

# MATHEMATIK IT-3

## ANALYSIS FÜR INFORMATIK-STUDIENGÄNGE

28. Oktober 2021

Prof. G. Averkov

Institut für Mathematik, Fakultät 1

Fachgebiet Algorithmische Mathematik

Brandenburgische Technische Universität Cottbus-Senftenberg

# **Inhaltsverzeichnis**

<b>I Folgen und Reihen</b>	<b>13</b>
1 Folgen . . . . .	13
2 Grenzwerte von Folgen . . . . .	27
3 Reihen . . . . .	59
<b>II Grenzwerte und Stetigkeit von Funktionen</b>	<b>115</b>

# *INHALTSVERZEICHNIS*

3

1	Grenzwert einer Funktion . . . . .	115
2	Stetige Funktionen und Unstetigkeiten . . . . .	127
3	Asymptotisches Verhalten . . . . .	145
<b>III</b>	<b>Differentialrechnung I</b>	<b>150</b>
1	Ableitung . . . . .	151
2	Monotonie und Extremwerte . . . . .	186
3	Konvexität . . . . .	206
4	Der Mittelwertsatz der Differentialrechnung . . . . .	212
<b>IV</b>	<b>Taylor- und Potenzreihen</b>	<b>219</b>
1	Taylor-Entwicklung und Taylor-Reihen . . . . .	219
2	Potenzreihen . . . . .	238
3	Exkurs in komplexwertige Funktionen . . . . .	276

<b>V Integralrechnung I</b>	<b>285</b>
1 Riemann-Integral . . . . .	285
2 Grundeigenschaften des Riemann-Integrals . . . . .	299
3 Substitutionsregel . . . . .	318
4 Partielle Integration . . . . .	324
5 Uneigentliche Integrale . . . . .	329
6 Exkurs in Lebesgue-Integrale . . . . .	335
7 Abschätzung von Summen und Reihen durch Integrale . . . . .	339
8 Integrale rationaler Funktionen . . . . .	351
9 Integrale mit trigonometrischen Funktionen . . . . .	367
<b>VI Fourier-Reihen</b>	<b>371</b>
1 Basics . . . . .	371

# INHALTSVERZEICHNIS

5

2	Der Raum $\mathcal{L}^2(0, T)$ . . . . .	379
3	Formeln für die Koeffizienten . . . . .	387
4	Differenzieren von Fourier-Reihen . . . . .	397
5	Pythagoras in $\mathcal{L}^2(0, T)$ . . . . .	404
6	Komplexwertige Funktionen . . . . .	410
7	Fourierreihen bei Signalverarbeitung . . . . .	421
8	Die Verwandten der Fourier-Reihe . . . . .	426

## VII Differentialrechnung II

436

1	Exkurs in topologische Räume . . . . .	436
2	Konvergenz und Stetigkeit im multivariaten Fall . . . . .	451
3	Partielle Ableitungen . . . . .	465
4	Kettenregel . . . . .	491

5	Totales Differential und implizite Funktionen . . . . .	512
6	Extremwertaufgaben . . . . .	516
7	Multidimensionale Taylorentwicklung . . . . .	527
8	Extremwertaufgaben: Bedingungen zweiter Ordnung . . . . .	540
9	Wiederholung: semidefinite und definite symmetrische Matrizen . . . . .	545
10	Restringierte Optimierungsaufgaben . . . . .	562
<b>VIII</b>	<b>Integralrechnung II</b> . . . . .	<b>563</b>
1	Riemann-Integral . . . . .	563
2	Der Satz von Fubini . . . . .	576
3	Substitution beim mehrdimensionalen Integrieren . . . . .	584
<b>IX</b>	<b>Anwendungen der Analysis</b> . . . . .	<b>588</b>
1	Numerische Verfahren . . . . .	588

## *INHALTSVERZEICHNIS*

7

2	Gewöhnliche Differentialgleichungen . . . . .	593
---	---	-----

# **Organatorisches**

**0.1 (Vorlesung).**

- (a) In Präsenz. Bei Bedarf Aufzeichnung.
- (b) Hilfsmittel: diese Präsentation (in der Form eines Skripts im Moodle verfügbar) und das Skript von Martin Henk (ebenfalls im Moodle).
- (c) Ziel der Vorlesung: Theorie erklären und mit vielen Beispielen bzw. Lösungen von Aufgaben illustrieren.
- (d) Fragen sehr willkommen. Nutzen Sie dafür den BigBlueButton-Chat.

**0.2 (Übung).**

- (a) Zwei Übungstermine pro Woche.
- (b) Die Teilnehmer werden nicht in zwei Übungsgruppen eingeteilt. Jede Woche kann eine der beiden Übungen frei gewählt werden.
- (c) Aufgabenblatt Nummer  $x$  (mit  $x \geq 1$ ) wird in der Vorlesungswoche  $x + 1$  diskutiert und zum Ende der Woche  $x$  abgegeben.
- (d) Abgabe der Blätter digital über Moodle.
- (e) Blatt Nummer 0 wird in der ersten Woche diskutiert und wird nicht abgegeben.

**0.3 (Prüfung).**

- (a) Zulassung zur Klausur bei mind. 50% aller Punkte für die Aufgabenblätter
- (b) Bei mind. 75% aller Punkte für die Aufgabenblätter gibt es einen 10% Bonus in der Klausur
- (c) Keine Rechengeräte in der Klausur zugelassen
- (d) Anregung: während des Semesters mitarbeiten

## Bezeichnungen

$\mathbb{Z}$  die Menge der ganzen Zahlen

$\mathbb{Z}_+ = \mathbb{N}_0$  die Menge der nichtnegativen ganzen Zahlen

$\mathbb{N}$  die Menge positiver ganzen Zahlen (wir nennen die Elemente von  $\mathbb{N}$  natürliche Zahlen).

$\mathbb{R}$  die Menge der reellen Zahlen

$\mathbb{R}_+ = \mathbb{R}_{\geq 0}$  die Menge der nichnegativen reellen Zahlen

# **Kapitel I**

## **Folgen und Reihen**

### **1 Folgen**

**1.1.** Die (klassische) **Mathematische Analysis** – deren dieser Kurs gewidmet ist – ist die Theorie, die auf dem Begriff des Grenzwertes aufgebaut ist. Wir werden Summen von verschiedenen Arten von Reihen, einfache und partielle Ableitungen, ein- und mehr-dimensionale Integrale betrachten. Das sind alles Grenzwerte. Wir beginnen mit dem Grenzwert in der einfachsten Situation: dem Grenzwert einer Folge.

**1.2 Def (Folge).** Eine unendliche **Folge** über einer Menge besteht aus unendlich vielen durchgängig indexierten Elementen der gegebenen Menge, den sogenannten **Gliedern der Folge**. Das Anfangsglied der Folge wird in der Regel mit 0 oder 1 indexiert. Eine Indexierung ab einer anderen ganzen Zahl ist auch möglich. Streng mathematisch formuliert ist eine ab 1 indexierte Folge über  $X$  eine Abbildung von  $\mathbb{N}$  nach  $X$ , welche für jeden Index  $n \in \mathbb{N}$  vorschreibt, was für einen Wert in  $X$  das Glied mit dem Index  $n$  hat. Die typischen Schreibweisen für eine Folge sind  $(a_n)_{n \in \mathbb{N}}$  mit  $n \in \mathbb{N}$  und  $a_n \in X$  oder auch  $a_1, a_2, a_3, \dots$ . Hierbei ist  $a_n$  das Glied zum Index  $n$ .

### 1.3 Bsp (Beispiele von Folgen).

- (a) Eine ziemlich langweile Folge ist die Folge, in der jedes Glied gleich ist, etwa die Folge  $(a_n)_{n \in \mathbb{N}}$  mit  $a_n = 1$  für alle  $n \in \mathbb{N}$ .
- (b) Wir können die Folge der immer besser werden Approximationen der Kreiszahl  $\pi$  betrachten,  $(a_n)_{n \in \mathbb{N}}$  mit

$$a_n = 3, \underbrace{\dots \dots \dots \dots \dots \dots}_{n \text{ Nachkommastellen der Zahl } \pi}$$

Bei dieser Folge ist  $a_1 = 3,1$ ,  $a_2 = 3,14$ ,  $a_3 = 3,141$  usw.

- (c) Die Folge  $a_n = \frac{1}{2^n}$  der Potenzen von  $\frac{1}{2}$ . Diese Folge kann auch **rekursiv** als

$$a_n = \begin{cases} \frac{1}{2}, & \text{für } n = 1 \\ \frac{1}{2}a_{n-1}, & \text{für } n > 1 \end{cases}$$

Die Glieder dieser Folge werden sehr schnell sehr klein, wenn man  $n$  wachsen lässt.

**1.4 Aufgabe.** Zeigen Sie, dass jedes Glied der Folge  $(a_n)_{n \in \mathbb{N}}$  mit  $a_n = 2^{2^n} + 2$  durch 3 teilbar ist.

**1.5 Aufgabe.** Beschreiben Sie die rekursiv definierte Folge  $(a_n)_{n \in \mathbb{N}}$

$$a_n = \begin{cases} 2, & \text{für } n = 1, \\ 1 + \prod_{i=1}^{n-1} a_i, & \text{für } n > 1. \end{cases}$$

durch die Rekursion

$$a_n = \begin{cases} 2, & \text{für } n = 1, \\ F(a_{n-1}), & \text{für } n > 1 \end{cases}$$

für eine geeignete Funktion  $F$ . Hinweis: Versuchen Sie das Produkt von  $n - 1$  Gliedern zu vermeiden, indem Sie die Rekursion (nochmals) benutzen.

**1.6 Aufgabe.** Man betrachte die Folge  $x_n = 3^n + 2^n$ . Geben Sie eine rekursive Beschreibung dieser Folge durch

$$x_n = \begin{cases} 5, & \text{für } n = 1, \\ 13, & \text{für } n = 2, \\ ax_{n-1} + bx_{n-2}, & \text{für } n > 2 \end{cases}$$

für geeignet gewählte Konstanten  $a, b \in \mathbb{R}$ . Vergewissern Sie sich, dass Ihre rekursive Darstellung tatsächlich die selbe Folge definiert. Hinweis: Die elementare Lineare Algebra kann Ihnen helfen.

**1.7 Aufgabe.** Die Folge der sogenannten Fibonacci-Zahlen ist rekursiv durch

$$f_n = \begin{cases} 1, & \text{für } n \in \{1, 2\}, \\ f_{n-1} + f_{n-2}, & \text{für } n > 2 \end{cases}$$

definiert. Haben Sie bereits eine nicht-rekursive Darstellung dieser Folge gesehen? Die Darstellung hat etwas mit den Lösungen der quadratischen Gleichung  $x^2 - x - 1 = 0$  zu tun (die positive Lösung heißt der goldene Schnitt). Leiten Sie diese Darstellung her. Hinweis: Berechnen Sie die beiden Nullstellen  $u, v$  von  $x^2 - x - 1$  und versuchen Sie dann  $f_n$  geeignet als Linearkombination von  $u^n$  und  $v^n$  darzustellen.

**1.8.** Wir werden im Folgenden vor allem reellwertige Folgen betrachten, da es in diesem Kurs vorrangig um die **Reelle Analysis** geht: das ist die Analysis, die auf reellen Zahlen aufgebaut ist. Das heißt, unsere Menge für die Werte der Folgen wird in den meisten Fällen  $\mathbb{R}$  sein.

**1.9.** Folgen betrachtet man unter anderem in iterativen Rechenprozessen, die theoretisch unendlich lang laufen können, bei denen man in jeder Iteration ein aktuelles Ergebnis  $a_n$  ausrechnet. Abhängig von der vorliegenden Situation kann dieses aktuelle Ergebnis eine Zahl oder ein Vektor, oder evtl. auch ein anderes, komplexeres Objekt sein.

**1.10 Bsp** (Iterative Lösung von Fixpunktgleichungen). Betrachten wir die Fixpunktgleichung

$$f(x) = x$$

für eine reellwertige Funktion  $f$  einer reellen Variablen  $x$ . In der Schule ist es so, dass jede Gleichung, die man betrachtet, eine Lösung in einer netten abgeschlossen Form darstellbar ist. In der realen Welt ist die Situation etwas anders. Betrachten wir zum Beispiel die Gleichung  $2 \sin(x) = x$ . Diese Gleichung hat eine positive Lösung, aber als eine nette Formel werden wir diese Lösung nicht hinschreiben können. Die Lösung ist  $x \approx 1,89$ .

Man kann versuchen, die Gleichung  $f(x) = x$  annähernd auf die folgende Weise zu lösen. Man beginnt mit einem Startwert  $x_0$  und setzt in der  $n$ -ten Iteration  $x_n = f(x_{n-1})$  für  $n > 0$ . Wenn man Glück hat (das heißt, gewisse Voraussetzungen an  $f$  und  $x_0$  sind erfüllt), ist  $x_n$  für genügend große  $n$  eine beliebig gute Approximation einer Lösung von  $f(x) = x$ . In diesem Beispiel reden wir also von einer Folge  $(x_n)_{n \in \mathbb{N}_0}$  von Approximationen einer Lösung. Aber was

heißt es eigentlich, dass die Folge  $(x_n)_{n \in \mathbb{N}_0}$  einen gewissen Wert  $a \in \mathbb{R}$  beliebig gut approximiert?  
Genau das wird durch den Grenzwertbegriff geklärt, den wir einführen wollen.

**1.11 Aufgabe.** Berechnen Sie die (eindeutige) positive Lösung von  $2 \sin(x) = x$  mit der Benutzung von  $x_0 = 2$  als Startwert wie oben dargestellt. Als Hilfsmittel können Sie beliebige Soft- bzw. Hardware benutzen. Was kriegen Sie für Werte für  $x_n$  für  $n \in \{1, \dots, 10\}$ ?

## 2 Grenzwerte von Folgen

**2.1 Def** (Grenzwert einer Folge). Man sagt, dass eine Folge  $(a_n)_{n \in \mathbb{N}}$  von reellen Zahlen gegen einen reellen Wert  $a$  konvergiert wenn Folgendes gilt: für jedes  $\epsilon > 0$  existiert ein  $n_0 \in \mathbb{N}$ , sodass für alle  $n \in \mathbb{N}$  mit  $n \geq n_0$  die Ungleichung  $|a_n - a| < \epsilon$  erfüllt ist.

In diesem Fall schreibt man  $a_n \rightarrow a$  für  $n \rightarrow \infty$  oder auch  $\lim_{n \rightarrow \infty} a_n = a$ . Man sagt:

- $a$  ist Grenzwert von  $a_n$  für  $n \rightarrow \infty$  oder
- $a$  ist Limes von  $a_n$  für  $n \rightarrow \infty$  oder
- $a_n$  geht gegen  $a$  für  $n \rightarrow \infty$  oder
- $a_n$  konvergiert gegen  $a$  für  $n \rightarrow \infty$ .

**2.2.** Wir schauen uns den Begriff des Grenzwerts in Ruhe an. Die Bedingung  $|a_n - a| < \epsilon$  können wir etwa als “ $a_n$  ist  $\epsilon$ -Approximation von  $a$ ” mit Worten beschreiben. Die Bedingung  $n \geq n_0$  heißt “ $n$  muss groß genug sein”, mindestens so groß wie ein passend gewähltes  $n_0$ . Also beschreibt unsere Grenzwert-Definition die folgende Situation: egal welche Güte  $\epsilon > 0$  der Approximation des Wertes  $a$  wir vorschreiben, wird diese Güte durch alle Glieder  $a_n$  der Folge erreicht, bei denen der Index  $n$  genügend groß ist.

**2.3.** Ab wann  $n$  genügend groß ist wird in der Definition durch eine natürliche Zahl  $n_0$  vorgeschrieben. Streng genommen brauchen wir für den Wert  $n_0$  nicht unbedingt Ganzzahligkeit. Daher findet man in der Literatur auch die folgende äquivalente Definition des Grenzwerts einer Folge: für jedes  $\epsilon > 0$  existiert ein  $N \in \mathbb{R}$ , sodass für alle  $n \in \mathbb{N}$  mit  $n \geq N$  die Ungleichung  $|a_n - a| < \epsilon$  erfüllt ist. In dieser Definition ist  $n_0$  durch eine reelle Zahl  $N$  ersetzt.

**2.4.** Der Begriff des Grenzwerts einer Folge kann ganz analog über den Mengen eingeführt werden, in denen ein Abstandsbegriff festgelegt ist: die Bedingung  $|a_n - a| < \epsilon$ , dass der Abstand von  $a_n$  und  $a$ . Für die komplexen Zahlen ist der Abstand von zwei Zahlen ebenso der Betrag der Differenz dieser Zahlen. Also kann unsere Definition des Grenzwerts ohne Änderungen für folgen komplexer Zahlen benutzt werden.

**2.5 Aufgabe.** Der Quotient  $a_n = f_{n+1}/f_n$  der aufeinanderfolgenden Fibonacci-Zahlen ist ebenfalls eine Folge, über deren Eigenschaften für große  $n$  man reflektieren könnte. Rechnen Sie die ersten Zehn Glieder  $a_1, \dots, a_{10}$  dieser Folge mit Computer/Taschenrechner in Gleitkomma-Arithmetik aus. Was stellen Sie experimentell fest? Sieht es wie eine konvergente Folge aus? Was könnte der Grenzwert sein?

## 2.6 Bsp.

- (a) In der Folge  $a_n = 5 + \frac{(-1)^n}{n}$  schwanken die Glieder um die Zahl 5 und liegen mal oberhalb mal unterhalb dieser Zahl. Der Grenzwert  $\lim_{n \rightarrow \infty} a_n$  ist gleich 5, weil die Abweichungen von 5 für große  $n$  beliebig klein werden. Wir können es auch formal aus der Definition des Grenzwerts herleiten. Sei  $\epsilon > 0$  gegeben. Wir können nun als  $N = \frac{2}{\epsilon}$  fixieren. Für alle  $n \geq N$  gilt  $|a_n - 5| \leq \frac{1}{2\epsilon} < \epsilon$ .
- (b) Die Folge  $a_n = (-1)^n$  besteht aus Werten  $-1$  und  $1$ , welche in dieser Folge abwechselnd auftreten. Diese Folge konvergiert gegen keinen Wert. Das ist intuitiv klar. Wenn die Folge für  $n$  überhaupt was approximieren kann, dann nur  $-1$  oder  $1$ , da der Wert beim wachsenden  $n$  ständig zwischen  $-1$  und  $1$  wechselt, approximiert die Folge weder  $-1$  noch  $1$ .

**2.7 Thm** (Eindeutigkeit des Grenzwertes). *Jede konvergente Folge hat genau einen Grenzwert.*

**2.8.** Dieser Satz ist intuitiv klar: Eine konvergente Folge kann nur einen Wert approximieren. Wenn  $a_n$  für große  $n$  gleich zwei verschiedene Werte, etwa 0 und 1, approximieren würde, dann wäre  $a_n$  gleichzeitig sehr nah an 0 und sehr nah an 1. Das geht natürlich nicht.

**2.9 Def** (Beschränktheit & Monotonie). Eine Folge  $(a_n)_{n \in \mathbb{N}}$  heißt:

- (a) **nach oben beschränkt**, wenn für eine Konstante  $C$  die Ungleichung  $a_n \leq C$  für alle  $n$  erfüllt ist;
- (b) **nach unten beschränkt**, wenn für eine Konstante  $c$  die Ungleichung  $a_n \geq c$  für alle  $n$  erfüllt ist.
- (c) **beschränkt**, wenn die Folge nach oben sowie nach unten beschränkt ist;
- (d) **monoton steigend**, wenn  $a_n \leq a_{n+1}$  für alle  $n$  erfüllt ist;
- (e) **strikt monoton steigend**, wenn  $a_n > a_{n+1}$  für alle  $n$  erfüllt ist;
- (f) **monoton fallend**, wenn  $a_n \geq a_{n+1}$  für alle  $n$  erfüllt;

- (g) **strikt monoton fallend**, wenn  $a_n > a_{n+1}$  für alle  $n$  erfüllt ist;
- (h) **monoton**, wenn die Folge monoton steigend oder monoton fallend ist.

**2.10 Thm** (Zusammenhang der Monotonie, Beschränktheit und Konvergenz). *Für Folgen reeller Zahlen gilt Folgendes:*

- (a) *Jede monotone, beschränkte Folge konvergiert.*
- (b) *Jede konvergente Folge ist beschränkt (aber im Allgemeinen, nicht umgekehrt).*

**2.11.** Wenn wir in einer Folge endlich viele Glieder ändern, hat so eine Änderung keine Auswirkung auf Konvergenz und Grenzwert. Denn in der Definition des Grenzwert geht es um die genügend großen  $n$ . Man kann sich also auf so große  $n$  einschränken, bei denen man nur die Glieder sieht die unverändert geblieben sind. Daraus folgt, dass wir in der Behauptung (a) die Monotonie durch die Monotonie ab irgendeinem Glied ersetzen können: wenn eine Folge beschränkt und ab einem Glied monoton ist, ist sie konvergent.

**2.12 Bsp.** Das Theorem gibt uns eine hinreichende und eine notwendige Bedingung für Konvergenz. Probieren wir es mal aus.

- (a) Die Folge  $a_n = 2^{1/n}$  ist monoton und beschränkt. Je höher die Wurzel ist, die wir aus 2 ziehen, desto kleiner ist das Ergebnis (daher die Monotonie). Beschränktheit ist ebenfalls klar, denn der Wert  $2^{1/n}$  liegt offensichtlich zwischen 1 und 2. Auf diese Weise sehen wir dass die Folge  $a_n$  konvergent ist, ohne dass wir den genauen Grenzwert dieser Folge angeben. Es kommen tatsächlich manchmal Situationen vor, in denen man weiß, dass der Grenzwert der vorliegenden Folge existiert, aber kein schöner einfacher Ausdruck für den Grenzwert vorhanden ist. So, weiß man zum Beispiel, dass der Grenzwert  $\lim_{n \rightarrow \infty} (1+1/n)^n$  existiert, den Grenzwert kann man aber nicht "elementar beschreiben". Und so hat man einfach für diesen Wert den Buchstaben  $e$  reserviert und den Wert die Euler-Zahl genannt.
- (b) Die Folge  $a_n = n$  konvergiert nicht, denn diese Folge ist unbeschränkt.

**2.13 Thm.** *Die Folge  $(1 + 1/n)^n$  ist konvergent. Der Grenzwert dieser Folge wird als e bezeichnet und die Euler-Zahl genannt.*

**2.14 Bsp.** Kehren wir zurück zur Folge  $2^{1/n}$ . Wir bestimmen nun den Grenzwert dieser Folge. Wir zeigen, dass der Grenzwert gleich 1 ist. Dafür sollen wir zeigen, dass für ein beliebiges festes  $\epsilon > 0$ , die Ungleichung  $|2^{1/n} - 1| < \epsilon$  für alle genügend großen  $n$  erfüllt ist. Schreiben wir unsere Ungleichung ohne Betrag als eine untere und obere Schranke an  $2^{1/n}$ :

$$1 - \epsilon < 2^{1/n} < 1 + \epsilon.$$

Die untere Schranke ist für alle  $n$  erfüllt, denn  $2^{1/n}$  ist sogar durch 1 nach unten beschränkt. Nun zur oberen Schranke: diese kann als  $1/n < \log_2(1 + \epsilon)$  und dann als  $n > \frac{1}{\log_2(1+\epsilon)}$  umformuliert werden. Also ist die obere Schranke tatsächlich für alle genügend großen  $n$  erfüllt.

**2.15 Def.** Eine nicht-konvergente Folge nennt man divergent.

**2.16.** Man muss aber an dieser Stelle etwas Ordnung in die Menge der divergenten Folgen bringen. Manche dieser Folgen gehen tatsächlich gegen keinen Wert, wie etwa  $(-1)^n$ , manche anderen gehen schon gegen etwas, aber dieses etwas ist kein reeller Wert.

**2.17 Def.** Man sagt, eine Folge  $(a_n)_{n \in \mathbb{N}}$  **konvergiert gegen**  $\infty$  (bzw. ist **bestimmt divergent gegen**  $+\infty$ ), wenn für jede Schranke  $C$  für genügend große  $n$  überschritten wird: das heißt, für jede Konstante  $C$  existiert ein  $n_0$  derart, dass die Ungleichung  $a_n > C$  für alle  $n \geq n_0$  erfüllt ist. Komplett analog definiert man auch die bestimmte Divergenz gegen  $-\infty$ .

**2.18 Bsp.**

- (a) Die Folge  $a_n = n$  konvergiert gegen  $\infty$ , das schreibt man als  $\lim_{n \rightarrow \infty} n = \infty$ .
- (b) Die Folge  $a_n = -n$  konvergiert gegen  $-\infty$ . Das schreibt man als  $\lim_{n \rightarrow \infty} (-n) = -\infty$ .
- (c) Die Folge  $a_n = (-1)^n n$  ist divergent, geht aber weder gegen  $\infty$  noch gegen  $-\infty$ , denn  $a_n$  schwankt für große  $n$  zwischen großen positiven Werten und betragsmäßig großen negativen Werten. Der Betrag  $|a_n| = n$  aber konvergiert gegen  $\infty$ .
- (d) Die Folge  $a_n = (1 + (-1)^n)n$  schwankt für große  $n$  zwischen 0 und sehr großen Werten. Die Folge ist divergent und sie geht weder gegen  $\infty$  noch gegen  $-\infty$ .

**2.19 Def.** Eine **Polynomfunktion** ist eine Funktion der Form

$$f(x) = c_d x^d + \cdots + c_0 x^0$$

mit Koeffizienten  $c_d, \dots, c_0$  und  $d \in \mathbb{N}_0$ . Sind alle Koeffizienten von  $f$  gleich 0, so setzen man den Grad  $\deg(f)$  gleich  $-\infty$ . Ist  $c_d \neq 0$ , so definiert man den Grad von  $f$  durch  $\deg(f) = d$  und nennt den Term  $c_d x^d$  den **Leitterm** von  $f$ .

**2.20 Bsp.**

- (a) Der Leiterterm von  $2x^3 - 7x^2 + 8x - 6$  ist  $2x^3$ .
- (b) Der Leiterterm von  $(2x - 1)^{10}$  ist  $(2x)^{10}$ . (Wieso?)

**2.21.** Eigentlich gibt es einen Unterschied zwischen Polynomen und Polynomalfunktionen. Polynome sind formale Ausdrücke und Polynomalfunktionen sind spezielle Abbildungen, aber im Kontext der Analysis ist dieser Unterschied nicht so wichtig, sodass wir Polynomalfunktionen auch einfach Polynome nennen werden.

**2.22 Thm.** Seien  $f, g$  Polynome mit Koeffizienten in  $\mathbb{R}$  und sei  $g$  nicht identisch gleich null.

Dann gilt

$$\lim_{n \rightarrow \infty} \frac{f(n)}{g(n)} = \lim_{n \rightarrow \infty} \frac{(\text{Leitterm von } f)(n)}{(\text{Leitterm von } g)(n)}.$$

**2.23 Bsp.**

(a)

$$\lim_{n \rightarrow \infty} \frac{5n^3 - 2n^2 + 7}{10n^3 + 3n^2 + 4n + 5} = \lim_{n \rightarrow \infty} \frac{5n^3}{10n^3} = \frac{5}{10} = \frac{1}{2}.$$

(b) Der Leiterterm des Produkts ist Produkt der Leiterme. Darauf basierend haben wir:

$$\lim_{n \rightarrow \infty} \frac{(2n^2 + 1)^3}{(5n^3 + n^2 - 1)^2} = \lim_{n \rightarrow \infty} \frac{(2n^2)^3}{(5n^3)^2} = \frac{8}{25}.$$

(c)

$$\lim_{n \rightarrow \infty} \frac{2n + 1}{3n^2 - 7n + 6} = \lim_{n \rightarrow \infty} \frac{2n}{3n^2} = 0.$$

(d)

$$\lim_{n \rightarrow \infty} \frac{n^2 + 100}{3n - 1000} = \lim_{n \rightarrow \infty} \frac{n^2}{3n} = \infty.$$

**2.24 Thm** (Rechenregeln für Folgen). *Für konvergente Folgen  $(a_n)_{n \in \mathbb{N}}, (b_n)_{n \in \mathbb{N}}$  gilt:*

(a)  $\lim_{n \rightarrow \infty} (a_n + b_n) = \lim_{n \rightarrow \infty} a_n + \lim_{n \rightarrow \infty} b_n$

(b)  $\lim_{n \rightarrow \infty} (a_n \cdot b_n) = (\lim_{n \rightarrow \infty} a_n) \cdot (\lim_{n \rightarrow \infty} b_n).$

(c)  $\lim_{n \rightarrow \infty} b_n \neq 0 \implies \lim_{n \rightarrow \infty} \frac{a_n}{b_n} = \frac{\lim_{n \rightarrow \infty} a_n}{\lim_{n \rightarrow \infty} b_n}.$

(d)  $a_n \geq 0 \text{ für alle } n \in \mathbb{N} \implies \lim_{n \rightarrow \infty} a_n^{b_n} = (\lim_{n \rightarrow \infty} a_n)^{\lim_{n \rightarrow \infty} b_n}.$

**2.25. WARNUNG!** Wenn man mit Folgen mit Hilfe der obigen Regeln rechnet, soll man das Folgende beachten: Es kann passieren, dass  $a_n + b_n$  konvergiert, auch wenn weder  $a_n$  noch  $b_n$  konvergieren. Zum Beispiel bei  $a_n = -b_n = (-1)^n$ . Die Summe  $a_n + b_n = 0$  konvergiert gegen 0, aber  $a_n$  und  $b_n$  konvergieren nicht. Wenn man also einer Folge etwa in die Summe von zwei Folgen ungünstig zerlegt hat, sodass die beiden Folgen divergent sind, heißt es noch lange nicht, dass die zugrundeliegende Folge divergent ist!

**2.26 Def** (Teilfolge und ihr Grenzwert). Eine **Teilfolge** über einer Menge  $X$  besteht aus den Gliedern  $a_{n_1}, a_{n_2}, a_{n_3}, \dots$  aus der Menge  $X$ , die durch ganze Zahlen  $n_1 < n_2 < n_3 < \dots$  indexiert sind.

Der einzige Unterschied zu Folgen: Im Gegensatz zu Folgen sind die Glieder nicht unbedingt durchgängig nummeriert. Mit anderen Worten: Teilfolgen unterscheiden sich von den Folgen nur durch die Art der Nummerierung.

Die für Folgen eingeführten Begriffe wie **Grenzwert**, **Konvergenz** und **Divergenz** werden für Teilfolgen völlig analog definiert.

**2.27 Bsp.**  $(a_p)_p$  ist Primzahl mit  $a_p = \frac{1}{p^2}$  ist eine Teilfolge, die durch die Menge der Primzahlen indexiert ist. Man hat  $a_2 = \frac{1}{4}, a_3 = \frac{1}{9}, a_5 = \frac{1}{25}$  usw.

**2.28 Def.** Ist  $(a_n)_{n \in \mathbb{N}}$  Folge und  $1 \leq n_1 < n_2 < n_3 < \dots$  unendlich viele Indizes, so nennt man  $(a_{n_k})_{k \in \mathbb{N}}$  eine **Teilfolge der Folge**  $(a_n)_{n \in \mathbb{N}}$ .

**2.29 Thm.** Konvergiert eine Folge gegen einen Wert  $a$ , so konvergiert jede Teilfolge dieser Folge gegen den selben Wert  $a$ .

**2.30 Bsp.** Das vorige Theorem kann benutzt werden, um Divergenz nachzuweisen.

(a) Die Folge  $a_n = (-1)^n$  ist divergent, denn

die Teilfolge  $a_{n_k} = -1$  mit  $n_k = 2k - 1$  konvergiert gegen  $-1$  und

die Teilfolge  $a_{n_k} = 1$  mit  $n_k = 2k$  konvergiert gegen den anderen Wert (die 1).

(b) Die Folge

$$a_n = \begin{cases} 1, & n \text{ Zweierpotenz,} \\ 0, & \text{sonst} \end{cases}$$

ist divergent, denn

die Teilfolge  $a_{n_k} = 1$  mit  $n_k = 2^k$  geht gegen 1 und

die Teilfolge  $a_{n_k} = 0$  mit  $n_k = 2k + 1$  geht gegen 0.

### 3 Reihen

**3.1.** Salopp gesagt sind Reihen unendliche Summen. Etwas genauer sind es formale Summen der Glieder einer gegebener Folge.

Wir reden im Folgenden meistens über Reihen mit reellwertigen Gliedern.

**3.2 Def** (Reihe). Eine **Reihe** ist ein formaler Ausdruck der Form  $\sum_{k=0}^{\infty} a_k$ . Die Elemente  $a_k$  nennt man **Glieder** der Reihe. Die Glieder werden bei uns in der Regel reelle Werte sein.

Der Wert  $s_n = \sum_{k=0}^n a_k$  ist **Partialsumme** zum Index  $n$ .

Eine Reihe nennt man **konvergent**, wenn die Folge  $(s_n)_{n \in \mathbb{N}_0}$  der Partialsummen einen endlichen Grenzwert  $s \in \mathbb{R}$  besitzt. In diesem Fall nennt man  $s$  die Summe der Reihe und schreibt

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

Man benutzt in Bezug auf Reihen die gleiche Terminologie wie bei Folgen wie **divergent**, **bestimmt divergent gegen**  $\pm\infty$ , indem man sich dabei auf die Folge der Partialsummen bezieht.

**3.3.** Neben den Reihen mit den Gliedern, die ab 0 indexiert sind, kann man natürlich auch Reihen betrachten, deren Gliedern ab einem anderen Wert indexiert sind. Neben der 0 ist die 1 ebenfalls sehr verbreitet.

**3.4 Def.** Eine **Nullfolge** ist eine Folge, die gegen 0 konvergiert.

**3.5 Thm** (Notwendige Bedingung für Konvergenz einer Reihe). *Die Folge der Glieder einer konvergenten Reihe ist eine Nullfolge.*

**3.6.** Die Begründung zu diesem Theorem ist total einfach: hat  $s_n = \sum_{k=0}^n a_k$  einen Grenzwert  $s$  für  $n \rightarrow \infty$ , dann hat natürlich auch  $s_{n-1}$  den Grenzwert  $s$ , für  $n \rightarrow \infty$ . Also ist  $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} s_n - \lim_{n \rightarrow \infty} s_{n-1} = s - s = 0$ .

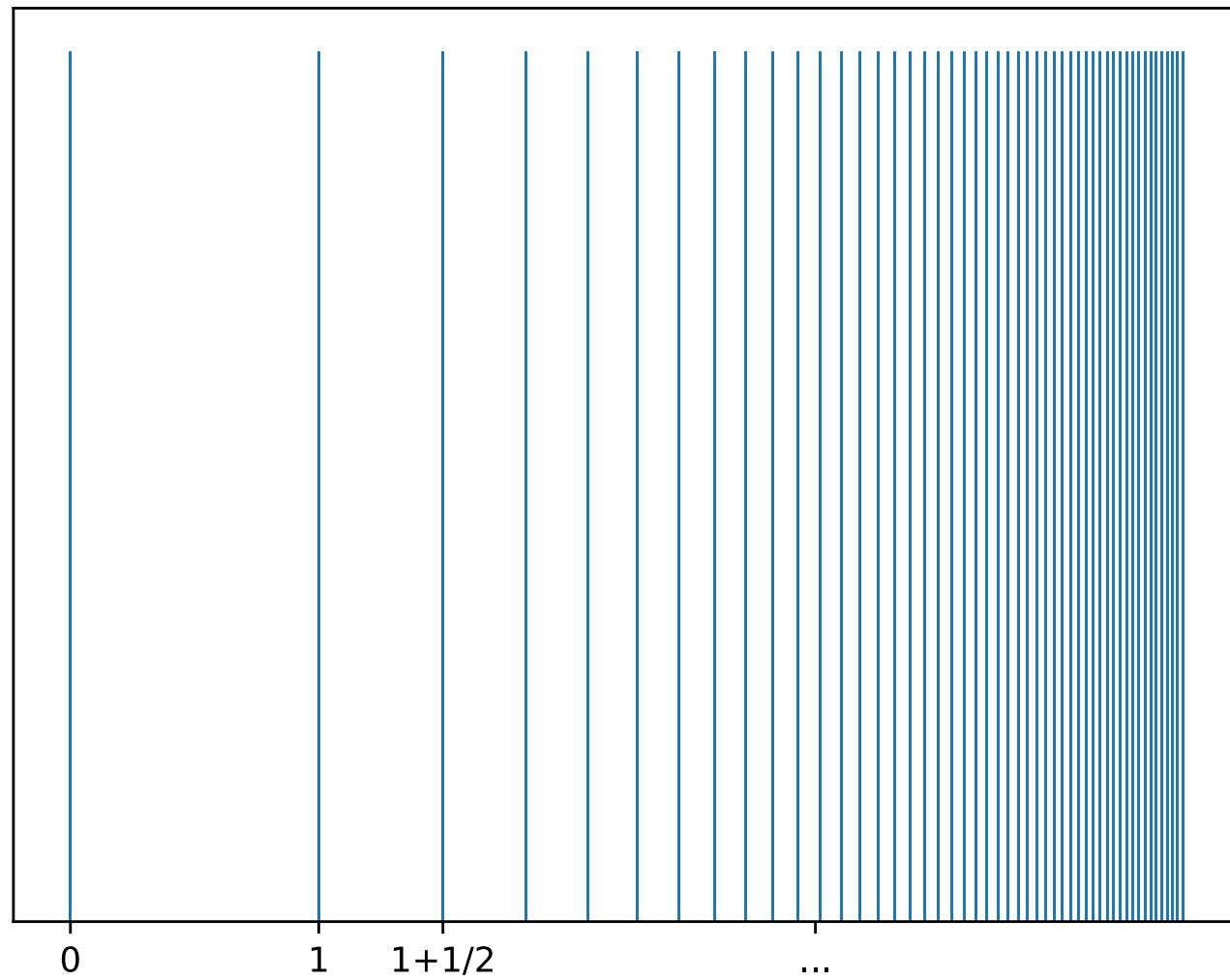
**3.7 Bsp.** Die Reihe  $\sum_{k=1}^{\infty} (-1)^k$  ist divergent, da die Folge der Glieder nicht gegen 0 konvergiert. In diesem Fall konvergiert die Folge der Glieder gar nicht. Wie sehen die Partialsummen dieser Reihe aus?

**3.8 Bsp.** Dass die Folge der Glieder gegen 0 konvergiert, ist im Allgemeinen keine hinreichende Bedingung für die Konvergenz der Reihe. Betrachten wir die Reihe

$$1 + \underbrace{\frac{1}{2} + \frac{1}{2}}_{2 \text{ mal}} + \underbrace{\frac{1}{3} + \frac{1}{3} + \frac{1}{3}}_{3 \text{ mal}} + \cdots + \underbrace{\frac{1}{k} + \cdots + \frac{1}{k}}_{k \text{ mal}} + \cdots$$

Das heißt, die Glieder sind  $a_0 = 1$ ,  $a_1 = a_2 = \frac{1}{2}$ ,  $a_3 = a_4 = a_5 = \frac{1}{3}$  usw. Die Folge der Glieder geht gegen 0. Die Summe der Reihe ist aber unendlich, denn die  $k$  Glieder mit dem Wert  $\frac{1}{k}$  ergeben insgesamt  $k \cdot \frac{1}{k} = 1$ . Da wir aber für jedes  $k$ , eine Gruppe aus  $k$  Gliedern mit dem Wert  $\frac{1}{k}$  haben, ist die Summe der Reihe unendlich.

**3.9 Bsp** (Harmonische Reihe). Die Reihe  $\sum_{k=1}^{\infty} \frac{1}{k}$  heißt harmonisch. Die Summe der Reihe ist unendlich!



Wir zerlegen die Glieder in Gruppen, indem wir ab jeder Zweierpotenz eine neue Gruppe starten:

Es gilt

### 3. REIHEN

71

$$\sum_{k=1}^n \frac{1}{k} \simeq e_n$$

$$\begin{aligned}
 \sum_{k=1}^{\infty} \frac{1}{k} &= \sum_{i=0}^{\infty} \sum_{k=2^i}^{2^{i+1}-1} \frac{1}{k} \\
 &\geq \sum_{i=0}^{\infty} \underbrace{\sum_{k=2^i}^{2^{i+1}-1} \frac{1}{2^{i+1}}}_{2^i \text{ Terme}} \\
 &= \sum_{i=0}^{\infty} 2^i \frac{1}{2^{i+1}} \\
 &= \sum_{i=0}^{\infty} \frac{1}{2} \\
 &= \infty.
 \end{aligned}$$

**3.10.** Es stellt sich heraus, dass die Partialsumme  $s_n = \sum_{k=1}^n \frac{1}{k}$  der harmonischen Reihe ziemlich gut durch den natürlichen Logarithmus  $\ln n$  approximiert werden kann.

