus den Zeilen $\text{grad } f_i(x^*)$

$$J(x^*) = \begin{pmatrix} \frac{\partial f_1(x^*)}{\partial x_1} & \dots & \frac{\partial f_1(x^*)}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial f_m(x^*)}{\partial x_1} & \dots & \frac{\partial f_m(x^*)}{\partial x_n} \end{pmatrix}$$

ergibt sich so $Df(x^*)(u)$ aus dem Produkt von $J(x^*)$ mit $u \in \mathbb{R}^n$ (als Spaltenvektor).

Die Matrix $J(x^*)$ heißt **Jacobi-Matrix** von f in x^* . Sie wird zumeist mit der linearen Abbildung $Df(x^*)$ identifiziert, da der Zusammenhang zwischen der Matrix J und der Ableitung $Df(x^*)$ als lineare Funktion in der obigen Weise klar ist.

Es sei nun speziell $m = n$.

Die Regularität der Jacobi-Matrix $J(x^*)$ (sie bedeutet, dass $Df(x^*)$ eine bijektive lineare Abbildung von \mathbb{R}^n in sich vermittelt, siehe auch Algebra-Vorlesung) lässt wichtige Aussagen über die Nullstellen der nichtlinearen Gleichung $f(x) = y$ (finde x bei gegebenem y) auch bei kleinen Änderungen der Funktion f selbst zu.

6.4.1 Satz über implizite Funktionen

Um dies genauer zu fassen, betrachten wir eine differenzierbare Funktion $F = F(x, t)$, die für $x \in \mathbb{R}^n$, $t \in \mathbb{R}^m$ erklärt ist und Werte in \mathbb{R}^n annimmt. Für festes t bildet also $F(\cdot, t)$ als Funktion von x den \mathbb{R}^n in sich ab. Die Ableitung dieser Funktion (bzgl. x) in (x^*, t^*) bezeichnen wir mit $D_x F(x^*, t^*)$. Analog sei $D_t F(x^*, t^*)$ definiert.

Satz 6.18 (Satz über implizite Funktionen).

Vor.: Es sei $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$ und $F(x^*, t^*) = 0$. Die Ableitungen $D_x F$ und $D_t F$ der Funktion F bzgl. x und t mögen existieren und stetig vom Argument $(x, t) \in \mathbb{R}^{n+m}$ (nahe (x^*, t^*)) abhängen. Weiter sei die (n, n) -Matrix $J = D_x F(x^*, t^*)$ regulär.

Beh.: Dann gibt es positive α und β , sodass die Gleichung

$$(1) \quad F(x, t) = 0$$

für alle t mit $\|t - t^*\| < \beta$ genau eine Lösung $x = x(t)$ mit $\|x - x^*\| < \alpha$ besitzt. Außerdem ist die Funktion $x = x(t)$ in t^* differenzierbar und besitzt die Ableitung (Jacobi-Matrix)

$$Dx(t^*) = -J^{-1}D_t F(x^*, t^*).$$

Beweis. Wir identifizieren J mit der durch diese Matrix erzeugten linearen Abbildung von \mathbb{R}^n in sich. Die zugehörige inverse Abbildung/Matrix sei J^{-1} .

Umformulierung von (1) als eine geeignete Fixpunktgleichung:

Gleichung (1) lässt sich äquivalent als

$$J^{-1}F(x, t) = 0$$

schreiben. Da F differenzierbar ist und $F(x^*, t^*) = 0$, kann sie weiter umgeformt werden:

$$J^{-1} (DF(x^*, t^*)(x - x^*, t - t^*) + o(x - x^*, t - t^*)) = 0. \quad (6.4.1)$$

Für die Ableitung (als lineare Funktion) gilt

$$DF(x^*, t^*)(x - x^*, t - t^*) = D_x F(x^*, t^*)(x - x^*) + D_t F(x^*, t^*)(t - t^*),$$

und da J^{-1} angewandt auf $J = D_x F(x^*, t^*)$ die identische Abbildung liefert, bedeutet $F(x, t) = 0$ auch

$$x - x^* + J^{-1} D_t F(x^*, t^*)(t - t^*) + J^{-1} o(x - x^*, t - t^*) = 0 \quad \text{bzw.}$$

$$(4) \quad x = [x^* - J^{-1} D_t F(x^*, t^*)(t - t^*)] - J^{-1} o(x - x^*, t - t^*).$$

Mit

$$g(t) = -x^* + J^{-1} D_t F(x^*, t^*)(t - t^*), \quad h(x, t) = J^{-1} o(x - x^*, t - t^*)$$

sowie

$$T(x, t) = - (g(t) + h(x, t)) = x^* - J^{-1} D_t F(x^*, t^*)(t - t^*) - J^{-1} o(x - x^*, t - t^*) \quad (6.4.2)$$

wird (4) und damit (1) mit jedem festen t zu einer Fixpunktgleichung mit gesuchtem x ,

$$(5) \quad x = T(x, t).$$

Kontraktivität: Die stetig differenzierbare Funktion o und damit h hat in (x^*, t^*) die Ableitung Null (speziell bzgl. x). Wendet man den Mittelwertsatz auf jede ihrer n Komponenten an, sieht man, dass zu jedem vorgegebenem $r > 0$ die Abschätzung

$$(6) \quad \|h(y, t) - h(x, t)\| \leq r \|y - x\|$$

für x und y nahe x^* und t nahe t^* richtig ist (die Abstände hängen dabei von r ab).

Da g nicht von x, y abhängt, gilt (6) auch für Abbildung T anstelle von h . Es folgt so mit $r = \frac{1}{2}$:

Es gibt ein $\alpha > 0$, sodass

$$(7) \quad \|T(y, t) - T(x, t)\| \leq \frac{1}{2} \|y - x\| \quad \forall x, y \in B(x^*, \alpha), \quad \forall t \in B(t^*, \alpha)$$

für die abgeschlossene Kugel $B(x^*, \alpha) = \{x \mid d(x, x^*) \leq \alpha\}$ und analog $B(t^*, \alpha)$ richtig ist.

Das Bild von $T(\cdot, t)$: Weiter gilt

$$(8) \quad \|x^* - T(x, t)\| = \|J^{-1} D_t F(x^*, t^*)(t - t^*) + J^{-1} o(x - x^*, t - t^*)\| \\ \leq \|J^{-1} D_t F(x^*, t^*)(t - t^*)\| + \frac{1}{2} \|x - x^*\|$$

Die affin-lineare Funktion $l(t) = J^{-1}D_tF(x^*, t^*)(t - t^*)$ ist Lipschitz-stetig. Damit gibt es eine Konstante K , sodass

$$\|J^{-1}D_tF(x^*, t^*)(t - t^*)\| \leq K\|t - t^*\|. \quad (6.4.3)$$

Ist $\beta < \alpha$ klein genug und $\|t - t^*\| < \beta$, so erfüllen deshalb alle $x \in B(x^*, \alpha)$ die Ungleichung

$$\|x^* - T(x, t)\| \leq \frac{\alpha}{2} + \frac{\alpha}{2} = \alpha.$$

Für so gewählte t wird also die abgeschlossene Kugel $B(x^*, \alpha)$ durch die Funktion $T(\cdot, t)$ in sich abgebildet.

Nach (7) ist weiter $T(\cdot, t)$ kontraktiv auf dem vollständigen metrischen Raum $X = B(x^*, \alpha)$.

Damit kann Banachs Fixpunktsatz mit $X = B(x^*, \alpha)$ angewandt werden, wonach Gleichung (5) für jedes $t \in B(t^*, \beta)$ genau eine Lösung $x = x(t) \in B(x^*, \alpha)$ besitzt.

Die erste Aussage des Satzes ist also richtig.

Stetigkeit des Fixpunktes $x = x(t) = T(x(t), t)$. Es ist für $s, t \in B(t^*, \beta)$,

$$\begin{aligned} \|x(s) - x(t)\| &= \|T(x(s), s) - T(x(t), t)\| \\ &= \|T(x(s), s) - T(x(t), s)\| + \|T(x(t), s) - T(x(t), t)\| \\ &\leq \|T(x(s), s) - T(x(t), s)\| + \|T(x(t), s) - T(x(t), t)\| \\ &\leq \frac{1}{2}\|x(s) - x(t)\| + \|T(x(t), s) - T(x(t), t)\|. \end{aligned}$$

Folglich gilt

$$\frac{1}{2}\|x(s) - x(t)\| \leq \|T(x(t), s) - T(x(t), t)\|.$$

Mit $t = t^*$ ist dann $x(t^*) = x^*$ und $\frac{1}{2}\|x(s) - x^*\| \leq \|T(x^*, s) - T(x^*, t^*)\|$.

Nach Definition von T gilt

$$T(x, s) - x^* = -J^{-1}D_tF(x^*, t^*)(s - t^*) - J^{-1}o(x - x^*, s - t^*). \quad (6.4.4)$$

Das gibt uns $T(x^*, s) - T(x^*, t^*) = -J^{-1}D_tF(x^*, t^*)(s - t^*) - J^{-1}o(0, s - t^*)$. Wegen

$$\lim_{s \rightarrow t^*} \frac{o(0, s - t^*)}{\|s - t^*\|} = 0$$

und Linearität der Abbildung $J^{-1}D_tF(x^*, t^*)$ (siehe (6.4.3)) gibt es deshalb eine Konstante C sodass

$$\|x(s) - x^*\| \leq C\|s - t^*\| \text{ wenn } \|s - t^*\| \text{ klein genug ist} \quad (6.4.5)$$

(z.B. $C = K + 1$). Also strebt $x(s)$ in (6.4.5) sogar Lipschitz stetig gegen x^* , wenn $s \rightarrow t^*$.

Differenzierbarkeit: Hierzu benutzen wir (6.4.4) mit $x = x(s)$

$$x(s) - x^* = T(x(s), s) - x^* = -J^{-1}D_tF(x^*, t^*)(s - t^*) - J^{-1}o(x(s) - x^*, s - t^*). \quad (6.4.6)$$

Nun nutzen wir

$$\frac{o(x(s) - x^*, s - t^*)}{\|s - t^*\|} = \frac{o(x(s) - x^*, s - t^*)}{\|x(s) - x^*, s - t^*\|} \frac{\|x(s) - x^*, s - t^*\|}{\|s - t^*\|}.$$

Die Norm des ersten Faktors strebt wegen Differenzierbarkeit von F und $x(s) \rightarrow x^*$ gegen Null. Für den zweiten erhalten wir mit (6.4.5)

$$\frac{\|x(s) - x^*, s - t^*\|}{\|s - t^*\|} \leq \frac{\|x(s) - x^*, 0\|}{\|s - t^*\|} + \frac{\|0, s - t^*\|}{\|s - t^*\|} \leq \frac{C\|s - t^*\|}{\|s - t^*\|} + \frac{\|s - t^*\|}{\|s - t^*\|} = C + 1.$$

Also ist in (6.4.6) $\lim_{s \rightarrow t^*} \frac{\|J^{-1}o(x(s) - x^*, s - t^*)\|}{\|s - t^*\|} = 0$, und es gilt tatsächlich

$$x(s) - x(t^*) = x(s) - x^* = -J^{-1}D_tF(x^*, t^*)(s - t^*) + r(s - t^*) \text{ wobei } \frac{\|r(s - t^*)\|}{\|s - t^*\|} \rightarrow 0. \quad (6.4.7)$$

Das ist die Behauptung: $Dx(t^*)$ ist durch die Matrix $-J^{-1}D_tF(x^*, t^*)$ definiert. \square

In der Vorlesung haben wir $x = y$ und $t = x$ gesetzt, um die implizite Funktion als $y = y(x)$ zu erhalten.

Inverse Funktion: Setzt man $F(x, t) = f(x) - t$ mit $x, t \in \mathbb{R}^n$, liefert der bewiesene Satz eine Aussage über Existenz und Differenzierbarkeit der Lösungen der Gleichung $f(x) = t$, d.h. über die inverse Funktion f^{-1} mit $x(t) = f^{-1}(t)$. Ihre Ableitung in t^* ist dann

$$D(f^{-1})(t^*) = -Df(x^*)^{-1}(-E) = Df(x^*)^{-1}.$$

Vergleichen Sie diese mit dem Fall einer linearen Abbildung f !

Der Satz über implizite Funktionen lässt sich in vieler Hinsicht allgemeiner formulieren und beweisen. Die Beweisidee (Reduktion auf eine Fixpunktaussage, zumeist mittels Banachs Fixpunktsatz), findet sich aber immer wieder.

Eine von vielen wichtigen Anwendungen des Satzes ist verbunden mit

6.4.2 Lagrange Multiplikatoren für Extrema mit Nebenbedingungen

Gesucht sei

$\min f(x)$, wobei $x \in \mathbb{R}^n$ ein Gleichungssystem $g(x) = 0$ erfüllen soll

$$g = \begin{pmatrix} g_1 \\ \dots \\ g_m \end{pmatrix} \quad m \leq n.$$

Die Menge der zulässigen x heiße M . Die Funktionen f, g seien stetig differenzierbar. Sei $x^* \in M$ eine lokale Lösung der Aufgabe, d.h., es gibt ein $\varepsilon > 0$, sodass $f(x) \geq f(x^*) \forall x \in M$ mit $\|x - x^*\| < \varepsilon$. Weiter habe die (m, n) -Matrix der partiellen Ableitungen von g (die Jacobi-Matrix von g)

$$Dg = \begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n} \\ \dots & \dots & \dots \\ \frac{\partial g_m}{\partial x_1} & \dots & \frac{\partial g_m}{\partial x_n} \end{pmatrix}$$

im Punkt x^* vollen Rang (dann ist $\text{rang} Dg(x^*) = m \leq n$).

Satz 6.19 (Existenz von Lagrange Multiplikatoren). *Unter obigen Voraussetzungen gibt es Zahlen $\lambda_1, \dots, \lambda_m$ mit*

$$Df(x^*) + \lambda_1 Dg_1(x^*) + \dots + \lambda_m Dg_m(x^*) = 0. \quad (6.4.8)$$

Bemerkung: Die reellen λ_i heißen *Lagrange Multiplikatoren*.

Zusammen mit $g(x^*) = 0$ ergeben diese Bedingungen, aufgeschrieben in Form der partiellen Ableitungen, ein System von $n + m$ reellwertigen Gleichungen mit genau $n + m$ Variablen, das die optimalverdächtigen Punkte analog charakterisiert wie $Df(x^*) = 0$, wenn keine Nebenbedingungen auftreten. Mit der Lagrange Funktion

$$L(x, \lambda) = f(x) + \lambda_1 g_1(x) + \dots + \lambda_m g_m(x)$$

wird das Paar (x^*, λ) auch durch $D_x L(x^*, \lambda) = 0$ und $D_\lambda L(x^*, \lambda) = 0$ beschrieben.

Beweis. Die Matrix $Dg(x^*)$ hat Rang m im Punkt x^* . Es gibt also genau m linear unabhängige Spalten. O.B.d.A. seien das die ersten m (sonst Umbenennen). Wenn $m = n$, bilden die Vektoren $Dg_1(x^*), \dots, Dg_m(x^*)$ eine Basis des \mathbb{R}^n . Damit hat das lineare System (6.4.8) eine Lösung λ und wir sind fertig. Sei $m < n$. Wir zerlegen x in die ersten m und die $n - m$ restlichen Komponenten und schreiben entsprechend $x = (x_1, x_2)$.

Analog schreibe man $Dg(x^*) = (J_1, J_2)$. Die (m, m) Matrix $J_1 = D_{x_1} g(x^*)$ ist dann regulär. Sei

$$\lambda = -D_{x_1} f(x^*) J_1^{-1} \quad (\lambda \text{ und } D_{x_1} f(x^*) \text{ als Zeilenvektoren der Länge } m). \quad (6.4.9)$$

Dann folgt $\lambda J_1 = -D_{x_1} f(x^*)$, also $D_{x_1} f(x^*) + \lambda Dg_{x_1}(x^*) = 0$. Die Behauptung (6.4.8) gilt demnach für die m partiellen Ableitungen zu x_1 .

Der Vektor x_1 hängt über die Gleichung $g = 0$ nach dem Satz über implizite Funktionen nahe x^* lokal eindeutig und differenzierbar von $x_2 (= t)$ ab. Sei u diese implizite Funktion. Ihre Ableitung ist

$$Du(x_2^*) = -J_1^{-1} D_{x_2} g(x^*). \quad (6.4.10)$$

Die Punkte $(u(x_2), x_2)$ sind alle aus M (wenn x_2 nahe x_2^* variiert). Also minimiert x_2^* die Funktion $h(x_2) := f(u(x_2), x_2)$ lokal. Damit gilt $Dh(x_2^*) = 0$. Mit Kettenregel heißt das:

$$0 = D_{x_2}f(u(x_2^*), x_2^*) = D_{x_1}f(x^*)Du(x_2^*) + D_{x_2}f(x^*).$$

Setzt man (6.4.10) und (6.4.9) ein, folgt wie verlangt auch

$$0 = D_{x_1}f(x^*)(-J_1^{-1})D_{x_2}g(x^*) + D_{x_2}f(x^*) = \lambda D_{x_2}g(x^*) + D_{x_2}f(x^*).$$

□

Die Rang-Voraussetzung im Satz ist wichtig. Man nehme sonst die Aufgabe $\min x$ unter der Bedingung $x^2 = 0$ mit der einzigen Lösung $x^* = 0$. Dann wird (6.4.8) zur Gleichung $1 + \lambda_1 * 0 = 0$.

6.4.3 Newton-Methode zur Lösung von $F(x) = 0$:

Es seien $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ stetig differenzierbar und x^* eine Lösung der Gleichung

$$(1) \quad F(x) = 0.$$

Zu vorgegebenem x (nahe x^*) betrachten wir die Linearisierung von F in x , also die in y (affin-) lineare Funktion G_x mit $G_x(y) = F(x) + DF(x)(y - x)$.

Die Idee der Methode besteht darin, die einfachere lineare Gleichung

$$(2) \quad G_x(y) = 0$$

zu lösen und sukzessiv x durch die (oder eine) Lösung $x' = y$ zu ersetzen. Beginnend mit einem Startpunkt x_0 soll so eine Folge $\{x_k\}$ entstehen, die gegen x^* konvergiert, nämlich - wegen der Gestalt von $G_x(y)$ -

$$x_{k+1} = x_k - DF(x_k)^{-1}F(x_k) \quad \text{sofern } DF(x_k)^{-1} \text{ existiert.}$$

Man wendet also auf $g(x) = x - DF(x)^{-1}F(x)$ einfach sukzessive Approximation an.

Satz 6.20 (Konvergenz der Newton-Methode). *Es sei $DF(x^*)$ regulär. Dann gibt es zu jedem $\varepsilon \in (0, 1)$ ein positives δ , sodass für alle Startpunkte x_0 mit $\|x_0 - x^*\| < \delta$ die Newton-Methode eine Folge $(x_k)_{k \in \mathbb{N}}$ liefert, die der Ungleichung*

$$(3) \quad \|x_{k+1} - x^*\| \leq \varepsilon \|x_k - x^*\| \quad \forall k \geq 0$$

genügt. Insbesondere bleiben die linearen Gleichungen (2) für $x = x_k$ sämtlich eindeutig lösbar.

Beweis. Um etwas zu vereinfachen, setzen wir stärker voraus, dass auch D^2F existiert und stetig ist. Man betrachte die Funktion

$$h(x, y) = F(x) + DF(x)(y - x).$$

Das Verfahren ordnet jedem x ein $y(x)$ mit $h(x, y(x)) = 0$ zu. Für x nahe x^* , etwa $\|x - x^*\| < \alpha$, bleibt $DF(x)$ regulär. Dann ist $y(x)$ eindeutig, und zu x^* gehört die Lösung $y^* = y(x^*) = x^*$. Weiter ist

$$D_x h = DF(x) + D^2F(x)(y - x) - DF(x) = D^2F(x)(y - x).$$

$$D_y h = DF(x) \quad (\text{regulär in } x^*).$$

Nach dem Satz über implizite Funktionen ist die Funktion $x \mapsto y(x)$ für $\|x - x^*\| < \alpha$ nun nicht nur eindeutig bestimmt, sondern besitzt in x^* auch eine Ableitung $Dy(x^*)$. Diese ist Null, denn wegen $y^* = x^*$ folgt

$$D_y h(x^*, y^*) = DF(x^*), \quad D_x h(x^*, y^*) = 0 \quad \text{und} \quad Dy(x^*) = -D_y h(x^*, y^*)^{-1} D_x h(x^*, y^*) = 0.$$

Folglich gilt $\|y(x) - x^*\| = \|y(x) - y^*\| = o(\|x - x^*\|)$, was offenbar die Behauptung sichert. □

Bemerkungen:

Die Newton-Methode lässt sich in vielerlei Hinsicht verallgemeinern und modifizieren, z.B:

1. Ist $\|x_0 - x^*\|$ klein genug, kann man stets mit derselben Matrix arbeiten; man setzt dann $DF(x^k) \equiv DF(x_0)$.
2. Man kann $DF(x^k)$ durch Näherungsmethoden (billiger als exakt) berechnen
3. Man kann zum Teil auf die C^1 -Eigenschaften von F verzichten.

6.4. ABLEITUNGEN FÜR $F : \mathbb{R}^N \rightarrow \mathbb{R}^M$ UND SATZ ÜBER IMPLIZITE FUNKTIONEN 73

4. Die Räume müssen nicht endlich-dimensional sein.
5. Die Voraussetzung $\|x_0 - x^*\| < \delta$, die man nicht überprüfen kann, weil x^* unbekannt ist, kann man für hinreichend „reguläre“ Funktionen F durch $\|F(x_0)\| < \delta$ ersetzen.
6. Die Konvergenzaussagen sind i.a. nur für Startpunkte nahe einer Nullstelle richtig (man setze etwa $F(x) = \sin x$, um das zu sehen (siehe Vorlesung). Für Konvergenz mit beliebigen Startpunkten braucht man zusätzlich Eigenschaften (gewisse Monotonie) von F .
7. Die Menge aller Startpunkte x_0 , für welche die Methode eine konvergente Folge erzeugt, kann eine sehr komplizierte Struktur besitzen (\rightarrow Fraktale).

Kapitel 7

Komplexe Zahlen und Funktionen

Rechnen mit komplexen Zahlen, e^z , $\sin z$, $\cos z$, komplexe Differenzierbarkeit, Fundamentalsatz, Darstellung in Polarkoordinaten, Potenzen und Wurzeln, Linearität bezüglich der komplexen Multiplikation, Komplexe Differenzierbarkeit und Cauchy-Riemannsche Differentialgleichungen. Die Funktionen e^z , $\sin z$, $\cos z$, $\sinh z$, $\cosh z$ und ihre Reihen. Die Eulersche Darstellung $e^{i\varphi} = \cos \varphi + i \sin \varphi$. Periodizität der e -Funktion und des Logarithmus, Fundamentalsatz.

7.1 Potenzreihen und spezielle komplexe Funktionen

7.1.1 Potenzreihen, Exponential-, Sinus- und Kosinus-Funktion

Eine komplexe Potenzreihe $p(z)$ hat dieselbe Form $\sum_{k \geq 0} a_k z^k$ wie im Reellen, nur sind jetzt a_k und z komplexe Zahlen. Schreibt man sie als

$$a_k = r_k(\cos \varphi_k + i \sin \varphi_k), \quad z = s(\cos \omega + i \sin \omega),$$

so wird

$$a_k z^k = r_k(\cos \varphi_k + i \sin \varphi_k) s^k (\cos k\omega + i \sin k\omega) = r_k s^k (\cos(\varphi_k + k\omega) + i \sin(\varphi_k + k\omega))$$

mit dem Betrag

$$|a_k z^k| = |r_k s^k| \cdot |\cos(\varphi_k + k\omega) + i \sin(\varphi_k + k\omega)| = r_k s^k.$$

Im Reellen wären φ_k und ω beide 0 und $a_k z^k = r_k s^k$.

Mit dem natürlichen Konvergenzbegriff komplexer Zahlen, der wie im Reellen auf $|z - \zeta|$ aufbaut und nun Konvergenz der Real- und Imaginärteile fordert, konvergiert die Reihe $\sum_{k \geq 0} a_k z^k$ gegen eine komplexe Zahl $\zeta = u + iv$ genau dann, wenn

$$u = \sum_{k \geq 0} r_k s^k \cos(\varphi_k + k\omega) \quad \text{und} \quad v = \sum_{k \geq 0} r_k s^k \sin(\varphi_k + k\omega).$$

Aus dem Konvergenzintervall wird nun ein Kreis, der durch diejenigen Radien (Beträge) s von z bestimmt ist, für die $r_k s^k \leq q^k$ mit einem gewissen $q \in [0, 1)$ und hinreichend großen k gilt. Damit kann man definieren:

$$e^z = \sum_{k \geq 0} \frac{z^k}{k!}$$

Nun ist a_k reell, also $\varphi_k = 0$, $r_k = \frac{1}{k!}$, $u = \sum_{k \geq 0} \frac{s^k}{k!} \cos(k\omega)$, $v = \sum_{k \geq 0} \frac{s^k}{k!} \sin(k\omega)$.

