Einführung in die Analysis

Prof. Dr. René Grothmann

2011

Vorwort

Es handelt sich bei diesem Skript nur um eine Zusammenfassung der Vorlesung. Beweise und Beispiele wurden auf ein Minimum reduziert. Auch eine Einführung in die Mengenlehre fehlt, ebenso wie Aufgaben. Es mag zwar inhaltliche Grundlage einer Vorlesung sein, kann aber nicht wörtlich übernommen werden.

Das Skript ist auch für den Bachelorstudiengang oder das Lehramt Gymnasium geeignet, da der Stoff Grundlage jeder Vorlesung in den ersten zwei Semestern Analysis sein sollte. Die Vorlesungen werden sich letztlich deutlich durch die Auswahl der Übungsaufgaben, das Gleichgewicht zwischen Beweis und Beispiel, und die Präsentation der Grundlagen der Mathematik unterscheiden.

Eichstätt, 2011
R. Grothmann

Inhaltsverzeichnis

1 Die reellen Zahlen **9**

1.1 Die Körperaxiome 9

1.2 Die Anordnung 12

1.3 Vollständigkeit 13

1.4 Die Betragsfunktion 16

1.5 Die Komplexen Zahlen 18

2 Folgen **21**

2.1 Die natürlichen Zahlen 21

2.2 Folgen 22

2.3 Endliche und abzählbare Mengen 23

2.4 Summen und Produkte 25

2.5 Die rationalen Zahlen 27

3 Konvergenz **29**

3.1 Grenzwert von Folgen 29

3.2 Häufungspunkte 32

3.3 Uneigentliche Grenzwerte 33

3.4 Grenzwert von Reihen 34

3.5 Absolute Konvergenz 35

3.6 Dezimalbrüche 38

3.7 Komplexe Folgen und Reihen 39

4 Stetigkeit **43**

4.1 Grenzwert von Funktionen 43

4.2 Rationale Funktionen 46

4.3 Der Zwischenwertsatz 47

4.4 Extrema von Funktionen 49

4.5 Stetigkeit in metrischen Räumen 50

5 Die Ableitung **51**

5.1 Differenzierbarkeit 51

5.2 Lokale Extrema und Monotonie 52

5.3 Die Kettenregel 55

5.4 Exponentialfunktion und Logarithmus 56

5.5 Der Satz von de l'Hospital 59

5.6 Konvexe Funktionen 60

6 Integralrechnung **63**

6.1 Das Riemann-Integral 63

6.2 Der Hauptsatz 66

6.3 Partielle Integration und Substitution 68

6.4 Uneigentliche Integrale 69

6.5 Allgemeine Funktionen 71

7 Potenzreihen **73**

7.1 Die Taylorreihe 73

7.2 Potenzreihen 75

7.3 Gleichmäßige Konvergenz 76

7.4 Trigonometrische Funktionen 79

7.5 Partialbruchzerlegung 82

7.6 Die Stirlingsche Formel 83

8 Der Euklidsche Raum **85**

8.1 Konvergenz . . . 85

8.2 Offene Mengen . . . 87

8.3 Stetige Funktionen . . . 89

8.4 Kompakte Mengen . . . 90

8.5 Zusammenhängende und konvexe Mengen . . . 93

8.6 Die Norm von linearen Abbildungen . . . 94

8.7 Der Banachsche Fixpunktsatz . . . 97

9 Differenzialrechung mehrerer Variablen **99**

9.1 Richtungsableitungen . . . 99

9.2 Lokale Extrema . . . 102

9.3 Die totale Ableitung . . . 104

9.4 Implizite Funktionen . . . 109

9.5 Extrema mit Nebenbedingungen . . . 111

9.6 Die Kurvenlänge . . . 113

9.7 Vektorfelder . . . 115

10 Integralrechung mehrerer Variablen **121**

10.1 Beispiele . . . 121

10.2 Treppenfunktionen . . . 125

10.3 Das Riemann-Integral . . . 126

10.4 Riemann-messbare Mengen . . . 127

10.5 Der Transformationssatz . . . 128

Kapitel 1

Die reellen Zahlen

Die reellen Zahlen sind eine Menge, in der man addieren und multiplizieren kann, und in der die Elemente linear geordnet sind. Eine gute Veranschaulichung ist geometrisch mit Hilfe des Zahlenstrahls möglich, wobei zur Multiplikation die Ebene und die Strahlensätze zu Hilfe genommen werden können.

Wir führen die reellen Zahlen axiomatisch ein, anstatt sie aus den natürlichen Zahlen zu konstruieren, wie das, wenn auch unexakt, in der Schule gemacht wird. Eine exakte Konstruktion würde sehr viel Zeit kosten und für Anfänger zu schwierig sein.

Die Axiome der reellen Zahlen bestehen aus drei Gruppen. Wir werden sie nach und nach behandeln, wobei das Vollständigkeitsaxiom die größten Schwierigkeiten im Verständnis bereiten wird.

- Die Axiome der Multiplikation und Addition (Körperaxiome)
- Die Axiome der Anordnung
- Das Vollständigkeitsaxiom

1.1 Die Körperaxiome

1.1. Definition: Eine Menge K mit einer Addition „+“ und einer Multiplikation „·“ heißt **Körper**, wenn in ihr zwei Konstanten $0 \neq 1$ gegeben sind, und wenn die folgenden Rechengesetze gelten.

(1) Die Addition ist **kommutativ**. D.h.

$$a + b = b + a \qquad \text{für alle } a, b \in K$$

(2) Die Addition ist **assoziativ**. D.h.

$$(a + b) + c = a + (b + c) \qquad \text{für alle } a, b, c \in K$$

(3) 0 ist das **neutrale Element** der Addition. D.h.

$$a + 0 = a \qquad \text{für alle } a \in K.$$

(4) Für jedes $a \in K$ existiert ein $b \in K$ (das **additive Inverse**) mit

$$a + b = 0.$$

(5) Die Multiplikation ist **kommutativ**. D.h.

$$a \cdot b = b \cdot a \qquad \text{für alle } a, b \in K$$

(6) Die Multiplikation ist **assoziativ**. D.h.

$$(a \cdot b) \cdot c = a \cdot (b \cdot c) \qquad \text{für alle } a, b, c \in K$$

(7) 1 ist das **neutrale Element** der Multiplikation. D.h.

$$a \cdot 1 = a \qquad \text{für alle } a \in K.$$

(8) Für jedes $a \in K$, $a \neq 0$, existiert ein $b \in K$ (das **multiplikative Inverse**) mit

$$a \cdot b = 1.$$

(9) Es gilt das **Distributivgesetz**. D.h.

$$a \cdot (b + c) = a \cdot b + a \cdot c \qquad \text{für alle } a, b, c \in K.$$

Axiom I. Die reellen Zahlen sind ein Körper.

1.2 Satz: *Es gilt*

$$a + b = a + c \;\Leftrightarrow\; b = c$$

für alle $a, b, c \in K$. Ebenso

$$a \cdot b = a \cdot c \;\Leftrightarrow\; b = c$$

für alle $a, b, c \in K$, vorausgesetzt dass $a \neq 0$ ist. Die neutralen Elemente und die Inversen der Addition und der Multiplikation sind eindeutig bestimmt.

1.3 Satz: *Ein Körper K ist **nullteilerfrei**. D.h., es gilt*

$$a \cdot b = 0 \;\Leftrightarrow\; a = 0 \text{ oder } b = 0$$

für alle $a, b \in K$.

1.4. Bemerkung: Aus diesem Grund kann 0 kein multiplikatives Inverses haben.

1.5. Definition: Wir schreiben das additive Inverse von a wie gewohnt als $-a$, und das multiplikative Inverse von $a \neq 0$ als

$$a^{-1} = \frac{1}{a} = 1/a.$$

Außerdem schreiben wir wie üblich

$$a - b := a + (-b),$$

sowie

$$\frac{a}{b} = a/b := a \cdot b^{-1}.$$

Zwischen verschiedenen Variablen oder zwischen geklammerten Ausdrücken ist es außerdem üblich, den Malpunkt wegzulassen, also

$$ab := a \cdot b$$

Natürlich setzen wir $a^2 = a \cdot a$ etc. Außerdem brauchen wir die Klammern beim Addieren und Multiplizieren nicht immer zu schreiben. Wir definieren

$$a + b + c = (a + b) + c = a + (b + c).$$

Analog für die Multiplikation.

1.6 Satz: *In jedem Körper K gelten die aus der Schule bekannten Rechenregeln für die Addition und die Multiplikation. Zum Beispiel ist $-0 = 0$ und $1^{-1} = 1$. Die Inversen der Inversen sind die Elemente selbst, also*

$$-(-a) = a \qquad \text{für alle } a \in K,$$

und

$$\left(a^{-1}\right)^{-1} = a \qquad \text{für alle } a \in K,\ a \neq 0.$$

Außerdem

$$-(a + b) = -a + (-b) \qquad \text{für alle } a, b \in K,$$

und

$$(ab)^{-1} = a^{-1}b^{-1} \qquad \text{für alle } a, b \in K,\ a, b \neq 0.$$

Außerdem

$$(-a)(-b) = -(a \cdot (-b)) = ab \qquad \text{für alle } a, b \in K.$$

Ebenso hat man die üblichen Rechenregeln für Brüche, z.B.

$$\frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}, \quad \frac{a/b}{c/d} = \frac{ad}{bc}, \quad \frac{a}{b} \pm \frac{c}{d} = \frac{ad \pm bc}{bd}, \quad \frac{a}{b} = \frac{ac}{bc}$$

für alle a, b, c, d, für die diese Brüche definiert sind. Natürlich gelten auch die binomischen Formeln

$$(a \pm b)^2 = a^2 \pm 2ab + b^2, \quad (a + b)(a - b) = a^2 - b^2$$

für alle $a, b \in K$.

1.7. Beispiel: (1) Das für uns wichtigste Beispiel ist $\mathbb{R}$, der Körper der reellen Zahlen.

(2) Aber für viele Rechnungen reicht der Körper $\mathbb{Q}$ der rationalen Zahlen aus. Er ist die Teilmenge von $\mathbb{R}$, die aus Brüchen der Form n/m mit $m \in \mathbb{N}$ und $n \in \mathbb{Z}$ besteht. Wir werden $\mathbb{Q}$ später genauer betrachten.

(3) Es gibt aber auch den Körper $\mathbb{C}$ der komplexen Zahlen, der eine Obermenge von $\mathbb{R}$ ist. Dazu später mehr.

(4) Praktisch wichtig und interessant sind auch Körper mit nur endlich vielen Elementen. So ist etwa der Körper $K = \{0, 1\}$, der nur die beiden notwendigen Elemente 0 und 1 enthält ein Körper. Man rechnet dort $1 + 1 = 0$. Allgemeiner sind endliche Körper die Primzahlrestkörper, die nicht Thema der Analysis sind.

1.2 Die Anordnung

1.8. Definition: Ein **angeordneter Körper** ist ein Körper K, in dem eine Relation $a < b$ definiert ist. Diese Anordnung muss die folgenden Rechenregeln erfüllen.

(1) Die Anordnung ist **transitiv**. D.h.

$$a < b \text{ und } b < c \;\Rightarrow\; a < c$$

für alle $a, b, c \in K$.

(2) Die Anordnung ist **vollständig**. D.h. für je zwei Elemente $a, b \in K$ gilt *genau eine* der folgenden drei Beziehungen

$$a < b, \quad a = b, \quad b < a.$$

(3) Die Anordnung ist **verträglich** mit der Addition. D.h.

$$a < b \;\Rightarrow\; a + c < b + c$$

für alle $a, b, c \in K$.

(4) Die Anordnung ist **verträglich** mit der Multiplikation mit *positiven* Zahlen. D.h.

$$a < b \text{ und } 0 < c \;\Rightarrow\; ac < bc.$$

Axiom II. Die reellen Zahlen sind ein angeordneter Körper.

1.9. Definition: Selbstverständlich schreiben wir wie üblich

$$\begin{aligned} a > b &: \Leftrightarrow\; b < a, \\ a \leq b &: \Leftrightarrow\; a < b \text{ oder } a = b. \\ a \geq b &: \Leftrightarrow\; a > b \text{ oder } a = b. \end{aligned}$$

Außerdem

$$a < b < c \;\Leftrightarrow\; a < b \text{ und } b < c$$

für alle $a, b, c \in K$ etc.

1.10. Bemerkung: Man beachte, dass in Bedingung (2) gefordert wird, dass *genau* eine der Relationen gilt. Es folgt daraus zum Beispiel

$$a \leq b \text{ und } b \leq a \;\Leftrightarrow\; a = b$$

für alle $a, b \in K$.

1.11 Satz: *In einem angeordneten Körper K gelten die aus der Schule gewohnten Rechenregeln für die Ungleichung. Zum Beispiel gilt*

$$a < 0 \;\Leftrightarrow\; -a > 0 \qquad \text{für alle } a \in K$$

oder allgemeiner

$$a < b \;\Leftrightarrow\; -a > -b \qquad \text{für alle } a, b \in K$$

Eine negative Multiplikation dreht die Ungleichung herum, also

$$a < b \text{ und } 0 > c \;\Rightarrow\; ac > bc \qquad \text{für alle } a, b, c \in K$$

für alle $a, b, c \in K$. Ungleichungen kann man addieren, also für alle $a, b, c, d \in K$

$$a < b \text{ und } c < d \;\Rightarrow\; a + c < b + d.$$

Bei der Multiplikation von Ungleichungen muss man vorsichtiger sein. Es gilt aber zum Beispiel

$$0 < a < b \text{ und } 0 < c < d \;\Rightarrow\; ac < bd.$$

Aus der Schule ist bekannt, wann ein Produkt positiv und wann es negativ ist. Es gilt für alle $a, b \in K$

$$ab > 0 \;\Leftrightarrow\; a, b < 0 \text{ oder } a, b > 0$$

Insbesondere folgt daraus

$$a^2 > 0 \qquad \text{für alle } a \in K,\ a \neq 0,$$

also insbesondere $1 = 1^2 > 0$. Für das multiplikative Inverse gilt

$$a > 0 \;\Leftrightarrow\; \frac{1}{a} > 0$$

für alle $a \in K$, $a \neq 0$.

1.12. Beispiel: (1) Unser wichtigstes Beispiel ist wieder $\mathbb{R}$, die Menge der reellen Zahlen.

(2) Aber auch $\mathbb{Q}$ ist als Teilkörper von $\mathbb{R}$ angeordnet.

(3) Die endlichen Körper und auch der Körper $\mathbb{C}$ der komplexen Zahlen lassen sich nicht anordnen.

1.3 Vollständigkeit

Das Vollständigkeitsaxiom ist etwas schwerer zu erklären. Wir benötigen zunächst einige Vorbereitungen, bevor wir das Axiom formulieren.

1.13. Definition: (1) Sei K ein angeordneter Körper und $M \subseteq K$ eine Teilmenge von K. Dann heißt $m \in K$ **obere Schranke** von M,

$$m \geq x \qquad \text{für alle } x \in M.$$

M heißt **nach oben beschränkt**, wenn es eine obere Schranke von M gibt. Falls die obere Schranke m Element von M ist, so heißt sie **Maximum** von M. Wenn m das Maximum von M ist, so schreiben wir

$$m = \max M,$$

oder

$$m = \max_{x \in M} x.$$

(2) Analog definieren wir eine **untere Schranke** von M und das **Minimum** von M.

(3) M heißt **beschränkt**, wenn es nach oben und unten beschränkt ist.

1.14. Bemerkung: Als Abkürzung dafür, dass m obere Schranke von M ist, ist auch

$$m \geq M$$

üblich. Dies ist so zu lesen, dass

$$m \geq x \qquad \text{für alle } x \in M$$

gilt. Für eine beschränkte Menge M gibt es also $c, d \in K$ mit

$$c \leq M \leq d.$$

1.15. Definition: Für a, b in einem angeordneten Körper K sind, so definieren wir die Teilmengen

$$[a, b] := \{x \in K : a \leq x \leq b\},$$
$$]a, b[:= \{x \in K : a < x < b\},$$

und in nahe liegender Weise die Mischformen $[a, b[$ und $]a, b]$. $[a, b]$ heißt **abgeschlossenes Intervall**, und $]a, b[$ **offenes Intervall**. Die Mischformen heißen **halboffene Intervalle**. Wir definieren außerdem die abgeschlossenen, nicht beschränkten Intervalle

$$[a, \infty[:= \{x \in A : x \geq a\},$$
$$]-\infty, a] := \{x \in A : x \leq a\},$$

sowie die offenen Intervalle $]a, \infty[$ und $]-\infty, a[$. Wir schreiben auch

$$\mathbb{R} =]-\infty, \infty[$$

als Intervall. Ein **Intervall** in $\mathbb{R}$ ist eine der hier angegeben Mengen.

1.16. Bemerkung: Für $a > b$ sind die Intervalle $[a, b]$ etc. leer.

1.17. Bemerkung: Offenbar

$$a = \min[a, b] = \min\,[a, \infty[, \quad b = \max[a, b] = \max\,]-\infty, b].$$

Aber das offene Intervall $]a, b[$ hat weder Minimum, noch Maximum.

1.18. Definition: Eine obere Schranke m einer nicht leeren Teilmenge $M \subseteq K$ eines angeordneten Körpers heißt **Supremum** von M, wenn sie die *kleinste* obere Schranke von M ist. D.h. jedes Element $\tilde{m} < m$ ist keine obere Schranke mehr. Analog ist das **Infimum** einer Menge M die *größte* untere Schranke von M. Wir schreiben

$$\sup M, \quad \inf M$$

oder

$$\sup_{x \in M} x, \quad \inf_{x \in M} x$$

für das Infimum und das Supremum, sofern es existiert. Falls M nicht nach oben, bzw. nicht nach unten beschränkt ist, so existiert kein eigentliches Supremum bzw. Infimum. Wir schreiben dennoch

$$\sup M = \infty, \quad \inf M = -\infty.$$

Supremum und Infimum der leeren Menge sind nicht definiert.

1.19. Bemerkung: Infimum und Supremum sind eindeutig, wenn sie existieren.

1.20. Bemerkung: Jedes Maximum ist ein Supremum. Aber umgekehrt ist ein Supremum nur Maximum, wenn es in der Menge liegt.

1.21 Satz: *Sei $a < b$. Dann gilt*

$$a = \inf\,]a, b[, \quad b = \sup\,]a, b[.$$

Analoges gilt für $]a, \infty[$ und $]-\infty, a[$.

1.22. Definition: Ein angeordneter Körper K heißt **vollständig angeordnet**, wenn jede nicht leere, nach oben beschränkte Teilmenge von K ein Supremum besitzt.

Axiom III. $\mathbb{R}$ ist ein vollständig angeordneter Körper.

1.23. Bemerkung: Dann hat auch jede nach unten beschränkte, nicht leere Menge ein Infimum.

1.24 Satz: *Für jede nicht leere Teilmenge $M \subseteq \mathbb{R}$ gibt es ein eindeutig bestimmtes kleinstes Intervall, dass die Menge umfasst. D.h. alle anderen Intervalle, die die M enthalten sind Obermengen dieses Intervalls.*

1.25. Beispiel: In $\mathbb{R}$ existiert zu jedem $a \geq 0$ ein eindeutiges s mit $s^2 = a$. Dies ist aus der Schule her bekannt, und die Wurzel von a kann am Funktionsgraphen von x^2 abgelesen werden. Natürlich bezeichnen wir

$$s = \sqrt{a}.$$

Zur Begründung der Existenz der **Quadratwurzel** betrachten wir die Menge

$$M = \{x \in \mathbb{R} : x > 0 \text{ und } x^2 < a\}.$$

Diese Menge ist nach oben beschränkt und nicht leer. Wir verwenden im Beweis die **Monotonie** der Funktion $x \mapsto x^2$, also den Sachverhalt

$$0 < x < y \;\Leftrightarrow\; 0 < x^2 < y^2$$

für alle $x, y \in \mathbb{R}$. Man kann nun $s = \sup M$ wählen.

1.26 Satz: *Eine Teilmenge $M \subseteq \mathbb{R}$ ist genau dann ein Intervall, wenn für alle a, b gilt*

$$a, b \in M \;\Rightarrow\; [a, b] \subseteq M.$$

*Intervalle sind also die einzigen Teilmengen von $\mathbb{R}$, die mit je zwei Punkten auch die Strecke zwischen den Punkten enthalten. Solche Mengen heißen **konvexe Mengen**.*

Insbesondere ist der Schnitt von Intervallen wieder ein Intervall.

1.4 Die Betragsfunktion

Wichtig für die Konvergenz in $\mathbb{R}$ ist der Begriff des Abstands zweier Punkte. Wir benötigen dazu zunächst die aus der Schule bekannte Betragsfunktion.

1.27. Definition: Für $x \in \mathbb{R}$ definieren wir

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

$|a|$ heißt der **Betrag** von a. Dementsprechend heißt

$$\operatorname{sign}(a) := \begin{cases} 1 & \text{für } a > 0, \\ 0 & \text{für } a = 0, \\ -1 & \text{für } a < 0 \end{cases}$$

das **Vorzeichen** von a.

1.28. Bemerkung: Offenbar gilt

$$a = \operatorname{sign}(a) \cdot |a|$$

und natürlich

$$|a| \geq a$$

für alle $a \in \mathbb{R}$.

1.29 Satz: *Es gilt $|a| \geq 0$ und*

$$|a| = 0 \;\Leftrightarrow\; a = 0$$

für alle $a \in \mathbb{R}$. Außerdem

$$|ab| = |a| \cdot |b|,$$

*sowie die **Dreiecksungleichung***

$$|a + b| \leq |a| + |b|$$

und die Ungleichung

$$||a| - |b|| \leq |a - b|$$

für alle $a, b \in \mathbb{R}$.

1.30. Definition: Mit Hilfe der Betragsfunktion definieren wir eine Abstandsfunktion

$$d(a, b) := |a - b|.$$

für $a, b \in \mathbb{R}$. Wir verwenden diese Schreibweise hier nur der Deutlichkeit halber, und werden in Zukunft einfach $|a - b|$ für den Abstand schreiben.

1.31 Satz: *Die eben definierte Abstandsfunktion ist eine* ***Metrik*** *auf* $\mathbb{R}$. *D.h., sie ist nicht negativ*

$$d(a, b) \geq 0 \qquad \text{für alle } a, b \in \mathbb{R}$$

sowie ***positiv definit****, also*

$$d(a, b) = 0 \iff a = b.$$

Außerdem ist sie ***symmetrisch****, also*

$$d(a, b) = d(b, a) \qquad \text{für alle } a, b \in \mathbb{R}.$$

Und es gilt die ***Dreiecksungleichung***

$$d(a, b) + d(b, c) \leq d(a, c) \qquad \text{für alle } a, b, c \in \mathbb{R}.$$

1.32. Definition: Wir benötigen später den Begriff der **ϵ-Umgebung** eines Punktes

$$U_\epsilon(a) := \{x \in \mathbb{R} : |x - a| < \epsilon\}$$

der für $\epsilon \geq 0$ und $a \in \mathbb{R}$ definiert ist.

1.33. Bemerkung: Es gilt

$$U_\epsilon(a) =]a - \epsilon, a + \epsilon[.$$

1.34 Satz: *Offene Intervalle sind* ***offene Mengen****. D.h. für jedes x in einem offenen Intervall I gibt es ein $\epsilon_x > 0$, so dass*

$$U_{\epsilon_x}(x) =]x - \epsilon_x, x + \epsilon_x[\subseteq I$$

ist.

1.35. Bemerkung: Es gibt natürlich noch andere offene Mengen in $\mathbb{R}$. So ist etwa die Vereinigung zweier oder mehr offener Intervalle wieder offen.

1.5 Die Komplexen Zahlen

1.36. Bemerkung: Wir wollen einen Körper konstruieren, der den Körper $\mathbb{R}$ als Teilkörper enthält, und außerdem ein Element i mit

$$i^2 = -1.$$

Wenn K ein solcher Körper ist, so muss K alle Elemente der Form

$$z = a + bi, \qquad a, b \in \mathbb{R}$$

enthalten. Es muss gelten

$$\begin{aligned}(a + ib) + (c + id) &= (a + c) + i(b + d)\\ (a + ib) \cdot (c + id) &= (ac - bd) + i(ad + bc).\end{aligned}$$

Damit ist klar, wie die Körperoperationen für diese Elemente definiert werden müssen. Wegen der Eindeutigkeit der neutralen Elemente müssen

$$0, 1 \in \mathbb{R} \subseteq K$$

auch die neutralen Elemente in K sein. Die Inversen von $a + ib$ berechnen sich zu

$$\begin{aligned}-(a + ib) &= (-a) + i(-b)\\ \frac{1}{a + ib} &= \frac{a}{a^2 + b^2} - i\frac{b}{a^2 + b^2}\end{aligned}$$

Damit ist klar dass die Teilmenge der Elemente $a + ib$ von K wieder ein Körper bildet, und wir können K auf diese Elemente beschränken.

1.37 Satz: *Definiert man auf*

$$\mathbb{R}^2 = \{(a, b) : a, b \in \mathbb{R}\}$$

die Operationen

$$\begin{aligned}(a, b) + (c, d) &= (a + c, b + d)\\ (a, b) \cdot (c, d) &= (ac - bd, ad + bc),\end{aligned}$$

so wird der $\mathbb{R}^2$ zu einem Körper, den wir den ***Körper der komplexen Zahlen*** *$\mathbb{C}$ nennen.*

1.38. Definition: Wir betten $\mathbb{R}$ als Teilmenge in $\mathbb{C}$ ein, indem wir

$$a = (a, 0) \qquad \text{für alle } a \in \mathbb{R}$$

identifizieren. Außerdem setzen wir

$$i = (0, 1).$$

1.39. Bemerkung: Mit diesen Bezeichnungen lassen sich alle Elemente von $\mathbb{C}$ eindeutig in der Form

$$z = a + ib, \qquad a, b \in \mathbb{R}$$

schreiben. Man nennt a den **Realteil** von z und b den **Imaginärteil** von z

$$a = \mathrm{Re}\,(z), \quad b = \mathrm{Im}\,(z).$$

1.40. Definition: Für $z = a + ib \in \mathbb{C}$ schreiben wir

$$|z| = \sqrt{a^2 + b^2}.$$

Den Abstand zweier komplexer Zahlen $z, w \in \mathbb{C}$ definieren wir durch

$$d(z, w) := |z - w|.$$

Außerdem setzen wir die **konjugiert komplexe Zahl** von $z = a + ib$ gleich

$$\overline{z} = a - ib.$$

1.41. Bemerkung: Damit gilt

$$|z|^2 = z\overline{z},$$

sowie

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

1.42. Bemerkung: Der Betrag und der Abstand stimmen für $z \in \mathbb{R}$ mit den reellen Definitionen überein. Wir definieren wieder die ϵ-Umgebung von z wie in jedem metrischen Raum durch

$$U_\epsilon(z) := \{w \in \mathbb{C} : d(z, w) < \epsilon\}.$$

Im Fall von $\mathbb{C}$ ist dies nach dem Satz von Pythagoras das Innere eines Kreises mit Radius ϵ um z.

1.43 Satz: *$\mathbb{C}$ wird mit dieser Abstandsfunktion zu einem metrischen Raum.*

1.44. Bemerkung: $\mathbb{C}$ lässt sich nicht anordnen. Denn wir wissen, dass in jedem angeordneten Körper $i^2 = -1 > 0$ gelten müsste. Andererseits ist in jedem angeordnente Körper $1 = 1^2 > 0$.

Kapitel 2

Folgen

Es ist intuitiv klar, was unter eine Folge

$$a_1, a_2, a_3, \ldots$$

von reellen Zahlen zu verstehen ist. Allerdings benötigen wir zur Konstruktion von Folgen und auch zum Beweis von komplizierteren Eigenschaften von Folgen das Induktionsprinzip. Wir müssen uns daher zunächst mit den wesentlichen Eigenschaften der natürlichen Zahlen beschäftigen.

2.1 Die natürlichen Zahlen

Intuitiv sind die natürlichen Zahlen diejenigen Elemente von $\mathbb{R}$, die man durch Zählen

$$1, 2, 3, \ldots,$$

also durch die Abbildung $n \mapsto n+1$ ausgehend von 1 erreichen kann. Wir fassen diese Vorstellung etwas genauer in Worte.

2.1. Definition: Eine Teilmenge $N \subseteq \mathbb{R}$ heißt **induktiv**, wenn $1 \in N$ ist und für alle $n \in \mathbb{R}$

$$n \in \mathbb{N} \;\Rightarrow\; n+1 \in \mathbb{N}$$

gilt. Die Menge $\mathbb{N}$ ist die *kleinste* induktive Teilmenge von $\mathbb{R}$. Wir bezeichnen ihre Elemente als **natürliche Zahlen**. Man definiert außerdem

$$\mathbb{N}_0 = \mathbb{N} \cup \{0\}.$$

2.2 Satz: *In $\mathbb{N}$ gilt das Prinzip der* ***vollständigen Induktion****. D.h., wenn eine Aussage für 1 gilt, und wenn man zeigen kann, dass aus der Gültigkeit der Aussage für $n \in \mathbb{N}$ auch die Gültigkeit für $n+1 \in \mathbb{N}$ folgt, so gilt die Aussage für alle $n \in \mathbb{N}$.*

2.3. Beispiel: Um ein nicht triviales Beispiel für Induktion zu geben, zeigen wir die **Ungleichung von Bernoulli**

$$(1+x)^n \geq 1 + nx \qquad \text{für alle } x \geq -1,\ n \in \mathbb{N}.$$

Wir fixieren ein $x \geq -1$. Für $n = 1$ gilt diese Aussage. Diesen Teil des Beweises bezeichnet man als **Induktionsanfang**. Wenn sie für $n \in \mathbb{N}$ gilt, so zeigen wir, dass sie auch für $n + 1$ gilt. Denn

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

Diesen Teil des Beweises bezeichnet man als **Induktionsschritt**.

2.4 Satz: *Es gilt*

$$n \geq 1 \qquad \text{für alle } n \in \mathbb{N}.$$

Außerdem

$$n + m \in \mathbb{N}, \quad nm \in \mathbb{N} \qquad \text{für alle } n, m \in \mathbb{N},$$

und

$$n > m \ \Rightarrow \ n - m \in \mathbb{N}$$

für alle $n, m \in \mathbb{N}$. *Wenn* $n \in \mathbb{N}$ *ist, dann enthält das Intervall* $]n, n+1[$ *kein Element von* $\mathbb{N}$.

2.5 Satz: $\mathbb{N}$ *ist nicht nach oben beschränkt.*

Beweis: Angenommen doch. Dann hätte $\mathbb{N}$ ein Supremum s. Also existiert ein $n \in \mathbb{N}$ mit

$$s - 1 < n < s.$$

Es folgt $s < n + 1$. Dies ist ein Widerspruch dazu, dass s Supremum von $\mathbb{N}$ sein soll. $\square$

2.6. Bemerkung: Es folgt, dass es zu $0 < a < b$ ein $n \in \mathbb{N}$ gibt mit

$$b < na.$$

Diesen Sachverhalt bezeichnet man als **Axiom von Archimedes**.

2.7 Satz: *Jede Teilmenge von* $\mathbb{N}$ *hat ein Minimum.*

2.2 Folgen

Wie schon gesagt, ist der Begriff der Folge intuitiv klar. Man hat einfach für jedes $n \in \mathbb{N}$ ein $a_n \in \mathbb{R}$, also eine Abbildung $n \mapsto a_n$.

2.8. Definition: Sei M eine Menge. Eine Abbildung $a : \mathbb{N} \to M$ heißt **Folge** in M. Wir schreiben meist

$$a_n := a(n).$$

Die gesamte Folge wird als $(a_n)_{n \in \mathbb{N}}$ bezeichnet. Wir werden manchmal einfach von der Folge der a_n, $n \in \mathbb{N}$, sprechen.

2.9. Definition: Eine Folge in M kann **induktiv** definiert werden. Wir legen dabei $x_1 \in M$ irgendwie fest, und nehmen an, dass x_{n+1} mit Hilfe der vorigen

Folgenglieder $x_1, \dots, x_n$ und $n \in \mathbb{N}$ eindeutig bestimmt werden kann. Man nennt solche Folgen auch **rekursiv** definierte Folgen.

2.10 Satz: *Auf diese Weise wird eine Folge eindeutig definiert.*

2.11. Bemerkung: Man kann die Induktion auch mit x_0 beginnen. Die Folge ist dann für alle $n \in \mathbb{N}_0$ definiert.

2.12. Definition: Die Folge der **Fakultäten** $n!$, $n \in \mathbb{N}$, ist induktiv durch

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

definiert. Die Folge der **Potenzen** x^n, $n \in \mathbb{N}$, ist für $x \in \mathbb{N}$ induktiv durch

$$x^0 = 1, \quad x^{n+1} = x^n \cdot x$$

definiert.

2.13. Bemerkung: Man beachte, dass mit dieser Definition $0^0 = 1$ definiert wird, aber $0^n = 0$ für alle $n \in \mathbb{N}$.

2.14. Definition: Wir setzen für $n \in \mathbb{N}$ und $x \neq 0$

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

2.15. Beispiel: Es gibt auch zweigliedrige rekursiv gegebene Folgen. So ist die **Fibonacci**-Folge durch

$$\begin{aligned} x_0 = 0, \quad x_1 = 1,& \\ x_{n+1} = x_n + x_{n-1} \qquad & \text{für alle } n \in \mathbb{N} \end{aligned}$$

gegeben. Da man zur Berechnung von x_{n+1} zwei vorherige Glieder braucht, ist es nötig, die ersten *beiden* Folgenglieder vorzugeben. Man kann durch vollständige Induktion beweisen, dass

$$x_n = \frac{1}{\sqrt{5}} \left(\left(\frac{1+\sqrt{5}}{2} \right)^n - \left(\frac{1-\sqrt{5}}{2} \right)^n \right)$$

für alle $n \in \mathbb{N}_0$ gilt.

2.3 Endliche und abzählbare Mengen

2.16. Definition: Für $n \in \mathbb{N}$ definieren wir das **Anfangsstück** von $\mathbb{N}$

$$\{1, \dots, n\} := \{m \in \mathbb{N} : 1 \leq m \leq n\}.$$

Eine Menge M heißt **endlich**, wenn sie leer ist oder wenn es eine bijektive Abbildung

$$f : \{1, \dots, n\} \to M$$

für ein $n \in \mathbb{N}$ gibt. In diesem Fall schreiben wir für die **Mächtigkeit** vom M

$$|M| := n.$$

Die Mächtigkeit der leeren Menge ist 0. Eine Menge M heißt **abzählbar**, wenn es eine bijektive Abbildung

$$f : \mathbb{N} \to M$$

gibt.