Dass der QuickSort im Durchschnitt die Laufzeit der Ordnung  $n \ln n$  hat, hängt auch mit der Partialsumme der harmonischen Reihe zusammen. Mehr Details dazu findet man im Buch von Cormen, Leiserson, Rivest und Stein “Algorithmen: Eine Einführung”.

**3.11 Bsp** (Geometrische Reihe). Die geometrische Reihe  $\sum_{k=0}^{\infty} q^k$  zu Basis  $q \in \mathbb{R} \setminus \{1\}$  ist eines der bekanntesten Beispiele einer Reihe. Wir analysieren diese Reihe auf Konvergenz. Es ist nicht besonders schwer, eine abgeschlossene Formel für die Partialsummen

$$s_n = q^0 + \cdots + q^n$$

zu erstellen. Der Wert  $qs_n$  ist die Summe  $q^1 + \cdots + q^{n+1}$ . Wenn wir also von  $qs_n$  die Summe  $s_n$  abziehen, kompensieren sich die meisten Terme und man erhält  $qs_n - s_n = q^{n+1} - 1$ . Somit ist

$$s_n = \frac{1 - q^{n+1}}{1 - q}. \quad (\text{I.1})$$

Nun sieht man, dass die geometrische Reihe genau dann konvergiert, wenn  $|q| < 1$  gilt. In der Tat:

- (a) Für  $|q| < 1$ , geht  $q^n$  in der Darstellung (I.1) von  $s_n$  gegen 0, woraus man  $s_n \rightarrow \frac{1}{1-q}$  erhält.

(b) Für  $|q| \geq 1$ , ist das Glied  $q^k$  der Reihe keine Nullfolge, weil man  $|q^k| \geq 1$  hat. Bei  $|q| \geq 1$  kann noch zwischen der bestimmten Divergenz und unbestimmten Divergenz unterscheiden, denn:

- Bei  $q > 1$ , geht  $s_n$  gegen  $\infty$ .
- Bei  $q = -1$  geht die Partiasumme  $s_n$  gegen “gar nichts”: sie zwischen zwei festen Werten.
- Bei  $q < -1$  schwankt die Partialsumme  $s_n$  zwischen “sehr positiven” und “sehr negativen” Werten und geht somit ebenfalls gegen “gar nichts”.

Zusammenfassend:  $\sum_{k=0}^{\infty} q^k$  ist

- (a) bei  $|q| < 1$  der endliche Wert  $\frac{1}{1-q}$  (die Reihe ist konvergent).
- (b) bei  $q > 1$  der unendliche Wert  $+\infty$  (die Reihe ist bestimmt divergent gegen  $+\infty$ )

(c) bei  $q \leq -1$  "gar nichts" (die Reihe ist unbestimmt divergent).

**3.12 Bsp** (Dezimaldarstellung von Zahlen). Eigentlich ist die Darstellung von reellen Zahlen mit Hilfe von potenziell unendlich vielen Nachkommastellen ebenfalls eine Reihe. Wenn wir zum Beispiel die Zahl  $0.\overline{1} = 0,1111\cdots$  betrachten, so meinen wir darunter die Summe der Reihe

$$\sum_{k=1}^{\infty} \frac{1}{10^k}.$$

Schreiben Sie diese Zahl als Bruch. Bei allgemeinen Zahlen der Form  $0,z_1z_2z_3\cdots$  mit den Nachkommastellen  $z_1, z_2, \dots \in \{0, \dots, 9\}$  ist die entsprechende Summe gleich

$$\sum_{k=1}^{\infty} \frac{z_k}{10^k}.$$

**3.13 Aufgabe.** Eine Programmfunction wirft innerhalb einer While-Schleife iterativ eine faire Münze. Sobald eine Zahl geworfen wird, terminiert Ihre Funktion. Dabei ist der Rückgabewert 1, wenn die Funktion nach einer ungeraden Anzahl der Iterationen terminiert hat, und 0 sonst. Wie hoch ist die Wahrscheinlichkeit, dass die Funktion mit dem Rückgabewert 1 terminiert? Was hat diese Wahrscheinlichkeit mit Reihen zu tun?

Hinweis: Die elementare Wahrscheinlichkeitstheorie hatten Sie in der Schule.

```
import random

KOPF=0
ZAHL=1

def unfaire_muenze():
    cntr=0
    while True:
        cntr+=1
        if random.randint(KOPF,ZAHL)==ZAHL:
```

```
    return cntr % 2

# TEST
print([unfaire_muenze() for i in range(50)])
```

**3.14 Aufgabe.** Bestimmen Sie die Werte  $z_1, z_2, z_3, z_4, \dots \in \{0, 1\}$  derart, dass die Gleichung

$$\sum_{k=1}^{\infty} \frac{z_k}{2^k} = \frac{1}{3}$$

erfüllt ist.

Hinweis: Im Binärsystem kann man sehr bequem rechnen. Da Sie Informatiker/innen/ sind, muss das Binärsystem eines Ihrer Lieblingssysteme sein, oder?

**3.15 Bsp.** Bei manchen elementaren Reihen, deren Glieder in rationaler Arithmetik erstellt werden, kommt erstaunlicherweise als Summe ein interessanter irrationaler Wert aus. Man hat zum Beispiel

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

Wer hätte gedacht, das in einem so einfachen Ausdruck wie  $\sum_{k=1}^{\infty} \frac{1}{k^2}$  die Kreiszahl versteckt ist! Wir könne diese Formel noch nicht herleiten, aber nach einigen Vorlesungen schon...

**3.16 Thm** (Linearkombination von Reihen). *Sind  $\sum_{k=1}^{\infty} a_k$  und  $\sum_{k=1}^{\infty} b_k$  konvergente Reihen, dann ist Ihre Linearkombination  $\sum_{k=1}^{\infty} (\alpha a_k + \beta b_k)$  mit Koeffizienten  $\alpha, \beta \in \mathbb{R}$  ebenfalls eine konvergente Reihe und es gilt:*

$$\sum_{k=1}^{\infty} (\alpha a_k + \beta b_k) = \alpha \sum_{k=1}^{\infty} a_k + \beta \sum_{k=1}^{\infty} b_k.$$

**3.17 Def** (Absolute Konvergenz). Eine Reihe  $\sum_{n=1}^{\infty} a_n$  heißt **absolut konvergent**, wenn die Reihe  $\sum_{n=1}^{\infty} |a_n|$  der Beträge der Glieder konvergent ist.

**3.18 Thm.** *Jede absolut konvergente Reihe ist konvergent, aber im Allgemeinen nicht umgekehrt.*

**3.19 Bsp.**  $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$  ist konvergent, aber nicht absolut konvergent.

**3.20. WARNUNG!** Die konvergenten Reihen, die nicht absolut konvergent sind, sind ziemlich eigenartig!!! Ihr Grenzwert hängt zum Beispiel von der Reihenfolge der Glieder ab.

Die absolut konvergenten Reihen sind dagegen ganz in Ordnung: sie sind im gewissen Sinne ähnlicher zu endlichen Summen.

**3.21 Bsp.** Aus den Werten der Menge  $M := \{\pm 1, \pm \frac{1}{2}, \pm \frac{1}{3}, \dots\}$  kann man durch eine geeignete Anordnung der Werte eine Reihe erstellen, die eine beliebige von uns gewünschte Summe hat. Zum Beispiel:

- Starte mit  $s_0 := 0$  (vorige Partialsumme) und Zähler  $n := 1$  (aktuelles Glied) sowie  $i := 1$  (Index zum Verbrauch der positiven Werte) und  $j := 1$  (Index zum Verbrauch der negativen Werte).
- Iteriere unendlich wie folgt:

Ist  $s_{n-1} \leq 20$ , setze  $a_n := 1/i$  und  $i := i+1$ , sonst setze  $a_n := -1/j$  und  $j := j+1$ .

Generiere die nächste Partialsumme  $s_n := s_{n-1} + a_n$  und setze  $n := n + 1$ .

- Ende der Iteration.

Dieses Verfahren produziert eine Reihe  $\sum_{n=1}^{\infty} a_n$  mit der Summe 20. Hierbei werden alle Elemente von  $M$  als Glieder von  $a_n$  auftauchen, weil die harmonische Reihe  $\sum_{k=1}^{\infty} \frac{1}{k}$  gleich  $\infty$  ist, was dazu führt, dass man sowohl bei den negativen Elementen als auch bei den positiven Elementen aus  $M$  nie aufhört, diese Elemente zu verbrauchen.

Genauso könnte man aber an der Stelle von 20 einen beliebigen Wert vorgeben, etwa 40, und eine Reihe  $\sum_{n=1}^{\infty} b_n$  produzieren, die bis auf die Änderung der Anordnung, dieselben Glieder aber eine andere Summe (die 40) hat!

**3.22.** Wir wollen nun Produkte von Reihen einführen. Wir sind also an einer Reihe interessiert, die auf irgendeine Weise das Produkt

$$\left( \sum_{n=0}^{\infty} a_n \right) \left( \sum_{k=0}^{\infty} b_k \right)$$

von zwei Reihen darstellen soll. Wenn wir diesen Ausdruck formal ausmultiplizieren, so erhalten wir die formale Summe

$$s := \sum_{n,k \in \mathbb{N}_0} a_n b_k,$$

in welcher Glieder nicht von einem sondern von zwei Indizes  $n, k$  abhängig sind. Das heißt, die Glieder sind nicht durch aufsteigend geordnete ganze Zahlen indexiert, sondern durch Paare aus  $\mathbb{N}_0^2$ . Für die Paare ganzer Zahlen haben wir aber keine feste Anordnung. Bei der Diskussion der absoluten Konvergenz haben wir aber festgestellt, dass die Anordnung der Glieder im Allgemeinen Auswirkungen auf den Wert der Reihe hat. Das bedeutet, dass man das Produkt von

allgemeinen konvergenten Reihen nicht (sinnvoll) einführen kann. Wenn man sich aber auf die absolut konvergenten Reihen einschränkt, hat man das beschriebene Problem nicht, sodass die formale Summe  $s$  durch eine absolut konvergente Reihe beschrieben werden kann, zum Beispiel so:

$$s := \underbrace{a_0 b_0}_{c_0} + \underbrace{a_0 b_1 + a_1 b_0}_{c_1} + \underbrace{a_0 b_2 + a_1 b_1 + a_2 b_0}_{c_3} + \underbrace{a_3 b_0 + a_2 b_1 + a_1 b_2 + a_0 b_3}_{c_3} + \cdots.$$

Man zerlegt hier die Gesamtsumme der Werte  $a_n b_k$  in Gruppen nach dem Wert von  $n + k$ .

**3.23 Def** (Produkt absolut konvergenter Reihen). Das **Cauchy-Produkt** von zwei absolut konvergenten Reihen  $\sum_{k=0}^{\infty} a_k$  und  $\sum_{k=0}^{\infty} b_k$  ist als die Reihe

$$\sum_{k=0}^{\infty} \underbrace{\left( \sum_{i=0}^k a_i b_{k-i} \right)}_{c_k}$$

mit den Gliedern  $c_0, c_1, c_2, \dots$  definiert.

**3.24 Thm.** Das Cauchy-Produkt von absolut konvergenten Reihen  $\sum_{k=0}^{\infty} a_k$  und  $\sum_{k=0}^{\infty} b_k$  ist eine absolut konvergente Reihe, für welche die Gleichung

$$\sum_{k=0}^{\infty} \left( \sum_{i=0}^k a_i b_{k-i} \right) = \left( \sum_{k=0}^{\infty} a_k \right) \left( \sum_{k=0}^{\infty} b_k \right)$$

erfüllt ist.

**3.25 Bsp.** Jede konvergente geometrische Reihe ist absolut konvergent.

**3.26 Bsp** (Quadrat einer geometrischen Reihe.). Es gilt  $\sum_{k=0}^{\infty} 2^{-k} = 2$ . Da diese Reihe absolut konvergent ist (das ist klar - die Glieder sind nichtnegativ), können wir das Quadrat dieser Reihe als das Cauchy-Produkt der Reihe mit sich selbst darstellen. Man hat

$$2^{-0}2^{-0} = 1 \cdot 2^{-0},$$

$$2^{-0}2^{-1} + 2^{-1}2^{-0} = 2 \cdot 2^{-1},$$

$$2^{-0}2^{-2} + 2^{-1}2^{-1} + 2^{-2}2^{-0} = 3 \cdot 2^{-2}$$

⋮

Es gilt also

$$4 = \left( \sum_{k=0}^{\infty} 2^{-k} \right)^2 = \sum_{k=0}^{\infty} (k+1)2^{-k}.$$

Das gibt uns die interessante Formel

$$\sum_{k=1}^{\infty} k2^{-k} = 2.$$

**3.27 Aufgabe.** Finden Sie eine Formel für  $\sum_{k=0}^{\infty} kq^k$  im Fall  $|q| < 1$ .

**3.28 Aufgabe.** In einem Casino zahlen Sie für ein Spiel den Eintrittspreis von  $g$  Euro und dürfen dann eine faire Münze solange werfen, bis sie eine Zahl geworfen haben. Haben Sie die Zahl das erste mal beim  $k$ -ten Wurf erhalten, dann zahlt Ihnen das Casino  $k$  Euro. Wenn Ihre Würfe zum Beispiel Kopf, Kopf, Kopf, Kopf, Kopf, Zahl waren, erhalten Sie 6 Euro.

Was ist der faire Preis für dieses Spiel? Oder anders formuliert: bestimmen Sie für dieses Spiel den Wert  $g^*$  derart, dass das Casino beim Eintrittspreis  $g > g^*$  Geld (im Durchschnitt) gewinnt und beim Eintrittspreis  $g < g^*$  Geld (im Durchschnitt) verliert.

Hinweise: Die Wahrscheinlichkeitstheorie hatten Sie in der Schule. Es gibt mehrere Weisen, diese Aufgabe zu lösen, unter anderem eine sehr direkte Weise, die auf den oben diskutierten Formeln für Reihen basiert.

**3.29 Def.** Eine Reihe  $\sum_{k=1}^{\infty} a_k$  mit  $a_k a_{k+1} \leq 0$  für alle  $k \geq 0$  nennt man **alternierend**.

**3.30 Bsp.** Die Reihe  $\sum_{k=1}^{\infty} \frac{(-1)^k}{k}$  ist ein Standardbeispiel einer alternierenden Reihe.

**3.31.** Bis auf das Vorzeichen kann jede alternierende reellwertige Reihe als

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

mit  $a_k \geq 0$  für alle geschrieben werden. Der Faktor  $(-1)^k$  im Glied der Reihe sorgt also für die Wahl des Vorzeichens des aktuellen Glieds  $(-1)^k a_k$ , während der Faktor  $a_k$  den Betrag des Glieds bestimmt.

**3.32 Thm** (Leibniz-Kriterium). *Ist  $(a_k)_{k \in \mathbb{N}}$  eine monoton fallende Nullfolge nicht-negativer reeller Zahlen, dann ist die alternierende Reihe*

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

*konvergent. Darüber hinaus wird die Summe  $s$  dieser Reihe durch die Partialsummen  $s_n$  mit der folgenden Präzision approximiert:*

$$|s - s_n| \leq a_{n+1}.$$

**3.33.** Die Glieder der alternierenden Reihe “schwanken” um den Nullwert. Das Leibniz-Kriterium besagt, dass im Fall von Schwankungen, die immer geringer werden und im Grenzwert verschwinden die zugrundeliegende Reihe konvergent ist.

**3.34 Def.** Die Reihe  $\sum_{k=0}^{\infty} b_k$  mit nichnegativen Gliedern  $b_k$  nennt man eine Majorante der Reihe  $\sum_{k=0}^{\infty} a_k$ , wenn für ein  $k_0$  die Ungleichung  $|a_k| \leq b_k$  für alle  $k \geq k_0$  erfüllt ist.

**3.35 Thm** (Majorantenkriterium). *Eine Reihe, die eine konvergente Majorante besitzt, ist absolut konvergent.*

**3.36 Def.** Die Reihe  $\sum_{k=0}^{\infty} b_k$  mit nichnegativen Gliedern  $b_k$  nennt man eine Minorante der Reihe  $\sum_{k=0}^{\infty} a_k$ , wenn für ein  $k_0$  die Ungleichung  $a_k \geq b_k$  für alle  $k \geq k_0$  erfüllt ist.

**3.37 Bsp.** Wir analysieren die Reihe  $\sum_{k=1}^{\infty} \frac{k}{3^k}$  auf Konvergenz:

$$\sum_{k=1}^{\infty} \frac{k}{3^k} \leq \sum_{k=1}^{\infty} \frac{2^k}{3^k} = \sum_{k=1}^{\infty} \left(\frac{2}{3}\right)^k < \infty.$$

$\Rightarrow$  die Reihe  $\sum_{k=1}^{\infty} \frac{k}{3^k}$  ist konvergent.

**3.38 Thm** (Minorantenkriterium). *Eine Reihe, die eine divergente Minorante besitzt, ist divergent.*

**3.39 Bsp.** Wir analysieren die Reihe  $\sum_{k=1}^{\infty} \frac{1}{\ln k}$  auf Konvergenz.

$$\sum_{k=1}^{\infty} \frac{1}{\ln k} \geq \sum_{k=1}^{\infty} \frac{1}{k} = \infty.$$

$\Rightarrow$  die Reihe  $\sum_{k=1}^{\infty} \frac{1}{\ln k}$  ist divergent.

**3.40 Thm** (Quotientenkriterium für Konvergenz). *Gibt es für  $\sum_{k=0}^{\infty} a_k$  Werte  $k_0 \in \mathbb{N}$  und  $0 \leq q < 1$  derart, dass  $\frac{|a_{k+1}|}{|a_k|} \leq q$  für alle  $k \geq k_0$  erfüllt ist, dann ist diese Reihe absolut konvergent.*

**3.41 Bsp.**  $\sum_{k=1}^{\infty} \underbrace{\frac{k}{3^k}}_{=a_k}$  ist nach dem Quotientenkriterium konvergent, denn

$$\frac{|a_{k+1}|}{|a_k|} = \frac{(k+1)3^k}{3^{k+1}k} = \left(1 + \frac{1}{k}\right) \frac{1}{3} \leq \underbrace{\frac{2}{3}}_{=q} < 1.$$

**3.42 Thm** (Quotientenkriterium für Divergenz). *Gibt es für  $\sum_{k=0}^n a_k$  Werte  $k_0 \in \mathbb{N}$  und  $q > 1$  derart, dass  $\frac{|a_{k+1}|}{|a_k|} \geq q$  für alle  $k \geq k_0$  erfüllt ist, dann ist diese Reihe divergent.*

**3.43 Thm** (Quotientenkriterium für Konvergenz in der Grenzwertform). *Gilt für die Reihe*

$\sum_{k=0}^n a_k$  die Ungleichung

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

*dann ist diese Reihe absolut konvergent.*

**3.44 Bsp.**  $\sum_{k=1}^{\infty} \underbrace{\frac{k}{q^k}}_{=a_k}$  ist nach dem Quotientenkriterium konvergent für jedes  $0 \leq q < 1$ , denn

$$\lim_{k \rightarrow \infty} \frac{|a_{k+1}|}{|a_k|} = \lim_{k \rightarrow \infty} \frac{kq^{k+1}}{(k+1)q^k} = \lim_{k \rightarrow \infty} \left(1 - \frac{1}{k+1}\right) q = q \in [0, 1).$$

**3.45 Thm** (Quotientenkriterium für die Divergenz in der Grenzwertform). *Gilt für die Reihe*

$\sum_{k=0}^n a_k$  die Ungleichung

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

*dann ist diese Reihe divergent.*

**3.46.** Bei  $\lim_{k \rightarrow \infty} \frac{|a_{k+1}|}{|a_k|} = 1$  kann man a priori nichts sagen:

Die Reihe  $\sum_{k=1}^{\infty} \frac{1}{k}$  ist divergent und

$\sum_{k=1}^{\infty} \frac{1}{k^2}$  ist konvergent.

# **Kapitel II**

## **Grenzwerte und Stetigkeit von Funktionen**

### **1 Grenzwert einer Funktion**

**1.1 Def.** Der topologische **Abschluss**  $\overline{M}$  einer Teilmenge  $M$  der reellen Zahlen ist die Menge der Grenzwerte konvergenter Folgen  $(x_n)_{n \in \mathbb{N}}$  mit der Eigenschaft  $x_n \in M$  für alle  $n \in \mathbb{N}$ . Eine Menge  $M$  mit  $M = \overline{M}$  nennt man **abgeschlossen**.

**1.2.**  $M$  ist stets eine Teilmenge von  $\overline{M}$ .

**1.3 Bsp.**

- (a)  $[0, 1]$  ist abgeschlossen. Mit anderen Worten: wenn eine Folge  $(x_n)_{n \in \mathbb{N}}$  mit Werten aus  $[0, 1]$  einen Grenzwert besitzt, dann liegt dieser Grenzwert ebenfalls in  $[0, 1]$ .
- (b)  $[0, 1)$  ist nicht abgeschlossen, denn  $\overline{[0, 1)} = [0, 1]$ .
- (c) Für  $M = \left\{ \frac{1}{t} : t \in \mathbb{N} \right\}$  gilt  $\overline{M} = \{0\} \cup M$ .

**1.4 Def.** Sei  $p \in \mathbb{R}$  und  $M$  Teilmenge von  $\mathbb{R}$ . Man nennt  $p$  **Häufungspunkt** von  $M$ , wenn  $p$  im Abschluss von  $M \setminus \{p\}$  liegt.

**1.5 Bsp.**

- (a) Jeder Punkt von  $[0, 1]$  ist ein Häufungspunkt von  $[0, 1]$ .
- (b) 0 ist der einzige Häufungspunkt der Menge  $M = \{1/t : t \in \mathbb{N}\}$ . Die Elemente von  $M$  häufen sich nur in 0, und nirgendwo sonst.
- (c) Alle Punkte aus  $[0, 1]$  sind Häufungspunkte von  $M = [0, 1] \cup \{2\}$ . Es gibt keine anderen Häufungspunkte. Insbesondere ist der Punkt 2 in  $M$  kein Häufungspunkt von  $M$ .

**1.6 Def.** Sei  $f : X \rightarrow \mathbb{R}$  Funktion und  $a$  Häufungspunkt von  $X$ . Man sagt, dass  $f$  für  $x \rightarrow a$  gegen einen reellen Wert  $y$  konvergiert, wenn für jedes  $\epsilon > 0$  ein  $\delta > 0$  existiert, für welches  $|f(x) - y| < \epsilon$  für alle  $x \in X$  mit  $0 < |x - a| < \delta$  erfüllt ist. Schreibweise:  
 $\lim_{x \rightarrow a} f(x) = y$ .

**1.7 Bsp.** Die Funktion  $f(x) = \frac{\sin x}{x}$  ist auf  $\mathbb{R} \setminus \{0\}$  definiert. Die Null ist der Häufungspunkt des Definitionsbereiches. Es gilt:

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1.$$

**1.8 Bsp.**  $\lim_{x \rightarrow 0} x \sin(1/x) = 0$ .

**1.9 Def.** Eine Funktion  $f : X \rightarrow \mathbb{R}$  heißt

- (a) **nach oben beschränkt**, wenn für eine Konstante  $C \in \mathbb{R}$  die Ungleichung  $f(x) \leq C$  für alle  $x \in X$  erfüllt ist.
- (b) **nach unten beschränkt**, wenn für eine Konstante  $c \in \mathbb{R}$  die Ungleichung  $f(x) \geq c$  für alle  $x \in X$  erfüllt ist.
- (c) **beschränkt**, wenn  $f$  nach oben und unten beschränkt ist.

**1.10 Thm.** Sei  $f : X \rightarrow \mathbb{R}$  Funktion und  $a$  Häufungspunkt von  $f$ . Konvergiert  $f$  gegen  $y$  für  $x \rightarrow a$ , so gilt  $\lim_{n \rightarrow \infty} f(x_n) = y$  für jede Folge  $(x_n)_{n \in \mathbb{N}}$  mit  $x_n \in X \setminus \{a\}$  und  $\lim_{n \rightarrow \infty} x_n = a$ .

**1.11 Bsp.** Die Funktion  $\sin(1/x)$  besitzt weder einen linksseitigen noch einen rechtseitigen Grenzwert im Punkt 0, denn:

Für  $x_n = \frac{1}{2\pi n}$ , gilt

$$\lim_{n \rightarrow \infty} x_n = 0,$$

$$\lim_{n \rightarrow \infty} \sin(1/x_n) = 0.$$

Für  $x_n = \frac{1}{2\pi n + \frac{\pi}{2}}$ , gilt

$$\lim_{n \rightarrow \infty} x_n = 0,$$

$$\lim_{n \rightarrow \infty} \sin(1/x_n) = 1.$$

## 2 Stetige Funktionen und Unstetigkeiten

**2.1 Def.** Sei  $f : X \rightarrow \mathbb{R}$  und ei  $a \in X$  Häufungspunkt von  $X$ . Mann nennt  $f$  **stetig** in  $a$ , wenn  $\lim_{x \rightarrow a} f(x) = f(a)$  gilt. Man nennt  $f$  eine stetige Funktion, wenn die Funktion  $f$  in jedem Punkt ihres Definitionsbereichs  $X$  stetig ist.

**2.2 Thm.** Ist  $f : X \rightarrow \mathbb{R}$  eine stetige Funktion so gilt für jede konvergente Folge  $(x_n)_{n \in \mathbb{N}}$  mit  $x_n \in X$  für alle  $n$  und

$$\lim_{n \rightarrow \infty} f(x_n) = f\left(\lim_{n \rightarrow \infty} x_n\right).$$

**2.3 Thm.** *Komposition, Produkt, Quotient und Summe stetiger Funktion sind stetige Funktionen.*

**2.4 Thm.** *Die folgenden Funktionen sind stetig:*

- (a) *Polynomialfunktionen und ihre Quotienten.*
- (b) *Funktionen  $f(x) = x^a$  mit  $a \in \mathbb{R}$  auf  $(0, +\infty)$*
- (c) *Exponentialfunktionen  $f(x) = a^x$  mit  $a > 0$ ,*
- (d) *Logarithmische Funktionen  $f(x) = \log_a x$  mit  $a > 0$  und  $a \neq 1$ ,*
- (e) *Sinus, Kosinus, Tangens*
- (f) *Arcus Sinus, Arcus Kosinus, Arcus Tangens.*

**2.5 Def** (Rechtsseitiger Grenzwert einer Funktion). Sei  $f : X \rightarrow \mathbb{R}$  Funktion auf  $X \subseteq \mathbb{R}$  und  $a \in \mathbb{R}$  Häufungspunkt von  $X \cap [a, +\infty)$ . Man nennt  $f$  konvergent für  $x \rightarrow a+$ , wenn ein Wert  $y \in \mathbb{R}$  existiert mit der Eigenschaft, dass für jedes  $\epsilon > 0$  ein  $\delta > 0$  existiert, sodass für alle  $x \in X$  mit  $a < x < a + \delta$  die Ungleichung  $|f(x) - y| < \epsilon$  erfüllt ist.

Den Wert nennt man den rechtsseitigen Grenzwert und bezeichnet als

$$y = \lim_{x \rightarrow a^+} f(x).$$

**2.6 Def** (Linksseitiger Grenzwert einer Funktion). Sei  $f : X \rightarrow \mathbb{R}$  Funktion auf  $X \subseteq \mathbb{R}$  und  $a \in \mathbb{R}$  ein Häufungspunkt von  $X \cap (-\infty, a]$ . Der rechtsseitiger Grenzwert

$$y = \lim_{x \rightarrow a^-} f(x)$$

ist ein Wert  $y$  mit der Eigenschaft, dass für jedes  $\epsilon > 0$  ein  $\delta > 0$  existiert, sodass für alle  $x \in X$  mit  $a - \delta < x < a$  die Ungleichung  $|f(x) - y| < \epsilon$  erfüllt ist.

**2.7.** Alternative Bezeichnungen für einseitige Konvergenz:

$x \rightarrow a+$  schreibt man auch als  $x \downarrow a$  und

$x \rightarrow a-$  schreibt man auch als  $x \uparrow a$ .

**2.8.** Bestimmte Divergenz gegen  $\infty$  und  $-\infty$  für  $x \rightarrow a+$  und  $x \rightarrow a^-$  kann analog zur bestimmten Divergenz für  $x \rightarrow a$  eingeführt werden.

**2.9 Thm** (Beschreibung der Konvergenz über die rechts- und linksseitige Konvergenz). Sei  $f : X \rightarrow \mathbb{R}$  und  $a$  ein Häufungspunkt von  $X \cap [a, +\infty)$  und  $(-\infty, a] \cap X$ . Dann sind die folgenden Bedingungen äquivalent:

(i)  $f(x)$  ist konvergent für  $x \rightarrow a$ .

(ii)  $f(x)$  ist konvergent für  $x \rightarrow a+$  und für  $x \rightarrow a-$ , und es gilt  $\lim_{x \rightarrow a+} f(x) = \lim_{x \rightarrow a-} f(x)$ .

Gegebenenfalls gilt  $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a+} f(x) = \lim_{x \rightarrow a-} f(x)$ .

**2.10 Bsp.**

$$\lim_{x \rightarrow 0^+} \arctan \frac{1}{x} = \frac{\pi}{2},$$
$$\lim_{x \rightarrow 0^-} \arctan \frac{1}{x} = -\frac{\pi}{2}.$$

$\Rightarrow f(x)$  ist divergent für  $x \rightarrow 0$ .

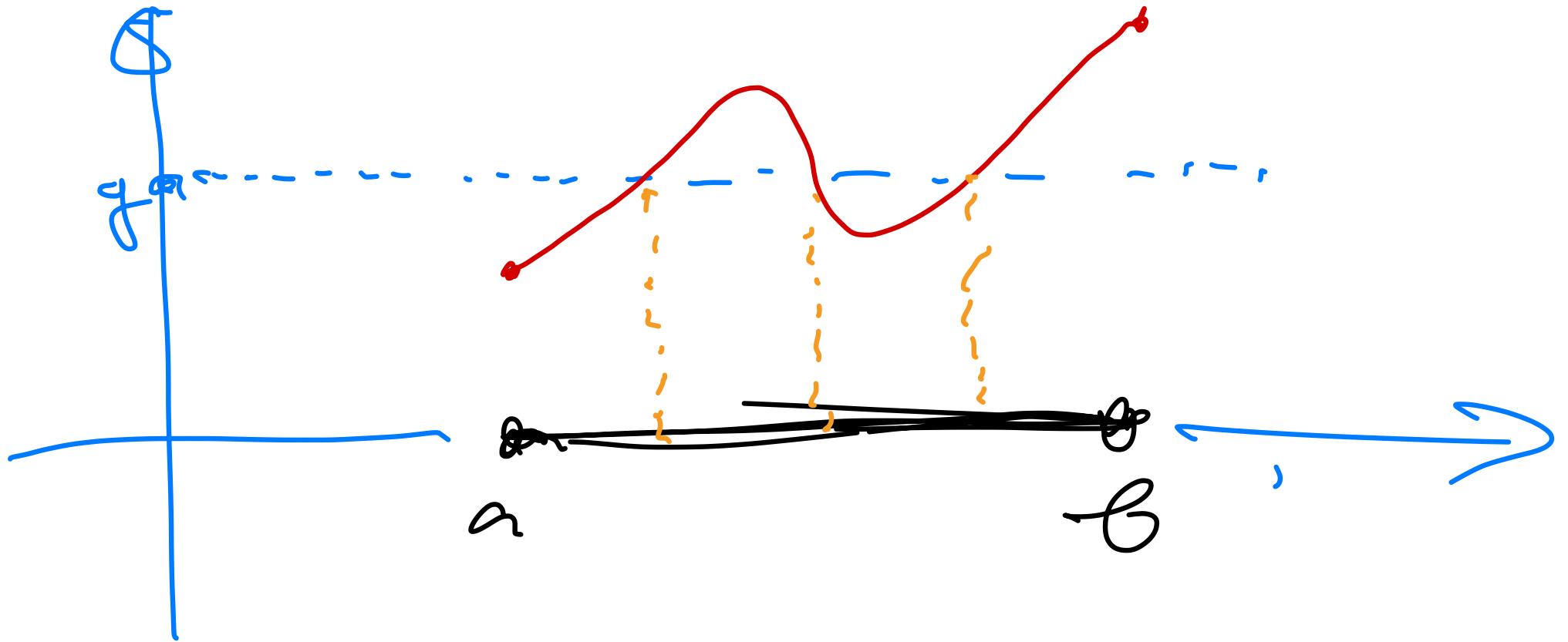
**2.11 Def.** Eine Menge  $M \subseteq \mathbb{R}$ , die beschränkt und abgeschlossen ist, nennt man **kompakt**.

**2.12 Bsp.**

- (a)  $[a, b]$  ist kompakt für alle  $a, b \in \mathbb{R}$  mit  $a \leq b$ .

**2.13 Thm** (Satz von Weierstraß). *Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine stetige Funktion auf einem kompakten Intervall. Dann erreicht die Funktion  $f$  auf  $[a, b]$  ihr Minimum und Maximum. Das heißt, es gibt  $s, t \in X$  mit  $f(s) \leq f(x) \leq f(t)$  für alle  $x \in [a, b]$ .*

**2.14 Thm (Zwischenwertsatz).** Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetige Funktion auf einem Intervall  $[a, b]$  mit  $a, b \in \mathbb{R}$  und  $a < b$ . Sei  $f(a) \leq y \leq f(b)$  oder  $f(b) \leq y \leq f(a)$ . Dann existiert ein  $\xi \in [a, b]$  mit  $f(\xi) = y$ .



**2.15.** Der Beweis vom Zwischenwertsatz basiert auf einem konstruktiven Ansatz. Man kann den Suchraum für  $\xi$  iterativ halbieren und ein  $\xi$  wie in der Behauptung beliebig gut approximieren.

```
def zero_through_binary_search(f,a,b,xerr=0.00001):
```

```
    """
```

*Sei  $f$  stetige Funktion, die in den Endpunkten von  $[a,b]$  das Vorzeichen wechselt.*

*Man berechnet eine Approximation eines  $x_i$  mit  $f(x_i)=0$  mit der Fehlertoleranz  $xerr$ .*

*Umsetzung mit dem Gedanken im Hinterkopf, dass ein Auswertung von  $f$  sehr teuer sein*

*kann (etwa eine Simulation, die man durchlaufen muss, und die mehrere Minuten in Anspruch nehmen kann)*

```
    """
```

```
assert(a<b) # a und b sollen ein Segment definieren
```

```
# setup
```

```
left=a
```

```
right=b
```

```
f_left=f(left)
f_right=f(right)

# Am anfang sollen wir ein Vorzeichenwechsel haben
assert( f_left * f_right <= 0 )

while right-left>xerr:
    middle = 0.5 * (left+ right)
    f_middle = f(middle)
    if f_left*f_middle<=0:
        right=middle
        f_right=f_middle
    else:
        left=middle
        f_left=f_middle

return left
```

```
f = lambda x : sin(x)+1 - x
a = 0.0
b=5.0

x=SR('x')
print(" Nullstelle dieser Funktion wird numerisch ermittelt:")
show(plot(f, (x,a,b)))

print(" Approximation einer Nullstelle mit Hilfe der binären Suche:",
      zero_through_binary_search(f, a, b))
```

So ein algorithmischer Ansatz (der mit der binären Suche in sortierten Arrays verwandt ist) benötigt nur eine schwache Annahme über  $f$  (Stetigkeit). Auch wenn  $f$  nicht durch eine explizite Formel verfügbar ist sondern als ein sogenanntes Orakel zur Verfügung steht, mit dem man für ein gegebenes  $x \in [a, b]$  den Wert von  $f$  an der Stelle  $x$  ermitteln kann, kann der Ansatz benutzt werden.

### **3 Asymptotisches Verhalten im Unendlichen**

**3.1 Def.** Eine Teilmenge  $M$  von  $\mathbb{R}$  heißt

- (a) **nach oben beschränkt**, wenn für eine Konstante  $C \in \mathbb{R}$  die Ungleichung  $x \leq C$  für alle  $x \in M$  erfüllt ist.
- (b) **nach unten beschränkt**, wenn für eine Konstante  $c \in \mathbb{R}$  die Ungleichung  $x \geq c$  für alle  $x \in M$  erfüllt ist.
- (c) **beschränkt**, wenn  $M$  nach oben und nach unten beschränkt ist.

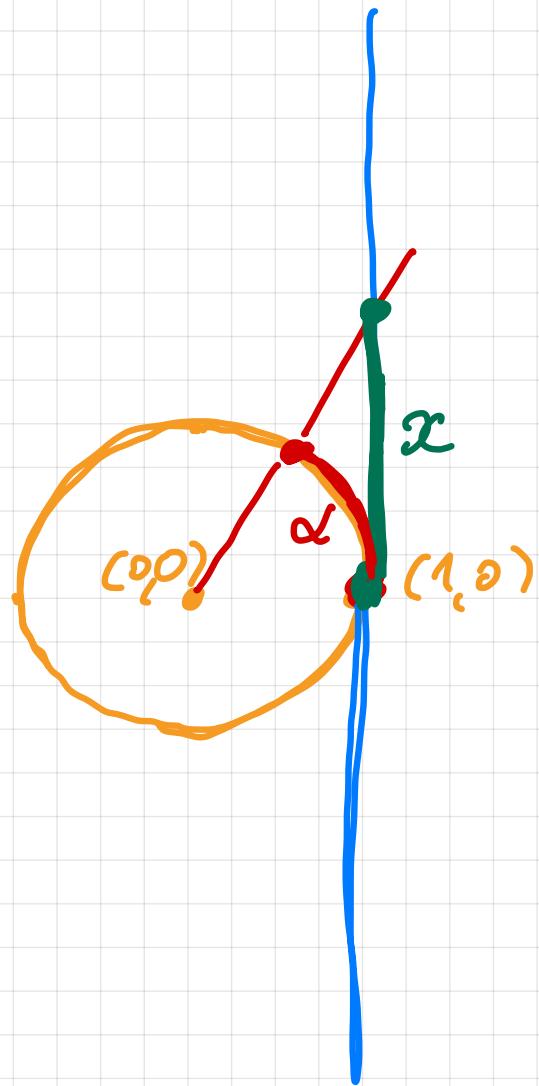
**3.2 Def.** Sei  $f : X \rightarrow \mathbb{R}$  Funktion auf einer Menge  $X$ , die nach oben nicht beschränkt ist. Man nennt  $y$  den Grenzwert der Funktion  $f(x)$  für  $x \rightarrow \infty$ , wenn für jedes  $\epsilon > 0$  ein  $N \in \mathbb{R}$  existiert, für welches die Ungleichung  $|f(x) - y| \leq \epsilon$  für alle  $x \in X$  mit  $x \geq N$  erfüllt ist. Schreibweise:  $\lim_{x \rightarrow \infty} f(x) = y$ .

**3.3 Def.** Sei  $f : X \rightarrow \mathbb{R}$  Funktion auf einer Menge  $X$ , die nach unten nicht beschränkt ist. Man nennt  $y$  den Grenzwert der Funktion  $f(x)$  für  $x \rightarrow -\infty$ , wenn für jedes  $\epsilon > 0$  ein  $M \in \mathbb{R}$  existiert, für welches die Ungleichung  $|f(x) - y| \leq \epsilon$  für alle  $x \in X$  mit  $x \leq M$  erfüllt ist. Schreibweise:  $\lim_{x \rightarrow -\infty} f(x) = y$ .