Wegen $\left| \frac{s^k}{k!} \cos(k\omega) \right| \leq \left| \frac{s^k}{k!} \right|$ und $\left| \frac{s^k}{k!} \sin(k\omega) \right| \leq \left| \frac{s^k}{k!} \right|$ konvergieren beide reellen Reihen absolut und

$$|u| \leq e^s, \quad |v| \leq e^s.$$

Analog definiert man (siehe reelle Taylorreihen von \sin und \cos) im Komplexen

$$\begin{aligned}\cos z &= \sum_{k \geq 0} (-1)^k \frac{z^{2k}}{(2k)!} = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} \dots \\ \sin z &= \sum_{k \geq 0} (-1)^k \frac{z^{2k+1}}{(2k+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} \dots\end{aligned}$$

Für reelle φ folgt dann (Einsetzen und Addieren) die Eulersche Darstellung komplexer Zahlen

$$e^{i\varphi} = \cos \varphi + i \sin \varphi \quad (\text{mit } z = i\varphi) \quad (7.1.1)$$

Sie zeigt insbesondere, dass $i = \cos(\frac{\pi}{2}) + i \sin(\frac{\pi}{2}) = e^{i\frac{\pi}{2}}$ und dass $e^{i\varphi}$ in φ periodisch ist:

$$e^{i\varphi} = e^{i(\varphi+2k\pi)} \quad \forall k \in \mathbb{Z}.$$

Mit der Polar-Darstellung $z = r(\cos \varphi + i \sin \varphi)$ für z folgt nun $z = re^{i\varphi}$ und mit $z = x + iy$ erhält man

$$e^z = e^{x+iy} = e^x e^{iy} = e^x (\cos y + i \sin y). \quad (7.1.2)$$

Hierzu ist nur die Formel

$$e^{x+iy} = e^x e^{iy}$$

zu beweisen, wozu wir $\binom{k}{j}/k! = \frac{1}{j!(k-j)!}$ nutzen: Nach Definition ist

$$e^{x+iy} = \sum_{k \geq 0} \frac{(x+iy)^k}{k!} = \sum_{k \geq 0} \sum_{j=0}^k \frac{\binom{k}{j} (iy)^j x^{k-j}}{k!} = \sum_{k \geq 0} \left(\sum_{j=0}^k \frac{(iy)^j}{j!} \frac{x^{k-j}}{(k-j)!} \right).$$

Hier tritt jedes Produkt der Form $\frac{(iy)^j}{j!} \frac{x^m}{m!}$ genau einmal auf, nämlich wenn $k-j = m \geq 0$. Die Partialsummen von e^{x+iy} bis $k = n$ (was $j \leq n$ und $m \leq n$ impliziert) haben deshalb die Form

$$S_n = \sum_{m=0}^n \left[\sum_{j=0}^n \frac{(iy)^j}{j!} \frac{x^m}{m!} \right] = \left(\sum_{m=0}^n \frac{x^m}{m!} \right) \left(\sum_{j=0}^n \frac{(iy)^j}{j!} \right) = u_n * v_n$$

Dabei sind u_n und v_n die konvergenten Partialsummen zu e^x und e^{iy} . Mit $n \rightarrow \infty$ folgt deshalb die Behauptung. \square

Folgerung: Mit der Ueb-Aufgabe $e^{i(\alpha+\beta)} = e^{i\alpha} e^{i\beta}$ erhalten wir so auch

$$\begin{aligned}e^z &= e^{z+i2k\pi} \quad \forall k \in \mathbb{Z}, \\ e^{z_1+z_2} &= e^{x_1+x_2+i(y_1+y_2)} = e^{x_1+x_2} e^{i(y_1+y_2)} = e^{x_1} e^{iy_1} e^{x_2} e^{iy_2} = e^{z_1} e^{z_2}.\end{aligned}$$

7.1.2 Logarithmus

Als Umkehrfunktion von e^z , d.h. $z = \ln \zeta$ genau dann, wenn $e^z = \zeta$, ist deshalb der Logarithmus im Imaginärteil nicht eindeutig (2π -mehrwertig), denn aus

$$\zeta = e^z = e^{x+iy} = e^x (\cos y + i \sin y)$$

folgt auch

$$\zeta = e^x (\cos(y+2k\pi) + i \sin(y+2k\pi)) = e^{x+i(y+2k\pi)} = e^{z+i2k\pi}.$$

Wenn also $z = \ln \zeta$, so ist auch $z + i2k\pi = \ln \zeta$.

Die Funktionen $\sinh z = \frac{1}{2}(e^z - e^{-z})$, $\cosh z = \frac{1}{2}(e^z + e^{-z})$ und ihre Umkehrfunktionen definiert man wie im reellen Fall (sie werden ebenfalls periodisch), ebenso $\tan z$ und $\tanh z$ als Quotient. Interessanterweise folgt nun (bei natürlicher Erweiterung der Potenzgesetze) wegen

$$i = \cos\left(\frac{\pi}{2}\right) + i \sin\left(\frac{\pi}{2}\right) = e^{i\left(\frac{\pi}{2}\right)},$$

dass $i^i = (e^{i\frac{\pi}{2}})^i = e^{i \cdot i \frac{\pi}{2}} = e^{-\frac{\pi}{2}}$ reell wird.

7.2 Der Fundamentalsatz der Algebra

Eine Funktion $P : \mathbb{C} \rightarrow \mathbb{C}$ des Typs

$$P(z) = \sum_{k=0}^n a_k z^k \quad \text{mit } a_k, z \in \mathbb{C}, \text{ sowie } a_n \neq 0 \quad (7.2.1)$$

heißt komplexes Polynom vom Grade (degree) n . Wir beweisen zuerst zwei Hilfsaussagen.

Lemma 7.1. *Wenn der Grad $\deg P$ nicht Null ist, so folgt $|P(z)| \rightarrow \infty$ aus $|z| \rightarrow \infty$.*

Beweis. Man benutze Polarkoordinaten $z = r(\cos \varphi + i \sin \varphi)$, wonach $r = |z|$,

$$P(z) = \sum_{k=0}^n a_k r^k [\cos(k\varphi) + i \sin(k\varphi)], \quad \text{und offenbar gilt}$$

$$(|P(z)| \geq) \quad |a_n| r^n - \sum_{k=0}^{n-1} |a_k| r^k \rightarrow \infty \quad \text{wenn } r \rightarrow \infty.$$

□

Damit sind zwei Polynome nur dann gleich, wenn alle ihre Koeffizienten identisch sind (wieso?)

Lemma 7.2. *(Andere Darstellung des Polynoms in Form der Taylorreihe): Ein Polynom (7.2.1) lässt sich mit jedem festen ζ auch in der Form*

$$P(z) = \sum_{k=0}^n b_k (z - \zeta)^k \quad \text{mit } b_k, \zeta \text{ komplex} \quad (7.2.2)$$

darstellen.

Beweis. Wenn $n = \deg P = 0$, ist das trivial. Sei schon jedes Polynom vom Grade $\leq n - 1$ wie oben darstellbar. Man schreibe $a_n z^n = a_n z z^{n-1} = a_n (z - \zeta) z^{n-1} + a_n \zeta z^{n-1}$ und damit

$$P(z) = a_n (z - \zeta) z^{n-1} + a_n \zeta z^{n-1} + \sum_{k=0}^{n-1} a_k z^k.$$

Für z^{n-1} und die beiden weiteren Summanden können wir $\sum_{k=0}^{n-1} b_k (z - \zeta)^k$ mit unterschiedlichen b_k schreiben. Multiplikation der ersten Summe mit $a_n (z - \zeta)$ und Addieren der beiden anderen liefert die Behauptung. □

7.2.1 Beweis des Fundamentalsatzes

Satz 7.3 (Fundamentalsatz). *Jedes komplexe Polynom einer Ordnung $n \geq 1$ besitzt wenigstens eine Nullstelle z .*

Beweis. Wir folgen im Wesentlichen dem Beweis aus [7]. Zunächst bemerken wir, dass die Funktion h mit $h(z) = |P(z)|$ stetig ist, da dies für P und den Betrag gilt.

Sei z_0 beliebig (z.B. $z_0 = 0$) und $v = \inf_{z \in \mathbb{C}} h(z)$. Wenn r hinr. gross ist und $|z| > r$ so folgt nach Lemma 7.1, $h(z) > h(z_0)$. Für das Infimum von h sind also nur z in einem hinreichend großen Kreis $K_r = \{z \mid |z| \leq r\}$ interessant, also z aus einer kompakten Menge $K_r \subset \mathbb{C}$. Daher gibt es (Satz von Weierstraß, Satz 3.11) ein ζ mit $h(\zeta) \leq h(z) \quad \forall z \in K_r$. Wegen $h(z) > h(z_0)$ für $z \in \mathbb{C} \setminus K_r$ folgt so auch

$$(*) \quad |P(\zeta)| \leq |P(z)| \quad \forall z \in \mathbb{C}; \quad (\zeta \text{ ist Minimalpunkt von } h).$$

Wir schreiben nun P in der Form aus Lemma 7.2,

$$P(z) = \sum_{k=0}^n b_k (z - \zeta)^k.$$

Für $z = \zeta$ erhalten wir $P(\zeta) = b_0$. Wenn $b_0 = 0$, ist ζ Nullstelle von P , wir wären fertig. Sei also $b_0 \neq 0$. Wir konstruieren daraus einen Widerspruch, indem wir zeigen, dass P konstant ($= b_0$) ist (also $b_k = 0 \forall k > 0$ gilt). Der Beweis zeigt zugleich, dass ein komplexes Polynom P konstant ist, wenn nur eine lokale Minimalstelle ζ von $|P(\cdot)|$ mit $|P(\zeta)| > 0$ existiert.

Dazu betrachte man Punkte z der Form $z = \zeta + tw$ mit kleinem, reellem $t > 0$ und speziellen komplexen w . Für diese gilt zunächst $z - \zeta = tw$ und

$$P(z) = \sum_{k=0}^n b_k (tw)^k = \sum_{k=0}^n b_k w^k t^k = b_0 + \sum_{k=1}^n b_k w^k t^k.$$

Angenommen, $b_1 \neq 0$. Dann hat die Gleichung

$$b_1 w^1 = -b_0$$

eine Lösung w . Mit ihr folgt offenbar

$$P(\zeta + tw) = b_0 - tb_0 + \sum_{k>1} t^k b_k w^k = (1-t)b_0 + \sum_{k>1} t^k b_k w^k.$$

Damit wird

$$|P(\zeta + tw)| \leq |(1-t)b_0| + \left| \sum_{k>1} t^k b_k w^k \right| \leq (1-t)|b_0| + \sum_{k>1} t^k |b_k w^k|,$$

also für kleine $t > 0$ auch $|P(\zeta + tw)| < |b_0|$ wegen

$$\frac{|P(\zeta + tw)| - |b_0|}{t} \leq -|b_0| + \sum_{k>1} t^{k-1} |b_k w^k| < 0$$

(Hierzu reicht $\sum_{k>1} t^{k-1} |b_k w^k| < |b_0|/2$). Wegen (*) muss deshalb $b_1 = 0$ sein.

Induktion: Es möge schon $b_1 = \dots = b_{m-1} = 0$ gelten, $m \leq n$. Angenommen, $b_m \neq 0$. Dann hat die Gleichung

$$b_m w^m = -b_0$$

eine Lösung w (die Existenz von Wurzeln $w = \sqrt[m]{-b_0/b_m}$ kann schon benutzt werden). Nun ist

$$P(\zeta + tw) = b_0 + \sum_{k \geq m} t^k b_k w^k = b_0 - t^m b_0 + \sum_{k > m} t^k b_k w^k = (1-t^m)b_0 + \sum_{k > m} t^k b_k w^k$$

(die Summe $\sum_{k > m} t^k b_k w^k$ ist Null, wenn $m = n$) und

$$|P(\zeta + tw)| \leq (1-t^m)|b_0| + \sum_{k > m} t^k |b_k w^k|.$$

Daraus folgt für kleine $t > 0$ ebenfalls $|P(\zeta + tw)| < |b_0|$ wegen

$$\frac{|P(\zeta + tw)| - |b_0|}{t^m} \leq -|b_0| + \sum_{k > m} t^{k-m} |b_k w^k| < 0.$$

Also ist wegen (*) auch $b_m = 0$. So folgt $b_1 = \dots = b_n = 0$. Das widerspricht wegen $b_n = a_n$ aber der Voraussetzung $a_n \neq 0$. \square

Folgerung 7.4. Jedes komplexe Polynom P vom Grade $n > 0$ und mit $a_n = 1$ hat eine Darstellung der Form

$$P(z) = (z - \zeta_1)(z - \zeta_2) \cdots (z - \zeta_n)$$

wobei die ζ_i die (eventuell mehrfachen) Nullstellen von P sind.

Beweis. Sei ζ in (7.2.2) eine Nullstelle ζ_1 von $P_1 := P$ aus (7.2.1). Dann ist

$$P_1(z) = b_0 + \sum_{k=1}^n b_k (z - \zeta_1)^k, \quad (b_n = a_n)$$

und wegen $P_1(\zeta_1) = 0$ also auch $b_0 = 0$. Damit folgt

$$P_1(z) = (z - \zeta_1) \sum_{k=1}^n b_k (z - \zeta_1)^{k-1} = (z - \zeta_1) P_2(z). \quad (7.2.3)$$

Hier ist P_2 das angegebene Polynom vom Grade $n-1$. Mit einer Nullstelle ζ_2 von P_2 erhalten wir analog

$$P_2(z) = (z - \zeta_2) P_3(z) \quad \text{also auch} \quad P_1(z) = (z - \zeta_1)(z - \zeta_2) P_3(z),$$

wobei nun P_3 ein Polynom vom Grade $n-2$ ist. Nach n Schritten folgt somit die Behauptung, weil der höchste Koeffizient stets 1 bleibt. \square

7.2.2 Polynomdivision und konjugiert komplexe Nullstellen

Polynomdivision: Um das Polynom P_2 aus Gleichung (7.2.3) explizit in der Form

$$P_2(z) = \sum_{k=0}^{n-1} c_k z^k$$

auszurechnen, kann man die Koeffizienten beider Seiten, d.h. von $P_1(z) = P(z)$ und $(z - \zeta_1) \sum_{k=0}^{n-1} c_k z^k$ vergleichen. Die Nullstelle heie jetzt ζ . Man erhlt durch Ausmultiplizieren

$$\begin{aligned} a_n &= c_{n-1} \\ a_{n-1} &= c_{n-2} - c_{n-1}\zeta \\ &\dots \\ a_k &= c_{k-1} - c_k\zeta \\ &\dots \\ a_1 &= c_0 - c_1\zeta \\ a_0 &= -c_0\zeta \end{aligned}$$

Aus den ersten n Gleichungen lassen sich c_{n-1}, \dots, c_0 sukzessiv ausrechnen:

$$c_{n-1} = a_n \quad \text{und} \quad c_{k-1} = a_k + c_k\zeta \quad \forall k = n, \dots, 1.$$

Diese Prozedur heit auch **Polynomdivision** (durch $(z - \zeta)$). Die letzte Gleichung $a_0 = -c_0\zeta$ gilt mit den so bestimmten c_{n-1}, \dots, c_0 , weil ζ Nullstelle von P ist: Man multipliziert jede Gleichung mit ζ^k ($k \geq 1$) und summiert auf:

$$\text{Linke Summe} = \sum_{k=1}^n a_k \zeta^k = -a_0, \quad \text{Rechte Summe} = c_0 \zeta.$$

Fr die Division zweier Polynome $P(z)/Q(z)$ mit $Q = (z - \zeta_1) \cdot \dots \cdot (z - \zeta_n)$, folgt induktiv mittels Division durch $z - \zeta$:

$$\frac{P(z)}{Q(z)} = \frac{P(z)}{(z - \zeta_1) \cdot \dots \cdot (z - \zeta_n)} = \frac{\frac{P(z)}{(z - \zeta_1)}}{(z - \zeta_2) \cdot \dots \cdot (z - \zeta_n)} = \frac{P_1(z)}{(z - \zeta_2) \cdot \dots \cdot (z - \zeta_n)}.$$

Konjugiert komplexe Nullstellen bei reellen Koeffizienten

Lemma 7.5. *Es sei $P(z) = \sum_{k=0}^n a_k z^k$ und $a_k \in \mathbb{R} \forall k$. Dann folgt aus $P(z) = 0$ fr die konjugierte \bar{z} auch $P(\bar{z}) = 0$.*

Beweis. Mit $z = r(\cos \varphi + i \sin \varphi)$ bedeutet $P(z) = 0$:

$$P(z) = \sum_k a_k r^k \cos(k\varphi) + i \sum_k a_k r^k \sin(k\varphi) = 0.$$

Wegen $a_k \in \mathbb{R}$ reprsentieren die Summen gerade den Real- bzw. den Imaginrteil der komplexen Zahl $P(z)$. Sie mssen also beide verschwinden. Damit wird

$$-i \sum_k a_k r^k \sin(k\varphi) = i \sum_k a_k r^k \sin(-k\varphi) = 0 \quad \text{und} \quad \sum_k a_k r^k \cos(-k\varphi) = \sum_k a_k r^k \cos(k\varphi) = 0.$$

Beides zusammen bedeutet: Auch \bar{z} , nmlich $r(\cos \varphi - i \sin \varphi) = r(\cos(-\varphi) + i \sin(-\varphi))$, ist Nullstelle. \square

7.3 Komplexe Differenzierbarkeit

Identifiziert man komplexe Zahlen mit Punkten des \mathbb{R}^2 , so werden lineare Abbildungen $L : \mathbb{C} \rightarrow \mathbb{C}$ durch $(2, 2)$ -Matrizen A definiert. Sie ordnen der komplexen Zahl $z = x + iy$ die Zahl $L(z) = \alpha + i\beta$ mit

$$\begin{aligned} \alpha &= L_1(z) = a_{11}x + a_{12}y \quad (\text{Realteil}) \\ \beta &= L_2(z) = a_{21}x + a_{22}y \quad (\text{Imaginrteil}) \end{aligned}$$

zu. Dann ist also

$$L(z) = L_1(x, y) + iL_2(x, y) := L_1(z) + iL_2(z).$$

Diese Abbildungen sind (bzgl. der durch $|z|$ erklärten Norm = Euklidische Norm im \mathbb{R}^2) stetig und additiv:

$$(1) \quad L(\zeta + z) = L(\zeta) + L(z) \quad \forall \zeta, z \in \mathbb{C}.$$

Außerdem sind sie homogen in Bezug auf die Multiplikation mit reellen Zahlen:

$$(2) \quad L(rz) = rL(z) \quad \forall r \in \mathbb{R}, \forall z \in \mathbb{C}.$$

Man kann daher differenzierbare $f : \mathbb{C} \rightarrow \mathbb{C}$ und differenzierbare Funktionen $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ miteinander identifizieren, solange neben (1) nur die Homogenität bzgl. \mathbb{R} , d.h. (2), verlangt wird.

Allerdings (und das ist interessanter) kann man jetzt stärker fordern, dass Homogenität (2) auch in Bezug auf alle komplexen r (wir nennen sie w) gelten soll, also

$$(3) \quad L(wz) = wL(z) \quad \forall w, z \in \mathbb{C}.$$

Diese Zusatzforderung kann die Menge der linearen Abbildungen L - und damit auch die der differenzierbaren Funktionen- verkleinern. Wir werden sehen, dass dies tatsächlich passiert. Zur Unterscheidung nennen wir die linearen Abbildungen $L : \mathbb{C} \rightarrow \mathbb{C}$, die (3) erfüllen, komplex linear.

Definition 7.6. Ist $f : \mathbb{C} \rightarrow \mathbb{C}$ differenzierbar in z (im \mathbb{R}^2 -Sinne) und die Ableitung $Df(z)$ sogar komplex linear, so nennen wir f in z komplex differenzierbar.

Um zu sehen, was (3) bedeutet, schreiben wir $w = u + iv$. Angenommen (3) gilt für alle rein komplexen $w' = iv$. Dann folgt (3) auch für alle w wegen (1) und (2), denn

$$L(wz) = L((u + iv)z) = L(uz + ivz) = L(uz) + L(ivz) = uL(z) + (iv)L(z) = wL(z).$$

Es reicht also, (3) allein für rein komplexe $w' = iv$ zu studieren. Deshalb schreiben wir jetzt $w = iv$ und rechnen beide Seiten von (3) aus. Mit $z = x + iy$ ist dann $wz = (-yv) + i(xv)$. Also soll gelten:

$$\begin{aligned} L(wz) &= L_1(wz) + iL_2(wz) = a_{11}(-yv) + a_{12}(xv) + i[a_{21}(-yv) + a_{22}(xv)], \\ wL(z) &= (iv)(L_1(z) + iL_2(z)) = -vL_2(z) + i[vL_1(z)] \\ &= -v(a_{21}x + a_{22}y) + i[v(a_{11}x + a_{12}y)]. \end{aligned}$$

Beide Ausdrücke sollen für alle z und $w = iv$, also für alle reellen v, x, y identisch sein. Das heißt

$$\begin{aligned} a_{11}(-yv) + a_{12}(xv) &= -v(a_{21}x + a_{22}y), \\ a_{21}(-yv) + a_{22}(xv) &= v(a_{11}x + a_{12}y). \end{aligned}$$

Die Gleichungen gelten offenbar genau dann für alle reellen v, x, y , wenn

$$\begin{aligned} a_{12}xv &= -a_{21}xv \quad \text{und} \quad a_{11}xv = a_{22}xv && \text{sowie} \\ a_{21}yv &= -a_{21}yv \quad \text{und} \quad a_{11}yv = a_{22}yv. \end{aligned}$$

gilt, also wenn

$$(4) \quad a_{12} = -a_{21} \quad \text{und} \quad a_{11} = a_{22}$$

richtig ist. Das heißt, es gilt das wichtige

Lemma 7.7. Die komplex-linearen Abbildungen von \mathbb{C} in \mathbb{C} werden genau durch diejenigen $(2, 2)$ -Matrizen A definiert, die (4) erfüllen. Sie sind schon durch die erste Spalte $(a_{11}, a_{21})^T$ bestimmt, die man auch als komplexe Zahl $c = a_{11} + i a_{21}$ lesen darf.

Es sei nun $f : \mathbb{C} \rightarrow \mathbb{C}$ eine komplex differenzierbare Funktion. Wir schreiben sie in der Form

$$f(z) = f_1(z) + if_2(z) = f_1(x, y) + if_2(x, y), \quad z = x + iy.$$

Ist A die der linearen Abbildung $Df(z^*)$ zugeordnete Matrix (z^* fixiert), d.h. ist

$$(5) \quad Df(z^*) \begin{pmatrix} x \\ y \end{pmatrix} = (a_{11}x + a_{12}y) + i(a_{21}x + a_{22}y) \quad \forall z = x + iy,$$

so gilt einerseits (4) und andererseits wegen der Gestalt von Ableitungen für Funktionen $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ in jedem festen Punkt $z^* = (x^*, y^*)$

$$\begin{aligned} a_{11} &= \frac{\partial f_1}{\partial x}, & a_{12} &= \frac{\partial f_1}{\partial y}, \\ a_{21} &= \frac{\partial f_2}{\partial x}, & a_{22} &= \frac{\partial f_2}{\partial y}. \end{aligned}$$

Damit sind die Bedingungen

$$(6) \quad \frac{\partial f_1}{\partial x} = \frac{\partial f_2}{\partial y} \quad \text{und} \quad \frac{\partial f_1}{\partial y} = -\frac{\partial f_2}{\partial x}$$

notwendig für komplexe Differenzierbarkeit von f in z^* . Sind darüberhinaus die partiellen Ableitungen stetig als Funktion von z^* , so ist f tatsächlich differenzierbar (im \mathbb{R}^2 -Sinn) und wegen (4) bzw. (6) auch komplex differenzierbar. Die Gleichungen (6) heißen *Cauchy-Riemannsche (Differential-) Gleichungen*.

Die komplexe Ableitung als komplexe Zahl:

Die komplexe Ableitung von f wird schon durch die Elemente a_{11} und a_{21} **der ersten Spalte von A** repräsentiert; sie entsprechen der Ableitung von $f = (f_1, f_2)$ nach der Realteil-Variablen x . Die partiellen Ableitungen nach y sind durch (6) festgelegt. Bildet man wie oben

$$c = a_{11} + i a_{21},$$

ist die rechte Seite in (5) das Produkt cz , denn

$$cz = (a_{11} + ia_{21})(x + iy) = (a_{11}x - a_{21}y) + i(a_{11}y + a_{21}x) = (a_{11}x + a_{12}y) + i(a_{21}x + a_{22}y).$$

Man vereinbart deshalb, dass für eine komplex differenzierbare Funktion f die komplexe Zahl

$$c = \frac{\partial f_1}{\partial x}(z^*) + i \frac{\partial f_2}{\partial x}(z^*)$$

als Ableitung $f'(z^*)$ von f in z^* verstanden wird. Hiermit gilt erneut die Approximationsbeziehung

$$(7) \quad f(z^* + z) = f(z^*) + f'(z^*)z + o(z)$$

und $f'(z^*)z$ bedeutet nun die komplexe Multiplikation von $f'(z^*)$ mit z . Es folgt so wie im Reellen aus (7) und $\lim_{z \rightarrow 0} \frac{o(z)}{z} = 0$:

$$(8) \quad \lim_{z \rightarrow 0} \frac{f(z^* + z) - f(z^*)}{z} = f'(z^*) + \lim_{z \rightarrow 0} \frac{o(z)}{z} = f'(z^*).$$

Das motiviert die gewöhnliche Limes-Definition der Ableitung komplexer Funktionen.

Aus letzterer sieht man aber noch nicht die Darstellung mittels partieller Ableitungen und die Cauchy-Riemannschen (Differential-) Gleichungen, kurz den Zusammenhang zu \mathbb{R}^2 -Ableitungen.

Geht man von (8) aus, so nimmt man die Existenz und Eindeutigkeit des Limes

$$f'(z^*) := \lim_{z \rightarrow 0} \frac{f(z^* + z) - f(z^*)}{z}$$

für alle Folgen $z = z_k \rightarrow 0$, $z \neq 0$ an. Multiplikation mit der Konjugierten führt zu reeller Division

$$\frac{f(z^* + z) - f(z^*)}{z} = \frac{(f(z^* + z) - f(z^*))\bar{z}}{|z|^2}.$$

Den Zähler kann man wieder mittels f_1 und f_2 in Real- und Imaginärteil $R(z)$ und $I(z)$ aufspalten.

Die Überlegung, was die Konvergenz beider Quotienten bedeutet, führt dann ebenfalls zu den Gleichungen (6).