2.17 Satz: *Die Mächtigkeit einer endlichen Menge ist eindeutig bestimmt. Wenn es also für $n, m \in \mathbb{N}$ eine bijektive Abbildung*

$$f : \{1, \ldots, n\} \to \{1, \ldots, m\}$$

gibt, so muss $n = m$ sein.

2.18 Satz: (1) *Eine abzählbare Menge ist nicht endlich. Jede endliche Teilmenge von $\mathbb{N}$ hat ein Maximum.*

(2) *Teilmengen von abzählbaren Mengen sind endlich oder abzählbar. Jede Teilmenge M einer endlichen Menge N ist endlich und es gilt*

$$|M| + |N \setminus M| = |N|.$$

(3) *Die Vereinigung von endlich vielen endlichen Mengen ist endlich. Insbesondere gilt für zwei endliche Mengen M_1, M_2*

$$|M_1 \cup M_2| = |M_1| + |M_2| - |M_1 \cap M_2|.$$

Die Vereinigung von endlich oder abzählbar vielen endlichen oder abzählbaren Mengen ist endlich oder abzählbar.

(4) *Das Bild einer endlichen Menge unter eine Abbildung ist endlich. Das Bild einer abzählbaren Menge unter eine Abbildung ist endlich oder abzählbar.*

2.19. Definition: Analog zur Definition einer Folge definieren wir ein **Tupel** von Elementen einer Menge als eine Abbildung

$$a : \{1, \ldots, n\} \to M$$

Wir schreiben für das Tupel einfach $(a_1, \ldots, a_n)$, oft auch ohne die Klammern, also einfach in der Form

$$a_1, \ldots, a_n \in M.$$

2.20. Bemerkung: Das Tupel $(a_1, \ldots, a_n)$ ist etwas anderes als die Menge $\{a_1, \ldots, a_n\}$. Denn erstens kommt es beim Tupel auf die Reihenfolge an, und zweitens werden Wiederholungen beachtet. Dasselbe gilt für den Unterschied von Folgen und abzählbaren Mengen. Zwar kann man jede abzählbare Menge M als Bild einer Folge

$$M = \{a_1, a_2, a_3, \ldots\}$$

schreiben, aber dies ist dennoch verschieden von der Folge $(a_n)_{n \in \mathbb{N}}$, bei der Wiederholungen von Folgengliedern wichtig sind.

2.4 Summen und Produkte

Dieser Abschnitt dient nur dazu, eine wichtige Schreibweise für Summen und Produkte von endlich oder abzählbar vielen Elementen in $\mathbb{R}$ zu definieren. Auf die Dauer ist die Schreibweise

$$a_1 + \ldots + a_n,$$

zu unpräzise und auch zu platzgreifend.

2.21. Definition: Wir definieren

$$\sum_{k=1}^{n} a_k := a_1 + \ldots + a_n$$
$$\prod_{k=1}^{n} a_k := a_1 \cdot \ldots \cdot a_n$$

2.22. Bemerkung: Man kann diese Schreibweisen exakt mit induktiver Definition einführen. Also etwa

$$\sum_{k=1}^{1} a_k = a_1$$
$$\sum_{k=1}^{n+1} a_k = \left(\sum_{k=1}^{n} a_k \right) + a_{n+1}.$$

2.23. Bemerkung: Die Summe und das Produkt kann auch mit einem anderen Index anfangen, also etwa

$$\sum_{k=m}^{n} a_k := a_m + a_{m+1} + \ldots + a_n$$

für $n \geq m$. Für $n < m$ wird die Summe als 0 definiert, das Produkt als 1.

2.24. Beispiel: Die Fakultät und die Potenz schreiben sich als

$$n! = \prod_{k=1}^{n} k, \quad x^n = \prod_{k=1}^{n} x.$$

2.25. Beispiel: (1) Man zeigt mit vollständiger Induktion

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

für alle $n \in \mathbb{N}$. Dies ist die **arithmetische Summe**.

(2) Ebenso die **geometrische Summe**

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

für alle $x \neq 1$ und $n \in \mathbb{N}$.

2.26. Definition: Für $n \in \mathbb{R}$ und $m \in \mathbb{N}_0$ definieren wir den **Binomialkoeffizient**

$$\binom{n}{m} = \frac{n \cdot (n-1) \cdot \ldots \cdot (n-(m-1))}{m!} = \frac{1}{m!} \prod_{k=n-m+1}^{n} k.$$

2.27. Bemerkung: Im Fall $m = 0$ ist das Produkt leer, und es gilt daher

$$\binom{n}{0} = 1$$

für alle $n \in \mathbb{R}$.

2.28 Satz: *Für $n, m \in \mathbb{N}_0$, $0 \leq m \leq n$, gilt*

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

Außerdem gilt für alle $n \in \mathbb{R}$, $m \in \mathbb{N}_0$

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

2.29. Bemerkung: Dieser Sachverhalt führt zum **Pascalschen Dreieck**

	$m=0$	$m=1$	$m=2$	$m=3$	$m=4$	$\ldots$
$n=0$	1					
$n=1$	1	1				
$n=2$	1	2	1			
$n=3$	1	3	3	1		
$n=4$	1	4	6	4	1	
$\vdots$						

2.30 Satz: *Es gilt die* ***Binomialentwicklung***

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

für alle $a, b \in \mathbb{R}$ und $n \in \mathbb{N}$.

2.31 Satz: *Eine endliche Menge der Mächtigkeit n hat 2^n verschiedene Teilmengen, und*

$$\binom{n}{m}$$

verschiedene Teilmengen der Mächtigkeit m, $0 \le m \le n$.

2.32. Definition: Die Menge

$$\mathcal{P}(M) := \{T : T \subseteq M\}$$

heißt **Potenzmenge** von M.

2.33 Satz: *Für eine endliche Menge M gilt*

$$|\mathcal{P}(M)| = 2^{|M|}.$$

2.5 Die rationalen Zahlen

Das Rechnen mit Brüchen ist aus der Schule bekannt, etwa in der Form von Dezimalbrüchen. Wir werden hier nur die wichtigsten Eigenschaften des Körpers der rationalen Zahlen festhalten. Die Besprechung der Dezimalbrüche heben wir uns auf, bis wir die Konvergenz zur Verfügung haben.

2.34. Definition: Wir definieren die Menge der **ganzen Zahlen**

$$\mathbb{Z} := \{m \in \mathbb{R} : m \in \mathbb{N} \text{ oder } m = 0 \text{ oder } -m \in \mathbb{N}\},$$

sowie die Menge der **rationalen Zahlen**

$$\mathbb{Q} := \{\frac{m}{n} : m \in \mathbb{Z} \text{ und } n \in \mathbb{N}\}.$$

Beides sind natürlich Teilmengen von $\mathbb{R}$.

2.35 Satz: *Die Summe, die Differenz und das Produkt von ganzen Zahlen ist wieder ganz. Die ganzen Zahlen bilden einen kommutativen* ***Ring****. D.h., es gelten alle Körperaxiome bis auf die Existenz des Inversen.*

2.36 Satz: *Die rationalen Zahlen bilden einen angeordneten Körper.*

2.37. Bemerkung: Es gibt zwischen $\mathbb{Q}$ und $\mathbb{R}$ noch weitere Körper. Zum Beispiel erfüllt auch

$$K := \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$$

die Axiome eines angeordneten Körpers und es gilt $\mathbb{Q} \subset K \subset \mathbb{R}$.

2.38 Satz: *Die rationalen Zahlen sind abzählbar, und sie sind* ***dicht*** *in $\mathbb{R}$. D.h., es gibt zu jedem $a \in \mathbb{R}$ und jedem $\epsilon > 0$ ein $q \in \mathbb{Q}$ mit*

$$|a - q| < \epsilon.$$

Daraus folgt, dass es in jedem nicht leeren, offenen Intervall $]a, b[$ unendlich viele rationale Zahlen gibt.

2.39. Bemerkung: Es gilt $\sqrt{2} \notin \mathbb{Q}$. Wir werden erst später zeigen, dass $\mathbb{R}$ nicht abzählbar (**überabzählbar**) ist.

Kapitel 3

Konvergenz

Konvergenz ist einer der wesentlichen Grundbegriffe der Analysis. Sie wird in der Schule meist anschaulich mit Hilfe des Begriffs der „beliebigen Annäherung“ eingeführt. Wir präzisieren hier diesen Begriff. Außerdem wollen wir mit Grenzwerten rechnen lernen.

Dieses Kapitel untersucht zunächst die Konvergenz von Folgen und Reihen. Den Grenzübergang bei Funktionen werden wir im nächsten Kapitel kennen lernen.

3.1 Grenzwert von Folgen

Was bedeutet es, dass sich eine reelle Folge $a_1, a_2, \ldots$ einem Wert $a \in \mathbb{R}$ „beliebig annähert“. Dazu kann es nicht genügen, dass einzelne Folgenglieder gleich a werden oder gar nur in der Nähe von a liegen, sondern es müssen alle Folgenglieder ab einen gewissen Index in beliebig kleinem, vorgegebenem Abstand zu a liegen.

3.1. Definition: Die Folge der $a_n \in \mathbb{R}$, $n \in \mathbb{N}$, **konvergiert** genau dann gegen $a \in \mathbb{R}$, wenn es zu jedem (beliebig kleinen) $\epsilon > 0$ ein $N_\epsilon \in \mathbb{N}$ gibt, so dass

$$|a_n - a| < \epsilon \qquad \text{für alle } n \geq N_\epsilon.$$

Man kann das so formulieren, dass für alle $\epsilon > 0$ außerhalb der Umgebung $(a-\epsilon, a+\epsilon)$ nur endlich viele Folgenglieder liegen dürfen. In Fall der Konvergenz bezeichnen wir a als den **Grenzwert** der Folge und schreiben

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

Wenn Klarheit besteht, dass $n \to \infty$ gehen soll, so lassen wir diesen Zusatz manchmal weg, oder wir schreiben auch

$$a_n \to a$$

mit dem Zusatz „für $n \to \infty$“.

3.2 Satz: *Der Grenzwert einer Folge ist eindeutig, wenn er existiert.*

3.3. Beispiel: (1) Es gilt

$$\lim_{n\to\infty} \frac{1}{n} = 0.$$

(2) Für $|q| < 1$ gilt

$$\lim_{n\to\infty} q^n = 0.$$

(3) Die Folge $a_n = (-1)^n$, $n \in \mathbb{N}$, konvergiert nicht.

3.4 Satz: (1) *Man kann „+“, „−“ und „·“ mit dem Grenzwert vertauschen, also*

$$\lim_{n\to\infty} (a_n \pm b_n) = \lim_{n\to\infty} a_n + \lim_{n\to\infty} b_n$$
$$\lim_{n\to\infty} a_n b_n = \left(\lim_{n\to\infty} a_n\right) \cdot \left(\lim_{n\to\infty} b_n\right).$$

Die Grenzwerte auf der linken Seite existieren, wenn die Grenzwerte auf der rechten Seite existieren. Falls $b_n \neq 0$ ist für alle $n \in \mathbb{N}$ und $\lim b_n \neq 0$, so gilt auch

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

(2) *Wenn a_n, b_n, $n \in \mathbb{N}$, konvergente Folgen sind und*

$$a_n \leq b_n \qquad \text{für alle } n \in \mathbb{N},$$

so gilt

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

(3) *Wenn a_n, b_n, c_n, $n \in \mathbb{N}$, Folgen sind mit*

$$a_n \leq b_n \leq c_n$$

und die Folgen a_n, c_n konvergieren mit

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} c_n,$$

so konvergiert auch die Folge der b_n gegen denselben Grenzwert. Dieses Prinzip nennt man das ***Sandwich-Prinzip****.*

3.5. Definition: Eine Folge von $a_n \in \mathbb{R}$, $n \in \mathbb{N}$, heißt **monoton** wachsend, wenn

$$a_n \leq a_{n+1} \qquad \text{für alle } n \in \mathbb{N}$$

gilt. Falls sogar „<“ gilt, so heißt die Folge **streng monoton** wachsend. Analog definieren wir monoton fallend und streng monoton fallend. Falls es ein $c \in \mathbb{R}$ gibt, so dass

$$a_n \leq c \qquad \text{für alle } n \in \mathbb{N}$$

gilt, so heißt die Folge nach oben **beschränkt**. Analog definieren wir nach unten beschränkte Folgen.

3.6 Satz: *Jede monoton wachsende, nach oben beschränkte Folge in $\mathbb{R}$ konvergiert. Jede monoton fallende, nach unten beschränkte Folge konvergiert.*

3.7. Beispiel: Die Folge

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

ist monoton fallend und nach unten durch 1 beschränkt. Sie konvergiert also gegen einen Grenzwert, den man **Eulersche Zahl** e nennt. Man kann e einfacher als

$$e := \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^{n}$$

definieren. Es gilt $e = 2.71828\ldots$ Es ist bekannt, dass $e \notin \mathbb{Q}$ ist.

3.8 Satz: *Als* ***Intervallschachtelung*** *bezeichnen wir eine Folge von Intervallen $[a_n, b_n]$ mit*

$$[a_{n+1}, b_{n+1}] \subseteq [a_n, b_n] \qquad \textit{für alle } n \in \mathbb{N}$$

und

$$\lim_{n\to\infty} b_n - a_n = 0.$$

Für solche Intervallschachtelungen gibt es genau ein $x \in \mathbb{R}$, das in allen Intervallen liegt und es gilt

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

3.9. Beispiel: Mit Hilfe der Intervallschachtelung kann man numerische Näherungen berechnen. Zur Berechnung von $\sqrt{2}$ starten wir etwa mit dem Intervall

$$[a_0, b_0] = [1, 2].$$

Wir wissen

$$a_0^2 < 2 < b_0^2.$$

Nun berechnen wir in jedem Schritt

$$m_n = \frac{1}{2}(a_n + b_n).$$

Dies ist die Mitte des Intervalls $[a_n, b_n]$. Wir setzen nun

$$[a_{n+1}, b_{n+1}] := \begin{cases} [a_n, m_n] & \text{falls } m_n^2 \le 2, \\ [m_n, b_n] & \text{falls } m_n^2 > 2. \end{cases}$$

Man erhält

$$[a_{n+1}, b_{n+1}] \subset [a_n, b_n]$$
$$b_n - a_n = \frac{1}{2^n}$$

Dies ist eine Intervallschachtelung. Man zeigt, dass sie gegen $\sqrt{2}$ konvergiert, indem man

$$\lim_{n\to\infty} a_n^2 \leq 2 \leq \lim_{n\to\infty} b_n^2$$

ausnutzt. Dieses numerische Verfahren heißt **Bisektionsverfahren**.

3.10. Bemerkung: Ein wichtiges Prinzip ist, dass im Fall

$$a_n \to a > b$$

ein $N \in \mathbb{N}$ existiert mit

$$a_n > b \qquad \text{für alle } n > N.$$

Wenn der Grenzwert einer Folge positiv ist, so ist die Folge ab einem Index positiv, ja sie ist sogar ab einem gewissen Index durch ein $c > 0$ nach unten beschränkt.

3.11. Beispiel: Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{R}$ mit $a_n \geq 0$ für alle $n \in \mathbb{N}$ Dann gilt

$$a_n \to a \;\Leftrightarrow\; \sqrt{a_n} \to \sqrt{a}.$$

Zum Beweis beachten wir im Fall $a > 0$

$$|\sqrt{a_n} - \sqrt{a}| = \frac{|a_n - a|}{\sqrt{a_n} + \sqrt{a}} \leq c|a_n - a|$$

für eine Konstante $c > 0$. Der Fall $a = 0$ ist einfach.

3.12. Beispiel: Bei rekursiv gegebenen Folgen kann man den Grenzwert durch Grenzübergang auf beiden Seiten der Rekursionsgleichung berechnen, wenn man weiß, dass er existiert. Wir betrachten für festes $a \in [0, 1]$ die rekursiv gegebene Folge mit $a_0 = 0$.

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

Man zeigt mit vollständiger Induktion

$$a_n \leq a_{n+1} \leq \sqrt{a} \qquad \text{für alle } n \in \mathbb{N}.$$

Dazu wird

$$\frac{\sqrt{a} - a_n}{a - a_n^2} = \frac{1}{\sqrt{a} + a_n} \geq \frac{1}{2}$$

verwendet. Also konvergiert a_n und wir können in der Rekursionsgleichung auf beiden Seiten den Grenzwert nehmen. Man erhält

$$\lim_{n\to\infty} a_n = \sqrt{a}.$$

3.2 Häufungspunkte

Die Reihe $(-1)^n$ konvergiert zwar nicht, sie nimmt aber nur zwei Werte ± 1 je unendlich oft an. Die Frage ist, ob sich nicht konvergente, aber beschränkte Folgen immer so verhalten.

3.13. Definition: Sei a_n, $n \in \mathbb{N}$, eine reelle Folge. Ein Punkt $a \in \mathbb{R}$ heißt **Häufungspunkt** der Folge, wenn es zu jedem $\epsilon > 0$ und jedem $N \in \mathbb{N}$ ein

$n > N$ gibt mit

$$|a - a_n| < \epsilon.$$

Es müssen also in jeder noch so kleinen Umgebung von a unendlich viele Folgenglieder liegen.

3.14. Definition: Eine Teilfolge der Folge a_n, $n \in \mathbb{N}$ ist eine Folge von $a_{k(n)}$, $n \in \mathbb{N}$, wobei $k(n)$ eine streng monoton wachsende Folge in $\mathbb{N}$ ist. Also werden aus der Folge a_n die Folgenglieder

$$a_{k(1)}, a_{k(2)}, \ldots$$

zu einer neuen Folge zusammen gefasst.

3.15 Satz: *a ist genau dann Häufungspunkt der Folge a_n, $n \in \mathbb{N}$, wenn es eine Teilfolge dieser Folge gibt, die gegen a konvergiert.*

3.16. Beispiel: Sei $a_n = (-1)^n$. Dann ist die Teilfolge $a_{2n} = 1$ für alle $n \in \mathbb{N}$ und konvergiert deswegen gegen 1, und die Teilfolge $a_{2n-1} = -1$ für alle $n \in \mathbb{N}$ und konvergiert deswegen gegen -1. 1 und -1 sind daher die einzigen Häufungspunkte von a_n.

3.17 Satz: *Wenn eine Folge gegen einen Grenzwert konvergiert, so gilt dies auch für jede Teilfolge.*

3.18 Satz: *Jede beschränkte Folge in $\mathbb{R}$ hat einen Häufungspunkt.*

Beweis: Man konstruiert eine Intervallschachtelung von Intervallen $[a_n, b_n]$, so dass jedes Intervall unendlich viele Folgenglieder enthält. Der Grenzwert der Intervallschachtelung ist dann Häufungspunkt der Folge. □

3.19. Bemerkung: Diese Bedingung wird gerne als Ersatz für unser Vollständigkeitsaxiom genommen.

3.3 Uneigentliche Grenzwerte

Es ist in diesem Zusammenhang sinnvoll $\pm\infty$ als Grenzwert zuzulassen. Mit Einschränkungen können wir sogar mit ∞ rechnen. Wir können $\mathbb{R}$ allerdings nicht einfach um diese Punkte erweitern, da $\infty - \infty$ etc. nicht definiert wäre.

3.20. Definition: Eine Folge von $a_n \in \mathbb{R}$, $n \in \mathbb{N}$, konvergiert **uneigentlich** gegen ∞, wenn es zu jedem (beliebig großen) $c > 0$ ein $N_c \in \mathbb{N}$ gibt mit

$$a_n \geq c \qquad \text{für alle } n \geq N_c.$$

Wir schreiben

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

Analog definieren wir uneigentliche Konvergenz gegen $-\infty$. Der Zusatz „uneigentlich" wird gelegentlich weg gelassen.

3.21 Satz: (1) *Eine Folge a_n positiver Zahlen konvergiert genau dann uneigentlich gegen ∞, wenn die Folge $1/a_n$ gegen 0 konvergiert.*

(2) *a_n konvergiert genau dann uneigentlich gegen ∞, wenn $-a_n$ uneigentlich gegen $-\infty$ konvergiert.*

3.22. Bemerkung: Beim Rechnen mit Grenzwerten kann man also

$$\frac{a}{\infty} = \frac{a}{-\infty} = 0$$

für alle $a \geq 0$ setzen. Allerdings kann man nicht $1/0 = \infty$ setzen, wie die Folge

$$a_n = \frac{(-1)^n}{n}$$

zeigt, deren Kehrwert $1/a_n$ nicht einmal uneigentlich konvergiert. Natürlich kann man aber

$$\infty \cdot \infty = \infty + \infty = \infty$$

verwenden. Aber auf keinen Fall darf man einfach $\infty - \infty = 0$ setzen!

3.23. Beispiel: Durch geeignetes Kürzen der Brüche zeigt man

$$\lim_{n\to\infty} \frac{n-1}{n+1} = 1, \quad \lim_{n\to\infty} \frac{n-1}{n^2+1} = 0, \quad \lim_{n\to\infty} \frac{n^2-1}{n+1} = \infty.$$

3.4 Grenzwert von Reihen

3.24. Definition: Sei a_k, $k \in \mathbb{N}$, eine Folge in $\mathbb{R}$. Dann kann man die Folge der **Partialsummen**

$$s_n = \sum_{k=1}^{n} a_k$$

bilden. Diese Folge von immer länger werdenden Summen nennt man eine **Reihe**. Wenn die Reihe (uneigentlich) konvergiert so schreiben wir für den Grenzwert

$$\sum_{k=1}^{\infty} a_k := \lim_{n\to\infty} \sum_{k=1}^{n} a_k.$$

Bisweilen wird allerdings für die Reihe selbst, also für die Folge der Partialsummen, diese Notation verwendet, obwohl die Reihe nicht konvergiert. Man nennt dies dann auch eine **formale** Reihe.

3.25. Beispiel: Ein wichtiges Beispiel ist die **geometrische** Reihe. Es gilt für $|x| \leq 1$

$$\sum_{k=1}^{\infty} x^k = \frac{1}{1-x},$$

und diese Reihe konvergiert auch nur für $|x| < 1$.

3.26 Satz: *Wenn die Reihe*

$$\sum_{k=1}^{\infty} a_k$$

konvergiert, so gilt

$$\lim_{k\to\infty} a_k = 0.$$

3.27. Beispiel: Die Umkehrung ist falsch. Es gilt

$$\sum_{k=1}^{\infty} \frac{1}{k} = \infty$$

Diese Reihe, die man **harmonische Reihe** nennt, konvergiert also nicht, sondern nur uneigentlich. Dies beweist man mit Hilfe des Cauchyschen Verdichtungskriteriums, das besagt, dass für eine monoton fallende Folge

$$a_1 \geq a_2 \geq \ldots > 0$$

von positiven Zahlen die Reihe

$$\sum_{k=1}^{\infty} a_k$$

genau dann konvergiert, wenn die Reihe

$$\sum_{k=1}^{\infty} 2^k a_{2^k}$$

konvergiert.

3.28 Satz: *Sei a_k, $k \in \mathbb{N}$, eine Folge in $\mathbb{R}$. Wenn a_n monoton fällt und*

$$\lim_{k\to\infty} a_k = 0$$

ist, so konvergiert die ***alternierende*** *Reihe*

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

Man nennt dieses Konvergenzkriterium das ***Leibniz-Kriterium****.*

3.5 Absolute Konvergenz

3.29. Definition: Sei a_k, $k \in \mathbb{N}$, eine reelle Folge. Eine Reihe

$$\sum_{k=1}^{\infty} a_k$$

konvergiert **absolut**, wenn

$$\sum_{k=1}^{\infty} |a_k|$$

konvergiert, wenn also

$$\sum_{k=1}^{\infty} |a_k| < \infty$$

gilt.

3.30 Satz: *Jede absolut konvergente Reihe konvergiert.*

Ziel dieses Abschnittes ist, diesen Satz zu beweisen. Wir werden dazu den Begriff einer **Cauchy-Folge** benötigen.

3.31. Definition: Eine Folge s_n, $n \in \mathbb{N}$, reeller Zahlen heißt **Cauchy-Folge**, wenn für jedes $\epsilon > 0$ ein $N_\epsilon \in \mathbb{N}$ existiert mit

$$|s_n - s_m| < \epsilon \qquad \text{für alle } n, m > N_\epsilon$$

3.32. Bemerkung: Äquivalent dazu ist, dass es zu jedem (beliebig kleinen) $\epsilon > 0$ ein $x_\epsilon \in \mathbb{R}$ gibt, so dass alle, bis auf endlich viele Folgenglieder in der ϵ-Umgebung von x_ϵ liegen, also

$$|x_\epsilon - s_n| < \epsilon \qquad \text{für alle } n \geq N.$$

Man bemerke, dass x_ϵ von ϵ abhängt. Dies ist der Unterschied zur Definition des Grenzwertes, und x_ϵ ist auch nicht unbedingt der Grenzwert der Folge. Es folgt daraus, dass jede konvergente Folge Cauchy-Folge ist.

3.33 Satz: *In $\mathbb{R}$ konvergiert jede Cauchy-Folge.*

Beweis: Cauchy-Folgen sind beschränkt und haben daher einen Häufungspunkt. Man zeigt leicht, dass eine Cauchy-Folge mit Häufungspunkt gegen diesen Häufungspunkt konvergiert. □

3.34. Bemerkung: Dieser Satz, zusammen mit dem Axiom von Archimedes, ist äquivalent zu unserem Vollständigkeitsaxiom. Er ist auch äquivalent dazu, dass alle absolut konvergenten Reihen konvergieren.

Beweis: (Satz über absolut konvergente Reihen) Man zeigt, dass die Folge der Partialsummen eine Cauchy-Folge ist, wenn die Reihe absolut konvergiert. Denn es gilt für $n < m$

$$|\sum_{k=1}^{m} a_k - \sum_{k=1}^{n} a_k| \leq \sum_{k=n+1}^{m} |a_k| = \sum_{k=1}^{m} |a_k| - \sum_{k=1}^{n} |a_k|$$

und die rechte Seite kann beliebig klein gemacht werden. □

3.35 Satz: *Es gilt das* ***Majoranten-Kriterium****. D.h., wenn eine Reihe*

$$\sum_{k=1}^{\infty} a_k$$

absolut konvergiert, und für die Folge b_k, $k \in \mathbb{N}$,

$$|b_k| \leq |a_k| \qquad \textit{für alle } k \geq K \in \mathbb{N}$$

gilt, dann konvergiert auch

$$\sum_{k=1}^{\infty} b_k$$

absolut.

3.36. Beispiel: Es gilt

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

Man bezeichnet solche Summen als **Teleskop-Summen**. Aus diesem Grund konvergiert die Reihe und es gilt

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

Aus dem Majorantenkriterium folgt, dass auch alle Reihen der Form

$$\sum_{k=1}^{n} \frac{\sigma_k}{k^2}$$

mit beliebigen $|\sigma_k| \leq 1$ konvergieren. Insbesondere konvergiert die Reihe

$$\sum_{k=1}^{n} \frac{1}{k^2},$$

für die die Konvergenz allerdings auch aus dem Cauchyschen Verdichtungskriterium folgt. Für $\sigma_k = (-1)^k$ folgt die Konvergenz (aber nicht die *absolute* Konvergenz) auch aus dem Leibnizkriterium.

3.37. Bemerkung: Es gilt umgekehrt ein **Minoranten-Kriterium**. Wenn a_k, b_k, $k \in \mathbb{N}$, reelle Folgen sind mit

$$0 \leq a_k \leq b_k \qquad \text{für alle } k \geq K \in \mathbb{N}$$

und

$$\sum_{k=1}^{\infty} a_k = \infty,$$

so gilt auch

$$\sum_{k=1}^{\infty} b_k = \infty.$$

3.38 Satz: *Sei a_k, $k \in \mathbb{N}$, eine Folge in $\mathbb{R}$, $a_k \neq 0$ für alle $k \geq K \in \mathbb{N}$. Wenn es ein $q \in \mathbb{R}$ gibt, so dass*

$$\frac{|a_{k+1}|}{|a_k|} < q < 1$$

für alle $k \geq K$ gilt, so konvergiert die Reihe

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

absolut. Dieses Kriterium nennt man das ***Quotienten-Kriterium*** *für Reihen. Es ist insbesondere erfüllt, wenn*

$$\lim_{k\to\infty} \frac{|a_{k+1}|}{|a_k|} < 1$$

ist. Im Fall

$$\lim_{k\to\infty} \frac{|a_{k+1}|}{|a_k|} > 1$$

divergiert die Reihe. Falls der Grenzwert gleich 1 *ist, so kann man keine Aussage treffen.*

3.39. Beispiel: Die Reihen

$$\sum_{k=0}^{\infty} k^m x^k$$

konvergieren für alle $m \in \mathbb{N}_0$ und alle $|x| < 1$. In diesem Fall gilt nämlich

$$\lim_{k\to\infty} \frac{|a_{k+1}|}{|k_n|} = |x| < 1.$$

Für $|x| > 1$ kann man auf Divergenz schließen.

Im Vorgriff auf die Definition der k-ten Wurzel beweisen wir hier den folgenden Satz.

3.40 Satz: *Wenn für eine Folge a_k, $k \in \mathbb{N}$, ein $q \in \mathbb{R}$ existiert mit*

$$\sqrt[k]{|a_k|} < q < 1$$

für alle $k \geq K \in \mathbb{N}$, dann konvergiert die Reihe

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

absolut. Man nennt dieses Kriterium das ***Wurzel-Kriterium****. Falls der Grenzwerte von $\sqrt[k]{|a_k|}$ existiert, gelten analoge Aussagen wie beim Quotientenkriterium.*

3.6 Dezimalbrüche

3.41. Definition: Eine Reihe der Form

$$b_n \ldots b_0.a_1 a_2 a_3 \ldots := \sum_{k=0}^{n} b_k 10^k + \sum_{k=1}^{\infty} a_k 10^{-k}$$

mit

$$a_k, b_k \in \{0, 1, \ldots, 9\}$$

heißt **Dezimalentwicklung**. Analog gibt es auch die Entwicklung in andere Basen als 10.

3.42. Bemerkung: Alle Dezimalentwicklungen konvergieren, da sie monoton wachsende, nach oben beschränkte Folgen darstellen. Allerdings haben verschiedene Dezimalentwicklungen den gleichen Grenzwert. So gilt etwa

$$0.999\ldots = 1.000\ldots = 1$$

3.43 Satz: *Zu jedem $x > 0$ gibt es eine Dezimalentwicklung, die gegen x konvergiert. Dabei kann man verhindern, dass $a_k = 9$ für $k \geq N$ gilt. Wenn man das tut, ist die Entwicklung sogar eindeutig bestimmt.*

3.44. Bemerkung: Mit Hilfe dieses Satzes kann man beweisen, dass die Menge der reellen Zahlen **überabzählbar** ist (**Cantorsches Diagonalargument**).

3.7 Komplexe Folgen und Reihen

3.45. Definition: In einen metrischen Raum M mit Metrik d definieren wir die Konvergenz der Folge $(x_n)_{n\in\mathbb{N}}$ gegen x durch

$$x_n \to x \;\Leftrightarrow\; d(x_n, x) \to 0.$$

Wir definieren, dass $(x_n)_{n\in\mathbb{N}}$ Cauchy-Folge sein soll, wenn es zu jedem $\epsilon > 0$ ein $N_\epsilon \in \mathbb{N}$ gibt mit

$$d(x_n, x_m) < \epsilon \qquad \text{für alle } n, m \geq N_\epsilon.$$

3.46. Bemerkung: Im Folgenden sei $\mathbb{C}$ mit der Metrik

$$d(z, w) := |z - w|$$

ausgestattet.

3.47 Satz: (1) *Sei $(z_n)_{n\in\mathbb{N}}$ eine Folge in $\mathbb{C}$. Dann gilt*

$$z_n \to z \;\Leftrightarrow\; Re\,(z_n) \to Re\,(z) \text{ und } Im\,(z_n) \to Im\,(z)$$

Also hat in $\mathbb{C}$ jede beschränkte Folge eine Häufungspunkt.

(2) *Eine komplexe Folge ist genau dann Cauchy-Folge, wenn die Folge der Realteile und die Folge der Imaginärteile Cauchy-Folgen in $\mathbb{R}$ sind. In $\mathbb{C}$ konvergiert jede Cauchy-Folge.*

Beweis: Der Beweis von (1) folgt aus den Abschätzungen

$$|\mathrm{Re}\,(z_n) - \mathrm{Re}\,(z)| \leq |z_n - z|, \quad |\mathrm{Im}\,(z_n) - \mathrm{Im}\,(z)| \leq |z_n - z|,$$

sowie

$$|z_n - z| = \sqrt{|\mathrm{Re}\,(z_n) - \mathrm{Re}\,(z)|^2 + |\mathrm{Im}\,(z_n) - \mathrm{Im}\,(z)|^2}.$$

Analog folgt die Aussage (2). □

3.48. Bemerkung: Einen metrischen Raum, in dem jede Cauchy-Folge konvergiert, heißt **vollständiger metrischer Raum**.

3.49 Satz: *In $\mathbb{C}$ konvergiert jede absolut konvergente Reihe. Es gilt das Majoranten-Kriterium, das Quotienten-Kriterium und das Wurzel-Kriterium für Reihen.*

3.50. Definition: Wir definieren

$$\exp(z) := \sum_{k=0}^{\infty} \frac{z^k}{k!}$$

für alle $z \in \mathbb{C}$. Diese Reihe konvergiert für alle $z \in \mathbb{C}$ absolut. Wir nennen diese Funktion die **Exponentialfunktion**.

3.51 Satz: *Sei*

$$\sum_{k=1}^{\infty} a_k$$

eine absolut konvergente Reihe in $\mathbb{C}$ oder $\mathbb{R}$ und

$$\pi : \mathbb{N} \to \mathbb{N}$$

bijektiv. Dann konvergiert auch

$$\sum_{k=1}^{\infty} a_{\pi(k)}$$

absolut. Außerdem gilt

$$\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} a_{\pi(k)}.$$

Dieser Satz heißt der ***Umordnungssatz****.*

Beweis: Die Konvergenz folgt einfach aus

$$\sum_{k=1}^{N} |a_{\pi(k)}| \leq \sum_{k=1}^{\infty} |a_k|$$

für alle $N \in \mathbb{N}$. Dass der Grenzwert gleich ist folgt daraus, dass bei gegebenem $N \in \mathbb{N}$ für M und K groß genug gilt

$$\left| \sum_{k=1}^{M} a_{\pi(k)} - \sum_{k=1}^{K} a_k \right| \leq \sum_{k=N+1}^{M} |a_{\pi(k)}| + \sum_{k=1}^{K} |a_k|.$$

$\square$

3.52 Satz: *Seien*

$$a = \sum_{k=0}^{n} a_k, \quad b = \sum_{k=1}^{n} b_k$$

absolut konvergente Reihen. Dann konvergiert auch das ***Cauchy-Produkt***

$$\sum_{k=0}^{\infty}\left(\sum_{l=0}^{k} a_l b_{k-l}\right)$$

gegen ab.