**3.4 Bsp.**

- (a)  $\lim_{x \rightarrow \infty} \left(1 + \frac{1}{x}\right)^x = e$ . Wir hatten bereits ein solches Beispiel mit einer Variablen  $n$  aus  $\mathbb{N}$ . Nun haben wir eine Variable  $x$ , die innerhalb der reellen Zahlen gegen  $\infty$  geht.
- (b)  $\lim_{x \rightarrow \infty} \arctan x = \frac{\pi}{2}$
- (c)  $\lim_{x \rightarrow -\infty} \arctan x = -\frac{\pi}{2}$

$$\alpha = \arctan x$$



# **Kapitel III**

## **Differentialrechnung I**

Differentialrechnung für Funktionen einer Variablen

# 1 Ableitung

**1.1 Def (Ableitung).** Sei  $f : X \rightarrow \mathbb{R}$  und  $a \in X$  Häufungspunkt von  $X$ . Man nennt  $f$  **differenzierbar** in  $a$ , wenn ein endlicher Grenzwert

$$f'(a) := \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

existiert. Man nennt  $f'(a)$  die **Ableitung** von  $f$  an der Stelle  $a$ . Somit ist die **Ableitung**  $f'$  eine Funktion auf der Menge aller  $a \in X$ , in denen die Funktion  $f$  differenzierbar ist.

$$x \rightarrow a \quad \Leftrightarrow \quad h \rightarrow 0$$

$$h := x - a$$

**1.2.** Wenn man  $h = x - a$  an der Stelle von  $x$  benutzt, kann man die Formel für die Ableitung als

$$f'(a) = \lim_{h \rightarrow 0} \frac{f(a + h) - f(a)}{h}$$

umschreiben.

Man schreibt auch in manchen Quellen so:

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{\Delta f}{\Delta x}$$

mit  $\Delta f := f(x + \Delta x) - f(x)$ . Hierbei ist

$\Delta x$  eine (beliebig klein werdende) Änderung von  $x$  und

$\Delta f$  die entsprechende Änderung von  $f$  bzgl. einer festen Stelle  $x$ .

$$\begin{array}{lcl} a & \leftrightarrow & x \\ h & \leftrightarrow & \Delta x \end{array}$$

$$f'(x) = \lim_{\Delta x \rightarrow 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

**1.3** (Leibniz-Notation für die Ableitung und Differentiale). Leibniz hat die Ableitung als einen formalen Quotienten geschrieben

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

Motivation zu dieser Bezeichnung: die Ableitung ist der Grenzwert eines Quotienten. Diese Gleichung kann man auch formal auch als

$$d f = f'(x) d x.$$

Hierbei nennt man  $d f$  Differential von  $f$  und  $d x$  Differential von  $x$ . Das sind formale Symbole. Die intuitive Bedeutung dieser Gleichung ist:

$$\Delta f \approx f'(x) \Delta x.$$

Das Vorige Bedeutet:  $\Delta f = f'(x) \Delta x + o(\Delta x)$  mit  $\frac{o(\Delta x)}{\Delta x} \rightarrow 0$  für  $\Delta x \rightarrow 0$ .

$f$  mit der Ableitung  $f'(a)$

Wir sind auf der Suche nach einer Funktion  $Ax + B$ ,  
die an  $a$  den selben Wert und die gleiche  
Ableitung als  $f$  hat.

Die Funktion:

$$f(a) + f'(a) \cdot (x - a)$$

**1.4** (Approximation, Tangente). Die (affin)lineare Funktion  $f(a) + f'(a)(x - a)$  hat den gleichen Wert und die gleiche Ableitung in  $a$  wie die Funktion  $f$ . Mit Hilfe des Wertes  $f(a)$  von  $f$  an der Stelle  $f$  und der Ableitung von  $f$  an der Stelle  $a$  erhält man eine Approximation der Funktion  $f$  in einer kleinen Umgebung der Stelle  $a$ .

Geometrisch gesehen beschreibt  $y = f(a) + f'(a)(x - a)$  den Graphen der Tangente zum Graphen von  $f$  an der Stelle  $(a, f(a))$ .

**1.5.** Einige physikalische Bedeutungen der Ableitung:

- (a) Ist  $f(t) \in \mathbb{R}$  die Position eines Objekts zum Zeitpunkt  $t \in \mathbb{R}$ , dann ist  $f'(t)$  die Geschwindigkeit des Objekts im Zeitpunkt  $t$ .
- (b) Hat man auf der reellen Achse einen Stab  $[0, L]$  der Länge  $L > 0$ , bei dem die Masse des Abschnitts  $[0, x]$  für  $0 < x < L$  gleich  $m(x)$  ist, dann ist die Ableitung  $\rho(x) = m'(x)$  die (lineare) Dichte im Punkt  $x$ .

**1.6 Bsp.** Wir berechnen einige Ableitungen direkt aus der Definition.

(a) Quadratische Funktion:

$$(x^2)' = \lim_{h \rightarrow 0} \frac{(x+h)^2 - x^2}{h} = \lim_{h \rightarrow 0} \frac{2xh + h^2}{h} = \lim_{h \rightarrow 0} (2x + h) = 2x.$$

(b) Quadratische Wurzel:

$$(\sqrt{x})' = \lim_{h \rightarrow 0} \frac{\sqrt{x+h} - \sqrt{x}}{h}$$

Die Wurzel im Nenner stört, daher wir der Bruch mit  $\sqrt{x+h} + \sqrt{x}$  ergänzt, um von der Wurzel im Zähler loszuwerden. Mit der Verwendung der dritten binomischen Formel erhalten wir dann

$$(\sqrt{x})' = \lim_{h \rightarrow 0} \frac{(x+h) - x}{h(\sqrt{x+h} + \sqrt{x})} = \lim_{h \rightarrow 0} \frac{1}{\sqrt{x+h} + \sqrt{x}} = \frac{1}{2\sqrt{x}}.$$

Hierbei sollen wir voraussetzen, dass man  $x > 0$  hat.

$$\begin{aligned}(x^3)' &= \lim_{h \rightarrow 0} \frac{(x+h)^3 - x^3}{h} = \\&= \lim_{h \rightarrow 0} \frac{\cancel{x^3} + 3x^2h + 3xh^2 + h^3 - \cancel{x^3}}{h} \\&= \lim_{h \rightarrow 0} \frac{3x^2h + 3xh^2 + h^3}{h} \\&= \lim_{h \rightarrow 0} (3x^2 + 3xh + h^2) \\&= 3x^2.\end{aligned}$$

(c) Die Funktion  $1/x$ :

$$(1/x)' = \lim_{h \rightarrow 0} \frac{1/(x+h) - 1/x}{h}.$$

Es bietet sich an, den Quotienten unter dem Grenzwert durch die Erweiterung mit  $(x+h)x$  zu vereinfachen:

$$(1/x)' = \lim_{h \rightarrow 0} \frac{x - (x+h)}{h(x+h)x} = \lim_{h \rightarrow 0} -\frac{1}{(x+h)x} = -\frac{1}{x^2}.$$

Hierbei sollen wir  $x \neq 0$  voraussetzen.

**1.7 Thm.** Jede differenzierbare Funktion ist stetig, aber im Allgemeinen nicht umgekehrt.

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} (f(x) - f(a)) + f(a)$$

$$= \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \cdot (x - a) + f(a)$$

(Rückwärtsregel)

$$= \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \cdot \lim_{x \rightarrow a} (x - a) + f(a)$$

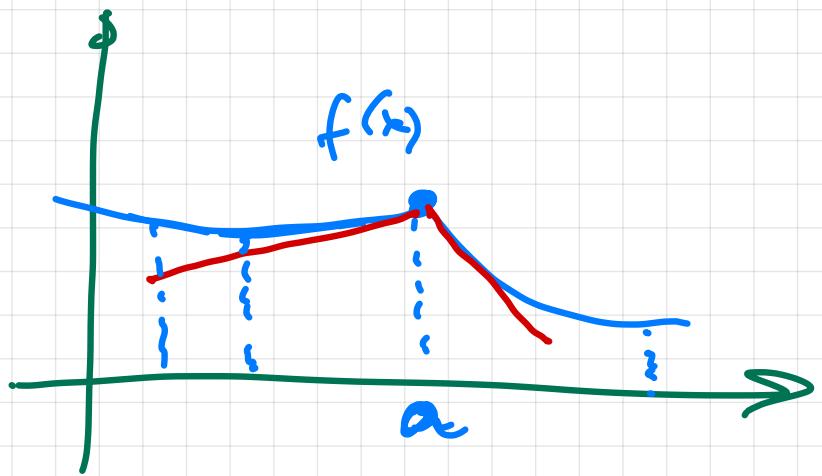
$$= f'(a) \cdot 0 + f(a) = f(a)$$

**1.8.** Begründung zur Stetigkeit einer differenzierbaren Funktion ist wie folgt. Ist  $f$  differenzierbar in  $a$  so ist

$$f'(a) = \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

ein endlicher Wert. Dann ist

$$\begin{aligned}\lim_{x \rightarrow a} f(x) &= \lim_{x \rightarrow a} (f(x) - f(a)) + f(a) \\ &= \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} (x - a) + f(a) \\ &= \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} \lim_{x \rightarrow a} (x - a) + f(a) \\ &= f'(a) \cdot 0 + f(a) \\ &= f(a)\end{aligned}$$



**1.9 Bsp.**  $|x|$  ist stetig aber in 0 nicht differenzierbar. Es gibt Beispiele von Funktionen, die auf  $\mathbb{R}$  stetig aber an keiner Stelle differenzierbar sind.

**1.10 Thm** (Rechenregeln für Differenzierbarkeit). Für differenzierbare Funktionen  $f$  und  $g$  gelten die folgenden Regeln:

(a) *Linearität*:  $(\alpha f + \beta g)' = \alpha f' + \beta g'$  mit  $\alpha, \beta \in \mathbb{R}$

(b) *Produktregel*:  $(fg)' = f'g + fg'$

(c) *Quotientenregel*:  $\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$

(d) *Kettenregel*:  $(f \circ g)' = (f' \circ g)g'$ .

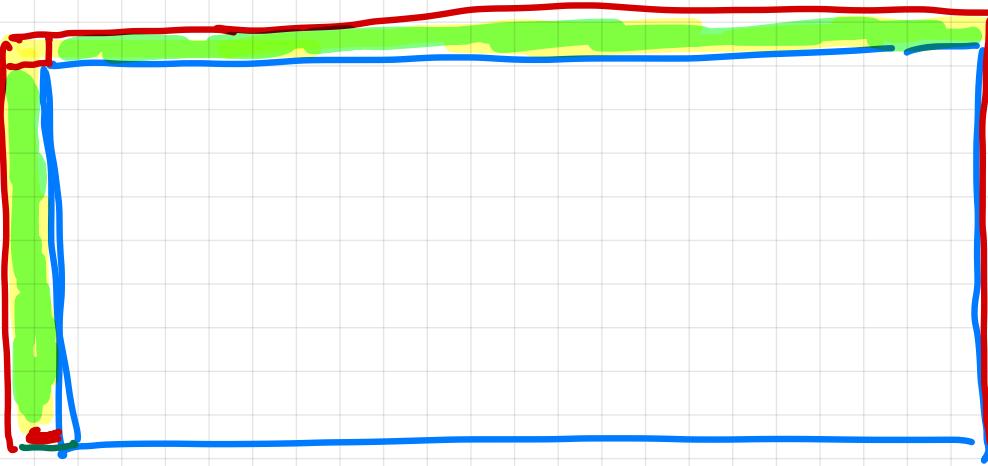
$$1000 \times 0,9$$

$$1001 \times 0,92 = (1000+1) \times (0,90+0,02)$$

$$= 1000 \cdot 0,9$$

$$+ 1000 \cdot 0,02 + 1 \cdot 0,90$$

$$+ 1 \cdot 0,02$$



**1.11.** Intuition hinter der Produktregel:

$$\begin{aligned}\Delta(fg) &= (f + \Delta f)(g + \Delta g) - fg \\ &= (\Delta f)g + f(\Delta g) + (\Delta f)(\Delta g).\end{aligned}$$

Wenn  $\Delta f$  und  $\Delta g$  klein sind, dann ist  $(\Delta f)(\Delta g)$  noch kleiner (zu klein). Wenn man durch  $\Delta x$  teilt, und dann  $\Delta x$  gegen 0 schickt, erhält man

$$(fg)' = f'g + fg'.$$

**1.12 Bsp** (zu Produktregel).

$$\begin{aligned}(x^2 \sin x)' &= (x^2)' \sin x + x^2 (\sin x)' && | \text{ Produktregel} \\ &= 2x \sin(x) + x^2 \cos(x). && | \text{ Formeln für } (x^2)' \text{ and } (\sin x)'\end{aligned}$$

**1.13.** Intuition hinter der Quotientenregel:

$$\begin{aligned}\Delta \left( \frac{f}{g} \right) &= \frac{f + \Delta f}{g + \Delta g} - \frac{f}{g} \\&= \frac{(f + \Delta f)g - f(g + \Delta g)}{g(g + \Delta g)} \\&= \frac{(\Delta f)g - f\Delta g}{g(g + \Delta g)}\end{aligned}$$

- | Aus der Def. von  $\Delta$  einer Funktion
- | Als ein Quotient umgeschrieben
- | Klammern aufgelöst,  $fg$  verschwindet.

Wenn man durch  $\Delta x$  teilt und dann  $\Delta x$  gegen 0 schickt, erhält man:

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

**1.14 Bsp** (zu Quotientenregel).

$$\begin{aligned}\left(\frac{\sin x}{x^2}\right)' &= \frac{(\sin x)'x^2 - \sin x(x^2)'}{(x^2)^2} \\ &= \frac{x^2 \cos x - 2x \sin x}{x^4} \\ &= \frac{x \cos x - 2x \sin x}{x^3}.\end{aligned}$$

**1.15 Bsp** (Intuition hinter der Kettenregel an einem Beispiel). Wir berechnen die Ableitung von  $z(x) = \sin(x^2)$ , wie es Leibniz gemacht hätte. Man hat  $z = \sin y$  und  $y = x^2$ .

$$\begin{aligned}\frac{d z}{d x} &= \frac{d z}{d y} \cdot \frac{d y}{d x}. && (\text{Ableitung ist "Quotient", wir ergänzen formal}) \\ &= \left(\frac{d}{d y} \sin y\right) \cdot \left(\frac{d}{d x} x^2\right) \\ &= (\cos y) \cdot (2x) \\ &= (\cos x^2) \cdot (2x) \\ &= 2x \cos x^2.\end{aligned}$$

Wie man es sonst schreibt ist:

$$(\sin x^2)' = (\cos x^2) \cdot (x^2)' = (\cos x^2) \cdot (2x) = 2x \cos x^2.$$

**1.16 Bsp** (Kettenregel für längere Ketten). Hier ein Beispiel, in dem man drei Funktionen ineinander geschachtelt hat (Sinus, e hoch, Quadrat):

$$\begin{aligned}(\sin e^{x^2})' &= (\cos e^{x^2}) \cdot (e^{x^2})' \\&= (\cos e^{x^2}) \cdot e^{x^2} \cdot (x^2)' \\&= (\cos e^{x^2}) \cdot e^{x^2} \cdot (2x).\end{aligned}$$

**1.17 Thm** (Ableitung der Umkehrfunktion). *Seien  $I, J$  Intervalle. Für eine bijektive differenzierbare Funktion  $f : I \rightarrow J$  ist die Umkehrfunktion überall, wo  $f' \circ f^{-1}$  ungleich 0 ist, differenzierbar, und es gilt:*

$$(f^{-1})' = \frac{1}{f' \circ f^{-1}}.$$

**1.18 Bsp** (Leibniz'sche Intuition an einem Beispiel). Wir kennen, dass  $(e^x)' = e^x$  gilt. Die Gleichung

$$y = e^x$$

kann man auch als  $x = \ln y$  um schreiben. Man hat

$$\frac{d y}{d x} = \frac{d}{d x} e^x = e^x.$$

Das schreiben wir wie Leibniz formal um:

$$\frac{d x}{d y} = e^{-x} = (e^x)^{-1} = y^{-1}.$$

Da wir  $x = \ln y$  haben, erhalten wir

$$\frac{d}{d y} \ln y = \frac{1}{y}.$$

Das heißt, aus der Formel für die Ableitung der Exponentialfunktion erhalten wir die Formel für die Ableitung für den Logarithmus.

Nach diesem Prinzip kann man für Umkehrfunktion bekannter Funktionen Formeln für die Ableitungen bestimmen.

**1.19 Thm.** Es gelten die folgenden Formeln für die Ableitungen:

$$(x^a)' = ax^{a-1} \text{ für } a \in \mathbb{R}.$$

$$(e^x)' = e^x$$

$$(\ln x)' = \frac{1}{x}.$$

$$(\sin x)' = \cos x$$

$$(\cos x)' = -\sin x$$

$$(\tan x)' = -\frac{1}{(\cos x)^2}.$$

$$(\arcsin x)' = \frac{1}{\sqrt{1-x^2}}$$

$$(\arccos x)' = -\frac{1}{\sqrt{1-x^2}}$$

$$(\arctan x)' = \frac{1}{1+x^2}.$$

**1.20 Aufgabe.** Bestimmen Sie:

- (a)  $(a^x)'$  aus  $(e^x)' = e^x$  und der Kettenregel.
- (b)  $(x^a)'$  aus  $(e^x)' = e^x$  und der Kettenregel.
- (c)  $(\sin x)'$  aus  $\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$  und den trigonometrischen Formeln.
- (d)  $(\cos x)'$  aus  $(\sin x)' = \cos x$  und den trigonometrischen Formeln.
- (e)  $(\tan x)'$  aus der Quotientenregel und den trigonometrischen Formeln.
- (f) die Formeln für die Ableitung der inversen trigonometrischen Funktionen aus den Formeln für die Ableitung der trigonometrischen Funktionen.

**1.21 Aufgabe.** In der Formel  $(\ln x)' = \frac{1}{x}$  ist  $x$  auf den Fall  $x > 0$ . Im Gegensatz dazu lässt  $\ln |x|$  für alle  $x \in \mathbb{R} \setminus \{0\}$  auswerten? Bestimmen Sie die Ableitung von  $\ln |x|$ . **Hinweis:** obwohl  $|x|$  durch eine Fallunterscheidung definiert ist, hat man eine Formel für die Ableitung von  $\ln |x|$ , die keine Fallunterscheidung benötigt.

**1.22 Thm** (Regel von l'Hospital). *Seien  $f, g : X \rightarrow \mathbb{R}$  differenzierbare Funktionen. Sei  $a \in \mathbb{R}$  Häufungspunkt von  $X$  oder  $a = \infty$  und  $X$  nach oben unbeschränkt oder  $a = -\infty$  und  $X$  nach unten unbeschränkt. Sei  $\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a} g(x) = 0$  oder  $\lim_{x \rightarrow a} |f(x)| = \lim_{x \rightarrow a} |g(x)| = \infty$ .*

*Wenn der Quotient  $\frac{f'(x)}{g'(x)}$  für  $x \rightarrow a$  konvergent ist, dann ist der Quotient  $\frac{f(x)}{g(x)}$  für  $x \rightarrow a$  ebenfalls konvergent ist, und es gilt:*

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \lim_{x \rightarrow a} \frac{f'(x)}{g'(x)}.$$

Bsp.  $\lim_{x \rightarrow 0} \frac{\sin x}{e^x - 1}$

Unbestimpter Ausdruck  $\frac{0}{0}$ .

R.H.  $= \lim_{x \rightarrow 0} \frac{(\sin x)'}{(e^x - 1)'} = \lim_{x \rightarrow 0} \frac{\cos x}{e^x} = \frac{1}{1} = 1$

Bsp.

$\lim_{x \rightarrow 0} \frac{x^2}{1 - \cos x} \stackrel{l.H.}{=} \lim_{x \rightarrow 0} \frac{(x^2)'}{(1 - \cos x)'} \quad$

$= \lim_{x \rightarrow 0} \frac{2x}{\sin x} \stackrel{l.H.}{=} \lim_{x \rightarrow 0} \frac{(2x)'}{(\sin x)'} \quad$

$= \lim_{x \rightarrow 0} \frac{2}{\cos x} = \frac{2}{1} = 2 \quad$

**1.23 Bsp.**

$$(a) \lim_{x \rightarrow 0} \frac{\sin x}{e^x - 1} \stackrel{l'H}{=} \lim_{x \rightarrow 0} \frac{(\sin x)'}{(e^x - 1)'} = \lim_{x \rightarrow 0} \frac{\cos x}{e^x} = \frac{1}{1} = 1.$$

$$(b) \lim_{x \rightarrow 0} \frac{x^2}{1 - \cos x} \stackrel{l'H}{=} \lim_{x \rightarrow 0} \frac{(x^2)'}{(1 - \cos x)'} = \lim_{x \rightarrow 0} \frac{2x}{\sin x} \stackrel{l'H}{=} \lim_{x \rightarrow 0} \frac{(2x)'}{(\sin x)'} = \lim_{x \rightarrow 0} \frac{2}{\cos x} = 2.$$

**1.24 Bsp** (Intuition hinter der Regel von L'Hospital an einem Beispiel).

$$\lim_{x \rightarrow 0} \frac{\sin x}{e^x - 1} = \lim_{x \rightarrow 0} \frac{\frac{\sin x - \sin 0}{x - 0}}{\frac{e^x - e^0}{x - 0}} = \frac{\lim_{x \rightarrow 0} \frac{\sin x - \sin 0}{x - 0}}{\lim_{x \rightarrow 0} \frac{e^x - e^0}{x - 0}} = \frac{(\sin x)'|_{x=0}}{(e^x)'|_{x=0}} = \frac{\cos 0}{e^0} = 1.$$

**1.25 Def** (Höhere Ableitungen). Für eine Funktion  $f : X \rightarrow \mathbb{R}$  auf  $X \subseteq \mathbb{R}$ , wird die  $n$ -te Ableitung von  $f$  rekursiv definiert:

$$f^{(n)} := \begin{cases} f, & \text{für } n = 0, \\ (f^{(n-1)})', & \text{für } n \in \mathbb{N}. \end{cases}$$

Das heißt:  $f^{(0)} = f$ ,  $f^{(1)} = f'$ ,  $f^{(2)} = f''$  usw.

$$\begin{array}{ccccc} f^{(0)} & f^{(1)} & f^{(2)} & f^{(3)} & f^{(4)} \\ || & || & || & || & || \\ 1 & f' & f'' & f''' & f'''' \end{array}$$

**1.26.** In Physik, Wirtschaftswissenschaft und anderen Anwendungsbereichen betrachtet man oft zeitliche Veränderungen verschiedener Größen.

Dabei benutzt man oft die Ableitungen der Größen nach der Zeit, um solche Veränderungen formal und/oder quantitativ zu beschreiben. Wenn wir  $f(t)$  betrachten und  $t$  Zeit ist, so ist:

- $f(t)$  - der Wert  $f$  zum Zeitpunkt  $t$ .
- $f'(t)$  - die Geschwindigkeit von  $f(t)$ ,  $f'(t) > 0$  - Wachstum,  $f'(t) < 0$  - Abnahme.
- $f''(t)$  - Beschleunigung von  $f(t)$ .

**1.27.** In den Nachrichten findet man immer wieder Berichte, die indirekt einen Bezug zu höheren Ableitungen haben. Z.B. "Wachstum beschleunigt sich" heißt  $f'(t) > 0$  (Wachstum) und  $f'''(t) > 0$  (Beschleunigung des Wachstums).

**1.28 Bsp.** Weil für  $f(x) = e^x$  man  $f'(x) = e^x = f(x)$  hat, gilt  $f^{(n)}(x) = e^x$  für alle  $n \geq 0$ .

**1.29 Bsp.** Für  $f(x) = x^a$  mit  $a \in \mathbb{R}$  gilt

$$f^{(n)}(x) = a \cdot (a - 1) \cdot \dots \cdot (a - n + 1) x^{a-n}.$$

Insbesondere: für  $a \in \mathbb{N}_0$  hat man  $f^{(a)}(x) = a!$  und  $f^{(n)}(x) = 0$  für  $n > a$ .

**1.30 Bsp.** Für  $f(x) = \sin x$  gilt  $f^{(4)}(x) = f(x)$ . Es folgt:  $f^{(n)}(x)$  4-periodisch in  $n$ . Das heißt,  $f^{(n+4)}(x) = f^{(n)}(x)$ .

**1.31.** Ist  $f(t) \in \mathbb{R}$  die Position eines Objekts auf der reellen Achse im Zeitpunkt  $t \in \mathbb{R}$ , dann ist  $f'(t)$  die Geschwindigkeit im Zeitpunkt  $t$  und  $f''(t)$  die Beschleunigung im Zeitpunkt  $t$ .

Zweite Ableitungen spielen eine wichtige Rolle in der Physik (das zweite Gesetz von Newton: Kraft = Masse  $\times$  Beschleunigung).

## 2 Monotonie und Extremwerte

**2.1 Thm.** Für eine differenzierbare Funktion  $f : I \rightarrow \mathbb{R}$  auf einem Intervall  $I$  sind die folgenden Bedingungen äquivalent:

- (i)  $f$  ist monoton steigend (bzw. fallend)
- (ii)  $f'$  ist nichtnegativ (bzw. nichtpositiv)

**2.2 Def** (globale Extrema). Sei  $f : X \rightarrow \mathbb{R}$ .

Wenn die Bedingung  $f(a) = \max_{x \in X} f(x)$  für ein  $a \in X$  gilt, so sagt man, dass das (globale) **Maximum** von  $f$  im Punkt  $a$  erreicht wird.

Gilt  $f(a) = \min_{x \in X} f(x)$  für ein  $a \in X$ , dann sagt man, dass das (globale) **Minimum** von  $f$  im Punkt  $a$  erreicht wird.

Wird im Punkt  $a \in X$  das Minimum oder Maximum von  $f$  über  $X$  erreicht, so sagt man, dass im Punkt  $a$  ein (globales) **Extremum** der Funktion  $f$  erreicht wird.

**WARNUNG!** Man beachte, dass sich diese Begriffe auf den Definitionsbereich  $X$  der Funktion  $f$  beziehen. Das heißt: wenn man die Begriffe benutzt, soll klar gemacht werden

(oder klar sein), welcher Definitionsbereich  $X$  festgelegt ist.

**2.3 Def.** Für  $\epsilon > 0$  und  $a \in \mathbb{R}$  heißt das Intervall  $(a - \epsilon, a + \epsilon)$  eine (offene)  $\epsilon$ -Umgebung von  $a$ .

**2.4 Def (Lokale Extrema).** Sei  $f : X \rightarrow \mathbb{R}$  Funktion auf einer Teilmenge von  $\mathbb{R}$  und sei  $a \in X$ .

Wenn die Einschränkung von  $f$  auf eine genügend kleine  $\epsilon$ -Umgebung von  $a$  das Maximum bzw. das Minimum bzw. ein Extremum im Punkt  $a$  erreicht, so sagt man, dass  $f$  im Punkt  $a$  ein **lokales Maximum** bzw. ein **lokales Minimum** bzw. ein **lokales Extremum** erreicht.

**2.5 Def.** Sei  $X \subseteq \mathbb{R}$ . Dann heißt  $a$  ein **innerer Punkt** von  $X$ , wenn eine genügend kleine  $\epsilon$ -Umgebung von  $X$  Teilmenge von  $X$  ist.

**2.6 Thm.** Sei  $f : X \rightarrow \mathbb{R}$  differenzierbare Funktion auf einer Teilmenge  $X$  von  $\mathbb{R}$  und sei  $a$  ein innerer Punkt von  $X$ , in welchem  $f$  ein lokales Extremum erreicht. Dann gilt  $f'(a) = 0$ .

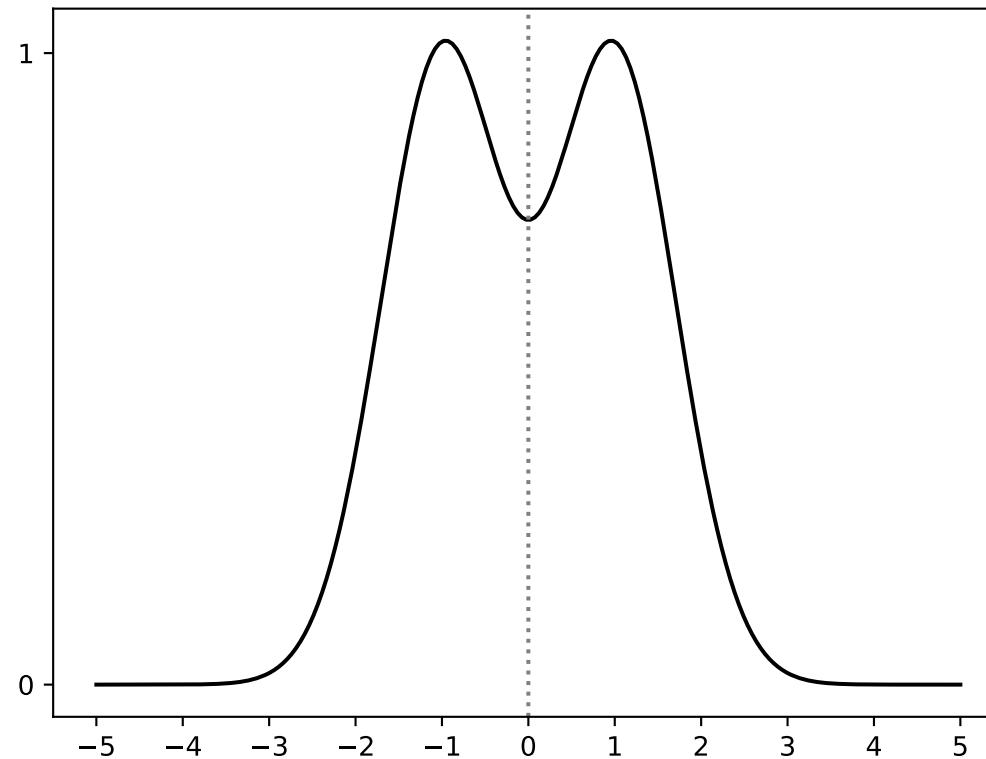
**2.7 Bsp.** Das vorige Theorem hilft, globale Optima einer differenzierbaren Funktion  $f : X \rightarrow \mathbb{R}$  auf einer kompakten Mengen  $X \subseteq \mathbb{R}$  zu finden.

Wenn wir etwa auf die globalen Extreme von

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

für  $x \in X := [-1, 1] \cup \{2\}$  bestimmen wollen, so ist jeder Punkte, in dem Extremum erreicht ist, entweder ein  $x \in (-1, 1)$  mit  $f'(x) = 0$  oder einer der drei Randpunkte  $-1, 1$  oder  $2$ . Da man in diesem Fall nur endlich viele  $x \in (-1, 1)$  mit  $f'(x) = 0$  hat, kommen lediglich endlich viele Punkte von  $X$  in Frage.

**2.8 Bsp.** Auf unbeschränkten Mengen, wie z.B., muss eine differenzierbare Funktion  $f$  weder ein globales Maximum noch ein globales Minimum erreichen, auch wenn die Funktion beschränkt ist und/oder lokale Extreme besitzt.



$f(x) = e^{-(x-1)^2} + e^{-(x+1)^2}$  erreicht ein lokales Minimum in 0, aber kein globales Minimum;  
denn  $f(0) = \frac{2}{e} > 0$ , während  $f(x)$  für große  $|x|$  beliebig klein positiv wird.

**2.9 Bsp.**  $f(x) = \arctan x$  erreicht weder lokales Minimum noch ein lokales Maximum.

**2.10 Def.**  $C \in [-\infty, \infty]$  heißt **obere Schranke** an  $f : X \rightarrow \mathbb{R}$ , wenn  $f(x) \leq C$  für alle  $x \in X$  gilt.  $c \in [-\infty, \infty]$  heißt **untere Schranke** an  $f : X \rightarrow \mathbb{R}$ , wenn  $f(x) \geq c$  für alle  $x \in X$  gilt.  $f : X \rightarrow \mathbb{R}$  heißt nach oben bzw. unten **beschränkt**, wenn  $f$  eine endliche obere bzw. untere Schranke hat.

Das **Supremum** von  $f$  auf  $X$  ist

$$\sup_{x \in X} f(x) := \text{die kleinste obere Schranke von } f : X \rightarrow \mathbb{R}$$

Das **Infimum** von  $f$  auf  $X$  ist

$$\inf_{x \in X} f(x) := \text{die größte untere Schranke von } f : X \rightarrow \mathbb{R}.$$

**2.11 Bsp.** Bei einer  $f$ , die ein endliches Supremum bzw. Infimum hat, muss das Maximum bzw. Minimum im Allgemeinen nicht erreicht werden. Z.B.:

$$\sup_{x \in \mathbb{R}} \arctan x = +\frac{\pi}{2}, \quad | \text{ Max. wird nicht erreicht, obwohl das Sup. endlich ist}$$

$$\inf_{x \in \mathbb{R}} \arctan x = -\frac{\pi}{2} \quad | \text{ Min. wird nicht erreicht, obwohl das Inf. endlich ist}$$

Man beachte:  $\arctan(x)$  ist sehr regelmäßig (beliebig oft differenzierbar).

Problem hier: unbeschränktes  $\mathbb{R}$ .

**2.12 Bsp.**

$$\sup_{-1 < x < 1} (5x + x^3) = 6 \quad | \text{ Max. wird nicht erreicht,}$$

$$\inf_{-1 < x < 1} (5x + x^3) = 4 \quad | \text{ Min. wird nicht erreicht.}$$

Problem hier: das Intervall  $(-1, 1)$  ist zwar beschränkt aber nicht abgeschlossen.

**2.13 Bsp.** Durch  $\lfloor x \rfloor$  wird das Abrunden bezeichnet.

$$\sup_{0 \leq x \leq 1} (x - \lfloor x \rfloor) = 1 \quad | \text{ Max. wird nicht erreicht.}$$

Problem hier: zwar ist  $[0, 1]$  kompakt (d.h., beschränkt und abgeschlossen), ist die Funktion nicht stetig.

Bei stetigen Funktionen auf einer kompakten Menge garantiert der Satz von Weierstraß, dass das Maximum sowie das Minimum erreicht werden.

**2.14 Thm.** Sei  $f : X \rightarrow \mathbb{R}$  zweimal differenzierbare Funktion auf einer Teilmenge  $X$  von  $\mathbb{R}$  und sei  $a$  ein innerer Punkt von  $X$ , in welchem  $f'(a) = 0$  gilt. Dann gilt:

- (a) Ist  $f''(a) > 0$ , dann erreicht  $f$  in  $a$  ein lokales Minimum.
- (b) Ist  $f''(a) < 0$ , dann erreicht  $f$  in  $a$  ein lokales Maximum.

**2.15.** Um die Unterscheidung zwischen  $f'(a) = 0, f''(a) < 0$  und  $f'(a) = 0, f''(a) > 0$  zu merken, denken Sie an die beiden Funktionen  $f(x) = \pm x^2$  und  $a = 0$ .

**2.16.** Für  $f(x) = e^x - x$ , ist  $x = 0$  die einzige Stelle mit  $f'(x) = 0$ . An dieser Stelle gilt  $f''(x) > 0$ . Das  $f$  erreicht also in 0 ein lokales Minimum. Man kann zeigen, dass  $f$  in 0 sogar das globale Minimum erreicht, denn

$$\lim_{x \rightarrow +\infty} f(x) = +\infty,$$

$$\lim_{x \rightarrow -\infty} f(x) = +\infty,$$

woraus man sieht, dass bei einem genügend großen  $|x|$  der Wert  $f(x)$  größer als  $f(0) = 1$  ist.

**2.17 Bsp.** Man kann nicht immer (in der Praxis so gut wie nie) Lösungen der Gleichung  $f'(x) = 0$  in einer abgeschlossenen Form finden. In der Praxis nutzt man numerische Verfahren so eine Art Gleichungen annähernd zu lösen.

### 3 Konvexität

**3.1 Def.** Eine Menge  $X \subseteq \mathbb{R}^n$  nennt man konvex, wenn für alle  $a, b \in X$  und  $0 \leq \lambda \leq 1$  die Bedingung  $(1 - \lambda)a + \lambda b \in X$  erfüllt ist. Mit anderen Worten: mit je zwei Punkten in  $X$  ist die Verbindungsstrecke dieser Punkte Teilemenge von  $X$ .

**3.2.** Konvexe Teilmengen von  $\mathbb{R}$ , die nicht aus einem einzigen Punkt bestehen sind **Intervalle**. Intervalle können beschränkt oder unbeschränkt, offen, abgeschlossen oder halboffen sein.

**3.3 Def.** Eine Funktion  $f : X \rightarrow \mathbb{R}$  auf einer konvexen Teilmenge  $X$  von  $\mathbb{R}^n$  nennt man **konvex**, wenn für alle  $a, b \in X$  und alle  $0 \leq \lambda \leq 1$  die Ungleichung

$$f((1 - \lambda)a + \lambda b) \leq (1 - \lambda)f(a) + \lambda f(b)$$

erfüllt ist.

Eine Funktion  $f$  nennt man **konkav**, wenn die Funktion  $-f$  konvex ist.

**3.4 Thm.** Für eine differenzierbare Funktion  $f : I \rightarrow \mathbb{R}$  auf einem Intervall  $I$  sind die folgenden Bedingungen äquivalent:

- (i)  $f$  ist konvex,
- (ii)  $f'$  ist monoton steigend.

**3.5 Thm.** Für eine zweimal differenzierbare Funktion  $f : I \rightarrow \mathbb{R}$  auf einem Intervall  $I$  sind die folgenden Bedingungen äquivalent:

- (i)  $f$  ist konvex,
- (ii)  $f'' \geq 0$  auf  $I$ .

## **4 Der Mittelwertsatz der Differentialrechnung**

**4.1 Thm.** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine stetige Funktion auf einem Intervall  $[a, b]$  mit  $a, b \in \mathbb{R}$  und  $a < b$ , die auf  $(a, b)$  differenzierbar ist. Dann existiert ein  $\xi \in (a, b)$  mit

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

**4.2.** Der Satz ist eine weitere Berechtigung für die Approximation der Ableitung durch einen Quotienten.

**4.3.** Geometrische Intuition sieht man am Graphen der Funktion: Die Sekante zu Punkten  $(a, f(a))$  und  $(b, f(b))$  wird durch eine geeignete Verschiebung zu einer Tangente an einem der Punkte  $(\xi, f(\xi))$  mit  $\xi \in (a, b)$ .

**4.4. Weitere Intuition:** Die Strecke von der BTU Cottbus bis zu Berlin Zoo ist 140 km lang. Wenn wir diese Strecke in 1,5 h geschafft haben, dann war unsere Geschwindigkeit in mindestens einem Zeitpunkt unserer Fahrt

$$\frac{140 \text{ km}}{1,5 \text{ h}} = 93\frac{1}{3} \frac{\text{km}}{\text{h}}.$$

Das ist unsere Durchschnittsgeschwindigkeit. Die Durchschnittsgeschwindigkeit wird mindestens ein mal erreicht.

Die Funktion in diesem Beispiel ist: Zurückgelegter Weg  $f(x)$  (in km) zum Zeitpunkt  $x$  (in h) mit  $0 \leq x \leq 1,5$ , und wir setzen  $a = 0$  und  $b = 1,5$ .

**4.5 Def.** Man sagt,  $\xi \in \mathbb{R}$  liegt zwischen  $a \in \mathbb{R}$  und  $b \in \mathbb{R}$ , wenn  $a \leq \xi \leq b$  oder  $b \leq \xi \leq a$  erfüllt ist.

**4.6.** Wir betrachten eine differentierbare Funktion  $f : I \rightarrow \mathbb{R}$  auf einem Intervall und fixieren ein  $a$  aus dem inneren von  $I$ . Machen wir das  $b$  variabel, und bezeichnen es als  $x$ . Dann ist das  $\xi$  zwischen  $a$  und  $x$  abhangig von der Wahl von  $x$ , und wir bezeichnen es als  $\xi_x$ . Das vorige Theorem ergibt:

$$\begin{aligned} f'(\xi_x) &= \frac{f(x) - f(a)}{x - a} & | \Rightarrow \\ f(x) &= f(a) + f'(\xi_x)(x - a) & | \text{ Nach } f(x) \text{ aufgelost} \end{aligned}$$

Dies ist die Taylorentwicklung der Ordnung 0. Dies ist ein Spezialfall einer viel allgemeineren Formel, die wir in Kurze diskutieren werden.