Welche Funktionen sind komplex differenzierbar ?

Test: Man hat partiell zu differenzieren und nachzusehen, ob die partiellen Ableitungen stetig sind und die Cauchy-Riemannschen Gleichungen (6) erfüllen.

Falls ja, liegt komplexe Differenzierbarkeit vor. Falls (6) nicht gilt, so nicht. Der Fall, dass (6) gilt, aber die partiellen Ableitungen nicht stetig sind, bleibt für uns offen.

Beispiel:

$f(z) = z^2 = (x + iy)^2 = x^2 - y^2 + 2ixy$. Also ist $f_1 = x^2 - y^2$ die Realteilfunktion und $f_2 = 2xy$ die Imaginärteilfunktion. Damit gilt

$$\begin{aligned} \frac{\partial f_1}{\partial x} &= 2x, & \frac{\partial f_1}{\partial y} &= -2y, \\ \frac{\partial f_2}{\partial x} &= 2y, & \frac{\partial f_2}{\partial y} &= 2x, \end{aligned}$$

also ist (6) erfüllt und die Ableitung ist

$$f'(z^*) = \frac{\partial f_1}{\partial x}(z^*) + i \frac{\partial f_2}{\partial x}(z^*) = 2x^* + 2iy^* = 2z^*.$$

Beispiel 2:

$f(z) = \dots$

Auch lassen sich Produkt- und Kettenregel unmittelbar auf den Fall der komplexen Differenzierbarkeit übertragen. Man beachte einfach, dass die komplexe Ableitung nur die Ableitung einer Funktion von \mathbb{R}^2 in \mathbb{R}^2 mit spezieller Gestalt (nämlich (6)) ist!

Damit gilt z.B. auch $(z^3)' = (z \cdot z^2)' = 1 \cdot z^2 + z \cdot 2z = 3z^2$ und analog $(z^n)' = nz^{n-1}$.

Somit ist jedes komplexe Polynom der Ordnung n ,

$$P(z) = a_0 + a_1z^1 + \dots + a_nz^n \quad (a_n \neq 0)$$

auch komplex differenzierbar. Dasselbe erhält man (über die Reihendarstellung und Konvergenz von Potenzreihen) für die Funktionen e^z , \sin^z , \dots , $\sinh z$, \dots

Kapitel 8

Bestimmte Integrale und Hauptsatz der Differential- und Integralrechnung

8.1 Das bestimmte Integral für eine stetige Funktion

Es sei $f : [a, b] \rightarrow \mathbb{R}$ eine reellwertige Funktion.

Um so etwas wie die Fläche zwischen der Kurve $f = f(x)$ und der x -Achse (bezogen auf $a \leq x \leq b$) zu definieren, betrachten wir eine Unterteilung des Intervalls

$$a = x_0 < x_1 < x_2 < \dots < x_n = b$$

in endlich viele Teilintervalle $I_k = [x_k, x_{k+1}]$ der Länge $\Delta x_k := x_{k+1} - x_k$ ($0 \leq k < n$).

Die Länge $\delta = \Delta x_k$ des/eines größten Teilintervalls charakterisiert die Feinheit der Unterteilung.

Kleines δ verlangt natürlich große n .

Weiter wählen wir zu jedem k einen Punkt $\xi_k \in [x_k, x_{k+1})$ und bilden die (Integral-) Summe

$$(1) \quad \sigma = \sum_{k=0}^{n-1} f(\xi_k) \Delta x_k.$$

Offenbar hängt σ ab von der Feinheit δ , der Lage und Anzahl der Punkte x_k und der Wahl der ξ_k .

Definition 8.1 (bestimmtes Integral). *Existiert ein gemeinsamer Grenzwert J für alle Summen σ , wenn $\delta \rightarrow 0$ strebt, und zwar für jede erlaubte Wahl der x_k und ξ_k , so heißt J das bestimmte Integral von f über $[a, b]$.*

Kurzschreibweise:

$$J = \int_a^b f(x) dx.$$

Diese Definition stammt von B. Riemann; es gibt viele weitere Integraldefinitionen. Funktionen f , für welche J existiert, heißen R -integrierbar (oft auch einfach integrierbar).

Satz 8.2 (Existenz des bestimmten Integrals). *Das Integral $J = \int_a^b f(x) dx$ existiert für jede auf $[a, b]$ stetige Funktion f .*

Beweis. Die Werte der stetigen Funktion f sind beschränkt $|f(\xi)| \leq C \quad \forall \xi \in [a, b]$. Für jede beschränkte Funktion f sind außerdem die Integralsummen beschränkt:

$$|\sigma| \leq C(b - a).$$

Für jede Verfeinerung der Unterteilung $\delta \rightarrow 0$ (d.h. δ durchläuft irgendeine positive Nullfolge δ_ν) hat deshalb wenigstens eine Teilfolge der Integralsummen einen Grenzwert (siehe Satz von Bolzano und

Weierstraß, Kap.2.10). Es bleibt zu zeigen, dass dieser Grenzwert immer derselbe ist: Die stetige Funktion f ist gleichmäßig stetig auf $[a, b]$. Also findet sich zu jedem $\varepsilon > 0$ ein $\delta > 0$, sodass

$$(2) \quad |f(\xi') - f(\xi)| < \varepsilon, \text{ wenn nur } \xi', \xi \in [a, b] \text{ und } |\xi' - \xi| < \delta \text{ gilt.}$$

Wir betrachten zwei beliebige Integralsummen σ und σ' , beide mit Feinheit $\leq \delta$ und mit unterschiedlichen Punkten x_k ($0 \leq k \leq n$) und $x'_{k'}$ ($0 \leq k' \leq n'$) sowie ausgewählten Elementen ξ_k und $\xi'_{k'}$.

Aus den $n + n' + 2$ Zerlegungspunkten (die auch zusammenfallen können) bilden wir die (verfeinerte) Unterteilung $a = y_0 < y_1 < \dots < y_m < \dots < y_N = b$, wählen $\lambda_m \in (y_m, y_{m+1})$ und bilden die (Integral-) Summe

$$s = \sum_m f(\lambda_m) \Delta y_m \quad (m = 0, 1, \dots, N - 1).$$

Ein (Original-) Intervall $I_k = [x_k, x_{k+1}]$ erscheint in der y -Zerlegung entweder unverändert oder wird weiter unterteilt durch einen oder mehrere Punkte $x'_{k'}$.

Der Summand $f(\xi_k) \Delta x_k$ zum Intervall I_k in der Original-Integralsumme σ in (1) ist in s ersetzt durch eine Teilsumme s_k , die der weiteren Unterteilung von I_k entspricht:

$$s_k = \sum_p f(\lambda_p) \Delta y_p, \text{ wobei } m' \leq p \leq m'' \text{ mit } y_{m'} = x_k \text{ und } y_{m''+1} = x_{k+1}.$$

Für alle diese p gilt $|\lambda_p - \xi_k| < \delta$, also wegen (2) auch

$$|f(\lambda_p) - f(\xi_k)| < \varepsilon.$$

Außerdem ist $\sum_p \Delta y_p = \Delta x_k$. Beides zusammen liefert uns

$$|f(\xi_k) \Delta x_k - s_k| < \varepsilon \Delta x_k.$$

Summiert man über alle k , erhält man

$$|\sigma - s| < \varepsilon \sum_k \Delta x_k = \varepsilon(b - a).$$

Denselben Vergleich können wir mit s und σ' anstellen. Damit unterscheiden sich die zwei Integralsummen σ und σ' beliebig wenig

$$|\sigma - \sigma'| \leq |\sigma - s| + |s - \sigma'| < 2\varepsilon(b - a),$$

wenn nur δ klein genug ist. Folglich ist der Limes aller Integralsummen für $\delta \rightarrow 0$ eindeutig. \square

Eigenschaften des Integrals

Aus den Eigenschaften der Integralsummen (1) folgt über Grenzwertbildung mit $h(x) = f(x) + g(x)$, sofern die Integrale zu f und g existieren:

$$\begin{aligned} \int_a^b h(x) dx &= \int_a^b f(x) dx + \int_a^b g(x) dx \\ \int_a^b r f(x) dx &= r \int_a^b f(x) dx \quad \forall r \in \mathbb{R}. \end{aligned}$$

Ist $c \in (a, b)$, so gilt außerdem mit dem zusammengesetzten Intervall $[a, b] = [a, c] \cup [c, b]$:

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Hieran ändert auch eine endliche Unstetigkeits- (Sprung-) stelle von f in c nichts.

Außerdem sieht man anhand der Integralsummen, dass

$$(3) \quad (b-a)f_{\inf} \leq \int_a^b f(x) dx \leq (b-a)f_{\sup}, \quad (a < b)$$

wenn $f_{\text{sup}} = \sup\{f(x) \mid x \in [a, b]\}$ und $f_{\text{inf}} = \inf\{f(x) \mid x \in [a, b]\}$.

Ist f stetig, so gibt es zu jedem $r \in [f_{\text{inf}}, f_{\text{sup}}]$ ein $\xi \in [a, b]$ mit $f(\xi) = r$. Insbesondere zu

$$r = \frac{1}{b-a} \int_a^b f(x) dx.$$

Als Folgerung ergibt sich damit der

Satz 8.3 (Mittelwertsatz der Integralrechnung). *Für stetiges f existiert stets ein $\xi \in [a, b]$ mit $\int_a^b f(x) dx = (b-a)f(\xi)$.*

Für $b < a$ definiert man $\int_a^b f(x) dx = -\int_b^a f(x) dx$, was der Integralsumme mit negativen Δx_k entspricht.

Nun bleibt $\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx$ für beliebige a, b, c richtig, sofern f auf ganz \mathbb{R} definiert und stetig ist.

Variable obere Grenze

Wir denken uns nun f als stetig auf \mathbb{R} und variieren die obere Grenze b des Integrals:

$$J(b) = \int_a^b f(x) dx.$$

Satz 8.4 (Ableitung des bestimmten Integrals). *Ist f stetig auf \mathbb{R} , so ist $J = J(b)$ differenzierbar und $DJ(b) = J'(b) = f(b)$.*

Beweis. Der Satz gilt sowohl für $a < b$ als auch $b \leq a$. Wir betrachten hier nur den ersten Fall; im zweiten muss man nur den Vorzeichenwechsel beachten. Es ist $J(b+\delta) - J(b) = \int_b^{b+\delta} f(x) dx$.

Damit folgt aus (3)

$$f_{\text{inf}}\delta \leq J(b+\delta) - J(b) \leq f_{\text{sup}}\delta,$$

wobei $f_{\text{sup}} = \sup\{f(x) \mid x \in [b, b+\delta]\}$ und $f_{\text{inf}} = \inf\{f(x) \mid x \in [b, b+\delta]\}$.

Für $\delta \rightarrow +0$ streben beide Werte gegen $f(b)$ wegen der Stetigkeit von f in b . Also gilt auch

$$\lim_{\delta \rightarrow +0} \frac{J(b+\delta) - J(b)}{\delta} = f(b).$$

Beachtet man die Definition von $\int_b^{b+\delta} f(x) dx$ für negative δ durch Vorzeichenwechsel, ist auch

$$\lim_{\delta \rightarrow -0} \frac{J(b+\delta) - J(b)}{\delta} = f(b)$$

richtig. Beides zusammen liefert die Behauptung. \square

Stammfunktionen

Seien $F, f : \mathbb{R} \rightarrow \mathbb{R}$ stetige Funktionen. Man nennt F Stammfunktion für f , wenn $F'(y) = f(y)$ für alle y gilt. Dann schreibt man auch

$$F(x) = \int f(x) dx$$

und nennt $\int f(x) dx$ ein *unbestimmtes Integral*.

Natürlich muss dazu F differenzierbar sein. Stammfunktionen sind nicht eindeutig bestimmt, denn ist $F = F(y)$ eine, so auch $G(y) = F(y) + C$ mit jeder Konstanten C . Wir zeigen, dass sich Stammfunktionen *nur* durch eine Konstante unterscheiden.

Lemma 8.5. *Sind F und G Stammfunktionen für f , so ist $F - G$ eine konstante Funktion.*

Beweis. Offenbar erfüllt $H = F - G$:

$$H'(y) = F'(y) - G'(y) = f(y) - f(y) = 0.$$

Wenn $H(y) = C \forall y$ ist, sind wir fertig. Sei also $H(a) \neq H(b)$ für gewisse $a \neq b$. Nach dem Mittelwertsatz der Differentialrechnung gibt es ein Θ zwischen a und b , sodass

$$H(b) - H(a) = H'(\Theta)(b - a).$$

Aus $H'(\Theta) = 0$ erhalten wir dann aber $H(b) - H(a) = 0$, also einen Widerspruch. \square

Eine Stammfunktion für f (stetig) finden wir immer, nämlich die in y durch das Integral

$$(6) \quad G(y) = \int_0^y f(x) dx$$

definierte Funktion. Sie erfüllt ja nach dem gerade bewiesenen Satz 8.4

$$G'(y) = f(y) \quad \forall y.$$

Damit kommen wir nun zum

Satz 8.6 (Hauptsatz der Differential- und Integralrechnung). *Für jede stetige reelle Funktion f und jede Stammfunktion F für f gilt*

$$(7) \quad \int_a^b f(x) dx = F(b) - F(a).$$

Beweis. Mit der speziellen Stammfunktion $G(y) = \int_0^y f(x) dx$ gilt

$$\int_a^b f(x) dx = \int_0^b f(x) dx - \int_0^a f(x) dx = G(b) - G(a).$$

Da jede Stammfunktion F für f sich von G nur durch eine Konstante unterscheidet, folgt auch

$$\int_a^b f(x) dx = F(b) - F(a).$$

\square

Der Satz verknüpft über den Begriff der Stammfunktion die Differential- und die Integralrechnung. Er lässt aber offen, wie man im konkreten Fall eine Stammfunktion finden kann, denn die Berechnung von G nach (6) ist nicht immer einfach.

Wir wollen ja eigentlich mittels einer Stammfunktion das bestimmte Integral $\int_a^b f(x) dx$ ermitteln.

Für die Differenz $F(b) - F(a)$ schreibt man (im context von Integralen) oft $[F(x)]_a^b$.

8.2 Berechnung von Stammfunktionen und Integrationsmethoden

In diesem Abschnitt geben wir Methoden an, um Stammfunktionen zu finden.

Die Wichtigste besteht darin, Ableitungen von Standardfunktionen zu kennen und damit natürlich auch viele Grundintegrale; siehe Wichtige Ableitungen, Kap. 6.1.

Allerdings gibt es auch verschiedene Hilfsmittel, die wir jetzt betrachten wollen.

8.2.1 Partielle Integration

(Umkehrung der Produktregel der Differentiation)

Ausgehend von der Produktregel für Differentiation

$$(u(x)v(x))' = u'(x)v(x) + u(x)v'(x),$$

wissen wir, dass für die Stammfunktionen der einzelnen Terme (bis auf eine Konstante) gilt

$$u(x)v(x) = \int u'(x)v(x)dx + \int u(x)v'(x)dx.$$

Das erlaubt oft, eine Stammfunktion nach der Gleichung

$$\int u'(x)v(x)dx = u(x)v(x) - \int u(x)v'(x)dx$$

zu finden, wenn sich das rechte Integral als einfach erweist und die zu integrierende Funktion f sich als $f(x) = u'(x)v(x)$ schreiben lässt ...

Beispiele 8.7.

1) $\int \ln x \, dx = \int 1 \cdot \ln x \, dx = x(\ln x - 1).$

Man setze $u'(x) = 1$ und $v(x) = \ln x$. Dann wird $u(x) = x$ und $v'(x) = \frac{1}{x}$ (für $x > 0$), und wegen $uv' = 1$,

$$\int \ln x \, dx = x \ln x - \int 1 \, dx = x \ln x - x = x(\ln x - 1).$$

Probe: Differenziert man $F(x) = x(\ln x - 1)$ nach der Produktregel, erhält man tatsächlich

$$F'(x) = (\ln x - 1) + x \left(\frac{1}{x} \right) = \ln x.$$

Bemerkung: Setzt man hier andersherum $u' = \ln x$ und $v = 1$, so braucht man u fuer die rechte Seite von $\int u'v \, dx = uv - \int v'u \, dx \quad (+C)$. Das hilft uns nichts.

2) $\int xe^x \, dx = (x - 1)e^x.$

Sei $u' = e^x$ und $v = x$. Dann folgt $u = e^x$ und $v' = 1$, $uv = xe^x$. Also

$$\int xe^x \, dx = xe^x - \int e^x \, dx = (x - 1)e^x.$$

u' und v sind klar bei:

3) $\int \sin^2 x \, dx = \frac{1}{2}(x - \sin x \cos x)$

4) $\int \cos^2 x \, dx = \frac{1}{2}(x + \sin x \cos x)$

5) $\int \sinh^2 x \, dx = \frac{1}{2}(\sinh x \cosh x - x)$

6) $\int \cosh^2 x \, dx = \frac{1}{2}(\sinh x \cosh x + x)$ (Nachrechnen !)

7) $\int e^x \sin x \, dx = \frac{1}{2}e^x(\sin x - \cos x)$

Sei $u' = e^x$ und $v = \sin x$. Dann folgt $u = e^x$, $v' = \cos x$. Also

$$(*) \quad \int e^x \sin x \, dx = e^x \sin x - \int e^x \cos x \, dx.$$

Nochmals partiell integrieren: Mit $u' = e^x$ und $v = \cos x$, $v' = -\sin x$ wird

$$\begin{aligned} \int e^x \cos x \, dx &= e^x \cos x - \int -e^x \sin x \, dx, \\ \Leftrightarrow -\int e^x \cos x \, dx &= -e^x \cos x - \int e^x \sin x \, dx, \\ (*) \Rightarrow \int e^x \sin x \, dx &= e^x \sin x - e^x \cos x - \int e^x \sin x \, dx \\ \Rightarrow \int e^x \sin x \, dx &= \frac{1}{2}e^x(\sin x - \cos x). \end{aligned}$$

Probe

$$[e^x(\sin x - \cos x)]' = e^x(\sin x - \cos x) + e^x(\cos x + \sin x) = 2e^x \sin x.$$

- 8) $\int x \arctan x \, dx = \frac{1}{2}(x^2 \arctan x - x + \arctan x)$.
 Setze $u' = x$, $v = \arctan x$. Dann wird

$$\begin{aligned} \int x \arctan x \, dx &= \frac{1}{2} x^2 \arctan x - \int \frac{1}{2} \frac{x^2}{1+x^2} dx \\ &= \frac{1}{2} x^2 \arctan x - \frac{1}{2} \int \frac{x^2 + 1 - 1}{1+x^2} dx \\ &= \frac{1}{2} x^2 \arctan x - \frac{1}{2} \int 1 - \frac{1}{1+x^2} dx \\ &= \frac{1}{2} x^2 \arctan x - \frac{1}{2} (x - \arctan x). \end{aligned}$$

Zur Übung: $\int \arcsin x \, dx$, $\int \arctan x \, dx$, $\int x \operatorname{artanh} x \, dx$

$$\begin{aligned} \int x \sin x \, dx &= \int vu' \, dx = x(-\cos x) - \int -\cos x \, dx = \sin x - x \cos x. \\ \int x \cos x \, dx &= \int vu' \, dx = x(\sin x) - \int \sin x \, dx = \cos x + x \sin x. \\ \int x \cos(kx) \, dx &= \int vu' \, dx = \frac{x \sin(kx)}{k} - \int \frac{\sin(kx)}{k} \, dx = \frac{x \sin(kx)}{k} + \frac{\cos(kx)}{k^2}. \end{aligned}$$

Integral bei den Grenzen 0 und π ist $-2k^{-2}$ für ungerade k und 0 für gerade k .
 Die Fourier-Reihe für $|x|$ wird so über $[-\pi, \pi]$:

$$|x| = \frac{1}{2}\pi - 4\pi^{-1} \sum_k k^{-2} \cos(kx) \quad (\text{Summe über ungerade } k = 1, 3, 5, \dots).$$

8.2.2 Substitutionsmethode

(Umkehrung der Kettenregel der Differentiation)

Herleitung: Wir betrachten $\Phi(x) = f(g(x))$ mit stetig differenzierbaren f und g , und das bestimmte Integral

$$I = \int_a^b f'(g(x))g'(x) \, dx.$$

Nach Kettenregel ist $\Phi(x)$ eine Stammfunktion des Integranden, denn $\Phi'(x) = f'(g(x))g'(x)$.
 Also gilt nach Hauptsatz:

$$I = \Phi(b) - \Phi(a) = f(g(b)) - f(g(a)).$$

Andererseits ist, mit $u = g(a)$ und $v = g(b)$, auch

$$I = f(g(b)) - f(g(a)) = \int_u^v f'(y) \, dy,$$

$$\text{womit folgt: } \int_a^b f'(g(x))g'(x) \, dx = \int_u^v f'(y) \, dy.$$

Hier war f' irgendeine stetige Funktion mit Stammfunktion f . Ersetzt man f' durch eine stetige Funktion h , folgt deshalb

$$(*) \quad \int_a^b h(g(x))g'(x) \, dx = \int_u^v h(y) \, dy, \quad \text{wenn } u = g(a) \text{ und } v = g(b).$$

Die Integranden entstehen *formal* auseinander, indem man rechts $y = g(x)$ setzt sowie in der Schreibweise $\frac{dy}{dx} = g'(x)$ für die Ableitung mit dx multipliziert, $dy = g'(x)dx$, und rechts einsetzt.

Die Integralgrenzen werden wie oben mittels g ersetzt (Wir brauchen nur, dass es a und b mit $u = g(a)$ und $v = g(b)$ tatsächlich gibt und g stetig differenzierbar ist). In dieser Weise gelangt man am einfachsten zum linken Integral von (*).

ABER: Das „Multiplizieren“ mit dx war nur wegen der Integral- und Ableitungs-Schreibweise möglich (weil man sich bei der Schreibweise des Integrals und der Ableitung etwas gedacht hat), es ist kein Beweis

für die Substitutionsformel! Trotzdem führt es zum Ergebnis, weil (*) wegen der Kettenregel gilt (für stetiges h und stetig differenzierbares g).

Bemerkung: Meist sind beim „normalen Rechnen“ und in vielen Formeln die Rollen von x , y sowie (a, b) und (u, v) vertauscht und h heißt f , was zu $x = g(y)$ und

$$(**) \quad \int_a^b f(x) dx = \int_u^v f(g(y))g'(y) dy \quad \text{mit} \quad a = g(u) \text{ und } b = g(v).$$

führt. Man beachte, dass sich $a = g(u)$ und $b = g(v)$ nicht mit jeder (stetig differenzierbaren) Funktion g und gegebenen a, b erfüllen lassen.

Außerdem hat das Ersetzen von x durch $g(y)$ sicher nur Sinn, wenn das zweite Integral einfacher wird. Um zu sehen, wann dies mit welchem g passiert, braucht man einige Erfahrung und Übung. Oft nennt man die neue Variable t statt y .

Beispiel 8.8. Zu berechnen sei $J = \int_2^9 \sqrt{3x-4} dx$. Wir setzen $3x-4 = t$, um \sqrt{t} zu erhalten.

Dann ist $x = \frac{t+4}{3} = g(t)$ und $\frac{dx}{dt} = \frac{1}{3}$, also wird dx durch $\frac{1}{3}dt$ ersetzt.

Die neuen Grenzen müssen erfüllen:

$$\begin{aligned} a = 2 = g(u) = \frac{u+4}{3} &\quad \Rightarrow \quad u = 6 - 4 = 2 \\ b = 9 = g(v) = \frac{v+4}{3} &\quad \Rightarrow \quad v = 27 - 4 = 23 \end{aligned}$$

Damit wird

$$J = \frac{1}{3} \int_2^{23} \sqrt{t} dt = \frac{1}{3} \int_2^{23} t^{\frac{1}{2}} dt = \frac{1}{3} \left[\frac{2}{3} t^{\frac{3}{2}} \right]_2^{23} = \frac{2}{9} \left(23^{\frac{3}{2}} - 2^{\frac{3}{2}} \right).$$

8.2.3 Anwendungen der Substitutionsmethode

Kreisfläche

Zu berechnen sei die Fläche eines Kreises mit Radius r .

$$f(x) = \sqrt{r^2 - x^2}, \quad x \in [0, r],$$

Fläche zwischen f und x -Achse ist $A = \int_0^r \sqrt{r^2 - x^2} dx$ und bildet $\frac{1}{4}$ der Kreisfläche.

Wir setzen mit einer neuen Variablen t , $x = r \sin t$ ($= g(t)$), $t \in [u, v] = [0, \frac{\pi}{2}]$.

Dann ist $\frac{dx}{dt} = r \cos t$ und $\sqrt{r^2 - x^2} = r \sqrt{1 - \sin^2 t} = r \cos t$. Wir erhalten

$$A = r^2 \int_0^{\frac{\pi}{2}} \cos^2 t dt$$

und mit partieller Integration (s.o.)

$$A = r^2 \left[\frac{1}{2} (x + \sin x \cos x) \right]_0^{\frac{\pi}{2}}$$

Damit folgt $A = r^2 \frac{\pi}{4}$. Die Fläche des Kreises ist also πr^2 .

Kugelvolumen

Zu berechnen sei das Volumen einer Kugel mit Radius r .

$$f(x) = \sqrt{r^2 - x^2}, \quad x \in [0, r],$$

Rotieren ergibt die Halbkugel und das Volumen

$$\begin{aligned} V_{\text{halb}} &= \pi \int_0^r f(x)^2 dx = \pi \int_0^r (r^2 - x^2) dx = \pi r^2 \int_0^r 1 dx - \pi \int_0^r x^2 dx \\ &= \pi r^3 - \pi \left[\frac{x^3}{3} \right]_0^r = \pi r^3 - \pi \frac{r^3}{3} = 2\pi \frac{r^3}{3}. \end{aligned}$$

Also ist das ganze Kugelvolumen $V = 2V_{\text{halb}} = \frac{4}{3}\pi r^3$.