Beweis: Die Folge

$$s_N = \left(\sum_{k=0}^{N} |a_k|\right) \cdot \left(\sum_{k=0}^{N} |b_k|\right) = \left(\sum_{0\le k,l\le N} |a_k b_l|\right)$$

konvergiert. Also ist die Reihe

$$s_0 + \sum_{N=0}^{\infty}(s_{N+1} - s_N) = \sum_{N=0}^{\infty}(a_0 b_N + \ldots + a_N b_N + \ldots + a_N b_0)$$

absolut konvergent, mit Grenzwert ab. Nach dem Umordnungssatz konvergiert jede Reihe, die die Glieder $a_k b_l$, $k, l \in \mathbb{N}_0$, enthält, gegen ab. Das Cauchy-Produkt ist eine solche Reihe, oder genauer eine Teilfolge der Partialsummen einer solchen Reihe. □

3.53 Satz: *Für die Exponentialfunktion gilt*

$$\exp(z + w) = \exp(z) \cdot \exp(w)$$

für alle $z \in \mathbb{C}$.

3.54. Bemerkung: Dies ist die fundamentale Identität der Exponentialfunktion. Außerdem gilt

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

und wir werden später zeigen, dass

$$\exp(1) = e = \lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n$$

gilt. Wir schreiben daher auch

$$e^z := \exp(z).$$

Für $n \in \mathbb{Z}$ folgt, dass diese Definition mit der alten Definition von e^n übereinstimmt.

Kapitel 4

Stetigkeit

Wir erweitern nun den Konvergenzbegriff auf Funktionen und werden dadurch in die Lage versetzt, den Begriff „Stetigkeit“ exakt zu fassen. Dieser Begriff wird in der Schule meist anschaulich anhand der fehlenden Sprünge im Funktionsgraphen beschrieben.

4.1 Grenzwert von Funktionen

Der Grenzwert einer Funktion bei Annäherung an einen Punkt kann mit Hilfe von Folgen oder mit dem ϵ-δ-Kriterium beschrieben werden.

Wenn M der Definitionsbereich der Funktion ist, so wollen wir allerdings nur Grenzwerte in Punkten berechnen, die in M oder am Rand von M liegen.

4.1. Definition: Ein Punkt a liegt im **Abschluss** einer Menge $M \subseteq \mathbb{R}$, wenn es eine Folge in M gibt, die gegen a konvergiert. Die Menge aller Punkt im Abschluss heißt die **abgeschlossene Hülle** von M. Sie wird mit $\overline{M}$ bezeichnet. M heißt **abgeschlossen**, wenn $M = \overline{M}$ gilt.

4.2. Bemerkung: Für uns sind zunächst hauptsächlich Intervalle interessant. Dort gilt zum Beispiel

$$\overline{]a,b[} = [a,b].$$

Die Randpunkte a, b liegen also im Abschluss des offenen Intervalls.

4.3. Definition: (1) Sei

$$f : M \to \mathbb{R}$$

eine Funktion und a im Abschluss von M. Dann heißt b der **Grenzwert** von f bei Annäherung an a, in Zeichen

$$b = \lim_{x \to a} f(x),$$

wenn es zu jedem $\epsilon > 0$ ein $\delta > 0$ gibt so dass gilt

$$|a - x| < \delta \ \Rightarrow \ |b - f(x)| < \epsilon$$

für alle $x \in M$.

(2) Die Funktion f heißt **stetig** in $a \in M$, wenn

$$\lim_{x \to a} f(x) = f(a)$$

gilt. Sie heißt stetig in M, wenn sie in allen Punkten $a \in M$ stetig ist.

4.4. Beispiel: Die Funktion $f(x) = x^2$ ist in allen $a \in \mathbb{R}$ stetig. Denn es gilt

$$|f(a+h) - f(a)| = |(a+h)^2 - a^2| = |h| \cdot |2a + h| \leq |h|(2|a| + |h|)$$

Sei nun $\epsilon > 0$ vorgegeben. Dann wählen wir $\delta > 0$ so klein, dass

$$\delta(2|a| + \delta) < \epsilon$$

gilt. Damit folgt

$$|h| < \delta \;\Rightarrow\; |f(a+h) - f(a)| < \epsilon.$$

Das ist äquivalent zu

$$\lim_{x \to a} x^2 = a^2,$$

und damit der Stetigkeit von x^2 in a.

4.5 Satz: *Sei $f : M \to \mathbb{R}$ und a im Abschluss von M. Genau dann gilt*

$$\lim_{x \to a} f(x) = b,$$

wenn für jede Folge a_n, $n \in \mathbb{N}$, in M gilt

$$\lim_{n \to \infty} a_n = a \;\Rightarrow\; \lim_{n \to \infty} f(a_n) = b.$$

Also ist f genau dann in $a \in M$ stetig, wenn

$$\lim_{n \to \infty} a_n = a \;\Rightarrow\; \lim_{n \to \infty} f(a_n) = f(a)$$

gilt.

4.6. Bemerkung: Wir können daher mit Hilfe unserer Kenntnisse über konvergente Folgen Aussagen über Grenzwerte von Funktionen gewinnen. Aus Beispiel 3.11 folgt nun etwa sofort, dass die Funktion

$$f(x) = \sqrt{x}$$

stetig ist.

4.7. Bemerkung: Sei $f : M \to \mathbb{R}$ eine Funktion und $a \in M$. Dann definieren wir in nahe liegender Weise

$$\lim_{x \to a, x \neq a} f(x).$$

f ist genau dann stetig, wenn dieser Grenzwert gleich $f(a)$ ist.

4.8 Satz: *Man kann auch beim Grenzwert von Funktionen „+“, „−“ und „·“ mit dem Limes vertauschen. Also*

$$\lim_{x \to a} (f(x) \pm g(x)) = \lim_{x \to a} f(x) \pm \lim_{x \to a} g(x)$$
$$\lim_{x \to a} (f(x) \cdot g(x)) = \lim_{x \to a} f(x) \cdot \lim_{x \to a} g(x)$$

Der Grenzwert auf der linken Seite existiert, falls die Grenzwerte auf der rechten Seite existieren. Falls $g(x) \neq 0$ für alle $x \in M$ ist, und auch der Grenzwert bei Annäherung an a nicht 0 ist, so gilt

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \frac{\lim\limits_{x \to a} f(x)}{\lim\limits_{x \to a} g(x)}$$

4.9. Bemerkung: Wenn

$$\lim_{x \to a} g(x) \neq 0$$

ist, so gibt es offenbar eine ϵ-Umgebung von x, wo $g(x) \neq 0$ ist. Da es beim Grenzübergang im Wesentlichen nur auf diese Umgebung ankommt, so gilt dann also die Aussage des Satzes über den Quotienten von Funktionen.

4.10. Beispiel: (1) Seien $a < b < c$ und

$$\begin{aligned} f &:]a, b[\to \mathbb{R} \\ g &:]b, c[\to \mathbb{R} \end{aligned}$$

stetig. Dann gibt es genau dann eine stetige **Fortsetzung** von f und g auf das Intervall $]a, c[$, wenn

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

ist. Dabei handelt es sich eigentlich um **einseitige Grenzwerte**, was man auch durch

$$\lim_{x \uparrow a} f(x) = \lim_{x \downarrow a} g(x)$$

andeutet. Auf diese Weise kann man **stückweise** gegebene stetige Funktionen zu einer stetigen Funktion zusammensetzen, wenn die Übergänge die richtigen Grenzwerte aufweisen.

(2) Ebenso ist jede Funktion $f : M \to \mathbb{R}$ in $a \in M$ stetig, wenn

$$\lim_{x \uparrow a} f(x) = \lim_{x \downarrow a} f(x)$$

gilt.

(3) Wenn eine Funktion in a nicht definiert ist, so kann man sie stetig nach a im Abschluss von M eindeutig **fortsetzen**, wenn

$$\lim_{x \to a} f(x)$$

existiert. Falls a nicht im Abschluss von M liegt, so ist diese Fortsetzung immer möglich, aber nicht eindeutig.

4.11. Definition: (1) Ganz analog zu den Folgen definieren wir **uneigentliche** Grenzwerte von Funktionen. Es gilt also

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

genau dann, wenn es zu jedem $c > 0$ ein $\epsilon > 0$ gibt mit

$$|x - a| < \epsilon \;\Rightarrow\; f(x) > c.$$

(2) In nahe liegender Weise definieren wir Grenzwerte für $x \to \infty$, wenn es eine Folge im Definitionsbereich gibt, die gegen ∞ konvergiert. Die Aussage

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

bedeutet also, dass es zu jedem $c > 0$ ein $d > 0$ gibt mit

$$x \geq d \;\Rightarrow\; f(x) \geq c.$$

4.2 Rationale Funktionen

Polynome sind die einfachsten Funktionen. Zur Auswertung muss man lediglich addieren und multiplizieren. Bei rationalen Funktion kommt eine Division dazu.

4.12. Definition: (1) Ein **Polynom** ist eine Funktion der Form

$$p(x) = a_0 + a_1 x + \ldots + a_n x^n = \sum_{k=0}^{n} a_k x^k.$$

Wenn $a_n \neq 0$ ist, so heißt n der Grad von p und wird mit $\deg(p)$ bezeichnet.

(2) Eine **rationale Funktion** ist der Quotient zweier Polynome

$$r(x) = \frac{p(x)}{q(x)}.$$

und ist definiert in allen Punkten, in denen $q(x) \neq 0$ ist.

4.13 Satz: *Polynome sind überall stetig. Ein Polynom p hat höchstens $deg(p)$* ***Nullstellen****. Man kann die Nullstellen aus p ausdividieren und erhält ein Polynom $\tilde{p}$ ohne Nullstellen, also*

$$p(x) = (x - x_1) \cdot \ldots \cdot (x - x_k) p(x)$$

wobei $x_1, \ldots, x_k$ die Nullstellen von p sind, und

$$deg(\tilde{p}) = deg(p) - k.$$

Dabei können Nullstellen mehrfach vorkommen, also mit höherer ***Vielfachheit****.*

4.14. Bemerkung: Rationale Funktion sind stetig in allen Punkten, in denen sie definiert sind.

4.15 Satz: (1) *Sei p ein Polynom mit höchstem Koeffizienten $a_n > 0$ und Grad n. Dann gilt*

$$\lim_{x \to \infty} p(x) = \infty,$$

und

$$\lim_{x \to -\infty} p(x) = \begin{cases} \infty & \text{falls } n \text{ gerade ist,} \\ -\infty & \text{falls } n \text{ ungerade ist.} \end{cases}$$

(2) *Sei $r = p/q$ eine rationale Funktion, $p(a) > 0$ und a eine Nullstelle der Vielfachheit k von q. Dann gilt*

$$\lim_{x \downarrow a} r(x) = \infty$$

und

$$\lim_{x \uparrow a} r(x) = \begin{cases} \infty & \text{falls } k \text{ gerade ist,} \\ -\infty & \text{falls } k \text{ ungerade ist.} \end{cases}$$

Wenn auch der Zähler $p(x)$ eine Nullstelle in a hat, so muss man zunächst die rationale Funktion ***kürzen****.*

4.3 Der Zwischenwertsatz

Der folgende Satz ist eines der wesentlichen Hilfsmittel in der Analysis.

4.16 Satz: *Sei $f : [a, b] \to \mathbb{R}$ stetig $c \in \mathbb{R}$ und*

$$f(a) \leq c \leq f(b).$$

Dann existiert ein $\xi \in [a, b]$ mit

$$f(\xi) = c.$$

Dieser Satz heißt ***Zwischenwertsatz****. Es folgt, dass das stetige Bild eines Intervalls ein Intervall ist.*

Beweis: Der Beweis kann konstruktiv durch eine Intervallschachtelung zur Berechnung von ξ geführt werden. Ausgehend von

$$[a_0, b_0] = [a, b]$$

konstruiert man Intervalle $[a_n, b_n]$ mit den Eigenschaften

$$\begin{aligned} f(a_n) \leq c, \quad f(b_n) &\geq c, \\ [a_{n+1}, b_{n+1}] &\subseteq [a_n, b_n], \\ b_{n+1} - a_{n+1} &= \frac{b_n - a_n}{2}. \end{aligned}$$

Nach dem Satz 3.8 über die Intervallschachtelungen gibt es einen gemeinsamen Grenzwert ξ von a_n und b_n. Da f stetig ist, folgt

$$\lim_{n \to \infty} f(a_n) = \lim_{n \to \infty} f(b_n) = f(\xi),$$

und daraus folgt

$$\begin{aligned} f(\xi) &= \lim_{n \to \infty} f(a_n) \leq c, \\ f(\xi) &= \lim_{n \to \infty} f(b_n) \geq c, \end{aligned}$$

also $f(\xi) = c$. □

4.17 Satz: *Sei I ein Intervall in $\mathbb{R}$, und $f : I \to \mathbb{R}$ streng monoton (wachsend oder fallend) und stetig. Dann gibt es eine* ***Umkehrfunktion***

$$f^{-1} : f(I) \to I,$$

die ebenfalls stetig ist. Wenn I ein offenes Intervall ist, so ist $f(I)$ ebenfalls ein offenes Intervall, und wenn I ein abgeschlossenes, beschränktes Intervall ist, so gilt das auch für $f(I)$.

Beweis: Da eine streng mononote Funktion injektiv ist, existiert f^{-1} auf $f(I)$. Sei $\epsilon > 0$. Falls $f(x) = y$ ist und x im Innern von I, so wählen wir eine δ-Umgebung von y in $]f(x - \epsilon), f(x + \epsilon)[$. Falls x ein Randpunkt von I ist, zum Beispiel der rechte Randpunkt, so wählen wir etwa $\delta = (f(x + \epsilon) - f(x))/2$. $\square$

4.18. Bemerkung: Das Bild des abgeschlossenen Intervalls $[a, \infty[$ unter einer stetigen Funktion kann ein halboffenes Intervall sein.

4.19. Beispiel: Die Funktion

$$f(x) = \frac{x^2 - 1}{x^2 + 1}$$

ist in $[0, \infty[$ stetig und streng monoton wachsend. Es gilt

$$f([0, \infty[) = [-1, 1[$$

Die Umkehrfunktion berechnet sich zu

$$f^{-1}(y) = \sqrt{\frac{1 + y}{1 - y}}.$$

für alle $-1 \leq y < 1$.

4.20. Beispiel: (1) Die Funktionen $f(x) = x^n$ sind auf $[0, \infty[$ monoton wachsend und stetig. Außerdem gilt $f(0) = 0$, sowie

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

Es folgt

$$f[0, \infty[= [0, \infty[.$$

Also gibt es eine Umkehrfunktion, die wir als n-te **Wurzel** bezeichnen. Wir schreiben

$$\sqrt[n]{y} = y^{\frac{1}{n}} := f^{-1}(y).$$

(2) Die Funktion

$$f(x) = x^3 - x$$

ist auf $[1, \infty[$ streng monoton wachsend. Aber die Umkehrfunktion

$$f^{-1} : [0, \infty[\to [1, \infty[$$

ist nicht so leicht zu berechnen (Formel von Cardano).

4.4 Extrema von Funktionen

Für die Praxis ist oft wichtig zu wissen, wo eine Funktion maximal wird. Man rechnet solche Maxima mit Ableitungen aus. Aber ihre Existenz wollen wir schon hier beweisen.

4.21. Definition: Sei $f : M \to \mathbb{R}$ eine Funktion ($M \subseteq \mathbb{R}$). Dann heißt $a \in M$ **Maximalpunkt** von f, wenn

$$f(a) \geq f(x) \qquad \text{für alle } x \in M$$

gilt. Der Wert

$$f(a) = \max f(M)$$

wird als **Maximum** von f auf M bezeichnet. a heißt **lokaler** Maximalpunkt, wenn f Maximalpunkt auf einer ϵ-Umgebung von a

$$U_\epsilon(a) \cap M$$

ist. Analog definieren wir (lokale) Minima. Die Minima und Maxima zusammengenommen nennt man **Extrema** von f.

4.22 Satz: (1) *Sei M abgeschlossen und beschränkt, sowie $f : M \to \mathbb{R}$ stetig in M. Dann besitzt f einen Maximalpunkt und einen Minimalpunkt in M.*

(2) *Das stetige Bild eines abgeschlossenen, beschränkten Intervalls $[a, b]$ ist ein abgeschlossenes, beschränktes Intervall $[c, d]$.*

Beweis: Zum Beweis von (1) zeigen wir zunächst dass, $f(M)$ beschränkt ist. Denn sonst gibt es zum Beispiel eine Folge $(x_n)_{n\in\mathbb{N}}$ mit $f(x_n) \to \infty$. Diese Folge hat eine konvergente Teilfolge $(x_{k(n)})_{n\in\mathbb{N}}$, deren Grenzwert x in M liegen muss. Wegen der Stetigkeit von f ergibt sich ein Widerspruch. Sei nun

$$s = \sup M.$$

Dann existiert eine Folge $(x_n)_{n\in\mathbb{N}}$ mit

$$\lim_{n\to\infty} f(x_n) = s.$$

Diese Folge hat eine konvergente Teilfolge $(x_{k(n)})_{n\in\mathbb{N}}$, die gegen $x \in M$ konvergiere. Es folgt wegen der Stetigkeit von f

$$f(x) = \lim_{n\to\infty} f(x_{k(n)}) = s.$$

Zum Beweis von (2) beachten wir lediglich, dass das stetige Bild eines Intervalls nach Satz 4.16 ein Intervall ist. □

4.23. Bemerkung: Wenn M diese Bedingungen nicht erfüllt ist, so braucht kein Extremalpunkt zu existieren, selbst wenn f stetig ist. Beispielsweise hat die rationale Funktion

$$f(x) = \frac{x^2 - 1}{x^2 + 1}$$

zwar den Minimalpunkt 0, aber keinen Maximalpunkt, obwohl $f(\mathbb{R})$ sogar beschränkt ist. Es gilt nämlich

$$f(\mathbb{R}) = [-1, 1[.$$

4.5 Stetigkeit in metrischen Räumen

Wir erinnern an die Definition einer Metrik aus 1.31.

Die Definition der Stetigkeit und des Grenzwertes von Funktionen lässt sich unmittelbar auf metrische Räume, und insbesondere auf $\mathbb{C}$ übertragen.

4.24. Definition: Seien A, B metrische Räume, deren Metriken wir beide mit d bezeichnen, sowie $a \in A$. Dann ist $f : A \to B$ in $a \in A$ stetig genau dann, wenn für jedes $\epsilon > 0$ ein $\delta > 0$ existiert mit

$$d(x, a) < \delta \;\Rightarrow\; d(f(x), f(a)) < \epsilon.$$

Der Abschluss $\overline{M}$ einer Teilmenge $M \subseteq A$ ist die Menge aller möglichen Grenzwerte von Folgen aus M. Wenn $a \in \overline{M}$ ist, so definieren wir $b \in B$ als Grenzwert

$$b = \lim_{x \to a} f(x)$$

genau dann, wenn die Funktion f stetig auf $M \cup \{a\}$ mit $f(a) = b$ erweiterbar ist.

4.25. Bemerkung: Die Konvergenz kann äquivalent durch Folgen definiert werden. Dazu muss für jede Folge $(x_n)_{n\in\mathbb{N}}$ gelten

$$x_n \to a \;\Rightarrow\; f(x_n) \to b.$$

4.26. Beispiel: Sei $A \subseteq \mathbb{C}$. Dann ist $f : A \to \mathbb{C}$ genau dann stetig, wenn die Funktionen

$$\mathrm{Re}\,(f), \mathrm{Im}\,(f) : A \to \mathbb{R}$$

stetig sind. Man beachte allerdings, dass $z_n \to z$ in $\mathbb{C}$ bedeutet, dass sich die komplexen Zahlen „aus allen Richtungen“ an z annähern können.

4.27. Bemerkung: Auch in $\mathbb{C}$ nimmt jede stetige Funktion $f : \mathbb{C} \to \mathbb{R}$ auf einer abgeschlossenen und beschränkten Menge ihre Extrema an. Wir werden darauf aber erst später eingehen.

Kapitel 5

Die Ableitung

Wir sind nun in der Lage, das von Newton und Leibniz entdeckte Kalkül der Analysis, nämlich die Ableitung, exakt zu fassen. Die Rechenregeln der Differenzialrechung und ihrer Umkehrung, der Intergralrechung, sind es, die die Analysis so mächtig machen.

5.1 Differenzierbarkeit

Anschaulich ist die Ableitung einer reellen Funktion in einem Punkt a die Steigung der Tangenten am Funktionsgraph im Punkt $(a, f(a))$. Wir fassen diesen Gedanken in der folgenden Definition exakt.

5.1. Definition: (1) Sei $f : I \to \mathbb{R}$ eine Funktion, I ein offenes Intervall. Dann heißt der Grenzwert

$$f'(x) := \lim_{t \to x, t \neq x} \frac{f(t) - f(x)}{t - x}$$

für $x \in I$ die **Ableitung** von f in x, falls er existiert. In diesem Fall heißt f **differenzierbar** in x. Wir schreiben auch

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

Dies ist insbesondere dann praktisch, wenn für $f(x)$ ein Rechenausdruck in x gesetzt wird.

(2) Falls f überall in I differenzierbar ist, so heißt f differenzierbare Funktion. Falls dann die Ableitungsfunktion f' stetig ist, so heißt f **stetig differenzierbar**.

(3) Falls f' wieder differenzierbar ist, so schreibt man für die Ableitung f'', bzw.

$$f''(x) = \frac{d^2}{dx^2} f(x).$$

f heißt in diesem Fall zweimal differenzierbar. Analog definieren wir n-mal (stetig) differenzierbare Funktionen, und schreiben für die Ableitung $f^{(n)}$.

5.2 Satz: *Jede differenzierbare Funktion ist stetig.*

5.3. Bemerkung: Die Umkehrung ist falsch, wie schon die Funktion $|x|$ zeigt.

5.4 Satz: *Seien $f, g : I \to \mathbb{R}$ differenzierbare Funktionen, I ein offenes Intervall. Dann ist auch $f \pm g$ und $f \cdot g$ differenzierbar und es gilt*

$$(f \pm g)' = f' \pm g',$$
$$(f \cdot g)' = f'g + fg'.$$

Falls $g(x) \neq 0$ ist für alle $x \in (a, b)$, so ist auch f/g differenzierbar und es gilt

$$\left(\frac{f}{g}\right)' = \frac{gf' - fg'}{g^2}.$$

Dabei bedeuten die Notationen $f+g$ etc., dass die Funktionen punktweise addiert werden.

5.5 Satz: *Es gilt*

$$\frac{d}{dx}x^n = nx^{n-1}.$$

für alle $n \in \mathbb{Z}$.

5.6. Bemerkung: Damit sind alle Polynome differenzierbar und alle rationalen Funktionen, wo sie definiert sind. Für Polynome gilt

$$\frac{d}{dx}(a_0 + a_1x + \ldots + a_nx^n) = a_1 + 2a_2x + \ldots + na_nx^{n-1}.$$

5.7. Definition: In nahe liegender Weise definieren wir einseitige Ableitungen für Funktionen $f : [a, b] \to \mathbb{R}$ in a und b, als

$$\lim_{x \downarrow a, x \neq a} \frac{f(x) - f(a)}{x - a}, \qquad \lim_{x \uparrow b, x \neq b} \frac{f(x) - f(b)}{x - b}.$$

5.2 Lokale Extrema und Monotonie

Um den Zusammenhang zwischen Ableitung, Monotonie und lokalen Extrema aufzuklären, benötigen wir den wichtigen **Mittelwertsatz** der Differenzialrechnung.

5.8 Satz: *Sei $f : I \to \mathbb{R}$ differenzierbar, I ein offenes Intervall, und $x \in I$ eine lokale Extremalstelle von f. Dann gilt $f'(x) = 0$.*

5.9. Bemerkung: Die Umkehrung ist falsch, wie die Funktion x^3 zeigt.

5.10. Beispiel: (1) Wir betrachten die Funktion $f(x) = x^3 - x$ auf $[-1, 1]$. Aufgrund des Satzes über die Extrema muss f eine Maximalstelle und eine

Minimalstelle in $[-1,1]$ haben. Um diese Stelle zu finden, suchen wir zunächst lokale Extrema in $]-1,1[$ durch die Bedingung $f'(x)=0$. Wir finden die Punkte

$$x_{1,2} = \pm\sqrt{\frac{1}{3}}.$$

Schließlich vergleichen wir die Werte

$$f(-1), f(x_1), f(x_2), f(1)$$

und finden auf diese Weise die Extrema in $[-1,1]$. Es gilt

$$f([-1,1]) = [f\left(-\frac{1}{\sqrt{3}}\right), f\left(\frac{1}{\sqrt{3}}\right)].$$

(2) Das gleiche kann man auch auf ganz $\mathbb{R}$ machen. Die Funktion

$$f(x) = \frac{x^2-1}{x^2+1}$$

hat bei $\pm\infty$ den Grenzwert 1 und die Ableitung hat nur einen kritischen Punkt $x=-1$, und $f(-1)=-1$. Wir schließen daraus

$$f(\mathbb{R}) = [-1,1[.$$

Denn f muss Extrema auf jedem Intervall $[-c,c]$ annehmen, wobei das Minimum immer in -1 liegen muss, und das Maximum in den Rändern.

5.11 Satz: *Seien $f,g:[a,b]\to\mathbb{R}$ stetige Funktionen, die in $]a,b[$ differenzierbar sind. Sei außerdem $g(a)\neq g(b)$ und $g'(x)\neq 0$ für alle $x\in\,]a,b[$. Dann gibt es ein $\xi\in\,]a,b[$ mit*

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

Beweis: Wir definieren die Funktion

$$h(x) = f(x) - f(b) - \frac{f(a)-f(b)}{g(a)-g(b)}\cdot(g(x)-g(b))$$

h ist ebenfalls in $]a,b[$ differenzierbar und in $[a,b]$ stetig. Außerdem gilt $h(a)=h(b)=0$. Also existiert ein Maximalpunkt $\xi\in\,]a,b[$. Es folgt $h'(\xi)=0$, und daraus die Behauptung. □

5.12. Bemerkung: Dieser Satz wird oft als **2. Mittelwertsatz** bezeichnet. Der **1. Mittelwertsatz** ist der Sonderfall $g(x)=x$, also

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

Der Sonderfall $f(a)=f(b)=0$ wird als **Satz von Rolle** bezeichnet. Er besagt, dass zwischen je zwei Nullstellen eine Nullstelle der Ableitung liegen muss.

5.13. Bemerkung: Die Bedingung $g'(x)\neq 0$ für alle $x\in\,]a,b[$ ist notwendig. Dazu betrachten wir

$$f(x)=x^2,\quad g(x)=x^3,\quad a=-1,\quad b=1.$$

In der Tat existiert kein Punkt $\xi \in]a, b[$, in dem $f'(\xi)/g'(\xi)$ definiert ist, und für den der Mittelwertsatz gültig wäre.

5.14. Bemerkung: Es folgt aus dem Mittelwertsatz

$$|f(b) - f(a)| \leq \left(\sup_{\xi \in]a,b[} |f'(\xi)| \right) \cdot |a - b|$$

Die Gleichheit gilt hier genau dann, wenn $f'(\xi)$ konstant ist.

5.15 Satz: (1) *Sei $f : I \to \mathbb{R}$ differenzierbar, I ein offenes Intervall. Genau dann ist $f'(x) \geq 0$ für alle $x \in I$, wenn f in I monoton wachsend ist. Eine analoge Bedingung gilt für monoton fallende Funktionen.*

(2) *Falls $f'(x) > 0$ für alle $x \in I$ ist, so ist f streng monoton wachsend auf I.*

5.16. Beispiel: (1) Eine Parabel ist ein reelles Polynom vom Grad 2

$$p(x) = ax^2 + bx + c$$

mit $a \neq 0$. Die Parabel hat für $a > 0$ ein absolutes Minimum auf $\mathbb{R}$ in

$$x_0 = -\frac{b}{2a}.$$

Für $a > 0$ ist sie auf $]-\infty, x_0]$ streng monoton fallend, und auf $[x_0, \infty[$ streng monoton steigend. Für $a < 0$ ist x_0 ein Maximum, und die Monotonie ist genau umgekehrt.

(2) Wir betrachten wieder $f(x) = x^3 - x$ und setzen das obige Beispiel fort. Da f' nur zwei Nullstellen in $\mathbb{R}$ hat und stetig ist, muss es in den verbleibenden drei Intervallen streng monoton sein. Wegen der Grenzwerte

$$\lim_{x \to -\infty} f(x) = -\infty, \quad \lim_{x \to \infty} f(x) = \infty,$$

und der Funktionswerte in x_1 und x_2 erhalten wir daher die **Monotonie-Intervalle**

$$]-\infty, x_1], \quad [x_1, x_2], \quad [x_2, \infty[,$$

wo f abwechselnd monoton steigt, fällt und wieder steigt. Wir wissen daher, dass x_1 ein lokales Maximum, und x_2 ein lokales Minimum ist. Beide sind aber keine absoluten Extremalstellen.

5.17 Satz: *Sei $f : I \to \mathbb{R}$ differenzierbar, I ein offenes Intervall, sowie in $x \in I$ zweimal differenzierbar. Es gelte*

$$f'(x) = 0, \quad f''(x) < 0.$$

Dann hat f in x ein lokales Maximum. Wenn statt dessen $f''(x) > 0$ gilt, so liegt ein lokales Minimum vor.

Beweis: Es gilt wegen $f'(x) = 0$

$$0 < f''(x) = \lim_{t \to x} \frac{f'(t)}{t - x}$$

Daraus folgt, dass ein $\epsilon > 0$ existiert, so dass f' in $]x-\epsilon, x[$ streng monoton fällt und in $]x, x+\epsilon[$ streng monoton wächst. □

5.18. Bemerkung: Dieser Satz ist bei der Suche nach *globalen* Extrema nicht sehr hilfreich, und selbst für lokale Extrema ist der Vergleich der Funktionswerte meist einfacher. Das Resultat ist eher theoretischer Natur. Stattdessen sollte man Monotonie-Intervalle und die Werte oder Grenzwerte in den Randpunkten verwenden.

5.19. Bemerkung: Ein lokales Minimum einer Funktion $f : \mathbb{R} \to \mathbb{R}$ ist allerdings ein absolutes Minimum, wenn f' nur eine Nullstelle hat.

5.3 Die Kettenregel

Wir betrachten hier die Ableitung der **Hintereinanderausführung** zweier Funktionen.

5.20 Satz: *Seien*

$$f : I \to J$$
$$g : J \to \mathbb{R}$$

differenzierbare Funktionen, wobei I und J offene Intervalle seien. Dann ist auch $g \circ f$ differenzierbar und es gilt

$$\frac{d}{dx} g(f(x)) = g'(f(x)) \cdot f'(x).$$

Diese Formel wird auch als ***Nachdifferenzieren*** *bezeichnet.*

5.21. Beispiel: Anstatt $p(x) = (x^2 + x + 1)^2$ auszumultiplizieren und dann zu differenzieren, erhalten mit $f(x) = x^2 + x + 1$ und $g(t) = t^2$

$$h'(x) = \frac{d}{dx} g(f(x) = 2(x^2 + x + 1)(2x + 1).$$

5.22 Satz: *Sei $f : I \to \mathbb{R}$ differenzierbar, I ein offenes Intervall, und*

$$f'(x) > 0 \qquad \text{für alle } x \in I$$

Dann ist f streng monoton wachsend auf I und hat eine Umkehrfunktion

$$f^{-1} : f(I) \to I,$$

die ebenfalls differenzierbar ist. Für die Ableitung gilt

$$f^{-1\prime}(y) = \frac{1}{f'(x)}$$

wobei $f(x) = y$ sei.

5.23. Beispiel: Die Funktion $g(x) = \sqrt{x}$ ist die Umkehrung von $f(x) = x^2$ auf $]0, \infty[$. Also ist sie differenzierbar und es gilt

$$\frac{d}{dy} \sqrt{y} = \frac{1}{2x} = \frac{1}{2\sqrt{y}}$$

wobei $y = f(x) = x^2$ ist.

5.4 Exponentialfunktion und Logarithmus

Diese beiden Funktionen sind zentrale Hilfsmittel in der Analysis. Ziel dieses Abschnittes ist es, Exponenten der Form a^x für $a > 0$ definieren zu können. Diese Exponenten sollen die Gleichung

$$a^{x+y} = a^x \cdot a^y$$

erfüllen, wie dies auch die Funktionen a^n, $n \in \mathbb{N}$, tun. Außerdem soll

$$a^1 = 1$$

sein. Die Umkehrfunktion $\log_a$ erfüllt dann automatisch die entsprechende Gleichung

$$\log_a(xy) = \log_a(x) + \log_a(y),$$

sowie

$$\log_a(a) = 1.$$

Unsere Vorgehensweise startet mit der Umkehrfunktion.

5.24 Satz: (1) *Für die Exponentialfunktion aus 3.50 gilt die Funktionalgleichung*

$$\exp(x+y) = \exp(x) \cdot \exp(y) \qquad \text{für alle } x, y \in \mathbb{R}$$

Also insbesondere

$$\exp(-x) = \frac{1}{\exp(x)} \qquad \text{für alle } x \in \mathbb{R}.$$

(2)

$$\exp(x) > 0 \qquad \text{für alle } x \in \mathbb{R}$$

(3) *Die Exponentialfunktion ist differenzierbar und*

$$\exp'(x) = \exp(x) \qquad \text{für alle } x \in \mathbb{R}$$

(4) exp *ist streng monoton wachsend.*

(5) *Es gilt*

$$\exp(x) \geq 1 + x \qquad \text{für alle } x \in \mathbb{R}.$$

Also insbesondere

$$\lim_{x \to \infty} \exp(x) = \infty.$$

sowie

$$\lim_{x \to -\infty} \exp(x) = 0.$$

Beweis: (1) Wurde bereits in 3.53 gezeigt. (2) folgt wegen

$$\exp(2x) = \exp(x)^2 \geq 0$$

und $\exp(x) \neq 0$ für alle $x \in \mathbb{R}$. Zum Beweis von (3) beachten wir

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

Eine einfache Abschätzung zeigt

$$\lim_{h\to 0}\sum_{k=0}^{\infty}\frac{h^k}{(k+1)!} = 1.$$

Daraus folgt (3). (4) folgt aus dem Satz über monotone Funktionen. (5) ist eine Kurvendiskussion von $f(x) = \exp(x) - (1+x)$. □

5.25. Bemerkung: Die Exponentialfunktion ist die einzige auf $\mathbb{R}$ stetige Funktion mit den Eigenschaften

$$\exp(0) = 1, \quad \exp(x+y) = \exp(x)\exp(y)$$

für alle $x, y \in \mathbb{R}$. Zum Beweis zeigt man, dass exp auf $\mathbb{Q}$ eindeutig definiert ist. Sie ist auch die einzige Funktion mit den Eigenschaften

$$f(0) = 1, \quad f'(x) = f(x)$$

für alle $x \in \mathbb{R}$. Zum Beweis betrachten wir die Funktion

$$h_y(x) = \frac{f(x+y)}{f(x)},$$

und zeigen durch Differenzieren, dass diese Funktion konstant sein muss. Es folgt die Funktionalgleichung für f, und damit $f = \exp$.