# **Kapitel IV**

## **Taylor- und Potenzreihen**

### **1 Taylor-Entwicklung und Taylor-Reihen**

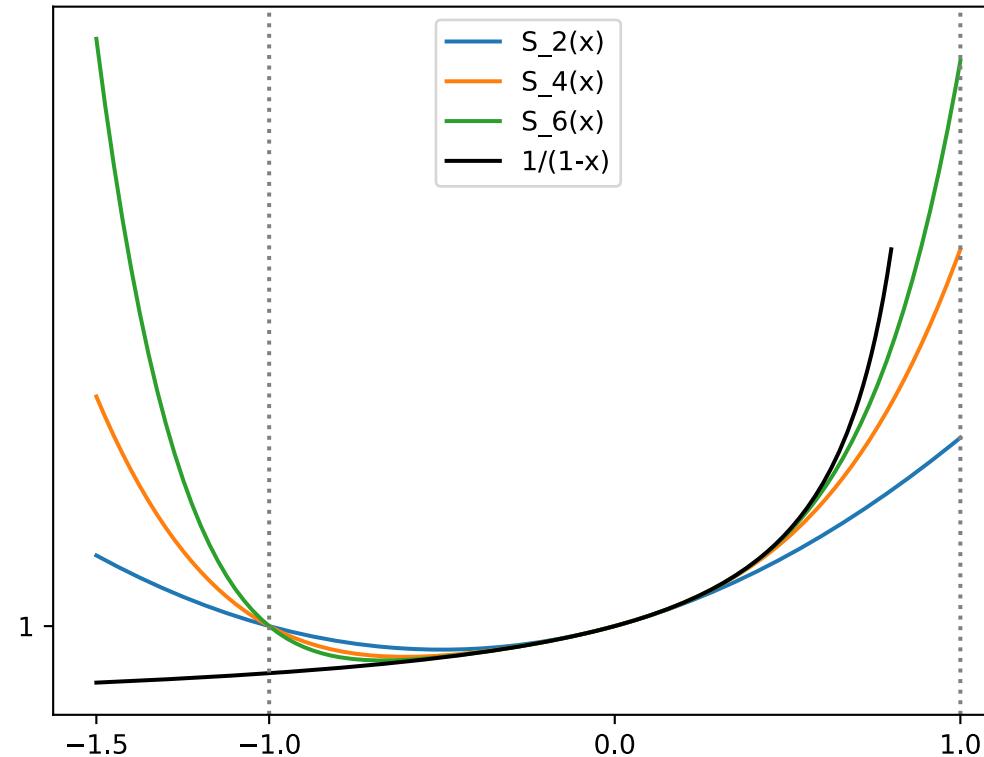
**1.1 Def.** Eine **Funktionsreihe** ist eine Ausdruck der Form  $\sum_{k=0}^{\infty} g_k(x)$ , in dem die Glieder  $g_0(x), g_1(x), g_2(x), \dots$  der Reihe Funktionen in  $x$  sind.

Die Funktionsreihen der Form  $\sum_{k=0}^{\infty} a_k(x - c)^k$ , die durch Konstanten  $c$  und die Folge von Konstanten  $a_0, a_1, a_2, \dots$  gegeben sind, nennt man **Potenzreihen**. Den Wert  $c$  nennt man den **Entwickelpunkt** der Potenzreihe.

**1.2 Bsp.** Die geometrische Reihe mit einer variablen Basis  $x$  ist eine Potenzreihe, für welche die Summe bekannt ist:

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

Diese Gleichung gilt für  $|x| < 1$ , das heißt, für  $x$  im offenen Intervall  $(-1, 1)$ .

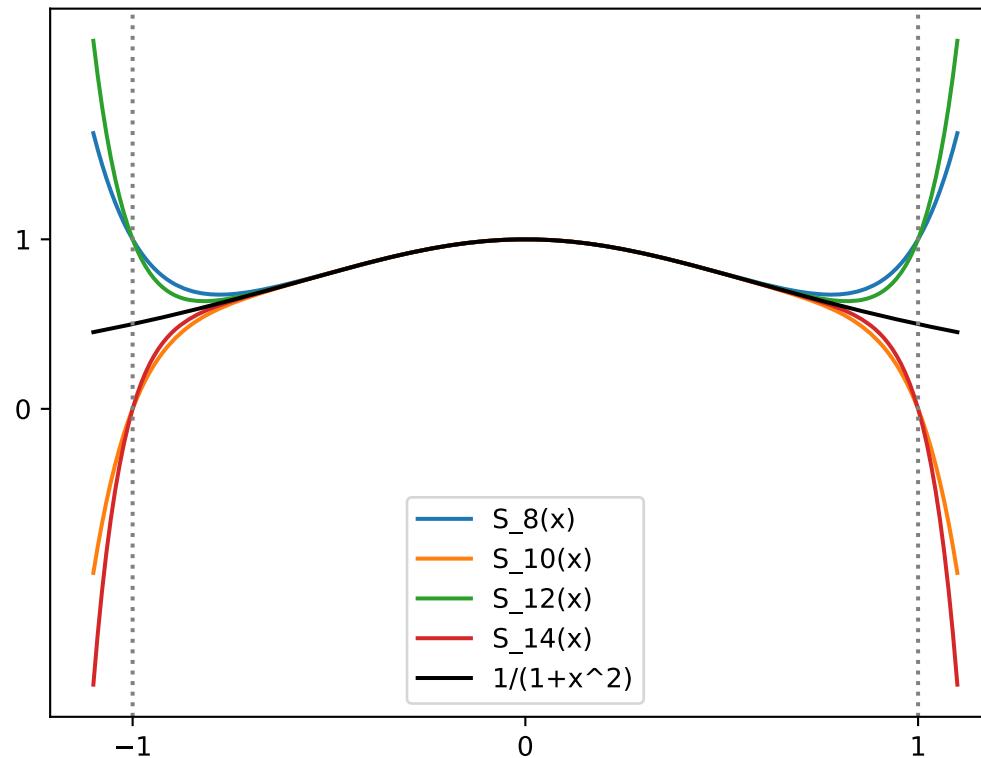


Partialsummen  $s_n(x) := \sum_{k=0}^n x^k$ .

**1.3 Bsp.** Wir entwickeln die Funktion  $\frac{1}{1+x^2}$  in die Potenzreihe zum Entwicklungspunkt 0:

$$\begin{aligned}\frac{1}{1+x^2} &= \sum_{k=0}^{\infty} (-x^2)^k && | \text{ Formel für die Summe der Geom. Reihe} \\ &= \sum_{k=0}^{\infty} (-1)^k x^{2k} && | \text{ Nur gerade Potenzen hier!} \\ &= 1 - x^2 + x^4 - x^6 + \dots && | \text{ Potenzreihe mit den Koef. } 1, 0, -1, 1, 0, -1, \dots\end{aligned}$$

Man beachte: wir wissen aus der Analyse der geometrische Reihe, dass diese Entwicklung nur für  $x^2 < 1$ , also für  $|x| < 1$ , eine gültige Gleichung ergibt. Für  $|x^2| \geq 1$ , ist die Reihe divergent, auch wenn die Formel  $\frac{1}{1+x^2}$ , für die wir die Entwicklung aufstellen, für alle  $x \in \mathbb{R}$  ausgewertet werden kann!



Partialsummen  $s_n(x)$ .

**1.4 Def.** Für eine Funktion  $f : X \rightarrow \mathbb{R}$  auf  $X \subseteq \mathbb{R}$  wird das **Taylor-Polynom** von  $f$  der Ordnung  $n$  bzgl. des Entwicklungspunkts  $a$  durch

$$T_n(x) := \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x - a)^k$$

definiert.

Die Funktion  $R_n(x) = f(x) - T_n(x)$  nennt man das **Restglied** der Taylor-Entwicklung der Ordnung  $n$  an der Stelle  $a$ .

Die Potenzreihe

$$T_\infty(x) := \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x - a)^k$$

nennt man die **Taylor-Reihe** von  $f$  bzgl. des Entwicklungspunkts  $a$ .

**1.5.** Was haben  $f(x)$  das Polynom  $T_n(x)$  gemeinsam? Die Ableitungen an der Stelle  $a$ . Wenn man die Ableitungen bis zur Ordnung  $n$  von  $T_n(x)$  ausrechnet, stellt man fest:

$$T_n^{(i)}(a) = f^{(i)}(a) \quad \text{für } i = 0, \dots, n.$$

Dadurch, dass das Polynom  $T_n(x)$  gleiche Ableitungen wie  $f(x)$  an der Stelle  $a$  hat, sieht es um die Stelle  $a$  ziemlich ähnlich zu  $f(x)$  aus. Diese informelle Bemerkung wird im Folgenden genau formuliert.

**1.6 Def** ( $O$ -Notation für  $x \rightarrow a$ ). Seien  $f, g : X \rightarrow \mathbb{R}$  Funktionen auf  $X \subseteq \mathbb{R}$  und  $a$  Häufungspunkt von  $X$ . Man schreibt,

$$f(x) = O(g(x)), \quad \text{für } x \rightarrow a,$$

wenn eine Konstante  $C \geq 0$  und ein  $\delta > 0$  existieren, für welche die Ungleichung

$$|f(x)| \leq C|g(x)|$$

für alle  $x \in X$  mit  $|x - a| < \delta$  erfüllt ist.

Mit Worten:  $f(x)$  hat die **Ordnung  $O$**  von  $g(x)$  für  $x \rightarrow a$ .

**1.7.** Sie kennen bestimmt verschiedene Formen der asymptotische Notation für Funktionen in einer Variablen  $n$  aus  $\mathbb{N}$  aus der Informatik, da diese Notation bei der Analyse der Laufzeit von Algorithmen sehr oft verwendet ist. All die Bezeichnungen, die man bei der Analyse der Laufzeit von Algorithmen benutzt, kann man auch für Funktionen in einer reellwertigen Variablen  $x$  für  $x \rightarrow a$  einführen.

Neben der O-Notation hat man zum Beispiel auch  $\Omega$ -,  $\Theta$ - und  $o$ -Notation für  $x \rightarrow a$ . Die Bedingung  $f(x) = o(g(x))$  für  $x \rightarrow a$  heißt zum Beispiel, dass

$$\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = 0$$

gilt.

**1.8 Def.** Eine Funktion  $f : X \rightarrow \mathbb{R}$  auf  $X \subseteq \mathbb{R}$  nennt man  **$n$ -mal stetig differenzierbar**, wenn die Ableitungen  $f, f', \dots, f^{(n)}$  stetige Funktionen auf  $X$  sind.

**1.9 Thm** (Über Approximation durch Taylor-Entwicklung). *Sei  $f : I \rightarrow \mathbb{R}$  ( $n + 1$ )-mal stetig differenzierbare Funktion auf einem Intervall  $I$  und  $a$  ein innerer Punkt von  $I$ . Dann gilt für das Restglied  $R_n(x)$  der Taylor-Entwicklung der Ordnung  $n$  bzgl.  $a$  die asymptotische Abschätzung*

$$R_n(x) = O((x - a)^{n+1}), \quad x \rightarrow a.$$

*Mit anderen Worten:*

$$f(x) = T_n(x) + O((x - a)^{n+1}), \quad x \rightarrow a.$$

**1.10.** Interessante Spezialfälle sind die lineare Approximation

$$f(x) = f(a) + f'(a)(x - a) + O((x - a)^2), \quad x \rightarrow a.$$

und die quadratische Approximation

$$f(x) = f(a) + f'(a)(x - a) + \frac{1}{2}f''(a)(x - a)^2 + O((x - a)^3), \quad x \rightarrow a.$$

Die notwendigen und hinreichenden der Bedingungen der Optimalität erhält man aus diesen Approximationen.

**1.11.** Das vorige Theorem ist Folgerung aus einem Theorem, welches etwas mehr Details über eine mögliche Darstellung des Restterms verrät.

**1.12 Thm** (Über die Darstellung des Restglieds). *Sei  $f : I \rightarrow \mathbb{R}$  ( $n + 1$ -mal stetig differenzierbare Funktion auf einem Intervall  $I$  und sei  $a$  innerer Punkt von  $I$ . Man betrachte das Taylor-Polynom  $T_n(x)$  von  $f$  der  $n$ -ten Ordnung bzgl. des Entwicklungspunkts  $a$ . Dann existiert für jedes  $x \in I$  ein  $\xi_x \in I$ , das zwischen  $a$  und  $x$  liegt, mit*

$$f(x) = T_n(x) + \frac{f^{(n+1)}(\xi_x)}{(n+1)!}(x-a)^{n+1}.$$

*Insbesondere gilt*

$$|f(x) - T_n(x)| \leq \frac{M_{n+1}}{(n+1)!}|x-a|^{n+1},$$

*wenn  $M_{n+1}$  eine Konstante ist, welche den Betrag der  $(n+1)$ -ten Ableitung auf dem Intervall  $I$  nach oben abschätzt.*

**1.13 Bsp.** Wir approximieren  $f(x)$  auf  $x \in [-1, 1]$  durch die Taylorentwicklung erster Ordnung im Punkt 0 und schätzen die Qualität der Approximation ab:

$$\begin{aligned} f(x) = f'(x) = e^x &\implies f(0) = f'(0) = 1 \\ &\implies T_1(x) = f(0) + f'(0)x = 1 + x. \end{aligned}$$

Es gilt:

$$|e^x - (1 + x)| \leq \frac{M_2}{2!} x^2 \quad \text{für } -1 \leq x \leq 1$$

mit

$$M_2 = \max_{-1 \leq x \leq 1} |f''(x)| = \max_{-1 \leq x \leq 1} |e^x| = e.$$

Fazit:

$$|e^x - (1 + x)| \leq \frac{e}{2} \cdot x^2 \quad \text{für } -1 \leq x \leq 1.$$

**1.14 Bsp.** Wir approximieren  $f(x) = \sin x$  auf  $x \in [-\pi/2, \pi/2]$  durch die Taylorentwicklung zweiter Ordnung im Punkt 0 und schätzen die Qualität der Approximation ab.

$$\begin{array}{l|l} f(x) = \sin x & f(0) = 0 \\ f'(x) = \cos x & f'(0) = 1 \\ f''(x) = -\sin x & f''(0) = 0 \end{array} \implies T_2(x) = x.$$

Es gilt

$$|x - \sin x| \leq \frac{M_3}{3!} |x|^3$$

mit

$$M_3 = \max_{-\pi/2 \leq x \leq \pi/2} |f'''(x)| = \max_{-\pi/2 \leq x \leq \pi/2} |- \cos x| = 1.$$

Fazit:

$$|x - \sin x| \leq \frac{1}{6} |x|^3 \quad \text{für } -\pi/2 \leq x \leq \pi/2.$$

**1.15 Aufgabe.** Rechnen Sie  $T_\infty(x)$  bzgl. der Entwicklungsstelle  $a = 0$  für die Funktionen  $e^x, \cos x, \sin x$  und  $\frac{1}{1-x}$  aus:

**1.16 Aufgabe.** Rechnen Sie  $T_\infty(x)$  bzgl. der Entwicklungsstelle  $a = 0$  für  $\ln x$  aus.

## 2 Potenzreihen

**2.1 Def.**  $C \in [-\infty, \infty]$  heißt obere Schranke an eine Menge  $X \subseteq \mathbb{R}$  wenn  $x \leq C$  für alle  $x \in X$  erfüllt ist.

$c \in [-\infty, \infty]$  heißt untere Schranke an eine Menge  $X \subseteq \mathbb{R}$ , wenn  $x \geq c$  für alle  $x \in X$  erfüllt ist.

Wir definieren das Supremum/Infimum einer Menge  $X$  als:

$\sup X :=$  kleinste obere Schranke an  $X$ ,

$\inf X :=$  größte untere Schranke an  $X$

**2.2 Bsp.** Supremum und Infimum für  $\mathbb{R}$ , leere Menge und das offene Intervall  $(0, 1)$ .

$$\sup \mathbb{R} = +\infty$$

$$\sup \emptyset = -\infty$$

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

$$\inf \mathbb{R} = -\infty$$

$$\inf \emptyset = +\infty$$

$$\inf(0, 1) = 0.$$

**2.3 Def.** Für die Potenzreihe  $\sum_{k=0}^{\infty} a_k(x - c)^k$  nennt man die Menge

$$X := \left\{ x \in \mathbb{R} : \sum_{k=0}^{\infty} a_k(x - c)^k \text{ konvergent} \right\}$$

den **Konvergenzbereich** der Reihe und den Wert

$$\rho := \sup \{|x - c| : x \in X\}$$

den **Konvergenzradius**.

**2.4.** Der Konvergenzradius ist ein Wert  $\rho \in [0, \infty]$ . Wenn man  $\rho = 0$  hat, dann konvergiert die Reihe nur für  $x = c$ .

**2.5 Bsp.** Für  $\sum_{k=0}^{\infty} x^k$  ist der Konvergenzbereich  $X = (-1, 1)$  und somit der Konvergenzradius  $\rho = \sup \{|x| : x \in X\} = 1$ . Im Folgenden Theorem wird behauptet, dass jeder Konvergenzbereich ein Intervall ist, das  $c$  genau in der Mitte hat...

**2.6 Thm.** Für die Potenzreihe  $\sum_{k=0}^{\infty} a_k(x - c)^k$  mit dem Konvergenzradius  $\rho$  gilt:

- (a) Die Reihe ist absolut konvergent für alle  $x \in \mathbb{R}$  mit  $|x - c| < \rho$
- (b) Die Reihe ist divergent für alle  $x \in \mathbb{R}$  mit  $|x - c| > \rho$ .
- (c) Im Fall  $0 < \rho < \infty$  kann man über die Konvergenz in den beiden Punkten  $x \in \mathbb{R}$  mit  $|x - c| = \rho$  a priori nichts sagen: das heißt, der Konvergenzbereich ist  $[c - \rho, c + \rho]$  oder  $(c - \rho, c + \rho]$  oder  $[c - \rho, c + \rho)$  oder  $(c - \rho, c + \rho)$  und im Allgemeinen ist jeder dieser vier Fälle möglich.

**2.7 Bsp.** Die Potenzreihe  $\sum_{k=0}^{\infty} \frac{(-1)^k}{k} x^k$  hat den Konvergenzbereich  $(-1, 1]$ .

Begründung:

- (a) Für  $|x| > 1$ , ist das Glied der Reihe keine Nullfolge, sodass die Reihe für diese Wahl von  $x$  nicht konvergent ist.
- (b) Für  $-1 < x < 1$  folgt die Konvergenz aus dem Quotientenkriterium.
- (c) Für  $x = 1$  folgt die Konvergenz aus dem Leibniz-Kriterium.
- (d) Für  $x = -1$  ist das die harmonische Reihe und somit eine divergente Reihe.

**2.8 Bsp.** Wie kann man den Konvergenzbereich bzw. -radius von  $\sum_{k=0}^{\infty} 2^{-k}x^k$  ohne (komplizierte) Hilfsmittel berechnen?

Wir wissen: die geometrische Reihe  $\sum_{k=0}^{\infty} x^k$  konvergiert genau dann, wenn  $|x| < 1$  gilt.

Es folgt:

$$\sum_{k=0}^{\infty} 2^{-k}x^k = \sum_{k=0}^{\infty} \left(\frac{x}{2}\right)^k$$

konvergiert genau dann wenn  $\left|\frac{x}{2}\right| < 1$  gilt.

Fazit: Konvergenzradius von  $\sum_{k=0}^{\infty} 2^{-k}x^k$  ist 2.

**2.9.** Wie kann man das vorige Beispiel auf allgemeinere Potenzreihen  $\sum_{k=0}^{\infty} a_k(x - c)^k$  verallgemeinern? Betrachten wir Einfachheit halber den Fall  $a_k > 0$ .

Man hat

$$a_k(x - c)^k = \left(a_k^{1/k}(x - c)\right)^k$$

Das heißt: Wenn  $a_k^{1/k}$  einen Grenzwert  $\gamma$  besitzt, so verhält sich  $a_k(x - c)^k$  ungefähr wie der  $k$ -te Term der geometrischen Reihe mit der Basis  $\gamma(x - c)$ . Für die Geometrische  $\sum_{k=0}^{\infty} (\gamma(x - c))^k$  hat man aber die Konvergenz genau dann, wenn  $|\gamma(x - c)| < 1$  erfüllt ist, sodass ihr Konvergenzradius gleich  $\frac{1}{\gamma}$  ist.

Dies ist die Intuition hinter dem folgenden Theorem...

**2.10 Thm.** Wenn die Potenzreihe  $\sum_{k=0}^{\infty} a_k(x - c)^k$  einen Grenzwert  $\lim_{k \rightarrow \infty} |a_k|^{1/k} \in [0, +\infty]$  besitzt (0 und  $+\infty$  sind beide zugelassen), dann ist der Konvergenzradius der Reihe gleich

$$\rho = \frac{1}{\lim_{k \rightarrow \infty} |a_k|^{1/k}}, \quad (\text{IV.1})$$

wobei man  $1/0$  bzw.  $1/\infty$  auf der rechten Seite von (IV.1) als  $\infty$  bzw.  $0$  interpretiert.

**2.11 Bsp.** Für  $\sum_{k=0}^{\infty} (2 + 1/k)^k x^k$  gilt

$$\lim_{k \rightarrow \infty} ((2 + 1/k)^k)^{1/k} = \lim_{k \rightarrow \infty} (2 + 1/k) = 2.$$

Der Konvergenzradius der Reihe ist 1/2.

**2.12 Bsp.** Der Konvergenzradius von  $\sum_{k=0}^{\infty} k^k x^k$  ist 0. Das heißt, die Reihe konvergiert nur für  $x = 0$ .

**2.13 Bsp.** Der Konvergenzradius von  $\sum_{k=0}^{\infty} k^{-k} x^k$  ist  $\infty$ . Das heißt, die Reihe konvergiert (absolut) für alle  $x$ .

**2.14.** Das vorige Theorem ist für die Beispiele wie z.B.

$$\sum_{t=0}^{\infty} x^{t^2} = 1 + x + x^4 + x^9 + x^{16} + \dots$$

nicht anwendbar, da hier die Folge  $a_k^{1/k}$  keinen Grenzwert hat. Es gibt aber eine allgemeinere Version des Theorems.

**2.15 Def.** Für eine Folge  $(x_k)_{k \in \mathbb{N}_0}$  ist der **Limes Superior**

$\limsup_{k \rightarrow \infty} x_k :=$  der maximale Grenzwert einer konvergenten Teilfolge von  $(x_k)_{k \in \mathbb{N}_0}$ .

wenn die Folge nach oben beschränkt ist, und  $\limsup_{k \rightarrow \infty} x_k = \infty$ , wenn die Folge nach oben unbeschränkt ist.

**2.16.** Sind alle  $x_k$  nicht negativ, und die Teilfolge aus den positiven Gliedern konvergent, dann gilt

$$\limsup_{k \rightarrow \infty} x_k = \text{der Grenzwert der Teilfolge der positiven Glieder von } (x_k)_{k \in \mathbb{N}_0}.$$

**2.17 Thm** (Update). *Der Konvergenzradius der Potenzreihe  $\sum_{k=0}^{\infty} a_k(x - c)^k$  ist*

$$\rho = \frac{1}{\limsup_{k \rightarrow \infty} |a_k|^{1/k}}, \quad (\text{IV.2})$$

*wobei man  $1/0$  bzw.  $1/\infty$  auf der rechten Seite von (IV.2) als  $\infty$  bzw.  $0$  interpretiert.*

**2.18 Bsp.** Der Konvergenzradius von  $\sum_{t=0}^{\infty} x^{t^2}$  ist 1, da der Limes Superior der Folge

$$a_k = \begin{cases} 1 & k = t^2 \text{ für ein } t \in \mathbb{N}_0 \\ 0 & \text{sonst} \end{cases}$$

gleich 1 ist.

**2.19.** Es gibt eine weitere nützliche Formel für den Konvergenzradius im Fall, dass man  $a_k = 0$  nur für endlich viele  $k \in \mathbb{N}_0$  hat. Die Formel ergibt als direkte Folgerung aus dem Quotientenkriterium (in der Grenzwert-Form) für Reihen...

**2.20 Thm.** Sei  $\sum_{k=0}^{\infty} a_k(x - c)^k$  Potenzreihe, in der  $a_k = 0$  nur für endlich viele  $k \in \mathbb{N}_0$  erfüllt ist und für welche der Grenzwert

$$\rho := \lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right|$$

als Wert aus  $[0, \infty]$  existiert. Dann ist  $\rho$  der Konvergenzradius dieser Potenzreihe.

**2.21 Bsp.** Der Konvergenzradius von  $\sum_{k=0}^{\infty} \frac{1}{k!} x^k$  ist  $\infty$ .

**2.22 Bsp.** Der Konvergenzradius von  $\sum_{k=1}^{\infty} \frac{(-1)^k}{k} x^k$  ist 1.

**2.23 Thm.** Sei  $\sum_{k=0}^{\infty} a_k(x - c)^k$  Potenzreihe und sei Konvergenzradius  $\rho$  dieser Reihe ungleich 0, d.h.,  $\rho \in (0, \infty]$ . Dann ist  $g(x) = \sum_{k=0}^{\infty} a_k(x - c)^k$  im Bereich  $(c - \rho, c + \rho)$  beliebig oft differenzierbar. Darüber hinaus gilt:

- Die  $i$ -te Ableitung  $g^{(i)}(x)$  mit  $i \in \mathbb{N}$  lässt als eine Potenzreihe durch gliederweises Differenzieren der Reihe  $\sum_{k=0}^{\infty} a_k(x - c)^k$  darstellen. Das heißt, die Reihe  $\sum_{k=0}^{\infty} (a_k(x - c)^k)^{(i)}$  ist absolut konvergent in  $(-\rho + c, c + \rho)$  mit

$$g^{(i)}(x) = \sum_{k=0}^{\infty} (a_k(x - c)^k)^{(i)}.$$

- Man hat  $a_k = \frac{g^{(k)}(c)}{k!}$  für alle  $k \in \mathbb{N}$ .

**2.24.** Das vorige Theorem zeigt: jede Potenzreihe (mit einem positiven Konvergenzradius) ist eine Taylorreihe einer unendlich oft differenzierbaren Funktion.

**2.25 Bsp.** Die Berechnung in diesem Beispiel machen wir für  $x \in (-1, 1)$ . Wir wissen, dass

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

gilt. Nach dem vorigen Theorem können wir diese Gleichung differenzieren, indem wir die Glieder der Reihe auf der rechten Seite differenzieren. Wir erhalten also

$$\frac{1}{(1-x)^2} = \left( \frac{1}{1-x} \right)' = \sum_{k=0}^{\infty} (x^k)' = \sum_{k=1}^{\infty} kx^{k-1}.$$

Die Formel

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} kx^{k-1}$$

kann man auch mit der Verwendung des Cauchy-Produkts erhalten. Wenn wir noch einmal

differenzieren, erhalten wir

$$\frac{2}{(1-x)^3} = \sum_{k=2}^{\infty} k(k-1)x^{k-2}.$$

Diese Gleichung kann man auch als

$$\frac{1}{(1-x)^3} = \sum_{k=2}^{\infty} \binom{k}{2} x^{k-2}$$

mit Hilfe von Binomialkoeffizienten umschreiben.

**2.26 Aufgabe.** Finden Sie eine Darstellung von  $(1 - x)^{-d}$  mit  $d \in \mathbb{N}$  und  $-1 < x < 1$  als eine Potenzreihe in  $x$ , indem Sie die Darstellung von  $\frac{1}{1-x}$  als Summe der geometrischen Reihegliederweise  $(d - 1)$ -mal differenzieren.

**2.27 Def.** Eine Funktion  $g$  heißt **analytisch** in  $X$ , wenn man für jedes  $c \in X$  ein  $\epsilon > 0$  und eine Potenzreihe  $\sum_{k=0}^{\infty} a_k(x - c)^k$  mit dem Konvergenzradius  $> \epsilon$  findet, für welche die Gleichung

$$g(x) = \sum_{k=0}^{\infty} a_k(x - c)^k$$

für alle  $x \in (c - \epsilon, c + \epsilon)$  erfüllt ist.

**2.28 Thm.** Ist  $g$  analytisch in  $X$ , dann ist  $g$  beliebig oft differenzierbar in  $X$ .

**2.29.** Die Umkehrung gilt im Allgemeinen nicht. Die Funktion

$$f(x) := \begin{cases} e^{-1/x} & \text{für } x > 0, \\ 0 & \text{für } x \leq 0 \end{cases}$$

ist beliebig oft differenzierbar auf  $\mathbb{R}$ , ist aber nicht analytisch (Problemstelle ist die 0).

Für diese Funktion hat man  $f^{(k)}(0)$  für alle  $k \in \mathbb{N}_0$ . Das heißt: die Taylorreihe  $T_\infty(x)$  dieser Funktion zum Entwicklungspunkt 0 definiert eine Funktion, die identisch gleich 0 ist, obwohl die Funktion  $f$  in keiner Umgebung von 0 identisch null ist.

**2.30 Thm.**  $e^x, \sin x, \cos x, \ln x$ , Polynome  $p(x)$  und die Quotienten  $p(x)/q(x)$  von Polynomen sind analytische Funktionen in ihren natürlichen Definitionsbereichen.

**2.31 Thm.** *Man hat*

$$e^x = \sum_{k=0}^{\infty} \frac{1}{k!} x^k,$$

$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k},$$

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

*für alle*  $x \in \mathbb{R}$  *und*

$$\ln(1+x) = \sum_{k=1}^{\infty} \frac{(-1)^k}{k} x^k$$

*für alle*  $x \in (-1, 1)$ .

**2.32.** Die vorigen Formeln werden aus der Abschätzung des Restgliedes hergeleitet. Zum Beispiel: für  $f(x) = \sin x$  hat man

$$|\sin x - T_n(x)| \leq \frac{M_{n+1}}{(n+1)!} |x|^{n+1}$$

mit

$$M_{n+1} \leq \max_{x \in \mathbb{R}} |f^{(n+1)}(x)| \leq 1,$$

denn  $f^{(n+1)}(x)$  ist  $\pm \sin x$  oder  $\pm \cos x$ . Es folgt:

$$|\sin x - T_n(x)| \leq \frac{|x|^{n+1}}{(n+1)!}.$$

Wegen  $\lim_{n \rightarrow \infty} \frac{|x|^{n+1}}{(n+1)!} = 0$ , erhält man

$$|\sin x - T_\infty(x)| = 0.$$

**2.33 Aufgabe.** Verifizieren Sie analog die angeführten Darstellungen als Reihe für  $e^x$ ,  $\cos x$  und  $\ln(1 + x)$ .

**2.34.** Potenzreihen für Quotienten von Polynomen kann man mit Hilfe von Ableitungen (als Taylorreihen) aber auch elementar ausrechnen...

**2.35 Bsp.** Wir entwickeln  $\frac{1}{x^2 - 5x + 6}$  in eine Potenzreihe bzgl. des Entwicklungspunkts 0 auf eine elementare Weise, mit der Verwendung der Formeln für die Summe der geometrischen Reihe:

$$\begin{aligned}
 \frac{1}{x^2 - 5x + 6} &= \frac{1}{(2-x)(3-x)} \\
 &= \frac{1}{2-x} - \frac{1}{3-x} \\
 &= \frac{1}{2} \cdot \frac{1}{1-x/2} - \frac{1}{3} \cdot \frac{1}{1-x/3} \\
 &= \frac{1}{2} \sum_{k=0}^{\infty} (x/2)^k - \frac{1}{3} \sum_{k=0}^{\infty} (x/3)^k \\
 &= \sum_{k=0}^{\infty} (2^{-k-1} - 3^{-k-1}) x^k.
 \end{aligned}$$

Die Gleichung

$$\frac{1}{x^2 - 5x + 6} = \sum_{k=0}^{\infty} (2^{-k-1} - 3^{-k-1})x^k$$

gilt für  $x \in (-2, 2)$ .

### **3 Exkurs in komplexwertige Funktionen**

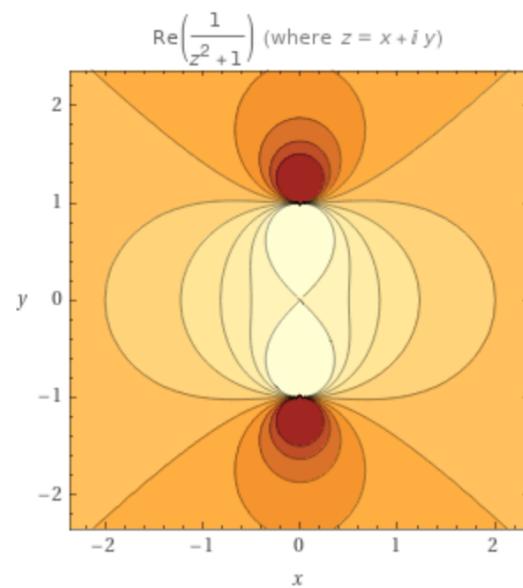
**3.1.** Potenzreihen und analytische Funktionen können analog bzgl. komplexer Zahlen eingeführt werden. Die Menge der komplexen Zahlen ist ein “natürlicher Lebensraum” für Potenzreihen und analytische Funktionen.

**3.2 Bsp.**  $\frac{1}{1+x^2}$  hat Konvergenzradius 1, auch wenn diese Funktion auf dem gesamten  $\mathbb{R}$  definiert und beliebig oft differenzierbar ist. Was ist der Grund? Wir erweitern diese Funktion auf die Menge der komplexen Zahlen.

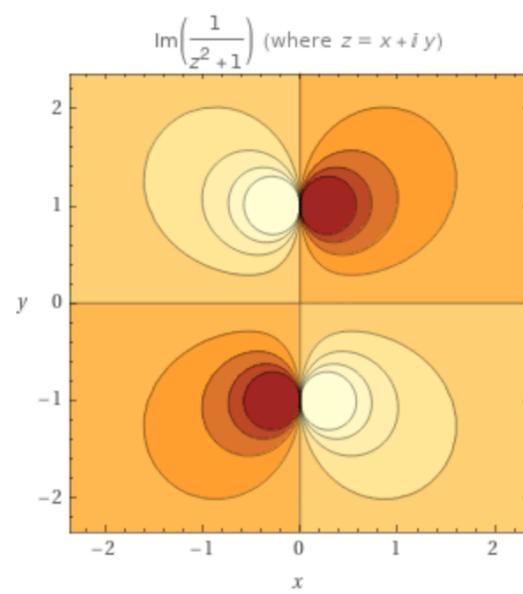
Die Funktion  $f(z) := \frac{1}{1+z^2}$  hat die Entwicklung

$$f(z) = \sum_{k=0}^{\infty} (-1)^k z^{2k}$$

für alle  $z \in \mathbb{C}$  mit  $|z| < 1$ . Die komplexen Polstellen  $z = \pm i$  mit  $|z| = 1$  sind Nachweis dafür, dass der Konvergenzradius nicht größer als 1 sein kann.



Imaginary part



**3.3.**  $\cos z, \sin z, e^z$  können durch entsprechende Reihen für alle  $z \in \mathbb{C}$  definiert werden. Man erhält:

$$\begin{aligned} e^{iz} &= \sum_{k=0}^{\infty} \frac{(iz)^k}{k!} \\ &= \sum_{k=0}^{\infty} \frac{(iz)^{2k}}{(2k)!} + \sum_{k=0}^{\infty} \frac{(iz)^{2k+1}}{(2k+1)!} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} \frac{(-1)^k z^{2k+1}}{(2k+1)!} \\ &= \cos z + i \sin z. \end{aligned}$$

Wir erhalten also die sogenannte Eulersche Formel:

$$e^{iz} = \cos z + i \sin z.$$

**3.4.** Durch Eulersche Formel kann Trigonometrie durch Algebra ersetzt werden. Wenn wir die Eulersche Formel zweimal benutzen

$$e^{\mathbf{i}z} = \cos z + \mathbf{i} \sin z,$$

$$e^{-\mathbf{i}z} = \cos z - \mathbf{i} \sin z$$

und dann die beiden Gleichungen nach  $\cos z$  und  $\sin z$  auflösen, bekommen wir die Darstellung von  $\cos z$  und  $\sin z$  als

$$\cos z = \frac{e^{\mathbf{i}z} + e^{-\mathbf{i}z}}{2},$$

$$\sin z = \frac{e^{\mathbf{i}z} - e^{-\mathbf{i}z}}{2\mathbf{i}}$$

**3.5 Bsp.** Was heißt oben “durch Algebra ersetzen”? Das heißt: trigonometrische Identitäten lassen sich aus den Algebraischen Identitäten herleiten. Nehmen wir die Formel

$$\cos 2\alpha = \cos^2 \alpha - \sin^2 \alpha.$$

für den Kosinus des doppelten Winkel als ein Beispiel. Um diese Gleichung mit Hilfe der Euler-Formel algebraisch zu verifizieren, schreiben wir diese Gleichung als

$$\frac{e^{i2\alpha} + e^{-i2\alpha}}{2} = \left( \frac{e^{i\alpha} + e^{-i\alpha}}{2} \right)^2 - \left( \frac{e^{i\alpha} - e^{-i\alpha}}{2i} \right)^2.$$

um. Die vorige Gleichung kann mit Hilfe von  $u := e^{i\alpha}$  als die algebraische Gleichung

$$\frac{u^2 + u^{-2}}{2} = \left( \frac{u + u^{-1}}{2} \right)^2 - \left( \frac{u - u^{-1}}{2i} \right)^2.$$

formuliert werden. Dass die Gleichung in  $u$  gilt, sieht man durch das Ausmultiplizieren der beiden Quadrate (mit der Verwendung der binomischen Formeln) auf der rechten Seite.

**3.6 Bsp** (Herleitung von Pythagoras).

$$\begin{aligned} 1 &= e^0 \\ &= e^{\mathbf{i}z} e^{-\mathbf{i}z} \\ &= (\cos z + \mathbf{i} \sin z)(\cos z - \mathbf{i} \sin z) && | \text{ nach der Euler-Formel} \\ &= \cos^2 z - (\mathbf{i} \sin z)^2 && | \text{ nach der dritten binomischen Formel} \\ &= \cos^2 z + \sin^2 z && | \text{ wegen } \mathbf{i}^2 = -1. \end{aligned}$$

**3.7 Aufgabe.** Finden Sie ein einfache Formel für die Summe

$$\sum_{k=0}^n \cos k\alpha$$

# **Kapitel V**

## **Integralrechnung I**

### **1 Riemann-Integral**

**1.1 Def.** Für  $a, b \in \mathbb{R}$  mit  $a < b$  nennt man eine endliche Teilmenge  $Z$  mit

$$\{a, b\} \subseteq Z \subseteq [a, b]$$

eine **Zerlegung** von  $[a, b]$ .

Mit anderen Worten: Zerlegung eines abgeschlossenen Intervalls ist eine endliche Teilmenge des Intervalls, welche die beiden Endpunkte des Intervalls enthält.

**1.2 Def.** Eine Zerlegung  $Z = \{x_0, \dots, x_n\}$  von  $[a, b]$  aus  $n + 1$  Elementen mit  $a = x_0 < x_1 < \dots < x_n = b$  ergibt eine Darstellung von  $[a, b]$  als Vereinigung der  $n$  Intervalle  $[x_{i-1}, x_i]$  mit  $i \in \{1, \dots, n\}$ .

Wir definieren für eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  die **Unter-** und **Obersumme** von  $f$  bzgl.  $Z$  durch die Gleichungen:

$$U_f(Z) := \sum_{i=1}^n \inf \{f(x) : x \in [x_{i-1}, x_i]\} \cdot (x_i - x_{i-1})$$
$$O_f(Z) := \sum_{i=1}^n \sup \{f(x) : x \in [x_{i-1}, x_i]\} \cdot (x_i - x_{i-1}).$$

In jeder der beiden Summe hat man einen Term pro jedes der  $n$  Intervalle.

**1.3 Def.** Direkt aus der Definition sieht man, dass  $U_f(Z) \leq O_f(Z)$  für alle Zerlegungen  $Z$  von  $[a, b]$  erfüllt ist.

**1.4 (Verfeinerung).** Für eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  und Zerlegungen  $Z_1, Z_2$  von  $[a, b]$  gilt:

$$Z_1 \subseteq Z_2$$

$$\implies$$

$$U_f(Z_1) \leq U_f(Z_2)$$

$$Z_1 \subseteq Z_2$$

$$\implies$$

$$O_f(Z_1) \geq O_f(Z_2)$$

Das heißt:

- (a) Je feiner Zerlegung, desto größer die Untersumme.
- (b) Je feiner Zerlegung, desto kleiner die Obersumme.