Bemerkung: Die Integrale von $\sin^2 x$, $\cos^2 x$, $\sinh^2 x$, $\cosh^2 x$ sind für Substitution bei Ausdrücken der Form $\sqrt{a^2 + x^2}$ und $\sqrt{a^2 - x^2}$ wichtig.

Fläche der Ellipse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1, \quad |x| \leq a, \quad |y| \leq b.$$

Nun ist

$$\frac{y}{b} = \sqrt{1 - \frac{x^2}{a^2}}, \quad y = \frac{b}{a} \sqrt{a^2 - x^2}.$$

Mit $J = \int_{-a}^a \sqrt{a^2 - x^2} dx$ wird die Fläche der Ellipse gerade $A = 2 \frac{b}{a} J$. Man setze nun wie beim Kreis

$$x = a \sin t, \quad dx = a \cos t dt; \quad t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Dann wird

$$J = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} a \cos t \sqrt{a^2 - a^2 \sin^2 t} dt = a^2 \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2 t dt = \frac{a^2}{2} [t + \sin t \cos t]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \frac{a^2}{2} \pi.$$

Die Fläche selbst wird dann $A = \frac{2b}{a} J = ab \pi$.

Fläche der Einheitshyperbel

$$x^2 - y^2 = 1, \quad y = \sqrt{x^2 - 1}, \quad 1 \leq x < b, \quad A = 2 \int_1^b \sqrt{x^2 - 1} dx.$$

Beachte: Nach Substitution $x = g(t)$ sollte $\sqrt{g(t)^2 - 1}$ möglichst einfach werden. Das geht jetzt nicht mit $g = \sin t$, weil erstens $\sin^2 t - 1 = -\cos^2 t$ falsches Vorzeichen hat und außerdem $b > g(t)$ gilt. Der hyperbolische Pythagoras

$$\cosh^2 t - \sinh^2 t = 1, \quad \text{d.h.,} \quad \cosh^2 t - 1 = \sinh^2 t$$

motiviert die zweckmässige Substitution

$$x = \cosh t \quad (= g(t), \text{ rechter Ast } t > 0), \quad dx = \sinh t dt.$$

Stammfunktion: $F = \int \sinh^2 t dt = \frac{1}{2}(\sinh t \cosh t - t)$;

$$A = 2 [\sinh t \cosh t - t]_u^v$$

Integrations-Grenzen u, v :

$$1 = g(u) = \cosh u, \quad b = g(v) = \cosh v;$$

d.h.,

$$u = \operatorname{arcosh}(1) = 0, \quad v = \operatorname{arcosh}(b) = \ln(b + \sqrt{b^2 - 1})$$

(der negative Zweig der ln-Darstellung $\ln(b - \sqrt{b^2 - 1})$ ist wegen $b, v > 0$ hier nicht relevant).

Dann gilt im rechten Randpunkt v :

$$\cosh v = b \quad \text{und} \quad \sinh v = \sqrt{\cosh^2 v - 1} = \sqrt{b^2 - 1},$$

und links

$$\sinh(u) = 0.$$

Das liefert:

$$\frac{1}{2} A = b\sqrt{b^2 - 1} - \operatorname{arcosh}(b) = b\sqrt{b^2 - 1} - \ln(b + \sqrt{b^2 - 1}).$$

Umrechnen von arcosh in Logarithmus-Form:

$a = \operatorname{arcosh}(b) \Leftrightarrow b = \cosh a \Leftrightarrow 2b = (e^a + e^{-a})$ (dann ist $b \geq 1$).

Weiter sei $u = e^a$. Dann kann man umformen:

$$2b = \left(u + \frac{1}{u}\right), \quad 2bu = u^2 + 1, \quad u^2 - 2bu + 1 = 0.$$

Also muss $u = b \pm \sqrt{b^2 - 1}$ sein, und somit $a = \operatorname{arcosh}(b) = \ln(b \pm \sqrt{b^2 - 1})$.
Analog zeigt man $\operatorname{arsinh}(b) = \ln(b + \sqrt{b^2 + 1})$ (nur eine Lösung).

Die Substitution $t = \tan(\frac{x}{2})$ zum Integrieren rationaler Funktionen von $\sin x$ und $\cos x$.
(= Quotienten von Polynomen in $\sin x$, $\cos x$ oder $\tan x$)

Man setze

$$t = \tan\left(\frac{x}{2}\right) \quad \text{bzw.} \quad x = 2 \arctan t.$$

Dann gilt wegen der Additionstheoreme

$$\begin{aligned} \sin(2x) &= 2 \sin x \cos x, \\ \cos(2x) &= \cos^2 x - \sin^2 x, \\ \tan(2x) &= \frac{2 \tan x}{1 - \tan^2 x}; \end{aligned}$$

die 3. Gleichung folgt nach Division durch $\cos^2 x$ aus $\tan(2x) = \frac{2 \sin x \cos x}{\cos^2 x - \sin^2 x}$. Also gilt auch

$$\begin{aligned} \tan x &= \frac{2 \tan\left(\frac{x}{2}\right)}{1 - \tan^2\left(\frac{x}{2}\right)}, \quad \text{d.h.,} \\ \tan x &= \frac{2t}{1 - t^2}. \end{aligned}$$

Schreibt man

$$\sin x = \frac{a}{1 + t^2}, \quad \cos x = \frac{b}{1 + t^2},$$

so folgt $a^2 + b^2 = (1 + t^2)^2$ und $\frac{a}{b} = \frac{2t}{1 - t^2}$.

Das definiert a und b eindeutig als $a = 2t$, $b = 1 - t^2$.

Denn jetzt ist $\frac{a}{b} = \frac{2t}{1 - t^2}$ offensichtlich, und es wird

$$a^2 + b^2 = 4t^2 + (1 - t^2)^2 = (1 + t^2)^2.$$

Also ist

$$\sin x = \frac{2t}{1 + t^2}, \quad \cos x = \frac{1 - t^2}{1 + t^2}.$$

Schließlich ist

$$\frac{dt}{dx} = \frac{1 + \tan^2\left(\frac{x}{2}\right)}{2}, \quad \text{d.h.} \quad dx = \frac{2dt}{1 + t^2}.$$

Mit dieser Substitution kann man also die Standard- Winkelfunktionen zum Verschwinden bringen. Statt dessen bekommt man, mit gewissen Faktoren, Ausdrücke der Form $\frac{t}{1-t^2}$, $\frac{t}{1+t^2}$, $\frac{1}{1+t^2}$ und

$$\frac{t^2}{1 + t^2} = \frac{1 + t^2 - 1}{1 + t^2} = 1 - \frac{1}{1 + t^2},$$

d.h. wieder Ausdrücke in Form von Konstanten und $\frac{1}{1+t^2}$.

Rationale Funktionen von $\sinh x$ und $\cosh x$.

Man setze analog $t = \tanh(\frac{x}{2})$.

Bogenlänge

Länge der Kurve zu $y = f(x)$ im Intervall $[a, b]$.

Summen der Form $L = \sum_i J_i$, mit $J_i = \sqrt{\Delta x_i^2 + \Delta y_i^2}$ sind zu betrachten, die wieder zu vorgegebenen

Zerlegungen des Intervalls gehören. Konvergieren sie gegen eine wohlbestimmte Zahl L^* mit jeder Verfeinerung jeder Zerlegung, so heißt L^* Länge der Kurve.

Angenommen f ist differenzierbar.

Dann ist nach Mittelwertsatz $\Delta y_i = f'(\xi_i) \Delta x_i$ mit einem ξ_i im i -ten Einteilungsintervall,

$$L = \sum_i \sqrt{(\Delta x_i)^2 + (f'(\xi_i) \Delta x_i)^2} = \sum_i \sqrt{1 + (f'(\xi_i))^2} \Delta x_i.$$

Ist f' auch noch stetig, so konvergieren diese Summen (nach Existenz des bestimmten Integrals für stetige Funktionen) gegen das Integral

$$I = \int_a^b \sqrt{1 + (f'(x))^2} dx.$$

Beispiel 8.9 (Länge des Kreisbogens). $y = f(x) = \sqrt{r^2 - x^2}$, $-r \leq x \leq r$. Für x im offenen Intervall gilt

$$\begin{aligned} y' &= -\frac{x}{\sqrt{r^2 - x^2}} = f'(x), \\ (y')^2 &= \frac{x^2}{r^2 - x^2} = (f'(x))^2, \\ 1 + (y')^2 &= \frac{r^2}{r^2 - x^2} = 1 + (f'(x))^2. \end{aligned}$$

Sei $a = -r + \varepsilon$, $b = r - \varepsilon$. Dann folgt damit

$$I(\varepsilon) = \int_a^b \sqrt{1 + (f'(x))^2} dx = \int_a^b \frac{r}{\sqrt{r^2 - x^2}} dx = \int_a^b \frac{dx}{\sqrt{1 - (\frac{x}{r})^2}}.$$

Mit der Substitution $x = ru$, $dx = rdu$ folgt

$$I(\varepsilon) = \int_{-1+\frac{\varepsilon}{r}}^{1-\frac{\varepsilon}{r}} \frac{r}{\sqrt{1-u^2}} du = r \arcsin u \quad \text{mit den Grenzen } \pm \left(1 - \frac{\varepsilon}{r}\right).$$

Das liefert: $I = r \left(\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right) = r\pi$ wegen $\varepsilon \rightarrow 0$, und der Kreisumfang ergibt sich als $2\pi r$.

Das Integral $I = \int \frac{Ax+B}{ax^2+bx+c} dx$. ($a = 0$ Übungsaufgabe).

1. Ausklammern von a .

$$I = \frac{1}{a} \int \frac{Ax+B}{x^2+b'x+c'} dx; \quad b' \text{ und } c' \text{ sind Konstanten!}$$

2. Nenner-Ableitung im Zähler Ausnutzen (falls $A \neq 0$)

$$I = \frac{1}{a} \frac{A}{2} \int \frac{(2x+b') + \left(\frac{2B}{A} - b'\right)}{x^2+b'x+c'} dx.$$

Das nun auftretende Integral heie J . Dann ist mit $k = \frac{2B}{A} - b'$:

$$J = \ln|x^2+b'x+c'| + k \int \frac{dx}{x^2+b'x+c'}.$$

Wir mssen noch $H := \int \frac{dx}{x^2+b'x+c'}$ behandeln:

3. Nenner mit $\beta = \sqrt{|c' - (\frac{b'}{2})^2|}$ auf quadratische Form bringen:

$$\begin{aligned} x^2 + b'x + c' &= \left(x + \frac{b'}{2}\right)^2 + \left(c' - \left(\frac{b'}{2}\right)^2\right) \\ &= (x + \alpha)^2 + \beta^2 \quad \text{wenn } c' - \left(\frac{b'}{2}\right)^2 > 0 \quad (= \text{Fall 1}) \\ \text{oder} &= (x + \alpha)^2 - \beta^2 \quad \text{wenn } c' - \left(\frac{b'}{2}\right)^2 < 0 \quad (= \text{Fall 2}) \end{aligned}$$

(der Fall $\beta = 0$ wird Übungsaufgabe)

Fall 1: Nenner $(x + \alpha)^2 + \beta^2$ in der Form

$$\beta^2 \left[\left(\frac{x + \alpha}{\beta} \right)^2 + 1 \right]$$

schreiben. Nach Substitution $u = \frac{x + \alpha}{\beta}$, $dx = \beta du$ wird

$$H = \int \frac{dx}{x^2 + b'x + c'} = \frac{\beta}{\beta^2} \int \frac{du}{1 + u^2} = \frac{1}{\beta} \arctan u.$$

Fall 2: Nenner $(x + \alpha)^2 - \beta^2$ in der Form

$$\beta^2 \left[\left(\frac{x + \alpha}{\beta} \right)^2 - 1 \right]$$

schreiben. Nach derselben Substitution $u = \frac{x + \alpha}{\beta}$, $dx = \beta du$ wird

$$H = \int \frac{dx}{x^2 + b'x + c'} = \frac{\beta}{\beta^2} \int \frac{du}{u^2 - 1} = -\frac{1}{\beta} \int \frac{du}{1 - u^2}$$

Man erhält nun Stammfunktionen $H = -\frac{1}{\beta} \operatorname{artanh} u$ bzw. $H = -\frac{1}{\beta} \operatorname{arcoth} u$, abhängig von $|u| < 1$ oder $|u| > 1$.

Zur Erinnerung:

$$(\operatorname{artanh} x)' = \frac{1}{1-x^2}; \quad (|x| < 1)$$

$$(\operatorname{arcoth} x)' = \frac{1}{1-x^2}; \quad (|x| > 1). \quad \text{Und was bei } \beta = 0 \text{ unter 3. ?}$$

8.2.4 Partialbruchzerlegung

Es sei das unbestimmte Integral

$$I = \int \frac{P(x)}{Q(x)} dx$$

zu finden, wobei P und Q Polynome mit reellen Koeffizienten vom Grade m bzw. n sind.

Ist $m \geq n$, so hat I nach Polynomdivision die Gestalt

$$I = \int G(x) dx + \int \frac{R(x)}{Q(x)} dx$$

wobei G ein Polynom und R ein Polynom vom Grade $\deg(R) < n$ ist. Da G leicht zu integrieren ist, dürfen wir daher gleich $m < n$ voraussetzen.

Q lässt sich als Produkt $Q(z) = (z - \zeta_1) \cdot \dots \cdot (z - \zeta_n)$ schreiben (nach Division durch $a_n \neq 0$).

Wir wollen den Quotienten in der Form

$$\frac{P}{Q} = \frac{A_1}{z - \zeta_1} + \dots + \frac{A_n}{z - \zeta_n}$$

schreiben (wozu natürlich $z \neq \zeta_i$ anzunehmen ist). Das bedeutet nach Multiplikation mit Q ,

$$(*) \quad P(z) = \sum_{k=0}^{n-1} b_k z^k = A_1(z - \zeta_2) \cdot \dots \cdot (z - \zeta_n) + \dots + A_n(z - \zeta_1) \cdot \dots \cdot (z - \zeta_{n-1}).$$

(jeweils $n - 1$ Produkte $(z - \zeta_j)$ werden mit A_i multipliziert. Für diesen Ansatz gibt es eine Chance.

Satz 8.10. *Sind alle (komplexen) Nullstellen ζ_1, \dots, ζ_n von Q paarweise verschieden, so gibt es eindeutige komplexe Koeffizienten A_i , die (*) erfüllen.*

Beweis. Durch Koeffizientenvergleich erhält man ein lineares Gleich.-system für die n Koeffizienten A_i :

$$\begin{aligned} b_{n-1} &= A_1 + \dots + A_n \\ b_{n-2} &= -A_1(\zeta_2 + \dots + \zeta_n) - \dots - A_n(\zeta_1 + \dots + \zeta_{n-1}) \quad (\text{Summen zu } A_i \text{ ohne } \zeta_i) \\ b_{n-3} &= A_1(\zeta_2\zeta_3 + \dots + \zeta_{n-1}\zeta_n) + \dots + A_n(\zeta_1\zeta_2 + \dots + \zeta_{n-2}\zeta_{n-1}) \quad (\text{Paare ohne } \zeta_i) \\ &\dots \\ b_0 &= (-1)^{n+1} [A_1(\zeta_2 \cdot \dots \cdot \zeta_n) + \dots + A_n(\zeta_1 \cdot \dots \cdot \zeta_{n-1})]. \end{aligned}$$

Gäbe es keine eindeutige Lösung A , so müsste das homogene System (linke Seite $b = 0$) eine nichttriviale Lösung A besitzen. Das würde aber mit (*) bedeuten

$$(**) \quad P(z) = 0 = A_1(z - \zeta_2) \cdots (z - \zeta_n) + \dots + A_n(z - \zeta_1) \cdots (z - \zeta_{n-1})$$

wobei etwa $A_1 \neq 0$ ist. Wählt man nun $z \rightarrow \zeta_1$, so strebt der erste Summand gegen

$$A_1(\zeta_1 - \zeta_2) \cdots (\zeta_1 - \zeta_n) \neq 0,$$

die übrigen aber alle gegen 0 im Widerspruch zu (**). \square

Mehrfach-Nullstellen bei Polynomen mit reellen Koeffizienten

Sei $f = \frac{P(x)}{Q(x)}$, $x \in \mathbb{R}$, Koeffizienten reell, $\deg P < \deg Q$.

Q besitzt i.a. reelle Nullstellen einer gewissen Häufigkeit und komplexe Nullstellen z, \bar{z} , die jeweils konjugiert-komplex auftreten, auch mit gewissen Häufigkeiten.

Dann hat $(x - z)(x - \bar{z})$ die Form $(x^2 + px + q)$ mit reellen p, q und $p^2 - 4q < 0$, und der Partialsummenansatz sieht so aus:

$$\begin{aligned} f &= \frac{A_k}{(x-a)^k} + \frac{A_{k-1}}{(x-a)^{k-1}} + \dots + \frac{A_1}{(x-a)^1} \quad (\text{bei } k\text{-facher reeller Nullstelle } a \text{ von } Q) \\ &+ \dots \quad (\text{bei weiteren solchen Nullstellen } b, c, \dots) \\ &+ \frac{B_k x + C_k}{(x^2 + px + q)^k} \quad [x^2 + px + q = (x - z)(x - \bar{z})] \\ &+ \frac{B_{k-1} x + C_{k-1}}{(x^2 + px + q)^{k-1}} \\ &+ \dots \\ &+ \frac{B_1 x + C_1}{(x^2 + px + q)^1} \quad (\text{bei } k\text{-facher konjugiert komplexer Nullstelle } z \text{ von } Q) \\ &+ \dots \quad (\text{bei weiteren solchen Nullstellen }). \end{aligned} \tag{8.2.1}$$

Wieder erhält man die Zahlen A_i, B_ν und C_ν durch Koeffizientenvergleich nach Multiplizieren beider Seiten der Gleichung mit Q (wonach links P stehen bleibt).

Beispiel:

$$\frac{P(x)}{Q(x)} = \frac{x - 1}{x^3 + x^2 + x + 1}.$$

Die Nullstelle $\zeta_1 = -1$ von Q kann man raten. Dann wird $\frac{P}{Q} = \frac{x-1}{(x+1)(x^2+1)}$ mit den beiden Nullstellen $\pm i$ zu $(x^2 + 1) = (x - i)(x + i)$.

Der Ansatz lautet nun

$$\frac{P(x)}{Q(x)} = \frac{A_1}{(x+1)} + \frac{B_1 x + C_1}{(x^2 + 1)}. \tag{8.2.2}$$

Multiplikation mit Q :

$$P(x) = x - 1 = A_1(x^2 + 1) + (B_1 x + C_1)(x + 1)$$

liefert das Gleich.-system

$$\begin{array}{lll} 0 & = A_1 + B_1 & \text{zu } x^2 \\ 1 & = B_1 + C_1 & \text{zu } x^1 \\ -1 & = A_1 + C_1 & \text{zu } x^0 \end{array}$$

mit den Lösungen $A_1 = -1, B_1 = 1, C_1 = 0$. So wird (8.2.2) zu

$$\frac{x - 1}{(x + 1)(x^2 + 1)} = \frac{-1}{(x + 1)} + \frac{x}{(x^2 + 1)}.$$

Die Integration von f (über Intervallen, die keine reellen Nullstellen des Nenners Q beinhalten) ist nun durch Integration der Summanden möglich.

Nötige Stammfunktionen allgemein:

1.

$$\int \frac{A}{(x-a)^k} dx = \begin{cases} (1-k)A(x-a)^{1-k}, & \forall k > 1 \\ A \ln|x-a|, & k = 1 \end{cases}$$

2.

$$\int \frac{Bx + \frac{1}{2}pB}{x^2 + px + q} dx = \frac{1}{2}B \ln|x^2 + px + q|$$

3. von $\frac{C}{x^2+px+q}$: Sei $d = -\frac{D}{4} > 0$. Setze

$$x^2 + px + q = \left(x + \frac{p}{2}\right)^2 + q - \frac{p^2}{4} = \underbrace{\left(q - \frac{p^2}{4}\right)}_{=d} \left[\underbrace{\left(\frac{x + \frac{p}{2}}{\sqrt{q - \frac{p^2}{4}}}\right)^2}_{=y^2} + 1 \right]$$

Also führt das zu

$$\int \frac{dy}{y^2 + 1} = \arctan y$$

und insgesamt zu

$$\int \frac{C}{x^2 + px + q} dx = \frac{C}{d} \arctan \left(\frac{x + \frac{p}{2}}{\sqrt{d}} \right)$$

Analog erhält man Funktionen vom Typ artanh bzw. arcoth (je nachdem, ob die entsprechenden Argumente kleiner oder größer als 1 sind), wenn $d < 0$ ist, siehe oben $I = \int \frac{Ax+B}{ax^2+bx+c} dx$.

4. Für den Fall $k > 1$ und $F = \int \frac{Ax+B}{(x^2+px+q)^k} dx$ gibt es eine **Rekursionsformel**:

$$F = \frac{1}{2} \frac{A}{1-k} (x^2 + px + q)^{1-k} + \left(B - A\frac{p}{2}\right) d^{\frac{1}{2}-k} \int \frac{dx}{(1+x^2)^k}.$$

Das letzte Integral ist rekursiv zu lösen (partielle Integration $\frac{x}{(1+x^2)^k} \cdot \frac{1}{x}$):

$$\int \frac{dx}{(1+x^2)^k} = \frac{x}{2(k-1)(x^2+1)^{k-1}} + \frac{2k-3}{2k-2} \int \frac{dx}{(1+x^2)^{k-1}}.$$

Sicherheitshalber Probe nie vergessen!

Satz 8.11 (Satz von Schwarz (Vertauschbarkeit -stetiger- zweiter partieller Ableitungen)). *Es sei $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ zweimal stetig differenzierbar für (x, y) nahe (x_0, y_0) .*

Dann stimmen die zweiten partiellen Ableitungen $\frac{\partial^2 f}{\partial x \partial y}$ und $\frac{\partial^2 f}{\partial y \partial x}$ in (x_0, y_0) überein.

Bemerkung: Analog für $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Der Satz steht erst hier, weil wir integrieren müssen.

Beweis. O.B.d.A. sei $(x_0, y_0) = (0, 0)$. Für kleine positive s schreiben wir die Differenz $d = f(s, s) - f(0, 0)$ als

$$d = f(s, s) - f(0, s) + f(0, s) - f(0, 0) = \int_0^s \frac{\partial f}{\partial x}(t, s) dt + \int_0^s \frac{\partial f}{\partial y}(0, t) dt = A + B$$

sowie

$$d = f(s, s) - f(s, 0) + f(s, 0) - f(0, 0) = \int_0^s \frac{\partial f}{\partial y}(s, t) dt + \int_0^s \frac{\partial f}{\partial x}(t, 0) dt = C + D.$$

Dann ist offenbar $A + B = C + D$ und $A - D = C - B$.

Andererseits lassen sich die Integrale mit derselben partiellen Ableitung vergleichen:

$$A - D = \int_0^s \frac{\partial f}{\partial x}(t, s) - \frac{\partial f}{\partial x}(t, 0) dt, \quad C - B = \int_0^s \frac{\partial f}{\partial y}(s, t) - \frac{\partial f}{\partial y}(0, t) dt$$

Sei t fixiert mit $0 \leq t \leq s$. Mit $u_t(s) = \frac{\partial f}{\partial x}(t, s)$ ist der erste Integrand gerade $u_t(s) - u_t(0)$ und u_t ist eine differenzierbare Funktion mit der Ableitung nach s : $u_t'(s) = \frac{\partial f}{\partial y} \left(\frac{\partial f}{\partial x} \right) (t, s)$. Der Mittelwertsatz sichert daher: Zu jedem festen t gibt es ein $\theta_t \in (0, s)$ so dass

$$\frac{\partial f}{\partial x}(t, s) - \frac{\partial f}{\partial x}(t, 0) = \frac{\partial f}{\partial y} \left(\frac{\partial f}{\partial x} \right) (t, \theta_t) s$$

Analog (Symmetrie) gibt es ein (anderes) $\theta'_t \in (0, s)$ so dass

$$\frac{\partial f}{\partial y}(s, t) - \frac{\partial f}{\partial y}(0, t) = \frac{\partial f}{\partial x} \left(\frac{\partial f}{\partial y} \right) (\theta'_t, t) s.$$

Beide partielle Ableitungen konvergieren (Stetigkeit!), mit $s \rightarrow 0$, gegen ihre Werte am Nullpunkt. Sind beide verschieden, etwa $\frac{\partial f}{\partial y}(\frac{\partial f}{\partial x})(0,0) < \frac{\partial f}{\partial x}(\frac{\partial f}{\partial y})(0,0) - \varepsilon$, so gilt auch mit kleinen s, t ,

$$\frac{\partial f}{\partial y}(\frac{\partial f}{\partial x})(t, \theta_t) < \frac{\partial f}{\partial x}(\frac{\partial f}{\partial y})(\theta'_t, t) - \varepsilon$$

was zum Widerspruch $A - D < C - B$ führt. \square

Bemerkung: Die Matrix der zweiten partiellen Ableitungen für $f: \mathbb{R}^n \rightarrow \mathbb{R}$ heißt auch Hesse-Matrix. Sie ist also symmetrisch, wenn 2. partielle Ableitungen existieren und stetig sind.

8.3 Uneigentliche Integrale

Uneigentliche Integrale sind definiert als Grenzwerte bestimmter Integrale integrierbarer Funktionen, wenn die entsprechenden Limes existieren und (wie bei Ableitungen und bestimmten Integralen) von der Wahl der jeweiligen (nach Voraussetzung) betrachteten Folgen unabhängig sind. Ein wichtiger Grundtyp hat die Form

$$\int_a^\infty f(t) dt = \lim_{b \rightarrow \infty} \int_a^b f(t) dt,$$

und die Existenz des uneigentlichen Integrals bedeutet (nach Definition), dass der rechte Limes für jede Folge divergierender $b_k \rightarrow \infty$ existiert und derselbe bleibt.