5.26. Definition: Die Umkehrfunktion von $\exp : \mathbb{R} \to]0, \infty[$ ist

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

Sie heißt **Logarithmus naturalis**.

5.27 Satz: (1) $\ln(x)$ *ist für* $x > 0$ *definiert, streng monoton wachsend, und es gilt*

$$\lim_{x\downarrow 0}\ln(x) = -\infty, \quad \lim_{x\to\infty}\ln(x) = \infty.$$

(2)

$$\ln'(x) = \frac{1}{x} \qquad \textit{für alle } x > 0.$$

(3) *Es gilt die Funktionalgleichung*

$$\ln(xy) = \ln(x) + \ln(y) \qquad \textit{für alle } x, y > 0,$$

insbesondere

$$\ln\left(\frac{1}{x}\right) = -\ln(x) \qquad \textit{für alle } x \in \mathbb{R}.$$

Beweis: (1) ist eine Folgerung aus der Definition als Umkehrfunktion von exp. (2) folgt aus dem Satz über die Ableitung der Umkehrfunktion. (3) folgt aus der Funktionalgleichung für exp. □

5.28. Bemerkung: Es gilt

$$\ln(e) = 1, \quad \exp(1) = e.$$

Dies kann man aus den Eigenschaften der Logarithmus-Funktion herleiten. Denn per Induktion zeigt man

$$\ln(x^n) = n \ln(x)$$

für alle $x > 0$ und $n \in \mathbb{N}$. Aus der Stetigkeit von ln folgt also

$$\ln(e) = \lim_{n\to\infty} n \ln(1 + 1/n) = \frac{\ln(1 + 1/n) - \ln(1)}{(1 + 1/n) - 1} = \ln'(1) = 1.$$

5.29. Definition: (1) Für $a > 0$, $x \in \mathbb{R}$ definieren wir

$$a^x = \exp(x \cdot \ln(a)).$$

Insbesondere also

$$e^x = \exp(x)$$

(2) Für $a > 0$, $a \neq 1$, $x > 0$ definieren wir

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

5.30 Satz: (1) *Die Funktion a^x ist $a > 1$ streng monoton wachsend, und für $0 < a < 1$ streng monoton fallend. Sie bildet in jedem Fall $\mathbb{R}$ auf $]0, \infty[$ bijektiv ab.*

(2) *Es gilt für $a > 0$*

$$a^{x+y} = a^x \cdot a^y$$

für alle $x, y \in \mathbb{R}$, sowie

$$a^0 = 1, \quad a^1 = a.$$

(3) *Für $a > 0$ und $x, y \in \mathbb{R}$ gilt*

$$a^{xy} = (a^x)^y.$$

(4) *Für $a > 0$ gilt*

$$\frac{d}{dx} a^x = \ln(a) \cdot a^x$$

für alle $x \in \mathbb{R}$

(5) *Die Funktion $\log_a(x)$ ist für $a > 1$ streng monoton wachsend, und für $a < 1$ streng monoton fallend. Sie bildet in jedem Fall $]0, \infty[$ auf $\mathbb{R}$ bijektiv ab.*

(6) *Für $a > 0$, $a \neq 1$, gilt*

$$\log_a(xy) = \log_a(x) + \log_a(y)$$

für alle $x, y > 0$ sowie

$$\log_a(1) = 0, \quad \log_a(a) = 1.$$

(7) *Für $a > 0$, $a \neq 1$, $x, y > 0$ gilt*

$$\log_a(x^y) = y \log_a(x).$$

(8) *Für $a > 0$, $a \neq 1$ gilt*

$$\frac{d}{dx} \log_a(x) = \frac{1}{x \ln(a)}$$

für alle $x > 0$.

5.31 Satz: *Für alle $a \in \mathbb{R}$ gilt*

$$\lim_{x \to \infty} \left(1 + \frac{a}{x}\right)^x = e^a.$$

Beweis: Analog zum Beweis von $\ln(e) = 1$. □

5.5 Der Satz von de l'Hospital

Dieser Satz mag auch schon aus der Schule bekannt sein. Er erlaubt die Berechnung von Grenzwerten durch Differenzieren.

5.32 Satz: (1) *Seien $f, g :]a, b[\to \mathbb{R}$ differenzierbare Funktionen. Es gelte*

$$\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0,$$

oder

$$\lim_{x \to a} f(x) = \pm\infty, \quad \lim_{x \to a} g(x) = \pm\infty.$$

Außerdem sei vorausgesetzt, dass $g'(x) \neq 0$ für alle $x \in]a, b[$ ist. Dann gilt

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)},$$

sofern der Grenzwert auf der rechten Seite existiert. Natürlich gilt derselbe Satz analog für b.

(2) *Der Satz gilt auch, wenn $a = -\infty$ oder $b = \infty$ ist.*

Beweis: Sei zunächst der Grenzwert von Zähler und Nenner gleich 0. Dann können wir f und g stetig auf a fortsetzen mit $f(a) = g(a) = 0$. Aus dem zweiten Mittelwertsatz folgt für ein $\xi \in \,]a, x[$

$$\left|\frac{f(x)}{g(x)}\right| = \left|\frac{f(x) - f(a)}{g(x) - g(a)}\right| = \left|\frac{f'(\xi)}{g'(\xi)}\right|.$$

Mit $x \to a$ folgt die Behauptung, da dann auch $\xi \to a$ gilt. Seien nun die Grenzwerte $\pm\infty$. Sei $a < y < x < b$. Es gilt dann für ein $\xi \in \,]y, x[$

$$\left|\frac{f(y)}{g(y)}\right| \cdot \left|\frac{1 - f(x)/f(y)}{1 - g(x)/g(y)}\right| = \left|\frac{f(y) - f(x)}{g(y) - g(x)}\right| = \left|\frac{f'(\xi)}{g'(\xi)}\right|$$

Die rechte Seite geht mit $x \to a$ gegen einen Grenwert η, unabhängig von der Wahl von y. Mit $y \to a$ geht aber der zweite Faktor links gegen 1. Es folgt die Behauptung.

(2) Dieser Fall kann durch $\tilde{f}(x) = f(1/x)$ und $\tilde{g}(x) = g(1/x)$ auf (1) zurück geführt werden. □

5.33. Bemerkung: Man achte stets genau darauf, ob die Voraussetzungen des Satzes auch wirklich erfüllt sind. Sonst rechnet man sehr leicht fehlerhaft.

5.34. Beispiel: Man berechnet

$$\lim_{x\to\infty} \frac{\ln(x)}{x} = \lim_{x\to\infty} \frac{1/x}{1} = 0.$$

Den Satz von de l'Hospital kann man deswegen anwenden, weil der Zähler und der Nenner den Grenzwert ∞ haben, und weil der Grenzwert des Quotienten der Ableitungen existiert.

5.6 Konvexe Funktionen

Die Entscheidung, ob eine Funktion konvex oder konkav ist, beschreibt nicht nur die Form einer Funktion, sondern sie erlaubt auch überraschende Abschätzungen der Funktion. Das Kriterium mit der zweiten Ableitung wird schon in der Schule behandelt.

5.35. Definition: Sei I ein Intervall in $\mathbb{R}$. Eine Funktion $f : I \to \mathbb{R}$ heißt **konvex** in I, wenn gilt

$$f(\lambda x + \mu y) \leq \lambda f(x) + \mu f(y)$$

für alle $x, y \in I$ und $\lambda, \mu \in \mathbb{R}$ mit

$$\lambda, \mu \geq 0, \quad \lambda + \mu = 1.$$

Im Fall „$\geq$" spricht man von einer **konkaven** Funktion. Wenn f links von einem Punkt x konvex, und rechts von einem Punkt konkav ist oder umgekehrt, so nennt man x einen **Wendepunkt** von x.

5.36. Bemerkung: Die Gerade

$$g(t) = f(x) + (t - x)\frac{f(y) - f(x)}{y - x}$$

ist die **Sekante** durch f in x und y. Die Tatsache, dass f konvex ist, besagt nun gerade $g \geq f$ in $[x, y]$ für $x < y$.

5.37 Satz: *Sei I ein Intervall und $f : I \to \mathbb{R}$ zweimal differenzierbar im Innern, am Rand stetig. Dann gilt*

$$f''(x) \geq 0 \qquad \textit{für alle } x \in I,$$

genau dann, wenn f konvex in I ist. Analog ist f genau dann konkav, wenn hier „$\leq$“ gilt.

Beweis: Sei $x, y \in I$, sowie $z \in I$ ein weiterer Punkt, $x < y$ und $z \in]x, y[$. Dann betrachtet man

$$h(t) = (f(t) - g(t))(z - x)(z - y) - (f(z) - g(z))(t - x)(t - y),$$

wobei g die Sekante durch f in x und y ist. Es gilt $h(x) = h(y) = h(z) = 0$. Nach dem Satz von Rolle, zweimal angewendet, gibt es ein $\xi \in]x, y[$ mit $h''(\xi) = 0$. Also

$$f(z) - g(z) = \frac{f''(\xi)}{2}(z - x)(z - y),$$

Es folgt die Behauptung. $\square$

5.38. Beispiel: Die Wurzelfunktion ist konkav auf $]0, \infty[$. Also

$$\sqrt{1 + \lambda h} \geq 1 + \lambda(\sqrt{1 + h} - 1)$$

für alle $\lambda \in [0, 1]$ und $1 + h \in]0, \infty[$. Denn die Sekante in 1 und $1 + h$ lautet

$$g(t) = 1 + (t - 1)\frac{\sqrt{t} - 1}{t - 1}.$$

Einsetzen von $t = 1 + \lambda h$ ergibt die obige Gleichung.

5.39. Bemerkung: Wenn $f'' > 0$ im Intervall ist, so gilt nach demselben Beweis

$$f(\lambda x + \mu y) < \lambda f(x) + \mu f(y)$$

für alle $\lambda, \mu > 0$ mit $\lambda + \mu = 1$, also $f < g$ in $]x, y[$. Solche Funktionen heißen (**strikt konvex**.

5.40. Bemerkung: Für die Tangente $T_x(z)$ an f im Punkt x beweist man analog

$$f(t) - T(t) = -\frac{f''(\xi)}{2}(t - x)^2,$$

Aus $f'' \geq 0$ in I folgt also $f \leq T$. Eine konvexe Funktion liegt also unterhalb ihrer Tangenten. Entsprechend liegt eine konkave Funktion oberhalb jeder Tangenten.

Kapitel 6

Integralrechnung

In diesem Kapitel behandeln wir das Riemann-Integral in einer Variablen. Schon aus der Schule ist bekannt, dass das Integral mit Hilfe der Umkehrung der Ableitung berechnet werden kann. Wir werden diesen Satz und diverse, daraus folgende Rechenregeln für Integrale hier herleiten.

6.1 Das Riemann-Integral

Die Idee des Riemann-Integrals ist es, die Fläche unter dem Graphen einer positiven Funktion durch Treppenfunktionen von oben und von unten anzunähern.

6.1. Definition: (1) Sei $[a, b]$ ein Intervall. Eine Funktion $T : [a, b] \to \mathbb{R}$ heißt **Treppenfunktion** auf $[a, b]$, wenn es eine **Unterteilung**

$$a = x_0 < x_1 < \ldots < x_n = b$$

des Intervalls gibt, so dass f auf $]x_{k-1}, x_k[$ für alle $k = 1, \ldots, n$ konstant ist. Wir setzen

$$\int_a^b T(t)\, dt := \sum_{i=1}^{n} T(\xi_k)(x_k - x_{k-1})$$

mit irgendwelchen Punkten

$$\xi_k \in \,]x_k, x_{k-1}[\qquad \text{für alle } k = 1, \ldots, n.$$

(2) Die Summe auf der rechten Seite heißt **Riemannsche Zwischensumme**, auch wenn T keine Treppenfunktion ist. In diesem Fall hängt die Summe natürlich von der Wahl der ξ_k und von der Unterteilung ab. Man nennt

$$\max_{1 \leq k \leq n} (x_{k+1} - x_k)$$

die **Feinheit** der Unterteilung.

(3) Für eine beschränkte Funktion $f : [a, b] \to \mathbb{R}$ und eine Treppenfunktion T auf $[a, b]$ mit $T \leq f$ heißt $\int T$ eine **Untersumme** von f. Das **Unterintegral** von f ist das Supremum aller Untersummen. Also

$$\underline{\int_a^b} f(t)\,dt := \sup\{\int_a^b T(t)\,dt : T \leq f \text{ ist Treppenfunktion}\}.$$

Analog definieren wir das **Oberintegral**

$$\overline{\int_a^b} f(t)\,dt := \inf\{\int_a^b T(t)\,dt : T \geq f \text{ ist Treppenfunktion}\}.$$

(4) Eine beschränkte Funktion $f : [a, b] \to \mathbb{R}$ heißt **Riemann-integrierbar**, wenn ihr Oberintegral gleich dem Unterintegral ist. In diesem Fall setzen wir

$$\int_a^b f(t)\,dt := \underline{\int_a^b} f(t)\,dt = \overline{\int_a^b} f(t)\,dt.$$

6.2. Bemerkung: Wir müssen uns auf *beschränkte* Funktionen zurück ziehen, weil sonst kein Ober- oder kein Unterintegral existieren würde.

6.3. Bemerkung: Es gilt für jede beschränkte Funktion $f : [a, b] \to \mathbb{R}$

$$\underline{\int_a^b} f(t)\,dt \leq \overline{\int_a^b} f(t)\,dt.$$

Um das zu zeigen, muss man die Monotonie des Integrals für Treppenfunktionen zeigen, also

$$\int_a^b T_1(t)\,dt \leq \int_a^b T_2(t)\,dt$$

für Treppenfunktionen $T_1 \leq T_2$.

6.4. Beispiel: (1) Jede Treppenfunktion ist Riemann-integrierbar, insbesondere die konstante Funktion $f = c$, und

$$\int_a^b c\,dt = c(b - a).$$

(2) Die Funktion $f(x) = x$ ist auf $[a, b]$ Riemann-integrierbar. Dazu konstruiert man Untersummen und Obersummen mit äquidistanter Unterteilung

$$a = x_0 < x_1 = a + h < \ldots < x_n = b = a + nh$$

mit $h = (b - a)/n$. Sei U_n eine maximale Untersumme mit dieser Unterteilung, so gilt

$$\begin{aligned} U_n &= h(a + (a + h) + \ldots + (a + (n - 1)h)) \\ &= hna + \frac{1}{2}h^2(n - 1)n \\ &\to (b - a)a + \frac{1}{2}(b - a)^2 \\ &= \frac{b^2}{2} - \frac{a^2}{2}, \end{aligned}$$

wobei die Konvergenz für $n \to \infty$ gemeint ist. Dieses Ergebnis ist auch aus der Schule bekannt. Die minimale Obersumme unterscheidet sich nur um $nh^2 = (b-a)^2/n \to 0$.

6.5 Satz: *Jede monotone Funktion $f : [a,b] \to \mathbb{R}$ ist Riemann-integrierbar.*

6.6 Satz: *Eine Funktion $f : [a,b] \to \mathbb{R}$ ist genau dann Riemann-integrierbar, wenn jede Folge von Riemannschen Zwischensummen, deren Feinheit der Unterteilung gegen 0 geht, konvergiert.*

Beweis: Sei f Riemann-integrierbar. Dann gibt es zu jedem $\epsilon > 0$ Treppenfunktionen $T_1 \leq f \leq T_2$ mit

$$\int T_1 \leq \int f \leq \int T_2 \leq \int T_1 + \epsilon.$$

Sei S_n eine Folge von Riemannschen Zwischensummen, deren Feinheiten der Unterteilung gegen 0 geht. Dann gibt es ein $N \in \mathbb{N}$ mit

$$\int T_1 - \epsilon \leq S_n \leq \int T_2 + \epsilon \qquad \text{für alle } n \geq N.$$

Es folgt die Konvergenz von S_n gegen $\int f$.

Es gelte umgekehrt die Konvergenz der Riemannschen Zwischensummen. Dann muss f beschränkt sein. Sei $\epsilon > 0$. Sei S_n eine Folge Riemannscher Zwischensummen mit Feinheiten, die gegen 0 konvergieren. Dann kann man Riemannsche Zwischensummen U_n, O_n von Treppenfunktionen über denselben Unterteilungen konstruieren, für die gilt

$$\underline{\int} f - \epsilon \leq U_n \leq S_n \leq O_n \leq \overline{\int} f + \epsilon.$$

Es folgt, dass f Riemann-integrierbar ist. □

6.7 Satz: (1) *Das Integral ist mit „+“ und „−“ vertauschbar. D.h., wenn $f, g : [a,b] \to \mathbb{R}$ Riemann-integrierbar sind, so sind es auch $f \pm g$ und*

$$\int_a^b (f(t) \pm g(t))\, dt = \int_a^b f(t)\, dt \pm \int_a^b g(t)\, dt.$$

(2) *Eine Konstante kann aus dem Integral herausgezogen werden. D.h., wenn $f : [a,b] \to \mathbb{R}$ Riemann-integrierbar ist, so ist es auch λf für jedes $\lambda \in \mathbb{R}$ und es gilt*

$$\int_a^b \lambda f(t)\, dt = \lambda \int_a^b f(t)\, dt.$$

Dies wird, zusammen mit Punkt (1), als ***Linearität*** *des Integrals bezeichnet.*

(3) *Wenn $f, g : [a,b] \to \mathbb{R}$ Riemann-integrierbar sind und*

$$f(t) \leq g(t) \qquad \text{für alle } t \in [a,b]$$

gilt („$f \le g$“), so gilt

$$\int_a^b f(t)\,dt \le \int_a^b g(t)\,dt.$$

Dies wird als ***Monotonie*** *des Integrals bezeichnet.*

6.8. Definition: Wir setzen für $b < a$

$$\int_a^b f(t)\,dt = -\int_b^a f(t)\,dt,$$

wenn dieses Integral existiert.

6.9 Satz: *Eine Riemann-integrierbare Funktion $f : I \to \mathbb{R}$ ist auch auf jedem Teilintervall von I integrierbar. Es gilt für alle $a, b, c \in I$ in beliebiger Reihenfolge*

$$\int_a^b f(t)\,dt + \int_b^c f(t)\,dt = \int_a^c f(t)\,dt.$$

Wenn die Integrale links existieren, so existiert auch das Integral rechts.

6.2 Der Hauptsatz

Der Hauptsatz der Differenzial- und Integralrechnung besagt, dass die Integralrechnung die Umkehrung der Differenzialrechung ist. Er gilt für stetige Funktionen. Wir müssen daher nebenbei nachweisen, dass alle stetigen Funktionen integrierbar sind.

6.10. Definition: Sei $M \subseteq \mathbb{R}$. Eine Funktion $f : M \to \mathbb{R}$ heißt **gleichmäßig** stetig auf M, wenn es zu jedem $\epsilon > 0$ ein $\delta > 0$ gibt, so dass gilt

$$|x - y| < \delta \;\Rightarrow\; |f(x) - f(y)| < \epsilon$$

für alle $x, y \in M$.

6.11. Bemerkung: Dies ist fast wie die Definition von Stetigkeit in x. Nur hängt dort δ von x und ϵ ab. Hier soll es nur von ϵ abhängen. Man kann die gleichmäßige Stetigkeit auch durch Folgen ausdrücken. f ist genau dann gleichmäßig stetig, wenn für je zwei Folgen x_n und y_n, $n \in \mathbb{N}$, aus M gilt

$$x_n - y_n \to 0 \;\Rightarrow\; f(x_n) - f(y_n) \to 0.$$

6.12. Beispiel: (1) Die Funktion $f(x) = 1/x$, definiert auf $]0, 1[$ ist auf diesem Intervall nicht gleichmäßig stetig.

(2) Für differenzierbare Funktionen $f : [a, b] \to \mathbb{R}$ kann man δ in Abhängigkeit von ϵ berechnen. Denn es gilt

$$|f(x) - f(y)| = |f'(\xi)|\,|x - y| \le |x - y| \max_{\xi \in]a,b[} |f'(\xi)|.$$

6.13 Satz: *Jede stetige Funktion $f : [a, b] \to \mathbb{R}$ ist gleichmäßig stetig.*

Beweis: Angenommen nicht. Dann gibt es zu einem $\epsilon > 0$ kein solches δ, also Folgen x_n, y_n in $[a, b]$ mit

$$|x_n - y_n| < \frac{1}{n}, \quad |f(x_n) - f(y_n)| \geq \epsilon.$$

Wenn $x \in [a, b]$ ein Häufungspunkt von x_n ist, so ensteht sofort ein Widerspruch zur Stetigkeit von f in x. □

6.14 Satz: (1) *Jede stetige Funktion $f : [a, b] \to \mathbb{R}$ ist Riemann-integrierbar.*

(2) *Sei $f : [a, b] \to \mathbb{R}$ stetig und*

$$F(x) = \int_a^x f(t)\, dt.$$

Dann ist F nach x differenzierbar (an den Rändern einseitig) und es gilt

$$F'(x) = f(x).$$

Man nennt F eine ***Stammfunktion*** *von f.*

(3) *Je zwei Stammfunktionen einer stetige Funktion $f :]a, b[\to \mathbb{R}$ unterscheiden sich nur durch eine Konstante, $F_1 = F_2 + c$.*

(4) *Seien $f, F : [a, b] \to \mathbb{R}$ stetig, und $F'(x) = f(x)$ für alle $x \in]a, b[$. Dann gilt*

$$\int_a^b f(t)\, dt = F(b) - F(a) =: [F(x)]_{x=a}^{x=b}.$$

Dieser Satz heißt ***Hauptsatz*** *der Differenzial- und Integralrechnung.*

Beweis: (1) folgt aus der gleichmäßigen Stetigkeit. Für (2) verwendet man, dass

$$\min f[x, x + h] \leq \frac{1}{h} \int_x^{x+h} f(t)\, dt \leq \max f[x, x + h]$$

gilt, bzw. die entsprechende Ungleichung für $[x-h, x]$. Für (3) beweist man mit Hilfe des Mittelwertsatzes, dass eine differenzierbare Funktion, deren Ableitung auf $]a, b[$ gleich 0 ist, überall 0 ist. Für (4) zeigt man, dass die Gleichung für die in (2) konstruierte Stammfunktion gilt. Mit Hilfe von (3) zeigt man, dass sie dann für jede Stammfunktion gilt. □

6.15. Bemerkung: Man schreibt, etwas salopp,

$$F(x) := \int f(x)\, dx + c$$

für eine Stammfunktion von f. Das Integral ohne die Grenzen beschreibt also eine Stammfunktion von f und wird **unbestimmtes Integral** genannt.

6.16. Beispiel: (1) Wir erhalten

$$\int x^n\, dx = \frac{1}{n+1} x^{n+1} + c$$

für alle $n \in \mathbb{N}_0$ und damit für Polynome

$$\int (a_0 + a_1 x + \ldots + a_n x^n)\, dx = C + a_0 x + a_1 \frac{x^2}{2} + \ldots + a_n \frac{x^{n+1}}{n+1}.$$

(2) Für die Funktionen x^{-n}, $n \in \mathbb{N}$, erhalten wir das Integral bis auf den Fall $n = -1$ wie oben. Also

$$\int x^n\, dx = \begin{cases} \frac{x^{n+1}}{n+1} + c, & n \neq -1, \\ \ln|x| + c, & n = -1. \end{cases}$$

6.3 Partielle Integration und Substitution

6.17 Satz: *Seien $f, g : [a,b] \to \mathbb{R}$ stetige Funktionen, die in $]a,b[$ differenzierbar seien, und deren Ableitungen stetig auf $[a,b]$ fortgesetzt werden können. Dann gilt*

$$\int_a^b f'(t)g(t)\, dt = [f(x)g(x)]_{x=a}^{x=b} - \int_a^b f(t)g'(t)\, dt.$$

*Man bezeichnet diese Art zu rechnen als **partielle Integration**, da man ja nur die Stammfunktion des einen Faktors f' ermitteln muss.*

6.18. Bemerkung: Wir können dies als unbestimmtes Integral in der Form

$$\int f'(x)g(x)\, dt = f(x)g(x) - \int f(x)g'(x)\, dx$$

schreiben.

6.19. Beispiel: (1) Es gilt

$$\int xe^x\, dx = \int x \exp'(x)\, dx = xe^x - \int e^x\, dx = (x-1)e^x + c.$$

In der Tat ist $(x-1)e^x$ eine Stammfunktion von xe^x.

(2) Es gilt

$$\int \ln(x)\, dx = x\ln(x) - \int 1\, dx = x(\ln(x) - 1) + c.$$

6.20 Satz: *Sei $f : [a,b] \to \mathbb{R}$ stetig. Sei außerdem $g : [c,d] \to [a,b]$ eine stetige, in $]c,d[$ differenzierbare Funktion, deren Ableitung auf $[c,d]$ stetig fortsetzbar sei. Dann gilt*

$$\int_{g(c)}^{g(d)} f(t)\, dt = \int_c^d f(g(s))g'(s)\, ds.$$

*Dies bezeichnet man als **Substitution**, da praktisch $t = g(s)$ substituiert wird.*

Beweis: Wenn F gemäß dem Hauptsatz eine Stammfunktion von f ist, so ist

$$\frac{d}{ds} F(g(s)) = f(g(s))g'(s).$$

Aus dem Hauptsatz folgt

$$\int_{g(c)}^{g(d)} f(t)\,dt = [F(t)]_{t=g(c)}^{t=g(d)} = [F(g(s))]_{s=c}^{s=d} = \int_c^d f(g(s))g'(s)\,ds.$$

$\square$

6.21. Bemerkung: Man beachte, dass man auch die Grenzen substituieren muss. Wenn $s = c$ ist, so ist in der Tat $t = g(c)$.

6.22. Bemerkung: Mit unbestimmten Integralen geschrieben, lautet der Satz

$$\int f(t)\,dt = \int f(g(s))g'(s)\,ds + c,$$

wobei $t = g(s)$ zu substituieren ist.

6.23. Beispiel: (1) Meist wendet man die Substitution von rechts nach links an. Mit $g(s) = s^2 + 1 = t$ gilt beispielsweise

$$\int_0^1 \frac{s}{\sqrt{s^2+1}}\,ds = \int_1^2 \frac{1}{2\sqrt{t}}\,dt = \left[\sqrt{t}\right]_{t=1}^{t=2} = \sqrt{2} - 1.$$

Als unbestimmtes Integral

$$\int \frac{s}{\sqrt{s^2+1}}\,ds = \int \frac{1}{2\sqrt{t}}\,dt + c = \sqrt{t} + c = \sqrt{s^2+1} + c.$$

(2) Mit stetigem $f : \mathbb{R} \to \mathbb{R}$ und $g(s) = s + d$ folgt

$$\int_a^b f(t)\,dt = \int_{a-d}^{b-d} f(s+d)\,ds$$

Bei dieser Formel kann man eigentlich auf Stetigkeit verzichten. Sie folgt lediglich durch Betrachtung von Riemannschen Zwischensummen.

6.24. Bemerkung: Bei der Substitution muss man genau aufpassen, dass f in allen Punkten $g(s)$ definiert ist. Sonst ist das Ergebnis falsch.

6.4 Uneigentliche Integrale

Das Riemann-Integral ist in der bisherigen Form an beschränkte Funktionen auf beschränkten, abgeschlossenen Intervallen gebunden. Um mehr damit anfangen zu können erweitern wir es nun.

6.25. Definition: (1) Sei $f :]a, b] \to \mathbb{R}$ auf jedem abgeschlossenen Teilintervall Riemann-integrierbar. Dann definieren wir

$$\int_a^b f(t)\,dt := \lim_{\epsilon \to 0} \int_{a+\epsilon}^b f(t)\,dt,$$

sofern dieser Grenzwert existiert. Analog für Intervalle $[a,b[$.

(2) Sei $f : [a,\infty[\to \mathbb{R}$ auf jedem abgeschlossenen Teilintervall Riemann-integrierbar. Dann definieren wir

$$\int_a^\infty f(t)\,dt := \lim_{c\to\infty} \int_a^c f(t)\,dt,$$

sofern dieser Grenzwert existiert. Analog für $]-\infty,b]$.

Diese Integrale werden als **uneigentliche** Integrale bezeichnet.

6.26. Beispiel: (1)

$$\int_1^\infty \frac{1}{x^2}\,dx = \lim_{c\to\infty} \left[\frac{-1}{x}\right]_{x=1}^{x=c} = 1.$$

(2)

$$\int_0^1 \frac{1}{x^2}\,dx = \lim_{\epsilon\to 0} \left[\frac{-1}{x}\right]_{x=\epsilon}^{x=1} = \infty.$$

Dieses Integral existiert also nur im Sinne der **uneigentlichen** Grenzwerte.

6.27. Bemerkung: Auf beidseitigen offenen Intervallen muss man die Grenzwerte auf jeder Seite *einzeln* nehmen. Also

$$\int_{-\infty}^\infty f(t)\,dt = \lim_{a\to-\infty} \lim_{b\to\infty} \int_a^b f(t)\,dt.$$

Nimmt man statt dessen

$$\lim_{a\to\infty} \int_{-a}^a f(t)\,dt,$$

so erhält man den **Hauptwert** des Integrals. Es kann sein, dass der Hauptwert existiert, die einzelnen Grenzwerte jedoch nicht. Wenn jedoch das uneigentliche Integral existiert, so ist es gleich dem Hauptwert. Ist die Funktion positiv, so kann man zeigen, dass aus der Existenz des Hauptwertes die Existenz des uneigentlichen Integrals folgt.

6.28 Satz: *Sei $f : [1,\infty[\to \mathbb{R}$ eine monoton fallende stetige, positive Funktion. Dann konvergiert die Reihe*

$$\sum_{k=1}^\infty f(k)$$

genau dann, wenn

$$\int_a^\infty f(t)\,dt < \infty$$

ist.

6.29. Beispiel: (1) Es gilt

$$\int_1^\infty \frac{1}{x}\,dx = \infty.$$

Daher divergiert die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k}.$$

(1) Es gilt

$$\int_2^{\infty} \frac{1}{x \log(x)^2} \, dx = \left[\frac{-1}{\log x} \right]_{x=2}^{x=\infty} = \frac{1}{\log 2} < \infty.$$

Daher konvergiert die Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k \log(k)^2}.$$

6.5 Allgemeine Funktionen

Für allgemeine, nicht unbedingt stetige Funktionen kann man einige Existenzaussagen machen. Meist ist die Funktion allerdings stückweise stetig, so dass man das Integral aus den Stücken berechnen kann.

6.30 Satz: *Sei $f : [a,b] \to \mathbb{R}$ Riemann-integrierbar. Dann sind auch die Funktionen*

$$f_+ = \max\{f, 0\}, \quad f_- = -\min\{f, 0\}, \quad |f|, \quad f^2$$

Riemann-integrierbar. Falls $g : [a,b] \to \mathbb{R}$ ebenfalls Riemann-integrierbar ist, so sind auch die Funktionen

$$fg, \quad \max\{f, g\}, \quad \min\{f, g\}$$

Riemann-integrierbar. Es gilt

$$\left| \int_a^b f(t) \, dt \right| \leq \int_a^b |f(t)| \, dt.$$

Beweis: Die Aussage für f_+ folgt aus der Betrachtung von Unter- und Obersummen. Bildet man die entsprechenden positiven Anteile von Treppenfunktionen $T_1 \leq f \leq T_2$, so wird die Differenz der Integrale kleiner. Es gilt $f_- = f_+ - f$, $|f| = f_+ + f_-$.

Die Aussage für f^2 folgt aus

$$T_2^2 - T_1^2 = (T_1 - T_2)(T_1 + T_2).$$

Es gilt

$$fg = \frac{1}{2}((f+g)^2 - f^2 - g^2),$$

sowie

$$\max\{f, g\} = \frac{1}{2}(f + g + |f - g|), \quad \min\{f, g\} = \frac{1}{2}(f + g - |f - g|).$$

Die letzte Aussage folgt mit der Dreiecksungleichung aus $|f| = f_+ + f_-$, $f = f_+ - f_-$. □

6.31 Satz: *Sei $f : [a,b] \to \mathbb{R}$ stetig, und $g : [a,b] \to \mathbb{R}$ Riemann-integrierbar mit $g \geq 0$. Dann existiert ein $\xi \in \,]a,b[$ mit*

$$f(\xi) \int_a^b g(t)\,dt = \int_a^b f(t)g(t)\,dt.$$

Speziell existiert ein $\xi \in \,]a,b[$ mit

$$f(\xi) \cdot (b-a) = \int_a^b f(t)\,dt.$$

*Dieser Satz heißt **Mittelwertsatz** der Integralrechnung.*

Beweis: Man hat wegen $g \geq 0$ und der Monotonie des Integrals

$$\left(\min_{x \in [a,b]} f(x) \right) \int_a^b g(t)\,dt \leq \int_a^b f(t)g(t)\,dt \leq \left(\max_{x \in [a,b]} f(x) \right) \int_a^b g(t)\,dt.$$

Im Fall

$$\int_a^b g(t)\,dt = 0$$

gilt wegen der Beschränktheit von f

$$|\int_a^b f(t)g(t)\,dt| \leq \int_a^b |f(t)g(t)|\,dt \leq c \int_a^b g(t)\,dt = 0.$$

Andernfalls folgt die Aussage aus dem Zwischenwertsatz für f. □

Kapitel 7

Potenzreihen

Zur numerische Berechnung von Integralen und Funktionen wie $\ln(x)$ und e^x kann man am schnellsten mit Hilfe von Potenzreihen erreichen.

7.1 Die Taylorreihe

Die Taylorreihe einer Funktion ist eine Reihe, deren Partialsummen im Funktionswert und in den Ableitungen bis zu einer gewissen Ordnung in einem gegebenen Punkt mit der Funktion übereinstimmen. Auf diese Weise hofft man, die Funktion möglichst genau zu approximieren.