**1.5.** Nehmen wir zwei Zerlegungen  $Z_1, Z_2$ , die nicht unbedingt kompatibel sind, d.h., weder  $Z_1 \subseteq Z_2$  noch  $Z_2 \subseteq Z_1$  ist vorausgesetzt. Dann ist die Zerlegung  $Z_1 \cup Z_2$  feiner als  $Z_1$  und  $Z_2$  also erhalten wir

$$U_f(Z_1) \leq U_f(Z_1 \cup Z_2) \leq O_f(Z_1 \cup Z_2) \leq O_f(Z_2).$$

Das zeigt

$$U_f(Z_1) \leq O_f(Z_2).$$

Also: jede beliebige Untersumme ist nicht höher als jede beliebige Untersumme.

**1.6 Def.** Wir definieren das **Riemannsche Unter- bzw. Oberintegral** als

$$U_f := \sup \{U_f(Z) : Z \text{ Zerlegung von } [a, b]\},$$

$$O_f := \inf \{O_f(Z) : Z \text{ Zerlegung von } [a, b]\}.$$

**1.7.** Untersummen sind nicht höher als die Obersummen von  $f$ . Also haben wir  $U_f \leq O_f$ . Es gibt “komische” Beispiele, für die man  $U_f < O_f$  hat. Etwa, bei  $f : [0, 1] \rightarrow \mathbb{R}$  mit

$$f(x) = \begin{cases} 1 & \text{für } x \in [0, 1] \setminus \mathbb{Q}, \\ 0 & \text{für } x \in [0, 1] \cap \mathbb{Q}. \end{cases}$$

Für dieses Beispiel hat man  $U_f = 0$  und  $O_f = 1$ .

**1.8 Def.**  $f : [a, b] \rightarrow \mathbb{R}$  heißt über  $[a, b]$  **Riemann-integrierbar**, wenn  $U_f = O_f$  gilt.  
Gegebenenfalls nennt man  $U_f = O_f$  das **Riemann-Integral** von  $f$  über  $[a, b]$ . Bezeichnung:

$$\int_a^b f(x) \, dx.$$

Hierbei heißt  $a$  die **untere Integrationsgrenze**,  $b$  die **obere Integrationsgrenze**,  $x$  die **Integrationsvariable** und  $f$  der **Integrand**.

**1.9.** Ist  $f$  über  $[a, b]$  Riemann-integrierbar, so ist  $\int_a^b f(x) \, dx$  der eindeutige Wert, der zwischen allen Untersummen und Obersummen liegt, also der Wert, für welchen die Ungleichungen

$$U_f(Z_1) \leq \int_a^b f(x) \, dx \leq O_f(Z_2)$$

für alle Zerlegungen  $Z_1, Z_2$  von  $[a, b]$  erfüllt sind.

**1.10.** Riemann hat  $\int_a^b f(x) dx$  auf eine etwas andere (aber zur vorigen Definition äquivalente) Weise definiert. Für einer Zerlegung  $Z = \{x_0, \dots, x_n\}$  mit  $a = x_0 < \dots < x_n = b$  und ein System  $T$  von Punkten  $t_1, \dots, t_n$ , mit  $t_i \in [x_{i-1}, x_i]$ , führt man die sogenannte *Riemann-Summe* ein:

$$S_f(Z, T) := \sum_{i=1}^n f(t_i) \cdot (x_i - x_{i-1})$$

Dann ist  $\int_a^b f(x) dx$  der Grenzwert aller solchen Riemann-Summen im Fall, dass die Feinheit

$$\Delta(Z) := \max \{x_i - x_{i-1} : i = 1, \dots, n\}$$

gegen 0 geht.

Man beachte: Die Anzahl der Elemente in  $Z$  ist nicht fest. Wenn die Feinheit von  $Z$  gegen 0 geht, geht  $n$  gegen  $\infty$ .

Ein weiterer Kommentar: Es gilt offensichtlich

$$U_f(Z) \leq S_f(Z, T) \leq O_f(Z).$$

Ist  $f$  Riemann-integrierbar gehen bei  $\Delta(Z) \rightarrow 0$  konvergieren alle drei Summe  $U_f(Z)$  (Untersumme)  $S_f(Z, T)$  (Riemann-Summe) und  $O_f(Z)$  (Obersumme), gegen den selben Wert  $\int_a^b f(x) dx$ . Somit hat die Wahl des Punkts  $t_i$  in  $[x_{i-1}, x_i]$  bei  $\Delta(Z) \rightarrow 0$  keine großen Auswirkungen auf  $S_f(Z, T)$ .

**1.11 Def** (Orientierte Version des Riemann-Integrals). Wir haben  $\int_a^b f(x) \, dx$  für  $a < b$  definiert. Nun setzen wir

$$\int_a^b f(x) \, dx = \begin{cases} 0, & \text{für } a = b, \\ -\int_b^a f(x) \, dx & \text{für } a > b. \end{cases}$$

**1.12 Thm.** Eine Riemann-integrierbare Funktion  $f$  auf  $[a, c]$  ist auf jedem Unterintervall von  $[a, c]$  ebenfalls Riemann-integrierbar und es gilt

$$\int_a^c f(x) \, dx = \int_a^b f(x) \, dx + \int_b^c f(x) \, dx$$

für alle  $b \in [a, c]$ .

## 2 Grundeigenschaften des Riemann-Integrals

**2.1 Thm.** Sind  $f, g$  Riemann-integrierbar über  $[a, b]$  und  $\alpha, \beta \in \mathbb{R}$ , dann sind auch  $f \cdot g$  und  $\alpha f + \beta g$  Riemann integrierbar, uns es gilt:

$$\int_a^b (\alpha f(x) + \beta g(x)) \, dx = \alpha \int_a^b f(x) \, dx + \beta \int_a^b g(x) \, dx.$$

**2.2 Thm.** *Jede stetige Funktion auf  $[a, b]$  ist Riemann-integrierbar.*

**2.3 Thm.** *Jede beschränkte Funktion auf  $[a, b]$ , die in nur in endlich vielen Punkten von  $[a, b]$  nicht stetig ist, ist Riemann-integrierbar.*

**2.4.** Ungleichungen können integriert werden...

**2.5 Thm.** Sind  $f$  und  $g$  über  $[a, b]$  Riemann-integrierbar, dann folgt aus  $f(x) \leq g(x)$  für alle  $x \in [a, b]$  die Ungleichung

$$\int_a^b f(x) \, dx \leq \int_a^b g(x) \, dx.$$

**2.6 Def.** Gilt für Funktionen  $f, F : I \rightarrow \mathbb{R}$  auf einem Intervall  $I$  die Relation  $F' = f$  so nennt man  $F$  *Stammfunktion* von  $f$ .

**2.7.** Sind  $F_1, F_2$  Stammfunktion von  $f : I \rightarrow \mathbb{R}$ , dann ist ihre Differenz  $F_1 - F_2$  eine Konstante Funktion auf dem Intervall  $I$ . Das folgt aus dem Mittelwertsatz der Differentialrechnung. Die Tatsache, dass  $I$  ein Intervall ist, ist wesentlich.

**2.8 Def.** Das **unbestimmte Integral** einer Funktion  $f : I \rightarrow \mathbb{R}$  auf einem Intervall  $I$  ist die parametrische Familie aller Stammfunktionen von  $f$ . Bezeichnung:  $\int f(x) \, dx$ .

**2.9.** Das Integral  $\int_a^b f(x) \, dx$  nennt man im Gegensatz zum unbestimmten Integral  $\int f(x) \, dx$  das **bestimmte Integral**.

**2.10 Bsp.** Es gilt

$$\int \cos x \, dx = \sin x + C$$

wegen

$$(\sin x + C)' = \cos x.$$

Hier steht  $C \in \mathbb{R}$  für eine Konstante.

Der Ausdruck  $\sin x + C$  steht hier eigentlich für die Menge  $\{\sin x + C : C \in \mathbb{R}\}$  aller Stammfunktion von  $\cos x$ , es ist aber üblich in solchen Situation die Schreibweise  $\sin x + C$  zu benutzen.

**2.11.** Die Formeltafeln für die Stammfunktionen sind im Wesentlichen die Formeltafeln für Ableitungen, die man “rückwärts” aufgeschrieben hat. Zum Beispiel:

$$\begin{aligned}
 (\sin x)' = \cos x &\iff \int \cos x \, dx = \sin x + C, \\
 (\cos x)' = -\sin x &\iff \int \sin x \, dx = -\cos x + C, \\
 (x^p)' = px^{p-1} &\iff \int x^a \, dx = \frac{x^{a+1}}{a+1} + C \quad (\text{für } a \neq -1), \\
 &\vdots
 \end{aligned}$$

**2.12.** Das unbestimmte und das bestimmte Integral hängen zusammen...

**2.13 Thm** (Hauptsatz der Differential- und Integralrechnung). *Ist  $F$  Stammfunktion von  $f : [a, b] \rightarrow \mathbb{R}$ , so gilt*

$$\int_a^b f(x) \, dx = F(b) - F(a).$$

**2.14.** Als Folgerung können wir die Ableitung von  $\int_a^x f(t) \, dt$  nach  $x$  bestimmen. Ist  $F$  Stammfunktion von  $f$ , so gilt

$$\int_a^x f(t) \, dt = F(x) - F(a).$$

Es folgt

$$\left( \int_a^x f(t) \, dt \right)' = (F(x) - F(a))' = F'(x) = f(x).$$

Wir haben gezeigt, dass  $\int_a^x f(t) \, dt$  als Funktion in  $x$  Stammfunktion von  $f(x)$  ist.

**2.15.** Für  $F(b) - F(a)$  benutzt man die folgenden Bezeichnungen:  $F(x)\Big|_{x=a}^b$  oder  $F(x)\Big|_a^b$  oder  $[F(x)]_a^b$  oder  $[F(x)]_{x=a}^b$ .

**2.16.** Im Folgenden werden Regeln diskutieren, die man beim Ausrechnen von Integralen benutzen kann. Man muss aber gewarnt werden, dass nicht jede elementare Funktion eine elementare Stammfunktion besitzt. Eines der bekanntesten Beispiele ist die Funktion  $e^{-x^2}$ . Es gilt

$$\int e^{-x^2} \mathrm{d}x = F(x) + C,$$

für die Stammfunktion  $F$  (das ist die Funktion mit  $F'(x) = e^{-x^2}$ ) hat man aber keine elementare Formel. Das bedeutet unter anderem, dass man zur Berechnung des bestimmten Integrals

$$\int_a^b e^{-x^2} \mathrm{d}x$$

für gegebene  $a$  und  $b$  numerische Approximationen benutzen muss (Stichwort: Numerisches Integrieren).

**2.17 Bsp.** Wir berechnen das Volumen eines Kegels  $K$  in  $\mathbb{R}^3$  in Abhängigkeit von der Grundfläche und der Höhe. Sei  $A$  der Flächeninhalt der Grundfläche des Kegels und  $h$  die Höhe.

Wir platzieren den Kegel im Raum  $\mathbb{R}^3$  so, dass die Spitze in  $(0, 0, 0)$  ist und die Grundfläche in der Ebene  $\{h\} \times \mathbb{R}^2$  liegt. Für  $x \in [0, h]$  hat der Querschnitt des Kegels mit der Ebene  $\{x\} \times \mathbb{R}^2$  die Fläche

$$f(x) := \left(\frac{x}{h}\right)^2 A,$$

Wenn wir im Bereich  $[0, h]$  eine Zerlegung  $Z = \{x_0, \dots, x_n\}$  mit  $0 = x_0 < \dots < x_n \leq n$  einführen, dann ist der “Ausschnitt”  $\{(x, y, z) \in K : x_{i-1} \leq x \leq x_i\}$  des Kegels  $K$  ein Kegelstumpf, den man zwischen zwei Zylindern schachteln kann: die beiden Zylinder haben Höhe  $x_i - x_{i-1}$ , der kleinere Zylinder hat die Fläche  $f(x_{i-1})$  und der größere Zylinder hat die Fläche  $f(x_i)$ . Wir haben also das Volumen  $V$  des Kegels durch

$$\sum_{i=1}^n (x_i - x_{i-1}) f(x_{i-1}) = U_f(Z) \leq V \leq O_f(Z) \leq \sum_{i=1}^n (x_i - x_{i-1}) f(x_i)$$

approximiert. Da man  $U_f(Z) \leq V \leq O_f(Z)$  hat und die Zerlegung  $Z$  beliebig ist, sehen wir dass  $V = \int_0^h f(x) \, dx$  ist.

Wir erhalten also für das Volumen die Beschreibung

$$\begin{aligned} V &= \int_0^h \left(\frac{x}{h}\right)^2 A \, dx \\ &= \frac{A}{h^2} \int_0^h x^2 \, dx \\ &= \frac{A}{h^2} \left[ \frac{1}{3}x^3 \right]_{x=0}^h \\ &= \frac{A}{h^2} \cdot \frac{h^3}{3} \\ &= \frac{1}{3}hA. \end{aligned}$$

### 3 Substitutionsregel

**3.1 (Substitution).** Sei  $F$  Stammfunktion von  $f$  und sei  $g$  differenzierbar. Die Kettenregel ergibt:

$$(F \circ g)' = (F' \circ g) \cdot g' = (f \circ g) \cdot g'.$$

Das heißt:  $F \circ g$  ist eine Stammfunktion von  $(f \circ g) \cdot g'$ . Das kann man als die Gleichung

$$\int f(g(x))g'(x) \, dx = F(g(x)) + C$$

für das unbestimmte Integral von  $f(g(x))g'(x)$  hinschreiben. Durch das Anwenden des Hauptsatzes ergibt sich daraus die Substitutionsregel für bestimmte Integrale:

$$\int_a^b f(g(x))g'(x) \, dx = \int_{g(a)}^{g(b)} f(y) \, dy.$$

Tatsächlich: Nach dem Hauptsatz kann man die linke Seite als  $(F \circ g)(b) - (F \circ g)(a)$  umschreiben und die rechte Seite als  $F(g(b)) - F(g(a))$ . Das ist aber dasselbe.

**3.2 (Leibniz'sche Sichtweise).** Erinnern wir uns an die Schreibweise von Leibniz:

$$q'(x) = \frac{d q(x)}{d x}.$$

Wenn wir nun diese Gleichung rein formal als  $d q(x) = q'(x) d x$  umschreiben, so werden wir motiviert, die folgende Bezeichnung einzuführen:

$$\int p(x) d q(x) := \int p(x) q'(x) d x$$

Hier noch die analoge Bezeichnung für bestimmte Integrale:

$$\int_a^b p(x) d q(x) := \int_a^b p(x) q'(x) d x$$

Mit diesen Bezeichnungen kann nun die Substitutionsregel so beschrieben werden.

- *Für unbestimmte Integrale.* Zum Berechnen von

$$\int f(g(x)) d g(x)$$

reicht es aus,

$$\int f(y) \, dy$$

zu berechnen ( $y$  ist neue Integrationsvariable) und dann

$$y = g(x)$$

einsetzen. Als Formel:

$$\int f(g(x)) \, dg(x) = \left( \int f(y) \, dy \right) \Big|_{y=g(x)}.$$

- *Bestimmte Integrale.*

$$\int_a^b f(g(x)) \, dg(x) = \int_{g(a)}^{g(b)} f(y) \, dy.$$

**3.3 Bsp** (Substitution für unbestimmte Integrale).

$$\begin{aligned} \int e^{-x^2} x \, dx &= -\frac{1}{2} \int e^{-x^2} d(-x^2) && | \text{ wegen } (x^2)' = 2x \\ &= -\frac{1}{2} \int e^y \, dy && | \text{ mit } y = -x^2. \\ &= -\frac{1}{2} e^y + C && | e^y \text{ Stammfunktion von } e^y \\ &= -\frac{1}{2} e^{-x^2} + C && | y = -x^2 \text{ einsetzen.} \end{aligned}$$

**3.4 Bsp** (Substitution für bestimmte Integrale).

$$\begin{aligned}\int_0^{\pi/3} \tan(x) \, dx &= \int_0^{\pi/3} \frac{\sin x \, dx}{\cos x} && | \text{ tan}(x) \text{ ausschreiben} \\ &= - \int_0^{\pi/3} \frac{d \cos x}{\cos x} && | \text{ wegen } (\cos x)' = -\sin x \\ &= - \int_{\cos 0}^{\cos \pi/3} \frac{dy}{y} && | \text{ mit } y = \cos x \\ &= - \int_1^{1/2} \frac{dy}{y} && | \text{ Grenzen ausgerechnet} \\ &= - \left[ \ln y \right]_1^{1/2} \\ &= -(\ln 1/2 - \ln 1) \\ &= \ln 2.\end{aligned}$$

## 4 Partielle Integration

**4.1.** Partielle Integration ist die Umkehrung der Produktregel für das Differenzieren. Die Produktregel besagt  $(u(x)v(x))' = u'(x)v(x) + v'(x)u(x)$ . Das bedeutet

$$\int (u'(x)v(x) + u(x)v'(x)) \, dx = u(x)v(x) + C.$$

Somit ergibt sich die Gleichung

$$\int u(x)v'(x) \, dx = u(x)v(x) - \int u'(x)v(x) \, dx$$

für zwei unbestimmte Integrale. Die Ableitung “wandert” von  $v$  auf  $u$ , was in manchen Situationen zur einer Vereinfachung des Integrals führt.

Wir können die Gleichung auch in den folgenden Bezeichnungen formulieren:

$$\int u(x) \, dv(x) = u(x)v(x) - \int v(x) \, du(x)$$

## 4.2 Bsp (Partielle Integration für unbestimmte Integrale).

$$\begin{aligned}
 \int x^2 e^x \, dx &= \int x^2 \, d(e^x) && | \text{ wegen } (e^x)' = e^x \\
 &= x^2 e^x - \int e^x \, dx^2 && | \text{ partiell integrieren} \\
 &= x^2 e^x - 2 \int x e^x \, dx && | \text{ wegen } (x^2)' = 2x \\
 &= x^2 e^x - 2 \int x \, d(e^x) && | \text{ wegen } (e^x)' = e^x \\
 &= x^2 e^x - 2x e^x + 2 \int e^x \, dx && | \text{ partiell integrieren} \\
 &= x^2 e^x - 2x e^x + 2e^x + C && | \text{ wegen } (e^x)' = e^x \\
 &= (x^2 - 2x + 2)e^x + C.
 \end{aligned}$$

**4.3 Bsp.**

$$\begin{aligned}\int \ln x \, dx &= x \ln x - \int x \, d \ln x && | \text{ partielle Integration} \\ &= x \ln x - \int x \frac{1}{x} \, dx && | \text{ wegen } (\ln x)' = \frac{1}{x} \\ &= x \ln x - \int \, dx \\ &= x \ln x - x + C\end{aligned}$$

**4.4.** Partielle Integration für bestimmte Integrale:

$$\int_a^b u(x)v'(x) \, dx = \left[ u(x)v(x) \right]_{x=a}^b - \int_a^b u'(x)v(x) \, dx.$$

## 5 Uneigentliche Integrale

**5.1 Bsp.** Das Integral  $\int_0^1 x^{-1/2} dx$  existiert im Rahmen der oben angeführten Theorie der Riemann-Integrale nicht. Die Funktion  $x^{-1/2}$  ist an der unteren Grenze 0 nicht definiert. Auch wenn wir die Funktion auf die Stelle erweitern würden, wäre eine solche Erweiterung über  $[0, 1]$  nicht Riemann-integrierbar. Wir können das Integral als ein sogenanntes **uneigentliches Integral** folgendermaßen einführen:

$$\int_0^1 x^{-1/2} dx = \lim_{h \rightarrow 0+} \int_h^1 x^{-1/2} dx.$$

Weil man

$$\int_h^1 x^{-1/2} dx = \left[ 2x^{1/2} \right]_{x=h}^1 = 2 - 2\sqrt{h}$$

hat, erhalten wir

$$\int_0^1 x^{-1/2} dx = \lim_{h \rightarrow 0+} (2 - 2\sqrt{h}) = 2.$$

**5.2 Bsp.** Wir können das Integrieren über unbeschränkte Intervalle ebenfalls als uneigentliche Integrale einführen. Zum Beispiel:

$$\int_0^\infty e^{-x} dx = \lim_{t \rightarrow \infty} \int_0^t e^{-x} dx.$$

Weil man

$$\int_0^t e^{-x} dx = \left[ -e^{-x} \right]_{x=0}^t = 1 - e^{-t}$$

hat, erhalten wir

$$\int_0^\infty e^{-x} dx = \lim_{t \rightarrow \infty} (1 - e^{-t}) = 1.$$

Diese Berechnung wird in der Regel folgendermaßen aufgeschrieben:

$$\int_0^\infty e^{-x} dx = \left[ -e^{-x} \right]_{x=0}^\infty = 1 - 0 = 1,$$

wobei man unter dem Einsetzen von  $\infty$  einen Grenzwertübergang meint.

**5.3 Def** (Verschiedene uneigentliche Integrale). Seien  $-\infty < a < b < \infty$ . Wir definieren die uneigentlichen Integrale:

$$\int_a^\infty f(x) dx := \lim_{t \rightarrow \infty} \int_a^t f(x) dx,$$

$$\int_{-\infty}^b f(x) dx := \lim_{t \rightarrow -\infty} \int_t^b f(x) dx,$$

$$\int_{-\infty}^\infty f(x) dx := \int_{-\infty}^0 f(x) dx + \int_0^\infty f(x) dx,$$

$$\int_a^b f(x) dx := \lim_{h \rightarrow 0+} \int_{a+h}^b f(x) dx,$$

$$\int_a^b f(x) dx := \lim_{h \rightarrow 0+} \int_a^{b-h} f(x) dx,$$

bei  $\lim_{h \rightarrow 0+} |f(a+h)| = \infty$ .

bei  $\lim_{h \rightarrow 0+} |f(b-h)| = \infty$ .

**5.4.** Da uneigentliche Integrale Grenzwerte sind, kann man für sie die Terminologie aus der Theorie der Grenzwert benutzen (Konvergenz, Divergenz, bestimmte Divergenz).

**5.5 Bsp.**

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{dx}{x^2 + 1} &= \left[ \arctan x \right]_{x=-\infty}^{\infty} \\&= \lim_{x \rightarrow \infty} \arctan x - \lim_{x \rightarrow -\infty} \arctan x \\&= \frac{\pi}{2} - \left( -\frac{\pi}{2} \right) \\&= \pi\end{aligned}$$

## 6 Exkurs in Lebesgue-Integrale

**6.1.** Uneigentliche Integrale sind eine “schnelle Reparatur” einiger Probleme, die man mit der Theorie der Riemann-Integrale hat. Man kann über unbeschränkten Bereichen nicht direkt nach Riemann integrieren, man kann viele (gute) unbeschränkte Funktionen nicht direkt nach Riemann integrieren.

Eine “grundätzliche Reparatur” bietet die Theorie von **Lebesgue-Integralen**. Lebesgue-Integrierbarkeit ist eine viel weniger einschränkende Eigenschaft einer Funktion. In der Theorie der Lebesgue-Integrale müssen die Integrale

$$\int_0^1 x^{-1/2} \, dx = 2, \quad \text{und} \quad \int_0^\infty e^{-x} \, dx = 1$$

nicht gesondert behandelt werden: die beiden Funktionen sind über die jeweiligen Bereichen Lebesgue-integrierbar.

## 6.2. Basics zu Lebesgue-Integralen:

- (a) In der Definition des Lebesgue-Integrals zerlegt man nicht den Defintionsbereich sondern den Wertebereich in Stückchen.
- (b) Definition des Lebesgue-Integrals für Funktion einer Variable basiert auf dem Lebesgue-Maß in  $\mathbb{R}$ , d.h., für mache Teilmengen von  $\mathbb{R}$  kann ihre Länge (im Lebesgue-Sinne) messen. Hierbei ist die Länge von  $[a, b]$  gleich  $b - a$ . Die Länge von  $[a, b] \cap \mathbb{Q}$  ist aber zum Beispiel 0, weil  $[a, b] \cap \mathbb{Q}$  aus abzählbar vielen einelementigen Mengen zusammengesetzt ist, deren Länge gleich 0 ist.
- (c) Wenn die Gleichung  $\int_a^b f(x) \, dx = I$  nach der Definition von Riemann gilt, nach gilt sie auch nach der Definition von Lebesgue. Das heißt, das Riemann-Integral ist eine Ereweiterung des Lebesgue-Integrals.

**6.3** (Lebesgue-Integral und Wahrscheinlichkeitstheorie). Für die Wahrscheinlichkeitstheorie reicht die Theorie der Riemann-Integrale nicht aus. Lebesgue-Integrale sind mit Hinblick auf Anwendungen in Wahrscheinlichkeitstheorie sehr natürliche Objekte.

## 7 Abschätzung von Summen und Reihen durch Integrale

**7.1.** Summen, Reihe und Integrale sind alles Varianten von Summen. Obwohl Integral umsetzungstechnisch die kompliziertesten Varianten von Summen sind, sind sie oft einfacher zu Berechnen als Summen oder Reihen. Was man also machen kann: Summen und Reihen durch Integrale Abschätzen.

**7.2 Bsp** (Partialsumme der harmonischen Reihe). Wir geben Abschätzungen für

$$s_n := \sum_{k=1}^n \frac{1}{k}$$

Obere Schranke:

$$\begin{aligned}
 s_n &= \sum_{k=1}^n \frac{1}{k} \\
 &= 1 + \sum_{k=2}^n \frac{1}{k} \\
 &= 1 + \sum_{k=2}^n \int_{k-1}^k \frac{1}{k} dx && | \text{ wegen } \int_{k-1}^k dx = 1 \\
 &\leq 1 + \sum_{k=2}^n \int_{k-1}^k \frac{1}{x} dx && | \text{ wegen } x \leq k \text{ für } x \in [k-1, k] \\
 &= 1 + \int_1^n \frac{1}{x} dx && | \text{ Integrale zusammenfassen} \\
 &= 1 + \left[ \ln x \right]_{x=1}^n \\
 &= 1 + \ln n.
 \end{aligned}$$

Das ergibt

$$s_n \leq 1 + \ln n.$$

Untere Schranke:

$$\begin{aligned}
 s_n &= \sum_{k=1}^n \int_k^{k+1} \frac{1}{k} dx && | \text{ wegen } \int_k^{k+1} dx = 1. \\
 &\geq \sum_{k=1}^n \int_k^{k+1} \frac{1}{x} dx && | \text{ wegen } x \geq k \text{ für } x \in [k, k+1]. \\
 &= \int_1^{n+1} \frac{1}{x} dx && | \text{ alle } n \text{ Integrale zusammengefügt} \\
 &= \left[ \ln x \right]_{x=1}^{n+1} \\
 &= \ln(n+1).
 \end{aligned}$$

Das ergibt:

$$s_n \geq \ln(n + 1).$$

**7.3 Bsp.** Wir wissen, dass die harmonische Reihe  $\sum_{k=1}^{\infty} \frac{1}{k}$  divergent ist (das vorige Beispiel war ein weiterer Nachweis). Die harmonische Reihe ist Spezialfall  $s = 1$  der Reihe  $\sum_{k=1}^{\infty} \frac{1}{k^s}$ .

Wir zeigen nun, dass  $\sum_{k=1}^{\infty} \frac{1}{k^s}$  für  $s > 1$  konvergent ist.

$$\begin{aligned}\sum_{k=2}^{\infty} \frac{1}{k^s} &= \sum_{k=2}^{\infty} \int_{k-1}^k \frac{1}{x^s} dx \\ &\leq \sum_{k=2}^{\infty} \int_{k-1}^k \frac{1}{x^s} dx \\ &= \int_1^{\infty} \frac{1}{x^s} dx \\ &= \left[ \frac{1}{1-s} x^{1-s} \right]_{x=1}^{\infty} \\ &= \frac{1}{s-1}\end{aligned}$$

Das ergibt die Abschätzung

$$\sum_{k=1}^{\infty} \frac{1}{k^s} = 1 + \sum_{k=2}^{\infty} \frac{1}{k^s} \leq 1 + \frac{1}{s-1} = \frac{s}{s-1}$$

**7.4 Aufgabe.** Geben Sie eine untere Abschätzung an  $\sum_{k=2}^{\infty} \frac{1}{k^s}$  für  $s > 1$  durch eine Funktion in  $s$ .

**7.5 Aufgabe.** Nehmen wir an, Sie sind mit der Qualität der Abschätzung

$$\sum_{k=1}^{\infty} \frac{1}{k^s} \leq \frac{s}{s-1}$$

für  $s > 1$  nicht ganz zufrieden. Sie erhalten zum Beispiel, für  $s = \frac{11}{10}$ .

$$\sum_{k=1}^{\infty} \frac{1}{k^{11/10}} \leq 11$$

und würden gerne die 11 durch eine bessere Schranke ersetzen. Wie würden Sie vorgehen?

**7.6.** Formulieren wir die vorigen Abschätzungen als ein allgemeines Prinzip. Sei  $f : [0, +\infty) \rightarrow [0, +\infty)$  monoton fallend und sei  $\int_0^\infty f(x) \, dx < +\infty$ . Dann gilt

$$\int_1^\infty f(x) \, dx \leq \sum_{k=1}^{\infty} f(k) \leq \int_0^\infty f(x) \, dx.$$

Man kann also die Summe  $\sum_{k=1}^{\infty} f(k)$ , die man evtl. nicht exakt berechnen kann, durch die Berechnung der beiden Integrale von  $f(x)$  (wenn man etwa die Stammfunktion von  $f(x)$  exakt bestimmen kann) abschätzen.

**7.7 Aufgabe.** Vergewissern Sie sich, dass die Abschätzungen aus der vorigen Bemerkung tatsächlich stimmen.

## 8 Integrale rationaler Funktionen

**8.1 Def. Rationale Funktion** ist Quotient von Polynomen.

**8.2 Bsp.** Wir lesen die Formeltafeln für Ableitungen und bestimmen dadurch die Stammfunktion von  $\frac{1}{x}$ :

$$\int \frac{dx}{x} = \ln|x| + C. \quad (\text{V.1})$$

- Man beachte, dass weder  $\frac{1}{x}$  noch  $\ln|x|$  in 0 definiert ist.
- Formel (V.1) bedeutet: fixiert man ein Intervall  $I$ , das 0 nicht enthält, so ist

$$\{\ln|x| + C : C \in \mathbb{R}\}$$

die Menge aller Funktionen auf  $I$ , deren Ableitung gleich  $\frac{1}{x}$  ist.

- Formel (V.1) kann für die Berechnung bestimmter Integrale benutzt werden, man sollte

aber die oben angeführten Kommentare beachten:

$$\int_1^e \frac{dx}{x} = \ln|e| - \ln|1| = 1 \quad | \text{ alles okay}$$

$$\int_{-4}^{-2} \frac{dx}{x} = \ln|-2| - \ln|-4| = -\ln 2 \quad | \text{ alles okay}$$

$$\int_{-3}^1 \frac{dx}{x} = \ln \cancel{|1|} - \ln \cancel{|-3|} \quad | \text{ nicht okay wegen } 0 \in [-3, 1]$$

**8.3 Bsp.** Für  $a \in \mathbb{R} \setminus \{0\}$  und  $b \in \mathbb{R}$  kriegt man

$$\int \frac{dx}{ax + b} = \frac{1}{a} \ln |ax + b| + C.$$

durch die Substitution  $y = ax + b$ .

**8.4 Bsp.** Für  $p \neq -1$  kriegt man

$$\int x^p \, dx = \frac{x^{p+1}}{p+1} + C,$$

indem man die Formeltafeln für Ableitungen rückwärts liest. Zum Beispiel:  $(x^{-3})' = -3x^{-4}$  heißt, dass  $-\frac{1}{3}x^{-3}$  Stammfunktion von  $x^{-4}$  ist. Das schreibt man als

$$\int x^{-4} \, dx = -\frac{1}{3}x^{-3} + C.$$

**8.5 Bsp.** Wir berechnen

$$\int \frac{x+1}{x^2+x-6} dx.$$

Mit der  $pq$ -Formel kann  $x^2 + x - 6$  als

$$x^2 + x - 6 = (x - 2)(x + 3)$$

faktorisiert werden. Das zeigt

$$\frac{x+1}{x^2+x-6} = \frac{x+1}{(x-2)(x+3)}.$$

Nun kann man den Integranden als die Linearkombination

$$\frac{x+1}{(x-2)(x+3)} = \frac{A}{x-3} + \frac{B}{x-2}$$

für gewisse  $A, B \in \mathbb{R}$  darstellen. Wie bestimmt man  $A$  und  $B$ ? Erstmal die Gleichung mit

$(x - 2)(x + 3)$  multiplizieren, sodass man auf die Gleichung

$$x + 1 = A \cdot (x - 2) + B \cdot (x - 3) \quad (\text{V.2})$$

kommt. Nun kann man  $A$  und  $B$  aus der Gleichung (V.2) durch den Koeffizientenvergleich bestimmen. Dafür müssen wir dann lineares Gleichungssystem mit zwei Gleichungen und den Unbekannten  $A$  und  $B$  lösen.

Noch einfacher geht es, wenn wir die Gleichung (V.2) an  $x = 2$  und  $x = 3$  auswerten. Dann erhalten wir

$$2 + 1 = A \cdot (2 - 2) + B \cdot (2 - 3) = -B, \quad | \text{ für } x = 2$$

$$3 + 1 = A \cdot (3 - 2) + B \cdot (3 - 3) = A. \quad | \text{ für } x = 3$$

Wir haben nun die Darstellung

$$\frac{x + 1}{(x - 2)(x + 3)} = \frac{4}{x - 3} - \frac{3}{x - 2}$$

ermittelt. Unser ursprüngliches Integral wird also eine linear Kombination von zwei einfacheren Integralen dargestellt, die wir bereits kennen:

$$\begin{aligned} \int \frac{x+1}{x^2+x-6} dx &= 4 \int \frac{dx}{x-3} - 3 \int \frac{dx}{x-2} \\ &= 4 \ln|x-3| - 3|x-2| + C. \end{aligned} \quad (\text{V.3})$$

Die Formel (V.3) kann auch zur Berechnung bestimmter Integrale benutzt werden, solange weder 3 noch 2 im Integrationsbereich enthalten sind:

$$\int_a^b \frac{x+1}{x^2+x-6} dx = \left[ 4 \ln|x-3| - 3|x-2| \right]_{x=a}^b$$

unter der Voraussetzung  $[a, b] \cap \{2, 3\} = \emptyset$ .

## 8.6 Bsp.

Wir berechnen

$$\int \frac{x^3 + 2x^2 - 1}{x^2 + x - 6} dx.$$

Hier hat das Zähler-Polynom einen größeren Grad als das Nenner-Polynom. Wir dividieren Polynome mit Rest:

$$\begin{array}{r} (x^3 + 2x^2 - 1) \div (x^2 + x - 6) = x + 1 + \frac{5x + 5}{x^2 + x - 6} \\ \hline -x^3 - x^2 + 6x \\ \hline x^2 + 6x - 1 \\ \hline -x^2 - x + 6 \\ \hline 5x + 5 \end{array}$$

Das Originalintegral wird durch zwei einfachere Integrale ersetzt.

$$\int \frac{x^3 + 2x^2 - 1}{x^2 + x - 6} dx = \int (x + 1) dx + 5 \int \frac{x + 1}{x^2 + x - 6} dx.$$

Die Stammfunktion von  $x+1$  ist  $\frac{1}{2}(x+1)^2$ . Die Stammfunktion von  $(x+1)/(x^2+x-6)$  haben wir im vorigen Beispiel ausgerechnet.

**8.7 Bsp.**

$$\int \frac{dx}{x^2 + 1} = \arctan x + C,$$

denn, wie wir wissen, ist  $(\arctan x)' = \frac{1}{1+x^2}$ .

**8.8 Bsp.**

$$\begin{aligned}
 \int \frac{dx}{x^2 + 4} &= \frac{1}{4} \int \frac{dx}{(x/2)^2 + 1} && | \text{ wir brauchen } (\dots)^2 + 1 \text{ im Nenner} \\
 &= \frac{1}{2} \int \frac{d(x/2)}{(x/2)^2 + 1} && | y = x/2 \\
 &= \frac{1}{2} \int \frac{dy}{y^2 + 1} \\
 &= \frac{1}{2} \arctan y + C \\
 &= \frac{1}{2} \arctan(x/2) + C.
 \end{aligned}$$

Natürlich kann man auf dieser Berechnung basierend auch eine allgemeine Formel erstellen:

$$\int \frac{dx}{x^2 + a} = \frac{1}{\sqrt{a}} \arctan \frac{x}{\sqrt{a}} + C \quad \text{für } a > 0.$$

## 8.9 Bsp.

Wir berechnen

$$\int \frac{dx}{x^2 - 6x + 18}$$

Die  $pq$ -Formel sagt uns, dass  $x^2 - 6x + 18$  keine reellen Nullstellen hat. Über reellen Zahlen können wir also  $x^2 - 6x + 18$  nicht in lineare Faktoren zerlegen. Wir machen die sogenannte quadratische Ergänzung. Dafür müssen wir die 18 so in eine Summe zerlegen, dass wir mit einem Summanden zusammen mit  $x^2 - 6x$  die binomische Formel einsetzen können:

$$\begin{aligned} x^2 - 6x + 18 &= x^2 - 6x + 9 + 9 \\ &= (x - 3)^2 + 9 \end{aligned} \quad | \text{ binomische Formel}$$

Diese Darstellung des Nennerpolynoms, hilft uns das Integral zu berechnen:

$$\begin{aligned} \int \frac{dx}{x^2 - 6x + 18} &= \int \frac{dx}{(x-3)^2 + 9} \\ &= \int \frac{d(x-3)}{(x-3)^2 + 9} \\ &= \int \frac{dy}{y^2 + 9} && | \text{ mit } y = x - 3 \\ &= \frac{1}{3} \arctan \frac{y}{3} + C && | \text{ nach dem vorigen Beispiel} \\ &= \frac{1}{3} \arctan \frac{x-3}{3} + C. \end{aligned}$$

**8.10 Bsp.** Es gilt

$$\int \frac{(x+1) \, dx}{x^2 + 1} = \int \frac{x \, dx}{x^2 + 1} + \int \frac{dx}{x^2 + 1}$$

mit

$$\int \frac{x \, dx}{x^2 + 1} = \frac{1}{2} \int \frac{d(x^2 + 1)}{x^2 + 1} = \frac{1}{2} \ln(x^2 + 1) + C.$$

und

$$\int \frac{dx}{x^2 + 1} = \arctan x + C.$$

## 9 Integrale mit trigonometrischen Funktionen

## 9.1 Bsp. Zur Berechnung von

$$\int \cos^2 x \, dx.$$

können wir  $\cos^2 x$  umformen. Durch

$$\begin{aligned} \cos 2x &= \cos^2 x - \sin^2 x && | \text{ Kosinus des doppelten Winkels} \\ &= \cos^2 x - (1 - \cos^2 x) = 2\cos^2 x - 1 && | \text{ Trigonometrischer Pythagoras} \end{aligned}$$

erhalten wir eine Darstellung von  $\cos^2 x$  mit Hilfe von  $\cos 2x$ . Diese Darstellung benutzten wir nun:

$$\int \cos^2 x \, dx = \frac{1}{2} \int (1 + \cos 2x) \, dx = \frac{1}{2}x + \frac{1}{4}\sin 2x + C.$$

Hier noch eine Nebenbemerkung, die wir später bei Fourier-Reihen gebrauchen werden: es folgt

$$\int_a^b \cos^2 x \, dx = \pi \quad \text{im Fall } b - a = 2\pi.$$

**9.2 Aufgabe.** Berechnen Sie

$$\int \sin^2 x \, dx.$$

### 9.3 Bsp. Wir berechnen

$$\int \cos^3 x \, dx.$$

Wegen  $\cos x \, dx = d(\sin x)$  und  $\cos^2 x = 1 - \sin^2 x$  haben wir

$$\begin{aligned} \int \cos^3 x \, dx &= \int (1 - \sin^2 x) \, d \sin x \\ &= \int (1 - y^2) \, dy && | \text{ mit } y = \sin x \\ &= y - \frac{1}{3}y^3 + C \\ &= \sin x - \frac{1}{3}\sin^3 x + C. \end{aligned}$$

# **Kapitel VI**

## **Fourier-Reihen**

### **1 Basics**

**1.1 Def.** Sei  $\omega > 0$ . Eine **Fourier-Reihe** ist Funktionsreihe der Form

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos k\omega x + b_k \sin k\omega x).$$

Partialsummen

$$\frac{a_0}{2} + \sum_{k=1}^n (a_k \cos k\omega x + b_k \sin k\omega x)$$

von Fourier-Reihen nennt man **trigonometrische Polynome**.