Analog ist erklärt:

$$\int_{-\infty}^b f(t) dt = \lim_{a \rightarrow -\infty} \int_a^b f(t) dt.$$

Beispiel 8.12.

$$\int_0^\infty e^{-t} dt = \lim_{b \rightarrow \infty} \int_0^b e^{-t} dt$$

Hier ist $I(b) := \int_0^b e^{-t} dt = [-e^{-t}]_0^b = -e^{-b} + e^{-0}$ mit Limes $e^0 = 1$ für beliebige $b = b_k \rightarrow \infty$.

Im Gegensatz dazu gilt für $I(b) := \int_0^b \cos t dt$ mit $b_k = 2k\pi$ für alle $k = 1, 2, 3, \dots$:
 $I(b_k) = 0$, aber mit $b_k = \frac{\pi}{2} + 2k\pi$: $I(b_k) = I(b_1) = \sin(\frac{\pi}{2}) = 1$.

Wie bei Konvergenz von Folgen/Reihen ist der Vergleich mit schon bekannten uneigentlichen Integralen zur Behandlung neuer eine wichtige Hilfe:

Existiert $\int_a^\infty f(t) dt$ mit $f(t) \geq 0 \forall t \geq a$ und ist g eine stetige (integrierbare) Funktion mit $0 \leq g(t) \leq$

$f(t) \forall t \geq a$, so existiert auch $\int_a^\infty g(t) dt$.

Man vergleiche die Folgen $I(b_k)$ für g und f .

8.3.1 Laplace Transformierte

Zu gegebenem f ist die Laplace Transformierte die Funktion

$$F(s) = \int_0^\infty e^{-st} f(t) dt,$$

sofern sie definiert ist. Hierbei kann s auch komplex sein. Hauptanwendung liegt in der Theorie von Differentialgleichungen.

8.3.2 Gamma-Funktion

In der Wahrscheinlichkeitsrechnung sind Berechnungen mit der Normalverteilung oft nichttrivial. Das liegt daran, dass für $f(t) = \exp(-t^2)$ keine Stammfunktion explizit angegeben werden kann. Es erweist sich, dass das uneigentlich Integral

$$\int_{-\infty}^{\infty} \exp(-t^2) dt = \sqrt{\pi}$$

mit Hilfe der Gamma-Funktion berechnet werden kann: Letztere besitzt für $x \neq 0, -1, -2, \dots$ mehrere interessante Darstellungen:

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt \quad (\text{Euler}),$$

$$\Gamma(x) = \lim_{n \rightarrow \infty} \frac{n! n^x}{x(x+1) \cdot \dots \cdot (x+n)} \quad (\text{Gauß}).$$

Man erhält so (nach intelligenter Rechnung)

$$\int_0^{\infty} \exp(-t^2) dt = \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = \frac{1}{2} \sqrt{\pi}$$

und

$$\int_{-\infty}^{\infty} \exp(-t^2) dt = \sqrt{\pi}.$$

Dies zeigt, nach Substitution $t = as$, dass für die Standard-Normalverteilung ($\mu = 0, \sigma = 1$) mit der Dichtefunktion

$$f(t) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right)$$

und für die Dichte der Normalverteilung mit Erwartungswert μ und Varianz σ^2

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right)$$

die „Summenbedingung“ $\int_{-\infty}^{\infty} f(t) dt = 1$ gilt.

Für $n = 1, 2, 3, \dots$ zeigt man

$$\Gamma(n) = (n-1)!$$

durch vollständige Induktion (Übungsaufgabe) und partielle Integration.

8.4 Näherungsformeln für Integrale

8.4.1 Trapezformel

$$I_T = (b-a) \frac{f(a) + f(b)}{2},$$

Um das Integral $I = \int_a^b f(x) dx$ näherungsweise zu bestimmen, kann man f durch die lineare Funktion

$L(x) = f(a) + (x-a) \frac{f(b)-f(a)}{b-a}$ ersetzen, die durch $L(a) = f(a)$, $L(b) = f(b)$ eindeutig bestimmt ist.

Das Integral $I_L = \int_a^b L(x) dx$ wird dann gerade

$$I_L = (b-a) \frac{f(a) + f(b)}{2},$$

die Fläche des durch die x -Achse, a, b und L begrenzten Trapezes, daher der Name.

Der Fehler lässt sich für $f \in C^2$ (über die Taylor-Formel für $n = 1$) wegen mit $C = \sup_{\theta \in [a,b]} |D^2 f(\theta)|$

abschätzen (s. Vorlesung) durch

$$|I_L - I| \leq (b-a)^3 C.$$

Die Approximation ist also recht gut bei kleinem $|b-a|$ und bei Funktionen, die „fast linear“ sind, d.h. wenn $|D^2 f(\theta)|$ klein bleibt.

8.4.2 Fassformel

$$I_F = (b-a) \frac{f(a) + f(b) + 4f(c)}{6}.$$

Analog kann man f durch eine quadratische Funktion

$$q = f(a) + \alpha(x-a) + \beta(x-a)^2, \quad (q(a) = f(a))$$

approximieren. Sie möge jetzt durch die Werte von f in a, b und dem Mittelpunkt $c = \frac{a+b}{2}$ des Intervalls festgelegt sein: $c-a = \frac{a+b}{2} - a = \frac{b-a}{2}$

$$\begin{aligned} f(b) &= f(a) + \alpha(b-a) + \beta(b-a)^2 \\ f(c) &= f(a) + \alpha \frac{b-a}{2} + \beta \frac{(b-a)^2}{4}. \end{aligned}$$

Das ergibt lineare Gleichungen für α und β :

$$\begin{aligned} \alpha &= \frac{f(b) - f(a) - \beta(b-a)^2}{b-a} \\ f(c) &= f(a) + \frac{f(b) - f(a) - \beta(b-a)^2}{2} + \beta \frac{(b-a)^2}{4}, \\ \beta \frac{(b-a)^2}{4} &= f(a) - f(c) + \frac{f(b) - f(a)}{2} \end{aligned}$$

und die Lösungen

$$\begin{aligned} \beta &= \frac{4(f(a) - f(c)) + 2(f(b) - f(a))}{(b-a)^2} = \frac{2f(a) + 2f(b) - 4f(c)}{(b-a)^2}, \\ \alpha &= \frac{f(b) - f(a) - (2f(a) + 2f(b) - 4f(c))}{b-a} = \frac{-f(b) - 3f(a) + 4f(c)}{b-a}. \end{aligned}$$

Das Näherungintegral wird damit zu

$$\begin{aligned} I_F &= \int_a^b f(a) + \alpha(x-a) + \beta(x-a)^2 dx = f(a)(b-a) + \alpha \frac{(b-a)^2}{2} + \beta \frac{(b-a)^3}{3} = \dots = \\ &= (b-a) \left(f(a) \left(1 - \frac{3}{2} + \frac{2}{3}\right) + f(b) \left(-\frac{1}{2} + \frac{2}{3}\right) + f(c) \left(2 - \frac{4}{3}\right) \right) \\ &= (b-a) \left(\frac{f(a) + f(b)}{6} + \frac{2f(c)}{3} \right) \\ &= (b-a) \frac{f(a) + f(b) + 4f(c)}{6}. \end{aligned}$$

Diese Formel ist offenbar exakt, wenn f schon quadratisch war, und heißt Fassformel (wenn man diese Approximation q von f auf den entsprechenden Rotationskörper anwendet).

8.4.3 Polynom-Approximation stetiger Funktionen: Bernstein-Polynome

Ist $f = f(t)$ auf einem Intervall $[a, b]$ stetig, kann man eine Funktionenfolge p_n von Polynomen

$$p_n(t) = \sum_{k=0}^n c_k t^k \quad \text{mit}$$

(*) $\sup_{t \in [a, b]} |p_n(t) - f(t)| \rightarrow 0$ (gleichmäßige Konvergenz von $p_n \rightarrow f$ auf $[a, b]$ für $n \rightarrow \infty$)

explizit angeben;

z.B. kann man für $b = a + 1$ setzen

$$p_n(t) = \sum_{k=0}^n \binom{n}{k} f\left(a + \frac{k}{n}\right) (t-a)^k (b-t)^{n-k}.$$

Diese Polynome p_n heißen Bernstein-Polynome.

$$n = 1: \quad p_1(t) = f(a)(b-t) + f(b)(t-a)$$

$$n = 2: \quad p_2(t) = f(a)(b-t)^2 + 2f\left(a + \frac{1}{2}\right)(t-a)(b-t) + f(b)(t-a)^2.$$

Der direkte Beweis dafür, dass (*) gilt, ist allerdings technisch aufwendig.

Die Berechnung des/eines besten Polynoms, d.h., eines Polynoms p_n , welches das Supremum in (*) für gegebenes f und n minimiert, ist möglich. Sie führt allerdings zur einer schon recht komplizierten nichtlinearen Aufgabe, der Tschebyschev-Approximation.

Sucht man analog p_n in der Weise, dass das Integral

$$(**) \quad F(c) = \int_a^b (p_n(t) - f(t))^2 dt = \int_a^b \left(\sum_{k=0}^n c_k t^k - f(t) \right)^2 dt$$

minimal wird (der Fehler nach Gauss, oder auch Approximation im quadratischen Mittel), hat man nur ein lineares Gleichungssystem zu lösen (welches?).

Es entsteht durch die Forderung, dass in der Lösung alle partiellen Ableitungen der von c_0, c_1, \dots, c_n abhängigen Integralfunktion $F(c)$ verschwinden müssen.

8.5 Satz von Green (Gebiets- und Kurvenintegral; Dim. 2)

Gebietsintegral:

Es sei $G \subset \mathbb{R}^2$ eine nichtleere, kompakte Menge, die die Abschliessung ihres Inneren sein soll: $G = \overline{\text{int } G}$. Eine solche Menge heisst auch Gebiet (im \mathbb{R}^2). Weiter sei $f: G \rightarrow \mathbb{R}$ stetig.

Mit einem Gitter, das durch Zerlegungen der x -Achse / y -Achse in segmente Δx_i bzw. Δy_j entsteht, betrachten wir die entsprechenden Rechtecke $R_{i,j}$ mit Fläche $\Delta x_i \Delta y_j$. Dabei gelte

$$0 < \inf_i \Delta x_i, \quad 0 < \inf_j \Delta y_j \quad \text{und} \quad \sup_i \Delta x_i < \delta, \quad \sup_j \Delta y_j < \delta \quad (8.5.1)$$

mit einem $\delta > 0$. Aus den (dann endlich vielen) Rechtecken mit $R_{i,j} \cap G \neq \emptyset$ wähle man einen Punkt $z_{i,j} \in R_{i,j} \cap G$ und bilde die (Integral-) summe

$$S = \sum_{i,j: R_{i,j} \cap G \neq \emptyset} f(z_{i,j}) \Delta x_i \Delta y_j.$$

Sie hängt von der Gitterwahl (speziell der Feinheit δ), sowie der Auswahl der $z_{i,j}$ ab. Die Voraussetzungen bezüglich f und G sichern dann (im wesentlichen analog zur Existenz des Riemann-Integrals über einem Intervall bei Stetigkeit des Integranden), dass die Integralsummen S einen gemeinsamen Grenzwert \bar{S} besitzen, wenn nur $\delta \rightarrow 0$ strebt. Man nennt \bar{S} das Gebietsintegral von f über G und schreibt dafür

$$\bar{S} = \iint_G f(z) \, dx dy. \quad (8.5.2)$$

Kurvenintegral (2. Art):

Nun sei K eine stetige Kurve im \mathbb{R}^2 , die 2 Punkte p^A, p^E (Anfangs/Endpunkt) miteinander verbindet. Wir wählen auf K (in natürlicher Reihenfolge) endlich viele Punkte p_0, p_1, \dots, p_n mit $p_0 = p^A, p_n = p^E$ und bilden $\Delta p_i = p_i - p_{i-1}$, $i = 1, \dots, n$. Mit $p = (x, y)$ ist

$$\Delta p_i = (\Delta x_i, \Delta y_i) = (x_i - x_{i-1}, y_i - y_{i-1}), \quad \text{wobei } \Delta x_i < 0 \text{ und } \Delta y_i < 0 \text{ möglich sind.}$$

Schliesslich sei z_i ein Punkt der Kurve zwischen p_i und p_{i-1} und P und Q seien reellwertige Funktionen, definiert und stetig auf einer offenen Menge $W \subset \mathbb{R}^2$, die K enthält. Wir bilden die (Integral-) summen

$$\sigma = \sum_{i=1}^n [P(z_i) \Delta x_i + Q(z_i) \Delta y_i].$$

Strebt $\delta := \max_i \|\Delta p_i\|$ gegen Null, haben diese (bei jeder Wahl der z_i) wiederum einen gemeinsamen Grenzwert $\bar{\sigma}$, den man Kurvenintegral (2. Art) nennt und schreibt als:

$$\bar{\sigma} = \int_K (P \, dx + Q \, dy). \quad (8.5.3)$$

Zwischen den Integralen (8.5.2) und (8.5.3) besteht ein wichtiger Zusammenhang, den wir hier nur für *einfache* Gebiete G und Kurven K studieren, die den Rand von G begrenzen. Dann ist K eine geschlossene Kurve ($p^E = p^A$).

Um *einfaches* G zu definieren, seien

$$y_{\min} = \min\{y \mid (x, y) \in G\}; \quad y_{\max} = \max\{y \mid (x, y) \in G\}.$$

Für $y \in [y_{\min}, y_{\max}]$ sei weiter $x_1(y)$ das kleinste x mit $(x, y) \in G$ und $x_2(y)$ das größte x mit $(x, y) \in G$. Symmetrisch dazu definiere man x_{\min} , x_{\max} und für $x \in [x_{\min}, x_{\max}]$ analog $y_1(x)$ und $y_2(x)$. Diese Größen gibt es für jedes Gebiet G .

Einfach soll nun bedeuten, dass alle Punkte (x, y) des Rechtecks $[x_{\min}, x_{\max}] \times [y_{\min}, y_{\max}]$ mit $x \in [x_1(y), x_2(y)]$ oder $y \in [y_1(x), y_2(x)]$ zu G gehören und die Funktionen x_1, x_2, y_1, y_2 stetig sind. Insbesondere darf G dann kein Loch besitzen. Die Bedingung ist speziell für konvexes G mit *glattem* Rand erfüllt und hat zur Folge, dass

$$\iint_G f(z) \, dx dy = \int_{y_{\min}}^{y_{\max}} \left(\int_{x_1(y)}^{x_2(y)} f(x, y) \, dx \right) dy = \int_{x_{\min}}^{x_{\max}} \left(\int_{y_1(x)}^{y_2(x)} f(x, y) \, dy \right) dx,$$

was man durch Umsortieren der Integralsummen S sieht.

Satz 8.13. (Green) Für einfache Gebiete G mit dem Rand K und stetig differenzierbare P und Q gilt

$$\iint_G (Q_x - P_y) \, dx dy = \int_K (P \, dx + Q \, dy) \quad (P_x = D_x P, \quad Q_y = D_y Q).$$

(Wir brauchen nur Existenz und Stetigkeit von Q_x und P_y auf einer offenen Menge, die G enthält. Offen, damit auch die Ableitungen im Rand von G definiert sind)

Beweis. Linkes Integral ausrechnen jeweils über „horizontale“ und „vertikale“ Integration:

$$J = \iint_G (Q_x - P_y) \, dx dy = \int_{y_{\min}}^{y_{\max}} \left(\int_{x_1(y)}^{x_2(y)} (Q_x - P_y) \, dx \right) dy \quad (\text{innen horizontal integrieren})$$

Ausrechnen des Integrals mittels $\int_{x_1(y)}^{x_2(y)} Q_x(x, y) \, dx = Q(x_2(y), y) - Q(x_1(y), y)$

$$\begin{aligned} J &= \int_{y_{\min}}^{y_{\max}} \left(\int_{x_1(y)}^{x_2(y)} (Q_x - P_y) \, dx \right) dy \\ &= \int_{y_{\min}}^{y_{\max}} \left([Q(x_2(y), y) - Q(x_1(y), y)] - \int_{x_1(y)}^{x_2(y)} P_y \, dx \right) dy \\ &= \int_K Q \, dy - \int_{y_{\min}}^{y_{\max}} \int_{x_1(y)}^{x_2(y)} P_y \, dx dy = \int_K Q \, dy - \iint_G P_y \, dx dy. \end{aligned}$$

Analog: Ausrechnen des Integrals mittels $\int_{y_1(x)}^{y_2(x)} P_y(x, y) dy = P(x, y_2(y)) - P(x, y_1(y))$.

$$\begin{aligned} J &= \int_{x_{\min}}^{x_{\max}} \left(\int_{y_1(x)}^{y_2(x)} (Q_x - P_y) dy \right) dx \quad (\text{innen vertikal integrieren}) \\ &= \int_{x_{\min}}^{x_{\max}} \left(-[P(x, y_2(x)) - P(x, y_1(x))] + \int_{y_1(x)}^{y_2(x)} Q_x dy \right) dx \\ &= \int_K P dx + \int_{x_{\min}}^{x_{\max}} \int_{y_1(x)}^{y_2(x)} Q_x dy dx = \int_K P dx + \iint_G Q_x dx dy. \end{aligned}$$

Vergleich beider Darstellungen ergibt:

$$\begin{aligned} \int_K Q dy - \iint_G P_y dx dy &= \int_K P dx + \iint_G Q_x dx dy; \\ \int_K (Q dy - P dx) &= \iint_G (P_y + Q_x) dx dy. \end{aligned}$$

Ersetzt man nun P durch $-\hat{P}$, folgt die Behauptung für \hat{P} und Q . \square

Als Folgerungen ergeben sich z.B.

$$\int_K (P dx + Q dy) = 0 \quad \text{wenn} \quad Q_x(z) = P_y(z) \quad \forall z \in G \quad (8.5.4)$$

bzw.

$$\iint_G (Q_x - P_y) dx dy = 0 \quad \text{wenn} \quad P(z) = Q(z) = 0 \quad \forall z \in K. \quad (8.5.5)$$

Für das komplexe Integral entlang einer stetigen Kurve $K \subset \mathbb{C}$ und $f: \mathbb{C} \rightarrow \mathbb{C}$,

$$I = \int_K f(z) dz \quad (8.5.6)$$

das man wie üblich als Limes der Integralsummen

$$S = \sum_i f(z_i) \Delta z_i$$

definiert, folgt:

Satz 8.14. Ist K eine geschlossene Kurve, die ein (einfaches) Gebiet G berandet, und ist f komplex differenzierbar (auch holomorph genannt) auf einer offenen Menge W , die G enthält, so gilt

$$\int_K f(z) dz = 0. \quad (8.5.7)$$

Beweis. Mit $z = x + iy$, $f(z) = f_1(z) + if_2(z)$ und $dz = dx + idy$ ist

$$f(z) dz = [f_1(z)dx - f_2(z)dy] + i[f_2(z)dx + f_1(z)dy].$$

Damit zerfällt I in das Realteil/Imaginärteil -Integral

$$I_1 = \int_K [f_1(z)dx - f_2(z)dy], \quad I_2 = \int_K [f_2(z)dx + f_1(z)dy].$$

Für I_1 setzen wir $P = f_1$ und $Q = -f_2$. Dann ist wegen der Cauchy-Riemann Gleichungen

$$P_y = \frac{\partial f_1}{\partial y} = -\frac{\partial f_2}{\partial x} = Q_x$$

und (8.5.4) liefert $I_1 = 0$. Für I_2 setzen wir $P = f_2$ und $Q = f_1$. Nun ist erneut

$$P_y = \frac{\partial f_2}{\partial y} = \frac{\partial f_1}{\partial x} = Q_x$$

und (8.5.4) liefert auch $I_2 = 0$. \square

Für komplizierteres G werden die Beweise technisch schwieriger. Man zerlegt dann z.B. G in geeignete Teilgebiete und beachtet die Orientierungen der Randkurven K beim Umlaufen der Ränder.

Kapitel 9

Differentialgleichungen (DGL)

Von einer Differentialgleichung spricht man, wenn man Funktionen sucht, die zusammen mit ihren Ableitungen gewisse Gleichungen erfüllen sollen. Hängen die Funktionen nur von einer reellen Variablen ab, spricht man von gewöhnlichen DGL, andernfalls - wenn also auch partielle Ableitungen auftreten - von partiellen DGL.

Hier betrachten wir nur die ersteren.

Die *Ordnung* einer DGL kennzeichnet die Ordnung der höchsten auftretenden Ableitung. Die gesuchten Funktionen heißen zumeist $x = x(t)$ (speziell in der allgemeinen Theorie) oder $y = y(x)$ (in vielen Formelsammlungen).

Grundfragen: (a) Existenz und Eindeutigkeit von Lösungen und (b) Methoden, sie zu finden. Wir beginnen mit (a) für DGL in spezieller, aber wichtiger Form. Die potentiellen Lösungen $x = x(t)$ betrachten wir als Elementen eines metrischen Raumes.

9.1 Existenz und Eindeutigkeit von Lösungen

9.1.1 Der Raum $C[a, b]$

Der Raum $C[a, b]$ besteht aus allen stetigen reellen Funktionen $x = x(t)$, $a \leq t \leq b$, mit der Norm $\|x\| = \sup_{a \leq t \leq b} |x(t)|$ (Supremum-Norm; mitunter schreibt man auch $\|\cdot\|_\infty$).

Dieser wichtige normierte Raum ist als metrischer Raum vollständig (vollständige normierte Räume heißen Banach-Räume).

Satz 9.1. *Der Raum $C[a, b]$ mit der obigen Norm $\|\cdot\|$ ist ein vollständiger Raum.*

Beweis. Man betrachte eine Cauchy-Folge $\{x_k\}$ in $C[a, b]$. Es ist zu zeigen, dass sie gegen ein x aus $C[a, b]$ konvergiert. Nach Definition einer Cauchy-Folge gilt:

$$\forall \varepsilon > 0 \exists N(\varepsilon) : \|x_m - x_n\| \leq \varepsilon \quad \forall n, m > N(\varepsilon). \quad (9.1.1)$$

Dann folgt mit allen festen t : $|x_m(t) - x_n(t)| \leq \varepsilon$. Also existiert (da \mathbb{R} vollständig) für jedes feste t :

$$x(t) = \lim_{m \rightarrow \infty} x_m(t) \in \mathbb{R}. \quad (9.1.2)$$

Abschätzung der Konvergenz $x_n \rightarrow x$ in $C[a, b]$: Seien dazu $\varepsilon > 0$ und $t \in [a, b]$ beliebig fixiert. Weil für $m, n > N(\varepsilon)$ gilt $|x_m(t) - x_n(t)| \leq \varepsilon$, folgt auch

$$|x(t) - x_n(t)| \leq |x(t) - x_m(t)| + |x_m(t) - x_n(t)| \leq |x(t) - x_m(t)| + \varepsilon$$

und hiermit, falls $m > m(t, \varepsilon)$ groß genug ist, sodass $|x(t) - x_m(t)| \leq \varepsilon$ (wegen (9.1.2) möglich),

$$|x(t) - x_n(t)| \leq 2\varepsilon \quad \forall n > N(\varepsilon).$$

Da dies für jedes $t \in [a, b]$ richtig ist, gilt also $\|x - x_n\| \leq 2\varepsilon \quad \forall n > N(\varepsilon)$.

Stetigkeit von x : Es bleibt zu zeigen, dass die Funktion $x = x(t)$ aus (9.1.2) stetig ist, also zu $C[a, b]$ gehört. Dazu hält man irgendein $n > N(\varepsilon)$ fest, und schätzt ab

$$|x(t) - x(t')| \leq |x(t) - x_n(t)| + |x_n(t) - x_n(t')| + |x_n(t') - x(t')| \leq 2\varepsilon + |x_n(t) - x_n(t')| + 2\varepsilon.$$

Da $x_n \in C[a, b]$ nach Voraussetzung stetig ist (also über dem Kompaktum $[a, b]$ auch gleichmäßig stetig), gilt auch

$$|x_n(t) - x_n(t')| < \varepsilon \quad \text{falls nur } |t' - t| \text{ hinreichend klein ist, etwa } < \delta.$$

Folglich ist dann $|x(t) - x(t')| < 5\varepsilon$ wenn $|t' - t| < \delta$ (und deshalb x gleichmäßig stetig). \square

Einige Bezeichnungen:

- Abbildungen von einem normierten oder auch nur metrischen Raum in einen anderen nennt man häufig Operatoren (weil oft schon die Elemente $x \in X$ Funktionen sind), man spricht insbesondere von linearen Operatoren, wenn man allgemeinere Räume als den \mathbb{R}^n im Auge hat.
- Ist der Bildraum Y einer solchen Abbildung die Menge der reellen Zahlen \mathbb{R} , nennt man sie zumeist „Funktional“. Von besonderer Bedeutung sind die linearen (additiven, homogenen und stetigen) Funktionalen eines normierten linearen Raumes X . Sie bilden nach Abschnitt 4 selbst wieder einen linearen Raum, in dem man auch eine Norm einführen kann (Kap. 6). Dieser Raum wird (topologischer) Dualraum von X genannt und mit X^* bezeichnet. Die eigentliche Funktionalanalysis beginnt dann, wenn man sich mit der Konvergenz von Operatoren und Funktionalen (außerhalb des \mathbb{R}^n) beschäftigt. Gewisse Elemente davon werden wir benötigen, um Aufgaben zu lösen, in welchen Funktionen gesucht werden, die eine „Operatorgleichung“ erfüllen müssen. Ein erstes wichtiges Beispiel ist die folgende (Standard-) Differentialgleichung

9.1.2 Eine reelle Funktion wird gesucht

Wir betrachten hier die DGL

$$x' = f(x, t), \quad x(t_0) = x_0. \quad (9.1.3)$$

Sie bedeutet, dass eine, für t nahe t_0 , differenzierbare reelle Funktion $x = x(t)$ gesucht wird, die dort mit ihrer Ableitung $x' = \frac{dx}{dt}$ der Gleichung $x'(t) = f(x(t), t)$ und einer Anfangsbedingung $x(t_0) = x_0$ mit gegebenem t_0 und x_0 genügt.