7.1. Definition: Sei $f : I \to \mathbb{R}$ in $a \in \mathbb{R}$ unendlich oft differenzierbar, wobei I ein Intervall ist und x im Innern von I liegt. Dann heißt die formale Reihe

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(x-a)^k = f(a) + f'(a)(x-a) + \frac{f''(a)}{2}(x-a)^2 + \frac{f'''(a)}{6}(x-a)^3 + \ldots$$

die **Taylorreihe** von f im Punkt a. Die Partialsumme bis n heißt **Taylorentwicklung** n-ten Grades im Punkt a, und wir bezeichnen mit

$$R_n(x) := f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(a)}{k!}(x-a)^k$$

das **Restglied** der Taylorentwicklung. Man beachte, dass wir hier $(x-a)^0 = 1$ setzen.

7.2. Bemerkung: Man rechnet nach, dass für die Taylorentwicklung n-ten Grades p_n gilt

$$f(a) = p_n(a), \quad f'(a) = p_n'(a), \quad \ldots, \quad f^{(n)}(a) = p_n^{(n)}(a).$$

p_n ist das einzige Polynom n-ten Grades mit dieser Eigenschaft. Es kommt im Folgenden darauf an, das Restglied abzuschätzen.

7.3. Beispiel: (1) Die Taylorreihe von e^x im Punkt $a = 0$ ist die **Exponentialreihe**

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k$$

Wir wissen, dass die Reihe für alle $x \in \mathbb{R}$ konvergiert.

(2) Mit Hilfe der geometrischen Summe weist man nach, dass

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k \qquad \text{für alle } |x| < 1$$

ist. In der Tat ist dies die Taylorreihe der Funktion in $a = 0$. Sie wird **geometrische Reihe** genannt. Die Taylorreihe konvergiert aber nicht mehr überall, wo die Funktion definiert ist, sondern nur in $]-1, 1[$.

7.4 Satz: *Sei $f : I \to \mathbb{R}$ $n{+}1$-mal stetig differenzierbar, I ein offenes Intervall mit $a, x \in I$. Dann gilt die **Integraldarstellung** des Restglieds*

$$R_n(x) = \int_a^x \frac{(x-t)^n}{n!} f^{(n+1)}(t)\, dt$$

bei der Entwicklung von f in a. Außerdem gibt es ein ξ zwischen a und x, so dass gilt

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1},$$

*was man als **Lagrange-Form** des Restglieds bezeichnet.*

7.5. Beispiel: Der Fall $n = 0$ ist uns schon bekannt. Denn es gilt nach dem Mittelwertsatz

$$f(x) = f(a) + f'(\xi)(x-a).$$

für ein ξ zwischen a und x.

7.6. Bemerkung: (1) Wenn f $n+1$-mal stetig differenzierbar ist, so nimmt $f^{(n+1)}$ in einer Umgebung U von a ein Maximum an. Wir erhalten

$$|R_n(x)| \le C_n |x-a|^{n+1}$$

für alle $x \in U$. Dies schreibt man auch in der Form

$$|R_n(x)| = O(|x-a|^{n+1}) \qquad (x \to a)$$

(2) Es gilt für n-mal stetig differenzierbare Abbildungen f

$$\frac{f^{(n)}(a)}{n!}(x-a)^n + R_n(x) = R_{n-1}(x) = \frac{f^{(n)}(\xi)}{n!}(x-a)^n$$

mit einem ξ zwischen a und x. Also

$$R_n(x) = (f^{(n)}(\xi) - f^{(n)}(a))\frac{(x-a)^n}{n!}.$$

Wir erhalten also

$$\lim_{x \to a} \frac{|R_n(x)|}{|x-a|^n} = 0,$$

was man auch in der Form

$$|R_n(x)| = o(|x-a|^n) \qquad (x \to a)$$

schreibt.

7.7. Beispiel: Die Taylorreihe von $f(x) = (1+x)^\alpha$, $\alpha \in \mathbb{R}$, bei der Entwicklung um 0 lautet

$$\sum_{k=0}^{\infty} \binom{\alpha}{k} x^k,$$

mit

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1) \cdot \ldots \cdot (\alpha-(k-1))}{k!}.$$

für $k \in \mathbb{N}$ und

$$\binom{\alpha}{0} = 1.$$

Man beachte, dass diese Reihe für $\alpha \in \mathbb{N}_0$ abbricht und dann mit der Binomialreihe übereinstimmt. Aus dem Quotientenkriterium folgt leicht, dass die Reihe für alle $|x| \leq 1$ konvergiert Die Reihe konvergiert für alle $|x| < 1$ in der Tat gegen $f(x)$, wie man mit Hilfe des Integralrestgliedes beweist.

7.2 Potenzreihen

Wir untersuchen nun Reihen vom Typ der Taylorreihen etwas allgemeiner. Insbesondere werden wir herausfinden, dass sich diese Reihen gliedweise differenzieren und integrieren lassen, wenn sie konvergieren.

7.8. Definition: Eine formale Reihe der Form

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

mit $a, x \in \mathbb{R}$ oder $a, x \in \mathbb{C}$ heißt **Potenzreihe** um a. Dabei ist $(a_k)_{k \in \mathbb{N}}$ eine Folge in $\mathbb{R}$ oder $\mathbb{C}$.

7.9. Bemerkung: Die Taylorreihen sind Potenzreihen. Also sind die Ergebnisse dieses Abschnittes auf reelle Taylorreihen anwendbar.

7.10. Definition: Für eine Folge von a_n, $n \in \mathbb{N}$, definieren wir den **Limes superior** als das Supremum aller Häufungspunkte der Folge, wobei wir die Häufungspunkte $\pm\infty$ zulassen. Analog definieren wir den **Limes inferior** als das Infimum aller Häufungspunkte. Wir schreiben

$$\limsup_{n \to \infty} a_n, \qquad \liminf_{n \to \infty} a_n.$$

7.11. Bemerkung: Jede reelle Folge hat einen Häufungspunkt, wenn man $\pm\infty$ zulässt.

7.12 Satz: *Sei*

$$s = \limsup_{k\to\infty} \sqrt[k]{|a_k|}.$$

Dann konvergiert die Potenzreihe

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

für alle

$$|x-a| < r := \frac{1}{s}$$

Sie konvergiert nicht für alle $|x-a| > r$. r heißt der ***Konvergenzradius*** *der Potenzreihe.*

7.13. Bemerkung: Es gilt für den Konvergenzradius also

$$r = \frac{1}{\limsup\limits_{k\to\infty} \sqrt[k]{|a_k|}} = \liminf_{k\to\infty} \frac{1}{\sqrt[k]{|a_k|}},$$

wobei wir in diesem Fall $1/0 = \infty$ und $1/\infty = 0$ setzen. Diese für den Konvergenzradius heißt Formel von **Cauchy und Hadamard**.

7.14. Beispiel: (1) Die Reihen

$$\sum_{k=0}^{\infty} k^m x^k$$

haben für alle $m \in \mathbb{Z}$ den Konvergenzradius 1.

(2) Da die Exponentialreihe für alle $x \in \mathbb{R}$ konvergiert, muss ihr Konvergenzradius ∞ sein. Wir schließen daraus

$$\limsup_{k\to\infty} \sqrt[k]{\frac{1}{k!}} = 0.$$

Daraus folgt

$$\lim_{k\to\infty} \sqrt[k]{k!} = \infty.$$

7.15. Bemerkung: Für komplexe Potenzeihen gilt der Satz über den Konvergenzradius ganz analog. In diesem Fall konvergieren die Potenzreihen innerhalb eines Kreises mit Radius r um den Entwicklungspunkt.

7.3 Gleichmäßige Konvergenz

Ziel dieses Abschnittes ist es, Potenzreihen zu differenzieren und zu integrieren. Wie sich herausstellt, ist dies innerhalb des Konvergenzradius problemlos möglich.

7.16. Definition: Eine Folge $f_n : [a, b] \to \mathbb{R}$ von Funktionen heißt auf $[a, b]$ **gleichmäßig konvergent** gegen eine Funktion $f : [a, b] \to \mathbb{R}$, wenn die Folge

$$\|f - f_n\| := \sup_{x \in [a,b]} |f(x) - f_n(x)|$$

für $n \to \infty$ gegen 0 konvergiert.

7.17. Bemerkung: Die Größe $\|f - f_n\|$ bezeichnet den maximalen Abstand der Funktionen. Man nennt $\|r\|$ die **Supremums-Norm** vom r auf $[a, b]$. Wenn die Funktionen stetig sind, ist das Supremum natürlich ein Maximum. Äquivalent zur gleichmäßigen Konvergenz ist, dass es zu jedem $\epsilon > 0$ ein $N \in \mathbb{N}$ gibt mit

$$|f(x) - f_n(x)| < \epsilon \qquad \text{für alle } n \geq N,\ x \in [a, b].$$

7.18. Beispiel: Die Funktionen $f(x) = x^n$ konvergieren auf $[0, a]$ für alle $a < 1$ gleichmäßig gegen 0, aber nicht auf $[0, 1]$.

7.19 Satz: *Der gleichmäßige Grenzwert einer Folge von stetigen Funktionen ist stetig.*

7.20 Satz: *Eine Potenzreihe um a habe den Konvergenzradius $r > 0$, wobei wir $r = \infty$ zulassen, und es gelte $\rho < r$. Dann konvergiert die Potenzreihe auf $[a - \rho, a + \rho]$ gleichmäßig.*

7.21. Bemerkung: Aufgrund des Beweises sieht man, dass die Reihe auch auf $[a - r, a + r]$ gleichmäßig konvergiert, sofern sie in einem Randpunkt absolut konvergiert.

7.22 Satz: *Wenn die Folge der Riemann-integrierbaren Funktionen $f_n : [a, b] \to \mathbb{R}$ gleichmäßig gegen $f : [a, b] \to \mathbb{R}$ konvergiert, so folgt*

$$\lim_{n\to\infty} \int_a^b f_n(t)\, dt = \int_a^b f(t)\, dt.$$

Das heißt, Integration und Grenzübergang sind ***vertauschbar****.*

7.23 Satz: *Sei*

$$f(x) = \sum_{k=0}^{\infty} a_k (x - a)^k$$

eine Potenzreihe mit Konvergenzradius $r > 0$, wobei wir $r = \infty$ zulassen. Dann ist f für $|x - a| < r$ differenzierbar und es gilt

$$f'(x) = \sum_{k=1}^{\infty} k a_k (x - a)^{k-1}$$

sowie für die Stammfunktion F von f mit $F(a) = 0$

$$F(x) := \int_a^x f(t)\, dt = \sum_{k=0}^{\infty} \frac{a_k}{k+1} (x - a)^{k+1}.$$

*Diese Potenzreihen haben ebenfalls den Konvergenzradius r. Man kann also eine Potenzreihe **gliedweise** differenzieren und integrieren.*

Beweis: Die Formel für das Integral folgt aus der gleichmäßigen Konvergenz. Die Potenzreihe, die für f' angegeben ist, hat den gleichen Konvergenzradius und konvergiert gleichmäßig in $[a-\rho, a+\rho]$ für $\rho < r$. Durch Integrieren erkennt man, dass sie gegen f konvergiert. □

7.24. Beispiel: (1) Es gilt

$$\sum_{k=1}^{\infty} kx^k = x\frac{d}{dx}\sum_{k=1}^{n} x^k = x\frac{d}{dx}\frac{1}{1-x} = \frac{x}{(1-x)^2}$$

für $|x| < 1$.

(2) Es gilt

$$\sum_{k=1}^{\infty} \frac{x^k}{k} = \sum_{k=0}^{\infty} \int_0^x t^k \, dt = \int_0^x \left(\sum_{k=0}^{\infty} t^k \right) \, dt = \int_0^x \frac{1}{1-t} \, dt = \ln\left(\frac{1}{1-x}\right).$$

für $|x| < 1$. Die Reihe konvergiert aufgrund des Leibniz-Kriteriums auch für $x = -1$.

7.25 Satz: *Eine Potenzreihe*

$$f(x) = \sum_{k=0}^{\infty} a_k (x-a)^k$$

*habe den Konvergenzradius $r > 0$. Wenn die Reihe in $a + r$ konvergiert, so konvergiert sie gleichmäßig in $[a - \rho, a + r]$ für alle $\rho < r$. Dieser Satz heißt **Abelscher Grenzwertsatz**.*

7.26. Beispiel: Aufgrund dieses Satzes konvergiert die Reihe im vorigen Beispiel in $x = -1$ gegen

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

gilt. In $x = 1$ konvergiert die Reihe uneigentlich gegen ∞, und in der Tat gilt ja auch

$$\lim_{x \uparrow 1} \ln\left(\frac{1}{1-x}\right) = \infty.$$

7.27 Satz: *Die Potenzreihe*

$$f(x) = \sum_{k=0}^{\infty} a_k (x-a)^k$$

habe den Konvergenzradius $r > 0$. Dann stimmt sie mit der Taylorreihe von f überein. D.h. es gilt

$$a_k = \frac{f^{(k)}(a)}{k!}$$

für alle $k \in \mathbb{N}$.

7.4 Trigonometrische Funktionen

Da wir nun die Theorie der Potenzreihen zur Verfügung haben, können wir sehr elegant die Sinus- und die Kosinus-Funktion einführen, sowie die Zahl π definieren. Außerdem liefern wir noch den Beweis von Satz 5.24 nach.

7.28. Definition: Wir setzen

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

Diese Funktionen heißen **Sinus** und **Kosinus**.

7.29. Bemerkung: Es folgt aus der Reihendarstellung der Exponentialfunktion

$$\exp(\bar{z}) = \overline{\exp(z)}$$

und deswegen

$$|\exp(ix)| = 1$$

für alle $x \in \mathbb{R}$. Setze man $z = ix$ in die Reihe ein, so erhält man

$$\cos(x) = \operatorname{Re}(\exp(ix)), \quad \sin(x) = \operatorname{Im}(\exp(ix)).$$

Es ist allerdings immer noch nicht klar, ob dies geometrisch Sinn macht. Das können wir erst begründen, wenn wir die Bogenlänge des Kreisbogens von $(1, 0)$ nach $(\cos(x), \sin(x))$ berechnen können.

7.30. Bemerkung: Aufgrund der Reihendarstellung sieht man sofort

$$\sin(0) = 0, \quad \cos(0) = 1,$$

sowie

$$\sin(-x) = -\sin(x), \quad \cos(-x) = \cos(x)$$

für alle $x \in \mathbb{R}$.

7.31 Satz: (1) *Es gilt*

$$\sin'(x) = \cos(x), \quad \cos'(x) = -\sin(x)$$

für alle $x \in \mathbb{R}$. Zudem gelten die ***trigonometrischen Identitäten***

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

für alle $x, y \in \mathbb{R}$.

Beweis: Die Ableitungen folgen durch Differenzieren der Potenzreihen. Aufgrund der Fundamentalgleichung für die komplexe Exponentialfunktion gilt

$$1 = e^0 = e^{ix}e^{-ix} = (\cos(x) + i\sin(x))(\cos(x) - i\sin(x)) = \cos(x)^2 + \sin(x)^2.$$

Die anderen beiden Identitäten erhält man durch Berechnen von $\exp(i(x+y))$ □

7.32. Bemerkung: Man kann diese Identitäten auch allein aus der Funktionalgleichung und den Ableitungseigenschaften herleiten. Differenziert man $\sin^2 + \cos^2$, so wird die Ableitung 0. Die Funktion muss daher konstant sein. Einsetzen von $x = 0$ ergibt die erste Identität. Ebenso zeigt man die Identitäten

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

durch Differenzieren nach x für festes y. Die anderen beiden Identitäten folgen daraus mit der Cramerschen Regel.

7.33. Definition: Die Zahl π ist die kleinste positive Nullstelle von $\sin(x)$. D.h.

$$\pi = \inf\{x : \sin(x) = 0,\, x > 0\}.$$

7.34. Bemerkung: Wegen der Stetigkeit von sin ist $\sin(\pi) = 0$. Man beachte, dass für $|x| < 1$ die Sinusreihe eine alternierende Reihe ist. Es folgt

$$x - \frac{x^3}{6} \leq \sin(x) \leq x \qquad \text{für alle } 0 \leq x \leq 1.$$

Also ist $\pi > 1$. Man rechnet mit der Reihe nach, dass $\sin(3) > 0$ und $\sin(4) < 0$ ist. Es gilt in der Tat $\pi = 3.1415\ldots$

7.35 Satz: *Die Funktion* sin *ist in* $]0, \pi[$ *positiv, konvex, und sie besitzt genau ein absolutes Maximum in* $\pi/2$. *Außerdem gilt*

$$\sin\left(\frac{\pi}{2} - x\right) = \sin\left(\frac{\pi}{2} + x\right) = \cos(x)$$

für alle $x \in \mathbb{R}$. sin *ist also* ***symmetrisch*** *bezüglich* $\pi/2$. *Die Funktion* cos *ist in* $[0, \pi]$ *streng monoton fallend und es gilt*

$$\cos(0) = 1, \quad \cos(\pi) = -1.$$

cos *hat genau eine Nullstelle in* $\pi/2$.

Beweis: Man verwendet

$$\cos(2x) = \cos(x)^2 - \sin(x)^2 = 1 - 2\sin(x)^2 = 2\cos(x)^2 - 1.$$

Es folgt $\cos(\pi/2) = 0$ und $\sin(\pi/2) = 1$. □

7.36 Satz: *Es gilt für alle* $x \in \mathbb{R}$

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

sowie

$$\sin(x + \frac{\pi}{2}) = \cos(x).$$

Außerdem

$$\sin(-x) = -\sin(x), \quad \cos(-x) = \cos(x).$$

Also ist sin *eine **ungerade** und* cos *eine **gerade** Funktionen.*

7.37. Bemerkung: Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ mit

$$f(x + a) = f(a)$$

für alle $x \in \mathbb{R}$ nennt man a-periodisch. Sinus und Kosinus, sowie die daraus abgeleiteten trigonometrischen Funktionen, sind also 2π-periodische Funktionen.

7.38. Bemerkung: Es folgt auch, dass die Punkte

$$m\pi$$

für $m \in \mathbb{Z}$ die Nullstellen des Sinus sind. Die ungeraden Vielfachen von $\pi/2$, also die Punkte

$$x = (2m + 1)\frac{\pi}{2} \qquad \text{für alle } m \in \mathbb{Z},$$

sind die Nullstellen des Kosinus.

7.39 Satz: *Sei $x, y \in \mathbb{R}$ mit*

$$x^2 + y^2 = 1.$$

Dann gibt es genau ein $t \in [0, 2\pi[$ mit

$$x = \cos(t), \quad y = \sin(t).$$

Folglich lässt sich jeder Punkt $(x, y) \in \mathbb{R}^2$ $(x, y) \neq (0, 0)$ eindeutig als

$$(x, y) = (r\cos(t), r\sin(t))$$

mit $r > 0$ und $t \in [0, 2\pi[$ schreiben. Jedes $z \in \mathbb{C}$, $z \neq 0$ lässt sich eindeutig in der Form

$$z = re^{it}$$

mit $r > 0$ und $t \in [0, 2\pi[$ schreiben.

7.40. Bemerkung: Es gilt

$$e^{i\pi} = -1.$$

7.41. Definition: Wir definieren den **Tangens**

$$\tan(x) := \frac{\sin(x)}{\cos(x)}$$

in allen Punkten, in denen $\cos(x) \neq 0$ ist.

7.42. Bemerkung: Man berechnet

$$\tan'(x) = \frac{1}{\cos(x)^2}.$$

tan ist also zwischen je zwei Definitionslücken streng monoton wachsend. Außerdem gilt

$$\lim_{x \downarrow -\pi/2} \tan(x) = -\infty, \qquad \lim_{x \uparrow \pi/2} \tan(x) = \infty.$$

Daher hat tan eine Umkehrfunktion

$$\arctan : \mathbb{R} \to \left] -\frac{\pi}{2}, \frac{\pi}{2} \right[.$$

Es gilt

$$\arctan'(y) = \frac{1}{\tan'(x)} = \cos(x)^2 = \cos(\arctan(y)) = \frac{1}{1+y^2}$$

mit $y = \tan(x)$.

7.5 Partialbruchzerlegung

Dies ist eine Technik, mit der man Integrale von rationalen Funktionen berechnen kann. Wir werden uns auf Beispiele beschränken.

7.43. Beispiel: (1) Gesucht ist

$$\int \frac{1}{x^2 - 1} \, dx$$

auf den Intervallen $]-\infty, -1[$, $]-1, 1[$, $]1, \infty[$, wo diese Funktion wohldefiniert ist. Eine Partialbruchzerlegung ist nun eine Darstellung der Form

$$\frac{1}{x^2 - 1} = \frac{A}{x+1} + \frac{B}{x-1}.$$

Es folgt durch Ausmultiplizieren und Koeffizientenvergleich $A = 1/2$, $B = -1/2$. Also

$$\int \frac{1}{x^2 - 1} = \frac{1}{2} \ln(x+1) - \frac{1}{2} \ln(x-1) + c.$$

Natürlich kann die Konstante c in jedem Teilintervall anders gewählt werden, und man erhält dennoch eine Stammfunktion.

(2) Analog erhält man

$$\frac{x+1}{x(x^2+1)} = \frac{1-x}{x^2+1} + \frac{1}{x}.$$

Es folgt

$$\int \frac{1}{x(x^2+1)} \, dx = \arctan(x) - \frac{1}{2} \ln(x^2+1) + \ln(|x|) + c = \ln \sqrt{\frac{x+1}{x+1}} + c.$$

Allerdings funktioniert das im folgenden Beispiel nicht mehr. Man kommt auf andere Art zum Ziel.

$$\int \frac{1}{x(x^2+1)} \, dx = \int \left(\frac{1}{x} - \frac{x}{x^2+1} \right) dx = \ln(|x|) + \frac{1}{2} \ln(x^2+1) + c.$$

(3) Man berechnet

$$\int \frac{1}{x^2+x+1}\,dx$$
$$= \frac{2}{\sqrt{3}} \int \frac{1}{y^2+1}\,dx = \frac{2}{\sqrt{3}} \arctan(y) + c = \frac{2}{\sqrt{3}} \arctan\left(\frac{1}{\sqrt{3}}(2x+1)\right) + c$$

mit Hilfe der Substitution
$$x = \frac{1}{2}(\sqrt{3}y - 1).$$

Damit kann man auch kompliziertere Nenner behandeln. Die meisten dieser Integrale stehen allerdings in Formelsammlungen bereit.

7.6 Die Stirlingsche Formel

Große Fakultäten $n!$ sind ebenso wie große Binomialkoeffizienten nicht leicht zu berechnen. Die Stirlingsche Formel bietet eine Approximation, die für viele Zwecke genau genug ist. Zum Beweis der Formel benötigen wir eine Produktdarstellung von π von Wallis. Insgesamt wendet dieses Kapitel die bisher erlangten Kenntnisse trickreich an.

7.44. Definition: Wir sagen, dass ein Produkt

$$\prod_{k=1}^{\infty} b_k$$

konvergiert, wenn die Folge der Partialprodukte gegen einen Wert $b \neq 0$ konvergiert.

7.45 Satz: *Sei a_k, $k \in \mathbb{N}$, eine Folge von positiven reellen Zahlen oder eine Folge von negativen reellen Zahlen. Dann konvergiert das Produkt*

$$\prod_{k=1}^{\infty} (1 + a_k)$$

genau dann, wenn die Reihe

$$\sum_{k=1}^{\infty} a_k$$

konvergiert.

7.46 Satz: *Es gilt*

$$\frac{\pi}{2} = \prod_{k=1}^{\infty} \frac{4k^2}{4k^2-1}.$$

Diese Darstellung nennt man das ***Wallissche Produkt****.*

Beweis: Wir definieren für $n > 0$

$$c_n := \int_0^{\pi} \sin(x)^n\,dx.$$

Dann erfüllen diese Zahlen die Rekursionsformel

$$c_0 = \pi, \quad c_1 = 2, \quad c_n = \frac{n-1}{n} c_{n-2}$$

für $n \geq 2$. Folglich

$$c_{2n} = \pi \prod_{k=1}^{n} \frac{2k-1}{2k}, \quad c_{2n+1} = 2 \prod_{k=1}^{n} \frac{2k}{2k+1}.$$

Es gilt aufgrund der Rekursionsformel

$$\lim_{n\to\infty} \frac{c_{2n+2}}{c_{2n}} = 1.$$

Wegen der Monotonie der Folge c_n folgt

$$\lim_{n\to\infty} \frac{c_{2n+1}}{c_{2n}} = 1.$$

Man erhält nun das Wallissche Produkt. □

7.47 Satz: *Es gilt*

$$n! \sim \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$

Dabei bedeutet $a_n \sim b_n$, dass $a_n/b_n \to 1$ konvergiert. Man nennt die Folgen ***asymptotisch gleich****. Die Asymptotik für $n!$ heißt* ***Stirlingsche Formel****.*

Beweis: Sei g_k die lineare Funktion, die $\ln(x)$ in k und $k+1$ interpoliert. Mittels zweimaliger partieller Integration stellt man für $r_k(x) = \ln(x) - g_k(x)$ fest

$$\int_k^{k+1} (x-k)(k+1-x) r_k''(x)\, dx = \int_k^{k+1} r_k(x)\, dx.$$

Wegen der Konkavität von ln ist $r_k > 0$. Wegen $r_k''(x) = 1/x^2$ ist

$$0 < \int_k^{k+1} r_k(x)\, dx = \frac{-r_k''(\xi)}{12} \leq \frac{1}{12k^2}.$$

Also existiert

$$\begin{aligned}
\sum_{k=1}^{\infty} \int_k^{k+1} r_k(x)\, dx &= \lim_{n\to\infty} \left(\int_0^n \ln(x)\, dx - \sum_{k=1}^{n-1} g_k(x)\, dx \right) \\
&= \lim_{n\to\infty} \left(n\ln(n) - n + 1 - \sum_{k=1}^{n-1} \frac{1}{2}(\ln(k+1) + \ln(k)) \right) \\
&= \lim_{n\to\infty} \left(n\ln(n) - n + 1 - \ln(n!) + \frac{\ln(n)}{2} \right).
\end{aligned}$$

Folglich existiert auch der Grenzwert $c > 0$ von

$$c_n = \frac{n!}{\sqrt{n}} \left(\frac{e}{n}\right)^n.$$

Es gilt

$$c = \lim_{n\to\infty} \frac{c_n^2}{c_{2n}} = \lim_{n\to\infty} \frac{2^{2n}(n!)^2\sqrt{2}}{(2n)!\sqrt{n}} = \sqrt{2\pi}$$

mit Hilfe des Wallisschen Produktes. □

Kapitel 8

Der Euklidsche Raum

In diesem Kapitel werden Grenzwerte von Folgen und von Funktionen in den mehrdimensionalen Raum verallgemeinert. Dazu müssen wir allerdings einige topologische Begriffe einführen.

8.1 Konvergenz

Der $\mathbb{R}^n$ ist für uns einfach die Menge der n-Tupel reeller Zahlen, die wir als Spaltenvektoren

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

schreiben. Zu $x \in \mathbb{R}^n$ ist also automatisch $x_1, \ldots, x_n$ definiert. Mit der Addition und der Multiplikation im $\mathbb{R}^n$ beschäftigt sich die lineare Algebra. Gelegentlich werden wir auch den $\mathbb{C}^n$ erwähnen, der auf die gleiche Weise definiert ist.

Im Folgenden sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

8.1. Definition: Wie definieren für eine Vektor $x \in \mathbb{K}^n$ die **Euklidsche Norm**

$$\|x\| := \sqrt{\sum_{i=1}^{n} |x_i|^2}.$$

Zu zwei Vektoren $x, y \in \mathbb{K}^n$ definieren wir den **Euklidschen Abstand**

$$d(x, y) := \|x - y\|.$$

Wir nennen den $\mathbb{R}^n$ mit dieser Norm den **Euklidschen Raum** der Dimension n. Mit Hilfe des **Euklidschen Skalarprodukts**

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i$$

kann man auch schreiben

$$\|x\| = \sqrt{\langle x, x \rangle}$$

für $x \in \mathbb{R}^n$. Im Fall $\mathbb{K} = \mathbb{C}$ verwendet man das Skalarprodukt

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i \overline{y_i}$$

8.2 Satz: *Die Euklidsche Norm ist eine* ***Norm*** *auf dem $\mathbb{K}^n$. Das heißt, es gilt*

$$\|x\| \geq 0 \qquad \text{für alle } x \in \mathbb{K}^n,$$

die ***Dreiecksungleichung***

$$\|x + y\| \leq \|x\| + \|y\| \qquad \text{für alle } x, y \in \mathbb{K}^n,$$

sowie die Linearität

$$\|\lambda x\| = |\lambda| \|x\| \qquad \text{für alle } \lambda \in \mathbb{K},\ x \in \mathbb{K}^n,$$

und die Positivität

$$\|x\| = 0 \ \Leftrightarrow \ x = 0.$$

Außerdem gilt die ***Schwarzsche Ungleichung***

$$|\langle x, y \rangle| \leq \|x\| \cdot \|y\|$$

für alle $x, y \in \mathbb{K}^n$, bei der Gleichheit genau dann gilt, wenn x und y linear abhängig sind.

Beweis: Zum Beweis der Schwarzschen Ungleichung verwendet man

$$0 \leq \langle x + \lambda y, x + \lambda y \rangle$$

und setzt

$$\lambda = -\frac{\langle x, y \rangle}{\langle y, y \rangle}$$

im Fall $y \neq 0$. Es folgt durch einfache Rechnung mit Skalarprodukten die Dreiecksungleichung. □

8.3. Bemerkung: Auch der Raum $\mathbb{K}^n$ wird durch

$$\|z\| = \sqrt{\sum_{k=1}^{n} |z|^2}$$

zu einem normierten, komplexen Vektorraum. In diesem Fall gilt die Linearität für alle $\lambda \in \mathbb{K}$.

8.4 Satz: *Der Euklidscher Abstand ist eine* ***Metrik*** *auf dem $\mathbb{K}^n$ im Sinne von 1.31.*

8.5. Bemerkung: Wir haben Stetigkeit und Konvergenz schon in allgemeinen metrischen Räumen betrachtet. Demgemäß ist eine Folge von x_k, $k \in \mathbb{N}$, in $\mathbb{K}^n$ **konvergent** gegen $x \in \mathbb{K}^n$, wenn

$$\lim_{n \to \infty} \|x_k - x\| = 0$$

gilt. Analog zu $\mathbb{R}$ definieren wir Häufungspunkte von Folgen in metrischen Räumen,

8.6 Satz: *Bezeichnen $x_{k,1}, \ldots, x_{k,n}$ die Komponenten von x_k, so gilt*

$$\lim_{k\to\infty} x_k = x$$

genau dann, wenn

$$\lim_{k\to\infty} x_{k,i} = x_i$$

für alle $i = 1, \ldots, n$ gilt.

8.7. Beispiel: Viele Rechengesetze mit Folgen ergeben sich daher auf einfache Weise aus den entsprechenden Gesetzen in $\mathbb{R}$. Man kann also zum Beispiel Summe und Grenzwert vertauschen, weil dies auch in $\mathbb{R}$ möglich ist. Das heißt

$$\lim_{k\to\infty} x_k \pm y_k = \lim_{k\to\infty} x_k \pm \lim_{k\to\infty} y_k$$

Der Grenzwert auf der linken Seite existiert, wenn der Grenzwert auf der rechten Seite existiert. Analog für die Multiplikation λx, und viele andere Rechenverknüpfungen mit Vektoren. Wenn etwa $x_k \to x$ konvergiert, dann konvergiert

$$\|x_k\| \to \|x\|,$$

sowie

$$\langle x_k, y\rangle \to \langle x, y\rangle$$

für beliebiges $y \in \mathbb{R}^n$.

8.8 Satz: *Jeder beschränkte Folge in $\mathbb{R}^n$ besitzt einen Häufungspunkt. Der $\mathbb{R}^n$ ist ein vollständiger metrischer Raum, das heißt, jede Cauchy-Folge in $\mathbb{R}^n$ konvergiert. Jede absolut konvergente Reihe im $\mathbb{R}^n$ konvergiert.*

Beweis: Dieser Satz folgt sofort aus den entsprechenden Sätzen für $\mathbb{R}$. Um einen Häfungspunkt einer beschränkten Folge zu gewinnen, kann man aus den Komponentenfolgen sukzessive konvergente Teilfolgen isolieren. $\square$

8.2 Offene Mengen

Als Ersatz für das offene Intervall könnte man das offene Rechteck nehmen. Jedoch benötigen wir auch allgemeinere Mengen, wie etwa Kreise.

8.9. Definition: (1) Wir definieren die Kugel ohne Rand um $x \in \mathbb{R}^n$ mit Radius $r > 0$ durch

$$B_r(x) := \{y \in \mathbb{R}^n : \|x - y\| < r\}.$$

(2) Eine Teilmenge $U \subseteq \mathbb{R}^n$ heißt **Umgebung** eine Punktes $x \in U$, wenn es ein $\epsilon > 0$ gibt, so dass

$$B_\epsilon(x) \subseteq U$$

ist.

(3) U heißt **offen**, wenn es Umgebung aller seiner Punkte ist.

(4) Eine Teilmenge $A \subseteq \mathbb{R}^n$ heißt **abgeschlossen**, wenn der Grenzwert jeder konvergenten Folge in A wieder in A liegt.

(5) A heißt **beschränkt**, wenn es ein $C > 0$ gibt mit

$$\|x\| \leq C \qquad \text{für alle } x \in A.$$

8.10 Satz: *$U \subseteq \mathbb{R}^n$ ist genau dann offen, wenn $\mathbb{R}^n \setminus U$ abgeschlossen ist. $A \subseteq \mathbb{R}^n$ ist genau dann abgeschlossen, wenn $\mathbb{R}^n \setminus A$ offen ist.*

8.11. Beispiel: (1) Nach Definition sind $\emptyset$ und $\mathbb{R}^n$ gleichzeitig offen und abgeschlossen. Es sind die einzigen Mengen, die diese Eigenschaft haben.

(2) Die Kugeln ohne Rand $B_r(x)$ sind offen. In der Tat gilt

$$B_{r-\|x-y\|}(y) \subseteq B_r(x)$$

für alle $y \in B_r(x)$.

(3) Die Kugeln mit Rand

$$D_r(x) := \{y \in \mathbb{R}^n : \|x - y\| \leq r\}$$

sind abgeschlossen. Sei nämlich x_n, $n \in \mathbb{N}$, eine gegen a konvergente Folge in $D_r(x)$, so gilt $\|x_n\| \leq r$ für alle $n \in \mathbb{N}$. Es folgt

$$\|x - a\| = \|\lim_{n\to\infty} x_n - a\| = \lim_{n\to\infty} \|x_n - a\| \leq r.$$

Also $a \in D_r(x)$.