**1.2 Def.** Sei  $T > 0$ . Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  heißt  **$T$ -periodisch** wenn  $f(x + T) = f(x)$  für alle  $x \in \mathbb{R}$  erfüllt ist. Den Wert  $T$  nennt man **Periode** von  $f$ .

**1.3.** Wieso  $\frac{a_0}{2}$  und nicht einfach  $a_0$ ? Die Konstante  $a_0$  würde natürlich auch gehen. Der Faktor  $\frac{1}{2}$  vor  $a_0$  macht aber später die Formeln einheitlicher.

**1.4.** Glieder der Fourierreihe sind  $T$ -periodisch mit  $T = \frac{2\pi}{\omega}$ .

**1.5.**  $\omega$  ist lediglich ein Skalierungsfaktor für  $x$ . Daher reicht es aus, den Fall  $\omega = 1$  zu betrachten.

**1.6.** Eine  $T$ -periodische Funktion ist durch die Angabe auf einem Intervall  $[a, b]$  der Länge  $T$  eindeutig bestimmt.

Umgekehrt: eine Funktion  $f : [a, b] \rightarrow \mathbb{R}$  auf einem Intervall der Länge  $T$  und mit  $f(a) = f(b)$ , kann  $T$ -periodisch auf die gesamte reelle Achse eindeutig erweitert werden.

**1.7.** Ist  $[a, b]$  Intervall der Länge  $T > 0$  und  $f : \mathbb{R} \rightarrow \mathbb{R}$  eine  $T$ -periodische Funktion, die über  $[0, T]$  integrierbar ist, so gilt

$$\int_a^b f(x) \, dx = \int_0^T f(x) \, dx,$$

Das bedeutet: das Integral über jedes Intervall der Länge  $T$  ist gleich.

## 2 Der Raum $\mathcal{L}^2(0, T)$

**2.1 Def.** Wir führen den Vektorraum  $\mathcal{L}^2(0, T)$  ein, als den  $\mathbb{R}$ -Vektorraum aller  $T$ -periodischen Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$ , die auf  $[0, T]$ , Lebesgue-integrierbar sind, und die Bedingung

$$\int_0^T |f(x)|^2 dx < +\infty.$$

erfüllen. In diesem Raum führen wir das Skalarprodukt

$$\langle f, g \rangle_{\mathcal{L}^2(0, T)} = \int_0^T f(x)g(x) dx$$

ein und die zugehörige Norm

$$\|f\|_{\mathcal{L}^2(0, T)} := \sqrt{\int_0^T f(x)^2 dx}.$$

Zwei Funktionen  $f, g \in \mathcal{L}^2(0, T)$  werden identifiziert, wenn man  $\|f - g\|_{\mathcal{L}^2(0, T)} = 0$  hat. Das ist z.B. der Fall, wenn sich  $f$  und  $g$  innerhalb von  $[0, T]$  nur in endlich vielen Stellen unterscheiden.

**2.2.** In der vorigen Definition benutzen wir das Lebesgue-Integral, das wir in diesem Kurs gar nicht eingeführt haben. Das ist für uns kein Problem: für die Beispiele, die wir betrachten, reicht die Theorie der Riemann-Integrale aus.

**2.3** (Exkurs in die Lineare Algebra - Orthogonalbasen). Orthogonalbasen sind sehr praktisch. Ist  $b_1, \dots, b_n$  Orthogonalbasis eines Euklidischen Raums so kann man die Zerlegung

$$f = \alpha_1 b_1 + \cdots + \alpha_n b_n$$

für einen gegebenen Vektor  $f$  durch die Berechnung der Skalarprodukte bestimmen. Skalarmultiplikation mit  $b_i$  ergibt

$$\alpha_i = \frac{\langle f, b_i \rangle}{\langle b_i, b_i \rangle}$$

Bei Orthonormalsystemen hat man  $\langle b_i, b_i \rangle = 1$ .

## 2.4. Die Funktionen des Systems

$$\frac{1}{2}, \cos \omega x, \cos 2\omega x, \cos 3\omega x, \dots, \sin \omega x, \sin 2\omega x, \sin 3\omega x, \dots \quad (\text{VI.1})$$

bilden eine Art ‘‘unendliche Orthogonalbasis’’ für den Raum  $\mathcal{L}^2(0, T)$  mit  $T = \frac{2\pi}{\omega}$ . Das bedeutet:

- (a) Je zwei Funktionen dieses Systems sind stets orthogonal bzgl. des Skalarprodukts von  $\mathcal{L}^2(0, T)$ .
- (b) Jedes  $f \in \mathcal{L}^2(0, T)$  ist eine ‘‘unendliche Linearkombination’’ der Funktionen des Systems.  
Das heißt: es gibt eindeutige  $a_0, a_1, \dots, b_0, b_1, \dots \in \mathbb{R}$  mit

$$f = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos k\omega x + b_k \sin k\omega x),$$

wobei man diese Gleichung im Raum  $\mathcal{L}^2(0, T)$  als den Grenzwert

$$\lim_{n \rightarrow \infty} \left\| f - \frac{a_0}{2} - \sum_{k=1}^n (a_k \cos k\omega x + b_k \sin k\omega x) \right\|_{\mathcal{L}^2(0, T)} = 0$$

interpretiert.

**2.5 Aufgabe.** Verifizieren Sie, dass (VI.1) tatsächlich ein orthogonales System bilden. Berechnen Sie auch die  $\mathcal{L}^2(0, T)$ -Norm jeder Funktion aus (VI.1).

### 3 Formeln für die Koeffizienten

**3.1** (Formeln für die Koeffizienten der Fourier-Reihe). Für

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos k\omega x + b_k \sin k\omega x) \in \mathcal{L}^2(0, T)$$

mit  $T = \frac{2\pi}{\omega}$  gilt

$$a_k = \frac{2}{T} \int_0^T f(x) \cos k\omega x \, dx \quad (k = 0, 1, 2 \dots),$$

$$b_k = \frac{2}{T} \int_0^T f(x) \sin k\omega x \, dx \quad (k = 1, 2, \dots).$$

**3.2** (Formeln für die Koeffizienten der Fourier-Reihe im Fall  $\omega = 1$ ). Für

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) \in \mathcal{L}^2(0, 2\pi)$$

gilt

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos kx \, dx \quad (k = 0, 1, 2, \dots),$$
$$b_k = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin kx \, dx \quad (k = 1, 2, \dots).$$

**3.3 (Tipps und Tricks).** Man beachte dass  $\cos x$  eine gerade und  $\sin x$  eine ungerade Funktion ist. Das hat die folgenden Auswirkungen auf die Entwicklung in die Fourier-Reihe:

- (a)  $f$  gerade, d.h.,  $f(x) = f(-x)$  für alle  $x \in \mathbb{R} \implies$  alle  $b_k$  gleich 0.
- (b)  $f$  ungerade, d.h.,  $f(x) = -f(-x)$  für alle  $x \in \mathbb{R} \implies$  alle  $a_k$  gleich 0.

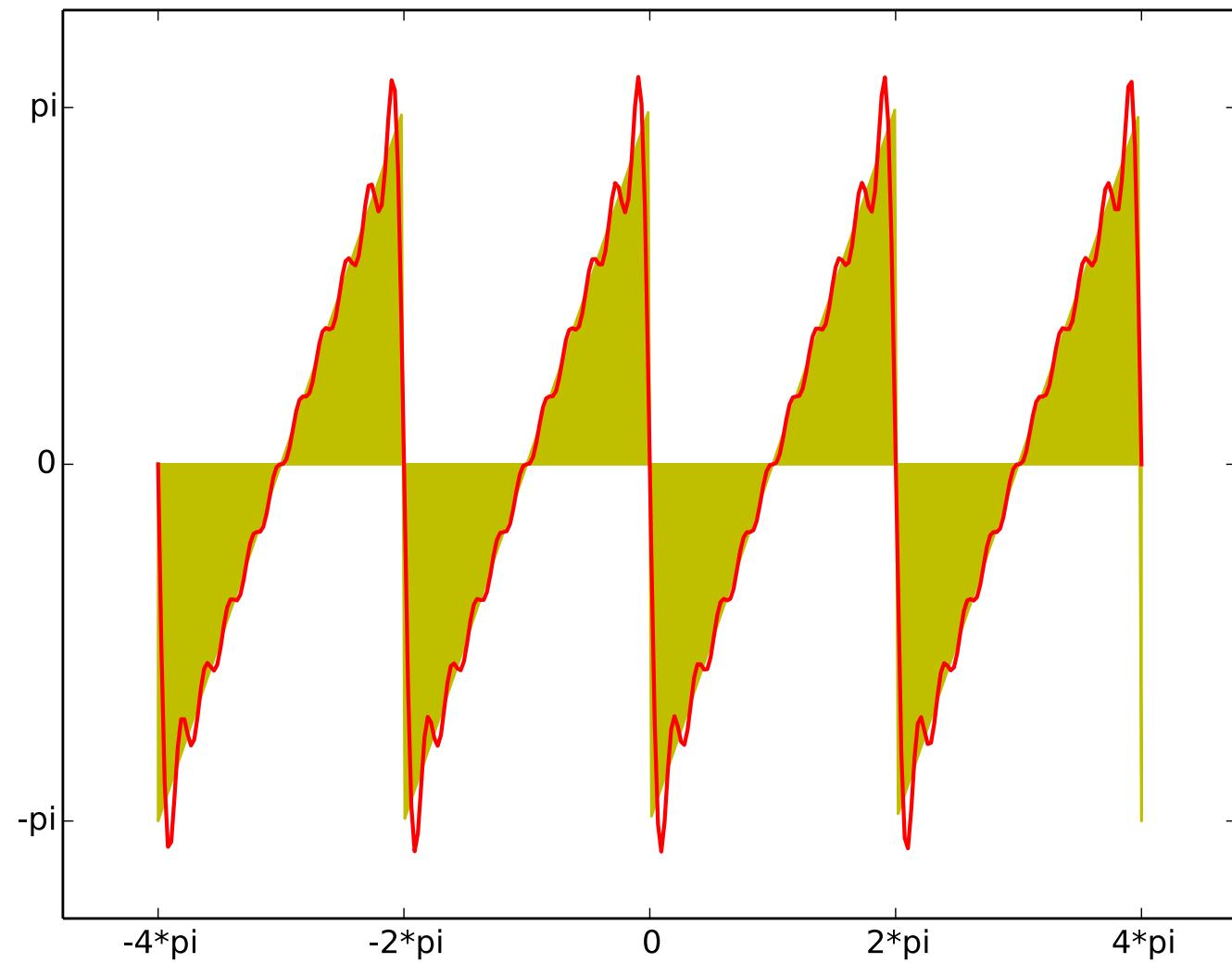
**3.4 Bsp.** Sei  $f(x)$  die Funktion mit  $f(x) = x - \pi$  für  $0 < x < 2\pi$ , die wir  $2\pi$ -periodisch auf  $\mathbb{R}$  erweitern. Da  $f$  ungerade ist, gilt

$$f(x) = \sum_{k=1}^{\infty} b_k \sin kx.$$

Die Koeffizienten:

$$b_k = \frac{1}{\pi} \int_0^{2\pi} (x - \pi) \sin kx \, dx = \frac{1}{\pi} \int_0^{2\pi} x \sin kx \, dx = -\frac{2}{k}$$

Approximation durch  $\sum_{k=1}^{10} b_k \sin kx$ :



**3.5 Aufgabe.** Berechnen Sie die Fourier-Entwicklung von  $f(x) = \text{sign}(\cos(x))$ .

**3.6 Aufgabe.** Berechnen Sie die Fourier-Entwicklung von  $f(x) = \sin x \cos x$ .

**3.7 Aufgabe.** Berechnen Sie die Fourier-Entwicklung von  $f(x) = \cos^4 x$ .

**3.8 Aufgabe.** Nehmen wir an, Sie haben die Fourier-Entwicklung einer Funktion  $f(x) \in \mathcal{L}^2(0, 2\pi)$  ausgerechnet. Wie kann man daraus die Fourier-Entwicklung von  $f(x + \alpha)$  für ein gegebenes  $\alpha \in \mathbb{R}$  bestimmen? Geben Sie die Formeln für die Koeffizienten  $\tilde{a}_0, \tilde{a}_1, \tilde{a}_2, \dots, \tilde{b}_1, \tilde{b}_2, \dots$  der Fourier-Entwicklung von  $f(x + \alpha)$  in Abhängigkeit von den Koeffizienten  $a_0, a_1, a_2, \dots, b_1, b_2, \dots$  der Fourier-Entwicklung von  $f(x)$ . Wie hängen  $\tilde{a}_k^2 + \tilde{b}_k^2$  und  $a_k^2 + b_k^2$  für  $k \in \mathbb{N}$  zusammen?

## 4 Differenzieren von Fourier-Reihen

**4.1 Thm.** Sei  $f$  die  $2\pi$ -periodische Erweiterung einer differenzierbaren Funktion  $g : \mathbb{R} \rightarrow \mathbb{R}$  mit  $g(0) = g(2\pi)$ . Wenn  $f$  und  $f'$  zu  $\mathcal{L}^2(0, 2\pi)$  gehören, dann ergibt sich die Fourier-Entwicklung von  $f'$  durch gliederweise Differenzieren der Fourier-Entwicklung von  $f$ .

**4.2** (Begründung). Wir betrachten die beiden Fourier-Entwicklungen:

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx),$$

$$f'(x) = \frac{a'_0}{2} + \sum_{k=1}^{\infty} (a'_k \cos kx + b'_k \sin kx).$$

Man hat

$$\begin{aligned} a'_k &= \frac{1}{\pi} \int_0^{2\pi} f'(x) \cos kx \, dx && | \text{ Formeln für die Koeffizienten} \\ &= \underbrace{\left[ f(x) \cos kx \right]_{x=0}^{2\pi}}_{=0} - \frac{1}{\pi} \int_0^{2\pi} f(x)(\cos kx)' \, dx && | \text{ Partielle Integration} \\ &= \frac{k}{\pi} \int_0^{2\pi} f(x) \sin kx \, dx \end{aligned}$$

Also ist  $a'_0 = 0$  und  $a'_k = kb_k$  für  $k \in \mathbb{N}$ .

Analog zeigt man auch  $b'_k = -ka_k$  für  $K \in \mathbb{N}$ . Das führt zur Formel:

$$f'(x) = \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)'.$$

**4.3 Bsp.** Wir erweitern  $f(x) = (x - \pi)^2$  periodisch von  $[0, 2\pi]$  auf  $\mathbb{R}$ . Da wir die Funktion  $x - \pi$  auf  $[0, 2\pi)$  bereits in die Fourier-Reihe entwickelt haben, wissen wir, dass man für  $f'(x) = 2(x - \pi)$  die Entwicklung

$$f'(x) = -4 \sum_{k=1}^{\infty} \frac{\sin kx}{k}$$

Nun können wir darauf basierend die Entwicklung

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \sin kx + b_k \cos kx)$$

bestimmen. Das Differenzieren dieser Entwicklung gliederweise gibt uns die Entwicklung von  $f'(x)$ . Daraus lassen sich die Koeffizienten  $a_k$  und  $b_k$  für alle  $k \in \mathbb{N}$  bestimmen. Wir erhalten

$$f(x) = \frac{a_0}{2} + 4 \sum_{k=1}^{\infty} \frac{\cos kx}{k^2}.$$

Zur Berechnung des Koeffizienten  $a_0$  können wir die Standard-Formel

$$a_0 = \frac{1}{\pi} \int_0^{2\pi} f(x) \, dx$$

benutzen. Wir erhalten

$$f(x) = \frac{\pi^2}{3} + 4 \sum_{k=1}^{\infty} \frac{\cos kx}{k^2}.$$

**4.4. Warnung!** Beim Anwenden von Theorem 4.1 sollen die Voraussetzungen beachtet werden!

Theorem 4.1 ist für unstetige  $2\pi$ -periodische Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$  **nicht** anwendbar.

## 5 Pythagoras in $\mathcal{L}^2(0, T)$

**5.1.** Wenn wir zwei Funktionen  $f_1(x), f_2(x) \in \mathcal{L}^2(0, 2\pi)$  haben:

$$f_s(x) = \frac{a_{s,0}}{2} + \sum_{k=1}^{\infty} (a_{s,k} \cos kx + b_{s,k} \sin kx) \in \mathcal{L}^2(0, 2\pi),$$

so gilt

$$\underbrace{\int_0^{2\pi} f_1(x) f_2(x) \, dx}_{\text{Skalarprodukt von Funktionen}} = \langle f_1, f_2 \rangle_{\mathcal{L}^2(0, 2\pi)} = \underbrace{\pi \left( \frac{1}{2} a_{1,0} a_{2,0} + \sum_{k=1}^{\infty} (a_{1,k} a_{2,k} + b_{1,k} b_{2,k}) \right)}_{\text{ein Skalarprodukt von zwei Folgen von Fourier-Koeffizienten}}$$

Im Fall  $f_1 = f_2 = f$  erhalten wir...

**5.2 Thm** (Parseval'sche Gleichung). *Für*

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) \in \mathcal{L}^2(0, 2\pi).$$

*gilt*

$$\|f\|_{\mathcal{L}^2(0,2\pi)}^2 = \int_0^{2\pi} f(x)^2 \, d\,x = \pi \left( \frac{a_0^2}{2} + \sum_{k=1}^{\infty} (a_k^2 + b_k^2) \right).$$

**5.3.** Parseval'sche Gleichung ist der Satz von Pythagoras in unendlich vielen Dimensionen.

**5.4 Bsp.** Wir berechnen  $\sum_{k=1}^{\infty} \frac{1}{k^2}$  mit Hilfe der Parseval'schen Gleichung. Wir wählen als  $f(x)$  die Funktion mit  $f(x) = x - \pi$  für  $0 < x < 2\pi$ , die wir  $2\pi$ -periodisch auf  $\mathbb{R}$  erweitern. Da  $f$  ungerade ist, gilt

$$\int_0^{2\pi} f(x)^2 dx = \pi \sum_{k=1}^{\infty} b_k^2.$$

Die linke Seite ist

$$\int_0^{2\pi} (x - \pi)^2 dx = \left[ \frac{(x - \pi)^3}{3} \right]_{x=0}^{2\pi} = \frac{2\pi^3}{3}.$$

Also gilt

$$\sum_{k=1}^{\infty} b_k^2 = \frac{2\pi^2}{3}.$$

Die Koeffizienten  $b_k$  sind

$$b_k = \frac{1}{\pi} \int_0^{2\pi} (x - \pi) \sin kx dx = \frac{1}{\pi} \int_0^{2\pi} x \sin kx dx = -\frac{2}{k}$$

Das ergibt:

$$\sum_{k=1}^{\infty} \left(\frac{2}{k}\right)^2 = \frac{2\pi^2}{3}$$

und somit

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6}.$$

## **6 Komplexwertige Funktionen**

**6.1.** Oft arbeitet man auch mit komplexwertigen  $T$ -periodischen Funktionen  $f : \mathbb{R} \rightarrow \mathbb{C}$ . In diesem Fall definiert man  $\mathcal{L}^2(0, T)$  entsprechend als einen  $\mathbb{C}$ -Vektorraum mit

$$\langle f, g \rangle_{\mathcal{L}^2(0, T)} = \int_0^T f(x) \overline{g(x)} \, dx$$

und

$$\|f\|_{\mathcal{L}^2(0, T)} := \sqrt{\int_0^T |f(x)|^2 \, dx}.$$

**6.2.** Das System  $(e^{ikx})_{k \in \mathbb{Z}}$  ist die Standardwahl einer “Orthogonalbasis” der komplexen Version von  $\mathcal{L}^2(0, 2\pi)$ . Man definiert die Fourier-Entwicklung dazu als:

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}.$$

Diese Gleichung interpretiert man als den Grenzwert

$$\lim_{n \rightarrow \infty} \left\| f(x) - \sum_{k=-n}^n c_k e^{ikx} \right\|_{\mathcal{L}^2(0, 2\pi)} = 0.$$

Die Formel für die Koeffizienten:

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx.$$

**6.3 Bsp.** Wir berechnen die Fourier-Entwicklung der Funktion

$$f(x) := \begin{cases} 1, & 0 \leq x < \pi \\ 0, & \pi \leq x < 2\pi, \end{cases},$$

die auf die gesamte reelle Achse  $2\pi$ -periodisch erweitert wird, in der “exponentiellen Basis”  $(e^{ikx})_{k \in \mathbb{Z}}$  und anschließend in der “trigonometrischen Basis”. Man hat

$$c_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x) \, dx = \frac{1}{2}.$$

Für  $k \in \mathbb{Z} \setminus \{0\}$  hat man

$$\begin{aligned} c_k &= \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx \\ &= \frac{1}{2\pi} \int_0^\pi e^{-ikx} dx \\ &= \left[ \frac{i}{2\pi k} e^{-ikx} \right]_{x=0}^\pi \\ &= \frac{i}{2\pi k} (e^{-ik\pi} - 1). \\ &= \begin{cases} -\frac{i}{\pi k} & k \text{ ungerade}, \\ 0 & k \text{ gerade} \end{cases} \end{aligned}$$

Das ergibt die Entwicklung

$$f(x) = \frac{1}{2} - \frac{i}{\pi} \sum_{\ell=-\infty}^{\infty} \frac{1}{2\ell+1} e^{i(2\ell+1)x}$$

in der exponentiellen Basis. Um die Entwicklung in der trigonometrischen Basis zu erhalten, werden die Glieder in Paare zerlegt, sodass  $e^{ikx}$  und  $e^{-ikx}$  in ein neues Glied aufgenommen werden, das als Summe von zwei Gliedern der vorigen Entwicklung entsteht.

$$f(x) = \frac{1}{2} - \frac{1}{\pi} \sum_{m=1}^{+\infty} \frac{1}{2m-1} \mathbf{i} (e^{\mathbf{i}(2m-1)x} - e^{-\mathbf{i}(2m-1)x}).$$

Durch die Anwendung der Euler-Formel erhalten wir nun auch die Darstellung

$$f(x) = \frac{1}{2} + \frac{2}{\pi} \sum_{m=1}^{\infty} \frac{\sin(2m-1)x}{2m-1}$$

in der trigonometrischen Basis. (Diese Darstellung kann man natürlich auch direkt erhalten.)

**6.4 (Pythagoras).** Für zwei Funktionen  $f_1, f_2 \in \mathcal{L}^2(0, 2\pi)$  mit

$$f_s(x) = \sum_{k=-\infty}^{\infty} c_{s,k} e^{ikx}$$

gilt

$$\int_0^{2\pi} f_1(x) \overline{f_2(x)} \, dx = \langle f_1, f_2 \rangle_{\mathcal{L}^2(0, 2\pi)} = 2\pi \underbrace{\sum_{k=-\infty}^{\infty} c_{1,k} \overline{c_{2,k}}} \quad .$$

ein Skalarprodukt von zwei komplexwertigen Folgen

Insbesondere hat man im Fall  $f = f_1 = f_2$ :

$$\int_0^{2\pi} |f(x)|^2 \, dx = 2\pi \sum_{k=-\infty}^{\infty} |c_k|^2.$$

**6.5.** Der Raum  $\mathcal{L}^2(0, \pi)$  ist also isometrisch zum komplexen Euklidischen Raum  $\ell^2$  aller Folgen  $c = (c_k)_{k \in \mathbb{Z}}$  mit  $c_k \in \mathbb{C}$  und mit der Eigenschaft

$$\sum_{k=-\infty}^{\infty} |c_k|^2 < \infty.$$

Die Norm in diesem Raum ist:

$$\|c\|_{\ell^2} := \sqrt{\sum_{k=-\infty}^{\infty} |c_k|^2}.$$

Das Skalarprodukt zu dieser Norm ist:

$$\langle c, d \rangle = \sum_{k=-\infty}^{\infty} c_k \overline{d_k}$$

für  $c = (c_k)_{k \in \mathbb{Z}}, d = (d_k)_{k \in \mathbb{Z}} \in \ell^2$ .

**6.6** (Übergang von der “Exponentialbasis” zur “trigonometrischen Basis”).

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx} = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx).$$

Man hat

$$a_0 = 2c_0$$

und, nach der Euler-Formel,

$$c_k \underbrace{(\cos kx + i \sin kx)}_{=e^{ikx}} + c_{-k} \underbrace{(\cos kx - i \sin kx)}_{=e^{-ikx}} = a_k \cos kx + b_k \sin kx$$

für  $k \in \mathbb{N}$ .

Also gilt

$$a_k = c_k + c_{-k}$$

$$b_k = i(c_k - c_{-k})$$

für  $k \in \mathbb{N}$ .

**6.7** (Übergang von “trigonometrischen Basis” zur “Exponentialbasis”).

$$c_0 = \frac{a_0}{2}.$$

Für  $k \in \mathbb{N}$  gilt:

$$c_k = \frac{a_k - \mathbf{i}b_k}{2}$$

$$c_{-k} = \frac{a_k + \mathbf{i}b_k}{2}$$

## 7 Fourierreihen bei Signalverarbeitung

**7.1.** In der Signalverarbeitung ist  $x$  der Zeitpunkt (wird deswegen auch oft als  $t$  bezeichnet).

Der Term

$$h_k(x) := a_k \cos k\omega x + b_k \sin k\omega x$$

bestimmt die Phase und die Amplitude der  $k$ -ten Frequenz. Der Vektor  $(a_k, b_k)$  kann in Polarkoordinaten als

$$a_k = A_k \cos \phi_k$$

$$b_k = A_k \sin \phi_k$$

mit  $A_k \geq 0$  beschrieben werden. Dann gilt

$$h_k(x) := A_k \cos(k\omega x - \phi_k).$$

Der Term  $h_k(x)$  ist der Beitrag der  $k$ -ten Frequenz,  $A_k$  ist die Amplitude und  $\phi_k$  die Phase der  $k$ -ten Frequenz.

## 7.2. Hier noch Kommentare zur Terminologie und Einheiten:

- $x$  ist Zeit. Wir nehmen an,  $x$  ist in Sekunden gegeben.
- $A_k$  ist die Amplitude. Die Einheiten von  $A_k$  hängen von der Natur des Signals ab (elektrisch, mechanisch, akustisch usw.).
- $k\omega$  ist die Kreisfrequenz von  $h_k(x)$ . Wenn man  $k\omega x$  als einen Winkel, und somit als einen Punkt auf dem Einheitskreis interpretiert, dann ist  $k\omega x$  die Geschwindigkeit dieses Punktes in Radianen pro Sekunde). Radianen sind dimensionslos, also ist die Einheit der Kreisfrequenz  $\frac{1}{\text{Sekunde}} = \text{Hz}$ .
- $\frac{2\pi}{k\omega}$  ist die Periode von  $h_k(x)$  in Sekunden. In so viel Zeit mach  $k\omega x$  die volle Runde auf dem Einheitskreis.

- $\frac{k\omega}{2\pi} = \left(\frac{2\pi}{k\omega}\right)^{-1}$  ist die Anzahl der Runden pro Sekunde, also die Frequenz. Die Einheit dazu ist  $\frac{1}{\text{Sekunde}} = \text{Hz}$ .

**7.3.** Bei der Darstellung in der “Exponentialbasis” ergibt sich die  $k$ -te Amplitude und Phase aus den Koeffizienten  $c_k$  und  $c_{-k}$ .

## 8 Die Verwandten der Fourier-Reihe

**8.1** (Diskrete Fourier-Transformation (=DFT)). Analog zu stetigen periodischen Funktionen auf  $\mathbb{R}$  kann man  $n$ -Periodische Funktionen auf  $\mathbb{Z}$  betrachten. Die Diskrete Fourier-Transformation (DFT) ist eine Art Fourier Reihe für solche Funktionen.

**8.2** (Diskrete harmonische Funktionen und ‘diskrete Fourier-Reihen’). Den analogen harmonischen Funktionen  $e^{ikx}$  entsprechen in der Welt der  $n$ -periodischen diskreten Signale die Folgen

$$p_k := \left( e^{\frac{2\pi i k j}{n}} \right)_{j \in \mathbb{Z}}$$

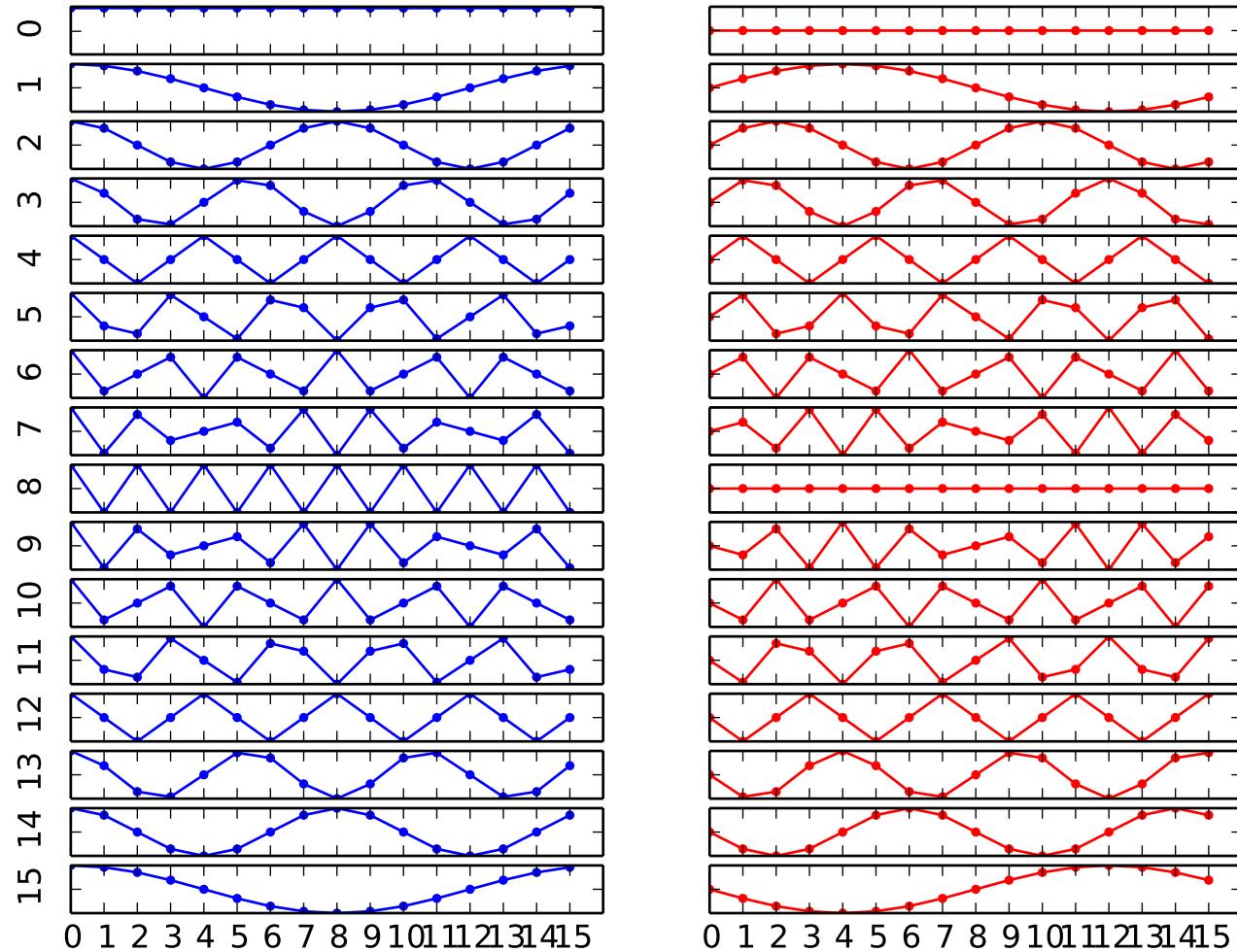
für  $k \in \{0, \dots, n-1\}$ . Ein  $n$ -periodisches Signal  $f = (f_j)_{j \in \mathbb{Z}}$  kann man mit dem Vektor  $(f_j)_{j=0, \dots, n} \in \mathbb{C}^n$  identifizieren, weil es ausreicht die Werte, in einer Periode zu notieren.

Die Zerlegung von  $f = (f_j)_{j=0, \dots, n} \in \mathbb{C}^n$  in harmonische Signale ist somit eine Darstellung der Form

$$f_j = \sum_{k=0}^{n-1} e^{\frac{2\pi i j k}{n}} c_k \tag{VI.2}$$

mit  $j = 0, \dots, n-1$  und  $c = (c_k)_{k=0, \dots, n-1} \in \mathbb{C}^n$ .

**8.3 Bsp.** Hier die Darstellung der komplexen harmonischen Funktionen. Der Realteil der harmonischen Funktionen wird in Blau und der Imaginärteil in Rot dargestellt. Fall  $n = 16$ :



**8.4.** FFT (= Fast Fourier Transform) ist der schnelle Algorithmus zur Berechnung der DFT. In Matlab als die Funktion `fft` verfügbar. Hier ein Beispiel zum “Entrauschen” mit Hilfe der DFT:  
<https://www.mathworks.com/help/matlab/ref/fft.html>

**8.5.** Fourier-Integrale:

$$f(x) = \int_{-\infty}^{\infty} c(\omega) e^{i\omega x} d\omega$$

für  $f : \mathbb{R} \rightarrow \mathbb{C}$  (hier muss  $f$  nicht periodisch sein). Im Gegenteil zu Fourier-Reihen ist hier das Spektrum (die Anzahl der Frequenzen in der Zerlegung) nicht mehr diskret.

**8.6.** Potenz-Reihen:

IM AUFBAU: ein Beispiel.

**8.7.** Multi-Dimensionalen Fourier-Reihen usw.

IM AUFBAU.

**8.8. Tschebyscheff-Polynome:**  $\cos n\phi$  ist ein Polynom in  $\cos \phi$  für alle  $n \in \mathbb{N}_0$ .

Man hat  $\cos n\phi = P_n(\cos \phi)$ . Das Polynom  $P_n$  wird das Tschebyscheff-Polynom genannt. Das Polynom  $P_n$  kann durch eine lineare Rekursion durch  $P_{n-1}$  und  $P_{n-2}$  dargestellt werden ( $P_0 = 1, P_1 = x$ ).

# **Kapitel VII**

## **Differentialrechnung II**

### **1 Exkurs in topologische Räume**

**1.1.** In diesem Kurs geben wir in Bezug auf Konvergenz/Stetigkeit sehr viele ähnliche Definitionen: Grenzwert einer Folge, links- und rechtsseitige Version davon, die Fälle mit  $\pm\infty$ , Grenzwert einer Funktion, Stetigkeit einer Funktion. Solche Definition können im Rahmen der Theorie topologischen Räume einheitlich eingeführt und behandelt werden.

**1.2 Def** (Topologischer Raum). Ein **topologischer Raum** ist eine Menge  $X$ , die mit einem System von Teilmengen von  $X$  ausgestattet ist, die man **die offenen Mengen** des topologischen Raums  $X$  nennt, und welche die folgenden Eigenschaften erfüllen:

- die leere Menge und das gesamte  $X$  sind offene Mengen,
- Vereinigung beliebiger Anzahl offener Mengen ist offen,
- Durchschnitt endlicher Anzahl offener Mengen ist offen.

**1.3 Def.** Eine Familie  $\mathcal{B}$  von Teilmengen von  $X$  erzeugt die Topologie auf  $X$ , in der die offenen Mengen Vereinigungen beliebig vieler Mengen aus  $\mathcal{B}$  sind. Man nennt in diesem Fall  $\mathcal{B}$  **Basis** der jeweiligen Topologie auf  $X$ .

Mit anderen Worten:  $U$  ist offen bzgl. der Basis  $\mathcal{B}$ , wenn für jedes  $x \in U$  ein  $B \in \mathcal{B}$  mit  $x \in B \subseteq U$  existiert.

## 1.4. Das System

$$\{(l, r) : l, r \in \mathbb{R}, l < r\}$$

aller offenen Intervalle ist die Basis der Standardtopologie von  $\mathbb{R}$ . Diese Topologie nennt man auch die **Euklidische Topologie** von  $\mathbb{R}$ .

**1.5 Def.** Sei  $\mathcal{F}$  eine nichtleere Familie von Teilmengen einer Menge  $N$ . Man nennt  $\mathcal{F}$  **Filter** auf  $N$ , wenn Folgendes gilt:

- Ist  $A$  Menge aus  $\mathcal{F}$  und  $B$  eine Teilmenge von  $N$  mit  $A \subseteq B$ , dann ist  $B$  ebenfalls Menge aus  $\mathcal{F}$ .
- Aus  $A, B \in \mathcal{F}$  folgt  $A \cap B \in \mathcal{F}$ .

**1.6 Def** (Konvergenz entlang eines Filters). Sei  $\mathcal{F}$  Filter auf einer Menge  $N$ , sei  $X$  topologischer Raum mit dem System der offenen Mengen  $\mathcal{O}_X$  und man betrachte eine Abbildung  $x : N \rightarrow X$  und ein  $p \in X$ . Wir sagen, dass  $x$  entlang des Filters  $\mathcal{F}$  und bzgl. Topologie  $\mathcal{O}_X$  gegen  $p$  konvergiert, wenn für jedes  $U \in \mathcal{O}_X$  mit  $p \in U$  ein  $A \in \mathcal{F}$  existiert, für welches die Bedingung  $x(n) \in U$  für alle  $n \in A$  erfüllt ist.

**1.7 Def.** Wir nennen eine Teilmenge  $A$  von  $\mathbb{N}$  **kofinit**, wenn  $\mathbb{N} \setminus A$  endlich ist. Mit anderen Worten: eine Teilmenge  $A$  von  $\mathbb{N}$  ist kofinit, wenn nur endlich viele Elemente von  $\mathbb{N}$  nicht zu  $A$  gehören.

**1.8.** Die Menge aller kofiniten Teilmengen von  $\mathbb{N}$  ist ein Filter. Konvergenz von Folgen  $(x_n)_{n \in \mathbb{N}}$  der Elemente aus  $\mathbb{R}$  entlang des Filters der kofiniten Mengen und bzgl. der Standardtopologie auf  $\mathbb{R}$  ist die Konvergenz, die wir am Anfang dieses Kurses mit Hilfe von  $\epsilon > 0$  und  $n_0 \in \mathbb{N}$  eingeführt haben.

**1.9 Def.** Sei  $X$  topologischer Raum mit dem System der offenen Mengen  $\mathcal{O}_X$  und sei  $a \in X$ . Man nennt Teilmenge  $A$  von  $X$  eine  $\mathcal{O}_X$ -**Umgebung** von  $a$ , wenn ein  $U \in \mathcal{O}_X$  mit  $a \in U \subseteq A$  existiert.

**1.10.** Das System aller Umgebungen von  $a \in X$  bzgl. der Topologie  $\mathcal{O}_X$  von  $X$  ist ein Filter.

**1.11 Def.** Sei  $X$  bzw.  $Y$  topologischer Raum mit dem Systemen von offenen Mengen  $\mathcal{O}_X$  bzw.  $\mathcal{O}_Y$ . Seien  $f : X \rightarrow Y$ ,  $a \in X$  und  $y \in Y$ . Man sagt,  $f(x) \rightarrow y$  für  $x \rightarrow a$ , wenn  $f(x)$  entlang des Filters der  $\mathcal{O}_X$ -Umgebungen von  $a$  und bzgl. der Topologie  $\mathcal{O}_Y$  von  $Y$  gegen  $y$  konvergiert.

**1.12 Def** (Stetige Abbildungen topologischer Räume). Sei  $X$  topologischer Raum mit dem System der offenen Mengen  $\mathcal{O}_X$  und  $Y$  topologische Raum mit dem System von offenen Mengen  $\mathcal{O}_Y$ .

Eine Abbildung  $f : X \rightarrow Y$  heißt **stetig** an der Stelle  $a \in X$ , wenn  $f(x) \rightarrow f(a)$  für  $x \rightarrow a$  gilt (bzgl. der Topologien  $\mathcal{O}_X$  und  $\mathcal{O}_Y$  von  $X$  bzw.  $Y$ ).