Insbesondere interessiert uns, ob $x = x(t)$ auf einem gegebenem Intervall $T_1 < t_0 < T_2$ existiert. Ein Blick auf die DGL

$$x' = 1 + x^2 \quad (9.1.4)$$

mit der/einer Lösung $x = \tan t$ zeigt nämlich, dass dies bei grossen Intervallen nicht sicher ist.

Umformulierung als (Fixpunkt-) Gleichung in $C[a, b]$:

Generell gelte: Die Funktion f sei stetig auf \mathbb{R}^2 , es sei $T_1 \leq a < t_0 < b \leq T_2$ und $x \in X = C[a, b]$. Wir ordnen x eine neue Funktion $y = F(x) \in X$ zu, indem wir y für jedes $t \in [a, b]$ definieren als

$$(1) \quad y(t) = x_0 + \int_{t_0}^t f(x(s), s) ds.$$

Da f und x stetig sind, ist es auch y . Außerdem ist

$$(2) \quad y(t_0) = x_0 \quad \text{und} \quad y'(t) = f(x(t), t) \quad \forall t \in (a, b).$$

Damit löst $x \in X$ die DGL (9.1.3) genau dann, wenn $x = y (= F(x))$ mit geeigneten a und b gilt.

Beispiel 9.2. $f(x, t) = \cos t$ (unabhängig von x) und Anfangs-Bedingung $x(0) = 2$. Die Funktion y mit $y(t) = 2 + \int_0^t \cos(s) ds = 2 + \sin t$ ist bereits eine Lösung der DGL.

Beispiel 9.3. $f(x, t) = x$, $x(0) = 2$. Die Funktionen y hat die Form $y(t) = 2 + \int_0^t x(s) ds$.

Mittels Stetigkeit von f folgt leicht die Stetigkeit von F als Abbildung von $C[a, b]$ in sich selbst. Sie sichert, dass man die Existenz einer Lösung (definiert für t nahe t_0) mittels Verallgemeinerungen des Brouwerschen Fixpunktsatzes beweisen kann. Dies ist gewöhnlich nicht Stoff des 1. Studienjahres.

Hier wollen wir unter der stärkeren Zusatzvoraussetzung

$$(3) \quad \text{Es existiere ein } L, \text{ sodass } |f(\xi', t) - f(\xi, t)| \leq L|\xi' - \xi| \quad \forall \xi', \xi \in \mathbb{R}, \forall t \in [T_1, T_2]$$

(einer Lipschitzbedingung bzgl ξ) die Existenz und Eindeutigkeit einer nahe t_0 definierten Lösung nachweisen. Dabei werden wir sehen, dass und wie Voraussetzung (3) abgeschwächt werden kann.

Mittels (3) kann man zeigen, dass F Lipschitz-stetig ist: Sind nämlich $y_1 = F(x_1)$, $y_2 = F(x_2)$ die Bilder zweier Funktionen $x_1, x_2 \in X$, so folgt wegen

$$|x_1(s) - x_2(s)| \leq \max_{t \in [a, b]} |x_1(t) - x_2(t)| = \|x_1 - x_2\| \quad \text{und} \quad t \in [a, b] :$$

$$(*) \quad \begin{aligned} |y_1(t) - y_2(t)| &= \left| \int_{t_0}^t f(x_1(s), s) - f(x_2(s), s) ds \right| \leq \left| \int_{t_0}^t |f(x_1(s), s) - f(x_2(s), s)| ds \right| \\ &\leq \left| \int_{t_0}^t L|x_1(s) - x_2(s)| ds \right| \leq \left| \int_{t_0}^t L\|x_1 - x_2\| ds \right| \leq L\|x_1 - x_2\| \left| \int_{t_0}^t ds \right| \\ &\leq (b-a)L\|x_1 - x_2\|. \end{aligned}$$

Da das für alle $t \in [a, b]$ gilt, erhält man

$$\|y_1 - y_2\| \leq (b-a)L\|x_1 - x_2\|; \quad F \text{ ist Lipschitz-stetig mit Konstante } K = (b-a)L. \quad (9.1.5)$$

Offenbar ist $K < 1$, wenn nur das Intervall klein ist: $(b-a) < \frac{1}{L}$.

Nach Banachs Fixpunktsatz (Kap. 3.3, S. 31) folgt dann die Existenz genau eines Fixpunktes von F :

$$x^* = F(x^*) \quad (= y^*).$$

Damit ist auf (kleinen) Intervallen $(a, b) \subset (T_1, T_2)$ mit $(b-a) < \frac{1}{L}$ und $a < t_0 < b$ die Lösung der DGL (mit Anfangs-Bedingung) eindeutig bestimmt.

Lösungen auf (T_1, T_2) :

Hat man bereits eine Lösung, etwa auf $[t_0 - c, t_0 + c] \subset (a, b)$, sie heiße x^c , kann man diese auf das Intervall $[t_0 - c, t_0 + 2c]$ fortsetzen, indem man nun die Differentialgleichung mit $t_1 = t_0 + c$ und der neuen Anfangsbedingung $x(t_1) = x^c(t_1)$ auf dem Intervall $[t_1 - c, t_1 + c]$ studiert. Damit erhält man eine auf $[t_0 - c, t_1 + c]$ erklärte (zusammengesetzte) Lösung x^{2c} (denn es gilt auf $[t_1 - c, t_1 + c]$ nach Voraussetzung dieselbe Lipschitz-Abschätzung wie vorher).

Anschließend kann man $t_2 = t_1 + c$ setzen und die eindeutige Lösbarkeit über $[t_2 - c, t_2 + c]$ mit Anfangsbedingung $x(t_2) = x^{2c}(t_2)$ ausnutzen. In dieser Weise folgt die Existenz und Eindeutigkeit auf jedem Intervall der Form $[t_0 - c, t_0 + kc]$, solange nur $t_0 + kc \leq T_2$, $k \in \mathbb{N}$. Ist $t_0 + (k+1)c > T_2$ betrachten wir nun einfach ein kleineres Intervall mit $c' = T_2 - (t_0 + kc)$. Setzt man die Lösung analog nach links fort, liefert dies den

Satz 9.4. [Existenzsatz 1] *Unter Voraussetzung (3) existiert genau eine Funktion $x = x(t)$ mit $x(t_0) = x_0$, die für alle $t \in (T_1, T_2)$ definiert ist und dort $x'(t) = f(x(t), t)$ erfüllt.* \square

Wir fordern nun (3) nur für ξ', ξ nahe dem Anfangswert x_0 :

$$(4) \quad \text{Für gewisse } L > 0, r > 0 \text{ gilt } |f(\xi', t) - f(\xi, t)| \leq L|\xi' - \xi| \quad \forall \xi', \xi \in [x_0 - r, x_0 + r] \quad \forall t \in [T_1, T_2]$$

Jetzt definiere man eine neue Funktion g mittels f und r durch

$$g(x, t) = \begin{cases} f(x, t), & \text{falls } |x - x_0| \leq r, \\ f(x_0 - r, t), & \text{falls } x < x_0 - r, \\ f(x_0 + r, t), & \text{falls } x > x_0 + r, \end{cases}$$

Sie erfüllt (wegen Konstantheit für große bzw. kleine x) die Voraussetzung (3) für $T_1 \leq t \leq T_2$.

Damit hat die DGL $x'(t) = g(x(t), t)$ mit $x(t_0) = x_0$ für $T_1 < t < T_2$ nach Existenzsatz 1 eine eindeutige Lösung x^g . Wegen Stetigkeit von x^g und $x^g(t_0) = x_0$ folgt für alle t in einem kleinen Intervall $(t_0 - \delta, t_0 + \delta) \subset (T_1, T_2)$ auch $|x^g(t) - x_0| < r$. Dort gilt aber $g(x^g(t), t) = f(x^g(t), t)$. Also löst x^g auf $(t_0 - \delta, t_0 + \delta)$ auch die Originalgleichung (9.1.3), und es folgt:

Satz 9.5. [Existenzsatz 2] *Unter Voraussetzung (4) gibt es auf hinreichend kleinen Intervallen $(t_0 - \delta, t_0 + \delta)$ stets eindeutig bestimmte Lösungen der DGL (9.1.3) mit Anfangs-Bedingung $x(t_0) = x_0$.*

Bemerkung: Das ist der Satz von Picard und Lindelöf für den Fall einer gesuchten reellen Funktion.

9.1.3 Ein System von Funktionen wird gesucht

Es sei jetzt $x = x(t) = (x_1(t), \dots, x_n(t))$ eine Funktion vom Type $x : \mathbb{R} \rightarrow \mathbb{R}^n$, und $f(x(t), t)$ besitze ebenfalls n reelle Komponenten $f_1(x(t), t), \dots, f_n(x(t), t)$ (also $f = (f_1, \dots, f_n)$).

Wir suchen nun $x_i = x_i(t)$ ($i = 1, \dots, n$) derart, dass

$$x'_i(t) = f_i(x(t), t) \quad \forall i \quad (9.1.6)$$

mit einer Anfangsbedingung $x_1(t_0) = x_{1,0}, \dots, x_n(t_0) = x_{n,0}$ für alle Komponenten erfüllt wird.

Der Unterschied zu vorher besteht darin, dass jetzt $x(t)$ und $f(x(t), t)$ Elemente des \mathbb{R}^n sind, und alle f_i hängen (stetig) von $n + 1$ reellen Argumenten ab.

Die Funktion x liegt im n -fachen kartesischen Produkt X von $C[a, b]$ mit sich selbst, bezeichnet mit $(C[a, b])^n$. Als Norm von $x \in (C[a, b])^n$ definieren wir jetzt

$$\|x\| = \max_{a \leq t \leq b} \|x(t)\|_2$$

wobei $\|x(t)\|_2$ die Euklidische Norm des \mathbb{R}^n ist, sie ersetzt hier den Betrag. Erneut ist X ein Banach-Raum (betrachte die Komponenten für Cauchy-Folgen einzeln). Man definiere nun $y = F(x)$ komponentenweise:

$$y_i(t) = x_{i,0} + \int_{t_0}^t f_i(x(s), s) ds \quad \forall i. \quad (9.1.7)$$

Zwei Funktionen x^1, x^2 aus $(C[a, b])^n$ sind nun - analog zu vorher - Funktionen y^1 und y^2 aus $(C[a, b])^n$ zugeordnet. Unter der Voraussetzung

$$(3)' \quad \text{Es existiere ein } L, \text{ sodass } \|f(\xi', t) - f(\xi, t)\|_2 \leq L\|\xi' - \xi\|_2 \quad \forall \xi', \xi \in \mathbb{R}^n, t \in [T_1, T_2],$$

kann man nun jede Komponente i von $y^1(t) - y^2(t) \in \mathbb{R}^n$ in derselben Weise wie oben abschätzen:

$$\begin{aligned} (**) \quad |y_i^1(t) - y_i^2(t)| &= \left| \int_{t_0}^t f_i(x^1(s), s) - f_i(x^2(s), s) ds \right| \leq \int_{t_0}^t |f_i(x^1(s), s) - f_i(x^2(s), s)| ds \\ &\leq \int_{t_0}^t L \|x^1(s) - x^2(s)\|_2 ds \leq \int_{t_0}^t L \|x^1 - x^2\| ds \leq (b-a)L \|x^1 - x^2\|. \end{aligned}$$

Für die Euklidische Norm folgt so

$$\|y^1(t) - y^2(t)\|_2 \leq \sqrt{2}(b-a)L \|x^1 - x^2\| \quad \text{und} \quad \|y^1 - y^2\| \leq \sqrt{2}(b-a)L \|x^1 - x^2\|. \quad (9.1.8)$$

Nun folgt Kontraktivität von $F = F(x)$ für $\sqrt{2}(b-a)L < 1$, und die obigen Fixpunkt-Argumente lassen sich sämtlich wiederholen.

Also gelten die obigen Existenzsätze 1 und 2 ganz analog für (2, 2) bzw. (n, n) Systeme von Differentialgleichungen mit entsprechenden Anfangsbedingungen.

Was passiert, wenn die Lipschitz-Bedingung nicht erfüllt ist ?

Beispiel 9.6. $x' = x^{\frac{1}{3}}, x(0) = 0$. Eine Lösung ist sicher die Nullfunktion.

Mit dem Ansatz $x = ct^p$ erhält man eine weitere: $x' = cpt^{p-1}, \quad x^{\frac{1}{3}} = c^{\frac{1}{3}}t^{\frac{p}{3}}$. Also muss gelten:

$$c^{\frac{1}{3}}t^{\frac{p}{3}} = cpt^{p-1}.$$

$p - 1 = \frac{p}{3}$ liefert $\frac{2}{3}p = 1, p = \frac{3}{2}$. Es bleibt also nur zu erfüllen: $c^{\frac{1}{3}} = \frac{3}{2}c$, d.h., $c^{\frac{2}{3}} = \frac{2}{3}$. Somit gibt es mindestens 2 Lösungen der DGL mit derselben Anfangsbedingung, die zweite ist

$$x = c t^{3/2}; \quad c = \sqrt[2/3]{2/3}, \quad \text{also} \quad x = \sqrt[2/3]{2/3} t^{3/2}.$$

9.2 Existenzsatz für Lineare Systeme

Mit reellen (n, n) Matrizen $A(t)$ und n -Vektoren $b(t)$, die als Funktionen von t stetig auf \mathbb{R} sind, hat das System mit $f = Ax + b$, d.h.,

$$x'_k(t) = A_k(t)x(t) + b_k(t), \quad A_k(t) = \text{Zeile } k \text{ von } A(t), \quad k = 1, 2, \dots, n \quad (9.2.1)$$

mit den Anfangsbedingungen

$$x_1(t_0) = x_{1,0}, \dots, x_n(t_0) = x_{n,0}$$

besondere Bedeutung: Es charakterisiert Systeme linearer DGL mit nicht konstanten Koeffizienten (A und b) und wird kurz beschrieben als

$$x' = A(t)x + b(t), \quad x(t_0) = x^0 \in \mathbb{R}^n. \quad (9.2.2)$$

Satz 9.7. (Existenz- und Eindeigkeitsatz für lineare DGL-Systeme.)

Sind A und b stetige Funktionen von t , so hat das System (9.2.1) mit jeder Anfangsbedingung eine auf ganz \mathbb{R} definierte eindeutige Lösung $x = x(t)$.

Beweis. Ist ein Intervall $T_1 < t_0 < T_2$ gegeben, findet man wegen Stetigkeit von A , dass

$$c_A = \max_{1 \leq i, j \leq n; t \in [T_1, T_2]} |a_{ij}(t)|$$

existiert. Damit folgt für beliebige $\xi', \xi \in \mathbb{R}^n$ und $i \in \{1, \dots, n\}$

$$|f_i(\xi', t) - f_i(\xi, t)| = |A_i(t)(\xi' - \xi)| \leq n c_A \|\xi' - \xi\|_2 \quad (\text{Euklidische Norm in } \mathbb{R}^n).$$

Deshalb gilt auch (3) in der entsprechenden Form

$$(3)' \quad \text{Es existiert ein } L, \text{ sodass } \|f(\xi', t) - f(\xi, t)\|_2 \leq L \|\xi' - \xi\|_2 \quad \forall \xi', \xi \in \mathbb{R}^n, T_1 \leq t \leq T_2.$$

Somit gibt es nach Satz 9.4 (in seiner \mathbb{R}^n -Form) zu jeder Anfangsbed. genau eine auf (T_1, T_2) definierte Lösung $x = x(t)$. Das Intervall darf beliebig groß sein. Folglich gibt es zu jeder Anfangsbed. sogar genau eine auf ganz \mathbb{R} definierte Lösung. \square

9.3 Lineare DGL-Systeme mit konstanten Koeffizienten

Wir setzen nun **konstante Koeffizienten** A und b in (9.2.2) voraus, obwohl mehrere Aussagen auch allgemeiner gelten.

$$x'_1(t) = A_1 x(t) + b_1, \dots, x'_n(t) = A_n x(t) + b_n; \quad \text{kurz } x' = Ax + b. \quad (9.3.1)$$

Ist $b = 0$ spricht man vom homogenen System, sonst vom inhomogenen. Die Lösungen u des homogenen Systems bilden offensichtlich einen linearen Unterraum U (in $C[T_1, T_2]^n$); und wenn x irgendeine spezielle Lösung des inhomogenen Systems (9.3.1) ist, so auch $x + u$ falls $u \in U$. So erhält man zugleich alle Lösung von (9.3.1); denn wenn y eine weitere Lösung von (9.3.1) ist, so liegt $y - x$ in U , d.h., $y \in x + U$ (vergl. lin. Gleichungen im \mathbb{R}^n).

Satz 9.8. Der Raum U besitzt die Dimension n .

Beweis. Jedem $x^0 \in \mathbb{R}^n$ ist nach Satz 9.7 eindeutig eine Funktion $u \in U$ mit $u(t_0) = x^0$ zugeordnet. Umgekehrt kann man jeder Funktion $u \in U$ den Wert $x^0 := u(t_0) \in \mathbb{R}^n$ zuordnen. Beide Abbildungen sind linear und stellen einen Isomorphismus zwischen U und \mathbb{R}^n her. Also ist $\dim U = \dim \mathbb{R}^n = n$. \square

Also brauchen wir ein System von n linear unabhängigen Funktionen $u^1, \dots, u^n \in U$, um ganz U mit Hilfe der Linearkombinationen

$$u = \sum_{i=1}^n c_i u^i; \quad c_i \in \mathbb{R} \quad (9.3.2)$$

zu beschreiben. Ein solches System nennt man *Fundamentalsystem*. Dabei bedeutet *linear unabhängig*, dass $u(t) = 0 \forall t$ in (9.3.2) nur mit $c_1 = \dots = c_n = 0$ möglich ist.

Ein Fundamentalsystem kann man mit Hilfe der Eigenwerte (und Eigenvektoren) der Matrix A bestimmen, wovon wir uns im Folgenden überzeugen.

9.3.1 Lösungsansätze

Um wenigstens eine Lösung $x \neq 0$ der homogenen DGL zu finden, wähle man speziell x als:

$$x(t) = ve^{\lambda t}, \quad v \in \mathbb{R}^n \quad (\text{als Spalte}). \quad (9.3.3)$$

Dies liefert (man differenziere jede Komponente) $x' = \lambda ve^{\lambda t} = \lambda E ve^{\lambda t}$, was für die DGL bedeutet

$$0 = Ave^{\lambda t} - \lambda E ve^{\lambda t} = (A - \lambda E)v e^{\lambda t}.$$

Da stets $e^{\lambda t} \neq 0$, gilt also $x'(t) = Ax(t)$ - mit unserem Ansatz - genau dann, wenn

$$(A - \lambda E)v = 0. \quad (9.3.4)$$

Eine Lösung $x \neq 0$ für die homogene DGL erhält man also aus (9.3.3) mittels Eigenvektoren v von A . Das bleibt auch richtig, wenn λ und v komplex sind; siehe unten. Zunächst denken wir uns λ reell.

Fall 1: n einfache reelle Eigenwerte Sind alle Eigenwerte $\lambda_1, \lambda_2, \dots, \lambda_n$ reell und verschieden, finden wir mit (9.3.3) n derartige Funktionen als Kandidaten für ein Fundamentalsystem

$$u^1 = v^1 e^{\lambda_1 t}, \dots, u^n = v^n e^{\lambda_n t} \in U. \quad (9.3.5)$$

Sie sind linear unabhängig, denn sonst würde mit einem $c \in \mathbb{R}^n \setminus \{0\}$ gelten

$$0 = c_1 v^1 e^{\lambda_1 t} + \dots + c_n v^n e^{\lambda_n t} = (c_1 e^{\lambda_1 t}) v^1 + \dots + (c_n e^{\lambda_n t}) v^n \quad \forall t.$$

Ist dies auch nur für ein t richtig, wären die Eigenvektoren linear abhängig, was unmöglich ist. Folglich liefert (9.3.5) eine Basis für U , und U selbst - auch *allgemeine Lösung* der homogenen DGL genannt - besteht aus den reellen Linearkombinationen

$$u = c_1 v^1 e^{\lambda_1 t} + \dots + c_n v^n e^{\lambda_n t}, \quad c \in \mathbb{R}^n \quad (9.3.6)$$

dieser Funktionen.

Fall 2: Polynomansatz bei Eigenwert der Ordnung $> m$. Kommt oben ein reeller Eigenwert, etwa $\lambda_1 = \lambda_2$, doppelt vor, liefert der Ansatz $u^1 = u^2$. Es fehlt eine Funktion für die Basis von U . Jetzt kann man wie oben $u^1(t) = v^1 e^{\lambda_1 t}$ setzen und $u^2(t) = w^1 t e^{\lambda_1 t}$ mit geeignetem $w^1 \in \mathbb{R}^n \setminus \{0\}$. Um dies zu begründen, fragen wir allgemeiner, wann es eine Lösung der Form

$$x(t) = P(t)e^{\lambda t} := [w_0 + w_1 t + \dots + w_m t^m] e^{\lambda t}; \quad w_\nu \in \mathbb{R}^n \quad (\text{seien Spaltenvektoren, } w_m \neq 0) \quad (9.3.7)$$

gibt. Mit $m = 0$ ist $w_0 = v$ aus (9.3.4). Mit $m = 1$ entspricht der Vektor w_1 obigem w^1 .

Nun gilt $x' = P'(t) e^{\lambda t} + \lambda P(t) e^{\lambda t}$ und $Ax = AP(t) e^{\lambda t}$, was für die DGL wieder wegen $e^{\lambda t} \neq 0$ bedeutet ($\forall t$)

$$0 = AP(t) - [P'(t) + \lambda P(t)] = -P'(t) + (A - \lambda E) P(t); \quad \text{d.h.,} \quad P'(t) = (A - \lambda E) P(t);$$

Setzen wir ein, dass

$$P'(t) = w_1 + 2w_2 t + \dots + m w_m t^{m-1}; \quad (A - \lambda E) P(t) = (A - \lambda E) [w_0 + w_1 t + \dots + w_m t^m],$$

ergibt Koeffizientenvergleich die für diesen Ansatz notwendigen und hinreichenden Bedingungen

$$\begin{aligned} w_1 &= (A - \lambda E)w_0 \\ 2w_2 &= (A - \lambda E)w_1 \\ &\dots \\ mw_m &= (A - \lambda E)w_{m-1} \\ 0 &= (A - \lambda E)w_m. \end{aligned} \quad (9.3.8)$$

Ihre Lösbarkeit mit $w_m \neq 0$ kann man algebraisch charakterisieren (für mehr Details siehe Jordansche Normalform und nilpotente Abbildungen). Aus $w_\nu = 0$ ($\nu < m$) folgt offenbar $w_{\nu+1} = \dots = w_m = 0$. Wegen $w_m \neq 0$ dürfen daher w_{m-1}, \dots, w_0 nicht im Eigenraum S_0 zum Wert λ liegen. Definiert man die (nicht fallenden) Teilräume

$$S_{\nu+1} = \{w \mid (A - \lambda E)w \in S_\nu\} = \{w \mid (A - \lambda E)^{\nu+1}w \in S_0\} \quad \nu = 0, 1, \dots, m-1$$

muss sogar gelten $w_{m-1} \in S_1 \setminus S_0$, $w_{m-2} \in S_2 \setminus S_1$, ... , $w_1 \in S_{m-1} \setminus S_{m-2}$, $w_0 \in S_m \setminus S_{m-1}$, insbesondere also

$$S_1 \setminus S_0 \neq \emptyset, S_2 \setminus S_1 \neq \emptyset, \dots, S_{m-1} \setminus S_{m-2} \neq \emptyset, S_m \setminus S_{m-1} \neq \emptyset. \tag{9.3.9}$$

Gilt (9.3.9) und ist $w_0 \in S_m \setminus S_{m-1}$ beliebig gewählt, sind w_1, \dots, w_m durch (9.3.8) bestimmt und erfüllen die letzte Gleichung mit $w_m \neq 0$. Wegen $\dim S_0 \geq 1$ und $n \geq \dim S_{\nu+1} > \dim S_\nu$, folgt dann auch $n \geq \dim S_m \geq 1 + m$. System (9.3.8) kann man mit dem gesuchten $(m+1)n$ Vektor $(w_0, \dots, w_m)^T$ auch in Matrixform mit totaler Zeilen/Spaltenzahl $(m+1)n$ schreiben:

$$\begin{pmatrix} (A - \lambda E) & -E & 0 & 0 & 0 \dots & 0 \\ 0 & (A - \lambda E) & -2E & 0 & 0 \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & (A - \lambda E) & -(m-1)E & 0 \\ 0 & 0 & 0 & 0 & (A - \lambda E) & -mE \\ 0 & 0 & 0 & 0 & 0 & (A - \lambda E) \end{pmatrix} \begin{pmatrix} w_0 \\ w_1 \\ \dots \\ w_{m-1} \\ w_m \end{pmatrix} = 0 \tag{9.3.10}$$

Induktiv kann man nun zeigen, dass eine Lösung mit $w_m \neq 0$ existiert (also die notw. und hinr. Bedingung (9.3.9) tatsächlich erfüllt ist), falls λ Eigenwert von A der Ordnung $m+1$ oder grösser ist. Dann kann also eine Lösung der Form (9.3.7) durch Lösen von (9.3.10) bzw. (9.3.8) gefunden werden.

Beispiel: Dreifache reelle Nullstelle λ .