(4) Als Produktmenge der Mengen $A_1, \ldots, A_n \in \mathbb{R}$ bezeichnet man die Menge

$$A_1 \times \ldots \times A_n := \{x \in \mathbb{R}^n : x_i \in A_i \text{ für alle } i = 1, \ldots, n\}.$$

Die Produkte von offenen Mengen sind offen, die Produkte von abgeschlossenen Mengen sind abgeschlossen. Die Produkte von beschränkten Mengen sind beschränkt.

(5) Der Schnitt von endlich vielen offenen Mengen ist offen. Die Vereinigung von beliebig vielen offenen Mengen ist offen.

(6) Die Vereinigung von endlich vielen abgeschlossenen Mengen ist abgeschlossen. Der Schnitt von beliebig vielen abgeschlossenen Mengen ist abgeschlossen.

8.12. Definition: (1) Für eine Teilmenge $M \subseteq \mathbb{R}^n$ ist das **offene Innere** M° die größte offene Menge, die in M enthalten ist.

(2) Der **Abschluss** $\overline{M}$ der Menge $M \subseteq \mathbb{R}^n$ ist die kleinste abgeschlossene Menge, die M umfasst.

(3) Der **Rand** ∂M ist die Differenzmenge von Abschluss und offenem Inneren.

8.13 Satz: *Sei $M \subseteq \mathbb{R}^n$. Dann gilt*

$$M^\circ = \{x \in M : \textit{Es gibt ein } \epsilon > 0 \textit{ mit } B_\epsilon(x) \subseteq M\},$$

und

$$\overline{M} = \{x \in \mathbb{R}^n : \textit{Für alle } \epsilon > 0 \textit{ gilt } B_\epsilon(x) \cap M \neq \emptyset\},$$

sowie

$$\partial M = \{x \in \mathbb{R}^n : \textit{Für alle } \epsilon > 0 \textit{ gilt } B_\epsilon \cap M \neq \emptyset \textit{ und } B_\epsilon \cap (\mathbb{R}^n \setminus M) \neq \emptyset\}.$$

Der Abschluss $\overline{M}$ ist außerdem die Menge aller Grenzwerte von Folgen in M.

8.14. Beispiel: Es gilt

$$\overline{B_r(x)} = D_r(x), \quad D_r(x)^\circ = B_r(x).$$

Außerdem

$$\partial D_r(x) = \partial B_r(x) = \{y \in \mathbb{R}^n : \|x - y\| = r\}$$

für alle $r > 0$, $x \in \mathbb{R}^n$.

8.3 Stetige Funktionen

Man kann die Definition von Stetigkeit und die Definition von Grenzwerten von Funktionen auf ganz nahe liegende Weise verallgemeinern.

8.15. Definition: Sei $f : M \to \mathbb{R}^m$ eine Abbildung, $M \subseteq \mathbb{R}^n$. Dann schreiben wir für $a \in M$, $b \in \mathbb{R}^m$

$$\lim_{x \to a} f(x) = b,$$

wenn es zu jedem $\epsilon > 0$ ein $\delta > 0$ gibt, so dass

$$\|x - a\| < \delta \;\Rightarrow\; \|f(x) - b\| < \epsilon$$

für alle $x \in M$ gilt. f heißt stetig in $a \in M$, wenn

$$\lim_{x \to a} f(x) = f(a)$$

gilt. f heißt stetig in M, wenn es in allen Punkten von M stetig ist.

8.16 Satz: *Wie in $\mathbb{R}$ kann man die Konvergenz und die Stetigkeit auch äquivalent durch Folgen definieren.*

8.17. Bemerkung: Eine Abbildung $f : M \to \mathbb{R}^m$, $M \subseteq \mathbb{R}^n$, kann man in **Komponentenfunktionen**

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$$

zerlegen. f ist genau dann stetig in $a \in M$, wenn alle Komponentenfunktionen stetig sind. Auch die Grenzwerte kann man komponentenweise berechnen. Aus diesem Grund beschäftigen wir uns hauptsächlich mit Funktionen $f : M \to \mathbb{R}$, $M \subseteq \mathbb{R}^n$.

8.18. Beispiel: (1) Die Funktion $x \mapsto x_i$ ist für alle $i = 1, \ldots, n$ stetig.

(2) Es gelten dieselben Sätze über Summen, Differenzen, Produkte und Quotienten von Funktionen wie in $\mathbb{R}$.

(3) Auch die Hintereinanderausführung von stetigen Abbildungen ist stetig. Als Beispiel ist etwa die Funktion

$$x \mapsto \|x\| = \sqrt{x_1^2 + \ldots + x_n^2}$$

aufgrund dieser Tatsachen stetig.

8.19 Satz: (1) *Sei $M \subseteq \mathbb{R}^n$ offen und $f : M \to \mathbb{R}^m$ stetig, sowie $U \subseteq \mathbb{R}^n$ offen. Dann ist $f^{-1}(U)$ offen.*

(2) *Dasselbe gilt, wenn man „offen“ durch „abgeschlossen“ ersetzt.*

8.20. Beispiel: (1) Die Kugeln mit Rand sind abgeschlossen, die Kugeln ohne Rand sind offen. Denn

$$D_r(x) = f^{-1}[0, r]$$

mit der Funktion

$$f(x) = \|x\|$$

die stetig ist.

(2) Die Niveaulinien von stetigen Funktionen

$$N_c = \{x \in \mathbb{R}^n : f(x) = c\}$$

sind die Urbilder der abgeschlossenen Menge $\{c\}$ unter f und daher abgeschlossen. Ebenso sind die Subniveaumengen

$$S_c = \{x \in \mathbb{R}^n : f(x) \leq c\}$$

abgeschlossen. Die Mengen

$$U_c = \{x \in \mathbb{R}^n : f(x) < c\}$$

sind dagegen offen. Allerdings muss S_c nicht immer der Abschluss von U_c sein, und N_c nicht immer der Rand von S_c.

8.21. Bemerkung: In der Tat folgt aus dem Beweis, dass es für jedes $M \subseteq \mathbb{R}^n$ und eine stetige Abbildung $f : M \to \mathbb{R}^m$, sowie eine offene Menge $U \subseteq \mathbb{R}^m$ eine offene Menge $V \subseteq \mathbb{R}^n$ gibt, so dass

$$f^{-1}(U) = M \cap V$$

ist. Man nennt Schnitte von offenen Mengen mit M **relativ offen** in M.

8.4 Kompakte Mengen

8.22. Definition: Eine Teilmenge M eines metrischen Raums A heißt folgenkompakt, wenn jede Folge aus M einen Häufungspunkt hat.

8.23. Bemerkung: Eine Folge in einer folgenkompakten Menge konvergiert also genau dann, wenn alle ihre Häufungspunkte gleich sind.

8.24 Satz: *Eine Teilmenge $K \subseteq \mathbb{R}^n$ ist genau dann folgenkompakt, wenn K abgeschlossen und beschränkt ist.*

Beweis: Wenn $(x_k)_{k\in\mathbb{N}}$ eine Folge in einer abgeschlossenen und beschränkten Menge $K \subseteq \mathbb{R}^n$ ist, so sind die Komponentenfolgen ebenfalls beschränkt. Man kann daher sukzessive Teilfolgen aus den Komponenten aussondern, so dass eine konvergente Teilfolge in K entsteht. Da K abgeschlossen ist, muss der Grenzwert in K liegen. Also ist jede abgeschlossene und beschränkte Menge in $\mathbb{R}^n$ folgenkompakt.

Die Umkehrung ist einfach. □

8.25. Bemerkung: Da der $\mathbb{C}^n$ mit der Euklidschen Norm identisch zum $\mathbb{R}^{2n}$ mit der Euklidschen Norm ist, gilt dieser Satz auch in $\mathbb{C}^n$, versehen mit der Euklidschen Metrik.

8.26 Satz: *Seien A und B metrische Räume und $K \subseteq A$ folgenkompakt, sowie $f : K \to B$ eine stetige Funktion. Dann ist auch $f(K)$ folgenkompakt.*

Beweis: Sei $(y_n)_{n\in\mathbb{N}}$ eine Folge in $f(M)$, und

$$f(x_n) = y_n$$

für alle $n \in \mathbb{N}$. Dann hat $(x_n)_{n\in\mathbb{N}}$ einen Häufungspunkt in K, und es folgt für eine Teilfolge

$$\lim_{n\to\infty} x_{k(n)} = x \in M.$$

Also

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

□

8.27 Satz: *Sei $K \subseteq A$ folgenkompakt. Dann nimmt eine stetige Funktion $f : K \to \mathbb{R}$ auf K ein Maximum und ein Minimum an.*

Beweis: Da $f(K)$ folgenkompakt ist, ist es abgeschlossen und beschränkt, und besitzt daher ein Maximum und ein Minimum. □

8.28. Definition: Seien A und B metrische Räume. Die gleichmäßige Stetigkeit einer Funktion $f : M \to B$, $M \subseteq A$, wird ganz analog zum eindimensionalen Fall definiert. Es muss also zu jedem $\epsilon > 0$ ein $\delta > 0$ geben, so dass gilt

$$d(x, y) < \delta \;\Rightarrow\; d(f(x), f(y)) < \epsilon$$

für alle $x, y \in M$.

8.29 Satz: *Jede stetige Funktion auf einer folgenkompakten Menge ist gleichmäßig stetig.*

Beweis: Angenommen nicht, dann gibt es Folgen $(x_n)_{n\in\mathbb{N}}$ und $(y_n)_{n\in\mathbb{N}}$, sowie ein $\epsilon > 0$ mit

$$d(x_n, y_n) < \frac{1}{n}, \quad d(f(x_n), f(y_n)) > \epsilon$$

Da die beiden Folgen, sowie jede Teilfolge, Häufungspunkte haben, erhalten wir eine Teilfolge

$$\lim_{n\to\infty} x_{k(n)} = \lim_{n\to\infty} y_{k(n)}.$$

Aus der Stetigkeit von f folgt

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

Dies ist ein Widerspruch. □

8.30. Definition: Eine **Überdeckung** einer Menge $K \subseteq \mathbb{R}^n$ ist eine Menge von Mengen U_i, $i \in I$, so dass

$$K \subseteq \bigcup_{i\in I} U_i$$

ist. Wenn die Indexmenge I endlich ist, so sprechen wir von einer endlichen Überdeckung.

8.31 Satz: *Eine Teilmenge K eines metrischen Raumes ist genau dann folgenkompakt, wenn zu jeder Überdeckung aus offenen Mengen eine endliche Teilüberdeckung existiert.*

Beweis: „⇒“: Angenommen, die Menge der U_i, $i \in I$, ist eine offene Überdeckung ohne endliche Teilüberdeckung. Dann gibt es ein $\epsilon > 0$, so dass für jeden Punkt $x \in K$ ein $i_x \in I$ existiert mit

$$U_\epsilon(x) \subseteq U_{i_x}.$$

Denn sonst könnte man eine Folge von x_n, $n \in \mathbb{N}$, finden so dass $U_{1/n}(x_n)$ in keinem U_i enthalten ist. Diese Folge kann keinen Häufungspunkt haben. Mit dem so gefundenen $\epsilon > 0$ kann man man nun eine Folge von y_n, $n \in \mathbb{N}$ konstruieren, so dass

$$d(y_i, y_j) \geq \epsilon$$

für alle $i \neq j$ gilt. Diese Folge kann keinen Häufungspunkt haben.

„⇐“: Sei $(x_n)_{n\in\mathbb{N}}$ eine Folge in K ohne Häufungspunkt. Dann gibt es zu jedem $x \in K$ ein $\epsilon_x > 0$, so dass $U_{\epsilon_x}(x)$ nur endlich viele Folgenglieder enthält. Die Menge dieser offenen Kugeln überdeckt K. Es kann aber keine endliche Teilüberdeckung geben, da die Folge sonst nur endlich viele Folgenglieder hätte.
□

8.32. Bemerkung: Man nennt Mengen mit dieser Überdeckungseigenschaft **kompakt**. Im $\mathbb{R}^n$ ist also äquivalent, dass K abgeschlossen und beschränkt ist, dass K folgenkompakt ist, und dass K kompakt ist.

8.33 Satz: *Seien A und B metrische Räume, M eine folgenkompakte Teilmenge von A und $f : M \to B$ injektiv. Dann ist $f(M)$ folgenkompakt, und f hat eine stetige Umkehrabbildung*

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

Beweis: Wir haben nur noch zu zeigen, dass f^{-1} stetig ist. Dazu sei $(y_n)_{n\in\mathbb{N}}$ eine Folge in $f(M)$ mit $y_n \to f(x)$. Sei

$$x_n = f^{-1}(y_n).$$

für alle $n \in \mathbb{N}$. Wenn x ein Häufungspunkt von $(x_n)_{n\in\mathbb{N}}$ ist, so folgt aus der Stetigkeit von f sofort $f(x) = y$, also $x = f^{-1}(y)$. Da M kompakt ist und alle Häufungspunkte von $(x_n)_{n\in\mathbb{N}}$ gleich sind, folgt

$$\lim_{n\to\infty} f^{-1}(y_n) = \lim_{n\to\infty} x_n = x = f^{-1}(y).$$

□

8.5 Zusammenhängende und konvexe Mengen

8.34. Definition: Eine **Kurve** ist eine stetige Abbildung

$$\gamma : I \to \mathbb{R}^n,$$

wobei I ein Intervall in $\mathbb{R}$ ist. Falls $I = [a, b]$, so heißt die Kurve **geschlossen**, wenn

$$\gamma(a) = \gamma(b)$$

ist, wenn also ihr Anfangspunkt $\gamma(a)$ gleich dem Endpunkt $\gamma(b)$ ist. Die Kurve heißt **einfach**, wenn γ injektiv ist. Das Bild der Kurve wird manchmal mit der Kurve identifiziert. In diesem Fall heißt γ die Parametrisierung der Kurve.

8.35. Beispiel:

$$\gamma(t) = (\cos(t), \sin(t)) \in \mathbb{R}^2$$

für $t \in [0, w\pi[$. Das Bild der Kurve ist in diesem Fall der Einheitskreis. Die Kurve ist geschlossen ($\gamma(0) = \gamma(2\pi)$) und einfach.

8.36. Beispiel:

$$\gamma_{v,w}(t) = v + t(w - v), \qquad t \in [0, 1],$$

parametrisiert die Strecke $S_{v,w}$ von v nach w. Jeder Punkt auf der Strecke hat eine eindeutige Darstellung

$$x = \lambda v + \mu w, \qquad \lambda + \mu = 1, \quad \lambda, \mu \geq 0.$$

λ, μ heißen **baryzentrische** Koordinaten von x.

8.37. Definition: (1) Eine Teilmenge $M \subseteq \mathbb{R}^n$ keißt **zusammenhängend**, wenn es zu je zwei Punkten aus M eine Kurve in M gibt, die die Punkte verbindet.

(2) M heißt **konvex**, wenn die Strecke zwischen je zwei Punkten aus M wieder in M liegt.

(3) Sie heißt **strernförmig**, wenn es einen Punkt in $v \in M$ gibt, so dass die Strecke $S_{v,w}$ für alle $w \in M$ in M liegt.

8.38. Beispiel: Kreise $D_r(x)$ sind konvex, ebenso wie Halbebenen

$$H = \{x \in \mathbb{R}^n : \phi(x) \leq c\}$$

wobei $\phi : \mathbb{R}^n \to \mathbb{R}$ linear ist. Natürlich sind auch Unterräume konvex, ebenso wie Strecken.

8.39. Beispiel: Die zusammenhängenden Teilmengen von $\mathbb{R}$ sind die Intervalle.

8.40. Bemerkung: Konvexe Mengen sind sternförmig, und sternförmige Mengen sind zusammenhängend.

8.41 Satz: *Der Durchschnitt von konvexen Mengen ist konvex. Die Vereinigung von konvexen Mengen ist strenförmig, sofern ein Element im Schnitt existiert. Die Vereinigung von zusammenhängenden Mengen ist ebenfalls zusammenhängend, sofern ein Element im Schnitt existiert.*

8.42 Satz: *Das stetige Bild von zusammenhängenden Mengen ist zusammenhängend.*

8.43. Beispiel: Als Folgerung erhalten wir, dass das stetige Bild einer zusammenhängenden Menge M unter eine Abbildung $f : M \to \mathbb{R}$ ein Intervall in $\mathbb{R}$ ist.

8.44. Definition: Sei $M \subseteq \mathbb{R}^n$ konvex. Eine Abbildung $f : M \to \mathbb{R}$ heißt konvex, wenn

$$f(\lambda x + \mu y) \geq \lambda f(x) + \mu f(y)$$

für alle $x, y \in M$ ist.

8.45. Bemerkung: Dies ist äquivalent dazu, dass die Menge

$$\{(x, f(x)) : x \in M\} \subseteq \mathbb{R}^{n+1}$$

über dem Graphen von f konvex ist.

8.46 Satz: *Sei $U \subseteq \mathbb{R}^n$ konvex und offen, sowie $f : U \to \mathbb{R}$ konvex. Dann ist f stetig.*

8.47 Satz: *Sei $M \subseteq \mathbb{R}^n$ zusammenhängend und $f : M \to \mathbb{R}$ stetig. Falls dann*

$$f(v) = a < b < c = f(w)$$

für $v, w \in M$ und $a, b, c \in \mathbb{R}$, so existiert ein $\xi \in M$ mit

$$f(\xi) = b.$$

Dies ist eine Verallgemeinerung des Zwischenwertsatzes im $\mathbb{R}^n$.

8.6 Die Norm von linearen Abbildungen

Wir betrachten in diesem Abschnitt normierte Vektorräume über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

8.48 Satz: *Seien V und W zwei normierte Vektorräume. Dann ist eine lineare Abbildung $\phi : V \to W$ genau dann stetig, wenn*

$$\|\phi\| := \sup_{\|v\| \leq 1} \|\phi(v)\| < \infty$$

gilt. In diesem Fall definiert $\|\phi\|$ eine ***Norm*** *auf dem Vektorraum $\mathcal{L}(V, W)$ der stetigen linearen Abbildungen von V nach W.*

Beweis: Wenn ϕ stetig ist, so ist es in 0 stetig, und es existiert zu $\epsilon = 1$ ein $\delta > 0$ mit

$$\phi(U_\delta(0)) \subseteq U_1(0).$$

Es folgt für alle v mit $\|v\| \leq 1$

$$\|\phi(v)\| = \frac{1}{\delta}\|\phi(\delta v)\| \leq \frac{1}{\delta}.$$

Es folgt $\|\phi\| < \infty$. Sei umgekehrt $\|\phi\| < \infty$. Dann gilt

$$\|w - v\| < \delta \;\Rightarrow\; \|\phi(w-v)\| = \delta\|\phi(\frac{1}{\delta}(w-v)) < \delta\|\phi\|.$$

Es folgt, dass
phi stetig in v ist. Der Beweis dafür, dass ϕ eine Norm ist, ist eine Übungsaufgabe. □

8.49. Bemerkung: Es gilt für die Hintereinanderausführung zweier stetiger linearer Abbildungen

$$\|\phi \circ \psi\| \leq \|\phi\| \cdot \|\psi\|.$$

Außerdem gilt für die identische Abbildung $\|\mathrm{id}\| = 1$. Es folgt für für alle $v \in V$

$$\|\phi(v)\| \leq \|\phi\| \cdot \|v\|$$

$\|\phi\|$ ist die kleinste Konstante, so dass dies für alle $v \in V$ gilt.

8.50. Bemerkung: Jede lineare Abbildung ϕ von $\mathbb{K}^n$ (mit der Euklidschen Norm) in einen normierten Vektorraum W ist stetig. Es gilt nämlich

$$\phi(\lambda) \leq |\lambda_1|\|\phi(e_1)\| + \ldots + |\lambda_n|\|\phi(e_n)\| \leq C\|\lambda\|$$

aufgrund der Schwarzschen Ungleichung, und man erhält

$$\|\phi\| \leq \| \begin{pmatrix} \|\phi(e_1)\| \\ \vdots \\ \|\phi(e_n)\| \end{pmatrix} \|.$$

8.51 Satz: *Seien V, W zwei normierte Vektorräume und V endlich dimensional. Dann ist jede lineare Abbildung $\phi : V \to W$ stetig.*

Beweis: Sei $v_1, \ldots, v_n$ eine Basis von V, und $\psi : \mathbb{K}^n \to V$ definiert durch

$$\psi(\lambda) = \lambda_1 v_1 + \ldots \lambda_n v_n.$$

Dann ist ψ eine bijektive lineare Abbildung. Es gilt außerdem für $v = \lambda_1 v_1 + \ldots + \lambda_n v_n$

$$\|\phi(v)\| \leq |\lambda_1|\|\phi(v_1)\| + \ldots + |\lambda_n|\|\phi(v_n)\|.$$

Es genügt also zu zeigen, dass es eine von v unabhängige Konstante $c > 0$ gibt, so dass gilt

$$\|v\| \leq 1 \;\Rightarrow\; \max_k |\lambda_k| \leq c.$$

Dies ist übrigens äquivalent zu $\|\psi^{-1}\| \leq c$. Angenommen, das ist nicht der Fall. Dann gibt es eine Folge $v_m \in V$, so dass $\|v_m\| = 1$ ist, aber $\|\psi^{-1}(v_m)\|$ nicht beschränkt ist. Daraus gewinnt man umgekehrt eine Folge $\lambda_m \in \mathbb{R}^m$

$$\|\lambda_m\| = 1, \quad \|\psi(\lambda_m)\| \to 0.$$

Wegen der Kompaktheit von

$$K = \{\lambda \in \mathbb{K}^n : \|\lambda\| = 1\}$$

in $\mathbb{K}^n$ und der Stetigkeit von $\|\psi(\lambda)\|$ ist das aber ein Widerspruch dazu, dass $\psi(\lambda)$ auf K nicht gleich 0 wird. □

8.52. Definition: Fixiert man eine Norm auf $\mathbb{K}^n$ und $\mathbb{K}^m$, so erhält man eine Norm auf dem Raum der Matrizen $\mathbb{K}^{m\times n}$ als Norm der zugehörigen linearen Abbildung.

8.53. Beispiel: (1) Nimmt man die Maximumsnorm

$$\|v\|_\infty = \max_k |v_k|$$

sowohl auf $\mathbb{K}^n$, als auch auf $\mathbb{K}^m$, so ist die **Zeilensummennorm**

$$\|A\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^{n} |a_{i,j}|$$

die zugehörige Matrixnorm.

(2) Für die l_1-Norm

$$\|v\|_1 = \sum_k |v_k|$$

auf Quell- und Zielraum ist es die Spaltensummennorm

$$\|A\|_\infty = \max_{1\leq j\leq m} \sum_{i=1}^{m} |a_{i,j}|$$

(3) Für die Euklidsche Norm

$$\|v\| = \sqrt{\sum_k |v_k|^2}$$

ist es der Spektralradius

$$\|A\| = \sqrt{\rho(A^T \cdot A)}$$

Dabei steht $\rho(M)$ für den betragsgrößten (komplexen) Eigenwert der Matrix M. (Man beachte, dass $A^T A$ allerdings nur nicht-negative reelle Eigenwerte hat.)

8.54 Satz: *Alle Normen auf einem endlich-dimensionalen Vektorraum V sind äquivalent. Das heißt, zu zwei Normen $\|v\|_1$ und $\|v\|_2$ existieren Konstanten $0 < c < d$ mit*

$$c\|v\|_1 \leq \|v\|_2 \leq d\|v\|_1$$

für alle $v \in V$. Die Metriken, die zu den beiden Normen gehören, erzeugen dieselben konvergenten Folgen, dieselben abgeschlossenen, offenen oder kompakten Mengen.

Beweis: Man kann

$$d = \|\mathrm{id}\|$$

wählen, wobei id die Abbildung von V versehen mit $\|v\|_2$ nach V versehen mit $\|v\|_1$ ist. □

8.7 Der Banachsche Fixpunktsatz

8.55 Satz: *Sei X ein vollständiger metrischer Raum und $A \subseteq X$ abgeschlossen, sowie*

$$f : A \to A$$

eine kontrahierende Abbildung, d.h. es gibt eine Konstante $c < 1$ mit

$$d(f(x), f(y)) \leq cd(x, y)$$

für alle $x, y \in A$. Dann hat f genau einen Fixpunkt

$$f(x) = x$$

in A, und jede Folge

$$x_0 \in A, \quad x_{n+1} = f(x_n)$$

konvergiert gegen x.

Beweis: Man erhält zunächst für $m \geq n$ per Induktion

$$d(x_{m+1}, x_m) \leq c^{m-n} d(x_{n+1}, x_n)$$

Daraus folgt

$$d(x_m, x_n) \leq d_(x_{n+1}, x_n) + d_(x_{n+2}, x_{n+1}) + \ldots \leq \frac{c^n}{1-c} d(x_1, x_0).$$

Es folgt, dass die x_n eine Cauchy-Folge bilden, und daher konvergieren. Da A abgeschlossen ist, liegt der Grenzwert x in A. Da f automatisch stetig ist, ist x ein Fixpunkt von f. □

8.56. Bemerkung: Es gilt die Abschätzung

$$d(x_n, x) \leq \frac{1}{1-c} d(x_{n+1}, x_n).$$

8.57. Bemerkung: Folgenkompakte metrische Räume sind automatisch vollständig.

Kapitel 9

Differenzialrechung mehrerer Variablen

Wir erweitern hier die Differenzialrechung auf Funktionen, die von mehreren Veränderlichen abhängig sind, die also auf Teilmengen des $\mathbb{R}^n$ gegeben sind. Dabei werden so weit wie möglich die Resultate der eindimensionalen Differenzialrechung verwendet.

9.1 Richtungsableitungen

Zunächst leiten wir Funktionen von mehreren Variablen einfach nach jeder Variablen einzeln ab, wobei die anderen fest bleiben. Danach erst setzen wir diese „partiellen“ Ableitungen zusammen.

9.1. Definition: Sei $U \subseteq \mathbb{R}^n$ offen. Dann heißt eine Funktion $f : U \to \mathbb{R}$ **partiell differenzierbar** nach x_i in x, wenn die Abbildung

$$g_{x,i}(t) := f(x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n)$$

definiert auf der offenen Menge

$$T_{x,i} := \{t \in \mathbb{R} : (x_1, \ldots, x_{i-1}, t, x_{i+1}, \ldots, x_n) \in U\}$$

nach t differenzierbar ist. In diesem Fall schreiben wir

$$D_i f(x) = \frac{\partial}{\partial x_i} f(x) := g'_{x,i}(x_i).$$

f heißt partiell differenzierbar in x wenn es nach allen Variablen partiell differenzierbar ist, und partiell differenzierbar in U, wenn es in allen Punkten $x \in U$ partiell differenzierbar ist. f heißt **stetig partiell differenzierbar** in U, wenn die Funktionen

$$\frac{\partial}{\partial x_1} f(x), \ldots, \frac{\partial}{\partial x_n} f(x)$$

stetig in U sind. Wenn diese Funktionen partiell differenzierbar sind, so schreiben wir für diese **zweifachen** partiellen Ableitungen

$$\frac{\partial^2}{\partial x_i \partial x_j} f(x), \quad \frac{\partial^2}{\partial x_i^2} f(x)$$

Entsprechend definiert man **zweimal stetig partiell** differenzierbare Funktionen etc. Man definiert den n-dimensionale Zeilenvektor, bzw. die $1 \times n$-Matrix

$$\operatorname{grad} f(x) := \left(\frac{\partial}{\partial x_1} f(x) \quad \ldots \quad \frac{\partial}{\partial x_n} f(x) \right),$$

die man den **Gradienten** von f in x nennt.

9.2. Bemerkung: Man kann genau wie im Eindimensionalen auch **einseitige** Ableitungen betrachten. Wir wollen jedoch hier davon abgesehen und annehmen, dass die Definitionsmenge U offen ist.

9.3. Definition: Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine Funktion. Für $v \in \mathbb{R}^n$ definieren wir

$$g_{x,v}(t) := f(x + tv),$$

sofern $x + tv \in U$. Dann heißt

$$D_v f(x) = g'_{x,v}(0)$$

die **Richtungsableitung** von f in x in Richtung v, sofern diese Ableitung existiert.

9.4. Bemerkung: (1) Die partielle Ableitung ist offenbar die Richtungsableitung in Richtung der Einheitsverktoren $e_1, \ldots, e_n$, also

$$D_i f(x) = D_{e_i} f(x).$$

(2) Es gilt mit den obigen Notationen

$$D_h f(x + tv) = g'_{x,v}(t).$$

(3) Die Richtungsableitung hängt von der Länge des Vektors v ab. Es gilt nämlich

$$D_{\lambda v} f(x) = \lambda D_v f(x).$$

Die „natürliche" Steigung der Funktion f in Richtung $v \neq 0$ definiert man daher am besten als

$$m_v := \frac{D_v f(x)}{\|v\|}.$$

Dies ist die anschaulich sichtbare Steigung am Graphen von f.

9.5 Satz: *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine stetig partiell differenzierbare Funktion. Dann existiert für alle $x \in U$ und $v \in \mathbb{R}^n$ die Richtungsableitung, und es gilt*

$$D_v f(x) = \sum_{i=1}^{n} v_i \frac{\partial}{\partial x_i} f(x) = \operatorname{grad} f(x) \cdot v.$$

Für die Funktion $R(h)$, die für festes x durch

$$f(x+h) = f(x) + \text{grad } f(x) \cdot h + R(h)$$

definiert ist, gilt

$$\lim_{h \to 0} \frac{R(h)}{\|h\|} = 0.$$

Man schreibt dafür $R(h) = o(h)$.

9.6. Bemerkung: Es gilt

$$\text{grad } f(x) \cdot v = \langle \text{grad } f(x), v \rangle.$$

Die Matrixmultipliation eines Zeilenvektors mit einem Spaltenvektor entspricht dem Skalarprodukt.

Beweis: Durch sukzessive Anwendung des Mittelwertsatzes erhält man

$$f(x+tv) - f(x) = \sum_{i=1}^{n} tv_i \frac{\partial}{\partial x_i} f(\xi_i)$$

mit Punkten

$$\xi_i \in B_{t\|v\|}(x) \qquad \text{für alle } i = 1, \ldots, n.$$

Mit $t \to 0$ folgt die Existenz der Richtungsableitung und die erste Behauptung aus dem Satz. Mit $t = 1$ erhalten wir

$$|R(h)| = \left| \sum_{i=1}^{n} h_i (D_i f(x) - D_i f(\xi_i)) \right|$$

mit Punkten $\xi_i \in B_{\|h\|}(x)$. Aufgrund der Schwarzschen Ungleichung und der Stetigkeit der partiellen Ableitungen gibt es für jedes $\epsilon > 0$ ein $\delta > 0$, so dass

$$\|h\| < \delta \ \Rightarrow \ \frac{|R(h)|}{\|h\|} \leq \epsilon.$$

Es folgt die Behauptung. □

9.7. Bemerkung: Der Satz besagt, dass es für jedes $x \in U$ eine lineare Abbildung $L(h)$ gibt, so dass

$$f(x+h) = f(x) + L(h) + R(h)$$

gilt, wobei $R(h) = o(h)$ ist. Die affine Abbildung

$$T(y) = f(x) + L(y - x)$$

nennt man **tangential** an die Funktion f. Den Graph der tangentialen affinen Abbildung nennt man **Tangentialebene** an f.

9.2 Lokale Extrema

Wir können nun die Richtungsableitung ausnutzen, um lokale Minima und Maxima zu erkennen. Für hinreichende Bedingungen müssen wir allerdings die zweite Ableitung zu Hilfe nehmen.

9.8 Satz: *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine partiell differenzierbare Funktion. In $x \in U$ liege ein lokales Extremum von f. Dann sind alle partiellen Ableitungen gleich 0, also*

$$\operatorname{grad} f(x) = 0.$$

9.9. Bemerkung: Genau wie im eindimensionalen ist dies nur eine **notwendige** Bedingung. Um eine hinreichende Bedingung zu erhalten, müssen wir die zweiten Richtungsableitungen berechnen.

9.10. Definition: Wir schreiben für die zweiten partiellen Ableitung, zuerst nach x_j, dann nach x_i

$$D_{i,j}f(x) = \frac{\partial^2}{\partial x_i \partial x_j} f(x).$$

Die Matrix

$$H(x) := \begin{pmatrix} \frac{\partial^2}{\partial x_1 \partial x_1} f(x) & \dots & \frac{\partial^2}{\partial x_n \partial x_1} f(x) \\ \vdots & & \vdots \\ \frac{\partial^2}{\partial x_n \partial x_1} f(x) & \dots & \frac{\partial^2}{\partial x_n \partial x_n} f(x) \end{pmatrix}$$

heißt **Hesse-Matrix** von f in x.

9.11 Satz: *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion. Dann gilt*

$$\frac{\partial^2}{\partial x_i \partial x_j} f(x) = \frac{\partial^2}{\partial x_j \partial x_i} f(x)$$

für alle $1 \leq i, j \leq n$. Die Hesse-Matrix ist also symmetrisch.

Beweis: Zur Vereinfachung der Notation sei $n = 2$. Man verwendet die Funktion

$$g(t) = f(t, x_2 + h_2) - f(t, x_2)$$

an, und erhält

$$g(x_1 + h_1) - g(x_1) = h_1 h_2 \frac{\partial^2}{\partial x_2 \partial x_1} f(\xi_1, \xi_2)$$

mit Zwischenwerten ξ_1, ξ_2. Es gilt aber auch

$$g(x_1 + h_1) - g(x_1) = \tilde{g}(x_2 + h_2) - \tilde{g}(x_2) = h_1 h_2 \frac{\partial^2}{\partial x_1 \partial x_2} f(\tilde{\xi}_1, \tilde{\xi}_2),$$

mit

$$\tilde{g}(t) = f(x_1 + h_1, t) - f(x_1, t).$$

Mit $h \to 0$ und der Stetigkeit der zweiten partiellen Ableitungen folgt die Behauptung. $\square$

9.12 Satz: *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion. Dann existiert für $x \in U$ und $h \in \mathbb{R}$ die zweite Ableitung der Funktion*

$$f_{x,h}(t) := f(x + th)$$

in allen Punkten mit t mit $x + th \in U$, und es gilt

$$f''_{x,h}(t) = \sum_{1 \le i,j \le n} h_i h_j \frac{\partial^2}{\partial x_i \partial x_j} f(x + th) = h' \cdot H(x + th) \cdot h.$$

Beweis: Gemäß dem vorigen Abschnitt ist

$$f'_{x,h}(t) = \sum_{i=1}^{n} h_i \frac{\partial}{\partial x_i} f(x + th).$$

Die Behauptung folgt durch erneutes Differenzieren. $\square$

9.13. Bemerkung: Wie verwenden nun den Satz von Hurwitz über positiv definite Matrizen. Eine reelle, quadratische $n \times n$-Matrix H heißt **positiv definit**, wenn

$$h' \cdot H \cdot h > 0 \qquad \text{für alle } h \neq 0,\ h \in \mathbb{R}^n$$

gilt. Nach dem Kriterium von Hurwitz ist dies genau dann der Fall, wenn

$$\det H_i > 0 \qquad \text{für alle } i = 1, \ldots, n$$

gilt, wobei H_i die linken oberen $i \times i$-Untermatrizen von H bezeichnet. Damit kann man nachrechnen, ob eine Matrix positiv definit ist. H heißt **negativ definit**, wenn $-H$ positiv definit ist. Dies ist genau dann der Fall, wenn

$$\det H_1 < 0, \quad \det H_2 > 0, \quad \det H_3 < 0, \quad \ldots$$

gilt.