**1.13.** Ohne Verwendung der Filter kann die Stetigkeit von  $f : X \rightarrow Y$  an der Stelle  $a$  folgendermaßen beschrieben werden:

Für jedes  $V \in \mathcal{O}_Y$  mit  $f(a) \in V$  existiert ein  $U \in \mathcal{O}_X$  mit  $a \in U$  derart, dass für alle  $x \in U$  die Bedingung  $f(x) \in V$  erfüllt ist.

**1.14** (Weitere nützliche Topologien auf  $\mathbb{R}$ ).

- (a) Die Topologie zur Basis  $\{[l, r) : l, r \in \mathbb{R}, l < r\}$  entspricht der rechtseitigen Konvergenz. In dieser Topologie konvergiert die Folge  $(1/n)_{n \in \mathbb{N}}$  gegen 0, die Folge  $(-1/n)_{n \in \mathbb{N}}$  konvergiert aber nicht.
- (b) Die Topologie zur Basis  $\{(l, r] : l, r \in \mathbb{R}, l < r\}$  entspricht der linkseitigen Konvergenz. In dieser Topologie konvergiert  $(-1/n)_{n \in \mathbb{N}}$  gegen 0, die Folge  $(1/n)_{n \in \mathbb{N}}$  konvergiert aber nicht.

## 2 Konvergenz und Stetigkeit im multivariaten Fall

**2.1** (Die Struktur des Raums  $\mathbb{R}^n$ ). Sei  $n \in \mathbb{N}$ .

- (a)  $\mathbb{R}^n$  ist ein  $n$ -dimensionaler Vektorraum über  $\mathbb{R}$ .
- (b) In  $\mathbb{R}^n$  ist Euklidischer Raum, mit dem Standardskalarprodukt

$$\langle x, y \rangle := x^\top y$$

und der zugehörigen Norm  $\|x\| := \sqrt{x^\top x}$ . Der Euklidische Abstand von  $x, y \in \mathbb{R}^n$  ist somit gleich  $\|x - y\|$ .

- (c) Die Norm von  $\mathbb{R}^n$  induziert die sogenannte Euklidische Topologie. Die offene Kugel

$$B(a, \epsilon) := \{x \in \mathbb{R}^n : \|x - a\| < \epsilon\}.$$

mit Zentrum  $a \in \mathbb{R}^n$  vom Radius  $\epsilon > 0$  nennen wir die  **$\epsilon$ -Umgebung** von  $a$ . Eine Menge  $X \subseteq \mathbb{R}^n$  nennt man **offen** (in der Euklidischen Topologie), wenn sie mit jedem  $a \in X$  für ein geeignet gewähltes  $\epsilon > 0$  die  $\epsilon$ -Umgebung von  $a$  als Teilmenge enthält.

**2.2.** Das System  $\{B(a, \epsilon) : a \in \mathbb{R}^n, \epsilon > 0\}$  aller  $\epsilon$ -Umgebungen von  $\mathbb{R}^n$  ist die Basis der so-nannten Euklidischen Topologie auf  $\mathbb{R}^n$ . Dies ist die Standardtopologie von  $\mathbb{R}^n$ .

**2.3 Def.** Für eine Folge  $(v_k)_{k \in \mathbb{N}}$  von Vektoren aus  $\mathbb{R}^n$  und ein  $v \in \mathbb{R}^n$  definiert man

$$v = \lim_{k \rightarrow \infty} v_k$$

Durch  $\lim_{k \rightarrow \infty} \|v_k - v\| = 0$ .

**2.4.** Die Gleichung  $v = \lim_{k \rightarrow \infty} v_k$  für eine Folge von Vektoren gilt genau dann, wenn sie komponentenweise gilt. Wenn man etwa  $n = 2$  hat,  $v = (x, y)$  und  $v_k = (x_k, y_k)$ , dann gilt die Gleichung genau dann wenn  $x = \lim_{k \rightarrow \infty} x_k$  und  $y = \lim_{k \rightarrow \infty} y_k$ .

**2.5 Def. Abschluss** einer Menge  $X \subseteq \mathbb{R}^n$ , **abgeschlossene Mengen**  $X \subseteq \mathbb{R}^n$  und **Häufungspunkte** von Mengen  $X \subseteq \mathbb{R}^n$  genauso wie im Fall  $n = 1$  definiert.

**2.6 Def.** Seien  $n, m \in \mathbb{N}$ . Sei  $f : X \rightarrow \mathbb{R}^m$  mit  $X \subseteq \mathbb{R}^n$  und sei  $a$  Häufungspunkt von  $X$ . Für  $y \in \mathbb{R}^m$  definiert man

$$y = \lim_{x \rightarrow a} f(x).$$

als die folgende Bedingung: Für jedes  $\epsilon > 0$  gib es ein  $\delta > 0$  derart, dass für alle  $x \in X$  mit  $0 < \|x - a\| < \delta$  die Ungleichung  $\|f(x) - y\| < \epsilon$  erfüllt ist.

**2.7.** Die vorige Definition ist wie im Spezialfall  $m = n = 1$ . Im Allgemeinen Fall ersetzt man die Beträge durch die Normen.

**2.8 Def.** Sei  $f : X \rightarrow \mathbb{R}^m$  mit  $X \subseteq \mathbb{R}^n$  und sei  $a \in X$  Häufungspunkt von  $X$ . Wir sagen, dass  $f$  **stetig** in  $a$  ist, wenn

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

erfüllt ist. Wir sagen  $f$  ist stetige Funktion auf  $X$ , wenn  $f$  in jedem Häufungspunkt  $a \in X$  stetig ist.

**2.9.** Für Grenzwert und Stetigkeit von multivariaten Funktionen gelten die gleichen Rechenregeln wie im univariaten Fall.

**2.10 Bsp.** Es gilt

$$\lim_{\substack{x \rightarrow 1 \\ y \rightarrow 0}} \frac{\ln(x + 2e^y)}{\sqrt{x^2 + y^2}} = \frac{\ln(1 + 2e^0)}{\sqrt{1^2 + 0^2}} = \ln 3$$

nach den Rechenregeln, da die Funktion  $f(x, y) = \frac{\ln(x+2e^y)}{\sqrt{x^2+y^2}}$  aus den stetigen Funktionen zusammengesetzt ist (durch die Anwendung von Komposition, Summe, Quotient usw.) und der Grenzwert des Nenners  $\sqrt{x^2 + y^2}$  ungleich 0 ist.

**2.11 Aufgabe.** Nehmen wir an, die Einschränkung von  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$  auf jede Gerade, die den Nullpunkt enthält, sei eine stetige Funktion. Ist  $f$  dann stetig im Nullpunkt? **Nicht unbedingt!**

Die Funktion

$$f(x, y) = \begin{cases} \frac{x^2y}{x^4+3y^4} & \text{wenn } (x, y) \neq (0, 0), \\ 0 & \text{wenn } (x, y) = (0, 0). \end{cases}$$

ist das Gegenbeispiel. Es reicht aus, den Grenzwert

$$\lim_{t \rightarrow 0} f(t \cos \alpha, t \sin \alpha)$$

für alle  $\alpha \in \mathbb{R}$  sowie den Grenzwert

$$\lim_{t \rightarrow 0} f(t, t^2)$$

auszurechnen.

**2.12 Def.** Wir nennen  $X \subseteq \mathbb{R}^n$  **beschränkt**, wenn für ein  $\rho > 0$  die Ungleichung  $\|x\| \leq \rho$  für alle  $x \in X$  erfüllt ist. Wir nennen  $X$  kompakt, wenn  $X$  beschränkt und abgeschlossen ist.

**2.13 Thm** (Der Satz von Weierstraß in Dimension  $n$ ). *Sei  $f : X \rightarrow \mathbb{R}$  stetige Funktion auf einer kompakten Teilmenge  $X$  von  $\mathbb{R}^n$ . Dann erreicht  $f$  auf  $X$  ihr Minimum und Maximum.*

### 3 Partielle Ableitungen

**3.1 Def** (Richtungsableitung). Ableitung in Richtung  $u \in \mathbb{R}^n$  an der Stelle  $x \in \mathbb{R}^n$  der Funktion  $f$  ist als

$$\partial_u f(x) := \lim_{h \rightarrow 0} \frac{f(x + hu) - f(x)}{h}.$$

**3.2.** Die vorige Formel kann auch als die folgende Rechenvorschrift formuliert werden: Man fixiert  $x$  und  $u$ , betrachtet ein variables  $h$ , differenziert  $f(x + hu)$  nach  $h$  und setzt danach  $h = 0$  ein.

**3.3.** In unserer Definition beeinflusst nicht nur die Richtung sondern auch die Länge des Vektors  $u$  den Wert  $\partial_u f(x)$ . Es gilt

$$\begin{aligned}\partial_{\alpha u} f(x) &= \lim_{h \rightarrow 0} \frac{f(x + h\alpha u) - f(x)}{h} \\ &= \alpha \cdot \lim_{h \rightarrow 0} \frac{f(x + h\alpha u) - f(x)}{\alpha h} \\ &= \alpha \cdot \partial_u f(x).\end{aligned}$$

für alle  $\alpha \in \mathbb{R} \setminus \{0\}$ . In manchen Quellen wird der Wert

$$\frac{1}{\|u\|} \partial_u f(x)$$

die Richtungsableitung genannt. Dieser Wert ist nur von der Richtung von  $u$  und nicht von der Länge von  $u$  abhängig:

$$\frac{1}{\|\alpha u\|} \partial_{\alpha u} f(x) = \frac{1}{\alpha \cdot \|u\|} \alpha \cdot \partial_u f(x) = \frac{1}{\|u\|} \partial_u f(x),$$

d.h., keine Änderung von  $\frac{1}{\|u\|}\partial_u f(x)$ , wenn wir  $u$  durch  $\alpha u$ , mit  $\alpha \in \mathbb{R} \setminus \{0\}$ , ersetzen.

**3.4.** Im Folgenden ist  $e_1, \dots, e_n$  die Standardbasis von  $\mathbb{R}^n$ : die  $i$ -te Komponenten von  $e_i$  ist gleich 1 und alle anderen gleich 0.

**3.5 Def** (Partielle Ableitungen). **Partielle Ableitungen** sind die Ableitungen in Koordinatenrichtungen  $e_1, \dots, e_n$

$$\frac{\partial}{\partial x_i} f(x) = \lim_{h \rightarrow 0} \frac{f(x + he_i) - f(x)}{h}.$$

**3.6 Bsp.** Wir berechnen die partiellen Ableitungen von  $f(x, y) = x \sin(x^2 - y^2)$ . Beim Ableiten nach  $x$ , sehen wir  $y$  als fest und  $x$  als variabel und benutzen dabei die Ableitungsregeln, die wir aus dem univariaten Fall kennen.

$$\begin{aligned}\frac{\partial f}{\partial x} &= \left( \frac{\partial}{\partial x} x \right) \sin(x^2 - y^2) + x \left( \frac{\partial}{\partial x} \sin(x^2 - y^2) \right) && | \text{ Produktregel} \\ &= \sin(x^2 - y^2) + x \cos(x^2 - y^2) \cdot (2x) && | \text{ Kettenregel} \\ &= \sin(x^2 - y^2) + 2x^2 \cos(x^2 - y^2).\end{aligned}$$

Beim Ableiten nach  $y$ , sehen  $x$  als fest und  $y$  als variable und gehen analog vor:

$$\frac{\partial f}{\partial y} = -2xy \cos(x^2 - y^2).$$

**3.7 Def.** Die **Ableitung** einer Abbildung  $f : X \rightarrow \mathbb{R}^m$  auf  $X \subseteq \mathbb{R}^n$  an der Stelle  $x \in X$  ist die Matrix  $A \in \mathbb{R}^{n \times n}$ , für welche die Entwicklung

$$f(x + h) = f(x) + A \cdot h + R(h),$$

existiert, sodass der Restterm  $R(h)$  die Bedingung

$$\lim_{h \rightarrow 0} \frac{R(h)}{\|h\|} = 0$$

erfüllt. Man sagt, dass  $f$  in  $x$  **differenzierbar** ist, wenn  $f$  in  $x$  eine Ableitung besitzt.

**3.8 Aufgabe.** Bestimmen Sie für eine beliebige affine Abbildung  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , d.h., eine Abbildung der Form  $F(x) = Ax + b$  mit  $A \in \mathbb{R}^{m \times n}$  und  $b \in \mathbb{R}^m$ , die Ableitung von  $F(x)$ .

**3.9.**

- (a) Laut der vorigen Definition ist  $A \in \mathbb{R}^{m \times n}$  als die Matrix mit der Eigenschaft

$$\lim_{h \rightarrow \infty} \frac{f(x + h) - f(x) - A \cdot h}{\|h\|} = 0$$

definiert. Für gegebene  $f$  und  $x$  gibt es höchstens eine solche Matrix (die Ableitung ist eindeutig).

- (b) In manchen Quellen benutzt man für die Ableitung die gleiche Bezeichnung wie im univariaten Fall:  $f'(x)$ .

**3.10 Thm.** Ist  $f : X \rightarrow \mathbb{R}^m$  mit

$$f(x) = \begin{pmatrix} f_1(x) \\ \vdots \\ f_m(x) \end{pmatrix}$$

differenzierbar in  $x \in X$ , dann ist

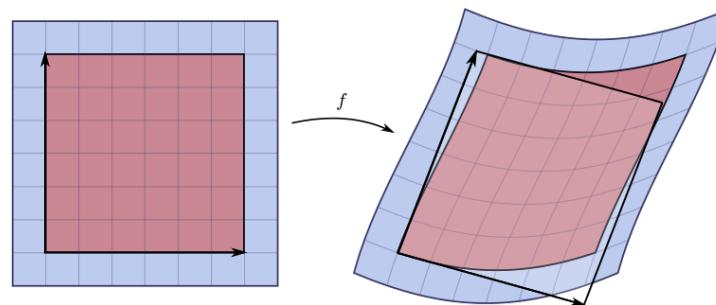
$$J_f(x) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \cdots & \frac{\partial f_1}{\partial x_n}(x) \\ & \vdots & \\ \frac{\partial f_m}{\partial x_1}(x) & \cdots & \frac{\partial f_m}{\partial x_n}(x) \end{pmatrix} \quad (\text{VII.1})$$

die Ableitung von  $f$  an der Stelle  $x$ .

**3.11.**

- (a) Die Matrix  $J_f(x)$  in (VII.1) nennt man die **Jacobi-Matrix**. Das Theorem besagt: die Ableitung von  $f$  (falls vorhanden) ist die Jacobi-Matrix. Theoretisch ist aber eine Situation möglich, dass  $f$  an der Stelle  $x$  die Jacobi-Matrix aber keine Ableitung besitzt.
- (b) Andere übliche Schreibweisen für die Jacobi-Matrix  $J_f(x)$  sind  $(Df)(x)$ ,  $\frac{\partial f}{\partial x}(x)$  und  $\frac{\partial(f_1, \dots, f_m)}{\partial(x_1, \dots, x_n)}(x)$

### 3.12. Geometrische Bedeutung:



Quelle: [https://en.wikipedia.org/wiki/Jacobian\\_matrix\\_and\\_determinant](https://en.wikipedia.org/wiki/Jacobian_matrix_and_determinant)

Die Jacobi-Matrix definiert eine lineare Abbildung, die  $e_i \in \mathbb{R}^n$  auf  $\frac{\partial f}{\partial x_i} \in \mathbb{R}^m$  abbildet. Diese lineare Abbildung ist eine Approximation der Abbildung  $h \mapsto f(x+h) - f(x)$  für kleine  $h \in \mathbb{R}^n$ .

**3.13.** Im Spezialfall  $m = 1$  erhält man

$$J_f(x) := \left( \frac{\partial f}{\partial x_1}(x), \dots, \frac{\partial f}{\partial x_n}(x) \right)$$

als die Ableitung von  $f$ .

Der Vektor

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}$$

ist der **Gradient** von  $f$  (gesprochen: Nabla  $f$ ) an der Stelle  $x$ . Eine weitere verbreite Bezeichnung für den Gradienten ist  $\text{grad } f(x)$ .

**3.14.** Der Fall  $m = 1$  von Theorem 3.10 kann nun so formuliert werden:

Ist  $f : X \rightarrow \mathbb{R}$  auf  $X \subseteq \mathbb{R}^n$  differenzierbar in  $x \in X$ , so gilt

$$f(x + h) = f(x) + \langle \nabla f(x), h \rangle + R(h)$$

mit  $\lim_{h \rightarrow 0} \frac{R(h)}{\|h\|} = 0$ .

**3.15 Bsp.** Wir betrachten die quadratische Form  $q_A(x) = \langle Ax, x \rangle$  für eine symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$ . Der Gradient von  $q_A(x)$  kann aus der vorigen Interpretation hergeleitet werden:

$$q_A(x + h) = \langle A(x + h), x + h \rangle = q_A(x) + 2 \langle Ax, h \rangle + q_A(h)$$

Hierbei ist  $q_A(h)$  der Restterm mit  $\lim_{h \rightarrow 0} \frac{q_A(h)}{\|h\|} = 0$ . Somit ist

$$\nabla q_A(x) = 2Ax$$

**3.16 Aufgabe.**

- (a) Zeigen Sie, dass  $\nabla\|x\| = \frac{x}{\|x\|}$  für alle  $x \in \mathbb{R}^n \setminus \{0\}$  erfüllt ist.
- (b) Zeigen Sie, dass  $\|x\|$  im Punkt  $x = 0$  nicht differenzierbar ist.
- (c) Sei  $f$  eine Funktion auf einer Teilmenge von  $\mathbb{R}^m$  und seien  $A \in \mathbb{R}^{m \times n}$  und  $b \in \mathbb{R}^m$ . Berechnen Sie den Gradienten von  $f(Ax + b)$  in Abhängigkeit vom Gradienten von  $f$ .

**3.17** (Geometrische Bedeutung des Gradienten einer differenzierbaren Funktion). Wenn  $f : X \rightarrow \mathbb{R}$  mit  $X \subseteq \mathbb{R}^n$  stetig differenzierbar ist, so gilt Folgendes:

- (a) Der Graph der affinen Funktion  $f(a) + \langle \nabla f(a), x - a \rangle$  ist die Tangentialebene zum Graphen der Funktion  $f$  an der Stelle  $(a, f(a))$ .
- (b) Der Vektor  $\nabla f(a)$  ist orthogonal zur  $(n - 1)$ -dimensionalen Niveau-Fläche

$$F := \{x \in X : f(x) = f(a)\}.$$

an der Stelle  $a$ . Die **Niveau-Fläche** von  $f$  zum Wert  $h$  ist die Menge

$$\{x \in X : f(x) = h\}.$$

Sie wird auch die **Niveau-Menge** und, für  $n = 2$ , die **Niveau-Linie** genannt.

**3.18.** Sei  $n \geq 2$  und  $X \subseteq \mathbb{R}^n$ . Auch wenn  $f : X \rightarrow \mathbb{R}$  in einem Punkt alle partiellen Ableitungen besitzt, muss  $f$  in diesem Punkt nicht differenzierbar und nicht mal stetig sein. Wir betrachten die folgende Funktion:

$$f(x_1, x_2) = \begin{cases} 0, & \text{if } x_1 = 0 \text{ oder } x_2 = 0, \\ 1, & \text{sonst.} \end{cases}$$

Es gilt offensichtlich  $\frac{\partial}{\partial x_1} f(0, 0) = \frac{\partial}{\partial x_2} f(0, 0) = 0$ , aber  $f$  ist nicht stetig in  $(0, 0)$ .

Diese Diskrepanz lässt sich folgendermaßen erklären. Die partiellen Ableitungen definiert man durch die Änderung des Eingabevektors  $x = (x_1, \dots, x_n)$  in Koordinatenrichtungen  $e_1, \dots, e_n$ . Die Ableitung von  $f$  an der Stelle  $x$  ist dagegen bzgl. einer beliebigen kleinen Änderung  $x + h$  von  $x$  definiert.

**3.19 Def** (Höhere partielle Ableitungen). Die  **$k$ -ten partiellen Ableitungen** von  $f : X \rightarrow \mathbb{R}$  auf  $X \subseteq \mathbb{R}^n$  sind

$$\frac{\partial^k}{\partial x_{i_1} \cdots \partial x_{i_k}} f := \frac{\partial}{\partial x_{i_1}} \frac{\partial}{\partial x_{i_2}} \cdots \frac{\partial}{\partial x_{i_k}} f.$$

mit  $i_1, \dots, i_k \in \{1, \dots, n\}$ . Man nennt  $f$   **$k$ -mal partiell differenzierbar bzw. stetig partiell differenzierbar**, wenn  $f$  alle  $k$ -ten partiellen Ableitungen existieren bzw. existieren und stetige Funktionen sind.

**3.20 Thm** (Schwarz). Sei  $f : X \rightarrow \mathbb{R}$  mit  $X \subseteq \mathbb{R}^n$ . Wenn  $f$   $k$ -mal stetig partiell differenzierbar ist, dann ist die  $k$ -te Ableitung

$$\frac{\partial^k}{\partial x_{i_1} \cdots \partial x_{i_k}} f$$

unabhängig von der Reihenfolge von  $x_{i_1}, \dots, x_{i_k}$ . Insbesondere gilt für  $k = 2$

$$\frac{\partial^2}{\partial x_i \partial x_j} f = \frac{\partial^2}{\partial x_j \partial x_i} f$$

für alle  $i, j \in \{1, \dots, n\}$ .

**3.21 Aufgabe.** Vergewissern Sie sich, dass für  $f(x, y) = x \sin(x^2 - y^2)$  die partiellen Ableitungen

$$\frac{\partial^2 f}{\partial x \partial y}$$

und

$$\frac{\partial^2 f}{\partial y \partial x}$$

übereinstimmen. Berechnen Sie alle partiellen Ableitungen von  $f$  der Ordnung zwei.

**3.22 Thm** (Hinreichende und Notwendige Bedingung für Differenzierbarkeit). Für  $f : X \rightarrow \mathbb{R}$  mit  $X \subseteq \mathbb{R}^n$  gelten die Implikationen:

*f ist stetig partiell differenzierbar*



*f ist differenzierbar*



*f ist stetig und ist partiell differenzierbar*

**3.23 Thm.** Sei  $f : X \rightarrow \mathbb{R}$  konvexe Funktion auf  $X \subseteq \mathbb{R}^n$  und sei  $a$  innerer Punkt von  $X$ . Dann sind die folgenden Bedingungen äquivalent:

- (i)  $f$  ist partiell differenzierbar in  $a$ .
- (ii)  $f$  ist differenzierbar in  $a$ .

**3.24 Aufgabe.** Da  $f = x \sin(x^2 - y^2)$  stetig partiell differenzierbar ist, hat man für jede Wahl von  $(x^*, y^*) \in \mathbb{R}^2$  eine Darstellung

$$f(x, y) = f(x^*, y^*) + \frac{\partial f}{\partial x}(x^*, y^*)(x - x^*) + \frac{\partial f}{\partial y}(x^*, y^*)(y - y^*) + R(x, y)$$

mit dem Restterm

$$R(x, y) := f(x, y) - \underbrace{\left( f(x^*, y^*) + \frac{\partial f}{\partial x}(x^*, y^*)(x - x^*) + \frac{\partial f}{\partial y}(x^*, y^*)(y - y^*) \right)}_{\text{lineare Approximation von } f \text{ am Punkt } (x^*, y^*)},$$

der die Bedingung

$$\lim_{\substack{x \rightarrow x^* \\ y \rightarrow y^*}} \frac{R(x, y)}{\sqrt{(x - x^*)^2 + (y - y^*)^2}} = 0$$

erfüllt. Bestimmen Sie die lineare Approximation von  $f$  für  $x^* = \sqrt{\pi/2}$  und  $y^* = \sqrt{\pi}$ .

## 4 Kettenregel

**4.1 Thm (Kettenregel).** Für  $f : Y \rightarrow \mathbb{R}^k$  und  $g : X \rightarrow Y$  mit  $X \subseteq \mathbb{R}^n$  und  $Y \subseteq \mathbb{R}^m$  gilt

$$J_{f \circ g}(x) = J_f(g(x)) \cdot J_g(x)$$

wenn  $f$  und  $g$  beide differenzierbar sind.

**4.2** (Informelle Beschreibung der Kettenregel). Nehmen wir an, dass wir eine Größe  $f$  in Abhängigkeit von zwei Größen  $u$  und  $v$  dargestellt haben, wobei  $u$  sowie  $v$  in Abhängigkeit von  $x$  und  $y$  dargestellt sind. Wenn wir  $u$  und  $v$  als Variablen auffassen und dabei  $x$  und  $y$  vergessen, ist  $f$  Funktion von  $u$  und  $v$ . Wenn wir aber  $u$  und  $v$  als Funktionen von  $x$  und  $y$  interpretieren, dann ist  $f$  Funktion von  $x$  und  $y$ . Man kann also  $f$  nach  $u$  und  $v$ , als Funktion in  $(u, v)$  ableiten, aber auch nach  $x$  und  $y$  als Funktion von  $(x, y)$  ableiten. Die Ableitungen nach  $(x, y)$  und nach  $(u, v)$  werden durch die Kettenregel miteinander verlinkt:

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial u} \cdot \frac{\partial u}{\partial x} + \frac{\partial f}{\partial v} \cdot \frac{\partial v}{\partial x}$$

$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial u} \cdot \frac{\partial u}{\partial y} + \frac{\partial f}{\partial v} \cdot \frac{\partial v}{\partial y}$$

In dieser Formel interpretiert man  $f$  als Funktion von  $u$  und  $v$  beim Berechnen von  $\frac{\partial f}{\partial u}$  und  $\frac{\partial f}{\partial v}$ .

und als Funktion in  $x$  und  $y$  beim Berechnen von  $\frac{\partial f}{\partial x}$  und  $\frac{\partial f}{\partial y}$ . Beim Berechnen der partiellen Ableitungen von  $u$  und  $v$  nach  $x$  sowie  $y$ , werden  $u$  und  $v$  als Funktionen in  $x$  und  $y$  aufgefasst.

Diese zwei Gleichungen werden als Gleichungen von Funktionen in Variablen  $x$  und  $y$  aufgefasst.

**4.3.** Ganz analog kann man auch die Kettenregel für allgemeine multivariate Funktionen formulieren. Sei  $f = f(u_1, \dots, u_m)$  mit  $u_i = u_i(x_1, \dots, x_n)$ : das heißt,  $f$  hängt von  $m$  Größen  $u_1, \dots, u_m$  ab, welche von  $n$  Größen  $x_1, \dots, x_n$  abhängig sind. Es gilt:

$$\frac{\partial f}{\partial x_j} = \sum_{i=1}^m \frac{\partial f}{\partial u_i} \cdot \frac{\partial u_i}{\partial x_j}.$$

Diese Formel ist rein formal nicht ganz korrekt, da man  $u_1, \dots, u_m$  in einem Ausdruck als Variablen und als Funktionen in  $x_1, \dots, x_n$  auffasst. Die Formel kann man aber mit etwas Sachverstand korrekt interpretieren. So ein Gebrauch vom mathematischen Formalismus nennt man *abuse of notation* im Englischen. Das ist so etwas wie sinnvoller Notationsmissbrauch: man benutzt Formalismus zwar etwas zweckentfremdet, aber in einer praktischen und berechtigten Weise.

**4.4 Bsp** (Ableitungen nach den polaren Koordinaten). Wir betrachten die Polarkoordinaten  $x = r \cos \phi, y = r \sin \phi$ . Wie kann die Ableitung nach  $\phi$  in den kartesischen Koordinaten  $(x, y)$  dargestellt werden?

$$\begin{aligned}\frac{\partial f}{\partial \phi} &= \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial \phi} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial \phi} = -r \sin \phi \frac{\partial f}{\partial x} + r \cos \phi \frac{\partial f}{\partial y} \\ &= x \frac{\partial f}{\partial y} - y \frac{\partial f}{\partial x}.\end{aligned}$$

Das zeigt

$$\frac{\partial f}{\partial \phi} = x \frac{\partial f}{\partial y} - y \frac{\partial f}{\partial x}$$

Die Ableitung nach  $\phi$  ist das Skalarprodukt des Gradienten von  $f$  mit der Drehung von  $(x, y)$  um den Winkel  $\pi/2$ .

Was ist mit der Ableitung nach  $r$ ?

$$\begin{aligned}\frac{\partial f}{\partial r} &= \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial r} + \frac{\partial f}{\partial y} \cdot \frac{\partial y}{\partial r} \\ &= \cos \phi \frac{\partial f}{\partial x} + \sin \phi \frac{\partial f}{\partial y} \\ &= \frac{1}{r} \left( x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} \right) \\ &= \frac{1}{\sqrt{x^2 + y^2}} \left( x \frac{\partial f}{\partial x} + y \frac{\partial f}{\partial y} \right).\end{aligned}$$

Der Vektor  $\frac{1}{\sqrt{x^2+y^2}}(x, y)$  hat Länge 1 und die gleiche Richtung wie der Vektor  $(x, y)$ . Die Ableitung nach  $r$  ist das Skalarprodukt des Gradienten mit dem Vektor  $\frac{1}{\sqrt{x^2+y^2}}(x, y)$ .

**4.5 Aufgabe.** Setzen Sie das vorige Beispiele fort und bestimmen Sie Darstellungen von  $\frac{\partial^2 f}{\partial \phi^2}$  und  $\frac{\partial^2 f}{\partial r^2}$  in kartesischen Koordinaten  $(x, y)$ .

**4.6** (Zusammenhang der Richtungsableitungen des Gradienten). Für  $f : X \rightarrow \mathbb{R}$  auf  $X \subseteq \mathbb{R}^n$  gilt

$$\begin{aligned}\partial_u f(x) &= \frac{\partial}{\partial t} f(x + tu) \Big|_{t=0} \\ &= J_f(x + tu) \frac{\partial x + tu}{\partial t} \Big|_{t=0} \\ &= \nabla f(x + t)^\top u \Big|_{t=0} \\ &= \langle f(x + tu), u \rangle \Big|_{t=0} \\ &= \langle f(x), u \rangle.\end{aligned}$$

Fazit:

$$\partial_u f(x) = \langle f(x), u \rangle.$$

**4.7 Aufgabe.** Benutzen Sie die Ungleichung von Cauchy-Schwarz um zu zeigen, dass die Richtung von  $\nabla f(x)$  die Richtung des größten Anstiegs von  $x$  an der Stelle  $f$  ist.

**4.8** (Wellengleichung in der Dimension eins). Seien  $c > 0$  die Geschwindigkeit,  $t$  die Zeitvariable und  $x$  die Raumvariable . Die Wellengleichung in der Dimension eins ist die Gleichung

$$\frac{\partial^2 u}{\partial t^2} = c^2 \frac{\partial^2 u}{\partial x^2}$$

für eine unbekannte Funktion  $u(t, x)$ .

Um diese Gleichung zu lösen, schreiben wir sie zuerst in einer Operator-Form um:

$$\left( c^2 \left( \frac{\partial}{\partial x} \right)^2 - \left( \frac{\partial}{\partial t} \right)^2 \right) u = 0. \quad (\text{VII.2})$$

Das soll heißen: Der Operator “ $c^2$  mal das zweifache Ableiten nach  $x$  minus das zweifache Ableiten nach  $t$ ” macht aus  $u$  die Nullfunktion. Der Begriff Operator bedeutet in unseren Kontext “eine Operation, die aus Funktionen (neue) Funktionen erzeugt”.

Nun können wir den Operator  $c^2 \left( \frac{\partial}{\partial x} \right)^2 - \left( \frac{\partial}{\partial t} \right)^2$  aus (VII.2) mit Hilfe der dritten binomischen

Formel als die Anwendung von zwei Operatoren umschreiben:

$$c^2 \left( \frac{\partial}{\partial x} \right)^2 - \left( \frac{\partial}{\partial t} \right)^2 = \left( c \frac{\partial}{\partial x} - \frac{\partial}{\partial t} \right) \left( c \frac{\partial}{\partial x} + \frac{\partial}{\partial t} \right)$$

Das ergibt:

$$\left( c \frac{\partial}{\partial x} - \frac{\partial}{\partial t} \right) \left( c \frac{\partial}{\partial x} + \frac{\partial}{\partial t} \right) u = 0. \quad (\text{VII.3})$$

Die Bedeutung von (VII.3) ist wie folgt: wenn wir zu  $u$  den Operator  $c \frac{\partial}{\partial x} + \frac{\partial}{\partial t}$  anwenden, und dann zum Ergebnis den Operator  $c \frac{\partial}{\partial x} - \frac{\partial}{\partial t}$  anwenden, kriegen wir die Nullfunktion.

Es bietet sich nun die folgende Substitution an, welche die Variablen  $(x, t)$  durch die Variablen  $(\xi, \eta)$  ersetzt:

$$\xi = x - ct,$$

$$\eta = x + ct.$$

Denn, nach der Substitutionsregel, die wir in der Operator-Form hinschreiben, erhalten wir

$$\frac{\partial}{\partial x} = \frac{\partial \xi}{\partial x} \cdot \frac{\partial}{\partial \xi} + \frac{\partial \eta}{\partial x} \cdot \frac{\partial}{\partial \eta} = \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta}$$

$$\frac{\partial}{\partial t} = \frac{\partial \xi}{\partial t} \cdot \frac{\partial}{\partial \xi} + \frac{\partial \eta}{\partial t} \cdot \frac{\partial}{\partial \eta} = -c \frac{\partial}{\partial \xi} + c \frac{\partial}{\partial \eta}$$

Die beiden Operatoren aus (VII.3) können in den neuen Variablen  $(\xi, t)$  dargestellt werden:

$$c \frac{\partial}{\partial x} - \frac{\partial}{\partial t} = c \left( \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta} \right) - \left( -c \frac{\partial}{\partial \xi} + c \frac{\partial}{\partial \eta} \right) = 2c \frac{\partial}{\partial \xi}$$

$$c \frac{\partial}{\partial x} + \frac{\partial}{\partial t} = c \left( \frac{\partial}{\partial \xi} + \frac{\partial}{\partial \eta} \right) + \left( -c \frac{\partial}{\partial \xi} + c \frac{\partial}{\partial \eta} \right) = 2c \frac{\partial}{\partial \eta}$$

Die Gleichung (VII.3) wird also in den neuen Variablen als

$$\frac{\partial}{\partial \xi} \frac{\partial}{\partial \eta} u = 0$$

dargestellt. Die vorige Gleichung bedeutet, dass  $\frac{\partial}{\partial \eta} u$  als Funktion von  $\xi$  und  $\eta$  von  $\xi$  unabhängig ist:

$$\frac{\partial}{\partial \eta} u = f(\eta)$$

Sei  $F$  Stammfunktion von  $f$ . Integration der vorigen Gleichung nach  $\eta$  ergibt "F( $\xi$ )+Konstante". Wir haben aber für jede Wahl von  $\eta$  potenziell eine andere Konstante. Wir bezeichnen also die "Konstante" als  $G(\eta)$ . Wir erhalten somit die Darstellung

$$u = F(\xi) + G(\eta) = F(x - ct) + G(x + ct).$$

Die Funktion  $u(x, t)$  setzt sich aus zwei Wellen zusammen: die Welle  $F(x - ct)$ , die sich mit der Geschwindigkeit  $c$  nach rechts bewegt und die Welle  $G(x + ct)$ , die sich mit der Geschwindigkeit

$c$  nach links bewegt.

**4.9** (Einhüllende Kurve). Betrachten wir eine parametrische Schar von Kurven

$$K_s := \{(x, y) \in \mathbb{R}^2 : f(s, x, y) = 0\},$$

die implizit durch eine Gleichung  $f(s, x, y) = 0$  mit einem Parameter  $s \in \mathbb{R}$  gegeben sind. Eine **einhüllende Kurve** zu dieser Schar ist eine Kurve  $K$ , die in jedem Punkt durch eine der Kurven der Schar  $(K_s)_{s \in \mathbb{R}}$  berührt wird (genaue Definition lassen wir weg).

Die Gleichung der einhüllenden Kurve kann mit Hilfe von partiellen Ableitungen bestimmt werden. Wir betrachten einen Schnittpunkt von zwei Kurven  $K_s$  und  $K_{s+\Delta s}$  der Schar mit  $\Delta s \neq 0$ . Für den Schnittpunkt  $(x, y)$  gilt

$$f(s + \Delta s, x, y) = f(s, x, y) = 0.$$

Diese beiden Gleichungen kann man als

$$\begin{aligned} f(s, x, y) &= 0, \\ \frac{f(s + \Delta s, x, y) - f(s, x, y)}{\Delta s} &= 0 \end{aligned}$$

umformulieren. Wenn wir im vorigen System für  $(x, y)$  den Wert  $\Delta s$  gegen 0 schicken, so geht die Lösung  $(x, y)$  dieses System gegen einen Punkt in der einhüllenden Kurve. Die Lösung erfüllt die Gleichungen

$$\begin{aligned} f(s, x, y) &= 0 \\ \frac{\partial}{\partial s} f(s, x, y) &= 0. \end{aligned}$$

**4.10** (Einhüllende Wurfparabel). Aus dem Punkt  $x = 0, y = 0$  kann Objekt mit Geschwindigkeit  $v$  in jede Richtung herausgeworfen werden. Welche Form hat der Bereich alle Punkte, die durch einen Wurf getroffen werden können?

Die Flugbahn in Abhängigkeit von dem Wurfwinkel  $\alpha$  zur horizontalen Richtung  $(1, 0)$  und der Zeit  $t$  ist durch die Gleichungen

$$x = tv \cos \alpha$$

$$y = tv \sin \alpha - \frac{g}{2}t^2,$$

beschrieben, wobei hier  $g > 0$  die Fallbeschleunigung ist. Wir können  $t$  eliminieren und erhalten die Formel

$$y = x \tan \alpha - \frac{g}{2v^2 \cos^2 \alpha} x^2,$$

die uns  $y$  in Abhängigkeit von  $x$  für ein gegebenes  $\alpha$  beschreibt. Um Trigonometrie zu vermeiden führen wir an der Stelle von  $\alpha$  einen anderen Parameter ein:  $s = \tan \alpha$ . So erhalten wir die

parametrische Schar von Parabeln

$$y = sx - \frac{g}{2v^2}(s^2 + 1)x^2$$

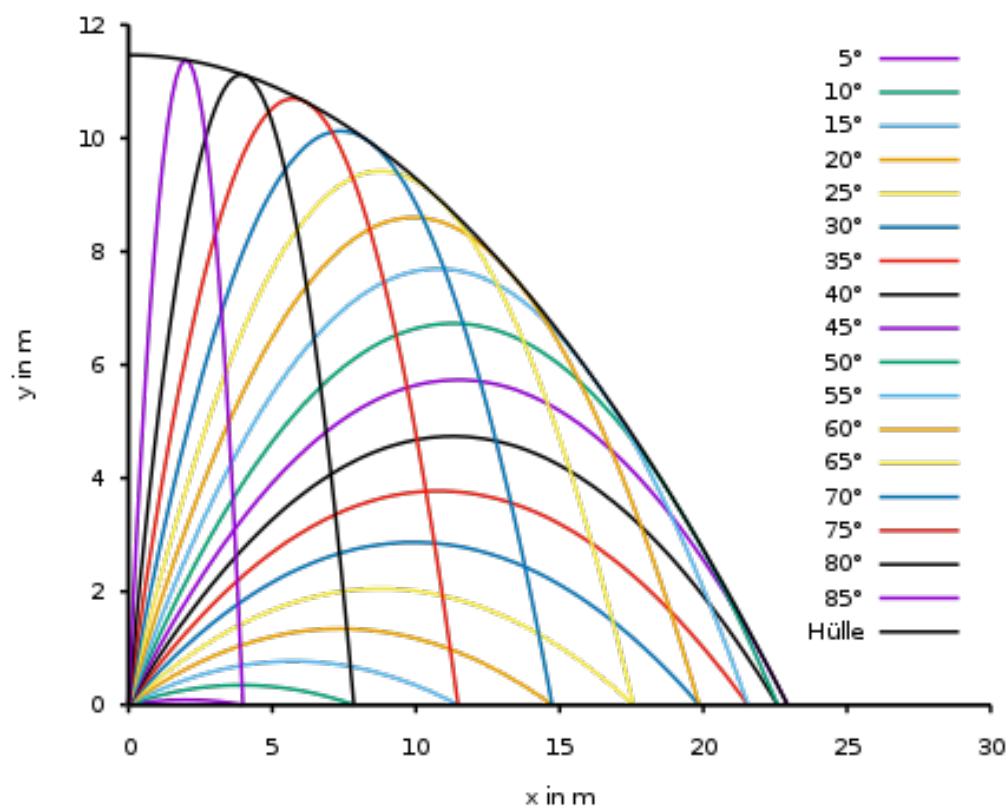
mit einem Parameter  $s \in \mathbb{R}$ . Der Bereich, der durch die Parabeln unserer Schar abgedeckt ist, ist durch die einhüllende Kurve dieser Schar beschränkt. Das Eliminieren von  $s$  im System

$$\begin{aligned} y &= sx - \frac{g}{2v^2}(s^2 + 1)x^2 \\ \frac{\partial}{\partial s}y &= \frac{\partial}{\partial s} \left( sx - \frac{g}{2v^2}(s^2 + 1)x^2 \right) \end{aligned}$$

gibt uns die Gleichung

$$y = \frac{v^2}{2g} - \frac{g}{2v^2}x^2$$

für die einhüllende Kurve.