Jetzt gilt offenbar $m < 3$ mit $m = 0, 1, 2$. Das erlaubt den Polynomansatz mit 3 Polynomen unterschiedlichen Grades < 3 ,

$$P_m(t) = [w_{0,m} + w_{1,m}t + \dots + w_{m,m}t^m], \quad w_{m,m} \neq 0 \in \mathbb{R}^n,$$

was mit (9.3.7) auch 3 (linear unabhängige) Funktionen als Kandidaten für ein Fundamentalsystem liefert

$$u^m(t) = P_m(t)e^{\lambda t}$$

Wegen des unterschiedlichen Polynomgrades ist leicht zu sehen, dass die u^m zusammen mit den Funktionen zu Fall 1 ein linear unabhängiges System bilden. Wechselt man zu einem anderen mehrfachen Eigenwert λ' , ändert sich $e^{\lambda t}$; die lineare Unabhängigkeit ist wieder offensichtlich.

Fall 3: Komplexer Eigenwert: Sei $\lambda = \alpha + \beta i$, $\beta \neq 0$ ein komplexer Eigenwert von A . Wir betrachten die komplexe homogene DGL

$$X'_k = A_k X, \quad (k = 1, \dots, n); \quad \text{mit } X_k(t) = X_k^r(t) + i X_k^i(t) \in \mathbb{C}; \tag{9.3.11}$$

Sie bedeutet:

$$(X_k^r)' + i (X_k^i)' = A_k X^r + i A_k X^i, \quad (k = 1, \dots, n), \tag{9.3.12}$$

d.h., $x = X^r$ und $x = X^i$ erfüllen beide die Originalgleichung. Der Ansatz (9.3.3) hat nun die Form

$$X(t) = V e^{\lambda t}, \quad V \in \mathbb{C}^n \text{ (als Spalte, konstant)}. \tag{9.3.13}$$

und liefert $X' = \lambda V e^{\lambda t} = \lambda E V e^{\lambda t}$, was für die hom. DGL bedeutet

$$0 = A V e^{\lambda t} - \lambda E V e^{\lambda t} = (A - \lambda E) V e^{\lambda t}.$$

Erneut ist $e^{\lambda t} \neq 0$, wir erhalten $(A - \lambda E)V = 0$, und wieder löst (9.3.13) die hom. DGL genau dann, wenn V Eigenvektor von A bzgl. λ ist. Da auch $\bar{\lambda}$ Eigenwert von A ist, können wir die Form der Lösungen vereinfachen, um wieder zu reellen Funktionen zu gelangen. Zunächst rechnet man direkt aus:

(*) Ist $V = V^r + i V^i$ EigVektor zu λ , so ist $-\bar{V} = -(V^r - i V^i)$ EigVektor zu $\bar{\lambda}$.

Damit wissen wir: Löst $X = X(t)$ (9.3.13) die homogene Gleichung (9.3.11), so auch Y mit

$$Y(t) := -\bar{V} e^{\bar{\lambda} t}.$$

Ferner sind auch die Linearkombinationen $\Phi = \frac{1}{2}(X + Y)/2$ und $\Psi = \frac{1}{2}(X - Y)$ Lösungen von (9.3.11). Wir rechnen aus, dass sie uns zu Funktionen $u^1, u^2 \in U$ mit

$$u^1(t) = e^{\alpha t} [V^r \sin(\beta t) + V^i \cos(\beta t)], \quad u^2(t) = e^{\alpha t} [V^r \cos(\beta t) - V^i \sin(\beta t)] \tag{9.3.14}$$

führen: Die Funktionen $p(t) = e^{\lambda t}$, $q(t) = e^{\bar{\lambda} t}$ erfüllen $q(t) = \bar{p}(t)$, womit in jeder Komponente k folgt

$$\begin{aligned} X_k(t) &= V_k p(t) = e^{\alpha t} (V_k^r + i V_k^i) [\cos(\beta t) + i \sin(\beta t)] \\ Y_k(t) &= -\bar{V}_k \bar{p}(t) = -e^{\alpha t} (V_k^r - i V_k^i) [\cos(\beta t) - i \sin(\beta t)]. \end{aligned} \tag{9.3.15}$$

Mit $z = x + iy$ und $\zeta = u + iv$ gilt $z\zeta = xu - yv + i[xv + yu]$, $\bar{z}\bar{\zeta} = xu - yv - i[xv + yu] = \overline{z\zeta}$.

Also ist $z\zeta + \bar{z}\bar{\zeta} = 2[xu - yv]$ reell und $z\zeta - \bar{z}\bar{\zeta} = 2i[xv + yu]$ rein imaginär.

Wir wenden das auf die Real/Imaginär-Teil Komponenten (z)[ζ] und (\bar{z})[$\bar{\zeta}$] in (9.3.15) an.

$$\begin{aligned}\Phi_k &= \frac{1}{2}(X_k + Y_k) = \frac{1}{2}(V_k p - \bar{V}_k \bar{p}) = i[xv + yu] = i e^{\alpha t} [V_k^r \sin(\beta t) + V_k^i \cos(\beta t)], \\ \Psi_k &= \frac{1}{2}(X_k - Y_k) = \frac{1}{2}(V_k p + \bar{V}_k \bar{p}) = [xu - yv] = e^{\alpha t} [V_k^r \cos(\beta t) - V_k^i \sin(\beta t)].\end{aligned}$$

Damit lösen Real- und Imaginärteil dieser Funktionen, also die Funktionen (mit k -ter Komponente)

$$\varphi_k = e^{\alpha t} [V_k^r \sin(\beta t) + V_k^i \cos(\beta t)], \quad \psi_k = e^{\alpha t} [V_k^r \cos(\beta t) - V_k^i \sin(\beta t)] \quad (9.3.16)$$

oder kürzer $u^1 = \varphi$, $u^2 = \psi$ beide die (reelle) homogene Gleichung (siehe (9.3.14)).

Fall 4: mehrfacher komplexer Eigenwert: Sei $\lambda = \alpha \pm \beta i$ ein mehrfacher konj.-komplexer Eigenwert. Der Polynomansatz (9.3.7) kann nun in der Form

$$X(t) = P(t)e^{\lambda t} := [W_0 + W_1 t + \dots + W_m t^m]e^{\lambda t}; \quad W_\nu \in \mathbb{C}^n \text{ (seien Spaltenvektoren, } W_m \neq 0) \quad (9.3.17)$$

wiederholt werden. Wir haben \mathbb{R}^n durch \mathbb{C}^n ersetzt, also die zugrunde liegenden Zahlkörper \mathbb{R} , \mathbb{C} ausgetauscht. Ansonsten lassen sich die Argumente wiederholen und liefern für den Ansatz die notw. und hinr. Bedingung, dass wir einen (konj.-kompl.) Eigenwert λ der Ordnung $> m$ brauchen.

Mittels Real/Imaginärteil der $W_\nu = W_\nu^r + i W_\nu^i$ gelangt man wie im Fall 3 zu den Winkelfunktionen

$$\varphi = e^{\alpha t} \sum_{\nu=0}^m t^\nu [W_\nu^r \sin(\beta t) + W_\nu^i \cos(\beta t)], \quad \psi = e^{\alpha t} \sum_{\nu=0}^m t^\nu [W_\nu^r \cos(\beta t) - W_\nu^i \sin(\beta t)] \quad (9.3.18)$$

als Lösungen des reellen homogenen Systems. Man muss nur analog zu (*) nachrechnen:

(**) Löst $W = (W_0, \dots, W_m)$ das komplexe System (9.3.8) zu λ , so löst $-\bar{W}$ das System zu $\bar{\lambda}$.

Mit demselben Argument wie bei einer reellen Dreifach-Nullstelle λ erhalten wir jetzt also jeweils 2 mal drei (lin. unabh.) Funktionen vom Typ φ bzw. ψ .

Zusammenfassung:

Wir finden in dieser Weise insgesamt n Funktionen u^k vom Typ $P(t)e^{\lambda t}$, die sich für konj. komplexe $\lambda = \alpha + i\beta$ mit Hilfe von $e^{\alpha t} \sin(\beta t)$ und $e^{\alpha t} \cos(\beta t)$ schreiben lassen. Sie erweisen sich (da Polynomgrad und λ nie beide zusammenfallen) in der Tat als linear unabhängig, was wir hier nicht explizit nachrechnen.

Damit folgt zugleich (durch Ausmultiplizieren in (9.3.2)):

Jede Lösung $u = u(t)$ der hom. Gleichung ist Summe von Funktionen des Typs

$$f = t^m e^{\alpha t} \sin(\beta t) \quad \text{und} \quad g = t^m e^{\alpha t} \cos(\beta t),$$

wenn man jede noch mit einem geeigneten n -Vektor $V(f)$ bzw. $V(g)$ multipliziert.

Die wirklich gebrauchten Funktionen ($V \neq 0$) hängen ab von den Eigenwerten $\lambda = \alpha + i\beta$ der Matrix A und deren Vielfachheit $O(\lambda)$, die m durch $0 \leq m < O(\lambda)$ beschränkt.

Für einfache Eigenwerte verschwindet also $t^m (= 1)$, für reelle bleibt nur $g = t^m e^{\alpha t}$.

Um die geeignete $V(f)$, $V(g)$ zu finden, kann man auch die Summe u in die hom. Gleichung einsetzen und mit den resultierenden Bedingungen Koeffizientenvergleich durchführen. Das führt (in anderer Weise) zu den obigen linearen Gleichungen.

Inhomogene Gleichung (Variation der Konstanten): Mit einem Fundamentalsystem bilde man

$$x(t) = \sum_k C_k(t) u^k(t)$$

und setze in die inhomogene Gleich. ein. Das liefert eine DGL für die reellwertigen Funktionen C_k , denn es ist $Ax + b = x'$ genau dann wenn

$$\begin{aligned}Ax + b &= \sum_k [(C_k)' u^k + C_k (u^k)'] &= \sum_k [(C_k)' u^k + C_k A u^k] \\ &= \sum_k (C_k)' u^k + \sum_k A C_k u^k &= \sum_k (C_k)' u^k + Ax,\end{aligned}$$

d.h., es bleibt zu lösen

$$b = \sum_k (C_k)' u^k(t) \quad (\text{auch wenn } b = b(t) \text{ stetig ist}). \quad (9.3.19)$$

Weil die Funktionen u^k linear unabhängig sind, sind es auch die \mathbb{R}^n -Vektoren $u^k(t)$ mit jedem $t = t_0$. Falls nämlich $0 = \sum_k c_k u^k(t_0)$, $c \neq 0$, ist die Funktion $u = \sum_k c_k u^k$ eine nicht-Null-Lösung der homogenen Gleichung, die die Anf. Bed. $u(t_0) = 0$ erfüllt. Da die Lösung eindeutig ist (Satz 9.7), kann das aber nur die Nullfunktion sein.

Mit der regulären Matrix $U(t)$ des Systems (9.3.19) existiert daher eine eind. stetige Lösung $C'(t) = \gamma(t) = U(t)^{-1}b$, und wir erhalten schliesslich $C(t)$ als

$$C_k(t) = \int \gamma_k(t) dt + d_k \quad \forall k.$$

Durch Wahl der n Integrationskonstanten d_k kann eine beliebige Anf.-Bed. erfüllt werden.

9.3.2 Beispiele

Beispiel 1 Finde $x: \mathbb{R} \rightarrow \mathbb{R}^2$ mit $x'_1 = x_1 + x_2$, $x'_2 = x_1 - x_2$. Hier ist

$$A - \lambda E = \begin{pmatrix} 1 - \lambda & 1 \\ 1 & -1 - \lambda \end{pmatrix} \text{ mit } \det = -(1 - \lambda)(1 + \lambda) - 1 = -(1 - \lambda^2) - 1 = \lambda^2 - 2.$$

Die Eigenwerte sind reell und verschieden $\lambda_{1/2} = \pm\sqrt{2}$. Die Eigenvekt. v sind nichttr. Lös. von

$$\begin{aligned} 0 &= (A - \lambda_1 E) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 - \sqrt{2} & 1 \\ 1 & -(1 + \sqrt{2}) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}; & v_1 = 1, v_2 = \sqrt{2} - 1; \\ 0 &= (A - \lambda_2 E) \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 1 + \sqrt{2} & 1 \\ 1 & -(1 - \sqrt{2}) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}; & v_1 = 1, v_2 = -(1 + \sqrt{2}). \end{aligned}$$

Damit erhalten wir ein Fundamentalsystem mit den Funktionen

$$u^1(t) = \begin{pmatrix} 1 \\ \sqrt{2} - 1 \end{pmatrix} e^{\sqrt{2}t}, \quad u^2(t) = \begin{pmatrix} 1 \\ -(1 + \sqrt{2}) \end{pmatrix} e^{-\sqrt{2}t}.$$

Die allgemeine Lösung der homogenen DGL besteht aus allen Funktionen

$$u(t) = c_1 \begin{pmatrix} 1 \\ \sqrt{2} - 1 \end{pmatrix} e^{\sqrt{2}t} + c_2 \begin{pmatrix} 1 \\ -\sqrt{2} - 1 \end{pmatrix} e^{-\sqrt{2}t}, \quad c \in \mathbb{R}^2.$$

Probe: Man setze u in die DGL ein und überzeuge sich, dass dies eine Lösung ist.

Beispiel 2 Finde $x: \mathbb{R} \rightarrow \mathbb{R}^2$ mit $x'_1 = x_1 + 4x_2$, $x'_2 = -4x_1 + x_2$.

$$A - \lambda E = \begin{pmatrix} 1 - \lambda & 4 \\ -4 & 1 - \lambda \end{pmatrix} \text{ mit } \det = (1 - \lambda)^2 + 16.$$

Die Eigenwerte sind $\lambda_{1/2} = 1 \pm 4i$. Eigenvektor V zu λ_1 :

$$A - \lambda_1 E = \begin{pmatrix} -4i & 4 \\ -4 & -4i \end{pmatrix}, \quad V = \begin{pmatrix} i \\ -1 \end{pmatrix} = V^r + i V^i = \begin{pmatrix} 0 \\ -1 \end{pmatrix} + i \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Also folgt komponentenweise nach (9.3.16)

$$\varphi = e^{\alpha t} [V^r \sin(\beta t) + V^i \cos(\beta t)], \quad \psi = e^{\alpha t} [V^r \cos(\beta t) - V^i \sin(\beta t)], \quad \text{d.h.}$$

$$\varphi = e^t \left[\begin{pmatrix} 0 \\ -1 \end{pmatrix} \sin(4t) + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cos(4t) \right], \quad \psi = e^t \left[\begin{pmatrix} 0 \\ -1 \end{pmatrix} \cos(4t) - \begin{pmatrix} 1 \\ 0 \end{pmatrix} \sin(4t) \right].$$

und $x = c_1 \varphi + c_2 \psi$, $c \in \mathbb{R}^n$.

Beispiel 3 Finde $x: \mathbb{R} \rightarrow \mathbb{R}^2$ mit $x'_1 = 2x_1 - x_2$, $x'_2 = x_1 + 1$. Hier ist

$$A = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

$$A - \lambda E = \begin{pmatrix} 2 - \lambda & -1 \\ 1 & 0 - \lambda \end{pmatrix}, \quad \det = -\lambda(2 - \lambda) + 1 = \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2.$$

$\lambda = 1$ ist doppelter EWert. Dann wird

$$A - \lambda E = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \quad \text{mit EVektor} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \text{also } u^1(t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^t.$$

Der Polynomansatz (9.3.7) $u(t) = [w_0 + w_1 t] e^t$ führt mit $m = 1$ und $w_0, w_1 \in \mathbb{R}^2$ zum System

$$\begin{aligned} w_1 &= \begin{pmatrix} w_{1,1} \\ w_{1,2} \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} w_{0,1} \\ w_{0,2} \end{pmatrix}, \\ 0 &= \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} w_{1,1} \\ w_{1,2} \end{pmatrix}. \end{aligned} \quad (9.3.20)$$

In Matrixform (9.3.10) mit totaler Zeilen/Spaltenzahl $(m+1)n$ ist das:

$$\left(\begin{pmatrix} \begin{bmatrix} 1 & -1 \\ 1 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} & \begin{bmatrix} -1 & 0 \\ 0 & -1 \\ 1 & -1 \\ 1 & -1 \end{bmatrix} \end{pmatrix} \begin{pmatrix} w_{0,1} \\ w_{0,2} \\ w_{1,1} \\ w_{1,2} \end{pmatrix} \right) = 0 \in \mathbb{R}^4. \quad (9.3.21)$$

Lösungen sind z.B. $\begin{pmatrix} w_{0,1} \\ w_{0,2} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\begin{pmatrix} w_{1,1} \\ w_{1,2} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. Damit erhalten wir

$$u^2(t) = \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} t \right] e^t$$

und die allgemeine Lösung der homogenen Gleichung mit bel. $c_1, c_2 \in \mathbb{R}$ als

$$\begin{aligned} u(t) &= c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^t + c_2 \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} t \right] e^t \\ &= \begin{pmatrix} c_1 + c_2 \\ c_1 \end{pmatrix} e^t + \begin{pmatrix} c_2 \\ c_2 \end{pmatrix} t e^t. \end{aligned}$$

Für das inhomogene System führt der Ansatz - Variation der Konstanten -

$$x(t) = C_1(t) u^1(t) + C_2(t) u^2(t), \quad C_i(t) \in \mathbb{R}$$

- wie wir schon wissen - zur Gleichung (9.3.19) $b = \sum_k (C_k)' u^k(t)$, hier also

$$\begin{aligned} \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= C_1' \begin{pmatrix} 1 \\ 1 \end{pmatrix} e^t + C_2' \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} t \right] e^t \\ e^{-t} \begin{pmatrix} 0 \\ 1 \end{pmatrix} &= C_1' \begin{pmatrix} 1 \\ 1 \end{pmatrix} + C_2' \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} t \right] \end{aligned} \quad (9.3.22)$$

d.h.,

$$C_1' = -C_2' (1+t), \quad e^{-t} = -C_2' (1+t) + C_2' t = -C_2'$$

bzw.

$$C_2' = -e^{-t}; \quad C_1' = -C_2' (1+t) = e^{-t} (1+t) = e^{-t} + t e^{-t}.$$

Wir integrieren und benutzen (partiell) $\int e^{-t} dt = -e^{-t} - e^{-t}$, was ergibt

$$C_2 = e^{-t} + d_2$$

$$C_1 = -e^{-t} + [-e^{-t}t - e^{-t}] + d_1 = -e^{-t} (t+2) + d_1$$

Das liefert mit den Integr.konst. $d_i = 0$ eine spezielle Lösung der inh. Gleichung

$$\begin{aligned} x(t) &= \begin{pmatrix} C_1(t) + C_2(t) \\ C_1(t) \end{pmatrix} e^t + \begin{pmatrix} C_2(t) \\ C_2(t) \end{pmatrix} t e^t, \\ &= \begin{pmatrix} -e^{-t} (t+2) + e^{-t} \\ -e^{-t} (t+2) \end{pmatrix} e^t + \begin{pmatrix} e^{-t} \\ e^{-t} \end{pmatrix} t e^t \\ &= \begin{pmatrix} -(t+2) + 1 \\ -(t+2) \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} t = \begin{pmatrix} -1 \\ -2 \end{pmatrix}. \end{aligned}$$

Damit wird tatsächlich

$$Ax + b = \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = x'.$$

Die allgemeine Lösung der inhomogenen Gleichung hat die Gestalt $x + U$, d.h.,

$$\begin{pmatrix} -1 \\ -2 \end{pmatrix} + \begin{pmatrix} c_1 + c_2 \\ c_1 \end{pmatrix} e^t + \begin{pmatrix} c_2 \\ c_2 \end{pmatrix} t e^t, \quad c_1, c_2 \in \mathbb{R}.$$

Schliesslich noch die Probe für die homogene Gleichung:

$$\begin{aligned} Au &= e^t \begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix} \left[\begin{pmatrix} c_1 + c_2 \\ c_1 \end{pmatrix} + \begin{pmatrix} c_2 \\ c_2 \end{pmatrix} t \right] = e^t \begin{pmatrix} c_1 + 2c_2 \\ c_1 + c_2 \end{pmatrix} + t e^t \begin{pmatrix} c_2 \\ c_2 \end{pmatrix}, \\ u' &= \begin{pmatrix} c_1 + c_2 \\ c_1 \end{pmatrix} e^t + \begin{pmatrix} c_2 \\ c_2 \end{pmatrix} e^t + \begin{pmatrix} c_2 \\ c_2 \end{pmatrix} t e^t = \begin{pmatrix} c_1 + 2c_2 \\ c_1 + c_2 \end{pmatrix} e^t + \begin{pmatrix} c_2 \\ c_2 \end{pmatrix} t e^t. \end{aligned}$$

9.4 Lösungsansätze für spezielle Typen von DGL

Gesucht sei stets eine reelle Funktion $y = y(x)$. Alle auftretenden Funktionen seien zumindest stetig.

9.4.1 Trennung der Variablen

$$y' = f(x)g(y) \quad (\text{bei } g(y) \neq 0)$$

Man schreibe dies als

$$\frac{dy}{g(y)} = f(x)dx, \quad \int \frac{dy}{g(y)} = \int f(x)dx.$$

Kennt man Stammfunktionen G und F beider Integrale, folgt $G(y) = F(x) + C$, und eine explizite Lösung verlangt noch die Auflösbarkeit nach y .

Z.B.

$$\begin{aligned} g(y) = y (\neq 0), \quad f(x) = x^2, \quad G(y) = \ln |y|, \quad F(x) = \frac{x^3}{3}, \\ |y| = e^C e^{\frac{x^3}{3}}, \end{aligned}$$

was etwa für x nahe 1 oder x nahe -1 zu positiven/negativen $y = y(x)$ führt.

9.4.2 Homogene DGL $y' = f(y/x)$

Sei $x \neq 0$. Man substituiere

$$z = y/x, \quad y = zx, \quad y' = z'x + z$$

Damit erhalten wir

$$z'x + z = f(z), \quad z' = \frac{f(z) - z}{x}.$$

Nun ist Trennung der Variablen möglich falls $f(z) \neq z$.

$$\frac{dz}{f(z) - z} = \frac{dx}{x}; \quad \int \frac{dx}{x} = \ln |x|;$$

Mit einer Stammfunktion F zu $\frac{1}{f(z)-z}$, folgt

$$F(z) = \ln |x| + C,$$

und eine explizite Lösung verlangt noch die Auflösbarkeit nach z (analog $y' = f(x/y)$).

9.4.3 Lin. DGL 1. Ordnung (var. Koeff.)

$$y' + a(x)y = b(x). \tag{9.4.1}$$

Homogene Gleichung

Man studiere zuerst die homogene Gleichung ($b = 0$). Trennung der Variablen macht sie mit einer Stammfunktion A für a zu

$$\frac{dy}{y} = -a(x)dx; \quad \ln |y| = -A(x) + c$$

womit folgt $|y| = e^{-A(x)+c} = e^c e^{-A(x)}$. Dies wird von

$$y = C e^{-A(x)} \quad \text{mit bel. } C \in \mathbb{R}$$

erfüllt. [Einsetzen in (9.4.1): $y' + a(x)y = C[-A'(x)e^{-A(x)}] + a(x)Ce^{-A(x)} = 0.$]

Durch Wahl von C kann man jede Anfangsbedingung $y(x_0) = y^0 \in \mathbb{R}$ für die homogene Gleichung erfüllen (warum?).

Inhomogene Gleichung

Man fasse C als Funktion $C = C(x)$ auf und setze jetzt

$$y(x) = C(x)e^{-A(x)} \quad (9.4.2)$$

in (9.4.1) (Variation der Konstanten). Das ergibt wegen

$$y' = C'(x)e^{-A(x)} + C(x)[-A'(x)e^{-A(x)}] = C'(x)e^{-A(x)} + C(x)[-a(x)e^{-A(x)}]$$

eine DGL für die Funktion C :

$$b(x) = C'(x)e^{-A(x)} + C(x)[-a(x)e^{-A(x)}] + a(x)C(x)e^{-A(x)} = C'(x)e^{-A(x)}.$$

Sie kann direkt gelöst werden, d ist eine Integrationskonstante:

$$C'(x) = b(x)e^{A(x)}, \quad C(x) = \int b(x)e^{A(x)} dx + d.$$

Damit sehen die Lösungen von (9.4.1) allgemein so aus:

$$y(x) = (d + \int b(x)e^{A(x)} dx) e^{-A(x)}; \quad d \in \mathbb{R}. \quad (9.4.3)$$

Diese Vorgehensweise ist typisch auch für andere lineare DGL (s. Systeme): Man bestimme zuerst die allgemeine Lösung der homogenen Gleichung. Dann betrachte man die Konstanten C_i als Funktionen $C_i(x)$ und gehe mit diesem Ansatz für y in die inhomogene Gleichung. Im Ergebnis erhält man eine lineare DGL für die Funktionen $C_i(x)$.

Probe: Mit (9.4.3) folgt

$$\begin{aligned} y' &= (d + \int b(x)e^{A(x)} dx) [-A'(x)]e^{-A(x)} + b(x)e^{A(x)} e^{-A(x)} \\ a(x)y &= A'(x) (d + \int b(x)e^{A(x)} dx) e^{-A(x)} \end{aligned}$$

also auch

$$y' + a(x)y = b(x).$$

Das Integral $\int b(x) e^{Ax} dx$ kann hässlich sein, z.B. wird es mit $b = 1$ und $a(x) = 2x$ zu $\int e^{(x^2)} dx$, was nicht mit elementare Funktionen angegeben werden kann. Mit $b(x) = x$ und $a(x) = \cos x$ wird es zu $\int x e^{\sin x} dx$. Partielle Integration liefert z.B.

$$\int x e^{\sin x} dx = \int v u' dx = uv - \int v' u dx = uv - \int u dx.$$

Man braucht dazu aber $u = \int e^{\sin x} dx$. Hier hilft die Substitution $y = \sin x$ wenig. Wegen $y' = \cos x = \sqrt{1 - \sin^2 x} = \sqrt{1 - y^2}$ und $dx = \frac{dy}{\sqrt{1 - y^2}}$ liefert sie nur $u = \int \frac{e^y}{\sqrt{1 - y^2}} dy$, was nicht leichter aussieht.