9.14 Satz: *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine zweimal stetig partiell differenzierbare Funktion. Sei für $x \in U$*

$$\text{grad } f(x) = 0$$

und sei $H(x)$ positiv definit. Dann liegt in x ein lokales Minimum vor. Wenn $H(x)$ negativ definit ist, so handelt es sich um ein lokales Maximum.

Beweis: Aufgrund des Kriterium von Hurwitz gilt die positive Definitheit in einer Umgebung V von x. Also sind alle oben berechneten zweiten Ableitungen positiv, wenn $x + th \in V$ ist. Deswegen hat f in allen Richtungen eindeutige lokale Maxima in x. Insgesamt ergibt sich die Behauptung. $\square$

9.15. Bemerkung: Ganz analog zum reellen Fall ist ein lokales Minimum von $f : \mathbb{R}^n \to \mathbb{R}$ gobal, wenn die Funktion überall positiv definite Hessematrizen besitzt. Dies kann man noch auf Funktionen $f : U \to \mathbb{R}$, $U \subseteq \mathbb{R}^n$ offen, verallgemeinern, wenn U die Eigenschaft hat, dass jeder Punkt vom Minimum aus durch eine Strecke zu erreichen ist (Solche Mengen heißen **sternförmig**).

9.16. Bemerkung: Eine zweimal stetig partiell differenzierbare Funktion $f : U \to \mathbb{R}$, die auf einer konvexen Menge definiert ist, ist genau dann konvex, wenn $Hf(x)$ über all positiv semi-definit ist. Wenn $Hf(x)$ überall positiv definit ist, so ist die Funktion strikt konvex.

9.3 Die totale Ableitung

In diesem Abschnitt erweitern wir die Differenzialrechnung auf Abbildung in den $\mathbb{R}^m$.

9.17. Definition: Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ eine Abbildung. Dann heißt f **total differenzierbar** (oder einfach **differenzierbar**) in $x \in U$, wenn es eine lineare Abbildung $L : \mathbb{R}^n \to \mathbb{R}^m$ gibt, so dass für die durch

$$f(x + h) = f(x) + L(h) + R(h)$$

definierte Restfunktion gilt

$$\lim_{h \to 0} \frac{R(h)}{\|h\|} = 0.$$

Die zu L gehörige Matrix heißt **Jocobi-Matrix** und wird mit $Df(x)$ oder mit $J(x)$ bezeichnet, also

$$f(x + h) = f(x) + Df(x) \cdot h + R(h).$$

f heißt differenzierbar in U, wenn es in allen $x \in U$ differenzierbar ist. f heißt **stetig differenzierbar**, wenn die Elemente der Jacobimatrix stetig von x abhängen.

9.18 Satz: *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^m$ eine Abbildung mit den Komponenten $f_1, \ldots, f_m$. Seien alle Komponentenabbildungen stetig partiell in $x \in U$ differenzierbar. Dann ist f in x stetig differenzierbar, und es gilt*

$$Df(x) = \begin{pmatrix} \frac{\partial}{\partial x_1} f_1(x) & \ldots & \frac{\partial}{\partial x_n} f_1(x) \\ \vdots & & \vdots \\ \frac{\partial}{\partial x_1} f_m(x) & \ldots & \frac{\partial}{\partial x_n} f_m(x) \end{pmatrix}.$$

Wenn umgekehrt f differenzierbar ist, dann sind auch alle Komponentenabbildungen partiell differenzierbar, und es existieren alle Richtungsableitungen.

9.19. Beispiel: Linear affine Abbildungen

$$f(x) = A \cdot x + b$$

sind differenzierbar mit Jacobimatrix A.

9.20. Bemerkung: Wenn $U \subseteq \mathbb{R}$ ist ($n = 1$), dann ist f genau dann differenzierbar, wenn alle f_i differenzierbar sind. In diesem Fall kann man also auf die Voraussetzung der stetigen Differenzierbarkeit verzichten.

9.21. Bemerkung: Wenn f in $x \in U$ differenzierbar ist, ist es in x stetig.

9.22 Satz: *Sei $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^k$ offen, sowie*

$$f : U \to V, \quad g : V \to \mathbb{R}^m$$

Abbildungen. Sei f in $x \in U$ differenzierbar, und g in $f(x)$. Dann ist $g \circ f$ in x differenzierbar und es gilt

$$D(g \circ f)(x) = Dg(f(x)) \cdot Df(x).$$

*Dies ist die **Kettenregel** für vektorwertige Abbildungen.*

Beweis: Sei $x \in U$, sowie R und $\tilde{R}$ die Restfunktionen des Differentials in x und $f(x)$. Dann gilt

$$\begin{aligned} g(f(x+h)) &= g(f(x) + Df(x)h + R(h)) \\ &= g(f(x)) + Dg(f(x))(Df(x)h + R(h)) + \tilde{R}(Df(x)h + R(h)) \\ &= g(f(x)) + (Dg(f(x))Df(x))h \\ &\qquad + Dg(R(h)) + \tilde{R}(Df(x)h + R(h)). \end{aligned}$$

Wir müssen diese Restfunktion noch untersuchen. Es gilt aber

$$\frac{\|Dg(R(h))\|}{\|h\|} \le \|Dg\| \frac{\|R(h)\|}{\|h\|} \to 0$$

mit $h \to 0$. Außerdem

$$\frac{\|\tilde{R}(Df(x)h + R(h))\|}{\|h\|} \le \frac{\|\tilde{R}(Df(x)h + R(h))\|}{\|Df(x)h + R(h)\|} \cdot \left(\|Df(x)\| + \frac{\|R(h)\|}{\|h\|} \right)$$

Mit $h \to 0$ geht der erste Faktor gegen 0, und der zweite ist beschränkt. $\square$

9.23. Beispiel: Das für uns wichtige Beispiel ist eine **Kurve** im $\mathbb{R}^n$, die durch eine Abbildung

$$f : I \to \mathbb{R}^n$$

mit Komponentenfunktionen

$$f_1, \ldots, f_n : I \to \mathbb{R}$$

gegeben ist, wobei I ein reelles Intervall ist. Man schreibt, wenn f differenzierbar ist,

$$f'(x) := Df(x) = \begin{pmatrix} f_1'(x) \\ \vdots \\ f_n'(x) \end{pmatrix}.$$

Wenn $f(I) \subseteq U$, U offen, und $g : U \to \mathbb{R}$ eine differenzierbare Funktion ist, so erhalten wir

$$\frac{d}{dx} g(f_1(x), \ldots, f_n(x)) = \text{grad } g(f(x)) \cdot f'(x) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i} g(x_i) f_i'(x).$$

9.24 Satz: *Sei $f : [a,b] \to \mathbb{R}^m$ stetig und in $]a,b[$ differenzierbar. Dann gilt*

$$\|f(b) - f(a)\| \le \left(\sup_{\xi \in]a,b[} \|f'(\xi)\| \right) |b - a|$$

Gleichheit gilt nur dann, wenn $f'(\xi)$ konstant ist.

Beweis: Mit $v = f(b) - f(a)$ setzen wir

$$g(t) := \langle f(t) - f(a), v \rangle$$

Dann gilt

$$g'(t) = \langle g'(t), v \rangle$$

Es folgt aus dem eindimensionalen Mittelwertsatz und der Schwarzschen Ungleichung

$$\|f(b) - f(a)\|^2 = g(b) - g(a) = g'(\xi)(b - a) \le \|f'(\xi)\| \cdot \|f(b) - f(a)\| \cdot (b - a).$$

Daraus folgt die Behauptung. Damit die Gleichheit gilt, muss aufgrund des eindimensionalen Falles $g'(t)$ konstant sein, und aufgrund der Schwarzschen Ungleichung $f'(\xi)$ ein Vielfaches von $f(b) - f(a)$. Daraus folgt, dass $f'(\xi)$ konstant sein muss. □

9.25. Bemerkung: Für $m = 1$ folgt der Satz aus dem Mittelwertsatz der Differentialrechnung. Im Fall $m = 3$ bedeutet er physikalisch, dass der zurückgelegte Weg mit der maximalen Geschwindigkeit abgeschätzt werden kann. Man kann den Satz also die „physikalische Dreiecksungleichung“ nennen.

9.26. Bemerkung: Also Konsequenz folgt für allgemeine differenzierbare Abbildungen $f : U \to \mathbb{R}^m$, $U \subseteq \mathbb{R}^n$ offen,

$$\|f(x_1) - f(x_2)\| \le \left(\sup_{\xi \in]0,1[} \|Df(x_1 + \xi(x_2 - x_1))\| \right) \|x_1 - x_2\|,$$

wobei wie annehmen, dass die Strecke zwischen x_1 und x_2 in U liegt.

9.27 Satz: *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}^n$ stetig differenzierbar, sowie $x \in U$ mit*

$$\det Df(x) \neq 0.$$

Dann existiert eine offene Umgebung $V \subseteq U$ von x, so dass $f(V)$ ebenfalls offen ist und $f : V \to f(V)$ bijektiv ist. f^{-1} ist in V stetig differenzierbar, und es gilt

$$Df^{-1}(f(x)) = Df(x)^{-1}.$$

*f ist also **lokal invertierbar**, wenn $Df(x)$ invertierbar ist.*

Beweis: Sei $y = f(x)$. Wir wählen $\epsilon, \delta > 0$ mit später zu spezifierenden Eigenschaften. Es soll zunächst gelten

$$U_{2\epsilon}(x) \subseteq U.$$

Sei $s \in U_\delta(y)$. Wir suchen ein $t \in U_\epsilon(x)$ mit Hilfe der Fixpunktiteration

$$t_0 = x, \quad t_{n+1} = \phi(t_n) := t_n - Df(x)^{-1}(f(t_n) - s).$$

Um den Banachschen Fixpunktsatz anzuwenden, müssen wir zunächst

$$\phi(D_\epsilon(x)) \subseteq D_\epsilon(x)$$

nachweisen. Dazu berechnen wir mit Hilfe der Differenzierbarkeit von f in x

$$\begin{aligned}
\|\phi(t) - x\| &= \|t - x - Df(x)^{-1}(f(x) + Df(x)(t-x) + R(t-x) - s)\| \\
&= \|Df(x)^{-1}(y - s + R(x-t))\| \\
&\leq \|Df(x)^{-1}\| \left(1 + \frac{\|R(t-x)\|}{\|t-x\|}\right) \delta \\
&< \epsilon.
\end{aligned}$$

wenn wir zunächst $\epsilon > 0$ so klein wählen, dass

$$\frac{\|R(t-x)\|}{\|t-x\|} < 1$$

ist, und dann

$$\delta < \frac{\epsilon}{2\|Df(x)^{-1}\|}.$$

Dann müssen wir nachweisen, dass ϕ auf $D_\epsilon(x)$ kontrahiert. Es gilt

$$\|\phi(t) - \phi(\tilde{t})\| \leq \max_{\xi \in D_\epsilon(x)} \|D\phi(\xi)\| \|t - \tilde{t}\|.$$

Außerdem

$$D\phi(\xi) = I - Df(x)^{-1} Df(\xi).$$

Diese Matrix konvergiert mit $\xi \to x$ gegen die Nullmatrix. Wählt man $\epsilon > 0$ klein genug, so ist also die Kontraktion gesichert.

Jeder Fixpunkt t von ϕ erfüllt offensichtlich $f(t) = s$. Wir haben also nachgewiesen, dass jedes $s \in U_\delta(y)$ genau ein Urbild in $D_\epsilon(x)$ hat. Wegen der obigen Abschätzung ist dieses Urbild in der Tat in $U_\epsilon(x)$. Also ist $f(U_\epsilon(x))$ Umgebung von $f(x)$.

Da die Determinante einer Matrix stetig von der Matrix abhängt, gibt nun eine offene Umgebung V von x, so dass $Df(\tilde{x})$ für alle $\tilde{x} \in V$ invertierbar ist. Indem wir die obigen Betrachtungen in jedem $\tilde{x} \in V$ anstellen, folgt, dass $f(V)$ offen ist. Wählen wir $V \subseteq U_\epsilon(x)$, so folgt, dass

$$f : V \to f(V)$$

bijektiv ist.

Wir müssen nun noch nachweisen, dass f^{-1} in y differenzierbar ist, und $Df^{-1}(y) = Df(x)^{-1}$ gilt. Dazu betrachten wir den Rest $\tilde{R}$ der Umkehrfunktion.

$$\begin{aligned}\tilde{R}(s-y) &= f^{-1}(s) - f^{-1}(y) - Df(x)^{-1}(s-y) \\ &= t - x - Df(x)^{-1}(f(t) - y) \\ &= t - x - Df(x)^{-1}(f(x) + Df(x)(t-x) + R(t-x) - y) \\ &= Df(x)^{-1}R(t-x).\end{aligned}$$

Also
$$\frac{\|R(s-y)\|}{\|s-y\|} \leq \|Df(x)^{-1}\| \frac{\|R(t-x)\|}{\|t-x\|} \frac{\|t-x\|}{\|s-y\|}.$$
Mit $s \to y$ geht auch $t \to x$, und daher geht der zweite Faktor gegen 0. Außerdem

$$s - y = f(t) - f(x) = Df(x)^{-1}(t-x) + R(t-x),$$

woraus

$$\begin{aligned}&\|t-x\| \left(1 - \|Df(x)^{-1}\| \frac{\|R(t-x)\|}{\|t-x\|}\right) \\ &\leq \|t-x\| \left(1 - \|Df(x)^{-1} \frac{1}{\|t-x\|} R(x-t)\|\right) \\ &\leq \|t-x\| \left\| \frac{1}{\|t-x\|}(t-x) - Df(x)^{-1} \frac{1}{\|t-x\|} R(t-x) \right\| \\ &= \|t - x - Df(x)^{-1}R(t-x)\| \\ &\leq \|Df(x)^{-1}\| \, \|s-y\|\end{aligned}$$

folgt. Es folgt
$$\|t-x\| \leq C\|s-y\|.$$
Daraus folgt die Differenzierbarkeit von f^{-1} in y. □

9.28. Bemerkung: Es folgt, dass das Bild von offenen Mengen offen ist, wenn die Abbildung überall eine invertierbare Jacobi-Matrix hat. Deswegen heißt dieser Satz auch Satz von der offenen Abbildung. Man beachte, dass solche Abbildung allerdings nur lokal injektiv sein müssen, nicht global.

9.29. Beispiel: Das einfachste Beispiel ist eine lineare Abbildung

$$f(x) = A \cdot x + b,$$

deren Ableitung
$$Df(x) = A$$
für alle $x \in \mathbb{R}^n$ ist. Wenn A invertierbar ist, so ist die Umkehrabbildung

$$f^{-1}(y) = A^{-1}y + A^{-1}b$$

mit Ableitung
$$Df^{-1}(y) = A^{-1} = Df(x)^{-1}$$
für alle $y \in \mathbb{R}^m$ und $f(x) = y$.

9.4 Implizite Funktionen

9.30. Beispiel: Versucht man die Gleichung

$$x_1^2 + x_2^2 = 1$$

nach x_2 aufzulösen, so erhält man zwei Zweige von Funktionen

$$x_2 = \pm\sqrt{1 - x_1^2}.$$

Lokal ist dies aber eindeutig möglich in jedem Punkt auf dem Einheitskreis mit $x_2 \neq 0$. Die entstehende Funktion ist sogar differenzierbar. Der folgende Satz verallgemeinert das.

9.31 Satz: *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine stetig partiell differenzierbare Funktion. Sei $x \in U$ mit*

$$f(x) = 0.$$

Wir setzen weiter voraus, dass

$$\frac{\partial}{\partial x_n} f(x) \neq 0$$

ist. Dann gibt es eine Umgebung V von x und eine Umgebung $\tilde{V}$ von

$$\tilde{x} = (x_1, \ldots, x_{n-1}) \in \mathbb{R}^{n-1},$$

auf der eine stetig partiell differenzierbare Funktion

$$g : \tilde{V} \to \mathbb{R}$$

definiert ist mit

$$f(t_1, \ldots, t_n) = 0 \;\Leftrightarrow\; t_n = g(t_1, \ldots, t_{n-1})$$

für alle $t \in V$. Es gilt

$$\frac{\partial}{\partial t_i} g(t_1, \ldots, t_{n-1}) = -\frac{\frac{\partial}{\partial t_i} f(t)}{\frac{\partial}{\partial t_n} f(t)}.$$

für alle $i = 1, \ldots, n-1$ mit

$$t = (t_1, \ldots, t_{n-1}, g(t_1, \ldots, t_{n-1})).$$

Der Satz gilt entsprechend für alle anderen Variablen.

Beweis: Für $t \in U$ definieren wir die Abbildung

$$h(t) = (t_1, \ldots, t_{n-1}, f(t)).$$

Man rechnet nach, dass $\det Dh(x) \neq 0$ ist. Daher besitzt h eine lokale Umkehrfunktion $h^{-1} : \tilde{V} \to \tilde{U}$, wobei $\tilde{U}$ eine Umgebung von x ist. Es gilt

$$h(x) = (y_1, \ldots, y_{n-1}, 0).$$

Wir definieren g durch

$$h^{-1}(t_1, \ldots, t_{n-1}, 0) = (t_1, \ldots, t_{n-1}, g(t_1, \ldots, t_{n-1})).$$

Es gilt dann also

$$f(t_1, \ldots, t_{n-1}, g(t_1, \ldots, t_{n-1})) = 0 \qquad \text{für alle } t \in \tilde{V}.$$

Differenzieren gemäß der Kettenregel für vektorwertige Abbildungen ergibt

$$\frac{\partial}{\partial t_i} f(t) + \frac{\partial}{\partial t_n} f(t) \cdot \frac{\partial}{\partial t_i} g(t_1, \ldots, t_{n-1}) = 0,$$

woraus die Behauptung folgt. □

9.32. Beispiel: Im obigen Beispiel gilt für $x_2 > 0$ und

$$f(x_1, x_2) := x_1^2 + x_2^2 = 0$$

dass

$$g(x_1) = \sqrt{1 - x_1^2},$$

ist, definiert auf $]-1, 1[$. In der Tat

$$g'(x_1) = -\frac{2x_1}{2x_2} = -\frac{x_1}{\sqrt{1 - x_1^2}}.$$

9.33. Bemerkung: Der Satz kann so gelesen werden, dass bei Lösbarkeit der Gleichung $f(x) = c$, die Gleichung für in einer Umgebung beliebig vorgegebene $x_1, \ldots, x_{n+1}$ ebenfalls eine eindeutige Lösung hat. Definiert man

$$\tilde{f}(x) = f(x) - c$$

so kann man sogar c in einer Umgebung beliebig vorgeben.

9.34. Bemerkung: Der Satz lässt sich erheblich verallgemeinern. Dazu sei $f : U \to \mathbb{R}^m$ eine stetig differenzierbare Funktion, und $U \subseteq \mathbb{R}^n$ offen mit $n > m$. Zur Abkürzung setzen wir $n = k + m$ und schreiben

$$f(x, y) = f(x_1, \ldots, x_k, y_1, \ldots, y_m)$$

Die Matrix

$$D_y f(x, y) = \begin{pmatrix} \frac{\partial}{\partial y_1} f_1(x) & \ldots & \frac{\partial}{\partial y_m} f_1(x) \\ \vdots & & \vdots \\ \frac{\partial}{\partial y_1} f_m(x) & \ldots & \frac{\partial}{\partial y_m} f_m(x) \end{pmatrix}$$

ist der rechte quadratische Teil von $Df(x)$. Sie ist die Ableitungsmatrix der Funktion

$$(y_{k+1}, \ldots, y_{k+m}) \mapsto f(x_1, \ldots, x_k, y_{k+1}, \ldots, y_{k+m}),$$

also quasi die partielle Ableitung nach y. Wenn $Mf(x)$ invertierbar ist, so gibt es eine lokale Auflösung g der Gleichung $f(x) = c$ nach diesen Variablen, also

$$f(x, y) = 0 \;\Leftrightarrow\; y = g(x),$$

und es gilt

$$Dg(y) = -M(x)^{-1} \cdot D_x f(x, y),$$

mit der analog definierten partiellen Ableitungsmatrix nach x. Im obigen Beispiel war $m = 1$, $k = n - 1$, und wir erhalten

$$\operatorname{grad} g(y) := -\frac{1}{\frac{\partial}{\partial x_{k+1}} f(x)} \operatorname{grad} f(x).$$

Ein Beispiel ist die lokale Paramtriesierung einer Kurve in $\mathbb{R}^3$ durch zwei Nebenbedingungen der Form

$$f_1(x) = c_1, \quad f_2(x) = c_2.$$

Wie in der vorigen Bemerkung erhält man sogar eine lokal eindeutige Abhängigkeit der Kurve von c_1 und c_2.

9.5 Extrema mit Nebenbedingungen

Wir haben bisher nur Extrema auf dem ganzen, offenen Definitionsgebiet einer Funktion f gesucht. Nun nehmen wir eine Nebenbedingung $g(x) = c$ hinzu.

9.35 Satz: *Sei $U \subseteq \mathbb{R}^n$ offen und $f : U \to \mathbb{R}$ eine differenzierbare Funktion. Sei weiter $g : U \to \mathbb{R}$ differenzierbar und $g(x) = c$ für ein $x \in U$. Wenn dann x ein lokales Extremum von f unter der Nebenbedingung $g(x) = c$, also auf der Menge*

$$N_c := \{t \in U : g(t) = c\}$$

ist, und wenn

$$\operatorname{grad} g(x) \neq 0$$

ist, so existiert ein $\lambda \in \mathbb{R}$, so dass

$$\operatorname{grad} f(x) = \lambda \operatorname{grad} g(x)$$

*ist. Man nennt diese Bedingung die **Lagrange**-Bedingung für lokale Extrema unter Nebenbedingungen.*

Beweis: Sei etwa

$$\frac{\partial}{\partial x_n} g(x) \neq 0.$$

Dann kann man $g(x) = c$ lokal nach x_n auflösen, also

$$g(t) = c \;\Leftrightarrow\; t_n = h(t_1, \ldots, t_{n-1})$$

in einer Umgebung von x mit einer entsprechenden Funktion h. Da x Extremum unter der Nebenbedingung $g(x) = c$ ist, folgt

$$0 = \frac{\partial}{\partial x_i} f(x_1, \ldots, x_{n-1}, h(x_1, \ldots, x_{n-1}))$$

für $i = 1, \ldots, n$. Dies bedeutet für $i = 1, \ldots, n-1$

$$0 = D_i f(x) + D_n f(x) D_i h(\tilde{t}). = D_i f(x) - D_n f(t) \frac{D_i g(t)}{D_n g(t)}.$$

Also

$$D_i f(x) = \frac{D_n f(t)}{D_n g(t)} D_i g(t).$$

Dasselbe gilt für $i = n$. □

9.36. Bemerkung: Der Satz lässt sich anschaulich sehr gut begründen, wenn man sich überlegt, das einem Maximum von f bezüglich N_c die Niveaulinie von f tangential zu der Niveaulinien N_c der Nebenbedingung g stehen muss.

9.37. Bemerkung: Aus der Bedingung über die Gradienten folgen n Gleichungen mit n Umbekannten. Eine zusätzliche Gleichung erhält man aus der Nebenbedingung.

9.38. Beispiel: Wir maximieren die Größe eines Rechtecks mit gegebenem Umfang, also

$$f(a, b) = ab$$

unter der Nebenbedingung

$$g(a, b) = 2(a + b) = c.$$

Die Definitionsmenge ist hierbei

$$U = \{(a, b) : a, b > 0\}.$$

Aus dem Lagrange-Ansatz

$$\text{grad } f(a, b) = (b, a) = \lambda \,\text{grad } g(a, b) = (2\lambda, 2\lambda)$$

folgt unmittelbar, dass $a = b$ gelten muss. Mit Hilfe der Nebenbedingung $2a + 2b = c$ lassen sich dann a und b berechnen.

9.39. Bemerkung: Wir haben einfach angenommen, dass ein Maximum existiert. Das kann man dadurch begründen, dass die Menge

$$K = \{(a, b) : 2(a + b) = c,\ a, b \geq 0\}$$

kompakt ist. Es muss also ein Maximum auf dieser Menge geben. Wenn $a = 0$ oder $b = 0$ ist, so liegt offenbar ein Minimum vor, so dass das Maximum in U liegen muss.

9.40. Bemerkung: Man kann solche Aufgaben auch durch Auflösen der Nebenbedingung nach einer Variablen und Einsetzen in die zu maximierende Funktion lösen. Aber oft führt das auf Fallunterscheidungen und komplexere Rechnungen. Im obigen Beispiel wäre

$$h(b) = \left(\frac{c}{2} - b\right) \cdot b$$

auf $]0, c[$ zu maximieren.

9.41. Bemerkung: Der Satz lässt sich auf mehrere Nebendingungen

$$g_1(x) = c, \quad \ldots, \quad g_k(x) = c$$

erweitern. In diesem Fall müssen lediglich die Gradienten

$$\text{grad } g_1(x), \ldots, \text{grad } g_k(x)$$

linear unabhängig sein. Als notwendige Bedingung erhält man

$$\text{grad } f(x) = \sum_{j=1}^{k} \lambda_j \text{ grad } g_k(x).$$

9.6 Die Kurvenlänge

Eine Kurve ist eine stetige Funktion

$$\gamma : [a, b] \to \mathbb{R}^m$$

Die Länge einer Kurve kann durch die Länge eines Streckenzuges mit Ecken auf der Kurve approximiert werden. Dies gibt Anlass zur folgenden Definition.

9.42. Definition: Für eine Unterteilung $T = (t_i)_{1 \leq i \leq n}$

$$a = t_0 < t_1 < \ldots < t_n = b$$

ist die Länge des Streckenzugs mit Ecken $\gamma(t_0), \ldots, \gamma(t_n)$ durch

$$L_T = \sum_{k=1}^{n} \|\gamma(t_k) - \gamma(t_{k-1})\|$$

gegeben. Die Länge der Kurve g ist dann durch

$$L(\gamma) := \sup\{L_T : T \text{ Unterteilung}\}$$

definiert. γ heißt **rektifizierbar**, wenn

$$L(\gamma) < \infty$$

ist.

9.43. Bemerkung: Ähnlich wie beim Riemann-Integral liegt es nahe, die Länge als Grenzwert von immer feineren Umterteilungen zu definieren. Für alle Folgen von Unterteilungen, deren Feinheit gegen 0 konvergiert, müssten die Längen dann denselben Grenzwert haben. Eine solche Definition ist äquivalent, und wir werden sie beim Beweis des folgenden Satzes verwenden.

9.44 Satz: *Sei $\gamma : [a, b] \to \mathbb{R}^m$ eine Kurve, die in $[a, b]$ stetig differenzierbar ist, einschließlich einseitiger Ableitungen in a und b. Dann gilt*

$$L(\gamma) = \int_a^b \|\gamma'(t)\| \, dt.$$

Insbesondere ist γ rektifizierbar.

Beweis: Für ein kleines Kurvenstück $\gamma : [t_{k-1}, t_k] \to \mathbb{R}^m$ verwenden wir die Hilfskurve

$$h(t) = \gamma(t) - (\gamma(t_{k-1}) + (t - t_{k-1})\gamma'(t_{k-1})), \qquad t \in [t_{k-1}, t_k].$$

Sei $\epsilon > 0$. γ' ist in $[a,b]$ gleichmäßig stetig. Für genügend feine Unterteilungen folgt also aus Satz 9.24 wegen $h'(t) = \gamma'(t) - \gamma'(t_{k-1})$

$$\begin{aligned} | \, \|\gamma(t_k) - \gamma(t_{k-1})\| - (t_k - t_{k-1})\|\gamma'(t_{k-1})\| \, | & \\ &\leq \|\gamma(t_k) - \gamma(t_{k-1}) - (t_k - t_{k-1})\gamma'(t_{k-1})\| \\ &= \|h(t_k) - h(t_{k-1})\| \\ &\leq (t_k - t_{k-1}) \sup_{t \in [t_{k-1}, t_k]} \|\gamma'(t) - \gamma'(t_{k-1})\| \\ &< (t_k - t_{k-1})\epsilon. \end{aligned}$$

Da $\|\gamma'\|$ Riemann-integrierbar ist, gilt für genügend feine Unterteilungen

$$\left| \sum_{k=1}^n (t_k - t_{k-1})\gamma'(t_{k-1}) - \int_a^b \|\gamma'(t)\| \, dt \right| < \epsilon$$

Also für alle genügend feinen Unterteilungen

$$\left| \sum_{k=1}^n \|\gamma(t_k) - \gamma(t_{k-1})\| - \int_0^1 \|\gamma'(t)\| \, dt \right| \leq 2(b-a)\epsilon.$$

Es folgt die Behauptung. □

9.45. Beispiel: (1) Die Strecke $S_{v,w}$ kann durch

$$\gamma(t) = v + t(w - v), \qquad t \in [0,1]$$

parametrisiert werden. Es ergibt sich wegen $\gamma'(t) = w - v$

$$L(\gamma) = \int_0^1 \|w - v\| \, dt = \|w - v\|$$

wie erwartet. Natürlich kann man die Strecke auch anders parametrisieren, etwa durch

$$\gamma(t) = v + t^2(w - v), \qquad t \in [0,1]$$

Dann erhält man

$$L(\gamma) = \int_0^1 \|2t(v - w)\| \, dt = \|v - w\| \int_0^1 2|t| \, dt = \|v - w\|.$$

(2) Der Funktionsgraph einer Funktion $f : [a,b] \to \mathbb{R}$ ist eine Kurve, die durch

$$\gamma(t) = (t, f(t)) \in \mathbb{R}^2$$

parametrisiert ist. Für die Länge des Graphen gilt also

$$L(\gamma) = \int_a^b \sqrt{1 + f'(t)^2}\, dt,$$

falls f stetig differenzierbar ist. Für die Parabel

$$f(t) = t^2, \qquad t \in [0, 1]$$

gilt beispielsweise

$$L(\gamma) = \int_0^1 \sqrt{1 + 4t^2}\, dt = \frac{\sinh^{-1}(2) + 2\sqrt{5}}{2} = 1.478942857544598\ldots$$

nach Formelsammlung. Dabei bezeichnet $\sinh : \mathbb{R} \to \mathbb{R}$ die Abbildung

$$\sinh(x) = \frac{e^x - e^{-x}}{2},$$

die man **Sinus Hyperbolicus** nennt.

9.7 Vektorfelder

9.46 Satz: *Sei $f : [a, b] \times I \to \mathbb{R}$ eine stetige Funktion, die nach der zweiten Variablen stetig partiell differenzierbar sein, I ein offenes Intervall in $\mathbb{R}$. Dann gilt für alle $c \in I$*

$$\frac{d}{dc} \int_a^b f(x, c)\, dx = \int_a^b \frac{\partial}{\partial c} f(x, c)\, dx.$$

Beweis: Es gilt

$$\left| \frac{1}{h} \left(\int_a^b f(x, c+h)\, dx - \int_a^b f(x, c)\, dx \right) - \int_a^b \frac{\partial}{\partial c} f(x, c)\, dx \right|$$
$$\leq \int_a^b \left| \frac{f(x, c+h) - f(x, c)}{h} - \frac{\partial}{\partial c} f(x, c) \right| dx$$

Außerdem

$$\frac{f(x, c+h) - f(x, c)}{h} - \frac{\partial}{\partial c} f(x, c) = \frac{\partial}{\partial c} f(x, \xi_{x,c}) - \frac{\partial}{\partial c} f(x, c)$$

für $|\xi_{x,c} - c| < |h|$. Die partielle Ableitung ist stetig auf der kompakten Menge $[a, b] \times [c - h, c + h]$, also gleichmäig stetig. Daher wird der obige Ausdruck für h klein genug kleiner als ϵ für alle $x \in [a, b]$. Es folgt die Behauptung. $\square$

9.47. Definition: Ein **Vektorfeld** ist eine Abbildung

$$f : U \to \mathbb{R}^n$$

definiert auf einer offenen Menge $U \subseteq \mathbb{R}^n$. Eine Funktion

$$F : U \to \mathbb{R}$$

heißt **Stammfunktion** von f, wenn

$$\text{grad } F(x) = f(x)$$

für alle $x \in U$ gilt.

9.48. Beispiel: Man kann Stammfunktionen durch Integration berechnen. Sei etwa

$$f(x,y,z) = \left(1 - \frac{z}{x^2}, 2yz, y^2 + \frac{1}{x}\right).$$

Für eine Stammfunktion $F(x,y,z)$ muss dann gelten

$$\frac{d}{dx}F(x,y,z) = f_1(x,y,z) = 1 - \frac{z}{x^2},$$

woraus

$$F(x,y,z) = x + \frac{z}{x} + c(y,z)$$

folgt. Man beachte, dass die Integrationskonstante von den anderen Variablen abhängen kann. Differenzieren nach y ergibt weiter

$$2yz = f_2(x,y,z) = \frac{d}{dy}F(x,y,z) = \frac{d}{dy}c(y,z),$$

also

$$c(y,z) = y^2 z + c(z).$$

Das ergibt

$$F(x,y,z) = x + \frac{z}{x} + y^2 z + c(z)$$

Differenzieren nach z ergibt

$$y^2 + \frac{1}{x} = f_3(x,y,z) = \frac{d}{dz}F(x,y,z) = y^2 + \frac{1}{x} + c'(z)$$

woraus $c'(z) = 0$ folgt, also

$$F(x,y,z) = 1 + \frac{z}{x} + y^2 z + C.$$

Aber nicht jedes Vektorfeld hat eine Stammfunktion.

9.49 Satz: *Sei $f : U \to \mathbb{R}^n$ ein stetig differenzierbares Vektorfeld auf einer sternförmigen, offenen Menge U. Dann besitzt f genau dann eine Stammfunktion, wenn die* ***Integrationsbedingungen***

$$\frac{\partial}{\partial x_i}f_j(x) = \frac{\partial}{\partial x_j}f_i(x)$$

für alle $1 \leq i,j \leq n$ und $x \in U$ erfüllt sind.

Zum Beweis benötigen wir das Integral eines Vektorfeldes längs eines Weges.