Quelle: <https://de.wikipedia.org/wiki/Wurfparabel>

**4.11 Aufgabe.** Bestimmen Sie die einhüllende der Wurfparabeln elementar mit der Verwendung der  $pq$ -Formel.

## **5 Totales Differential und implizite Funktionen**

**5.1.** Beispiel im Fall von zwei Variablen. Das **totale Differential** von  $F = F(x, y)$  ist der Ausdruck

$$dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy.$$

Definition im Fall von mehr als zwei Variablen analog. Für jede Wahl von  $(x, y)$  ist  $dF$  eine lineare Funktion in  $(dx, dy)$ . Bedeutung

$$\Delta F \approx \frac{\partial F}{\partial x} \Delta x + \frac{\partial F}{\partial y} \Delta y.$$

Zweite Bedeutung, das Ableitung nach einer anonymen Variablen. Wenn wir für  $x$  und  $y$  Funktionen  $x = x(t)$  und  $y = y(t)$  in  $t$  einsetzen, ist  $F(x(t), y(t))$  ebenfalls Funktion in  $t$ . Nach der Kettenregel gilt:

$$\frac{dF}{dt} = \frac{\partial F}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial F}{\partial y} \cdot \frac{dy}{dt}.$$

Im Ausdruck für das totale Differential wird “das Teilen durch  $dt$  weggekürzt” (das eine informelle Beschreibung).

**5.2.** Angenommen wir haben eine Gleichung  $F(x, y) = 0$ , die uns  $y$  als Funktion von  $x$  beschreibt. Man sagt in diesem Fall, dass  $y$  eine implizite Funktion ist (oder,  $y$  von  $x$  ist als eine implizite Funktion definiert). Aus dieser Gleichung kann man mit der Verwendung des totalen Differential die Ableitung von  $y$  nach  $x$  ausrechnen. Aus  $F(x, y) = 0$  erhalten wir

$$0 = dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy$$

Wenn  $\frac{\partial F}{\partial y} \neq 0$  ist, erhalten wir also

$$\frac{dy}{dx} = -\frac{\frac{\partial F}{\partial x}}{\frac{\partial F}{\partial y}}.$$

Diese rein formale Berechnung hat eine korrekte Interpretation als die Anwendung der Kettenregel zur Gleichung  $F(x, y(x)) = 0$  für eine Funktion  $y(x)$ .

## 6 Extremwertaufgaben

**6.1 Def. Lokale/globale Minima/Maxima/Extrema** einer multivariaten Funktion werden genauso wie im univariaten Fall definiert.

**6.2 Thm.** Sei  $f : X \rightarrow \mathbb{R}$  konvexe Funktion auf  $X \subseteq \mathbb{R}^n$  und sei  $a \in X$ . Dann sind die folgenden Bedingungen äquivalent:

- (i)  $f$  erreicht in  $a$  ein lokales Minimum.
- (ii)  $f$  erreicht in  $a$  ein globales Minimum.

**6.3 Thm** (Notwendige Bedingung). *Sei  $f : X \rightarrow \mathbb{R}$  partiell differenzierbar und sei  $a \in X$  innerer Punkt von  $X \subseteq \mathbb{R}^n$ . Erreicht  $f$  in  $a$  ein lokales Extremum, so gilt  $\nabla f(a) = 0$ .*

**6.4 Def.** Für  $f : X \rightarrow \mathbb{R}$  nennen wir einen Punkt  $a \in X$  von  $f$  kritisch bzw. stationär, wenn  $f$  in  $a$  differenzierbar ist und der Gradient von  $f$  an der Stelle  $a$  gleich 0 ist.

**6.5 Bsp.** Wir bestimmen das Minimum  $f(x, y) = 2y^2 + (x^2 - 2)x^2 + 2x^2y$  über  $\mathbb{R}^2$ . Die Funktion ist überall differenzierbar und auf dem gesamten  $\mathbb{R}^2$  definiert. Darüber hinaus gilt  $f \rightarrow \infty$  für  $x^2 + y^2 \rightarrow \infty$ , sodass die Funktion ihr Minimum erreicht. Es reicht also aus, alle stationären Punkte auszurechnen, und die Werte der Funktion an den stationären Punkten gegeneinander zu vergleichen.

Man hat

$$\begin{aligned}\frac{\partial f}{\partial x} &= 4x^3 - 4x + 4xy \\ \frac{\partial f}{\partial y} &= 2y + 2x^2\end{aligned}$$

Wir lösen das System

$$4x^3 - 4x + 4xy = 0,$$

$$4y + 2x^2 = 0$$

um die stationären Punkte auszurechnen.

Dabei unterscheiden wir zwischen  $x = 0$  und  $x \neq 0$ . Bei  $x = 0$  ist  $x = 0, y = 0$  die einzige Lösung. Bei  $x \neq 0$  können wir die erste Gleichung durch  $x$  teilen. Das ergibt (durch das zusätzliche Teilen durch Konstante) das System.

$$x^2 - 1 + y = 0$$

$$2y + x^2 = 0.$$

Das vorige System ist linear in  $x^2$  und  $y$ . Wir erhalten  $x^2 = 2, y = -1$ .

Insgesamt erhalten wir also die drei stationären Punkte:  $(0, 0), (-\sqrt{2}, 0), (\sqrt{2}, 0)$ .

Die Werte der Funktion  $f$  an diesen Punkten sind:

$$f(0, 0) = 0$$

$$f(\sqrt{2}, -1) = -2,$$

$$f(-\sqrt{2}, -1) = -2.$$

Das Minimum wird somit an zwei Stellen  $(\sqrt{2}, -1)$  und  $(-\sqrt{2}, -1)$  erreicht.

**6.6 Thm.** Sei  $f : X \rightarrow \mathbb{R}$  konvexe Funktion auf  $X \subseteq \mathbb{R}^n$  und sei  $a$  innerer Punkt von  $X$ , in dem die Funktion  $f$  differenzierbar ist. Dann sind die folgenden Bedingungen äquivalent:

- (a)  $f$  erreicht in  $a$  ein lokales Minimum.
- (b)  $f$  erreicht in  $a$  ein globales Minimum.
- (c)  $\nabla f(a) = 0$ .

**6.7 Aufgabe.** Seien  $a, b, c \in \mathbb{R}^2$  Ecken eines Dreiecks. Wir minimieren  $f(x) = \|x - a\| + \|x - b\| + \|x - c\|$  über  $\mathbb{R}^2$ . Da die Funktion  $f$  stetig ist und man  $f(x) \rightarrow \infty$  für  $\|x\| \rightarrow \infty$  hat, wird das Minimum erreicht. Welche Punkte von  $\mathbb{R}^2$  kommen als optimale Lösungen in Frage? Beschreiben Sie die Punkte geometrisch. Man beachte, dass  $f(x)$  nicht überall differenzierbar ist.

**6.8 Aufgabe.** Nehmen wir an, dass ein Roboter in der Lage ist, sich in der oberen Halbebene

$$H^+ = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \geq 0\}$$

mit einer Geschwindigkeit  $v > 0$  Einheiten pro Sekunde zu bewegen, und in der unteren Halbebene

$$H^- = \{(x_1, x_2) \in \mathbb{R}^2 : x_2 \leq 0\}$$

mit der Geschwindigkeit  $w > 0$  Einheiten pro Sekunde. Der Roboter soll von einem Punkt  $a \in H^+$  in der oberen Halbebene einen Punkt  $b \in H^-$  in der unteren Halbebene schnellstmöglich erreichen. In welchem Punkt  $(s, 0)$  mit  $s \in \mathbb{R}$  soll der Übergang von der oberen in die untere Halbebene stattfinden? Der Roboter bewegt sich geradlinig von  $a$  zu  $(s, 0)$  mit Geschwindigkeit  $v$  und anschließend von  $(s, 0)$  zu  $b$  geradlinig mit Geschwindigkeit  $w$ . Beschreiben Sie die Optimale Wahl des Übergangpunkts  $(s, 0)$  geometrisch.

## 7 Multidimensionale Taylorentwicklung

**7.1.** Zuerst formulieren wir die Entwicklung mit Hilfe des Operators für die Berechnung der Richtungsableitung:

$$\partial_h := h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n}.$$

**7.2 Thm** (Multidimensionale Taylorentwicklung, Hilfsformulierung). Sei  $f : X \rightarrow \mathbb{R}^n$  eine  $(N + 1)$ -mal stetig differenzierbare Funktion auf einer offenen konvexen Teilmenge  $X \subseteq \mathbb{R}^n$  und sei  $x \in X$ . Dann gilt

$$f(x + h) = \sum_{k=0}^N \frac{1}{k!} \left( h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n} \right)^k f(x) + O(\|h\|^{N+1}), \quad (\text{VII.4})$$

für  $h = (h_1, \dots, h_n) \rightarrow 0$ .

**7.3.** Theorem 7.2 folgt aus dem univariaten Fall der Taylorentwicklung. Man beachte, dass die Ableitungen  $(\partial_u)^k f(x)$  mit  $0 \leq k \leq N$ , die man in dieser Entwicklung benutzt Ableitungen der Ordnung bis zu  $N$  für die selbe Richtung  $u$  sind. Das erstellt den Bezug zur Dimension eins. Nun wollen wir aber Formel (VII.4) in eine explizite Form mit partiellen Ableitung der Ordnung höchstens  $N$  überführen, indem wir den Operator-Ausdruck

$$\left( h_1 \frac{\partial}{\partial x_1} + \cdots + h_n \frac{\partial}{\partial x_n} \right)^k$$

ausmultiplizieren. Dafür benötigen wir die Multinomialformel

**7.4** (Multinomialformel).

$$(u_1 + \cdots + u_n)^k = \sum_{\substack{p_1, \dots, p_n \geq 0 : \\ p_1 + \cdots + p_n = k}} \frac{k!}{p_1! \cdots p_n!} u_1^{p_1} \cdots u_n^{p_n}.$$

Hierbei ist  $\frac{k!}{p_1! \cdots p_n!}$  der sogenannte **Mutlinomialkoeffizient**; Bezeichnung dafür:

$$\binom{k}{p_1, \dots, p_n}.$$

Der **binomische Lehrsatz** ist der Fall  $n = 2$  der Multinomialformel.

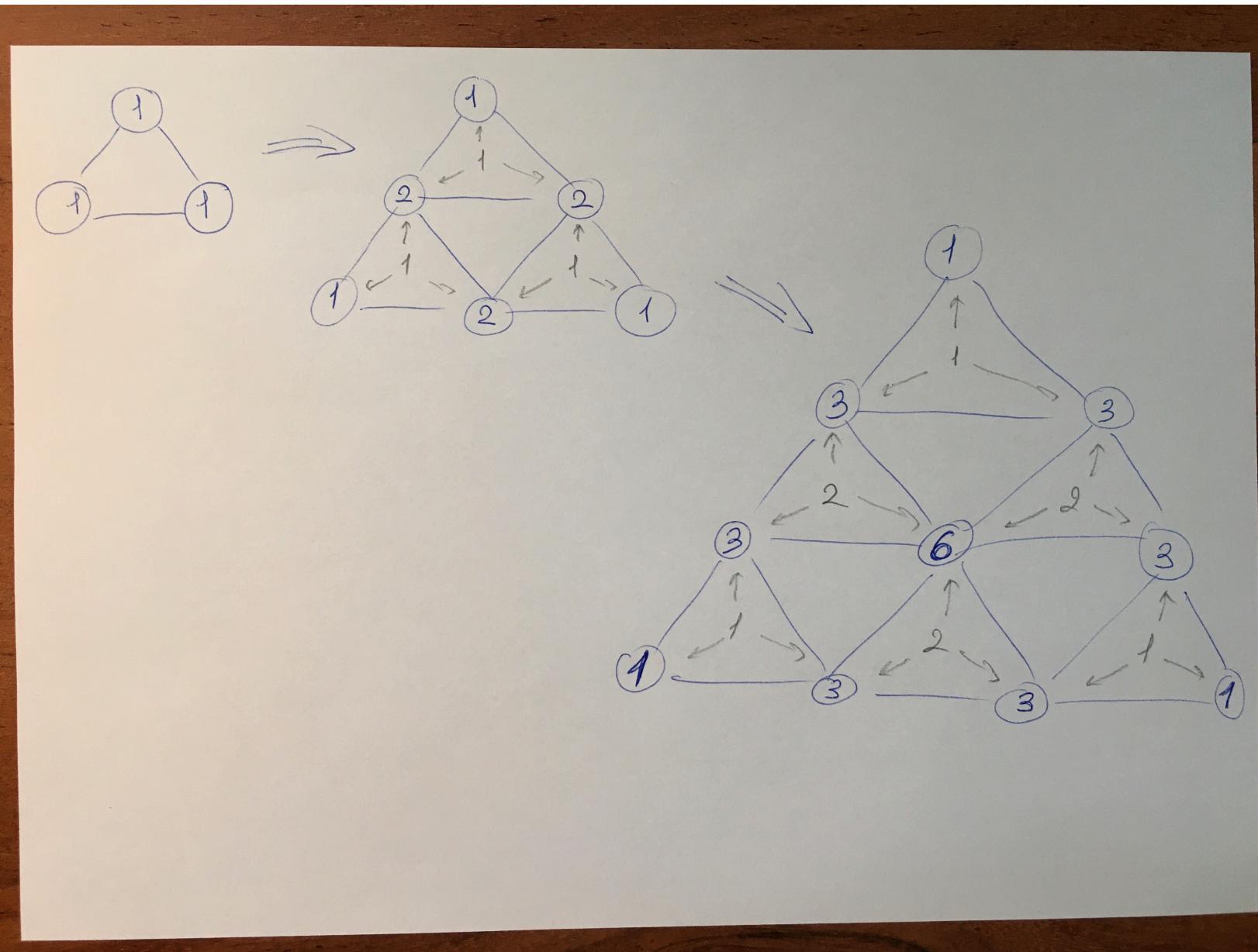
## 7.5 (Kombinatorische Bedeutung der multinomialen Koeffizienten).

Die Multinomiale Koeffizienten haben eine kombinatorische Bedeutung. Dementsprechend hat die Multinomialformel eine kombinatorische Begründung.

Hier ein Beispiel zur kombinatorischen Bedeutung der Multinomialkoeffizienten: Für jedes der 10 gegebenen Fächer bekommt eine Abschlussnote. Die Anzahl der Möglichkeiten, 4 Einsen, 3 Zweien und 2 Dreien zu bekommen ist

$$\frac{10!}{4! \cdot 3! \cdot 2!}$$

**7.6** (Ein Analogon des pascalschen Dreiecks.).



**7.7** (Multidimensionale Taylorentwicklung, Formulierung A).

$$f(x + h) = \sum_{\substack{p_1, \dots, p_n \geq 0 : \\ p_1 + \dots + p_n \leq N}} \frac{1}{p_1! \cdots p_n!} \cdot \frac{\partial^{p_1 + \dots + p_n} f}{\partial x_1^{p_1} \cdots \partial x_n^{p_n}}(x) h_1^{p_1} \cdots h_n^{p_n} + O(\|h\|^{N+1})$$

für  $h = (h_1, \dots, h_n) \rightarrow 0$ .

**7.8** (Multidimensionale Taylorentwicklung, Formulierung B).

$$f(x) = \sum_{\substack{p_1, \dots, p_n \geq 0 : \\ p_1 + \dots + p_n \leq N}} \frac{1}{p_1! \cdots p_n!} \cdot \frac{\partial^{p_1 + \dots + p_n} f}{\partial x_1^{p_1} \cdots \partial x_n^{p_n}}(a) (x_1 - a_1)^{p_1} \cdots (x_n - a_n)^{p_n} + O(\|x - a\|^{N+1})$$

für  $x = (x_1, \dots, x_n) \rightarrow a = (a_1, \dots, a_n)$ .

**7.9 Def.** Das  $n$ -variate Polynom

$$T_N(x) := \sum_{\substack{p_1, \dots, p_n \geq 0 : \\ p_1 + \dots + p_n \leq N}} \frac{1}{p_1! \cdots p_n!} \cdot \frac{\partial^{p_1 + \dots + p_n} f}{\partial x_1^{p_1} \cdots \partial x_n^{p_n}}(a) (x_1 - a_1)^{p_1} \cdots (x_n - a_n)^{p_n}$$

nennt man das **Taylorpolynom** der Ordnung  $N$  zum Entwicklungspunkt  $a$  für die Funktion  $f$ .

**7.10** (Spezialfall: Ordnung 1).

$$f(x + h) = f(x) + \langle \nabla f(x), h \rangle + O(\|h\|^2)$$

für  $h \rightarrow 0$ .

Lokale Approximation durch Polynom vom Grad 1. Relevant für notwendige Bedingung  $\nabla f(x) = 0$  der lokalen Optimalität.

**7.11** (Spezialfall: Ordnung 2).

$$f(x + h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \sum_{i,j=1,\dots,n} \frac{\partial^2}{\partial x_i \partial x_j} f(x) h_i h_j + O(\|h\|^3)$$

für  $h = (h_1, \dots, h_n) \rightarrow 0$ .

Lokale Approximation durch Polynom vom Grad 2. Relevant für notwendige und hinreichende Bedingungen der Optimalität (die wir im Folgenden diskutieren).

## **8 Extremwertaufgaben: Bedingungen zweiter Ordnung**

**8.1 Def.** Die **Hesse-Matrix** von  $f : X \rightarrow \mathbb{R}$  mit  $X \subseteq \mathbb{R}^n$  ist die  $n \times n$  Matrix der zweiten partiellen Ableitungen

$$\nabla^2 f := \left( \frac{\partial^2 f}{\partial x_i \partial x_j} \right)_{i,j=1,\dots,n}$$

Weitere Bezeichnung:  $H_f(x)$ .

**8.2** (Spezialfall: Taylorentwicklung der Ordnung 2 mit Hesse-Matrix).

$$f(x + h) = f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle h, \nabla^2 f(x)h \rangle + O(\|h\|^3)$$

oder mit etwas anderen Bezeichnungen

$$f(x + h) = f(x) + \nabla f(x)^\top h + \frac{1}{2} h^\top \nabla^2 f(x)h + O(\|h\|^3)$$

**8.3 Thm** (Notwendige Bedingungen zweiter Ordnung). *Sei  $f : X \rightarrow \mathbb{R}$  2-mal stetig partiell differenzierbar und sei  $a \in X$  innerer Punkt von  $X$ . Dann gilt:*

- (a) *Erreicht  $f$  in  $a$  ein lokales Minimum, so ist  $\nabla f(a) = 0$  und  $\nabla^2 f(a)$  positiv semidefinit.*
- (b) *Erreicht  $f$  in  $a$  ein lokales Maximum, so ist  $\nabla f(a) = 0$  und  $\nabla^2 f(a)$  negativ semidefinit.*

**8.4 Thm** (Hinreichende Bedingungen zweiter Ordnung). *Sei  $f : X \rightarrow \mathbb{R}$  2-mal stetig partiell differenzierbar und sei  $a \in X$  innerer Punkt von  $X$ . Dann gilt:*

- (a) *Ist  $\nabla f(a) = 0$  und  $\nabla^2 f(a)$  positiv definit, dann erreicht  $f$  in  $a$  ein lokales Minimum.*
- (b) *Ist  $\nabla f(a) = 0$  und  $\nabla^2 f(a)$  negativ definit, dann erreicht  $f$  in  $a$  ein lokales Maximum.*

## 9 Wiederholung: semidefinite und definite symmetrische Matrizen

**9.1 Def.** Eine Matrix  $A = (a_{ij})_{i,j} \in \mathbb{R}^{n \times n}$  heißt **symmetrisch**, wenn  $a_{ij} = a_{ji}$  für alle  $i, j \in \{1, \dots, n\}$  gilt. Mit Transposition schreibt man diese Bedingung als  $A = A^\top$ .

Die Funktion  $q_A(x) := x^\top Ax$  heißt **quadratische Form**. zur symmetrischen Matrix  $A = (a_{ij})_{i,j} \in \mathbb{R}^{n \times n}$ . Mit Koordinaten:

$$q_A(x_1, \dots, x_n) = \sum_{i,j=1,\dots,n} a_{ij}x_i x_j.$$

**9.2 Def.** Symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  bzw. die quadratische Form  $q_A(x)$  heißt:

- **positiv semidefinit**, wenn  $q_A(x) \geq 0$  für alle  $x \in \mathbb{R}^n$  gilt,
- **positiv definit**, wenn  $q_A(x) > 0$  für alle  $x \in \mathbb{R}^n \setminus \{0\}$  gilt,
- **negativ semidefinit**, wenn  $q_A(x) \leq 0$  für alle  $x \in \mathbb{R}^n$  gilt,
- **negativ definit**, wenn  $q_A(x) < 0$  für alle  $x \in \mathbb{R}^n \setminus \{0\}$  gilt.
- **indefinit**, wenn es  $x \in \mathbb{R}^n$  mit  $q_A(x) < 0$  sowie  $x \in \mathbb{R}^n$  mit  $q_A(x) > 0$  existieren.

**9.3 Thm.** *Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch. Dann sind alle Eigenwerte von  $A$  reell. Darüber hinaus besitzt  $\mathbb{R}^n$  eine Orthogonalbasis aus Eigenvektoren von  $A$ .*

**9.4.** Die Eigenwerte sind Nullstellen des characteristischen Polynoms  $\det(\lambda I - A)$ . Im Allgemeinen sind es komplexe Zahlen, das vorige Theorem garantiert aber, dass für symmetrische Matrizen  $A$ , die Eigenwerte stets reell sind.

**9.5 Thm.** Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch. Dann gilt:

- (a)  $A$  positiv semidefinit  $\Leftrightarrow$  alle Eigenwerte von  $A$  sind  $\geq 0$ .
- (b)  $A$  positiv definit  $\Leftrightarrow$  alle Eigenwerte von  $A$  sind  $> 0$ .
- (c)  $A$  negativ semidefinit  $\Leftrightarrow$  alle Eigenwerte von  $A$  sind  $\leq 0$ .
- (d)  $A$  negativ definit  $\Leftrightarrow$  alle Eigenwerte von  $A$  sind  $< 0$ .
- (e)  $A$  indefinit  $\Leftrightarrow A$  besitzt positive sowie negative Eigenwerte.

**9.6 Bsp.**

$$A = \begin{pmatrix} 1 & -3 \\ -3 & 9 \end{pmatrix}.$$

Wir lösen die charakteristische Gleichung

$$\begin{vmatrix} 1 - \lambda & -3 \\ -3 & 9 - \lambda \end{vmatrix} = 0$$

und erhalten 0 und 10 als Lösungen. Somit ist  $A$  positiv semidefinit (aber nicht positiv definit, wegen der 0).

**9.7 Bsp.**

$$A = \begin{pmatrix} -4 & 10 \\ 10 & -25 \end{pmatrix}$$

Wir lösen die charakteristische Gleichung

$$\begin{vmatrix} -4 - \lambda & 10 \\ 10 & -25 - \lambda \end{vmatrix} = 0$$

und erhalten 0 und  $-29$  als Lösungen. Somit ist  $A$  negativ semidefinit (aber nicht negativ definit, wegen der 0).

**9.8 Bsp.**

$$A = \begin{pmatrix} 5 & -1 \\ -1 & 2 \end{pmatrix}.$$

Wir lösen die charakteristische Gleichung

$$\begin{vmatrix} 5 - \lambda & -1 \\ -1 & 2 - \lambda \end{vmatrix} = 0$$

und erhalten  $3\frac{1}{2} \pm \sqrt{(3\frac{1}{2})^2 - 9}$  als Lösungen. Da die beiden Lösungen positiv sind, ist die Matrix  $A$  positiv definit.

**9.9.** Es gibt ein Resultat, das uns ermöglicht, die Berechnung der Eigenwerte zu vermeiden...

**9.10 Thm.** Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch. Dann gilt:

(a)  $A$  positiv semidefinit  $\Leftrightarrow$  die Koeffizienten  $c_1, \dots, c_n$  vom Polynom

$$\det(\lambda I + A) = \lambda^n + c_1 \lambda^{n-1} + \cdots + c_n \lambda^0$$

sind alle  $\geq 0$ .

(b)  $A$  positiv definit  $\Leftrightarrow$  die Koeffizienten  $c_1, \dots, c_n$  vom Polynom

$$\det(\lambda I + A) = \lambda^n + c_1 \lambda^{n-1} + \cdots + c_n \lambda^0$$

sind alle  $> 0$ .

(c) *A negativ semidefinit  $\Leftrightarrow$  die Koeffizienten  $d_1, \dots, d_n$  vom Polynom*

$$\det(\lambda I - A) = \lambda^n + d_1\lambda^{n-1} + \cdots + d_n\lambda^0$$

*sind alle  $\geq 0$ .*

(d) *A negativ definit  $\Leftrightarrow$  die Koeffizienten  $d_1, \dots, d_n$  vom Polynom*

$$\det(\lambda I - A) = \lambda^n + d_1\lambda^{n-1} + \cdots + d_n\lambda^0$$

*sind alle  $> 0$ .*

**9.11.**  $f_A(\lambda) := \det(\lambda I - A)$  ist das charakteristische Polynom und  $\det(\lambda I + A)$  ergibt sich aus dem charakteristischen Polynom:

$$\det(\lambda I + A) = (-1)^n f_A(-\lambda)$$

für  $A \in \mathbb{R}^{n \times n}$ . Das bedeutet, dass man die Art von  $A$  (positiv/negativ definit bzw. semidefinit or indefinit) anhand der Vorzeichen der Koeffizienten des charakteristischen Polynoms  $f_A(\lambda)$  entscheiden kann.

**9.12.** Für eine Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{R}^{2 \times 2}.$$

ist

$$\det(\lambda I - A) = \lambda^2 - \text{tr}(A)\lambda + \det(A)$$

Das führt zum folgenden Resultat:

$$\text{tr}(A) = a_{11} + a_{22} \quad (\text{Spur}),$$

$$\det(A) = a_{11}a_{22} - a_{12}a_{21} \quad (\text{Determinante}).$$

**9.13 Thm.** Sei  $A \in \mathbb{R}^{2 \times 2}$  symmetrisch. Dann gilt:

- (a)  $A$  positiv semidefinit  $\Leftrightarrow \text{tr}(A) \geq 0, \det(A) \geq 0$ .
- (b)  $A$  positiv definit  $\Leftrightarrow \text{tr}(A) > 0, \det(A) > 0$ .
- (c)  $A$  negativ semidefinit  $\Leftrightarrow \text{tr}(A) \leq 0, \det(A) \geq 0$ .
- (d)  $A$  negativ definit  $\Leftrightarrow \text{tr}(A) < 0, \det(A) > 0$ .
- (e)  $A$  indefinit  $\Leftrightarrow \det(A) < 0$

**9.14 Aufgabe.** Kann eine symmetrische Matrix  $A \in \mathbb{R}^{3 \times 3}$  mit  $\det(A) < 0$ :

- (a) positiv semidefinit sein?
- (b) positiv definit sein?
- (c) negativ semidefinit sein?
- (d) negativ definit sein?
- (e) indefinit sein?

Beantworten Sie die selben Fragen für

- eine symmetrische Matrix  $A \in \mathbb{R}^{4 \times 4}$  mit  $\det(A) < 0$ .
- eine symmetrische Matrix  $A \in \mathbb{R}^{4 \times 4}$  mit  $\text{tr}(A) > 0$  ( $\text{tr}(A)$  ist die Summe der Diagonalelemente von  $A$ ).

**9.15.** Eine andere Weise, die Art einer symmetrischen Matrix  $A \in \mathbb{R}^{n \times n}$  zu Bestimmen ist durch Diagonalisierung er quadratischen Form  $q_A(x)$ . Man kann eine reguläre Matrix  $U$  bestimmen (ein anderes Koordinatensystem) für welche die quadratische Form  $q_A(Ux)$  diagonal ist, das heißt

$$q_A(Ux) = \gamma_1 x_1^2 + \cdots + \gamma_n x_n^2$$

gilt für gewisse  $\gamma_1, \dots, \gamma_n \in \mathbb{R}$ . Das Diagonalisierungsverfahren sieht dem Gauß-Verfahren sehr ähnlich. TO DO: Mehr Details.

## 10 Restringierte Optimierungsaufgaben

TODO: KKT für eine Gleichheitsbedingung. KKT für mehrere Gleichheitsbedingungen.

# **Kapitel VIII**

## **Integralrechnung II**

### **1 Riemann-Integral**

**1.1 Def.** Ein  $n$ -dimensionaler (achsenparalleler) **Quader**  $Q \subseteq \mathbb{R}^n$  ist kartesisches Produkt

$$Q := [a_1, b_1] \times \cdots \times [a_n, b_n]$$

von  $n$  abgeschlossenen Intervallen  $[a_i, b_i]$  mit  $a_i < b_i$  für  $i \in \{1, \dots, n\}$ . Für  $n = 1$  ist  $Q$  ein abgeschlossenes Intervall, für  $n = 2$  ein achsenparalleles Rechteck, und für  $n = 3$  ein 3-dimensionaler Quader. Das **Volumen** von  $Q$ , wird als

$$\text{vol}(Q) = (b_1 - a_1) \cdot \dots \cdot (b_n - a_n)$$

definiert.

**1.2.** Der Durchschnitt von zwei Quadern

$$Q = I_1 \times \cdots \times I_n$$

und

$$R = J_1 \times \cdots \times J_n$$

ist gleich

$$Q \cap R = (I_1 \cap J_1) \times \cdots \times (I_n \cap J_n).$$

Dieser Durchschnitt ist nur dann wieder ein Quader, wenn alle Durchschnitte  $I_k \cap J_k$  mit  $k \in \{1, \dots, n\}$  nichtentartete Segmente sind.

**1.3 Def.** Wir sagen, dass sich zwei  $n$ -dimensionale Quader  $Q$  und  $R$  **überlappen**, wenn ihr Durchschnitt  $Q \cap R$  wieder ein  $n$ -dimensionaler Quader ist.

**1.4 Def.** Eine **Zerlegung**  $Z$  eines  $n$ -dimensionalen Quaders  $Q$  ist eine Menge  $Z = \{Q_1, \dots, Q_m\}$  von endlich vielen Quadern, die sich paarweise nicht überlappen, und deren Vereinigung  $Q_1 \cup \dots \cup Q_m$  mit  $Q$  übereinstimmt. Diese Definition erweitert den Begriff der Zerlegung von der Dimension  $n = 1$  auf den Fall einer allgemeinen Dimension  $n \in \mathbb{N}$ .

**1.5 Def.** Für eine Funktion  $f : Q \rightarrow \mathbb{R}$  auf einem Quader  $n$  und eine Zerlegung  $Z = \{Q_1, \dots, Q_m\}$  von  $Q$ , definiert man die **Ober-** bzw. die **Untersumme** von  $f$  bzgl.  $Z$  als

$$O_f(Z) := \sum_{i=1}^m \sup_{x \in Q_i} f(x) \operatorname{vol}(Q_i),$$

$$U_f(Z) := \sum_{i=1}^m \inf_{x \in Q_i} f(x) \operatorname{vol}(Q_i).$$

Des weiteren definiert man das Ober- bzw. das **Unterintegral** von  $f$  auf  $Q$  als

$$O_f := \inf_{Z \text{ Zerlegung von } Q} O_f(Z),$$

$$U_f := \sup_{Z \text{ Zerlegung von } Q} U_f(Z)$$

Mann nennt die Funktion  $f$  **Riemann-integrierbar** über  $Q$ , wenn  $O_f = U_f$  gilt. In diesem nennt man den Wert  $O_f = U_f$  das Riemann integral von  $f$  über  $Q$ . Bezeichnung:

$$\int_Q f(x) \, d\,x.$$

Bei  $x = (x_1, \dots, x_n)$  schreibt man  $d\,x$  in Koordinaten als  $d\,x_1 \cdot \dots \cdot d\,x_n$ .

**1.6 Def.** Ist  $f : D \rightarrow \mathbb{R}$  auf einem beliebigen beschränkten Bereich  $D \subseteq \mathbb{R}^n$  definiert, führt man das **Riemann-Integral** als

$$\int_D f(x) \, dx := \int_Q g(x) \, dx,$$

für einen Quader  $Q$  mit  $D \subseteq Q$  und die Funktion

$$g(x) := \begin{cases} f(x) & \text{für } x \in D, \\ 0 & \text{sonst.} \end{cases}$$

Die Funktion  $f$  nennt man **Riemann-integrierbar** über  $D$ , wenn das Integral  $\int_Q g(x) \, dx$  oben existiert (in diesem Fall hängt der Wert  $\int_Q g(x) \, dx$  von der konkreten Wahl  $Q$  des Umgebungsquader nicht ab).

**1.7 Def.** Wenn für eine beschränkte Menge  $D \subseteq \mathbb{R}^n$ , die Konstante 1 über  $D$  **Riemann-Integrierbar** ist, dann nenn wir

$$\text{vol}(D) := \int_D 1 \, d\,x$$

das Volumen von  $D$ .

**1.8 Thm.** Wenn eine Funktion  $f : A \cup B \rightarrow \mathbb{R}$  über  $A \subseteq \mathbb{R}^n$  und  $B \subseteq \mathbb{R}^n$  Riemann-integrierbar ist, und  $A \cap B$  keine innere Punkte hat, dann gilt

$$\int_{A \cup B} f(x) \, dx = \int_A f(x) \, dx + \int_B f(x) \, dx.$$

**1.9 Thm.** Sind  $f, g : D \rightarrow \mathbb{R}$  beide Riemann-integrierbar über  $D \subseteq \mathbb{R}^n$ , dann ist  $\alpha f + \beta g$  für jede Wahl von  $\alpha, \beta \in \mathbb{R}$  ebenfalls Riemann-integrierbar über  $D$ , und es gilt:

$$\int_D (\alpha f(x) + \beta g(x)) \, dx = \alpha \int_D f(x) \, dx + \beta \int_D g(x) \, dx.$$

**1.10 Thm.** *Jede stetige Funktion über einem Quader ist Riemann-integrierbar.*

**1.11 Thm.** Seien  $f, g : D \rightarrow \mathbb{R}$  Riemann-integrierbare Funktion mit  $f(x) \leq g(x)$  für alle  $x \in D$ . Dann gilt

$$\int_D f(x) \, dx \leq \int_D g(x) \, dx.$$

## 2 Der Satz von Fubini

**2.1 Thm** (Fubini für zwei-dimensionale Quadere). *Ist  $f : [a_1, b_1] \times [a_2, b_2] \rightarrow \mathbb{R}$  über das Rechteck  $[a_1, b_1] \times [a_2, b_2]$  mit  $a_1 < b_1, a_2 < b_2$  Riemann-integrierbar, so gilt:*

$$\begin{aligned} \int_{[a_1, b_1] \times [a_2, b_2]} f(x_1, x_2) \, d(x_1 \, d x_2) &= \int_{a_2}^{b_2} \left( \int_{a_1}^{b_1} f(x_1, x_2) \, d x_1 \right) \, d x_2 \\ &= \int_{a_1}^{b_1} \left( \int_{a_2}^{b_2} f(x_1, x_2) \, d x_2 \right) \, d x_1. \end{aligned}$$

**2.2.** Das Theorem von Fubini für  $n$ -dimensionale Quader ist komplett Analog: ein  $n$ -dimensionales Integral lässt sich als durch  $n$  geschachtelte ein-dimensionale Integrale ausschreiben. Beim Ausrechnen geht man von innen nach außen vor.

**2.3 Aufgabe.** Berechnen Sie

$$\int_{[-\pi/2, \pi/2] \times [0, \pi]} \cos\left(x - \frac{y}{2}\right) dx dy.$$

**2.4 Thm** (Fubini für Normalbereiche). Seien  $c, d : [a, b] \rightarrow \mathbb{R}$  stetige Funktionen mit  $c(x) \leq d(x)$  für alle  $x \in \mathbb{R}$  und sei  $f$  Riemann-integrierbare Funktion auf

$$D = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, c(x) \leq y \leq d(x)\}.$$

Dann gilt

$$\int_D f(x, y) \, dx \, dy = \int_a^b \left( \int_{c(x)}^{d(x)} f(x, y) \, dx \right) dy.$$

**2.5.** Den Bereich  $D$  wie in dem vorigen Theorem nennt man den Normalbereich bzgl. der  $x$ -Achse. Man kann das Theorem von Fubini analog für Normalbereiche bzgl.  $y$ -Achse formulieren. Des Weiteren kann man im Fall von mehr Variablen multidimensionale Normalbereiche einführen und dafür den Satz von Fubini formulieren, etwa für  $n = 3$

$$D = \{(x_1, x_2, x_3) \in \mathbb{R}^3 : a_1 \leq x_1 \leq b_1, a_2(x_1) \leq x_2 \leq b_2(x_1), a_3(x_1, x_2) \leq x_3 \leq b_3(x_1, x_2)\}.$$

hat man

$$\int_D f(x_1, x_2, x_3) d x_1 d x_2 d x_3 = \int_{a_1}^{b_1} \left( \int_{a_2(x_1)}^{b_2(x_1)} d x_2 \left( \int_{a_3(x_1, x_2)}^{b_3(x_1, x_2)} f(x_1, x_2, x_3) \right) \right) d x_1.$$

**2.6 Aufgabe.**

$$\int_{x^2 \leq y \leq x} (xe^y - y^2) \, d\,x \, d\,y.$$

**2.7.** Wenn man über einen Bereich integriert, der nicht als Normalbereich darstellbar ist, kann man versuchen, den Bereich in eine Vereinigung nicht-überlappender Normalbereiche zu zerlegen.

### **3 Substitution beim mehrdimensionalen Integrieren**

**3.1 Thm.** Seien  $D, R \subseteq \mathbb{R}^n$ . Sei  $f : D \rightarrow \mathbb{R}$  Riemann-integrierbar und sei  $T : D \rightarrow R$  eine bijektive Abbildung, für welche  $f$  und  $f^{-1}$  beide differenzierbar sind (insbesondere, man hat  $R = T(D)$ ). Dann gilt:

$$\int_{T(D)} f(y) \, dy = \int_D f(T(x)) |\det(J_T(x))| \, dx.$$

**3.2.** Bijektivität von  $T : D \rightarrow R$  kann durch eine schwächere Voraussetzung ersetzt werden, dass  $T$  keine “Selbstüberlappungen” erzeugt. Das heißt, dass die Menge aller  $y \in R$ , für welche man mehr als ein  $x \in D$  mit  $f(x) = y$  findet, Volumen null hat.

### 3.3 Bsp. Polarkoordinaten. TODO

TODO

# **Kapitel IX**

## **Anwendungen der Analysis**

### **1 Numerische Verfahren**

**1.1.** Fixpunkt-Gleichung TODO

**1.2.** Newton-Verfahren TODO

**1.3.** Sekanten-Verfahren

**1.4.** Newton-Verfahren zur Lösungen von Optimierungsaufgaben

## 2 Gewöhnliche Differentialgleichungen

## Literaturverzeichnis

- [Rud] Walter Rudin: Analysis. 4. Auflage, Oldenbourg 2009, 408 Seiten. (Eines der Standard-lehrbücher zur Analysis. Englisches Original “Principles of Mathematical Analysis”)
- [Dem] Boris Pavlovich Demidovitch: A Collection of Problems and Exercises in Mathematical Analysis, 11. Auflage, 1995, in Russisch (Über 4000 Aufgaben zu verschiedenen Themen aus der Analysis)