Partielle Integration von $\int 1 * e^{\sin x} dx$ sieht auch nicht besser aus.

9.4.4 Bernoulli-Gleichung

$$a(x)y' + b(x)y = c(x)y^n, \quad n > 1.$$

Man dividiere durch $y^n (\neq 0)$ und substituiere in der neuen Gleichung

$$a(x) \frac{y'}{y^n} + b(x) \frac{1}{y^{n-1}} = c(x), \quad z = \frac{1}{y^{n-1}}.$$

Dann ist (Kettenregel da $y = y(x)$!)

$$z' = -(n-1) \frac{1}{y^n} y'.$$

Das liefert

$$-\frac{1}{n-1} a(x) z' + b(x) z = c(x), \quad (9.4.4)$$

also nach Division durch $-\frac{1}{n-1} a(x) (\neq 0)$ eine lineare DGL 1. Ordnung mit variablen Koeffizienten.

9.4.5 Riccati-Gleichung

$$y' + f(x)y + g(x)y^2 = h(x) \quad (9.4.5)$$

Ist eine Lösung $y = y(x)$ bekannt, kann man eine weitere $z = z(x)$ mit dem linearen Ansatz

$$z(x) = y(x) + v(x)$$

suchen. Einsetzen von z in (9.4.5) ergibt links -da y (9.4.5) löst - den Ausdruck

$$y' + v' + f(x)(y + v) + g(x)(y^2 + 2yv + v^2) = h(x) + v' + f(x)v + g(x)(2yv + v^2).$$

Das liefert eine Gleichung, um v zu bestimmen:

$$v' + f(x)v + g(x)(2yv + v^2) = 0, \quad d.h.$$

$$v' + [f(x) + 2g(x)y(x)]v = -g(x)v^2, \quad (\text{Bernoulli-Typ, s. oben})$$

$$\frac{v'}{v^2} + \frac{[f(x) + 2g(x)y(x)]}{v} = -g(x). \quad \text{Die Substit. } w = 1/v; \quad w' = -v'/v^2$$

ergibt nun die lin. DGL

$$w' - [f(x) + 2g(x)y(x)] w = g(x).$$

Hier ist $a = -[f(x) + 2g(x)y(x)]$ und $b = g$. Die allg. Lös. sieht dann nach (9.4.3) mit

$$A(x) = - \int [f(x) + 2g(x)y(x)] dx$$

so aus:

$$w(x) = \left(d + \int g(x)e^{A(x)} dx \right) e^{-A(x)}; \quad d \in \mathbb{R}; \quad z = y + \frac{1}{w}. \quad (9.4.6)$$

9.4.6 Lineare DGL n-ter Ordnung (feste Koeff.)

$$y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y^{(1)} + a_0 y = b(x), \quad n > 1. \quad (9.4.7)$$

Man führe neue Funktionen für die höheren Ableitungen ein, indem man

$$z_1 = y \quad \text{und} \quad z_{k+1} = z'_k, \quad k = 1, \dots, n-1$$

setzt. Dann ist

$$z_2 = z'_1 = y^{(1)}, \quad z_3 = z'_2 = y^{(2)}, \quad \dots, \quad z_n = z'_{n-1} = y^{(n-1)} \quad \text{und} \quad z'_n = y^{(n)}.$$

Aus Gleichung (9.4.7) wird so ein lineares DGL-System mit gesuchter Funktion $z = (z_1, \dots, z_n)^T$.

$$\begin{aligned} z'_k &= z_{k+1} & k &= 1, \dots, n-1 \\ z'_n &= -[a_{n-1} z_n + \dots + a_1 z_2 + a_0 z_1] + b(x). \end{aligned}$$

Die (n, n) - Matrix A zum homogenen System $z' = Az$ hat also die Form

$$A = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-1} \end{pmatrix}.$$

Die Vorgabe eines Anfangswertes $z(t_0) = z^0$ bedeutet nun die Vorgabe von Werten

$$y(t_0) = z_1^0, \quad y'(t_0) = z_2^0, \quad \dots, \quad y^{(n-1)}(t_0) = z_n^0$$

für den Funktionswert und die Ableitungen bis Ordnung $n-1$ von y and der Stelle t_0 . Nach Satz 9.7 besitzt dann die Gleichung (9.4.7) mit jedem Anfangswert stets eine eindeutige Lösung (auf ganz \mathbb{R} definiert). Die Eigenwerte von A - für den Lösungsansatz (9.3.3) - ergeben sich aus der (charakteristischen) Gleichung

$$\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0\lambda^0 = 0. \quad (9.4.8)$$

Das sieht man induktiv, wenn man die Determinante von $A - \lambda E$ nach der ersten Spalte entwickelt. Die Lösungsansätze ergeben sich aus der Diskussion des (n, n) - Systems.

Kapitel 10

Anhang

10.1 Konvexität und Norm-Äquivalenz im \mathbb{R}^n

Konvexität:

Eine Teilmenge $M \subset \mathbb{R}^n$ heisst konvex, wenn für alle $x, y \in M$ gilt: Die Punkte $z = \lambda x + (1 - \lambda)y$; $0 \leq \lambda \leq 1$ der Verbindungsstrecke zwischen x und y liegen ebenfalls in M .

Eine Funktion $f : \mathbb{R}^n \rightarrow \mathbb{R}$ heisst konvex, wenn für alle $x, y \in \mathbb{R}^n$ und $\lambda \in (0, 1)$ gilt

$$f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y).$$

Speziell sind also affin-lineare $f : \mathbb{R}^n \rightarrow \mathbb{R}$ konvex. Wenn $-f$ konvex ist, heisst f konkav. Der folgende Fakt ist trivial, aber wichtig.

Lemma 10.1. *Für konvexe Funktionen f ist jeder lokale Minimalpunkt ein globaler, und die Menge $\operatorname{argmin} f$ aller Minimalpunkte ist konvex (leer nicht ausgeschlossen).*

Weiterhin gilt, wenn man die max-Norm des \mathbb{R}^n zugrunde legt,

Satz 10.2. *(Stetigkeit konvexer Funktionen) Wenn $f : \mathbb{R}^n \rightarrow \mathbb{R}$ konvex ist, so ist f stetig.*

Beweis. Um Stetigkeit in $y \in \mathbb{R}^n$ zu zeigen (bezüglich max-Norm of \mathbb{R}^n), nehme man einen Würfel S mit $y \in \operatorname{int} S$. Mit dessen (endl. vielen) Ecken p_k , folgt für jedes $x \in S$ (wenn man x als Konvex-Kombination der Ecken schreibt) $x = \sum_k \lambda_k p_k$, $\sum_k \lambda_k = 1$, $\lambda_k \geq 0$ und $f(x) \leq \sum_k \lambda_k f(p_k)$. Also ist f nach oben beschränkt auf S : $f \leq C$ on S .

Nun gelte $\varepsilon > 0$, $x \rightarrow y$, (i.e. $\|x - y\|_{\max} \rightarrow 0$), $x \neq y$. Wir wählen das einzige z im Rand von S so dass mit einem $\lambda \in (0, 1)$ gilt

$$x = \lambda y + (1 - \lambda)z, \quad x \text{ zwischen } y \text{ und } z, \quad \lambda = \lambda(x).$$

Dann folgt $\lambda \rightarrow 1$ aus $x \rightarrow y$. Konvexität und $f(z) \leq C$ ergeben weiter

$$f(x) \leq \lambda f(y) + (1 - \lambda)f(z) \leq \lambda f(y) + (1 - \lambda)C < f(y) + \varepsilon \quad \text{wenn } \|x - y\|_{\max} \text{ klein genug.}$$

Nun wählen wir das einzige z im Rand von S so dass mit einem $\lambda \in (0, 1)$ gilt

$$y = \lambda x + (1 - \lambda)z, \quad y \text{ zwischen } x \text{ und } z, \quad \lambda = \lambda(x).$$

Wieder folgt $\lambda \rightarrow 1$ aus $x \rightarrow y$, und wir erhalten mit $f(z) \leq C$,

$$f(y) \leq \lambda f(x) + (1 - \lambda)f(z) \leq \lambda f(x) + (1 - \lambda)C.$$

Wenn $\|x - y\|_{\max}$ klein genug ist, sichert das

$$f(y) - \varepsilon < \frac{f(y) - (1 - \lambda)C}{\lambda} \leq f(x).$$

Damit folgt also -wie verlangt- mit jedem $\varepsilon > 0$: $|f(x) - f(y)| < \varepsilon$ falls $\|x - y\|_{\max}$ klein genug ist. \square

Ausserdem ist jede Norm $\|\cdot\|$ eine konvexe Funktion: Mit $f(x) = \|x\|$ gilt nämlich für beliebige $x, y \in \mathbb{R}^n$ und $\lambda \in (0, 1)$:

$$f(\lambda x + (1 - \lambda)y) = \|\lambda x + (1 - \lambda)y\| \leq \|\lambda x\| + \|(1 - \lambda)y\| = \lambda\|x\| + (1 - \lambda)\|y\| = \lambda f(x) + (1 - \lambda)f(y).$$

Normäquivalenz:

Es sei $\|\cdot\|$ eine beliebige Norm des \mathbb{R}^n und $\|\cdot\|_{max}$ die Maximum-Norm $\|x\|_{max} = \max_i |x_i|$. Wir zeigen die Existenz positiver Konstanten α, β mit

$$\alpha\|x\| \leq \|x\|_{max} \leq \beta\|x\| \quad \forall x.$$

Als konvexe Funktion f auf dem \mathbb{R}^n ist $f(x) = \|x\|$ stetig bzgl. der MaxNorm $\|\cdot\|_{max}$. Ausserdem ist die Menge $B_1 = \{x \mid \|x\|_{max} = 1\}$ kompakt (bzgl. MaxNorm des \mathbb{R}^n).

1. Damit existiert $K := \max \{\|x\| \mid x \in B_1\}$ (Weierstrass). Folglich gilt (wegen $\|rx\| = |r|\|x\|$)

$$\|x\| \leq K \|x\|_{max}.$$

2. Wir zeigen weiter: Es gibt ein C sodass $\|x\|_{max} \leq C$ aus $\|x\| = 1$ folgt. D.h. die Einheitskugel bzgl. $\|\cdot\|$ ist beschränkt bzgl. $\|\cdot\|_{max}$ und es gilt $\|x\|_{max} \leq C \|x\|$.

Andernfalls gibt es $x^k \in \mathbb{R}^n$ mit $\|x^k\| = 1$ und $\|x^k\|_{max} \rightarrow \infty$ (die $\|\cdot\|$ -Einheitskugel passt in keinen Würfel). Die Punkte

$$u^k := \frac{x^k}{\|x^k\|_{max}} \quad \text{erfüllen dann} \quad \|u^k\|_{max} = 1 \quad \text{bzw.} \quad u^k \in B_1$$

und (nach Bolzano-Weierstrass) existiert eine bzgl. MaxNorm konvergente Teilfolge $u^{k_\nu} \rightarrow u$ ($\nu \rightarrow \infty$). Dann ist auch $\|u\|_{max} = 1$, also $u \neq 0$ und somit $\|u\| = c > 0$. Da $f = \|\cdot\|$ stetig ist, folgt so $\|u^{k_\nu}\| > \frac{c}{2}$ für grosse Indizes k_ν . Das liefert aber für $k = k_\nu$,

$$\|x^k\| = \|x^k\|_{max} \|u^k\| \geq \|x^k\|_{max} \frac{c}{2} \rightarrow \infty,$$

was $\|x^k\| = 1$ widerspricht. □

10.2 Satz von Arzela-Ascoli

Der folgende Satz ist nicht Bestandteil der Analysis II - Vorlesung 2011, zeigt aber, dass die kompakten Mengen in $C[a, b]$ nicht einfach die beschränkten und abgeschlossenen Mengen wie im \mathbb{R}^n sind.

Satz 10.3 (Satz von Arzela-Ascoli (1883, 1893); Kompaktheit in $C[a, b]$). *Sei $X \subset C[a, b]$ abgeschlossen. X ist kompakt genau dann wenn:*

(1) *alle $x \in X$ sind gleichmäßig beschränkt, d.h. $\exists K : |x(t)| \leq K \forall x \in X \forall t \in [a, b]$*

und

(2) *alle $x \in X$ sind gleichgradig gleichmäßig stetig,*

d.h. $\forall \varepsilon > 0 \exists \delta : |x(t) - x(t')| \leq \varepsilon \forall x \in X \forall t, t' \in [a, b] \text{ mit } |t - t'| \leq \delta$.

Beweis. (\Leftarrow) Sei $\varepsilon > 0$ beliebig fixiert. Wir wählen δ nach (2) und teilen $[a, b]$ in endlich viele, etwa m , Intervalle $I_k = [t_k, t_{k+1}]$ der Länge $< \delta$. Sei L die Menge aller stückweise linearen (stetigen) Funktionen y zur Intervallzerlegung, die $|y(t)| \leq K \forall t \in [a, b]$ erfüllen. Sie lassen sich durch $2m$ reelle Zahlen des Intervalls $[-K, K]$ charakterisieren (die Werte von y in den Randpunkten von I_k).

Also ist L kompakt in C , denn jede unendliche Folge solcher Zahlentupel hat eine unendliche konvergente Teilfolge, und die entsprechenden Funktionen y konvergieren dann in C .

Sei nun $x \in X$. Wir ordnen x diejenige Funktion y zu, die in allen t_k gerade $y(t_k) = x(t_k)$ leistet. Dann folgt:

$$|x(t) - y(t)| \leq \varepsilon \quad \forall t \Rightarrow \|x - y\| \leq \varepsilon.$$

Damit ist L ein kompaktes ε -Netz für X . Aus der Vollständigkeit von C folgt so die Kompaktheit von X .

(\implies) Sei M ein endliches ε -Netz für X , bestehend aus irgendwelchen Funktionen $y^1, \dots, y^m \in C$.

Mit $K = \max_k \|y^k\| + \varepsilon$ folgt dann (1). Sei $\varepsilon > 0$ fixiert. Dann gibt es ein $\delta > 0$, sodass (2) für alle endlich vielen Funktionen $x = y^k$ gilt, denn jede einzelne ist gleichmäßig stetig über dem kompakten Intervall.

Sei nun $x \in X$ beliebig gewählt. Wir finden $y = y^k$, sodass $\|x - y\| \leq \varepsilon$.

Dann gilt für $|t - t'| < \delta$:

$$|x(t) - x(t')| \leq |x(t) - y(t)| + |y(t) - y(t')| + |y(t') - x(t')| \leq 3\varepsilon.$$

Also ist (2) mit $\varepsilon' = 3\varepsilon$ und δ erfüllt. □

10.3 Daten einiger bekannter Mathematiker

Isaak Newton 04.01. 1643
 Brook Taylor 18.08. 1685
 Leonhard Euler 15.04. 1707
 Maurice René Fréchet 2. 9 1878
 George Green 1793-1841
 Joseph-Louis Lagrange 25.01. 1736
 Pierre-Simon Laplace 28.03. 1749
 Jean-Baptiste-Joseph de Fourier 21.03. 1768
 Carl Friedrich Gauss 30.04. 1777
 Augustin-Louis Cauchy 21.08. 1789
 Niels Henrik Abel 05.08. 1802
 Carl Gustav Jacob Jacobi 10.12. 1804
 Karl Weierstrass 31.10. 1815
 Bernhard Riemann 17.09. 1826
 David Hilbert 23.01. 1862
 Luitzen Egbertus Jan Brouwer 27.02. 1881
 Stefan Banach 30.03. 1892
 Kazimierz Kuratorwski 02.02. 1896
 Shizuo Kakutani 28. 8. 1911 in Osaka
 Erwin Schroedinger 12. 8. 1887 in Wien geboren.
 George Gabriel Stokes 13. 8. 1819 in Skreen (Irland)
 William Rowan Hamilton 4. 8. 1805 in Dublin.

Bernoullis (nach Geburtsjahr geordnet)

Leon Bernoulli (1570), Arzt
 Jacob Bernoulli (1598-1634),
 Niklaus Bernoulli (1623-1708), Gewuerzhaendler
 Jakob I. Bernoulli (1655-1705), Mathematiker
 (Bernoulli-Differentialgleichung, Bernoulli-Verteilung, Bernoulli-Zahl (und Polynom),
 Bernoullische Annahmen (Bernoulli-Balken), Bernoullische Ungleichung, Infinitesimalrechnung)
 Nikolaus Bernoulli (1662-1716), Mathematiker
 Johann Bernoulli (1667-1748), Mathematiker
 (Regel von LHospital, Infinitesimalrechnung)
 Nikolaus I. Bernoulli (1687-1759), Mathematiker (Sankt-Petersburg-Paradoxon)
 Nikolaus II. Bernoulli (1695-1726), Mathematiker
 Daniel Bernoulli (1700-1782), Mathematiker
 (Stroemung nach Bernoulli und Venturi (Bernoulli-Effekt),
 Entscheidung unter Risiko (Bernoulli-Prinzip), Bernoullische Energiegleichung).
 Johann II. Bernoulli (1710-1790), Mathematiker
 Johann III. Bernoulli (1744-1807), Astronom
 Daniel II. Bernoulli (1751-1834), Mathematiker
 Jakob II. Bernoulli (1759-1809), Physiker
 Christoph Bernoulli (1782-1863), Naturwissenschaftler und Statistiker
 Johann Jakob Bernoulli (1831-1913), Archaeologe
 Carl Gustav Bernoulli (1834-1878), Arzt, Forschungsreisender und Archaeologe

Wilhelm Bernoulli-Sartorius (1838-1914), Botaniker
August Bernoulli-Burckhardt (1839-1921), Historiker
Carl Christoph Bernoulli (1861-1923), Bibliothekar
Johannes Bernoulli (1864-1920), Bibliothekar
Eduard Bernoulli (1867-1927), Musikwissenschaftler
Carl Albrecht Bernoulli (1868-1937), evangelischer Theologe
Maria Bernoulli (1868 - 1963), Fotografin, 1. Ehefrau von Hermann Hesse und Mutter des
schweizer Malers Bruno Hesse
Elisabeth Bernoulli (1873-1935), Sozialistin
Ludwig Bernoulli (1873-1928), Architekt
Hans Bernoulli (1876-1959), Architekt (Siloturm Basel (Bernoullisilo))
August Leonhard Bernoulli (1879-1939), Physiker
Rudolf Bernoulli (1880-1944), Kunsthistoriker
Walter Bernoulli (1885-1946), Geologe
Eva Bernoulli (1903-1995), Paedagogin und erste Logopdin der Schweiz

Index

- Abbildung, 9
- abgeschlossen, 24
- Ableitung
 - tangens*, 53
 - cosinus*, 53
 - sinus*, 53
 - Fréchet, 48
 - partielle, 58
- Abschließung, 24
- absolut konvergent, 35
- additiv, 49
- alternierend
 - Reihe, 37
- Äquivalenzklasse, 9
- Äquivalenzrelation, 9
- Archimedisches Axiom, 17
- Archimedisches Axiom, 17
- arcus cosinus*, 55
- arcus cotangens*, 55
- arcus sinus*, 55
- arcus tangens*, 55
- area cosinus hyperbolicus*, 56
- area sinus hyperbolicus*, 56
- Assoziativität, 16

- Banach-Raum, 103
- Bernoullische Ungleichung, 11
- Bernstein-Polynome, 99
- beschränkt, 19
- bestimmtes Integral, 83
- bijektiv, 9
- Bild, 9
- Bildmenge, Bild, 9
- Binomialkoeffizient, 12
- Bogenlänge, 91

- Cauchy-Folge, 30
- Cauchy-Riemannsche Gleichungen, 81
- closure, 24
- cosinus hyperbolicus*, 56

- Dedekindscher Schnitt, 14
- Definitionsbereich, 9
- DGL
 - linear, 107
- dicht, 24
- Differential
 - vollständig, 59
- Differentialgleichung, 104
- Differenz, 7
- differenzierbar
 - komplex, 80
 - partiell, 58
- diskrete Metrik, 23
- Distributivgesetz, 16
- divergent, 19
- Dreiecksungleichung, 23
- Dualraum, 104
- Durchschnitt, 7, 24

- Eulersche Darstellung komplexer Zahlen, 76
- Eulersche Zahl, 39
- Evolute, 68

- Fassformel, 98
- fast alle, 19
- Fixpunkt, 31
- Folge, 19
- Fréchet-Ableitung, 48
- Fundamentalsatz der Algebra, 77
- Fundamentalsystem, 107
- Funktion, 9
 - implizite, 69
 - inverse, 9, 71
- Funktional, 104

- Gebietsintegral, 99
- geometrische Reihe, 33
- gleichmäßig stetig, 25
- Gradient, 58
- Grenzwert, 19

- Halbordnung, 8
- harmonische Reihe, 33
- Hauptwert, 18
- Hesse-Matrix, 96
- homogen, 49
- Hyperbel-Funktionen, 56
- hyperbolischer Pythagoras, 56

- Imaginärteil, 17
- implizite Funktion, 69
- Infimum, 15
- injektiv, 9
- innerer Punkt, 24
- Inneres, 24
- Integral
 - bestimmtes, 83
 - komplexes, 101
 - unbestimmtes, 85
- interior, 24
- Intervallschachtelung, 14

- inverse Funktion, 9, 71
- isolierte Punkte, 24
- Jacobi-Matrix, 68
- Kommutativität, 16
- komplex differenzierbar, 80
- komplex linear, 80
- komplexe Zahl, 17
- komplexes Integral, 101
- komplexes Polynom, 77
- konjugiert komplex, 17
- Kontraktionskonstante, 31
- konvergent, 19
 - absolut, 35
- Konvergenzradius, 38
- konvexe Menge, 117
- konvexe, konkave Funktion, 65, 117
- Krümmung, 67
- Krümmungskreis, 67
- Krümmungsradius, 67
- Kreisumfang, 92
- Kugel
 - offen, 23
- Kurvenintegral, 99
- Lagrange Funktion, 71
- Lagrange Multiplikatoren, 71
- Laplace Transformierte, 96
- Limes, 19
- lin. unabh. Funktionen, 107
- linear
 - komplex, 80
 - DGL, 106
- Linearisierung, 50
- Logarithmus, 45
 - natürlicher, 45
 - zur Basis a , 45
- Maximum, Minimum, 26
- Metrik, 23
 - diskrete, 23
- metrischer Raum, 23
- monoton, 20
 - fallend, 20
 - streng, 20
 - strikt, *siehe* streng monoton
 - wachsend, 20
- Newton-Methode, 72
- Norm, 48
- Norm-Äquivalenz, 118
- normierter Raum, 48
- offen, 24
- offene Kugel, 23
- Operator, 104
- Ordnungsrelation, 8
- Partialbruchzerlegung, 93
- Partialsumme, 33
- partiell differenzierbar, 58
- partielle Ableitung, 58
- Pascalsches Dreieck, 12
- Permutation, 11
- Polarkoordinaten, 18
- Polynom
 - komplex, 77
- Polynomdivision, 79
- Potenzmenge, 7
- Potenzreihen, 38
- Produktregel, 52
- Produktreihen, 38
- Pythagoras
 - hyperbolischer, 56
- Quotientenkriterium, 34
- Quotientenregel, 52
- Realteil, 17
- Reihe, 33
 - geometrische, 33
 - harmonisch, 33
- Restglied-Darstellung, 60
- Satz über implizite Funktionen, 69
- Satz von Arzela-Ascoli, 118
- Satz von Picard und Lindelöf, 105
- Satz von Schwarz, 95
- Schmiegekreis, 67
- separabel, 24
- sinus hyperbolicus*, 56
- Stammfunktion, 85
- stationär, 65
- stetig, 25
 - gleichmäßig, 25
- stetige partielle Ableitungen, 58
- streng monoton, 20
 - fallend, 20
 - wachsend, 20
- strikt monoton, *siehe* streng monoton
- Substitutionsmethode, 88
- sukzessive Approximation, 31, 72
- Supremum, 15
- Supremum-Norm, 103
- surjektiv, 9
- tangens hyperbolicus*, 56
- Taylorentwicklung, 60
- Taylorreihe, 60
- Teilmenge, 7
- totales Differential, 68
- Überdeckung, 27
- Urbildmenge, 9
- Vereinigung, 7, 24
- vollständig, 30
- vollständiges Differential, 59, 68
- Wendepunkte, 65
- Wertebereich, 9
- Wohlordnung, 8
- Wurzelkriterium, 34

Bibliographie wird noch ergänzt, gute weitere Anregungen finden sich in

[http : //www.mathematik.uni-stuttgart.de/studium/informat/Analysis_Mielke_WS03/AnaLit.pdf](http://www.mathematik.uni-stuttgart.de/studium/informat/Analysis_Mielke_WS03/AnaLit.pdf)
aber, wenn man nötige Voraussetzungen verstehen will, auch im Klassiker [1].

Literaturverzeichnis

- [1] B.R. Gelbaum , J.M.H. Olmsted. Counterexamples in analysis. Holden-Day, San Francisco, London, Amsterdam, 1964.
- [2] Reiffen, H., Trapp, H. Differentialrechnung. Wissenschaftsverlag, Mannheim 1989.
- [3] Jaenich. Funktionentheorie. Springer-Verlag, Berlin 2004.
- [4] K. Endl, W. Luh. Analysis II. AULA-Verlag, Wiesbaden 1989.
- [5] H. Heuser. Lehrbuch der Analysis, Teil 1 und 2.
- [6] O. Forster. Analysis 1. Vieweg Sohn Verlag, Wiesbaden 2004.
- [7] K. Königsberger. Analysis 1. Springer Verlag, Berlin 2001.