9.50. Definition: Sei $f : U \to \mathbb{R}^n$ ein stetiges Vektorfeld, $U \subseteq \mathbb{R}^n$ offen, und $\gamma : [a, b] \to U$ ein stetig differenzierbarer Weg in U. Dann definieren wir

$$\int_\gamma f := \int_a^b \langle f(\gamma(t)), \gamma'(t) \rangle \, dt.$$

Falls γ stückweise aus stetig differenzierbaren Wegen besteht, so definieren wir dieses Integral als die Summe der Integrale über alle diese Teilwege.

9.51. Bemerkung: Wenn γ aus den hintereinander gehängten Wegen γ_1 und γ_2 besteht, so folgt

$$\int_\gamma f = \int_{\gamma_1} f + \int_{\gamma_2} f.$$

Wenn γ_1 und γ_2 zwei Wege mit gleichem Start- und Endpunkt sind, so kann man γ_2 umkehren und hinter γ_1 hängen. Es entsteht ein geschlossener Weg γ, und

$$\int_\gamma f = \int_{\gamma_1} f - \int_{\gamma_2} f.$$

9.52. Bemerkung: Wenn f eine Stammfunktion F hat, und γ von $\tilde{x}$ nach x läuft, so gilt

$$\begin{aligned} F(x) - F(\tilde{x}) &= [F(\gamma(t))]_a^b \\ &= \int_a^b \frac{d}{dt} F(\gamma(t)) \, dt \\ &= \int_a^b \operatorname{grad} F(\gamma(t)) \cdot \gamma'(t) \, dt \\ &= \int_a^b \langle f(\gamma(t)), \gamma'(t) \rangle \, dt \\ &= \int_\gamma f. \end{aligned}$$

Also kann man F auch auf diese Weise berechnen. Außerdem folgt, dass in diesem Fall das Integral nicht vom Weg γ abhängt, der die Punkte verbindet. Für jeden geschlossenen Weg gilt dann

$$\int_\gamma f = 0.$$

Beweis: Wenn f eine Stammfunktion hat, so gelten die Integrationsbedingungen wegen

$$\frac{\partial}{\partial x_i} f_j(x) = \frac{\partial^2}{\partial x_i x_j} F(x) = \frac{\partial^2}{\partial x_j x_i} F(x) = \frac{\partial}{\partial x_j} f_i(x)$$

Nehmen wir an, dass umgekehrt die Integrationsbedingungen gelten. Sei U sternförmig bezüglich $\tilde{x}$. Wir nehmen zur Vereinfachung der Schreibweise an, dass $\tilde{x} = 0$ ist. Dann definieren wir

$$F(x) := \int_0^1 \langle f(tx), x \rangle \, dt.$$

Dies ist gleich $\int_\gamma f$, wobei γ der lineare Weg von 0 nach x ist, also

$$\gamma(t) = tx.$$

Wir erhalten

$$F(x) = \sum_{i=1}^{n} \int_0^1 f_i(tx)x_i \, dt.$$

Zur Berechnung der partiellen Ableitung von F nach x_j differenzieren wir unter dem Integral. Für $j \neq i$ erhält man wegen der Integralbedingung

$$\begin{aligned}
\frac{\partial}{\partial x_j} \int_0^1 f_i(tx)x_i \, dt &= \int_0^1 x_i \frac{d}{dx_j} \left(f_i(tx) \right) \, dt \\
&= \int_0^1 x_i t \frac{\partial}{\partial x_j} f_i(tx) \, dt \\
&= \int_0^1 t x_i \frac{\partial}{\partial x_i} f_j(tx) \, dt
\end{aligned}$$

Für $j = i$ erhält man

$$\frac{\partial}{\partial x_j} \int_0^1 f_j(tx)x_j \, dt = \int_0^1 t x_j \frac{\partial}{\partial x_j} f_j(tx) \, dt + \int_0^1 f_j(tx) \, dt.$$

Also, mittels partieller Integration

$$\begin{aligned}
\frac{\partial}{\partial x_j} F(x) &= \int_0^1 t \left\langle \operatorname{grad} f_j(tx), x \right\rangle dt + \int_0^1 f_j(tx) \, dt \\
&= \int_0^1 t \frac{d}{dt} \left(f_j(tx) \right) \, dt + \int_0^1 f_j(tx) \, dt \\
&= \left[t f_j(tx) \right]_0^1 \\
&= f_j(x).
\end{aligned}$$

Es folgt, dass F eine Stammfunktion von f ist. $\square$

9.53. Bemerkung: Damit kann man lokal in einer Umgebung jeder Funktion, die den Integralbedingungen genügt, eine Stammfunktion erhalten. Es ist aber keineswegs gesagt, dass es eine globale Stammfunktion auf U geben muss.

9.54 Satz: *Sei $f : U \to \mathbb{R}^n$ ein stetiges Vektorfeld. Genau dann, wenn für jeden geschlossenen Weg*

$$\int_\gamma f = 0$$

gilt, hat f auf U eine globale Stammfunktion.

Beweis: Die eine Richtung wurde schon gezeigt. Wir nehmen nun an, dass das Integral über jeden geschlossenen Weg gleich 0 ist. Es folgt, dass das Integral unabhängig vom Weg ist. Wir können daher auf jeder Zusammenhangskomponenten von U

$$F(x) = \int_\gamma f$$

definieren, wobei γ ein Weg von einem festen $\tilde{x}$ nach x ist. In jeder Kugel mit Mittelpunkt z gilt daher

$$F(x) = F(x) - F(z) + F(z) - F(\tilde{x}) = \int_{\gamma_1} f + \int_{\gamma_2} f,$$

wobei γ_1 der lineare Wege zwischen z und x sei, und γ_2 ein Weg von $\tilde{x}$ nach z. Der zweite Summand ist konstant für x, und nach dem vorigen Beweis ist der erste Summand eine Stammfunktion von f. □

9.55. Definition: Zwei Wege γ_0 und γ_1, definiert auf einem Parameterinterval $[a, b]$, in einer offenen Menge $U \subseteq \mathbb{R}^n$ mit gleichem Start- und Endpunkt heißen **homotop**, wenn es eine stetige Abbildung

$$\gamma : [a, b]x[0, 1] \to U$$

gibt, so dass die Wege $\gamma_s(t) = \gamma(t, s)$ alle denselben Anfangs- und Endpunkt haben. Ein geschlossener Weg heißt **nullhomotop**, wenn er homotop zu dem konstanten Weg ist, der auf dem Anfangspunkt stehen bleibt. Eine zusammenhängende Menge U heißt **einfach zusammenhängend**, wenn jeder geschlossene Weg in U nullhomotop ist.

9.56 Satz: *Wenn das stetige Vektorfeld $f : U \to \mathbb{R}^n$, $U \subseteq \mathbb{R}^n$ offen, überall lokal eine Stammfunktion hat, dann stimmen die Integrale längs homotopen Wegen überein, und die Integrale längs nullhomotopen Wegen sind gleich* 0.

Beweis: Da f lokal integrierbar ist, existieren endlich viele Kugeln $B_\epsilon(x_k)$, so dass f auf $B_{3\epsilon}(x_k)$ eine Stammfunktion hat, und diese Kugeln sämtliche Wege γ_s überdecken. Aufgrund von gleichmäßige Stetigkeit kann man sicherstellen, dass es eine Unterteilung

$$0 \leq s_0 < \ldots < s_m = 1$$

und eine Unterteilung

$$a \leq t_0 < \ldots < t_l = b$$

gibt, so dass

$$|\gamma_{s_i}(t_j) - \gamma_{s_i}(t_{j+1})| < \epsilon$$

und

$$|\gamma_{s_i}(t_j) - \gamma_{s_{i+1}}(t_j)| < \epsilon$$

für alle i, j ist. Mit Hilfe von lokalen Stammfunktionen F_j gilt dann

$$\begin{aligned}\int_{\gamma_{s_j}} f &= \sum_{j=1}^{l} F_j(\gamma_{s_i}(t_j)) - F_j(\gamma_{s_i}(t_{j-1})) \\ &= F_l(b) + \sum_{j=1}^{l-1}(F_j - F_{j+1})(\gamma_{s_i}(t_j)) - F_1(a).\end{aligned}$$

Dasselbe gilt für $\gamma_{s_{j+1}}$, und zwar mit denselben lokalen Stammfunktionen. Beachtet man, dass die Differenz zweier Stammfunktionen konstant sein muss, so folgt

$$\int_{\gamma_{s_j}} f = \int_{\gamma_{s_{j+1}}} f$$

und damit die Behauptung. $\square$

9.57. Bemerkung: Als Folgerung erhält man, dass lokal integrierbare Vektorfelder auf einfach zusammenhängenden Mengen global integrierbar sind.

9.58. Beispiel: Nicht einfach zusammenhängend ist beispielsweise

$$U = \mathbb{R}^2 \setminus \{0\}$$

Denn das Vektorfeld

$$f(x, y) = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right)$$

erfüllt die Integrationsbedingung, ist also lokal integrierbar. Integriert man es aber längs dem Einheitskreis, so ist das Integral verschieden von 0. Also ist f nicht global integrierbar.

Kapitel 10

Integralrechung mehrerer Variablen

In diesem Kapitel erweitern wir das Riemann-Integral auf Funktionen mehrerer Veränderlicher. Zunächst geben wir einen intuitiven Zugang, mit dem schon die meisten Integrale berechnet werden können. Danach erst wird das Riemann-Integral exakt begründet.

10.1 Beispiele

Die folgenden Beispiele zeigen, dass die Idee des Hauptsatzes der Differential- und Integralrechung in nahe liegender Weise auf Integrale in der Ebene und im Raum übertragen werden kann. Wir werden allerdings die folgenden Berechnungen erst exakt begründen können, wenn wir die Definition des Riemann-Integrals in den kommenden Abschnitten festgelegt haben.

10.1. Beispiel: Um das Volumen der Kugel mit Radius $R > 0$ im $\mathbb{R}^3$ zu berechnen, stellen wir uns Schnitte durch die Kugel in Höhe $z \in [-1, 1]$ vor und bezeichnen das Volumen der Menge

$$M_z := \{(x, y, t) : x^2 + y^2 + t^2 \leq R^2 \text{ und } -1 \leq t \leq z\}$$

unterhalb dieses Schnitte als $V(z)$. Die Schnittfläche ist

$$A(z) = \pi(R^2 - z^2).$$

Wir stellen uns nun vor, dass z sich um Δz verändert. Das Volumen verändert sich dabei in erster Näherung um

$$\Delta V \approx A(z)\, \Delta z.$$

Tatsächlich gilt mit Hilfe anschaulicher Überlegungen für $z \geq 0$ und $\Delta z > 0$

$$A(z)\, \Delta z \leq \Delta V \leq A(z + \Delta z)\, \Delta z$$

Daher überzeugen wir uns, dass

$$V'(z) = \lim_{\Delta z \to 0} \frac{\Delta V}{\Delta z} = A(z)$$

gilt. Es folgt für das Gesamtvolumen aus dem Hauptsatz der Differential- und Integralrechung

$$V(R) = V(R) - V(-R) = \int_{-R}^{R} V'(z)\, dz = \int_{-R}^{R} A(z)\, dz = \frac{4\pi R^3}{3}.$$

Dies ist die bekannte Formel für das Volumen der Kugel mit Radius R.

10.2. Bemerkung: Nach dem gleichen Muster können wir Schnitte durch andere Körper M anlegen. Wenn $A(z)$ die Schnittfäche bezeichnet, so gilt für das Volumen von

$$M_{a,b} = \{(x, y, t) \in M : a \leq t \leq b\}$$

stets

$$\operatorname{vol}(M_{a,b}) = \int_a^b A(z)\, dz.$$

Das Volumen hängt also nur von der Schnittfläche in Höhe z ab. Dies ist das Prinzip von **Cavalieri**. Natürlich gilt dasselbe auch für Schnitte in Richtung der anderen Koordinatenachsen, also mit festem x oder festem y.

10.3. Beispiel: Auch das Integral einer Funktion $f : M \to \mathbb{R}$ kann mit Hilfe dieser Idee berechnet werden, wenn es als Volumen unter dem Funktionsgraphen, also als Volumen von

$$K = \{(x, y, z) : (x, y) \in M,\ 0 \leq z \leq f(x, y)\}$$

versteht. Für einen Quader

$$M = [a, b] \times [c, d]$$

erhält man

$$\int_M f := \operatorname{vol}(K) = \int_a^b \left(\int_c^d f(x, y)\, dy \right) dx,$$

wobei wir Schnitte mit konstantem x verwendet haben. Schneidet man mit konstantem y, so erhält man

$$\int_M f := \operatorname{vol}(K) = \int_c^d \left(\int_a^b f(x, y)\, dx \right) dy.$$

Dass diese Intergrale gleich sind, bezeichnet man als Satz von **Fubini**. Wenn die Menge M unregelmäßig ist, so muss man die Schnitte

$$M_x = \{y : (x, y) \in M\}$$

berechnen, und das innere Integral über diese Schnitte ausrechen. Also

$$\int_M f = \int_a^b \left(\int_{M_x} f(x, y)\, dy \right) dx.$$

Dabei ist a der minimale Wert von x, so dass M_x nicht leer ist, und b der maximale.

10.4. Beispiel: Wenn $f : [a,b] \to [0,\infty[$ eine Riemann-integrierbare Funktion ist, so kann man den Graphen von f um die x-Achse drehen und erhält den Drehkörper

$$M = \{(x,y,z) : y^2 + z^2 \leq f(x)^2,\ a \leq x \leq b\}.$$

Das Volumen von M ist

$$\mathrm{vol}\,(M) = \pi \int_a^b f(x)^2\,dx,$$

da die Schnittflächen für festes x Kreise mit Radius $f(x)$ sind. So entsteht beispielsweise durch $f(x) = cx$ ein Kegelstumpf mit Volumen

$$\pi \int_a^b (cx)^2\,dx = \frac{\pi c^2 (b^3 - a^3)}{3}.$$

Speziell für $a = 0$ entsteht ein Kegel $K(r,h)$ mit Grundradius $r = cb$ und Höhe b. Also

$$\mathrm{vol}\,(K) = \frac{\pi c^2 b^3}{3} = \frac{\pi r^2 h}{3},$$

was der bekannten Formel „Grundfläche mal Höhe durch 3" entspricht. Allgemeiner hat ein Kegel K mit Grundfläche A und Höhe h in der Höhe $0 \leq t \leq h$ die Schnittfläche der Größe

$$A(t) = \left(\frac{h-t}{h}\right)^2 A,$$

da die Fläche A mit dem Faktor $(h-t)/h$ skaliert wird. Man berechnet

$$\mathrm{vol}\,(K) = \int_0^h \left(\frac{h-t}{h}\right)^2 A\,dh = \frac{Ah}{3},$$

was der bekannten Formel entspricht.

10.5. Beispiel: Eine Funktion $f : \mathbb{R}^2 \to \mathbb{R}$ vom Typ

$$f(x,y) = g(r), \qquad r = \sqrt{x^2 + y^2}$$

ist auf Kreisen konstant und heißt deswegen **rotationsinvariant**. Wir wollen

$$I(R) = \int_{D_R} f$$

berechnen, wobei D_R die Kreisscheibe mit Radius R sei. Analog zu den obigen Überlegungen erhält man

$$I'(r) = 2\pi r g(r).$$

Also

$$\int_{D_R} f = 2\pi \int_0^R r g(r)\,dr.$$

Die Fläche des Kreises πR^2 erhält man mit $g(r) = 1$. Ein weiteres Beispiel ist $g(r) = e^{-r^2}$, wo man

$$\int_{K_R} f = 2\pi \int_0^R re^{-r^2}\,dr = \pi(1 - e^{-R^2})$$

erhält. Mit $R \to \infty$ erhalten wir das uneigentliche Integral

$$\int_{\mathbb{R}^2} f = \pi.$$

Andererseits gilt für das Rechteck $Q_R = [-R, R] \times [-R, R]$

$$\int_{Q_R} f = \int_{-R}^R \left(\int_{-R}^R e^{-x^2-y^2}\,dx \right) dy = \left(\int_{-R}^R e^{-x^2}\,dx \right)^2.$$

Mit $R \to \infty$ folgt

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}.$$

Dies ist eine bekannte Formel von Gauß.

10.6. Beispiel: Geht man etwas weiter, so kann man mit der gleichen Idee auch Oberflächen berechnen. Bezeichnen wir die Oberfläche der Kugel mit Radius R im $\mathbb{R}^3$ als $O(R)$ und das Volumen als $V(R)$, so überlegt man sich ganz analog

$$O(R) = \lim_{\Delta R \to 0} \frac{\Delta V}{\Delta R} = V'(R) = 4\pi R^2.$$

Dies stimmt mit der bekannten Formel überein.

10.7. Beispiel: Mit der gleichen Idee kann Oberflächen von Drehkörpern berechnen. Der zur Funktion $f : [a, x] \to]0, \infty[$ gehörige Drehkörper habe die Oberfläche $O(x)$. Dann ist der Zuwachs angenähert gleich dem Umfang mal dem Längenzuwachs des Graphen. Daraus ergibt sich

$$O'(x) = 2\pi f(x)\sqrt{1 + f'(x)^2}.$$

Im Beispiel der Kugel erhalten wir für die Oberfläche O_R der Kugel mit Radius R mit $f(x) = R\sqrt{1-x^2}$

$$O_R = 2\pi \int_{-R}^R R\sqrt{1-x^2}\sqrt{1 + \frac{x^2}{1-x^2}}\,dx = 4\pi R^2.$$

Also wieder das gleicher Resultat.

10.8. Beispiel: Auch die von dem Fahrstrahl einer Spirale überfahrene Fläche kann mit derselben Idee berechnet werden. Sei dazu etwa $r : [0, 2\pi] \to]0, \infty[$ der Radius beim Winkel r und die Spirale durch

$$\gamma(t) = r(t)(\cos(t), \sin(t))$$

gegeben. Der Fahrstrahl ist der Strahl von 0 nach $\gamma(t)$. Er überstreicht in der Zeit Δt Näherungsweise einen Kreissektor der Größe

$$\Delta A = \frac{1}{2}\gamma(t)^2\,\Delta t.$$

Also ergibt sich für die Gesamtfläche nach einem Umlauf

$$A = \frac{1}{2}\int_0^{2\pi} \gamma(t)^2\,dt.$$

Beim Kreis ist Beispielsweise $\gamma(t) = R$, so dass man wieder $A = \pi R^2$ erhält.

10.2 Treppenfunktionen

Wir verwenden denselben Zugang wie im eindimensionalen Fall. Allerdings ist es etwas umständlicher, Treppenfunktionen zu definieren.

10.9. Definition: (1) Das Produkt von halboffenen Intervallen $I_i = [a_i, b_i[$, $i = 1, \ldots, n$ nennen wir halboffenen **Quader** im $\mathbb{R}^n$, also

$$Q = I_1 \times \ldots \times I_n.$$

Wir definieren den **Inhalt** von diesem halboffenen Quader durch

$$\mu(Q) := \prod_{i=1}^{n}(b_i - a_i).$$

(2) Die **charakteristische Funktion** einer Menge Q ist die Funktion

$$\mathbf{1}_Q(x) := \begin{cases} 1, & x \in Q, \\ 0, & x \notin Q. \end{cases}$$

Funktionen der Form $\lambda\mathbf{1}_Q$ mit halboffenen Quadern $Q \subset \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ nennen wir einfache Treppenfunktionen. Summen von einfachen Treppenfunktionen

$$T = \sum_{k=1}^{m} \lambda_k \mathbf{1}_{Q_k}$$

nennen wir **Treppenfunktionen**.

(3) Wir definieren das Integral von solchen Treppenfunktionen durch

$$\int T := \sum_{k=1}^{m} \lambda_k \mu(Q_k).$$

10.10. Bemerkung: Man beachte, dass Treppenfunktionen einen beschränkten Träger haben. Außerhalb sind sie identisch 0. Der folgende Satz wird ergeben, dass die obige Definition mit dem eindimensionalen Integral von Treppenfunktionen übereinstimmt.

10.11 Satz: *Für das Integral von Treppenfunktionen gilt der **Satz von Fubini**. Das heißt,*

$$\int T = \int_{a_1}^{b_1}\left(\int_{a_2}^{b_2}\left(\ldots\left(\int_{a_n}^{b_n} T(x_1, \ldots, x_n)\,dx_n\right)\ldots\right)\right).$$

Dabei existieren alle diese Integrale im Riemannschen Sinn, und es kommt nicht auf die Reihenfolge der Integration an. Die Integrationsgrenzen sind so gewählt, dass T außerhalb des Quaders

$$[a_1, b_1] \times [a_n, b_n]$$

identisch 0 *ist.*

Beweis: Es genügt, den Satz für einfache Treppenfunktionen nachzurechnen. □

10.12. Bemerkung: Aufgrund dieses Satzes schreibt man

$$\int T = \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} T(x_1, \dots, x_n)\, dx_1 \dots dx_n.$$

Man lässt also die oben aufgeführten Klammern weg. Natürlich muss die Reihenfolge der Grenzen der Reihenfolge der Integration entsprechen. Also zum Beispiel

$$\int_a^b \int_c^d T(x,y)\, dy\, dx = \int_c^d \int_a^b T(x,y)\, dx\, dy.$$

Wobei T auf $[a,b] \times [c,d]$ definiert ist.

10.13 Satz: *Das Integral von Treppenfunktionen ist linear und monoton.*

10.3 Das Riemann-Integral

Wir können nun genau wie im eindimensionalen Fall das Riemann-Integral definieren. Wieder müssen wir uns zunächst auf beschränkte Funktionen beschränken, die zudem noch einen beschränkten Träger haben müssen.

10.14. Definition: Eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ heißt **Riemann-integrierbar**, wenn sie beschränkt ist und einen beschränkten Träger hat, und wenn Unterintegral und Oberintegral übereinstimmen. Das heißt

$$\sup\{\int T : T \leq f,\, T \text{ Treppenfunktion}\}$$
$$= \inf\{\int T : f \leq t,\, T \text{ Treppenfunktion}\}.$$

In diesem Fall definiert man diese Größe als das Riemann-Integral von f.

10.15. Bemerkung: Wenn $f : \mathbb{R}^n \to \mathbb{R}$ keinen beschränkten Träger hat, so kann man f auf einen Quader

$$Q = [a_1, b_1] \times \ldots \times [a_n, b_n]$$

einschränken und dort integrieren. Man integriert dann also die Funktion $f \times \mathbf{1}_Q$. Wir verwenden die nahe liegende Schreibweise

$$\int_Q f = \int_{a_1}^{b_1} \dots \int_{a_n}^{b_n} f(x_1, \dots, x_n)\, dx_1 \dots dx_n.$$

10.16. Bemerkung: Das eindimensionale Riemann-Integral stimmt damit überein, und es gilt

$$\int_{[a,b]} f = \int_a^b f(x)\,dx$$

für $a < b$. Für $b < a$ ist das aber nicht richtig. Dies liegt daran, dass Teilmengen des $\mathbb{R}^n$ keine Orientierung haben.

10.17 Satz: *Die Summe, das Vielfache, das Maximum, das Minimum, der Betrag und das Produkt von Riemann-integrierbaren Funktionen sind Riemann-integrierbar. Das Riemann-Integral ist linear und monoton.*

Beweis: Der Beweis ist genau wie im eindimensionalen Fall. □

10.18 Satz: *Wenn eine Funktion $f : \mathbb{R}^n \to \mathbb{R}$ Riemann-integrierbar ist und alle Integrale im Satz von Fubini existieren, so gilt der Satz von Fubini für f.*

10.19. Definition: Wir übertragen die Definition der gleichmäßigen Stetigkeit wörtlich auf den Fall von mehreren Variablen.

10.20 Satz: *Eine stetige Funktion, eingeschränkt auf einen kompakten Quader, ist gleichmäßig stetig und Riemann-integrierbar.*

10.4 Riemann-messbare Mengen

Wir möchten eine stetige Funktion f außer auf Quader auch auf andere Mengen M einschränken können. Dies entspricht der Integration von $f \cdot \mathbf{1}_M$.

10.21. Definition: Eine Teilmenge $M \subseteq \mathbb{R}^n$ heißt **Riemann-messbar**, wenn $\mathbf{1}_M$ Riemann-integrierbar ist. Man setzt

$$\mu(M) = \int \mathbf{1}_M(x)\,dx.$$

10.22. Beispiel: Wir wissen schon, dass Quader Riemann-messbar sind.

10.23 Satz: *Endliche Vereinigungen, Schnitte und Differenzen von Riemann-messbaren Funktionen sind Riemann-messbar.*

10.24 Satz: *Sei $K \subset \mathbb{R}^n$ kompakt und Riemann-messbar, $K \subseteq U$ für eine offene Menge $U \subseteq \mathbb{R}^n$, sowie*

$$g : U \to \mathbb{R}$$

stetig partiell differenzierbar, so dass grad $g(x)$ *nur für endlich viele $x \in K$ gleich 0 wird. Dann ist*

$$M := \{x \in K : g(x) \leq 0\}$$

kompakt und Riemann-messbar.

10.25. Beispiel: Der Einheitskreis mit Rand ist Riemann-messbar. Man kann die Fläche mit dem Satz von Fubini zu π berechnen. Mit Hilfe des Komplementes zeigt man ebenso, dass er Einheitskreis ohne Rand Riemann-messbar ist.

10.26. Beispiel: Ein **Rotationskörper** $M \subset \mathbb{R}^3$ ist ein Menge

$$M := \{(x, y, z) \in \mathbb{R}^3 : y^2 + z^2 \leq f(x)^2,\ a \leq x \leq b\}$$

mit einer Funktion $f : [a, b] \to \mathbb{R}$. Zumindest für stetig differenzierbare Funktionen f ist der obige Satz anwendbar und wir erhalten

$$\mu(M) = \pi \int_a^b f(x)^2 \, dx.$$

In der Tat ist M auch für Riemann-integrierbare Funktionen f Riemann-messbar.

10.5 Der Transformationssatz

Dieser Satz erlaubt die Berechnung von Integralen in Kreisen, Kugeln und Zylindern. Der Beweis liegt außerhalb dieser Vorlesung.

10.27 Satz: *Sei $M \subset \mathbb{R}^n$ Riemann-messbar und kompakt, $U \subseteq \mathbb{R}^n$ offen und $M \subset U$. Sei $g : U \to \mathbb{R}^n$ stetig differenzierbar, injektiv und*

$$\det Dg(x) \neq 0$$

für alle $x \in U$. Dann ist $g(M)$ Riemann-messbar und für jede Riemann-messbare Funktion $f : g(M) \to \mathbb{R}$ gilt

$$\int_{f(M)} f(y) \, dy = \int_M f(g(x)) |\det(Dg(x))| \, dx.$$

*Diese Formel heißt **Transformationsformel** für Integrale.*

10.28. Beispiel: (1) Der Satz erinnert an die Substitutionsregel. In diesem Fall ist $M = [a, b]$. Man muss aber beachten, dass Riemann-Integrale im $\mathbb{R}^n$ keine Orientierung haben. Außerdem stimmt die eindimensionale Formel auch für nicht streng monotone g.

(2) Die Anwendungen auf Kreise

$$K_R = D_R(0) \subseteq \mathbb{R}^2$$

sind nicht wörtlich möglich, sondern man muss zusätzliche Argumente verwenden. Man erhält die Formel

$$\int_{K_R} f(y) \, dy = \int_0^R \int_0^{2\pi} r f(r\cos(t), r\sin(t)) \, dt \, dr.$$

Um die Kreisfläche zu berechnen, setzt man $f = 1$ und erhält

$$\mu(K_R) = \pi R.$$

Falls $f(y)$ nur von $\|y\|$ abhängt, dann heißt f **rotationsinverariant** und man erhält

$$\int_{K_R} f(y) \, dy = 2\pi \int_0^R r f(r) \, dr.$$

Der Faktor $2\pi r$ ist der Umfang der Kugel. Damit kann man diese Formel anschaulich motivieren.

(3) Für Integrale auf Kugeln

$$B_R = D_R(0) \subseteq \mathbb{R}^3$$

geben wir nur die Formel für rotationsinvariante f an. Es gilt

$$\int_{B_R} f(y)\,dy = 4\pi \int_0^R r^2 f(r)\,dr.$$

Mit $f = 1$ ergibt sich für das Volumen der Kugel

$$\mu(B_R) = \frac{4R^3}{3}.$$

Man beachte, dass der Faktor $4\pi r^2$ die Oberfläche der Kugel ist.

(4) Ein Zylinder $Z_{R,H}$ mit Radius R und Höhe H ergibt

$$\int_Z f(y)\,dy = \int_0^H \int_0^R \int_0^{2\pi} r f(r\cos(t), r\sin(t), h)\,dt\,dr\,dh.$$

Index

ϵ-Umgebung, 17
Überdeckung, 92
überabzählbar, 27, 39
1. Mittelwertsatz, 53
2. Mittelwertsatz, 53

Abelscher Grenzwertsatz, 78
abgeschlossen, 43, 88
abgeschlossene Hülle, 43
abgeschlossenes Intervall, 14
Ableitung, 51
Abschluss, 43, 88
absolut, 35
abzählbar, 24
additive Inverse, 10
alternierende, 35
Anfangsstück, 23
angeordneter Körper, 12
arithmetische Summe, 25
assoziativ, 9, 10
Axiom von Archimedes, 22

baryzentrische, 93
beschränkt, 14, 31, 88
Betrag, 16
Binomialentwicklung, 26
Binomialkoeffizient, 26
Bisektionsverfahren, 32

Cantorsches Diagonalargument, 39
Cauchy und Hadamard, 76
Cauchy-Folge, 36
Cauchy-Produkt, 41
Cavalieri, 122
charakteristische Funktion, 125

Dezimalentwicklung, 39
dicht, 27
differenzierbar, 51, 104
Distributivgesetz, 10
Dreiecksungleichung, 16, 17, 86

einfach, 93
einfach zusammenhängend, 119
einseitige Grenzwerte, 45
endlich, 23
Euklidsche Norm, 85
Euklidschen Skalarprodukts, 85
Eulersche Zahl, 31
Exponentialreihe, 74
Exponentialfunktion, 40
Extrema, 49

Fakultäten, 23
Feinheit, 63
Fibonacci, 23
Folge, 22
formale, 34
fortsetzen, 45
Fortsetzung, 45
Fubini, 122

ganzen Zahlen, 27
geometrische, 34
geometrische Reihe, 74
geometrische Summe, 26
gerade, 81
geschlossen, 93
gleichmäßig, 66
gleichmäßig konvergent, 77
gliedweise, 78
Gradienten, 100
Grenzwert, 29, 43

Häufungspunkt, 32
halboffene Intervalle, 14
harmonische Reihe, 35
Hauptsatz, 67
Hauptwert, 70
Hesse-Matrix, 102
Hintereinanderausführung, 55
homotop, 119

Imaginärteil, 19

Induktionsanfang, 22
Induktionsschritt, 22
induktiv, 21, 22
Infimum, 15
Inhalt, 125
Integraldarstellung, 74
Integrationsbedingungen, 116
Intervallschachtelung, 31

Jocobi-Matrix, 104

Körper, 9
Körper der komplexen Zahlen, 18
kürzen, 47
Kettenregel, 105
kommutativ, 9, 10
kompakt, 92
Komponentenfunktionen, 89
konjugiert komplexe Zahl, 19
konkaven, 60
Konvergenzradius, 76
konvergiert, 29
konvex, 60, 93
konvexe Mengen, 16
Kosinus, 79
Kurve, 93, 105

Lagrange, 111
Lagrange-Form, 74
Leibniz-Kriterium, 35
Limes inferior, 75
Limes superior, 75
Linearität, 65
Logarithmus naturalis, 57
lokal invertierbar, 107
lokaler, 49

Mächtigkeit, 24
Majoranten-Kriterium, 36
Maximalpunkt, 49
Maximum, 14, 49
Metrik, 17, 86
Minimum, 14
Minoranten-Kriterium, 37
Mittelwertsatz, 52, 72
monoton, 30
Monotonie, 16, 66
Monotonie-Intervalle, 54
multiplikative Inverse, 10

nach oben beschränkt, 14
Nachdifferenzieren, 55
natürliche Zahlen, 21
negativ definit, 103
neutrale Element, 10
Norm, 86, 95
notwendige, 102
nullhomotop, 119
Nullstellen, 46
nullteilerfrei, 10

obere Schranke, 13
Oberintegral, 64
offen, 88
offene Innere, 88
offene Mengen, 17
offenes Intervall, 14

Partialsummen, 34
partielle Integration, 68
Pascalschen Dreieck, 26
Polynom, 46
positiv definit, 17, 103
Potenzen, 23
Potenzmenge, 27
Potenzreihe, 75

Quader, 125
Quadratwurzel, 15
Quotienten-Kriterium, 38

Rand, 88
rationale Funktion, 46
rationalen Zahlen, 27
Realteil, 19
Reihe, 34
rektifizierbar, 113
rekursiv, 23
Restglied, 73
Richtungsableitung, 100
Riemann-integrierbar, 64, 126
Riemann-messbar, 127
Riemannsche Zwischensumme, 63
Ring, 27
rotationsinvariant, 123
rotationsinverariant, 128
Rotationskörper, 127

Sandwich-Prinzip, 30
Satz von Fubini, 125
Satz von Rolle, 53
Schwarzsche Ungleichung, 86

Sekante, 61
Sinus, 79
Sinus Hyperbolicus, 115
stückweise, 45
Stammfunktion, 67, 116
sternförmig, 104
stetig, 44
stetig differenzierbar, 51, 104
stetig partiell differenzierbar, 99
streng monoton, 30
strernförmig, 93
strikt konvex, 61
Substitution, 68
Supremum, 15
symmetrisch, 17, 80

Tangens, 81
tangential, 101
Taylorentwicklung, 73
Taylorreihe, 73
Teleskop-Summen, 37
total differenzierbar, 104
Transformationsformel, 128
transitiv, 12
Treppenfunktion, 63
Treppenfunktionen, 125
trigonometrischen Identitäten, 79
Tupel, 24

Umgebung, 87
Umkehrfunktion, 48
Umordnungssatz, 40
unbestimmtes Integral, 67
uneigentlich, 33
uneigentliche, 45, 70
uneigentlichen, 70
ungerade, 81
Ungleichung von Bernoulli, 21
untere Schranke, 14
Unterintegral, 64
Untersumme, 64
Unterteilung, 63

Vektorfeld, 115
vertauschbar, 77
verträglich, 12
Vielfachheit, 46
vollständig, 12
vollständig angeordnet, 15
vollständigen Induktion, 21
vollständiger metrischer Raum, 40
Vorzeichen, 16

Wendepunkt, 60
Wurzel, 48
Wurzel-Kriterium, 38

Zeilensummennorm, 96
zusammenhängend, 93
zweifachen, 100
zweimal stetig partiell, 100
Zwischenwertsatz, 47